TECHNICKÁ UNIVERZITA V LIBERCI

Fakulta mechatroniky, informatiky a mezioborových studií

BAKALÁŘSKÁ PRÁCE

LIBEREC 2009 David Vápenka

TECHNICKÁ UNIVERZITA V LIBERCI

Fakulta mechatroniky a mezioborových inženýrských studií

Studijní program: B 2612 Elektrotechnika a informatika

Studijní obor: 2612R011 Elektronické informační a řídící systémy

DC/DC měnič pro napájení piezoelektrických aktuátorů

DC/DC converter for piezoelectric actuators

Bakalářská práce

Autor: **David Vápenka**

Vedoucí práce: Ing. Jan Václavík

Konzultant: doc. Ing. Pavel Rydlo, Ph.D.

V Liberci

Prohlášení

Byl jsem seznámen s tím, že na mou bakalářskou práci (BP) se plně vztahuje zákon č. 121/2000 o právu autorském, zejména §60 (školní dílo).

Beru na vědomí, že TUL má právo na uzavření licenční smlouvy o užití mé BP, a prohlašuji, že **souhlasím** s případným užitím mé BP (prodej, zapůjčení apod.).

Jsem si vědom toho, že užít svou BP či poskytnout licenci k jejímu využití mohu jen se souhlasem TUL, která má právo ode mne požadovat přiměřený příspěvek na úhradu nákladů, vynaložených univerzitou na vytvoření díla (až do jejich skutečné výše).

Bakalářskou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím bakalářské práce a konzultantem.

Datum: 20. května 2009

Podpis:

Poděkování

Rád bych využil této příležitosti a poděkoval svým rodičům za umožnění studia a za jejich vytrvalou podporu, bez níž bych se nedostal tak daleko a to jak ve studiu tak v osobním životě. Dále bych chtěl poděkovat Ing. Janu Václavíkovi za cenné rady, poskytnutý čas a odborné vedení.

Abstrakt

Zadáním této práce je konstrukce DC/DC měniče pro napájení piezoelektrických aktuátorů. Parametry měniče a konstrukce zdroje je uzpůsobena pro využití komerčně dostupného transformátoru pro síťové adaptéry. Návrh tedy neprobíhá standardním způsobem. Je uzpůsoben potřebám vybraného transformátoru. Důraz je kladen především na jednoduchost řešení a funkčnost výsledné realizace.

Celá práce je kvůli přehlednosti členěna do tematických celků. V úvodní části jsou vysvětleny základní principy a rozdělení týkající se problematiky spínaných zdrojů. Následuje vlastní návrh a konstrukce DC/DC měniče. V závěru je uvedeno srovnání předpokládaných parametrů s parametry zkonstruovaného měniče.

Klíčová slova: piezoelektrický aktuátor, DC/DC měnič, transformátor

Abstract

This assignment is about the design DC/DC converters for power piezo actuators. Parameters of the converter structure and resources are designed for the use of commercially available transformer for network adapters. Therefore, the proposal is not standard and it is adapted to needs of the selected transformer. The emphasis is placed on simplicity and functionality of the final implementation. It is easy to follow the assignment because it is divided into thematic units.

The introduction explains the basic principles of division on the issue of switching power supply. The design and construction of DC/DC converter is coming after. In conclusion, there is a comparison between the estimated parameters and the parameters of the designed converter.

Keywords: piezoelectric actuators, DC/DC converter, transformer

Obsah

Seznam obrázků 8

Seznam tabulek 9

Seznam použitých symbolů 10

Seznam použitých zkratek 11

ÚVOD 12

1 PIEZOELEKTRICKÉ AKTUÁTORY 13

1.1 Piezoelektrický efekt 13

1.2 Vlastnosti piezoelektrických aktuátorů 13

2 PROBLEMATIKA SPÍNANÝCH ZDROJŮ 15

2.1 Srovnání s lineárními zdroji 15

2.1.1 Výhody spínaných zdrojů 17

2.1.2 Nevýhody spínaných zdrojů 17

2.2 Princip funkce 17

2.2.1 Přerušovaný režim 18

2.2.2 Souvislý režim 19

2.2.3 Kritický režim 19

2.3 PWM modulace 20

3 ZÁKLADNÍ TOPOLOGIE SPÍNANÝCH ZDROJŮ 21

3.1 Blokové schéma 21

3.2 Rozdělení podle funkce 22

3.2.1 Snižující měnič (BUCK, Step-Down) 22

3.2.2 Zvyšující měnič (BOOST, Step-Up) 23

3.2.3 Invertující měnič 24

3.3 Rozdělení podle způsobu přenosu energie 25

3.3.1 Propustný měnič (FORWARD) 25

3.3.2 Blokující měnič (FLYBACK) 25

4 NÁVRH DC/DC MĚNIČE 27

4.1 Stanovení parametrů zdroje 27

4.2 Výběr transformátoru 28

4.2.1 Vybraný transformátor 28

4.3 Výběr řídícího obvodu 29

4.3.1 Vybraný řídící obvod 30

4.4 Výběr tranzistoru 31

4.4.1 Vybraný tranzistor 33

4.5 Výběr výstupního filtru 33

4.6 Výběr výstupní diody 34

4.7 Návrh schématu 34

4.7.1 Předpoklady pro návrh 35

4.7.2 Navržené schéma 36

4.8 Návrh desky plošných spojů 39

4.8.1 Navržená DPS 39

5 VÝSLEDKY 41

5.1 Měření zátěžové charakteristiky 41

5.2 Měření parametrů 42

5.2.1 Napěťové rozsahy 42

5.2.2 Výstupní zvlnění 43

5.2.3 Odezva na skok 44

5.2.4 Účinnost 45

5.2.5 Porovnání parametrů 47

ZÁVĚR 48

Seznam použité literatury 49

Seznam příloh 50

Příloha A 51

Seznam obrázků

Obr. 1: Uspořádání piezoelektrických aktuátorů [9] 14

Obr. 2: Náhradní schéma lineárního regulátoru 15

Obr. 3: Skutečné schéma lineárního regulátoru 16

Obr. 4: Náhradní schéma spínaného regulátoru 16

Obr. 5: Princip funkce spínaného zdroje 17

Obr. 6: Přerušovaný režim – průběhy 19

Obr. 7: Souvislý režim – průběhy 19

Obr. 8: Rekonstrukce sinusového signálu pomocí PWM 20

Obr. 9: Blokové schéma spínaného zdroje 21

Obr. 10: Snižující měnič 22

Obr. 11: Zvyšující měnič 23

Obr. 12: Invertující měnič 24

Obr. 13: Propustný měnič 25

Obr. 14: Blokující měnič 26

Obr. 15: Schéma zapojení napájecí části 36

Obr. 16: Schéma zapojení řídící části 37

Obr. 17: Schéma zapojení ošetřující protizkratovou ochranu 38

Obr. 18: Rozložení součástek na DPS 39

Obr. 19: Obraz DPS ze strany BOTTOM 40

Obr. 20: Obraz DPS ze strany TOP 40

Obr. 21: Schéma zapojení při měření zatěžovací charakteristiky 41

Obr. 22: Výstupní zvlnění 43

Obr. 23: Stanovení rychlosti reakce (naprázdno → 50 %) 44

Obr. 24: Stanovení rychlosti reakce (25 % → 75 %) 45

Seznam tabulek

Tab. 1: Vlastnosti piezoelektrických materiálů 14

Tab. 2: Porovnání lineárních a spínaných zdrojů 16

Tab. 3: Stanovené parametry 27

Tab. 4: Parametry transformátoru MYRRA 74010 29

Tab. 5: Parametry řídícího obvodu TDA4605-3 30

Tab. 6: Ztráty na zvolených tranzistorech 32

Tab. 7: Parametry tranzistoru IRF540 33

Tab. 8: Parametry diody SF38 34

Tab. 9: Rozsah výstupního napětí 43

Tab. 10: Rozsah vstupního napětí 43

Tab. 11: Měření účinnosti - hodnoty 46

Tab. 12: Srovnání parametrů měniče 47

Tab. 13: Zatěžovací charakteristika – naměřené hodnoty 51

Seznam použitých symbolů

R	[Ω]	elektrický odpor
U	[V]	elektrické napětí
I	[A]	elektrický proud
P	[W]	výkon
L	[H]	indukčnost
C	[F]	kapacita
t	[s]	čas
T	[s]	perioda
f	[Hz]	frekvence
η	[%]	účinnost
p		převod transformátoru

Seznam použitých zkratek

T_{ON}	čas po který je spínač sepnut
T_{OFF}	čas po který je spínač vypnut
PWM	pulzně šířková modulace
ESR	ekvivalentní sériový odpor kondenzátoru
DPS	deska plošných spojů
Obr.	obrázek
Tab.	tabulka
Eq.	rovnice
SMD	technologie povrchové montáže součástek
IO	integrovaný obvod
R_Z	odpor zátěže
GND	uzemnění
DC	stejnosměrné napětí (proud)
AC	střídavé napětí (proud)
RC	RC článek
LC	LC článek

Úvod

Jak spínané zdroje, tak piezoelektrické aktuátory jsou v dnešní době velice aktuální téma. Jde o poměrně nedávno vyvinuté technologie, které ještě zdaleka nevyčerpaly svůj potenciál a nalézají uplatnění ve stále nových aplikacích.

Zvlášť poslední dobou je kladen veliký důraz na zefektivňování, vyšší ekonomičnost a pokud možno objemovou minimalizaci veškerých zařízení. Klasické zdroje jsou tak ve velké míře nahrazovány menšími efektivnějšími, a tím pádem i hospodárnějšími spínanými zdroji. Nahrazování lineárních zdrojů je navíc podporováno faktem, že cena spínaných zdrojů je vzhledem k velkým výrobním sériím více než příznivá. Konstrukční různorodost a vysoká přizpůsobitelnost zapříčiňují využití spínaných zdrojů ve stále širším okruhu aplikací. Navíc je možné pomocí DC/DC měničů napětí nejen snižovat, ale i zvyšovat.

Úkolem této práce je navrhnout spínaný zdroj pro piezoelektrické aktuátory. Jde o testovací výrobek, který má ověřit získané teoretické poznatky a možnosti uplatnění spínaných zdrojů v této oblasti. Jelikož nejde o průmyslovou aplikaci, ale o testovací kus, nebyl kladen důraz na ošetření výrobku od rušení a jiných parazitních vlivů reálného světa. Důraz byl kladen především na to, aby zkonstruovaný měnič splňoval podmínky stanovené vlastnostmi aktuátorů a co nejlépe této aplikaci vyhovoval.

Hlavním rysem odlišujícím tento návrh od běžného návrhu spínaného zdroje spočívá ve využití komerčně vyráběného transformátoru. Není tak možné postupovat obvyklým způsobem. Návrh je potřeba přizpůsobit vybranému transformátoru. Tím se na jedné straně návrh částečně zjednodušuje, na straně druhé to však přináší některá omezení, která komplikují nejen výběr součástek, ale i návrh jako celek.

1 Piezoelektrické aktuátory

Piezoelektrické aktuátory jsou akční členy využívající inverzní piezoelektrický efekt. Mají nezastupitelnou úlohu v oblasti mikromanipulace. Uplatňují se také v systémech aktivního potlačování hluku a vibrací. Další aplikace lze nalézt v reproduktorech, vstřikování paliva do válců systémem Common Rail, miniaturních ultrazvukových motorech, pohybových ústrojích zrcadel laserů, inkoustových tiskárnách a mikroskopech.

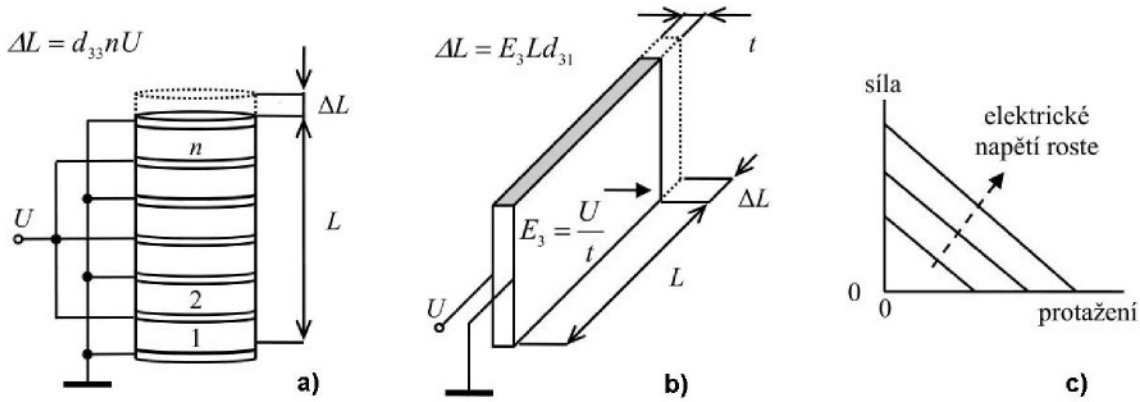
1.1 Piezoelektrický efekt

Přímý piezoelektrický efekt využívá schopnosti některých krystalických materiálů generovat elektrický náboj úměrně vnější síle, která na ně působí. Tento jev je využíván především ve snímačích síly. Naproti tomu **inverzní jev** spočívá v tom, že mechanické roztažení či smrštění je vybuzeno elektrickým polem, orientovaným rovnoběžně ke směru polarizace. Piezoelektrický efekt je anizotropní, tj. působí jen v určitém směru.

1.2 Vlastnosti piezoelektrických aktuátorů

Nejrozšířenějšími průmyslově využívanými piezoelektrickými materiály jsou keramika PZT (olovo, zirkonium, titan) a polymer PVDF (polyvinyliden fluorid). Aktuátory vyrobené z PVDF jsou ohebnější než z PZT, mají však mnohonásobně nižší citlivost. Maximální protažení je v rozmezí 0,1 až 0,13 %. Tyto prvky vykazují hysterezi až 15 %. Právě na hysterezi materiálu, z něhož je aktuátor vyroben, je silně závislá jeho pracovní frekvence. Mezi další parametry, činící piezoelektrické aktuátory tak atraktivními, patří jejich extrémně krátká doba odezvy.

Lineární piezoelektrický aktuátor je buď vrstven z tenkých keramických destiček (tloušťka mezi 0,1 až 1 mm) oddělených elektrodami (Obr. 1a), nebo je laminárně uspořádán (Obr. 1b). Potřebné napětí pro jednu destičku diskového tvaru tloušťky 0,1 mm je asi 100 V. Destička o tloušťce 1 mm vyžaduje 1000 V. K výpočtu protažení je podle vzorců uvedených v obrázcích potřeba znát kromě vyznačených délek také piezoelektrické konstanty *d33* a *d31*. Obrázek (Obr. 1c) ukazuje závislost blokovací síly aktuátoru na jeho protažení.



Obr. 1: Uspořádání piezoelektrických aktuátorů [9]

K dimenzování je navíc třeba znát také maximální mechanické napětí a intenzitu elektrického pole. Tyto údaje jsou pro srovnání obou materiálů uvedeny v tabulce (Tab. 1).

Tab. 1: Vlastnosti piezoelektrických materiálů

Materiál	PZT	PVDF
Piezoelektrické konstanty		
d33	300	-25
d31	-150	15
Maximální napětí (Mpa)		
tah	80	200
tlak	600	200
Max. Intenzita el. Pole (V/mm)	2000	5.10^5
Hustota (Kg/m^3)	7600	1800

Pro doplnění jsou zde uvedeny lineární stavové rovnice, popisující piezoelektrický jev.

$$D_i = d_{i,j,k} \cdot T_{j,k} + \varepsilon^T{}_{i,j,k} \cdot E_j$$

$$S_{i,j} = S^E{}_{i,j,k,l} \cdot T_{k,l} + d_{k,i,j} \cdot E_k$$

Kde $S_{i,j}$, $T_{j,k}$ jsou složky mechanické deformace, mechanického napětí a D_i, E_k jsou složky elektrické indukce a intenzity.

2 Problematika spínaných zdrojů

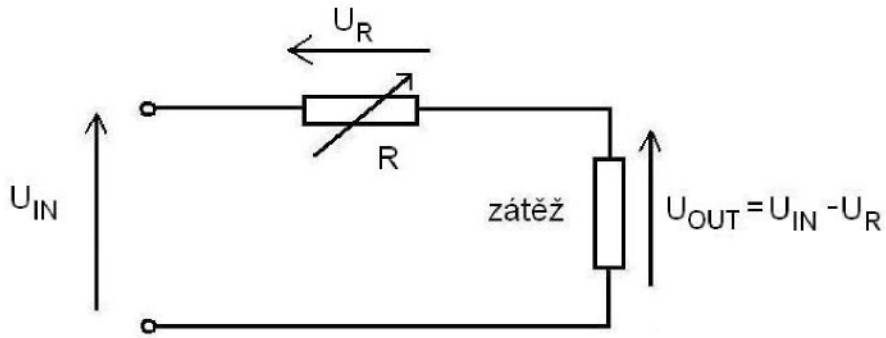
Tato kapitola se zabývá základním principem funkce spínaných zdrojů. Dále srovnáním s lineárními zdroji a v neposlední řadě také jejich výhodami a nevýhodami.

Rozmach těchto zdrojů mohl, díky potřebě vysokofrekvenčního spínacího prvku, začít až s rozvojem tranzistorů a polovodičů obecně. Návrháři potřebovali zdroje zefektivnit a hlavně zmenšit. Postupným upravováním a zefektivňováním nabírala zapojení na složitosti. I přes četné výhody spínaných zdrojů mají stále lineární zdroje nezastupitelné místo a to především v nízkovýkonových aplikacích a také tam kde je potřeba nízkošumový zdroj.

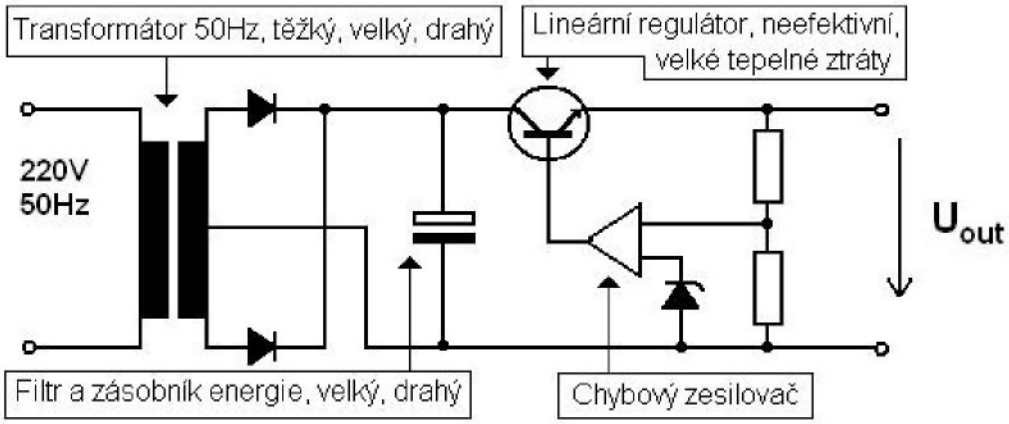
2.1 Srovnání s lineárními zdroji

Hlavní rozdíl mezi spínací a lineární regulací je v hmotnosti, efektivitě a váze zkonstruovaného zdroje. Navíc lineární regulací, na rozdíl od spínací, lze z principu napětí pouze snižovat.

Lineární regulátory (Obr. 2) dokážou udržet stabilní výstupní napětí i při kolísavém výkonu, avšak za cenu vysokých ztrát v regulační jednotce, které se dále zvyšují s rozdílem vstupního a výstupního napětí. Nadbytečný výkon je v tomto případě mařen na proměnném odporu *R*, který musí být dostatečně chlazen, aby nedošlo k jeho zničení. Skutečné schéma lineárního regulátoru je zde (Obr. 3).

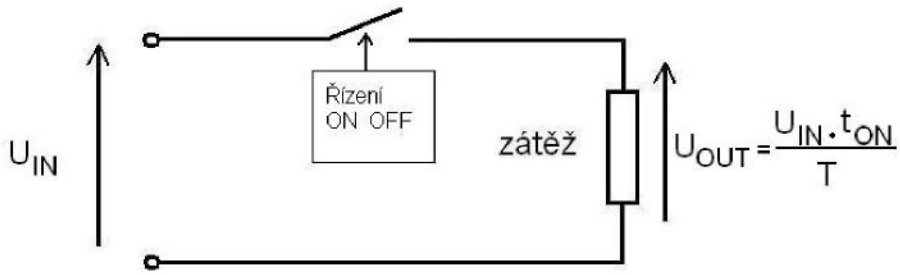


Obr. 2: Náhradní schéma lineárního regulátoru



Obr. 3: Skutečné schéma lineárního regulátoru

Naproti tomu spínané regulátory (Obr. 4) nemaří přebytečný výkon na své spínací jednotce, ale omezují jeho přísun do zátěže střídou spínání napájení. Srovnání obou typů zdrojů je přehledně uvedeno níže (Tab. 2).



Obr. 4: Náhradní schéma spínaného regulátoru

Z uvedené tabulky je patrné, že rozdíly mezi spínanými a lineárními zdroji nejsou zanedbatelné. Co se týče objemových a energetických úspor, dosahují lepších parametrů jednoznačně spínané zdroje. Z hlediska kvality výstupního napětí mají však lepší parametry zdroje lineární.

Tab. 2: Porovnání lineárních a spínaných zdrojů

parametr	spínaný zdroj	lineární zdroj
účinnost	75%	30%
velikost	0,2[W/cm^3]	0,05[W/cm3]
váha	100[W/Kg]	20[W/Kg]
výstupní zvlnění	50[mV]	5[mV]
šumové napětí	200[mV]	50[mV]
odezva na skok	1[ms]	20[µs]
doba náběhu	20[ms]	2[ms]
cena	přibližně konstantní	roste s výkonem

2.1.1 Výhody spínaných zdrojů

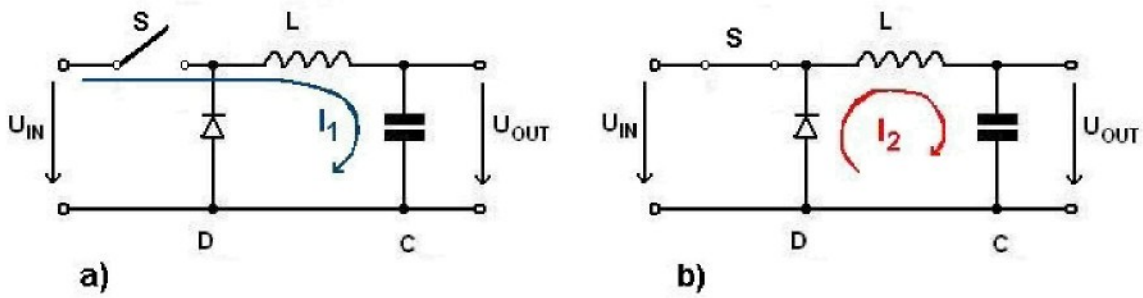
Největší výhodou je vysoká účinnost pohybující se běžně v rozmezí od 70 % do 80 %. Pro srovnání, obvyklá účinnost lineárních stabilizátorů s podobnými parametry se pohybuje kolem 30 %. Další neméně důležité výhody s sebou přináší vysoká pracovní frekvence (10 kHz až 1 MHz), která je tak snadno odfiltrovatelná. Díky této frekvenci přenáší transformátor energii po mnohonásobně nižších dávkách, může tak být daleko menší a levnější. Z principu činnosti se u spínaných zdrojů neobjevují takové problémy s odvodem tepla.

2.1.2 Nevýhody spínaných zdrojů

S vysokou pracovní frekvencí souvisí také nevýhody spínaných zdrojů. Na použité součástky, které musí při takovýchto frekvencích spolehlivě pracovat, jsou kladeny vyšší nároky (mezní kmitočty tranzistorů a diod, rozptylové kapacity transformátorů a stejnosměrné odpory elektrolytických kondenzátorů). Další nevýhodou, provázející návrh těchto zdrojů, je jejich složitost.

2.2 Princip funkce

Jak již bylo řečeno, regulují spínané zdroje energii dodávanou do zátěže střídou spínání napájení. Jako spínací prvek bývá zpravidla použit výkonový MOSFET tranzistor. Je-li však požadováno, lze použít i IGBT nebo tyristor. Nejdůležitější komponentou, z hlediska funkce, je u spínaných zdrojů transformační člen (indukčnost nebo transformátor). Na její volbě závisí správná funkce zdroje. Princip funkce spínaného zdroje je vysvětlen na jednoduchém příkladu. Předpokládejme jednoduchý obvod (schéma na Obr. 5).



Obr. 5: Princip funkce spínaného zdroje

Sepneme-li spínač *S*, začne téct indukčností *L* proud I_1 (Obr. 5a). Ten narůstá podle vztahu (Eq. 1), kde I_{01} je proud cívkou na počátku intervalu. Je patrné, že strmost

růstu proudu I_1 je dána velikostí indukčnosti cívky L a napětí na této cívce. Proud roste po celou dobu sepnutí spínače (dále jen T_{ON}).

$$I_1 = I_{01} + \int_0^{T_{ON}} \left(\frac{U_{IN}-U_{OUT}}{L}\right) dt \quad \text{(Eq. 1)}$$

Rozepneme-li nyní spínač (Obr. 5b), bude se cívka snažit zachovat směr a velikost proudu I_1, v důsledku čehož se na ní, dle vztahu (Eq. 2), naindukuje napětí U_{IND}.

$$U_{IND} = -L \cdot \frac{dI}{dt} \quad \text{(Eq. 2)}$$

Toto indukované napětí vyvolá tok proudu I_2, jehož velikost se zmenšuje dle vztahu (Eq. 3), kde I_{02} je proud cívkou na počátku intervalu. Strmost klesání I_2 závisí na velikosti U_{IND} a naopak velikost U_{IND} závisí na strmosti klesání I_2.

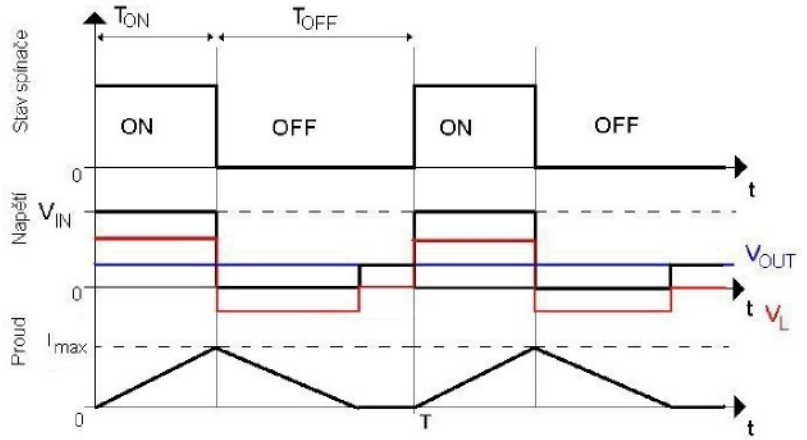
$$I_2 = I_{02} - \int_0^{T_{OFF}} \left(\frac{U_{OUT}}{L}\right) dt \quad \text{(Eq. 3)}$$

Proud klesá po celou dobu rozepnutí spínače (dále jen T_{OFF}), pokud ovšem dříve nedosáhne nuly. Dosáhne-li I_2 nuly dříve, než opět nastane T_{ON}, pracuje obvod v takzvaném přerušovaném režimu. Nastane-li T_{ON} dříve, než dosáhne I_2 nuly pracuje obvod v takzvaném souvislém režimu. Je-li spínač sepnut přesně ve chvíli, kdy dosáhne I_2 nuly, pracuje obvod v kritickém režimu.

2.2.1 Přerušovaný režim

Jak bylo již výše naznačeno, pracuje obvod v přerušovaném režimu, pokud proud indukčností klesl na nulu ještě před opětovným sepnutím spínače. Jinými slovy, veškerá naakumulovaná energie v indukčnosti za dobu T_{ON} byla dodána do zátěže před sepnutím spínače.

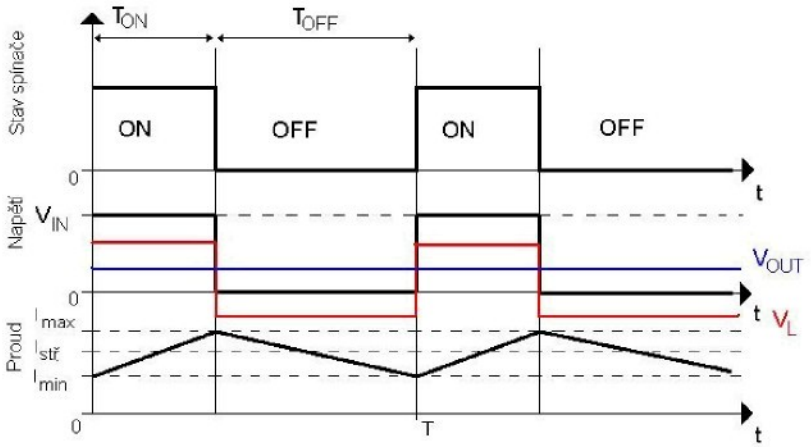
Činnost měniče v tomto režimu s sebou přináší několik výhod. Vzhledem k tomu, že cívka v každém cyklu odevzdá veškerou svou naakumulovanou energii do zátěže, tj. žádná energie se v cívce nestřádá, je možno použít cívku s nižší hodnotou indukčnosti, než u ekvivalentního obvodu pracujícího v souvislém režimu. Nižší indukčnost cívky má za následek i rychlejší odezvu na změny na výstupu. Průběhy napětí a proudu v přerušovaném režimu jsou na obrázku (Obr. 6)



Obr. 6: Přerušovaný režim – průběhy

2.2.2 Souvislý režim

Pracuje-li obvod v souvislém režimu, znamená to, že proud cívkou neklesl do okamžiku sepnutí spínače až na nulu. V cívce je tak stále naakumulovaná energie. Je tedy třeba použít cívku s vyšší hodnotou indukčnosti. Tím se zpomaluje odezva na změny na výstupu. Nese to s sebou ale i určitou výhodu. Proud indukčností je nižší, než u přerušovaného režimu. Průběh napětí a proudu v souvislém režimu je na obrázku (Obr. 7).



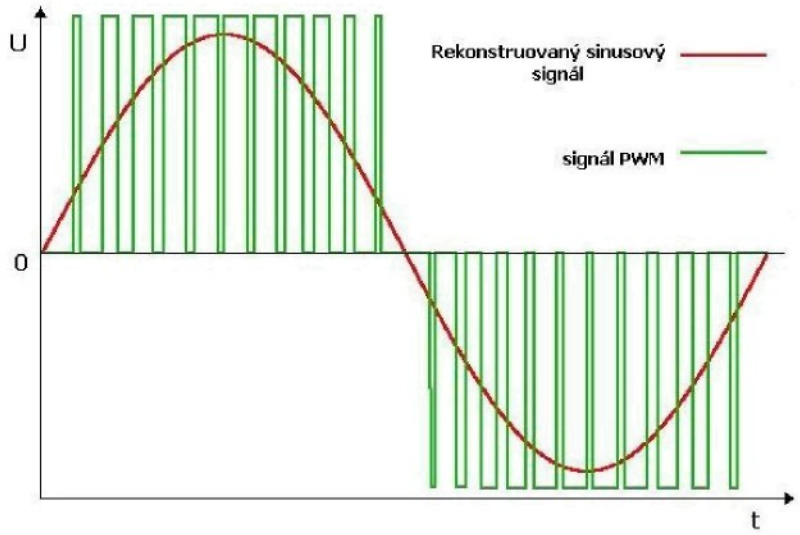
Obr. 7: Souvislý režim – průběhy

2.2.3 Kritický režim

Tento režim spojuje výhody obou předešlých. Dosáhne-li proud indukčností nuly, je spínač sepnut. Tím je zajištěno, že cívka dodá do zátěže všechnu svoji akumulovanou energii a zároveň zde nevzniká žádný mrtvý čas (doba mezi časem kdy dosáhne proud nuly a časem opětovného sepnutí spínače).

2.3 PWM modulace

Pulzně šířková modulace je princip řízení, založený na modulaci požadovaného napětí na šířku pulzu. Přičemž je po průchodu tohoto signálu dolnopropustním filtrem střídavá složka odfiltrována. Na výstupu se tak objeví pouze střední hodnota signálu. Šířkou pulzu je tak možné dosáhnout požadovaného napětí. Pro názornost je na obrázku (Obr. 8) ukázána rekonstrukce části sinusového signálu pomocí PWM modulace.



Obr. 8: Rekonstrukce sinusového signálu pomocí PWM

Výhody tohoto principu jsou zřejmé. Tranzistor pracuje pouze v režimech sepnutí a rozepnutí. V režimu sepnutí sice teče tranzistorem vysoký proud, avšak v tomto režimu má tranzistor minimální přechodový odpor, úbytek na něm je tak minimální. V režimu rozepnutí teče tranzistorem jen minimální zbytkový proud. Největší ztráty vznikají při přechodech mezi těmito dvěma stavy. Přechody neprobíhají nekonečně krátkou dobu, trvají určitý čas (řádově μs). Během tohoto času se tranzistor pohybuje v téže oblasti charakteristiky jako ve spojitých zesilovačích. Je tedy zřejmé, že s rostoucí frekvencí spínání, ztráty na tranzistoru rostou. I přes to jsou ztráty oproti spojité regulaci mnohonásobně nižší.

3 Základní topologie spínaných zdrojů

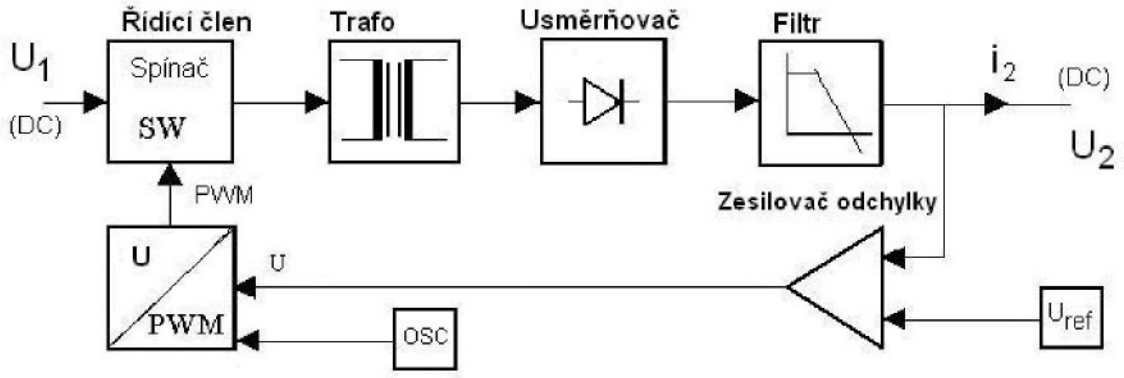
Na úvod je třeba pomocí obecného blokového schématu nastínit princip funkce spínaných zdrojů. Princip je u všech zapojení shodný. Jednotlivé topologie se liší především ve způsobu přenosu energie z primárních obvodů do sekundárních.

Existuje celá řada různých typů zapojení. Každé z nich má své specifické vlastnosti, které ho předurčují pro určité aplikace. Některá jsou vhodnější pro nízký výstupní výkon, některá pro vyšší. Některá jsou lepší pro vysoké střídavé napětí (220V), některá pro nižší (120V nebo méně). Některá je výhodnější použít pro vyšší výstupní stejnosměrné napětí nebo tam, kde je potřeba mít více různých výstupních napětí. Časté kritérium při výběru typu zapojení je zvlnění výstupu a šum.

Základem pro dobrý návrh zdroje je určení podmínek a potřebných parametrů a na základě toho výběr nejoptimálnější topologie.

3.1 Blokové schéma

Podmínkou funkce spínaného zdroje je stejnosměrné vstupní napětí. Pokud je zdroj napájen střídavým napětím, je třeba do blokového schématu (Obr. 9) ještě zahrnout vstupní usměrňovač a filtr. Vstupní filtr má zde především funkci akumulátoru energie. Musí mít odpovídající kapacitu, aby byla zaručena energetická kompenzace při krátkodobém podpětí.



Obr. 9: Blokové schéma spínaného zdroje

Vzhledem k tomu, že transformovat lze jen střídavé napětí, je třeba převést U_1 na střídavý tvar. K tomu zde slouží spínač (vysokofrekvenční tranzistor), který převede stejnosměrné napětí na obdélníkový průběh. Řízení spínače (regulace požadovaného napětí) bylo vysvětleno v kapitole PWM modulace.

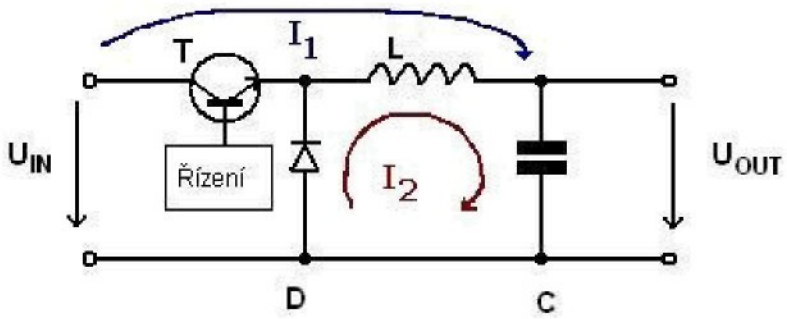
Vlastní transformace probíhá buď na indukčnosti, nebo na transformátoru. Transformované napětí je třeba dále usměrnit a vyfiltrovat od střídavé složky. Na rozdíl od vstupního usměrňovače jsou zde kladeny vysoké nároky na usměrňovací diody, které musí spolehlivě pracovat při vysokých pracovních kmitočtech. Na výstupní filtr už tak vysoké požadavky kladeny nejsou, protože takto vysoký kmitočet lze snadno a velmi účinně odfiltrovat.

3.2 Rozdělení podle funkce

Spínané zdroje lze na základě jejich funkce dělit na řadu zapojení. Nejzákladnější z nich jsou: snižující, zvyšující a invertující měnič. Vysvětlení jejich funkce a principu činnosti je uvedeno níže.

3.2.1 Snižující měnič (BUCK, Step-Down)

Jedná se o nejranější topologii vůbec. Schéma je na (Obr. 10). Spínač, ve formě tranzistoru T, je zde v sérii se vstupem U_{IN} a indukčností L, která má zde spolu s kondenzátorem C funkci filtru a akumulátoru energie.



Obr. 10: Snižující měnič

Pokud je spínač T sepnutý, roste proud I_1 indukčností L dle vztahu (Eq. 1). Kondenzátor C je tímto proudem nabíjen a roste na něm napětí U_C dle vztahu (Eq. 4). Po rozepnutí spínače se indukčnost snaží zachovat směr i velikost proudu. Energie naakumulovaná v indukčnosti za dobu, co byl spínač sepnutý, se mění na proud I_2. Aby tento proud mohl téct, musí být uzavřený obvod. K tomu zde slouží dioda D.

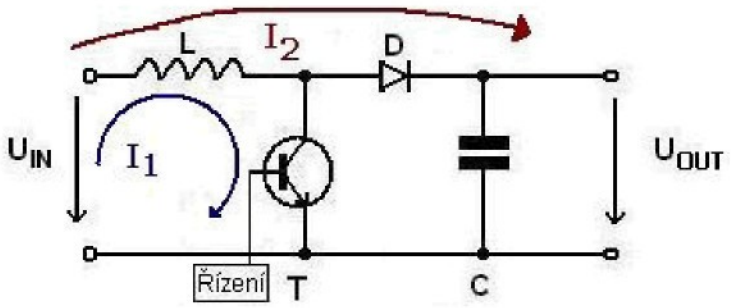
$$U_C = \int_0^t \frac{I_1}{C} dt \qquad \text{(Eq. 4)}$$

Je patrné, že po dobu T_{ON} napětí na kondenzátoru roste, zatímco po dobu T_{OFF} klesá. Je-li však frekvence spínání a rozpínání dostatečná, lze tuto střídavou složku snadno odfiltrovat. Je také patrné, že čím delší bude T_{ON} a čím kratší T_{OFF}, tím vyšší bude napětí na výstupu. Totéž platí i naopak, chceme-li napětí snížit, zkrátíme T_{ON} nebo prodloužíme T_{OFF}. Pro výstupní napětí platí vztah (Eq. 5). Kde T je perioda spínání. Z uvedeného vztahu vyplývá, že toto zapojení může napětí pouze snižovat.

$$U_{OUT} = \frac{U_{IN}.T_{ON}}{T} \qquad \text{(Eq. 5)}$$

3.2.2 Zvyšující měnič (BOOST, Step-Up)

Předchozí topologii můžeme použít pouze tam, kde je třeba napětí snížit. Potřebujeme-li aplikaci, kde je potřeba vyšší napětí než je napájecí, lze použít zapojení zvyšujícího měniče (Obr. 11).



Obr. 11: Zvyšující měnič

Indukčnost L je zde opět v sérii s U_{IN}. Spínací tranzistor T je připojen paralelně k zemi. Dioda D brání vybíjení kondenzátoru C přes sepnutý tranzistor do země. Je-li spínač T sepnutý, je dioda orientována závěrně. Proud I_1 tak teče pouze indukčností a spínačem T. Hodnota proudu, dle vztahu (Eq. 1), narůstá a energie se akumuluje v magnetickém poli cívky. Po dobu sepnutí spínače je proud do zátěže dodáván z kondenzátoru C. Ten musí být dost velký na to, aby tento proud dokázal dodávat s vyhovujícím zvlněním.

Je-li v tomto okamžiku spínač rozepnut, snaží se indukčnost zachovat směr a velikost proudu I_1 a vzniká na ní indukované napětí podle vzorce (Eq. 2). Jelikož je cívka v sérii s U_{IN}, tak se obě napětí sčítají. Dioda D je tímto orientována propustně, což umožňuje průtok proudu I_2 do kondenzátoru C. Tento proud hradí jak zátěžový proud, tak dobíjí kondenzátor C. Velikost indukovaného napětí závisí, dle vzorce, na hodnotě indukčnosti, velikosti proudu I_1 a rychlosti rozepnutí. Může tedy nabývat

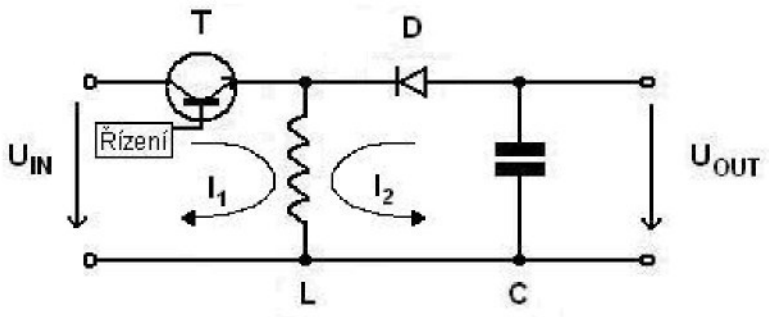
vysokých hodnot. Je zřejmé, že součtem vstupního napětí U_{IN} a napětí indukovaného na cívce, nám vždy vyjde hodnota vyšší, než je hodnota U_{IN}. Tento typ zapojení tedy napětí pouze zvyšuje.

V předchozím případě jsme zvyšováním doby T_{ON} zvyšovali napětí na výstupu. Zde už to tak jednoznačně neplatí. Zvýšením doby T_{ON} energie akumulovaná v cívce roste, čímž roste i indukované napětí. Poroste-li však T_{ON} příliš, kondenzátor dodávající po tuto dobu proud do zátěže se bude vybíjet a napětí na něm poklesne. V součtu se tedy může stát, že zvýšením T_{ON} napětí na výstupu poklesne. Pro výstupní napětí zvyšujícího měniče platí vztah (Eq. 6)

$$U_{OUT} = \frac{U_{IN} \cdot T}{t_{OFF}} \qquad \text{(Eq. 6)}$$

3.2.3 Invertující měnič

Jak je již z názvu patrné, tento měnič napětí invertuje. Schéma zapojení je na obrázku (Obr. 12).



Obr. 12: Invertující měnič

Spínací tranzistor T je v sérii se vstupním napětím, indukčnost je paralelně na zem. Dioda D brání přebíjení kondenzátoru na kladnou polaritu během doby T_{ON}. Je-li spínač sepnutý, teče jím a cívkou proud I_1. Ten roste až do doby, než je spínač rozepnut. Po rozepnutí spínače se snaží cívka zachovat směr a velikost proudu. Naindukuje se na ní napětí opačné polarity, než jaká na ní byla po dobu T_{ON}. Na toto napětí je tedy kondenzátor C dobíjen proudem I_2. Z rovnice (Eq. 7) vyplývá, že výstupní napětí může být jak vyšší tak nižší, než je napětí napájecí.

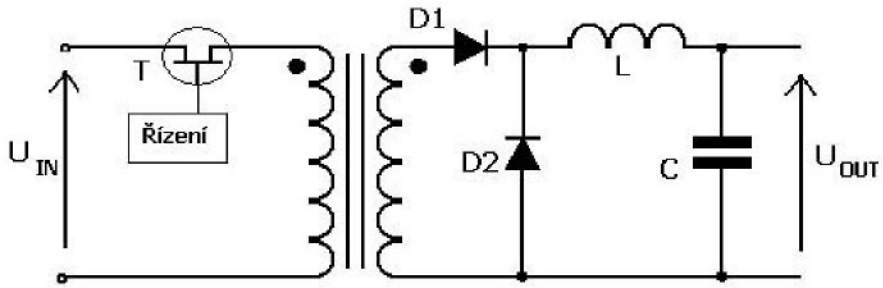
$$U_{OUT} = \frac{-U_{IN} \cdot t_{ON}}{t_{OFF}} \qquad \text{(Eq. 7)}$$

3.3 Rozdělení podle způsobu přenosu energie

Podle způsobu přenosu energie z primárních obvodů do sekundárních lze spínané zdroje rozdělit na několik základních typů. Je zbytečné představovat všechny, budou vysvětleny dva nejzákladnější. Propustný a blokující měnič. Ty je možné dále dělit podle použitého transformačního členu, a to na zdroje s transformátorem nebo indukčností. Vzhledem k tomu, že cílem této práce je navrhnout zvyšující měnič s poměrně značným rozdílem mezi vstupním a výstupním napětím, bude jako transformační člen použit transformátor. Z toho důvodu budou vysvětleny pouze topologie blokujícího a propustného měniče s transformátorem.

3.3.1 Propustný měnič (FORWARD)

Schéma propustného měniče je na obrázku (Obr. 13). Tato topologie je široce používána pro výstupní výkony od 150 do 200W. Tranzistor je zapojen do série s primárním vinutím transformátoru. Tečky vyznačují počátek vinutí. Dioda D znemožňuje vybíjení kondenzátoru přes vinutí transformátoru. Indukčnost L zde spolu s kondenzátorem C tvoří LC filtr, jež má zároveň funkci zásobníku energie.



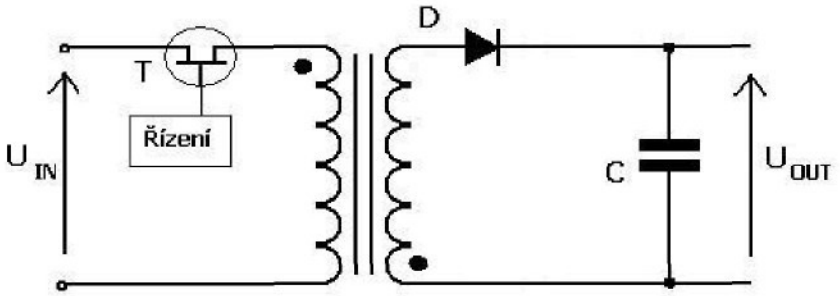
Obr. 13: Propustný měnič

Jak napovídá název, energie se u tohoto zapojení přenáší na výstup přímo. To znamená, že pokud je spínač sepnut, teče proud jak primárním vinutím, tak i sekundárním. Takovéto chování určuje vzájemná orientace teček (počátků vinutí) a orientace polarity diody. Jednou z nevýhod této topologie je nutnost výstupní indukčnosti L.

3.3.2 Blokující měnič (FLYBACK)

Schéma je na obrázku (Obr. 14). Všechny topologie, vyjma zvyšujícího a invertujícího měniče, dodávají energii do výstupu během času T_{ON}. Blokující měnič funguje opačně. Po dobu T_{ON} akumuluje energii do magnetického pole cívky. Veškerý odebíraný proud je po tuto dobu dodáván z výstupního kondenzátoru C. Tuto funkci

zajišťuje vzájemná orientace počátků vinutí transformátoru a orientace polarizace diody. Po rozepnutí spínacího tranzistoru se díky energii naakumulované v transformátoru za dobu T_{ON}, naindukuje na vinutích napětí opačných polarit. Dioda D se tedy stává vodivou. Vzniká proud I_2, který dobíjí kondenzátor C a pokrývá zátěžový proud.



Obr. 14: Blokující měnič

Tato topologie je široce užívána pro vysoká výstupní napětí při relativně malém výkonu ($5000\,V < 15\,W$). Největší výhodou blokujícího měniče je, že výstupní indukčnost potřebná u všech propustných měničů, zde potřebná není. Toto je výhodné zvláště v případech, kdy je potřeba mít více výstupních vinutí. Z tohoto plyne významná finanční a objemová úspora. Z nepotřebnosti výstupní indukčnosti plyne další nesporná výhoda, u zdrojů s více výstupy se napětí a změny zátěže navzájem lépe sledují.

4 Návrh DC/DC měniče

Tato kapitola se zabývá vlastním návrhem DC/DC měniče. Důležité je nejprve stanovit požadované vlastnosti a parametry zdroje. Dle toho pak vybírat transformátor, řídící obvod, spínací prvek a výstupní, popřípadě vstupní filtr. Topologie, v jaké má být zdroj navržen, je dle zadání blokující měnič.

4.1 Stanovení parametrů zdroje

Vzhledem k tomu, že je měnič určen pro napájení piezoelektrických aktuátorů, musí splňovat podmínky dané jejich potřebami. Nejprve je třeba stanovit výstupní napětí. Jak již bylo výše uvedeno, pohybuje se potřebné napětí běžně od 100 do 180 V. Je tedy samozřejmě lepší stanovit požadované výstupní napětí někde v okolí vyššího z uvedených. Bylo tedy zvoleno 180 V. Tento parametr ovlivňuje jak výběr transformátoru, tak výběr výstupní diody a výstupního kondenzátoru.

Dále je třeba stanovit výkon. Výkony piezoelektrických aktuátoru se pohybují ve velmi širokém pásmu od 0,001 W do 1 kW. Po dohodě s vedoucím práce, ohledně nároků na konstruovaný měnič, byl zvolen požadovaný výkon 10 W. Tímto parametrem je ovlivňován především výběr transformátoru.

Důležitý parametr, již daný zadáním, je napájecí napětí měniče. To by se mělo pohybovat v intervalu 8 až 25 V. Pro toto rozmezí napájecích napětí musí tedy být zaručena funkčnost zkonstruovaného zdroje.

Výstupní zvlnění je jakýmsi parametrem kvality výstupu spínaného zdroje. Čím je zvlnění nižší, tím kvalitnější měnič je. Není však dobré stanovit zvlnění o moc nižší, než je pro zadanou aplikaci potřeba. Byly by tak kladeny zbytečně vysoké nároky na výstupní filtr.

Stanovené parametry jsou z důvodu přehlednosti uvedeny v tabulce (Tab. 3).

Tab. 3: Stanovené parametry

parametr	hodnota
výstupní napětí	180 V
vstupní napětí	8 - 25 V
výstupní proud	0,5 A
výkon	10 W
zvlnění na výstupu	200 mV

4.2 Výběr transformátoru

Vzhledem k tomu, že je transformátor jedna z nejdůležitějších součástí měniče, je třeba dbát při jeho návrhu zvýšené pozornosti. Určujícími faktory jsou: potřebný výstupní výkon zdroje, maximální proud primárním a sekundárním vinutím, požadovaný transformační poměr, topologie v jaké bude zdroj navržen a potřebná velikost indukčnosti.

Dle zadání má být ovšem měnič zkonstruován s komerčně vyráběným transformátorem, je tedy zbytečné provádět jeho návrh. Musí být vybrán ze široké škály výrobků. Jelikož je zdroj navrhován ve flyback topologii, ostatně v této výkonové třídě se trafa pro propustné měniče téměř nevyrábějí, je nutné omezit výběr transformátoru na flyback transformátory. Tato skupina se od ostatních odlišuje jinou konstrukcí jádra, což je dáno principem funkce flyback topologie. Jak již bylo výše uvedeno, veškerá energie se za dobu T_{ON}, u blokujícího měniče, akumuluje v magnetickém poli transformátoru. Jádro tak musí mít hodnotu sycení dostatečně vysokou, aby bylo schopno tuto energii akumulovat a nedostalo se přitom do saturace.

Dalším velmi důležitým kritériem, které musí být splněno, jsou výkonové a elektrické specifikace. Výkon transformátoru musí být vyšší než výkon zdroje (10 W). Dále musí trafo splňovat proudové a napěťové limity dané návrhem.

Primární vinutí musí být dimenzováno na proud vypočtený z následujícího vztahu (Eq. 8).

$$I_{IN} = \frac{P_{OUT_{MAX}}}{U_{IN_{MIN}}} = \frac{10}{5} = 2\,A \qquad \text{(Eq. 8)}$$

Sekundární vinutí musí být dimenzováno na výstupní proud daný návrhem (0,5 A). Napětí jak primárního tak sekundárního vinutí musí být samozřejmě vyšší, než jsou hodnoty vstupního a výstupního napětí navrhovaného zdroje.

4.2.1 Vybraný transformátor

Po důkladném zvážení kritérií uvedených výše, byl vybrán transformátor MYRRA 74010. Jeho parametry jsou přehledně uvedeny v tabulce (Tab. 4). Transformátor má čtyři vinutí: primární Pri, sekundární S1 a S2 a regulační vinutí Aux. I když byl původně transformátor vybírán bez ohledu na to, zda má, či nemá regulační vinutí, ukázalo se, že lze tohoto vinutí při návrhu elegantně využít.

Z důvodu návrhu zvyšujícího měniče, budou při konstrukci funkce některých vinutí prohozeny. Původně sekundární S1 a S2 budou spojena a použita jako primární vinutí, naopak původně primární vinutí Pri bude použito jako sekundární. Funkce regulačního vinutí bude zachována.

Tab. 4: Parametry transformátoru MYRRA 74010

maximální výstupní výkon	vinutí	Piny	počet závitů	maximální proud	indukčnost
12 W	Pri	4 – 6	120	0.55	1660 µH
	Aux	2 – 1	14	0.1	
	S1	9 – 10	7	2	
	S2	7 – 8	17	1	

4.3 Výběr řídícího obvodu

Existuje nepřeberné množství řídících obvodů pro spínané zdroje. Liší se cenou, způsobem regulace napětí, počtem pinů, integrovanými ochranami, počtem potřebných externích součástek atd. Při výběru je třeba zvážit veškerá pro a proti vybíraného obvodu. Jelikož je na trhu možno vybírat z ohromného množství různých obvodů, je třeba nejprve určit parametry, podle kterých bude obvod vybírán.

Po důkladném zvažování vstoupily do užšího výběru, jako potenciálně vhodné, dva diametrálně odlišné řídící obvody. LT1172 od firmy Linear Technology a TDA4605-3 od firmy Infineon Technologies.

Jak již bylo řečeno, jedná se o naprosto rozdílné obvody. LT1172 je pětipinový obvod s integrovaným výkonovým tranzistorem. Ke své funkci potřebuje minimální počet externích součástek. Další výhodou tohoto obvodu je skutečnost, že firma Linear Technology vyvinula návrhový program Switcher Cad, v němž je možno si zapojení spínaného zdroje navrhnout a odsimulovat. To lze bohužel učinit pouze s jejich výrobky. Nevýhoda popisovaného obvodu spočívá v jeho poměrně vysoké ceně.

Naproti tomu TDA4605-3 je osmipinový, poměrně levný řídící obvod, určený především pro napájecí zdroje televizorů. Ke své funkci potřebuje větší množství externích součástek, než je tomu u LT1172, a to včetně externího tranzistoru. Potřeba externího tranzistoru je však v případě tohoto návrhu spíše výhodou. Dojde-li totiž ke zkratu či jinému destrukčnímu činiteli, postihne destrukce nejpravděpodobněji právě tranzistor, který pokud je externí, lze snadno vyměnit. Navíc je jeho cena ve srovnání

s obvody s interním tranzistorem výrazně nižší. Další nespornou výhodou tohoto obvodu je velké množství integrovaných ochran.

4.3.1 Vybraný řídící obvod

I přes četné výhody obvodu LT1172 byl nakonec vybrán TDA4605-3. Rozhodujícími faktory, které sehrály při výběru zásadní roli, byly: nízká cena obvodu, potřeba externího tranzistoru, spousta ochran a skutečnost, že byl obvod přímo k dispozici. Elektrické specifikace jsou uvedeny v tabulce (Tab. 5).

Tab. 5: Parametry řídícího obvodu TDA4605-3

značka	parametr	hodnota	jednotka
UIN_{max}	maximální napájecí napětí	20	V
U6E	napětí potřebné pro start	12	V
U6A	vypínací napětí	5	V
$U6_{max}$	minimální pracovní napětí	7,25	V
$U6_{min}$	maximální pracovní napětí	16	V
U1R	referenční napětí zpětné vazby	400	mV
I5	maximální výstupní proud	1,5	A
I2B	napětí pinu pro simulaci primárního proudu	1	V
Tstg	maximální provozní teplota	-40, 125	°C

Obvod má implementovány následující ochrany:

- ochrana proti přetížení externích součástek
- ochrana proti přerušené či zkratované regulační smyčce
- vypnutí v případě, že je napájecí napětí příliš nízké či vysoké
- soft start – pro omezení proudových nárazů vznikajících při náběhu měniče do pracovního režimu
- tepelná ochrana a potlačení parazitních oscilací transformátoru.

Vzhledem k tomu, že je tento integrovaný obvod konstruován pro široké spektrum zatížení, je vhodný jak pro průmyslové, tak pro spotřebitelské aplikace. Obvod funguje následovně. Externí tranzistor a primární vinutí transformátoru jsou zapojeny sériově. Během času T_{ON} je tranzistor zapnut a energie je tak ukládána v transformátoru. Během doby T_{OFF} zásobuje transformátor prostřednictvím sekundárního vinutí zátěž. Integrovaný obvod řídí tranzistor proměnlivým časem T_{ON}. Tím je ovládaný přenos energie na sekundární stranu transformátoru tak, aby bylo výstupní napětí téměř nezávislé na zátěži. Potřebné informace jsou snímány z vstupního napětí v čase T_{ON}

a regulačního vinutí v čase T_{OFF}. Nový cyklus začne teprve tehdy, bude-li veškerá energie, uložená v transformátoru za dobu T_{ON}, převedena do zátěže.

Jedním z důvodů, proč je možné tento obvod využít pro tak širokou škálu zátěží je skutečnost, že může pracovat ve čtyřech režimech činnosti:

- **bez zátěže** – obvod pracuje na frekvenci 20 – 40 kHz, přičemž výstupní napětí může být trochu vyšší nebo nižší než je tomu u běžného provozu.
- **běžný provoz** – frekvence spínání se snižuje s rostoucí zátěží a snižujícím vstupním napětím.
- **bod přetížení** – v tomto bodě je nejvyšší dovolený výstupní výkon
- **přetížení** – energie transformovaná za pracovní cyklus je omezena, proto klesne výstupní napětí o sekundární přetížení.

Další vlastnosti IO a funkce jednotlivých pinů budou vysvětleny při popisu schématu.

4.4 Výběr tranzistoru

Tranzistor by měl být vybírán především na základě maximálního povoleného závěrného napětí, maximálního špičkového proudu, maximální povolené provozní teploty a velikosti celkových výkonových ztrát.

Maximální napětí D-S (U_{DSmax}) musí být vyšší než $\mathbf{U_{max}}$ spočtené podle tohoto vztahu (Eq. 9):

$$U_{max} = U_{IN_{max}} + \frac{(U_{out}+U_d)}{p} = 25 + \frac{(180+0{,}5)}{5} = 61{,}1\ V \qquad \text{(Eq. 9)}$$

Kde p je transformační poměr primárního vinutí k sekundárnímu a U_d je napětí na výstupní diodě v propustném směru. Napětí $U_{DS_{max}}$ je potřeba ještě o 20 až 30 % předimenzovat. Výsledné napětí, na které musí být tranzistor dimenzován, je tedy: $\mathbf{U_{DS_{max}} \cdot 1{,}3 = 80\ V}$.

Maximální proud kolektoru (I_d) musí být vyšší než napájecí proud zdroje vypočtený ze vztahu (Eq. 8). Maximální napětí G-S (U_{GSmax}) musí vyhovovat řídícímu obvodu. TDA4605-3 řídí tranzistor napětím, které má hodnotu napájecího napětí obvodu, tedy až 20 V. Zároveň musí být vyhovující prahové napětí. To by mělo být dostatečně nízké, aby byla vyloučena možnost operace tranzistoru v lineární oblasti.

Základním úkolem při výběru tranzistoru je minimalizovat jeho provozní teplotu minimalizací výkonových ztrát. Ty totiž určují účinnost zdroje, přičemž čím nižší tyto ztráty jsou, tím vyšší je účinnost. Kromě maximálního napětí D-S, maximálního proudu kolektoru a maximálního napětí G-S uvedených výše, jsou tedy další důležité parametry: Rds(on), prahové napětí a parazitní kapacita.

Na spínacím tranzistoru vznikají tyto tři typy ztrát: vodivostní ztráty, spínací ztráty a kapacitní ztráty.

- **Vodivostní ztráty** jsou ztráty způsobené průtokem proudu tranzistorem přes odpor Rds_{on}. Jsou rovny: $\boldsymbol{I \cdot R^2}$, proto by měl být odpor tranzistoru mezi S-D v sepnutém stavu (Rds_{on}), co nejnižší.
- **Spínací ztráty** jsou rovny: $\boldsymbol{t_{sw} \cdot U_{DS} \cdot I \cdot f}$. Kde $\boldsymbol{t_{sw}}$ je spínací čas, který je funkcí G-D Millerovy konstanty (Q_{gd}), vnitřního odporu řízení tranzistoru a prahového napětí Ugs_{th} (minimálního napětí na svorce G, při kterém je tranzistor sepnut).
- **Kapacitní ztráty** jsou způsobené nabíjením kapacity Qg_{tot} a jejím rychlým vybíjením do země v každém cyklu. Jsou rovny: $\boldsymbol{f \cdot Q_{g(tot)} \cdot U_{DR}}$.

Jelikož je tranzistorů na trhu nepřeberné množství, probíhal výběr následovně: Nejprve bylo, na základě podmínek daných návrhem, vybráno několik potenciálně vhodných tranzistorů (IRF530, IRF540, IRFD120, IRLR014, IRLR120 a STP16N010). Z této skupinky bylo dále vybíráno na základě co nejnižších výkonových ztrát. Spočtené ztráty jednotlivých tranzistorů jsou pro srovnání uvedeny v tabulce (Tab. 6). Uvedené hodnoty byly spočteny pro proud 2 A, napětí 100V a střídu 0,5.

Tab. 6: Ztráty na zvolených tranzistorech

	ztráty				
tranzistor	vodivostní	spínací	kapacitní	celkem	jednotka
IRF530	0,32	0,42	0,016	0,76	Ω
IRF540	0,14	0,37	0,009	0,52	Ω
IRFD120	0,60	0,21	0,025	0,84	Ω
IRLR014	0,40	0,21	0,005	0,61	Ω
IRLR120	0,54	0,24	0,007	0,78	Ω
STP16N06	0,28	0,33	0,018	0,63	Ω

4.4.1 Vybraný tranzistor

Na základě výše provedené analýzy, byl jako nejvhodnější tranzistor vybrán IRF540. Jeho parametry jsou ukázány v tabulce (Tab. 7).

Tab. 7: Parametry tranzistoru IRF540

značka	parametr	hodnota	jednotka
U_{DSmax}	maximální D -S napětí (při VGS = 0)	100	V
U_{GS}	maximální G -S napětí	+- 20	V
ID	maximální proud kolektorem (souvislý)	30	A
IDM	maximální proud kolektorem (pulzní)	120	A
P_{tot}	celkové výkonové ztráty při T = 25°C	100	W
T_{STG}	maximální povolená teplota	220	°C
Tj	maximální provozní teplota	-55 až 175	°C
R_{DSon}	odpor v sepnutém stavu	0,065	Ω
$U_{GS(th)}$	Prahové napětí	3	V

Vzhledem k tomu, že je tranzistor poměrně vysoce nadhodnocen, by měl být velmi odolný proti destrukčním činitelům. Jelikož je zdroj konstruován pro výstupní výkon 10 W, neměl by být žádný problém s jeho chlazením.

4.5 Výběr výstupního filtru

Vzhledem k tomu, že je zdroj navrhován ve flyback topologii, není potřeba zařadit na výstup indukčnost. Jako výstupní filtr postačí kondenzátor. K jeho výběru jsou určující dvě různá kritéria:

- Minimální kapacita potřebná k stanovenému zvlnění vychází ze vzorce $I = C \cdot dV/dt$ kde dt je čas T_{OFF} a dV je 25 % stanoveného zvlnění. [8]
- Ekvivalentní sériový odpor (ESR) kondenzátoru by měl způsobovat méně než 75 % výstupního zvlnění ($V_{zv} = dI \cdot ESR$). [8]

Kapacita kondenzátoru tedy musí být větší než:

$$C = \Delta Is \cdot \frac{T_{OFF}}{V_{zv} \cdot 0.25} = 0{,}11 \cdot \frac{3 \cdot 10^{-5}}{0{,}2 \cdot 0.25} = 51\ \mu F$$

Ekvivalentní sériový odpor musí být menší než:

$$ESR = \frac{V_{zv} \cdot 0.75}{\Delta Is} = \frac{0{,}2 \cdot 0.75}{0{,}11} = 1{,}36\ \Omega$$

Vzhledem k nepatrnému rozdílu mezi hodnotami 47 a 51 μF byla zvolena hodnota 47 μF. Pokud by bylo meřením zjištěno, že je tato hodnota nevyhovující, lze kondenzátor bez problémů vyměnit.

Mimo kritéria uvedená výše je bezpodmínečně nutné, aby měl kondenzátor jmenovitou hodnotu provozního napětí vyšší, než bude výstupní napětí měniče.

4.6 Výběr výstupní diody

Jelikož je měnič navržen jako zvyšující a jeho výstupní napětí by mělo dosahovat až sto osmdesáti voltů, je třeba se zaměřit i na výběr vhodné diody. Určující parametry pro výběr jsou: maximální napětí v závěrném směru, maximální propustný proud, kapacita v propustném směru a doba zotavení.

Maximální napětí v závěrném směru diody musí být vyšší než:

$$Vd_{max} = Vi_{max} \cdot p + V_{out} = 25 \cdot 5 + 180 = 305\,V$$

Maximální propustný proud musí být vyšší než výstupní proud navrhovaného měniče $I_{out} = 0{,}5\,A$.

Na kvalitách vybrané diody závisí částečně i účinnost měniče. Je tedy nutné vybírat i na základě parametrů určujících velikost výkonových ztrát. Tyto parametry jsou: kapacita v propustném směru a doba zotavení. Přičemž čím nižší jsou, tím nižší jsou výkonové ztráty.

Všem uvedeným kritériím nejlépe vyhovuje Schottkyho dioda SF38. Její parametry jsou uvedeny v tabulce (Tab. 8)

Tab. 8: Parametry diody SF38

značka	parametr	hodnota	jednotka
U_{RRM}	maximální závěrné napětí	600	V
IF(AV)	maximální propustný proud	3	A
Trr	maximální doba zotavení	35	μs
CJ	kapacita přechodu	50	pF
Tstg	pracovní teplota	-65 až 150	°C

4.7 Návrh schématu

Nyní, když už jsou nejdůležitější komponenty potřebné k návrhu schématu vybrány, lze přistoupit k samotnému návrhu zapojení měniče. Ten se musí samozřejmě řídit potřebami, parametry a způsobem funkce použitých součástek.

4.7.1 Předpoklady pro návrh

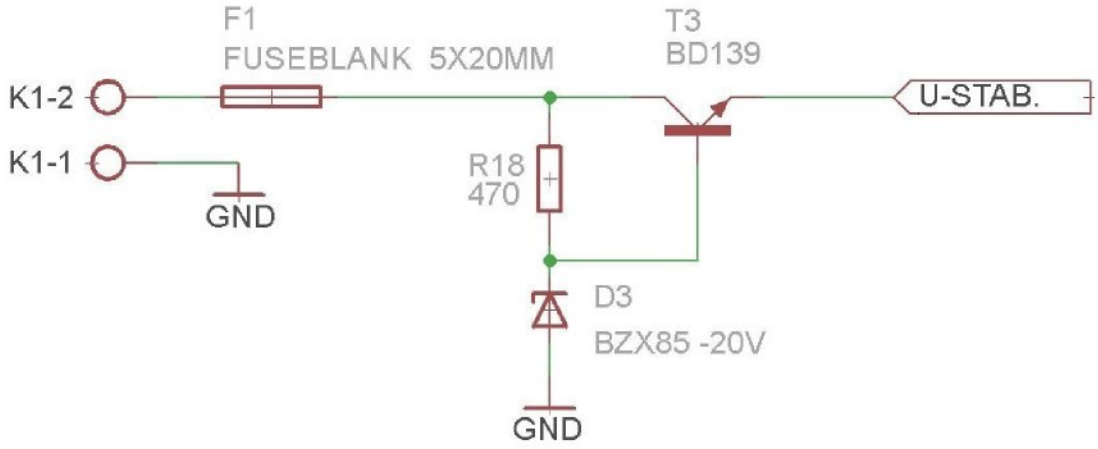
První věc, podle které je třeba se řídit, je funkce jednotlivých pinů řídícího obvodu a jeho potřeby z hlediska externích součástek. Nejprve si tedy představíme jednotlivé piny s příslušnými funkcemi.

- **Pin 1** – vstup pro informaci o sekundárním napětí – porovnáním regulovaného napětí, získaného z regulačního vinutí transformátoru, s vnitřním referenčním napětím, je šířka výstupního pulsu na pinu 5 upravena dle zatížení na sekundární straně transformátoru (bez zátěže, normální režim, přetížení, zkrat).
- **Pin 2** – vstup pro informaci o primárním proudu – rostoucí proud na primárním vinutí transformátoru je simulován externím RC členem. Pokud napětí překročí velikost hodnoty odvozené z pinu 1, je výstupní puls na pinu 5 ukončen. RC člen slouží k nastavení maximálního výkonu v režimu *bod přetížení.*
- **Pin 3** – vstup pro monitorování napětí na primárním vinutí transformátoru – je-li obvod v normálním režimu, pohybuje se toto napětí mezi mezemi V_{IL} a V_{IH}. Klesne-li napětí pod V_{IL} tak se řídící obvod vypne (napětí je příliš nízké). Překročí-li napětí V_{IH} začne pracovat obvod v režimu *přetížení.*
- **Pin 4** – uzemnění.
- **Pin 5** – výstup pro ovládání výkonového mosfet tranzistoru – tento výstup dokáže poskytnout proud až ±1 A pro rychlé nabytí/vybytí parazitní kapacity tranzistoru.
- **Pin 6** – vstup pro napájecí napětí – od tohoto napětí je odvozené nejen vnitřní referenční napětí, ale i veškerá prahová napětí *V6A*, *V6E*, $V6_{max}$ a $V6_{min}$. Je-li napětí na pinu 6 vyšší než *V6E*, pak je referenční napětí zapnuto. Je-li napětí na pinu 6 nižší než V6A pak je vypnuto. Tato logika je použitelná pouze pokud je $V6_{min} < V6 < V6_{max}$.
- **Pin 7** – vstup pro soft start – propojením tohoto vstupu a země kondenzátorem nabíhá obvod krátkými pulzy, což omezuje proudové rázy způsobené vybitými kapacitami při náběhu měniče do činnosti.
- **Pin 8** – vstup pro oscilační zpětnou vazbu – po startu oscilací sepne každý přechod zpětnovazebního napětí přes nulu výstupní puls na pinu 5.

Z popisu pinů je vidět, že je ke správné funkci potřeba poměrně velké množství externích součástek.

4.7.2 Navržené schéma

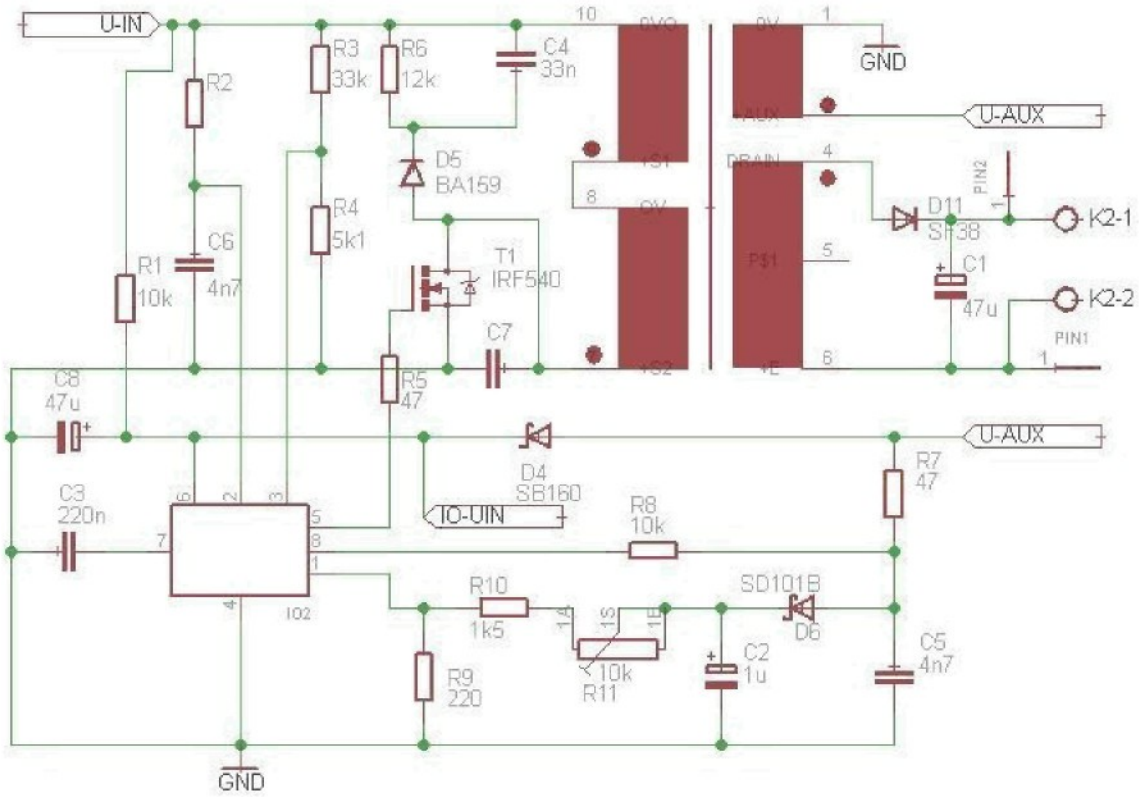
Kvůli větší přehlednosti a lepšímu popisu, je schéma rozděleno do několika funkčních bloků. Jednotlivé bloky jsou popsány níže.



Obr. 15: Schéma zapojení napájecí části

Na obrázku (Obr. 15) je schéma zapojení napájecí části. Ta slouží k napájení celého obvodu. Napětí je do obvodu přivedeno prostřednictvím konektoru *K1*. Kvůli ochraně proti zkratu, je za konektor umístěna tavná pojistka *F1*. Tranzistorem *T3*, odporem *R18* a Zenerovou diodou *D3* je realizován jednoduchý stabilizátor stabilizující napětí *U-STAB* na hodnotu 20 V. *U-STAB* slouží k dalšímu napájení v obvodu.

Schéma zapojení řídícího obvodu je na obrázku (Obr. 16). Napájení je přivedeno na svorku *U-IN*. Před započetím prvního pracovního cyklu *IO1*, je obvod napájen energií uloženou v kondenzátoru *C8*. Ten je dobíjen z napájení *U-IN* přes odpor *R1*. Po proběhnutí několika pracovních cyklů je již *IO1* napájen přes diodu *D4* z regulačního vinutí transformátoru. Vzhledem k tomu, že je potřeba minimalizovat úbytek napětí na diodě, byla jako *D4* použita Schottkyho dioda. Odpor *R2* a kondenzátor *C6* tvoří *RC* člen, prostřednictvím něhož je na pin 2 řídícího obvodu přivedena informace o průběhu primárního proudu. Odpory *R3* a *R4* tvoří napěťový dělič poskytující pinu 3 *IO1* informaci o napětí na primáru. Pin 5, výstup integrovaného obvodu, řídí pulsy přes odpor *R5* výkonový mosfet tranzistor *T1*, který spíná primární vinutí transformátoru. Kondenzátor *C7* slouží ke stanovení rezonanční frekvence obvodu. Odpor *R6*, dioda *D5* a kondenzátor *C4* mají funkci ochrany tranzistoru *T1*. Pohlcují napěťové špičky vznikající spínáním a rozpínáním primární indukčnosti transformátoru a zmenšují jeho napěťové namáhání.



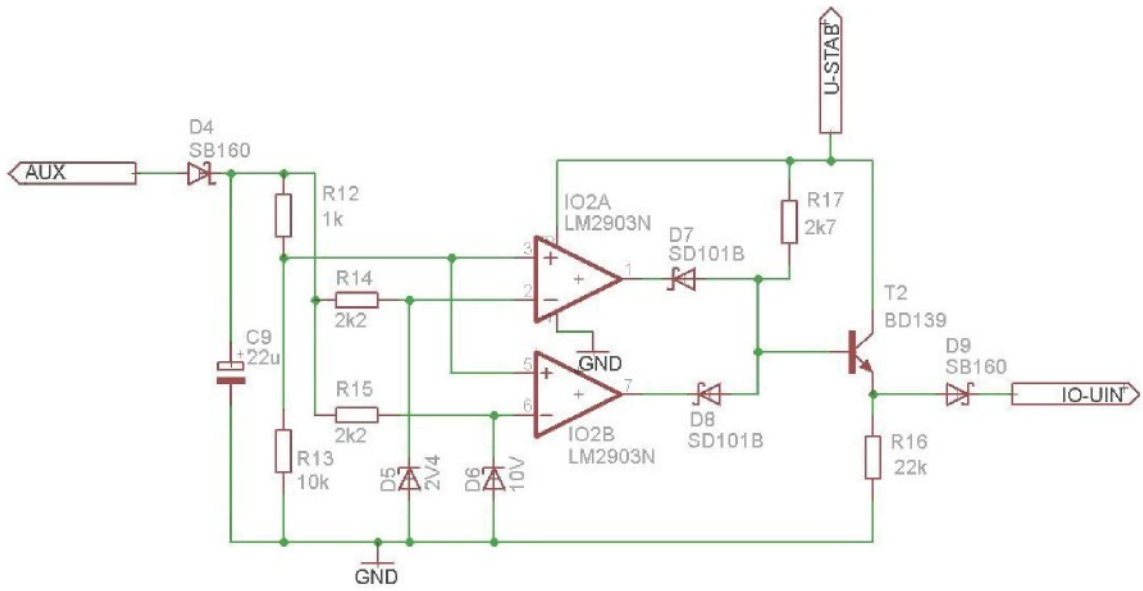
Obr. 16: Schéma zapojení řídící části

Z regulačního vinutí transformátoru U-AUX je přes odpor *R8* na pin 8 přivedena informace o průchodech sekundárního napětí nulou. Jelikož je transformátor v závislosti na požadovaném výstupním napětí spínán a rozpínán, je na všech jeho sekundárních vinutích napětí obdélníkového průběhu. Tato napětí je třeba usměrnit. U regulačního vinutí je tak provedeno prostřednictvím Schottkyho diody *D6* a filtračního kondenzátoru *C2*. Tímto usměrněným napětím je napájen odporový dělič tvořený potenciometrem *R11*, odpory *R9* a *R10*. Výstup z děliče je přiveden na regulační vstup *IO*. Potenciometrem je tak možné měnit výstupní napětí *U-OUT*. To je získáno z výstupu transformátoru usměrněním diodou *D11* a vyfiltrováním kondenzátorem *C1*. Toto napětí je připojené na konektor *K2*. Jumpry 1 a 2 slouží k případnému připojení panelového voltmetru.

Co zde ještě nebylo zmíněno, je způsob ochrany proti zkratu na výstupu. Lze totiž využít toho, že je integrovaný obvod po téměř celou dobu své činnosti napájen z regulačního vinutí. Dojde-li tedy ke zkratu, klesne napětí na sekundárních vinutích (včetně regulačního) téměř až k nule. *IO1* tak nebude napájen a vypne se. Zapne se teprve tehdy, až se kondenzátor *C8* přes odpor *R1* nabije na dostatečnou úroveň.

Pokud zkrat na výstupu trvá, spotřebuje *IO1* při startu všechnu energii z *C8* a opět se vypne. Tyto neúspěšné starty se budou opakovat po celou dobu trvání zkratu.

Aby však bylo možné řešit ochranu proti zkratu tímto způsobem a zároveň ponechat měniči možnost potenciometrem měnit výstupní napětí, je třeba doplnit k dosavadnímu návrhu obvod ošetřující ochranu proti zkratu. Důvod je následující: bude-li na výstupu měniče napětí *Uout*, pak na regulačním vinutí transformátoru bude napětí *Uout/p*. *IO1* ke své činnosti potřebuje minimální napětí 7 V, což již pro *Uout = 60 V* není splněno. Není tak možné měnit výstupní napětí pod 60 voltů. Takovýto zdroj by však postrádal svou funkci, je teda třeba tento stav vhodně ošetřit. Schéma zapojení obvodu ošetřujícího ochranu proti zkratu je na obrázku (Obr. 17).



Obr. 17: Schéma zapojení ošetřující protizkratovou ochranu

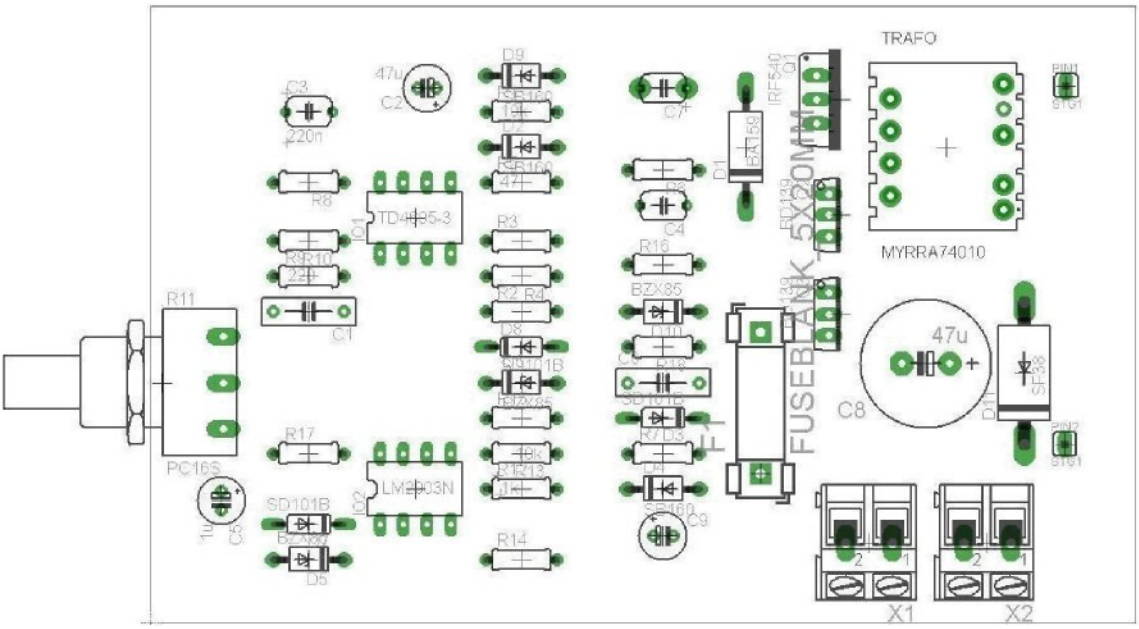
Obvod má za úkol vyhodnocovat napětí na regulačním vinutí. To je přivedeno na vstupní dělič tvořený odpory *R12* a *R13*. Tímto děličem upravené napětí je pak pomocí dvou komparátorů porovnáváno s referenčními napětími tvořenými Zenerovými diodami *D5*, *D6* a odpory *R14* a *R15*. Je-li napětí vyšší, než je mez stanovená referencí *R14-D5* má komparátor 1 výstup v nejnižší poloze. Totéž platí pro druhý komparátor, pokud je napětí nižší, než je mez stanovená referencí *R15-D6*. Oba výstupy jsou ve vysoké úrovni pouze tehdy, je-li napětí v rozmezí obou referencí. Výstupy komparátorů jsou zapojeny do jakéhosi hradla AND realizovaného odporem *R1* a diodami *D2* a *D3*. Výstupem tohoto hradla je řízený tranzistor T_2, zapojený jako emitorový sledovač. Je-li tedy napětí v mezích daných referenčními napětími komparátorů, napájí tento emitorový sledovač *IO1*.

4.8 Návrh desky plošných spojů

Po navržení a důkladné kontrole schématu je možné přistoupit k návrhu desky plošných spojů. DPS byla navrhována v programu EAGLE. Nejprve bylo potřeba navržené schéma překreslit do Schematic Editoru, definovat pouzdra součástek a vytvořit si vlastní knihovnu se součástkami, které nejsou v EAGLU obsaženy. Takto byl vytvořen transformátor a řídící obvod. Po překreslení schématu následovalo přepnutí do Layout Editoru. Zde bylo třeba součástky rozmístit a příslušně propojit. Při propojování součástek byl kladen největší důraz na tloušťku vodičů a dostatečnou vzdálenost mezi nimi.

4.8.1 Navržená DPS

Schéma je realizováno na dvoustranné desce s plošnými spoji o rozměrech 106×60 mm. Rozložení součástek na desce je na obrázku (Obr. 18).

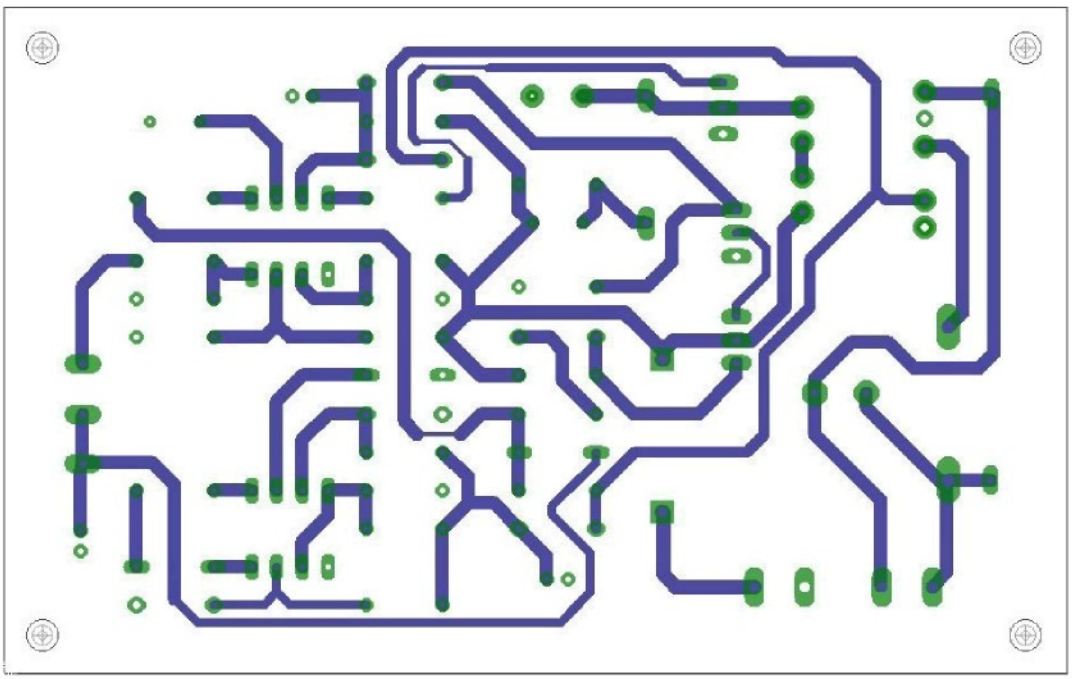


Obr. 18: Rozložení součástek na DPS

Kritéria při rozmisťování součástek byla především tato: vhodné umístění pojistkového držáku, jumperů a konektorů *K1* a *K2*. Držák by měl být umístěn pokud možno na snadno přístupném místě, aby byla zaručena snadná výměna pojistky. Jumpery, jelikož je na nich výstupní napětí a to může dosahovat až 180 voltů, je třeba umístit dále od sebe, aby byla co nejvíce eliminována možnost nechtěného propojení a tím zkratu na výstupu.

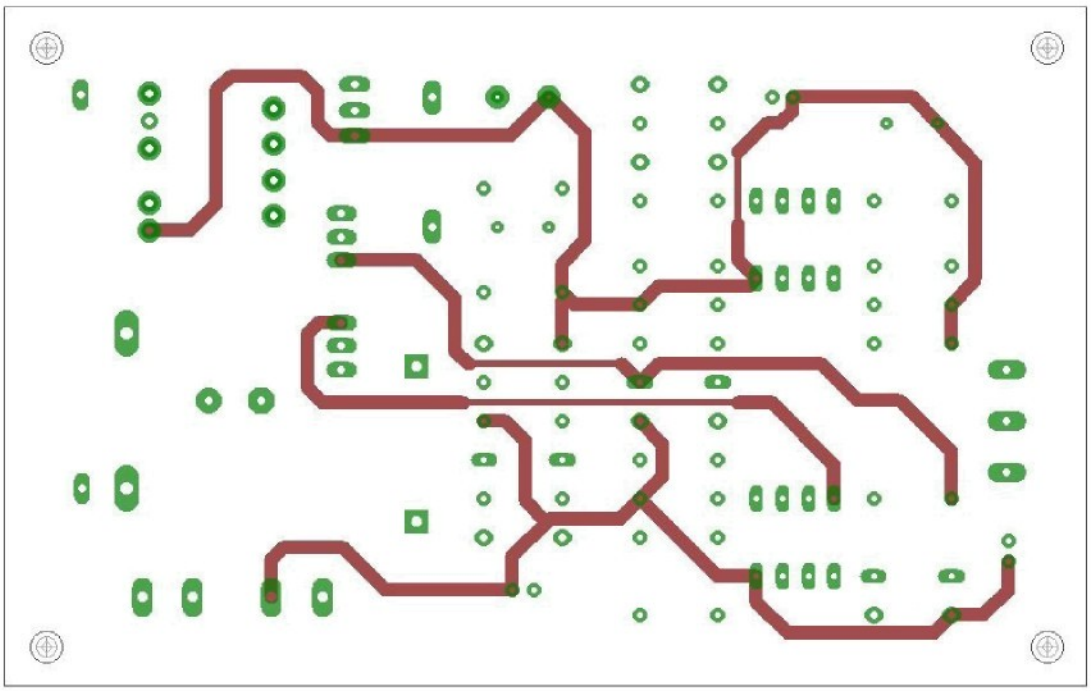
Jelikož je předpokládaný výkon měniče do 10 Wattů, není potřeba připevňovat na tranzistory přídavné chladiče. Ty jsou však rozmístěny tak, aby bylo i toto v případě potřeby možné.

Obraz desky plošných spojů ze strany BOTTOM je na obrázku (Obr. 19).



Obr. 19: Obraz DPS ze strany BOTTOM

Obraz desky plošných spojů ze strany TOP je na obrázku (Obr. 20).



Obr. 20: Obraz DPS ze strany TOP

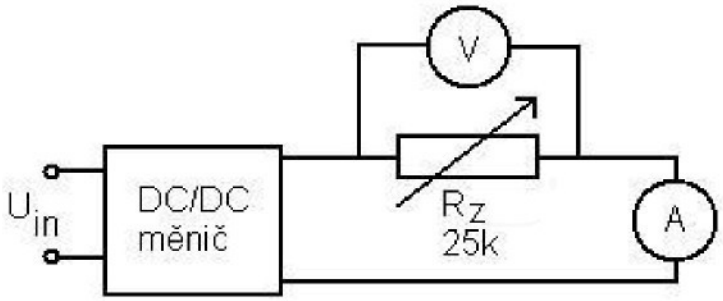
5 Výsledky

Po návrhu a konstrukci měniče přichází na řadu zhodnocení hotového výrobku. Je potřeba ověřit jeho funkčnost, určit podmínky, za kterých je možné zařízení provozovat a porovnat jeho parametry s návrhem.

5.1 Měření zátěžové charakteristiky

Jedním z nejdůležitějších popisů zdroje je popis z hlediska reakce na měnící se zátěž, takzvaná zatěžovací charakteristika. Jde o závislost svorkového napětí na proudu, který tento zdroj do zátěže dodává. V ideálním případě je tato charakteristika konstantní. Napětí ideálního zdroje tak není ovlivňováno dodávaným proudem. V reálném případě je však toto nemožné. Zdroj má vždy nějaký vnitřní odpor R_i, který je příčinou závislosti svorkového napětí na dodávaném proudu. Podle velikosti odporu se tedy zdroje dělí na tvrdé a měkké.

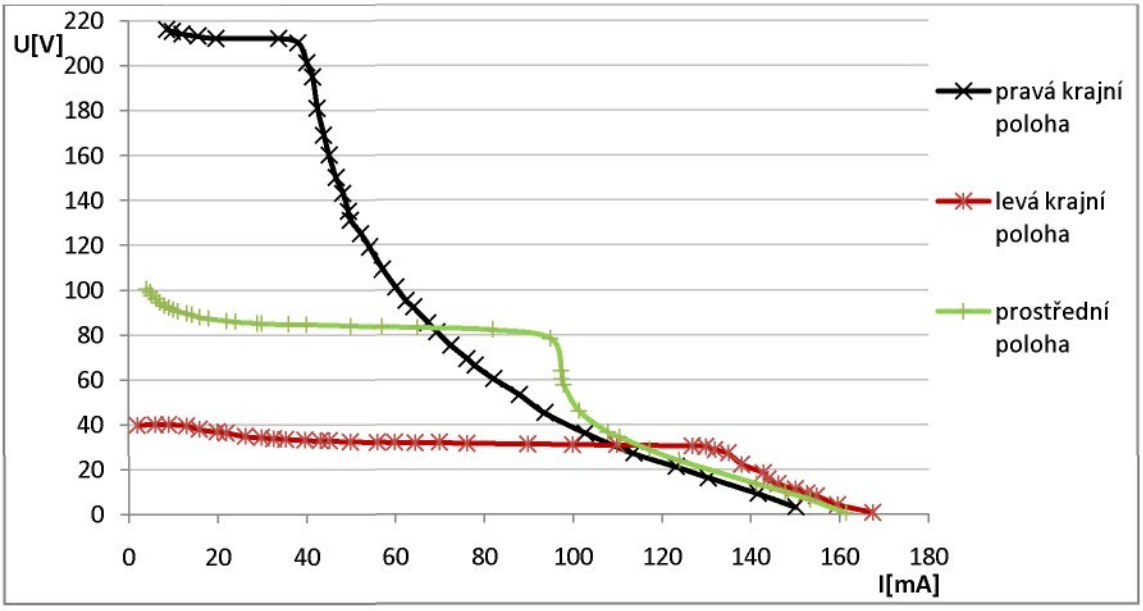
Při měření zatěžovací charakteristiky bylo postupováno následovně: měnič byl zapojen dle schématu, uvedeném na obrázku (Obr. 21).



Obr. 21: Schéma zapojení při měření zatěžovací charakteristiky

Byly naměřeny tři sady parametrů, při třech různých polohách regulačního potenciometru. První bod v charakteristice odpovídá vždy napětí při maximálním odporu R_Z. Další body jsou získávány postupným zvyšováním odebíraného proudu (snižováním odporu zátěže R_Z).

Tabulka naměřených hodnot je uvedena v příloze A. Příslušný graf (graf. 1) je ukázán níže.



graf 1: Zatěžovací charakteristika

Jak je z grafu (graf 1) vidět, jsou charakteristiky z počátku poměrně lineární. Poté co zdroj překročí bod přetížení (výkonové maximum), začne řídící obvod výkon omezovat. Dostane se tak do režimu *přetížení*. Tento režim je v grafu reprezentován prudkým poklesem napětí.

Při měření je zároveň možné otestovat protizkratovou ochranu. Na zkratování výstupních svorek, by měl řídící obvod včas zareagovat a vypnout se. Díky špatně nastavenému obvodu protizkratové ochrany se tak nestalo. Řídící obvod omezoval výkon až do zkratu, k vypnutí však nedošlo.

5.2 Měření parametrů

Určujícími parametry spínaného zdroje jsou: vstupní a výstupní napěťové rozsahy, výstupní zvlnění, rychlost odezvy na skokovou změnu zátěže a účinnost. Většina těchto parametrů je závislá na výstupním napětí a výkonu. Z toho důvodu bylo, pokud to bylo možné, měřeno několikrát, pokaždé za jiných podmínek.

5.2.1 Napěťové rozsahy

Kromě stanovení proudových a napěťových maxim návrhem, je potřeba tato maxima experimentálně ověřit. Minimální a maximální výstupní napětí bylo změřeno po nastavení regulačního potenciometru do obou krajních poloh. Byly naměřeny 3 sady parametrů, odpovídající třem různým zátěžovým proudům. Všechny tři naměřené sady jsou v tabulce (Tab. 9).

Tab. 9: Rozsah výstupního napětí

Iz [mA]	Umin [V]	Iz [mA]	Umax [V]
8,6	215,0	1,5	35,0
24,8	211,0	20,0	33,0
50,0	130,0	131,0	27,0

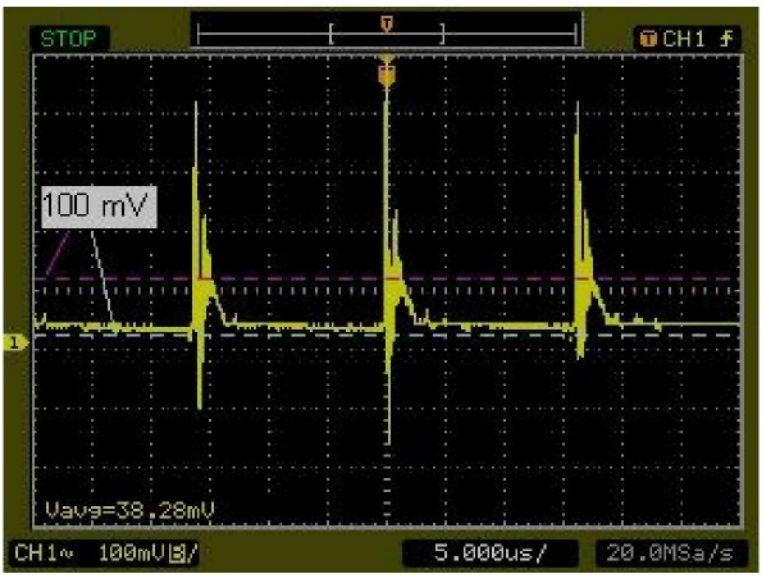
Maximální vstupní napětí stanovené návrhem by se nemělo překračovat, jelikož je to zároveň napájecí napětí řídícího obvodu. Bylo tedy jen ověřeno, zda je toto maximum reálné. Minimální vstupní napětí může být teoreticky velmi malé, po rozběhu měniče je totiž řídící obvod napájen z regulačního vinutí transformátoru. Bylo tedy ověřeno, jak velké napětí je potřebné k rozběhu měniče a jak velké napětí je potřebné k udržení měniče v chodu. Všechny tyto hodnoty jsou v tabulce (Tab. 10).

Tab. 10: Rozsah vstupního napětí

značka	parametr	hodnota	jednotka
Ustart	minimální napětí potřebné ke startu obvodu	13,18	U [V]
Umin	minimální pracovní napětí	9,20	U [V]
Umax	maximální pracovní napětí	18,02	U [V]

5.2.2 Výstupní zvlnění

Velikost výstupního zvlnění byla zjištěna pomocí osciloskopem naměřeného průběhu. Na obrázku (Obr. 22) je vidět, že výsledný průběh je složen ze zvlnění způsobeného kondenzátorem a napěťovými špičkami způsobenými spínáním.

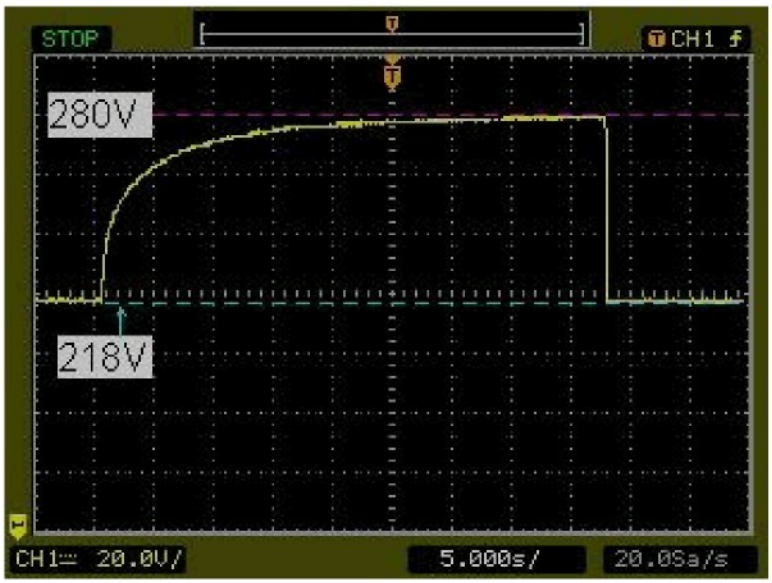


Obr. 22: Výstupní zvlnění

Zatím co zvlnění dosahuje hodnoty 100 mV, napěťové špičky dosahují až 500 mV. Avšak vzhledem k tomu, že je výstupní napětí 210 V, jsou tyto hodnoty uspokojivé. Pokud by bylo třeba hodnotu zvlnění snížit, lze tak provést zlepšením parametrů výstupního filtru (náhradou původního kondenzátoru za kondenzátor s lepšími parametry).

5.2.3 Odezva na skok

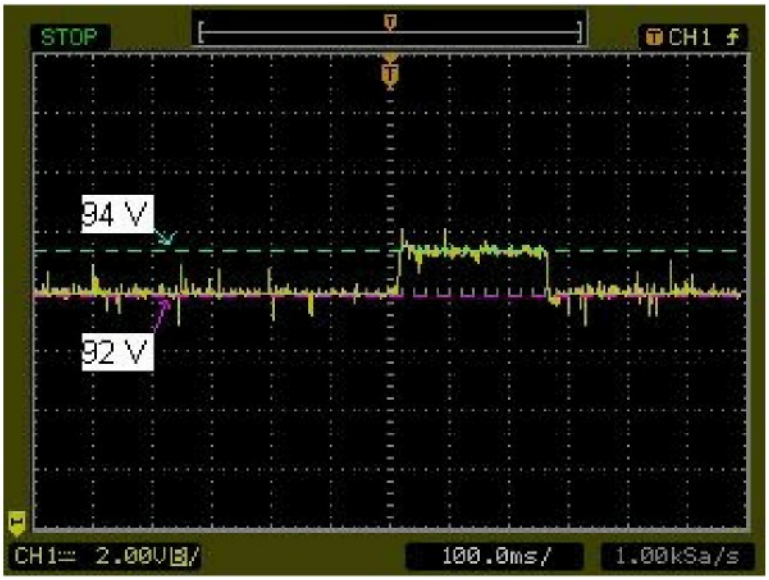
Další naměřený parametr se týká kvality regulace požadovaného napětí. Jde o rychlost reakce na skokovou změnu zátěže. Měření probíhalo následujícím způsobem: na výstup měniče byl připojen spínač a odpor odpovídající 50 % jmenovitého výkonu. Naměřený průběh (Obr. 23) tedy odpovídá přepnutím z chodu při 50 % výkonu do chodu naprázdno a zpět.



Obr. 23: Stanovení rychlosti reakce (naprázdno → 50 %)

Z uvedeného obrázku (Obr. 23) je vidět, že mezi chodem naprázdno a chodem se zátěží jsou poměrně vysoké napěťové rozdíly. To je způsobeno změnou operačních režimů řídícího obvodu, z režimu *bez zátěže* do *normálního režimu*. Přičemž, jak je v datasheetu uvedeno, napětí v režimu *bez zátěže* se může od napětí v *normálním režimu* lišit. Tento nedostatek by bylo možné jednoduše odstranit. Přidáním odporu paralelně k výstupnímu kondenzátoru, by bylo docíleno toho, že i při odpojení zátěže by byl měnič minimálně, avšak dostatečně zatížen. Obvod by tak po celou dobu pracoval v normálním režimu. Napěťové rozdíly by tak měly být minimální.

Z výše uvedených důvodů bylo, jako více vypovídající, zvoleno měření průběhu při přepnutí mezi 25% a 75% výkonem. Naměřený průběh s vyznačenými hodnotami je na obrázku (Obr. 24)



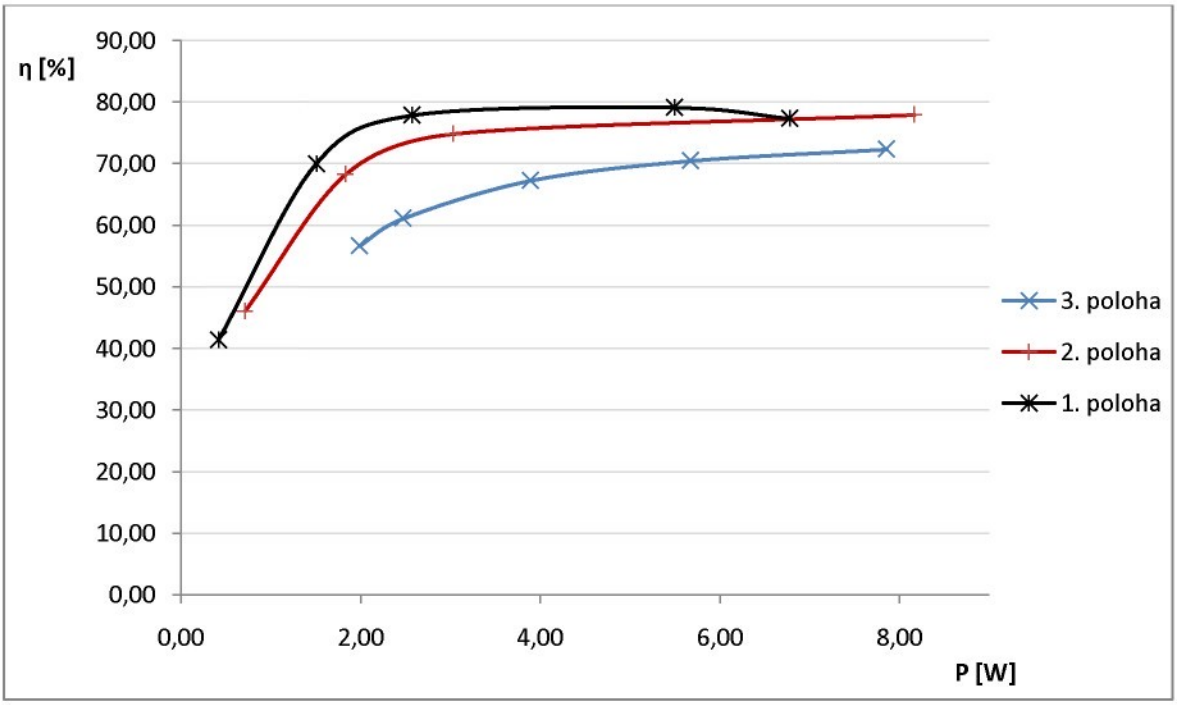
Obr. 24: Stanovení rychlosti reakce (25 % → 75 %)

Jak je vidět z obrázku (Obr. 24), reakce je na skokovou změnu zátěže poměrně rychlá (> 1 ms). Obvod také nepřechází z jednoho pracovního režimu do druhého, změna výstupního napětí je tak mnohonásobně nižší, než v prvním případě.

5.2.4 Účinnost

Posledním parametrem, vypovídajícím především o efektivitě měniče, je účinnost. Při měření bylo vycházeno z definičního vztahu pro účinnost (***účinnost = výkon/příkon***). Byl tedy měřen vstupní příkon a výstupní výkon, a to pro tři různé polohy regulačního potenciometru. Z těchto hodnot byly dále spočítány příslušné účinnosti.

Naměřené hodnoty (Tab. 11) a jsou příslušný graf (graf 2) jsou uvedeny níže.



graf 2: Závislost účinnosti na výkonu

Jak je z grafu (graf 2) patrné, roste účinnost s blížícím se výkonovým maximem. To je v případě tohoto zdroje 8 W. Tuto hodnotu lze dodatečně zvýšit použitím vyšší hodnoty kondenzátoru *C6*. Po překročení výkonového maxima, začne řídící obvod výkon omezovat a klesá tak i účinnost zdroje.

Tab. 11: Měření účinnosti - hodnoty

out			in			
1. poloha						
U[V]	I[mA]	P[W]	U[V]	I[A]	P[W]	η[%]
78,00	5,37	0,42	13,50	0,08	1,01	41,35
74,00	20,42	1,51	13,50	0,16	2,16	69,97
72,00	35,74	2,57	13,50	0,25	3,31	77,81
70,00	78,56	5,50	13,50	0,52	6,95	79,10
68,00	99,75	6,78	13,50	0,65	8,78	77,30
2. poloha						
133,00	5,38	0,72	13,50	0,12	1,55	46,06
127,00	14,44	1,83	13,50	0,20	2,69	68,24
125,00	24,25	3,03	13,50	0,30	4,05	74,83
122,00	66,94	8,17	13,50	0,78	10,48	77,96
3. poloha						
216,00	9,20	1,99	13,50	0,26	3,51	56,62
214,00	11,57	2,47	13,50	0,30	4,05	61,11
213,00	18,28	3,89	13,50	0,43	5,79	67,23
211,00	26,91	5,68	13,50	0,60	8,06	70,44
201,00	39,10	7,86	13,50	0,81	10,87	72,32

5.2.5 Porovnání parametrů

Pro vyšší přehlednost a lepší porovnatelnost návrhu s výslednou realizací jsou v tabulce (Tab. 12) uvedeny jak navrhnuté, tak naměřené parametry zkonstruovaného měniče. Jak již bylo uvedeno výše, lze hodnotu 8 W dodatečně upravit změnou hodnoty kondenzátoru *C6*, přičemž zvýšením kapacity se zvýší maximální výstupní výkon.

Tab. 12: Srovnání parametrů měniče

parametr	navržená hodnota	naměřená hodnota
výstupní napětí	180 V	38 - 218 V
vstupní napětí	8 - 25 V	9 - 18 V
výstupní proud	0,5 A	0,2 A
výkon	10 W	10 W
zvlnění na výstupu	200 mV	100 mV
odezva na skok		1 ms
účinnost		76%

Závěr

Cílem této práce bylo navrhnout spínaný zdroj pro napájení piezoelektrických aktuátorů. Návrh parametrů měniče, výběr součástek a výsledná konstrukce se tedy řídila potřebami těchto aktuátorů. Byly stanoveny parametry a potřebné vlastnosti zdroje. Z naměřených hodnot a uvedených grafů vyplývá, že měnič, až na drobné výjimky, požadované podmínky splňuje.

Většina parametrů výsledného zkonstruovaného zdroje dosahuje, či předčuje parametry dané návrhem. Výraznějšími nedostatky jsou pouze: poměrně úzký rozsah vstupního napětí a nesprávná funkčnost obvodu ošetřujícího protizkratovou ochranu. To by bylo možné vyřešit přenastavením snímacího napěťového děliče na 3. pinu *IO1* a drobnou úpravou zapojení napájecí části obvodu. Co se týče protizkratové ochrany, je z grafu (graf 1) patrné, že i bez plně funkční ochrany obvod, po překročení bodu přetížení, spolehlivě a velmi účinně omezuje výkon. Obvod ošetřující ochranu proti zkratu, je tak vlastně zbytečný.

Možnosti vylepšení jsou poměrně vysoké. Z důvodu pojetí měniče jako testovacího kusu, nebyly do návrhu zařazeny prvky omezující různé druhy rušení. Použitím SMD technologií, by bylo navíc možné zařízení poměrně výrazně zmenšit. To jsou změny, které by se týkaly především DPS. Z hlediska vlastností a parametrů zdroje by bylo dobré zlepšit především kvalitu regulace a účinnost. Jelikož je informace o výstupním napětí snímána nepřímo (z regulačního vinutí), není výsledná regulace optimální. Lepší volba by tedy byla snímání výstupního napětí pomocí napěťového děliče. Co se týče účinnosti (dosahuje 79 %), bylo by dobré její hodnotu zvýšit. Šlo by tak provést použitím kvalitnějších součástek a minimalizací ztrát na DPS.

Seznam použité literatury

[1] *Datasheetcatalog : datasheet* [online]. 2009 [cit. 2009-05-17]. Dostupný z WWW: <http://www.datasheetcatalog.org/datasheet/infineon/1-tda46053.pdf>.

[2] *Datasheetcatalog : datasheet* [online]. 2009 [cit. 2009-05-17]. Dostupný z WWW: <http://www.datasheetcatalog.org/datasheet/stmicroelectronics/9387.pdf>.

[3] *Datasheetscatalog : datasheet* [online]. 2009 [cit. 2009-05-17]. Dostupný z WWW: <http://www.datasheetcatalog.org/datasheet/eic/SF33.pdf>.

[4] *Farnell : Datasheet* [online]. [cit. 2009-05-17]. Dostupný z WWW: <http://www.farnell.com/datasheets/41319.pdf>.

[5] KREJČIŘÍK, Alexandr. *Napájecí zdroje 1 : Základní zapojení analogových a spínaných napájecích zdrojů*. 2. vyd. Praha : BEN, 1997. 351 s. ISBN 80-86056-02-3.

[6] KREJČIŘÍK, Alexandr. *Napájecí zdroje 2 : Integrované obvody ve spínaných zdrojích*. 2. vyd. Praha : BEN, 1997. 351 s. ISBN 80-86056-03-1.

[7] PRESSMAN, Abraham I. . *Switching power supply design*. 2nd edition. New York : McGraw-Hill, 1997. 682 s. ISBN 0-07-052236-7.

[8] SCLOCCHI, Michele. *Switching power supply design : continuous mode flyback converter* [online]. [cit. 2009-05-17]. Dostupný z WWW: <www.national.com/appinfo/power/files/flyaback-cont2000-5000-new.pdf>.

[9] TŮMA, Jiří, et al. *Základy automatizace*. 1. vyd. Ostrava : Ediční středisko VŠB – TUO , 2007. 283 s. ISBN 978-80-248-1523-7.

Seznam příloh

A: Tab. 13: Zatěžovací charakteristika – naměřené hodnoty

B: CD

Obsah CD

- Elektronická podoba bakalářské práce
- Literatura
  - Základy automatizace.pdf
  - Switching power supply design.pdf
  - IRF540.pdf
  - Myrra 74010.pdf
  - TDA4605-3.pdf
  - SF38.pdf
- Soubory programu EAGLE

Příloha A

Tab. 13: Zatěžovací charakteristika – naměřené hodnoty

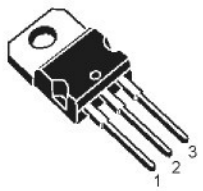
pravá krajní poloha			prostřední poloha			levá krajní poloha		
I[mA]	U[V]	P[W]	I[mA]	U[V]	P[W]	I[mA]	U[V]	P[W]
8,67	216,00	18,73	4,04	100,50	4,06	2,00	39,70	0,79
9,90	215,00	21,29	4,92	99,00	4,87	6,00	40,00	2,40
12,00	214,00	25,68	5,10	97,60	4,98	9,00	39,90	3,59
15,63	213,00	33,29	6,09	96,20	5,86	13,00	39,30	5,11
19,67	212,00	41,70	7,00	94,30	6,60	16,00	37,80	6,05
33,80	212,00	71,66	8,00	93,00	7,44	20,00	36,70	7,34
38,22	210,00	80,26	9,00	92,00	8,28	22,00	36,20	7,96
40,20	201,00	80,80	10,00	91,20	9,12	26,40	34,60	9,13
41,42	195,00	80,77	11,00	90,50	9,96	30,00	34,10	10,23
42,50	181,00	76,93	13,00	89,50	11,64	32,70	33,70	11,02
44,00	169,00	74,36	14,20	88,90	12,62	35,40	33,40	11,82
45,20	160,00	72,32	16,00	87,80	14,05	39,70	33,00	13,10
46,80	150,00	70,20	18,00	87,20	15,70	43,20	32,70	14,13
48,30	143,00	69,07	22,00	86,20	18,96	45,20	32,60	14,74
49,50	135,00	66,83	24,00	85,80	20,59	50,00	32,20	16,10
50,00	131,00	65,50	29,00	85,00	24,65	56,00	32,00	17,92
52,30	125,00	65,38	30,00	84,90	25,47	60,00	31,80	19,08
54,30	119,00	64,62	36,00	84,60	30,46	64,60	31,70	20,48
57,20	109,00	62,35	40,00	84,40	33,76	70,00	31,80	22,26
60,20	101,00	60,80	50,00	83,90	41,95	76,20	31,40	23,93
62,50	95,00	59,38	57,00	83,50	47,60	90,00	31,10	27,99
64,20	92,00	59,06	65,00	83,00	53,95	100,00	30,80	30,80
67,48	85,00	57,36	82,00	82,00	67,24	110,00	30,60	33,66
69,50	81,00	56,30	95,00	78,20	74,29	127,00	30,20	38,35
72,62	75,00	54,47	97,50	63,80	62,21	130,00	29,90	38,87
76,30	69,00	52,65	97,60	60,20	58,76	132,00	28,50	37,62
78,10	66,00	51,55	97,94	57,60	56,41	135,00	27,00	36,45
82,30	60,00	49,38	101,50	45,80	46,49	138,00	22,00	30,36
88,00	53,00	46,64	108,00	36,40	39,31	142,80	18,20	25,99
93,80	45,00	42,21	110,50	34,30	37,90	143,90	15,30	22,02
102,90	36,00	37,04	117,30	28,60	33,55	146,20	13,20	19,30
113,70	27,00	30,70	123,90	23,80	29,49	150,10	10,90	16,36
123,20	21,00	25,87	141,20	13,40	18,92	152,60	9,10	13,89
130,50	16,00	20,88	147,70	9,80	14,47	154,80	7,90	12,23
141,70	9,00	12,75	153,40	6,20	9,51	159,60	4,00	6,38
150,20	3,00	4,51	161,40	0,40	0,65	167,50	0,60	1,01

IRF540

N-CHANNEL 100V - 0.065 Ω - 30A TO-220
LOW GATE CHARGE STripFET™ POWER MOSFET

TYPE	V_{DSS}	$R_{DS(on)}$	I_D
IRF540	100 V	<0.070 Ω	30 A

- TYPICAL $R_{DS}(on)$ = 0.065Ω
- EXCEPTIONAL dv/dt CAPABILITY
- 100% AVALANCHE TESTED
- LOW GATE CHARGE
- APPLICATION ORIENTED CHARACTERIZATION



TO-220

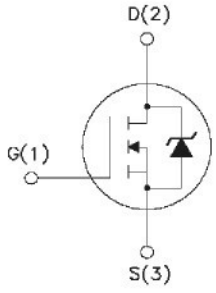
DESCRIPTION

This MOSFET series realized with STMicroelectronics unique STripFET process has specifically been designed to minimize input capacitance and gate charge. It is therefore suitable as primary switch in advanced high-efficiency, high-frequency isolated DC-DC converters for Telecom and Computer applications. It is also intended for any applications with low gate drive requirements.

APPLICATIONS

- HIGH-EFFICIENCY DC-DC CONVERTERS
- UPS AND MOTOR CONTROL

INTERNAL SCHEMATIC DIAGRAM



ABSOLUTE MAXIMUM RATINGS

Symbol	Parameter	Value	Unit
V_{DS}	Drain-source Voltage (V_{GS} = 0)	100	V
V_{DGR}	Drain-gate Voltage (R_{GS} = 20 kΩ)	100	V
V_{GS}	Gate- source Voltage	± 20	V
I_D	Drain Current (continuos) at T_C = 25°C	30	A
I_D	Drain Current (continuos) at T_C = 100°C	21	A
I_{DM}(•)	Drain Current (pulsed)	120	A
P_{tot}	Total Dissipation at T_C = 25°C	100	W
	Derating Factor	1	W/°C
T_{stg}	Storage Temperature	-65 to 175	°C
T_j	Max. Operating Junction Temperature	-55 to 175	°C

(•) Pulse width limited by safe operating area.

THERMAL DATA

$R_{thj\text{-}case}$	Thermal Resistance Junction-case	Max	1.5	°C/W
$R_{thj\text{-}amb}$	Thermal Resistance Junction-ambient	Max	62.5	°C/W
T_l	Maximum Lead Temperature For Soldering Purpose	Typ	300	°C

ELECTRICAL CHARACTERISTICS (T_{case} = 25 °C unless otherwise specified)

OFF

Symbol	Parameter	Test Conditions	Min.	Typ.	Max.	Unit
$V_{(BR)DSS}$	Drain-source Breakdown Voltage	I_D = 250 µA, V_{GS} = 0	100			V
I_{DSS}	Zero Gate Voltage Drain Current (V_{GS} = 0) Zero Gate Voltage Drain Current (V_{GS} = 0)	V_{DS} = Max Rating V_{DS} = Max Rating T_C = 125°C			1 10	µA µA
I_{GSS}	Gate-body Leakage Current (V_{DS} = 0)	V_{GS} = ± 20V			±100	nA

ON (1)

Symbol	Parameter	Test Conditions	Min.	Typ.	Max.	Unit
$V_{GS(th)}$	Gate Threshold Voltage	V_{DS} = V_{GS} I_D = 250 µA	2	3	4	V
$R_{DS(on)}$	Static Drain-source On Resistance	V_{GS} = 10 V I_D = 15 A		0.065	0.070	Ω

DYNAMIC

Symbol	Parameter	Test Conditions	Min.	Typ.	Max.	Unit
g_{fs} (*)	Forward Transconductance	V_{DS} > $I_{D(on)}$ x $R_{DS(on)max}$, I_D = 15 A		10		S
C_{iss} C_{oss} C_{rss}	Input Capacitance Output Capacitance Reverse Transfer Capacitance	V_{DS} = 25V, f = 1 MHz, V_{GS} = 0		870 125 52		pF pF pF

ELECTRICAL CHARACTERISTICS (continued)

SWITCHING ON

Symbol	Parameter	Test Conditions	Min.	Typ.	Max.	Unit
$t_{d(on)}$ t_r	Turn-on Delay Time Rise Time	V_{DD} = 50 V I_D = 15 A R_G = 4.7 Ω V_{GS} = 10 V (Resistive Load, Figure 3)		58 45		ns ns
Q_g Q_{gs} Q_{gd}	Total Gate Charge Gate-Source Charge Gate-Drain Charge	V_{DD}= 80 V I_D= 30 A V_{GS}= 10V		30 6 10	41	nC nC nC

SWITCHING OFF

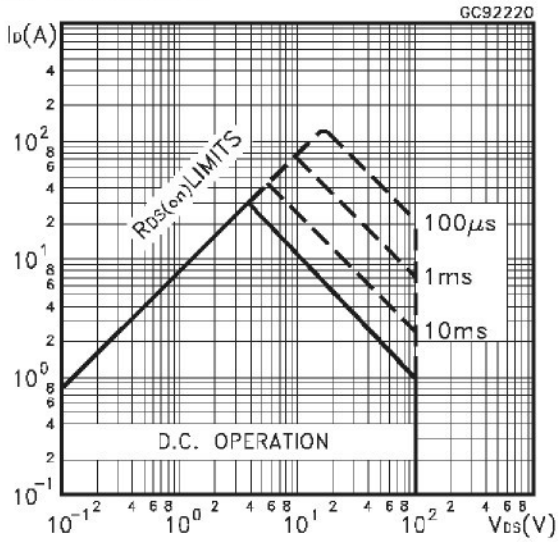
Symbol	Parameter	Test Conditions	Min.	Typ.	Max.	Unit
$t_{d(off)}$ t_f	Turn-off Delay Time Fall Time	V_{DD} = 50 V I_D = 30 A R_G = 4.7Ω, V_{GS} = 10 V (Resistive Load, Figure 3)		49 17		ns ns
$t_{r(Voff)}$ t_f t_c	Off-voltage Rise Time Fall Time Cross-over Time	V_{clamp} = 80 V I_D = 30 A R_G = 4.7Ω, V_{GS} = 10 V (Inductive Load, Figure 5)		43 21 39		ns ns ns

SOURCE DRAIN DIODE

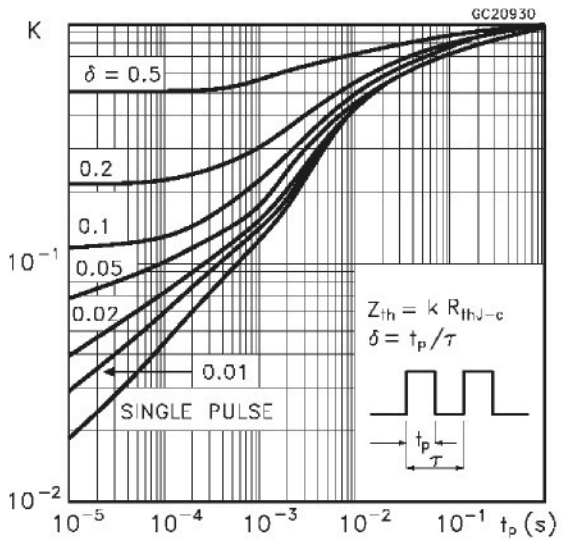
Symbol	Parameter	Test Conditions	Min.	Typ.	Max.	Unit
I_{SD} I_{SDM} (•)	Source-drain Current Source-drain Current (pulsed)				30 120	A A
V_{SD} (*)	Forward On Voltage	I_{SD} = 30 A V_{GS} = 0			1.3	V
t_{rr} Q_{rr} I_{RRM}	Reverse Recovery Time Reverse Recovery Charge Reverse Recovery Current	I_{SD} = 30 A di/dt = 100A/µs V_{DD} = 30 V T_j = 150°C (see test circuit, Figure 5)		100 375 7.5		ns nC A

(*)Pulsed: Pulse duration = 300 µs, duty cycle 1.5 %.
(•)Pulse width limited by safe operating area.

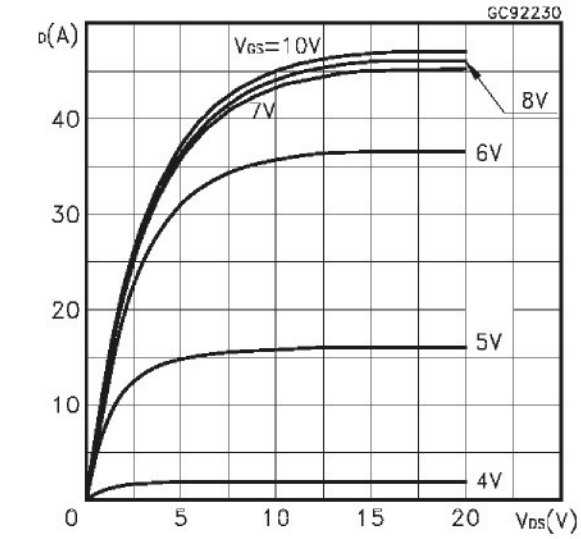
Safe Operating Area



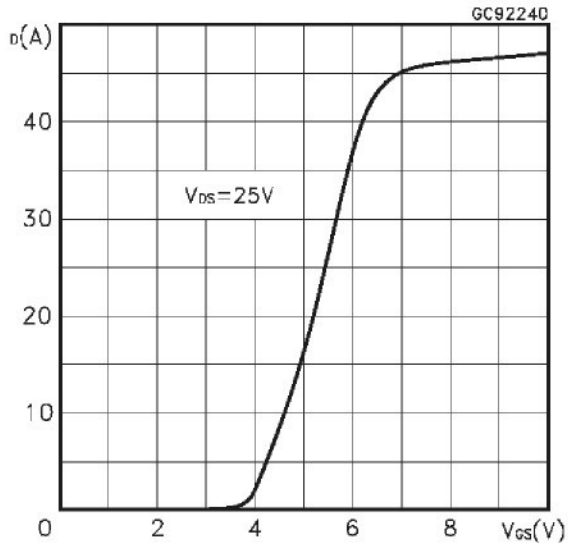
Thermal Impedance



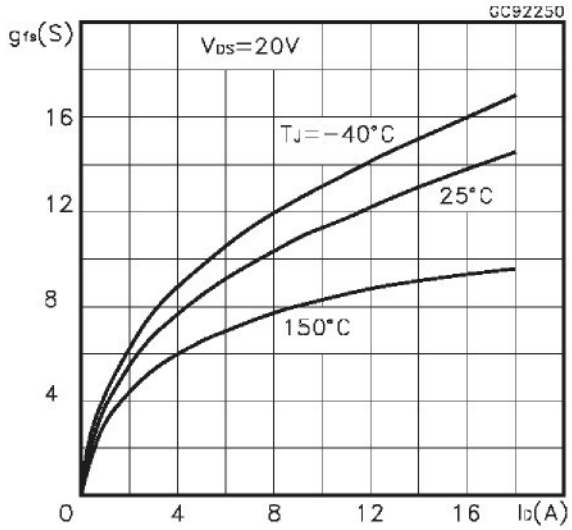
Output Characteristics



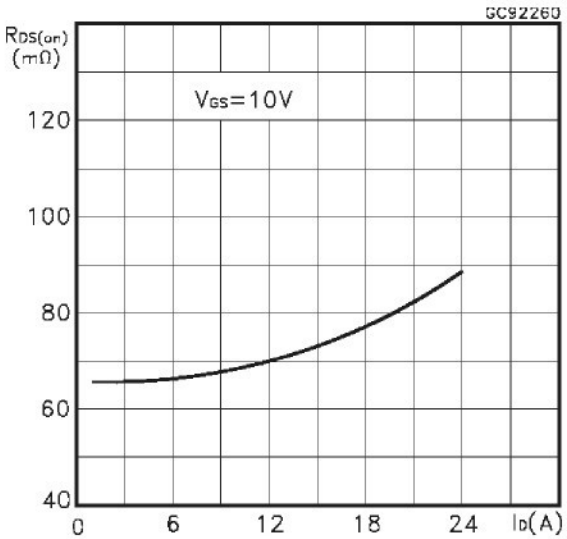
Transfer Characteristics



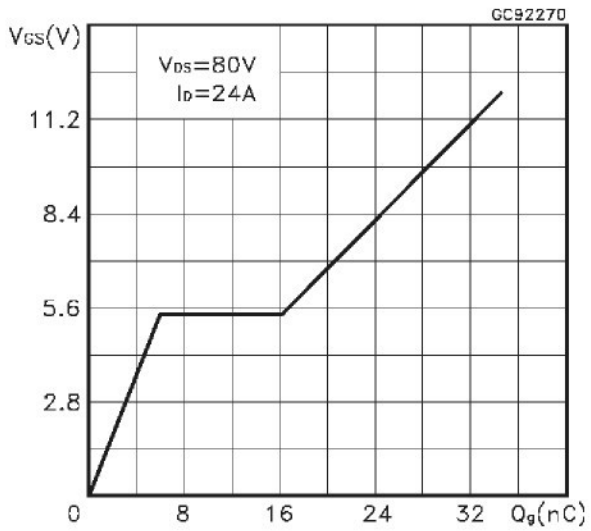
Transconductance



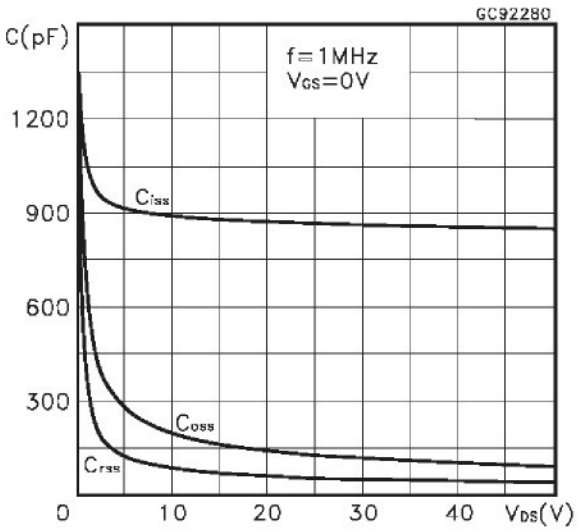
Static Drain-source On Resistance



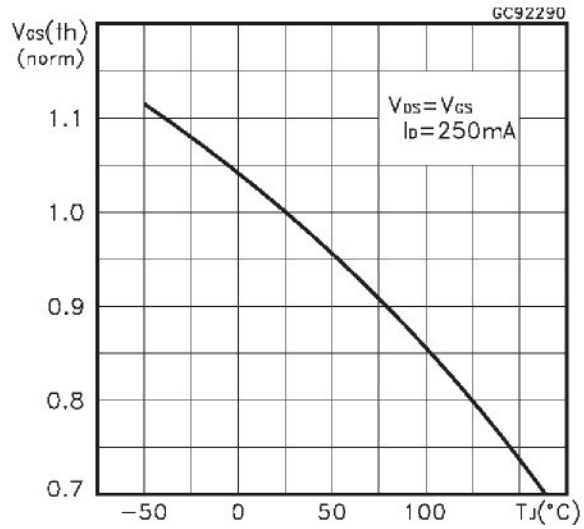
Gate Charge vs Gate-source Voltage



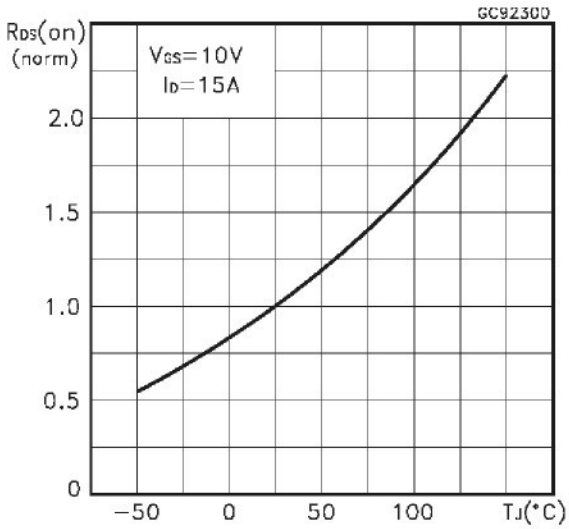
Capacitance Variations



Normalized Gate Threshold Voltage vs Temperature



Normalized on Resistance vs Temperature



Source-drain Diode Forward Characteristics

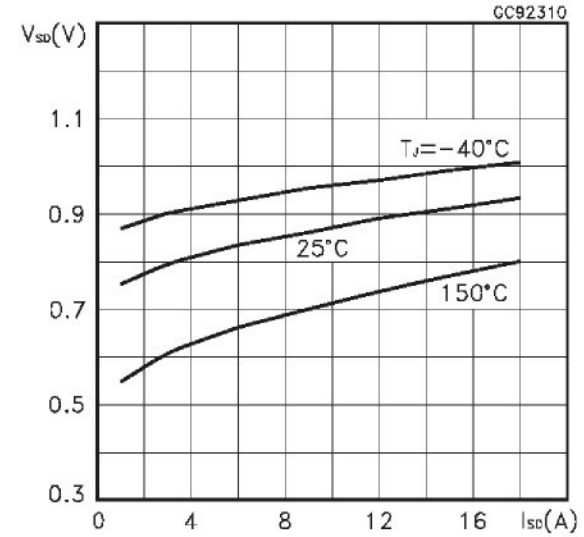


Fig. 1: Unclamped Inductive Load Test Circuit

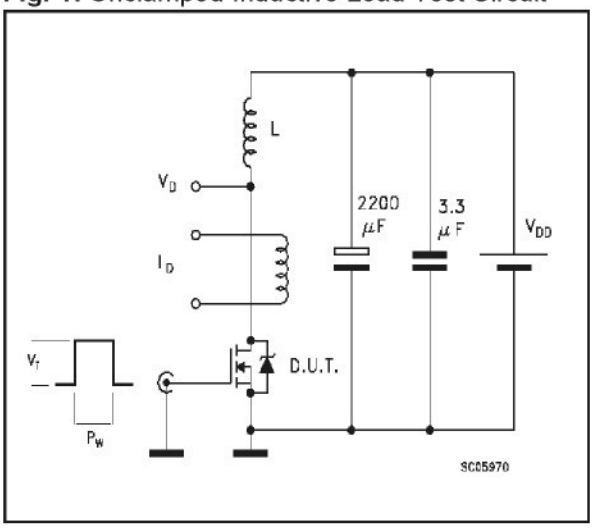


Fig. 2: Unclamped Inductive Waveform

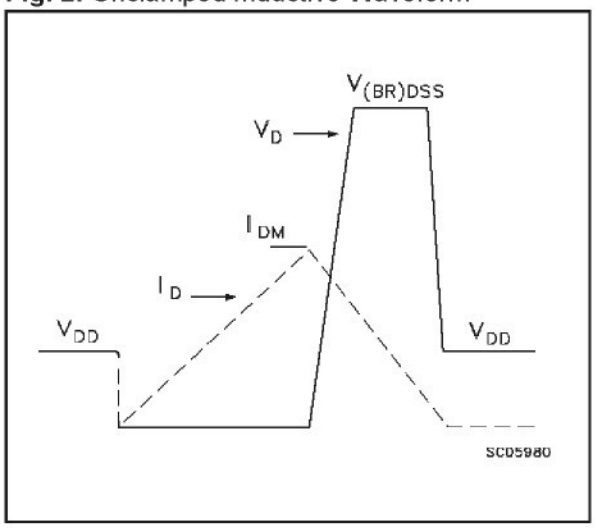


Fig. 3: Switching Times Test Circuits For Resistive Load

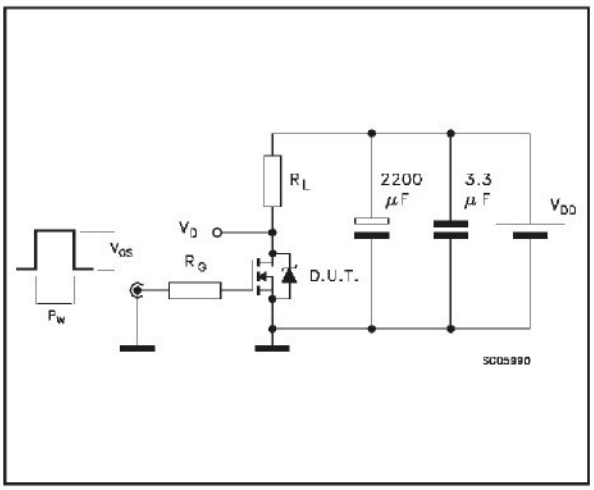


Fig. 4: Gate Charge test Circuit

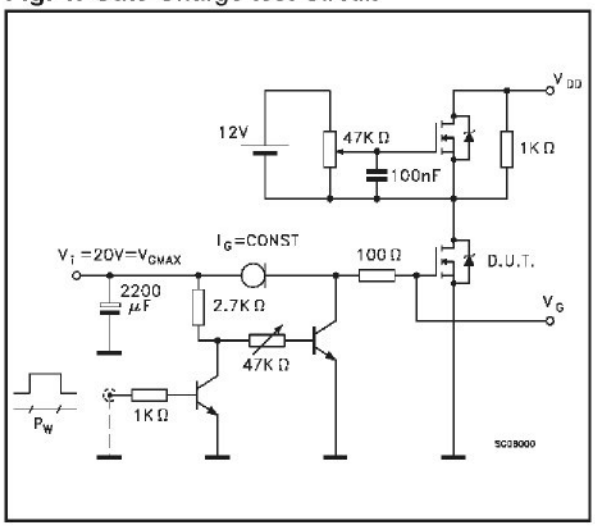
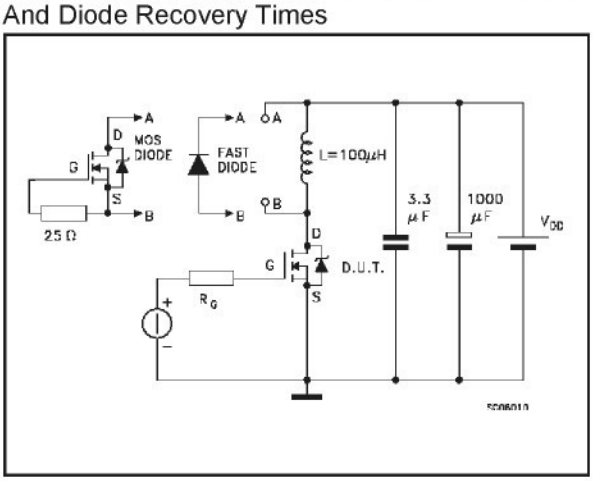
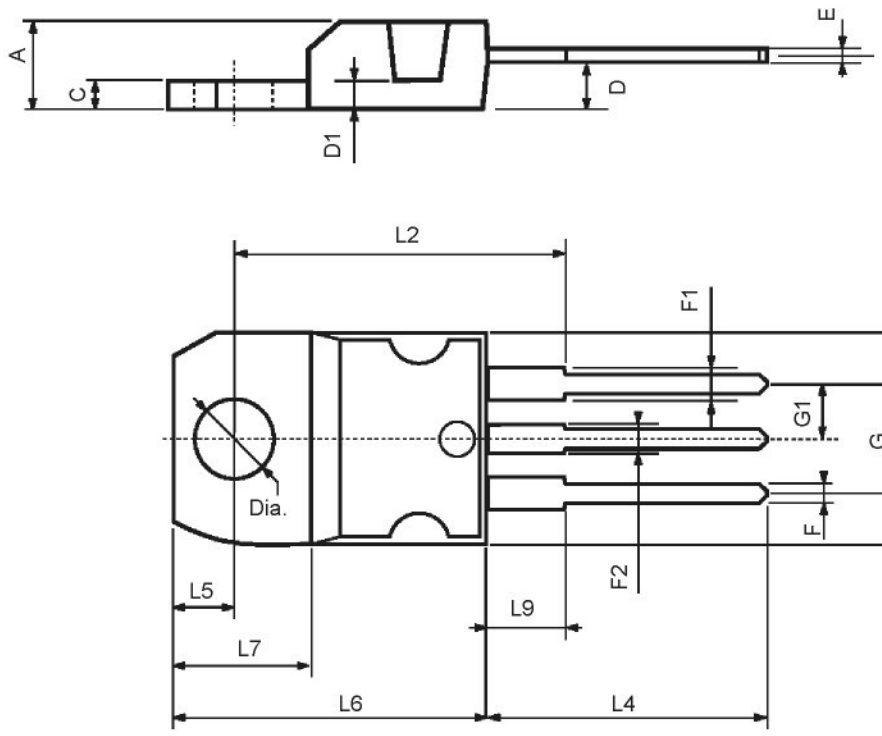


Fig. 5: Test Circuit For Inductive Load Switching And Diode Recovery Times



TO-220 MECHANICAL DATA

DIM.	mm			inch		
	MIN.	TYP.	MAX.	MIN.	TYP.	MAX.
A	4.40		4.60	0.173		0.181
C	1.23		1.32	0.048		0.051
D	2.40		2.72	0.094		0.107
D1		1.27			0.050	
E	0.49		0.70	0.019		0.027
F	0.61		0.88	0.024		0.034
F1	1.14		1.70	0.044		0.067
F2	1.14		1.70	0.044		0.067
G	4.95		5.15	0.194		0.203
G1	2.4		2.7	0.094		0.106
H2	10.0		10.40	0.393		0.409
L2		16.4			0.645	
L4	13.0		14.0	0.511		0.551
L5	2.65		2.95	0.104		0.116
L6	15.25		15.75	0.600		0.620
L7	6.2		6.6	0.244		0.260
L9	3.5		3.93	0.137		0.154
DIA.	3.75		3.85	0.147		0.151



Information furnished is believed to be accurate and reliable. However, STMicroelectronics assumes no responsibility for the consequences of use of such information nor for any infringement of patents or other rights of third parties which may result from its use. No license is granted by implication or otherwise under any patent or patent rights of STMicroelectronics. Specifications mentioned in this publication are subject to change without notice. This publication supersedes and replaces all information previously supplied. STMicroelectronics products are not authorized for use as critical components in life support devices or systems without express written approval of STMicroelectronics.

The ST logo is registered trademark of STMicroelectronics
® 2001 STMicroelectronics - All Rights Reserved

All other names are the property of their respective owners.

STMicroelectronics GROUP OF COMPANIES
Australia - Brazil - China - Finland - France - Germany - Hong Kong - India - Italy - Japan - Malaysia - Malta - Morocco - Singapore - Spain - Sweden - Switzerland - United Kingdom - U.S.A.

http://www.st.com

This datasheet has been download from:

www.datasheetcatalog.com

Datasheets for electronics components.

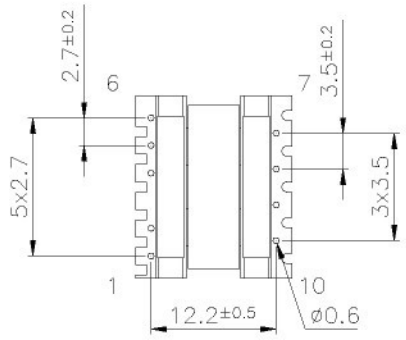
FLYBACK TRANSFORMERS
EE16 1 to 10 W

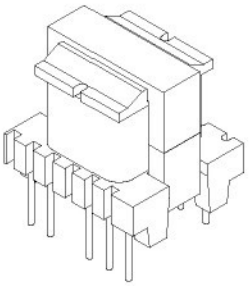
- Ambient Temperature ≤ 50°C
- Primary Reflected Voltage = 90 to 120V
- Dielectric Strength ≥ 3750Vac
- Creepage Distances ≥ 6mm
- Construction conforms to CEI950, CEI335, CEI61558 for reinforced insulation
- Secondaries may be series connected
- Output power can be delivered with any combination of secondaries within the max current limits.

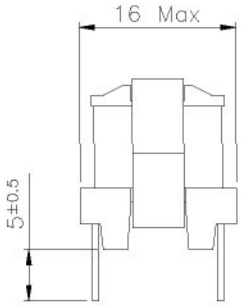
MYRRA Part N°	Control IC	Mains	Total output Power (max)	Outputs				Frequency	Primary Inductance	Pinout	Remarks
				S1		S2					
		Voltage Range Vac	Watts	Voltage Range Vdc	Max Current Adc	Voltage Range Vdc	Max Current Adc	kHz	µH		
74000	VIPer20	85 - 265	4	4,5 - 6	0,95	11 - 15	0,35	70	3900	Fig.1	
	TOP221P	85 - 265	4	3,3 - 7	1,2	8 - 17	0,35	100			
74001	VIPer20	85 - 265	6	5	1,2			70	3000	Fig.2	Note 1
	VIPer20	85 - 265	3	5	0,6			40			Note 1
	TOP221P	185 - 265	6	3 - 6	1,2			100			
	TOP221P	85 - 265	4,5	3 - 6	0,9			100			
	MC33369	85 - 265	6	3,3 - 6	1,2			100			
	TDA16831	185 - 265	6	4 - 6	1,2			100			
74002	VIPer20	85 - 265	6	14	0,43			70	3000	Fig.2	Note 1
	VIPer20	85 - 265	3	14	0,21			40			Note 1
	TOP221P	185 - 265	6	9 - 18	0,5			100			
	TOP221P	85 - 265	4,5	9 - 18	0,4			100			
	MC33369	85 - 265	6	10 - 18	0,5			100			
	TDA16831	185 - 265	6	11,5 - 17	0,5			100			
74003	VIPer20	85 - 265	6	2,5 - 3,5	1,8	4 - 5,2	1,2	70	3000	Fig.3	
	VIPer20	85 - 265	3	2,5 - 3,5	0,9	4 - 5,2	0,6	40			
	TOP221P	85 - 265	4,5	2 - 4	1,8	3 - 6	1,2	100			
	TOP221P	185 - 265	5	2 - 4	1,8	3 - 6	1,2	100			
	MC33369	85 - 265	6	2,5 - 4	1,8	3,8 - 6	1,2	100			
	TDA16831	185 - 265	6	3 - 4	1,8	4,5 - 6	1,2	100			
74010	VIPer20	85 - 265	8	4,5 - 6	1,6	11 - 15	0,7	70	1660	Fig.1	
	VIPer20	185 - 265	10	4,5 - 6	2	11 - 15	0,85	70			
	TOP222P	85 - 265	8	3,3 - 7	1,6	8 - 17	0,7	100			
	TOP222P	185 - 265	10	3,3 - 7	2	8 - 17	0,85	100			
	MC33369	85 - 265	8	4 - 7	1,6	10 - 17	0,7	100			
	MC33369	185 - 265	10	4 - 7	2	10 - 17	0,85	100			
	TDA16831	92 - 265	7,5	4,8 - 6	1,5	12 - 15	0,62	100			
	TDA16831	185 - 265	10	4,8 - 6	2	12 - 15	0,85	100			
	KA5L0165R	85 - 265	7	5 - 7	1,4	12,5 - 17	0,6	50			
	KA5H0165RN	185 - 265	10	5,6 - 7	1,8	14 - 17	0,7	100			

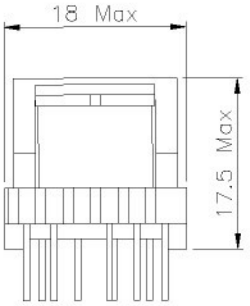
Note 1 : Regulation with auxiliary winding

FLYBACK TRANSFORMERS
1 to 10 W EE 16



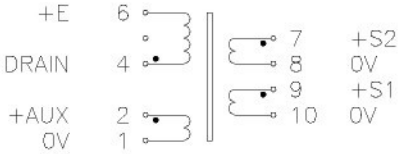






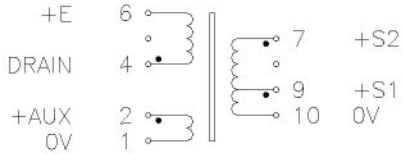
VIEW FROM PIN SIDE
PCB DRILLING DIAMETER 1.1mm

FIG.1



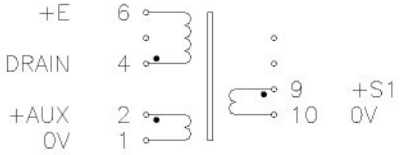
PIN 3 REMOVED

FIG.3



PIN 3 REMOVED

FIG.2



PIN 3 REMOVED

Certificate Number: Q10561 Certificate Number: E17276

SF31 - SF39

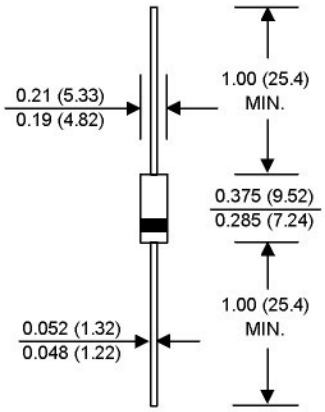
SUPER FAST RECTIFIER DIODES

PRV : 50 - 1000 Volts
Io : 3.0 Amperes

FEATURES :
* High current capability
* High surge current capability
* High reliability
* Low reverse current
* Low forward voltage drop
* Super fast recovery time

MECHANICAL DATA :
* Case : DO-201AD Molded plastic
* Epoxy : UL94V-O rate flame retardant
* Lead : Axial lead solderable per MIL-STD-202, Method 208 guaranteed
* Polarity : Color band denotes cathode end
* Mounting position : Any
* Weight : 1.21 grams

DO-201AD



Dimensions in inches and (millimeters)

MAXIMUM RATINGS AND ELECTRICAL CHARACTERISTICS

Rating at 25 °C ambient temperature unless otherwise specified.
Single phase, half wave, 60 Hz, resistive or inductive load.
For capacitive load, derate current by 20%.

RATING	SYMBOL	SF31	SF32	SF33	SF34	SF35	SF36	SF37	SF38	SF39	UNITS
Maximum Recurrent Peak Reverse Voltage	V_{RRM}	50	100	150	200	300	400	600	800	1000	Volts
Maximum RMS Voltage	V_{RMS}	35	70	105	140	210	280	420	560	700	Volts
Maximum DC Blocking Voltage	V_{DC}	50	100	150	200	300	400	600	800	1000	Volts
Maximum Average Forward Current 0.375"(9.5mm) Lead Length Ta = 55 °C	$I_{F(AV)}$	3.0									Amps.
Peak Forward Surge Current, 8.3ms Single half sine wave Superimposed on rated load (JEDEC Method)	I_{FSM}	125									Amps.
Maximum Peak Forward Voltage at I_F = 3.0 A.	V_F	0.95				1.4		1.7			Volts
Maximum DC Reverse Current Ta = 25 °C	I_R	5									µA
at Rated DC Blocking Voltage Ta = 100 °C	$I_{R(H)}$	50									µA
Maximum Reverse Recovery Time (Note 1)	Trr	35									ns
Typical Junction Capacitance (Note 2)	C_J	50									pf
Junction Temperature Range	T_J	- 65 to + 150									°C
Storage Temperature Range	T_{STG}	- 65 to + 150									°C

Notes :

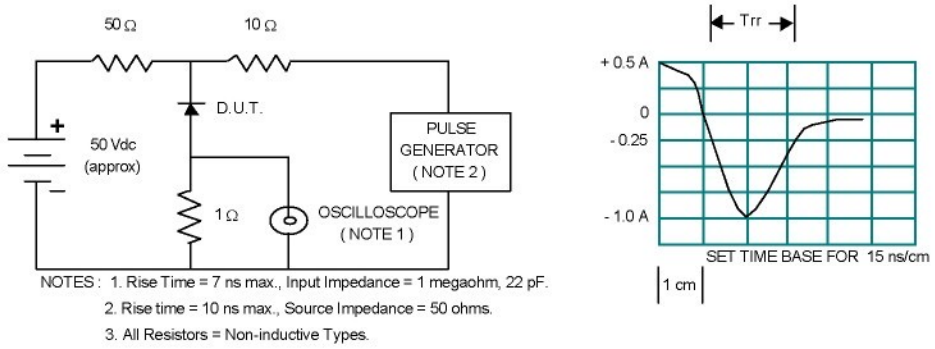
(1) Reverse Recovery Test Conditions : I_F = 0.5 A, I_R = 1.0 A, Irr = 0.25 A.

(2) Measured at 1.0 MHz and applied reverse voltage of 4.0 V_{DC}

UPDATE : APRIL 23, 1998

RATING AND CHARACTERISTIC CURVES (SF31 - SF39)

FIG.1 - REVERSE RECOVERY TIME CHARACTERISTIC AND TEST CIRCUIT DIAGRAM



NOTES : 1. Rise Time = 7 ns max., Input Impedance = 1 megaohm, 22 pF.
2. Rise time = 10 ns max., Source Impedance = 50 ohms.
3. All Resistors = Non-inductive Types.

FIG.2 - DERATING CURVE FOR OUTPUT RECTIFIED CURRENT

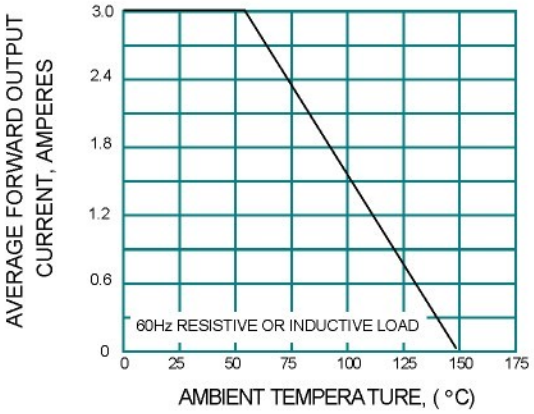


FIG.3 - MAXIMUM NON-REPETITIVE PEAK FORWARD SURGE CURRENT

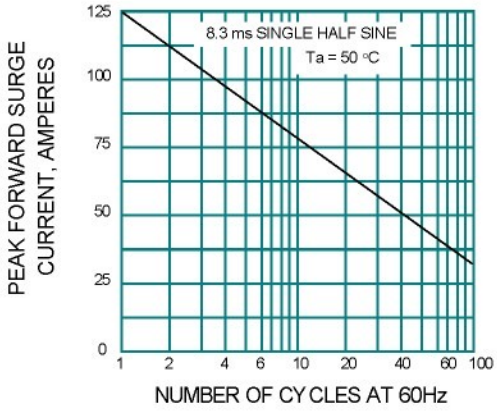


FIG.4 - TYPICAL FORWARD CHARACTERISTICS

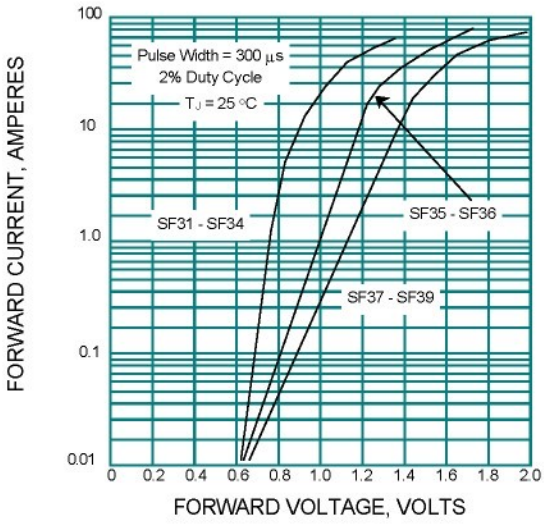
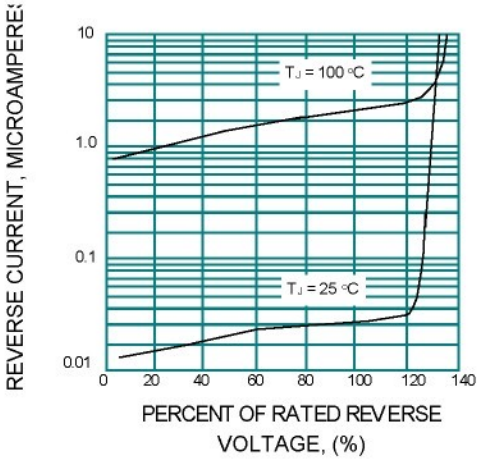


FIG.5 - TYPICAL REVERSE CHARACTERISTICS



This datasheet has been download from:

www.datasheetcatalog.com

Datasheets for electronics components.

Second Edition

Switching Power Supply Design

Abraham I. Pressman

Copyrighted material

Switching Power Supply Design

Abraham I. Pressman
President, Switchtronix Power, Inc.
Waban, Massachusetts

Second Edition

McGraw-Hill
New York San Francisco Washington, D.C. Auckland Bogotá
Caracas Lisbon London Madrid Mexico City Milan
Montreal New Delhi San Juan Singapore
Sydney Tokyo Toronto

Library of Congress Cataloging-in-Publication Data

Pressman, Abraham I.
Switching power supply design / Abraham I. Pressman. — 2nd ed.
p. cm.
Includes index.
ISBN 0-07-052236-7
1. Switching power supplies. 2. Electronic apparatus and appliances—Power supply. 3. Microelectronics—Power supply. 4. Electric current converters. I. Title.
TK7868.P6P75 1998
621.381'044—dc21 97-31668
CIP

McGraw-Hill

A Division of The McGraw-Hill Companies

Copyright © 1998 by The McGraw-Hill Companies, Inc. All rights reserved. Printed in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher.

5 6 7 8 9 0 DOC/DOC 0 2 1 0

ISBN 0-07-052236-7

The sponsoring editor for this book was Steve Chapman, the editing supervisor was Scott Amerman, and the production supervisor was Sherri Souffrance. It was set in Century Schoolbook by Ron Painter of McGraw-Hill's Professional Book Group composition unit.

Printed and bound by R. R. Donnelley & Sons Company.

McGraw-Hill books are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. For more information, please write to the Director of Special Sales, McGraw-Hill, Professional Publishing, Two Penn Plaza, New York, NY 10121-2298. Or contact your local bookstore.

This book is printed on recycled, acid-free paper containing a minimum of 50% recycled, de-inked fiber.

Information contained in this work has been obtained by The McGraw-Hill Companies, Inc. ("McGraw-Hill") from sources believed to be reliable. However, neither McGraw-Hill nor its authors guarantee the accuracy or completeness of any information published herein and neither McGraw-Hill nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that McGraw-Hill and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

Contents

Preface xix

Part 1 Topologies

Chapter 1. Fundamental Switching Regulators—Buck, Boost, and Inverter Topologies 3

1.1 Introduction 3
1.2 Linear Regulators—Swtiching Regulator Ancestors 4
1.2.1 Basic operation—merits and drawbacks 4
1.2.2 Linear regulator drawbacks 5
1.2.3 Power dissipation in the series-pass transistor 5
1.2.4 Linear regulator efficiency versus output voltage 6
1.2.5 Linear regulators with PNP series-pass transistors for lesser required headroom 8
1.3 "Buck" Switching Regulator Topology 9
1.3.1 Basic operation 9
1.3.2 Significant current waveforms in buck regulator 12
1.3.3 Buck regulator efficiency neglecting AC switching losses 13
1.3.4 Buck regulator efficiency including AC switching losses 13
1.3.5 Optimum switching frequency in buck regulator 16
1.3.6 Design relations—output filter inductor selection 17
1.3.7 Design relations—output filter capacitor selection 21
1.3.8 DC-isolated, regulated voltage from a buck regulator 23
1.4 Boost Switching Regulator Topology 24
1.4.1 Basic operation 24
1.4.2 Quantitative relations—boost regulator 26
1.4.3 Discontinuous and continuous modes in boost regulator 27
1.4.4 Discontinuous-mode boost regulator design relations 28
1.4.5 Boost regulator applications and flyback comparison 31
1.5 Polarity Inverting Switching Regulator Topology 32
1.5.1 Basic operation 32
1.5.2 Design relations in polarity inverter 34
Reference 35

Chapter 2. Push-Pull and Forward Converter Topologies 37

2.1 Introduction 37
2.2 Push-Pull Topology 37
2.2.1 Basic operation—master/slave outputs 37
2.2.2 Slave line-load regulation 40
2.2.3 Slave absolute output voltage levels 41
2.2.4 Master output inductor minimum current limitations 41
2.2.5 Flux imbalance in push-pull topology 42
2.2.6 Indications of flux imbalance 45
2.2.7 Testing for flux imbalance 48
2.2.8 Coping with flux imbalance 49
2.2.9 Power transformer design relations 51
2.2.10 Primary, secondary peak and rms currents 55
2.2.11 Transistor voltage stress and leakage inductance spikes 58
2.2.12 Power transistor losses 60
2.2.13 Output power and input voltage limitations in push-pull topology 63
2.2.14 Output filter design relations 64
2.3 Forward Converter Topology 66
2.3.1 Basic operation 66
2.3.2 Design relations: output/input voltage, on time, turns ratios 70
2.3.3 Slave output voltages 71
2.3.4 Secondary load, free-wheeling diode, and inductor currents 72
2.3.5 Relations between primary current, output power, and input voltage 73
2.3.6 Maximum off-voltage stress in power transistor 74
2.3.7 Practical input voltage/output power limits 74
2.3.8 Forward converter with unequal power and reset winding turns 75
2.3.9 Forward converter magnetics 78
2.3.10 Power transformer design relations 81
2.3.11 Output filter design relations 84
2.4 Double-Ended Forward Converter Topology 86
2.4.1 Basic operation 86
2.4.2 Design relations and transformer design 88
2.5 Interleaved Forward Converter Topology 89
2.5.1 Basic operation—merits, drawbacks, and output power limits 89
2.5.2 Transformer design relations 92
2.5.3 Output filter design 92

Chapter 3. Half- and Full-Bridge Converter Topologies 93

3.1 Introduction 93
3.2 Half-Bridge Converter Topology 93
3.2.1 Basic operation 93
3.2.2 Half-bridge magnetics 95
3.2.3 Output filter calculations 97
3.2.4 Blocking capacitor to avoid flux imbalance 97
3.2.5 Half-bridge leakage inductance problems 98
3.2.6 Double-ended forward converter versus half bridge 99

3.2.7 Practical output power limits in half bridge 100
3.3 Full-Bridge Converter Topology 101
3.3.1 Basic operation 101
3.3.2 Full-bridge magnetics 103
3.3.3 Output filter calculations 104
3.3.4 Transformer primary blocking capacitor 104

Chapter 4. Flyback Converter Topologies 105

4.1 Introduction 105
4.2 Flyback Converter—Areas of Application 105
4.3 Discontinuous-Mode Flybacks—Basic Operation 107
4.3.1 Relation between output voltage versus input voltage, on time, output load 108
4.3.2 Design relations and sequential decision requirements 109
4.3.3 Flyback magnetics 114
4.3.4 Flyback disadvantages 121
4.3.5 Flybacks for 120- or 220-V-AC operation with no doubler and full-wave rectifier switching 125
4.4 Continuous-Mode Flybacks—Basic Operation 127
4.4.1 Discontinuous-mode to continuous-mode transition 130
4.4.2 Design relations—continuous-mode flybacks 132
4.5 Interleaved Flybacks 137
4.5.1 Summation of secondary currents in interleaved flybacks 138
4.6 Double-Ended Discontinuous-Mode Flyback 138
4.6.1 Area of application 138
4.6.2 Basic operation 139
4.6.3 Leakage inductance effect in double-ended flyback 140
References 141

Chapter 5. Current-Mode and Current-Fed Topologies 143

5.1 Introduction 143
5.2 Current-Mode Advantages 144
5.2.1 Avoidance of flux imbalance in push-pull converters 144
5.2.2 Instantaneous correction against line voltage changes without the delay in an error amplifier (voltage feedforward characteristics) 145
5.2.3 Ease and simplicity of feedback-loop stabilization 145
5.2.4 Paralleling outputs 146
5.2.5 Improved load current regulation 146
5.3 Current-Mode versus Voltage-Mode Control Circuits 146
5.3.1 Voltage-mode control circuitry 146
5.3.2 Current-mode control circuitry 149
5.4 Detailed Explanation of Current-Mode Advantages 154
5.4.1 Line voltage regulation 154
5.4.2 Elimination of flux imbalance 154
5.4.3 Simplified loop stabilization resulting from elimination of output inductor in small-signal analysis 155
5.4.4 Mechanism of load current regulation 156
5.5 Current-Mode Deficiencies and Problems 158
5.5.1 Constant peak versus constant average output inductor problems 158

5.5.2 Response to an output inductor current disturbance 161
5.5.3 Slope compensation to correct problems in current mode 161
5.5.4 Slope compensation with a positive-going ramp voltage 163
5.5.5 Implementing slope compensation 164
5.6 Voltage-Fed and Current-Fed Topologies 165
5.6.1 Introduction and definitions 165
5.6.2 Deficiencies of voltage-fed, width-modulated full-wave bridge 166
5.6.3 Buck-voltage-fed full-wave bridge topology-basic operation 170
5.6.4 Buck-voltage-fed full-wave bridge advantages 172
5.6.5 Drawbacks in buck-voltage-fed full-wave bridge 174
5.6.6 Buck-current-fed full-wave bridge topology-basic operation 175
5.6.7 Flyback-current-fed push-pull topology 189
References 208

Chapter 6. Miscellaneous Topologies 211

6.1 SCR Resonant Topologies—Introduction 211
6.2 SCR Basics 213
6.3 SCR Turnoff by Resonant Sinusoidal Anode Currents—Single-Ended Resonant Inverter Topology 216
6.4 SCR Resonant Bridge Topologies—Introduction 222
6.4.1 SCR half-bridge resonant converter, series-loaded—basic operation 224
6.4.2 Design calculations—SCR half-bridge resonant converter, series-loaded 226
6.4.3 Design example—SCR half-bridge resonant converter, series-loaded 229
6.4.4 SCR half-bridge resonant converter, shunt-loaded 230
6.4.5 Single-ended SCR resonant converter topology design 230
6.5 Cuk Converter Topology—Introduction 237
6.5.1 Cuk converter—basic operation 237
6.5.2 Relation between output/input voltage and Q on time 239
6.5.3 Rates of change of currents in $L1$, $L2$ 240
6.5.4 Reducing input ripple currents to zero 241
6.5.5 Isolated outputs in the Cuk converter 242
6.6 Low Output Power "Housekeeping" or "Auxillary" Topologies—Introduction 242
6.6.1 Housekeeping power supply—on output or input ground 243
6.6.2 Housekeeping supply alternatives 244
6.6.3 Specific housekeeping supply block diagrams 245
6.6.4 Royer oscillator housekeeping supply—basic operation 248
6.6.5 Minimum-parts-count flyback as a housekeeping supply 260
6.6.6 Buck regulator with DC-isolated output as a housekeeping supply 262
References 263

Part 2 Magnetics and Circuits Designs

Chapter 7. Transformer and Magnetics Design 267

7.1 Introduction 267
7.2 Transformer Core Materials and Geometries and Peak Flux Density Selection 268
7.2.1 Ferrite core losses versus frequency and flux density for widely used core materials 268
7.2.2 Ferrite core geometries 271
7.2.3 Peak flux density selection 275
7.3 Maximum Transformer Core Output Power, Peal Flux Density, Core and Bobbin Areas, and Coil Current Density 277
7.3.1 Derivation of output power relations for forward converter topology 277
7.3.2 Derivation of output power relations for push-pull topology 280
7.3.3 Derivation of output power relations for half-bridge topology 286
7.3.4 Output power relations in full-bridge topology 287
7.3.5 Conversion of output power equations into charts permitting core and operating frequency selection at a glance 288
7.4 Transformer Temperature Rise Calculations 294
7.5 Transformer Copper Losses 298
7.5.1 Introduction 298
7.5.2 Skin effect 300
7.5.3 Skin effect—quantitative relations 301
7.5.4 AC/DC resistance ratio for various wire sizes at various frequencies 303
7.5.5 Skin effect with rectangular current waveshapes 306
7.5.6 Proximity effect 308
References 317

Chapter 8. Bipolar Power Transistor Base Drives 319

8.1 Introduction 319
8.2 Objectives of Bipolar Base Drive Circuits 320
8.2.1 Sufficiently high current throughout the on time 320
8.2.2 A spike of high base input current I_{b1} at instant of turnon 321
8.2.3 A spike of high reverse base current I_{b2} at the instant of turnoff 322
8.2.4 A base-to-emitter reverse voltage spike -1 to -5 V in amplitude at the instant of turnoff 324
8.2.5 A scheme to permit the circuit to work equally well with high- or low-beta transistors 324
8.2.6 High efficiency 325
8.3 Baker Clamps 325

8.3.1 Baker clamp operation 328
8.3.2 Transformer coupling into a Baker clamp 330
8.3.3 Transformerized Baker clamp 336
8.3.4 Inherent Baker clamping in Darlington transistors 338
8.3.5 Proportional base drive 329
8.3.6 Miscellaneous base drive schemes 345
References 351

Chapter 9. MOSFET Power Transistors and Input Drive Circuits 353

9.1 Introduction 353
9.2 MOSFET Basics 354
9.2.1 MOSFET drain current versus drain-to-source voltage characteristics ($I_d - V_{dc}$) 356
9.2.2 MOSFET input impedance and required gate currents 359
9.2.3 MOSFET gate rise and fall times for desired drain current rise and fall times 361
9.2.4 MOSFET gate drive circuits 362
9.2.5 MOSFET R_{dc} temperature characteristics and safe operating area limits 365
9.2.6 MOSFET gate threshold voltage and temperature characteristic 369
9.2.7 MOSFET switching speed and temperature characteristics 370
9.2.8 MOSFET current ratings 371
9.2.9 Parallelling MOSFETs 374
9.2.10 MOSFETs in push-pull topology 377
9.2.11 MOSFET maximum gate voltage specifications 378
9.2.12 MOSFET drain-to-source ("body") diode 379
References 380

Chapter 10. Magnetic-Amplifier Postregulators 381

10.1 Introduction 381
10.2 Linear and Buck Regulator Postregulators 382
10.3 Magnetic Amplifiers—Introduction 383
10.3.1 Square hysteresis loop magnetic core as a fast acting on/off switch with electrically adjustable on/off time 385
10.3.2 Blocking and firing times in magnetic-amplifier postregulators 388
10.3.3 Magnetic-amplifier core resetting and voltage regulation 389
10.3.4 Slave output voltage shutdown with magnetic amplifiers 390
10.3.5 Square hysteresis loop core characteristics and sources 391
10.3.6 Core loss and temperature rise calculations 396
10.3.7 Design example—magnetic-amplifier postregulator 402
10.3.8 Magnetic-amplifier gain 407
10.3.9 Magnetic amplifiers for a push-pull output 408
10.4 Magnetic Amplifier Pulse-Width Modulator and Error Amplifier 408
10.4.1 Circuit details, magnetic amplifier pulse-width modulator-error amplifier 409
References 412

Chapter 11. Turnon, Turnoff Switching Losses and Snubbers 413

11.1 Introduction 413
11.2 Transistor Turnoff Losses without a Snubber 414
11.3 *RCD* Turnoff Snubber Operation 416
11.4 Selection of Capacitor Size in *RCD* Snubber 418
11.5 Design Example—*RCD* Snubber 419
11.5.1 *RCD* snubber returned to positive supply rail 420
11.6 Nondissipative Snubbers 421
11.7 Snubber Reduciton of Leakage Inductance Spike to Avoid Second Breakdown 422
11.8 Transformer-Aided Snubber 425
References 426

Chapter 12. Feedback-Loop Stabilization 427

12.1 Introduction 427
12.2 Mechanism of Loop Oscillation 429
12.2.1 Gain citerion for a stable circuit 429
12.2.2 Gain slope criteria for a stable circuit 430
12.2.3 Gain characteristic of LC output filter with and without equivalent series resistance (ESR) in output capacitor 433
12.2.4 Pulse-width-modulator gain 436
12.2.5 Total output LC filter plus modulator and sampling network gain 437
12.3 Shaping Error-Amplifier Gain-versus-Frequency Characteristic 437
12.4 Error-Amplifier Transfer Fuction, Poles and Zeros 440
12.5 Rules for Gain Slope Changes Due to Zero and Pole Frequencies 442
12.6 Derivation of Transfer Function of an Error Amplifier with Single Zero and Single Pole from Its Schematic 444
12.7 Calculation of Type 2 Error-Amplifier Phase Shift from Its Zero and Pole Locations 445
12.8 Phase Shift through LC Filter Having ESR in Its Output Capacitor 446
12.9 Design Example—Stabilizing a Forward Converter Feedback Loop with a Type 2 Error Amplifier 448
12.10 Type 3 Error Amplifier—When Used and Transfer Function 451
12.11 Phase Lag through a Type 3 Error Amplifier as Function of Zero and Pole Locations 453
12.12 Type 3 Error Amplifier Schematic, Transfer Function, and Zero and Pole Locations 454
12.13 Design Example—Stabilizing a Forward Converter Feedback Loop with a Type 3 Error Amplifier 456
12.14 Component Selection to Yield Desired Type 3 Error-Amplifer Gain Curve 458
12.15 Conditional Stability in Feedback Loops 459
12.16 Stabilizing a Discontinuous-Mode Flyback Converter 461
12.16.1 DC gain from error-amplifier output to output voltage node 461
12.16.2 Discontinuous-mode flyback transfer function or AC voltage gain from error-amplifier output to output voltage node 462

12.17 Error-Amplifier Transfer Function for Discontinuous-Mode Flyback 464
12.18 Design Example—Stabilizing a Discontinuous-Mode Flyback Converter 465
12.19 Transconductance Error Amplifiers 468
References 470

Chapter 13. Resonant Converters 471

13.1 Introduction 471
13.2 Resonant Forward Converter 473
13.2.1 Measured waveforms in a resonant forward converter 476
13.3 Resonant Converter Operating Modes 479
13.3.1 Discontinuous and continuous; above resonance and below resonance operating modes 479
13.4 Resonant Half Bridge in Continuous-Conduction Mode 481
13.4.1 Parallel resonant converter (PRC) and series resonant converter (SRC) 481
13.4.2 AC equivalent circuits and gain curves for series- and parallel-loaded half bridges in continuous-conduction mode 482
13.4.3 Regulation with series-loaded half bridge in continuous-conduction mode (CCM) 484
13.4.4 Regulation with parallel-loaded half bridge in continuous-conduction mode 485
13.4.5 Series-parallel resonant converter in continuous-conduction mode 486
13.4.6 Zero-voltage-switching quasi-resonant (CCM) converters 488
13.5 Resonant Power Supplies—Conclusion 490
References 492

Part 3 Typical Switching Power Supply Waveforms

Chapter 14. Waveforms 495

14.1 Introduction 495
14.2 Forward Converter Waveshapes 496
14.2.1 V_{dc}, I_d photos at 80 percent of full load 497
14.2.2 V_{dc}, I_{dc} photos at 40 percent of full load 499
14.2.3 Overlap of drain voltage and drain current at turnon/turnoff transitions 499
14.2.4 Relative timing of drain current, drain-to-source voltage, and gate-to-source voltage 502
14.2.5 Relative timing of input voltage to output inductor, output inductor current rise and fall times, and power transistor drain-to-source voltage 502
14.2.6 Relative timing of critical waveforms in PWM driver chip (UC3525A) for forward converter of Fig. 14.1 503
14.3 Push-Pull Topology Waveshapes—Introduction 504
14.3.1 Transformer center tap currents and drain-to-source voltages at maximum load currents for maximum, nominal, and minimum supply voltages 505

14.3.2 Opposing V_{dc} waveshapes, relative timing, and flux locus during dead time 509
14.3.3 Relative timing of gate input voltage, drain-to-source voltage, and drain currents 511
14.3.4 Drain current as measured with a current probe in series in the drain as compared with measurement with a current probe in series in the transformer center tap 511
14.3.5 Output ripple voltage and rectifier cathode voltage 511
14.3.6 Oscillatory ringing at rectifier cathodes at transistor turnon 513
14.3.7 AC switching loss due to overlap of falling drain current and rising drain voltage at turnoff 515
14.3.8 Drain currents as measured in the transformer center tap and drain-to-source voltages at one-fifth of maximum output power 515
14.3.9 Drain current as measured in drain lead and drain voltage at one-fifth of maximum output power 519
14.3.10 Relative timing of opposing drain voltages at one-fifth of maximum output currents 519
14.3.11 5-V output inductor current and rectifier cathode voltage 519
14.3.12 5-V rectifier cathode voltage at output current higher than minimum 520
14.3.13 Gate input and drain current timing 520
14.3.14 Rectifier diode and transformer secondary currents 520
14.3.15 Apparent double turnon per half period arising from too high a magnetizing current or too low DC output currents 520
14.3.16 Drain currents and drain-to-source voltage at output power 15 percent above specified maximum 523
14.3.17 Ringing at drain during transistor dead time 523
14.4 Flyback Topology Waveshapes 523
14.4.1 Introduction 523
14.4.2 Drain current/drain-to-source waveshapes at 90 percent of full load for minimum, nominal, and maximum input voltages 525
14.4.3 Voltage and currents at output rectifier inputs 527
14.4.4 Snubber capacitor current at transistor turnoff 527

Part 4 Newer Applications for Switching Power Supply Technique

Chapter 15. Power Factor, Power Factor Correction 533

15.1 Power Factor—What Is It and Why Must It Be Corrected? 533
15.2 Power Factor Correction in Switching Power Supplies 535
15.3 Power Factor Correction—Basic Circuit Details 536
15.3.1 Continuous- versus discontinuous-mode boost topology for power factor correction 539
15.3.2 Line input voltage regulation in continuous-mode boost converters 541
15.3.3 Load current regulation in continuous-mode boost regulators 543
15.4 Integrated-Circuit Chips for Power Factor Correction 544

15.4.1 Unitrode UC 3854 power factor correction chip 544
15.4.2 Forcing sinusoidal line current with the UC 3854 546
15.4.3 Maintaining constant output voltage with UC 3854 548
15.4.4 Power output with the UC 3854 548
15.4.5 Boost switching frequency with the UC 3854 550
15.4.6 Selection of boost output inductor *L*1 550
15.4.7 Selection of boost output capacitor 552
15.4.8 Peak current limiting n the UC 3854 553
15.4.9 Stabilizing the UC 3854 feedback loop 554
15.5 Motorola MC 34261 Power Factor Correction Chip 554
15.5.1 Details of the Motorola MC 34261 (Fig. 15.11) 556
15.5.2 Logic details of the MC 34261 (Figs. 15.11 and 15.12) 557
15.5.3 Calculations for frequency and magnitude of inductor *L*1 558
15.5.4 Selection of current-sensing resistor (*R*9) and multiplier input resistor netowrk (*R*3, *R*7) for the MC 34261 560
References 561

Chapter 16. High-Frequency Power Sources for Fluorescent Lamps 563

16.1 Why High-Frequency Power Sources? 563
16.2 Fluorescent Lamp—Physics and Types 567
16.3 Electric Arc Characteristics 569
16.3.1 Arc characteristics with DC supply voltage at the electrodes 571
16.3.2 AC-driven fluorescent lamps 573
16.3.3 Fluorescent lamp volt/ampere characteristics 574
16.4 Electronic Ballast Circuits 576
16.5 DC/AC Inverter—General Characteristics 579
16.6 DC/AC Inverter Topologies 580
16.6.1 Current-fed push-pull topology 582
16.6.2 Voltages and currents in current-fed push-pull topology 585
16.6.3 Magnitude of "current feed" inductor in current-fed topology 586
16.6.4 Specific core selection for current feed inductor 587
16.6.5 Coil design for current feed inductor 593
16.6.6 Ferrite core transformer for current-fed topology 594
16.6.7 Toroidal core transformer for current-fed topology 600
16.7 Voltage-Fed Push-Pull Topology 601
16.8 Current-Fed Parallel Resonant Half-Bridge Topology 604
16.9 Voltage-Fed Series Resonant Half-Bridge Topology 606
16.10 Electronic Ballast Packageing 608
References 608

Chapter 17. Low-Input-Voltage Regulators for Laptop Computers and Portable Electronics 611

17.1 Introduction 611
17.1.1 Low-input-voltage regulator suppliers 612
17.2 Linear Technology Boost and Buck Regulators 613
17.2.1 Linear Technology LT1170 Boost regulator 613
17.2.2 Significant Polaroid waveforms in the LT1170 boost regulator 615
17.2.3 Thermal considerations in IC regulators 623

17.2.4 Alternative uses for the LT1170 boost regulator 625
17.2.5 Additional LTC high-power boost regulators 630
17.2.6 Component selection for boost regulators 631
17.2.7 Linear Technology buck regulator family 634
17.2.8 Alternative uses for the LT1074 buck regulator 640
17.2.9 Higher-efficiency LTC high-power buck regulators 643
17.2.10 Summary of high-power Linear Technology buck regulators 654
17.2.11 Linear Technology micropower regulators 654
17.2.12 Feedback loop stabilization 654
17.3 Maxim IC Regulators 661
17.4 Distributed Power Systems with IC Building Blocks 665
References 668

Appendix 669
Bibliography 670
Index 671

Preface

Much has changed in the power conversion field since the author's previous book on switching power supply design (*Switching and Linear Power Supply, Power Converter Design*, 1978).

The main goal of these changes has been to make the power supply smaller. As integrated circuits packed more features in a smaller volume, it became essential also to decrease the size of a system's power supply. Switching power supply packing density currently ranges from 2 to 6 load watts per cubic inch (W/in^3) as compared to about 1 W/in^3 a decade ago. Newer resonant converter techniques offer a possibility of 20 to 40 W/in^3.

Higher switching frequencies, made possible by power MOSFET transistors, newer topologies, and integrated-circuit pulse-width-modulating chips, which pack more control and supervisory features in a smaller volume, have contributed to making present-day power supplies smaller. All these new technologies are covered in this book.

It has been the author's experience, in teaching a course on modern switching power supply design to all levels of engineers from most of the major American electronics companies, that those who have a good understanding of the fundamental principles can easily solve their day-to-day design problems as well as assess and adapt to new technologies.

Thus, this book covers the new technologies in a tutorial way so that the reader can understand the fundamental reasons for various effects. Explanations for various significant waveshapes and explanations for alternative design decisions are given. Care has been taken to avoid offering "handy-dandy" design equations without showing how they were derived. All equations that effect a design decision are derived from fundamental relations.

Magnetics design is emphasized; for, the majority of power supply designers are primarily circuits-oriented. They appreciate and can analyze problems when they can see DC voltage levels, voltage spikes,

and waveshapes on an oscilloscope. But the locus of operation on a BH loop cannot be seen on an oscilloscope. Thus, circuits designers too often shy away from or do not fully understand magnetics design. They leave that to a magnetics specialist who may not appreciate how circuits characteristics effect magnetics design decisions. It is hoped the emphasis on magnetics herein may help correct that problem.

This second edition of the book also includes chapters on the currently hottest topics in the field: power factor correction, high-frequency ballasts for fluorescent lamps, and low-input voltage supplies for laptop computers.

Throughout this text and throughout the literature, idealized voltage-current waveforms are often shown. But I feel it is very valuable from a tutorial viewpoint to view actual Polaroid waveforms taken at critical points on working circuits. Seeing an actual photographed waveform with its spikes, rings, and oddities conveys a great deal more information and confidence about a circuit than does an idealized, hand-drawn waveform. Such photographed waveforms are presented at critical points at various frequencies on some of the most commonly used topologies.

This book is directed primarily to design engineers and engineering students at the undergraduate and graduate level. It may also be of significant value to people who are not directly involved in start-from-scratch designs, and whose main interest is power supply design analysis, design review, test, and debugging.

The material contained herein is a consolidation and, hopefully, a simpler, clearer explanation and logical reorganization of the high-points of all the proven and practical aspects of modern switching power supply technology. As such, credit for the material belongs to the innumerable engineers, designers in industry, and universities that have brought switching power supply technology to its present advanced state.

Abraham I. Pressman

Part

1

Topologies

Chapter

1

Fundamental Switching Regulators—Buck, Boost, and Inverter Topologies

1.1 Introduction

There are about 14 basic topologies (basic block diagrams) commonly used to implement a switching power supply. Each topology has unique properties which make it best suited for certain applications.

Some are best used for AC/DC off-line converters at low (< 200 W) output power, some at higher output power. Some are a better choice for high AC input voltages (≥ 220 V AC); some are better for an AC input of 120 V or less. Some have advantages for higher DC output voltages (> ~ 200 V) or in applications where there are more than four or five different output voltages. Some have a lower parts count than do others for the same output power or offer a better tradeoff in parts count versus reliability. Lesser input or output ripple and noise is a frequent factor in a topology selection.

Similarly, some topologies are better used for DC/DC converters—again with high and low output power, high or low output voltages, and minimum parts count as significant selection criteria. Further, some topologies have inherent design drawbacks which require additional circuitry or more complex circuitry, not completely analyzable in various worst-case situations.

Thus to make the best choice of a topology, it is essential to be familiar with the merits, drawbacks, and areas of usage of all topologies. A poor choice of a topology can doom a new power supply design at the very outset.

In this chapter, the schematics of the fundamental and earliest topologies—the buck, boost, and polarity inverter—are described. Their basic operation is described; critical waveforms are shown and ex-

plained; and merits, drawbacks, and application areas are discussed. Peak transistor current and voltage stresses versus output power and input voltage and their tolerances are discussed. The dependence of input current on output power and input voltage is shown. Efficiency and DC and AC switching losses are discussed.

To show the origin and need for switching regulators, the discussion starts with what preceded them—linear or *series-pass* regulators.

1.2 Linear Regulators—Switching Regulator Ancestors

1.2.1 Basic operation—merits and drawbacks

The basic circuit is shown in Fig. 1.1*a*. It consists of an electrically variable resistance in the form of a transistor (operating in the linear mode) in series with the output load. An error amplifier senses the DC output voltage via a sampling resistor network $R1$, $R2$ and compares it with a reference voltage V_{ref}. The error amplifier output voltage drives the base of the series-pass power transistor via a current amplifier.

The phasing is such that if the DC output voltage goes up (as a result of either an increase in input voltage or a decrease in output load

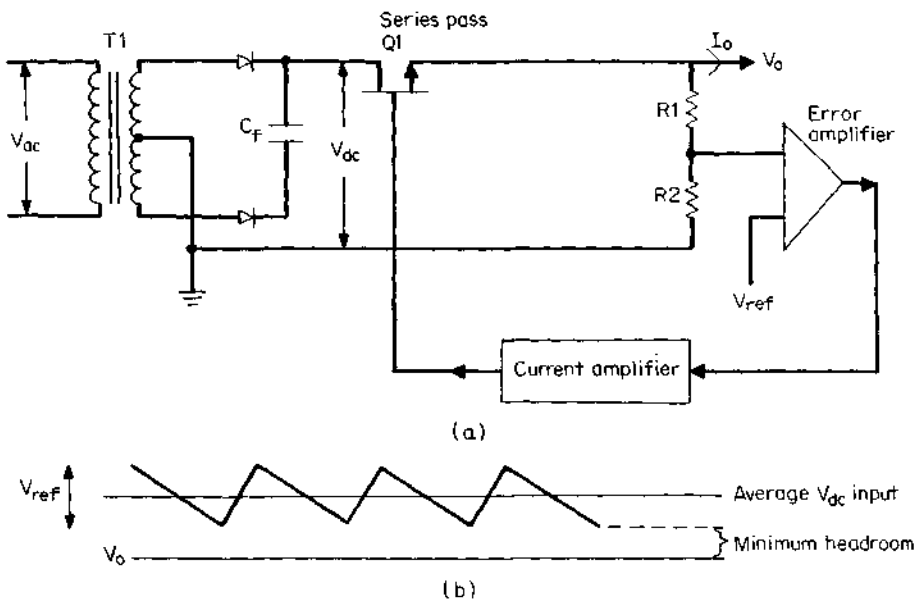


Figure 1.1 (*a*) Linear regulator $Q1$ in series with output load is an electrically variable resistance. The negative-feedback loop alters its resistance to keep V_o constant. (*b*) Headroom in linear regulator. With an NPN series pass transistor, a minimum input-output differential or headroom of 2.5 V between V_o and the bottom of the input ripple triangle must exist at minimum V_{ac} input.

current), the base of the series-pass transistor (assuming an NPN transistor) goes down. This increases the resistance of the series-pass element and hence brings the output voltage back down so that sampled output equals the reference voltage. This negative-feedback loop works in a similar fashion for decreases in output voltage due to a drop in input voltage or an increase in output load current. In that case, the error output voltage moves the series-pass base slightly more positive, decreases its collector-to-emitter resistance, and raises the DC output voltage so that the sampled output again equals the reference.

Essentially, any change in input voltage—due to either AC input line voltage ripple, steady-state changes in the input over its tolerance band, or dynamic changes resulting from rapid load changes—is absorbed across the series-pass element leaving the output voltage constant to an extent determined by the gain in the open-loop feedback amplifier.

The feedback loop is entirely DC-coupled. There are no elements switching on and off and within the loop; all DC voltage levels are predictable and calculable. There are no transformers within the loop and no fast-rise-time voltage or current spikes to cause RFI noise output. With no transistors switching on and off, there are no AC switching losses due to momentary overlap of falling current and rising voltage across the transistors. All power losses are due to DC currents in and DC voltages across the various elements, and these are easily calculable.

1.2.2 Linear regulator drawbacks

This simple, DC-coupled series pass regulator was the basis for a multi-billion-dollar power supply industry until the early 1960s or so. It could produce only lower voltage from a higher one; its output voltage always had one end DC common with the input voltage, yet frequently DC isolation between input and output was required.

The raw DC input voltage (V_{dc} in Fig.1.1*a*) was usually derived from the rectified secondary of a 60-Hz transformer whose weight and volume was often a serious system constraint.

1.2.3 Power dissipation in the series-pass transistor

The major drawback of a linear regulator is mainly the excessive dissipation in the series-pass element. Since all the load current must pass through the pass transistor, its dissipation is $(V_{dc} - V_o)(I_o)$. The minimum differential $(V_{dc} - V_o)$, which is often referred to as "headroom," in most cases is 2.5 V for NPN pass transistors.

Now when the raw DC input comes from the rectified secondary of a 60-Hz transformer, the secondary turns can always be chosen so that

the rectified secondary voltage is V_o + 2.5 V (assume for the moment that the filter capacitor is large enough to yield insignificant ripple) when the input AC is at its low tolerance limit. But when the input AC voltage is at its high tolerance limit, the headroom will be much greater, series-pass element dissipation will be greater, and the power supply efficiency will be significantly lower. This effect of very poor efficiency at high tolerance limit input voltage when the raw DC input has been chosen for a minimum 2.5 V headroom at lowest DC input is much more pronounced at lower output voltages.

This can be seen dramatically in the following three examples of linear regulators: the first with an output of 5 V at 10 A, the second with an output of 15 V at 10 A, and a third at 30 V at 10 A.

Assume a large secondary filter capacitor so that ripple input to the regulator is negligible. The rectified secondary voltage DC voltage tolerances will be identical to the AC input tolerances, which are assumed to be ±15 percent. The transformer secondary voltages will be chosen to yield V_o + 2.5 V when the AC input is at its low tolerance limit of −15 percent. Then the maximum DC input is 1.30 (V_o + 2.5) V when the AC input is at its maximum tolerance limit of +15 percent. Thus

V_o	I_o, A	$V_{dc(min)}$, V	$V_{dc(max)}$, V	Headroom, max, V	$P_{in(max)}$, W	$P_{o(max)}$, W	Dissipation $Q1_{max}$	Efficiency, % $P_o/P_{in(max)}$
5.0	10	7.5	9.75	4.75	97.5	50.0	47.5	51.25
15.0	10	17.5	22.8	7.75	228	150	78.0	65.9
30.0	10	32.5	42.25	12.25	423	300	123	71.0

It thus can be seen that at higher DC output voltages, the efficiency is significantly higher than at lower voltages. When realistic input line ripple voltages are assumed, efficiency for 5-V output for input line tolerances of ±15 percent are in the range 32 to 35 percent.

1.2.4 Linear regulator efficiency versus output voltage

When ripple is taken into account, the minimum headroom of 2.5 V must be guaranteed at the bottom of the ripple triangle at the low tolerance limit of the input AC voltage as shown in Fig. 1.1*b*. Regulator efficiency is then calculated as follows for various assumed input AC tolerances and output voltages.

Assume an input voltage tolerance of ±T percent about its nominal. Transformer secondary turns will be selected so that the voltage at the bottom of the ripple triangle will be 2.5 V above the desired output

voltage (2.5 V minimum headroom) when the AC input is at its low tolerance limit. For a peak-to-peak ripple voltage of V_r volts, the average or DC voltage at the input to the pass transistor is $(V_o + 2.5 + V_r/2)$ volts when the input AC is at its low tolerance limit as can be seen in Fig. 1.1*b*. But when the AC input is at its high tolerance limit, the DC voltage at the input to the series-pass element is

$$V_{dc(max)} = \frac{1 + 0.01T}{1 - 0.01T}(V_o + 2.5 + V_r/2)$$

The maximum achievable efficiency, which occurs at maximum input voltage and hence maximum input power, is

$$\text{Efficiency}_{max} = \frac{P_o}{P_{in(max)}} = \frac{V_o I_o}{V_{dc(max)} I_o} = \frac{V_o}{V_{dc(max)}} \tag{1.1}$$

$$= \frac{1 - 0.01T}{1 + 0.01T}\left(\frac{V_o}{V_o + 2.5 + V_r/2}\right) \tag{1.2}$$

This is plotted in Fig. 1.2 for an assumed peak-to-peak (p/p) ripple voltage of 8 V (it will later be shown that in a 60-Hz full-wave recti-

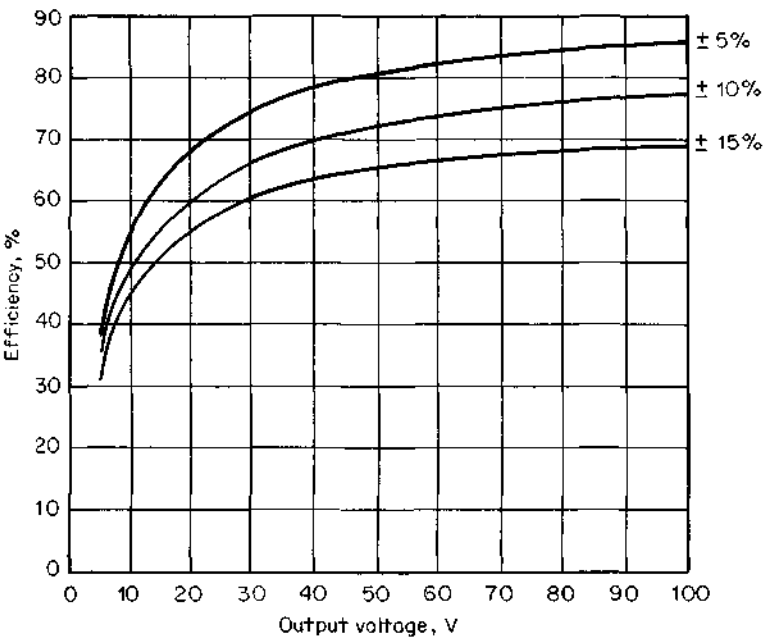


Figure 1.2 Linear regulator efficiency versus output voltage. Minimum efficiency at maximum V_{ac} input assuming 2.5 V headroom at bottom of ripple triangle at minimum V_{ac} input. Assumes 8 V peak-to-peak ripple at top of filter capacitor. (*From Eq. 1.2*)

fier, the p/p ripple voltage is 8 V if the filter capacitor is chosen at the rate of 1000 microfarads per ampere of DC load current).

It can be seen in Fig. 1.2 that even for 10-V outputs, efficiency is under 50 percent for reasonable AC line tolerance of ±10 percent. It is this poor efficiency and the weight and size of the 60-Hz input transformer which led to the revolution of switching power supplies.

Nevertheless, if a preregulated (to ±5 percent or so) raw DC input is available (as frequently is the case in some of the later configurations to be shown), a linear regulator is a reasonable choice if still better regulation is required. Complete integrated-circuit linear regulators are available up to 1.5 A output in single plastic packages at prices in the 50¢ range. They are also available at 3 to 5 A in metal-case integrated-circuit packages at higher prices, but the dissipation across the internal series-pass transistor becomes a serious problem, especially at the 5-A output level and higher.

1.2.5 Linear regulators with PNP series-pass transistors for lesser required headroom

Linear regulators using PNP transistors as the series-pass element can operate with a minimum headroom down to 1 or even 0.5 V and hence achieve better efficiency. The reason for this can be seen in Fig. 1.3.

With an NPN series-pass element as in Fig. 1.3*a*, base current must be inward directed and hence must come from some point at a potential higher than $V_o + V_{be}$ or roughly $(V_o + 1)$ volts. If the base drive comes through a resistor, the input end of that resistor must come from a point higher than $V_o + 1$ to force current through it. The least expensive scheme is to supply the base current resistor from the same point which feeds the series-pass collector, i.e., the raw DC input.

However, now the raw DC input—the bottom of the ripple triangle at low input tolerance—cannot be permitted to come too close to $V_o + 1$ V—nominally the base input voltage. For then, the base resistor R_b would have to be very small to allow sufficient base current at high output current. Then at high input tolerance, when $V_{dc} - V_o$ is much greater, R_b would deliver an excessive current to the base and a significant amount would have to be diverted away into the current amplifier, adding to its dissipation.

It is this effect which dictates that the minimum voltage at the bottom of the ripple triangle not be required to come to less than 2.5 V at minimum input tolerance, i.e., 2.5-V minimum headroom. It renders R_b more of a constant-current resistor and results in relatively constant current through R_b over the range of input voltage tolerances.

With a PNP series-pass transistor (as in Fig. 1.3*b*), however, this problem does not exist. For drive current is outward-directed (sinked)

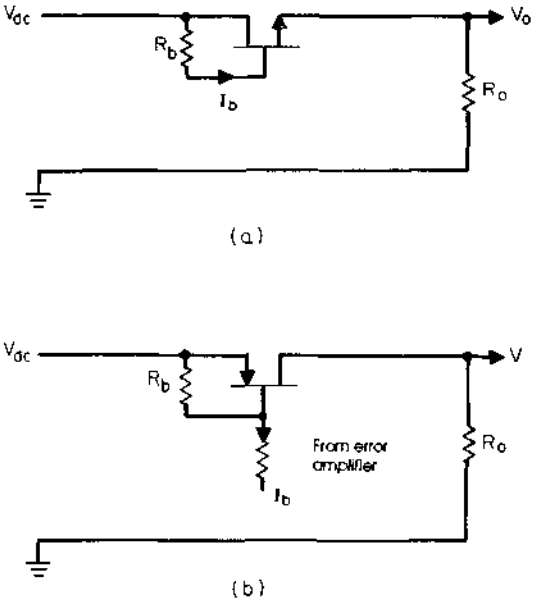


Figure 1.3 *(a)* Linear regulator with an NPN series pass transistor. If base drive is taken from V_{dc} via a resistor R_b, a minimum voltage or headroom must exist across R_b to supply inward-directed base current at the bottom of the input ripple triangle. *(b)* Linear regulator with a PNP series pass transistor. With a PNP series pass transistor, the required base current does not come from V_{dc} but flows outward from the base to the current amplifier; V_{dc} is no longer restricted to a minimum of 2.5 V above V_o. Headroom of 1.0 to 0.5 V is possible.

from the base and is supplied from the current amplifier. Nothing dictates how close V_{dc} may get to V_o except the knee of the I_c versus V_{ce} characteristic of the pass transistor. Minimum headroom with a PNP pass transistor thus can be as low as 1.0 or 0.5 V and low dissipation and higher efficiency results.

Integrated-circuit linear regulators with PNP for transistors are now available but are considerably more expensive. In the fabrication technique of integrated circuits, it is more difficult to make a high-current PNP transistor on the same chip as the lower-current NPN control transistors.

1.3 "Buck" Switching Regulator Topology

1.3.1 Basic operation

The high dissipation across the series-pass transistor in a linear regulator made the use of the linear regulator prohibitive for output cur-

rents over 5 A or so. The high dissipation required large-volume heat sinks. This and the large volume and weight of the input 60-Hz transformer rendered the power supply disproportionately large as integrated circuits made the rest of an electronic system smaller. Linear regulators, with their large input transformer and heat sinks, could achieve output load power densities of only 0.2 to 0.3 W/in^3, and this was not good enough for the smaller systems made possible by integrated circuits.

Alternatives to linear regulators started being widely used in the early 1960s. These so-called switching regulators used a fast-operating transistor switch to switch a DC input voltage through to the output at an adjustable duty cycle. By varying the duty cycle, the average DC voltage delivered to the output could be controlled. Such "average" voltage consisted of rectangular voltage pulses of adjustable width whose average value was the required DC output voltage (Fig. 1.4*b*).

With the use of appropriately chosen *LC* filters, however, the square-wave modulation could be eliminated and ripple-free DC voltages equal to the average of the duty-cycle-modulated raw DC input resulted. By sensing the DC output and controlling the switch duty cycle in a negative-feedback loop, the DC output could be regulated against input line and output load changes.

Such switching regulators are currently achieving 1 to 4 load watts per cubic inch and are capable of generating a multiplicity of output voltages from a single input—all DC isolated from the input. They require no input 50/60-Hz power transformers. Some DC/DC converter designers are claiming load power densities of 40 to 50 W/in^3.

The earliest of these switching regulators, the "buck" regulator, is shown in Fig. 1.4. There, a single-pole single-throw switch in the form of a transistor $Q1$, is in series with the DC input V_{dc}. It is closed for a time T_{on} out of the switching period T. When it is on, the voltage at $V1$ is V_{dc} (assuming for the moment the "on" drop across $Q1$ is zero). When it is open, the voltage at $V1$ falls very rapidly to ground and would have gone dangerously negative had it not been caught and held at ground by the so-called free-wheeling or clamp diode $D1$.

Assume for the moment that the "on" drop of diode $D1$ is zero also. Then the voltage at $V1$ (Fig. 1.4*b*) is rectangular, ranging between V_{dc} and ground with a "high" time of T_{on}. The average or DC value of this voltage is $V_{dc}T_{on}/T$. The L_oC_o filter is added in series between $V1$ and V_o and yields a clean, ripple-free DC voltage at V_o whose magnitude is $V_{dc}T_{on}/T$.

Now V_o is sensed by sampling resistors $R1$, $R2$ and compared to a reference voltage V_{ref} in the error amplifier (EA). The amplified DC

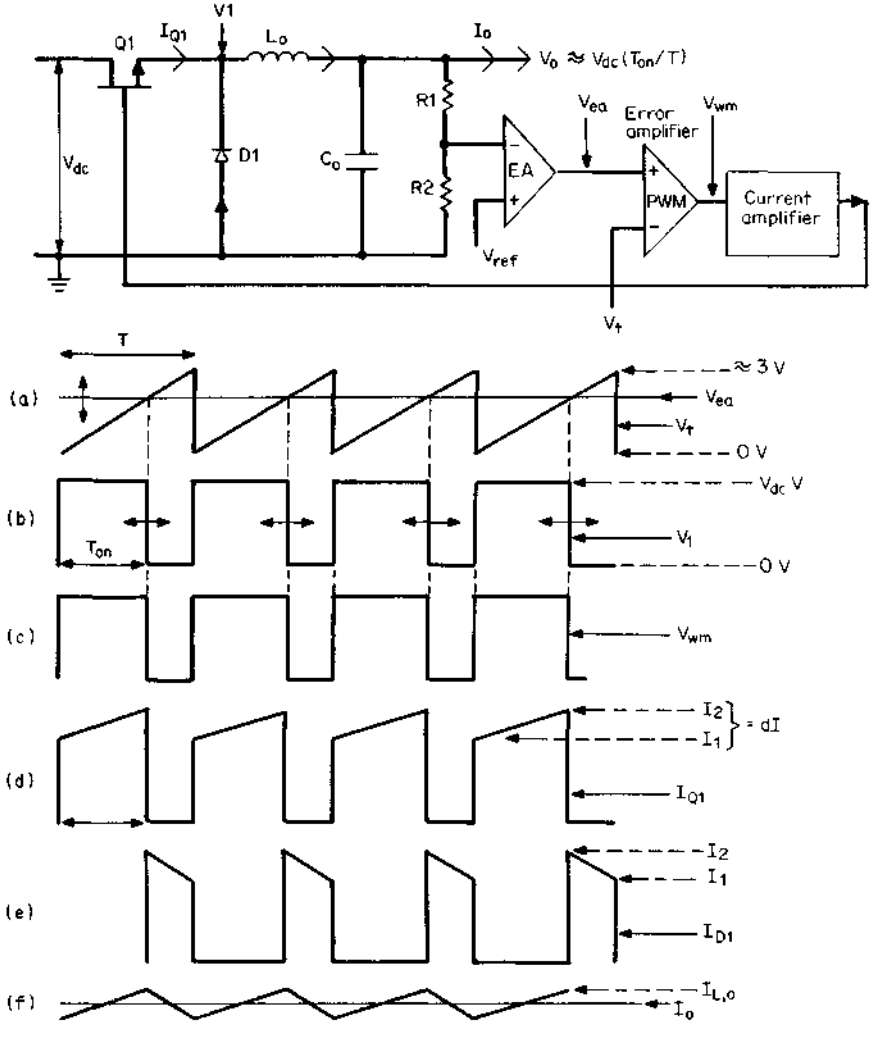


Figure 1.4 Buck switching regulator and critical waveforms.

error voltage V_{ea} is fed to a pulse-width-modulator (PWM) which is essentially a voltage comparator. Another input to the PWM is a sawtooth (Fig. 1.4*a*) of period T and usually 3 V in amplitude (V_t). The PWM voltage comparator generates a rectangular waveform (V_{wm}, Fig. 1.4*c*) which goes high at the start of the triangle and low at the instant the triangle crosses the DC voltage level of the error-amplifier output. The PWM output pulse width T_{on} is thus proportional to the EA amplifier output DC voltage level.

The PWM output pulse is fed to a current amplifier and used to control the "on" time of switch transistor $Q1$ in a negative-feedback loop. The phasing is such that if V_{dc} goes slightly high, the EA DC level goes closer to the bottom of the PWM triangle, the triangle crosses the EA output level earlier in time and the $Q1$ on time decreases, bringing V_o (= $V_{dc}T_{on}/T$) back down. Similarly, if V_{dc} goes low by a certain percentage, the on time increases by the same percentage to maintain V_o constant. The $Q1$ on time is controlled so as to make the sampled output $V_oR_2/(R_1 + R_2)$ always equal to the reference voltage V_{ref}.

1.3.2 Significant current waveforms in buck regulator

The major advantage of the buck regulator is its low internal losses and high efficiency. It is necessary to understand the waveshape and magnitude of the currents throughout the circuit to permit calculating efficiency and some of the subtleties of its operation.

When $Q1$ is on, assuming for the moment zero drop across it, there is a constant voltage ($V_{dc} - V_o$) across L_o. With a constant voltage across an inductor L_o, current in it rises linearly at a rate given by $dI/dT = (V_{dc} - V_o)/L_o$. This accounts for the ramp which sits on a step in Fig. 1.4*d*. Now it is not possible to change the current in an inductor instantaneously. Thus when $Q1$ turns off, the voltage polarity across L_o immediately reverses trying to maintain the same current I_2 which had been flowing just prior to turnoff. This polarity reversal is the so-called inductive kick. If not for diode $D1$, V_1 would have gone very far negative in an attempt to maintain L_o current in the same direction. But diode $D1$ latches in and holds the front end of L_o at one diode drop or about 1 V below ground.

Now the same current I_2 which had been flowing in $Q1$ just prior to turnoff is picked up and flows through diode $D1$ (Fig. 1.4*e*). But now the voltage polarity across L_o has reversed and its magnitude is (V_o + 1). The current in L_o then ramps down linearly at a rate given by $dI/dT = (V_o + 1)/L_o$. This is the downward-going ramp which sits on a step in Fig. 1.4*e*. At the end of the $Q1$ off time, the L_o current has fallen to I_1 and is still flowing through $D1$. Now $Q1$ turns on again and starts supplying current into the cathode of $D1$, displacing its forward current. When $Q1$ current has risen to I_1, all the forward $D1$ current has been displaced and V_1 rises to within about 1 V below V_{dc} backbiasing $D1$.

The current in L_o is then the sum of the $Q1$ current when it is on (Fig. 1.4*d*) and the $D1$ current when $Q1$ is off. This is the current $I(L_o)$ in Fig. 1.4*f*. It is seen to ramp an amount ($I_2 - I_1$) around the DC output current I_o. Thus the value of the current at the center of the

ramp in Fig. 1.4*d* and 1.4*e* is simply the DC output current I_o. As the DC output current I_o is varied, the center of the ramps in either Fig. 1.4*d* or *e* moves, but the slope of the ramps remain parallel to their original value. During the $Q1$ on time the ramp rate in L_o is still $(V_{dc} - V_o)/L_o$, and during the $Q1$ off time it is $(V_o + 1)/L_o$.

It will be seen shortly that when the DC current I_o is reduced to the point where I_1 in Fig. 1.4*d* and 1.4*e* just reaches zero, a drastic change occurs; but this will be discussed shortly.

1.3.3 Buck regulator efficiency neglecting AC switching losses

With the currents in the switch transistor $Q1$ and free-wheeling diode $D1$ (Fig. 1.4*d* and 1.4*e*), losses and efficiency can be calculated. If the preceding currents flowed through $Q1$ and $D1$ at zero voltage drop, total losses would be zero and efficiency would be 100 percent. When $Q1$ is off, it operates at a maximum voltage of V_{dc} but at zero current and hence dissipates no power. When $Q1$ is on, $D1$ operates at a reverse voltage of V_{dc} and draws negligible current and also dissipates no power.

Thus the only losses in the circuit are the DC conduction losses in $Q1$ and $D1$ plus the AC switching losses in $Q1$. There is an AC switching loss in $Q1$ during the turnon transition as a result of the momentary overlap of rising current and falling voltage across it. Also at the turnoff transition, there is an AC switching loss in $Q1$ due to the momentary overlap of falling current and rising voltage across it.

It can be seen in Fig. 1.4*d* and 1.4*e* that the average currents in $Q1$ and $D1$ during their conduction times of T_{on} and T_{off} are the values at the center of the ramps or I_o, the DC output current. These currents flow at a forward voltage of about 1 V over a very large range of currents. Thus conduction losses are

$$P_{dc} = L(Q1) + L(D1) = 1I_o \frac{T_{on}}{T} + 1I_o \frac{T_{off}}{T} = 1I_o$$

Thus if AC switching losses could be neglected, efficiency would be

$$\text{Efficiency} = \frac{P_o}{P_o + \text{losses}} = \frac{V_o I_o}{V_o I_o + 1I_o} = \frac{V_o}{V_o + 1} \tag{1.3}$$

1.3.4 Buck regulator efficiency including AC switching losses

Alternating-current switching or voltage/current overlap loss calculations depend on the scenario assumed in the relative timing and

slopes of the rising current and falling voltage. A best-case scenario—which rarely exists in actual cases—is shown in Fig. 1.5*a*.

Here in the best-case scenario, at turnon, voltage and current start moving simultaneously and reach their endpoints simultaneously. Current goes from 0 to I_o and voltage across $Q1$ goes from a maximum of V_{dc} to zero. The average power during this switching time is $P(T_{on}) = \int_0^{T_{on}} IV\,dt = I_o V_{dc}/6$ and the power averaged over one complete period is $(I_o V_{dc}/6)(T_{on}/T)$.

Assuming the same scenario of simultaneous starting and ending points of the current fall and voltage rise waveforms at the turnoff transition, voltage/current overlap dissipation at this transition is given by $P(T_{off}) = \int_0^{T_{off}} IV\,dt = I_o V_{dc}/6$, and this power averaged over one complete cycle is $(I_o V_{dc}/6)(T_{off}/T)$.

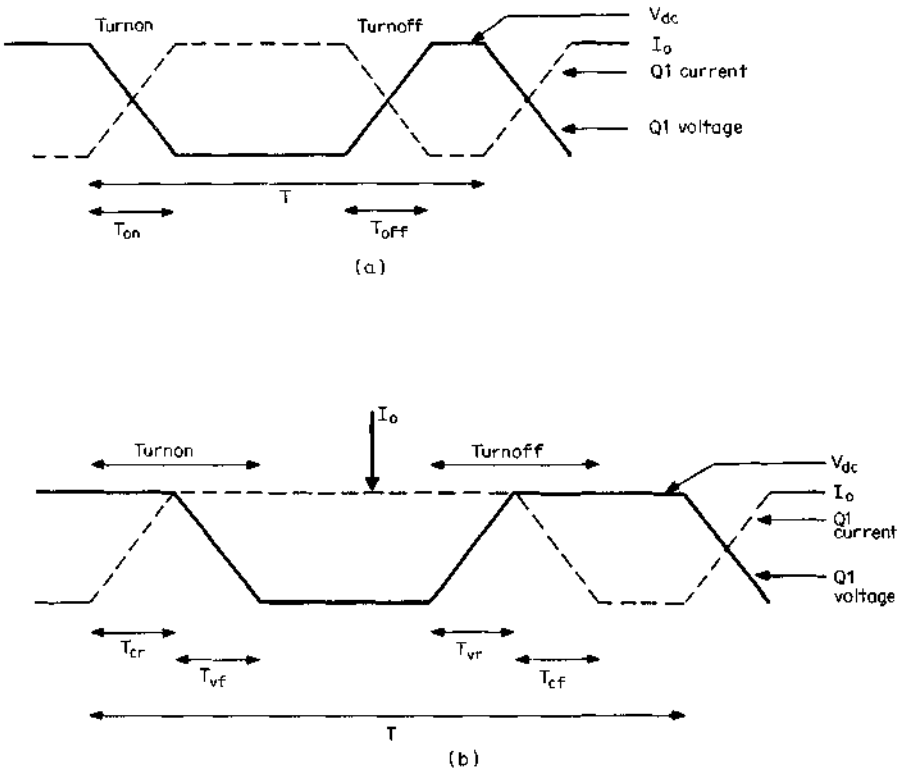


Figure 1.5 (*a*) Transistor switching—best-case scenario. Voltage and current transitions start and end simultaneously. (*b*) Transistor switching—worst-case scenario. At turnon, voltage remains constant at $V_{dc(max)}$ until current reaches its maximum. At turnoff, current remains constant at I_o until $Q1$ voltage reaches its maximum of V_{dc}.

Assuming $T_{on} = T_{off} = T_s$, total switching losses (the sum of turnoff and turnon losses) are $P_{ac} = (V_{dc}I_oT_s)/3T$ and efficiency is

$$\text{Efficiency} = \frac{P_o}{P_o + \text{DC losses} + \text{AC losses}}$$

$$= \frac{V_oI_o}{V_oI_o + 1I_o + V_{dc}I_oT_s/3T} \tag{1.4}$$

$$= \frac{V_o}{V_o + 1 + V_{dc}T_s/3T}$$

It is of interest to calculate the efficiency of a buck regulator and compare it to that of a linear regulator. Assume bucking down from 48 to 5 V, 50-kHz switching frequency ($T = 20$ μs).

If there were no AC switching losses and a switching time T_s of 0.3 μs were assumed, Eq. 1.3 would give

$$\text{Efficiency} = \frac{5}{5 + 1} = 83.3\%$$

If switching losses and the best-case scenario as shown in Fig. 1.5*a* were assumed, Eq. 1.4 would give for $T_s = 0.3$ μs and $T = 20$ μs:

$$\text{Efficiency} = \frac{5}{5 + 1 + 48 \times 0.3/3 \times 20}$$

$$= \frac{5}{5 + 1 + 0.24} = \frac{5}{5 + 1.24}$$

$$= 80.1\%$$

If a worst-case scenario—as shown in Fig. 1.5*b*, which is closer to reality—is assumed, efficiencies are lower. In Fig. 1.5*b*, it is assumed that at turnon, the voltage across the transistor remains at its maximum value (V_{dc}) until the on-turning current reaches its maximum value of I_o. Then voltage starts falling, and to a close approximation, current rise time T_{cr} will equal voltage fall time. Turnon switching losses will then be

$$P(T_{on}) = \frac{V_{dc}I_o}{2}\frac{T_{cr}}{T} + \frac{I_oV_{dc}}{2}\frac{T_{vf}}{T}$$

and for $T_{cr} = T_{vf} = T_s$, $P(T_{on}) = V_{dc}I_o(T_s/T)$.

At turnoff as seen in Fig. 1.5*b*, it is assumed that current hangs on at its maximum value I_o until the voltage has risen to its maximum value of V_{dc} in a time T_{vr}. Then current starts falling and reaches zero in a time T_{cf}. Total turnoff dissipation is then

$$P(T_{off}) = \frac{I_o V_{dc}}{2}\frac{T_{vr}}{T} + \frac{V_{dc} I_o}{2}\frac{T_{cf}}{T}$$

and for $T_{vr} = T_{cf} = T_s$, $P(T_{off}) = V_{dc} I_o (T_s/T)$. Total AC losses (the sum of turnon plus turnoff losses) are then

$$P_{ac} = 2V_{dc} I_o \frac{T_s}{T} \tag{1.5}$$

Total losses (the sum of DC plus AC losses) are

$$P_t = P_{dc} + P_{ac} = 1I_o + 2V_{dc} I_o \frac{T_s}{T} \tag{1.6}$$

and efficiency is

$$\text{Efficiency} = \frac{P_o}{P_o + P_t} = \frac{V_o I_o}{V_o I_o + 1I_o + 2V_{dc} I_o T_s/T}$$

$$= \frac{V_o}{V_o + 1 + 2V_{dc} T_s/T} \tag{1.7}$$

For the same 50-kHz, 48- to 5-V buck regulator as above with T_s = 0.3 μs, efficiency from Eq. 1.7 is

$$\text{Efficiency} = \frac{5}{5 + 1 + 2 \times 48 \times 0.3/20} = \frac{5}{5 + 1 + 1.44}$$

$$= \frac{5}{5 + 2.44}$$

$$= 67.2\%$$

If a linear regulator were used to bring 48 V down to 5 V, its efficiency which is $V_o/V_{dc(max)}$, would be unacceptably lower at 5/48 = 10.4 percent (from Eq. 1.1).

1.3.5 Optimum switching frequency in buck regulator

The output voltage of the buck regulator is $V_o = V_{dc} T_{on}/T$ independent of whatever the period T is. The question arises as to whether there is an optimum period and on what basis the period is selected. At first thought it might seem best to minimize the size of the filter components L_o, C_o by going to as high a frequency as possible.

However, going to higher frequencies does not necessarily minimize the overall size of the regulator when all factors are considered. This can be appreciated by examining the expression for the AC losses in the circuit seen in Eq. 1.5, where the AC losses are inversely propor-

tional to the switching period T. Decreasing T results in increased losses and possibly necessitates the use of a larger-volume heat sink to keep the switching transistor temperature within desired limits.

Further, Eq. 1.5 gives losses in the switching transistor alone and neglects losses in the free-wheeling diode $D1$ due to its finite reverse recovery time. This time is the time required for the diode to cease drawing reverse leakage current measured from the instant it has been subjected to reverse voltage. Free-wheeling diodes as $D1$ should be specified as ultra-fast recovery types which have recovery times as low as 35 to 50 ns. But even such types can dissipate significant power which is inversely proportional to the switching period. Thus increasing switching frequency (decreasing switching period) does decrease size of the filter elements L_oC_o but adds to the total losses and contributes to the requirement for a larger heat sink.

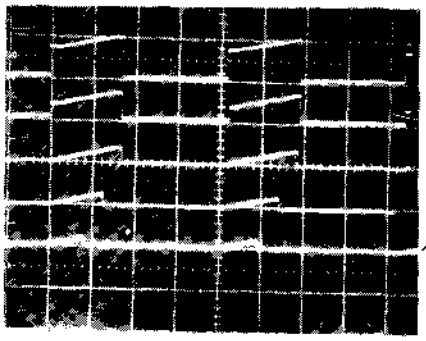
In general, then, between 25 and 50 kHz, overall volume of the buck regulator is lower at a higher frequency. But going to frequencies above 50 kHz offers only marginal advantages because of the increased AC switching losses and possibly larger heat sink.

1.3.6 Design relations—output filter inductor selection

Current waveform of the output inductor is shown in Fig. 1.4*d* and its characteristic "ramp-on-a-step" shape is defined in Sec. 1.3.2, where it was noted that the current at the center of the ramp is equal to the DC output current I_o. As the DC output current decreases, the slope of the ramp remains constant as the voltage across L_o remains constant. But the current at the center of the ramp decreases. At a DC current of $I_o = (I_2 - I_1)/2$ or half the magnitude of the ramp slope dI, the front end of the ramp reaches zero. At this point, the inductor is said to "run dry" or enter into the "discontinuous" operating mode and a drastic change occurs in the current and voltage waveforms.

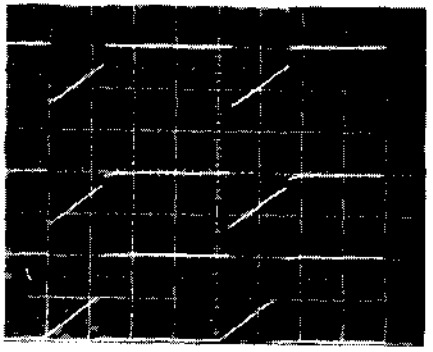
The effect of operation in the discontinuous mode can be seen in Fig. 1.6*a*. These are the inductor or switch transistor on time current waveforms of a 20- to 5-V buck regulator operating at 25 kHz designed for a nominal output current of 5 A and a minimum current of 1 A. The top two waveforms have the characteristic ramp-on-a-step waveshape with the step smaller in amplitude at the lower current and the DC current level at the center of the ramp very closely equal to the DC output current.

In the third waveform in Fig. 1.6*a* at $I_o = 0.95$ A, the step has just disappeared and the front end of the ramp starts at zero current. This is the start of the discontinuous or "run-dry" mode for the inductor. It can be seen in the first three waveforms, the $Q1$ on time is constant



I_{dc} = 4.19 A. Note step at start of ramp
I_{dc} = 2.31 A. Note smaller step at start of ramp
I_{dc} = 0.95 A. Note no step; L_0 has run dry
I_{dc} = 0.49 A. Note much lower on time
I_{dc} = 0.20 A. Note still lower on time

(a)



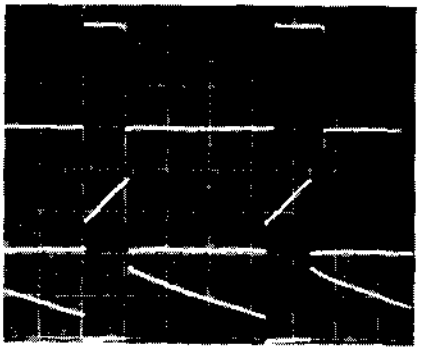
(1) Emitter Q1 = + 20 V, 10 μs/cm

(2) I(Q1) = 2 A, 10 μs/cm; I_{dc} = 4.4 A

(3) I(Q1) = 2 A, 10 μs/cm; I_{dc} = 2.30 A

(4) I(Q1) = 2 A, 10 μs/cm, I_{dc} = 0.95 A

(b)



(1) Emitter Q1 = 10 V, 10 μs/cm

(2) I(Q1) 2 A, 10 μs/cm;

(3) I_{D1} = 2 A, 10 μs/cm

(c)

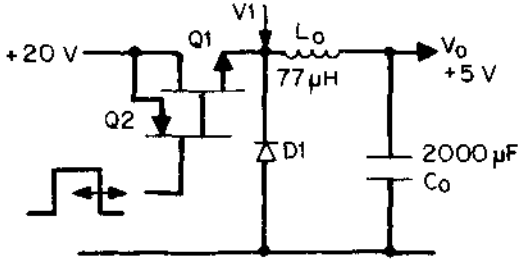


Figure 1.6 A 25-kHz buck regulator, showing the effect of the output inductor L_o running dry or moving from the continuous mode to the discontinuous mode. In panel (*a*), note that on time remains constant only so long as inductor is in the continuous mode.

but decreases drastically after operation has become discontinuous. Yet the DC output voltage remains constant to within 5 mV even after the inductor has gone discontinuous.

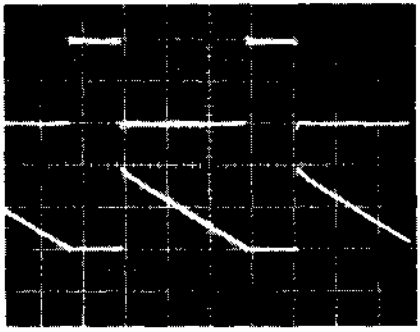
It would seem then that there is no drawback in permitting the inductor to go discontinuous. There actually is no drawback for a buck regulator. Up to the start of the discontinuous mode, the DC output voltage is given by $V_o = V_1 T_{on}/T$. The on time decreases just a very small amount as I_o decreases because the forward drop across $Q1$ decreases slightly with decreasing current. This increases $V1$ slightly, requiring a small decrease in T_{on}.

After going into the discontinuous mode, however, the relation $V_o = V_1 T_{on}/T$ no longer holds, as can be seen in the bottom two waveforms of Fig. 1.6*a*, where the $Q1$ on time has markedly decreased and depends mainly DC output current. Before the start of the discontinuous mode, the circuit coped with DC output current changes by changing the amplitude of the step part of the ramp-on-step waveforms in the $Q1$ and $D1$ waveforms of Figs. 1.4*d* and 1.4*e*. It could do this at relatively constant $Q1$ on time.

Now the average of the sum of the $Q1$ and $D1$ currents [Fig. 1.6*c*(2) and 1.6*c*(3)] is the DC output current. Once the inductor has gone discontinuous and the step part of the latter waveforms has gone to zero, the only way the average of their sum can decrease further is to decrease the $Q1$ on time. The negative-feedback loop automatically sets the $Q1$ on time so that the average of the sum of $Q1$ and $D1$ currents is equal to the DC output current at the correct output voltage as fixed by the error amplifier, the sampling resistors, and the reference voltage of Fig. 1.4.

The loss of the step in the $Q1$ and $D1$ currents between nominal and the start of the discontinuous mode can be seen more clearly in Fig. 1.7*a,b*. When the inductor has gone discontinuous, the $D1$ current has fallen to zero before the next $Q1$ turn on as seen in Fig. 1.7*b*(2). This causes a decaying "ring" at the $Q1$ emitter at a frequency determined by L_o and the capacitance looking into the $D1$ cathode and $Q1$ emitter [Fig. 1.7*b*(1)]. This causes no problem; the $Q1$ on time simply is readjusted by the feedback loop so that the average voltage at the emitter of $Q1$ is equal to the commanded DC output voltage. The ring in Fig. 1.7*b*(1) is no problem; it is easily eliminated by the L_oC_o filter.

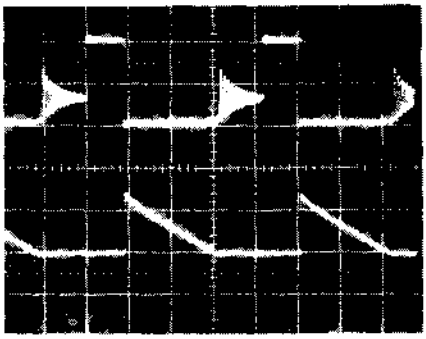
So though operating in the discontinuous mode is no problem in the buck regulator, it is in later topologies with buck-type output L_oC_o filters. Hence in all such output filters, the inductor will be chosen so that it does not go discontinuous until the DC output current falls to its specified minimum value, which in most cases is one-tenth the nominal value or 0.1 (I_{on}). Since the onset of the discontinuous mode occurs at a DC current equal to half the amplitude of the inductor cur-



(a)

(1) Emitter Q1 = 10 V, 10μs/cm

(2) I_{D1} = 1A, 10 μs/cm
Note L0 has just barely gone discontinuous as evidenced by D1 current just falling to zero at instant of next turn-on; I_{dc} = 1.0 A



(b)

(1) Emitter Q1 = 10 V, 10 μs/cm
Note ringing at instant inductor current has fallen to zero; despite ringing, feedback loop corrects main on time and supply still regulates

(2) I_{D1} = 1A, 10 μs/cm
Inductor has gone discontinuous or fallen to zero current before next turn-on; this occurs because DC current is less than half the peak-to-peak ramp amplitude in the photo; I_{dc} = 0.45 A

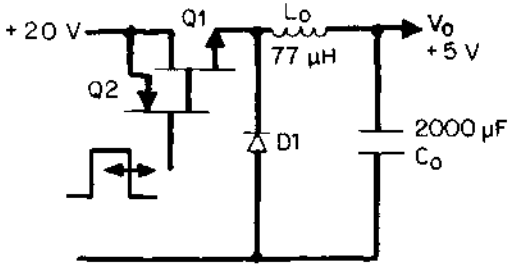


Figure 1.7 A 25-kHz buck regulator, showing $Q1$ emitter waveforms for the case of L_o at the brink of going into the discontinuous mode and also beyond that point.

rent ramp [$(I_2 - I_1)$ of Fig. 1.4*d*], then $I_{o(min)} = 0.1I_{on} = (I_2 - I_1)/2$ or $(I_2 - I_1) = dI = 0.2I_{on}$. But dI, the slope ramp amplitude, is $dI = V_L T_{on}/L = (V_1 - V_o)T_{on}/L$, where V_1 is voltage at the input end of I_o and is very closely equal to V_{dc}. Then

$$L = \frac{(V_{dc} - V_o)T_{on}}{dI} = \frac{(V_{dc} - V_o)T_{on}}{0.2I_{on}}$$

But $T_{on} = V_o T/V_{dc}$; then

$$L = \frac{5(V_{dcn} - V_o)V_o T}{V_{dcn} I_{on}} \tag{1.8}$$

where V_{dcn} and I_{on} are their nominal values.

Thus if L is selected from Eq. 1.8, $(I_2 - I_1) = 0.2I_{on}$ and I_{on} is the inductor current at the center of the ramp at nominal DC output current, the inductor current swings ±10 percent around its value at the center. If the inductor remains constant in value up to $I_2 = 1.1I_{on}$, the ramp will be linear. Thus the inductor must be designed so that it does not saturate or significantly saturate at DC current bias of $1.1I_{on}$.

Such inductors, which remain relatively constant in value under DC bias conditions, can be built with gapped ferrite cores or powdered iron or powdered Permalloy (a magnetic alloy of nickel and iron). Such powdered cores have an effective distributed air gap as they are made from a sludge of powdered magnetic particles embedded in a resin which forms the uniformly distributed air gap. Design of such inductors will be discussed in a later chapter.

The inductor selected from Eq. 1.8 can tolerate higher DC output currents than the specified I_{on} if it is designed not to saturate at these higher currents. The only restriction on maximum current in the buck regulator is the increased DC and AC losses given by Eq. 1.6. The minimum current restriction on the inductor chosen from Eq. 1.8 is that below I_{on} it will go discontinuous and load regulation will slightly degrade.

If the inductor is chosen half the value given by Eq. 1.8, it will go discontinuous at one-fifth rather than one-tenth the nominal DC output current. This will slightly degrade the load regulation commencing at the higher minimum current. But since it is a smaller inductor, the buck regulator will respond more quickly and with a smaller transient output spike to dynamic load changes.

1.3.7 Design relations—output filter capacitor selection

The filter capacitor of Fig. 1.4 is chosen to meet the output ripple specifications. It is not an ideal capacitor, as shown in the figure. It has a resistance R_o and inductance L_o in series with it as shown in Fig. 1.8. These are referred to as the *equivalent series resistance* (ESR) and equivalent series inductance or (ESL). Below about 300 or 500 kHz, L_o can be neglected and output ripple is determined by R_o and C_o.

There are two ripple components, due to R_o and C_o. They are not in phase as that due to R_o is proportional to $I_2 - I_1$, the peak-to-peak inductor ramp current of Fig. 1.4*f*, and that due to C_o is proportional to

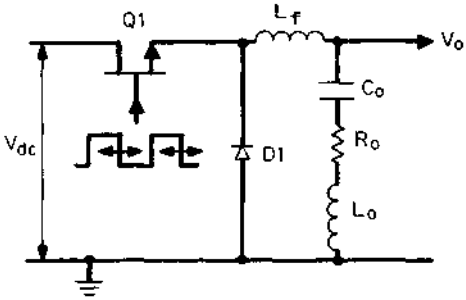


Figure 1.8 Output ripple in the buck regulator. For a given inductor current ripple amplitude, output ripple is determined by the value of the output filter capacitor, C_o, its equivalent series resistance (ESR) R_o, and its equivalent series inductance (ESL) L_o. Below about 500 kHz, L_o can be neglected. In most cases, the ripple component contributed by C_o is small compared to that contributed by R_o. Under 500 kHz, output ripple is closely equal to the AC ripple current in L_o times R_o.

the integral of that current. For a worst-case comparison, however, it can be assumed that they are in phase.

To assess these ripple components and permit a capacitor selection, it is necessary to know the values of R_o, which are seldom given by capacitor manufacturers. But examination of a number of manufacturers' catalogs shows that for the most frequently used types (aluminum electrolytic) over a large range of voltage ratings and capacitance values, R_oC_o tends to be constant. It ranges from 50 to 80 $\times$ 10^{-6} F.

It is instructive to calculate the capacitive and resistive ripple components for a typical buck regulator. Assume a 20- to 5-V regulator with I_{on} = 5 A operating at 25 kHz. Assume $I_{o(min)} = 0.1I_{on}$ = 0.5 A and choose L from Eq. 1.8:

$$L = \frac{5(V_{den} - V_o)V_oT}{V_{den}I_{on}} = \frac{5(20 - 5)5 \times 40 \times 10^{-6}}{20 \times 5} = 150 \quad \mu H$$

and $I2 - I1$, the peak-to-peak ramp amplitude, is $0.2I_{on}$ = 1 A.

Assume to start, a resistive ripple component V_{rr}, of 0.05 V peak-to-peak. Then $0.05 = (I2 - I1)R_o = 1R_o$ or R_o = 0.05 Ω and for R_oC_o = 50 $\times$ 10^{-6}, C_o = 50 $\times$ 10^{-6}/0.05 = 1000 μF.

Now calculate the capacitive ripple voltage V_{cr} from Fig. 1.4*d*. There it is seen that the ripple current is positive from the center of the offtime to the center of the on time or for one-half of a period or 20 μs. The average value of this triangle of current is $(I_2 - I_1)$/

4 = 0.25 A. Then this average current produces across C_o a ripple voltage:

$$V_{cr} = \frac{It}{C_o} = \frac{0.25 \times 20 \times 10 \times^{-6}}{1000 \times 10^{-6}} = 0.005 \quad \text{V}$$

The ripple current below the I_o line in Fig. 1.4*f* yields another 0.005-V ripple for a total peak-to-peak capacitive ripple voltage of 0.01 V. Thus in this particular case, the ripple due to the capacitance is small compared to that due to the ESR resistor R_o and may be ignored. The filter capacitor is then chosen to yield the desired peak-to-peak ripple voltage output by choosing R_o from

$$R_o = \frac{V_{or}}{I_2 - I_1} = \frac{V_{or}}{0.2I_{on}} \tag{1.9}$$

and selecting C_o from the average value of the R_oC_o product, which is closely 65×10^{-6} or

$$C_o = \frac{65 \times 10^{-6}}{R_o} = (65 \times 10^{-6}) \frac{0.2I_{on}}{V_{or}} \tag{1.10}$$

The preceding result—that the output ripple is determined mainly by the ESR of the filter capacitor R_o—has been demonstrated more generally in a paper by K. V. Kantak.[1] There the author shows that if R_oC_o is is larger than half the transistor on time and half the transistor off time—which is the usual case—the output ripple is determined by the ESR resistor itself as demonstrated above for a particular case.

1.3.8 DC-isolated, regulated voltage from a buck regulator

The return end of the regulated output voltage is DC-common with the return end of the raw DC input as can be seen in Fig. 1.4. It is frequently necessary to have a second regulated DC output voltage which is DC-isolated from the main regulated output voltage. This can be done with negligibly fewer additional components as shown in Fig. 1.9 if output variations of about 2 to 3 percent in this second voltage are acceptable.

In Fig. 1.9, a second winding with $N2$ turns is added to the output filter choke. Its output is peak-rectified with diode $D2$ and capacitor $C2$. The polarity of the $N1$, $N2$ windings are shown by the dots. When $Q1$ turns off, the no-dot end of $N1$ goes negative and is caught at one diode drop below ground by free-wheeling diode $D1$. Since V_o is regulated against line and load changes, the voltage across $N1$ is constant as long as the current in free-wheeling diode $D1$ remains constant.

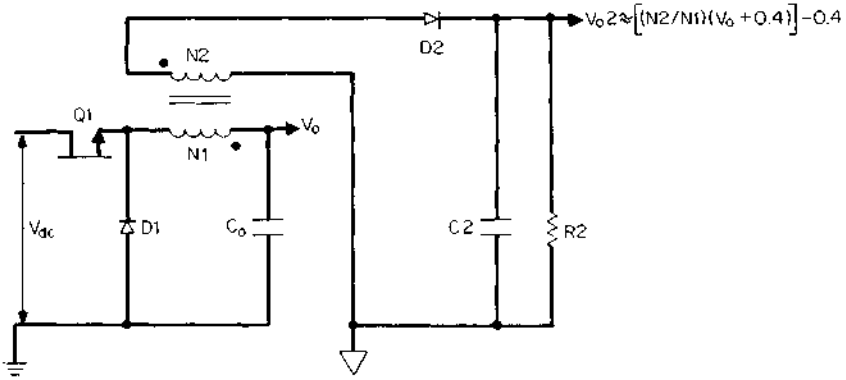


Figure 1.9 Deriving a second output from a buck regulator by using the output choke as a transformer. The second output is DC-isolated from input ground and is regulated to within about 2 to 3 percent as its primary is powered from the regulated V_o and the cathode of $D1$ when $Q1$ turns off.

With $D1$ a low-forward-drop Schottky diode, its forward drop remains constant at about 0.4 V over a large change in DC output current.

Thus when $Q1$ turns off, the voltage across $N2$ is relatively constant at $N_2/N_1(V_o + 0.4)$ volts with its dot end positive. This is peak-rectified by $D2$ and $C2$ to yield $V_o2 = N_2/N_1(V_o + 0.4) - 0.4$ if $D2$ is also a Schottky diode. This output is independent of V_{dc} as $D2$ is reverse-biased when $Q1$ turns on. Capacitor $C2$ should be selected large enough so that its output does not droop for the maximum $Q1$ on time. The 2 to 3 percent variation in V_o2 is due to changes in $D1$ forward drop as I_o changes. Since $N2$ and $N1$ are DC-isolated from one another, this auxiliary regulated output voltage can be referenced to any desired DC voltage level.

1.4 Boost Switching Regulator Topology

1.4.1 Basic operation

The switching technique of Fig. 1.4 can only produce a lower voltage from a higher voltage. Figure 1.10 shows how switching techniques can produce a higher regulated voltage from a lower unregulated voltage. The circuit is called a "boost regulator" or a "ringing choke" converter. It works as follows. An inductor $L1$ is placed in series with V_{dc} and a switching transistor $Q1$. The bottom end of $L1$ feeds the output capacitor C_o and load resistor through rectifying diode $D1$.

The output voltage V_o is higher than the DC input V_{dc}, as can be seen qualitatively thus: When $Q1$ is on for a time t_{on}, $D1$ is reverse-

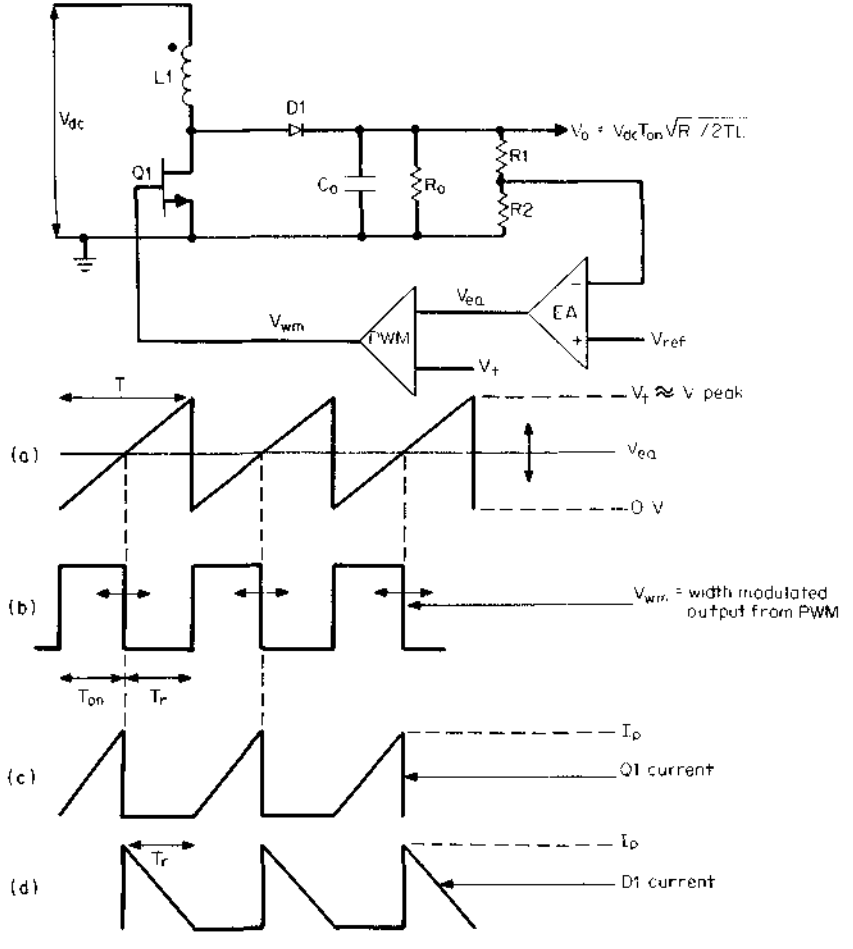


Figure 1.10 Boost regulator and critical waveforms. Energy stored in $L1$ during the $Q1$ on time is delivered to the output via $D1$ at a higher output voltage when $Q1$ turns off and the polarity across $L1$ reverses.

biased and current ramps up linearly in $L1$ to a peak value $I_p = V_{dc}t_{on}/L1$. This represents an amount of stored energy:

$$E = \frac{1}{2L_1(I_p)^2} = 0.5\,L_1 I_p^2 \tag{1.11}$$

where E is in joules, L_1 is in henries, and I_p is in amperes.

During the $Q1$ on time, the output current is supplied entirely from C_o, which is chosen large enough to supply the load current for the time t_{on} with the minimum specified droop.

When $Q1$ turns off, since the current in an inductor cannot change

instantaneously, the current in $L1$ reverses in an attempt to maintain the current constant. Now the no-dot end of $L1$ is positive with respect to the dot end, and since the dot end is at V_{dc}, $L1$ delivers its stored energy to C_o and charges it up via $D1$ to a higher voltage (a boosted-up voltage) than V_{dc}. This energy supplies the load current and replenishes the charge drained away from C_o when it alone was supplying load current. Energy is also supplied to the load from V_{dc} during the $Q1$ off time as will be seen quantitatively below.

The output voltage is regulated by controlling the $Q1$ on time in a negative-feedback loop. If DC load current increases, the on time is automatically increased to deliver the greater required energy to the load. If V_{dc} decreases, and if t_{on} were not changed, the peak current and hence also the energy stored in $L1$ would decrease and the DC output voltage would decrease. But the negative-feedback loop senses a slightly decreased output voltage and increases t_{on} to maintain output voltage constant.

1.4.2 Quantitative relations—boost regulator

If the current through $D1$ has fallen to zero before the next $Q1$ turnon, all the energy stored in $L1$ (Eq. 1.11) during the last $Q1$ on time has been delivered to the output load and the circuit is said to be operating in the "discontinuous" mode.

An amount of energy E delivered to a load in a time T represents power. With E in joules and T in seconds, P is in watts. Thus if all the joules of Eq. 1.11 are delivered to the load once per period T, the power to the load from $L1$ alone (assuming for the moment 100 percent efficiency) is

$$P_L = \frac{\frac{1}{2}L(I_p)^2}{T} \qquad W \tag{1.12}$$

During the time the current in $L1$ is ramping down to zero (T_r in Fig. 1.10*d*), however, that same current is flowing through V_{dc} and is also delivering to the load an amount of power P_{dc} equal to the average current during T_r multiplied by its duty cycle and V_{dc}:

$$P_{dc} = V_{dc}\frac{I_p}{2}\frac{T_r}{T} \tag{1.13}$$

The total power delivered to the load is then

$$P_t = P_L + P_{dc} = \frac{\frac{1}{2}L_1(I_p)^2}{T} + V_{dc}\frac{I_p}{2}\frac{T_r}{T} \tag{1.14}$$

But $I_p = V_{dc}T_{on}/L_1$. Then

$$P_t = \frac{(\frac{1}{2}L_1)(V_{dc}T_{on}/L_1)^2}{T} + V_{dc}\frac{V_{dc}T_{on}}{2L_1}\frac{T_r}{T}$$

$$= \frac{V_{dc}{}^2T_{on}}{2TL_1}(T_{on} + T_r) \tag{1.15}$$

Now to make sure that the current in $L1$ has ramped down to zero before the next $Q1$ turnon, set $(T_{on} + T_r) = kT$, where k is fraction less than 1. Then $P_t = (V_{dc}{}^2T_{on}/2TL_1)(kT)$.

But for an output voltage V_o and output load resistor R_o

$$P_t = \frac{V_{dc}{}^2T_{on}}{2TL_1}(kT) = \frac{V_o^2}{R_o}$$

or

$$V_o = V_{dc}\sqrt{\frac{kR_oT_{on}}{2L1}} \tag{1.16}$$

Thus the negative-feedback loop keeps the output constant against input voltage changes and output load R_o changes in accordance with Eq. 1.16. As V_{dc} and R_o go down or up, the loop will increase or decrease T_{on} so as to keep V_o constant.

1.4.3 Discontinuous and continuous modes in boost regulator

It was mentioned in the previous section that if the $D1$ current (Fig. 1.10*d*) fell to zero before the next turnon, the circuit is said to operate in the discontinuous mode. If the current has not fallen to zero at the end of the off time, the $Q1$ current at the next turnon will have a front-end step as the current in an inductor cannot change instantaneously. Currents in $Q1$ and $D1$ will have a characteristic ramp-on-a-step waveshape as in Fig. 1.11. The circuit is now said to be operating in the continuous mode.

If the feedback loop had been successfully stabilized for operation in the discontinuous mode, as R_o or V_{dc} decreases, the feedback loop increases T_{on} to maintain the output voltage constant. As R_o or V_{dc} continues to decrease, a point is reached such that T_{on} is so large that the decaying current through $D1$ has not yet fallen to zero before the next turnon (Figs. 1.10, 1.11) and the circuit is in the continuous mode.

Now the error-amplifier circuit, which had successfully stabilized the loop while it was operating in the discontinuous mode, is not able to keep the loop stable in the continuous mode and begins to oscillate. In the jargon of feedback-loop analysis, the continuous-mode boost

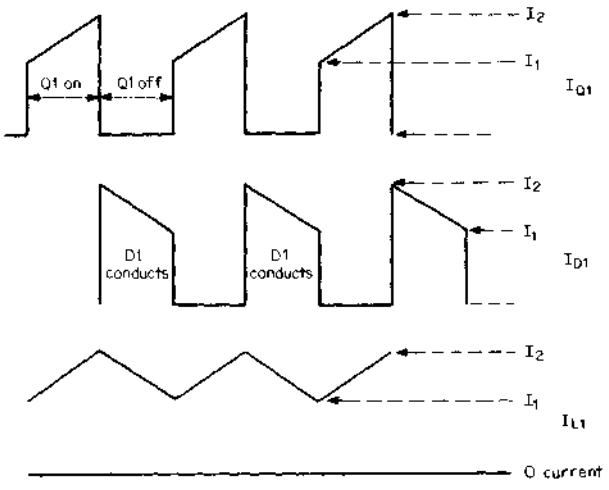


Figure 1.11 Current in $Q1$ and $D1$ for boost regulator in continuous mode. Inductor $L1$ has not had enough time to release all its current to the load before the next $Q1$ turnon.

regulator has a *right-half-plane zero.* The only way to stabilize a loop with a right-half-plane zero is to drastically reduce the error-amplifier bandwidth. Right-half-plane zeros will be discussed in more detail in the chapter on feedback stabilization.

1.4.4 Discontinuous-mode boost regulator design relations

In Fig. 1.10*d*, the decaying $D1$ current is seen to just come down to zero at the start of the next turnon. That was the very end of discontinuous-mode operation. It is seen from Eq. (1.16) to occur at certain combinations of V_{dc}, T_{on}, R_o, and $L1$ for a given operating period T.

If certain combinations of these result in $D1$ current falling to zero just immediately prior to the next $Q1$ turnon as shown in Fig. 1.10*d*, any further decrease in V_{dc} or R_o (increase in load current) will force the circuit into the continuous mode and oscillation will occur if the error amplifier has not been designed for this mode. This can be seen from Eq. (1.16). If the circuit is just at the end of discontinuous-mode operation, any further decrease in V_{dc} or R_o will cause the error amplifier to increase T_{on} in order to keep V_o constant. Then $D1$ starts ramping down later in time (and from a higher current as T_{on} has increased). Now at the start of the next $Q1$ on time, the $D1$ current has not yet fallen to zero and the circuit has entered the continuous mode.

To avoid this problem (see Eq. 1.16) T_{on} must be selected so that

when it is a maximum (which is when V_{dc} and R_o are at their minimum specified values), and then when the current in $D1$ has fallen back to zero, there is an arbitrarily selected dead-time margin (T_{dt}) before $Q1$ turns on again. In the time the current in $D1$ returns to zero, the $L1$ core must be restored or "reset" to its starting place on its hysteresis loop.

Now to understand how the core may fail to be reset to its starting place and to comprehend the basic concept of the core moving up (being set) or moving down (being reset) on its hysteresis loop, the most fundamental law in magnetics—Faraday's law—must be appreciated. Faraday's law states

$$E = NA_e\,(dB/dt) \times 10^{-8} \tag{1.17}$$

where E = voltage across a core (inductor or transformer winding) volts
N = number of turns on the winding
A_e = iron area of core (cm^2)
dB = core flux change (gauss) (which has a plus or minus direction depending on polarity of the voltage across the winding)
dt = time for that flux change (seconds)

Now if a voltage E is applied across a winding of N turns for a time dt, it produces from Faraday's law a flux change of

$$dB = \frac{E\,dt \times 10^{+8}}{NA_e} \quad \text{(gauss)} \tag{1.18}$$

Thus, as shown in Fig. 1.12, if a core is subjected to a positive flux change of $dB = B2 - B1$, when, say, the dot end of the winding had been positive, the core must be restored exactly to B_1 by a reversed voltage polarity before original polarity voltage may be applied. If not,

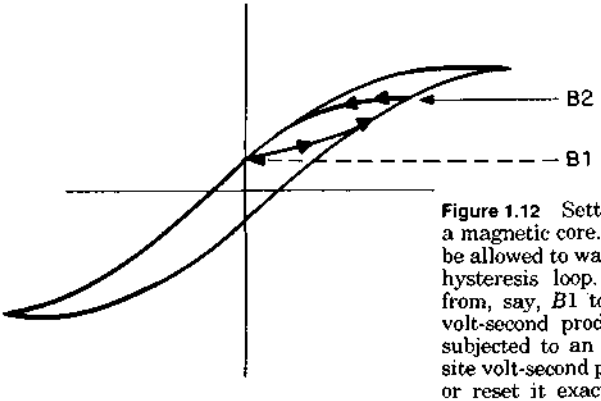


Figure 1.12 Setting and resetting a magnetic core. A core must not be allowed to walk up or down its hysteresis loop. If it is driven from, say, $B1$ to $B2$ by a given volt-second product, it must be subjected to an equal and opposite volt-second product to restore or reset it exactly to $B1$ before driving it upward again.

if a core moves from, say, $B1$ to $B2$ and with a reverse polarity voltage is restored to only a few percent above $B1$, then after many such cycles, it will drift up its hysteresis loop and saturate the core. Since a saturated core cannot sustain voltage, the voltage at the transistor collector will suddenly move up to the supply voltage and the transistor will be destroyed.

From Eq. (1.18), the flux change dB caused by a voltage E applied for a time dt is proportional to the volt-second product $E\,dt$. Thus if a core is moved up or "set" by a given volt-second product, it must be "reset" by an opposite polarity and exactly equal the volt-second product.

Now in Fig. 1.10, during the transistor on time T_{on}, the dot end of $L1$ will be positive and the core will be set or moved—say, upward—on the hysteresis loop by the volt-second product $V_{dc}T_{on}$ (assuming negligible on drop across $Q1$). When $Q1$ turns off, polarity across $L1$ reverses, the no-dot end of $L1$ is now positive, and the core is reset in a time T_r to its exact starting point on the hysteresis loop. During the reset time, the voltage across $L1$ (assuming here also negligible drop across $D1$) is $(V_o - V_{dc})$. To ensure that the core is always reset to its initial starting point and does not drift either way on the hysteresis loop, the on or set volt-second product must be equal to the off or reset volt-second product.

Now to ensure that the circuit remains in the discontinuous mode, a dead-time T_{dt} of 20 percent of a full period will be established such that the sum of the maximum on time, the core reset time, and the dead time will equal a full period as in Fig. 1.13. This ensures that the stored current in $L1$ has fallen to zero 20 percent of a full period before the next $Q1$ turnon.

Now hereafter, a line appearing below a term will indicate the minimum permitted or specified or required value of that term, and a line appearing over a term will indicate the maximum value of that term. Then $\overline{T_{on}} + T_r + T_{dt} = T$, $\overline{T_{on}} + T_r + 0.2T = T$, or

$$\overline{T_{on}} + T_r = 0.8T \tag{1.19}$$

From Eq. (1.16), maximum on time $\overline{T_{on}}$ occurs at minimum $\underline{V_{dc}}$ and minimum $\underline{R_o}$. Then for the on or set volt-second product to equal the off or reset volt-second product at minimum $\underline{V_{dc}}$ and minimum $\underline{R_o}$:

$$\underline{V_{dc}}\overline{T_{on}} = (V_o - \underline{V_{dc}})T_r \tag{1.20}$$

Now in Eqs. (1.19) and (1.20) there are only two unknowns, $\overline{T_{on}}$ and T_r, and thus both are determined; $\overline{T_{on}}$ is then

$$\overline{T_{on}} = \frac{0.8T(V_o - \underline{V_{dc}})}{V_o} \tag{1.21}$$

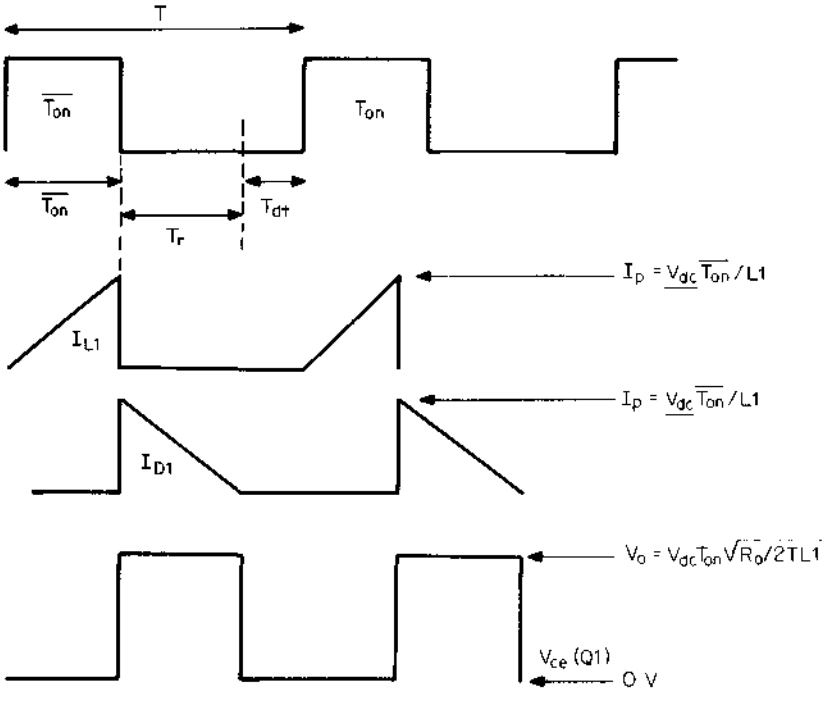


Figure 1.13 Ensuring discontinuous-mode operation in a boost regulator. For discontinuous-mode operation, the current in $D1$ (Fig. 1.10) must have decayed to zero before the next turnon. To ensure this, with a minimum input voltage, a maximum on time, and a maximum output current at the desired output voltage, the inductor $L1$ is chosen such that $T_{on(max)} + T_r = 0.8T$, leaving a dead time T_{dt} of $0.2T$.

Now in Eq. 1.16, with V_{dc} and R_o (maximum load current) specified, $\overline{T_{on}}$ calculated from Eq. 1.21 and $\overline{k}$ [$= (\overline{T_{on}} + T_r)/T$)] = 0.8 as set from Eq. 1.19, inductor $L1$ is fixed and the circuit is guaranteed not to enter the continuous mode. If output load current is inadvertently increased beyond its specified maximum value (R_o decreased below its specified minimum) or V_{dc} is decreased below its specified minimum, the feedback loop will attempt to increase T_{on} in order to keep V_o constant. This will eat into the dead time T_{dt} and move the circuit closer to the continuous mode. To avoid this, a clamp to limit maximum on time or maximum peak current must be introduced.

With $L1$ determined as above from Eq. 1.16, V_{dc} specified and $\overline{T_{on}}$ calculated from Eq. 1.21, the peak current in $Q1$ can be calculated from Eq. 1.14 and a transistor selected to have adequate gain at I_p.

1.4.5 Boost regulator applications and flyback comparison

The boost regulator is not as widely used as the buck simply because there are fewer instances where a higher from a lower voltage is re-

quired. In those instances where the higher voltage is required, there are alternative and preferable ways to obtain it.

Nevertheless, the boost regulator has been treated in rather great detail because by replacing the inductor $L1$ by a transformer, a very similar, valuable, and widely used topology—the "flyback" is realized. As for the boost, the flyback stores energy in its magnetics in one part of the switching period and, in a subsequent part, transfers it to the output load. Because of the transformer, output return is DC-isolated from input return. Also by using multiple secondaries, a multi-output-voltage power supply is possible with outputs at a higher or lower voltage than the input.

The problems of discontinuous or continuous operation and the design relationships and procedures for the flyback are similar to those of the boost and will be discussed in detail in a later chapter.

The boost regulator is most frequently used at low power levels (<10 W). Its greatest usage at such low power levels is on printed-circuit boards where it is desired to step up a 5-V computer logic level supply to 12 or 15 V for operational amplifiers. It is frequently encountered at higher power levels in battery input power supplies. As a battery discharges, its output voltage drops significantly and suddenly. Many systems whose prime power is a nominal 12- or 28-V battery will present problems when the battery voltage falls to about 9 or 22 V, respectively. Boost regulators are frequently used in such applications to boost the voltages back up to the 12- and 28-V level. Power level in such applications can be in the 50- to 200-W level.

1.5 Polarity Inverting Switching Regulator Topology

1.5.1 Basic operation

Polarity inverting switching regulator topology is shown in Fig. 1.14. It uses the same basic principle of energy storage in an inductor in one part of the operating period and then transferring it to the output load in the latter part of the period.

Comparing Figs. 1.14 and 1.10, it is seen that the transistor and inductor have changed places. In the polarity inverter, the transistor is above the inductor rather than below it as in the boost. Also the orientation of the rectifying diode has been reversed.

In Fig. 1.14, when $Q1$ turns on, diode $D1$ is reverse-biased as its cathode is at V_{dc} (assuming to a close approximation that the on drop across $Q1$ is zero) and assuming for the moment that C_o has charged down to some negative voltage. With a fixed voltage V_{dc}

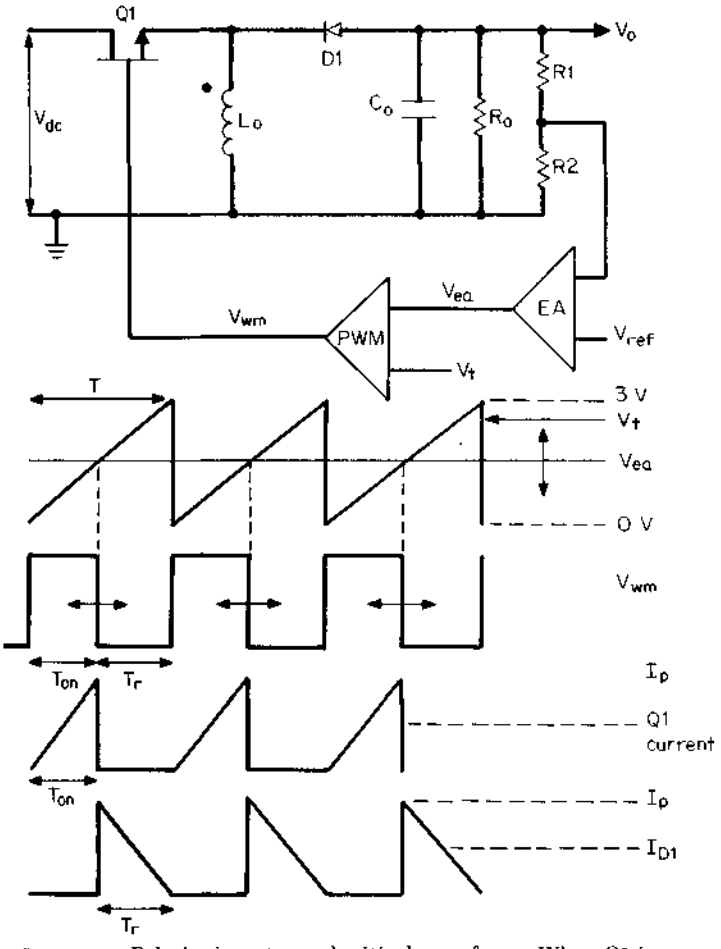


Figure 1.14 Polarity inverter and critical waveforms. When $Q1$ is on, current ramps up linearly in L_o and stores current I_p in L_o. When $Q1$ turns off, polarity across L_o reverses. Current I_p now decays linearly through C_o and D1, charging the top end of C_o negatively. If all the current I_p stored in L_o during the $Q1$ on time has decayed to zero before the next $Q1$ turnon, power delivered to the load is $P_o = \frac{1}{2}L_o(I_p)^2/T$.

across L_o, current in it ramps up linearly at a rate $di/dt = V_{dc}/L_o$. After an on time T_{on}, current in L_o has reached $I_p = V_{dc}t_{on}/L_o$ and represents a stored energy (in joules) of $E = \frac{1}{2}L_oI_p^2$. When $Q1$ turns off, polarity across L_o reverses in an attempt to maintain its current unchanged. Thus at the instant of turnoff, the same I_p which flowed through $Q1$ before it turned off now flows through C_o and $D1$. This current flowing upward through C_o charges the top end of it to a negative voltage as assumed.

After a number of cycles, the error amplifier sets the $Q1$ on time T_{on} so that the sampled output voltage $[V_o R2/(R1 + R2)]$ is equal to the reference voltage V_{ref}. Further, if all the energy stored in L_o is delivered to the load before the next $Q1$ turnon (by I_{D1} having fallen to zero before turnon), the circuit operates in the discontinuous mode and the power delivered to the load is

$$P_t = \frac{\frac{1}{2}L_o I_p^2}{T} \tag{1.22}$$

It should be noted that unlike the case of the boost regulator, when $Q1$ turns off, the stored current does not flow through the supply source (see Eq. 1.13). Hence the only power to the load is that given by Eq. 1.22. Thus assuming 100 percent efficiency, and a minimum output resistance R_o, output power is

$$P_o = \frac{V_o^2}{R_o} = \frac{\frac{1}{2}L_o I_p^2}{T} \tag{1.23}$$

and for $I_p = V_{dc}T_{on}/L_o$

$$V_o = V_{dc}T_{on}\sqrt{\frac{R_o}{2TL_o}} \tag{1.24}$$

1.5.2 Design relations in polarity inverter

As in the boost, it is desirable to keep the circuit operating in the discontinuous mode by ensuring that the current stored in L_o during the $Q1$ maximum on time has decayed to zero at the end of T_r with a dead time T_{dt} margin of $0.2T$ before the next $Q1$ turnon. Thus $\overline{T_{on}} + T_r + T_{dt} = T$, and for $T_{dt} = 0.2T$, we obtain

$$\overline{T_{on}} + T_r = 0.8T \tag{1.25}$$

In addition, as in the boost regulator, the on volt-second product must equal the reset volt-second product to prevent the core from drifting up or down its hysteresis loop. Since maximum T_{on} occurs for minimum V_{dc} and minimum R_o (as can be seen from Eq. 1.24) it follows that

$$V_{dc}\overline{T_{on}} = V_o T_r \tag{1.26}$$

Thus in both Eqs. 1.25 and 1.26 there are two unknowns: $\overline{T_{on}}$ and T_r. This fixes $\overline{T_{on}}$ at

$$\overline{T_{on}} = \frac{0.8V_oT}{\underline{V_{dc}} + V_o} \tag{1.27}$$

Now, with $\overline{T_{on}}$ calculated from Eq. 1.27, $\underline{V_{dc}}$, $\underline{R_o}$, V_o, and T specified, Eq. 1.24 fixes L_o. This fixes $I_p = \underline{V_{dc}}\,\overline{T_{on}}/\overline{L_o}$ and transistor $Q1$ can be selected to have adequate gain at $\overline{I_p}$.

Reference

1. K. V. Kantak, "Output Voltage Ripple in Switching Power Converters," *Power Electronics Conference Proceedings*, Boxborough, Mass., pp. 35–44, April 1987.

Chapter

2

Push-Pull and Forward Converter Topologies

2.1 Introduction

The three earliest switching regulator topologies discussed in the previous chapter had the significant drawbacks that the output returns were DC-common with the input returns and that multiple outputs were not possible (except for the restricted case discussed in Sec. 1.3.8).

In this chapter, the most widely used switching regulator topologies are discussed. These topologies—the push-pull, single-ended forward converter, and two modifications of the latter, the double-ended and interleaved forward converter—have sufficient similarity to consider them as a family. All these topologies deliver their power to the loads via a transformer; hence output voltage returns are DC-isolated from input returns, and with multiple transformer secondaries, multiple DC output voltages are possible.

2.2 Push-Pull Topology

2.2.1 Basic operation—master/slave outputs

Push-pull topology is shown in Fig. 2.1. It consists of a transformer $T1$ with multiple secondaries. Each secondary supplies a pair of 180° out-of-phase square-wave power pulses whose amplitude is fixed by the number of secondary turns. Pulse widths at all secondaries are, of course, identical and are determined by the control circuit in the negative-feedback loop around a main secondary V_m. That control circuit is basically much like that for the buck and boost regulators of Figs. 1.4 and 1.10 except that two equal and adjustable-width, 180°

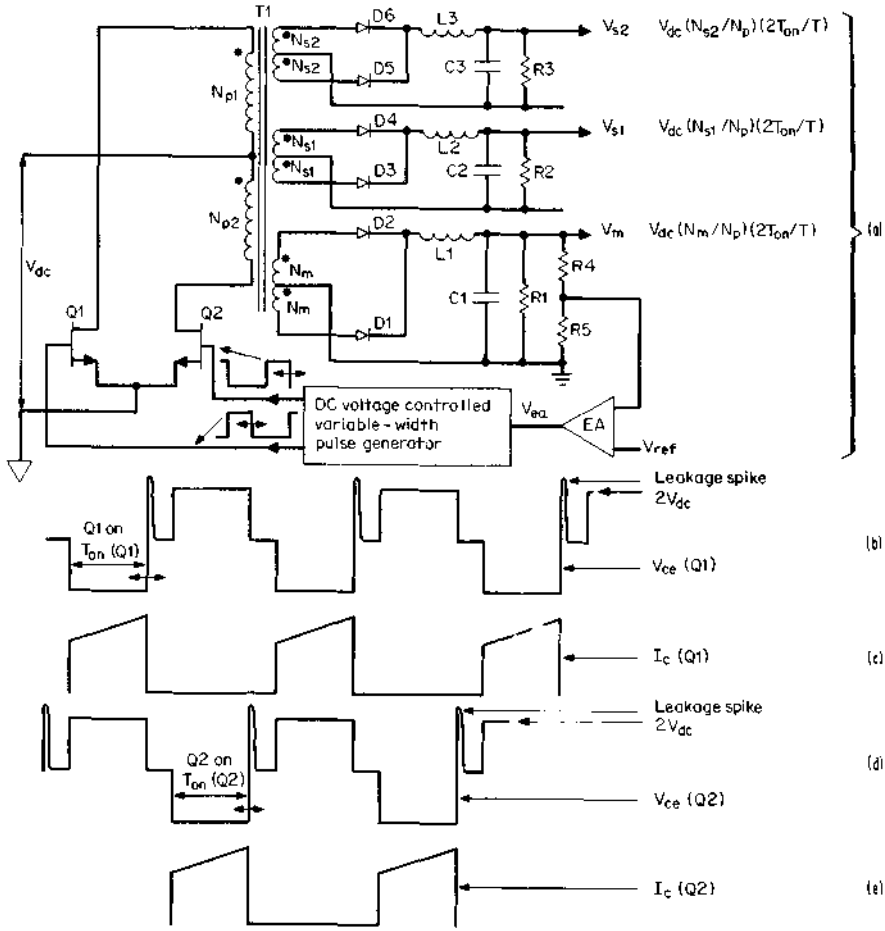


Figure 2.1 Push-pull width-modulated converter. Transistors $Q1$ and $Q2$ are driven by 180 out-of-phase width-modulated signals. Outputs are one master (V_{sm}) and two slaves (V_{s1} and V_{s2}). Feedback loop is closed around V_{sm} and T_{on} is controlled to regulate it against line and load changes. Slaves are regulated against line but only partially against load.

out-of-phase pulses are supplied to drive the bases of $Q1$, $Q2$. Secondaries N_{s1}, N_{s2} are referred to as "slaves."

Transistor base drives at turnon are sufficient to bring the bottom end of each half primary down to the $V_{ce(sat)}$ level of about 1 V for the transistors over the full specified current range. Hence when either transistor turns on, it applies to its half primary a square voltage pulse of magnitude V_{dc} − 1.

Then for output rectifier forward drops of V_d, outputs at the rectifier cathodes will be flat-topped square waves of amplitude $(V_{dc} - 1)(N_s/N_p) - V_d$ and of duration T_{on}. Here, V_d, the rectifier diode forward drop, will be taken as 1 V for a conventional fast-recovery diode and 0.5 V for a Schottky diode (usually used for 5-V high-current outputs as V_m). These output pulses at the rectifier cathodes have a duty cycle of $2T_{on}/T$ as there is one such pulse of duration T_{on} per half period $T/2$.

Thus the waveform at the input to the LC filters in Fig. 2.1 is very much like that at the input to the buck regulating LC filter of Fig. 1.4. It has a unique flat-topped amplitude and adjustable width. The LC filters of Fig. 2.1 serve the same purpose as that in Fig. 1.4. They provide a DC output which is the average of the square wave at their inputs and filter out the square-wave ripple. The analysis of the inductor and capacitor functions proceeds exactly as for the buck regulator, and the relations for calculating their magnitudes are exactly as for the buck.

Thus the DC or average voltage at the V_m output in Fig. 2.2 (assuming $D1$, $D2$ are 0.5-V forward-drop Schottky diodes) is

$$V_m = \left[(V_{dc} - 1)\left(\frac{N_m}{N_p}\right) - 0.5\right]\frac{2T_{on}}{T} \tag{2.1}$$

Waveforms at the V_m output rectifiers are shown in Fig. 2.2. Now, if a negative-feedback loop is closed around V_m as in Fig. 2.1 to control

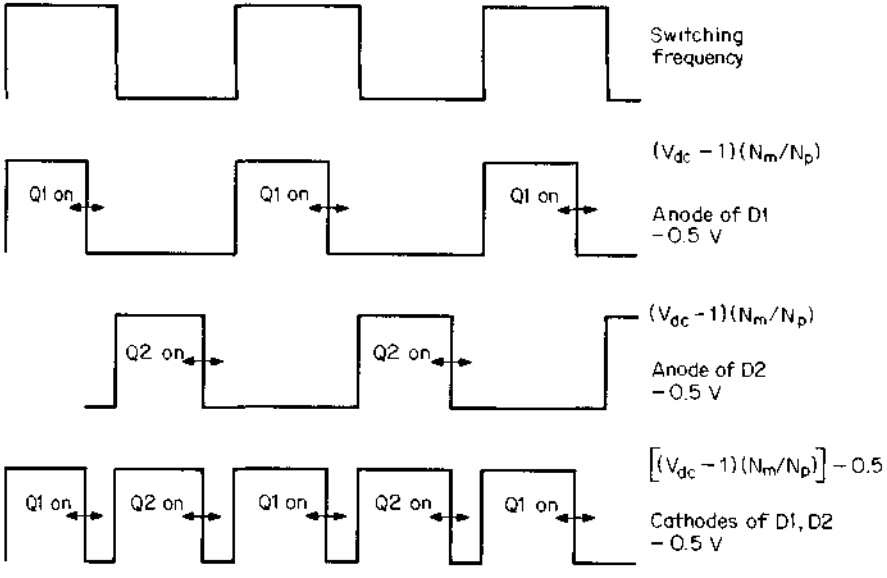


Figure 2.2 Voltages at main (N_m) secondary winding. Output LC filter is an averaging filter. It yields DC out voltage $= V_m = [(V_{dc} - 1)(N_m/N_p) - 0.5](2T_{on}/T)$. As V_{dc} varies, the negative-feedback loop corrects T_{on} in such a direction as to keep V_m constant.

T_{on}, V_m will be regulated against DC input voltage and load current output changes. Although load current does not appear in Eq. 2.1, whatever causes V_m to change, that change will be sensed by the error amplifier and T_{on} will be altered in such a direction as to correct that change. So long as $L1$ (Fig. 2.1) does not go discontinuous, changes in T_{on} will be small and the absolute value of T_{on} will be given by Eq. 2.1 for any turns ratio N_m/N_p, V_{dc}, and period T.

Now the voltage at the cathodes of the slave rectifying diodes is fixed by the number of slave secondary turns, and the duration of the square waves is the same T_{on} as was fixed by the master (V_m) feedback loop. Thus the slave outputs (assuming non-Schottky rectifier diodes) are

$$V_{s1} = \left[(V_{dc} - 1)\frac{N_{s1}}{N_p} - 1\right]\frac{2T_{on}}{T} \tag{2.2}$$

$$V_{s2} = \left[(V_{dc} - 1)\frac{N_{s2}}{N_p} - 1\right]\frac{2T_{on}}{T} \tag{2.3}$$

2.2.2 Slave line-load regulation

It can now be seen from Eqs. 2.1, 2.2, and 2.3 that the slaves are regulated against V_{dc} input changes by the negative-feedback loop which keeps V_m constant. As V_{dc} changes, the loop changes T_{on} so that $(V_{dc} - 1)T_{on}$ is kept constant to maintain V_m constant in accordance with Eq. 2.1. But the same product $(V_{dc} - 1)T_{on}$ also appears in Eqs. 2.2 and 2.3, and thus V_{s1}, V_{s2} are also kept constant as V_{dc} changes.

But if load current in the master (V_m) changes, forward drop across its rectifying diodes will change slightly, decreasing the peak voltage at their cathodes slightly. The forward drop across a Schottky rectifying diode was assumed as 0.5 V as a simplifying approximation. That drop can change by ±0.1 V with large load changes. Thus the negative-feedback loop to correct for V_m load changes, will alter T_{on} to keep V_m constant as dictated by Eq. 2.1.

Then in Eqs. 2.2 and 2.3, T_{on} will change without corresponding changes in V_{dc} and changes in V_{s1}, V_{s2} will result. Such changes in the slave output voltages due to changes in the master output current is referred to as *cross regulation*. Slave output voltages will also change as a result of changes in their own output currents. Slave current changes will cause forward voltage drop changes in their rectifying diodes, lowering the peak voltage slightly at the input to their voltage-averaging LC filter. These, of course, are not corrected by the main feedback loop, which senses only V_m. But so long as the slave output inductors, $L2$, $L3$ and especially main inductor $L1$ do not go discontinuous, slave output voltages can be depended on to remain constant to about ±5 to ±8 percent.

2.2.3 Slave absolute output voltage levels

Although slave output voltage changes are relatively small, their absolute values are not accurately adjustable. As seen in Eqs. 2.2 and 2.3, they are fixed by T_{on} and their corresponding secondary turns N_{s1},N_{s2}. But T_{on} is fixed by the feedback loop in keeping the master constant. Also, since the number of slave secondary turns can be changed only by integral numbers, the absolute value of slave output voltages is not finely settable. The change in secondary voltage for a single turn change in N_s is given by Faraday's law (Eq. 1.17). From that relation, the secondary volts per turn is

$$\frac{E}{N} = A_e\,(dB/dt) \times 10^{-8} \tag{2.4}$$

Here A_e is the core iron area in cm^2 and dB is the flux change in gauss in a time $dt(= T_{on})$ in seconds. Since T_{on} is some fraction of a half period (see Fig. 2.1), the volts per turn E/N is directly proportional to the switching frequency. In practice, the volts per turn ranges from about 2 V at a switching frequency of 25 kHz to 5 or 6 V at 100 kHz. In Eq. 2.4, since E/N is also proportional to dB, at 100 kHz it is less than four times its value at 25 kHz. This is so because, as will be seen in the chapter on magnetics, higher core losses at the higher frequency force a lower permissible dB at the high frequency.

In most cases, the absolute value of slave output voltages is not too important. Slaves usually drive operational amplifiers or motors, and these most often can tolerate DC voltage levels within about 2 V of a desired value. If absolute magnitude of a slave output is important, it is usually designed higher than required and brought back down to a desired exact value with a linear regulator or buck regulator. Because a slave output is semiregulated (against line changes only), it is not too inefficient to use a linear regulator. In such cases, the slave absolute voltage level is designed to be at about the minimum headroom (2.5 to 3 V) above the desired exact output voltage when DC input voltage is at its minimum.

2.2.4 Master output inductor minimum current limitations

The selection of the output inductor for a buck regulator was discussed in Sec. 1.3.6. It was mentioned that at a DC current level when the step at the front end of the inductor current waveform has fallen to zero (Figs 1.6*a* and 1.6*b*), the inductor is said to run dry or go discontinuous. Up to this point and beyond, the feedback loop maintains the buck regulator's output voltage constant, but with somewhat poorer load regulation after the inductor has gone discontinuous.

In Fig. 1.6*a*, however, it can be seen that before going discontinuous the on time is very nearly constant over large output current changes.

After run-dry, the on time changes drastically. In the buck regulator, this does not matter at all as only the one voltage is involved and the feedback loop keeps the output voltage constant. But in the push-pull width-modulated converter with a master and some slaves, the slave output voltages are directly proportional to the master's on time as shown by Eqs. 2.2 and 2.3.

It is thus very important when slaves are involved that the master output inductor not be permitted to go discontinuous until it reaches its minimum specified output current level. If the master's minimum output current is specified at one-tenth its nominal value, an output inductor selected from Eq. 1.8 will prevent the discontinuous mode condition until the master output current has fallen to its minimum value. Up to this point, the slave output voltages will remain constant to within about 5 percent. After the master inductor has gone discontinuous—below its minimum DC current specification—T_{on} will decrease significantly, and so will slave output voltages. But the feedback loop will keep the master output voltage constant.

Slave outputs, of course, must not be permitted to go discontinuous until they reach their own specified minimum current limits. Slave output inductors should also be selected from Eq. 1.8 if their minimum currents are as low as one-tenth their nominal value. If minimum currents are only one-fifth their nominal value, inductors can be half the Eq. 1.8 value before the discontinuous mode situation occurs.

2.2.5 Flux imbalance in push-pull topology

The push-pull converter is one of the oldest topologies and a still valuable one. It has multiple outputs whose returns are DC-isolated from input ground and from one another. Output voltages can be higher or lower than DC input voltages. The master is regulated against line and load variations. The slaves are equally well regulated against line changes and to within about 5 percent for load changes as long as output inductors are not permitted to go discontinuous.

The topology was widely used and, in very large numbers, performed successfully in the field. Yet occasionally and in greater numbers as output powers increased, power transistor failures started occurring in the field for unknown reasons. Waveforms prior to shipping looked normal and field performance prior to the failures showed no problems. Generally, no obvious design flaw would be found and after just simple replacement of the power transistors, the supplies would continue to work well for months until another unexplained transistor failure would occur.

Eventually, a rather subtle reason for these failures became obvious and widely known throughout the industry. It led to less frequent use

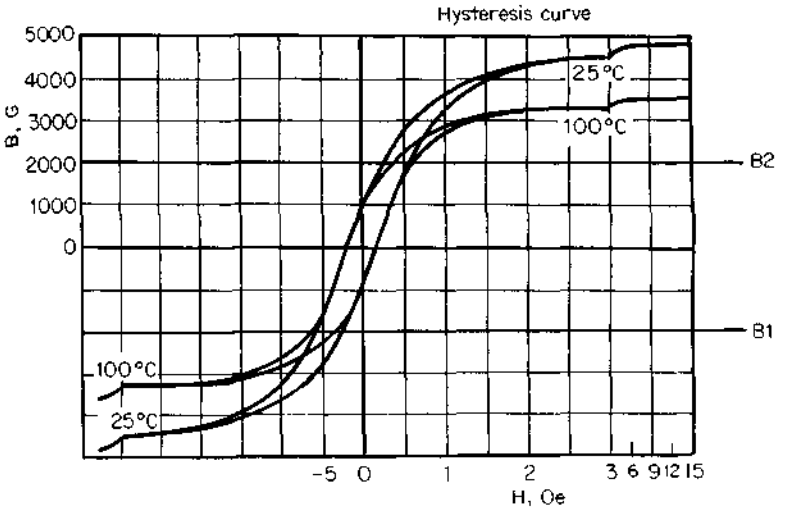


Figure 2.3 Hysteresis loop of a typical ferrite core material (Ferroxcube 3C8). Flux excursions are generally limited to ±2000 G up to about 30 kHz by requirement to stay on the linear part of the loop. At frequencies of 100 to 300 kHz, peak flux excursions must be reduced to about ±1200 or ±800 G because of core losses at these higher frequencies.

of the push-pull topology or to applications where simple and inexpensive "fixes" could avoid the problem.

This subtle failure mode in push-pull converters is known as *flux imbalance*. It can be understood by examination of the hysteresis loop (Fig. 2.3)* of a typical ferrite core material used in the power transformer.

In normal operation, the core's flux excursion is between levels such as B_1 and B_2 gauss in Fig. 2.3. It is desirable to stay on the linear part of the hysteresis loop below ±2000 G. At frequencies of up to 25 kHz, core losses are low and it may be permitted to move between plus and minus 2000 G. Yet other considerations discussed in Sec. 2.2.9.4 dictate that in a conservative design, excursions between ±1600 G are preferable even though losses at ±2000 G are easily acceptable. But core losses go up rapidly with frequency, and above 100 kHz they may limit peak flux density to 1200 or even 800 G.

Now (see Fig. 2.1) when $Q1$ is on, the no-dot end of N_{p1} is positive with respect to the dot end and the core moves up the hysteresis loop—say, from B_1 toward B_2. The actual amount it moves up is proportional to the product of the voltage across N_{p1} and the $Q1$ on time (from

*Hysteresis loop of 3C8, a typical ferrite from Ferroxcube Corporation. Other materials from the same or other manufacturers are very similar. Other materials differ mainly in core losses and Curie temperature.

Faraday's law; Eq. 1.18). When $Q1$ turns off and $Q2$ turns on, the dot end of N_{p2} is positive with respect to the no-dot end and the core moves back down from B_2 toward B_1. The actual amount it moves down is now proportional to the voltage across N_{p2} and the $Q2$ on time.

Further, if the volt-second product across N_{p1} when $Q1$ is on is equal to the volt-second product across N_{p2} when $Q2$ is on, after one complete period, the core will have moved up from B_1 to B_2 and returned exactly to B_1. But if those volt-second products differ by only a few percent and the core has not returned to its exact starting point in one period, after a number of periods, the core will "walk" or drift off center of the hysteresis loop into saturation. In saturation, of course, the core cannot sustain voltage and the next time the appropriate transistor turns on, it is subjected to high current and high voltage and is destroyed.

Now there are a number of factors which cause the on or set volt-second product to be different from the off or reset volt-second product. The $Q1$ and $Q2$ collector voltage on times are not exactly equal even if their base voltage on times are equal. Base voltage on times are very closely equal as the usual integrated-circuit control chip which generates them makes them so.

But if $Q1$, $Q2$ are bipolar transistors, they have "storage" times which keep the collector on for a longer time than does the base on voltage time. Storage times can range from 0.3 to 6 μs and have large production spread. They also are very temperature-dependent, increasing significantly as temperature increases; and even if, fortunately, $Q1$ and $Q2$ have equal storage times at the same temperature, they may be very unequal if located far apart on a heat sink and operate at different temperatures.

Further, if one transistor has a slightly larger volt second product than the other, it will start the core drifting slightly off center toward saturation. This will cause that transistor to draw slightly more current than the other as the core moves onto the curved part of the hysteresis loop (Fig. 2.3) and the core's magnetizing current on that half period starts to become a significant part of the load current. Now the transistor which draws more current will run slightly warmer, increasing its storage time. With a longer storage time in that transistor, the volt-second product it applies to the core in its on half period increases, the current in that half period increases, storage time in that transistor increases still further. Thus a runaway condition arises which quickly drives the core into saturation and destroys the transistor.

The on volt-second product between $Q1$ and $Q2$ can also differ because of their initially unequal "on" or $V_{ce(sat)}$ voltages, which have a significant production spread. As described above, with bipolar tran-

sistors, any initial difference in on voltage is magnified because the on voltage with bipolars decreases as temperature decreases.

If $Q1$, $Q2$ are MOSFET transistors, the flux imbalance problem is much less serious. To start with, MOSFETs have no storage time and with equal input on (gate) times, output (drain) times are equal—and, importantly, the on voltage of a MOSFET transistor increases as temperature increases. Thus the runaway condition described above cannot occur.

To the contrary, the fact that MOSFET on voltage increases with temperature provides negative feedback, which tends to correct any tendency to flux imbalance. If there were any initial inequality in volt-second product, in the half period in which it was larger, transistor current would be greater as the core started moving up the curved part of the hysteresis loop. That transistor with the larger current would run warmer, and its on voltage would increase and rob voltage from its half primary. This would decrease the volt-second product in that half period and bring the transistor current back down.

2.2.6 Indications of flux imbalance

A push-pull converter can continue to operate for some time with a certain amount of flux imbalance without immediately saturating its core and destroying its transistors. The mechanism described above may imply a slight imbalance in volt-second product between each half cycle causes certain failure. But failure is not certain.

Obviously, unless there was an inherent corrective mechanism, core saturation and transistor failure would always occur. It would simply take more switching cycles with a small volt-second imbalance. Thus if there were an initial volt-second imbalance of 0.01 percent (which would be practically impossible to achieve), it would take only 10,000 cycles before the core would move from a lowermost starting point of B_1 (Fig. 2.3) to a starting point of B_2, and the transistors would be destroyed considerably before that.

The corrective mechanism which permits the converter to survive and operate with no external indications of possible imminent failure is simply the primary wiring resistance. If there is an initial volt-second imbalance in the half period with the larger product, the transistor will commence taking more current as the core commences walking up the curved part of its hysteresis loop. But as that side takes more current, it produces a larger voltage drop across its half primary resistance. That voltage drop robs volt-seconds from the winding and restores the volt-second balance.

The converter can then remain in the unbalanced state without immediately going into run away and completely saturating the core. An

indication of where the core is on the hysteresis loop and how close to disaster the converter is can be obtained by placing a current probe in transformer center tap as shown in Fig. 2.4*d*.

The waveform indicating a volt-second balance is shown in Fig. 2.4*a*, where alternate current peaks are equal. Primary load current

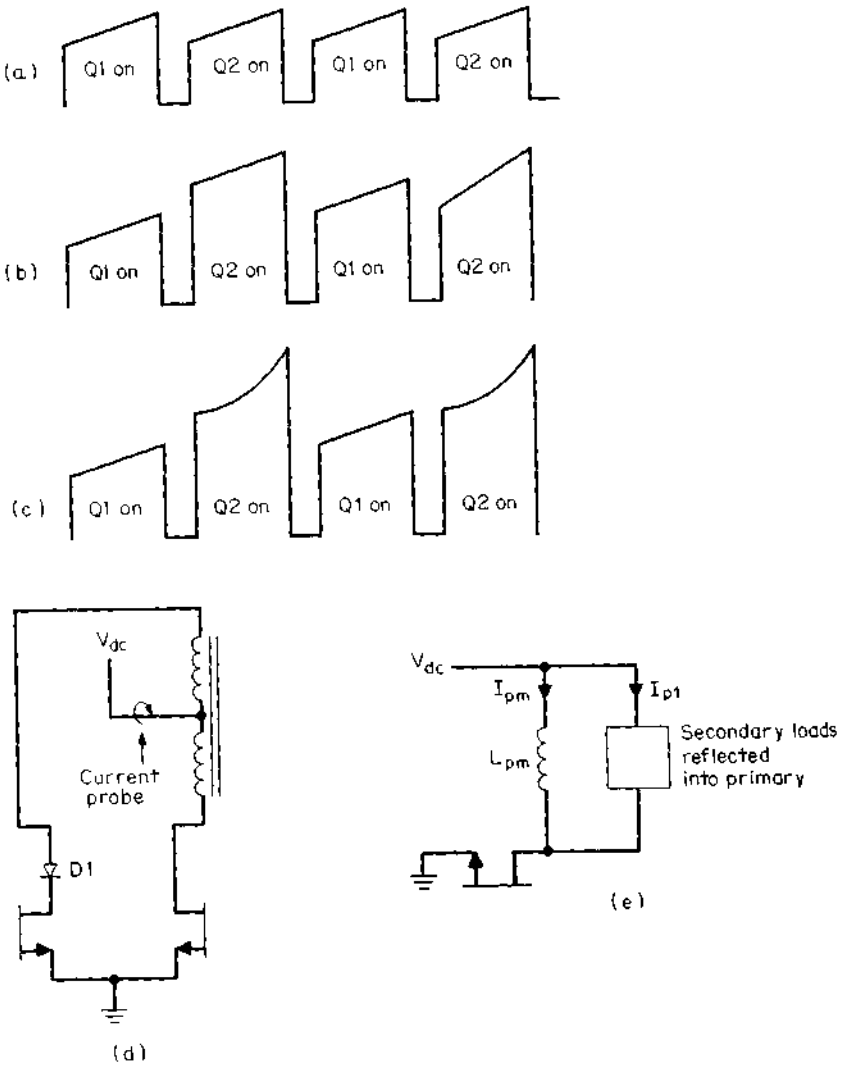


Figure 2.4 (*a*) Current in transformer center tap. Equal volt-second product on both halves of transformer primary. (*b*) Transformer center tap current. Unequal volt-second product on both halves of transformer primary. Core is not yet on curved part of hysteresis loop. (*c*) *Transformer center tap current.* Unequal volt-second product. Upward concavity indicates dangerous situation. *Core is far up on curved part of hysteresis loop.* (*d*) Adding a diode in series with one side of primary to test how serious a volt-second inequality exists. (*e*) Total primary current is the sum of the ramp-on-a-step reflected secondary load currents plus the linear ramp of magnetizing current.

pulses have the characteristic shape of a ramp on a step just as for the buck regulator in Fig. 1.4*d*. They have this shape because all the secondaries have output *LC* filters which generate such waveshapes as described in Sec. 1.3.2.

Now the primary load current waveshape is the sum of all the secondary ramps on a step reflected into the primary by their respective turns ratios.

However, the total primary current is the sum of these secondary currents reflected into the primary plus the so-called magnetizing current. This magnetizing current is the current drawn by the *magnetizing inductance,* which is the inductance seen looking into the primary with all secondaries open-circuited. This inductance is always present in a loaded transformer and is effectively in parallel with the secondary currents reflected into the primary as in Fig. 2.4*e*.

The waveshape of the total primary current is then the sum of the ramp-on-a-step reflected load currents and the magnetizing current. But the magnetizing current is a linear ramp starting from zero current level. When a transistor turns on, it applies a step of voltage of approximately $V_{dc} - 1$ across the magnetizing inductance L_m. Magnetizing current then ramps up linearly at a rate $dI/dt = (V_{dc} - 1)/L_m$ and for the transistor on time of T_{on} reaches a peak of

$$I_{pm} = \frac{(V_{dc} - 1)(T_{on})}{L_{pm}} \tag{2.5}$$

This I_{pm} is kept small compared to the sum of the load currents reflected into the primary by ensuring that L_m in Eq. (2.5) is large. Maximum permissible peak magnetizing current should be no greater than 10 percent of the primary load current.

Since the ramp of magnetizing current is small, added to the ramp on a step of load current, it simply increases the slope of the latter slightly and is not noticeable. Also, if volt-seconds are equal on alternate half cycles, peak currents will also be equal on each half cycle as in Fig. 2.4*a* as operation is centered around the zero-oersted (0-Oe) origin of the hysteresis loop of Fig. 2.3.

However, if volt-second products on alternate half cycles are unequal, core operation is not centered around the origin of the hysteresis loop. Since oersteds are proportional to magnetizing current, this shows up as a DC current bias as in Fig. 2.4*b*, making alternate current pulses unequal in amplitude.

As long as the DC bias does not drive the core appreciably up the hysteresis loop, the slope of the ramp still remains linear (Fig. 2.4*b*) and operation is still safe. Primary wiring resistance keeps the core from moving further up into saturation.

But if there is a large inequality in volt-seconds on alternate half cycles, the core is biased closer toward saturation and enters the curved part of the hysteresis loop. Now, magnetizing inductance, which is proportional to the slope of the hysteresis loop, decreases and magnetizing current increases significantly. This shows up as an upward concavity in the current slope as in Fig. 2.4*c*.

This is a dangerous and imminent failure situation. Now small temperature increases can bring on the runaway scenario described above. The core will be driven hard into saturation and destroy the power transistor. A push-pull converter design should certainly not be considered safe if current pulses in the primary center tap show any upward concavity in their ramps. Even linear ramps as in Fig. 2.4*b* with anything greater than 20 percent inequality in peak currents are unsafe and should not be accepted.

2.2.7 Testing for flux imbalance

A simple test to determine how close to a dangerous flux imbalance situation a push-pull converter may be is shown in Fig. 2.4*d*. Here a silicon diode with about 1 V forward drop is placed in series with the end of either half of the transformer. Now in the on state, that half with the diode in series has 1 V less voltage across it than the other half and there is an artificially produced volt-second unbalance. The center tap waveform will then look like either Fig. 2.4*b* or 2.4*c*. The current ramp corresponding to the side which does not have the diode will have the larger volt-second product and the larger peak current. By switching the diode to the other side, the larger peak current will be seen to switch to the opposite transformer half primary.

Now it can be determined how close to the upward concave situation of Fig. 2.4*c* the circuit is. If one series diode can make a current ramp go concave, the circuit is too close to imminent failure. Going to two series diodes on one side will give an indication of how much margin there is.

It should be noted that primary magnetizing current contributes no power to the secondaries. It will not be reflected by any turns ratio into the secondaries. It simply swings the magnetic core across the hysteresis loop of Fig. 2.3 and is related to the driving force in oersteds in that figure by the fundamental magnetic relation

$$H = \frac{0.4\pi N_p I_m}{L_m} \tag{2.6}$$

where H is in oersteds
N_p is in turns
I_m is in amperes
L_m is in henries

2.2.8 Coping with flux imbalance

The flux imbalance problem has caused the push-pull topology to fall out of favor in recent years. There are a number of ways to circumvent the problem, but they all more or less have drawbacks in terms of increased cost or components. Some schemes to combat flux imbalance are described in the following subsections.

2.2.8.1 Gapping the core. Flux imbalance becomes serious when the core moves out onto the curved part of the hysteresis loop (Fig. 2.3) and magnetizing current starts increasing exponentially as in Fig. 2.4c. This can be avoided by moving the curved part of the hysteresis loop out to greater oersted levels. The core can then tolerate a larger DC current bias or volt-second product inequality.

An air gap introduced in series with the magnetic flux lines has the effect shown in Fig. 2.5. It tilts the slope of the hysteresis loop, keeping the point where it crosses the zero-gauss (0-G) level (the so-called coercive force H_c) fixed. An air gap of 2 to 4 mils brings the curved portion of the loop far out in oersteds so that the core can accept a reasonably large volt-second product inequality for cores used in the 10- to 500-W power level.

The air gap for a prototype EE or cup core is usually done with plastic shims in the center and outer legs. Since the flux passes through the center leg and returns through the outer legs, the total gap is twice the shim thickness. In a production transformer, it is not very much more expensive to have the center leg ground down to twice the

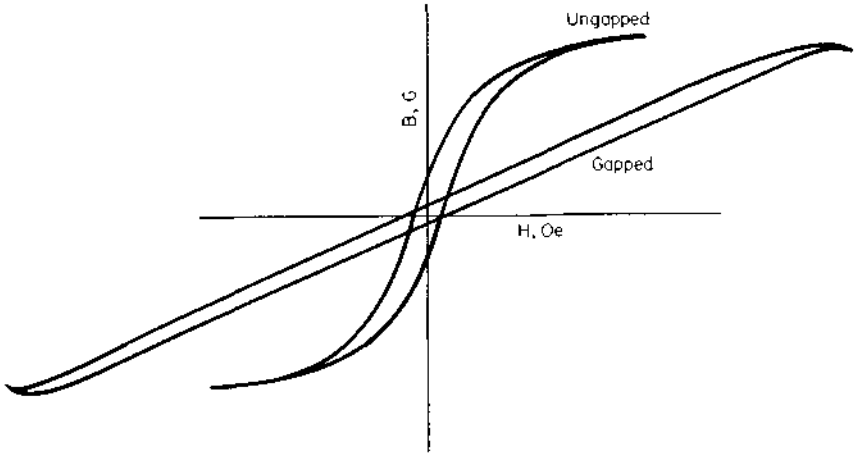


Figure 2.5 Gapping a core tilts its hysteresis loop. The gapped core shifts the curved part of the hysteresis loop out to higher oersted region and permits operation with a greater flux imbalance or DC current bias without saturation.

shim thickness. This will achieve pretty much the same effect as shims in the center and outer legs but is preferable as the gap remains more constant over time and results in less RFI interference.

2.2.8.2 Adding primary resistance. It was pointed out in Sec. 2.2.6 that primary wiring resistance keeps the core from rapidly being driven into saturation if there is a volt-second inequality. If there is such an inequality, the half primary with the larger volt-second product draws a larger peak current. That larger current causes a larger voltage drop across the wiring resistance and robs volt-seconds from that half primary, restoring the current balance.

This effect can be augmented by adding additional resistance in series with both primary halves. The added resistors can be located in either the collectors or emitters of the power transistors. They are best determined empirically by observing the current pulses in the transformer center tap. The resistors are usually under 0.25 Ω.

2.2.8.3 Matching power transistors. Since a volt-second inequality arises mainly from an inequality in storage time or on voltage in the power transistors, if those parameters are matched, it adds confidence that together with the above two "fixes" there will be no problem with flux imbalance.

This is not a certain or inexpensive fix. It is quite expensive to match transistors in two parameters. It requires a specialized test set up to do the matching and this is not available if field replacements become necessary.

It also must be ascertained that if the matching is done at certain load currents and room temperature, the matching still holds when these vary. Further, a storage time match is difficult to make credible as it depends strongly on forward and reverse base input currents in the bipolar transistors. Generally, any matching is usually done by matching V_{ce} and V_{be} (the on collector-to-emitter and base-to-emitter voltages) at the maximum operating current. Matching is not a viable solution for high-volume commercial supplies.

2.2.8.4 Using MOSFET power transistors. Since most of the volt-second inequality arises from storage time inequality between the two power transistors, this problem largely disappears if MOSFETs which have no storage times are used.

Further, there is an added advantage as the on voltage of a MOSFET transistor increases with temperature. Thus if one half primary tends to take a larger current, its transistor runs somewhat warmer and its on voltage increases and steals voltage from the winding. This reduces the volt-second product on that side and tends to re-

store balance. This, of course, is qualitatively in the right direction but cannot be depended on quantitatively and with certainty to solve the flux-imbalance problem at all power levels and with a worst-case combination of the two power transistors.

Yet with power MOSFETs, at power levels under 100 W and with a gapped core, push-pull converters can be built with a high degree of confidence that there will be no flux imbalance problem.

2.2.8.5 Using current-mode topology. The best solution to the flux-imbalance problem is not to use the push-pull topology at all but a modified two-transistor push-pull version of it. This version—*current-mode topology*—has all the advantages of the conventional push-pull, solves the flux-imbalance problem with certainty, and has significant advantages of its own.

In conventional push-pull, there is always a residual fear that despite all the fixes, a flux-imbalance problem will arise in some worst-case situations. One transistor will start drawing more current as its half primary moves into saturation, and the transistor will be destroyed. Current-mode topology solves this problem by monitoring the current in each of the push-pull transistors on a pulse-by-pulse basis and forcing alternate current pulses to have equal amplitude.

Details of current-mode topology will be discussed in a later chapter.

2.2.9 Power transformer design relations

2.2.9.1 Core selection. Design of the transformer starts with an initial decision on the selection of a core for the desired total output power. The available output power from a core depends on operating frequency, peak operating flux density (B_1 and B_2 in Fig. 2.3), the core's iron area A_e and bobbin area A_b, and the current density in each winding.

Decisions on each of these parameters are interrelated, and choices are made to minimize the transformer size and its temperature rise. In Chap. 7 on magnetics, an equation will be derived showing the maximum available output power from a given core as a function of the parameters mentioned above.

The equation can be used in a set of iterative calculations, first making a tentative selection of a specific core, peak flux density, and operating frequency and calculating the available output power. Then if the available power is insufficient, a larger-sized core is selected and the calculations repeated until a core with the required output power is found.

This is a long and cumbersome procedure, and instead, the equation is turned into a set of charts which permit a core and operating fre-

quency to be selected at a glance for any desired output power. Such equations and charts will be presented for most of the commonly used topologies in the chapter on magnetics.

But for the present, it will be assumed that these charts will be used to select a core and its iron area A_e is known. The rest of the transformer design involves calculation of the number of turns on the primary and secondaries, selection of wire sizes, and calculation of core and copper losses and transformer temperature rise.

Optimum sequencing of the various layers of wire on the core bobbin is important in improving coupling between the windings and reducing copper losses due to "skin" and "proximity" effects. Sequencing and skin and proximity effects will be discussed in Chap. 7. For the present, the design will proceed from the known value of the core iron area A_e for the core chosen from the selection charts described above.

2.2.9.2 Maximum power transistor on-time selection. Equation 2.1 shows that the converter keeps output voltage V_m constant by increasing T_{on} as V_{dc} decreases. Thus the maximum on time T_{on} occurs at the minimum specified DC input voltage V_{dc}. But the maximum on time must not exceed half the switching period. For if it did, the reset volt-second product would be less than the set volt-second product (Sec. 2.2.5) and after a very few cycles, the core would drift into saturation and destroy the power transistor.

Moreover, because of storage time in bipolar transistors, the base drive on time cannot be as large as a full half period. For storage time would cause the collector voltage on time to overlap with the on time of the opposite transistor on time. With simultaneous conduction in the two power transistors, the two half primaries could not support voltage. For (see Fig. 2.1) the dot end of N_{p1} would be trying to go negative and the no-dot end of N_{p2} would be trying to go negative also. Consequently, there would be no dB/dT in the core and from Faraday's law, there could be no voltage across the windings. Thus each transistor would take large currents at the full supply voltage and would immediately be destroyed.

Thus, to ensure that the core will always be reset within one period and there is never any possibility of simultaneous conduction, whenever the DC input voltage is at its minimum V_{dc} and the feedback loop is trying to increase T_{on} to maintain V_m constant, the on time will be constrained by some kind of a clamp to never be more than 80 percent of a half period. Then in Eq. 2.1, for specified V_{dc}, T and for $\overline{T_{on}} = 0.8T/2$, the ratio N_m/N_p will be fixed to yield the desired output V_m.

2.2.9.3 Primary turns selection. The number of primary turns is determined by Faraday's law (Eq. 1.17). From it N_p is fixed by the minimum voltage across the primary (V_{dc} − 1) and the maximum on time, which as above, is to be no more than $0.8T/2$. Then

$$N_p = \frac{(V_{dc} - 1)(0.8T/2) \times 10^{+8}}{A_e\, dB} \tag{2.7}$$

Since A_e in Eq. 2.7 is fixed by the selected core, V_{dc} and T are specified and the number of primary turns is fixed as soon as dB (the desired flux change in $0.8T/2$) is decided on. This decision is made as follows.

2.2.9.4 Maximum flux change selection. From Eq. 2.7, it is seen that the number of primary turns is inversely proportional to dB, the flux swing. It would seem desirable to maximize dB so as to minimize N_p, for fewer turns would mean that larger wire size could be used, resulting in higher permissible currents and more output from a given core. Also, fewer turns would result in a less expensive transformer and lower stray parasitic capacities.

From the hysteresis loop of Fig. 2.3, however, it is seen that in ferrite cores, the loop enters the curved portion above ±2000 G. It is desirable to stay below that point as there the magnetizing current starts increasing exponentially. It might seem, then, that a good choice is ±2000 G if that is not limited by core losses.

Now ferrite core losses increase at about the 2.7th power of the peak flux density and at about the 1.6th power of the operating frequency. Up to about 50 kHz, core losses do not prohibit operation to ±2000 G and it would seem desirable at least up to that frequency to operate at that high peak flux level.

However, for a not easily evident reason, it is preferable to restrict operation to ±1600 G even at frequencies where core losses are not prohibitive. The explanation is as follows. Faraday's law solved for the flux change dB is

$$dB = \frac{(V_{dc} - 1)(T_{on}) \times 10^{+8}}{N_p A_e} \tag{2.8}$$

Now Eq. 2.8 says that if N_p is chosen for a given dB—say, from −2000 to +2000 G or a dB of 4000 G, as long as the product of $(V_{dc} - 1)(T_{on})$ is constant, dB will be constant at 4000 G. Further, if the feedback loop is working and keeping the output voltage V_m constant, Eq. 2.1 says that $(V_{dc} - 1)(T_{on})$ is constant and dB will truly remain constant. That is, the feedback loop always ensures that when-

ever V_{dc} is a minimum, T_{on} is a maximum and there can never be a situation when T_{on} is a maximum and V_{dc} is simultaneously a maximum.

Now in some transient or fault condition even if for a single or possibly a few cycles, if T_{on} had been a maximum and V_{dc} had a transient step to 50 percent above its normal value and the feedback loop failed to reduce the on time in accordance with Eq. 2.1, the condition of maximum V_{dc} and maximum T_{on} could exist. Equation 2.8 then says that dB would be 1.5(4000) or 6000 G.

Then if the core had started from the −2000-G point, at the end of that on time the core would have been driven 6000 G above that or to +4000 G. Also, the hysteresis loop (Fig. 2.3) shows that at temperatures somewhat above 25°C, it would be deep in saturation and could not support the applied voltage. The transistor would be subject to high current and high voltage and would immediately fail.

This situation could arise as can be seen from the circuit in Fig. 2.1 if there were a delay in the response of the error amplifier. The error amplifier shown in that figure will always eventually correct the on time so as to keep the product $(V_{dc} - 1)(T_{on})$ constant in accordance with Eq. 2.1, for that is how the circuit regulates. But as will be seen in the chapter on feedback analysis, the error amplifier has a delay in its response time as its bandwidth is limited to stabilize the feedback loop.

Thus if even for a single cycle, if the core is subjected to maximum input voltage and maximum on time as a result of error-amplifier delay, it will saturate and destroy the transistor.

However, if N_p in Eq. 2.8 is chosen to yield dB of 3200 G at $\underline{V_{dc}}$ and $\overline{T_{on}}$ the design is safer and can tolerate a 50 percent transient step in input voltage. With dB = 3200 G, if the error amplifier is too slow to correct the on time, the transient dB will be 1.5(3200) or 4800 G; and if the core started from its normal location of −1600 G, it will be driven up to only −1600 + 4800 or +3200 G. The hysteresis loop of Fig. 2.3 shows that the core can tolerate that even at 100°C.

Thus the number of primary turns is selected from Eq. 2.7 for dB = 3200 G even at frequencies where larger flux changes would not yield excessive core losses. Above 50 kHz, increasing core losses force a decrease in peak flux density. At 100 to 200 kHz, peak flux density may have to be limited to 1200 or possibly to 800 G, respectively, to achieve an acceptably low core temperature rise.

2.2.9.5 Secondary turns selection. Secondary turns for the main and slave output voltages are selected from Eqs. 2.1, 2.2, and 2.3. In those equations all parameters have already been specified, arbitrarily set, or calculated. Output voltages, $\underline{V_{dc}}$ and T have been specified. The

maximum on time $\overline{T_{on}}$ has been arbitrarily set at $0.8T/2$ and N_p has been calculated from Faraday's law (Eq. 2.7) for the known A_e from the selected core. Flux change dB has been set at 3200 G for frequencies under 50 kHz and to minimize core losses, at lower values at higher frequencies as discussed above.

2.2.10 Primary, secondary peak and rms currents

Selection of wire sizes for all the windings requires knowledge of the rms currents they carry. For wire sizes will be selected on the basis of a conservative value of their operating current density in terms of circular mils* of area per rms ampere.

2.2.10.1 Primary peak current calculation. Current drawn from the DC input source V_{dc} may be monitored in the transformer center tap and has the waveshape shown in Fig. 2.1*b* and 2.1*d*. The pulses have the characteristic ramp-on-a-step waveshape because the secondaries all have output LC filters as discussed in Sec. 1.3.2. The primary current is simply the sum of all the secondary ramp-on-a-step currents reflected by their turns ratios into the primary plus the magnetizing current.

As discussed in Sec. 2.2.9.2, at minimum DC input voltage V_{dc}, transistor on times will be 80 percent of a half period and since there is one pulse for each half period, the duty cycle of the pulses in Fig. 2.1 is 0.8 at V_{dc}. To simplify calculation, the pulses in the figure are assumed to have an equivalent flat-topped waveshape whose amplitude I_{pft} is the value of the current at the center of the ramp.

Then the input power at V_{dc} is that voltage times the average current, which is $0.8I_{pft}$, and assuming 80 percent efficiency (which is usually achievable up to 200 kHz), $P_o = 0.8P_{in}$ or

$$P_{in} = 1.25P_o = V_{dc}\, 0.8I_{pft}$$

Then

$$I_{pft} = 1.56\,\frac{P_o}{V_{dc}} \tag{2.9}$$

This is a useful relation, as it gives the equivalent flat-topped primary current pulse amplitude in terms of what is known—the output power and the specified minimum DC input voltage. It permits select-

*A circular mil is the area of a circle 1 mil in diameter. Thus, area in square inches = $(\pi/4)10^{-6}$ (area in circular mils).

ing a transistor with an adequate current rating and choosing primary wire size from the corresponding primary rms current.

2.2.10.2 Primary rms current calculation and wire size selection. Each half primary carries only one of the I_{pft} pulses per period and hence its duty cycle is $(0.8T/2)/T$ or 0.4. It is well known that the rms value of a flat-topped pulse of amplitude I_{pft} at a duty cycle D is

$$I_{rms} = I_{pft}\sqrt{D} = I_{pft}\sqrt{0.4}$$

or

$$I_{rms} = 0.632 I_{pft} \tag{2.10}$$

and from Eq. 2.9

$$I_{rms} = 0.632 \frac{1.56P_o}{V_{dc}} = \frac{0.986P_o}{V_{dc}} \tag{2.11}$$

This again gives the rms current in each half primary in terms of what is known: output power and the specified minimum DC input voltage.

Now conservative practice in transformer design is to operate the windings at a current density of 500 circular mils per rms ampere. There is nothing absolute about this; current densities of 300 circular mils per rms ampere are frequently used for windings with few turns. As a general rule, however, densities greater than 300 circular mils per rms ampere should be avoided as that will cause excessive copper losses and temperature rise.

Thus at 500 circular mils per rms ampere, the required number of circular mils for the half primaries is

$$\text{Primary circular mil requirement} = 500 \frac{0.986P_o}{V_{dc}}$$

$$= 493 \frac{P_o}{V_{dc}} \tag{2.12}$$

This, too, is in terms of known values—output power and specified minimum DC input voltage. Proper wire size can then be chosen from wire tables at the circular mils given by Eq. 2.12.

2.2.10.3 Secondary peak, rms current, and wire size calculation. Currents in each half secondary are shown in Fig. 2.6. Note the ledge at the end of the transistor on time. This ledge of current exists because there is no free-wheeling diode $D1$ at the input to the filter inductor as

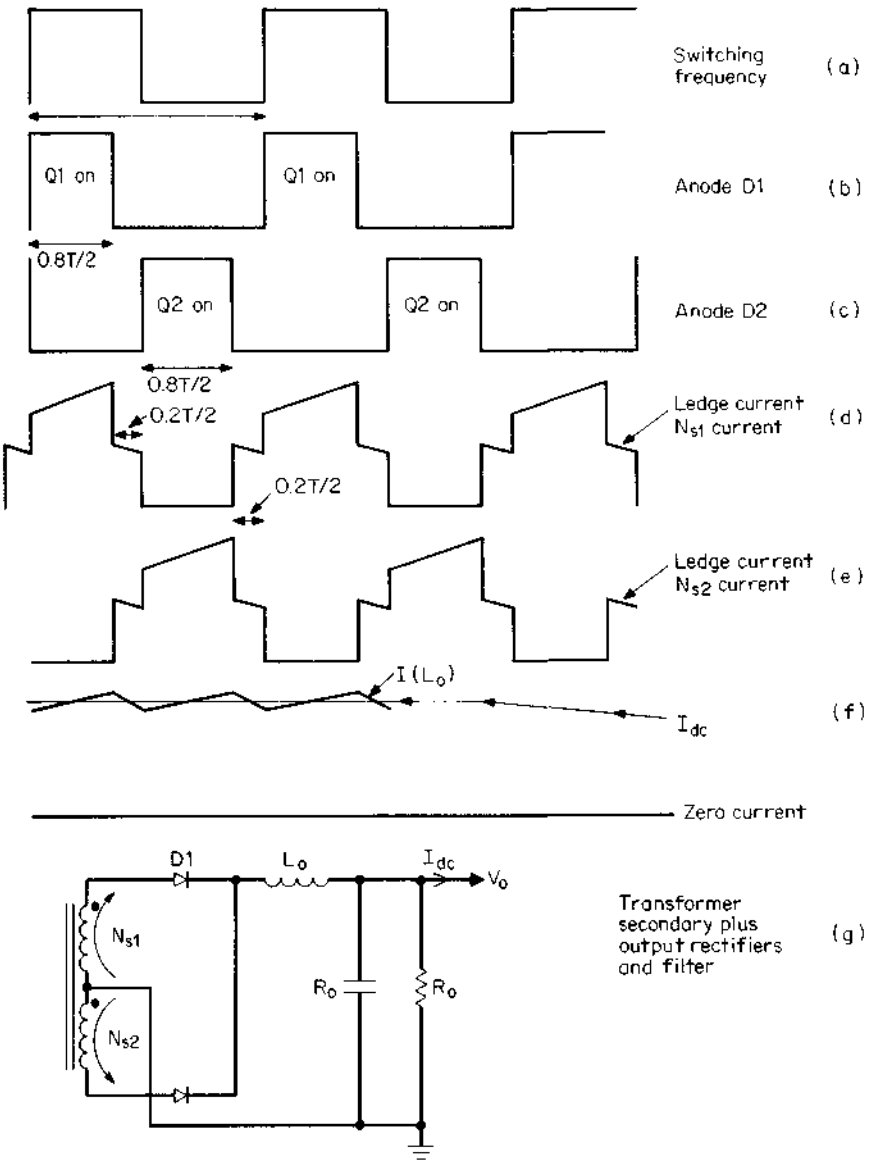


Figure 2.6 Output rectifiers serve as free-wheeling diodes in a push-pull circuit. Secondaries carry the normal free-wheeling "ledge" current during the 20 percent dead time. This should be considered in estimating secondary copper losses.

in the buck regulator of Fig. 1.4. In the buck, the free-wheeling diode was essential as a return path for inductor current when the transistor turned off. When the transistor turned off, the polarity across the output inductor reversed and its front end would have gone disastrously negative if it had not been caught by the free-wheeling diode at about 1 V below ground. Inductor current then continued to flow through the free-wheeling diode $D1$ of Fig. 1.4*e*.

In the push-pull output stages, however, the function of the free-wheeling diode is performed by the output rectifiers. When either transistor turns off, the front end of the inductor tries to go negative. As soon as it goes about one diode drop below ground, both rectifiers latch in, each drawing roughly half the total current the inductor had been drawing just prior to turnoff (Fig. 2.6*d* and 2.6*e*). Since the impedance of each half secondary is small, there is negligible drop across them and the rectifier diode cathodes are caught at about 1 V below ground.

Thus if half-secondary rms currents are to be calculated exactly, the ledge currents during the 20 percent dead time should be taken into account. But as can be seen, they are only about half the peak inductor current and have a duty cycle of $(0.4T/2)/T$ or 0.2. With such small amplitudes and duty cycle they can be ignored in calculating the half-secondary rms currents. Each half secondary can then be considered to have the characteristic ramp-on-a-step waveform, which at minimum DC input comes at a duty cycle of $(0.8T/2)/T$ or 0.4. The magnitude of the current at the center of the ramp is the DC output current I_{dc}, as can be seen from Fig. 2.6*f*.

As for the primary currents, to simplify calculating rms currents, the ramp-on-a-step pulses will be approximated by "equivalent flat-topped" pulses I_{sft} whose amplitude is that at the center of the ramp or the DC output current I_{dc} and which come at a duty cycle of 0.4.

Thus rms current in each half secondary is

$$I_{s(rms)} = I_{dc}\sqrt{D} = I_{dc}\sqrt{0.4} = 0.632I_{dc} \tag{2.13}$$

At 500 circular mils per rms ampere, the required number of circular mils for each half secondary is

$$\begin{aligned}\text{Secondary circular mil requirement} &= 500(0.632)I_{dc} \\ &= 316I_{dc}\end{aligned} \tag{2.14}$$

2.2.11 Transistor voltage stress and leakage inductance spikes

It can be seen from the transformer dots in Fig. 2.1 that when either transistor is on, the opposite transistor's collector is subject to at least

twice the DC supply voltage since both half primaries have an equal number of turns.

However, the maximum stress is somewhat more than twice the maximum DC input voltage. The added contribution comes from the so-called leakage inductance spikes shown in Fig. 2.1*a* and 2.1*c*. These come about because there is an effective small inductance (leakage inductance L_l) in series with each half primary as shown in Fig. 2.7*a*.

At the instant of turnoff, current in the transistor falls rapidly at a rate dI/dT causing a positive-going spike of amplitude $E_{ls} = L_l\, dI/dT$ at the bottom end of the leakage inductance. Conservative design practice is to assume the leakage inductance spike may be as much as

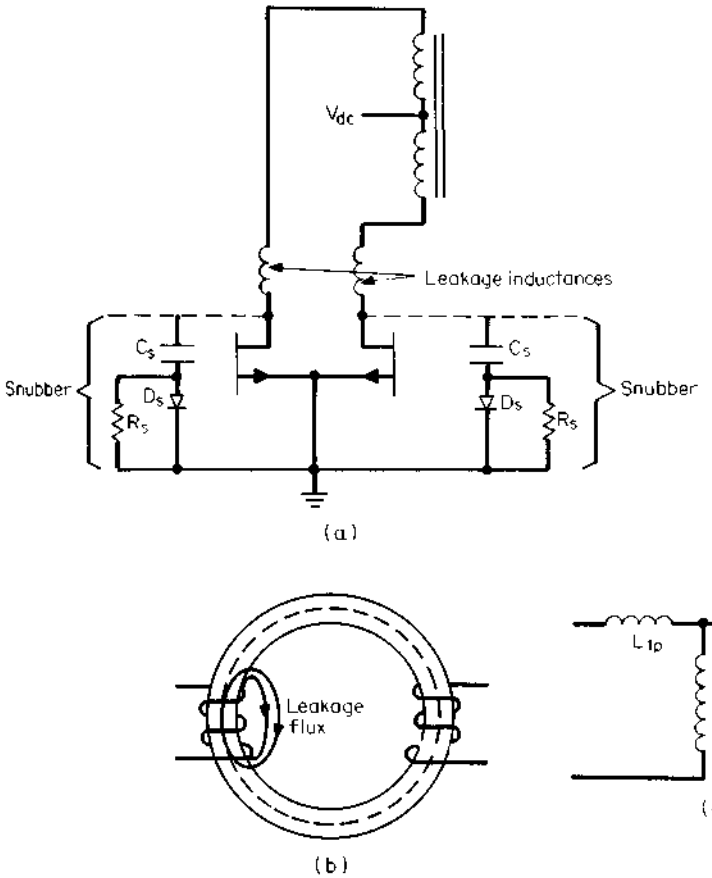


Figure 2.7 (*a*) Leakage inductances cause the spikes of Fig. 2.1*a*. (*b*) Leakage inductance stems from the fact that some of the magnetic flux lines return through a local air path rather than linking the secondary through the core. (*c*) Low-frequency equivalent circuit of a transformer showing magnetizing inductance L_m and primary and secondary leakage inductances L_{1p} and L_{1s}.

30 percent more than twice the maximum DC input voltage. The transistors should then be chosen so that they can tolerate with some safety margin, a maximum voltage stress (V_{ms}) of

$$V_{ms} = 1.3(2\overline{V_{dc}}) \tag{2.15}$$

Magnitude of the leakage inductance is not easily calculable. It can be minimized by use of a transformer core with a long center leg and sandwiching the secondary windings (especially the higher current ones) in between halves of the primary. A good transformer should have leakage inductance no more than 2 to 4 percent of its magnetizing inductance.

Leakage inductance spikes can be minimized by addition of a capacitor, resistor, and diode (CRD) combination to the transistor collector as shown in Fig. 2.7*a*. Such CRD configurations (referred to as "snubbers") also serve the important function of reducing AC switching losses due to the overlap of falling transistor current and rising voltage at the collector. Detailed design of snubbers and some associated penalties they incur will be discussed in Chap. 11.

Leakage inductance arises from the fact that some of the primary's magnetic flux lines do not return through the core and couple with the secondary windings. Instead, they return around the primary winding through a local air path as seen in Fig. 2.7*b*.

The equivalent circuit of a core with its magnetizing inductance L_m (Sec. 2.2.6) and primary L_{lp} and secondary L_{ls} leakage inductances is shown in Fig. 2.7*c*. Secondary leakage inductance arises from the fact that some of the secondary current's magnetic flux lines also do not couple with the primary but instead link the secondary windings via a local air path. But in most cases, the number of secondary turns is smaller than the primary and L_{ls} can be neglected.

The transformer's equivalent circuit (Fig. 2.7*c*) is very valuable in understanding many odd circuit effects. That equivalent circuit is an accurate representation up to about 300 or 500 kHz, where shunt parasitic capacitors across and between windings must be taken into account.

2.2.12 Power transistor losses

2.2.12.1 AC switching or current-voltage "overlap" losses. Leakage inductance in the power transformer causes a very rapid collector voltage fall time because for a short time at $T = 0$, for a fast current step, an inductor is an infinite impedance. Since the current cannot change instantaneously through an inductor, current rises slowly through it.

Thus there is very little overlap of falling voltage and rising current at turnon and negligible switching losses.

At turnoff, however, there is significant overlap and a worst-case scenario such as that for the buck regulator of Fig. 1.5*b* is assumed. The exact situation is shown in Fig. 2.8, where it is assumed that the current hangs on at its equivalent flat-topped peak value I_{pft} (Sec. 2.2.10.1) for the time it takes the voltage to rise T_{vr} to its maximum value of $2\overline{V_{dc}}$. Then the voltage remains at $2\overline{V_{dc}}$ during the time it takes the current to fall from I_{pft} to zero T_{cf}. Assuming $T_{vr} = T_{cf} = T_s$ and a switching period T, the total switching dissipation per transistor per period $P_{t(ac)}$ is

$$P_{t(ac)} = I_{pft}\frac{2\overline{V_{dc}}}{2}\frac{T_s}{T} + 2V_{dc}\frac{I_{pft}}{2}\frac{T_s}{T}$$

$$= 2(I_{pft})(\overline{V_{dc}})\frac{T_s}{T}$$

and from Eq. 2.9, $I_{pft} = 1.56(P_o/\overline{V_{dc}})$:

$$P_{t(ac)} = 3.12\frac{P_o}{\overline{V_{dc}}}\overline{V_{dc}}\frac{T_s}{T} \tag{2.16}$$

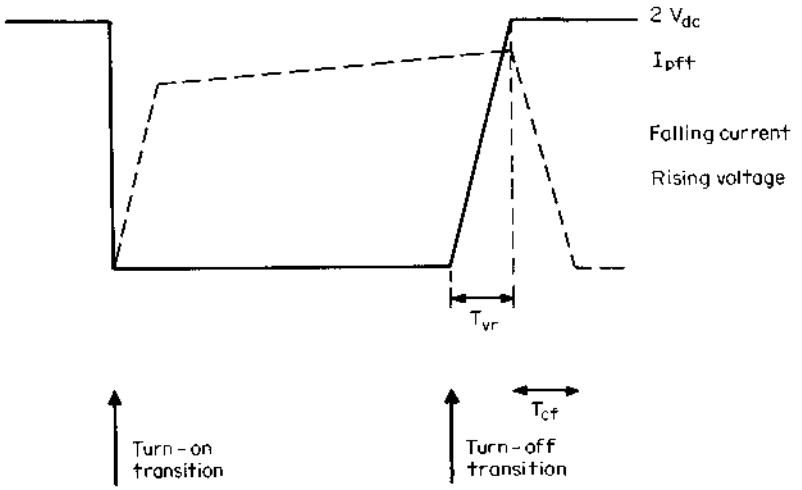


Figure 2.8 AC or "overlap" switching losses in push-pull transistor. There are negligible switching losses at turnon because transformer leakage inductance causes a very fast voltage fall down and a slow current rise time. This results in very little overlap of falling voltage and rising current. Worst-case scenario is assumed at turnoff. Current remains constant at its peak I_{pft} until voltage rises to $2\overline{V_{dc}}$. The voltage remains at $2\overline{V_{dc}}$ for the duration of the current fall time T_{cf}.

2.2.12.2 DC conduction losses. Maximum DC losses per transistor are simply the on voltage times the on current times the on duty cycle per transistor, or

$$P_{dc} = I_{pft} V_{on} \frac{0.8T/2}{T} = 0.4 I_{pft} V_{on}$$

It will be seen in Chap. 8 on bipolar base drives that a technique called *Baker clamping* will be used to reduce transistor storage times. This forces the collector on potential to be about 1 V over a large range of current. Then for I_{pft} from Eq. 2.9 we obtain

$$P_{dc} = 0.4 \frac{1.56P_o}{V_{dc}} = \frac{0.624P_o}{V_{dc}} \tag{2.17}$$

and total losses per transistor are

$$\begin{aligned} P_{total} &= P_{(ac)} + P_{dc} \\ &= 3.12 \frac{P_o}{V_{dc}} \overline{V_{dc}} \frac{T_s}{T} + \frac{0.624P_o}{V_{dc}} \end{aligned} \tag{2.18}$$

2.2.12.3 Typical losses: 150-W, 50 kHz push-pull converter. It is instructive to get a feel for the dissipation per transistor in a 150-W push-pull converter at, say, 50 kHz operating from a standard telephone industry power source. Standard telephone industry power sources provide a nominal voltage of 48 V, minimum (V_{dc}) of 38 V and maximum ($\overline{V_{dc}}$) of 60 V. It is assumed that at 50 kHz, bipolar transistors will be used and a reasonable value of the switching time (T_s as defined above) is 0.3 μs.

The DC conduction losses from Eq. 2.17 are

$$P_{dc} = \frac{0.624 \times 150}{38} = 2.46 \quad \text{W}$$

and the AC switching losses from Eq. 2.16 are

$$P_{(ac)} = 3.12 \times \frac{150}{38} \times 60 \times \frac{0.3}{20} = 11.08 \quad \text{W}$$

Thus the AC overlap or switching losses are about 4.5 times as great as the DC conduction losses. If MOSFET transistors with switching times T_s of about 0.05 μs, are used, the falling current and rising voltage intersect so low down on the current waveform that switching losses are negligible.

2.2.13 Output power and input voltage limitations in push-pull topology

Aside from the flux-imbalance problem in the push-pull topology (which does not exist in its close cousin, viz., current-mode topology), its area of usage in output power is limited by Eq. 2.9 and in input voltage by Eq. 2.15.

Equation 2.9 gives the peak current required of the transistor for a desired output power, and Eq. 2.15 gives the maximum voltage stress on the transistor in terms of the maximum DC input voltage. Both these relations limit the push-pull to about 300 W. Above that level, it is difficult to find a bipolar transistor which can meet the peak current and voltage stresses and be fast enough with adequate gain and low on voltage.

The faster MOSFET transistor with adequately high voltage and current ratings and sufficiently low on voltage may be too expensive for commercial applications.

Thus consider a 400-W push-pull converter operating from telephone industry prime voltage source which is 48 V (nominal), 38 V (minimum), and 60 V (maximum).

Equation 2.9 gives the peak current requirement as $I_{pft} = 1.56P_o/V_{dc} = 1.56(400)/38 = 16.4$ A, and Eq. 2.15 gives the maximum off voltage stress as $V_{ms} = 2.6V_{dc} = 2.6 \times 60 = 156$ V. To provide a margin of safety, a transistor with at least a 200-V rating would be sought.

A potential bipolar transistor candidate for the above supply is possibly the MJ13330. It has a 20-A peak current rating, V_{ceo} rating of 200 V, and V_{cev} (the voltage it can sustain if, at the instant of turnoff, it has a negative bias of -1 to -5 V) rating of 400 V. It can thus meet the peak voltage and current stresses.

At the 16-A level, it has a maximum on potential of about 3 V, a minimum gain of about 5, and a storage time of about 1.3 to 4 μs. With these limitations, it would have high DC and AC switching losses, have difficulty with flux imbalance (unless the current-mode equivalent version of push-pull were used) and would have difficulty operating above 40 kHz because of the long storage times.

A potential MOSFET transistor for such an application might be a MTH30N20. This is a 30-A, 200-V device which at 16 A would have 1.3 V of on drop and hence half the DC conduction losses of the preceding bipolar transistor. With its rapid current fall time, it would have negligible AC switching losses. But at $11.50 in quantities above 100, it would be too expensive for commercial applications. The 15-A, 200-V device (MTH15N200) could be used and would cost only $4.10 in quantities above 100. But it would have twice the on drop and twice the DC conduction losses.

For off-line converters the push-pull topology is almost always ruled out by the maximum off voltage stress of $2.6V_{dc}$ (Eq. 2.15). Thus for a 120-V-AC line input with ±10 percent tolerance, peak rectified DC voltage is 1.41 × 1.1 × 120 = 186 V. At the top of the leakage spike, Eq. 2.15 gives peak off stress as 2.6 × 186 = 484 V.

Transients above the maximum steady-state value must also be considered. Transients seldom are specified for commercial power supplies, in which case conservative design practice is to assume that they are 15 percent above the maximum steady-state value. This would bring the above maximum stress to 1.15 × 484 or 557 V.

Input voltage transients in special cases can be greater than 15 percent above maximum steady-state value. Thus AC input voltage specifications on military aircraft are given by Military Standard 704. That specification gives the nominal voltage as 113 V AC but with a 10-ms transient to 180 V AC. At 180 V AC, the peak off stress from Eq. 1.42 is 180 × 1.41 × 2.6 or 660 V.

There are many fast bipolar transistors, which can safely sustain voltages as high as 850 V if they have a reverse input bias of −1 to −5 V at turnoff (V_{cev} rating). But it is unnecessarily dangerous to use a topology which subjects the off transistor to twice the maximum DC input plus a 30 percent leakage spike. There are topologies to be described below which subject the off transistor to only the maximum DC input voltage with no leakage spike. These are a better choice—not only because of the lesser voltage stress but also because the smaller voltage excursion at turnoff produces less electromagnetic interference (EMI).

2.2.14 Output filter design relations

2.2.14.1 Output inductor design. It was pointed out in Sec. 2.2.4 that in both master and slave outputs the output inductors should not be permitted to go discontinuous. Recall that the discontinuous mode situation commences when the front end of the inductor current ramp of Fig. 1.6*b* has dropped to zero and that this occurs when the DC current has dropped to half the ramp amplitude dI (Sec. 1.3.6). Then (see Fig. 2.9)

$$dI = 2I_{dc} = V_L \frac{T_{on}}{L_o} = (V_1 - V_o) \frac{T_{on}}{L_o} \tag{2.19}$$

But $V_o = V_1(2T_{on}/T)$. Then

$$T_{on} = \frac{V_o T}{2V_1}$$

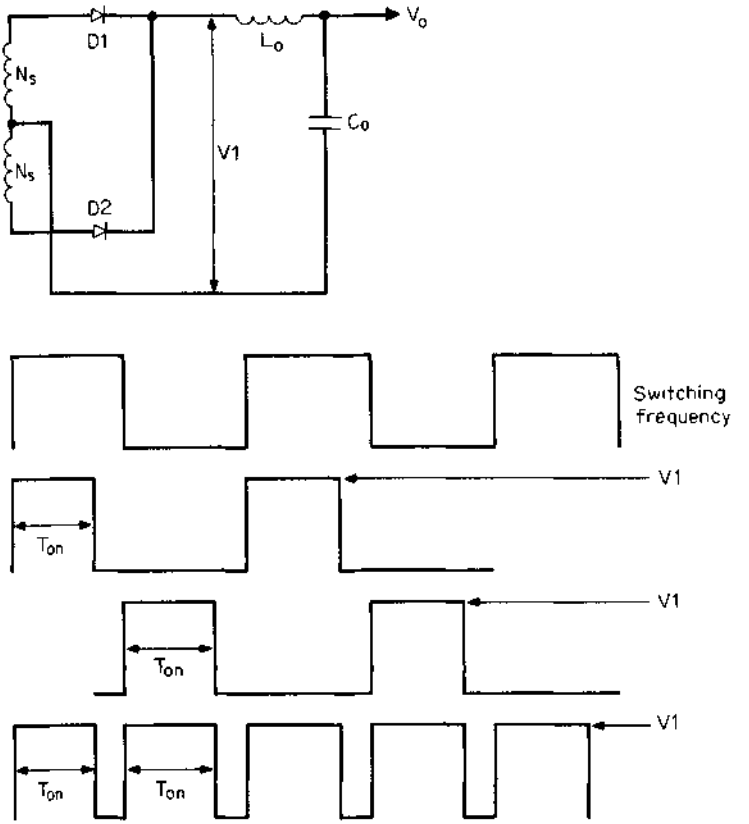


Figure 2.9 Output circuit for calculation and L_o, C_o. When V_{dc} at its minimum, N_s will be chosen so that as $V1$ is at its minimum, T_{on} will not have to be greater than $0.8T/2$ to yield the specified value of V_o.

But N_s will be chosen so that T_{on} will be $0.8T/2$ when V_{dc} and consequently V_1 are at their minima. Then

$$\overline{T_{on}} = \frac{0.8T}{2} = \frac{V_o T}{2\underline{V_1}} \quad \text{or} \quad \underline{V_1} = 1.25V_o$$

and

$$dI = \frac{(1.25V_o - V_o)(0.8T/2)}{L_o} = 2\underline{I_{dc}} \quad \text{and} \quad L_o = \frac{0.05V_o T}{\underline{I_{dc}}}$$

Then if the minimum current $\underline{I_{dc}}$ is specified as one-tenth the nominal current I_{on} (the usual case)

$$L_o = \frac{0.5V_oT}{I_{on}} \tag{2.20}$$

where L_o is in henries
V_o is in volts
T is in seconds
I_{dc} is minimum output current in amperes
I_{on} is nominal output current in amperes

2.2.14.2 Output capacitor design. The output capacitor C_o is selected to meet the maximum output ripple voltage specification. In Sec. 1.3.7 it was shown that the output ripple is determined almost completely by the magnitude of the ESR (equivalent series resistance, R_o) in the filter capacitor and not by the magnitude of the capacitor itself. The peak-to-peak ripple voltage V_r is very closely equal to

$$V_r = R_o\, dI \tag{2.21}$$

where dI is the selected peak-to-peak inductor ramp amplitude.

However, it was pointed out that (for aluminum electrolytic capacitors) the product R_oC_o has been observed to be relatively constant over a large range of capacitor magnitudes and voltage ratings. For aluminum electrolytics, the product R_oC_o ranges between 50 and 80×10^{-6}. Then C_o is selected as

$$C_o = \frac{80 \times 10^{-6}}{R_o} = \frac{80 \times 10^{-6}}{V_r/dI} = \frac{(80 \times 10^{-6})(dI)}{V_r} \tag{2.22}$$

where C_o is in farads for dI in amperes (Eq. 2.19) and V_r is in volts.

2.3 Forward Converter Topology

2.3.1 Basic operation

Forward converter topology is shown in Fig. 2.10. It is probably the most widely used topology for output powers under 150 to 200 W when maximum DC input voltage is in the range of 60 to 200 V. Below a maximum of 60 V, for the associated minimum input voltage, the required primary input current becomes uncomfortably large. Above a maximum DC input voltage of about 250 V, the maximum voltage stress on the transistor becomes too large. Above a required output power of 200 W, at any DC input voltage, the required primary input current becomes too large. Mathematical explanations for these limitations will be shown in the discussions to follow.

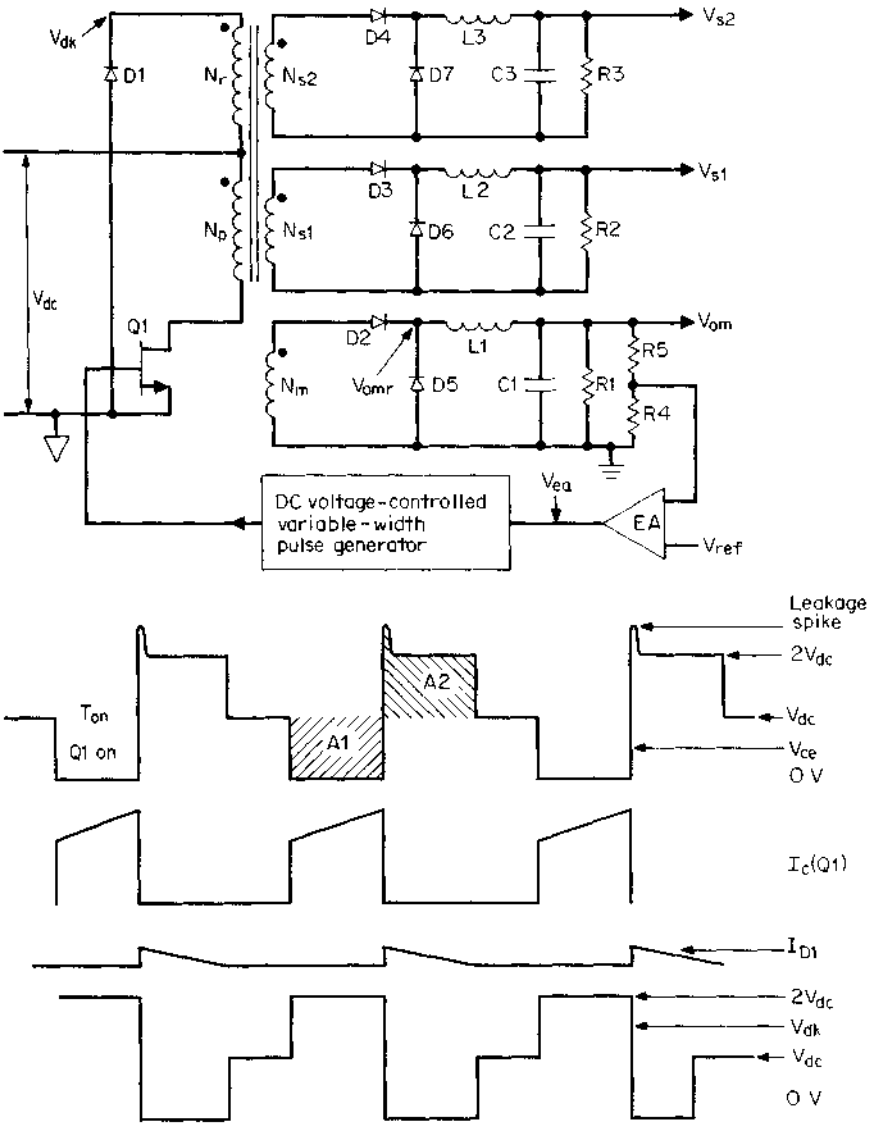


Figure 2.10 Forward converter topology. Showing feedback loop closed around one master N_{om} which is regulated against line and load changes. Also two semiregulated slaves (V_{s1} and V_{s2}) which are regulated against line changes only.

The topology is an outgrowth of the push-pull circuit of Fig. 2.1 but does not suffer from the latter's major shortcoming of flux imbalance. Since it has one rather than two transistors, compared to the push-pull, it is more economical in dollars and required space.

As for the push-pull, Fig. 2.10 shows a master output V_{om} and two

slaves, V_{s1} and V_{s2}. A negative-feedback loop is closed around the master and controls the $Q1$ on time so as to keep V_{om} constant against line and load changes. With an on time fixed by the master feedback loop, the slave outputs V_{s1}, V_{s2} are kept constant against input voltage changes but only partly (to about 5 to 8 percent) against load changes either in themselves or in the master. The circuit works as follows.

In the forward converter, one of the transistors of the push-pull of Fig. 2.1 has been replaced by the diode $D1$. When $Q1$ is turned on, the dot end of the primary power winding N_p and of all secondaries go positive with respect to their no dot ends. Current and power flows into the dot end of N_p. All rectifier diodes $D2$ to $D4$ are forward-biased and current and power flows out of the dot ends of all secondaries to the LC filters and the loads.

Note that power flows to the loads when the power transistor $Q1$ is turned on—thus the term *forward converter*. The push-pull and buck regulator also deliver power to the loads when the power transistors are on and are also forward types. In contrast, the boost regulator, the polarity inverter of Figs. 1.10 and 1.14, and the flyback type (to be discussed in a later chapter) store energy in an inductor or transformer primary when the power transistor is on and deliver it to the load when the transistor turns off. Such energy storage topologies can operate in either the discontinuous or continuous mode. They are fundamentally different from forward topologies, are discussed in Secs. 1.4.2 and 1.4.3, and will be taken up again in Chap. 4 on flyback topology.

Now, for a $Q1$ on time of T_{on}, the voltage at the master rectifier cathode (Fig. 2.10*b*) is at a high level for T_{on}. Assuming a 1-V on voltage for $Q1$ and a rectifier on forward drop of V_{D2}, that high-level voltage V_{omr} is

$$V_{omr} = \left[(V_{dc} - 1)\frac{N_m}{N_p}\right] - V_{D2} \tag{2.23}$$

The circuitry after the rectifier diode cathodes is exactly like that of the buck regulator of Fig. 1.4. Diodes $D5$ to $D7$ act like the freewheeling diode $D1$ of that figure. When $Q1$ turns off, current stored in the magnetizing inductance of $T1$ (recall the equivalent circuit of a transformer as in Fig. 2.7*c*) reverses the polarity of the voltage across N_p. Now all the dot ends of primary and secondary windings go negative with respect to their no-dot ends. If not for "catch" diode $D1$, the dot end of N_r would go very far negative and since N_p and N_r have equal turns (usually), the no-dot end of N_p would go sufficiently positive to avalanche $Q1$ and destroy it.

However, catch diode $D1$ catches the dot end of N_r at one diode drop below ground. If there were no leakage inductance in $T1$ (recall again

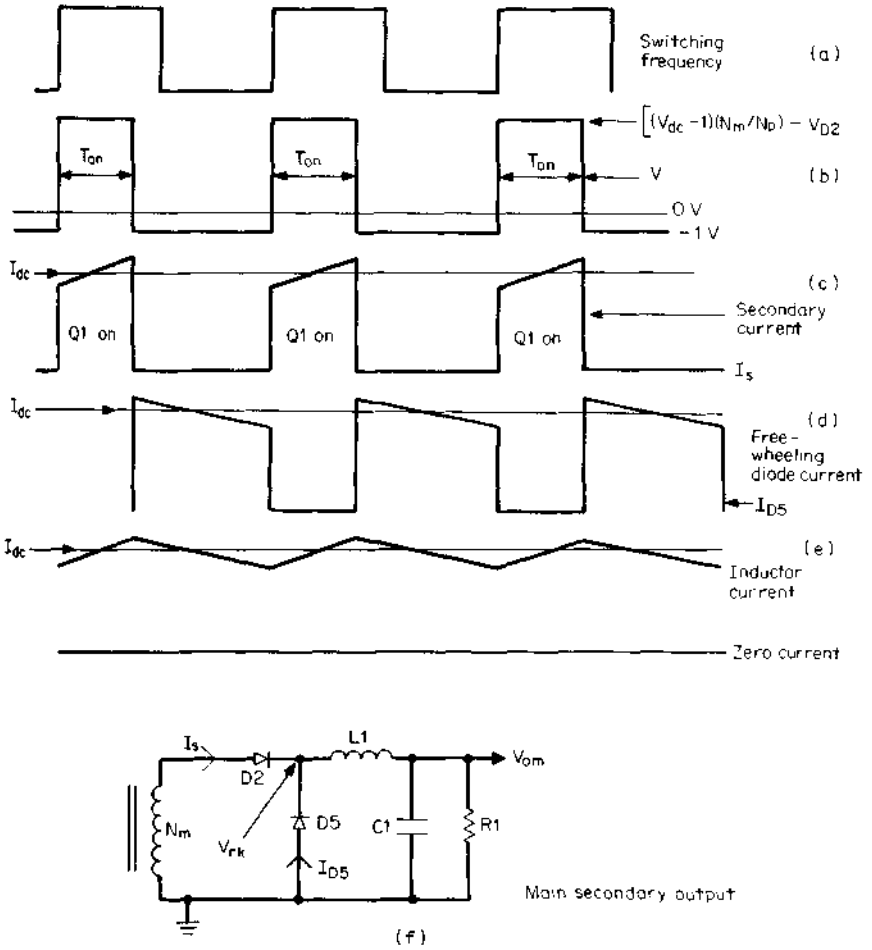


Figure 2.11 Critical secondary currents in forward converter. Each secondary has the characteristic ramp-on-a-step waveshape because of the fixed voltage across the output inductor and its constant inductance. Inductor current is the sum of the secondary plus the free-wheeling diode current. It ramps up and down about the DC output current. Primary current is the sum of all the ramp-on-a-step secondary currents reflected by their turns ratios into the primary. Primary current is then also at ramp-on-a-step waveform.

the equivalent circuit of a transformer as in Fig. 2.7*c*), the voltage across N_p would equal that across N_r. Assuming that the 1-V forward drop across $D1$ can be neglected, the voltage across N_r and N_p is V_{dc} and the voltage at the no-dot end of N_p and at the $Q1$ collector is then $2V_{dc}$.

Recall, now, that within one cycle, if a core has moved in one direction on its hysteresis loop, it must be restored to exactly its original position on the loop before it moves in the same direction again in the

next cycle. For otherwise, after many cycles, the core will be pushed in one direction or another into saturation. It will not be able to support the applied voltage and the transistor will be destroyed.

In Fig. 2.10, it is seen that when $Q1$ is on for a time T_{on}, N_p is subjected to volt-second product $V_{dc}T_{on}$ with its dot end positive. That volt-second product is the area $A1$ in Fig. 2.10. By Faraday's law (Eq. 1.17), that volt-second product causes—say, a positive—flux change $dB = (V_{dc}T_{on}/N_pA_e)10^{-8}$ gauss.

At turnoff, when the magnetizing inductance has reversed the polarity across N_p and kept its no dot-end at $2V_{dc}$ long enough for the volt-second area product $A2$ in Fig. 2.10 to equal area $A1$, the core has been restored to its original position on the hysteresis loop and the next cycle can safely start. In the common jargon expression, the "reset volt-seconds" has equaled the "set volt-seconds."

Now when $Q1$ has turned off, the dot ends of all secondaries go negative with respect to their no-dot ends. Current in all output inductors $L1$ to $L3$ try to decrease. Since current in inductors cannot change instantaneously, the polarity across all inductors reverses in an attempt to maintain constant current. The front end of the inductors try to go far negative but are caught at one diode drop below output ground by free-wheeling diodes $D5$ to $D7$ (Fig. 2.10) and rectifier diodes $D2$ to $D4$ are reverse-biased. Inductor current now continues to flow in the same direction through its output end, returning through the load, partly through the filter capacitor, and up through the free-wheeling diode back into the inductor.

Voltage at the cathode of the main diode rectifier $D2$ is then as shown in Fig. 2.11*b*. It is high at a level of $[(V_{dc} - 1)(N_m/N_p)] - V_{D2}$ for time T_{on}, and for a time $T - T_{on}$ it is one free-wheeling diode($D5$) drop below ground. The LC filter averages this waveform, and assuming that the forward drop across $D5$ equals that across $D2(=V_d)$, the DC output voltage at V_{om} is

$$V_{om} = \left[\left((V_{dc} - 1)\frac{N_m}{N_p}\right) - V_d\right]\frac{T_{on}}{T} \tag{2.24}$$

2.3.2 Design relations: output/input voltage, on time, turns ratios

The negative-feedback loop senses a fraction of V_{om}, compares it to a reference voltage V_{ref}, and varies T_{on} so as to keep V_{om} constant for any changes in V_{dc} or load current.

From Eq. 2.24, it can be seen that as V_{dc} changes, the feedback loop keeps the output constant by keeping the product $V_{dc}T_{on}$ constant.

Thus maximum $T_{on}(\overline{T_{on}})$ will occur at minimum specified $V_{dc}(\underline{V_{dc}})$ and Eq. 2.24 can be rewritten for minimum DC input voltage as

$$V_{om} = \left[\left(\left(\underline{V_{dc}} - 1\right)\frac{N_m}{N_p}\right) - V_d\right]\frac{\overline{T_{on}}}{T} \tag{2.25}$$

In relation 2.25, a number of design decisions must be made in the proper sequence. First, the minimum DC input voltage V_{dc} is specified. Then the maximum permitted on time T_{on} which occurs at $\underline{V_{dc}}$ (minimum V_{dc}) will be set at 80 percent of a half period.

This is done to ensure (Fig. 2.10) that the area $A2$ can equal $A1$. For if on time were permitted to go to a full half period, $A2$ would just barely equal $A1$ at the start of the next full cycle. Now small increases in on time due to storage time changes with temperature or production spreads in storage time would not permit $A2$ to equal $A1$. The core would not be completely reset to its starting point on the hysteresis loop, it would drift up into saturation after some cycles and destroy the transistor.

Next the number of primary turns N_p will be established from Faraday's law (Eq. 1.17) for V_{dc} and a certain specified flux change dB in the time T_{on}. Limits on that flux change are similar to those described for the push-pull topology in Sec. 1.5.9 and will also be discussed below.

Thus in Eq. 2.25, $\underline{V_{dc}}$, $\overline{T_{on}}$, T, and V_d are specified and N_p is calculated from Faraday's law. That fixes the number of main secondary turns N_m to achieve the required main output voltage V_{om}.

2.3.3 Slave output voltages

The slave output filters $L2$, $C2$ and $L3$, $C3$ average the width-modulated rectangular waveforms at their respective rectifier cathodes. The waveform upper levels are $[(V_{dc} - 1)(N_{s1}/N_p)] - V_{d3}$ and $[V_{dc} - 1)(N_{s2}/N_p)] - V_{d4}$, respectively. At their lower level, voltages are one diode drop below ground. They are at their high level for the same maximum T_{on} as the main secondary when the input DC input voltage is at its specified minimum V_{dc}. Then again assuming that the forward rectifier and free-wheeling diode drops are equal V_d, the slave output voltages at low line V_{dc} are

$$V_{s1} = \left[\left(\left(\underline{V_{dc}} - 1\right)\frac{N_{s1}}{N_p}\right) - .V_d\right]\frac{\overline{T_{on}}}{T} \tag{2.26}$$

$$V_{s2} = \left[\left(\left(\underline{V_{dc}} - 1\right)\frac{N_{s2}}{N_p}\right) - V_d\right]\frac{\overline{T_{on}}}{T} \tag{2.27}$$

Now in regulating V_{om}, the feedback loop keeps $V_{dc}T_{on}$ constant. But that same product appears in Eqs. 2.26 and 2.27 and hence the slave outputs remain constant as V_{dc} varies.

It can be seen from Eq. 2.24 and Fig. 2.14 that the negative-feedback loop keeps the main output constant for either line or load changes by appropriately controlling T_{on} so that the sampled output is equal to the reference voltage V_{ref}. This is not obvious for load changes as load current does not appear directly in Eq. 2.24. But it does appear indirectly. For load changes will change the on drop of $Q1$ (assumed as 1 V heretofore) and the forward drop in the rectifier diode. These, though small, will cause small changes in the output voltage, be sensed by the error amplifier, and be corrected for by a small change in T_{on}.

Moreover, as can be seen in Eqs. 2.26 and 2.27, changes in T_{on} without corresponding changes in V_{dc} will cause changes in the slave output voltages. Slave output voltages also change with changes in their own load currents. As those currents change, the rectifier forward drops change, causing a change in the peak voltage at the input to the LC averaging filter. Slave output voltages will change as peak voltages to the averaging filter change without a corresponding change in T_{on}.

Such changes in the slave outputs for load changes in the master and slave can be kept to within 5 to 8 percent. But as discussed in Sec. 2.2.4, neither master nor slave output inductors must be permitted to go discontinuous at their minimum load currents. This is done by choosing the output inductors appropriately large, as will be described below.

The number of slave secondary turns N_{s1}, N_{s2} is calculated from Eqs. 2.26 and 2.27 as all parameters there are either specified or calculated from specified values. The parameters V_{dc}, T, and V_d are all specified and $\overline{T_{on}}$ is set at $0.8T/2$ as discussed above; N_p is calculated from Faraday's law (Eq. 1.17) as described above.

2.3.4 Secondary load, free-wheeling diode, and inductor currents

Knowledge about the amplitude and waveshape of the various output currents is needed to select secondary and output inductor wire sizes and current ratings of the rectifier and free-wheeling diodes.

As described for the buck regulator in Sec. 1.3.2, secondary current during the $Q1$ on time has the shape of an upward-sloping ramp sitting on a step (Fig. 2.11c) because of the constant voltage across the inductor with its input end positive with respect to the output end.

Now when $Q1$ turns off, the input end of the inductor is negative with respect to the output end and inductor current ramps downward.

The free-wheeling diode, at the instant of turnoff, picks up exactly the inductor current which had been flowing in it just prior to turnoff. That diode current then ramps downward (Fig. 2.11*d*) as it is in series with the inductor. Inductor current is the sum of the secondary current when $Q1$ is on plus the free-wheeling diode current when $Q1$ is off and is shown in Fig. 2.11*e*. Current at the center of the ramp in either of Fig. 2.11*c*, 2.11*d*, or 2.11*e* is equal to the DC output current.

2.3.5 Relations between primary current, output power, and input voltage

Assume an efficiency of 80 percent from the DC input voltage node to the total output power from all secondaries. Then $P_o = 0.8P_{in}$ or $P_{in} = 1.25P_o$. Now calculate P_{in} at minimum DC input voltage $\underline{V_{dc}}$. That input power is $\underline{V_{dc}}$ times the average primary current at minimum DC input.

Now all secondary currents have the waveshape of a ramp sitting on a step as all secondaries have output inductors. These ramp-on-a-step waveforms have a width of $0.8T/2$ at minimum DC input voltage. All these secondary currents reflect into the primary by their turns ratio, and hence the primary current pulse is a single ramp-on-a-step waveform of width $0.8T/2$. There is only one such pulse per period (see Fig. 2.10) as this is a single-transistor circuit. The duty cycle of this primary pulse is then $(0.8T/2)/T$ or 0.4.

Now as for the push-pull topology, approximate this ramp-on-a-step by an equivalent flat-topped pulse I_{pft} of the same width and whose amplitude is that at the center of the ramp. The average value of this current is then $0.4I_{pft}$. Then

$$P_{in} = 1.25P_o = \underline{V_{dc}}(0.4I_{pft}) \quad \text{or} \quad I_{pft} = \frac{3.13P_o}{\underline{V_{dc}}} \tag{2.28}$$

This is a valuable relation. It gives the equivalent peak flat-topped primary current pulse amplitude in terms of what is known at the outset—the minimum DC input voltage and the total output power. This permits an immediate selection of a transistor with adequate current rating and gain if it is a bipolar transistor or with sufficiently low on resistance if it is a MOSFET type.

It can be seen that for a forward converter, Eq. 2.28 shows I_{pft} is twice the value required in a push-pull topology (Eq. 2.9) at the same output power and minimum DC input voltage.

This is obvious as in the push-pull there are two pulses of current or power per period as compared to a single pulse in the forward converter. In the forward converter from Eqn. 2.25, if the number of sec-

ondary turns is chosen large enough, then the maximum on time at minimum DC input voltage will not need to be greater than 80 percent of a half period. Then, as seen in Fig. 2.10, the area $A2$ can always equal $A1$ before the start of the next period. The core is then always reset to the same point on its hysteresis loop within one cycle and can never walk up into saturation.

But the penalty paid for this guarantee in the forward converter, that flux walking cannot occur is that the primary peak current is twice that for a push-pull at the same output power. In the push-pull, with all the precautions described in Sec. 2.2.8, there is never complete certainty that flux imbalance does not occur under unusual dynamic load or line conditions.

2.3.6 Maximum off-voltage stress in power transistor

In the forward converter, with the number of turns on the reset N_r winding equal to that on the power winding N_p, maximum off-voltage stress on the power transistor is twice the maximum DC input voltage plus a leakage spike. Leakage inductance spikes, and their origin and minimization have been discussed in Sec. 2.2.11. Conservative design, even with all precautions to minimize leakage spikes, should assume they may be 30 percent above twice the maximum DC input voltage. Maximum off-voltage stress is then the same as in the push-pull and is

$$V_{ms} = 1.3(2\overline{V_{dc}}) \tag{2.29}$$

2.3.7 Practical input voltage/output power limits

At the outset in Sec. 2.3.1 it was stated the practical maximum output power limit for a forward converter whose maximum DC input voltage is under 60 V is about 150 to 200 W. This is so because the peak primary current as calculated from Eq. 2.28 becomes excessive as there is only a single pulse per period as compared to two in the push-pull topology.

Thus consider a 200-W forward converter for the telephone industry where the specified minimum and maximum input voltages are 38 and 60 V, respectively. Peak primary current from Eq. 2.28 is $I_{pft} = 3.13P_o/V_{dc} = 3.13(200)/38 = 16.5$ A and from Eq. 2.29, maximum off voltage stress is $\overline{V_{ms}} = 2.6\overline{V_{dc}} = 2.6 \times 60 = 156$ V.

To provide safety margin, a device with at least a 200-V rating would be used to provide protection against input voltage transients

which could drive the DC input above the maximum steady-state value of 60 V.

Transistors with 200-V, 16-A ratings are available, but they all have drawbacks as was discussed in Sec. 2.2.13. Bipolar transistors are slow, and MOSFETs are easily fast enough but expensive. For such a 200-W application, a push-pull version guaranteed to be free from flux imbalance would be preferable; with two pulses of current per period, peak current would be only 8 A. With the lower peak current, noise spikes on ground buses and the radio-frequency interference (RFI) would be considerably less—a very important consideration for a telephone industry power supply. Such a flux imbalance free topology is "current mode," which is discussed later.

The forward converter topology, like the push-pull (discussed in Sec. 2.2.13), has the same difficulty in coping with maximum voltage stress in an off-line converter where the AC input voltage is 120 ±10 percent. At high line, the rectified DC input is $1.1 \times 120 \times 1.41 = 186$ V minus 2 V for the rectifier diode drops or 184 V. From Eq. 2.29 the maximum voltage stress on the transistor in the off state is $\overline{V_{ms}} = 2.6 \times 184 = 478$ V.

At minimum AC input voltage, the rectified DC output is $V_{dc} = (0.9 \times 120 \times 1.41) - 2 = 150$ V, and from Eq. 2.28, the peak primary current is $I_{pft} = 3.13 \times 22/150 = 4.17$ A.

Thus, for a 200-W off-line forward converter, the greater problem is more the 478-V maximum voltage stress than the 4.17-A peak primary current stress—for, as was seen in Sec. 2.2.13, when a 15 percent input transient is taken into account, the peak off voltage stress is 550 V. With a bipolar transistor operating under V_{cev} conditions (reverse input bias of −1 to −5 V at the instant of turnoff), a voltage stress of even 550 V is not a serious restriction. There are many devices with 650- to 850-V V_{cev} ratings and which have high gain, low on drop, and high speed at 4.17 A. But, as discussed in Sec. 2.2.13, there are preferable topologies, discussed below, which subject the off transistor to only V_{dc} and not twice V_{dc}.

2.3.8 Forward converter with unequal power and reset winding turns

Heretofore, it has been assumed that the numbers of turns on the power winding N_p and the reset winding N_r are equal. Some advantages result if N_r is made smaller or larger than N_p.

The number of primary power turns N_p is always chosen by Faraday's law and will be discussed in Sec. 2.3.10.2. If N_r is chosen less than N_p, the peak current required for a given output power is less than that calculated from Eq. 2.28, but the maximum $Q1$ off-voltage stress is

greater than that calculated from Eq. 2.29. If N_r is chosen larger than N_p, the maximum $Q1$ off-voltage stress is less than that calculated from Eq. 2.29, but the peak primary current for a given output power is greater than that calculated from Eq. 2.29. This can be seen from Fig. 2.12 as follows. When $Q1$ turns off, polarities across N_p and N_r reverse, the dot end of N_r goes negative and is caught at ground by catch diode $D1$. Transformer $T1$ is now an autotransformer. There is a

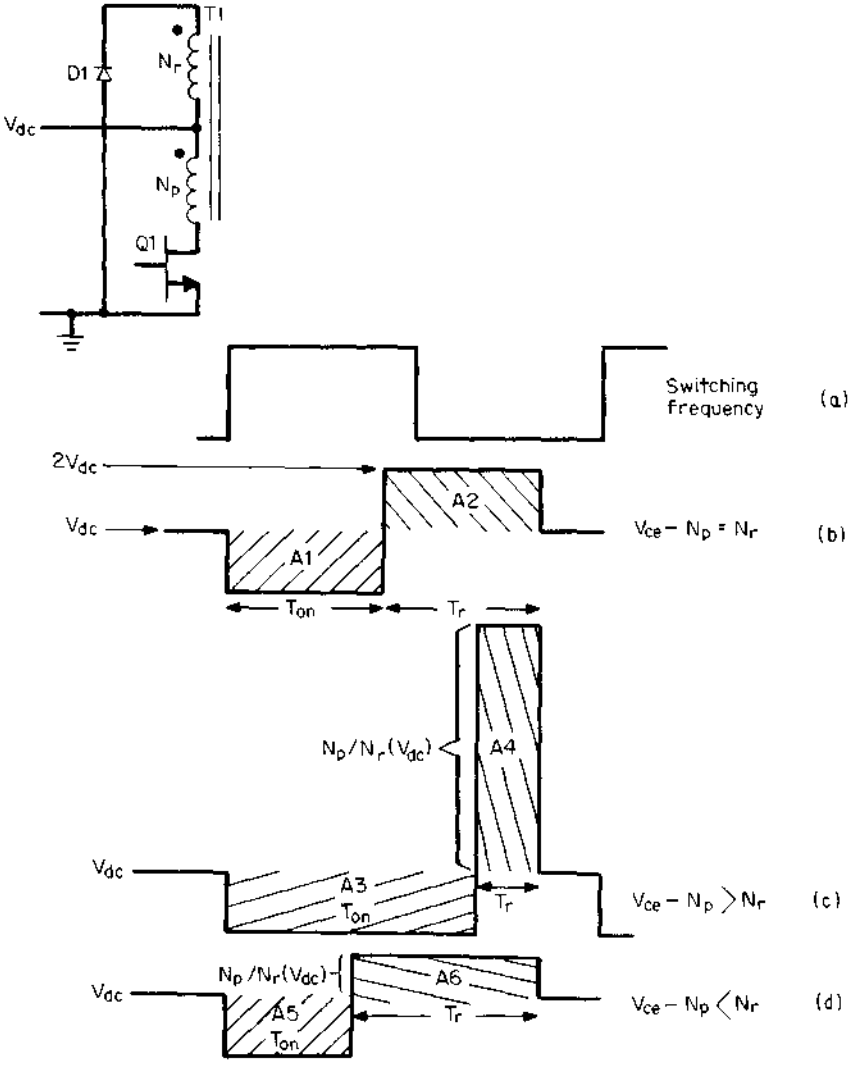


Figure 2.12 Forward converter–collector-to-emitter voltages for $N_p = N_r$. Note in all cases that reset volt-second product equals set volt-second product. (*a*) Switching frequency; (*b*) $N_p = N_r$; (*c*) $N_p > N_r$; (*d*) $N_p < N_r$.

voltage V_{dc} across N_r and hence a voltage $N_p/N_r(V_{dc})$ across N_p. The core is set by the volt-second product $V_{dc}T_{on}$ during the on time and must be reset to its original place on the hysteresis loop by an equal volt-second product. That reset volt-second product is $N_p/N_r(V_{dc})T_r$.

When N_r equals N_p, the reset voltage equals the set voltage and the reset time is equal to the set time (area $A1$ = area $A2$) as seen in Fig. 2.12*b*. For $N_r = N_p$, the maximum $Q1$ on time which occurs at minimum DC input voltage is chosen as $0.8T/2$ to ensure that the core is reset before the start of the next period; $T_{on} + T_r$ is then $0.8T$.

Now if N_r is less than N_p, the resetting voltage is larger than V_{dc} and consequently T_r can be smaller (area $A3$ = area $A4$) as shown in Fig. 2.12*c*. With a shorter T_r, T_{on} can be longer than $0.8T/2$ and $T_{on} + T_r$ can still be $0.8T$ so that the core is reset before the start of the next period. With a longer T_{on}, the peak current is smaller for the same average current and the same average output power. Thus in Fig. 2.12*c*, as compared to Fig. 2.12*b*, a smaller peak current stress has been traded for a larger voltage stress.

With N_r greater than N_p, the reset voltage is less than V_{dc}. Then if $T_{on} + T_r$ is still to equal $0.8T$, and the reset volt-seconds is to equal the set volt-seconds (area $A5$ = area $A6$ in Fig. 2.12*d*), T_r must be longer and T_{on} must be shorter than $0.8T/2$ as the reset voltage is less than the set voltage. With T_{on} less than $0.8T/2$, the peak current must be higher for the same average current. Thus in Fig. 2.12*d*, a lesser voltage stress has been achieved at the cost of a higher peak current for the same output power as in Fig. 2.12*b*. This can be seen quantitatively as

$$\text{Set } T_{on} + T_r = 0.8T; \qquad \text{reset voltage} = V_r = \frac{N_p}{N_r} V_{dc} \tag{2.30}$$

For on volt-seconds equal to reset volt-seconds

$$V_{dc}T_{on} = \frac{N_p}{N_r} V_{dc}T_r \tag{2.31}$$

Combining Eqs. 2.30 and 2.31

$$\overline{T_{on}} = \frac{0.8T}{1 + N_r/N_p} \tag{2.32}$$

For 80 percent efficiency $P_{in} = 1.25P_o$ and P_{in} at $\underline{V_{dc}} = \underline{V_{dc}}(I_{av}) = \underline{V_{dc}}\, I_{pft}(\overline{T_{on}})/T$ or $I_{pft} = 1.25(P_o/\underline{V_{dc}})(T/T_{on})$. Then from Eq. 2.32

$$I_{pft} = 1.56 \left(\frac{P_o}{\underline{V_{dc}}}\right)(1 + N_r/N_p) \tag{2.33}$$

and the maximum $Q1$ off voltage stress $\overline{V_{ms}}$—exclusive of the leakage spike—is the maximum DC input voltage V_{dc} plus the reset voltage (voltage across N_p when the dot end of N_r is at ground). Thus

$$\overline{V_{ms}} = \overline{V_{dc}} + \frac{N_p}{N_r}(\overline{V_{dc}}) = \overline{V_{dc}}(1 + N_p/N_r) \tag{2.34}$$

Values of I_{pft} and V_{ms} calculated from Eqs. 2.33 and 2.34 are

N_r/N_p	I_{pft} (from Eq. 2.33)	V_{ms} (from Eq. 2.34)
0.6	$2.50(P_o/\overline{V_{dc}})$	$2.67\overline{V_{dc}}$ + leakage spike
0.8	$2.81(P_o/\overline{V_{dc}})$	$2.25\overline{V_{dc}}$ + " "
1.0	$3.12(P_o/\overline{V_{dc}})$	$2.00\overline{V_{dc}}$ + " "
1.2	$3.43(P_o/\overline{V_{dc}})$	$1.83\overline{V_{dc}}$ + " "
1.4	$3.74(P_o/\overline{V_{dc}})$	$1.71\overline{V_{dc}}$ + " "
1.6	$4.06(P_o/\overline{V_{dc}})$	$1.62\overline{V_{dc}}$ + " "

2.3.9 Forward converter magnetics

2.3.9.1 First-quadrant operation only. The transformer core in the forward converter operates in the first quadrant of the hysteresis loop only. This can be seen in Fig. 2.10. When $Q1$ is on, the dot end of $T1$ is positive with respect to the no-dot end and the core is driven, say, in a positive direction on the hysteresis loop and the magnetizing current ramps up linearly in the magnetizing inductance.

When $Q1$ turns off, stored current in the magnetizing inductance reverses the polarity of voltages on all windings. The dot end of N_r goes negative until it is caught one diode drop below ground by catch diode $D1$. Now the magnetizing current which really is stored in the magnetic core continues to flow. It simply transfers from N_p, where it had ramped upward during the $Q1$ on time, into N_r. It flows out of the no-dot end of N_r into the positive end of the supply voltage V_{dc}, out of the negative end of V_{dc}, through the anode, then through the cathode of $D1$, and back into N_r.

During the $Q1$ off time, as the dot end of N_r is positive with respect to its no-dot end, the magnetizing current I_d ramps linearly downward as can be seen in Fig. 2.10. When it has ramped down to zero (at the end of area $A2$ in Fig. 2.10), there is no longer any stored energy in the magnetizing inductance and nothing to hold the dot end of N_r below the $D1$ cathode. The voltage at the dot end of N_r starts rising toward V_{dc}, and that at the no-dot end of N_p ($Q1$ collector) starts falling from $2V_{dc}$ back down toward V_{dc}.

Thus operation on the hysteresis loop is centered about half the peak magnetizing current ($V_{dc}T_{on}/2L_m$). Nothing ever reverses the direction of the magnetizing current—it simply builds up linearly to a peak and relaxes back down linearly to zero.

This first-quadrant operation has some favorable and some unfavorable consequences. First, compared to a push-pull circuit, it halves the available output power from a given core. This can be seen from Faraday's law (Eq. 1.17), which fixes the number of turns on the primary.

Solving Faraday's law for the number of primary turns, $N_p = E\,dt/A_e\,dB \times 10^{-8}$. If dB in the forward converter is limited to an excursion from zero to some B_{max} instead of from $-B_{max}$ to $+B_{max}$ as in a push-pull topology, the number of primary turns for the forward converter will be twice that in each half primary for a push-pull operating from the same V_{dc}. Although the push-pull has two half primaries, each of which must support the same volt-second product as the forward converter, the push-pull offers two power pulses per period as compared to one for the forward converter. The end result is that a core used in a forward converter offers only half the available output power it is capable of in a push-pull configuration.

But the push-pull core at twice the output power will run somewhat warmer as its flux excursion is twice that of the forward converter. Since core losses are proportional to the area of the hysteresis loop traversed, the push-pull core losses are twice that of the forward converter.

Yet total copper losses in both half primaries of a push-pull are no greater than that of the forward converter of half the output power. For the rms current in each push-pull half primary is equal to that in the forward converter of half the output power. But since the number of turns in each push-pull half primary is half that of the forward converter of half the output power, each push-pull half primary has half the resistance of the forward converter. Thus, total copper losses of a forward converter is then equal to the total losses of the two half primaries in a push-pull of twice the output power.

2.3.9.2 Core gapping in a forward converter. In Fig. 2.3, the hysteresis loop of a ferrite core with no air gap is shown. It is seen that at zero magnetizing force (0 Oe) there is a residual magnetic flux density of about ±1000 G. This residual flux is referred to as *remanence*.

In a forward converter, if the core started at 0 Oe and hence at 1000 G, the maximum flux change dB possible before the core is driven up into the curved part of the hysteresis loop is about 1000 G. It is desirable to stay off the curved part of the hysteresis loop, and hence the forward converter core with no air gap is restricted to a maximum dB of 1000 G. As shown above, the number of primary turns is inversely proportional to dB. Such a relatively small dB requires a relatively large number of primary turns. A larger number of primary turns requires smaller wire size and hence decreases the current and power available from the transformer.

By introducing an air gap in the core, the hysteresis loop is tilted as shown in Fig. 2.5 and magnetic remanence is reduced significantly.

The hysteresis loop tilts over and still crosses the zero flux density at the same point on the H axis (referred to as *coercive force*. Coercive force for ferrites is seen to be about 0.2 Oe in Fig. 2.3. An air gap of about 2 to 4 mils will reduce remanence to about 200 G for most cores used at a 200- to 500-W output power level. With remanence of 200 G, the permissible dB before the core enters the curved part of the hysteresis loop is now about 1800 G and fewer turns are permissible.

However, there is a penalty paid in introducing an air gap. Figure 2.5 shows the slope of the hysteresis loop tilted over. The slope is dB/dH or the core permeability, which has been decreased by adding the gap. Decreasing permeability has decreased the magnetizing inductance and increased the magnetizing current ($I_m = V_{dc}T_{on}/L_m$). Magnetizing current contributes no output power to the load; it simply moves the core across the hysteresis loop and is wasteful if it exceeds 10 percent of the primary load current.

2.3.9.3 Magnetizing inductance with gapped core. Magnetizing inductance with a gapped core can be calculated as follows. Voltage across the magnetizing inductance is $L_m\, dI_m/dt$ and from Faraday's law:

$$V_{dc} = \frac{L_m dI_m}{dt} = \frac{N_p A_e\, dB}{dt}\, 10^{-8} \quad \text{or} \quad L_m = \frac{N_p A_e\, dB}{dI_m}\, 10^{-8} \tag{2.35}$$

where L_m = magnetizing inductance, H
N_p = number of primary turns
A_e = core area, cm^2
dB = core flux change, G
dI_m = change in magnetizing current, A

Now a fundamental law in magnetics is Ampere's law:

$$\int H \cdot dl = 0.4\pi NI$$

This states that if a line is drawn encircling a number of ampere turns NI, the dot product $H \cdot dl$ along that line is equal to $0.4\pi NI$. If the line is taken through the core parallel to the magnetic flux lines and across the gap, since H is uniform at a level H_i within the core and uniform at a value H_a across the gap, then

$$H_i l_i + H_a l_a = 0.4\pi N I_m \tag{2.36}$$

where H_i = magnetic field intensity in iron (ferrite), Oe
l_i = length of iron path, cm
H_a = magnetic field intensity in air gap, Oe
l_a = length of the air gap, cm
I_m = magnetizing current, A

However, $H_i = B_i/u$, where B_i is the magnetic flux density in iron and u is the iron permeability; $H_a = B_a$ as the permeability of air is 1; and $B_a = B_i$(flux density in iron = flux density in air) if there is no fringing flux around the air gap. Then Eq. 2.36 can be written as

$$\frac{B_i}{u} l_i + B_i l_a = 0.4\pi N_p I_m \qquad \text{or} \qquad B_i = \frac{0.4\pi N I_m}{l_a + l_i/u} \tag{2.37}$$

Then $dB/dI_m = 0.4\pi N/(l_a + l_i/u)$, and substituting this into Eq. 2.35

$$L_m = \frac{0.4\pi (N_p)^2 A_e \times 10^{-9}}{l_a + l_i/u} \tag{2.38}$$

Thus, introducing an air gap of length l_a to a core of iron path length l_i reduces the magnetizing inductance in the ratio of

$$\frac{L_{m\,(\text{with gap})}}{L_{m\,(\text{without gap})}} = \frac{l_i/u}{l_a + l_i/u} \tag{2.39}$$

It is instructive to consider a specific example. Take an international standard core such as the Ferroxcube 783E608-3C8. It has a magnetic path length of 9.7 cm and an effective permeability of 2300. Then if a 4-mil (= 0.0102-cm) gap were introduced into the magnetic path, from Eq. 2.39

$$L_{m\,(\text{with gap})} = \frac{9.7/2300}{0.0102 + 9.7/2300} L_{m\,(\text{without gap})}$$

$$= 0.29 L_{m\,(\text{without gap})}$$

A useful way of looking at a gapped core is to examine the denominator in Eq. 2.38. In most cases, u is so high that the term l_i/u is small compared to the air gap l_a and the inductance is determined primarily by the length of the air gap.

2.3.10 Power transformer design relations

2.3.10.1 Core selection. As discussed in Sec. 2.2.9.1 on core selection for a push-pull transformer, the amount of power available from a core for a forward converter transformer is related to the same parameters—peak flux density, core iron and bobbin areas, frequency, and coil current density in circular mils per rms ampere.

In Chap. 7, an equation will be derived giving the amount of available output power as a function of these parameters. This equation will be converted to a chart which permits selection of core size and operating frequency at a glance.

For the present, it is assumed that a core has been selected and its iron and bobbin area are known.

2.3.10.2 Primary turns calculation. The number of primary turns is calculated from Faraday's law as given in Eq. 2.7. Recall (Sec. 2.3.9.2) that in the forward converter, with a gapped core, flux moves from about 200 G to some higher value B_{max}. As in the push-pull topology, this peak value will be set at 1600 G for ferrites even at low frequencies where core losses are not a limiting factor.

This, as discussed in Sec. 2.2.9.4, is to avoid the problem of a much larger and dangerous flux swing resulting from rapid changes in DC input voltage or load current. Such rapid changes are not immediately corrected for because the limited error-amplifier bandwidth does not permit a rapid correction in the power transistor on time.

During the error-amplifier delay, the peak flux density can, for a number of cycles, go up to as high as 50 percent above the calculated for normal steady-state operation. This can be tolerated if the normal peak flux density in the absence of a line or load transient is set to a low value of 1600 G. As discussed above, the excursion from approximately zero to 1600 G will take place in 80 percent of a half period to ensure that the core can be reset before the start of the next period (Fig. 2.12*b*).

Thus the number of primary turns is set by Faraday's law at

$$N_p = \frac{(\underline{V_{dc}} - 1)(0.8T/2) \times 10^{+8}}{A_e\, dB} \tag{2.40}$$

where V_{dc} = minimum DC input, V
T = operating period, s
A_e = iron area, cm^2
dB = 1600 G

2.3.10.3 Secondary turns calculation. Secondary turns are calculated from Eqs. 2.25 to 2.27. In those relations, all values except the secondary turns are specified or already calculated. Thus (Fig. 2.10):

V_{dc} = minimum DC input, V
T_{on} = maximum on time, s (= $0.8T/2$)
N_m, N_{s1}, N_{s2} = numbers of main and slave turns
N_p = number of primary turns
V_d = rectifier forward drop

Usually the main output is a high-current 5-V one for which a Schottky diode of 0.5 V is used. The slaves usually have higher output voltages which require the use of diodes with higher reverse-voltage

ratings. Such fast-recovery higher-reverse-voltage diodes have forward drops of 1.0 V over a large range of currents.

2.3.10.4 Primary rms current and wire size selection. Primary equivalent flat-topped current is given by Eq. 2.28. That current flows a maximum of 80 percent of a half period out of each full period, and hence its maximum duty cycle is 0.4. Recalling that the rms value of a flat-topped pulse of amplitude I_p is $I_{rms} = I_p\sqrt{T_{on}/T}$, the rms primary current is

$$I_{\text{rms (primary)}} = \frac{3.12P_o}{V_{dc}}\sqrt{0.4}$$

$$= \frac{1.97P_o}{V_{dc}} \tag{2.41}$$

If the wire size is chosen on the basis of 500 circular mils per rms ampere, the required number of circular mils is

$$\text{Circular mils needed} = \frac{500 \times 1.97P_o}{V_{dc}}$$

$$= \frac{985P_o}{V_{dc}} \tag{2.42}$$

2.3.10.5 Secondary rms current and wire size selection. It is seen in Fig. 2.11 that the secondary current has the characteristic shape of a ramp on a step. The pulse amplitude at the center of the ramp is equal to the DC output current. Thus the equivalent flat-topped secondary current pulse at V_{dc} (when its width is a maximum) has an amplitude of I_{dc}, a width of $0.8T/2$, and a duty cycle of $(0.8T/2)/T$ or 0.4. Then

$$I_{\text{rms (secondary)}} = I_{dc}\sqrt{0.4}$$

$$= 0.632I_{dc} \tag{2.43}$$

and at 500 circular mils per rms ampere, the required number of circular mils for each secondary is

$$\text{Circular mils needed} = 500 \times 0.632I_{dc}$$

$$= 316I_{dc} \tag{2.44}$$

2.3.10.6 Reset winding rms current and wire size selection. The reset winding carries only magnetizing current, as can be seen by the dots

in Fig. 2.10. When $Q1$ is on, diode $D1$ is reverse-biased and no current flows in the reset winding. But magnetizing current builds up linearly in the power winding N_p. When $Q1$ turns off, that magnetizing current must continue to flow. When $Q1$ current ceases, the current in the magnetizing inductance reverses all winding polarities. When the dot end of N_r reaches ground, the magnetizing current transfers from N_p to N_r and continues flowing through the DC input voltage source V_{dc} out from the negative end of V_{dc} through $D1$ and back into N_r. Since the no-dot end of N_r is positive with respect to the dot end, the magnetizing current ramps downward to zero as seen in Fig. 2.10.

The waveshape of this N_r current is the same as that of the magnetizing current which ramped upward when $Q1$ was on, but it is reversed from left to right. Thus the peak of this triangle of current is $I_{p\ (\text{magnetizing})} = V_{dc}\overline{T_{on}}/L_{mg}$, where L_{mg} is the magnetizing inductance with an air gap as calculated from Eq. 2.39. The inductance without the gap is calculated from the ferrite catalog value of A_l, the inductance per 1000 turns. Since inductance is proportional to the square of the number of turns, inductance for n turns is $L_n = A_l(n/1000)^2$. The duration of this current triangle is $0.8T/2$ (the time required for the core to reset), and it comes at a duty cycle of 0.4.

It is known that the rms value of a repeating triangle (no spacing between successive triangles) of peak amplitude I_p is $I_{rms} = I_p/\sqrt{3}$. But this triangle comes at a duty cycle of 0.4, and hence its rms value is

$$I_{rms} = \frac{V_{dc}\overline{T_{on}}}{L_{mg}}\frac{\sqrt{0.4}}{\sqrt{3}}$$

$$= 0.365\,\frac{V_{dc}\overline{T_{on}}}{L_{mg}}$$

and at 500 circular mils per rms ampere, the required number of circular mils for the reset winding is

$$\text{Circular mils required} = 500 \times 0.365\,\frac{V_{dc}\overline{T_{on}}}{L_{mg}} \tag{2.45}$$

Most frequently, magnetizing current is so small that the reset winding can be done with wire sizes smaller than No. 30 wire (AWG).

2.3.11 Output filter design relations

The output filters $L1C1$, $L2C2$, and $L3C3$ average the peak voltage at the rectifier cathodes. The inductor is selected so as not to go discontinuous (Sec. 1.3.6) at the minimum DC output current. The capacitor is selected to yield a minimum specified output ripple voltage.

2.3.11.1 Output inductor design. It will be recalled from Sec. 1.3.6 that the discontinuous mode condition occurs when the front end of the inductor current ramp drops to zero (Fig. 2.10). Since the DC output current is the value at the center of the ramp, discontinuous mode starts at a minimum current I_{dc} equal to half the ramp amplitude dI as can be seen in Fig. 2.10.

Now referring to Fig. 2.11:

$$dI = 2\underline{I_{dc}} = \frac{(\underline{V_{rk}} - V_o)\overline{T_{on}}}{L1} \quad \text{or} \quad L1 = \frac{(\underline{V_{rk}} - V_o)(\overline{T_{on}})}{2\underline{I_{dc}}}$$

But $V_o = \underline{V_{rk}}\overline{T_{on}}/T$. Then

$$L1 = \left(\frac{V_o T}{\overline{T_{on}}} - V_o\right)\frac{\overline{T_{on}}}{2\underline{I_{dc}}}$$

$$= \frac{V_o(T/\overline{T_{on}} - 1)\overline{T_{on}}}{2\underline{I_{dc}}}$$

But $\overline{T_{on}} = 0.8T/2$. Then

$$L1 = \frac{0.3V_o T}{\underline{I_{dc}}} \tag{2.46}$$

and if minimum DC current $\underline{I_{dc}}$ is one-tenth the nominal output current I_{on}, then

$$L1 = \frac{3V_o T}{I_{on}} \tag{2.47}$$

2.3.11.2 Output capacitor design. It was seen in Sec. 1.3.7 that the output ripple is almost completely determined by the equivalent series resistance R_o of the filter capacitor. The peak-to-peak ripple amplitude V_{or} is $V_{or} = R_o\, dI$, where dI is the peak-to-peak ripple current amplitude chosen by the selection of the ripple inductor as discussed above. Then assuming as in Sec. 1.3.7 that the average value of $R_o C_o$ for aluminum electrolytic capacitors of a large range of voltage and capacitance ratings is given by $R_o C_o = 65 \times 10^{-6}$:

$$C_o = 65 \times 10^{-6}/R_o$$

$$= 65 \times 10^{-6}\frac{dI}{V_{or}} \tag{2.48}$$

where dI is in amperes and V_{or} is in volts for C_o in farads.

2.4 Double-Ended Forward Converter Topology

2.4.1 Basic operation

Double-ended forward converter topology is shown in Fig. 2.13. Although it has two transistors rather than one compared to the single-ended forward converter of Fig. 2.10, it has a very significant advantage. In the off state, both transistors are subjected to only the DC input voltage rather than twice that as in the single-ended converter. Further, at turnoff, there is no leakage inductance spike.

It was pointed out in Sec. 2.3.7 that the off-voltage stress in the single-ended forward converter operating off a nominal 120-V-AC line can be as high as 550 V when there is a 15 percent transient above a 10 percent steady-state high line and a 30 percent leakage spike.

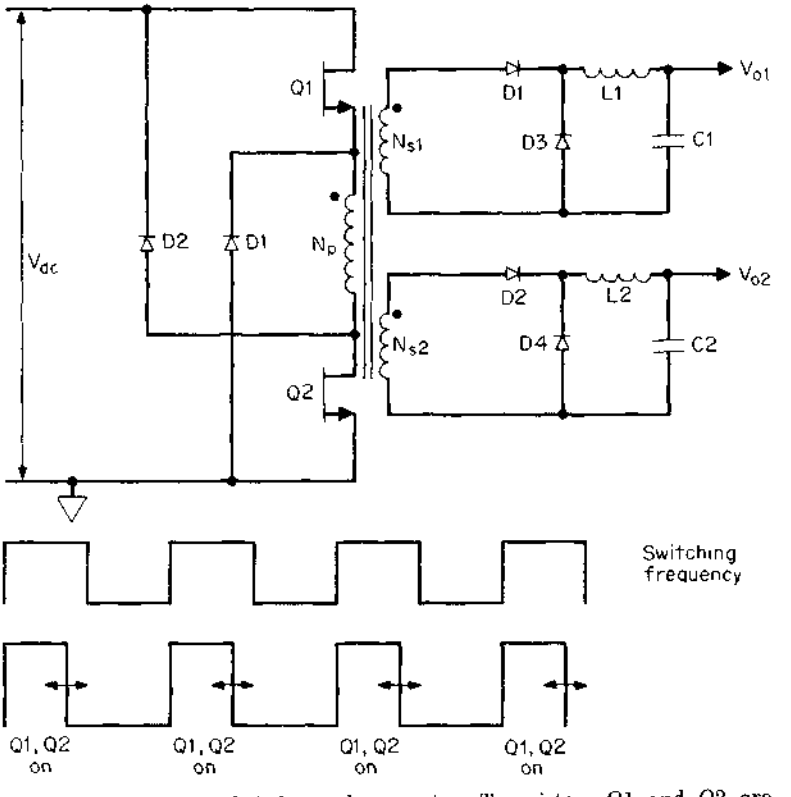


Figure 2.13 Double-ended forward converter. Transistors $Q1$ and $Q2$ are turned on and off simultaneously. Diodes $D1$ and $D2$ keep the maximum off voltage stress on $Q1$, $Q2$ at a maximum of V_{dc} as contrasted with $2V_{dc}$ plus a leakage spike for the single-ended forward converter of Fig. 2.10.

Although there are a number of bipolar transistors with V_{cev} ratings up to 650 and even 850 V which can take that stress, it is a far more reliable design to use the double-ended forward converter with half the off-voltage stress. Reliability is of overriding importance in a power supply design, and in any weighing of reliability versus initial cost, the best and—in the long run—least expensive choice is reliability.

Further, for power supplies to be used in the European market where the AC voltage is 220 V (rectified DC voltage is nominally about 308 V), the single-ended forward converter is not usable at all because of the excessive voltage stress on the off transistor (Eq. 2.29). The double-ended forward converter and the half bridge and the full bridge (to be discussed in Chap. 3) are the only choices for equipment to be used in the European market.

The double-ended forward converter works as follows. In Fig. 2.13, $Q1$ and $Q2$ are in series with the top and bottom of the transformer primary. Both of these transistors are turned on simultaneously and turned off simultaneously. When they are on, all primary and secondary dot ends are positive and power is delivered to the loads. When they turn off, current stored in the $T1$ magnetizing inductance reverses polarity of all windings. The dot end of N_p tries to go far negative but is caught at ground by diode $D1$. The no-dot end of N_p tries to go far positive but is caught at V_{dc} by diode $D2$.

Thus the emitter of $Q1$ can never be more than V_{dc} below its collector, and the collector of $Q2$ can never be more than V_{dc} above its emitter. Leakage inductance spikes are clamped so that the maximum voltage stress on either transistor can never be more than the maximum DC input voltage.

The further significant advantage is that there is no leakage inductance energy to be dissipated. Any energy stored in the leakage inductance is not lost by dissipation in some resistive element or in the power transistors. Instead, energy stored in the leakage inductance during the on time is fed back into V_{dc} via $D1$ and $D2$ when the transistors turn off. The leakage inductance current flows out of the no-dot end of N_p, through $D2$, into the positive end of V_{dc}, out of its negative end, and up through $D1$ back into the dot end of N_p.

Examination of Fig. 2.13 reveals that the core is always reset with the reset time equal to the on time. For the reverse polarity voltage across N_p when the transistors are off is equal to the forward polarity voltage across it when the transistors are on. Thus the core will always succeed in being fully reset with a 20 percent safety margin before the start of a next half cycle if the maximum on time is never required to be greater than 80 percent of a half period. This is done by

choosing enough secondary turns so that the peak secondary voltage at minimum V_{dc} times the maximum duty cycle of 0.4 equals the desired output voltage (Eq. 2.25).

2.4.1.1 Practical output power limits. It should, however, be noted that this topology still yields only one power pulse per period just as the single-ended forward converter. Thus the power available from a specific core is pretty much the same regardless of whether the single- or double-ended configuration is used. As noted in Sec. 2.3.10.6, the reset winding in the single-ended circuit carries only magnetizing current during the power transistor off time. Since that current is small, the reset winding can most often be done with No. 30 wire or narrower. Thus the absence of a reset winding in the double-ended circuit does not permit significantly larger power winding wire size and any more output power from a given core.

But because the maximum off transistor voltage stress cannot be greater than the maximum DC input voltage, the 200-W practical power limit for the single-ended forward converter discussed in Sec. 2.3.7 no longer holds. With the reduced voltage stress offered by the double-ended forward converter, output powers of 400 to 500 W are obtainable and transistors with the required voltage and current capability are easily available at adequate gain and low price.

Thus consider a double-ended forward converter operating off a nominal 120-V-AC line with ±10 percent steady-state tolerance and a ±15 percent tolerance on top of that. Maximum rectified DC voltage is $1.41 \times 120 \times 1.1 \times 1.15 = 214$ V, minimum rectified DC voltage is $1.41 \times 120 \times 0.9 \times 0.85 = 130$ V, and equivalent flat-topped primary current from Eq. 2.28 is $I_{pft} = 3.13P_o/V_{dc}$, and for $P_o = 400$ W, $I_{pft} = 9.6$ A.

Many transistors (bipolar and MOSFET) with a 214-V, 9.6-A rating are easily and inexpensively available with adequately high gain.

A better alternative using the double-ended forward converter would be to use a voltage doubler off the 120-V-AC line (Fig. 3.1). This would double the voltage stress to 428 V and half the peak current to 4.8 A. With 4.8 A of primary current, RFI problems would be less severe. A bipolar transistor with a 400-V V_{ceo} rating could easily take 428 V if it had a −1- to −5-V reverse bias at the instant of turnoff (V_{cev} rating).

2.4.2 Design relations and transformer design

2.4.2.1 Core selection—primary turns and wire size. The transformer design for the double-ended forward converter proceeds exactly as for the single-ended converter. A core is selected from the aforementioned

selection charts (to be presented in Chap. 7 on magnetics) for the required output power and operating frequency.

The number of primary turns is chosen from Faraday's law as in Eq. 2.40. There the minimum primary voltage is $(V_{dc} - 2)$ as there are two transistors rather than one in series with the primary—but that is insignificant as V_{dc} is usually 130 V (120 V AC). Maximum on time is taken as $0.8T/2$ and dB is taken as 1600 G up to about 50 kHz or even to a higher frequency if not limited by core losses.

As mentioned, in the region of 100 to 300 kHz, peak flux density may have to be lowered down to about 1400 or possibly to 800 G as core losses increase with frequency. But the exact peak flux density chosen depends on whether the newer, lower-loss materials are available. It also depends to some extent on transformer size—smaller cores can generally operate at higher flux density as they have a larger ratio of radiating surface area to volume and hence can more easily get rid of the heat they generate (which is proportional to volume).

Since there is only one current or power pulse per period as in the single-ended forward converter, the primary current for a given output power and minimum DC input voltage is given by Eq. 2.28 and the primary wire size is chosen from Eq. 2.42.

2.4.2.2 Secondary turns and wire size. Secondary turns are chosen exactly as in Secs. 2.3.2 and 2.3.3 from Eqs. 2.25 to 2.27. Wire sizes are calculated as in Sec. 2.3.10.5 from Eq. 2.44.

2.4.2.3 Output filter design. The output inductor and capacitor magnitudes are calculated exactly as in Sec. 2.3.11 from Eqs. 2.46 to 2.48.

2.5 Interleaved Forward Converter Topology

2.5.1 Basic operation—merits, drawbacks, and output power limits

This topology is simply two identical single-ended forward converters operating on alternate half cycles with their secondary currents adding through rectifying OR diodes. The topology is shown in Fig. 2.14.

The advantage, of course, is that now there are two power pulses per period as seen in Fig. 2.14 and each converter supplies half the total output power.

Equivalent flat-topped peak transistor current is given by Eq. 2.28 as $I_{pft} = 3.13\, P_{ot}/2V_{dc}$ where P_{ot} is the total output power. This is half the transistor current if the same total output power were being produced with only a single forward converter. Thus the expense of two transistors is somewhat balanced by the fact that each has a lower

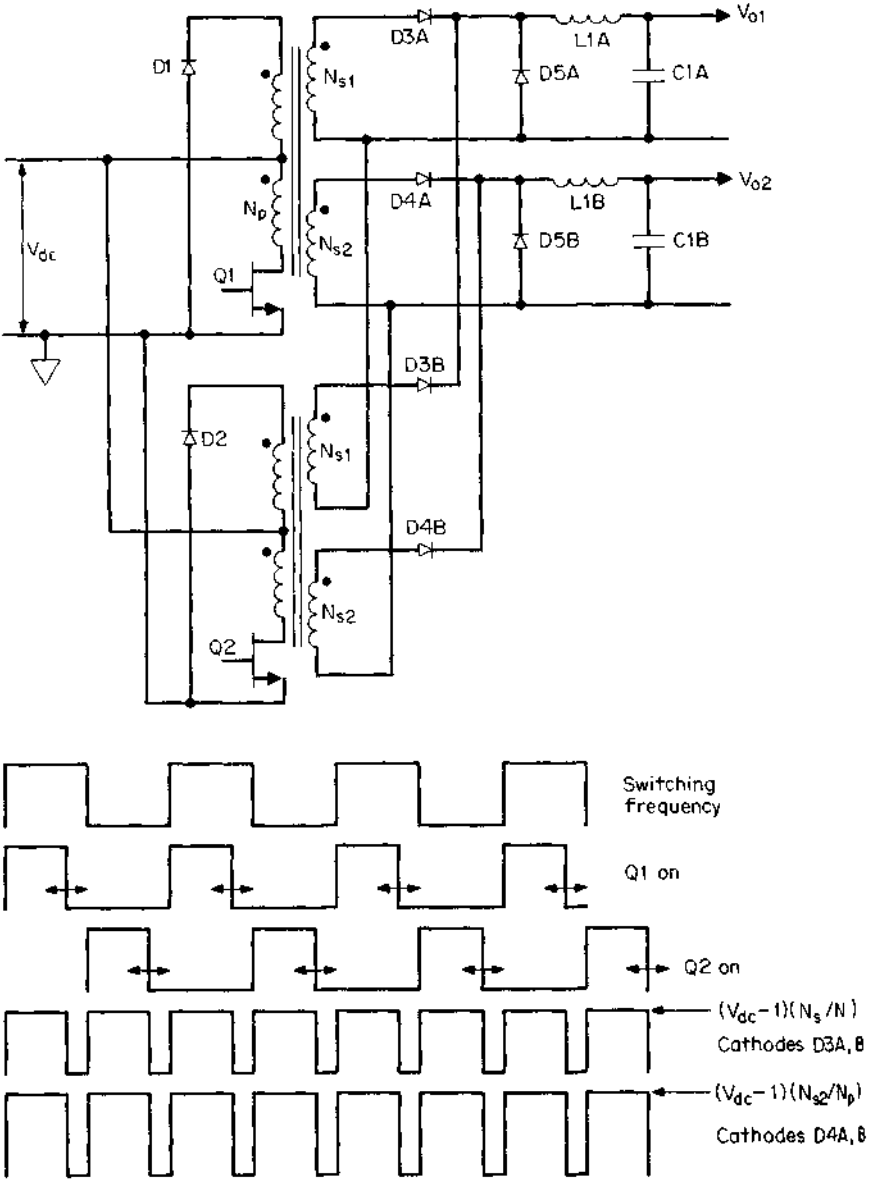


Figure 2.14 Interleaved forward converter. Interleaving the on time of $Q1$ and $Q2$ on alternate half cycles and summing their secondary output gives two power pulses per period but avoids the flux imbalance problem of the push-pull topology.

peak current rating and may be less expensive than the single one at twice the current rating.

Or looking at it another way, if two transistors of the same current rating were used and at the same peak current as one single-ended converter which had been designed for a given output power, the interleaved converter would yield twice the output power of the single converter.

Also, since the intensity of EMI generated is proportional to the peak currents, not to the number of current pulses, an interleaved converter of the same total output power as a single forward converter will generate less EMI.

The two transformers are each smaller than one for a single forward converter of the same output power. But the two together will most likely occupy more volume and cost more than the single transformer.

Thus, although there are some pluses and some minuses, a major advantage of the interleaved forward converters remains that it yields twice the output power of a single, single-ended forward converter operated at the same peak current in the transistor.

If this topology is compared to a push-pull, it might be thought that the push-pull is preferable. Although both are two-transistor circuits, the two transformers in the interleaved forward converter are most probably more expensive and occupy more space than a single large one in a push-pull circuit. But there is the ever-present uncertainty that the flux imbalance problem in the push-pull has not disappeared under odd transient line and load conditions.

The certainty that there is no flux imbalance problem in the interleaved forward converter is probably the best argument for its use.

There is one special, although not frequent case, where the interleaved forward converter is a much more desirable choice than a single forward converter of the same output power. This occurs when a DC output voltage is high—over about 200 V. In a single forward converter, the peak reverse voltage experienced by the output freewheeling diodes (*D5A* or *D5B*) is twice that for an interleaved forward converter as the duty cycle in the latter is twice that in the former.

This is no problem when output voltages are low, as can be seen in Eq. 2.25. Transformer secondary turns are always selected (for the single forward converter) so that at minimum DC input, when the secondary voltage is at its minimum, the duty cycle T_{on}/T need not be more than 0.4 to yield the desired output voltage. Then for a DC output of 200 V, the peak reverse voltage experienced by the freewheeling diode is 500 V. At the instant of power transistor turnon, the free-wheeling diode has been carrying a large forward current and will suddenly be subjected to reverse voltage. If the diode has slow re-

verse recovery time, it will draw a large reverse current for a short time at 500-V reverse voltage and run dangerously hot.

Generally, diodes with larger reverse voltage ratings have slower recovery times and can be a serious problem in such a case. The interleaved forward converter runs at twice the duty cycle and for a 200-V-DC output, subjects the free-wheeling diode to only 250 V. This permits a lower voltage, faster-recovery diode and reduces its dissipation considerably.

2.5.2 Transformer design relations

2.5.2.1 Core selection. The core will be selected from the aforementioned charts, to be presented in Chap. 7. But they will be chosen for half the total power output as each transformer supplies half the output power.

2.5.2.2 Primary turns and wire size. The number of primary turns in the interleaved forward converter is still given by Eq. 2.40 as each converter's on time will still be $0.8T/2$ at minimum DC input. The core iron area A_e will be read from the catalogs for the selected core. Primary wire size will be chosen from Eq. 2.42 at half the total output power.

2.5.2.3 Secondary turns and wire size. The number of secondary turns will be chosen from Eqs. 2.26 and 2.27, but therein the duty cycle will be 0.8 as there are two voltage pulses, each of duration $0.8T/2$ at V_{dc}. Wire size will still be chosen from Eq. 2.44 where I_{dc} is the actual DC output current as each secondary carries that current at a maximum duty cycle of 0.4.

2.5.3 Output filter design

2.5.3.1 Output inductor design. The output inductor sees two current pulses per period exactly like the output inductor in the push-pull topology. These pulses have the same width, amplitude, and duty cycle as the push-pull inductor at the same DC output current. Hence the magnitude of the inductance is calculated from Eq. 2.20 as for the push-pull inductor.

2.5.3.2 Output capacitor design. Similarly, the output capacitor "doesn't know" whether it is filtering a full-wave push-pull waveform from a push-pull topology secondary or from an interleaved forward converter. Thus for the same inductor current ramp amplitude and permissible output ripple as the push-pull circuit, the capacitor is selected from Eq. 2.22.

Chapter

3

Half- and Full-Bridge Converter Topologies

3.1 Introduction

Half- and full-bridge topologies subject their transistors in the off state to a voltage stress equal to the DC input voltage and not to twice that as do the push-pull, single-ended, and interleaved forward converter topologies. Thus the bridge topologies are used mainly in off-line converters where twice the rectified DC would be more than the usual switching transistors could safely tolerate. Bridge topologies hence are almost always used where the nominal AC input voltage is 220 V or higher and frequently for 120 V AC.

An additional valuable feature of the bridge topologies is that primary leakage inductance spikes (Figs. 2.1, 2.10) are easily clamped to the DC supply bus and any energy stored in the leakage inductance is returned to the input bus instead of having to be dissipated in some resistive element.

3.2 Half-Bridge Converter Topology

3.2.1 Basic operation

Half-bridge converter topology is shown in Fig. 3.1. Its major advantage is that, just as the double-ended forward converter, it subjects the off transistor to only V_{dc} and not twice that as do the push-pull and singled forward converter. It is thus widely used in equipment intended for the European market, where the AC input voltage is 220 V.

First consider the input rectifier and filter in Fig. 3.1. It is the universally used scheme when equipment is to be used with minimum changes either on 120-V-AC American or 220-V-AC European power. The circuit always yields roughly 320 V of rectified DC voltage

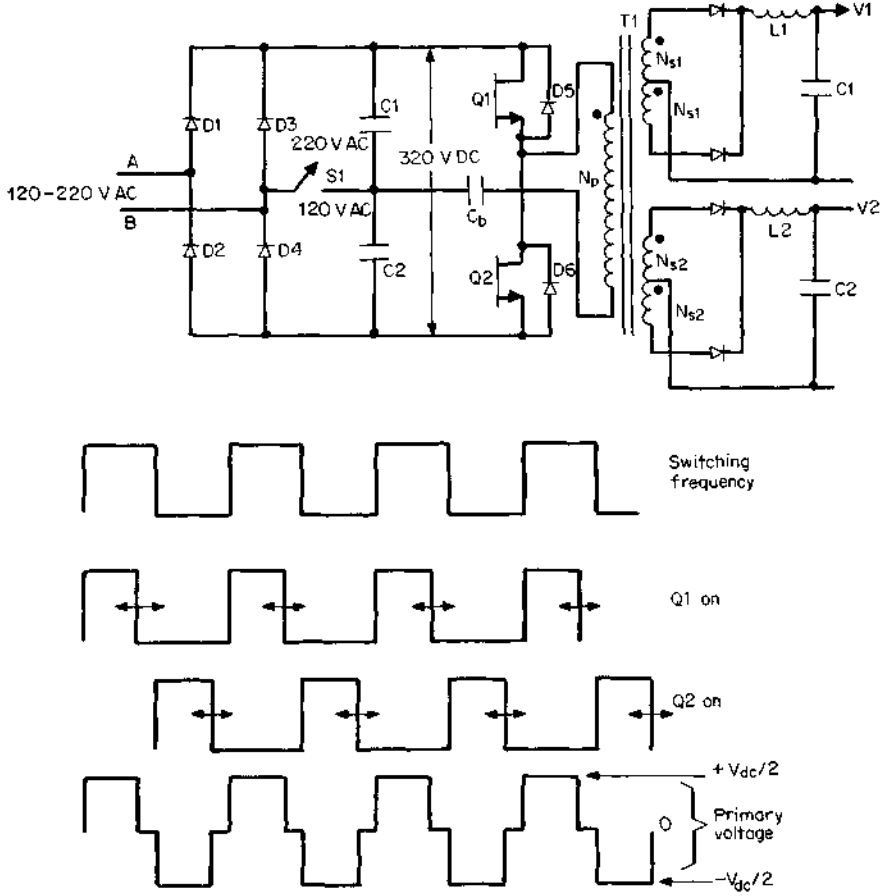


Figure 3.1 Half-bridge converter. One end of the power transformer is connected to the junction of filter capacitors C1, C2 via a small DC blocking capacitor C_b. The other end is connected to the junction of Q1, Q2 which turn on and off on alternate half cycles. With S1 in the lower position, the circuit is a voltage doubler; in the upper position, it is a full wave rectifier. Rectified output is thus always about 308 to 336 VDC.

whether input is 120 or 220 V AC. It does this by setting switch S1 to the upper position for 220-V-AC input or to the lower position for 120-V-AC input. The S1 component is not actually a switch; it is a linkage which is either made or not made to the lower position depending on whether the input is 120 or 220 V AC.

With the switch in the upper, 220-V-AC position, the circuit is a full-wave rectifier with filter capacitors C1 and C2 in series. It produces a peak rectified DC voltage of about (1.41 × 220) − 2 or 308 V.

When the switch is in the lower 120-V-AC position, the circuit acts as a voltage doubler. On the first half cycle of the input voltage, when, say, A is positive relative to B, the top end of $C1$ is charged positively via $D1$ to a peak of $(1.41 \times 120) - 1$ or to 168 V. On the next negative half cycle, when A is negative with respect to B, the top end of capacitor $C2$ is charged positively via $D2$ to 168 V. The total output across $C1$ and $C2$ in series is then 336 V. It can be seen in Fig. 3.1 that with either transistor on, the off transistor is subjected only to the maximum DC input voltage and not twice that.

Now since the topology subjects the off transistor to only V_{dc} and not $2V_{dc}$, there are very many bipolar and MOSFET inexpensive transistors which can tolerate a nominal 336 DC V or even its maximum—a 15 percent high value of 386 V. Thus the equipment can be used off either 120 or 220 V AC with a simple switch or linkage change.

Assuming a nominal rectified DC voltage of 336 V, the topology works as follows. Ignore, for the moment, the small series blocking capacitor C_b. The bottom end of N_p can thus be assumed connected to the junction of $C1$ and $C2$. To a close approximation, if leakages in $C1$, $C2$ are assumed equal, that point will be at half the rectified DC voltage or about 168 V. It is generally good practice to place equal bleeder resistors across $C1$ and $C2$ to equalize their drop. Now $Q1$ and $Q2$ are switched on at alternate half cycles. When $Q1$ is on and $Q2$ off (Fig. 3.1), the dot end of N_p is 168 V positive with respect to its no-dot end and the off stress on $Q2$ is only 336 V. When $Q2$ is on and $Q1$ off, the dot end of N_p is 168 V negative with respect to its no-dot end and the emitter of $Q1$ is 336 V negative with respect to its collector.

This AC square-wave primary voltage produces full-wave square waveshapes on all secondaries—exactly like the secondary voltages in the push-pull topology. The selection of secondary voltages, wire sizes, and the output inductor and capacitor proceeds exactly as for the push-pull circuit.

3.2.2 Half-bridge magnetics

3.2.2.1 Maximum on time, magnetic core, primary turns selection. It can be seen in Fig. 3.1, that if $Q1$ and $Q2$ are simultaneously on—even for a very short time—there is a dead short circuit across the supply voltage and the transistors will be destroyed. The maximum $Q1$ or $Q2$ on time, which occurs at minimum DC supply voltage, will then be set at 80 percent of a half period to ensure that this does not happen. Secondary turns will be chosen so that the desired output voltages are obtained at an on time of no more than $0.8T/2$. An on-time clamp will be

provided to ensure that under fault conditions, on time can never be greater than $0.8T/2$.

The core is selected from the earlier mentioned tables (to be presented in Chap. 7). These tables give maximum available output power as a function of operating frequency, peak flux density, core and iron area, and coil current density.

With a core selected and its iron area known, the number of primary turns is calculated from Faraday's law (Eq. 1.17) for the minimum primary voltage ($V_{dc}/2$) − 1, the maximum on time of $0.8T/2$. The value of dB in Faraday's law is taken as twice the desired peak flux density (1600 G below 50 kHz or less at higher frequency). The flux excursion dB is twice the desired peak value as the half-bridge core operates in the first and third quadrants of its hysteresis loop—unlike the forward converter, which (Sec. 2.3.9) operates in the first quadrant only.

3.2.2.2 Relation between primary current, output power, and input voltage. Assume an efficiency of 80 percent. Then

$$\text{Input power} = P_{in} = 1.25P_o$$

But the input power at minimum supply voltage is the minimum primary voltage times the average primary current at minimum DC input. At minimum DC input, on time in each half period will be a maximum and will be selected as $0.8T/2$ as discussed above. Since at minimum DC input, the primary has two current pulses per period T, of width $0.8T/2$, and at a voltage $V_{dc}/2$, the input power is $1.25P_o = (V_{dc}/2)(I_{pft})(0.8T/T)$, where I_{pft} is the peak equivalent flat-topped primary current pulse. Then

$$I_{pft(\text{half bridge})} = \frac{3.13P_o}{V_{dc}} \tag{3.1}$$

3.2.2.3 Primary wire size selection. Primary wire size in the half bridge is much larger than in a push-pull circuit of the same output power. But there are two half primaries for the push-pull, each of which has to support twice the voltage of a half bridge operated from the same supply voltage. Consequently, coil sizes for the two topologies are not much different. Half-bridge primary rms current is $I_{rms} = I_{pft}\sqrt{0.8T/T}$, and from Eq. 3.1

$$I_{rms} = \frac{2.79P_o}{V_{dc}} \tag{3.2}$$

At 500 circular mils per rms ampere, the required number of circular mils is

$$\text{Circular mils needed} = \frac{500 \times 2.79P_o}{V_{dc}}$$

$$= \frac{1395P_o}{V_{dc}} \tag{3.3}$$

3.2.2.4 Secondary turns and wire size selection. The number of secondary turns will be selected from Eqs. 2.1 to 2.3 for $\overline{T_{on}} = 0.8T/2$ and the term $V_{dc} - 1$ replaced by the minimum primary voltage, which is $(V_{dc}/2) - 1$.

The secondary rms currents and wire sizes are calculated from Eqs. 2.13 and 2.14 exactly as for the full-wave secondaries of a push-pull circuit.

3.2.3 Output filter calculations

The output inductor and capacitor are selected from Eqs. 2.20 and 2.22 for the same inductor current ramp amplitude and desired output ripple voltage as in a push-pull circuit.

3.2.4 Blocking capacitor to avoid flux imbalance

In Fig. 3.1, a small capacitor C_b is seen in series with the primary. Its purpose is to avoid the flux-imbalance problem discussed in connection with the push-pull circuit (Sec. 2.2.5). Recall that flux imbalance occurs if the volt-second product across the primary when the core was set (moved in one direction along the hysteresis loop) is different from the volt-second product when it had moved in the opposite direction.

Thus, if the junction of $C1$ and $C2$ is not exactly at half the supply voltage, the voltage across the primary when $Q1$ is on will differ from the voltage across it when $Q2$ is on. The core can then walk up or down the hysteresis loop into saturation and destroy the transistors.

The preceding description of the cause of flux imbalance is simply another way of stating that it comes about because there is a DC current bias in the primary. To avoid the DC current bias, the small DC blocking capacitor is placed in series in the primary. Selection of the magnitude of the capacitor is done as follows. If the capacitor is made too small, it will charge up as the primary current I_{pft} flows into it.

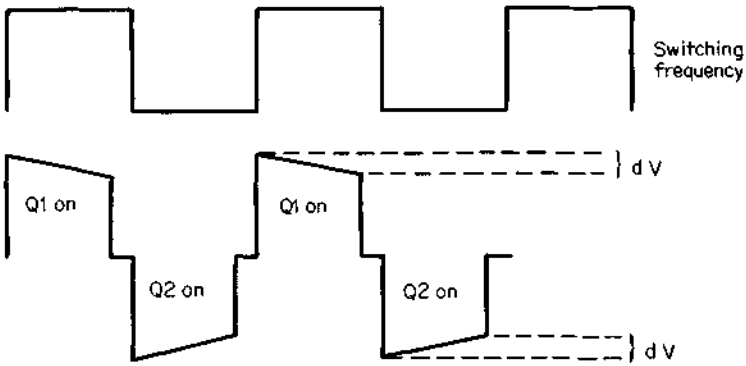


Figure 3.2 The small blocking capacitor C_b in series with the half-bridge primary (Fig. 3.1) is needed to prevent flux imbalance if the junction of the filter capacitors is not exactly at the midpoint of the supply voltage. Primary current charges the capacitor, causing a droop in the primary voltage waveform. Droop should be kept to no more than 10 percent. (dV is permissible droop in primary voltage due to charging of blocking capacitor.)

Whatever voltage it charges up to robs voltage from the flat-topped primary pulse shown in Fig. 3.2.

This robs volt-seconds from all secondary windings and forces a longer on time to achieve the desired output voltage. In general, it is desirable to keep the primary voltage pulses as flat-topped as possible.

Assume a permissible droop dV. The equivalent flat-topped current pulse which causes that droop is I_{pft} of Eq. 3.1. Then since that current flows for $0.8T/2$, the required capacitor magnitude is simply

$$C_b = \frac{I_{\text{pft}} \times 0.8T/2}{dV} \tag{3.4}$$

It is instructive to consider an example. Assume a 150-W half bridge operating at 100 kHz from a nominal DC input of 320 V. At 15 percent low line, it is 272 V and the primary voltage is $\pm 272/2$ or ± 136 V.

A tolerable droop in the flat-topped primary voltage pulse is about 10 percent or about 14 V. For 150 W and V_{dc} of 272 V, I_{pft} is, from Eq. 3.1, $I_{\text{pft}} = 3.13 \times 150/272 = 1.73$ A. Then from Eq. 3.4, $C_b = 1.73 \times 0.8 \times 5 \times 10^{-6}/14 = 0.49$ μF. The capacitor must, of course, be a nonpolarized type.

3.2.5 Half-bridge leakage inductance problems

Leakage inductance spikes, which are so troublesome in the single-ended forward converter and push-pull topology, do not exist in the

half bridge. This is so because any such spikes become clamped to V_{dc} by diodes $D5$, $D6$ across transistors $Q1$, $Q2$.

Thus, if $Q1$ is on, load and magnetizing current flow through it, through the $T1$ leakage inductance, the paralleled $T1$ magnetizing inductance, and the secondary load impedances reflected by their turns ratio squared into the primary and thence through C_b into the $C1$, $C2$ junction. The dot end of N_p is positive with respect to its no-dot end.

Now, when $Q1$ turns off, the magnetizing inductance forces all winding polarities to reverse. The dot end of $T1$ tries to go far negative and if it did, would put more than V_{dc} across $Q1$ and might damage it and also might damage $Q2$ by imposing a reverse voltage across it. But the dot end of $T1$ is caught by diode $D6$ and can go no more negative than the negative end of V_{dc}.

Similarly, when $Q2$ is on, it stores current in the magnetizing inductance. The dot end of N_p is negative with respect to the no-dot end (which is close to $V_{dc}/2$). When $Q2$ turns off, the magnetizing inductance reverses all winding polarities and the dot end of N_p tries to go far positive but is caught at V_{dc} by clamp diode $D5$. Any energy stored in the leakage inductance during the on times is returned to V_{dc} via diodes $D5$, $D6$.

3.2.6 Double-ended forward converter versus half bridge

Both the half-bridge and double-ended forward converter (Fig. 2.13) subject their off transistors to only V_{dc} and not twice that. They are thus both candidates for power supplies to be used in the European market where the prime power is 220 V AC. They both have been used in such applications in enormous numbers, and it is instructive to consider the relative merits and drawbacks of each approach.

The major and probably the most significant difference between the two approaches is that the half-bridge secondary provides a full-wave output as compared to a half wave in the forward converter. Thus the square-wave frequency in the half-bridge secondary is twice that in the forward converter, and hence the output LC inductor and capacitor are smaller with the half bridge.

Peak secondary voltages are higher with the forward converter as their duty cycle is half that of the half bridge. But this is significant only if DC output voltages are high—above 200, as discussed in Sec. 2.5.1.

There are twice as many turns on the forward converter primary as on the half bridge as the former must sustain the full supply voltage as compared to half that voltage for the half bridge. The fewer turns

on the half bridge primary may reduce its winding cost and result in lower parasitic capacities.

One final marginal factor in favor of the half bridge is that the coil losses in the primary due to the proximity effect (Sec. 7.5.6.1) are slightly lower than in the forward converter.

Proximity effect losses are due to induced eddy currents in one winding layer by currents in adjacent layers. Proximity losses increase rapidly with the number of winding layers and the forward converter may have more layers. The half-bridge primary has half the turns of a double-ended forward converter primary of equal output power operating from the same DC supply voltage. But this is balanced somewhat by the larger required wire size for the half bridge. Thus the required number of circular mils for a forward converter primary is given by Eq. 2.42 as $985P_o/\underline{V_{dc}}$ and for a half bridge by Eq. 3.3 as $1395P_o/\underline{V_{dc}}$.

In a practical case, the lower proximity effect losses for the half bridge may be only a marginal advantage. Proximity effect losses will be discussed in detail in Chap. 7.

3.2.7 Practical output power limits in half bridge

Peak primary current and maximum transistor off-voltage stress determine the practical maximum available output power in the half bridge. This limit is about 400 to 500 W for a half bridge operating off 120-V-AC input in the voltage-doubling mode of Fig. 3.1. This is equal to that for the double-ended forward converter as discussed in Sec. 2.4.1.1 and can be seen as follows. The peak equivalent flat-topped primary current is given by Eq. 3.1 as $I_{pft} = 3.13P_o/\underline{V_{dc}}$. For a ±10 percent steady-state tolerance with a 15 percent transient on top of that, the maximum off-voltage stress is $V_{dc} = 1.41 \times 120 \times 2 \times 1.1 \times 1.15$ or 428 V. The minimum DC input voltage is $\underline{V_{dc}} = 1.41 \times 120 \times 2 \times 0.9 \times 0.85 = 259$ V.

Thus, for—say—500 W output, Eq. 3.1 gives the peak primary current as $I_{pft} = 3.13 \times 500/259 = 6.04$ A, and there is a large choice of transistors—either MOSFETs or bipolars—with 428-V, 6-A ratings. Bipolars, of course, would have to have a −1- to −5-V reverse bias (to permit V_{cev} rating) at turnoff to permit a safe off voltage of 428 V. Most adequately fast transistors at that current rating have only a V_{ceo} rating of 400 V.

The half bridge can be pushed to 1000-W output, but at the required 12-A rating, most available bipolar transistors with adequate current ratings and speed have too low a gain. MOSFET transistors at the required current and voltage rating have too large an on drop and are

too expensive for most commercial applications at the time of this writing.

Above 500 W, the topology to consider is the full bridge, a small modification of the half bridge, capable of twice its output power.

3.3 Full-Bridge Converter Topology

3.3.1 Basic operation

Full-bridge converter topology is shown in Fig. 3.3. It is shown with the same voltage-doubling–full-wave bridge rectifying scheme as was shown for the half bridge (Sec. 3.2.1). It can be used as an off-line converter from a 440-V-AC line.

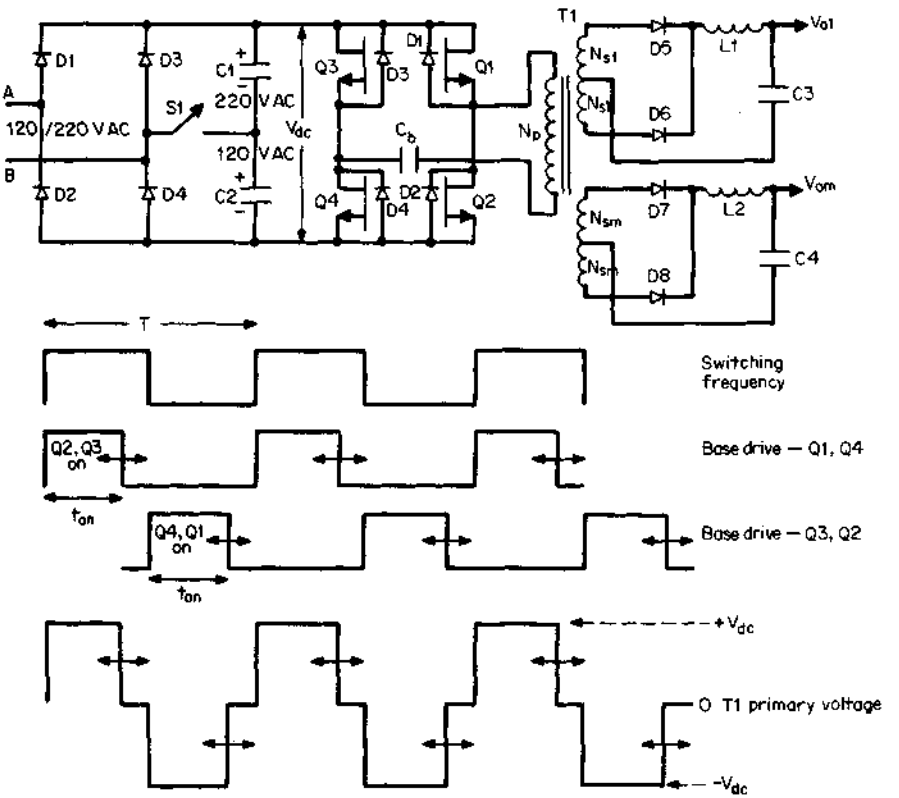


Figure 3.3 Full-bridge converter topology. Power transformer $T1$ is bridged between the junction of $Q1$, $Q2$ and $Q3$, $Q4$. Transistors $Q2$, $Q3$ are switched on simultaneously for an adjustable time during one half period; then transistors $Q4$, $Q1$ are simultaneously on for an equal time during the alternate half period. Transformer primary voltage is a square wave of $\pm V_{dc}$. This compares to the $\pm V_{dc}/2$ primary voltage in the half bridge and yields twice the available power of the half bridge.

Its major advantage is that the voltage impressed across the primary is a square wave of $\pm V_{dc}$ compared to $\pm V_{dc}/2$ for the half bridge. Further, the maximum transistor off-voltage stress is only the maximum DC input voltage—just as for the half bridge. Thus, for transistors of the same peak current and voltage ratings, the full bridge is able to deliver twice the output power of the half bridge.

The transformer primary turns must, of course, be chosen as twice that of the half bridge as it must sustain twice the primary voltage. But for the same output power as a half bridge, from the same DC supply voltage, the peak and rms currents are half that of the half bridge as the transformer primary supports twice the voltage of that in the half bridge. With twice the primary turns but half the rms current, the full-bridge transformer size is identical to that of the half bridge at equal output powers. But with a larger transformer, the full bridge can deliver twice the output of the half bridge with transistors of identical voltage and current ratings.

In Fig. 3.3 is shown a master output, V_{om} and a single slave, V_{o1}. The circuit works as follows. Diagonally opposite transistors ($Q2$ and $Q3$ or $Q4$ and $Q1$) are turned on simultaneously during alternate half cycles. Assuming that the on drop of the transistors is negligible, the transformer primary is driven with an alternating polarity square wave of amplitude V_{dc} and on time t_{on} determined by the feedback loop.

The feedback loop senses a fraction of V_{om} and with its width modulator controls t_{on} so as to keep V_{om} constant against line and load changes. The slave outputs as in all other topologies are kept constant against AC line input changes, but to within only 5 to 8 percent against load changes. Output voltages, assuming 1-V on drop in each switching transistor and 0.5 V forward drop in the master output Schottky rectifiers, 1.0-V forward drops in the slave output rectifiers are

$$V_{om} = \left[(V_{dc} - 2)\frac{N_{sm}}{N_p} - 0.5\right]\frac{2t_{on}}{T} \tag{3.5a}$$

$$V_{om} \approx V_{dc}\frac{N_{sm}}{N_p}\frac{2t_{on}}{T} \tag{3.5b}$$

$$V_{o1} = \left[(V_{dc} - 2)\frac{N_{s1}}{N_p} - 1.0\right]\frac{2t_{on}}{T} \tag{3.6a}$$

$$V_{o1} \approx V_{dc}\frac{N_{s1}}{N_p}\frac{2t_{on}}{T} \tag{3.6b}$$

As in all the other regulators, of course, as V_{dc} goes up or down by a given percentage, the width modulator decreases or increases the on

time by the same percentage so as to keep the product $(V_{dc})(t_{on})$ and hence output voltages constant.

3.3.2 Full-bridge magnetics

3.3.2.1 Maximum on time, core, and primary turns selection. In Fig. 3.3 it can be seen that, just as in the half bridge, if two transistors vertically stacked above one another (as $Q3$ and $Q4$ or $Q1$ and $Q2$) are simultaneously on, there would be a dead short circuit across the DC supply bus and the transistors would fail. To ensure against this, the maximum on time $\overline{t_{on}}$ which occurs at minimum DC input voltage $\underline{V_{dc}}$—as can be seen in Eqs. 3.5b and 3.6b—will be chosen as 80 percent of a half period. This, of course, is "chosen" by choosing the turns ratios N_{sm}/N_p, N_{s1}/N_p so that in those equations for $\underline{V_{dc}}$ and $\overline{t_{on}}$ equal to $0.8T/2$, the correct output voltages—V_{om}, V_{o1}—are obtained.

The magnetic core and operating frequency are chosen from the core-frequency selection chart in Chap. 7. With a core selected and its iron area A_e known, the primary turns number N_p is chosen from Faraday's law (Eq. 1.17). In Eq. 1.17, E is the minimum primary voltage ($\underline{V_{dc}}$ − 2) and dB is the flux change desired in the time dt of $0.8T/2$. And as discussed in Sec. 2.2.9.4, dB will be chosen as 3200 G (−1600 to +1600 G) for frequencies of up to 50 kHz and lesser excursions at higher frequencies as core losses increase.

3.3.2.2 Relation between primary current, output power, and input voltage. Assume an efficiency of 80 percent from the primary input to the total output power. Then $P_o = 0.8P_{in}$ or $P_{in} = 1.25P_o$.

But at minimum DC input voltage $\underline{V_{dc}}$, on time per half period is $0.8T/2$ and duty cycle over a complete period is 0.8. Then neglecting the power transistor on drops, input power at $\underline{V_{dc}}$ is

$$P_{in} = \underline{V_{dc}}(0.8)I_{pft} = 1.25P_o$$

or

$$I_{pft} = \frac{1.56P_o}{\underline{V_{dc}}} \tag{3.7}$$

where I_{pft} is the equivalent primary flat-topped current as described in Sec. 2.2.10.1.

3.3.2.3 Primary wire size selection. Current I_{pft} flows at a duty cycle of 0.8 and hence its rms value is $I_{rms} = I_{pft}\sqrt{0.8}$. Then, from Eq. 3.7 $I_{rms} = (1.56P_o/\underline{V_{dc}})\sqrt{0.8}$ or

$$I_{rms} = \frac{1.40P_o}{V_{dc}} \tag{3.8}$$

And at 500 circular mils per rms ampere, the required number of circular mils is

$$\text{Circular mils needed} = \frac{500 \times 1.40P_o}{V_{dc}}$$

$$= \frac{700P_o}{V_{dc}} \tag{3.9}$$

3.3.2.4 Secondary turns and wire size. The number of turns on each secondary is calculated from Eq. 3.5a and 3.5b, where $\overline{t_{on}}$ is $0.8T/2$ for the specified minimum DC input V_{dc}, N_p is as calculated in Sec. 3.3.2.1, and all DC outputs are specified.

Secondary rms currents and wire sizes are chosen exactly as for the push-pull secondaries as described in Sec. 2.2.10.3. The secondary rms currents are given by Eq. 2.13 and the required circular mils for each half secondary is given by Eq. 2.14.

3.3.3 Output filter calculations

As for the half-bridge and push-pull topologies which also have full-wave output rectifiers, the output inductor and capacitors are calculated from Eqs. 2.20 and 2.22. Equation 2.20 specifies the output inductor for minimum DC output currents equal to one-tenth the nominal values. Equation 2.22 specifies the output capacitor for the specified peak-to-peak output ripple V_r and the selected peak-to-peak inductor current ripple amplitude.

3.3.4 Transformer primary blocking capacitor

Figure 3.3 shows a small nonpolarized blocking capacitor C_b in series with the transformer. It is needed to avoid the flux imbalance problem as discussed in Sec. 3.2.4.

Flux imbalance in the full bridge is less likely than in the half bridge, but still is possible. With bipolars, an on pair in one half cycle may have different storage times than the pair in the alternate half cycle. With MOSFETs, the on-voltage drops of the pairs in alternate half cycles may be unequal. In either case, the volt-second product applied to the transformer primary in alternate half cycles will be unequal and the core may walk off center of its hysteresis loop, saturate the core, and destroy the transistors.

Chapter

4

Flyback Converter Topologies

4.1 Introduction

All topologies discussed thus far, with the exception of the boost regulator (Sec. 1.4) and the polarity inverter (Sec. 1.5) deliver power to their loads during the time when the power transistor is turned on.

Flyback topologies described in this chapter operate in a fundamentally different way. During their power transistor on time, they store energy in their power transformer while load current is supplied from an output filter capacitor. When the power transistor turns off, the energy stored in the power transformer is transferred to the output as load current and to the filter capacitor to replenish the charge it lost when it alone was delivering load current.

The topology has advantages and drawbacks, discussed in detail below. The major advantage is that the output filter inductors required for all forward topologies is not required for flybacks. Especially for multioutput power supplies, this is a significant saving in cost and space.

4.2 Flyback Converter—Areas of Application

Flyback converter topology is shown in Fig. 4.1. It is very widely used for output powers from about 150 down to under 5 W. Its great initial attraction—although it is not strictly so, as will soon be seen—is that it has no secondary output inductors as have all topologies discussed thus far. The consequent savings in cost and volume of the output inductors is a significant advantage.

It is widely used for high output voltages at relatively low power ($\leq$5000 V at $<$15 W). It can also be used at powers of up to 150 W if DC supply voltages are high enough ($\geq$160 V) so that primary cur-

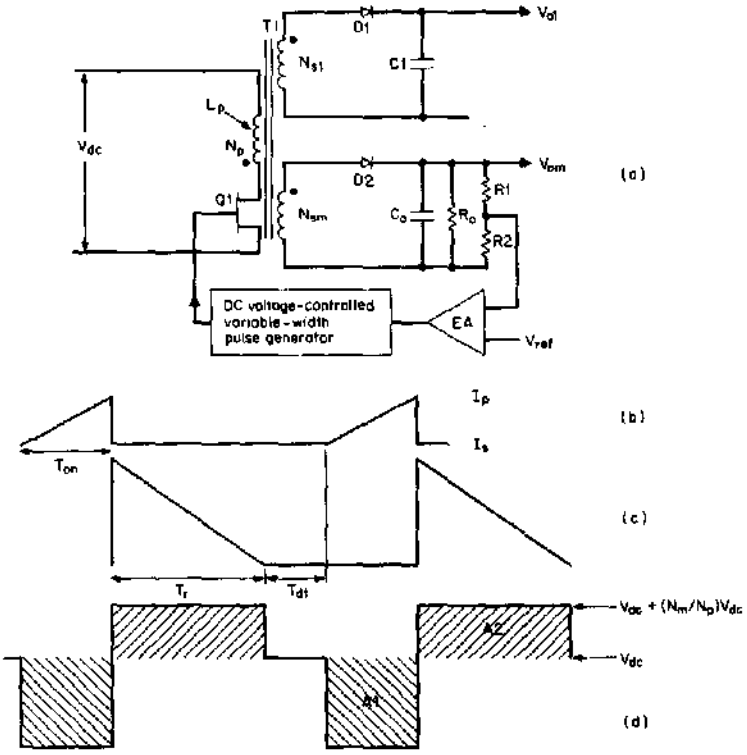


Figure 4.1 Discontinuous-mode flyback converter. When $Q1$ is on, all rectifier diodes are reverse-biased and all output capacitors supply load currents. N_p acts like a pure inductor and load current builds up linearly in it to a peak I_p. When $Q1$ turns off, the primary stored energy $\frac{1}{2}LI_p^2$ is delivered to the secondaries to supply load current and replenish the charge on output capacitors which they had lost when $Q1$ was on. The circuit is discontinuous if the secondary current has decayed to zero before the start of the next turnon.

rents are not excessive. The feature which makes it valuable for high output voltages is that it requires no output inductor. In forward converters, discussed above, output inductors become a troublesome problem at high output voltages because of the large voltages they have to sustain. Not requiring a high voltage free-wheeling diode is also a plus for the flyback in high-voltage supplies.

It is also a frequent choice for a supply with many output voltages (≤ 10 is not uncommon) in the region of 50 to 150 W. It is attractive for a multioutput supply because the output voltages track one another with line and load changes far better than they do in the forward

types described earlier. The better tracking is due to the absence of output inductors.

It can be used from DC input voltages as low as 5 V up to the usual rectified 160 V from a 115-V-AC power line by proper choice of the secondary to primary turns ratios. By careful design of these turns ratios, it can also be used off the rectified DC of 160 V from a 115-V-AC power line or from the rectified 320 V DC from a 220-V-AC power line without resorting to the voltage doubling–full-wave rectifying scheme (switch $S1$) of Fig. 3.1.

The latter scheme, although very widely used, has the objectionable feature that to do the switching from 115 to 220 V AC, both ends of the switch in Fig. 3.1 have to be accessible on the outside of the supply which is a safety hazard. Or the supply must be opened to change the switch position. Both these alternatives have drawbacks. An alternative scheme not requiring switching will be discussed in Sec. 4.3.5.

4.3 Discontinuous-Mode Flybacks—Basic Operation

Referring to Fig. 4.1, the topology works as follows. Figure 4.1 shows a master and one slave output. As in all other topologies shown previously, a negative-feedback loop will be closed around the master V_{om}. A fraction of V_{om} will be compared to a reference, and the error signal will control the $Q1$ on time pulse width so as to make the sampled output equal to the reference voltage for line and load changes. The slaves will be well regulated against line changes and somewhat less against load changes. But slave changes with line and load will be better than in the previous forward-type topologies.

In Fig. 4.1, flyback operation can be immediately recognized from the dots on the transformer primary and secondary. When $Q1$ is on, dot ends of all windings are negative with respect to their no-dot ends. Output rectifier diodes $D1$ and $D2$ are reverse-biased and all the output load currents are supplied from storage filter capacitors $C1$ and $C2$. These will be chosen as described below to deliver the load currents with the maximum specified ripple or droop in output voltage.

During the $Q1$ on time, there is a fixed voltage across N_p and current in it ramps up linearly (Fig. 4.1*b*) at a rate of $dI/dt = (V_{dc} - 1)/L_p$, where L_p is the primary magnetizing inductance. At the end of the on time, the primary current has ramped up to $I_p = (V_{dc} - 1)T_{on}/L_p$. This current represents a stored energy of

$$E = \frac{L_p(I_p)^2}{2} \tag{4.1}$$

where E is in joules
L_p is in henries
I_p is in amperes

Now when $Q1$ turns off, current in the magnetizing inductance forces reversal of polarities on all windings. Assume for the moment that there are no slave windings and only the master secondary N_m. Since the current in an inductor cannot change instantaneously, at the instant of turnoff, the primary current transfers to the secondary at an amplitude $I_s = I_p(N_p/N_m)$.

After a number of cycles, the secondary DC voltage has built up to a magnitude (calculated below) of V_{om}. Now with $Q1$ off, the dot end of N_m is positive with respect to its no-dot end and current flows out of it, but ramps down linearly (Fig. 4.1c) at a rate $dI_s/dt = V_{om}/L_s$, where L_s is the secondary inductance. If the secondary current has ramped down to zero before the start of the next $Q1$ on time, all the energy stored in the primary when $Q1$ was on has been delivered to the load and the circuit is said to be operating in the discontinuous mode. Since an amount of energy E in joules delivered in a time T in seconds represents input power in watts, at the end of one period, power drawn from V_{dc} is

$$P = \frac{\frac{1}{2}L_p(I_p)^2}{T} \text{ W} \tag{4.2a}$$

But $I_p = (V_{dc} - 1)T_{on}/L_p$. Then

$$P = \frac{[(V_{dc} - 1)T_{on}]^2}{2TL_p} \approx \frac{(V_{dc}T_{on})^2}{2TL_p} \text{ W} \tag{4.2b}$$

As can be seen from Eq. 4.2b, the feedback loop maintains constant output voltage by keeping the product $V_{dc}T_{on}$ constant.

4.3.1 Relation between output voltage versus input voltage, on time, output load

Assume an efficiency of 80 percent.

$$\text{Input power} = 1.25 \text{ (output power)}$$

$$= \frac{1.25(V_o)^2}{R_o} = \frac{\frac{1}{2}(L_p I_p{}^2)}{T}$$

But $I_p = V_{dc}\overline{T_{on}}/L_p$ since maximum on time $\overline{T_{on}}$ occurs at minimum supply voltage $\underline{V_{dc}}$, as can be seen from Eq 4.2b.

Then $1.25(V_o)^2/R_o = \frac{1}{2}L_p\underline{V_{dc}}^2\,\overline{T_{on}}^2/L_p{}^2T$ or

$$\underline{V_o} = \underline{V_{dc}}\overline{T_{on}}\sqrt{\frac{R_o}{2.5TL_p}} \tag{4.3}$$

Thus the feedback loop will regulate the output by decreasing T_{on} as V_{dc} or R_o goes up, increasing T_{on} as V_{dc} or R_o goes down.

4.3.2 Design relations and sequential decision requirements

4.3.2.1 Establishing primary/secondary turns ratio. There are a number of decisions which should be made in the proper sequence. The first is to choose the primary/master secondary turns ratio N_p/N_{sm}; this determines the maximum off-voltage stress $\overline{V_{ms}}$ on the power transistor in the absence of a leakage inductance spike. Neglecting the leakage spike, the maximum transistor voltage stress at maximum DC input $\overline{V_{dc}}$ and for a 1-V rectifier drop is

$$\overline{V_{ms}} = \overline{V_{dc}} + \frac{N_p}{N_{sm}}(V_o + 1) \tag{4.4}$$

where $\overline{V_{ms}}$ is chosen sufficiently low so that a leakage inductance spike of $0.3V_{dc}$ on top of that still leaves a safety margin of about 30 percent below the maximum pertinent transistor rating (V_{ceo}, V_{cer}, or V_{cev}).

4.3.2.2 Ensuring core does not saturate, circuit remains discontinuous. Recall that to ensure that the core does not drift up or down its hysteresis loop, the on volt-second product ($A1$ in Fig. 4.1*d*) must equal the reset volt-second product ($A2$ in Fig. 4.1*d*). Assume that the on drop of $Q1$ and the forward drop of the rectifier $D2$ are both 1 V:

$$(\underline{V_{dc}} - 1)\overline{T_{on}} = (V_o + 1)\frac{N_p}{N_{sm}}T_r \tag{4.5}$$

where T_r is the reset time shown in Fig. 4.1*c* and is the time required for the secondary current to return to zero.

To ensure the circuit operates in the discontinuous mode, a dead time (T_{dt} in Fig. 4.1*c*) is established so that the maximum on time $\overline{T_{on}}$, which occurs when V_{dc} is a minimum plus the reset time T_r is only 80 percent of a full period. This leaves $0.2T$ margin against unexpected decreases in R_o, which according to Eq. 4.3 would force the feedback loop to increase T_{on} in order to keep V_o constant.

As for the boost regulator (which is also a flyback type), Secs. 1.4.2

and 1.4.3 pointed out that if the error amplifier had been designed to keep the loop stable in the discontinuous mode, if the circuit momentarily entered the continuous mode, it will break into oscillation.

Thus, increasing DC load current or decreasing V_{dc} causes the error amplifier to increase T_{on} in order to keep V_o constant (Eq. 4.3). This increased T_{on} eats into the dead time T_{dt} and eventually results in the secondary current not having fallen to zero by the start of the next $Q1$ on time. This is the start of the continuous mode, and if the error amplifier has not been designed for this with a drastically lower bandwidth, the circuit will oscillate. To ensure that the circuit remains discontinuous, the maximum on time which will generate the desired maximum output power is established thus:

$$\overline{T_{on}} + T_r + T_{dt} = T$$

or

$$\overline{T_{on}} + T_r = 0.8T \tag{4.6}$$

Now in Eqs. 4.5 and 4.6, there are two unknowns as N_p/N_{sm} has been calculated from Eq. 4.4 for specified $\overline{V_{dc}}$ and $\overline{V_{ms}}$. Then from the last two relations

$$\overline{T_{on}} = \frac{(V_o + 1)(N_p/N_{sm})(0.8T)}{(\overline{V_{dc}} - 1) + (V_o + 1)(N_p/N_{sm})} \tag{4.7}$$

4.3.2.3 Primary inductance versus minimum output resistance and DC input voltage. From Eq. 4.3, the primary inductance is

$$L_p = \frac{R_o}{2.5T}\left(\frac{\overline{V_{dc}}\,\overline{T_{on}}}{V_o}\right)^2 = \frac{(\overline{V_{dc}}\,\overline{T_{on}})^2}{2.5T\overline{P_o}} \tag{4.8}$$

4.3.2.4 Transistor peak current, maximum voltage stress. The transistor, if bipolar, should have an acceptably high gain at a peak current of

$$I_p = \frac{\overline{V_{dc}}\,\overline{T_{on}}}{L_p} \tag{4.9}$$

where V_{dc} is specified and $\overline{T_{on}}$ is calculated from Eq. 4.7 and L_p is calculated from Eq. 4.8. If it is a MOSFET transistor, it should have a peak current rating about 5 to 10 times the value calculated from Eq. 4.9 so that its on resistance is low enough to yield a low on drop.

4.3.2.5 Primary rms current and wire size. Primary current is a triangle of peak amplitude I_p (Eq. 4.9) at a maximum duration T_{on} out of every period T. Its rms value then (see Sec. 2.2.10.6) is

$$I_{rms(primary)} = \frac{I_p}{\sqrt{3}} \sqrt{\frac{\overline{T_{on}}}{T}} \tag{4.10}$$

where I_p and $\overline{T_{on}}$ are as given by Eqs. 4.9 and 4.7.

At 500 circular mils per rms ampere, the required number of circular mils is

$$\text{Circular mils required (primary)} = 500 I_{rms(primary)}$$

$$= 500 \frac{I_p}{\sqrt{3}} \sqrt{\frac{\overline{T_{on}}}{T}} \tag{4.11}$$

4.3.2.6 Secondary rms current and wire size. Secondary current is a triangle of peak amplitude $I_s = I_p(N_p/N_s)$ and duration T_r. Primary/secondary turns ratio N_p/N_s is given by Eq. 4.4 and $T_r = (T - T_{on})$. Secondary rms current is then

$$I_{rms(secondary)} = \frac{I_p(N_p/N_s)}{\sqrt{3}} \sqrt{\frac{T_r}{T}} \tag{4.12}$$

At 500 circular mils per rms ampere, the required number of circular mils is

$$\text{Secondary circular mils required} = 500 I_{rms(secondary)} \tag{4.13}$$

4.3.2.7 Design example—discontinuous-mode flyback. Design a flyback converter with the following specifications:

V_o	5.0 V
$P_{o(max)}$	50 W
$I_{o(max)}$	10 A
$I_{o(min)}$	1.0 A
$V_{dc(max)}$	60 V
$V_{dc(min)}$	38 V
Switching frequency	50 kHz

First select the voltage rating of the transistor as this mainly determines the transformer turns ratio. Choose a device with a 200-V rat-

ing. In Eq. 4.4 choose the maximum stress V_{ms} on the transistor in the off state (excluding the leakage inductance spike) as 120 V. Then even with a 25 percent or 30-V leakage spike, this leaves a 50-V margin to the maximum voltage rating. Then from Eq. 4.4:

$$120 = 60 + \frac{N_p}{N_{sm}}(V_o + 1) \qquad \text{or} \qquad \frac{N_p}{N_{sm}} = 10$$

Now choose maximum on time from Eq. 4.7:

$$\overline{T_{on}} = \frac{(V_o + 1)(N_p/N_{sm})(0.8T)}{(V_{dc} - 1) + (V_o + 1)N_p/N_{sm}}$$

$$= \frac{6 \times 10 \times 0.8 \times 20)}{(38 - 1) + 6 \times 10}$$

$$= 9.9\ \mu s$$

From Eq. 4.8

$$L_p = \frac{(V_{dc}\overline{T_{on}})^2}{2.5TP_o}$$

$$= \frac{(38 \times 9.9 \times 10^{-6})^2}{2.5 \times 20 \times 10^{-6} \times 50}$$

$$= 56.6\ \mu H$$

From Eq. 4.9

$$I_p = \frac{V_{dc}\overline{T_{on}}}{L_p}$$

$$= \frac{38 \times 9.9 \times 10^{-6}}{56.6 \times 10^{-6}}$$

$$= 6.6 \quad \text{A}$$

From Eq. 4.10, primary rms current is

$$I_{rms(primary)} = \frac{I_p}{\sqrt{3}}\sqrt{\frac{\overline{T_{on}}}{T}}$$

$$= \frac{6.6}{\sqrt{3}} \times \sqrt{\frac{9.9}{20}}$$

$$= 2.7\ \text{A}$$

From Eq. 4.11, primary circular mils requirement is

$$I_{\text{(primary circular mils)}} = 500 \times 2.7 = 1350 \text{ circular mils}$$

This calls for No. 19 wire of 1290 circular mils, which is close enough.

From Eq. 4.12, secondary rms current is

$$I_{\text{rms(secondary)}} = \frac{I_p(N_p/N_s)}{\sqrt{3}} \sqrt{\frac{T_r}{T}}$$

But reset time T_r is

$$(0.8T - \overline{T_{\text{on}}}) = (16 - 9.9) = 6.1 \ \mu\text{s}$$

Then

$$I_{\text{rms(secondary)}} = \frac{6.6 \times 10}{\sqrt{3}} \sqrt{\frac{6.1}{20}}$$

$$= 21 \text{ A}$$

and from Eq. 4.12, the required number of circular mils is $500 \times 21 = 10{,}500$. This calls for No. 10 wire, which, of course, is impractically large in diameter. A foil winding or a number of smaller diameter wires in parallel with an equal total circular-mil area would be used.

Output capacitor is chosen on the basis of specified output ripple. At maximum output current, during the transistor on time of 9.9 μs, the filter capacitor C_o carries the 10 A for 10.1 μs. It slopes in voltage an amount $V = I(T - t_{\text{on}})C_o$. Then for a voltage droop of 0.05 V

$$C_o = \frac{10 \times 10.1 \times 10^{-6}}{0.05}$$

$$= 2000 \ \mu\text{F}$$

From Sec. 1.4.7, the average value of the ESR of a 2000-μF aluminum electrolytic capacitor is $R_{\text{esr}} = 65 \times 10^{-6}/C_o = 0.03 \ \Omega$.

At the instant of transistor turnoff, the peak secondary current of 66 A flows through this ESR, causing a thin spike of $66 \times 0.03 = 2$ V. This large-amplitude thin spike at transistor turnoff is a universal problem with flybacks having a large N_p/N_s ratio. It is usually solved by using a larger filter capacitor than calculated as above (since R_{esr} is inversely proportional to C_o) and/or integrating away the thin spike with a small LC circuit.

The selection of a transformer core for a flyback topology circuit is

significantly different from nonflyback core selection. Recall in the flyback that when current flows in the primary, no load current flows in the secondary to buck out the primary ampere turns. Thus in the flyback, all the primary ampere turns tend to saturate the core.

In contrast, in nonflyback topologies, secondary load current flows when primary current flows and is in the direction (by Lenz' law) to cancel the ampere turns of the primary. It is only the primary magnetizing current which drives the core over its hysteresis loop and moves it toward saturation. That magnetizing current is kept a small fraction of the primary load current by providing a large magnetizing inductance and hence core saturation is never a problem with nonflyback topologies.

Thus for flyback transformers, cores must have some means of carrying large primary currents without saturating. This is done with gapped ferrite cores (Sec. 2.3.9.3) or with MPP (molybdenum permalloy powder) cores which have an inherent internal air gap. This is discussed further in the following section.

4.3.3 Flyback magnetics

Referring to Fig. 4.1*a*, it is seen from the winding dots that when the transistor is on and current flows in the primary, no secondary currents flow. This is totally different from forward-type converters where current flows in the secondary when it flows in the primary. Thus in a forward-type converter, primary current flows into a dot end and secondary current flows out of a dot end.

Primary and secondary load ampere turns then cancel each other out and do not move the core across its hysteresis loop. In the forward-type converters, it is only the magnetizing current which drives the core across the hysteresis loop and may potentially saturate it. But this magnetizing current is a small fraction (rarely >10 percent) of the total primary current.

In a flyback converter, however, the entire triangle of primary current shown in Fig. 4.1*b* drives the core across the hysteresis loop as it is not canceled out by any secondary ampere turns. Thus even at very low output power, the core would almost immediately saturate and destroy the transistor if nothing were done to prevent it.

To prevent core saturation in the flyback transformer, the core is gapped. The gapped core can be of either of two types. It can be a solid ferrite core with a known air-gap length obtained by grinding down the center leg in EE or cup-type cores. The known gap length can also be obtained by inserting a plastic shim between the two halves of an EE, cup, or UU core.

A more usual gapped core for flyback converters is the MPP or

molypermalloy powder core. Such cores are toroids made of baked and hardened mix of magnetic powdered particles. These powdered particles are mixed in a slurry with a plastic resin binder and cast in the shape of toroids. Each magnetic particle in the toroid is thus encapsulated within a resin envelope which behaves as a "distributed air gap" and acts to keep the core from saturating. The basic magnetic material which is ground up into a powder is Square Permalloy 80, an alloy of 79% nickel, 17% iron, and 4% molybdenum made by Magnetics Inc. or Arnold Magnetics.

The permeability of the resulting toroid is determined by controlling the concentration of magnetic particles in the slurry. Permeabilities are controlled to within ±5 percent over large temperature ranges and are available in discrete steps ranging from 14 to 550. Toroids with low permeability behave like gapped cores with large air gaps. They require a relatively large number of turns to yield a desired inductance but tolerate many ampere turns before they saturate. Higher permeability cores require relatively fewer turns but saturate at a lower number of ampere turns.

Such MPP cores are used not only for flyback transformers where all the primary current is a DC bias current. They are also used for forward converter output inductors where, as has been seen, a unique inductance is required at the large DC output current bias (Sec. 1.3.6).

4.3.3.1 Gapped ferrite cores to avoid saturation. Adding an air gap to the solid ferrite core achieves two results. First, it tilts the hysteresis loop as shown in Fig. 2.5 and hence decreases its permeability, which must be known to select the number of turns for a desired inductance. Second, and more important, it increases the number of ampere turns it can tolerate before it saturates.

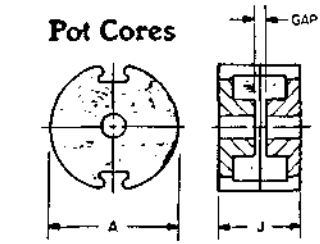
Core manufacturers often give curves which permit calculation of the number of turns for a desired inductance and the number of ampere turns at which saturation commences. Such curves are shown in Fig. 4.2 and show A_{lg}, the inductance per 1000 turns with an air gap and the number of ampere turns (NI_{sat})where saturation starts to set in. Since inductance is proportional to the square of the number of turns, the number of turns N_l for any inductance L is calculated from

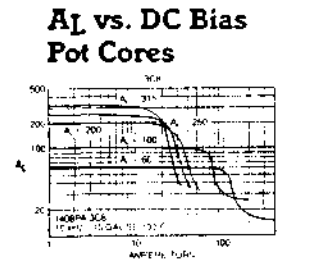
$$N_l = 1000\sqrt{\frac{L}{A_{lg}}} \tag{4.14}$$

Figure 4.2 shows A_{lg} curves for a number of different air gaps and the "cliff" point at which saturation starts. It can be seen that the larger the air gap, the lower the value of A_{lg} and the larger the number of ampere turns at which saturation starts. If such curves were

Part No	A	J	Approx Gap	A_L mH/1000 Turns ±3%	μ_e (Ref.)
1408PA 60-3C8	.551	.328	.027	60	37
1408PA100-3C8	.551	.328	.013	100	63
1408PA200-3C8	.551	.328	.005	200	126
1408PA250-3C8	.551	.328	.004	250	156
1408PA315-3C8	.551	.328	.003	315	198
1811PA 75-3C8	.705	.416	.037	75	35
1811PA130-3C8	.705	.416	.018	130	62
1811PA250-3C8	.705	.416	.008	250	119
1811PA315-3C8	.705	.416	.006	315	150
1811PA400-3C8	.705	.416	.004	400	190
2213PA 85-3C8	.846	.528	.050	85	33
2213PA145-3C8	.846	.528	.025	145	57
2213PA315-3C8	.846	.528	.009	315	123
2213PA400-3C8	.846	.528	.007	400	157
2213PA500-3C8	.846	.528	.005	500	196
2616PA100-3C8	1.004	.634	.064	100	31
2616PA170-3C8	1.004	.634	.032	170	53
2616PA400-3C8	1.004	.634	.009	400	125
2616PA500-3C8	1.004	.634	.007	500	156
2616PA630-3C8	1.004	.634	.005	630	197
3019PA125-3C8	1.181	.740	.070	125	32
3019PA210-3C8	1.181	.740	.035	210	54
3019PA500-3C8	1.181	.740	.011	500	129
3019PA630-3C8	1.181	.740	.008	630	163
3019PA800-3C8	1.181	.740	.007	800	206
3622PA160-3C8	1.398	.855	.079	160	33
3622PA275-3C8	1.398	.855	.040	275	57
3622PA630-3C8	1.398	.855	.016	630	131
3622PA800-3C8	1.398	.855	.012	800	166
3622PA1000-3C8	1.398	.855	.008	1000	208
4229PA160-3C8	1.669	1.164	.103	160	33
4229PA275-3C8	1.669	1.164	.051	275	56
4229PA630-3C8	1.669	1.164	.020	630	128
4229PA800-3C8	1.669	1.164	.015	800	162
4229PA1000-3C8	1.669	1.164	.011	1000	202

Nominal dimensions in inches.





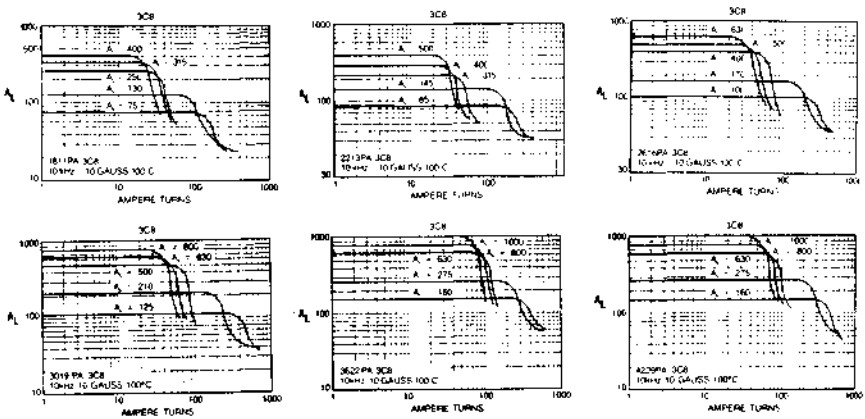


Figure 4.2 Inductance per 1000 turns (A_{lg}) for various ferrite cores with various air gaps. Also the point in ampere turns "cliff" point where saturation commences. (*Courtesy Ferroxcube Corporation.*)

available for all cores at various air gaps, Eq. 4.14 would give the number of turns for any selected air gap from the value of A_{lg} read from the curve. The cliff point on the curve would tell whether, at those turns and for the specified primary current, the core had fallen over the saturation cliff.

Such curves, though, are not available for all cores and all air gaps. But this is no problem, for A_{lg} can be calculated with reasonable ac-

curacy from Eq. 2.39 and A_l with no gap, which is always given in the manufacturers' catalogs. The cliff point at which saturation starts can be calculated from Eq. 2.37 for any air gap. The cliff point corresponds to the flux density in iron B_i, where the core material itself starts bending over into saturation.

From Fig. 2.3, it is seen that this is not a very sharply breaking point but occurs around 2500 G for this ferrite material (Ferroxcube 3C8). Thus the cliff in ampere turns is found by substituting 2500 G in Eq. 2.37. As noted in connection with Eq. 2.37, in the usual case, the air-gap length l_a is much larger than l_i/u as u is so large. Then the iron flux density as given by Eq. 2.37 is determined mainly by the air-gap length l_a.

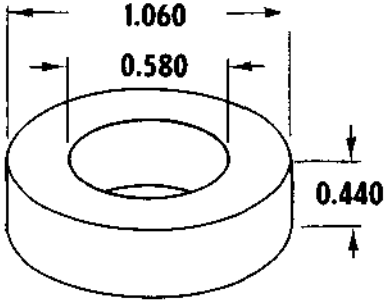
4.3.3.2 Powdered permalloy (MPP) cores to avoid saturation. These toroidal cores are widely used and made by Magnetics Inc. (data in catalog MPP303S) and by Arnold Co. (data in catalog PC104G).

The problem in designing a core of desired inductance at a specified maximum DC current bias is to select a core geometry and material permeability, such that the core does not saturate at the maximum ampere turns to which it is subjected. There are a limited number of core geometries, each available in permeabilities ranging from 14 to 550. Selection procedures are described in the catalogs mentioned above, but the following has been found more direct and useful.

In the Magnetics Inc. catalog, one full page (Fig. 4.3) is devoted to each size toroid and for each size, its A_l value (inductance in millihenries per 1000 turns) is given for each discrete permeability. Figure 4.4, also from the Magnetics Inc. catalog, gives the falloff in permeability (or A_l value) for increasing magnetizing force in oersteds for core material of the various available permeabilities. (Recall the oersteds–ampere-turns relation in Eq. 2.6.)

A core geometry and permeability can be selected so that at the maximum DC current and the selected number of turns, the A_l and hence inductance has fallen off by any desired percentage given in Fig. 4.4. Then at zero DC current the inductance will be greater by that percentage. Such inductors or chokes are referred to as "swinging chokes" and in many applications are desirable. For example, if an inductor is permitted to swing a great deal, in an output filter, it can tolerate a very low minimum DC current before it goes discontinuous (Sec. 1.3.6). But this greatly complicates the feedback-loop stability design and most often, the inductor in an output filter or transformer in a flyback will not be permitted to "swing" or vary very much between its zero and maximum current value.

Referring to Fig. 4.3, it is seen that a core of this specific size is available in permeabilities ranging from 14 to 550. Cores with perme-



WINDOW AREA	308,000 cir. mils	
CROSS SECTION	0.1014 in²	0.654 cm²
PATH LENGTH	2.50 in	6.35 cm
WEIGHT	1.3 oz (.08 lb.)	36. gm

WINDING TURN LENGTH

WINDING FACTOR	LENGTH/TURN	
100% (UNITY)	0.1714 ft	5.23 cm
60%	0.1526 ft	4.66 cm
40%	0.1344 ft	4.10 cm
20%	0.1263 ft	3.85 cm
0%	0.1233 ft	3.76 cm

CORE DIMENSIONS AFTER FINISH

OD (MAX.)	1.090 in	27.7 mm
ID (MIN.)	0.555 in	14.10 mm
HT (MAX.)	0.472 in	11.99 mm

WOUND COIL DIMENSIONS

UNITY WINDING FACTOR

OD (MAX.)	1.468 in	37.3 mm
HT (MAX.)	0.944 in	24.0 mm

MAGNETIC INFORMATION

PART NO.	PERM μ	INDUCTANCE @ 1000 TURNS MH±8%	NOMINAL DC RESISTANCE, OHMS/MH**	FINISHES AND STABILIZATIONS*	GRADING STATUS 2% BANDS	A/N GAUSS PER AMP. TURN
55933–	14	18	0.457	A2	*	2.77 (< 1500 gauss)
55932–	26	32	0.257	A2	*	5.15 (< 1500 gauss)
55894–	60	75	0.110	ALL	YES	11.9 (< 1500 gauss)
55930–	125	157	0.0524	ALL	YES	24.8 (< 1500 gauss)
55929–	147	185	0.0444	ALL	YES	29.1 (< 1500 gauss)
55928–	160	201	0.0409	ALL	YES	31.7 (< 1500 gauss)
55924–	173	217	0.0379	ALL	YES	34.3 (< 1500 gauss)
55927–	200	251	0.0327	ALL	YES	39.6 (< 600 gauss)
55925–	300	377	0.0218	A2 and L6	YES	59.4 (< 300 gauss)
55926–	550	740	0.0111	A2	YES	109 (< 50 gauss)

14, 26, 60 and 125μ types are available as high flux cores. See p. 84, "How to Order".

WINDING INFORMATION

FOR UNITY WINDING FACTOR

AWG WIRE SIZE	TURNS	Ω/in	AWG WIRE SIZE	TURNS	Ω/in
9	21	0.00291	24	587	2.58
10	27	0.00459	25	725	4.02
11	34	0.00726	26	906	6.37
12	42	0.01148	27	1,141	10.05
13	53	0.01805	28	1,400	15.67
14	66	0.0284	29	1,711	24.1
15	82	0.0447	30	2,139	38.1
16	103	0.0707	31	2,633	59.1
17	127	0.1102	32	3,209	89.1
18	159	0.1739	33	3,980	140.5
19	197	0.272	34	5,066	227
20	246	0.428	35	6,286	357
21	308	0.676	36	7,759	552
22	380	1.056	37	9,478	832
23	474	1.649	38	11,847	1,316

Figure 4.3 A typical MPP core. With its large distributed air gap, it can tolerate a large DC current bias without saturating. It is available in a large range of different geometries. (*Courtesy Magnetics Inc.*)

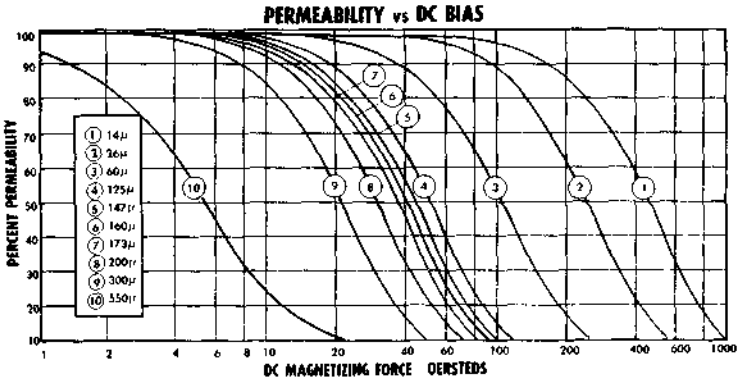


Figure 4.4 Falloff in permeability or A_1 for MPP cores of various permeabilities versus DC magnetizing force in oersteds. (*Courtesy Magnetics Inc.*)

ability above 125 have large values of A_l and hence require fewer turns for a specified inductance at zero DC current bias. But in Fig. 4.4 it is seen that the higher-permeability cores saturate at increasingly lower oersteds or ampere turns of bias. Hence in power supply usage, where DC current biases are rarely under 1 A, cores of permeability greater than 125 are rarely used. And an inductance swing or change of 10 percent from zero to the maximum specified current is most often easily acceptable.

Thus in Fig. 4.4 it is seen that for a permeability dropoff or swing of 10 percent, core materials of permeabilities 14, 26, 60, and 125, can sustain maximum magnetizing forces of only 170, 95, 39, and 19 Oe, respectively. These maximum magnetizing forces in oersteds can be translated into maximum ampere turns by Eq. 2.6 ($\overline{H} = 0.4\pi(\overline{NI})/l_m$), where l_m is the magnetic path length in centimeters, given in Fig. 4.3 for this particular core geometry as 6.35 cm.

From these maximum number of ampere turns ($\overline{NI}$), beyond which inductance falls off more than 10 percent, the maximum number of turns ($\overline{N}$) is calculated for any peak current. And from $\overline{N}$, the maximum inductance possible for any core at the specified peak current is calculated as $L_{\max} = 0.9A_l\,(N_{\max}/1000)^2$.

Tables 4.1, 4.2, 4.3 show $N_{\max}$ and $L_{\max}$ for three often used core geometries in permeabilities of 14, 26, 60, and 125 at peak currents of 1, 2, 3, 5, 10, 20, and 50 amperes. These tables permit core geometry and permeability selection at a glance without iterative calculations.

Thus, Table 4.1 is used in the following manner. Assume that this particular core has the acceptable geometry. The table is entered horizontally to the first peak current greater than specified value. At that peak current, move down vertically until the first inductance L_{max}

TABLE 4.1 Maximum Number of Turns and Maximum Inductance at Those Turns for Various Peak Currents I_p at a Maximum Inductance Falloff of 10 Percent from Zero Current Level

Magnetics Inc. core number	Permeability	A_l, mH per 1000 turns	Maximum oersted level H for 10% falloff in inductance	NI Maximum permissible ampere turns corresponding to H	Maximum permissible turns and inductance at those turns for a 10% inductance falloff at indicated peak currents							
					N_{max}/L_{max}							
Core	μ	A_l	H	NI	1A	2A	3A	5A	10A	20A	50A	I_p
55930	125	157	19	96	96	48	32	19	10	5	2	N_{max}
					1,382	339	145	56	15	3.5	0.6	L_{max}
55894	60	75	39	197	197	99	66	39	20	10	4	N_{max}
					2,620	662	294	103	27	7	1	L_{max}
55932	26	32	95	480	480	240	160	96	48	24	10	N_{max}
					6,635	1,659	737	265	66	17	3	L_{max}
55933	14	18	170	859	859	430	286	172	86	43	17	N_{max}
					11,954	2,995	1,325	479	120	30	5	L_{max}

Note: Magnetics Inc. MPP cores. All cores have outer diameter (OD) = 1.060 in, inner diameter (ID) = 0.58 in, height = 0.44 in, l_m = 6.35 cm. All inductances in microhenries.

greater than the desired value is reached. The core at that point is the only one which can yield the desired inductance with only a 10 percent swing. The number of turns N_d on that core for a desired inductance L_d is within 5 percent given by

$$N_d = 1000 \sqrt{\frac{L_d}{0.95A_l}}$$

where A_l is the value in column 3 in Table 4.1. If, moving vertically, no core can be found whose maximum inductance is greater than the desired value, the core with the next larger geometry (greater OD or greater height) must be used. The core ID must, of course, be large enough to accommodate the number of turns of wire selected at the rate of 500 circular mils per rms ampere or the next larger size core must be used.

Tables 4.2 and 4.3 show similar data for a smaller (OD = 0.80 in) and larger (OD = 1.84 in) family of cores. Similar charts can be generated for all the other available core sizes, but Tables 4.1 to 4.3 bracket about 90 percent of the possible designs for flyback transformers under 500 W or output inductors of up to 50 A.

A commonly used scheme for correcting the number of turns on a core when an initial selection has resulted in too large an inductance falloff should be noted. If, for an initially selected number of turns and a specified maximum current, the inductance or permeability falloff from Fig. 4.4 is down by P percent, the number of turns is increased by P percent.

This moves the operating point further out by P percent as the magnetizing force in oersteds is proportional to the number of turns. The core slides further down its saturation curve, and it might be thought that the inductance would falloff even more. But since inductance is proportional to the square of the number of turns, and magnetizing force is proportional only to the number of turns, the zero current inductance has been increased by $2P$ percent but magnetizing force has gone up only by P percent. The inductance is then correct at the specified maximum current. If the consequent swing is too large, a larger core must be used.

4.3.4 Flyback disadvantages

Despite its many advantages, the flyback has the following drawbacks.

4.3.4.1 Large output voltage spikes. At the end of the on time, the peak primary current is given by Eq. 4.9. Immediately after the end of

TABLE 4.2 Maximum Number of Turns and Maximum Inductance at Those Turns for Various Peak Currents I_p at a Maximum Inductance Falloff of 10 Percent from Zero Current Level

Magnetics Inc. core number	Permeability	A_l, mH per 1000 turns	Maximum oersted level $\overline{H}$ for 10% falloff in inductance	$\overline{NI}$ Maximum permissible ampere turns corresponding to $\overline{H}$	Maximum permissible turns and inductance at those turns for a 10% inductance falloff at indicated peak currents							
					N_{max}/L_{max}							
Core	μ	A_l	H	$\overline{NI}$	1A	2A	3A	5A	10A	20A	50A	I_p
55206	125	68	19	77	77	39	26	15	8	4	2	N_{max}
					363	93	41	14	4	1	0.24	L_{max}
55848	60	32	39	158	158	79	53	32	16	8	3	N_{max}
					719	180	81	29	7	2	0.26	L_{max}
55208	26	14	95	385	385	193	128	77	39	19	8	N_{max}
					1,868	469	206	75	19	4.5	0.8	L_{max}
55209	14	7.8	170	689	689	345	230	138	69	34	14	N_{max}
					3,333	836	371	134	33	8	1.4	L_{max}

Note: Magnetics Inc. MPP cores: OD = 0.8 in, ID = 0.5 in, height = 0.25 in, l_m = 5.09 cm. All inductances in microhenries.

TABLE 4.3 Maximum Number of Turns and Maximum Inductance at Those Turns for Various Peak Currents I_p at a Maximum Inductance Falloff of 10 Percent from Zero Current Level

Core	μ	A_l	$\overline{H}$	$\overline{NI}$	1A	2A	3A	5A	10A	20A	50A	I_p
55438	125	281	19	162	162	81	54	32	16	8	3	N_{max}
					6,637	1,659	737	259	65	16	2	L_{max}
55439	60	135	39	333	333	167	111	67	33	17	7	N_{max}
					13,473	3,389	1,497	545	132	35	6	L_{max}
55440	26	59	95	812	812	406	271	162	81	41	16	N_{max}
					35,011	8,753	3,900	1,394	348	89	14	L_{max}
55441	14	32	170	1454	1,454	727	485	291	145	73	29	N_{max}
					60,744	15,222	6,774	2,439	605	153	24	L_{max}

Note: Magnetics Inc. MPP cores: OD = 1.84 in, ID = 0.95 in, height = 0.71 in, l_m = 10.74 in.
All inductances in microhenries.

the on time, that primary peak current, multiplied by the turns ratio N_p/N_s is driven into the secondary where it decays linearly as shown in Fig. 4.1c. In most cases, output voltages are low and input voltages are higher, resulting in a large N_p/N_s ratio and a consequent large secondary current.

At the initial start of turnoff, the impedance looking down into C_o is much lower than R_o (Fig. 4.1) and all the large secondary current flows down into C_o and its equivalent series resistor R_{esr}. This produces a large, thin output voltage spike, $I_p(N_p/N_s)R_{esr}$. The spike is generally less than 0.5 μs in width as it is differentiated with a time constant of $R_{esr}C_o$.

Frequently a power supply specification specifies output voltage ripple only as an rms or peak-to-peak fundamental value. Such a large, thin spike has a very low rms value and if a sufficiently large output filter capacitor is chosen, the supply can easily meet its rms ripple specification but can have disastrously high, thin output spikes. It is common to see a 50-mV fundamental peak-to-peak output ripple with a 1-V thin spike sitting on top of it.

Thus, almost always, a small *LC* filter is added after the main storage capacitor in flybacks. The *L* and *C* are quite small as they have to filter out a spike generally less than 0.5 μs in width. The inductor is considerably smaller than the inductor in forward-type converters, but it still has to be stocked and have space provided for it. Output voltage sensing for the error amplifier is taken before this *LC* filter.

4.3.4.2 Large output filter capacitor and high ripple current requirement.

A flyback filter capacitor is much larger than one for a forward-type converter. In a forward converter, when the power transistor turns off (Fig. 2.10), load current is supplied from the energy stored in both the filter inductor and the filter capacitor. But in the flyback, that capacitor is necessarily larger as it is the stored energy in it alone that supplies current to the load during the transistor on time. Output ripple is determined mostly by the ESR of the filter capacitor (see Sec. 1.3.7). An initial selection of the filter capacitor is made on the basis of output ripple specification from Eq. 1.10.

Frequently, however, it is not the output ripple voltage requirement that determines the final choice of the filter capacitor. Ultimately it may be the ripple current rating of the capacitor selected initially on the basis of the output ripple voltage specification.

In a forward-type converter (as in a buck regulator), the capacitor ripple current is greatly limited by the output inductor in series with it (Sec. 1.3.6). But in a flyback, the full DC current flows up from ground through the capacitor during the transistor on time. During the transistor off time, a current of equal (ampere)(time) product must

flow down into the capacitor to replenish the charge it lost during the on time. Assuming, as in Fig. 4.1, a sum of on time plus reset time of 80 percent of a full period, the rms ripple current in the capacitor is closely

$$I_{rms} = I_{dc}\sqrt{\frac{t_{on}}{T}} = I_{dc}\sqrt{0.8} = 0.89I_{dc} \tag{4.15}$$

If the initial capacitor selected on the basis of output ripple voltage specifications did not also have the ripple current rating of Eq. 4.15, a larger capacitor or more units in parallel are chosen.

4.3.5 Flybacks for 120- or 220-V-AC operation with no doubler or full-wave rectifier switching

In Sec. 3.2.1, a universally used scheme that permitted operation off either a 120-V-AC or 220-V-AC line with minimum changes was described. As seen in Fig. 3.1, at 120 V AC, switch $S1$ is thrown to the lower position and the circuit is a voltage doubler yielding a rectified voltage of 336 V. With 220 V AC, $S1$ is thrown to the upper position and the circuit is a full-wave rectifier with $C1$ and $C2$ in series yielding closely about 308 V. The converter is now designed to always work from a rectified nominal 308 to 336 V DC by proper choice of the transformer turns ratio.

In some usages, it is undesirable to have to throw $S1$ from one position to the other in changing from 120 to 220 V AC operation. To change switch position without opening the power supply case, the switch must be accessible externally, and this is a safety hazard. The alternative is not to do the switch changing externally, but this requires opening the power supply case to make the change, and this is a nuisance. Further, there is always the possibility that the switch is mistakenly thrown to the voltage doubling position when operated off 220 V AC. This, of course, would cause significant failures—the power transistor, rectifiers, and filter capacitors would be destroyed.

An alternative is not to do any switching at all whether the input is 115 or 220 V AC. Accept a rectified 160 V DC when the input is 115 V AC or a rectified 310 V DC from 220 V AC. Now a flyback converter with a small primary/secondary turns ratio can be used to ensure that the off-voltage stress at high AC input does not overstress the power transistor.

The maximum on time $\overline{T_{on}}$ at the minimum value of the 220-V-AC input is calculated from the corresponding minimum rectified DC input as in Eq. 4.7 and the rest of the magnetics design can proceed as shown in the text following Eq. 4.7. The minimum on time occurs at

the maximum value of the 220-V-AC input. Since the feedback loop keeps the product of $V_{dc}T_{on}$ constant (Eq. 4.3), minimum on time is $T_{on} = \overline{T_{on}}(V_{dc}/\overline{V_{dc}})$ where V_{dc} and $\overline{V_{dc}}$ correspond to the minimum and maximum values of the 220-V-AC line.

The maximum on time with 115-V-AC input is still given by Eq. 4.7 and will be greater as the term $V_{dc} - 1$ is smaller. But the primary inductance L_p given by Eq. 4.8 is seen to be proportional to the product $V_{dc}T_{on}$ is still the same as that product is kept constant by the feedback loop. So long as the transistor can operate with the minimum on time calculated for the maximum DC corresponding to high AC input, there is no problem. With bipolar transistors operating at a high frequency, transistor storage time could prevent operation at too low an on time. An example will clarify this.

Thus Eq. 4.4 gives the maximum off stress in terms of the maximum DC input voltage, the output voltage, and the N_p/N_s turns ratio. Assume in that equation that $\overline{V_{ms}}$ is 500; many bipolar transistors can safely sustain that voltage with a negative base bias at turnoff (V_{cev} rating). At 220 V AC, nominal V_{dc} is 310 V. Assume that the maximum at high line and a worst-case transient is 375 V. Then for a 5-V output, Eq. 4.4 gives a turns ratio of 21.

Now assume that minimum DC supply voltage is 80 percent of nominal. Assume a switching frequency of 50 kHz (period T of 20 μs). Maximum on time is calculated from Eq. 4.7 at the minimum DC input corresponding to minimum AC input of 0.8 × 115 or 92 V AC. For the corresponding DC input of 1.41 × 92 or about 128 V, maximum on time calculated from Eq. 4.7 is 7.96 μs.

Minimum on time occurs at maximum DC input at highest AC input. Assuming a 20 percent high line, maximum DC input is 1.2 × 220 × 1.41 = 372 V. Since the feedback loop keeps the product of $V_{dc}T_{on}$ constant, (Eq. 4.3), on time at the 20 percent high line of 264 V AC is (128/372)(7.96) or 2.74 μs. The circuit can thus cope with either a 20 percent low AC line input of 92 V AC from a nominal 115 V AC or a 20 percent high AC input of 264 V AC from the nominal 220 V AC line by readjusting its on time from 7.96 to 2.74 μs.

But it is obvious that if this were attempted at higher switching frequencies, the minimum on time at a 220-V-AC line condition would become so low as to prohibit the use of bipolar transistors which could have 0.5- to 1.0-μs storage time. The upper-limit switching frequency at which the above scheme can be used with bipolar transistors is about 100 kHz.

It is instructive to complete the above design. Assume an output power of 150 W at 5-V output. Then $R_o = 0.167\ \Omega$ and the primary inductance from Eq. 4.8 is

$$L_p = \left(\frac{0.167}{2.5 \times 20 \times 10^{-6}}\right)\left(\frac{128 \times 7.96 \times 10^{-6}}{5}\right)^2$$
$$= 139\ \mu H$$

and the peak primary current from Eq. 4.9 is

$$I_p = \frac{128 \times 7.96 \times 10^{-6}}{139 \times 10^{-6}} = 7.33\ \text{A}$$

There are many reasonably priced bipolar transistors with a V_{cev} rating above 500 V having adequate gain at 7.33 A.

Table 4.1 shows that the 55932 MPP core can tolerate a maximum of 480 ampere turns beyond which its inductance will fall off by more than 10 percent (at 5 A, column 9 shows that its maximum turns is 95 for a maximum inductance of 265 μH). For a maximum ampere turns of 480 at 7.33 A, its maximum turns is 480/7.33 = 65.4. For 66 turns, its inductance is $32{,}000 \times 0.9(66/1000)^2 = 125\ \mu H$. If (as discussed in Sec. 4.2.3.2) 10 percent more turns are added, the inductance at 7.33 A will increase by 10 percent to 138 μH, but at zero current, the inductance will "swing" up to 20 percent above that.

If the 20 percent inductance swing is undesirable, the lower permeability core 55933 of Table 4.1 can be used. Table 4.1 shows that its maximum ampere turns is 859. For 7.33 A, the maximum number of turns is 859/7.33 or 117. Its maximum inductance for only a 10 percent swing is $(0.117)^2 \times 18000 \times 0.9$ or 222 μH. For the desired 139 μH, required turns are $1000\sqrt{0.139/18 \times 0.95} = 90$ turns.

Thus a design not requiring voltage doubling/full-wave rectifier switching when operation is changed from 115 to 220 V AC is possible. But this subjects the power transistor to about 500 V plus a leakage inductance spike at turnoff. The reliability of this must be weighed against the use of a double-ended forward converter or half bridge—both of which subject the off transistor to only the maximum DC input (375 V in the preceding example) with no leakage spike. Of course, for 115/220-V-AC operation, the rectifier switching of Fig. 3.1 must be accepted.

4.4 Continuous-Mode Flybacks—Basic Operation

There are two distinctly different modes of flyback converters (discontinuous and continuous) whose waveforms are shown in Fig. 4.5.

Both modes have an identical circuit diagram, shown in Fig. 4.1, and it is only the transformer's magnetizing inductance and output load current which determines its operating mode. With a given mag-

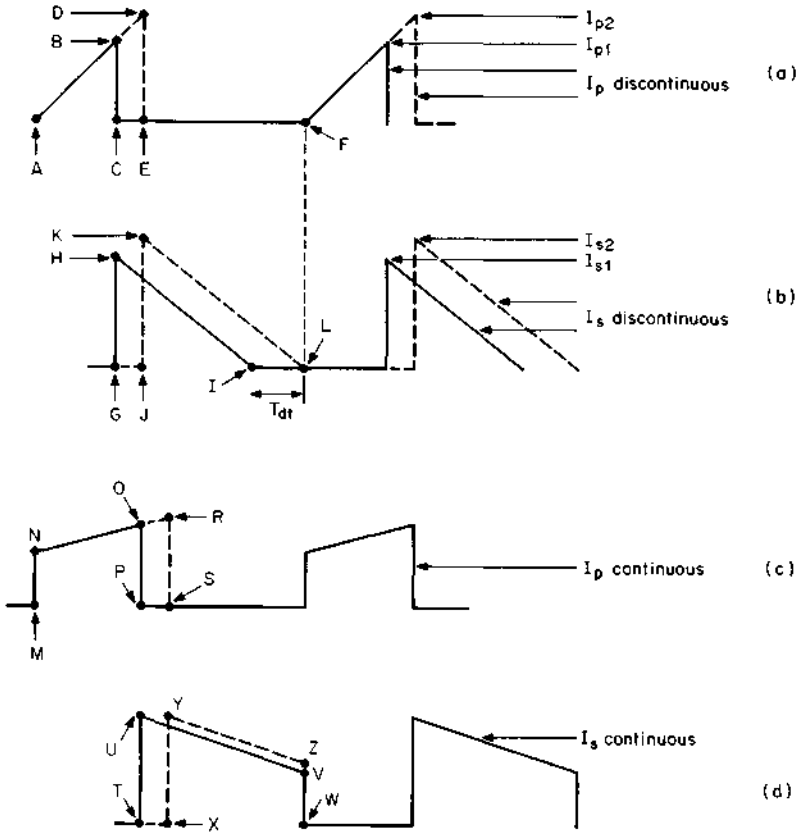


Figure 4.5 A discontinuous-mode flyback in transition to continuous-mode. Circuit is discontinuous if there is a dead time (T_{dt}) between the instant the secondary current reaches zero and the start of the next turnon. If loaded beyond this point, there is still some current left in the secondary at the end of the turnoff and the next turnon has a step at its front end. This is characteristic of the continuous mode, and if the error-amplifier bandwidth has not been drastically reduced, the circuit will oscillate.

netizing inductance, a circuit which has been designed for the discontinuous mode will move into the continuous mode when the output load current is increased beyond a unique boundary. The mechanism for this and its consequence is discussed below.

The discontinuous mode as shown in Fig. 4.5*a* has no front-end step in its primary current, and at turnoff as shown in Fig. 4.5*b*, the secondary current is a decaying triangle which has ramped down to zero before the next turnon. All the energy stored in the primary during the on time has been completely delivered to the secondary and thus to the load before the next turnon.

In the continuous mode, as seen in Fig. 4.5*c*, the primary current has a front-end step and the characteristic appearance of a rising ramp on a step. During the transistor off time (Fig. 4.5*d*), the second-

ary current has the shape of a decaying triangle sitting on a step with current still remaining in the secondary at the instant of the next turnon. There is thus still some energy left in the secondary at the instant of the next turnon.

The two modes have significantly different operating properties and usages. The discontinuous mode responds more rapidly and with a lower transient output voltage spike to sudden changes in load current or input voltage.

But the secondary peak current in the discontinuous mode can be between two and three times that in the continuous mode, as can be seen in Fig. 4.5*b* and 4.5*d*. Secondary DC load current is the average of the current waveshapes in those figures. Also, assuming closely equal off times, it is obvious that the triangle in the discontinuous mode must have a much larger peak than the trapezoid of the continuous mode for the two waveshapes to have equal average values.

With a larger peak secondary current, the discontinuous mode will have a larger transient output voltage spike at the instant of turnoff (Sec. 4.3.4.1) and will require a larger *LC* spike filter to remove it. Also, the larger secondary peak current at the start of turnoff in the discontinuous mode will cause a greater RFI problem. Even for moderate output powers, the very large initial spike of secondary current at the instant of turnoff causes a much more severe noise spike on the output ground bus because of the large *di/dt* into the output bus inductance.

Secondary rms currents in the discontinuous mode can be up to twice that in the continuous mode. This requires larger secondary wire size and output filter capacitors with larger ripple current ratings for the discontinuous mode. Rectifier diodes will also have larger temperature rise in the discontinuous mode because of the larger secondary rms currents.

Primary peak currents for the discontinuous mode are about twice those in the continuous mode—again because for equal average values, the triangle of Fig. 4.5*a* must have a larger peak than the trapezoid of Fig. 4.5*c*. The consequence is that the discontinuous mode with its larger peak primary current requires a higher current rating and possibly a more expensive power transistor. Also, the higher primary current in the discontinuous mode results in a greater RFI problem.

Despite all these relative disadvantages for the discontinuous mode, it is much more widely used than the continuous mode. This is so for two reasons. First, as mentioned above, the discontinuous mode, with an inherently smaller transformer magnetizing inductance, responds more quickly and with a lower transient output voltage spike to rapid changes in output load current or input voltage. Second, because of a unique characteristic of the continuous mode (its transfer function has

a right-half-plane-zero, to be discussed in a later chapter on feedback loop stabilization), the error amplifier bandwidth must be drastically reduced to stabilize the feedback loop.

4.4.1 Discontinuous-mode to continuous-mode transition

In Fig. 4.5*a* and 4.5*b* the solid lines represent primary and secondary currents in the discontinuous mode. Primary current is a triangle starting from zero and rising to a level I_{p1} (point B) at the end of the power transistor on time (see Fig. 4.5*a*).

At the instant of turnoff, the current I_{p1} stored in the primary is dumped into the secondary at point H as $I_{s1} = (N_p/N_s)I_{p1}$. Since the dot end of the secondary is positive during the off time, secondary current ramps downward at a rate $dI_s/dt = (V_o + 1)/L_s$, where L_s is the secondary inductance, which is $(N_s/N_p)^2$ times the primary magnetizing inductance. This current reaches zero at time I, leaving a dead time T_{dt} before the start of the next turnon at point F. All the current or energy stored in the primary during the on time has now been completely delivered to the load before the next turnon. The average or DC output current is the average of the triangle GHI multiplied by its duty cycle of T_{off}/T.

Now, to remain in the discontinuous mode, there must be a dead time T_{dt} (Fig. 4.5*b*) between the time the secondary current has dropped to zero and the start of the next power transistor on time. As more power is demanded (by decreasing R_o), T_{on} must increase to keep output voltage constant (see Eq. 4.3). As T_{on} increases (at constant V_{dc}), primary current slope remains constant and peak current rises from B to D as in Fig. 4.5*a*. Secondary peak current ($= I_pN_p/N_s$) increases from H to K in Fig. 4.5*b* and starts later in time (from G to J).

As output voltage is kept constant by the feedback loop, the secondary slope V_o/L_s remains constant and the point at which the secondary current falls to zero moves closer to the start of the next turnon. This reduces T_{dt} until a point L is reached where the secondary current has just fallen to zero at the instant of the next turnon. This is the load current which marks the end of the discontinuous mode. Dead time can also be completely lost at constant output power (current) as V_{dc} decreases. The on time T_{on} must also increase as V_{dc} decreases to maintain constant output voltage.

Note that as long as the circuit is in the discontinuous mode and so long as there remains a dead time, increasing the on time increases the area of the primary current triangle and also the area of the secondary triangle GHI up to the limit of the area JKL. Further, since DC output current is the average of the secondary current triangle

multiplied by its duty cycle, then the very next off time after an increase in on time, more secondary current is available to the load.

When the dead time has been lost, however, any further increase in load current demand will increase the on time and decrease the off time as the back end of the secondary current can no longer move to the right. The secondary current will start later than point *J* (Fig. 4.5*b*) and from a higher point than *K*. Then at the start of the next on time (*F* in Fig. 4.5*a* or *L* in Fig. 4.5*b*), there is still some current or energy left in the secondary.

Now the front end of the primary current will have a small step. The feedback loop tries to deliver the increased DC load current demand by keeping the on time later than point *J*. Now at each successive off time, the current remaining at the end of the off time and hence the current step at the start of the next on time increases.

Finally after many switching cycles, the front-end step of primary current and the back end current at the end of the off time in Fig. 4.5*d* are sufficiently high so that the area *XYZW* is somewhat larger than that sufficient to supply the output load current. Now the feedback loop starts to decrease the on time so that the primary trapezoid lasts from *M* to *P* and the secondary current trapezoid lasts from *T* to *W* (Figs. 4.5*a* and 4.5*b*).

At this point, the on volt-seconds across the transformer primary when the power transistor is on is equal to the off volt-seconds across it when the transistor is off. This is the condition that the transformer core is always reset to its original point on the hysteresis loop at the end of a full cycle. Or, what amounts to the same thing, it is the condition that the average or DC voltage across the primary is zero since it is assumed that the DC resistance in the primary is zero and it is not possible to support a DC voltage across zero resistance.

Once in the continuous mode, increased load current is supplied initially by an increase in on time (from *MP* to *MS* in Fig. 4.5*c*). This, oddly enough, results in a decrease in off time from *TW* to *XW* (Fig. 4.5*d*) as the back end of the secondary current pulse cannot move further to the right in time because the dead time has vanished. Although the peak of the secondary current has increased somewhat (from point *U* to *Y*), the area lost in the decreased off time (*T* to *X*) is greater than the area gained in the slope change from *UV* to *YZ* in Fig. 4.5*d*.

Thus, in the continuous mode, a sudden increase in DC output current initially causes a decrease in width and a smaller increase in height of the secondary current trapezoid. After many switching cycles, the average trapezoid height builds up and the width relaxes back to the point where the on volt-seconds again equals the off volt-seconds across the primary.

In addition, since the DC output voltage is proportional to the area of the secondary current trapezoid, the feedback loop, in attempting to keep the output voltage constant against an increased current demand, first drastically decreases the output voltage and then, after many switching cycles, corrects it by building up the amplitude of the secondary current trapezoid. This is the physical-circuits significance of the so-called right-half-plane-zero which forces the drastic reduction in error-amplifier bandwidth to stabilize the feedback loop. The right-half-plane-zero will be discussed further in the chapter on loop stabilization.

4.4.2 Design relations—continuous-mode flybacks

4.4.2.1 Relation between output voltage and on time. Refer to Fig. 4.1. When the transistor is on, the voltage across the primary is closely $V_{dc} - 1$ with the dot end negative with respect to the no-dot end and the core is driven—say, up the hysteresis loop. With the transistor off, the magnetizing current, which is trying to remain constant, reverses the polarity of all voltages. The primary and secondary try to go far positive. But the secondary is caught and clamped to $V_{om} + 1$—assuming a 1-V forward drop across $D2$.

This reflects across to the primary as a voltage $(N_p/N_s)(V_{om} + 1)$ with the dot end now positive with respect to the no-dot end. All the current flowing in the primary (I_{PO} in Fig. 4.5*c*) now transfers to the secondary as I_{TU} in Fig. 4.5*d*. The initial magnitude of the secondary current I_{TU} is equal to the final primary current at the end of the on time (I_{PO}) times the turns ratio N_p/N_s. Since the dot end of the secondary is now positive with respect to the no-dot end, the secondary current ramps downward as the slope UV in Fig. 4.5*d*.

Since the primary is assumed to have zero DC resistance, it cannot sustain a DC voltage averaged over many cycles. Thus in the steady state, the volt-second product across it when the transistor is on must equal that across it when the transistor is off—i.e., the voltage across the primary averaged over a full cycle must equal zero. This is equivalent to saying the core's downward excursion on the BH loop during the off time is exactly equal to the upward excursion during the on time. Then

$$(\underline{V_{dc}} - 1)\,\overline{t_{on}} = (V_{om} + 1)\frac{N_p}{N_s}\,t_{off}$$

or

$$V_{om} = \left[\left(\underline{V_{dc}} - 1\right)\frac{N_s}{N_p}\frac{\overline{t_{on}}}{t_{off}}\right] - 1 \qquad (4.16)$$

and since now there is no dead time, $\overline{t_{on}} + t_{off} = T$, then

$$V_{om} = \left[\frac{(V_{dc} - 1)(N_s/N_p)(\overline{t_{on}}/T)}{1 - \overline{t_{on}}/T}\right] - 1 \tag{4.17a}$$

$$= \left[\frac{(V_{dc} - 1)(N_s/N_p)}{(T/\overline{t_{on}}) - 1}\right] - 1 \tag{4.17b}$$

The feedback loop regulates against DC input voltage changes by decreasing T_{on} as V_{dc} increases, or increasing t_{on} as V_{dc} decreases.

4.4.2.2 Input, output current–power relations. In Fig. 4.6, the output power is equal to the output voltage times the average of the secondary current pulses. For I_{csr} equal to the current at the center of the ramp in the secondary current pulse

$$P_o = V_o I_{csr} \frac{t_{off}}{T}$$

$$= V_o I_{csr} (1 - \overline{t_{on}}/T) \tag{4.18}$$

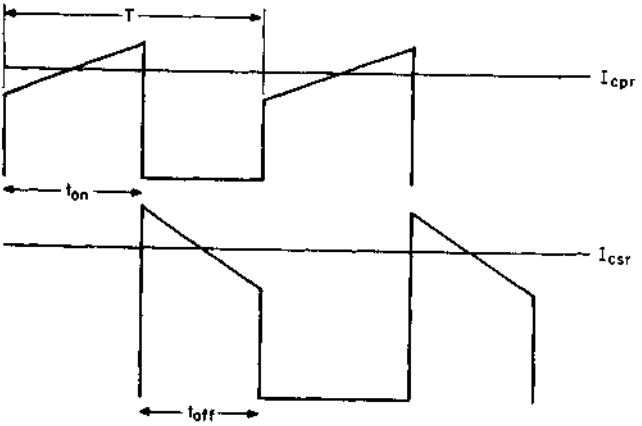


Figure 4.6 Current–on-time relations in a continuous-mode flyback. Current is delivered to the load only during the off time. At a fixed DC input voltage, t_{on} and t_{off} remain constant. Output load current changes are accommodated by the feedback loop changing the magnitude of the current at the center of the primary current ramp I_{cpr} and at the center of the secondary current ramp (I_{csr}). This occurs over many switching cycles by temporary increases in on time until the average current pulse amplitudes build up and then relax to the steady-state values of t_{on} and t_{off}.

or
$$I_{csr} = \frac{P_o}{V_o(1 - \overline{t_{on}}/T)} \tag{4.19}$$

In Eqs. 4.18 and 4.19, $\overline{t_{on}}/T$ is given by Eq. 4.17 for specified values of V_{om} and V_{dc} and a turns ratio N_s/N_p from Eq. 4.4 chosen for an acceptably low maximum off-voltage stress at maximum DC input.

Further, for an assumed efficiency of 80 percent, $P_o = 0.8P_{in}$ and I_{cpr} equal to the current at the center of the ramp in the primary current pulse:

$$P_{in} = 1.25P_o = V_{dc}I_{cpr}\frac{\overline{t_{on}}}{T}$$

or
$$I_{cpr} = \frac{1.25P_o}{(V_{dc})(\overline{t_{on}}/T)} \tag{4.20}$$

4.4.2.3 Ramp amplitudes for continuous mode at minimum DC input. It has been pointed out that the start of continuous-mode operation occurs when there is just the beginning of a step at the front end of the primary current ramp. Referring to Fig. 4.6, the step begins to appear when the current at the center of the primary ramp I_{cpr} is equal to half the ramp amplitude dI_p. That value of I_{cpr} ($\underline{I_{cpr}}$) is then the minimum value at which the circuit is still in the continuous mode. From Eq. 4.20, I_{cpr} is proportional to output power and hence for the minimum output power $\underline{P_o}$ corresponding to $\underline{I_{cpr}}$

$$\underline{I_{cpr}} = \frac{dI_p}{2} = \frac{1.25\underline{P_o}}{(\underline{V_{dc}})(\overline{t_{on}}/T)}$$

or
$$dI_p = \frac{2.5\underline{P_o}}{(\underline{V_{dc}})(\overline{t_{on}}/T)} \tag{4.21}$$

In Eq. 4.21, $\overline{t_{on}}$ is taken from Eq. 4.17 at the corresponding value minimum of V_{dc} ($\underline{V_{dc}}$). The slope of the ramp dI_p is given by $dI_p = (\underline{V_{dc}} - 1)\overline{t_{on}}/L_p$, where L_p is the primary magnetizing inductance. Then

$$L_p = \frac{(\underline{V_{dc}} - 1)\overline{t_{on}}}{dI_p} = \frac{(\underline{V_{dc}} - 1)(\underline{V_{dc}})(\overline{t_{on}})^2}{2.5P_oT} \tag{4.22}$$

Here again, P_o is the minimum specified value of output power and $\overline{t_{on}}$ is the maximum on time calculated from Eq. 4.17 at the minimum specified DC input voltage V_{dc}.

4.4.2.4 Discontinuous and continuous-mode flyback design example. It is instructive to compare a discontinuous- and continuous-mode flyback design at the same output power levels and input voltages. The magnitudes of the currents and primary inductances will be revealing.

Assume a 50-W, 5-V output flyback converter operating at 50 kHz from a telephone industry prime power source (38 V DC minimum, 60 V maximum). Assume a minimum output power of one-tenth the nominal or 5 W.

Consider first a discontinuous-mode flyback. Assume a bipolar transistor with a 150-V V_{ceo} rating and, wishing to be very conservative, do not bank on the V_{cer} or V_{cev} rating which permits a larger voltage. Then in Eq. 4.4, assume that the maximum off-voltage stress $\overline{V_{ms}}$ without a leakage spike is 114 V, which permits a 36-V leakage spike before the V_{ceo} limit is reached. Then Eq. 4.4 gives $N_p/N_s = (114 - 60)/6 = 9$.

Now Eq. 4.7 gives the maximum on time as

$$\overline{t_{on}} = 6 \times 9 \times 0.8 \frac{20 \times 10^{-6}}{37 + 6 \times 9}$$

$$= 9.49\ \mu s$$

and primary inductance for $R_o = 5/10 = 0.5\ \Omega$ from Eq. 4.8 is

$$L_p = \frac{0.5}{2.5 \times 20^{-6}} \left(\frac{38 \times 9.49}{5}\right)^2 \times 10^{-12}$$

$$= 52\ \mu H$$

Peak primary current from Eq. 4.9 is

$$I_p = \frac{38 \times 9.49 \times 10^{-6}}{52 \times 10^{-6}}$$

$$= 6.9\ A$$

and the front end of the secondary current triangle is

$$I_{s(peak)} = (N_p/N_s)I_p = 9 \times 6.9 = 62\ A$$

Recall that in the discontinuous flyback, the reset time T_r or the time for the secondary current to decay back to zero plus the maximum on time is equal to $0.8T$ (Eq. 4.6). Reset time is then

$T_r = (0.8 \times 20) - 9.49 = 6.5$ μs, and the average value of the secondary current triangle (which should equal the DC output current of 10 A is

$$I \text{ (secondary average)} = \frac{I_{s(\text{peak})}}{2}\frac{T_r}{T}$$

$$= \left(\frac{62}{2}\right)\frac{6.5}{20} = 10 \text{ A}$$

which truly is the DC output current.

Now consider the continuous-mode flyback for the same frequency, input voltages, and output power output voltage and the same N_p/N_s ratio of 9. From Eq. 4.17b, calculate $\overline{t_{on}}/T$ for $\underline{V_{dc}} = 38$ V as

$$5 = \left[\frac{(37/9)(\overline{t_{on}}/T)}{1 - \overline{t_{on}}/T}\right] - 1$$

or $\overline{t_{on}}/T = 0.5934$ and $\overline{t_{on}} = 11.87$ μs, $t_{off} = 8.13$ μs and from Eq. 4.19

$$I_{csr} = \frac{50}{(5)(1 - 0.5934)} = 24.59 \text{ A}$$

and the average of the secondary current pulse which should equal the DC output current is I (secondary average) $= I_{csr}(t_{off}/T) = 24.59 \times 8.13/20 = 10.0$ A, which checks. From Eq. 4.20 $I_{cpr} = 1.25 \times 50/(38)(11.86/20) = 2.77$ A.

From Eq. 4.22, for the minimum input power of 5 W at the minimum DC input voltage of 38 V, $L_p = 37 \times 38(11.86)^2 \times 10^{-12}/2.5 \times 5 \times 20 \times 10^{-6} = 791$ μH.

The contrast between the discontinuous and continuous mode can now be seen in the table below in the comparison between the required primary inductances and primary and secondary currents at minimum DC input of 38 V.

	Discontinuous	Continuous
Primary inductance, μH	52	791
Primary peak current, A	6.9	2.77
Secondary peak current, A	62.0	24.6
On time, μs	9.49	11.86
Off time, μs	6.5	8.13

The lower primary current and especially the secondary current for the continuous mode are certainly an advantage. But the much larger primary inductance which slows up response to load current changes and the right-half-plane-zero, which requires a very low error-amplifier bandwidth to achieve loop stabilization, make the continuous mode a less desirable choice.

4.5 Interleaved Flybacks

Interleaved flyback topology is shown in Fig. 4.7. It consists simply of two discontinuous-mode flybacks whose power transistors are turned on at alternate half cycles and whose secondary currents are summed through their rectifying diodes.

It can be used at power levels up to 300 W, limited mainly by the high peak primary and especially secondary currents. Although that power level can be reached with a single continuous-mode flyback at reasonable currents, it is more advantageous to accept the greater cost

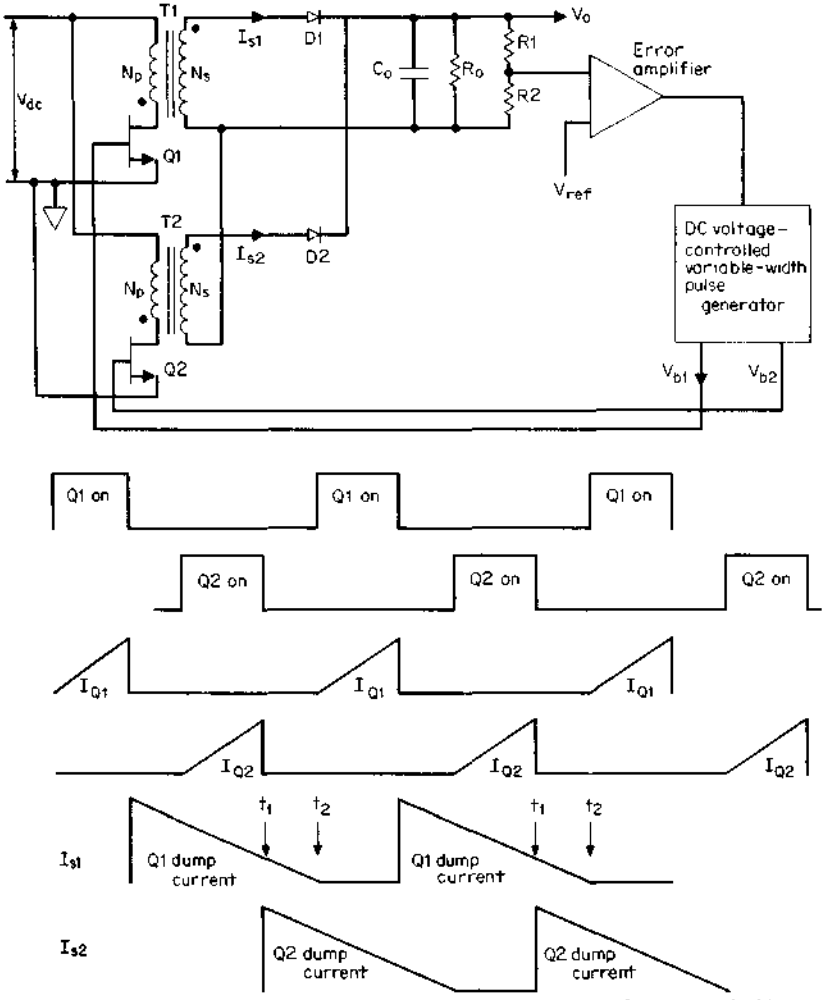


Figure 4.7 Interleaving two discontinuous-mode flyback on alternative half cycles to reduce peak currents. Output powers of up to 300 W are possible with reasonably low peak currents.

and required volume of two interleaved discontinuous-mode flybacks—if the flyback topology is chosen for some particular reason.

This is so because of the discontinuous mode's faster response to load current changes, greater error-amplifier bandwidth, and lesser feedback loop stabilization problems. A single discontinuous-mode flyback at the 300-W level is impractical because of the high peak primary and secondary currents as can be seen from Eqs. 4.2, 4.7, and 4.8.

Actually, at the 150-W level, a single forward converter is very likely a better choice than the two interleaved flybacks because of the considerably lower secondary current of the forward converter. But the interleaved flyback is shown here for the sake of completeness and its possible use at lower power levels than 150 W, but where many (over five) outputs are required.

4.5.1 Summation of secondary currents in interleaved flybacks

The magnetics design of each flyback in an interleaved flyback proceeds exactly as for a single flyback at half the power level. For the secondary currents dump into the output through their "ORing" rectifier diodes and add.

Even when both secondary diodes dump current simultaneously (as from t_1 to t_2), there is no possibility that one diode can back-bias the other and supply all the load current. This can happen when two low-impedance voltage sources attempt to sum currents simultaneously. For with low-impedance voltage sources, one source with only say 0.1 V higher open-circuit voltage or with a lower forward-drop OR diode, will back-bias the other diode and supply all the load current by itself. This can overdissipate the diode or the transistor supplying that diode.

Looking back into the secondary of a flyback, however, there is seen a high-impedance current source. Thus the current dumped out into the common load by either diode is unaffected by the other diode simultaneously supplying load current.

4.6 Double-Ended Discontinuous-Mode Flyback

4.6.1 Area of Application

The topology is shown in Fig. 4.8*a*. Its major advantage is that, using the scheme of the double-ended forward converter of Fig. 2.13, its power transistors in the off state are subjected to only the maximum DC input voltage. This is a significant advantage over the single-ended forward converter of Fig. 4.1, where the maximum off-voltage stress is the maximum DC input voltage plus the reflected secondary

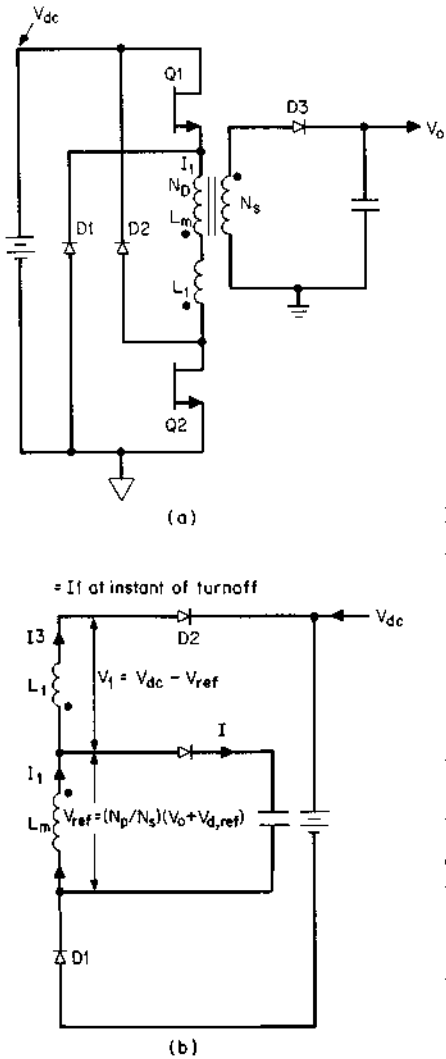


Figure 4.8 Circuit during $Q1$ and $Q2$ off time. Current I_1, stored in L_m during $Q1$, $Q2$ on time, also flows through leakage inductance L_l. During the off time, energy stored in L_m must be delivered to the secondary load as reflected into the primary across L_m. But I_1 also flows through L_l, and during the off time, the energy it represents $\frac{1}{2}L_l I^2$) is returned to the input source V_{dc} through diodes $D1$, $D2$. This robs energy which should have been delivered to the output load and continues to rob energy until I_1, the leakage inductance current, falls to zero. To minimize the time for I_1 in L_l to fall to zero, V_r is made significantly large by keeping the reflected voltage $V_r (= N_p/N_s)(V_o + V_{D3})$ low by setting a low N_p/N_s turns ratio. A usual value for V_r is two-thirds of the minimum V_{dc}, leaving one-third for V_l.

voltage $(N_p/N_s)(V_o + 1)$ plus a leakage inductance spike which may be as high as one third of the DC input voltage.

The price paid for this advantage is, of course, the requirement for two transistors and the two clamp diodes, $D1$, $D2$.

4.6.2 Basic operation

The lower off-voltage stress comes about in the same way as for the double-ended forward converter of Fig. 2.13. Power transistors $Q1$, $Q2$ are turned on simultaneously. When they are on, the dot end of the secondary is negative, $D3$ is reverse-biased, and no secondary current flows. The primary is then just an inductor, and current in it ramps up

linearly at a rate of $dI_1/dt = V_{dc}/(L_m + L_l)$, where L_m and L_l are the primary magnetizing and leakage inductances, respectively.

When $Q1$ and $Q2$ turn off, as in the previous flybacks, all primary and secondary voltages reverse polarity, $D3$ becomes forward-biased, and the stored current or energy in $L_m(= 1/2L_m(I_1)^2)$ is delivered to the load.

As always, the on or set volt-second product across the primary must equal the off or reset volt-second product. At the instant of turnoff, the bottom end of L_l attempts to go very far positive but is clamped to the positive end of V_{dc}. The top end of L_m attempts to go far negative but is clamped to the negative end of V_{dc}. Thus the maximum voltage stress at either $Q1$ or $Q2$ can never be more than V_{dc}.

The actual resetting voltage V_r across the mutual inductance L_m during the off time is given by the voltage reflected from the secondary load $(N_p/N_s)(V_o + V_{D3})$. The voltage across L_m and L_l in series is the DC supply voltage and hence as seen in Fig. 4.8*b*, the voltage across the leakage inductance L_l is $V_l = (V_{dc} - V_r)$.

This division of the V_{dc} supply voltage across L_m and L_l in series during the off time is a very important point in the circuit design and establishes the transformer turns ratio N_p/N_s as discussed below.

4.6.3 Leakage inductance effect in double-ended flyback

Figure 4.8*b* shows the circuit during the $Q1$, $Q2$ off time. The voltage across L_m and L_l in series is clamped to V_{dc} through diodes $D1$, $D2$. The voltage V_r across the magnetizing inductance is clamped against the reflected secondary voltage and equals $(N_p/N_s)(V_o + V_{D3})$. The voltage across L_l is then $V_l = V_{dc} - V_r$.

At the instant of turnoff, the same current I_1 flows in L_m and $L_l(I_3 = I_1$ at instant of turnoff). That current in L_l flows through diodes $D1$, $D2$ and returns its stored energy to the supply source V_{dc}. The L_l current decays at a rate of $dI_1/dt = V_l/L_l$ as shown in Fig. 4.9*a* as slope AC or AD. The current in L_m (initially also equal to I_1) decays at a rate V_r/L_m and is shown in Fig. 4.9*a* as slope AB.

Now the current actually delivering power to the load is I_2—the difference between the currents in L_m and L_l. This is shown as current RST in Fig. 4.9*b* if the L_1 current slope is AC of Fig. 4.9*a* or the larger area current UVW in Fig. 4.9*c* if the L_1 current slope is faster as AD of Fig. 4.9*a*. It is obvious, then, in Fig. 4.9*b* and 4.9*c* that so long as current still flows in leakage inductance L_l, through $D1$ and $D2$ back into the supply source, all the current available in L_m does not flow into the reflected load but is partly diverted back into the supply.

It can thus be seen from Fig. 4.9*b* and 4.9*c* that to maximize the transfer of L_m current to the reflected load and to avoid a delay in the

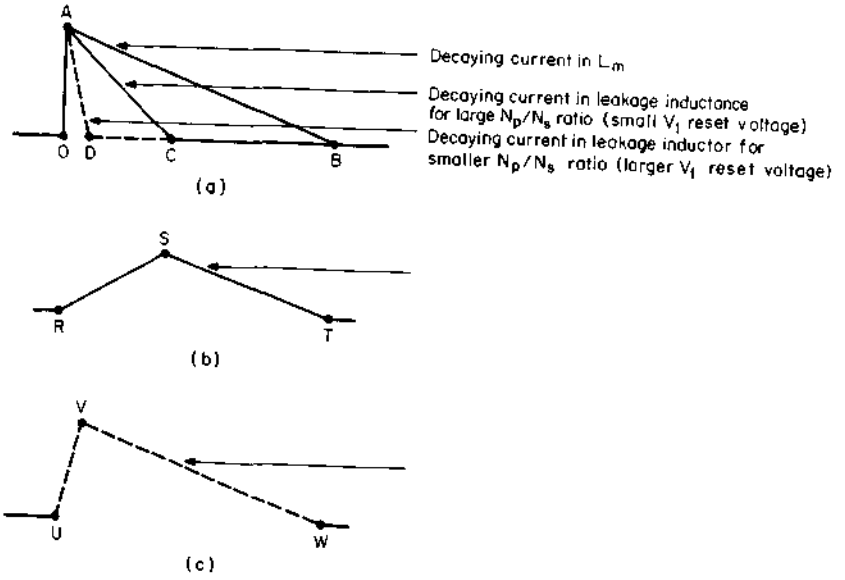


Figure 4.9 (*a*) Currents in magnetizing and leakage inductances in double-ended flyback. (*b*) Current into reflected load impedance for large N_p/N_s ratio. *AB–AC* of Fig. 4.9*a*. (*c*) Current into reflected load impedance for smaller N_p/N_s ratio. *AB–AD* of Fig. 4.9*a*.

transfer of current to the load, the slope of the leakage inductance current decay should be maximized (slope *AD* rather than *AC* in Fig. 4.9*a*). Or in magnetics–power supply jargon, the leakage inductance current should be rapidly reset to zero.

Since the rate of decay of the leakage inductance current is V_l/L_l and $V_l = V_{dc} - (N_p/N_s)(V_o + V_{D3})$, choosing lower values of N_p/N_s increases V_l and hastens leakage current reset. A usual value for the reflected voltage $(N_p/N_s)(V_o + V_{D3})$ is two-thirds of V_{dc}, leaving one-third for V_l. Too low a value for V_r will require a longer time to reset the magnetizing inductance, rob from the available $Q1$, $Q2$ on time, and decrease the available output power.

Once N_p/N_s has been fixed to yield $V_l = V_{dc}/3$, the maximum on time for discontinuous operation is calculated from Eq. 4.7, L_m is calculated from Eq. 4.8 and I_p from Eq. 4.9, just as for the single-ended flyback.

References

Billings, K., *Switchmode Power Supply Handbook*, McGraw-Hill, New York, 1989.

Chryssis, G., *High Frequency Switching Power Supplies*, 2d ed., pp. 122–131, McGraw-Hill, New York, 1989.

Dixon, L., "The Effects of Leakage Inductance on Multi-Output Flyback Circuits," *Unitrode Power Supply Design Seminar Handbook*, Unitrode Corp., Lexington, Mass., 1988.

Patel, R., D. Reilly, and R. Adair, "150 Watt Flyback Regulator," *Unitrode Power Supply Design Seminar Handbook*, Unitrode Corp., Lexington, Mass., 1988.

Chapter

5

Current-Mode and Current-Fed Topologies

5.1 Introduction

Current-mode[1–7] and current-fed[9–20] topologies are grouped into one family despite their very significant differences. They are grouped together because they gain advantages by controlling both input current (although in different ways) and output voltage.

Current mode (Fig. 5.3) has two feedback loops: an outer one which senses DC output voltage and delivers a DC control voltage to an inner loop which senses peak power transistor currents and keeps them constant on a pulse-by-pulse basis. The end result is that it solves the magnetic flux imbalance problem in the current-mode version of the push-pull topology and restores push-pull as a viable approach in applications where the uncertainty of other solutions to flux imbalance is a drawback (Sec. 2.2.8). Further, the constant power transistor current pulses simplify the feedback-loop design.

Current-fed topology drives its power train (center tap of a push-pull or top end of a forward converter transformer) through an input inductor. Thus the power train is driven from high impedance of a current source (the input inductor) rather than the low impedance of a rectifier filter capacitor or the low source impedance of an input battery. This high output impedance of the prime power source very largely solves the flux imbalance problem when a push-pull converter is driven from the input inductor and offers other significant advantages.

In all topologies discussed thus far (voltage-mode circuits), output voltage alone is monitored and controlled directly. In those circuits, regulation against load current changes occur because current changes cause small output voltage changes which are sensed by a voltage-monitoring error amplifier which then corrects the power

transistor on time to maintain output voltage constant. But output current itself is not monitored directly.

About 7 years ago, a new topology, *current mode,* appeared in which both voltage and current could be monitored. The scheme had been known previously but not widely used as it required discrete circuit components to implement it. But when, also 7 years ago, a new pulse-width-modulating (PWM) chip—the UC1846—appeared with all the features needed to implement current mode, the advantages of the technique were quickly recognized and it was widely adopted.

Where two 180° out-of-phase width-modulated outputs are required—as in the push-pull, half-bridge, full-bridge, interleaved forward converter, or flyback—the UC1846—can be used to implement current-mode control. A lower-cost, single-ended PWM controller—the UC1842—is currently available to implement current mode in single-ended circuits such as forward converters, flybacks, and buck regulators.

5.2 Current-Mode Advantages

5.2.1 Avoidance of flux imbalance in push-pull converters

Flux imbalance was discussed in Sec. 2.2.5. It occurs in a push-pull converter when the transformer core operates unsymmetrically about the origin of its hysteresis loop. The consequence is that the core moves up toward saturation and one transistor draws more current during its on time than does the opposite transistor (Fig. 2.4*c*).

As the core drifts further off center of the origin, it goes into deep saturation and may destroy the power transistor. A number of ways to cope with flux imbalance have been described in Sec. 2.2.8. These schemes work, but under unusual line and load transient conditions and especially with higher output powers, there is never complete certainty that flux imbalance cannot occur.

Current mode monitors each current pulse on a pulse-by-pulse basis and forces alternate pulses to have equal peak amplitudes by correcting each transistor's on time so that current amplitudes must be equal. This puts push-pull back into the running in any proposed new design and is a valuable contribution to the reservoir of possible topologies. For example, if to ensure certainty of no flux imbalance, in the absence of current mode, a forward converter which has no flux imbalance problem were chosen, a severe penalty would be paid.

Thus Eq. 2.28 shows the peak primary current in a forward converter is $3.13(P_o/V_{dc})$. But Eq. 2.9 shows it is only half that or $1.56(P_o/V_{dc})$ for the push-pull. At low output powers, it is not a serious drawback to use the forward converter with twice the peak current of a

push-pull at equal output power, especially as the forward converter has only one transistor. But at higher output power, twice the peak primary current in the forward converter as compared to the push-pull becomes prohibitive.

The push-pull is a very attractive choice for telephone industry power supplies where the maximum DC input voltage is specified as only 60 V (38 V minimum). Thus having it available in its current-mode version with a certainty that flux imbalance cannot exist is very valuable.

5.2.2 Instantaneous correction against line voltage changes without the delay in an error amplifier (voltage feedforward characteristics)

It is inherent in the details of how current mode works that a line voltage change immediately causes a change in power transistor on time. This change is corrected for without having to wait for a sensed output voltage change to pass through the relatively long delay in a conventional voltage error amplifier. The details of how this comes about will be discussed below.

5.2.3 Ease and simplicity of feedback-loop stabilization

All the topologies discussed above (with the exception of flybacks) have an output LC filter. An LC filter has a possible maximum phase shift of 180° not far beyond its resonant frequency of $f_o = \frac{1}{2\pi}\sqrt{LC}$, and gain between output and input falls very rapidly with increasing frequency. As frequency increases, the impedance of the series L arm increases and that of the shunt arm decreases.

This large possible phase shift and rapid change of gain with frequency complicates the feedback-loop design. More important, the elements around the error amplifier required to stabilize the loop are more complex and can give rise to problems for rapid changes in input voltage or output current.

In current mode, however, even with an output inductor physically in series with the output shunt capacitor, in a small-signal analysis which calculates gain and phase shift to consider the possibility of oscillation, the circuit behaves as if the inductor were not there.

The circuit behaves as if there were a constant current feeding the parallel combination of the output capacitor and the output load resistor. Such a network can yield only a 90° rather than a 180° phase shift, and the gain between output and input falls half as rapidly as for a true LC filter (−20 dB per decade rather than −40 dB per decade).

This simplifies the feedback-loop design, simplifies the circuitry around the error amplifier required for stabilization, and avoids the problems arising from rapid line or load changes. The details of why this is so will be discussed below.

5.2.4 Paralleling outputs

A number of current-mode power supplies may be operated in parallel with each sharing the total load current equally. This is achieved by sensing current in each supply with equal current sensing resistors which convert peak power transistor current pulses to voltage pulses. These are compared in a voltage comparator to a common error-amplifier output voltage which forces peak current sensing voltages and hence peak currents in the paralleled supplies to be equal.

5.2.5 Improved load current regulation

Current mode has better load current regulation than does voltage mode. But the improvement is not as great as voltage regulation, which is greatly enhanced by the feedforward characteristic inherent in current mode. The improved load current regulation comes about because of the greater error-amplifier bandwidth possible in current mode.

5.3 Current-Mode versus Voltage-Mode Control Circuits

To understand the differences and advantages of current mode, it is essential to first see how voltage-mode control circuitry works. The basic elements of a typical voltage mode, PWM control circuit is shown in Fig. 5.1. That block diagram shows most of the elements of the SG1524, the very first of the integrated-circuit control chips which have revolutionized the switching power supply industry. The SG1524, originally made by Silicon General Corporation, is now manufactured by many other companies and in improved versions such as the UC1524A (Unitrode) and SG1524B (Silicon General).

5.3.1 Voltage-mode control circuitry

In Fig. 5.1, a sawtooth oscillator generates a 3-V sawtooth triangle V_{st}. The DC voltage level at the triangle base is about 0.5 V and 3.5 V at its peak. The period of the triangle is set by external discrete components R_t and C_t and is approximately equal to $T = R_tC_t$.

An error amplifier compares a fraction of the output voltage KV_o to a voltage reference V_{ref} and produces an error voltage V_{ea}; V_{ea} is compared to the sawtooth V_{st} in a voltage comparator (PWM). Note in the error

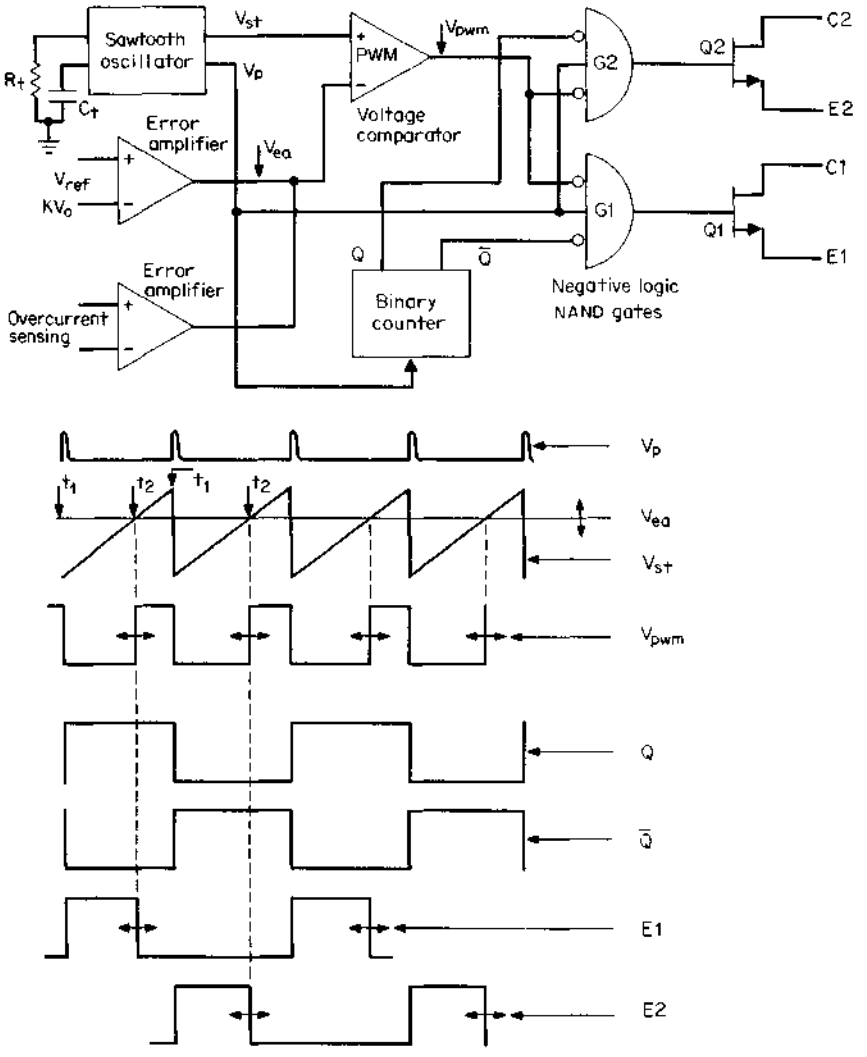


Figure 5.1 A basic voltage-mode PWM controller. Output voltage only is directly sensed by the error amplifier. Regulation against load current changes occurs only after the current changes cause small output voltage changes. The current-limit amplifier operates only to shut down the supply when a maximum current limit is exceeded. Transistor on time is from start of sawtooth until the sawtooth crosses V_{ea}.

amplifier that the fraction of the output KV_o is fed to the inverting input so that when V_o goes up, the error-amplifier output V_{ea} goes down.

In the PWM voltage comparator, the sawtooth is fed to the noninverting input and V_{ea} is fed to the inverting input. Thus the PWM output is a negative-going pulse of variable width. The pulse is negative for the entire time the sawtooth is below the DC level of the

error-amplifier output V_{ea} or from t_1 to t_2. As the DC output voltage goes—say—slightly positive, KV_o goes slightly positive, and V_{ea} goes negative and closer to the bottom of the sawtooth. Thus the duration of the negative-going pulse V_{pwm} decreases.

The duration of this negative-going pulse will be the duration of the power transistor on time. Further, since in all the topologies discussed above, the DC output voltage is proportional to the power transistor on time, decreasing the on time brings the DC output voltage back down in the negative-feedback loop. The duration of the negative pulse V_{pwm} increases, of course, as the output DC voltage decreases.

The 1524 is designed for push-pull-type topologies, so the single negative-going pulse of adjustable width, coming once per sawtooth period, must be converted to two 180° out-of-phase pulses of the same width. This is done with the binary counter and negative logic NAND gates $G1$ and $G2$. A positive-going pulse V_p occurring at the end of each sawtooth is taken from the sawtooth oscillator and used to trigger the binary counter.

Outputs from the binary counter Q and $\overline{Q}$ are then out-of-phase square waves at half the sawtooth frequency. These square waves, when they are negative, are used to steer negative V_{pwm} pulses alternatively through negative logic NAND gates $G1$ and $G2$. Those gates produce a positive output only for the duration of time that all inputs are negative. Thus the bases (and emitters) of output transistors $Q1$ and $Q2$ are positive only on alternative half cycles and only for the same duration as the V_{pwm} negative pulses.

The on time of the power transistors must correspond to the time the V_{pwm} pulse is negative for the complete circuit to have negative feedback as KV_o is connected to the inverting terminal of the error amplifier. Thus if the power transistors are of the NPN type, they must be fed from the emitters of $Q1$, $Q2$; or if of the PNP type, from the collectors. Or if current amplifiers are interposed between the bases of the power transistors and $Q1$, $Q2$, polarities must be such that the power transistors are on when $Q1$, $Q2$ are on.

The narrow positive pulse V_p is fed directly into gates $G1$, $G2$. This forces both gate outputs to be simultaneously low for the duration of V_p and both output transistors to be off for that duration. This ensures that if the pulse width of V_{pwm} ever approached a full half period, it would be guaranteed that both power transistors could never be simultaneously on at the end of the half period. In a push-pull topology, if both transistors are simultaneously on even for a short time, they are subjected to high current and the full supply voltage and could be destroyed.

This, then, is a voltage-mode circuit. Power transistor or output current is never sensed directly. The power transistors are turned on at

the beginning of a half period and turned off when the sawtooth V_{st} crosses the DC level of the error-amplifier output, which is a measure of output voltage only.

The complete details of the SG1524 are shown in Fig. 5.2*a*. The negative logic NAND gates $G1$, $G2$ of Fig. 5.1 are shown in Fig. 5.2*a* as positive logic NOR gates. These perform the same function for requiring all "lows" to make a "high" is identical to any one "high" forcing a "low."

In Fig. 5.2*a*, when pin 10 goes high, the associated transistor collector goes low and brings the error-amplifier output (pin 9) down to the base of the sawtooth. This reduces output transistor on times to zero and shuts down the supply. In the current limit comparator, if pin 4 is 200 mV more positive than pin 5, the error-amplifier output is also brought down to ground (there is an internal phase inversion, not shown) and the supply is shut down. Pins 4 and 5 are bridged across a current-sensing resistor in series with the current being monitored. If current is to be limited to I_m, the resistor is selected as $R_s = 0.2/I_m$.

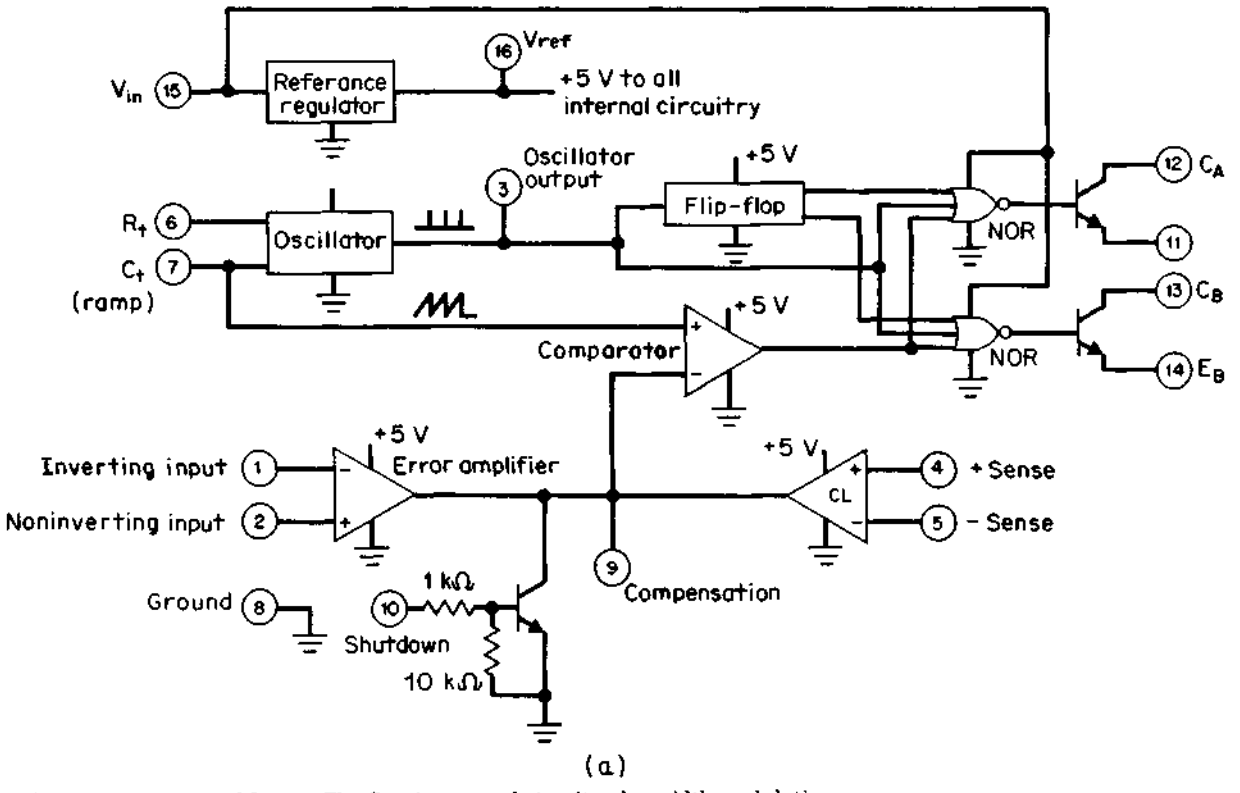
5.3.2 Current-mode control circuitry

Circuitry of the first integrated-circuit current-mode control chip (Unitrode UC1846) is shown in Fig. 5.2*b*. Figure 5.3 shows its basic elements controlling a push-pull converter.

It is seen in Fig. 5.3 that there are two feedback loops—an outer loop consisting of output voltage sensor (EA) and an inner loop comprising primary peak current sensor (PWM) and current-sensing resistor R_i which converts ramp-on-a-step transistor currents to ramp-on-a-step voltages.

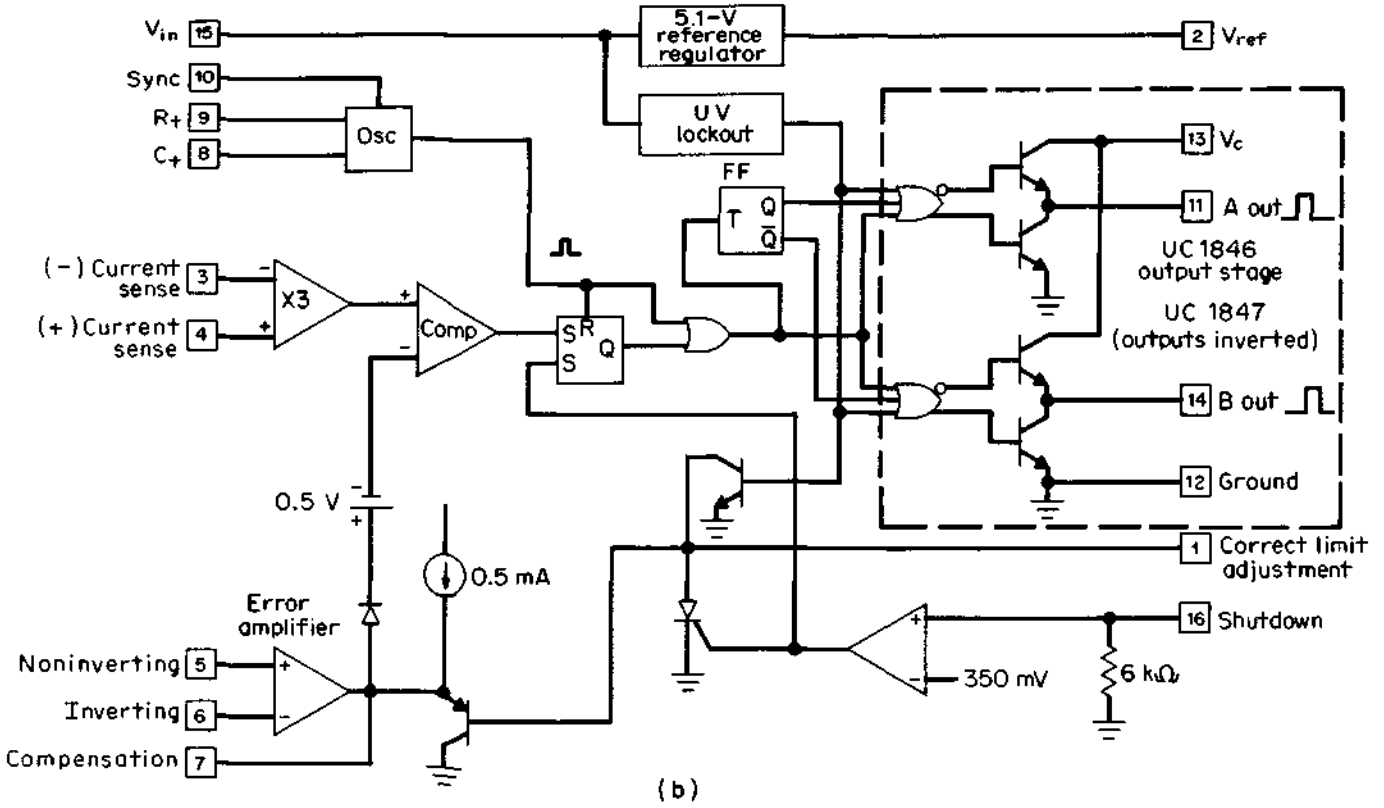
Regulation against line and load current changes is done by variation of the power transistor on time. On time is determined by both the voltage-sensing error-amplifier output V_{eao} and the PWM voltage comparator which compares V_{eao} to the ramp-on-a-step voltage at the top of the current-sensing resistor R_i.

Currents in the output transistors have the characteristic shape of a ramp on a step as the secondaries all have output inductors. This causes ramp on a step secondary currents which reflect as identical-shaped currents in the primary, lower in the ratio of N_s/N_p. Those currents flowing through R_i in the common emitters produce the ramp-on-a-step voltage waveshape shown as V_i. Power transistor on time is then determined as follows. An internal oscillator, whose period is set by external discrete components R_t, C_t, generates narrow clock pulses C_p. Oscillator period is approximately $0.9R_tC_t$. At every clock pulse, feedforward $FF1$ is reset, causing its output Q_{pw} to go "low." The duration of the low time at Q_{pw}, it will soon be seen, is the duration of



(a)

Figure 5.2 (*a*) PWM chip SG1524. The first integrated-circuit pulse-width-modulating control chip. (*Courtesy Silicon General Corp.*)



(b)

Figure 5.2 (*b*) PWM UC1846. The first integrated-circuit current-mode control chip. (*Courtesy Unitrode Corp.*)

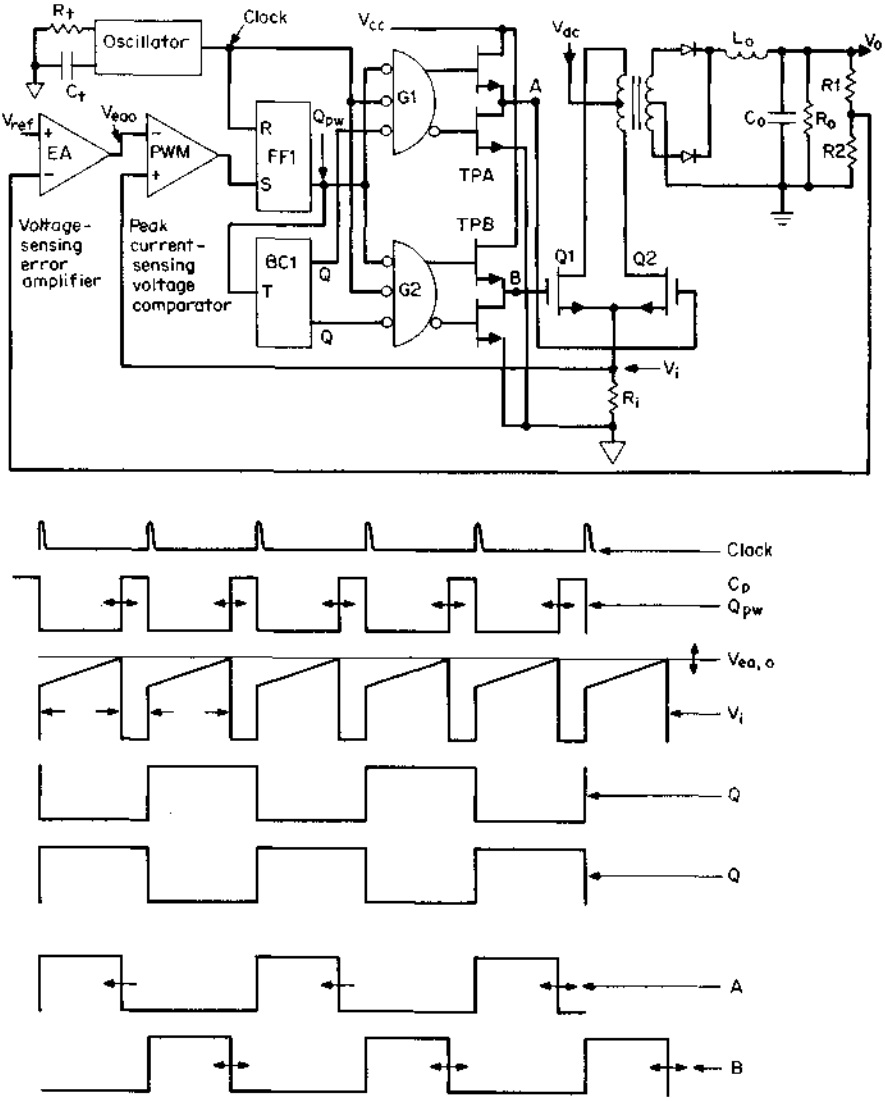


Figure 5.3 Current-mode controller (UC1846). Driving a push-pull MOSFET converter. Transistors are turned on alternatively at every clock pulse. They are turned off when the peak voltage across the common current-sensing resistor equals the output voltage of the voltage-sensing error amplifier. PWM forces all $Q1$, $Q2$ current pulses to have equal amplitude.

the "high" time at either of the chip outputs A or B and hence the duration of the power transistor on times.

Now when the PWM voltage comparator output goes high, $FF1$ is set, thus terminating the Q_{pw} low and hence the high time at A or B and turns off the power transistor which had been on. Thus the instant at which the PWM comparator output goes high determines the end of the on time.

Now the PWM comparator compares the ramp-on-a-step current-sensing voltage V_i to the output of the voltage error-amplifier EA. Hence when the peak of V_i equals V_{eao}, the PWM output goes positive, sets $FF1$, Q_{pw} goes high, and whichever of A or B had been high, goes back down low. The power transistor which had just been on is now turned off.

A low output from $FF1$ occurs once per clock period. It starts low at every clock pulse and goes back high when the PWM noninverting input equals the DC level of the EA output. Most frequently, power transistors $Q1$, $Q2$ will be N types which require positive-going signals for turnon. Thus these equal-duration negative-going pulses are steered alternately through negative logic NAND gates $G1$ and $G2$, becoming 180° out-of-phase, positive-going pulses at the chip output points A and B.

Chip output stages TPA and TPB are "totem poles." When the bottom transistor of a totem pole is on, the top one is off and vice versa. Output nodes A and B have very low output impedance. When the bottom transistor is on, it can "sink" (absorb inward-directed current) 100 mA statically and 400 mA during the high-to-low transition. When the top transistor is on, it can "source" (emit outward-directed current) 100 mA statically and 400 mA during the low-to-high transition.

The steering is done by binary counter $BC1$, which is triggered once per clock pulse by the leading edge of the pulse. The negative-going Q pulses steer the negative Q_{pw} pulses alternately through negative logic NAND gates $G1$, $G2$. The chip outputs at points A and B are 180° out-of-phase positive pulses whose duration is the same as that of the negative pulses shown as Q_{pw}.

Note that Q_{pw} is positive from the end of the on time until the start of the next turnon. This forces the bubble outputs of $G1$, $G2$ high and brings points A and B both low. This low at both power transistor inputs during the dead time between the turnoff of one transistor and the turnon of the other is a valuable feature. It presents a low impedance at the off-voltage level and prevents noise pickup from falsely turning the power transistors on. While the bubble outputs of $G1$, $G2$ are both high, their no-bubble outputs are both low. This turns off the upper transistors of the totem poles TPA and TPB and avoids overdissipating them.

It can also be seen that the narrow positive clock pulse is fed as a third input to NAND gates $G1$, $G2$. This makes bubble outputs from $G1$, $G2$ high and points A, B simultaneously low for the duration of the clock pulse. This guarantees that under fault conditions, if the controller attempted a full half period on time (Q_{pw} low and either A or B high for a full half period), there would be a dead time between the end of one on time and the start of the opposite on time. There could thus be no simultaneous conduction in the power transistors.

5.4 Detailed Explanation of Current-Mode Advantages

5.4.1 Line voltage regulation

Now consider how the controller regulates against line voltage changes. Assume that line voltage (and hence V_{dc}) goes up. Since secondary DC voltages are proportional to secondary winding peak voltages and power transistor on time, the on time must decrease as the peak secondary voltage has increased. Now as V_{dc} goes up, the peak secondary voltage will go up and after a delay in L_o, V_o will eventually go up. Then after a delay in getting through the error amplifier, V_{eao} will go down and in the PWM comparator, the ramp in V_i will equal the lowered value of V_{eao} earlier in time. Thus, on time will be decreased and the output voltage will be brought back down.

However, if this were the mechanism to correct against line voltage changes, that correction would be slow as it is slowed up by the delay in L_o and in getting through the error amplifier. But there is a shortcut around those delays. As V_{dc} goes up, the peak voltage at the input to the output inductor V_{sp} increases, the slope of inductor current dI_s/dt increases and hence the slope of the ramp of V_i increases. Now the faster ramp equals V_{eao} earlier in time and the on time is shortened without having to wait for V_{eao} to move down to shorten the on time. Output voltage transients resulting from input voltage transients will then be smaller in amplitude and shorter in duration because of this feedforward characteristic.

5.4.2 Elimination of flux imbalance

Consider the waveform V_i in Fig. 5.3. It is taken off the top of the current-sensing resistor and is hence proportional to power transistor currents. The on time ends when the peak of the ramp in V_i equals the output voltage of the error amplifier V_{eao}. It can be seen in Fig. 5.3 that peak currents on alternate half cycles cannot be unequal as in Fig. 2.4*b* and 2.4*c* because the error-amplifier output V_{eao} is pretty

much a straight line and cannot change within one cycle because of its limited bandwidth.

If the transformer core got slightly off center and started walking up into saturation on one side, the voltage V_i would go slightly concave upward close to the end of that on time. It would then equal V_{eao} earlier in time and terminate that on time sooner. Flux increase in that half cycle would then cease, and in the next half cycle, since the opposite transistor does not have a foreshortened on time, the core would be brought back down in flux and away from saturation.

Since the peak of the voltage ramps in Fig. 5.3 (V_i) are equal, peak currents on alternate half cycles must be equal. Thus flux imbalance and the inequality of alternate currents shown in Fig. 2.4*b* is not possible.

5.4.3 Simplified loop stabilization resulting from elimination of output inductor in small-signal analysis

Refer to Fig. 5.3. In a small-signal analysis to determine whether the loop is stable, it is assumed that the loop is opened at some point and a sinusoidal signal of variable frequency is inserted at that point. The gain and phase shift versus frequency is calculated through all the elements starting from the point at which the loop was assumed opened around to the point at the other end of the loop break. By tailoring the error-amplifier gain and phase shift properly in relation to the other elements in the open loop (primarily the output *LC* filter), the closed loop can be made stable.

The variable frequency is generally assumed inserted in series at the input end of the error amplifier. In Chap. 12 on feedback loop stability analysis, it will be shown how gain and phase shift through the error amplifier may be calculated and tailored to achieve desired results.

In Fig. 5.3, the concept of gain and phase shift of a sinusoidal signal from the error-amplifier output to the input of the *LC* filter may be difficult to grasp. At the error-amplifier output V_{eao} there is a DC voltage level which, when it equals the peak of the ramp-on-a-step pulse sequence V_i, results in a sequence of negative-going pulses at Q_{pw} whose duration is dependent on the DC voltage level of V_{eao}. The Q_{pw} negative pulses result in a sequence of positive-going pulses at the input to the *LC* filter.

It may seem puzzling to speak of gain and phase shift of sinusoidal signals in view of this odd operation of converting a voltage level to a sequence of pulses at the switching frequency. The situation may be clarified as follows.

If there is a sinusoidal signal at the error-amplifier input, the same

signal is amplified and phase-shifted at its output. Then the DC voltage level at V_{eao} is sinusoidally modulated at that frequency. Hence the pulse width of the Q_{pw} negative pulses is similarity modulated at that frequency. So is the pulse width or on time of the positive-going pulses at the output rectifiers modulated at that frequency. Hence the voltage at the output rectifier cathodes, which is proportional to the pulse widths, when averaged over a time long compared to the switching period, is simply the same modulating frequency as was inserted at the error-amplifier input.

The modulation operation (so long as the modulation period is long compared to the switching period) is then simply a sinusoid-to-pulse width-to-sinusoid converter. The gain of this modulation operation will be discussed further in the chapter on feedback loop stability.

Now in Fig. 5.3, there remains only the problem of calculating the gain and phase shift versus frequency for the sinusoid through the *LC* filter. If there were a sine wave of voltage at the rectifier cathodes, the phase shift through the *LC* filter would be 90° at its resonant frequency $\frac{1}{2}\pi\sqrt{LC}$ and 180° shortly after that, and gain from input to output would fall at a fast rate—40 dB/decade shortly after resonance.

But in current mode, the PWM comparator forces the output at the rectifier cathodes to be a sequence of width-modulated constant-current pulses—not voltage pulses. Thus at the input to the *LC* filter, the averaged waveform is a constant-current, not a constant-voltage, sinusoid.

With constant-current sinusoids, the filter inductor cannot act to change phase. The circuit behaves, in this small-signal analysis, as if the inductor were not present. Thus after the rectifier cathodes, the gain and phase shift correspond to that of constant-current sinusoids flowing into the paralleled combination of the output capacitor and load resistor. Such a circuit can yield only a maximum phase shift of 90° and a gain-versus-frequency characteristic which falls at 20 dB rather than 40 dB per decade.

Chapter 12 on feedback stability analysis, will show that this greatly simplifies the error-amplifier design, yields greater bandwidth, and improves the response of the closed-loop circuit to step changes in load and line. For now, Fig. 5.4*a* and 5.4*b* shows a comparison of the error-amplifier feedback networks required to stabilize a voltage-mode circuit (Fig. 5.4*a*) and a current-mode circuit (Fig. 5.4*b*).

5.4.4 Mechanism of load current regulation

In Fig. 5.3, the V_i voltage waveforms are proportional to power transistor current waveforms which are related to secondary current waveforms by the transformer turns ratio.

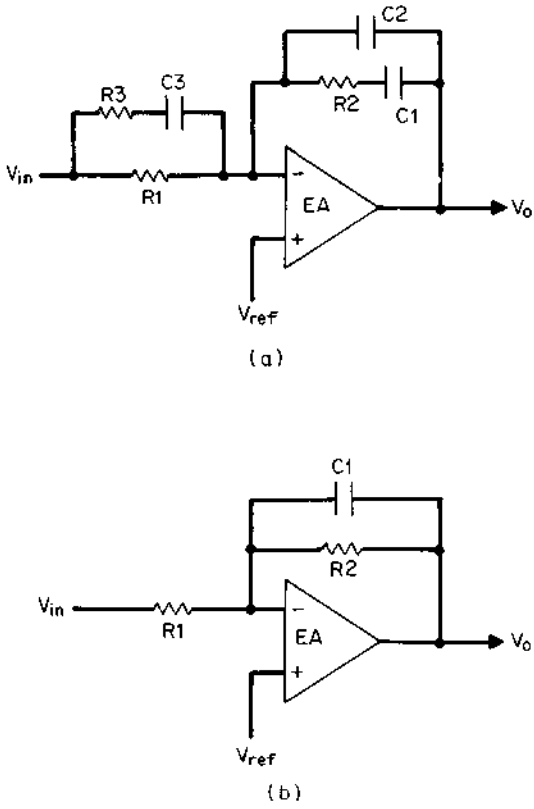


Figure 5.4 (*a*) Typical compensating network for a voltage-mode power supply. The complex input-feedback network in voltage mode is necessary because the output inductor and with the filter capacitor together yield a 180° phase shift and a 40 dB/decade gain versus frequency characteristic which make loop stabilization more difficult. (*b*) Typical compensation network for a current-mode power supply. In current mode, the source driving the output inductor is an effective "current source." The output inductor thus does not contribute to phase shift. The circuit acts at its output as if there were a constant current driving the parallel combination of the output filter capacitor and the output load resistance. Such a network yields a maximum 90° phase shift and a −20 dB/decade gain versus frequency characteristic. This permits the simpler input-feedback network for loop stabilization. It also copes much more easily with large-amplitude load and line changes.

Now at a DC input voltage V_{dc}, the peak secondary voltage is $V_{sp} = V_{dc}(N_s/N_p)$. For an on time per transistor of t_{on}, the DC output voltage is $V_o = V_{sp}(2t_{on}/T)$—just as for a voltage-mode push-pull circuit. The on time starts at the clock pulse as shown and ends when the V_i ramp equals the level of the voltage error-amplifier output as shown in Fig. 5.3.

If the DC voltage goes up, initially as described, the V_i ramp rate increases and shortens the on time as it reaches the original V_{eao} level earlier in time. This yields a fast correction for a step change in input voltage and the on time remains shorter as required by the preceding relation for the peak secondary voltage has gone up.

The mechanism of load current regulation, though, is different. For a fast step—say, an increase—in DC load current, the DC output voltage drops somewhat momentarily as the LC output filter has a certain surge impedance (approximately $\sqrt{L/C}$). After the delay in the error amplifier, V_{eao} moves up an amount determined by the EA gain.

Now the V_i ramp must go out further in time and hence in amplitude for it to reach equality with the higher V_{eao}. The secondary peak current and hence the output inductor current are thus larger in amplitude. The upslope of the inductor current lasts longer and eats somewhat into the dead time before the opposite transistor turnon.

With a shorter dead time, and starting from a higher current level at the beginning of the dead time, at the opposite transistor turnon, the current remaining in the inductor will be larger than it had been in the previous cycle. Thus the front-end step in each current pulse represented by V_i is greater than in the previous cycle.

This process continues for a number of switching cycles until the step part of the ramp-on-a-step current waveform builds up sufficiently to supply the increased demand for DC load current. As this current builds up, the DC output voltage gradually builds backup and V_{eao} relaxes back down, returning the on time to its original value. The time to respond to a change in DC load current is thus seen to be dependent on the size of the output inductor (a smaller value permits more rapid current changes in it) and the bandwidth of the error amplifier.

5.5 Current-Mode Deficiencies and Problems

5.5.1 Constant peak versus constant average output inductor problems[1–4]

Current mode forces peak power transistor currents to be constant at a level needed to supply the DC load current at the DC output voltage dictated by the voltage error amplifier as shown in Fig. 5.3.

The DC load current is the average of the output inductor current,

however, and keeping the peak transistor current constant keeps the peak output inductor current constant but does not keep the average inductor current constant. Because of this, in the unmodified current-mode scheme described thus far, changes in the DC input voltage will cause momentary changes in the DC output voltage. The output voltage change will be corrected by the voltage error amplifier outer feedback loop, as this is the loop which ultimately sets output voltage.

Again, however, the inner loop, in keeping peak inductor current constant, does not supply the correct average inductor current and output voltage changes again. The effect is then an oscillation which commences at every change in input voltage and which may continue for some time. The mechanism can be better understood from an examination of the upslope and downslope of the output inductor currents in Fig. 5.5.[3]

Figure 5.5*a* shows the upslope and downslope of the output inductor current for two different DC input voltages in current mode. Slope m_2 is the down slope = $dI_1/dt = V_o/L_o$. It is seen to be constant for the two different DC input voltages. At the high input voltage, on time is short at $t_{on,h}$ and at the lower DC input, on time is longer at $t_{on,l}$.

The peak inductor currents are constant as (see Fig. 5.3) the peak power transistor currents are kept constant by the PWM comparator. The DC voltage input V_{eao} to that comparator is constant as the outer feedback loop is keeping V_o constant. The constant V_{eao} then keeps V_i peaks constant, and hence transistor and output inductor peak currents are constant.

In Fig. 5.5*a*, in the steady state, the currents are drawn so that the current change during an on time are equal and opposite to those during an off time. This must be so because if not, there would be a DC voltage across the inductor; and since it is assumed that the inductor has negligible resistance, it cannot support DC voltage.

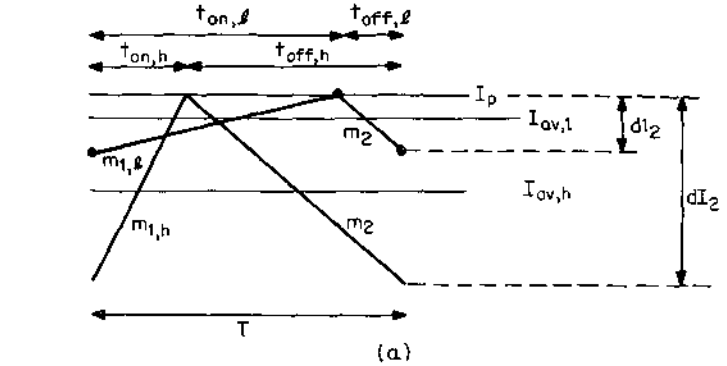
It can be seen in Fig. 5.5*a*, that the average inductor current at low DC input is higher than it is at high DC input voltage. This can be seen quantitatively as

$$I_{av} = I_p - \frac{dI_2}{2} = I_p - \left(\frac{m_2 t_{off}}{2}\right)$$

$$= I_p - \left[\frac{m_2(T - t_{on})}{2}\right]$$

$$= I_p - \left(\frac{m_2 T}{2}\right) + \left(\frac{m_2 t_{on}}{2}\right) \tag{5.1}$$

Since the voltage feedback loop keeps the product of $V_{dc}t_{on}$ constant, at lower DC input voltage, where the on time is higher, the average



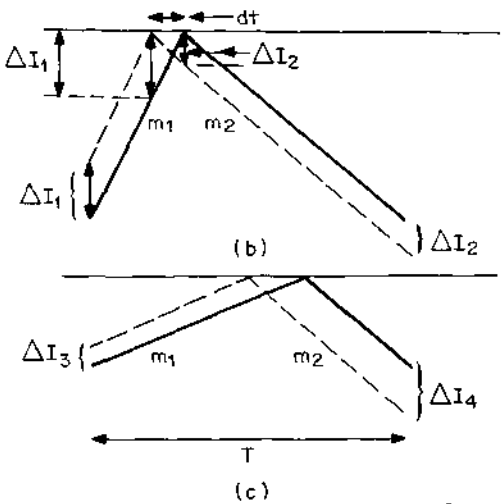


Figure 5.5 Problems in current mode. (*a*) Output inductor currents at high and low input voltages. In current mode, peak inductor currents are constant. At low DC input, t_{on} is maximum, yielding average inductor current I_{avl}. At high DC input, on time decreases to keep output voltage constant. But average inductor current I_{avh} is lower at high DC input. Since output voltage is proportional to average—not peak—inductor current, this causes oscillation when input voltage is changed. Slopes m_2 are inductor current downslopes. Slope m_{1l} is inductor current upslope at low line; m_{1h} is inductor current upslope at high line. (*b*) For a duty cycle less than 50 percent, an initial inductor current disturbance I_1 results in smaller I_2 disturbances in successive cycles until the disturbances die out. (*c*) For a duty cycle greater than 50 percent, an initial inductor current disturbance I_3 results in larger I_4 disturbances in successive cycles. The disturbances grow, and then decay resulting in an oscillation.

output inductor current I_{av} is higher, as can be seen from Eq. 5.1 and Fig. 5.5*a*.

Further, since the DC output voltage is proportional to the average, not the peak, inductor current, as DC input goes down, DC output voltage will go up. DC output voltage will then be corrected by the outer feedback loop and a seesaw action or oscillation will occur.

This phenomenon does not occur in voltage-mode control, where only DC output voltage is controlled. Also, since DC output voltage is proportional to average, not peak, inductor current, keeping output voltage constant maintains average inductor current constant.

5.5.2 Response to an output inductor current disturbance

A second problem which gives rise to oscillation in current mode is shown in Fig. 5.5*b* and 5.5*c*. In Fig. 5.5*b*, it is seen that at a fixed DC input voltage, if for some reason there is an initial current disturbance ΔI_1, after a first downslope the current will be displaced by an amount ΔI_2.

Further, if the duty cycle is less than 50 percent ($m_2 < m_1$) as in Fig. 5.5*b*, the output disturbance ΔI_2 will be less than the input disturbance ΔI_1, and after some cycles, the disturbance will die out. But if the duty cycle is greater than 50 percent ($m_2 > m_1$) as in Fig. 5.5*c*, the output disturbance after one cycle ΔI_4 is greater than the input disturbance ΔI_3. This can be seen quantitatively from Fig. 5.5*b* as follows. For a small current displacement ΔI_1, the current reaches the original peak value earlier in time by an amount dt where $dt = \Delta I_1/m_1$.

On the inductor downslope, at the end of the on time, the current is lower than its original value by an amount ΔI_2 where

$$\Delta I_2 = m_2 dt = \Delta I_1 \frac{m_2}{m_1} \tag{5.2}$$

Now with m_2 greater than m_1, the disturbances will continue to grow but eventually decay, giving rise to an oscillation.

5.5.3 Slope compensation to correct problems in current mode[1–4]

Both current-mode problems mentioned above can be corrected as shown in Fig. 5.6, where the original, unmodified output of the error amplifier is shown as the horizontal voltage level *OP*. The scheme for correcting the preceding problems ("slope compensation") consists of adding a negative voltage slope of magnitude m to the output of the error amplifier. By proper selection of the m in a manner discussed

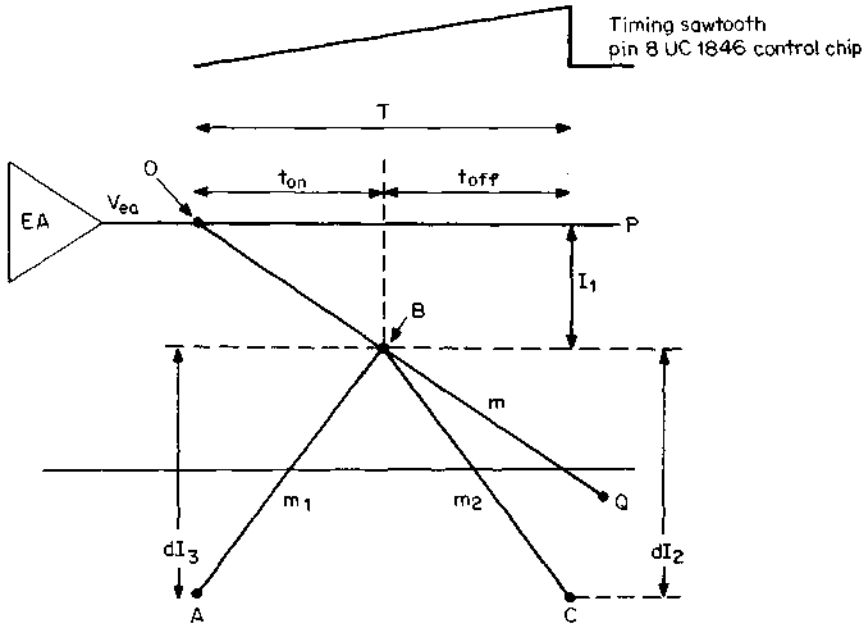


Figure 5.6 Slope compensation. By adding a negative voltage slope of magnitude $m = N_s/N_p(R_i)(m_2/2)$ to the error-amplifier output (Fig. 5.3), the two problems shown in Fig. 5.5 are corrected.

below, the output inductor average DC current can be made independent of the power transistor on time. This corrects the problems indicated by both Eqs. 5.1 and 5.2.

In Fig. 5.6, the upslope m_1 and downslope m_2 of output inductor current is shown. Recall that in current mode, the power transistor on time starts at every clock pulse and ends at the instant the output of the PWM comparator reaches equality with the output of the voltage error-amplifier as shown in Fig. 5.3. In slope compensation, a negative voltage slope of magnitude $m = dV_{ea}/dt$ starting at clock time is added to the error-amplifier output. The magnitude of m is calculated thus. In Fig. 5.6, the error-amplifier output at any time t_{on} after a clock pulse is

$$V_{ea} = V_{eao} - mt_{on} \tag{5.3}$$

where V_{eao} is the error-amplifier output at t_{on} equal to zero. The peak voltage V_i across the primary current-sensing resistor R_i in Fig. 5.3 is

$$V_i = I_{pp}R_i = I_{sp}\frac{N_s}{N_p}R_i$$

where I_{pp} and I_{sp} are the primary and secondary peak currents, respectively. But $I_{sp} = I_{sa} + dI_2/2$, where I_{sa} is the average secondary or average output inductor current and dI_2 in Fig. 5.6 is the inductor current change during the off time ($= m_2 t_{off}$). Then

$$I_{sp} = I_{sa} + \frac{m_2 t_{off}}{2}$$

$$= I_{sa} + \frac{m_2}{2}(T - t_{on})$$

Then
$$V_i = \frac{N_s}{N_p} R_i \left[I_{sa} + \frac{m_2}{2}(T - t_{on}) \right] \tag{5.4}$$

Equating Eqs. 5.3 and 5.4, which is what the PWM comparator does, we obtain

$$\frac{N_s}{N_p} R_i I_{sa} = V_{eao} + t_{on}\left(\frac{N_s}{N_p} R_i \frac{m_2}{2} - m\right) - \left(\frac{N_s}{N_p} R_i \frac{m_2}{2} T\right)$$

It can be seen in this relation that if

$$\frac{N_s}{N_p} R_i \frac{m_2}{2} = m = \frac{dV_{ea}}{dt} \tag{5.5}$$

then the coefficient of the t_{on} term is zero and the average output inductor current is independent of the on time. This then corrects the above two problems arising from the fact that without compensation, current mode maintains constant the peak, and not the average, output inductor current.

5.5.4 Slope compensation with a positive-going ramp voltage[3]

In the previous section it was shown that if a negative ramp of magnitude given by Eq. 5.5 is added to the error-amplifier output, the two current-mode problems described above are corrected.

The same effect is obtained by adding a positive-going ramp to the output of the current-sensing resistor V_i (Fig. 5.3) and leaving the error-amplifier output voltage V_{eao} (Fig. 5.3) unmodified. Adding the positive ramp to V_i is simpler and is the more usual approach. That adding the appropriate positive ramp to V_i also makes the average output inductor current independent of the on time can be shown as follows. A ramp voltage of slope dV/dt will be added to the voltage V_i of Fig. 5.3 and the resultant voltage will be compared in the PWM to the error-amplifier output V_{eao} of that figure. When the PWM finds

equality of those voltages, its output terminates the on time. Then $V_i + dV/dt = V_{eao}$. Substitute V_i from Eq. 5.4:

$$\frac{N_s}{N_p} R_i \left[I_{sa} + \frac{m_2}{2} (T - t_{on}) \right] + \frac{dV}{dt} t_{on} = V_{eao}$$

Then $$\frac{N_s}{N_p} R_i I_{sa} + \frac{N_s}{N_p} R_i \frac{m_2}{2} T + t_{on} \left(\frac{dV}{dt} - \frac{N_s}{N_p} R_i \frac{m_2}{2} \right) = V_{eao}$$

From the above, it is seen that if the slope dV/dt of the voltage added to V_i is equal to $(N_s/N_p)R_i m_2/2$, the terms involving t_{on} in the preceding relation vanish and the secondary average voltage I_{sa} is independent of the on time. Note that $m_2 (= V_o/L_o)$ is the current downslope of the output inductor as defined earlier.

5.5.5 Implementing slope compensation[3]

In the UC1846 chip, a positive-going ramp starting at every clock pulse is available at the top of the timing capacitor (pin 8 in Fig. 5.2*b*). The voltage at that pin is

$$V_{osc} = \frac{\Delta V}{\Delta t} t_{on} \tag{5.6}$$

where $\Delta V = 1.8$ V and $\Delta t = 0.45 R_t C_t$

As seen in Fig. 5.7, a fraction of that voltage whose slope is $\Delta V/\Delta t$ is added to V_i (the voltage across the current-sensing resistor). That slope is set to $(N_s/N_p)R_i(m_2/2)$ by resistors $R1$, $R2$. Thus in Fig. 5.7, since R_i is much less than $R1$, the voltage delivered to the current sensing terminal (pin 4) is

$$V_i + \frac{R1}{R1 + R2} V_{osc} = V_i + \frac{R1}{R1 + R2} \frac{\Delta V}{\Delta t} t_{on} \tag{5.7}$$

and setting the slope of that added voltage equal to $(N_s/N_p)R_i m_2/2$; we obtain

$$\frac{R1}{R1 + R2} = \frac{(N_s/N_p)(R_i)(m_2/2)}{\Delta V/\Delta t} \tag{5.8}$$

where $\Delta V/\Delta t = 1.8/(0.45 R_t C_t)$.

Since $R1 + R2$ drains current off the top of the timing capacitor, it changes operating frequency. Then either $R1 + R2$ is made large enough so that the frequency change is small or an emitter-follower is

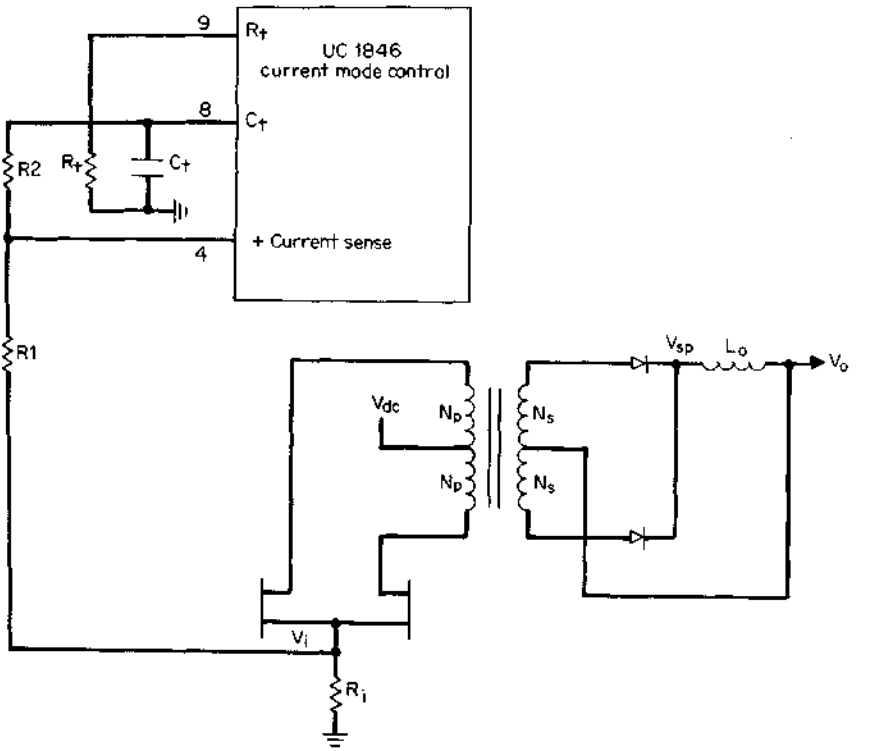


Figure 5.7 Slope compensation in the UC1846 current-mode control chip. A positive ramp voltage is taken off the top of the timing capacitor, scaled by resistors R_1, R_2 and added to the voltage at the top of the current resistor R_i. By choosing R_1, R_2 to make the slope of the voltage added to V_i equal to half the downslope of the output inductor current, reflected into the primary and multiplied by R_i, the output inductor average current is rendered independent of power transistor on times.

interposed between pin 8 and the resistors. Usually $R1$ is preselected and $R2$ is calculated from Eq. 5.8.

5.6 Voltage-Fed and Current-Fed Topologies

5.6.1 Introduction and definitions

All topologies discussed thus far have been of the voltage-fed type. *Voltage-fed* implies that the source impedance of whatever drives the topology is low and hence there is no way of limiting the current drawn from it during the unusual conditions at power switch turnon or turnoff or under various fault conditions in the topology.

Of course, there are various ways of implementing "current limit-

ing" with additional circuitry which senses an overcurrent condition and takes some kind of corrective action such as narrowing the controller's switching pulse width or stopping it completely. But all such schemes are not instantaneous; they involve a delay over a number of switching cycles during which there can be excessive dissipation in either the power transistors or output rectifiers and dangerous voltage or current spiking. Thus such overcurrent-sensing schemes are of no help in the case of high transient currents at the instant of the power switch turnon and turnoff.

The low source impedance in voltage-fed topologies is the output impedance of the filter capacitor in off-line converters or the low output impedance of the battery in battery-powered converters. Or in compound schemes where a buck regulator preregulates the rectified DC voltage of the AC line rectifier, it is the very low output impedance of the buck regulator itself.

In current-fed topologies, the high instantaneous impedance of an inductor is interposed between the power source and the topology itself. This provides a number of significant advantages, especially in high power supplies (>1000 W), high output voltage supplies (> ~200 V), and multioutput supplies where close tracking between slaves and a master output voltage is required.

Advantages of the current-fed technique can be appreciated by examining the shortcomings of the usual high-power, high-output-voltage, multioutput voltage-fed topologies.

5.6.2 Deficiencies of voltage-fed, width-modulated full-wave bridge[9]

Figure 5.8 shows a conventional voltage-fed full bridge—the usual choice for a switching supply at 1000-W output. At higher output powers and high output voltages and for many output voltages, it has the following significant shortcomings.

5.6.2.1 Output inductor problems in voltage-fed, width-modulated full-wave bridge. For high output voltages, the size and cost of the output inductor L_o (or inductors in a multioutput supply) become prohibitive, as can be seen from the following. The inductor is selected to prevent its going into the discontinuous mode or running dry at the specified minimum DC load current (Secs. 1.3.6, 2.2.14.1). For a minimum DC load current of one-tenth its nominal value I_{on}, Eq. 2.20 gives the magnitude of the inductor as $L_o = 0.5V_oT/I_{on}$.

Now consider a 2000-W supply at $V_o = 200$ V, $I_{o(nominal)} = 10$ A, and a minimum DC output current of 1 A. To minimize the size of the out-

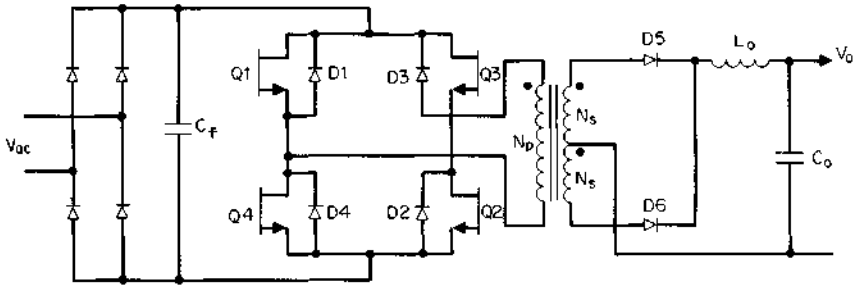


Figure 5.8 Conventional voltage-fed full bridge communally used for output powers at 1000 W. The low source of impedance seen looking back into filter capacitor C_f and the need for output inductor L_o are significant drawbacks for output powers over 1000 W and output voltages over a few hundred volts. Also, the requirement for the output inductor at each output of a multioutput power supply makes the topology expensive in dollars and required space.

put inductor, T should be minimized and a switching frequency of 50 kHz might be considered. At 50 kHz, for V_o = 200 V, I_{on} = 10 A, Eq. 2.20 yields L_o of 200 μH.

The inductor would have to carry the nominal current of 10 A without saturating. Such inductors capable of carrying large DC bias currents without saturating are discussed in a later chapter and are made either with gapped ferrite or powdered iron toroidal cores. Such a 200-μH 10-A inductor using a powdered iron toroid would be about 2.5 in in diameter and 1.0 in in height.

Although this is not too prohibitive a size for a single-output 2-kW supply, if there were many outputs, higher output voltages or higher output power, the size and cost of many such large inductors would be a serious drawback. For high output voltages (>1000 V), even at low output currents, the output inductor is far more troublesome because of the large number of turns required to support the high voltage across the inductor. The high voltage across the inductor—especially during the dead time when cathodes of $D5$, $D6$ of Fig. 5.8 are both low—can produce corona and arcing.

A further problem with a topology as shown in Fig. 5.8 requiring output inductors is the poor cross regulation or change in output voltage of a slave when current changes in the master (Sec. 2.2.2). The output inductors in both the master and slave must be large enough to prevent them from going into the discontinuous mode and causing large output voltage changes at minimum load currents.

The current-fed circuit (Fig. 5.10) to be discussed below, avoids all these problems as it does not require output inductors. It uses a single-input inductor $L1$ before the bridge and only capacitors after second-

ary rectification. Thus DC output voltages are the peak rather than the average of the transformer secondary voltages. Voltage regulation is achieved by width modulation of the transistor switch $Q5$ ahead of the input inductor.

5.6.2.2 Turnon transient problems in voltage-fed, width-modulated full-wave bridge.[9] In Fig. 5.8 diagonally opposite transistors are simultaneously on at alternate half cycles. The maximum on time of each pair is designed to never have to be more than 80 percent of a half period. This ensures a $0.2T/2$ dead time between the turnoff of one transistor pair and the turnon of the other. This dead time is essential, for if the on time of alternate pairs overlapped even if only for a fraction of a microsecond, there would be a dead short circuit across the filter capacitor, and with nothing to limit current flow, the transistors would fail immediately.

Now during the dead time, all four transistors are off, the anodes of output rectifiers $D5$, $D6$ are at ground and the input end of filter inductor L_o has swung down in an attempt to keep its current constant. The input end of L_o is caught at one diode drop below ground by the cathodes of $D5$, $D6$, which now act as free-wheeling diodes. The current which had been flowing in L_o before the dead time (roughly equal to the DC output current) now continues to flow in the same direction. It flows out through the ground terminal into the secondary center tap, where it divides equally, with each half flowing through $D5$, $D6$ back into the input end of L_o.

Now at the start of the next half cycle when, say, $Q1$, $Q2$ turn on, the no-dot end of the $T1$ primary is high and the no-dot end of the $T1$ secondary (anode of $D6$) attempts to go high. But the cathode of $D6$ is looking into the cathode of $D5$, which is still conducting half the DC output current. Until $D6$ supplies a current equal to and canceling the $D5$ forward current, it is looking into the low impedance of a conducting diode ($\sim 10\ \Omega$).

This secondary low impedance reflects as a low impedance across the transformer. But this low impedance is in series with the transformer's leakage inductance, which limits the primary current during the time required to cancel the $D5$ free-wheeling current. Because of the high-impedance current-limiting effect of the leakage inductance, transistors $Q1$ and $Q2$ remain in saturation until the $D5$ free-wheeling current is canceled.

When the $D5$ current is canceled, it still has a low impedance because of its reverse-recovery time, which may range from 35 ns (ultra-fast-recovery type) to 200 ns (fast-recovery type). For a reverse recovery time of t_r, supply voltage of V_{cc}, and transformer primary leakage inductance of L_l, the primary current overshoots to $V_{cc}t_r/L_l$. This over-

shoot current can pull the transistors out of saturation and either partially damage or destroy them.

Finally, when the output rectifier recovers abruptly, there is a damped oscillatory ring at its cathode. The first positive half cycle of this ring can more than double the reverse voltage stress on the diode and possibly destroy it. Even in lower power supplies, it is often necessary to put a series RC arm across the rectifiers to damp the oscillation. The penalty paid for this is, of course, dissipation in the resistor.

5.6.2.3 Turnoff transient problems in voltage-fed, width-modulated full-wave bridge.[9] In Fig. 5.8, there is a spike of high power dissipation at turnoff as a result of the instantaneous overlap of falling current and rising voltage across the off-turning transistors.

Thus, consider that $Q3$ and $Q4$ are on and have received turnoff signals at their bases. As $Q3$, $Q4$ commence turning off, current stored in the leakage and magnetizing inductance of $T1$ force a polarity reversal across its primary. The bottom end of $T1$ goes immediately positive and is clamped via $D1$ to the positive rail at the top of $C1$. The top end of $T1$ goes immediately negative and is clamped via $D2$ to the negative rail at the bottom end of $C1$. Now voltages across $Q3$ and $Q4$ are clamped at V_{cc} so long as diodes $D1$, $D2$ conduct. There are no leakage inductance voltage spikes across $Q3$, $Q4$ as in push-pull or single-ended forward converter topologies. Energy stored in the leakage inductance is returned without dissipation to the input capacitor C_f.

However, while voltage across $Q3$, $Q4$ is held at V_{cc}, the current in these two transistors falls linearly to zero in a time t_f determined by their reverse base drives. This overlap of a fixed voltage V_{cc} and a current falling linearly from a value I_p results in a dissipation averaged over a full period T of

$$PD = V_{cc} \frac{I_p}{2} \frac{t_f}{T} \tag{5.9}$$

It is instructive to calculate this dissipation for, say, a 2-kW supply operating at 50 kHz from a nominal V_{cc} of 336 V (typical V_{cc} for an off-line inverter operating from a 120-V-AC line in the voltage-doubling mode as in Sec. 3.1.1). Assume a minimum V_{cc} of 0.9 (336) or 302 V. Then from Eq. 3.7, the peak current is

$$I_p = \frac{1.56 P_o}{V_{dc}} = 1.56 \frac{2000}{302} = 10.3 \text{ A}$$

A bipolar transistor at this power level could be expected to have a fall time of 0.3 μs. Then since peak currents are independent of DC input voltage, calculate overlap dissipation from Eq. 5.9 at a high line

of 1.1 × 336 = 370 V. For the dissipation in either $Q3$ or $Q4$, Eq. 5.9 gives $PD = V_{cc}(I_p/2)(t_f/T) = 370(10.3/2)(0.3/20) = 28.5$ W, and for the four transistors in the bridge, total overlap losses would be 114 W.

It is of interest to calculate the dissipation per transistor during the on time. This is $V_{ce(sat)}I_cT_{on}/T$ and for a typical $V_{ce(sat)}$ of 1.0 V and an "on" duty cycle of 0.4 is only 1 × 10.3 × 0.4 or 4.1 W.

Even though the 28.5 W of overlap dissipation per transistor can be reduced with four load- and line-shaping "snubbers" (to be discussed in a later chapter), these snubbers reduce transistor losses only by diverting them to the snubber resistors with no improvement in efficiency. It will be shown that in the current-fed topology, only two snubbers will be required, reducing transistor overlap dissipation to a negligible value. There will be a price paid for this as the dissipation in each of the two snubber resistors will be somewhat greater than that for the voltage-fed full bridge.

5.6.2.4 Flux-imbalance problem in voltage-fed, width-modulated full-wave bridge. Flux imbalance or operation not centered about the origin of the transformer's *BH* loop was discussed in Sec. 2.2.5 in connection with the push-pull and in Sec. 3.2.4 for the half bridge. It arises because of an unequal volt-second product applied to the transformer primary on alternate half cycles. As the core drifts further and further off center on the *BH* loop, it can move into saturation, be unable to sustain the supply voltage, and destroy the transistor.

Flux imbalance can also arise in the conventional full-wave bridge because of a volt-second imbalance on alternate half cycles. This can come about with bipolar transistors because of unequal storage times on alternate half cycles or with MOSFET transistors because of unequal MOSFET on-voltage drops. The solution for the full-wave bridge is to place a DC blocking capacitor in series with the primary. This prevents a DC current bias in the primary and forces operation to be centered about the *BH* loop origin. The size of such a DC blocking capacitor is calculated just as in Sec. 3.2.4 for the half bridge.

Current-fed circuits, to be discussed below, do not require such DC blocking capacitors and thus offer another advantage over voltage-fed circuits despite the relatively small size and cost of blocking capacitors.

5.6.3 Buck-voltage-fed full-wave bridge topology—basic operation

This topology is shown in Fig. 5.9. It avoids many of the deficiencies of the voltage-fed width-modulated full-wave bridge in high-voltage high-power multioutput supplies.

Consider first how it works. It is simply a buck regulator preceding

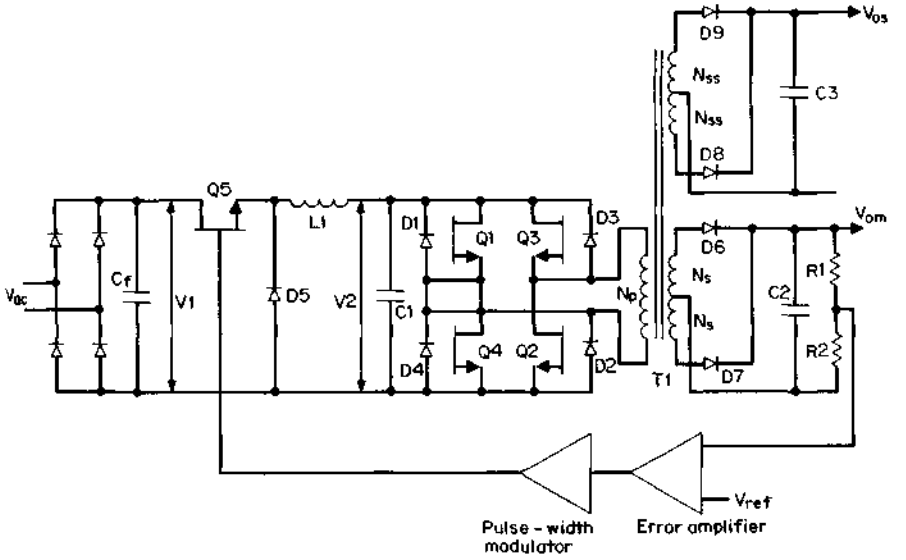


Figure 5.9 Buck-voltage-fed full bridge. The buck regulator preceding the full bridge eliminates the output inductors in a multioutput supply, but the low source impedance offered by the buck capacitor and the low output impedance of the buck regulator still leaves many drawbacks for this approach. $Q5$ is width-modulated, but $Q1$ to $Q4$ are operated at a fixed on time at about 90 percent of a half period to avoid simultaneous conduction. The output filters $C2$ are peak rather than averaging rectifiers. Practical output powers of about 2 kW to 5 kW are realizable.

a square-wave inverter which has only capacitors after the secondary rectifying diodes. Thus the DC output voltage at the filter capacitor is the peak of the secondary voltage (less, of course, the negligible rectifying diode drop). Also neglecting inverter transistor on drop, the DC output voltage is $V_o = V_2(N_s/N_p)$, where V_2 is the output of the buck regulator. Now the inverter transistors are not width-modulated. They are operated at a fixed on time—roughly 90 percent of a half period to avoid simultaneous conduction in the two transistors positioned vertically above one another. Diagonally opposite transistors are switched on and off simultaneously.

Feedback is taken from one of the secondary outputs (usually the highest current one or the one specified at tightest output voltage tolerance) and used to width-modulate the buck transistor $Q5$. This bucks down the rectified, unregulated DC voltage V_1 to a DC value V_2 arbitrarily selected to be about 25 percent lower than the lowest rectified voltage V_1 corresponding to the lowest specified AC input voltage. The turns ratio N_s/N_p is then chosen so that for the prespecified value of V_2, the correct master output voltage $V_{om} = V_2(N_s/N_p)$ is obtained.

The feedback loop, in keeping V_{om} constant against line and load changes, then keeps V_2 constant (neglecting relatively constant rectifier diode drops) at $V_2 = V_o(N_p/N_s)$. Now additional secondaries, rectifier diodes, and peak rectifying filter capacitors can be added for slave outputs.

Feedback can alternatively be taken from the top of C1 to keep V_2 constant. From V_2 to the outputs, the circuit is open-loop. But all output voltages are still quite constant for line and load changes as they are all proportional to V_2 and change only slightly as the forward drop in diode rectifiers and the on drop of the transistors changes only slightly with output currents.

Taking feedback from V2 thus results in somewhat less constant output voltages but avoids the problem of transmitting a width-modulated control voltage pulse across the boundary from output to input ground. Or if an error amplifier is located on output ground with a pulse-width modulator on input ground, it avoids the problem of transmitting the amplified DC error voltage across the output ground–input ground boundary. Such a scheme usually involves the use of an optocoupler, which has wide tolerances in gain and is not too reliable a device.

5.6.4 Buck-voltage-fed full-wave bridge advantages

5.6.4.1 Elimination of output inductors. The first obvious advantage of the topology is that for a multioutput supply, it replaces many output inductors with a single-input inductor with consequent savings in dollars and space.

Since there are no output inductors in either the master or slaves, there is no problem of the inductors going into the discontinuous mode or running dry with consequent large output voltage changes (Secs. 1.3.6, 2.2.4). Slave output voltages will then track the master over a large range of output currents to about ±2 percent rather than the ±6 percent to ±8 percent or substantially more if output inductors were present and were permitted to go into the discontinuous mode.

The input inductor will be designed to go discontinuous just at the point where the total power or current is at its minimum. Since it is unlikely that all outputs are at minimum current simultaneously, this indicates a higher total minimum current and a smaller input inductor (Sec. 1.3.6).

Further, even after the input inductor has gone discontinuous, the master output voltage will remain substantially constant but with somewhat more output ripple and somewhat poorer load regulation. The feedback loop keeps the main output voltage constant even after onset of the discontinuous mode point by large decreases in on time of the buck transistor (Fig. 1.6*a*).

Further, since the slave outputs are clamped to the main output in the ratio of their respective turns ratios, slaves will also remain constant over large line and load changes.

Elimination of output inductors with their many turns required to sustain high AC voltages for high DC outputs makes 2000- to 3000-V outputs easily feasible. Higher output voltages—15,000 to 30,000 V—at relatively low output currents as for cathode-ray tubes or high-voltage–high-current outputs as for traveling-wave tubes are easily obtained by conventional diode-capacitor voltage multipliers after the secondaries.[8]

5.6.4.2 Elimination of bridge transistor turnon transients. With respect to the full-wave width-modulated bridge of Fig. 5.8, Sec. 5.6.2.2 discussed turnon transient current stresses in the bridge transistors ($Q1$ to $Q4$) and excessive voltage stresses in the rectifying diodes ($D5$, $D6$).

It was pointed out in Sec. 5.6.2.2 that these stresses arose because the rectifier diodes were also acting as free-wheeling diodes. Thus at the instant of turnon of one diagonally opposite pair (say, $Q1$, $Q2$), $D6$ was still conducting as a free-wheeling diode. Then until the forward current in $D6$ was canceled, the impedance seen by $Q1$, $Q2$ was the leakage inductance of $T1$ in series with the low forward impedance of $D6$ reflected into the primary.

Subsequently, when $Q1$, $Q2$ forced a current into the primary sufficient to cancel the $D6$ forward current, there was still a low impedance reflected into the primary because of poor reverse recovery time in $D6$. This caused a large primary current overshoot which overstressed $Q1$, $Q2$. At the end of the recovery time, when the large secondary current overshoot terminated, it caused an oscillation and excessive voltage stress on $D6$.

But this current overstress on bridge transistors and voltage overstress on output rectifiers does not occur with the buck-voltage-fed topology of Fig. 5.9. The inverter transistors are operated with a dead time ($\sim 0.1T/2$) between the turnoff of one pair of transistors and the turnon of the other pair.

During this dead time when none of the bridge transistors are on, no current flows in the output rectifiers and output load current is supplied from the filter capacitors alone. Thus at the start of the next half period, the on-turning rectifier diode is not loaded down with a conducting free-wheeling diode as in Fig. 5.8. The opposite diode has long since ceased conducting; thus there is no current overstress in the bridge transistors and no recovery time problem in the rectifier diodes and hence no overvoltage stress in them.

5.6.4.3 Decrease of bridge transistor turnoff dissipation. In Sec. 5.6.2.3 it was calculated that for a 2000-W supply operating from a nominal

input of 120-V-AC in the input voltage-doubling mode, the bridge dissipation is 28.5 W at maximum AC input for each of the four transistors in the voltage-fed width-modulated bridge circuit of Fig. 5.8.

In the buck-voltage-fed full-wave bridge (Fig. 5.9), this dissipation is somewhat less. This is so because even at maximum AC input, the off-turning bridge transistors are subjected to bucked-down voltage V_2 (Fig. 5.9) of about 0.75 times the minimum rectified voltage as discussed in Sec. 5.6.3. For the minimum rectified DC of 302 V (Sec. 5.6.2.3), this is 0.75 × 302 or 227 V. This compares to 370 V DC at maximum AC input as calculated in Sec. 5.6.2.3.

The peak current from the bucked-down 227 V will not differ much from the calculated 10.3 A in Sec. 5.6.2.3. Thus assume a total efficiency of 80 percent as for the circuit of Fig. 5.8. Assume that half the losses are in the bridge and half in the buck regulator of Fig. 5.9.

Then for a bridge efficiency of 90 percent, its input power is 2000/0.9 or 2222 W. The bridge, with a preregulated input, can operate its transistors at a 90 percent duty cycle without concern about simultaneous conduction in two transistors vertically above one another. Input power is then $0.9I_pV_{dc} = 2222$. For V_{dc} of 227 V as above, this yields I_p of 10.8 A. Then, calculating bridge transistor dissipation as in Sec. 5.6.2.3 for a current fall time t_f of 0.3 μs out of a period T of 20 μs, dissipation per transistor is $(I_p/2)(V_{dc})(t_f/T) = (10.8/2)227 \times 0.3/20 =$ 18.4 W. This is 74 W for the entire bridge as compared to 114 W for the circuit of Fig. 5.8 as calculated in Sec. 5.6.2.3.

5.6.4.4 Flux imbalance problem in bridge transformer. This problem is still the same as in the topology of Fig. 5.8. A volt-second unbalance can occur because of unequal storage times for bipolar bridge transistors or because of unequal on voltages for MOSFET transistors. The solution for both the Fig. 5.8 and Fig. 5.9 topologies is to insert a DC blocking capacitor in series with the transformer primary.

5.6.5 Drawbacks in buck-voltage-fed full-wave bridge[9,10]

Despite the advantages over the width-modulated full-wave inverter bridge, the buck-voltage-fed full-wave bridge has a number of significant drawbacks.

First, of course, there is the added cost, volume, and power dissipation of the buck transistor $Q5$ (Fig. 5.9) and the cost and volume of the buck LC filter ($L1$,$C1$). The added cost and volume of these elements is partly canceled by the saving of an inductor at each output voltage. The added dissipation of the buck regulator $Q5$ plus the free-wheeling diode $D5$ is most often a small percentage of the total losses for a ≥2000-W power supply.

Second, there are turnon and turnoff transient losses in the buck transistor which can be greater than its DC conduction losses. These can be reduced in the transistor by diverting them to passive elements in snubbers. But the losses, cost, and required space of the snubbers is still a drawback. Such turnon-turnoff snubbers will be discussed in the later section on the buck-current-fed full-wave bridge.

Turnoff transient losses in the bridge transistors, although less (as discussed in Sec. 5.6.4.3) than for the width-modulated bridge of Fig. 5.8, are still significant.

Finally, under fault conditions (unusually long storage time at high temperature and low load or low line conditions), at the turnon of one transistor pair, the opposite pair may still be on. With the low source impedance of the buck filter capacitor and the momentary short circuit across the supply bus, this will cause immediate failure of at least one and possibly all of the bridge transistors.

5.6.6 Buck-current-fed full-wave bridge topology—basic operation[9,10]

This topology is shown in Fig. 5.10.[6] It has no output inductors and is seen to be exactly like the buck-voltage-fed full-wave bridge of Fig. 5.9 with the exception that the buck filter capacitor $C1$ of Fig. 5.9 is not physically present. Instead, there is present at that point a virtual capacitor $C1V$ which is the sum of all the secondary filter capacitors reflected by the square of their respective turns ratios into the $T1$ primary. The filtering by this virtual capacitor $C1V$ is exactly the same as if a real capacitor of equal magnitude were present at the $L1$ output.

Thus, by replacing all the output inductors of the width-modulated full-wave bridge of Fig. 5.8, with a single primary side inductor as in Fig. 5.9, all the advantages described in Sec. 5.6.2.1 for the Fig. 5.9 circuit are also obtained for the circuit of Fig. 5.10.

Now bridge transistors $Q1$ to $Q4$ are not width-modulated as in Fig. 5.8. Diagonally opposite transistors are operated simultaneously on at alternate half cycles, but without the "off dead time" between the turnoff of one pair and the turnon of the next pair as was required for the voltage-fed circuit of Fig. 5.9. Each pair in Fig. 5.10 is kept deliberately on for slightly more than a half period by depending on the storage time of slow bipolar transistors or somehow delaying the turnoff time (by ~1 μs) of faster bipolars or MOSFETs. Output voltage regulation is achieved by width-modulating the on time of the buck transistor $Q5$ as was done for the buck-voltage-fed circuit of Fig. 5.9.

Significant advantages accrue from the physical removal of buck filter capacitor $C1$ of Fig. 5.9 and the deliberate overlapping on times of alternate transistor pairs. These advantages can be seen as follows.

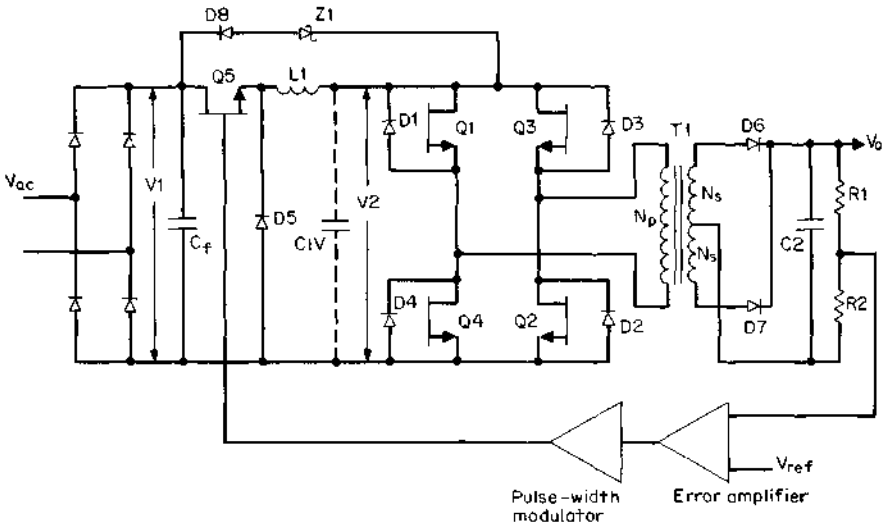


Figure 5.10 Buck-current-fed full-wave bridge. The usual buck filter capacitor $C1$ is omitted. There is a virtual capacitor $C1V$ there—it is all the output capacitors of the master and slaves, reflected back across to the primary. Diagonally opposite transistors are turned on simultaneously. By causing the off-turning and on-turning pair to overlap in the on state for a short time (~1μ s), significant advantages are obtained. During the overlap of the off- and on-turning pairs, the high impedance looking into $L1$ (with $C1$ missing) forces all input and output nodes of the bridge to collapse to ground. It is this high impedance looking back into $L1$ which makes the source driving the bridge have the characteristic of a constant current generator. $Z1$, $D8$ constitute an upper clamp to limit $V2$ when the previously on transistors turn off.

5.6.6.1 Alleviation of turnon-turnoff transient problems in buck-current-fed bridge.[9,10]

For the width-modulated full-wave bridge of Fig. 5.8, Sec. 5.6.2.2 described excessive current, power dissipation stresses in the bridge transistors, and voltage stresses in output rectifier diodes at the instant of turnon. Such stresses do not occur in the current-fed circuit of Fig. 5.10 because of the overlapping on time of alternate transistor pairs and the high impedance seen looking back into $L1$ with no filter capacitor physically present at that node.

This can be seen from Figs. 5.11 and 5.12. Consider in those figures that $Q3$, $Q4$ had been on and $Q1$, $Q2$ commence turning on at T_1. Transistors $Q3$, $Q4$ remain on until T_2 (Fig. 5.12), resulting in an overlap time of $T_2 - T_1$. At T_1, as $Q1$, $Q2$ come on, a dead short circuit appears across the output of $L1$; and since the impedance looking into $L1$ is high, the voltage V_2 collapses to zero (Fig. 5.12*c*). But $L1$ is a large inductor and current in it must remain constant at its initial value I_L. Thus as current in $Q1$, $Q2$ rises from zero toward I_L (Fig. 5.12*f* and 5.12*g*), current in $Q3$, $Q4$ falls from I_L toward zero (Fig. 5.12*d* and 5.12*e*).

But the rising current in $Q1$, $Q2$ occurs at zero voltage at V_2 and

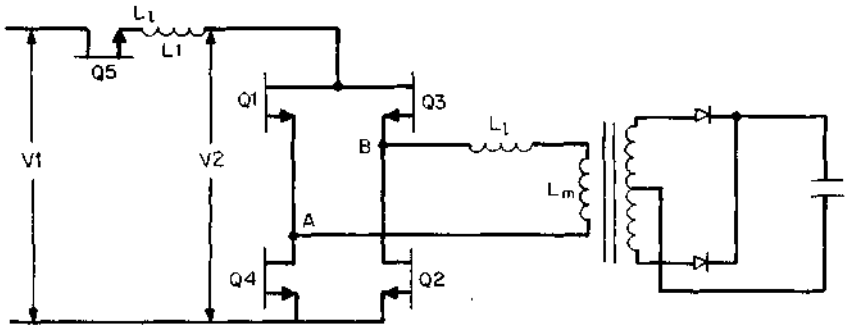


Figure 5.11 During the overlap instant, when all four transistors are on, the voltage $V2$ and that across nodes A, B collapse to zero. Energy stored in leakage inductance L_l is fed to the load via the transformer instead of being dissipated in a snubber resistor or being returned to the input bus as in conventional circuits. Hence, there is no turnon transient dissipation in the bridge transistors or overvoltage stress in output rectifiers.

hence also zero voltage between nodes A and B in Fig. 5.11. There is thus no voltage across $Q1$, $Q2$ as their current rises and there is no dissipation in them. At some time T_3, currents in $Q1$, $Q2$ have risen to $I_L/2$ and currents in $Q3$, $Q4$ have fallen from I_L to $I_L/2$, thus summing to the constant current I_L from inductor $L1$.

Now at T_2, $Q3$ and $Q4$ are commanded off. Assume, as a worst-case scenario, that $Q3$ is slower and $Q4$ turns off first. Since voltage V_2 is zero, $Q4$ turns off with zero voltage across it and hence no turnoff dissipation in it. As I_{Q4} falls from $I_L/2$ toward zero (T_2 to T_4), I_{Q2} rises from $I_L/2$ toward I_L to maintain the constant I_L demanded by $L1$. As I_{Q2} rises from $I_L/2$ toward I_L, I_{Q3} rises from $I_l/2$ to I_L to supply I_{Q2}. Again, since $L1$ demands a constant current I_L, as I_{Q3} rises toward I_L, I_{Q1} falls from $I_L/2$ to zero at T_4.

During the time T_1 to T_4, V_2 has collapsed to zero volts and so has the voltage across the transformer primary (A to B in Fig. 5.11). But current had been stored in the transformer leakage inductance L_L while $Q3$, $Q4$ were on. As voltage A to B collapses, voltage across the primary leakage inductance reverses in an attempt to keep the current constant. Thus the leakage inductance acts like a generator and delivers this stored current or energy through the transformer to the secondary load instead of returning it to the input supply bus or to dissipative snubbers as in conventional circuits.

Now at a later time T_5, the slower transistor $Q3$ starts turning off. As current in it falls from I_L to zero (Fig. 5.12*d*), current I_{Q1} tries to rise from zero to I_L to maintain the constant current I_L demanded by $L1$. But I_{Q1} rise time is limited by the transformer leakage inductance (Fig. 5.12*f*). Since I_{Q3} fall time is generally greater than I_{Q1} rise time,

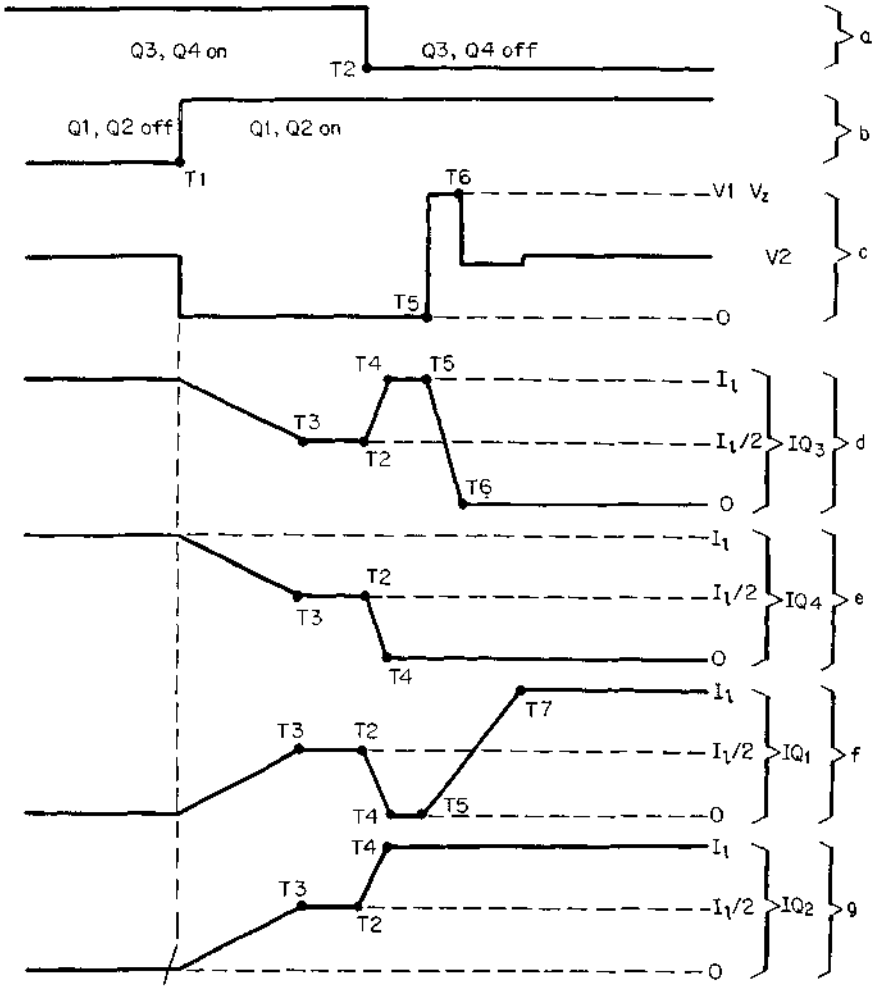


Figure 5.12 Currents in bridge transistors and voltage at bridge input. During overlapping on time of all four bridge transistors in buck-current-fed topology.

voltage V_2 will overshoot its quiescent value and must be clamped to avoid overstressing $Q3$ as its emitter is now clamped to ground by the conducting $Q2$. The clamping is done by a zener diode $Z1$ as shown in Figs. 5.10 and 5.13.

Overshooting of V_2 during the slower transistor ($Q3$) turnoff time results in somewhat more dissipation in it than in the circuit of the conventional width-modulated bridge (Fig. 5.8). This dissipation is $(V_1 + V_z)(I_L/2)(T_6 - T_5)/T$ for Fig. 5.10 but only $V_1(I_L/2)(T_6 - T_5)/T$ for Fig. 5.8. But in Fig. 5.8, there are four transistors which have relatively high turnoff dissipation. In Figs. 5.10 and 5.13, only the two

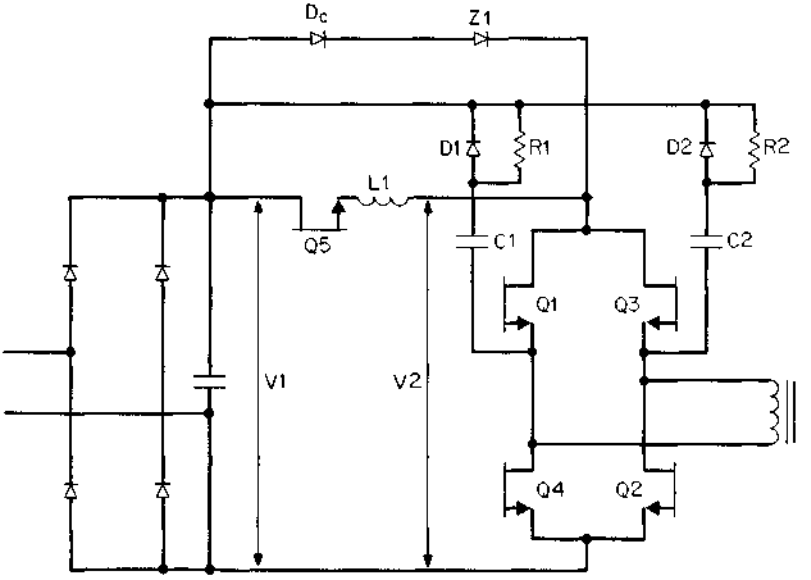


Figure 5.13 In the buck-current-fed bridge, only two turnoff snubbers ($R1$, $C1$, $D1$ and $R2$, $C2$, $D2$) are required. An upper voltage clamp ($Z1$, D_c) is required to limit $V2$ when the last of the off-turning transistors turns off.

transistors with slow turnoff time have high dissipation. As discussed above, the faster transistor suffers no dissipation at turnoff as it turns off at zero voltage; and at turnon, all transistors have negligible dissipation as they have the transformer leakage inductance in series and hence turn on at zero voltage.

The increased dissipation of the two transistors at turnoff can be diverted from the transistors to resistors by adding the snubbing networks $R1$, $C1$, $D1$ and $R2$, $C2$, $D2$ of Fig. 5.13. Design of such turnoff snubbing circuits will be discussed in the later chapter on snubbers.

5.6.6.2 Absence of simultaneous conduction problem in buck-current-fed bridge. In the buck-voltage-fed bridge of Fig. 5.9, care must be taken to avoid simultaneous conduction in transistors positioned vertically above one another ($Q1$, $Q4$ or $Q3$, $Q2$). Such simultaneous conduction provides a low impedance across $C1$. But $C1$ has a low impedance and can supply large currents without its output (V_2) dropping in voltage. Thus the bridge transistors could be subjected to simultaneous high voltage and high current and one or more would immediately fail.

Even if a dead time between the turnoff of one transistor pair and the turnon of the other is designed in to avoid simultaneous conduction, it still may occur under various odd circumstances, such as high temperature and/or high load conditions when transistor storage time

may be much lower than data sheets indicate, or low input voltage (in the absence of maximum on time clamp or undervoltage lockout) as the feedback loop increases on time to maintain constant output voltage.

But in the buck-current-fed bridge, simultaneous conduction is actually essential to its operation and provides the advantages discussed above. Further, in the buck-current-fed bridge, since the on time is slightly more than a half period for each transistor pair, the peak current is less than in the buck-voltage-fed bridge, whose maximum on time is usually set at 90 percent of a half period to avoid simultaneous conduction.

5.6.6.3 Turnon problems in buck transistor of buck-current- or buck-voltage-fed bridge.[10] The buck transistor in either the voltage- or current-fed bridge suffers from a large spike of power dissipation at the instant either of turnon and turnoff, as can be seen in Fig. 5.14.

Consider first the instantaneous voltage and current of $Q5$ during the turnon interval. The locus of rising current and falling voltage during that interval is shown in Fig. 5.14*b*. Just prior to $Q5$ turning on, free-wheeling diode $D5$ is conducting and supplying inductor current I_L. As $Q5$ commences turning on, its collector is at V_1, its emitter is at one diode ($D5$) drop below ground. The emitter does not move up from ground until the current in $Q5$ has risen from zero to I_L and canceled the $D5$ forward current.

Thus during the current rise time of t_r, the I_c-V_{ce} locus is along points A to B. During t_r, the average current supplied by $Q5$ is $I_L/2$ and the voltage across it is V_1. Once current in $Q5$ has risen to I_L, assuming negligible capacitance at the $Q5$ emitter node and fast recovery time in $D5$, the voltage across $Q5$ rapidly drops to zero along the path B to C. If there is one turnon of duration t_r in a period T, the dissipation in $Q5$, averaged over T, is

$$PD_{\text{turnon}} = V_1 \frac{I_L t_r}{2T} \tag{5.10}$$

It is of interest to calculate this dissipation for a 2000-W buck-current-fed bridge operating from the rectified 220-V-AC line. Nominal rectified DC voltage (V_1) is about 300 V, minimum is 270 V, and maximum is 330 V. Assume that the bucked-down DC voltage V_2 is 25 percent below the minimum V_1 or 200 V.

Further, assume the bridge inverter operates at 80 percent efficiency, giving an input power of 2500 W. This power comes from a V_2 of 200 V, and hence the average current in $L1$ is 12.5 A. Assume that $L1$ is large enough so that the ripple current in I_L can be neglected.

Then, for an assumed current rise time of 0.3-μs rise time (easily

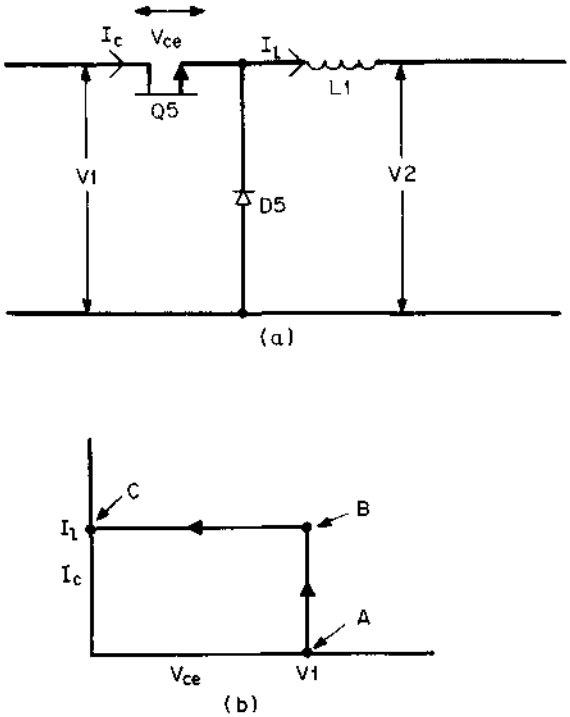


Figure 5.14 (*a*) The buck transistor in the buck-current- or voltage-fed topology has a very unfavorable voltage-current locus at the instant of turnon. It operates throughout its current rise time, at the full input voltage $V1$ until the forward current in free-wheeling diode $D5$ has been canceled. This generates a large spike of dissipation at turnon. (*b*) $I_{ce} = V_{ce}$ locus during turnon of buck transistor $Q5$. Voltage V_{ce} remains constant at $V1$ until the current in $Q5$ has risen to I_l (A to B) and canceled the forward current I_l in free-wheeling diode $D5$. Then, if capacitance at the $Q5$ emitter is low and $D5$ has a fast recovery time, it moves very rapidly to its on voltage of about 1 V (B to C).

achieved with modern bipolar transistors) and a $Q5$ switching frequency of 50 kHz, turnon dissipation at maximum AC input voltage is (from Eq. 5.10)

$$PD = 330\left(\frac{12.5}{2}\right)\left(\frac{0.3}{20}\right) = 31 \text{ W}$$

Note in this calculation that the effect of poor recovery time in $D5$ has been neglected. This has been discussed in Sec. 5.6.2.2 in connection with the poor recovery time of output rectifiers of the bridge inverter. This problem can be far more serious for the free-wheeling di-

ode of the buck regulator, for $D5$ must have a much higher voltage rating—at least 400 V for the maximum V_1 of 330 V—and high-voltage diodes have poorer recovery time than do lower-voltage ones. Thus the $Q5$ current can considerably overshoot the peak of 12.5 A that $D5$ had been carrying. Further, the oscillatory ring after the recovery time, discussed in Sec. 5.6.2.2, can cause a serious voltage overstress in free-wheeling diode $D5$.

This turnon dissipation in $Q5$ and voltage overstress of $D5$ can be eliminated with the turnon snubber of Fig. 5.15.

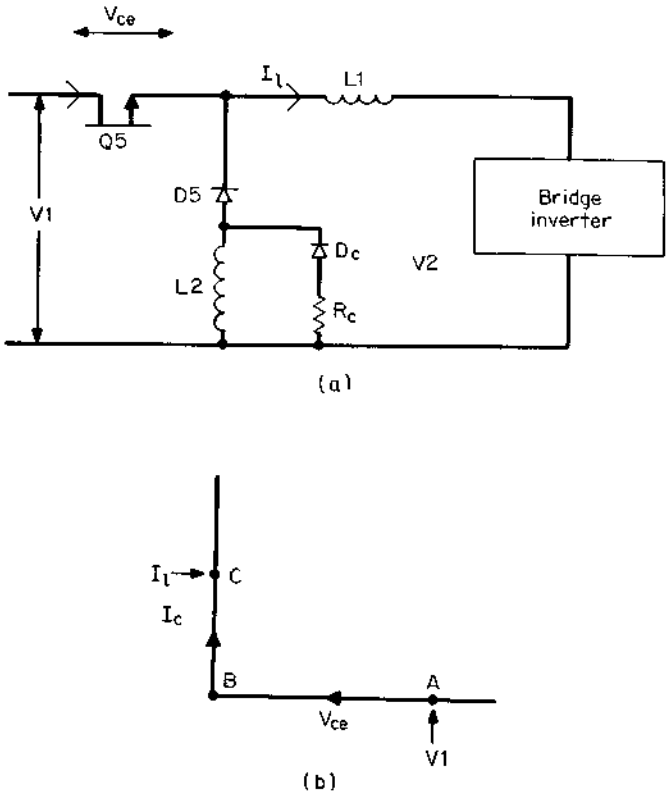


Figure 5.15 (*a*) Turn on snubber—$L2$, D_c, R_c—eliminates turnon dissipation in $Q5$, but at the price of an equal amount of dissipation in R_c. When $Q5$ commences turning on, the presence of $L2$ drives the $Q5$ emitter up to within 1 V of its collector. As $Q5$ current rises toward I_l, the current in $L2$ which has been stored in it by $L1$ during the $Q5$ off time decreases to zero. Thus the voltage across $Q5$ during its turnon time is about 1 V rather $V1$ V. During the next $Q5$ off time, $L2$ must be charged to a current I_l without permitting too large a drop across it. Resistor R_c limits the voltage across $L2$ during its charging time. (*b*) $Q5$'s locus of falling voltage (A to B) and rising current (B to C) during turnon with the snubber of Figure 5.15*a*.

5.6.6.4 Buck transistor turnon snubber—basic operation. The turnon snubber of Fig 5.15*a* does not reduce total circuit dissipation. It only diverts it from the vulnerable semiconductor $Q5$ where it is a potential failure hazard to the passive resistor R_s, which can far more easily survive the heat. It works as follows. An inductor $L2$ is added in series with the free-wheeling diode $D5$. While $Q5$ is off, the inductor load current I_L flows out of the bottom of the bridge transistors, up through the bottom end of $L2$, through free-wheeling diode $D5$, and back into the front end of $L1$. This causes the top end of $L2$ to be very slightly more negative than its bottom end.

Now as $Q5$ commences turning on, it starts delivering current into the cathode of $D5$ in an attempt to cancel its forward current. But this current flows down into $L2$, opposing the load current it is carrying. Since the current in an inductor ($L2$) cannot change instantaneously, the voltage polarity across it reverses instantaneously in an attempt to maintain constant current.

The voltage at the top end of $L2$ rises, pushing the free-wheeling diode up with it until its cathode meets the on-turning $Q5$ emitter. The $Q5$ emitter is forced up to within $V_{ce(sat)}$ of its collector and now $Q5$ continues increasing its current, but at a V_{ce} voltage of about 1 V rather than the V_1 voltage of 370 V it had to sustain in the absence of $L2$.

When the $Q5$ current has risen to I_L(in a time t_r), the forward current in $D5$ has been canceled and $Q5$ continues to supply the load current I_L demanded by $L1$. Since voltage across $Q5$ during the rise time t_r is only 1 V, its dissipation is negligible. Further, because of the high impedance of $L2$ in series with the $D5$ anode, there is negligible recovery time current in $D5$. The current-voltage locus of $Q5$ during the turnon time is shown in Fig. 5.15*b*.

5.6.6.5 Selection of buck turnon snubber components. For the preceding sequence of events to proceed as described, the current in $L2$ must be equal to the load current I_L at the start of $Q5$ turnon and must have decayed back down to zero in the time t_r that current from $Q5$ has risen to I_L. Since the voltage across $L2$ during t_r is clamped to V_1, the magnitude of $L2$ is calculated from

$$L2 = \frac{V_1 t_r}{I_L} \tag{5.11}$$

For the above example, V_1 was a maximum of 370 V, t_r was 0.3 μs, and I_L was 12.5 A. From Eq. 5.11, this yields $L2 = 370 \times 0.3/12.5 = 8.9$ μH.

Now the purpose of R_c, D_c in Fig. 5.15*a* is to ensure that at the start of $Q5$ turnon, current in $L2$ truly is equal to I_L and that it had reached that value without overstressing $Q5$.

Consider for the moment that R_c, D_c were not present. As $Q5$ turned off, since current in $L1$ cannot change instantaneously, the input end of $L1$ goes immediately negative in an attempt to keep current constant. If $L2$ were not present, $D5$ would catch the front end of $L1$ (and hence the $Q5$ emitter) at ground and permit across $Q5$ a voltage of only V_1. But with $L2$ present, the impedance looking out of the $D5$ anode is the high instantaneous impedance of $L2$. Then as $Q5$ turned off, I_L would be drawn through $L2$, pulling its top end very far negative. This would put a very large negative voltage at the $Q5$ emitter, and with its collector at V_1 (370 V in this case), it would immediately fail.

Thus R_c and D_c are shunted around $L2$ to provide a path for I_L at the instant $Q5$ turns off; and R_c is selected low enough so that the voltage drop across it at a current I_L plus V_1 is a voltage stress that $Q5$ can safely take. Thus

$$V_{Q5(\max)} = V_1 + R_c I_L \tag{5.12}$$

In the preceding example, $V_{1(\max)}$ was 370 V. Assume that $Q5$ had a V_{ceo} rating of 450 V. With a −1- to −5-V reverse bias at its base at the instant of turnoff, it could safely sustain the V_{cev} rating of 650 V. Then to provide a margin of safety, select R_c so that $V_{Q5(\max)}$ is only 450 V. Then from Eq. 5.12, $450 = 370 + R_c \times 12.5$ or $R_c = 6.4\ \Omega$.

5.6.6.6 Dissipation in buck transistor snubber resistor. Examination of Fig. 5.15*a* shows that essentially the constant current I_L is charging the parallel combination of R_c and $L2$. The Thevinin equivalent of this is a voltage source of magnitude $I_L R_c$ charging a series combination of R_c and $L2$. It is well known that in charging a series inductor L to a current I_p or energy $\frac{1}{2}L(I_p)^2$, an equal amount of energy is delivered to the charging resistor. Then if L is charged to I_p once per period T, the dissipation in the resistor is $\frac{1}{2}L(I_p)^2/T$.

In the preceding example where $T = 20\ \mu s$, $L = 8.9\ \mu H$, $I_p = 12.5\ A$

$$PD_{\text{snubber resistor}} = \frac{(1/2)(8.9)(12.5)^2}{20}$$

$$= 35\ W$$

Thus, as mentioned above, this snubber has not reduced circuit dissipation; it has only diverted it from the transistor $Q5$ to the snubbing resistor.

5.6.6.7 Snubbing inductor charging time. The snubbing inductor must be fully charged to I_L during the off time of the buck transistor. The charging time constant is L/R, which in the above example is 8.9/

6.4 = 1.39 μs. The inductor is 95 percent fully charged in three time constants or 4.17 μs.

In the preceding example, switching period T was 20 μs. To buck down the input of 370 V to the preregulated 200 V, on time is T_{on} = 20(200/370) = 10.8 μs. This leaves a $Q5$ off time of 9.2 μs, which is sufficient, as the snubbing inductor is 95 percent fully charged in 4.17 μs.

5.6.6.8 Lossless turnon snubber for buck transistor.[10,21,22] Losses in the snubbing resistor of Fig. 5.15*a* can be avoided with the circuit of Fig. 5.16. Here, a small transformer $T2$ is added. Its primary turns N_p and gap are selected so that at a current I_L, its inductance is the same $L2$ of Fig. 5.15*a*. The polarities at the primary and secondary are as shown by the dots.

When $Q5$ turns off, the front end of $L1$ goes negative in an attempt to keep I_L constant. Then I_L is pulled through $D5$ and N_p, producing a negative voltage V_n at the dot end of N_p and producing a voltage stress across $Q5$ of $V_1 + V_n$. Voltage V_n is chosen so that $V_1 + V_n$ is a voltage that $Q5$ can safely sustain. To maintain the voltage across N_p at V_n when $Q5$ has turned off, the turns ratio N_s/N_p is selected equal to V_1/V_n. Now when $Q5$ turns off, as the dot end of N_p goes down to V_n, the no-dot end of N_s goes positive and is clamped to V_1, holding the voltage across N_p to V_n.

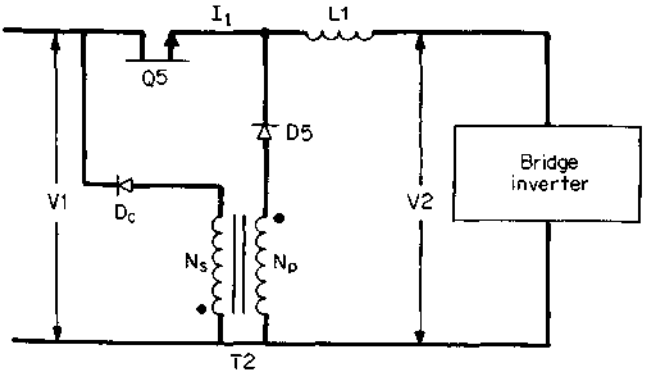


Figure 5.16 Dissipationless turnon snubber. When $Q5$ turns off, $L1$ stores a current I_L in N_p of $T1$. But the negative voltage at the dot end of N_p during this charging time is fixed by the turns ratio N_s/N_p. If the top end of N_p is to be permitted to go only V_n negative when $Q5$ turns off, the voltage stress on $Q5$ is $V1 + V_n$. When the dot end of N_p has gone negative by V_n, the no-dot end of N_s has been driven up to $V1$, D_c latches in and clamps to $V1$, clamping the voltage across N_p to the preselected V_n. Thus N_s/N_p is chosen as $V1/V_n$. The charging of N_p is thus not limited by a resistor as in Fig. 5.14 and hence there is no snubber dissipation.

Prior to Q5 turn on, L1 current flows through N_p, D5. As Q5 commences turning on, its emitter looks into the high impedance of N_p and immediately rises to within one volt of its collector. Thus, current in Q5 rises with only one volt across it, and its dissipation is negligible. All the energy stored in N_p when Q5 was off, is returned via L1 to the load with no dissipation. Q5 turn off dissipation can be minimized with a turn off snubber (Chap. 11).

5.6.6.9 Design decisions in buck-current-fed bridge. The first decision to be made on the buck-current-fed bridge is when to use it. It is primarily a high-output-power high-output-voltage topology.

In terms of cost, efficiency, and required space, it is a good choice for output powers in the range of 1 to 10 or possibly 20 kW. For high output voltages—above about 200 V—and above about 5 A output current—the absence of output inductors makes it a good choice. For output powers above 1 kW, the added dissipation, volume, and cost of the buck transistor is not a significant increase above what is required in a competing topology such as a width-modulated full-wave bridge.

It is an especially good choice for a multioutput supply consisting of one or more high output voltages (5000 to 30,000 V). In such applications, the absence of output inductors permits the use of capacitor-diode voltage multiplier chains.[8,13] Also, the absence of output inductors in the associated lower output voltages partly compensates for the cost and volume of the buck transistor and its output inductor.

The next design decision is the selection of the bucked-down voltage (V_2 of Fig. 5.10). This is chosen at about 25 percent below the lowest ripple trough at V_1 (Fig. 5.10) at the lowest specified AC input. Inductor $L1$ is chosen so as not to go discontinuous at the calculated minimum inductor current I_L corresponding to the minimum total output power at the preselected value of V_2. It is chosen exactly as in Sec. 1.3.6 for a conventional buck regulator.

The output capacitors are chosen, not to provide storage or reduce ripple directly at the output nodes because the overlapping conduction of bridge transistors minimizes this requirement. Rather they are chosen so that when reflected into the primary, the equivalent series resistance R_{esr} of all reflected capacitors is sufficiently low as to minimize ripple at V_2. Recall from Sec. 1.3.7, in calculating the magnitude of the output capacitor, it was pointed out that output ripple in a buck regulator V_{br} is given by

$$V_{br} = \Delta I R_{esr}$$

where ΔI is the peak-to-peak ripple current in the buck inductor and is usually set at twice the minimum DC current in it so that the inductor

just about goes discontinuous at its minimum DC current. Minimum DC current in this case is the current at minimum specified output power at the preselected value of V_2. Thus with R_{esr} selected so as to yield the desired ripple at V_2, ripple at each secondary is

$$V_{sr} = V_{br} \frac{N_s}{N_p}$$

There is an interesting contrast in comparing a current to a voltage-fed bridge at the same bucked-down voltage (V_2 of Figs. 5.9, 5.10).

For the voltage-fed bridge, a maximum on time of 80 percent of a half period must be established to ensure that there is no simultaneous conduction in the two transistors positioned vertically above one another. With the low impedance looking back into the buck regulator of the voltage-fed circuit, such simultaneous conduction would subject the bridge transistors to high voltage and high current and destroy one or more of them.

In the current-fed circuit, such slightly overlapping simultaneous conduction is essential to its operation and on time of alternate transistor pairs is slightly more than a full half period at any DC input voltage. In addition, since the on time of a voltage fed bridge (Fig. 5.9) is only 80 percent of a half period, its peak current must be 20 percent greater than that of the current-fed bridge at the same output power.

It should also be noted that the number of primary turns as calculated from Faraday's law (Eq. 2.7) must be 20 percent greater in the current-fed bridge as the on time is 20 percent greater for a flux change equal to that in a voltage-fed bridge at the same V_2.

5.6.6.10 Operating frequencies—buck and bridge transistors. The buck transistor is synchronized to and operates at twice the square-wave switching frequency of the bridge transistors. It will be recalled that it alone is width-modulated, and that the bridge devices are operated at a 50 percent duty cycle with a slightly overlapping on time.

Frequently, however, the scheme of Fig. 5.17*a* is used to reduce dissipation in the buck transistor. It is seen there that two buck transistors (*Q5A* and *Q5B*) are used. They are synchronized to the bridge transistor frequency and are turned on (and width-modulated) at alternate half cycles of the bridge square-wave frequency. Thus the DC and whatever AC switching losses there may be are shared between two transistors with a resulting increase in reliability.

5.6.6.11 Buck-current-fed push-pull topology. The buck-current-fed circuit can also be used to drive a push-pull circuit as in Fig. 5.18 with the consequent saving of two transistors over the buck-current-fed

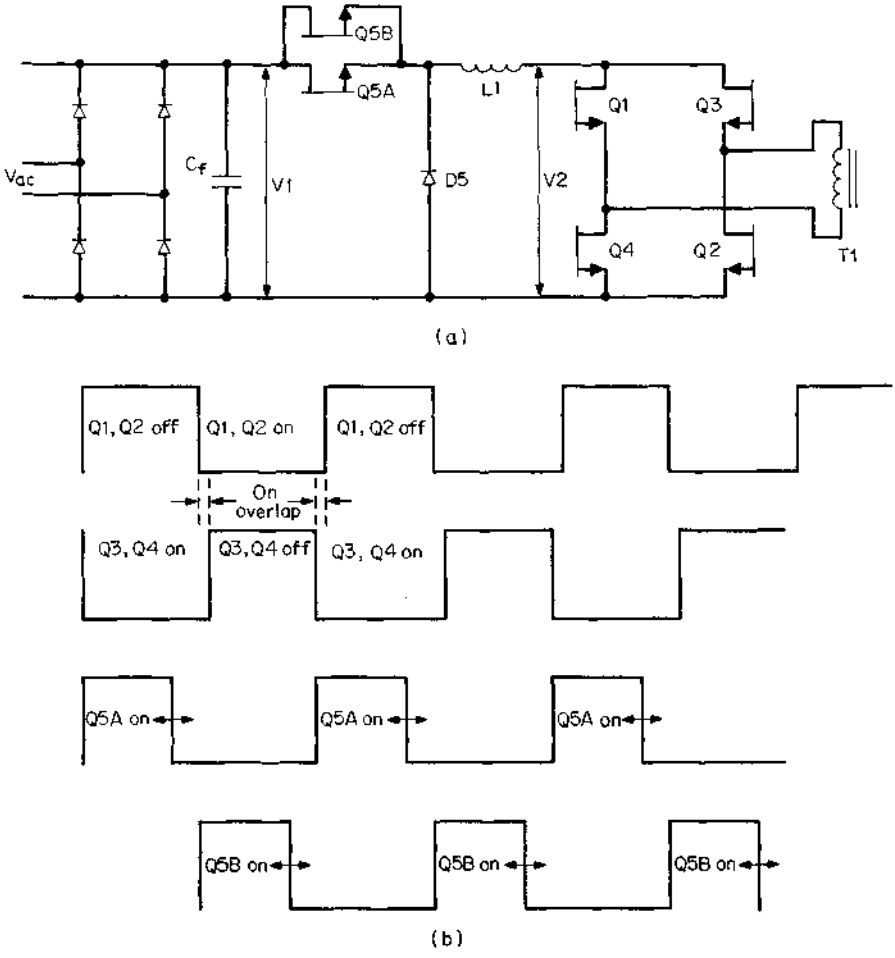


Figure 5.17 (*a*) Buck transistor *Q5* may be a single transistor operating at twice the frequency of the bridge transistors and synchronized to them or, more usually, two synchronized transistors, which are width-modulated and are each on at alternate half periods of the bridge transistors. (*b*) To reduce dissipation in the buck transistor, it is usually implemented as two transistors, each synchronized to the bridge transistors and operated at the same square-wave frequency as the bridge devices. Transistors *Q5A*, *Q5B* are width-modulated. Bridge transistors are not and are operated with a small "on" overlap time.

bridge. Most of the advantages of the buck-current-fed bridge are realized and the only disadvantage is that the push-pull circuit puts greater voltage stress on its power transistors. This voltage stress is now twice V_2 rather than V_2 as in the bridge circuit. But V_2 is the preregulated and bucked-down input voltage—usually only 75 percent of the minimum V_1 input. This is usually about the same as the max-

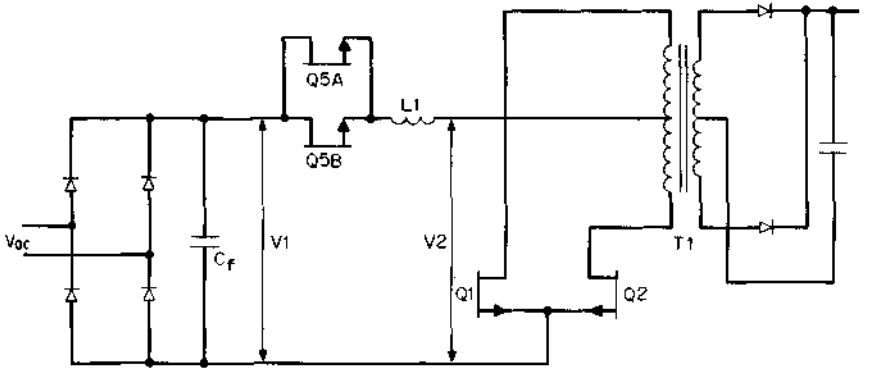


Figure 5.18 The current-fed topology can also be implemented as a buck–push-pull circuit. As in the buck-current-fed bridge, the capacitor after the buck inductor $L1$ is omitted and $Q1$, $Q2$ are operated with a deliberately overlapping on time. Only buck transistors $Q5A$, $Q5B$ are width-modulated. Output inductors are not used. All the advantages of the buck-current-fed bridge are retained. Although off-voltage stress is twice $V2$ (plus a leakage spike) instead of $V2$ as in the bridge, that is still significantly less than twice $V1$ as $V2$ is bucked down to about 75 percent of the minimum value of $V1$. This circuit is used at lower power levels than the buck-current-fed bridge and offers the saving of two transistors.

imum DC input of a competing topology—the width-modulated full-wave bridge (Fig. 5.8).

However, the major advantages of the current-fed technique—no output inductors and no possibility of flux—imbalance still exist.

The topology can be used to greatest advantage in supplies of 2 to 5 kW and especially if there are multiple outputs or at least one high voltage output.

5.6.7 Flyback-current-fed push-pull topology (Weinberg circuit; Ref. 23)

This topology[1,23] is shown in Fig. 5.19. It is seen to have a flyback transformer in series with a push-pull inverter. It has many of the valuable attributes of the buck-current-fed push-pull (Fig. 5.18) topology, and since it requires no width-modulated input transistor ($Q5$), it has lesser dissipation, cost, volume, and greater reliability.

It might be puzzling at first glance to see how the output voltage is regulated against line and load changes since there is no LC voltage-averaging filter at the output. The diode-capacitor at the output is a peak, rather than an averaging, circuit. The answer, of course, is that the averaging or regulating is done at the $T1$ center tap to keep V_{ct} relatively constant. The output voltage (or voltages) is (are) kept constant by width modulating the $Q1$, $Q2$ on time. Output voltage is simply $(N_s/N_p)V_{ct}$ and a feedback loop sensing V_o controls the $Q1$, $Q2$ on times to keep V_{ct} at the correct value to maintain V_o constant. The

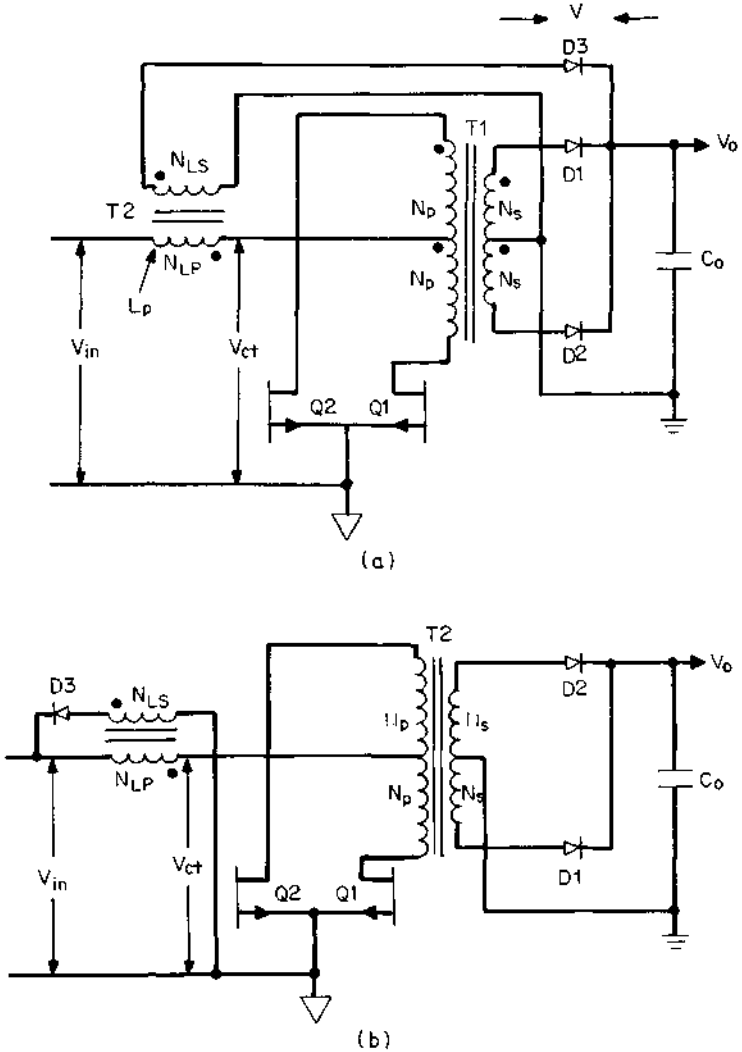


Figure 5.19 (*a*) Flyback-current-fed push-pull topology (Weinberg circuit; Ref. 23). This is essentially a flyback transformer in series with a width-modulated push-pull inverter. Its usage is primarily as a multioutput supply with one or more high-voltage outputs as it requires no output inductors and only the one input flyback transformer $T1$. The high impedance seen looking back into the primary of $T2$ makes it a "current fed" with all the advantages shown in Fig. 5.18. Here the $T2$ secondary is shown clamped to V_o. Transistors $Q1$, $Q2$ may be operated either with a "dead time" between their on times or with overlapping on times. Its advantage over Fig. 5.18 is that it requires no input switching transistors. Usual output power level is 1 to 2 kW. (*b*) Same circuit as Fig. 5.19*a* but with $D3$ clamped to V_{in}. This results in lesser input current ripple but more output voltage ripple.

relation between the $Q1$, $Q2$ on times and output voltage will be shown below.

The circuit retains the major advantage of the current-fed technique—a single-input inductor but no output inductors, which makes it a good choice for a multioutput supply with one or more high-voltage outputs. Further, because of the high source impedance of the flyback transformer primary $L1$, the usual flux-imbalance problem of voltage-fed push-pulls does not result in transformer saturation and consequent transistor failure. Its major usage is at the 1- to 2-kW power level.

Two circuit configurations of the flyback-current-fed push-pull topology are shown in Fig. 5.19*a* and 5.19*b*. Figure 5.19*a* shows the $T2$ secondary returned to the output voltage through diode $D3$; in Fig. 5.19*b* the diode is returned to the input voltage. When the diode is returned to V_o, output ripple voltage is minimized; when it is returned to V_{in}, input ripple current is minimized. Consider first the configuration of Fig. 5.19*a*, where the diode is returned to the output.

The configuration of Fig. 5.19*a* can operate in two significantly different modes. In the first mode, $Q1$ and $Q2$ are never permitted to have overlapping on times at any DC input voltage. In the second mode, $Q1$ and $Q2$ may have overlapping on times throughout the entire range of specified DC input voltage or may have overlapping on times through some lower range of input voltage and automatically move (under control of the feedback loop) into the nonoverlapping mode at higher DC input voltage.

It will be shown below that in the nonoverlapping mode, power is delivered to the secondaries at a center tap voltage V_{ct} lower than the DC input voltage (bucklike operation) and in the overlapping mode, power is delivered to the secondaries at a center tap voltage V_{ct} higher than the DC input voltage (boostlike operation). Thus in the nonoverlapping mode, since V_{ct} is relatively low, $Q1$, $Q2$ currents are relatively high for a given output power. But with the lower V_{ct} voltage, off-voltage stress in $Q1$, $Q2$ is relatively low. In the overlapping mode, since V_{ct} is higher than V_{in}, $Q1$, $Q2$ currents are lower for a given output power but off-voltage stress in $Q1$, $Q2$ is higher than that for the nonoverlapping mode.

The circuit is usually designed not to remain in one mode throughout the full range of input voltages. Rather, it is designed to move from the overlapping mode with an on duty cycle T_{on}/T greater than 0.5 through to the nonoverlapping mode T_{on}/T less than 0.5 as the DC input voltage shifts from its minimum to its maximum specified values. This permits proper operation throughout a larger range of DC input voltages than if operation remained within one mode throughout the entire range of DC input voltage.

5.6.7.1 Absence of flux-imbalance problem in flyback-current-fed push-pull topology. Flux imbalance is not a serious problem in this topology because of the high-impedance current-fed source feeding the push-pull transformer center tap.

The current-fed nature of the circuit arises from the flyback transformer $T2$, which is in series with the $T2$ center tap. The impedance looking back from the $T1$ center tap is the high impedance of the magnetizing inductance of the $T2$ primary.

Now in a conventional voltage-fed push-pull inverter, unequal volt-second product across each half primary causes the flux-imbalance problem (Sec. 2.2.5). The transformer core moves off center of its hysteresis loop and moves toward saturation. Because of the low impedance of a voltage source, current to the $T2$ center tap is unlimited and the voltage at that point (V_{ct}) remains high. The core then moves further into saturation, where its impedance eventually vanishes and transistor currents increase drastically. With high currents and high voltage across the transistor, it very soon fails.

However, as shown in Fig. 5.19, with the high impedance looking back into the dot end of N_{LP}, as the $T1$ core tries to move far into saturation drawing larger currents than the load itself demands, the high currents cause a voltage drop at V_{ct}. This reduces the volt-second product on the primary half which is moving toward saturation and prevents complete core saturation.

Thus the high source impedance of N_{LP} does not fully prevent core saturation. In the worst case, it keeps the core close to the knee of the BH loop, which is sufficient to keep transistor currents from rising to disastrous levels. The major drawback of push-pull circuit flux imbalance is thus not a problem with this inverter.

5.6.7.2 Decreased push-pull transistor currents in flyback-current-fed topology. In conventional width modulated push-pulls driven at their center tap from a low-impedance voltage source, it is essential to avoid simultaneous conduction in the transistors by providing a dead time of about 20 percent of a half period between turnoff of one transistor and turnon of the other. Thus if maximum on time per transistor must be cut back from 100 percent to 80 percent of a half period, it results in a higher peak transistor current for the same output power as the output power is proportional to the average transistor current.

This dead time is essential in the voltage-fed push-pull, for if $Q1$, $Q2$ were simultaneously on, the half primaries could not sustain voltage. Then the transistor collectors would rise to the supply voltage, which would remain high, and with high voltage and high current, the transistors would fail.

But with the high impedance looking back into the dot end of N_{LP}, if

both transistors were simultaneously on (under fault conditions when the DC input voltage is momentarily lower than specified and with storage times greater than specified), there is no problem. For with both transistors simultaneously on for a brief instant, V_{ct} simply drops to zero and current drawn from the input source is limited by the impedance of the $T2$ primary.

Thus even in the "nonoverlap" mode, no dead time need be provided between the turnoff of one transistor and the turnon of the other. If there is a momentary overlap because of storage time, V_{ct} simply collapses to zero and no harm results. Hence by eliminating the 20 percent dead time, the peak current required for a given output power is decreased by 20 percent at the same value of V_{ct}. Further, as discussed above, in the overlap mode, overlapping on time need not be the small amount arising from transistor storage time but can be a deliberately large fraction of a half period.

5.6.7.3 Nonoverlapping mode in flyback-current-fed push-pull topology—basic operation. The circuit operation can be understood from examination of the significant voltage and current waveforms shown in Fig. 5.20.

Operation will be explained on the assumption that the "on" potentials of transistors $Q1$, $Q2$ are negligibly small and can be neglected. For considering their actual on drop of about 1 V would only complicate the design equations and hamper understanding of really important circuit behavior. Also, the forward drop V_d of diodes $D1$, $D2$, $D3$ are assumed equal.

Now in Fig. 5.19*a*, when either $Q1$ or $Q2$ is on, the voltage across the corresponding half secondary is clamped to $V_o + V_d$. Then (assuming that transistor on drop is zero) the voltage at the $T1$ center tap V_{ct} is clamped to $(N_p/N_s)(V_o + V_d)$ as can be seen in Fig. 5.20*d*. Now N_p/N_s is chosen so that V_{ct} is 25 percent lower than the bottom of the input ripple trough at the lowest specified value of V_{in}.

Thus when either transistor is on, the dot end of N_{LP} is negative with respect to its no-dot end and current flows through to the $T1$ center tap which is clamped to the voltage V_{ct}. The waveshapes of the currents which flow are shown in Fig. 5.20*g* and 5.20*h*. These currents have the ramp-on-a-step waveshape characteristic of any buck regulator operating in the continuous conduction mode as discussed in Sec. 1.3.2.

Now when the on transistor turns off, the dot end of N_{LP} goes positive in an attempt to maintain the L_p current constant. The dot end of N_{LS} also goes positive until $D3$ gets forward-biased and clamps to V_o. The ratio N_{LP}/N_{LS} is set equal to N_p/N_s of $T1$ (hereafter those ratios will be designated as N). The reflected voltage on the $T2$ primary is

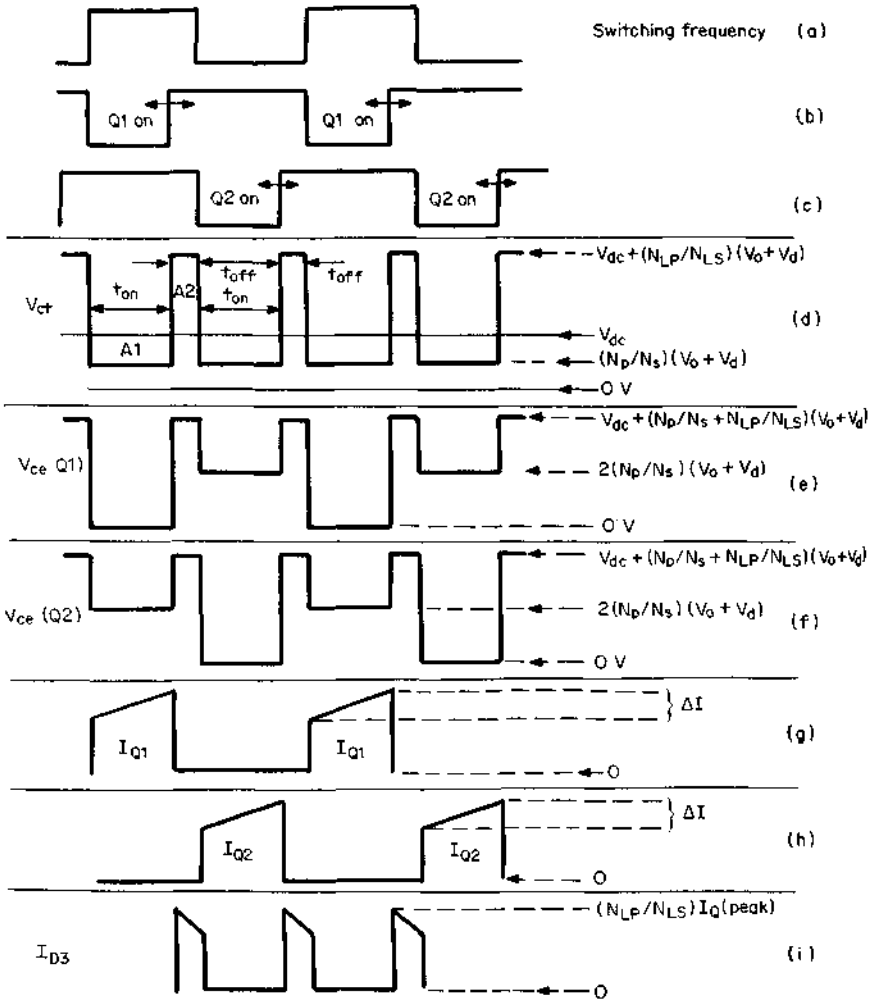


Figure 5.20 Significant voltage-current waveforms in flyback-current-fed topology when operating in nonoverlapping conduction mode. Power is delivered to the load only when either $Q1$ or $Q2$ is on. Power is delivered at a supply voltage of $(N_p/N_s)(V_o + V_d)$, which is less than V_{dc}.

then $N(V_o + V_d)$. Thus when either transistor turns off, V_{ct} rises to $V_{dc} + N(V_o + V_d)$ and stays there until the opposite transistor turns on as shown in Fig. 5.20*d*. The waveshape of Fig. 5.20*d* permits calculation of the relation between output voltage and on time as follows.

5.6.7.4 Output voltage–on-time relation in nonoverlapping mode of flyback-current-fed push-pull topology. Refer to Fig. 5.20*d*. It is seen that during t_{on}, V_{ct} is at a voltage level $N(V_o + V_d)$ and during t_{off} it is at a level of $V_{dc} + N(V_o + V_d)$. But the average voltage at this point

must be equal to the V_{dc}, the DC voltage at the front end of L_p. This is because L_p is assumed to have negligible DC resistance, and hence the voltage across it, averaged over a full or half period must equal to zero. Or another way of expressing this is that in Fig. 5.20*d*, the volt-second area $A1$ must equal to the volt-second area $A2$.

But area $A1 = [V_{dc} - N(V_o + V_d)]t_{on}$; and

$$V_{dc}t_{on} - NV_o t_{on} - NV_d t_{on} = NV_o t_{off} + NV_d t_{off}$$

or $NV_o(t_{on} + t_{off}) = V_{dc}t_{on} - NV_d(t_{on} + t_{off})$ but $t_{on} + t_{off} = \dfrac{T}{2}$

Then
$$V_o = \left(\frac{2V_{dc}t_{on}}{NT}\right) - V_d$$

or
$$V_o = \left(2V_{dc}\frac{N_s}{N_p}\right)\frac{t_{on}}{T} - V_d \tag{5.13}$$

Thus the feedback loop regulates V_o by width-modulating t_{on}, just as in all previous circuits—to keep the product $V_{dc}t_{on}$ constant.

5.6.7.5 Output voltage and input current ripple in nonoverlapping mode. Choosing N_{LP}/N_{LS} equal to N_p/N_s in Fig. 5.19*a* results in negligible ripple at V_o, as can be seen in Fig. 5.21. The voltages delivered to the anodes of either $D1$, $D2$, or $D3$ are all equal in amplitude. The currents delivered through $D1$, $D2$ are NI_{Q1} and NI_{Q2}, which are equal. Further, since $N_{LP}/N_{LS} = N_p/N_s$ during t_{off}, the current delivered through $D3$ is also NI_{Q1}. Thus there is no gap in time during which C_o must supply or absorb current. The total load current is at all times being supplied through $D1$, $D2$, or $D3$ and C_o need serve no energy storage function.

Output ripple is then the product of the secondary ripple current amplitude ΔI_s times the equivalent series resistance R_{esr} of C_o. Further, $\Delta I_s = N\,\Delta I_p$ where ΔI_p is set as twice the minimum current at the center of the $Q1$, $Q2$ ramps at minimum output power; ΔI_p is set at the desired value as discussed in Sec. 1.3.6 by choosing L_p sufficiently large that it does not run dry or go into the discontinuous conduction mode before minimum output power. Then C_o is chosen to minimize R_{esr} as discussed in Sec. 1.3.7.

There may be very narrow (<1 μs) voltage spikes at each turnon/turnoff transition as can be seen in Fig. 5.21. These may occur if the voltage fall time at the anode of $D1$, $D2$, or $D3$ is slightly faster than the voltage rise time of the next on-turning diode. Such spikes are easily eliminated with very small LC integrators.

Current drawn from V_{in} is discontinuous. As seen in Fig. 5.20*i*, although $T2$ secondary current flows during the off time, making output

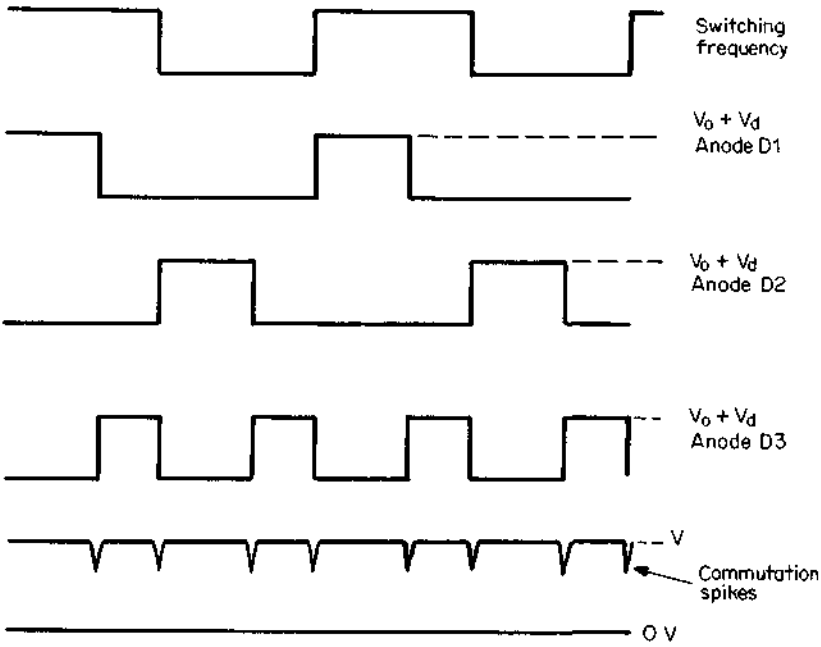


Figure 5.21 In Fig. 5.19*a*, peak voltages applied to the anodes of $D1$, $D2$, $D3$ are all equal if N_p/N_s is chosen equal to N_{LP}/N_{LS}. But if fall time at one anode is slightly faster than the rise time of the next on-coming anode, there will be narrow commutation spikes at the output.

currents continuous, input current falls to zero during each off time. These discontinuous input currents require the addition of a space-consuming RFI input filter to keep large transient currents off the input lines. By returning the $T2$ secondary and $D3$ to the input lines as in Fig. 5.19*b*, input currents will never fall to zero but will ramp up and down with the amplitudes shown in Fig. 5.20*g* and 5.20*h*. This will lower the size of the required RFI input filter greatly or may even make it unnecessary.

5.6.7.6 Output stage–transformer design example—nonoverlapping mode. It is instructive to run through a typical design example of the output stage and transformer for the flyback-current-fed push-pull topology operating in the nonoverlapping mode.

Only a single master secondary will be considered. Additional secondaries may be added, and their required turns will be related to the master in the ratio of their output voltages. Slave secondary output voltages will track the master to within about 2 percent—much more closely than is possible in supplies with output LC filters.

The design example will assume the following:

Output power	2000 W
Output voltage	48 V
Efficiency	80 percent
Switching frequency	50 kHz (T = 20 μs)
Diode voltage drops	1 V
DC input voltage (from 115-V, ±15 percent AC line)	Maximum, 184 V; nominal, 160 V, minimum, 136 V

The first decision to be made is selection of V_{ct} during the on time. As in Sec. 5.6.7.3, V_{ct} is set at 75 percent of the minimum DC input voltage or $0.75 \times 136 = 102$ V.

The turns ratio is now selected to yield 102 V during the on time. From Fig. 5.20*d*, V_{ct} during the on time is

$$V_{ct} = \frac{N_p}{N_s}(V_o + V_d) \qquad \text{or} \qquad \frac{N_p}{N_s} = N = \frac{102}{48 + 1} = 2$$

Now transistor current amplitudes I_{Q1}, I_{Q2} and on times are calculated to permit calculation of primary and secondary rms currents and hence wire sizes. From the on times and preselected V_{ct}, the number of primary turns will be calculated from Faraday's law once a transformer core area is selected for the specified output power.

Thus from Eq. 5.13:

$$\frac{t_{on}}{T} = \frac{(V_o + V_d)(N_p/N_s)}{2V_{dc}}$$

From this, Table 5.1 can be constructed.

Now when either transistor is on, it delivers a current whose waveshape is as shown in Fig. 5.20*g* and 5.20*h*. This current is delivered to the $T1$ center tap at V_{ct} of 102 V. When the transistor turns off, the current shown in Fig. 5.20*i* is delivered via $D3$ to the secondary load. This last current (NI_Q) is delivered during the off time at a voltage $V_o + V_d$ but is equivalent to current I_Q being delivered at a

TABLE 5.1

V_{dc}, V	t_{on}/T	t_{on}, μs
200	0.245	4.9
184	0.266	5.3
160	0.306	6.12
136	0.360	7.2

voltage $N(V_o + V_d)$. Thus, effectively, this current is being delivered at 100 percent duty cycle at a voltage of $N(V_o + V_d)$ or at 102 V.

Assuming an efficiency of 80 percent outward from the $T1$ center tap, input power at the center tap is 2000/0.8 or 2500 W. Average current into the center tap is then 2500/102 = 24.5 A. This is very closely the current at the center of the current ramp in Fig. 5.20*g* and 5.20*h*.

Current in each $T1$ half primary is thus very closely approximated by an equivalent flat-topped pulse whose amplitude I_{pk} is 24.5 A and whose duration is given in Table 5.1. The rms value of this is $I_{pk}\sqrt{t_{on}/T}$. This is a maximum at minimum V_{dc}, and hence at V_{dc} = 136 V the rms current in each half secondary is $I_{rms} = 24.5\sqrt{0.36} = 14.7$ A. At a current density of 500 circular mils per rms ampere, the required number of circular mils for each half primary is 500 × 14.7 = 7350 circular mils.

The transformer core will be selected from charts in the coming section on magnetics design, which has been discussed in Sec. 2.2.9.1. Jumping ahead then, these charts will show that a Ferroxcube EC70 (also an international standard type) core with an area of 2.79 cm^2 can deliver 2536 W at 48 kHz and can be used.

The number of turns per half primary is calculated from Faraday's law (Eq.1.17) at the maximum on time (7.2 μs) in Table 5.1 and at a primary voltage of 102 V. At 50 kHz, using Ferroxcube material type 3F3 core, losses are only 60 mW/cm^3 at a peak flux density of 1600 G. For a core volume of 40.1 cm^3, total core losses are only 2.4 W. This is low enough so that copper losses of even twice that much will still leave the transformer at a safely low temperature. Then from Faraday's law $N_p = V_p t_{on} \times 10^{+8}/A_e \Delta B = 102(7.2 \times 10^{-6})10^{+8}/2.79 \times 3200 = 8$ turns, and for $N_p/N_s = 2$, each half secondary has 4 turns.

Finally, the secondary wire size must be calculated. Each half secondary delivers the characteristic ramp-on-a-step waveform shown in Fig. 5.20*g* and 5.20*h*. The current at the center of the ramp is the DC output current. Their maximum width occurs at minimum DC input of 136 V and is seen in Table 5.1 to be 7.2 μs. For purposes of calculating their rms value, they can be approximated by a rectangular pulse of amplitude I_{dc} and width 7.2 μs. For an output power of 2000 W, the DC output current is 2000/48 = 41.6 A. The rms value of this rectangular current pulse of 41.6 A, 7.2-μs width, and coming once per 20 μs is $41.6\sqrt{7.2/20} = 24.9$ A.

At a current density of 500 circular mils per rms ampere, the required wire area for the half secondaries is 500 × 24.9 = 12,450 circular mils. For such large area, the secondary would most likely be wound with metal foil of thickness and width to yield the required circular mil area. The primary with a required 7350-circular-mil area would be wound with a number of paralleled small diameter wires.

5.6.7.7 Flyback transformer for design example of Sec. 5.6.7.6. The preceding design was based on the inductor L_p operating in the continuous mode or not running dry until the minimum DC current passed through it (Sec. 1.3.6). It was shown there that the discontinuous mode starts when ΔI, the peak-to-peak ramp amplitude, is equal to twice the minimum DC current, which is the current in the center of the ramp at minimum output power.

In the design example above, assume that minimum output power is one-tenth the nominal output power. At nominal output power, current at the center of the ramp was calculated above as 24.5 A. Then current at the center of the ramp at minimum output power is 2.45 A and the peak-to-peak ramp amplitude (ΔI of Fig. 5.20*g* and 5.20*h*) is 4.9 A. But $\Delta I = V_L t_{on}/L_p$, where V_L is the voltage across L_p during t_{on}; and from Table 5.1, at V_{dc} = 136 V, t_{on} = 7.2 μs. Then

$$L_p = \frac{(136 - 102)(7.2 \times 10^{-6})}{4.9} = 50\ \mu H$$

Wire size for the $T2$ secondary must be calculated at high DC input, for it is then that its current pulse width t_{off} is greatest (Table 5.1). Further, as above, the equivalent flat-topped pulse amplitude is I_{dc}. Then from Table 5.1, its maximum width is 10 − 5.3 = 4.7 μs. Maximum rms current in the $T2$ secondary is then

$$I_{rms(T2\ secondary)} = 41.6\sqrt{\frac{t_{off}}{0.5T}}$$

$$= 41.6\sqrt{\frac{4.7}{10}}$$

$$= 28.5\ A$$

At 500 circular mils per rms ampere, that winding requires a circular-mil area of 500 × 28.5 = 14,260 circular mils.

Wire size for the $T2$ primary is calculated for minimum DC input voltage, for it is then that primary current has greatest width and hence largest rms value. Then from Table 5.1

$$I_{rms(T2\ primary)} = 24.5\sqrt{\frac{t_{on}}{0.5T}} = 24.5\sqrt{\frac{7.2}{10}} = 20.7\ A$$

At 500 circular mils per rms ampere, it requires 500 × 20.7 = 10,394 circular mils.

Both primary and secondary would most likely be wound with metal foil rather than round wire for those large required areas.

And of course, since $T2$ is a flyback transformer with no secondary

current flowing when primary current flows, all the primary current drives the core toward saturation. To maintain the required 50-μH primary inductance at 24.5 A of primary current, the core must be a gapped ferrite, powdered Permalloy, or powdered iron type (Sec. 4.3.3).

5.6.7.8 Overlapping mode in flyback-current-fed push-pull topology—basic operation.[14] In the nonoverlapping mode of Fig. 5.20 ($t_{on}/T < 0.5$) it is difficult to accommodate a large ratio of maximum to minimum DC input voltage. Since the maximum on time is $0.5T$ at minimum DC, then at high DC inputs, on time will be a small fraction of a period and may approach 1 to 3 μs at a 100- to 50-kHz switching rate. But bipolar transistors with their appreciable storage times cannot work reliably down to such low on times.

By operating with overlapping on times ($T_{on}/T > 0.5$) as in Fig. 5.22, however, a much larger range of maximum to minimum DC input voltages is possible.

The usual integrated-circuit pulse-width-modulating chips cannot be used for the overlapping mode as their two 180° out-of-phase outputs have a maximum duty cycle D of only 0.5. But a number of schemes using several discrete integrated-circuit packages and capable of a duty cycle of 0 to 100 percent have been described in the literature.[14,19]

Overlapping-mode operation is achieved using the same circuit as shown in Fig. 5.19 but by proper choice of the turns ratios N_p/N_s and N_{LP}/N_{LS} (hereafter designated N_1, N_2, respectively). Circuit operation will be described for the scheme of Fig. 5.19*a*, where diode $D3$ is returned to the output voltage rather than the input. It will be recalled that connection minimizes output voltage ripple rather than input current ripple.

The circuit operation can be understood from the waveforms of Fig. 5.22 and the circuit of Fig. 5.19*a*. For the overlapping mode, $Q1$ and $Q2$ are simultaneously on during intervals T_1 and only one of these is on during T_{off} intervals (t_2 to t_3 when $Q2$ is off and t_4 to t_5 when $Q1$ is off). Power is delivered to the load only during the T_{off} times.

When both transistors are on, the half primaries cannot support voltage and the $T1$ center tap voltage falls to zero as seen in Fig. 5.22*d*. The full input voltage V_{dc} is applied across the $T2$ primary inductance L_p and current ramps up it linearly at a rate $dI/dT = V_{dc}/L_p$. This current divides roughly evenly between $Q1$ and $Q2$ and is seen as the upward-going ramps at the ledges at t_3 to t_4 and t_5 to t_6. During T_1, $D3$ is reverse-biased, and there is no voltage across the $T1$ secondaries. Hence all the output power is supplied from the output filter capacitor C_o during T_1 intervals.

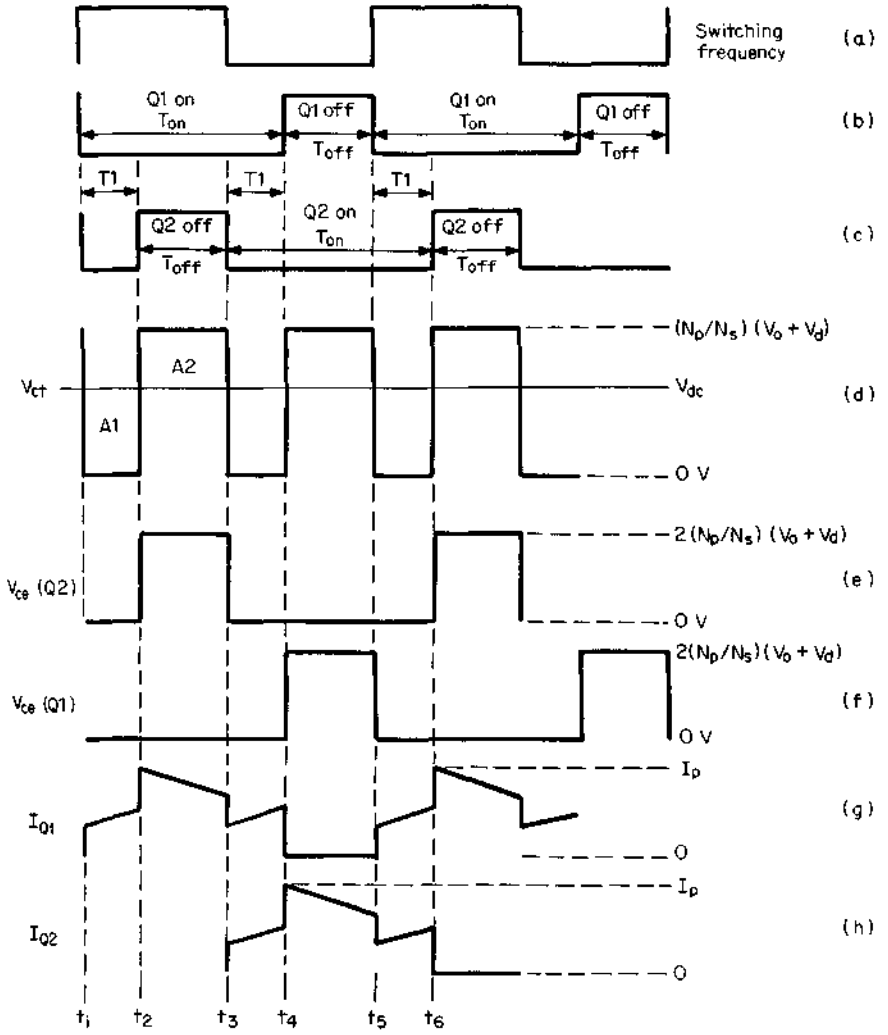


Figure 5.22 Significant voltage-current waveform in flyback-current-fed topology when operating in overlapping mode ($T_{on} > T_{off}$). This mode permits a much larger range of input voltages. Power is delivered to the load only in the interval when one transistor is on and one is off. It is delivered at voltage V_{ct}, which is higher than DC input voltage (boost operation). In the nonoverlapping mode of Fig. 5.20 it is delivered at a voltage lower then the DC input (buck operation).

When $Q2$ turns off at t_2, $Q1$ is still on. Now the $Q1$ half primary can support voltage and V_{ct} begins to rise, as does the $D1$ anode. The $D1$ anode rises until its cathode reaches V_o, the secondary is clamped to $(V_o + V_d)$ and V_{ct} is clamped to $N_1(V_o + V_d)$ as seen in Fig. 5.22*d*.

The turns ratio $N_{LP}/N_{LS}(= N2)$ is chosen so large that when one transistor is on and one off (intervals T_{off}), with the maximum voltage across N_{LP}, the voltage across N_{LS} is insufficient to forward-bias $D3$. This permits V_{ct} to be clamped to $N_1(V_o + V_d)$ during T_{off}. Some of the current stored in L_p during T_1 intervals and current from V_{dc} is delivered via the $T1$ primary to the load at a voltage $N_1(V_o + V_d)$. It will soon be seen that $D3$ will be forward-biased and deliver load power at some sufficiently higher DC input voltage.

The current in $Q1$ at the instant t_2 is equal to the sum of the $Q1$ and $Q2$ currents at the instant just prior to $Q2$ turnoff as the current in L_p cannot change instantly. During (t_2 to t_3), the current ramps downward (Fig. 5.22*g*) as N_1 will be chosen so high that $N_1(V_o + V_d)$ is greater than V_{dc}. With the dot end of L_p positive with respect to its no-dot end, current in it (and in $Q1$) ramps downward.

Now when $Q2$ turns on at t_3, again both transistors are on and their half primaries cannot support voltage, V_{ct} again drops to zero and remains there until t_4 when $Q1$ turns off (Fig. 5.22*d*). From t_4 to t_5, V_{ct} is again clamped to $N_1(V_o + V_d)$.

Now from Fig. 5.22*d*, the relation between output/input voltage and on time can be calculated as follows.

5.6.7.9 Output/input voltage–on-time relation in overlapping mode. Refer to Fig. 5.19*a*. When both transistors were on, the dot end of L_p was negative with respect to the no-dot end. When one transistor turned off (during T_{off}), the polarity across L_p reversed in an attempt to keep current in it constant. The voltage at the dot end of L_p rose until it was clamped to $N_1(V_o + V_d)$ by the clamping action at the secondary.

Now since L_p is assumed to have negligible DC resistance, it cannot support a DC voltage across it. Thus the voltage across it averaged over a full or half period must equal zero. Since the input end of L_p is at V_{dc}, so must be the output end averaged over a half period. Or another way of stating this is that in Fig. 5.22*d*, the area $A1$ must equal area $A2$. Or

$$V_{dc}T_1 = [N_1(V_o + V_d) - V_{dc}]T_{off}$$

$$= N_1V_oT_{off} + N_1V_dT_{off} - V_{dc}T_{off}$$

$$V_oN_1T_{off} = V_{dc}(T_1 + T_{off}) - N_1V_dT_{off}$$

But $T_1 + T_{off} = T/2$; then

$$V_o = \left(\frac{V_{dc}T}{2N_1T_{off}}\right) - V_d \qquad \text{and} \qquad T_{off} = T - T_{on}$$

Then for $D = T_{on}/T$

$$V_o = \left[\frac{V_{dc}}{2N_1(1 - D)}\right] - V_d \tag{5.14a}$$

and from Eq. 5.14a, the duty cycle for any DC input is

$$D = \frac{2N_1(V_o + V_d) - V_{dc}}{2N_1(V_o + V_d)} \tag{5.14b}$$

5.6.7.10 Turns ratio selection in overlapping mode. Equation 5.14a gives the relation between output/input voltage and on time for the overlapping mode for a preselected choice of T_1 turns ratio N_1. As good a choice for N_1 is the value calculated from Eq. 5.14a, which makes $D = 0.5$ at the nominal value of input voltage V_{dcn}. Then for all DC input voltages less than V_{dcn}, there will be overlapping on times ($D > 0.5$) and the output voltage–on-time relation is given by Eq. 5.14a for that calculated value of N_1.

For input voltages greater than V_{dcn}, D is less than 0.5, there is no overlapping on time, and Eq. 5.14a no longer holds. The output voltage–on-time relation will now involve N_2. It did not involve N_2 for D greater than 0.5 in Eq. 5.14a because N_2 had been made so large that during T_{off}, $D3$ was reverse-biased and the peak voltage at V_{ct} involved only N_1 (see Fig. 5.22*d*).

Thus the first choice is the value of N_1 from Eq. 5.14a, which makes $D = 0.5$ for nominal input voltage V_{dcn}. From Eq. 5.14a

$$N_1 = \frac{V_{dcn}}{2(V_o + V_d)(1 - 0.5)} = \frac{V_{dcn}}{V_o + V_d} \tag{5.15}$$

Next $N_2 (= N_{LP}/N_{LS})$ must be selected so that during T_{off} in Fig. 5.22*d*, the maximum voltage across N_{LS} does not forward-bias $D3$. But the maximum N_{LS} voltage occurs at the maximum voltage across N_{LP}; and that latter is a maximum when V_{dc} is a minimum (see Fig. 5.22*d*). Maximum $T2$ secondary voltage is then $[N_1(V_o + V_d) - V_{dc(min)}]/N_2$. Further, since the $D3$ cathode is at V_o, in order for $D3$ not to be forward-biased

$$\frac{N_1(V_o + V_d) - V_{dc(min)}}{N_2} < V_o + V_d$$

or

$$N_2 > \frac{[N_1(V_o + V_d) - V_{dc(min)}]}{V_o + V_d} \tag{5.16a}$$

To avoid problems arising from $T1$ leakage inductance spikes, N_2 is usually selected twice this minimum value.[14] Thus

$$N_2 = \frac{2\,[N_1(V_o + V_d) - V_{dc(min)}]}{V_o + V_d} \tag{5.16b}$$

5.6.7.11 Output/input voltage–on-time relation for overlapping-mode design at high DC input voltages where operation has been forced nonoverlapping. Now with N_1 selected from Eq. 5.15 and N_2 from 5.16b, when V_{dc} is less than its nominal value, the relation between output voltage and on time is given by Eq. 5.14a. At nominal input V_{dcn}, $D = T_{on}/T$ is 0.5; and at DC input voltages greater than V_{dcn}, D is less than 0.5 and there is no overlapping on time. Waveforms for this input voltage range are shown in Fig. 5.23.

For the conditions of Fig. 5.23, whenever $Q1$ or $Q2$ are on, the secondaries are clamped to $(V_o + V_d)$ and the center tap is then clamped to $N_1(V_o + V_d)$ where N_1 is the value calculated from Eq. 5.15. When either transistor turns off, the dot end of L_p rises in an attempt to keep current in it constant. As the dot end of L_p rises, so does the dot end of N_{LS} until it clamps to $V_o + V_d$ via $D3$. This then clamps the dot end of L_p (or V_{ct}) to $V_{dc} + N_2(V_o + V_d)$ as seen in Fig. 5.23*d*.

Again in Fig. 5.23, since the DC voltage averaged over a half cycle must equal zero, the area $A1$ must equal area $A2$. Or

$$[V_{dc} - N_1(V_o + V_d)]T_{on} = [N_2(V_o + V_d)]\,[(T/2) - T_{on}]$$

From this, for $T_{on}/T = D$,

$$V_o = \frac{V_{dc}D - N_2V_d(0.5 - D) - N_1V_dD}{N_2(0.5 - D) + N_1D} \tag{5.17a}$$

In Eq. 5.17a, since V_d, the diode forward drops, are about 1 V, the last two terms in the numerator are small compared to $V_{dc}D$ and can be neglected. The equation can then be rewritten as

$$V_o = \frac{V_{dc}D}{N_2(0.5 - D) + N_1D} \tag{5.17b}$$

And from Eq. 5.17b, the duty cycle at any DC input is

$$D = \frac{0.5V_oN_2}{V_{dc} - V_o(N_1 - N_2)} \tag{5.18}$$

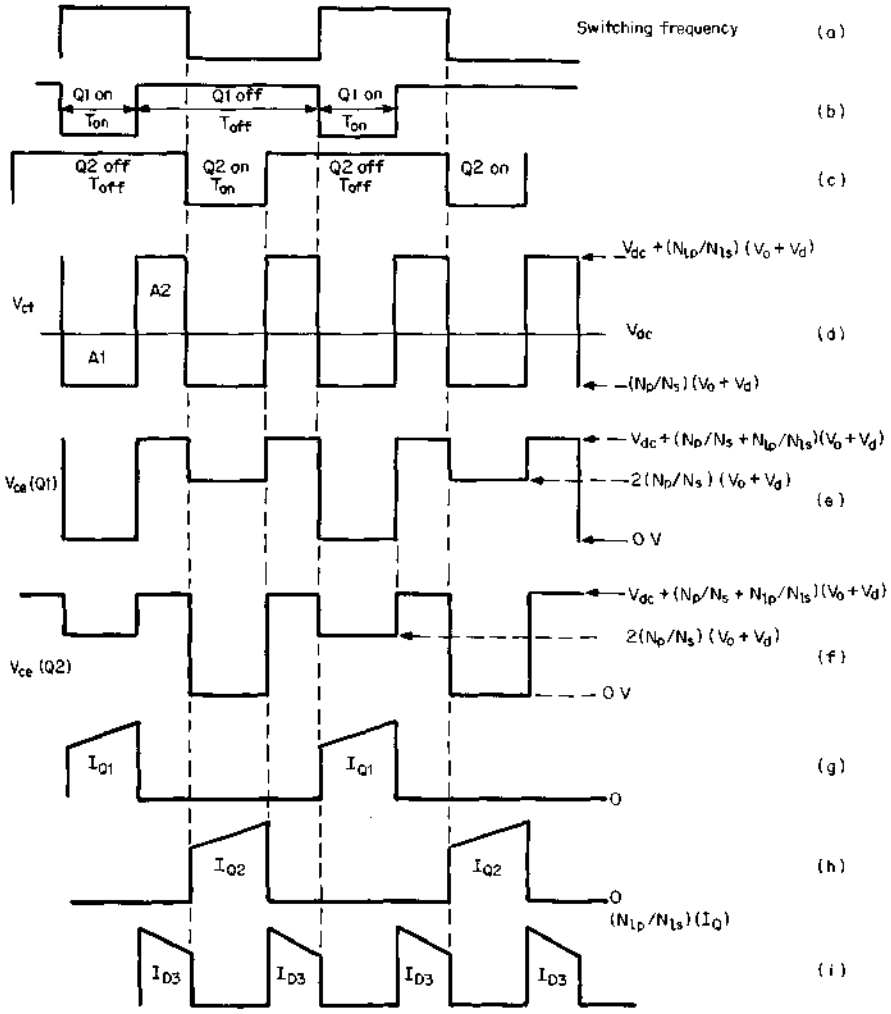


Figure 5.23 Circuit of Fig. 5.20 in overlapping mode but when DC input voltage has risen to a level which has forced it into the nonoverlapping mode. Transition between overlapping and nonoverlapping mode is continuous if turns ratios are chosen correctly.

Thus, when the design is for the overlapping mode, N_1 is calculated from Eq. 5.15 and N_2 from 5.16b. Then at DC input voltages less than nominal V_{dcn}, the feedback loop sets the duty cycle in accordance with Eq. 5.14a to maintain V_o constant. This duty cycle will be greater than 0.5.

When input voltage has risen to V_{dcn}, the duty cycle has decreased to 0.5 to keep V_o at the same value. When DC input voltage has risen above V_{dcn}, the feedback loop sets the duty cycle in accordance with Eq. 5.17b to maintain the output constant. This duty cycle will now be less than 0.5.

The transition from $D > 0.5$ to $D < 0.5$ will be smooth and continuous as V_{dc} rises through V_{dcn}. A much larger range of DC input voltages can now be tolerated than if the design were restricted to the nonoverlapping mode entirely as in Sec. 5.6.7.6.

5.6.7.12 Design example—overlapping mode. It is instructive, using the overlapping mode design, to calculate on times for a range of DC input voltages. This will be done for the design example of Sec. 5.6.7.6 which had restricted operation to the nonoverlap mode entirely. Recall in that design example that V_o was 48 V, nominal input voltage V_{dcn} was 160 V, minimum input voltage $V_{dc(min)}$ was 100 V, switching frequency was 50 kHz and P_o was 2000 W. From Eq. 5.15

$$N_1 = \frac{V_{dcn}}{V_o + V_d} = \frac{160}{48 + 1} = 3.27$$

and from Eq. 5.16b

$$N_2 = \frac{2[(3.2)(49) - 100]}{49}$$

$$= 2\left(3.27 - \frac{100}{49}\right) = 2.46$$

and for $V_{dc} < V_{dcn}$, from Eq. 5.14b

$$D = \frac{[2N_1(V_o + V_d) - V_{dc}]}{2N_1(V_o + V_d)}$$

$$= 1 - \left(\frac{V_{dc}}{2 \times 3.27 \times 49}\right)$$

$$= 1 - \left(\frac{V_{dc}}{320.5}\right) \tag{5.19}$$

and for $V_{dc} > V_{dcn}$, from Eq. 5.17b

$$D = \frac{0.5V_oN_2}{V_{dc} - V_o(N_1 - N_2)}$$

$$= \frac{0.5 \times 48 \times 2.46}{V_{dc} - 48(3.27 - 2.46)}$$

$$= \frac{59}{V_{dc} - 38.9} \tag{5.20}$$

and from Eqs. 5.19 to 5.22, we can construct Table 5.2.

Comparing this (Table 5.2) to Table 5.1, where operation was re-

TABLE 5.2

V_{dc}, V	D	T_{on}, μs	T_{off}, μs	I_p, A (Eq. 5.21)	I_{rms}, A (Eq. 5.22)
50	0.840	16.9	3.1	50.2	24.6
100	0.688	13.8	6.2	25.2	15.1
136	0.576	11.5	8.5	18.3	12.2
160	0.500	10.0	10.0	15.6	11.0
175	0.433	8.67	11.3	13.8	
185	0.404	8.08	11.9	13.1	
200	0.366	7.32	12.7	12.3	

stricted to the nonoverlapping mode, it is seen that the design for overlapping and crossing into nonoverlapping mode permits a larger range of DC input voltages and allows a larger on time at high input voltages. This latter permits the use of bipolar transistors, which would have difficulty operating reliably at the low on times close to the vicinity of their storage times.

5.6.7.13 Voltages, currents, and wire size selection for design example in overlap mode. Transistor currents and transformer rms currents can be calculated from the waveforms of Figs. 5.22 and 5.23. Wire sizes will be selected from the rms currents at the rate of 500 circular mils per rms ampere.

Consider first operation at V_{dc} less than nominal so that there will be overlapping conduction with the waveforms of Fig. 5.22. Assume an efficiency of 80 percent as in the design example of Sec. 5.6.7.6. Input power is then $P_o/0.8 = 2000/0.8 = 2500$ W. Now note that whether V_{dc} is less or greater than nominal, power is supplied to the load through $T1$ at a center tap voltage of $N_1(V_o + V_d)$. For V_{dc} less than nominal (Fig. 5.22*d*), the center tap voltage is boosted up to $N_1(V_o + V_d)$. For supply voltages greater than nominal, center tap voltage is bucked down (Fig. 5.23*d*) to the same value. In this design example, $N_1(V_o + V_d) = 3.27(48 + 1) = 160$ V.

The equivalent flat-topped current pulse I_p into the $T1$ center tap will be calculated. This is closely the current in the center of the ramp in Fig. 5.22*g*. Then power into the $T1$ center tap is

$$P_{in} = 2500 = 160 I_p \frac{2T_{off}}{T} \quad \text{or} \quad I_p = \frac{156}{T_{off}} \tag{5.21}$$

Peak currents for supply voltages less than nominal are calculated from Eq. 5.23 and shown in Table 5.2. If $Q1$, $Q2$ are bipolar transistors, base drives will have to be adequate to saturate them at that peak current level. The transistors will be chosen for a maximum

collector-emitter voltage of $V_{dc(max)} + (N_1 + N_2)(V_o + V_d)$ from Fig. 5.23*e* and 5.23*f* as this is greater than $2N_1(V_o + V_d)$ of Fig. 5.22*e* and 5.22*f*. Allowance should be made for a leakage inductance spike.

Although power and current is delivered to the load only during the two off times per period (Fig. 5.22*g* and 5.22*h*), each half secondary carries I_p during one off time but also $I_p/2$ during two T_1 times per period. Note in Fig. 5.22, $T_1 = (T/2) - T_{off}$. Then the rms current carried by each half secondary is

$$I_{rms} = I_p \left(\frac{T_{off}}{T}\right)^{1/2} + \left(\frac{I_p}{2}\right)\left(\frac{T - 2T_{off}}{T}\right)^{1/2} \tag{5.22}$$

These rms currents are shown in Table 5.2. Wire size for each half secondary will be selected at the rate of 500 circular mils per rms ampere, and it is seen in Table 5.2 that maximum rms current occurs at minimum DC input. RMS currents for supply voltages above V_{dcn} are lower than those below V_{dcn}, so the rms currents of Table 5.2 dictate the wire sizes.

The flyback transformer $T2$ secondary carries the pulses I_{D3} shown in Fig. 5.23. As DC voltage goes up, the transistor on times decrease toward zero and the I_{D3} pulses widen until they reach a full half period each. All the output load current is then fully supplied by flyback action from the $T2$ secondary. Since the $D3$ pulses at the center of the ramp constitute the DC output current, the $T2$ secondary winding should be sized for the worst-case condition, to carry the DC output current at a 100 percent duty cycle.

Finally, wire size for the $T2$ primary must be chosen. Table 5.2 gives the rms currents per half primary at supply voltages of less than nominal. Since the $T2$ primary carries the currents of both half primaries, its rms current is twice that shown in the table.

Examination of Fig. 5.22 shows that the astonishingly high currents at low DC input should be expected, for in Fig. 5.22, as the supply voltage goes lower, the T_{off} times become narrower. And since power is delivered to the load only during the T_{off} times when voltage exists at the $T1$ center tap, the very narrow T_{off} times demand high peak and rms currents to supply the output power.

References

Current Mode:

1. B. Holland, "A New Integrated Circuit for Current Mode Control," *Proceedings Powercon 10,* 1983.
2. W. W. Burns and A. K. Ohri, "Improving Off Line Converter Performance with Current Mode Control," *Proceedings Powercon 10,* 1983.
3. "Current Mode Control of Switching Power Supplies" Unitrode Power Supply Design Seminar Manual SEM 400, 1988 Unitrode Corp., Lexington, MA.

4. T. K. Phelps, "Coping with Current Mode Regulators," *Power Control and Intelligent Motion (PCIM Magazine)*, April 1986.
5. C. W. Deisch, "Simple Switching Control Method Changes Power Converter into a Current Source," 1978 IEEE.
6. R. D. Middlebrook, "Modelling Current Programmed Regulators," *APEC Conference Proceedings*, March 1987.
7. G. Fritz, "UC3842 Provides Low Cost Current Control," Unitrode Corporation Application Note U-100, Unitrode Corp. Lexington, MA.

Current Fed:

8. A. I. Pressman, *Switching and Linear Power Supply, Power Converter Design*, p. 146, Switchtronix Press, Waban, MA, 1977.
9. E. T. Calkin and B. H. Hamilton, "A Conceptually New Approach for Regulated DC to DC Converters Employing Transistor Switches and Pulse Width Control," *IEEE Transactions on Industry Applications*, 1A: 12, July 1986.
10. E. T. Calkin and B. H. Hamilton, "Circuit Techniques for Improving the Switching Loci of Transistor Switches in Switching Regulators," *IEEE Transactions on Industry Applications*, 1A: 12, July 1986.
11. K. Tomaschewski, "Design of a 1.5 kW Multiple Output Current Fed Converter Operating at 100 kHz," *Proceedings Powercon 9*, 1982.
12. B. F. Farber, D. S. Goldin, C. Siegert, and F. Gourash, "A High Power TWT Power Processing System," *PESC Record*, 1974.
13. R. J. Froelich, B. F. Schmidt, and D. L. Shaw, "Design of an 87 Per Cent Efficient HVPS Using Current Mode Control," *Proceedings Powercon 10*, 1983.
14. V. J. Thottuvelil, T. G. Wilson, and H. A. Owen, "Analysis and Design of a Push Pull Current Fed Converter," *IEEE Proceedings*, 1981.
15. J. Lindena, "The Current Fed Inverter—A New Approach and a Comparison with the Voltage Fed Inverter," *Proceedings 20th Annual Power Sources Conference*, pp. 207–210, 1966.
16. P. W. Clarke, "Converter Regulation by Controlled Conduction Overlap," U.S. Patent 3,938,024, issued Feb. 10, 1976.
17. B. Israelson, J. Martin, C. Reeve, and A. Scown, "A 2.5 kV High Reliability TWT Power Supply: Design Techniques for High Efficiency and Low Ripple, *PESC Record*, 1977.
18. J. Biess and D. Cronin, "Power Processing Module for Military Digital Power Sub System," *PESC Record*, 1977.
19. R. Redl and N. Sokal, "Push Pull Current Fed, Multiple Output DC/DC Power Converter with Only One Inductor and with 0 to 100% Switch Duty Ratio," *IEEE Proceedings*, 1980.
20. R. Redl and N. Sokal, "Push Pull, Multiple Output, Wide Input Range DC/DC Converter—Operation at Duty Cycle Ratio Below 50%," *IEEE Proceedings*, 1981.
21. L. G. Meares, "Improved Non-Dissipative Snubber Design for Buck Regulator and Current Fed Inverter," *Proceedings Powercon 9*, 1982.
22. E. Whitcomb, "Designing Non-Dissipative Snubber for Switched Mode Converters," *Proceedings Powercon 6*, 1979.
23. A. H. Weinberg, "A Boost Regulator with a New Energy Transfer Principle," *Proceedings of the Spacecraft Power Conditioning Electronics Seminar*, European Space Research Organization Publication Sp-103, September 1974.

Chapter

6

Miscellaneous Topologies

6.1 SCR Resonant Topologies—Introduction

The silicon controlled rectifier (SCR) has been used in DC/AC inverters and DC/DC power supplies for over 25 years.[1,2] It is available with higher voltage and current ratings at lower cost than bipolar (single or Darlington types) or MOSFET transistors. Because of the higher voltage-current ratings, it is used primarily for supplies of over 1000 W.

It will be recalled that an SCR is a solid-state switch which is easily turned on by a narrow pulse at its input terminal (gate) and latches and stays on after the input is removed. The literature is replete with schemes for turning an SCR off ("commutating" it). Essentially all these schemes involve reducing its on current to zero by diverting it to an alternate path for a certain minimum time t_q. SCR turnoff will be discussed below.

Earlier DC/AC or DC/DC SCR supplies could not operate reliably at switching frequencies over about 8 to 10 kHz. This was because even the fastest inverter-type SCRs available then did not have a reliable high-impedance open circuit at their output terminals until about 10 to 20 μs after they had been turned off and their currents had dropped to zero. Thus they could not be subjected to high voltage until the 10 to 20 μs after their currents had been reduced to zero.

Further, early SCRs could not tolerate a large dV/dt across their output terminals even beyond the 10 to 20 μs after their output currents had fallen to zero. Most were specified at a maximum rate of change of output voltage of 200 V/μs. At a dV/dt faster than that, they would spontaneously turn back on independently of the input control voltage.

Early inverter-type SCRs also could not tolerate a large rate of change of output current dI/dt at the instant of turnon. Most were

specified in the range of 100 to 400 A/μs. At *dI/dt* faster than their specified values, average junction temperatures would rise, local hot spots would develop on the chips, and the SCR would either fail immediately or degrade to the failure point in a short period of time.

With switching frequencies thus limited to 10 kHz, transformers, inductors, and capacitors were still relatively large and made the overall size of a DC/AC or DC/DC converter too large in many applications. Further, switching frequencies under 10,000 kHz are in the middle of the audio range, and the audible noise emitted from such converters made them unacceptable in an office or even factory environment. To be acceptable in such environments, switching frequencies must be above the highest usual audible frequency of about 18 kHz.

About 1977, RCA developed the asymmetrical silicon controlled rectifier (ASCR), which solved most of these problems and made possible DC/AC and DC/DC converters operating up to 40 and possibly 50 kHz.

Conventional SCRs can sustain or block reverse voltages across their output terminals equal to their forward-voltage blocking capability. But in a large number of SCR circuits, reverse voltage at the output terminals is clamped to one or two diode drops or to a maximum of about 2 V, making large reverse-voltage blocking capability unnecessary. Thus by certain solid-state changes on the SCR chip, RCA achieved turnoff times (t_q or the time after SCR forward current has dropped to zero until the full-rated forward voltage can be reapplied) of 4 μs. The price paid for this was a reduction of the reverse voltage blocking capability to 7 V—easily adequate in most inverter circuits.

Thus with t_q times of 4 μs, this RCA device (S7310) made possible inverters at switching frequencies of 40 to 50 kHz in a host of circuit configurations. The S7310 had many other astonishingly useful features. Its *dV/dt* and *dI/dt* ratings were 3000 V/μs (with a 1-V negative bias on its input terminal) and 2000 A/μs.

This compared to 200 V/μs and 400 A/μs for conventional SCRs. Further, the device was available in voltage ratings of 800 V and rms current ratings of 40 A. These voltage, current, and t_q ratings made possible an inverter (or DC/DC power supply) with 4000 W of output power in a two-SCR 40-kHz half-bridge circuit configuration.

In its first few years on the market the S7310 sold for about $5. No other circuit topology using transistors could remotely approach the achievement of 4000 W of output power with two switching devices costing a total of $10. Unfortunately, the S7310 is no longer manufactured by RCA, but equivalent unsymmetrical SCRs are made by other manufacturers. A very similar Marconi ASCR, type ACR25U, has a t_q of 4 μs, blocking voltage of up to 1200 V, and rms current rating of 40 A at a cost of about $17 at this writing.

A final significant feature of an SCR approach to high-power inverters is that the SCR does not suffer from the most frequent failure mode of transistors—second breakdown.

6.2 SCR Basics

The SCR symbol is shown in Fig. 6.1. Its input terminal (terminal 1) is designated the *gate,* terminal 2 is the *anode,* and 3 is the *cathode.* When it is on, current flows down from the anode to the cathode. When it is off, the maximum voltage it can sustain or block from anode to cathode is designated V_{DRM}. ASCR types are available with V_{DRM} voltage ratings ranging from 400 to 1200 V.

Once turned on, the anode-to-cathode current is determined by the supply voltage and load impedance from anode to supply source. Anode to cathode voltage versus anode current characteristics are given in the data sheets for the specific device. For the Marconi ACR25U, a 40-A rms device, anode-to-cathode voltage at 100-A anode current is typically 2.2 V (Fig. 6.2).

The device is turned on by a gate-to-cathode current pulse which is not too well specified in amplitude and duration in the data sheets. Figure 6.3 shows anode current delay and rise time (to 100 A) as a function of gate-trigger current. Gate current pulse width also determines to some extent the current rise time, but this is seldom given. Typically, for the ACR25U, gate current should be in the range of 90 to 200 mA for an anode current of 100 A.

Gate pulse duration should be greater than 400 ns for 100 A of anode current. The gate-to-cathode voltage during the duration of the gate current pulse is shown in Fig. 6.4 and is in the range of 0.9 to 3 V for a large range of gate currents. Once turned on, the anode will latch on and stay conducting after the gate turnon pulse is gone. The anode "on" potential ranges from about 1.2 to 2.2 V over an anode current range of 20 to 100 A as seen in Fig. 6.2.

Anode current rise time (shown in Fig. 6.3) is not as important for an SCR as is anode-to-cathode voltage fall time. This is obvious from Fig. 6.5, which shows that even with a 500-mA gate current pulse, with a half-sinusoid 8-μs anode current pulse, 125 A in amplitude, the

Figure 6.1 Silicon controlled rectifier symbol.

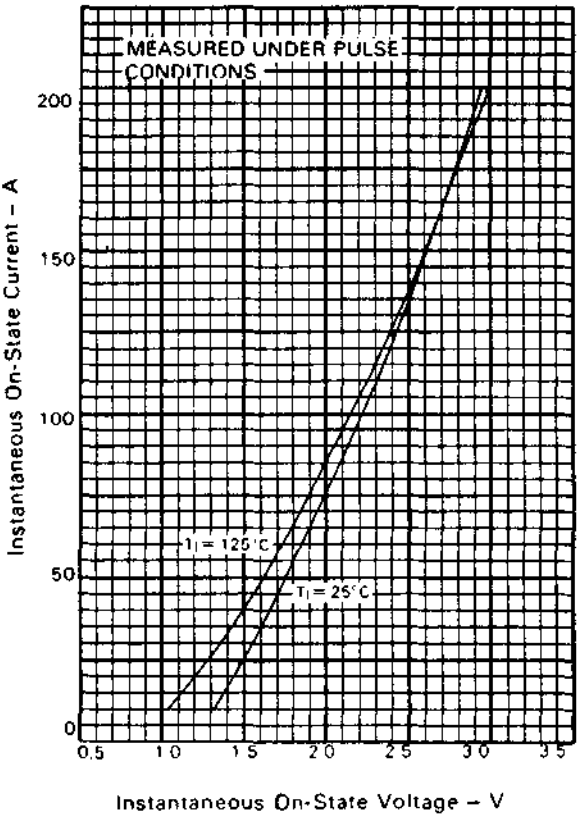


Figure 6.2 Anode current versus voltage.

anode-to-cathode voltage has fallen to only 5 V in 2.5 μs. This is still twice the anode-to-cathode quiescent voltage at that current as shown in Fig. 6.2.

The reason for this long anode voltage fall time is that it takes a long time for the anode current carriers to spread uniformly throughout the chip area. Initially, the current carriers are concentrated in only a small fraction of the chip area and instantaneous anode-to-cathode resistance is high, causing a high instantaneous on voltage. After a time, the current carriers spread uniformly throughout the chip and the on voltage drops down to the quiescent level given in Fig. 6.2.

Thus most of the dissipation in an SCR occurs during the turnon time. This dissipation is the integral $\int I_a V_a \, dt$. In most cases with SCR circuits, currents will be in the shape of half sinusoids rather than square waves, which is helpful. As seen in Fig. 6.5, if anode current pulses were square waves, the front end of the current pulse would

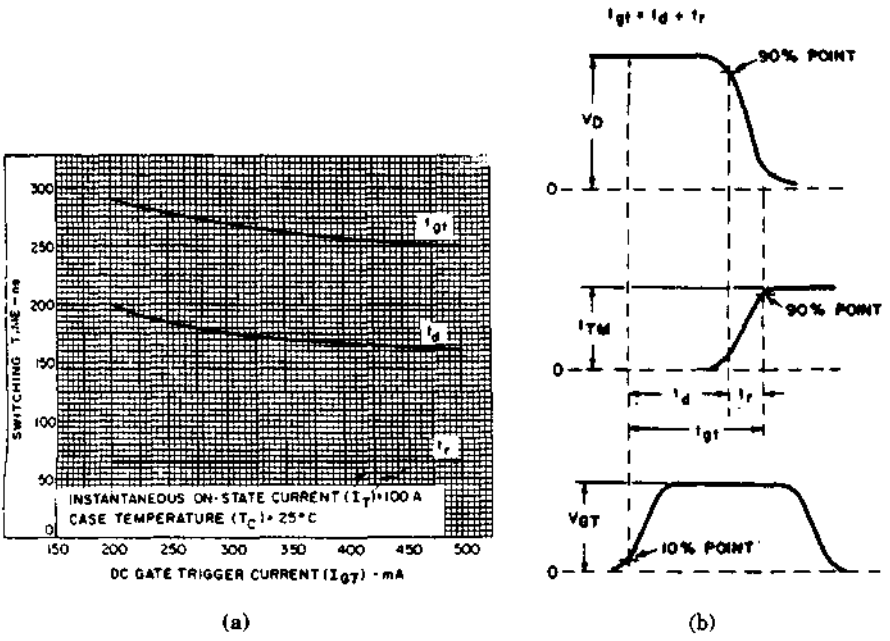


Figure 6.3 (*a*) Switching times. Typical switching time t_{gt}, t_d, t_r versus gate-trigger current. (*b*) Relationship between off-state voltage, on-state current, and gate-trigger voltage showing reference points for definition of turnon time t_{gt}. (*Note:* Figure 6.3*a*, 6.3*b* illustrate the original RCA S7310—a type of SCR very similar to the Marconi type ACR25U.)

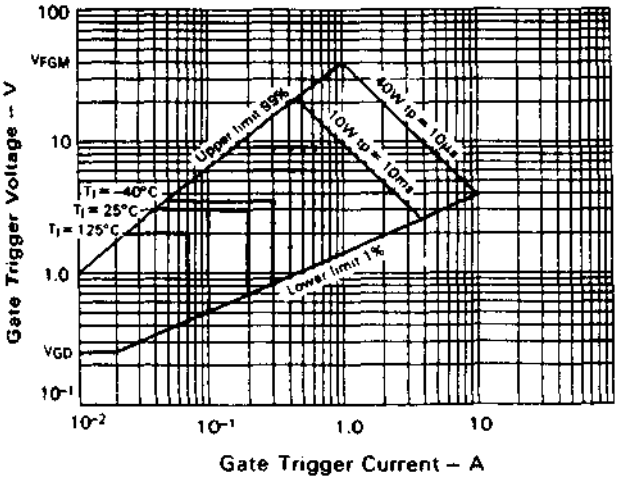


Figure 6.4 Gate voltage versus current.

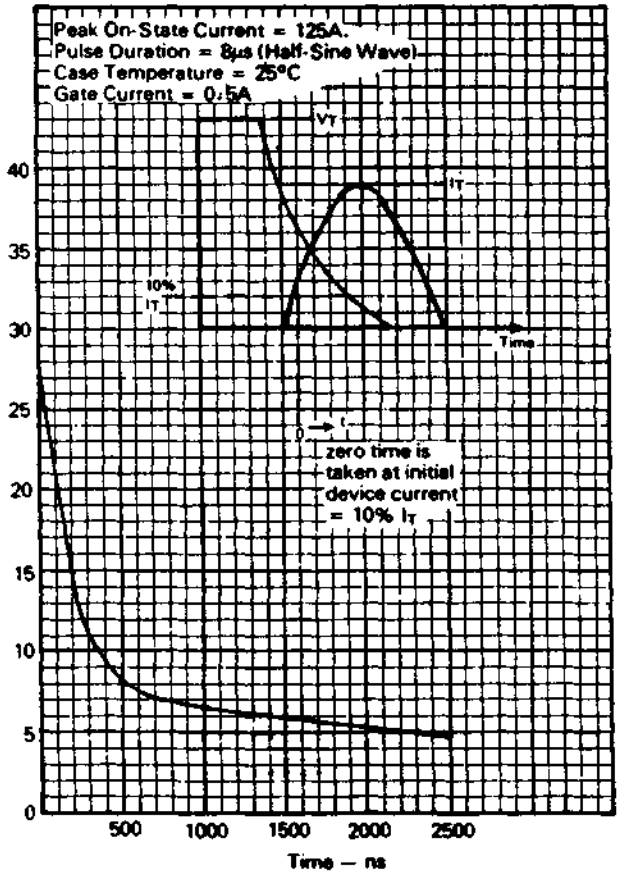


Figure 6.5 Anode-to-cathode voltage fall time. Marconi ASCR type ACR25U characteristics.

flow at an anode voltage in the vicinity of 25 V and dissipation would be high. Figure 6.5 also shows that if anode current pulses are half sinusoid, their base width should be longer than 2.5 μs to avoid an "on" anode potential greater than 5 V throughout the entire half sinusoid.

Figures. 6.6 and 6.7 show maximum dV/dt and t_q for the Marconi ACR25U.

6.3 SCR Turnoff by Resonant Sinusoidal Anode Currents—Single-Ended Resonant Inverter Topology

It was pointed out above that an SCR is easily turned on with a narrow pulse but stays latched on after the pulse has gone. To turn it off,

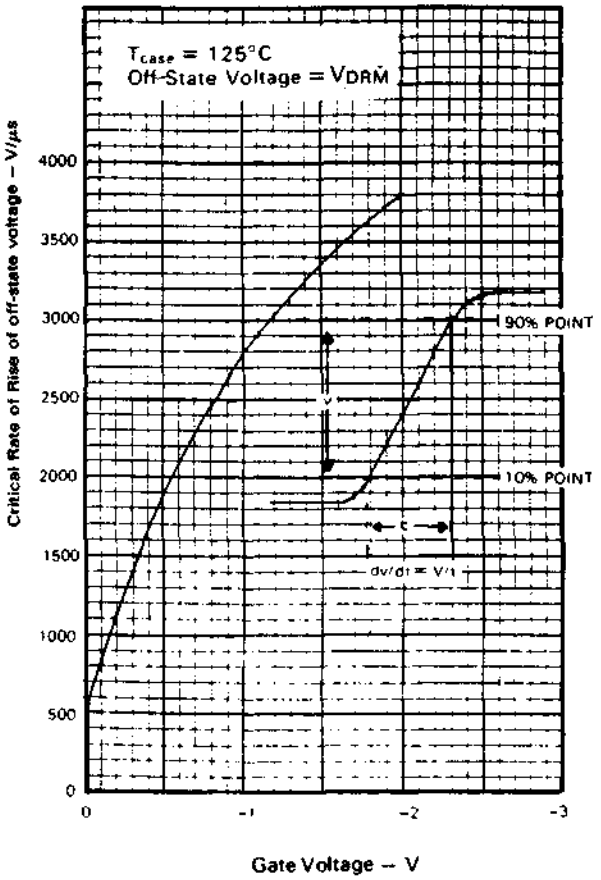


Figure 6.6 Marconi ASCR type ACR25U characteristics. Minimum linear critical rate or rise of off-state voltage versus gate voltage.

anode current must be reduced to zero for a time equal to at least the specified t_q time of the device. Further, after the t_q time, the reapplied anode voltage rise-time rate must be less than the specified dV/dt rating of the SCR.

All this and other significant advantages are easily achieved by forcing the SCR anode current to be sinusoidal in shape. The basic scheme and its advantages are most easily described with a typical single-ended SCR resonant converter such as that shown in Fig. 6.8.[3–8]

The SCR, an inductor L, and a capacitor C are arranged in series. Before the SCR is fired, capacitor C is charged to some positive voltage through the larger constant-current inductor L_c. Now when the SCR is triggered on with a narrow gate pulse, the equivalent circuit is that of

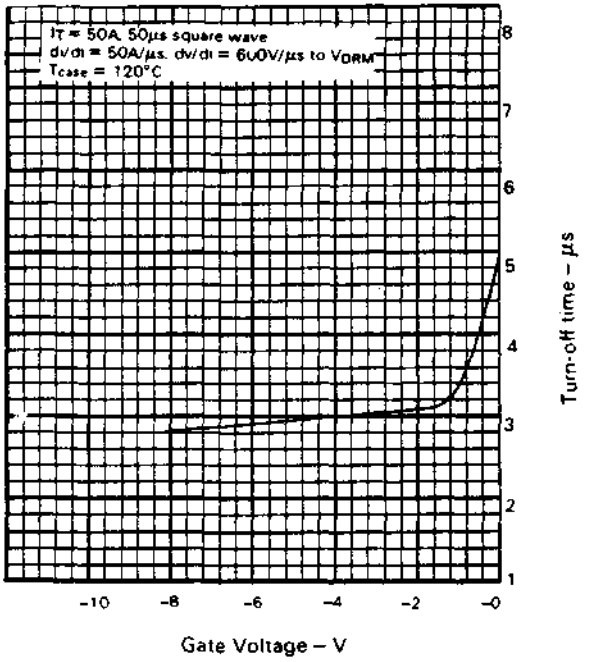


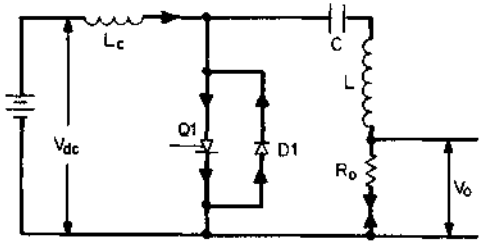
Figure 6.7 Marconi ASCR type ACR25U characteristics. Typical circuit commutated turnoff time versus gate voltage at turnoff.

a switch closure applying a step waveform to a series resonant LC circuit. Current in the circuit is shocked into a resonant "ring" whose period $t_r = 2\pi\sqrt{LC}$

Current increases in the anode sinusoidally goes through its first negative peak and decreases sinusoidally to zero at the end of a half period ($= \pi\sqrt{LC}$). Now with the antiparallel diode $D1$ across the SCR, as the sine wave of current in the SCR reaches zero at t_1, it reverses direction and flows sinusoidally for the next half cycle through $D1$. During this half cycle of diode conduction time T_d, the SCR is clamped with a reverse voltage of about 1 V by $D1$. This maximum reverse voltage is safely below the 7- to 10-V reverse-voltage specification of the asymmetrical SCR.

Now if T_d is greater than the specified t_q time of the device, at the end of T_d, the SCR has safely extinguished itself without the need for any external "commutation" circuitry and forward voltage may be safely reapplied. The half sinusoids of current through the SCR and the antiparallel diode provide power to the load resistor R_c.

After t_2, when the $D1$ current has fallen back to zero, both $Q1$ and $D1$ are safely off and the constant current from L_c commences charg-



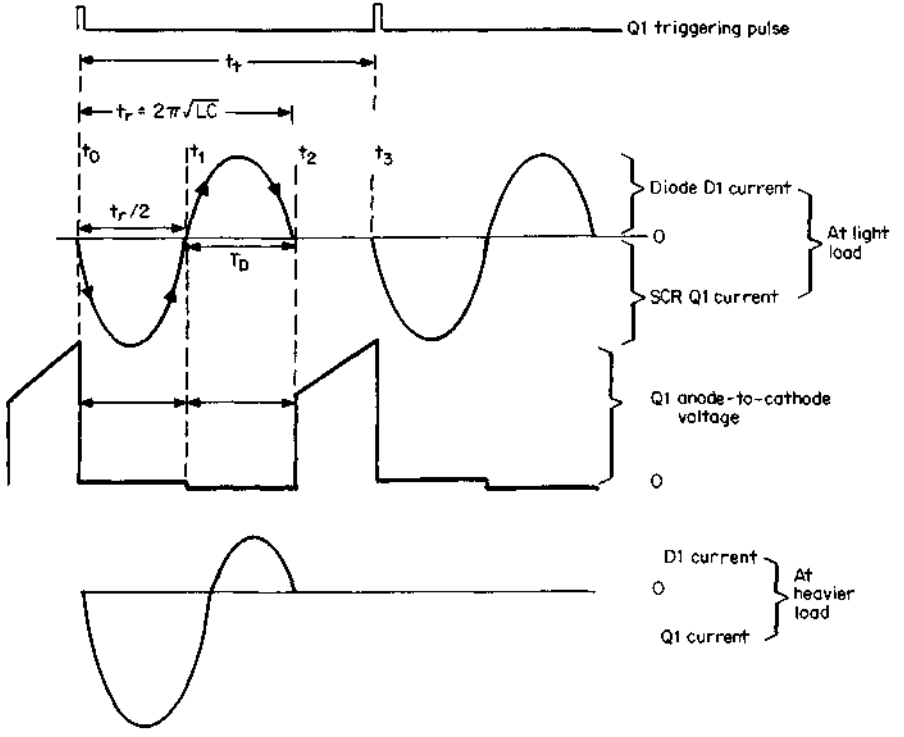


Figure 6.8 Single-ended SCR resonant converter. Inductor L_c charges C to a voltage higher than V_{dc}. When $Q1$ is fired, a sinusoidal flows through $Q1$, delivering power to R_o. At t_1, the sinusoidal current reverses direction and flows through $D1$ again, delivering power to R_o. If T_d is greater than t_q time of $Q1$, the SCR self-extinguishes. During $(t_t - t_r)$ L_c recharges C and the cycle repeats.

ing C back up with its left end positive. During the interval t_3 to t_2, C is charged up to such a DC level that some fraction of its stored energy ($CV^2/2$) is equal to the energy delivered to the load the next time $Q1$ is triggered. After a time t_t (the triggering period), $Q1$ can be triggered on again and the cycle repeats.

As the load is increased (resistor R_o is decreased), the amplitude of the first half cycle increases and its duration increases somewhat. This decreases the duration of the second half cycle T_d, and care must be taken that the load is not increased to the point where T_d is shorter than t_q, the SCR turnoff time, or the SCR will not turn off successfully.

By choosing the triggering period t_t in the range of 1.5 to 2 times the resonant period t_r at minimum line input and maximum load, the off time $t_t - t_r$, shown in Fig. 6.8 is not too large and consequently the output across R_o is a fairly distortion-free AC sine wave over a large range of load resistors. Capacitance across R_o can reduce the distortion due to the time gap $t_t - t_r$. The circuit can thus be used as a DC/AC converter. As the supply voltage V_{dc} is increased, output power would be increased as the sine-wave peak amplitudes would increase. To maintain constant output power or output voltage, a feedback loop senses output voltage and decreases the triggering frequency (increases t_t) as V_{dc} or R_o goes up. This maintains a roughly constant peak AC voltage at the output, although distortion increases as the time (gap $t_t - t_r$) increases.

The circuit is more useful as DC/DC converter with regulated and isolated output as in Fig. 6.9., where R_o is replaced by the primary of a power transformer with rectifying diodes and a capacitor filter at the secondary. Regulation against line and load changes is done by varying the switching or triggering frequency f_t.

As the line voltage or load resistance increases, the sine-wave peaks in Fig. 6.8 increase. But their base widths remain roughly constant and equal to $\pi\sqrt{LC}$. To regulate against line and load changes, the

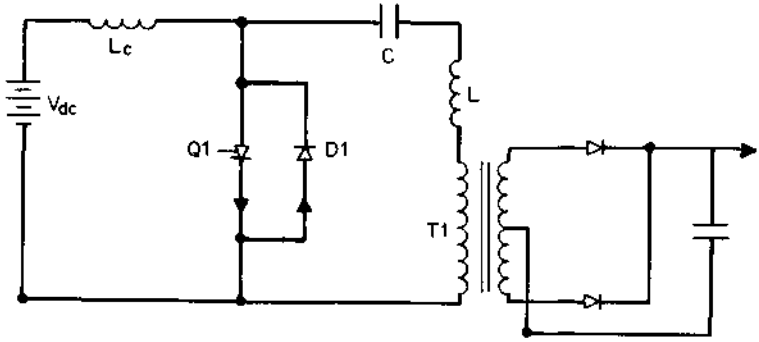


Figure 6.9 A transformer-coupled, series-loaded, single-ended SCR resonant converter.

rectified DC output is sensed with a voltage error amplifier which alters the switching frequency to maintain constant output. If output voltage rises because of an increase in line voltage or a decrease in DC load current, the switching frequency is decreased so as to take less power per unit time from the input. Similarly, a small decrease in output is corrected by an increase in switching frequency.

This method of regulating the output voltage—by variation of the switching frequency—is common to most resonant power supplies as the output current or voltage pulses are constant in width. In nonresonant topologies, regulation is done by operating at a constant switching frequency and varying pulse width to control the DC output voltage.

In such constant-frequency schemes, the power supply switching frequency is generally synchronized to the horizontal line rate in an associated display terminal or to the system's clock rate. This makes it easier to tolerate any RFI noise pickup on the display screen and lessens the possibility of computer logic errors due to noise pickup.

Since this advantage is lost in resonant-topology, variable-frequency, voltage-regulating schemes, these schemes are not acceptable in many applications. Yet it can be argued that although RFI noise pickup, if it exists, is more troublesome with variable-frequency switching supplies, it is less likely to exist. For the sinusoidal currents in resonant supplies have much slower di/dt than do the square-wave currents of fixed frequency, adjustable pulse width supplies, and hence emit less RFI.

The SCR supply was originally made resonant to ensure turnoff of the SCR at the zero crossing of the current sinusoid. This added a further significant advantage. It was seen in Sec. 1.2.4 that with square waves of current, most of the losses in the switching device occur at turnoff as a result of the overlap of falling current in it and rising voltage across it. But with sinusoidal currents, turnoff occurs at zero voltage across the device and such losses are almost nonexistent. Yet the turnon losses due to the relatively slow voltage fall time (Sec. 6.1 text and Fig. 6.5) can be high. If the sine-wave base width $\pi\sqrt{LC}$ is not less than about 8 μs, those losses are also not too great (Fig. 6.5).

This DC/DC converter, used with a single 800-V, 45-A rms SCR can generate 1-kW of output power.[4] An inductor is not needed in the secondary output. For the SCR and diode currents shown in Fig. 6.8 are constant-current pulses whose magnitude is closely equal to the voltage applied across the series LC elements divided by $\sqrt{L/C}$. These constant currents reflect into the secondary as constant currents. Those constant-current pulses flowing into the output resistor produce constant-voltage pulses of similar waveshape. The voltage averaging to obtain a constant ripple-free output voltage is done by the filter ca-

pacitor alone. Thus, not requiring an output inductor, the circuit can be used as a high output voltage supply.

A design example showing quantitative design decisions will be presented after the following discussion of more widely used two SCR resonant-bridge-type DC/DC inverters.

6.4 SCR Resonant Bridge Topologies—Introduction

The resonant half bridge (Fig. 6.10) and full bridge (Fig. 6.11) are the most useful SCR circuits. The half bridge, with two 800-V 45-A rms SCRs (Marconi ACR25UO8LG) can deliver up to 4 kW of AC or rectified DC power from a rectified 220-V AC line. A 1200-V version of that device can generate up to 8 kW in a full-bridge circuit. Full-bridge operation is much the same as the half-bridge except, of course, that SCR voltage stresses are twice as great and current levels are

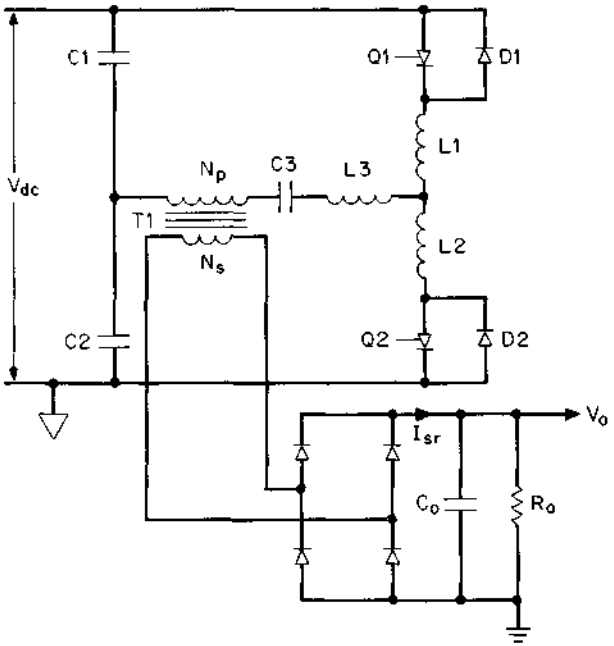


Figure 6.10 SCR resonant half-bridge series loaded. Transistors $Q1$ and $Q2$ are fired on alternate half cycles. Capacitor $C3$ resonates with $L3$ and $L1$ when $Q1$ is on and with $L3$ and $L2$ when $Q2$ is on. At the firing of either SCR, its current goes through a half sinusoid, reverses direction, and flows through the associated antiparallel diode. If the duration of the diode current is greater than the t_q time of the SCR, the SCR self-extinguishes.

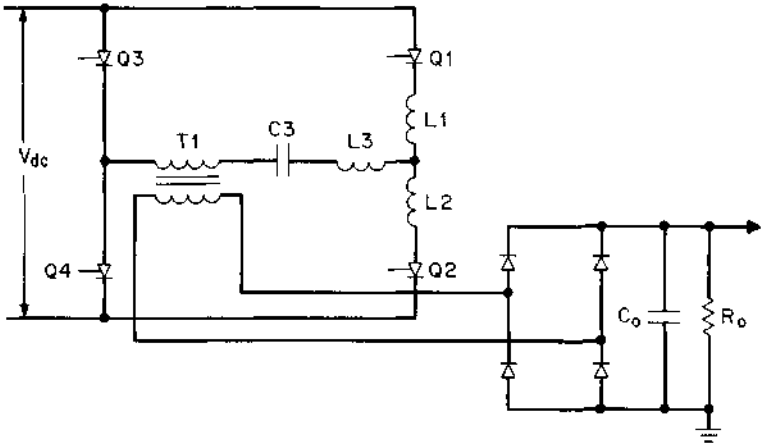


Figure 6.11 SCR resonant full bridge. This can deliver twice the output power of the half bridge.

half those of the half bridge for equal output power. Hence only the half bridge will be discussed here in detail.

The half bridge can be operated *series-loaded* as in Fig. 6.10 with the secondary load reflected via transformer $T1$ in series with a series resonant ($C3$ resonating with the series combination of $L3$ and $L1$ when $Q1$ is on or with the series combination of $L3$ and $L2$ when $Q2$ is on).

In the series-loaded circuit, the secondary load, when reflected into the primary, must not not appear as too high an impedance or the resonant circuit Q will be low and the "on" SCR may not be safely commutated off by the above described resonant current reversal. No output inductor is required in the series-loaded configuration, and hence it can be used for either high or low output DC voltages. The series-loaded circuit can safely survive an output short circuit as the normal load impedance in series with the resonating LC is already small compared to the impedance of the LC elements. But as discussed below, there is a problem when the output load is open-circuited.

Alternatively, the half bridge can be operated *shunt-loaded* as in Fig. 6.12. Here, the output load reflected into the $T1$ primary is fed in shunt with the voltage across the resonating capacitor $C3$. In this case, the output load reflected across $C3$ must not be too low or the resonant circuit Q will be so low as to prevent resonant turnoff of the SCR. Thus, this configuration can easily tolerate an open-loaded but not a short-circuited secondary.

The series-loaded configuration is analyzed as a current source driving the $T1$ primary whereas the shunt-loaded circuit is better analyzed as a voltage source driving the primary. The shunt-loaded cir-

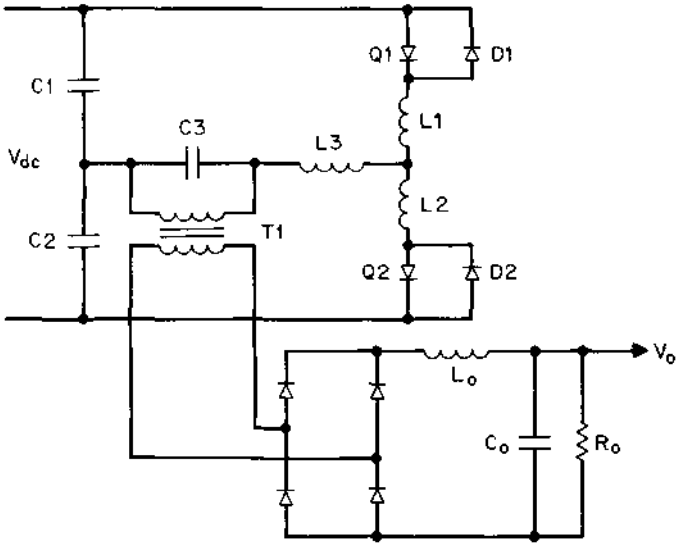


Figure 6.12 A shunt loaded SCR resonant half bridge. Power is taken in shunt across the resonant capacitor. The load reflected into the primary from the secondary must be high so as not to lower the resonant circuit Q. Too low a Q can prevent successful turnoff of the SCR.

cuit does require secondary output inductors when DC output is required, although they may be omitted if large output voltage ripple can be tolerated. For a DC/DC converter, the series-loaded circuit is a better choice.

6.4.1 SCR half-bridge resonant converter, series-loaded—basic operation[9,10]

The series-loaded SCR resonant half-bridge circuit is shown in Fig. 6.10 and its significant waveforms in Fig. 6.13. In Fig. 6.10, when $Q1$ is triggered on, the equivalent circuit is that of a voltage step of magnitude $V_{dc}/2$ applied to a series combination of $L3 + L1$ resonating with $C3$. That resonant circuit is series-loaded with the $T1$ secondary resistance reflected into the primary and that reflected primary resistance is shunted by the $T1$ magnetizing inductance.

If the Q of that equivalent circuit is not too low, current in it is shocked into a sinusoidal "ring" as shown in Fig. 6.13*a*. During the first half cycle, the half sinusoid of current flows through $Q1$. At the end of that half cycle at t_1, the current reverses and continues flowing through the antiparallel diode $D1$. From t_1 to t_2, current in the SCR is zero, and if that time is greater than the specified t_q time of the device, the SCR extinguishes itself and can safely sustain forward voltage again.

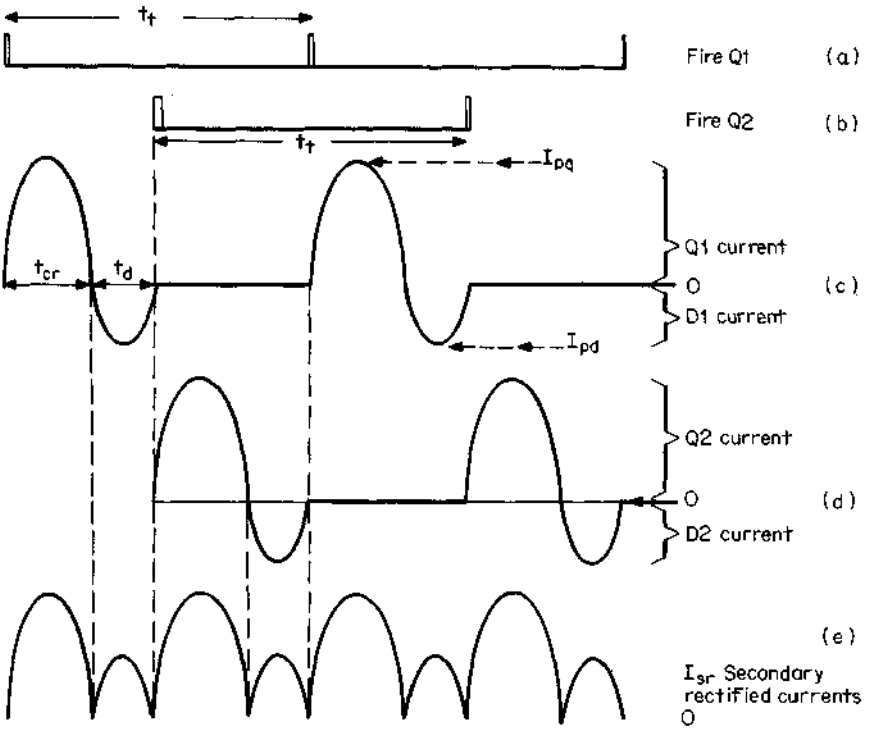


Figure 6.13 Significant currents for series loaded SCR half bridge of Fig. 6.10 for condition of minimum DC input voltage, maximum load output. As line voltage or load resistance increases, the feedback loop will decrease, triggering frequency so as to space the $Q1$, $A2$ sine waves further apart and maintain the average output current and voltage constant.

The period of the resonant ring is $t_r = 2\pi\sqrt{(L_1 + L_3)C_3}$. At light loads, the SCR and diode conduction times (t_{cr} and t_d, respectively) are equal. As the secondary load increases (R_o decreases), t_{cr} and I_{pq} increase and t_d, I_{pd} decrease. Load power must not be increased beyond the point where t_d is less than the maximum specified value of t_q, or the SCR will not turn off successfully.

Now at the next half cycle of the triggering period t_f with $Q1$ safely off, $Q2$ is fired. The resonant period is now $2\pi\sqrt{(L_2 + L_3)C_3}$. Current waveforms of $Q2$, $D2$ are similar to those of $Q1$, $D1$ and are shown in Fig. 6.13*b*. The $Q1$, $D1$ and $Q2$, $D2$ current waveforms, shown in Fig. 6.13c, are multiplied in the $T1$ secondary by the turns ratio N_p/N_s, rectified and summed by the output diodes.

The DC output voltage is the average of the Fig. 6.13c current waveforms multiplied by the secondary load resistor R_o. Averaging of the output current waveforms to yield a constant and ripple-free DC

output voltage is done by the output filter capacitor without requiring an output inductor. After t_t the entire cycle repeats.

The timing relations shown in Fig. 6.13 will hold at minimum DC input voltage and minimum output load resistor (maximum output power). The values of resonating inductance and capacitance will be chosen so that the current amplitudes and the spacings shown in Fig. 6.13c will yield the correct average output current at the desired output voltage at minimum line and maximum load. Calculations to achieve this are shown below. As line and load change, a feedback-loop-sensing output voltage will adjust SCR triggering frequency f_t to keep output voltage constant.

Since the peak currents in Fig. 6.13c are proportional to DC input voltage, as that goes up, the feedback loop will have to decrease f_t to maintain constant secondary average current and hence constant output voltage. Further, since at a fixed DC input voltage and fixed values of L and C, peak currents (Fig. 6.13c, d) are constant, as R_o goes up, the feedback loop will also have to decrease f_t to maintain constant average output voltage. Waveshapes at DC inputs higher than minimum or output loads higher than minimum are shown in Fig. 6.14.

6.4.2 Design calculations—SCR half-bridge resonant converter, series-loaded[9,10]

The circuit configuration is shown in Fig. 6.10 and its significant waveforms, in Figs. 6.13 and 6.14. The discussion herein is based on a paper by D. Chambers.[9]

The first decision to be made is the resonant period or frequency. Assume the use of the Marconi ASCR type ACR25. Figure 6.7 shows its typical turnoff time t_q as 5 μs from an on current of 50 A with a gate bias of 0 V. Assume the worst case is 20 percent higher or 6 μs. Also assume desired operation at minimum DC input and maximum output power at the edge of the continuous mode (as in Fig. 6.13 with no gap between the zero crossing of the diode current and the start of the

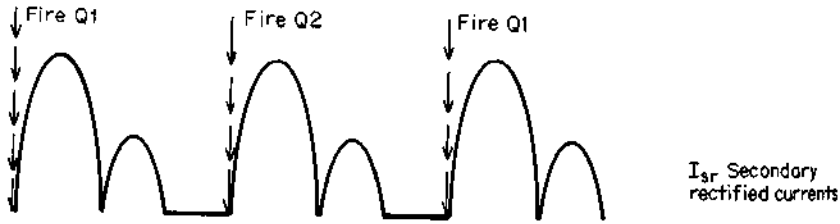


Figure 6.14 Bridge output currents I_{sr} at higher input voltage. Peak currents are higher, forcing feedback loop to decrease triggering frequency to maintain constant output voltage.

opposite SCR current). Then the absolute minimum resonant period as can be seen in Fig. 6.13*c* would be 12 μs or a resonant frequency of 83 kHz.

However, it was noted in Sec. 6.3, that at high output power t_d the diode conduction time shortens to a not easily calculable value. If it shortened to a value less than t_q, the SCR might not turn off successfully. Thus there should be more margin in the t_d time.

Further, Fig. 6.5 shows that the anode-to-cathode on voltage does not fall very quickly to its quiescent value of 2 to 3 V. To keep the high anode-to-cathode voltage time a small fraction of the SCR conduction time t_{cr}, the resonant half period should be increased to at least four times 2.5 μs. This, as can be seen in Fig. 6.5, puts the anode-to-cathode voltage at the peak of the sinusoidal anode current at about 3 V and is a reasonable compromise. Thus the resonant period is chosen as 20 μs (f_r = 50 kHz). Or

$$T_r = 2\pi\sqrt{(L_3+L_1)(C_3)} = 20 \times 10^{-6} \tag{6.1}$$

Next the peak voltage on the $T1$ primary (Fig. 6.10) must be determined. Following the suggestion in the Chambers paper,[9] this will be chosen as 60 percent of the minimum voltage across one of the bridge capacitors or

$$\text{Minimum primary voltage} = V_{p(\min)} = \frac{0.6\,V_{dc(\min)}}{2} \tag{6.2}$$

Assuming a bridge output rectifier with a 1-V drop across each rectifier diode, this fixes the $T1$ turns ratio at

$$\frac{N_p}{N_s} = \frac{0.6V_{dc(\min)}}{2(V_o + 2)} \tag{6.3}$$

At minimum line input, maximum output current, the secondary currents are as shown in Fig. 6.13*e* with no time gaps between the termination of current in one antiparallel diode and turnon of current in the opposite SCR. Assume, to a close approximation, that even at maximum DC current, SCR and diode currents have equal widths of a half period. Also assume—as in the Chambers paper[9]—that the diode peak current is one-fourth that of the SCR current. Then the average of these Fig. 6.13*e* SCR plus diode currents is

$$I_{\text{secondary average}} = I_{o(dc)} = \frac{2I_{ps}}{\pi}\frac{T}{2T} + \frac{2I_{ps}}{4\pi}\frac{T}{2T}$$

$$= \frac{1.25I_{ps}}{\pi}$$

where I_{ps} is the peak primary SCR current after reflection into the secondary. Then

$$I_{ps} = 0.8\pi I_{o(dc)} \tag{6.4}$$

$$I_{pp} = I_{ps}\frac{N_s}{N_p} = 0.8\pi I_{o(dc)}\frac{N_s}{N_p} \tag{6.5}$$

where I_{pp} is the peak primary SCR current.

In Fig. 6.10, V_{ap}, the voltage applied to the series resonant elements when, say, $Q1$ turns on, is the voltage across the bridge capacitor $C1$ plus the transformer voltage peak of $0.6V_{dc(min)}/2$. Or

$$V_{ap} = \frac{V_{dc(min)}}{2} + \frac{0.6V_{dc(min)}}{2} = 0.8V_{dc(min)} \tag{6.6}$$

It can be shown that, to a close approximation, when a step voltage V_{ap} is applied to a series LC circuit, the peak amplitude of the first resonant current pulse is

$$I_{pp} = \frac{V_{ap}}{\sqrt{L/C}} \tag{6.7}$$

Then in this case, where $L = (L_1 + L_3)$, $C = C_3$:

$$\sqrt{(L_1 + L_3)/C_3} = \frac{V_{ap}}{I_{pp}}$$

or

$$\sqrt{(L_1 + L_3)/C_3} = \frac{0.8V_{dc(min)}}{0.8\pi I_{o[dc(min)]}(N_s/N_p)}$$

$$= \frac{V_{dc(min)}(N_p/N_s)}{\pi I_{o(dc)}} \tag{6.8}$$

Thus from Eq. 6.1, which gives the product of $(L_3 + L_1)$, (C_3), and Eq. 6.8, which gives their ratio, values of the resonating elements—$(L_3 + L_1)$ or $(L_3 + L_2)$ and C_3—are fixed for specified values of $V_{dc(min)}$ and the maximum output current $I_{o(dc)}$. The transformer turns ratio is fixed from Eq. 6.3.

The ratio of L_3/L_1 is chosen to minimize off-voltage stress on the SCRs. A smaller ratio is in the direction of lesser off stress and lesser dV/dt stress. The precise ratio is best determined empirically. Inductance L_3 constitutes the transformer primary leakage inductance plus some external inductance. It is best not to rely too much on the leak-

age inductance alone as it usually has a wide tolerance, and that would result in large tolerances in the resonant period.

6.4.3 Design example—SCR half-bridge resonant converter, series-loaded

The preceding relations will now be used in a design example. Assume the circuit of Fig. 6.10 with the following specifications:

Output power	2000 W
Output voltage	48 V
Output current $I_{o(dc)}$	41.7 A
Input DC voltage, nominal	310 V
Input DC voltage, maximum	370 V
Input DC voltage, minimum	270 V

Then from Eq. 6.3 $N_p/N_s = 0.6 \times 270/2(48 + 2) = 1.62$ and from Eq. 6.1

$$2\pi\sqrt{(L_3 + L_1)C_3} = 20 \times 10^{-6}$$

or

$$\sqrt{(L_3 + L_1)C_3} = 3.18 \times 10^{-6} \tag{6.1a}$$

From Eq. 6.8

$$\sqrt{\frac{(L_3 + L_1)}{C_3}} = \frac{V_{dc(min)}(N_p/N_s)}{\pi I_{o(dc)}}$$

$$= \frac{270 \times 1.62}{\pi \times 41.7}$$

$$= 3.34 \tag{6.8a}$$

From Eqs. 6.1a and 6.8a $C_3 = 0.95\ \mu F$ and $L_3 + L_1 = 10.6\ \mu H$ and from Eqs. 6.6 and 6.7

$$I_{pp} = \frac{V_{ap}}{\sqrt{(L_3 + L_1)/C_3}}$$

$$= \frac{0.8 \times 270}{3.34} = 64.7 \text{ A}$$

and since the maximum duty cycle of this peak SCR current is $t_r/2t_{f(min)} = 0.25$ (Fig. 6.13*a* and 6.13*c*), the SCR rms current is $I_{rms(SCR)} = 64.7 \times \sqrt{0.25}/\sqrt{2} = 22.9$ A. This is easily within 40 peak rms capability of the Marconi ACR25U. Also, as assumed above, the antiparallel diode peak current is one-fourth of the SCR current or

64.7/4 = 16.2 A. With a 1.62 turns ratio, the peak rectifier diode currents will be 104.8 and 26.2 A corresponding to the peak SCR and antiparallel diode currents. With such high rectifier diode currents, a full-wave rectifier with one series diode rather than a bridge with two diodes might be preferable.

6.4.4 SCR half-bridge resonant converter, shunt-loaded[6,12]

The ancestor of most of the practical SCR resonant bridge power supplies is Neville Mapham's classical paper.[6] That paper described an SCR shunt-loaded resonant half bridge used as a DC to AC inverter. The circuit was essentially that of Fig. 6.12 with a resistive load at the transformer secondary instead of the rectifier plus output filter.

Mapham's paper made no attempt to regulate the output but showed that over a 10/1 range of output load currents, with no feedback, the output AC voltage was constant to within 1 percent and generated a relatively distortion-free sine wave. A detailed and elegant computer analysis showed that the peak sine-wave output could be regulated against line input changes by changing the SCR triggering frequency.

The computer analysis is presented in terms of normalized relations ($R_o/\sqrt{L/C}, I_{scr}\sqrt{L/C}/E, I_{diode}\sqrt{L/C}/E, V_o/E$), which permits a simplified design of the circuit with a resistance loaded secondary. Although the analysis is not easily extended to the configuration of a rectifier and LC filter at the output, it provides a very good guide to that circuit.

If the shunt-loaded configuration of Fig. 6.12 is used, output regulation is also achieved by varying the SCR triggering frequency. Output voltage waveshapes at the cathodes of the rectifier in Fig. 6.12 are similar to the current waveshapes of Figs. 6.13*c* and 6.14 for the series-loaded circuit. The output LC filter averages the voltage waveshapes for the shunt-loaded circuit, whereas the output capacitor alone is needed to average the current waveshapes for the series-loaded case.

Since the shunt-loaded circuit requires output inductors, it has no advantages over the series-loaded configuration. The shunt circuit has only been touched on here to bring attention to the Mapham paper with its useful computer analysis, which is basic to a full understanding of the series-loaded circuit. The shunt-loaded circuit has been used with MOSFETs as a much-higher-frequency DC/DC converter.[13]

6.4.5 Single-ended SCR resonant converter topology design[3,5]

The single-ended SCR resonant converter (Fig. 6.8) has been discussed in qualitative terms in Sec. 6.3. A more quantitative analysis will be

presented here. The discussion herein contains some simplifying approximations but is not yet accurate enough to permit a workable design. A more rigorous discussion would require computer analysis of the circuit.

The circuit is shown again in Fig. 6.15 in a more useful configuration with the load coupled to the series resonant circuit through an

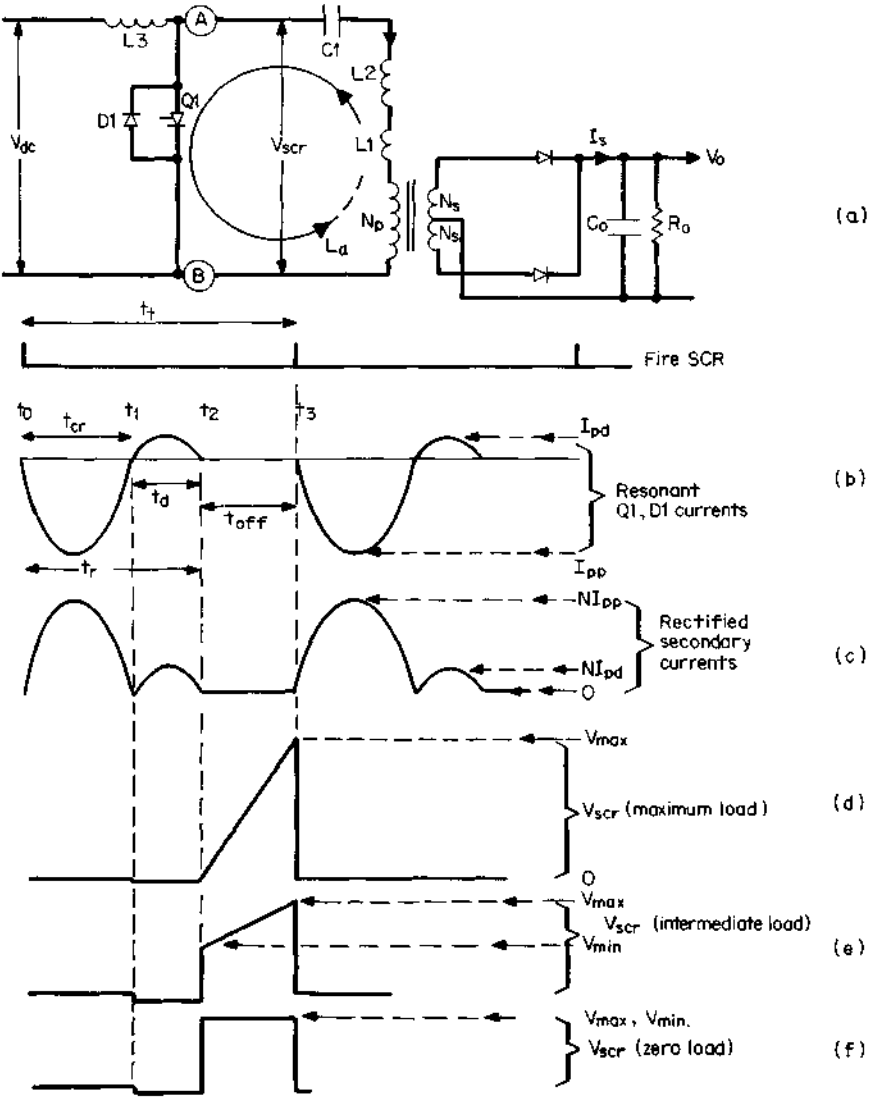


Figure 6.15 Single-ended SCR resonant converter.

isolating transformer. The circuit is a series-loaded circuit with the secondary load reflected into the primary in series with the resonating elements (C_1 and $L_1 + L_2$). The resonating inductance is the leakage inductance L_1 of $T1$ plus another discrete inductance L_2 to make up the total inductance required for the desired resonant period. If L_2 is relatively large compared to L_1, unavoidable tolerances in L_1 are not significant.

The $T1$ magnetizing inductance, although in series with the resonating elements, does not, in most circumstances, affect the resonant period or circuit operation as it is shunted by the secondary load resistance reflected into the primary. But with an open-circuited secondary, the high impedance of the magnetizing inductance can kill the Q of the resonant circuit and prevent SCR turnoff. By gapping the $T1$ core, the circuit can operate successfully with extremely large values of R_o.

Recall the basic circuit operation. Assume that the left-hand side of $C1$ has been charged up to some DC voltage V_{max} higher than V_{dc} (as the boost regulator of Sec. 1.4.1). Now when $Q1$ is fired with a narrow trigger pulse, $Q1$ turns on and will stay on until its current reduces to zero. With $Q1$ on, a voltage step of amplitude V_{max} is applied to all the elements in series to the right of $Q1$. This causes a half sine wave of current of duration t_{cr} to flow through $Q1$ in loop L_a. When the $Q1$ primary current falls back down to zero at t_1, it reverses and continues flowing as a half sine wave of duration t_d through $D1$.

At low loads, the duration of the $Q1$ and $D1$ half sine waves (t_{cr} and t_d) are closely equal to $\pi\sqrt{(L_1 + L_2)C_1}$. At higher loads, t_{cr} increases somewhat and t_d decreases. If t_d at its minimum is still longer than the t_q time of the SCR, the SCR will automatically extinguish and can safely sustain forward voltage without falsely refiring. During the off time $t_{off} = (t_3 - t_2)$, current in $L3$[L_3 is chosen at least 20 times $(L_1 + L_2)$] charges $C1$ back to its original voltage V_{max} and $Q1$ can be fired again. The cycle repeats with a triggering period $t_t = t_r + t_{off} = (2\pi\sqrt{(L_1 + L_2)C_1}) + t_{off}$.

The SCR and diode primary currents produce similarly shaped sinusoids, $N(= N_s/N_p)$ times as large in the secondary. They are rectified and summed by rectifier diodes $D3$, $D4$ as shown in Fig. 6.15c. The average of these secondary current pulses (averaging of the pulses is done by C_o without requiring an output inductor) is equal to the DC output current.

Direct-current output voltage regulation is accomplished by varying the triggering period to maintain a constant average output current and hence constant output voltage. As the DC input voltage or output

resistance R_o increase, t_t is increased by a feedback loop to maintain constant output voltage. As the input voltage or output resistance decrease, t_t is decreased.

6.4.5.1 Minimum trigger period (maximum trigger frequency) selection. The peak SCR and diode currents (I_{pcr}, I_{pd}) of Fig. 6.15*b* are determined by the peak voltage to which the resonating capacitor $C1$ is charged. These peak currents must be known to fix the triggering period and hence the average output current and average output voltage.

Waveshapes of voltage V_{scr} at the left-hand side of $C1$ (or across the SCR) are shown in Fig. 6.15*d* to 6.15*f* for maximum, intermediate, and zero output power. At the instant of a SCR trigger pulse, $C1$ has been charged up to a voltage V_{max} and hence has a stored energy $C_1(V_{max})^2/2$. At the end of the next diode half sine wave t_2 some of this energy has been delivered to the load by the SCR and diode half sine waves.

At intermediate powers, there still is some energy and hence voltage left on the capacitor. Thus when the diode current has fallen back to zero t_2, the left side of $C1$ is unclamped and current from $L3$ starts charging $C1$ up. With some voltage still across $C1$, V_{scr} first steps up by the amount of that voltage V_{min} and then starts rising more slowly as shown in Fig. 6.15*e*.

At maximum load, all the energy stored at the end of an off time t_0 has been delivered to the load, and hence at the start of the next off time t_3, there is no remaining voltage on $C1$ and hence no front-end step as shown in Fig. 6.15*d*.

At zero DC load, all the energy stored on $C1$ at the end of an off time still remains on it at the start of the next off time. Hence V_{max} is equal to V_{min} as in Fig. 6.15*f*. From the waveshapes in Fig. 6.15*d* to 6.15*f*, the minimum triggering period t_t can be established thus. Since the inductor $L3$ cannot support a DC voltage, the average voltage at its output end, must equal that at its input end, which is V_{dc}. Thus

$$V_{dc} = \frac{V_{max} + V_{min}}{2} \frac{t_{off}}{t_t}$$

or

$$V_{max} + V_{min} = \frac{2V_{dc}t_t}{t_{off}} \tag{6.9}$$

and from Fig. 6.15*d*, at maximum load, $V_{min} = 0$, then

$$V_{max} = \frac{2V_{dc}t_t}{t_{off}}$$

But from Fig. 6.15*b*, $t_{off} = t_t - t_r$, then

$$V_{max} = \frac{2V_{dc}}{1 - t_r/t_t} \tag{6.10}$$

From Eq. 6.10, V_{max} is calculated for various ratios of t_r/t_t. Calculation will be for an off-line converter and maximum power will be assumed to occur at minimum AC line input and maximum DC output current as those are the conditions for maximum SCR current. Assume that nominal, minimum line inputs are 115 and 98 V AC, respectively, giving approximately 160 and 138 V of rectified DC. Then for the minimum V_{dc} of 138 V from Eq. 6.10, we can construct Table 6.1.

TABLE 6.1

t_r/t_t	V_{max} (at V_{dc} = 138 V)	V_{max}, V
0.7	$6.6V_{dc}$	911
0.6	$5.0V_{dc}$	690
0.5	$4.0V_{dc}$	552
0.4	$3.3V_{dc}$	455
0.3	$2.9V_{dc}$	393

From Table 6.1, a good compromise choice for t_r/t_t is 0.6. It still permits using a reasonably inexpensive 800-V SCR. For t_r = 16 μs, the resonant half period is 8 μs, which still permits the on-turning SCR to spend most of its on time at a low anode-to-cathode voltage (Fig. 6.5). Then t_t, the minimum trigger period, is 26.6 μs (maximum trigger frequency is 38 kHz). In regulating down to lower output power, the minimum trigger frequency will be about one-third of that or about 13 kHz—not too far down into the audible range.

6.4.5.2 Peak SCR current choice and *LC* component selection. The peak SCR and diode *D*1 currents are shown in Fig. 6.15*b*. Assume, as for the half bridge (Sec. 6.4.2), that the peak diode current is one-fourth the SCR current. Then rectified secondary currents are as shown in Fig. 6.15*c*. The average of those currents for a *T*1 turns ratio *N* is

$$I_{s(av)} = \frac{2I_{pp}N}{\pi}\frac{t_r}{2t_t} + \frac{2I_{pp}N}{4\pi}\frac{t_r}{2t_t}$$

$$= 1.25I_{pp}N\frac{t_r}{\pi t_t} \tag{6.11}$$

where I_{pp} is the peak primary current (Fig. 6.15*b*). This is equal to the average or DC output current at minimum line input and maximum current output. The output voltage at minimum output load resistor R_o is

$$V_o = 1.25NI_{pp}R_o\frac{t_r}{\pi t_t} \tag{6.12}$$

Now the transformer turns ratio must be chosen and the magnitudes of the resonant LC components selected to yield the peak resonant currents given above.

Choose, as in Sec. 6.4.2 as per the suggestion in the Chambers paper,[9] that the $T1$ primary voltage is taken as 60 percent of the voltage across the entire loop (points A to B in Fig. 6.15*a*). This is the $0.6V_{max}$ of Eq. 6.10. This voltage is clamped through the turns ratio N against the output voltage plus one diode rectifier drop. Then

$$N = \frac{0.6V_{max}}{V_o + 1} \tag{6.13}$$

Now when the SCR is fired, the voltage applied to the series resonant circuit elements is $V_{ap} = V_{max} + 0.6V_{max} = 1.6V_{max}$ and the peak amplitude of the half-sine-wave SCR current pulse is

$$I_{pp} = \frac{V_{ap}}{\sqrt{(L_1 + L_2)/C_1}} = \frac{1.6V_{max}}{\sqrt{(L_1 + L_2)/C_1}} \tag{6.14}$$

Since, in Eq. 6.11, maximum DC output current is specified and all other terms in that relation are known (Eq. 6.13 and Table 6.1), Eq. 6.14 gives the ratio of $(L_1 + L_2)/C1$. Also since the resonant period t_r was chosen in Sec. 6.4.5.1 as 16 μs, we obtain

$$t_r = 2\pi\sqrt{(L_1 + L_2)C_1} = 16 \times 10^{-6} \tag{6.15}$$

Between Eqs. 6.14 and 6.15, there are two unknowns and two equations and so both C_1 and $(L_1 + L_2)$ are determined.

6.4.5.3 Design example. Design a single-ended SCR resonant converter with the following specifications:

Output power	1000 W
Output voltage	48 V
Output current	20.8 A

AC input, nominal	115 V AC rms
AC input, minimum	98 V AC rms
Rectified DC, nominal	160 V
Rectified DC, minimum	138 V

From Table 6. 1, chose $t_r/t_t = 0.6$, giving $V_{max} = 690$ V. From Eq. 6.13

$$N = \frac{0.6V_{max}}{V_o + 1}$$
$$= \frac{0.6 \times 690}{49}$$
$$= 8.44$$

From Eq. 6.11

$$I_{s(av)} = I_{o(dc)} = 20.8 = 1.25I_{pp}N\frac{t_r}{\pi t_t}$$
$$= 1.25I_{pp} \times 8.44\left(\frac{0.6}{\pi}\right)$$

or $I_{pp} = 10.3$ A. Then, from Eq. 6.14

$$I_{pp} = 10.3 = \frac{1.6V_{max}}{\sqrt{L_1 + L_2)/C_1}}$$
$$= \frac{1.6 \times 690}{\sqrt{(L_1 + L_2)/C_1}}$$

or
$$\sqrt{\frac{(L_1 + L_2)}{C_1}} = 107.8 \qquad (6.14a)$$

From Eq. 6.15

$$t_r = 16 \times 10^{-6} = 2\pi\sqrt{(L_1 + L_2)C_1}$$

or
$$\sqrt{(L_1 + L_2)C_1} = 2.55 \times 10^{-6} \qquad (6.15a)$$

and from Eqs. 6.14a, 6.15a, $C_1 = 0.024$ μF, $L_1 + L_2 = 275$ μH.

A choice of a lower value for t_r/t_t would have yielded a lower maximum SCR voltage stress (Table 6.1) and hence possibly greater reliability. But it would have resulted in larger values of t_t (lower trigger frequencies), and at low output power, the resulting value of t_t (Eq. 6.12) would have brought trigger frequency down far into the audible frequency range.

6.5 Cuk Converter Topology—Introduction[14–16]

In its specialized area of application, this is a very imaginative and valuable topology. Its major asset is that at both input and output, ripple currents are continuous, i.e., there is no time gap where the ripple current falls to zero. Thus in the buck regulator of Fig. 1.4, it is seen that if L_o is made sufficiently large, output current is continuous (Fig. 1.4*f*) but input current is discontinuous (Fig. 1.4*d*). In the boost regulator (Figs. 1.10, 1.11), input current is continuous (Fig. 1.11, I_{L1}) but the output current through the rectifier diode is discontinuous (Fig. 1.11, I_{D1}).

In applications where very low input and output noise is essential, it is important to have ripple currents at input and output ramp up and down without going back down to zero as in Fig. 1.4*f*. At the inputs of most topologies (forward converters, push-pulls, bucks, flybacks, and bridges), this is usually done by adding an RFI input filter. But this adds cost and space.

6.5.1 Cuk converter—basic operation

The Cuk converter (Fig. 6.16*a*) has continuous current at both inputs and outputs as seen in Fig. 16.6*e* and 6.16*f*. By making $L1$ and $L2$ sufficiently large, the amplitude of the current ramps can be made extremely small. And as discussed below, by winding $L1$ and $L2$ on the same core, the ripple amplitude can be reduced to zero.

The circuit is shown in Fig. 6.16*a* in its basic form with its input and output having a common DC return. Input and output returns can be DC isolated by adding a transformer, as will be discussed below.

In its basic nonisolated version, the circuit works as follows (see Fig. 6.16*a*). When $Q1$ turns on, $V1$ goes steeply negative. Since the voltage across a capacitor cannot change instantaneously, $V2$ goes negative an equal amount, reverse-biasing $D1$ as seen in Fig. 6.16*c* and 6.16*d*. With V_{dc} across $L1$, current ramps up linearly in it, adding to its stored energy. But before $Q1$ had turned on, the left-hand end of $C1$ had been charged up to a voltage V_p and its right-hand end had been clamped to ground via $D1$. This had stored an amount of energy $C_1V_p^2/2$ on it.

Now when $Q1$ has turned on, $C1$ acts as battery delivering current down through $Q1$, up through the bottom end of R_o, and back through $L2$ into the right-hand end of $C1$. Thus the stored energy in $C1$ plus the stored current (or energy) in $L2$ delivers power into R_o and charges the top end of output filter capacitor to a negative voltage $-V_o$.

When $Q1$ turns off, $V1$ goes positive to some voltage V_p, and V_2 follows it up but gets clamped to ground by $D1$ as discussed above. Now

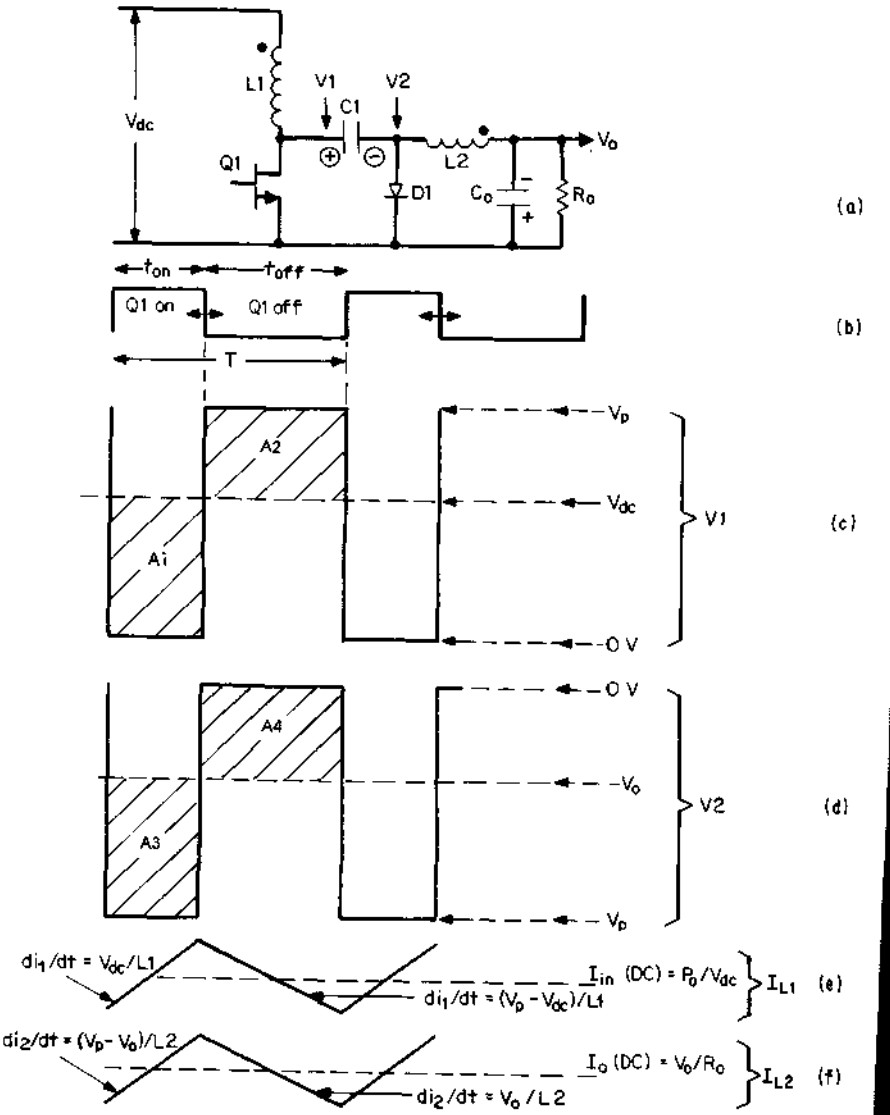


Figure 6.16 Basic Cuk converter with input and output not isolated significant voltage and currents.

with the left-hand end of $L2$ at ground and its right-hand end at $-V_o$, the stored energy or current in $L2$ flows down through $D1$, up through R_o, and back into the right-hand end of $L2$.

When $Q1$ is on, current in $L1$ ramps up at a rate $dI/dt = V_{dc}/L_1$ (Fig. 6.16*e*). Since $V2$ has gone down by the same amount as $V1$ (V_p), the

left-hand end of $L2$ is at $-V_p$ and current in $L2$ ramps up at a rate $dI/dt = (V_o - V_p)/L_2$ as seen in Fig. 6.16*f*.

When $Q1$ is off, $V1$ has risen to a voltage V_p higher than V_{dc} and current in $L1$ ramps downward at a rate $dI/dt = (V_p - V_{dc})/L_1$ (Fig. 6.16*e*).

Then, with $Q1$ off, the left-hand end of $L2$ is clamped to ground through $D1$ and with its right-hand end at $-V_o$, current in $L2$ ramps downward at a rate $di/dt = V_o/L_2$ as in Fig. 6.16*f*. If $L1$, $L2$ are made large enough, these ramp currents rest on some DC level and never fall to zero.

6.5.2 Relation between output/input voltage and *Q* on time

Since $L1$ is assumed to have zero DC resistance, it cannot support a DC voltage. Hence the voltage at its bottom end ($V1$), averaged over one cycle, must equal the DC voltage at its top end (V_{dc}). This is equivalent to stating that in Fig. 6.16*c*, the area $A1$ in volt-seconds is equal to the area $A2$ in volt-seconds. Thus

$$V_p \frac{t_{off}}{T} = V_{dc} \tag{6.16a}$$

or

$$V_p = V_{dc} \frac{T}{t_{off}} \tag{1.16b}$$

Now since the voltage change at $V2$ equals the voltage change at $V1$, the bottom end of the $V2$ voltage is at $-V_p$ during t_{on}. During t_{off}, the top end of $V2$ is clamped to 0 V by diode $D1$.

Similarly, since $L2$ cannot support DC voltage, the voltage averaged over one cycle at its left-hand end ($V2$) must equal the DC voltage at its right-hand end ($-V_o$). Here, too, this is equivalent to stating that in Fig. 6.16*d*, the area $A3$ in volt-seconds is equal to the area $A4$ in volt-seconds. Thus

$$V_p \frac{t_{on}}{T} = V_o \tag{6.17a}$$

or

$$V_p = V_o \frac{T}{t_{on}} \tag{6.17b}$$

Equating relations 6.16b and 6.17b yields

$$V_o = \frac{V_{dc} t_{on}}{t_{off}} \tag{6.18}$$

It is seen from Eq. 6.18 that the magnitude of the DC output voltage can be less than, equal to, or greater than the DC input voltage depending on the ratio t_{on}/t_{off}.

6.5.3 Rates of change of currents in *L*1, *L*2

It is interesting to note that (for $L_1 = L_2$) the upslopes of current in $L1$ and $L2$ during t_{on} are equal and their downslopes during t_{off} are also equal. This is shown in Fig. 6.16*e* and 6.16*f*. It is this fact that makes it possible to reduce current ripple at the input completely to zero and makes this Cuk converter an ultra-low-noise circuit.

Equality of upslope and downslope currents can be shown as follows. During t_{on}, the upslope of $L1$ current is

$$\frac{+di_1}{dt} = \frac{V_{dc}}{L_1}$$

and again during t_{on}, the upslope of $L2$ current is

$$\frac{+di_2}{dt} = \frac{V_p - V_o}{L_2}$$

From Eqs. 6.16a and 6.18

$$\begin{aligned}\frac{+di_2}{dt} &= \frac{1}{L_2}\left(\frac{V_{dc}T}{t_{off}} - \frac{V_{dc}t_{on}}{t_{off}}\right)\\ &= \frac{V_{dc}}{L_2 t_{off}}(T - t_{on})\\ &= \frac{V_{dc}}{L_2}\end{aligned}$$

and for $L_1 = L_2$ the upslopes of current in the inductors are equal during t_{on}.

Now consider the downslopes of current during t_{off}. In $L2$, $-di_2/dt = V_o/L2$, and from Eq. 6.18

$$\frac{-di_2}{dt} = \frac{V_{dc}}{L_2}\frac{t_{on}}{t_{off}} \tag{6.19a}$$

and in $L1$, $-di_1/dt = (V_p - V_{dc})L_1$. But from Eq. 6.16b

$$\frac{-di_1}{dt} = \frac{(V_{dc}T/t_{off}) - V_{dc}}{L_2}$$

$$= \frac{V_{dc}}{L_2} \frac{T - t_{off}}{t_{off}}$$

$$= \frac{V_{dc}}{L_2} \frac{t_{on}}{t_{off}} \tag{6.19b}$$

Thus again, for $L_1 = L_2$, the downslopes of the inductor currents during t_{off} are equal and of magnitude given by Eq. 6.19a and 6.19b. This is seen in Fig. 6.16*e* and 6.16*f*.

6.5.4 Reducing input ripple currents to zero

Now the current ramps at the input and output can be reduced to zero, yielding pure DC in those lines if $L1$ and $L2$ are wound on the same core with the polarities as shown in Fig. 6.17.

If $L1, L2$ are wound on the same core as in Fig. 6.17, that core is now a transformer. During the on time, current flows from V_{dc} into the dot end of $L1$ in an upgoing ramp as seen in Fig. 6.16*e*. But in $L2$, its dot end is positive with respect to its no-dot end, and by transformer action, it forces a voltage across $L1$, making its dot end positive with respect to its no-dot end. This forces a "secondary" current from the dot end of $L1$ back into V_{dc}. This current is a positive-going ramp of the same upslope as the upslope of current from V_{dc}.

Since these two upslopes are of equal magnitude (as demonstrated in Sec. 6.5.3) but with the currents flowing in opposite directions, the net current change during t_{on} is zero; i.e., current flow in V_{dc} is pure DC. That current is $P_o/V_{dc}E$, where P_o is the output power and E is the efficiency.

A similar line of reasoning demonstrates that if the coupling be-

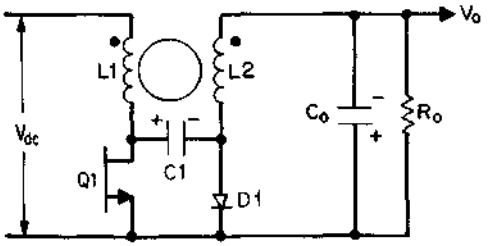


Figure 6.17 If $L1, L2$ are wound on the same core, input and output ripple currents are very closely equal to zero. The ramp currents into the primary are bucked out by an equal and oppositely directed ramp rate reflected into the primary from the secondary. A similar cancellation occurs in the secondary.

tween $L1$ and $L2$ is 100 percent, the current in the load is also pure DC with no ripple component. That current is V_o/R_o.

Further, since it has been shown that the current downslopes in $L1$ and $L2$ during t_{off} are equal, similar reasoning indicates that there are also no ripple currents in the input or output lines during t_{off}.

6.5.5 Isolated outputs in the Cuk converter

In most instances, output returns must be DC-isolated from input returns. This can be done with the addition of a 1/1 isolating transformer as shown in Fig. 6.18. Output voltage is still determined by the ratio t_{on}/t_{off}, and the output polarity can be either positive or negative, depending on which end of the secondary circuit is grounded.

Thus, although the Cuk converter of Fig. 6.18 is a very clever way of producing pure DC currents in input and output lines, the requirement of two pieces of magnetics (the $L1$,$L2$ inductor core and $T1$) is a high price to pay for the advantage.

6.6 Low Output Power "Housekeeping" or "Auxiliary" Topologies—Introduction[15–17]

These are not strictly "topologies" having a broad range of usages; rather, they are specialized circuits for unique applications. But since they generate an output voltage which serves a vital function in any switching power supply design, they are discussed here as separate topologies.

All the topologies discussed thus far require a low power (1 to 3 W) supply of about 10- to 15-V output. It is used to feed the usual

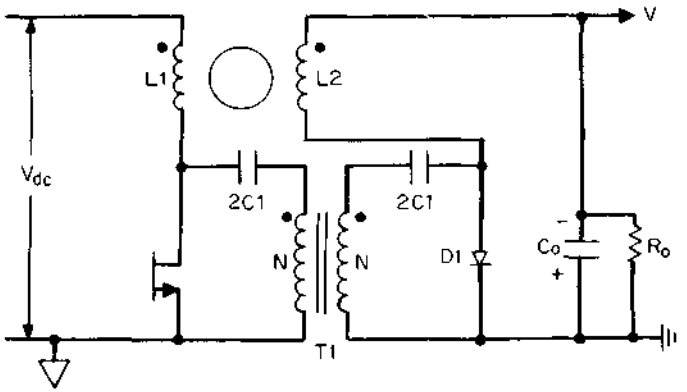


Figure 6.18 By adding a 1/1 isolating transformer, output return is DC-isolated from input return. Either a positive or negative output is now possible depending on which end of the secondary is grounded.

pulse-width-modulating (PWM) chip for the main "power train" and power the logic, sensing circuits which perform various housekeeping functions. Such housekeeping functions may include overcurrent, overvoltage sensing and protection, remote signaling, and correct turnon, and turnoff sequencing for each output in a multioutput supply.

These housekeeping supplies need not always be regulated as the usual loads can tolerate a relatively large range of supply voltage (± 15 percent maximum). But reliability is improved, and more predictable operation of the main power train results if the housekeeping supply is regulated—usually ±2 percent is adequate. The important objectives for these supplies is that they be low in parts count and cost and occupy only a small fraction of the space that the main power train, with all its outputs, occupies.

6.6.1 Housekeeping power supply—on output or input ground?

In any new design, an initial and major decision must be made as to whether the housekeeping power supply with the PWM chip which it powers should be located on output or input ground. The load-driving outputs are located on output ground, which is, in most cases, DC-isolated from input ground. The main switching power transistors are located on input ground—one end of the rectified AC line in off-line converters or at one end of the DC prime power source in battery-operated DC/DC converters.

Thus, in regulating the output voltage, a DC error amplifier must be located on output ground to sense voltage, compare it to a reference voltage, and produce an amplified error voltage. This amplified error voltage is the difference between the reference and a fraction of the output voltage. The amplified error voltage is then used to control width of the pulse which drives the main power transistor or transistors which are located on input ground. A typical example of this is shown in Fig. 2.1.

Since output and input ground are DC-isolated and may be at DC voltage levels tens or hundreds of V apart, the width-modulated pulse cannot be DC-coupled to the power transistor.

Thus if the error amplifier and pulse-width modulator are on output ground (usually in the PWM chip), the width-modulated pulse is transferred across the output ground–input ground barrier by a pulse transformer. It is the function of the housekeeping supply, whose input power comes from the prime power source at input ground, to produce, referenced to output ground, the usual 10 to 15 V for the 1 to 3 W of housekeeping power.

Such a housekeeping supply is also often used when the PWM chip is located on input ground. Although power for the chip may be derived from an auxiliary winding on the main transformer when the main power transistor is being driven, if the drive is shut down (for overvoltage or overcurrent reasons), that power goes away and it is no longer possible to energize remote indicators. Further, on shutdown, as voltage from the auxiliary winding goes away, supply voltage for the PWM chip decays, and odd time races causing excessive pulse width can occur and may cause failures.

In general, it is far more reliable to have a housekeeping supply which is always present instead of deriving it ("bootstrapping" it) from an auxiliary winding on the main power transformer.

An alternative method of transmitting a measure of the output voltage across the output ground–input ground barrier to width-modulate the power transistor on input ground is via an optical coupler. But this also requires a housekeeping supply on input ground if bootstrapping from an auxiliary winding on the main power transformer is considered undesirable.

Some schemes for implementing these housekeeping supplies are discussed below.

6.6.2 Housekeeping supply alternatives

In an attempt to minimize the parts count of housekeeping supplies, many designers have resorted to single- or two-transistor transformer coupled feedback, self-oscillating circuits. This saves the space and cost of a PWM chip or a stable multivibrator for generating the required AC drive frequency to some kind of a driven converter. Further, by using a transformer coupled feedback oscillator, adding a separate winding on the transformer, provides output power referenced to any desired DC voltage level. The feedback from a collector to a base winding which keeps the circuit oscillating also provides sufficient drive to deliver the required power from the output winding.

Such self-oscillating housekeeping converters are very well covered in Keith Billings' handbook.[16]

These self-oscillating converters appear at first glance to be very attractive because of their simple circuit and low parts count. But without additional circuitry, most do not produce regulated DC output voltage and have various other shortcomings. Adding circuitry to regulate the output and overcome any other shortcomings increases internal dissipation, parts count, and circuit complexity.

At some point it becomes debatable whether these self-oscillating converters are a better choice than a conventional single-transistor,

low-power, low-component-count-driven converter such as a flyback, fed from its own PWM chip.

With the increasingly lower price of a PWM chip, no need for an output inductor and a simple scheme for regulating its output voltage on output ground by sensing the voltage of a slave winding on input ground, the flyback is a credible alternative to a self-oscillating circuit.

Nevertheless, two of the most frequently considered self-oscillating types—the Royer and Jensen oscillating converters—are considered below and compared to a simple flyback.

6.6.3 Specific housekeeping supply block diagrams

Figures 6.19 to 6.21 show the block diagrams of three reasonable approaches for a housekeeping supply where the error amplifier and pulse width modulator are in a PWM chip on output ground.

6.6.3.1 Housekeeping supplies for AC prime power. Figure 6.19 shows the simplest and most frequently used scheme when the prime power is AC. A small (usually 2- to 6-W) 50/60-Hz transformer is powered from the AC input and has its secondary referenced to output ground. Such transformers are available from a large number of manufacturers and have tapped primaries so that they can be fed from either 115 or 220 V AC—either 50 or 60 Hz.

Typical sizes are 1.88 × 1.56 × 0.85 in for a 6VA unit or 1.88 × 1.56 × 0.65 in for a 2VA unit. They come with a large range of standard secondary voltages and usually have two secondaries which can be wired in series for full-wave center-tapped output rectifier or in

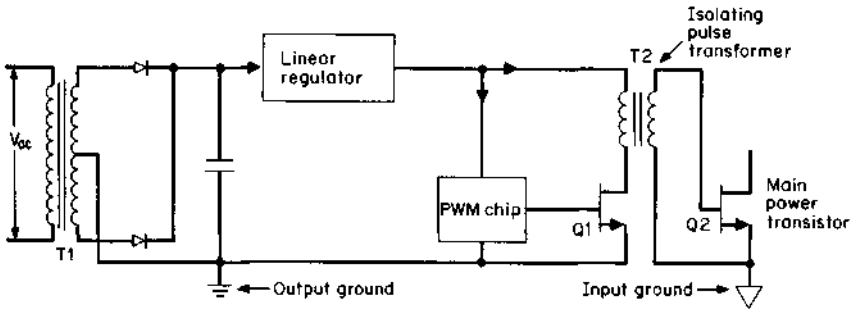


Figure 6.19 For an off-the-AC-line converter, the simplest housekeeping supply is a small (2-W) 60-Hz isolating transformer with its secondary generating a rectified 15 V, referenced to output ground. This is followed by an inexpensive linear regulator (12-V output) which is referenced to output.

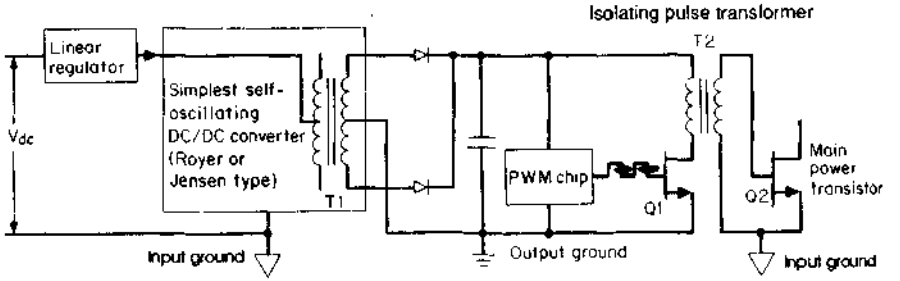


Figure 6.20 A housekeeping supply for DC input voltage. When AC voltage is not present to provide a rectified DC at output ground, a simple magnetically coupled feedback oscillator fed from the DC at input ground is used as a DC/DC converter to provide output voltage reference to output ground to power the PWM chip on output ground. The housekeeping supply output is proportional to the input voltage. Input regulator may be unnecessary if main PWM chip can tolerate ±10 percent input voltage variation.

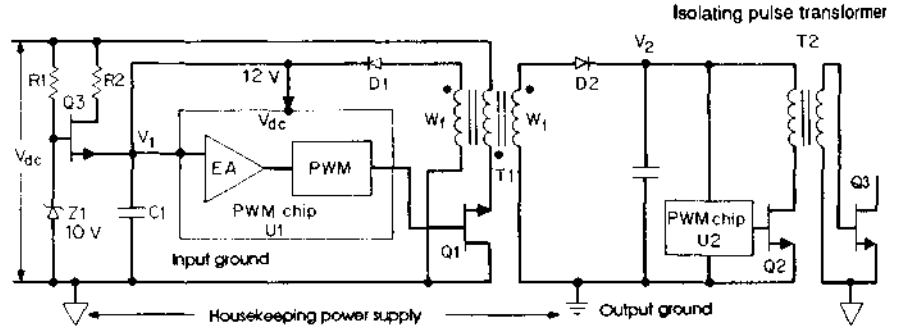


Figure 6.21 A minimum-parts-count flyback as a housekeeping supply. Although this approach may have more parts than Fig. 6.20, it generates a regulated output V_2 without requiring any linear regulators and at the 2- to 3-W power level, may require less input power than that shown in Figs. 6.19 and 6.20.

parallel for a bridge output rectifier. The secondary is rectified and filtered with a capacitor input filter, and the rectified DC is returned to output or input ground as desired.

The rectified DC will have the same tolerance as the AC input—usually ±10 percent. Since most PWM chips can accept DC inputs ranging from 8 to 40 V, it is not essential to regulate the output. But in general, safer and more predictable performance results if the output is regulated. Thus, the transformer secondary voltage frequently is chosen to yield a rectified DC voltage of about 3 V above the desired regulated DC voltage and an inexpensive integrated-circuit linear regulator in a TO 220 package is added in series after the filter capacitor as shown in Fig. 6.19. The configuration is usually designed to yield a regulated +12-V output and achieves an efficiency of about 55 percent at 3-W output at a 10 percent high-line input.

6.6.3.2 Oscillator-type housekeeping supply for AC prime power. But when the prime input power is DC, there is no AC voltage easily available to produce a rectified DC at output ground. Figure 6.20 shows a usual configuration for this case.[17] A simple magnetically coupled feedback oscillator fed from the DC input produces high-frequency square-wave output in a secondary referenced to output ground.

The secondary is rectified and filtered with a capacitor input filter, and the resultant DC is returned to output ground. Because of the high-frequency square wave, the filter capacitor after secondary rectification is far smaller than is required for the 60-Hz rectifier-filter of Fig. 6.19.

The rectified DC output in such a scheme is almost often proportional to DC input voltage. Thus, if DC output voltage variation of about ±10 percent is acceptable, this is a very efficient and low-component-count scheme if an efficient oscillator is available. There are a number of oscillator configurations which can achieve efficiencies of 75 to 80 percent at an output power level of 3 W.[16] One such particularly useful oscillator, the Royer circuit, will be discussed below.

If regulated DC output is required, the oscillator is preceded by an integrated-circuit linear regulator as shown in Fig. 6.20. Adding the linear preregulator would drop the total efficiency down to about 44 percent (assuming a telephone industry supply where the maximum DC input is 60 V and the linear preregulator drops that to 35 V—just below the minimum specification for a telephone industry supply).

At a small increase in parts count and cost, by replacing the linear regulator of Fig. 6.20 by a simple buck regulator, efficiency can be brought up to 70 and possibly 75 percent at a 3-W output power level.

6.6.3.3 Flyback-type housekeeping supplies for DC prime power. Now, however, with the added cost and component count needed to produce regulated DC in the scheme of Fig. 6.20, simple self-oscillator schemes begin to lose their attractiveness. Figure 6.21 shows a third alternative for the housekeeping supply. It is a simple flyback driven from one of the very many currently inexpensive PWM chips. It is powered from the DC prime power at input ground and a secondary winding $W1$ delivers DC voltage referenced to output ground.

In Figure 6.21, the PWM chip is powered via emitter-follower $Q3$ at initial turnon. At initial turnon, the $Q3$ output voltage is one base-emitter voltage drop less than the 10-V zener diode $Z1$. The resulting 9-V output at the $Q3$ emitter is enough to power the PWM chip ($U1$) which commences driving flyback transistor $Q1$ and delivers, via $W1$, output power to the main PWM chip ($U2$) on output ground.

As $U1$ is now powered via $Q3$, a bootstrap winding W_f on the flyback transformer starts generating output voltage. The number of turns on W_f is chosen to produce at filter capacitor $C1$, a voltage of about 12 V, which is higher than the 9 V at that point when it is being fed from the $Q3$ emitter. Thus with the emitter of $Q3$ at 12 V and its base at 10 V, its base is reverse-biased and it turns off. The auxiliary PWM chip on input ground now continues to be powered from W_f, and the main PWM chip on output ground continues to be powered via $W1$.

Resistor $R2$ is chosen small enough to deliver the current required by $U1$ during the initial turnon interval. Dissipation in $R2$ is negligible as that interval lasts only a few tens of microseconds. Resistor $R1$ carries current continuously but dissipates very little power as it carries only the $Q3$ base current which need be only about 1 mA to supply the initial startup current of about 10 to 20 mA for $U1$.

The error amplifier in $U1$ senses its own bootstrapped supply voltage V_1 and keeps it constant as a "master." The output voltage V_2, from $W1$ is a "slave" and is also quite constant (within 1 to 2 percent) as slaves track the master very well in a flyback topology.

6.6.4 Royer oscillator housekeeping supply—basic operation[17,18]

This configuration is shown in Fig. 6.22*a*. It was one of the earliest applications of transistors to power electronics and was conceived in 1955, only a few years after transistors were invented.[1]

It had been used to generate square-wave AC and with rectification, as a DC/DC converter up to a power level of a few hundred watts. But in its original form, it had two significant drawbacks which limited its usefulness and in some cases made it unreliable. But with the addition of three small changes and with more modern components, it has become a valuable circuit—at power levels as low as <10 W and up to the 200- and possibly 300-W power level.

The original and basic Royer oscillator shown in Fig. 6.22*a* works as follows. It is a push-pull circuit with positive feedback from collectors to the base windings to keep it oscillating. The positive feedback can be seen from the dots on the collector and base windings.

Thus, assume that $Q1$ has been turned on and is in saturation. The no-dot end of its half primary N_{p1} is positive, and hence the no-dot end of its base winding N_{b1} is also positive. Voltage across N_{p1} is V_{dc} (assuming negligible V_{ce} drop); N_{p1} delivers output current to the load via N_{s1} and also enough current to the $Q1$ base via N_{b1} and $R1$ to keep $Q1$ on and in saturation at the maximum current reflected into the primary by the minimum R_o.

The $T1$ transformer core is made of material with a square hyster-

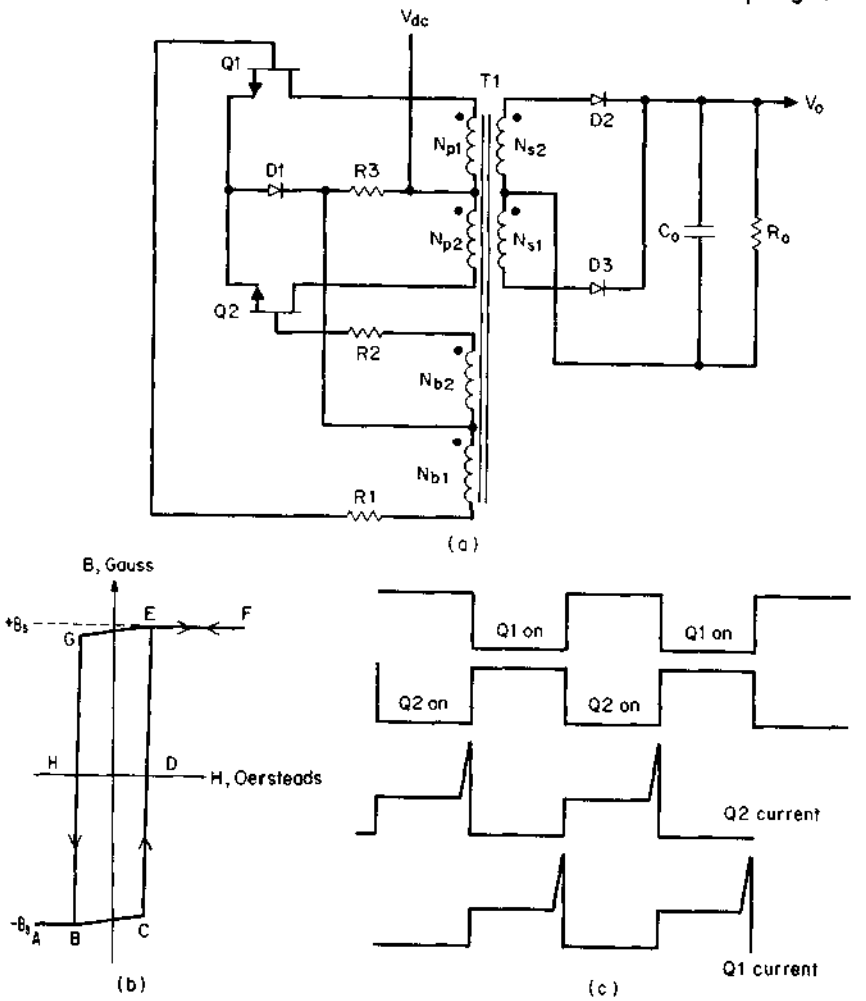


Figure 6.22 (*a*) Basic Royer oscillator. (*b*) Square hysteresis loop of $T1$ core. (*c*) Characteristic high current spikes at end of on time. These spikes are major drawbacks in Royer oscillators. As long as the core is on the vertical part of its hysteresis loop, the positive feedback from N_p to N_b widings keeps a transistor on and in saturation. When the core has moved to either the top or the bottom of its hysteresis loop, coupling between the collector and base windings immediately drops to zero as the core permeability in such a square loop material is unity. The on transistor's base voltage and current drop to zero, and its collector voltage starts to rise. Some small residual air coupling couples this rising collector voltage into the opposite base, and by positive feedback, the opposite transistor turns on. In one half period, the core moves along the path *CDEF*, then in the next half period along the path *FEGHBA*.

esis loop as seen in Fig. 6.22*b*. Assume that when $Q1$ turned on initially, the core was at point C on its hysteresis loop. With a voltage V_{dc} across N_{p1}, the rate of change of flux density in the core is given by Faraday's law as

$$\frac{dB}{dt} = \frac{V_{dc} \times 10^{+8}}{N_{p1}A_e} \tag{6.20}$$

The core moves up the hysteresis loop from negative saturation $-B_s$ to positive saturation $+B_s$ along the path CDE. The time required for this is given by Eq. 6.20 as

$$T1 = \frac{T}{2} = \frac{dB\, N_{p1}A_e \times 10^{-8}}{V_{dc}}$$

$$= \frac{2B_s N_{p1}A_e \times 10^{-8}}{V_{dc}} \tag{6.21}$$

When the core has reached point E, it is saturated, its permeability is close to unity, and coupling between the $Q1$ collector and base windings suddenly drops to zero. The $Q1$ base current quickly drops to zero and $Q1$ collector voltage starts rising.

A small residual air coupling from the dot end of N_{p1} to the dot end of N_{b2} starts turning $Q2$ on. As it commences turning on, positive feedback from the N_{p2} to the N_{b2} winding speeds up the turnon process until $Q2$ is fully on. When $Q1$ had been on, the no-dot end of N_{p1} had been positive and the core had moved up the hysteresis loop. But now with $Q2$ on, the dot end of N_{p2} is positive and the core is driven back down the hysteresis loop along the path EGHBA.

It requires the same $T/2$ given by Eq. 6.21 to move back down the hysteresis loop. The preceding cycles repeat and thus the circuit oscillates at a frequency given by

$$F = \frac{1}{T} = \frac{V_{dc} \times 10^{+8}}{4B_s N_p A_e} \tag{6.22}$$

In Fig. 6.22*a*, the function of resistor $R3$ is to start the circuit oscillating. When V_{dc} is first applied, neither $Q1$ nor $Q2$ is on and the above cycles cannot commence. But current from V_{dc} flows down through $R3$ to the base windings center tap, then through either half base winding, through its base resistor and then its base, and the cycle can now start.

In general, the transistor with the highest gain will be the one to turn on first. Once the circuit is oscillating, base current flows from the outer end of one base winding, through its base resistor, its base,

out of the transistor emitter, through $D1$, and then back into the base winding center tap.

The circuit of Fig. 6.22*a* shows only one of many possible base drive configurations. The base resistors shown serve to limit base current which may be excessive at high temperature and cause long transistor storage delay. Collector current limiting may be done with emitter resistors. Baker clamps (to be discussed in a later chapter on bipolar base drives) may be used to make the circuit less sensitive to load changes, production spread in transistor gain, and temperature.

6.6.4.1 Royer oscillator drawbacks[17]. The basic Royer oscillator has two major drawbacks, but these can be corrected by quite simple means. The effect of the first drawback can be seen in Fig. 6.22*c* as an ultra-high-current spike at the end of a transistor on time. This spike may last only 1 to 2 μs but may be up to three or five times the current prior to the spike.

The spike occurs at a collector voltage about equal to the supply voltage and thus adds significantly to the transistor dissipation. Since it comes at simultaneously high current and voltage, it may exceed the safe operating area (SOA) boundary and cause failure by "second breakdown" even if the average dissipation is low.

The spike is inherent to the very nature of the Royer oscillator, and it can be explained as follows. During the time—say—$Q1$ is on, $Q2$ has a reverse bias and is held off (observe the dots at the base windings). When the core has moved up—say, to the top of its hysteresis loop—it saturates, the windings can no longer support voltage and the $Q1$ collector voltage rises.

However, its base voltage does not immediately go negative to turn-off collector current. It goes negative only after the stored base charges drain away and $Q2$ has turned on sufficiently to produce a solidly negative voltage at the no-dot end of N_{p2} and hence at the no-dot end of N_{b1}. During this delay between core saturation at the end of one on time and the flipover to the opposite transistor turnon, the off-turning transistor operates with high collector voltage and a high-current spike and may fail.

The second drawback is really part of and a partial cause of the first. It is the long delay between core saturation on one side and turnon of the opposite transistor. During this delay, voltage at the base of the off-turning transistor hangs on at about the 0.5-V level and drifts slowly negative before being pulled down abruptly by opposite transistor turning solidly on.

While the off-turning base is drifting slowly down from the 0.5-V level, the transistor is still partially on at a high collector-to-emitter voltage. Frequently, while the base hangs on thus, the partially on

transistor will oscillate at a very high frequency. This can easily be corrected by small capacitors—empirically chosen between 100 and 500 pF—cross-coupled from collectors to opposite bases.

There is a further drawback in that the oscillator square-wave frequency is directly proportional to supply voltage (Eq. 6.22). There are many systems-related objections to a variable-frequency switching power supply. They all relate to the fact any RFI generated will cover a wider and more continuous frequency spectrum with a variable frequency as opposed to a fixed-frequency switching power supply.

Figure 6.23*a* shows the schematic and critical waveforms for a typical Royer oscillator DC/DC converter for 2.4-W output operating from 38 V DC—the minimum specified input for a telephone industry power supply.

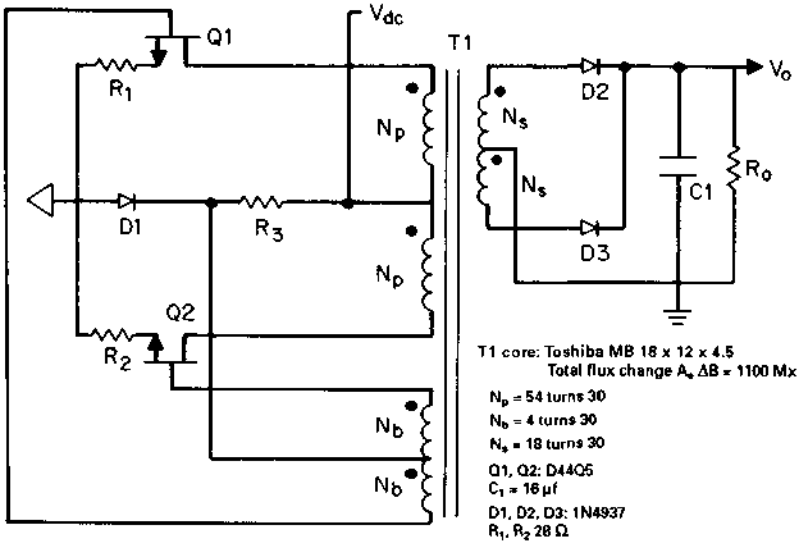
The Royer oscillator is clearly low in parts count, but Fig. 6.23*c* shows the aforementioned spikes at the end of turnon and also in this case at the start of turnon. Collector voltages are shown in Fig. 6.23*b*. The low efficiency of only 50.6 percent is a consequence of the dissipation due to the current spikes at turnon and turnoff.

6.6.4.2 Current-fed Royer oscillator[19]. By simply adding an inductor in series with Royer transformer center tap, the aforementioned current spikes at turnoff and turnon are eliminated and efficiency is greatly increased.

The addition of the series inductor makes the circuit constant-current-fed as opposed to voltage-fed and achieves all the advantages of current-fed topologies discussed in Sec. 5.6. The series inductor helps in the following way. When the core has saturated on one side, the associated transistor commences having a large current spike with a large di/dt. Since the current in an inductor cannot change instantaneously, the voltage at the transformer center tap drops down to ground and the collector current is limited to the value it had just prior to core saturation.

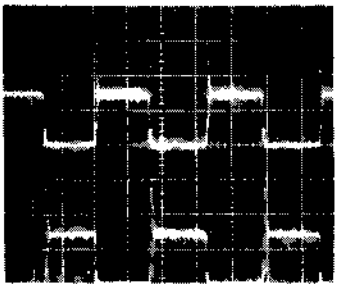
Now the start current from $R3$ turns on the opposite transistor and both transistors are on simultaneously at close to zero collector-to-emitter voltage—certainly for the duration of the storage time in the off-turning transistor. Now the off-turning transistor turns off at zero collector-to-emitter voltage—the condition for minimum transient turnoff dissipation (Sec. 2.2.12.1). The on-turning transistor turns on at zero collector-to-emitter voltage, which also minimizes transient turnon losses.

The benefits from this inductor or constant-current-fed Royer can be seen in Fig. 6.24. There the same Royer circuit of Fig. 6.23 was fed from an adjustable voltage power supply through a series 630-μH inductor (50 turns on 1408-3C8 ferrite core with a total 2-mil air gap).

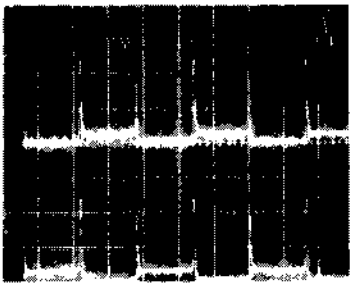


T1 core: Toshiba MB 18 x 12 x 4.5
Total flux change $A_e \Delta B$ = 1100 Mx

N_p = 54 turns 30
N_b = 4 turns 30
N_s = 18 turns 30
Q1, Q2: D44Q5
C_1 = 16 µf
D1, D2, D3: 1N4937
R_1, R_2 28 Ω
R_3, 5.1 Ω
R_o 50.0 Ω
V_{dc} =40 V V_o =11 V



Top: V_{ce} (Q1) 50V, 10µs/cm
Bottom: V_{ce} (Q2) 50V, 10µs/cm



Top: $V_{(R1)}$ 10 V, 10 µs/cm
I_o = 10/28 = 357 mA/cm
Bottom: $V_{(R2)}$ 10 V, 10 µs/cm
I_c = 10/28 = 357 mA/cm

V_{dc} = 38.6 V	R_o = 50.0 Ω
I_{dc} = 124 mA	I_o = 219 mA
P_{in} = 4.78 W	P_o = 2.42 W
V_o = 10.99 V	Efficiency = 50.6%

Figure 6.23 A typical Royer oscillator using a square hysteresis loop core, frequently used as a low-power "housekeeping supply" to power a PWM chip on output ground with its own power derived from power source on input ground. It is low in parts count, but the high-current spikes at the end of on time and often at start of turnon make it unreliable and undesirable despite its low parts count.

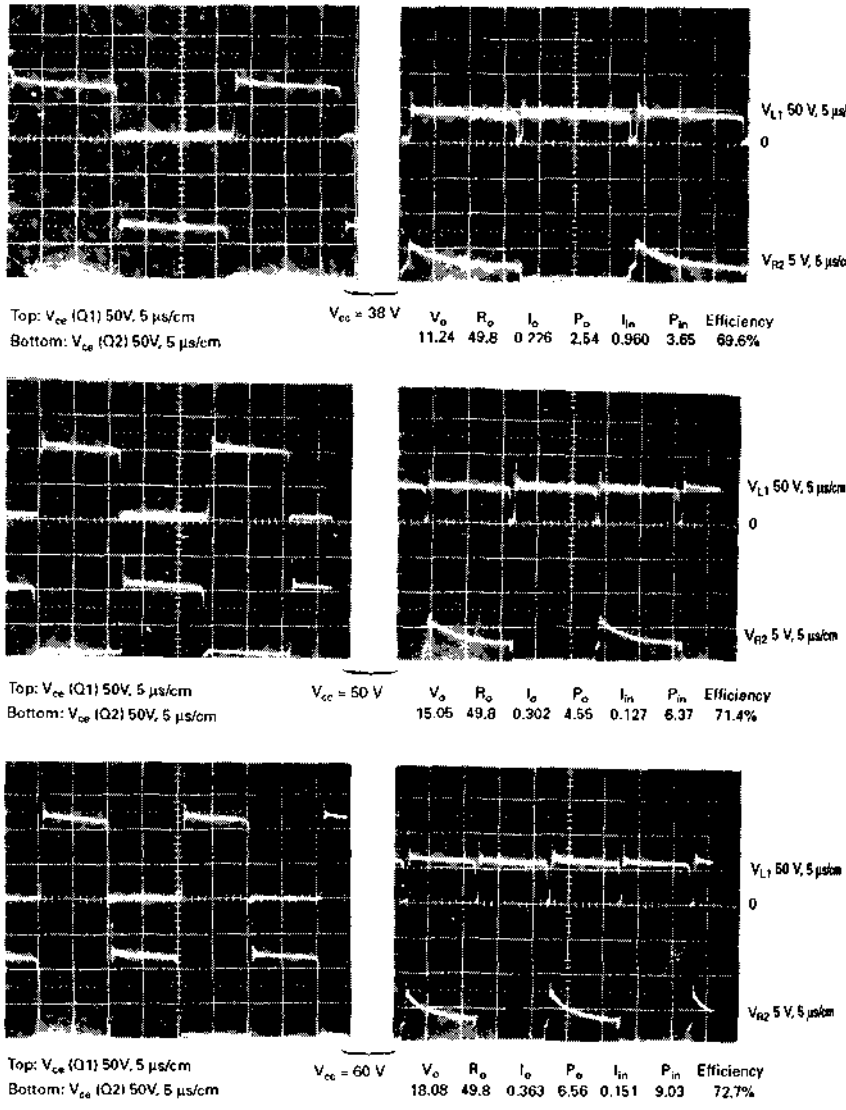


Figure 6.24 Waveform in a current-fed Royer oscillator. By adding an inductor in series between V_{cc} and the transformer center tap, the high-current spikes at the start and end of the transistor on time (Fig. 6.22*c*) are eliminated and efficiency improves greatly. This occurs because the center tap voltage drops to zero when both transistors are simultaneously on for a brief instant at each transition. Waveforms are for the circuit of Fig. 6.22*a* with an inductor of 630 μH in series with V_{cc} and at 38, 50, and 60 V.

TABLE 6.2

$V_{dc(in)}$, V	$I_{dc(in)}$, mA	P_{in}, W	V_{out}, V	R_o, Ω	P_{out}, W	Efficiency, %
38.0	96	3.65	11.24	49.8	2.54	69.6
50.0	127	6.37	15.05	49.8	4.55	71.4
60.0	151	9.03	18.08	49.8	6.56	72.7

The transistor currents shown in Fig. 6.24 show no sign of an end of on-time spike. The numerical data of Fig. 6.24 are summarized in Table 6.2.

It is seen from Table 6.2 that efficiency has averaged about 71 percent with a constant load over the 38- to 60-V range of telephone industry specifications for power supplies. This compares to the 50.6 percent efficiency for the same Royer with the same 49.8-Ω load resistor without the series input inductor (Fig. 6.23).

The voltage drop down to zero at the transformer center tap, due to the input inductor, is clearly seen in Figs. 6.24 and 6.25.

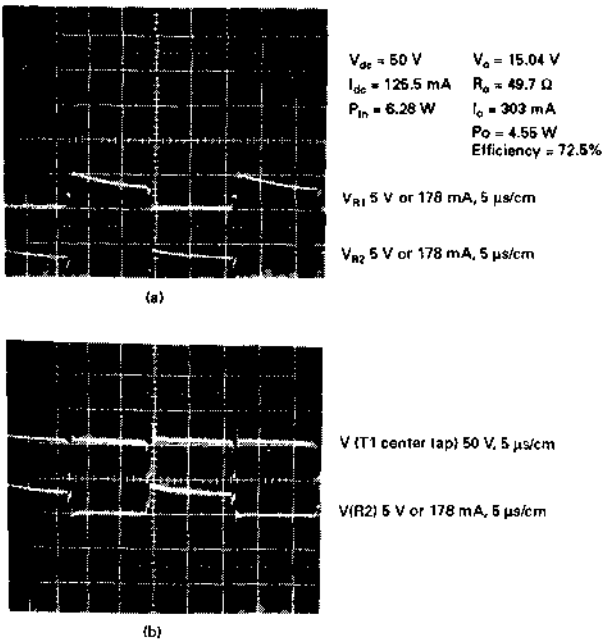


Figure 6.25 (*a*) Voltage across emitter resistors in Fig. 6.22 with 1630 μH in series *T*1 center tap. (*b*) Circuit as in Fig. 6.24*a*, showing *T*1 center tap voltage dropping to zero at each transmission.

Thus if output voltage variations of 11 to 18 V for input changes of 38 to 60 V were acceptable, the unregulated, current-fed Royer DC/DC converter would be a very good choice because of its very low parts count.

6.6.4.3 Buck preregulated current-fed Royer converter. In many applications, however, regulated output voltage is required. Output voltage regulation can be achieved with very little more complexity and cost by preceding the Royer with a buck regulator as in Figs. 6.26 and 6.27. Since buck regulators can quite easily be built with efficiencies of 90 percent, total efficiency does not suffer too much even though the power is handled twice—in the buck and in the Royer. Figure 6.27 shows the composite efficiency ranges from 57.9 to 69.5 percent over an input voltage range of 38 to 60 V and an output power range of 2.3 to 5.7 W.

For Fig. 6.27, the buck regulator sensed output of the buck itself, keeping it constant and running the Royer open loop. This was good enough as the Royer output voltage is constant for constant input voltage and has quite good open-loop load regulation. In Fig. 6.27 it is seen that output voltage change over the above-mentioned line and load change was only from 9.79 to 10.74 V—easily adequate for a housekeeping power supply.

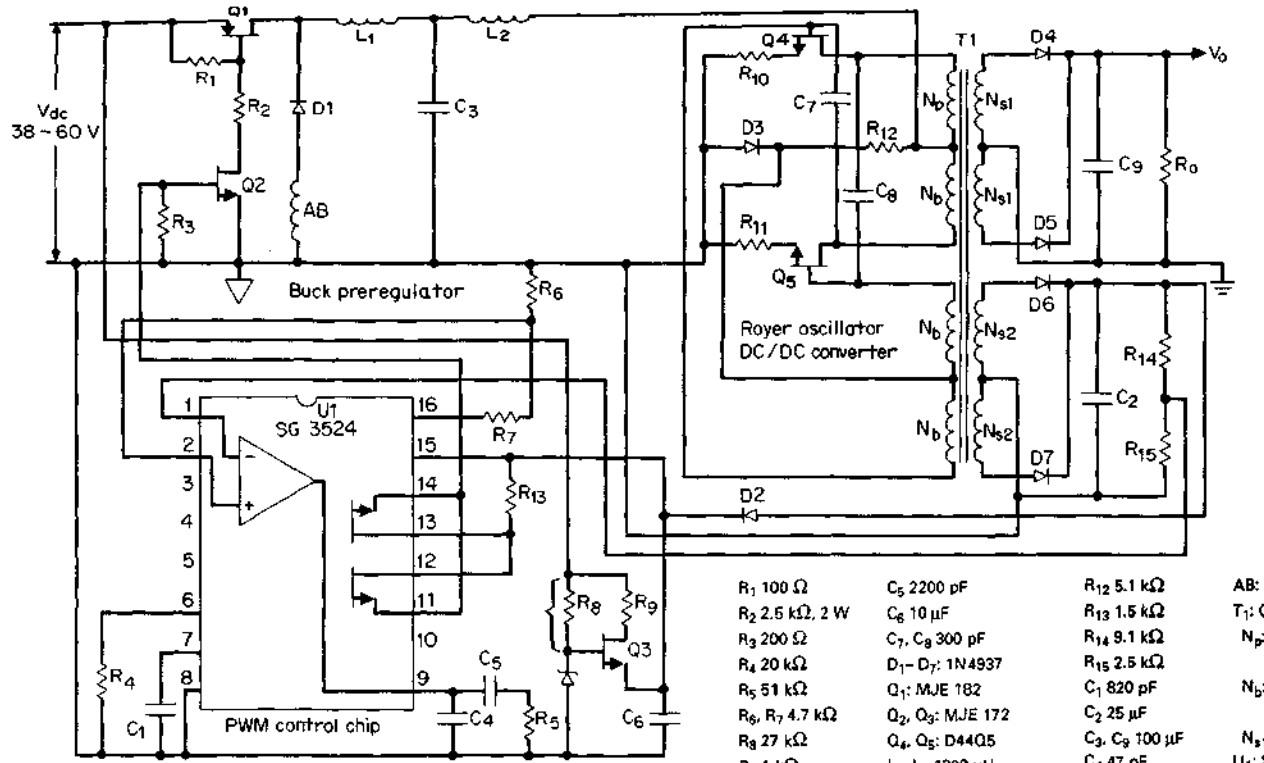
But if better load regulation is desired, the error amplifier in the buck can sense a bootstrapped slave secondary off the main power transformer as described in Sec. 6.6.3.3 and Fig. 6.21.

The circuit details for the data of Fig. 6.27 are shown in Fig. 6.26.

6.6.4.4 Square hysteresis loop materials for Royer oscillators. The transformer core for a Royer oscillator must have a square hysteresis loop. If the loop is not square, turnon flipover from one transistor to the other will be sluggish and in the worst case flipover may not occur.

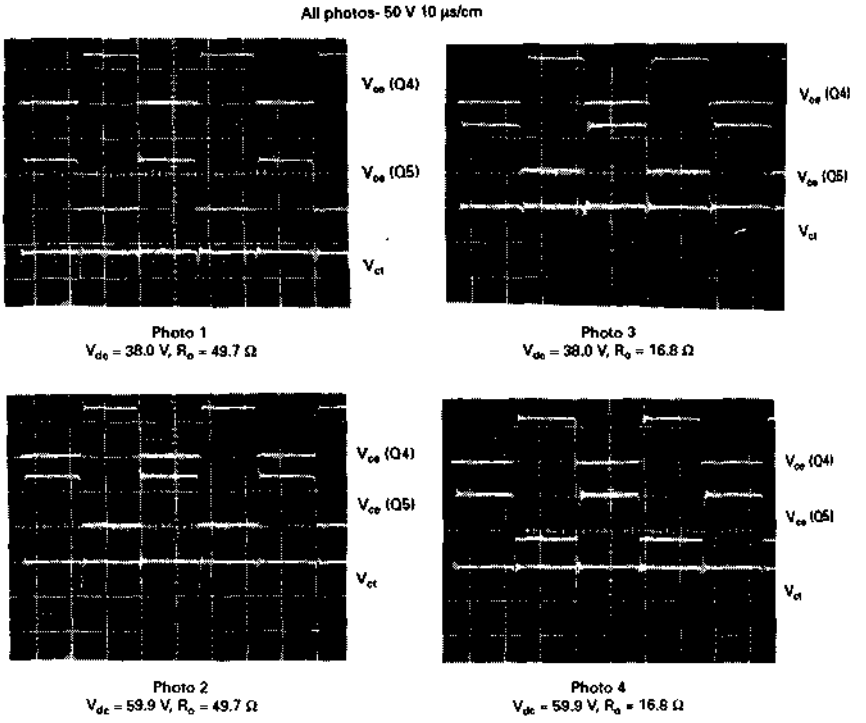
The on transistor may push the core to the top of the hysteresis loop and hang up there, delivering sufficient base drive to keep itself on, yet not turning the opposite transistor on. If this occurs, the partially on transistor will fail in a few tens of microseconds.

Most ferrite core materials do not have a sufficiently square hysteresis loop, but there are various other materials which do. The earliest material was an alloy of 79% nickel, 17% iron, and 4% molybdenum available from a number of manufacturers under various trade names. Magnetics Inc. refers to its material as Square Permalloy 80 and has probably the largest selection of standard core sizes in it. Other manufacturers' trade names for roughly the same material are 4-79 Permalloy, Square Mu 79, and Square Permalloy. The material has a saturation flux density ranging between 6600 and 8200 G.



R_1 100 Ω	C_5 2200 pF	R_{12} 5.1 kΩ	AB: Toshiba "Ammobead"
R_2 2.5 kΩ, 2 W	C_6 10 μF	R_{13} 1.5 kΩ	T_1: Core—Toshiba MB 18 x 12 x 4.5
R_3 200 Ω	C_7, C_8 300 pF	R_{14} 9.1 kΩ	N_p: 54 turns #30
R_4 20 kΩ	D_1–D_7: 1N4937	R_{15} 2.5 kΩ	
R_5 51 kΩ	Q_1: MJE 182	C_1 820 pF	N_b: 4 turns #30
R_6, R_7 4.7 kΩ	Q_2, Q_3: MJE 172	C_2 25 μF	
R_8 27 kΩ	Q_4, Q_5: D44Q5	C_3, C_9 100 μF	N_{s1}, N_{s2}: 18 turns #30
R_9 1 kΩ	L_1, L_2 1300 μH	C_4 47 pF	U_1: SG 3524
R_{10}, R_{11} 15 Ω	Z_1: 1N4103, 9.1 V		

Figure 6.26 A buck regulator driving a current-fed Royer DC/DC converter for constant V_o load. Feedback can be taken in from the buck output mode with the Royer DC/DC-converter-operated open loop. This yields regulation of better than 0.5 percent over input change from 38 to 60 V, but load regulation of only ±5 percent. For better load regulation, feedback is taken from the bootstrapped slave output, which is referenced to input ground as shown.



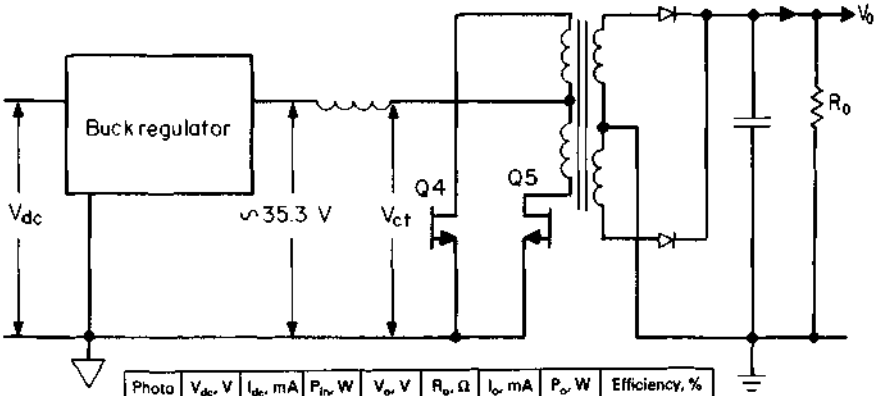


Photo	V_{dc}, V	I_{dc}, mA	P_{in}, W	V_o, V	R_o, Ω	I_o, mA	P_o, W	Efficiency, %
1	38.0	98.0	3.72	10.74	49.7	216.1	2.32	62.4
2	59.9	67.4	4.04	10.78	49.7	216.9	2.34	57.9
3	38.0	216	8.21	9.79	16.8	582.7	5.71	69.5
4	59.9	146	8.75	9.82	16.8	584.5	5.74	65.8

Figure 6.27 Waveforms and data on the current-fed Royer driven from a buck regulator. Feedback is from the buck output. Same circuit as in Fig. 6.26.

The material is produced in a thin ribbon tape, wound on a toroidal core and then encased in either aluminum or some nonmetallic case. The tape is available in either 1- or ½-mil thickness. Core losses increase rapidly with frequency, and just as with power transformers, where higher frequencies require thinner laminations to minimize losses, beyond 50 kHz, the ½-mil tape cores should be used. But beyond 100 kHz, losses in even the ½-mil cores become prohibitive.

In the past 5 years, a new "amorphous" magnetic material with low losses at high frequencies and for flux swings between $+B_s$ and $-B_s$ has become available. It permits building Royer oscillators of up to 200 kHz with acceptably low losses and core temperature rise.

Such amorphous cores are manufactured in the United States under the trade name Metglas by Allied Corporation and Magnetics Inc. and by Toshiba under the trade name Amorphous-MB. These core materials have saturation flux densities ranging between 5700 and 6200 G.

All of these square-loop cores, since they are made of thin tapes, have relatively low iron area compared to ferrite cores. Thus at the same frequency they might be expected to require more turns than would a ferrite core (if a square-loop ferrite core were available).

However, the number of turns for a Royer is no problem with the small-area tape-wound cores as can be seen from Eq. 6.22. There it is seen that the required number of turns is inversely proportional to the saturation flux density, frequency, and iron area. Thus, although iron area of the tape-wound cores is small, their saturation flux density is close to twice that of any available square hysteresis loop ferrite. And since Royers with tape-wound cores can be built at 50 to 200 kHz, where fewer turns are necessary, there is no problem of excessive number of turns with low-iron-area tape-wound cores.

Nevertheless, if it is desired to use a ferrite core because of its larger iron area, there are some few sources for a square hysteresis loop ferrite core. One such is material Type 83 from the Fairite Corporation (Walkill, New York). It has a saturation flux density of 4000 G. But its losses operating at ±4000 G are sufficiently high that the absolute maximum operating frequency is 50 kHz. Available square hysteresis core materials are listed in Table 6.3.

It should be noted that most of these tape-wound cores have relatively small radiating surface area and hence high thermal resistance (in the range of 40 to 100°C/W). Unless bound to some kind of a heat sink, their total losses should be kept under 1 W.

Hysteresis loops for the above materials are shown in Fig. 6.28.

6.6.4.5 Future potential for current-fed Royer and buck preregulated current-fed Royer. It may perhaps seem surprising in a text on modern power supply design to devote much space to the Royer circuit

TABLE 6.3

Core material	Saturation flux density, G	Core losses,* W/cm³	
		50 kHz	100 kHz
Toshiba MB	6000	0.49	1.54
Metglas 2714A	6000	0.62	1.72
Square Permalloy 80 (½ mil)	7800	0.98	2.26
Square Permalloy 80 (1 mil)	7800	4.2	9.6
Fairite Type 83	4000	4.0†	30.0

*For flux excursions between positive and negative saturation.
†1 W/cm³ at 25 kHz.

which had been cast aside 30 years ago. But the Royer, operated in the current-fed mode, with small collector-to-opposite-base flipover capacitors and with the new low-loss amorphous cores is very attractive in many applications.

If line regulation is not required, it is extremely low in parts count (Figs. 6.23, 6.24). Its major fields of application are where prime input is low-voltage DC—48 V for telephone industry supplies, 28 V for aircraft supplies, and 12 or 24 V for automotive supplies.

With the new available cores, they can generate up to 200 of possibly 300 W. Since they require no output inductors, they can easily generate high voltage—either with a multiturn secondary and/or an output voltage multiplier. If regulated output voltage is required, they can be preceded by high efficiency buck regulator (Figs. 6.26, 6.27).

It appears at this writing that in the coming years, there will be widespread renewed interest in the current-fed Royer oscillator DC/DC converter.

6.6.5 Minimum-parts-count flyback as a housekeeping supply

The low-power flyback scheme of Fig. 6.21 as a housekeeping supply is shown detailed in Fig. 6.28.

The circuit was designed as discontinuous-mode flyback from the design relations presented in Chap. 4. It was designed for 6 W of output power at a switching frequency of 50 kHz from a supply voltage of 38 to 60 V—the usual range for a telephone industry power supply. The circuit as is can easily deliver twice the output power without overstressing any components. But beyond 6 W at less than 38 V of input, it will enter the continuous mode and oscillate unless the feedback loop is changed to cope with that (Sec. 4.3).

In Fig. 6.28, the regulated output is a master secondary V_{om} referred to input ground. The housekeeping output V_o is a slave re-

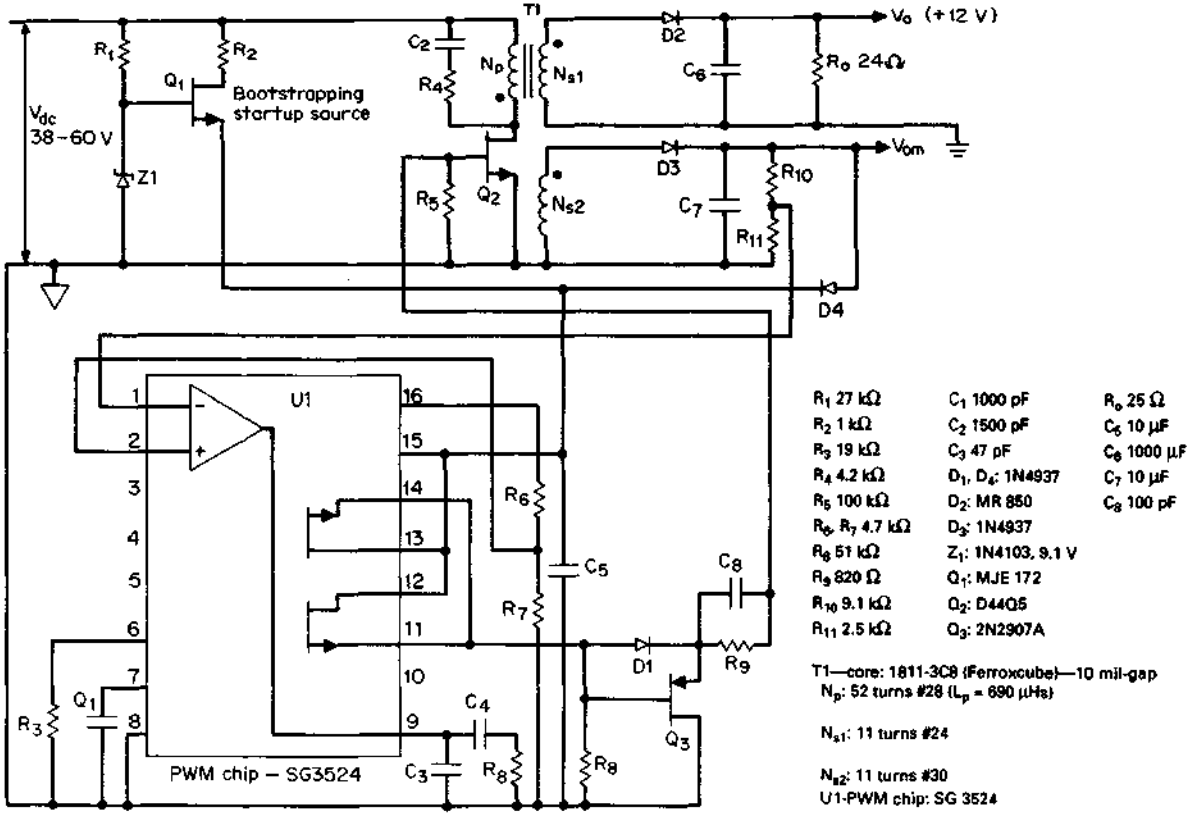


Figure 6.28 A low-power flyback housekeeping power supply with an isolated output V_o.

turned to output ground where it can drive the PWM chip for the main power supply. As discussed in Sec. 4.2.1, flybacks having no output inductors have the useful property that slaves track the master very closely. Thus regulating V_{om} on input ground keeps V_o on output ground sufficiently constant for a housekeeping supply.

In Fig. 6.28, $Q1$ provides supply voltage to the PWM chip at initial turnon. After the supply is up and delivering its output voltages, V_{om} takes over and supplies the chip via diode $D4$. The voltage at the $D4$ cathode is about 11 V. Since the base of $Q1$ is kept at 9.1 V by $Z1$, it is biased off as soon as its emitter rises to about 9 V via $D4$.

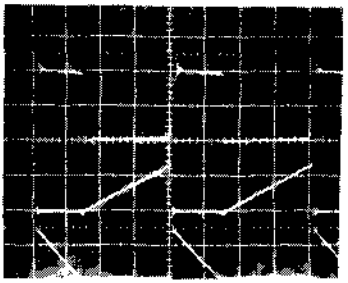
Significant waveforms and performance data for the circuit are shown in Fig. 6.29*a*, 6.29*b*, and 6.29*c* for input voltages of 38, 50, and 60 V, respectively. Efficiencies are seen to be about 70 percent. This is not spectacular, and no effort has been made to optimize the circuit and its efficiency.

It was intended here only to show the significant classical waveforms of an actual operating, discontinuous-mode flyback. Figure 6.29*a*, for a DC input voltage of 38 V, shows the circuit to be just at the end of the discontinuous mode (Sec. 4.4.1 and Fig. 4.8). Primary current is seen to start ramping up with no dead time after the instant the previous secondary current has ramped down to zero. Figure 6.29*b* and 6.29*c* shows the same waveforms at 50 and 60 V, respectively, and it is seen that there is now a dead time from the instant secondary current has ramped down to zero and the start of the next turnon.

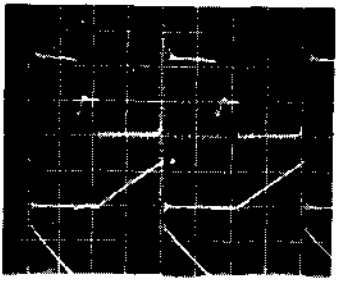
The flyback is seen to have fewer components than the buck-current-fed Royer of Fig. 6.26. But it is seen in Fig. 6.29 that at 6 W of output power, peak secondary current for the flyback is 3 A. This compares to 0.36 A for the buck-current-fed Royer (Fig. 6.25*a*). The higher flyback secondary current would produce a greater RFI problem, requires a larger output filter, and possibly requires a small LC filter after the main filter capacitor to eliminate the output spike at the instant of turnoff due to ESR in the main capacitor (Sec. 4.3.4.1).

6.6.6 Buck regulator with DC-isolated output as a housekeeping supply

Figure 1.9 shows another possible, inexpensive, and low-parts-count scheme for generating a DC-isolated power supply. It is described in Sec. 1.3.8. Care must be taken in this scheme that the current drawn from the secondary is not sufficient to cause the primary current to go into the discontinuous mode, or regulation will suffer.



Note no dead time from end of off time to start of next on time; operation is just at end of discontinuous mode



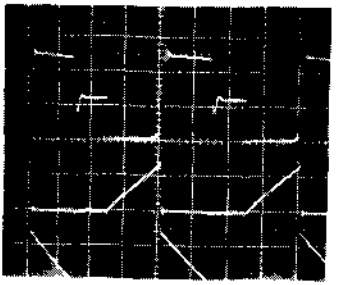


Figure 6.29 Significant waveforms for low-power flyback of Fig. 6.29.

References

1. B. D. Bedford and R. G. Hoft, *Principles of Inverter Circuits*, Wiley, New York, 1964.
2. *General Electric SCR Manual*, 6th ed., General Electric Co., Auburn, NY, 1979.
3. I. Martin, *Operating Characteristics of Self-Commutated Sinewave SCR Inverters*, RCA Application Note AN-6745, RCA, Somerville, NJ, 1978.
4. I. Martin, *Regulating the SCR Inverter Power Supply*, RCA Application Note AN-6856, RCA, Somerville, NJ, 1980.
5. Z. F. Chang, *Application of ASCR in 40 kHz Sine Wave Converter*, RCA Application Note ST-6867, RCA, Somerville, NJ, 1980.

6. N. Mapham, "An SCR Inverter with Good Regulation and Sine Wave Output," *IEEE Transactions on Industry and General Applications*, IGA-3 (2), 1967.
7. N. Mapham, "Low Cost Ultrasonic Frequency Inverter Using Single SCR," *IEEE Transactions on Industry and General Applications*, IGA-3 (5), 1967.
8. I. Martin, "Application of ASCR's to High Frequency Inverters," *Proceedings Powercon 4*, May 1977.
9. D. Chambers, "Designing High Power SCR Resonant Converters for Very High Frequency Operation," *Proceedings Powercon 9*, 1982.
10. D. Chambers, "A 30 kW Series Resonant X Ray Generator," *Powertechnics Magazine*, January 1986.
11. K. Check, "Designing Improved High Frequency DC/DC Converters with a New Resonant Thyristor Technique," Intel Corporation, Hillsboro, OR.
12. See Ref. 1, Chaps. 8, 10.
13. D. Amin, "Applying Sinewave Power Switching Techniques to the Design of High Frequency Off-Line Converters," *Proceedings Powercon 7*, 1980.
14. S. Cuk and D. Middlebrook, *Advances in Switched Mode Power Converters*, Vols. 1, 2, Teslaco, Pasadena, CA, 1981.
15. G. Chryssis, *High Frequency Switching Power Supplies*, 2d ed., McGraw-Hill, New York, 1989.
16. K. Billings, *Switchmode Power Supply Handbook*, McGraw-Hill, New York, 1990.
17. A. Pressman, *Switching and Linear Power Supply, Power Converter Design*, Switchtronix Press, Waban, MA, 1977.
18. G. H. Royer, "A Switching Transistor AC to DC Converter," *AIEE Transactions*, July 1955.
19. D. V. Jones, "A Current Sourced Inverter with Saturating Output Transformer," *IEEE Proceedings*, 1981.

Part

2

Magnetics and Circuits Design

Chapter

7

Transformer and Magnetics Design

7.1 Introduction

In Part 1, characteristics of most of the frequently used topologies were discussed in sufficient depth to permit the choice of a topology most suited to the power supply specifications.

Most frequently, the topology will be selected to minimize the power transistor's off-voltage stress at high line and peak current stress at maximum output power. Other considerations are to minimize parts count, cost, and required volume of the complete supply. Minimizing potential RFI problems is also a frequent factor in choice of topology.

After a topology is selected, the next major decisions are to select an operating frequency and transformer core size which yields the specified maximum output power for the smallest-size transformer.

To make the frequency and transformer core selection, it is necessary to know the numerical relations between available output power and transformer parameters such as iron area, core window or bobbin winding area, peak flux density, operating frequency, and coil current density. In the following sections, equations giving these relations will be derived for most of the frequently used topologies.

These relations can be used in equation form to select a transformer core and operating frequency. But in equation form, they require a tentative guess as to the required core and frequency, a calculation as to the available power from that core for some parameters, and corrected choices of a core and frequency and some of the parameters if the desired power is not available. Since all the parameters are interrelated, such iterative calculations may have to be done several dozen times before a satisfactory combination of the parameters is found.

This cumbersome procedure is avoided by throwing the equations

into a chart form. The charts show frequency increasing (in multiples of 8 kHz to the right) in vertical columns and specific core sizes from various manufacturers in horizontal rows with available output power (calculated from the equations) at the column-row intersections.

The cores are arranged in horizontal rows of increasing output power. Thus at a glance one can choose an operating frequency and move vertically through the rows until the first core of sufficient power is found. Or if a specific core whose dimensions fit the available space is chosen, one can move horizontally through columns of increasing frequency until the frequency which yields the desired output power is found.

Such charts are shown for various core geometries of four major core manufacturers.

Core losses versus frequency and peak flux density for widely used core materials from various core manufacturers are shown. The available ferrite core geometries and their usages are discussed. Core and copper loss calculations are presented. A significant contributor to copper losses—proximity effect—is described. Transformer temperature rise calculation from the sum of core and copper losses is demonstrated.

7.2 Transformer Core Materials and Geometries and Peak Flux Density Selection

7.2.1 Ferrite core losses versus frequency and flux density for widely used core materials[1–4]

Most switching power supply transformers use ferrite cores. Ferrites are ceramic ferromagnetic materials having a crystalline structure consisting of mixtures of iron oxide with either manganese or zinc oxide. Their eddy current losses are negligible as their electrical resistivity is very high. Core losses are then due mainly to hysteresis losses which are so low as to permit use of some materials up to a frequency of 1 MHz.

Ferrite cores are available from about five major American manufacturers (Ferroxcube-Philips, Magnetics Inc., Ceramic Magnetics Inc., Ferrite International, Fairite) and some major overseas manufacturers (TDK, Siemens, Thomson-CSF, Tokin).

Each manufacturer has a number of different mixes of the various oxides processed in various ways to achieve different advantages. Some materials are tailored to yield the minimum core loss point at higher frequency (>100 kHz) or to shift the minimum core loss temperature point to a higher value (90°C) or to achieve minimum core

losses at the most usual combination of high frequency and peak flux density.

However, most vendors' ferrites tailored for power transformer applications are quite similar in their DC hysteresis loops. They are within 10 percent of complete saturation in the region of 3000 to 3200 G at 100°C, have a coercive force of 0.10 to 0.15 Oe at 100°C, and have a residual flux density of 900 to 1200 G at 100°C.

The major factors affecting a material selection are the curves of core loss (usually in milliwatts per cubic centimeter) versus frequency and peak flux density. A typical curve showing these data plus the DC hysteresis loop is shown in Fig. 7.1 for a Ferroxcube-Philips high-frequency material 3F3. A tabulation of core losses for some widely used materials is given in Table 7.1.

Losses in Table 7.1 are taken from core manufacturer data sheets and, although seldom pointed out, are for bipolar magnetic circuits where the flux excursion extends into the first and third quadrants

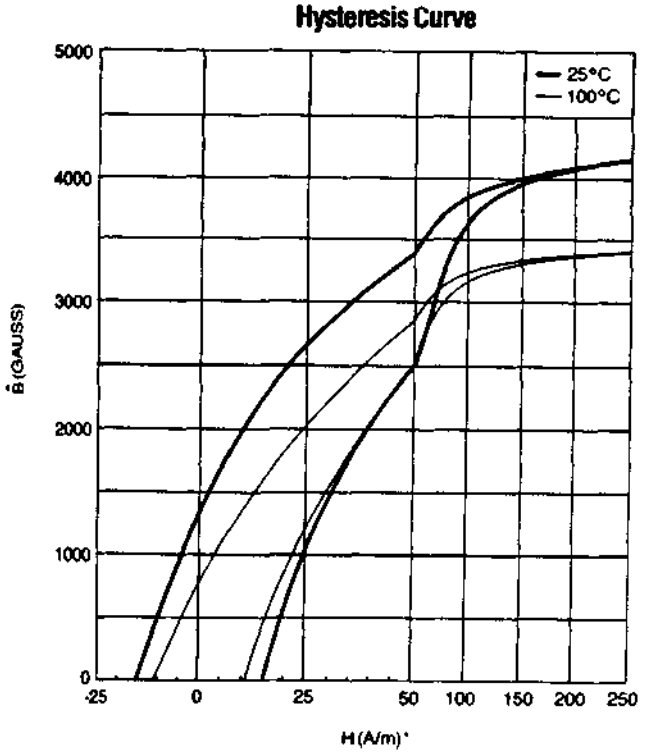


Figure 7.1 Significant characteristics of the 3F3—a high-frequency, low-loss core material. (*Courtesy of Ferroxcube-Philips Corp.*)

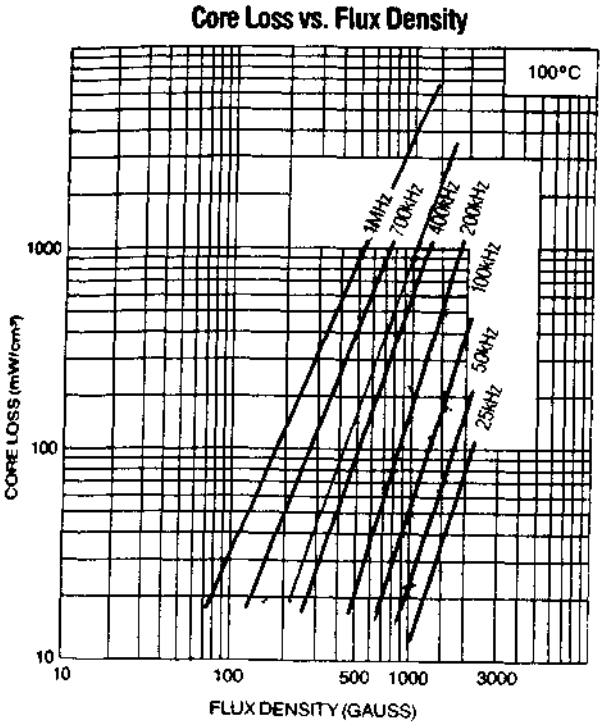


Figure 7.1 *Continued.*

(push-pulls, half and full bridges). But in forward converters and flybacks, operation is over the first quadrant only.

Since ferrite core losses are hysteresis losses only and these losses are proportional to the area of the hysteresis loop, it might be thought that in unipolar magnetic circuits where only half of the hysteresis loop is traversed, core losses would be half those given for bipolar circuits at the same peak flux density.

There is considerable difference of opinion among manufacturers on this. Some say unipolar circuit losses are one-fourth the quoted and measured values for bipolar circuits at the same peak flux density. They base this on the fact that if a unipolar circuit swings from 0 to B_{max} gauss, it is equivalent to a bipolar circuit swinging around a mean value of $B_{max}/2$ with a peak excursion of $B_{max}/2$. Further, since core losses are roughly proportional to the square of the peak flux excursion in a bipolar circuit, halving the peak flux excursion reduces losses by a factor of 4.

TDK offers a curve showing that unipolar circuit losses for a zero to B_{max} excursion are a factor K_{fc} times as great as losses in a bipolar circuit swinging from $-B_{max}$ to $+B_{max}$. The factor K_{fc} is shown as be-

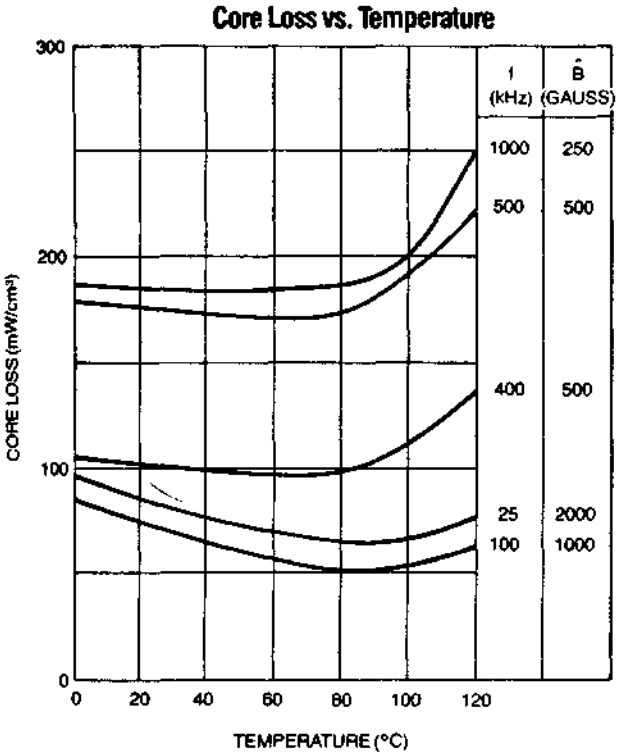


Figure 7.1 *Continued*

ing frequency-dependent. It is 0.35 at 60 kHz, 0.39 at 20 kHz, and 0.34 at 100 kHz.

A conservative approach and accepting the argument that losing half the area of a hysteresis loop (for unipolar circuits) should reduce losses measured for a bipolar circuit by a factor of 2, it will be assumed herein that unipolar circuits at the same peak flux density as listed in Table 7.1 will have losses half those shown in the table.

7.2.2 Ferrite core geometries

Ferrite cores are manufactured in a relatively small number of geometric shapes and varying dimensions within these shapes. Their shapes and dimensions are described in four core manufacturers' catalogs.[1–4]

Many of the core shapes and dimensions in those catalogs are international standards and are available from the various manufacturers in their proprietary core material. Cores which are international standards are listed in publications from the Magnetic Material Producers

TABLE 7.1 Core Losses for Various Core Materials at Various Frequencies and Peak Flux Densities at 100°C

Frequency, kHz	Material	Core loss, mW/cm³ for various peak flux densities, G					
		1600	1400	1200	1000	800	600
20	Ferroxcube 3C8	85	60	40	25	15	
	Ferroxcube 3C85	82	25	18	13	10	
	Ferroxcube 3F3	28	20	12	9	5	
	Magnetics Inc.-R	20	12	7	5	3	
	Magnetics Inc.-P	40	18	13	8	5	
	TDK-H7C1	60	40	30	20	10	
	TDK-H7C4	45	29	18	10		
	Siemens N27	50			24		
50	Ferroxcube 3C8	270	190	130	80	47	22
	Ferroxcube 3C85	80	65	40	30	18	9
	Ferroxcube 3F3	70	50	30	22	12	5
	Magnetics Inc.-R	75	55	28	20	11	5
	Magnetics Inc.-P	147	85	57	40	20	9
	TDK-H7C1	160	90	60	45	25	20
	TDK-H7C4	100	65	40	28	20	
	Siemens N27	144			96		
100	Ferroxcube 3C8	850	600	400	250	140	65
	Ferroxcube 3C85	260	160	100	80	48	30
	Ferroxcube 3F3	180	120	70	55	30	14
	Magnetics Inc.-R	250	150	85	70	35	16
	Magnetics Inc.-P	340	181	136	96	57	23
	TDK-H7C1	500	300	200	140	75	35
	TDK-H7C4	300	180	100	70	50	
	Siemens-N27	480			200		
	Siemens-N47				190		
200	Ferroxcube 3C8				700	400	190
	Ferroxcube 3C85	700	500	350	300	180	75
	Ferroxcube 3F3	600	360	250	180	85	40
	Magnetics Inc.-R	650	450	280	200	100	45
	Magnetics Inc.-P	850	567	340	227	136	68
	TDK-H7C1	1400	900	500	400	200	100
	TDK-H7C4	800	500	300	200	100	45
	Siemens-N27	960			480		
	Siemens-N47				480		
500	Ferroxcube 3C85				1800	950	500
	Ferroxcube 3F3		1800	1200	900	500	280
	Magnetics Inc.-R		2200	1300	1100	700	400
	Magnetics Inc.-P		4500	3200	1800	1100	570
	TDK-H7F						100
	TDK-H7C4		2800	1800	1200	980	320
1000	Ferroxcube 3C85						2000
	Ferroxcube 3F3				3500	2500	1200
	Magnetics Inc.-R				5000	3000	1500
	Magnetics Inc.-P						6200

Note: Data are for bipolar magnetic circuits (first- and third-quadrant operation). For unipolar circuits (forward converter, flyback), divide by 2.

Association (MMPA)[5,6] and IEC publications from the American National Standards Institute.[7]

The various core geometries shown in Fig. 7.2 are pot or cup cores, RM cores, EE cores, PQ cores, UU or UI cores. The pot core is shown in Fig. 7.2*e*. It is used mostly at low power levels up to 125 W and usually in DC/DC converters. Its major advantage is that the coil on the bobbin around the center post is almost entirely enclosed by ferrite material. This decreases its radiating magnetic field and hence is used when EMI or RFI problems must be minimized.

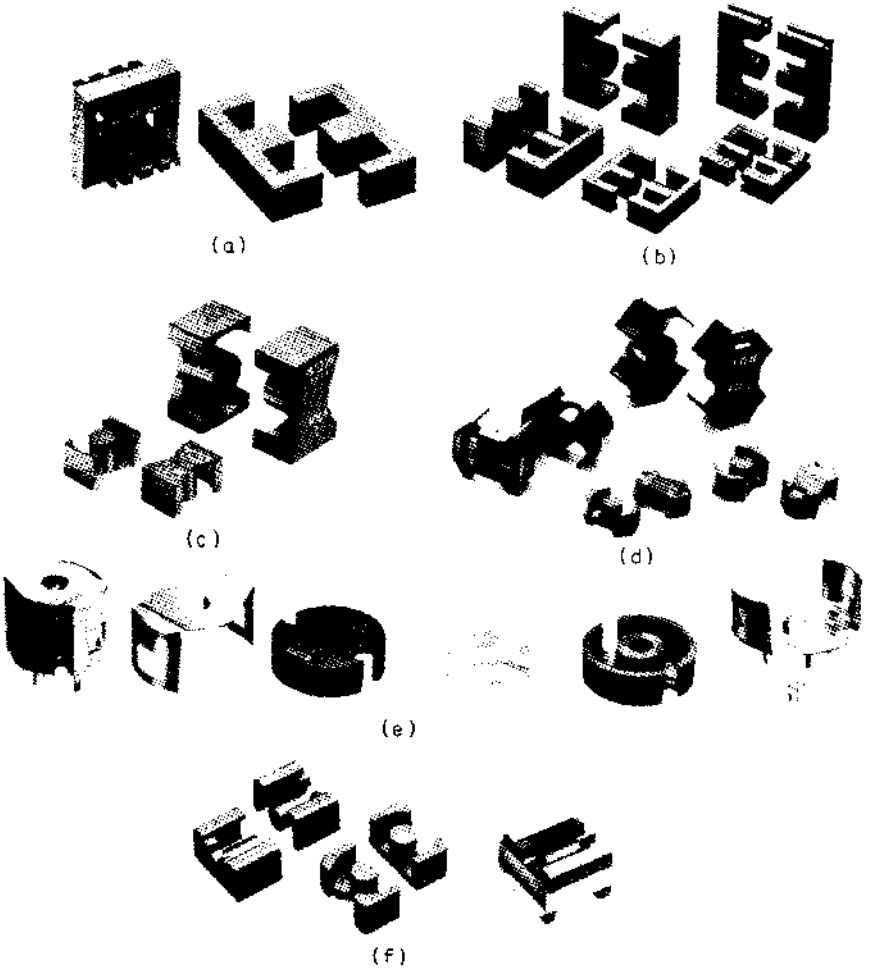


Figure 7.2 Various core geometries for power transformers: (*a*) EE cores; (*b*) EC and ETD cores; (*c*) PQ cores; (*d*) RM cores; (*e*) pot cores; (*f*) LP cores. (*Courtesy of TDK Corp.*)

The major disadvantage of the pot core is the narrow slot in the ferrite through which the coils leads exit. This makes it difficult to use it at high input or output current levels where the wire diameter is large or in multioutput supplies with many wires exiting.

It is also not a good choice for a high voltage supply, even at low power. Leads carrying a high voltage may arc because of the close spacing in the narrow exit notch in the ferrite.

Many of the pot cores are available with gaps of various sizes in the center leg so that they may carry a DC bias current without saturating. This permits their use as output inductors in buck regulators (Sec. 1.3.6), forward converters (Secs. 2.3.9.2, 2.3.9.3), push-pulls (Sec. 2.2.8.1), and flybacks (Sec. 4.3.3.1).

Most often, when a core is available with a gapped center leg, the manufacturer gives its A_l value (inductance in millihenries per 1000 turns) and the cliff point in ampere turns at which it falls over its saturation cliff.

If a gapped core is required, it is more cost effective—and from a performance viewpoint, preferable—to use a core with a gapped center leg rather than using ungapped core halves separated with the proper thickness of plastic shims. Shimming core halves will not yield reproducible A_l values over time, temperature, and production spread. Also, shimming the outer leg will increase EMI.

The most widely used cores are EE cores (Fig. 7.2*a*) because there is no narrow ferrite notch restricting coil leads entering or leaving the bobbin. But since the coil is not as fully surrounded by ferrite, it does produce a larger EMI-RFI field. But because the coil is not completely surrounded by ferrite, airflow past it is unimpeded and it tends to run cooler. EE cores are available with either a square or round center leg. Round-center-leg cores (EC or ETD types; Fig. 7.2*b*) have a small advantage in that the mean length of a turn is about 11 percent shorter than for a square-legged core of equal center-leg area. Coil resistance is thus about 11 percent less for equal numbers of turns, and copper loss and temperature rise is somewhat less for the round center leg.

There is a large range of EE core sizes, and depending on frequency and peak flux density, they can deliver output powers from under 5 W up to 5 or possibly 10 kW. By using two square-center-leg EE cores side by side, the core area is doubled, requiring half the number of turns for a fixed primary voltage, peak flux density, and frequency (Faraday's law; Eq. 2.7). This doubles the available power from a single core and may result in a smaller transformer than using a single core of the next-larger size.

The RM or "Square" core, a compromise between a pot and an EE core, is shown in Fig. 7.2*d*. It is effectively a pot core with a much

wider notch cut out of the ferrite. It thus makes it easier to bring larger diameter or many wires in and out of the coil, and hence this core is usable for much higher output power levels and for multi-output transformers. The larger ferrite notch also provides easier access for convection air currents than in a pot core and results in lower temperature rise than that for the pot.

Because the coil is not as fully surrounded by ferrite as in a pot core, it causes more EMI-RFI radiation than does the pot core. Yet, because the ferrite surrounding the coil is still greater than in the case of an EE core, its EMI radiation is less than that for an EE of equal output power.

RM cores are available with or without a center-leg hole. The center-leg hole is used for mounting with a bolt through the hole or in frequency-sensitive applications. By inserting an adjustable ferrite "tuning rod" into the center hole, the A_l value may be adjusted up to as much as 30 percent. Although this tuning feature is not usable in power transformers because of increased energy losses, it is usable for frequency-sensitive filters.

The geometry of the PQ core (Fig. 7.2*c*; Magnetics Inc. and TDK) is such that it provides an optimum ratio of volume to radiating surface and coil winding area. Since core losses are proportional to core volume, and heat radiation capability is proportional to radiating surface area, these cores have a minimized temperature rise for a given output power. Further, since the PQ core volume-to-coil-winding-area ratio is optimized, the volume is minimized for a given output power.

LP cores (Fig. 7.2*e*; TDK) are specifically designed for low-profile transformers. They have long center legs which minimize leakage inductance.

UU or UI cores (shown in Fig. 7.2*f*) are used mainly for high-voltage or ultra-high-power applications. They are rarely used at power levels under 1 kW. Their large window area compared to an EE core of equal core area permits much larger wire sizes or many more turns. But their much larger magnetic path length does not yield as close primary-secondary coupling as in an EE core and results in larger leakage reactances.

7.2.3 Peak flux density selection

As discussed in Sec. 2.2.9.3, the number transformer primary turns will be selected from Faraday's law (Eq. 1.17, 2.7) for a preselected peak flux density B_{max}. It is seen in Eq. 1.17, that the larger the flux excursion (larger value of B_{max}), the fewer the primary turns and hence the larger the permissible wire size, and therefore the greater the available output power.

There are two limitations to peak flux density in the ferrite cores. The first is core losses and hence increased core temperature rise. Core losses in most ferrite materials increase as the 2.7th power of the peak flux density, and thus high peak flux densities cannot be permitted, especially at higher frequencies. But most ferrites—even the lossiest ones—have such low losses at 25 kHz and below, that core losses are not a limiting factor at those frequencies (see Table 7.1). Then at these low frequencies, peak flux excursion may possibly extend far up the hysteresis loop into the curved area of the *BH* loop. But care must be taken that the core does not move so far into saturation that the primary cannot support the applied voltage, or the power transistor will be destroyed.

Ferrite core losses also increase roughly as the 1.7th power of the switching frequency. Thus at higher frequencies, for the more lossy materials (Table 7.1), attempting a high peak flux density to minimize the number of turns results in such high losses (core losses of course are equal to the loss factor in milliwatts per cubic centimeter times the core volume in cubic centimeters) that temperature rise will be excessive.

Hence at 50 kHz and above, the less lossy (somewhat more expensive) core material must be used or peak flux density must be reduced. Reducing the peak flux density, of course, requires increasing the number of primary turns (Eq. 2.7) and hence requires smaller wire size for the same core–bobbin winding area. With smaller wire size, primary and secondary currents are smaller and available output power is decreased.

Thus at high frequencies ($> \sim 50$ kHz) the least lossy core material will be used and peak flux density will have to be reduced to such a value that total core and copper losses result in an acceptably low temperature rise. Temperature rise calculation for the sum of core plus coil losses will be demonstrated in Sec. 7.4.

Even at low frequencies, however, where core losses are not a limiting factor, peak flux density cannot be permitted to move very high up on the *BH* loop to minimize the number of primary turns.

It is seen in Fig. 2.3 that the *BH* loop is still linear up to 2000 G (this is closely the end of the linear portion for most ferrite materials). Going beyond this will increase the magnetizing current near the end of the transistor on time and will unnecessarily increase coil losses. But designing (choosing N_p in Eq. 2.7) for a maximum of 2000 G for most ferrites is risky. For fast transient line or load steps, if the feedback error amplifier is not fast enough for a few switching cycles, peak flux density may move up to hard saturation (>3200 G at 100°C) and destroy the power transistor. This is discussed in more detail in Sec. 2.2.9.4.

Thus, for ferrites even below 50 kHz where core loss is not a limiting factor, in all following designs herein, peak flux density will be

chosen at 1600 G. At higher frequencies, where for 1600 G even the least lossy material results in excessive power losses, peak flux density will have to be reduced. The "available output power tables" to be developed below will show how easily output power may be calculated for these reduced peak flux densities.

7.3 Maximum Transformer Core Output Power, Peak Flux Density, Core and Bobbin Areas, and Coil Current Density

7.3.1 Derivation of output power relations for forward converter topology

Refer to Fig. 7.3 for the forward converter topology. The following output power relation will be based on the following assumptions:

1. Efficiency of the *power train*—from V_{dc} to the sum of all outputs and neglecting control circuit dissipation—is 80 percent.

2. The *space factor SF,* the fraction of total bobbin winding area occupied by the coil, is 0.4. This includes primary and all secondaries plus layer insulations plus any RFI or Faraday shields. This is a usual value (usually in the range 0.4 to 0.6) in transformer design because of the many factors contributing to waste of the total bobbin winding area. One significant factor is that turns within a coil layer often are widely spaced to make all bobbin layers of equal width to improve magnetic coupling between layers and reduce leakage inductances. Also, adherence to European safety specifications (VDE) requires leaving a 4-mm space from each end of a layer to the end of the bobbin. There is also the thickness of the insulation layers themselves.

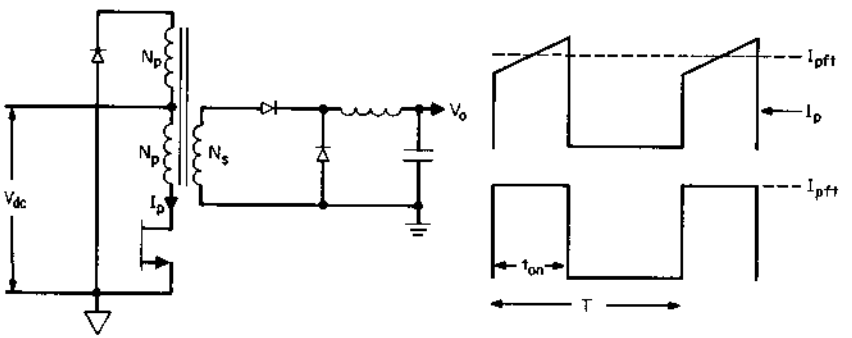


Figure 7.3 Forward converter topology and primary current waveshape I_p. Equivalent flat-topped current waveshape I_{pft} is used to calculate output power versus B_{max}, frequency, A_e, A_b, and D_{cmg} relation. Turns ratio N_s/N_p is chosen to yield $t_{on} = 0.8T/2$ at minimum V_{dc} for specified V_o.

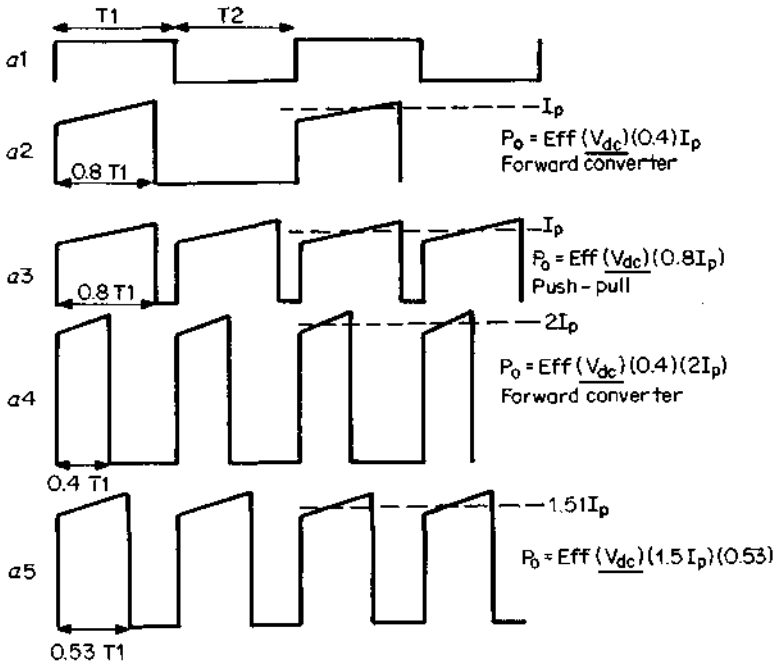


Figure 7.3a (see Section 7.3.2.2)

VDE specifications generally require three layers of 1-mil-thick insulating material; if secondaries are to be sandwiched in between halves of the primary (as usually done to reduce proximity-effect copper losses), that wastes 6 mils of bobbin height. Finally, there is the practical problem that it is difficult to safely assemble the core and bobbin if the full bobbin height is fully utilized.

3. Primary current waveshape is as shown in Fig. 7.3. At minimum V_{dc} input, on time is a maximum at $0.8T/2$ (Secs. 2.3.2, 2.3.5). Primary current has the ramp-on-a-step waveshape because of inductors at each secondary output. The ramp swings ±10 percent about the value at its center I_{pft}. The primary current waveshape can be accurately approximated by a rectangular pulse of peak amplitude I_{pft} of duty cycle $0.8T/2T$ or 0.4 (Sec. 2.3.5). Then for minimum $V_{dc} = V_{dc(min)}$

$$\begin{aligned} P_o &= 0.8P_{in} = 0.8V_{dc(min)}[I_{av} \text{ at } V_{dc(min)}] \\ &= 0.8V_{dc(min)}(I_{pft})\frac{0.8T}{2T} \\ &= 0.32V_{dc(min)}I_{pft} \end{aligned} \tag{7.1}$$

But the rms value of a rectangular waveform of amplitude I_{pft}, duty cycle of 0.4, is $I_{rms} = I_{pft}\sqrt{0.4}$ or $I_{pft} = 1.58I_{rms}$. Then

$$P_o = 0.32V_{dc(min)}(1.58I_{rms})$$
$$= 0.506V_{dc(min)}I_{rms} \tag{7.2}$$

From Faraday's law

$$V_p = N_pA_e\frac{\Delta B}{\Delta T} \times 10^{-8}$$

where V_p = primary volts ($\approx V_{dc}$)
N_p = number of primary turns
A_e = core area, cm^2
ΔB = flux change, G (0 to B_{max})
ΔT = time, s for this flux change = $0.4T$

At $V_{dc(min)}$, $\Delta B/\Delta T = B_{max}/0.4T$. Then for $f = 1/T$ from Eq. 7.2

$$P_o = \frac{0.506I_{rms}N_pA_eB_{max}f}{0.4} \times 10^{-8}$$
$$= 1.265N_pB_{max}A_ef \times 10^{-8}(I_{rms}) \tag{7.3}$$

Now assume that primary and all secondaries operate at the same current density D_{cma} in circular mils per rms ampere. Bobbin area occupied by the reset winding is assumed negligible as it carries only magnetizing current and its wire size is usually smaller than No. 30 (AWG).

Let A_b = bobbin winding area, in^2
A_p = primary winding area, in^2
A_s = secondary winding area (total of all secondaries), in^2
A_{ti} = area of one turn of primary power winding, in^2

Then for a space factor SF of 0.4 and $A_p = A_s$

$$A_p = 0.20A_b = N_pA_{ti}$$

or
$$A_{ti} = \frac{0.2A_b}{N_p} \tag{7.4}$$

Now current density D_{cma} in circular mils per rms ampere is

$$D_{cma} = \frac{A_{tcm}}{I_{rms}}$$

where A_{tcm} is primary wire area in circular mils. Then

$$I_{rms} = \frac{A_{tcm}}{D_{cma}} \tag{7.5}$$

But area in square inches equals area in circular mils times $(\pi/4)\ 10^{-6}$. Then

$$A_{tcm} = \frac{4A_{ti} \times 10^{+6}}{\pi} = \frac{4(0.20A_b)10^{+6}}{\pi N_p}$$

Then from Eq. 7.5

$$I_{rms} = \frac{0.8A_b \times 10^{+6}}{\pi N_p D_{cma}} \tag{7.6}$$

and putting Eq. 7.6 into Eq. 7.3

$$P_o = (1.265N_pB_{max}A_ef \times 10^{-8}) \frac{0.8A_b \times 10^{+6}}{\pi N_p D_{cma}}$$

$$= \frac{0.00322B_{max}fA_eA_b}{D_{cma}}$$

But A_b is in square inches. If it is to be expressed in square centimeters, divide by 6.45. Then

$$P_o = \frac{0.00050B_{max}fA_eA_b}{D_{cma}} \tag{7.7}$$

where P_o is in watts for B_{max} in gauss, A_e and A_b are in square centimeters, f is in hertz, and D_{cma} in circular mils per rms ampere.

7.3.2 Derivation of output power relations for push-pull topology

For the same assumptions as in Sec. 7.3.1, $P_o = 0.8P_{in} = 0.8V_{dc(min)}$ $[I_{av}$ at $V_{dc(min)}]$. In a push-pull, at $V_{dc(min)}$, each transistor is on a maximum of $0.8T/2$ within its half period. For two such pulses per period, the total duty cycle of current drawn from $V_{dc(min)}$ is 0.8, and during its on time, each transistor and half primary carries an equivalent flat-topped current pulse of amplitude I_{pft}. Then

$$P_o = 0.8V_{dc(min)}(0.8I_{pft})$$

$$= 0.64V_{dc(min)}I_{pft} \tag{7.8}$$

But each half primary carries current at a duty cycle of only 0.4, so rms current in each half primary is $I_{\text{rms}} = I_{\text{pft}} \sqrt{0.4}$ or $I_{\text{pft}} = 1.58 I_{\text{rms}}$. Then

$$P_o = 0.64 V_{\text{dc(min)}}(1.58 I_{\text{rms}})$$
$$= 1.01 V_{\text{dc(min)}} I_{\text{rms}} \tag{7.9}$$

Again for $SF = 0.4$, half the total coil area devoted to the primary and half to the secondary and each winding operating at a current density of D_{cma} circular mils per rms ampere, we obtain

$$A_p = 0.20 A_b = 2 N_p A_{\text{ti}}$$

or
$$A_{\text{ti}} = \frac{0.1 A_b}{N_p} \tag{7.10}$$

where N_p = number of turns per half primary
A_b = total bobbin winding area, in^2
A_{ti} = area of a single turn of primary wire, in^2

Now
$$D_{\text{cma}} = \frac{A_{\text{tcm}}}{I_{\text{rms}}}$$

where A_{tcm} = wire area in circular mils and I_{rms} = rms current per half primary

$$I_{\text{rms}} = \frac{A_{\text{tcm}}}{D_{\text{cma}}} \tag{7.11}$$

Now $A_{\text{ti}} = A_{\text{tcm}}(\pi/4)10^{-6}$. Putting this into Eq. 7.10, we obtain

$$A_{\text{tcm}} = 0.1273 \frac{A_b}{N_p} 10^{+6}$$

and putting this into Eq. 7.11

$$I_{\text{rms}} = 0.1273 \frac{A_b}{N_p D_{\text{cma}}} 10^{+6}$$

and putting this into Eq. 7.9

$$P_o = 1.01 V_{\text{dc(min)}} \frac{0.1273 A_b}{N_p D_{\text{cma}}} 10^{+6}$$
$$= 0.129 \frac{V_{\text{dc(min)}} A_b}{N_p D_{\text{cma}}} 10^{+6} \tag{7.12}$$

and again from Faraday's law

$$V_{\text{primary min}} \approx V_{\text{dc(min)}} = \left(\frac{N_p A_e \, \Delta B}{\Delta T}\right) 10^{-8}$$

where here in a push-pull, flux swing is $2B_{\max}$ in a time $0.4T$ at $V_{\text{dc(min)}}$. Put this into Eq. 7.12:

$$P_o = 0.129(N_p A_e) \frac{2B_{\max}}{0.4T} \frac{A_b}{N_p D_{\text{cma}}} 10^{-2}$$

$$= \frac{0.00645 B_{\max} f A_e A_b}{D_{\text{cma}}}$$

Again, A_b is in square inches. If it is expressed in square centimeters, divide this last relation by 6.45. Then

$$P_o = \frac{0.0010 B_{\max} f A_e A_b}{D_{\text{cma}}} \tag{7.13}$$

where P_o is in watts for $B_{\max}$ in gauss, A_e and A_b are in square centimeters, f is in hertz, and D_{cma} is in circular mils per rms ampere.

This result—specifically, that the available power from a given core in a push-pull topology is twice that for the same core in a forward converter topology—could have been quite easily foreseen.

In the push-pull, each transformer half must sustain the same voltage as a forward converter fed from the same supply voltage. But in the push-pull, the available flux change is $2B_{\max}$ as compared to $B_{\max}$ in the forward converter. Thus from Faraday's law, the number of turns on a half primary of the push-pull is half that in the forward converter for the same $B_{\max}$. But since there are two half secondaries in the push-pull, the total number of turns in the push-pull is equal to the number of turns in the forward converter for equal V_{dc} and $B_{\max}$ in both circuits (neglecting the insignificant space occupied by the reset winding in the forward converter).

In the push-pull, however, half the total output power is delivered through each half secondary. Thus, for equal output powers in a push-pull and a forward converter, the peak and rms currents in each push-pull transformer half is half that in the forward converter (compare Eqs. 2.11 to 2.28 and 2.9 to 2.41). Thus the required circular-mil wire area and hence the required bobbin winding space for each half push-pull are half that for a forward converter of equal output power. Then for equal bobbin winding space, the push-pull can deliver twice the power of a forward converter from the same core, as indicated by Eqs. 7.17 and 7.13.

7.3.2.1 Core and copper losses in push-pull, forward converter topologies. Comparing Eqs. 7.7 and 7.13, it is seen that the push-pull topology can yield twice the output power of a forward converter using the same sized (same $A_e A_b$ product) core.

There is a slight penalty paid in that the push-pull transformer, at twice the forward converter output power, will run warmer than the forward converter. Doubling a forward converter output power by going to a push-pull does double the core losses, but copper losses remain unchanged. This can be seen as follows: In the push-pull (Fig. 7.3*a*3), each transformer half must sustain the same voltage as the primary of the forward converter fed from the same supply voltage. In the forward converter, the flux density change is zero to some preselected value B_{max} in a time $0.8T1$ (Fig. 7.3*a*2). In the push-pull, the flux density change is from $-B_{max}$ to $+B_{max}$ or $2B_{max}$ in the same time $0.8T1$.

In Faraday's law, the number of turns is directly proportional to the applied voltage and inversely proportional to the flux density change. Thus, the number of turns in each half primary of the push-pull is half that of the forward converter primary for equal peak-flux densities.

For the push-pull to have twice the output power of the original forward converter, the peak current in each half period must be equal to that of the forward converter. Since the duty cycle of current flow in each half primary of the push-pull is equal to that of the forward converter of half the output power, rms current in each push-pull half primary is equal to that of the forward converter primary.

Thus, the forward converter and each push-pull primary half will use equal wire size (equal number of circular mils) since the wire size will be selected on the basis of 500 circular mils per rms ampere, and since the number of half primary turns is half that of the full primary in the forward converter of half the output power, the two half primaries of the push-pull occupy a volume equal to that of the forward converter.

The push-pull half primary has half the turns and equal rms current compared to the forward converter primary, and hence has half the I^2R copper loss of the full-forward converter primary. Thus, total copper loss for the two half-primaries of the push-pull equals that of the forward converter of half the output power.

But, core losses for the push-pull will be twice that for the forward converter. For in the push-pull, the flux-density change is from $-B_{max}$ to $+B_{max}$ and only form zero to B_{max} in the forward converter. But core losses are proportional to the area of the hysteresis loop transversed.

Thus, the push-pull operating over the first and third quadrant of the hysteresis loop will have twice the losses of the forward converter

that operates only over the first quadrant of the hysteresis loop (see Sec. 7.2.1). But with the newer low-loss core material, the increased core losses will not become the limiting factor in transformer temperature rise—certainly at frequencies below about 100 kHz.

7.3.2.2 Doubling output power from a given core without resorting to a push-pull topology. In the previous section, it is pointed out that for the same core, a push-pull topology can yield twice the output power of a forward converter. The only penalty paid is a doubling of the core losses ; copper losses remain unchanged.

But, there may be a reluctance to going to a push-pull. For push-pull with its two transistors has the added cost and space drawbacks. Also, the push-pull transformer is harder and, hence, more expensive to wind as 3 wires must be taken out of each winding as compared to 2 for a forward converter. Also, there is the always present possibility of flux imbalance (Sec. 2.25) if current-mode version of push-pull is not used.

If it is desired to double the output power of a given forward converter without increasing its core size, the following is a possible alternative to resorting to a push-pull (see Fig. 7.3*a*).

The push-pull yields twice the output of a forward converter from a given core because there are two current pulses per period rather than one for the forward converter. Thus, a possible alternative to the push-pull is to retain the forward converter topology, and have it give one pulse (whose amplitude is twice that of the push-pull) per half period of the originally considered push-pull. This can be seen in Fig. 7.3*a*4.

This, of course, is another way of saying that the forward converter frequency is doubled (forward converter frequency is defined as the inverse of the time required for a complete traversal of the hysteresis loop back to its original starting point). Equation 7.7 states the available output power from a core is directly proportional to frequency.

Thus, going to twice the original forward converter frequency would unquestionably yield twice the output power from the core. But it would result in somewhat more than twice (close to three times) the core losses. For core losses are roughly proportional to the 1.7th power of the frequency (Table 7.1 and Sec. 7.2.3).

Copper losses in a forward converter at twice the frequency and peak primary current (and hence twice the output power) would remain unchanged. This will be discussed below.

Thus, if doubling the frequency and peak primary-current amplitude of a forward converter doubles its output power and costs only a threefold increase in core losses, this may be a credible alternative to going to push-pull at the original frequency. For the push-pull at the

original frequency also achieves a doubled output power but at a cost of only doubling the core losses.

Doubling the forward converter frequency and its peak current amplitude is practical only at original frequencies below 50 to 80 kHz. For, if it is attempted at higher original frequencies (say 100 kHz), the increased AC switching and snubber losses (see Chapter 11) at the doubled frequency would result in poor efficiency.

In Fig. 7.3*a*4, the original forward converter frequency and peak primary current were doubled to achieve a doubling of output power. But, this has its own drawbacks. First, the higher peak current will cause more severe RFI problems, and it may force the selection of a higher current, more expensive transistor.

If the doubled peak primary current is unnattractive, an alternative might be to increase the maximum transistor on-time. In Fig. 7.3*a*4, it was chosen as $0.8T1/2$ so as to ensure that the transformer core can be fully reset with an equal reset time and a guaranteed "dead" time before the start of the next turnon (Section 2.3.2). This is achieved by setting the ratio N_r/N_p (transformer reset winding to power winding turns ratio) equal to 1.0. (See Fig. 7.3).

In Sec. 2.3.8, it was shown that making N_r/N_p less than 1.0 permits a larger on-time, though at a cost of a larger reset voltage and, hence, higher transistor off-voltage stress. Thus, in Fig. 7.3*a*5, an N_r/N_p ratio of 0.5 was chosen. This yielded a larger maximum on-time of $0.53T1$ and a peak current of $1.51I_p$ as compared to $0.4T1$ and a peak current of $2I_p$ when N_r/N_p is 1.0. But the penalty is that now the peak transistor off-voltage stress is $3V_{dc}$ rather than $2V_{dc}$ for the case of $N_r/N_p = 1.0$.

Thus, as discussed in Sec. 2.3.8, setting the N_r/N_p ratio less than 1.0 decreases transistor peak current stress, but increases its peak voltage stress. Ratios of N_r/N_p less than 0.5 generally lead to unacceptably high off-voltage stress.

As stated above, going to twice the forward converter frequency and peak transistor current doubles the output power but does not increase copper losses. This can be seen below as follows:

Figure No.	Frequency	I_p	"On" duty cycle	I_{rms}	N	Wire area	Wire resistance	$(I_{rms})^2R$	P_0
7.3*a*2	F1	I_p	0.4	$0.632I_p$	N	$A1$	$R1$	$(I_{rms})^2R$	P_0
7.3*a*4	2F1	$2I_p$	0.4	$1.264I_p$	$0.5N$	$2A1$	$0.25R1$	$(I_{rms})^2R$	$2P_0$

Since the doubled-frequency forward converter has twice the rms current, it will have twice the wire area of the original converter of

half the output power. And since it has half the number of primary turns, its resistance is one-fourth the resistance of the original forward converter. With twice the rms current, its I^2R losses are equal to that of the original forward converter of half the output power.

7.3.3 Derivation of output power relations for half-bridge topology

The half bridge is shown in Fig. 3.1. Again assume that minimum DC input voltage is $V_{dc(min)}$. Maximum on time per transistor is $0.8T/2$ and occurs at $V_{dc(min)}$. Thus

Efficiency = 80%
A_e, A_b = core, bobbin area, cm^2
A_{bi} = bobbin area, in^2
A_p = primary area, in^2
SF = 0.4, primary and total secondary areas equal
D_{cma} = current density, circular mils/rms A (all windings operate at same current density)
A_{ti} = wire area, in^2
A_{tcm} = wire area, circular mils
N_p = number of primary turns
I_{pft} = equivalent flat-topped primary current pulse

As I_{pft} comes at a duty cycle of 0.8, its rms value is

$$I_{rms} = I_{pft}\sqrt{0.8} = 0.894 I_{pft} \qquad \text{or} \qquad I_{pft} = 1.12 I_{rms}$$

Then
$$P_o = 0.8P_{in} = 0.8\frac{V_{dc(min)}}{2}[I \text{ average at } V_{dc(min)}]$$
$$= 0.4V_{dc(min)}0.8I_{pft}$$
$$= 0.32V_{dc(min)}I_{pft}$$
$$= 0.358V_{dc(min)}I_{rms} \tag{7.14}$$

and
$$A_p = 0.2A_{bi} = N_pA_{ti}$$
$$A_{ti} = \frac{0.2A_{bi}}{N_p}$$

But $A_{ti} = A_{tcm}(\pi/4)10^{-6}$. Then

$$A_{tcm} = 0.255\left(\frac{A_{bi}}{N_p}\right)10^{+6} \tag{7.15}$$

$$I_{\text{rms}} = \frac{A_{\text{tcm}}}{D_{\text{cma}}}$$

$$= 0.255\frac{A_{\text{bi}}}{N_p D_{\text{cma}}}10^{+6} \tag{7.16}$$

Putting Eq. 7.16 into Eq. 7.14 we obtain

$$P_o = 0.0913\,\frac{V_{\text{dc(min)}}A_{\text{bi}}}{N_p D_{\text{cma}}}\,10^{+6} \tag{7.17}$$

From Faraday's law [since $V_{\text{dc(min)}}/2$ is applied to the primary]

$$V_{p(\text{min})} = \frac{V_{\text{dc(min)}}}{2} = N_p A_e \frac{\Delta B}{\Delta T}\,10^{-8}$$

where ΔB is $2B_{\text{max}}$ and ΔT is $0.4T$. Then

$$V_{\text{dc(min)}} = 10N_p f A_e B_{\text{max}} 10^{-8}$$

Putting this into Eq. 7.17, we have

$$P_o = \frac{0.00913B_{\text{max}} f A_e A_{\text{bi}}}{D_{\text{cma}}}$$

and dividing by 6.45 for the bobbin area in square centimeters:

$$P_o = \frac{0.0014B_{\text{max}} f A_e A_b}{D_{\text{cma}}} \tag{7.18}$$

where P_o is in watts, B_{max} is in gauss, f is in hertz, and A_e and A_b are in square centimeters.

7.3.4 Output power relations in full-bridge topology

A given core used in a full-bridge topology can yield no more output power than the same core used in a half bridge. The full bridge does deliver twice the output power of the half bridge, but it requires a larger core to do so. This comes about as follows. A full-bridge primary must sustain twice the supply voltage of the half bridge and hence must have twice the number of primary turns (Faraday's law and Sec. 3.3.2.1).

If the same fraction of the total bobbin area as in the half bridge is to be utilized, the wire size (area) must be halved. If the wire area is halved, operating at the same current density (circular mils per rms ampere), the permissible rms current must be halved. But the full-bridge transformer core, operating at twice the voltage and half the

primary current of a half bridge, delivers the same output power as does the half bridge with the same core.

It is, of course, true that a full-bridge primary, operating at twice the voltage of a half-bridge primary but at the same primary current as a half bridge, delivers twice the half-bridge output power. But the full bridge would require a larger core winding area and hence a larger core to contain twice the number of turns at the same current density of a half bridge.

7.3.5 Conversion of output power equations into charts permitting core and operating frequency selection at a glance

Equations 7.7, 7.13, and 7.18 are valuable in selecting a core and operating frequency for a desired output power. But as discussed in Sec. 7.1, using them requires a number of time-consuming iterative calculations.

The charts of Table 7.2*a* and 7.2*b* avoid such calculations by showing the available output power as calculated from these equations for a peak flux density B_{max} of 1600 G and a coil current density D_{cma} of 500 circular mils per rms ampere.

The reason for the selection of 1600 G is discussed in Secs. 7.2.3 and 2.2.9.4. At frequencies above about 50 kHz, excessive core losses for some of the more lossy materials may dictate a lower value $B_{max,l}$ for the peak flux density. The charts make a rapid calculation of the available power simple. The lowered available power at the lowered peak flux density is then the values shown in the charts multiplied by $(B_{max,l}/1600)$.

The selection of D_{cma} = 500 circular mils per rms ampere is a compromising choice common in transformer design. A higher density (lower value of D_{cma}) would result in more copper losses, and a lower density would unnecessarily increase the coil size. Values down to 400 or possibly as low as 300 circular mils per rms ampere can be acceptable, but densities below 300 should definitely be avoided.

Actually, the choice of D_{cma} values specifies only the DC wire resistance. In subsequent sections skin and proximity effects will be discussed. These effects produce eddy currents in the wires, cause the currents to flow in only a fraction of the wire area, and hence may increase the wire resistance considerably above the values shown in wire tables for wires of a specified number of circular-mil area. Nevertheless, choosing current density of 500 circular mils per rms ampere is a good starting point.

The charts of Table 7.2 are used as follows. First, choose a topology which yields the best combination of minimum power transistor off-

TABLE 7.2a Maximum Available Ouput Power in Forward Converter Topology

Core	A_e, cm^2	A_b, cm^2	A_eA_b, cm^4	Output power in watts at 20 kHz	24 kHz	48 kHz	72 kHz	96 kHz	150 kHz	200 kHz	250 kHz	300 kHz	Volume, cm^3
EE Cores, Ferroxcube-Philips													
814E250	0.202	0.171	0.035	1.1	1.3	2.7	4.0	5.3	8.3	11.1	13.8	16.6	0.57
813E187	0.225	0.329	0.074	2.4	2.8	5.7	8.5	11.4	17.8	23.7	29.6	35.5	0.89
813E343	0.412	0.359	0.148	4.7	5.7	11.4	17.0	22.7	35.5	47.3	59.2	71.0	1.64
812E250	0.395	0.581	0.229	7.3	8.8	17.6	26.4	35.3	55.1	73.4	91.8	110.2	1.93
782E272	0.577	0.968	0.559	17.9	21.4	42.9	64.3	85.8	134.0	178.7	223.4	268.1	3.79
E375	0.810	1.149	0.931	29.8	35.7	71.5	107.2	143.0	223.4	297.8	372.3	446.7	5.64
E21	1.490	1.213	1.807	57.8	69.4	138.8	208.2	277.6	433.8	578.4	722.9	867.5	11.50
783E608	1.810	1.781	3.224	103.2	123.8	247.6	371.4	495.1	773.7	1031.6	1289.4	1547.3	17.80
783E776	2.330	1.810	4.217	135.0	161.9	323.9	485.8	647.8	1012.2	1349.5	1686.9	2024.3	22.90
E625	2.340	1.370	3.206	102.6	123.1	246.2	369.3	492.4	769.4	1025.9	1282.3	1538.8	20.80
E55	3.530	2.800	9.884	316.3	379.5	759.1	1138.6	1518.2	2372.2	3162.9	3953.6	4744.3	43.50
E75	3.380	2.160	7.301	233.6	280.4	560.7	841.1	1121.4	1752.2	2336.3	2920.3	3504.4	36.00
EC Cores, Ferroxcube-Philips													
EC35	0.843	0.968	0.816	26.1	31.3	62.7	94.0	125.3	195.8	261.1	326.4	391.7	6.53
EC41	1.210	1.350	1.634	52.3	62.7	125.5	188.2	250.9	392.0	522.7	653.4	784.1	10.80
EC52	1.800	2.130	3.834	122.7	147.2	294.5	441.7	588.9	920.2	1226.9	1533.6	1840.3	18.80
EC70	2.790	4.770	13.308	425.9	511.0	1022.1	1533.1	2044.2	3194.0	4258.7	5323.3	6388.0	40.10
ETD Cores, Ferroxcube-Philips													
ETD 29	0.760	0.903	0.686	22.0	26.4	52.7	79.1	105.4	164.7	219.6	274.5	329.4	5.50
ETD 34	0.971	1.220	1.185	37.9	45.5	91.0	136.5	182.0	284.3	379.1	473.8	568.6	7.64
ETD 39	1.250	1.740	2.175	69.6	83.5	167.0	250.6	334.1	522.0	696.0	870.0	1044.0	11.50
ETD 44	1.740	2.130	3.706	118.6	142.3	284.6	427.0	569.3	889.5	1186.0	1482.5	1779.0	18.00
ETD 49	2.110	2.710	5.718	183.0	219.6	439.2	658.7	878.3	1372.3	1829.8	2287.2	2744.7	24.20
Pot Cores, Ferroxcube-Philips													
704	0.070	0.022	0.002	0.0	0.1	0.1	0.2	0.2	0.4	0.5	0.6	0.7	0.07
905	0.101	0.034	0.003	0.1	0.1	0.3	0.4	0.5	0.8	1.1	1.4	1.6	0.13
1107	0.167	0.054	0.009	0.3	0.3	0.7	1.0	1.4	2.2	2.9	3.6	4.3	0.25

TABLE 7.2*a* **Maximum Available Ouput Power in Forward Converter Topology** (*Continued*)

Core	A_e, cm^2	A_b, cm^2	A_eA_b, cm^4	Output power in watts at 20 kHz	24 kHz	48 kHz	72 kHz	96 kHz	150 kHz	200 kHz	250 kHz	300 kHz	Volume, cm^3
Pot Cores, Ferroxcube-Philips (*continued*)													
1408	0.251	0.097	0.024	0.8	0.9	1.9	2.8	3.7	5.8	7.8	9.7	11.7	0.50
1811	0.433	0.187	0.081	2.6	3.1	6.2	9.3	12.4	19.4	25.9	32.4	38.9	1.12
2213	0.635	0.297	0.189	6.0	7.2	14.5	21.7	29.0	45.3	60.4	75.4	90.5	2.00
2616	0.948	0.407	0.386	12.3	14.8	29.6	44.4	59.3	92.6	123.5	154.3	185.2	3.53
3019	1.380	0.587	0.810	25.9	31.1	62.2	93.3	124.4	194.4	259.2	324.0	388.8	6.19
3622	2.020	0.774	1.563	50.0	60.0	120.1	180.1	240.2	375.2	500.3	625.4	750.5	10.70
4229	2.660	1.400	3.724	119.2	143.0	286.0	429.0	572.0	893.8	1191.6	1489.6	1787.5	18.20
RM Cores, Ferroxcube-Philips													
RM5	0.250	0.095	0.024	0.8	0.9	1.8	2.7	3.6	5.7	7.6	9.5	11.4	0.45
RM6	0.370	0.155	0.057	1.8	2.2	4.4	6.6	8.8	13.8	18.4	22.9	27.5	0.80
RM8	0.630	0.310	0.195	6.2	7.5	15.0	22.5	30.0	46.9	62.5	78.1	93.7	1.85
RM10	0.970	0.426	0.413	13.2	15.9	31.7	47.6	63.5	99.2	132.2	165.3	198.3	3.47
RM12	1.460	0.774	1.130	36.2	43.4	86.8	130.2	173.6	271.2	361.6	452.0	542.4	8.34
RM14	1.980	1.100	2.178	69.7	83.6	167.3	250.9	334.5	522.7	697.0	871.2	1045.4	13.19
PQ Cores, Magnetics, Inc.													
42016	0.620	0.256	0.159	5.1	6.1	12.2	18.3	24.4	38.1	50.8	63.5	76.2	2.31
42020	0.620	0.384	0.238	7.6	9.1	18.3	27.4	36.6	57.1	76.2	95.2	114.3	2.79
42620	1.190	0.322	0.383	12.3	14.7	29.4	44.1	58.9	92.0	122.6	153.3	183.9	5.49
42625	1.180	0.502	0.592	19.0	22.7	45.5	68.2	91.0	142.2	189.6	236.9	284.3	6.53
43220	1.700	0.470	0.799	25.6	30.7	61.4	92.0	122.7	191.8	255.7	319.6	383.5	9.42
43230	1.610	0.994	1.600	51.2	61.5	122.9	184.4	245.8	384.1	512.1	640.1	768.2	11.97
43535	1.960	1.590	3.116	99.7	119.7	239.3	359.0	478.7	747.9	997.2	1246.6	1495.9	17.26
44040	2.010	2.490	5.005	160.2	192.2	384.4	576.6	768.8	1201.2	1601.6	2002.0	2402.4	20.45

Note: From Eq. 7.7, $P_o = 0.00050B_{max}fA_eA_b/D_{cma}$, where P_o is in watts, B_{max} in gauss, A_e and A_b in square centimeters, f in hertz, D_{cma} in circular mils per rms ampere, bobbin winding space factor = 40 percent. For B_{max} = 1600 G. For other B_{max}, multiply by $B_{max}/1600$. For D_{cma} = 500 circular mils/rms ampere. For other D_{cma}, multiply by $500/D_{cma}$. For push-pull topology, multiply powers by a factor of 2.

TABLE 7.2b Maximum Available Output Power in Half- or Full-Bridge Topology

Core	A_e, cm^2	A_b, cm^2	A_eA_b, cm^4	Output power in watts at 20 kHz	24 kHz	48 kHz	72 kHz	96 kHz	150 kHz	200 kHz	250 kHz	300 kHz	Volume, cm^3
EE Cores, Ferroxcube-Philips													
814E250	0.202	0.171	0.035	3.1	3.7	7.4	11.2	14.9	23.2	30.9	38.7	46.4	0.57
813E187	0.225	0.329	0.074	6.6	8.0	15.9	23.9	31.8	49.7	66.3	82.9	99.5	0.89
813E343	0.412	0.359	0.148	13.3	16.0	31.8	47.8	63.6	99.4	132.5	165.7	198.8	1.64
812E250	0.395	0.581	0.229	20.6	24.8	49.3	74.1	98.7	154.2	205.6	257.0	308.4	1.93
782E272	0.577	0.968	0.559	50.0	60.3	120.1	180.4	240.2	375.3	500.4	625.6	750.7	3.79
E375	0.810	1.149	0.931	83.4	100.5	200.1	300.6	400.2	625.4	833.9	1042.4	1250.8	5.64
E21	1.490	1.213	1.807	161.9	195.2	388.6	583.8	777.2	1214.6	1619.4	2024.3	2429.1	11.50
783E608	1.810	1.781	3.224	288.8	348.1	693.1	1041.2	1386.2	2166.2	2888.4	3610.4	4332.5	17.80
783E776	2.330	1.810	4.217	377.9	455.5	906.7	1362.2	1813.4	2834.0	3778.7	4723.4	5668.1	22.90
E625	2.340	1.370	3.206	287.2	346.2	689.2	1035.5	1378.5	2154.3	2872.4	3590.5	4308.6	20.80
E55	3.530	2.800	9.884	885.6	1067.5	2125.1	3192.5	4250.1	6642.0	8856.1	11070.1	13284.1	43.50
E75	3.380	2.160	7.301	654.2	788.5	1569.7	2358.2	3139.3	4906.1	6541.5	8176.9	9812.3	36.00
EC Cores, Ferroxcube-Philips													
EC35	0.843	0.968	0.816	73.1	88.1	175.4	263.6	350.9	548.4	731.2	913.9	1096.7	6.53
EC41	1.210	1.350	1.634	146.4	176.4	351.2	527.6	702.4	1097.7	1463.6	1829.5	2195.4	10.80
EC52	1.800	2.130	3.834	343.5	414.1	824.3	1238.4	1648.6	2576.4	3435.3	4294.1	5152.9	18.80
EC70	2.790	4.770	13.308	1192.4	1437.3	2861.3	4298.6	5722.6	8943.2	11924.2	14905.3	17886.4	40.10
ETD Cores, Ferroxcube-Philips													
ETD 29	0.760	0.903	0.686	61.5	74.1	147.6	221.7	295.1	461.2	614.9	768.6	922.4	5.50
ETD 34	0.971	1.220	1.185	106.1	127.9	254.7	382.6	509.4	796.1	1061.4	1326.8	1592.1	7.64
ETD 39	1.250	1.740	2.175	194.9	234.9	467.6	702.5	935.3	1461.6	1948.8	2436.0	2923.2	11.50
ETD 44	1.740	2.130	3.706	332.1	400.3	796.8	1197.1	1593.7	2490.6	3320.8	4150.9	4981.1	18.00
ETD 49	2.110	2.710	5.718	512.3	617.6	1229.4	1846.9	2458.8	3842.6	5123.4	6404.3	7685.1	24.20
Pot Cores, Ferroxcube-Philips													
704	0.070	0.022	0.002	0.1	0.2	0.3	0.5	0.7	1.0	1.4	1.7	2.1	0.07
905	0.101	0.034	0.003	0.3	0.4	0.7	1.1	1.5	2.3	3.1	3.8	4.6	0.13

TABLE 7.2*b* **Maximum Available Output Power in Half- or Full-Bridge Topology** (*Continued*)

Core	A_e, cm^2	A_b, cm^2	A_eA_b, cm^4	Output power in watts at 20 kHz	24 kHz	48 kHz	72 kHz	96 kHz	150 kHz	200 kHz	250 kHz	300 kHz	Volume, cm^3
Pot Cores, Ferroxcube-Philips (*continued*)													
704	0.070	0.022	0.002	0.1	0.2	0.3	0.5	0.7	1.0	1.4	1.7	2.1	0.07
905	0.101	0.034	0.003	0.3	0.4	0.7	1.1	1.5	2.3	3.1	3.8	4.6	0.13
1107	0.167	0.054	0.009	0.8	1.0	1.9	2.9	3.9	6.1	8.1	10.1	12.1	0.25
1408	0.251	0.097	0.024	2.2	2.6	5.2	7.8	10.4	16.3	21.8	27.2	32.7	0.50
1811	0.433	0.187	0.081	7.3	8.7	17.4	26.2	34.8	54.4	72.6	90.7	108.8	1.12
2213	0.635	0.297	0.189	16.9	20.4	40.5	60.9	81.1	126.7	169.0	211.2	253.5	2.00
2616	0.948	0.407	0.386	34.6	41.7	83.0	124.6	165.9	259.3	345.7	432.1	518.6	3.53
3019	1.380	0.587	0.810	72.6	87.5	174.2	261.6	348.3	544.4	725.8	907.2	1088.7	6.19
3622	2.020	0.774	1.563	140.1	168.9	336.1	505.0	672.3	1050.7	1400.9	1751.1	2101.3	10.70
4229	2.660	1.400	3.724	333.7	402.2	800.7	1202.9	1601.3	2502.5	3336.7	4170.9	5005.1	18.20
RM Cores, Ferroxcube-Philips													
RM5	0.250	0.095	0.024	2.1	2.6	5.1	7.7	10.2	16.0	21.3	26.6	31.9	0.45
RM6	0.370	0.155	0.057	5.1	6.2	12.3	18.5	24.7	38.5	51.4	64.2	77.1	0.80
RM8	0.630	0.310	0.195	17.5	21.1	42.0	63.1	84.0	131.2	175.0	218.7	262.5	1.85
RM10	0.970	0.426	0.413	37.0	44.6	88.8	133.5	177.7	277.7	370.2	462.8	555.4	3.47
RM12	1.460	0.774	1.130	101.3	122.0	243.0	365.0	485.9	759.4	1012.5	1265.6	1518.8	8.34
RM14	1.980	1.100	2.178	195.1	235.2	468.3	703.5	936.5	1463.6	1951.5	2439.4	2927.2	13.19
PQ Cores, Magnetics, Inc.													
42016	0.620	0.256	0.159	14.2	17.1	34.1	51.3	68.2	106.7	142.2	177.8	213.3	2.31
42020	0.620	0.384	0.238	21.3	25.7	51.2	76.9	102.4	160.0	213.3	266.6	320.0	2.79
42620	1.190	0.322	0.383	34.3	41.4	82.4	123.8	164.8	257.5	343.3	429.2	515.0	5.49
42625	1.180	0.502	0.592	53.1	64.0	127.4	191.3	254.7	398.1	530.8	663.4	796.1	6.53
43220	1.700	0.470	0.799	71.6	86.3	171.8	258.1	343.6	536.9	715.9	894.9	1073.9	9.42
43230	1.610	0.994	1.600	143.4	172.8	344.1	516.9	688.1	1075.4	1433.9	1792.4	2150.9	11.97
43535	1.960	1.590	3.116	279.2	336.6	670.0	1006.6	1340.1	2094.2	2792.3	3490.4	4188.4	17.26
44040	2.010	2.490	5.005	448.4	540.5	1076.1	1616.6	2152.1	3363.3	4484.4	5605.5	6726.6	20.45

Note: From Eq. 7.18, $P_o = 0.0014B_{max}fA_eA_b/D_{cma}$, where P_o is in watts, B_{max} in gauss, A_e and A_b in square centimeters, f in hertz, D_{cma} in circular mils per rms ampere, bobbin winding space factor = 40 percent. For B_{max} = 1600 G. For other B_{max}, multiply by $B_{max}/1600$. For D_{cma} = 500 circular mils/rms ampere. For other D_{cma}, multiply by $500/D_{cma}$.

voltage and peak-current stress. Or another topology selection criterion might be one that minimizes the number or cost of components.

Now note that the cores are arranged vertically in order of increasing A_eA_b product and hence increasing output power capability. If familiarity or experience dictates or suggests a particular operating frequency, that vertical frequency column is entered. Now move vertically downward and choose the first core whose output power is greater than the specified maximum power.

Or if a specific core is first chosen because it fits the available space, go to that core and move horizontally to the right to the first frequency which yields somewhat more than the specified maximum output power.

If a desired core does not yield the required output power at a selected frequency in—say—the forward converter topology, the push-pull might be considered. The push-pull (voltage or current mode) topology with the same core offers twice the output power at the same frequency. If voltage-mode push-pull is selected, all the precautions relating to flux imbalance (Sec. 2.2.8) should be kept in mind.

Thus, by moving upward to smaller cores and to the right to higher frequencies, an optimum core-frequency combination can be found. For a given output power, at higher frequencies, the core gets smaller but core losses and transformer temperature rise, and transistor switching losses increase.

7.3.5.1 Peak flux density selection at higher frequencies. Care should be taken in the use of Table 7.2*a* and 7.2*b*. The powers shown therein are available only if operation at a peak flux density of 1600 G at the selected frequency does not cause excessive temperature rise. At frequencies in the range of 20 to 50 kHz, core losses, even for the most lossy materials of Table 7.1 are so low that temperature rise at a peak flux density of 1600 G is negligibly small.

However, core losses increase roughly as the 1.6th power of the frequency and the 2.7th power of the peak flux density. Thus at higher frequencies (starting at about 50 kHz) peak flux density may have to be reduced below 1600 G (by increasing the number of primary turns) to keep the transformer temperature rise acceptably low.

In general, smaller cores can more easily tolerate a higher peak flux density at high frequencies than can larger cores. This is so because core losses are proportional to volume, but core cooling is proportional to its radiating surface area. Thus, as a core gets larger, its volume increases faster than its surface area and the internal heat generated increases more rapidly than its surface area which cools it. Note, for example, that the volume of a sphere is proportional to the third

power of its radius but its radiating surface area is proportional to only the second power of its radius).

A specific example can easily demonstrate this. Consider the Ferroxcube-Philips E55 core in Table 7.2*b*. It is seen in that table that if operation at 200 kHz and 1600 G were possible, it would be capable of 8856 W of output power in a half-bridge topology. But from Table 7.1, for 3C85 material, its losses are 700 mW/cm^3 at 1600 G and 200 kHz. Then for its volume of 43.5 cm^3, its dissipation is 0.7(43.5) = 30.5 W. Coil losses (considered in the following section) probably equal this.

Consider a smaller core, the 813E343. From Table 7.2*b*, its output current capability at 1600 G, 200 kHz is 133 W in a half bridge. For a volume of 1.64 cm^3 and the same 700 mW/cm^3, its core losses are only 1.15 W. Thus, neglecting coil losses, the 813E343 with a volume of 1.64 cm^3 and core losses of 1.15 W would run at a far lower temperature than the E55 core with a volume of 43.5 cm^3 and 30.5 W of core losses.

Calculation of actual transformer temperature rise due to core plus coil losses will be demonstrated in the following section.

Thus for the larger cores, the powers shown in Table 7.2*a* and 7.2*b* may not be obtainable at frequencies above 50 kHz as operation at 1600 G may result in excessive temperature rise. Peak flux density may have to be reduced to B_{max} of somewhere in the range 1400 to 800 G. Available output powers are then the values shown in Table 7.2*a* and 7.2*b* multiplied by $B_{max}/1600$.

Table 7.2*a* and 7.2*b* shows output powers for the two major American core manufacturers. Many of their cores are interchangeable in their geometries and A_e values. But they are made in their own proprietary core materials which have different core losses in milliwatts per cubic centimeter. Core interchangeability (with regard to geometry and A_e only) and their type numbers are shown in Table 7.3 for Ferroxcube-Philips, Magnetics Inc., and TDK. Manufacturers' catalogs[1–4] also show a variety of core accessories—bobbins and assembly and mounting hardware.

7.4 Transformer Temperature Rise Calculations[6]

Transformer temperature rise above its ambient-air environment depends on its total core plus coil (copper) losses and its radiating surface area. Forced air flowing past the transformer can lower its temperature rise considerably, depending on the airflow rate in cubic feet per minute.

There is no way of calculating transformer temperature rise with

TABLE 7.3 Core Type Numbers for Geometrically Interchangeable Cores

Ferroxcube-Philips	Magnetics Inc.	TDK
	EE Cores	
814E250	41205	
813E187	41808	EE19
813E343		
812E250		
782E272		
E375	43515	
E21	44317	
783E608		EE42/42/15
783E776		
E625	44721	
E55		EE55/55/21
E75	45724	
	EC Cores	
EC35	43517	EC35
EC41	44119	EC41
EC52	45224	EC52
EC70	47035	EC70
	ETD Cores	
ETD29		
ETD34	43434	ETD34
ETD39	43939	ETD39
ETD44	44444	ETD44
ETD49	44949	ETD49
	Pot Cores	
704	40704	P7/4
905	40905	P9/5
1107	41107	P11/17
1408	41408	P14/8
1811	41811	P14/8
2213	42213	P22/13
2616	42616	P26/16
3019	43019	P30/19
3622	43622	P36/22
4229	44229	P42/29
	RM Cores	
RM4	41110	RM4
RM5	41510	RM5
RM6	41812	RM6
RM7		RM7
RM8	42316	RM8
RM10	42819	RM10
RM12	43723	RM12
RM14		RM14
	PQ Cores	
	42016	PQ20/16
	42020	PQ20/20
	42620	PQ2620
	42625	PQ26/25
	43220	PQ32/20
	43230	PQ32/30
	43535	PQ32/30
	44040	PQ40/40
		PQ50/50

Each manufacturer produces these cores in their own proprietary core materials having their unique core loss characteristic in milliwatts per cubic centimeter.

great accuracy analytically. But it can be done to within about 10°C with some empirical curves based on the concept of thermal resistance of a radiating surface area. Recall that the definition of thermal resistance R_t of a heat sink is the temperature rise (usually in degrees Celsius) per watt of dissipation in it. Then temperature rise dT for a power dissipation P is simply $dT = PR_t$.

Some core manufacturers list R_t for their various cores, implying that factor multiplied by the total core plus copper losses yields the temperature rise of the outer surface of the core. Some educated guess of about 10 to 15°C is often assumed for the temperature rise of internal hot spot (usually the core center leg) above the core's outer surface.

Temperature rise is dependent not only on the radiating surface area but also on the total dissipation. The greater the power dissipation into a radiating surface, the greater the temperature differential between the surface area and the ambient air and the more easily the surface area loses its heat or the lower the thermal resistance.

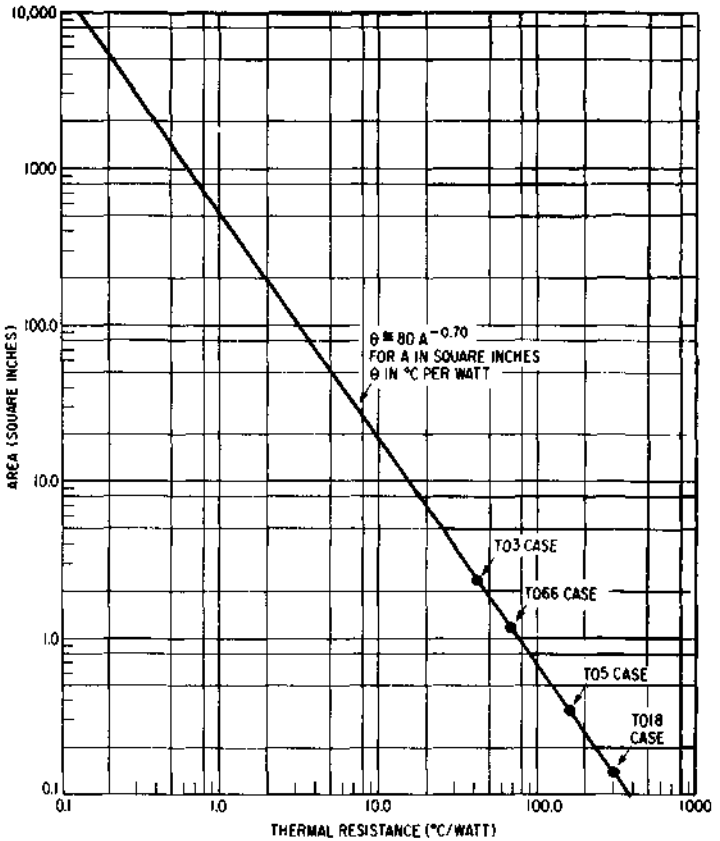
Thus transformer temperature rise will be estimated herein[8] as if the transformer's total outer surface area (2 × width × height + 2 × width × thickness + 2 × height × thickness) was the radiating area of an equivalent heat sink. The thermal resistance of this equivalent heat sink will be modified by the total dissipation (total of core plus copper losses).

A curve of heat-sink thermal resistance versus total surface area is shown in Fig. 7.4*a*. It is an empirical curve and is taken from the average of a large number of heat sinks of different sizes and shapes from different heat-sink manufacturers. It is the thermal resistance at a 1-W power level and is seen to lie on an averaging straight line on a log-log graph.

Although thermal resistance of a finned heat sink depends somewhat on the fin's shape and spacing and whether the surface is blackened or aluminized, these are second-order effects. To a close approximation, the thermal resistance of a heat sink depends almost entirely on only its radiating surface area.

Also from various heat-sink manufacturers' catalogs, an averaging empirical curve giving the variation of thermal resistance with power dissipation is shown in Fig. 7.4*b*.

From Fig. 7.4*a* and 7.4*b*, the more directly useful curves of Fig. 7.4*c* are derived. Figure 7.4*c* gives the temperature rise above ambient for various heat-sink areas (diagonal lines) and power dissipation. Thus the transformer's outer surface temperature rise will hereafter be read from Fig. 7.4*c* for the sum of its core plus copper losses and total radiating surface area as defined above. It is interesting to read from Fig. 7.4*c* the temperature rise of the two cores discussed in Sec.

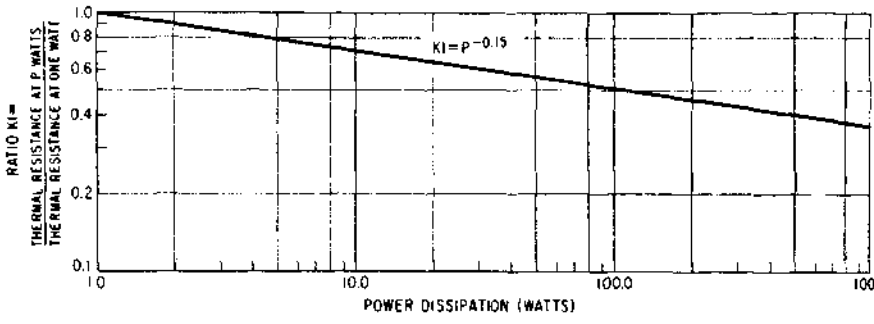


(a)

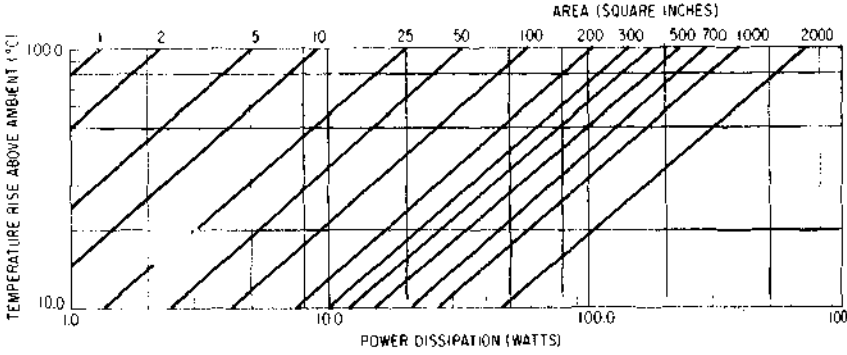
Figure 7.4 Calculating transformer temperature rise from its equivalent heat-sink area (total area of both faces plus core edge area). (*a*) Thermal resistance versus total heat-sink area. Total area means area of both sides of a flat plate or includes both sides of a fin in a finned heat sink. Curve is at a power dissipation of 1 W. Use multiplying factor of Fig. 7.4*b* for other power levels. (*b*) Normalized thermal resistance versus power dissipation in a heat sink. (*c*) Heat-sink temperature rise versus power dissipation for various heat-sink areas. From Eq. 7.19: $T = 80A^{-0.70}P^{+0.85}$. Figure 7.4*b* is represented analytically by $K_1 = P^{-0.15}$. Combining Fig. 7.4*a* ($= 80A^{-0.70}$) and 7.4*b* gives the temperature rise for any transformer power dissipation and radiating surface areas as $T = 80A^{-0.70}P^{+0.85}$.

7.3.5.1. There it was calculated that an E55 of 43.5 cm^3 volume, operated at 1600 G and 200 kHz, dissipated 30.5 W. Its radiating surface area as defined above is 16.5 in^2. From Fig. 7.4c, neglecting copper losses entirely, at 30.5 W of core losses, its temperature rise is 185°C.

The smaller 813E343 core of 1.64-cm^3 volume also operated at 1600 G and 200 kHz has 1.15 W of core losses. Its radiating surface area,



(b)



(c)

Figure 7.4 (*Continued*)

also calculated as above, is 1.90 in^2. Also from Fig. 7.4*c*, neglecting copper losses, its temperature rise is only 57°C. It is thus verified that it is easier for smaller cores to deliver the powers shown in Fig. 7.2*a* and 7.2*b* at 1600 G and high frequency.

It is of interest to compare the thermal resistance of some cores as measured by the manufacturer and as calculated ($R_t = 80A^{-0.70}$) from Fig. 7.4*a* (see Table 7.4).

7.5 Transformer Copper Losses

7.5.1 Introduction

In Sec. 7.3 it was stated that wire size for all windings would be chosen to yield a current density of 500 circular mils per rms ampere. It was assumed there that copper losses would be calculated as $(I_{rms})^2 R_{dc}$, where R_{dc} is the winding's DC resistance as calculated from

TABLE 7.4 Core Resistance

Core	Radiating surface, area, in^2	Thermal resistance, °C/W	
		Measured by manufacturer	Calculated from Fig. 7.4*a*
EC35	5.68	18.5	23.7
EC41	7.80	16.5	19.0
EC52	10.8	11.0	12.6
EC70	22.0	7.5	9.2

its length and resistance in ohms per foot as read from the wire tables for the selected wire size. It was also assumed that I_{rms} is the rms current as calculated from its waveshape (Secs. 2.2.10.2, 2.3.10.4).

There are two effects, *skin* and *proximity* effects, which can cause the winding losses to be significantly greater than $(I_{rms})^2 R_{dc}$.

Both skin and proximity effects arise from eddy currents which are induced by varying magnetic fields in the coil. Skin effect is caused by eddy currents induced in a wire by the magnetic field of the current carried by the wire itself. Proximity effect is caused by eddy currents induced in wires by magnetic fields of currents in adjacent wires or adjacent layers of the coil.

Skin effect causes current in a wire to flow only in a thin skin on the outer periphery of the wire. The depth of this skin or annular conducting area is inversely proportional to the square root of the frequency. Thus, as frequency increases, a progressively larger part of solid wire area is lost, increasing the AC resistance and hence copper losses.

One might not expect skin effect to increase wire resistance significantly for low frequency (say, 25 kHz) switchers as the skin depth is still a fairly large 17.9 mils at 25 kHz.

However, currents in conventional switching supplies have rectangular waveshapes which have a considerable amount of energy in their higher-frequency Fourier components. Thus high skin resistance at these high-frequency Fourier components of even a 25-kHz rectangular current waveform makes skin effect a concern even for a low-frequency 25 kHz switcher.

Skin effect will be discussed quantitatively in the following sections.

Proximity effect, caused by eddy currents induced by varying magnetic fields from adjacent conductors or adjacent coil layers, causes considerably more copper losses than does skin effect.

Proximity-effect losses can be especially high in multilayer coils. This is so only partly because the induced eddy currents crowd the net current into a small fraction of the copper wire area, increasing its resistance. What is more serious in proximity effect is that these induced eddy currents can be many times greater in amplitude than the

net current flowing in the individual wire or wire layers. This will be discussed quantitatively in the following sections.

7.5.2 Skin effect

Skin effect[9–15] has been known and equations had been derived for skin depth versus frequency as far back as 1915.[9] The means by which induced eddy currents cause current to crowd into the thin outer skin of a conductor can be seen in Fig. 7.5.

A section of a round conductor sliced through a diameter is shown in Fig. 7.5. It carries its main current in the direction of *OA*. If not for skin effect, the current would be uniformly distributed throughout the wire area.

Now all elements of current flowing along the direction *OA* are encircled by magnetic flux lines normal to *OA*. Consider a thin filament of current flowing along the axis *OA*. By Fleming's right-hand rule,

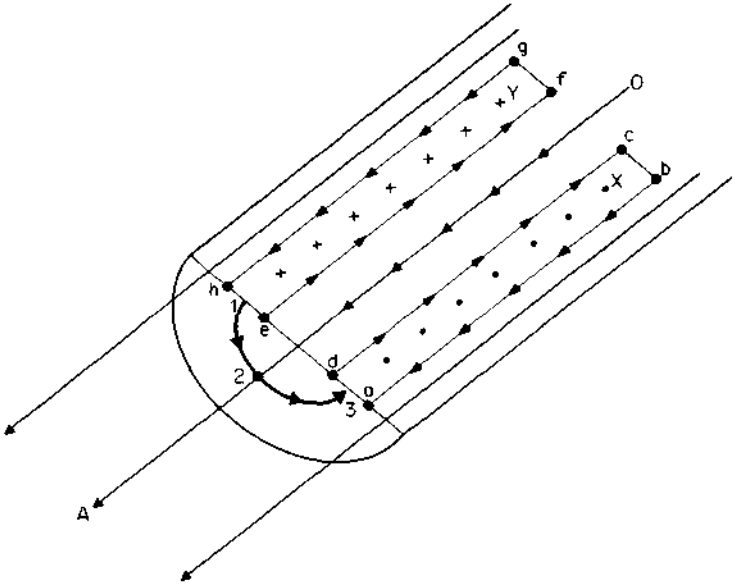


Figure 7.5 Eddy currents in a round wire causes skin effect—the canceling of current in the center of the wire and its crowding into the outer skin. The magnetic field of the current in the wire induces voltages in loops such as *abcd* and *efgh*. The polarity of these voltages is such as to cause eddy currents to flow around the boundary of these loops. The direction of these eddy currents is opposite to the main current flow on the inside of the loops (*d* to *c* and *e* to *f*) and in the same direction as the main current flow on the outside of the loops (*b* to *a* and *g* to *h*). The consequence is the canceling of current flow on the inside of the wire and its concentration in a thin skin on its outside.

its magnetic flux lines are in the direction shown by the arrow in the figure—from 1 to 2 to 3 and around back to 1.

Consider two flat loops (X and Y) within the wire. They are on a wire diameter and extend the full length of the wire. These loops are symmetrically disposed to either side of the wire axis. The magnetic flux lines flow up through loop X (shown by the dots in the center of the loop), around, and back down through loop Y (shown by the crosses in the center of loop Y).

By Faraday's law, when a varying magnetic field flows through an area, a voltage is induced in a line encircling that area. By Lenz' law, the polarity of that voltage is such as to cause an eddy current flow in such a direction that its magnetic field will be in a direction opposite to that which caused the eddy current.

Thus voltages will be induced in loops X and Y and eddy currents will flow in those loops in the direction shown. By the right-hand rule, the current in loop X flows clockwise in the direction d to c to b to a to d. That direction of current flow causes a magnetic field to go down through the center of the loop and in opposition to the field from the main filament of current along OA. Similarly, in loop Y, the eddy current flows counterclockwise ($efgh$ and back to e) so as to cause a magnetic field to come up through the center of the loop in opposition to the magnetic field of the main current filament along OA.

Now note that the eddy currents along arms dc and ef are in the direction opposite to that of the main current filament OA and tend to cancel it. Further, eddy currents along arms ab and gh (along the outer skin of the wire) are in the same direction of the main current filaments and tend to reinforce them.

Thus the net current—the sum of the eddy currents and the main current which caused them—is canceled at the center of the wire and crowded into the outer skin. Thus at high frequency, the total current-carrying area is less than the full wire area and the AC resistance is greater than the DC resistance by an amount determined by the skin thickness.

7.5.3 Skin effect—quantitative relations

Skin depth is defined as distance below the surface where the current density has fallen to $1/e$ or 37 percent of its value at the surface. The relation between skin depth and frequency has been derived by many sources[9] and for copper wire at 70°C is

$$S = \frac{2837}{\sqrt{f}} \tag{7.19}$$

where S is the skin depth in mils and f is frequency in hertz.

Table 7.5 shows skin depth for copper wire at 70°C at various frequencies as calculated from Eq. 7.19.

Now assume conductors of circular cross section. For a DC resistance R_{dc}, AC resistance due to skin effect of R_{ac}, and ΔR, the resistance change due to skin effect is

$$R_{ac} = R_{dc} + \Delta R = R_{dc}\left(1 + \frac{\Delta R}{R_{dc}}\right) = R_{dc}(1 + F)$$

or

$$\frac{R_{ac}}{R_{dc}} = 1 + F \tag{7.20}$$

From the skin depth relation of Eq. 7.19, the ratio R_{ac}/R_{dc} can be calculated for any wire size at any frequency. Since the resistances are inversely proportional to wire conducting area, for any skin depth S and wire radius r and diameter d, since the conducting area of the wire is the annular ring whose inner radius is $(r - S)$, then

$$\begin{aligned}\frac{R_{ac}}{R_{dc}} &= \frac{\pi r^2}{\pi r^2 - \pi(r - S)^2} \\ &= \frac{(r/S)^2}{(r/S)^2 - (r - S)^2/S^2} \\ &= \frac{(r/S)^2}{(r/S)^2 - (r/S - 1)^2} \\ &= \frac{(d/2S)^2}{(d/2S)^2 - (d/2S - 1)^2}\end{aligned} \tag{7.21}$$

TABLE 7.5 Skin Depth in Copper Wire at 70°C

Frequency, kHz	Skin depth, mils*
25	17.9
50	12.7
75	10.4
100	8.97
125	8.02
150	7.32
175	6.78
200	6.34
225	5.98
250	5.67
300	5.18
400	4.49
500	4.01

*From Eq. 7.19. Skin depth $S = 2837/\sqrt{F}$; S in mils for F in hertz.

Thus Eq. 7.21 indicates that the wire's AC-to-DC resistance $R_{ac}/R_{dc} = (1 + F)$ is dependent only on the ratio of wire diameter to skin depth. Figure 7.6 plots R_{ac}/R_{dc} against the ratio d/S from Eq. 7.21.

7.5.4 AC/DC resistance ratio for various wire sizes at various frequencies

Because of skin effect, the AC-to-DC resistance of round wire is dependent on the ratio of the wire diameter to skin depth (Eq. 7.21). Further, since skin depth is inversely proportional to the square root of frequency, different-sized wires will have different AC-to-DC resistance ratios, and these ratios will increase with frequency.

Table 7.6 shows this for all even-numbered wire sizes at 25, 50, 100, and 200 kHz. In this table, d/S (wire diameter/skin depth ratio) is calculated from the maximum bare wire diameter as given in the wire tables and skin depth is calculated from Eq. 7.19 (Table 7.5). From these d/S ratios, R_{ac}/R_{dc} is calculated from Eq. 7.21 or read from Fig. 7.6.

It is obvious from Table 7.6 that large-diameter wires have a large AC/DC resistance ratio, which increases greatly with frequency. Thus No. 14 wire has a diameter of 64.7 mils and a skin depth of 17.9 mils at 25 kHz (Table 7.5). This yields a d/S ratio of 3.6, and from Fig. 7.6, its AC/DC resistance ratio is already 1.25. But the same wire at 200 kHz has a skin depth of 6.34 mils and a d/S ratio of 10.2—and from Fig. 7.6, its AC/DC resistance ratio is 3.3!

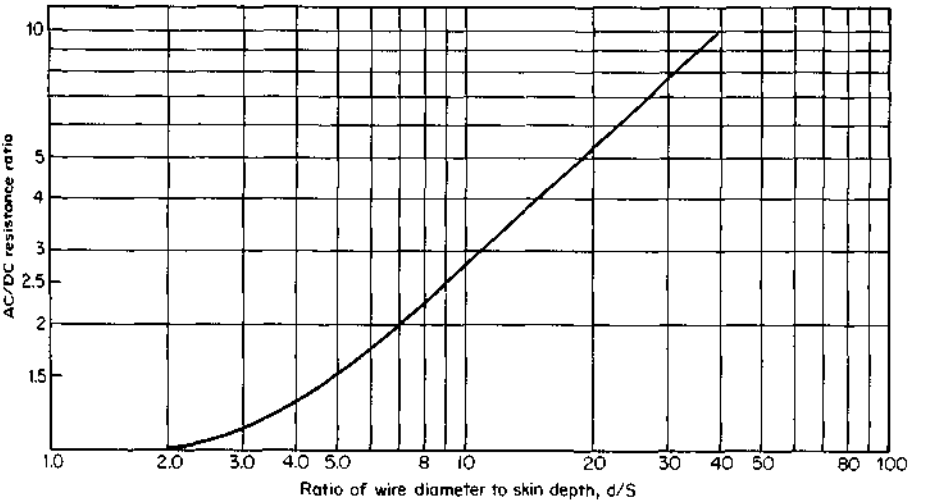


Figure 7.6 AC-to-DC resistance ratio for round wires versus ratio of wire diameter to skin depth (d/s). (*From Eq. 7.21.*)

TABLE 7.6 AC/DC Resistance Ratios Due to Skin Effects for All Even-Numbered Wire Sizes at Four Commonly Used Frequencies

		25 kHz			50 kHz			100 kHz			200 kHz		
Wire no.	Diameter d, mils	Skin depth S, mils	d/S	R_{ac}/R_{dc}	Skin depth S, mils	d/S	R_{ac}/R_{dc}	Skin depth S, mils	d/S	R_{ac}/R_{dc}	Skin depth, S, mils	d/S	R_{ac}/R_{dc}
12	81.6	17.9	4.56	1.45	12.7	6.43	1.85	8.97	9.10	2.55	6.34	12.87	3.50
14	64.7	17.9	3.61	1.30	12.7	5.09	1.54	8.97	7.21	2.00	6.34	10.21	2.90
16	51.3	17.9	2.87	1.10	12.7	4.04	1.25	8.97	5.72	1.70	6.34	8.09	2.30
18	40.7	17.9	2.27	1.05	12.7	3.20	1.15	8.97	4.54	1.40	6.34	6.42	1.85
20	32.3	17.9	1.80	1.00	12.7	2.54	1.05	8.97	3.60	1.25	6.34	5.09	1.54
22	25.6	17.9	1.43	1.00	12.7	2.02	1.00	8.97	2.85	1.10	6.34	4.04	1.30
24	20.3	17.9	1.13	1.00	12.7	1.60	1.00	8.97	2.26	1.04	6.34	3.20	1.15
26	16.1	17.9	0.90	1.00	12.7	1.27	1.00	8.97	1.79	1.00	6.34	2.54	1.05
28	12.7	17.9	0.71	1.00	12.7	1.00	1.00	8.97	1.42	1.00	6.34	2.00	1.00
30	10.1	17.9	0.56	1.00	12.7	0.80	1.00	8.97	1.13	1.00	6.34	1.59	1.00
32	8.1	17.9	0.45	1.00	12.7	0.64	1.00	8.97	0.90	1.00	6.34	1.28	1.00
34	6.4	17.9	0.36	1.00	12.7	0.50	1.00	8.97	0.71	1.00	6.34	1.01	1.00

Note: Skin depths are taken from Table 7.5; R_{ac}/R_{dc} from Eq. 7.21.

However, Table 7.6 should not be misinterpreted. Although Fig. 7.6 shows that R_{ac}/R_{dc} increases as wire diameter increases (d/S increases), R_{ac} actually decreases as wire diameter increases and larger wire sizes will yield lesser core losses. This, of course, is because R_{dc} is inversely proportional to d^2 and decreases more rapidly than R_{ac} decreases as a result of increase in d. This is because R_{ac} is inversely proportional to the area of the annular skin, which is equal to dS. Thus as d increases, R_{ac} decreases.

As is well known, large-diameter wire is much too lossy to use at high frequencies. Rather than using a single large-diameter wire, a number of paralleled smaller-diameter wires with a total equal circular-mil areas are used. This increases the total area of conducting annular skin zones. This can be seen as follows. If, say, two paralleled wires are to have a total circular mil area of a single wire, the diameter of the two smaller wires must be $D/\sqrt{2}$ (where D is the diameter of the single original wire).

Then for a skin depth of S, the single original wire has a total area in its annular skin zone of πDS. In the two smaller wires, skin depth is not related to wire diameter—only to frequency (Eq. 7.20). Thus with the two smaller wires, the total area of the two skins is $2[\pi(D/\sqrt{2})S] = \pi\sqrt{2}DS$; and the conducting area of the two skins with two smaller wires has been increased by $\sqrt{2}$ or 41 percent over that of a single wire with the same skin depth and equal circular-mil area at DC.

An extension of this is to use Litz wire.[16] This consists of a number of fine, insulated wires twisted into a bundle. Individual wires are woven together in a bundle so that each wire spends an equal distance at the center of the bundle and at its periphery. This minimizes both skin and proximity effects.[12]

Litz wire is about 5 percent more expensive than solid wire. But this is not as much of a drawback as is the difficulty of handling it in a production environment. Care must be taken that all the fine strands (usually from No. 28 to No. 50 wire) are soldered together at each end. It is reported that if some of the fine wires are broken or for some reason not picked up in the soldered connection at each end, losses increase significantly. Also, odd effects such as audible noise or vibration can occur.

General practice is to avoid Litz wire up to switching frequencies of 50 kHz. It is occasionally used at 100 kHz, and its use should be weighed against the use of two or possibly up to four paralleled smaller-diameter wires.

Some appreciation of this is obtained from Table 7.6, where it is seen that AC resistance of No. 18 wire is only 5 percent greater than No. 18 wire DC resistance at 25 kHz because of skin effect. This is not

too significant. But at 50 kHz, R_{ac} is already 15 percent higher and at 100, 200 kHz, it is 40 and 85 percent higher, respectively.

These numbers do not take into account that in most switching supply topologies, currents have rectangular waveshapes in which much of the energy is in higher harmonics. When the losses of higher harmonics of the current square wave are considered, the AC/DC resistance ratios of Table 7.6 will be seen to increase significantly. This will be taken up in the following section.

For high currents (usually in secondaries above ~ 15 to 20 A), thin copper foil rather than Litz wire or multiple strands of solid wire is often used. The foil width is cut to the bobbin width (or appropriately less if VDE safety specifications must be observed). Foil thickness is chosen as about 37 percent greater thickness than the skin depth at the fundamental switching frequency. The foil is covered with a 1-mil layer of plastic (Mylar) and is wrapped as a ribbon around the bobbin for the required number of turns.

7.5.5 Skin effect with rectangular current waveshapes[14]

The ratio of AC-to-DC resistance is strongly dependent on the wire diameter/skin depth ratio (Fig. 7.6). But skin depth is dependent on frequency (Eq. 7.19). In most switching power supply topologies, current waveshapes are rectangular with much of the energy residing in the harmonics of the square-wave current waveshapes. The question thus arises as to the frequency at which skin depth should be calculated. Venkatramen has rigorously analyzed this issue.[14] A simplifying approximation has been made herein to permit estimation of the ratio of AC to DC resistance and hence calculation of copper losses.

It is assumed that the majority of the energy in the square current waveshape resides in the first three harmonics of the square-wave frequency. Skin depth S is then calculated from Eq. 7.19 for each of the three harmonics of the most usual switching frequencies (25, 50, 100, 200 kHz).

An average of these S_{av} is taken for the simplifying average skin depth for those switching frequencies. From this average skin depth, the average ratio d/S_{av} is calculated for all even-numbered wire sizes. This d/S_{av} is then used in reading R_{ac}/R_{dc} from Fig. 7.6. The results are shown in Table 7.7.

If this simplifying approximation is valid (it may be optimistic, depending on how much of the square-wave energy is contained at harmonics above the third), Table 7.6 may be giving an optimistic estimate of skin effect losses for square current waveforms. Table 7.8

TABLE 7.7 Skin Effect AC/DC Resistance Ratios for Square-Wave Currents at Four Commonly Used Switching Frequencies

		25 kHz			50 kHz			100 kHz			200 kHz		
Wire no.	Diameter d, mils	Skin depth S, mils	d/S	R_{ac}/R_{dc}	Skin depth S, mils	d/S	R_{ac}/R_{dc}	Skin depth S, mils	d/S	R_{ac}/R_{dc}	Skin depth, S, mils	d/S	R_{ac}/R_{dc}
12	81.6	13.2	6.18	1.85	9.66	8.45	2.40	6.83	11.95	3.30	4.83	16.89	4.50
14	64.7	13.2	4.90	1.50	9.66	6.70	1.90	6.83	9.47	2.65	4.83	13.40	3.70
16	51.3	13.2	3.89	1.25	9.66	5.31	1.59	6.83	7.51	2.12	4.83	10.62	2.90
18	40.7	13.2	3.08	1.13	9.66	4.21	1.35	6.83	5.96	1.75	4.83	8.43	2.36
20	32.3	13.2	2.45	1.05	9.66	3.34	1.17	6.83	4.73	1.45	4.83	6.69	1.90
22	25.6	13.2	1.94	1.00	9.66	2.65	1.07	6.83	3.75	1.25	4.83	5.30	1.56
24	20.3	13.2	1.54	1.00	9.66	2.10	1.01	6.83	2.97	1.12	4.83	4.20	1.35
26	16.1	13.2	1.22	1.00	9.66	1.67	1.00	6.83	2.36	1.04	4.83	3.33	1.17
28	12.7	13.2	0.96	1.00	9.66	1.31	1.00	6.83	1.86	1.00	4.83	2.63	1.07
30	10.1	13.2	0.77	1.00	9.66	1.05	1.00	6.83	1.48	1.00	4.83	2.09	1.01
32	8.1	13.2	0.61	1.00	9.66	0.84	1.00	6.83	1.19	1.00	4.83	1.68	1.00
34	6.4	13.2	0.48	1.00	9.66	0.66	1.00	6.83	0.94	1.00	4.83	1.33	1.00

Note: This is a simplifying approximation. It is assumed that most of the square-wave energy resides in the first three harmonics of the fundamental square-wave frequency. Average skin depth for the square-wave current is then taken as the average of the values for the first three harmonics of each fundamental as read in Table 7.5. From these average skin depths, d/S is calculated and from this, R_{ac}/R_{dc} is read from Fig. 7.6.

TABLE 7.8 AC-to-DC Resistance for No. 18 Wire

Frequency, kHz	R_{ac}/R_{dc} (Table 7.6) sine-wave currents	R_{ac}/R_{dc} (Table 7.7) square-wave currents
25	1.05	1.13
50	1.15	1.35
100	1.40	1.75
200	1.85	2.36

presents a comparison of R_{ac}/R_{dc} for No. 18 wire as read from Tables 7.6 and 7.7.

7.5.6 Proximity effect

Proximity effect[11–15] is caused by alternating magnetic fields arising from currents in adjacent wires or, more seriously, from currents in adjacent winding layers in a multilayer coil.

It is more serious than skin effect because the latter increases copper losses only by restricting the conducting area of the wire to a thin skin on its surface. But it does not change the magnitude of the currents flowing—only the current density at the wire surfaces. In contrast, in proximity effect, eddy currents caused by magnetic fields of currents in adjacent coil layers increase exponentially in amplitude as the number of coil layers increases.

7.5.6.1 Mechanism of proximity effect. Figure 7.7 shows how proximity effect comes about. There currents are shown flowing in opposite directions (AA' and BB') in two parallel conductors. For simplicity, the conductors are shown as having thin rectangular cross section and are closely spaced. The conductors could just as well be round wires or flat layers of closely spaced round wires such as adjacent layers in a transformer coil.

The bottom conductor is surrounded by a magnetic field which is shown coming out of its edge 1234, passing into the edge of the upper conductor, out of the opposite edge, and returning back into the far edge of the lower conductor. By Fleming's right-hand rule, the direction of the magnetic field is into edge 5678 of the upper conductor.

By Faraday's law, this varying magnetic field flowing through the area of loop 5678 induces a voltage in series with any line bounding the area. By Lenz' law, the direction of this induced voltage is to produce a current flow in the area boundary such that its magnetic field is in the direction opposite to that of the magnetic field which induced the current flow.

Thus the current flow is counterclockwise in loop 5678. It is seen

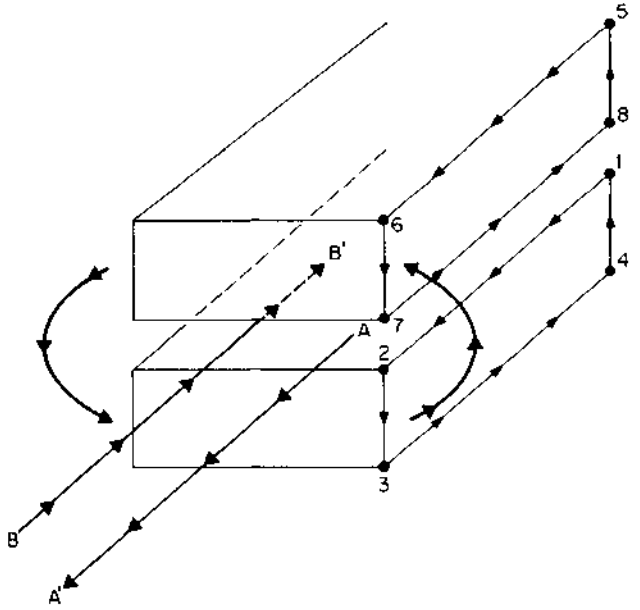


Figure 7.7 Magnetic fields of currents in one conductor induce voltages in loops of adjacent conductors. The resultant eddy currents shown in this diagram flow along the full length of the wires on their top and bottom surfaces. On the bottom conductor, eddy current is in the same direction as the main current flow (AA') on its top surface and reinforces it. On the bottom surface of the bottom conductor, the eddy current is in the direction opposite to that of the main current flow and cancels it. On the bottom surface of the upper conduction, the eddy current is in the same direction as that of the main current and reinforces it. On the top surface of the top conductor, the direction of the eddy current is opposite to that of the main current and cancels it. The consequence is that current in each conductor is confined to thin skins in the surfaces facing each other.

that on the bottom end of the loop, the current flow is in the same direction of current flow (7 to 8) as the main current in the upper conductor (B to B') and reinforces that current. On the top edge of the loop, current flow is opposite (5 to 6) to the main current and tends to cancel it. This occurs at all the loops parallel to 5678 throughout the conductor width.

The consequence of this is an eddy current flowing along the full length of the bottom surface of the upper conductor in the direction 7 to 8 and returning along the upper surface of the upper conductor where it is canceled by the main current.

A similar analysis shows an eddy current flows along the full length of the upper surface of the bottom conductor in the direction to rein-

force the main current. In the bottom surface of the lower conductor, it is in the direction opposite to the main current flow and tends to cancel it.

Thus currents in the two conductors are confined to a thin skin on the surfaces of the conductor which face one another. The depth of the skin is related to frequency as in normal skin effect.

7.5.6.2 Proximity effect in adjacent layers in a transformer coil. Current in individual wires in a layer of a transformer coil flow parallel to one another. The current in a layer can then be considered to flow in a thin rectangular sheet whose height is the wire diameter and whose width is the bobbin width. Thus there will be induced eddy currents flowing the full length of the winding. They will flow in thin skins on the interfaces between adjacent coil layers just as described in the previous section for proximity effect in two adjacent flat conductors.

However, it is very significant that the amplitude of these eddy currents increases exponentially with the number of layers. It is this that makes proximity current effect much more serious than skin effect.

A widely referenced classic paper by Dowell[13] analyzes proximity effect in transformers and derives curves showing the ratio of AC to DC resistance R_{ac}/R_{dc} as a function of the number of winding layers and the ratio of wire diameter to skin depth. A detailed summary of Dowell's results is beyond the scope of this text, but it is very well covered by Snelling.[12] A very good discussion of Dowell's curves, showing physically why R_{ac}/R_{dc} increases exponentially with the number of layers, is given by Dixon.[11]

Herein, Dowell's curves will be presented and a discussion of their use and significance on the basis of Dixon's treatment will be given.

Figure 7.8*a* shows an EE core with three primary layers. Each layer can be considered as a single sheet carrying a current $I = NI_t$ (say, 1 A), where N is the number of turns in the layer and I_t is the current per turn. Now recall Ampere's law which states that $\oint H\,dl = 0.4\pi I$ or the line integral of $H\,dl$ around any closed loop is equal to $0.4\pi I$, where I is the total current enclosed by the loop. This is the magnetic equivalent of Ohm's law, which states that the applied voltage to a closed loop is equal to the sum of all the voltage drops around that loop.

Now if the line integral is taken around the loop *abcd* in Fig. 7.8*b*, the magnetic reluctance (magnetic analog of resistance) along the path *bcda* is low as that path is in ferrite material which has very high permeability. Thus all the magnetic field intensity appears along the path *ab* which lies between sheets 1 and 2 and no magnetic field intensity lies along the leftmost surface of sheet 1. But since it is the

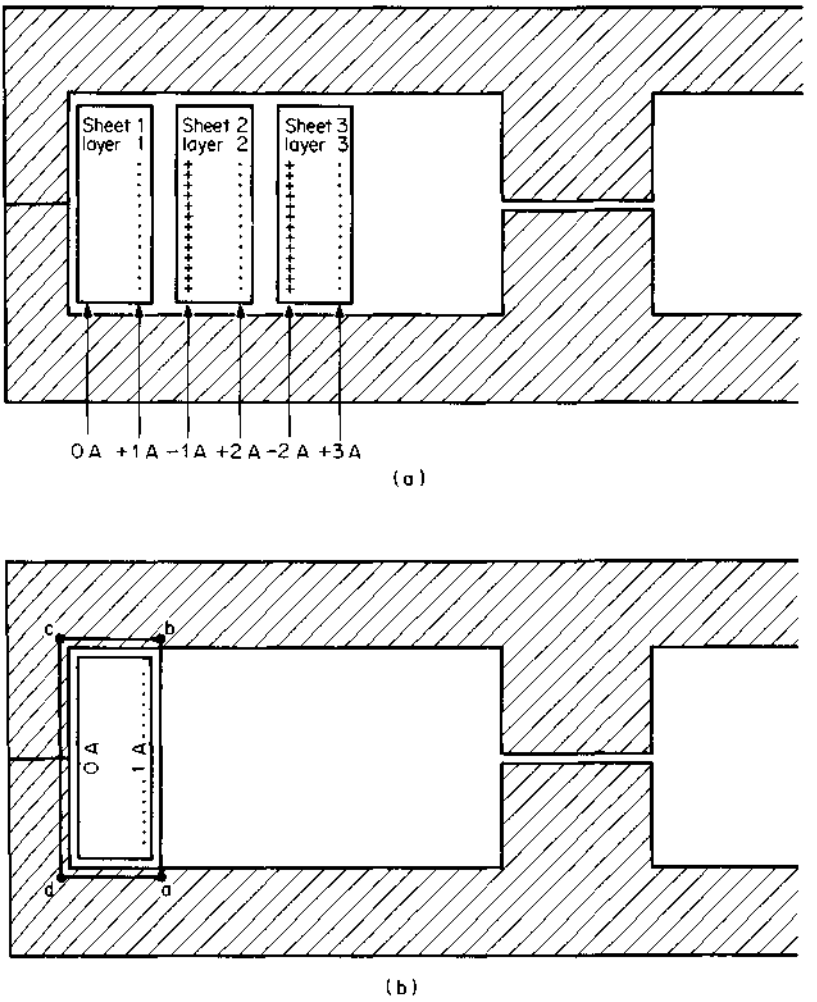


Figure 7.8 (*a*) Showing exponential buildup of surface eddy currents in a multilayer coil. (*b*) Current in the first layer is confined to a thin skin on its surface facing away from the ferrite material as dictated by Ampere's law.

magnetic field intensity along the surfaces which cause the skin currents to flow, all the current I which is carried by sheet 1 flows on its rightmost surface in say the plus direction (indicated by the dots) and no current flows on its leftmost surface.

Now consider the currents in sheet 2 (Fig. 7.8*a*). Proximity effect as described for Fig. 7.7 will cause eddy currents to flow on its left- and rightmost surfaces to a depth equal to the skin depth for that fre-

quency. Now the magnetic field intensity cannot penetrate more than a skin depth below the right-hand surface of sheet 1 or the left-hand surface of sheet 2.

Hence, if the integral $\oint H\,dl$ is taken around the loop *efgh* (through the centers of sheets 1 and 2), since there is zero field intensity along that path, the net current enclosed by that path must be zero by Ampere's law. Further, since the current on the rightmost surface of sheet 1 is 1 A in the plus direction, the current in the left-hand skin of sheet 2 must be 1 A, but in the minus direction (indicated by crosses).

However, the net current in each of the three sheets is 1 A. Hence, with a −1 current in the left skin of sheet 2, the current in its right-hand skin must be +2 A.

In a similar argument, the current in the left-hand skin of sheet 3 is −2 A, forcing the current in its right-hand skin to be +3 A.

It can thus be seen from this physical or intuitive reasoning that proximity effect causes eddy currents in the skins of a multilayer coil to increase exponentially with the number of layers. The Dowell[13] analysis covered in the next section verifies this quantitatively.

7.5.6.3 Proximity effect AC/DC resistance ratios from Dowell curves. Dowell's analysis[13] yields the widely referenced curves of Fig. 7.9. They show the ratio of AC/DC resistance ($F_R = R_{ac}/R_{dc}$) versus a factor

$$\frac{h\sqrt{F_l}}{\Delta}$$

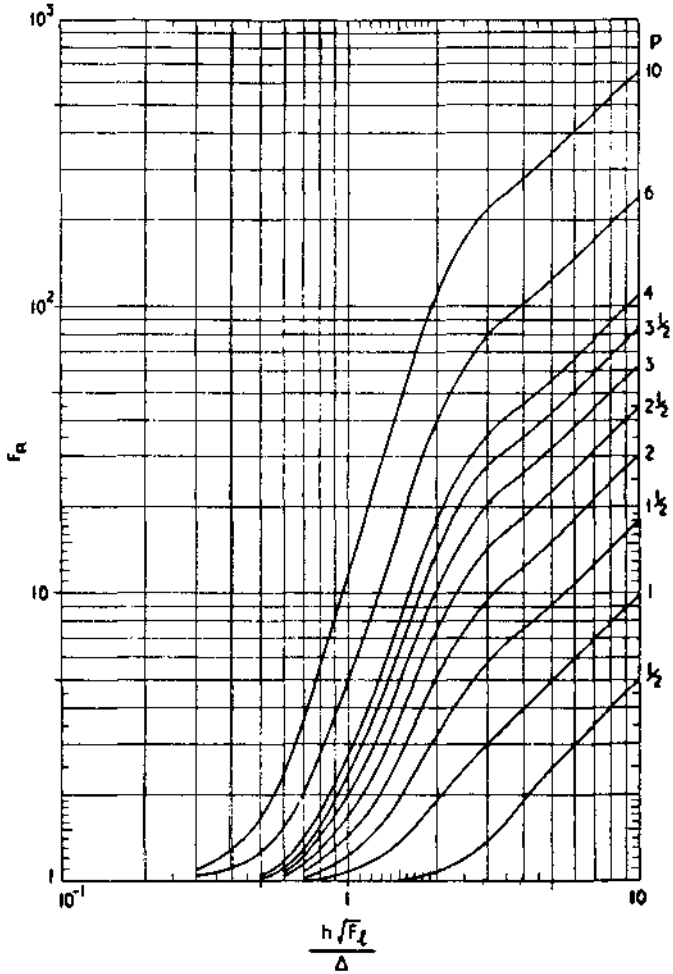
where h = effective round wire height = 0.866 (wire diameter d) = $0.866d$

Δ = skin depth (from Table 7.5)

F_l = copper layer factor = $N_l d/w$ (where N_l = number of turns per layer, w = layer width, d = wire diameter; note $F_l = 1$ for foil)

The ratio is given for a number of different values of a variable p, which is the number of coil layers per portion. A "portion" is defined as a region where the low-frequency magnetomotive force ($\oint H\,dl = 0.4\pi NI$) ranges from zero to a peak.

This "portion"—often misinterpreted—is clarified thus. Consider that the primary and secondary are both multilayer windings and stacked on the bobbin with the primary layers innermost, followed by the secondary layers on top as in Fig. 7.10*a*. Now moving outward from the bottom of the primary, the magnetomotive force ($\oint H\,dl = 0.4\pi NI$) increases linearly as shown in Fig. 7.10*a*.



F_R as a function of $h\sqrt{F_t}/\Delta$ with number of layers, p per winding portion as a parameter.

$$R_{AC} = R_{dc}F_R \quad When \frac{h\sqrt{F_t}}{\Delta} > 5 \quad F_R \to \frac{2p^2+1}{3} \cdot \frac{h\sqrt{F_t}}{\Delta}$$

Figure 7.9 Ratio of AC to DC resistance due to proximity effect. (*From Dowell, Ref. 13.*)

As the line integral is taken over an increasing distance out from the innermost primary layer, it encloses more ampere turns. Then at the secondary-primary interface, $\oint H\,dl$ has reached a peak and starts falling linearly. In a conventional transformer (unlike a flyback), the secondary ampere turns are always simultaneously in the direction opposite to that of the primary ampere turns. Or, stated otherwise,

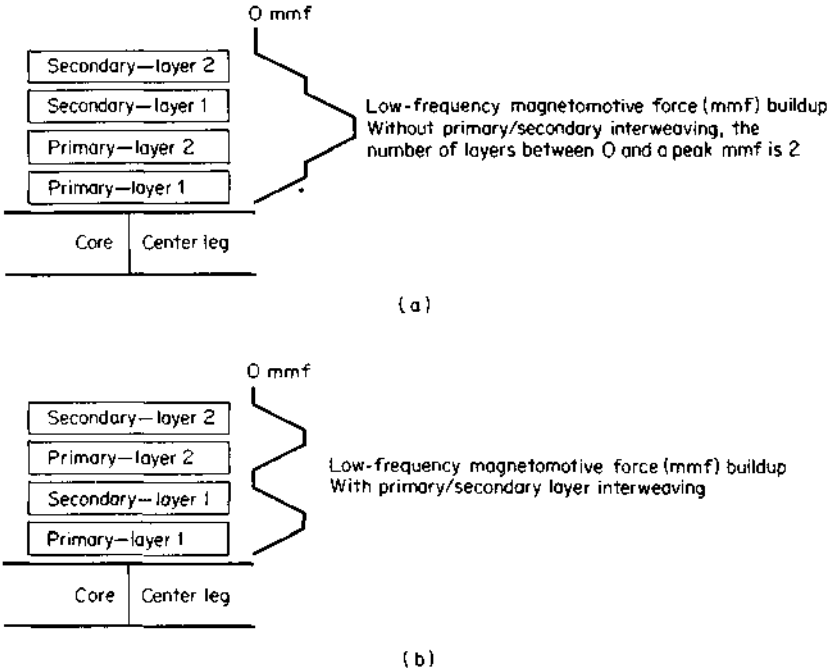


Figure 7.10 (*a*) Low-frequency magnetomotive force buildup in two-layer primary, secondary when primary and secondary layers are not interleaved. A "portion" is defined as the region between zero and a peak magnetomotive force. Here there are two layers per portion. See Fig. 7.9 for significance of "layers per portion." (*b*) By interleaving primary and secondary layers, the number of layers per portion has decreased to 1 and the AC/DC resistance ratio has decreased significantly (Fig. 7.9).

when current in a primary flows into a "dot" end, it flows in the secondary out of the dot end.

Now when the line integral is taken over the last secondary layer, $\oint H\,dl$ has fallen back to zero. This is just another way of stating "the total secondary ampere turns bucks out the primary ampere turns"—except for the small primary magnetizing current.

Thus a "portion" is the region between the zero and the peak magnetizing force and for two secondary and two primary layers, sequenced as in Fig. 7.10*a*, the number of layers per portion p is 2. In Fig. 7.9, for—say—a $(h\sqrt{F_l})/\Delta$ ratio of 4, R_{ac}/R_{dc} is about 13!

Now the number of ampere turns in each half secondary is equal to half the total ampere turns of the primary. If the two primary and secondary layers are interleaved as shown in Fig. 7.10*b*, the low-frequency magnetomotive forces are as shown in the associated diagram.

Now the number of layers between zero and a peak of magnetomotive force is only 1. For the same $(h\sqrt{F_l})/\Delta$ ratio of 4 (from Fig 7.9),

the ratio R_{ac}/R_{dc} per portion is now only 4 instead of 13! Thus the total AC resistance of either the primary or secondary is only 4 instead of 13 times its DC resistance.

Note in Fig. 7.9 that the number of layers per portion is shown going down to ½. The significance of ½ layer per portion can be seen in Fig. 7.11. There it is seen that if the secondary consisted of only one layer, the point at which the low-frequency magnetomotive force came back down to zero is half way through the thickness of the secondary layer. Thus Fig. 7.9 shows that for the same $(h\sqrt{F_l})/\Delta$ ratio of 4, R_{ac}/R_{dc} is 2 instead of 4 for the case if the single-layer secondary had not been interleaved between half primaries as in Fig. 7.11.

Figure 7.9 is very valuable in selecting a primary wire size or a secondary foil thickness at a rate other than the previously quoted "500 circular mils per rms ampere." That choice usually leads to large values of h/Δ at high frequencies and, as seen in Fig. 7.9, to very large values of R_{ac}/R_{dc}.

It is often preferable to choose a smaller wire diameter or foil thickness yielding a $(h\sqrt{F_l})/\Delta$ in the region of, say, 1.5. This, of course, would increase R_{dc}. But even though R_{dc} has increased, the smaller ratio of R_{ac}/R_{dc} may yield a lesser R_{ac} and lesser copper losses.

For the case of a push-pull circuit with two layer primaries and two layer secondaries, when interleaving, the simultaneously conducting half primary and secondary should be adjacent to each other as in Fig. 7.12*a*. Placing the nonconducting secondary adjacent to the conducting primary (Fig. 7.12*b*) would induce eddy currents in it even when it is nonconducting. Placing the nonconducting secondary outside the conducting one places it in a region where the half primary and half

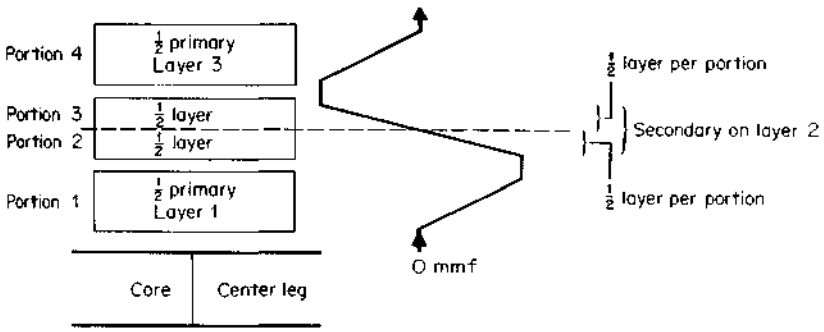


Figure 7.11 Half layers per portion (see Fig. 7.9). With a single-layer secondary sandwiched in between two half primaries, the ampere terms of a half primary are bucked out by half the current in the secondary. In the definition of "layers per portion" of Fig. 7.9, each half segment of the secondary operates at one-half layer per portion.

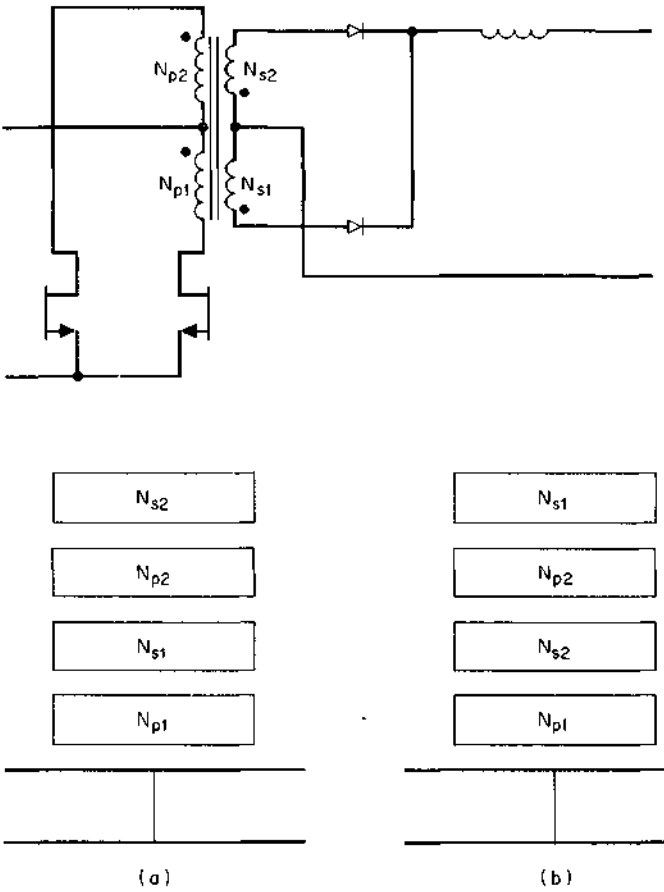


Figure 7.12 Correct (*a*) and incorrect (*b*) layer sequencing in a push-pull transformer. The sequencing of Fig. 2.12*b* will produce significantly more eddy current losses than that of Fig. 7.12*a* because of proximity effect.

secondary ampere turns cancel during conduction time at each half cycle. Then at that region, the line integral $\oint H\ dl$ is zero, magnetomotive force is zero and hence no eddy currents flow in it during the conduction time of the opposite half secondary.

Note that in a flyback circuit, primary and secondary currents are not simultaneous. Thus, interleaving of primary and secondary windings does not reduce proximity effect in flybacks. That can be done only by keeping the number of layers to a minimum and using finer wire than obtained from the rule of "500 circular mils per rms ampere." Although that increases DC resistance, it decreases R_{ac}/R_{dc}, from Fig. 7.9.

References

1. Ferroxcube-Philips Catalog, *Ferrite Materials and Components*, Saugerties, NY.
2. Magnetics Inc. Catalog, *Ferrite Cores*, Butler, PA.
3. TDK Corp. Catalog, *Ferrite Cores*, M. H. & W International, Mahwah, NJ.
4. Siemens Corp. Catalog, *Ferrites*, Siemens Corp., Iselin, NJ.
5. MMPA Publication PC100, *Standard Specifications for Ferrite Pot Cores*, MMPA, 800 Custer St., Evanston, IL.
6. MMPA Publication UE 1300 *Standard Specifications for Ferrite U, E And I Cores*, MMPA, 800 Custer St., Evanston, IL.
7. IEC Publications 133, 133A, 431, 431A, 647, American National Standards Institute, 1430 Broadway, New York.
8. A. I. Pressman, *Switching and Linear Power Supply, Power Converter Design*, pp. 116–120, Switchtronix Press, Waban, MA, 1977.
9. A. Kennelly, F. Laws, and P. Pierce, "Experimental Researches on Skin Effect in Conductors," *Transactions of AIEE*, 34: 1953, 1915.
10. F. Terman, *Radio Engineer's Handbook*, p. 30, McGraw-Hill, New York, 1943.
11. L. Dixon, *Eddy Current Losses in Transformer Windings and Circuit Wiring*, Unitrode Corp. Power Supply Design Seminar Handbook, Unitrode Corp., Watertown MA, 1988.
12. E. Snelling, *Soft Ferrites*, pp. 319–358, Iliffe, London, 1969.
13. P. Dowell, "Effects of Eddy Currents in Transformer Windings," *Proceedings IEE* (U.K.), 113(8): 1387–1394, 1966.
14. P. Venkatramen, "Winding Eddy Currents in Switchmode Power Transformers Due to Rectangular Wave Currents," *Proceedings Powercon 11*, 1984.
15. B. Carsten, "High Frequency Conductor Losses in Switchmode Magnetics," *High Frequency Power Converter Conference*, pp. 155–176, 1986.
16. A. Richter, "Litz Wire Use in High Frequency Power Conversion Magnetics," *Powertechnics Magazine*, April 1987.

Chapter

8

Bipolar Power Transistor Base Drive Circuits

8.1 Introduction

Over the past 5 to 8 years, bipolar power transistors have increasingly been replaced by MOSFET transistors in switching power supplies. New designs in the coming years will most frequently be done with MOSFETs.

There will nevertheless remain some niche areas (perhaps low power applications, because of their lower cost) which will continue using bipolars. Thus because of some small remaining areas where bipolar transistors still offer some advantages and because the vast majority of the switching supplies still operating in the field were originally done with bipolars, it is important for designers to be familiar with their significant characteristics, in the event field failures occur with these older but still operating designs.

The first considerations for the designer of a switching supply based on bipolar transistors are, of course, selection of a device with the proper voltage and current ratings. Maximum voltage and current stresses are dependent on the topology chosen, input voltage and its tolerances, and output power. Equations giving these stresses have been derived and are presented in the discussions of each topology.

However, the means and detailed design of the bipolar base drive circuit is as important for overall reliability as choice of a transistor with adequate voltage and current ratings. The general principles of what constitutes a good base drive circuit and some widely used techniques are discussed herein.

8.2 Objectives of Bipolar Base Drive Circuits

A good base drive circuit should have the criteria and parameters described in Secs. 8.2.1 to 8.2.6.

8.2.1 Sufficiently high current throughout the on time

Base input current should be adequate to keep the on collector-to-emitter potential at a sufficiently low level (0.5 to 1.0 V) for maximum output power at minimum input voltage for the lowest gain (lowest beta) transistor at the peak current instant during the on time.

Thus, design should allow for a four-to-one production spread in transistor beta. The usual current-voltage curves (I_{ce} vs. V_{ce}) in the manufacturer's data sheets (Fig. 8.1) are for a transistor of typical or nominal beta. It should be assumed that the minimum beta is one-half and the maximum beta is twice the typical one shown in the curves.

Power transistor currents in most topologies have the shape of a ramp starting from zero or a ramp on a step. Thus the input base current should be adequate to keep V_{ce} "bottomed" to about 0.5 to 1.0 V at the peak of the I_c ramp at maximum output power at minimum input voltage for the minimum beta transistor. This is especially true for discontinuous-mode flybacks where the ratio of peak to average current is high.

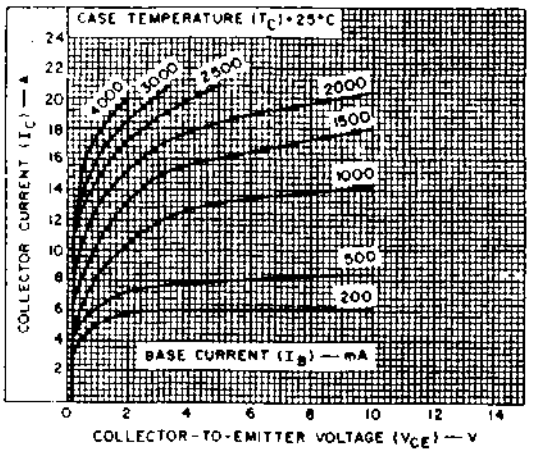


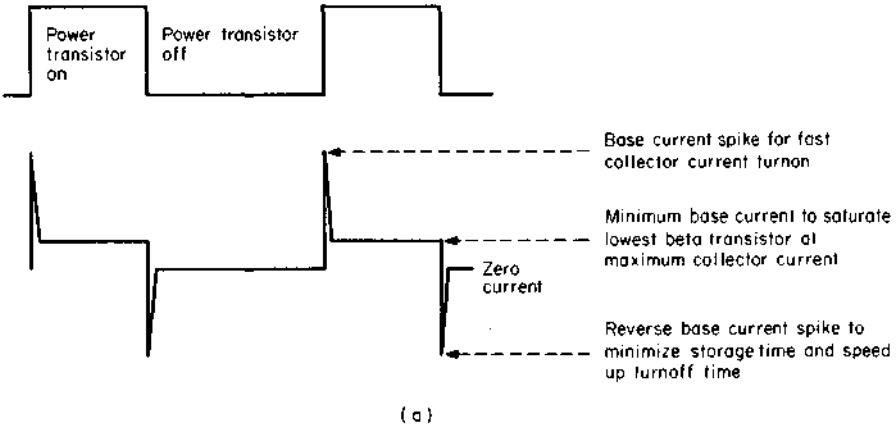
Figure 8.1 Curves of I_c versus V_{ce} for a typical high-voltage high-current transistor. 2N6676, 15 A, 450 V. (*Courtesy RCA*) Curves such as this are usually for a typical device. Depending on the manufacturer's production spread, lowest-beta device may have half the indicated beta; highest-beta device, twice that shown.

8.2.2 A spike of high base input current I_{b1} at instant of turnon

To ensure fast collector current turnon, there should be a short spike of base current about two to three times the average value during the on time. This spike need last only about 2 to 3 percent of the minimum on time (Fig. 8.2*a*).

The effect of this "turnon overdrive" can be seen in Fig. 8.2*b*. If turnon speed is not a factor, base input current I_{b1}) for a desired collector current (I_{c1})need be only that required to bottom V_{ce} to the saturation voltage $V_{ce(sat)}$ at the intersection of the collector load line and the I_c/V_{ce} curve.

At that I_{b1}, collector current will rise exponentially with some time constant t_a and get to within 5 percent of I_{c1} in three time constants $3t_a$.



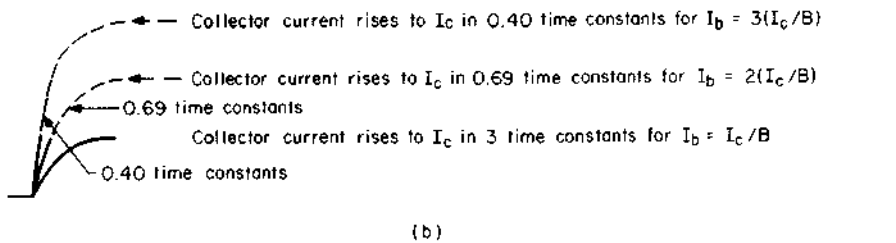


Figure 8.2 (*a*) Optimum base current waveforms. (*b*) Accelerating collector current rise time by base current overdrive.

If base input current is $2I_{b1}$ (overdrive factor of 2), however, collector current will rise as if it were heading for $2I_{c1}$ and if not limited by the load impedance and supply voltage, would reach $2I_{c1}$ in the same $3t_a$. But collector current is limited to I_{c1} by the supply voltage and load impedance. Hence the current reaches I_{c1} (the desired value) in $0.69t_a$ instead of $3t_a$.

Similarly, if an overdrive factor of 3 is used ($I_b = 3I_{b1}$), collector current will rise as if it were headed to $3I_{c1}$ and if not limited by the load resistance and supply voltage, would reach $3I_{c1}$ in the same $3t_a$. But since it is limited, it reaches the desired I_{c1} in $0.4t_a$ instead of $3t_a$.

Overdrive factors of 2 to 3 are usually used to speed up turnon time. This is an overdrive above the required base current for a nominal-beta transistor. Low-beta transistors are faster and do not require as high an overdrive factor. High-beta transistors are slower, and an overdrive of 2 for a nominal beta device corresponds to an overdrive of 4 for a high-beta transistor as its beta value is generally twice that of the nominal-beta transistor.

8.2.3 A spike of high reverse base current I_{b2} at the instant of turnoff (Fig. 8.2*a*)

If base input current is simply dropped to zero when it is desired to turn off, collector current will remain unchanged for a certain time (storage time t_s). Collector voltage will remain at its low $V_{ce(sat)}$ value of about 0.5 V and when it finally rises, will have a relatively slow rise time.

This comes about because the base-to-emitter circuit acts like a charged capacitor. Collector current keeps flowing until the stored base charges drain away through the external base-to-emitter resistor. There is generally a large excess of stored base charges because the base input current must be chosen sufficiently large to bottom the collector-to-emitter voltage to about 0.5 V for the lowest-beta transistor. Thus the highest or even the nominal-beta transistor has an excess of base current and long storage time.

A momentary spike of reverse base current I_{b2} is required to pull out the stored base charges. This reduces storage time and permits higher switching frequencies. It also significantly reduces power dissipation during the turnoff interval.

This can be seen (Fig. 8.3*a*) in the instantaneous I_c/V_{ce} curves during the turnoff interval. There it is seen (t_1 to t_2) that before V_{ce} starts rising very rapidly, it moves up out of saturation slowly while collector current hangs on at its peak value. During this interval, current is at its peak and collector-to-emitter voltage is considerably higher than its saturation level of 0.5 to 1.0 V.

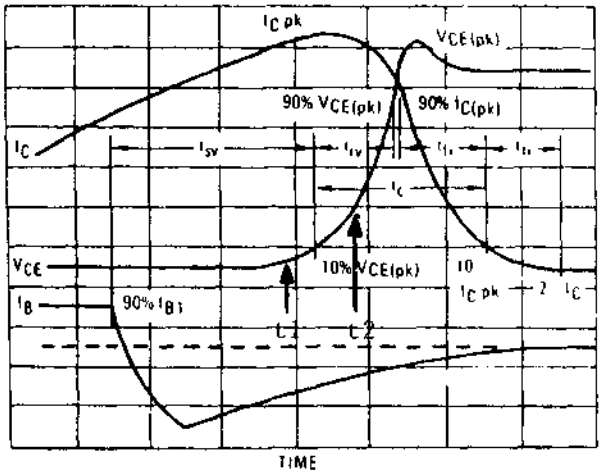


Figure 8.3 (*a*) Typical turnoff transition falling current and rising collector-to-emitter voltage for a power transistor; no snubber at collector. (*b*) Typical switching times for a typical high-current high-voltage transistor 2N6836, 15 A, 850 V (V_{cev}). (*Courtesy Motorola Inc.*)

This high spike of power dissipation can be a large fraction of the total dissipation in the transistor. The spike of reverse base current (Fig. 8.2*a*) reduces this dissipation by shortening the interval t_1 to t_2 and permits higher-frequency operation by reducing the storage time.

Manufacturers usually offer curves showing storage, rise time, and fall time for their power transistors for values of I_c/I_{b1} and I_c/I_{b2} ranging from 5 to 10 at various values of collector current (Fig. 8.3*b*, *c*).

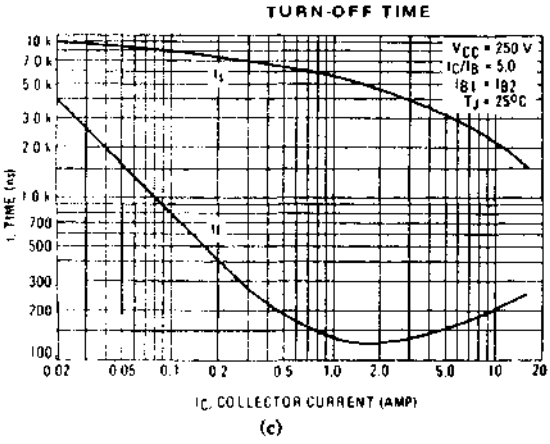


Figure 8.3 (*Continued*)

8.2.4 A base-to-emitter reverse voltage spike −1 to −5 V in amplitude at the instant of turnoff

Bipolar transistors have three significant collector-to-emitter voltage ratings: V_{ceo}, V_{cer}, and V_{cev}. The rating V_{ceo}, is the maximum collector-to-emitter voltage at turnoff when the base to emitter is open-circuited at the instant of turnoff. It is the lowest voltage rating for the device.

The transistor can tolerate a higher voltage (V_{cer} rating) during the off state if it has a "low" (usually 50 to 100 Ω) resistance from base to emitter.

The highest voltage the transistor can safely tolerate is its V_{cev} rating. This is the maximum voltage the transistor can tolerate at the instant of turnoff during the leakage inductance spike (Figs. 2.1, 2.10). It can tolerate this voltage only if there is a −1- to −5-V reverse voltage spike at the base during the instant of turnoff (Fig 8.4). This reverse-bias voltage or voltage spike must be supplied by the base drive circuit and must last at least as long as the leakage inductance spike.

8.2.5 A scheme to permit the circuit to work equally well with high- or low-beta transistors

Since production spread in beta may be four to one, if base input currrent is sufficient to safely turn on a low-beta transistor, it will greatly overdrive a high-beta transistor and result in excessive storage time. Storage times may be so long as to require unacceptably large reverse base currents to yield acceptably low storage times.

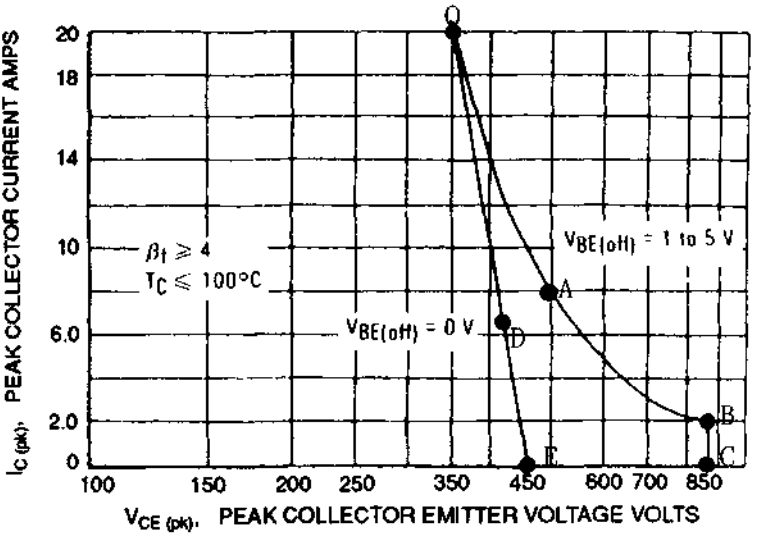


Figure 8.4 Reverse-bias safe operating area curves (RBSOA) for 2N6836. During turnoff, the I_c-V_{ce} locus must not cross the boundaries shown. Even a single crossing may destroy the transistor because of the current crowding into a small part of the chip area and causing local hot spots. With a −1- to −5-V reverse bias at the instant of turnoff, the V_{cev} boundary *OABC* applies. For V_{be} = 0 at turnoff, the boundary *ODE* applies. (*Courtesy Motorola Inc.*)

8.2.6 High efficiency

Since high-collector-currrent transistors generally have low beta, the base current driver may be called on to deliver high currents. If these currents come directly from high-voltage sources without the benefit of current gain through a voltage step-down transformer, efficiency will be low.

Thus a widely used fast power transistor—the 2N6836—is rated at 15 A, V_{cev} of 850 V. It is widely used in off-the-AC-line switching supplies. At 15 A, it has a minimum current gain of 5, requiring at least 3 A of base current.

If this current came from a 6-V source for a push-pull circuit at, say, 80 percent duty cycle of base drive, base drive dissipation (at the source) would be an unacceptably high 14.4 W (for a total duty cycle of 0.8). Thus a good base drive scheme should couple the drive pulse from a DC housekeeping voltage source through a voltage step-down–current step-up transformer.

8.3 Baker clamps

The transformer-driven Baker clamp[1–4] is a widely used base drive scheme. It is inexpensive in dollars, is low in component count, and provides all six features described in Secs. 8.2.1 to 8.2.6.

Since it is transformer-driven, it also nicely solves the problem of coupling a width-modulated pulse originating on output ground (where the PWM chip and housekeeping supply are best located for off-the-AC-line supplies) to the base of the power transistor on input ground (see Fig. 6.19 and Sec. 6.6.1).

Since a transformer is involved, it is relatively simple to get a voltage step-down–current step-up ratio of 10 or more. The secondary delivers its voltage to the base at a voltage of about 1 to 1.8 V and the primary takes its current from the housekeeping power supply, which is usually 12 to 18 V.

Thus in the preceding example, for a 10/1 voltage step-down, current step-up, the 3 A of base current at the transformer secondary is obtained at a cost of only 300 mA from its housekeeping supply located at output ground.

Baker clamping is done after the secondary of $T2$ in Fig. 6.19 between the collector and the base of the power transistor. The operation of a Baker clamp is discussed below.

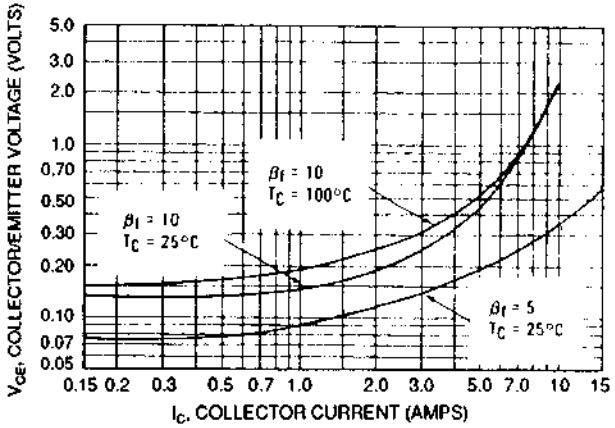
Normally an NPN transistor operates with its collector-to-base junction in reverse bias or with its collector positive with respect to its base. But when it is hard on and in saturation, the collector is negative with respect to its base and the base-to-collector junction is in forward bias, acting like a conducting diode.

This can be seen in Fig. 8.5*a* and 8.5*b* for the 2N6386, which is a fast 15-A, 450-V transistor. Figure 8.5*a* shows the on collector-to-emitter voltage at various collector currents for two values of forced beta $B_f (= I_c/I_b)$ and temperature. It is seen that V_{ce} depends strongly on I_c, B_f, and temperature; but at usual operating conditions of I_c = 10 A, B_f = 5, and 100°C junction temperature, the on V_{ce} potential is about 0.2 V. Figure 8.5*b* shows that at 10 A, the on base-to-emitter potential is about 0.9 V at 100°C.

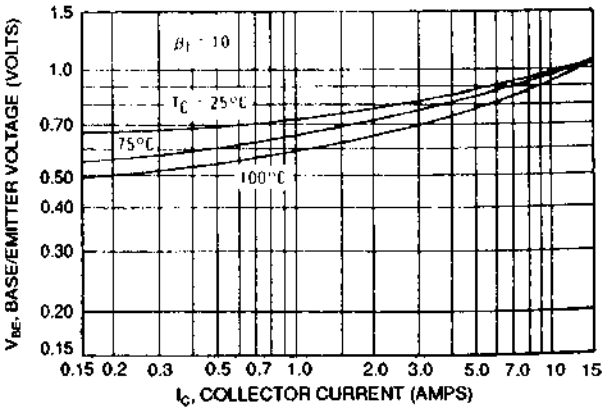
With the resulting 0.7-V forward bias on the base-to-collector junction, there is an excess of stored base charges. Further, when the base current is simply reduced to zero, storage time for this very fast bipolar transistor is still a very long 3 μs (Fig. 8.5*c*).

Baker clamping corrects this problem by not permitting the base-to-collector junction to take on a forward bias or at the worst, by allowing only a 0.2- to 0.4-V forward bias—low enough to prevent significant storage times.

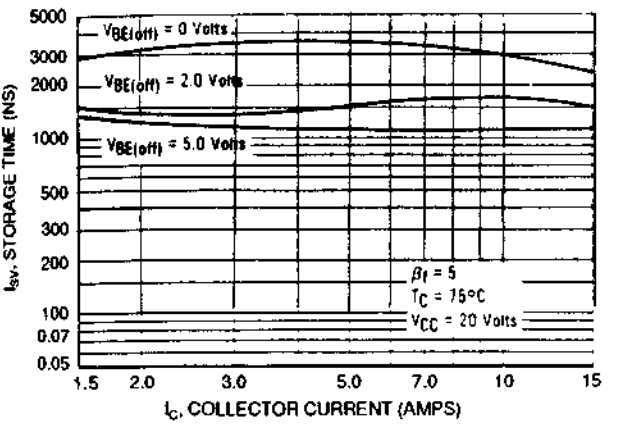
The Baker clamp can reduce storage times by a factor of 5 to 10.[4] It works nicely over a large temperature and collector current range, and very importantly, it permits equally good circuit operation with transistors whose production spread in beta can be as large as 4 to 1. Its operation is described in the following section (see also Fig. 8.6).



(a)



(b)



(c)

Figure 8.5 (*a*, *b*) Junction voltages for 2N6836 transistor at two values of forced beta. (*c*) Storage time for 2N6836 without Baker clamping. (*Courtesy Motorola Inc.*)

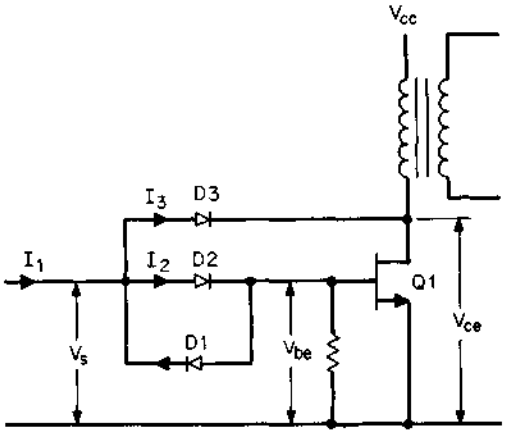


Figure 8.6 Baker clamping. The objective of the Baker clamp is to prevent a forward bias on the $Q1$ collector-to-base junction during the on time and hence minimize or eliminate storage time. If the voltage rise across $D2$ equals the drop across $D3$, V_{ce} is exactly equal to V_{be}. Small variations in the $D2$, $D3$ forward drops permit a small forward bias on the collection-to-base junction, but not enough to result in significant storage times. Currents I_2 and I_3 redistribute themselves by negative feedback so that I_2 is sufficient only to keep V_{ce} low enough to keep $D3$ in forward bias. The balance of I_1 ($= I_3$) flows through $D3$ and the $Q1$ collector to ground.

8.3.1 Baker clamp operation

In Fig. 8.6, a large current I_1 of the desired pulse width is provided at the anode of $D2$. The current is large enough and has a sufficient overdrive to turn on the maximum current in $Q1$ with the desired speed when $Q1$ is a minimum beta transistor. As $Q1$ commences turning on, $D3$ is reverse-biased, draws no current, and is effectively out of the circuit. All the I_1 flows through $D2$ into the base, yielding very fast rise time.

However, when the collector voltage has fallen low enough to forward-bias $D3$, the current I_1 redistributes itself. Only that fraction of it I_2 flows through $D2$ into the $Q1$ base to keep $D3$ forward-biased. The balance I_3 flows around through $D3$ into the $Q1$ collector and then through the emitter to ground.

The circuit now operates with negative feedback. As load current changes or as transistors with different beta are used, the $Q1$ base demands from I_1 only enough current through $D2$ to keep $D3$ forward-biased. Since the forward drop in $D2$ and $D3$ changes only by a few

tens of millivolts over very large forward current changes, the potential at the $Q1$ collector does not change significantly.

Now consider the $Q1$ junction potentials. Diode $D3$ must be a high-voltage fast-recovery diode as it is subjected to twice the supply voltage plus a leakage inductance spike. Diode $D2$ is a type of the same speed but is never subjected to a reverse voltage of greater than about 0.8 V (forward voltage of $D1$). Thus, assume that $D3$ is an MUR450 (450 V, 3 A, and 75 ns of recovery time) and $D2$ is an MUR 405 (50 V, 4 A, 35 ns of recovery time).

Assume for the moment that the $D2$, $D3$ forward voltages are independent of forward current and temperature and are equal to 0.75 V. This is a good enough approximation as seen in Table 8.1. The small variations are not sufficient to change the reasoning described below.

Thus in Fig. 8.6, when $D3$ latches in and conducts, the voltage rise in $D2$ is closely equal to the drop in $D3$ and is assumed to be 0.75 V. Then for V_{be} of 1.0 V in $Q1$, its base is at +1.0 V with respect to ground, the $D2$ anode (V_s) is at +1.75, and the $Q1$ collector is back down to 1.0 V. Thus there is no forward bias on the $Q1$ base-to-collector junction and negligible storage time. As temperature rises, the forward drop in $D2$ decreases, but so does the drop in $D3$ and the collector-to-base junction still has no forward bias.

Now assume that I_1 is 3.5 A and that $Q1$ is a maximum beta transistor and requires only 0.5 A of base current, leaving 3.0 A for $D3$. Table 8.1 shows the rise in $D2$ as 0.61 V at 100°C and the drop in $D3$ at 3 A, 100°C as 0.95 V. This leaves a forward bias of only 0.34 V across the $Q1$ base-to-collector junction—not enough to cause diode-type conduction in it. Storage time is still negligible at that forward bias.

Note that Baker clamping holds the collector-to-emitter potential to about 1 V as compared to the 0.2 to 0.5 V without the Baker clamp. This increases transistor losses during the on time, but the decreased AC loss during the transition from on to off more than makes up for it (Sec. 8.2 and Fig. 8.3*a*).

TABLE 8.1 Forward Voltage of Two Ultra-Fast-Recovery Diodes (*D2*, *D3* in Fig. 8.6)

	MUR450 V_f, V		MUR405 V_f, V	
I_f, A	25°C	100°C	25°C	100°C
0.5	0.89	0.75	0.71	0.61
1.0	0.93	0.80	0.74	0.65
2.0	1.01	0.90	0.78	0.70
3.0	1.10	0.95	0.80	0.73

Thus the Baker clamp has satisfactorily solved two significant problems. It prevents a sufficient forward bias on the base-to-collector junction to cause appreciable storage time. It also permits the circuit to work equally well with large changes in load current and over a large production spread in transistor beta because of the redistribution of input currents between $D2$ and $D3$ as base current demands change.

However, it is still desired to provide reverse base current to $Q1$ at the instant of turnoff to speed up turnoff time and provide a −1- to −5-V reverse base bias. This is prevented by the blocking action of $D2$, but by adding the "reach-around" diode $D1$, across $D2$, it becomes possible. It now only remains to find a low-parts-count scheme to provide the large forward current I_{b1} for turnon, an equally large reverse current I_{b2}, and a reverse voltage bias at the base for turnoff. This is easily achieved with the transformer-coupled scheme of the following section.

8.3.2 Transformer coupling into a Baker clamp

8.3.2.1 Transformer supply voltage, turns ratio selection, and primary and secondary current limiting. The circuit of Fig. 8.7 provides all the required drive characteristics for the Baker clamp—high forward and reverse base drive for $Q2$ at relatively low primary current drawn from the housekeeping supply V_h. It also provides the reverse $Q2$ base voltage which permits it to tolerate its V_{cev} rating. It works as follows.

First, the $T1$ turns ratio N_p/N_s is chosen as large as conveniently possible so as to provide the desired secondary current with a reasonably low primary current and voltage. Since the primary current will be taken from the usual housekeeping supply V_h, which also feeds the PWM chip, V_h should be kept low to keep chip dissipation low.

Choosing a high N_p/N_s ratio to get a large current gain in $T1$ may also force too high a value of V_h. Thus a reasonable choice for V_h is the 15 to 18 V usually selected for a PWM chip. This largely fixes N_p/N_s. It will soon be seen that the $T1$ primary voltage V_p should be considerably less than V_h because of $R1$, which plays a significant part in the circuit.

The $T1$ secondary voltage V_s is clamped to $V_{be(Q_2)}$ plus the $D2$ forward drop V_{D2}. Or, in other words, $V_s = V_{be(Q2)} + V_{D2} = 1.0 + 0.75 = 1.75$ V. The voltage at the top of the $T1$ primary is

$$V_{pt} = \frac{N_p}{N_s} V_s + V_{ce(Q1)}$$

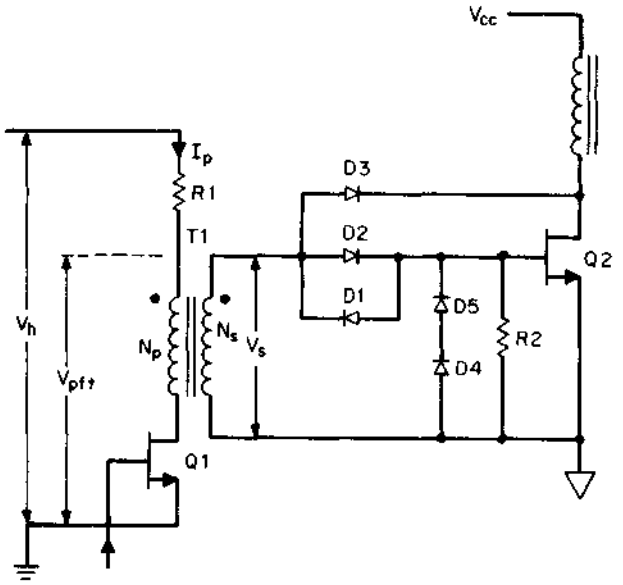


Figure 8.7 A transformer-driven Baker clamp. Transformer $T1$ provides a large current gain from primary to secondary so that secondary currents of 2 to 3 A of secondary current can be obtained from primary currents of about 600 mA, permitting use of small, inexpensive transistors in TO18 cans for $Q1$. By storing magnetizing current in N_p, large reverse currents are available in N_s at the instant of $Q2$ turnoff. Resistor $R1$ serves as a primary current limiter.

$$= \frac{N_p}{N_s} 1.75 + V_{ce(Q1)} \quad \text{V}$$

$$\approx \frac{N_p}{N_s} 1.75 + 1.0 \text{ V} \tag{8.1}$$

Now V_h will be kept constant by an inexpensive linear regulator fed from the secondary of a small 60-Hz transformer returned to output ground as in Fig. 6.19.

However, V_{pt} should be kept considerably lower than V_h so as to provide a relatively constant voltage across $R1$ (V_{R1}) when V_s varies as a result of temperature-induced variations in V_{be} and V_{D2}. The reason for keeping V_{R1} constant is that $R1$ serves to limit primary current from

$$I_{p(Q1)} = \frac{V_h - V_{pt}}{R_1}$$

$$= \frac{V_h - (N_p/N_s)V_s - 1.0}{R_1} \tag{8.2}$$

Thus $R1$ is chosen to limit $T1$ primary and hence its secondary current and current into $D2$. Although negative feedback through the Baker clamp diode $D3$ allots current to the $Q2$ base only sufficient to supply the maximum collector current and to keep $D3$ in conduction, an excess current into $D2$ from the $T1$ secondary is simply wasted by being diverted via $D3$ into the $Q2$ collector.

Thus by choosing V_h large compared to $(N_p/N_s)V_s$, $R1$ becomes more of a constant-current source relatively independent of temperature variations in V_s.

For an initial guess, assume $N_p/N_s = 5$. For a nominal $V_s (= V_{sn})$ of 1.75 V, the nominal I_p is

$$I_{pn} = \frac{V_h - 5V_{sn} - 1.0}{R_1}$$

and the change in $I_{pn} (= dI_{pn})$ is $dI_{pn} = 5\,dV_s/R_1$, where dV_s is the anticipated change in V_s due to temperature changes. The fractional change in I_{pn} is

$$\frac{dI_{pn}}{I_{pn}} = \frac{5\,dV_s}{V_h - 5 \times 1.75 - 1.0}$$

$$= \frac{5\,dV_s}{V_h - 9.75} \tag{8.3}$$

Then, for an anticipated temperature variation dV_s of 0.1 V, and a permissible fractional change of 0.1 in I_{pn}, from Eq. 8.3 we obtain

$$0.1 = \frac{5 \times 0.1}{V_h - 9.75}$$

or

$$V_h \cong 14.75 \text{ V}$$

$$= 15.0 \text{ V}$$

Thus if it were desired to limit the $Q1$ primary current to I_{pn}, $R1$ would be chosen as

$$R_1 = \frac{15 - (5 \times 1.75) - 1.0}{I_{pn}} \tag{8.4a}$$

$$R_1 = \frac{5.25}{I_{pn}} \tag{8.4b}$$

8.3.2.2 Power transistor reverse base current derived by flyback action in drive transformer. A large reverse current to the base of power transistor $Q2$ (Fig. 8.7) can be obtained by choosing a low magnetizing inductance in the $T1$ primary. During the $Q1$ on time, total current to the $T1$ primary is limited by $R1$. Part of that current is multiplied by the N_p/N_s turns ratio and delivered to the secondary to turn the $Q2$ base on.

But part of that current flows to the primary magnetizing inductance L_m and does not contribute to the secondary current. It simply ramps up linearly at a rate $dI_m/dt = (V_{pt} - 1)/L_m$. At the end of the on time t_{on} it has reached a peak $I_{pm} = (V_{pt} - 1)t_{on}/L_m$ and is stored as energy in the magnetizing inductance.

Now when $Q1$ is turned off, that magnetizing current I_{pm}, multiplied by the turns ratio N_p/N_s, is delivered by flyback action as a negative-going pulse to the secondary (note the $T1$ primary and secondary dots). At the secondary, it pulls reverse current from the base of $Q2$ through reach-around diode $D1$.

After the base current charges have been fully swept out, the base impedance is very high. As there usually is significant energy left in the $T1$ secondary, the $Q2$ base may be pulled sufficiently far negative to damage or destroy it. This is prevented by the two series diodes $D4$, $D5$ which clamp the $Q2$ base to a negative bias of about 1.6 V—far enough negative to permit $Q2$ to sustain its V_{cev} rating.

8.3.2.3 Drive transformer primary current limiting to achieve equal forward and reverse base currents in power transistor at end of the on time. Significant $T1$ current waveshapes are shown in Fig. 8.8. Current through $R1$ is shown in Fig. 8.8. For $V_h = 15$ V and $N_p/N_s = 5$, the nominal peak current is given by Eq. 8.4b as $I_{pn} = 5.25/R_1$.

Now if the $Q2$ base reverse current at the instant of turnoff is to be equal to its forward base current just the instant before turnoff, the $T1$ primary magnetizing current at the end of on time should be permitted to ramp up linearly to $I_{pn}/2$ by proper choice of its magnetizing inductance.

The $Q1$ magnetizing current is shown in Fig. 8.8*b*; $Q1$ primary load current is then the difference between Fig. 8.8*a* and 8.8*b* and is as shown in Fig. 8.8*c*. The $T1$ secondary current during the $Q2$ on time is then the $T1$ primary current of Fig. 8.8*c* multiplied by N_p/N_s. Its amplitude at the end of the on time, $(N_p/N_s)I_{pn}/2$, is chosen to turn on the maximum $Q2$ collector current for the minimum-beta transistor. Its amplitude at the start of the on time is twice that at its end, and consequently turnon time is very fast.

At the end of the $Q1$ on time, $T1$ magnetizing current in L_m is $I_{pn}/2$. At the very instant of $Q1$ turnoff, this reflects into the $T1$ secondary as

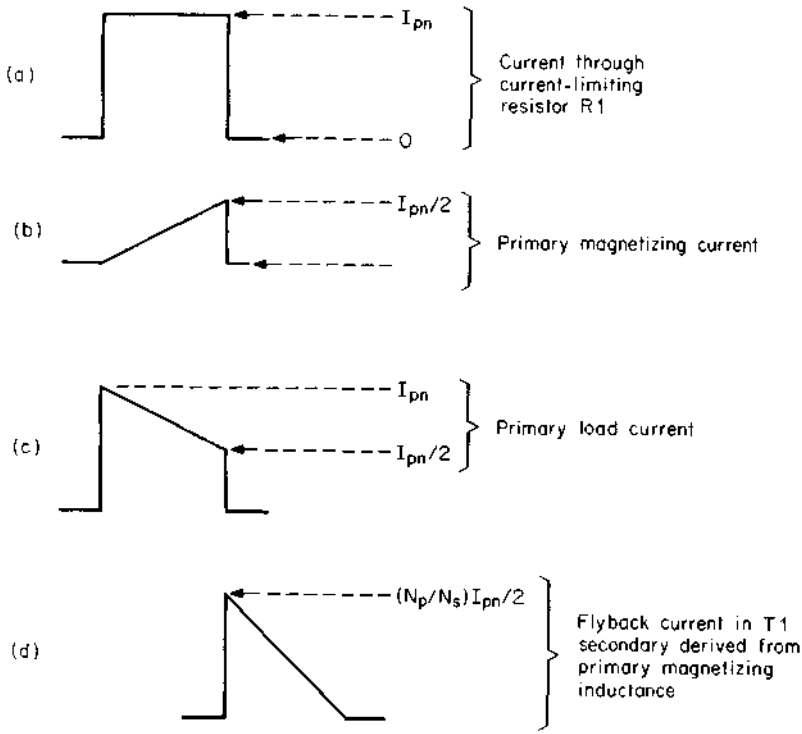


Figure 8.8 By appropriately choosing primary magnetizing inductance of $T1$, the reverse base current to $Q2$ is made equal to its forward current by flyback action in $T1$.

$(N_p/N_s)I_{pn}/2$ and provides a $Q2$ reverse base current (Fig. 8.8*d*) equal to its forward current just prior to turnoff.

8.3.2.4 Design example—transformer-driven Baker clamp. Assume a 500-W forward converter operating from the rectified 115-V-AC line. Assume a nominal rectified DC input of 160 V at minimum load and a minimum rectified DC input at maximum load of 0.85 × 160 = 136 V. From Eq. 2.28, peak primary current is

$$I_{pft} = \frac{3.13P_o}{V_{dc(min)}} = \frac{3.13 \times 500}{136} = 11.5 \text{ A}$$

Assume that $Q2$ is a 2N6836—a 15-A, 450-V (V_{ceo}) device with a V_{cev} rating of 850 V. Its minimum beta at 10 A is 8—assume 7 at 11.5 A. Maximum base current is 11.5/7 = 1.64 A. For an N_p/N_s ratio of 5 in $T1$, primary load current is 1.64/5 = 0.328 A.

Also from Fig. 8.8*a*, current in $R1$ (Fig. 8.7) must be twice that so as to store 0.328 A of $T1$ magnetizing current for the $Q2$ reverse base drive. Then from Eq. 8.4b, for $V_h = 15$ V, $R1 = 5.25/2 \times 0.328 = 8.0\ \Omega$.

The $T1$ magnetizing inductance must permit a peak magnetizing current of 0.328 A at the end of the on time. Calculate the magnetizing inductance from the minimum on time of $Q1$.

Assume a switching frequency of 50 kHz and a maximum on time at minimum AC input of $0.8T/2$ or 8 μs. For ±10 percent variation in AC input, the minimum on time of $Q1$ will be $0.8 \times 8 = 6.4$ μs. Then

$$L_{m(T1)} = \frac{(N_p/N_s)V_s t_{on(min)}}{0.328}$$

$$= \frac{5 \times 1.75 \times 6.4 \times 10^{-6}}{0.328}$$

$$= 171\ \mu H$$

Use a Ferroxcube 1408PA3153C8 core. It is small, pill-box in shape, with diameter = 0.551 in and height = 0.328 in. It has an A_l (inductance in millihenries per 1000 turns) of 315 (Fig. 8.9). For 0.171 mH, the required number of turns is $N_p = 1000\sqrt{0.171/315} = 23$ turns—say, 25 and for a turns ratio of 5, $N_s = 5$ turns. The magnetizing ampere-turns product is $0.328 \times 25 = 8.2$ ampere turns. Figure 8.9 shows that the cliff at which this specific core ($A_l = 315$) starts saturating is at about 12 ampere turns, which is safe.

Transistor $Q1$ need carry a peak current of only $2 \times 328 = 656$ mA.

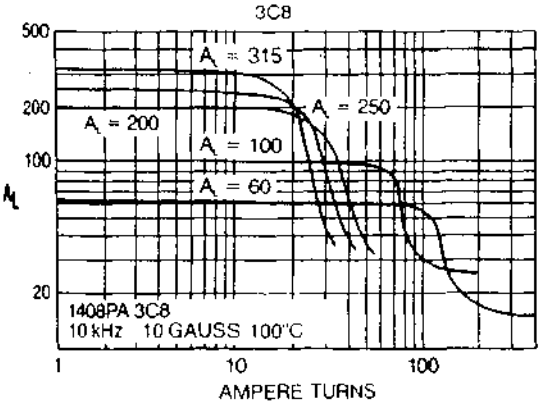


Figure 8.9 Inductance per 1000 turns A_l for Ferroxcube 1408PA3C8 pot core. A small core suitable for the transformer $T1$ of Fig. 8.7.

It can be a 2N2222A—an 800-mA 40-V device whose rise and fall times are under 60 ns. It comes in a small TO18 package and cost is under 25¢!

8.3.3 Transformerized Baker clamp[5]

By changing the circuit of Fig. 8.7 to the simpler one of Fig. 8.10, greatly improved performance with all the advantages of Baker clamping is achieved. Current gain in $T1$ can be doubled without increasing V_h, and better performance over a wider temperature range results. Also, the problem inherent in Fig. 8.7—that the forward rise in $D2$ does not track with the drop in $D3$ over large current changes—no longer exists. The circuit of Fig. 8.10 works as follows. The secondary of $T1$ in Fig. 8.10 is now center-tapped ($N_{s1} = N_{s2}$). Thus

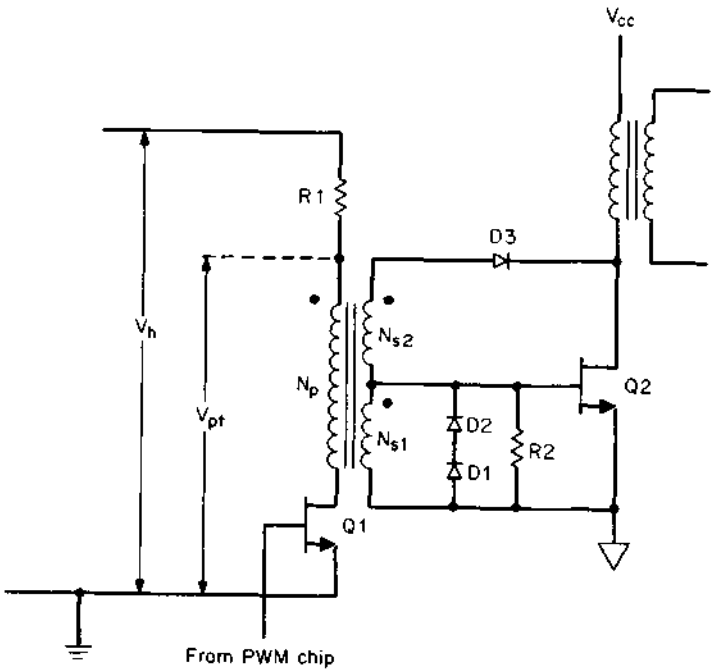


Figure 8.10 Transformerized Baker clamp. An improvement over Fig. 8.7; $N_{s1} = N_{s2}$. Thus voltage rise across N_{s2} is very closely equal to drop across $D3$ and $Q2$ collector is at same potential as its base and to a close approximation tracks it as temperature varies. With $V_{Ns1} = 1.0$ V, N_p/N_s can be 10 rather than 5 for the circuit of Fig. 8.7 at roughly the same value of V_h as for Fig. 8.7. Now current gain between N_p and N_{s1} is 10, and compared to Fig. 8.7, $Q1$ current is halved for the same $Q2$ base current.

$V_{Ns1} = V_{Ns2} = V_{be(Q2)}$, and when $D3$ has latched in and is conducting, we obtain

$$\begin{aligned} V_{ce(Q)} &= V_{Ns1} + V_{Ns2} - V_{D3} \\ &= 2V_{be(Q2)} - V_{D3} \\ &= 2 \times 1.0 - 1.0 = 1.0 \qquad \text{(Table 8.1, Fig. 8.5}b\text{)} \end{aligned}$$

Since $V_{be} = 1.0$ also, there is no forward bias across the base-to-collector junction and storage time is minimized.

There are, of course, changes in V_{be} and V_{D3} with current and temperature, but since there is only one diode involved as compared to two for Fig. 8.7, the maximum forward bias on the $Q2$ collector-to-base junction is considerably less than that for Fig. 8.7.

Now that $D2$ is no longer needed (it was needed only in Fig. 8.7, to provide a forward-voltage rise equal to the $D3$ forward drop), the reach-around diode $D1$ is not needed.

The greatest advantage of Fig. 8.10 is that now since V_s is approximately 1.0 rather than the 1.75 V of Fig. 8.7, the turns ratio N_p/N_{s1} can be roughly doubled without increasing V_h. Thus the same $Q2$ base current can be delivered at about half the current in $Q1$ and the $T1$ primary. The advantage of Fig. 8.10 can be appreciated by repeating the design example of Sec. 8.3.2.4 with the Fig. 8.10 circuit as below.

8.3.3.1 Design example—transformerized Baker clamp. In Fig. 8.10, $V_{pt} = (N_p/N_{s1})V_{Ns1} + V_{ce(Q1)}$. Now choose $N_p/N_{s1} = 10$ rather than 5 as for the Fig. 8.7 circuit. With a larger $T1$ turns ratio, I_{Q1} will be smaller. Hence assume that $V_{ce(Q1)}$ will be 0.5 V rather than the 1.0 V assumed for the Fig. 8.7 circuit. Then $V_{pt} = 10 \times 1.0 + 0.5 = 10.5$ V, and keeping the same nominal 5.25 V across $R1$ (Eq. 8.4b), $V_h = 15.75$ V. For the same 1.64 A to the $Q2$ base as in the previous design example, with $N_p/N_{s1} = 10$, $Q1$ primary load current is now only 164 mA.

Also as in Sec. 8.3.2.4, choose the $T1$ magnetizing current at the end of the on time to equal the primary load current (Fig. 8.8). Thus $R1$ is chosen to limit current to 328 mA and $R_1 = 5.25/0.328 = 16\ \Omega$. And with 10 V across the $T1$ primary, for the same 6.4 μs, the magnetizing inductance $L_m = (10 \times 6.4 \times 10^{-6})/0.164 = 390\ \mu H$.

For the same Ferroxcube 1408PA3153C8 core with an A_l value of 315 mH per 1000 turns, the required number of primary turns is $1000\sqrt{0.390/315} = 35$. For a turns ratio of 10 N_{s1}, N_{s2} would, of course, require 3.5 turns.

Half turns are possible with pot cores, but introduce other odd, undesirable effects. Hence choose $N_{s1} = N_{s2} = 4$ turns and $N_p = 40$

turns. This makes the magnetizing inductance $(40/35)^2 \times 390$ or 509 μH and the peak magnetizing current equal to (390/509)164 = 126 mA. For a 10/1 turns ratio, the reverse base current to $Q2$ is now 1.26 instead of 1.64 A. This would still yield a sufficiently low storage time.

The number of magnetizing ampere turns in $T1$ is now $0.126 \times 40 = 5.04$ ampere turns—still safely below the saturation knee for the $A_l = 315$ core of Fig. 8.9.

8.3.4 Inherent Baker clamping in Darlington transistors

In a Darlington transistor, the output transistor $Q2$ is automatically Baker clamped by the base-emitter diode of the drive transistor $Q1$ acting as $D2$ of Fig. 8.7 and the base-collector diode of the drive transistor acting as $D3$ in Fig. 8.7. This can be seen in Fig. 8.11*a* to 8.11*c*.

Thus the output transistor in a Darlington has negligible forward bias on its base-collector junction and should have low storage times. But data sheets on integrated-circuit Darlingtons show storage times up to 3 and 4 μs. This is due mainly to storage times in the Darlington drive transistor, which does saturate and has a forward-biased base-collector junction.

If a Darlington configuration with lesser storage time is needed, the 3- to 4-μs storage time can be lowered by using discrete transistors for the drive and output transistors.

By implementing the Darlington with discrete drive and output transistors, the drive transistor can be chosen as an ultra-high-frequency device which can have minimized storage time despite its forward-biased base-collector junction.

Most integrated-circuit Darlingtons have a built-in reach-around

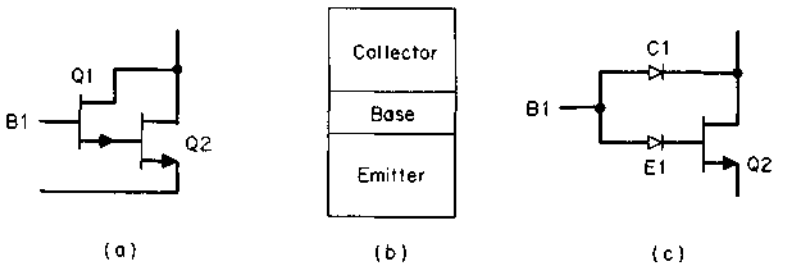


Figure 8.11 In a Darlington configuration, output transistor $Q2$ is inherently Baker clamped by the base-collector diode of $Q1$ acting as $D3$ and the base-emitter diode of $Q1$ acting as $D2$ in Fig. 8.7. Storage time in a Darlington configuration is due to saturation of the Darlington driver which is not Baker clamped. Figure 8.11*b* shows the junctions in a junction transistor as $Q1$.

diode as $D1$ in Fig. 8.7 which permits pulling reverse current from the output transistor base to improve its switching time.

8.3.5 Proportional base drive[2–4]

This base drive scheme (shown in Fig. 8.12) is widely used for high output power or power transistor currents over about 5 to 8 A.

The circuit does not attempt to keep the power transistor from saturating by use of Baker clamping. Rather, it ensures a large reverse base current to minimize storage time; and it always generates a base drive current proportional to its output current.

Thus even with a high base current required for high output currents, when output current is reduced from maximum to minimum, so is the input base current. Consequently, the base is never overdriven at low output currents and storage time is kept reasonably low throughout the load current range.

A particularly valuable feature is that the base current is obtained by positive feedback from the collector. For large output currents which require large base currents, this results in far less base source dissipation than if the base currents were derived from some housekeeping supply.

Circuit details are described in the following paragraphs.

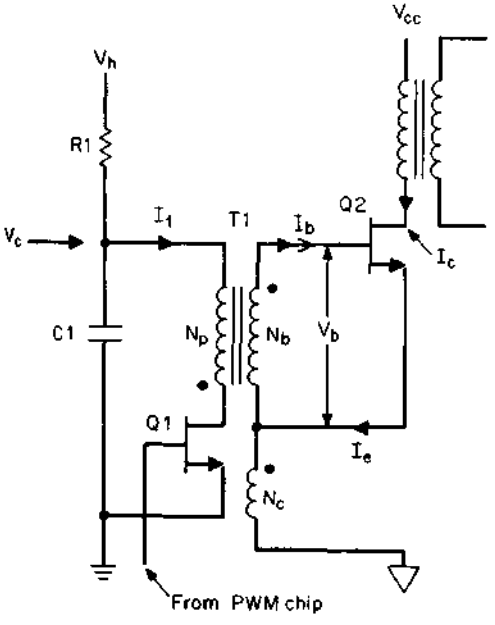


Figure 8.12 Proportional base drive. When $Q1$ turns off, the magnetizing current stored in N_p provides a short impulse to turn $Q2$ on by flyback action in $T1$. Thereafter, positive feedback from N_c to N_b holds $Q2$ on. The ratio N_b/N_c is chosen equal to $Q2$ minimum beta. When $Q1$ turns on, it couples a negative impulse to N_b which starts a regenerative turnoff sequence between N_c and N_b.

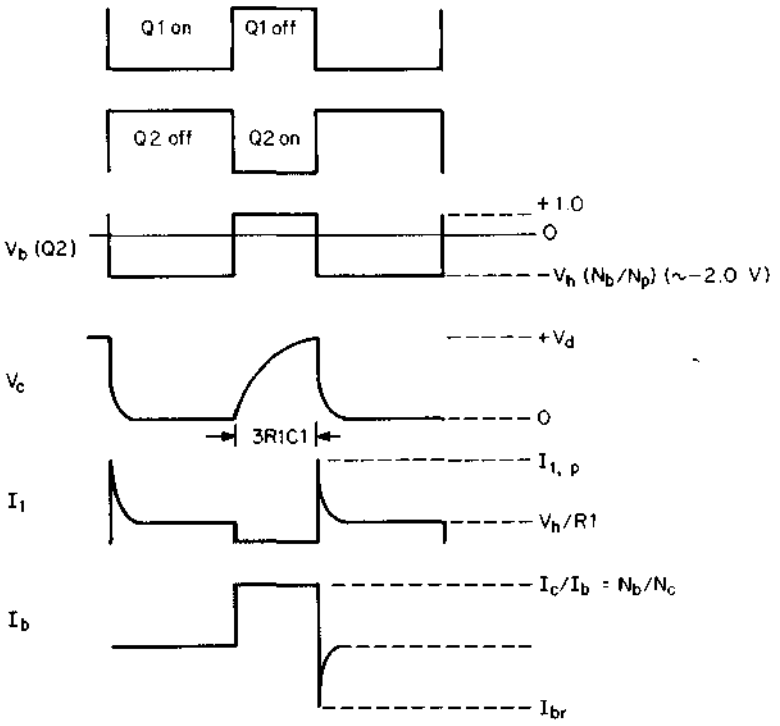


Figure 8.12 *(Continued)*

8.3.5.1 Detailed circuit operation—proportional base drive. In Fig. 8.12, there is positive feedback between windings N_c and N_b (note the dots) of drive transformer $T1$. These windings act as a current transformer with a turns ratio N_c/N_b. If $Q2$ is turned on somehow and a current $I_c(\approx I_e)$ flows in N_c, a base current $(N_c/N_b)I_c$ flows in N_b.

Thus for $(N_c/N_b) = 0.1$ if I_c is 10 A, $Q2$ base current is 1 A. If I_c is reduced to 1 A, $Q2$ base current is only 0.1 A. Now when $Q2$ is to be turned off, the stored base charges which must be removed correspond only to 0.1 and not to 1.0 A and turnoff at the 1.0-A output current level is still rapid.

The problem now is how to initiate $Q2$ turnon and at turnoff, how to break the tightly coupled positive-feedback loop between N_c and N_b and supply a large reverse base current to minimize $Q2$ storage and turnoff time.

Transistor $Q2$ can be turned on and off in a number of ways. Some designers turn on an auxiliary transistor $Q1$ to turn $Q2$ on and turn $Q1$ off to turn $Q2$ off. But herein, $Q2$ is turned on by turning $Q1$ off. Then, by flyback action (note the dots in N_p and N_b), the stored mag-

netizing current in the $T1$ primary is multiplied by the N_p/N_b ratio and delivered to N_b to initiate the turnon operation.

The impulse from N_p to N_b need last only a short time—long enough for $Q2$ collector current to have risen sufficiently to establish a solid positive-feedback loop between N_c and N_b. Then, for the duration of the on time, base current is supplied by transformer coupling between N_c and N_b.

Now to turn $Q2$ off, $Q1$ is turned on. It is assumed that V_c, the voltage at the top of the capacitor $C1$, is fully charged up to the supply voltage V_h. When $Q1$ comes on, the dot end of N_b goes negative with respect to its no-dot end (see dot orientation) and turns $Q2$ off. The turns ratio N_p/N_b is chosen to yield a 2-V negative pulse at N_b so as to permit $Q2$ to sustain its V_{cev} rating as it turns off.

The magnitudes of $R1$, $C1$, V_h, N_p magnetizing inductance, and all the $T1$ turns ratios are critical to the design. The quantitative design of all these components are covered in the following section.

8.3.5.2 Quantitative design of proportional base drive scheme. The first decision to be made is selection of the turns ratio N_b/N_c (Fig. 8.12). This is chosen equal to the minimum $Q2$ beta or

$$\frac{N_b}{N_c} = \beta_{\min(Q2)} \tag{8.5}$$

Proportional base drive is used mostly for $Q2$ currents over 5 to 8 A, at which level minimum transistor betas range between 5 and 10. Thus N_b/N_c is usually chosen in the range of 5 to 10.

The next decision is the choice of N_p/N_b. Below, it will be ensured that when $Q1$ is turned on, V_c, the voltage at the top of capacitor $C1$ is at the supply voltage V_h. Now when $Q1$ is turned on, it is desired to couple a 2-V negative impulse to the $Q2$ base. This permits $Q2$ to tolerate its V_{cev} voltage rating and prevents "second breakdown." Thus

$$\frac{N_p}{N_b} = \frac{V_h}{2} \tag{8.6}$$

With these ratios fixed, the transformer is almost designed. Then N_c is chosen at one turn which fixes N_b and N_p. A core will be chosen and a gap will be specified as below.

Now when $Q2$ has been turned on and is carrying its maximum current $I_{c(\max)}$, the base current to support that collector current comes from positive feedback between N_c and N_b. But to initiate that regenerative process, the current impulse delivered from N_p to N_b by flyback action when $Q1$ turns off must be large enough to ensure that

the positive-feedback loop latches in securely. If the impulse from N_p is too small in amplitude or width, $Q2$ may turn on and fall back off.

To ensure that this does not happen, the short current impulse delivered to N_b from N_p will be made equal to that delivered to N_b from N_c after solid positive feedback has been established. Thus I_1 (Fig. 8.12) will be chosen, by proper selection of $R1$, to yield

$$I_1 = \frac{N_b}{N_p}\frac{I_{c(\max)}}{\beta_{\min}}$$

$$= \frac{N_b}{N_p}\left[I_{c(\max)}\left(\frac{N_c}{N_b}\right)\right] \tag{8.7a}$$

$$I_1 = I_c\frac{N_c}{N_p} \tag{8.7b}$$

This value of I_1 can be obtained by properly choosing the N_p inductance. It is chosen low enough (as will be calculated below) so that at the end of the minimum $Q1$ on time, the voltage across it will have collapsed to zero, and I_1 is fixed by V_h and R_1 as:

$$I_1 = \frac{V_h}{R_1} \tag{8.8a}$$

and from Eq. 8.7b

$$R_1 = \frac{V_h}{I_c}\frac{N_p}{N_c} \tag{8.8b}$$

8.3.5.3 Selection of holdup capacitor (C1; Fig. 8.12) to guarantee power transistor turnoff. Now at the end of the $Q1$ on time, when $Q1$ has been on long enough for the voltage across N_p to collapse to zero, N_p draws a current of only V_h/R_1. This is sufficient that at $Q1$ turnoff, the current impulse delivered to N_b by flyback action is sufficient to safely close the regenerative feedback loop between N_c and N_b.

At the initial instant when $Q1$ has just turned on, however, it is necessary for $Q1$ to draw more current than I_1. The initial current in N_p has two tasks. When $Q1$ turns on initially, it must first break the positive-feedback loop between N_c and N_b by canceling I_b, the $Q2$ forward base current. Then to turn $Q2$ off quickly with minimum storage time, it should deliver an equal reverse current I_b to the $Q2$ base. Thus to deliver $-2I_b$ to N_b requires $2I_b(N_b/N_p)$ in N_p.

Further, in delivering $-2I_b$ to N_b, the voltage at the top of N_p must remain at V_h to deliver the aforementioned momentary 2-V reverse-bias pulse to the $Q2$ base. If not for capacitor $C1$, $R1$ would not be able

to remain at V_h and supply $2I_b(N_b/N_p)$ to the primary. Thus capacitor $C1$ is added at the junction of $R1$ and N_p to hold V_c up long enough to turn $Q2$ off.

This last requirement fixes the value of C_1. Capacitor $C1$ must supply the current $2I_b(N_b/N_p)$ at a voltage V_h for a time equal to the $Q2$ turnoff time t_{off}. This requires a stored energy in $C1$ of

$$\begin{aligned} \tfrac{1}{2}C_1(V_h)^2 &= 2I_b \frac{N_b}{N_p} V_h t_{off} \\ &= 2I_c \frac{N_c}{N_b}\frac{N_b}{N_p} V_h t_{off} \end{aligned}$$

or

$$\tfrac{1}{2}C_1(V_h)^2 = 2I_c \frac{N_c}{N_p} V_h t_{off} \tag{8.9a}$$

or

$$C_1 = 4\left(\frac{I_c}{V_h}\right)\frac{N_c}{N_p} t_{off} \tag{8.9b}$$

For the usual average turnoff time t_{off} for high-current bipolar transistors of 0.30 μs, this fixes the value of C_1.

Now that $R1$ and $C1$ have been selected, care must be taken that at the instant $Q1$ is turned on to turn $Q2$ off, the voltage V_c (Fig. 8.12) has risen to V_h. For during the previous on time (Fig. 8.12), the inductance of N_p has been chosen so low that V_c has fallen to zero to permit a buildup of I_1 in $R1$. Hence at the start of the $Q1$ off time, V_c is at ground and has only the complete minimum off time to charge back up to V_h again.

Thus, to recharge $C1$ to within 5 percent of V_h, it must be true that $3R_1C_1$ = minimum $Q1$ off time.

If the preselected R_1C_1 time constant is too large, $C1$ can be recharged rapidly with an emitter-follower as in the scheme of Fig. 8.13 suggested by Dixon.[6]

8.3.5.4 Base drive transformer primary inductance and core selection. At the start of the $Q1$ on time, V_c is at V_h and at the end of the on time, it is desired that V_c collapse to ground so as to store a current I_1 in N_p. Now assume that V_c falls linearly from V_h to ground in the minimum $Q1$ on time $t_{on(min)}$. Then at the end of the on time, N_p must be carrying a current:

$$\begin{aligned} I(N_p) &= I_{R1} + I_{C1} \\ &= \frac{V_h}{R1} + \frac{C_1 V_h}{t_{on(min)}} \end{aligned} \tag{8.10}$$

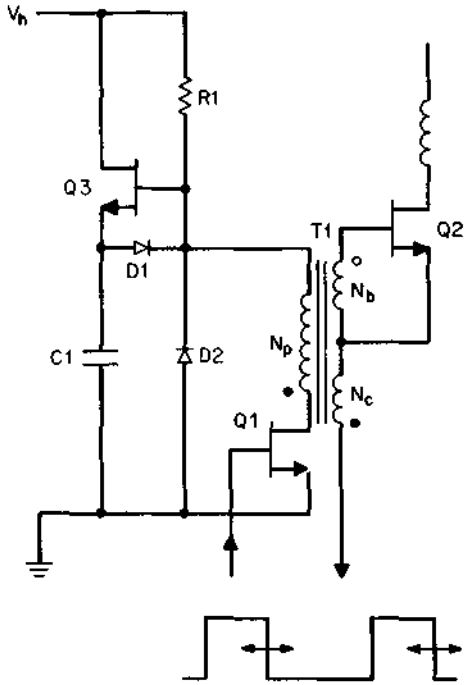


Figure 8.13 Fast $C1$ recharge circuit for proportional base drive. In Fig. 8.12, if $C1$ cannot be recharged to V_h in the minimum $Q2$ on time, emitter-follower $Q3$ is interposed between $R1$ and $C1$ for fast recharge.

Now in the time t_{on}, the voltage across N_p falls linearly from V_h to ground. This is the equivalent of a step of voltage $V_h/2$ applied to the N_p inductance L_p for a time t_{on}. The current rise in the inductor in a time t_{on} is then $I_1 = V_h t_{on}/2L_p$. This current rise must equal the current $I(N_p)$ of Eq. 8.10. Then

$$\frac{V_h t_{on}}{2L_p} = \frac{V_h}{R_1} + \frac{C_1 V_h}{t_{on(min)}}$$

or

$$L_p = \frac{t_{on}}{2(1/R_1 + C_1/t_{on(min)})} \tag{8.11}$$

8.3.5.5 Design example—proportional base driver. Consider a forward converter with the base drive circuit of Fig. 8.12. Assume a $Q2$ collector current of 12 A. Assume the circuit is an off-line converter from the 115-V-AC line with a minimum rectified DC supply voltage of 145 V. Then from Eq. 2.28, 12 A of collector current corresponds to an output power of $12 \times 145/3.13 = 556$ W.

Assume that $Q2$ is a 15-A device with a minimum beta of 6 at 12 A. Then from Eq. 8.5, $N_b/N_c = 6$ and for $N_c = 1$, $N_b = 6$ turns. Now as-

sume a housekeeping supply voltage V_h of 12 V. From Eq. 8.6, $N_p/N_b = V_h/2 = 6$ and $N_p = 6N_b = 36$ turns.

From Eq. 8.7b, $I_1 = I_c(N_c/N_p) = 12/36 = 0.33$ A; and from Eq. 8.8b, $R_1 = (V_h/I_c)(N_p/N_c) = (12/12)(36/1) = 36\ \Omega$.

From Eq. 8.9b, for a $Q2$ turnoff time t_{off} of 0.3 μs

$$C_1 = 4\frac{I_c}{V_h}\frac{N_c}{N_p}t_{off}$$

$$= 4\left(\frac{12}{12}\right)\left(\frac{1}{36}\right)(0.3 \times 10^{-6})$$

$$= 0.033\ \mu F$$

Now assume a 50-kHz switching frequency. From Fig. 8.11, minimum $Q1$ on time occurs at maximum $Q2$ on time, which will be assumed to be a half period or 10 μs. Then from Eq. 8.11

$$L_p = \frac{t_{on(min)}}{2(1/R_1 + C_1/t_{on})}$$

$$= \frac{10 \times 10^{-6}}{2[1/36 + (0.033 \times 10^{-6}/10 \times 10^{-6})]}$$

$$= 162\ \mu H$$

Further, since N_p is calculated as 36 turns, A_l (inductance per 1000 turns) is $(1000/36)^2(0.162) = 125$ mH per 1000 turns.

Finally, from Fig. 8.9, the Ferroxcube 1408PA3C8-100 core can be used. Its A_l is 100 mH per 1000 turns, which is close enough.

8.3.6 Miscellaneous base drive schemes

A wide variety of specialized bipolar base drive schemes have evolved through the years. They are more often used at lower power levels and, by various circuit "tricks," seek to achieve two common goals. (1) a low-parts-count scheme to obtain substantial reverse base voltage, reverse base current, or a base-emitter short circuit at turnoff and at turnon and (2) forward base current adequate to drive lowest beta transistors at maximum current without long storage times for high beta transistors at lowest current. Some examples are shown below.

P. Wood devised the circuit of Fig. 8.14*a* for a 1000-W off-line power supply.[7] Its major features are current gain through transformer $T1$ and a 2-V reverse base bias for power transistor $Q2$ which is automatically switched on by turning $Q1$ on at the instant that $Q2$ is to be turned off. It can be used either at ground level for the bottom power transistors in a bridge or at the power supply voltage source level for the upper bridge transistors. It has also been widely used at lower

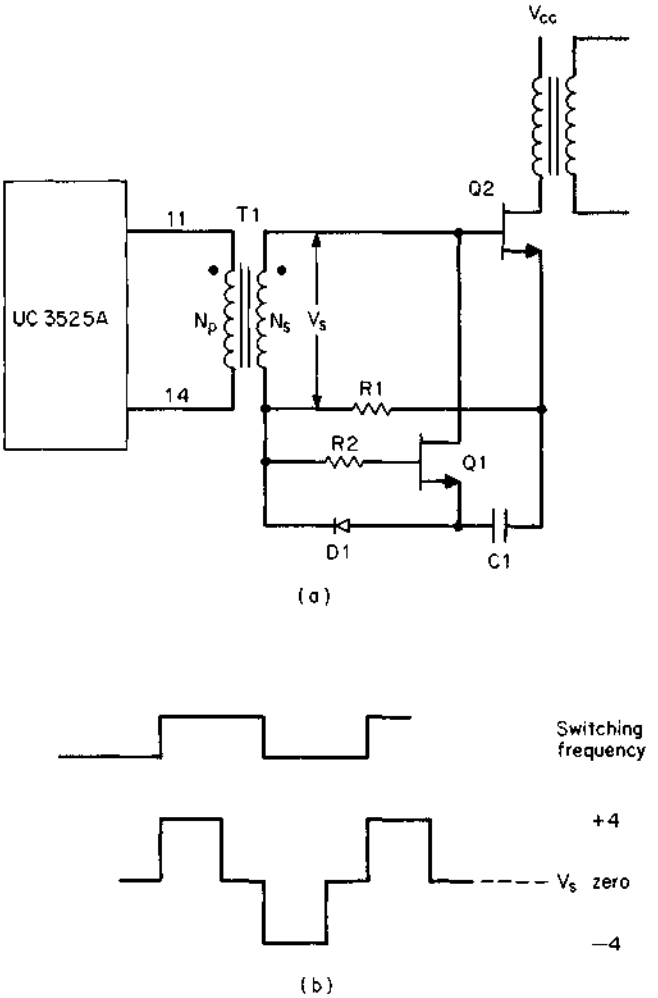


Figure 8.14 (*a*) Wood base drive circuit. When dot end of N_s goes positive, $Q2$ turns on with its base current limited by $R1$. Voltage V_s is chosen to provide about 4 V across $R1$ for the known base current. With 4 V across $R1$, $C1$ charges to 3 V through $D1$. While voltage exists across N_s, $Q1$ is reverse-biased and is off. When voltage across N_p drops to zero, so does voltage across N_s. Now 3-V charge on $C1$ turns on $Q1$ via $R1$ and $R2$. This brings $Q2$ base sharply down to a −3-V bias and turns it off rapidly. (*c*) Adding $D3$, $Z1$, and $D2$ permits driving $T1$ from the on collector in the output transistor of a 3524 chip.

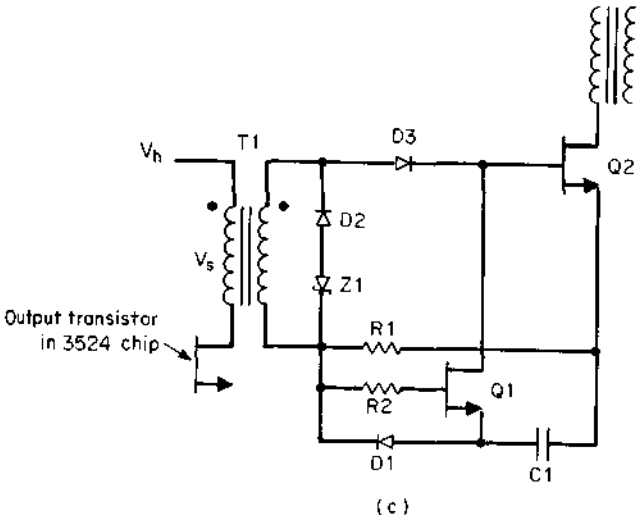


Figure 8.14 *(Continued)*

power levels. It works as follows. Assume (Fig. 8.14*b*) the voltage across N_s is high for the part of a half period that $Q2$ is supposed to be on. Then for the balance of the half period that $Q2$ is to be off ($Q2$ dead time), the voltage across N_s is clamped to zero. Then, in the following half period, the N_s voltage reverses polarity to reset the core on its hysteresis loop.

Such a waveform can be obtained using the Unitrode UC3525A PWM chip by connecting the $T1$ primary across output pins 11 and 14. Now the $T1$ turns ratio is chosen to yield a secondary voltage V_s of about 4 V. Resistor $R1$ is chosen to limit $Q2$ base current to a value sufficient to turn on and saturate the lowest beta transistor at the maximum I_c with the required speed. Then

$$R_1 = \frac{V_s - V_{be}}{I_{b(max)}} = \frac{(V_s - 1)\beta_{(min)}}{I_{c(max)}} \tag{8.12}$$

Now for 4 V across N_s, voltage across $R1$ is about 3 V and the left-hand side of $C1$ charges down to approximately −2 V with respect to the $Q2$ emitter. With the left-hand side of $R1$ at −3 V with respect to the $Q2$ emitter, $Q1$ has a −1-V bias so long as current flows through $R1$ and is kept off.

At the start of the dead time, when V_s falls to zero, $Q2$ should turn

off as rapidly as possible and with as low a storage delay as possible. At the instant V_s falls to zero, the left-hand side of $C1$ is at -2 V with respect to the $Q2$ emitter and $C1$ is a floating power supply of -2 V feeding the series combination of $R1$, $R2$ and the base-emitter of $Q1$.

Thus at the instant V_s falls to zero, $Q1$ turns on and pulls the $Q2$ base down to -2 V and turns it off. The reverse base current to $Q2$ is not uniquely determined. But $Q1$ is a small, fast 2N2222AA with a minimum gain of 50 at 500 mA. This is sufficient to yield quite low storage and turnoff time in $Q2$. As soon as the dot end of N_s goes positive again, the $Q1$ base is again reverse-biased and $Q1$ turns off.

A circuit which permits the use of the less expensive SG3524 instead of the UC3525 chip is shown in Fig. 8.14*c*. Here, when the V_s polarity reverses, zener diode $Z1$ (3.3 V) and diode $D2$ clamp it and reset the core in a time approximately equal to its setting time. Diode $D3$ blocks the resetting voltage from being clamped by $Q1$.

Now the voltage to which $C1$ had charged via $D1$ is the floating bias voltage which turns $Q1$ on via $R1$ and $R2$ in series. As $Q1$ turns on, it pulls the $Q2$ base down hard to -2 V and turns it off. The addition of diode $D3$ requires the $T1$ turns ratio to be set to a value which yields a V_s of about 5 V.

If power transistor reverse base bias is not critical, the components $Q1$, $C1$, and $D1$ of the Wood circuit can be eliminated by the use of a UC3525 chip as shown in Fig. 8.15. Although the scheme does not offer a reverse base bias, it does provide a dead short circuit at the power transistor base immediately after turnoff. This minimizes storage and turnoff time but does not provide the reverse base bias which enables the transistor to sustain its V_{cev} rating.

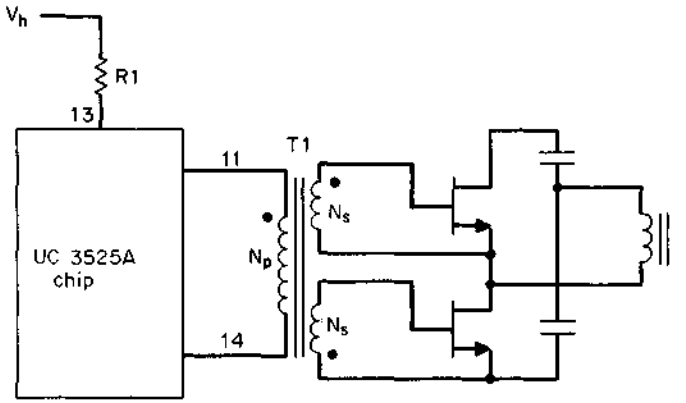


Figure 8.15 Base drive to upper and lower transistors in a bridge through a transformer whose primary is fed directly from UC3525A output.

The circuit is shown in Fig. 8.15 providing base drive to the upper and lower transistors in a half bridge—or, alternatively, to a pair of push-pull transistors.[8] Resistor $R1$ is used for current limiting, and to be a good constant-current source, the voltage across it should be relatively independent of voltage drops in the internal source and sink drivers at pins 11 and 14.

These source and sink drivers in the UC3525 are each specified to have about a 2-V drop at 200 mA. Thus for $V_s (= V_{be}) = 1$ V and a current gain of 10, the primary voltage at pins 11 and 14 is 10 V. Further, since the primary current flows out of the source driver and returns through the sink driver to ground, for 10 V across the primary, pin 13, the top of the source driver is at 14 V. For $R1$ to be a fairly good constant-current source, the voltage across it is set at 6 V, making the housekeeping supply $V_h = 20$ V.

Then for primary current limiting at 200 mA, $R_1 = 6/0.2 = 30\ \Omega$. For a 10/1 turns ratio, this yields 2 A of power transistor base current. At maximum power transistor current, turnoff and storage delay are minimized because with the UC3525, both output terminals are short-circuited together immediately after the end of an on time.

Note, though, that unlike with Baker clamping or proportional base drive, the same power transistor base current is delivered at maximum and minimum collector current. Thus, if base currents are adequate at maximum collector current, bases will be very heavily overdriven at minimum current and storage and turnoff time will be long at those currents.

Figure 8.16 shows another alternative with some desirable features. Transistor $Q4$ is the power transistor which is being Baker clamped

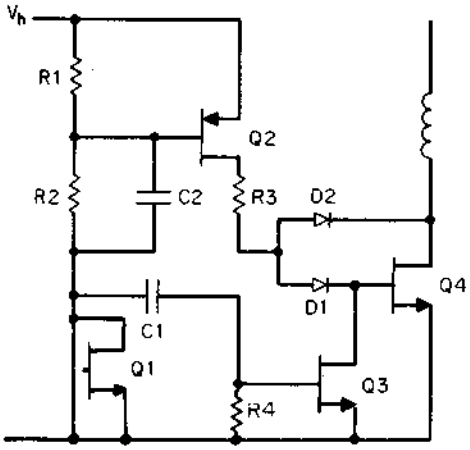


Figure 8.16 DC-coupled power transistor base drive. When $Q1$ in PWM chip turns on, it turns $Q2$ on, which, via $R3$, turns $Q4$ on. When $Q1$ turns off, a positive-going differentiated spike is coupled via $C1$ into $Q3$ base, turning $Q4$ base to ground to turn it off rapidly; $Q1$ turnoff also couples a positive spike into $Q2$ base to turn if off rapidly.

by diodes $D1$, $D2$. Turnon drive comes from the PNP emitter-follower $Q2$ (2N2907A—a small, fast 800-mA device). Resistor $R3$ is chosen to provide the desired maximum $Q4$ base current [$I_{R3} = (V_h - 2)/R3$]. When $Q1$, the PWM output transistor, turns on, it turns on $Q2$, which then turns $Q4$ on.

Normally $Q3$ is off and does not rob any base drive from $Q4$. When it is desired to turn $Q4$ off, $Q1$ is turned off, removing base drive from $Q2$. Its collector current ceases and forward drive is removed from $Q4$, which then starts turning off.

As $Q1$ starts turning off, its collector voltage rises steeply. A differentiated positive pulse is coupled via $C1$ into the $Q3$ base and turns it on momentarily. This short-circuits the $Q4$ base to ground, minimizing its storage and current fall time.

As $Q1$ starts turning off, $Q2$ does not turn completely off until the bottom end of $R2$ has risen almost to V_h. To speed up the $Q4$ turnoff time, in addition to the differentiated positive pulse coupled to the $Q3$ base via $C1$, a differentiated positive pulse is coupled to the $Q2$ base via $C2$ to accelerate its turnoff time.

A final base drive scheme is shown in Fig. 8.17, where $Q1$ is the output transistor in the PWM chip, $Q4$ is the power transistor, and $Q2$ and $Q3$ constitute an NPN-PNP emitter-follower totem pole. For $Q2$ a 2N2222A, $Q3$ a 2N2907A, it is capable of "sourcing" up to 800 mA when $Q2$ is on and "sinking" 800 mA when $Q3$ is on. Both $Q2$ and $Q3$ are fast 300-mHz transistors.

On the way up, when $Q2$ is turning on, $Q3$ has a 0.6-V reverse base-

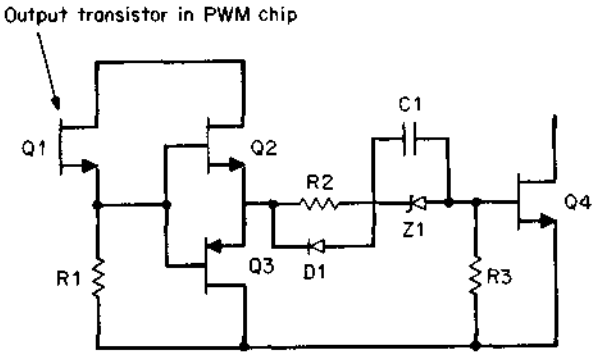


Figure 8.17 Direct coupling from emitter of output transistor in PWM chip. When $Q1$ is on, totem-pole driver $Q2$ turns on power transistor $Q4$ with base current limiting determined by $R2$, $Z1$. Capacitor $C1$ takes on a charge equal to the zener voltage (~ 3.3 V). When $Q1$ turns off, $Q3$ emitter falls to about +0.6 V and right-hand side of $C1$ forces $Q4$ base down to about −3 V, turning it off rapidly.

emitter bias and is off. On the way down, when $Q3$ is turning on, $Q2$ has a 0.6-V reverse base-emitter bias and is off.

In Fig. 8.17 $Z1$ is a 3.3-V zener diode. When $Q1$ and $Q2$ are on, $C1$ charges up to 3.3 V with its $Q4$ base end negative with respect to its $R2$ end. $R2$ serves to limit $Q4$ base current ($= [V_h - V_{ce(Q2)} - V_{be(Q4)} - V_{Z1}]/R_2$)

When $Q1$ turns off, $R1$ brings the $Q2$, $Q3$ base quickly to ground as the capacity at the bases of the totem-pole transistors is very low. The $Q3$ emitter comes to +0.6 and $Q3$ is hard on. Now the 3.3-V negative charge at the right-hand side of $C1$ brings the $Q3$ base down to −3.3 V and quickly turns $Q4$ off.

References

1. A. I. Pressman, *Switching and Linear Power Supply, Power Converter Design*, pp. 322–323, Switchtronix Press, Waban, MA, 1977.
2. K. Billings, *Switchmode Power Supply Handbook*, pp. 1.132–1.133, McGraw-Hill, New York, 1989.
3. G. Chryssis, *High Frequency Switching Power Supplies*, 2d ed., pp. 68–71, McGraw-Hill, New York, 1989.
4. P. Wood, *High Efficiency, Cost Effective Off Line Inverters*, TRW Power Semiconductor Application Note 143-1978, Lawndale, CA, 1978.
5. R. Carpenter, "A New Universal Proportional Base Drive Technique for High Voltage Switching Transistors," *Proceedings Powercon 8*, 1981.
6. L. Dixon, "Improved Proportional Base Drive Circuit," Unitrode Power Supply Design Seminar, Unitrode Corp., 1985.
7. P. Wood, "Design of a 5 Volt, 1000 Watt Power Supply," TRW Power Semiconductor Application Note 122, Lawndale, CA, 1975.
8. Unitrode Corp. Application Note 89-1987, Unitrode Corp., Watertown, Mass.

MOSFET Power
Inpu

9.1 Introduction

Over about the past 10 years, the pov
the power supply industry and in an
of the electronics industry. The faste
has permitted increasing power sup
about 50 to 200 or 400 kHz. It has th
and made possible a host of new pro
with smaller power supplies. The incr
computers is a prime example.

The higher frequencies spurred a
components industries. The semic
changed drastically first. More of th
MOSFETs, and their voltage and cur
cally and prices dropped, making a la
possible—even at lower frequencies.

Magnetic materials with lesser loss
magnetic flux density were developed
capable of operating at higher frequ
frequencies and thus smaller transfo
tors required everything else to get
greater emphasis on manufacturing
techniques.

A new industry and field for res
plies—was developed. Resonant pow
trolled rectifiers at about 20 to 30 kHz
But with the high frequencies possible

Chapter

9

MOSFET Power Transistors and Input Drive Circuits

9.1 Introduction

Over about the past 10 years, the power MOSFET has revolutionized the power supply industry and in an enormous way changed the rest of the electronics industry. The faster switching speed of MOSFETs has permitted increasing power supply switching frequencies from about 50 to 200 or 400 kHz. It has thus made power supplies smaller and made possible a host of new products which were feasible only with smaller power supplies. The increasingly smaller size of personal computers is a prime example.

The higher frequencies spurred a wide variety of changes in the components industries. The semiconductor industry, of course, changed drastically first. More of the research funds were spent on MOSFETs, and their voltage and current ratings improved dramatically and prices dropped, making a large number of new applications possible—even at lower frequencies.

Magnetic materials with lesser losses at high frequencies and high magnetic flux density were developed. Pulse-width-modulating chips capable of operating at higher frequencies were introduced. Higher frequencies and thus smaller transformers and smaller filter capacitors required everything else to get smaller, and there was thus a greater emphasis on manufacturing processes such as surface-mount techniques.

A new industry and field for researchers—resonant power supplies—was developed. Resonant power supplies using silicon controlled rectifiers at about 20 to 30 kHz had been in use for many years. But with the high frequencies possible with MOSFETs, an army of re-

searchers stepped in to develop dozens of new resonant circuit topologies operating at 0.3 to 5 MHz.

Phenomena which could be ignored at frequencies up to 100 kHz now had to be considered by designers. Thus skin-effect and especially proximity-effect losses in transformer coils had to be taken into account as these became a larger fraction of total transformer losses at higher frequencies. With faster-rise-time current waveforms, $L\ di/dt$ spikes on ground buses and supply rails became more troublesome and more attention had to be paid to wiring layout, low-inductance buses and supply rails, and capacitive decoupling at critical points.

Yet with all these attractive new possibilities and problems, the power supply designer, familiar with bipolar design, could quickly learn to design with MOSFETs by acquiring a surprisingly small amount of information about their characteristics. Material on the internal, solid-state physics structure of the MOSFET which determines its circuits behavior is not of great importance to the circuit designer and will not be considered herein. The MOSFET's various DC volt-ampere characteristics, terminal capacitances, temperature characteristics, and turnon-turnoff speed information are all that is needed for circuit design, and only such material will be considered here.

In very many ways, design with MOSFETs is simpler than with bipolars. MOSFET input terminal (gate) drive circuitry is far simpler than the complex base drive schemes for bipolar transistors discussed in the previous chapter. Specifically, because MOSFETs have no storage times, the complexities of Baker clamps and proportional base drive are unnecessary. The problems arising from the four-to-one production spread of transistor beta in bipolars do not exist.

During the turnoff transition, because of the fast MOSFET current fall time, the overlap of falling current and rising voltage at the output terminal occurs at lower current levels and decreases overlap or AC switching losses (Sec. 1.3.4). This simplifies or makes unnecessary the design of load-line-shaping circuits (snubbers, to be discussed in Chap. 11).

Basic MOSFET characteristics and MOSFET design will be discussed in the following sections.

9.2 MOSFET Basics

The MOSFET[1,2] is a three-terminal voltage-controlled switch—in contrast to the bipolar transistor, which is three-terminal current-controlled switch. In switching power supply circuits, it is used just as a bipolar transistor—either fully on with an input drive which should minimize its on-voltage drop or fully off at zero current and sustaining the supply voltage or some multiple of it.

The symbol for a MOSFET is shown in Fig. 9.1*a*, but herein, the simplified version shown in Fig. 9.1*b* will be used.

The three MOSFET terminals are called the *drain, gate,* and *source,* corresponding to the collector, base, and emitter of a bipolar transistor. MOSFETs, just as bipolars, are available for operation from either positive or negative power supply buses.

The N-channel type (equivalent to a bipolar N type) is fed from a positive supply voltage. The load impedance is connected between the positive supply source and the drain. Current, controlled by a positive gate-to-source voltage, flows from the positive rail through the load impedance into the drain and returns from the source to the negative rail (Fig. 9.2*a*).

The P-channel type (equivalent to a bipolar P type) is fed from a negative supply source. The load impedance is connected between the drain and negative rail. Current, controlled by a negative gate-to-source voltage, flows into the source from the positive rail, out through the drain, and through the load impedance to the negative rail (Fig. 9.2*b*).

Most power transistors are N-channel types. There are two further distinct types—either *enhancement* or *depletion* types. In the N enhancement channel type, drain-to-source current is zero at zero gate-to-source voltage. It requires a positive gate-to-source voltage to turn on drain to source current.

In the N depletion type, drain-to-source current is a maximum at zero gate-to-source voltage. It requires a negative gate-to-source voltage to turn off drain-to-source current. Depletion-type MOSFETs are never used for power transistors. They are still—but increasingly rarely—used in low-current signal-type MOSFETs.

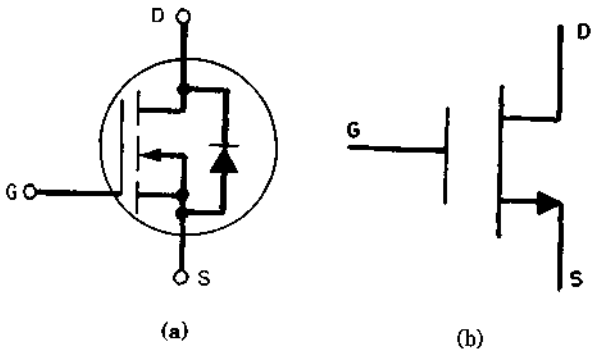


Figure 9.1 (*a*) Symbol for an N-type MOSFET. The diode is inherent in its structure and has forward current, reverse voltage ratings identical to the MOSFET itself. (*b*) Simplified N-type MOSFET symbol used herein.

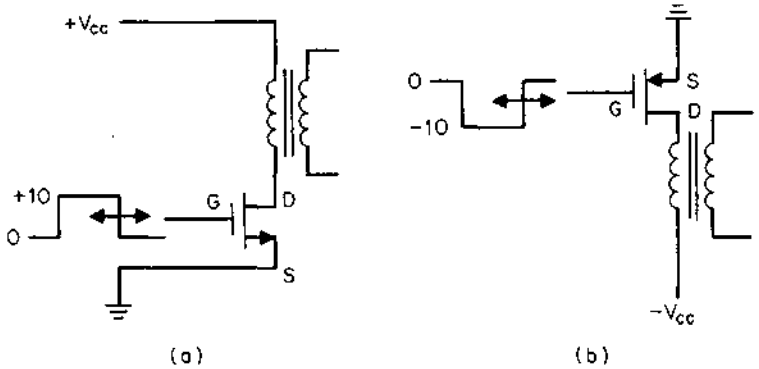


Figure 9.2 (*a*) A drain-loaded N-type MOSFET. (*b*) A drain loaded P-type MOSFET.

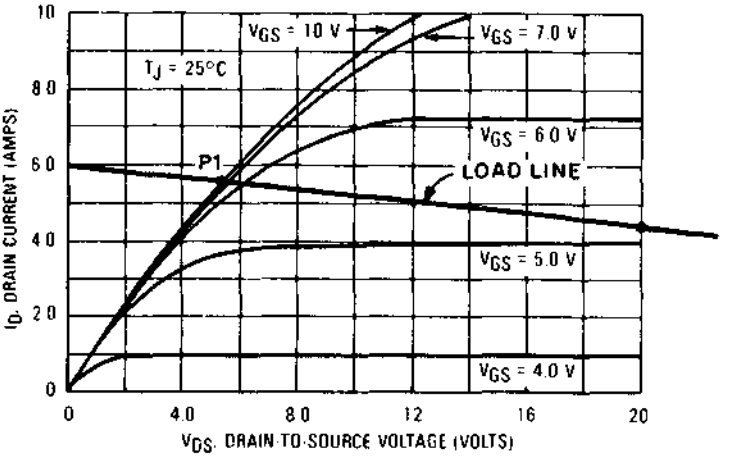
9.2.1 MOSFET drain current versus drain-to-source voltage characteristics ($I_d - V_{ds}$)

These are shown in Fig. 9.3*a* for a typical 7-A 450-V device (Motorola MTM7N45). They correspond to the $I_c - V_{ce}$ curves of a bipolar transistor (see Fig. 8.1). Note that they are voltage-controlled and that drain current is turned on by a positive gate-to-source voltage. This can be seen more clearly in the "transfer" characteristics of Fig. 9.3*b*, which shows drain-to-source current versus gate-to-source voltage.

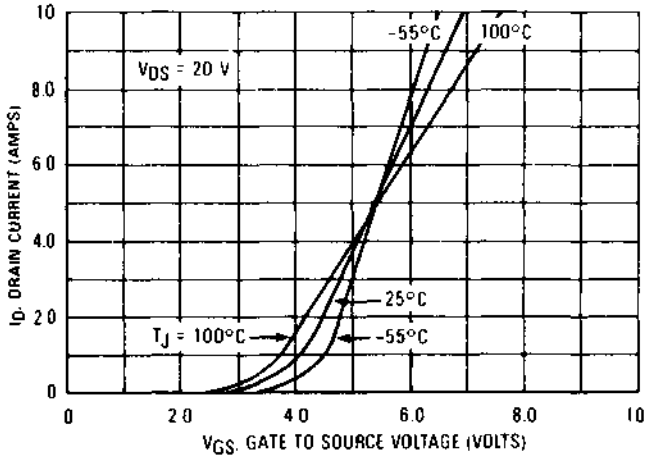
The transfer characteristics show a very significant advantage of the MOSFET over the bipolar transistor. Note in Fig. 9.3*b* that drain current does not begin to turn on until the gate-to-source voltage reaches about 2.5 V. Thus positive noise pickup spikes at the gate terminal cannot falsely turn drain current on until the 2.5-V threshold is reached. This is in contrast to bipolar transistors, where the diodelike base-to-emitter input voltage characteristic permits false collector current turnon with noise voltage spikes as low as 0.6 V or even lower at temperatures above 25°C.

Note in Fig. 9.3*a*, drain current characteristics have a "knee" somewhat like the bipolar transistor knee. Beyond the knee, drain current is constant over a large range of drain-to-source voltage and is determined only by the gate-to-source voltage. Below the knee, the $I_d - V_d$ curves asymptotically converge to a constant slope. The slope of that asymptote (dV/dI) is referred to as r_{ds}.

It is seen in Fig. 9.3*a* that a gate-to-source voltage of 10 V is sufficient to drive the drain-to-source voltage down to the voltage at the intersection of a load line with the r_{ds} slope (point *P*1). Higher gate voltages will not decrease the on V_{ds} significantly unless operating

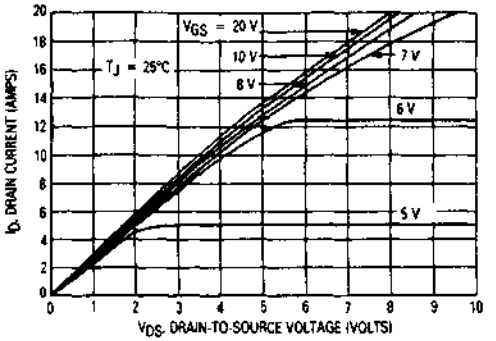


(a)



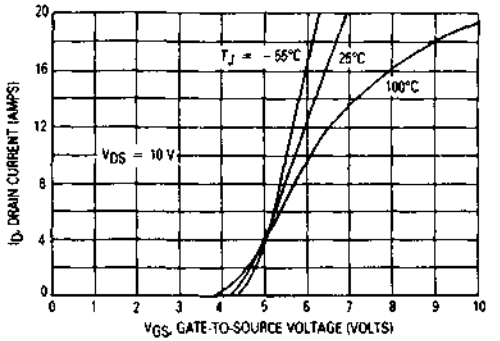
(b)

Figure 9.3 (*a*) The $I_d - V_{ds}$ characteristics of the MTM7N45—a typical 7-A 450-V MOSFET. (*b*) The $I_d - V_{gs}$ or "transfer" characteristics of the MTM7N45. (*c, d*) $I_d - V_d$ and transfer characteristics of the MTM15N40. (*Courtesy of Motorola Inc.*)



On-Region Characteristics

(c)



Transfer Characteristics

(d)

Figure 9.3 (*Continued*)

current is close to the maximum-rated current where the (I_d - V_d) curve bends away from the r_{ds} slope.

Thus, in contrast to the bipolar transistor, where the collector to emitter on voltage is about 0.3 to 0.5 V over a very large range of collector currents, in a MOSFET, the drain "on" potential is equal to $I_{ds}R_{ds}$. Generally, it will be found that to have an on V_{ds} voltage of about 1 V, at a current I_d, a MOSFET with a maximum continuous current rating of about $3I_d$ to $5I_d$ should be selected as r_{ds} is inversely proportional to maximum current rating.

This can be seen in Fig. 9.3*a* for the 7-A MTM7N45 (r_{ds} = 0.8 Ω). If it were to be used at 7 A, for a gate-to-source voltage V_{gs} of 10 V at 7 A, Fig. 9.3*a* shows its V_{ds} to be 7 V. This would yield an unacceptable 49 W of dissipation during the on time. The MTM15N40—a 15-A, 400-

V Motorola device (r_{ds} = 0.4 Ω) is shown in Fig. 9.3*c* and 9.3*d*. There it is seen that at V_{gs} of 10 V and 7 A, its V_{ds} is still about 2.5 V.

It is customary to keep bipolar V_{ce} on voltages at 1 V or less to keep their total dissipation low. With a bipolar, the on dissipation (I_cV_{ce}) may be only half to a third of the total—the balance being the overlap or AC switching dissipation. But MOSFETs generally can be operated with an on V_{ds} of up to 2 or 3 V. Drain current turnoff time is so fast that the dissipation due to the overlap of falling current and rising voltage is generally negligible. Total dissipation is thus $I_{ds}V_{ds}(t_{on}/T)$ dissipation alone.

9.2.2 MOSFET input impedance and required gate currents

The MOSFET's DC input impedance is extremely high. At V_{gs} of 10 V, the gate draws only nanoamperes of current. Thus once the gate has been driven up to, say, the 10 V "on" level, the current it draws is negligible.

However, there is considerable capacity at the gate-to-source terminal. This requires relatively large currents to drive the gate up and down the 10 V required to switch drain current on and off with the required speed. The required gate drive currents are calculated as follows (Fig. 9.4).

In Fig. 9.4, the current I_g required to move the gate up 10 V from ground consists of the two currents I_1 and I_2. Two capacitors, *C*1 and *C*2, are involved. Capacitance C_1 is the capacitance from gate to source and is referred to in the data sheets as "C_{iss}," the input capacitance. Capacitance C_2 is the capacitance from gate to drain and is referred to in the data sheets as "C_{rss}," the reverse-transfer capacitance.

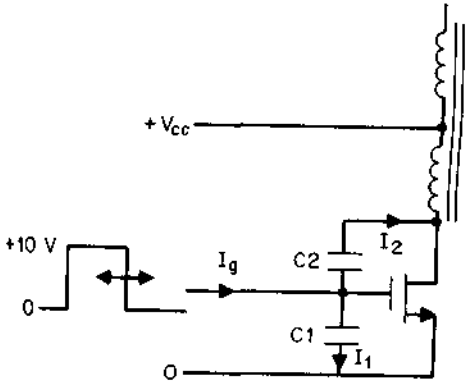


Figure 9.4 The gate driver must supply I_1 to C_1 and I_2 to C_2. Because of the Miller effect, I_2 may be greater than I_1 even though C_1 may be 10 times as large as C_2.

To drive the gate up to the 10-V "on" level in a time t_r, the required current I_1 is

$$I_1 = \frac{C_1\,dV}{dt} = \frac{C_1 \times 10}{t_r} \tag{9.1}$$

However, in driving the gate up 10 V, the drain turns on and drops from the supply voltage V_{dc} (Fig. 2.10) to the V_{ds} on voltage, which for simplicity will be taken as ground. Thus, as the top end of $C2$ moves down V_{dc} volts, its bottom end moves up 10 V. The current required to achieve this is

$$I_2 = C_2 \frac{dV}{dt}$$

$$= \frac{C_2(V_{dc} + 10)}{dt} \tag{9.2}$$

It is revealing to calculate these currents in a typical case. Assume an off-line forward converter operating from a nominal 115-V-AC line with ±10 percent tolerance. Then the maximum rectified DC voltage is $1.1 \times 115 \times 1.41 = 178$ V. Assume an MTM7N45 whose C_1 is 1800 pF and whose C_2 is 150 pF. Then from Eq. 9.12, for a gate rise time of 50 ns, we obtain

$$I_1 = \frac{1800 \times 10^{-12} \times 10}{50 \times 10^{-9}}$$

$$= 360 \text{ mA}$$

and from Eq. 9.2

$$I_2 = \frac{(150 \times 10^{-12})(178 + 10)}{50 \times 10^{-9}}$$

$$= 564 \text{ mA}$$

and the total current required from the gate driving source is $I_g = I_1 + I_2 = (0.36 + 0.564) = 0.924$ A. Thus, the smaller capacitor, $C2$, requires about 50 percent more current than does $C1$. This, of course, is the well-known Miller effect, which multiplies the input-to-output capacity by the gain from input to output. By a similar calculation, the current required to move the gate down 10 V is the same 0.924 A.

Generally, the effective input capacity of low-voltage MOSFETs will be lower than high-voltage devices. The input capacitance C_{iss} will be higher and the reverse-transfer capacitance C_{rss} will be lower than

those for higher-voltage devices. But more importantly, the Miller effect multiplication of C_{rss} will be less as the voltage change across C_{rss} is less as the transistor operates from a lower supply voltage.

Thus, consider an MTH15N20—a 15-A, 200-V device. Assume that it operates from a nominal supply voltage of 48 V in a telephone industry supply where minimum voltage is usually assumed to be 38 V and maximum, 60 V. Capacities are C_{iss} = 2000 pF and C_{rss} = 200 pF. Then total input capacity, assuming turnon from a supply voltage of 60 V, is

$$C_{in} = C_{iss} + \left(\frac{70}{10}\right)C_{rss} = 2000 + 7 \times 200 = 3400 \text{ pF}$$

and to move the effective input capacity up 10 V in the same 50 ns requires a gate input current of $I_g = C_{in}\, dv/dt = 3400 \times 10^{-12}$ $(10/50 \times 10^{-9}) = 0.68$ A.

Such input currents, although large, are not too serious as they must last only 50 ns (or whatever is desired for the gate rise time t_r).

9.2.3 MOSFET gate rise and fall times for desired drain current rise and fall times

Very rapid drain current rise and fall times are undesirable as they cause large $L\, di/dt$ spikes on ground buses, supply rails, and large $C\, dV/dt$ capacitatively coupled spikes into adjacent wires or nodes.

The question thus arises as to what gate voltage rise time is required to yield a desired drain current rise time. This can be seen from the transfer characteristics shown in Fig. 9.3*b* and 9.3*d*.

In a MOSFET, switching time between zero and a drain current I_d is only the times required for the gate voltage to traverse the DC voltage level from the threshold to the gate voltage V_{gI} in Fig. 9.3*b*, which yields I_d. The time for the gate to move up from ground to the threshold voltage of about 2.5 V is simply a delay time. Drain current turnon time is not accelerated by overdriving the input terminal as with bipolars (Sec. 8.2 text and Fig. 8.2*b*).

Further, MOSFETs have no storage time. There is only a turnoff delay corresponding to the time required for the gate to fall from its uppermost level of about 10 V to the DC level V_{gI} in Fig. 9.3*b*, which corresponds to the current I_d. Once gate voltage V_{gI} is reached, drain current fall time is the time required for the gate to fall from V_{gI} to the threshold.

Thus, consider turning on an MTM7N45 to 2.5 A. If a 50-ns gate rise time were used for the first 2.5 V, no drain current at all would result (see Fig. 9.3*b*). There would only be a delay equal to the time

required for the gate to rise the first 2.5 V, and by the time the gate rose to about 5 V, most of the 2.5 A would have been turned on.

Then, for a 0- to 10-volt gate rise time in 50 ns, the drain current would rise from 0 to 2.5 A in about (2.5/10)50 or 12.5 ns (Fig. 9.5). Thus, 10-V gate voltage transition times can be two to three times the desired drain current transition times because of the narrow range of gate voltage across which drain current changes. The gate currents calculated above thus need be only about one-half to one-third the values calculated as above.

9.2.4 MOSFET gate drive circuits

As described above, the gate driver must be able to deliver outward-directed current or be a "source" to drive gates in a positive direction to turn drain current on. It must also be able to absorb inward-directed current or to be a "sink" to pull gates in a negative direction to ground to turn drain current off.

Most of the earlier PWM chips (SG1524 family) could not both source and sink large currents—they could only source or sink fairly large currents, as can be seen from their output stage circuit in Fig. 9.6*a*.

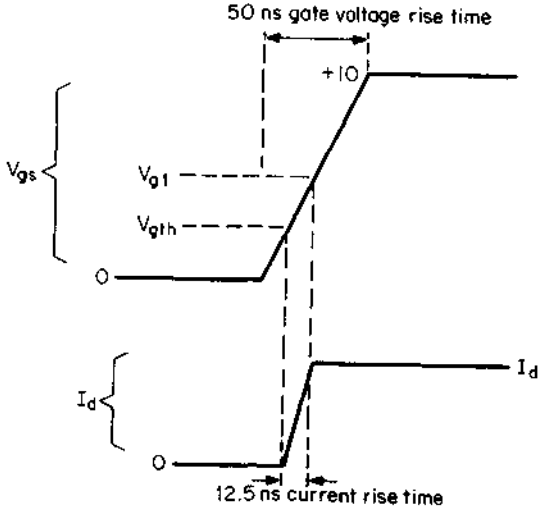
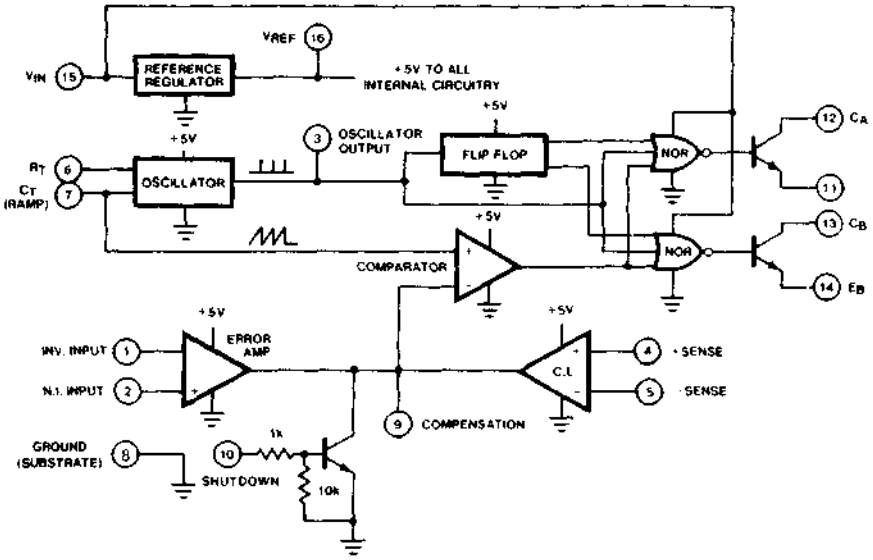


Figure 9.5 A 0- to 50-ns gate voltage rise time causing a 12.5-ns drain current rise time. The first 2.5 V on the gate voltage rise time to the gate threshold voltage is simply a delay time. At a gate voltage level of about 5 to 7 V, most of the drain current has already been turned on.



(a)

Figure 9.6 (*a*) The SG1524 PWM chip with its output transistors (pins 11, 12 and 13, 14). The output transistors can either source or sink (but not both) 200 mA. (*b*) An output transistor-emitter driving the high input capacity at a MOSFET gate. The 200-mA source current can drive the high MOSFET gate rapidly in a positive direction, but it is only R_1 which pulls the gate negative in a time $3R_1C_{in}$. (*c*) A PNP emitter-follower for a high-current sink driver and the chip output transistor as 200-mA source driver. Resistor R_2 can now be fairly large.

The output stage of an SG1524-type PWM chip consists simply of a transistor with an "uncommitted" emitter and collector. When the transistor is turned on, it can sink 200 mA into its collector or source 200 mA out of its emitter. When used as a source driver, it is only the emitter resistor which can sink current from the gate to pull it low and turn the MOSFET off. Most often, the power transistor must be on when the chip's output transistor is on. Thus for an N-type MOSFET, its gate is driven from the emitter of the output stage as shown in Fig. 9.6*b*.

Such an emitter driving stage can source or deliver 200 mA of outward-directed current to turn a MOSFET gate on relatively rapidly. Thus, as mentioned in Sec. 9.2.2, it required 0.924 A to drive the MTM7N45 gate up 10 V in 50 ns.

The 200 mA of SG1524 source current can thus drive the gate up 10 V in only (0.924/0.2)50 = 231 ns. Also, as mentioned in Sec. 9.2.3, because drain current rise time is within the 2.5- to 5.0-V level of the gate voltage rise time, drain current rise time will be only (2.5/

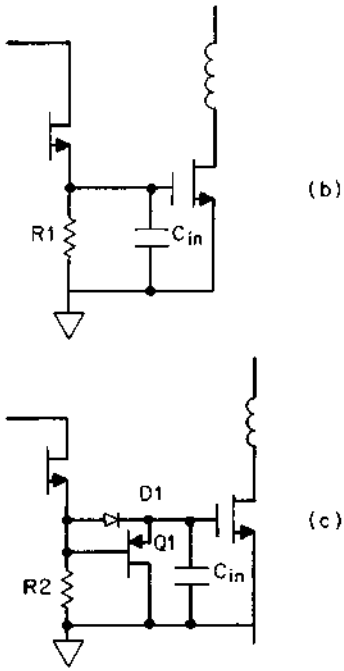


Figure 9.6 (*Continued*)

10) = 58 ns. The first 2.5 V of gate rise time until the threshold is reached yields a delay of only 58 ns. Although this slows up maximum switching frequency, it does not, of course, cause any switching losses.

The gate voltage rise time of 231 ns within the 0- to 10-V level is thus fast enough. But in the circuit of Fig. 9.6*b*, gate voltage fall time is determined only by the emitter resistor R_e and the gate input capacity as there is no constant current discharging that capacity.

In Fig. 9.6*b*, when the base of the chip output transistor goes low (internal to the chip), the output transistor emitter is held up by the large MOSFET input capacity. This biases the chip output transistor fully off and leaves only the emitter resistor to discharge the MOSFET input capacity. That capacity (Sec. 9.2.2) is the input plus the Miller capacity or 1800 + (180/10)150 = 4500 pF!

Thus, even with a—say—200-Ω emitter resistor, MOSFET gate fall time is $3R_1C = 3 \times 200 \times 4500 = 2.70$ μs—obviously impossibly long if switching frequencies of over 100 kHz are to be achieved. Hence driving the high MOSFET gate input capacitances requires the use of active transistors which can source and sink currents of about 200 to 400 mA.

For use with SG1524-type PWM chips which have only source or sink capability, but not both, a simple but, in most cases, adequate

MOSFET gate driver is the NPN-PNP emitter-follower totem pole of Fig. 8.17, where $Q2$, $Q3$ are 2N2222A, 2N2907A transistors in TO18 cans costing under 10¢. They are capable of sourcing and sinking 800 mA with current rise times of about 60 ns. The upper transistor of the totem pole can be eliminated by using the PWM chip's output transistor as a source driver as shown in Fig. 9.6c. Although it has less source capability (200 mA) than the external 2N2222A, in many instances it is adequate.

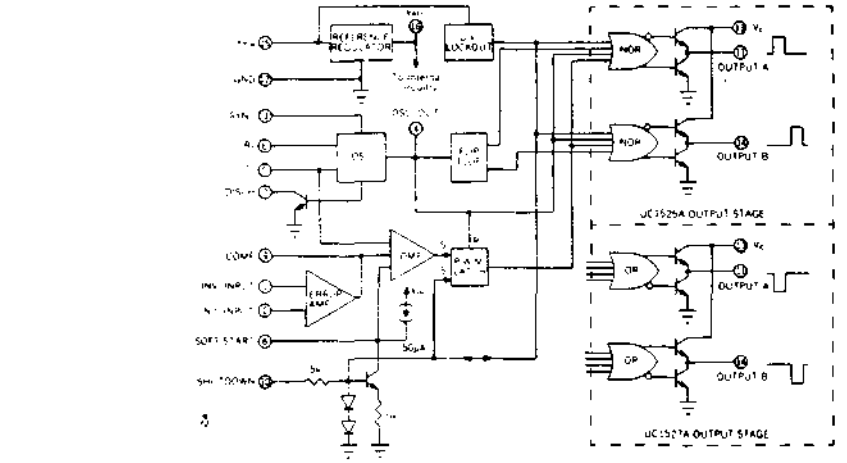
Second-generation PWM chips such as the Unitrode Corp. UC1525A[3] have a built-in totem pole consisting of an NPN emitter-follower resting on top of an NPN inverter. These transistors can source and sink over 200 mA. The emitter-follower and the inverter are driven by 180° out-of-phase signals so that when either is on, the other is off as shown in Fig. 9.7*a*. By using a transformer with two isolated secondaries as in Fig. 9.7*b*, the top and bottom transistors of a half or full bridge which are at different DC voltage levels can be driven. Similarly, of course, the MOSFET gates which are at the same DC voltage level in a push-pull circuit can be driven.

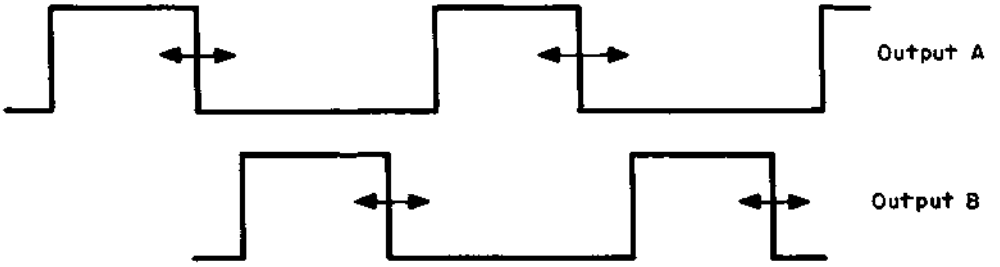
In Fig. 9.7*b*, the transformer primary is connected between pins 11 and 14 of the PWM chip. During the on time in one half period, when pin 11 is, say, positive with respect to pin 14 and is sourcing 200 mA, it is at about $+(V_h - 2)$ V with respect to ground and pin 14 is at about +2 V with respect to ground. During the dead time within that half period, pins 1, 14 are both short-circuited to ground. In the next half period, the polarity across pins 11, 14 reverses for the on time in that half period. Thus for ±10 V across the primary, the supply voltage V_h should be about 14 V.

9.2.5 MOSFET R_{ds} temperature characteristics and safe operating area limits[4,5]

The most common failure mode in bipolar transistors—second breakdown—comes about because their on-voltage drop $V_{ce(sat)}$ decreases with temperature. This imposes limits (RBSOA curve of Fig. 8.4) on the $I_c - V_{ce}$ trajectory the transistor may not cross during the turnoff transition. Manufacturers state that only a single crossing of this limit curve may cause the bipolar to fail in the second-breakdown mode.

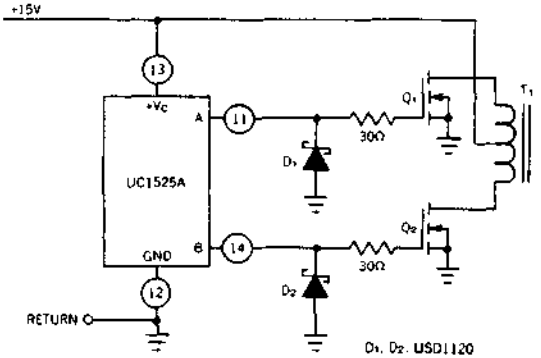
However, MOSFETs, because their on-voltage drop or r_{ds} increases with temperature, do not suffer from second breakdown and consequently have a much larger switching SOA as shown in Fig. 9.8. This is the boundary which if crossed for over 1 μs during either the turnon or turnoff trajectory may damage or destroy the transistor. The bound-

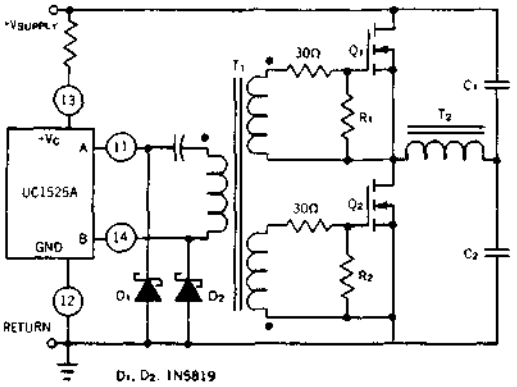




(a)

Figure 9.7 (*a*) Second-generation PWM chip UC1525A with totem-pole outputs (outputs *A* and *B*) for sourcing and sinking high currents. (*Courtesy Unitrode Corp.*) (*b*) Totem-pole outputs at *A*, *B* can drive high input capacity of MOSFETs either directly when gates are as same DC voltage level as in a push-pull or via a transformer when gates are at different DC voltage levels as in a half or full bridge.





(b)

Figure 9.7 *(Continued)*

ary limits are the maximum pulsed current ($I_{dm,pulsed}$) and maximum drain-to-source V_{dss} voltage ratings for the device.

An explanation of why the negative temperature coefficient of V_{ce} causes second breakdown in the bipolar and the V_{ds} positive temperature prevents it in the MOSFET is as follows. Second breakdown in bipolars comes about because of local hot spots on the chip. These hot spots are considerably hotter than the average junction temperature as calculated from the chip's junction to case thermal resistance and total transistor dissipation. Such a calculation assumes a uniform distribution of current carriers throughout the collector area.

If current is not uniformly distributed but is, for some reason, initially bunched up or crowded into a small fraction of the collector area, that area runs slightly hotter than the rest of the chip. Further, since the collector-to-emitter resistance decreases with increasing tempera-

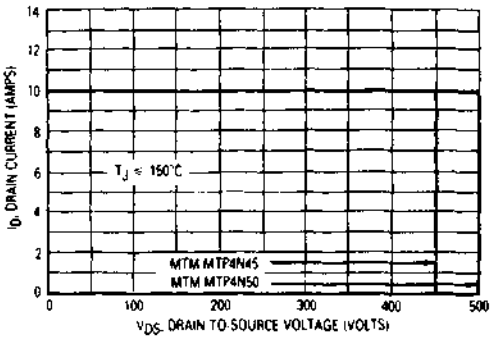


Figure 9.8 Reverse-biased safe operating area of 4 A, 400-V MOSFET. This is a far larger area than that for a 4-A 400-V bipolar transistor. For the MOSFET it is a rectangle bounded by I_{dm} (maximum pulsed current rating) which is two to three times I_d (maximum continuous current rating) and V_{dss}, the maximum drain-to-source voltage rating. The switching safe operating area is the boundary that the load line may traverse without incurring damage to the device. The fundamental limits are the maximum rated peak drain current I_{dm}, the minimum drain-to-source breakdown voltage $V_{(br)dss}$ and the maximum rated junction temperature. The boundaries are applicable for both turnon and turnoff of the devices for rise and fall times of less than 1 μs. (*Courtesy Motorola Inc.*)

ture, any incipient local hot spot has slightly less resistance than its surrounding areas and robs current carriers from adjacent areas. This causes the hot spot to get even hotter, decrease even more in resistance and rob even more current from adjacent areas. This process builds up rapidly until the local hot spot reaches a high current density and temperature sufficient (>200°C) to cause failure.

However, MOSFETs, with their positive r_{ds} temperature coefficient, tend to wipe out or cool off incipient local hot spots. If a point on the chip started operating at a slightly higher current density than its neighbors, its temperature would rise slightly. Because r_{ds} has a positive temperature, its resistance would increase and it would shift off some of its current carriers to neighboring areas and cool down. The result is the much larger SOA of Fig. 9.8 for the MOSFET as compared to Fig. 8.2 for the bipolar.

Curves showing the variation of r_{ds} with temperature and drain current are seen in Fig. 9.9 for a typical 15-A, 450-V MOSFET (MTM15N45). The variation of r_{ds} with temperature is also dependent on the voltage rating of the MOSFET as shown in Fig. 9.10. It is seen

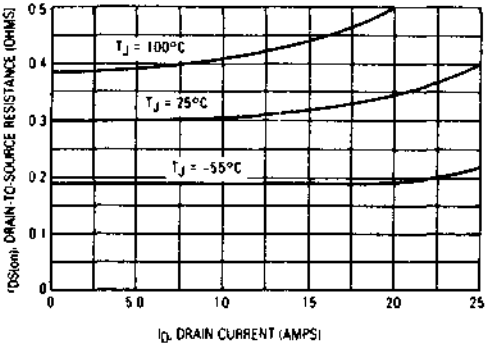


Figure 9.9 Variation of $r_{ds(on)}$ with drain current and temperature for the MTM15N45 (*Courtesy Motorola Inc.*)

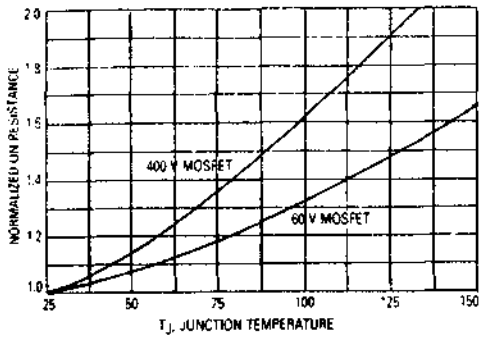


Figure 9.10 The influence of junction temperature on resistance varies with breakdown voltage. (*Courtesy Motorola Inc.*)

there that higher-voltage MOSFETs have larger r_{ds} temperature coefficients than do lower voltage devices.

9.2.6 MOSFET gate threshold voltage and temperature characteristic[4,5]

The major American MOSFET manufacturers—Motorola and International Rectifier—specify the gate threshold voltage V_{gsth} differently. Motorola specifies it as the gate-to-source voltage for which I_{ds} equals 1 mA at $V_{ds} = V_{gs}$. International Rectifier (IR) defines it as the gate-to-source voltage for which $I_{ds} = 0.25$ mA at $V_{ds} = V_{gs}$. There appears to be a two-to-one production spread in V_{gsth} in either Motorola or IR devices.

The gate threshold voltage V_{gsth} has a negative temperature coefficient and falls about 5 percent for each 25°C rise in temperature (Fig. 9.11) and quickly falls to zero in a radiation environment.[6] A negative gate-to-source voltage is required to keep the MOSFET off under radiation. The transfer characteristic is strongly dependent on type, duration, and intensity of radiation (Fig. 9.12). Discussion of MOSFETs in a radiation environment is beyond the scope of this text.

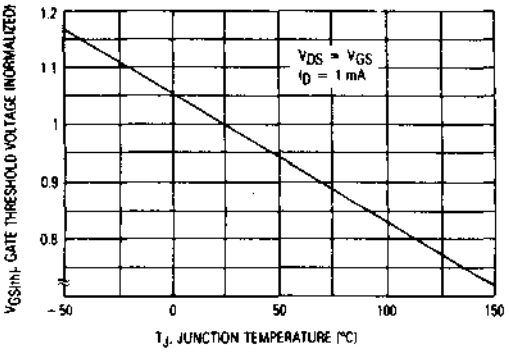


Figure 9.11 Gate threshold voltage variation with temperature. (*Courtesy Motorola Inc.*)

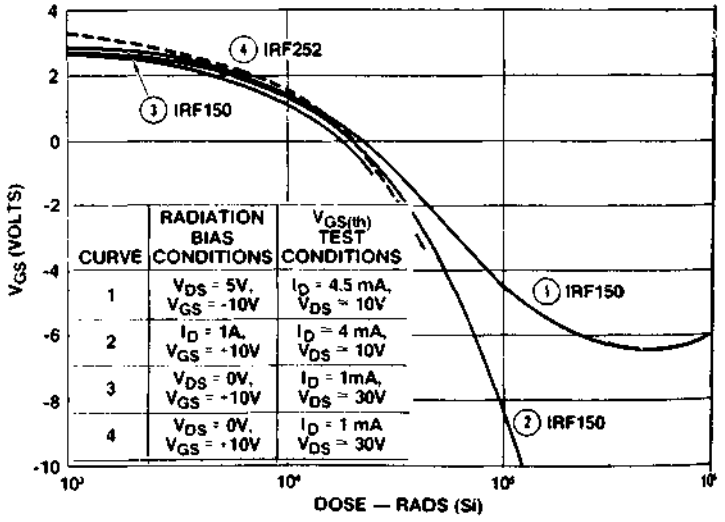


Figure 9.12 Typical gate threshold voltage versus total gamma dose. Under radiation, a negative gate bias is required to keep a MOSFET turned off. Gate threshold voltage depends strongly on the type, intensity, and duration of radiation. (*Courtesy Unitrode Corp.*)

9.2.7 MOSFET switching speed and temperature characteristics

MOSFET switching speed is significantly independent of temperature. Drain current rise and fall times depend only on the time required for the gate voltage to cross the narrow boundary between the gate threshold voltage (V_{gsth}) and V_{gl} in Fig. 9.3*b*. This depends on the output resistance of the source-sink driver and the effective gate input capacity. The source-sink output resistance is usually a discrete exter-

nal resistance which has a low-temperature coefficient. Further, since gate input capacity is so independent of temperature, drain current rise and fall times are also independent of temperature.

Turnon and turnoff delays are somewhat temperature-dependent. Turnon delay is the time for the gate voltage to rise from 0 to the threshold voltage V_{gsth}. Since V_{gsth} falls 5 percent for each 25°C rise in temperature, turnon delay will decrease with temperature. Also since there is a two-to-one production spread in V_{gsth}, turnon delay will vary from device to device even at the same temperature. Note, though, that devices with large variation in V_{gsth} may not have a large variation in delay to turn on to a relatively large specific current. The lower tails of the transfer characteristic can vary significantly without changing V_{gl}—the gate voltage for a given current I.

Turnoff delay is the time required for the gate to fall from its usual on voltage of 10 V to V_{gl} (Fig. 9.3*b*). Since gate threshold voltage and transconductance vary with temperature, so will turnoff delays.

Turnon and turnoff delays are significant only when paralleling MOSFETs.

9.2.8 MOSFET current ratings

For bipolar transistors, maximum output current is limited by the fact that current gain falls drastically as output current rises. Thus unacceptably high base input currents are required as output current increases. This is shown in Fig. 9.13 for the 2N6542—a typical 5-A, 400-V bipolar transistor.

With MOSFETs, however, output-input gain (transconductance or dI_{ds}/dV_{gs}) does not decrease with output current, as can be seen in Fig. 9.14. Thus the only limitation on drain current is power dissipation or maximum MOSFET junction temperature.

Maximum MOSFET junction temperature is 150°C. Good standard design practice is to derate this to 105°C or at most to 125°C.

Manufacturers rate the current carrying capability of their device in terms of I_d, the maximum continuous drain current.

Apparently Motorola and IR, the major American MOSFET manufacturers, have rated I_d as that current, which at the maximum $V_{ds(on)}$ voltage for that current yields a power dissipation at 100 percent duty cycle such that when multiplied by the thermal resistance brings the MOSFET junction temperature to the maximum of 150°C when the transistor case is at 100°C. Thus

$$dT = 50 = PDR_{th} = V_{ds(on)} I_d R_{th} \qquad \text{or} \qquad I_d = \frac{50}{V_{ds(on)} R_{th}}$$

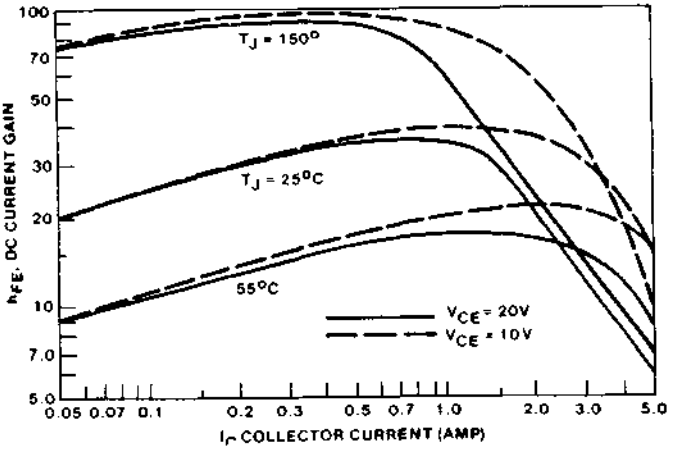


Figure 9.13 Typical DC current gain for the 2N6542/3 bipolar transistor. Gain of a bipolar transistor falls off with increasing output current, but that of a MOSFET does not. Maximum current in a MOSFET is limited only by junction temperature rise. (*Courtesy Motorola Inc.*)

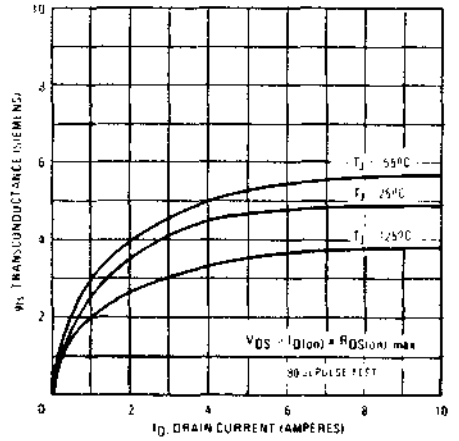


Figure 9.14 Transconductance versus drain current of the IRF330. Gain of a bipolar transistor falls off with increasing output current, but that of a MOSFET does not. Maximum current in a MOSFET is limited only by junction temperature rise. (*Courtesy International Rectifier.*)

where $V_{ds(on)}$ is the maximum drain-to-source on voltage at 150°C and R_{th} is the thermal resistance in degrees Celsius per watt.

This I_d rating is not a useful guide for selecting a MOSFET for a given peak current in a switching supply application. In such usage, duty cycle is never 100 percent. For reliability, it is desired to operate at junction temperatures derated to 125°C or the usual military-specified 105°C. But it does show the relative current-carrying capability of various MOSFETs operated at 100 percent duty cycle.

In general, there are two ways of selecting a MOSFET for a specified output power in a switching supply. First, the equivalent flat-topped primary current I_{pft} is calculated for specified output power and minimum DC input voltage. This current is given in Eqs. 2.9, 2.28, 3.1, and 3.7 for the push-pull, forward converter, half and full bridge, respectively. Then for these currents, a MOSFET of r_{ds} is chosen so that the on drain-to-source voltage $I_{pft}r_{ds}$) is a small percentage (usually no more than 2 percent) of the minimum supply voltage so as to rob no more than 2 percent of the transformer's minimum primary voltage.

In selecting a device with a desired r_{ds}, it should be recalled that the data sheets give it at a case temperature of 25°C. Also noteworthy is the fairly large variation of r_{ds} with temperature and device voltage rating as shown in Figs. 9.9 and 9.10. Thus, Fig. 9.10 shows that an r_{ds} of a 400-V MOSFET at 100°C is 1.6 times its value at 25°C.

As a design example, consider a 150-W forward converter operating from a nominal 115-V-AC line. Assume that maximum and minimum rectified DC voltages are 184 and 136 V, respectively. Peak flat-topped current from Eq. 2.28 is $I_{pft} = 3.13(150/136) = 3.45$ A, then, for the MOSFET on voltage to be 2 percent of the minimum supply voltage, $V_{on} = 0.02 \times 136 = 2.72 = I_{pft}r_{ds} = 3.45r_{ds}$ or $r_{ds} = 0.79\Omega$ at, say, 100°C or $0.79/1.6 = 0.49\ \Omega$ at 25°C.

Possible plastic-cased 450-V MOSFET choices from the Motorola catalog are shown in Table 9.1.

The choice would be made on an engineering judgment of the relative importance of cost and performance. The 7-A MTH7N45 is not quite good enough. It would have more than the sought on drop of 2.72 V at 3.45 A. That itself is not prohibitive; it would run at a junction temperature somewhat higher than the MTH13N45. That would have to be weighed against the higher cost of the 13-A MTH13N45.

An alternative way to chose the MOSFET is to preset its maximum junction temperature—say, to 100°C for reliability. Then assume a reasonably low MOSFET junction-to-case temperature rise so as not to require too low a heat-sink temperature—for, of course, the further the case temperature is below the MOSFET junction, the larger must be the heat sink on which the case rests.

TABLE 9.1 Motorola MOSFETs

Device	I_d, A	V_{dss}, V	r_{ds}, Ω at 25°C
MTH7N45	7	450	0.8
MTH13N45	13	450	0.4

Thus assume a reasonably low 5°C case-to-junction temperature rise. Then

dT = 5 = power dissipation × thermal resistance (junction to case)

and assuming that there are negligible AC switching losses

$$dT = (I_{rms})^2(r_{ds})R_{th} \qquad \text{or} \qquad r_{ds} = \frac{5}{(I_{rms})^2 R_{th}}$$

where I_{rms} is the rms current in the MOSFET.

For a forward converter where the maximum on time per period is $0.8T/2$, the rms current is $I_p(\sqrt{t_{on}/T}) = 0.632\, I_p$. Then for the preceding design example of a 150-W forward converter where I_p is 3.45 A, we obtain

$$r_{ds} = \frac{5}{(0.632 I_p)^2 R_{th}}$$

and for the usual thermal resistance of 0.83°C/W for MOSFETs of this current and package size

$$r_{ds} = \frac{5}{(0.632 \times 3.45)^2 \times 0.83}$$
$$= 1.26\ \Omega \text{ at } 100°C$$

This is the r_{ds} at 100°C junction temperature, which causes a 5°C junction-to-case temperature differential with a 3.45-A peak current pulse at 0.4 percent duty cycle. The r_{ds} at 25°C junction temperature is then 1.26/1.6 or 0.78 Ω.

Thus, on the basis of a 5°C junction-to-case temperature differential, Table 9.1 shows that the MTH7N45 would be an adequate choice. But it would rob 1.26 × 3.45 = 4.34 V from the primary voltage at 100°C. The MTH13N45 would be a better choice if its somewhat higher cost were acceptable.

9.2.9 Paralleling MOSFETs[7]

In paralleling MOSFETs, two situations must be considered: (1) whether the paralleled devices share current equally—in the static case when they are fully on and (2) whether they share current equally during the dynamic turnon-to-turnoff transition.

With paralleled MOSFETs, in either the static or dynamic case, the concern is that if one MOSFET hogs a disproportionate part of the current, it will run hotter, and long-term reliability will decrease or, in the short run, will fail.

Unequal static current sharing comes about because of unequal r_{ds}

between paralleled devices. The lower r_{ds} device draws more than its numerical share of the total current—just as with a group of paralleled discrete resistors, the lowest resistor draws the most current.

In describing the absence of second breakdown with MOSFETs, it was pointed out that this is so because of the positive temperature coefficient of r_{ds} in MOSFETs. Thus, if a small portion of the chip tends to hog a disproportionate part of the total current, it runs hotter, its r_{ds} increases, and it shifts off some of its current to some neighboring areas to equalize current density.

This mechanism also works to some extent with paralleled discrete MOSFETs. By itself, however, it is not sufficient to minimize the temperature of the hottest device. This is so because the temperature coefficient of r_{ds} is not very large and a large temperature differential between devices is required to shift off excess current. But with a large required temperature differential, the hottest device is at a high temperature, and this is exactly what reduces reliability and is to be avoided.

The mechanism does work well within a chip because all elementary areas of the chip are thermally coupled. But it does not work as well if thermal coupling between the parallel MOSFETs is poor as with discrete MOSFETs in separate cases and physically separated on a common or in the worst case on separate heat sinks.

To improve the equality of static current sharing, MOSFETs, if in discrete cases, should be located as close as possible on the same heat sink. If possible, MOSFETs in discrete chips should be located close together on the same substrate within the same package. Such packages containing multiple paralleled MOSFETs on a common substrate–heat sink are currently available from a number of manufacturers. As a last resort, if MOSFETs in discrete packages must be used, and close location on a common heat sink does not suffice, matching the r_{ds} of paralleled devices will guarantee equal current sharing.

For equal dynamic current sharing, the transconductance curves of paralleled devices must lie on top of one another. This is obvious from Fig. 9.15. If all gates have identical voltage at the same time, and the transconductance curves are not superimposed, the drains carry different currents at that time, on either turnon or turnoff.

It is not so important that the gate thresholds match. If n devices are to be paralleled for a total current of I_t, they should be matched to have as closely as possible the same I_t/n at the same gate voltage—even if there is a large mismatch in gate threshold voltages.

Symmetrical circuit layout is also important for equal dynamic current sharing (Fig. 9.16). Lead lengths from the common output point of the gate driver to the gate terminals should be equal. Leads from

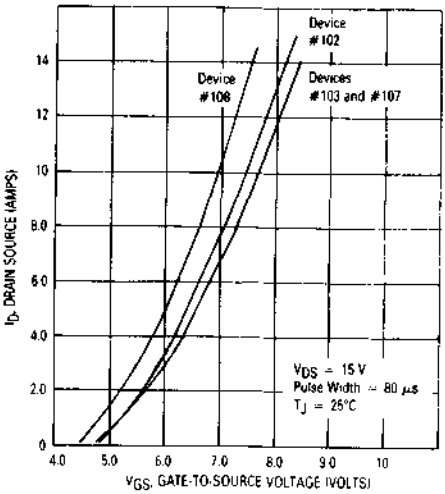


Figure 9.15 Widest variation in transconductance curves of 250 additional MTPBN2O. For equal dynamic current sharing, MOSFET transconductance curves must coincide. (*Courtesy Motorola Inc.*)

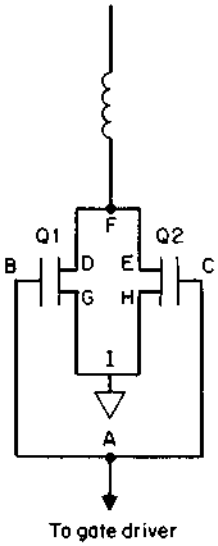


Figure 9.16 Symmetrical circuit layout for equal dynamic current sharing in parallel MOSFETs. To ensure equal dynamic current sharing in parallel MOSFETs, circuit layout should be symmetrical. Thus $AB = AC$, $GI = HI$, $DF = EF$.

the source terminals directly at the MOSFETs to a common tie point should be equal, and that common tie point should be brought directly as possible to a common tie point on the ground bus. That ground bus tie point should be common (with as short a lead as possible) with the negative rail of the housekeeping supply. Then, finally, to prevent oscillations with paralleled MOSFETs, resistors of 10 to 20 Ω or ferrite beads should be placed in series with gate leads.

9.2.10 MOSFETs in push-pull topology

MOSFETs in the push-pull topology minimize the transformer flux-imbalance problem significantly. It will be recalled (Sec. 2.2.5) that if the volt-second product applied to the transformer during one half period is not equal to that in the next half period, the transformer core moves off center of its hysteresis loop. After a number of such periods, the core drifts very far off center, saturates, and cannot support the supply voltage, and the transistor is destroyed.

MOSFETs alleviate this problem for two reasons. First, there is no storage time with MOSFETs and for equal gate on times, drain voltage on times are always equal in alternate half periods. There is hence no inequality in volt-second product applied to the transformer due to unequal transistor on times. In contrast, with bipolars, the main reason for unequal volt-seconds on alternate half periods is the inequality in storage times. Second, with MOSFETs, the positive temperature coefficient of r_{ds} acts in a negative-feedback way to prevent severe flux imbalance. If there is a certain amount of flux imbalance, the core walks partly up its hysteresis loop. This causes the magnetizing current and hence total current on one half period to be larger than that on the alternate half period (Figs. 2.4*b* and 2.4*c*). But now the MOSFET with the larger peak current runs hotter and its r_{ds} increases, increasing its on drop. This robs voltage from its half primary, its volt-second product decreases, and the core moves back down toward the center of its hysteresis loop.

Qualitatively, both these effects are in the direction to prevent a catastrophic flux imbalance. But it is not easily feasible to demonstrate quantitatively that it works at all power levels, at all temperatures, and for all core materials. A certain solution to the flux-imbalance problem is to go to current-mode topology (Sec. 2.2.8.5). But if for some reason it is not desired to go to current mode, MOSFETs in a conventional push-pull circuit may be used with no flux-imbalance problem up to power levels of 150 W; many designers have used this approach successfully.

One final interesting advantage of MOSFETs in a push-pull circuit is shown in Fig. 9.17. This simple scheme makes it possible to totally eliminate any residual flux imbalance due possibly to differences in r_{ds}. If there is an incipient flux imbalance, it manifests itself by the current in one transistor being somewhat larger than the current in the other (Figs.2.4*b* and 2.4*c*). The MOSFET with the higher peak current has, for some reason, a larger volt-second product than the other device.

This is corrected as illustrated in Fig. 9.17 by adding a small, empirically selected resistor r_b in series with the gate of the MOSFET with the larger peak current. This integrates away some of the front

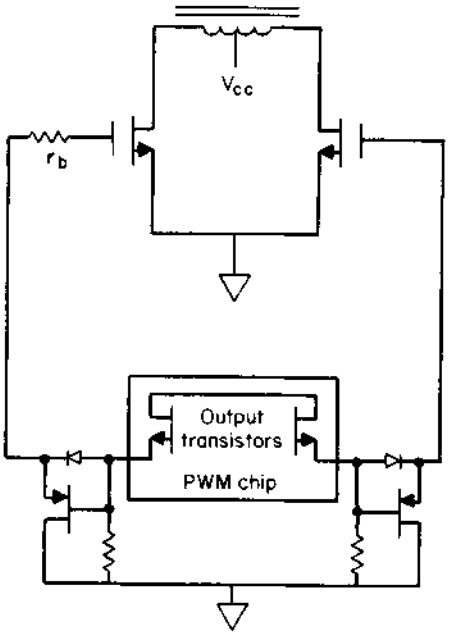


Figure 9.17 Flux-imbalance correction in a MOSFET push-pull circuit. An empirically selected resistor inserted in series in the gate of the MOSFET which originally had the higher peak current forces currents in Q_1, Q_2 to be exactly equal. It does this by integrating away part of the front end of the gate drive pulse and forces the volt-second product on each transformer half to be equal.

edge of its gate turnon pulse, narrows it, equalizes the volt-second products on each side, and equalizes peak currents on alternate periods.

Although some (military) programs do not permit empirical "Select At Test" components, once a resistor is selected, it corrects peak current imbalance independently of temperature, supply voltage, and load current. Thus the only objection to the scheme is that if either transistor is changed, a new "Select At Test" resistor must be chosen. This may be unacceptable in the field where the proper test equipment is not available.

9.2.11 MOSFET maximum gate voltage specifications

Most MOSFET gates have a maximum gate-to-source voltage specification of ±20 V. The gate is easily destroyed if that limit is exceeded.

A problem arises in this respect when a MOSFET has a gate input resistor and is turned off in a circuit with a large supply voltage. Thus, consider a forward converter operating from a nominal supply voltage of 160 DC V (maximum of 186 V). Then, when the MOSFET turns off (at maximum supply voltage), its drain goes to twice the supply voltage or 372 V. A fraction of this positive-going front edge is coupled back and voltage divided down by C_{rss} and C_{iss}. For the

MTH7N45, C_{rss} = 150 pF, C_{iss} = 1800 pF. Then the voltage coupled back down to the gate is 372 × 150/(150 + 1800) = 29 V.

This may not damage the gate even though it exceeds the 20-V limit. The gate resistor will also load down and decrease its amplitude. But it is close to the point of causing a possible failure as line transients and the leakage inductance spike have not yet been considered. Thus, good design practice is to shunt the gate with an 18-V zener diode. Note that a high frequency oscillation may often result from the capacitive drain to gate feedback.

9.2.12 MOSFET drain-to-source ("body") diode

Inherently in solid-state structure of a MOSFET, a parasitic "body" diode is located across the drain-source terminals as shown in Fig. 9.18.

The diode polarity is such as to prevent reverse voltage across the MOSFET. The forward current handling capability and reverse voltage rating of the diode are identical to those of the MOSFET itself. Its reverse recovery time is faster than a conventional AC power rectifier diode, but not as fast as discrete fast-recovery types. Manufacturers' data sheets show the diode reverse recovery time for each specific MOSFET.

The diode is of no importance for most switching supply topologies

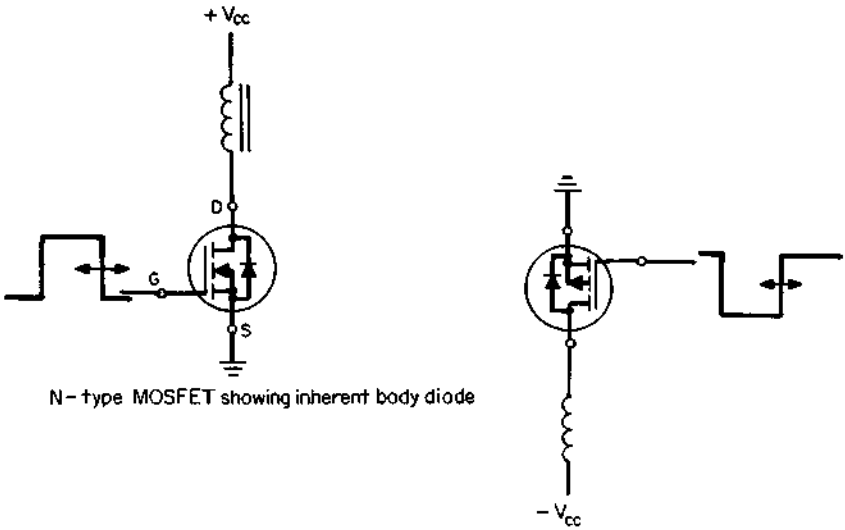


Figure 9.18 Inherent body diodes in N- and P-type MOSFETs. In the N-channel MOSFET, the diode prevents a negative drain-to-source voltage. In the P-channel MOSFET, the diode prevents a positive drain-to-source voltage.

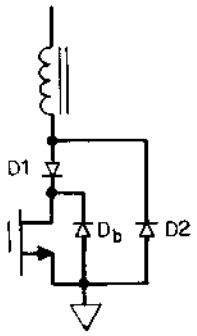


Figure 9.19 Scheme to avoid using a MOSFET body diode when reverse current must be allowed to flow around a MOSFET. To prevent forward current from flowing through the MOSFET's body diode which does not have good reverse recovery characteristics, blocking diode D_1 is added in series in the drain. D_2 with better reverse recovery characteristics is shunted around the D_1 anode and the Q1 source.[9]

as the drain to source is never subjected to a reverse voltage (drain negative with respect to source for an N-type MOSFET, positive with respect to source for a P-type MOSFET).

There are some exceptions—specifically, the half- or full-bridge topologies of Figs. 3.1 and 3.3. But in these circuits there is almost always a dead time between the time the diode conducts (when it is returning to the supply source, energy stored in the transformer leakage inductance) and the time it is subjected to reverse voltage. Because of this delay between forward current and reverse voltage, the relatively poor reverse recovery time of the MOSFET body diode is not harmful.

However, if a specific new circuit configuration requires reverse voltage across the MOSFET, a blocking diode must be placed in series with the drain. Various motor drive circuits or circuits with highly inductive loads may have problems because of the body diode.[8]

High-frequency resonant circuit topologies (Chap. 13) frequently must be able to support reverse voltage immediately after carrying forward current. When this is necessary, the circuit of Fig. 9.19 is used. Diode D_1 prevents forward current from flowing through the MOSFET body diode, and the fast-reverse-recovery-time diode D_2 carries the required forward current.

References

1. *International Rectifier Power MOSFET Data Book,* HDB-3, International Rectifier, El Segundo, CA.
2. *Power MOSFET Transistor Data,* Motorola Inc., Motorola Literature Distribution, Phoenix, AZ.
3. *Linear Integrated Circuits Databook—PWM chip UC3525A,* Unitrode Corp., Merrimack, NH.
4. Reference 1, Chaps. 1, 4.
5. Reference 2, Chap. 2.
6. Radiation Resistance of Hexfets, Ref. 1, p. B10.
7. Reference 1, Chap. 1; Ref. 2, Chap. 7; International Rectifier Application Note AN-941.
8. "Hexfet's Integral Body Diode," Ref. 1, p. A65.
9. Reference 1, page A65.

Chapter

10

Magnetic-Amplifier Postregulators

10.1 Introduction

In Secs. 2.2.1 and 2.3.3, multi-output-voltage push-pull and forward converter topologies were discussed. As was described, in either circuit, a feedback loop is closed around a main or "master" (usually the highest-current or 5-V) output. The feedback loop keeps the master output constant against line or load changes.

Additional secondaries on the power transformer yield "slave" output voltages whose magnitudes are proportional to their respective numbers of turns. These slaves are operated open-loop. Their on time, or duration of high voltage at the secondaries, is fixed by the master feedback loop so as to keep the master output voltage constant. These on times are unaffected by the slave output currents and are proportional to the DC supply voltage.

These slave voltages are thus as well regulated against line input changes as is the master. But the slaves are not well regulated against load current changes—either in the master or in themselves. Slave output change due to current changes in the master is referred to as *cross regulation* and may be as high as ±8 percent for the maximum specified current change in the master.

Slave output changes due to current change in themselves are considerably less so long as the master or slave output inductor does not go discontinuous or run dry (Sec. 2.2.4). If either the master or an individual slave output inductor is permitted to go discontinuous by decreasing its output current, that slave DC output voltage may change up to 50 percent.

If their magnitudes of inductance are chosen very large, the output inductors can be kept from going discontinuous. But larger output in-

ductors cause larger and longer lasting output voltage transients in response to step output current changes.

A final drawback to open-loop slave output voltages is that they cannot be set precisely to a desired value. Their DC voltage level can be set to within only 1 or 2 V of a specific value.

The preciseness of setting depends on the volts per turn of the transformer core. Further, since either primary or secondary can be changed only in whole multiples of a single turn, output voltage can only be changed in coarse steps. Moreover, since, from Faraday's law, the volts per turn is directly proportional to switching frequency, the coarseness of these steps increases with increasing frequency.

This scheme of a multioutput supply with a master, well regulated against line and load changes but with slaves poorly regulated against master or slave load current changes is nevertheless widely used.

Usually it is only the master—the 5-V output that feeds crucial logic circuits—which must be well regulated against line and load changes. Slaves usually feed motors on disk or tape drives or error amplifiers. Such loads can most often tolerate a DC voltage which is as much as 1 to 2 V off a specified nominal value. For motor drives, this changes the motor acceleration times only slightly. For various linear circuits, it changes internal dissipation only somewhat.

Yet there are numerous applications where the slave outputs must be precisely set to a specified value and must be well regulated (better than 1 percent) against line and load changes. Heretofore, when slaves well regulated against line and load changes were required, the solution was to postregulate a semiregulated slave with either a linear regulator for output currents of under 1.5 A or a buck regulator for higher output currents.

These approaches have their merits and drawbacks, which are discussed below. A better solution for poorly regulated slaves is the magnetic-amplifier postregulator.[1,2] It uses an old basic technique, but with a simpler circuit and better magnetic material, it made a dramatic reappearance about 7 years ago and has rapidly been adopted throughout the industry.

10.2 Linear and Buck Regulator Postregulators

A linear regulating postregulator is the best approach for output currents up to 1.5 A. The 1.5-A limit comes about because of low cost and low internal dissipation.

Linear regulators with up to 1.5 A of output current are available as integrated circuits in plastic TO220 packages at a cost of about 50¢. They require no additional external components other than a small filter capacitor.

They are usually specified as requiring 2 V (typically, 3 V for the worst case) minimum input-output differential or headroom as discussed in Sec. 1.2.3). Thus, for 1.0 A of output current at 3 V of headroom, their internal dissipation is 3.0 W.

They are also available as integrated circuits at much higher currents in metal TO66 or TO3 cases. But they are not widely used at currents above 1.0 A—not so much because of excessive junction temperature, for that can be handled by heat-sinking—but rather because of the excessive dissipation and consequent total supply inefficiency due to the 3-V headroom requirement, which limits their use at currents above 1.0 A.

Integrated-circuit linear regulators with only 0.5- to 1.0-V headroom requirement (Sec. 1.2.5) are available, but these are considerably more expensive.

Postregulators for output currents above 1.5 or possibly 3.0 A are most often implemented as buck regulators. The slave output voltage is usually set at a minimum of about 4 V above the desired output and is then bucked down (Sec. 1.3.1) to the desired output.

This yields better efficiency than does a linear step-down regulator. But it is more expensive in dollars, number of components, and required space. Further, the new frequency of the buck transistor introduces an additional source of RFI and may produce beats with the main switching frequency causing problems in other parts of the frequency spectrum.

The magnetic-amplifier postregulator, discussed below, is a better approach than a buck regulator at currents over 1.5 A and is even a credible alternative at lower currents.

10.3 Magnetic Amplifiers—Introduction

Referring to Figs. 2.1 and 2.10, it is seen that the slave output voltages are not well regulated, primarily because they are operated open-loop and their "high" time, or time that their secondary voltages are high, is determined only by the master feedback loop. If there were independent control of the slave secondary high times, through separate feedback loops, they, too, could be constant.

High time of the slave secondary of a forward converter can be controlled by a symbolic switch $S1$ as in Fig. 10.1. Switch $S1$ is in series between the slave secondary output terminal and its rectifying diode $D1$. By independently controlling its on/off-time ratio, the duty cycle of voltage supplied to the slave LC filter and hence the slave DC output voltage can be controlled.

Assume that the high time directly at the slave secondary output (before $S1$) is t_h out of a period T as in Fig. 10.1. The time t_h is set by

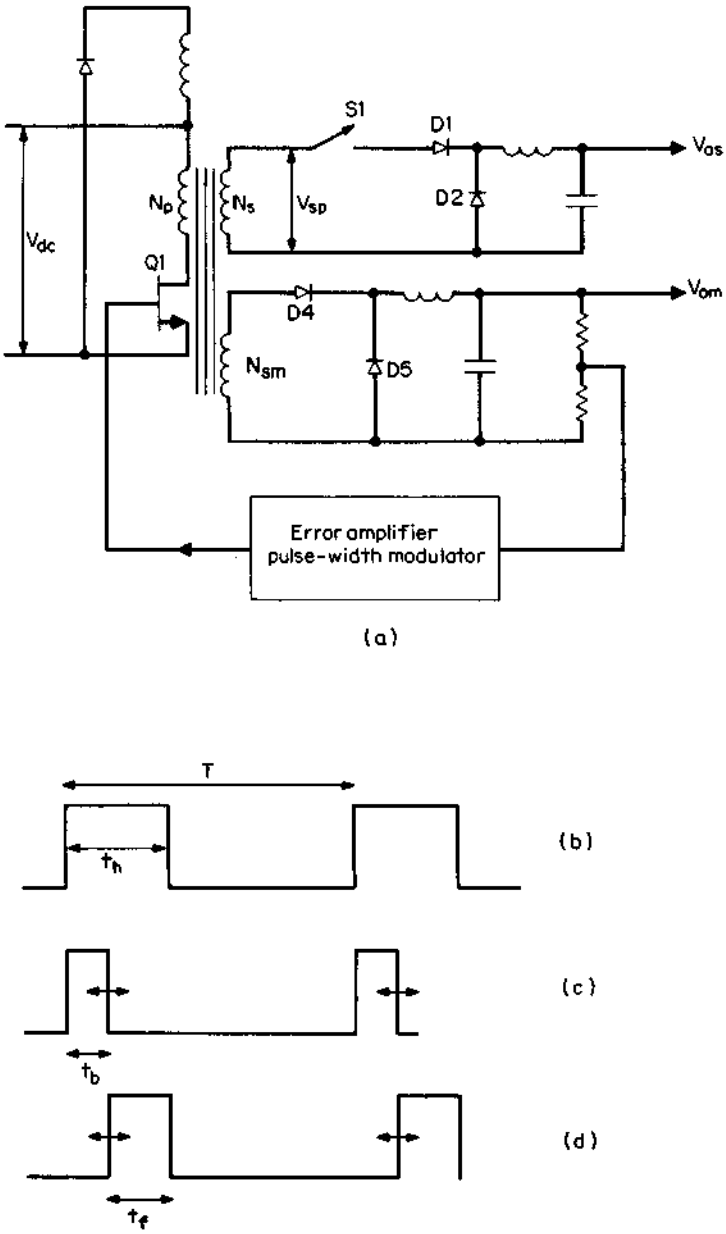


Figure 10.1 Width modulation of a slave secondary pulse with a symbolic switch *S1*. If the open and closed times of *S1* could be controlled in a separate feedback loop independently of the on and off times of *Q1*, the slave output voltage could be regulated independently of the master.

the main feedback loop to maintain constant main output DC voltage V_{om}. Thus

$$V_{om} = \left[(V_{dc} - V_{ce})\frac{N_{sm}}{N_p} - V_{D4}\right]\frac{t_h}{T} \tag{10.1a}$$

$$V_{om} \approx \left[(V_{dc} - 1)\frac{N_{sm}}{N_p} - 1\right]\frac{t_h}{T} \tag{10.1b}$$

Of this time, t_h, assume that $S1$ is open and blocking the slave secondary voltage V_{sp} from getting through to the $D1$ anode for a time t_b (Fig. 10.1*c*) and that $S1$ is closed with zero resistance for a firing time t_f (Fig. 10.1*d*). The slave DC output voltage is then

$$V_{os} = (V_{sp} - V_{D1})\frac{t_f}{T} \tag{10.2a}$$

$$V_{os} \approx (V_{sp} - 1)\frac{t_f}{T} \tag{10.2b}$$

Now $t_f + t_b = t_h$, or

$$t_f = t_h - t_b \tag{10.3}$$

The slave output voltage will be kept constant, as indicated by Eq. 10.2, by controlling t_f. But the physical nature of the switch $S1$ is such that it is not t_f directly that will be controlled, but rather t_b, the blocking time.

Thus in Eq. 10.2a, if the peak secondary voltage $V_{sp} = (V_{dc} - V_{Q1})(N_s/N_p)$ varies—say, increases—because the supply voltage V_{dc} increases, t_f will be decreased by increasing t_b or cutting away a larger piece of the front end of the t_h pulse.

The switch $S1$ is a *magnetic amplifier*. It consists simply of a toroidal magnetic core of square hysteresis loop material with a few turns of wire. It works as follows.

10.3.1 Square hysteresis loop magnetic core as a fast acting on/off switch with electrically adjustable on/off time

Figure 10.2 shows the BH loop of a typical square hysteresis loop material (Toshiba MB amorphous core material[3]). Other square-loop materials usable in magnetic amplifiers will be discussed below.

The slope of the BH loop ($\mu = dB/dH$) is its permeability. A coil wound around the core will have an impedance proportional to the core permeability μ. Thus, so long as the core is on the vertical of its

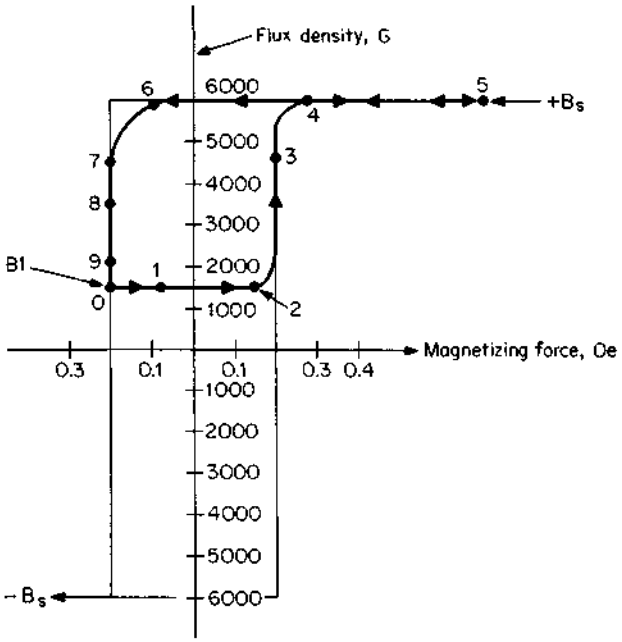


Figure 10.2 Toshiba MB amorphous core, *BH* loop at 100 kHz. In magnetic-amplifier operation, the core moves along a minor loop O1234567890. In going from 1 to 4, the core is on the steep part of the hysteresis loop and the MA has high impedance. At point 4, the core saturates and the MA has essentially zero impedance. At the end of the $Q1$ on time (Fig. 10.1), the core is reset to B_1. The time to move from B_1 to $+B_s$ is the switch-open time. The further down B_1 is pushed, the longer the blocking or switch-open time. The level to which B_1 is reset is determined by the current forced into the no-dot end of MA by $Q2$ (Fig. 10.3). That current is controlled by the error amplifier.

hysteresis loop, its permeability and hence the coil impedance is very high. It is effectively a single-pole switch in the open position.

When the core is in saturation on the horizontal axis of its hysteresis loop (beyond point 4 in Fig. 10.2), the *BH* loop is so square that the slope dB/dH or permeability is unity. The coil impedance is thus the very low impedance of an air core coil of an equal number of turns. The coil is thus effectively a single-pole switch in the closed position.

Such a core with a few turns of wire constitutes the switch $S1$ of Fig. 10.1. It is shown in Fig. 10.3 with the error amplifier and the scheme for controlling the on/off time of the switch.

Throughout one switching cycle (t_0 to t_3 in Fig. 10.4) the core moves around a so-called minor hysteresis loop—the path 01234567890 shown in Fig. 10.2. The blocking (t_b) and firing (t_f) times of the magnetic core switch are controlled as follows. Assume that at the start of a cycle (t_0),

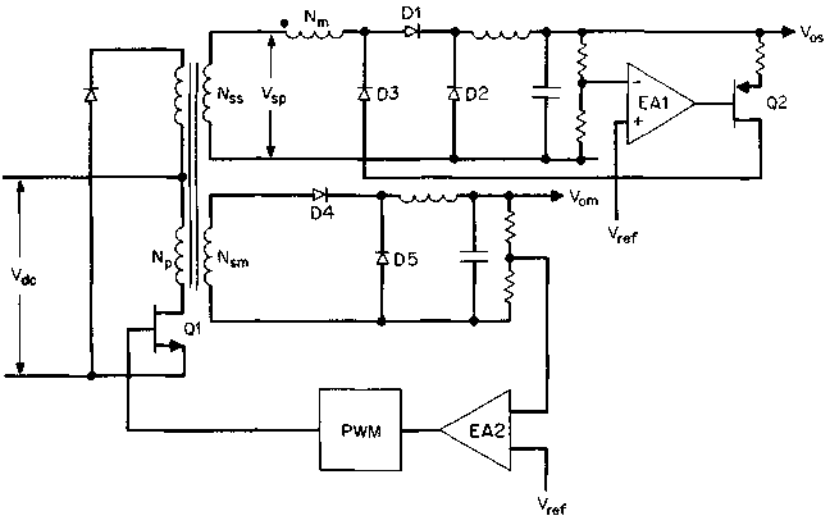


Figure 10.3 Magnetic-amplifier regulation of a slave output. The switch $S1$ of Fig. 10.1 is implemented with magnetic amplifier MA. The MA has high impedance as long as it is below saturation on the steep part of its hysteresis loop. That high-impedance time is determined by the amount of reset current pushed into the no-dot end of MA during the $Q1$ off time. The MA is simply a square hysteresis loop core with a few turns of wire.

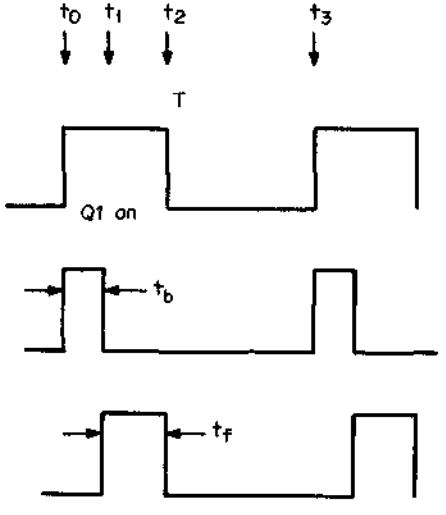


Figure 10.4 Critical timing intervals for magnetic amplifier (Fig. 10.3).

the core had been pushed down to B_1 on Fig. 10.2. When $Q1$ turns on, a voltage $V_{sp} \approx (V_{dc} - 1)N_{ss}/N_p$ appears across the slave secondary.

Just prior to t_0, $D2$ had been conducting and its cathode was one diode drop below ground. At t_0, a voltage $(V_{sp} + 1)$ appears across the magnetic amplifier MA and $D1$ in series in the direction to drive the core up toward saturation at point P4 (Fig. 10.2).

Until the core gets up to point 4, only a very small "coercive" current flows through MA and $D1$ into $D2$. This coercive current is considerably smaller than the free-wheeling current $D2$ is carrying. Thus the $D2$ cathode does not move substantially; it remains at one diode drop below ground and the voltage V_{sp} is blocked from getting through and raising the voltage at the front end of L_s.

Diode $D1$ carries the MA coercive current, and although it is much smaller than the $D2$ current, assume that forward drops in $D1$ and $D2$ are equal (and equal to 1 V). The full V_{sp} voltage thus appears across MA, whose right-hand end remains at ground. The MA is now in its blocking phase and remains there for a time t_b equal to the time required to move the core up from B_1 to saturation at B_s (= point 4).

Once the core has reached saturation at point 4, the MA impedance becomes negligible within a few nanoseconds and it cannot support the voltage V_s. A large current flows from the top end of N_{ss} into $D2$ and unclamps it, and its cathode voltage rises to one diode drop ($D1$) below V_{sp} and remains there for the duration of the $Q1$ on time.

The time from the onset of saturation to the instant $Q1$ turns off is the firing time t_f. The LC output filter is thus presented with a square pulse of amplitude $(V_{sp} - V_{D1})$, duration t_f, and duty cycle t_f/T. It averages this pulse to produce the slave DC output voltage of $(V_{sp} - V_{D1})t_f/T$.

Regulation against line and load variation is achieved by controlling the firing time t_f. This is done indirectly by controlling the blocking time t_b.

10.3.2 Blocking and firing times in magnetic-amplifier postregulators

Blocking time t_b is calculated from Faraday's law:

$$V_{sp} = N_m A_e \frac{dB}{dT} 10^{-8}$$

$$= N_m A_e \frac{(B_s - B_1)}{t_b} 10^{-8}$$

or

$$t_b = N_m A_e \frac{(B_s - B_1)}{V_{sp}} 10^{-8} \qquad (10.4)$$

where N_m = number of turns on magnetic amplifier MA
A_e = MA core area, cm^2
B_s = saturation flux density, G
B_1 = starting point in flux density, G
V_{sp} = peak slave secondary voltage, V
t_b = blocking time, s (Fig. 10.4)

Now note (from Eq. 10.2b) $V_{os} = (V_{sp} - 1)t_f/T$ and (from Eq. 10.3) $t_f = t_h - t_b$.

If V_{sp} increases for any reason, V_{os} can be kept constant by decreasing the firing time t_f (Eq. 10.3), and from Eq. 10.2, t_f can be decreased by increasing t_b. From Eq. 10.4, t_b can be increased by decreasing B_1, that is, by pushing B_1 further down on the hysteresis loop. Similarly, if for some reason V_{os} decreases, t_f must be increased by decreasing the blocking time t_b. This is accomplished by increasing the flux level B_1 so that less time is required for the voltage V_{sp} to drive the core up into saturation.

10.3.3 Magnetic-amplifier core resetting and voltage regulation

Thus far only the transition of the core from its starting point B_1 up to saturation has been discussed. During t_f, the MA impedance is essentially zero and it delivers the characteristic ramp-on-a-step current waveform demanded by the output LC filter.

At the end of the $Q1$ on time and prior to the start of the next on time, the core must be restored to the B_1 level on the hysteresis loop, which yields the correct blocking time on the next switching cycle. If no current is delivered into MA in the reverse direction (into its no-dot end) immediately after the end of the ramp-on-a-step pulse, the MA core would return to $+B_s$ at 0 Oe.

The core is reset to the appropriate B_1 by a current-reset technique. In magnetic-amplifier practice, cores are reset to desired flux levels by either a voltage-reset or current-reset scheme. In voltage reset, the core is reset to any desired flux level by applying a unique amount of volt-seconds product to it (recall from Faraday's law, $dB = E\,dt/NA$). In this particular circuit, core reset is more simply done by current reset.

Current reset to the appropriate B_1 and line and load regulation is achieved with the voltage error amplifier, blocking diode $D3$, and the voltage-controlled current source $Q2$ of Fig. 10.3. The core is reset to that B_1 level which yields a blocking t_b and hence firing time t_f which gives the desired DC output voltage.

The resetting is done by pushing the appropriate DC current into the no-dot end of MA at the end of the $Q1$ on time.

When $Q1$ is on, $D3$ is reverse-biased and blocks $Q2$ from loading

down the secondary. When $Q1$ turns off, the top end of N_{ss} goes negative and current from the $Q2$ collector is driven into the no-dot end of MA. The amount of this current is controlled by the voltage error amplifier (error amplifier 1) such as to make the sampled fraction of the DC output voltage equal to the reference voltage.

If, for any reason, the DC output voltage goes up, t_f must go down, which means that t_b must go up. Thus, if the DC output voltage goes up, the error-amplifier output goes down and the $Q2$ collector current goes up, pushing more reset current into the no-dot end of MA. This pushes the initial flux level B_1 down further and, by Eq. 10.4, increases t_b, thus decreasing t_f and bringing the DC output voltage back down.

Similarly, of course, a decrease in V_{os} causes an increase in error-amplifier output voltage and a decrease in $Q2$ reset current. As B_1 flux level rises, t_b decreases, t_f increases, and V_{os} is brought back up.

All this, of course, occurs over a number of switching cycles and is accomplished in a time dependent on the error-amplifier bandwidth.

Stabilization of this negative-feedback loop will not be considered here. It is treated by C. Jamerson in Ref. 4 and C. Mullett in Ref. 2.

10.3.4 Slave output voltage shutdown with magnetic amplifiers

Heretofore, the magnetic amplifier was presented only as a means of voltage regulating the slave output voltage. That was done by controlling the flux level B_1 to which the core is reset at the end of the power transistor on time. The further down B_1 was pushed, the longer the blocking time t_b, the shorter the firing time t_f, and hence the lower the DC output voltage.

The magnetic amplifier can also be used to shut down the DC output voltage completely. This is done by pushing the initial flux level $+B_1$ down to $-B_s$. Blocking time, from Eq. 10.4, is $t_b = N_m A_e(2B_s)10^{-8}/V_{sp}$. The core area A_e and N_m will be chosen so that this blocking time is greater than the maximum $Q1$ on time.

Flux level B_1 can be brought down to $-B_s$ in a number of different ways by forcing the $Q2$ current to be sufficiently large. Thus the error-amplifier output can be overridden by forcing it down to a sufficiently low-impedance, low-voltage source. Or V_{ref} at the error-amplifier input terminal can be short-circuited to ground with an isolating resistor from the reference voltage source to the error-amplifier input terminal to avoid overloading the reference source.

It should be noted that if the magnetic-amplifier core is used only for voltage regulation (as it generally is), it moves around a minor hysteresis loop as in Fig. 10.2. The area of this loop may be a small

fraction of the total hysteresis loop area—depending on the minimum B_1 flux level, which is dependent on the maximum and minimum supply voltage and load current specifications.

Generally, for a voltage-regulation-only design, the minor loop area will be about one-fourth the total hysteresis loop area. Since core losses are proportional to the area of the hysteresis loop traversed, this ordinarily will result in quite low core dissipation and temperature rise.

However, if the design is to shut the slave voltage down completely to zero and the core flux excursion is over the full area of the major hysteresis loop from $+B_s$ to $-B_s$, higher core losses and temperature rise will result. Losses and core temperature rise should thus be calculated from the manufacturer's curves of core loss versus total flux excursion. This will be demonstrated below.

Note also that in Fig. 10.3, supply voltage for the error amplifier and $Q2$ is the slave output voltage source itself. If it is desired to shut the slave output down completely to zero, supply voltage for the error amplifier and $Q2$ will have to be taken from a source which is always present—possibly another slave output.

10.3.5 Square hysteresis loop core characteristics and sources

When interest in magnetic amplifiers resumed about 7 years ago, only one material having the required characteristics for an efficient high-frequency magnetic amplifier was available.

This material is an alloy of 79% nickel, 17% iron, and 4% molybdenum. It is available from various manufacturers under their particular brand names. Magnetics Inc. of Butler, Pennsylvania[3] is the foremost American supplier and has the largest range of available core sizes and iron areas. It calls its material Square Permalloy 80. Other manufacturers' brand names for similar material are 4-79 Molypermalloy, Square Mu 79, Square Permalloy, and Hy Ra 80.

Square Permalloy 80 is available from Magnetics Inc.[5] in tapes of various thicknesses wound on toroidal bobbins. Available tape thicknesses are 0.5, 1.0, 2.0, 4.0, 6.0, and 14.0 mils. Since the material is electrically conductive, eddy currents contribute a large fraction of the total losses, and at high frequencies, the thinner tapes must be used to keep losses down.

Generally, 1-mil tape thickness is used up to a switching frequency of 50 kHz, and ½ mil for frequencies of 50 to 100 kHz. Beyond 100 kHz, the newer "amorphous" core materials, to be discussed below, are used because of their lower losses.

The two characteristics required for an efficient, high-frequency

magnetic amplifier are a very square hysteresis loop and low losses at high peak flux density. The ½-mil tape is more expensive than 1-mil tape and should be used only if low loss is more important than low cost. If the design is only to provide voltage regulation by traversing minor loops rather than the full major loop, the higher loss 1-mil tape can be used even above 50 kHz.

The square hysteresis loop requirement is necessary for a very low impedance in the saturated state. If the hysteresis loop is not sufficiently square, its permeability (dB/dH in saturation, as shown in Fig. 10.2) is appreciable and its impedance at the top of the loop will be significantly greater than that of an air core coil of an equal number of turns.

Thus, in Fig. 10.3, the voltage at the output of the MA will be less than V_{sp} during the firing time and the drop across MA will be dependent on the secondary current. Further, if the loop is not square, the transition from high to low impedance may take a considerable time, prohibiting its use at high frequencies.

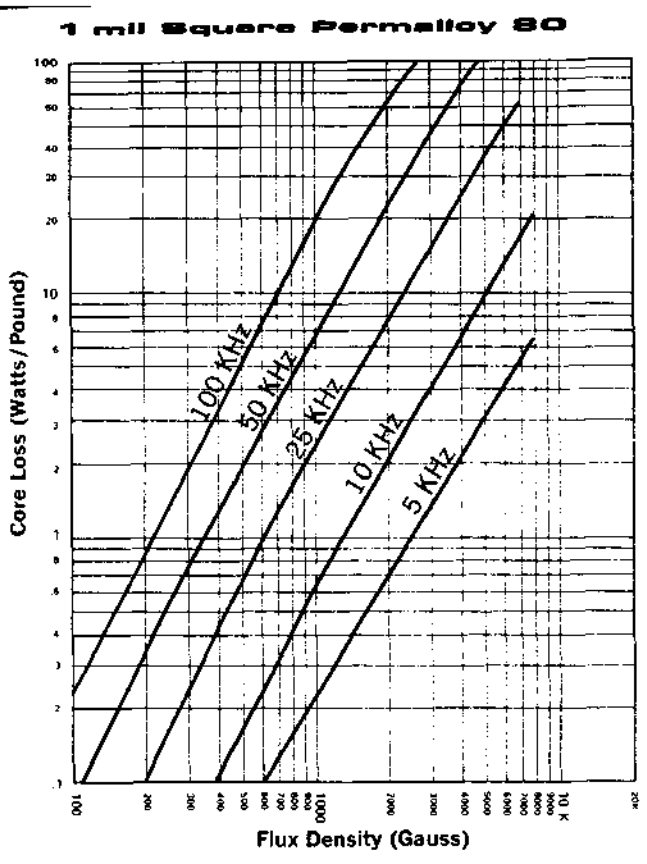
Core losses in watts per pound as a function of peak flux density and operating frequency are shown in Fig. 10.5*a* and 10.5*b* for 1- and ½-mil Square Permalloy tapes. Flux density on those curves are half the peak-to-peak excursion. The Magnetics Inc. catalog does not give the tape weight for each core, but this can be calculated from the core area, mean path length, and material density of 8.75 g/cm^3 which are given in the catalogs.

Temperature rise of the outer surface of the case enclosing the magnetic toroid can be estimated from its thermal resistance as read from Fig. 7.4*a* or 7.4*c* and the core losses. A reasonable estimate of the temperature differential between the magnetic core itself and the outer surface of the case is about 15°C.

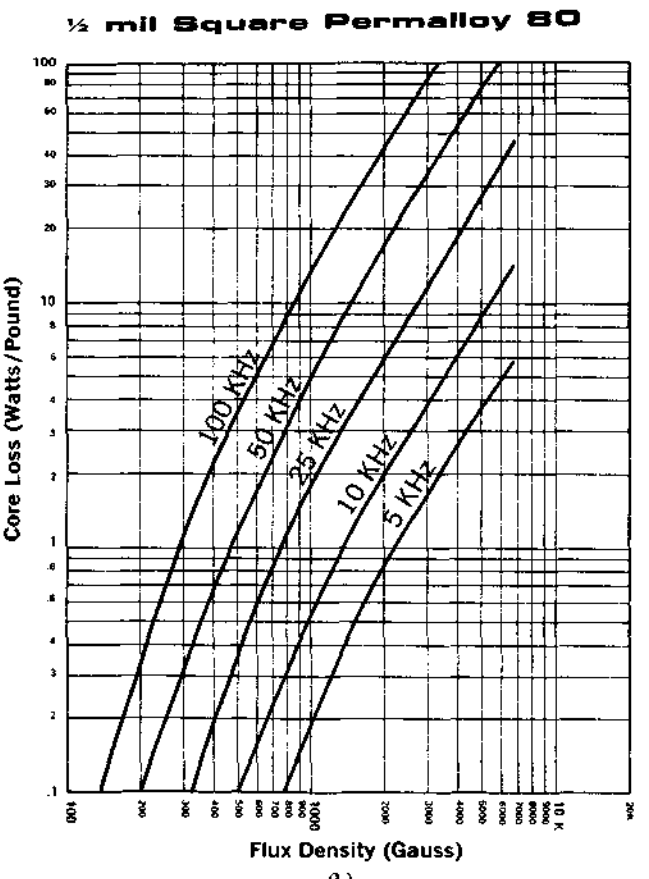
The core temperature rise can be quite high since the Curie temperature of Square Permalloy is 460°C. Thus the limiting factor for core losses is either the temperature rating of the wire or the specified magnetic-amplifier efficiency.

About 4 years after the renewed interest in magnetic amplifiers commenced, a new type of magnetic material was introduced. It is not crystalline in structure, but amorphous, and has lower core losses and a squarer hysteresis loop at higher frequencies than does Square Permalloy. Although ½-mil Permalloy can be used up to 100 kHz and possibly somewhat higher, beyond that frequency, this new material is preferable.

This amorphous material is available from two sources—Allied Signal in Parsippany, New Jersey,[6] and Toshiba Corporation[7] (the American sales agent is Mitsui in New York).



(a)



(b)

Figure 10.5 (*a*) Core losses, 1-mil Square Permalloy. (*Courtesy Magnetics Inc.*) (*b*) Core losses, ½-mil Square Permalloy. (*Courtesy Magnetics Inc.*) (c) Core losses, amorphous core material, Metglas 27144. (*Courtesy Magnetics Inc.*) (*d*) Core losses, Toshiba MA (watts/LB = 56.8 × watts/CC). (*Courtesy Toshiba Corp.*)

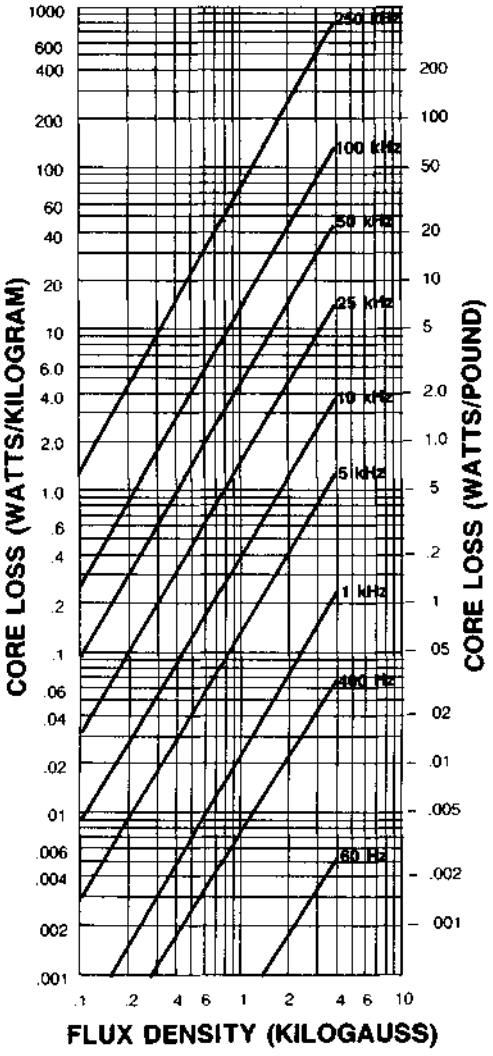


Figure 10.5c (*Continued*)

Allied calls its product Metglas 2714A. Toshiba has two amorphous core materials: MA and MB. Toshiba MB material is closely identical to Metglas 2714A in core losses, coercive force and squareness of its hysteresis loop. Toshiba MA material is midway between ½-mil Permalloy and MB material. Metglas 2714A material is also used by Magnetics Inc. for magnetic amplifier cores in its own line of standard-sized cores.

Curves of core loss versus peak flux density at various frequencies

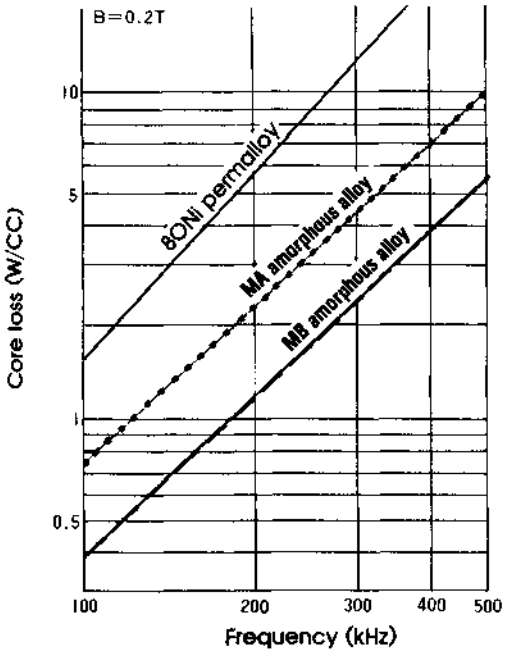


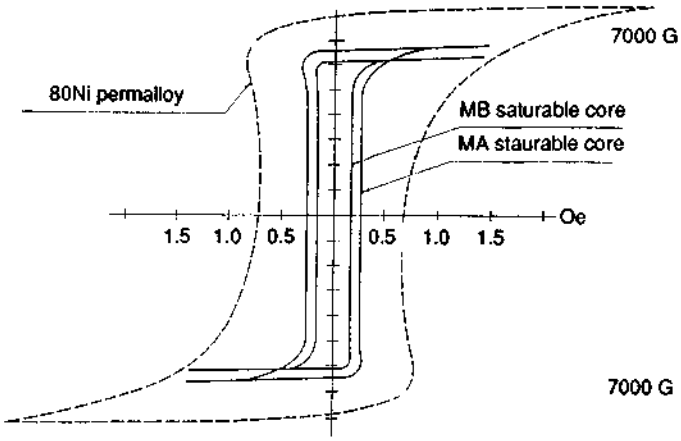
Figure 10.5d (*Continued*)

for Metglas 2714A are shown in Fig. 10.5*c*. Data comparing core loss versus frequency at a peak flux density of 2000 G for Toshiba MA, MB material and 1-mil Permalloy are given in Fig. 10.5*d*. Note that figure gives loss in watts per cubic centimeter. For MA, MB density of 8.0 g/cm^3, loss in watts per pound is 56.8 × loss in watts per cubic centimeter.

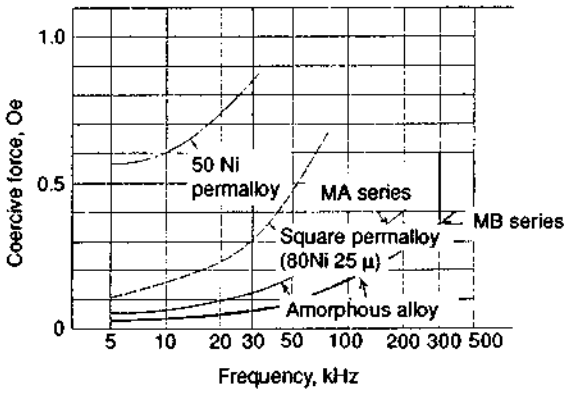
The *BH* loops at 100 kHz for Toshiba MA, MB material and ½-mil Permalloy are shown in Fig. 10.6*a*. Figure 10.6c compares 100-kHz *BH* loops for Metglas 2417A, ½- and 1-mil Permalloy.

The *BH* loops for increasing frequency have a characteristic appearance. As frequency increases, coercive force increases but saturation flux density remains fixed. This, of course, explains the increase in core loss with frequency as the loss is proportional to the area of the hysteresis loop. The increase in coercive force with frequency is shown in Fig. 10.6*b* for Toshiba MA, MB and Permalloy material.

Standard-sized MA and MB cores available from Toshiba are shown in Fig. 10.7*a* and 10.7*b*. Metglas 2714A cores available from Magnetics Inc.[8] are shown in Fig. 10.8*a* and from Allied Signal, in Fig. 10.8*b*.



(a)



(b)

Figure 10.6 (*a*) *BH* loops at 100 kHz. (*Courtesy Toshiba Corp.*) (*b*) Coercive force versus frequency. (*Courtesy Toshiba Corp.*) (*c*) *BH* loops at 100 kHz, 1-mil and ½-mil Permalloy, Metglas 2714A. (*Courtesy Magnetics Inc.*)

10.3.6 Core loss and temperature rise calculations

Toshiba provides curves useful in calculating core temperature rise for each of its cores. Figures 10.9, 10.10, and 10.11 show core loss versus total flux change in maxwells for its three largest MB cores.

Recall that flux change in maxwells equals flux density change in

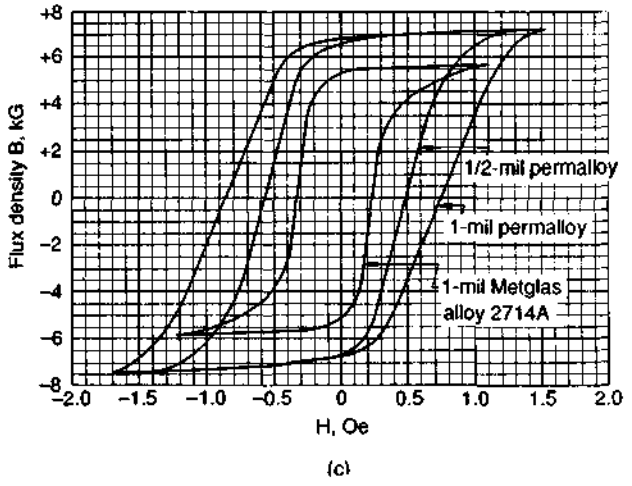


Figure 10.6 (*Continued*)

gauss multiplied by core area in square centimeters. Thus, dividing the maxwells shown in the curves by the core area gives the flux density change in gauss. The maximum maxwells shown on the curves then correspond to the total maximum flux density change from negative to positive saturation flux density ($-B_s$ to $+B_s$). Toshiba's MA, MB materials have B_s of 6500 and 6000 G, respectively.

Losses at the maximum maxwells shown in Figs. 10.9 to 10.11 then correspond to operation around the total major loop as when the magnetic amplifier is used to fully shut down the slave output voltage. Losses at the lesser maxwell level correspond to operation around a minor loop as in Fig. 10.2, where the magnetic amplifier is used only for voltage regulation.

Thus, whatever the total maxwell excursion is, Figs. 10.9 to 10.11 give core loss for each core at that flux excursion. From this core loss, Fig. 10.12 gives the core temperature rise. Actually, Fig. 10.12 is a measure of the thermal resistance of each core and is related to the radiating surface area of the core which can be calculated from its outer dimensions as given in Fig. 10.7*b*.

It is interesting to note that the temperature rise given by Toshiba in Fig. 10.12 coincides to within a very few degrees with the rise as calculated from Fig. 7.4*c* for any power dissipation and core area (calculated from Fig. 10.7*b*).

Type No.	Standard core dimensions (mm)			Finished dimensions (mm)			Effective core cross section (mm^2)	Mean flux path length (mm)	Total flux $2\Phi_1$ ($\times 10^{-8}$WB)	Insulating covers
	Outer dia.	Inner dia.	Height	Outer dia.	Inner dia.	Height				
MA 26×16×4.5W	26	16	4.5	29.5 max.	13.0 min.	8.0 max.	16.9	66.0	1800	Resin casing
" 22×14×4.5W	22	14	4.5	25.5 "	11.0 "	8.0 "	13.5	56.5	1440	"
" 18×12×4.5X	18	12	4.5	20.5 "	10.2 "	8.0 "	10.1	47.1	1080	Epoxy resin coating
" 14× 8×4.5X	14	8	4.5	16.3 "	6.3 "	7.5 "	10.1	34.6	1080	"
" 10× 6×4.5X	10	6	4.5	12.3 "	4.4 "	7.5 "	6.75	25.1	720	"
" 8× 6×4.5X	8	6	4.5	10.0 "	4.4 "	7.5 "	3.38	22.0	360	"
" 7× 6×4.5X	7	6	4.5	9.0 "	4.4 "	7.5 "	1.69	20.4	180	"

(a)

Type No.	Standard core dimensions (mm)			Finished dimensions (mm)			Effective core cross section (mm^2)	Mean flux path length (mm)	Total flux $2\Phi_1$ ($\times 10^{-8}$WB)	Insulating covers
	Outer dia.	Inner dia.	Height	Outer dia.	Inner dia.	Height				
MB 21×14×4.5	21	14	4.5	24.0 max.	12.0 min.	8.5 max.	11.8	55.0	1105	Epoxy resin coating
" 18×12×4.5	18	12	4.5	20.5 "	10.2 "	8.0 "	10.1	47.1	935	"
" 15×10×4.5	15	10	4.5	17.5 "	8.3 "	7.5 "	8.44	39.3	790	"
" 12× 8×4.5	12	8	4.5	14.3 "	6.3 "	7.5 "	6.75	31.4	629	"
" 10× 7×4.5	10	7	4.5	12.3 "	5.4 "	7.5 "	5.06	26.7	476	"
" 9× 7×4.5	9	7	4.5	11.0 "	5.4 "	7.5 "	3.38	25.1	315	"
" 8× 7×4.5	8	7	4.5	10.0 "	5.4 "	7.5 "	1.69	23.6	158	"
" 15×10×3W	15	10	3.0	17.0 "	8.0 "	5.0 "	5.63	39.3	526	Resin casing
" 12× 8×3W	12	8	3.0	14.0 "	6.0 "	5.0 "	4.5	31.4	420	"

(b)

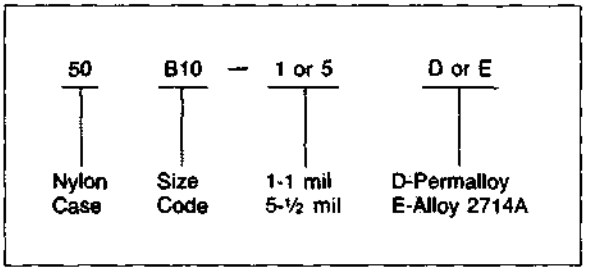
Figure 10.7 (*a*) **Available MA amorphous cores for magnetic amplifiers.** (*Courtesy Toshiba Corp.*) (*b*) **Available MB amorphous cores for magnetic amplifiers.** (*Courtesy Toshiba Corp.*)

Part Number*	DIMENSIONS						Core loss (w) @ 50KHz, 2000 gauss (Max.)	ml cm	Ac cm^2	Wa cr m	Core wt. grams	Wa Ac cr m cm^2 ($\times 10^{-4}$)
	I.D. (in.)		O.D. (in.)		Ht. (in.)							
	core	case (Min.)	core	case (Max.)	core	case (Max.)						
50B10-5D	.650	.580	.900	.970	.125	.200	.118	6.18	.051	348,000	2.7	0177
50B10-1D	.650	.580	.900	.970	.125	200	.22	6.18	.076	348,000	4.0	.0264
50B10-1E	650	.580	.900	.970	.125	.200	.092	6.18	.076	348,000	3.5	.0264
50B11-5D	.500	.430	.625	.695	.125	.200	.044	4.49	.025	194,000	1.0	.0048
50B11-1D	.500	.430	.625	.695	.125	.200	.083	4.49	.038	194,000	1.5	.0074
50B11-1E	.500	.430	.625	.695	.125	.200	.034	4.49	.038	194,000	1.3	.0074
50B12-5D	.375	.305	.500	570	.125	.200	.035	3.49	.025	99,000	8	.0025
50B12-1D	.375	.305	.500	570	.125	.200	.066	3.49	.038	99,000	1.2	.0038
50B12-1E	.375	.305	.500	570	.125	.200	.027	3.49	.038	99,000	1.04	.0038
50B45-5D	.500	.430	.750	.820	.250	.325	.194	4.99	.101	194,000	4.4	.0143
50B45-1D	.500	.430	.750	.820	.250	.325	.363	4.99	.151	194,000	6.6	.0214
50B45-1E	.500	.430	.750	.820	.250	.325	.149	4.99	.151	194,000	5.7	.0214
50B66-5D	.500	.430	.750	820	.125	.200	.097	4.99	.050	194,000	2.2	.0071
50B66-1D	.500	.430	.750	.820	.125	.200	.182	4.99	.076	194,000	3.3	0108
50B66-1E	.500	430	.750	.820	.125	.200	.075	4.99	.076	194,000	2 9	.0108

*For other sizes, refer to factory.

Figure 10.8 Available Square Permalloy 80 and Metglas 2714A cores for magnetic amplifiers. (*Courtesy Magnetics Inc.*) (*b*) Standard Metglas 2714A amorphous magnetic-amplifier cores. (*Allied Signal.*)

PART NUMBER CODE



MATERIAL CHARACTERISTICS

	Alloy 2714A	½ mil Permalloy	1 mil Permalloy
Bm (gauss min.)	5000	7000	7000
Br/Bm (min.)**	.9	.83	.80
H_1 (oersted max.)**	.025	.045	.040
Core loss (w/lb. max. @ 50 kHz, 2000 gauss)	12	20	25

**Measured @ 400 Hz, CCFR Test

Figure 10.8 (*Continued*)

CORE NUMBER	DIMENSIONS (mm)*	OD	ID	HT	(cm)	A (cm²)	Mass (g)	Flux Capacity 2σ (μ Wb)	W. (cm³)	W A (cm³)
MP1303	CORE CASE	12.8 14.6	9.5 7.9	3.2 5.1	3.50	0.041	1.1	4.7	0.49	0.021
MP1603	CORE CASE	15.9 17.8	12.7 11.1	3.2 5.1	4.50	0.041	1.4	4.7	0.96	0.039
MP1903	CORE CASE	19.2 21.0	12.7 11.1	3.2 5.1	5.00	0.082	3.1	9.3	0.96	0.079
MP2303	CORE CASE	22.9 25.0	16.5 14.6	3.2 5.1	6.19	0.081	3.8	9.2	1.68	0.14
MP1305	CORE	12.5	9.5	4.8	3.46	0.057	1.5	6.5	0.49	0.028
MP1505	CORE	15.1	9.5	4.8	3.87	0.11	3.1	12	0.49	0.049
MP1805	CORE CASE	18.4 20.8	12.7 10.8	4.8 6.7	4.88	0.11	4.0	12	0.92	0.10
MP1906	CORE CASE	19.1 21.3	12.7 10.7	6.4 8.4	4.99	0.16	6.1	18	0.90	0.14
MP3506	CORE CASE	35.2 37.3	25.4 23.4	6.4 8.4	9.52	0.241	17.4	27.5	4.29	1.04
MP2510	CORE CASE	25.6 27.8	19.1 17.0	9.5 11.8	7.01	0.241	12.8	27.5	2.28	0.552

PART NUMBER CODE:

NOTE: 1 circular mil = 5.067 × 10 cm

MP	**13 03**	**P or E**	**4A**	**S**
Metglas	OD HT	P - Plastic Box	Alloy	Square
Products		E - Encapsulated	2714A	Loop

(b)

Figure 10.8 *(Continued)*

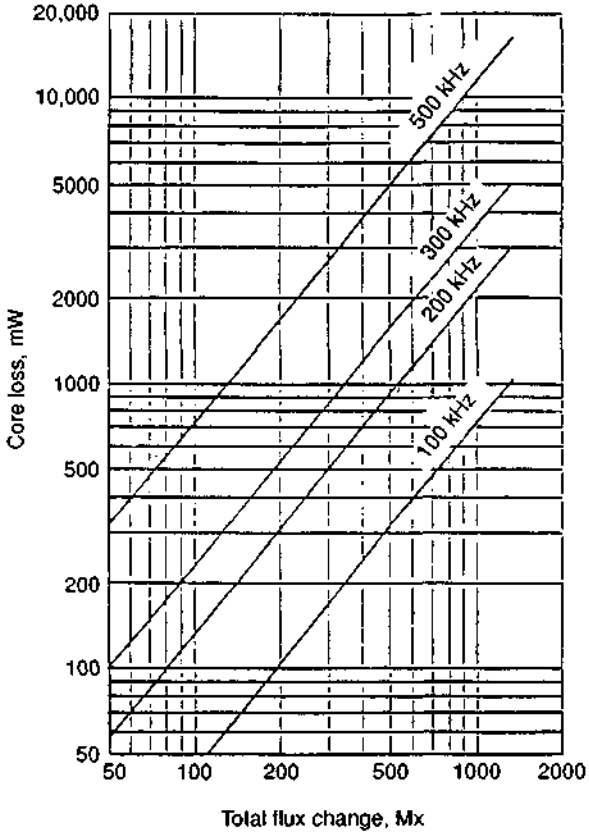


Figure 10.9 Core loss versus total flux change. Toshiba MB 21 × 14 × 4.5 core. Core area = 0.118 cm^2; $\Delta B = \Delta\phi$(maxwells)/0.118. (*Courtesy Toshiba Corp.*)

10.3.7 Design example—magnetic-amplifier postregulator

Design a magnetic-amplifier postregulator for the output of the forward converter shown in Fig. 10.13*a*. Specifications are

Forward converter switching frequency	100 kHz
Slave output voltage	15 V
Slave output current	10 A

The main output voltage is $V_{om} = V_{dc}(N_{sm}/N_p)(t_{on}/T)$. The main feedback loop, in keeping V_{om} constant, then must keep the product $V_{dc}t_{on}$ constant. Thus t_{on} is a maximum when V_{dc} is a minimum.

In the usual case, the number of turns on the $T1$ reset winding N_r is set equal to the turns on the power winding N_p. This forces the voltage

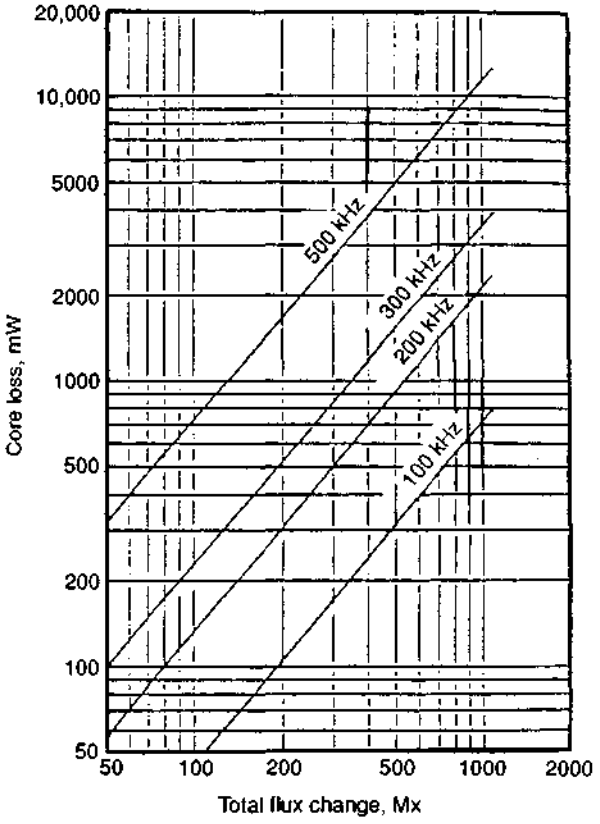


Figure 10.10 Core loss versus total flux change. Toshiba MB 18 × 12 × 4.5 core. Core area = 0.101 cm^2; $\Delta B = \Delta\phi$(maxwells)/0.101. (*Courtesy Toshiba Corp.*)

across N_p when $Q1$ is off to be equal (and opposite, of course) to the voltage across it when $Q1$ is on.

Within one cycle, the volt-second product across N_p when $Q1$ is on must be equal and opposite to the volt-second product across it when $Q2$ is off. Otherwise, from Faraday's law, the core would move in gauss in one direction on the hysteresis loop and at the start of the next cycle, would not have been returned to its starting point on the loop. After a number of such cycles, the core would drift up the hysteresis loop, saturate, and—being unable to support voltage—destroy the power transistor the next time it is turned on.

Thus the absolute maximum $Q1$ on time at minimum DC input voltage is $0.5T$ or 5 μs so that $+V_{dc}t_{on}$ can equal $-V_{dc}t_{off}$.

To ensure that the core can always be reset during transient power-line dips below its specified minimum, in the above expression for

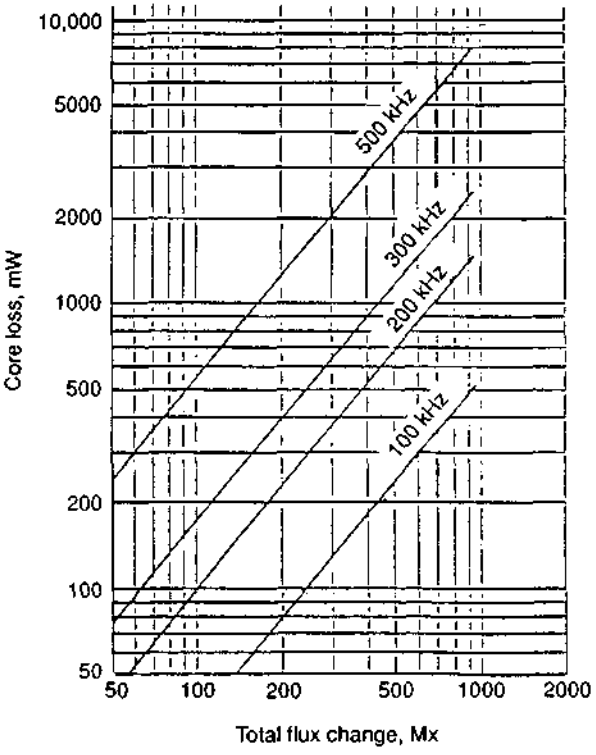


Figure 10.11 Core loss versus total flux change. Toshiba MB 15 × 10 × 4.5 core. Core area = 0.0843 cm^2; ΔB = $\Delta\phi$(maxwells)/0.0843. (*Courtesy Toshiba Corp.*)

V_{om}, N_{sm} will be chosen so that V_{om} is obtained for a maximum on time of $0.4T$ or 4 μs at the specified minimum V_{dc}. This yields a guard band of $0.1T$ or 1 μs to allow for line input dips.

The slave peak voltage V_{sp} (at the input end of MA) is then high for 4 μs at minimum DC input voltage. The MA will block this voltage for a time t_b (Fig. 10.13*c*), leaving a high voltage V_{sp1} for a time t_f at the no-dot end of the MA after it has saturated or fired.

If the MA has zero impedance when it has saturated, then $V_{sp1} = V_{sp}$. Then, assuming a 1-V drop across rectifier diode $D1$, the peak voltage at the front end of L_s when MA has fired is $(V_{sp} - 1)$. Now of the 4-μs duration of V_{sp} at low V_{dc} input, arbitrarily set $t_f = 3$ μs and $t_b = 1$ μs. This permits t_1 (Fig. 10.13*b*) to move either left or right, increasing or decreasing t_f to regulate against load changes.

Then for a slave output voltage of the specified 15 V

$$15 = (V_{sp} - 1)\left(\frac{t_f}{T}\right) = (V_{sp} - 1)\left(\frac{3}{10}\right) \qquad \text{or} \qquad V_{sp} = 51 \text{ V}$$

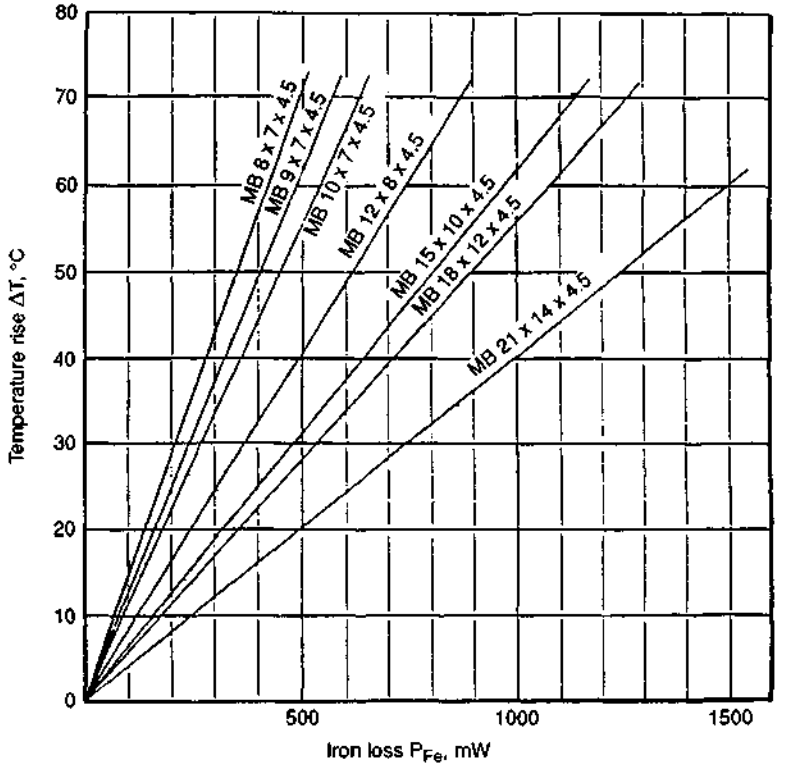


Figure 10.12 Toshiba amorphous MB cores. Temperature rise versus core losses. (*Courtesy Toshiba Corp.*)

Now the number of turns on MA and its iron area must be chosen to block 51 V. But they will be chosen to block not for the 1 μs of minimum blocking time, but for the full 4 μs corresponding to the maximum duration of V_{sp} on the assumption that the MA may be used to force the slave output voltage completely down to zero.

Assume that for a 100-kHz magnetic amplifier, an amorphous core such as the Toshiba MB (Fig. 10.7*b*) will be used. Arbitrarily select the one with the largest iron area as it will require the least turns to block the 51 V for 4 μs for a given flux change. This minimizes the residual air core inductance of the MA in its saturated state. Thus the MB 21 × 14 × 4.5, of iron area 0.118 cm^2, is chosen.

From Faraday's law, to minimize the number of turns, the flux change should be maximized or the core should traverse the full *BH* loop from $-B_s$ to $+B_s$. From Fig. 10.6*a*, B_s for the Toshiba MB material is 6000 G. Then from Faraday's law, to block 51 V for 4 μs with a core area of 0.118 cm^2:

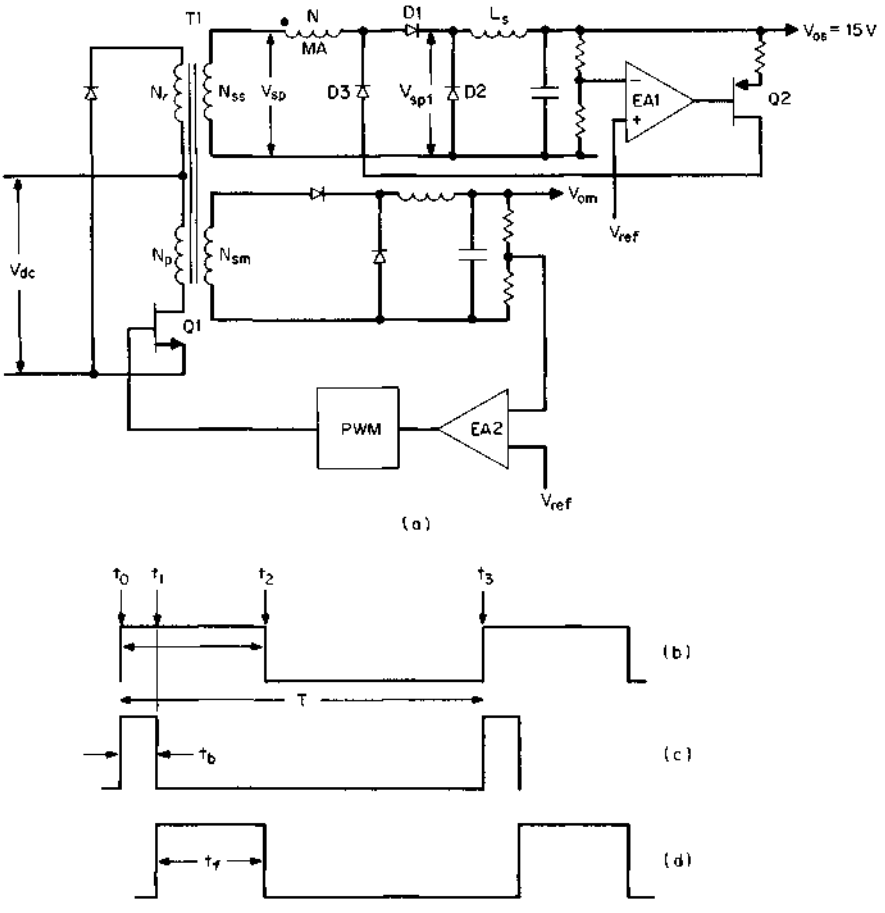


Figure 10.13 Design example of magnetic-amplifier postregulator. Magnetic amplifier blocks V_{so} for a time t_b and is a zero-impedance short circuit for a time t_r. Then $V_{os} = V_{so}(t_f/T)$. Time t_f is controlled in a negative-feedback loop by the current that $Q2$ forces into MA via D_m. That current is controlled by error amplifier 1.

$$E = 51 = NA_e \frac{dB}{dt} 10^{-8}$$

$$= 51 = N(0.118) \frac{2B_s}{4 \times 10^{-6}} 10^{-8}$$

or $N = 14$ turns.

For DC output current of 10 A, MA will be carrying the 10 A for a maximum of only 3 μs (maximum t_f) at low DC input. Maximum rms current is then $10\sqrt{3/10} = 5.48$ A. At 500 circular mils per rms ampere, a wire area of 2739 circular mils is required.

Two No. 19 wires in parallel provide 2 × 1290 or 2580 circular mils,

which is close enough. The inner periphery of the Toshiba MB21 × 14 × 4.5 core is $\pi \times 0.55 = 1.73$ in. For the 0.0391-in diameter of No. 19 wire, the inner periphery can hold 1.73/0.0391 or 44 turns on a single layer. The 14 turns of two paralleled No. 19 wires can thus easily be accommodated in a single layer on the inner periphery.

If the core is operated in the shutdown mode, the full major loop is traversed each cycle. Figure 10.9 shows that for the MB 21 × 14 × 4.5, this corresponds to a total flux change of 1400 Mx or 12000 G ($-B_s$ to $+B_s$) and core loss of 1 W, and Fig. 10.12 shows that the core temperature rise is only 40°C.

10.3.8 Magnetic-amplifier gain

When the MA has fired, it has close to zero impedance and the DC current through it is determined only by the DC output impedance and the slave output voltage. That is simply the specified DC output current. But to bring the MA to its fired state, a current equal to twice the coercive current I_c is required to force the core from the left to the right side of the hysteresis loop (Fig. 10.2). That current comes from the transformer secondary V_{sp}. Similarly, when the core is reset to the left side of the hysteresis loop (Fig. 10.2), a current equal to the coercive current must be supplied from Q2 via D3.

Magnetic-amplifier gain from $Q2$ to the output is then I_o/I_c. From Ampere's law, the coercive force $H_c = 0.4\pi N_m I_c/L_p$, where N_m is the number of turns, I_c is the coercive current, and L_p is the mean path length in centimeters.

For the MB core at 100 kHz, the coercive force is 0.18 Oe (Fig. 10.6*b*). Mean path length L_p of the MB 21 × 14 × 4.5 core is 5.5 cm, from Fig. 10.7*b*. Then for N_m of 14 turns, the coercive current is

$$I_c = \frac{H_c L_p}{0.4\pi N_m}$$

$$= \frac{0.18 \times 5.5}{0.4\pi \times 14}$$

$$= 56.3 \text{ mA}$$

Thus a current of 56.3 mA from $Q2$ can control the 10-A output—or, stated another way, the magnetic-amplifier gain is 10/0.056 = 178.

It is of interest to realize the physical significance of this and to appreciate that a magnetic amplifier is very different from a usual saturable reactor. In the magnetic amplifier, when the control current (from $Q2$) flows, no load current flows as $D1$ is reverse-biased. Thus the control current does not have to "buck out" the load current.

In a saturable reactor (a variable-inductance reactor controlled by current in a control winding) in series with a load, however, current

flows to the load at the same time the control current flows. Thus the control winding ampere turns must buck out the load ampere turns and gain is low.

10.3.9 Magnetic amplifiers for a push-pull output

For a full-wave output (push-pull or half-bridge topology), the circuit of Fig. 10.14 is often presented. But it has serious problems in that the dead time between transitions, the magamps, carry primary magnetizing current (Fig. 2.6). This is fully discussed in Ref. 15.

10.4 Magnetic-Amplifier Pulse-Width Modulator and Error Amplifier

Thus far, in this chapter, magnetic amplifiers as postregulators only have been considered. In this section, an interesting example of the use of a magnetic amplifier used simultaneously as a pulse-width modulator and error amplifier is described.[9]

It may be puzzling why, with the current enormous proliferation of inexpensive semiconductor pulse-width modulating chips with their built-in width modulators and error amplifiers, there is interest in magnetic elements to perform these functions.

The circuit to be described below does have advantages in specialized applications where a semiconductor integrated circuit cannot be used for some reason. The circuit consists of only square hysteresis loop core material and wire which is far more reliable than that in an integrated circuit.

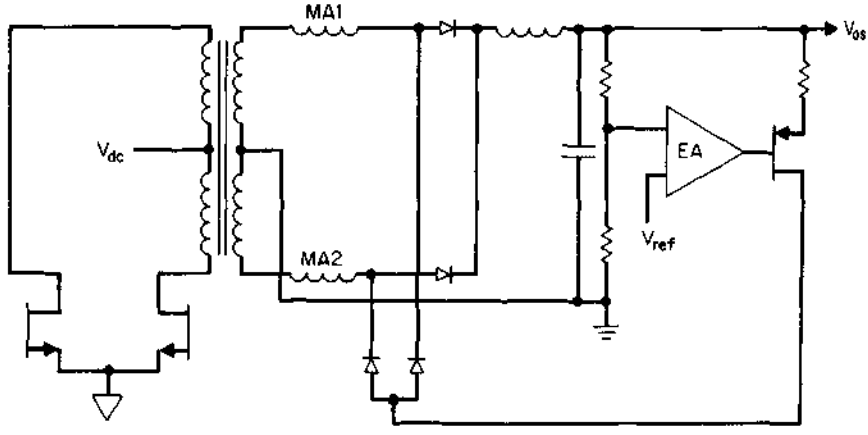


Figure 10.14 Two magnetic-amplifier cores driven from the same EA and PNP current source are required for a push-pull output.

There are some environmental conditions—such as excessive temperature—where the magnetic amplifier can survive more easily than an integrated circuit. Since discrete transistors which can survive high temperatures can be found easily, a circuit consisting of discrete transistors and a magnetic amplifier PWM–error amplifier is a more survivable circuit than one with discrete transistors and an integrated-circuit PWM chip.

Finally, the circuit described below offers a simple solution to an omnipresent problem in switching power supplies. That problem is how to sense a voltage on output ground and deliver to the power transistor on input ground, the appropriate width-modulated pulse, without requiring a housekeeping power supply to power the error amplifier on output ground.

The circuit works as follows.

10.4.1 Circuit details, magnetic amplifier pulse-width modulator–error amplifier[9]

The circuit is shown in Fig. 10.15. It was originally devised by Dulskis and Estey (cited in Ref. 9). It is a conventional push-pull topology using Darlington power transistors. A 40-kHz, ±8-V, 50/50 duty cycle square wave is applied at points *AB* via transformer *T*1.

The heart of the design is the magnetic amplifier *M*1. It consists of two square hysteresis loop cores sitting on top of one another. There are two equal turn gate windings N_{g1}, N_{g2} and a control winding N_c. Gate winding N_{g1} links core *A* only, N_{g2} links core *B* only, and N_c links both cores.

On alternate half cycles, a +8-V, 12.5-μs appears at points *A*, *B*. Consider the half period when, say, *A* is positive. That 8 V is divided across *D*1, the base-emitters of *Q*3, *Q*1, and *R*5. Resistor *R*5 is a current limiter.

At the start of the half period, magnetic-amplifier core *M*1*A* is on the steep part of its hysteresis loop (B_0 of Fig. 10.16*a*) and has a sufficiently high impedance so that it does not short-circuit out the voltage from the *Q*3 base to ground. The Darlington, *Q*3, *Q*1, is energized and a voltage of roughly $V_{dc} - 1$ is applied to its half primary of *T*2.

The voltage across *M*1*A* is the sum of the base-emitter drops of *Q*1 and *Q*3 (~1.6 V). This voltage drives *M*1*A* upward toward saturation along the minor hysteresis loop *B*0–*B*1–*B*2–*B*3–B_s. At B_s, *M*1*A* saturates, *Q*3 loses its base drive, and Darlington *Q*3, *Q*1 turn off. The *Q*3, *Q*1 on time is given by Faraday's law as

$$t_{on(Q1,\,Q3)} = \frac{N_g A_c (B_s - B_0)}{1.6}\,10^{-8}$$

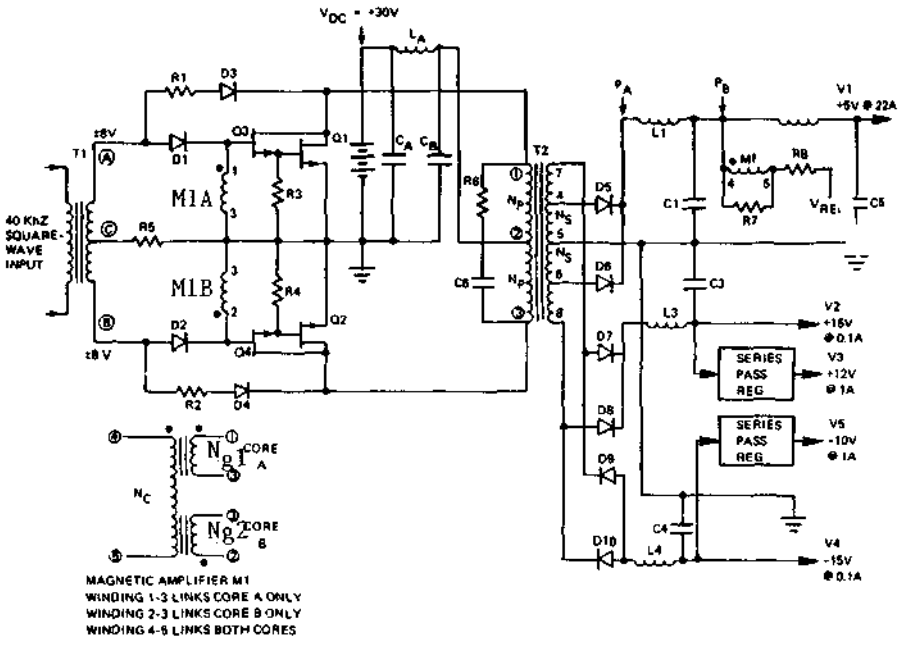


Figure 10.15 Magnetic-amplifier-controlled switching regulator. The magnetic amplifier serves as both a pulse-width modulator and an error amplifier. DC current through control winding N_c determines the initial flux bias in gate windings and hence their time to saturate, which fixes the Darlington on times. (*From R. Dulskis, J. Estey, and A. Pressman, "A Magnetic Amplifier Controlled 40 kHz Switching Regulator," Wescon Proceedings, San Francisco, 1977.*)

where N_g = number of turns of $M1A$
A_e = core iron area
B_0 = initial starting point on the hysteresis loop
1.6 = voltage across $M1A$ driving it toward saturation

It is seen that the further down B_0 is from B_s, the longer the on time. On-time control will be accomplished by controlling the level B_0 to which the core is reset at the start of the positive half cycle.

Now, as $M1A$ is being push up toward saturation with a voltage of 1.6 V across its N_g turns, it induces a voltage of $(N_c/N_g)1.6$ across N_c. That couples a voltage back down to 1.6 V across $M1B$. But because of the winding polarities, that voltage is in the direction to push $M1B$ down to B_0 if in the previous half period it had been pushed up to $+B_s$ by point B having been positive.

Thus in one half period, as—say—$M1A$ goes from B_0 to B_s, its Darlington is on and $M1B$ is driven down from B_s to B_0. But the $M1B$ Darlington is off as its input diode $D2$ is reverse-biased. During the

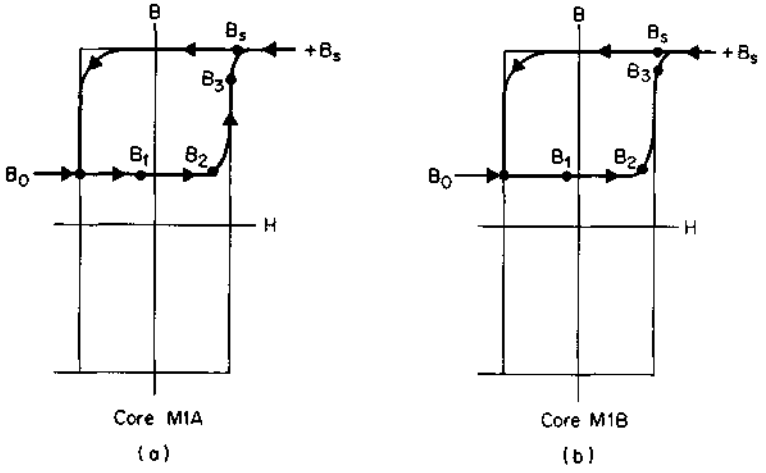


Figure 10.16 The *BH* loop operating locus of magnetic-amplifier cores (Fig. 10.15). As core *M1A* moves from B_0 to $+B_s$, core *M1B* is pushed down from $+B_s$ to B_0. Time to move from B_0 to B_s is the on time of the power transistors. On time is determined by how far down from B_s the initial starting ping B_0 has been pushed. This is determined by the DC current in control winding N_c, which is proportional to the output voltage (Fig. 10.15).

next half period the sequence reverses; *M1B* is driven up from B_0 to $+B_s$, its Darlington is on, and *M1A* is forced down from B_s to B_0. Power Darlington on time is fixed by B_0, the initial starting point on the *BH* loop, and that is determined by the DC current in the control winding N_c.

Because of the winding polarities, there is no net fundamental AC voltage across N_c. Now N_c is bridged between the DC output voltage and a reference voltage as in Fig. 10.15. The polarity of N_c is such that if the DC output voltage goes up, the gate windings are biased further up toward B_s, time to saturate the core decreases and the Darlington on time decreases, bringing the DC output voltage back down.

Thus the control winding serves as the error voltage sensor. It is referenced to output ground and DC current in it serves to control the Darlington on time on input ground by controlling the gate winding's initial (starting) flux density B_0. It is this last feature which is perhaps the most useful characteristic of the circuit. Resistors *R7*, *R8* serve to set the control winding error-amplifier gain.

In the circuit of Fig. 10.15, the MA cores were of ¼-mil Square Permalloy wound on bobbins of 0.290 in OD, 0.16 in ID, and 0.175 in height. The gate windings—winding 1–3, which linked core A only, and winding 2–3, which linked core B only, consisted of 40 turns of No. 35 wire. The control winding 4–5, which linked both cores, consisted of 250 turns of No. 37 wire.

References

1. C. Mullett, "Design and Analysis Of High Frequency Saturable Core Magnetic Regulators," *Proceedings Powercon 10,* 1983.
2. C. Mullett, "Design of High Frequency Saturable Reactor Output Regulators," *High Frequency Power Conversion Conference Proceedings,* 1986.
3. Toshiba Corp. Bulletin, *Toshiba Amorphous Magnetics Parts,* Mitsui & Co., New York.
4. C. Jamerson, "Calculation of Magnetic Amplifier Post Regulator Voltage Control Loop Parameters," *High Frequency Power Conversion Conference Proceedings,* 1987.
5. Magnetics, Inc. Bulletin TWC 300S, Butler, PA.
6. Allied Signal Technical Bulletin, *Metglas Amorphous Alloy Cores,* Parsippany, NJ.
7. Toshiba Corp. Bulletin, *Toshiba Amorphous Saturable Cores,* Mitsui & Co., New York.
8. Magnetics Inc. Bulletin, *New Magnetic Amplifier Cores and Materials,* Butler, PA.
9. R. Dulskis, J. Estey, and A. Pressman, "A Magnetic Amplifier Controlled 40 kHz Switching Regulator," *Wescon Proceedings,* San Francisco, CA, 1977.
10. R. Taylor, "Optimizing High Frequency Control Magamp Design," *Proceedings Powercon 10,* 1983.
11. R. Hiramatsu, K. Harada, and T. Ninomiya, "Switchmode Converter Using High Frequency Magnetic Amplifier," *International Telecommunications Energy Conference,* 1979.
12. S. Takeda and K. Hasegawa, "Designing Improved Saturable Reactor Regulators with Amorphous Magnetic Materials," *Proceedings Powercon 9,* 1982.
13. R. Hiramatsu and C. Mullett, "Using Saturable Reactor Control in 500 kHz Converter Design," *Proceedings Powercon 10,* 1983.
14. A. Pressman, "Amorphous Magnetic Parts for High Frequency Power Supplies," *APEC Conference 1990* (Toshiba Seminar), Mitsui & Co., New York.
15. C. Jamerson and D. Chen, "Magamp Post Regulators For Symmetrical Topologies with Emphasis on Half Bridge Configuration," *Applied Power Electronics Conference (APEC),* 1991.

Chapter

11

Turnon, Turnoff Switching Losses and Snubbers

11.1 Introduction

In topologies which have a transformer in series with their power transistor (all except the buck regulator), switching losses due to the overlap of falling current and rising voltage across the power transistor during the turnoff interval account for most of the losses. The integral $\int I(t)V(t)dt$ over the turnoff interval, which may last anywhere from 0.2 to 2 μs (for bipolar transistors), generally has a very high peak value.

Even averaged by the turnoff duty cycle, it can be two to four times as large as dissipation during the transistor conduction time. This spike of AC dissipation coming once per period is, of course, a larger fraction of the total transistor dissipation at higher frequencies. It is one of the prime limitations to the use of bipolar transistors above 100 kHz.

Circuits to minimize these overlap losses at turnoff are called turnoff snubbers and are the prime subject of discussion in this chapter.

Switching losses at the instant of turnon are considerably smaller for topologies with a power transformer because of transformer leakage inductance. At the instant of turnon, the large instantaneous impedance of the leakage inductance forces the voltage across the transistor to drop rapidly to zero while the leakage inductance slows up current rise time.

Thus, throughout most of the current rise time, voltage across the transistor is close to zero and switching losses due to voltage-current overlap are negligible.

In the buck regulator (Fig 1.4), however, there are large voltage-current overlap losses in the transistor at both turnon and turnoff. In the buck, the power transistor turns on into the negligibly low imped-

ance of the conducting free-wheeling diode $D1$ and the overlap of rising current and falling voltage amounts to a large spike of instantaneous power dissipation. Such turnon losses in the buck transistor and circuits to minimize them (turnon snubbers) have been discussed in Secs. 1.3.4 and 5.6.6.3 to 5.6.6.8.

Switching losses in the buck transistor at turnoff are minimized by the same turnoff snubber circuit usually used in transformer-loaded topologies.

With MOSFETs, turnoff switching losses are considerably less than those with bipolar transistors. Current fall time with a MOSFET is so rapid that current in it will have fallen almost to zero by the time the voltage across it has risen significantly.

Thus, although turnoff snubbers are used with MOSFETs, their prime function is not to reduce overlap dissipation, which is already low. Rather, the function of the MOSFET turnoff snubber is to reduce the amplitude of the leakage inductance voltage spike. Since leakage inductance voltage spikes are proportional to dI/dt in the transistor, a MOSFET with much faster current turnoff time than a bipolar will have a larger voltage leakage spike. Thus, although a MOSFET also requires a turnoff snubber, it is a more benign one and has less drawbacks than does one for a bipolar transistor.

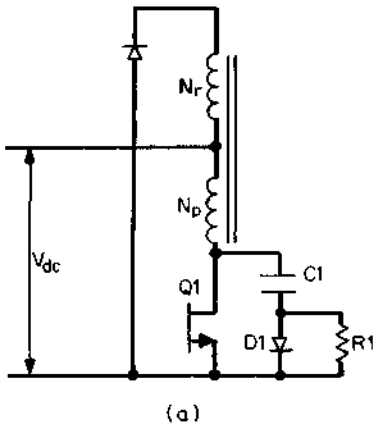
With MOSFETs, there is considerable dissipation at turnon; this is due not to overlap of simultaneous high voltage and current, but to the relatively high MOSFET output capacity C_o. This capacity is charged (usually to twice the supply voltage) and stores an energy $\frac{1}{2}C_o(2V_{dc})^2$ at turnoff. At the subsequent turnon, this energy is dissipated in the MOSFET and results in a dissipation, averaged over a period T, of $\frac{1}{2}C_o(2V_{dc})^2/T$.

Unfortunately, this dissipation is increased by the small, benign snubber the MOSFET requires for minimizing the leakage inductance voltage spike as it adds capacity at the transistor output.

All this will become more obvious after considering the design of the usual RCD (resister, capacitor, diode) turnoff snubber.

11.2 Transistor Turnoff Losses without a Snubber

Consider the elements of a forward converter shown in Fig. 11.1*a*. The turnoff snubber consists of $R1$, $C1$, $D1$. Assume that its output power is 150 W and it operates from a nominal DC supply voltage of 160 V (maximum 184, minimum 136 V). These are roughly the rectified DC supply voltages for an off-line power supply operating from a nominal 115-V-AC line.



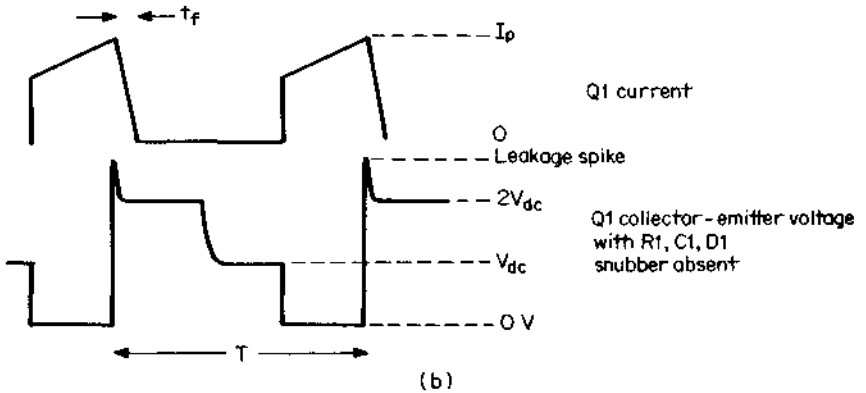


Figure 11.1 (*a*) A forward converter with snubber $R1$, $C1$, $D1$. When $Q1$ turns off, collector voltage starts to rise. But immediately $D1$ latches in and $C1$ slows up voltage rise time to minimize the overlap of rising voltage and falling current and thus reduce $Q1$ switching loss. The next time $Q1$ turns off, $C1$ must have lost the $2V_{dc}$ charge it picked up the previous turnoff. It looses that charge in the previous $Q1$ on time. Capacitor $C1$ discharges through $Q1$ and $R1$. (*b*) With snubber absent, collector voltage rises almost instantaneously; $Q1$ dissipation is $(2V_{dc}I_p/2)(t_f/T)$.

From Eq. 2.28, $Q1$ peak current is $I_p = 3.13P_o/V_{dc} = 3.13 \times 150/136 = 3.45$ A. Assume that $Q1$ is a very fast bipolar transistor such as the third-generation Motorola 2N6836. It has a V_{ceo} rating of 450 V (850 V V_{cev}) and a 15-A collector current rating. Its data sheet shows that collector current fall time from 3.5 A with a 5-V reverse bias is 0.15 μs. This presumably is nominal—assume that the worst case is twice that or 0.3 μs.

Assume that the snubber network $R1$, $C1$, $D1$ is absent. At the instant of turnoff, the collector rises as in any forward converter with equal power and reset winding turns to twice V_{dc}. The force which drives the collector to $2V_{dc}$ is the current that had been stored in the magnetizing inductance and leakage inductances in series. As with any inductor which is carrying a current, when a series switch opens to open that current, the polarity across those inductors reverses and drives the collector toward $2V_{dc}$.

Since collector output capacity is low, assume that this voltage rise is instantaneous. Assuming a maximum V_{dc} of 184 V, the voltage rises instantaneously to 368 V and the current falls linearly from 3.45 A in 0.3 μs as seen in Fig. 11.1*b*.

This amounts to a dissipation, averaged over the 0.3 μs, of 368 × 3.45/2 = 635 W. Further, assuming a 100-kHz switching frequency, power dissipation, averaged over one period, is 635 × 0.3/10 = 19 W. This is excessive and would require, in most cases, an unacceptably large heat sink to keep the transistor junction temperature to a reasonable value.

Note also that this is an optimistic dissipation calculation. As discussed in Sec. 1.3.4, the estimated dissipation depends on the scenario assumed in the timing of current turnoff. Current generally hangs on at its peak for a relatively incalculable time before starting to fall. Thus the 19-W dissipation calculated above may very well be about 50 percent greater in actuality.

The $R1$, $C1$, $D1$ snubber of Fig. 11.1*a* reduces transistor dissipation by slowing up the collector voltage rise time so that the current fall-time waveform intersects the voltage rise-time waveform as low down as feasible on the rising voltage waveform. How it works and how the component magnitudes are calculated can be seen as follows.

11.3 *RCD* Turnoff Snubber Operation

In Fig. 11.1*a*, when the $Q1$ base receives its turnoff command, the transformer leakage inductance attempts to maintain the peak on current which had been flowing just before the turnoff command. That peak current divides in some way between the off-turning collector and $C1$ through diode $D1$ which has latched in.

The amount of current I_{C1} flowing into $C1$ slows up the collector voltage rise time, and by making $C1$ large enough, the rising collector voltage and falling collector current intersect so low down on the rising collector voltage waveform that transistor dissipation is decreased significantly.

There is a limitation on how large $C1$ can be made, as can be seen in

Fig. 11.2, where it is shown that at the start of any turnoff—say, at $A1$—$C1$ must have no charge on it. At the end of the turnoff (at B), $C1$ has slowed up the voltage rise time but has accumulated a voltage $2V_{dc}$ (neglecting for the moment the leakage inductance spike).

At the start of the next turnoff at $A2$, $C1$ must again have no voltage across it. Thus, at some time between B and $A2$, $C1$ must be discharged. It is discharged in the interval C to $A2$ by resistor $R1$. When $Q1$ turns on at C, the top end of $C1$ goes immediately to ground and $C1$ discharges through $Q1$ and $R1$.

Thus, once $C1$ has been selected large enough to yield a sufficiently slowed up collector voltage rise time, $R1$ is chosen to discharge $C1$ to within 5 percent of its full charge in the minimum t_{on}. Or

$$3R1C1 = t_{on(min)} \tag{11.1}$$

But if $C1$ accumulates a voltage $2V_{dc}$ at each turnoff, it stores an energy of $0.5C1(2V_{dc})^2$ joules. If it dissipates that energy in $R1$ during each on time, power dissipation in $R1$ in watts (for T in seconds) is

$$PD_{R1} = \frac{0.5C1(2V_{dc})^2}{T} \tag{11.2}$$

It will be seen in the following section that power dissipation in $R1$ given by Eq. 11.2 seriously limits how much collector voltage rise time can be slowed up by increasing $C1$.

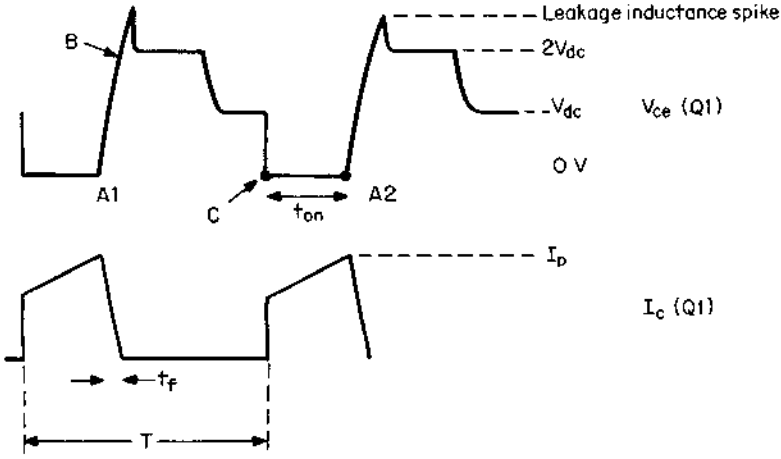


Figure 11.2 Snubber capacitor $C1$ slows up collector voltage rise time from $A1$ to B at turnoff. At the next turnoff at $A2$, $C1$ must lose the $2V_{dc}$ charge it picked up at time B. It loses that charge in the on time t_{on} just before a turnoff. With snubber in, $Q1$ dissipation is $(2V_{dc}I_p/12)(t_f/T)$.

11.4 Selection of Capacitor Size in *RCD* Snubber

Equation 11.2 shows that power dissipation in $R1$ is proportional to $C1$. Hence there must be some procedure to select $C1$ large enough to adequately slow up collector voltage rise time and yet not cause excessive dissipation in $R1$.

There is no best way to select $C1$. It is done in different ways by different designers who make different assumptions as to how much of the peak current in $Q1$ is available to charge $C1$, how much the $Q1$ collector voltage rise time is to be slowed up, and how fast the collector current fall time to zero is. The last depends strongly on the turnoff speed of the transistor actually used and the amount of its reverse base drive.

The following has been found to be a satisfactory scheme to select $C1$ for bipolar transistors.

When the $Q1$ base receives its turnoff command, the peak current the collector had been carrying is partly diverted into $C1$ as the voltage across it starts to increase. It is assumed that half the initial peak current I_p is diverted into $C1$ and half remains flowing into the gradually off-turning collector (since the transformer leakage inductance keeps the total current constant at I_p for some time).

Capacitor $C1$ is then selected so that the collector voltage will be permitted to rise to $2V_{dc}$ in the same time t_f that the collector current will fall to zero. This last is read or estimated from data sheets for the specific transistors. Thus

$$C1 = \frac{I_p}{2}\frac{t_f}{2V_{dc}} \tag{11.3}$$

Now capacitance $C1$ can be calculated from Eq. 11.3, resistance $R1$ can be calculated from Eq. 11.1 from the known minimum on time, and dissipation in resistor $R1$ can be calculated from Eq. 11.2.

Overlap dissipation in $Q1$ can be estimated as in Sec. 1.3.4 on the assumption that during the turnoff interval t_f, $Q1$ current $I_p/2$ starts falling toward zero at the same time its voltage starts rising toward $2V_{dc}$ and that the collector current reaches zero at the same time the collector voltage reaches $2V_{dc}$. In Sec. 1.3.4, it was noted that the collector dissipation during that interval t_f is

$$\frac{I_{max}V_{max}}{6} = \frac{I_p}{2}\frac{2V_{dc}}{6}$$

and averaged over one period T, transistor dissipation is

$$\frac{(I_p/2)(2V_{dc})t_f}{6T} \tag{11.4}$$

11.5 Design Example—*RCD* Snubber

Design the *RCD* snubber for the forward converter of Sec. 11.2. Recall that the peak current I_p just before turnoff was 3.45 A, transistor fall time was 0.3 μs, and transistor dissipation without a snubber was 19 W. From Eq. 11.3

$$C1 = \frac{(I_p/2)t_f}{2V_{dc}}$$

$$= \frac{(3.45/2)(0.3 \times 10^{-6})}{2 \times 184}$$

$$= 0.0014\ \mu F$$

and from Eq. 11.1

$$R1 = \frac{t_{on(min)}}{3C1}$$

But recall that a forward converter transformer is designed so that maximum transistor on time is at minimum DC voltage and will be forced to be $0.8T/2$. For a switching frequency of 100 kHz, this is 4 μs. But in Sec. 11.2, maximum and minimum input voltages were 15 percent above and below the nominal value. Then $t_{on(min)}$ is 4/1.3 or 3 μs. Then from Eq. 11.1

$$R1 = \frac{3 \times 10^{-6}}{3 \times 0.0014 \times 10^{-6}}$$

$$= 714\ \Omega$$

and from Eq. 11.2

$$PD_{R1} = \frac{0.5C1(2V_{dc})^2}{T}$$

$$= \frac{0.5(0.0014 \times 10^{-6})(2 \times 184)^2}{10 \times 10^{-6}}$$

$$= 9.5\ W$$

and from Eq. 11.4, $Q1$ overlap dissipation is

$$PD_{Q1} = \frac{(I_p/2)(2V_{dc})t_f}{6T}$$

$$= \frac{(3.45/2)(2 \times 184)(0.3 \times 10^{-6})}{6 \times 10 \times 10^{-6}}$$

$$= 3.2\ W$$

Thus, although 9.5 W has been added in a resistor $R1$, dissipation in the transistor, which is far more failure-prone, has been reduced from 19 to 3.2 W. Actually overlap dissipation in the transistor may be more than the calculated 3.2 W, as a best-case scenario has been assumed for the relative timing of the current fall time and voltage rise time (Sec. 1.3.4).

If temperature measurement on the $Q1$ case indicates that it is running too warm, $C1$ can be increased, but only at the expense of more dissipation in $R1$. But this is far more acceptable than dissipation in $Q1$.

It is sometimes thought that decreasing $R1$ decreases its dissipation. This, of course, is not so, as Eq. 11.2 indicates. Making $R1$ smaller only ensures that $C1$ is totally discharged earlier than needed (earlier than $A2$ in Fig. 11.2). Dissipation is equal to the energy stored on $C1$ (and released through $R1$) per cycle, and energy stored is proportional only to the magnitude of $C1$ and the square of the voltage to which it is charged at turnoff. Dissipation in $R1$ is thus unrelated to its magnitude.

11.5.1 *RCD* snubber returned to positive supply rail

The *RCD* snubber is often (and preferably) returned to the positive supply rail as shown in Fig. 11.3. It works exactly in the same way as when it is returned to ground as in Fig. 11.1. At turnoff, $D1$ latches in and $C1$ slows up collector voltage rise time with its charging current flowing into V_{dc}. At the following turnon, $C1$ is discharged through $Q1$ and the supply source V_{dc}.

The advantage of returning $R1$, $D1$ to V_{dc} instead of to ground is

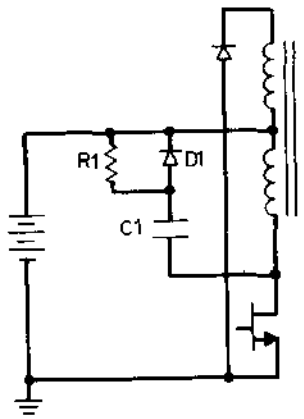


Figure 11.3 When snubber is returned to positive rail, voltage stress on $C1$ is half that when it is returned to ground.

that the maximum voltage stress on $C1$ is now only V_{dc} instead of $2V_{dc}$ when it is returned to ground.

11.6 Nondissipative Snubbers[1–8]

The conventional *RCD* snubber most often ends up dissipating 10 W or more for off-line switching supplies operating at over 50 kHz. This is troublesome not only for the added dissipation but also because of the size and required location of the snubber resistor. General practice is to derate power resistors by a factor of 2 and thus 10 W of dissipation usually requires the use of a 20-W resistor.

A 20-W resistor is quite large, and finding a location for it is often difficult. Also, in dissipating 10 W, it heats any surrounding nearby components, which further complicates the selection of a satisfactory location for it.

"Dissipationless snubbers," as in Fig. 11.4, although more complex, provide a good solution to this problem. Just as does the conventional *RCD* dissipative snubber, they use a capacitor to slow up collector voltage rise time. But they do not get rid of the resulting charge on the capacitor by discharging it through a resistor, which would waste power.

Instead, they change the stored electrostatic energy on the capacitor into electromagnetic energy in the form of stored current in an inductor. Then later in time (but before the next cycle when the capacitor must again be discharged), the inductor current is forced to discharge its stored current back into the DC input bus. Thus no energy is wasted—it is first stored on a slow-up capacitor, then returned back with negligible loss to the input bus.

The details can be seen in Fig. 11.4. When $Q1$ turns off, its collector voltage starts rising, $D1$ latches in, and $C1$ slows up the voltage rise time just as in Fig. 11.3. The bottom end of $C1$ is driven up to $2V_{dc}$, and its top end is clamped to V_{dc} through $D1$.

Now when $Q1$ turns on, capacitor $C1$ with a voltage V_{dc} across it is switched across $L1$ and $D2$ in series. An oscillatory current "ring" commences with a frequency $F_r = 1/2\pi\sqrt{L1C1}$.

At the end of a half period of this ring, the electrostatic energy in the form of voltage on $C1$ has been changed into electromagnetic energy in the form of stored current in $L1$. During the next half period of this ring, the voltage at the top end of $L1$ rings positive enough for $D1$ to conduct and the current in $L1$ flows through it back into the supply bus. If $L1$ is a high-Q inductor, all the energy stored in it in the first half cycle of the ring is returned back to V_{dc} in the second half cycle.

Capacitor $C1$ is first chosen large enough to slow up the $Q1$ voltage

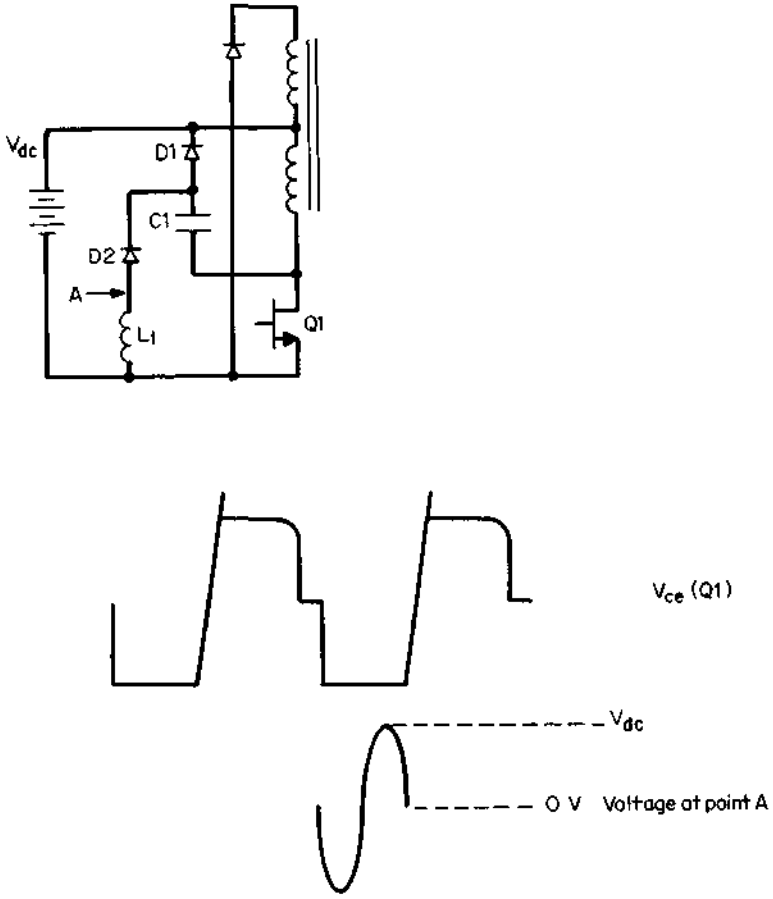


Figure 11.4 A dissipationless snubber. Capacitor $C1$ slows up $Q1$ collector voltage rise time as with *RCD* snubber. But at $Q1$ turnon, stored energy on $C1$ is transferred to stored magnetic energy in the form of current in $L1$ during the first half cycle of a "ring." During the second half cycle of the ring, point A goes positive and returns this energy to V_{dc} without loss.

rise time as much as required. Then $L1$ is chosen so that the full ring period is somewhat less than the minimum $Q1$ on time.

11.7 Snubber Reduction of Leakage Inductance Spike to Avoid Second Breakdown

The snubber offers a second very important advantage in addition to slowing up voltage rise time and thus decreasing average transistor dissipation. It prevents second breakdown, which occurs if the instantaneous voltage and current cross the reverse-bias safe operating area

(RBSOA) boundary given in the manufacturer's data sheets (Fig. 11.5).

This boundary can be crossed by the omnipresent leakage inductance spike (Fig. 2.10) which occurs at the instant of turnoff. Transistor manufacturers state that if the boundary is crossed even once, second-breakdown failure may occur.

An exact analysis of the sequence of collector voltage changes and their magnitudes at the instant of turnoff is not possible without computer analysis. But the following not-too-exact discussion illustrates the magnitude of the problem and how the snubber capacitor reduces the leakage inductance spike.

When $Q1$ turns off, the transformer leakage inductance keeps the current which had been flowing in it constant for some short time. It was assumed above that, to an approximation, half that current continues to flow into the slowly off-turning transistor and half flows into the snubber capacitor $C1$.

Now in Fig. 11.6, when $Q1$ turns off, the voltage across the magnetizing inductance reverses and this reverses the voltage across the reset winding N_r. The top end of N_r goes immediately negative and is clamped to ground by $D4$. Since $N_p = N_r$, this clamps the voltage across L_m to V_{dc} and the voltage at point A rises to $2V_{dc}$.

But half the original peak current $I_p/2$ flowed into the series combination of the leakage inductance L_l, the snubber capacitor $C1$, and diode $D1$. This causes a sinusoidal ring whose half period is $\pi\sqrt{L_lC_1}$. To a close approximation, the amplitude of the first half cycle is $1/2\pi(\sqrt{L_l/C1})$.

This first half-cycle ring, which is the leakage inductance spike, sits on top of the voltage $2V_{dc}$ at point A. The quantity $\sqrt{L_l/C1}$ is often referred to as the *characteristic impedance* of the LC circuit. Thus increasing $C1$ to slow up collector voltage rise time also decreases the leakage inductance spike.

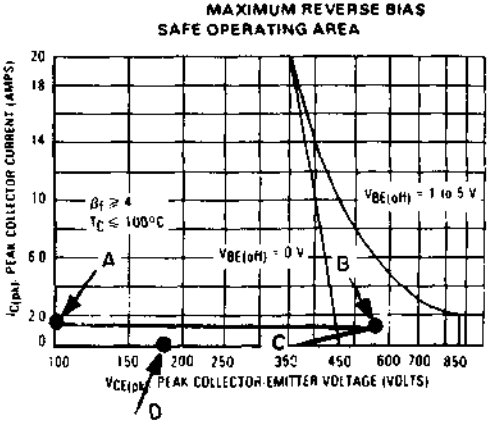


Figure 11.5 Reverse-bias safe operating area (RBSOA) for fast 15 A, 450-V transistor type 2N6836. At turnoff, because leakage inductance in transformer keeps current constant, operation is along A to B at the tip of the leakage inductance spike. If not for the reverse base bias, operation would have crossed the RBSOA boundary and transistor would have failed in second-breakdown mode. The *RCD* snubber, in addition to decreasing overlap dissipation, decreases amplitude of the leakage inductance spike. (*Courtesy Motorola Inc.*)

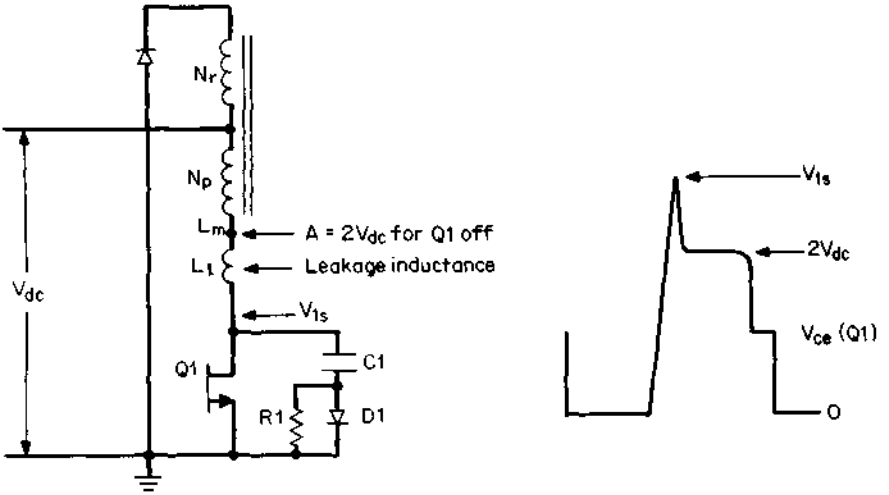


Figure 11.6 The leakage spike is roughly $(I_p/2)(L_1/C1)$ above $2V_{dc}$.

It is of interest to calculate the magnitude of the leakage inductance spike on the basis of these observations. The leakage inductance of a 100-kHz transformer for the design example at the end of Sec. 11.4 is expected to be about 15 μH. Then in the preceding example $\sqrt{L_l/C1} = \sqrt{15 \times 10^{-6}/0.0014 \times 10^{-6}} = 103\ \Omega$, and for the 3.45-A peak current in $Q1$ just prior to turnoff, the leakage inductance spike amplitude is (3.45/2)103 = 178 V. Peak voltage at the top of the leakage inductance spike is then $2V_{dc}$ + 178 = 547 V.

Although not too precise an analysis, this yields the amplitude of the leakage inductance spike sufficiently accurately to explain the possibility of second breakdown.

In Fig. 11.5, just prior to turnoff, the location of $Q1$ on its volt-ampere curve is at point A—3.45 A at close to 0 V. As the transistor is turned off, its load-line locus is along the path $ABCD$. The leakage inductance keeps total current constant. But half starts flowing into $C1$, leaving 1.73 flowing into the transistor itself. The transistor itself then moves horizontally along the line AB. At B, the current is still 1.73 A and the voltage is 547 V. After the short duration of the leakage spike, the locus drops down to $2V_{dc}$, until the transformer core resets, and then back to V_{dc}, where it stays until the next turnon (Fig. 2.10.).

Now Fig. 11.5 shows that if the transistor is subject to a 5-V reverse bias at turnoff, the point at 547 V, 1.73 A is still within the manufacturer's RBSOA curve and second breakdown should not occur. With no

reverse bias, the transistor is outside the RBSOA curve and second breakdown failure will occur.

Thus, although the initial selection of the snubber capacitor $C1$ (Fig. 11.1) was to slow up collector voltage rise time, it may have to be increased above the value calculated from Eq. 11.3 to reduce the leakage inductance spike. At higher output powers, even with a 5-V reverse base bias at turnoff, the leakage inductance spike may cross the RBSOA curve as the turnoff locus is at a higher current along its horizontal portion. This then would require a larger snubber capacitor.

11.8 Transformer-Aided Snubber

Figure 11.7 shows a scheme which reduces and even totally eliminates the leakage inductance spike for an *RCD* snubber. It does this without requiring an increase in the snubber capacitor—which would have increased snubber resistor dissipation.

The tradeoff for this is the addition of the very small transformer $T1$. It still requires the conventional *RCD* snubber, but $C1$ can be considerably less than if the scheme were not used. It works as follows. As shown in Fig. 11.7, $T1$ is a small 1/1 transformer. Its core area and number of turns must be such as to sustain the maximum volt-second product across the $T2$ primary when $Q1$ is off. Since the core size should be as small as possible, this means a small core area and hence a relatively large number of turns. But since the power which $T1$ will be carrying will be quite small, primary and secondary wire size can be small, thus permitting a small core.

Note the dot polarity on the $T1$ primary and secondary. Diode $D1$ does not conduct until the dot end of N_s reaches V_{dc}, and it does not reach V_{dc} until the voltage across the $T1$ primary rises to V_{dc} with its

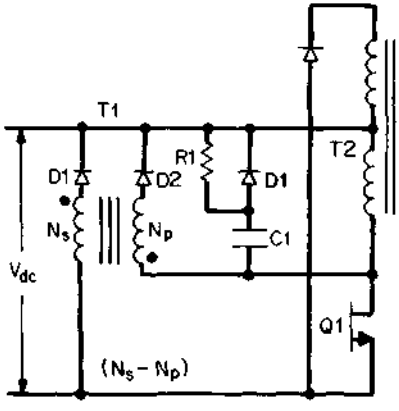


Figure 11.7 A leakage spike clipping aid to an *RCD* snubber. Transformer $T1$ is a small 1/1 transformer. When $Q1$ collector voltage reaches $2V_{dc}$, $D2$ latches in, forcing $D1$ to latch in and clamp $V_{ce(Q1)}$ to $2V_{dc}$. Thus no leakage spike at the $Q1$ collector above $2V_{dc}$ if $T1$ has very little leakage inductance. This minimizes the size of $C1$.

dot end positive with respect to its no-dot end. But that occurs only when the $Q1$ collector reaches $2V_{dc}$.

Thus, when the $Q1$ collector reaches $2V_{dc}$, $D1$ latches in and clamps the voltage across N_s to V_{dc}. Since the transformer turns ratio is 1/1, the collector voltage is clamped at $2V_{dc}$—i.e., no leakage spike at the collector above $2V_{dc}$. For this to be very effective, the leakage inductance of $T1$ itself must be very small.

When $D1$ has latched in and conducts, the energy stored in the $T2$ leakage inductance is returned with no loss to the V_{dc} supply bus.

References

1. E. Whitcomb, "Designing Non-Dissipative Current Snubbers for Switched Mode Converters," *Proceedings Powercon 6*, 1979.
2. W. Shaunessy, "Modeling and Design of Non-Dissipative *LC* Snubber Networks," *Proceedings Powercon 7*, 1980.
3. L. Meares, "Improved Non-Dissipative Snubbers for Buck Regulators and Current Fed Inverters," *Proceedings Powercon 9*, 1982.
4. M. Domb, R. Redl, and N. Sokal, "Non-Dissipative Turn Off Snubber Alleviates Switching Power Dissipation, Second Breakdown Stress," *PESC Record*, 1982.
5. T. Ninomiya, T. Tanaka, and K. Harada, "Optimum Design of Non-Dissipative Snubber by Evaluation of Transistor's Switching Loss," *PESC Record*, 1985.
6. T. Tanaka, T. Ninomiya, and K. Harada, "Design of a Non Dissipative Snubber in a Forward Converter," *PESC Record*, 1988.

Chapter

12

Feedback-Loop Stabilization

12.1 Introduction

Before going into the details of stabilizing a feedback loop, it is of interest to consider in a semiquantitative way, why a feedback loop may oscillate.

Consider the negative-feedback loop for a typical forward converter in Fig. 12.1. The essential error-amplifier and PWM functions are contained in all pulse-width-modulating chip. The chip also provide many other functions, but for understanding the stability problem, only the error amplifier and pulse-width modulator need be considered.

For slow or DC variations of the output voltage V_o, the loop is, of course, stable. A small, slow variation of V_o due to either line input or load changes will be sensed by the inverting input of error amplifier EA via the sampling network $R1$, $R2$ and compared to a reference voltage at the noninverting EA input. This will cause a small change in the DC voltage level V_{ea} at the EA output and at the A input to the pulse-width-modulator PWM.

The PWM, as described heretofore, compares that DC voltage level to a roughly 0- to 3-V triangle V_t at its B input. It generates a rectangular pulse whose width t_{on} is equal to the time from the start of the triangle t_0 until t_1, the time the triangle crosses the DC voltage level at the B input of the PWM. That pulse fixes the on time of the output transistor of the chip and should also fix the on time of the power transistor.

Thus a slow increase (e.g.) in V_{dc} causes a slow increase in V_y and hence a slow increase in V_o since $V_o \cong V_y t_{on}/T$. The increase in V_o causes an increase in V_s and hence a decrease in V_{ea}. Since t_{on} is the time from the start of the triangle to t_1, this causes a decrease in t_{on} and restores V_o to its original value. Similarly, of course, a decrease in V_{dc} causes an increase in t_{on} to maintain V_o constant.

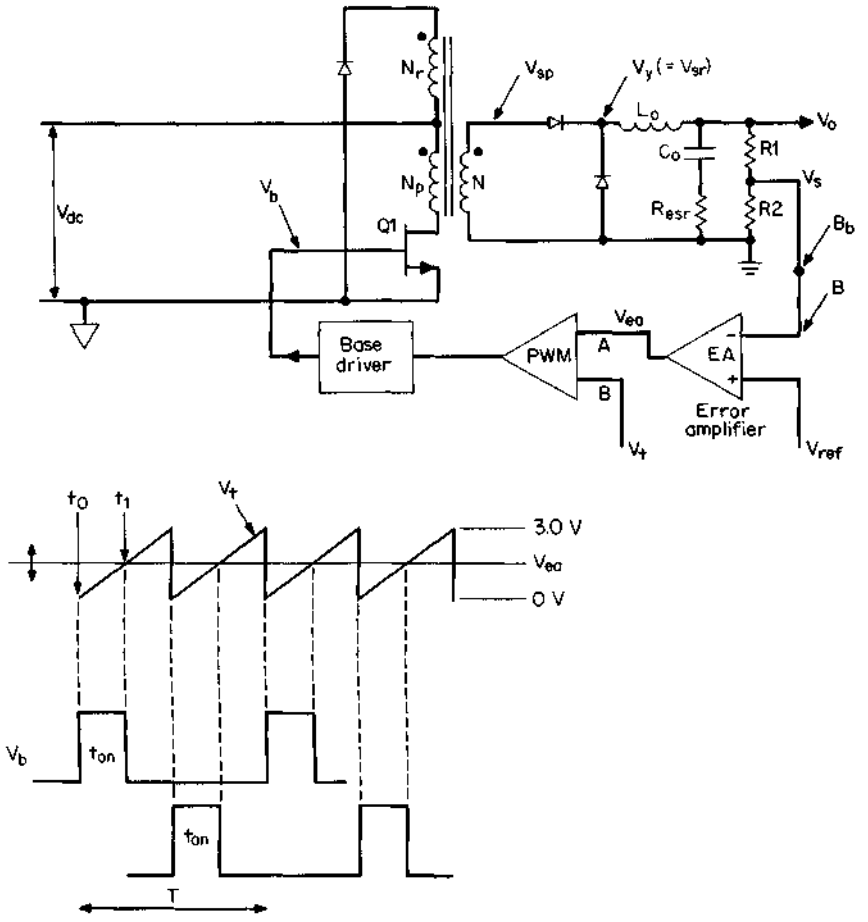


Figure 12.1 A closed feedback loop in a typical forward converter.

Drive to the power transistor may be taken from either the emitter or the collector of the chip's output transistor via a current-amplifying base driver. But from whichever point—emitter or collector—output is taken, it must be ensured that polarities are such that an increase in V_o causes a decrease in t_{on}.

Note that most PWM chips have their output transistors on for the time from t_0 to t_1. With such chips, V_s is fed to the inverting EA input and for an NPN power transistor, its base (or gate if a MOSFET) is driven from the emitter of the chip's output transistor.

In some PWM chips (TL494 family), however, the output transistor is on from the time the triangle at the V_t input to the PWM crosses the V_{ea} level (t_1) until the end of the triangle at t_2. With such chips, if the

NPN power transistor were to be turned on when the chip output transistor was on (drive from the chip transistor's emitter), that would cause the power transistor on time to increase as V_{dc} increased. This, of course, would be positive rather than negative feedback.

Thus, with chips of the TL494 family, V_s is fed to the noninverting input to the EA. This causes the output transistor on time to decrease as V_o increases and permits the power transistor to be driven from the chip transistor's emitter.

The circuit in Fig. 12.1 thus provides negative feedback and a stable circuit at low frequencies. But within the loop, there exist low-level noise voltages or possible voltage transients which have a continuous spectrum of sinusoidal Fourier components. All these Fourier components suffer gain changes and phase shifts in the L_o, C_o output filter, the error amplifier, and the PWM from V_{ea} to V_{sr}. At one of these Fourier components, the gain and phase shifts can result in positive rather than negative feedback and thereby result in oscillation as described below.

12.2 Mechanism of Loop Oscillation

Consider the forward converter feedback loop of Fig. 12.1. Assume for a moment that the loop is broken open at point B,the inverting input to the error amplifier. At any of the Fourier components of the noise, there is gain and phase shift from B to V_{ea}, from V_{ea} to the average voltage at V_{sr}, and from the average voltage at V_{sr} through the L_o, C_o filter around back to B_b (just before the loop break).

Now assume that a signal of some frequency f_1 is injected into the loop at B and comes back around as an echo at B_b. The echo is modified in phase and gain by all the previously mentioned elements in the loop. If the modified echo has returned exactly in phase with and is equal in amplitude to the signal which started the echo, if the loop is now closed (B_b closed to B) and the injected signal is removed, the circuit will continue to oscillate at the frequency f_1. The initial signal which starts the echo and maintains the oscillation is the f_1 Fourier component in the noise spectrum.

12.2.1 Gain criterion for a stable circuit

Thus the first criterion for a stable loop is that at the frequency where the total open-loop gain is unity (the *crossover frequency*), the total open-loop phase shift of all elements involved must be less than 360°. The amount by which the total phase shift is less than 360° (at the frequency where the total open-loop gain is unity) is called the *phase margin*.

To ensure a stable loop under worst-case tolerances of the associated components, the usual practice is to design for at least a 35° to 45° phase margin with nominal components.

12.2.2 Gain slope criteria for a stable circuit

At this point, a universally used jargon expression describing the gain slope is introduced. Gain versus frequency is usually plotted in decibels (dB) on semilog paper as in Fig. 12.2. If the scales are such that a linear distance of 20 dB (numerical gain of 10) is equal to the linear distance of a factor of 10 in frequency, lines representing gain variations of ±20 dB/decade have slopes of ±1. Circuit configurations having a gain variation of ±20 dB per decade are thus described as having " ± 1 gain slopes."

An elementary circuit having a gain slope of −1 (beyond the frequency $f_p = 1/2\pi R1C1$) between output and input is the *RC* integrator of Fig. 12.2*a*. The *RC* differentiator of Fig. 12.2*b* has a +1 gain slope (below the frequency $f_z = 1/2\pi R2C2$) or a gain variation of +20 dB/decade between output and input. Such circuits have only 20 dB/decade gain variations because as frequency increases or decreases by a factor of 10, the capacitor impedance decreases or increases by a factor of 10 but the resistor impedance remains constant.

A circuit which has a −2 or −40 dB/decade gain slope (beyond the frequency $F_o = 1/2\pi\sqrt{L_oC_o}$) is the output *LC* filter (Fig. 12.2*c*), which has no resistance (ESR) in its output capacitor. This, of course, is because as frequency increases by a factor of 10, the inductor impedance increases and the capacitor impedance decreases by a factor of 10.

Now gain and phase shift versus frequency for an L_oC_o filter are plotted in Fig. 12.3*a* and 12.3*b* for various values of output resistance R_o[2]. The gain curves are normalized for various ratios of $k_1 = f/F_o$ where $F_o = 1/2\pi\sqrt{L_oC_o}$ and for various ratios $k_2 = R_o/\sqrt{L_o/C_o}$.

Figure 12.3*a* shows that whatever the value of k_2, all gain curves, beyond the so-called corner frequency of $F_o = 1/2\pi\sqrt{L_oC_o}$ asymptotically approach a slope of −2 (−40 dB/decade). The circuit for $k_2 = 1.0$ is referred to as the *critically damped* circuit. The critically damped circuit has a very small resonant "bump" in gain and at the corner frequency F_o, starts immediately falling at a −2 slope.

For k_2 greater than 1 the circuit is described as *underdamped*. It is seen that underdamped *LC* filters can have a very large resonant bump in gain at F_o.

Circuits of k_2 less than 1.0 are *overdamped*. It is seen in Fig. 12.3*a* that overdamped *LC* filters also asymptotically approach a gain slope of −2. But for a heavily overdamped ($k_2 = 0.1$) filter, the frequency at

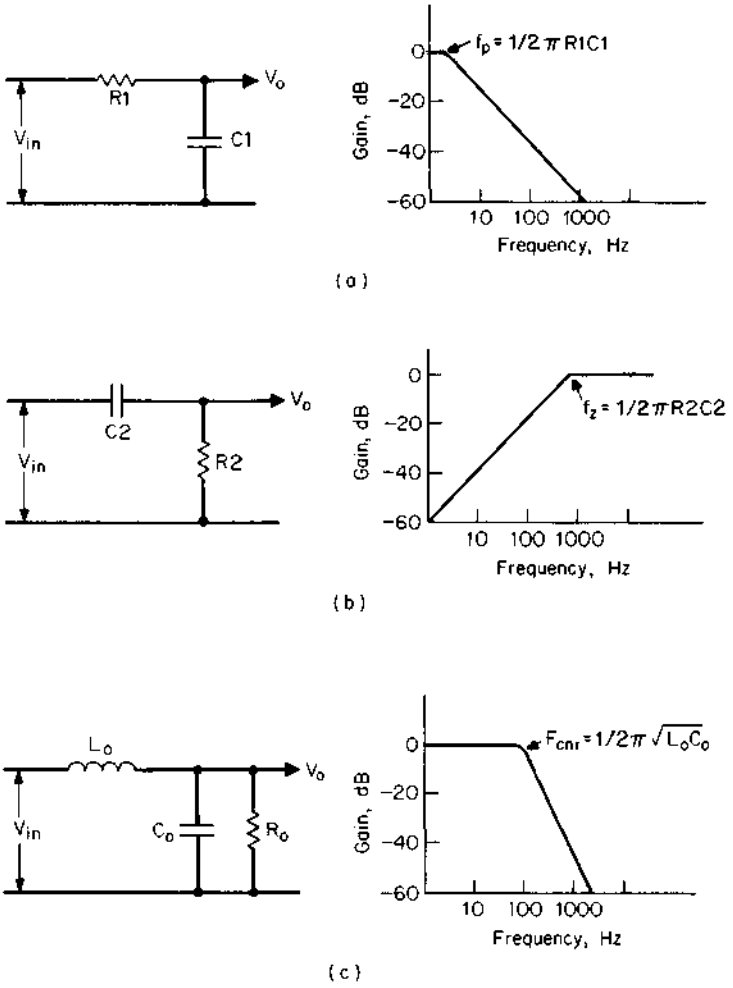
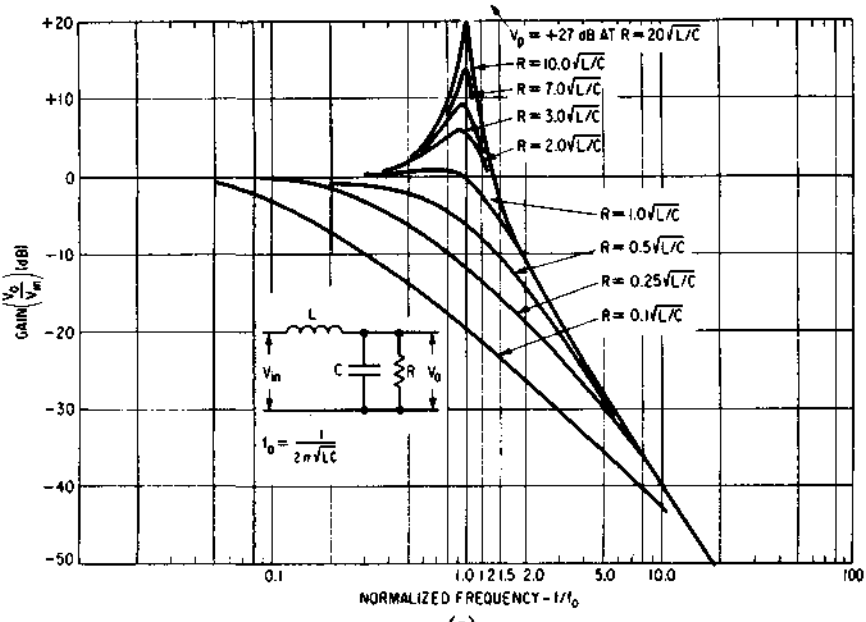
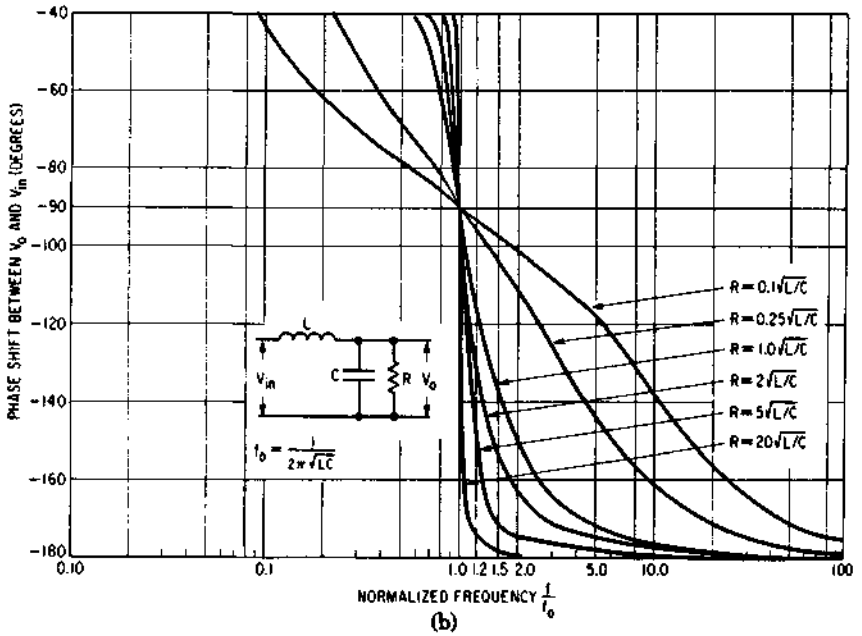


Figure 12.2 (*a*) An *RC* integrator has a gain dV_o/dV_{in} of −20 dB/decade beyond $F_p = \frac{1}{2}\pi R1C1$. If the scales are such that 20 dB is the same linear distance as 1 decade in frequency, a gain slope of −20 dB/decade has a −1 slope. Such a circuit is referred to as a −1 *slope* circuit. (*b*) An *RC* differentiator has a gain of +20 dB/decade. At $F_z = 1/2\pi R2C2$, where $X_{C2} = R_2$, gain asymptotically approaches 0 dB. If scales are such that 20 dB is the same linear distance as 1 decade in frequency, a gain slope of +20 dB/decade has a +1 slope. Such a circuit is referred to as a +1 *slope* circuit. (*c*) An *LC* filter has a gain (dV_o/dV_{in}) of unity (0 dB) up to its corner frequency of $F_{cnr} = 1/2\pi\sqrt{L_oC_o}$ when critically damped ($R_o = \sqrt{L_o/C_o}$). Beyond F_{cnr} it commences falling at a rate of −40 dB/decade. This is so because for every decade increase in frequency, X_L increased and X_c decreases in impedance by a factor of 10. If scales are such that 20 dB is the same linear distance as 1 decade in frequency, a gain slope of −40 dB/decade has a −2 slope. Such a circuit is referred to as a −2 *slope* circuit.



(a)



(b)

Figure 12.3 (*a*) Gain versus frequency for switching regulator *LC* filter. (*b*) Phase shift versus frequency for switching regulator *LC* filter. (*Courtesy Switchtronix Press*).

which the gain slope has come close to −2 is about 20 times the corner frequency F_o.

Figure 12.3*b* shows phase shift versus normalized frequency (f/F_o) again for various ratios of $k_2 = R_o/\sqrt{L_o/C_o}$. It is seen for any value of k_2, that the phase shift between output and input is 90° at the corner frequency ($F_o = 1/2\pi\sqrt{L_oC_o}$). And for highly underdamped filters (R_o greater than $\sim 5\sqrt{L_o/C_o}$), phase shift varies very rapidly with frequency. The shift is already 170° at a frequency of $1.5F_o$ for $R_o = 5\sqrt{L_o/C_o}$.

In contrast, a circuit with a −1 gain slope can never yield more than a 90° phase shift, and its rate of change of phase shift with frequency is far slower than that of a circuit −2 gain slope as exemplified in Fig. 12.3*b*.

This leads to the second criterion for a stable circuit. The first criterion was that the total phase shift at the crossover frequency (frequency where total open-loop gain is unity or 0 dB) should be short of 360° by the "phase margin," which is usually taken as at least 45°.

This second criterion for a stable circuit is that to prevent rapid changes of phase shift with frequency characteristic of a circuit with a −2 gain slope, the slope of the open-loop gain–frequency curve of the entire circuit (arithmetic sum in decibels of all the gain elements involved) as it passes through crossover frequency should be −1. This is shown in Fig. 12.4.

It is not an absolute requirement that the total open-loop gain curve must come through crossover at a −1 gain slope. But it does provide insurance that if any phase-shift elements have been overlooked, the small phase shift and relatively slow phase-shift–frequency curve characteristic of a −1 gain slope element will still preserve an adequate phase margin.

The third criterion for a stable loop is to provide the desired phase margin, which will be set at 45° herein (Fig. 12.4).

To satisfy all three criteria, it is necessary to know how to calculate gains and phase shifts of all the elements in Fig. 12.1. This is shown below.

12.2.3 Gain characteristic of *LC* output filter with and without equivalent series resistance (ESR) in output capacitor

Aside from the flyback (which has an output capacitor filter only), all topologies discussed herein have an output *LC* filter.[1] The gain-versus-frequency characteristic of this output *LC* filter is of fundamental importance. It must be calculated first as it determines how the gain and phase shift-versus-frequency characteristics of the error amplifier must be shaped to satisfy the three criteria for a stable loop.

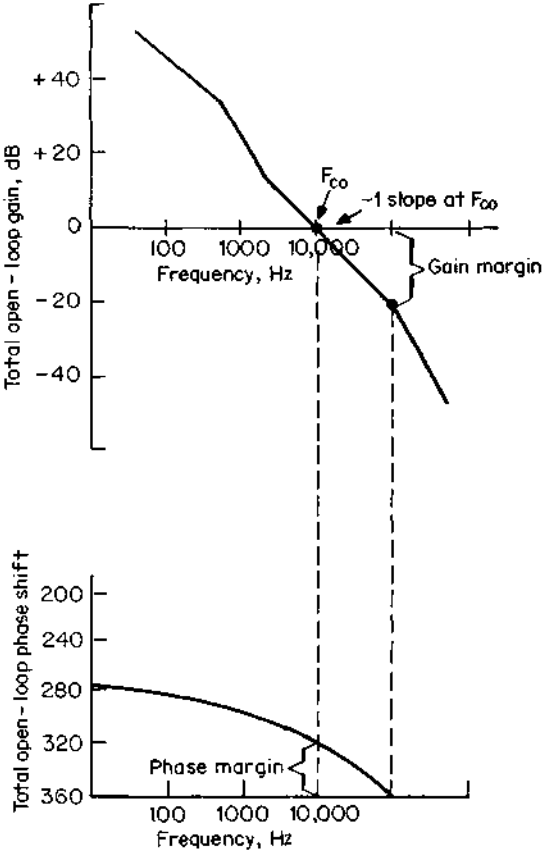


Figure 12.4 Total open-loop gain and phase shift. The frequency at which the total open-loop gain is 0 dB (F_{co}) is forced to be 0 dB is usually chosen one-fourth and one-fifth the switching frequency. For loop to be stable, total open-loop phase shift at F_{co} should be short of 360° by as large a value as possible. The amount by which it is short of 360° is the *phase margin*. A usual phase margin to strive for is 45°. A second criterion for a stable loop is that the total open-loop gain pass through F_{co} at a −1 slope.

The gain characteristic of an output LC filter with various output load resistances is shown in Fig. 12.3*a*. This curve assumes that the output capacitor has zero equivalent series resistance (ESR). For the purpose of this discussion, it is sufficiently accurate to assume that the filter is critically damped, that is, $R_o = 1.0\sqrt{L_o/C_o}$. If the circuit is made stable for the gain curve corresponding to $R_o = 1.0\sqrt{L_o/C_o}$, it will be stable at other loads. Nevertheless, the circuit merits exami-

nation for light loads ($R_o \gg 1.0\sqrt{L_o/C_o}$) because of the resonant bump in gain at the LC corner frequency $F_o = 1/2\pi\sqrt{L_oC_o}$. This will be considered below.

Thus the gain characteristic of the output LC filter with zero ESR will be drawn as curve 12345 in Fig. 12.5*a*. There it is seen that the gain is 0 dB (numerical gain of 1) at DC and low frequencies up to the

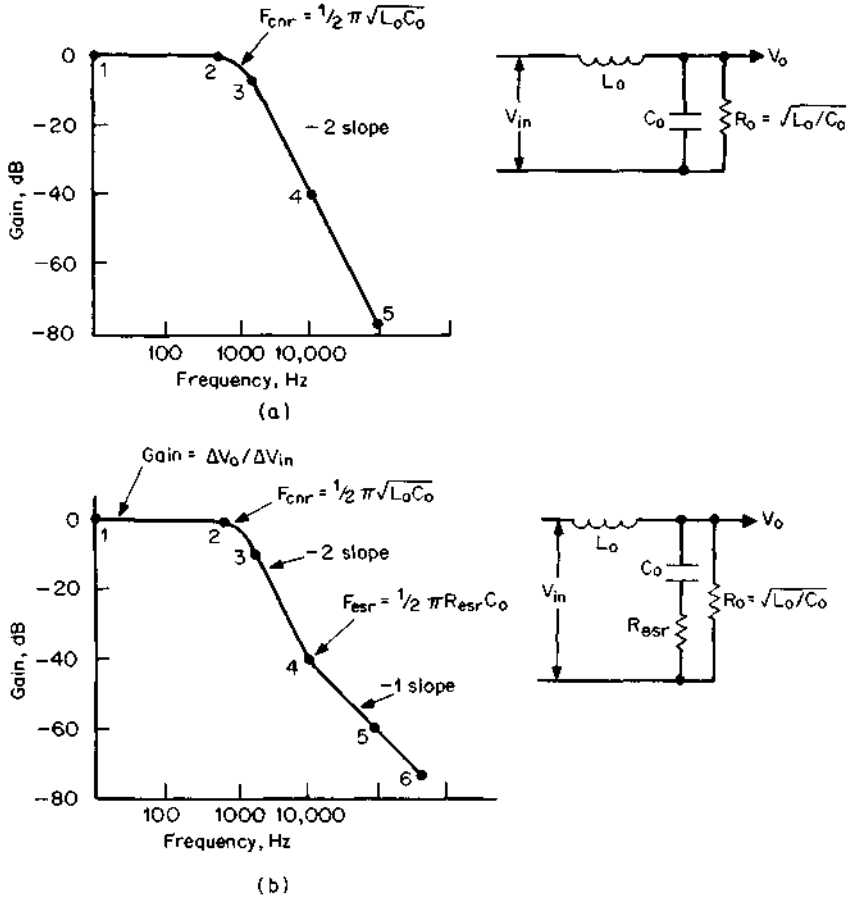


Figure 12.5 (*a*) Gain versus frequency for a critically damped LC filter in which the output capacitor has zero ESR. (*b*) Gain versus frequency for a critically damped LC filter in which the output capacitor has an equivalent series resistor (R_{esr}). When C_o has an ESR, the gain slope still breaks from horizontal to a −2 slope at F_{cnr}. But at a frequency $F_{esr} = 1/2\pi\,(R_{esr}C_o)$, it breaks into a −1 slope. This is because at F_{esr}, $X_{co} = R_{esr}$, the impedance of C_o becomes increasingly small compared to R_{esr}. The circuit is now an LR rather than an LC circuit. An LR circuit falls at a −1 slope because as frequency increases, the impedance of the series L increases, but that of the shunt R remains constant.

corner frequency $F_o = 1/2\pi\sqrt{L_oC_o}$. At DC and frequencies less than F_o, the impedance of C_o is much greater than that of L_o and the output-input gain is unity.

Beyond F_o, the impedance of C_o decreases and that of L_o increases at the rate of 20 dB/decade, making the gain slope fall at the rate of −40 dB/decade or at a −2 slope. Of course, the transition to a −2 slope at F_o is not as abrupt as shown. The actual gain curve leaves 0 dB smoothly just before F_o and asymptotically approaches the −2 slope shortly after F_o. But for the purposes of this discussion, gain will be shown with a relatively abrupt transition as curve 12345 in Fig. 12.5*a*.

Most filter capacitor types have an internal resistance R_{esr} in series with their output leads as shown in Fig. 12.5*b*. This modifies the gain characteristic between the output and input terminals in a characteristic way.

Beyond F_o, in the lower frequency range where the impedance of C_o is much greater than R_{esr}, looking down to ground from V_o, the only effective impedance is that of C_o. In this frequency range, the gain still falls at a −2 slope. At higher frequencies, where the impedance of C_o is less than R_{esr}, the effective impedance looking down from V_o to ground is that of R_{esr} alone. Hence in that frequency range, the circuit is an *LR* rather than an *LC* circuit. In that frequency range, the impedance of L_o increases at the rate of 20 dB/decade but that of R_{esr} remains constant. Thus in that frequency range, gain falls at a −1 slope.

The break from a −2 to a −1 gain slope occurs at the frequency $F_{esr} = 1/2\pi R_{esr}C_o$, where the impedance of C_o is equal to R_{esr}. This is shown as G_{lc} in curve 123456 in Fig. 12.5*b*. The break in slope from −2 to −1 is, of course, asymptotic, but it is sufficiently accurate to assume it to be abrupt as shown.

12.2.4 Pulse-width-modulator gain

In Fig. 12.1, the gain from the error-amplifier output to the average voltage at V_{sr} (input end of the output inductor) is the PWM gain and is designated as G_{pwm}.

It may be puzzling how this can be referred to as a *voltage gain*. For at V_{ea}, there are DC voltage level variations proportional to the error-amplifier input at point *B*, and at V_{sr}, there are fixed-amplitude pulses of adjustable width.

The significance and magnitude of this gain can be seen as follows. In Fig. 12.1, the PWM compares the DC voltage level from V_{ea} to a 3-V triangle at V_t. In all PWM chips which produce two 180° out-of-phase adjustable-width pulses (for driving push-pulls, half or full bridges) these pulses occur once per triangle and have a maximum on or high time of a half period. After the PWM, the pulses are binary-

counted and alternately routed to two separate output terminals (see Fig. 5.2*a*). In a forward converter, only one of these outputs is used.

Now (Fig. 12.1*b*), when V_{ea} is at the bottom of the 3-V triangle, on time or pulse width at V_{sr} is zero. The average voltage V_{av} at V_{sr} is then zero as $V_{av} = (V_{sp} - 1)(t_{on}/T)$, where V_{sp} is the secondary peak voltage. When V_{ea} has moved up to the top of the 3-V triangle, $t_{on}/T = 0.5$ and $V_{av} = 0.5\,(V_{sp} - 1)$. The modulator DC gain G_m, then, between V_{av} and V_{ea} is

$$G_m = \frac{0.5(V_{sp} - 1)}{3} \tag{12.1}$$

This gain is independent of frequency.

There is also a loss G_s due to the sampling network R_1, R_2 in Fig. 12.1. Most of the frequently used PWM chips cannot tolerate more than 2.5 V at the reference input to the error amplifier (point *A*). Thus, when sampling a +5-V output, $R_1 = R_2$ and gain G_s between V_s and V_o in Fig. 12.1 is −6 dB.

12.2.5 Total output *LC* filter plus modulator and sampling network gain

From the above, the total gain G_t (in decibels) of the output *LC* filter gain G_f plus modulator gain G_m plus sampling network gain G_s is plotted as in Fig. 12.6. It is equal to $G_m + G_s$ from DC up to $F_o = 1/2\pi\sqrt{L_oC_o}$. At F_o, it breaks into a −2 slope and remains at that slope up to the frequency F_{esr} where the impedance of C_o equals R_{esr}. At that frequency, it breaks into a −1 slope.

From this curve, the error-amplifier gain and phase-shift-versus-frequency characteristic is established to meet the three criteria for a stable loop as described below.

12.3 Shaping Error-Amplifier Gain-Versus-Frequency Characteristic

Recall that the first criterion for a stable loop is that at the frequency F_{co} where the total open-loop gain is unity (0 dB), total open-loop phase shift must be short of 360° by the desired *phase margin*, which will herein be taken as 45°.

The sequence of steps is then first to establish the crossover frequency F_{co}, where the total open-loop gain should be 0 dB. Then choose the error-amplifier gain so that the total open-loop gain is forced to be 0 dB at that frequency. Next design the error-amplifier gain slope so that the total open-loop gain comes through F_{co} at a −1 slope (Fig. 12.4). Finally, tailor the error-amplifier gain versus frequency so that the desired phase margin is achieved.

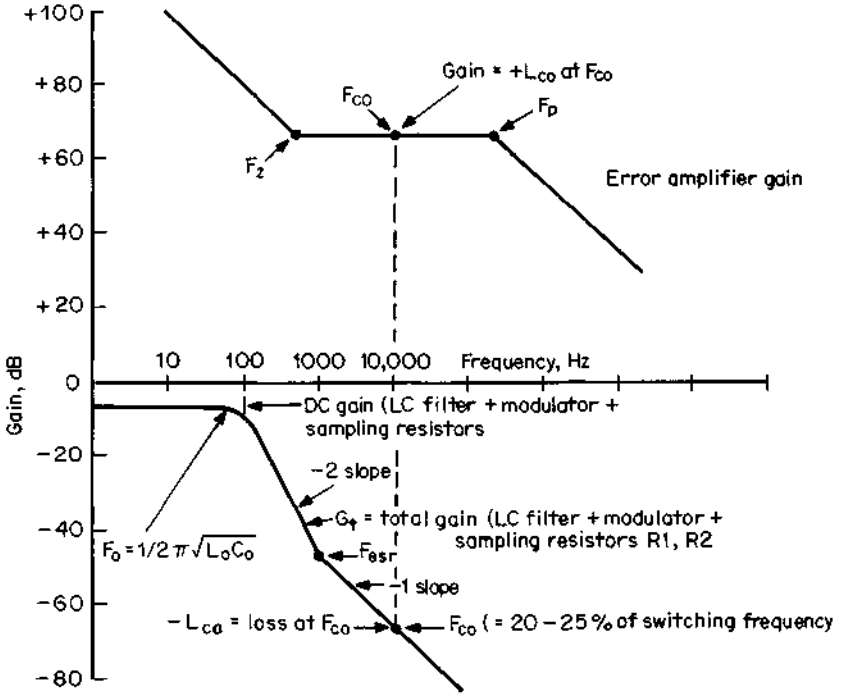


Figure 12.6 Gain G_t = sum of (*LC* filter + modulator + output voltage sampling resistors) gains determines error-amplifier gain. Error-amplifier gain at F_{co} is made equal and opposite to loss of G_t. Error-amplifier gain slope at F_{co} is made horizontal with upward and downward breaks at F_z and F_p. Location of F_z and F_p in frequency determines total circuit phase margin.

Sampling theory shows that F_{co} must be less than half the switching frequency for the loop to be stable. But it must be considerably less than that, or there will be large-amplitude switching frequency ripple at the output. Thus, the usual practice is to fix F_{co} at one-fourth to one-fifth the switching frequency.

Thus, refer to Fig. 12.6, which is the open-loop gain of the *LC* filter plus the PWM modulator plus the sampling network. The capacitor in the output filter is assumed in Fig. 12.6 to have an ESR which causes a break in the slope from −2 to −1 at $F_{esr} = 1/2\pi R_{esr}C_o$. Assume that F_{co} is one-fifth the switching frequency and read the loss in decibels at that point.

In most cases, the output capacitor will have an ESR and F_{esr} will come at a lower frequency than F_{co}. Thus at F_{co} the $G_1 = (G_{lc} + G_{pwm} + G_s)$ curve will already have a −1 slope.

Now when gains are plotted in decibels, both gains and gain slopes of gain elements in cascade are additive. Hence, to force crossover frequency to be at the desired one-fifth the switching frequency, choose

the error-amplifier gain at F_{co} to be equal and opposite in decibels to the $G_t = (G_{lc} + G_{pwm} + G_s)$ loss at that frequency.

That forces F_{co} to occur at the desired point. Then, if the error-amplifier gain slope at F_{co} is horizontal, since the G_t curve at F_{co} already has a -1 slope, the sum of the error amplifier plus the G_t curve comes through crossover frequency at the desired -1 slope and the second criterion for a stable loop has been met.

Now the error-amplifier gain has been fixed as equal and opposite to the G_t loss at F_{co} and to have a horizontal slope as it passes through F_{co} (Fig. 12.6). Such a gain characteristic can be achieved with an operational amplifier with a resistor input and resistor feedback as in Fig. 12.7*a*. Recall that the gain of such an operational amplifier is $G_{ea} = Z2/Z1 = R2/R1$. But how far in frequency to the left and right of F_{co} should it continue to have this constant gain?

Recall that the total open-loop gain is the sum of the error-amplifier gain plus G_1 gain. If the error-amplifier gain remained constant down to DC, the total open-loop gain would not be very large at 120 Hz—the frequency of the AC power line ripple.

Yet it is desired to keep power line ripple attenuated down to a very low level at the output. To degenerate the 120-Hz ripple sufficiently, the open-loop gain at that frequency should be as high as possible. Thus at some frequency to the left of F_{co}, the error-amplifier gain should be permitted to increase rapidly.

This can be done by placing a capacitor $C1$ in series with $R2$ (Fig. 12.7*b*). This yields the low-frequency gain characteristic shown in Fig. 12.6. In the frequency range where the impedance of $C1$ is small compared to $R2$, the gain is horizontal and simply equal to $R2/R1$. At

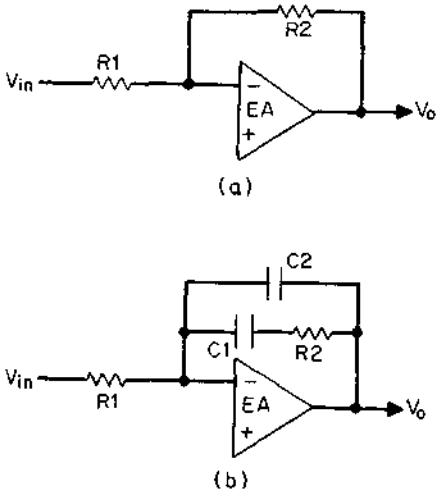


Figure 12.7 (*a*) Error amplifier with resistor feedback $R2$ and resistor input $R1$ arms have a gain equal to $R2/R1$ which is independent of frequency up to the frequency where the open-loop error amplifier inside the loop (EA) commences falling off the gain. (*b*) Using complex feedback and input arms permits shaping the gain-versus-frequency and phase-shift-versus-frequency curves. The configuration above has the gain-versus-frequency characteristic of Fig. 12.6.

lower frequencies where the impedance of $C2$ is much higher than $R2$, effectively $R2$ is out of the circuit and the gain is X_{c1}/R_1. This gain increases at the rate of +20 dB/decade (+ 1 slope) toward lower frequencies and yields the higher gain at 120 Hz. Going in the direction of higher frequency, the −1 gain slope breaks and becomes horizontal at a frequency of $F_z = 1/2\pi R2C1$.

Now going to the right of F_{co} toward higher frequency (Fig. 12.6), if the error-amplifier gain curve were permitted to remain horizontal, total open-loop gain would remain relatively high at the high frequencies. But high gain at high frequencies is undesirable as thin, high-frequency noise spikes would be picked up and transmitted at large amplitudes to the output. Thus gain should be permitted to fall off at high frequencies.

This is easily done by placing a capacitor $C2$ across the series combination of $R2$ and $C1$ (Fig. 12.7*b*). At F_{co}, X_{c1} is already small compared to $R2$ and $C1$ is effectively out of the circuit.

At higher frequencies where X_{c2} is small compared to $R2$, however, $R2$ is effectively out of the circuit and gain is X_{c2}/R_1. Now the gain characteristic beyond F_{co} is horizontal up to a frequency F_p ($=1/2\pi R2C2$), where it breaks and thereafter falls at a −1 slope as can be seen in Fig. 12.6. This lower gain at high frequency keeps high-frequency noise spikes from coming through to the output.

Now the break frequencies F_z and F_p must be chosen. They will be chosen so that $F_{co}/F_z = F_p/F_{co}$. The farther apart F_z and F_p are, the greater the phase margin at F_{co}. Large phase margins are desirable, but if F_z is chosen too low, low-frequency gain will be lower at 120 Hz than if a higher frequency were chosen (Fig. 12.8). Thus 120-Hz attenuation will be poorer. If F_p is chosen too high, gain at high frequencies is higher than if a lower F_p is chosen (Fig. 12.8). Thus high-frequency noise spikes would come through at a higher amplitude.

Thus a compromise between separating F_z and F_p by a large amount to increase phase margin and decreasing the separation to achieve better 120-Hz attenuation and lower-amplitude high-frequency noise spikes is sought.

This compromise and a more exact analysis of the problem is made easy by introducing the concept of transfer functions, poles, and zeros as shown below.

12.4 Error-Amplifier Transfer Function, Poles, and Zeros

The circuit of an operational amplifier with a complex impedance Z_1 input arm and a complex impedance Z_2 feedback arm is shown in Fig. 12.9. Its gain is Z_2/Z_1. If Z_1 is a pure resistor $R1$ and Z_2 is a pure re-

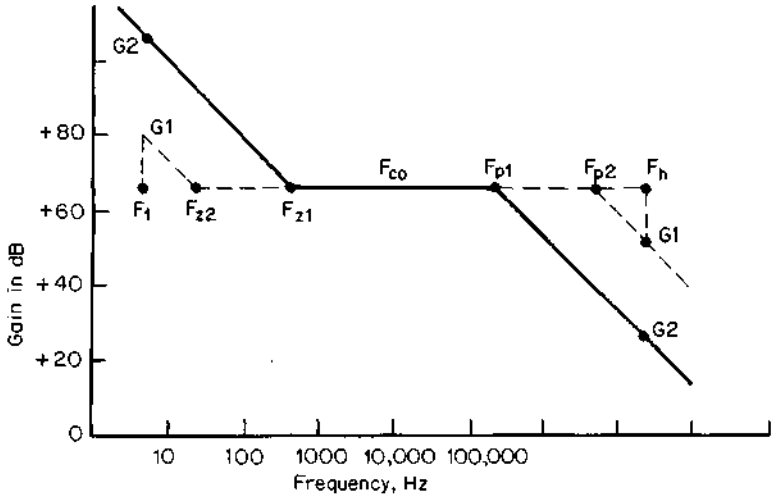


Figure 12.8 Where to locate break frequencies F_z and F_p. The farther apart F_z and F_p are spread, the greater the phase margin. But spreading them further apart reduces low-frequency gain, which reduces the degeneration of low-frequency line ripple. It also increases high gain, which permits high-frequency, thin noise spikes to come through at greater amplitude. If F_z were at F_{z2} instead of F_{z1}, gain at some low frequency F_1 would be G_1 instead G_2. And if F_p were at F_{p2} instead of F_{p1}, gain at some high frequency F_h would be G_1 instead of G_2.

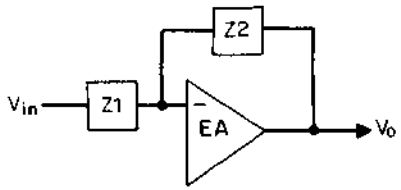


Figure 12.9 If inputs and feedback arms are various combinations of *R*s and *C*s, various gain-versus-frequency and phase-shift-versus-frequency curves are possible. By expressing impedances Z_1 and Z_2 in terms of an s operator ($= jw$), and performing a number of algebraic manipulations, a simplified expression for the gain arises. From this simplified gain expression (transfer function) the gain-versus-frequency and phase-shift-versus-frequency curves can be drawn at a glance.

sistor $R2$ as in Fig. 12.7*a*, gain is R_2/R_1 and is independent of frequency. Phase shift between V_o and V_{in} is 180°, as the input is to the inverting terminal.

Now impedances in Z_1, Z_2 are expressed in terms of the complex variable $s = j(2\pi f) = jw$. Thus the impedance of a capacitor $C1$ is then $1/sC1$ and that of a resistor $R1$ and capacitor $C1$ in series is $(R1 + 1/sC1)$.

The impedance of an arm consisting of a capacitor $C2$ in parallel with a series combination of $R1$ and $C1$ is then

$$Z = \frac{(r + 1/sC1)(1/sC2)}{r + 1/\text{sC1} + 1/\text{sC2}} \tag{12.2}$$

Now the gain or transfer function of the operational error amplifier is written in terms of its Z_1, Z_2 impedances, which are expressed in terms of the complex variable s. Thus $G(s) = Z_2(s)/Z_1(s)$, and by algebraic manipulation, $G(s)$ is broken down into a simplified numerator and denominator which are functions of s: $G(s) = N(s)/D(s)$. The numerator and denominator, again by algebraic manipulation, are factored and $N(s)$, $G(s)$ are expressed in terms of these factors. Thus

$$G(s) = \frac{N(s)}{D(s)} = \frac{(1 + sz_1)(1 + sz_2)(1 + sz_3)}{sp_0(1 + sp_1)(1 + sp_2)(1 + sp_3)} \tag{12.3}$$

These z and p values are RC products and represent frequencies. These frequencies are obtained by setting the factors equal to zero. Thus

$$1 + sz_1 = 1 + s(j2\pi fz_1) = 1 + j2\pi fR1C1 = 0 \qquad \text{or} \qquad f_1 = 1/2\pi R1C1$$

The frequencies corresponding to the z values are called *zero* frequencies, and those corresponding to the p values are called *pole* frequencies. There is always a factor in the denominator which has the "1" missing (note sp_0 above). This represents an important pole frequency, $F_{po} = 1/2\pi R_oC_o$, which is called *the pole at the origin.*

From the location of the pole at the origin and the zero and pole frequencies, the gain-versus-frequency characteristic of the error amplifier can be drawn as discussed below.

12.5 Rules for Gain Slope Changes Due to Zero and Pole Frequencies

The zero and pole frequencies represent points where the error-amplifier gain slope changes.

A zero represents a +1 change in gain slope. Thus (Fig. 12.10*a*), if a zero appears at a point in frequency where the gain slope is zero, it turns the gain into a +1 slope. If it appears where the original gain slope is −1 (Fig. 12.10*b*), it turns the gain slope to zero. Or if there are two zeros at the same frequency (two factors in the numerator of Eq. 12.3 having the same RC product) where the original gain slope is −1,

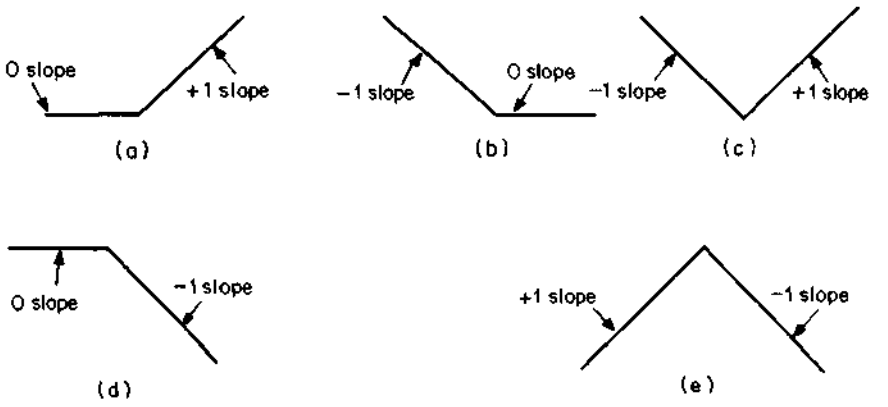


Figure 12.10 (*a*) A zero appearing on a gain curve where the original slope was horizontal turns that gain slope to +1 or +20 dB/decade. (*b*) A zero appearing on a gain curve where the original slope was −1, terms that gain slope horizontal. (*c*) Two zeros at the same frequency appearing on a gain curve where the original slope was −1, turns that slope to +1. The first zero turns the −1 slope horizontal; the second zero at the same frequency turns the horizontal slope to +1. (*d*) A pole appearing on a gain curve where the original slope is horizontal turns that slope to −1 or 20 dB/decade. (*e*) Two poles at the same frequency appearing on a gain curve where the original slope is +1 turns that slope to −1. The first pole turns the original +1 slope horizontal; the second one at the same frequency turns the horizontal slope to −1 or −20 dB/decade.

the first zero turns the gain slope horizontal, and the second zero at the same frequency turns the gain into a +1 slope (Fig. 12.10*c*).

A pole represents a −1 change in gain slope. If it appears in frequency where the original gain slope is horizontal (slope is zero), it turns the gain into a −1 slope (Fig. 12.10*d*). Or if there are two poles at the same frequency at a point where the original gain slope is +1, the first pole turns the slope horizontal and the second at that same frequency turns the slope to −1 (Fig. 12.10*e*).

The pole at the origin, as does any pole, represents a gain slope of −1. It also indicates the frequency at which the gain is 1 or 0 dB. Thus, drawing the total gain curve for the error amplifier starts with the pole at the origin as follows. Go to 0 dB at the frequency of the pole at the origin $F_{po} = 1/2\pi R_o C_o$. At F_{po}, draw a line backward in frequency with a slope of +1 (Fig. 12.11). Now if somewhere on this line, which has a −1 slope in the direction of higher frequency, the transfer function has a zero at a frequency $F_z = 1/2\pi R1C1$, turn the gain slope horizontal at F_z. Extend the horizontal gain indefinitely. But if at some higher frequency there is a pole in the transfer function at a frequency $F_p = 1/2\pi R2C2$, turn the horizontal slope into a −1 slope at F_p (Fig. 12.11).

The gain along the horizontal part of the transfer function is $R2/R1$

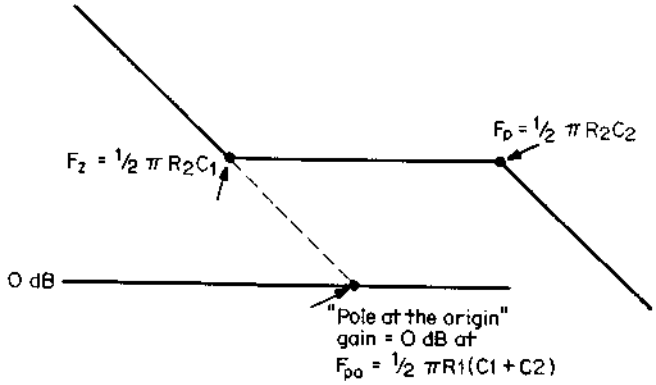


Figure 12.11 Drawing the gain curve of the error amplifier of Fig. 12.7*b* directly from its transfer function of Eq. 12.3.

and is made (in decibels) equal and opposite to the loss of the G_f curve (Fig. 12.6) at F_{co}.

Thus an error-amplifier gain curve having a single pole at the origin, a single zero, and another single pole has the desired shape shown in Fig. 12.11. It is implemented by the circuit of Fig. 12.7*b*. It remains only to select the location of the zero and pole frequencies which yield the desired phase margin. This will be discussed below.

12.6 Derivation of Transfer Function of an Error Amplifier with Single Zero and Single Pole from Its Schematic

It has been shown above that if an error amplifier had a single zero, a single pole, and a pole at the origin, its gain-versus-frequency curve would be as in Fig. 12.11.

Now it will be demonstrated how the transfer function of an error amplifier is derived and that the circuit of Fig. 12.7*b* does truly have a single zero, a single pole, and a pole at the origin. Gain of the circuit in Fig. 12.7*b* is

$$G = \frac{dV_o}{dV_i}$$

$$= \frac{Z_2}{Z_1}$$

$$= \frac{(R2 + 1/jwC1)(1/jwC2)}{R1(R2 + 1/jwC1 + 1/jwC2)}$$

Now introduce the complex variable $s = jw$. Then

$$G = \frac{(R2 + 1/sC1)(1/sC2)}{R1(R2 + 1/sC1 + 1/sC2)}$$

And by algebraic manipulation

$$G = \frac{1 + sR2C1}{sR1(C1 + C2)(1 + sR2C1C2/C1 + C2)}$$

And since generally $C2 << C1$

$$G = \frac{1 + sR2C1}{sR1(C1 + C2)(1 + R2C2)} \tag{12.4}$$

The error amplifier of Fig. 12.7*b*, having the transfer function of Eq. 12.4, is commonly referred to as a *Type 2 amplifier* in conformance with the designation introduced by Venable in his classic paper.[1] A Type 2 error amplifier is used when the output filter capacitor has an ESR so that F_{co} lies on a -1 slope of the G_1 curve (Fig. 12.6).

Examination of this transfer function for the circuit of Fig. 12.7*b* permits immediate drawing of its gain characteristic as follows (Fig. 12.11). Equation 12.4 shows that this circuit (Fig. 12.7*b*) has a pole at the origin at a frequency of $F_{po} = 1/2\pi R1(C1 + C2)$. Thus, go to 0 dB at that frequency and draw a line backward toward lower frequency with a slope of $+1$.

From Eq. 12.4, the circuit has zero at a frequency of $F_z = 1/2\pi R2C1$. Go to that sloped line and turn it horizontal at F_z. Again from Eq. 12.4, the circuit has a pole at a frequency of $F_p = 1/2\pi R2C2$. Now go out along that horizontal line and turn it into a -1 slope at F_p.

Now that the transfer function of the Type 2 error amplifier can be drawn from its pole and zero frequencies, it remains to be able to locate them (choosing $R1$, $R2$, $C1$, $C2$) so as to achieve the desired phase margin. This is demonstrated below.

12.7 Calculation of Type 2 Error-Amplifier Phase Shift from Its Zero and Pole Locations

Adopting Venable's scheme,[1] the ratio $F_{co}/F_z = K$ will be chosen equal to $F_p/F_{co} = K$.

Now a zero, like an RC differentiator (Fig. 12.2*b*), causes a phase lead. A pole, like an RC integrator (Fig. 12.2*a*), causes a phase lag.

The phase lead at a frequency F due to a zero at a frequency F_z is

$$\Theta_{ld} = \tan^{-1}\frac{F}{F_z}$$

But we are interested in the phase lead at F_{co} due to a zero at a frequency F_z. This is

$$\Theta_{ld}\,(\text{at } F_{co}) = \tan^{-1} K \tag{12.5}$$

The phase lag at a frequency F due to a pole at a frequency F_p is

$$\Theta_{lag} = \tan^{-1} \frac{F}{F_p}$$

and we are interested in the lag at F_{co} due to the pole at F_p. This is

$$\Theta_{lag}\,(\text{at } F_{co}) = \tan^{-1} \frac{1}{K} \tag{12.6}$$

The total phase shift at F_{co} due to the lead of the zero at F_z and the lag due to the pole at F_p is the sum of Eqs. 12.5 and 12.6.

These shifts are in addition to the inherent low-frequency phase shift of the error amplifier with its pole at the origin. The error amplifier is an inverter and at low frequency causes a 180° phase shift.

At low frequencies, the pole at the origin causes a 90° phase shift. This is just another way of saying that at low frequencies, the circuit is just an integrator with resistor input and capacitor feedback. This is seen from Fig. 12.7*b*. At low frequencies, the impedance of $C1$ is much greater than $R2$. The feedback arm is thus only $C1$ and $C2$ in parallel.

Thus the inherent low-frequency phase lag is 180° because of the phase inversion plus 90° due to the pole at the origin or a total lag of 270°. Total phase lag, including the lead due to the zero and lag due to the pole, is then

$$\Theta_{(\text{total lag})} = 270° - \tan^{-1} K + \tan^{-1} \frac{1}{K} \tag{12.7}$$

Note that this is still a net phase lag as when K is a very large number (zero and pole frequencies very far apart), the lead due to the zero is a maximum of 90° and the lag due to the pole is 0°.

Total phase lag through the error amplifier, calculated from Eq. 12.7 is shown in Table 12.1.

12.8 Phase Shift through *LC* Filter Having ESR in Its Output Capacitor

The total open-loop phase shift consists of that through the error amplifier plus that through the output *LC* filter. Figure 12.3*b* showed for $R_o = 20\sqrt{L_o/C_o}$ and no ESR in the filter capacitor, the lag through the filter itself is already 175° at $1.2F_o$.

This lag is modified significantly if the output capacitor has an ESR

TABLE 12.1 Phase Lag through a Type 2 Error Amplifier for Various Values of K(= $F_{co}/F_z = F_p/F_{co}$)

K	Lag (from Eq. 12.7)
2	233°
3	216°
4	208°
5	202°
6	198°
10	191°

as in Fig. 12.5*b*. In that figure, the gain slope breaks from a −2 to a −1 slope at the so-called ESR zero frequency of $F_{esr} = 1/2\pi R_{esr}C_o$. Recall that at F_{esr}, the impedance of C_o equals that of R_{esr}. Beyond F_{esr}, the impedance of C_o becomes smaller than R_{esr} and the circuit becomes increasingly like an *LR* rather than an *LC* circuit. Moreover, an *LR* circuit can cause only a 90° phase lag as compared to the possible maximum of 180° for an *LC* circuit.

Thus the ESR zero creates a boost in phase over a possible maximum of 180°. Phase lag at a frequency F due to an ESR zero at F_{esro} is

$$\Theta_{lc} = 180° - \tan^{-1}\frac{F}{F_{esro}}$$

and since we are interested in the phase lag at F_{co} due to the zero at F_{esro}

$$\Theta_{lc} = 180° - \tan^{-1}\frac{F_{co}}{F_{esro}} \tag{12.8}$$

Phase lags through the *LC* filter (having an ESR zero) are shown in Table 12.2 for various values of F_{co}/F_{esro} (from Eq. 12.8).

TABLE 12.2 Phase Lag through an *LC* Filter at F_{co} Due to a Zero at F_{esro}

F_{co}/F_{esro}	Phase lag	F_{co}/F_{esro}	Phase lag
0.25	166°	2.5	112°
0.50	153°	3	108°
0.75	143°	4	104°
1.0	135°	5	101°
1.2	130°	6	99.5°
1.4	126°	7	98.1°
1.6	122°	8	97.1°
1.8	119°	9	96.3°
2.0	116°	10	95.7°

Thus, by setting the error-amplifier gain in the horizontal part of its gain curve (Fig. 12.6) equal and opposite to the G_t (Fig. 12.6) loss at F_{co}, the location of F_{co} is fixed where it is desired. Since F_{co} will in most cases be located on the -1 slope of the G_t curve, the total gain curve will come through F_{co} at a -1 slope. From Tables 12.1 and 12.2, the proper value of K (locations of the zero and pole) is established to yield the desired phase margin.

12.9 Design Example—Stabilizing a Forward Converter Feedback Loop with a Type 2 Error Amplifier

The design example presented below demonstrates how much of the material discussed in all previous chapters is interrelated.

Stabilize the feedback loop for a forward converter with the following specifications:

V_o	5.0 V
$I_{o(nom)}$	10 A
Minimum I_o	1 A
Switching frequency	100 kHz
Minimum output ripple (peak to peak)	50 mV

It is assumed that the filter output capacitor has an ESR and F_{co} will occur on the -1 slope of the LC filter. This permits the use of a Type 2 error amplifier with the gain characteristics of Fig. 12.6. The circuit is shown in Fig. 12.12.

First L_o, C_o will be calculated and the gain characteristic of the output filter will be drawn. From Eq. 2.47

$$L_o = \frac{3V_oT}{I_{on}}$$

$$= \frac{3 \times 5 \times 10^{-6}}{10}$$

$$= 15 \times 10^{-6}\ \text{H}$$

and from Eq. 2.48

$$C_o = 65 \times 10^{-6} \frac{dI}{V_{or}}$$

where dI is twice the minimum output current = $2 \times 1 = 2$ A and V_{or} is the output ripple voltage = 0.05 V. Then $C_o = 65 \times 10^{-6} \times 2/0.05 =$ 2600 microfarads.

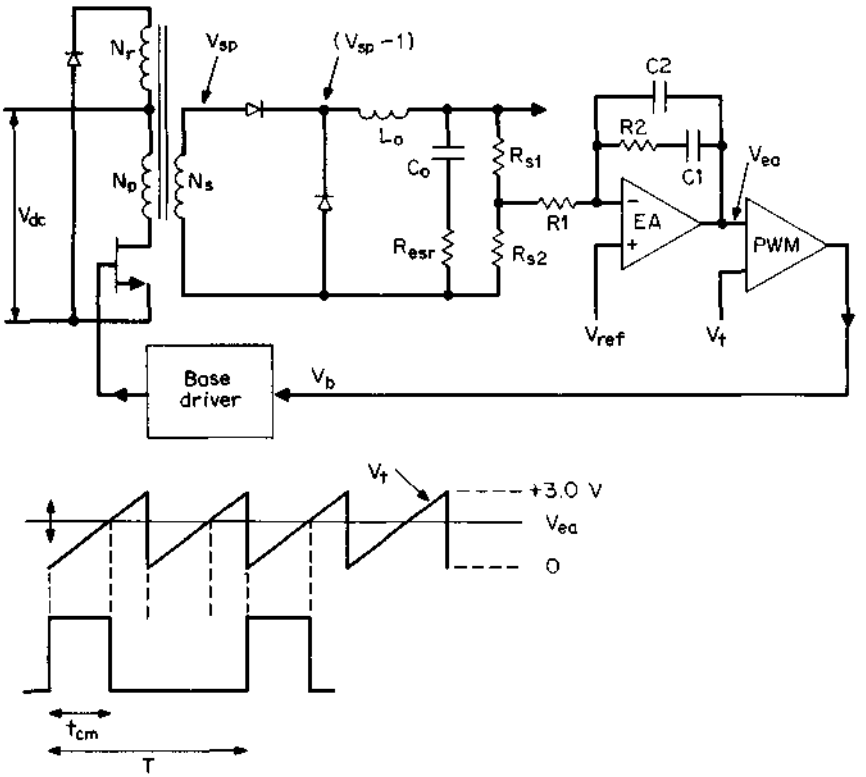


Figure 12.12 Forward converter design example schematic stabilizing the feedback loop.

Corner frequency of the output LC filter, from Sec. 12.2.3, is

$$F_o = 1/2\pi\sqrt{L_oC_o}$$
$$= 1/2\pi\sqrt{15 \times 10^{-6} \times 2600 \times 10^{-6}}$$
$$= 806 \text{ Hz}$$

Again from Sec. 12.2.3, the frequency of the ESR zero (frequency where the gain slope abruptly changes from a -2 to a -1 slope) is

$$F_{esr} = 1/2\pi R_{esr}C_o$$
$$= 1/2\pi(65 \times 10-6)$$ (assuming, as in Sec. 2.3.11.2, that over a large range of aluminum electrolytic capacitor magnitudes and voltage ratings, that $R_{esr}C_o$ is constant and equal to $65 \times 10-6$)

$$= 2500 \text{ Hz}.$$

From Eq. 12.1, the modulator gain G_m is $G_m = 0.5(V_{sp} - 1)/3$, and when the duty cycle is 0.5, for $V_o = 5V$, $V_{sp} = 11$ V since $V_o = (V_{sp} - 1)T_{on}/T$. Then $G_m = 0.5(11 - 1)/3 = 1.67 = +4.5$ dB.

For the usual SG1524-type PWM chip, which can tolerate only 2.5 V at the reference input to the error amplifier, for $V_o = 5$ V, $R_{s1} = R_{s2}$. Sampling network gain (loss) is then $G_s = -6$ dB. Then $G_m + G_s = +4.5 - 6.0 = -1.5$ dB.

The open-loop gain curve of everything but the error amplifier is then $G_t = G_{lc} + G_m + G_s$ and is drawn in Fig. 12.13 as curve *ABCD*. From *A* to the corner frequency at 806 Hz (*B*) it has a value of $G_m + G_s = -1.5$ dB. At *B*, it breaks into a −2 slope and continues at that slope up to the ESR zero at 2500 Hz (*C*). At point *C*, it breaks into a −1 slope.

Now crossover frequency is taken at one-fifth the switching frequency or 20 kHz. From the G_t curve, loss at 20 kHz is −40 dB (numerical loss of 1/100). Hence, to make 20 kHz the crossover frequency, the error-amplifier gain at that frequency is made +40 dB. Since the total open-loop gain of the error amplifier plus curve *ABCD*) must come through crossover at a −1 slope, the error-amplifier gain curve must have zero slope between points *F* and *G* in curve *EFGH* of Fig. 12.13 since *ABCD* already has a −1 slope at 20 kHz.

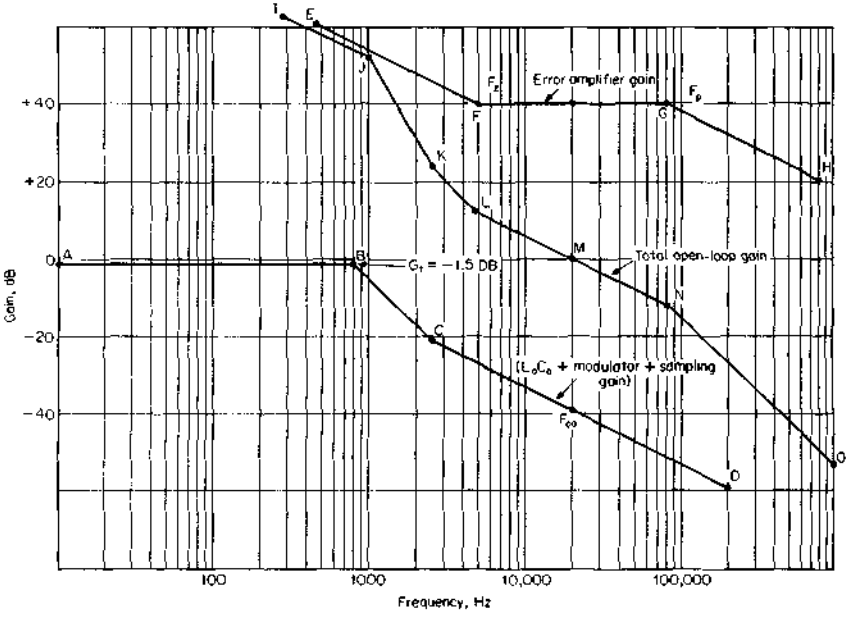


Figure 12.13 Design example stabilizing the feedback loop for Fig. 12.12.

This horizontal gain slope between points F and G is obtained as described above with a Type 2 error amplifier. The gain of the Type 2 error amplifier in the horizontal part of its slope is $R2/R1$. If $R1$ is arbitrarily taken as 1000 Ω, R_2 is 100,000 Ω.

Now a zero is located at F_z to increase low-frequency gain to degenerate 120-Hz line ripple and a pole is located at G to commence decreasing high-frequency gain so as to minimize thin noise spikes at the output. The zero and the pole will be located to give the desired phase margin.

Assume a 45° phase margin. Then total phase shift around the loop at 20 kHz is 360 − 45 = 315°. But the LC filter by itself causes a phase lag given by Eq. 12.7. From that equation, the lag for F_{co} = 20 kHz and F_{esro} = 2500 Hz is 97° (Table 12.2). Thus the error amplifier is permitted only 315 − 97 or 218°. Table 12.1 shows that for an error-amplifier lag of 218°, a K factor of slightly less than 3 would suffice.

To provide somewhat more insurance, assume a K factor of 4 which yields a phase lag of 208°. This, plus the 97° lag of the LC filter, yields a total lag of 305° and a phase margin of 360 − 305° or 55° at F_{co}.

For a K factor of 4, the zero is at F_z = 20/4 = 5 kHz. From Eq. 12.3, $F_z = 1/2\pi R2C1$. For $R2$ determined above as $100K$, $C1 = 1/2\pi(100{,}000)(5000) = 318 \times 10^{-12}$.

Again for the K factor of 4, the pole is at $F_{po} = 20 \times 4 = 80$ kHz. From Eq. 12.3, $F_{po} = 1/2\pi R2C2$. For $R2 = 100K$, $F_{po} = 80$ kHz, $C2 = 1/2\pi(100{,}000)(80{,}000) = 20 \times 10^{-12}$. This completes the design; the final gain curves are shown in Fig. 12.13. Curve $IJKLMO$ is the total open-loop gain. It is the sum of curves $ABCD$ and $EFGH$.

12.10 Type 3 Error Amplifier—When Used and Transfer Function

In Sec. 2.3.11.2, it was pointed out that the output ripple $V_{or} = R_o\, dI$ where R_o is the ESR of the filter output capacitor C_o and dI is twice the minimum DC current. Now most aluminum electrolytic capacitors do have an ESR. Study of many capacitor manufacturers' catalogs indicates that for such capacitors, R_oC_o is constant and equal to an average value of 65×10^{-6}.

Thus, using conventional aluminum electrolytic capacitors, the only way to reduce output ripple is to decrease R_o, which can be done only by increasing C_o. This, of course, increases size of the capacitor, which may be unacceptable.

Within the past few years, capacitor manufacturers have been able (at considerably greater cost) to produce aluminum electrolytic capacitors with essentially zero ESR for those applications where output ripple must be reduced to an absolute minimum.

When such zero ESR capacitors are used, it affects the design of the error amplifier in the feedback loop significantly. When an ESR was present in the output capacitor, F_{co} usually was located on the −1 slope of the output filter. This required a Type 2 error amplifier with a horizontal slope at F_{co} in its gain-versus-frequency characteristic (Fig. 12.6).

With a capacitor with zero ESR, the *LC* gain-versus-frequency curve continues falling at a −2 slope after the corner frequency $F_o = 1/2\pi\sqrt{L_oC_o}$ (curve *ABCD* in Fig. 12.14*a*). Now an error amplifier can be designed to have its gain equal and opposite to the *LC* loss at the desired F_{co}. But for the total gain to come through F_{co} at a −1 slope, the error-amplifier gain curve must be designed to have a +1 slope in its central region at F_{co} (curve *EFGHI* in Fig. 12.14*b*).

Now the error-amplifier gain cannot be permitted to fall in the direction of lower frequencies. If it did, it would not provide enough gain at 120 Hz to degenerate line frequency input ripple sufficiently.

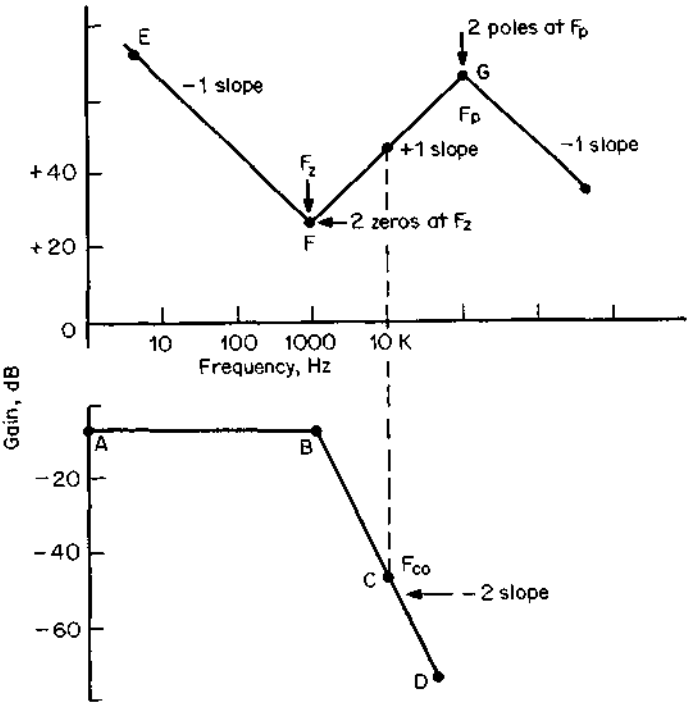


Figure 12.14 When the output capacitor has no ESR, its gain continues to fall at a −2 slope. This requires an error amplifier with a +1 slope at F_{co} for the total open-loop gain to come through F_{co} at a −1 slope. To achieve the above error-amplifier gain curve, two zeros are located at F_z and two poles at F_p.

Thus at some point F_z (Fig. 12.14*b*), the gain curve must be turned around to proceed upward at a +1 slope in the direction of lower frequencies. As described in Sec. 12.5, this is done by providing two zeros at the same frequency (F_z) in the error amplifier's transfer function. Below F_z, the gain falls at a −1 slope (in the direction of higher frequencies) because of a pole at the origin which will be provided. At F_z, the first zero turns the gain slope horizontal; the second one turns it to a +1 slope.

Now the gain cannot be permitted to continue upward at a +1 slope much beyond F_{co}. If it did, gain would be high at high frequencies and thin noise spikes would get through to the output at high amplitudes. Thus, as described in Sec. 12.5, two poles are provided at point H at the frequency F_p. The first pole turns the +1 gain slope horizontal; the second pole turns it to a −1 slope.

An error amplifier with the gain-versus-frequency *EFGHI* in Fig. 12.14 is referred to as a *Type 3 error amplifier* (again following the widely used Venable designation.[1]

As for the Type 2 error amplifier, location of the two zeros at F_z and the two poles at F_p determines the phase lag at F_{co}. The wider the separation between F_p and F_z, the greater the phase margin.

Also with respect to the Type 2 error amplifier, shifting F_z down to too low a frequency reduces low-frequency gain and prevents sufficient degeneration of 120-Hz line ripple. Moving F_p to too high a frequency offers higher gain at high frequencies and permits high-frequency, thin noise spikes to come through at greater amplitude.

Again a K factor is introduced to describe locations of F_z and F_p. This ratio is set equal to $K = F_{co}/F_z$ and $K = F_p/F_{co}$. In the following section, phase boost at F_{co} due to the double zero at F_z and phase lag at F_{co} due to the double pole at F_p will be calculated.

12.11 Phase Lag through a Type 3 Error Amplifier as Function of Zero and Pole Locations

In Sec. 12.7, it was pointed out that the phase boost at a frequency F_{co} due to a zero at a frequency F_z is $\Theta_{zb} = \tan^{-1}(F_{co}/F_z) = \tan^{-1} K$ (Eq. 12.4). If there are two zeros at the frequency F_z, the boosts are additive. Thus boost at F_{co} due to two zeros at the same frequency F_z is $\Theta_{2zb} = 2 \tan^{-1} K$.

Similarly, the lag at F_{co} due to a pole at F_p is $\Theta_{lp} = \tan^{-1}(1/K)$ (Eq. 12.5). The lags due to two poles at F_p are also additive. Thus lag at F_{co} due to two poles at F_p is $\Theta_{l2p} = 2 \tan^{-1}(1/K)$. The lag and boost are in addition to the inherent low-frequency 270° lag due to the 180° phase inversion plus the 90° due to the inherent pole at the origin.

Thus total phase lag through a Type 3 error amplifier is

$$\Theta_{tl} = 270° - 2\tan^{-1}K + 2\tan^{-1}(1/K) \tag{12.9}$$

Total phase lag through the Type 3 error amplifier is calculated from Eq. 12.9 for various values of K (see Table 12.3).

Comparing Tables 12.3 and 12.1, it is seen that a Type 3 error amplifier with two zeros and two poles yields considerably less phase lag than does the Type 2 error amplifier, which has only a single zero and a single pole.

However, the Type 3 error amplifier is used with an LC filter which has no ESR zero to decrease the lag down from 180°. Thus the lower lag of the Type 3 error amplifier is essential because of the higher lag of an LC filter with no ESR.

12.12 Type 3 Error Amplifier Schematic, Transfer Function, and Zero and Pole Locations

The schematic of a circuit which has the gain-versus-frequency characteristic of Fig. 12.14*b* is shown in Fig. 12.15. Its transfer function can be derived in the manner described in Sec. 12.6 for the Type 2 error amplifier. Impedances of the feedback and input arm are expressed in terms of the s operator, and the transfer function is simply $G(s) = Z_2(s)/Z_1(s)$. Algebraic manipulation yields the following expression for the transfer function:

$$G(s) = \frac{dV_o}{dV_{in}} = \frac{(1 + sR2C1)[1 + s(R1 + R3)C3]}{sR1(C1 + C2)(1 + sR3C3)[1 + sR2(C1C2/(C1 + C2)]} \tag{12.10}$$

This transfer function is seen to have

(*a*) A pole at the origin at a frequency of

$$F_{po} = 1/2\pi R1(C1 + C2) \tag{12.11}$$

This is the frequency where the impedance of $R1$ is equal to that of $(C1 + C2)$ in parallel.

TABLE 12.3 Phase Lag through Type 3 Error Amplifier for Various Values of $K(= F_{co}/F_z = F_p/F_{co})$

K	Lag (from Eq. 12.9)
2	196°
3	164°
4	146°
5	136°
6	128°

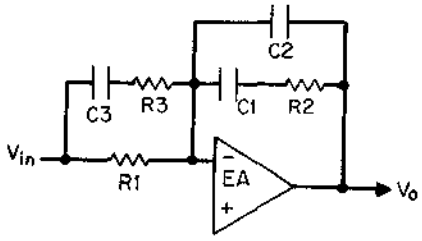


Figure 12.15 A Type 3 error amplifier. It has a pole at the origin, two zeros, and two poles. Its transfer function is

$$G = dV_o/dV_{in}$$

$$= \frac{(1 + sR2C1)[1 + s(R1 + R3)C3]}{sR1(C1 + C2)(1 + sR3C3)[1 + sR2(C1C2/(C1 + C2)]}$$

(*b*) A first zero at a frequency of

$$F_{z1} = 1/2\pi R2C1 \tag{12.12}$$

This is the frequency where the impedance of $R2$ equals that of $C1$.

(*c*) A second zero at a frequency of

$$F_{z2} = 1/2\pi(R1 + R3)C3$$

$$\approx 1/2\pi R1C3 \tag{12.13}$$

This is the frequency where the impedance of $(R_1 + R_3)$ equals that of C_3.

(*d*) A first pole at a frequency of

$$F_{p1} = \frac{1/2\pi R2(C1C2)}{C1 + C2}$$

$$\approx 1/2\pi R2C2 \tag{12.14}$$

This is the frequency where the impedance of $R2$ equals that of the series combination of $C1$ and $C2$ in series.

(*e*) A second pole at a frequency of

$$F_{p2} = 1/2\pi R3C3 \tag{12.15}$$

This is the frequency where the impedance of $R3$ equals that of $C3$.

To yield the gain-versus-frequency curves of Fig. 12.14*b*, the RC products will be chosen so that $F_{z1} = F_{z2}$ and $F_{p1} = F_{p2}$. The location of the double-zero and double-pole frequencies will be fixed by the K factor, which yields the desired phase margin. Gain of the error am-

plifier on the +1 slope of Fig. 12.14*b* will be set equal and opposite to the loss of the *LC* filter (Fig. 12.14*a*) at the desired F_{co}.

From Table 12.3 and the transfer function of Eq. 12.10, the *RC* products which set the zero and pole frequencies at the desired points are determined as in the design example below.

12.13 Design Example—Stabilizing a Forward Converter Feedback Loop with a Type 3 Error Amplifier

Design the feedback loop for a forward converter having the following specifications:

V_o	5.0 V
$I_{o(\text{nom})}$	10 A
$I_{o(\text{min})}$	1.0 A
Switching frequency	50 kHz
Output ripple (peak to peak)	< 20 mV

Assume that the output capacitor is of the type advertised as having zero ESR.

First the output *LC* filter and its corner frequency are calculated. Refer to Fig. 12.15. From Eq. 2.47

$$L_o = \frac{3V_oT}{I_o}$$

$$= \frac{3 \times 5 \times 20 \times 10^{-6}}{10}$$

$$= 30 \times 10^{-6}\ \text{H}$$

Now it was assumed that the output capacitor had zero ESR so that ripple due to ESR should be zero. But there is a small capacitive ripple component (Sec. 1.2.7). This is usually very small, and hence a filter capacitor much smaller than the 2600-μF capacitor used in the Type 2 error-amplifier design example can be used. But to be conservative, for this design assume the same 2600-μF capacitor is used and that it has zero ESR. Then

$$F_o = 1/2\pi\sqrt{L_oC_o}$$

$$= 1/2\pi\sqrt{30 \times 10^{-6} \times 2600 \times 10^{-6}}$$

$$= 570\ \text{Hz}$$

Assume, as for the Type 2 error-amplifier design example, that the modulator plus sampling resistor gain is −1.5 dB. The gain of the *LC* filter plus modulator plus sampling resistor gain is plotted in Fig. 12.16 as curve *ABC*. It is horizontal at a level of −1.5 dB up to the corner frequency of 570 Hz at point *B*. There it changes abruptly to a −2 slope and remains at that slope since the capacitor has no ESR.

Frequency F_{co} is chosen as one-fifth the switching frequency or 50/5 = 10 kHz. On curve *ABC* of Fig. 12.16, loss at 10 kHz is −50 dB. Hence to force 10 kHz to be F_{co}, the error-amplifier gain at 10 kHz is set at +50 dB (point *F* in Fig. 12.16). But the error amplifier must have a +1 slope at F_{co} to yield a net −1 slope when added to the −2 slope of the *LC* filter. Thus, at point *F* draw a line of +1 slope. Extend this in the direction of lower frequencies to F_z—the frequency of the double zero. Extend it in the direction of higher frequencies to F_p, the frequency of the double pole. Then determine F_z and F_p from the *K* factor (Table 12.3) required to yield the desired phase margin.

Assume a phase margin of 45°. Then total phase lag of the error amplifier plus the *LC* filter is 360 − 45 = 315°. But the *LC* filter, not having an ESR zero, has a lag of 180°. This leaves a permissible lag of 315 − 180 = 135° for the error amplifier.

From Table 12.3, a *K* factor of 5 yields a lag of 136°, which is close

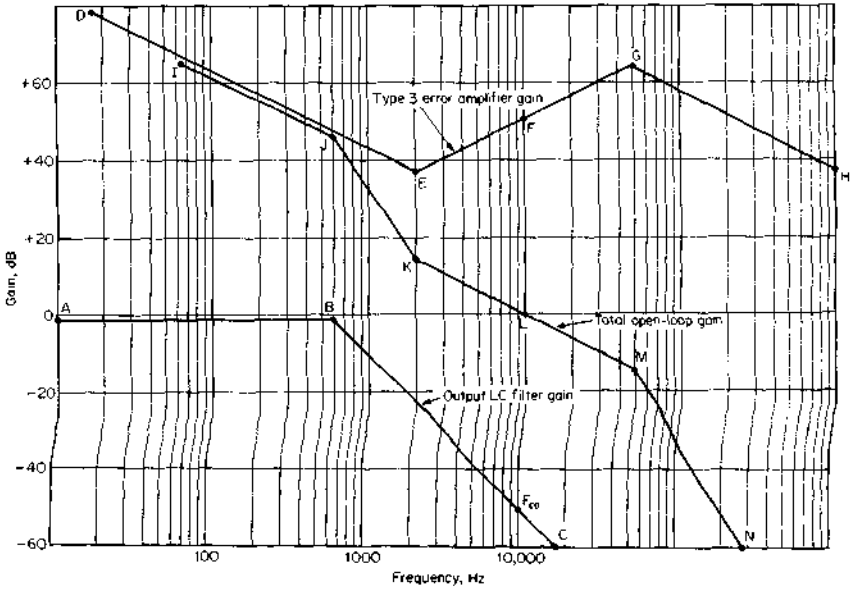


Figure 12.16 Gain curves—design example of Sec. 12.13. Output capacitor with zero ESR and Type 3 error amplifier.

enough. For F_{co} = 10 kHz, K = 5, F_z is 2 kHz and F_p is 50 kHz. Thus in Fig. 12.16, the +1 sloped line is extended down to 2 kHz at E, where it breaks upward to a +1 slope (−1 slope in direction of higher frequencies due to the pole at the origin). It is extended on a +1 slope from F to the double-pole frequency at 50 kHz. There it turns down to a −1 slope because of the two poles.

The curve *IJKLMN* is the total open-loop gain and is the sum of curves *ABC* and *DEFGH*. It is seen to have a gain of 0 dB at 10 kHz (the crossover frequency F_{co}) and to come through F_{co} at a −1 slope. The K factor of 5 yields the required 45° phase margin. Components must now be selected to yield the error-amplifier gain curve *DEFGH* in Fig. 12.16.

12.14 Component Selection to Yield Desired Type 3 Error-Amplifier Gain Curve

There are six components to be selected ($R1$, $R2$, $R3$, $C1$, $C2$, $C3$) and four equations for zero and pole frequencies (Eqs. 12.12 to 12.15).

Arbitrarily choose $R1$ = 1000 Ω. Now the first zero (at 2000 Hz) occurs when $R2 = X_{c1}$ and the impedance of the feedback arm at that frequency is mainly that of $R2$ itself. Thus gain at 2000 Hz is $R2/R1$. From Fig. 12.16, gain of the error amplifier at 2000 Hz is +37 dB or a numerical gain of 70.8. Then for $R1 = 1K$, $R2 = 70.8K$, from Eq. 12.12, we obtain

$$\begin{aligned} C1 &= 1/2\pi R2F_z \\ &= 1/2\pi(70{,}800)2000 \\ &= 0.011\ \mu\text{F} \end{aligned}$$

from Eq. 12.14

$$\begin{aligned} C2 &= 1/2\pi R2F_p \\ &= 1/2\pi(70{,}800)(50{,}000) \\ &= 45\ \text{pF} \end{aligned}$$

from Eq. 12.13

$$\begin{aligned} C3 &= 1/2\pi R1F_z \\ &= 1/2\pi(1000)(2000) \\ &= 0.08\ \mu\text{F} \end{aligned}$$

and finally from Eq. 12.15

$$\begin{aligned} R3 &= 1/2\pi C3F_p \\ &= 1/2\pi(0.08 \times 10^{-6})(50{,}000) \\ &= 40\ \Omega \end{aligned}$$

12.15 Conditional Stability in Feedback Loops

A feedback loop may be stable under normal operating conditions when it is up and running, but can be shocked into continuous oscillation at turnon or by a line input transient. This odd situation, called *conditional stability*, can be understood from Fig. 12.17*a* and 12.17*b*.

Figure 12.17*a*and 12.17*b* contains plots of total open-loop phase shift and total open-loop gain versus frequency, respectively. Conditional stability may arise if there are two frequencies (points *A* and *C*) at which the total open-loop phase shift reaches 360° as in Fig. 12.17*a*.

Recall that the criterion for oscillation is that at the frequency where the total open-loop gain is unity or 0 dB, the total open-loop

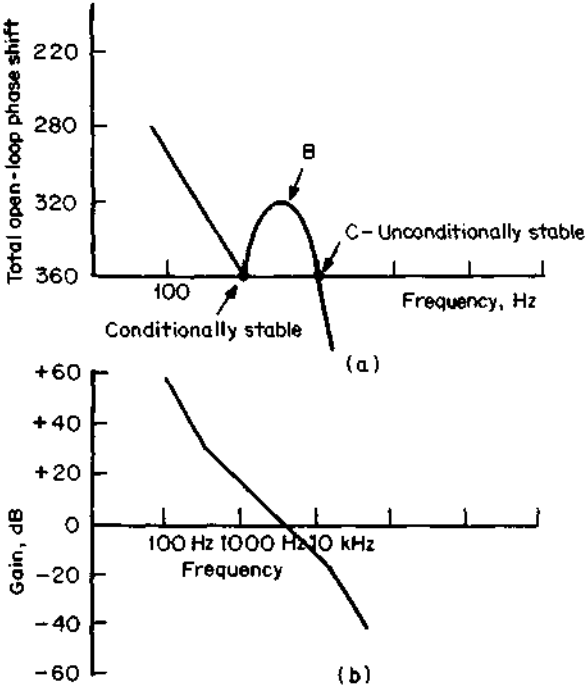


Figure 12.17 A loop may be conditionally stable if there are two frequencies where the total open-loop phase shift is 360°. Loop is conditionally stable at point *A* as a momentary drop in gain to 0 dB such as may occur at initial turnon may result in the conditions for oscillation, i.e., 360° total open-loop phase shift and 0 dB gain. Once oscillation breaks out, it will continue. Circuit is unconditionally stable at *B* as momentary increases in gain are very unlikely.

phase shift is 360°. The loop is still stable if the total open-loop phase shift is 360° at a given frequency but the total open-loop gain at that frequency is greater than 1.

This may be difficult to grasp, as it would appear that if at some frequency the echo of a signal coming around the loop is exactly in phase with the original signal but larger in amplitude, it would grow larger in amplitude each time around the loop. It would thus build up to a level where the losses would be such to limit the oscillation to some high level and remain in oscillation. This does not occur, as can be demonstrated mathematically. But for the purposes herein, it will simply be accepted that oscillations do not occur if the total open-loop gain is greater than unity at the frequency where the total open-loop phase shift is 360°.

Thus in Fig. 12.17*a*, the loop is unconditionally stable at *B* as there the open-loop gain is unity but the open-loop phase shift is less than 360° by about 40°—i.e., there is a phase margin at point *B*. The loop is also stable at point *C* as there the open-loop phase shift is 360° but gain is less than unity—i.e., there is gain margin at point *C*. But at point *A* the loop is conditionally stable. Although the total open-loop phase shift is 360°, the gain is greater than unity (about +16 dB) and, as stated, the loop is stable for those conditions.

However, if under certain conditions—say, at initial turnon when the circuit has not yet come to equilibrium and open-loop gain momentarily drops 16 dB at the frequency of point *A*—the condition for oscillation exists. Gain is unity and phase shift is 360°. The circuit will break into oscillation and remain oscillatory. Point *C* is not a likely location for such conditional oscillation as it is not possible for gain to increase momentarily.

If conditional stability exists (most likely at initial turnon), it is likely to occur at the corner frequency of the output *LC* filter under conditions of light load. It is seen in Fig. 12.3*a* and 12.3*b* that a lightly loaded *LC* filter has a large resonant bump in gain and very fast phase shifts at its corner frequency. The large phase shifts can result in a total of 360° at the *LC* corner frequency. If total open-loop gain (which is not easily predictable during the turnon transient may be unity or may momentarily be unity—the loop may break into oscillation.

It is rather difficult to calculate whether this may occur. The safest way to avoid this possibility is to provide a phase boost at the *LC* corner frequency by introducing a zero there to cancel some of the phase lag in the loop. This can be done easily by adding a capacitor in shunt with the upper resistor in the output voltage sampling network (Fig. 12.12).

12.16 Stabilizing a Discontinuous-Mode Flyback Converter

12.16.1 DC gain from error-amplifier output to output voltage node

The essential elements of the loop are shown in Fig. 12.18*a*. The first step in designing the feedback loop is to calculate its DC or low-frequency gain from the error-amplifier output to the output voltage node. Assume an efficiency of 80 percent. Then from Eq. 4.2a

$$P_o = \frac{0.8(1/2L)(I_p)^2}{T} = \frac{(V_o)^2}{R_o} \tag{12.16}$$

But $I_p = V_{dc}\overline{T_{on}}L_p$; then

$$P_o = \frac{0.8L_p(V_{dc}\overline{T_{on}}/L_p)^2}{2T} = \frac{(V_o)^2}{R_o} \tag{12.17}$$

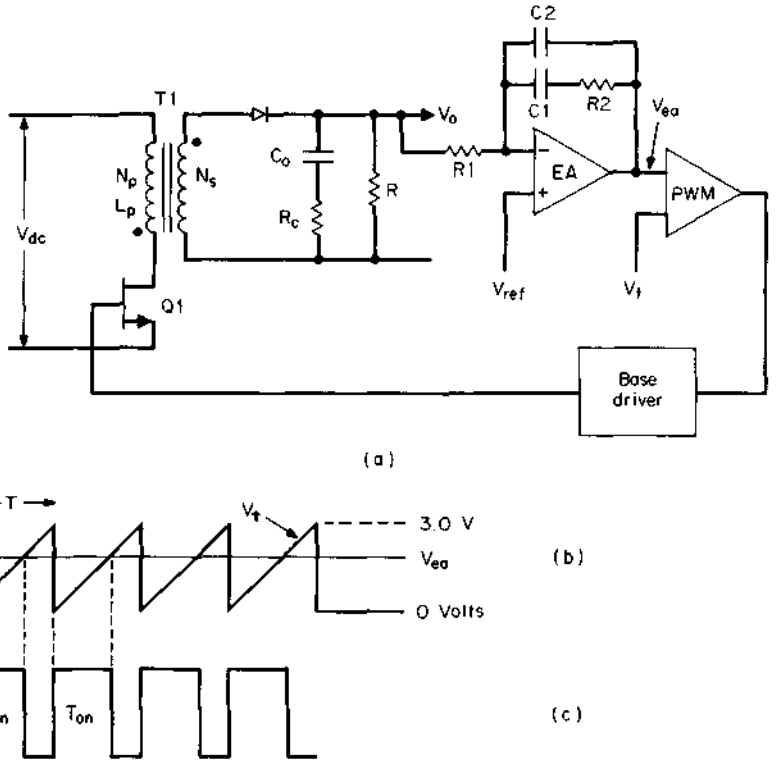


Figure 12.18 Discontinuous-mode flyback feedback loop.

Referring to Fig. 12.18*b*, it is seen that the PWM compares the output of the error amplifier V_{ea} to a 0- to 3-V triangle. It generates a rectangular pulse whose width (T_{on}; Fig. 12.18*c*) is equal to the time from the start of the triangle to its intersection with DC voltage level V_{ea}. This T_{on} will be the on time of power transistor $Q1$. It is seen in Fig. 12.18*b* that $V_{ea}/3 = T_{on}/T$ or $T_{on} = V_{ea}T/3$. Putting this into Eq. 12.16, we have

$$P_o = \frac{0.8L_p(V_{dc}/L_p)^2(V_{ea}T/3)^2}{2T} = \frac{(V_o)^2}{R_o}$$

or

$$V_o = \frac{V_{dc}V_{ea}}{3}\sqrt{\frac{0.4R_oT}{L_p}} \tag{12.18}$$

and the DC or low-frequency gain from the error-amplifier output to the output node is

$$\frac{\Delta V_o}{\Delta V_{ea}} = \frac{V_{dc}}{3}\sqrt{\frac{0.4R_oT}{L_p}} \tag{12.19}$$

12.16.2 Discontinuous-mode flyback transfer function or AC voltage gain from error-amplifier output to output voltage node

Now assume a small sinusoidal signal of frequency f_n inserted in series at the error-amplifier output point. This will cause a sinusoidal modulation in amplitude of the triangular current pulses (of peak amplitude I_p) in the $T1$ primary. Consequently, there is a sinusoidal amplitude modulation in the triangular secondary current pulses (whose instantaneous amplitude is I_pN_p/N_s).

The average value of these triangular secondary current pulses then is modulated at the same sinusoidal frequency f_n. There is thus a sinusoidal current of frequency f_n flowing into the top of the paralleled combination of R_o and C_o.

But this alternating current flows into the Thevenin equivalent of R_o and C_o in series. It is thus seen that the output voltage across C_o falls off in amplitude at the rate of -20 dB/decade or at a -1 slope starting from the frequency of $F_p = 1/2\pi R_oC_o$.

This is simply another way of saying that the transfer function from the error-amplifier output to the output voltage node has a pole at

$$F_p = 1/2\pi R_oC_o \tag{12.20}$$

and DC gain below the pole frequency is given by Eq. 12.19.

This is in contrast to topologies which have an *LC* output filter. In such topologies, a sinusoidal voltage inserted at the error-amplifier output node results in a sinusoidal voltage at the input to the *LC* filter. That voltage, coming through the *LC* filter, falls off in amplitude at −40 dB/decade rate or −2 slope. Or to use the common jargon expression, the *LC* filter has a *two-pole rolloff* at the output node.

This −1 slope or *single-pole rolloff* of the flyback topology output circuit, of course, changes the error-amplifier transfer function required to stabilize the feedback loop. The flyback converter output filter capacitor, in most cases, also has an ESR zero at a frequency of

$$F_z = 1/2\pi R_c C_o \tag{12.21}$$

Now, a complete analysis of the stabilization problem should consider maximum and minimum values of both DC input voltage and of R_o. Equation 12.19 shows DC gain as proportional to V_{dc} and to the square root of R_o. Further, the output circuit pole frequency is inversely proportional to R_o.

Thus in the following graphical analysis, all four combinations of V_{dc} and R_o should be considered as the output circuit transfer function may vary significantly with them.

For one output circuit transfer function (one set of line and load conditions), the error-amplifier transfer function is designed to establish F_{co} at a desired frequency and to have the total gain curve come through F_{co} at a −1 slope. Care must be taken, then, that at another output circuit transfer function (different load and line conditions), the total gain curve does not come through F_{co} at a −2 slope and possibly cause oscillation.

For this example, consider that V_{dc} variations are small enough to be neglected. Thus calculate DC gain from Eq. 12.19 and output circuit pole frequency from Eq. 12.20. Assume $R_{o(max)} = 10R_{o(min)}$.

Now in Fig. 12.19, curve *ABCD* is the output circuit transfer function for $R_{o(max)}$. It has a gain given by Eq. 12.19 from *A* to *B*. At *B*, it breaks into a −1 slope because of the output pole given by Eq. 12.20. At *C*, its slope turns horizontal because of the ESR zero of the output capacitor. Frequency at point *C* is given by Eq. 12.21, whereas in Sec. 1.3.7, R_cC_o is 65×10^{-6} for an aluminum electrolytic capacitor over a large range of voltage and capacitance ratings.

Also in Fig. 12.19, curve *EFGH* is the output circuit transfer function for $R_{o(min)} = R_{o(max)}/10$. Its pole frequency is 10 times that for R_o as F_p is inversely proportional to R_o. DC gain at *F* is 10 dB below that for $R_{o(max)}$ as gain is proportional to the square root of R_o (20 log $\sqrt{10} = 10dB$).

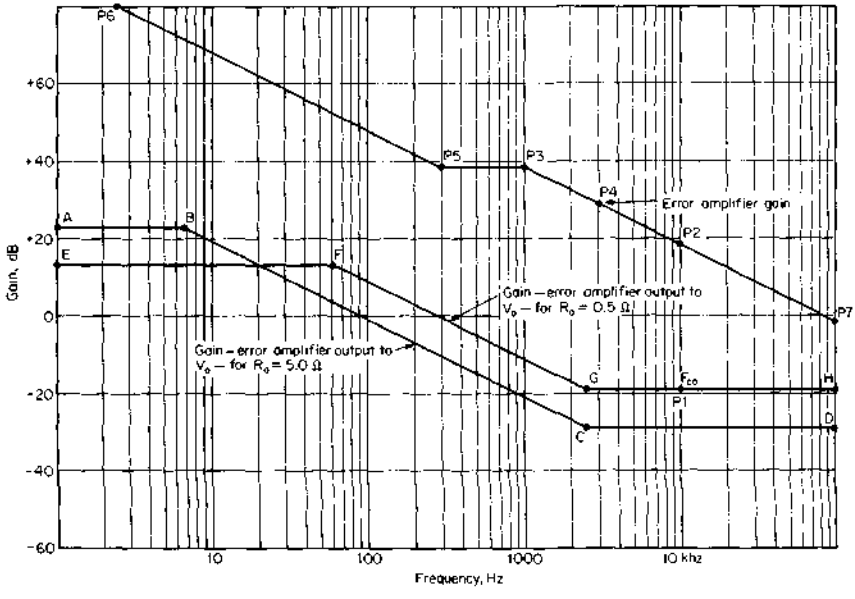


Figure 12.19 Gain curves for stabilizing the feedback loop or the discontinuous-mode flyback of design example in Sec. 12.18.

Thus the output circuit transfer function for $R_{o(\min)}$ is drawn as follows. Go to point F, which is at a frequency 10 times that of point B and 10 dB below point B. Draw a horizontal line back toward DC for the low-frequency gain (line FE). At F, draw a line of -1 slope ($-$ 20 dB/decade) and continue it to the ESR zero frequency at G. At G, draw a line of horizontal slope out toward higher frequency.

From output circuit transfer functions $ABCD$ and $EFGH$ of Fig. 12.19, the error-amplifier gain or transfer function curve is drawn as described below (Sec. 12.17).

12.17 Error-Amplifier Transfer Function for Discontinuous-Mode Flyback

In Fig. 12.19, for $R_{o(\min)}$, on curve $EFGH$, F_{co} will be established at one-fifth the switching frequency (point $P1$) as stated in Sec. 12.3. Most often, F_{co} will occur on the horizontal slope of the output circuit transfer function.

To force F_{co} to be at the desired point, the error amplifier will be designed to have a gain at F_{co} (point $P2$) equal and opposite to the output circuit loss at point $P1$. Since the slope of $EFGH$ at F_{co} is horizon-

tal, the error-amplifier gain slope must be −1 (in the direction of higher frequencies) at point *P*2.

Thus, go to point *P*2 and draw a line backward with a slope of +1 in the direction of lower frequencies. Extend it to a frequency (point *P*3) somewhat lower than the frequency at *C*. For at $R_{o(\max)}$, the output circuit transfer function will be *ABCD*. Since the total gain curve must come through the new F_{co} for $R_{o(\max)}$ at a −1 slope, this new F_{co} will occur at the frequency where the loss along the horizontal line *CD* is equal and opposite (at point *P*4) to the gain of the error amplifier on its −1 slope.

The exact frequency for point *P*3 is not critical. It must be lower than the frequency at *C* to ensure that for the absolutely maximum R_o, where point *C* will be depressed to its greatest loss, this maximum loss can be matched by the equal and opposite gain of the error amplifier somewhere along its −1 slope.

Thus a pole is located at the frequency F_p corresponding to the pole at point *P*3. A Type 2 error amplifier is used. The input resistor *R*1 (Fig. 12.18*a*) is arbitrarily selected sufficiently high so as not to load down the sampling resistor network.

Gain along the horizontal arm (points *P*3–*P*5) is read from the graph and made equal to *R*2/*R*1 (Fig. 12.18*a*). This fixes resistor *R*2. From the pole frequency F_p and *R*2, the value of *C*2 in Fig. 12.18*a* ($= 1/2\pi F_p R2$) is fixed.

Now the gain is extended along the horizontal line *P*3–*P*5 and a zero is introduced at point *P*5 to increase low-frequency gain and offer a phase boost. Frequency of the zero F_z at point *P*5 is not critical; it should be about a decade below F_p. To locate the zero at F_z, *C*1 (Fig. 12.18*a*) is chosen as $C_1 = 1/2\pi R2 F_z$.

The design example of the following section will clarify all the above.

12.18 Design Example—Stabilizing a Discontinuous-Mode Flyback Converter

Stabilize the feedback loop of the design example in Sec. 4.3.2.7. It is assumed that the output capacitor has an ESR and so a Type 2 error amplifier will be used. The circuit is shown in Fig. 12.18*a*. Recall its specifications:

V_o	5.0 V
$I_{o(\text{nom})}$	10 A
$I_{o(\min)}$	1.0 A
$V_{dc(\max)}$	60 V

$V_{dc(min)}$	38 V
$V_{dc(av)}$	49 V
Switching frequency	50 kHz
L_p (calculated in Sec. 4.3.2.7)	56.6 μH

Recall from Sec. 4.3.2.7 that C_o was calculated as 2000 μF. But it was pointed out there that at the instant of turnoff, the peak secondary current of 66 A would cause a thin spike of 66 × 0.03 = 2 V across the anticipated ESR of 0.03 V for a 2000-μF capacitor. It was noted that either this thin spike could be integrated away with a small *LC* circuit or C_o could be increased to lower its ESR.

Here, both will be done. Capacitance C_o will be increased to 5000 μF to decrease R_c to (2/5)0.03 or 0.012 Ω (since R_c is inversely proportional to C_o). The initial spike at $Q1$ turnoff is then 66 × 0.012 or 0.79 V peak. This can easily be integrated down to an acceptable level with a small *LC* which will be outside the feedback loop.

Now the output circuit gain curve can be drawn—first for $R_{o(min)}$ of 5/10 = 0.5 Ω. The DC gain from Eq. 12.19 is

$$G = \frac{V_{dc}}{3}\sqrt{\frac{0.4R_oT}{L_p}}$$

$$= \frac{49}{3}\sqrt{\frac{0.4 \times 0.5 \times 20 \times 10^{-6}}{56.6 \times 10^{-6}}}$$

$$= 4.3$$

$$= +12.8 \text{ dB}$$

Pole frequency, from Eq. 12.20, is

$$F_p = 1/2\pi R_oC_o$$

$$= 1/2\pi 0.5 \times 5000 \times 10^{-6}$$

$$= 63.7 \text{ Hz}$$

and ESR zero frequency, from Eq. 12.20, is

$$F_{esro} = 1/2\pi R_oC_o$$

$$= 1/2\pi 65 \times 10^{-6}$$

$$= 2500 \text{ Hz}$$

The output circuit gain curve for R_o = 0.5 Ω is then drawn as *EFGH* in Fig. 12.19. It is horizontal at a level of +12.8 dB up to F_p = 63.7 Hz. There it breaks to a −1 slope down to the ESR zero at 2500 Hz. The error-amplifier gain curve can now be drawn.

Then choose F_{co} as one-fifth the switching frequency or as 50/5 = 10

kHz. On *EFGH*, the loss is −19 dB at 10 kHz. Hence make the error-amplifier gain +19 dB at 10 kHz. Go to 10 kHz and +19 dB (point *P*2) and draw a line backward with a slope of +1 (+20 dB/decade) in the direction of lower frequency. Now extend that line to a frequency somewhat lower than F_{esro}—say, to point *P*3 at 1 kHz, +39 dB. At point *P*3, draw a horizontal line back to—say—300 Hz at point *P*5 (where a zero will be located).

The location of the zero is not critical. In Sec. 12.17, it was suggested the zero at point *P*5 should be one decade below point *P*3. Some designers actually omit the zero at point *P*5.[5] But here it is added to gain some phase boost. Thus, for a zero at point *P*5, at that point turn the gain slope upward to a +1 slope (again in the direction of lower frequency).

Now verify that for $R_{o(max)}$ of 5 Ω, the total gain curve (output circuit plus error-amplifier transfer function) comes through F_{co} at a −1 slope.

For R_o = 5 Ω, Eq. 12.19 gives a DC gain of 13.8 or +23 dB. And Eq. 12.20 gives the pole frequency as 6.4 Hz. The frequency of the ESR zero remains at 2500 Hz. Thus the output circuit transfer function for R_o = 5 Ω is *ABCD*.

The new F_{co} is then the frequency where the gain of the error amplifier on *P*6–*P*5–*P*3–*P*7 equals the loss on *ABCD*. This is seen to be at point *P*4 (3200 Hz), where the output filter loss is −29 dB and the error-amplifier gain is +29 dB. The sum of the error-amplifier gain curve and *ABCD* (equal to total gain curve) is seen to have a −1 slope as it passes through F_{co}.

It should be noted, however, that if R_o were somewhat larger, the curve *ABCD* would be depressed to a lower value along its entire length. Then the point at which the previously fixed error-amplifier gain curve is equal and opposite to the output filter loss curve would occur on the −1 slope of each curve.

The total gain curve would then come through the new F_{co} at a −2 slope and oscillations could occur. Thus, as a general rule, discontinuous-mode flybacks should be tested carefully for stability at minimum load current (maximum R_o).

The error-amplifier transfer function of *P*6–*P*5–*P*3–*P*7 is implemented as follows. In Fig. 12.18*a*, arbitrarily choose R_1 = 1000 Ω. Gain at point *P*3 is seen in Fig. 12.19 to be +38 dB or numerical gain of 79. Thus R_2/R_1 = 79 or R_2 = 79,000 Ω. For the pole at point *P*3 at 1 kHz, $C_2 = 1/(2\pi F_p R_2)$ or C_2 = 2000 pF. For the error-amplifier zero at 300 Hz, $C_1 = 1/(2\pi F_z R_2)$ = 6700 pF.

Because of the single-pole rolloff characteristic of the output circuit, its absolute maximum phase lag is 90°. But because of the ESR zero, it is much less and there rarely is a phase-margin problem in the discontinuous-mode flyback.

Thus consider the situation for R_o = 0.5 Ω. Lag at F_{co} (10 kHz) due to the pole at 64 Hz and the ESR zero at 2500 Hz is

$$\begin{aligned}\text{Output circuit lag} &= \tan^{1}\left(\frac{10{,}000}{64}\right) - \tan^{1}\left(\frac{10{,}000}{25{,}000}\right)\\ &= 89.6 - 76.0\\ &= 13.6^\circ\end{aligned}$$

and the error-amplifier lag at 10,000 Hz due to the zero at 300 Hz and the pole at 1000 Hz (see Fig. 12.20, curve *P6–P5–P3–P7*) is

$$270 - \tan^{-1}\left(\frac{10{,}000}{300}\right) + \tan^{-1}\left(\frac{10{,}000}{1000}\right) = 270 - 88 + 84 = 266^\circ$$

Total phase lag at 10,000 is then 13.6 + 266 = 280°.This yields a phase margin at F_{co} of 360 − 280 = 80°.

12.19 Transconductance Error Amplifiers

Many of the commonly used PWM chips (1524, 1525, 1526 family) have *transconductance* error amplifiers. Transconductance g_m is the change in output current per unit change in input voltage. Thus

$$g_m = \frac{dI_o}{dV_{in}}$$

Then for shunt impedance Z_o at the output node to ground

$$dV_o = dI_o Z_o = g_m\, dV_{in} Z_o$$

or gain G is

$$G = \frac{dV_o}{dV_{in}} = g_m Z_o$$

The unloaded, open-loop gain characteristic of the 1524, 1525-family amplifiers have a DC gain of nominally +80 dB, have a pole at 300 Hz, and thereafter fall at a −1 or −20 dB/decade slope. This is seen as curve *ABCD* in Fig. 12.20*a*.

A pure resistance R_o shunted from output node to ground yields a gain curve which is constant and equal to $g_m R_o$ from DC up to the frequency where it intersects the curve *ABCD* in Fig. 12.20*a*. For the 1524, 1525 family, g_m is nominally 2 mA/V. Thus gains for R_o = 500*K*, 50*K*, and 30*K* are respectively 1000, 100, and 60 and are shown as curves *P1–P2*, *P3–P4*, and *P5–P6* in Fig. 12.20*a*.

In most cases, Type 2 error-amplifier gain characteristics are re-

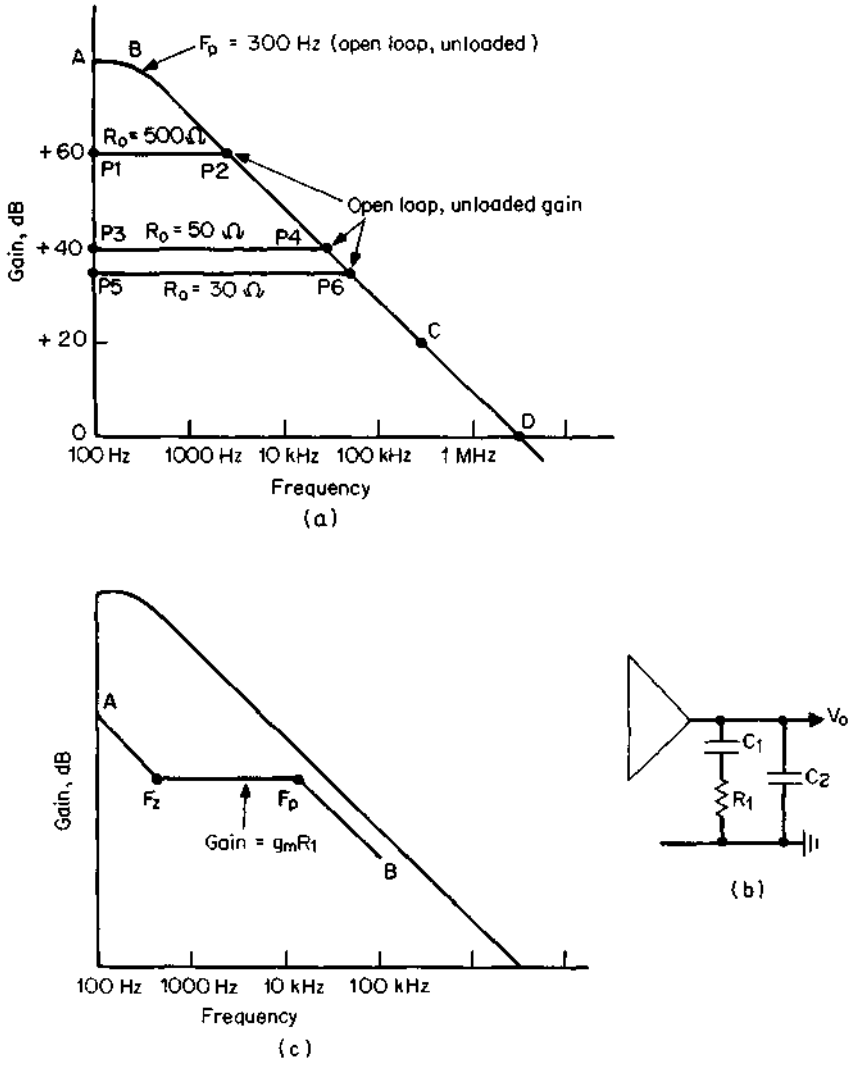


Figure 12.20 (*a*) Open-loop unloaded gain curve for PWM chip 1524, 1525 error amplifiers in *ABCD*. When loaded with indicated shunt resistors to ground, gain is constant at $G = g_m R_o$. (*Courtesy Silicon General*). (*b*) A Type 2 error-amplifier gain curve with shunt network to ground. (c) Gain with circuit of Fig. 12.20*b* is *A*–F_z–F_p–*B*: $F_z = 1/2\pi R1C1$; $F_p = 1/2\pi R1C2$.

quired. This is easily obtained with the network shown in Fig. 12.20*b* shunted to ground.

At low frequencies, X_{C1} is much greater than $R1$ and effectively $C1$ and $C2$ are in parallel with the internal 100 pF to ground which caused the open-loop 300-Hz pole. This shifts the 300-Hz pole to a

lower frequency, and after that lower frequency, gain resumes falling at a −1 slope. At a frequency F_z (= $1/2\pi R1C1$), where $X_{C1} = R_1$, there is a zero and gain slope turns horizontal at a magnitude $g_m R1$. Further on in frequency at $F_p = 1/2\pi R1C2$ where $X_{C2} = R1$, the pole turns the gain slope to −1.

The gain curve with the circuit configuration of Fig. 12.20*b* is shown in Fig. 12.20*c*.

Most frequently, in the 1524, 1525 family of PWM chips the error-amplifier gain curves are shaped in the above-mentioned fashion with the network of Fig. 12.20*b* shunted to ground rather than being used in the conventional operational-amplifier mode.

Whether a network as Fig. 12.20*b* is shunted to ground or returned around to the inverting input terminal as in a conventional operational amplifier, there is a restriction on the magnitude of $R1$ arising from the following. The internal error amplifiers in the above-mentioned chips cannot emit or absorb more than 100 μA. With a 3-V triangle at the PWM comparator, the error-amplifier output may have to move the 3 V from the bottom to the top of the triangle for sudden line or load changes. Thus for $R1$ less than 30,000 Ω, this 3-V fast swing would demand more than the available 100 μA. Response time to fast load or line changes would then be sluggish.

Because of this 100-μA limit on output current, many designers prefer not to use the error amplifier internal to the PWM chip. Since the chip's output node is brought out to one of the output pins, some prefer to use a better external error amplifier and connect it to the chip's error-amplifier output node at the appropriate output pin.

However, it may be essential from a cost viewpoint to use the chip's internal error amplifier. Calculation of the output filter may show that its loss at F_{co} is so low that to match the error-amplifier gain to it, R_1 must be less than 30,000 Ω. If this situation arises, R_1 can be increased to 30,000 Ω to match an artificially increased output filter loss at F_{co}. This increased output filter loss at F_{co} can easily be achieved by shifting its pole frequency to a lower value by increasing the output filter's inductance or capacitance.

References

1. D. Venable, "The *K* Factor: A New Mathematical Tool for Stability Analysis and Synthesis," *Proceedings Powercon 10*, 1983.
2. A. Pressman, *Switching and Linear Power Supply, Power Converter Design*, pp. 331–332, Switchtronix Press, Waban, Mass., 1977.
3. K. Billings, *Switchmode Power Supply Handbook*, Chap. 8, McGraw-Hill, New York, 1989.
4. G. Chryssis, *High Frequency Switching Power Supplies*, 2d ed., Chap. 9, McGraw-Hill, New York, 1989.
5. Unitrode Corp., *Power Supply Design Seminar Handbook*, Apps. B, C, Watertown, Mass., 1988.

Chapter

13

Resonant Converters

13.1 Introduction

As newer integrated circuits lead to more electronic functions in smaller packages, it becomes essential for power supplies also to become smaller. The power supplies mainly get smaller by increasing their operating frequency to decrease the size of the power transformer and output LC or capacitive filter. The supplies also get smaller by increasing their efficiency so as to require smaller heat sinks.

Thus the major objective in present-day power supply technology is to operate at switching frequencies higher than the currently commonplace 100 to 200 kHz.

However, going to higher switching frequencies with the conventional square current waveform topologies discussed up to this point increases transistor switching losses at both turnoff and turnon. Turnon losses, due to charging and discharging MOSFET output capacitances (Sec. 11.1), become important only at frequencies over 1 MHz.

As discussed in Chaps. 11 and 1, the overlap of falling current and rising collector voltage at turnoff yields a high spike of dissipation during the turnoff interval. As switching frequency increases, the more frequent the occurrence of high-dissipation spikes results in higher average transistor dissipation.

The higher losses require larger heat sinks to the point where there may be no net size decrease despite the smaller power transformer and output filter.

At sufficiently high frequencies ($\geq$1 MHz) transistor switching losses may become so high that even with a sufficiently large heat sink to keep the transistor case temperature at a low value, there still is a problem. The junction-to-case temperature rise with the usual

1°C/W junction-to-case thermal resistance may still raise the transistor junction temperature to dangerous levels.

Adding snubbers (Chap. 11) at collector-drain outputs reduces transistor switching losses. But if a dissipative *RCD* snubber (Sec. 11.3) is used, that does not decrease total dissipation—it simply shifts losses from the transistor to the snubber resistor. Nondissipative snubbers (Sec. 11.6) do reduce transistor switching losses but are troublesome at frequencies over 200 kHz.

Thus to operate at higher frequencies which will permit smaller power supplies, transistor switching losses at turnoff and turnon must be decreased.

This is achieved in *resonant converters* by associating a resonating *LC* circuit with the switching transistor, to render its current sinusoidal rather than square wave in shape. It is then arranged to turn the transistor on and off at the zero crossings of the current sine wave. There is thus no overlap of falling current and rising voltage at turnoff or rising current and falling voltage at turnon and hence no switching losses.

Circuits which turn on and off at zero current are referred to as *zero current switching* (ZCS) types.[1] But recall (Sec. 11.1) that switching losses can occur at turnon even though there is no overlap of rising voltage and falling current at the zero crossing of the current sine wave.

Thus in Sec. 11.1 it was pointed out that considerable energy $[0.5C_o(2V_{dc})^2]$ is stored on the relatively large output capacity of a MOSFET. When the MOSFET is turned on once per period T, it dissipates $0.5C_o(2V_{dc})^2/T$ watts in the MOSFET.

Circuits which cope with this are called *zero-voltage-switching* (ZVS) types.[2] They work by ensuring that the transistor output capacity is made the capacitor of a resonant *LC* circuit. Then the voltage or energy stored on it when the transistor is off is changed to stored current or energy in the inductor of the resonant circuit. Then later in the cycle, this energy is returned without loss to the power supply bus. Operation is much like that of the nondissipative snubber of Sec. 11.6.

The intense industry interest in resonant converters started about 10 years ago. In the past decade, there has been a vast army of researchers producing a vast amount of articles on the subject.

Dozens of new resonant converter topologies have been proposed and mathematically analyzed. Some of these have actually been built and have achieved high efficiency (80 to 92 percent) and good power density (some claims of ≤ 50 W/in^3). Such quoted high power densities are presumably for DC/DC converters which do not have the large input filter capacitor required of all off-line converters. Undoubtedly, they also are for cooling by an external "cold plate" whose size and means of cooling are seldom reckoned in calculating power densities.

Coverage of all the (very large number of) resonant converter topologies and their large tree of operating modes is beyond the scope of this text. An overview of only those topologies and their operating modes that are long-lived and not yet cast aside will be presented herein.

It is worth noting here that very frequently in this field, articles on a new approach often end with comments on its restricted usage in line and load variations and excessive component stress. Surely at the next conference the following year, solutions to last year's problems are offered by numerous other investigators, but again with the admonition that there are other restrictions on their usage.

This simply reflects the fact that despite their significant advantage in somewhat restricted areas of application, resonant converters do not yet have the flexibility of PWM converters to cope with large line and load changes, short-circuited or unloaded outputs, and component tolerances. Also, they operate mainly at higher peak transistor currents for the same output power than do conventional PWM square-wave inverters and in some circuit configurations at larger voltage stresses. But this is to be expected in an area where totally new ground is being broken.

13.2 Resonant Forward Converter

First the simplest resonant circuit, the resonant forward converter will be discussed—primarily to see how it is arranged for transistor turnoff to occur at zero current and to see how critical the exact turnoff time may be.

Figure 13.1 shows a simple resonant forward converter[3] operating in the discontinuous mode. *Discontinuous mode* implies that current in the resonant LC circuit is not a continuous sine wave, but a sequence of a burst of one half or one full cycle of sine-wave current separated by a large time interval T_s as in Fig. 13.1.

The resonant frequency of the circuit F_r is $1/2(t_1 - t_0)$ and is fixed by the passive resonant elements L_r and the value of C_r reflected into the primary, $(C_r)(N_s/N_p)^2$. The resonant frequency of the circuit is then

$$F_r = 1/2\pi\sqrt{L_rC_r(N_s/N_p)^2} \tag{13.1}$$

where L_r is the transformer leakage inductance or that plus some externally added small inductance to make the total L_r relatively independent of production variations in the leakage inductance itself.

Transistor $Q1$ is turned on at a switching frequency rate of $F_s = 1/T_s$. The circuit works as follows. First, L_r and $C_r(N_s/N_p)^2$ form a series resonant circuit with DC secondary current reflected by the transformer turns ratio in shunt across the reflected capacity in the pri-

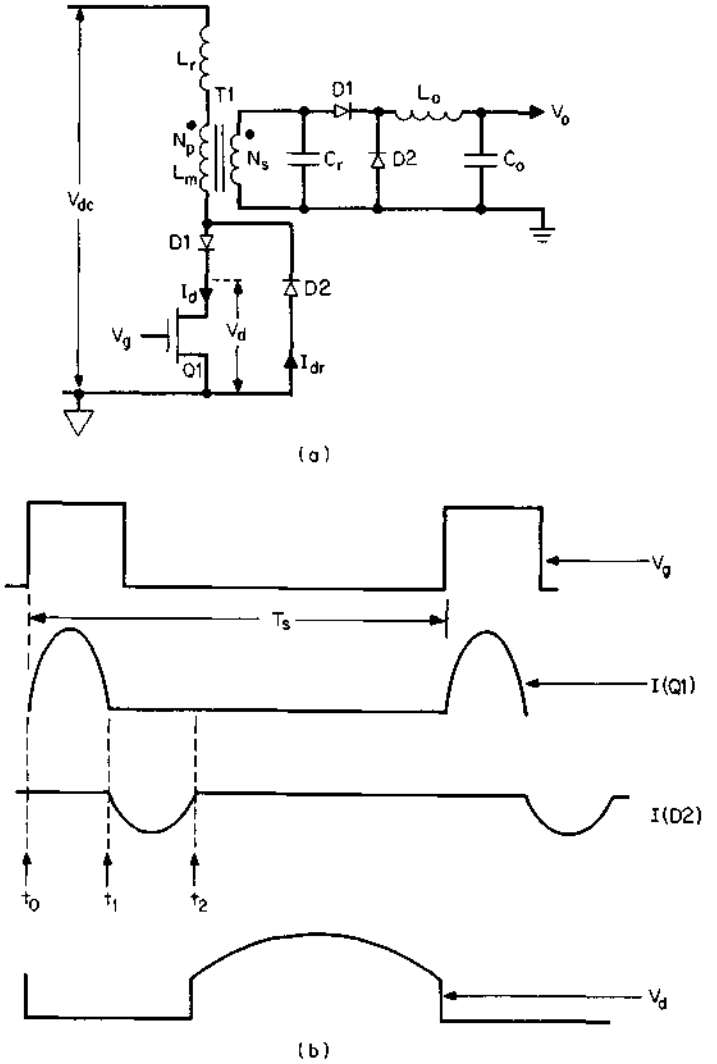


Figure 13.1 Resonant forward converter. Capacitance C_r reflected into primary resonates with L_r, the leakage inductance. The MOSFET gate is turned off shortly after the zero crossing of the first positive half cycle of drain current. When the drain rises after $D2$ stops conducting at the end of the first negative half cycle of current in $D2$, a half cycle of resonant ring of the magnetizing inductance L_m with the capacitance C_r reflected into the primary resets the $T1$ core.

mary. In resonant converter jargon, this is a *parallel resonant converter* (PRC) as the load is placed in parallel or shunt with the resonating capacitor. Other circuits, to be discussed later, place the load in series with the series resonating *LC* elements and are called *series resonant converters* (SRCs).

Now just prior to $Q1$ turnon, no current is flowing in the resonant circuit as the circuit is discontinuous—i.e., there is a long time gap between bursts of half sine waves (Fig. 11.1). Thus when $Q1$ does turn on at t_0, a half sine wave of current starts flowing through it. The current amplitude is zero at time t_0 since there has been no current flow in the resonant circuit just prior to turnon.

The current goes through its first positive half sine wave. At t_1, it passes through zero again and reverses, trying to generate the first negative half sine wave. This current is forced to flow by the voltage stored on the resonating capacitor in the primary.

This negative half cycle of current flows up through the anode of $D2$ and completes its loop through the supply source V_{dc}. For the half cycle that current flows through $D2$, the $Q1$ drain remains clamped to about -1 V (the $D2$ forward drop). Now any time after t_1 (between t_1 and t_2) there is no current in $Q1$ and its gate may be turned off.

The $Q1$ on-time duration is thus not at all critical. It must be greater than a half period of the resonant current sine wave and less than a full period. At t_2, the negative half sine wave of current in $D2$ has returned to zero and now current in the $T1$ magnetizing inductance drives the drain up toward $2V_{dc}$ to reset the core.

The magnetizing inductance and the capacitance reflected across it from the secondary form another resonant circuit. When the drain finally rises at t_2, a negative half sine wave of voltage across that resonant circuit resets the core exactly to its starting point on the *BH* loop.

The objective of zero current turnoff has thus been achieved, and there is no turnoff dissipation. The reverse recovery time of diode $D2$ is short enough so that dissipation in it is not significant—especially since current in it is already zero or close to it when the drain rises toward $2V_{dc}$.

For the duration of the first positive half resonant cycle ($t_1 - t_0$), the primary is delivering a burst of power to the secondary and the output load. The DC output voltage is regulated by varying the spacing between these discontinuous bursts—i.e., varying the switching frequency $F_sC = 1/T_s$. If V_{dc} goes up or the DC output load current goes down, the spacing between bursts must increase (F_s must decrease)—and vice versa, of course; as V_{dc} goes down or load current increases, F_s must increase.

This method of voltage regulation—varying the switching fre-

quency F_s—is a major drawback of resonant converters. In conventional pulse-width-modulated converters, switching frequency remains constant and on-time pulse width is varied. For most resonant topologies, pulse width is constant (the half period of the resonant LC circuit) and its repetition frequency is varied.

Variable-frequency regulation is objectionable in many circumstances. Where a computer is involved, the computer operators often require the power supply switching frequency synchronized to a submultiple of their computer clock. This reduces the probability of any power supply noise spike pickup generating false ones or zeros in the computer logic circuits.

Also where a cathode-ray tube (CRT) screen is involved, it is desirable to have the power supply switching frequency locked in phase with the CRT horizontal line frequency. Unless it is, any power supply switching frequency noise pickup appears running continuously and in a random fashion across the screen. When power supply switching frequency is constant and locked in phase to the CRT horizontal line rate, any noise pickup remains at a fixed location on the screen and is far less disconcerting to an operator.

13.2.1 Measured waveforms in a resonant forward converter

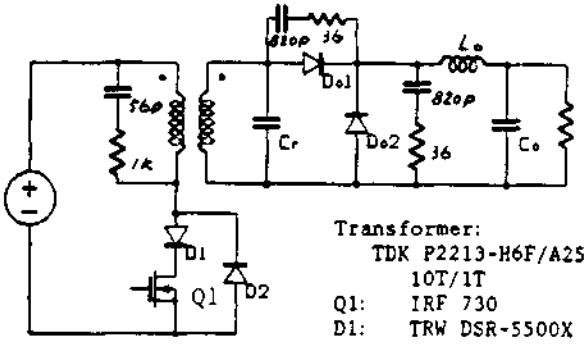
It is of interest to see actual measured waveforms in a resonant forward converter of the type discussed above. The waveforms in Fig. 13.2*b* are from Ref. 3 and are reproduced with the courtesy of the authors, F. Lee and K. Liu.

The circuit in Fig. 13.2*a* is a forward converter with an output power of 32 W (5.2 V, 6.2 A). Input voltage V_{dc} is 150 V and the circuit is shown at a switching frequency of 856 kHz. The transformer turns ratio is 10/1.

In Fig. 13.2*b*, waveform 3, drain current is seen to be a half sine wave whose resonant half period is 0.2 μs (F_r = 2.5 MHz). The secondary capacitor is 0.15 μF, which reflects into the primary as $0.15(0.1)^2 = 0.0015$ μF. Then for F_r = 2.5 MHz, the transformer leakage inductance plus any added discrete inductance must be

$$L_r = 1/4\pi^2(F_r)^2C = 1/4\pi(2.5)^2(0.0015)10^{-6}$$
$$= 2.7\ \mu\text{H}$$

It is seen in Fig. 13.2*b*, waveform 1, that the MOSFET gate is turned off shortly after the first positive half cycle of the current sine wave has passed through zero, thus meeting the objective of zero current turnoff. It is also seen in waveform 4 that the drain voltage has

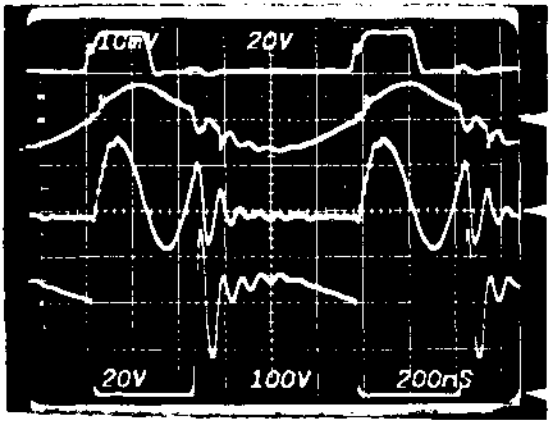


Transformer:
TDK P2213-H6F/A250
10T/1T

Q1:	IRF 730
D1:	TRW DSR-5500X
D2:	TRW DSR-5500X
Do1:	IR 31DQ06 *2
Do2:	IR 31DQ06 *2
Cr:	0.15μ
Lo:	200μ
Co:	1.0μ + 470μ

A 800kHz, 30W forward quasi-resonant converter

(a)



Vin = 120V, Vo = 3.56V
fs = 856kHz, Io = 4.2A
1st waveform: Vgs(20V/Div)
2nd waveform: Vcr(20V/Div)
3rd waveform: Ip (1A /Div)
4th waveform: Vds(100V/Div)

(b)

Figure 13.2 Measured waveforms on an actual circuit as Fig. 13.1. (*Courtesy F. Lee and K. Liu*)

started rising shortly after the negative diode current ($D2$) has come back up again to zero as discussed above.

It is also seen that the voltage across the secondary capacitor in waveform 2 (which is a replica of the voltage across the primary magnetizing inductance) has reversed polarity to reset the core as discussed above. That waveform also indicates the maximum switching frequency (F_s). The minimum spacing between $Q1$ turnon pulses must be such as to permit the negative half cycle in waveform 2 to return to zero, as that is an indication that the core has been fully reset.

The foregoing is a small indication of some of the difficulties in making resonant converters work over large line and load variations. Consider the 2.7 μH calculated for the resonant inductor above. This is so small a value that changes in wire lengths and routing to the transformer can result in large percentage changes in the total inductance. For as small a value as 2.7 μH, production spread in the transformer will cause large percentage changes in its leakage inductance. If the inductance increases, the resonant half period increases. Then the MOSFET on time, which should be greater than a resonant half period, may be too short. The MOSFET may find itself turning off before the end of a resonant half period—i.e., before the sine-wave zero crossing.

Then L_r may be increased to make it less susceptible to large percentage changes arising from transformer production spread and wire length and routing. This increases the resonant half period and hence also decreases the maximum switching frequency as that must be sufficiently low to permit complete core resetting.

Now if L_r is increased but C_r is decreased to maintain the same resonant half period, the ratio L_r/C_r is increased. But the peak of the $Q1$ sine-wave current (Fig. 13.2*b*, waveform 3) is roughly inversely proportional to $\sqrt{L_r/C_r}$. If that is increased, maximum required peak currents may not be obtainable and maximum DC output current may not be obtainable.

Of course, many such problems may be solved in specific cases. The aforesaid is mentioned only as an indication that even such a simple resonant converter has less flexibility to cope with varying specifications, line and load conditions, and manufacturing tolerances.

It is of interest to note that a conventional forward converter operating at a relatively low frequency could yield equal output power at less primary current than shown in Fig. 13.2*b*. Thus, that figure shows (for full-wave mode) peak current is 1.5 A for V_o = 3.56 V, I_o = 4.2 A or 15.0 W at V_{in} of 120 V.

Equation 2.28 gives the peak primary current for a conventional PWM square-wave forward converter as $I_p = 3.13P_o/V_{dc} = 3.13 \times 15/120 = 0.39$ A. Table 7.2*a* shows that the next-smaller-sized core than

the 2213 pot core used for Fig. 13.2*b* could easily be used, and at a lower frequency. Thus Table 7.2*a* shows that the 1811 core at 150 kHz could put out 19.4 W in a forward converter operating at only 150 kHz.

This is not to put down resonant converters, but only to point out the need to weigh them against conventional, proven topologies.

13.3 Resonant Converter Operating Modes

13.3.1 Discontinuous and continuous; above resonance and below resonance operating modes

Operating modes can be discontinuous as in Fig. 13.1. In the discontinuous mode (DCM), as noted, output voltage regulation is accomplished by varying the switching frequency. Power is delivered to the load as sequence of discrete current or power pulses separated by times long compared to their duration.

If the output voltage must be raised because V_{dc} has gone down or DC load current has been increased, the switching frequency or repetition rate of the discrete pulses is increased. Conversely, if output voltage has gone up because input voltage or output load resistance has gone up, the repetition rate of the discrete pulses is decreased.

Silicon controlled rectifier resonant converters, discussed in Chap. 6, have been operating successfully in the discontinuous mode for many years at frequencies in the range of 20 to 30 kHz.

In discontinuous-mode operation, however, large changes in load and line result in large changes in switching frequency. An improvement in this respect is obtained by operating in the continuous-conduction mode (CCM). In CCM, there is no or negligible gap between successive square-wave voltage pulses from the switching transistors or between successive current sinusoids.

The fundamental component of this square wave is a relatively distortionless, continuous sine wave whose frequency is that of the transistor square wave switching frequency. The resonant LC circuit has the characteristic Q curve (impedance vs. frequency) shown in Fig. 13.3. The average switching frequency is set either above or below resonance on the side of the Q curve.

Direct-current output voltage in the CCM mode is proportional to either the peak AC voltage on the resonating capacitor or the peak AC current in series in the resonating LC circuit.

Now, output voltage regulation is accomplished by moving the switching frequency along the side of the Q curve to change the amplitude of either the voltage output across the resonating capacitor (if

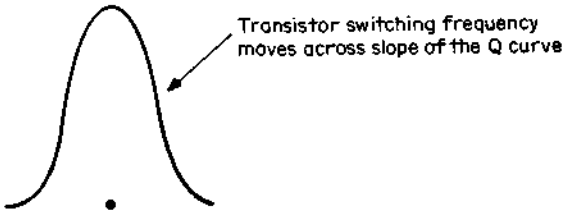


Figure 13.3 Output voltage regulation by shifting switching frequency across the slope of the Q curve of the resonating elements.

output is taken as the voltage across the capacitor) or resonant circuit current (if output is in the form of a resonant circuit current).

When operation is set above the resonant peak of the Q curve, it is referred to as *above-resonance mode* (ARM). Operation below the resonant peak is referred to as *below-resonance mode* (BRM).

In Fig. 13.3, it is seen that with a relatively steep (high Q) Q curve, it requires only small changes in frequency to obtain large changes in output amplitude.

Note that if the average switching frequency is above the resonant peak (ARM mode), to increase the output voltage or current, switching frequency must decrease and that if the average switching frequency is below the resonant peak (BRM mode), to increase the output voltage or current, switching frequency must increase.

This points out one of the problems with resonant converters. Much of the current literature indicates that the CCM is becoming the preferred mode of operation because it results in a smaller frequency range to achieve the usually desired load and line regulation.

If the feedback system has been designed on the assumption that operation is always above the resonant peak (ARM mode), however, a disastrous problem can arise if operation for some reason shifts to the BRM mode. Of course, if in the ARM mode, a decrease in the DC output voltage would be corrected by the variable-frequency oscillator in the control loop decreasing switching frequency to move higher up on the Q curve.

Since the Q curve is relatively steep, small changes in the magnitudes of the resonant L and C due to their production tolerances can shift the resonant peak. If the resonant peak were shifted sufficiently, operation could fall on the other side of the Q curve, and now the feedback loop, sensing a decrease in output voltage, would still try to correct it by decreasing switching frequency. This, of course, would now result in a further decrease in output voltage—i.e., positive rather than negative feedback.

13.4 Resonant Half Bridges in Continuous-Conduction Mode[4]

Much of the current resonant converter development effort seems to be for half-bridge topologies, which will now be considered. The following discussion is based on a classic article by R. Steigerwald.[4]

13.4.1 Parallel resonant converter (PRC) and series resonant converter (SRC)

Output power can be taken from the resonant LC circuit in either of two ways. When the output load (reflected into the power transformer primary) is reflected in parallel with the resonating capacitor, the circuit is referred to as a *parallel resonant converter* (PRC). When the load is reflected in series with the resonating LC circuit, it is referred to as a *series resonant converter* (SRC).

A parallel loaded resonant half bridge is shown in Fig. 13.4*b*; C_{f1}, C_{f2} are the input filter capacitors used in the scheme for generating a rectified 320 V whether operation is from 120 or 220 V AC (Fig. 5.1). They are large capacitors used only to split the rectified DC and have nothing to do with the resonant LC circuit.

The capacitor C_r is across the power transformer primary and resonates with an external inductor L_r at a frequency $F_r = 1/2\pi\sqrt{L_rC_r}$. The output inductor L_o is large and has a high impedance at F_r so that it does not load down C_r and kill the Q of the resonant LC circuit; L_o is sufficiently large that it runs in the continuous-conduction mode (Sec. 1.3.6). The impedance seen across C_r is the output load resistance multiplied by the turns ratio squared. The $T1$ magnetizing inductance is much larger than this and does not affect circuit operation.

A series-loaded resonant half-bridge converter (SRC) is shown in Fig. 13.4*a*. Here an external inductor L_r resonates with the equivalent capacity at the junction of the two $C_r/2$ capacitors ($= C_r$). Again the C_f capacitors are large line frequency filter capacitors and have nothing to do with the resonant circuit operation. The load in a SRC is the secondary load resistor reflected by the turns ratio squared in series with the resonating LC elements. In the series resonant circuit, the secondary side inductor is omitted.

The SRC is used for high-voltage supplies as it requires no output inductor. An output inductor for high output voltage would have to support large voltages across themselves and would be bulky. The PRC is used for low-voltage, high-current supplies as the output inductor limits ripple current in the output capacitor.

Both the series and parallel half bridges to be discussed below will operate in the continuous-conduction mode. Since DC voltage will be

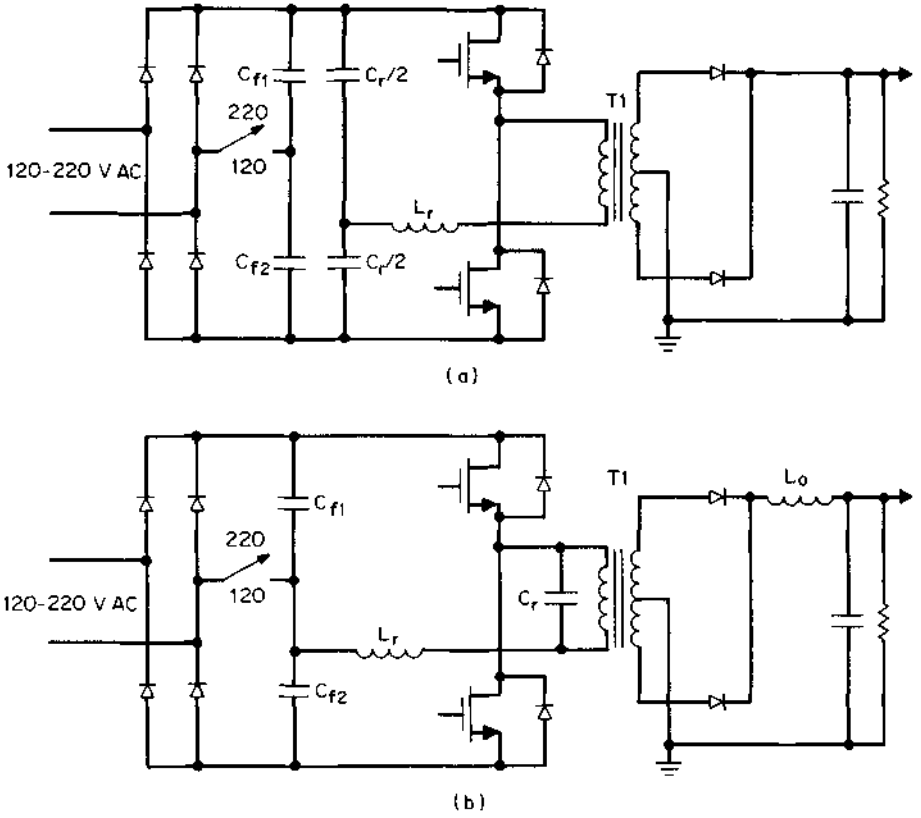


Figure 13.4 (*a*) A series-loaded resonant half bridge. Inductance L_r resonates with capacitance C_r. The load is reflected by $T1$ in series with the resonant circuit. Transistors are turned off directly after the end of the first half cycle of resonant current to achieve zero current switching. In series loading, the output filter is capacitive. (*b*) A parallel-loaded resonant half bridge. Inductance L_r resonates with capacitance C_r. The load is reflected by $T1$ in shunt with the resonating capacitor. In parallel loading, the output filter has a high-impedance inductor input to avoid lowering the Q of the resonant circuit.

regulated by varying the switching frequency, it is necessary to know how AC voltage across the reflected load varies with frequency as operation moves along the side of the resonant circuit's Q curve. The rectified DC output voltage will be proportional to the AC voltage across the reflected resistance.

13.4.2 AC equivalent circuits and gain curves for series- and parallel-loaded half bridges in continuous-conduction mode[4]

Figure 13.5*a* and 13.5*b* shows the equivalent AC circuits for the series- and parallel-loaded half bridges of Fig. 13.4*a* and 13.4*b*, respec-

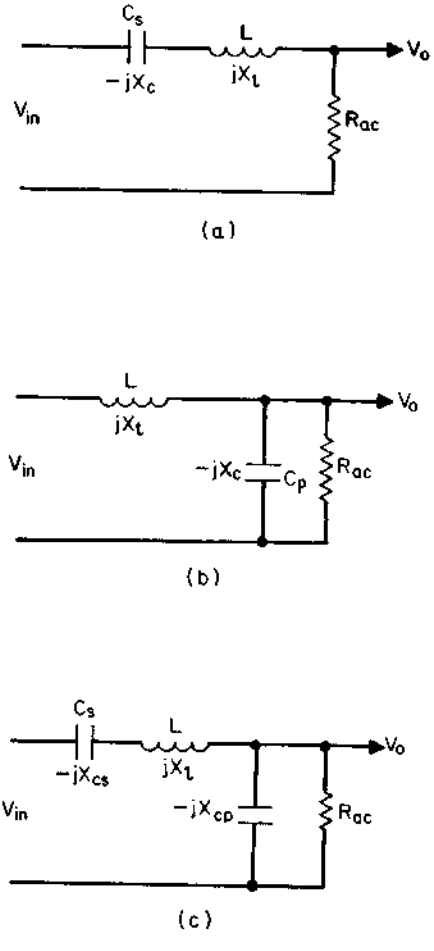


Figure 13.5 (*a*) AC equivalent circuit for series-loaded resonant half bridge (Fig. 13.4*a*). (*b*) AC equivalent circuit for parallel-loaded resonant half bridge (Fig. 13.4*b*). (*c*) AC equivalent circuit for series-parallel-loaded resonant half bridge, often called *LCC circuit*. AC gain between outputs and inputs above fixes the ratio between DC output voltage and half the DC input voltage. (*Courtesy R. Steigerwald*)

tively. Inputs to these circuits are the square waves of amplitude $\pm V_{dc}/2$ generated by the switching transistors. Following the analysis presented by Steigerwald,[4] consider only the fundamental of this square-wave frequency and calculate gain or ratio of output to input voltage as a function of frequency.

From the equivalent circuits of Fig. 13.5, the ratios for series and parallel loaded circuits are

$$\text{Series-loaded:} \quad \frac{V_o}{V_{in}} = \frac{1}{1 + j[(X_l/R_{ac}) - (X_c/R_{ac})]} \tag{13.2}$$

$$\text{Parallel-loaded:} \quad \frac{V_o}{V_{in}} = \frac{1}{1 - (X_l/X_c) + j(X_l/R_{ac})} \tag{13.3}$$

where R_L is the secondary load reflected into the primary and $R_{ac} = 8R_L/\pi^2$ for series loading, $R_{ac} = \pi^2 R_L/8$ for parallel loading.

From these relations, Steigerwald plots the ratio $NV_{odc}/0.5V_{in}$, where N is the power transistor turns ratio, V_{in} is the input supply voltage, and V_{odc} is the DC input voltage.

For the series-loaded case in Fig. 13.6, $Q = w_o L/R_L$ and $w_o = 1/\sqrt{LC_s}$. For the parallel-loaded case in Fig. 13.7, $Q = R_L/w_o L$ and $w_o = 1/\sqrt{LC_p}$.

From Figs. 13.6 and 13.7, some of the problems with resonant converters can be seen.

13.4.3 Regulation with series-loaded half bridge in continuous-conduction mode (CCM)

For a number of reasons, Steigerwald states that operating above the resonant peak (ARM) is preferable to below resonance.

Figure 13.6 shows how the continuous conduction mode SRC half bridge regulates. If initial operation were at A at $Q = 2$ at normalized frequency 1.3, the output/input voltage ratio would be 0.6. Now if the load resistance R_L decreased so as to increase Q to 5, normalized switching frequency would have to be decreased to about 1.15 at point B to yield the same output voltage. If R_L increased to a value which yielded $Q = 1$, for the same output voltage, operation would have to shift to point C, where normalized frequency is about 1.62.

It is obvious that as load R_L increases, Q decreases and the gain curve approaches a horizontal line, making the required frequency change so large as to be impractical. Obviously, regulation at open circuit is impossible as the Q curve has no resonant peak or *selectivity.*

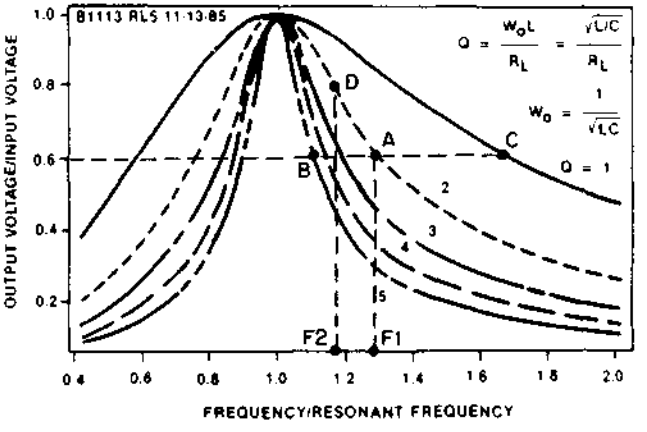


Figure 13.6 Series resonant converter gain curve from AC equivalent circuit of Fig. 13.5*a*. (*Courtesy R. Steigerwald*)

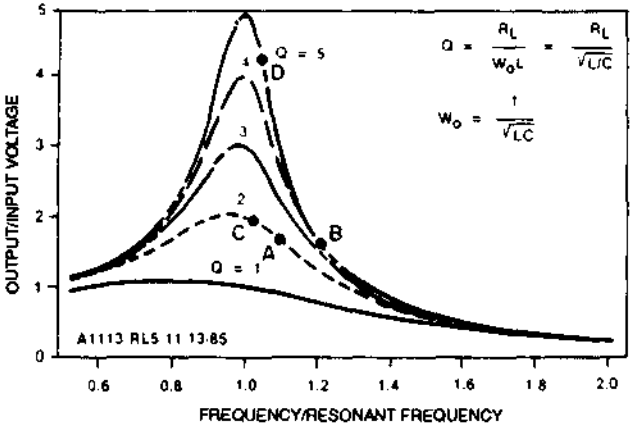


Figure 13.7 Parallel resonant converter gain curve from AC equivalent circuit of Fig. 13.5*b*. (*Courtesy R. Steigerwald*)

The actual operating point on the Fig. 13.6 gain curve depends on the DC input supply and output voltage. For a given output voltage, if input voltage dropped from *A* to *D*, for example, operation would shift out on the same *Q* curve to a lower normalized frequency at F1 to F2.

Operation is thus very nonlinear over the slopes of the *Q* curve. Some combinations of line and load may even not be possible. It is also obvious that choosing the exact location to operate on the *Q* curve is very tricky.

The feedback loop automatically selects the location on the *Q* curve which yields the correct output voltage. But that location may be too dangerously close to the resonant peak, at a low-*Q* curve where selectivity or a resonant peak is minimal or on the bottom tail of a high-*Q* curve where large frequency changes are required.

The scheme is sensitive to tolerances in the resonant components. Operation close to the resonant peak may cause switchover to the opposite side of the peak on load and line transients and result in positive rather than negative feedback.

13.4.4 Regulation with parallel-loaded half bridge in continuous-conduction mode

From Eq. 13.3, Steigerwald plots the gain curve of the parallel-loaded half bridge (Fig. 13.7). Here it is seen that operation at high and even no load is possible. If operation were initially at point *A* ($Q = 2$, normalized frequency = 1.1), and R_L increased so that Q ($= R_L/w_oL$) was 5, operation would shift to *B* at a normalized frequency of about 1.23. Obviously from the shape of the *Q* curve, even larger values of *Q* or open-circuit operation could be tolerated.

One problem with the PRC circuit is that if operation is close to the resonant peak at low Q (say, point C) and load momentarily opens or gets much larger to, say, $Q = 5$ before the feedback loop can correct frequency, output voltage could rise dangerously high at point D.

It is sometimes stated that the PRC is naturally short-circuit proof, as even when there is an output short circuit, there is also a short circuit across the transformer primary (Fig. 13.4*b*) and the resonating inductor limits transistor current.

Further examination indicates that this would not occur. If operation is above resonance, a short circuit at the output would force the feedback loop to move higher up on the Q curve to increase output (Fig. 13.3).

This would move switching frequency lower and eventually over the peak of the Q curve into the positive-feedback region. But the loop would not "know" that; it would continue trying to drive the switching frequency lower, "thinking" that lower frequency would move operation higher up to the top or the resonant peak.

Further, at lower frequency, the resonant inductor would have lower impedance, which would result in increased current drawn from the transistors. A clamp to limit the minimum frequency would not be practical as the resonant peak frequency is subject to large variations because of production spread in the resonant L and C.

13.4.5 Series-parallel resonant converter in continuous-conduction mode

One disadvantage of the PRC is that at light load (large shunt resistor across C_r in Fig. 13.4*b*), the circulating currents (and currents in the transistor) are no less than those at heavy load. In either case, the effective resistance reflected across C_r must be high so as not to kill the Q of the circuit. If this is true at heavy loads, it is even more true at light loads. This is simply another way of saying that the current in the reflected load resistor across C_r at light load is a small fraction of the current in C_r. Thus power losses in the transistors do not decrease at light load, and efficiency at such loads is poor.

This is not so in the SRC, where, for constant output voltage across the load resistor in series in the resonating LC circuit, as load current decreases, current through the load resistor (which is also current in the transistors) also decreases. Thus efficiency remains high at light load (low output power) in the SRC.

A circuit which takes advantage of the good light-load efficiency of the SRC and the ability to regulate at light or open load of the PRC is the series-parallel continuous-conduction mode converter of Fig. 13.5*c*. It is also referred to as the *LCC circuit*.[5]

As seen in Fig. 13.5*c* it has both a series capacitor C_s and a shunt

capacitor C_p. For proper selection of C_s and C_p, the advantages of both SRC and PRC can be obtained to an acceptable degree.

It is seen that if C_p is zero, the circuit is that of an SRC. As C_p becomes larger than C_s, it takes on more of the characteristics of a PRC. If it becomes higher than C_s, it develops the poor efficiency at light load characteristic of a PRC.

Again from the Steigerwald article,[4] the AC gain characteristics of the *LCC* is calculated from its AC equivalent circuit in Fig. 13.5*c* as

$$\frac{V_o}{V_{in}} = \frac{1}{1 + (X_{cs}/X_{cp}) - (X_l/X_{cp}) + j[(X_l/R_{ac}) - (X_{cs}/R_{ac})]}$$

and for $Q_s = X_l/R_l$, $R_{ac} = 8R_l/\pi^2$, $w_o = 1/\sqrt{LC_s}$, and R_l equal to the DC output resistance reflected by the square of the turns ratio into the primary, the gain from half the input supply voltage to the DC output voltage reflected into the primary is

$$\frac{NV_{odc}}{0.5V_{in}} = \frac{8/\pi^2}{1 + (C_p/C_s) - (w^2LC_p) + jQ_s[(w/w_s) - (w_s/w)]}$$

This is plotted in Fig. 13.8*a* and 13.8*b* for $C_s = C_p$ and $C_s = 2C_p$.

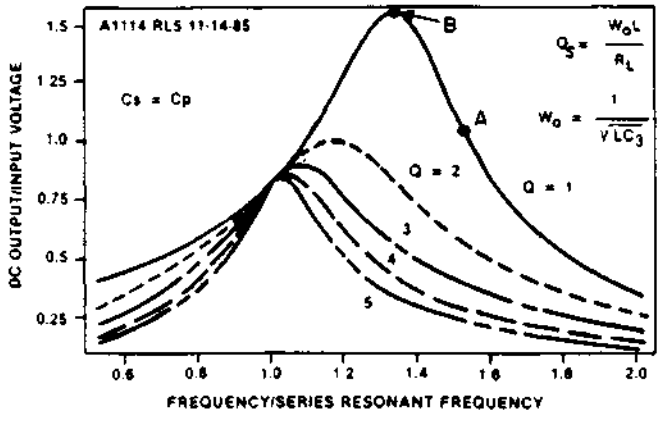
It is seen in Fig. 13.8*a* and 13.8*b*, for both cases, for $Q = 1$ and less, that there still remains a resonant bump and some selectivity on the gain curves so that no load operation is possible. This removes one of the drawbacks of the pure SRC circuit—that of no regulation for no-load or light-load conditions (compare to Fig. 13.6 for the pure SRC circuit).

Figure 13.8*a* and 13.8*b* illustrates some of the subtleties and complexities of continuous-mode operation. Thus consider that operation is at point *A* in Fig. 13.8*a*, DC load resistance and *Q* are constant, and DC input voltage decreases causing a slight decrease in output voltage.

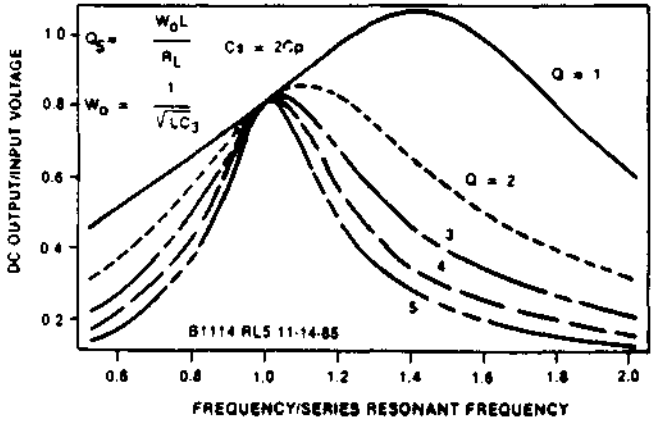
The feedback loop will attempt to correct this by moving higher up on the same *Q* curve to get more output. To do this, the switching frequency will be decreased. But if it decreases too much, it will fall over the top of the resonant peak at *B* and thereafter further frequency decrease will decrease the output.

To avoid this, it must be ensured that at lowest DC input and minimum *Q*, operation never requires a frequency lower than that corresponding to point *B* in Fig. 13.8*a*. A better appreciation of the practical problems in CCM is gained when it is attempted to make the circuit work over all tolerances in the *LC* product which shift the resonant peak in frequency and amplitude.

Steigerwald concludes that $C_p = C_s$ is a best-compromise design.



(a)



(b)

Figure 13.8 *(a) Gain curves for series-parallel half-bridge resonant converter (LCC) of Fig. 13.5c for $C_s = C_p$. (b) Gain curves for series-parallel half-bridge resonant converter (LCC) of Fig. 13.5c for $C_s = 2C_p$. (Courtesy R. Steigerwald)*

Additional analyses of the *LCC* resonant converter, discussed above, can be found in Refs. 6 to 9.

13.4.6 Zero-voltage-switching quasi-resonant (CCM) converters[2]

Zero-current-switching (ZCS) converters, discussed above, force the transistor current to be sinusoidal by having the square-wave drive from the transistor switches drive a resonant *LC* circuit. Further, by turning the transistor off as the sine wave of current passes through zero at the end

of the first half period (or shortly after when current has reversed and flows into the antiparallel diode as Fig. 13.1), there is no overlap of high voltage and current and turnoff losses are eliminated.

However, there are turnon losses even though there is no overlap of high voltage and current at turnon because the leakage inductance speeds up drain (or collector) voltage fall time and slows up drain current rise time. These losses occur because the drain capacity stores an energy (at turnoff) of $0.5C(V_{max})^2$ once per cycle and dissipates it in the transistor at the next turnon. This happens once per cycle and results in a dissipation of $0.5C(V_{max})^2/T$. This dissipation becomes significant over 500 kHz to 1 MHz, and as designers move to these higher frequencies to reduce size, this becomes a problem.

A new technique—ZVS—has been proposed to circumvent this problem. It has been proposed for single-ended (flyback or buck) circuits,[10] but its value is mainly for half bridges.

A ZVS half bridge is shown in Fig. 13.9.[2] Its basic principle is that the MOSFET output capacitor is made the capacitor of a resonant LC circuit. It stores a voltage (and hence energy) in one part of the switching cycle. In a following part of the switching cycle, the energy on the capacitor is discharged through the resonant inductor back into the supply bus without dissipation. The following discussion is based on Ref. 2.

Thus in Fig. 13.9, when $Q1$ is on and $Q2$ is off, $C2$ is charged up to V_s. Transistor $Q1$ is first turned off. Transformer $T1$ magnetizing current continues to flow through $C1$, pulling it down. Half way down, there is no voltage across the primary and in the $T1$ secondary, the output inductor tries to maintain a constant current. Both rectifier diodes latch in, which provides a short circuit across the secondary and hence across the primary.

At this point, the energy stored on the top end of $C2$ discharges through the short-circuited primary, through the resonant L, through

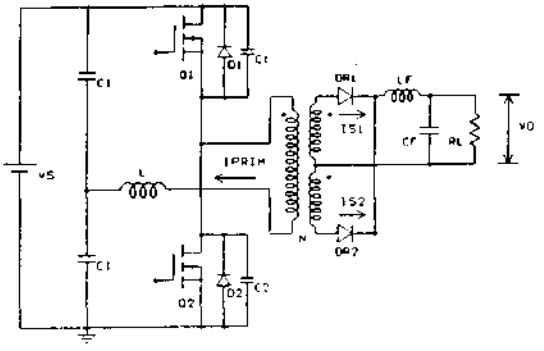


Figure 13.9 A zero-voltage-switching half-bridge resonant converter. *(Courtesy Jovanovic, Tabisz, Lee)*

the bottom filter capacitor C_{r2}, and back into the bottom end of $C2$. Since there are no resistors in this path, the discharge is lossless. The negative resonant voltage impulse on the right hand of L pulls the junction of $C1$, $C2$ down to ground, and now $Q2$ is turned on at zero voltage.

Capacitor $C1$ has slowed up $Q1$ voltage fall time sufficiently, so that there is no simultaneous high voltage and current during its turnoff.

The circuit, a CCM type, operates on the slope of the Q curve of the resonant circuit consisting of L and $C1$, $C2$ in parallel. Jovanovic et al. give the DC voltage conversion curve for the circuit (the equivalent of Figs. 13.5 to 13.8) in Fig. 13.10.[2]

Despite the many articles written on CCM, where regulation is by changing frequency along the steep slope of a resonant Q curve, this author feels it is not a reliable scheme. It will not yield similar results over a large production run of supplies with all the various production spreads in the resonant components and with the limited region of V_{dc} and R_l in the conversion curves such as those in Figs. 13.5 to 13.8 and 13.10.

13.5 Resonant Power Supplies—Conclusion

Despite the enormous number of articles and development work over the past decade or so in many industry and academic circles, there is

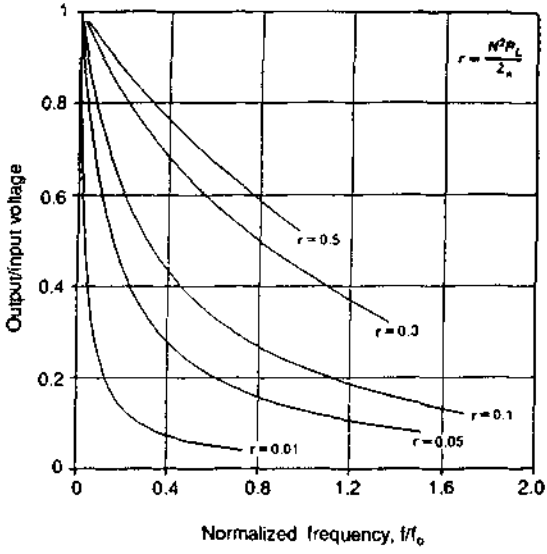


Figure 13.10 Gain curve for the zero-voltage-switching circuit of Fig. 13.9. (*Courtesy Jovanovic, Tabisz, Lee*)

no current industry-wide consensus on the future—or even the present—value of resonant power supplies.

It is of interest to consider the pros and cons on the subject. First, it is interesting to note that to the author's knowledge, resonant power supplies are not listed in any of the major manufacturers' catalogs.

This may indicate that resonant supplies are presently not sufficiently smaller, sufficiently more efficient, and sufficiently flexible in their ability to cope with large line and load changes. It may also indicate that they are not easily and inexpensively manufacturable for a variety of reasons. Possibly they cannot cope with the usual production spread in component values, parasitics in transformers, stray inductances in wiring, and greater susceptibility to problems in wire routing and component layout. It may indicate that it costs more in dollars per watt to produce and sell resonant power supplies.

To the engineer who must decide whether to consider a resonant supply for the next design, the preceding issues are important. The engineer must also ask:

May I do almost as well in watts per cubic inch with a conventional PWM supply operating at 200 to 300 kHz? Are the added complexity, limited line and load capability, and difficulty to design for worst-case conditions worth the extra 3 to 6 percent better efficiency possible with resonant supplies?

Will all units coming off a production line have identical characteristics? Or because of tolerances and production spread in component values, will fine-tuning of each unit or a large fraction of the production run be required?

Do resonant supplies generate less RFI actually because their currents are sinusoidal rather than square wave with their higher di/dt? Considering that most resonant supplies have sine wave currents three to four times the amplitude of PWM square-wave supplies of equal power, may not the RFI problem with resonant supplies be as severe since di/dt at the zero crossing of a sine wave is proportional to its peak value?

How serious is the problem that most resonant supplies regulate by varying frequency? Will users accept this or insist that the switching supply be synchronized to a clock signal which they will supply?

If the decision is to attempt a resonant design, which of the bewildering number of topologies advocated should be the safest approach? Are continuous-mode designs too unpredictable because they require operation on the slippery slope of a narrow resonant Q curve for regulation? There seems to be general agreement that discontinuous-mode operation is more predictable and reproducible. Is the larger frequency range required in discontinuous mode an important drawback?

There is no question that high frequency resonant converters will continue to be studied and improved. But, until circuit configurations are found which lend themselves to simple, worst-case designing and which are as insensitive as PWM circuits to component tolerances, wiring layout, parasitic inductances, and capacitances, high frequently resonant converters will not be widely adopted—certainly not in programs with a large production run. They will occupy only a narrow niche in the power supply field where higher cost, finely tuned component selection, and wiring layout is acceptable.

In summary, in switching power supplies, complexity equals higher cost and unreliability.

References

1. F. Lee, "High Frequency Quasi-Resonant and Multi Resonant Converter Topologies," *Proceedings of the International Conference on Industrial Electronics*, 1988.
2. M. Jovanovic, W. Tabisz, and F. Lee, "Zero Voltage Switching Technique in High Frequency Inverters," Applied Power Electronics Conference, 1988.
3. K. Liu and F. Lee, "Secondary Side Resonance for High Frequency Power Conversion," Applied Power Electronics Conference, 1986.
4. R. Steigerwald, "A Comparison of Half Bridge Resonant Topologies," *IEEE Transactions on Power Electronics*, 1988.
5. R. Severn, "Topologies for Three Element Resonant Converters," Applied Power Electronics Conference, 1990.
6. F. Lee, X. Batarseh, and K. Liu, "Design of the Capacitive Coupled LCC Parallel Resonant Converter," *IEEE IECON Record*, 1988.
7. X. Bhat and X. Dewan, "Analysis and Design of a High Frequency Converter Using LCC Type Commutation," *IEEE IAS Record*, 1986.
8. X. Batarseh, K. Liu, F. Lee, and X. Upadhyay, "150 Watt, 140 kHz Multi Output LCC Type Parallel Resonant Converter," IEEE APEC Conference, 1989.
9. B. Carsten, "A Hybrid Series-Parallel Resonant Converter for High Frequencies And Power Levels," *HFPC Conference Record*, 1987.
10. W. Tabisz, P. Gradski, and F. Lee, "Zero Voltage Switched Quasi-Resonant Buck and Flyback Converters," PESC Conference, 1987.
11. F. Lee, Ed., "High Frequency Resonant Quasi-Resonant and Multi-Resonant Converters," Virginia Power Electronics Center, 1989.
12. Y. Kang, A. Upadhyay, and D. Stephens, "Off Line Resonant Power Supplies," *Powertechnics Magazine*, May 1990.
13. Y. Kang, A. Upadhyay, and D. Stephens, "Designing Parallel Resonant Converters," *Powertechnics Magazine*, June 1990.
14. P. Todd, "Practical Resonant Power Converters—Theory and Application," *Powertechnics Magazine*, April–June 1986.
15. P. Todd, "Resonant Converters: To Use or Not to Use? That Is the Question," *Powertechnics Magazine*, October 1988.

Part

3

Typical Switching Power Supply Waveforms

Chapter

14

Waveforms

14.1 Introduction

In all previous chapters, voltage and current waveforms have been shown at critical points for the various topologies (e.g., Figs. 2.1, 2.10, 3.1, 4.1 for the major topologies).

These are idealized waveforms, and newcomers to switching power supply design may wonder how idealized these waveforms really are. They may wonder how closely actual waveforms resemble the theoretical ones on which much of the circuit design is based.

Newcomers may question how these waveshapes vary with line voltage and load current variations; whether there are noise spikes on ground buses; whether there is decaying, oscillatory, ringing waveform at sharp transitions in voltage and current; and whether there is time jitter at the leading or trailing edges of the waveforms. Other questions may be as follows: How closely does the on volt-second product equal the reset volt-second product in a transformer or an inductor? What does a leakage inductance spike really look like? Since output inductors and flyback transformers are designed to yield certain current waveshapes, how close are these actual waveshapes to the theoretical ones? Since much of the power transistor dissipation at high frequencies comes from simultaneous high voltage and current at turnoff and turnon, can this be observed on a fast time base? What sort of waveform oddities may be expected?

It is thus felt that it would be very instructive and provide a feeling of confidence, especially to designers seeing high-frequency switching waveforms for the first time, if some actual oscilloscope waveforms at critical points for some of the major topologies were presented.

The topologies selected are the major ones—the forward converter, the push-pull, and the flyback. Some of the critical waveforms for the buck regulator which show the importance of operating output induc-

tors in the continuous mode were shown in the oscilloscope photographs of Figs. 1.6 and 1.7.

The waveforms are taken mostly of points in the so-called power train—the circuitry from the input to the power transistors to the output of the output filters, as that is where most of the energy is handled and where most of the potential failures occur.

The selected circuits operate at switching frequencies above 100 kHz and are powered from telecommunications industry DC supply voltages—nominal 48 V, minimum 38 V, and maximum 60 V. Output powers in all cases are under 100 W as voltage-current waveforms at higher power change only in amplitude but not significantly in shape.

Off-line converters operating from AC voltages of 120 or 220 V will not be considered here. Supplies operating from those rectified voltages have similar waveshapes at corresponding points. But since the rectified DC supply voltages are higher, current amplitudes are lower and voltage waveshape amplitudes are higher than in telecommunications power supplies.

Since the waveshapes to be shown herein are from DC-powered circuits, they have somewhat less time and amplitude jitter than do those on circuits powered from rectified alternating current. The rectified DC in an off-line converter has line frequency ripple which will cause amplitude ripple. The feedback loop, in degenerating the input line ripple, causes time jitter in the pulse widths.

All waveshapes are taken on circuits operating above 100 kHz, use MOSFETs, and thus show no power transistor storage delay.

In all cases, the feedback loop was open and the power transistor(s) was (were) driven with a pulse of the desired frequency. But the pulse width was manually adjusted at each supply voltage input (38, 48, 60 V) so as to maintain the 5-V output within millivolts of 5 V. The slave output voltage was accepted and recorded at the pulse width, which made the 5-V output equal to 5.00 V—just as if the feedback loop were closed.

The power transistor driver for all photos is the UC3525 PWM control chip. Pulse widths were set by feeding the error-amplifier output node (which is available at one of the chip output pins) from a well-regulated DC voltage source adjusted to yield the pulse width required to set the +5-V output at 5.00 V at each DC input voltage.

14.2 Forward Converter Waveshapes

The circuit schematic for these waveshapes is shown in Fig. 14.1. It is a 125-kHz forward converter designed for 100 W, and waveshapes are shown at 80 and 40 percent of full load. Full-load outputs are 5 V at 10 A and 13 V at 3.8 A. Waveshapes are shown for nominal input volt-

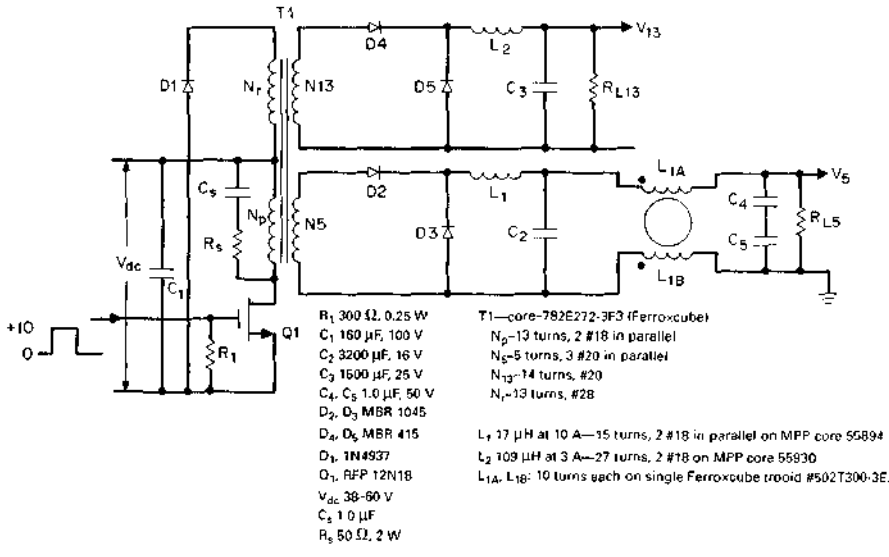


Figure 14.1 The 125-kHz forward converter— + 5.0 V at 10 A, +13 V at 3.8 A.

age of 48 V, minimum of 38 V, and maximum of 60 V. The transformer core was selected from Table 7.2*a* and the numbers of turns and wire sizes from Secs. 2.3.2 to 2.3.10. The output filters (*L*1, *C*2 and *L*2, *C*3) were chosen from the relations given in Sec. 2.3.11.

14.2.1 V_{ds}, I_d photos at 80 percent of full load

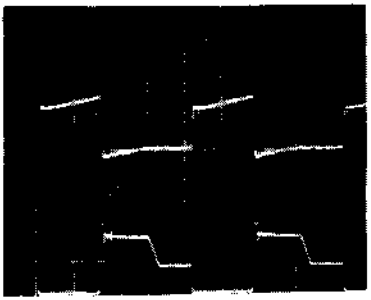
These photographs (photos 1 to 3 in Fig. 14.2) show drain-to-source voltages and drain currents at low, nominal, and maximum DC supply voltages.

Drain currents have the ramp-on-a-step waveshape characteristic of secondaries with output *LC* filters. Drain current is the sum of the secondary currents reflected by their turns ratios into the primary. Also, since both secondaries have output inductors which yield ramp-on-a-step currents (Sec. 1.3.2), these reflect into the primary as ramp-on-a-step currents.

Drain current amplitude at the center of the ramp should be (from Eq. 2.28) equal to $I_{pft} = 3.12P_o/V_{dc}$. For 80-W output and a minimum DC voltage of 38 V, peak current should be 6.57 A. Photos 1, 2, and 3 (Fig. 14.2) show that to be the current at the center of the ramp as accurately as can be read.

As DC supply voltage is increased, the photos show that pulse width (transistor on time) decreases but peak current and current at the ramp center remain unchanged—as theoretically they should.

Photo 1: V_{dc} = 38.0 V

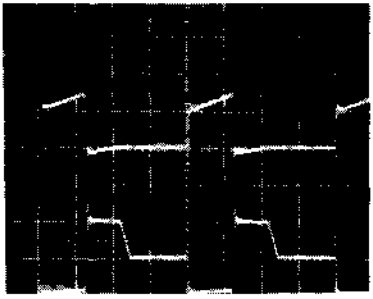


I_{ds} = 5 A, 2 μs/cm

V_{ds} = 50 V, 2 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_5, Ω	I_5, A	P_5, W	V_{13}, V	R_{13}, Ω	I_{13}, A	P_{13}, W	$P_{(total)}$, W	Efficiency, %
38.0	2.45	93.1	5.000	0.597	8.375	41.9	13.64	5.00	2.728	37.2	79.1	84.9

Photo 2: V_{dc} = 48.0 V

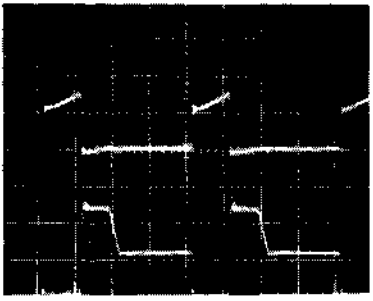


I_{ds} = 5 A, 2 μs/cm

V_{ds} = 50 V, 2 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_5, Ω	I_5, A	P_5, W	V_{13}, V	R_{13}, Ω	I_{13}, A	P_{13}, W	$P_{(total)}$, W	Efficiency, %
48.0	1.38	87.8	5.003	0.597	8.38	41.9	13.66	5.00	2.73	37.3	79.2	90

Photo 3: V_{dc} = 60.0 V



I_{ds} = 5 A, 2 μs/cm

V_{ds} = 50 V, 2 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_5, Ω	I_5, A	P_5, W	V_{13}, V	R_{13}, Ω	I_{13}, A	P_{13}, W	$P_{(total)}$, W	Efficiency, %
59.5	1.55	92.2	5.004	0.597	8.382	41.9	13.70	5.00	2.74	37.5	79.4	86.1

Figure 14.2 The 125-kHz 100-W forward converter of Fig. 14.1 at 80 percent full load.

Drain-to-source voltages also appear as they should be theoretically. Transistor on time at low line (V_{dc} = 38 V) is seen to be very close to 80 percent of a half period as discussed in Sec. 2.3.2. It is not always exactly that because of the inevitable rounding up or rounding down of fractional secondary turns to the nearest integral number. In this transformer, the calculated 4.5 turns on the secondary were rounded up to 5 turns. This yielded a larger peak secondary voltage and a somewhat shorter on time than called for by Eq. 2.25.

Narrow and barely discernible leakage spikes at the instant of turnoff are seen in photos 1, 2, and 3. At V_{dc} of 60 V, the leakage spike is only about 21 percent above $2V_{dc}$. But at V_{dc} = 38 V, it is about 64 percent above $2V_{dc}$.

The drain-to-source voltage waveshapes also show that V_{ds} at turnoff falls back down to $2V_{ds}$ immediately after the leakage inductance spike and remains there until the on volt-second product equals the reset volt-second product $(V_{dc}t_{on}) = (2V_{dc} - V_{dc})t_{reset}$. After those volt-second areas are equal, the drain voltage drifts back down to V_{dc}.

Figure 14.2 shows the average efficiency from V_{dc} = 38 to 60 V to be 87 percent. This is achieved at a peak flux density of 1600 G with Ferroxcube 3F3 core material. Core temperature rise was under 25°C.

Such high efficiency at 125-kHz and 1600-G peak flux density could not have been achieved with the very widely used, higher-loss core material of only 5 or so years ago—the 3C8. But even with the lower-loss 3F3 material, as discussed in Sec. 7.3.5.1, larger-sized cores could not operate at 1600 G at 125 kHz. Peak flux density would probably have to be reduced to the region of 1400 or possibly 1200 G.

14.2.2 V_{ds}, I_{ds} photos at 40 percent of full load

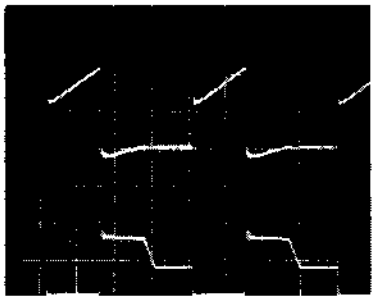
Photographs 4 to 6 in Fig. 14.3 give much the same information as do photos 1 to 3. At the lower output currents and power, efficiencies average 90 percent over the low, nominal, and maximum DC input voltages.

Leakage inductance spikes are considerably smaller, and transistor on times are slightly shorter as the forward drop in the output rectifier diodes is somewhat less as a result of the lesser output current. This increases the peak square-wave voltage at the cathode of the 5-V rectifier diode and permits a shorter on time to generate the 5 V of output.

14.2.3 Overlap of drain voltage and drain current at turnon/turnoff transitions

The simultaneous high voltage and current at turnon/turnoff transitions result in spikes of high power dissipation. Even though the high-dissipation spikes are very narrow (especially with MOSFETs), when

Photo 4
V_{dc} = 38.0 V

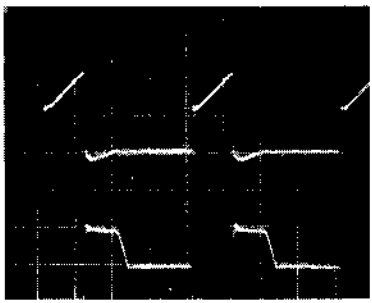


I_{ds} = 2 A, 2 μs/cm

V_{ds} = 50 V, 2 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_5, Ω	I_5, A	P_5, W	V_{13}, V	R_{13}, Ω	I_{13}, A	P_{13}, W	$P_{(total)}$, W	Efficiency, %
38.0	1.15	43.5	4.998	1.18	4.24	21.17	13.3	10.0	1.33	17.69	38.86	89

Photo 5
V_{dc} = 48.0 V

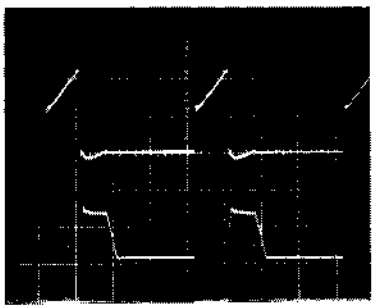


I_{ds} = 2 A, 2 μs/cm

V_{ds} = 50 V, 2 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_5, Ω	I_5, A	P_5, W	V_{13}, V	R_{13}, Ω	I_{13}, A	P_{13}, W	$P_{(total)}$, W	Efficiency, %
48.0	0.853	40.9	5.007	1.18	4.24	21.24	12.88	10.0	1.28	16.46	37.7	92

Photo 6
V_{dc} = 60.0 V



I_{ds} = 2 A, 2 μs/cm

V_{ds} = 50 V, 2 μs/cm

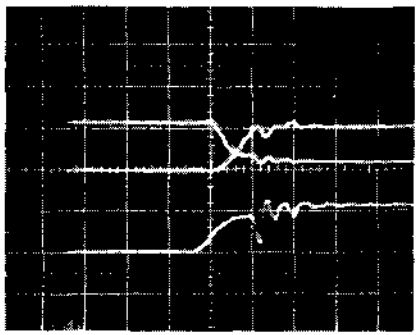
V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_5, Ω	I_5, A	P_5, W	V_{13}, V	R_{13}, Ω	I_{13}, A	P_{13}, W	$P_{(total)}$, W	Efficiency, %
59.6	0.736	43.8	5.014	1.18	4.49	22.5	13.2	10.0	1.32	17.5	40.0	91

Figure 14.3 The 125-kHz 100-W forward converter of Fig. 14.1 at 40 percent full load.

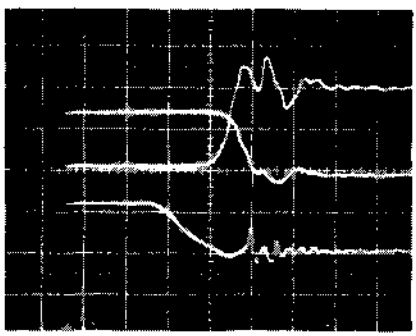
they come at a high repetition rate, their average dissipation can be high and can exceed the "conduction" dissipation of $V_{ds}I_{ds}t_{on}/T$.

The overlap dissipation at turnon is not as serious as at turnoff. At turnon, the power transformer leakage inductance presents an infinite impedance for a short time and causes a very fast drain-to-source voltage fall time. The same leakage inductance does not permit a very fast current rise time. Thus the falling V_{ds} intersects the rising I_{ds} quite low on the current waveform and the integral $\int V_{ds}I_{ds}\,dt$ taken over the turnon time is small and dissipation averaged over a full cycle is small. This can be seen on a fast time base of 0.1 μs/cm in photo 7 in Fig. 14.4.

At turnoff (photo 8 in Fig. 14.4), however, the drain current remains constant at its peak for a while (because leakage inductance tends to maintain constant current) as V_{ds} rises to about V_{dc}. Then V_{ds} contin-



Overlap of rising drain current and falling Vds voltage at turnon
at V_{dc} = 48 Volts, R_S = 0.59Ω, R_{13} = 5 Ω



Overlap of falling drain current and rising V_{ds} voltage at turnoff

Figure 14.4 Overlap of rising drain current and falling V_{ds} at turnon and falling drain current and rising V_{ds} at turnoff. Simultaneous high voltage and current during these turnon/turnoff transitions causes the AC switching losses. Losses are greater at the turnoff transition because current falls more slowly and V_{ds} rises more rapidly at turnoff.

ues to rise at a fast rate and reaches $2V_{dc}$ before I_d has fallen significantly below its peak. Thus, as seen in photo 8, the integral $\int V_{ds} I_d \, dt$ over the turnoff time is much greater than the same integral over the turnon time.

Performing an approximate "eyeball" integration of the above two integrals, it is seen that the average dissipation due to the overlap of falling current and rising voltage at turnoff is 2.18 W. At turnon, the average dissipation due to rising current and falling voltage is 1.4 W.

14.2.4 Relative timing of drain current, drain-to-source voltage, and gate-to-source voltage

Photo 9 in Fig. 14.5 is presented to show the negligible delay between the gate input voltage transitions and the drain voltage–drain current transitions at turnon and turnoff.

14.2.5 Relative timing of input voltage to output inductor, output inductor current rise and fall times, and power transistor drain-to-source voltage

Photo 10 in Fig. 14.6 is presented to show the output inductor upramp of current during the transistor on time and the downramp of current during the transistor off time. For L_1 = 17 μH, input voltage to the inductor of 16 V and on time of 2.4 μs (last two from Fig. 14.6), the ramp amplitude should be dI = (16 − 5)2.4/17 = 1.55 A. As accurately as photo 10 (Fig. 14.6) can be read, the ramp amplitude is 1.4 A. Because of the scales on the photo, this is the best correlation which can be obtained between the calculated and measured values.

Photo 9

Figure 14.5 Relative timing of drain current, drain-to-source voltage, and gate-to-source voltage. For 125-kHz forward converter of Fig. 14.1 at V_{dc} = 48 V, R_5 = 0.597 Ω, R_{13} = 5 Ω.

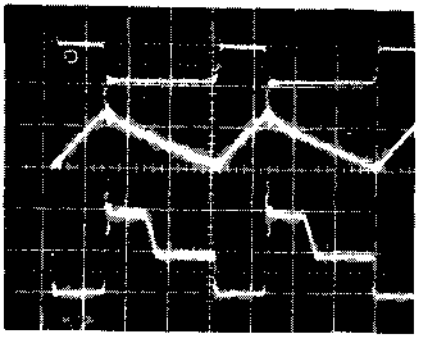


Photo 10

Figure 14.6 Relative timing input voltage to output inductor and its ripple current and drain-to-source voltage.

14.2.6 Relative timing of critical waveforms in PWM driver chip (UC3525A) for forward converter of Fig. 14.1

This demonstrates how the UC3525A PWM chip generates output pulses whose widths are inversely proportional to the DC supply voltage (see Fig. 14.7).

Internal to the chip, a 3-V peak-to-peak triangle (~0.5 to 3.5 V DC level) is generated and occurs once per half period of the switching frequency. This triangle is compared in a voltage comparator to the DC

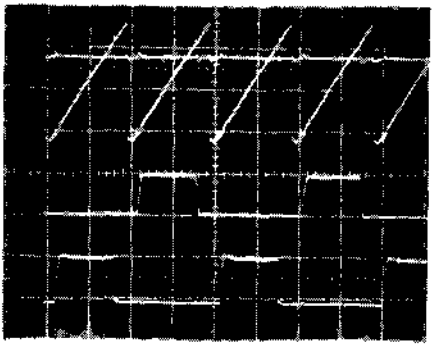


Photo 11

Figure 14.7 Significant waveforms in UC3525 pulse-width-modulating control chip. The PWM comparator internal to the chip compares an internally generated 3-V triangle to the DC voltage at its internal error-amplifier output node. The PWM comparator generates two 180° out-of-phase pulses at the chip *A* and *B* output pins. The width of these pulses is the time duration between the start of the triangle and the instant the triangle reaches the DC voltage level at the error-amplifier output node. As the error voltage output moves across the 3-V height of the triangle (in response to the difference between an internal reference voltage and a fraction of the output voltage being regulated), the output pulse widths at outputs *A* and *B* are varied.

voltage level at the output of the internal error amplifier. The error amplifier compares a fraction of the DC voltage to be regulated (at its inverting input) to a reference voltage.

The voltage comparator generates two 180° out-of-phase rectangular pulses. These pulses commence at the start of the triangle and terminate when the triangle crosses the DC voltage level at the error-amplifier output.

Thus, when the sampled fraction of the voltage to be regulated goes slightly positive, the error-amplifier output goes slightly negative, the triangle crosses the lowered DC voltage level earlier in time, and the PWM comparator output pulse widths decrease. Similarly, a decrease in sampled input to the error amplifier raises the error-amplifier output voltage slightly, the triangle crosses that higher DC voltage level later in time, and the pulse widths increase.

These adjustable-width pulses are alternately routed by a binary counter to two chip output pins for driving the transistors in a push-pull topology. For the forward converter of Fig. 14.1, only one of these pulse outputs was used.

14.3 Push-Pull Topology Waveshapes—Introduction

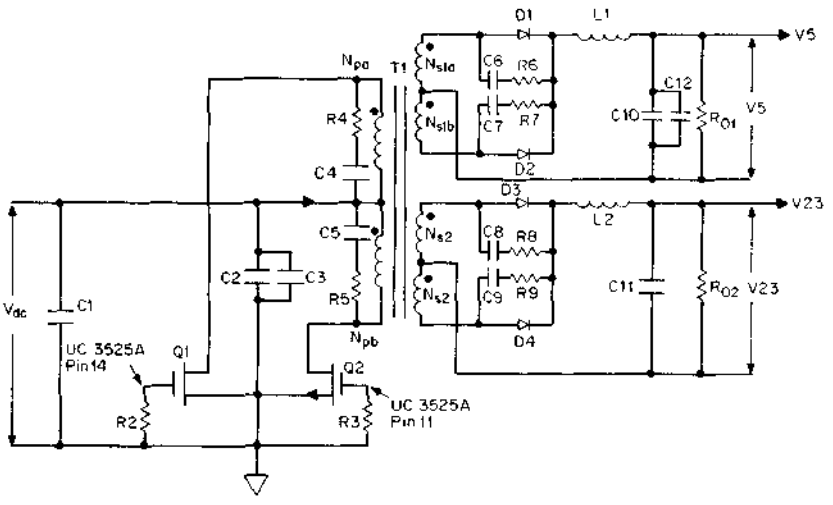
The next topology for which significant waveforms are presented is a push-pull whose circuit schematic is shown in Fig. 14.8.

The circuit is a 200-kHz DC/DC converter of 85-W maximum output power, and minimum of one-fifth of that. It was designed to operate from standard telecommunications industry supply voltages of 48 V nominal, 60 V maximum, and 38 V minimum. Outputs are +5 V at a maximum of 8 A and +23 V at a maximum of 1.9 A.

Significant waveforms are shown at maximum, nominal, and minimum supply voltages at maximum output currents and also at the same voltages at output currents closely one-fifth of maximum. Some waveforms are also shown at 100-W output power (15 percent above maximum) to demonstrate that at this power level, efficiency is still high, transformer temperature rise is acceptable, and waveshapes remain as expected.

For the design, the transformer core was selected from Table 7.2*a* and the transformer was designed (peak flux density selection, numbers of turns, wire sizes) from Secs. 2.2.9 and 2.2.10. The output filters were designed from the equations in Sec. 2.2.14.

A 200-kHz operating frequency was chosen arbitrarily as this is close to the currently maximum value of 300 to 400 kHz for square-wave converters. Above 200 kHz, although transformer and output filter sizes still decrease, efficiencies fall more rapidly and transformer



R_2, R_3 300 Ω 1/2 W
R_4 to R_9 24 Ω, 1 W

C_1 60 μF, 100 V
C_2 22 μF, 100 V
C_3 6800 pF
C_4 to C_9 1000 pF
C_{10} 3200 μF, 16 V
C_{11} 1000 μF, 50 V
D_1, D_2 MBR 1545
D_3, D_4 MUR415
Q_1, Q_2 Buz 31

L_1 1.65 mH (5 turns of 3 parallel No. 20 wire on Magnetics Inc. MPP core 55120
L_2 14.1 mH (14 No. 18 wire on Magnetics Inc. MPP core 55120

T1 Core: Ferroxcube 813E3433F3 1-mil gap

N_{pa} 6 turns of 2 paralleled No. 22, layer 1; Npb 6 turns of 2 paralleled No. 22, layer 4

N_{s1a} 1 turn of 4 paralleled No. 22, layer 2
N_{s1b} 1 turn of 4 paralleled No. 22, layer 2
} bipolar wound

N_{s2} 4 turns of 2 paralleled No. 24, wind bipolar, layer 3

Figure 14.8 A 200-kHz 85-W DC/DC converter: +5 V at 8.0 A, +23 V at 1.9 A.

core and copper losses increase more sharply. Thus, above 200 kHz, it is questionable whether the advantage of a small decrease in size is worth the penalty of increased dissipation and transformer temperature rise.

Here, as for the forward converter of Fig. 14.1, the feedback loop was not closed. At each DC input voltage, the pulse width was manually set (as described in Sec. 14.1) to yield 5.00 V at the 5-V output. The resulting slave secondary voltage was accepted and recorded at the pulse width which yielded 5.00 V at the 5-V output.

14.3.1 Transformer center tap currents and drain-to-source voltages at maximum load currents for maximum, nominal, and minimum supply voltages

Alternate current pulses, as monitored in the transformer center tap, correspond to alternate transistors ($Q1$, $Q2$) turning on.

These waveshapes are shown to demonstrate that with MOSFETs, having no storage times and with equal-width pulses at the two gate in-

puts, there is no sign of flux imbalance (Sec. 2.2.5). As discussed in Sec. 2.2.5, flux imbalance would show up as an inequality in amplitude of AC pulses monitored in the transformer center tap (as in Fig. 2.4*b* and 2.4*c*).

It is seen in photos PP1 to PP3 in Fig. 14.9 that alternate current pulses at any supply voltage are of equal amplitude (as closely as can be read in the photos). No attempt was made to match r_{ds} of MOSFETs $Q1$, $Q2$.

The V_{ds} waveshapes show negligible leakage inductance spikes at the end of the on times. The largest leakage inductance spike (at V_{dc} = 41 V) is only about 5 V above $2V_{dc}$ (photo PP5).

Negligible leakage inductance spikes result from the fact that at high frequencies, leakage inductance is minimal because the numbers of turns is small and coupling between primary and secondaries is better than would be possible at lower frequencies. Also, sandwiching secondaries between the two half primaries (Fig. 14.8) has helped reduce leakage inductance.

Photos PP1 to PP3 show that as supply voltage increases, on times must decrease to maintain a constant 5.00 output voltage. Peak currents remain constant for constant DC output currents. It is only pulse widths that change as supply voltage changes.

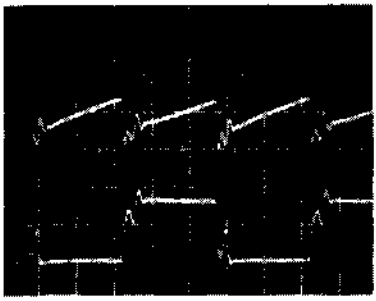
The primary currents are seen to have the characteristic shape of a ramp on a step. They have this shape because they are the sum of the ramp-on-a-step secondary currents, reflected into the primaries by the respective turn ratios, plus the primary magnetizing current which is a triangle (Fig. 2.4*e*). The secondary currents have ramp-on-a-step waveshape because there are inductors in all outputs.

It should, however, be noted that although the equal current peaks in photos PP1 to PP3 indicate equality of peak currents in $Q1$, $Q2$ during their on times, those photos give a fictitious picture of the absolute value of the transistor currents (especially during the dead time between turnons). This comes about because of the nature of the magnetically coupled current probe used to monitor currents in the transformer center tap.

Apparently because of the short dead time between alternate transistor turnons, the current probe does not have time to recover to its original starting point on its hysteresis loop and hence gives a false picture of absolute current.

A true measure of absolute current amplitude in each transistor is obtained by monitoring current with the same probe in series with each transistor drain as in photo PP6 in Fig. 14.10. In that photo, it is seen with the longer time between current pulses (more than a half period), the peak current is 4.4 A. This compares to 2.4 A in photo PP2 (in Fig. 14.9), which is taken at the same supply voltage and output load currents.

Photo PP1
V_{dc} = 41 V

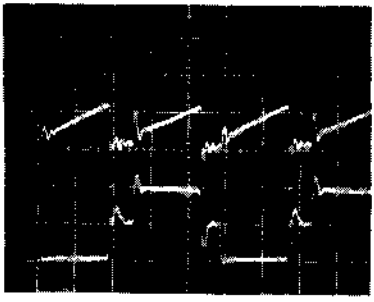


$I_{center\ tap}$ 2 A, 1 μs/cm

$V_{ds(Q2)}$, 50 V, 1 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_S, V	R_{01}, Ω	I_S, A	P_S, W	V_{23}, V	R_{02}, Ω	I_{23}, A	P_{23}, W	$P_{(total)}$, W	Efficiency, %
41.0	2.46	100.9	5.00	0.615	8.09	40.3	23.55	12.46	1.91	44.9	85.2	84.4

Photo PP2
V_{dc} = 48 V

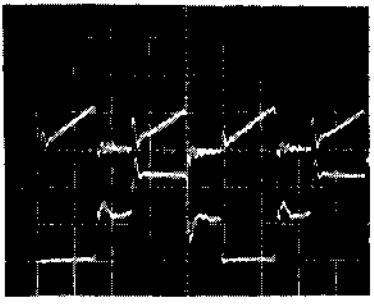


$I_{center\ tap}$ 2 A, 1 μs/cm

$V_{ds(Q2)}$, 50 V, 1 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_S, V	R_{01}, Ω	I_S, A	P_S, W	V_{23}, V	R_{02}, Ω	I_{23}, A	P_{23}, W	$P_{(total)}$, W	Efficiency, %
48.0	2.16	103.7	5.00	0.615	8.13	40.6	23.74	12.46	1.91	45.2	85.8	82.7

Photo PP3
V_{dc} = 58.9 V



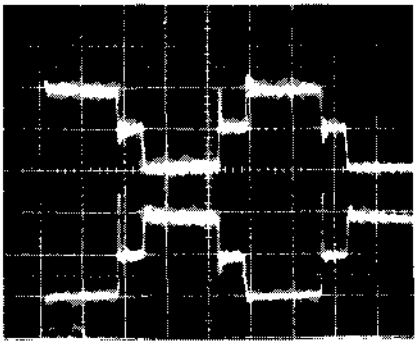
$I_{center\ tap}$ 2 A, 1 μs/cm

$V_{ds(Q2)}$, 50 V, 1 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_S, V	R_{01}, Ω	I_S, A	P_S, W	V_{23}, V	R_{02}, Ω	I_{23}, A	P_{23}, W	$P_{(total)}$, W	Efficiency, %
58.9	1.79	195.4	5.00	0.615	8.13	40.6	23.72	12.46	1.90	45.1	85.7	81.3

Figure 14.9 Transformer center tap current and drain-to-source voltage (Q2) at minimum (photo PP1), nominal (photo PP2), and maximum (photo PP3) input voltage for maximum output currents.

Photo PP4
V_{dc} = 48.0 V

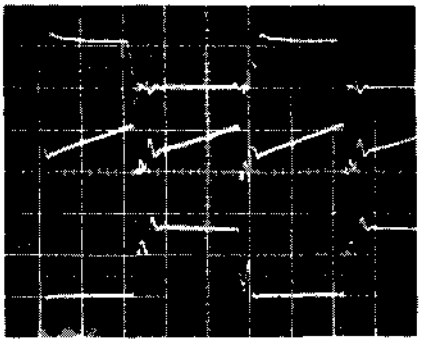


$V_{ds(Q1)}$, 50 V, 1 μs/cm

$V_{ds(Q2)}$, 50 V, 1 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	I_5, A	P_5, W	V_{23}, V	I_{23}, A	P_{23}, W	$P_{o\ (total)}$, W	Efficiency, %
48.0	2.17	104.1	5.00	8.13	40.6	23.73	1.90	45.1	85.7	82.3

Photo PP5
V_{dc} = 41.0 V



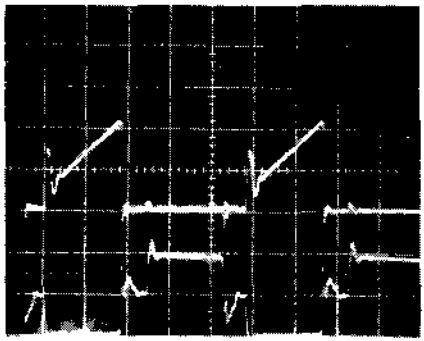
$V_{gs(Q1)}$, 10 V, 1 μs/cm

$I_{center\ tap}$ 2 A, 1 μs/cm

$V_{ds(Q1)}$, 50 V, 1 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	I_5, A	P_5, W	V_{23}, V	I_{23}, A	P_{23}, W	$P_{o\ (total)}$, W	Efficiency, %
41.0	2.46	100.9	5.00	8.13	40.6	23.55	1.19	44.9	84.9	84.1

Photo PP6
V_{dc} = 48.0 V



$I_{ds(Q2)}$, 2 A, 1 μs/cm

$V_{ds(Q2)}$, 50 V, 1 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	I_5, A	P_5, W	V_{23}, V	I_{23}, A	P_{23}, W	$P_{o\ (total)}$, W	Efficiency, %
48.0	2.19	105.1	5.00	8.13	40.7	23.79	1.91	45.4	86.1	81.9

Figure 14.10 Significant waveforms in 200-kHz 85-W converter of Fig. 14.8.

The assumption that the amplitude measured with a current probe in the drain as for photo PP6 is more valid than the measurement with the same probe in the transformer center tap (as photos PP1 to PP3) is verified by measuring voltage drop across a small current-monitoring resistor in series in the transistor source. Measurement with a current probe in series in the drain gives exactly the same absolute currents as does measuring the voltage drop across a known resistor in the source lead.

Figure 14.10 shows efficiency to exceed 81.9 percent at any supply voltage for maximum current in the two output nodes. If it were seriously attempted, efficiency could be raised by 3 to 4 percent by going to larger wire size and perhaps decreasing peak flux density. But the efficiency achieved is quite good for an operating frequency of 200 kHz and a peak flux density of 1600 G.

14.3.2 Opposing V_{ds} waveshapes, relative timing, and flux locus during dead time

Photo PP4 (in Fig. 14.10) shows the relative timing of the $Q1$, $Q2$ drain voltages. This is classic and just what would be expected. As one transistor turns on, its drain voltage falls closely to zero and the other drain rises to $2V_{dc}$.

Delays between turnon of one drain and turnoff of the other are seen to be negligible. The highest leakage inductance spike is seen to rise about 20 V above $2V_{dc}$ (Fig. 14.17, photo PP24). Note that for this photo, total output power was increased to 112 W (5 V at 4.85 A and the slave at 21 V, 4.05 A).

Note that at the start of dead time between turnons, after one transistor turns off, there is a leakage inductance spike and immediately thereafter, drain voltage falls back to V_{dc}. It remains at V_{dc} until the opposite transistor turns on, driving it to $2V_{dc}$.

It is important to understand this as it is basic to the operation of a push-pull (also to a half or full bridge, for that matter).

One may ask: What is the transformer core's flux density during the dead time between transistor turnons? At the end of a turnon, the flux density has been driven through a change of $2B_{max}$—say, up from some B_{max} to $+B_{max}$. During the dead time, neither transistor is on. Hence, does the flux density remain at $+B_{max}$ or fall back to remanence (Fig. 2.3) at 0 Oe?

If flux density fell back to remanence (~ 1000 G from Fig. 2.3), at the end of next turnon, an application of the same $2B_{max}$ change would drive the flux density to a peak of $(1000 + 2B_{max})$. This would saturate the core and destroy the transistors.

Thus, obviously, at the end of a transistor turnon, the core flux density, having been driven up to—say— $+ B_{max}$, remains locked there

throughout the dead time. At the end of the dead time, the opposite transistor turns on and drives the flux density down from $+B_{max}$ to $-B_{max}$ and the cycle repeats.

One may then ask: Since neither transistor is on during the dead time, where is the current coming from to keep the flux density locked at $+B_{max}$ or $-B_{max}$? To keep the core there, there must be some oersteds holding it up—and since oersteds are proportional to ampere turns, there must be some current flowing to hold them up during the dead time.

Obviously, since both transistors are off during the dead time, the current holding the core up at $+B_{max}$ or $-B_{max}$ must be flowing in the secondaries. This current is the core primary magnetizing current reflected into the secondaries.

During the transistor on time (see Fig. 2.4*e*), the DC input voltage supplies primary load current (all the secondary currents reflected by their turns ratios into the primary) plus magnetizing current which flows into the primary magnetizing inductance.

However, since current in an inductor cannot change instantaneously, when the on transistor turns off, that stored current or stored energy must continue elsewhere. The magnetizing current then simply continues to flow where it finds a closed path.

That closed path is in the secondaries. As a transistor turns off, the current in each secondary output inductor cannot change. Thus all output inductors reverse polarity at the end of an on time. If there were free-wheeling diodes at the output as for forward converters (see Fig. 2.10), the output inductor current would continue to flow through those diodes.

Since there are no free-wheeling diodes in a push-pull, the output rectifiers serve a similar function (Sec. 2.2.10.3). As the output inductors reverse polarity, when they reach a point one diode drop below ground (or above for a negative output voltage), the rectifier diodes latch in and serve as free-wheeling diodes carrying the output inductor current.

However, the diodes carry more than the output inductor current. By flyback action, the magnetizing current built up in the primary during the on time is reflected by the turns ratio into one of the secondary halves (the one which just previously had not been carrying current). It is this current flowing in one of the secondary halves which keeps the core locked up at B_{max} during the dead time.

Thus, during the dead time, the output inductor current continues to flow through the half-secondary windings and the rectifier diodes continue acting as free-wheeling diodes. That current divides roughly equally as "ledge" currents (Fig. 2.6*d* and 2.6*e*) between the two half secondaries as described in Sec. 2.10.3. But one of those ledge currents

is always larger than the other, as can be seen dramatically in photo PP21 (Fig. 14.15).

Thus primary magnetizing current does not result in dissipation. It simply increases in one direction during one transistor on time, is stored (and decreases slightly) in the secondary during the dead time, and reverses direction in the primary during the next transistor on time, repeating the previous half cycle.

As discussed, during the dead time, output inductor current divides roughly equally between the two half secondaries. But since those half secondaries have low impedance, current flowing through them produces no voltage drop across them. Hence there is no voltage drop across either of the two half primaries. Thus voltage at the two off drains during the dead time must equal V_{dc} as seen in photos PP4 and PP1 to PP3.

14.3.3 Relative timing of gate input voltage, drain-to-source voltage, and drain currents

This timing is shown in photo PP5 (Fig. 14.3). It shows negligible delays between gate input voltage rise and fall times and the corresponding drain-current–drain-voltage transitions.

14.3.4 Drain current as measured with a current probe in series in the drain as compared with measurement with a current probe in series in the transformer center tap

As discussed in Sec. 14.3.1, drain current measured with a current probe in series in the drain as for photo PP6 gives a correct measure of absolute drain current.

However, when—as in photo PP1—measured with the current probe in series with the transformer center tap (where both transistor currents are seen), the short transistor dead time between alternate on times does not permit the flux in the current probe transformer to reset. Consequently in such a measurement, the current indicated during the dead time is not zero and absolute values of current cannot be obtained from it.

14.3.5 Output ripple voltage and rectifier cathode voltage

Photo PP7 (in Fig. 14.11) shows a 5-V output ripple voltage and noise. It is seen to be about 80 mV peak to peak. Such measurements are often difficult to make and often fictitious. They are frequently masked by common-mode noise. Common-mode noise voltage is volt-

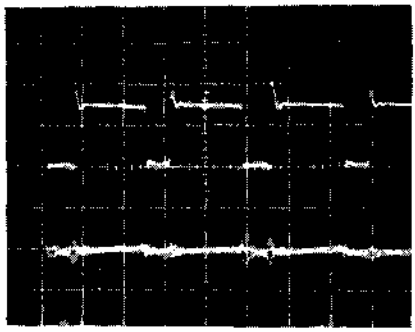


Photo PP7
V_{dc} = 48.0 V

V (cathodes, D1 D2) 5 V, 1 μs

5 V output ripple,
100 mV, 1 μs/cm

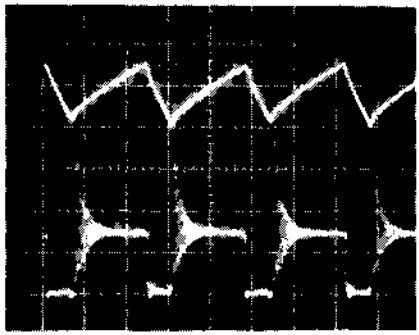


Photo PP8
V_{dc} = 48.0 V

$I_{(L1)}$ 1 A, 1 μs/cm

V (cathodes, D1, D2)
5 V, 1 μs/cm

With snubbers R6, C6 and R7, C7 out to
show ringing at transistor turnon

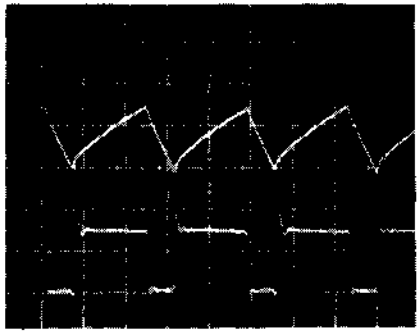


Photo PP9
V_{dc} = 48.0 V

$I_{(L1)}$ 1 A, 1 μs/cm

V (cathodes, D1, D2)
5 V, 1 μs/cm

With snubbers R6, C6 and R7, C7 in to
show ringing at transistor turnon is
damped out by snubbers

Figure 14.11 Significant waveforms in 200-kHz converter of Fig. 14.8.

age simultaneously common to the negative and positive rail across which the measurement is being made.

Such common-mode noise can cause problems at the loads and can be minimized by a common-mode filter or *balun* as shown in Fig. 14.1.

To determine whether noise voltage is truly differential or common-mode, the "hot" end of the voltage probe is short-circuited to its short-

est ground lead and those two points are then touched to the return rail of the output voltage.

If the oscilloscope still indicates almost the same large noise voltage (as it most often will), that noise is common-mode noise. It will change in amplitude as oscilloscope ground connections are changed to various points on the power supply ground rail and as grounding lead lengths are changed.

Output ripple voltage measurements should be made with a differential probe which has a good common-mode rejection ratio at high frequencies. With the fast rise and fall times of MOSFETs, the common-mode "ringing" or noise on ground buses can occur at frequencies over 50 MHz.

Photo PP7 also shows the voltage at the 5-V output rectifier cathodes. It is this voltage which is averaged by the output LC filter to yield the desired DC output voltage.

If this waveform has notches, bumps, or odd ledges along its supposedly vertical sides, those contribute area to the voltage being averaged. The feedback loop will then alter the transistor on time so that the averaged volt-second area at the rectifier cathodes yields the desired DC voltage.

Thus the master output voltage will always be correct regardless of whether there are odd bumps, notches, or ledges during the dead time or along the vertical sides of the rectifier cathode waveshapes. But then any slave rectifier cathode voltage may not have the same proportion of extraneous bumps or notches as the master and the slave DC output voltage differ from what would be expected from the relative turns ratio.

For the slave DC output voltages to be what is expected from their turns ratios, the voltages at the rectifier cathodes of the master and all slaves should have steep vertical sides (photos PP7, PP9, Fig. 14.11) and no bumps, notches, or bumps during the dead time which alter their volt-second area. Examples of such aberrations of the rectifier cathode voltages are shown in photos PP11 (in Fig. 14.12) and PP18 (Fig. 14.14). These waveform aberrations will be explained below.

14.3.6 Oscillatory ringing at rectifier cathodes at transistor turnon

This is shown in photos PP8 and PP10 (Figs. 14.11, 14.12) and its elimination in photos PP9 and PP11 (Figs. 14.11, 14.12).

At the instant of transistor turnon, the on-turning rectifier diode (say, diode $D1$ in Fig. 14.8) cancels the free-wheeling current in the opposite diode—say, $D2$. As the $D2$ forward current is canceled and its cathode voltage starts rising, there is an exponentially decaying oscil-

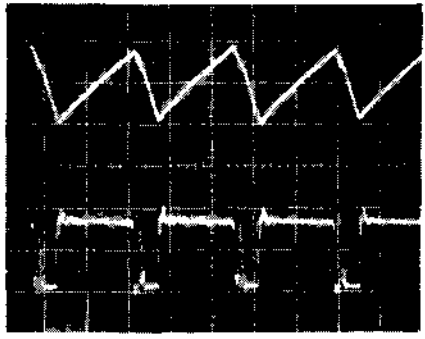


Photo PP10
V_{dc} = 48.0 V

$I_{(L2)}$ 500 mA, 1 μs/cm

V (cathodes, D3 D4)
20 V, 1 μs/cm

With snubbers R8, C8 and R9, C9 out to show ringing at transistor turnon

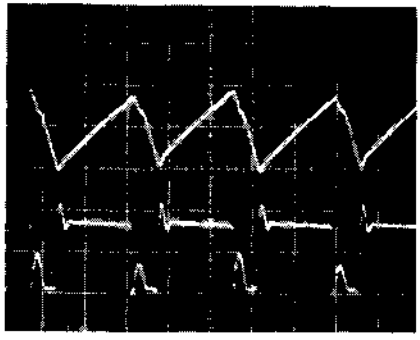


Photo PP11
V_{dc} = 48.0 V

$I_{(L2)}$ 500 mA, 1 μs/cm

V (cathodes, D3, D4)
20 V, 1 μs/cm

With snubbers R8, C8 and R9, C9 in to show ringing at transistor turn on is damped out by snubbers

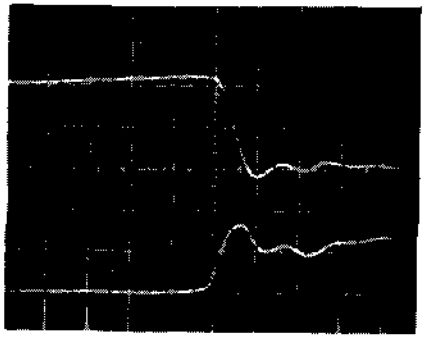


Photo PP12
V_{dc} = 48.0 V

Falling current in Q2 at turnoff
2 A, 50 ns/cm

Rising drain-to-source voltage (Q2) at turnoff 50 V, 50 ns/cm

To show overlap of falling current and risingV_{ds} voltage at turnoff; even with the fast I_{ds} fall time of MOSFETs, this overlap yields AC switching losses

Figure 14.12 Significant waveforms in 200-kHz converter of Fig. 14.8.

lation or "ring" at the common cathodes as seen in photos PP8 and PP10.

The oscillation is at a frequency determined by the inherent capacity of the off-turning diode *D*2 and the value of the output inductor. The amplitude and duration of the ring are determined by the rectifier diode reverse recovery times and DC output current.

The ring can cause RFI problems, drive the rectifier diodes too close to their maximum reverse voltage rating, and increase their dissipation. The oscillation can easily be eliminated by *RC* snubbers (*R*6, *C*6; *R*7, *C*7; *R*8, *C*8; and *R*9, *C*9) across the diodes as shown in Fig. 14.8. Cathode waveforms before the snubbers were added are shown in photos PP8 and PP10 and after the snubbers were added, in photos PP9 and PP11.

14.3.7 AC switching loss due to overlap of falling drain current and rising drain voltage at turnoff

This is shown in photo PP12 (Fig. 14.12). Because MOSFETs have negligible storage and very fast current turnoff time, this is close to the best-case scenario described in Sec. 1.3.4.

In that scenario, current starts falling at the very same instant the drain voltage starts rising, and current has fallen to zero at the same instant voltage has risen to its maximum. As mentioned in Sec. 1.3.4, for this case, the AC switching loss averaged over the current fall time is $\int_0^t IV\,dt = I_{max}V_{max}/6 = 4.2 \times 85/6 = 59.5$ W. For a current fall time of about 40 ns at a switching period of 5 μs, the AC switching loss averaged over a full cycle is only 59.5 × 0.04/5 or 0.48 W.

14.3.8 Drain currents as measured in the transformer center tap and drain-to-source voltages at one-fifth of maximum output power

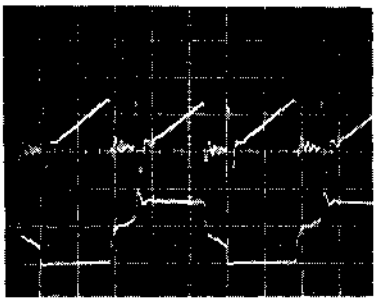
Waveshapes for output currents of one-fifth of maximum value (a usual power supply specification for minimum output currents) are shown in photos PP13 to PP15 (in Fig. 14.13).

Efficiencies are still seen in Fig. 14.13 to be close to 80 percent. For the worst efficiency of 78.7 at V_{dc} of 59.8, total internal losses are only 4.3 W—which is quite good.

These last photos and the ones to follow point out an interesting and subtle problem. The problem is not at all a catastrophic failure mode but one which may cause slave voltages to depart significantly from specified values.

The problem arises from too large a transformer primary magnetiz-

Photo PP13
V_{dc} = 41.0 V

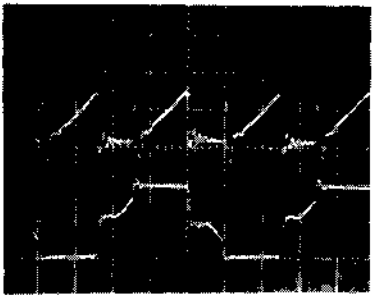


$I_{center\ tap}$ 500 mA, 1 μs/cm

$V_{ds\ (Q2)}$ 50 V, 1 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_{01}, Ω	I_5, A	P_5, W	V_{23}, V	R_{02}, Ω	I_{23}, A	P_{23}, W	$P_{(total)}$, W	Efficiency, %
41.0	0.52	21.3	5.00	3.05	1.64	8.20	21.46	50.0	0.429	9.21	17.4	81.7

Photo PP14
V_{dc} = 48.0 V

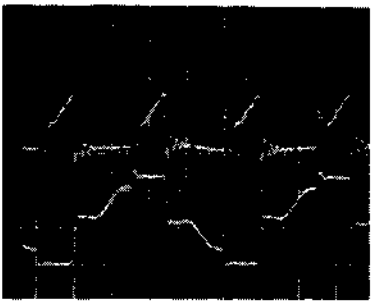


$I_{center\ tap}$ 500 mA, 1 μs/cm

$V_{ds\ (Q2)}$ 50 V, 1 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_{01}, Ω	I_5, A	P_5, W	V_{23}, V	R_{02}, Ω	I_{23}, A	P_{23}, W	$P_{(total)}$, W	Efficiency, %
48.0	0.44	21.1	5.00	3.05	1.64	8.20	21.52	50.0	0.430	9.25	17.5	82.7

Photo PP15
V_{dc} = 59.8 V



$I_{center\ tap}$ 500 mA, 1 μs/cm

$V_{ds\ (Q2)}$ 50 V, 1 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_{01}, Ω	I_5, A	P_5, W	V_{23}, V	R_{02}, Ω	I_{23}, A	P_{23}, W	$P_{(total)}$, W	Efficiency, %
59.8	0.37	22.1	5.00	3.05	1.64	8.20	21.88	50.0	0.438	9.58	17.8	78.0

Figure 14.13 Transformer center tap current and drain-to-source voltage ($Q2$) at minimum (photo **PP13**), nominal (photo PP14), and maximum (photo PP15) input voltage for one-fifth of maximum output currents.

ing current or for a given maximum magnetizing current, from too low a total DC output current. The magnetizing current can become larger than originally specified if the two transformer halves inadvertently separate slightly, thus decreasing the magnetizing inductance and increasing the magnetizing current; or possibly if too large a transformer gap was used to achieve a desired magnetizing inductance; or if the minimum load current is made lower than the value originally specified when the magnetizing inductance was selected.

A clue to the problem can be seen in the DC voltage levels at the drain during the transistor dead times in photos PP13 to PP15. In those photos, the drain voltage during the dead time (which should be firmly at V_{dc} as in photo PP4) starts at V_{dc} but gradually pulls above it before the end of the dead time.

This drain voltage pulling above V_{dc} during the dead time is not the cause of the problem, but only another manifestation of the basic cause. The basic cause is seen in photo PP18 (in Fig. 14.14), which shows the voltage at the cathodes of the 5-V output rectifiers. There it is seen that during the dead time, the rectifier cathodes do not remain clamped to ground as they should be, but gradually pull off ground. This is the bump or ledge discussed in Sec. 14.3.5.

It is this voltage at the rectifier cathodes which is averaged by the output *LC* filter to yield the 5-V DC output voltage. If that voltage has such a bump in the dead time before the normal transistor turnon time, its volt-second area is increased. Thus the feedback loop (which is forcing the 5-V output to be precisely 5.00 V) will decrease the normal on time. Then, since the slave outputs do not have this increased volt-second area due to a bump at their rectifier cathodes, their DC output voltage will decrease.

This can be seen in a comparison of photos PP9 (Fig. 14.11) and PP18 (Fig. 14.14). For PP9, the rectifier cathodes have no bump during the dead time and the voltage rises vertically from ground at the start of the normal on time. For this case, which is at maximum load current in both outputs, the "23"-V output is 23.74 at V_{dc} of 48.0 V (photo PP2, Fig. 14.1). For photo PP18, at the same V_{dc} of 48 V and at minimum load currents, when the 5-V output is 5.00 V, the "23"-V output is 21.52 V (tabular data for photo PP14 in Fig. 14.13).

The final question to be answered is why this bump during the dead time occurs for too high a magnetizing current or too low a DC output current. This can be understood from the discussion in Sec. 14.3.2 as follows.

As the on transistor turns off, some fraction of the total primary magnetizing current, multiplied by the turns ratio, continues flowing in one of each of the half secondaries. But rectifier diodes of all secondaries are now free-wheeling and carrying about half their associ-

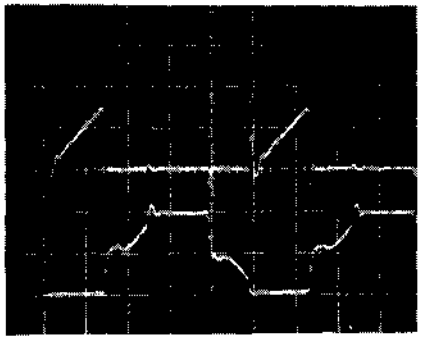


Photo PP16
V_{dc} = 48.0 V

$I_{ds\ (Q2)}$ 1A, 1 μs/cm

$V_{ds\ (Q2)}$ 50 V, 1 μs/cm

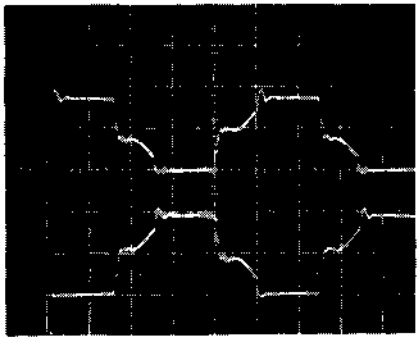


Photo PP17
V_{dc} = 48.0 V

$V_{ds\ (Q1)}$ 50 mA, 1 μs/cm

$V_{ds\ (Q2)}$ 50 V, 1 μs/cm

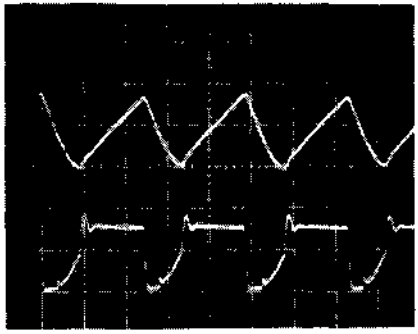


Photo PP18
V_{dc} = 48.0 V

$I_{(L1)}$ 1A, 1 μs/cm

V (cathodes, D1, D2)
5 V, 1 μs/cm

Figure 14.14 Significant waveforms in 200-kHz converter of Fig. 14.8 at minimum (photo PP16), nominal (photo PP17), and maximum (photo PP18) input voltages for one-fifth of maximum output current.

ated output inductor currents. In each secondary, one rectifier diode also carries a portion of the primary magnetizing current which has been reflected into that secondary.

Now in Fig. 14.8, consider the 5-V output rectifiers. Assume that $Q1$ is on. This makes the $D2$ anode positive and $D2$ delivers current to the load via $L1$. When $Q1$ turns off, the dead time starts and the $L1$ current is roughly equally divided between the $D1$ and $D2$ acting as free-

wheeling diodes. So long as the sum of the $D1 + D2$ currents is equal to the $L1$ current, their cathodes remain clamped one diode drop below ground.

However, as $T1$ turns off, some fraction of the total primary magnetizing current is transferred by flyback action into N_{s1a} and $D1$ into $L1$. So long as this current is less than the current in $L1$, the $D1$, $D2$ cathodes remain clamped one diode drop below ground and the balance of the $L1$ current is supplied via $D1$, $D2$.

But when the current reflected by flyback action from the primary into N_{s1a} exceeds more than the current in $L1$, the impedance at the input end of $L1$ becomes high and the common cathodes of $D1$, $D2$ pull up from ground before the end of the dead time as seen in photo PP18.

Now with a closed feedback loop, the increased volt-second area due to the bump at the rectifier cathodes during the dead time causes the PWM chip to decrease the on time so that the voltage averaged by the LC filter yields the desired 5.00-V output. This decreased on time then results in a lowered DC voltage for the slave outputs.

14.3.9 Drain current as measured in drain lead and drain voltage at one-fifth of maximum output power

This is shown in photo PP16 (Fig. 14.14). Note that with the current probe in the transistor drain lead, a true measure of drain current is obtained (peak of 1.5 A). When drain current is measured in the transformer center tap as for photo PP14 (Fig. 14.13), a false value of 700 mA is observed. This has been discussed in Secs. 14.3.1 and 14.3.4.

14.3.10 Relative timing of opposing drain voltages at one-fifth of maximum output currents

This is shown in photo PP17 (Fig. 14.14). Nothing new here. It is presented to show relative timings.

14.3.11 5-V output inductor current and rectifier cathode voltage

The inductor current waveform in photo PP18 shows the upslope of inductor current during the transistor on time $[di/dt = (V_{\text{cathode}} - V_o)/L_1]$ and the downslope of its current during the dead time $(di/dt = V_o/L_1)$.

It is of interest to calculate $L1$ to verify that the inductor is as designed. Thus as closely as can be read from the cathode voltage waveform, $V_{\text{cathode}} = 7.5$ V, and from the inductor current waveform

during the upslope di is 1.8 A in 1.45 μs. Then $L_1 = (7.5 - 5)1.45 \times 10^{-6}/1.8 = 2.0$ μH.

The inductor has 5 turns on an MPP 55120 core (see Magnetics Inc. MPP core catalog) which has an A_l of 72 mH/1000 turns at the low-oersted level of PP18. Then a 5-turn winding should have an inductance of $(0.005)^2 \times 72000 = 1.8$ μH. This is as close a correlation as can be expected from reading di/dt from photo PP18.

14.3.12 5-V rectifier cathode voltage at output current higher than minimum

This is shown in photo PP19 (in Fig. 14.15). It is presented to show that when DC output current is increased, the voltage ledge during the dead time shown in photo PP18 vanishes. The voltage remains clamped at ground throughout the dead time and rises steeply to its peak at transistor turnon. As a consequence (as discussed in Sec. 14.3.8, the "23"-V output voltage rises from 21.50 V for photo PP18 to 22.97 V for photo PP19.

14.3.13 Gate input and drain current timing

This is presented (see photo PP20 in Fig. 14.15) to show the relative timing of both gate input voltages and both drain currents. Delays between gate voltage transitions and corresponding drain current rise and fall times are seen to be negligible.

14.3.14 Rectifier diode and transformer secondary currents

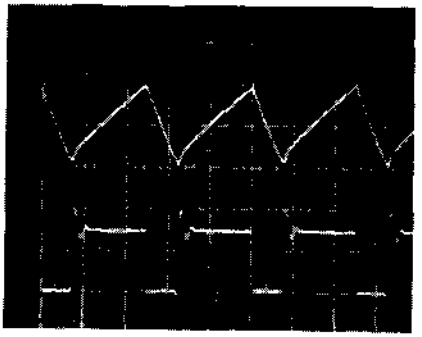
This has been discussed in Sec. 14.3.2 (see also photo PP21 in Fig. 14.15). It shows the ledge currents during the transistor dead time. Note that the ledge current immediately after an on time is greater than that before an on time. This is because the previously on diode carries the primary magnetizing current plus half the output inductor current during the dead time immediately after turnoff.

14.3.15 Apparent double turnon per half period arising from too high a magnetizing current or too low DC output currents

This is shown in photo PP22 (in Fig. 14.16). It is an extreme example of the case of photo PP13 (Fig. 14.13). It is seen in photo PP22 that there are apparently two turnons per half period—at $A1$ and $A2$ for $Q1$ and $B1$ and $B2$ for $Q2$.

Actually only at $A2$ and $B2$ do the transistors turn on and bring the

Photo PP19
V_{dc} = 48.0 V

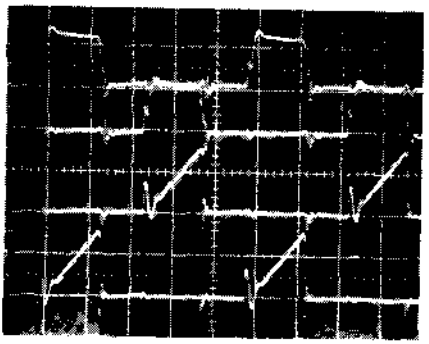


$I_{(L1)}$ 1A, 1 μs/cm

V (cathodes)
5 V, 1 μs/cm

For I_5 = 5 A, I_{23} = 0.425 A
Showing the D1, D2 cathode voltage remaining clamped down to ground during the transistor off time for larger output currents because the larger diode free-wheeling current is nit unclamped by the transformer magnetizing current (compare to photo PP18)

Photo PP20
V_{dc} = 48.0 V



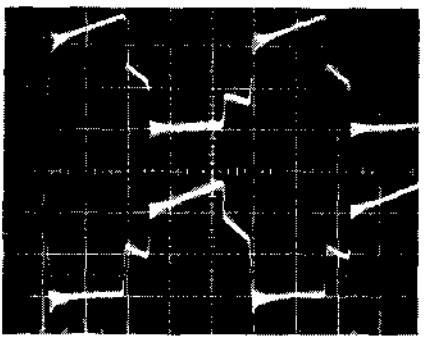
$V_{gs\ (Q1)}$ 10 V, 1 μs/cm

$V_{gs\ (Q2)}$ 10 V, 1 μs/cm

$I_{ds\ (Q2)}$ 1 A, 1 μs/cm

$I_{ds\ (Q1)}$ 1 A, 1 μs/cm

Photo PP21
V_{dc} = 48.0 V

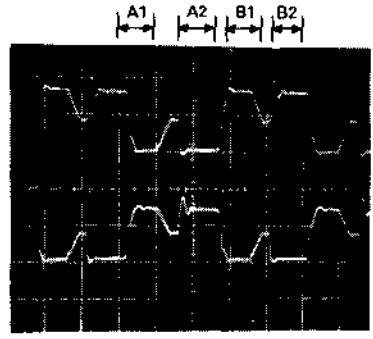


I (diode D1) 2A, 1 μs/cm

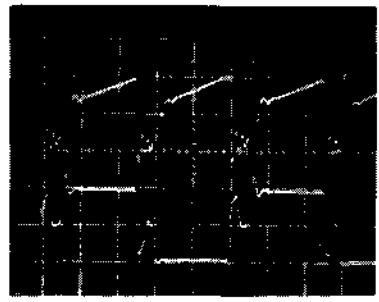
I (diode D2) 2 V, 1 μs/cm

For I_{dc} (5 V output) of 5 A
Showing "ledge" currents during dead time between transistor on times (refer to Section 2.2.10.3 and Fig. 2.6

Figure 14.15 Significant waveforms in 200-kHz converter of Fig. 14.8.



Showing an apparent failure mode at minimum output current or too high a magnetizing current: each transistor apparently turns on twice per half period (at A1, A2 for Q1 and at B1, B2 for Q2); the drains dropping to zero at A1, B1 are not transistor turnons—they occur for too large a transformer magnetizing current or too low a DC load current, as explained in the accompanying text



Showing critical waveforms on Fig. 14.8 circuit at output power 15% above specified maximum; converter efficiency is still a relatively high 83.8% and transformer temperature rise is only 54°C

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_{Q1}, Ω	I_5, A	P_5, W	V_{23}, V	R_{Q2}, Ω	I_{23}, A	P_{23}, W	$P_{(total)}$, W	Efficiency, %
48.0	2.51	120.5	5.00	0.339	14.7	73.7	26.12	25.0	1.04	27.3	101.0	83.8

Figure 14.16 Significant waveforms in 200-kHz converter of Fig. 14.8.

drains down to ground. However, $A1$ and $B1$ are fictitious turnons. At those times, the drain is driven down to ground by the positive drain bump after turnoff of the opposite transistor. This phenomenon is an extreme example of the situation seen in photo PP13 obtained by increasing the primary magnetizing current by increasing the transformer gap at low DC output currents.

As discussed in Sec. 14.3.8, this phenomenon occurs for a large magnetizing current reflected into the secondary. When this exceeds the DC load current in the output inductor, the rectifier diodes are unclamped from ground and produce a large positive bump above V_{dc} at the drain.

This, through the transformer coupling, produces a negative dip at

the opposite drain. When that negative dip at the opposite drain falls as low as ground, the inherent body diode of that MOSFET clamps in and holds the drain at ground. It appears thus that the MOSFET has turned on.

The circuit may continue to work in this odd mode, but with the feedback loop closed, oscillations may occur. On time will jump erratically from a true $B2$ aided by the fictitious $B1$ to a true on time of $B1 + B2$ at a slightly higher load current or slightly lower primary magnetizing current.

The problem can be avoided by ensuring that there can be no inadvertent increase of the transformer gap which would increase magnetizing current and that the lowest DC load current is always larger than the magnetizing current reflected into the master secondary.

14.3.16 Drain currents and drain-to-source voltage at output power 15 percent above specified maximum

This is presented (see also photo PP23, Fig. 14.16) to show that efficiency is still above 83 percent, transformer temperature rise is still only 54°C, and critical waveforms are still clean at this higher power level. The circuit of Fig. 14.8 has also been run up to 112 W of output power (photo PP24 in Fig. 14.17) with an efficiency of over 86 percent with clean waveforms and transformer temperature rise of only 65°C.

14.3.17 Ringing at drain during transistor dead time

To avoid ringing, RC snubbers ($R4$, $C4$ and $R5$, $C5$) are required across each half primary. Without them, a high-frequency oscillatory ring occurs throughout the transistor dead time (photo PP25 in Fig. 14.17). This worsens the RFI problem and affects slave DC output voltage in the same way as the dead time ledge of photo PP18.

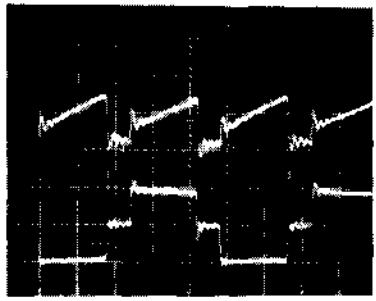
14.4 Flyback Topology Waveshapes

14.4.1 Introduction

Typical significant waveforms will be presented for a relatively low-power discontinuous mode, single-ended flyback.

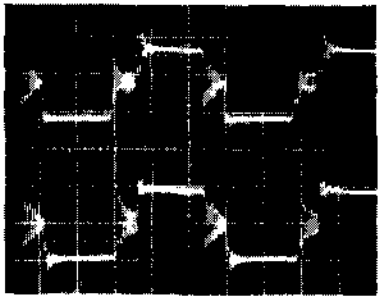
A discontinuous mode single-ended flyback is selected as it is the simplest for output powers up to about 60 W—the area of greatest usage for flybacks.

About the most serious drawback of the single-ended flyback is that energy stored in the transformer leakage inductance must be ab-



V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	R_{01}, Ω	I_5, A	P_5, W	V_{23}, V	R_{02}, Ω	I_{23}, A	P_{23}, W	$P_{(total)}$, W	Efficiency, %
48.0	2.71	130	5.00	1.03	4.85	24.3	21.7	5.36	4.05	87.9	112.1	86.2

Drain currents and drain-to-source voltage at 30% above maximum specified output power



Showing ringing during dead time when RC snubbers (R4, C4 and C5) are removed from primary

Figure 14.17 Ringing at drain during transistor dead time.

sorbed and subsequently dissipated in an *RCD* snubber (Chap. 11). If it is not, the leakage spike amplitude can rise to a voltage high enough to destroy the transistor. But above 60 W (especially for low-input voltage supplies where input currents are high), the high snubber dissipation is a significant drawback.

The double-ended flyback (Sec. 4.6) solves this problem not by storing leakage inductance energy in a snubber capacitor and then dissipating it in a resistor but by returning the leakage inductance energy without dissipation to the input supply bus.

Thus, the double-ended flyback is a desirable and widely used approach for output powers above 60 to 75 W. But significant waveforms in the single-ended flyback are much like those of the superior double-ended flyback, and hence waveshapes of the single-ended circuit only will be shown.

Waveforms shown herein were taken on the circuit of Fig. 14.18. The circuit is a 50-W, 50-kHz supply with one master output and one slave. Here again, the feedback loop was not closed. The transistor in-

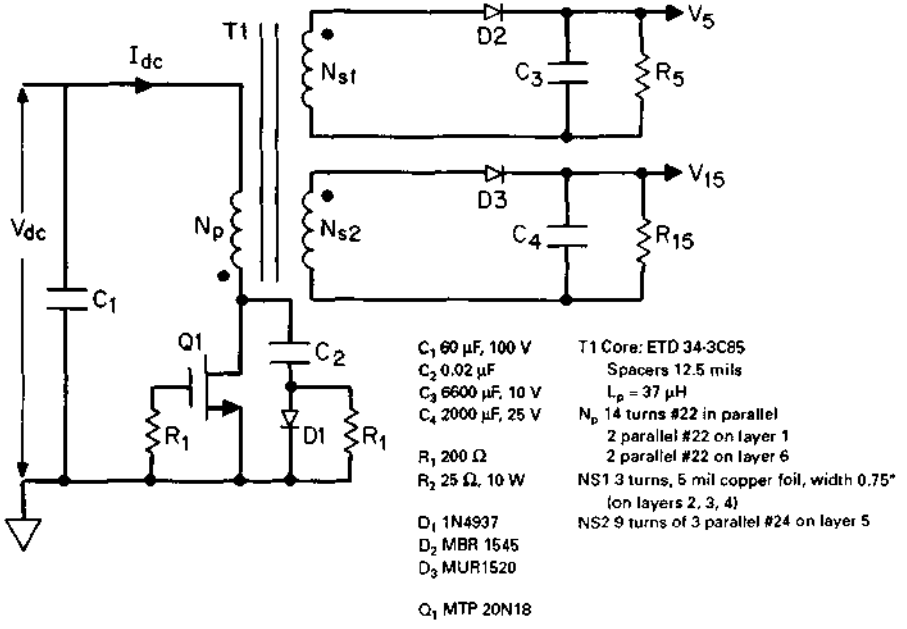


Figure 14.18 A 50-kHz 50-W flyback supply.

put was driven by a 50 kHz, manually adjustable width, pulse generator (using a 3525A PWM chip).

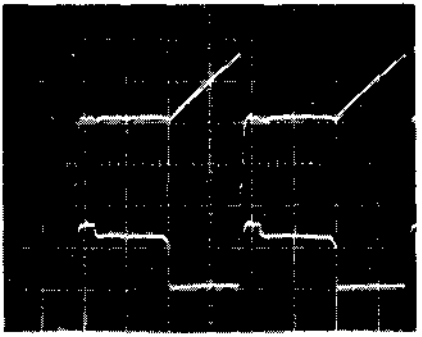
The pulse width at any input voltage or output load condition was manually set to the value which put the master output exactly at 5.00 V. The voltage of the slave was read and recorded at whatever value it then turned out to be.

The slave turns were chosen to yield 15 V when the master was 5.00 V. As will be noted, the slave output voltage is not entirely determined by the master/slave turns ratio. Due to secondary leakage inductances and interaction between master and slave windings, the slave output voltage is also dependent on the master DC output current.

14.4.2 Drain current/drain to source waveshapes at 90 percent of full load for minimum, nominal, and maximum input voltages

These are shown in Fig. 14.19. They show the characteristic linear ramp of primary current ($di_p/dt = V_{\text{primary}}/L_{\text{primary}}$) during the transistor on time. At the instant of turnoff, they show the leakage inductance voltage spike at the rising drain. The amplitude of this spike is controlled by the *RCD* snubber capacitor *C*2. That capacitor is chosen

Photo FB1
V_{dc} = 38 V

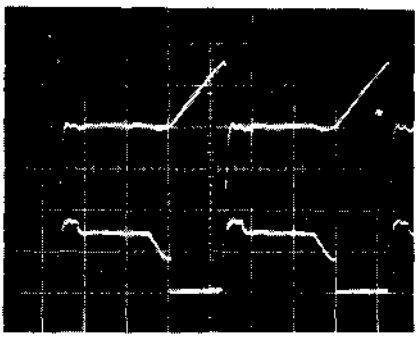


I_{ds} (Q1) 5 A, 5 μs/cm

V_{ds} (Q1) 50 V, 5 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	I_5, A	P_5, W	V_{15}, V	I_{15}, A	P_{15}, W	$P_{o\ (total)}$, W	Efficiency, %
38.0	1.67	63	5.002	6.58	32.9	25.3	0.504	12.8	45.7	72.1

Photo FB2
V_{dc} = 48.0 V

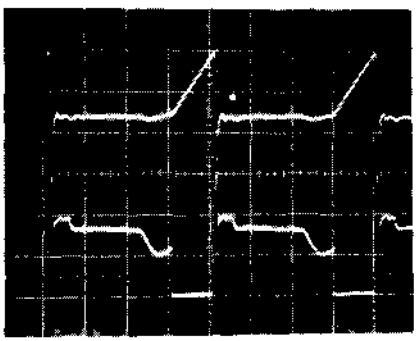


I_{ds} (Q1) 5 A, 5 μs/cm

V_{ds} (Q1) 50 V, 5 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	I_5, A	P_5, W	V_{15}, V	I_{15}, A	P_{15}, W	$P_{o\ (total)}$, W	Efficiency, %
48.0	1.31	62.6	5.001	6.57	32.9	25.42	0.505	12.8	45.7	72.0

Photo FB3
V_{dc} = 59.2 V



I_{ds} (Q1) 5 A, 5 μs/cm

V_{ds} (Q1) 50 V, 5 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	I_5, A	P_5, W	V_{15}, V	I_{15}, A	P_{15}, W	$P_{o\ (total)}$, W	Efficiency, %
59.2	1.07	63.1	5.003	6.58	32.9	25.36	0.504	12.8	45.7	72.4

Figure 14.19 Significant waveforms in 50-kHz flyback supply of Fig. 14.18.

large enough to limit the spike to a safe amplitude without causing too much dissipation (= $0.5C2(V_{peak})^2/T$) in snubber resistor $R1$.

The waveforms show that as V_{dc} increases the pulse width required to maintain a constant master output voltage decreases. This of course is what the feedback loop will do when connected in. It is also seen that the peak-ramp current is constant for all input voltages at constant output power.

It is also seen that after the leakage spike, the drain falls (except for the small pedestal for a short time after the spike) to a level of $V_{dc} + (N_p/N_{s1})\ (V_5 + V_{D2})$. It remains at that level until the reset volt-second product equals the set volt-second product $[V_{dc}t_{on} = (N_p/N_s)]\ (V_5 + V_{D2})$ and then falls back to V_{dc}.

Figure 14.20 shows the same waveforms at the three input voltages for a lower total output power (17 W). All on time pulse widths are narrower, and consequently all peak primary currents ($= di = V_{primary}t_{on}/L_p$) are lower. This of course is because less power $[0.5L_p(I_p)^2/T]$ is required from the input bus, and because the peak primary current at the end of the on time is less—so is the leakage inductance spike.

14.4.3 Voltage and currents at output rectifier inputs

These are shown to track down the reason why slave output voltages are not entirely dependent on the master/slave turns ratios, and why they are dependent on the master output current.

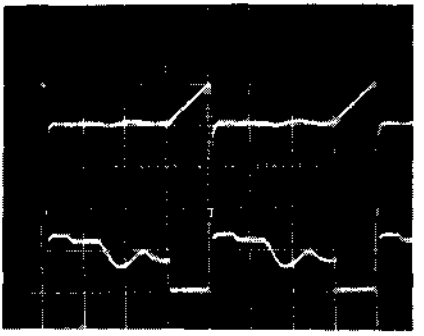
When the transistor turns off, all the energy stored in the primary $(0.5L_pI_p)^2$ is delivered to the secondaries (except for some stored in the leakage inductance which is diverted to the snubber capacitor $C2$). That energy is delivered to the master and slave outputs in the form of currents shown in Fig. 14.21. But it can be seen that the currents to the master and slave outputs via their rectifiers are not of equal duration and certainly not of similar waveshapes.

Note that when DC output current of the master is increased from 2.08 A (photo FB8) to 6.58 A (photo FB10), the small voltage pedestal on the slave output rectifier moves up from about 20 to 28 V (photo FB7 and FB9). This occurs because of secondary leakage inductance and mutual coupling between master and slave secondaries, and with the larger voltage at the pedestal at the slave rectifier anode (compare photos FB7 and FB9), the slave DC output voltage is increased.

14.4.4 Snubber capacitor current at transistor turnoff

Figure 14.22 shows that at the instant of transistor turnoff, all the transformer primary current (5 A from photo FB5) is immediately

Photo FB4
V_{dc} = 38 V

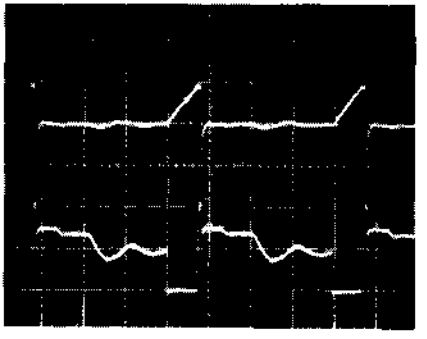


I_{ds} (Q1) 5 A, 5 μs/cm

V_{ds} (Q1) 50 V, 5 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	I_5, A	P_5, W	V_{15}, V	I_{15}, A	P_{15}, W	$P_{o\ (total)}$, W	Efficiency, %
38.0	0.595	22.6	5.003	2.08	10.4	19.1	0.380	7.27	17.7	78.3

Photo FB5
V_{dc} = 48.0 V



I_{ds} (Q1) 5 A, 5 μs/cm

V_{ds} (Q1) 50 V, 5 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	I_5, A	P_5, W	V_{15}, V	I_{15}, A	P_{15}, W	$P_{o\ (total)}$, W	Efficiency, %
48.0	0.481	23.1	5.004	2.09	10.4	19.2	0.381	7.29	17.7	76.6

Photo FB6
V_{dc} = 59.7 V

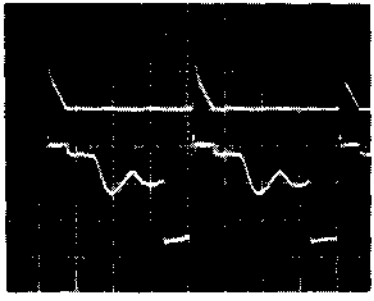
I_{ds} (Q1) 5 A, 5 μs/cm

V_{ds} (Q1) 50 V, 5 μs/cm

V_{dc}, V	I_{dc}, A	P_{in}, W	V_5, V	I_5, A	P_5, W	V_{15}, V	I_{15}, A	P_{15}, W	$P_{o\ (total)}$, W	Efficiency, %
59.7	9.397	23.7	5.004	2.08	10.4	19.15	0.381	7.29	17.7	74.7

Figure 14.20 Significant waveforms in 50-kHz flyback supply of Fig. 14.18.

Photo FB7



Voltage and current into 15 V rectifier (D3, Fig. 14.18)

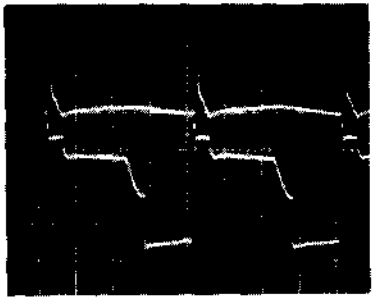
V_{15} current 0.39 A, voltage = 19.39 V
V_5 current 2.08 A, voltage = 5.00 V

Photo FB8

Voltage and current into 5 V rectifier (D2, Fig. 14.18)

V_{15} current 0.39 A, voltage = 19.39 V
V_5 current 2.08 A, voltage = 5.00 V

Photo FB9



Voltage and current into 15 V rectifier (D3, Fig. 14.18)

V_{15} current 0.50 A, voltage = 25.8 V
V_5 current 6.58 A, voltage = 5.00 V

Photo FB10

Voltage and current into 5 V rectifier (D2, Fig. 14.18)

V_{15} current 0.50 A, voltage = 25.8 V
V_5 current 6.58 A, voltage = 5.00 V

Figure 14.21 Showing why slave DC output voltage is not related to master DC output voltage by their respective turns ratio. Master secondary leakage inductance diverts some of its peak secondary current into the slave rectifier anode immediately after turn-off. As long as this diverted current persists, it couples a pedestal voltage into the slave rectifier anode. The slave output capacitor charges up to the pedestal peak which is higher than the induced secondary voltage and is dependent on the master load current. The remedy is to minimize secondary leakage inductances. The effect is minimized by an inductor in series in the slave secondaries.

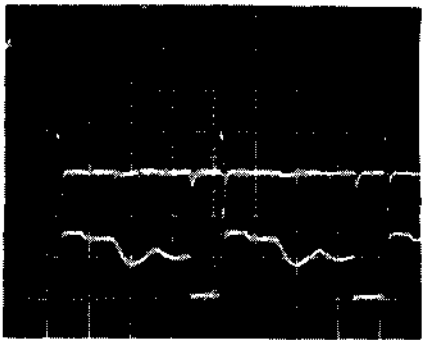


Photo FB11

Figure 14.22 At the instant of turnoff, all the current which had been flowing in series in the transformer magnetizing and leakage inductances is immediately driven into the snubber capacitor of Fig. 14.18c.

transferred to the snubber capacitor $C2$ and through snubber diode $D1$ to ground. That current represents energy stored in the primary leakage inductance $[0.5L_{\text{leakage}}(I_p)^2]$. Until it is transferred to $C2$ as electrostatic energy $[0.5C2(V_p)^2]$, all the current or energy stored in the transformer magnetizing inductance cannot be delivered to the secondaries.

If $C2$ is made too small, the current reflected from the primary at the instant of turnoff may charge it to a dangerously high leakage inductance spike. Thus, $C2$ is selected large enough to limit the leakage spike to a safe value; yet, not so large as to result in excessive dissipation.

Part

4

Newer Applications for Switching Power Supply Technique

Chapter

15

Power Factor, Power Factor Correction

15.1 Power Factor—What Is It and Why Must It Be Corrected?

The term *power factor* is borrowed from elementary AC circuit theory. When a sinusoidal AC power source feeds either an inductive or a capacitive load, the load current is also sinusoidal but lags or leads the input voltage by some angle x. For an rms input voltage V_i and rms input current I_i, the *apparent power* taken from the line is $V_i I_i$.

But the actual power delivered to the load is only $V_i I_i \cos x$. It is only the component of input current which is in phase with the voltage across the load resistance ($I_i \cos x$) which contributes to the load power. The component of input current normal to the voltage across the load resistor ($I_i \sin x$) does not contribute to the actual load power. At one phase of the AC input waveform, it represents power drawn from the input source which is temporarily stored in the reactive component of the load device. In a later phase of the input voltage, this stored current or energy is returned to the input source. But this excessive current, which does not contribute to load power, wastes power in the winding resistance of the input power source and the resistance of the input power lines.

Reactive and power generating components of the input line current can be seen in Fig. 15.1.

In AC power circuit jargon, the magnitude of $\cos x$ is referred to as the *power factor*. It is desirable to keep the power factor as close to unity as possible, i.e., to keep the input line current sinusoidal and in phase with the sinusoidal input line voltage. The means to achieve this is referred to as *power factor correction*.

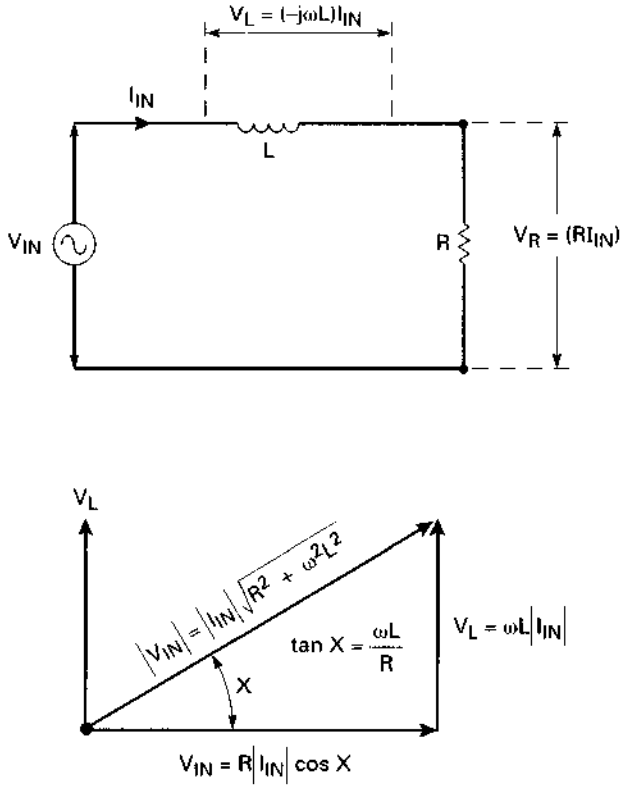


Figure 15.1 Power factor. In AC circuit terminology, *power factor* is defined as the cosine of the phase angle between input voltage and input current. In circuits with pure resistive load, there is no phase difference between input voltage and current, and the power factor is unity. If current lags or leads input voltage, it is only the component of input current which is in phase with the voltage across the load resistance which contributes power to the load. The component of input current normal to voltage across the load resistance wastes power in the resistance of the source generator. In power supplies with a capacitor filter across the input bridge rectifier, the input line current consists of very narrow spikes with fast rise and fall time. These current spikes have a high rms value, waste power, and give rise to RFI/EMI problems. Borrowing a term from AC circuit theory, power supplies with such input line currents are said to have poor "power factor." *Power factor correction* seeks to eliminate such line current spikes and force input current to be sinusoidal, in phase with input voltage, and to generate a fairly well regulated DC output voltage somewhat greater than the peak of the incoming sine wave.

15.2 Power Factor Correction in Switching Power Supplies

In the field of switching regulators, any circuit configuration which causes input line current to be nonsinusoidal or even sinusoidal but out of

phase with the sinusoidal input voltage or to have harmonics of the line voltage results in a lowered power factor and consequent waste of power.

The circuit technique described in this chapter maximizes or "corrects" the power factor by forcing the input line current to be sinusoidal, in phase with the input line voltage, and as free from line harmonics as possible. These techniques are especially useful—indeed, mandatory in new designs—in "off-the-AC power line" power supplies. This is so because of the severely distorted input line current waveform caused by the usual capacitor filter following the input bridge rectifier. This can be seen in Fig. 15.2*c*.

In Fig. 15.2*a* and *b*, if the filter capacitor C_o were absent and the load were a pure resistor, the voltage at V_o would have the half sinusoidal waveshape *ABXCDYEF*. Rectified current coming out of the rectifier would have the same half sinusoidal waveshape, and the line current drawn from the input source would be purely sinusoidal and in phase with the sinusoidal input voltage. The power factor would be unity; and if V_i and I_i were the input voltage and current measured with rms meters, the input and output load power would be $V_i I_i$.

But, of course, half sinusoids of rectified output voltage such as *ABXCDYEF* (Fig. 15.2*b*) could not be tolerated. The sole purpose of the rectifier is to convert the input AC voltage to a DC voltage with as

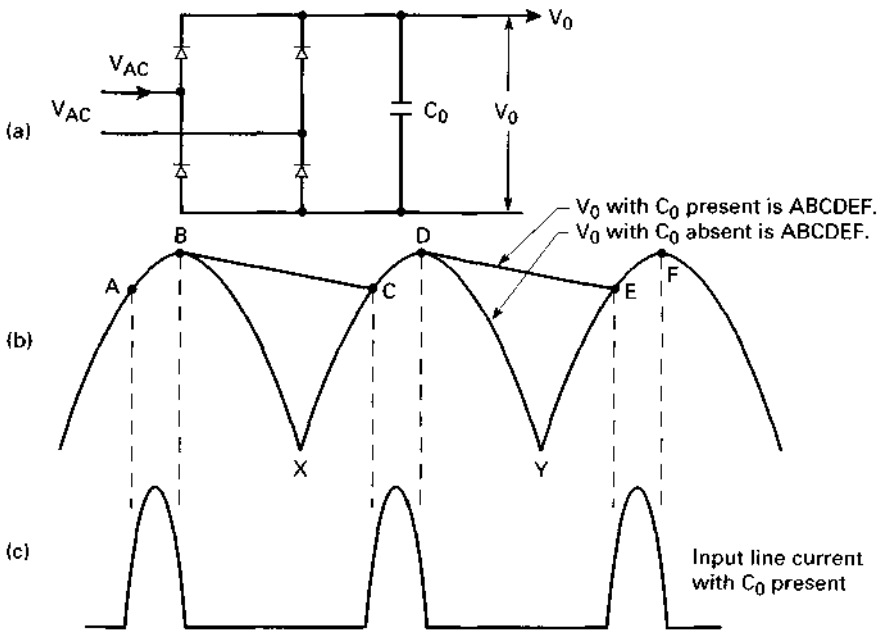


Figure 15.2 (*a* and *b*) Output voltages with and without a capacitor filter after the input bridge rectifier. (*c*) Input line current with capacitor C_o present.

low a ripple content as possible. The filter capacitor C_o is thus added to yield the waveform *ABCDEF*. This results in a higher DC component (midway between amplitudes at B and C or D and E) and a lowered peak-to-peak ripple of $B - C$ or $C - D$. Now between instants B and C or D and E, all rectifiers are reverse-biased, no line current flows, and all load current is drawn from the filter capacitor C_o. At instants like A, C, and E, the rising input voltage forward-biases the rectifiers, and line current now flows to the load and into C_o to replenish the charge lost when it alone was supplying load current.

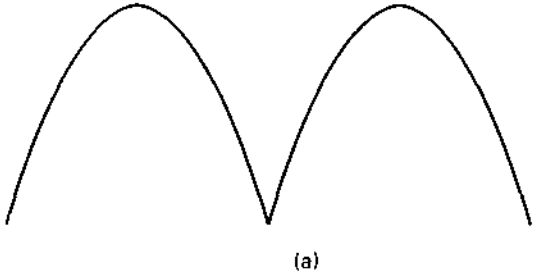
Line current with the filter capacitor is shown in Fig. 2.15*c*. It is a sequence of narrow current pulses on the front end of each half sinusoid of input voltage. The larger the filter capacitor, the shorter the duration of the pulses, the higher their peak, the larger their rms values, and the sharper their rise and fall times.

It is these narrow, fast-rise-time line current pulses that power factor correction aims to eliminate. Their fast rise time causes radio-frequency interference (RFI) problems, and more important, their rms equivalent value is higher than what is needed to supply the required output load power. The higher-than-required rms current causes not only excessive temperature rise in the input generator, but also higher temperature rise and decreased reliability in the filter capacitor.

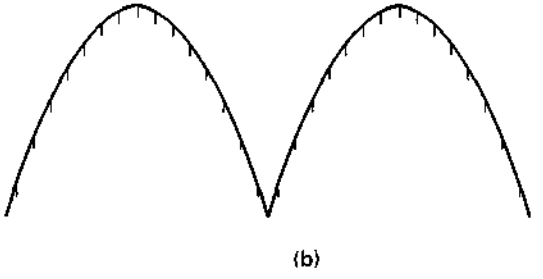
Power factor correction eliminates the large filter capacitor C_o after the bridge rectifier and thereby permits voltage after the rectifier to rise and fall in half-sinusoidal fashion. It then converts these half sinusoids of input voltage to a constant DC output voltage and by monitoring the input line current forces it to be sinusoidal and proportional to the instantaneous input line voltage. It does this by use of an "on" time-modulated boost regulator (Sec. 1.4). Throughout the half sinusoids, the on time of the boost regulator is modulated by a PWM control chip in such a way as to force the input line current to be sinusoidal and the output voltage to remain constant at a value somewhat greater than the incoming sine wave peak.

15.3 Power Factor Correction—Basic Circuit Details

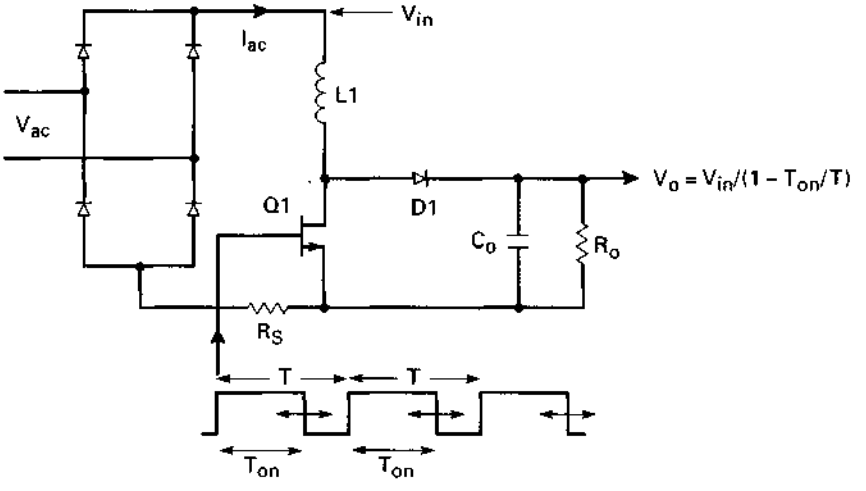
The basic scheme used to implement power factor correction (PFC) is shown in Fig. 15.3. First, it eliminates the input filter capacitor of Fig. 15.2*a* and permits voltage immediately after the bridge rectifier to fall to zero at the input sine wave zero crossings (Fig. 15.3*a*). By removing the input capacitor C_o, the line current flows continuously and sinusoidally, avoiding the narrow current pulses of Fig. 15.2*c*. The resulting half sinusoids of voltage drive a continuous-mode boost converter.



(a)



(b)



(c)

Figure 15.3 Use of boost regulator after the input bridge rectifier to force sinusoidal line current in phase with the AC line voltage and to yield a regulated DC output voltage somewhat higher than the peak of the input sine wave voltage.

The first task of the power factor correction circuit is to use the boost converter to convert the varying input voltage up and down the half sinusoids to a constant, fairly well-regulated DC voltage somewhat higher than the input sine wave peak. It does this by using a continuous-conduction-mode *boost converter* (Sec. 1.4) in the following way.

A boost converter, it is recalled, "boosts" a low voltage to a higher voltage. It does this by turning on the transistor $Q1$ for a time T_{on} out of a period T and storing energy in inductor $L1$. When $Q1$ turns off, the polarity across $L1$ reverses, and the dot end of $L1$ rises to a voltage (V_o) higher than the input voltage—V_{in}. Energy stored in $L1$ during T_{on} is transferred via $D1$ to the load and $C1$ during the $Q1$ off time. It can be shown that the output-input voltage relation of such a boost converter is given by

$$V_o = \frac{V_{in}}{1 - T_{on}/T} \tag{15.1}$$

Now throughout the half sinusoids of V_{in}, the $Q1$ on time, denoted by T_{on}, is width-modulated in accordance with Eq. 15.1 to yield a constant DC voltage V_o somewhat higher than the peak of the input voltage sine wave. The on time throughout the sinusoidal half periods is controlled by a PFC control chip which senses V_o, compares it to an internal reference in a DC voltage error amplifier, and in a negative feedback loop sets T_{on} to keep V_o constant at the selected value.

From Eq. 15.1 it is seen that at the lower voltage portions of the half sinusoids of Fig. 15.3*a*, the $Q1$ on time will be large to boost low input voltage to a value higher than the peak of the sinusoid. And as V_{in} rises toward its peak, the PFC control chip will automatically decrease the $Q1$ on time so that each voltage along the rising half sinusoid is boosted to that same DC level somewhat higher than the sinusoid peak. The progression of the on times throughout the half sinusoids is seen in Fig. 15.4.

The second task of the power factor correction circuit is to sense input line current and force it to have a sinusoidal waveshape in phase with the input line voltage. This, too, is done by width modulation of the same boost regulator's on time. The on time is determined in a negative feedback loop which compares a sample of the actual input line current to the amplitude of a clean reference current sine wave. The difference between these two sine waves is an error voltage which is used to modulate the on time to force the two sine waves to be equal in amplitude.

The final voltage which controls the boost regulator's on time must be a mix of the DC output voltage error and the input line current error voltage. This mixing is done in a multiplier building block

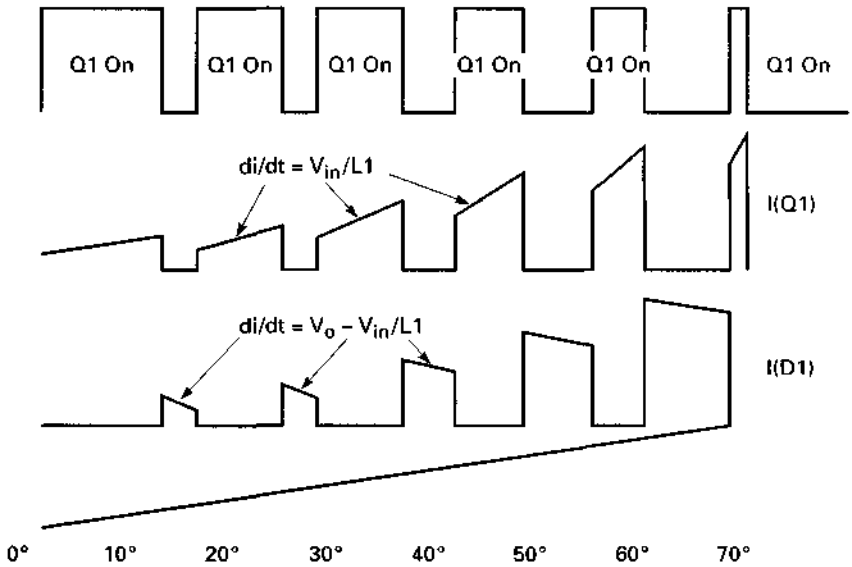


Figure 15.4 Significant currents in boost converter of Fig. 15.3 at points progressively higher up on the half sinusoidal waveforms.

whose output is proportional to the product of the output voltage error voltage and input current error voltage.

15.3.1 Continuous- versus discontinuous-mode boost topology for power factor correction

Boost converters can be operated in either the discontinuous or continuous mode (pages 27–31). But the continuous-mode boost topology is far better suited to yield relatively smooth, ripple-free half sinusoids of input line current in this application. This can be seen from Fig. 15.5 which shows a continuous-mode boost converter fed from a constant DC input voltage. The continuous-mode boost topology differs significantly from the discontinuous-mode (Fig. 1.10).

In the discontinuous mode, the inductor $L1$ is made small to yield a steep ramp ($di/dt = V_{in}/L1$) of input current (Fig. 1.10c) to $Q1$. When $Q1$ turns off, all the current or energy stored in $L1$ is transferred via $D1$ to the load (Fig. 1.10*d*). Since $L1$ is small, the downward ramp of current through $D1$ [$di/dt = (V_o - V_{in})/L1$)] is also steep and $D1$ current falls to zero before the next $Q1$ turn-on. The input line current then, which is the sum of the $Q1$ current when it is on and the $D1$ current when $Q1$ is off, is not at all constant over one complete switching cycle. It consists of steep up and down ramps with zero current gaps between the end of a turnoff and the next turn-on.

But in the continuous-mode of Fig. 15.5, the inductor $L1$ is made sig-

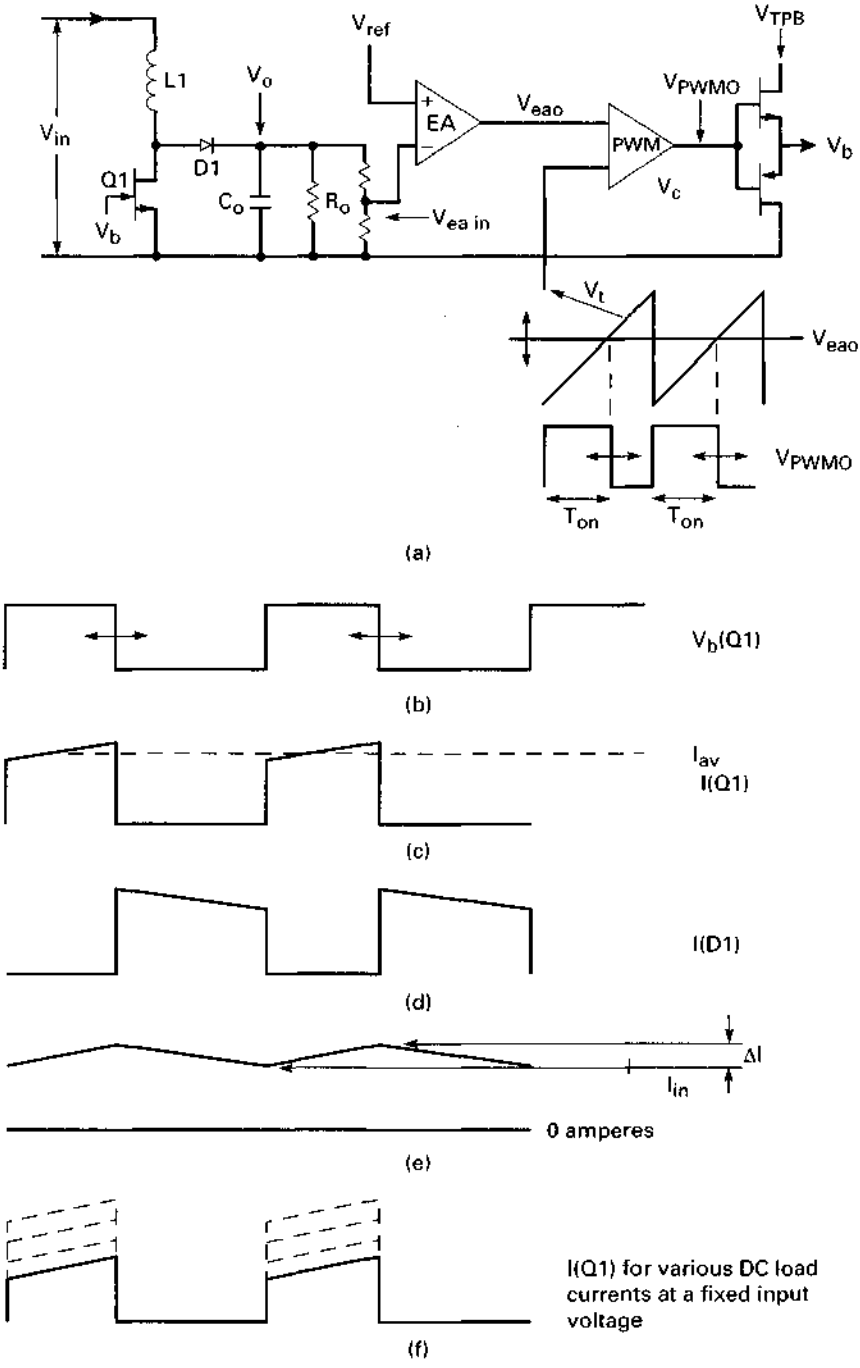


Figure 15.5 A continuous conduction mode boost converter fed from a fixed DC input voltage. The circuit regulates against DC input voltage changes by varying $Q1$ on time. It regulates against load current changes by maintaining a fixed on time, while building up the average current delivered by $Q1$ over a number of cycles (Fig. 15.5*f*).

nificantly larger. Then the $Q1$ current (Fig. 15.5*c*) has the shape of a large step of current with a slow upward ramp on it, and the $D1$ current has the shape of a large step with a slow downward ramp. And importantly, there is no gap of zero current between the end of a turnoff and the next turn-on. Input line current (Fig. 15.5*e*) is now the sum of the I_{Q1} and I_d currents; and if the ramps are made small by using a large $L1$, the line input current over one switching cycle is then a constant I_{av} with very small peak-to-peak ripple of ΔI. The input power is $V_{in}I_{av}$.

Now for an AC input, such a continuous-mode boost converter is used after the input bridge rectifier output, as in Fig. 15.3. At any point on the half sinusoid input voltage, the $Q1$ on time will be forced by a PWM control chip to boost that instantaneous voltage to the desired DC output voltage. A DC voltage error amplifier, a DC reference voltage, and a pulse width modulator in the control chip modulate the $Q1$ on time in a negative feedback loop, to yield a constant DC output voltage throughout the half sinusoidal input voltage.

The instantaneous input line current will be sensed by R_s and will be proportional to the instantaneous voltage throughout the half sinusoid. During any one on time, current flows through $L1$, $Q1$, and R_s back to the negative end of the bridge; and during the following off time it flows through $L1$, $D1$, R_o, and C_o in parallel and R_s back to the negative end of the bridge.

By making $L1$ large, the peak-to-peak ripple current throughout each switching cycle is kept small. Depending on the switching speed of $Q1$, there may be very narrow spikes on the half sinusoids of current monitored in R_s (seen in Fig. 15.3*b*). If present, these may cause an RFI problem. But a very small capacitor (in the vicinity of 1.0 mF) across R_s easily eliminates them.

15.3.2 Line input voltage regulation in continuous-mode boost converters

Before we consider the details of boost regulators with half sinusoidal input voltage as in power factor correction schemes, it is of interest to see how a continuous-mode boost regulator corrects against line and load changes with a constant DC input voltage.

First, consider how the output/input voltage relation of Eq. 15.1 comes about. In Fig. 15.5, the switch transistor $Q1$ is on for a time T_{on} and off for T_{off} out of the total period T. Neglect the on voltage drops in $Q1$ and $D1$. Since inductor $L1$ has negligible resistance, the voltage across it, averaged over one switching cycle, must be zero. And since the voltage at the top end of $L1$ is V_{in}, the voltage at its bottom end, averaged over one cycle, must also be V_{in}. This means, in Fig. 15.6*a*, that area $A1$ must equal area $A2$. Since during T_{off} the top end of $L1$ is at V_o,

$$V_{in}T_{on} = (V_o - V_{in})T_{off}$$

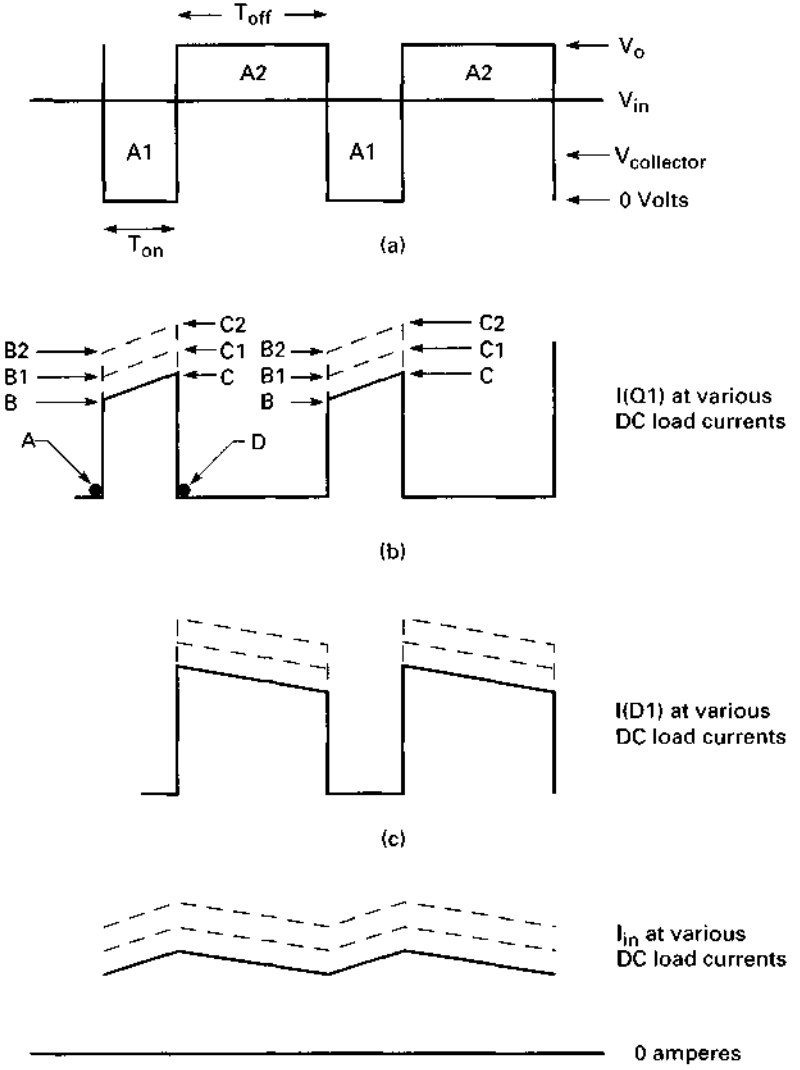


Figure 15.6 Regulation against load current changes in a continuous conduction mode boost converter.

$$= (V_o - V_{in})(T - T_{on})$$

or solving for V_o, we have

$$V_o = \frac{V_{in}}{1 - T_{on}/T}$$

which is the previously mentioned Eq. 15.1.

Now in Fig. 15.5*a*, output voltage regulation is achieved by chang-

ing T_{on} in accordance with Eq. 15.1 as V_{in} changes. This is done with the pulse width modulator (Fig. 15.5*a*). If V_{in} momentarily changes, so does V_o. A fraction of V_o ($= V_{ea\ in}$) is sensed by error amplifier EA and compared to a reference voltage V_{ref} to yield an error voltage $V_{ea\ o}$. This DC error voltage is compared to a built-in triangle voltage V_t in voltage comparator *Vc*. The V_c output is a square wave which is high for the time from the start of the triangle up to the instant that the triangle crosses the error voltage output $V_{ea\ o}$. And *Q*1 is turned on via totem pole driver (TPD) for the high time of the *Vc* output.

Thus, if V_{in} goes momentarily low, so does V_o and the inverting input to V_{ea}. Then the V_{ea} output goes high, the triangle V_t crosses the error amplifier output later, the on time increases, and V_o moves back up in accordance with Eq. 15.1. And obviously, if V_{in} goes high, V_o goes high, $V_{ea\ o}$ goes low, T_{on} decreases, and V_o moves back down.

15.3.3 Load current regulation in continuous-mode boost regulators

Continuous-mode boost converters operate in an odd way to correct for load current changes. From Eq. 15.1, note that V_o and T_{on} are independent of the load current. Yet if the DC load current changes, it is obvious that the transistor and output diode currents must change despite a constant on time.

The circuit responds to a change in load current in the following way. Prior to, say, an increase in load current, assume the *Q*1 current is like *ABCD* in Fig. 15.6*b*. Now for a small increase in load current, in the steady state, *Q*1 current will move up to, say, *AB*1*C*1*D*. And for a larger load current change, the *Q*1 current will move up to *AB*2*C*2*D*. To cause these changes, T_{on} changes over a few switching cycles but returns to its original value in the steady state. The steady-state diode *D*1 current for these three different load currents is shown in Fig. 15.6*c*. The output load current is the sum of I_{Q1} and I_d, and its peak-to-peak ripple I_{or} can be made small as desired by simply increasing *L*1.

The increased value of the "ramp on a step" waveforms in Fig. 15.6*b* and *c* for increased DC loads occurs over a number of switching cycles as follows (Fig. 15.5*a*). If the DC load current increases, V_o goes down momentarily because of the source impedance looking back into it. Then $V_{ea\ in}$ goes down, $V_{ea\ o}$ goes up, the V_t triangle crosses $V_{ea\ o}$ later in time, and T_{on} increases. Now the I_{Q1} current ramps up for a longer time to a higher value. Then I_d starts later in time from a higher value and, with a shorter off time, has a higher value at the end of the off time. Hence the initial I_t current at the start of next turn-on is larger.

This progresses over a number of cycles with the average currents at the center of the I_{Q1}, I_d ramps in Fig. 15.5*c* and *d* increasing until they equal the increased DC load, at which time T_{on} and T_{off} slowly fall back to their initial values, as called for by Eq. 15.1. Thus for any

change in DC load, the T_{on} and T_{off} times temporarily change, but slowly relax back to their original value.

Thus in a qualitative way, it can be seen that the bandwidth of the output voltage error amplifier must not be too large. If it is large, it responds too quickly and does not permit the output voltage to shift for too long a time from its normal value at a fixed input voltage. The output voltage must be permitted to shift from its value dictated by the input voltage for a time sufficient for the above-described current buildup to occur over a number of switching cycles.

All the required functions described above are currently doable with a power-factor-correcting integrated-circuit chip available from a number of manufacturers. The functions that these chips implement are mostly the voltage and current sensing error amplifiers, mixing of the error signals, and generating of the width-modulated boost transistor turn-on control pulses.

Thus adding power factor correction to a power supply requires only removing the usual filter capacitor (C_o) of Fig. 15.2*a*, adding one of the available chips, and adding the inductor, boost transistor, current-sensing resistor, and output capacitor of Fig. 15.3*c* plus about half a dozen small resistors and capacitors.

15.4 Integrated-Circuit Chips for Power Factor Correction

A number of major manufacturers have available integrated-circuit (IC) chips to perform all the functions required for power factor correction. They all use substantially the same block diagram based on a continuous-mode boost regulator and a scheme to sense and control DC output voltage and input line current by width modulation of the boost on time.

The earliest and probably most widely used of these chips, the Unitrode UC 3854 is typical of most of the others and is discussed in detail. Other chips—the Motorola MC 34261 and MC 34262—will be mentioned briefly. Microlinear ML 4821, Linear Technology LT 1248, and Toko 83854 Linear are mostly identical to the Unitrode 3854.

15.4.1 Unitrode UC 3854 power factor correction chip

A simplified block diagram showing the major elements of the chip is seen in Fig. 15.7. The functions of its various components are as follows. The discussion here is based on Unitrode Application Note U-125 by Claudio de Silva.

Transistor $Q1$, inductor $L1$, diode $D1$, and output capacitor C_o com-

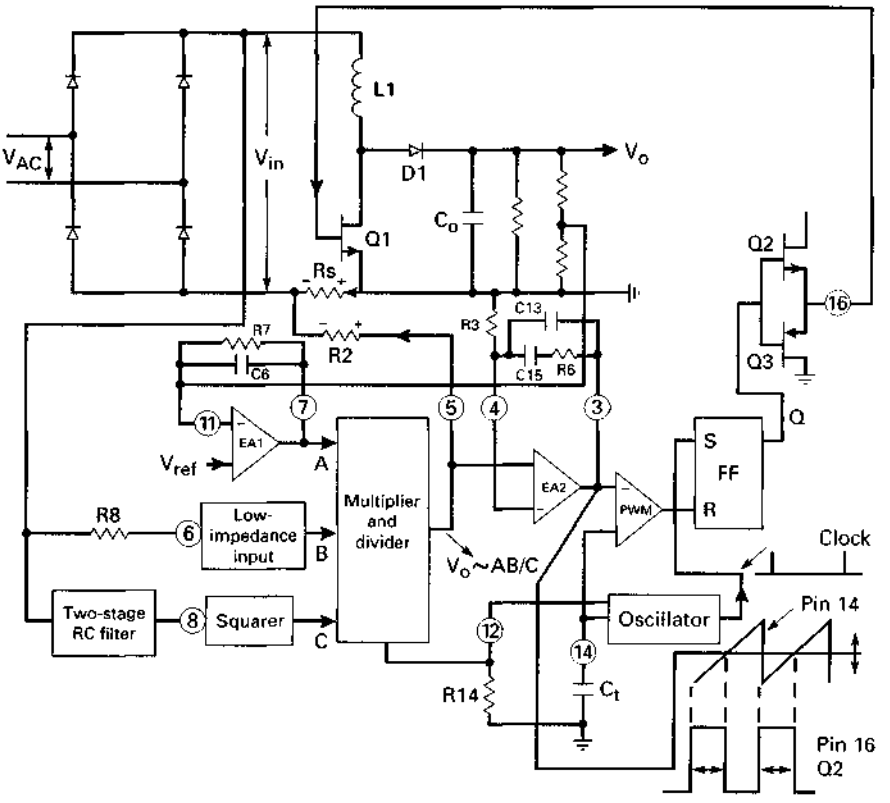


Figure 15.7 Simplified block diagram of the Unitrode UC 3854 Power Factor Control Chip.

prise the boost converter. A sawtooth voltage generator, operating at a frequency $F_s = 1.25/(R_{14}C_t)$ sets its switching frequency. Power switch $Q1$ is turned on and off by totem pole output drivers $Q2$ and $Q3$.

An on time commences when FF1 (flip-flop 1) is set by a narrow spike at the start of each sawtooth from the oscillator. The end of the on time occurs when FF1 is reset by the pulse width modulator (PWM) output at the instant the sawtooth at its noninverting input terminal crosses the DC voltage level at the output (pin 3) of linear current amplifier EA2. Voltage at pin 3 is simply a noninverted, amplified version of the instantaneous difference between voltage drop across R_s and voltage rise across $R2$.

Width modulation of this on time by the PWM boosts the half sinusoids of input voltage from the bridge rectifier to a constant output voltage. It also forces the input line current to be accurately sinusoidal and in phase with the input line voltage.

15.4.2 Forcing sinusoidal line current with the UC 3854

Current out of pin 5 is a continuous sequence of positive-going half sinusoids whose amplitude at any instant is proportional to the product of the DC voltage at point A and the current into pin 6. The input at pin 6 is a reference half sinusoid of current in phase with the half sinusoid line voltage after the bridge. Voltage at pin 5 is then a continuing sequence of half sinusoids in phase with the half sinusoids of line voltage at the bridge output. The amplitude of the sinusoids is proportional to the voltage at the output of error amplifier EA1.

Line current is made sinusoidal by making, at all instants throughout each half sinusoid, the voltage drop in R_s (from right to left, Fig. 15.8*c*) very closely equal to the voltage rise (from left to right, Fig. 15.8*b*) across $R2$.

Now the current in R_s, averaged over one switching cycle, is equal to the input line current averaged over that cycle. This is so, of course, because the line current is equal to the sum of the $Q1$ current when it is on and the $D1$ current when $Q1$ is off.

Hence when the voltage drop across R_s is forced to equal the voltage rise in $R2$, the line current is also half sinusoidal and in phase with the voltage waveshape after the bridge rectifier. It can be seen from Figs. 15.5*c, d,* and *e* that since the boost operates in the continuous mode with a large inductor, the ripple current over one switching cycle is small. Thus when the voltage drop across R_s is made equal to the rise across $R2$ throughout a half period, since the voltage across $R2$ is a smooth half sinusoid, the line current in R_s will also be smooth with very little switching frequency ripple.

Now throughout the 60-Hz half cycle, the voltage rise across $R2$ is slightly higher than the drop across R_s, as the voltage drop across R_s is continually racing to keep up with the reference voltage rising across $R2$. This difference—the instantaneous error voltage—is shown in Fig. 15.8*d*. It is a voltage positive with respect to ground throughout the half sinusoid and has a cusplike concave-upward waveshape. It is amplified by noninverting current amplifier EA2 and has the concave-upward waveshape shown in Fig. 15.8*e*.

Now in the PWM comparator, the waveform at pin 3 is compared to the roughly 5-V peak triangle at pin 14. At points like X and Y (Fig. 15.8*e*), the triangle crosses the higher voltage levels there late in time and the on time is long. At the sine wave peak (point P), the voltage level is lower and hence the triangle crosses the lower voltage earlier in time and the on time is less. Thus throughout the half cycle, the cusplike pin 3 waveform yields an on time which is maximum close to the positive-going zero crossing, decreases as input voltage from the bridge rises toward its peak, and falls as the input voltage falls toward its negative-going zero crossing. These varying on times boost

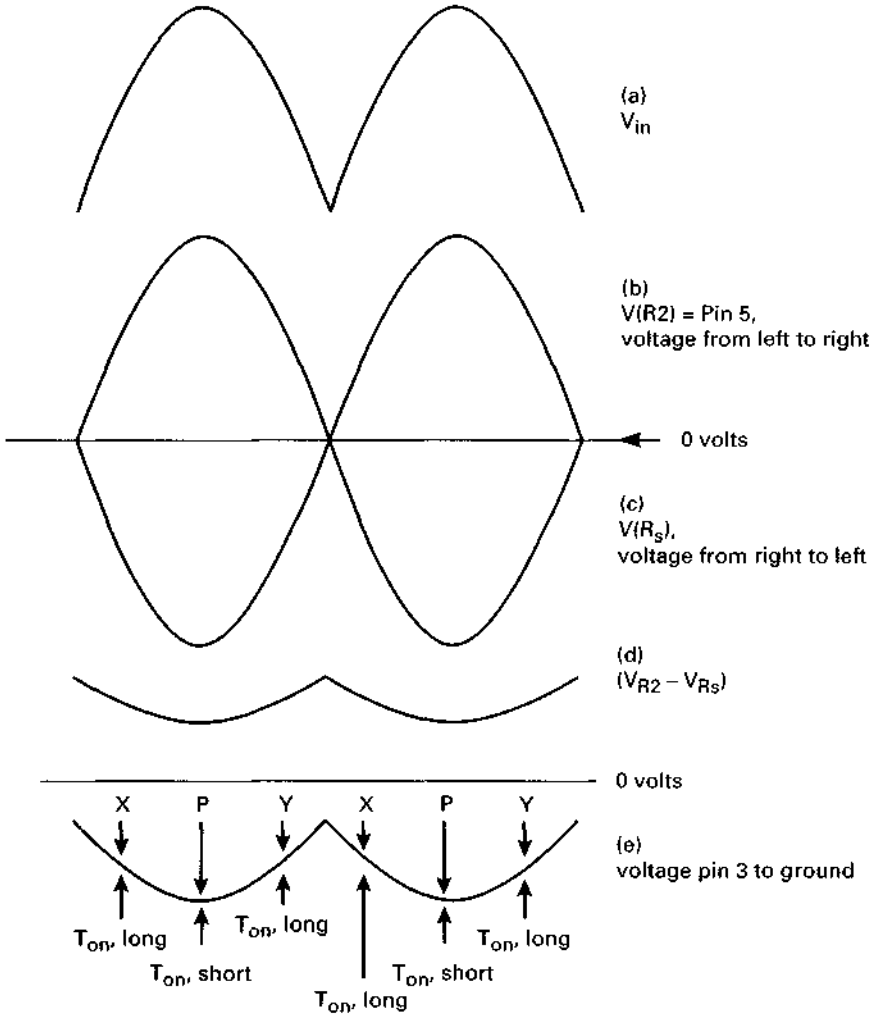


Figure 15.8 Critical waveforms in Unitrode 3854 Power Factor Control Chip.

the half sinusoid input waveforms to the constant DC output voltage at C_o as called for by Eq. 15.1.

But these on times dictated by the error voltage signal at pin 3 are the average taken over a few switching cycles. As the current demanded by the sinusoidal voltage across $R2$ changes, so must the ramp-on-a-step current pulses in R_s change. This occurs as discussed in Sec. 15.3.3 by momentarily altering the error voltage at pin 5 and hence at pin 3. The PWM comparator then momentarily alters the on time so that the ramp-on-a-step current pulse flowing through R_s

causes the average voltage across it (averaged over one switching cycle) to equal the voltage across $R2$. And after a few cycles when those voltages are equal, the on time relaxes back to the value required by Eq. 15.1 to boost the instantaneous input voltage to the desired constant DC output voltage.

15.4.3 Maintaining constant output voltage with UC 3854

Without going into the details of the multiplier and divider, it is sufficient to say that the output voltage at pin 5 is simply the product of voltage at input point A and current at input point B. Regulation against V_o changes as follows (Fig. 15.7):

Point A is the output voltage of the V_o error amplifier which compares a fraction of V_o to a fixed reference voltage. The voltage at point 5 is a sequence of distortion-free half sinusoids of voltage whose amplitude is proportional to the DC level at pin 7 (the output of error amplifier EA1). Then, if V_o, say, goes up, the voltage at pin 7 goes down and the half sinusoid voltages at pin 5 get smaller in amplitude. Now the difference error voltage between pin 5 and ground (Fig. 15.8*d*) goes closer to ground, and so does the voltage at pin 3. Thus in the PWM comparator, the sawtooth crosses the voltage at pin 3 earlier in time, the on time for each switching cycle goes down throughout each half period, and in accordance with Eq. 15.1, V_o goes back up.

Thus the output at pin 5 contains the information needed to keep both V_o, the output voltage, constant and the input line current sinusoidal.

Current into pin 6 is sinusoidal and in phase with the input voltage because the impedance at that point is low and the large resistor $R8$ is driven by the sinusoidal voltage after the bridge rectifier.

15.4.4 Power output with the UC 3854

Figure 15.9 is the schematic for a 250-W power factor corrector using the UC 3854. The maximum obtainable output wattage is determined by setting the peak of the sinusoidal current I_{p1} which flows through sensing resistor R_s. This determines the maximum obtainable rms line input current and hence the maximum obtainable output power at any rms input voltage. Thus in Fig. 15.9, for lines above and below any terms indicating the maximum and minimum values for those terms,

$$P_o = EP_{\text{in}} = E\underline{V_{\text{rms}}}\,\overline{I_{\text{rms}}} = E\underline{V_{\text{rms}}}(0.707I_{p1}) \tag{15.2}$$

Where E is the efficiency and I_{p1} is the current flowing in the current-sensing resistor R_s at the peak of the sine wave at V_{rms}.

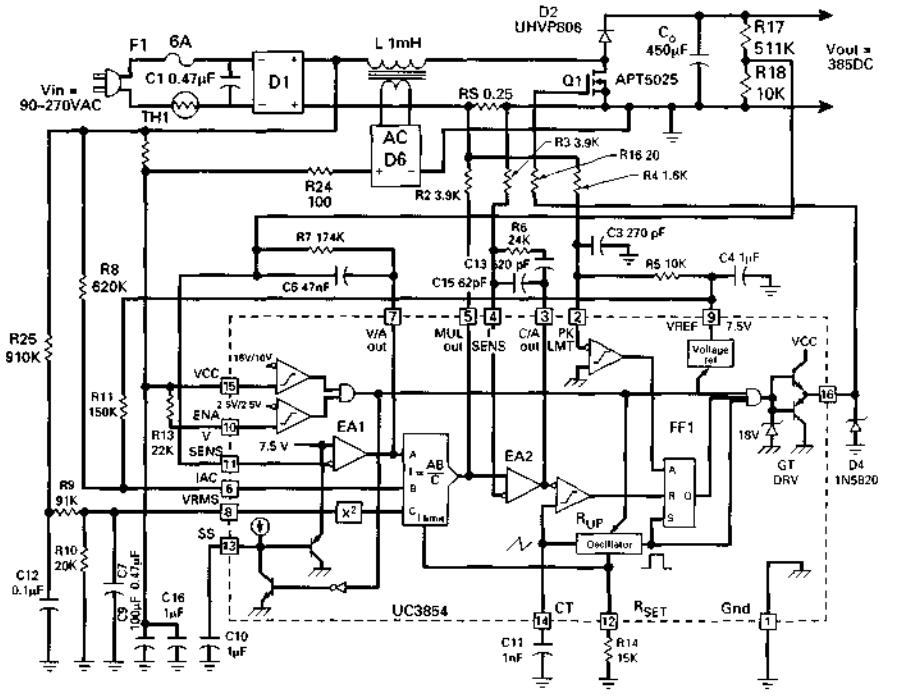


Figure 15.9 Detailed schematic for a 250-watt power factor controller using the Unitrode UC 3854 Control Chip. (*Courtesy Unitrode Integrated Circuits, Inc.*)

First I_{p1} is chosen from Eq. 15.2. Then R_s is selected to minimize its dissipation at low line and maximum load and yet to have a peak voltage drop at low line not less than 1 V. Then for, say, a 1-V peak drop across R_s,

$$R_s = \frac{1}{I_{p1}} \tag{15.3}$$

Now because of internal design details, I_{pmd}, the maximum current available at the *MD* output (pin 5), is fixed by

$$I_{pmd} = \frac{3.75}{R_{14}} \tag{15.4}$$

And I_{pmd} can be up to 0.5 mA but is usually set at 0.25 mA.

Now at all instants, the feedback loop keeps the voltage drop across R_s ($= R_s I_1$) equal to the rise across $R2$ ($= R2I_{md}$). And for maximum line and *MD* currents of I_{p1} and I_{pmd},

$$R2 = \frac{I_{p1}R_s}{I_{pmd}} \tag{15.5}$$

Thus for P_o = 250 W, V_{rms} = 90 V, and E = 0.85: From Eq. 15.2, I_{rms} = 250/0.85 × 90 = 3.27 A and I_{p1} = 1.41 × 3.27 = 4.61 A. And for a 1-V drop in R_s, from Eq. 15.3:

$$R_s = \frac{1.0}{I_{p1}} = \frac{1}{4.61} = 0.22, \text{ say, } 0.25\ \Omega$$

Then from Eq. 15.4 for an I_{pmd} of 0.25 mA,

$$R14 = \frac{3.75}{0.00025} = 15\ \text{k}\Omega$$

And from Eq. 15.5, for an I_{pmd} of 0.25 mA, an I_{p1} of 4.61 A, and an R_s of 0.25 Ω:

$$R2 = \frac{4.61 \times 0.25}{0.00025} = 4.61\ \text{k}\Omega$$

And to minimize drift in EA2 (Fig. 15.7), $R3$ is set equal to $R2$.

15.4.5 Boost switching frequency with the UC 3854

In addition to determining the current out of the M&D at pin 5, because of internal circuit details, $R14$ also sets the boost switching frequency. Once $R14$ has been fixed, the boost switching frequency is fixed by

$$F_s = \frac{1.25}{R14C11} \tag{15.6}$$

where $C11$ is the capacitor to ground at pin 14. For $R14$ in ohms and $C11$ in farads, F_s is in hertz. The UC 3854 can be used up to somewhat above 200 kHz but is generally used at closer to 100 kHz.

15.4.6 Selection of boost output inductor *L1*

Boost inductor $L1$ of Fig. 15.5 or 15.7 is chosen for a desired minimum line ripple current at the peak of the sinusoidal input voltage. It is chosen at maximum input power, minimum input voltage when the sine wave current peak I_{p1} is at a maximum. This ripple is ΔI; the ramp amplitude is either I_{Q1} or I_{D1} in Fig. 15.5. It is the current change in $L1$ for minimum voltage across $L1$ (V_p) when the $Q1$ on time T_{on} is a

maximum. Thus again for lines above and below the terms signifying their maximum and minimum values:

$$L1 = \frac{V_p \overline{T_{on}}}{\Delta I} = \frac{1.41 \underline{V_{rms}}\, \overline{T_{on}}}{\Delta I} \tag{15.7}$$

But

$$I_{p1} = \frac{1.41\, P_o}{E \underline{V_{rms}}} \tag{15.8}$$

Arbitrarily choose ΔI as 20 percent of I_p.

$$\Delta I = \frac{0.2 \times 1.41 P_o}{E \underline{V_{rms}}} = \frac{0.282 P_o}{E \underline{V_{rms}}} \tag{15.9}$$

Then from Eqs. 15.7 and 15.9,

$$L1 = \frac{5.0(\underline{V_{rms}})^2 E \overline{T_{on}}}{P_o} \tag{15.10}$$

Now from Eq. 15.1,

$$\overline{T_{on}} = T\left(\frac{1 - \underline{V_p}}{V_o}\right) \tag{15.11}$$

Now select V_o as 10 percent above $\overline{V_p}$—the sine wave peak at maximum rms input voltage. Then

$$\overline{T_{on}} = T\left(1 - \frac{\underline{V_p}}{1.1\overline{V_p}}\right) \tag{15.12}$$

But $V_p/\overline{V_p} = \underline{V_{rms}}/V_{rms}$. And taking $\underline{V_{rms}} = 90$ V and $\overline{V_{rms}} = 250$ V from Eq. 15.12,

$$\overline{T_{on}} = T\left(1 - \frac{90}{1.1 \times 250}\right) = 0.673T \tag{15.13}$$

And from Eqs. 15.10 and 15.13,

$$L1 = \frac{3.37(\underline{V_{rms}})^2 TE}{P_o} \tag{15.14}$$

Thus for V_{rms} = 90 V, frequency = 100 kHz (T = 10 μs), E = 85 percent, and P_o = 250 W, from Eq. 15.14

$$L1 = \frac{3.37(90^2)(10 \times 10^{-6})(0.85)}{250} = 928\ \mu H$$

15.4.7 Selection of boost output capacitor

Refer to Fig. 15.10. The boost capacitor C_o usually feeds a DC/DC converter—generally a half bridge for output powers under 600 W and a full bridge for higher powers. Recall that the nominal output voltage V_{on} is generally set at 10 percent above the peak at maximum rms input voltage V_{rms}. Then for $\overline{V_{rms}}$ = 250 V, $V_{on} = 1.1 \times 1.41 \times 250 = 388$ V. This voltage is not well regulated as the voltage error amplifier gain per bandwidth is kept low to improve the response to changes in load current. Hence assume the minimum output voltage $\underline{V_o}$ with AC input present is 370 V.

Now if AC input is lost at the instant V_o is at its minimum value, it is desired that C_o be large enough to hold up the output voltage to a value (V_{mhu}) still permitting all DC/DC converter outputs to remain within specifications for a time T_{mhu} (usually specified as 30 ms).

As V_o droops from $\underline{V_o}$ toward V_{mhu}, the DC/DC converter on time increases to maintain all its outputs within specification. A small droop of V_o would require a large capacitor, and a large permitted droop would force an increased on time too close to the maximum permissible value of a half period for most converter topologies. A usual compromise is for V_{mhu} to be chosen 60 to 80 V below $\underline{V_o}$ and to design the converter transformer to have secondary voltages high enough that the on time for an input of V_{mhu} is still only 80 percent of a half period (Sec. 3.2.2.1).

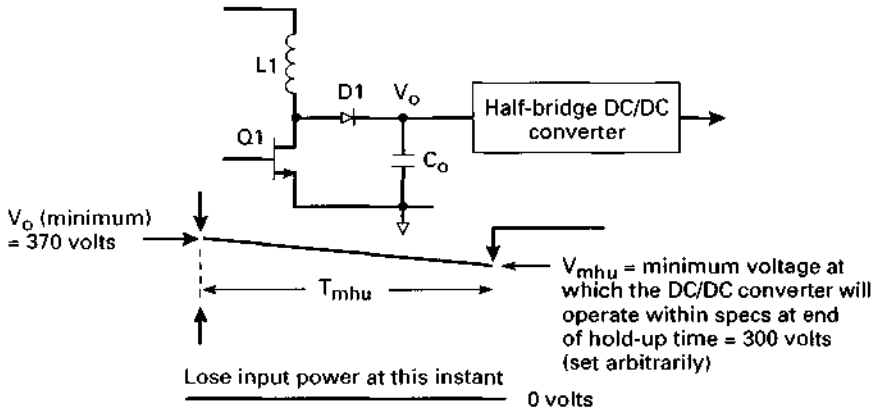


Figure 15.10 Selection of Capacitor C_o to meet a specified hold-up time.

Thus C_o is selected from

$$C_o = \frac{I_{av} T_{hu}}{\Delta V} = \frac{I_{av} T_{hu}}{V_o - V_{mhu}} \tag{15.15}$$

where I_{av} is the average current taken out of V_o during the droop from V_o to V_{mhu}. For converter output power of P_c and an efficiency of E_c,

$$I_{av} = \frac{2P_c}{E_c(V_o + V_{mhu})} \tag{15.16}$$

Thus for $V_{mhu} = V_o - 70 = 370 - 70 = 300$ V, $T_{hu} = 30$ ms, from Eq. 15.15,

$$C_o = \frac{I_{av} \times 0.03}{370 - 300} = 429 \times 10^{-6} I_{av}$$

And for $P_c = 250$ W and $E_c = 0.85$, from Eq. 15.16,

$$I_{av} = \frac{2 \times 250}{0.85(370 + 300)} = 0.88 \text{ A}$$

$$C_o = 0.88(429 \times 10^{-6}) = 378, \text{ say, } 400 \text{ mF}$$

Now the transformer in the DC/DC converter must be designed so that the outputs remain within specification at the above-specified V_{mhu} or 300 V. This will be safely achieved, as discussed in Sec. 3.2.2.1, if the number of secondary turns is sufficient to yield high enough secondary voltages that the required output voltages are obtained at an on time of 80 percent of a half period for $V_{mhu} = 300$ V.

But capacitor C_o, in addition to being specified to yield the desired holdup time, must be specified to have an adequate ripple current rating. It can be shown that the boost diode $D2$ (Fig. 15.7) current consists of a DC component (which is the DC load flowing out of V_o) plus a 120-Hz component whose peak amplitude is equal to the DC load. The DC component flows, of course, to the load, but the 120-Hz component flows into capacitor C_o. Thus the rms ripple current rating for C_o is $I_{rms} = 0.707 I_{dc}$. Thus for a DC/DC converter output of 250 W at $V_o = 388$ V and 85 percent efficiency, $I_{dc} = 250/[388(0.85)] = 0.76$ A. The rms ripple current rating for C_o is then $0.707 \times 0.76 = 0.54$ A.

15.4.8 Peak current limiting in the UC 3854

The peak limiting comparator at pin 2 and resistors $R4$ and $R5$ (Fig. 15.9) provide peak current limiting. Note in Fig. 15.9 that the power transistor $Q1$ is turned off and no further current is taken from $Q1$

when the FF1 is reset. The flip-flop is reset when the output of the current limit comparator goes positive. This occurs when the DC voltage level at the inverting input to the comparator (pin 2) falls below the voltage (ground) at its noninverting input. Resistors $R4$ and $R5$, fed from the +7.5-V reference at pin 9, provide an upward level shift so that pin 2 falls to ground and limits at a peak current I_{1p} such that

$$I_{p1}R_s = I_{R4}R4$$

But with pin 2 at ground:

$$I_{R4} = I_{R5} = \frac{7.5}{R5}$$

Then for, say, $R5 = 10\ \text{k}\Omega$, $I_{R4} = 0.75$ mA:

$$R4 = \frac{R_s I_{1p}}{I_{R4}} \tag{15.17}$$

Now for a P_o of 250 W, it was calculated above that the peak current I_p was 4.61 A. Then for peak current limiting at, say, 5.5 A, from Eq. 15.17, $R4 = 0.25 \times 5.5/0.00075 = 1.8\ \text{k}\Omega$.

15.4.9 Stabilizing the UC 3854 feedback loop

Stabilizing the feedback loops is beyond the scope of this discussion. It is sufficient to say that there are two feedback loops—a high-bandwidth inner one (EA2) which forces the input line current to be sinusoidal and a low-bandwidth outer one (EA2) that maintains constant output voltage.

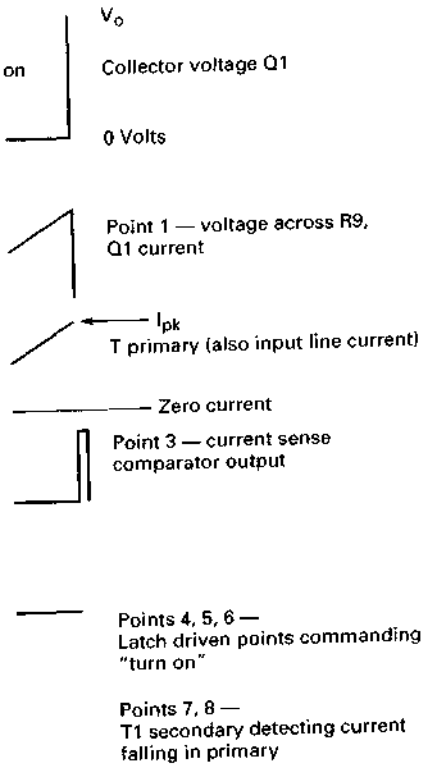
The EA2 (Fig. 15.7 or 15.9) is a type 2 linear amplifier (Sec. 12.6). It has a zero at $F_z = 1/(2\pi \cdot R6C15)$, a pole at $F_p = 1/(2\pi \cdot R6C13)$, and a pole at the origin at $1/[(2\pi \cdot R3(C13 + C15)]$.

Voltage error amplifier EA1, in addition to maintaining constant DC output voltage, minimizes harmonic distortion of the 60-Hz line current by having low bandwidth and low gain at and beyond the third harmonic of the line frequency.

Detailed design of these feedback loops is discussed in Refs. 1 through 4.

15.5 Motorola MC 34261 Power Factor Correction Chip

The Motorola MC 34261 is another widely used power factor correction chip (see Ref. 5) and is shown in Fig. 15.11. As the Unitrode chip,



la MC 34261 Power Factor Controller shown haversine. Points refer to locations in Fig.

4261 (Fig. 15.11)

fter the input bridge rectifier is re- (*C*6) which permits the voltage out- input line voltage in haversine fash- nd. This haversine must be cleanly e input line voltage. For it is this by the EA1 output voltage sensor, sense comparator to the line current aversines by *R*7) which forces input line current to be sinusoidal input line voltage.

The voltage at pin 3 at the h to prevent the internal circuitr rent into *R*9 is a sequence of d is equal to the peak of the up-s erage of the current triangle cycle is the average of the inp forcing the average of the curr gles by *R*9) of Fig. 15.12*c* to e reference haversines at the mu to be sinusoidal and in phase w

15.5.2 Logic details for the MC (Figs. 15.11 and 15.12)

Assume the RS latch had beer point 5. Output pin 7 goes hi The *Q*1 collector falls to grou starts rising at a rate $dI/dt = V$ rent flowing through *R*9 reach instantaneous voltage at the r sense comparator output (poin and point 4 goes low. This pu turning off *Q*1.

Now the *Q*1 collector (Fig. 1 the *T*1 primary commences fall 15.12*c*). Also as the dot end of t dot end of its secondary (point RS latch are positive, so both it are low. Importantly, point 4 re point 8) remains high. The lo point 5 low, and hence point 7 i as point 7 is high or as long as c

Now, when current in the *T*1 dot end of the secondary (point 7 the zero current detector (point output at point 4 goes high, driv to repeat the cycle. But the lat point 10 low after point 7 falls lo

There is a delay between poin point 8 falls to ground, point 9 Now after the delay, when point over at the input to NOR gate *X* latch with point 4 high, point 10

the voltage at point 1 above the instantaneous voltage from the multiplier at point 2.

15.5.3 Calculations for frequency and magnitude of inductor *L1*

In Fig. 15.12*c*, when $Q1$ is on, the voltage across $L1$ is V_{in} and current in it ramps up at a rate of $dI_r/dt = V_{in}/L1$. When $Q1$ is off, the voltage across $L1$ is $V_o - V_{in}$ and current in it (and in the AC line input) ramps downward at a rate $dI_f/dt = (V_o - V_{in})/L1$. If it rises to a peak I_p in a time T_{on}, then before it is turned on again, it must fall the same I_p in a time T_f, or

$$\frac{V_{in}T_{on}}{L1} = \frac{(V_o - V_{in})T_{off}}{L1}$$

then

$$T_{off} = \frac{T_{on}V_{in}}{V_o - V_{in}} \tag{15.18}$$

Here V_{in} is the instantaneous haversine input voltage as time progresses along the haversine. It simplifies matters to operate at a fixed on time and permit the off time to vary as in Eq. 15.18. Frequency $[-1/(T_{on} + T_{off})]$ will then vary across the haversine.

Inductance L will be chosen so as to require a reasonably small inductor which can tolerate the peak current for the required output power. For large values of inductances (say, over 1 mH) which do not saturate at currents over 2 A are large and expensive. Then

$$P_{in} = \frac{P_o}{E} = \frac{V_{rms}\,\overline{I_{rms}}}{E}$$

Current is a maximum for minimum voltage.

Then again for lines under and over a term signifying the minimum and maximum values of the term,

$$I_{rms} = \frac{P_o}{E\underline{V_{rms}}} = 0.707I_{pk} \qquad \text{or} \qquad I_{pk} = \frac{1.41P_o}{E\underline{V_{rms}}}$$

where I_{pk} is the peak 60-Hz input line current at $\underline{V_{rms}}$. But it is seen in Fig. 15.12*c* that since the current averaged over one switching cycle is only one-half of $I_{pk\,t}$, then $I_{pk\,t}$ must be $2I_{pk}$ or

$$I_{pk\,t} = \frac{2.82P_o}{E\underline{V_{rms}}} \tag{15.19}$$

This is the peak transistor ramp current at the peak of the input haversine at $\underline{V_{rms}}$. When the transistor turns on at the peak of the haversine, its current ramps up in a time T_{on} to

$$I_{pk\,t} = \frac{\underline{V_p} T_{on}}{L}$$

where $\underline{V_p} = 1.41\underline{V_{rms}}$. Then L is:

$$L = \frac{1.41\underline{V_{rms}}\, T_{on}}{I_{pk\,t}}$$

Then from Eq. 15.19, $I_{pk\,t} = 2.82P_o/E\underline{V_{rms}}$:

$$L = \frac{(V_{rms})^2 T_{on} E}{2P_o} \tag{15.20}$$

Now assume nominal, minimum, and maximum rms input voltages of 120, 92, and 138 V. For $\underline{V_{rms}} = 92$ V, $P_o = 80$ W, and $E = 0.95$,

$$L1 = \frac{(92)^2 \times 0.95 T_{on}}{2 \times 80} = 50T_{on} \tag{5.21}$$

If a T_{on} of 10 μs is selected, $L1$ is 500 μH from Eq. 15.21 which is reasonably small for a peak ramp current of $I_{pk\,t} = 1.41 \times 92 \times 10 \times 10^{-6}/(500 \times 10^{-6}) = 2.59$ A.

Now the boosted voltage must be above the sine wave peak at maximum line input. For $\underline{V_{rms}} = 138$ V, the sine wave peak is $1.41 \times 138 = 195$ V. If it is not boosted far enough above this, Eq. 15.18 shows the off time will be many times the on time and frequency will be low. This large frequency change must be balanced against the higher voltage stress on the off transistor and more recovery time losses in rectifier diode $D5$ for higher boost voltages.

Thus at low line of 92 V, the sine wave peak is $1.41 \times 92 = 129$V. If this is boosted as required to only 50 V above the sine wave peak of the high-line peak of 195 V, Eq. 15.18 shows the off time $T_{off} = T_{on}V_{in}/(V_o - V_{in}) = 10 \times 129/(245 - 129) = 11.1$ μs. This would yield a period of 21.1 μs, or 48 kHz. But at a high line of $V_{AC} = 138$, where $V_{in} = 1.41 \times 138 = 195$ V, at the peak of the haversine from Eq. 15.18

$$T_{off} = \frac{T_{on}V_{in}}{V_o - V_{in}} = \frac{10 \times 195}{245 - 195} = 39\ \mu s$$

Switching frequency F_s would then be $1/(T_{on} + T_{off}) = 20$ kHz. This change from 48 kHz at the peak of the 92-V AC line to 20 kHz at the peak of a 138-V line may not be desirable.

Yet the frequency change throughout a haversine is much greater. Assume $V_{AC} = 92$ V and the boost had been to 245 V—only 50 V above the sine wave peak at the high line of 138 V. As shown above, this yields a switching frequency of 48 kHz at the peak of the haversine. Now calculate the switching frequency close to the notch of the haversine, say, at the 10° point. Then at the 10° point on a 92-V AC line, V_{in} is $1.41 \times 92 \sin 10 = 23$ V. At that point, Eq. 15.18 shows $T_{off} = T_{on} \times 23/(245 - 23) = 0.1\ T_{on}$. And the switching frequency at the 10° point on the haversine is $1/10.1 = 99$ kHz.

Such frequency variations over high to low AC input voltages and especially over the haversine at a 120-Hz rate could present an RFI/EMI problem as they cover a wide frequency spectrum. Yet the MC 34261 chip has found very widespread acceptance throughout the industry.

15.5.4 Selection of current-sensing resistor (*R9*) and multiplier input resistor network (*R3*,*R7*) for the MC 34261

A Motorola data sheet recommends $R9$ be selected as $R9 = V_{cs}/I_{pk\,t}$, where $I_{pk\,t}$ is the peak transistor current (Eq. 15.19) and the current threshold V_{cs} is 0.5 V for $V_{AC} = 92$ to 138 V rms. From Eq. 15.19, for $P_o = 80$ W, $V_{rms} = 92$ V, and an efficiency of 95 percent,

$$I_{pk\,t} = \frac{2.82 \times 80}{0.95 \times 92} = 2.58\ A$$

then

$$R9 = \frac{0.5}{2.58} = 0.19\ \Omega$$

The data sheet suggests the multiplier voltage (chip pin 3) be 3.0 V (V_M) at the haversine peak at high line. Thus,

$$3 = \frac{1.41V_{AC}R3}{R3 + R7}$$

And for $V_{AC} = 138$ V,

$$\frac{R3}{R7} = 0.015\ \omega$$

It is suggested that this be a starting point. First set V_M at 3 V with a high value of $R7$ (say, 1 MΩ) and vary $R3$ for the lowest distortion in the AC line current.

References

1. Claudio De Silva, "Power Factor Correction with the UC 3854," Unitrode Corporation Application Note U-125.
2. B. Mammano and L. Dixon, "Choose the Optimum Topology for High Power Factor Supplies," PCIM, March 1991.
3. Motorola Inc. data sheet, "MC 34261 Power Factor Controller."
4. Motorola Inc. data sheet, "MC 34262 Power Factor Controller."
5. N. Nalbant and W. Chom, "Theory and Application of the ML 4821 Average Current Mode Controller," Application Note 16, Micro Linear Inc.
6. Toko Inc. data sheet, "TK 83854 High Power Factor Pre-regulator."
7. Silicon General data sheet, "Power Factor Controller."

High-Frequenc
for Flu

16.1 Why High-Frequency Power S

High-frequency power sources for
enormously high-volume market for
supplies. This may be appreciated
the mid-1980s, shipment of fluores
million annually.[1]

From the introduction of the fluor
1970s, fluorescent lamps were drive
line via a series inductor or via a 60
combination. The inductor or tran
ferred to as *magnetic ballast* (Fig. 16

The series inductor is essential a
ment because the lamp's volt/ampe
tive input impedance. The signific
input current increases, its voltage
from a low-output-impedance sou
lamp current would cause a decrea
lamp voltage would cause a furthe
would continue thus until the incre
to fail. Or the lamp voltage would fa
rupt the current flow, as the lamp
age to maintain current flow.

The 300 million fluorescent lamps
mid-1980s thus required an equal n
which cost only $5 each at that volu
drawbacks to the simple scheme of p

Chapter

16

High-Frequency Power Sources for Fluorescent Lamps

16.1 Why High-Frequency Power Sources?

High-frequency power sources for fluorescent lamps represent an enormously high-volume market for the technique of switching power supplies. This may be appreciated from the statistic that, as about the mid-1980s, shipment of fluorescent lamps of all types was 300 million annually.[1]

From the introduction of the fluorescent lamp in 1938 until the late 1970s, fluorescent lamps were driven directly from the 60-Hz power line via a series inductor or via a 60-Hz step-up transformer/inductor combination. The inductor or transformer/inductor element is referred to as *magnetic ballast* (Fig. 16.1).

The series inductor is essential as a lossless current-limiting element because the lamp's volt/ampere characteristic exhibits a negative input impedance. The significance of this is that as the lamp input current increases, its voltage drop decreases. Hence if it is fed from a low-output-impedance source, any momentary increase in lamp current would cause a decrease in lamp voltage and the lower lamp voltage would cause a further increase in lamp current. This would continue thus until the increased lamp current would cause it to fail. Or the lamp voltage would fall to such a low value as to interrupt the current flow, as the lamp requires a certain minimum voltage to maintain current flow.

The 300 million fluorescent lamps shipped annually from 1938 to the mid-1980s thus required an equal number of such magnetic ballasts which cost only $5 each at that volume. But it was realized there were drawbacks to the simple scheme of powering the fluorescent lamp at 60

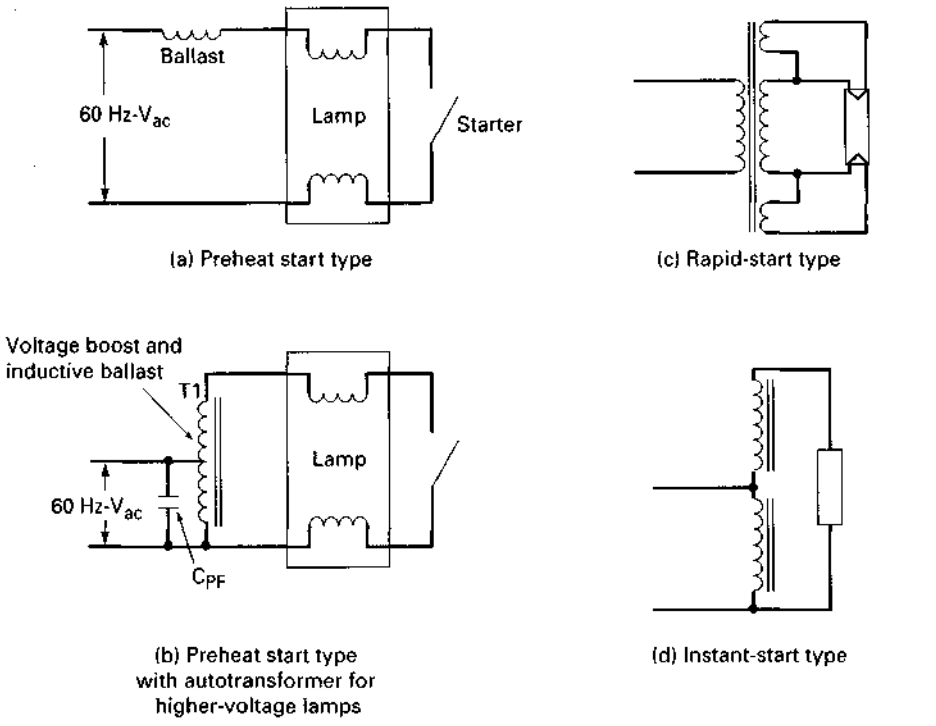


Figure 16.1 Fluorescent lamp types with magnetic ballasts. When the lamp is fed from the 60-Hz power line, "ballast" or current limiting, is required in series with the lamp. Current limiting is required because the lamp itself has a negative input impedance. For 60-Hz-driven lamps, the ballast is an inductor, as in Fig. 16.1*a*, or the leakage inductance of the autotransformer, as in Fig. 16.1*b*. In "preheat" lamps, the switch is momentarily closed to heat the filament and then released. Rapid-start lamps have their filaments powered at starting and also afterward. Instant-start lamps have no heated filaments but depend on field emission to supply the electrons to initially start the lamp. This saves filament heating power, but the necessarily higher starting voltage for instant-start types results in shorter lifetimes.

Hz through an inexpensive, passive magnetic element. First, because of the nature of the lamp, it extinguishes at a 120-Hz rate at each zero crossing of the 60-Hz waveform and produces a noticeable flicker and hence decreased average light intensity. Second, iron laminations of the 60-Hz inductor/transformer tend to vibrate and cause an audible buzz which requires potting in tar to suppress. The typical weight of such a 60-Hz, potted magnetic ballast is 3 to 4 lb, and although this is acceptable in the industry, it is a significant shortcoming. Also the inductor/transformer and potting tar increase dissipation.

These drawbacks led to studying the use of higher-frequency alternating current to power the lamp. The advantages sought were increased lamp efficacy (light output power in lumens per watt of input

power) and the possibility of replacing the noisy and bulky 60-Hz magnetic ballast with a smaller and lighter current-limiting capacitor ballast.

Early experiments showed that lamp efficacy increased with increasing frequency, reached a peak of about 14 percent at 20 kHz, and stayed constant thereafter (Fig. 16.2). Also the capacitor ballast was

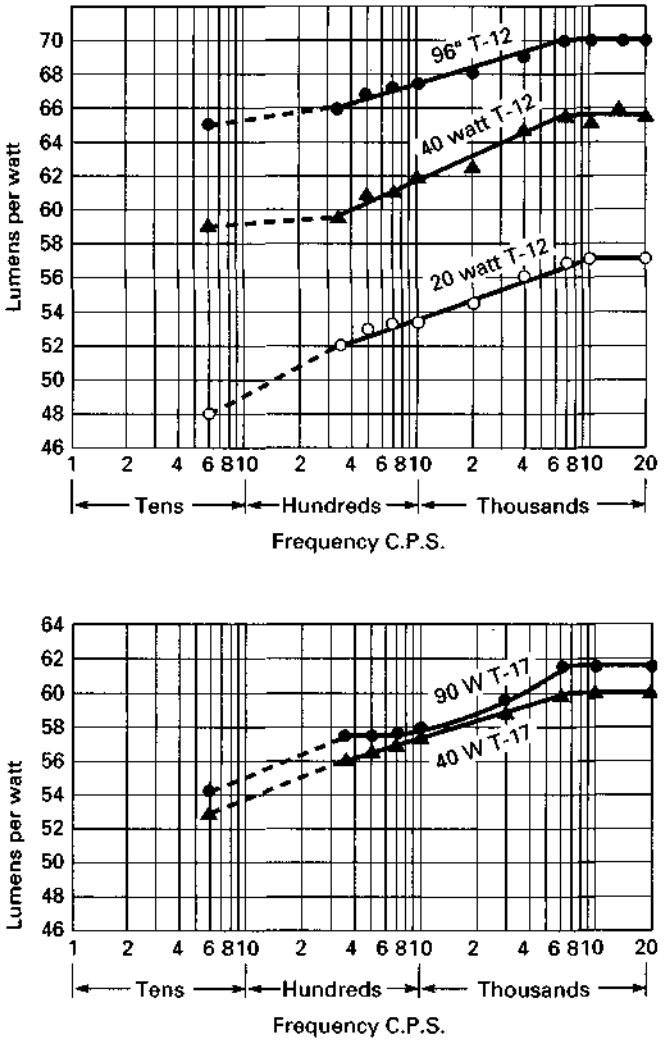


Figure 16.2 Fluorescent lamp light output in lumens per watt. For T12(1.5-in lamp diameter) and T17(1.88-in lamp diameter) versus frequency. (*From High-Frequency Fluorescent Lamps,* Campbell, Schultz, Kershaw, Illumination Engineering, *Feb. 1953.*)

smaller and lighter, had no audible noise, and was less expensive. And at this high frequency, the lamp showed no flicker and conducted and radiated EMI was easier to suppress.

These advantages of high-frequency operation, though significant, could not be realized until a practical, inexpensive method of generating a high-frequency source was available. Small local SCR inverters for each lamp were considered as well as larger ones or even high-frequency rotating AC generators for the large banks of lamps in factories or large office buildings.

But with the rapid drop in transistor and ferrite core prices, instead of a large central high-frequency power source, local DC/AC inverters for each lamp or each two-to-four lamp assembly, powered from the rectified AC power line, became an obvious possibility.

Most of these goals have been achieved with these high-frequency electronic ballasts, but their individual cost (about $15) is still greater than that of an equivalent magnetic ballast ($5). But when lower power cost because of the greater lumens per watt, longer lamp lifetime, and improved performance characteristics are considered, the high-frequency electronic ballast is increasingly replacing magnetic types for new installations.

And because the electronic ballast is interchangeable with magnetic ones without replacing the fluorescent lamps in old installations, there is a large market in replacing the enormous number of magnetic ballasts in large office buildings or factories.

It is estimated that the cost of replacing the magnetic ballast with an electronic ballast is recovered by the reduced power costs in about 1 year. The total reduction in power costs with the new electronic ballasts can amount to 20 to 25 percent when all factors are considered. Contributions to the lower power cost come from the improved lamp efficacy due to flicker elimination, improved efficacy arising from the ability to operate lamps at higher current where efficiency is higher, and lower losses with a capacitor ballast rather than a magnetic one.

And the continuing lower power costs after changeover make changing to electronic ballasts financially very attractive.

One major manufacturer reports replacing magnetic ballasts with electronic ones for a building with 960 eight-foot fluorescent lamps. Each lamp/ballast fixture now drew 87 W rather than 227 W for the same lumens of light power. At 6000 h of lamp operation per year and $0.106 per kilowatthour, the annual savings in power cost for the 960 lamps amounted to $86,088. And even taking into account the higher cost of an electronic ballast and the changeover labor cost, the manufacturer recovered costs in 1 year. The manufacturer also estimates another $8814 saving in decreased burden on the air-conditioning system.[2]

16.2 Fluorescent Lamp—Physics and Types

Most fluorescent lamps come in tubes of standard diameters, lengths, and wattage ratings. Diameters come in increments of $\frac{1}{8}$ in, and standard lengths are 2, 3, 4, and 8 ft. The various manufacturers all use a common four-character type code *FPTD,* where *F* designates fluorescent, *P* designates input power in watts, *T* signifies tubular, and *D* is diameter in eighths of an inch. Thus an F32T8 is a 32-W, 1-in-diameter lamp; an F40T12 is a 40-W, 1.5-in-diameter lamp. Fluorescent lamps are also available in circular and U shapes.

Electronic ballasts available from the various manufacturers are designed to operate the various lamp types from the same line voltage as appears in the type code for their original magnetic ballast usage.

Fluorescent lamps consist of glass tubes in the above-described standard lengths, diameters, and shapes and are filled with argon or krypton gas at low pressure. A small amount of liquid mercury, which will vaporize when warmed by an initial low-energy arc in the argon or krypton gas, is also enclosed (Fig. 16.3).

The inside of the lamp is coated with various phosphors which emit the desired visible light when irradiated with ultraviolet light generated by a high-current mercury *arc* flowing through the lamp. The high arc current is drawn through the lamp by a high voltage applied across electrodes at each end.

These electrodes are simply passive unheated coils of wire in instant-start lamps or current-carrying, oxide-coated filaments in rapid-start lamps.

Before electrode voltage is applied, there are relatively few current carriers in the lamp, as none of the gas molecules are ionized. The current carriers will be supplied in quantity when a few free electrons in the gas are accelerated to high speeds by the voltage across the end electrodes. An accelerated electron, colliding with a neutral gas atom, ionizes it, providing a free electron and a massive positively charged ion. The electron accelerated toward the anode and the positive ion toward the cathode now produce more *ionization by collision.* Each collision produces more current carriers, and each current carrier causes more ionizing collisions. An avalanche of current or *arc* results.

Instant-start (Fig. 16.1*d*) lamps do not have a large initial supply of free electrons to start the above process. They depend on the few free electrons produced by cosmic rays and a large voltage gradient at the end electrodes to strip enough electrons to start the ionization-by-collision process. They consequently require a high electrode voltage to start or "strike" the arc, but start instantly on application of the voltage.

Rapid-start (Fig. 16.1*c*) lamps have current-carrying, oxide-coated filaments which supply a large reservoir of electrons to initiate the arc. They require a lower accelerating voltage to light the lamp, do not light

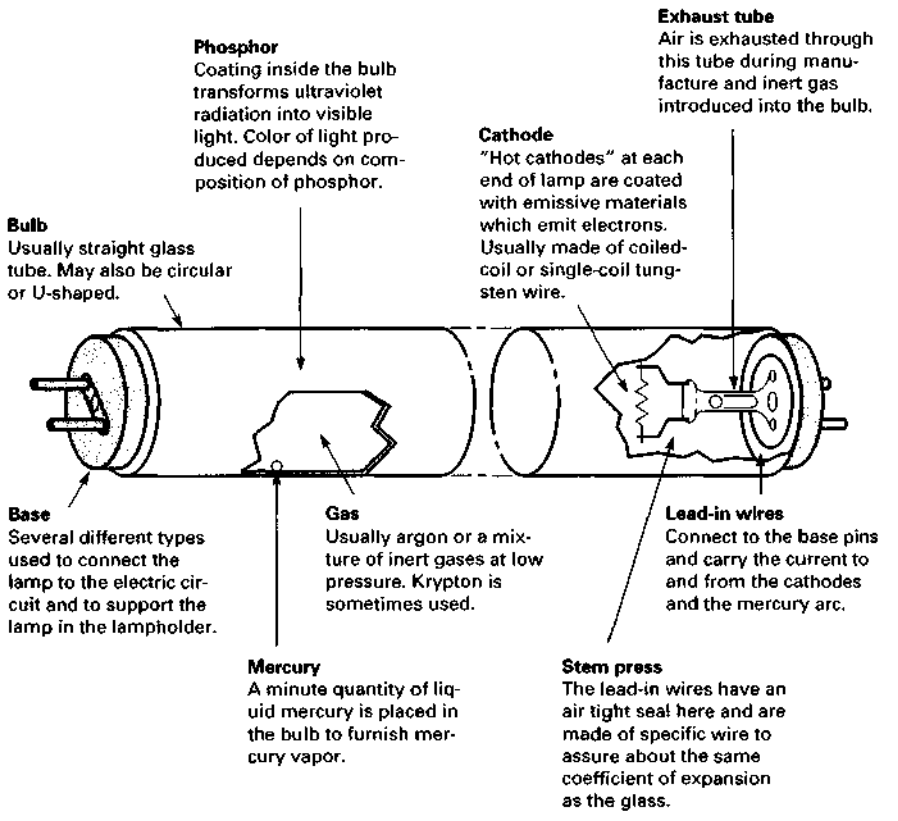


Figure 16.3 A typical rapid-start fluorescent lamp. Heated filaments at each end boil off a large supply of free electrons to permit starting the lamp at lower voltages. Instant-start lamps do not use a heated filament to supply the pool of electrons for starting. Instead, they use a higher voltage across the end electrodes to supply the electrons by field emission. This saves filament heating power but results in lesser lamp lifetime. (*Courtesy GTE/Sylvania, "Fluorescent Lamps."*)

as quickly as instant-start types, but are yet acceptably fast. Rapid-start lamps remain powered by filament current after the lamp starts. Older "preheat" lamps had filaments which were heated at turn-on but used a *starter* device to automatically turn off filament power after the lamp lighted, to conserve power.

Instant- and rapid-start lamps have their pluses and minuses. The rapid-start type is slower in starting, yet acceptably fast and requires lower starting voltage than the instant-start type. But its filament makes it more expensive to manufacture and needs a source of filament power which may require a separate filament transformer or separate winding on the power transformer. The instant-start lamp with its higher striking voltage strips material from the cathode at each start. After many starts, this darkens the lamp at its ends and reduces

its lifetime. But considering the infrequent starts in its usual applications, its lower cost makes it competitive. Both types are commercially available and require a ballast tailored to their characteristics.

The purpose of the low-pressure argon or krypton gas in the lamp is to achieve faster lamp starting. Initially, the ionization by collision or arc is started in those gases. But as the temperature rises due to the low-energy arc in the argon or krypton gas, the mercury vaporizes and ionization by collision of its atoms produces significantly more electrons and positively charged ions. But more importantly, it is the mercury vapor atoms which produce the ultraviolet light that stimulates the phosphors to emit visible light.

Now electrons driven to the anode and positively charged mercury ions driven to the cathode collide with other mercury atoms and excite some of their orbital electrons to various higher energy levels. On falling back to their original energy levels, radiation is emitted at frequencies corresponding to the energy level differences between transitions. Transitions across lower-energy-level differences produce relatively long-wavelength, visible light; those across higher-energy-level differences produce higher-energy, short-wavelength ultraviolet light.

One particular transition produces high-energy ultraviolet light at a wavelength of 2537 angstroms (1 Å = 10^{-8} cm). This is below the visible spectrum of about 4000 to 7000 Å, but the high energy at such short wavelengths is very effective in producing visible light at the phosphors.

Figure 16.4 shows the spectral distribution of energy in watts per nanometer (1 nm = 10 Å) from a 40-W white fluorescent lamp. The smooth curve represents the continuous spectrum emitted by the phosphor powders stimulated by the ultraviolet light. The discrete bands, 10 nm wide, represents radiation emitted by the mercury atoms in transitions between low-energy-level differences. These transitions produce visible light but are not as effective in generating light from the phosphors as the mercury atom energy level transition at 2537 Å. Most of the visible light output comes from the phosphors stimulated by the 2537-Å ultraviolet light.

16.3 Electric Arc Characteristics

An overview of further details of an electric arc in a gas is presented here. Although not essential, it is of value to the electronic ballast circuit designer in making some design decisions.

The nature of electric conduction in gases began to be studied intensively at the end of the 19th century. This led to understanding the nature and properties of the electron, the beginning of the understanding of atomic structure, and to X rays for medical diagnosis.

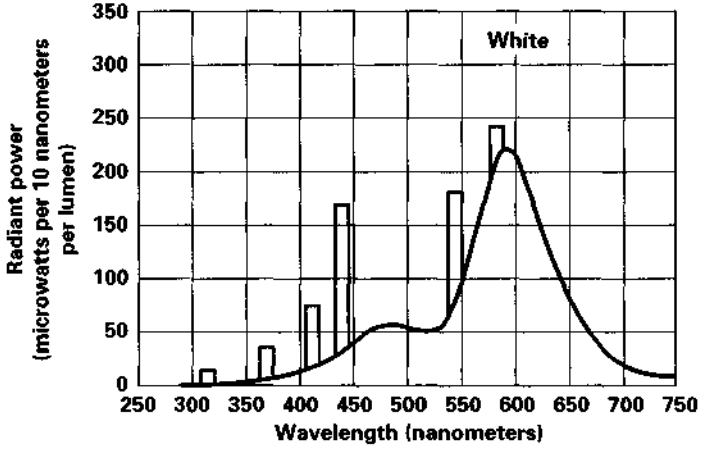


Figure 16.4 Spectral energy distribution from a 40-W white fluorescent lamp. In microwatts per nanometer (1 nm = 10 Å). The smooth curve is the continuous spectrum of energy generated by the white phosphorous. The discrete bands, 10 nm wide, represent energy generated by the mercury atoms in transition from a high to a low energy level. (*Courtesy "General Electric Bulletin Fluorescent Lamps."*)

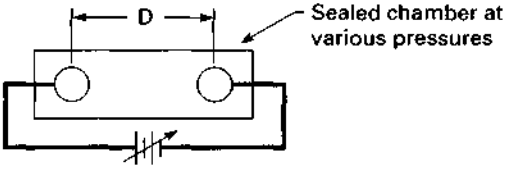


Figure 16.5 Paschen's classic experiment for measuring sparking potential between a pair of spherical electrodes at various spacings as a function of air pressure in the chamber.

Paschen, in 1889, studied the DC voltage required to initiate a spark between a pair or electrodes in an enclosed gas (air) as a function of pressure. He used spherical electrodes of diameter larger than the electrode separation (Fig. 16.5), to avoid high-voltage gradients in the vicinity of sharp points or edges. His result—the well-known Paschen's law—is shown in Fig. 16.6. For an electrode spacing of about 0.3 to 0.5 cm, at atmospheric pressure the voltage required to initiate a spark is close to 1000 V. As pressure is decreased, the sparking potential falls continuously toward a minimum of about 300 V and then rises steeply again. Gases other than air exhibit the same general characteristics—the minimum or critical pressure may be different for other gases.

Paschen's law offers an experimental explanation of the phenomena described above—ionization by collision. At high pressures, the mean spacing between neutral atoms is so small that before an ion—an

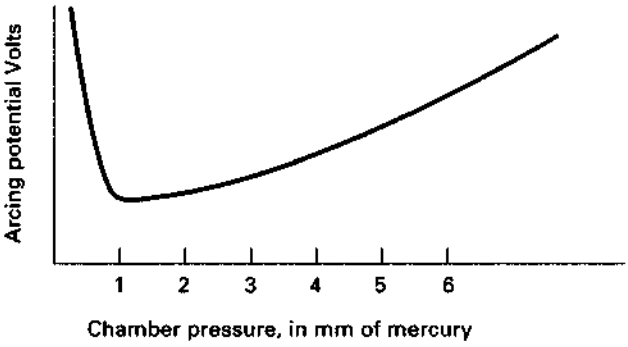


Figure 16.6 Classical Paschen curve. Required potential between electrodes to start an arc discharge. At about 3 to 5 mm of spacing in air, arcing potential minimum is about 300 V.

electron or a positive ion—can be accelerated to a velocity sufficient to ionize a neutral atom, it is slowed down or deflected by collision with some other atom. As pressure is reduced, the mean free path between atoms is increased and accelerating electrons or positive ions can travel longer distances and gather speed before being slowed down by collisions. Now when they finally collide, their energy is sufficient to ionize neutral atoms, there is an avalanche of charge carriers, and a spark or arc occurs.

16.3.1 Arc characteristics with DC supply voltage at the electrodes

As early as the late 19th century, physicists studied the visible appearance of an arc discharge with DC voltage at the electrodes. Their observations were as follows:

Their original experiments used solid electrodes—not heated, electron-emitting cathodes—at each end of a glass tube. A high voltage (several hundred volts) was applied across the electrodes through a current-limiting resistor. As the air pressure inside the tube was decreased sufficiently, they observed the following pattern of light and dark regions stretching out from the cathode to the anode (Fig. 16.7). Up against the cathode there was a short glowing region CG followed by a longer dark segment CDS. This was followed by a longer glowing region NG, then by an equally long dark region FDS. Following the FDS and up to the anode, there were alternate bands (PC) of luminosity separated by dark spaces. These regions were subsequently named *cathode glow* (CG), *Crookes' dark space* (CDS), *negative glow* (NG), *Faraday dark space* (FDS), and *positive column* (PC).

This was explained in subsequent years as follows:

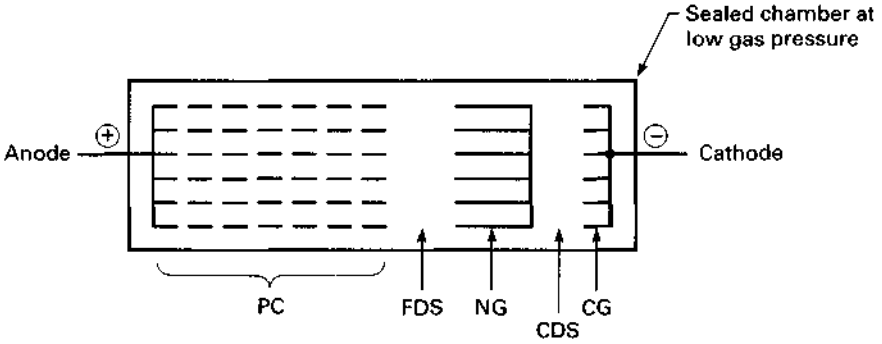


Figure 16.7 Pattern of light and dark regions in an arc discharge.

As the minimum pressure on the Paschen curve was approached, stray free electrons (from cosmic rays or by a high-voltage gradient close to the cathode) were accelerated sufficiently to ionize neutral gas atoms. The resulting positive ions, being massive, do not move very rapidly or far from the cathode and build up a positive space charge and hence a high-voltage gradient near the cathode. This voltage gradient—nowadays called the *cathode fall*—accelerates positive ions and drives them into the cathode, sputtering off some of its material.

When neutral or ionized mercury atoms close to the cathode are bombarded with electrons of sufficient energy, some of their orbital electrons absorb this energy and are driven up to a higher energy level still within the atom. When these electrons fall back to their initial energy level, they emit the visible light seen in the CG region.

At the outer edge of the cathode glow, all the electrons on their way to the anode have given up their energy and slowed up to the point where they can no longer excite atoms to higher energy levels. Throughout the Crookes dark space, they are being accelerated again; and at the edge of the negative glow region, they again have sufficient energy to excite atoms to higher energy levels. Throughout the negative glow, those atoms emit visible radiation as they fall back to their initial energy state. This happens again as the electrons speed up through the Faraday dark space, where they do not have sufficient energy to excite the atoms to luminosity.

And at the start of the positive column, there is again a luminous region. Thereafter, throughout the positive column, there are alternate dark and luminous regions. The dark spaces are the speed-gathering regions, the bright ones occur where the electrons have achieved enough energy to stimulate the atoms to emit visible light on their fallback to their initial energy level. Most of the voltage applied across the electrodes is dropped across the positive column, which occupies 80 to 90 percent of the distance between them.

16.3.2 AC-driven fluorescent lamps

Electrodes in a fluorescent lamp are driven by an AC voltage—60 Hz with a magnetic ballast and 20 to 40 kHz for an electronic ballast. Thus throughout a cycle, each end of the lamp is alternately an anode and a cathode, and the above pattern of bright and dark regions flips and flops right and left so that the brightness striations are not visible through the phosphor coating.

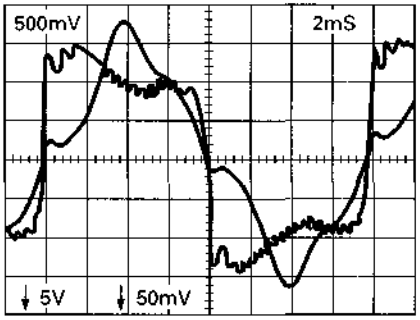
Some have proposed to use the secondary of a flyback transformer to power a fluorescent lamp. This appears attractive at first as the output voltage of a flyback secondary after the rectifying diode is unlimited in amplitude and is easily high enough to strike an ark in most fluorescent lamps. After the arc is struck, the lamp voltage falls back to its operating voltage, which is generally in the range of 100 to 300 V.

This could be done, but the lamp lifetime would be short because the same end of the lamp would always be the cathode and in not many hours of usage, that end would darken. It has been pointed out that there is a large voltage gradient at a cathode because of the positive space charge accumulating close to it. This voltage gradient would drive heavy positive ions into the cathode, sputtering away some of its material, and would soon darken the lamp close to the cathode. Thus opposite ends of a fluorescent lamp must always alternate as anode and cathode.

It can be appreciated from all the above that a fluorescent lamp driven from a high-frequency source produces more light and is easier to start than one driven from a 60-Hz source at the same input power level. At 60 Hz, there is no voltage across the lamp at the zero crossings of the input sine wave. The lamp thus extinguishes, and the arc must be restruck twice per cycle at the zero crossings of the input sine wave. This lowers the average light output and makes it more difficult to restart, especially at low temperatures. When driven at frequencies above 20 kHz, the ionized atoms do not have time to recombine at the zero crossings and the lamp does not extinguish, but maintains its light output.

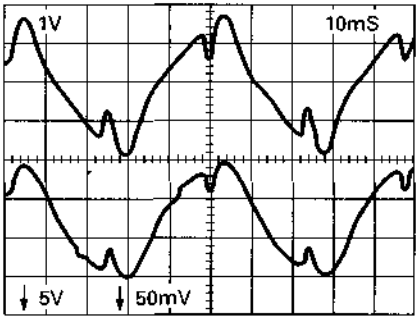
The greater dissipation in a lamp driven from a standard 60-Hz magnetic ballast compared to a 25-kHz electronic ballast can be seen in Fig. 16.8*a* and *b*. In Fig. 16.8*a* the lamp is driven by a magnetic ballast at 60 Hz. There it can be seen that shortly after the zero crossing of the current sine wave, the lamp voltage rises steeply to ignite the lamp.

These high-voltage moments shortly after the zero crossings waste power. In contrast, Fig. 16.8*b* shows for a 25-kHz electronic ballast, there are no such high-voltage intervals immediately after the current zero crossings. Also with the electronic ballast, voltage and current waveforms are fairly sinusoidal and in phase. Further, in the 60-Hz magnetic ballast, the *crest factor*, or ratio of peak to rms current, is



SB 60Hz
Lamp 100V/.4A/cm

(a)
Voltage across and current in a fluorescent lamp fed from 60-Hz power with a magnetic ballast.



EB 25kHz
Lamp 100V/.35A/cm

(b)
Voltage across and current in the same lamp driven from a 25-kHz power with a capacitor ballast.

Figure 16.8 Lamp waveforms. Note with a 60-Hz ballast, at every zero crossing of the current waveform, voltage across the lamp rises steeply. For at the zero crossing, the lamp extinguishes, its impedance increases, and the voltage across the lamp must rise to a high value to reignite the lamp. With a high-frequency power source, the lamp never extinguishes at the zero crossings and the lamp voltage is instantaneously proportional to the lamp current. With a 60-Hz power source, the lamp extinguishing at every zero crossing results in diminished light output power efficiency. (*From R. J. Haver,* Power Conversion/Intelligent Motion, *April 1987.*)

much higher than that for the 25-kHz electronic ballast. It has been widely reported in the literature[3] that high current crest factors yield poor lamp efficiencies. A perfect sine wave has a crest factor of 1.41.

It is estimated (Fig. 16.9) that in a 40-W fluorescent lamp at 60 Hz with a magnetic ballast, 23 percent of the input energy or 9.3 W is converted to visible light via the conversion to ultraviolet light and the consequent stimulation of the lamp phosphors. And 41 percent or 16.3 W is converted to convected and conducted heat, leaving 36 percent or 14.4 W radiated as infrared energy. For comparison, an incandescent light bulb (300 W) yields 11 percent of its input power as visible light and 89 percent as heat. In terms of *light efficacy,* a fluorescent lamp delivers 75 lm/W (up to 90 to 100 lm/W for the newest lamps with electronic ballasts[4]) compared to 18 lm/W for an incandescent lamp.

16.3.2 Fluorescent lamp volt/ampere characteristics

Before a fluorescent lamp lights or the arc is struck, it has high impedance as there are few current carriers. It takes a high voltage—V_{ns}, the *nominal striking voltage*—to light the lamp. But after lighting, the voltage across it falls to a lower *operating voltage* V_{op}. The operat-

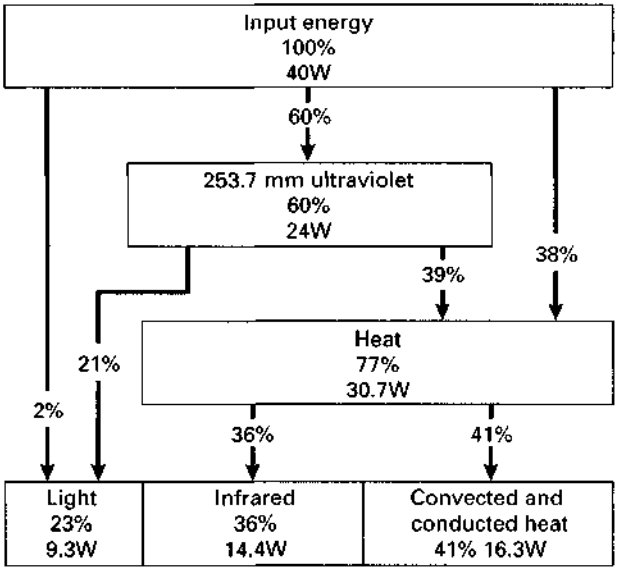


Figure 16.9 Distribution of energy in a 40-W fluorescent lamp. (*Courtesy GTE/Sylvania Bulletin, "Fluorescent Lamps."*)

ing current I_{op} drawn by the lamp is determined by X_b (the impedance of the ballast at operating frequency) and is given by

$$I_{op} = \frac{V_{ns} - V_{op}}{X_b} \tag{16.1}$$

The voltages and currents here are the rms values, and hence the actual power drawn by the lamp is simply

$$P_{in} = V_{op} I_{op} \tag{16.2}$$

The nominal striking voltage V_{ns} is given by the lamp manufacturer, usually at 50°F. But to ensure lighting the hardest-to-start lamp, the striking voltage at its minimum should be about 10 percent higher to accommodate manufacturing tolerances. The American National Standards Institute (ANSI), in "Fluorescent Lamp ANSI Specifications,"[5] sets the values for V_{op} and I_{op} for each specific lamp type so that the product meets the maximum wattage rating for that lamp.

Thus from V_{op} and I_{op} obtained from the ANSI specifications and V_{ns} on the manufacturer's data sheet, Eq. 16.1 fixes the ballast impedance X_b. And for electronic ballasts where the ballast is a capacitor, its value is calculated from

$$X_b = \frac{1}{2\pi f C_{bT}} \tag{16.3}$$

where C_{bT} is the value of the thevenized capacitance in series with the lamp, as the lamp may be driven from the junction of two capacitors.

Note that by choice of the ballast impedance any lamp can be operated at power levels higher or lower than the maximum value rated for that type—or at any given power level, including the maximum, at different combinations of V_{op} and I_{op}. Thus, for any desired power level, I_{op} can be arbitrarily selected and V_{op} calculated from Eq. 16.2. Then from V_{ns} and these values of I_{op} and V_{op}, the ballast impedance X_b is calculated from Eq. 16.1. Then for a capacitor ballast, C_{bT} is calculated from Eq. 16.3.

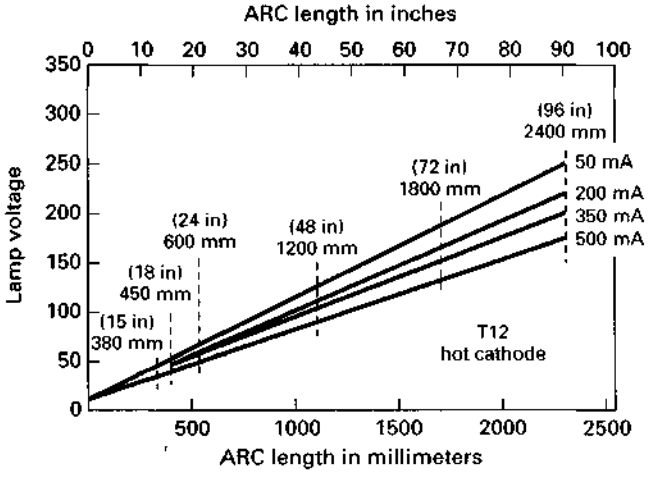
But although any fluorescent lamp can be operated at input power levels greater than those specified in the ANSI specification to yield more lumens of light output power, its lifetime will be decreased. Lamp lifetime may also be less than the manufacturer's specification even if it is operated at the specified wattage, but at currents different from the specified value. The manufacturer's specified lifetime is based on life tests operating at her or his specified currents and voltages.

Figure 16.10*a* and *b* shows lamp operating voltages and currents for a number of T8 and T12 hot-cathode (rapid-start) lamps. The negative input impedance of a fluorescent lamp can be seen. For a constant arc length (about 90 percent of the lamp length), as the lamp current increases, lamp voltage decreases, but input power increases. Figure 16.11*a* and *b* show the important ANSI specifications (V_{op}, I_{op}, and V_{ns}) for a number of instant- and rapid-start lamps. Note for instant-start lamps in Fig. 16.11*a*, again the negative input impedance, that as operating input current increases, operating voltage decreases.

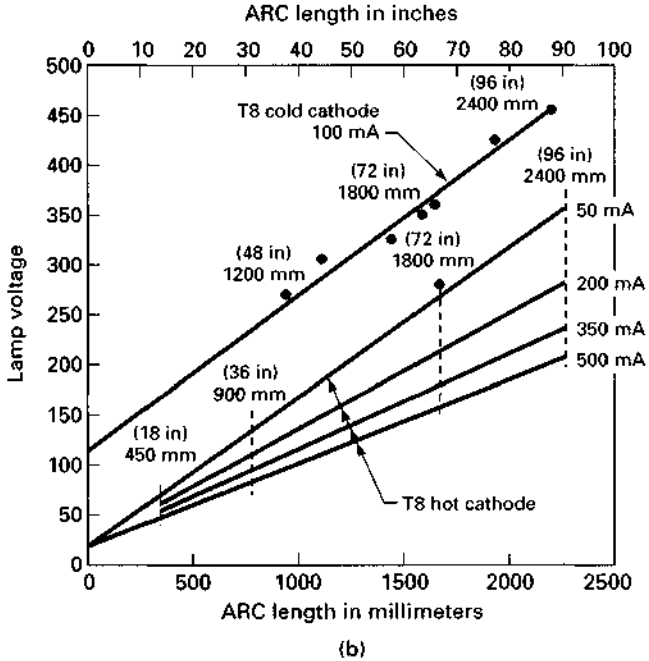
16.4 Electronic Ballast Circuits

The basic block diagram of a modern electronic ballast is shown in Fig. 16.12. The DC/AC converter which drives the lamp is not powered from the AC power line directly, but through a power factor correction building block (Chap. 15).

Recall from Chap. 15 that without power factor correction, the input bridge rectifier requires a large filter capacitor directly at its output. This capacitor results in line current pulses which are high in amplitude, narrow with fast rise and fall times, and with high harmonic content. The fast risetime line current pulses cause troublesome EMI and RFI problems at adjacent electronic equipment. The rms value of the nonsinusoidal line current is higher than that required to supply the actual DC load and consequently causes exces-



(a)



(b)

Figure 16.10 Fluorescent lamp operating voltages and currents. The source impedance and ballast impedance determine the ballast operating voltage at the manufacturer's specified operating current. Operating at currents lower than the manufacturer's specified value results in less input and less light power output. At higher currents than specified, power input and light power output increase but lamp lifetime is decreased. (*From "Fluorescent Lamp Light Sources,"* Illumination Engineering Magazine.)

Nominal Power (Watts)	T Size	Length (Inches)	V op (Volts)	I op (Amperes)	Vop x I op (Watts)	V st (Volts)
Rapid Start:						
25	T8	36	100	0.265	26.5	230
30	T12	36	77	0.43	33.1	205
32	T8	48	137	0.265	36.3	204
34	T12	48	79	0.46	36.3	260
37	T12	24	41	0.80	32.8	325
95	T12	96	126	0.83	106	
100	T12	84	135	0.80	108	280
113	T12	96	153	0.79	121	295
113	T12	96	139	1.00	139	295
116	T12	48	84	1.50	126	160
168	T12	72	125	1.50	188	225
215	T12	96	163	1.50	245	300
Instant Start:						
40	T12	48	104	0.425	44	385
40	T12	60	107	0.425	39	385
60	T12	96	157	0.44	69	565
38	T8	72	195	0.300	59	540
38	T8	72	220	0.200	44	540
38	T8	72	245	0.120	29	540
51	T8	96	263	0.300	79	675
51	T8	96	295	0.200	59	675
51	T8	96	325	0.120	39	675

(a)

Lamp Type	MA – RMS Current	Volts-RMS Voltage	Watts Power
48" Hot Cathode – T12	500	75	37.5
"	350	95	33.3
"	250	110	22.0
"	50	120	6.0
72" Hot Cathode – T12	500	120	60
"	350	140	49
"	200	160	32
"	50	185	9.3
96" Hot Cathode – T12	500	160	80
"	350	196	67
"	200	210	42
"	50	246	12
72" Hot Cathode – T8	500	150	75
"	350	175	61
"	200	210	42
"	50	265	13
96" Hot Cathode – T8	500	200	100
"	350	240	84
"	200	280	56
"	50	350	18
48" Cold Cathode – T8	100	260	26
72" Cold Cathode – T8	100	380	38
96" Cold Cathode – T8	100	460	46

(b)

Figure 16.11 (*a*) American National Standards Institute (ANSI) specifications for various fluorescent lamps; (*b*) volt/ampere characteristics at different operating currents for various hot/cold cathode lamps.

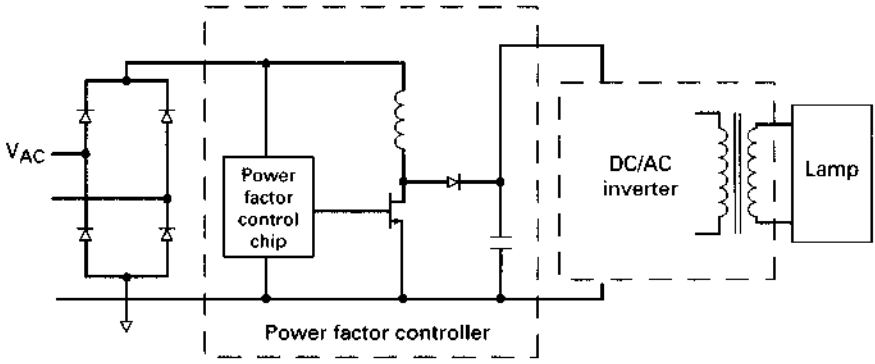


Figure 16.12 Block diagram of a modern fluorescent lamp light source. Output frequency of the DC/AC inverter is set by a series or parallel self-resonant oscillator in the range of 20 to 50 kHz. The ballast is usually a capacitor or the controlled source impedance of a series *LC* resonant circuit.

sive heating of the input power lines and generator windings. For the many fluorescent lamps in a large office building or smaller generators as on naval vessels, this last could be a problem.

Power factor correction (PFC) solves this problem by eliminating the large input capacitor and forcing the input line current to be sinusoidal and in phase with the input line voltage. It also yields a fairly well-regulated DC output voltage about 10 to 15 V higher than the peak of the AC input at high line. It is universally done with a continuous-conduction-mode boost converter controlled by a PFC chip, frequently the inexpensive Motorola 34262. That chip is essentially a 34261 with some improvements, as described in Chap. 15. Lamp ballast manufacturers are currently required to have power factor correction meeting IEC555-2 specification which limits the harmonic line content of the input power line. Electronic ballasts are also required to meet EMI/RFI limits set by the FCC (CFR 47, part 18).

16.5 DC/AC Inverter—General Characteristics

The DC/AC inverter topologies usually used in electronic ballasts are the push-pull for 120-V AC line input and the half-bridge for 220-V line input. Unlike switching power supplies which use fixed-frequency, chip-driven, square wave amplifiers for their DC/AC inverters, ballast inverters are series or parallel, self-resonant LC oscillators. There are a number of reasons for this.

Probably most important is that the lamp light power efficacy is highest with sinusoidal current drive. Thus there is an obvious cost

and space savings in initially generating the alternating current as a sine rather than a square wave and adding filter components to remove higher harmonics. Although using a chip is probably the simplest and most direct way of generating alternating current from the rectified AC line, the cost would be prohibitive. The *LC* oscillators actually used are not too constant in frequency, but that is no problem as long as the frequency remains above 20 kHz. Beyond that, there is little further gain in light power efficacy from the lamp.

Also, with sinusoidal voltages or currents, there are very significant advantages to reduced switching losses at turn-on and turnoff and a higher permissible "off" voltage stress with a sinusoidal base drive voltage. The negative half cycle of the base drive automatically provides the reverse bias, which permits the transistor to sustain the V_{cer} voltage rating rather than the lower V_{ceo} rating.

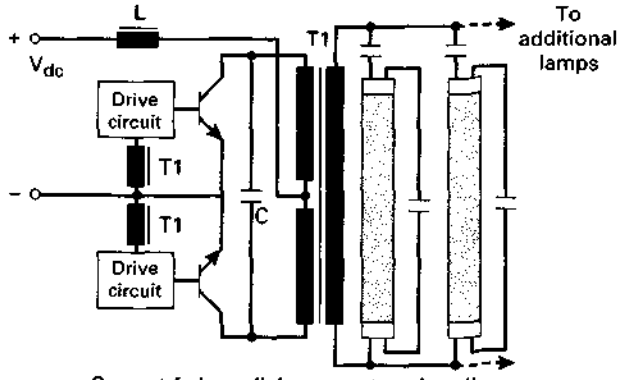
Higher permissible off collector voltages with sinusoidal base drive can be appreciated from Fig. 16.14*a* for one of the above topologies. There it can be seen that throughout the collector off time, the negative half sinusoid of the base drive winding, aided by the negative bias on the input capacitor, automatically forces a negative bias on the off base. This permits the usual high-voltage transistor to safely sustain its V_{cev} rather than its V_{ceo} rating. The V_{ceo} rating, you will recall, is the maximum off voltage that a bipolar transistor can withstand if it has a high impedance or open circuit at its base in the off condition. On the other hand, the V_{cev} rating for most high-voltage bipolar transistors is generally 100 to 300 V higher than the V_{ceo} rating and pertains if, in the off state, the base has a 2- to 5-V negative bias. Such a negative bias is automatically obtained from the negative half cycle of the turn-on voltage from the feedback winding (Fig. 16.14*a*) of the oscillation transformer. This negative bias on the off transistor is also automatically obtained in the alternative base drive scheme in Fig. 16.14*a*.

It will be shown in later waveforms (Figs. 16.14*b* and 16.22*b*) that two of the most frequently used topologies achieve transistor switching at the zero crossing of the collector voltage waveform and thus minimize transistor switching loss.

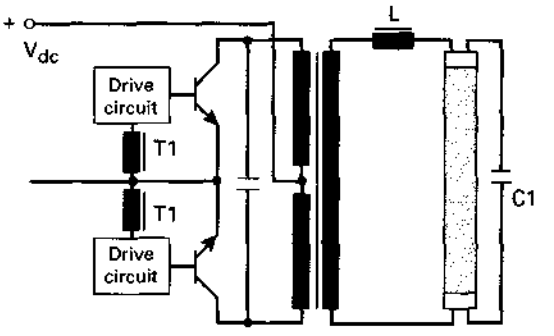
16.6 DC/AC Inverter Topologies

Figure 16.13 (Ref. 6) shows the four most commonly used topologies for electronic ballasts. They comprise two versions—current-fed and voltage-fed variations of the push-pull for 120-V AC line input, and current- and voltage-fed for 220-V AC line input.

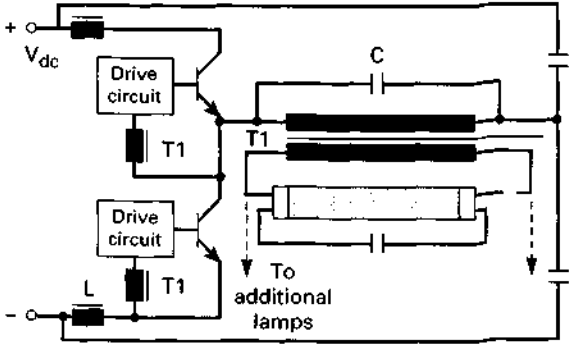
For both push-pull and half bridge, the current-fed version requires either one or two extra inductors and subjects the off transistors to higher off voltage stress than the voltage-fed types.



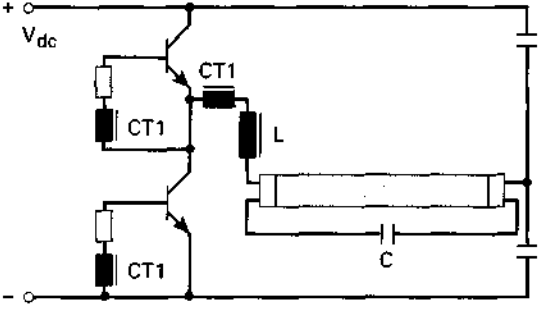
Current-fed parallel resonant push-pull ballast with isolated output.



Voltage-fed push-pull ballast with isolated output



Current-fed parallel resonant half bridge ballast with isolated output.



Voltage-fed half bridge ballast.

Figure 16.13 Usual topologies for DC/AC inverters in AC power sources for fluorescent lamps. Half bridges are used for 220-V AC input, push-pulls for 120-V AC input. (*Courtesy Ferroxcube / Philips, Inc.*)

But reliable designs are harder to achieve with voltage-fed types because they suffer from large start-up voltage and current transients whose magnitude depends on the resonant circuit Q. Start-up current transients in voltage-fed circuits may be 5 to 10 times the operating currents.[7] Further, voltage-fed circuits have greater difficulty coping with open- or short-circuited lamp loads while current-fed schemes can operate indefinitely with such loads.[6,7]

Further, on the plus side, current-fed circuits yield cleaner sinusoidal waveforms and hence offer higher lamp light power efficacy.[6] And importantly, the current-fed circuit can drive a number of lamps in parallel, compared to only a single lamp drive capability with the voltage-fed scheme.

With the increasing availability of ever higher-voltage and lower-cost bipolar transistors, the higher-voltage stress with current-fed topologies is not a drawback to their use. Their major drawback remains the few added components and hence somewhat higher cost.

16.6.1 Current-fed push-pull topology

This is shown in Fig. 16.14*a* and *b*. It was first described in detail in a classic article by R. J. Haver.[8] The Haver version did not use an isolation transformer to drive the lamps but rather a center-tapped choke connected directly across from collector to collector with the lamps across the choke. Currently, most usages require the DC isolation transformer as lamps are frequently changed while power is on.

In Fig. 16.14*a*, the power transformer center tap is fed, not directly from the low-impedance DC output of the boost regulator, but via a *constant-current* inductor L_{cf}. The circuit is an oscillator with positive feedback from a pair of windings (N_{fb}) on the power transformer to drive the bases.

Capacitor $C1$ in shunt with the magnetizing inductance L_m of both halves of the primary comprises the parallel resonant tank circuit having a resonant frequency of $f_r = 1/(2\pi \sqrt{L_m C1})$. This is the resonant frequency for a short time before the lamps light, as ballast capacitors $C4$ do not reflect back into the primary until the lamps are lighted. When the lamps light, the ballast capacitors draw current, and some, not easily calculable, capacitance reflects back across to the primary in shunt with C_4 and lowers the operating frequency.

Two alternative base drive schemes are shown. The upper one uses an RC combination ($R2C2$, $R3C3$) at each base in series with the drive winding. The negative bias at the base side of the capacitors helps provide more rapid turnoff and reduced storage time for the off-turning transistor. It also adds to the negative bias which helps ensure it is the higher off voltage specification V_{cev} rather than the V_{ceo} rating which pertains.

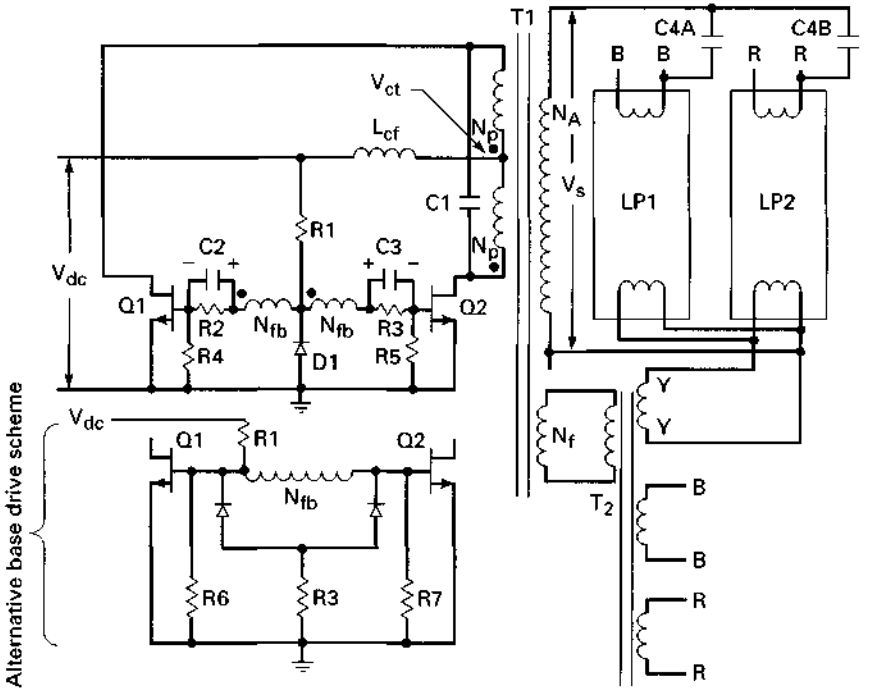


Figure 16.14*a* Current-fed parallel resonant DC/AC inverter. *C*1 plus the reflected ballast capacitors *C*4*A* and *C*4*B* form the parallel resonant tank together with the *T*1 magnetizing inductance in shunt. L_{CF} is a constant-current inductor.

The secondary is shown driving two rapid-start lamps in parallel. Filament power for the lamps is derived from the secondaries of a second transformer *T*2 whose primary is an additional winding on *T*1. Though *T*2 is an added-cost item, it may be preferable to the alternative of having the filament windings themselves as additional secondaries on *T*1. This latter would make for a bulkier, less efficient and harder/more expensive transformer to wind *T*1.

Drawing a small current from V_{dc}, *R*1 supplies turn-on current to start the oscillation. The always-present inequality in beta between transistors Q_1 and Q_2 ensures that the higher-gain device turns on first, and thereafter the base drive for the continuing oscillation comes from positive feedback windings N_{fb} via *D*1.

Collector voltage and current waveforms are shown in Fig. 16.14*b*2-*b*5. Turn-on and turnoff are achieved at the zero voltage points on the collector voltage waveform, minimizing AC switching losses.

The peak secondary voltage across N_s will be set to the specified nominal lamp striking voltages V_{ns} at 50°F for the specific lamp type.

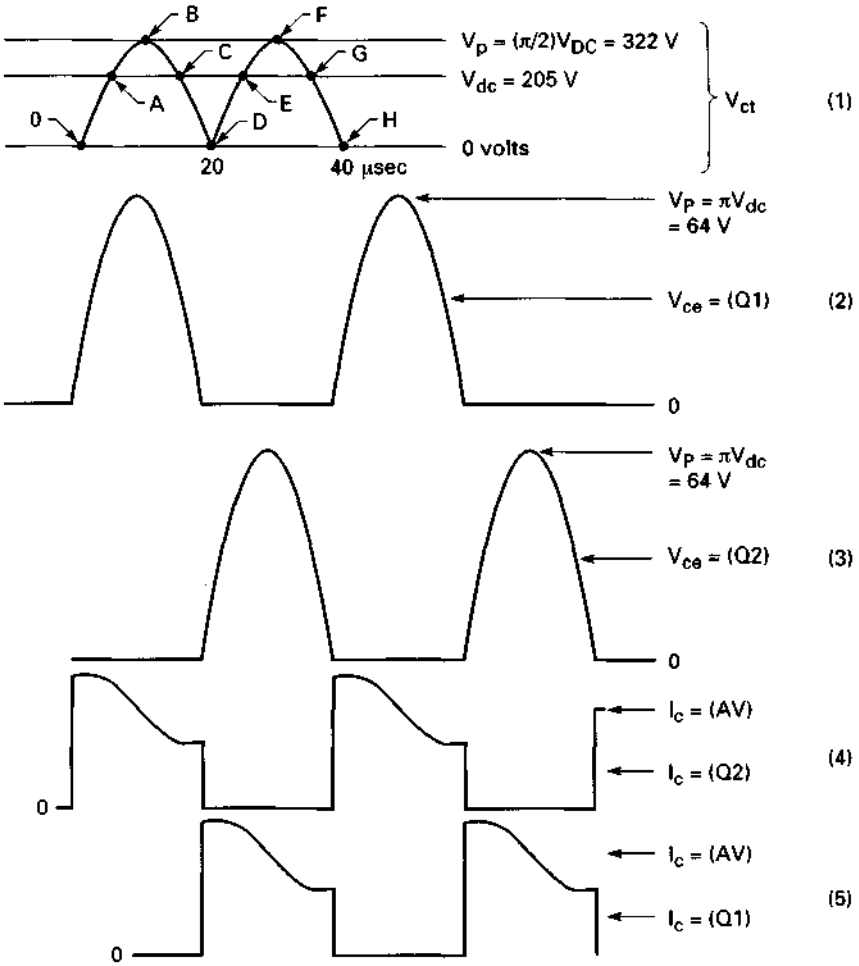


Figure 16.14*b* Critical waveforms for the current-fed parallel resonant DC/AC inverter of Fig. 16.14*a*.

Capacitor $C4$ is the ballast capacitor which limits the lamp rms operating current (I_{lrms}) to its specified value at its specified rms operating voltage (V_{lrms}) to yield the specified lamp power $P = (I_{\text{lrms}})(V_{\text{lrms}})$. Thus $C4$ will be selected from

$$I_{\text{lrms}} = I_{C4\,\text{rms}} = \frac{0.707V_{ns} - V_{\text{lrms}}}{X_{c4}}$$

$$= (0.707V_{ns} - V_{\text{lrms}})\,(2\pi f_r C4) \tag{16.4}$$

16.6.2 Voltages and currents in current-fed push-pull topology

Figure 16.14*b*1 shows the waveform at the center tap of power transformer $T1$ in Fig. 16.14*a*. It comes about because of the current feeding inductor L_{cf} and is a full wave rectified sine wave going down to 0 V at the zero crossings. Since L_{cf} has negligible DC resistance and hence cannot support a DC voltage, the average voltage at the output end of L_{cf} must equal that at its input end, or V_{dc}. And since the average of a full wave rectified sine wave of amplitude V_p is $V_{av} = V_{dc} = (2/\pi)V_p$, the peak voltage at the center tap is $V_p = (\pi/2)V_{dc}$. And since the peak center tap voltage at the center of each on time is $(\pi/2)V_{dc}$, the opposite transistor, which is off, is subjected to a peak V_{ce} voltage of πV_{dc} (Fig. 16.14*b*2, *b*3).

Now assume a nominal AC input of 120 V rms with ±15 percent tolerance yielding a peak of $1.41 \times 1.15 \times 120 = 195$ V. Recall that a PFC circuit generates a fairly well-regulated DC output voltage about 20 V above the peak AC input (Chap. 15). Then V_{dc} will be about $195 + 20 = 205$ V, and from the above, the transistors must safely sustain a peak off voltage of π (205), or 644, V. There are currently a number of transistors with adequate voltage, current, and f_t ratings to meet this requirement (MJE18002, MJE18004, with V_{ceo} of 1000 V, f_t of 12 mHz, and minimum betas of 14). Transistors with a lower f_t of 4 mHz are most probably also suitable, as the negative bias immediately after turnoff serves to minimize transistor storage time.

Figure 16.14*b*2–*b*5 shows collector currents rising and falling at the zero voltage points of the collector voltage waveforms. This minimizes AC switching losses. And since the half sine waves of voltage across each half primary are equal in volt-second area (Fig. 16.14*b*1) and with negligible storage time, there is no possibility of *flux imbalance* (Sec. 2.2.5) or momentary simultaneous condition.

Collector currents in each half cycle are shown in Figs. 16.14*b*4, *b*5. The sinusoidal "wiggle" at the top of the square pulse of current is characteristic and will be discussed below. The average value of the current at the center of the sinusoidal wiggle ($I_{c\,av}$) is calculated from the lamp power. Assuming two lamps of specified power P_1, a DC/AC converter efficiency of E, and input voltage V_{dc}, the collector current is

$$I_{c\,av} = \frac{2P_1}{EV_{dc}} \tag{16.5}$$

Thus for two 40-W lamps, a converter efficiency of 90 percent, and an input voltage V_{dc} of 205 V from the power factor corrector circuit, $I_{c\,av} = 2 \times 40/(0.9 \times 205) = 434$ mA.

16.6.3 Magnitude of "current feed" inductor in current-fed topology

This inductor (L_{CF} in Fig. 16.14*a*) is calculable from Fig. 16.14*b*1. There it is seen that the output end of L_{CF} swings in half sinusoidal fashion around V_{dc}. Voltage across L_{CF} at any instant is $V_1 = L_{CF}\, dI/dt$. Hence dI, the change in current in L_{CF} between any two instants t_1 and t_2, is

$$dI = \frac{1}{L_{CF}} \int_{t_1}^{t_2} V_1\, dt \tag{16.6}$$

Now it can be seen in Fig. 16.14*b*1 that from times A to C the voltage at the output end of L_{CF} is above V_{dc}. Thus the volt-second area across the inductor in Eq. 16.6 is positive, signifying that the inductor current increases from times A to C. And from times C through E, the volt-second area across the inductor is negative, indicating the current in it decreases in that time interval.

Hence the inductor L_{CF} is chosen so that the current change dI in Eq. 16.6 is an arbitrarily chosen small fraction of the current in Eq. 16.5. The inductor driving the DC/AC inverter can then be considered a *constant-current* source. Assume then that dI of Eq. 16.6 is ±20 percent of the current $I_{c\ av}$ in Eq. 16.5. Then from the above value of 434 mA for $I_{c\ av}$, $di = 0.4 \times 434 = 174$ mA. Then from Eq. 16.6, $L_{CF} = \int_A^C V_1\, dt/0.174$ H. Now in Fig. 16.14*b*1, $\int_A^C V_1\, dt$ is the area (in volt-seconds) of the region lying between V_{ct} and the V_{dc} line.

By "eyeball" integration, that area is about 806×10^{-6} V · s. Then from Eq. 16.6, $L_{CF} = 806 \times 10^{-6}/0.174 = 4.6$, say, 4.0 mH. It will be wound on either a powered iron or gapped ferrite core, so it does not saturate at the maximum direct current it will draw. Although the normal direct current it draws has been calculated above as 434 mA, it should be designed for about twice that to allow for turn-on transients.

Examination of waveforms in Fig. 16.14*b*1, *b*2, and *b*3 offers a semiquantitative explanation for the characteristic sinusoidal wiggle at the top of the current waveforms in Fig. 16.14*b*4, *b*5. The explanation follows from the consideration that the voltage across L_{CF} is $V_1 = L_{CF}\, di/dt$. At points A and C in Fig. 16.14*b*1, those voltages are zero, hence di/dt is zero at those points, and they correspond to points I and K in Fig. 16.14*b*4, where di/dt is also zero. At point B in Fig. 16.14*b*1, V_1 is a maximum and hence di/dt is also a maximum. Point B in Fig. 16.14*b*1 thus corresponds to point J in Fig. 16.14*b*4, where di/dt is also a maximum.

16.6.4 Specific core selection for current feed inductor

The 4.0-mH inductor L_{CF} can be designed with MPP cores (Sec. 4.3.3.2), KoolMu cores[9] which are less expensive and more modern versions of MPPs (Fig. 16.15), gapped ferrite cores (Sec. 4.3.3.1), or inexpensive but very lossy powdered iron (Micrometal)[10] cores (Fig. 16.17). The choice will be made on the basis of cost, core losses, and size. Table 16.1 serves as a guide to the choice.

The high price of MPP cores obviously rules them out. But the core loss comparison at 1000 G in Table 16.1 is deceptive, as it will be seen that for the large number of turns for 4.0 mH, the peak flux density is so low that core losses even for the inexpensive Micrometals core are not significant. The following calculations will demonstrate the final core selection.

The required number of turns (N_r) will be calculated from A_1 (millihenries per 1000 turns) as

$$N_r = 1000\sqrt{\frac{L}{A_1}} = 1000\sqrt{\frac{4}{A_1}} \tag{16.7}$$

The peak flux density (B_m can be calculated from Faraday's law:

$$V_1 = NA_e\, dB/dt \times 10^{-8}$$

or

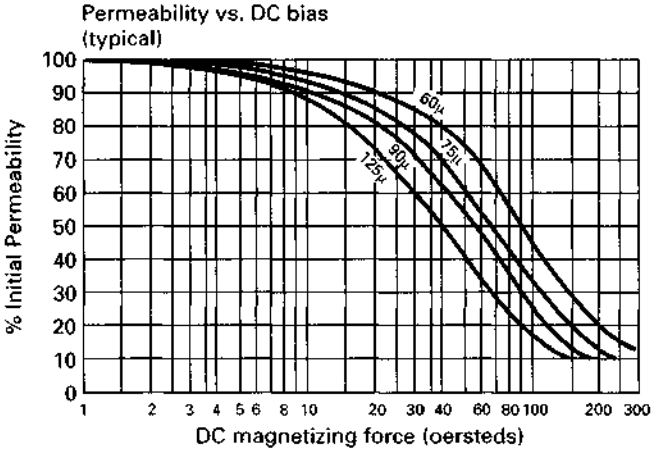
$$B_m = \int_A^C V_1\, dt/NA_e \times 10^{+8}$$

where $\int_A^C V_1\, dt$ is the area in volt-seconds between V_{ct} and V_{dc} between points A and C in Fig. 16.14*b*1. That area is 806×10^{-6} V · s, as estimated above. Then

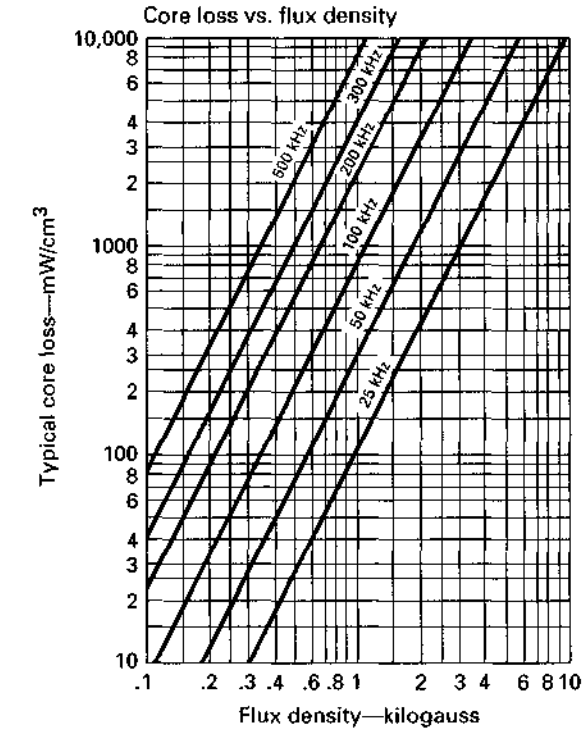
$$B_m = \frac{806 \times 10^{+2}}{NA_e} \tag{16.8}$$

Core losses will be calculated from the manufacturer's curves of losses versus B_m and frequency. Frequency will be taken at 50 kHz, as Fig. 16.14*b*1 shows that for 25-kHz oscillation frequency, inductor frequency is 50 kHz.

Tentative core selections will be for the three types—KoolMu, Micrometals, and gapped ferrite. The initial selection will be from Eq. 16.8 seeking an A_1 which yields minimized turns and does not satu-



(a)



(b)

Figure 16.15 Characteristics of KoolMu, a powdered magnetic core less lossy than powdered iron. (*Courtesy Magnetics, Inc.*)

TABLE 16.1 Cost/Losses Comparison of Various Core Materials

Core type	Cost (500 quantity), $	Core loss, mW/cm^3, at 1000 G, 50 kHz
MPP	14.00	180
KoolMu	4.20	300
Gapped ferrite (3C85)	2.20 (two halves, 3019 pot)	30
Micrometal	0.34 (26 material)	2000

rate the core, as determined from the manufacturer's curves of percent of falloff (%Fall) in inductance versus oersteds (Fig. 16.16*a*). It will be seen that cores with higher values of A_1 (higher permeability) which minimize the number of turns saturate at lower value of DC oersteds. Maximum oersteds will be calculated from

$$H_m = \frac{0.4\,\pi\,NI}{l_m} \tag{16.9}$$

where l_m is the magnetic path length in centimeters and I is twice the normal maximum current of 434 mA as calculated in Sec. 16.6.2, because of the possibility of turn-on transients twice the normal maximum current (Sec. 16.6.3).

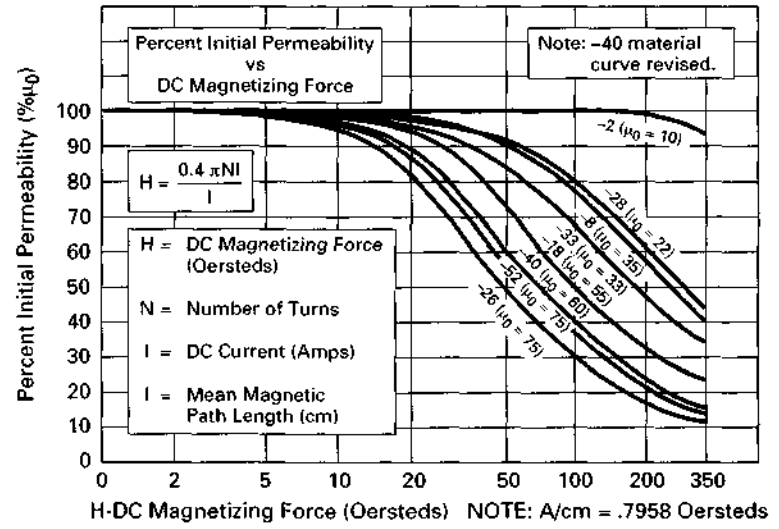
Then for such tentatively selected cores, one will be chosen which yields minimum core losses as calculated from N and B_m of Eq. (16.8) and the manufacturer's curves of core loss versus B_m and frequency (Fig. 16.15*b*). This procedure will first be done for KoolMu cores. See Table 16.2.

The inductor can be done with any of the above four cores. Total core loss is obviously insignificant. The percentage of falloff in inductance is negligible and can be recovered by increasing the number of turns by the square root of the falloff percentage. Core 77439 would be the best choice as it has the fewest turns.

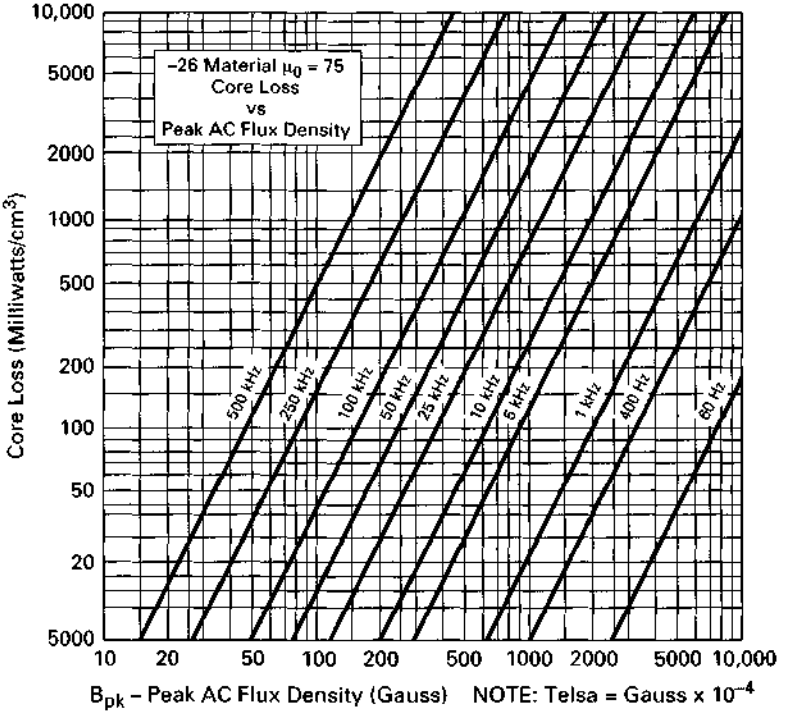
Repeating this procedure for Micrometal toroidal cores where loss versus B_m and frequency is taken from Fig. 16.16*b* and %Fall in inductance is taken from Fig. 16.16*a*, gives the results shown in Table 16.3.

Table 16.3 shows that total core losses with the inexpensive Micrometal cores are five to seven times higher than those with KoolMu cores but yet not prohibitively high. The required number of turns does not differ greatly. The Micrometal 157-26 and 175-26, having smaller magnetic path length and more turns than the first two Micrometal types, have higher peak oersteds and consequently a higher "swing" or falloff in inductance.

The best choice among Micrometal cores is probably the 250-26 because of its lower inductance falloff and despite the higher core dissipation.



(a)



(b)

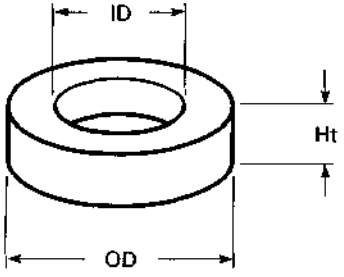
Figure 16.16 Characteristics of Micrometal powdered iron cores, a more lossy but far less expensive core than KoolMu or MPP material.

TABLE 16.2 Possible KoolMu Cores for Current Feed Inductor

KoolMu type	A_l, mH/1KT	N turns	A_c, cm^2	B_m, G	l_m, cm	H_m, Oe	%Fall	Loss, MW/cm^3	Volume, cm^3	Loss, total mW
77110	75	231	1.44	121	14.3	17.6	8	4	20.6	84
77214	94	206	1.44	136	14.3	15.7	11	5	20.6	103
77094	107	193	1.34	159	11.6	18.5	13	8	15.6	125
77439	135	172	1.99	117	10.74	17.4	9	4	21.3	85

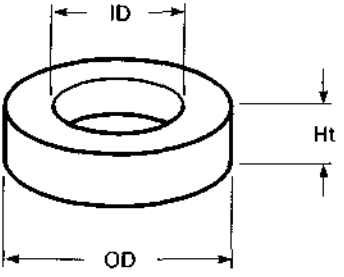
TABLE 16.3 Possible Micrometal Cores for Current Feed Inductor

Micrometal type	A_l, mH/1KT	N turns	A_c, cm^2	B_m, G	l_m, cm	H_m, Oe	%Fall	Loss, MW/cm^3	Volume, cm^3	Loss, total mW
225-26B	160	158	2.59	98	14	12.3	10	17	38	646
250-26	242	129	3.84	81	15	9.38	6	15	57	855
157-26	100	200	1.06	190	10	21.8	20	60	11	660
175-26	105	195	1.34	154	11	19.3	18.0	40	15	600



KoolMu 77439 Core:
OD = 1.84 in
ID = 0.94 in
Ht = 0.71 in

(a)



Micrometal T250-26 Core:
OD = 2.5 in
ID = 1.25 in
Ht = 1.0 in

(b)

Figure 16.17 Geometry of candidates for core for current feed inductor L_{CF} of Fig. 16.14*a*.

The final choice, then, is based on a cost versus engineering performance comparison. The KoolMu 77439 (Fig. 16.17*a*) is smaller, has lower dissipation, but costs more. Its outside diameter (OD) is 1.84 in, it dissipates 85 mW, and it costs $4.20 in 500 quantity. The Micrometal 250-26 (Fig. 16.17*b*) has an OD of 2.5 in, dissipates 855 mW, but costs 34 cents in 500 quantity.

A gapped-ferrite core might be the best compromise. A core can be selected from an initial educated guess and two to three repeated calculations. Required are curves showing A_l vs ampere-turns for various air gaps and the *cliff point* in ampere-turns at which the DC bias commences. Such curves are shown in Sec. 4.3.3.1 in Fig. 4.2.

The procedure is as follows. Guess at a core and thus its A_l, and calculate the required number of turns for the desired inductance (4 mH) from Eq. 16.7. Calculate the maximum number of ampere-turns at the anticipated maximum direct current (1 A). If these ampere-turns exceed the saturation cliff point, the core is too small or its gap is not sufficiently large. Try again for the same core with larger gap (smaller A_l). Now proceed until a core or gap or A_l is found for which $(NI)_{max}$ is less than the saturation cliff point.

For this selected core and A_e, calculate the peak flux density B_m from Eq. 16.8. Then from the manufacturer's curve of core loss, mW/cm^3, versus B_m (Fig. 16.18), read that loss at B_m and the current ripple frequency (50 kHz for 25-kHz switching frequency).

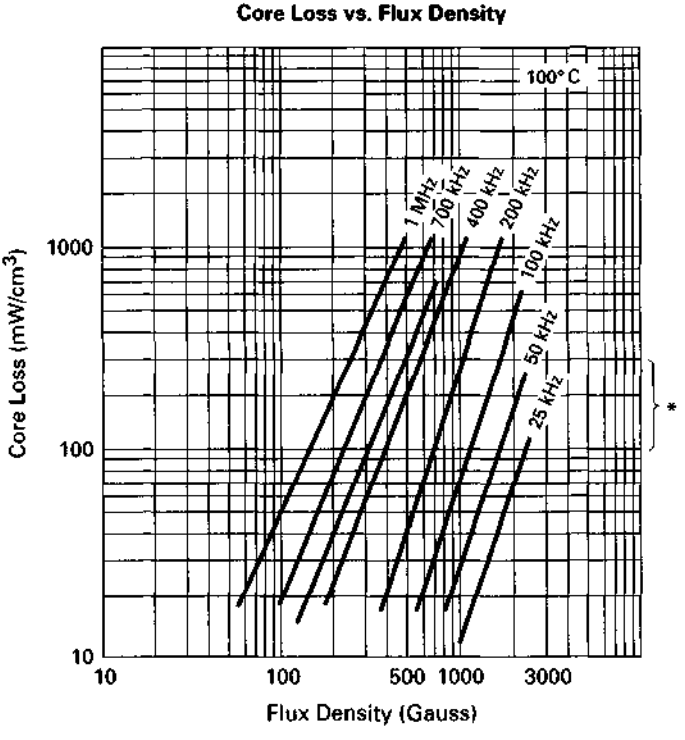


Figure 16.18 Core loss vs. peak flux density and frequency, Ferroxcube/Philips Ferrite 3C85 material. (*Courtesy Ferroxcube/Philips, Inc.*)

TABLE 16.4 Possible Gapped-Ferrite Cores for Current Feed Inductor

Core Philips	A_l, MH/1KT	Gap, mils	N turns	NI_{max}, A · turns	Cliff point, A · turns
2616	170	32	153	153	129
2616	100	64	200	200	200
3019	500	11	89	89	60
3019	210	35	138	138	170

Thus, guessing first at the 2616 pot core and proceeding to the 3019 pot core: yields Table 16.4.

Table 16.4 shows that the 2616 cannot be used because even with a 32-mil gap which requires 200 turns, the maximum ampere-turns at 1 A of current just barely rest at the saturation cliff. Increasing the gap would probably put NI_{max} inside the cliff, but would require significantly more than 153 turns. Most likely, the core bobbin could not then accommodate the required turns of the appropriate wire size.

It is then seen that the next-larger pot core—the 3019—would work with an appropriate gap. For an 11-mil gap, it doesn't make it, as it is subjected to 89 A · turns at 1 A but the saturation cliff is at 60 A · turns. But with a 35-mil gap, it requires 138 turns, and NI_{max} at 1 A is 138 A · turns, which is safely inside the saturation cliff point of 170 A · turns.

The 3019 has an iron area of 1.38 cm^2 and, from Eq. 16.8, has a peak flux density of 423 G. At that flux density, Fig. 16.18 shows its loss is about 5 mW/cm^3. For its 6.19-cm^3 volume, its total loss is 31 mW.

Finally, the cost/performance comparison of the three possible cores is seen in Table 16.5. Obviously, if minimum cost is the major criterion, the Micrometal 250-26 is the best choice despite its 850-mW core loss. If that core loss is not acceptable, the ferrite 3019 pot core is the best choice despite its higher cost.

16.6.5 Coil design for current feed inductor

The rms current (equals direct current) for the coil is the above-calculated 434 mA. At 500 circular mils/rms A, the required wire area is $500 \times 0.432 = 217$ cmils. Number 26 wire of 253-cmil area is ade-

TABLE 16.5 Comparison of Contending Current Feed Inductor Cores

Core type	Cost (500 qty.), $	Total core loss, mW	OD, in	Height, in
KoolMu 77439	4.20	85	1.84	0.71
Mic-Met 250-26	0.34	850	2.5	1.00
Ferrite pot 3019	2.20	31	1.18	0.74

quate. If the 3019 ferrite core is selected, its bobbin width is 0.459 in and height is 0.198 in. For a no. 26 wire diameter of 0.0182 in, the number of turns per width is 0.459/0.0182 = 25. The number of layers per bobbin height is 0.198/0.0182 = 10. Thus the 138 turns could be accommodated within six layers.

If any of the above toroidal cores were selected, the 138 turns could easily be accommodated in three layers.

The skin effect is no problem as the AC amplitude is small, and Table 7.6 shows that the AC-to-DC resistance ratio is unity at 50 kHz.

16.6.6 Ferrite core transformer for current-fed topology

For two 40-W lamps as shown in Fig. 16.14*a*, transformer primary input power at an assumed efficiency of 85 percent is 94 W.

The transformer core will most likely be ferrite, and its size will be selected from Table 7.2*a*, page 289. A toroidal KoolMu core design will be attempted, but it will be shown that its design is less flexible and more lossy. Table 7.2*a* shows the maximum power available in a forward converter from a core at various frequencies and a peak flux density of 1600 G (page 288). For a push-pull topology, the available power is twice that.

Table 7.2*a* shows that at 24 kHz the smallest core which can deliver more than 94 W in a push-pull topology is the *E*21 which is an international standard size and is available from a number of manufacturers. Its maximum available output power in a push-pull configuration is 2 × 69.4 = 138 W at a maximum flux density B_m of 1600 G. At B_m = 2000 G it should be able to deliver (2000/16,000)138 = 172 W. It has an iron area A_e of 1.49 cm^2.

But because of the unique nature of this resonant converter, it is only marginally capable of being usable. It is instructive in a tutorial way to see why this is so. See Fig. 16.14*a* and 16.14*b*1.

The number of primary turns N_p is calculated from Faraday's law—$E = N_p A_e\, dB/dt \times 10^{-8}$—or

$$N_p = \int_0^{\pi} \frac{V\,dt\,(10^{+8})}{A_e\,dB} = \int_0^{\pi} \frac{V\,dt\,(10^{+8})}{2A_e B_m} \tag{16.10}$$

Here $\int_0^{\pi} V\,dt$ is the area of a half period in volt-seconds, and dB is the total flux change in that time (4000 G for a peak flux density B_m of 2000 G). For the peak of a half sinusoid of V_p, the area of the half period is $(2/\pi)(V_p)(T/2)$V · s. Then from Eq. 16.11, for V_p = 322 V as shown in Fig. 16.14*b*1, *b*2, and for a 20-μs half period:

$$N_p = \frac{(2/\pi)(322)(20 \times 10^{+2})}{1.49 \times 4000}$$

$$= 68 \text{ turns}$$

The wire size will be calculated on the basis of 500 cmil/rms. Figure 16.14*b*1 shows that the peak center tap voltage at the center of the on time is 322 V. The transformer is delivering its power at an rms voltage of $0.707 \times 322 = 228$ V. For the above-calculated 94 W of input power, the rms input current is then $94/228 = 0.412$ A. But each half primary carries this rms current only for a half period out of every period. Hence the rms current for each half primary should be $0.412 \times 0.5 = 0.291$ A. And at 500 cmil/rms A, wire of 146-cmil area is required.

For 146 cmils, no. 27 gage wire of 202 cmils and 0.0164-in diameter would be used. The E21 core bobbin has width of 0.734 in and a height of 0.256 in. Its width can accommodate $0.734/0.0164 = 44$ turns, and its height can accommodate $0.256/0.0164 = 15$ layers of no. 27 wire. Thus each half primary of 68 turns will consist of two layers of 34 turns each, and the full primary will occupy only four layers. Assuming $N_s/(2N_p) = 1$ and a no. 27 wire secondary, the primary and secondary together occupy only half the bobbin height. This would easily leave sufficient room for the filaments or a primary for a separate filament transformer, as shown in Fig. 16.14*a*.

Transformer core material could be either Magnetics Inc. type P or Ferroxcube/Philips 3F3, both of which have 75 mW/cm^3 loss at 2000 G and 25 kHz (Fig. 16.18). The E21 size core (Fig. 16.19) at a volume of 11.5 cm^3 will dissipate only $0.075 \times 11.5 = 863$ mW.

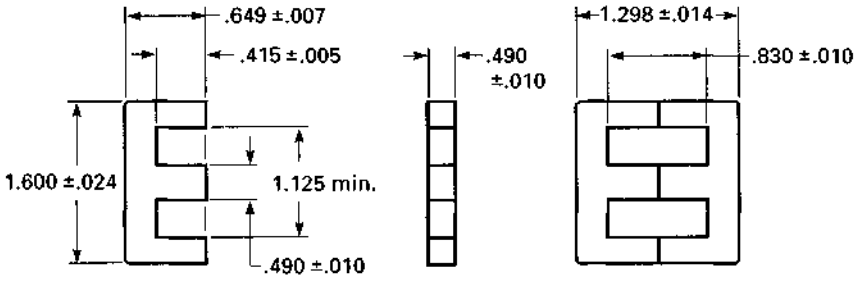
This appears to be a reasonable design, and the E21 core appears usable. But the fallacy is that it was assumed the primary carries current needed only to supply the output load power. The above calculation showed that to supply the load power of 94 W, the current in each half primary was a half sinusoid of 0.412-A peak amplitude with an rms value of 0.291 A.

But it will be seen below that the current actually carried by the primary is considerably greater than this. For the two half primaries in series are in shunt with *C*1 and the reflected ballast capacitors *C*4A, *C*4B (Fig. 16.14*a*) to form the resonant tank circuit. The primary current is then fixed by the amplitude of the circulating resonant tank current which is considerably more than the 0.291 rms A shown above.

The actual primary current which is the circulating tank current is calculated as follows:

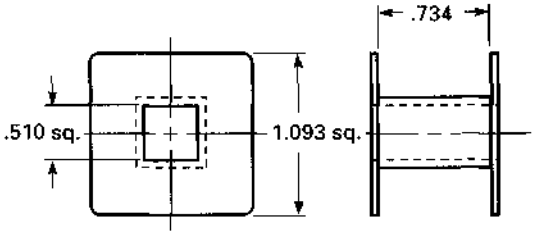
The voltage across this tank (across the two half primaries) is sinusoidal with a peak amplitude of 644 V and an rms voltage of $0.707 \times 644 = 455$ V. Thus each half primary carries current throughout a full period rather than only one-half period. And the rms value of that current is $I_{rms} = V_{rms}/X_1$, where $X_1 = 2fL_t$ and L_t is the inductance of both primary halves in series.

Thus the current carried by the primary depends on L_t, and L_t cannot be made arbitrarily large to keep its current low so as to permit



$A_e = 1.49\ cm^2$

Standard Bobbin



PART NO. E21F1N
Material: Nylon
Max. Operating Temp.: 105°C
Winding Area: .188 in^2.
Mean Length of Turn: 3.08 in.
Flammability: UL94-V2

Figure 16.19 Dimensions of the E21 core and bobbin. This core is a potential candidate for the current-fed parallel resonant push-pull topology of Fig. 16.14*a*. But its magnetizing inductance is the resonant inductor of the resonant tank circuit. Since there are constraints on how small the resonant capacitor is, this limits the magnetizing inductance to a relatively small value for resonance at 25 kHz. With a small magnetizing inductance, the circulating tank current is large, requiring such large wire size as to make it impossible to fit the coil inside the bobbin. Thus, larger cores than might be expected for the given power level are required.

small wire size. For L_t together with its total shunt capacity C_t sets the resonant frequency at $1/(2\pi\sqrt{L_tC_t}) = 25$ kHz.

But C_t is the sum of $C1$ and the reflected ballast capacitors $C4A,4B$. And C_t cannot be made small to permit a large value for L_t and consequent low current in L_t. At initial turn-on, when the lamp impedances are high, $C4A$ and $C4B$ are essentially out of the circuit, and $C1$ alone and L_t set the resonant frequency at its highest value. When the lamps light, the total resonating capacitance increases to $C1 + C4A + C4B$ [assuming $N_s/(2N_p) = 1$] and the frequency drops to its desired

value of 25 kHz. But if one lamp is defective and out of the circuit, only one ballast capacitor reflects into the primary and the operating frequency is somewhere in between. Thus to keep the frequency from changing too much under these three operating conditions, $C1$ should be large—or at least not small compared to $C4A + C4B$.

Now Eqs. 16.1 and 16.3 set restrictions on $C4A$, 4B. In Eq. 16.1, V_{ns} may be any value greater than the highest striking voltage for the given lamp. The higher it is, the greater must be the impedance X_b of the ballast capacitor C_4 to limit lamp current to its specified value, as seen from Eq. 16.1. And if V_{ns} is set high to permit a high value for X_b and hence low value for C_b, this makes it easier to achieve the desired goal of having $C1$ large compared to the reflected ballast capacitors. But for a high V_{ns} (rms secondary voltage V_s), $N_s/(2N_p)$ is high and any ballast capacitors reflect across into the primary as larger capacitors. There is thus no advantage in going to any turns ratio higher than $N_s/(2N_p) = 1$.

Thus for a specified lamp operating voltage V_{op} of 101 V rms, an operating current of 430 mA rms, and $V_s = 455$ V as above, Eq. 16.1 shows

$$X_b = \frac{455 - 101}{0.430} = 823\ \Omega$$

And Eq. 16.3 shows

$$C_b = \frac{1}{2\pi F_r X_b} = \frac{1}{(2\pi)(25 \times 10^{+3})(823)} = 0.0077\ \text{mF}$$

Thus the two ballast capacitors reflect across into the primary as $2 \times 0.0077 = 0.015$ mF. Now $C1$ is chosen equal to the sum of the two reflected ballast capacitors to achieve the goal of having not too great a frequency change under the above operating conditions.

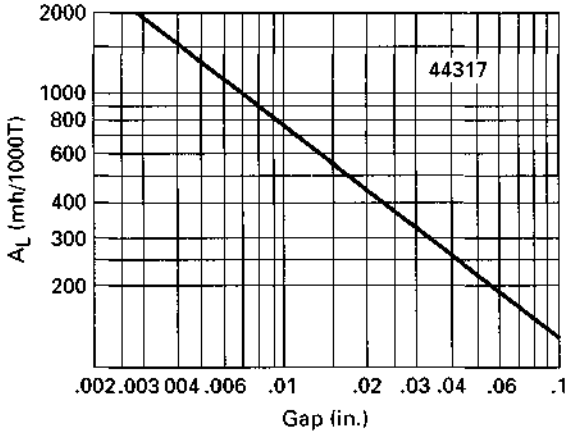
Then for $C_t = C1 + C4A + C4B = 0.03$ mF and a resonant frequency of 25 kHz,

$$L_t = \frac{1}{4\pi^2 F_r^2 C_t} = \frac{1}{4\pi^2 (2.5^2) \times 10^{+8} (0.03 \times 10^{-6})} \qquad (16.11)$$

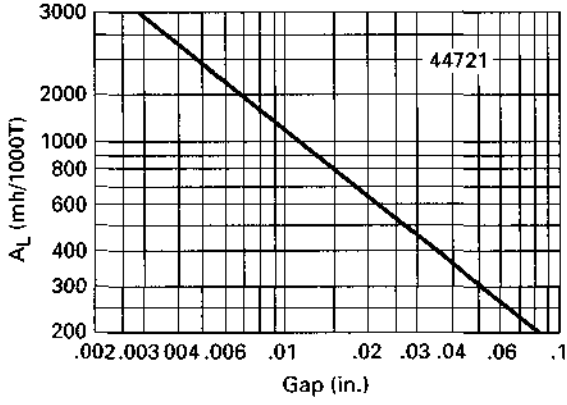
$$= 0.00135, \text{ or } 0.0015\ \text{H}$$

The core would be gapped so that the above-calculated 68 turns per half primary (Eq. 16.10) or 136 turns for both halves would have an inductance of 0.0015 H. This corresponds to an A_1 (inductance in millihenries per 1000 turns) of $(1000/136)^2 \times 1.5 = 81$ mH/1Kt. Figure 16.20 shows this could be obtained with an air gap of 80 mils.

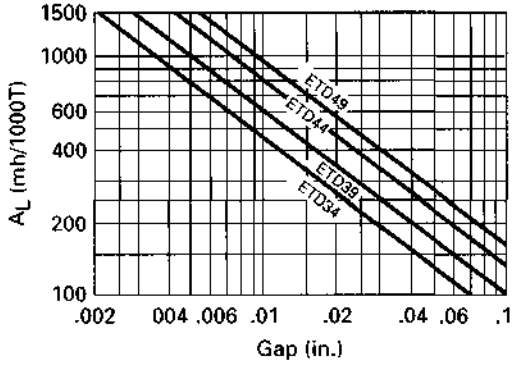
And now with 455 rms V across L_t, it draws a current of



(a) Magnetics Inc. E21 Cores



(b) Magnetics Inc. E625 Core



(c) Magnetics Inc. ETD Cores

Figure 16.20 A1 (Inductance in mH per 1000 turns). (*Courtesy Magnetics, Inc.*)

TABLE 16.6 Possible Coil Design for an E21 Transformer Core

Wire no.	Area, cmils	Diameter, in	Bobbin width, in	Bobbin height, in	Turns/ width	Layers/ height
20	1020	0.0351	0.734	0.256	20	7
21	812	0.0314	0.734	0.256	23	8

$$I_{\text{rms}} = \frac{455}{2\pi(2.5 \times 10^{+4})(0.0015)}$$

$$= 1.93 \text{ rms A}$$

And at 500 cmils/rms A, this would require wire of 966-cmil area. Possible wire choices are shown in Table 16.6.

It is seen in Table 16.6 that the T_1 primary could only barely fit inside the E21 bobbin using no. 20 wire which has somewhat more than the required circular-mil area. Since it requires 136 turns and using no. 20 wire, the bobbin can hold only $20 \times 7 = 140$ turns. With no. 21 wire, each half primary could be handled in three layers (23, 23, and 22 turns). This would leave $0.256 - 6 \times 0.0314 = 0.068$ in for the secondary plus any filament windings. Even though the secondary carries only the operating current for two lamps ($2 \times 0.43 = 0.86$ A) and can use wire smaller than no. 21, there is far from enough space for the secondary, filaments, and insulation between primary and secondary. Thus a larger core must be chosen.

The core selection procedure is as follows:

Step 1 A tentative core selection will be made with an A_e large enough to yield a reasonably small number of half primary turns (N_p) as calculated from Eq. 16.10 for a peak flux density of 2000 G [dB in Eq. 16.10 = 4000 G].

Step 2. Values of resonating inductor L_t and capacitor C_t will be 1.5 mH and 0.03 mF as calculated above for the E21 core. From L_t and N_p, the inductance per 1000 turns A_l will be calculated. From A_l and the manufacturer's curves of A_l versus air gap (Fig. 16.22), the air gap will be read.

Step 3. The rms current in L_t will be the same 1.93 A as calculated above for the E21 core since the rms voltage across L_t is the same 455 V. Wire size then will be either no. 20 or 21 based on a current density of 500 cmils/rms A.

Step 4. From the manufacturer's data sheets, bobbin width and height will be read and calculation will be made to see if the total number of turns in L_t ($2N_p$) can be accommodated in the bobbin with enough space remaining for the secondary and filament windings.

Step 5. If the windings fit within the bobbin, the choice of core will be made on the basis of its core loss (Fig. 16.18) and cost.

TABLE 16.7 Characteristics of Possible Cores for T1 (Fig. 16.14*a*)

	E625	783E608	ETD44
A_e, cm^2	2.34	1.81	1.74
B_m, G	2000	2000	2000
N_p, turns/half primary	43	56	58
L_l, mH	1.5	1.5	1.5
C_t, mF	0.03	0.03	0.03
A_l, mH/1Kt	203	120	111
Gap, mils	72	110	120
Bobbin width, in	0.85	1.024	1.165
Bobbin height, in	0.251	0.270	0.283
Turns/W, no. 20 wire	24	29	33
Reqd. layers, full primary	4	4	4
Layers/height for no. 20 wire	7	7	8
Core loss, mW/cm^3 at 2000 G, 25 kHz	80	80	80
Core volume, cm^3	20.8	17.8	18.0
Total core loss, W	1.66	1.42	1.44

This procedure was followed for three tentatively usable cores with the results shown in Table 16.7. It can be seen that although all three cores could be used, the E625 would be questionable because there would be marginally sufficient vertical height left in the bobbin for the secondary and filament winding after laying down the two half primaries. Thus, the secondary feeds two lamps at 0.43 rms A each. For a total of 0.86 rms A at 500 cmils/rms A, the required area for the secondary wire is 430 cmils. Number 24 wire of 404-cmil area and 0.0227-in diameter would be adequate. Then

	E625	783E608	ETD44
No. secondary turns (= $2N_p$)	86	112	116
Max. turns per layer (no. 24 wire)	37	45	51
Required no. layers—secondary	3	3	3
Total height primary + secondary, in (4 layers no. 20 + 3 layers no. 24)	0.208	0.208	0.208
Remaining height in bobbin, in	0.043	0.062	0.075

The remaining height in bobbin must accommodate primary-to-secondary insulation plus the filament wires (about one turn of no. 22 wire at 0.0281-in diameter). Obviously the E625 core is marginal; of the other two, the ETD44 appears preferable.

16.6.7 Toroidal core transformer for current-fed topology

The circumference of the inside diameter (ID) of a toroid ($\pi \times$ ID) is much greater than the bobbin width of an EE core of roughly equal A_e.

Hence the toroid permits more turns per layer than the EE and will have fewer winding layers (in many cases, only two layers will suffice).

This reduces and almost completely eliminates *proximity-effect* losses (Sec. 7.5.6). It also results in a more reliable design since turns having a high voltage between them are spread farther apart and the possibility of arcing is far less. Further, the entire bobbin width in an EE core must not be utilized if VDE (European safety specifications) must be observed. This means more winding layers and a larger core. In some cases, VDE specifications permit using the entire bobbin width if triple-insulated wire is used, but this may still require a larger core.

But a toroidal core (KoolMu or MPP) is more expensive and more lossy and has less design flexibility. This decreased design flexibility can be seen from the following.

The number of turns per half primary N_p is calculated from Eq. 16.10 once a tentative core selection has been made, A_e has been established, and a B_m has been chosen which yields acceptable core loss (Fig. 16.15*a*). But the number of turns on the full primary $2N_p$ must yield the desired inductance L_t calculated from Eq. 16.7. Thus with L_t and $2N_p$ fixed, A_1, the inductance per 1000 turns, is fixed. But MPP, KoolMu, or even Micrometal toroidal cores come in discrete values for A_1, which are proportional to the core permeability, and core permeabilities are available in only five to six discrete values.

This is unlike the situation with EE cores where the number of turns per half primary N_p can be set to yield a desired peak flux density [B_m from Eq. 16.10]. Yet with that fixed value of $2N_p$ there still remains the liberty to have any desired value of A_1 by setting the air gap (Fig. 16.22).

This problem—that A_1 with the above toroidal cores is available in only five to six discrete values for any given A_e—can be solved by using gapped toroids. KoolMu and Micrometal cores are available with customized gaps to yield any desired A_1. But if a toroid is to be gapped, it has no significant advantage over a gapped ferrite EE core which is less expensive. The sole reason for considering a gapped toroid is that the coil winding length is longer than that with an EE core. This offers fewer coil layers and makes it easier to meet VDE safety specifications.

16.7 Voltage-Fed Push-Pull Topology[6,7,8,11]

This topology is shown in greater detail in Fig. 16.21. The $T1$ center tap is fed directly from the low output impedance of the rectified input line voltage or the output of the power factor corrector building block following it, with no intervening inductor as in the current-fed topology.

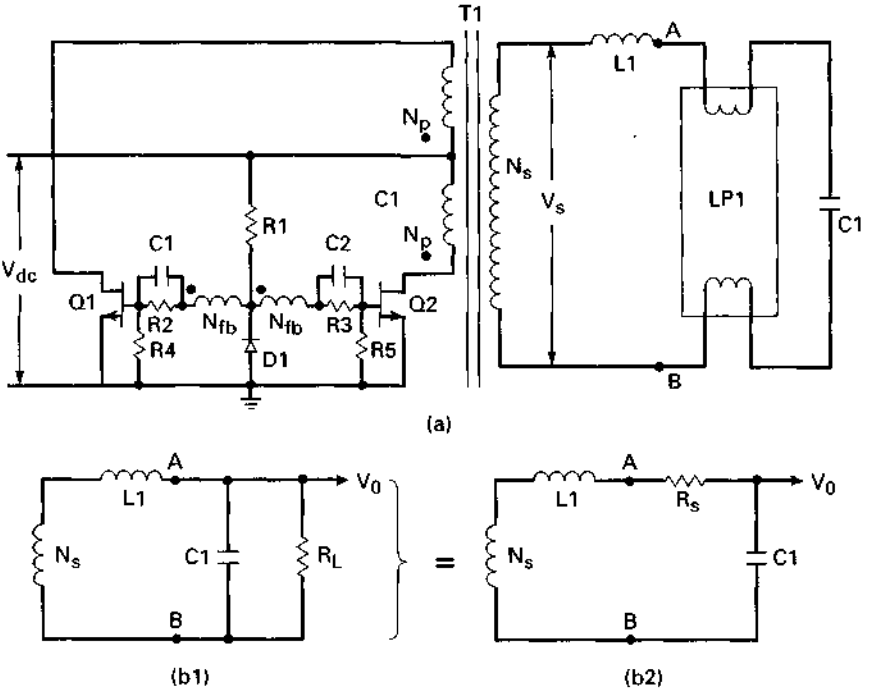


Figure 16.21 Voltage-fed push-pull topology. With no inductor between DC input and the $T1$ centertap, voltage stress on the off transistor is 2 V_{dc} rather than πV_{dc}, as is the case for the current-fed push-pull topology.

The circuit is also a resonant oscillator rather than a driven inverter. Windings N_{FA} and N_{FB} on transformer $T1$ provide the positive feedback to keep the circuit oscillating. Resistor $R1$ draws a small current from V_{dc} to start the oscillation. After the start, the feedback windings supply current through the base emitter of the transistors and D1 to keep them on.

The $T1$ secondary feeds the series resonant combination of $L1$ and $C1$ in series with the lamp filaments. The lamp resistance is in shunt with $C1$. Voltages at the collectors are square waves moving between V_{cesat} (about 1 V) and $2V_{dc}$. With square waves at the primary, voltage at the secondary is also a square wave. But with the secondary a resonant LC circuit, secondary current is sinusoidal and hence so is the primary current. When the secondary sine wave current reverses, so does the current in the feedback windings reverse and the on transistor turns off. This circuit thus also achieves turnoff and turn-on at the zero-current points of collector current and consequently has negligible switching losses.

With no inductor at the center tap, when either transistor is on, the

voltage across its half primary is V_{dc}, as is the voltage across the other half primary. Thus the off transistor is subjected to $2V_{dc}$ rather than πV_{DC} as was the case for current-fed circuit of Fig. 16.14*a*. For a V_{dc} of 205 V as assumed for the current-fed circuit this might mean a maximum voltage rating for the transistor of somewhat more than 410 V rather than 205 or 644 V for the current-fed circuit. The reduced cost of the lower-voltage transistor and saving the input inductor cost might seem a decisive advantage for this topology. But high transient currents at turn-on are a significant drawback to the voltage-fed topology. The higher required transistor current ratings outweigh the lower voltage ratings. This can be seen as follows:

In Fig. 16.21*a*, the equivalent circuit beyond points *AB* is shown in Fig. 16.21*b*1 as the impedance of the lamp *L*1 filaments can be ignored. In Fig. 16.21*b*1, R_L, the effective lamp impedance, is very high before the lamp has lighted and falls to a low value after it has lighted. The circuit behavior can better be understood by converting the parallel combination of $R_L C1$ in Fig. 16.21*b*1 to its equivalent series $R_s C_s$ circuit of Fig. 16.21*b*2. Thus in Fig. 16.21*b*1, the impedance across points *AB* is

$$Z_{AB} = \frac{R_L X_{C1}}{R_L + X_{C1}} = \frac{R_L[1/(j\omega C1)]}{R_L + 1/(j\omega C1)}$$

$$= \frac{R_L}{1 + j\omega R_L C1}$$

And multiplying the numerator and denominator by $1 - j\omega R_L C1$ yields

$$Z_{AB} = \frac{R_L}{(1 + \omega^2 R_L{}^2 C1^2)} - \frac{j\omega R_L{}^2 C1}{1 + \omega^2 R_L{}^2 C_1{}^2}$$

Now set $\omega R_L C1 = Q$. Then

$$Z_{AB} = \frac{R_L}{1 + Q^2} = \frac{j\omega R_L{}^2 C1}{1 + Q^2}$$

For $Q \gg 1$,

$$Z_{AB} = \frac{R_L}{Q^2} + \frac{1}{j\omega C1} \tag{16.12}$$

Equation 16.12 and Fig. 16.21*b*2 show that the transformer secondary N_s drives a series *LC* circuit which resonates at a frequency of $1/(2\pi\sqrt{L1C1})$ and has an equivalent series resistance R_L/Q^2. And for

$Q \gg 1$, that series resistance is very small. Now in a series resonant circuit at resonance, with input voltage V_{in}, current is $V_{in}/(R_L/Q^2)$.

Now at turn-on, before the lamp lights, its resistance R_L is high, Q ($= \omega R_L C1$) is high, and the equivalent series resistance ($= R_L/Q^2$) is very low. This results in high turn-on currents which may be 5 to 10 times higher than operating currents.[7] And with such high turn-on currents, the normal base drive currents may be insufficient to keep the collector low and in saturation and transistor failure may occur. Further, there is excessive voltage across the lamp before it lights, and lifetime may be reduced with frequent turn-ons.

Despite its drawbacks, some ballast manufacturers have used the voltage-fed circuit because of its lower required voltage rating on the transistor. But with the current availability of inexpensive high-voltage transistors, the current-fed scheme of Fig. 16.14*a* is preferable.

16.8 Current-Fed Parallel Resonant Half-Bridge Topology[7]

This topology is shown in Fig. 16.22. It is used when the AC input is 230 V and a power factor correction building block (Chap. 15) boosts rectified output to a voltage higher than the highest peak of the rectified 230 V AC.

For a ±15 percent tolerance on the AC input, that peak rectified voltage is $1.15 \times 1.41 \times 230 = 373$ V DC. The usual power factor correction circuit boosts that to a slightly higher 400 V, and the resonant half bridge is needed to cope with such a high input which is more than a current-fed push-pull can handle.

For an AC input of 120 V, for ±15 percent input tolerance, the peak rectified input is $1.15 \times 1.41 \times 120 = 195$ V, and the usual power factor correction building block boosts that to about 205 V, (Sec. 16.6.2). It was seen in that section that the off transistor in current-fed resonant push-pull topology is subjected to $\pi(V_{dc}) = \pi(205) = 644$ V. For the 230 V AC, a current-fed push-pull circuit would subject the off transistor to $\pi \times 400 = 1257$ V DC which would require too expensive a transistor.

It will be seen below that the current-fed parallel resonant half-bridge topology subjects the off transistor to only $(\pi/2)V_{dc} = (\pi/2)400 = 628$ V. There are numerous inexpensive candidates for such a transistor.

It also is self-oscillatory with windings $T2$A, $T2$B on $T2$ providing the positive feedback. Here $T1$ is the main power transformer, and its magnetizing inductance in shunt with C_r and the reflected ballast capacitor C_b forms the parallel resonant tank circuit. Inductors $L1$ and $L2$ are the constant-current drive elements for the resonant tank. The

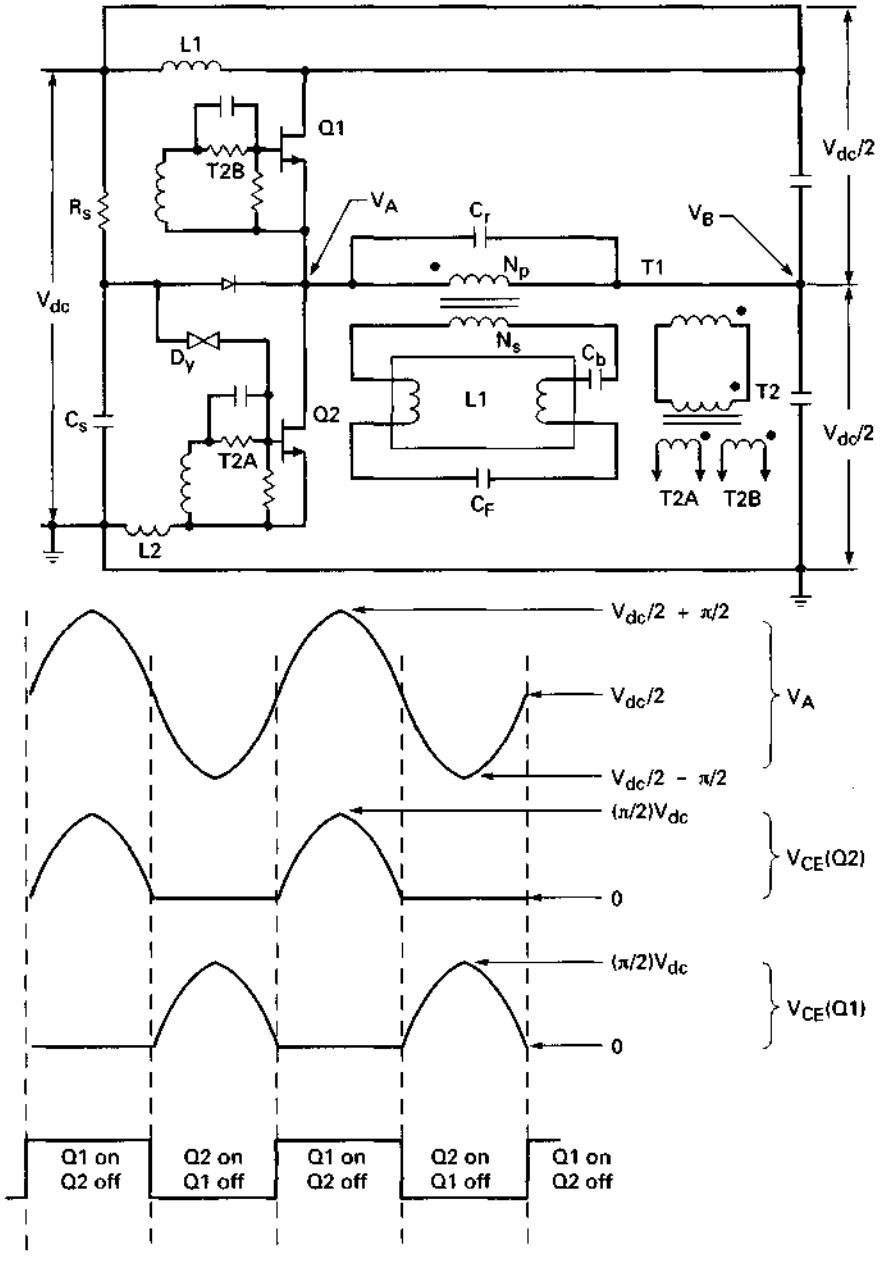


Figure 16.22 Current-fed half bridge topology.

one lamp shown is a rapid-start type driven by the secondary for isolation.

Filament current for the lamp (or paralleled lamps, as this topology permits driving lamps in parallel) is taken in series from the transformer secondary and is limited by C_f and C_b. The circuit starts oscillating when the voltage across C_s has risen above the breakdown voltage of dyac D_y. When the dyac fires, R_s supplies current into the base of Q_2, turning it on. Thereafter, whenever $Q2$ turns on, it discharges C_s, keeping it from interfering with the normal turn-on voltage of the half sinusoid across $T2A$.

Waveforms at significant points are shown in Fig. 16.14*b* to *e*. When $Q1$ turns on, it produces a half sinusoid of positive voltage at V_A. When it turns off, it produces a half sinusoid of negative voltage at V_A. The peak-to-peak voltage across the tank (from V_A to V_B) is πV_{dc}, which is $\pi(400) = 1256$ V for $V_{dc} = 400$ V. The rms voltage across the tank is $V_{rms} = 0.707 \times 1256/2 = 444$ V. Current in the primary is $I_{rms} = V_{rms}/X_{Lt}$, where L_t is the inductance of the transformer primary.

Now L_t and C_t (= C_r + reflected ballast capacitor C_b are calculated as in Sec. 16.6.6 for the current-fed push-pull circuit. The number of primary turns N_p is calculated from Faraday's law = (Eq. 16.10) for a tentatively selected core and as high a peak flux density B_m as possible with still reasonably low core losses. Once N_p is selected, the core gap is chosen from curves (as in Fig. 16.22) so that the selected N_p yields the chosen L_t. The ballast capacitor C_b is chosen from Eqs. 16.1, 16.2, and 16.3.

Figure 16.22*c*, *d*, and *e* shows the maximum voltage stress across an off transistor is $(\pi/2)V_{dc}$, and for $V_{dc} = 400$ V it is $(\pi/2)V_{dc} = 628$ V. There are many inexpensive 700-V transistors to meet this requirement.

16.9 Voltage-Fed Series Resonant Half-Bridge Topology[5,6,7,8]

This topology is shown in Fig. 16.23. It is used for an AC input line voltage of 230 V. Its advantage is that by eliminating any inductors in series with the rectified AC input or the power factor corrected DC voltage, voltage stress on the off transistor is only V_{dc} instead of $(\pi/2)V_{dc}$ V, as in the current-fed half-bridge circuit. For a power factor corrected voltage of 400 V, this means a maximum stress on the off transistor of 400 rather than 628 V and a much lower transistor cost.

Figure 16.23 shows a transformer driving the lamp to provide DC isolation. The circuit has been widely discussed in the literature in its nonisolated version, but currently nonisolated, transformerless circuits are not widely accepted.

It is apparent from Fig. 16.23 that when either transistor is on, the off transistor is subjected to a maximum voltage stress of V_{dc}. The

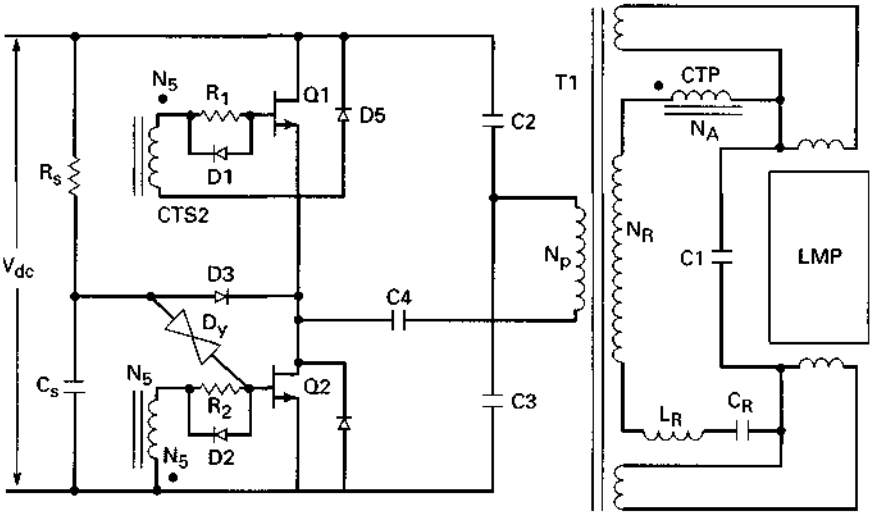


Figure 16.23 Voltage-fed series resonant half-bridge topology. The series resonant circuit comprises L_r, and C_r + $C1$ in series at initial turn on. When the lamp is lit, its low impedance shorts $C1$ out and the resonant frequency decreases. Current transformer primary CTP acts as the current limiting ballast impedance. It also acts as a proportional base drive transformer as its turns ratio N_A/N_S is set equal to the minimum beta of the transistors.

price paid for this advantage is that this series resonant circuit has the same problem of large-amplitude current spikes at turn-on, as discussed for the voltage-fed series resonant push-pull topology.

It is seen in Fig. 16.23 that the series resonant circuit is in the $T1$ secondary and comprises L_r, C_r, the primary of current transformer and $C1$ shunted by the lamp resistance. The high-current turn-on spikes occur because at turn-on, the lamp resistance R_L is high and the equivalent series resistance R_s [$= R_L/(1 + Q^2)$] of Fig. 16.21*b*2 is very low. Resistance R_s is low because [Eq. 16.12] Q ($= R_L C1$) is high. Thus at turn-on, the impedance of the series resonant circuit is the low R_s alone (neglecting for the moment the impedance of the current transformer primary). It is this low R_s at turn-on that is the cause of the high-current spikes. After the lamp is lit, its impedance R_L falls low, Q goes down, and R_s goes up, resulting in the normal current pulses for desired output power.

Current transformer primary (CTP) in series with the resonant $L_r C_r$ elements adds impedance at initial turn-on and helps reduce the amplitude of the turn-on current spikes. The current transformer is a "proportional base drive" transformer (Sec. 8.3.5). Its turns ratio N_A/N_S is set equal to the minimum transistor betas. This always ensures adequate base drive at all current levels as the collector/base current ratio

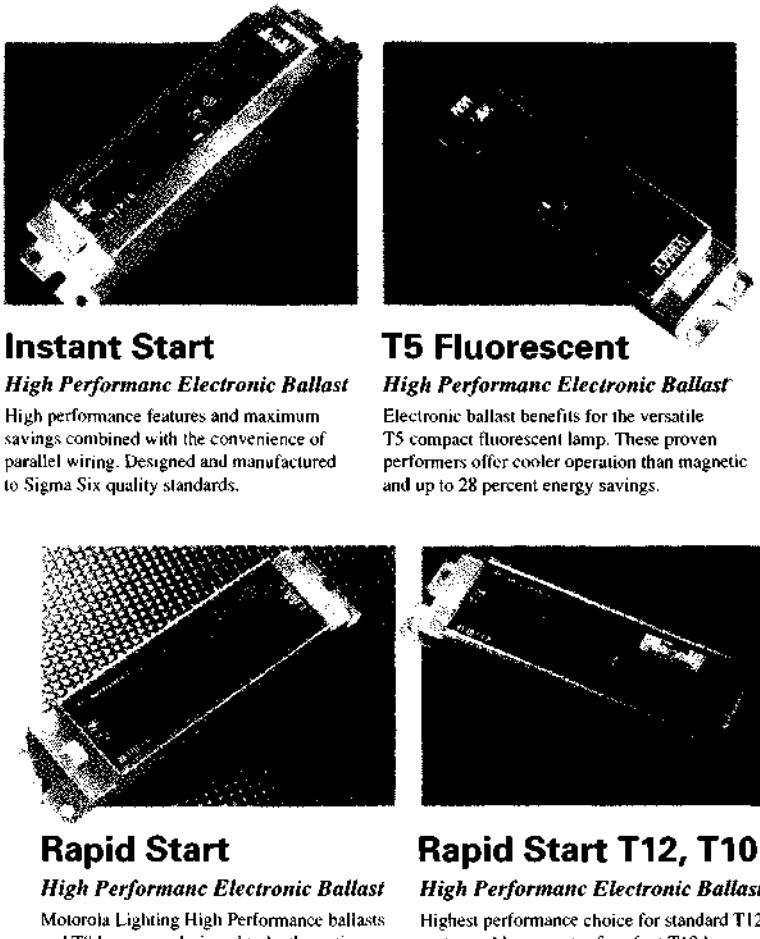


Figure 16.24 Packaged ballast from one major manufacturer (Motorola). (*Courtesy of Motorola Lighting Inc.*)

will always equal the primary/secondary turns ratio. This guarantees adequate base drive at high output current and yet a reduced base drive at low current levels, which minimizes storage time.

At initial turn-on, the circuit oscillates at a frequency of $1/(2\pi\sqrt{L_r C_e})$, where C_e is the capacitance of C_r and $C1$ in series $(C1C_r)/(C1 + C_r)$. When the lamp has lit, its low impedance shunts out $C1$ and the oscillation frequency drops to $1/(2\pi\sqrt{L_r C_r})$. The start circuit comprising R_s,

C_s, dyac D_y, and $D3$ works just as for the parallel resonant half bridge of Sec. 16.8. And C_4 is simply a DC voltage blocking capacitor.

In general, the circuit is not as easy to analyze and design as the current-fed circuit. It is not as easy to ensure that a worst-case design has been obtained and all units coming off a production line will be alike. This is so because of the odd transient conditions at turn-on and because exact equivalent circuits for various phases of the operating cycle are not easily definable, and the exact nature of just how the lamp current is determined is not easily calculable.

16.10 Electronic Ballast Packaging

The entire electronic ballast—the input rectifier, power factor correction building block, and DC/AC converter—must be packaged to fit within the standard housing of the conventional magnetic ballast. Figure 16.24 shows the packaged ballast of one major manufacturer.

References

1. "Fluorescent Lamps," General Electric Bulletin, General Electric Lighting, Cleveland, OH.
2. Motorola Lighting Inc., Technical Publication, Buffalo Grove, IL.
3. G. Meyers and J. C. Heffernan, "The Role of Crest Factor in Fluorescent Lamp Starting," Sylvania Lighting Products, Danvers, MA.
4. GTE/Sylvania Publication, Sylvania Lighting Center, Danvers, MA.
5. *ANSI Fluorescent Lamp Specifications*, American National Standards Institute, New York.
6. "Efficient Fluorescent Lighting Using Electronic Ballasts," Philips Semiconductor Inc., Saugerties, NY.
7. R. J. Haver, "Electronic Ballasts—Power Conversion and Intelligent Motion," April 1987.
8. R. J. Haver, "Solid State Ballasts Are Here," *Electronic Design News*, November 1976.
9. "KoolMu Powder Cores," Magnetics Inc., Butler, PA.
10. "Iron Powder Cores," Micrometals Inc., Anaheim, CA.
11. M. Bairanzade, J. Nappe, and J. Spangler, "Electronic Control of Fluorescent Lamps," Motorola Inc., Application Note AN 1049.
12. "Electronic Ballast Fundamentals," Motorola Lighting Inc., Buffalo Grove, IL.

Chapter

17

Low-Input-Voltage Regulators for Laptop Computers and Portable Electronics

Introduction

The explosion in the use of laptop computers and portable electronics in recent years has led to the formation of a new sector of the power conversion industry. This sector consists of low-input-voltage, battery-fed boost, buck, and polarity-inverting configurations (Secs. 1.3 to 1.5). They are almost entirely contained in one integrated-circuit (IC) package, and most require externally only a single inductor, capacitor, and diode plus about three to five small resistors. Since they operate at frequencies from 60 to 500 kHz, their external capacitors and inductors are small.

They differ from the commonly used PWM control circuits in that they have the main power switch transistor inside the package.

Since they are battery-operated, the output ground need not be isolated from input ground. This saves all the components (optocouplers, pulse transformers, or housekeeping supplies on output ground) usually associated with sensing a voltage on output ground and controlling a pulse width on input ground.

Output powers range from 0.5 up to 100 W depending on the topologies and input and output voltages in the manufacturer's specific type numbers. Efficiencies range from 80 to 95 percent, thus minimizing heat sink size and, for a large range of output powers, obviating their need entirely. For the manufacturer's various specific types, input voltages for boost regulators range from 3 V (two battery cell minimum) up to 60 V and for buck regulators from 4 to 60 V.

In addition to their use as stand-alone, battery-operated DC/DC converters, these devices make possible, without too much design effort, in quick turnaround time and at relatively low cost, a multi-output off-line power supply in a distributed power supply system.

Thus, a conventional off-line power supply can be designed to generate the usual +5-V, high-current isolated secondary. The usual ±12-V and ±15-V or other occasional random slave voltages can be generated at the point of use in either of two ways, depending on their output powers.

For relatively low-power output in the slaves, they can be generated by busing the +5-V output around to the point of use and boosting the +5 V up to the required output voltage with one of the above IC boost regulators. The boost regulators can also generate the negative slave voltages from the +5-V input. For higher powers in the slaves, it would be more efficient (though slightly more expensive) to add another slave of about +24 V and bus it around to the point of use. There, some of the above IC buck regulators would buck it down to the desired lower positive slave voltages or invert it down to the desired negative slave outputs.

At first glance, compared to the conventional scheme of generating slave outputs from added slave secondary windings on the main power transformer, this may appear more expensive and less efficient from a power dissipation viewpoint. But the advantages to this scheme of building an off-line, multi-output supply may outweigh the drawbacks. This will be discussed in detail below.

17.1 Low-input-voltage IC regulator suppliers

The above-discussed IC building blocks are available from two major U.S. manufacturers that have dominated the market over about the last 8 years: Linear Technology in Milpitas, California, and Maxim Integrated Products in Sunnyvale, California. Their products will be discussed in detail. Texas Instruments and Motorola offer some products for this market, but they are not covered here as they generally require more external components.

The discussion here will cover only the applications of these devices with, wherever necessary, some description of their inside design and why some uses are as they are. The basic material comes from the supplier's catalogs. These catalogs show an enormous number of products—boost and buck regulators and polarity inverters. Some devices have variable input and adjustable outputs, fixed input and fixed output (+5-V input, +3.3-output), variable input and fixed outputs (bucking anything from +8- to +40-V input down to usual voltages of +5, +12, or +15 V. A large group are tailored for boosting low-

battery voltages (anything from one to four series cells of 1.5 V up to the usually required +5, +12, or +15 V). The devices are most often pulse-width-modulated, fixed-frequency types at 40, 60, 100, or more recently 200 to 500 kHz. Some operate at a fixed on time and vary frequency to achieve regulation.

17.2 Linear Technology Boost and Buck Regulators[1]

Examples of typical LTC boost and buck regulators are shown in Fig. 17.1*a* and *b*. Boost regulators will be considered first.

Linear Technology offers two families of boost regulators—a high-current one with switch output currents ranging from 1.25 to 10.0 A and a micropower family with output switch currents ranging from 95 to 350 mA.

The boost regulator in the discontinuous and continuous conduction modes is described in Secs. 1.4.1 to 1.4.4 and 15.3.2 to 15.3.3. Over most of the current range of the LTC chips, the circuit operates in the continuous mode whose V_o/V_{in} relation is derived in Sec. 15.3.2 (Eq. 15.1) as $V_o = V_{in}/(1 - T_{on}/T)$. Here T_{on} is the time the internal power switch is on out of a total period T. Throughout the continuous mode, the on time remains constant (Fig. 15.5). At some point, as the output load current decreases, the circuit moves from the continuous to the discontinuous mode. But that is no problem, for if the feedback loop has been stabilized for the continuous mode (by selection of $R3$ and $C1$ in Fig. 17.1*a*), it will remain stable after entering the discontinuous mode.

To build a boost regulator with the LTC chips, the only external components required are shown in Fig. 17.1*a*—the inductor $L1$, capacitor $C2$, diode $D1$, small sampling resistors $R1$ and $R2$, and feedback loop stabilizing components $R3$ and $C1$. The design of the complete circuit requires only selection of these components.

Recalling how the boost regulator works (Secs. 1.4.2 and 1.4.3), we see that what is required at the bottom end of $L1$ (Fig. 15.5*a*) is a sink for current into ground. This is provided in the LTC chip as an *NPN* power switch collector at the bottom end of $L1$ which sinks current into ground for the switch's on time. Thus chips referred to as *boost* chips have *NPN* collectors at the output terminal marked V_{sw} and their emitters at the GND terminal, as shown in Fig. 17.1*a*.

Recalling how a buck regulator works (Sec. 1.3.1), we see in Fig. 1.4 that what is required at the input to $L1$ is a "source" of outward-directed current coming from the source voltage, which is to be bucked down. As seen in Fig. 17.1*b*, this is provided in the LTC buck chips by the emitter of an *NPN* power switch transistor whose emitter is internally connected to the V_{sw} pin and whose collector is internally connected to the V_{in} pin.

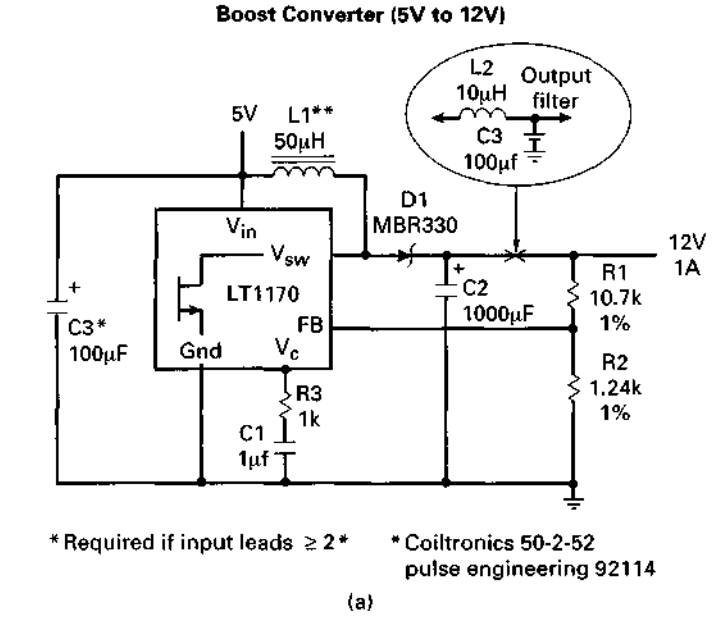


Figure 17.1 Basic boost and buck regulators from Linear Technology, Inc. These regulators have all the usual PWM control circuitry built in plus the power output switch transistor. For most applications, the only large external components required are *L*, *D*1, *C*1, and *C*3.

Although boost chips can be connected up as buck regulators, this requires additional components and it is best to buck with buck chips. Buck chips can also be used to boost a voltage negative with respect to ground to a voltage more negative with respect to ground. These alternate configurations will be shown below.

The devices within each family have nearly identical internal block diagrams and differ only in operating frequency, maximum switch and input voltage, and maximum switch current ratings. The later devices in each family operate at higher frequencies (100 to 500 kHz) to minimize the output inductor size, as that inductor is usually the largest component of the regulator. Since the devices are mostly fed from rechargeable batteries, to maximize time between recharging, a major objective is to maximize efficiency. Most of the dissipation with these devices comes from the volt/ampere drop across the internal *NPN* power transistor and the external diode (Fig. 17.1*a*). Hence in the newer devices, those components are often low R_{ds} power MOSFETS.

The discussion to follow will cover in detail only the most often used members of the buck and boost families. Other members of each family will be listed only in tabular form with their significant specifications shown to permit a quick device selection.

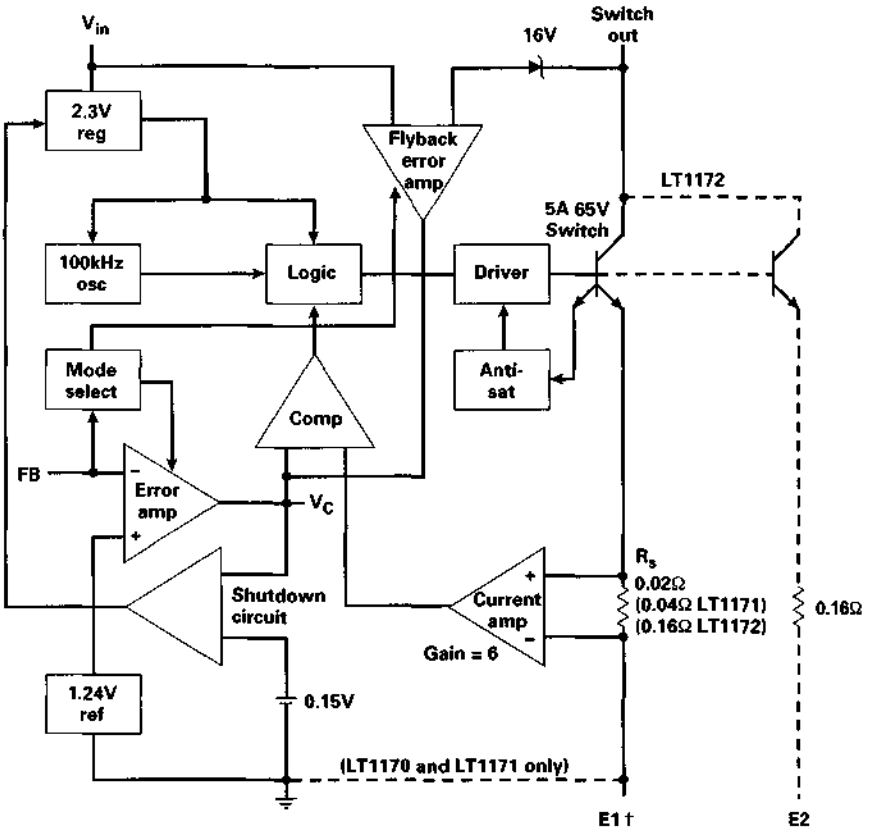
17.2.1 Linear Technology LT1170 boost regulator.[3]

Its use as a basic boost regulator is shown in Fig. 17.1*a*, and its internal circuitry in block diagram form is shown in Fig. 17.2. The device uses current-mode topology (Sec. 5.3 and Fig. 5.3), and its advantages are discussed in Sec. 5.3. Essentially, referring to Fig. 17.2, in current mode, we see the power switch on time commences when the oscillator pulse sets a flip-flop in the logic section and ends when it is reset by the comparator output. This reset instant is determined by the peak current in the power transistor, hence the description as *current mode.*

The comparator compares the DC output of the voltage error amplifier and the ramp-on-a-step voltage output of the current amplifier. At the instant the peak of the ramp-on-a-step current waveform of the current amplifier (converted to a voltage by R_s) exceeds the DC voltage level of the voltage error amplifier output, the comparator output goes positive and resets the flip-flop in the logic block, and the power switch turns off.

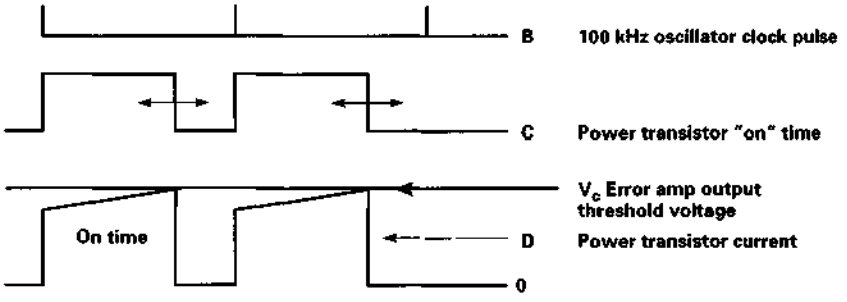
There are thus two feedback loops. The voltage error amplifier senses DC output to keep it constant by setting the threshold level which the ramp-on-a-step voltage waveform from the current amplifier must cross to reset the flip-flop. The second feedback loop monitors peak power switch current on a cycle basis and keeps it constant.

The ramp-on-a-step waveform at the current amplifier output comes about because the ramp-on-a-step current waveform coming



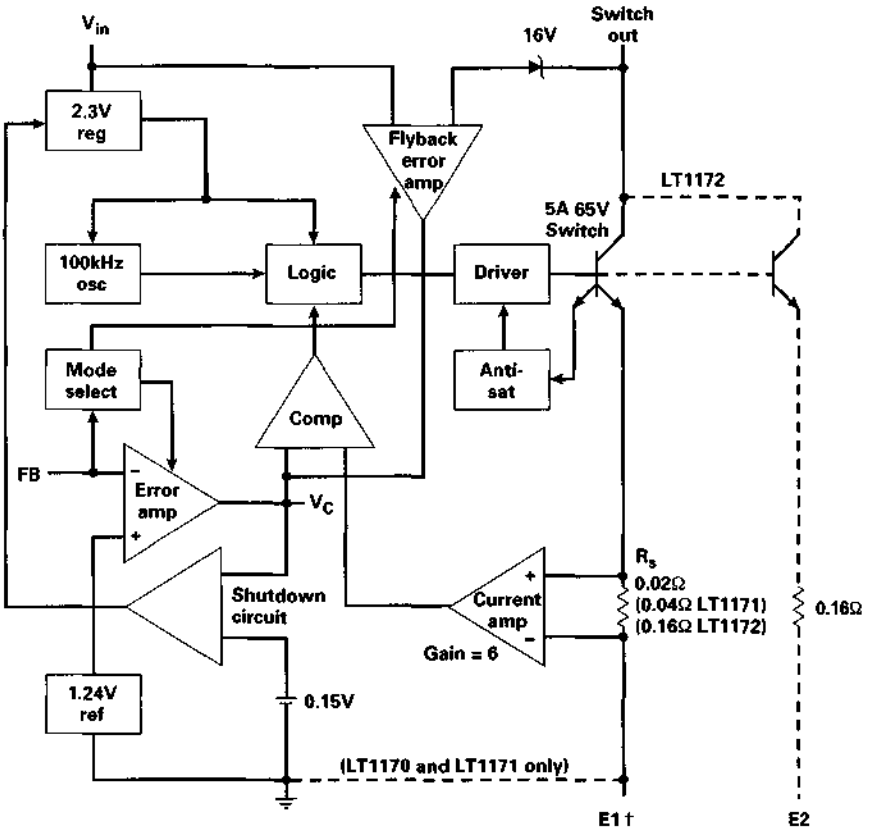
† Always connect E1 to the ground pin on minidip, 8-and 16-pin surface mount packages E1 and E2 internally tied to ground on TO3 and TO220 packages.

(a)



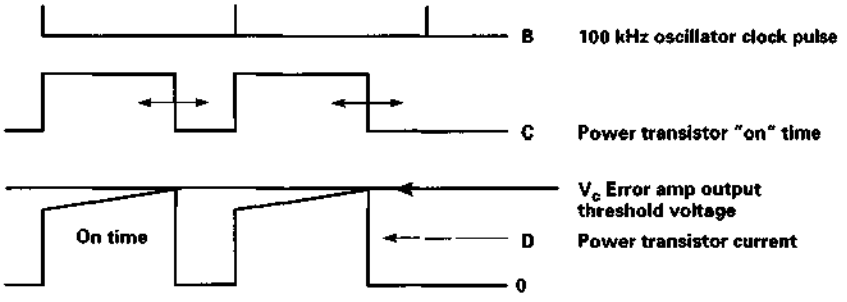
(b)

Figure 17.2 (*a*) LT1170 100-kHz, 5-A boost regulator; (*b*) LT1170 boost regulator waveforms. This is a classical current-mode boost regulator. Output transistor on time is initiated by the clock pulse. It is terminated when the ramp-on-a-step voltage waveform across sensing resistor R_s crosses the threshold set by V_c, the output voltage error amplifier.



† Always connect E1 to the ground pin on minidip, 8-and 16-pin surface mount packages E1 and E2 internally tied to ground on TO3 and TO220 packages.

(a)



(b)

Figure 17.2 (*a*) LT1170 100-kHz, 5-A boost regulator; (*b*) LT1170 boost regulator waveforms. This is a classical current-mode boost regulator. Output transistor on time is initiated by the clock pulse. It is terminated when the ramp-on-a-step voltage waveform across sensing resistor R_s crosses the threshold set by V_c, the output voltage error amplifier.

down through the power switch emitter is turned into a ramp-on-a-step voltage waveform by the internal current-sensing resistor R_s.

The power switch current has a ramp shape because when it turns on, there is a fixed voltage across inductor $L1$ (Fig. 17.1*a*) and current in it rises at a rate of $dI/dT = (V_{in} - V_{cesat})/L1$. The current amplifier is simply an amplifier with a constant gain of 6. Its purpose is to increase the slope of the ramp without having to increase R_s, as the latter would increase dissipation. A relatively steep slope at the comparator input is desirable as small-amplitude noise spikes on top of a flat slope can prematurely reset the flip-flop and shorten the power transistor's on time.

Voltage at the error amplifier output, which sets the instant the flip-flop resets and the power transistor turns off, is brought out at the V_c pin. The DC level at V_c ranges from 0.9 to 2.0 V and is internally clamped at 2 V to limit the power switch peak current. This peak current limit point can be reduced externally by clamping V_c to a regulated voltage less than 2 V through a Schottky diode. This peak current should be selected for the maximum duty cycle, for at maximum duty cycle (which occurs at minimum input voltage), average transistor current and hence transistor dissipation are maximum. This maximum permissible transistor current is determined by thermal considerations which are discussed below.

The reference voltage at the error amplifier input is 1.24 V. Note that in all LT1170 applications, the output voltage sampling resistor string has a resistance of 1.24 kΩ from its FB pin to ground ($R2$, Fig. 17.1*a*). This is done simply to facilitate the calculation of the value of the resistor from the FB pin to the output voltage node ($R1$, Fig. 17.1*a*). Since the DC voltage at the FB pin must equal (within millivolts) the 1.24-V reference, current drawn by $R2$ is 1.24/1240 = 1 mA. And since the same 1 mA must flow through $R1$, the voltage across it is $0.001R1$ and the output voltage is $V_o = 1.24 + 0.001R1$.

The LT1170 operates at a switching frequency of 100 kHz and has a maximum power switch current rating of 5 A. Minimum and maximum voltage ratings at their V_{in} pin are, respectively, 3 and 40 V. Maximum switch output voltage is 65 V.

17.2.2 Significant Polaroid waveforms in the LT1170 boost regulator

It is instructive to examine some significant voltage and current waveforms in an LT1170 test circuit and note how clean and glitch-free they are at the 100-kHz switching rate. Figure 17.3 shows the actual circuit from which the waveforms were taken. The circuit is a basic boost regulator with +12-V output. Waveforms will show line regulation at maximum and minimum load for an input variation

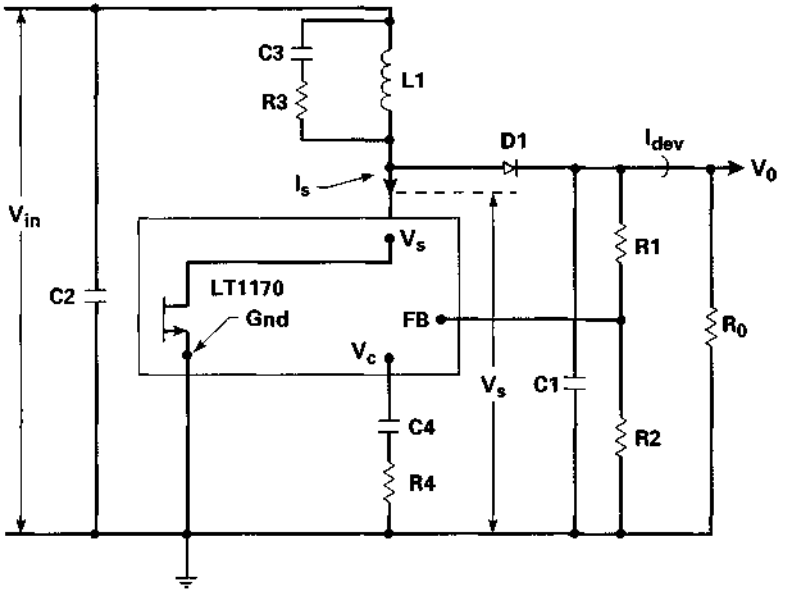


Figure 17.3 Actual test circuit for LT1170 on which waveforms of Figs. 17.4 and 17.5 were taken.
L1–50 mH, 18 turns #20 on MPP 55930 core
R1–13.35 K
R2–1.5 K
R3–2.2 K
R4–440 V
C1–1000 mF, 25 V
C2–1000 mF, 16 V
C3–1000 pF
C4–1.0 μF
D1–MBR 340P

from 4 to 8 V at a constant boosted output voltage of +12 V (Fig. 17.4) as well as load regulation at a constant +5-V input, +12-V output (Fig. 17.5) with a 10:1 load variation (0.082 to 0.82 A).

Figure 17.4*a* and *b* shows power switch voltage and current at +12 V output and a minimum current of 82 mA for input voltages of +4 to +8 V. Note the current waveforms are ramps—characteristic of discontinuous-mode operation. Output voltage increased only by 0.02 V for an input voltage variation of +4 to +8 V. Efficiency was 84 percent, which is reasonably good at an output power level of 1.2 W.

Figure 17.4*c* and *d* shows transistor power switch voltage and current at +12-V output for a 10-fold increase in output current (to 823 mA) for input voltages of +4 to +8 V. Note the output voltage changed by only 0.06 V over these last four input voltage, output current conditions. Waveforms are still equally glitch-free, and the worst-case efficiency is still 81 percent over the above four voltage

and current conditions. Note also now that the switch current waveform is a ramp-on-a-step type which is characteristic of continuous-mode operation. And the transistor on times are exactly what is called for by the relation $V_o = V_{in}/(1 - T_{on}/T)$.

Figure 17.4*e* shows transistor switch current during the on time and output diode current during the off time. These are the classical expected waveforms characteristic of continuous-mode operation. Output diode current at the start of the off time is exactly equal to the transistor switch current at the end of the on time. And output diode current at the end of the off time is exactly equal to the transistor current at the start of the on time. The ramp slopes are determined by the size of the inductor and will be discussed below.

Figure 17.5*a* through *e* shows transistor currents and voltages over a 10:1 load change from 823 to 82 mA for a constant input voltage of +5 V boosted up to an output of + 12 V. Load regulation is excellent—output voltage varies only 0.03 V over the 10:1 load change. And it is seen in Fig. 17.5*a* through *d* that the on time remains constant with the step part of the ramp-on-a-step decreasing as DC load decreases. At Fig. 17.5*e*, the step has been lost, and the transistor on time has decreased as operation entered the discontinuous mode.

17.2.3 Thermal considerations in IC regulators[3]

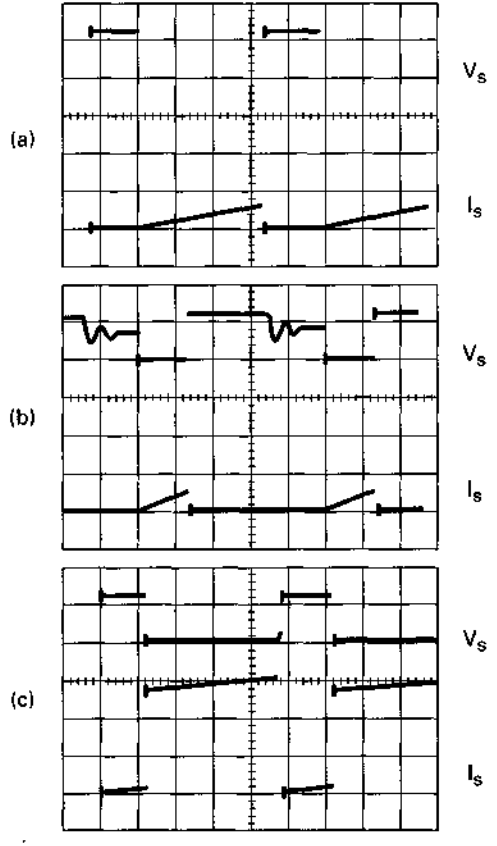
Integrated-circuit regulators considered here differ from the widely used PWM control chips (UC3525 family and similar voltage and current mode controllers) only in that the devices considered here contain the power switch transistor inside the chip. This transistor is the major source of power dissipation ($V_{cesat} I_{switch} \times$ duty cycle) for devices carrying more than about 1 A. Quite often, operating the regulator at its maximum specified transistor switch current results in a need for too large a heat sink for the intended application. But generally, the LTC regulators with output transistor current ratings under 1 A can operate with no or very minimally sized heat sinks.

Thus the thermal calculations should be done early in the design, and they are quite simple. They calculate total power dissipation inside the regulator. This consists of the following two parts:

1. Switch transistor dissipation = $I_{sw} V_{cesat} \times$ duty cycle
2. Internal control circuit dissipation = $V_{in} \times$ average current drawn from the V_{in} pin, where the average current drawn from input pin is

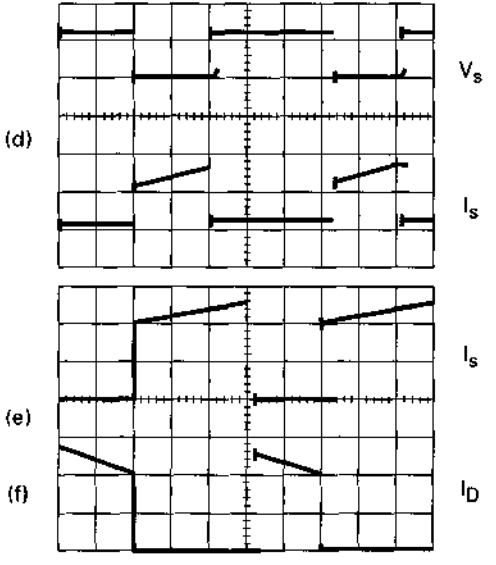
$$I_{av} = 0.006 + I_{sw}(0.0015) + \frac{I_{sw}}{40}(\text{duty cycle})$$

Here 6 mA is the steady-state current drawn by the internal control



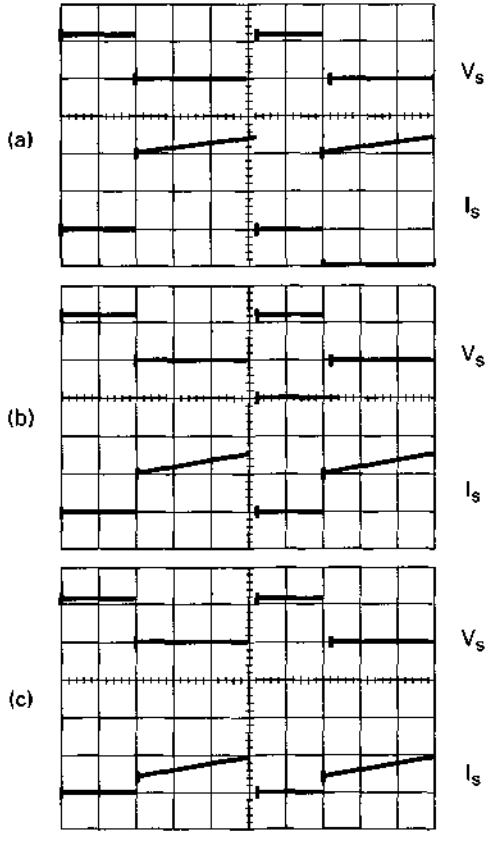
Volts V_{in}	Amps I_{in}	Volts V_0	Ohms R_0	Amps I_0	Watts P_0	Watts P_{in}	% Eff.
4.00	0.30	12.30	150	0.082	1.01	1.21	84.2
8.00	0.15	12.32	150	0.082	1.01	1.20	84.2
4.00	3.14	12.36	15.0	0.824	10.18	12.56	81.0

Figure 17.4 LT1170 boost circuit of Fig. 17.3. (*a, b*) Line regulation at minimum load; (*c, d*) line regulation at maximum load; (*e*) transistor switch current; (*f*) output diode *D*1 current. All V_s waveforms = 10 V, 2 μs/cm. All I_s waveforms = 1 A, 2 μs/cm.



Volts V_{in}	Amps I_{in}	Volts V_0	Ohms R_0	Amps I_0	Watts P_0	Watts P_{in}	% Eff.
8.0	1.35	12.32	15.0	0.823	10.64	10.88	93.9
5.00	2.30	12.35	15.0	0.824	10.2	11.5	88.7

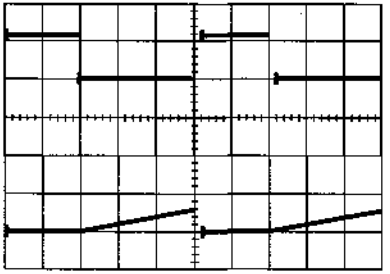
Figure 17.4 *(Continued)*



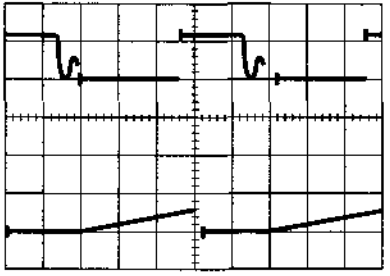
Volts V_{in}	Amps I_{in}	Volts V_0	Ohms R_0	Amps I_0	Watts P_0	Watts P_{in}	% Eff.
5.00	2.30	12.35	15.0	0.823	10.16	11.50	88.4
5.00	1.36	12.34	24.9	0.496	6.12	6.80	90.0
5.00	0.68	12.33	49.4	0.249	3.07	3.40	90.3

Figure 17.5 Load regulation and efficiency at V_{in} = 5.0 V; for boost test circuit of Fig. 17.3. All V_s waveforms = 10 V, 2 μs/cm. All I_s waveforms = 1 A, 2 μs/cm. *Note*: "On" time remains constant throughout the continuous mode (waveforms *a* to *d*).

(d)



(e)



Volts V_{in}	Amps I_{in}	Volts V_0	Ohms R_0	Amps I_0	Watts P_0	Watts P_{in}	% Eff.
5.00	0.34	12.33	100	0.123	1.52	1.70	89.2
5.00	0.23	12.32	150	0.082	1.01	1.15	87.8

Figure 17.5 (*Continued*)

circuitry, and the $0.0015I_{sw}$ term is an increase in this steady-state current which is proportional to I_{sw}. And $(I_{sw}/40) \times$ duty cycle is the average of the power switch base drive current, assuming an average switch transistor beta of 40.

Total chip dissipation (PD_{tot}) is the sum of items 1 and 2. Now most of these LTC regulators are specified to have an absolute maximum operating temperature of 100°C. This is already a manufacturer's derated value, but for greater margin, a safe design should derate this to about 90°C.

Now assume a maximum operating ambient temperature of 50°C, and, for example, calculate for an LT1170 in a TO220 package the thermal resistance of the heat sink required for the peak specified transistor switch current of 5 A.

Assume the regulator boosts up to +5 to +15 V. Then from Eq. 15.1, for boosting up by a factor of 3, duty cycle $\times T_{on}/T$ is 0.66. Item 1 above gives the power transistor dissipation as

$$PD_{sw} = I_{sw}V_{cesat} \times \text{duty cycle}$$

LTC data sheets give V_{cesat} for all their regulators. This is shown in Fig. 17.6 for the LT1170; and for 5-A peak current at, say, 100°C, it is 0.8 V. Then for a duty cycle of 0.66, the power transistor dissipation is

$$PD_{sw} = 5 \times 0.8 \times 0.66 = 2.6 \text{ W}$$

And for the LT1170 in a TO220 plastic package, the thermal resistance from junction to case O_{jc} is 2°C/W.

Then for a transistor dissipation of 2.6 W, the transistor junction is $2 \times 2.6 = 5.2$°C above the case temperature. For 90°C maximum junction temperature, the transistor case must be at $90 - 5.2 = 85$°C.

Now from item 2 above, the average current drawn from V_{in} is

$$0.006 + 5 \times 0.0015 + {}^{5}/_{40} \times 0.66 = 0.096 \text{ A}$$

And for $V_{in} = 5$ V, dissipation due to this average current is $5 \times 0.096 = 0.70$ W. Total dissipation in the chip is then the sum of the dissipation in items 1 and 2, or $2.6 + 0.7 = 3.3$ W.

And finally for a TO220 case at the same temperature as the heat sink on which it is mounted and for an ambient temperature of 50°C, the heat sink to ambient thermal resistance O_{hsamb} must be

$$O_{hsamb} = \frac{85 - 50}{3.3} = 10.6\text{°C/W}$$

And referring to an Aham heat sink catalog, we see this could be achieved with a heat sink like the 342 – 1PP which has a footprint area

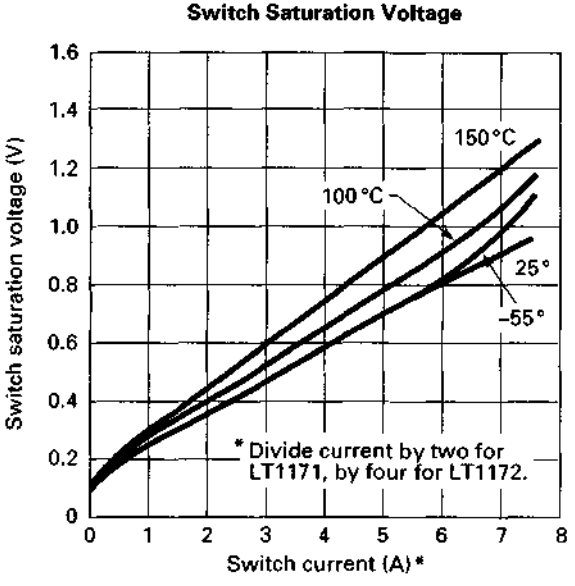


Figure 17.6 Power switch "on" voltage, LT1170 boost regulator. It is the "on" dissipation ($V_{on}I_{on}T_{on}/T$) rather than the peak current rating which determines how much peak current may be drawn. Attempting to operate at the device peak current rating may result in requiring a heat sink very much greater than the device package itself.

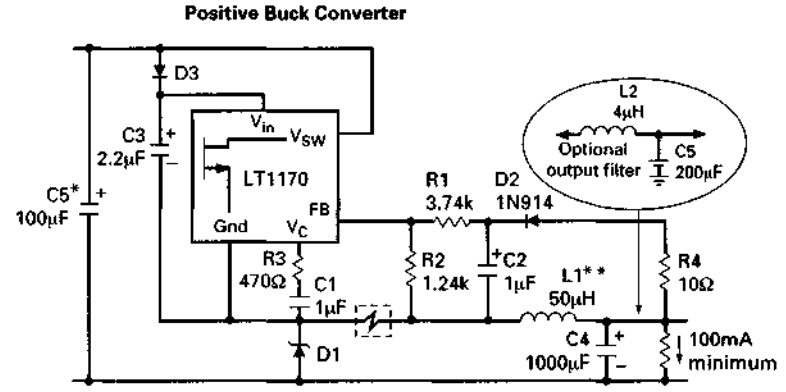
of 1.69 in by 0.75 in with four vertical fins of 0.75-in width and 0.60-in height. This is significantly larger than the device package itself.

Obviously, it is the heat sink size rather than the transistor peak current rating itself which determines the permissible peak operating current level.

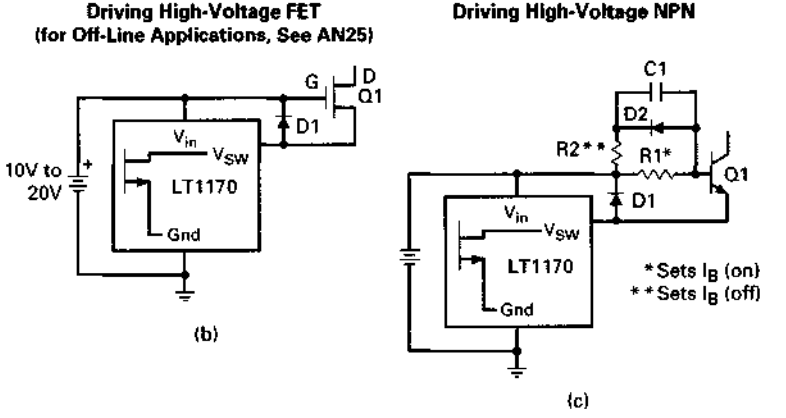
17.2.4 Alternative uses for the LT1170 boost regulator[5]

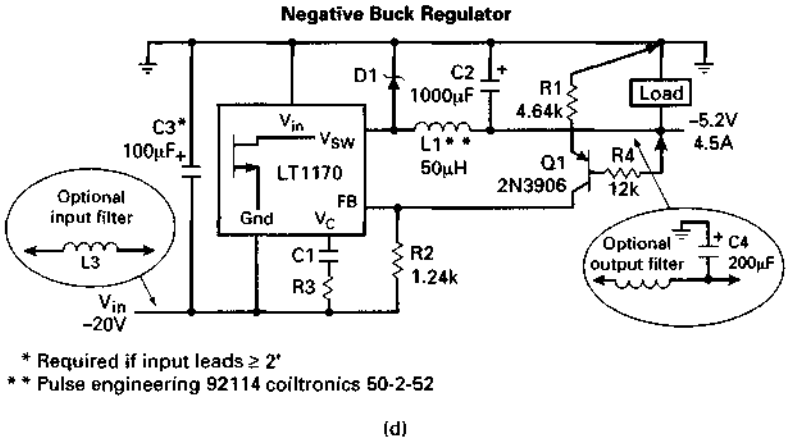
17.2.4.1 LT1170 buck regulator. Although mainly used as a boost converter, the LT1170 can perform many other types of voltage conversion. Some of the most useful ones are shown in Fig. 17.7 (courtesy of Linear Technology Corporation).

Figure 17.7*a* shows a positive buck regulator. Although doable, this requires more external components than a device designed specifically as a buck converter (Fig. 17.1*b*). Recall that the input to $L1$ in a buck converter must be a current source. Thus in Fig. 17.7*a*, it is the emitter of the internal power transistor which drives $L1$. But that emitter is connected internally to the GND pin, and so is the negative end of the internal +1.24-V reference voltage. Thus the +1.24-V reference voltage at the input to the internal error amplifier moves up and down



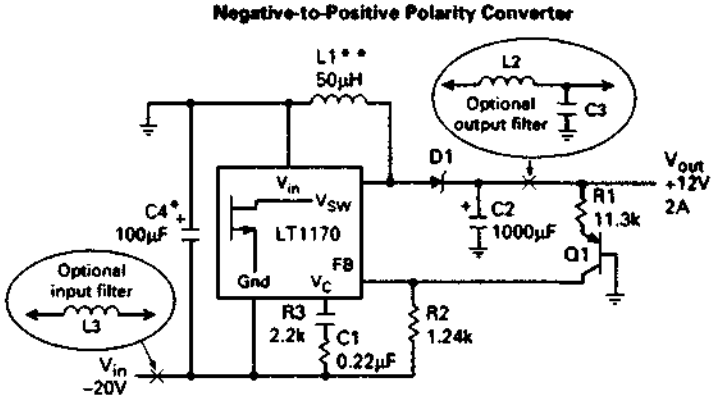
(a)





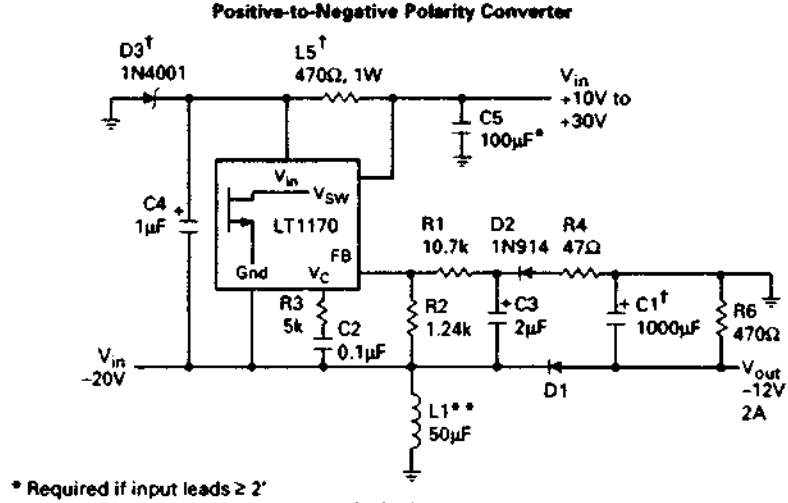
(d)

Figure 17.7 Alternative usages for the LT1170 boost regulator.



* Required if input leads ≥ 2"
** Pulse engineering 92114 coiltronics 50-2-52
This circuit is often used to convert -48V to 5V to guarantee full short-circuit protection the current limit circuit shown in AN19. Figure 39 should be added with C1 reduced to 200pF.

(e)



* Required if input leads ≥ 2"
** Pulse engineering 92114 coiltronics 50-2-52
† To avoid startup problems for input voltages below 10V. Connect anode of D3 to V_{in}, and remove R5. C1 may be reduced for lower output currents. C1 = (500μF)(I_{out}). For 5V outputs, reduce R3 to 1.5k, increase C2 to 0.3μF, and reduce R6 to 100Ω.

(f)

Figure 17.7 *(Continued)*

between ground and V_{in}. Hence to use the internal error amplifier, what is needed is a sample of the output voltage which moves up and down with the GND pin.

This is achieved in Fig. 17.7*a* with *R*1, *R*2, *R*4, *C*2, and *D*2. When *Q*1 inside the chip turns off, GND falls to one diode (*D*1) drop below

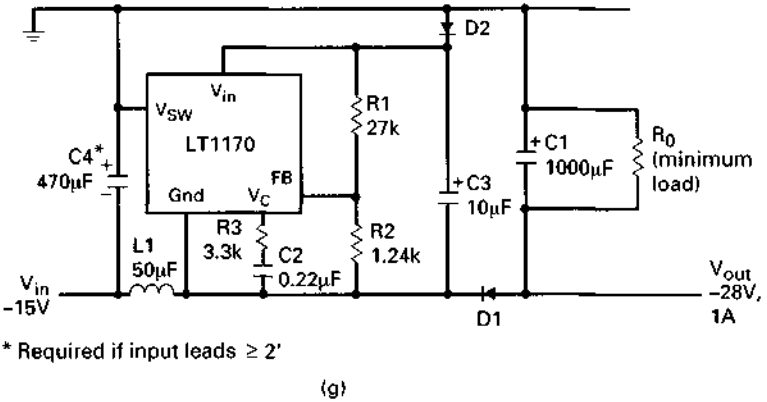


Figure 17.7 (*Continued*)

ground, and $R4$ and $D2$ in series charge the positive end of $C2$ up to one diode drop below V_o. Thus if the drops in $D1$ and $D2$ are almost equal, $C2$ is charged up to a voltage closely equal to V_o. Now the voltage across $C2$ moves up and down with the negative end of the internal reference voltage (GND pin), and it is that voltage which is regulated.

Regulators like those in Fig. 17.1*b* are designed as buck regulators with the internal power transistor emitter connected to the V_{sw} pin so that it can move up and down to drive the external inductor. The internal reference is connected to the GND pin, which is tied to ground and does not move. Hence the sample of the output voltage which is also referenced to ground can be directly tied to the FB pin. Such devices then do not require the artifice of Fig. 17.7*a* to function as a buck regulator. Such devices as the LT1074 will be discussed below.

17.2.4.2 LT1170 driving high-voltage MOSFETS or NPN transistors. This is shown in Fig. 17.7*b* and *c*. In Fig. 17.7*b*, the MOSFET gate is connected to V_{in}, which is fixed with respect to ground and may have any value between +10 V and the maximum safe gate-to-source voltage of the MOSFET (about 15 V).

When the internal transistor in the regulator turns on, it pulls the MOSFET source down close to ground ($I_{ds}R_{ds}$). But the gate is held at $+V_{in}$, and the MOSFET turns on with a gate-to-source voltage of V_{in}. Turn-on is fast because a large current is available from V_{in} to charge the gate-to-source capacitance of the MOSFET. Turnoff is slower—as shown as the internal switch transistor's turnoff time. For the gate-to-source capacitance holds the MOSFET on as the internal transistor turns slowly off. This capacitance is eventually discharged by the decreasing MOSFET current coming through its source.

In driving an NPN transistor as in Fig. 17.7*c*, the base of the transistor is driven by V_{in} through $R1$ which limits the base current. When the

internal transistor has turned on, $C1$ is charged via $R2$ to V_{in} with its right-hand end negative. When the internal transistor turns off, that negative charge on $C1$ speeds up the external transistor turnoff time.

17.2.4.3 LT1170 negative buck regulator. This is shown in Fig. 17.7*d*. When the internal power transistor turns on, V_{stv} is shorted to GND at a voltage of $-V_{in}$. Its collector load is $L1$, which is bridged between $-V_o$ and $-V_{in}$, and the left-hand end of $L1$ is negative with respect to its right-hand end. When the internal power transistor turns off, polarity across $L1$ reverses and the left-hand end of $L1$ attempts to shoot sky high but is clamped to ground by $D1$. Thus the input end of $L1$ switches between $-V_{in}$ and ground. For an on time of T_{on} out of a period T, the average voltage at the input to $L1$ is $-V_{in}(T_{on}/T)$. The $L1$ and $C2$ filter this as in conventional positive buck regulators. The LC components are selected as in Secs. 1.3.6 and 1.3.7. Emitter follower $Q1$ and $R1$ force a constant current proportional to V_o into $R2$. This provides a constant voltage proportional to V_o across $R2$ and is referenced to the chip's ground pin, thus enabling it to serve as the voltage to be regulated.

17.2.4.4 LT1170 negative-to-positive polarity inverter. This is shown in Fig. 17.7*e*. It behaves just as a conventional positive boost regulator. But the 1.24-V reference voltage at the error amplifier input is referred to the GND terminal which is at a voltage of $-V_{in}$ with respect to ground. And the output voltage is positive with respect to ground. Hence an artifice must be used to deliver to the chip, between pins FB and GND, a sample of the output voltage. And this sample must be at a DC voltage level referred to GND.

This is done with *PNP* transistor $Q1$, $R1$, and $R2$. Neglecting the base-emitter drop of $Q1$ and assuming its beta is high, we get $I_{R1} = I_{R2}$. Then $V_{R2} = R2I_{R1} = R2V_o/R1$. And as required, this voltage is referred to the GND pin. But note that there will be errors because of the variations with temperature of $Q1$'s beta and base-to-emitter voltage.

The output/input voltage relation is derived just as in Sec. 15.3.2 for the case when a positive voltage is boosted to a higher positive voltage. Thus assume the internal power transistor is on for a time T_{on} out of a period T. Since the volt-second product across $L1$ when the power transistor is on ($V_{in}T_{on}$) must be equal to that across it when the transistor is off (V_oT_{off}),

$$V_o = \frac{V_{in}T_{on}}{T_{off}} = \frac{V_{in}T_{on}}{T - T_{on}} = \frac{-V_{in}}{T/T_{on} - 1} \tag{17.1}$$

17.2.4.5 Positive-to-negative polarity inverter. This is shown in Fig. 17.7*f*. When the internal power transistor is on, GND is pulled up to

V_{in} via the saturated transistor, and current flows from V_{in} through $L1$ into ground. When the transistor turns off, the voltage across $L1$ reverses polarity and pulls the bottom end of $C1$ negative via $D1$. Again the output/input relation is fixed by equating the volt-second product across $L1$ when the transistor is on to that across it when the transistor is off. This yields the same expression as Eq. 17.1.

Here, as in Fig. 17.7*a*, since the GND pin of the chip moves up and down between ground and V_o, a sample of the output voltage must be brought up between the FB and GND pins, because the internal reference voltage is referred to the GND pin. This is achieved by $R4$, $D2$, and $C3$. When the power transistor turns off, the top end of $L1$ goes negative, pulls the bottom end of $C3$ negative, and clamps it with $D1$ to one diode drop below $-V_o$. The top end of $C3$ is clamped one diode drop below ground by $D2$; hence assuming equal drops in those diodes, the voltage across $C3$ equals $-V_o$ and moves up and down with the GND pin.

Now again, as in Fig. 17.7*a*, with $R2 = 1.2$ kΩ across the FB and GND pins, the sampling circuit $R1$, $R2$ is a 1-mA circuit, and $V_o = 1.2 + 0.001R1$ volts.

17.2.4.6 LT1170 negative boost regulator. This is shown in Fig. 17.7*g* and is much like Fig. 17.7*f*. When the internal power transistor turns on, its emitter at GND pulls the top end of $L1$ up to ground, storing current in it. When the transistor turns off, polarity across $L1$ reverses and pulls the bottom end of $C1$ negative via $D1$. Thus the GND pin swings between ground $-V_{cesat}$ and one diode drop below $-V_o$. Again, equating the positive and negative volt-second products across $L1$ yields the output/input voltage relation

$$V_o = \frac{-V_{in}}{1 - T_{on}/T} \tag{17.2}$$

The same artifice as in Fig. 17.7*a* and *f* is used to transfer a sample of V_o across to pins FB and GND as the GND pin switches between roughly ground and $-V_o$.

17.2.5 Additional LTC high-power boost regulators[5]

As mentioned in Sec. 7.1, the very large number of other boost regulators Linear Technology offers differ only in frequency, voltage, and current ratings. A number of them which appear to be of widest possible use are presented here but only in tabular form to permit quick selection. See Table 17.1.

Specific selection will be made, of course, on the basis of voltage and current ratings. The lower-current devices have a higher on resistance. They are generally less expensive but require a larger, perhaps more expensive heat sink. The higher-current devices, though more

TABLE 17.1

LTC boost regulator	Input voltage, V		Switch voltage, V, max.	Frequency, kHz	Switch	
	Min.	Max.			Current, A, max.	Resistance, Ω
LT1170	3.0	60	75	100	5	0.15
LT1172	3.0	60	65	100	1.25	0.60
LT1171HV	3.0	60	75	100	2.5	0.30
LT1270	3.5	30	60	60	8.0	0.12
LT1270A	3.5	30	60	60	10.0	0.12
LT1268	3.5	30	60	150	7.5	0.12
LT1373	2.4	30	35	250	1.5	0.50
LT1372	2.4	30	35	500	1.5	0.50
LT1371	2.4	30	35	500	3.0	0.25
LT1377	2.4	30	35	1000	1.5	0.50

expensive, may be able to operate at the desired current with no heat sink at all. The next criterion for selection is the operating frequency. Higher-frequency devices can use smaller inductors ($L1$, Fig. 17.1*a*), which is important as this is likely to be the largest and, next to the regulator, the most expensive component.

17.2.6 Component selection for boost regulators[3]

Once the regulator chip has been selected, the major components whose magnitude is to be chosen are (Fig. 17.1*a*), $L1$, $D1$, and $C2$.

17.2.6.1 Output inductor *L1* selection. It is desirable to keep the circuit in the continuous mode from maximum down to minimum load. It is seen in Fig. 17.5*a* through *d* that throughout that mode, as the load current decreases, the transistor on time remains constant while the step part of the ramp-on-a-step decreases. At a lower current, when the front end of the step has disappeared (Fig. 17.5*e*) and the discontinuous mode has started, the on time begins to decrease. To decrease the on time, the sampling voltage at the internal error amplifier input must change somewhat, and hence the output voltage must change. In most cases, changes in output voltage are easily acceptable—they may amount only to 10 to 30 mV as the error amplifier DC gain is very large.

Yet the somewhat poorer load regulation in the discontinuous mode and the somewhat greater input line current ripple make it desirable to keep the circuit in the continuous mode from nominal down to minimum load. This is done through $L1$ selection as follows:

In Fig. 17.4*e*, the DC input is the sum of the transistor current when it is on and the $D1$ current when the transistor is off. The actual DC input is that at the center of the ramps in Fig. 17.4*e*. Thus in Figs. 17.5*a* or 17.5*d*, we see that at the end of the continuous mode, transistor current is a triangle starting from zero and the DC input is

the average current at the center of the triangle. This current is the direct input current ($I_{\text{dc min}}$) at minimum specified input voltage ($V_{\text{in min}}$). This then corresponds to DC input at minimum input power at minimum input voltage.

Thus the change in the transistor input current (dI in Fig. 17.5d) is $2I_{\text{dc min}}$ and

$$L = V_{\text{in min}} \frac{dt}{dI} = \frac{V_{\text{in min}} T_{\text{on}}}{2I_{\text{dc min}}}$$

And from Eq. 5.1, $T_{\text{on}} = T(V_o - V_{\text{in}})/V_o$, so

$$L = \frac{V_{\text{in min}}(V_o - V_{\text{in min}})T}{2V_o I_{\text{dc min}}} \tag{17.3}$$

And $I_{\text{dc min}}$, the direct current at the end of the continuous mode, is usually taken as 10 percent of the current at maximum input power.

Thus for the circuit of Fig. 17.3 where +5 V was boosted up to +12 V, $T_{\text{on}} = T(12.3 - 5)/12.3 = 0.59T$, and from Eq. 17.3

$$L = \frac{5 \times 0.59T}{2 \times 0.1 \times 2.3} = 64\ \mu\text{H}.$$

The closest value, a 50-μH inductor, was used—the exact value is not critical, it only sets the direct current at which the discontinuous mode commences.

17.2.6.2 Output capacitor *C1* selection.[3] The output filter capacitor C1 in a boost regulator is selected for minimum equivalent series resistance (ESR) to minimize peak-to-peak ripple at the switching frequency. This ripple is very large in a boost, as can be seen as follows.

Thus when the internal transistor switch turns on, diode $D1$ is reverse-biased and all the DC load ($I_{\text{dc }o}$) comes out of $C1$. This causes a negative dip in outside voltage whose magnitude is ESR $\times I_{\text{dc }o}$ volts and whose duration is T_{on}. When the transistor turns off, there is a positive spike of amplitude ESR($I_{\text{dc }o} V_o/V_{\text{in}}$). For $V_o/V_{\text{in}} = 3$, the peak-to-peak ripple is $4I_{\text{dc }o}$ ESR.

The simplest way to minimize this ripple voltage is, of course, to select a $C1$ with minimized ESR. But generally capacitor vendors do not list that parameter for their devices.

It is of interest to calculate the peak-to-peak ripple voltage in a specific case. LTC Application Note AN 19-60 offers an empirical approximation to ESR as follows:

Mallory type VPR aluminum electrolytics:

$$\text{ESR} = \frac{200 \times 10^{-6}}{\text{CV}^{0.6}}\ \Omega$$

Sprague 673D, 674D aluminum electrolytics:

$$\text{ESR} = \frac{400 \times 10^{-6}}{\text{CV}^{0.6}}\ \Omega$$

Now assume a regulator boosting from +5 to +15 V with 25-W (1.66-A) output. Assume a Mallory VPR 200 = mF capacitor rated at 25 V. Then from the above relation

$$\text{ESR} = \frac{200 \times 10^{-6}}{200 \times 10^{-6} \times 25^{0.6}} = 0.145\ \Omega$$

Then from the above relation, the peak-to-peak ripple is $V_{o\,\text{rp}} = 4I_{\text{dc}\,o}$ (ESR) = $4 \times 1.66 \times 0.145 = 0.963$ V.

Modern tantalum capacitors may have lower ESR and yield lower ripple. Ripple may be reduced by increasing capacitance, using capacitors with higher voltage ratings or paralleling capacitors. All these techniques increase the required space. Compared to any of the above approaches, a smaller volume may result if a smaller capacitor with larger ESR is used and the resulting larger peak-to-peak ripple is eliminated with a small *LC* filter. Since the ripple frequency is twice the switching frequency, such a filter would be very small.

It is interesting to note, in a buck regulator, that the switching frequency ripple is not as serious. There, as seen in Fig. 17.1*b*, the output capacitor never supplies all the output load current by itself, whether the transistor is on or off. That current is always being supplied by $L1$ and $C1$ in parallel. Ripple current in the inductor flows in a loop through $C1$, diode $D1$, and back into the input end of $L1$. And that ripple current is minimized by selection of a relatively large inductor, as described in Sec. 1.3.6.

Because the output capacitor in a boost regulator supplies all the DC load every time the power transistor turns on, it is important to verify that the capacitor does not exceed its ripple current rating. Manufacturers often do not specify maximum ripple current limits for their devices.

17.2.6.3 Output diode selection (*D*1, Fig. 17.3). This diode is most often a Schottky type as it is second only to the power transistor inside the

chip in power dissipation. Its power dissipation, assuming a 0.5-V drop across it when it conducts, is

$$\mathrm{PD}_{D1} = \frac{0.5 I_{\text{in max}} T_{\text{off}}}{T} = \frac{0.5 P_{in}}{V_{\text{in min}}} \frac{T_{\text{off}}}{T} \qquad \text{W}$$

17.2.7 Linear Technology buck regulator family

Buck regulators (Fig. 1.4) were the earliest types of switching regulators and are discussed in detail in Sec. 1.3. The theory of operation, significant waveforms, and component selection are described. The concept of continuous-use and discontinuous-mode operation is introduced, and Polaroid waveforms showing the gradual transition between the two modes are shown as the output load current is decreased (Figs. 1.6 and 1.7).

LTC integrated-circuit regulators are, in basic operation, much like the circuit of Fig. 1.4. They produce a lower V_o from a higher V_{in} voltage by introducing a low-impedance, single-pole, single-throw switch in the form of a saturating transistor between the source and the load. By modulating its T_{on}/T ratio, this yields, after filtering with an LC circuit, a DC output voltage $V_o = V_{in}(T_{on}/T)$.

The LTC integrated-circuit buck regulators have all the circuitry of Fig. 1.4, including the power switching transistor, inside the package. And in the simplest case, the only external components are L_o, C_o, $D1$, and $R1$, $R2$. Most often they operate at fixed frequencies from 100 to 500 kHz (some to 1 MHz to minimize the size of L_o and C_o) and regulate by modulating T_{on}, as shown in Fig. 1.4. Some regulate the output voltage by operating with a constant off time and varying frequency. Most use current-mode topology (Sec. 5.2) rather than the voltage-mode scheme of Fig. 1.4.

LTC offers an enormous number of new designs, and there is a major driving force for them to improve efficiency, as their major use is for rechargeable, battery-operated equipment—laptop computers and portable consumer-type electronics. For such equipment, even about 2 to 8 or possibly 10 percent improvement in efficiency offers a large increase in the time between battery rechargings. And since a major reason for lowered efficiency is the volt-ampere drop across the saturated switch and across the freewheeling diode ($D1$, Fig. 1.4), those components in the new designs are replaced by low R_{ds} power MOSFETS.

17.2.7.1 LT1074 buck regulator. A typical example of high-power LTC positive buck regulator, the LT1074, is shown in Fig. 17.1*b*, and its internal block diagram is depicted in Fig. 17.8*a*. It is seen in Fig. 17.7*a* that a "boost" regulator chip can also be used as a positive buck regu-

lator but at the cost of added components ($D2$, $C2$, $R1$, and $R2$). But a chip designed primarily as a buck, such as the LT1074, has the emitter of its power transistor not fixed to anything and comes out at the V_{sw} pin. There it can move up and down and source current into $L1$ from its collector, which is internally connected to V_{in}. It thus does not need the artifice of Fig. 17.7*a* to function as a buck. The essentials of the LT1074 are given below (see Fig. 17.8).

The power transistor is a Darlington to supply the high specified output current of 5 A. The Darlington base drive needs a source of input current, and this is supplied from the collector of *PNP* transistor $Q3$. The $Q3$ base needs a current sink, and this is supplied by the output of positive NAND gate $G1$. The $G1$ output goes low when both its inputs go high. That occurs shortly after the R/S latch has been set and its Q output has gone high. The positive pulse from the oscillator, which sets the latch, also inhibits $G1$ for its duration by a "low" at the $G2$ output. This limits the maximum on time of the power transistor to a full period less the trigger pulse width, as a number of harmful things can occur if the power transistor never turns off. This then limits how close V_{in} may come to V_o and still regulate.

The on time is terminated when the output of comparator $C1$ goes positive to reset the latch. The comparator compares the DC level from the multiplier output to an internally generated 3-V triangle. The instant the triangle exceeds the multiplier output, the comparator output goes positive and resets the latch, turning off the power transistor. The circuit is thus a *voltage-mode* circuit—the power switch's on time is a function only of output voltage, not of that and the peak switch current (Sec. 5.2).

The multiplier output DC level is proportional to the error voltage amplifier output and inversely proportional to the input voltage. Thus an increase in V_{in} or an increase in V_o (increase in voltage at the FB pin) shortens the on time and keeps the output constant. Making the multiplier responsive to V_{in} changes directly (instead of via a change in V_o), which results in faster correction of V_{in} changes.

Power transistor current limiting is done with $C2$. If its output goes positive, it resets the latch and turns off the power transistor. The threshold at which current limiting occurs is set by the small negative bias across $R1$, and that is set by the current drawn through it by emitter-follower $Q4$. That current is kept relatively constant by the relatively constant voltage at the $Q4$ base. When the voltage drop across R_s exceeds the bias across $R1$, the output of comparator $C2$ goes positive, resets the latch, and turns off the power transistor.

Significant waveforms (idealized but yet quite accurate) in a typical LT1074 are shown in Fig. 17.9.[6]

Block Diagram

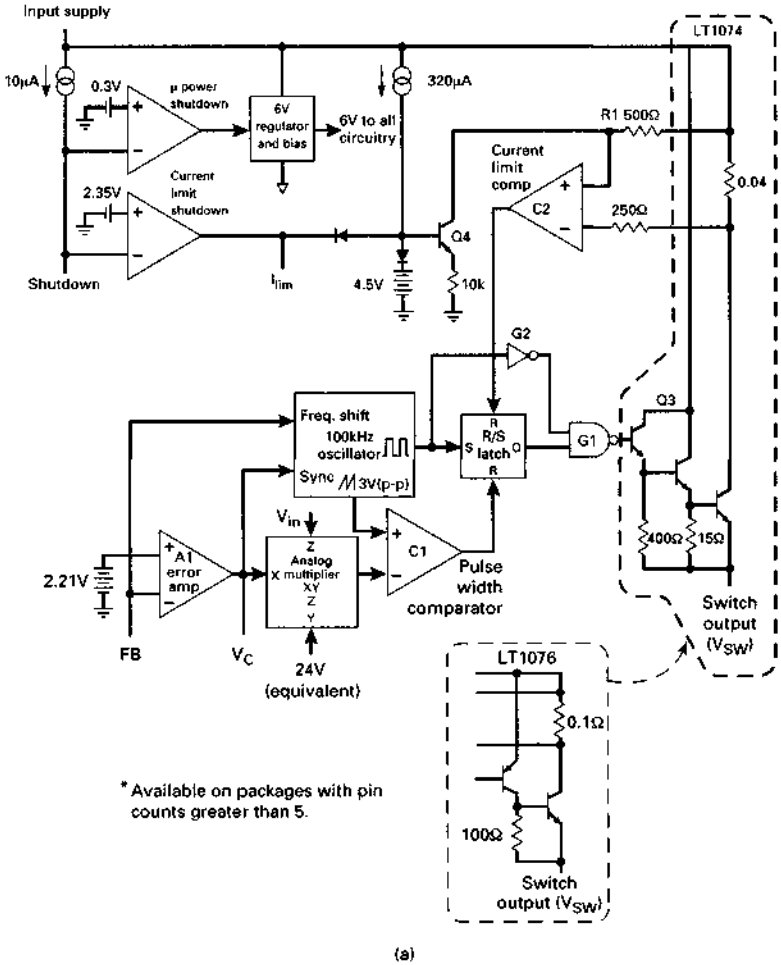


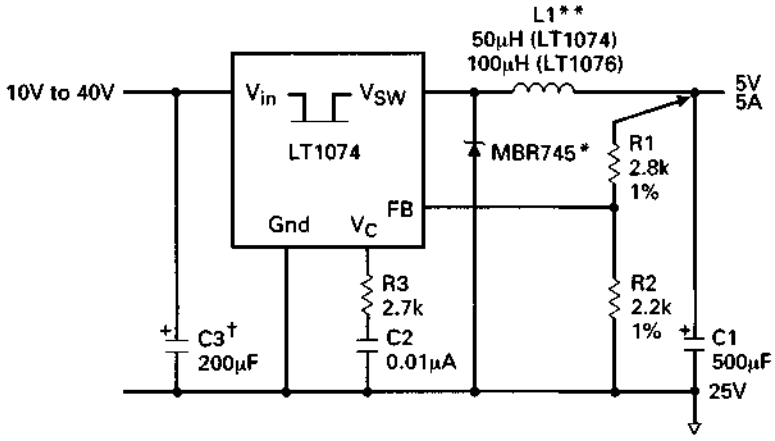
Figure 17.8 Linear Technology LT1074 100-KHz, 5-A buck regulator.

17.2.8 Alternative uses for the LT1074 buck regulator

17.2.8.1 LT1074 positive-to-negative polarity inverter. This is shown in Fig. 17.10*a*. The output/input relation can be derived, as always, by equating the volt-second product across $L1$ when the power transistor is on to that across it when the transistor is off. Thus when the transistor is on, it brings the V_{sw} pin up to V_{in} and the volt-seconds across $L1$ is $V_{in}T_{on}$. When the transistor turns off, the polarity across $L1$ reverses and the bottom end of $L1$ plunges negative until $D1$ conducts

Typical Application

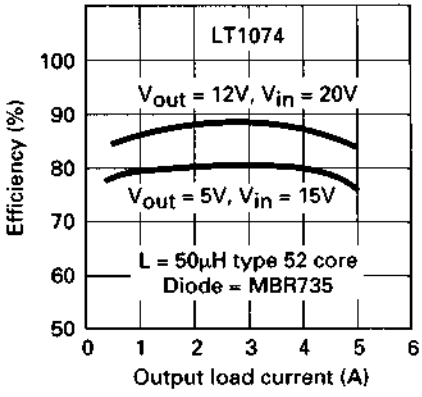
Basic Positive Buck Converter



* Use MBR340 for LT1076
* * Coiltronics #50-2-52 (LT1074)
#100-1-52 (LT1076)
Pulse Engineering, Inc.
#PE-92114 (LT1074)
#PE-92102 (LT1076)
Hurricane #HL-AK147QQ (LT1074)
#HL-AG21OLL (LT1076)
† Ripple current rating ≥ $I_{out}/2$

(b)

Buck Converter Efficiency



(c)

Figure 17.8 *(Continued)*

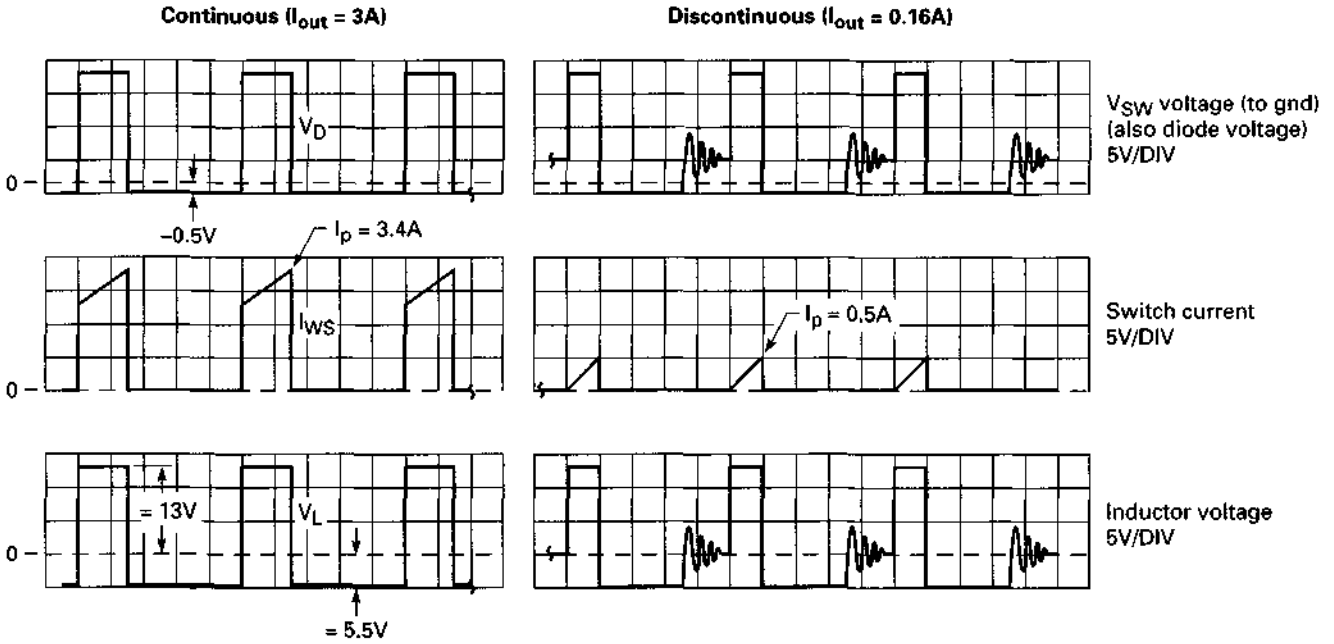
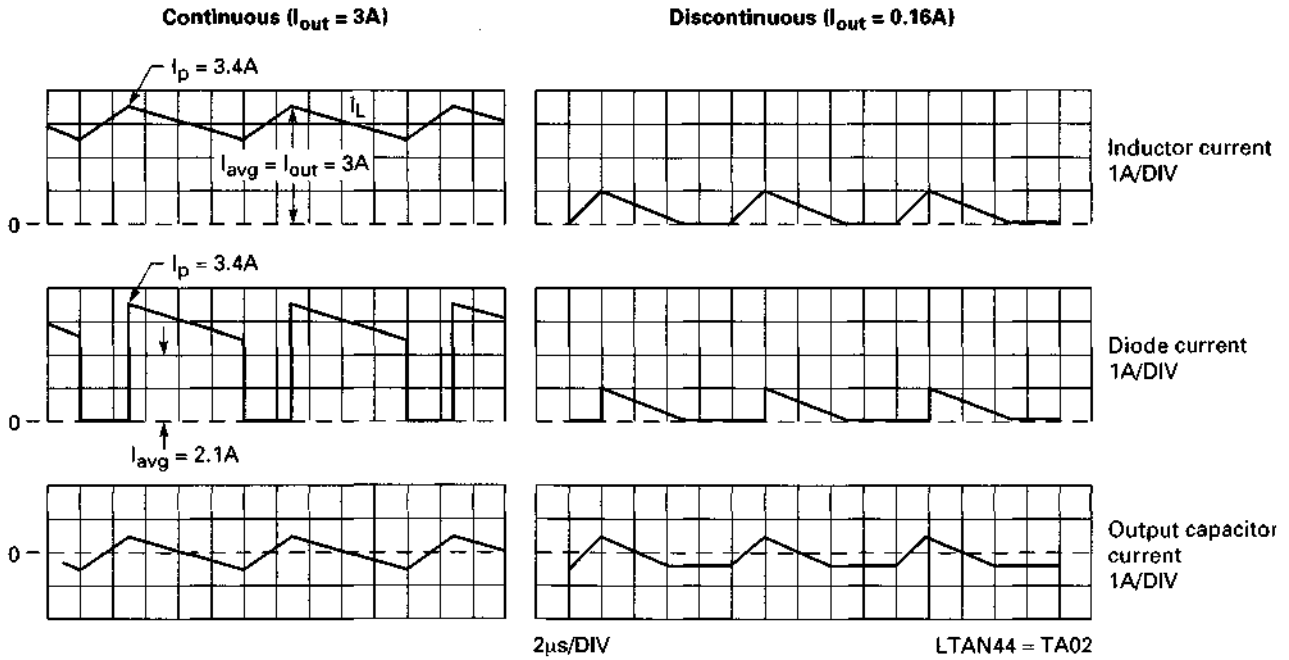


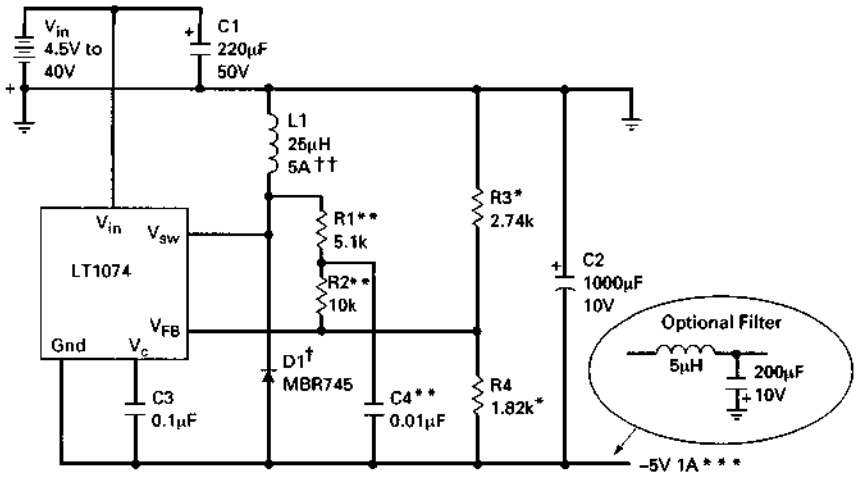
Figure 17.9 Significant waveforms for LT1074 buck regulator of Fig. 17.8.



Buck Converter Waveforms with V_{in} = 20V, L = 50μH

Figure 17.9 (*Continued*)

Positive-to-Negative Converter with 5V Output



† Lower reverse voltage rating may be used for lower input voltages. Lower current rating is allowed for lower output current. See AN44.

†† Lower current rating may be used for lower output current. See AN44.

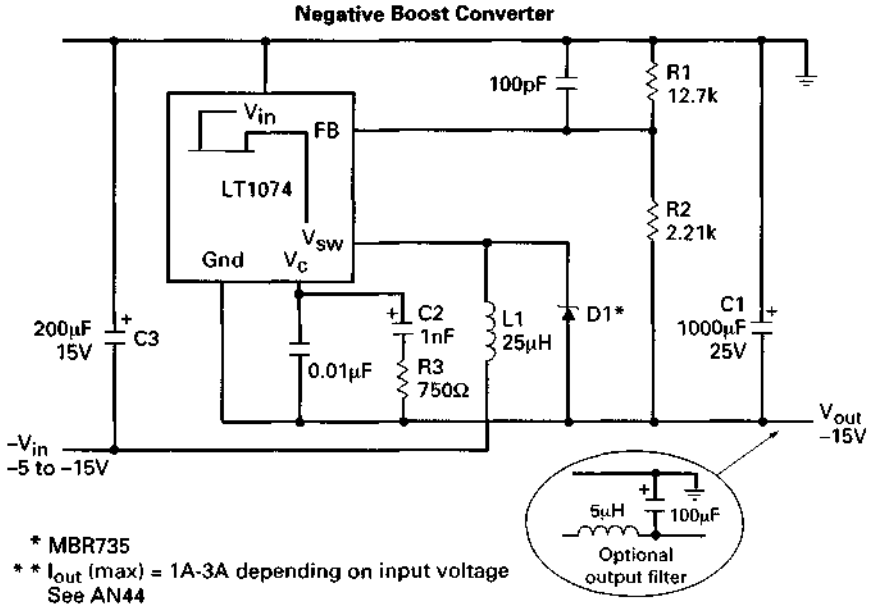
** R1, R2, and C4 are used for loop frequency compensation with low input voltage, but R1 and R2 must be included in the calculation for output voltage divider values. For higher output voltages, increase R1, R2, and R3 proportionately. For input voltage > 10V, R1 R2, and C4 can be eliminated, and compensation is done totally on the V_C pin.
R3 = V_{out} –2.37 (KΩ)
R1 = (R3) (1.86)
R2 = (R3) (3.65)

*** Maximum output current of 1A is determined by minimum input voltage of 4.5V higher minimum input voltage will allow much higher output currents. See AN44.

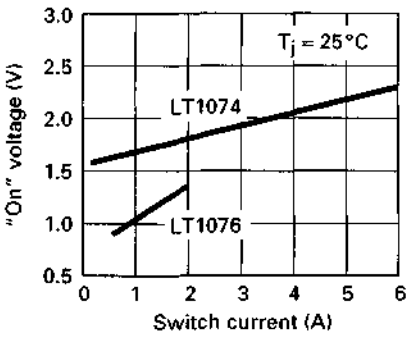
* = 1% film resistors
D1 = Motorola-MBR745
C1 = Nichicon-UPL1C221MRH6
C2 = Nichicon-UPL1A102MRH6
L1 = Coiltronics-CTX25-5-52

(a)

Figure 17.10 Alternative usages for the LT1074 buck regulator. Note the triple Darlington output configuration of Fig. 17.8 yields a 2.2-V "on" drop at 5 A and may require a quite large heat sink for operation at the 5-A peak current rating.



(b)



(c)

Figure 17.10 *(Continued)*

and pulls the bottom end of C2 to a voltage V_o with respect to ground for a time T_{off}. Then

$$V_{in}T_{on} = V_o T_{off} = V_o(T - T_{on})$$

or

$$V_o = \frac{V_{in} T_{on}}{T - T_{on}} = \frac{-V_{in}}{T/T_{on} - 1} \tag{17.4}$$

17.2.8.2 LT1074 negative boost regulator. This is shown in Fig. 17.10*b*. Again, we equate volt-second products across $L1$ in the on and off states. When the power transistor is on, V_{sw} is brought up to ground and the volt-second product across $L1$ is $V_{in} T_{on}$. When the transistor is off, polarity across $L1$ reverses and the top end of $L1$ dives negative and pulls the bottom end of $C1$ to negative voltage such that

$$V_{in} T_{on} = (V_o - V_{in}) T_{off} = (V_o - V_{in})(T - T_{on})$$
$$= V_o(T - T_{on}) - V_{in} T + V_{in} T_{on}$$

or

$$V_o = \frac{V_{in} T}{T - T_{on}} = \frac{V_{in}}{1 - T_{on}/T} \tag{17.5}$$

17.2.8.3 Thermal considerations for LT1074.[6] If a buck regulator for, say +24-V input, +15-V and 5-A output was being considered, the prospect of doing this 75-W regulator with an LT1074 in a TO220 package with only LCD components, as in Fig. 17.1*a*, might seem very attractive at first glance. But thermal calculations will show that the internal power dissipation, rather than the power transistor peak current rating, is very often the limiting factor to a practical design.

For although internal dissipation varies with input and output voltages, attempting to operate at the transistor's peak current at almost any input/output voltage often results in a heat sink many times that of the TO220 regulator package itself. This can be seen as follows for the above typical buck regulator: $V_{in} = +24$ V and $V_o = 15$ V at 5.0 A:

For $I_o = 5$ A, Fig. 17.10*c* shows the transistor's on voltage is 2.2 V. Then for $V_{in\ nominal} = +24$ V, $V_{in\ min} = 22$ V, $V_o = +15$ V, duty cycle (DC) = $T_{on}/T = V_o/V_{in\ min} = 15/22 = 0.68$. Then transistor switch dissipation is

$$PD_{sw} = V_{ce} I_o \times DC = 2.2 \times 5 \times 0.68 = 7.5 \text{ W}$$

And the average dissipation of the control circuitry in addition to that is

$$PD_{cc} = V_{in\ min} I_{in\ cc}$$

where

$$I_{in\ cc} = 0.007 + 0.005 \times DC + 2 I_o T_s F$$

Here the first term is the steady-state current drawn from V_{in}, and the second term is the increase in that proportional to output current. The last term is the average of the spikes of current lasting for the switching time T_s which are drawn from V_{in} at the instants of turn-on and turnoff. Then for $F = 100$ kHz and $T_s = 0.06$ μs,

$$PD_{cc} = 22(0.007 + 0.005 \times 0.68 + 2 \times 5 \times 0.06 \times 10^{-6} \times 1 \times 10^{+5})$$

$$= 1.7 \text{ W}$$

The total internal dissipation is then 7.5 + 1.7 = 9.2 W.

Now assuming 50°C ambient temperature, calculate the size of the required heat sink for a desired maximum power transistor junction temperature of 90°C. Assume a TO220 package whose thermal resistance is 2.5°C/W giving a transistor case temperature of 90 − (7.5 × 2.5) = 71°C. Assuming no temperature difference between the transistor case and heat sink, the permissible heat sink temperature rise above ambient is 71 − 50 = 21°C.

And referring to an AHAM heat sink catalog, for a 21°C temperature rise above ambient, we see that a typical heat sink will be the AHAM type S1100 5.5. This is a heat sink with eight fins of 0.461-in height and a footprint area of 5.5 in by 4.5 in. It is obvious, then, that no significant advantage has been gained by putting the power transistor inside the package.

LTC has addressed this problem by offering other regulators (LT1142, 1143, 1148, 1149, 1430) which have less internal dissipation by use of external MOSFET transistors for the transistor switch and freewheeling diode. These have very low R_{ds}, consequent lower on voltage and dissipation. They are available in surface-mounted packages and so still permit a small overall regulator size. Such and other high-efficiency regulators will be described below.

17.2.9 Higher-efficiency LTC high-power buck regulators

17.2.9.1 LT1376 high-frequency, low switch drop buck regulator. This is shown in a typical application in Fig. 17.11*a* and in block diagram form in Fig. 17.11*d*. It is a current-mode circuit (Secs. 5.1 to 5.5) in which the DC output voltage and transistor switch peak current are controlled and determine the transistor's on time.

For output currents from 1 to 1.5 A, it has greater efficiency and requires a smaller heat sink than the LT1074. For output currents under 1 A it may require no heat sink at all. Because it operates at 500 kHz (rather than 100 kHz for the LT1074), its output inductor and capacitor ($L1$ and $C1$ in Fig. 17.11*a*) are much smaller.

Its greater efficiency comes about primarily because of the low volt-

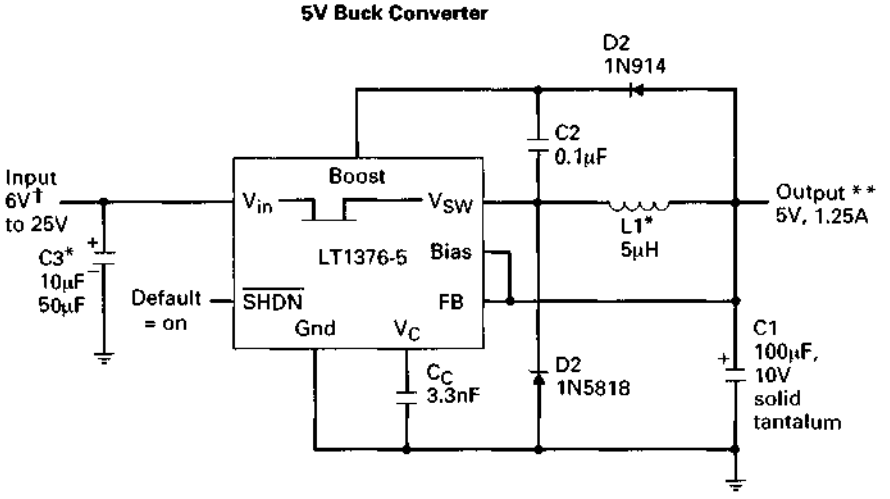


Figure 17.11 Linear Technology LT1375 500-kHz high-efficiency buck regulator. The single Darlington output transistor and the voltage boost provided by *D2*, *C2* result in low-output transistor "on" drop and in efficiency close to 90 percent over a large current range.

age drop across the output transistor (0.5 V at 1.5 A in Fig. 17.11*c*). This compares to 1.7 V at 1.5 A for the LT1074 or 1.25 V for the LT1076 (Fig. 17.10*c*).

There are two reasons for this low drop. First, the output transistor is

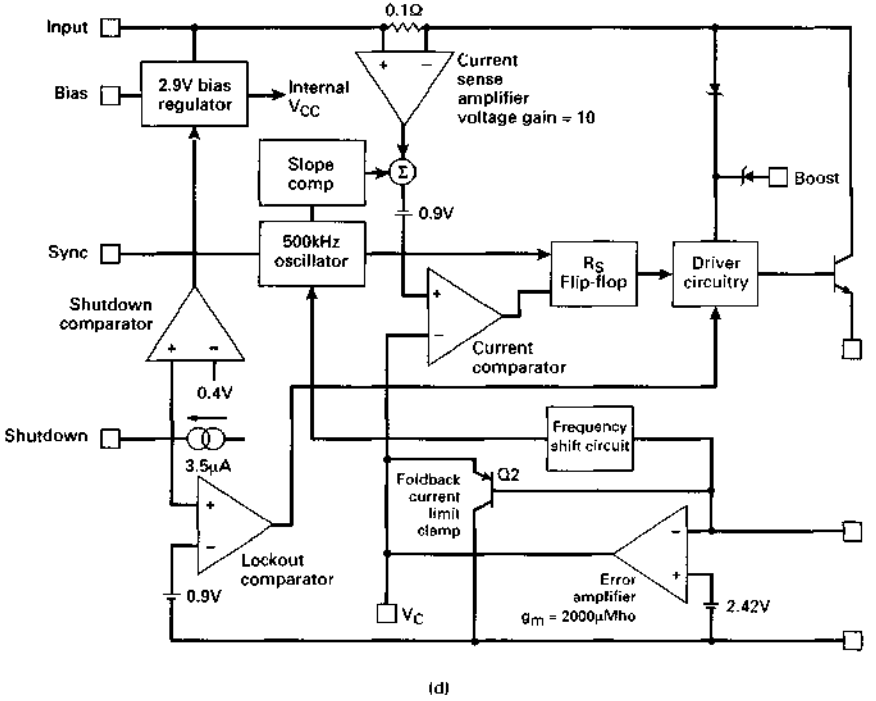
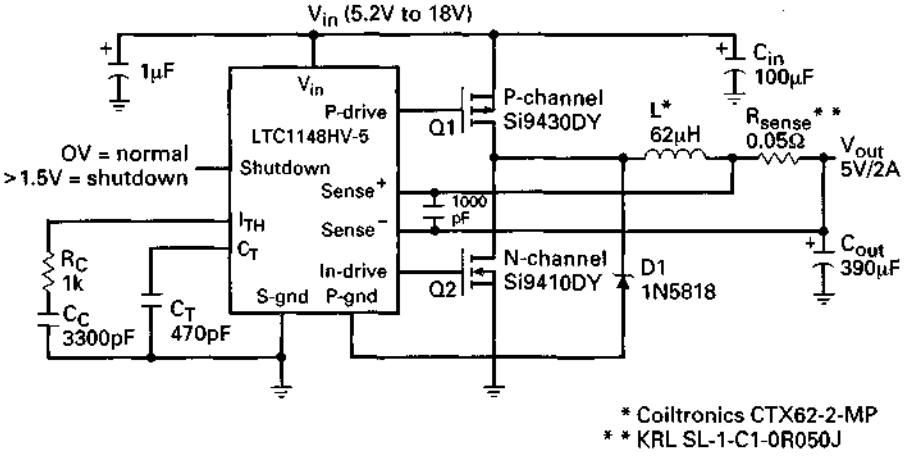


Figure 17.11 (*Continued*)

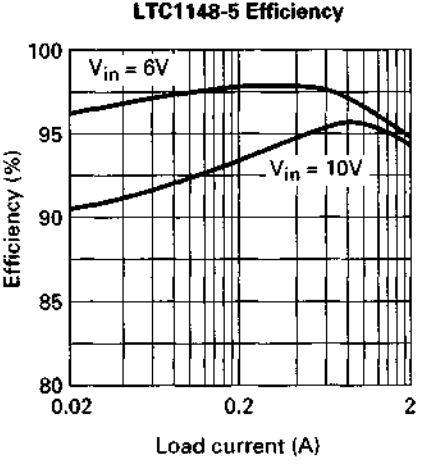
a single transistor rather that a triple Darlington (Fig. 17.8). Second, the output transistor is driven harder into saturation because a voltage higher than the input is made available to provide the hard drive. This higher voltage is produced with few components and inexpensively by $D2$ and $C2$ (Fig. 17.11*a*). When the internal power transistor turns off, the V_{sw} node falls to one diode drop ($D2$) below ground. Now the top end of $C2$ is charged up positively to one diode drop below $+V_o$ by $D2$. And when the internal power transistor turns back on, the top end of $C2$ provides the positive boost voltage for the internal output transistor driver.

17.2.9.2 LTC1148 high-efficiency buck with external MOSFET switches. This is shown in Fig. 17.12*a* in a typical application and in Fig. 17.12*c* in block diagram form. It achieves its high efficiency because of the low-R_{ds}, low on drop P-channel MOSFET ($Q1$) which switches the input end of L up to V_{in} during the on time. And during the off time when $Q1$ is off, the low-R_{ds}, low on drop N-channel MOSFET $Q2$ is on and acts as the freewheeling diode in shunt with $D1$ pulling the input end of L down to ground.

Efficiency is close to 95 percent because of these low on time drops, as can be seen in Fig. 17.12*b*.

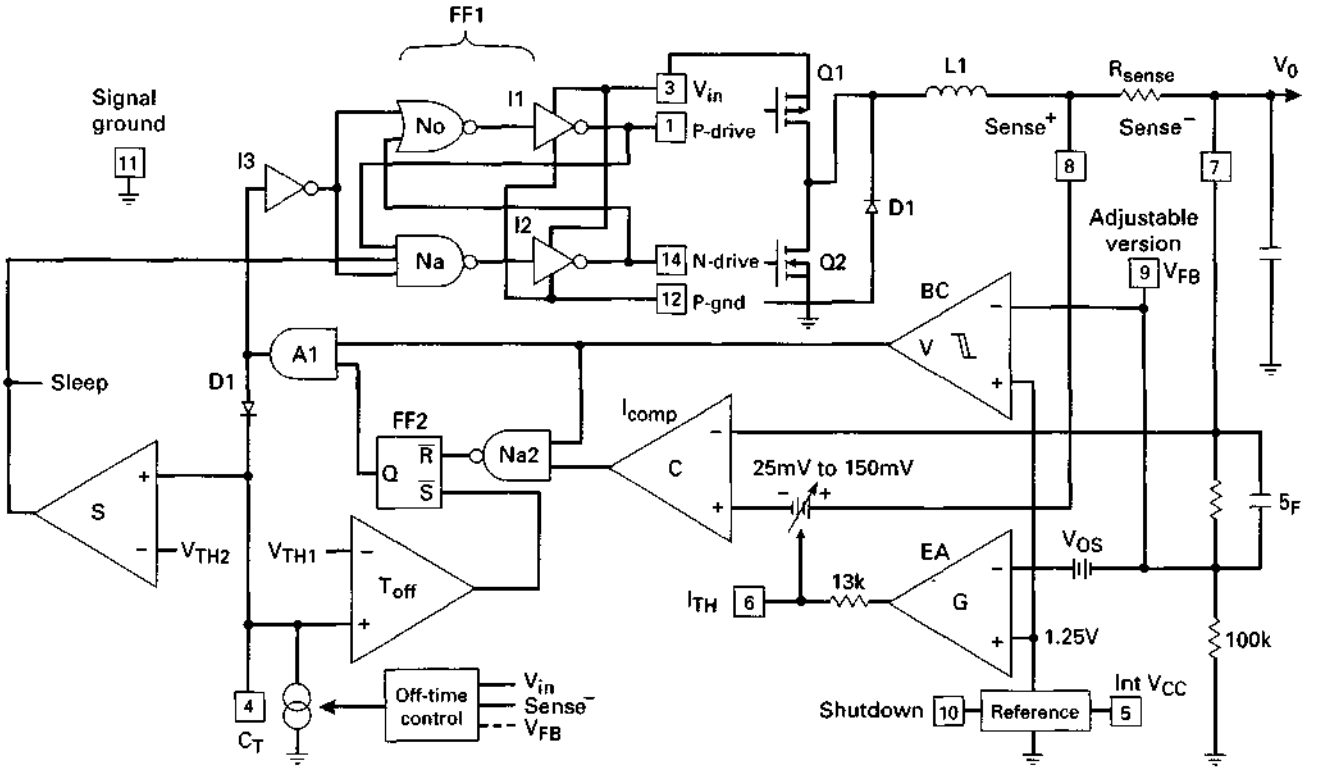


(a)



(b)

Figure 17.12 LIC1148 high-efficiency buck regulator with external MOSFETs. This regulator operates with affixed off time (set by C_t to ground). It regulates by varying switching frequency. The external P and N channel MOSFETs, with their low "on" drops, offer efficiencies exceeding 90 percent over a large current range.



(c)

Figure 17.12 *(Continued)*

The circuit is unusual in that unlike most other LTC products, it operates not at a constant frequency, which regulates by varying the on time. Rather it operates with a fixed off time, which regulates by varying the switching frequency to control the duty cycle.

Historically, there has been an aversion to using any other than a constant-frequency, on time-modulated scheme for regulation. Often the switching power supply in a large system was fed a synchronizing pulse, and it was required that the power supply switching frequency be locked to a submultiple of this pulse, which was synchronized and locked in phase either to a central computer clock or to the horizontal line rate in a CRT display.

There was fear that if the power supply generated either conducted or radiated RFI, it could be picked up and interfere with nearby electronics such as CRT displays or computers. If any generated noise was picked up by a CRT display, if locked in phase with the horizontal line rate of the display, the noise would be stationary on the screen and not as disconcerting to the operator as when it was wandering all over the screen. Also, if picked up by a computer, if not locked in phase to the central computer clock, the noise would have a greater probability of falsely turning 1s into 0s. And further, the Fourier spectrum of a variable-frequency scheme is much wider than a fixed-frequency one and could possibly be more easily picked up by nearby sensitive radio receivers.

But for low-power supplies for laptop computers and portable electronics, where most often there is no other nearby electronic equipment, there is no valid reason for rejecting a variable-frequency voltage-regulating scheme.

17.2.9.3 LTC1148 block diagram operation. Despite regulating the output voltage by varying the frequency rather than by varying the pulse width at a constant frequency, the output/input voltage relation is exactly the same as that for a PWM scheme.

This can be seen (Fig. 17.12*a*) by equating the volt-second product across L when its input is high to that when its input is low. Thus when $Q1$ is on, the input end of L is essentially at V_{in}; and when $Q2$ is on, the input end is at ground. Then

$$(V_{in} - V_o)t_{on} = V_o t_{off} = V_{in} t_{on} - V_o t_{on}$$

or

$$V_o = \frac{V_{in} t_{on}}{t_{on} + t_{off}} = V_{in} \frac{t_{on}}{T} \tag{17.6}$$

From this the frequency versus output/input relation is

$$V_o = \frac{V_{in} t_{on}}{T} = \frac{V_{in}(T - t_{off})}{T} = V_{in}\left(1 - \frac{t_{off}}{T}\right)$$

$$= V_{in}(1 - f t_{off}) \tag{17.7}$$

which can be written as

$$f = \frac{1 - V_o}{V_{in}/t_{off}} \tag{17.8}$$

It can be seen from Eq. 17.7 or 17.8 that for a constant t_{off}, as V_{in} goes up, to keep V_o constant, frequency f must go up. Regulation can be seen as follows (Fig. 17.12).

When $Q1$ is on, P drive, pin 1 is low, keeping the P-channel MOSFET on. And N drive, pin 14 is also low, keeping the N-channel MOSFET off. The four elements N_o, $I1$, N_a, $I2$ form a set-reset flip-flop (FF1), and hence the P and N drive outputs remain locked in the low state, keeping $Q1$ on and $Q2$ off until the flip-flop is reset.

Current in L ramps up with the usual ramp-on-a-step waveform characteristic of an LC output filter. This inductor current is monitored by sensing the voltage drop across R_{sense}. That voltage drop is added to the negative bias voltage at the noninverting input to comparator C. That negative bias is the output of voltage error amplifier G which compares a fraction of V_{out} to the internal 1.25-V reference.

When the voltage across R_{sense} exceeds the bias voltage at pin 6, the output of comparator C goes high, and the output of positive true logic NAND gate N_a goes negative (since its other input is high at this time). This resets flip-flop FF2, causing its Q output to go low. This causes the output of AND gate A1 to go low and commences the off time, which requires the P and N drive outputs to both go high. That will turn the P MOSFET off and the N MOSFET on, bringing the input end of L down to ground.

When the output of AND gate A1 had gone low at the start of the off time, it had two significant effects. First, the output of inverter $I3$ went high. Now element N_o is a positive logic NOR gate (any one input going high drives its output low). And element N_a is a positive logic NAND gate (its output is low only when all its inputs are high). Thus when the output of $I3$ went high, the N_o output went low and $I1$ output went high, causing P drive to go high to turn off the P MOSFET. And when the output of $I3$ went high, the output of N_a went low (all its other inputs were already high). That caused the N_a output to

go low, the $I2$ output to go high, and N drive to go high. That finally turned on the N MOSFET. The flip-flop thus remained locked in that set state with $Q1$ off and $Q2$ on because of the cross-coupling.

This off state remains until flip-flop FF2 is reset. The design requires a constant off time. This comes about as follows. When the $A1$ output went low, resetting FF1, the second significant result was that the anode of diode $D1$ went low, disconnecting it from timing capacitor C_t at pin 4. Now before the anode of $D1$ went low, that diode had clamped the voltage across $C4$ to a reference of 3 V and also forced that same voltage to the noninverting input to the off time comparator $T_{\text{off}\,c}$. But the inverting input of $T_{\text{off}\,c}$ is fixed at a threshold voltage V_{th1} (about 0.5 V). Thus the output of $T_{\text{off}\,c}$ had been high.

Then when the $D1$ anode went low, turning $Q1$ off and $Q2$ on and putting a reverse bias on $D1$, this started the off time. Now the capacitor C_t started discharging with a nominal current of about 0.25 mA. When it had discharged down to slightly below V_{th1}, the output of $T_{\text{off}\,c}$ went low and set FF2. This drive the output of $A1$ high again and via FF1 drove P drive and N drive low again, turning the P MOSFET on and the N MOSFET off, ending the off time.

Thus the off time is the time required to discharge C_t about 3 V with the internal 0.25 mA. This 0.25 mA is modulated somewhat internally to keep the frequency from going to low at low input voltage (Eq. 17.7). Then selection of C_t is as follows: Choose the desired switching frequency f, at nominal input voltage V_{in}. Then choose t_{off} from Eq. 17.8 for nominal input and output voltages (V_{in}, V_o), and choose C_t from

$$C_t = i\,\frac{dt}{dV} = \frac{0.0025\,t_{\text{off}}}{3}$$

17.2.9.4 LTC1148 line and load regulation. From the foregoing, it can be seen that line and load regulation occurs by changing the switching frequency by changing the on time for a fixed off time. But consider again in detail how the on time changes.

If V_{in} goes, say, high, V_o goes slightly high, the inverting input to voltage error amplifier G goes high, and its negative bias in series with the ramp voltage across R_{sense} decreases. Then the positive ramp at the noninverting terminal of C crosses the threshold at its inverting terminal earlier in time. Thus comparator C output goes positive earlier in time, the on time decreases, frequency increases as t_{off} is constant, and output voltage goes back up.

The same on time modulation occurs for load changes except that those changes are only temporary and then move back to the on time called for by Eq. 17.6. The temporary changes in the on time permit the step part of the ramp-on-a-step waveform (Fig. 1.6*a*) to build up or down over a number of switching cycles, as the current in the center of the ramps in Fig. 1.6*a* is the DC output.

17.2.9.5 LTC1148 peak current and output inductor selection. Because of internal details, the threshold voltage variation at the output of error amplifier G (Fig. 17.12*c*, pin 6) ranges between -0.025 and -0.15 V. Recall that the peak output inductor current is reached when the voltage across R_{sense} equals those threshold voltages. Then to consider tolerances, assume a maximum threshold of 0.100 V. For a specified maximum output current of I_{max}, R_{sense} is selected as

$$R_{sense} = \frac{0.100}{I_{max}} \tag{17.9}$$

The output inductor is selected so that it is at the end of the continuous mode (Sec. 1.3.6) at the minimum bias of 0.025 V. At the end of the continuous mode, the inductor on current is a triangle rising from 0 to a peak of $0.025/R_{sense}$ amperes in a time t_{on}. And that same peak current falls to zero in the t_{off} calculated from Eq. 17.8. Then

$$L = V\frac{dt}{dI} = \frac{V_o t_{off}}{0.025/R_{sense}} \tag{17.10}$$

17.2.9.6 LTC1148 burst-mode operation for low output current. At low output currents, as the output voltage starts rising, rather than permitting the inductor to go into the discontinuous mode, the circuit is designed to stop switching the inductor to V_{in} completely. Load current is then supplied entirely from the output capacitor. This discharges the output capacitor after a time, and when it falls back to its desired output voltage, switching commences again.

This is achieved with comparator BC. In normal operation, its output is high as its inverting input is below its noninverting one. This enables $A1$, and its output is then controlled by whatever is the state of the Q output of FF2. But when load current falls and V_o starts rising above its regulated value, the inverting terminal of BC rises above the reference at the noninverting terminal. The BC output goes low, the $A1$ output goes low, the $I3$ output goes high and forces (via N_o and $I1$) the P drive high to turn off $Q1$. And N drive is kept high to keep $Q2$ on via N_a and $I2$. But to lower the current drain in this nonswitching mode, as C_t falls below V_{TH2}, it turns off the N drive also and the circuit is in the *sleeping* mode.

Thus no switching occurs, and this off time persists even though C_t has discharged below V_{TH1}, for the low at the BC output keeps the $A1$ output low.

When the output capacitor has discharged back down to its regulated value, the BC inverting input goes low, its output goes high, it releases the inhibit on $A1$, and the circuit returns to its switching mode. This burst-mode operation and the fact that the P drive is high and N

drive is low in the sleep mode, reducing internal dissipation, permit high efficiency down to a load current level of one-thousandth its nominal value.

17.2.10 Summary of high-power Linear Technology buck regulators

The three buck regulators discussed above are typical and probably the most useful ones in the LTC buck family. There are numerous others, but they are specialized versions of the ones discussed. These others are higher-voltage ones, fixed-output-voltage types, and various lower-peak-current ones. They use the same block diagram and component selection as discussed in previous chapters (Secs. 1.3.6 and 1.3.7 for buck regulators, Secs. 15.4.6 and 15.4.7 for boost regulators).

Hence it is sufficient to provide only a tabular summary of the available LTC types. Such a summary is shown in Fig. 17.13.

17.2.11 Linear Technology micropower regulators

Battery-powered regulators with the all semiconductors inside one package for powering laptop computers represent a large segment of the market for low-input voltage regulators. Linear Technology has a wide variety of products for such use, but their topologies and circuitry are no different from what has been discussed above. Topologies are boost, buck, and polarity inverters; and they differ mainly only in that output currents are all under 1 A and most often they come in surface-mounted packages.

Since no new circuitry is involved, the description here will be confined to a tabular listing of the devices available. This is presented in Fig. 17.14*a* and *b*.

17.2.12 Feedback loop stabilization[3]

Feedback loop stabilization has been discussed mathematically in Chap. 12. There the significance of poles and zeros and their location along the frequency axis to stabilize the feedback loop was described.

But LTC prefers an empirical approach because the mathematical analysis depends on assumptions about various quantities which are not precisely specified by parts' manufacturers. Such quantities could be filter capacitor ESR or those which vary with use (unknown, user-added filter capacitance). LTC prefers to observe, on an oscilloscope, the power supply output voltage response to a step change in load current. By observing and optimizing the waveshape of the output voltage response to the step load for various potential *RC* combina-

High Efficiency 5V to 3.3V Step-Down Selection Guide

LOAD CURRENT	DEVICE	SYNCHRONOUS	SHUTDOWN CURRENT	EFFICIENCY
200mA to 400mA	LTC1174-3.3	NO	1μA	90%
200mA to 425mA	LTC1574-3.3	NO	2μA	90%
450mA	LTC1433	NO	15μA	93%
0.5mA to 2A	LTC1147-3.3	NO	10μA	92%
1A	LTC1265-3.3	NO	5μA	90%
1A to 5A	LTC1148-3.3	YES	10μA	94%
5A to 10A	LTC1266	YES	30μA	94%
5A to 20A	LTC1158	YES	2.2μA	91%

Step-Down Switching Regulator Selection Guide

	INPUT VOLTAGE		MAXIMUM	MAXIMUM RATED	PACKAGES	
DEVICE	MIN	MAX	SWITCH VOLTAGE	SWITCH CURRENT (A)	AVAILABLE	COMMENTS
LT1074	8	40	65	5.5	K, T, Y	Integrated 5A Switch
LT1074HV	8	60	75	5.5	K, T, Y	64V Max Input Voltage
LT1076	8	40	65	2.0	K, Q, R, T, Y	Integrated 2A Switch
LT1076HV	8	60	75	2.0	K, R, T, Y	64V Max Input Voltage
LTC1142	5	16	*	*	28-SSOP	Dual Output Synchronous Switching Controllers
LTC1142HV	5	20	*	*	28-SSOP	Dual Output Synchronous Switching Controllers
LTC1143	5	16	*	*	N16, S16	Dual 3.3V/5VLTC1147
LTC1147	4	16	*	*	N8, S8	Ext. P-Channel MOSFET, 90% Efficiency
LTC1148	4	16	*	*	N14, S14	Synchronous Switching, 93% Efficiency
LTC1148HV	4	20	*	*	N14, S14	Synchronous Switching, 93% Efficiency
LTC1149	5	60	*	*	N16, S16	Synchronous Switching, 48V Input
LTC1158**	5	30	*	*	N16, S16	Up to 15A at 90% Efficiency
LTC1159	4	40	*	*	N16, S16, G20	Synchronous Switching, 94% Efficiency
LT1160**	8.8	20	*	*	N14, S14	Up to 10A at 90% Efficiency
LTC1174	4	13.5	13.5	1	N8, S8	Integrated Switch, 90% Efficiency
LTC1174HV	4	18.5	18.5	1	N8, S8	Integrated Switch

Figure 17.13 Linear Technology buck regulators.

Step-Down Switching Regulator Selection Guide

DEVICE	INPUT VOLTAGE MIN	INPUT VOLTAGE MAX	MAXIMUM SWITCH VOLTAGE	MAXIMUM RATED SWITCH CURRENT (A)	PACKAGES AVAILABLE	COMMENTS
LT1176	8	38	50	1.2	N8, S20	Integrated 1.25A Switch, SO Package
LTC1265	3.5	13	V_{IN}-13	1.6	S14	Integrated Switch, 90% Efficiency
LTC1266	4	20	*	*	N16, S16	Synchronous Switching Uses N-Channel MOSFETs
LTC1267	4	40	*	*	28-SSOP	Dual Output Synchronous Switching Controller
LT1375/LT1376	6	25	30	1.5	N8, S8	500kHz Buck Switcher
LTC1430	4	9	*	*	S8, S16, N16	Synchronous Switching w/High Efficiency Up to 50A
LTC1435	3.5	36	*	*	S16	Constant Frequency w/Wide Input Voltage Range
LTC1436	3.5	36	*	*	24-SSOP	Constant Frequency, AUX Linear Regulator Controller
LTC1436-PLL	3.5	36	*	*	24-SSOP	Synchronizable Constant Frequency
LTC1437	3.5	36	*	*	28-SSOP	Synchronizable Constant Frequency, AUX Linear Regulator Controller
LTC1438	3.5	36	*	*	28-SSOP	Constant Frequency, Dual Output Synchronous Switching Regulator
LTC1439	3.5	36	*	*	36-SSOP	Constant Frequency, Dual Output Synchronous Switching Controller
LT1507	4	16	25	1.5	N8, S8	500kHz Buck Switcher, 5V to 3.3V
LTC1538	3.5	36	*	*	28-SSOP	LTC1438 with 5V Standby Active in Shutdown
LTC1539	3.5	36	*	*	36-SSOP	LTC1439 with 5V Standby Active in Shutdown
LTC1574	4.5	18.5	18.5	*	S16	1A Schottky Diode in Package

* Uses external FET switch

** Half-bridge driver requires external control circuit

Figure 17.13 (*Continued*)

Battery-Powered DC/DC Conversion Solutions Selection Guide

The following tables are a short form component selection guide for a collection of commonly used battery-powered DC/DC conversion applications. No design is required since inductor, capacitor and resistor values are completely specified.

Choose the appropriate LTC DC/DC converter for your application from the following tables.

The LT1073, LT1107, LT1108, LT1110, LT1111, LT1173, LTC1174 and LT1303, LT1307 all have low-battery detection capability.

Step-Up From One Cell (1V)

V_{out} (V)	I_{out} (mA)	DEVICE	I_Q (μA)	L (μH)	C (μF)	R (Ω)	Fig	COMMENTS
3.3V	75	LT1307	50	10	10	–	4	
5	40	LT1073-5	95	82	100	0	1	Lowest I_Q
	40	LT1110-5	350	27	33	0	1	Best For Surface Mount
12	40	LT1307	50	10	10	–	4	
	15	LT1073-12	95	82	100	0	1	Lowest I_Q
	15	LT1110-12	350	27	33	0	1	Best For Surface Mount

Adjustable versions also available for V_{out} up to 50V

Basic Step-Up Converters

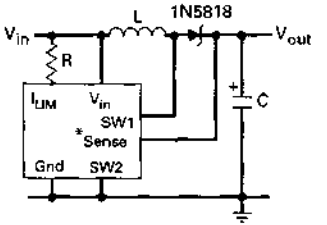
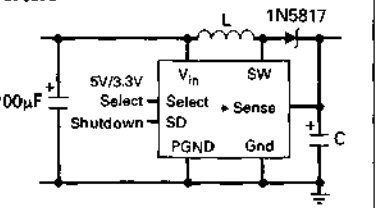


Figure 1



* See tables for recommended part, inductor, capacitor and resistor values

Figure 2

Figure 17.14 (*a*) Linear Technology micropower (I_o under 1 A) boost regulators.

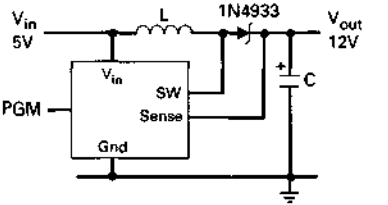
Step-Up From Two Cell (2V)

V_{out} (V)	I_{out} (mA)	DEVICE	IQ (μA)	L (μH)	C (μF)	R (Ω)	Fig	COMMENTS
3.3	400	LT1300**	120	10	100	–	2	Selectable 3.3V/5V Out
	250	LT1500	200	22	220	–	2	Low Spectral Noise
5	90	LT1173-5	110	47	100	47	1	Lowest I_Q
		LT1111-5	300	18	33	47	1	Surface Mount
	150	LT1107-5	300	33	33	47	1	Surface Mount
		LT1108-5	110	100	100	47	1	Lowest I_Q
	200	LT1304	120	22	100	–	2	Low-Battery Detector
		LT1501	200	22	220	–	2	Low Spectral Noise
	220	LT1300**	120	10	100	–	2	Selectable 3.3V/5V Out
		LT1301**	120	10	100	–	2	Selectable 5V/12V Out
	400	LT1305	120	10	220	–	2	High Output Power
	600	LT1302	200	10	100	–	*	Highest Power Output
12	20	LT1173-12	110	47	47	47	1	Lowest I_Q
		LT1111-12	300	18	22	47	1	Surface Mount
	40	LT1107-12	300	27	33	47	1	Surface Mount
		LT1108-12	110	82	100	47	1	Lowest I_Q
	50	LT1301**	120	10	100	–	2	Selectable 5V/12V Out
	120	LT1302	200	3.3	66	–	*	Highest Output Power

*See LT1302 data sheet **For low-battery detection use LT1303 or LT1304

Figure 17.14 (*a*) (*Continued*)

Flach Memory VPP Generator



* See tables for recommended inductor and capacitor values

Figure 3

Battery-Powered DC/DC Conversion Solutions Selection Guide (Continued)

Step-Up From 5V to 12V

V_{in} (V)	I_{out} (mA)	DEVICE	I_Q (μA)	L (μH)	C (μF)	R (Ω)	Fig	COMMENTS
12	90	LT1173-12	110	120	100	0	1	Lowest I_Q
		LT1111-12	300	47	33	0	1	Surface Mount
	175	LT1107-12	300	60	32	0	1	Surface Mount
		LT1108-12	110	180	100	0	1	Lowest I_Q
	200	LT1301**	120	33	47	–	2	True Shutdown
		LT1500	200	22	220	–	2	Low Spectral Noise
	350	LT1373	1000	22	100	–	4	250kHz

** For low-battery detection use LT1303

Step-Up From 5V to 12V

V_{in} (V)	I_{out} (mA)	I_{out} (mA)	DEVICE	I_Q (μA)	L (μH)	C (μF)	Fig	COMMENTS
3 or 5	12	60	LT1106	750	10	1	***	TSSOP
			LT1309	500	10	1	3	Small, Shutdown
5	12	60	LT1109-12	320	33	22	3	Small, SMT
		120	LT1109A-12	320	27	47	3	Small, SMT
		200	LT1301**	120	27	47	2	True Shutdown
		60	LT1109A-12	320	10	22	1	All Surface Mount
2 Cells	12	80	LT1301**	120	10	47	2	True Shutdown

** For low-battery detection use LT1303 ***See LT1106 data sheet

Figure 17.14 (*a*) (*Continued*)

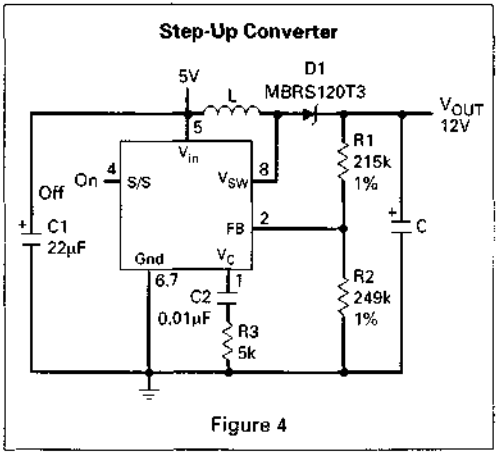


Figure 4

Step-Down Conversion to 3.3V

V_{in} (V)	I_{out} (mA)	DEVICE	I_Q (μA)	L (μH)	C (μF)	IPGM	Fig	COMMENTS
4.5 to	200	LTC1174-3.3	450	50	2 x 33	To Gnd	6	Low Dropout
12.5	425	LTC1174-3.3	450	50	2 x 33	To V_{in}	6	Surface Mount
4.5 to	200	LTC1574-3.3	600	50	2 x 33	To V_{in}	6	Low Component Count
16	425	LTC1574-3.3	600	50	2 x 33	To V_{in}	6	Low Component Count
5 to	2A	LTC1148-3.3	160	–	–		–	See Step Down
16						–		Switchers
5 to	450	LTC1433	470	100	100		6	Constant Frequency
12						–		
12 to	2A	LTC1149-3.3	600	–	–		–	See Step Down
60						–		Switchers

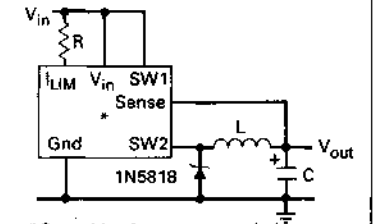
Step-Down Conversion to 5V

5.5 to	200	LTC1174-5	450	100	2 X 33	To Gnd	6	Low Dropout,
12	400	LTC1174-5	450	100	2 X 33	To V_{in}	6	Surface Mount
5.5 to	200	LTC1574-5	600	100	2 X 33	To V_{in}	6	Low Component Count
16	425	LTC1574-5	600	100	2 X 33	To V_{in}	6	Low Component Count
5.8 to	450	LTC1433	470	100	100	–	6	Constant Frequency
12								
12 to	300	LT1107-5	300	60	100	100	5	Surface Mount
20	300	LT1108-5	110	180	330	100	5	Lowest IQ
20 to	300	LT1173-5	110	470	470	100	5	Lowest IQ
30	300	LT1111-5	300	180	220	100	5	Surface Mount
6 to	2A +	LTC1147/8-5	160	–	–	–	–	See Step Down
16								Switchers
12 to	2A +	LTC1149-5	600		–		–	See Step Down
60								Switchers

Adjustable output voltages up to 6.2V can be obtained with the adjustable versions If LT1173. LT1111, LT1107, LT1108, or LT1110.

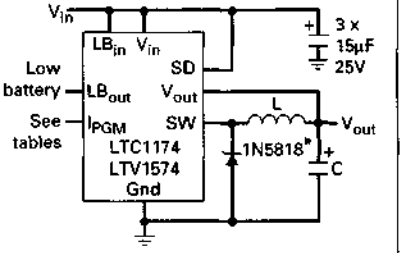
Figure 17.14 (*b*) Linear Technology micropower buck regulators.

Step-Down Converters



*See tables for recommended part inductor capacitor and resistor values

Figure 5



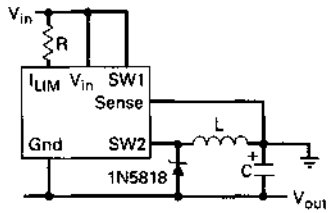
* Diode not required for LTC1574

Figure 6

Positive-to-Negative Voltage Conversion

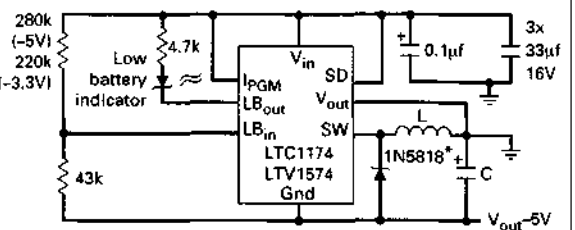
V_{in} (V)	V_{out} (V)	I_{out} (mA)	DEVICE	I_Q (μA)	L (μH)	C (μF)	R (Ω)	Fig	COMMENTS
5	–5	75	LT1108-5	110	100	100	100	7	Lowest I_Q
			LT1107-5	300	33	33	100	7	Surface Mount
		150	LTC1174-5	450	50	2 x 33	–	8	Surface Mount
		290	LTC1433	470	68	100	–	8	Constant Frequency
12	–5	250	LT1173-5	110	470	220	100	7	Lowest I_Q
		250	LT1111-5	300	180	82	100	7	Surface Mount
1	–5	110	LTC1574-5	450	50	100	–	8	SMT, No Ext.
8		170							Schottky Diode
12.5		235							Required

Positive-to-Negative Converters



*See tables for recommended part inductor capacitor and resistor values

Figure 7



* Diode not required for LTC1574

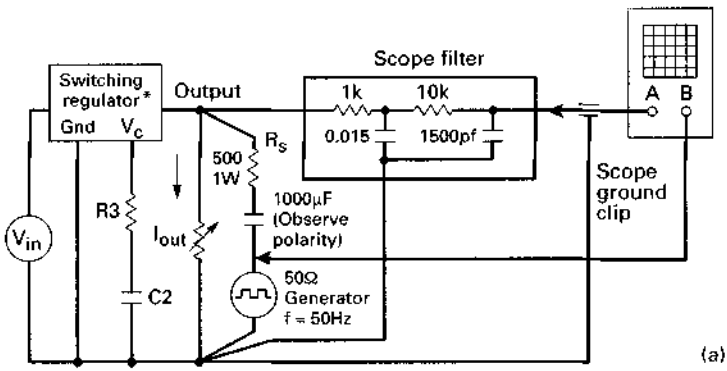
Figure 8

Figure 17.14 (*b*) (*Continued*)

tions (zeros) at the error amplifier output, the loop is stabilized. There is now no ambiguity as to the validity of any assumptions.

All LTC boost, buck, and inverting configurations shown in their data sheets are so stabilized. Their scheme discussed here comes from its description by C. Nelson and J. Williams in LTC Application Note 19.

The stabilization scheme is shown in Fig. 17.15*a*. There with the nominal DC load at the output, an additional step load (Fig. 17.15*b*) of about 10 percent ($R1$) is AC-coupled to the output through a large capacitor $C1$. The step frequency is not critical (usually 50 Hz) and is generated by any square wave generator having a low source impedance. The regula-



*All input and output filters must be in place. Input source (V_{in}) must be actual source used in final design to account for finite source impedance.

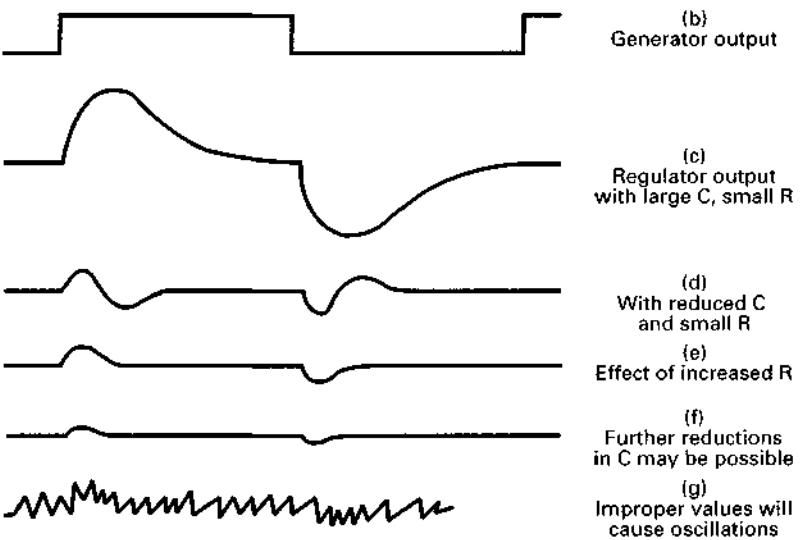


Figure 17.15 Stabilizing the feedback loop by varying R3C2 product to optimize the transient response waveshape when subjected to a steploaded current via R_sC_b.

tor transient response to the step is observed by the oscilloscope through the filter shown to keep switching frequencies out of the display.

The series RC network used to stabilize the supply is connected from the output of the internal voltage error amplifier (V_c pin) to ground. Initially, a large $C2$ (greater than 2 mF) and relatively small $R3$ (about 1 kΩ) is used. This almost always yields a stable DC loop. But with $C2$ so large, the supply responds to the step load with large transient overshoots and a slow decay back to the DC output, as seen in Fig. 17.15*c*. Now $C2$ is decreased in steps yielding the response of Fig. 17.15*d*. The overshoots are smaller in amplitude and fall back more quickly to the quiescent level. After the overshoot, there is a reverse-polarity *ring*. Now if $R3$ is increased, the waveshape of Fig. 17.15*e* with the reverse ring eliminated results. Now decreasing C may decrease the amplitude of the overshoots. More details are provided in the aforementioned LTC Application Note 19.

In terms of the feedback analysis of Chap. 12, the combination of $C2$ and $R3$ provides only a single zero, but no pole in the amplifier transfer function (Secs. 12.3 and 12.19). Recall that (Sec. 12.3) the pole (a shunt capacitor across the series RC combination) was added to reduce the gain at high frequencies, so that any high-frequency noise spikes picked up would not get through to the output. Such noise spikes can come from external noise sources or simply from internal switching noise sources.

But the LTC analysis and data sheets show no such poles or shunt C across the series RC combination. This may be due to the fact that there is already some shunt C from V_c to ground internally, or that in these relatively low-power devices with no large dV/dt or di/dt sources, there are no such noise spikes. If spikes are observed at the outputs, they can very likely be eliminated by empirically selected small capacitors from V_c to ground.

17.3 Maxim IC Regulators

Maxim Inc. is the one other major manufacturer of all the devices discussed here. Its product line also consists of boost, buck, and polarity inverters and matches many of the LTC uses in maximum voltage and current ratings. Generally, though, they tend to specialize in devices with lower current ratings than those of LTC devices.

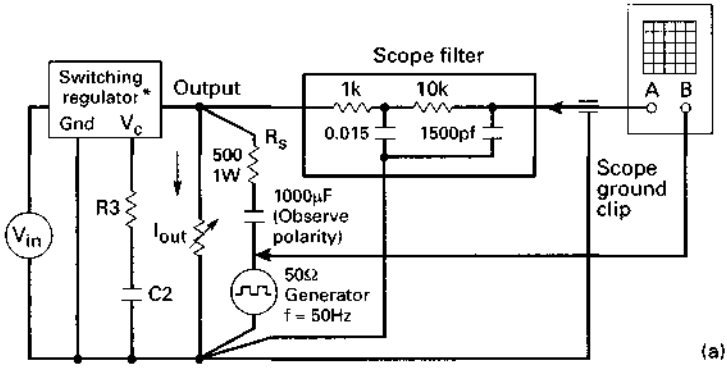
Since no new circuit techniques are involved, the description here is also limited to a tabular listing of the devices, showing all their significant voltage and current ratings. This is shown in Fig. 17.16.

But the fact that most of the discussion here has concerned LTC devices, with only a tabular presentation of Maxim types, should not be taken as the author's assessment of the relative merits of the two

tions (zeros) at the error amplifier output, the loop is stabilized. There is now no ambiguity as to the validity of any assumptions.

All LTC boost, buck, and inverting configurations shown in their data sheets are so stabilized. Their scheme discussed here comes from its description by C. Nelson and J. Williams in LTC Application Note 19.

The stabilization scheme is shown in Fig. 17.15*a*. There with the nominal DC load at the output, an additional step load (Fig. 17.15*b*) of about 10 percent ($R1$) is AC-coupled to the output through a large capacitor $C1$. The step frequency is not critical (usually 50 Hz) and is generated by any square wave generator having a low source impedance. The regula-



*All input and output filters must be in place. Input source (V_{in}) must be actual source used in final design to account for finite source impedance.

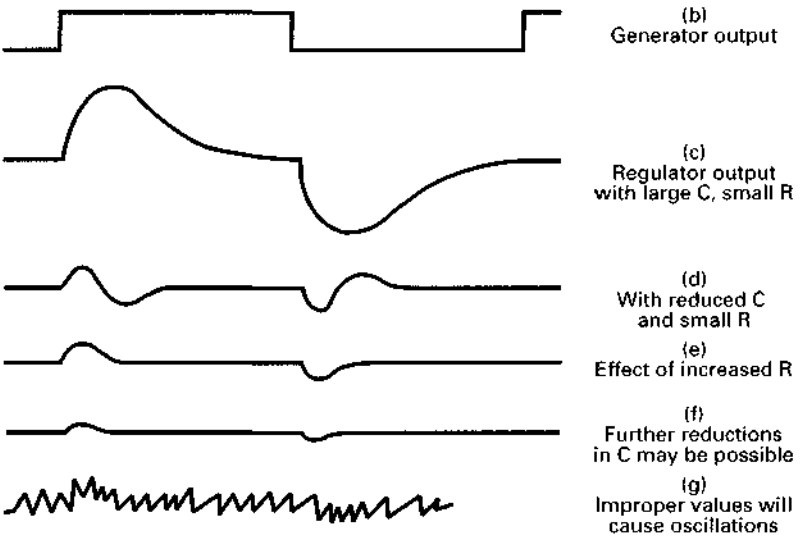


Figure 17.15 Stabilizing the feedback loop by varying R3C2 product to optimize the transient response waveshape when subjected to a stepload current via R_sC_b.

tor transient response to the step is observed by the oscilloscope through the filter shown to keep switching frequencies out of the display.

The series RC network used to stabilize the supply is connected from the output of the internal voltage error amplifier (V_c pin) to ground. Initially, a large $C2$ (greater than 2 mF) and relatively small $R3$ (about 1 kΩ) is used. This almost always yields a stable DC loop. But with $C2$ so large, the supply responds to the step load with large transient overshoots and a slow decay back to the DC output, as seen in Fig. 17.15*c*. Now $C2$ is decreased in steps yielding the response of Fig. 17.15*d*. The overshoots are smaller in amplitude and fall back more quickly to the quiescent level. After the overshoot, there is a reverse-polarity *ring*. Now if $R3$ is increased, the waveshape of Fig. 17.15*e* with the reverse ring eliminated results. Now decreasing C may decrease the amplitude of the overshoots. More details are provided in the aforementioned LTC Application Note 19.

In terms of the feedback analysis of Chap. 12, the combination of $C2$ and $R3$ provides only a single zero, but no pole in the amplifier transfer function (Secs. 12.3 and 12.19). Recall that (Sec. 12.3) the pole (a shunt capacitor across the series RC combination) was added to reduce the gain at high frequencies, so that any high-frequency noise spikes picked up would not get through to the output. Such noise spikes can come from external noise sources or simply from internal switching noise sources.

But the LTC analysis and data sheets show no such poles or shunt C across the series RC combination. This may be due to the fact that there is already some shunt C from V_c to ground internally, or that in these relatively low-power devices with no large dV/dt or di/dt sources, there are no such noise spikes. If spikes are observed at the outputs, they can very likely be eliminated by empirically selected small capacitors from V_c to ground.

17.3 Maxim IC Regulators

Maxim Inc. is the one other major manufacturer of all the devices discussed here. Its product line also consists of boost, buck, and polarity inverters and matches many of the LTC uses in maximum voltage and current ratings. Generally, though, they tend to specialize in devices with lower current ratings than those of LTC devices.

Since no new circuit techniques are involved, the description here is also limited to a tabular listing of the devices, showing all their significant voltage and current ratings. This is shown in Fig. 17.16.

But the fact that most of the discussion here has concerned LTC devices, with only a tabular presentation of Maxim types, should not be taken as the author's assessment of the relative merits of the two

Part Number	Input Voltage Range (V)	Output Voltage (V)	Supply Current (mA), max (typ)	Output (mA typ)	Control Scheme	Package Options *	EV Kit	Temp. Ranges **	Features	Price † 1000-up ($)
STEP-UP/STEP-DOWN SWITCHING REGULATORS										
MAX877/878/879	1 to 6.2	5/(3.3 or 3)/Adj.	0.310(0.220)	240	PFM	DIP,SO	Yes	C,E,M.	Gives regulated output when input above and below the output; no transformer	2.95
STEP-UP SWITCHING REGULATORS										
MAX4193	2.4 to 16.5	Adj.	0.200(0.090)	300mW	PFM	DIP,SO		C,E,M.	Improved RC 4193 2nd source	1.74
MAX630	2 to 16.5	Adj.	0.125(0.070)	300mW	PFM	DIP,SO		C,E,M.	Improved RC 4193 2nd source	2.88
MAX631	1.5 to 5.6	5, Adj.	0.4(0.135)	40	PFM	DIP,SO		C,E,M.	Only 2 external components	2.56
MAX632	1.5 to 12.6	12, Adj.	2.0(0.5)	25	PFM	DIP,SO		C,E,M.	Only 2 external components	2.56
MAX633	1.5 to 15.6	15, Adj.	2.5(0.75)	20	PFM	DIP,SO		C,E,M.	Only 2 external components	2.56
MAX641	1.5 to 5.6	5, Adj.	0.4(0.135)	300	PFM	DIP,SO		C,E,M.	PFM controller	2.87
MAX642	1.5 to 12.6	12, Adj.	2.0(0.5)	550	PFM	DIP,SO		C,E,M.	PFM controller	2.87
MAX643	1.5 to 15.6	15, Adj.	2.5(0.75)	325	PFM	DIP,SO		C,E,M.	PFM controller	2.87
MAX731	1.8 to 5.25	5	4(2)	200	PWM	DIP,SO	Yes	C,E,M.		2.60
MAX732	4 to 9.3	12	3(1.7)	200	PWM	DIP,SO	Yes	C,E,M.	Flash memory programmer, ± 4% output voltage tolerance	2.76
MAX733	4 to 11	15	3(1.7)	125	PWM	DIP,SO	Yes	C,E,M.		2.60
MAX734	1.9 to 12	12	2.5(1.2)	120	PWM	DIP,SO	Yes	C,E,M.	Flash memory programmer	2.23
MAX741U	1.8 to 15.5	5, 12, 15, Adj.	3.5(1.6)	5W	PWM	DIP,SSOP	Yes	C,E,M.	PWM step-up controller, $3V_{in}$ to $5V_{out}$ at 1A, 85% efficient	3.64
MAX751	1.2 to 5.25	5	3.5(2)	175	PWM	DIP,SO	Yes	C,E,M.		2.35
MAX752	1.8 to 16	Adj.	3(1.7)	2.4W	PWM	DIP,SO	Yes	C,E,M.		2.94
MAX756/757	1.1to 5.5	(3.3 or 5)/Adj.	0.060(0.045)	250	PFM	DIP,SO	Yes	C,E	Best combination of low I_Q & high 86% efficiency	1.95
MAX761/762	2 to 16.5	12/15 or Adj. to 16.5	0.1(0.080)	120	PFM	DIP,SO	Yes	C,E,M.	12V flash programmer, high efficiency over wide I_{out} range	2.23
MAX770/771/772	2 to 16.5	5/12/15 or Adj.	0.1(0.085)	1A	PFM	DIP,SO	Yes	C,E,M.	Controllers, high efficiency over wide I_{out} range	1.80
MAX1771	2 to 16.5	12 or Adj.	0.1(0.085)	1A	PFM	DIP,SO	Yes	C,E,M.	Same as MAX771, but only 100mV current-sense limit	1.80
MAX773	2 to 16.5	Adj. to 48	0.1(0.085)	1A	PFM	DIP,SO	Yes	C,E,M.	Controller, high-voltage output, high efficiency over wide I_{out} range	1.80
MAX777/778/779	1 to 6	5/(3 or 3.3)/Adj.	0.31(0.220)	300	PFM	DIP,SO	Yes	C,E,M.	On-chip active diode, true turn off in shutdown	2.65
MAX856/857	0.8 to 6	5/(3.3 or 5)/Adj.	0.060(0.025)	100	PFM	DIP,SO,μMAX	Yes	C,E	Best combination of low I_Q & high 85% efficiency	1.72
MAX858/859	0.8 to 6	5/(3.3 or 5)/Adj.	0.060(0.025)	25	PFM	DIP,SO,μMAX	Yes	C,E	Small, best combination of low I_Q & high efficiency	1.72

Figure 17.16 Maxim IC regulators. Devices similar to Linear Technology regulators from the other major supplier of such IC regulators requiring only a minimum number of components external to the package.

Part Number	Input Voltage Range (V)	Output Voltage (V)	Quiescent Supply Current (mA), max (typ)	Output (mA typ)	Control Scheme	Package Options [a]	EV Kit	Temp. Ranges [b,c]	Features	Price [†] 1000-up ($)
STEP-DOWN SWITCHING REGULATORS										
MAX638	2.6 to 16.5	5, Adj.	0.6(0.135)	75	PFM	DIP,SO		C,E,M.	Only 3 external components	2.56
MAX639/640/653	4 to 11.5	5/3.3/3 or Adj.	0.02(0.01)	225	PFM	DIP,SO	Yes	C,E,M.	>90% efficiencies over wide I_{out} range (1mA to 225mA)	2.96
MAX649/651/652	4 to 16.5	5/3.3/3 or Adj.	0.1(0.080)	2A	PFM	DIP,SO	Yes	C,E,M.	>90% efficiencies over wide I_{out} range, drives external P-channel FET	1.60
MAX649/1651	4 to 16.5	5/3.3/3 or Adj.	0.1(0.080)	2A	PFM	DIP,SO	Yes	C,E,M.	Same as MAX649/651 but 96.5% duty cycle and only 100mV current sense limit	1.60
MAX724/724H	3.5 to 40/60	Adj. (2.5 to 40)	12(8.5)	5A	PWM	TO/220, TO-3		C,E,M.	High power, few external components	4.52/6.83
MAX726	3.5 to 40/60	Adj. (2.5 to 40)	12(8.5)	2A	PWM	TO/220, TO-3		C,E,M.	High power, few external components	3.00
MAX727	3.5 to 40/60	5	12(8.5)	2A	PWM	TO/220, TO-3		C,E,M.	High power, few external components	3.00
MAX728	3.5 to 40/60	3.3	12(8.5)	2A	PWM	TO/220, TO-3		C,E,M.	High power, few external components	3.00
MAX729	3.5 to 40/60	3	12(8.5)	2A	PWM	TO/220, TO-3		C,E,M.	High power, few external components	3.00
MAX730A	5.2 to 11	5	3(1.7)	300	PWM	DIP,SO	Yes	C,E,M.	90% efficiencies, no subharmonic switching noise	2.15
MAX738A	6 to 16	5	3(1.7)	750	PWM	DIP,SO	Yes	C,E,M.	>85% efficiencies, no subharmonic switching noise	2.60
MAX741D	2.7 to 15.5	5, Adj.	4.25(2.8)	3A	PWM	DIP,SSOP	Yes	C,E,M.	PWM step-down controller, 6.5V_{in} to 5V_{out} at 3A, 90% efficient	3.74
MAX744A	4.75 to 16	5	2.5(1.2)	750	PWM	DIP,SO	Yes	C,E,M.	Optimized cellular communications, no subharmonic switching noise	2.90
MAX746	4 to 15	5/Adj.	1	2.5A	PWM	DIP,SO	Yes	C,E,M.	5V to 3.3V green PC apps., drives external N-channel FET	2.25
MAX747	4 to 15	5/Adj.	1.3(0.8)	2.5A	PWM	DIP,SO	Yes	C,E,M.	5V to 3.3V green PC apps., drives external N-channel FET	2.25
MAX748A	3.3 to 16	3.3	3(1.7)	750	PWM	DIP,SO	Yes	C,E,M.	>85% efficiencies, no subharmonic switching noise	2.60
MAX750A	4 to 11	Adj.	3(1.7)	1.5W	PWM	DIP,SO	Yes	C,E,M.	90% efficiencies, no subharmonic switching noise	2.15
MAX758A	4 to 16	Adj.	3(1.7)	3.75W	PWM	DIP,SO	Yes	C,E,M.	>85% efficiencies, no subharmonic switching noise	2.60
MAX763A	3.3 to 11	3.3	2.5(1.4)	500	PWM	DIP,SO	Yes	C,E,M.	>85% efficiencies, no subharmonic switching noise	2.15
MAX767	4.5 to 5.5	3.3, 3.45(R, or 3.6(S)	0.75	7A	PWM	SSOP	Yes	C,E	Green PC apps., high efficiency, small-size controller	3.49
MAX787/787H	3.5 to 40/60	5	12(8.5)	5A	PWM	TO-220, TO-3		C,E,M.	High power, few external components	††
MAX788/788H	3.5 to 40/60	3.3	12(8.5)	5A	PWM	TO-220, TO-3		C,E,M.	High power, few external components	††
MAX789/789H	3.5 to 40/60	3	12(8.5)	5A	PWM	TO-220, TO-3		C,E,M.	High power, few external components	††
MAX796-799	4.5 to 30	5.05/3.3/2.9 Adj.	1(0.7)	50W	PWM	DIP,SO	Yes	C,E,M.	Synchronous rectifier, secondary output regulation, Idle-mode PWM, high efficiency	††
MAX830-833	3.5 to 40	Adj./5/3/3.3	11(8)	1A	PWM	SO	Yes (MAX831)	C	High power, SOIC	††

Figure 17.16 (*Continued*)

Part Number	Input Voltage Range (V)	Output Voltage (V)	Quiescent Supply Current (mA), max (typ)	Output (mA typ)	Control Scheme	Package Options *	EV Kit	Temp. Ranges **	Features	Price † 1000-up ($)
INVERTING SWITCHING REGULATORS										
MAX4391	4 to 16.5	up to –20	0.25(0.09)	400mW	PFM	DIP,SO		C,E,M,	Improved RC4391 2nd source	2.09
MAX634	2.3 to 16.5	up to –20	0.15(0.07)	400mW	PFM	DIP,SO		C,E,M,	Improved RC4391 2nd source	2.61
MAX635	2.3 to 16.5	–5, Adj.	0.15(0.08)	50	PFM	DIP,SO		C,E,M,	Only 3 external components	2.56
MAX636	2.3 to 16.5	–12 Adj.	0.15(0.08)	40	PFM	DIP,SO		C,E,M,	Only 3 external components	2.56
MAX637	2.3 to 16.5	–15 Adj.	0.15(0.07)	25	PFM	DIP,SO		C,E,M,	Only 3 external components	2.56
MAX650	–54 to –42	5	10(0.5)	250	PFM	DIP,SO		C,E,M,	Telecom applications	3.50
MAX735	4 to 6.2	–5	3(1.6)	275	PWM	DIP SO		C,E,M,	>80% efficiencies	2.15
MAX736	4 to 8.6	–12	3(1.6)	125	PWM	DIP,SO	Yes	C,E,M,	>80% efficiencies	2.75
MAX737	4 to 5.5	–15	4.5(2.5)	100	PWM	DIP,SO	Yes	C,E,M,	>80% efficiencies	2.75
MAX739	4 to 15	–5	3(1.6)	500	PWM	DIP,SO	Yes	C,E,M,	>80% efficiencies	2.75
MAX741N	2.7 to 15.5	–5, –12, –15, Adj.	4.0(2.2)	5W	PWM	DIP,SSOP		C,E,M,	PWM inverting controller, high efficiency	3.64
MAX749	2 to 6	Adj	0.06	5W	PFM	DIP,SO	Yes	C,E,M,	Digital adjust for negative LCD	2.49
MAX755	2.7 to 9	Adj.	3.5(1.8)	1.4W	PWM	DIP,SO		C,E,M,	80% efficiencies	2.15
MAX759	4 to 15	Adj.	4(2.1)	1.5W	PWM	DIP,SO	Yes	C,E,M,	LCD driver, >80% efficiencies	2.75
MAX764/765/766	3 to 16.5	–5,/–12/–15 or Adj. to 21ΔV	0.1	200	PFM	DIP,SO	Yes	C,E,M,	High efficiency over wide I_{out} range	2.38
MAX774/775/776	3 to 16.5	–5/–12/–15 or Adj.	0.1	1A	PFM	DIP,SO	Yes	C,E,M,	Controllers, high efficiency over wide I_{out} range	2.20

* Package Options: DIP = Dual-In-Line Package, SO = Small Outline, SSOP = Shrink Small-Outline Package, TO___ = Can

** Temperature Ranges: C = 0°C to +70°C, E = 40°C to +85°C, M = –55°C to +125°C

† Prices provided are for design guidance and are FOB USA. International prices will differ due to local duties, taxes, and exchange rates.

†† Future product—contact factory for pricing and availability. Specifications are preliminary.

Figure 17.16 *(Continued)*

manufacturers' devices. This discrepancy is due only to the author's earlier familiarity with LTC products.

17.4 Distributed Power Systems with IC Building Blocks[7]

Figure 17.17*a* shows a conventional off-line, multi-output power supply. The alternating current is rectified (with or without power factor correction), and some topology—half, full bridge, forward converter, or flyback—is used to generate a precisely controlled *master* output voltage on output ground. This master is generally the highest current output—usually +5 V—and is very well regulated against line and load current changes. It is very well regulated because of the feedback loop around it which controls the on time of the power switching devices on input ground.

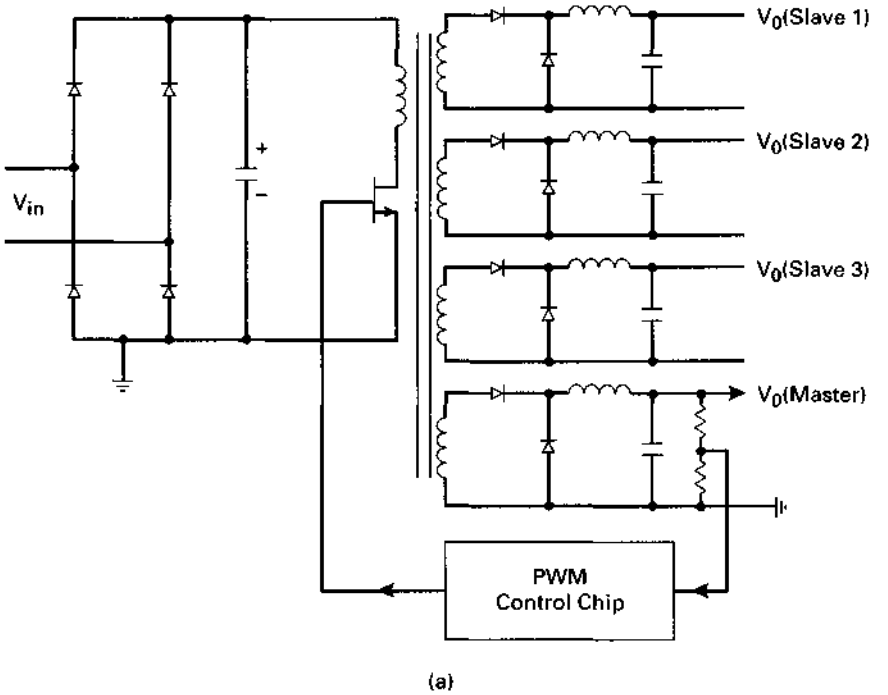
Additional outputs (referred to as *slaves*) are obtained by adding more secondaries whose turns are selected to yield the desired secondary DC voltages. Since the time the slave secondary outputs are high is the same as that of the master secondary, the DC voltages of the slaves after their *LC* filters are also very well regulated against line input voltage changes.

But slaves are not well regulated against load current changes in either the master or the slave outputs (Secs. 2.2.1 to 2.2.3). Slave regulation against line and load changes is generally only as good as ±5 to 8 percent and may change as much as 50 percent if either the master or slave inductors are permitted to go into discontinuous mode (Sec. 2.2.4). Further, the slave absolute DC voltages cannot be set precisely, as the smallest amount by which they can be changed is that corresponding to adding or removing a single turn. Since volts per turn is proportional to the switching frequency from Faraday's law $[E/N = A_e\ (dB/t_{on})(10^{-8}) = A_e(dB \times 0.4f)(10^{-8})]$ at high frequencies, E/N can amount to 2 to 3 V per turn depending on the flux change and core iron area.

These DC slave voltages are, of course, generated physically close to the main power transformer and are piped around to their point of use.

But generally the poor regulation of and the inability to set accurately the absolute DC voltage of the slaves are not too much of a problem. Slaves are used to power operational amplifiers or motors for various computer peripherals, and these can tolerate large output voltage changes.

But in cases where slaves must be accurate and well regulated, they must be controlled by a separate feedback loop around them. This often is done by following the poorly regulated slave with either a linear regulator for low current, a buck regulator for higher current, or a magnetic amplifier (Sec. 10.3).



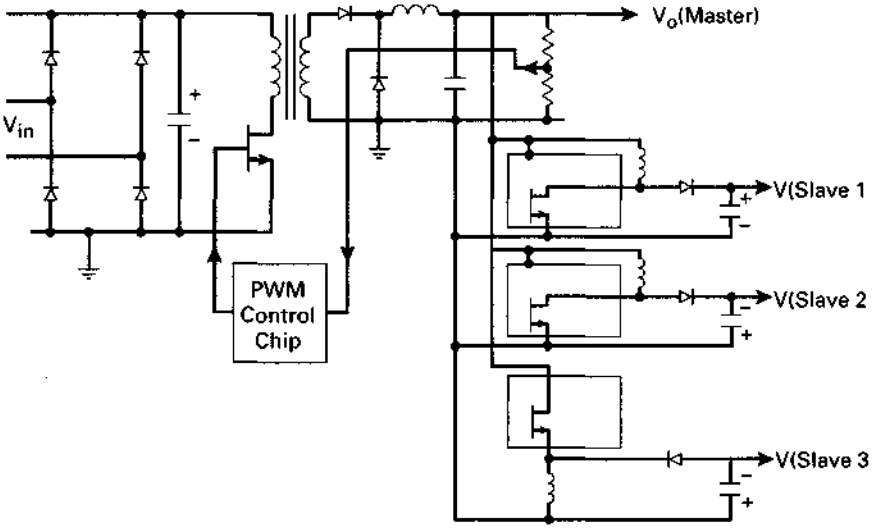
(a)

Figure 17.17 (*a*) Conventional scheme for building a multioutput power supply. A feedback loop around a master (highest current output) regulates it against line and load changes. Secondaries on the power transformer yield slaves which are well regulated against line changes but only ±8 percent for load changes.

Distributed power is an alternative which corrects this problem of inaccurate slaves, and it has additional significant advantages, as shown in one of its versions in Fig. 17.17*b*.

The essence of distributed power is that it generates a common DC voltage (not necessarily well regulated) at a central point and buses it around to the points of use. There, standardized well-regulated DC/DC converters—buck, boost, or polarity inverters—convert it to the desired voltage. The availability of the above-described LTC and Maxim standard regulators makes them immediate candidates for such a distributed power scheme.

Consider the following advantages of one version of distributed power shown in Fig. 17.17*b*. In that figure, the highest current output, usually +5 V, is generated directly by the main power transformer which has only the one secondary shown. It is controlled by a feedback loop from a sample on the output ground around to control the on time of the power transistor on the input ground. All other outputs are derived from boost or polarity inverters, each having their own feedback



(b)

Figure 17.17 (*Continued*) (*b*) Distributed power scheme. Here the transformer has only one secondary which is regulated very well against line and load changes. That output is bussed around and generates the slave voltages at the point of use through standardized DC/DC converters operating either as boost, buck, or polarity inverters. The master may not even require regulation by PWM control of the primary side power transistor. That transistor may operate at a fixed on time, and the master regulation may be derived from its own DC/DC converter for sufficiently low output currents.

loop. In the figure, all the slaves are generated by LTC1174 boost regulators. The advantages are as follows:

1. Simpler and less expensive main power transformer. That transformer is generally the largest and most expensive element in a switching power supply. With fewer secondary windings, it is easier to meet VDE safety specifications.

2. Ease of changing electrical parameters of the transformer. In the conventional scheme of Fig. 17.17*a*, the initial transformer design often requires repeated versions. Some windings may require adding or removing turns. Leakage or magnetizing inductance may be improper in an initial prototype. Winding sequence may have to be changed to improve coupling between windings or reduce proximity-effect losses.

In the scheme of Fig. 17.17*b*, each output has its own feedback loop, and the exact input voltage is unimportant as the output voltages are constant generally over a 3:1 input voltage range.

3. Ease of changing output voltages or currents without changing the main transformer design.

4. Ease of adding one or a number of new output voltages. Often in a large system design, it is found late in the design that some new voltages must be added.

5. Possibility of totally eliminating feedback from one master on output ground around to controlling the pulse width on input ground. This avoids all the problems of sensing an output voltage on output ground and controlling a pulse width on input ground.

It thus makes unnecessary optocouplers with their troublesome gain variation with temperature, low-power housekeeping supplies on output ground, and a small power transformer to couple a pulse on output ground around to the power transistor on input ground.

All this becomes possible by generating the usual +5-V high current output also entirely from a secondary standard buck converter. Thus, instead of generating +5 V directly from a secondary winding, generate an unregulated voltage of about +20 to +24 V and bus this around to the point of use. There the lower-current slaves could be bucked down to their desired value with LT1074-type buck regulators. The higher-current 5 or 3.3 V could be bucked down from +24 V by high-efficiency LT1270A or LTC1159 regulators.

This scheme using only secondary side regulation would not need pulse width modulation of the power transistor. It would just require setting the controller pulse width to about a constant 85 percent of a half period and using peak rectification with only a capacitor filter in the single secondary. Variation of the peak rectified DC voltage due to line and load changes and ripple at the filter capacitor would be taken care of by the secondary regulators.

In general, in any such distributed power scheme, it is best to choose a relatively high bus voltage such as +20 to +25 V and buck it down to the desired values, rather than a low bus of +5 V and boost it up to the desired voltages. The only time a low bus of +5 V makes sense is when the +5-V current is high—in the vicinity of, say, 10 to 100 A—as such high currents are not usually from DC/DC converters.

Although a distributed power scheme may be more expensive and may dissipate somewhat more power (since the power is handled twice), the above advantages and the quick turnaround time in developing a new multi-output power supply may outweigh its disadvantages.

References

1. Linear Technology LT1070 and LT1074 data sheets.
2. LT1170 data sheet.
3. Carl Nelson and Jim Wilson, LTC Application Note 19, Linear Technology.
4. LT1170 data sheet.
5. Linear Technology Power Solutions, 1997.
6. Carl Nelson, LTC Application Note 44, Linear Technology.
7. R. Mammano, "Distributed Power Systems," Unitrode Corp. publication.

Symbols, Units, and Conversion Factors

Symbols Frequently Used in This Book

A_b Winding area of a core bobbin (usually given in square inches)

A_e Effective core area (usually given in square centimeters)

A_l Inductance of core (usually given in millihenries per 1000 turns

B_r Remanence flux density of a core material (flux density at zero oersteds)

B_s Saturation flux density of a core material

D_{cma} Current density in wire (usually expressed in circular mils per rms ampere)

l_m Effective path length of a magnetic core (usually quoted in centimeters)

Φ Flux in a magnetic core (in CGS/EMU units, it is usually given in maxwells and is equal to flux density in gauss × core area in square centimeters)

TABLE A.1 Conversion Table for Frequently Used Magnetic Units

Quantity	CGS/EMU units	MKS units	To convert from: CGS/EMU to MKS multiply by	To convert from: MKS to CGS/EMU multiply by
Flux	maxwell	weber	1×10^{-8}	$1 \times 10^{+8}$
Flux density	gauss	tesla	1×10^{-4}	$1 \times 10^{+4}$
Flux density	gauss	millitesla	1×10^{-1}	$1 \times 10^{+1}$
Flux density	gauss	weber/meter^2	1×10^{-4}	$1 \times 10^{+4}$
Magnetic field intensity	oersted	amprere turns/meter	79.5	1.26×10^{-2}

TABLE A.2 Other Conversion Factors Used in This book

Quantity	A	B	To convert from: A to B multiply by	To convert from: B to A multiply by
Current density	Circular mils	Square inches	7.85×10^{-7}	$1.27 \times 10^{+6}$
Current density	Circular mils	Square centimeters	5.07×10^{-6}	$1.98 \times 10^{+5}$

SIEMENS

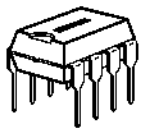
Control IC for Switched-Mode Power Supplies using MOS-Transistor

TDA 4605-3

Bipolar IC

Features

- Fold-back characteristics provides overload protection for external components
- Burst operation under secondary short-circuit condition implemented
- Protection against open or a short of the control loop
- Switch-off if line voltage is too low (undervoltage switch-off)
- Line voltage depending compensation of fold-back point
- Soft-start for quiet start-up without noise generated by the transformer
- Chip-over temperature protection implemented (thermal shutdown)
- On-chip ringing suppression circuit against parasitic oscillations of the transformer
- AGC-voltage reduction at low load



P-DIP-8-1

Type	Ordering Code	Package
TDA 4605-3	Q67000-A5066	P-DIP-8-1

The IC TDA 4605-3 controls the MOS-power transistor and performs all necessary control and protection functions in free running flyback converters. Because of the fact that a wide load range is achieved, this IC is applicable for consumer as well as industrial power supplies.

The serial circuit and primary winding of the flyback transformer are connected in series to the input voltage. During the switch-on period of the transistor, energy is stored in the transformer. During the switch-off period the energy is fed to the load via the secondary winding. By varying switch-on time of the power transistor, the IC controls each portion of energy transferred to the secondary side such that the output voltage remains nearly independent of load variations. The required control information is taken from the input voltage during the switch-on period and from a regulation winding during the switch-off period. A new cycle will start if the transformer has transferred the stored energy completely into the load.

In the different load ranges the switched-mode power supply (SMPS) behaves as follows:

No load operation

The power supply is operating in the burst mode at typical 20 to 40 kHz. The output voltage can be a little bit higher or lower than the nominal value depending of the design of the transformer and the resistors of the control voltage divider.

Nominal operation

The switching frequency is reduced with increasing load and decreasing AC-voltage.
The output voltage is only dependent on the load.

Overload point

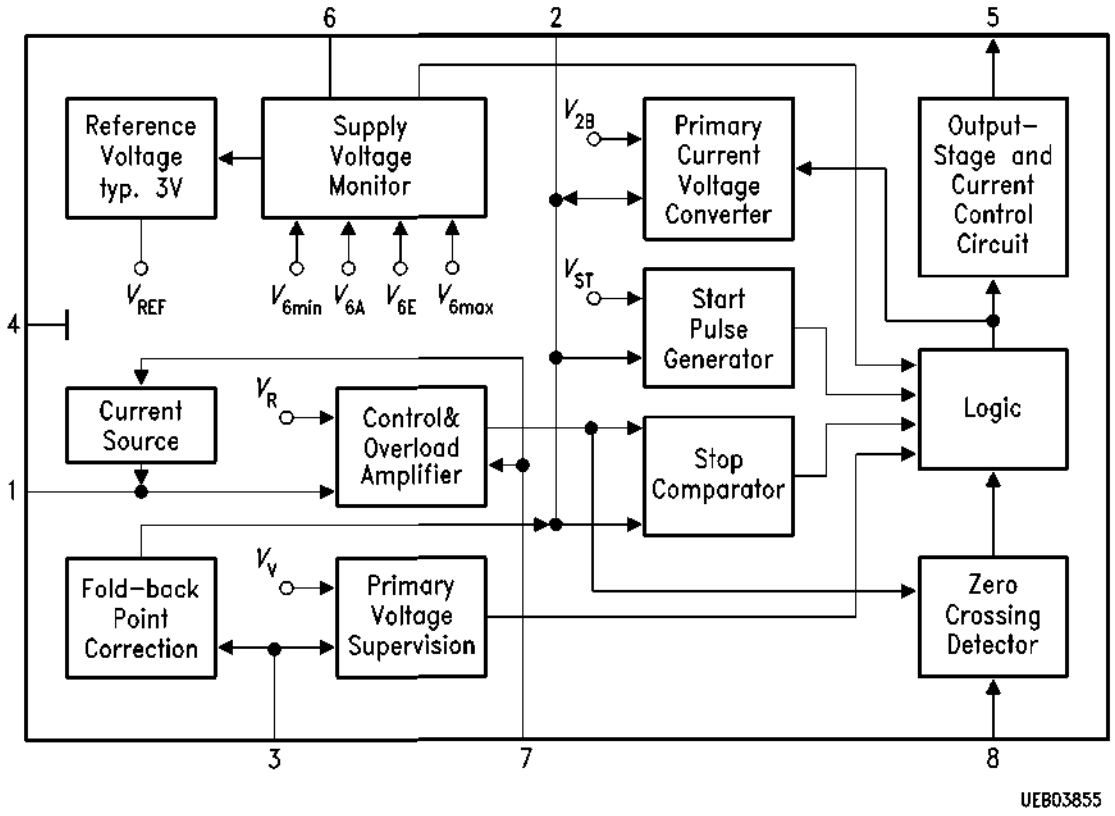
Maximal output power is available at this point of the output characteristic.

Overload

The energy transferred per operation cycle is limited at the top. Therefore the output voltages declines by secondary overloading.

Pin Definitions and Functions

Pin No.	Function
1	**Information Input Concerning Secondary Voltage.** By comparing the regulating voltage - obtained trom the regulating winding of the transformer - with the internal reference voltage, the output impulse width on pin 5 is adjusted to the load of the secondary side (normal, overload, short-circuit, no load).
2	**Information Input Regarding the Primary Current.** The primary current rise in the primary winding is simulated at pin 2 as a voltage rise by means of external RC-element. When a voltage level is reached thats derived from the regulating voltage at pin 1, the output impulse at pin 5 is terminated. The RC-element serves to set the maximum power at the overload point set.
3	**Input for Primary Voltage Monitoring:** In the normal operation V_3 is moving between the thresholds V_{3H} and V_{3L} ($V_{3H} > V_3 > V_{3L}$). $V_3 < V_{3L}$: SMPS is switched OFF (line voltage too low). $V_3 > V_{3H}$: Compensation of the overload point regulation (controlled by pin 2) starts at $V_{3H} : V_{3L} = 1.7$.
4	**Ground**
5	**Output:** Push-pull output provides ± 1 A for rapid charge and discharge of the gate capacitance of the power MOS-transistor.
6	**Supply Voltage Input:** A stable internal reference voltage V_{REF} is derived from the supply voltage also the switching thresholds V_{6A} , V_{6E} , $V_{6\,max}$ and $V_{6\,min}$ for the supply voltage detector. If $V_6 > V_{6E}$ then V_{REF} is switched on and swiched off when $V_6 < V_{6A}$. In addition the logic is only enable for $V_{6\,min} < V_6 < V_{6\,max}$.
7	**Input for Soft-Start.** Start-up will begin with short pulses by connecting a capacitor from pin 7 to ground.
8	**Input for the Oscillation Feedback.** After starting oscillation, every zero transition of the feedback voltage (falling edge) through zero (falling edge) triggers an output pulse at pin 5. The trigger threshold is at + 50 mV typical.



Block Diagram

Circuit Description

Application Circuit

The application circuit shows a flyback converter for video recorders with an output power rating of 70 W. The circuit is designed as a wide-range power supply for AC-line voltages of 180 to 264 V. The AC-input voltage is rectified by the bridge rectifier GR1 and smoothed by C_1. The NTC limits the rush-in current.

In the period before the switch-on threshold is reached the IC is suppled via resistor R_1; during the start-up phase it uses the energy stored in C_2, under steady state conditions the IC receives its supply voltage from transformer winding n_1 via diode D1. The switching transistor T1 is a BUZ 90. The parallel connected capacitor C_3 and the inductance of primary winding n_2 determine the system resonance frequency. The R_2-C_4-D2 circuitry limits overshoot peaks, and R_3 protects the gate of T1 against static charges.

During the conductive phase of the power transistor T1 the current rise in the primary winding depends on the winding inductance and the mains voltage. The network consisting of R_4-C_5 is used to create a model of the sawtooth shaped rise of the collector current. The resulting control voltage is fed into pin 2 of the IC. The RC-time constant given by R_4-C_5 must be designed that way that driving the transistor core into saturation is avoided.

The ratio of the voltage divider R_{10}/R_{11} is fixing a voltage level threshold. Below this threshold the switching power supply shall stop operation because of the low mains voltage. The control voltage present at pin 3 also determines the correction current for the fold-back point. This current added to the current flowing through R_4 and represents an additional charge to C_5 in order to reduce the turn-on phase of T1. This is done to stabilize the fold-back point even under higher mains voltages.

Regulation of the switched-mode power supplies via pin 1. The control voltage of winding n_1 during the off period of T1 is rectified by D3, smoothed by C_6 and stepped down at an adjustable ratio by R_5, R_6 and R_7. The R_8-C_7 network suppresses parasitic overshoots (transformer oscillation). The peak voltage at pin 2, and thus the primary peak current, is adjusted by the IC so that the voltage applied across the control winding, and hence the output voltages, are at the desired level.

When the transformer has supplied its energy to the load, the control voltage passes through zero. The IC detects the zero crossing via series resistors R_9 connected to pin 8. But zero crossings are also produced by transformer oscillation after T1 has turned off if output is short-circuited. Therefore the IC ignores zero crossings occurring within a specified period of time after T1 turn-off.

The capacitor C_8 connected to pin 7 causes the power supply to be started with shorter pulses to keep the operating frequency outside the audible range during start-up.

On the secondary side, five output voltages are produced across winding n_3 to n_7 rectified by D4 to D8 and smoothed by C_9 to C_{13}. Resistors R_{12}, R_{14} and R_{19} to R_{21} are used as bleeder resistors. Fusable resistors R_{15} to R_{18} protect the rectifiers against short circuits in the output circuits, which are designed to supply only small loads.

Block Diagram

Pin 1

The regulating voltage forwarded to this pin is compared with a stable internal reference voltage V_R in the **regulating and overload amplifier**. The output of this stage is fed to the stop comparator. If the control voltage is rather small at pin 1 an additional current is added by means of current source which is controlled according the level at pin 7. This additional current is virtually reducing the control voltage present at pin 1.

Pin 2

A voltage proportional to the drain current of the switching transistor is generated there by the external RC-combination in conjunction with the **primary current transducer**. The output of this transducer is controlled by the logic and referenced to the internal stable voltage V_{2B}. If the voltage V_2 exceeds the output voltage of the regulations amplifier, the logic is reset by the stop comparator and consequently the output of pin 5 is switched to low potential. Further inputs for the logic stage are the output for the **start impulse generator** with the stable reference potential V_{ST} and the **supply voltage motor**.

Pin 3

The down divided primary voltage applied there stabilizes the overload point. In addition the logic is disabled in the event of low voltage by comparison with the internal stable voltage V_V in the **primary voltage monitor** block.

Pin 4

Ground

Pin 5

In the output stage the output signals produced by the logic are shifted to a level suitable for MOS-power transistors.

Pin 6

From the supply voltage V_6 are derived a stable internal references V_{REF} and the switching threshold V_{6A}, V_{6E}, $V_{6\,max}$ and $V_{6\,min}$ for the **supply voltage monitor**. All references values (V_R, V_{2B}, V_{ST}) are derived from V_{REF}. If $V_6 > V_{VE}$, the V_{REF} is switched on and switched off when $V_6 < V_{6A}$. In addition, the logic is released only for $V_{6\,min} < V_6 < V_{6\,max}$.

Pin 7

The output of the overload amplifier is connected to pin 7. A load on this output causes a reduction in maximal impulse duration. This function can be used to implement a soft start, when pin 7 is connected to ground by a capacitor.

Pin 8

The zero detector controlling the logic block recognizes the transformer being discharged by positive to negative zero crossing of pin 8 voltage and enables the logic for a new pulse. Parasitic oscillations occurring at the end of a pulse cannot lead to a new pulse (double pulsing), because an internal circuit inhibits the zero detector for a finite time t_{UL} after the end of each pulse.

Start-Up Behaviour

The start-up behaviour of the application circuit per sheet 88 is represented an sheet 90 for a line voltage barely above the lower acceptable limit time t_0 the following voltages built up:

- V_6 corresponding to the half-wave charge current over R_1
- V_2 to $V_{2\,max}$ (typically 6.6 V)
- V_3 to the value determined by the divider R_{10}/R_{11}.

The current drawn by the IC in this case is less than 1.6 mA.

If V_6 reaches the threshold V_{6E} (time point t_1), the IC switches on the internal reference voltage. The current draw max. rises to 12 mA. The primary current- voltage reproducer regulates V_2 down to V_{2B} and the starting impulse generator generates the starting impulses from time point t_5 to t_6. The feedback to pin 8 starts the next impulse and so on. All impulses including the starting impulse are controlled in width by regulating voltage of pin 1. When switching on this corresponds to a short-circuit event, i.e. V_1 = 0. Hence the IC starts up with "short-circuit impulses" to assume a width depending on the regulating voltage feedback (the IC operates in the overload range). The IC operates at the overload point. Thereafter the peak values of V_2 decrease rapidly, as the starting attempt is aborted (pin 5 is switched to low). As the IC remains switched on, V_6 further decreases to V_6. The IC switches off; V_6 can rise again (time point t_4) and a new start-up attempt begins at time point t_1. If the rectified alternating line voltage (primary voltage) collapses during load, V_3 can fall below V_{3A}, as is happening at time point t_3 (switch-on attempt when voltage is too low). The primary voltage monitor then clamps V_3 to V_{3S} until the IC switches off ($V_6 < V_{6A}$). Then a new start-up attempt begins at time point t_1.

Regulation, Overload and No-Load Behaviour

When the IC has started up, it is operating in the regulation range. The potential at pin 1 typically is 400 mV. If the output is loaded, the regulation amplifier allows broader impulses (V_5 = H). The peak voltage value at pin 2 increases up to $V_{2S\,max}$. If the secondary load is further increased, the overload amplifier begins to regulate the pulse width downward. This point is referred to as the overload point of the power supply. As the IC-supply voltage V_6 is directly proportional to the secondary voltage, it goes down in accordance with the overload regulation behaviour. If V_6 falls below the value $V_{6\,min}$, the IC goes into burst operation. As the time constant of the half-wave charge-up is relatively large, the short-circuit power remains small. The overload amplifier cuts back to the pulse width t_{pk}. This pulse width must remain possible, in order to permit the IC to start-up without problems from the virtual short-circuit, which every switching on with V_1 = 0 represents. If the secondary side is unloaded, the loading impulses (V_5 = H) become shorter. The frequency increases up to the resonance frequency of the system. If the load is further reduced, the secondary voltages and V_6 increase. When $V_6 = V_{6\,max}$ the logic is blocked. The IC converts to burst operation.This renders the circuit absolutely safe under no-load conditions.

Behaviour when Temperature Exceeds Limit

An integrated temperature protection disables the logic when the chip temperature becomes too high. The IC automatically interrogates the temperature and starts as soon as the temperature decreases to permissible values.

Absolute Maximum Ratings

T_A = – 20 to 85 °C

Parameter		Symbol	Limit Values		Unit	Remarks
			min.	max.		
Voltages	pin 1	V_1	– 0.3	3	V	
	pin 2	V_2	– 0.3		V	
	pin 3	V_3	– 0.3		V	
	pin 5	V_5	– 0.3	V_6	V	
	pin 6	V_6	– 0.3	20	V	Supply voltage
	pin 7	V_7	– 0.3		V	
Currents	pin 1	I_1		3	mA	
	pin 2	I_2		3	mA	
	pin 3	I_3		3	mA	
	pin 4	I_4	– 1.5		A	$t_p \leq 50$ µs; v ≤ 0.1*)
	pin 5	I_5	– 0.5	1.5	A	$t_p \leq 50$ µs; v ≤ 0.1
	pin 6	I_6		0.5	A	$t_p \leq 50$ µs; v ≤ 0.1
	pin 7	I_7		3	mA	
	pin 8	I_8	– 5	3	mA	
Junction temperature		T_j		125	°C	
Storage temperature		T_{stg}	– 40	125	°C	

Operating Range

Supply voltage	V_6	7.5	15.5	V	IC "on"
Ambient temperature	T_A	– 20	85	°C	
Heat resistance					
Junction to environment	$R_{th\ JE}$		100	K/W	
Junction case	$R_{th\ JC}$		70	K/W	measured at pin 4

*) t_p = pulse width
v = duty circle

Characteristics

T_A = 25 °C; V_S = 10 V

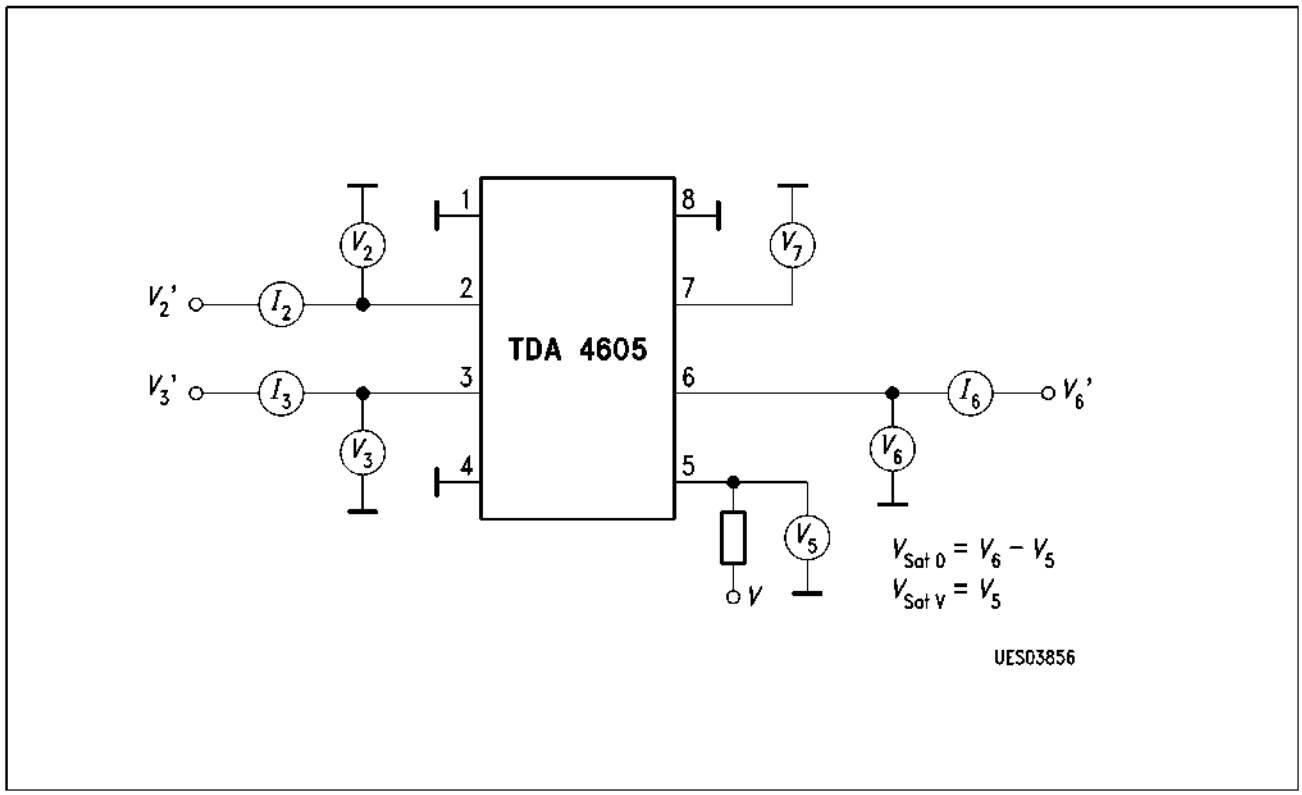
Parameter	Symbol	Limit Values			Unit	Test Condition	Test Circuit
		min.	typ.	max.			
Start-Up Hysteresis							
Start-up current drain	I_{6E0}		0.6	0.8	mA	$V_6 = V_{6E}$	1
Switch-on voltage	V_{6E}	11	12	13	V		1
Switch-off voltage	V_{6A}	4.5	5	5.5	V		1
Switch-on current	I_{6E1}	7	11	14	mA	$V_6 = V_{6E}$	1
Switch-off current	I_{6A1}	5	10	13	mA	$V_6 = V_{6A}$	1
Voltage Clamp (V_6 = 10 V, IC switched off)							
At pin 2 ($V_6 \leq V_{6E}$)	$V_{2\,max}$	5.6	6.6	8	V	I_2 = 1 mA	1
At pin 3 ($V_6 \leq V_{6E}$)	$V_{3\,max}$	5.6	6.6	8	V	I_3 = 1 mA	1
Control Range							
Control input voltage	V_{1R}	390	400	410	mV		2
Voltage gain of the control circuit in the control range	$-V_R$	30	43	60	dB	V_R = d ($V_{2S} - V_{2B}$) / – dV_1 f = 1 kHz	2
Primary Current Simulation Voltage							
Basic value	V_{2B}	0.97	1.00	1.03	V		2
Overload Range and Short-Circuit Operation							
Peak value in the range of secondary overload	V_{2B}	2.9	3.0	3.1	V	$V_1 = V_{1R}$ – 10 mV	2
Peak value in the range of secondary short-circuit operation	V_{2K}	2.2	2.4	2.6	V	V_1 = 0 V	2
Fold-Back Point Correction							
Fold-back point correction current	$-I_2$	300	500	650	µA	V_3 = 3.7 V	1

Characteristics (cont'd)
T_A = 25 °C; V_S = 10 V

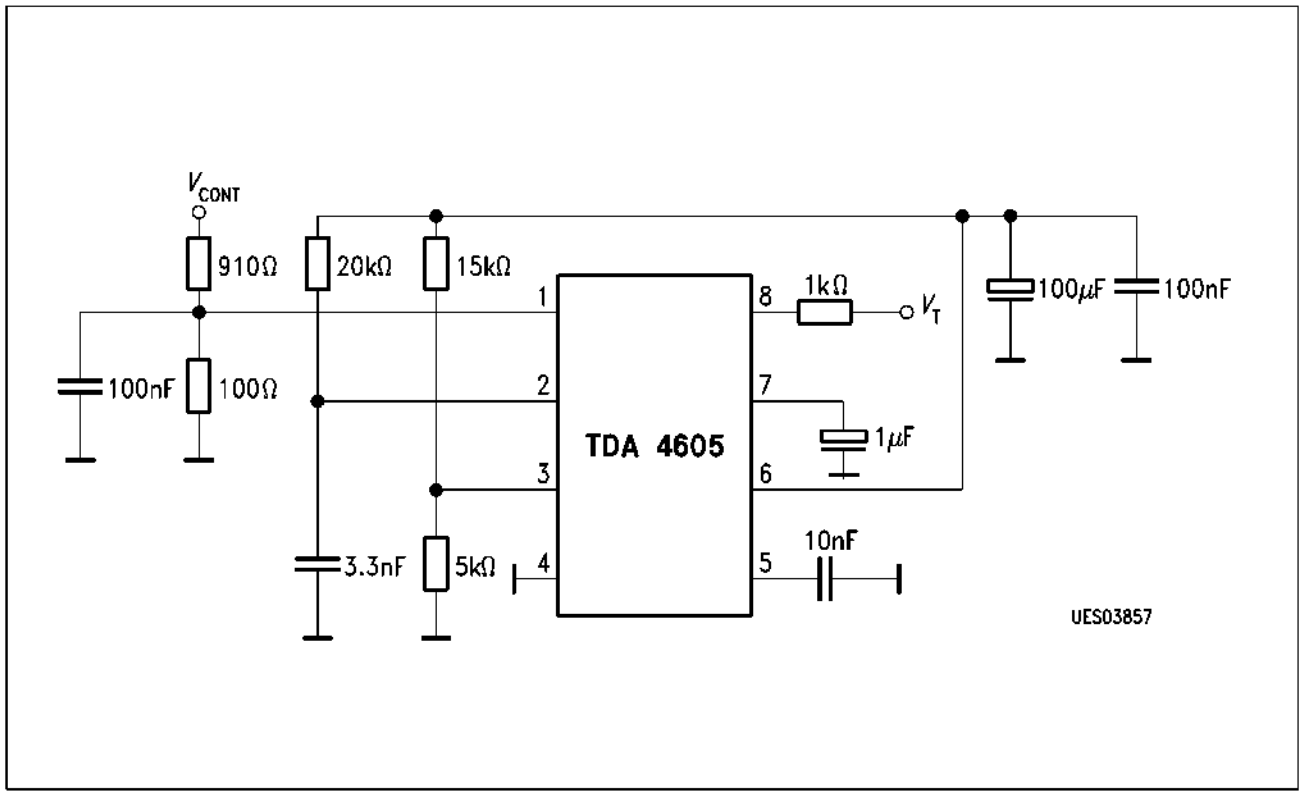
Parameter	Symbol	Limit Values			Unit	Test Condition	Test Circuit
		min.	typ.	max.			
Generally Valid Data (V_6 = 10 V)							
Voltage of the Zero Transition Detector							
Positive clamping voltage	V_{8P}	0.7	0.75	0.82	V	I_8 = 1 mA	2
Negative clamping voltage	V_{8N}	– 0.25	– 0.2	– 0.15	V	I_8 = – 1 mA	2
Threshold value	V_{8S}	40	50	76	mV		2
Suppression of transformer ringing	t_{UL}	3.0	3.5	3.8	µs		2
Input current	$-I_8$	0		4	µA	V_8 = 0	2
Push-Pull Output Stage							
Saturation voltages							
Pin 5 sourcing	V_{Sat0}		1.5	2.0	V	I_5 = – 0.1 A	1
Pin 5 sinking	V_{SatV}		1.0	1.2	V	I_5 = + 0.1 A	1
Pin 5 sinking	V_{SatV}		1.4	1.8	V	I_5 = + 0.5 A	1
Output Slew Rate							
Rising edge	+ dV_5/dt		70		V/µs		2
Falling edge	+ dV_5/dt		100		V/µs		2
Reduction of Control Voltage							
Current to reduce the control voltage	$-I_1$		50	130	µA	V_7 = 1.1 V, V_1 = 0.4 V	

Characteristics (cont'd)
T_A = 25 °C; V_S = 10 V

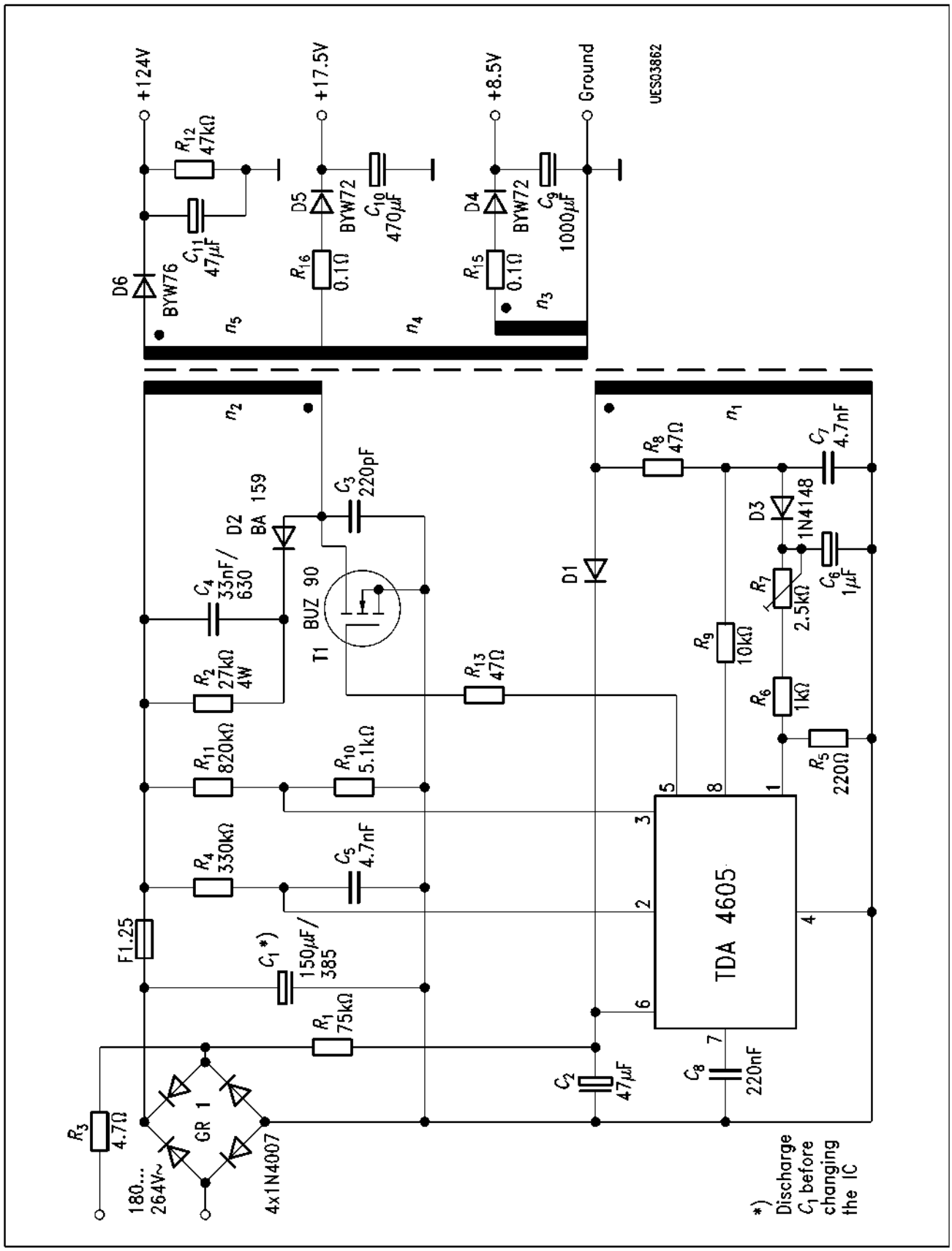
Parameter	Symbol	Limit Values			Unit	Test Condition	Test Circuit
		min.	typ.	max.			
Protection Circuit							
Undervoltage protection for V_6 : voltage at pin 5 = $V_{5\,min}$ if $V_6 < V_{5\,min}$	$V_{6\,min}$	7.0	7.25	7.5	V		2
Undervoltage protection for V_6 : voltage at pin 5 = $V_{5\,min}$ if $V_6 > V_{6\,max}$	$V_{6\,max}$	15.5	16	16.5	V		2
Undervoltage protection for V_{AC} : voltage at pin 4 = $V_{5\,min}$ if $V_3 < V_{3A}$	V_{3A}	985	1000	1015	mV	V_2 = 0 V	1
Over temperature at the given chip temperature the IC will switch V_5 to $V_{5\,min}$	T_j		150		°C		2
Voltage at pin 3 if one of the protection function was triggered; (V_3 will be clamped until $V_6 < V_{6A}$)	V_{3Sat}		0.4	0.8	V	I_3 = 750 µA	1
Current drain during burst operation	I_6		8		mA	$V_3 = V_2$ = 0 V	1



Test Circuit 1

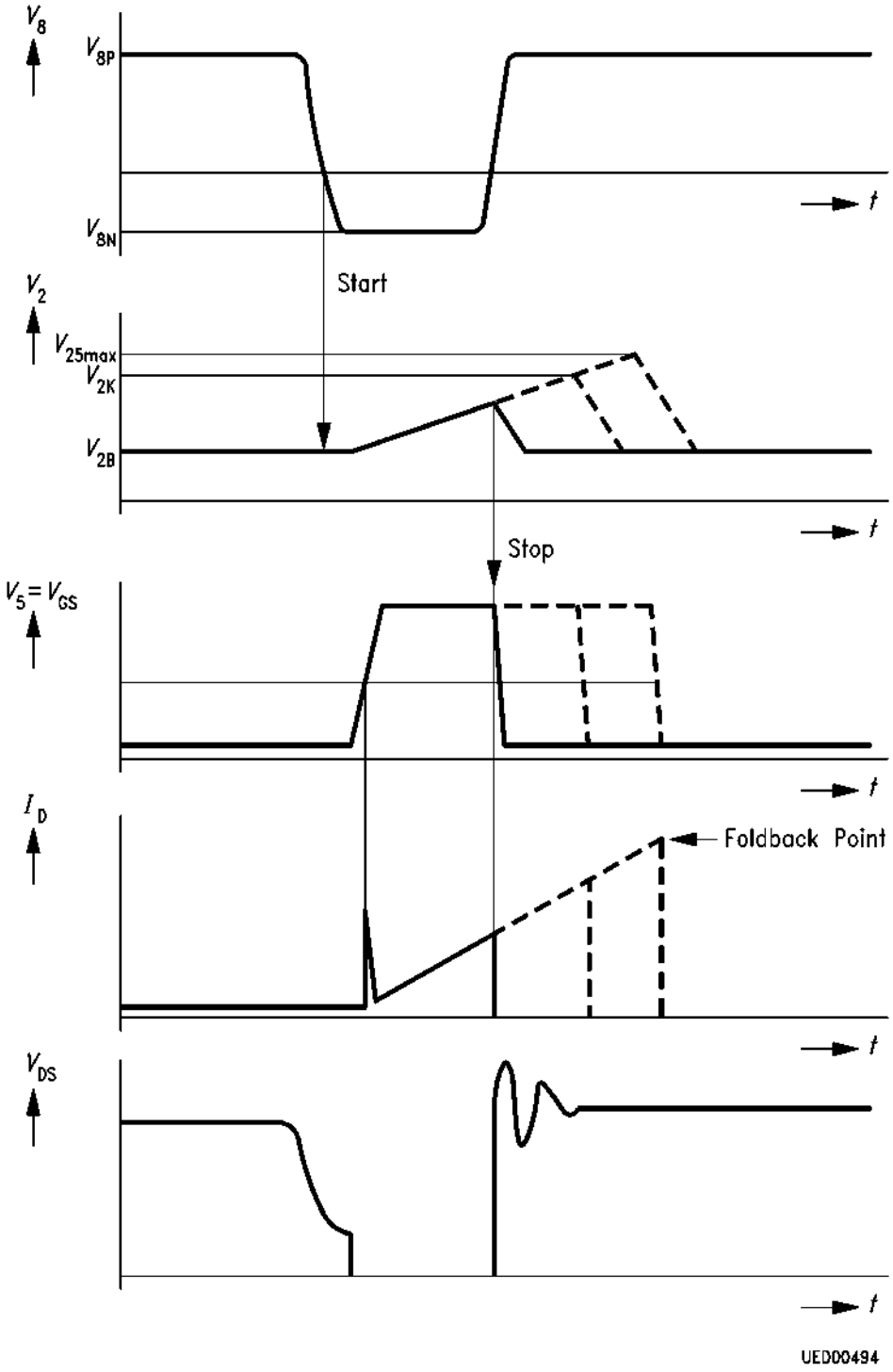


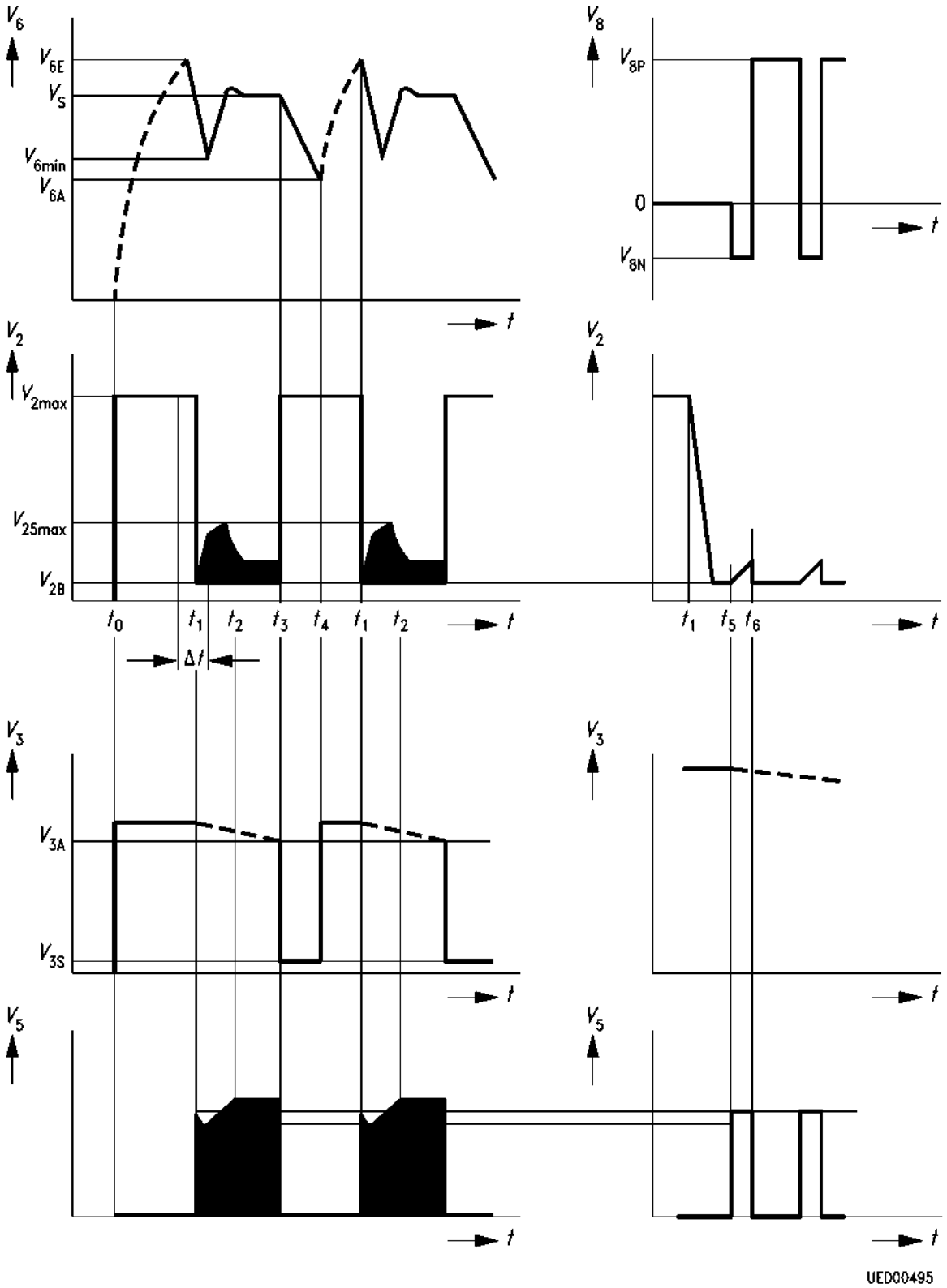
Test Circuit 2

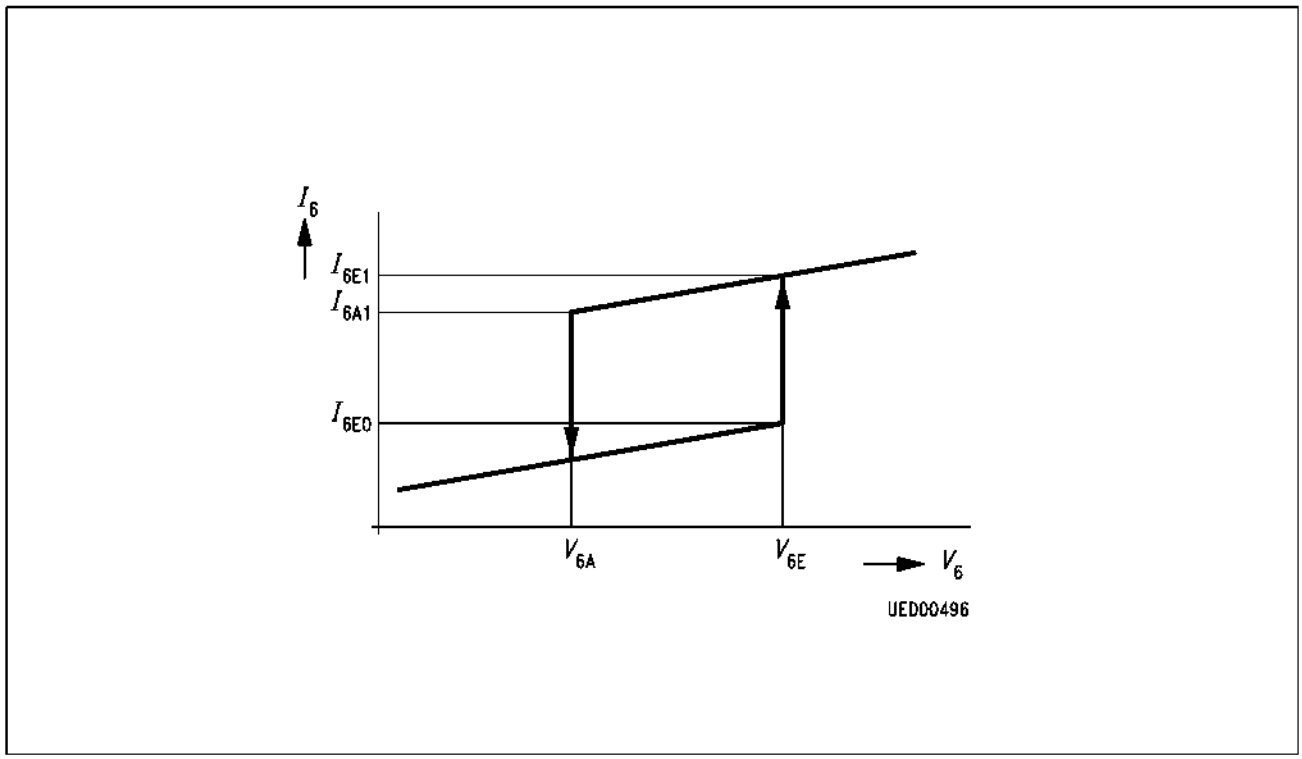


Application Circuit

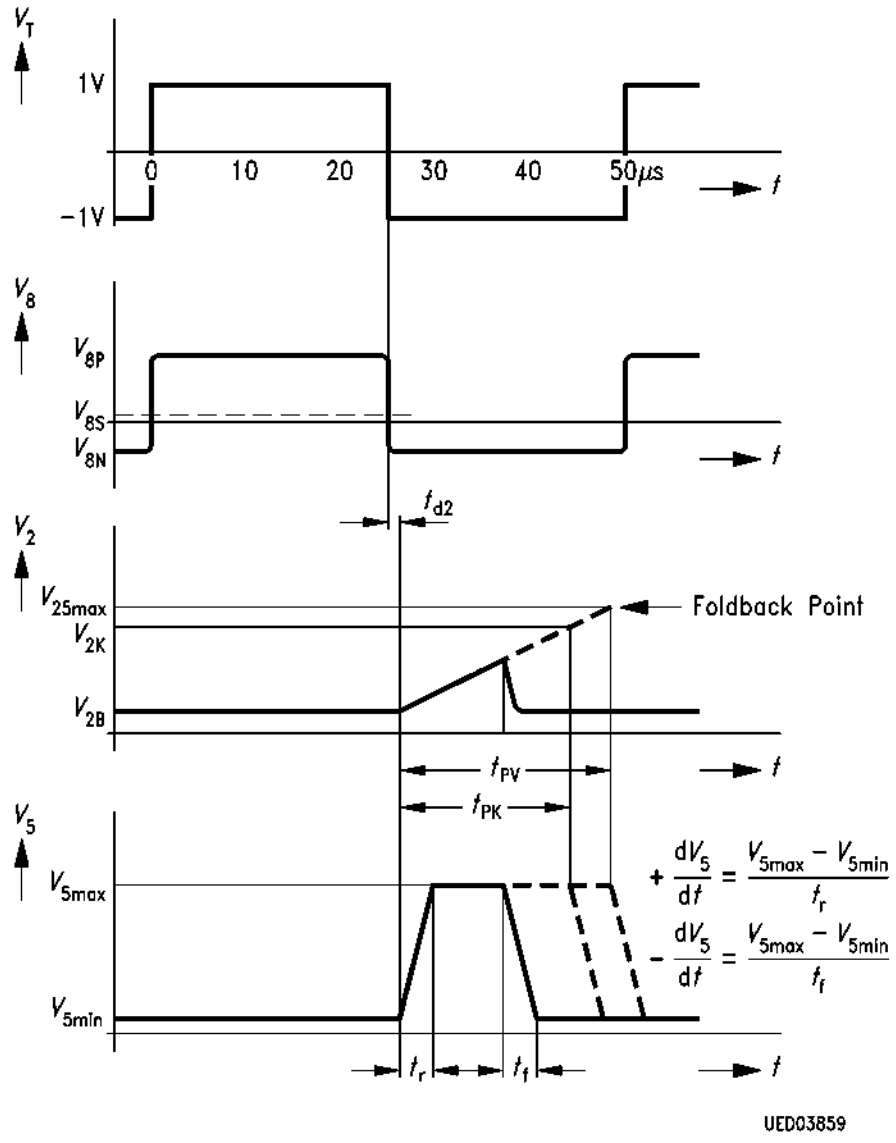
Diagrams


V_8
V_{8P}
V_{8N}
t
Start
V_2
V_{25max}
V_{2K}
V_{2B}
t
Stop
$V_5 = V_{GS}$
t
I_D
Foldback Point
t
V_{DS}
t
UED00494


V_6
V_{6E}
V_S
V_{6min}
V_{6A}
t
V_8
V_{8P}
0
V_{8N}
t
V_2
V_{2max}
V_{25max}
V_{2B}
t_0 t_1 t_2 t_3 t_4 t_1 t_2 t
Δt
V_2
t_1 t_5 t_6 t
V_3
V_{3A}
V_{3S}
t
V_3
t
V_5
t
V_5
t
UED00495

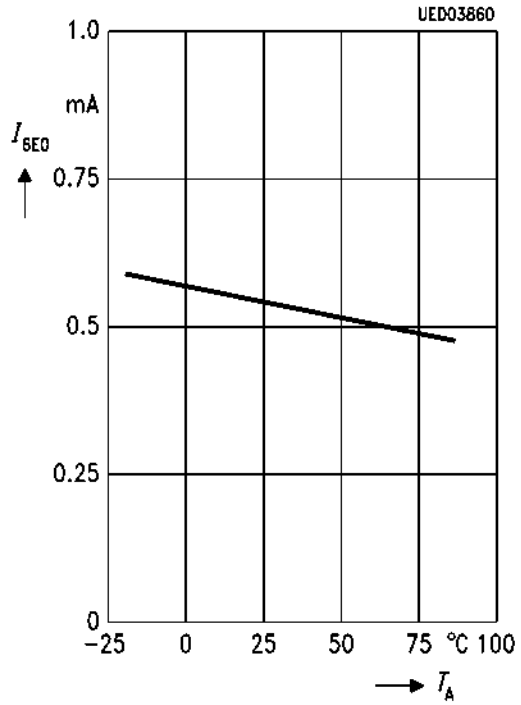


Start-Up Hysteresis

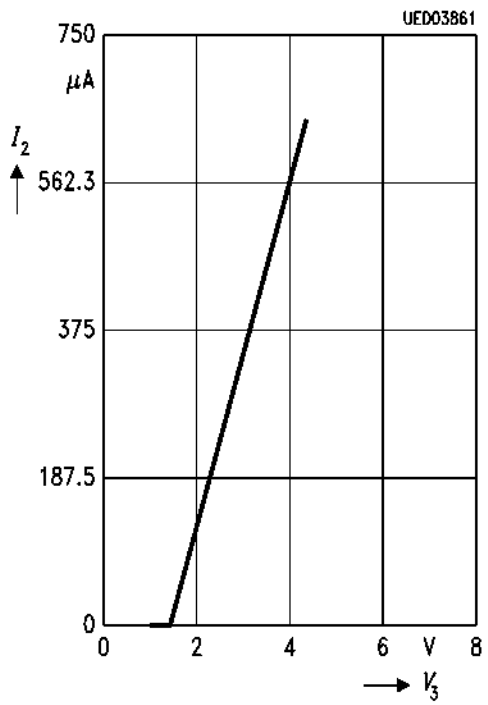


Operation in Test Circuit 2

Start-Up Current as a Function of the Ambient Temperature

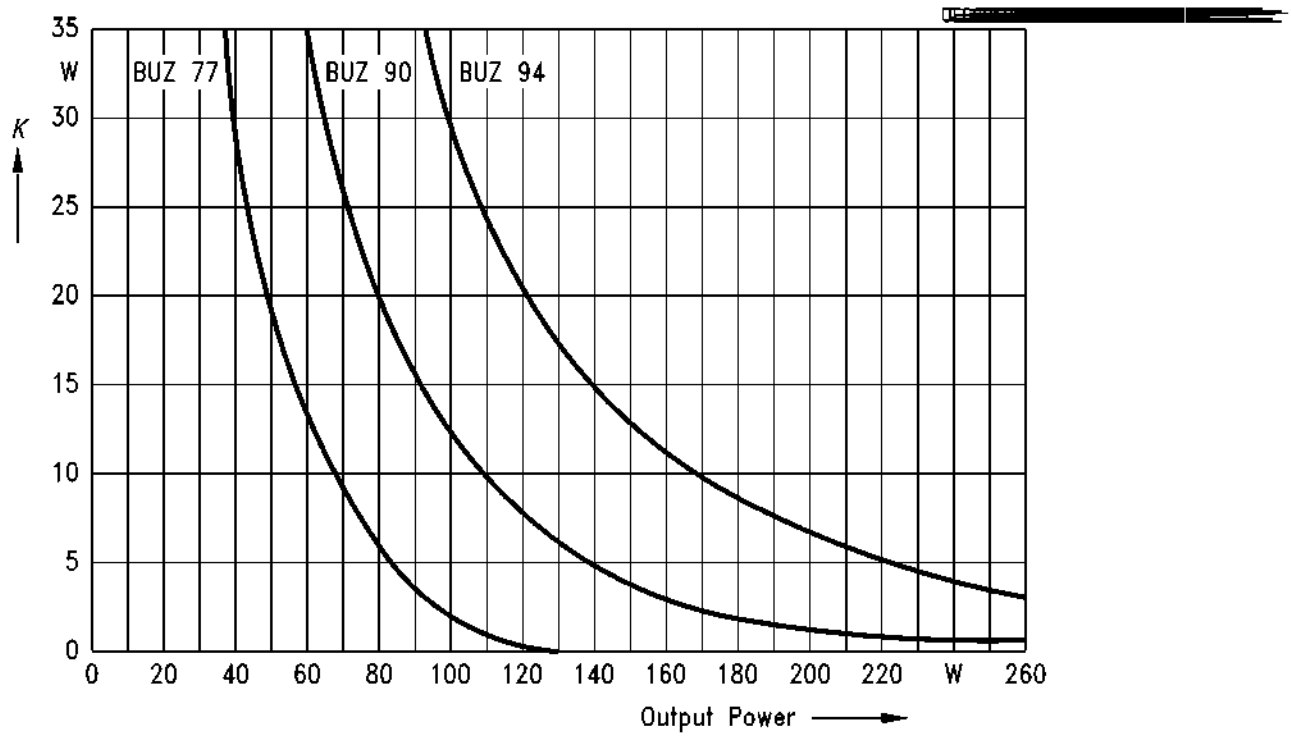


Overload Point Correction as a Function of the Voltage at Pin 3

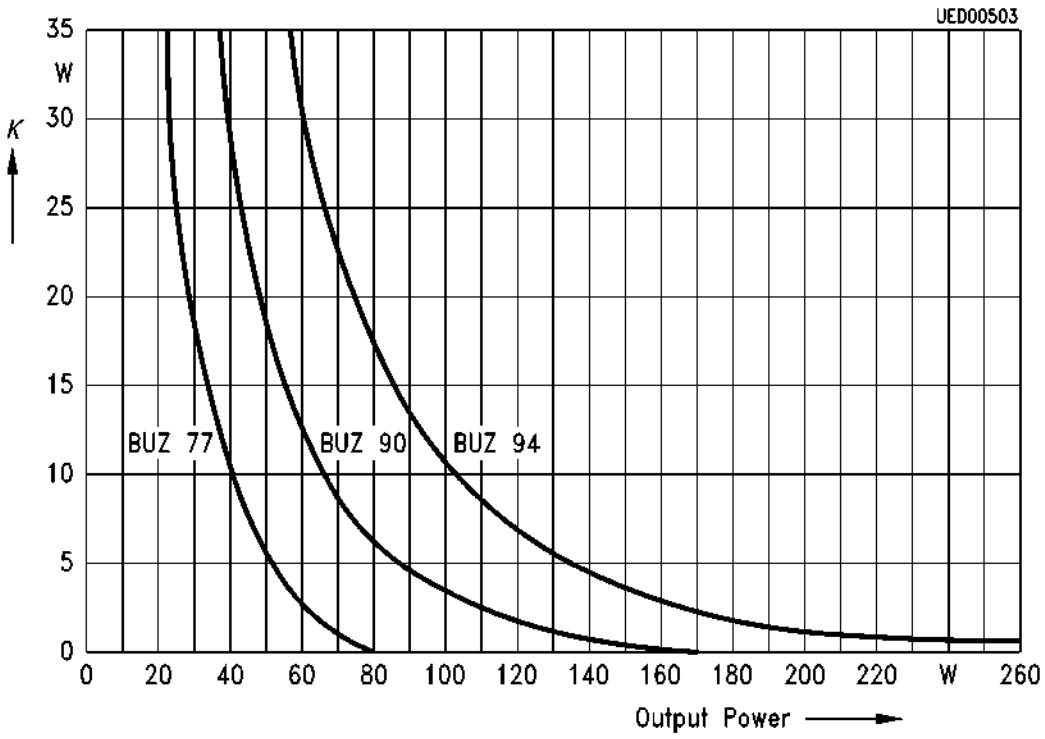


Recommended Heat Sink by 60 °C Ambient Temperature

Narrow Range 180 V ... 120 V ~



Narrow Range 90 V ... 270 V ~



This datasheet has been download from:

www.datasheetcatalog.com

Datasheets for electronics components.

Vysoká škola báňská – Technická univerzita Ostrava

Základy automatizace

učební text

Jiří Tůma, Renata Wagnerová, Radim Farana, Lenka Landryová

Ostrava 2007

Recenze: prof. RNDr. Alena Lukasová, CSc.
Ing. Jaromír Škuta, Ph.D.

Název: Základy automatizace
Autor: Jiří Tůma, Renata Wagnerová, Radim Farana, Lenka Landryová
Vydání: první, 2007
Počet stran: 283
Náklad: -
Vydavatel a tisk: Ediční středisko VŠB – TUO

Studijní materiály pro studijní obor Bakalářské strojírenství fakulty strojní
Jazyková korektura: nebyla provedena.

Určeno pro projekt:
Operační program Rozvoj lidských zdrojů
Název: E-learningové prvky pro podporu výuky odborných a technických předmětů
Číslo: CZ.O4.01.3/3.2.15.2/0326
Realizace: VŠB – Technická univerzita Ostrava
Projekt je spolufinancován z prostředků ESF a státního rozpočtu ČR

© Jiří Tůma, Renata Wagnerová, Radim Farana, Lenka Landryová
© VŠB – Technická univerzita Ostrava

ISBN 978-80-248-1523-7

POKYNY KE STUDIU

Základy automatizace

Pro předmět 3. semestru oboru Bakalářské strojírenství jste obdrželi studijní balík obsahující

- integrované skriptum pro distanční studium obsahující i pokyny ke studiu
- CD-ROM s doplňkovými animacemi vybraných částí kapitol

Prerekvizity

Pro studium tohoto předmětu se předpokládá absolvování předmětů teoretického základu 1. ročníku BS na strojní fakultě v ČR.

Cílem předmětu

je seznámení se základními pojmy z oblasti automatizace.

Pro koho je předmět určen

Modul je zařazen do bakalářského studia studijního programu Bakalářské strojírenství, ale může jej studovat i zájemce z kteréhokoliv jiného oboru, pokud splňuje požadované prerekvizity.

Skriptum se dělí na části, kapitoly, které odpovídají logickému dělení studované látky, ale nejsou stejně obsáhlé. Předpokládaná doba ke studiu kapitoly se může výrazně lišit, proto jsou velké kapitoly děleny dále na číslované podkapitoly a těm odpovídá níže popsaná struktura.

Při studiu každé kapitoly doporučujeme následující postup:

Čas ke studiu: xx hodin

Na úvod kapitoly je uveden **čas** potřebný k prostudování látky. Čas je orientační a může vám sloužit jako hrubé vodítko pro rozvržení studia celého předmětu či kapitoly. Někomu se čas může zdát příliš dlouhý, někomu naopak. Jsou studenti, kteří se s touto problematikou ještě nikdy nesetkali a naopak takoví, kteří již v tomto oboru mají bohaté zkušenosti.

Cíl: Po prostudování tohoto odstavce budete umět

- popsat ...
- definovat ...
- vyřešit ...

Ihned potom jsou uvedeny cíle, kterých máte dosáhnout po prostudování této kapitoly – konkrétní dovednosti, znalosti.

VÝKLAD

Následuje vlastní výklad studované látky, zavedení nových pojmů, jejich vysvětlení, vše doprovázeno obrázky, tabulkami, řešenými příklady, odkazy na animace.

Shrnutí pojmů 2.1.

Na závěr kapitoly jsou zopakovány hlavní pojmy, které si v ní máte osvojit. Pokud některému z nich ještě nerozumíte, vraťte se k nim ještě jednou.

Otázky 1.1.

Pro ověření, že jste dobře a úplně látku kapitoly zvládli, máte k dispozici několik teoretických otázek.

Úlohy k řešení 1.1.

Protože většina teoretických pojmů tohoto předmětu má bezprostřední význam a využití v databázové praxi, jsou Vám nakonec předkládány i praktické úlohy k řešení. V nich je hlavní význam předmětu a schopnost aplikovat čerstvě nabyté znalosti při řešení reálných situací hlavním cílem předmětu.

KLÍČ K ŘEŠENÍ

Výsledky zadaných příkladů i teoretických otázek výše jsou uvedeny v závěru učebnice v Klíči k řešení. Používejte je až po vlastním vyřešení úloh, jen tak si samokontrolou ověříte, že jste obsah kapitoly skutečně úplně zvládli.

Úspěšné a příjemné studium s touto učebnicí Vám přejí autoři výukového materiálu

Jiří Tůma, Renata Wagnerová, Radim Farana, Lenka Landryová

Obsah

1. LOGICKÉ ŘÍZENÍ 8

1.1 Úvod 8

- Booleova algebra 9
- Zadání logické funkce 9
- Základní logické funkce 11
- Zákony Booleovy algebry 14
- Syntéza kanonického tvaru kombinační logické funkce 15
- Sestavení logické funkce pomocí Karnaughovy mapy 17

1.2 Realizace kombinačních logických obvodů 21

- Kontaktní schémata 22
- Bloková schémata 23
- Programovatelné logické automaty 26

1.3 Časové operace v logickém řízení 37

- Sekvenční logické obvody 38

1.4 Klopné obvody 41

- Klopný obvod typu RS 41
- Klopný obvod typu D 43
- Klopný obvod typu JK 43

1.5 Zapojení klopných obvodů v sekvenčních obvodech 45

- Použití klopného obvodu typu RS 45
- Použití klopného obvodu typu JK pro čítání impulsů 46

1.6 Sekvenční funkční diagramy 48

- Příklad na sestavení sekvenčního funkčního diagramu 51

1.7 Programovatelné logické automaty 53

- Galvanické oddělení PLC od vnějších obvodů 54
- Funkce sběrnice pro komunikaci 55
- Vnitřní struktura PLC 56
- Vstupy a výstupy PLC 56
- Paměti PLC 57
- Programování PLC 57
- Provedení PLC 58
- Sběrnice PROFIBUS 58

1.8 Laboratorní úlohy na logické řízení 60

- Zařízení na ochranu před úrazem (Bezpečnostní zařízení) 61
- Řízení pásových dopravníků (Dva dopravníky, R-S klopný obvod) 63

2. SPOJITÉ SYSTÉMY 69

2.1. Základní pojmy 69

- Příklad sestavení blokového schématu 70
- Základní činnosti se systémy 71

2.2. Laplaceova transformace 77

2.3. Matematický popis statických a dynamických vlastností spojitých systémů 86

- Statická charakteristika 86
- Diferenciální rovnice 87
- Obrazový přenos 87
- Přechodová funkce 88
- Impulsní funkce 90
- Kmitočtový přenos 92

2.4. Logaritmické kmitočtové charakteristiky 99

- LKCH základních zapojení 100
- Základní členy přenosu 103

2.5. Základní rozdělení spojitých dynamických členů 110

❑ Proporcionální systémy ... 113
❑ Derivační systémy ... 116
❑ Integrační systém ... 118
❑ Dopravní zpoždění ... 120
2.6. Algebra blokových schémat ... 125
❑ Základní přenosy RO ... 132
2.7. Stabilita ... 137
❑ Hurwitzovo kriterium stability ... 139
❑ Routhovo-Schurovo kriterium stability ... 141
❑ Nyquistovo kriterium stability ... 144
2.8. Regulátory ... 148
2.9. Kvalita regulačního pochodu ... 152
2.10. Syntéza regulačních obvodů ... 155
❑ Jednoduché aproximační metody ... 164
3 TECHNICKÉ PROSTŘEDKY ŘÍZENÍ ... 168
3.1. Základní pojmy ... 168
3.2. Pneumatické a hydraulické systémy pro řízení a ovládání ... 171
❑ Hydraulické a pneumatické obvody ... 171
3.3. Analogové regulátory pro zpětnovazební řízení ... 175
❑ Elektronické regulátory ... 175
❑ Mechanické regulátory ... 178
3.4. Pneumatické regulátory ... 180
3.5. Hydraulické zesilovací prvky ... 183
❑ Příklad provedení proporcionálního ventilu (rozváděče) ... 185
❑ Servoventily ... 186
3.6. Snímače ... 187
❑ Některé obecné principy měření fyzikálních veličin ... 189
❑ Snímače teploty pro regulaci ... 192
❑ Snímače průtoku plynů a kapalin pro regulaci ... 194
❑ Snímače síly a kroutícího momentu pro regulaci ... 197
❑ Snímače tlaku a diferenčního tlaku pro regulaci ... 199
❑ Snímače hladiny pro regulaci ... 200
❑ Snímače polohy pro regulaci s analogovým výstupem ... 200
❑ Snímače polohy pro regulaci s číslicovým výstupem ... 201
❑ Dotykové sondy ... 203
❑ Průtok sypkých hmot ... 204
3.7. Akční členy ... 205
❑ Převody pohybů ... 206
❑ Regulační orgány pro ovládání průtoku plynů par a kapalin ... 208
❑ Kontinuální dávkování sypkých hmot ... 211
3.8. Elektrické pohony ... 212
❑ Přehled typů elektrických motorů ... 216
❑ Pracovní oblasti pohonů ... 216
❑ Stejnosměrný motor ... 217
❑ Střídavý asynchronní motor ... 220
❑ Jednofázový komutátorový motor ... 223
❑ Elektronicky komutovaný motor (EC motor) ... 223
❑ Synchronní střídavé motory ... 224
❑ Krokové motory ... 224
❑ Napájení elektrických motorů řízeným usměrňovačem ... 227
❑ Měniče frekvence ... 229
3.9. Hydraulické a pneumatické pohony ... 231
3.10. Piezoelektrické aktuátory ... 236

3.11. Počítačové sítě ... 240
- Síť Internet ... 242
- IP adresy ... 242
- Symbolické adresy ... 244
- Některé služby internetu ... 244

3.12. Vizualizace a řízení z operátorského pracoviště ... 246
- Vizualizace ... 246
- Systémy řízení z operátorského pracoviště ... 247
- Architektura vizualizačního systému ... 248
- Vizualizace dat v Internet/Intranet sítích ... 249
- Alarmy ... 250
- Trendy ... 250
- Příklad řešení vizualizace laboratorní úlohy ... 251
- Řešení vizualizace dat z laboratorní úlohy do prostředí internetu ... 253

4 APLIKACE AUTOMATIZAČNÍ TECHNIKY V TECHNICKÉ PRAXI 258

4.1. Regulace v tepelné technice ... 258
- Regulace teploty ... 258
- Kaskádová regulace v tepelné technice ... 259

4.2. Řízení dávkování směsi a řízení zásoby sypkých hmot ... 260
- Dávkování směsi sypkých hmot ... 260
- Regulace zásoby ... 261
- Kontinuální regulace vlhkosti sypkých hmot ... 262
- Integrační smyčková regulace ... 263

4.3. Řízení polohy ... 263
4.4. Soubor řídicích jednotek automobilu a sběrnice CAN ... 264
4.5. Řízení vstřikování a zapalování paliva zážehových motorů ... 267
4.6. Elektrohydraulické servořízení ... 271
4.7. Inteligentní dům ... 273

KLÍČ K ŘEŠENÍ ... 277

REJSTŘÍK ... 286

SEZNAM ANIMACÍ ... 288

1. LOGICKÉ ŘÍZENÍ

1.1 Úvod

Čas ke studiu: 2 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat logickou proměnou, funkci logické proměnné, zákony Boolovy algebry, Karnaughovu mapu, minimalizaci logických funkcí pomocí Karnaughovy mapy, kontaktní schémata, hradla
- popsat způsoby zadání logické funkce
- vyřešit minimalizaci logických funkcí pomocí Karnaghovy mapy, převádět jednotlivé způsoby zadání logických funkcí mezi sebou (vzorec, tabulka hodnot, Karnaughova mapa)

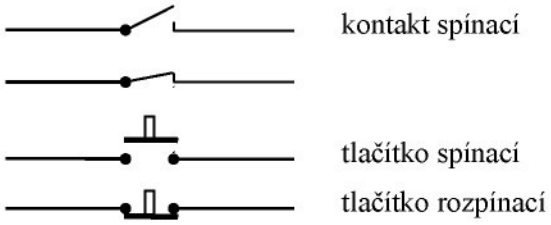
Výklad

Logické řízení je založeno na dvou stavech ovládaného prvku. Praktické označení těchto stavů je následující:

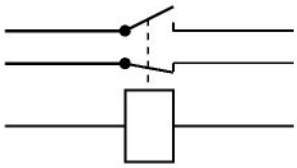
zapnuto / vypnuto, otevřeno / zavřeno, vede proud / nevede proud. atd.

Mezi těmito dvěma stavy není žádný mezistupeň. I když lze u některých prvků třeba připustit otevření napůl, funkce logického řízení to nepředpokládá, může tedy nastat jen jedna ze dvou zmíněných alternativ. Jedním z prvků, který tuto funkci plní je elektrický kontakt (spínač). Ten je buď zapnut nebo vypnut. Při kreslení schémat platí pravidlo, že se kreslí jejich poloha v klidovém stavu (pokud není uvedeno jinak), viz obrázek 1.1. Mezi kontakty řadíme také tlačítka spínací nebo rozpínací.

Kontakty se ovládají buď ručně nebo automaticky dálkově. Příkladem dálkově ovládaného kontaktu je elektromagnetické relé. Relé může obsahovat několik kontaktů dimenzovaných na různá proudová zatížení. Mohou tedy spínat malé proudy nebo velké proudy, což je běžné například u stykačů. Schématické znázornění relé je na obrázku 1.2. Průchodem elektrického proudu cívkou se přitáhne pohyblivá kotva k jádru cívky. Kotva svým pohybem způsobí sepnutí spínacích kontaktů nebo rozepnutí rozpínacích kontaktů. Relé mohou být různého druhu: střídavé, stejnosměrné, jazýčkové,.atd.



Obrázek 1.1 - Spínací a rozpínací kontakt



Obrázek 1.2 - Elektrické relé

Logické obvody budou v této učebnici rozděleny na:

- Kombinační logické obvody
- Sekvenční logické obvody.

Kombinační obvody se liší od sekvenčních tím, že jejich aktuální dvouhodnotový výstup kombinačních logických obvodů je závislý na aktuálních vstupních dvouhodnotových signálech. Naproti tomu výstup sekvenčních logických obvodů je navíc závislý na předcházejících výstupech. Toto dělení se odrazí v dělení textu kapitoly o logickém řízení.

Booleova algebra

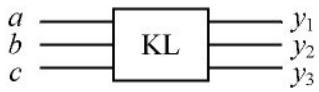
Logické řízení má svoji matematickou teorii, která se nazývá Booleova algebra (George Boole, 1815-1864, irský matematik, hlavní dílo Analysis of Logic). Vzorce nepotřebují reálná čísla, ale jen proměnné (veličiny), , které nabývají dvou hodnot:

ano / ne, pravda / nepravda, true / false (anglicky), logická 1 / logická 0.

Nejběžněji používané hodnoty jsou 1 a 0. Vzorce představují matematickou funkci. Nezávislé proměnné této funkce jsou například pojmenovány: a, b, c, ...atd. Proměnná, která je nazvána např. a může tedy nabýt jedné ze dvou možných hodnot: 1 nebo 0. Pro název proměnné lze použít několika písmen, omezením je jen náš vkus, ve jménu proměnné může být obsažen její technický význam. Základní vlastností funkcí v matematice je jednoznačnost přiřazení závislé proměnné, např. y, určité kombinaci hodnot nezávisle proměnných. Obecná funkce se označuje

$$y = f(a,b,c,...) \tag{1.1}$$

Logická funkce y, (nebo skupina logických funkcí y_1, y_2, y_3, ...) se realizuje kombinační logický obvod (KL), který je znázorněn jako blokové schéma na obr. 1.3. Vstupy jsou kresleny vlevo a výstupy vpravo.



Obrázek 1.3 - Kombinační logická funkce

Zadání logické funkce

Logickou funkci lze zadat následujícími způsoby:

- slovně, jakýkoliv smysluplný text popisující souvislost vstupních a výstupních dvouhodnotových veličin
- tabulkou přiřazující kombinace hodnot vstupních (nezávisle proměnných) k výstupním hodnotám (závisle proměnným)
- Karnaughovou mapou (K-mapou)
- Vzorcem obsahujícím operátory základních logických funkcí.

Z popsaných způsobů zadání si nejprve všimneme tabulky. Tabulka přiřazuje kombinacím vstupních proměnných hodnotu logické funkce, tj. buď 1 nebo 0. Počet kombinací příslušných k počtu *k* vstupních proměnných je

$$N = 2^k \tag{1.2}$$

Například pro tři vstupní proměnné (*a, b, c*) jsou $2^3 = 8$ kombinací hodnot.

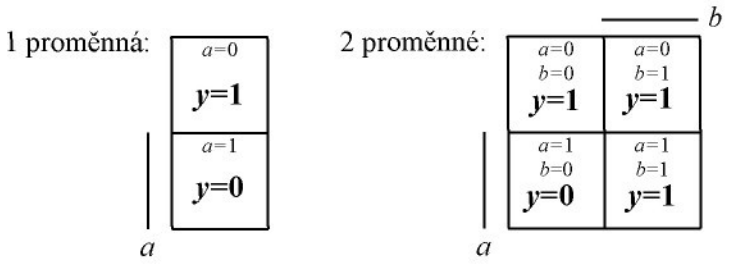
Tabulka 1.1 – Zadání logické funkce tabulkou

Pořadí	Nezávisle proměnné			Závisle proměnná
	a	***b***	***c***	***y***
1	0	0	0	1
2	0	0	1	1
3	0	1	0	0
4	0	1	1	1
5	1	0	0	0
6	1	0	1	0
7	1	1	0	1
8	1	1	1	0

V tabulce udává první sloupec pořadí kombinace, aby byl dokumentován jejich počet. Shodou okolností je pořadí o jednotku vyšší hodnota než je daná kombinace ve dvojkové soustavě. Hodnoty logické funkce doplňuje návrhář logického obvodu nebo automatu a řídí se jeho požadovanou funkcí.

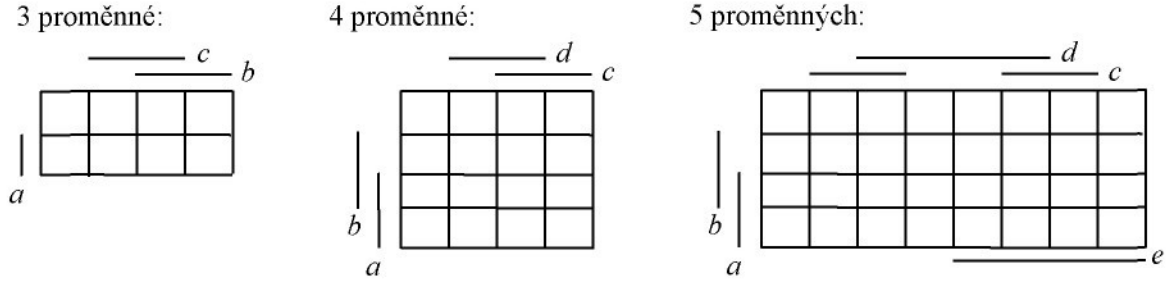
Je třeba zdůraznit, že funkční hodnoty nemusí být definovány pro všechny kombinace vstupních proměnných. Z definice mohou být vyňaty kombinace, které nemohou nastat. Například logická jednotka z obou koncových snímačů polohy jeřábu na jeřábové dráze nemůže nastat (jeden jeřáb nemůže být současně na obou koncích dráhy). Tato skutečnost bude využita v úloze optimalizace zápisu logické funkce, která bude vysvětlen v jedné z příštích kapitol.

Zápis formou tabulky představuje také Karnaughova mapa (K-mapa), jejíž pole jsou účelně uspořádána tak, aby bylo možné provádět jednoduše již zmíněnou optimalizaci zápisu logické funkce. Kombinaci hodnot proměnných určuje poloha pole a příslušná hodnota logické funkce je vepsána na plochu tohoto pole. Nejprve budou uvedeny Karnaughovy mapy pro jednu a dvě proměnné. Pro pochopení jejich uspořádání jsou do jednotlivých polí vepsány také příslušné hodnoty logických proměnných. Svislé a vodorovné pruhy po stranách tabulky označují řádek (nebo skupinu řádků) a sloupec nebo (skupinu sloupců), jejíž pole přísluší hodnotě 1 pro logickou proměnnou, která je k tomuto pruhu připsána.



Obrázek 1.4 - Karnaughova mapa pro jednu a dvě proměnné

Pro přehlednost, logická nula není do polí vpisována. Pole s nedefinovanou hodnotou logické funkce je vyplněno písmenem X. Uspořádání Karnaughovya mapy pro více proměnných je na následujícím obrázku. Tyto mapy jsou vytvářeny podle pravidla, které připouští změnu pouze jedné proměnné u sousedních sloupců nebo řádků. Mapa pro jednu a dvě proměnné toto pravidlo pro svou jednoduchost nepotřebuje. Pouze mapy pro více proměnných se tímto pravidlem řídí.



Obrázek 1.5 - Karnaughova mapa pro 3 až 5 proměnných

Poslední varianta zápisu logické funkce je vzorec, tj. funkční předpis logické funkce, která popisuje jednotlivé operace s nezávisle proměnnými.

Základní logické funkce

V definici základních logických funkcí (tzv. funktory) budou použity předchozí definice (bez k-mapy). Slovní vyjádření představuje klíčové slovo při definici hodnoty logické funkce. Například *opak* znamená, že hodnota logické funkce je opakem hodnoty logické proměnné. Další příklad *nebo* znamená, že jestliže jedna *nebo* druhá proměnná a *nebo* současně obě proměnné jsou rovny logické hodnotě 1, pak je logická funkce rovna také 1.

Negace	**Logický součet (disjunkce)**
Slovní vyjádření: opak	Slovní vyjádření: nebo
Anglický název: NOT	Anglický název: OR
Operátor: pruh nad proměnnou	Operátor: $+, \cup$
Vzorec: $y = \overline{a}$ ($y = \text{NOT}\, a$)	Vzorec: $y = a + b$ ($y = a \text{ OR } b$)
Tabulka:	Tabulka:

a	*y*
0	1
1	0

a	*b*	*y*
0	0	0
0	1	1
1	0	1
1	1	1

Logický součin (konjunkce)

Slovní vyjádření: a

Anglický název: AND

Operátor: $\bullet$, $\cap$

Vzorec: $y = a \bullet b$ ($y = a \text{ AND } b$)

Tabulka:

a	b	y
0	0	0
0	1	0
1	0	0
1	1	1

Výlučný (exkluzivní) součet

Slovní vyjádření: právě jedna z n

Anglický název: XOR

Operátor: +, $\cup$

Vzorec: $y = a \oplus b$

Tabulka:

a	b	y
0	0	0
0	1	1
1	0	1
1	1	0

Další dvě důležité funkce z hlediska prvků elektronických logických obvodů jsou negovaný součin a součet.

Negovaný logický součin (Shefferův funktor)

Slovní vyjádření: opak součinu

Anglický název: NAND

Vzorec: $y = \overline{a \bullet b}$ ($y = a \text{ NAND } b$)

Tabulka:

a	*b*	*y*
0	0	1
0	1	1

Negovaný logický součet (Piercův funktor)

Slovní vyjádření: opak součtu

Anglický název: NOR

Vzorec: $y = \overline{a + b}$ ($y = a \text{ NOR } b$)

Tabulka:

a	*b*	*y*
0	0	1
0	1	0

1	0	1
1	1	0

1	0	0
1	1	0

Pro úplnost jsou uvedeny poslední dvě funkce.

Shoda (ekvivalence)

Slovní vyjádření: identita

Anglický název: EQ

Vzorec: $y = a \text{ EQ } b$

Tabulka:

a	*b*	*y*
0	0	1
0	1	0
1	0	0
1	1	1

Neshoda (nonekvivalence)

Slovní vyjádření: opak identity

Anglický název: NEQ

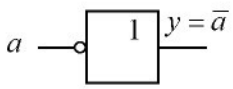
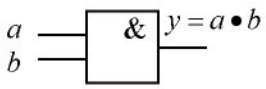
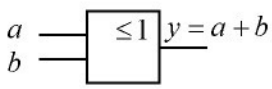
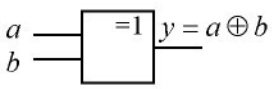
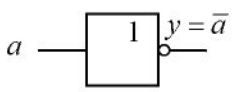
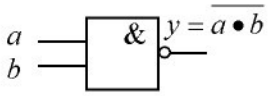
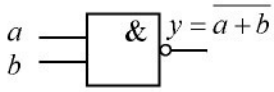
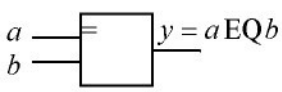
Vzorec: $y = a \text{ NEQ } b$

Tabulka:

a	*b*	*y*
0	0	0
0	1	1
1	0	1
1	1	0

Logické funkce se vyskytují v popisech logických obvodů. Speciální značky mají jen bloková schémata. Obsah značky napoví funkci obvodu. Pro AND je užíváno &. Logický součet naznačuje, že pro výstup logické jednotky je třeba, aby aspoň jeden vstup byla logická jednotka. Značky jsou dvojí, nové a zastaralé, které budou uvedeny pod značkou novou v následující tabulce. Počet vstupů jednoho bloku není omezen. Podle elektronických obvodových součástek se však používají prvky se dvěma, čtyřmi, osmi vstupy. Značky znázorněné v tabulce 1.2 budou se dvěma vstupy.

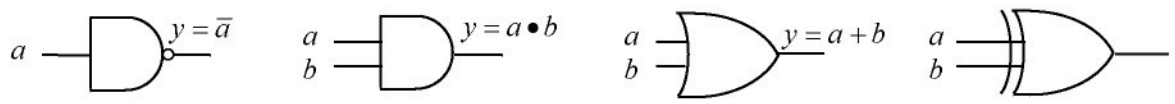
Tabulka 1.2 – Značky pro logické funkce (nové)

Negace vstupu NOT	Logický součin AND	Logický součet OR	Výlučný součet XOR
			
Negace výstupu NOT	**Negovaný součin NAND**	**Negovaný součet NOR**	**Ekvivalence EQ**
			

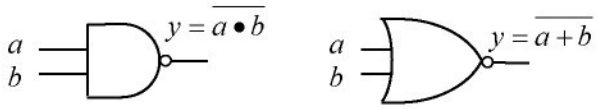
Při studiu staré dokumentace je třeba znát také dříve používané značení jednotlivých bloků jak ukazuje tabulka 1.3.

Tabulka 1.3 – Značky pro logické funkce (staré)

Negace výstupu NOT Logický součin AND Logický součet OR Výlučný součet XOR



Negovaný součin NAND Negovaný součet NOR



❑ Zákony Booleovy algebry

Základními zákony Booleovy algebry je zákon komutativní o změně pořadí jednotlivých proměnných

$$a+b=b+a\,,\ a\bullet b=b\bullet a\,, \tag{1.3}$$

asociativní o závorkách

$$a+(b+c)=(a+b)+c\,,\ a\bullet(b\bullet c)=(a\bullet b)\bullet c \tag{1.4}$$

a distributivní o vytýkání proměnných před závorku

$$a\bullet b+a\bullet c=a\bullet(b+c),\ (a+b)\bullet(a+c)=a+b\bullet c\,. \tag{1.5}$$

Důležitým je také zákon o dvojí negaci

$$\overline{\overline{a}}=a\,. \tag{1.6}$$

Další skupinou vztahů důležitých při úpravách vzorců s logickými proměnnými ukazuje tabulka 1.4 s duálními páry představující záměnu logického součtu za součin a záměnu logických 0 za 1 a naopak. Principu duality vyhovují také již uvedené základní zákony“ komutativní, asociativní a distributivní.

Nejčastěji jsou k úpravám vzorců a to zejména při převodu logického součtu na součin a naopak použity zákony (pravidla) de Morganovy v kombinaci se zákonem dvojí negace.

$$a+b=\overline{\overline{a+b}}=\overline{\overline{a}\bullet\overline{b}}\,,\ a\bullet b=\overline{\overline{a\bullet b}}=\overline{\overline{a}+\overline{b}} \tag{1.7}$$

Zákony lze ověřit s využitím definicí jednotlivých funkcí. Protože správnost de Morganových zákonů není na první pohled zřejmá, bude předveden postup, jak je ověřit. Nejjednodušší postup je kontrola platnosti pro všechny možné kombinace vstupních proměnných. V důsledku dvou hodnot, které každá proměnná může nabýt, je počet kombinací konečný, tj. 4. Nejprve je vyčíslena pravá strana vzorců a pak jejich levá strana. Ověření je dokončeno porovnáním příslušných sloupců (v tabulce 1.5 jsou vyznačeny tučně).

Tabulka 1.4 – Duální páry zákonů

Název zákona	Logický součet	Logický součin	Číslo rovnice
vyloučení třetího	$a+\overline{a}=1$	$a \bullet \overline{a}=0$	(1.7)
neutrálnost	$a+0=a$	$a \bullet 1=a$	(1.8)
agresivnost	$a+1=1$	$a \bullet 0=0$	(1.9)
idempotence	$a+a=a$	$a \bullet a=a$	(1.10)
absorpce	$a+a \bullet b=a$	$a \bullet (a+b)=a$	(1.11)
	$a+\overline{a} \bullet b=a+b$	$a \bullet (\overline{a}+b)=a \bullet b$	(1.12)
de Morganovy	$\overline{a+b}=\overline{a} \bullet \overline{b}$	$\overline{a \bullet b}=\overline{a}+\overline{b}$	(1.13)

Tabulka 1.5 – Ověření de Morganových zákonů

a	b	$\overline{a}$	$\overline{b}$	$a+b$	$\overline{a+b}$	$\overline{a} \bullet \overline{b}$	$a \bullet b$	$\overline{a \bullet b}$	$\overline{a}+\overline{b}$
0	0	1	1	0	**1**	**1**	0	**1**	**1**
0	1	1	0	1	**0**	**0**	0	**1**	**1**
1	0	0	1	1	**0**	**0**	0	**1**	**1**
1	1	0	0	1	**0**	**0**	1	**0**	**0**

❑ Syntéza kanonického tvaru kombinační logické funkce

Tabulka hodnot logické funkce se sestaví podle požadované funkce logického řízení. Dalším krokem analýzy funkce je odvození vzorce, který tabulkovým hodnotám logických proměnných vyhovuje. Podle principu duality lze definovat dva tvary (tzv. kanonické tvary dané aplikaci zvoleného pravidla), které jsou vytvořeny jako

- součet součinů (úplná disjunktní normální forma)
- součin součtu (úplná konjunktivní normální forma).

Dualita spočívá ve vzájemné záměně logického součtu a součinu. Postup vytvoření vzorce bude vysvětlen na příkladu.

Příklad 1.1. *Příklad na určení vzorce logické funkce zadané pravdivostní tabulkou.*

Nechť je dána část tabulky funkčních hodnot logické funkce y. Určete vzorec logické funkce.

a	b	c	y
...	...	...	...
1	1	0	1
1	1	1	0

Řešení:

A) Ve tvaru součtu součinů se logická funkce vytváří podle hodnot 1 logické funkce. Hodnotu 1 nabývá logická funkce mimo jiné kombinace pro $a = 1$, $b = 1$ a $c = 0$. V součtu dílčích příspěvků k hodnotě 1 logické funkce pro jiné kombinace hodnot nezávisle proměnných (určujících hodnotu funkce 1) musí být přítomen logický součin $a \bullet b \bullet \overline{c}$, který je pro tuto kombinaci proměnných roven 1, proto část logické funkce má tvar

$$y = ... + a \bullet b \bullet \overline{c} + ... \qquad (1.8)$$

Každý sčítanec součtu obsahuje součin všech logických proměnných, které jsou vhodně negovány tak, aby jejich součin byl jednotkový.

B) V duálním tvaru k případu ad A), tj. součinu součtů, se logická funkce vytváří podle hodnot 0 logické funkce. Hodnotu 0 nabývá logická funkce, jestliže aspoň jeden ze součinitelů v součinu je nulový. Pro danou kombinaci nezávisle proměnných $a = 1$, $b = 1$ a $c = 1$ je součet nulový, jestliže je tvořen z negací všech proměnných, tj. $\overline{a} + \overline{b} + \overline{c}$, proto logická funkce musí nabýt tvar

$$y = (...) \bullet ... \bullet (\overline{a} + \overline{b} + \overline{c}) \bullet ... \bullet (...) \qquad (1.9)$$

Každý součinitel součinu obsahuje součet, jehož hodnota je rovna logické nule pro kombinaci hodnot vstupních logických proměnných, kterým odpovídá v tabulce nulová hodnota logické funkce.

Příklad 1.2. ***Příklad na určení vzorce logické funkce zadané pravdivostní tabulkou.***

Nechť je logická funkce dána tabulkou. Určete vzorec funkce jak pomocí součtu součinů, tak i pomocí součinu součtů.

Řešení:

a	*b*	*c*	*y*
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

A) Součet součinů

$$y = \overline{a} \bullet \overline{b} \bullet \overline{c} + \overline{a} \bullet \overline{b} \bullet c + \overline{a} \bullet b \bullet \overline{c} + \overline{a} \bullet b \bullet c + a \bullet \overline{b} \bullet c + a \bullet b \bullet c$$

B) Součin součtů

$$y = (\overline{a} + b + c) \bullet (\overline{a} + \overline{b} + c)$$

Úpravy výrazů:

Ad A)

Z dvojic výrazů lze vytknout a uplatnit vzorec (1.7)

$$\overline{a} \bullet \overline{b} \bullet \overline{c} + \overline{a} \bullet \overline{b} \bullet c = \overline{a} \bullet \overline{b} \bullet (\overline{c} + c) = \overline{a} \bullet \overline{b}$$
$$\overline{a} \bullet b \bullet \overline{c} + \overline{a} \bullet b \bullet c = \overline{a} \bullet b \bullet (\overline{c} + c) = \overline{a} \bullet b$$
$$a \bullet \overline{b} \bullet c + a \bullet b \bullet c = a \bullet c \bullet (\overline{b} + b) = a \bullet c$$

$$y = \overline{a} \bullet \overline{b} + \overline{a} \bullet b + a \bullet c = \overline{a} + a \bullet c = \overline{a} + c$$

Ad B)

Vhodným uzávorkováním a uplatněném vzorce (1.5)

$$y = ((\overline{a} + c) + b) \bullet ((\overline{a} + c) + \overline{b}) = \overline{a} + c + b \bullet \overline{b} = \overline{a} + c$$

Oba postupy dávají shodný výsledek.

❑ Sestavení logické funkce pomocí Karnaughovy mapy

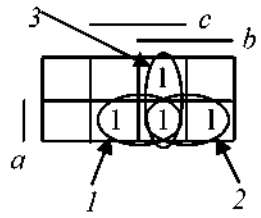
Sestavování vzorce v předcházející kapitole s využitím varianty součtu součinu vedlo na zápis dodržující určitá pravidla, avšak výsledný vzorec obsahuje zbytečně mnoho členů. Jeho zápis vyžadoval úpravy. K těmto úpravám byl využit vzorec $a + \overline{a} = 1$. Shodou okolností nalezení dvou členů umožňujících vytknutí součinu dvou proměnných, které obsahují zbylou proměnnou a její negaci (např. a a $\overline{a}$), nebylo těžké. Vhodné skupiny součinů tří proměnných byly v součtu vedle sebe.

Úpravy vzorců s dvouhodnotovými logickými proměnnými lze s pomocí Karnaughovy mapy algoritmizovat. Tato mapa je sestavena tak, že dvojice sousedních polí ve směru řádků

nebo sloupců se liší v hodnotě jedné logické proměnné. Jestliže kombinacím hodnot logických proměnných v sousedních polích mapy odpovídají logické jedničky, pak v součtu součinů se vyskytují dva členy, které se liší jen v negaci této logické proměnné. Tyto dva členy součtu lze nahradit členem jedním s počtem proměnných menším o jednotku. Z výrazu se vyloučí logická proměnná, která se liší v použití operace negace.

Algoritmus bude demonstrován na příkladu zápisu logické funkce pomocí Karnaughovy mapy. Následující příklad vede na kanonický tvar součtu 4 členů se součiny 3 logických proměnných

$$y = a \bullet \overline{b} \bullet c + a \bullet b \bullet c + a \bullet b \bullet \overline{c} + \overline{a} \bullet b \bullet c$$



Na mapě vlevo lze určit tři dvojice jedniček v sousedních polích mapy ve směru řádků a sloupců, které jsou označeny *1*, *2* a *3*. Tyto kombinace proměnných nahrazují členy

$$1\text{: } a \bullet \overline{b} \bullet c + a \bullet b \bullet c = a \bullet c \bullet \left(\overline{b} + b\right) = a \bullet c$$

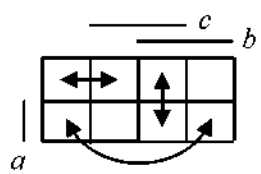
$$2\text{: } a \bullet b \bullet c + a \bullet b \bullet \overline{c} = a \bullet b \bullet \left(\overline{c} + c\right) = a \bullet b$$

$$3\text{: } a \bullet b \bullet \overline{c} + \overline{a} \bullet b \bullet c = b \bullet c \bullet \left(a + \overline{a}\right) = b \bullet c$$

Výsledný vzorec pro logickou funkci definovanou Karnaughovou mapou je následující

$$y = a \bullet c + a \bullet b + b \bullet c$$

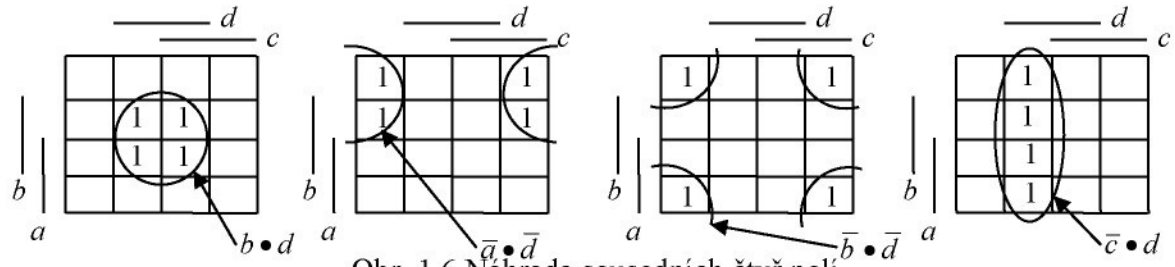
Zápis této funkce je oproti kanonickému tvaru zjednodušen a jeho realizace technickými prostředky bude jednoduchá. Karnaughova mapa je výhodným nástrojem k optimalizaci zápisu logických funkcí.



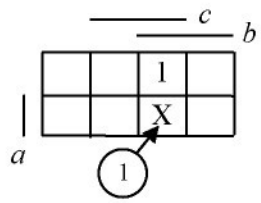
Sousední pole

Sousední pole v mapě ve významu odlišení o hodnotu jedné logické proměnné spolu nemusí bezprostředně sousedit. Také pole na protilehlých okrajích tabulky ve směru svislém nebo vodorovném se liší také jen v hodnotě jedné logické proměnné, proto jsou považovány také za sousední. Například ve vedlejším obrázku pole odpovídající hodnotám $a = 1$, $b = 0$ a $c = 0$ a pole s hodnotami $a = 1$, $b = 1$ a $c = 0$ se liší jen v hodnotě logické proměnné b.

Jestliže čtyři sousední pole jsou obsazena v Karnaughově mapě logickými jedničkami, pal lze ukázat, že v součtu součinů lze nahradit čtyři členy členem jedním se sníženým počtem logických proměnných o dvě. Příklad pro čtyři proměnné je na obrázku 1.6.



Obr. 1.6 Náhrada sousedních čtyř polí



Poslední poznámka k sestavení logické funkce se tyká polí s nedefinovanou hodnotou logické funkce. Tato pole Karnaughovy mapy se označují písmenem X. Nedefinované pole se doplňuje funkční hodnotou, která zjednoduší výslednou funkci. Na příkladu vlevo náhrada X nulou vede na zápis funkce $y = \bar{a} \bullet b \bullet c$, zatímco náhrada jedničkou zjednoduší zápis funkce na tvar $y = b \bullet c$.

Shrnutí pojmů 1.1

V této kapitole se studenti seznámili s úvodními pojmy z oblasti logického řízení, tzn. umí definovat logickou proměnnou, funkci logické proměnné a její možné zadání ať už vzorcem nebo tabulkou hodnot (pravdivostní tabulkou) nebo Karnaughovou mapou. Dále jsou zde popsány základní logické operace jako logický součet či součin, negace, atd. a postupy minimalizace logických funkcí buď pomocí zákonů Boolovy algebry nebo pomocí Karnaughovy mapy. Na závěr je ukázán převod jednotlivých forem zadání logické funkce mezi sebou (vzorec, tabulka, K mapa).

Otázky 1.1.

1. Co je to logické řízení?
2. Jak je definována logická funkce?
3. Co je to NAND a NOR?
4. Jak vypadá značka negace NOT?
5. Na co se používá Boolova algebra?
6. Jak vypadají De Morganovy zákony?
7. Můžeme v Karanughově mapě sdružovat 3 políčka?
8. Může obsahovat Karnaughova mapa nedefinované hodnoty?
9. Muže být logická funkce zadána slovně?
10. Pomocí jakých vyjádření můžeme vypsat logickou funkci z pravdivostní tabulky?

Úlohy k řešení 1.1.

1. Minimalizujte logickou funkci čtyř proměnných zadanou pravdivostní tabulkou pomocí Karnaughovy mapy.

a	b	c	d	y
0	0	0	0	0
0	0	0	1	1
0	1	0	1	1
0	1	1	1	1
1	1	0	1	1
1	1	1	0	1
1	1	0	0	1
1	1	1	0	1
1	1	0	0	1
1	1	1	1	0
1	0	1	0	0

2. Minimalizujte logickou funkci pěti proměnných zadanou pravdivostní tabulkou pomocí Karnaughovy mapy.

a	b	c	d	e	y
0	0	0	0	0	0
0	1	0	0	0	1
0	1	1	0	0	1
0	1	1	1	0	1
0	1	0	1	0	1
0	1	0	1	1	1
0	1	1	1	1	1

a	b	c	d	e	y
0	1	1	0	1	1
0	1	0	0	1	1
1	1	1	1	0	1
1	1	0	1	0	1
1	1	1	1	1	1
1	0	1	0	1	1
1	0	0	0	1	1
1	1	1	0	0	0

1.2 Realizace kombinačních logických obvodů

Čas ke studiu: 2 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat realizaci a syntézu logických funkcí.
- popsat prky pro realizaci logických funkcí.
- vyřešit minimalizaci a realizaci logických funkcí.

Výklad

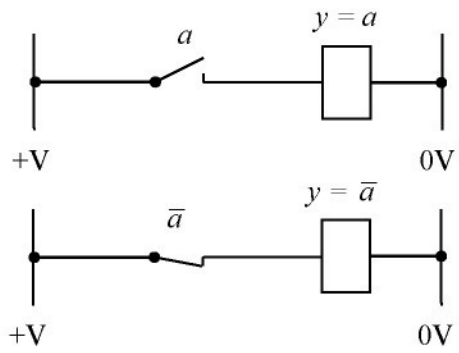
Kombinační logické funkce lze realizovat technickými prostředky, které jsou založeny na následujících technických prvcích.

- Mechanická relé
- Integrované obvody
- Programovatelné logické automaty.

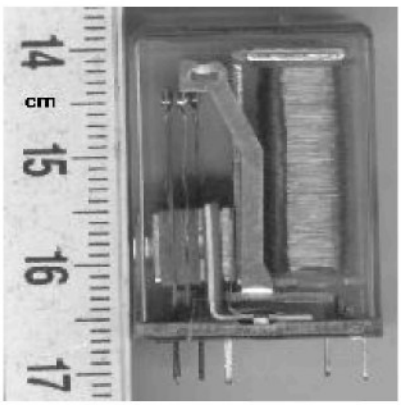
Každému z těchto prvků odpovídá specifický způsob znázorňování schéma zapojení těchto prvků. Reléové obvody se znázorňují v liniových kontaktních schématech. Elektronické logické prvky využívají blokových schémat a funkce programový logických automatů je popsána některou formou dokumentace software.

Kontaktní schémata

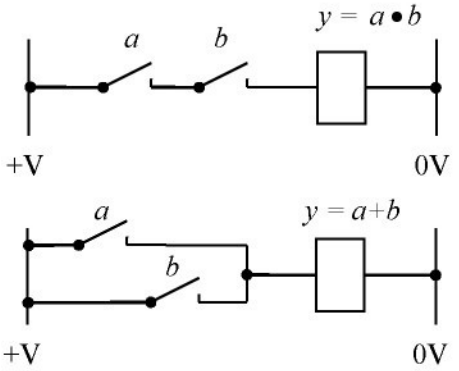
Jiné označení je liniové schéma. Toto schéma znázorňuje zapojení elektrických kontaktů, které umožňují průchod elektrického proudu mezi dvěma póly zdroje obvykle kreslenými jako svislé vodiče. Mezi těmito vodiči prochází proud přes různě zapojené kontakty do zátěže. Na pravé straně je kladné napájecí napětí a levá strana je obvykle uzemněna. Zátěží může být cívka elektromagnetického relé nebo jiný spotřebič, například žárovka nebo houkačka. Kontakty relé dimenzované na potřebné proudové zatížení mohou být napojeny i na elektrický motor nebo soustavu osvětlení. Nejjednodušší logické funkce $y = a$ a $y = \bar{a}$ jsou znázorněny kontaktními schématy na obrázku 1.6. Kontakty mají shodná označení jako logické proměnné. Jestliže je logická proměnná rovna logické jedničce, tak je spínací kontakt sepnut a pro logickou nulu rozepnut. Kontaktů se stejným označením může být ve schématu více. Jak je zřejmé z obrázku 1.6, je rozlišena negovaná proměnná, která přísluší rozpínacímu kontaktu. Kontakty označené stejnou proměnnou, včetně její negace, jsou ovládány jedním relé, které je obvykle vybaveno několika kontakty různého druhu (spínací, rozpínací).



Obrázek 1.7 - Kontaktní schéma realizující logické funkce $y = a$ a $y = \bar{a}$



Obrázek 1.8 - Relé s přepínacími kontakty



Obrázek 1.9 – Kontaktní schéma realizující logický součin a součet

Funkci logického součinu a součtu plní zapojení kontaktů na obrázku 1.9. Logický součin umožňuje průchod proudu zátěží jen když jsou oba kontakty sepnuty, tj. příslušné proměnné jsou rovny logické jedničce. Naproti tomu u logického součtu stačí jen jeden sepnutý kontakt, aby proud protékal zátěží. Obrázky na schématech jsou zobrazena vektorovou počítačovou grafikou. Dříve, kdy displeje počítačů umožňovaly jen textový režim, bylo běžné zobrazování spojení ze znaků. Příkladem spínacího kontaktu a zátěže je následující posloupnost znaků

|-----] [-----(=)-----|

rozpínací kontakt je znázorněn takto

|-----]/[-----(=)-----|

Příklad 1.3. ***Příklad na určení kontaktního schématu logické funkce zadané vzorcem.***

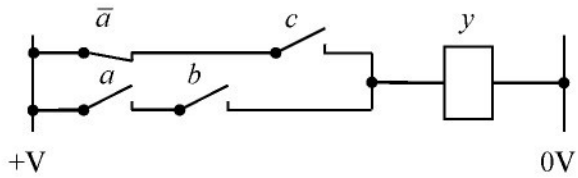
Vytvořte kontaktní schéma logické funkce dané vzorcem (1.10).

Řešení:

Kontaktní schéma logické funkce dané vzorcem (1.10) je znázorněno na obrázku 1.10.

$$y = \overline{a} \bullet c + a \bullet b \qquad (1.10)$$

Na kontaktním schématu jsou kresleny kontakty příslušné jedné logické proměnné být pod sebou proto, aby bylo demonstrováno jejich společné ovládání příslušným jedním relé.



Obrázek 1.10 – Řešení příkladu

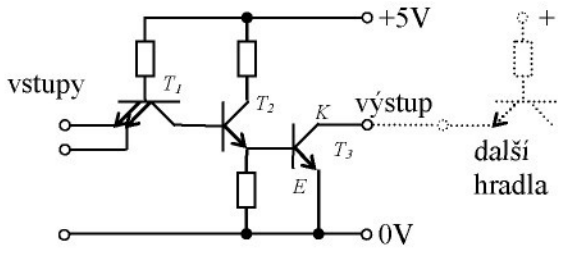
Bloková schémata

Tento typ schématu odpovídá zapojení elektronických součástek plnících logické funkce. Tyto součástky se nazývají hradla. Jsou především typu NAND a NOR. Signál logické jedničky v tzv. pozitivní logice TTL (Tranzistor-Tranzistor-Logic) je kladné elektrické napětí +5V, naproti tomu logické nule odpovídá 0V. K označení úrovně napětí pro logickou jedničku se používá H (High) a L (Low) pro logickou nulu. Samozřejmě, že je přípustná tolerance velikosti napětí. U integrovaných obvodů (IO) řady MH74000 je úroveň L do 0,8 V a H nad 2V. Mezi těmito hranicemi není obvykle chování hradel předvídatelné a vznikají hazardy (neurčité stavy).

Logická hradla se integrují do integrovaných obvodů (IO) a několik z nich tvoří jednu součástku. Počet vstupů těchto hradel je obvykle 2, 4 nebo 8 a podle tohoto počtu vstupů pro jedno hradlo se s ohledem na počet dostupných vývodů (pinů) slučují na jeden integrovaný obvod. Příklad hradla typu NAND se dvěma vstupy a třemi tranzistory typu NPN je na obrázku 1.11. Výstupem tohoto obvodu je třetí tranzistor, který se připojuje k dalšímu hradlu. Základem tohoto obvodu je tranzistor pracující ve spínacím režimu. Jak známo, tranzistor je zesilovač proudu. Velikost proudu kolektorem (K) je mnohonásobně zesílený proud do báze (B). Proud přechodem báze-emitor je dán napětím na tomto přechodu tranzistoru. Zhruba řečeno: kladné napětí na bázi oproti editoru otvírá tranzistor typu NPN pro průchod proudu, zatímco nulové nebo záporné napětí na tomto přechodu tranzistor typu NPN pro proud kolektorem uzavírá.

Pro obvod na obrázku 1.11 platí, že nulové napětí na jednom nebo obou vstupech obvodu otevře tranzistor T_1 průchodu proudu a napětí na jeho kolektoru poklesne k úrovni 0V, což uzavře tranzistor T_2 , jehož kolektorem přestane téci proud a na odporech, které jsou připojeny ke kolektoru a editoru, bude velmi malý úbytek napětí. Napětí na editoru tranzistoru

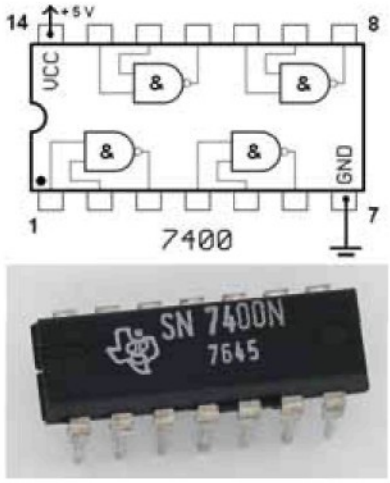
T_2 je téměř na napětí 0V, což uzavře tranzistor T_3, jehož kolektor je tzv.otevřený vzhledem k popisovanému obvodu. Kolektor tohoto tranzistoru T_3 je totiž na napájecí zdroj připojen až přes editor a odpor v bázi dalšího připojeného obvodu [hradla].



Obrázek 1.11 – Hradlo NAND se dvěma vstupy

Protože přes tento třetí tranzistor T_3 neteče žádný proud, je na výstupu téměř plné napájecí napětí. Shrnuto: nulové napětí na kterémkoliv vstupu tohoto obvodu (logická nula) vyvolá napětí blízké napájecímu na jeho výstupu. Jestliže jsou oba vstupy na úrovni napájecího napětí, pak tranzistor T_1 se zavře a jeho zbytkový proud otevře tranzistor T_2, což vyvolá zvýšení úbytku napětí na odporu, který je připojen k jeho emitoru. Tento úbytek napětí otevře průchodu proudu tranzistor T_3 s potenciálem na jeho kolektoru blízkým nule. Hradlo tedy plní funkci prvku NAND. Maximální počet vstupů jiných integrovaných obvodů, které lze připojit na výstup jednoho hradla, se nazývá logický zisk obvodu. U zmíněné řady IO je logický zisk 10, tj. na výstup tohoto hradla může být zapojeno až 10 dalších hradel a proudové zatížení výstupu je v přípustných mezích.

Ukázka integrovaného obvodu SN 7400N, který obsahuje čtyři hradla NAND se dvěma vstupy, je na obrázku 1.12. Kromě zapojení vývodů je přiloženo také foto tohoto integrovaného obvodu.



Obrázek 1.12 – Integrovaný obvod SN 7400N

Základními operacemi jsou negace, součet a součin. Způsob zapojení hradel NOR a NAND pro tyto operace je předmětem dalšího popisu. Zapojení hradel se dvěma vstupy je znázorněno na obrázku 1.13.

Negace jedné proměnné hradlem NAND nebo NOR se řeší jedním hradlem s propojenými vstupy. Podle výše vysvětlené funkce by se volný vstup choval jako připojená logická jednička, což by nebylo vhodné pro hradlo NOR.

Součin dvou proměnných lze zrealizovat dvěma hradly NAND s použitím vzorce

$$y = a \bullet b = \overline{\overline{a \bullet b}} = \left(\overline{a \bullet b}\right) \tag{1.11}$$

Dvojnásobná negace odpovídá postupným operacím: $\overline{a \bullet b}$ a pak opakovaná negace tohoto dílčího výsledku. Součet proměnných se vytvoří podle tohoto vzorce

$$y = a + b = \overline{\overline{a + b}} = \overline{\overline{a} \bullet \overline{b}} \tag{1.12}$$

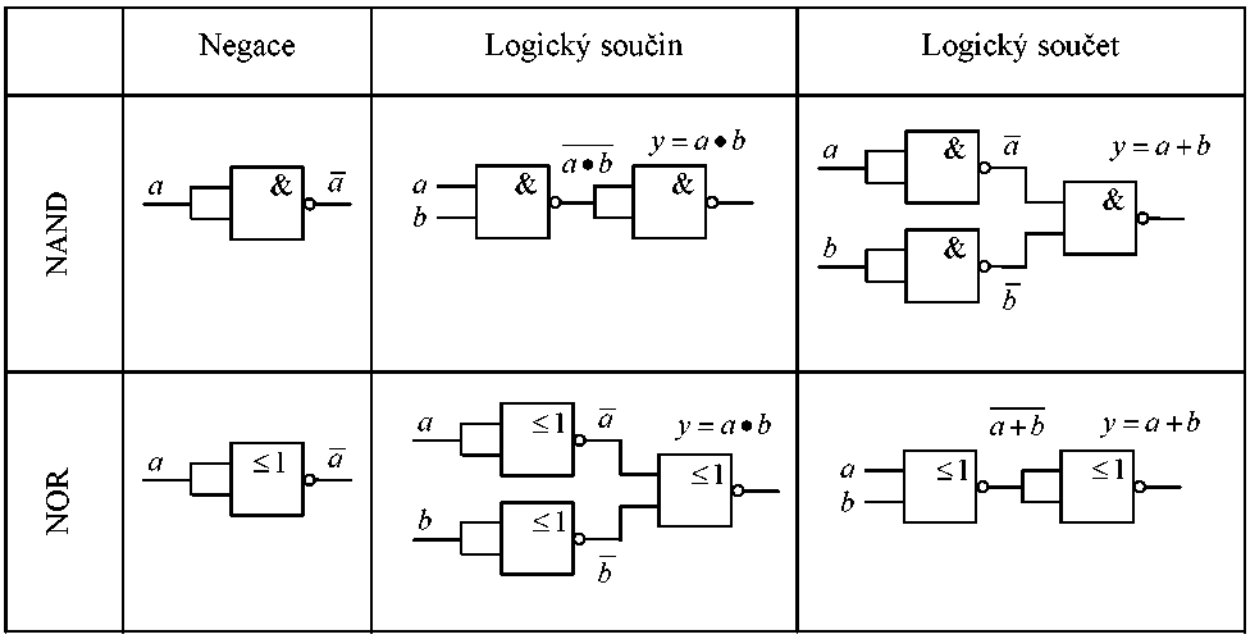
Dvojnásobná negace součtu se rozloží použitím de Morganova zákona na součin negovaných proměnných, který je opět negován.

Součin dvou logických proměnných, který je realizován dvěma hradly NOR, používá vzorce

$$y = a \bullet b = \overline{\overline{a \bullet b}} = \overline{\overline{a} + \overline{b}} \tag{1.13}$$

Součet proměnných se vytvoří jako dvojitá negace

$$y = a + b = \overline{\overline{a + b}} = \left(\overline{a + b}\right) \tag{1.14}$$



Obrázek 1.13 – Zapojení hradel NAND a NOR pro negaci, součin a součet

Jak již bylo zmíněno, počet vstupů hradel je 2 nebo 4 a nebo 8. Příklady k řešení jsou zadány tak, že se požaduje použít hradla s jednotným počtem vstupů, nejčastěji dvou. Nejprve bude demonstrován postup realizace logického součtu a součinu tří proměnných *a, b, c* hradly typu NAND se dvěma vstupy. Postup popisují následující vzorce. Ze tří proměnných se vytvoří dvě skupiny. První obsahuje dvě proměnné (*a, b*) a druhá jednu (*c*). Výsledek výpočtu obou těchto skupin může vstupovat do uvedeného typu hradla s počtem vstupů, který je roven počtu zmíněných skupin. V následujících vzorcích jsou mezivýsledky, které odpovídají výstupu jednotlivých hradel, označeny písmenem *H* s indexem:

$$y = a + b + c = (a + b) + c = \overline{\overline{(a + b)} + c} = \overline{\overline{(a \bullet b)} + c} = \overline{\overline{(\bar{a} \bullet \bar{b})} \bullet \bar{c}} = \overline{\overline{(\bar{a} \bullet \bar{b})} \bullet \bar{c}} =$$
$$= \overline{\overline{(H_1 \bullet H_2)} \bullet H_3} = \overline{\bar{H}_4 \bullet H_3} = \overline{H_5 \bullet H_3} = H_6$$

$$y = a \bullet b \bullet c = (a \bullet b) \bullet c = \overline{(a \bullet b)} \bullet c = \overline{\left(\overline{(a \bullet b)} \bullet c\right)} = \overline{\left(\overline{(H_1)} \bullet c\right)} = \overline{(H_2 \bullet c)} = \overline{(H_3)} = H_4$$

Logický součet a součin tří proměnných se realizuje hradly typu NOR se dvěma vstupy podle následujících vzorců, ve kterých lze uplatnit výše popsaný princip duality spočívající v náhradě logického součtu součinem a naopak:

$$y = a + b + c = (a + b) + c = \overline{(a + b)} + c = \overline{\left(\overline{(a + b)} + c\right)} = \overline{\left(\overline{(H_1)} + c\right)} = \overline{(H_2 + c)} = \overline{(H_3)} = H_4$$

$$y = a \bullet b \bullet c = (a \bullet b) \bullet c = \overline{\overline{(a \bullet b)} \bullet c} = \overline{\overline{(\bar{a} + \bar{b})} \bullet c} = \overline{\overline{(\bar{a} + \bar{b})} + \bar{c}} = \overline{\overline{(\bar{a} + \bar{b})} + \bar{c}} =$$
$$= \overline{\overline{(H_1 + H_2)} + H_3} = \overline{\bar{H}_4 + H_3} = \overline{H_5 + H_3} = H_6$$

❑ Programovatelné logické automaty

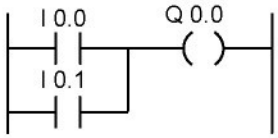
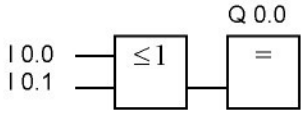
Programovatelné logické automaty představují speciální počítač. Jejich funkce se řídí programem, což je posloupnost instrukcí provádějící aritmeticko-logické operace a přesuny mezi svou operační pamětí, aritmeticko- logickou jednotkou (ALU) a vnějším prostředím.

V praxi se používají tři způsoby programování logických automatů pro kombinační logické úlohy. Jsou to jazyky

- liniových, resp. kontaktních schémat (Ladder Diagram – LAD), resp. Kontaktplan – KOP)
- blokových schémat logické funkce (Control System Flowchart – CSF, Funktions plan – FUP)
- symbolických instrukcí (STATMENT LIST – STL, resp. Anweisungsliste – AWL)

Prví dva způsoby využívají grafiku, zatímco třetí je založen textově. V tabulce 1.6 je ukázka uvedených způsobů programování logických automatů na příkladu logického součtu dvou logických proměnných. Jména vstupních logických proměnných jsou ve skutečnosti označení symbolických adres I 0.0 a I 0.1, důležitých pro obsazení paměti počítače. Výsledek je uložen na adrese Q 0.0. Všechny tři proměnné jsou typu BOOL (Boolean). U blokového schématu je použit blok výsledku s adresou nad jeho značkou. Jazyk symbolických instrukcí používá posloupnost příkazů v řádcích, ve kterých na prvním místě je typ operace a na druhém místě operand, což je u příkazu použitých v příkladu adresa, ale neplatí to obecně. V příkladu je typ operace označen symbolem O (OR), tj. logický součet. Pro logický součin se používá zkratka A (AND). Symbol = označuje uložení výsledků na adresu, která je operandem tohoto příkazu.

Tabulka 1.6 – Způsoby programování logických automatů

Jazyk kontaktních schémat	Jazyk blokových schémat	Jazyk symbolických instrukcí
		O I 0.0 O I 0.1 = Q 0.0

Příklad 1.4. ***Příklad na syntézu kombinačního logického obvodu***

Zadáním příkladu je odvodit minimální tvar zápisu logické funkce, vytvořit kontaktní schéma a bloková schémata buď jen s hradly NAND nebo NOR, přičemž počet jejích vstupů je omezen na dva.

Řešení:

Počáteční krok řešení úlohy kombinační logiky je sestavení tabulky funkce na základě slovních požadavků na její funkci. Nechť je požadováno, aby výstup kombinačního logického obvodu *y* závisel na vstupech *a, b, c, d* podle následující tabulky, která je rozdělena do dvou částí. V tabulce jsou dvě kombinace vstupních logických proměnných, pro které není funkční hodnota *y* definována – jejich pole je prázdné. Tato kombinace ve slovním zadání chyběla, protože byla nesmyslná.

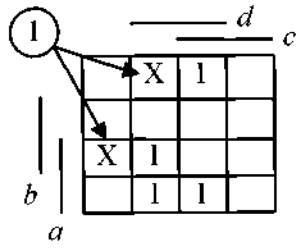
První část tabulky

a	*b*	*c*	*d*	*y*
0	0	0	0	0
0	0	0	1	
0	0	1	0	0
0	0	1	1	1
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0

Druhá část tabulky

a	*b*	*c*	*d*	*y*
1	0	0	0	0
1	0	0	1	1
1	0	1	0	0
1	0	1	1	1
1	1	0	0	
1	1	0	1	1
1	1	1	0	0
1	1	1	1	0

Druhým krokem syntézy je sestavení Karnaughovy mapy. Do této mapy jsou přeneseny jen logické jedničky a nedefinované funkční hodnoty z tabulky jsou označeny X. Za účelem optimalizace zápisu vzorce pro logickou funkci je vhodné, aby logická funkce byla dodefinována logickými jedničkami, protože vznikne dvojice sousedních logických proměnných a v zápise takto lze tyto dvojice součinů čtyř proměnných nahradit jedním součinem tří proměnných. Doplněním jedničky v prvním řádku tabulky dokonce vznikne čtveřice sousedních jedniček s efektem dalšího zjednodušení zápisu vzorce náhradou čtyř součinů čtyř proměnných jen jedním součinem dvou proměnných.

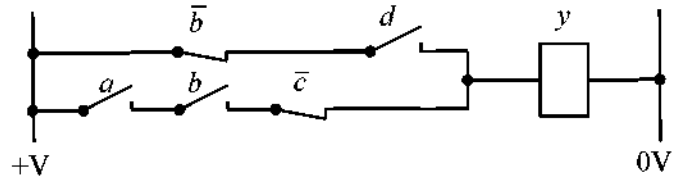


Karnaughova mapa

Třetím krokem je sestavení vzorce ve tvaru minimálního počtu členů součtu součinů logických proměnných. Tento tvar nelze dále zjednodušit

$$y = \overline{b} \bullet d + a \bullet b \bullet \overline{c}$$

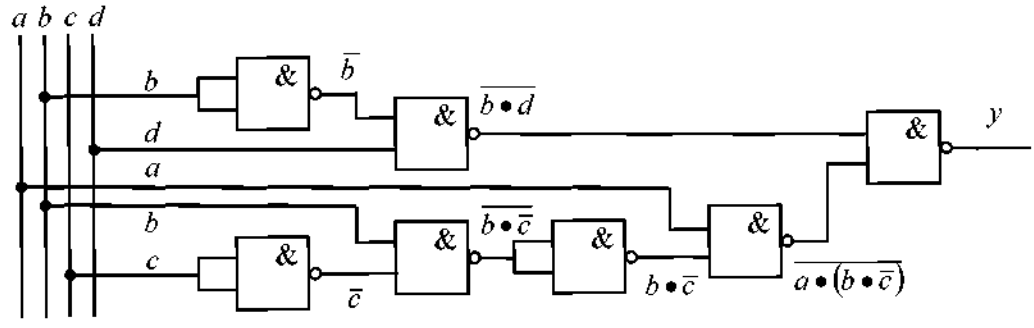
Čtvrtým krokem je sestavení kontaktního schématu. Kontakty *a, b, c, d* spínají elektrický proud do cívky relé *y*, jehož kontakty mohou sepnout přívod proudu k libovolnému dalšímu zařízení.



Pátým a také posledním krokem řešení příkladu je vytvoření blokových schémat pro hradla NAND a NOR. K tomuto účelu je třeba zápis vzorce upravit do tvaru obsahujícího negace a buď logické součty nebo logické součiny. Pro hradla NAND jsou úpravy následující

$$y = \overline{b} \bullet d + a \bullet b \bullet \overline{c} = \overline{\overline{\overline{b} \bullet d + a \bullet b \bullet \overline{c}}} = \overline{\overline{\overline{b} \bullet d} \bullet \overline{a \bullet (b \bullet \overline{c})}}$$

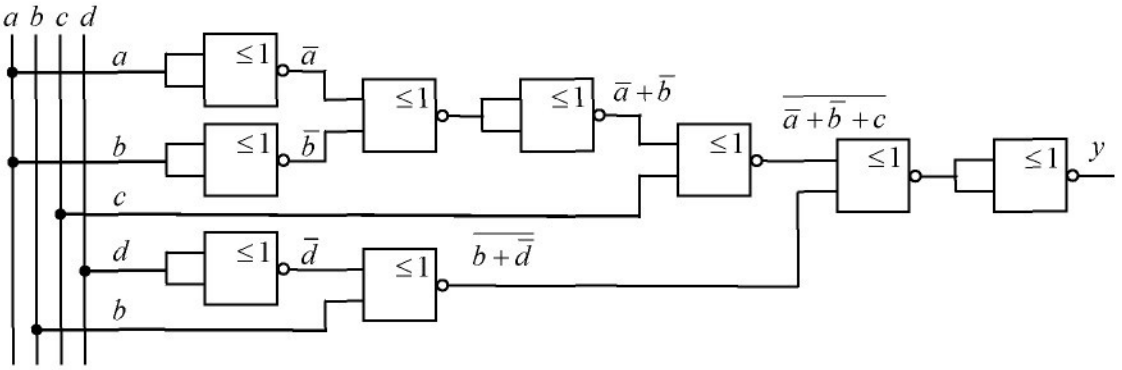
Příslušný blokový diagram sestavený z hradel NAND má následující podobu



Drobná modifikace zapojení by byla možná seskupením $(a \bullet b) \bullet \overline{c}$.

Dříve než bude blokové schéma s hradly typu NOR nakresleno, je třeba vzorec pro výpočet logické funkce upravit do tvaru, který obsahuje jen negace a součty

$$y = \overline{b} \bullet d + a \bullet b \bullet \overline{c} = \overline{\overline{b} \bullet d} + \overline{(a \bullet b) \bullet \overline{c}} = \overline{b + \overline{d}} + \overline{\overline{(a \bullet b)} + c} = \overline{b + \overline{d}} + \overline{(\overline{a} + \overline{b}) + c}$$



Příklad 1.5. ***Příklad na syntézu kombinačního logického obvodu***

Ze zadané pravdivostní tabulky vytvořte logickou funkci. Tu minimalizujte a nakreslete zapojení pomocí kontaktních prvků a logických členů NAND.

a	b	c	y
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

Řešení:

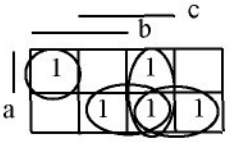
Logická funkce:

Vypíšeme z tabulky pomocí úplné disjunktní normální formy.

$y = \bar{a}\bar{b}\bar{c} + \bar{a}\bar{b}c + \bar{a}b\bar{c} + a\bar{b}\bar{c} + ab\bar{c}$

Karnaughova mapa:

Zapíšeme do Karnaughovy mapy jedničky podle logické funkce.



Z Karnaughovy mapy vypíšeme všechny seskupení jedniček. Mezi každým seskupením dáváme plus. Takto minimalizovanou logickou funkci můžeme ještě upravit pomocí Boolovy algebry.

$$y = \bar{a}\bar{b} + c\bar{b} + c\bar{a} + ab\bar{c} = \bar{b}(\bar{a} + c) + \bar{a}c + ab\bar{c} = \overline{\bar{b}(\bar{a} + c)(\bar{a}c + ab\bar{c})}$$

Kontaktní schéma:
Sestavíme ho podle minimalizované logické funkce a pomocí zásad, která jsou v přednášce.

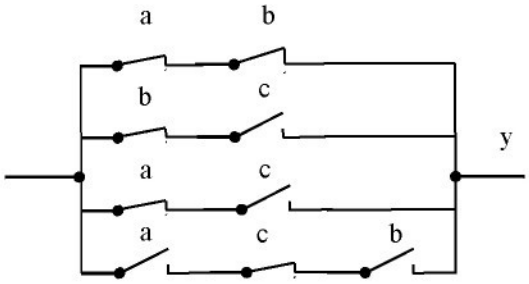
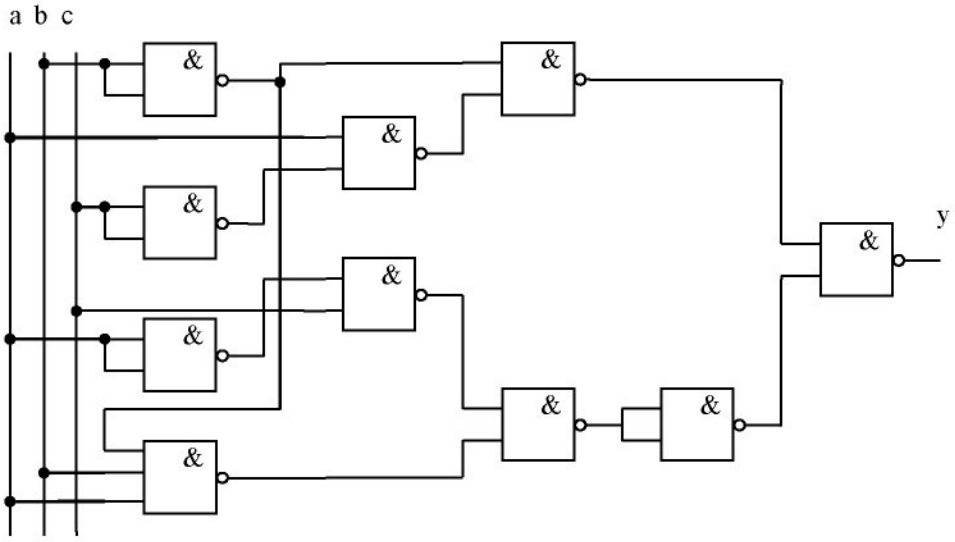
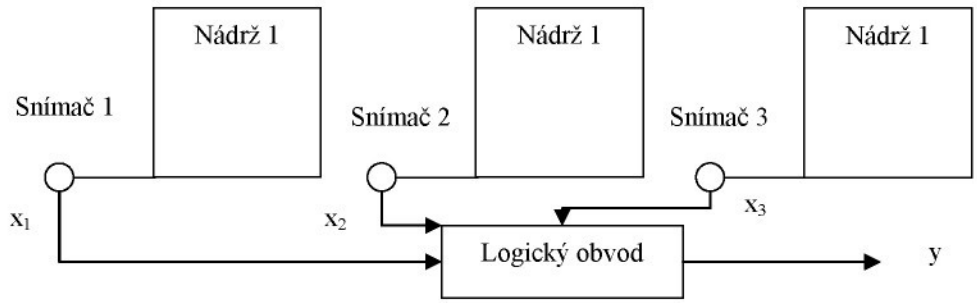


Schéma logických členů NAND:
Sestavíme ho podle minimalizované logické funkce a pomocí zásad, která jsou v přednášce.



Příklad 1.6. ***Příklad na syntézu kombinačního logického obvodu***

V dílně jsou tři nádrže provozního oleje. Jestliže dvě z nich jsou prázdné, musí se rozsvítit nápis DOPLNIT OLEJ!, aby obsluha provedla doplnění. Navrhněte příslušný logický obvod, sestavený z kontaktních prvků a logických členů NAND



Řešení:

Definice proměnných:

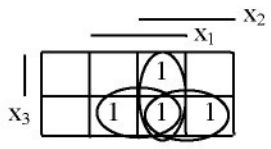
$x_1, x_2, x_3 = 0$ nádrž není prázdná $x_1, x_2, x_3 = 1$ nádrž je prázdná

$y=0$ nesvítí displej $y=1$ svítí displej

Pravdivostní tabulka:

x_1	x_2	x_3	y
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

Karnaughova mapa:



$$y = x_1x_2 + x_1x_3 + x_2x_3 = \overline{\overline{x_1x_2}\,\overline{x_1x_3}\,\overline{x_2x_3}}$$

Kontaktní schéma:

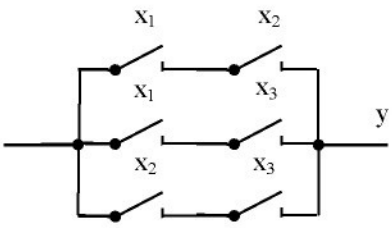
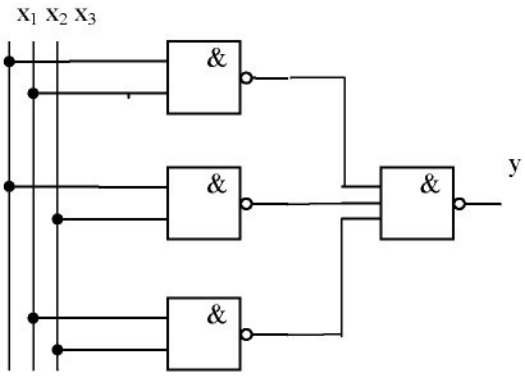
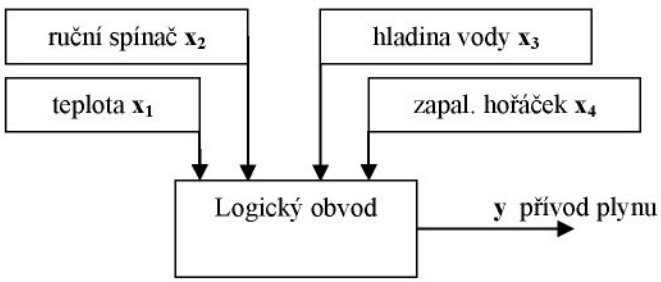


Schéma logických členů NAND:



Příklad 1.7. ***Příklad na syntézu kombinačního logického obvodu***

Automatika kotle pro vytápění rodinného domku má otevírat přívod plynu do kotle, když venkovní teplota klesne pod 15° anebo je sepnut ruční spínač a samozřejmě když je voda v kotli nad minimální hodnotou a hoří zapalovací hořáček. Navrhněte logický obvod, sestavený z kontaktních prvků a logických členů NANDD



Řešení:

$x_1=1$ teplota <15° $x_2=1$ zapnuto $x_3=1$ voda nad min $x_4=1$ hoří $y=1$ plyn otevřen
$x_1=0$ teplota ≥15° $x_2=0$ vypnuto $x_3=0$ voda pod min $x_4=0$ nehoří $y=0$ plyn zavřen

Pravdivostní tabulka:

x1	x2	x3	x4	y
0	0	0	0	0
0	0	0	1	0

0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	1
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	1
1	1	0	0	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	1

Karnaughova mapa:
Je zbytečná, protože výsledná funkce je jednoduchá.

$$y = \overline{x_1}x_2x_3x_4 + x_1\overline{x_2}x_3x_4 + x_1x_2x_3x_4 = x_1x_3x_4 + x_2x_3x_4$$

$$y = \overline{\overline{x_1x_3x_4}\,\overline{x_2x_3x_4}}$$

Kontaktní schéma:

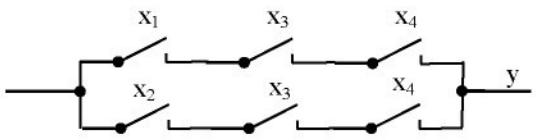
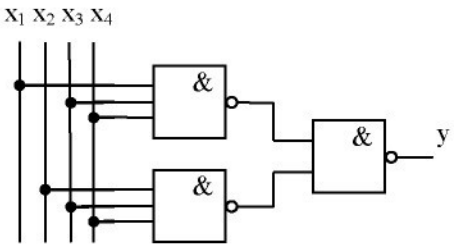


Schéma logických členů NAND:



Shrnutí pojmů 1.2.

V této kapitole se studenti seznámili s možnostmi realizace logických funkcí jak pomocí kontaktních schémat, tak pomocí hradel NAND nebo NOR se dvěma vstupy. Tyto postupy jsou ukázány na čtyřech vzorových řešených příkladech.

Otázky 1.2.

1. Jak se zapojuje logický součin pomocí kontaktních prvků?
2. Dá se kombinační logický obvod zapojit pomocí dvoustopého hradla NAND?
3. Jak se zapojuje logický součet pomocí hradel NOR?

Úlohy k řešení 1.2.

1. Ze zadané pravdivostní tabulky vytvořte logickou funkci. Tu minimalizujte a nakreslete zapojení pomocí kontaktních prvků a logických členů NAND.

X_1	x_2	x_3	y
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

2. Ze zadané pravdivostní tabulky vytvořte logickou funkci. Tu minimalizujte a nakreslete zapojení pomocí kontaktních prvků a logických členů NOR.

X_1	x_2	x_3	x_4	y
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	1
0	1	0	0	1
0	1	0	1	1
0	1	1	0	0
0	1	1	1	1
1	0	0	0	1
1	0	0	1	0
1	0	1	0	0
1	0	1	1	0
1	1	0	0	1
1	1	0	1	0
1	1	1	0	1
1	1	1	1	1

3. Ze zadané logické funkce vytvořte pravdivostní tabulku a minimalizujte ji. Nakreslete ji pomocí kontaktních prvků.

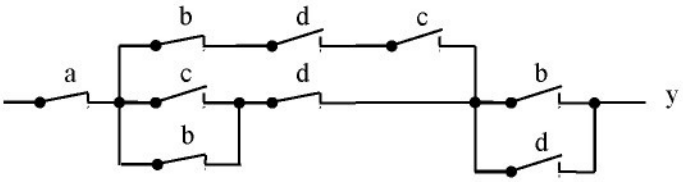
$$y = a\bar{b}c + \bar{a}\bar{b}cd + \bar{a}c + a\bar{b}cd + ab\bar{c}d$$

4. Ze zadané pravdivostní tabulky nakreslete zapojení pomocí logických členů NAND.

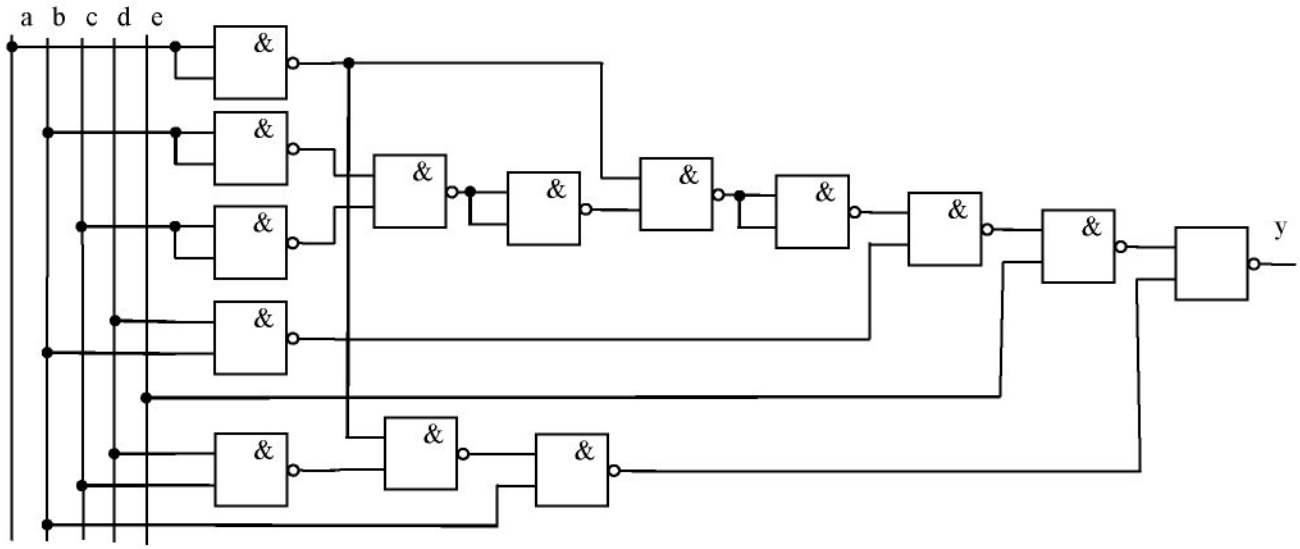
X_1	x_2	x_3	y
0	0	0	1
0	0	1	1
0	1	0	0

X_1	x_2	x_3	y
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

5. Přiřaďte k danému zapojení pomocí kontaktních prvků výslednou logickou funkci. (negace a = NOTa)



6. Přiřaďte k danému zapojení pomocí dvouvstupých hradel NAND výslednou logickou funkci. (negace a = NOTa)



1.3 Časové operace v logickém řízení

Čas ke studiu: 1 hodina

Cíl Po prostudování tohoto odstavce budete umět

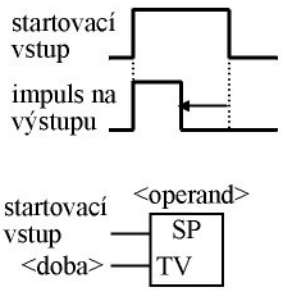
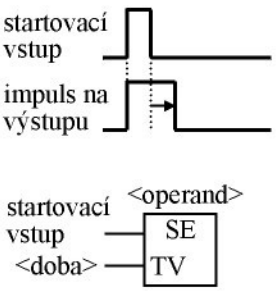
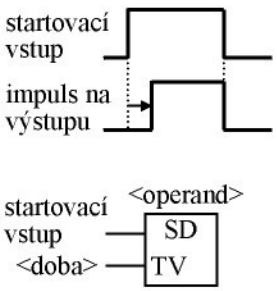
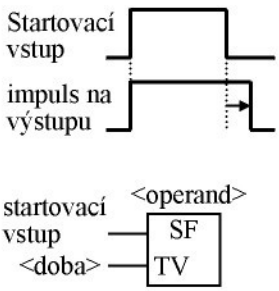
- definovat pojem sekvenční logické funkce
- popsat syntézu sekvenčních logických funkcích.

Výklad

V předcházejících kapitolách byly popsány způsoby jak řešit úlohy kombinační logiky a jak realizovat logické obvody. Do logického řízení však také patří úlohy, kdy je sledováno, co se stane za určitý časový interval. Například je spuštěn kompresor, jehož úkolem je natlakovat určitou nádobu nad tlak, který je indikován například manometrem s elektrickým kontaktem. V případě poruchy kompresoru nebo netěsnosti systému, požadovaný tlak nemůže být nikdy dosažen. Bylo by žádoucí tuto situaci signalizovat například kontrolkou nebo zvukovým výstražným signálem. V souboru dosud popsaných prostředků žádný prvek, který by umožňoval tento problém řešit, popsán nebyl.

V logických systémech s relé bylo k dispozici časové relé, které se dělí na relé se zpožděným přítahem a zpožděným odpadem. U programovacích logických automatů jsou prvky pro řešení časových operací znázorněny v tabulce 1.7. Značka časové operace v blokových schématech pro programovatelné automaty Siemens obsahuje vstupy, kterými je vstupní (startovací) logický signál a délka časového intervalu (TV – Time Value) pro jednu ze čtyř časových operací. Logický signál je datový typ BOOL (Boolean). Délka časového intervalu není logický signál, ale jiný datový typ označovaný TIMER. Nad značkou je operand, který je výstupem tohoto bloku, a je datový typ BOOL. Popsané značky se používají v obecnějším typu blokových schémat než bylo dříve popsáno. V tabulce jsou uvedeny čtyři operace SP, SE, SD nebo SF. Operace SP zkrátí případně delší startovací impuls na dobu TV a operace SE naopak udrží délku impulsu na hodnotě TV nezávisle na délce startovacího impulsu. Operace SD zpozdí náběžnou hranu startovacího impulsu bez vlivu na sestupnou hranu. Operace SF zpozdí sestupnou hranu startovacího impulsu. Při tvorbě programu pro programovací automat je třeba nejdříve zadat délku intervalu TV instrukcí L <doba>. Další instrukce určí typ časové operace a operand, tj. signál ve tvaru impulsu, který je výstupem této operace.

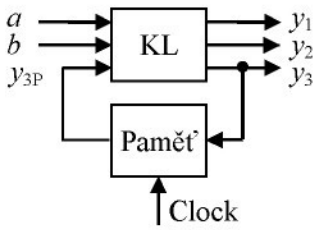
Tabulka 1.7 – Časové funkce

Funkce: SP (start timer as pulse)	Funkce: SE (start timer as extended pulse)	Funkce: SD (start timer as ON delay)	Funkce: SF (start timer as OFF delay)
			
Zápis programu: L <doba> SP <operand>	Zápis programu: L <doba> SE <operand>	Zápis programu: L <doba> SD <operand>	Zápis programu: L <doba> SF <operand>

Sekvenční logické obvody

Kombinační logická funkce popisuje relaci mezi závisle proměnnou (výstup KL) a skupinou nezávisle proměnných (vstupy KL), přičemž se předpokládá, že výstupy neovlivňují vstupy. Naproti tomu výstup sekvenčního logického obvodu y závisí nejen na zmíněných nezávislých vstupech, ale také na minulých výstupech y_P tohoto obvodu. Platí

$$y = f(y_P, a, b, c, ...) \tag{1.15}$$



Obrázek 1.14 – Sekvenční logický obvod

V kombinačním logickém obvodu na obrázku 1.14 je přidána ke kombinačnímu logickému obvodu zpětná vazba, která je zprostředkována paměťovým blokem. Přepis hodnot vstupních proměnných paměťového bloku na jeho výstup je řízen tzv. hodinovými impulsy (CLOCK). Důvodem zařazení paměťového bloku je zpoždění dvouhodnotových signálů při průchodem bloku kombinační logiky. U mechanických relé je zapotřebí určité doby na ustálení kontaktů stejně tak elektronické obvody mají svoje zpoždění třeba v nanosekundách.

Interval mezi hodinovými impulsy, které řídí paměť, je nastaven tak, že dojde k ustálení všech přechodových dějů.

Paměťový blok není nezbytně zapotřebí pokud oddělení minulosti a přítomnosti zajistí setrvačnost řízeného procesu. Jestliže je tento blok potřebný, pak je realizován klopnými obvody, které budou popsány později.

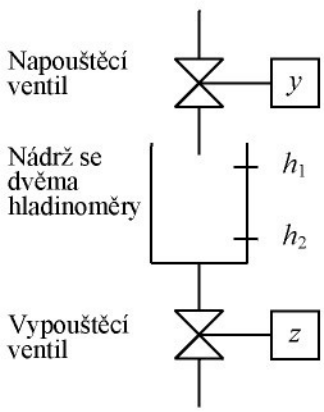
Syntéza sekvenčního logického řízení bude demonstrována na sekvenčním řízení zaplnění nádrže kapalinou, jehož funkční schéma je znázorněno na obrázku 1.15. Odtok kapaliny ovládá vypouštěcí ventil. Jeho otevření ($z = 1$) nebo uzavření ($z = 0$) představuje poruchu, jehož vlivu na zaplnění nádrže má čelit sekvenční řídicí obvod přítoku kapaliny do nádrže. Tento ventil je buď otevřen ($y = 1$) nebo uzavřen ($y = 0$), nabývá tedy dvou hodnot. Zaplnění nádrže indikují dva hladinoměry (stavoznaky) s dvouhodnotovým výstupem. Logická jednička znamená překročení hladiny ($h_1 = 1$, $h_2 = 1$) a logická nula signalizuje, že hladina je nižší. Hladinoměr h_1 je nad hladinoměrem h_2.

Příklad 1.8. ***Příklad syntézy sekvenčního logického obvodu***

Slovní zadání funkce sekvenčního logického obvodu je následující:

- Jestliže je napouštěcí ventil otevřen ($y_P = 1$) a je překročena hladina na horním hladinoměru ($h_1 = 1$), pak je třeba napouštěcí ventil uzavřít ($y = 0$).
- Jestliže je napouštěcí ventil uzavřen ($y_P = 0$) a hladina kapaliny je nižší než úroveň dolního hladinoměru ($h_2 = 0$), pak je třeba napouštěcí ventil otevřít ($y = 1$).
- V ostatních případech nechť se nastavení napouštěcího ventilu nemění.

Řešení:



Obrázek 1.15 – Schéma řízené soustavy

Ve slovním zadání se rozeznává stávající nastavení napouštěcího ventilu (index P u veličiny y) a nové nastavení (y), které má být výstupem sekvenčního obvodu. Vstupními proměnnými je tedy stávající stav napouštěcího ventilu (y_P) a dvouhodnotové výstupy obou hladinoměrů (h_1, h_2)Dalším krokem je sestavení tabulky přiřazující kombinace hodnot vstupních proměnných k výstupní závisle proměnné, kterým je žádoucí nastavení napouštěcího ventilu y. Za účelem optimalizace funkčního předpisu je výhodné vyplnit Karnaughovou mapu.

y \ h_2 h_1				
	1	0	0	X
y_P	1	1	0	X

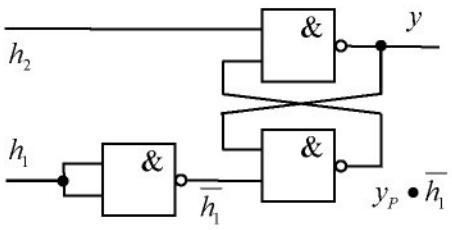
Jednotlivá pole mapy lze podle pravidel řízení vyplnit až na kombinaci údajů hladinoměrů předpokládající, že horní hladinoměr signalizuje překročení hladiny a současně dolní hladinoměr signalizuje nedosažení spodní hladiny, což je zjevný nesmysl. Tato pole je vhodné doplnit logickými jedničkami, aby zápis funkce byl co nejjednodušší.

$$y = \overline{h_2} + y_P \bullet \overline{h_1}$$

Pro realizaci této funkce hradly NAND je třeba vzorec upravit do tvaru

$$y = \overline{\overline{\overline{h_2} + y_P \bullet \overline{h_1}}} = \overline{h_2 \bullet \overline{y_P \bullet \overline{h_1}}}$$

Blokové schéma příslušného sekvenčního obvodu je znázorněno na obrázku 1.16. Ve schématu je propojen výstup obvodu y s jeho minulou hodnotou y_P.



Obrázek 1.16 – Zapojení sekvenčního obvodu

Toto propojení může být zdrojem nestability. Protože hladinoměry obvykle vykazují hysterezi, lze předpokládat, že sekvenční obvod bude stabilní. Jestliže změna výstupního signálu hloubkoměrů, tj. přechod $0 \to 1$ a $1 \to 0$ nastává při různých hladinách, pak nemůže dojít k rozkmitání obvodu. Jiným řešením stabilizace obvodu je zpoždění otevření a zavření ventilu pomocí dříve definovaných časových funkcí.

Shrnutí pojmů 1.3

Kapitola je věnována časovým operacím v logických funkcích a definici sekvenčních logických funkcí, včetně řešeného příkladu.

Otázky 1.3.

1. Co je to sekvenční logická funkce?
2. Co obsahují sekvenční logické obvody?

1.4 Klopné obvody

Čas ke studiu: 1 hodina

Cíl Po prostudování tohoto odstavce budete umět

- definovat klopné obvody a jejich základní typy.

Výklad

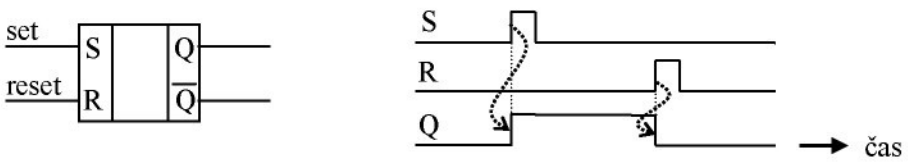
Klopné obvody (KO) jsou prvky s vnitřní pamětí dvou stavů. Těchto prvků je několik typů. Ve skriptech budou popsány klopné obvody typu RS, D a JK. Tyto klopné obvody mají dva výstupy, jeden výstup je tzv. přímý Q a druhý výstup je jeho negací $\overline{Q}$. Uvedené typy klopných obvodů se liší počtem a účinkem svých vstupů. Klopný obvod střídá v důsledku změn vstupů logickou hodnotu přímého a negovaného výstupu. Jak je zřejmé přímý výstup může nabýt hodnoty logické nuly nebo jedničky a negovaný výstup opačně logické jedničky nebo nuly. Tento jev se označuje jako klopení.

Klopný obvod lze zkonstruovat například z elektronických součástek. Zařízení vykazující dva stabilní stavy lze vytvořit také na mechanickém, pneumatickém nebo hydraulickém principu. Mechanickým klopným obvodem je obyčejný kolíbkový světelný vypínač, se kterým se setkáváme každý den. Každý ví, že má dvě stabilní pozice (vypnuto a zapnuto) a že ke změně pozice je třeba zásahu zvenčí, tj. stlačení jedné nebo druhé strany pohyblivé části spínače.

Klopný obvod typu RS

Klopný obvod tohoto typu má dva vstupy: nastavovací S (set) a resetovací R (reset). Impuls přivedený na vstup S klopného obvodu způsobí jeho překlopení ze stavu $Q = 0$ do stavu $Q = 1$. Jestli je klopný obvod již ve stavu $Q = 1$, pak se nestane nic. Účinek impulsu na vstupu R je opačný. V případě, že klopný obvod je ve stavu $Q = 1$, pak se překlopí do stavu $Q = 0$. Jestliže v tomto stavu klopný obvod již byl, pak se také neděje nic.

Vlastnosti klopných obvodů se popisují časovými diagramy, které znázorňují časový průběh vstupních signálů a změny stavu klopného obvodu. Obvodová značka a časový diagram funkce klopného obvodu RS je na obrázku 1.17. Změna stavu klopného obvodu nastává v reakci na náběžnou hranu impulsů. V diagramu na zmíněném obrázku je zdůrazněn časový okamžik klopení šipkou směřující od nástupné hrany impulsů S a R k nástupné a sestupné hraně přímého výstupu Q. Negovaný výstup není v diagramu nakreslen, avšak jeho časový průběh si lze domyslet.



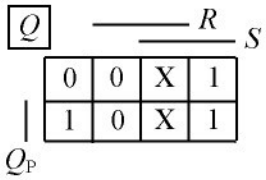
Obrázek 1.17 – Obvodová značka a časový diagram funkce klopného obvodu RS

Funkci klopné obvodu lze popsat dříve definovanými nástroji, tj. tabulkou nebo Karnaughovou mapou, která může být podkladem k odvození vzorce.

Tabulka funkčních hodnot KO typu RS

S	R	Q
0	0	předchozí stav
0	1	0
1	0	1
X	X	dáno zapojením

Karnaughova mapa popisující KO typu RS



Nedefinovaný výstup je pro situaci, kdy oba vstupy, tj. R a S, nabývají hodnotu logické jedničky. To odpovídá případu, kdy je současně požadováno nastavení a resetování klopného obvodu, což se samo sebou jeví jako nesmyslný požadavek. K odvození vzorce je přesto potřeba Karnaughovu mapu doplnit. Možnosti doplnění jsou čtyři, z nichž jen dva způsoby jsou z hlediska jednoduchosti výsledného vzorce výhodné. Pro obě nedefinované funkční hodnoty (znak X v mapě) lze zvolit logické jedničky nebo nuly. Jestliže se použijí pro obě pole logické jedničky, pak je vzorec popisující chování klopného obvodu následující.

$$Q = S + \overline{R} \bullet Q_P \tag{1.16}$$

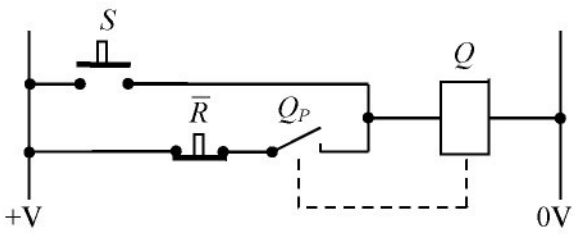
Realizovat logickou funkci podle vzorce (1.21) lze například pomocí kontaktů relé, viz obrázek 1.18. Kontakt s označením Q_P je ovládán cívkou relé Q. Stiskem tlačítka S se do relé přivede proud, který sepne spínací kontakt Q_P. Stiskem resetovacího tlačítka $\overline{R}$ se relé uvede do klidového stavu. Popsaný obvod představuje paměť nejmenšího možného množství informace, kterým je jeden bit. Tato jednotka znamená, že si lze vybrat ze dvou možností (sepnuto/rozepnuto).

Druhou možností doplnění Karnaughovy mapy je doplnění logickými nulami. Výsledný vzorec má po úpravách tvar

$$Q = Q_P \bullet \overline{R} + S \bullet \overline{R} = \overline{R} \bullet (Q_P + S) \tag{1.22}$$

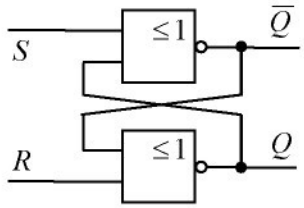
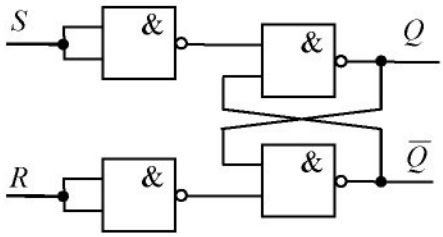
Také podle tohoto vzorce lze sestavit klopný obvod z kontaktů relé.

Klopný obvod lze sestavit také z integrovaných obvodů NOR a NAND.



Obrázek 1.18 – Kontaktní schéma KO typu RS

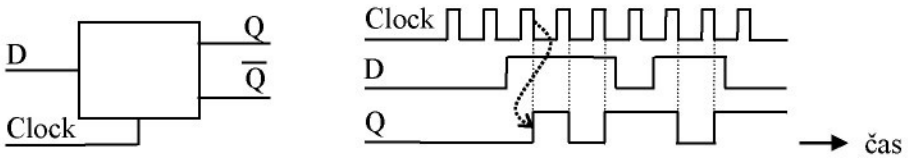
Vzorec (1.21) a (1.22) obsahuje jeden logický součet a jeden součin. S využitím de Morganových pravidel (1.13) lze dvojitou negací součtu nebo součinu, například u vzorce (1.22), obdržet zápis, podle kterého lze snadno sestavit zapojení zmíněných hradel na obrázku 1.19. Výraz $\overline{(Q_P + S)}$ u zapojení s hradly NOR označuje výstup $\overline{Q}$ a výraz $\overline{\overline{Q_P} \bullet \overline{S}}$ u zapojení s NAND označuje Q. Pro $S = 0$ platí identity $Q = Q_P$ a $\overline{Q} = \overline{Q_P}$ a pro $S = 1$ je $\overline{Q} = 0$ a $Q = 1$ pro zapojení s NOR, respektive NAND.

Hradla NOR:	Hradla NAND:
$Q = \overline{\overline{R} \bullet \overline{(Q_P + S)}} = \overline{R + \overline{(Q_P + S)}}$	$\overline{Q} = \overline{\overline{R} \bullet \overline{(Q_P + S)}} = \overline{\overline{R} \bullet \overline{Q_P} \bullet \overline{S}}$
	

Obrázek 1.19 –Klopné obvody typu RS z hradel NOR a NAND

Klopný obvod typu D

Klopný obvod tohoto typu má dva vstupy. Na jeden vstup se zapojí zdroj (tzv. hodinových) impulsů (Clock) a na druhý vstup D libovolný dvouhodnotový signál. Výstup tohoto KO se klopí jen když je na vstup D přivedena logická jednotka. Okamžik klopení je na časovém diagramu na obrázku 1.20 dán sestupnou hranou impulsů Clock. Označení Clock má svůj původ ve generátoru vysokofrekvenčních impulsů o stálé frekvenci, který je součástí digitálních hodin a každého počítače.



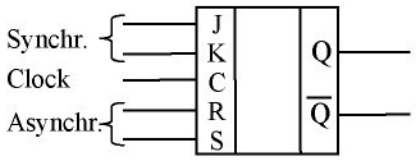
Obrázek 1.20 – Obvodová značka a časový diagram klopného obvodu typu D

Klopný obvod typu JK

Klopný obvod JK má pět vstupů. Jsou to vstupy R a S se stejnou funkcí jako u klopného obvodu typu RS. Kdykoliv po příchodu impulsu na vstup S je výstup Q překlopen na logickou jedničku a také kdykoliv po příchodu impulsu na vstup R je tento výstup Q překlopen na logickou nulu. Další dvojice vstupů je označena J a K. Účinek vstupního signálu na těchto dvou vstupech je vázán na vstup Clock s hodinovými impulsy. Signál logické jedničky na vstupu J klopí výstup Q do logické jedničky jen v okamžiku sestupné hrany hodinového impulsu. Podobně signál logické jedničky na vstupu K klopí výstup Q do logické nuly.

Tabulka:

J	K	Q_{n+1}
0	0	Q_n
0	1	0
1	0	1
1	1	$\overline{Q_n}$



Obrázek 1.21 – Obvodová značka klopného obvodu typu JK

Chování klopného obvodu lze popsat tabulkou. Narozdíl klopnému obvodu RS je definován také stav, kdy oba vstupy J a K jsou na logické jedničce. V tomto případě se klopný obvod chová stejně jako klopný obvod typu D. Jestliže se vstupy J a K nezapojí a zůstanou volné, pak se klopný obvod chová jakoby na těchto vstupech byla logická jednička.

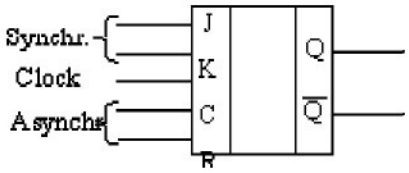
Obvodová značka klopného obvodu JK je na obrázku 1.21. Jeho vstupy jsou podle výše popsané funkce rozděleny na synchronní, asynchronní a hodiny (Clock).

Shrnutí pojmů 1.4.

V kapitole jsou popsány klopné obvody, jejich základní typy a realizace a časový diagram.

Otázky 1.4.

1. Kolik mají stavů klopné obvody?
2. Kdy se klopí výstup KO typu D?
3. Jakého obvodu je značka na obrázku?



1.5 Zapojení klopných obvodů v sekvenčních obvodech

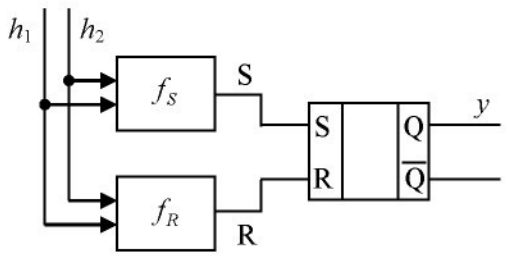
Čas ke studiu: 1 hodina

Cíl Po prostudování tohoto odstavce budete umět

- popsat použití klopných obvodů pro realizaci sekvenčních logických funkcí.

Použití klopného obvodu typu RS

Použití klopného obvodu bude demonstrováno na příkladu řízení zaplnění nádrže na obrázku 1.15. K řízení je použit napouštěcí ventil a dva hladinoměry h_1, h_2. Dvouhodnotový výstup y ovládá otevření a zavření napouštěcího ventilu. Nechť jako paměť nastavení tohoto ventilu je použit klopný obvod typu RS. Syntéza tohoto sekvenčního obvodu spočívá v návrhu kombinačních obvodů, jejichž výstup bude ovládat nastavovací a resetovací vstup klopného obvodu na obrázku 1.22. Hledá se tedy funkce



Obrázek 1.22 – Použití klopného obvodu RS v sekvenčním obvodu

$$S = f_S(h_1, h_2)$$
$$R = f_R(h_1, h_2) \qquad (1.17)$$

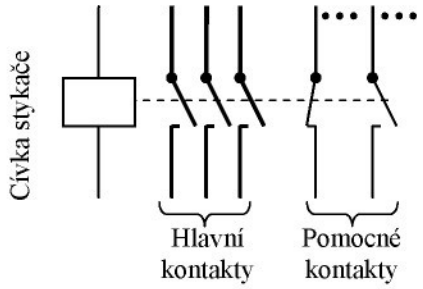
Návrh obou kombinačních obvodů bude proveden s využitím Karnaughovy mapy. Prvním krokem syntézy je vyplnění polí mapy podle pravidel pro ovládání napouštěcího ventilu ze zmíněného obrázku. Pro generování signálu nastavovacího (S) je třeba, aby dolní hladinoměr h_2 signalizoval pokles hladiny pod jeho úroveň, přičemž na výstupu horního hladinoměru h_1 nezáleží. Pro vznik signálu resetovacího je požadováno, aby byla překročena hladina horního hladinoměru h_1. Formální postup řešení této elementární úlohy s použitím Karnaughovy mapy je znázorněn na obrázku 1.23.

S \ h_2 / h_1: 1 0 / X 0	Funkce pro nastavení KO $S = \overline{h_2}$	R \ h_2 / h_1: 0 0 / X 1	Funkce pro resetování KO $R = h_1$

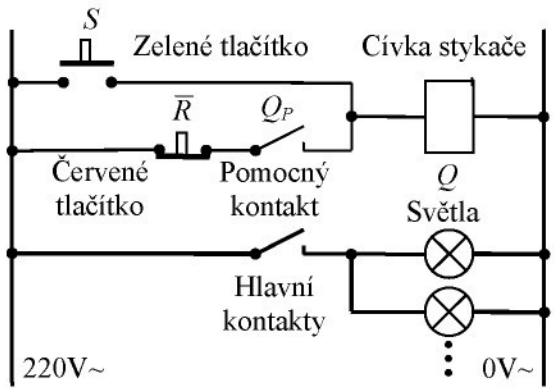
Obrázek 1.23 – Návrh KL obvodů pro vstup S a R klopného obvodu pro řízení zaplnění nádrže

Další příklad použití klopného obvodu RS bude ke spuštění a zastavení libovolného elektrického zařízení, například osvětlení. Často je zapnutí a vypnutí přívodu elektřiny vykonáno změnou polohy vypínače, například změnou polohy páčky nebo otočením

ovládacího prvku. Pohodlnější je však stisk zeleného tlačítka pro rozběh stroje nebo stisk červeného tlačítka pro jeho zastavení. Tuto funkci řeší jednoduchý sekvenční obvod se stykačem. Stykač je relé s kontakty, které jsou dimenzovány na rozdílné proudové zatížení, viz obrázek 1.24. Hlavní kontakty stykače jsou dimenzovány na velký proud, zatímco jeho pomocné kontakty jsou pro malý proud. Stykače se používají k připojování spotřebičů o výkonu 1 kW až 500 kW. Kontaktové schéma ovládacího obvodu je na obrázku 1.25.



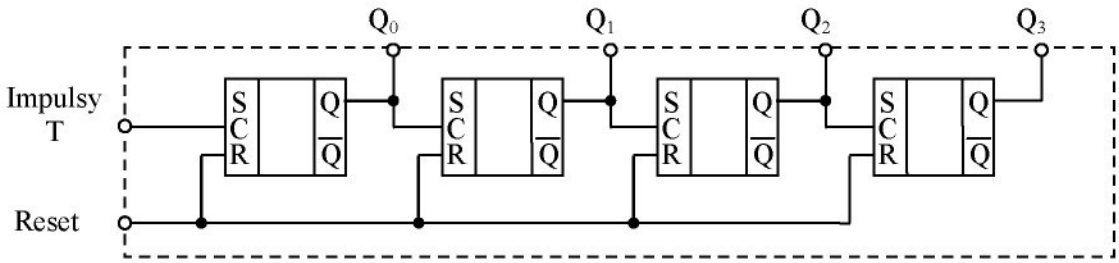
Obrázek 1.24 – Stykač



Obrázek 1.25 – Stykač

❑ Použití klopného obvodu typu JK pro čítání impulsů

Každý ví, co je to počítadlo ujetých kilometrů v automobilu. Dříve bylo toto počítadlo mechanické, s dělením po dekádách. Jedna dekáda odpovídala 10 stavům s čísly od nuly do devíti. Mechanicky bylo zajištěno, že jedno otočení kolečka s nižší dekádou vyvolalo otočení kolečka s vyšší dekádou o jednu desetinu. Konstrukce elektronických počítadel, nazývaných čítače, se opírá o elektronické prvky, které mají jen dva stavy. Začalo to u prvních pokusů sestavit počítač s elektromechanickými relé v roce 1934 v Německu (Konrád Zuse, počítače Z1, Z2 a Z3 – první použitelný počítač s 2600 relé). Jednoduchost realizace logických operací tento princip počítání se dvěma stavy udržela až do současné doby. Výpočty a funkce jsou založeny na dvojkové soustavě čísel, které obsahující jen dvě cifry (nulu a jedničku) pro jedno místo čísla. Tyto cifry dvojkové soustavy se mohou pamatovat v jednom klopném obvodu. Jedna cifra pak obsahuje informaci o velikosti jednoho bitu. Dva klopné obvody uchovávají informaci o velikosti 2 bitů, přičemž počet různých stavů je roven 2^2 (tj. 00, 01, 10, 11). Obecně počet *m* klopných obvodů představuje paměť *m* bitů s počtem stavů 2^m.



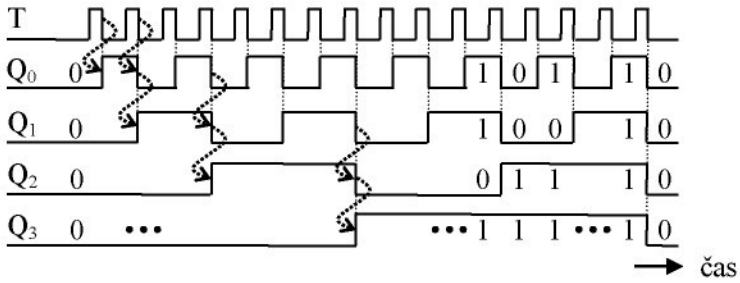
Obrázek 1.26 – Zapojení 4bitového dvojkového čítače o kapacitě 16 stavů

V následujícím příkladu bude demonstrováno použití klopného obvodu JK k návrhu 4bitového dvojkového čítače. Ve schématu zapojení bude předpokládáno, že vstupy JK nejsou zapojeny (stejný efekt má přivedení logické jedničky), proto bude KO pracovat jako typ D. Čítač bude obsahovat tedy čtyři klopné obvody. Celkový počet kombinací čtveřic nul

a jedniček je $2^4 = 16$. Čítač má kapacitu 16 stavů a je schopen spočítat 0 až 15 impulsů. Jednotlivé bity se budou pamatovat klopnými obvody s výstupy Q_0, Q_1, Q_2 a Q_3. Počítány budou impulsy T na vstupu Clock prvního klopného obvodu. Čítač lze vynulovat vstupem Reset. Zapojení čítače je na obrázku 1.26.

Jak je zřejmé z časového diagramu na obrázku 1.27, sekvenční obvod prochází postupně 16 stavy, které jsou dány výstupy obvodů Q_3, Q_2, Q_1, Q_0. Nechť logická nula odpovídá dvojkové nule 0_2 (a také dekadické nule 0_{10}) a logická jednička odpovídá dvojkové jedničce 1_2 (a také dekadické jedničce 1_{10}). Bity v pořadí od Q_3 do Q_0 představují číslo ve dvojkové soustavě. Zmíněných 16 stavů obvodu odpovídá hodnotám od 0000_2 až po 1111_2. Přepočet výstupů klopných obvodů na dekadickou hodnotu je podle následujícího vzorce, do kterého se za veličiny Q_0 až Q_3 dosazují desítkové nuly nebo jedničky, protože jejich numerická hodnota je shodná s dvojkovými nulami a jedničkami

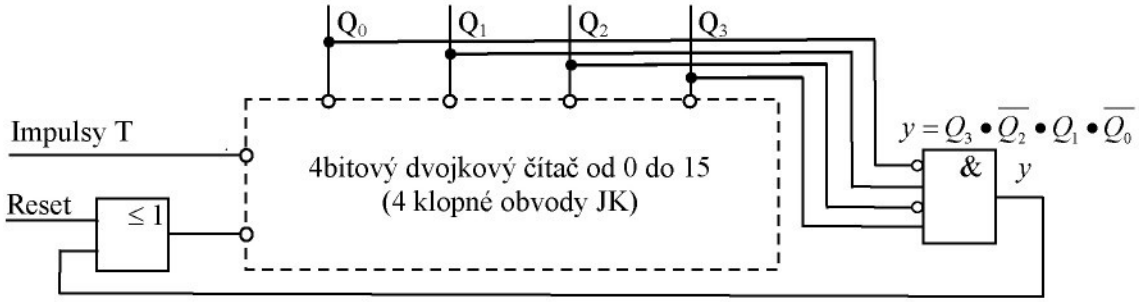
$$K = Q_3 2^3 + Q_2 2^2 + Q_1 2^1 + Q_0 2^0 \qquad (1.18)$$



Obrázek 1.27 – Časový diagram posloupnosti stavů dvojkového čítače

V šestnáctkové soustavě jsou tyto stavy označeny ciframi 0, 1, 2, ..., 9, A, B, ..., F. Dvě šestnáctkové cifry představují desítkové číslo od 0 do $2^8 - 1 = 255$ desítkově, tj. od 00 do FF šestnáctkově. Tyto dvě cifry obsahují kombinaci osmi nul nebo jedniček a tvoří jeden byte (čti bajt).

Čítač omezený na 10 stavů, tj. pro desítkové cifry 0 až 9, musí být při dosažení stavu $1010_2 = 10_{10}$ resetován. Zapojení čítače z obrázku 1.23 je třeba upravit. Výsledek úprav je znázorněn na obrázku 1.28. Toto zapojení může představovat jednu dekádu čítače, který pracuje po desítkách. Výsledek funkce tohoto obvodu představuje čtveřici bitů o počtu rovném počtu dekád. Takto organizovaným datům se říká BCD kód, tj. binárně kódované desetinné místo (anglicky Binary Coded Decimal).



Obrázek 1.28 – Zapojení dvojkového čítače o kapacitě 10 stavů

Shrnutí pojmů 1.5.

Kapitola popisuje možností použití jednotlivých typů klopných obvodů k realizaci sekvenčních logických funkcí. Jsou zde popsány konkrétní aplikace (stykač, čítač impulsů, dvojkový čítač,.

Otázky 1.5.

1. Z čeho se skládá čítač?

1.6 Sekvenční funkční diagramy

Čas ke studiu: 1 hodina

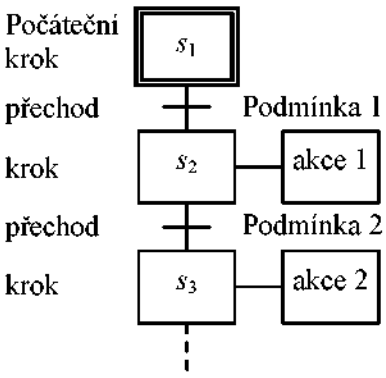
Cíl Po prostudování tohoto odstavce budete umět

- definovat sekvenční funkční diagram a jeho použití.
- popsat tvorbu funkčního schématu.
- vyřešit převod funkčního schématu na sekvenční funkční diagram.

Výklad

Mnohé technologické procesy nejsou kontinuální, ale představují periodicky se opakující posloupnost akcí, jako je otevření a uzavření ventilů, rozběh nebo doběh pohonů, čekání na dosažení hladiny nebo teploty, a pod.. K popisu takového algoritmu řízení je třeba srozumitelný jazyk a pravidla, jako například při kreslení strojařských výkresů. Zmíněná funkce sekvenčního logického systému řízení se popisuje sekvenčním funkčním diagramem (anglicky Sequential Function Chart). Tyto diagramy jsou standardizovány normami, například ČSN EN 61131-3, DIN EN 61131-3 (obě jsou evropskými normami) nebo IEC 1131-1 (mezinárodní elektrotechnická norma). Sekvenční funkční diagramy doplňují dříve uvedené tři způsoby programování programovatelných automatů.

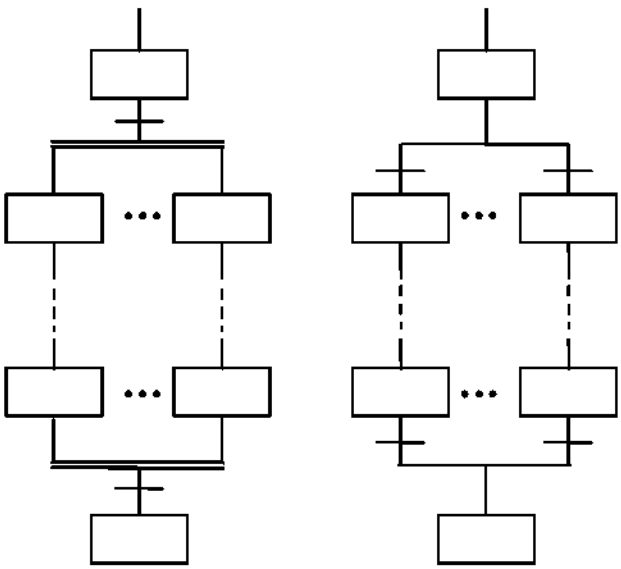
Příklad sekvenčního funkčního diagramu je na obrázku 1.29. Sekvenční diagram obsahuje dva typy prvků: kroky a přechody, které tyto kroky spojují. Krok je překlad anglického označení step, které má stejný význam jako české slovo stav. Pod označením krok nebo stav se rozumí určitá kombinace polohy kontaktů, logické úrovně na výstupu hradel nebo logické hodnoty proměnných programů. Každému kroku jsou přiřazeny akce (actions), které představují výstup logického systému řízení jako například otevření ventilů atd.. Kroky jsou aktivní nebo neaktivní. Aktivní krok odpovídá stavu, který v daném okamžiku v systému řízení nastal. Neaktivní kroky jsou stavy ostatní. Je zřejmé, že krok který je v určitý okamžik aktivní, se po splnění podmínky pro přechod (transition) mezi kroky stane krokem neaktivní a aktivním krokem se stane krok následující.



Obrázek 1.29 – Sekvenční funkční diagram

Jeden z kroků má označení počáteční krok. V sekvenčním diagramu se vyznačuje dvojitým orámováním bloku. Označení kroku se vpisuje dovnitř bloku. V našem diagramu je to logická proměnná (nabývá hodnoty logické nuly nebo jedničky). Bloky jsou spojeny přechody v podobě hrany grafu (prostá úsečka). U každého přechodu je vyznačena podmínka přechodu. Je to logická proměnná nebo výraz složený z několika logických proměnných. Jestliže se logická hodnota podmínky stane logickou jedničkou, přechod (přiřazení nového aktivního stavu) se uskuteční. Diagram se kreslí tak, že tyto přechody probíhají shora dolů, tudíž musí dosáhnout svého konce. Z konce diagramu se opět za splnění předepsaných podmínek přechodu vrací aktivní krok na krok počáteční.

Reálné výrobní procesy nejsou obvykle tak jednoduché, že tvoří jen jednoznačnou posloupnost kroků. Algoritmy řízení se větví do paralelních posloupností kroků a opět spojují. Pro tyto případy jsou definována určitá pravidla. Rozlišují se různé způsoby větvení a to na divergentní a konvergentní AND a divergentní a konvergentní OR. Tyto dva způsoby větvení jsou znázorněny na obrázku 1.30.



Divergentní a konvergentní AND

Divergentní a konvergentní OR

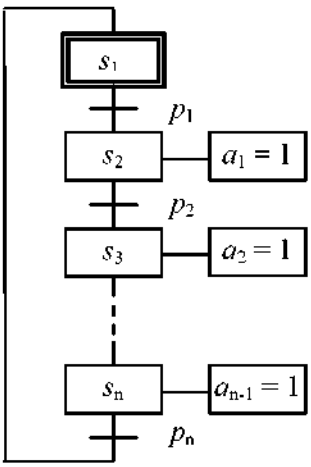
Obrázek 1.30 – Pravidla větvení sekvenčních diagramů

Při rozvětvení (divergenci) se způsob označený AND liší od způsobu označeného OR umístěním podmínky přechodu. Při divergentním AND po splnění podmínky přechodu přechází aktivní krok do všech větví diagramu současně a posloupnosti se vykonávají paralelně ovšem s přechody danými v každé větvi zvlášť. Před spojením paralelních posloupností kroků do prvního kroku po větvení AND je třeba splnit jen jednu podmínku.

Narozdíl od větvení AND je při větvení OR vykonávána posloupnost kroků jen v jedné větvi. Jestliže je při tomto způsobu (OR) rozvětvení splněno více podmínek přechodů do různých větví, pak se vybere větev nejvíce vlevo. Z rozvětvené části sekvenčního diagramu dojde k přechodu na společný krok po splnění podmínky přechodu na konci větve, jejíž posloupnost kroků se vykonává.

V předchozím textu kapitoly o logickém řízení byly popsány způsoby realizace logických obvodů pomocí kontaktů a hradel. Nyní bude popsáno počítačová realizace funkce sekvenčního logického řízení. Počítačový program je složen z posloupnosti instrukcí, které se v rychlém sledu vykonávají. Nejjednodušší je jednorázový start programu nebo dávky programů, který nebo které se po vykonání určité skupiny instrukcí ukončí. Na rozdíl od tohoto režimu práce počítače, například pro zpracování dat, jsou počítače pro řízení využívány jiným způsobem. Program běží nepřetržitě tak, že se posloupnost instrukcí opakuje s periodou velmi malého zlomku sekundy, například 5 ms, přičemž samotné instrukce se vykonávají o několik řádů rychleji. To znamená, že se opakovaně propočítávají všechny vzorce pro logické funkce, přičemž se respektují všechny změny na vstupech počítače.

Algoritmizace logického sekvenčního řízení do počítačového programu bude vysvětlena na příkladu, jehož sekvenční funkční diagram je znázorněn na obrázku 1.31. Logické proměnné v tomto diagramu jsou použity také v programu. V diagramu jsou tyto logické proměnné:



Obrázek 1.31 – Příklad pro programování

- Vnitřní proměnné označující jednotlivé kroky s_1, s_2, ..., s_n. Tyto proměnné nabývají hodnotu nula pro neaktivní krok a hodnotu jedna pro aktivní krok. V příkladu sekvence podle obrázku 1.27 může být jen jeden aktivní krok na rozdíl od sekvenčních diagramů, kde nastává divergence AND. Z důvodu algoritmického řešení programu, jak bude vysvětleno dále, mohou být po dobu jednoho opakování skupiny instrukcí dva aktivní kroky

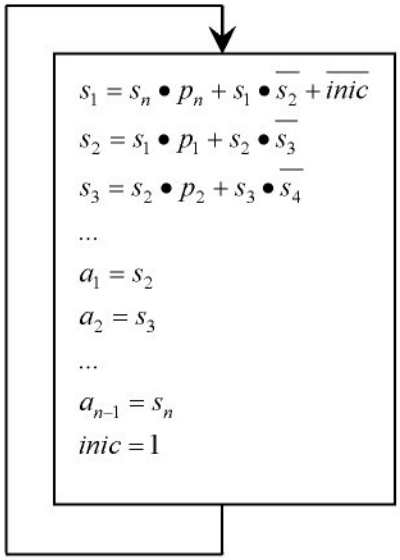
- Podmínky přechodů p_1, p_2, ..., p_n jsou logické proměnné spojené se vstupy počítače. Nabývají logických hodnot signalizujících překročení hladiny nebo teploty.

- Logické proměnné představující akce a_1, a_2, ..., a_{n-1}, které určují výstupy počítače. Jestliže tato proměnná nabude hodnoty jedna, pak se například otevře příslušný ventil.

- Pro úplnost je mezi proměnnými také proměnná *inic*, která je po startu počítače bezpečně vynulována. Zajistí správný start algoritmu tím, že umožní, aby proměnná s_1, která je spojena s počátečním krokem, byla nastavena na logickou hodnotu jedna.

Posloupnost logických funkcí k výpočtu pro aktuální kombinaci hodnot logických proměnných je uvedena ve vývojovém diagramu na obrázku 1.32. Bezprostředně po spuštění programu jsou všechny proměnné vynulovány (dosazeny logické nuly), tj. také proměnná *inic*. Žádný krok není také v tomto stavu programu aktivní. Dále je předpokládáno, že není splněna žádná podmínka přechodu.

S výjimkou prvního vzorce mají všechny ostatní vzorce pro výpočet vnitřních proměnných s_i dva členy logického součtu. První člen $s_i \bullet p_i$ tohoto logického součtu zajišťuje za splnění podmínky přechodu ($p_i = 1$) předání funkce aktivního kroku také následujícímu kroku, tj. nastavení logické proměnné s_{i+1} na logickou jedničku. V tomto průchodu jsou podle hodnot proměnných aktivní dva kroky (s_i a s_{i+1}). Následující výpočet však původní proměnnou vynuluje, protože ji předchází proměnná s_{i-1} s nulovou logickou hodnotou (ovlivňuje první člen součtu) a druhý člen logického součtu je také nulový, protože v součinu obsahuje negaci proměnné s_{i+1}, která je již na logické jedničce z minulého opakování programu. Vzhledem k velmi rychlému opakování výpočtů tato dvojznačnost není na závadu. Udržení hodnoty logické proměnné s_i na logické jedničce při opakování výpočtů je zajištěno druhým členem logického součtu $s_i \bullet \overline{s}_{i+1}$. Hodnota s_i je logická jednička a hodnota s_{i+1} je logická nula a ovšem její negace je logická jednička, proto je uvedený součin roven jedné.



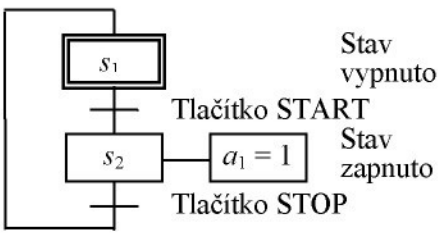
Obrázek 1.32 – Posloupnost vzorců pro přepočítávání proměnných

Přiřazení akcí, které jsou dány hodnotami výstupních proměnných a_i, na konec přepočtu vnitřních proměnných s_i je jen otázkou přehlednosti programu.

Počítač neustále přepočítává hodnoty logických funkcí ve velmi rychlém sledu a tak reaguje na změny podmínek pro přechody s odezvou v akcích, které představují výstupy počítače do technologického procesu. Postupné splnění podmínek přechodu zajišťuje také posouvání aktivního kroku a provádění příslušných akcí. Je pravděpodobné, že čtenář těchto skript nebude program upravovat, ale měl by jeho řešení vzít v úvahu při obsluze tohoto řídicího systému.

❑ Příklad na sestavení sekvenčního funkčního diagramu

V kapitole o realizaci klopných obvodů pomocí relé byl uveden příklad spuštění a zastavení stroje stiskem tlačítka start a stop. Tuto úlohu lze snadno řešit pomocí sekvenčního logického automatu na obrázku 1.33. Akce a_1 představuje logický stav jednoho z výstupů počítače.



Obrázek 1.33 – Klopný obvod RS

Způsob sestavení sekvenčního funkčního diagramu s větvením bude demonstrován na příkladu, který byl převzat ze skript [1]. Funkční schéma technologického procesu je na obrázku 1.34.

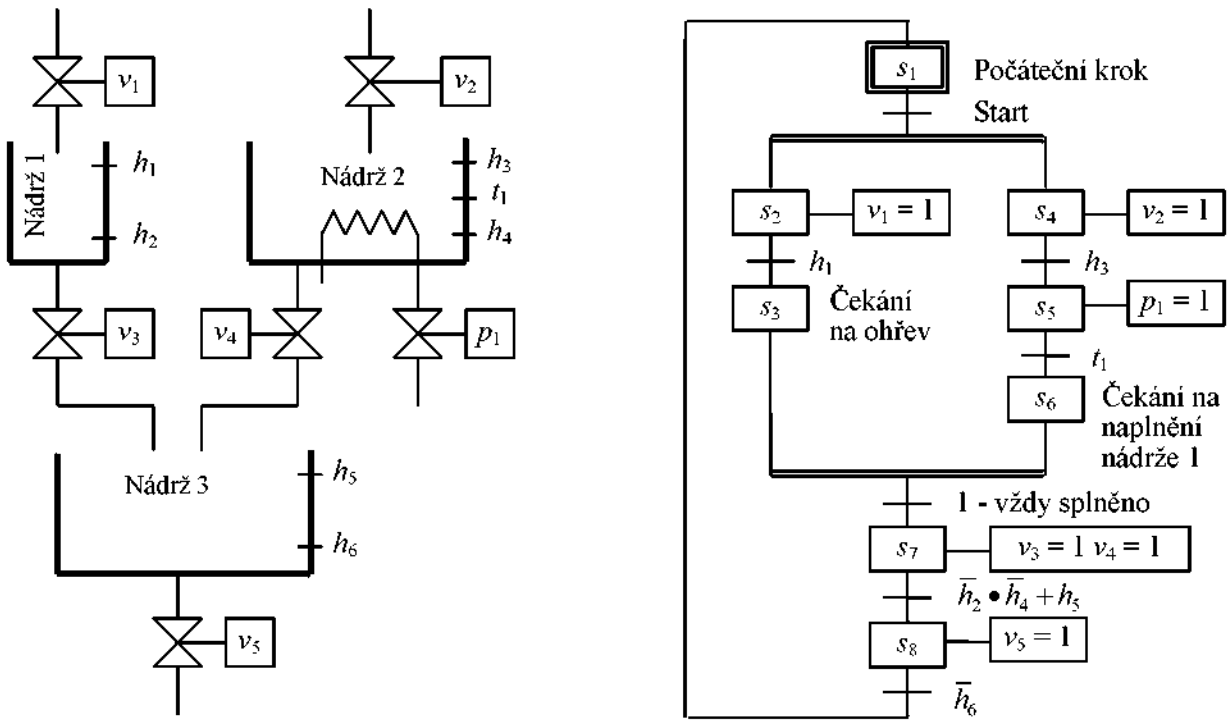
Jde o soustavu tří nádrží, které se plní a vyprazdňují, přičemž v jedné nádrži se navíc její obsah ohřívá na požadovanou teplotu. Nejprve se naplní nádrž 1 a 2. Konec plnění u obou nádrží je dán překročením nastavené hladiny, což se sleduje u každé nádrže zvlášť. Současně

s dokončením plnění nádrže 2 se spustí ohřev její náplně výměníkem, do kterého se vpustí ohřívací médium, jehož přítok je třeba po dokončení ohřevu zastavit.Po dokončení ohřevu a dokončení plnění nádrže 1 se obě nádrže vyprázdní přepuštěním jejich obsahu do nádrže 3. Až se nádrže 1 a 2 vyprázdní nebo se překročí zaplnění nádrže 3, začne se tato nádrž vyprazdňovat. Po vyprázdnění nádrže 3 se celý cyklus plnění, přepouštění a ohřevu opakuje znova.

Překročení zaplnění signalizuje logická jednička u logických proměnných h_1, h_3 a h_5. zatímco nedokončení plnění je signalizováno logickou nulou.

Neukončení vyprazdňování signalizuje logická jednička u logických proměnných h_2, h_4 a h_6. přičemž ukončení vyprazdňování je signalizováno logickou nulou u těchto proměnných. Nedokončení ohřevu je signalizováno logickou nulou u proměnné t_1 a dokončení logickou jedničkou.

Funkční schéma technologického procesu Sekvenční funkční diagram



Obrázek 1.34 – Funkční schéma a sekvenční funkční diagram příkladu technologického procesu

Všechny uvedené proměnné patří ke vstupům řídicího sekvenčního algoritmu. Jeho výstupy ovládají ventily tak, že při hodnotě proměnných v_1 až v_5 a p_1 rovné logické jedničce jsou ventily otevřeny a pro logickou nulu uzavřeny.

Sekvenční funkční diagram k příkladu v levé části obrázku 1.34 je v pravé části tohoto obrázku. Po startu cyklu výše popsaných operací nastavením podmínky přechodu Start na logickou jedničku se diagram rozvětví do dvou posloupností kroků a přechodů (divergentní AND). Levá větev řídí naplnění nádrže 1 a pravá větev navíc k řízení naplnění řídí také ohřev náplně nádrže 2. Po dokončení kroku plnění nádrže 1 přejde levá větev do kroku čekání na dokončení ohřevu. Pravá větev přejde po dokončení plnění nádrže 2 na krok ohřevu jejího

obsahu. Po dokončení kroku ohřevu přejde pravá větev do kroku čekání na naplnění nádrže 1. Krok čekání v obou větvích je s ohledem na vzájemné dokončení operací v sousedních větvích. Čeká však vždy jen jedna větev, zatímco u druhé větve není krok čekání delší než jeden cyklus opakování výpočtu programu. Je to proto, že obě větve se spojí v důsledku vždy platné podmínky před krokem zahájení vyprazdňování nádrže 1 a 2 do nádrže 3.

Shrnutí pojmů 1.6

V kapitole jsou vysvětleny sekvenční funkční diagramy, které se používají k popisu technologických procesů. Jsou popsány jejich prvky a způsoby větvení. Na závěr je praktická ukázka tvorby sekvenčního funkčního diagramu.

Otázky 1.6.

1. Jak se označuje počáteční krok v sekvenčním diagramu?
2. Jak se liší způsoby větvení AND a OR?

1.7 Programovatelné logické automaty

Čas ke studiu: 1,5 hodiny

Cíl Po prostudování tohoto odstavce budete umět

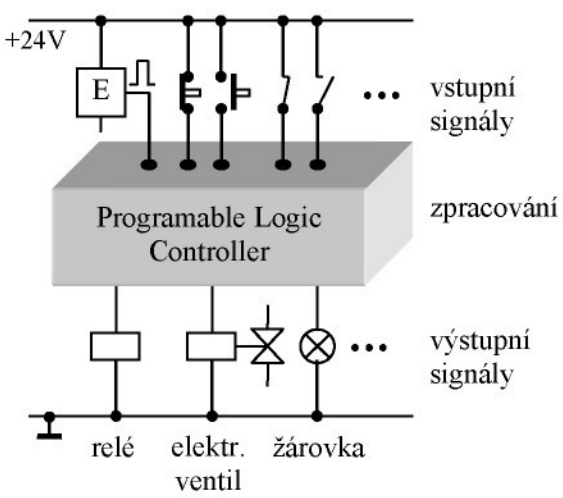
- definovat galvanické oddělení.
- popsat programovatelný logický automat a jeho základní prvky, způsoby programování, připojení na řízený proces.

Výklad

Dříve bylo řešení systémů sekvenčního logického řízení zajišťováno velíny s přidruženými skříněmi, které byly plné reléových obvodů, později s rošty se zásuvnými deskami s diskrétními obvody s tranzistory a ještě později s integrovanými obvody. Rozvoj počítačů přinesl inovace i do tohoto odvětví průmyslové elektroniky. Po prvních pokusech řešit problém sálovými počítači se dospělo do stavu kdy počítač třídy programovatelný logický automat - PLC (anglicky Programable Logic Controller) se stal univerzálním nástrojem k řešení řídicích systémů na úrovni blízké technologickému procesu. Kromě úloh logického řízení tyto PLC řeší také jednoduchou úlohu tzv. spojitého řízení, které bude popisováno v další důležité kapitole těchto skript.

Galvanické oddělení PLC od vnějších obvodů

Schématicky je připojení PLC na řízený systém nebo technologický proces znázorněno na obrázku 1.35. Vstupy tohoto řídicího systému představují kontakty, tlačítka a také elektrické dvouhodnotové impulsní nebo statické signály. Připojení spínačů a tlačítek k napájecímu napětí symbolizuje přivedení signálu o logické hodnotě jednička. Výstupní funkcí PLC je spojovat akční prvky jako jsou relé, stykače, elektrické ventily nebo jen signalizační osvětlení nebo zvukové signály na napájecí napětí a tímto způsobem je spustit.

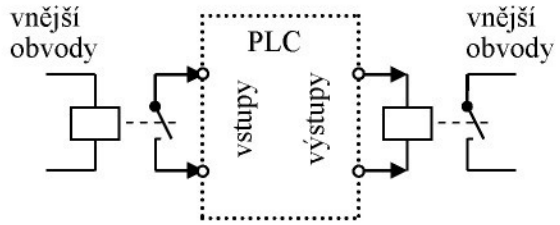


Obrázek 1.35 – Vazba PLC na řízený proces

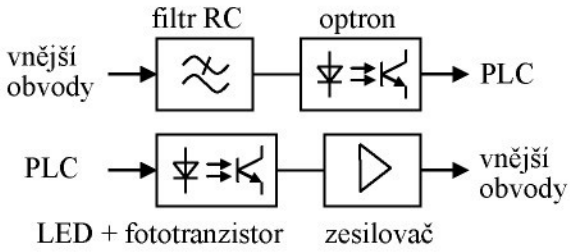
Napojení vnějších elektrických signálů vyžaduje jejich galvanické oddělení. Je nevýhodné, aby slaboproudé elektrické obvody PLC byly přímo propojeny se silovými obvody, které napájí například výkonné pohony. Důvodem je ochrana před jejich poruchami.

Parazitní impulsy, které se šíří po elektrickém vedení, mohou rušit PLC nebo ji dokonce zničit. Galvanické oddělení znamená přerušení přímého propojení elektrických obvodů. Nejjednodušším řešením galvanického oddělení obvodů je mechanické relé, viz. obrázek 1.36. Mezi cívkou relé a spínanými kontakty nemůže protékat elektrický proud a také průrazné napětí je velmi vysoké. Pro vstup se použijí kontaktní vstupy PLC a pro výstup dvouhodnotový napěťový výstup s úrovní, která stačí pro ovládání cívky relé. Nevýhodou mechanických relé jsou odskoky kontaktů při sepnutí. Tento děj je velmi krátký a u kontaktních obvodů nevadí. Avšak velmi rychlé elektronické obvody mohou změny stavu mechanických kontaktů zaznamenat a přivést sekvenční algoritmus do chybového stavu. Tuto nevýhodu lze odstranit Schmittovým klopným obvodem. Jiným řešením je použití klopného obvodu RS.

Optické galvanické oddělení zajišťuje LED dioda a fototranzistor podle obrázku 1.37 v prvku, který se nazývá optron. Tato dvojice představuje běžný obvodový prvek. Ke snížení rušení indukovanými vysokofrekvenčními signály je na vstupech z řízeného procesu dolnopropustný filtr RC. Naopak při oddělení výstupu je třeba signál pro ovládání připojených zařízení zesílit.



Obrázek 1.36 – Galvanické oddělení pomocí relé



Obrázek 1.37 – Galvanické oddělení pomocí LED diody a fototranzistoru

❑ Funkce sběrnice pro komunikaci

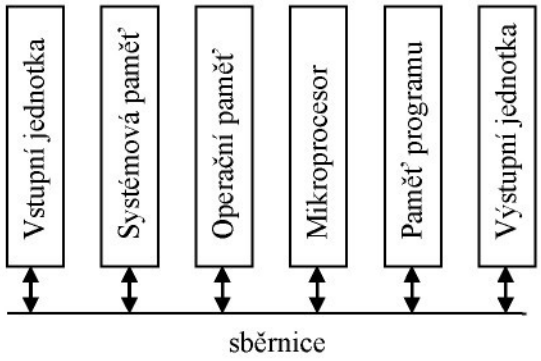
Obecným úkolem sběrnice je zajistit přenos dat a řídicích povelů mezi dvěma a více elektronickými zařízeními. Přenos dat na sběrnici se řídí stanoveným protokolem, což je souhrn pravidel, podle kterých se předávání informací řídí. V případě modulární architektury elektronického zařízení nebo počítače je sběrnice po mechanické stránce vybavena konektory, které jsou uzpůsobeny pro připojení modulů. Jednotlivé sběrnice se také liší hodnotami elektrického napětí pro úrovně H a L. Sběrnice se dělí na sériové a paralelní. V případě tzv. paralelní sběrnice přenášejí některé vodiče současně řídicí povely, další vodiče přenášejí adresu a data. U sériové sběrnice jsou tyto povely sdíleny na společném vodiči nebo vodičích a přenášejí se postupně.

Vnitřní jednotky počítače (paměti, procesor, atd.) třídy PC komunikují vzájemně prostřednictvím sběrnic. Propojení PC s vnějškem, tj. jednotkami, které jsou mimo skříňku počítače, je prostřednictvím rozhranní (anglicky interface), z nichž některé se také nazývají sběrnice. Příkladem paralelního rozhranní je třeba Centronics pro připojení tiskáren a nebo sběrnice IEE 488 pro připojení měřicích přístrojů. U sériového rozhranní (sběrnice) jsou bity zpráv přenášeny postupně. V obecném případě lze u sériového rozhranní rozlišit signály řídicí a datové. Datový přenos je zprostředkován jedním vodičem. Příkladem sériového rozhranní je proudová smyčka a sběrnice RS 232C, RS 422, RS 423 a RS 485. Proudovou smyčkou se v minulosti připojovaly k počítači dálnopisy. Sběrnice RS 232C patřila také k velmi rozšířeným typům rozhranní, které přenášelo data asynchronně po zprávách v řetězcích znaků. Toto rozhranní sloužilo k propojení jednoho vysílače s jedním přijímačem nebo v případě RS 422 s více přijímači. Sběrnice RS 485 umožňuje obousměrný přenos informací s více vysílači, které jsou připojeny k jedné sběrnici.

V případě nejnovějších počítačů začíná dominovat rozhranní USB (Universal Serial Bus), což je sériová polyfunkční sběrnice se 2 diferenciálními datovými vodiči a 2 napájecími vodiči (5 V/500 mA). Ve verzi 1.1 je přenosová rychlost 12 Mbit/s a ve verzi 2.0 je tato rychlost 480 Mbit/s. USB rozhraní používá dva typy konektorů a to plochý konektor „typ A“, který lze dnes připojit k prakticky každému novému PC aspoň do dvou přípojek, a „typ B“, který je určen pro periferní zařízení. Maximální délka kabelu mezi dvěma zařízeními je 5 metrů. USB dovoluje připojit až 127 zařízení pomocí jednoho konektoru. Připojeným zařízením poskytuje USB zároveň i napájecí napětí 5V s omezeným proudovým odběrem do 500 mA.

Vnitřní struktura PLC

PLC představuje mikropočítač se sběrnicovou strukturou, která je znázorněna na obrázku 1.38. Na společnou sběrnici jsou připojeny následující jednotky:



- Vstupní jednotka na kterou jsou připojeny vstupní signály číslicové a analogové
- Systémová paměť pro vstupy a výstupy, mezivýsledky, meze pro časovače a meze pro čítače
- Operační paměť s aktuální verzi programu, tvoří ji rychlá paměť typu RAM

Obrázek 1.38 – Vnitřní struktura PLC

- Mikroprocesor
- Paměť programů, které se zavádějí po restartu nebo resetu, tvoří ji obvykle paměť typu EPROM
- Výstupní jednotka členěná na analogové a digitální výstupy

Vstupy a výstupy PLC

Vstupní jednotka zajišťuje převzetí informací z řízeného procesu do PLC. Tato kapitola pojednává především o dvouhodnotových signálech. Tyto signály vstupují do počítače prostřednictvím tzv. číslicových (digitálních) vstupů. Každý vstupní dvouhodnotový (binární) signál obsahuje informaci o velikosti jednoho bitu. Číslicové vstupy zachycují nejen daný stav hodnot logických signálů, ale některé z nich jsou tzv. čítačového typu, tj. počítají počet příchozích impulsů.

Napěťová úroveň digitálních vstupů odpovídá nejčastěji TTL, tj. L/H odpovídá 0V/5V. Například komunikace přes sériový port počítače PC (RS 232) pracuje s úrovněmi L v mezích -30V až +5V a H v mezích +13V až +60V. Číslicové výstupy pracují s 24V a proudovou zatížitelností 2A.

Kromě číslicových vstupů je u PLC omezené množství tzv. analogových vstupů, které měří fyzikální veličiny jako teplota, tlak, a pod., tj. výsledkem měření je číslo. Převod signálu, nejčastěji napěťového, na číslo se uskutečňuje v tzv. analogově-číslicovém (analogově-digitálním) převodníku (convertor) ADC, česky A/Č. Rozsah převodníku je určen počtem bitů, což definuje rozsah celého čísla, které je výsledkem převodu. Pro převodník s m bity lze obdržet 2^m hodnot. Rozsah snímače ve fyzikálních jednotkách a počet různých hodnot na výstupu převodníku určuje rozlišitelnost měření a také jeho absolutní přesnost. Obvyklé počty bitů u převodníků pro účely řízení jsou 8 nebo 12 bitů, tj. rozsah celých čísel, které jsou výsledkem převodu, je pro unipolární převodníky od 0 do 255 nebo od 0 do 4096 a nebo pro bipolární převodníky od -128 do 127 nebo od -2048 do 2047. Samozřejmě je třeba tato celá

čísla přepočítat na skutečné fyzikální jednotky pro jejich zobrazení na displeji, který je použit jako řídicí a informační panel.

Analogové vstupy se rozdělují na

- symetrické - diferenční pro měření rozdílových napětí
- nesymetrické – se společnou nulou, počet nesymetrických vstupů je dvakrát větší než počet vstupů symetrických. Nesymetrické jsou analogové výstupy.

Další rozdělení vstupů je podle polarity vstupního napětí

- Unipolární, např. 0 až +5V, 0 až+ 10V
- Bipolární, např. ±5V, ±10V.

Výstupní jednotka zajišťuje předání dvouhodnotových a vícehodnotových informací ven z PLC do řízeného procesu. Dvouhodnotové veličiny vystupují staticky nebo také jako sled impulsů o zadané frekvenci. Kromě dvouhodnotových veličin je výstupem také analogová veličina jako výsledek číslico-analogového (DAC digitálně analogového) převodu. Počet těchto výstupů je obvykle jeden nebo dva.

Vlastnosti PLC se charakterizují především počty různých druhů vstupů a výstupů

- Binárních vstupů – DI (Digital Input), počet je násobkem 16, např. 16 DI
- Binárních výstupů – DO (Digital Output) , počet je shodný s počtem DI, např. 16 DO
- Analogových vstupů – AI (Analog Input), počet je obvykle 8 AI, zřídka 16 AI
- Analogových výstupů – AO (Analog Output), počet je obvykle jeden nebo dva, 1 AO nebo 2 AO.

❑ Paměti PLC

Jak již bylo zmíněno, součástí PLC jsou paměti. Jejich typy jsou jako u ostatních číslicových zařízení nebo počítačů následující:

ROM – paměť, kterou jejíž obsah lze jen číst (Read Only Memory). Její funkce byla zmíněna v popisu vnitřní struktury PLC. Potřeba přeprogramování vynutila zavést programovatelný ROM s označením PROM (Programable ROM). Mazání obsahu paměti pro nové nahrání dat u typu EPROM (Erasable PROM) se zajišťovalo ultrafialovým zářením. Mazání elektrické je u typu EEPROM (Electricaly EPROM).

RAM – paměť s náhodným přístupem (Random Access Memory) pro opakovaný zápis a čtení paměťové buňky s libovolnou adresou. Tyto paměti se dělí na dynamické RAM (DRAM) a statické RAM (SRAM). Paměť jednoho bitu u SRAM představuje klopný obvod se čtyřmi tranzistory. Obvody s SRAM pamětí jsou dražší ve srovnání s DRAM pamětí, ale rychlé a používají se u počítačů jako cache (vyrovnávací) paměti. Paměťová buňka pamětí SRAM je tvořena tranzistorem a elektrickým kondenzátorem. Protože se kondenzátor časem vybíjí, je třeba, aby paměť obsahovala podpůrné obvody pro cyklickou obnovu náboje kondenzátoru. Paměť SRAM má obrovskou kapacitu (stovky MB).

❑ Programování PLC

Programování PLC využívá softwarové pomůcky, takže uživatel v podstatě zadává dříve popsaným způsobem svou specifickou úlohu logického řízení. Tvůrci jazyků pro PLC využívají kontaktních schémat, blokových schémat a sekvenčních funkčních diagramů, ale

také i jazyka symbolických adres. Každý výrobce řídicích systémů zavádí drobné změny v terminologii a označování, které lze při znalosti obecných teoretických základů snadno překonat. Smysl odchylek je udržet starého zákazníka neodradit nového zákazníka pro své produkty.

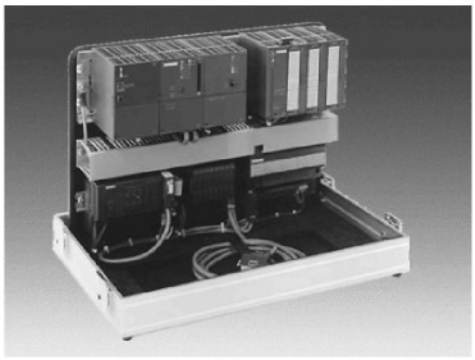
Ke studiu lze doporučit také další skripta, které katedra ATŘ vydala [2] a které obsahují řadu aplikačních příkladů jako vzor pro řešení mnoha úloh logického řízení.

❑ Provedení PLC

Jako mnoho jiných zařízení, PLC se dodává ve dvou verzích, a to v provedeních, které jsou označovány

- Kompaktní, jehož provedení má unifikovaný počet vstupů a výstupů a rozsah operační paměti.
- Modulární PLC má volitelnou konfiguraci v počtu všech druhů vstupů. Vstupy a výstupy se sdružují do modulů, jejichž počet je volitelný.

Konfigurace PLC se navrhuje s bezpečnou rezervou v počtech vstupů a výstupů a dalšími parametry. Příkladem PLC na trhu jsou v Evropě výrobky firem Siemens, Festo, AEG, Bosh a další. Firma Siemens vyrábí například PLC řady SIMATIC S7-300 a S7-400. Ukázka je na obrázku 1.39.



Obrázek 1.39 – PLC SIMATIC S7-300

Další možností je použít síť PLC propojených sběrnicí nebo průmyslové (Industrial) provedení PC (Personal Computer). Pro PC jsou na trhu zásuvné karty, které umožňují

❑ Sběrnice PROFIBUS

Složité nebo prostorově rozprostřené technologické procesy se řídí několika PLC, tj. jejich sítí. Ušetří se tím na kabeláži, protože vzájemné propojení PLC je jednodušší než dlouhá kabelová vedení od všech snímačů a akčních členů. Systém řízení je takto decentralizován a tím také odolnější proti poruchám. Komunikace mezi jednotlivými PLC se uskutečňuje po sběrnici (anglicky bus), což je skupina signálových vodičů. Funkční, elektrické a mechanické vlastnosti sběrnice PROFIBUS (PROcess FIeld BUS) pro vzájemné propojení PLC byly normalizovány v Německu v roce 1987. (norma DIN 19 245). Jako fyzické komunikační medium je sběrnicí RS-485, jejíž základní vlastnosti byly zmíněny již dříve.

Přístupová metoda sítě může být

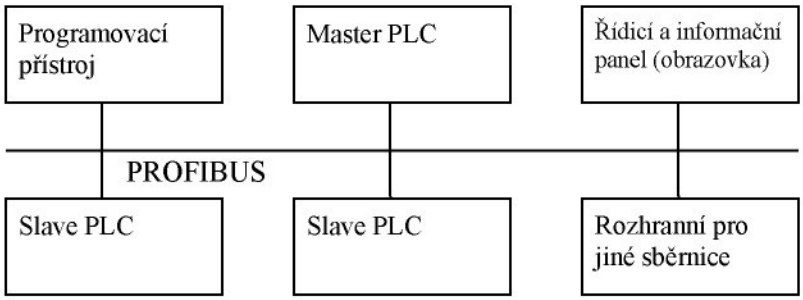
- token passing, u kterého se pověření řídit předává postupně jednotlivým stanicím sítě PLC v kruhu
- master-slave (pán-otrok) pro pevně stanovené řídicí a podřízené PLC

- hybridní, kdy se ve funkce řídicí stanice střídá u některých stanic jako v případě token passing, zatímco zbytek stanic jsou představuje stanice podřízené.

Maximální délka sběrnice je 1,2 km, s použitím opakovače (zesilovače) až 4,8 km. Přenosová rychlost byla nejdříve normalizována na 9,6, 19,2, 187 a 500 kbit/s, později došlo k rozšíření na 1,5, 3, 6, 9 a 12 Mbit/s.

Pod označením PROFIBUS jsou normovány tři sběrnice s označením Profibus DP, Profibus FMS a Profibus PA, které se liší v délce rámce, což je porce dat, které si PLC vzájemně posílají.

Příklad propojení PLC sběrnicí PROFIBUS je na obrázku 1.40. V této síti je jedno PLC řídicí (master) a dvě PLC podřízené (slave). Ke sběrnici je rovněž připojen programovací přístroj k úpravám (editaci) programu a kontrole jeho činnosti, případně k dalším servisním službám. Ke sběrnici je připojen také řídicí a informační panel (monitor). Sběrnice PROFIBUS může být napojena prostřednictvím rozhraní (interface) na sítě odlišného typu.



Obrázek 1.40 – Propojení několika PLC sběrnicí PROFIBUS

Funkci řídicího počítače může vykonávat také PC se zásuvnou kartou umožňující komunikaci se sběrnicí PRPFIBUS (viz řešení SIMATIC WinAC od firmy Siemens AG).

Σ Shrnutí pojmů 1.7.

Začátek kapitoly je věnován galvanickému oddělení PLC od vnějších obvodů. Dále pak je zda popsána vnitřní struktura PLC, vstupy a výstupy, paměti, způsoby programování, provedení PLC, používané sběrnice.

? Otázky 1.7.

1. Co patří mezi stavební prvky logického řízení?
2. Co jsou to programovatelné logické automaty?
3. Co je to PLC v modulárním provedení?
4. Jaké vstupy a výstupy obsahují PLC?
5. Jaké jazyky používáme k programování PLC?

1.8 Laboratorní úlohy na logické řízení

Čas ke studiu: 1 hodina

Cíl Po prostudování tohoto odstavce budete umět

- popsat dvě aplikace programovatelných logických automatů – ApphaController - v laboratořích katedry.

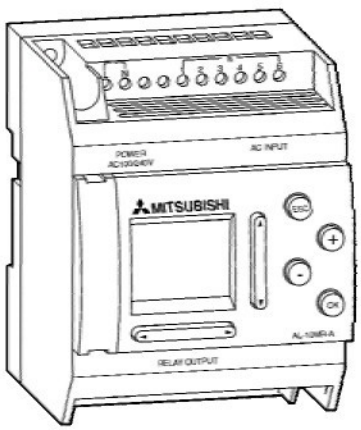
Výklad

Laboratoř k procvičování základů logických obvodů na katedře ATŘ, Fakulty strojní, je vybavena logickými automaty (PLC) Alpha-Controller od firmy Mitsubishi Electric, viz. obrázek 1.41. Jedná se o moderní kompaktní logické automaty určené pro přímé nasazení do automatizačních úloh.

Automat může být vybaven až dvanácti vstupy a osmi reléovými výstupy, praktické nasazení podporuje napěťový rozsah všech vstupů a výstupů 100–240 V AC (50/60 Hz), reléové výstupy jsou zatížitelné až do proudu 8 A.

Každé pracoviště je vybaveno kromě PLC také počítačem PC a jednoduchým modelem řízené soustavy. V počítači se řešená úloha nejprve naprogramuje a po překladů do strojového kódu se zavede (nahraje) do paměti PLC. Programování úlohy se uskuteční v grafickém prostředí, propojováním jednotlivých funkčních bloků. Zobrazování bloků se v detailech liší od bloků znázorněných v teoretické části tohoto učebního textu, které pojednávají o logickém řízení. Označení bloků pro standardní logické funkce, tj. logický součet a součin, a negování signálů zůstalo zachováno. Programátoři firmy, dodávající zmíněné PLC s příslušným software pro tvorbu zapojení úlohy řízení, dávají přednost staršímu grafickému znázornění bloků logických funkcí a kombinují je s obdélníkovými bloky nového značení, které tyto starší značky orámují. Na rozdíl od realizace bloku pro negaci pomocí hradel NOR nebo NAND je použit speciální blok s označením NOT (trojúhelník uvnitř bloku se znakem negace na vrcholu, který je spojen s výstupem bloku). Navíc jsou k dispozici speciální bloky, které plní například funkci čítače (Counter). Vstupy a výstupy bloků označuje patřičně orientovaná šipka.

Tvůrce zapojení logického obvodu, který plní požadovanou funkci logického řízení, vkládá na plochu monitoru jednotlivé bloky a vzájemně je propojuje spoji jako kdyby kreslil zapojení tužkou na papír. Křížení nebo odbočka se spojením signálů se označuje tečkou. Vstupní signály jsou označeny identifikátorem, který začíná písmenem I s připojeným číslem. Identifikátor výstupu začíná písmenem O, ke kterému je připojeno číslo. Typy vstupů a výstupů popisuje ikonka s upřesňujícím textem. Vnitřní signály začínají identifikátorem B a k tomuto písmenu je



připojeno číslo.

Obrázek 1.41 - PLC Alpha-Controller

Blokové schéma se podpůrným software převede (zkompiluje) do strojového kódu, který se nahraje do paměti PLC a zároveň spustí.

V tomto učebním textu budou demonstrovány dvě laboratorní úlohy, které řeší reálné problémy technické praxe. Je to blokování rozběhu motoru jako ochrana před úrazem obsluhy a řízení dopravní cesty, kterou tvoří pásové dopravníky.

❑ Zařízení na ochranu před úrazem (Bezpečnostní zařízení)

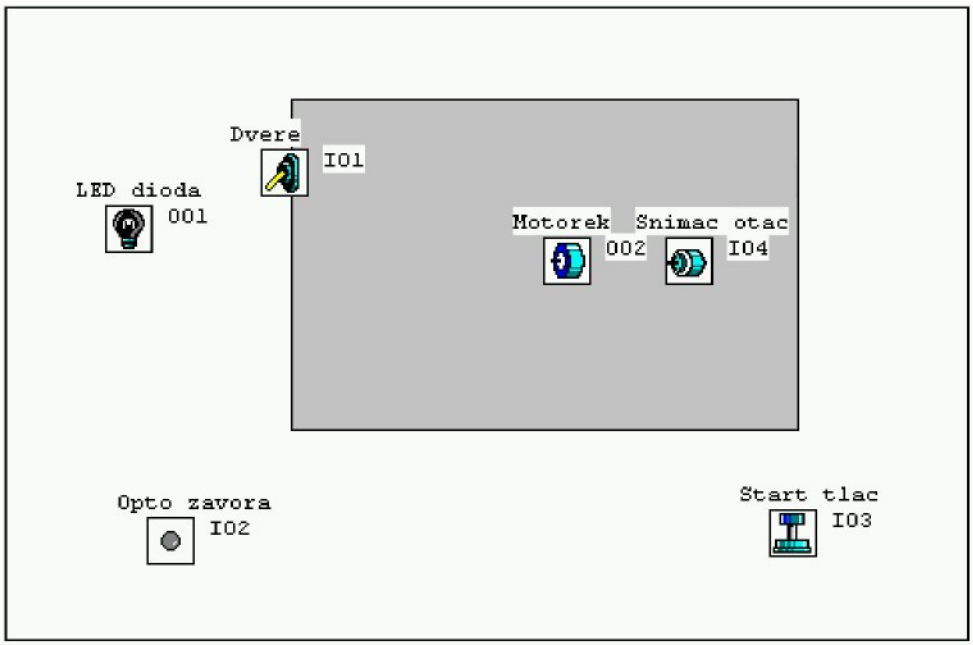
Zadání

Navrhněte způsob blokování spuštění chodu motorku tlačítkem za podmínek, že dveře do pomyslné strojovny budou otevřeny a světelná závora nebude přerušena. Jestliže budou uvedené podmínky splněny, pak se motorek rozběhne a v opačném případě se zase rozsvítí kontrolní světelná dioda (LED). Motorek má vykonat po stisku startovacího tlačítka stanovený počet otáček a pak se zastavit.

První varianta řešení

Na modelu strojovny je otevření dvířek indikováno jazýčkovým relé, které je ovládáno magnetickým polem. Otočení motorku je snímáno inkrementálním snímačem, který generuje posloupnost impulsů. Schéma laboratorního modelu je uvedeno na obrázku 1.42. Kromě rozmístění součástek jsou v obrázku identifikátory vstupních a výstupních logických proměnných.

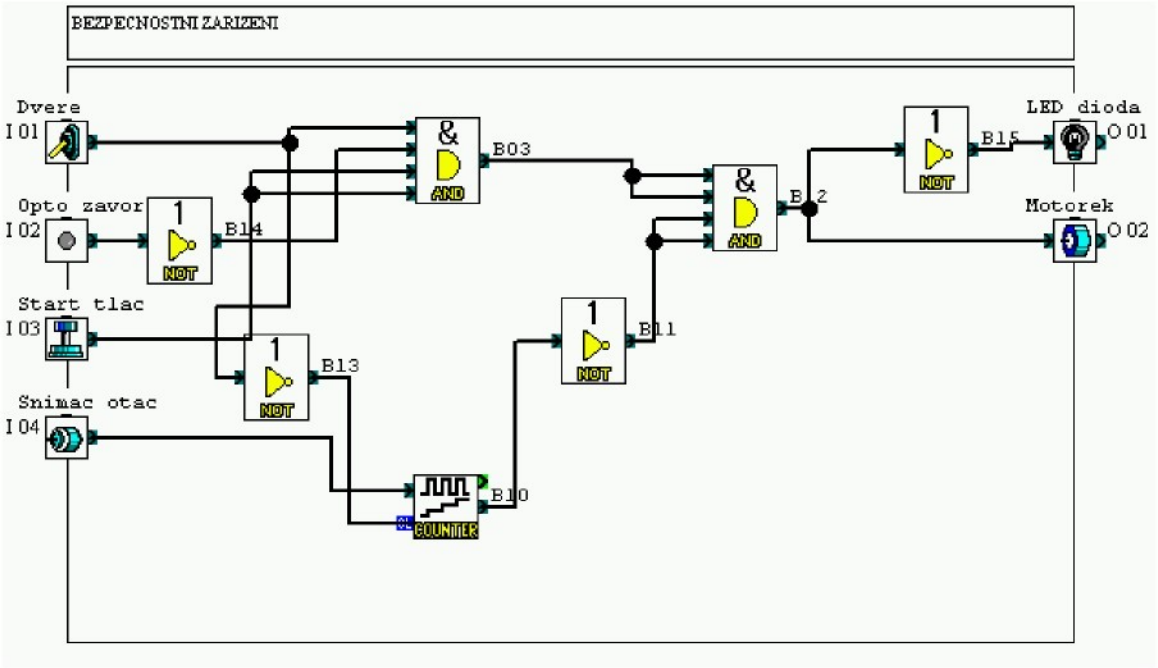
Řešení úlohy na logickém automatu Alfa je znázorněno na obrázku 1.43. Program je vytvořen s použitím standardních logických funkcí pro testování spouštěcích podmínek (logické jedničky na vstupech I01, I02 a I03), které generují logický signál B03 na výstupu hradla AND a předdefinovaného bloku Counter ve funkci čítače impulsů, které jsou generovány inkrementálním snímačem při otáčením motorku. Po dosažení přednastaveného počtu impulsů od otáček motorku se změní logický výstup čítače (B10) na logickou jedničku. Čítač reaguje na nástupnou hranu vstupního signálu I04 tak, že zvětší hodnotu čítače o jedničku. Negace logické proměnné B10 v logickém součinu s podmínkami pro spuštění motorku (signál B03) vynuluje výstup hradla AND, který je označen B12), a napájení motorku z výstupu O02 se odpojí. Negace signálu B12 ovládá výstup O01 pro svícení LED diody. Nevýhodou zapojení na obrázku 1.43 je to, že po celou dobu otáčení motorku je nutné přidržovat startovací tlačítko stlačené (vstup I03) a udržovat přerušený světelný tok v optické závoře (vstup I02). Pro další rozběh motorku je zapotřebí otevřít dvířka modelu (vstup I01), čímž se vynuluje čítač pulsů a motorek je připraven vykonat další sérii otáček.



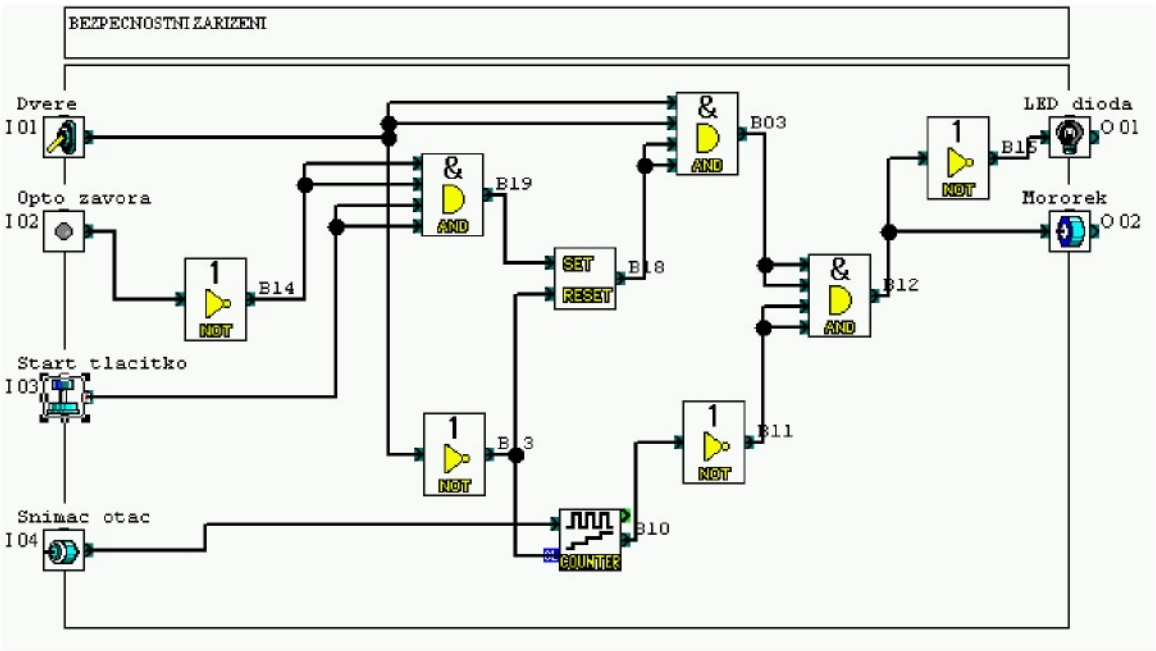
Obrázek 1.42 - Schéma laboratorního modelu

Druhá varianta řešení

Další řešení ovládání modelu pracuje na stejném principu jako předchozí řešení tím, že při zavřených dvířkách stačí krátkodobě stlačit startovací tlačítko a současně krátkodobě přerušit světelný tok v optické závoře. Ve schématu na obrázku 1.44 je použit klopný obvod (SET-RESET) typu RS, jehož vstupní signál B19 pro nastavení (SET) je odvozen od zadaných vstupních podmínek. K překlopení výstupu klopného B18 obvodu do logické jednotky stačí krátkodobý stisk startovacího tlačítka a současné přerušení světelného toku. Logický součin AND signálu B18 a signálu zavření dveří I01 ovládá přes další logický součin AND s výstupem B12, který určuje také přes výstup celého obvodu O02 běh motorku samozřejmě za podmínky, že nebylo dosaženo stanoveného počtu otáček (signál B11). Otevření dvířek nuluje čítač pulsů a prostřednictvím vstupu RESET do klopného obvodu nastavuje jeho výstup B18 na logickou nulu. Logická nula na výstupu B03 bloku AND nuluje také výstup B12 dalšího bloku AND a také výstup celého obvodu O02, což způsobí zastavení motorku pokud je ještě v chodu.



Obrázek 1.43 - Schéma zapojení funkčních bloků



Obrázek 1.44 - Řešení bez nutnosti držení tlačítka a zaclonění závory

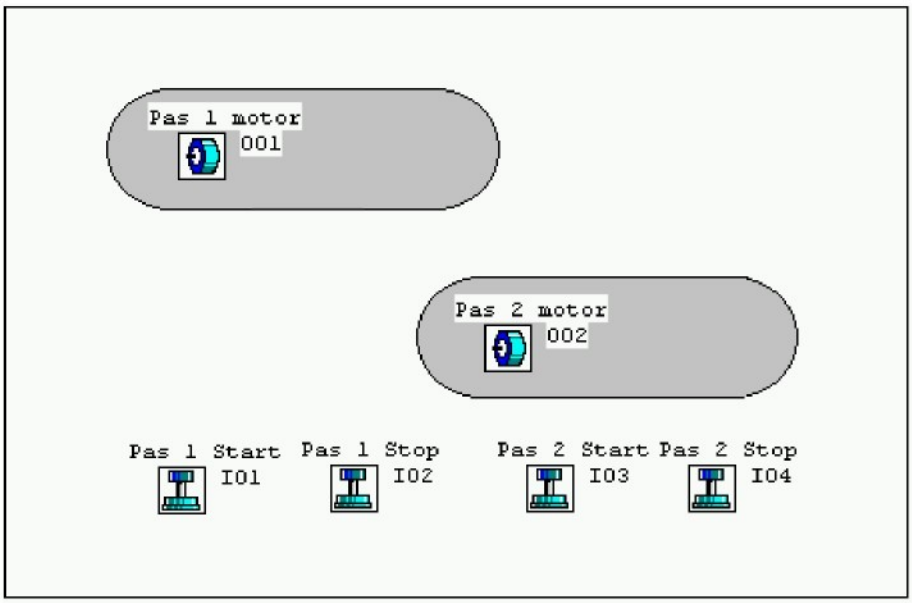
❑ Řízení pásových dopravníků (Dva dopravníky, R-S klopný obvod)

Zadání

Je třeba navrhnout ovládání dopravní linky, která se skládá ze dvou navazujících pásových dopravníků. Aby nedošlo k zavalení přesypu dopravovaným materiálem nesmí být první pás spuštěn pokud navazující pás stojí a také nesmí být druhý pás zastaven pokud první pás zůstane běžet.

První varianta řešení

Schéma uspořádání dopravní linky s automatem Alfa je na obrázku 1.45. Pohon pásového dopravníku je ovládán pomocí tlačítkových spínačů. Každý z obou dopravníků má svá spouštěcí a vypínací tlačítka. Celou dopravní linku dáme do pohybu pomocí spínače PAS 1 Start. Spínačem PAS 1 Stop zastavíme pouze dopravník č.1, přičemž dopravník č.2 pokračuje v chodu. Spínač PAS 2 Start slouží pouze ke spouštění dopravníku č.2 a spínačem PAS 2 Stop zastavíme celou dopravní linku (pokud je v provozu) nebo jen dopravník č.2 (pokud dopravník č. 1 stojí).



Obrázek 1.45 - Uspořádání dopravní linky

Řešení úlohy je znázorněno na obrázku 1.46. Ze struktury zapojení je patrný způsob programování automatu Alfa pomocí funkčních bloků NOR, kde výstup z tohoto bloku je v logické jedničce (ON), jsou-li všechny vstupy na logické nule (OFF) a výstup je na logické nule (OFF), je-li alespoň jeden vstup na logické jedničce (ON). Dvojice bloků NOR s výstupy B07 a B08 realizuje první klopný obvod pro ovládání pohonu dopravníku č.1 a druhá dvojice bloků B12 a B13 realizuje druhý klopný obvod pro ovládání dopravníku č. 2. Oba klopné obvody jsou typu RS, tj. signál SET nastavuje klopný obvod (jeho přímý výstup) do logické nuly a signál RESET nastavuje jeho výstup do logické nuly.

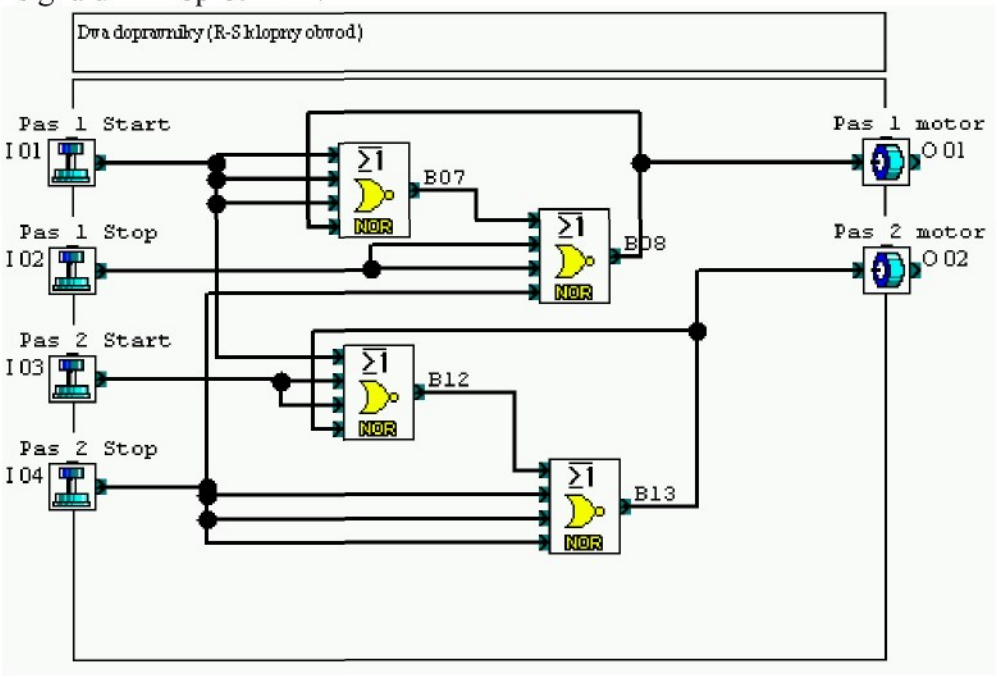
Dopravní cesty, které zabezpečují dopravní pásy, se spouštějí obvykle od koncového dopravníku, přičemž se zpožděním se uvádí do chodu předcházející dopravník postupně až k prvnímu pásu ve směru toku materiálu. Materiál se na dopravní pásy vpouští obvykle až všechny jsou v pohybu. Odstavení dopravní cesty se provádí od začátku dopravní cesty, přičemž další dopravníky se zastavují se zpožděním, aby se vyprázdnily. Je nebezpečné spouštět pás s materiálem na stojící dopravník z důvodu nebezpečí závalu přesypu. Ostatně přesypy jsou vybaveny snímači závalu, také s dvouhodnotovým výstupem.

Druhá varianta řešení

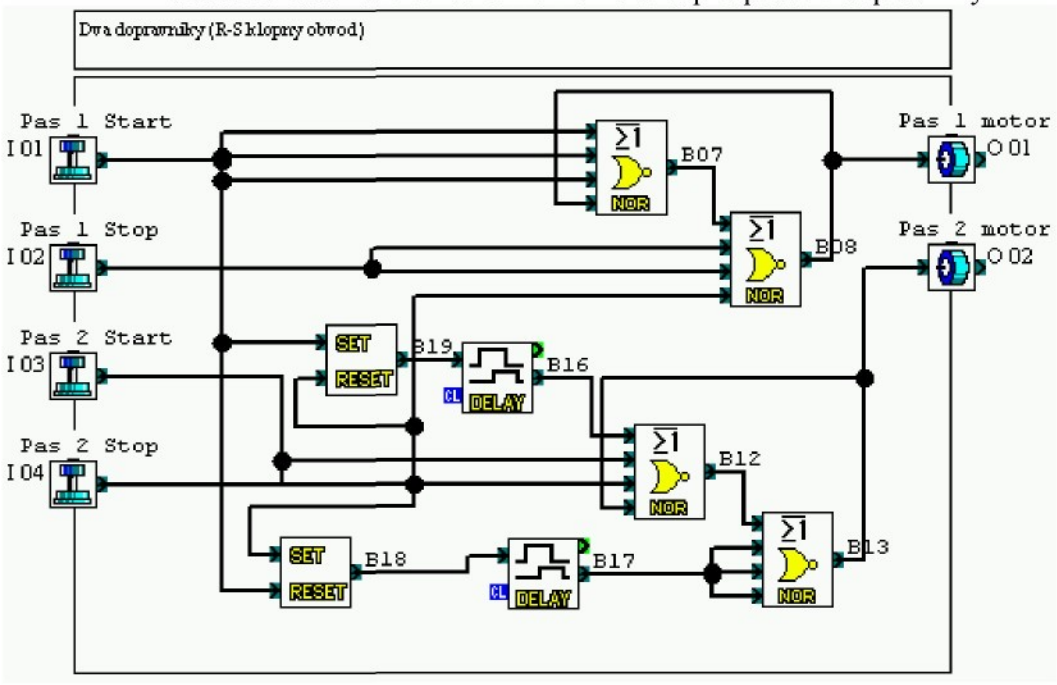
Funkce systému řízení na obrázku 1.47 je upravena použitím prvku zpoždění impulsu (Delay) a částečně řeší problém nebezpečí závalu přesypu. Po startu pásu č.1 se pás č.2 rozjede až po uplynutí definované doby, což není nejvhodnější způsob řízení. Obdobně po

zastavení pásu č.1 se pás č.2 zastaví po uplynutí stanovené doby, tj. až fiktivní materiál na páse vyběhne.

Příklad ukazuje možnost postupného zdokonalování funkce logického řízení a je zaměřen na použití bloků Delay, které realizují časové zpoždění signálu B16 oproti B19 a zpoždění signálu B17 oproti B18.



Obrázek 1.46 - Schéma funkčních bloků pro pásové dopravníky

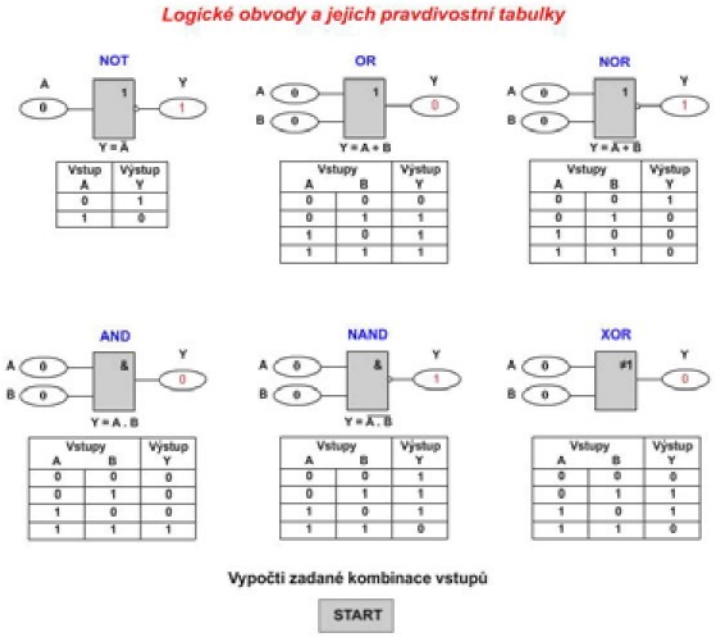


Obrázek 1.47 - Zapojení se zpožděním zapnutí a vypnutí pásu č. 2

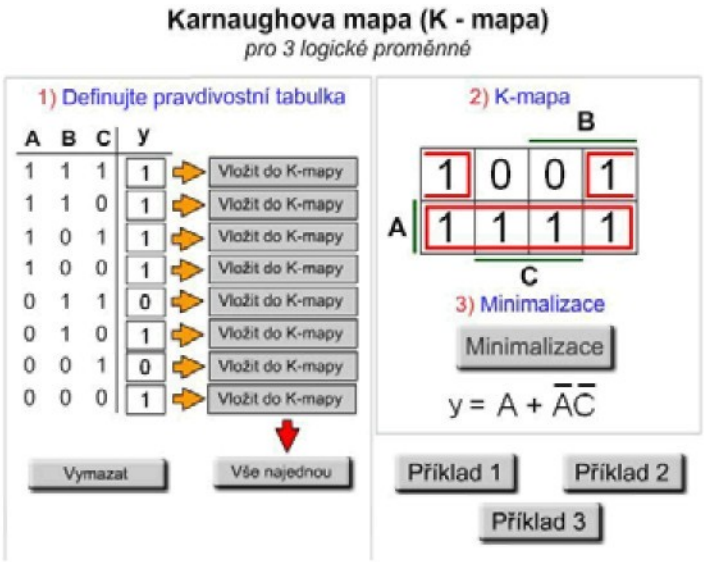
Shrnutí pojmů 1.8

Kapitola popisuje realizované úlohy logického řízení v laboratořích katedry. Jedná se o dvě úlohy a to: bezpečnostní zařízení, řízení pásových dopravníků. Pro obě úlohy je uvedeno zadání a dva způsoby řešení.

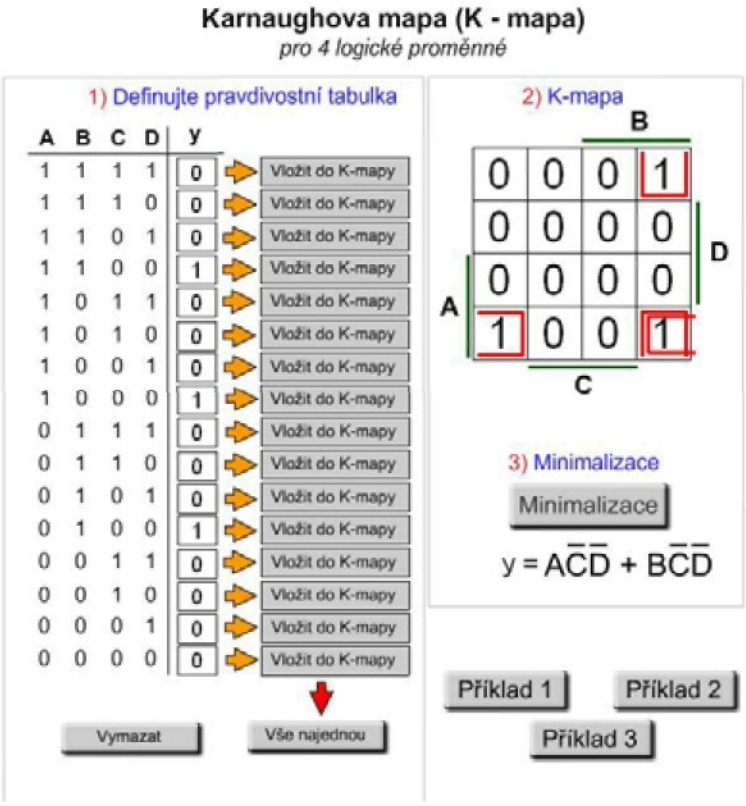
Kliknutím na následující obrázky můžete spustit animace vytvořené pro oblast logického řízení.



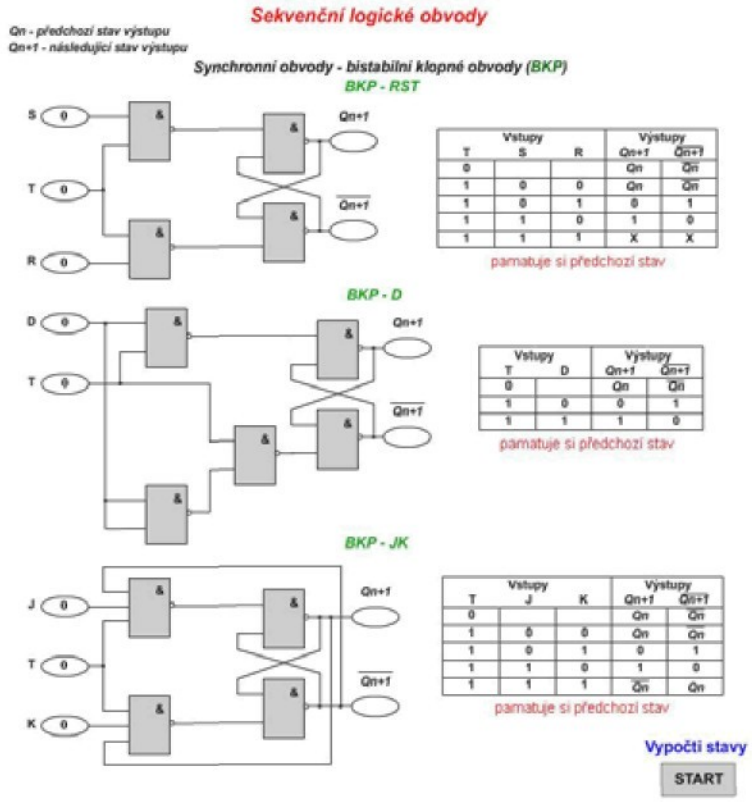
Obrázek 1.48 – Animace popisující základní logické operace



Obrázek 1.49 – Animace popisující převod mezi tabulkou hodnot a K-mapou



Obrázek 1.50 – Animace popisující převod mezi tabulkou hodnot a K-mapou



Obrázek 1.51 – Animace popisující základní logické obvody

Další zdroje

[1] ZÍTEK, P., HOFREITER, M. & HLAVA, J. *Automatické řízení.* Praha : skripta ČVUT, 2000. 148 s. ISBN 80-01-02044-4.

[2] FORMÁNEK, I. *Základy automatizace. logické řízení.* Ostrava : skripta VŠB-TU Ostrava, 2006, 86 s. ISBN 80-248-1012-3.

[3] ŠMEJKAL, L. & MARTINÁSKOVÁ, M. *PLC a automatizace. 1. díl : Základní pojmy, úvod do programování.* Praha : BEN – technická literatura, 1999, 223 s. ISBN 80-86056-58-9.

2. SPOJITÉ SYSTÉMY

2.1. Základní pojmy

Čas ke studiu: 1,5 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat blokové schéma, funkční schéma, pojmy systém, prvek, vazba.
- definovat pojmy řízení, ovládání a regulace.
- popsat dělení signálů a systémů z různých hledisek.
- popsat rozdíl mezi spojitým a diskrétním regulačním obvodem.

Výklad

Řízením stroje, technologického zařízení nebo procesu – zkrátka objektu – rozumíme dále jakýkoli cílevědomý způsob přinucení jej k činnosti od něj požadované – za pomoci zvlášť k tomu uzpůsobených technických prostředků. Pojem řízení také však úzce souvisí s pochopením kauzálních – příčinných vztahů, jimiž se činnost dotyčného objektu řídí. Jde tedy o seznámení se s principy pomocí nichž se řídí řešení problémů řízení obecně.

Nevyhnutelným východiskem pro skutečně správné řešení úlohy řízení daného objektu je pochopení jeho činnosti jako vztahu mezi působícími příčinami a jejich důsledky, jež jsou výsledky příslušné činnosti. Tyto vztahy nazýváme jako kauzální (orientovaná) relace. Schématicky můžeme tento vztah zapsat:

příčiny → důsledek,

kde šipkou je vyznačen neměnný "směr" závislosti, které proto říkáme také orientovaná. Graficky znázorňujeme kauzální relaci jako blok, viz obrázek 2.1, tedy v podstatě jako operaci, jíž se k zadaným hodnotám veličin reprezentujícím příčiny určují hodnoty veličin představujících důsledky. Zadané příčiny označujeme jak vstupy a vyplývající důsledky jako výstupy tohoto bloku.

Obrázek 2.1 - Grafické znázornění kauzálního vztahu

Pro řízení je příznačné, že se uskutečňuje jako součinnost dvou nebo více objektů (např. řízeného a řídicího), takže jejich kauzální relace se odpovídajícím způsobem kombinují. Spojením vzniká složený objekt, obecně zvaný **systém**. Za systém považujeme jakékoli účelové uspořádání jednodušších objektů ve složitější celek vyznačující se interakcí těchto objektů a výslednými vlastnostmi, jež tento celek charakterizují.

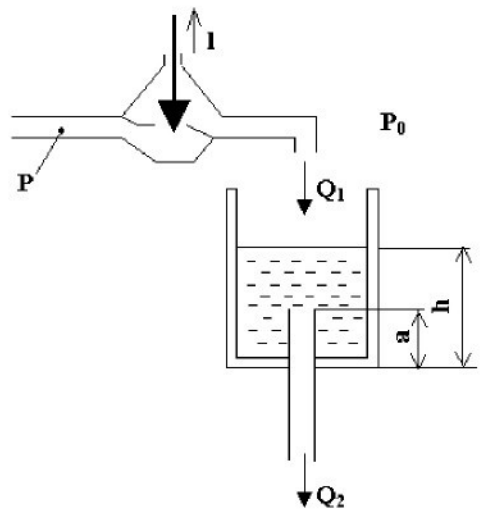
Při spojování dílčích objektů ve funkční celek – systém dílčí kauzální relace na sebe charakteristickým způsobem navazují a vytvářejí určité uspořádání zvané **struktura** systému.

Strukturu nejlépe znázorňujeme orientovaným grafem – nejčastěji tzv. **blokovým schématem**. V blokovém schématu je každá z dílčích relací reprezentována svým blokem podle obrázku 2.3 a orientovanými spojnicemi mezi bloky je vyjádřena vstupní, resp. výstupní úloha jednotlivých veličin v těchto blocích – relacích. Orientované spojnice u blokových schémat regulačních obvodů představují směr šířeného signálu. Jestliže se signál rozdvojuje do více bloků, označí se rozdvojovaní místo tečkou. Jestliže se naopak několik signálů algebraicky sčítá v jeden signál, označí se součtovým členem, při odečítání se příslušná část vyplní.

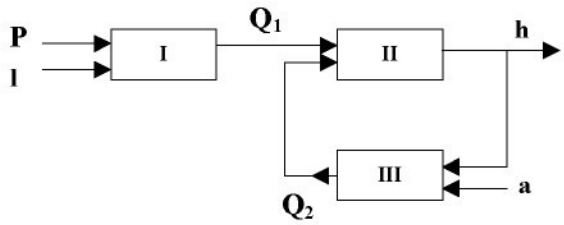
Vlastnosti bloků jsou nejčastěji popsány jejich přenosy $G(s)$. Může však také jít o popis diferenciální rovnicí, funkční závislostí, přechodovou funkcí nebo charakteristikou apod.

- **Příklad sestavení blokového schématu.**

Mějme nádobu k zadržení určitého množství protékající kapaliny danou výškou hladiny h – obrázku 2.2. Přítok Q_1 je závislý na otevření ventilu l a na tlaku před ním P. Odtok Q_2 závisí na výšce hladiny h a na poloze odtokového otvoru a. Označme závislost přítoku Q_1 římskou číslicí I, závislost přítoku Q_2 II a závislost změny výšky hladiny III.



Obrázek 2.2 - Nádrž s volným odtokem



Obrázek 2.3 - Blokové schéma nádrže s volným odtokem

Blokové schéma struktury systému, vytvořené na základě vzájemné návaznosti relací I, II, a III je znázorněno na obrázku 2.3.

Přenos se na blokových schématech označuje vepsáním funkční závislosti do příslušného bloku. Do bloku se taky může zakreslit přechodová charakteristika. Přenosem lze však popsat i větší části regulačních obvodů případně celé obvody. K tomu musíme ovšem znát metodiku, jak určit přenos celku, známe-li přenosy jednotlivých členů, z nichž se skládá. [Zítek, P. 1993]

Systémem rozumíme obecně soubor prvků, mezi nimiž existují vzájemné vazby a jako celek má určité vztahy ke svému okolí. Každý systém je charakterizován dvěma základními vlastnostmi:

- chováním systému, charakterizujícím jeho vnější vztahy k okolí; chování systému je závislost mezi podněty okolí systému působícími na jeho vstup a příslušnými odezvami objevujícími se na jeho výstupu,
- strukturou systému, charakterizující jeho vnitřní funkční vztahy; strukturou systému rozumíme jednak způsob uspořádáni (organizace) vzájemných vazeb mezi prvky systému a jednak chování těchto prvků.

Obě tyto vlastnosti systému jsou ve velmi úzkém vztahu, který lze charakterizovat jednak, že určité struktuře odpovídá jednoznačně určité chováni a naopak, že určitému chování odpovídá třída struktur, definovaná tímto chováním.

Podaří-li se nám definovat (vyčlenit) na zvoleném objektu jeho části a vztahy mezi nimi, pak jsme definovali na daném objektu systém. Uvažované části objektu nazýváme **prvky** (je tvůrcem dynamického chování systému; dělí se na: *vnitřní* - nemá spojení s okolním systémem, *vnější* - má spojení s okolním systémem) a vzájemné působení mezi těmito částmi **vazbami** (orientované čáry, slouží k přenosu informací z jednoho prvku na druhý nebo do okolí systému, dělí se na: *vnitřní* - přenáší informace mezi prvky, *vnější* - jsou to vstupy a výstupy, přenáší informace mezi systémem a okolím systému). Každý definovaný systém může být prvkem systému vyššího řádu – nadsystému (supersystému). Právě tak každý prvek systému může být systémem nižšího řádu -podsystémem (**subsystémem**). Tato vlastnost souvisí s rozlišovací úrovni. **Rozlišovací úroveň** je míra podrobnosti s jakou systém popisujeme.

Modelovaní je tvořeno dvěmi činnostmi a to:

- identifikaci - určení matematického popisu systému (*matematicky model* - soubor abstraktních rovnic a vztahů, které popisují dynamické chování systému)
- simulace - to je práce s matematickým modelem.

❑ Základní činnosti se systémy

1. Analýza systému

Na základě matematického popisu, který známe, určíme chování systému. Systém při analýze zůstává nezměněn. Do analýzy patří vykreslování charakteristik, kontrola stability.

2. Syntéza systému (inverzní k analýze)

Na základě známého chování a dané omezení struktury určíme matematický popis systému. Pomocí syntézy volíme typy regulátorů a hodnoty jejich stavitelných parametrů.

3. Černá schránka

Jakýkoliv systém, u kterého neznáme jeho strukturu a ani chování, nazýváme Černou schránkou. Máme stanovené omezení struktury a definované vstupní signály, na základě sledování chování systémů určíme matematický popis. Příkladem je experimentální identifikace. [BALÁTĚ 2003]

❑ Třídění signálů

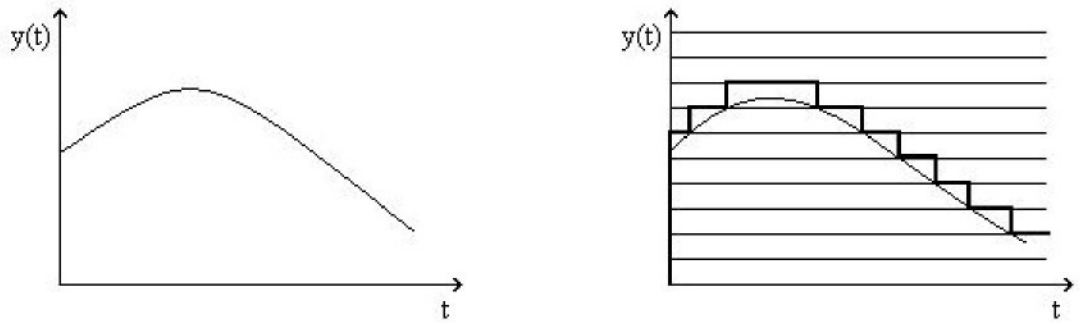
Signály lze rozdělovat na:

- spojité v čase (v každém časovém okamžiku víme jakou hodnotu má signál)
  - spojité v úrovni - analogový signál
  - diskrétní v úrovni - kvantovaný signál (Č/A převodníky)

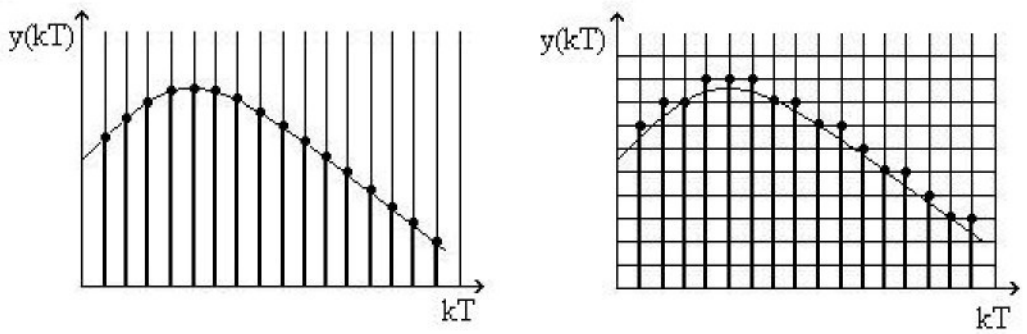
- diskrétní v čase (hodnotu signálu víme pouze v násobku vzorkovací periody)
  - spojité v úrovni – diskrétní signál získaný vzorkováním
  - diskrétní v úrovni – číslicový signál (A/Č převodník, měřící karty)

Dále lze signály dělit dle:

1) průběhu
   - periodické (goniometrické funkce)
   - neperiodické
2) matematického popisu
   - deterministické
   - stochastické (v daném čase jsou hodnoty náhodné)



Obrázek 2.4 - Analogový a kvantovaný signál



Obrázek 2.5 - Diskrétní a číslicový signál
(k - celé číslo, T - vzorkovací perioda)

Kliknutím na obrázek 2.6 se spustí animace, která vysvětluje převod od analogového signálu k číslicovému. Nejdříve se vykreslí vstupní/analogový signál (tlačítko vstupní signál), pak stisknutím tlačítek Vzorkování, Kvantování a Výstupní signál vykreslíme jednotlivé kroky převodu.

Obrázek 2.6 – Animace vysvětlující vzorkování a kvantování

❑ Třídění systému

Vždy je třeba určit hledisko, podle kterého chceme systémy třídit. Systémy lze dělit dle:

1) matematického modelu

- lineární (všechny rovnice, které popisují chování systému musí být lineární)
- nelineární (stačí aby 1 rovnice byla nelineární a celý systém bude nelineární)

2) typu signálu

- spojitý (všechny signály jsou spojité)
- diskrétní (všechny signály jsou diskrétní)
- hybridní (oba typy signálů, v praxi – PLC, čipy)

3) počtu signálu

- jednorozměrné SISO (jeden vstup a vystup)
- mnohorozměrné MIMO (více vstupů a výstupů)

4) změn vlastností v čase

- stacionární (funkce popsaná rovnicemi s konstantními koeficienty)
- nestacionární (rovnice jsou s proměnnými koeficienty)

5) změny stavu v čase

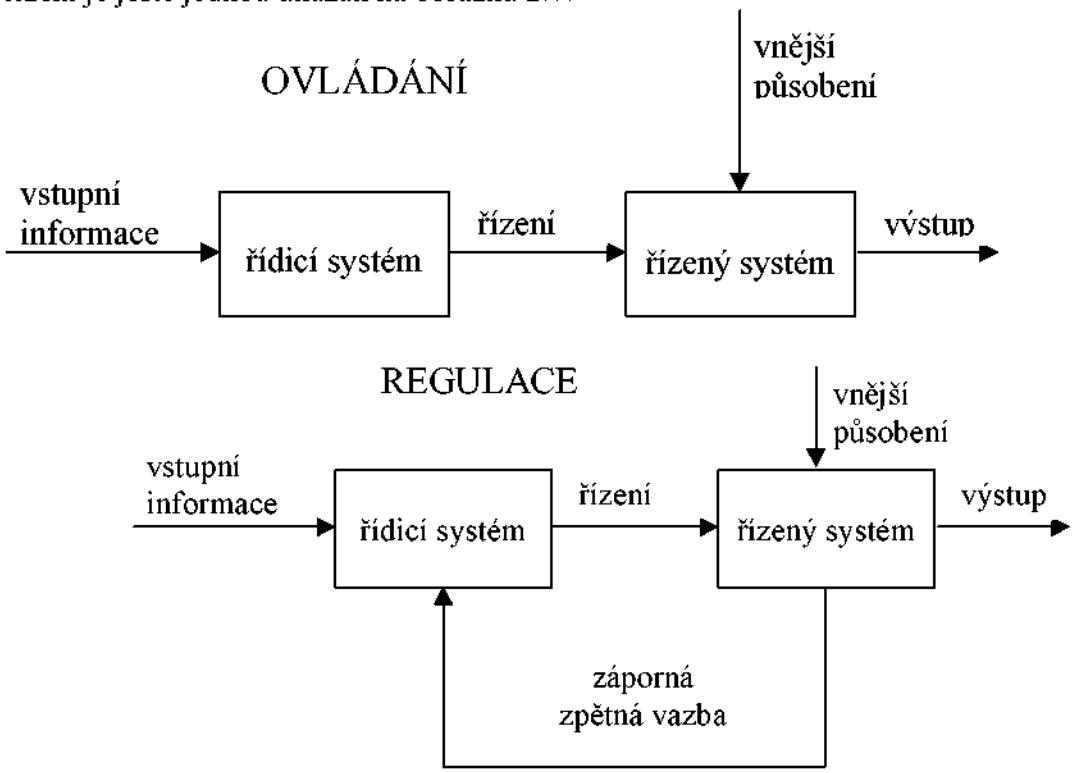
- statický (hodnota výstupu závisí pouze na okamžité hodnotě vstupů, chování systému lze popsat pomocí algebraických rovnic)
- dynamický (hodnota výstupu závisí na okamžité hodnotě vstupů ale i na historii hodnot vstupů a výstupů, chování systému lze popsat pomocí diferenciálních rovnic)

❑ Regulační obvody

Řízení je definováno jako cílevědomá činnost, při níž se hodnotí a zpracovávají informace o řízeném procesu nebo objektu i informace o dějích vně tohoto procesu a podle

nich se ovládají příslušná zařízení tak, aby se dosáhlo určitého cíle řízení. V zásadě může být ruční (např. řízení automobilu) nebo automatické (které bude dále blíže popsáno).

Řízení dělíme na ovládání a regulaci. **Ovládání** je řízení bez zpětné vazby. **Regulace** je udržování zvolené fyzikální veličiny na předem určené hodnotě. Během regulace se zjišťují hodnoty této veličiny a srovnávají se s hodnotou, kterou má mít. Podle zjištěných regulačních odchylek, které jsou mírou přesnosti regulace, se zasahuje do regulačního procesu v tom směru, aby tyto odchylky byly minimální. Charakteristickým rysem regulace je záporná zpětná vazba, která zajišťuje stabilizaci regulačního obvodu (RO). Rozdíl mezi oběma druhy řízení je ještě jednou ukázán na obrázku 2.7.

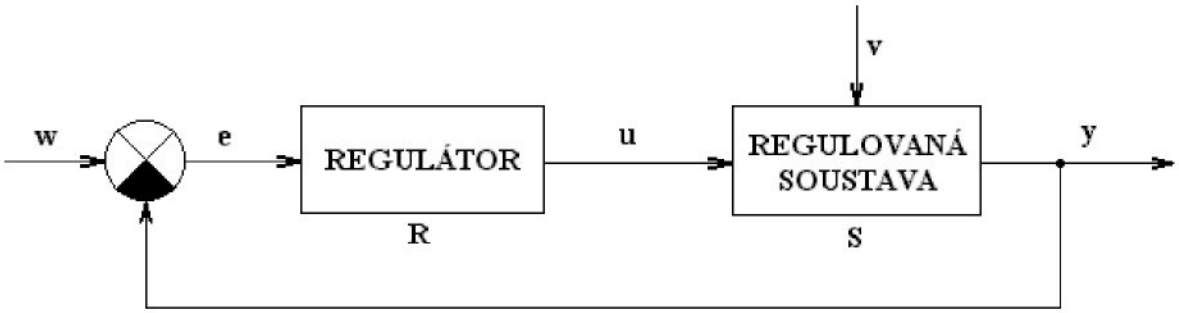


Obrázek 2.7 - Blokové schéma ovládání a regulačního obvodu

Regulační obvod (viz obrázek 2.8) je systém, který je tvořen dvěmi prvky, a to regulátorem a regulovanou soustavou. Regulátor na základě rozdílu mezi žádanou a regulovanou veličinou nastaví řídicí veličinu tak, aby byl splněn cíl regulace, tzn. $y \to w$ resp. $e \to 0$.

Dynamické chování celého regulačního obvodu můžeme popsat např. pomocí diferenciálních rovnic určující vztah regulované veličiny a žádané veličiny, nebo regulované veličiny a poruchy. Dynamické vlastnosti regulačního obvodu závisí na dynamických vlastnostech jeho jednotlivých členů. Za lineární regulační obvod budeme považovat takový obvod, jehož vlastnosti můžeme matematicky popsat s dostatečnou přesností lineárními diferenciálními rovnicemi s konstantními koeficienty.

Základní vlastností lineárních regulačních obvodů je princip superpozice, podle něhož reakce regulačního obvodu na součet podnětů se rovná součtu reakcí na jednotlivé podněty. Regulační obvody můžeme rozdělit na spojité a diskrétní [BOLTON 1992].



w - žádaná veličina; e - regulační odchylka; y - regulovaná veličina; v - poruchová veličina; u - řídicí veličina; R – regulátor; S - regulovaná soustava

Obrázek 2.8 - Blokové schéma jednoduchého regulačního obvodu

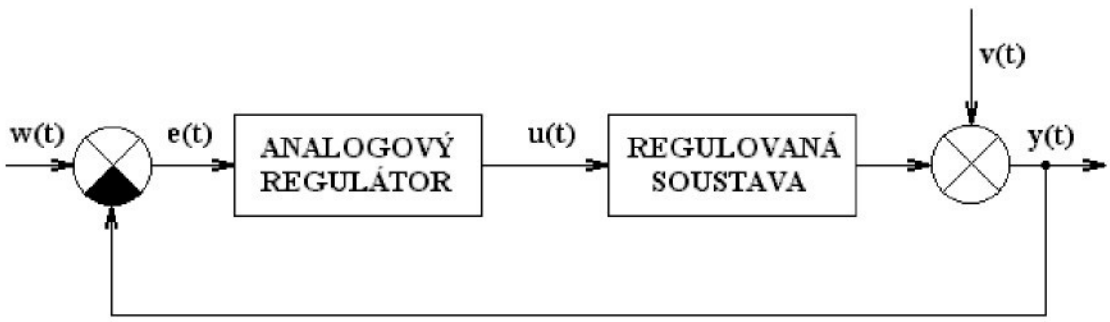
Spojité řízení je tam, kde jak akční zásah je spojitě nastavován, tak i údaje o řízeném systému jsou měřeny jako veličiny spojitě proměnné v čase.

Diskrétní řízení je dnes důsledkem nasazení počítačů jako regulátorů. U řídicích počítačů, které ani nedovedou zpracovávat spojitý signál, je nutný spojitý signál převádět na diskrétní.

Spojité regulační obvody

Jedná se o takové lineární regulační obvody, u nichž jsou všechny signály zpracovávány spojitě, tzn. všechny veličiny jsou funkcemi spojitého času t.

Blokové schéma lineárního regulačního obvodu znázorňuje obrázek 2.9. Regulátor na základě hodnoty regulační odchylky $e(t)$ nastaví akční veličinu $u(t)$ tak, aby byl splněn cíl regulace, tzn. $e(t) \to 0$.



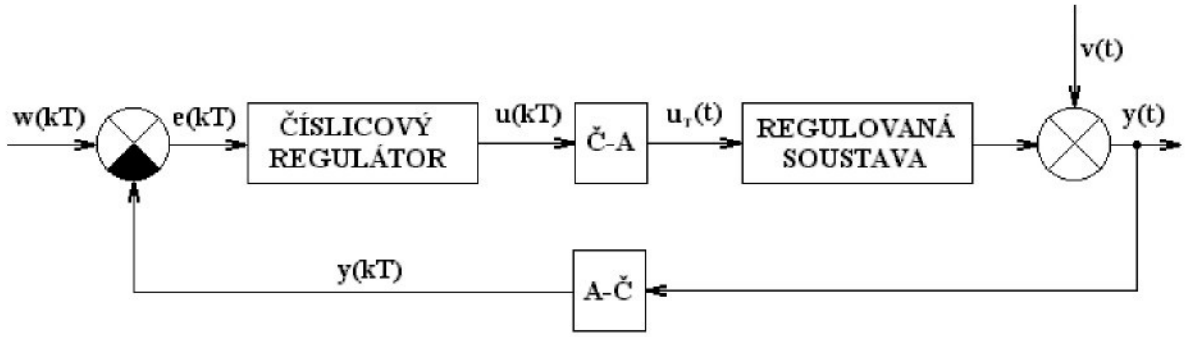
Obrázek 2.9 - Obecné blokové schéma jednorozměrného lineárního spojitého regulačního obvodu

Diskrétní regulační obvody

Diskrétní regulační obvod je takový obvod, ve kterém alespoň jedna veličina má tvar posloupnosti diskrétních hodnot vytvářených např. v časově ekvidistantních okamžicích. Obvykle si představujeme, že diskrétní regulační obvod využívá k výpočtu akční veličiny číslicový počítač, což však nemusí platit vždy.

Diskrétní regulační obvod,viz obr. 2.10, lze znázornit blokovým schématem skládajícím se ze spojitě pracujících členů, ze vzorkovacích členů, z tvarovacích členů,

nespojitě (diskrétně) pracujících členů, analogově-číslicových převodníků, z číslicově-analogových převodníků a popřípadě i z paměťových členů [BOLTON 1992].

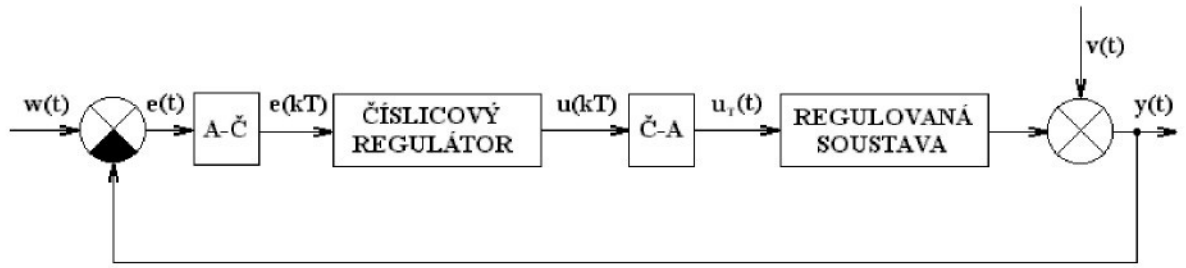


ČR - číslicový regulátor; A/Č - analogově číslicový převodník; Č/A - číslicově analogový převodník; k, kT - diskrétní čas $k = (0,1,2,...)$; u_T - tvarovaná akční veličina; T - vzorkovací perioda

Obrázek 2.10 - Obecné blokové schéma lineárního diskrétního regulačního obvodu

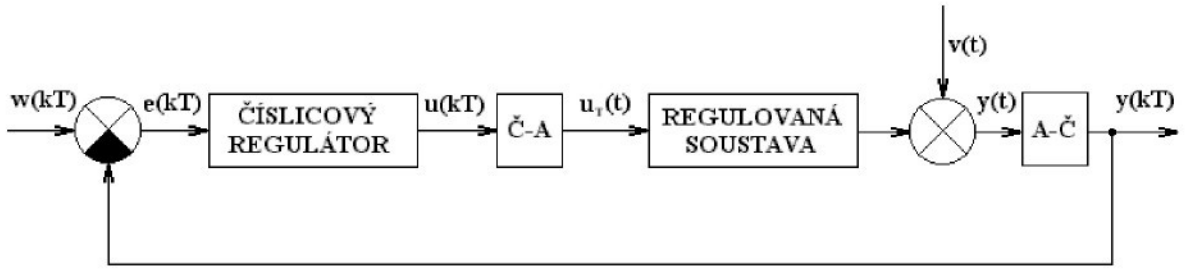
Regulovaná soustava je vždy spojitá. Při řešení syntézy diskrétních regulačních obvodů se obecně používají dva postupy:

a) předsunutí A-Č – převodníky se uvažují jako součást regulátoru, pak pracujeme s kvazispojitým regulátorem, viz obrázek 2.11. Tento postup je možný při malé vzorkovací periodě. Až 80% číslicových regulačních obvodů můžeme převést na analogové regulační obvody (analogový RO můžeme zastoupit číslicovým RO).



Obrázek 2.11 - Blokové schéma lineárního diskrétního obvodu s kvazianalogovým regulátorem

b) podsunutí A-Č – převodníky se přidají k regulované soustavě, kterou považujeme jako kvazidiskrétní, viz obrázek 2.12. Jestliže dáme A/Č za regulovanou soustavu, považujeme obvod za diskrétní. Tento postup se hlavně využívá při velkých vzorkovacích periodách. Používáme Z-transformaci, která je dosti složitá. Nedochází ke zkreslení informace.



Obrázek 2.12 - Blokové schéma lineárního diskrétního obvodu s diskrétní soustavou

Shrnutí pojmů 2.1.

Jsou zde definovány základní pojmy jako systém, prvek, vazba, matematický model, struktura systému. Dále jsou zde uvedeny základní činnosti se systémy a dělení signálů a systémů dle vybraných hledisek. Následuje vysvětlení pojmů řízení, ovládání a regulace, je zde popsána struktura regulačních obvodů, základní prvky a veličiny. Dále pak je vysvětlen rozdíl mezi spojitým a diskrétním regulačním obvodem a jsou popsány přístupy k syntéze pro diskrétní regulační obvody.

Otázky 2.1.

6. Co je to prvek?
7. Co je to vazba?
8. Co je to nestacionární systém?
9. Co je to dynamický systém?
10. Z čeho se skládá regulační obvod?
11. Co je to regulovaná veličina?
12. Co je to akční veličina?

2.2. Laplaceova transformace

Čas ke studiu: 1,5 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat Laplaceovu transformaci.
- popsat rozklad na parciální zlomky.
- určit obraz a originál funkcí.
- vyřešit lineární diferenciální rovnice pomocí Laplaceovy transformace.

Výklad

Laplaceova transformace (L transformace) je účinný nástroj pro popis, analýzu a syntézu lineárních spojitých dynamických systémů. Usnadňuje řešení lineárních diferenciálních rovnic s konstantními koeficienty, které se převádí na algebraické rovnice. Patří mezi integrální transformace, kde oblast originálů je oblast reálných proměnných a oblast obrazů je oblast komplexní proměnné. Definiční vztahy Laplaceovy transformace:

$$F(s) = \mathrm{L}\{f(t)\} = \int_0^{\infty} f(t)e^{-st}dt\,, \tag{2.1}$$

$$f(t) = \mathrm{L}^{-1}\{F(s)\} = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} F(s)e^{st}ds\,, \tag{2.2}$$

kde je: L, L^{-1} – operátor přímé/zpětné Laplaceovy transformace; t – reálná proměnná (čas); s – komplexní proměnná; $f(t)$ – funkce reálné proměnné, originál L transformace; $F(s)$ – funkce komplexní proměnné, obraz L transformace; j – imaginární jednotka, c – reálná konstanta zvolená tak, aby v polorovině Re $s > c$ funkce $F(s)$ neměla žádné singulární body.

Spojitá funkce $f(t)$ musí splňovat tři podmínky a to:

- musí být nulová pro záporný čas,
- musí být exponencionálního řádu,
- musí být spojitá, resp. po částech spojitá.

Pomocí definičních vztahů L transformace byla vytvořena tabulka základních párů obrazů a originálů vybraných funkcí, podrobnější slovníky lze najít např. [DORF&BISHOP 1998:], [VÍTEČKOVÁ 2005.].

Tabulka 2.1 Slovník L-transformace

	Originál $f(t)$	**Obraz $F(s)$**
	Základní vlastnosti	
1.	$af(t)$	$aF(s)$
2.	$f'(t) = \frac{\mathrm{d}f(t)}{\mathrm{d}t}$	$sF(s) - f(0)$
3.	$f^n(t) = \frac{\mathrm{d}^n f(t)}{\mathrm{d}t^n}$	$s^nF(s) - \sum_{i=1}^{n} s^{n-i}\frac{\mathrm{d}^{i-1}f(0)}{\mathrm{d}t^{i-1}}$
4.	$\int_0^t f(\tau)\mathrm{d}\tau$	$\frac{1}{s}F(s)$
	Základní slovník	
5.	$\delta(t)$	1
6.	$\eta(t)$	$\frac{1}{s}$
7.	t	$\frac{1}{s^2}$

	Originál $f(t)$	Obraz $F(s)$
8.	t^n	$\frac{n!}{s^{n+1}}$
9.	$e^{\pm at}$	$\frac{1}{s \mp a}$
10.	$te^{\pm at}$	$\frac{1}{(s \mp a)^2}$
11.	$t^2 e^{\pm at}$	$\frac{1}{(s \mp a)^3}$
12.	$1 - e^{-at}$	$\frac{1}{s(s+a)}$
13.	$t - \frac{1 - e^{-at}}{a}$	$\frac{1}{s^2(s+a)}$
14.	$1 - e^{-at} - ate^{-at}$	$\frac{1}{s(s+a)^2}$
15.	$(1 - at)e^{-at}$	$\frac{s}{(s+a)^2}$
16.	$\frac{e^{-at} - e^{-bt}}{b-a}$	$\frac{1}{(s+a)(s+b)}$
17.	$1 - \frac{b}{b-a}e^{-at} + \frac{a}{b-a}e^{-bt}$	$\frac{ab}{s(s+a)(s+b)}$
18.	$\frac{e^{-at}}{(b-a)(c-a)} + \frac{e^{-bt}}{(c-a)(a-b)} + \frac{e}{(a-c}$	$\frac{1}{(s+a)(s+b)(s+c)}$
19.	$\sin \omega t$	$\frac{\omega}{s^2 + \omega^2}$
20.	$\cos \omega t$	$\frac{s}{s^2 + \omega^2}$
21.	$a^{-at} \sin \omega t$	$\frac{\omega}{(s+a)^2 + \omega^2}$
22.	$e^{-at} \cos \omega t$	$\frac{s+a}{(s+a)^2 + \omega^2}$

Příklad 2.1. *Laplaceova transformace.*

Určete originál k obrazu:

$$a.)\ \frac{2}{s+3}$$

$$b.)\ \frac{s}{s^2+4s+13}.$$

Řešení:

a.) V základním slovníku musíme najít příslušný obraz funkce (určující je jmenovatel, ten se nedá změnit ani upravit). V našem případě obraz odpovídá výrazu na řádku 9, takže můžeme určit originál (násobení konstantou se zachovává, viz řádek 1):

$$\frac{2}{s+3} \doteq 2e^{-3t}$$

b.) V tomto případě máme ve jmenovateli polynom druhého stupně, takže musíme určit jeho kořeny a podle toho určíme řádek tabulky. Polynom druhého stupně ve jmenovateli mají výrazy 10, 16, 21 a 22. V našem případě jsou kořeny komplexně sdružené, tomu odpovídají vztahy 21 a 22, porovnáním s čitatelem obrazu vyplývá, že budeme pracovat s řádkem 22. Nejdříve určíme hodnoty a a ω

$$\frac{s}{s^2+4s+13} = \frac{s}{(s+2)^2+3^2}.$$

Teď musíme upravit čitatel, tzn. přičteme a odečteme hodnotu a, a rozepíšeme původní výraz na dva zlomky

$$\frac{s+2-2}{(s+2)^2+3^2} = \frac{s+2}{(s+2)^2+3^2} - \frac{2}{(s+2)^2+3^2}.$$

První zlomek odpovídá řádku 22 a druhý zlomek převedeme dle řádku 21, ten má ale v čitateli hodnotu ω, a proto výraz rozšíříme jedničkou a můžeme určit originál k danému obrazu

$$\frac{s+2}{(s+2)^2+3^2} - \frac{2}{(s+2)^2+3^2}\frac{3}{3} \doteq e^{-2t}\cos 3t - \frac{2}{3}e^{-2t}\sin 3t.$$

V předchozím příkladě jsme si ukázali, jak lze určit k danému obrazu $F(s)$ jeho originál $f(t)$ pouhým převodem na základě tabulky 2.1. Pokud však obraz má složitější vztah musíme ho zjednodušit a to pomocí rozkladu na parciální zlomky. Předpokládejme, že obraz $F(s)$ je ryze racionálně lomená funkce (tzn. stupeň polynomu v čitateli je nižší než stupeň polynomu ve jmenovateli). Nejdříve musíme určit kořeny polynomu ve jmenovateli a ty

mohou být reálné jednonásobné, reálné vícenásobné, komplexně sdružené a komplexně sdružené vícenásobné. Rozklad na parciální zlomky můžeme popsat vztahy:

- Reálné kořeny jednonásobné

$$F(s)=\frac{M(s)}{(s-s_1)(s-s_2)(s-s_3)}=\frac{A}{s-s_1}+\frac{B}{s-s_2}+\frac{C}{s-s_3}$$

- Reálné kořeny vícenásobné

$$F(s)=\frac{M(s)}{(s-s_1)^4}=\frac{A}{s-s_1}+\frac{B}{(s-s_1)^2}+\frac{C}{(s-s_1)^3}+\frac{D}{(s-s_1)^4}$$

- Komplexně sdružené kořeny

$$F(s)=\frac{M(s)}{s^2+as+b}=\frac{As+B}{s^2+as+b}$$

- Násobné komplexně sdružené kořeny

$$F(s)=\frac{M(s)}{(s^2+cs+d)^4}=\frac{C_1s+D_1}{s^2+cs+d}+\frac{C_2s+D_2}{(s^2+cs+d)^2}+\frac{C_3s+D_3}{(s^2+cs+d)^3}+\frac{C_4s+D_4}{(s^2+cs+d)^4}$$

Odstraněním zlomků dostaneme rovnici pro dopočet konstant rozkladu A, B, C, D a pak můžeme použít dosazovací metodu nebo metodu neurčitých koeficientů, viz. příklad 2.2.

Příklad 2.2. *Laplaceova transformace.*

Určete originál k obrazu: $F(s)=\dfrac{2s+4}{(s+2)(s^2+4)}$

Řešení:

$F(s)$ je ryze racionální lomená funkce, ale tento výraz ve slovníku nenajdeme. Musíme provést rozklad na parciální zlomky a pak teprve provést převod.

$$\frac{2s+4}{(s+2)(s^2+4)}=\frac{A}{s+2}+\frac{Bs+C}{s^2+4}=\frac{2}{s^2+4}\;\hat{=}\;\sin 2t$$

$$A(s^2+4)+(Bs+C)(s+2)=2s+4$$

$$s=-2\Rightarrow 6A=0\Rightarrow A=0$$

$$s=0\Rightarrow 2C=4\Rightarrow C=2$$

$$s=1\Rightarrow 3B+3C=6\Rightarrow B=0$$

Příklad 2.3. *Řešení diferenciálních rovnic.*

Pomocí L transformace vyřešte diferenciální rovnici
$y''(t)+2y'(t)+2y(t)=2, y(0)=0, y'(0)=0.$

Řešení:

Nejdříve určíme obraz diferenciální rovnice, tzn. každý výraz převedeme zvlášť, součet resp. rozdíl se zachovává stejně jako násobení konstantou. Obraz diferenciální rovnice je

$$s^2Y(s)+2sY(s)+2Y(s)=\frac{2}{s}.$$

V oblasti obrazu se diferenciální rovnice změnila na algebraickou, takže jednoduše vyjádříme obraz řešení diferenciální rovnice $Y(s)$

$$Y(s)=\frac{2}{s\left(s^2+2s+2\right)}.$$

Teď musíme pomocí zpětné L transformace najít originál řešení diferenciální rovnice.

$$y(t)=\mathrm{L}^{-1}\left\{\frac{2}{s\left(s^2+2s+2\right)}\right\}.$$

Tento výraz v našem slovníku L transformace nenajdeme, takže ho musíme zjednodušit a to rozkladem na parciální zlomky, v našem případě je rozklad na parciální zlomky následující

$$\frac{2}{s\left(s^2+2s+2\right)}=\frac{\mathrm{A}}{s}+\frac{\mathrm{B}s+\mathrm{C}}{s^2+2s+2}.$$

Konstanty parciálních zlomků dopočítáme pomocí metody neurčitých koeficientů nebo dosazovací metodou na základě rovnice

$$2=\mathrm{A}\left(s^2+2s+2\right)+\mathrm{B}s^2+\mathrm{C}s.$$

Např. dosazovací metoda:

$$\begin{aligned} &s=0: \quad 2=2A && \Rightarrow \quad A=1\\ &s=1: 2=A+B+C\\ &s=-1: 2=A+B-C && \Rightarrow B=2-3A=-1\\ &C=A+B-2=-2 \end{aligned}$$

Nyní můžeme určit řešení diferenciální rovnice:

$$y(t)=\mathrm{L}^{-1}\left\{\frac{1}{s}-\frac{s+2}{(s+1)^2+1}\right\}=\mathrm{L}^{-1}\left\{\frac{1}{s}-\frac{s+1}{(s+1)^2+1}-\frac{1}{(s+1)^2+1}\right\}=\eta(t)-e^{-t}\cos t-e^{-t}\sin t.$$

Správnost řešení můžeme ověřit výpočtem počátečních podmínek:

$$y(0) = 1 - 1 = 0$$

$$y'(t) = e^{-t}(\cos t + \sin t) - e^{-t}(-\sin t + \cos t) \quad \Rightarrow y'(0) = 0$$

Příklad 2.4. *Řešení diferenciálních rovnic.*

Vyřešte zadanou diferenciální rovnici $y''(t) + 7y'(t) + 10y(t) = \eta(t)$ s nulovými počátečními podmínkami.

Řešení:

Derivace v diferenciální rovnici odpovídá příslušné mocnině komplexní proměnné *s*. *η(t)* se převede podle Laplaceova slovníku na *1/s*.

$$s^2Y(s) + 7sY(s) + 10Y(s) = \frac{1}{s}$$

$$Y(s)\left[s^2 + 7s + 10\right] = \frac{1}{s}$$

Vyjádříme si $Y(s)$ a postupujeme obdobně, jako v minulém příkladu. Rozložíme zlomek na parciální zlomky.

$$Y(s) = \frac{1}{s(s^2 + 7s + 10)} = \frac{A}{s} + \frac{B}{s+2} + \frac{C}{s+5}$$

$$1 = A(s+2)(s+5) + Bs(s+5) + Cs(s+2)$$

$$s = 0 \Rightarrow A = \frac{1}{10}$$

$$s = -2 \Rightarrow B = -\frac{1}{6}$$

$$s = -5 \Rightarrow C = \frac{1}{15}$$

$$Y(s) = \frac{1}{10s} - \frac{1}{6(s+2)} + \frac{1}{15(s+5)}$$

Tuto funkci můžeme zpátky převést pomocí L-transformace.

$$y(t) = \frac{1}{10}\eta(t) - \frac{1}{6}e^{-2t} + \frac{1}{15}e^{-5t}$$

Správnost řešení můžeme ověřit výpočtem počátečních podmínek:

$$y(0) = \frac{1}{10} - \frac{1}{6} + \frac{1}{15} = \frac{3 - 5 + 2}{30} = 0$$

$$y'(t) = \frac{1}{3}e^{-2t} - \frac{1}{3}e^{-5t} \quad \Rightarrow y'(0) = 0$$

Příklad 2.5. *Řešení diferenciálních rovnic.*

Vyřešte zadanou diferenciální rovnici $y''(t)+7y'(t)+10y(t)=\eta(t)$
$y(0)=2, y'(0)=1$

Řešení:

Tento příklad se řeší podobně jako příklad 2., ale musíme v prvním kroku přidat počáteční podmínky. Použijeme k tomu vlastnost L-transformace derivace n-tého řádu
$L\left\{f^{(n)}(t)\right\}=s^nF(s)-s^{n-1}f(0)-s^{n-2}f'(0)-\ldots-f^{n-1}(0)$

$$s^2Y(s)-sy(0)-y'(0)+7\left[Y(s)-y(0)\right]+10Y(s)=\frac{1}{s}$$

$$Y(s)(s^2+7s+10)-2s-15=\frac{1}{s}$$

$$Y(s)(s^2+7s+10)=\frac{1}{s}+2s+15$$

$$Y(s)(s^2+7s+10)=\frac{2s^2+15s+1}{s}$$

$$Y(s)=\frac{2s^2+15s+1}{s(s+2)(s+5)}=\frac{A}{s}+\frac{B}{s+2}+\frac{C}{s+5}$$

$$2s^2+15s+1=A(s+2)(s+5)+Bs(s+5)+Cs(s+2)$$

$$s=0\Rightarrow A=\frac{1}{10}$$

$$s=-2\Rightarrow B=\frac{7}{2}$$

$$s=-5\Rightarrow C=-\frac{24}{15}$$

$$Y(s)=\frac{1}{10s}+\frac{7}{2(s+2)}-\frac{24}{15(s+5)}$$

$$y(t)=\frac{1}{10}\eta(t)+\frac{7}{2}e^{-2t}-\frac{24}{15}e^{-5t}$$

Správnost řešení můžeme ověřit výpočtem počátečních podmínek:

$$y(0)=\frac{1}{10}+\frac{7}{2}-\frac{24}{15}=\frac{3+105-48}{30}=2$$

$$y'(t)=-7e^{-2t}+8e^{-5t}\quad\Rightarrow y'(0)=1$$

Shrnutí pojmů 2.2.

Kapitola je věnována Laplaceově transformaci. Jsou zde popsány definiční vztahy a základní vlastnosti L transformace. Dále kapitola obsahuje zkrácený slovník L transformace, který je nutný pro práci s touto transformací. Závěr kapitoly tvoří řešené příklady, které

popisují práci s L transformací a hlavně ukazují postup řešení lineárních diferenciálních rovnic.

Otázky 2.2.

1. Jaká funkce je obraz funkce u Laplaceovy transformace?
2. Jak zní vlastnost Laplaceovy transformace pro n-tou derivaci?
3. Jaké typy rovnic se dají pomocí Laplaceovy transformace výhodně řešit?

Úlohy k řešení 2.1.

13. Určete originál funkce

a) $F(s)=\dfrac{2s+1}{(s+2)(s^2+10s+21)}$ b) $F(s)=\dfrac{1}{s^4+2s^3+s^2}$

c) $F(s)=\dfrac{s^3+s^2+9s-2}{s^3+s^2+2s}$ d) $F(s)=\dfrac{5s+3}{(s-1)(s^2+2s+3)}$

14. Vyřešte zadanou diferenciální rovnici $y''(t)+2y'(t)+3y(t)=\eta(t)$ s nulovými počátečními podmínkami.

15. Vyřešte zadanou diferenciální rovnici $y''(t)+9y(t)=\eta(t)$ s nulovými počátečními podmínkami.

16. Vyřešte zadanou diferenciální rovnici $y'''(t)+10y'(t)+8y(t)=2\eta(t)\quad y(0)=1, y'(0)=3$.

2.3. Matematický popis statických a dynamických vlastností spojitých systémů

Čas ke studiu: 2 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat způsoby popisu statických a dynamických vlastností spojitých systémů - oblast použití, obecný tvar, podmínky fyzikální realizovatelnosti, ustálený stav, charakteristiky, převodní vztahy.
- popsat měření přechodové, impulsní a kmitočtové charakteristiky.
- vypočítat matematické modely z diferenciální rovnice.
- nakreslit přechodovou, impulsní a amplitudo-fázovou kmitočtovou charakteristiku.

Výklad

Analýza dynamických systémů je nedílnou součástí takových oborů jako je identifikace, simulace a syntéza systémů. Analýzou zjišťujeme a zkoumáme chování systému při znalosti jeho struktury a vlastností vazeb mezi jednotlivými prvky. Analýzu můžeme chápat také jako rozbor zkoumaného systému s cílem specifikace jeho podstatných vlastností jak statických, tak i dynamických.

Statická charakteristika

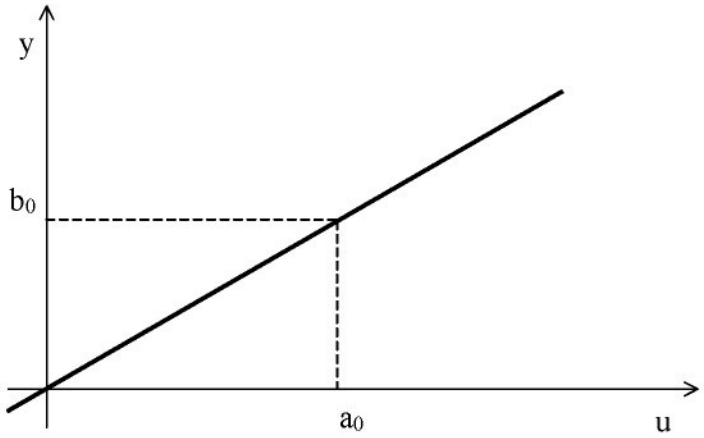
Statické vlastnosti lineárních spojitých dynamických systémů popisují systém v ustáleném stavu, což je stav, kdy se vstupy a výstupy v čase nemění, tzn. $t \to \infty$. Vyjdeme-li z popisu systému definovaného diferenciální rovnicí ve tvaru

$$a_n y^{(n)}(t) + a_{n-1} y^{(n-1)}(t) + \cdots + a_0 y(t) = b_m u^{(m)}(t) + \cdots + b_0 u(t) \qquad (2.3)$$

kde je a_i, b_j – konstantní koeficienty, $u(t)$ – vstupní veličina, $y(t)$ – výstupní veličina.
Z definice statické charakteristiky (závislost výstupní veličiny na vstupní veličině v ustáleném stavu) dostaneme statickou funkci ve tvaru

$$y(\infty) = \frac{b_0}{a_0} u(\infty), \qquad (2.4)$$

kde b_0/a_0 je směrnice přímky procházející počátkem. [Švarc, I 1992, Hanuš,B.,Balda, M. & kolektiv 1989]



Obrázek 2.13 - Příklad statické charakteristiky

Nyní uvedeme způsoby jak popsat dynamické vlastnosti spojitých systémů, jak v oblasti časové, kmitočtové, tak i v oblasti komplexní proměnné.

❑ Diferenciální rovnice

Je to matematický popis v časové oblasti. Řád diferenciální rovnice n [nejvyšší derivace výstupní veličiny $y(t)$] udává řád systému. obecný tvar lineárně diferenciální rovnice s konstantními koeficienty je

$$a_n \frac{d^n y(t)}{dt^n} + \ldots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = b_m \frac{d^m u(t)}{dt^m} + \ldots + b_1 \frac{du(t)}{dt} + b_0 u(t) . \quad (2.5)$$

Podmínky fyzikální realizovatelnosti

n > m	**silná**
n = m	**slabá**
n < m	**nerealizovatelné**

Ustálený stav - výstupní i vstupní veličina jsou v čase konstantní

$$a_0 y(\infty) = b_0 u(\infty) ,$$

$$y(\infty) = \lim_{t \to \infty} y(t) = \frac{b_0}{a_0} u(\infty) .$$

❑ Obrazový přenos

Je to popis oblasti komplexně proměnné. Je definován jako podíl Laplaceova obrazu výstupní veličiny k Laplacoeovu obrazu vstupní veličiny při nulových počátečních podmínkách.

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_m s^m + \ldots + b_1 s + b_0}{a_n s^n + \ldots + a_1 s + a_0} . \quad (2.6)$$

Podmínky fyzikální realizovatelnosti:

n > m	**silná** (stupeň polynomu jmenovatele n je vetší než stupeň polynomu čitatele m)
n = m	**slabá**

n < m **nerealizovatelné**

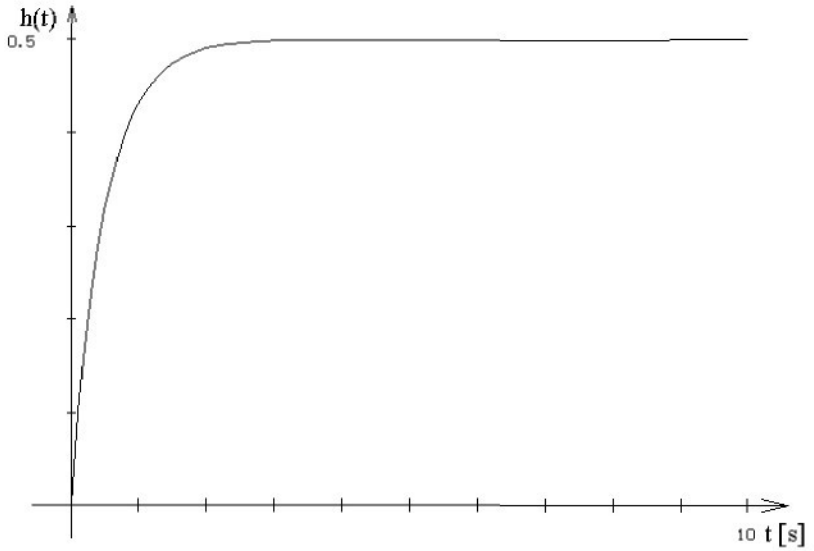
Ustálený stav:

$$y(\infty) = \lim_{s \to 0} G(s)u(\infty) \Rightarrow \frac{b_0}{a_0} u(\infty), \qquad a_0 \neq 0 .$$

❑ Přechodová funkce

Přechodová funkce systému je jeho odezva na vstup ve tvaru Heavisideova jednotkového skoku. Značí se $h(t)$ a její časová závislost se nazývá přechodová charakteristika. Přechodovou funkci můžeme určit analyticky z diferenciální rovnice nebo z obrazového přenosu systému $G(s)$.

$$h(t) = L^{-1}\{H(s)\} = L^{-1}\left\{\frac{1}{s} G(s)\right\} \tag{2.7}$$



Obrázek 2.14 - Příklad přechodové charakteristiky

Podmínky fyzikální realizovatelnosti:

$$h(t) = \begin{cases} 0 & t \leq 0 \\ \text{h(t)} & t > 0 \end{cases} \quad \text{slabá fyzikální realizovatelnost,} \tag{2.8}$$

$$h(t) = \begin{cases} 0 & t < 0 \\ \text{h(t)} & t \geq 0 \end{cases} \quad \text{silná fyzikální realizovatelnost.} \tag{2.9}$$

Příklad 2.6.

Určete přechodovou funkce a sestrojte přechodovou charakteristiku systému s přenosem $G(s) = \dfrac{3}{0{,}5s+1}$.

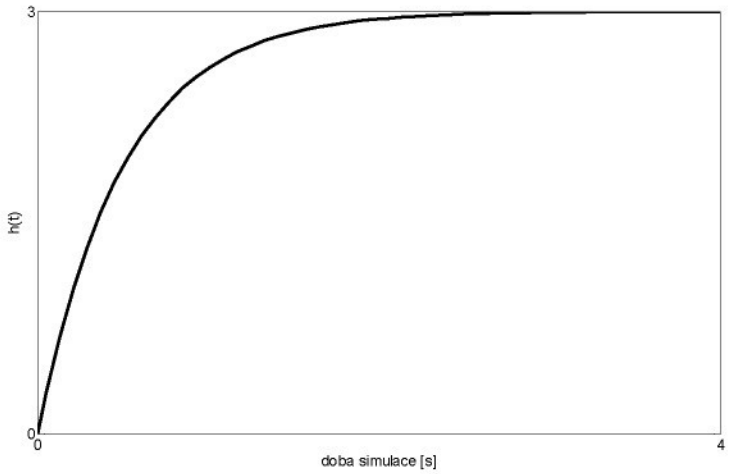
Řešení:

Přechodová funkce je dána (2.7), můžeme tedy psát

$$h(t) = \mathrm{L}^{-1}\left\{\frac{1}{s}\ \frac{3}{0{,}5s+1}\right\} = 2.3\mathrm{L}^{-1}\left\{\frac{1}{s}\ \frac{1}{s+2}\right\} = 3(1-e^{-2t}).$$

Pro vykreslení přechodové charakteristiky musíme spočítat počáteční a koncovou hodnotu přechodové funkce, tzn.

$$h(0) = 0,$$
$$h(\infty) = 3.$$



Obrázek 2.15 - Přechodová charakteristika k příkladu 2.6

Měření přechodové charakteristiky plyne z její definice, tedy na vstup soustavy přivedeme vstupní signál ve formě jednotkového skoku. Při této metodě určování přechodové charakteristiky musí být před zavedením změny vstupního signálu soustava v ustáleném stavu a vstupují-li do soustavy další signály je třeba je udržovat v průběhu měření na konstantní hodnotě. Účelné je měření několikrát opakovat (pro snížení vlivu náhodných poruch). Výslednou přechodovou charakteristiku lze potom stanovit podle následujícího vzorce:

$$f_i = \frac{\sum_{k=1}^{N} \mathrm{sign}(\Delta u_k) y_{ik}}{\sum_{k=1}^{N} |\Delta u_k|}$$

kde je N - počet opakovaných měření přechodové charakteristiky při obecně nestejně velkých skokových změnách vstupní veličiny soustavy, Δu_k – přírůstek řídicí veličiny při k-tém měření přechodové charakteristiky, f_i - pořadnice vyhodnocené přechodové charakteristiky v čase $t = i*\Delta t$, kde Δt je interval vzorkování, y_{ik} - hodnota odezvy vstupní veličiny soustavy v i-tém intervalu vzorkování při k-tém měření, i - pořadí vzorkovaných bodů přechodové charakteristiky $i = 0$ až m.

Impulsní funkce

Impulsní (někdy též váhová) funkce $g(t)$ je odezva systému na vstup ve tvaru Diracova impulsu. Diracův nebo též jednotkový impuls $\delta(t)$ se nedá fyzikálně realizovat. Při jeho přibližné realizaci se udává, že musí mít co největší amplitudu a co nejkratší dobu trvání. Matematicky je $\delta(t) = 0$ pro $t \neq 0$ přičemž $\delta(t)$ není definováno pro $t = 0$, ale platí

$$\int_{-\infty}^{\infty} \delta(t)dt = 1. \qquad (2.10)$$

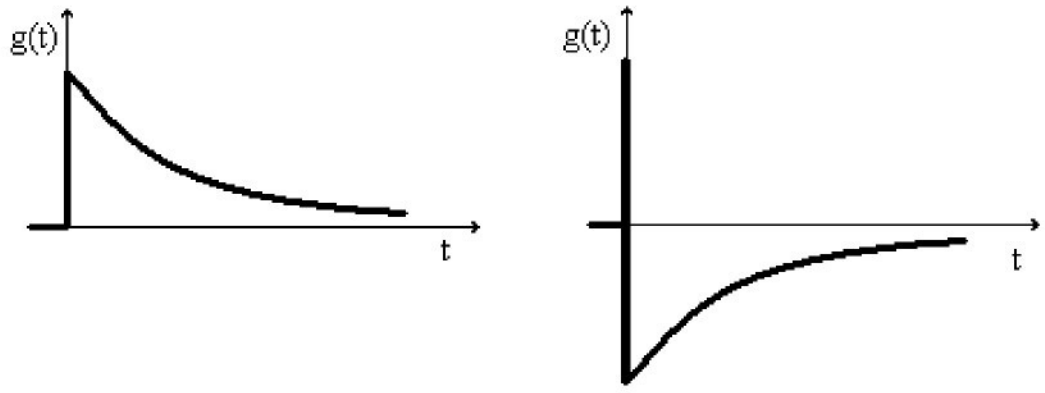
Časová závislost impulsní funkce se nazývá impulsní charakteristika. Impulsní funkce lze analyticky určit dle vztahu

$$g(t) = \mathrm{L}^{-1}\{G(s)\}. \qquad (2.11)$$

Podmínky fyzikální realizovatelnosti:

$$g(t) = \begin{cases} 0 & t < 0 \\ g(t) & t \geq 0 \end{cases} \quad \text{silná fyzikální realizovatelnost,} \qquad (2.12)$$

$$g(t) = \begin{cases} 0 & t < 0 \\ g(t) + \delta(t) & t \geq 0 \end{cases} \quad \text{slabá fyzikální realizovatelnost.} \qquad (2.13)$$



Obrázek 2.16 - Příklad impulsní charakteristiky: z leva silně fyzikálně realizovatelná, slabě fyzikálně realizovatelná

Ustálený stav:

$$y(\infty) = \left[\lim_{t \to \infty} \int_0^t g(\tau) d\tau \right] \cdot u(\infty) .$$

Vztah mezi přechodovou a impulsní funkcí

$$H(s) = \frac{G(s)}{s} \Rightarrow h(t) = \int_0^t g(\tau) d\tau , \tag{2.14}$$

$$G(s) = sH(s) \Rightarrow g(t) = \frac{dh(t)}{dt} . \tag{2.15}$$

Příklad 2.7.

Určete impulsní funkci a sestrojte přechodovou charakteristiku systému s přenosem $G(s) = \dfrac{3}{0{,}5s+1}$.

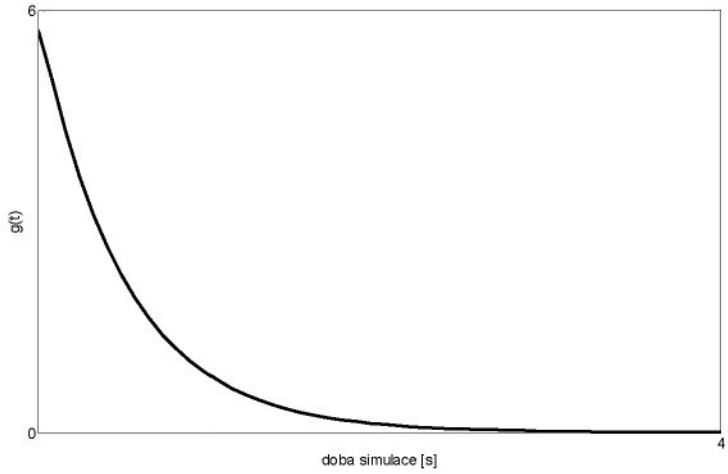
Řešení:

Impulsní funkce je dána (2.11) , můžeme proto psát

$$g(t) = \mathrm{L}^{-1}\left\{ \frac{3}{0{,}5s+1} \right\} = \mathrm{L}^{-1}\left\{ \frac{6}{s+2} \right\} = 6e^{-2t} .$$

Pro vykreslení impulsní charakteristiky musíme spočítat počáteční a koncovou hodnotu impulsní funkce, tzn.

$$g(0) = 6,$$
$$g(\infty) = 0.$$



Obrázek 2.17 - Přechodová charakteristika k příkladu 2.7

Měření impulsní charakteristiky plyne taktéž z její definice, tedy na vstup soustavy přivedeme vstupní signál ve formě jednotkového impulsu Požadavky při měření jsou stejné jako při zjišťování odezvy na skok u přechodové charakteristiky, metoda určení odezvy na impuls, i když je experimentálně náročnější, má přednost v tom, že zásah do provozu je při stejné výšce vstupního signálu podstatně menší.

❑ Kmitočtový přenos

Kmitočtový přenos popisuje vlastnosti systému v kmitočtové oblasti a je popsán vztahem

$$G(j\omega) = G(s)\big|_{s=j\omega} = \frac{b_m(j\omega)^m + \cdots b_1(j\omega) + b_0}{a_n(j\omega)^n + \cdots a_1(j\omega) + a_0} = \frac{M(j\omega)}{N(j\omega)}, \tag{2.16}$$

kde ω je reálný kmitočet [s^{-1}], tzn. v intervalu $(0,\infty)$.
Podmínky fyzikální realizovatelnosti:

$m = n$	slabá fyzikální realizovatelnost,
$m < n$	silná fyzikální realizovatelnost.

Kmitočtový přenos $G(j\omega)$ je funkce komplexní proměnné, takže jej můžeme vyjádřit ve všech základních tvarech:

- algebraický tvar:
  $$G(j\omega) = Re\{G(j\omega)\} + j\,Im\{G(j\omega)\} = P(\omega) + jQ(\omega), \tag{2.17}$$
  kde je $P(\omega)$ – reálná část kmitočtového přenosu, $Q(\omega)$ – imaginární část kmitočtového přenosu, j – imaginární jednotka, $j = \sqrt{-1}$,

- exponenciální tvar:
  $$G(j\omega) = A(\omega)e^{j\varphi(\omega)}, \tag{2.18}$$
  $A(\omega)$ – modul kmitočtového přenosu, $A(\omega) = \sqrt{P^2(\omega) + Q^2(\omega)}$, $\varphi(\omega)$ – fáze kmitočtového přenosu, $\varphi(\omega) = \operatorname{arctg}\frac{Q(\omega)}{P(\omega)}$.

- goniometrický tvar:
  $$G(j\omega) = A(\omega)\left[\cos\varphi(\omega) + j\sin\varphi(\omega)\right]. \tag{2.19}$$

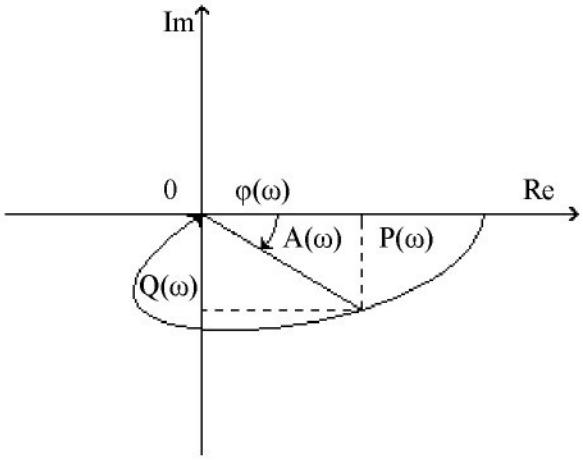
kde je: $P(\omega)$- reálná část, $Q(\omega)$ - imaginární část.

Ustálený stav:

$$y(\infty) = \lim_{\omega \to 0} G(j\omega)u(\infty)$$

Pro vyšetření vlastností systému v kmitočtové oblasti můžeme použít tři charakteristiky a to:

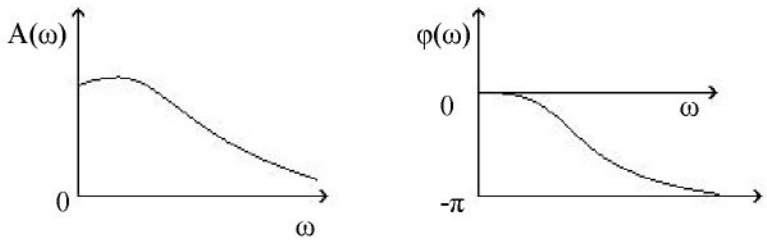
1) AFKCH (amplitudo-fázová kmitočtová charakteristika)



Obrázek 2.18 - Charakteristika AFKCH.

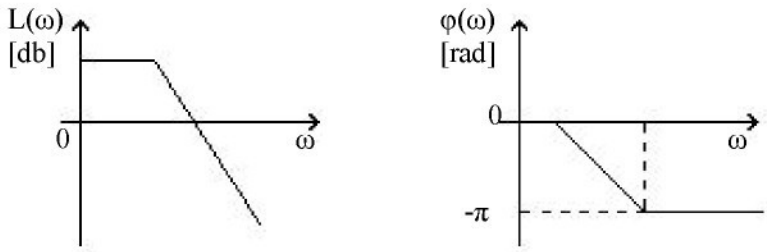
Silná podmínka – systém je silně fyzikálně realizovatelný, pokud charakteristika končí v nule. Jinak je systém slabě fyzikálně realizovatelný.

2) AKCH + FKCH (amplitudová a fázová kmitočtová charakteristika)



Obrázek 2.19 - Charakteristika AKCH+FKCH

3) LKCH (logaritmické kmitočtové charakteristiky)



Obrázek 2.20 - Charakteristika LKCH.

Měření kmitočtové charakteristiky. Měřením se určuje odezva soustavy na harmonicky proměnlivý vstupní signál konstantní amplitudy – budící kmity. Jenž lze popsat harmonickou funkcí

$$\Delta u(t) = \Delta u_A \cos \omega t$$

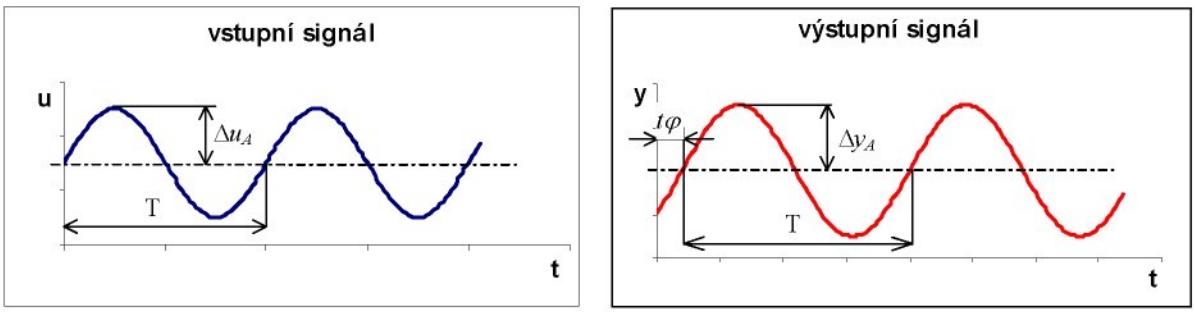
Měření se opakuje s různými kruhovými frekvencemi u vstupního signálu. Je nutno je volit tak, abychom pokryli celé kmitočtové pásmo harmonických kmitů, které zkoumaná soustava přenáší. U lineárních systémů lze výstupní signál soustavy – vynucené kmity, po odeznění přechodového děje, který jej zkresluje taktéž popsat harmonickou funkcí

$$\Delta y(t) = \Delta y_A \cos(\omega t + \varphi)$$

Výstupní signál – vynucené kmity má tedy stejnou frekvencí jako vstupní signál – budící kmity, avšak s různou amplitudou a oproti vstupnímu signálu je fázově posunut viz. obr. 2.21. Fázové posunutí lze popsat vztahem

$$\varphi = \omega \cdot t_\varphi,$$

kde t_φ představuje časové posunutí úhlu.



Obrázek 2.21 - Tvar a vzájemný vztah vstupního a výstupního signálu

Vztah mezi budícími a vynuceními kmity vyjadřují dva parametry kmitočtové charakteristiky

- amplitudová určující poměr amplitud v závislosti na frekvenci

$$\frac{\Delta y_A}{\Delta u_A} = A(\omega)$$

- fázová určující fázové posunutí v závislosti na frekvenci

$$\varphi = \varphi(\omega)$$

Použití periodických impulsů

Určování kmitočtové charakteristiky z harmonických signálů je časově i technicky velmi náročné. Tento nedostatek lze zmírnit nahrazením harmonického signálů periodickými pravoúhlými impulsy s frekvencí $f_0 = \frac{\omega}{2\pi}$ a amplitudou A_0. Rozložením tohoto signálu do Fourierovy řady obdržíme

$$u(t) = \frac{2A_0}{\pi}\left[\sin\omega_0 t + \frac{1}{3}\sin 3\omega_0 t + \frac{1}{5}\sin 5\omega_0 t + ...\right]$$

Odezvu na tento signál pomocí harmonické analýzy je možno rovněž rozložit do Fourierovy řady

$$y(t) = \frac{2A_0}{\pi}\left[A_1 \sin(\omega_0 t + \varphi_1) + \frac{A_3}{3}\sin(3\omega_0 t + \varphi_3) + \frac{A_5}{5}\sin(5\omega_0 t + \varphi_5) + ...\right]$$

Porovnáním odpovídajících harmonických složek vstupního a výstupního signálu obdržíme pro $\omega_0, 3\omega_0, 5\omega_0, ...$ hodnoty amplitud a fází kmitočtové charakteristiky. Vzhledem k tomu, že hodnoty amplitud rychle klesají, využívají se jen první tři složky a další body kmitočtové charakteristiky se získají dalším měřením s pravoúhlými impulsy jiné frekvence.

Příklad 2.8. *Matematické modely*

Určete matematické modely (přenos, přechodová charakteristika, impulsní charakteristika, frekvenční přenos) pro systém popsaný diferenciální rovnicí

$$T_1 y'(t) + y(t) = u(t)$$

Řešení:

Pro určení obrazového přenosu musíme najít L- obraz diferenciální rovnice **při nulových počátečních podmínkách**, který v tomto případě je popsán vztahem

$T_1 sY(s) + Y(s) = U(s)$.

Vyjádřením podílu L-obrazu výstupní proměnné k L-obrazu vstupní proměnné dostaneme vztah pro **obrazový přenos**

$$G(s) = \frac{1}{T_1 s + 1}.$$

Přechodová funkce se vypočte dle definičního vztahu:

$$h(t) = L^{-1}\left\{\frac{G(s)}{s}\right\} = L^{-1}\left\{\frac{1}{s(T_1 s + 1)}\right\} = 1 - e^{-\frac{1}{T_1}t}$$

$$h(0) = 0$$

$$h(\infty) = 1$$

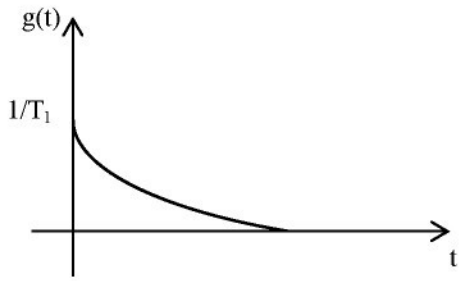
Obrázek 2.22 - Výsledný tvar přechodové charakteristiky.

Impulsní funkce se vypočte dle vztahu:

$$g(t) = L^{-1}\{G(s)\} = L^{-1}\left\{\frac{1}{T_1 s + 1}\right\} = \frac{1}{T_1} e^{-\frac{1}{T_1}t}$$

$$g(0) = \frac{1}{T_1}$$

$$g(\infty) = 0$$



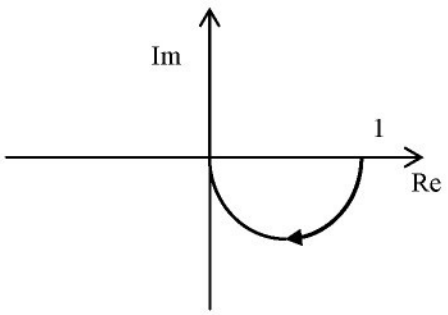
Obrázek 2.23 - Výsledný tvar impulsní charakteristiky

Kmitočtový přenos určíme tak, že do obrazového přenosu dosadíme za proměnnou *s* výraz $j\omega$ a určíme jeho reálnou a imaginární část:

$$G(j\omega) = \frac{1}{T_1 j\omega + 1} = \frac{1}{T_1(j\omega)+1}\frac{-T_1 j\omega + 1}{-T_1 j\omega + 1} = \frac{-T_1 j\omega + 1}{T_1^2\omega^2 + 1} = \frac{1}{T_1^2\omega^2 + 1} - j\frac{T_1\omega}{T_1^2\omega^2 + 1}$$

Amplitudo-fázová kmitočtová charakteristika:

ω	**0**	∞	**1**
P	1	0	+
Q	0	0	-



Obrázek 2.24 - Výsledný tvar amplitudo-fázové kmitočtové charakteristiky

---

Příklad 2.9. *Matematické modely*

Určete matematické modely (přenos, přechodová charakteristika, impulsní charakteristika, frekvenční přenos) pro systém popsaný diferenciální rovnicí.

$$T_1 y'(t) + y(t) = T_1 u'(t)$$

Řešení:

Přenos:

$$\mathrm{G(s)} = \frac{T_1 s}{T_1 s + 1}.$$

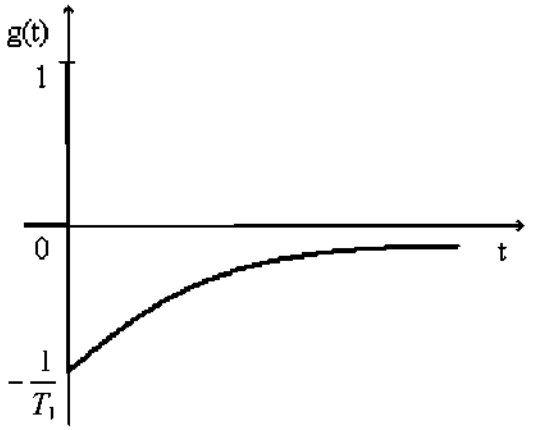
a) Impulsní funkce:

$$g(t) = L^{-1}\{G(s)\} = L^{-1}\left\{\frac{T_1 s}{T_1 s + 1}\right\} = \delta(t) - \frac{1}{T_1} e^{-\frac{1}{T_1}t},$$

$$g(0) = \delta(0) - \frac{1}{T_1},$$

$$g(\infty) = 0.$$

Podle podmínek je systém slabě fyzikálně realizovatelný.



Obrázek 2.25 - Výsledný tvar impulsní charakteristiky.

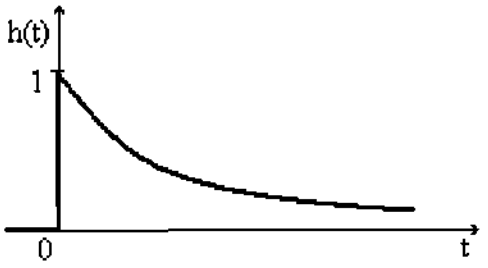
b) Přechodová funkce:

$$h(t) = L^{-1}\left\{\frac{G(s)}{s}\right\} = L^{-1}\left\{\frac{T_1}{T_1 s + 1}\right\} = e^{-\frac{1}{T_1}t},$$

$$h(0) = 1,$$

$$h(\infty) = 0.$$

Podle podmínek je systém slabě fyzikálně realizovatelný.



Obrázek 2.26 - Výsledný tvar přechodové charakteristiky.

c) Kmitočtový přenos:

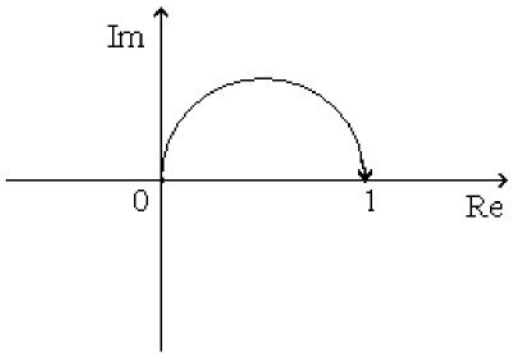
$$G(j\omega) = G(s)\big|_{s=j\omega} = \frac{T_1 j\omega}{T_1 j\omega + 1} \cdot \frac{1 - T_1 j\omega}{1 - T_1 j\omega} = \frac{T_1 j\omega + T_1^2 \omega^2}{T_1^2 \omega^2 + 1},$$

$$P(\omega) = \frac{T_1^2 \omega^2}{T_1^2 \omega^2 + 1} = \frac{T_1^2}{T_1^2 + \frac{1}{\omega^2}},$$

$$Q(\omega) = \frac{T_1 \omega}{T_1^2 \omega^2 + 1}.$$

ω	**0**	∞	**1**
P	0	1	+
Q	0	0	+

Podle podmínek je systém slabě fyzikálně realizovatelný.



Obrázek 2.27 - Výsledný tvar amplitudo-fázové kmitočtové charakteristiky.

Shrnutí pojmů 2.3.

Jsou zde uvedeny způsoby matematického popisu statických a dynamických vlastností spojitých systémů. Jedná se o statickou charakteristiku, obrazový přenos, přechodovou charakteristiku, impulsní charakteristiku, kmitočtovou charakteristiku. Dále jsou popsány převodní vztahy mezi jednotlivými popisy a způsoby měření charakteristik. Na závěr jsou uvedeny řešené příklady z dané problematiky.

Otázky 2.3.

1. Definujte statickou charakteristiku.
2. Kterými charakteristikami lze popsat dynamické vlastnosti systémů?
3. Jak je definován přenos?
4. Co je to Diracův impuls?
5. Čím se nahradí komplexní proměnná *s* ve vzorci pro kmitočtový přenos?

2.4. Logaritmické kmitočtové charakteristiky

Čas ke studiu: 1,5 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat logaritmické charakteristiky, logaritmickou amplitudu.
- popsat logaritmické charakteristiky základních šesti členů.
- vypočítat logaritmickou amplitudu ze zesílení.

Výklad

Kmitočtové charakteristiky jsou grafickým znázorněním kmitočtového přenosu. Pro vykreslování logaritmických kmitočtových charakteristik LKCH rovnice (2.18) se musí zlogaritmovat:

$$\ln G(j\omega) = \ln A(j\omega) + j\varphi(\omega) = \ln|G(j\omega)| + j\arg G(j\omega). \qquad (2.20)$$

Ve skutečnosti se však využívá dekadického logaritmu pro osu úhlového kmitočtu ω, tj. log ω, a na osu pořadnic amplitudové charakteristiky se vynáší absolutní hodnota kmitočtového přenosu v decibelech, což jsou jednotky definované jako dekadický logaritmus poměru obrazu výstupního signálu násobeného dvaceti, tj.

$$|G(j\omega)|_{\mathrm{dB}} = 20\log|G(j\omega)|. \qquad (2.21)$$

Tabulka 2.2 – Příklady vybraných hodnot amplitudy a logaritmické amplitudy

$A(\omega)$	0,001	0,01	0,1	1	10	100	1000	$\frac{1}{\sqrt{2}}$	$\sqrt{2}$
$L(\omega)$[dB]	-60	-40	-20	0	20	40	60	3	-3

Kmitočtové charakteristiky v logaritmických souřadnicích se znázorňují dvěmi charakteristikami: logaritmickou amplitudovou kmitočtovou charakteristikou (LAKCH) $L(\omega)$, logaritmickou fázovou kmitočtovou charakteristikou (LFKCH) $\varphi(\omega)$.
Význam zavedení logaritmických kmitočtových charakteristik (LKCH) spočívá ve zjednodušení vykreslení charakteristik složitých systémů a v jejich jednoduchém sestrojování. [Balátě, J. 2003]

Při vykreslování logaritmických kmitočtových charakteristik je důležité určit rozdělení na základní členy, pro které tyto charakteristiky umíme sestrojit, a dále určit zapojení těchto základních členů. V následujícím textu bude odvozeno sestrojení logaritmických charakteristik pro základní zapojení a základní členy.

LKCH základních zapojení

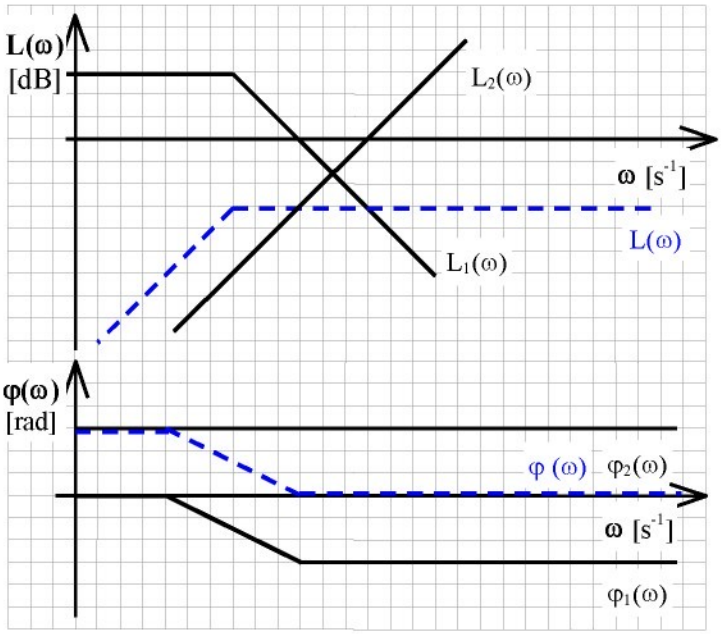
Sériové zapojení

Celkový přenos lineárních dynamických členů zapojených sériově (obrázek 2.75) je dán součinem přenosů jednotlivých členů, viz vztah (2.22). Amplituda a fáze výsledného přenosu se vypočte

$$A(\omega) = A_1(\omega)A_2(\omega) \quad \Rightarrow L(\omega) = L_1(\omega) + L_2(\omega),$$

$$\varphi(\omega) = \varphi_1(\omega) + \varphi_2(\omega).$$

Z výše uvedených vztahů vyplývá, že výsledná LKCH je dána součtem LKCH jednotlivých členů (viz obrázek 2.28).



Obrázek 2.28 - LKCH sériově zapojených členů

Paralelní zapojení

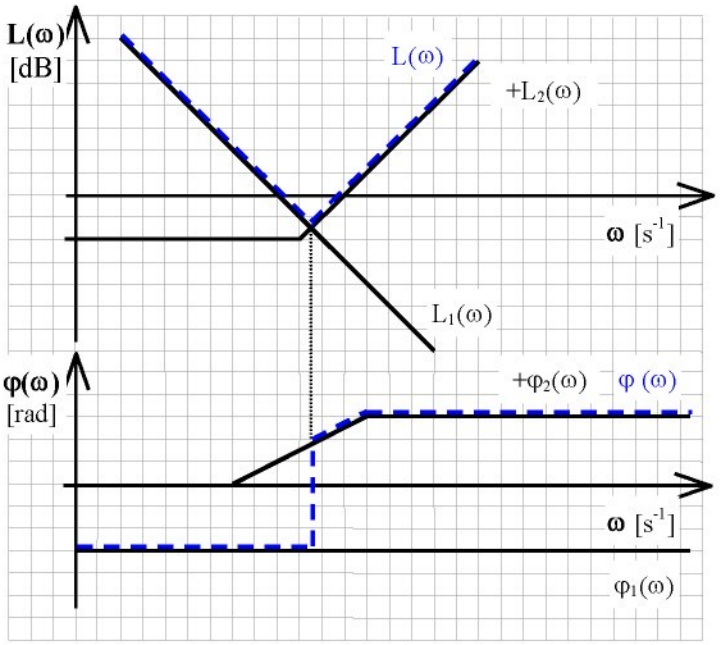
Celkový přenos lineárních dynamických členů zapojených sériově (obrázek 2.76) je dán součtem, resp. rozdílem přenosů jednotlivých členů, viz vztah (2.23). Také kmitočtový přenos je dán součtem, resp. rozdílem kmitočtových přenosů jednotlivých členů. Pokud v obvodu budou zařazeny jen členy s minimální fází (existuje jednoznačné přiřazení mezi LAKCH a LFKCH), stačí jen sestrojit LAKCH a LFKCH bude sestrojena na základě výsledné LAKCH.

Pro konstrukci LAKCH paralelně zapojených členů se uvažuje přibližný kmitočtový přenos

$$G(j\omega) = \begin{cases} G_1(\omega) \text{ pro} & A_1(\omega) >> A_2(\omega) \\ G_2(\omega) \text{ pro} & A_1(\omega) << A_2(\omega) \end{cases},$$

kde $A_1(\omega)$, $A_2(\omega)$ jsou moduly dílčích členů. Pak přibližná výsledná LAKCH je rovna vždy LAKCH členu s větší amplitudou.

$$L(\omega) = \begin{cases} L_1(\omega) \text{ pro} & L_1(\omega) >> L_2(\omega) \\ L_2(\omega) \text{ pro} & L_1(\omega) << L_2(\omega) \end{cases}.$$



Obrázek 2.29 - LKCH paralelního zapojení

Zpětnovazební zapojení

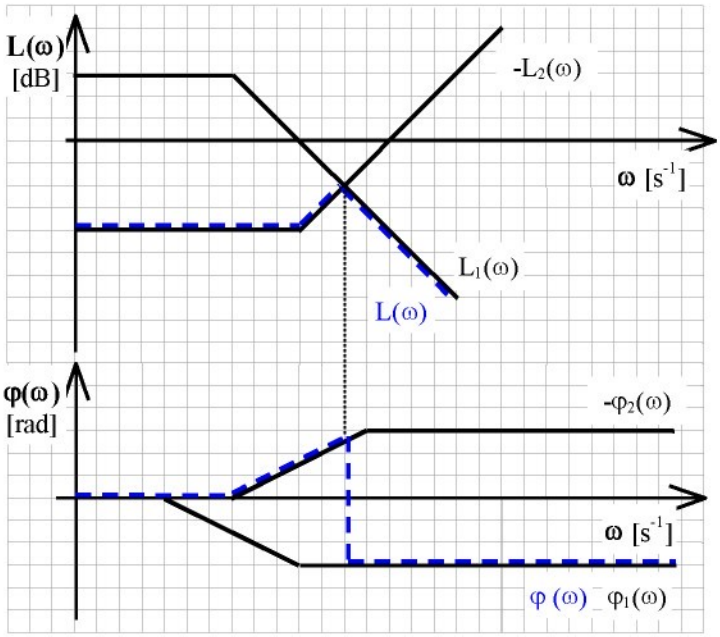
Rovněž u tohoto zapojení (obrázek 2.77) lze sestrojit přibližné LAKCH, za předpokladu, že v obvodu jsou jen dynamické členy s minimální fázi. Přibližný kmitočtový přenos je popsán rovnicí

$$G(j\omega) = \begin{cases} G_1(\omega) & \text{pro} \quad A_1(\omega) << A_2^{-1}(\omega) \\ G_2^{-1}(\omega) & \text{pro} \quad A_1(\omega) >> A_2^{-1}(\omega) \end{cases}.$$

Přibližná výsledná LAKCH zpětnovazebního obvodu je rovna

$$L(\omega) = \begin{cases} L_1(\omega) & \text{pro} \quad L_1(\omega) << -L_2(\omega) \\ -L_2(\omega) & \text{pro} \quad L_1(\omega) >> -L_2(\omega) \end{cases}.$$

Příklad sestrojení LKCH je na obrázku 2.30.



Obrázek 2.30 - LKCH zpětnovazebního zapojení

Inverzní člen

Pokud známe LKCH členu o přenosu $G_1(s)$, pak můžeme snadno sestrojit charakteristiku členu s inverzními dynamickými vlastnostmi, tj. s přenosem

$$G(s) = \frac{1}{G_1(s)}.$$

Kmitočtový přenos má tvar

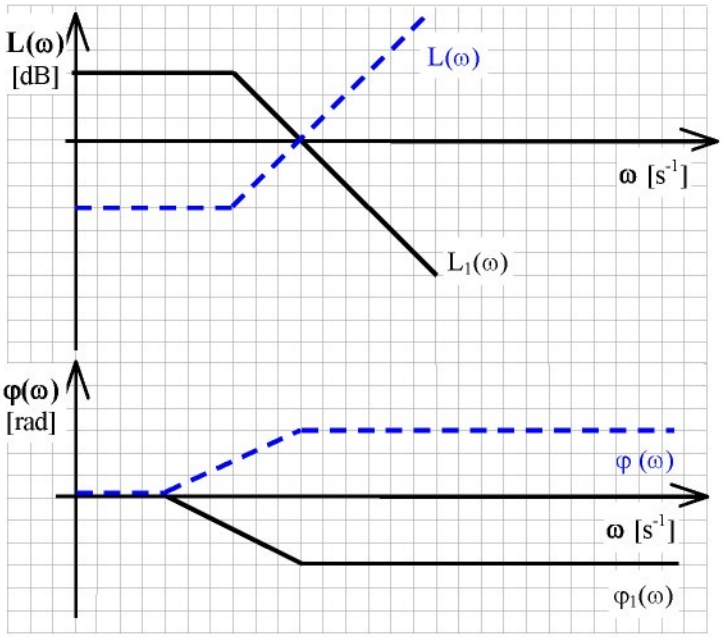
$$G(j\omega) = A(\omega)e^{j\varphi(\omega)} = \frac{1}{A_1(\omega)e^{j\varphi_1(\omega)}},$$

pak amplituda a fáze je dána vztahy

$$A(\omega) = \frac{1}{A_1(\omega)} \Rightarrow \qquad L(\omega) = -L_{-1}(\omega) = -20\log A_1(\omega)$$

$$\varphi(\omega) = -\varphi_1(\omega).$$

LKCH členů s inverzními vlastnostmi se liší jen znaménkem, tvar mají stejný, viz. obrázek 2.31.



Obrázek 2.31 - LKCH členů s inverzními vlastnostmi

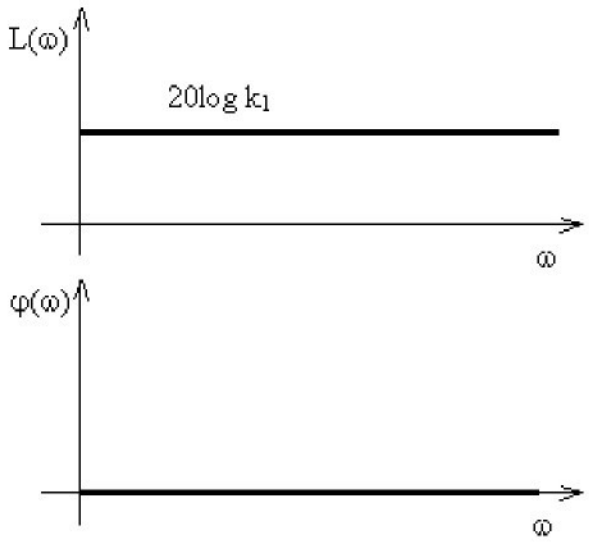
Základní členy přenosu

Ideální proporcionální člen, který lze popsat přenosem $G(s) = k_1$. Modul a fázi vypočteme podle vztahů

$$A(\omega) = k_1 \qquad \Rightarrow L(\omega) = 20 \log k_1 ,$$

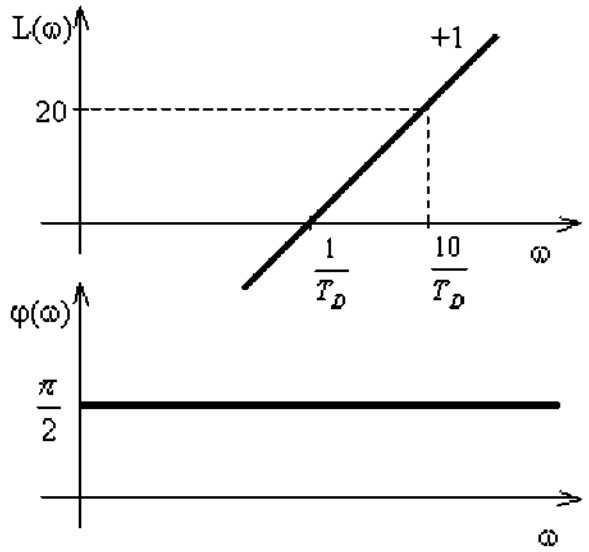
$$\varphi(\omega) = \text{arctg}\, 0 = 0.$$

Průběh LKCH je znázorněn na obrázku 2.32.



Obrázek 2.32 - LKCH pro ideální proporcionální člen

Ideální derivační člen, jehož vlastnosti popisuje přenos $G(s) = T_1 s$.



Obrázek 2.33 - LKCH kmitočtové charakteristiky pro ideální derivační člen

Vypočteme modul a fázi členu s kmitočtovým přenosem

$G(j\omega) = T_1 j\omega$. **Pak**

$P(\omega) = 0, Q(\omega) = T_1\omega$

proto

$A(\omega) = T_1\omega \quad \Rightarrow \quad L(\omega) = 20\log T_1\omega$,

$$\varphi(\omega) = \operatorname{arctg}\frac{T_1\omega}{0} = \operatorname{arctg}\infty = \frac{\pi}{2}.$$

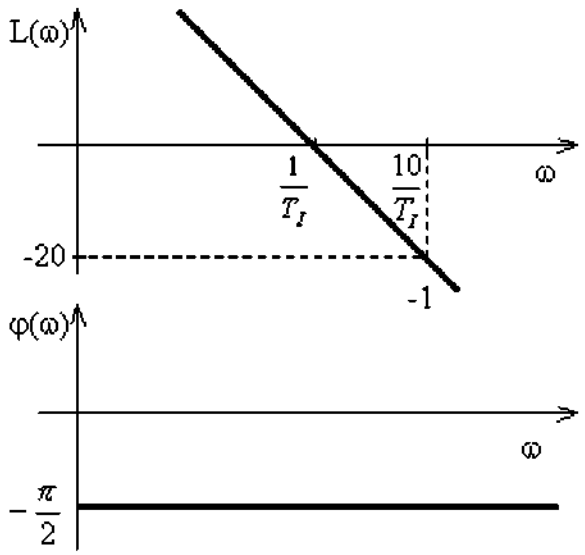
Je vidět, že fáze tohoto členu je konstantní, logaritmická amplituda je přímka protínající osu kmitočtu v $\omega_1 = \frac{1}{T_1}$ se sklonem 20 dB/dek [+1]. Tvar LKCH je na obrázku 2.33.

Ideální integrační člen, který je popsán přenosem $G(s) = \frac{1}{T_1 s}$. Pro výpočet a konstrukci jeho charakteristik použijeme vztahů (2.20), proto můžeme psát

$L(\omega) = -20\log T_1\omega$,

$$\varphi(\omega) = -\frac{\pi}{2}.$$

Výsledná LKCH tohoto členu je na obrázku 2.34.



Obrázek 2.34 - LKCH pro ideální integrační člen

Přenos typu $G(s) = T_1 s + 1$. Kmitočtový přenos, jeho reálná a imaginární část jsou popsány vztahy

$$G(j\omega) = 1 + T_1 j\omega \qquad \Rightarrow \qquad P(\omega) = 1, Q(\omega) = T_1\omega.$$

Proto modul a fáze jsou

$$L(\omega) = 20\log\sqrt{T_1^2\omega^2 + 1},$$

$$\varphi(\omega) = \text{arctg}\frac{T_1^2\omega^2}{1}$$

Pro LKCH platí následující vztahy:

$\omega << \frac{1}{T_1}$: $L(\omega) = 20\log 1 = 0$

$\varphi(\omega) = \text{arctg}0 = 0$

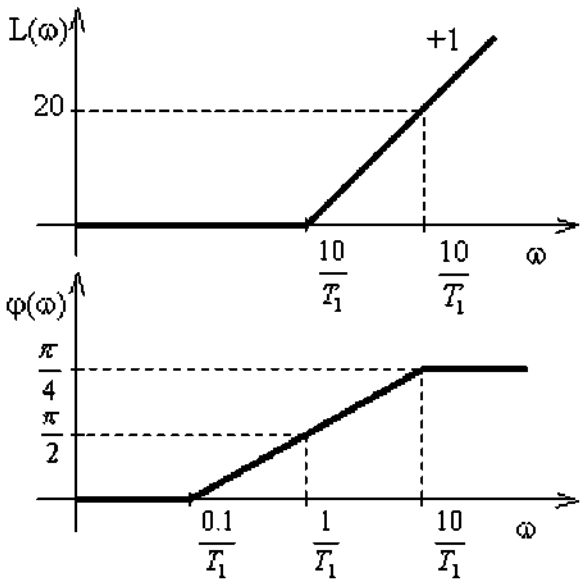
$\omega = \frac{1}{T_1}$: $L(\omega) = 20\log\sqrt{2} = 3 \approx 0$

$\varphi(\omega) = \text{arctg}1 = \frac{\pi}{4}$

$\omega >> \frac{1}{T_1}$: $L(\omega) = 20\log T_1\omega$

$\varphi(\omega) = \text{arctg}T_1\omega = \frac{\pi}{2}$

Pro jednoduchost se konstruuje LKCH tak, že se sestrojí pouze její asymptoty, viz obrázek 2.35. Toto přiblížení je poměrně přesné, pouze při kmitočtu zlomu $\omega_1 = \frac{1}{T_1}$ je odchylka od skutečného průběhu LAKCH největší a je 3 dB. Rovněž LFKCH se nesestrojuje přesná, ale při její konstrukci se používá aproximace sečnou.



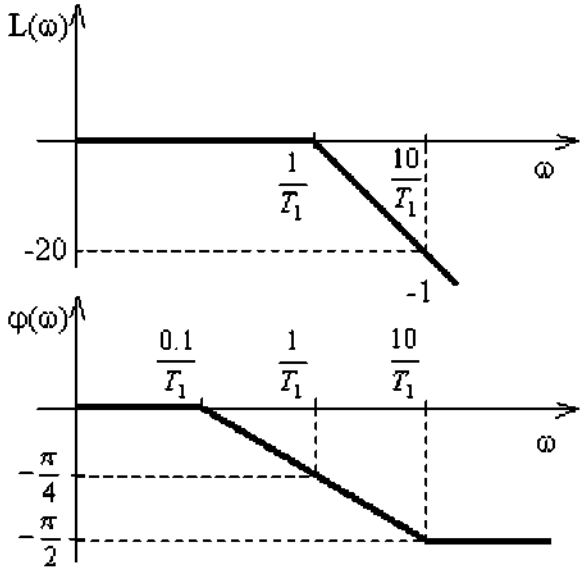
Obrázek 2.35 - LKCH pro přenos $G(s) = T_1 s + 1$

Proporcionální člen se setrvačností 1. řádu, který je popsán přenosem $G(s) = \frac{1}{T_1 s + 1}$. Pro určení LKCH opět použijeme vztahy (2.20), takže pro amplitudu a fázi dostaneme rovnice

$$L(\omega) = -20 \log \sqrt{T_1^2 \omega^2 + 1},$$

$$\varphi(\omega) = -\operatorname{arctg} \frac{T_1^2 \omega^2}{1}.$$

Odpovídající charakteristiky jsou na obrázku 2.36.



Obrázek 2.36 - LKCH pro proporcionální člen se setrvačností 1. řádu

Proporcionální člen se setrvačností 2. řádu, který je popsán přenosem ve tvaru $G(s) = \frac{1}{(T_0 s)^2 + 2\xi T_0 s + 1}$. Jde o případ, kdy jmenovatel přenosu $G(s)$ obsahuje kvadratický trojčlen s komplexně sdruženými kořeny. Amplituda a fáze tohoto členu se vypočte podle vztahů

$$A(\omega) = \frac{1}{\sqrt{(1 - T_0^2\omega^2)^2 + 4T_0^2\xi_0^2\omega^2}} \Rightarrow L(\omega) = -20\log\sqrt{(1 - T_0^2\omega^2)^2 + 4T_0^2\xi_0^2\omega^2},$$

$$\varphi(\omega) = \operatorname{arctg}\frac{2T_0\xi\omega}{1 - T_0^2\omega^2}.$$

Pro LKCH platí následující vztahy:

$\omega << \frac{1}{T_0}$: $L(\omega) = -20\log 1 = 0$,

$\varphi(\omega) = -\operatorname{arctg} 0 = 0$,

$\omega = \frac{1}{T_0}$: $L(\omega) = -20\log 2\xi$,

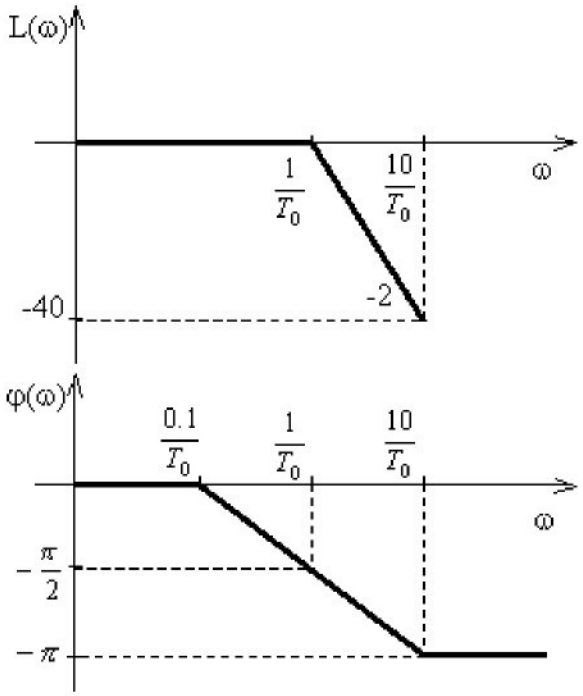
$\varphi(\omega) = -\operatorname{arctg}\infty = -\frac{\pi}{2}$,

$\omega >> \frac{1}{T_0}$: $L(\omega) = -20\log T_0^2\omega^2 = -40\log T_0^2\omega^2$,

$\varphi(\omega) = -\operatorname{arctg} 0 = -\pi$.

Z vypočtených hodnot vidíme, že asymptota LAKCH bude mít pro $\omega << \frac{1}{T_0}$ sklon 0 dB/dek, v bodě $\omega = \frac{1}{T_0}$ dojde ke zlomu a dále bude asymptota klesat se sklonem –40 dB/dek [-2]. Ve zlomovém kmitočtu bude velikost chyby od skutečného průběhu záviset na hodnotě

poměrného tlumení ξ, její maximální hodnota je pro $\xi = 1$ a má hodnotu 6 dB. Pro konstrukci LFKCH se opět použila aproximace sečnou. Tvar LKCH je vidět na obrázku 2.37. [VÍTEČKOVÁ&POLÁKOVÁ 1989]



Obrázek 2.37 - LKCH pro kmitavý člen

Příklad 2.10. *Logaritmické kmitočtové charakteristiky*

Nakreslete logaritmické kmitočtové charakteristiky pro systém popsaný obrazovým přenosem $G(s) = \dfrac{100(s+1)}{(s+10)(s^3+s^2+100s)}$

Řešení:

Zadaný přenos musíme rozdělit na základní členy, jejichž logaritmické kmitočtové charakteristiky umíme sestrojit. V našem případě přenos můžeme rozdělit na 4 členy, které musíme upravit tak, aby absolutní člen se rovnal 1.

$$G(s) = \frac{100(s+1)}{10\left(\frac{s}{10}+1\right)100s\left(\frac{s^2}{100}+\frac{s}{100}+1\right)} = \frac{1}{10s}\cdot(s+1)\cdot\frac{1}{\frac{s}{10}+1}\cdot\frac{1}{\frac{1}{10^2}s^2+2\cdot\frac{1}{10}\cdot\frac{1}{20}\cdot s+1}$$

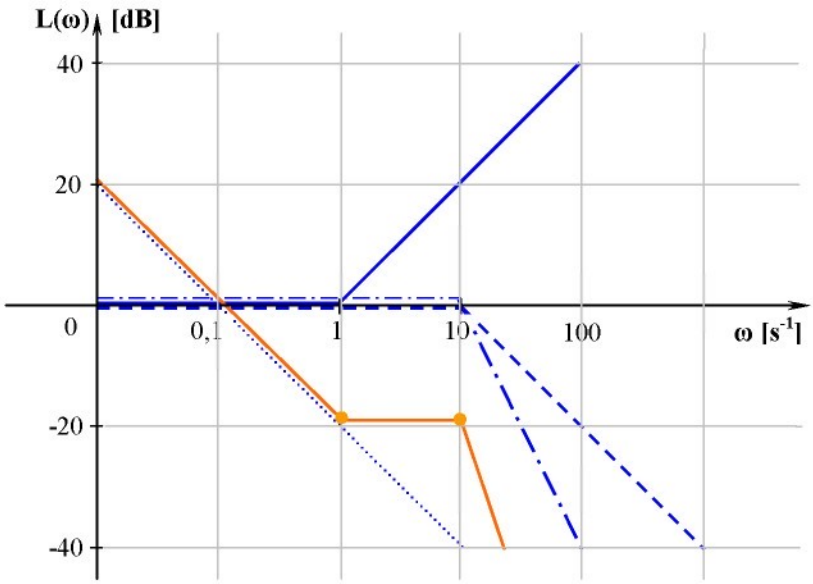
Určení rozsahu kmitočtu pro dílčí členy:

$$G_1(s) = \frac{1}{10s} \quad \rightarrow \quad \omega \in <0{,}01;1>$$

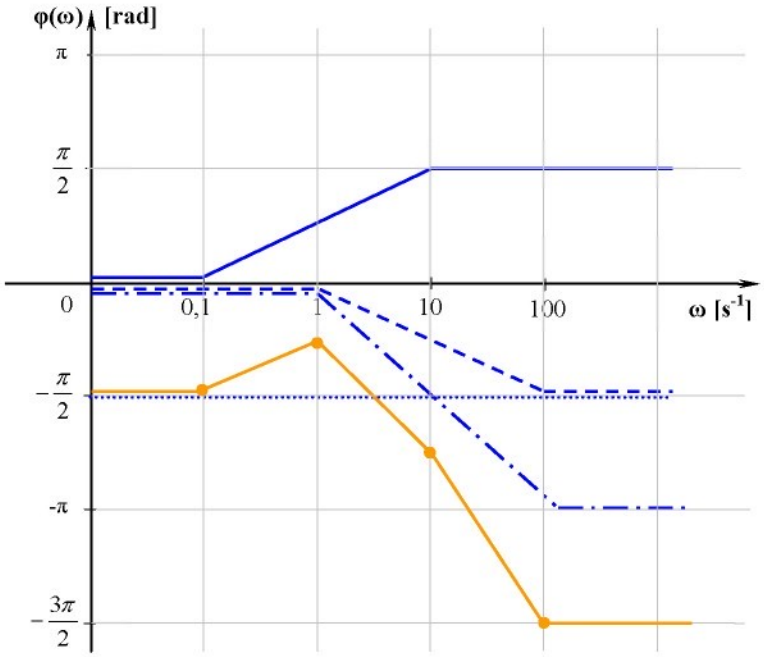
$$G_2(s) = s+1 \quad \rightarrow \quad \omega \in <0{,}1;10>$$

$$G_3(s) = \frac{1}{\frac{s}{10}+1} \qquad \rightarrow \qquad \omega \in <1;100>$$

$$G_4(s) = \frac{1}{\frac{1}{10^2}s^2 + 2\cdot\frac{1}{10}\cdot\frac{1}{20}\cdot s + 1} \qquad \rightarrow \qquad \omega \in <1;100>$$



a.)



b.)

Obrázek 2.38 - Aproximované logaritmické kmitočtové charakteristiky
a.) amplitudová, b.) fázová

Výsledné logaritmické kmitočtové charakteristiky jsou vidět na obrázku 2.38, modrou čárou jsou vykresleny dílčí členy a žlutá čáre je výsledná charakteristika, která vznikla součtem dílčích charakteristik (jedná se o sériové zapojení dílčích členů).

Shrnutí pojmů 2.4.

Kapitola definuje logaritmické kmitočtové charakteristiky a pravidla pro jejich vykreslování. Na konci je uveden příklad vykreslení těchto charakteristik.

Otázky 2.4.

1. V jakých hodnotách se počítá s reálným kmitočtem?
2. Kolik dB odpovídá zesílení 10?
3. Jak se vykreslí LKCH pro systémy v sériovém zapojení?

2.5. Základní rozdělení spojitých dynamických členů

Čas ke studiu: 2 hodiny

Cíl Po prostudování tohoto odstavce budete umět

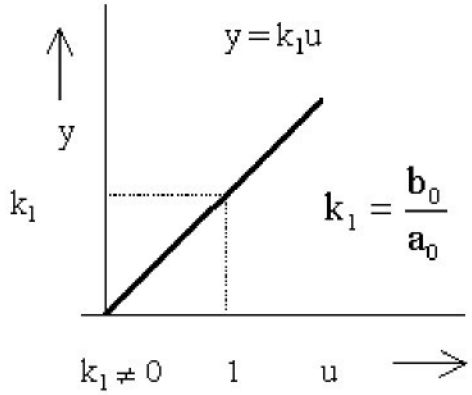
- definovat tři základní typy systémů – proporcionální, integrační, derivační – dle statické charakteristiky, přenosu, přechodové charakteristiky, amplitudo.fázové kmitočtové charakteristiky.
- definovat dopravní zpoždění a způsoby jeho aproximace.
- popsat minimálně přenos a dynamické charakteristiky pro první dva základní systémy každého typu.
- uvést reálné systémy patřící ke každému základnímu typu.

Výklad

Spojité dynamické systémy lze rozdělit na tři základní typy a to: proporcionální, derivační a integrační. V následující kapitole si ukážeme, jak lze jednotlivé typy poznat podle statické charakteristiky, obrazového přenosu, přechodové charakteristiky a amplitudo-fázové kmitočtové charakteristiky.

Proporcionální systémy

Statická charakteristika - existuje a je nenulová.

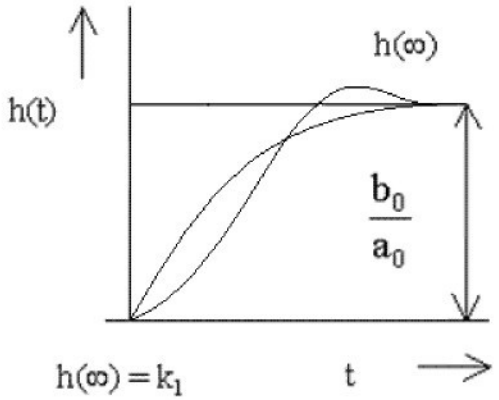


Obrázek 2.39 - Statická charakteristika.

Obrazový přenos – nelze vytknout proměnnou *s*.

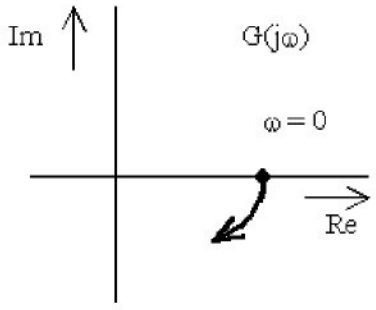
$$G(s) = \frac{b_m s^m + \ldots + b_1 s + b_0}{a_n s^n + \ldots + a_1 s + a_0}.$$

Přechodová charakteristika – ustálí se na nenulové hodnotě.



Obrázek 2.40 - Přechodová charakteristika.

Amplitudo-fázová kmitočtová charakteristika – začíná na kladné reálné poloose a jde do 4-tého kvadrantu.

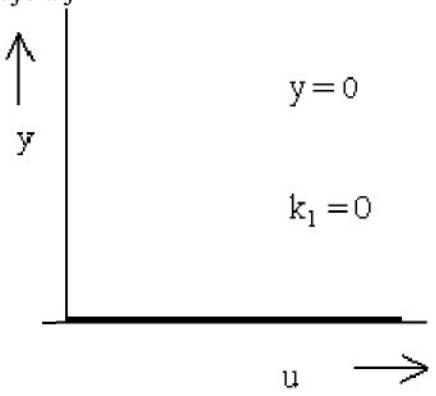


fáze $\varphi(0) = 0$
amplituda $A(0) = \text{konst.} = k_1$

Obrázek 2.41 - Amplitudo fázová kmitočtová charakteristika.

- **Derivační systémy**

Statická charakteristika – existuje a je nulová.

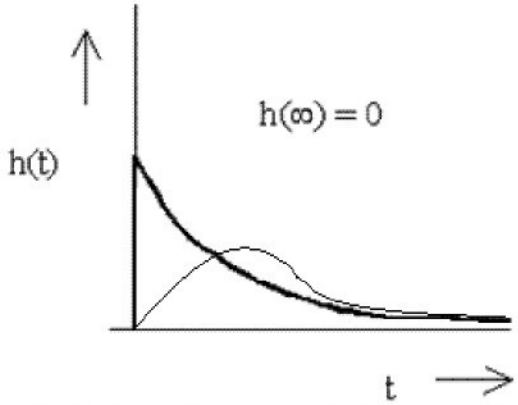


Obrázek 2.42 - Statická charakteristika.

Obrazový přenos – lze vytknout s v čitateli.

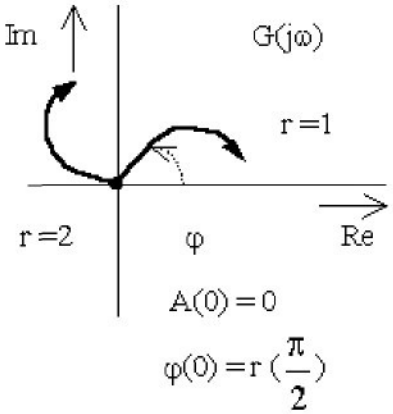
$$G(s) = \frac{s^r (b_m s^m + \ldots + b_1 s + b_0)}{a_n s^n + \ldots + a_1 s + a_0}.$$

Přechodová charakteristika – ustálí se na nule.



Obrázek 2.43 - Přechodová charakteristika.

Amplitudo-fázová kmitočtová charakteristika – začíná v nule.



Obrázek 2.44 - Amplitudo fázová kmitočtová charakteristika.

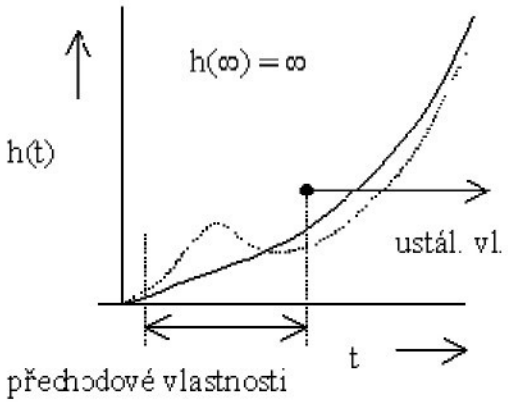
- **Integrační systémy**

Statická charakteristika – NEEXISTUJE.
$k_1 = \infty$

Obrazový přenos – proměnnou *s* lze vytknout ve jmenovateli; *q* je stupeň astatismu (mocnina vytknutého *s*).

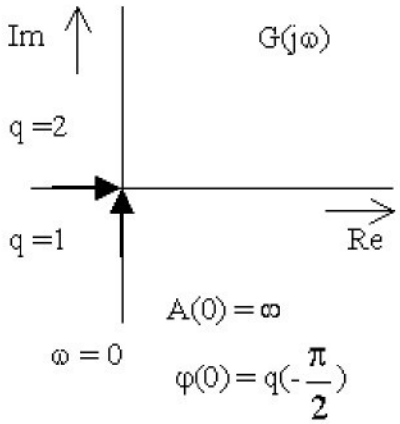
$$G(s) = \frac{b_m s^m + \ldots + b_1 s + b_0}{s^q (a_n s^n + \ldots + a_1 s + a_0)}.$$

Přechodová charakteristika – neustálí se.



Obrázek 2.45 - Přechodová charakteristika.

Amplitudo-fázová kmitočtová charakteristika – začíná v nekonečnu, stupeň astatismu určuje kvadrant.



Obrázek 2.46 - Amplitudo fázová kmitočtová charakteristika.

Nyní si uvedeme základní členy jednotlivých typů systémů, včetně tvaru přenosu, charakteristik a příklady reálných systémů.

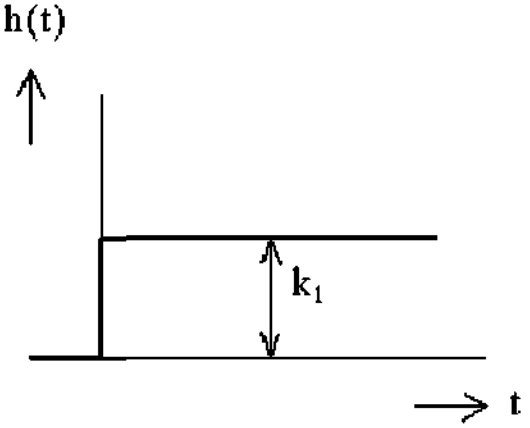
- **Proporcionální systémy**

Ideální proporcionální systém

Obrazový přenos:

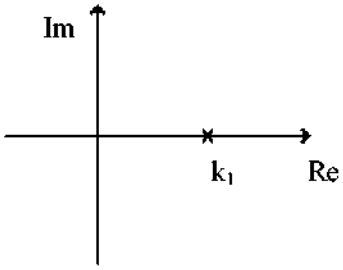
$G(s) = k_1$ k_1 – zesílení přenosu.

Přechodová charakteristika:



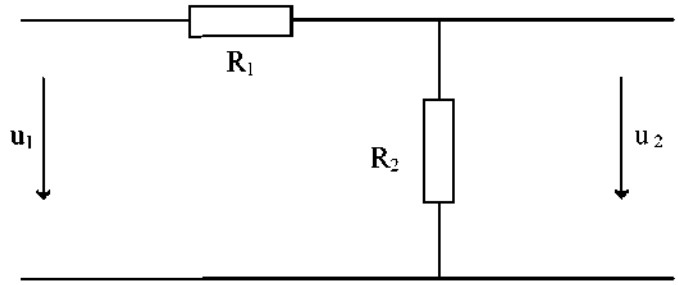
Obrázek 2.47 - Přechodová charakteristika ideálního proporcionálního členu.

Amplitudo-fázová kmitočtová charakteristika:



Obrázek 2.48 - Kmitočtová charakteristika ideálního proporcionálního členu.

Příklad: páka, dělič napětí s odpory, ...

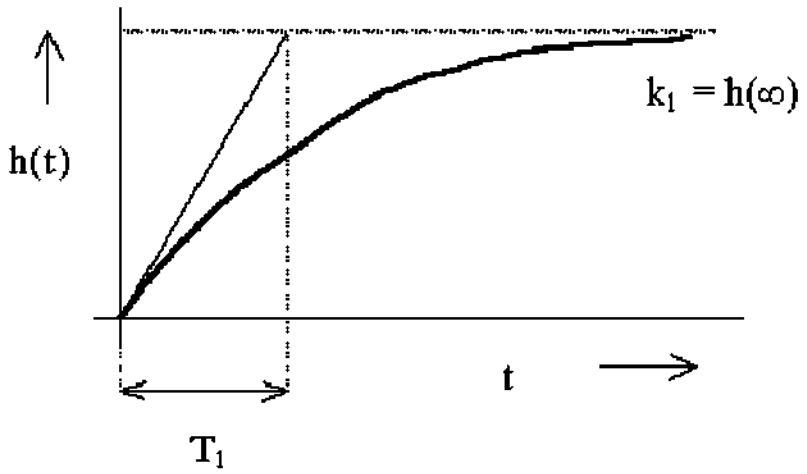


Obrázek 2.49 – Příklad ideálního proporcionálního členu.

Proporcionální systém se setrvačností 1. řádu (setrvačný článek)
Obrazový přenos:

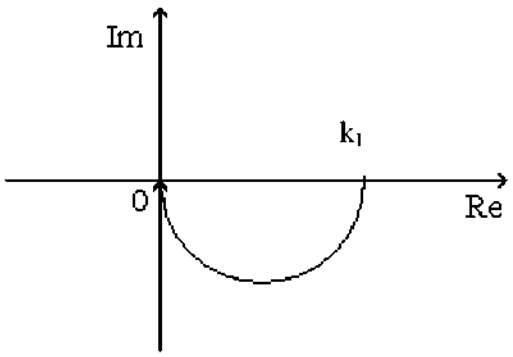
$$G(s) = \frac{k_1}{T_1 s + 1}.$$

Přechodová charakteristika:
$t = T_1 \rightarrow h(t) = 63{,}2\%$ ustálené hodnoty,
$t = 3T_1 \rightarrow h(t) = 95\%$ ustálené hodnoty,
$t = 4T_1 \rightarrow h(t) = 98\%$ ustálené hodnoty.

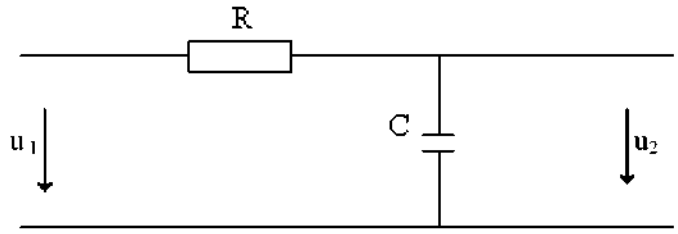


Obrázek 2.50 - Přechodová charakteristika proporcionálního členu se setrvačností 1. řádu.

Amplitudo-fázová kmitočtová charakteristika: nepřekročí IV kvadrant



Obrázek 2.51 - Kmitočtová charakteristika proporcionálního členu se setrvačností 1. řádu.
Příklad: nádrž s volným odtokem, dělič napětí s odporem a kondenzátorem, …



Obrázek 2.52 – Příklad proporcionálního členu se setrvačností 1. řádu.

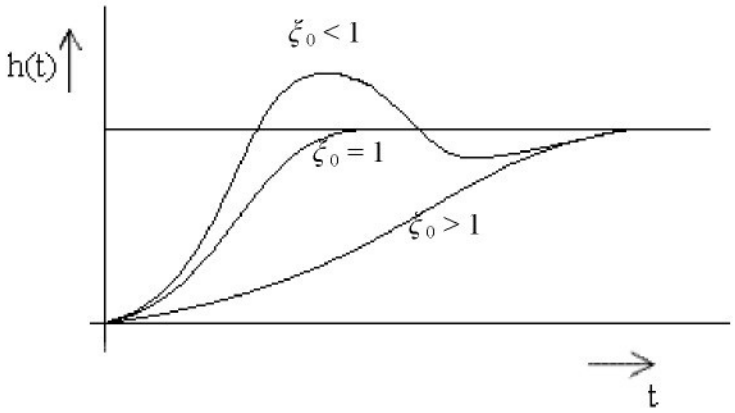
Proporcionální systém se setrvačností 2. řádu (kmitavý článek)

Obrazový přenos:

$$G(s) = \frac{k_1}{(T_1 s + 1)^2} = \frac{k_1}{(T_1 s + 1)(T_2 s + 1)} = \frac{k_1}{T_0^2 s^2 + 2T_0 \xi_0 s + 1}$$

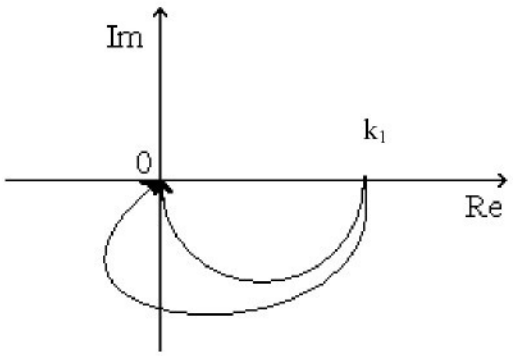
Přechodová charakteristika:
- podle koeficientu poměrného tlumení ξ_0 dělíme proporcionální členy se setrvačností 2. řádu.
$\xi_0 > 1$ aperiodický

$\xi_0 = 1$ mezní aperiodický
$0 \leq \xi_0 < 1$ kmitavý



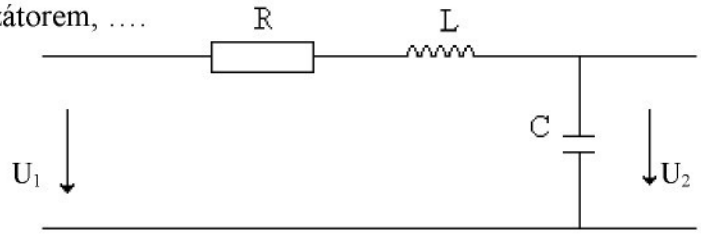
Obrázek 2.53 - Přechodová charakteristika proporcionálního členu se setrvačností 2. řádu.

Amplitudo-fázová kmitočtová charakteristika:



Obrázek 2.54 - Kmitočtová charakteristika proporcionálního členu se setrvačností 2. řádu.

Příklad: hmota na pružině a tlumiči, dvě nádrže s volným odtokem, dělič napětí s odporem, cívkou a kondenzátorem, ….



Obrázek 2.55 – Příklad proporcionálního členu se setrvačností 2. řádu.

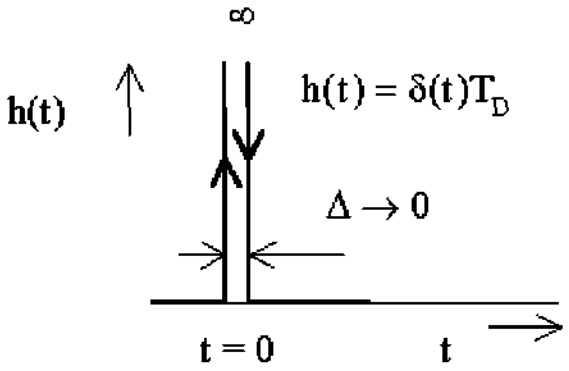
❑ Derivační systémy

Ideální derivační systém

Obrazový přenos:

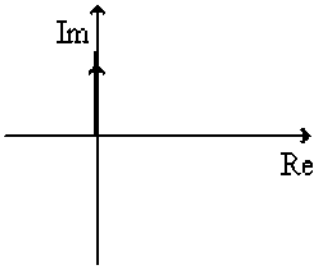
$G(s) = T_D s$ T_D - derivační časová konstanta.

Přechodová charakteristika:



Obrázek 2.56 - Přechodová charakteristika ideálního derivačního členu.

Amplitudo-fázová kmitočtová charakteristika:



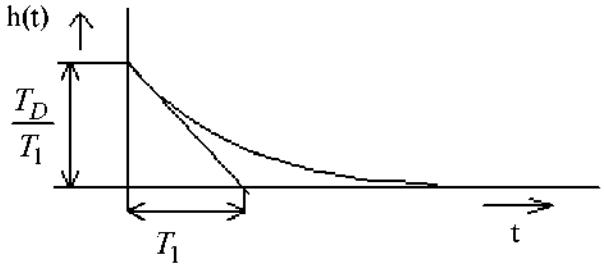
Obrázek 2.57 - Kmitočtová charakteristika ideálního členu.

Derivační systém se setrvačností 1. řádu

Obrazový přenos:

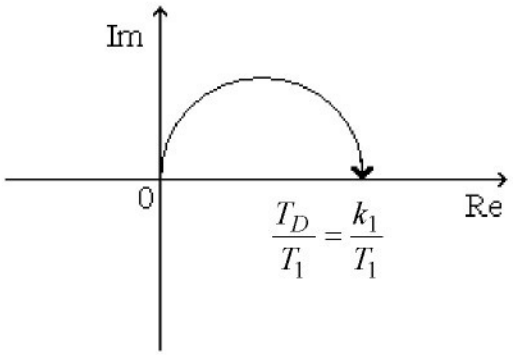
$$G(s) = \frac{T_D s}{T_1 s + 1} = \frac{k_1 s}{T_1 s + 1}.$$

Přechodová charakteristika:



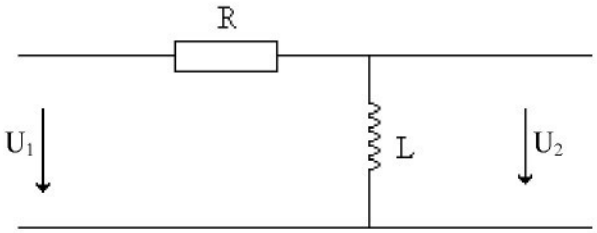
Obrázek 2.58 - Přechodová charakteristika derivačního členu se setrvačností 1. řádu.

Amplitudo-fázová kmitočtová charakteristika:



Obrázek 2.59 - Kmitočtová charakteristika derivačního členu se setrvačností 1. řádu.

Příklad: dělič napětí s odporem a cívkou, …



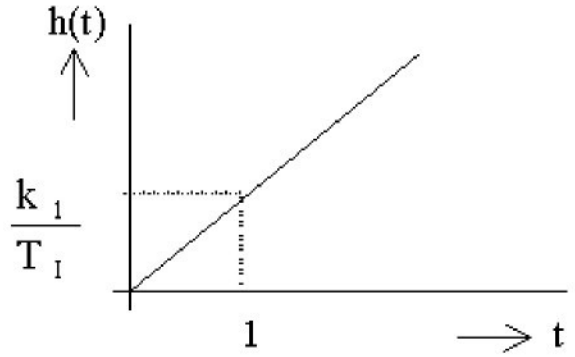
Obrázek 2.60 – Příklad derivačního členu se setrvačností 1. řádu.

❑ Integrační systém

Ideální integrační systém (integrátor)
Obrazový přenos:

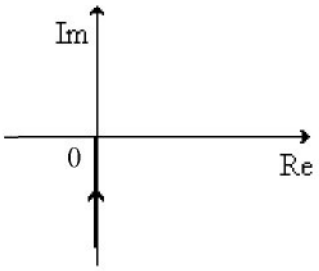
$$G(s) = \frac{1}{T_I s} = \frac{k_1}{s} \qquad T_I \text{ - integrační časová konstanta.}$$

Přechodová charakteristika:



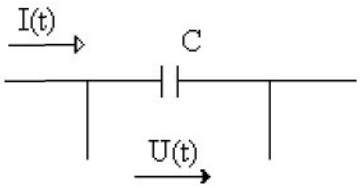
Obrázek 2.61 - Přechodová charakteristika ideálního integračního členu.

Amplitudo-fázová kmitočtová charakteristika:



Obrázek 2.62 - Kmitočtová charakteristika ideálního integračního členu.

Příklad: kondenzátor, nádrž bez odtoku, …



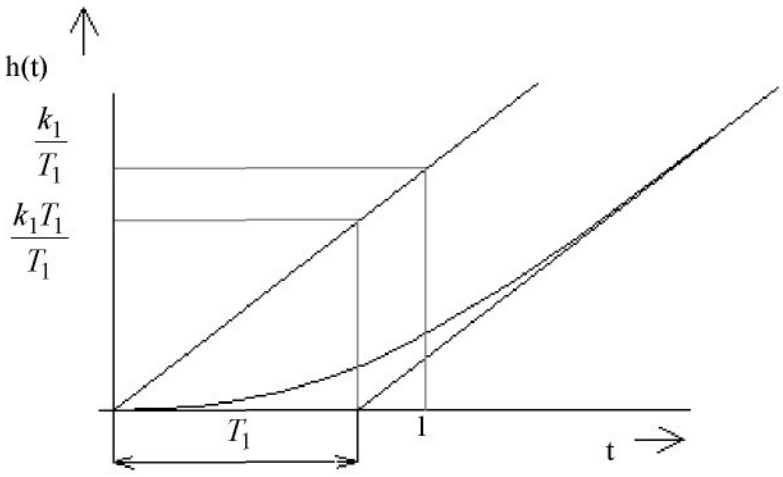
Obrázek 2.63 – Příklad ideálního integračního členu.

Integrační systém se setrvačností 1. řádu

Obrazový přenos:

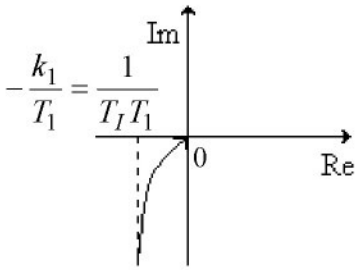
$$G(s) = \frac{1}{T_I s(T_1 s + 1)} = \frac{k_1}{s(T_1 s + 1)}.$$

Přechodová charakteristika:



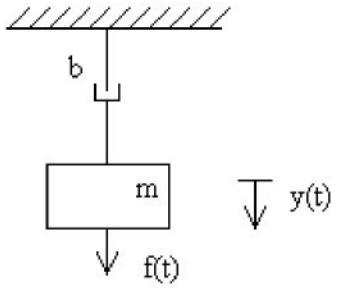
Obrázek 2.64 - Přechodová charakteristika integračního členu se setrvačností 1. řádu.

Amplitudo-fázová kmitočtová charakteristika:



Obrázek 2.65 - Kmitočtová charakteristika integračního členu se setrvačností 1. řádu.

Příklad: hmota na tlumiči, ….



Obrázek 2.66 – Příklad integračního členu se setrvačností 1. řádu.

❑ Dopravní zpoždění

V chování systému, kromě toho že je můžeme dělit na základní typy, lze vysledovat vlastnost dopravního zpoždění, tzn. že hodnota výstupu je v čase zpožděná; změna se neprojeví okamžitě na výstupu systému. Následující příklad proporcionálního systému se setrvačností 1. řádu ukazuje, jak rozpoznáme vlastnost dopravního zpoždění v diferenciální rovnici, obrazovém přenosu, přechodové charakteristice a amplitudo-fázové kmitočtové charakteristice.

Obrazový přenos:

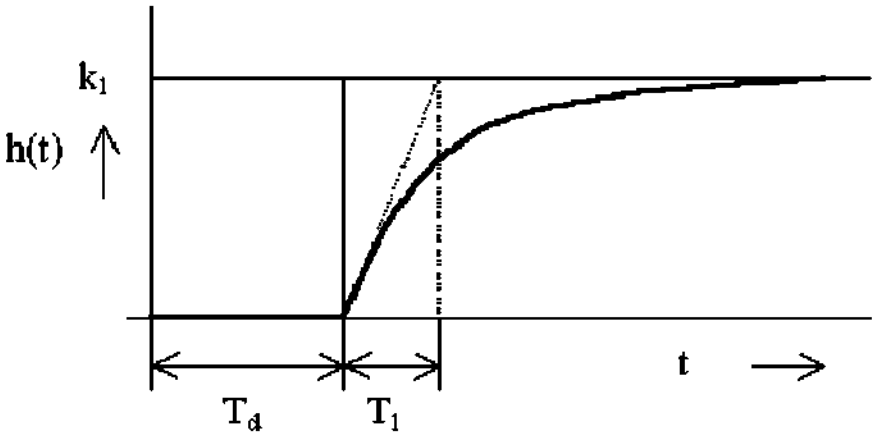
$$G(s) = \frac{k_1}{1 + T_1 s} e^{-sT_d}$$

T_d - časová konstanta dopravního zpoždění.

Diferenciální rovnice:

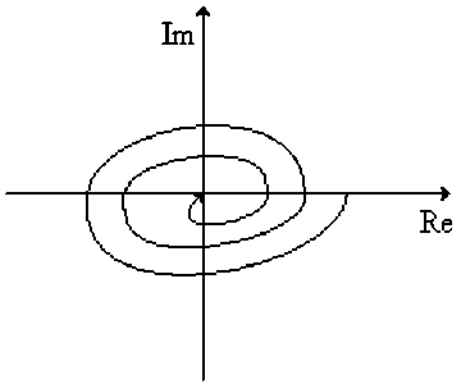
$$T_1 \frac{dy}{dt} + y = k_1 u(t - T_D).$$

Přechodová charakteristika:



Obrázek 2.67 - Přechodová charakteristika proporcionálního členu se setrvačností 1. řádu a s dopravním zpožděním.

Amplitudo-fázová kmitočtová charakteristika:



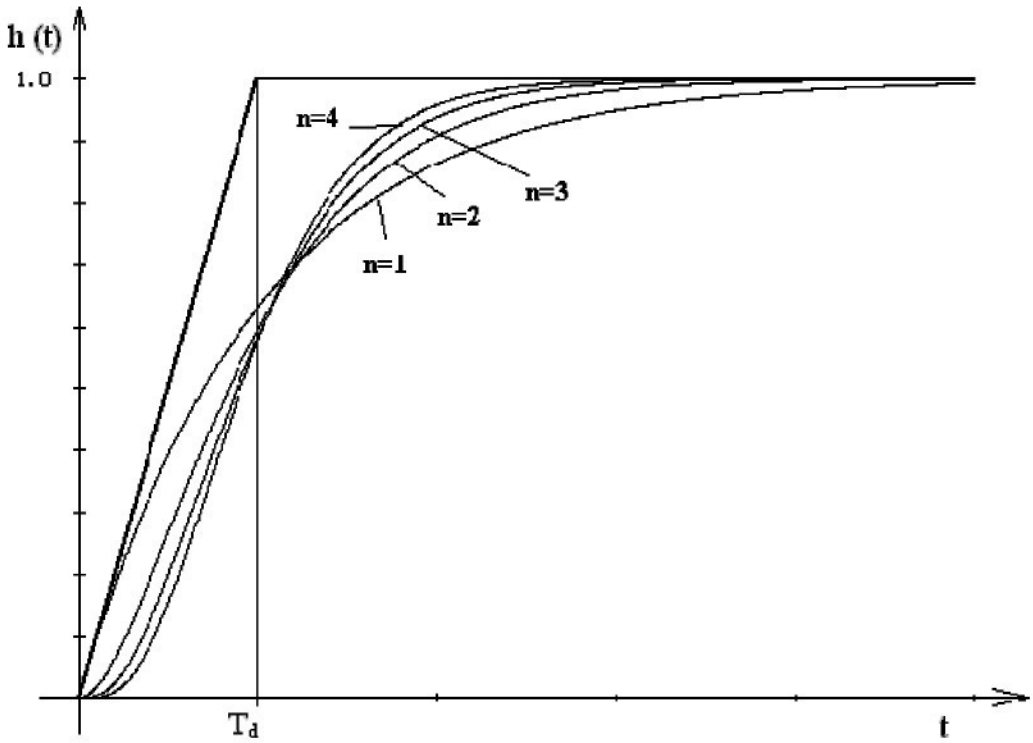
Obrázek 2.68 - Kmitočtová charakteristika proporcionálního členu se setrvačností 1. řádu a s dopravním zpožděním.

Člen dopravního zpoždění v obrazovém přenosu představuje exponenciální výraz $e^{-T_d s}$, v diferenciální rovnicí je to výraz $(t\text{-}T_d)$ u vstupní nebo výstupní veličiny. Jak je vidět porovnáním obrázku 2.59 a 2.42, tvar přechodové charakteristiky se nezmění, jen začátek je z nulové hodnoty posunut do hodnoty T_d. Oproti tomu tvar amplitudo-fázové kmitočtové charakteristiky se významně změnil, pro systém s dopravním zpožděním tato charakteristika krouží kolem počátku, takže v kmitočtové oblasti se dopravní zpoždění významně projevuje a to tím, že fáze systému klesá k nekonečné hodnotě, a proto tyto členy označujeme jako *členy s neminimální fází*.

Aproximace dopravního zpoždění

V regulačních obvodech se často vyskytuje člen dopravního zpoždění. Tento člen je buď vlastností regulované soustavy nebo se vyskytne v případě použití číslicového regulátoru. Tím máme vlastnosti regulačního obvodu popsány vztahem, který obsahuje exponenciální funkci a pokud chceme použít metody syntézy regulačních obvodů pro lineární regulační obvody neobsahující dopravní zpoždění, musíme toto dopravní zpoždění aproximovat. Dále jsou uvedeny používané aproximace.

$$e^{-T_d s} \approx \frac{1}{\left(\frac{T_d}{n} s + 1\right)^n}$$ → nahrazujeme zpoždění setrvačností.



Obrázek 2.69 - Aproximace dopravního zpoždění

V praxi se používá $n = 1$, $n = 2$. Čím větší je n, tím je aproximace přesnější. Tato aproximace se vyznačuje jednoduchostí, avšak málo se používá, neboť člen s neminimální fází (člen s dopravním zpožděním) se nahrazuje členem s minimální fází.

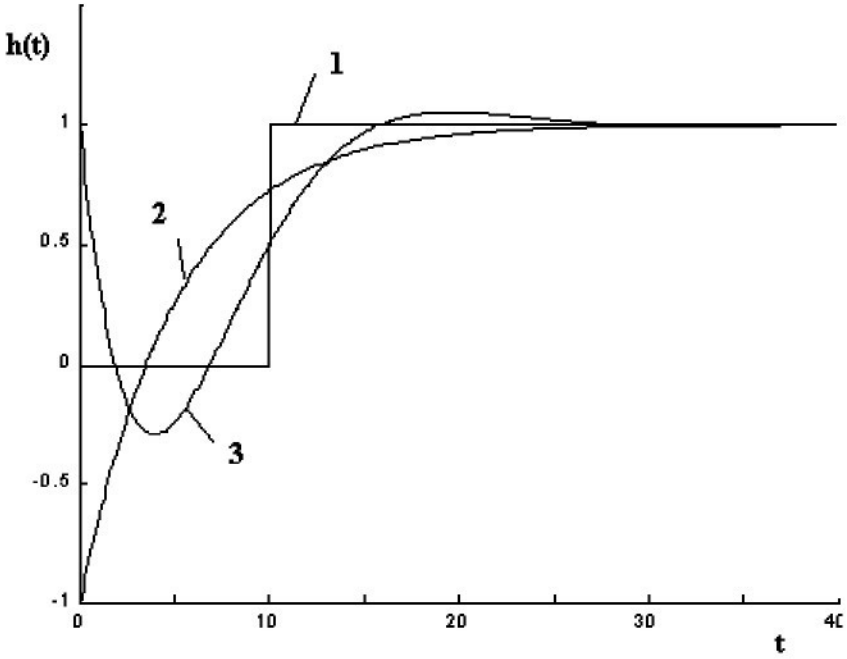
Aproximace Taylorovým rozvojem:

I. řádu.

$$e^{-T_d s} \approx \frac{1-\frac{T_d s}{2}}{1+\frac{T_d s}{2}},$$

II. řádu

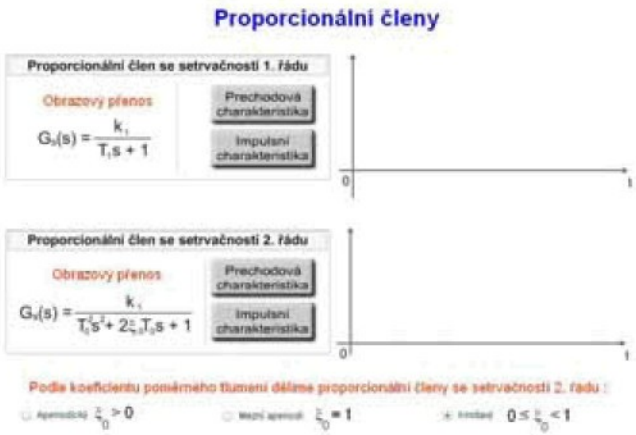
$$e^{-T_d s} \approx \frac{1-\frac{T_d s}{2}+\frac{T_d{}^2 s^2}{8}}{1+\frac{T_d s}{2}+\frac{T_d{}^2 s^2}{8}}.$$



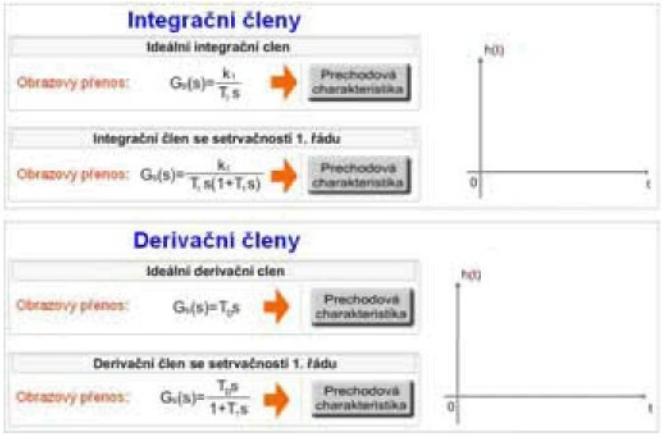
1 – dopravní zpoždění e^{-Td} s, 2 – aproximace Taylorovým rozvojem I. řádu, 3 – aproximace Taylorovým rozvojem II. řádu

Obrázek 2.70 – Aproximace dopravního zpoždění Padého a Taylorovým rozvojem

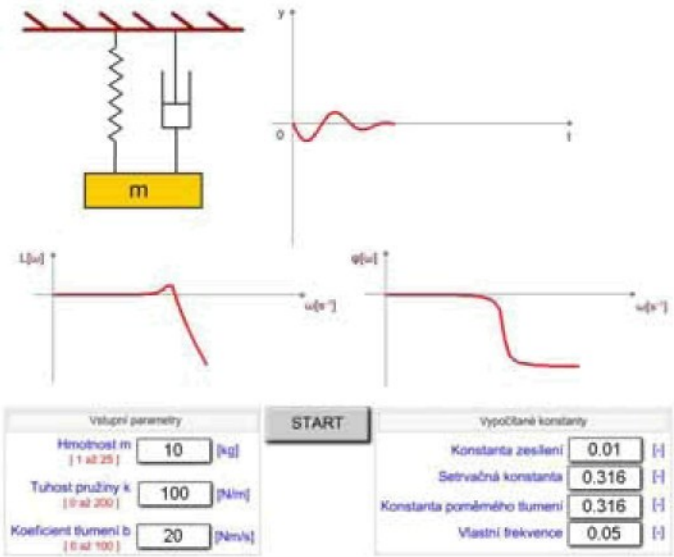
Na následujících obrázcích 2.64 až 2.67 lze spustit animace, které umožňují vykreslovat charakteristiky základních typů systémů nebo animují chování vybraných systémů.



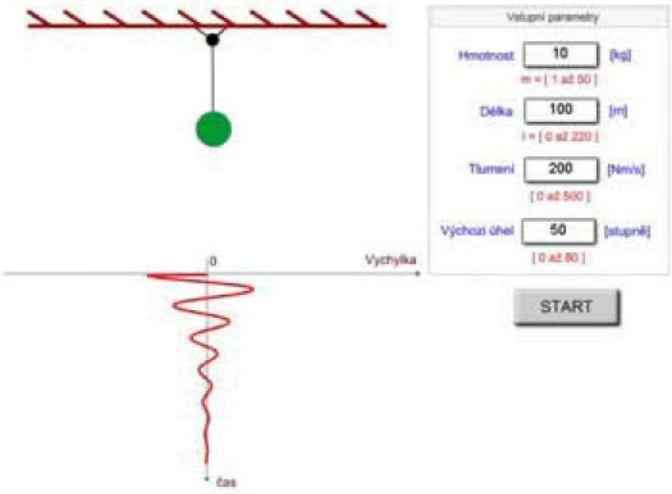
Obrázek 2.71 – Animace chování proporcionálních systémů



Obrázek 2.72 – Animace chování integračních a derivačních systémů



Obrázek 2.73 – Animace chování systému hmota na tlumiči a pružině



Obrázek 2.74 – Animace chování kyvadla

Shrnutí pojmů 2.5.

Definice tří základní typů systémů (proporcionální, integrační, derivační) a jejich určení na základě statické charakteristiky, obrazového přenosu, přechodové charakteristiky a amplitudo-fázové kmitočtové charakteristiky. Pak jsou uvedeny tvary charakteristik pro základní členy a příklady reálných systémů s danými vlastnostmi. Na závěr je vysvětlena vlastnost dopravního zpoždění a jeho vliv na tvar charakteristik.

Otázky 2.5.

1. Jak vypadá přechodová charakteristika setrvačného článku s časovou konstantou T_1?
2. Nakreslete přechodovou charakteristiku integračního článku.
3. Uveďte příklad ideálního derivačního systému.
4. Jaká je ustálená hodnota přechodové charakteristiky pro derivační systémy?
5. V jaké hodnotě začíná amplitudo-fázová kmitočtová charakteristika proporcionálního systému?
6. Lze vytknout proměnnou s v čitateli přenosu integračních systémů?

2.6. Algebra blokových schémat

Čas ke studiu: 2 hodiny

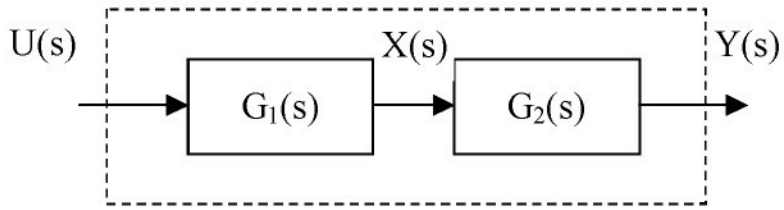
Cíl Po prostudování tohoto odstavce budete umět

- definovat výsledné přenosy základních zapojení systémů, regulační obvod a jeho základní prvky.
- popsat rozdíl mezi spojitým a diskrétním regulačním obvodem.
- vyřešit výsledný přenos libovolně složitého blokového schématu.
- vypočítat základní přenosy regulačního obvodu.

Výklad

Blokovým schématem lze popsat vlastnosti systémů, je to ekvivalentní zápis k rovnicím. Blokové schéma je tvořeno bloky (vlastnosti lze popsat matematickými výrazy, nejčastěji obrazovými přenosy) a šipkami (veličiny). Algebra blokových schémat jsou pravidla, podle nichž vytváříme matematický model systému tvořeným z více bloků. Nejdříve je nutné v blokových schématech určit základní zapojení bloků. Rozlišujeme tři základní zapojení bloků v blokových schématech a to: sériové, paralelní a zpětnovazební (antiparalelní) zapojení. Budeme předpokládat, že vlastnosti systémů jsou popsány obrazovými přenosy.

- **Sériové zapojení**



Obrázek 2.75 - Blokové schéma sériového zapojení

Pro jednotlivé členy platí vztahy:

$$X(s) = G_1(s)U(s)\,; Y(s) = G_2(s)X(s)\,.$$

Vztah pro výstupní veličinu odvodíme eliminací proměnné *X(s)*

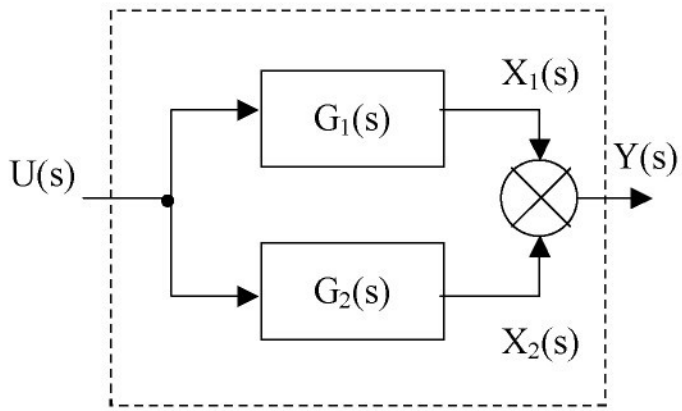
$$Y(s) = G_2(s)\,X(s) = G_1(s)G_2(s)U(s)\,.$$

Je tedy možno sériově zapojené členy nahradit jedním členem s přenosem

$$G(s) = \frac{Y(s)}{U(s)} = G_1(s).G_2(s) \qquad (2.22)$$

Při zapojení členů za sebou je výsledný přenos dán součinem přenosů jednotlivých členů.

- **Paralelní zapojení**



Obrázek 2.76 - Blokové schéma paralelního zapojení

Pro jednotlivé členy a pro součtový uzel platí vztahy:

$$X_1(s) = G_1(s)U(s)\,, X_2(s) = G_2(s)U(s),. Y(s) = X_1(s) + X_2(s)$$

Eliminací $X_1(s)$, $X_2(s)$ vypočítáme výstupní veličinu *Y(s)*

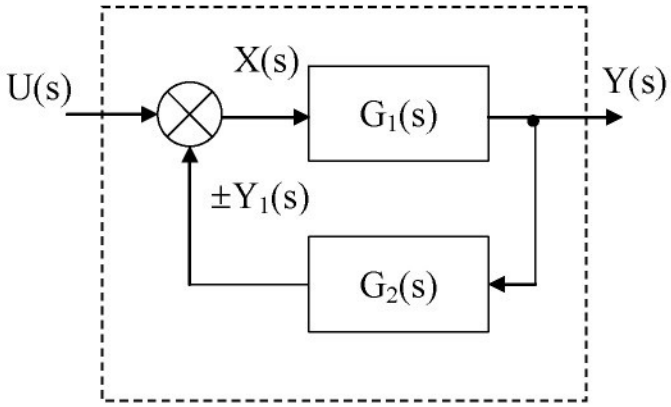
$$Y(s) \;=\; G_1(s)\,U(s) \;+\; G_2(s)\,U(s) = \;\left[G_1(s) \;+\; G_2(s)\right]U(s)\,.$$

Při paralelním zapojení členů je možné je nahradit jedním členem s přenosem

$$G(s) = \frac{Y(s)}{U(s)} = G_1(s) + G_2(s). \qquad (2.23)$$

Při paralelním zapojení je výsledný přenos dán součtem (respektive rozdílem) přenosů jednotlivých členů.

- **Zpětnovazební zapojení (antiparalelní)**



Obrázek 2.77 - Blokové schéma zpětnovazebního zapojení

Pro jednotlivé členy a rozdílový uzel platí následující vztahy:

$$Y(s) = G_1(s)X(s), Y_1(s) = G_2(s)Y(s), \; X(s) = U(s) \pm Y_1(s).$$

Eliminací $X(s)$ a $Y_1(s)$ obdržíme vztah pro výstupní veličinu $Y(s)$

$$Y(s) = G_1(s)\,X(s) = [U(s) \pm Y_1(s)]G_1(s) = [U(s) \pm G_2(s)Y(s)]G_1(s),$$

$$Y(s) \mp G_1(s)G_2(s)Y(s) = G_1(s)U(s).$$

Při zpětnovazebním zapojení je možno takto zapojené členy nahradit jedním členem s přenosem

$$G(s) = \frac{Y(s)}{U(s)} = \frac{G_1(s)}{1 \mp G_1(s)G_2(s)}. \qquad (2.24)$$

Při zpětnovazebním zapojení je výsledný přenos dán zlomkem, kdy v čitateli je tzv. přenos přímé větve a ve jmenovateli $1 \mp$ součin přenosů přímé větve a zpětné vazby:

$$\text{celkový přenos} = \frac{\text{přenos přímé větve}}{1 \mp (\text{přenos přímé větve}).(\text{přenos zpětné vazby})}$$

Znalost blokové algebry umožňuje zjednodušovat složitá bloková schémata a stanovit výsledný přenos zapojení. Zjednodušovat je nejlépe tak, že uvnitř blokového schématu hledáme některé z uvedených základních zapojení a postupně nahazujeme tato zapojení jediným členem.[ŠVARC 1992]

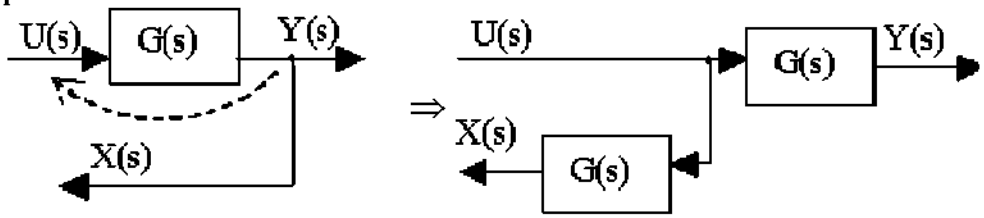
Pravidla pro úpravu blokového schématu

Při zjednodušení blokového schématu se vychází z náhradních přenosů pro základní zapojení (sériové, paralelní, zpětnovazební). Avšak často je nutno upravit blokové schéma,

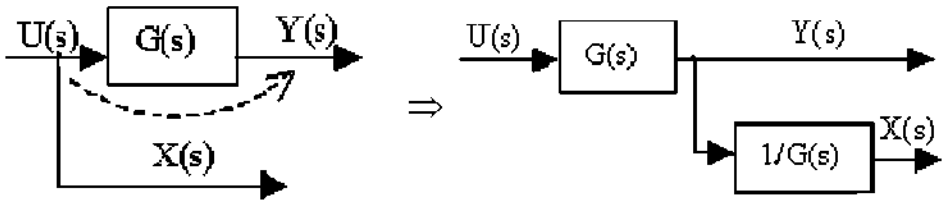
aby byla základní zapojení jednoznačně určena. Při úpravě blokových schémat se využívají dvě základní pravidla, a to:

1. Pravidlo pro přesun uzlu

- před blok

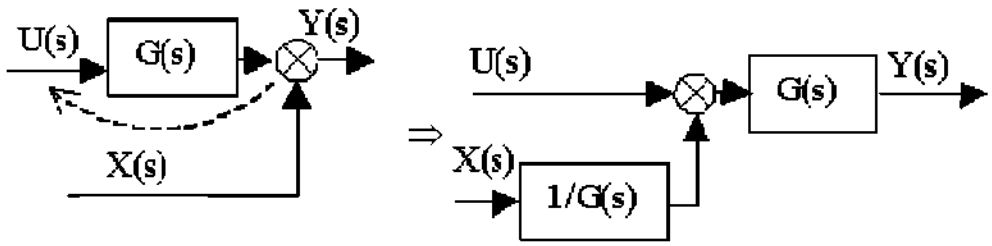


- za blok

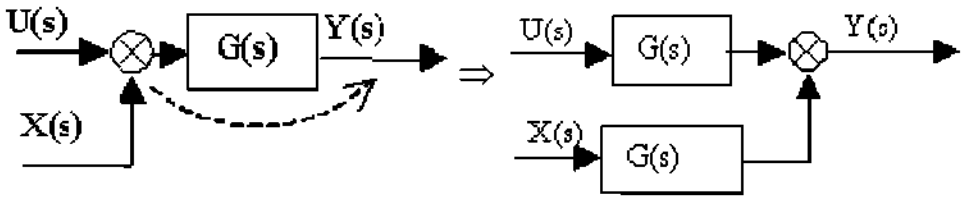


2. Pravidlo pro přesun sumačního členu:

- před blok

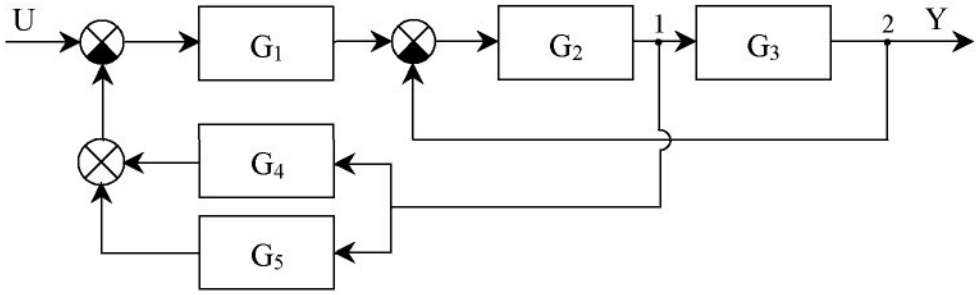


- za blok



Příklad 2.11. *Algebra blokových schémat*

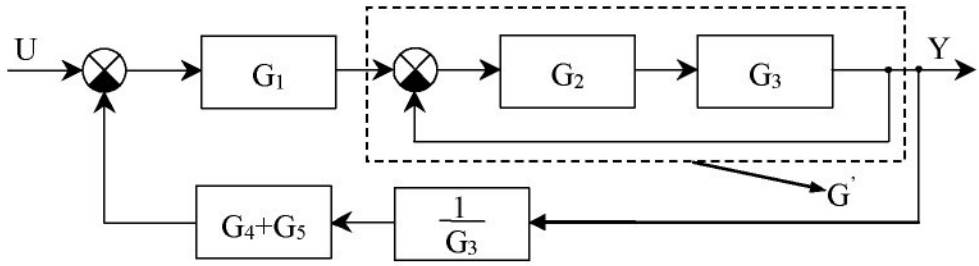
Určete výsledný přenos systému, jehož blokové schéma je na obr. 2.71.



Obrázek 2.78 - Blokové schéma

Řešení:

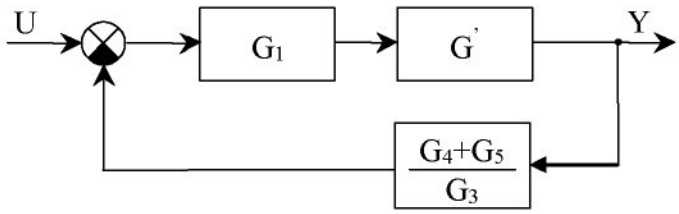
Zadané blokové schéma je tvořeno paralelním (bloky G_4,G_5), sériovým a dvěmi zpětnovazebními zapojeními. Zpětné vazby se nám však kříží, proto budeme muset přesunout uzel (1) za G_3 nebo uzel (2) před G_3. Dostaneme upravené blokové schéma:



Nyní lze vypočítat náhradní přenos G' vnitřního zpětnovazebního zapojení:

$$G' = \frac{G_2 G_3}{1 + G_2 G_3}$$

Pak blokové schéma má tvar

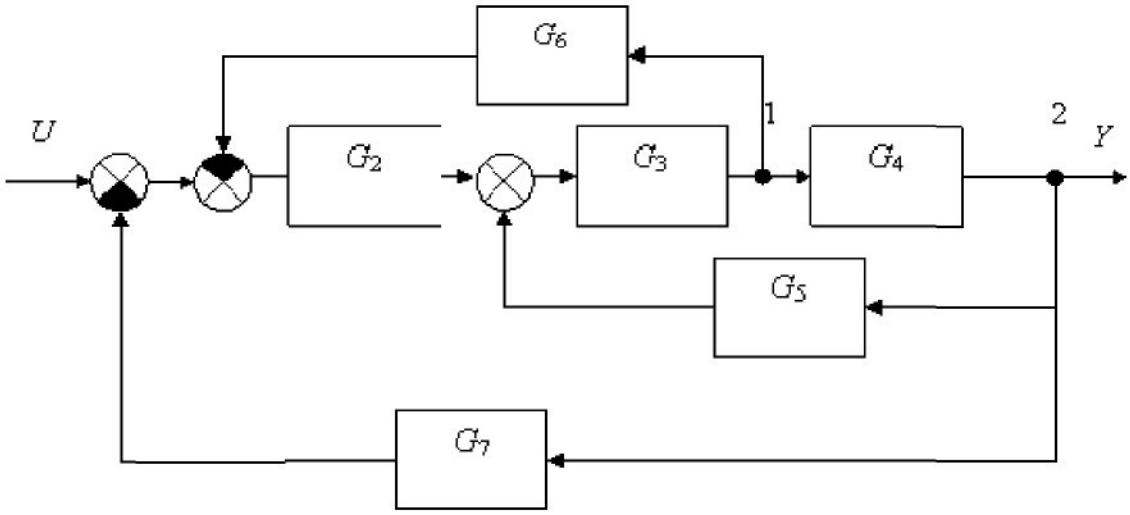


Výsledný přenos systému je:

$$G = \frac{G_1 G'}{1 + G_1 G' (G_4 + G_5)\frac{1}{G_3}} = \frac{G_1 G_2 G_3}{1 + G_2 G_3 + G_1 G_2 G_4 + G_1 G_2 G_5}$$

Příklad 2.12. *Algebra blokových schémat*

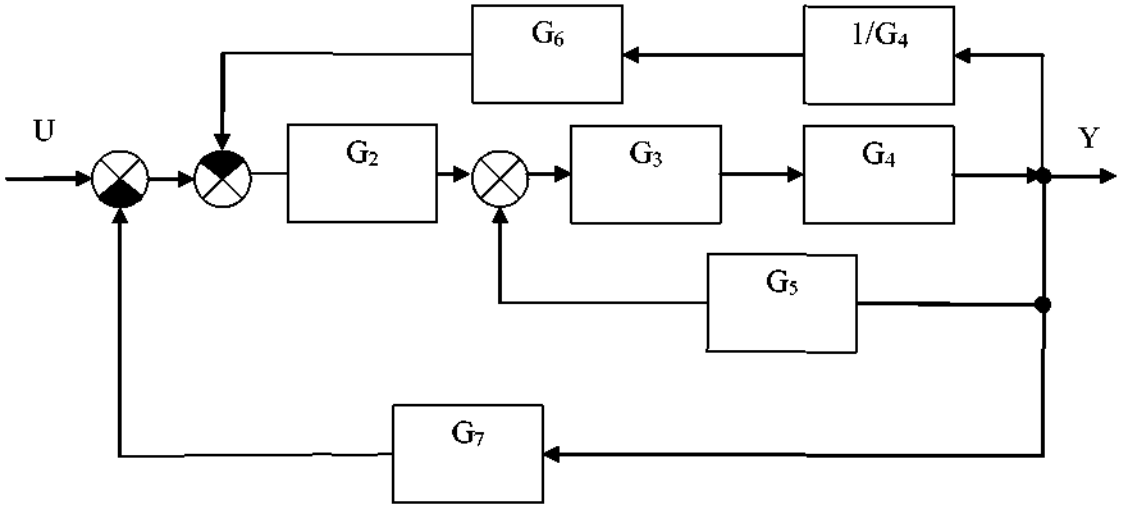
Určete výsledný přenos systému, jehož blokové schéma je na obr. 2.8.



Obrázek 2.79 - Blokové schéma

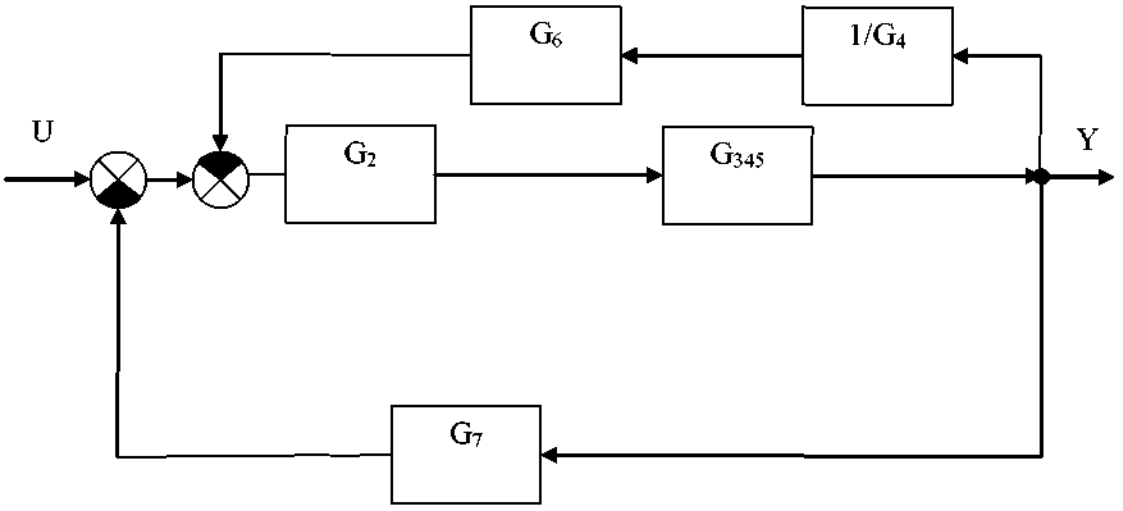
Řešení:

Zadané blokové schéma je tvořeno sériovým a třemi zpětnovazebními zapojeními. Zpětné vazby se nám však kříží, proto budeme muset přesunout uzel (1) za G_3 nebo uzel (2) před G_3. Dostaneme upravené blokové schéma:Přesuneme uzel od přenosu G_4.



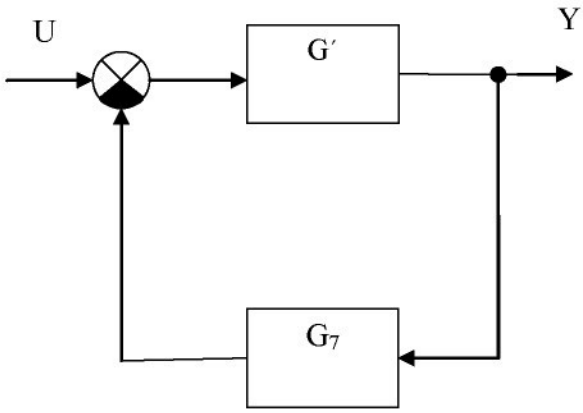
Vypočítáme přenos G_{345}

$$G_{345} = \frac{G_3 G_4}{1 - G_3 G_4 G_5}$$



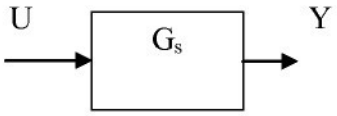
Vypočítáme přenos G´.

$$G' = \frac{G_2 G_{345}}{1 + G_2 G_{345} G_6 \frac{1}{G_4}}$$



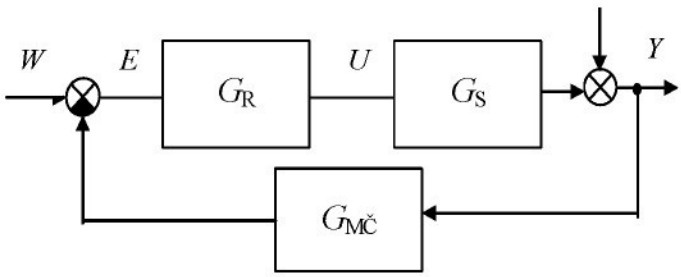
Vypočítáme konečný přenos G_s.

$$G_s = \frac{G'}{1 + G_7 G'}$$



❑ Základní přenosy RO.

Předpokládejme, že vlastnosti jednotlivých členů RO jsou popsány obrazovými přenosy, viz obrázek 2.71.



W – žádaná veličina, Y – regulovaná veličina, U – řídicí veličina, E – regulační odchylka, V – poruchová veličina, G_R – přenos regulátoru, G_S – přenos regulované soustavy, $G_{MČ}$ – přenos měřícího členu

Obrázek 2.80 - Blokové schéma regulačního obvodu

Uvažujme vstupní veličinu regulačního obvodu žádanou veličinu w a výstupní veličinu regulačního obvodu regulovanou veličinu y. Za předpokladu, že poruchová veličina bude rovna nule, tedy $v = 0$, dostáváme přenos mezi řídící veličinou a regulovanou veličinou, nazývaný jako **přenos řízení**

$$G_{wy} = \frac{Y}{W} = \frac{G_S G_R}{1 + G_S G_R G_{MČ}} . \qquad (2.25)$$

Jestliže uvažujeme jako vstupní veličinu regulačního obvodu poruchovou veličinu v a výstupní veličinu opět regulovanou veličinu y, za předpokladu, že žádaná veličina je rovna

nule, tedy $w = 0$, dostáváme přenos mezi poruchovou a regulovanou veličinou, nazývaný jako **přenos poruchy**.

$$G_{vy} = \frac{Y}{V} = \frac{1}{1 + G_S G_R G_{MČ}}. \tag{2.26}$$

Pro vyšetřování vlastností regulačních obvodů je nutná znalost obou přenosů, přenosu řízení i poruchy. Druhou možností popisu vlastností regulačních obvodů je znalost odchylkových přenosů. Uvažujme vstupní veličinu regulačního obvodu žádanou veličinu w a výstupní veličinu regulačního obvodu regulační odchylku e. Za předpokladu, že poruchová veličina bude rovna nule, tedy $v = 0$, dostáváme přenos mezi řídící veličinou a regulační odchylkou, nazývaný jako **odchylkový přenos řízení**

$$G_{we} = \frac{E}{W} = \frac{1}{1 + G_S G_R G_{MČ}}. \tag{2.27}$$

Jestliže uvažujeme jako vstupní veličinu regulačního obvodu poruchovou veličinu v a výstupní veličinu opět regulační odchylku e, za předpokladu, že žádaná veličina je rovna nule, tedy $w = 0$, dostáváme přenos mezi poruchovou a regulační odchylkou, nazývaný jako **odchylkový přenos poruchy**.

$$G_{ve} = \frac{Y}{V} = \frac{-G_{MČ}}{1 + G_S G_R G_{MČ}}. \tag{2.28}$$

Regulátor působí na regulovanou soustavu tak, aby byl splněn **cíl regulace**, tj. $y \to w$, resp. $e(t) \to 0$. Při znalosti základních přenosů řízení lze vypočítat regulovanou veličinu a regulační odchylku podle vztahů

$$Y = G_W W + G_V V, \; E = G_{WE} W + G_{VE} V. \tag{2.29}$$

Aby byl splněn cíl regulace, musí pro základní přenosy platit:

$$G_W \to 1, G_V \to 0, G_{WE} \to 0, G_{VE} \to 0. \tag{2.30}$$

Ve všech základních přenosech regulačního obvodu se vyskytuje součin přenosů regulované soustavy, regulátoru a měřicího členu, který je nazýván jako **přenos otevřeného regulačního obvodu** (ORO)

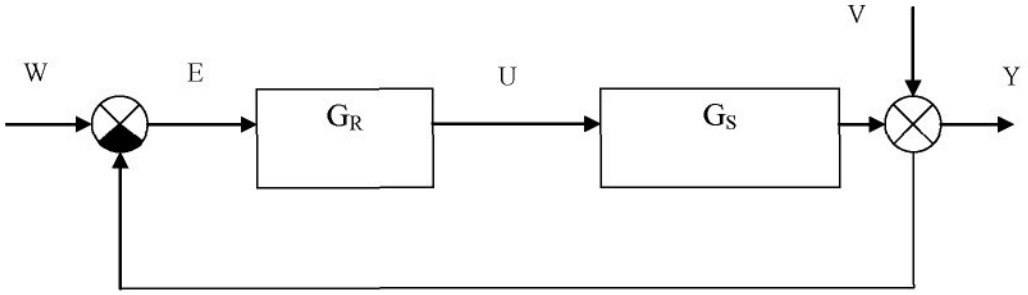
$$G_0 = G_S G_R G_{MČ}. \tag{2.31}$$

Ve všech základních přenosech regulačního obvodu je stejný výraz ve jmenovateli, který rozhoduje o stabilitě a nazývá se **charakteristický mnohočlen**. Pokud jej položíme roven nule, dostaneme **charakteristickou rovnici**. [ŠVARC 1992]

$$1 + G_0 = 0. \tag{2.32}$$

Příklad 2.13. *Základní přenosy regulačního obvodu*

Vypočítejte všechny základní přenosy regulačního obvodu dle schématu.



$$G_R = 5(1+0{,}25) \qquad G_S = \frac{2}{(s+0{,}3)^2}$$

Řešení:

Přenos řízení:

$$G_{wy} = \frac{G_R G_S}{1+G_R G_S} = \frac{(5+s)\dfrac{2}{(s+0{,}3)^2}}{1+(5+s)\dfrac{2}{(s+0{,}3)^2}} = \frac{2(5+s)}{(s+0{,}3)^2+2(5+s)} = \frac{10(0{,}2s+1)}{s^2+2{,}6s+10{,}09}$$

Přenos poruchy:

$$G_{vy} = \frac{1}{1+G_R G_S} = \frac{1}{1+(5+s)\dfrac{2}{(s+0{,}3)^2}} = \frac{(s+0{,}3)^2}{10+2s+(s+0{,}3)^2} = \frac{(s+0{,}3)^2}{s^2+2{,}6s+10{,}09}$$

Odchylkový přenos řízení:

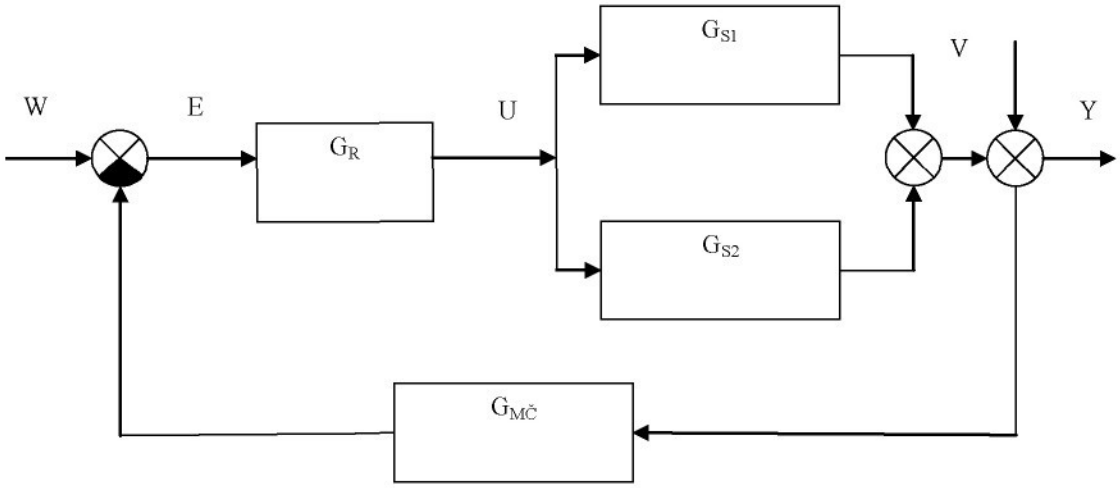
$G_{we} = G_{vy}$

Odchylkový přenos poruchy:

$G_{ve} = -G_{vy}$

Příklad 2.14. *Základní přenosy regulačního obvodu*

Vypočítejte všechny základní přenosy regulačního obvodu dle schématu.



Řešení:

Přenos regulované soustavy:

$$G_S = G_{S1} + G_{S2}$$

Přenos řízení:

$$G_{wy} = \frac{G_R G_S}{1 + G_R G_S G_{MČ}}$$

Přenos poruchy:

$$G_{vy} = \frac{1}{1 + G_R G_S G_{MČ}}$$

Odchylkový přenos řízení:

$$G_{we} = \frac{1}{1 + G_R G_S G_{MČ}}$$

Odchylkový přenos poruchy

$$G_{ve} = -\frac{G_{MČ}}{1 + G_R G_S G_{MČ}}$$

Shrnutí pojmů 2.6.

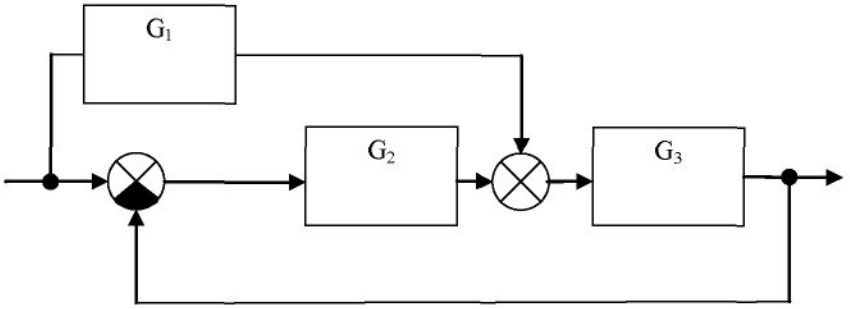
V této kapitole jsou studenti seznámeni s algebrou blokových schémat, se základními zapojeními bloků a výpočtem jejich výsledného přenosu. Pak se naučí zjednodušovat složité blokové schémata a počítat jejich výsledný přenos. Závěr je věnován popisu struktury regulačních obvodů a způsobu výpočtu jejich základních přenosů.

Otázky 2.6.

1. Jak se spočítá náhradní přenos paralelně zapojených soustav?
2. Co je to zesílení otevřené smyčky a jak se vypočte?
3. Jak se vypočítá odchylkový přenos řízení?
4. Co obsahují číslicové regulační obvody navíc než analogové?

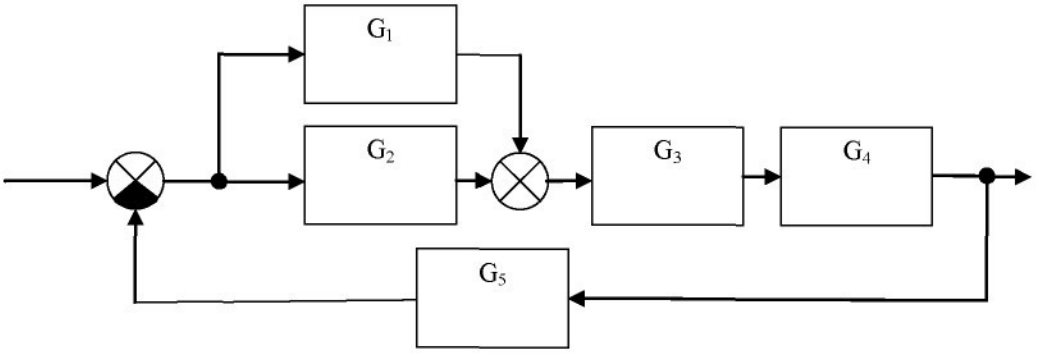
Úlohy k řešení 2.2.

1. Zjednodušte obvod a napište výsledný přenos.



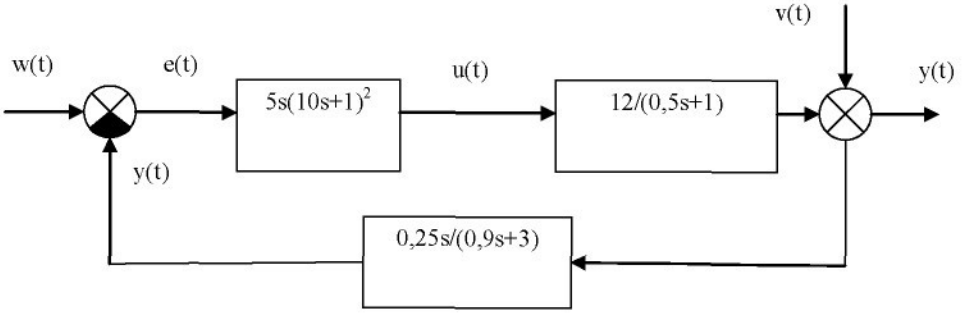
1.

2. Zjednodušte obvod a napište výsledný přenos.



1.

3. Vypočítejte přenos řízení a odchylkový přenos řízení.



1.

2.7. Stabilita

Čas ke studiu: 2 hodiny

Cíl Po prostudování tohoto odstavce budete umět

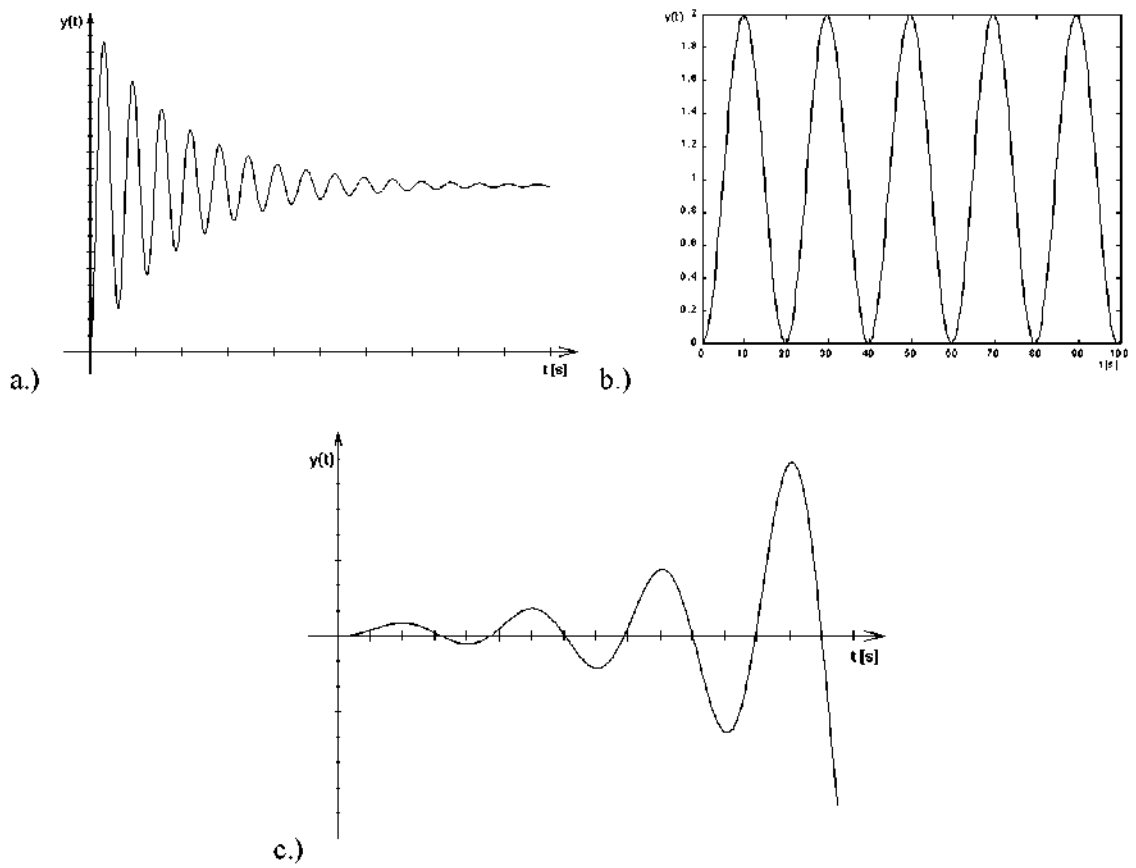
- definovat pojem stabilita spojitých systémů, nutnou a postačující podmínku stability, nutnou podmínku stability.
- popsat kritéria stability a jejich použití.
- Určit, zda regulační obvod je stabilní.

Výklad

Stabilita je jedním ze základních požadavků, které klademe na regulační obvod. Regulační obvod je stabilní, jestliže pro vychýlení regulačního obvodu z rovnovážného stavu a odeznění vnějších sil, které tuto odchylku způsobily, se regulační obvod během času znovu vrátí do původního rovnovážného stavu. Jinak řečeno je stabilita vlastnost regulačního obvodu udržet se v okolí rovnovážného stavu nebo se do něj vrátit po odeznění vnějších **působících sil.** Matematicky lze stabilitu definovat

$$\lim_{t\to\infty} y(t) = 0. \tag{2.33}$$

Z hlediska stability rozlišujeme RO stabilní, na mezi stability a nestabilní. RO na mezi stability se považují za stabilní. Průběh regulačního výše uvedené typy RO je znázorněn na obr. 2.81. Vždy se vyžaduje, aby regulační obvod byl za všech okolností stabilní. Zatímco parametry a dynamické vlastnosti regulované soustavy jsou dány konstrukcí soustavy, technologickým procesem apod. a nemůžeme je tudíž měnit, můžeme měnit dynamické vlastnosti regulátoru nastavováním volitelných parametrů regulátoru. Tím můžeme dosáhnout stability (a dalších vlastností) regulačního obvodu.



a.) stabilní, b.) mez stability, c.) nestabilní

Obrázek 2.81 - Příklady průběhů regulačního pochodu – shora: stabilní, na mezi stability, nestabilní RO

Při vyšetřování stability se vychází z tzv. charakteristického mnohočlenu, který je ve jmenovateli všech základní přenosů RO a je popsán vztahem (2.30). Má-li charakteristická rovnice (2.30) pouze nenásobené reálné kořeny s_i, je řešení rovnice (2.31) – tedy hledaný přechodný děj – ve tvaru

$$y = c_1 e^{s_1 t} + c_2 e^{s_2 t} + \ldots\ldots + c_n e^{s_n t} = \sum_{i=1}^{n} c_1 e^{s_i t}, \tag{2.34}$$

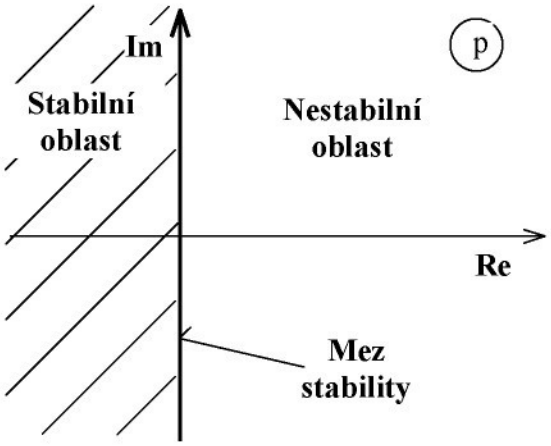
kde c_i jsou konstanty vypočtené z počátečních podmínek.

Jestliže má charakteristická rovnice (2.30) dvojici komplexně sdružených kořenů, $s_k = a + jb$, $s_{k+1} = a - jb$, bude řešení rovnice (2.31) mít tvar

$$y = \ldots\ldots + e^{at}(c_k \sin bt + c_{k+1} \cos bt) + \ldots\ldots \tag{2.35}$$

Ze vztahů (2.32), (2.33) vyplývá nutná a postačující podmínka stability RO, která se týká kořenů charakteristické rovnice (2.30). Regulační obvod je stabilní právě tehdy, když všechny kořeny charakteristické rovnice mají záporné reálné části, tedy leží-**li v levé komplexní polorovině**, viz obr. 2.82, [ŠVARC 1992]

$$\mathrm{Re}\, s_i \leq 0 \qquad \text{pro} \qquad i = 1,\ldots,n. \tag{2.36}$$



Obrázek 2.82 - Oblast stability spojitých RO

Kontrola stability RO spočívá v určení rozložení kořenů charakteristické rovnice v komplexní rovině kořenů. Pokud lze kořeny vyčíslit, použijeme nutnou a postačující podmínku stability. Jinak je nutno použít pravidla, která umožní rozhodnout o stabilitě bez přímého výpočtu kořenů, tyto pravidla se nazývají kritéria stability. V následujících odstavcích jsou uvedeny kritéria, která se často používají.

❑ Hurwitzovo kriterium stability

Vycházíme z charakteristické rovnice (3.8) a platí zde tzv. Stodolova nutná podmínka stability, která zní: „Všechny koeficienty charakteristické rovnice a_i musí existovat a musí mít stejné znaménko“. Je-li charakteristický mnohočlen řádu $n \le 2$, Stodolova podmínka přechází v nutnou a postačující podmínku stability. Jinak je nutno sestrojit z koeficientů charakteristické rovnice Hurwitzův determinant *n*-tého stupně ve tvaru

$$H_{n-1} = \begin{vmatrix} a_{n-1} & a_{n-3} & a_{n-5} & \cdots & 0 \\ a_{n-1} & a_{n-2} & a_{n-4} & \cdots & 0 \\ 0 & a_{n-1} & a_{n-3} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_0 \end{vmatrix} \qquad (2.37)$$

Hlavní rohové subdeterminanty jsou

$$H_1 = a_{n-1}; H_2 = \begin{vmatrix} a_{n-1} & a_{n-3} \\ a_n & a_{n-2} \end{vmatrix}; H_3 = \begin{vmatrix} a_{n-1} & a_{n-3} & a_{n-5} \\ a_n & a_{n-2} & a_{n-4} \\ 0 & a_{n-1} & a_{n-3} \end{vmatrix}; \ \ldots; H_n = a_0 H_{n-1}. \qquad (2.38)$$

Podle Hurwitzova kriteria je RO stabilní právě tehdy, když hlavní subdeterminanty (7.6) jsou kladné,

$$H_i > 0 \quad pro \quad i = 2,\ldots,n-1. \qquad (2.39)$$

První a poslední rohový subdeterminant se počítat nemusí, protože je-li splněna Stodolova podmínka, jsou dané subdeterminanty vždy kladné. Pokud je jeden z rohových

subdeterminantů roven nule, RO je na **kmitavé mezi stability**, tzn. charakteristická rovnice má dvojici ryze imaginárních kořenů. Je-li koeficient charakteristické rovnice a_0=0 a všechny rohové subdeterminanty jsou kladné, RO je na **nekmitavé mezi stability**, tzn. charakteristická rovnice má nulový kořen.

Příklad 2.15. *Stabilita*

Pomocí Hurwitzova kriteria stability určete, zda je obvod stabilní.

$$G_O(s) = \frac{4s^2 + 1}{10s^4 + 3s^3 + 14s^2}$$

Řešení:

V zadání je přenos otevřeného regulačního obvodu, a proto charakteristický mnohočlen je součet čitatele a jmenovatele.

$$N(s) = 10s^4 + 3s^3 + 18s^2 + 1$$

$a_i > 0$ splněno

$$H = \begin{vmatrix} 3 & 1 & 0 \\ 10 & 18 & 0 \\ 0 & 3 & 1 \end{vmatrix} \qquad H_2 = \begin{vmatrix} 3 & 1 \\ 10 & 18 \end{vmatrix} = 54 - 10 = 44 > 0$$

Regulační obvod je stabilní.

Příklad 2.16. *Stabilita*

Rozhodněte, zda RO je stabilní, když přenos regulované soustavy a regulátoru mají tvar

$$G_S(s) = \frac{1}{0.3s^3 + s^2 + 2s + 1}; G_R(s) = 1 + \frac{1}{3s}.$$

Řešení:

Přenos otevřeného RO je

$$G_0(s) = \frac{3s + 1}{0{,}9s^4 + 3s^3 + 6s^2 + 6s + 1}.$$

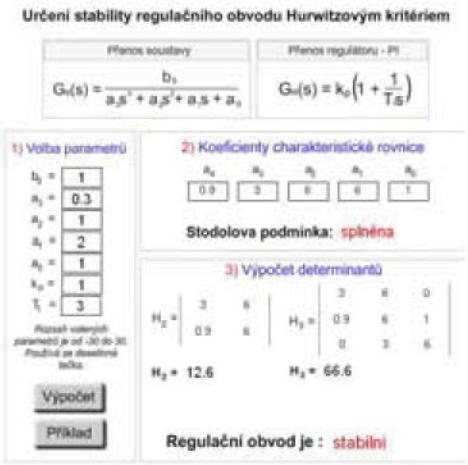
Pak charakteristická rovnice má tvar

$$0{,}9s^4 + 3s^3 + 6s^2 + 9s + 2 = 0.$$

Nutná Stodolova podmínka je splněna, musíme sestavit tedy determinant H_4 a subdeterminanty H_2 a H_3 vyčíslit.

$$H_4 = \begin{vmatrix} 3 & 9 & 0 & 0 \\ 0{,}9 & 6 & 2 & 0 \\ 0 & 3 & 9 & 0 \\ 0 & 0{,}9 & 6 & 2 \end{vmatrix} \Rightarrow H_2 = \begin{vmatrix} 3 & 9 \\ 0{,}9 & 6 \end{vmatrix} = 9{,}9;\ H_3 = \begin{vmatrix} 3 & 9 & 0 \\ 0{,}9 & 6 & 2 \\ 0 & 3 & 9 \end{vmatrix} = 71{,}1\,.$$

Oba subdeterminanty jsou kladné a RO je tedy stabilní.



Obrázek 2.83 – Animace kontroly stability pomocí Hurwitzova kritéria stability

❑ Routhovo-Schurovo kriterium stability

Vycházíme opět z charakteristické rovnice ve tvaru (3.8). Podle daného algoritmu provádíme postupnou redukci charakteristické rovnice na rovnici nižšího stupně, až se dostaneme ke kvadratické rovnici. Podle Routhova-Schurova kriteria je **RO stabilní právě tehdy, když při redukci se neobjeví záporné koeficienty a poslední tři koeficienty výsledné rovnice jsou kladné**. Koeficienty napíšeme v řadě od nejvyšších nebo nejnižších mocnin a podtrhneme sudé koeficienty v pořadí. Každý sudý koeficient v pořadí násobíme podílem prvních dvou a napíšeme pod předcházející řadu posunutý o člen vlevo. Tuto výslednou řadu, která má členy vždy ob jeden předcházející řady, od ní odečteme. Výsledná řada koeficientů je pak o první člen kratší (stupeň předchozí rovnice se snížil o jednotku). Stejným postupem pokračujeme dále, až zbudou koeficienty rovnice druhého stupně – tři koeficienty.
Výpočet lze znázornit schématem

$$\begin{array}{llllll} a_n & a_{n-1} & a_{n-2} & a_{n-3} & a_{n-4} & a_{n-5} \quad \cdots\cdots \Big/ \left(\frac{a_n}{a_{n-1}}\right) \\ \frac{a_n}{a_{n-1}}a_{n-1} & & \frac{a_n}{a_{n-1}}a_{n-3} & & \frac{a_n}{a_{n-1}}a_{n-5} & \\ \hline 0 & a_{n-1} & a_{n-2} - \frac{a_n}{a_{n-1}}a_{n-3} & a_{n-3} & a_{n-4} - \frac{a_n}{a_{n-1}}a_{n-5} & a_{n-5} \quad \cdots\cdots \end{array} \qquad (2.39)$$

Toto kritérium je účelné použít hlavně tehdy, jsou koeficienty charakteristické rovnice vyjádřené číselně. Stodolova podmínka zůstává v platnosti a je podmínkou nutnou.

Příklad 2.17. *Stabilita*

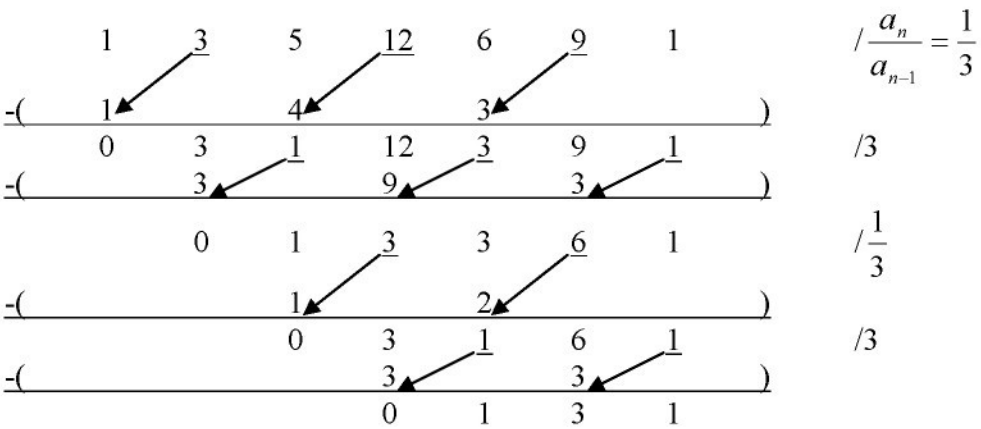
Pomocí Routhova-Schurova kriteria stability určete, zda je obvod stabilní.

$$N(s) = s^6 + 3s^5 + 5s^4 + 12s^3 + 6s^2 + 9s + 1$$

Řešení:

Routhovo-Schurovo kriterium opět vychází z charakteristického mnohočlenu. Regulační obvod je stabilní, když jsou koeficienty všech rovnic při postupné redukci charakteristické rovnice kladné.
Schéma redukce je následující:

- napíšeme koeficienty redukované rovnice do řádku od nejvyšší mocniny k nejnižší (možno i naopak)
- podtrhneme sudé koeficienty v pořadí (každý druhý)
- každý podtržený koeficient násobíme podílem dvou nejvyšších koeficientů a_n / a_{n-1} a výsledek napíšeme do druhého řádku posunutý o jedno místo vlevo
- druhý řádek (který má členy vždy ob jeden prvního řádku) odečteme od prvého řádku a dostaneme třetí řádek
- koeficienty třetího řádku jsou koeficienty rovnice o jeden stupeň nižší, než byla redukovaná rovnice, neboť na místě nejvyššího koeficientu jsme dostali nulu
- redukci provádíme tímto způsobem dále až na rovnici 2. stupně (tři koeficienty). Nulu na začátku řady koeficientů neuvažujeme. Koeficienty u všech redukovaných rovnic musí být kladné. To je podmínka stability.



Regulační obvod je stabilní, protože jsou koeficienty kladné.

Příklad 2.18. *Stabilita*

Pomocí Routhova-Schurova kriteria stability určete, zda je obvod stabilní.

$$G_w(s) = \frac{3s^2 + 2s + 1}{s^4 + 5s^3 + 13s^2 + 19s + 10}$$

Řešení:

V zadání je přenos řízení, a proto charakteristický mnohočlen je jen jmenovatel.

$N(s) = s^4 + 5s^3 + 13s^2 + 19s + 10$

	1	5	13	19	10
-(	1		19/5		)
	0	5	46/5	19	10
-(		5		250/46	)
	0	0	46/5	624/46	10

$/\frac{a_n}{a_{n-1}} = \frac{1}{5}$

$/\frac{35}{46}$

Regulační obvod je stabilní, protože jsou koeficienty kladné.

Příklad 2.19. *Stabilita*

K regulaci statické soustavy o přenosu

$$G_S(s) = \frac{1}{0{,}33s^5 + s^4 + 1{,}66s^3 + 4s^3 + 2s + 1},$$

je použit PI regulátor s přenosem

$$G_R(s) = 1 + \frac{1}{3s}.$$

Rozhodněte, zda RO je stabilní.

Řešení:

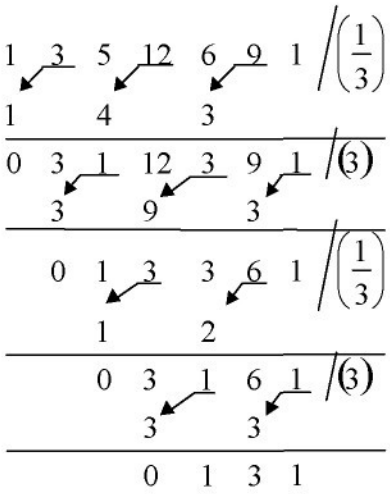
Přenos ORO je

$$G_{O(s)} = \frac{3s + 1}{s^6 + 3s^5 + 5s^4 + 12s^3 + 6s^2 + 6s}.$$

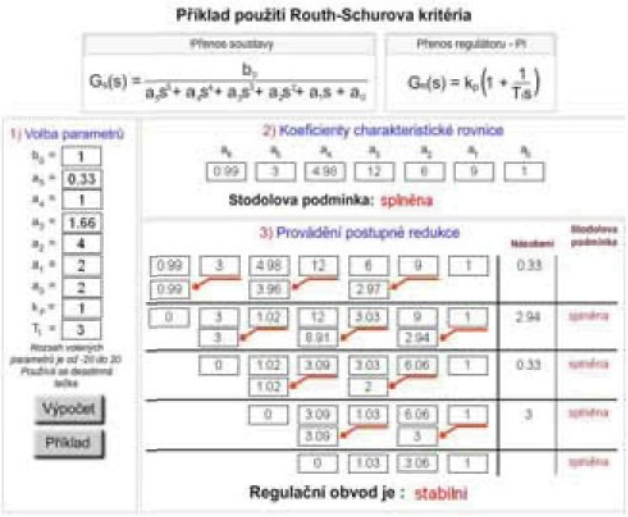
Charakteristická rovnice RO je

$$s^6 + 3s^5 + 5s^4 + 12s^3 + 6s^2 + 9s + 1 = 0$$

Podle algoritmu (6.8) provádíme nyní postupnou redukci stupně charakteristické rovnice:



Koeficienty u všech stupňů rovnic jsou kladné a proto je RO **stabilní**.



Obrázek 2.84 – Animace kontroly stability pomocí Routhova-Schurova kritéria stability

Předchozí dvě kritéria stability spadají do **algebraických kritérií stability**, tyto kritéria nedávají dosti dobrý názor o vlivu jednotlivých součinitelů na stabilitu regulačního pochodu a nemohly být použity pro řešení stability u systémů s dopravním zpožděním. Tyto chybějící vlastnosti nacházíme až u **kmitočtových kritérií stability**.

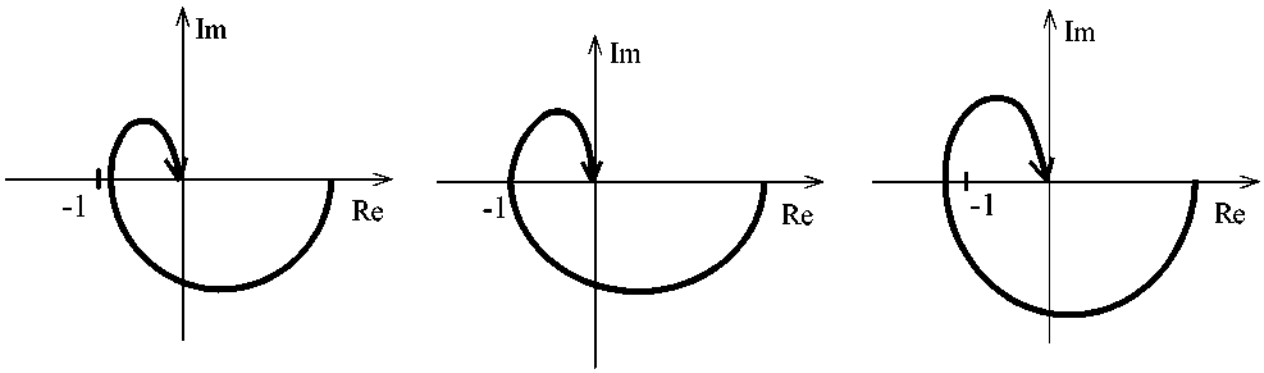
❑ Nyquistovo kriterium stability

Je to kmitočtové kriterium, které rozhoduje o stabilitě RO na základě průběhu kmitočtové charakteristiky otevřeného regulačního obvodu. Tuto charakteristiku lze určit analyticky nebo experimentálně. Skutečnost, že při vyšetřování stability pomocí Nyquistova kritéria lze vycházet z experimentálně získaných hodnot, zvýhodňuje toto kritérium v porovnání s ostatními kritérii stability. Další výhodou je, že ho lze použít i pro RO s dopravním zpožděním.

Charakteristická rovnice uzavřeného RO je dána vztahem

$$N(s) = 1 + G_o(s) = \frac{N_o(s) + M_o(s)}{N_o(s)} = 0 . \qquad (2.40)$$

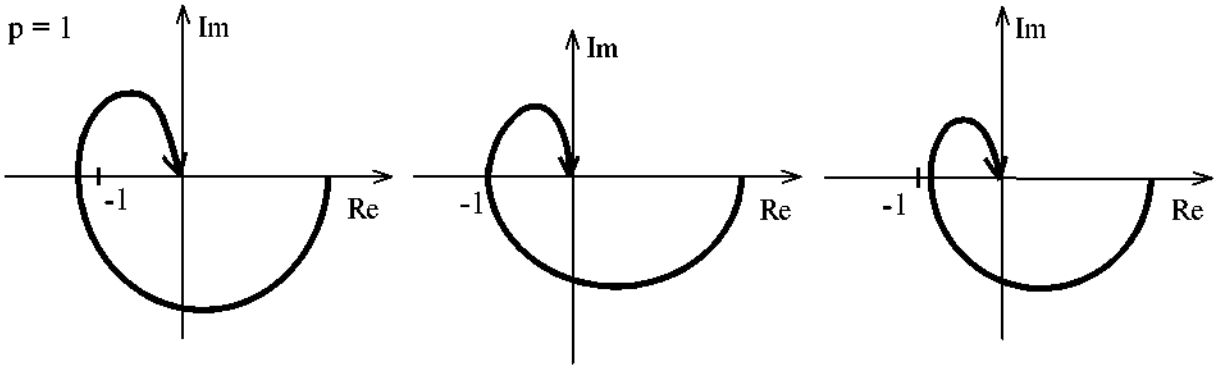
Nyquistovo kritérium stability: **„Je-li otevřený RO stabilní, pak uzavřený RO bude stabilní právě tehdy, když amplitudo-fázová kmitočtová charakteristika otevřeného regulačního obvodu $G_O(j\omega)$ neobklopuje kritický bod [-1, j0].“**



stabilní systém systém na mezi stability nestabilní systém

Obrázek 2.85 - Příklady průběhů amplitudo-fázové kmitočtové charakteristiky

Nyquistovo kritérium stability: **„Je-li otevřený RO nestabilní a má p nestabilních kořenů, pak uzavřený RO bude stabilní právě tehdy, když amplitudo-fázová kmitočtová charakteristika otevřeného RO obklopí kritický bod [-1, j0] p/2 krát.”**



stabilní systém systém na mezi stability nestabilní systém

Obrázek 2.86 - Příklady průběhů amplitudo-fázové kmitočtové charakteristiky

V případě, že určujeme stabilitu lineárních RO a otevřený RO je stabilní, můžeme použít **analytické formulace** Nyquistova kritéria. To spočívá v určení modulu a fáze kmitočtového přenosu otevřeného RO nebo v určení reálné a imaginární části kmitočtového přenosu otevřeného RO. Uzavřený RO bude stabilní právě tehdy, když

$$\arg G_o(j\omega_k) = -\pi \quad \Rightarrow \quad \mathrm{mod}\, G_o(j\omega_k) < 1, \qquad (2.41)$$

resp.

$$\arg G_o(j\omega_k) = -\pi \quad \Rightarrow \quad L(\omega_k) < 0 \qquad (2.42)$$

nebo

$$\mathrm{Im}\left\{G_o(j\omega_k)\right\}=0 \qquad \Rightarrow \qquad \mathrm{Re}\left\{G_o(j\omega_k)\right\}>-1. \tag{2.43}$$

Příklad 2.20. *Stabilita*

Vyšetřete stabilitu regulačního obvodu, jenž je tvořen regulovanou soustavou a regulátorem o přenosech

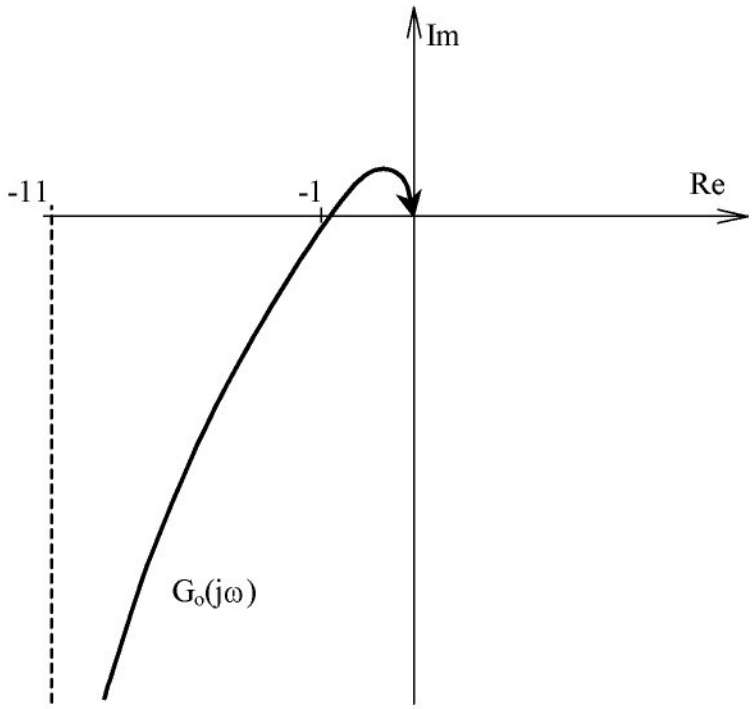
$$G_S(s)=\frac{1}{(10s+1)(s+1)}, G_R(s)=\frac{1}{s}.$$

Řešení:

Přenos otevřeného RO je

$$G_O(s)=\frac{1}{s(10s+1)(s+1)}.$$

Kořeny charakteristického mnohočlenu $N_o(s)$ jsou 0; -0,1; -1, tzn. otevřený RO je stabilní.



Obrázek 2.87 - Amplitudo-fázová KCH pro příklad

Na obr. 2.87 je nakreslena kmitočtová charakteristika rozpojeného obvodu $G_O(j\omega)$. Kritický bod –1 leží vlevo od kmitočtové charakteristiky $G_O(j\omega)$ a je proto obvod **stabilní.** [ŠVARC 1992], [BALÁTĚ 2003]

Shrnutí pojmů 2.7.

Definice stability a způsoby jejího ověřování buď na základě odezvy nebo pomocí nutné a postačující podmínky stability nebo pomocí kritérií stability.

Otázky 2.7.

1. Co to je stabilita?
2. S čím je svázáno jméno Nyquist?
3. Která kritéria stability patří mezi algebraická?
4. Jak určíme charakteristický mnohočlen regulačního obvodu popsaného přenosem otevřené smyčky?

Úlohy k řešení 2.3.

1. Ověřte, zda je daný regulační obvod stabilní.

$$G_R(s) = 4s + 1 \qquad G_S(s) = \frac{1}{5s(0{,}5s+1)(8s+1)}$$

2. Ověřte, zda je daný regulační obvod stabilní.

$$G_O(s) = \frac{2s^2 + 8}{s^4 + 3s^3 + 4s^2 + 12s}$$

3. Ověřte, zda je daný regulační obvod stabilní.

$$G_R(s) = k_p(1 + T_D s) \quad G_S(s) = \frac{2}{3s(0{,}5s+1)(3s+1)}$$

4. Ověřte, zda je daný regulační obvod stabilní.

$$G_R(s) = k_p\left(1 + \frac{1}{T_I s}\right) \quad G_S(s) = \frac{k_1}{s(T_1 s + 1)}$$

2.8. Regulátory

Čas ke studiu: 1,5 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat základní typy regulátorů – spojitých, číslicových, nespojitých.
- popsat přenosy regulátorů odvozených z PID regulátorů.
- nakreslit průběh regulované veličiny pro regulační obvod s nespojitým regulátorem.

Výklad

V této kapitole si uvedeme základní typy regulátorů, které se používají běžně v technické praxi. Jedná se o analogové, číslicové a nespojité regulátory.

Analogové regulátory

Analogový (spojitý) regulátor je dán obecným vztahem

$$u(t) = k_p\left[e(t) + \frac{1}{T_I}\int_0^t e(\tau)d\tau + T_D\frac{de(t)}{d(t)}\right], \tag{2.44}$$

kde k_p - zesílení regulátoru T_I - integrační časová konstanta regulátoru T_D. derivační časová konstanta regulátoru.
Přenos spojitého regulátoru je dán vztahem

$$G_R(s) = \frac{U(s)}{E(s)} = k_p\left(1 + \frac{1}{T_1 s} + T_D s\right). \tag{2.45}$$

Konvenční algoritmy regulace popsané vztahem (2.45) buď obsahují kombinaci dvou nebo tří složek, popř. jednotlivé složky samostatně. Tím vzniká pět základních typů regulátorů, viz. tabulka 2.3.

Tabulka 2.3 - Popis používaných typů spojitých regulátorů

Typ spojitého regulátoru	Přenos regulátoru
P	$G_R(s) = k_p$
I	$G_R(s) = \frac{1}{T_I s}$
PD	$G_R(s) = k_p(T_D s + 1)$
PI	$G_R(s) = k_p\left(1 + \frac{1}{T_I s}\right)$
PID	$G_R(s) = k_p\left(1 + \frac{1}{T_I s} + T_D s\right)$

P-reg - v uzavřeném obvodu pracuje s trvalou regulační odchylkou. Má dobrou stabilitu. Používá se velmi často např. na stabilizaci pevných bodů, stabilizaci napětí, proudu. Nezáleží na kvalitě.

I-reg - v uzavřeném obvodu pracuje pouze s přechodovou regulační odchylkou. Regulační pochod se ustálí tehdy, kdy regulační odchylka $e(t) = 0$. Nevyhoví podmínkám stability regulačního obvodu, když by měl regulovat astatickou regulovanou soustavu.

D-člen - není schopen samostatné funkce, jako regulátor připojený k regulované soustavě, protože vstupním signálem je derivace regulační odchylky a neví tedy nic o velikosti (hodnotě) odchylky $e(t)$. Připustí libovolně velkou ustálenou regulační odchylku. V kombinovaném regulátoru zlepšuje stabilitu regulačního obvodu. Natáčí fázi amplitudové fázové charakteristiky v komplexní rovině o +90°. Informuje regulátor o změně regulační odchylky a tedy regulátor může v "předstihu" na tuto změnu reagovat. Rozpojí regulační obvod

PI-reg - v uzavřeném regulačním obvodu odstraňuje trvalou regulační odchylku, kterou bychom měli při použití P-regulátoru. Zlepšuje stabilitu vzhledem k použití čistě I-regulátoru. Pro určitá nastavení stavitelných parametrů regulátoru vyhovuje z hlediska stability i pro astatické regulované soustavy. V počátku regulačního pochodu převládá vliv proporcionální složky, s narůstajícím časem převládá vliv integrační složky. Nejrozšířenější spojitý regulátor.

PD-reg - zlepšuje stabilitu regulačního obvodu ve srovnání s použitím čistě P-regulátoru. Je tedy možné pracovat s vyšším zesílením regulátoru a tedy menší trvalou regulační odchylkou vzhledem k použití čistě P-regulátoru. V počátku regulačního pochodu převládá vliv derivační složky, s narůstajícím časem převládá vliv proporcionální složky; regulátor pracuje s přechodným zvýšeným zesílením.

PID-reg - v uzavřeném regulačním obvodu odstraňuje vlivem I složky trvalou regulační odchylku a vlivem D složky zlepšuje stabilitu regulačního obvodu. V počátku přechodového děje převládá derivační složka regulátoru, s narůstajícím časem převládá integrační složka regulátoru [BALATĚ 1991].

Mechanické a pneumatické regulátory jsou drahé, elektrické jsou levné a většinou PID. Fyzikální realizace regulátorů a jejich charakteristiky jsou popsány v kapitole 3.

Číslicové regulátory

Číslicový (diskrétní) regulátor je dán obecným vztahem

$$u(kT) = k_p\left[e(kT) + \frac{T}{T_I}\sum_{i=0}^{k} e(kT) + \frac{T_D}{T}\nabla e(kT)\right], \tag{2.46}$$

kde $\nabla e(kT) = e(kT) - e[(k-1)T]$.

Přenos diskrétního regulátoru je dán vztahem

$$G_R(z) = \frac{U(z)}{E(z)} = k_p\left(1 + \frac{T}{T_I}\frac{z}{z-1} + \frac{T_D}{T}\frac{z-1}{z}\right). \tag{2.47}$$

Tabulka 2.4 - Popis jednotlivých diskrétních regulátorů

Typ diskrétního regulátoru	Přenos regulátoru
P	$G_R(z) = k_p$
S (číslicový I)	$G_R(z) = \frac{T}{T_I}\frac{z}{z-1}$
PD	$G_R(z) = k_p\left(1+\frac{T_D}{T}\frac{z-1}{z}\right)$
PS (číslicový PI)	$G_R(z) = k_p\left(1+\frac{T}{T_I}\frac{z}{z-1}\right)$
PSD (číslicový PI)	$G_R(z) = k_p\left(1+\frac{T}{T_I}\frac{z}{z-1}+\frac{T_D}{T}\frac{z-1}{z}\right)$

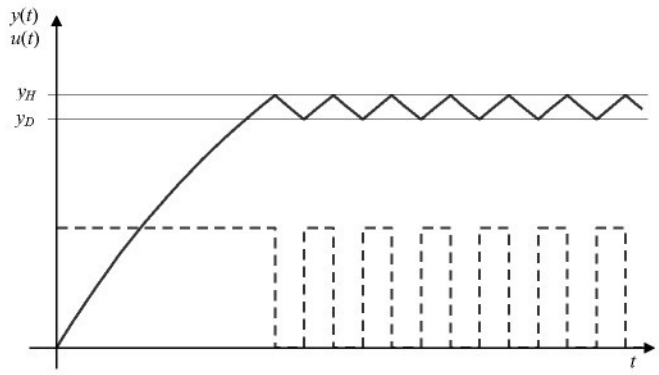
Pozn.: Diskrétní **regulátor S** nelze použít na astatické soustavy.

Nespojité regulátory

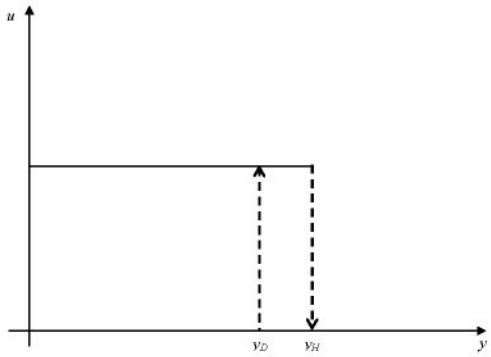
Nejjednodušší úlohy regulace spojitých proměnných veličin se velmi často realizují pomocí nespojitých regulátorů, nejčastější způsob je **dvoupolohová regulace**. To je zvláštní aplikace logického řízení. Hlavní důvod je jeho technická jednoduchost – nízká cena a dobrá spolehlivost. Dvoupolohový regulátor má pouze nastavení dvou hodnot, tj.

$$u(t) = \begin{cases} \text{zapnuto, v činnosti} \\ \text{vypnuto, zastaveno} \end{cases}.$$

Typický průběh regulované a akční veličiny při dvoupolohové regulaci je vidět na obr. Statická charakteristika dvoupolohového regulátoru je vidět na obr. Akční veličina nabývá pouze dvou hodnot a to maximální a nulovou. Přepínání mezi mezními hodnotami se řídí dvěmi hodnotami regulované veličiny, tzn spínací y_D a vypínací y_H. rozdíl mezi oběma hodnotami se nazývá hystereze a je stavitelným parametrem dvoupolohových regulátorů. Jak je vidět, regulovaná veličina se pohybuje v okolí žádané hodnoty, jehož velikost závisí jednak na dynamice řízeného systému, tak i na hysterezi regulátoru.

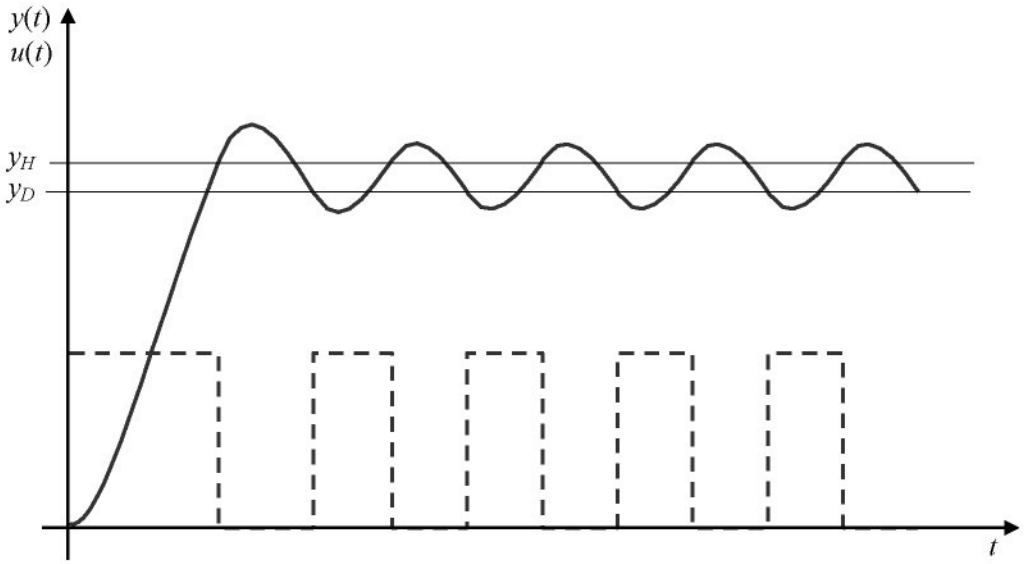


Obrázek 2.88 - Průběh regulované a akční veličiny při dvoupolohové regulaci



Obrázek 2.89 - Statická charakteristika dvoupolohového regulátoru

Pokud má regulovaná soustava dopravní zpoždění nebo se jedná a soustavu s vyššího řádu, pak lze předpokládat, že regulovaná veličina po přepnutí u nezmění směr tak náhle, jak je vidět v obr. , ale rozkmit regulované veličiny bude větší.



Obrázek 2.90 - Průběh regulované a akční veličiny při dvoupolohové regulaci regulované soustavy vyššího řádu

Shrnutí pojmů 2.8.

Kapitola je věnována základním typům regulátorů používaným v technické praxi – jedná se o analogové, číslicové a nespojité regulátory. Dále jsou zde uvedeny tvary jejich matematického popisu a charakteristik, realizace a stavitelné parametry, tvary přechodových charakteristik.

Otázky 2.8.

1. Existují regulátory typu PD?
2. Co je to pásmo proporcionality?
3. Jaké jsou stavitelné parametry PID regulátoru?
4. Napište zkratky pro regulátory odvozené od typu PID?
5. Existuje regulátor typu D?
6. Ustálí se regulovaná veličina v regulačním obvodu s nespojitým regulátorem?
7. Jak se realizují analogové regulátory?

2.9. Kvalita regulačního pochodu

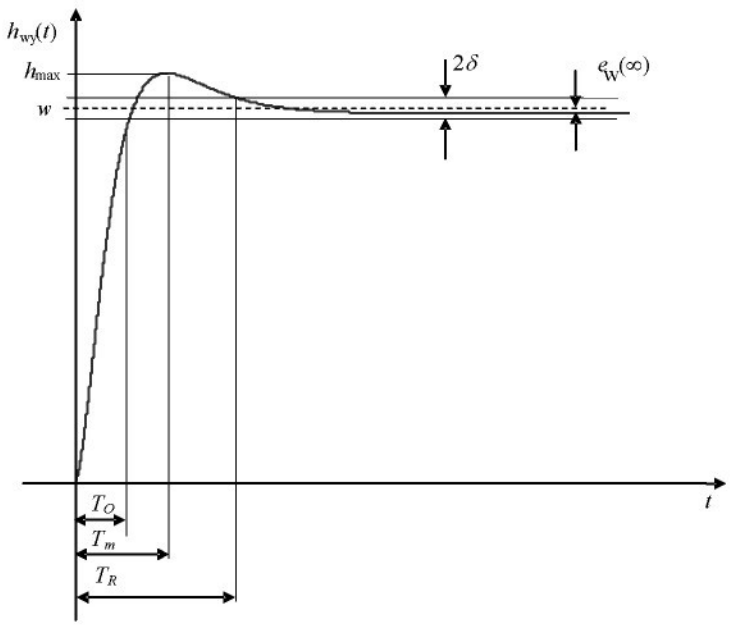
Čas ke studiu: 1 hodin

Cíl Po prostudování tohoto odstavce budete umět

- definovat integrační kritéria.
- popsat hodnocení kvality regulačního pochodu v časové oblasti.
- vypočítat trvalé regulační odchylky.

Výklad

Kvalitu regulace lze posuzovat ve třech oblastech: časové, kmitočtové a komplexní proměnné. Časová oblast je velmi oblíbená, protože umožňuje rychlé zhodnocení kvality regulace a základě průběhu odezvy regulované veličiny $y(t)$ vyvolané skokovou změnou žádané $w(t)$ nebo poruchové $v(t)$ veličiny, proto se budeme věnovat jen této oblasti.



Obrázek 2.91 - Průběh regulované veličiny na skokovou změnu žádané veličiny

Nejdříve se určuje, zda se jedná o kmitavý děj nebo nekmitavý, s překmitem nebo bez překmitu, viz. obrázek 2.91. U děje s překmitem jsou měřítkem kvality velikost relativního překmitu κ v procentech a rychlost odezvy T_0, která je určena dobou potřebnou k prvnímu dosažení ustálené hodnoty regulované veličiny. Doba regulace T_R je hodnota potřebná k dosažení 95% žádané hodnoty regulované veličiny. Dále se uvádí počet překmitů n za dobu regulace T_R.

Dalším důležitým parametrem kvality je velikost trvalé regulační odchylky. Určujeme trvalou regulační odchylku pro žádanou nebo poruchovou veličinu a pro tři základní testovací signály a to skok polohy, rychlosti a zrychlení. Odchylku lze velmi jednoduše spočítat pomocí následujících vztahů

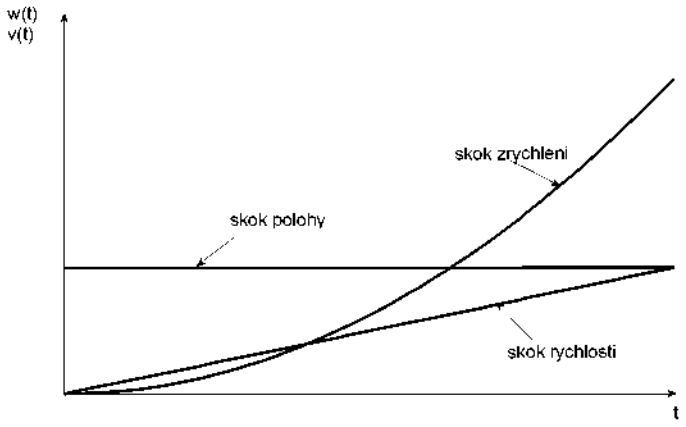
$$e_w(\infty) = \lim_{s\to 0} sE_w(s),\ e_v(\infty) = \lim_{s\to 0} sE_v(s), \tag{2.48}$$

$$E_w(s) = G_{we}(s)W(s),\ E_v(s) = G_{ve}(s)V(s). \tag{2.49}$$

Skok polohy $W(s) = \frac{w_0}{s},\ V(s) = \frac{v_0}{s}$. (2.50)

Skok rychlosti $W(s) = \frac{w_1}{s^2},\ V(s) = \frac{v_1}{s^2}$. (2.51)

Skok zrychlení $W(s) = \frac{2w_2}{s^3},\ V(s) = \frac{2v_2}{s^3}$. (2.52)



Obrázek 2.92 - Průběh testovacích signálů pro výpočet trvalých regulačních odchylek

Pro komplexní zhodnocení kvality regulačního pochodu jsou velmi vhodná integrální kritéria, která udávají velikost regulační plochy. Čím bude regulační plocha menší, tím bude kvalita větší. Nejčastěji se používají následující integrální kritéria.
Lineární regulační plocha (IE – Integral Error) – pouze pro pochody bez překmitu, lze řešit analyticky.

$$I_{IE} = \int_0^\infty e(t)dt$$

Absolutní regulační plocha (IAE – Integral Absolute Error) – vhodná i pro pochody s překmitem, nelze řešit analyticky.

$$I_{IAE} = \int_0^\infty |e(t)|dt$$

Kvadratická regulační plocha (ISE – Integral Square Error) – vhodná pro pochody s překmitem, lze řešit analyticky, může nabývat vysokých hodnot

.

$$I_{ISE} = \int_0^\infty e^2(t)dt$$

ITAE - Integral Time Absolute Error – zohledňuje dobu regulace.

$$I_{ITAE} = \int_0^\infty t|e(t)|dt$$

Příklad 2.21. *Výpočet trvalých regulačních odchylek*

Vypočítejte trvalé regulační odchylky řízení pro skok polohy, rychlosti a zrychlení.

$$G_R(s) = k_p(1 + \frac{1}{T_I s}) \quad G_S(s) = \frac{k_1}{(T_1 s + 1)(T_2 s + 1)}.$$

Řešení:

$$e_w(\infty) = \lim_{s\to 0} sE_w(s)$$

$$E_w(s) = G_{we}(s)W(s)$$

$$G_{we}(s) = \frac{1}{1 + G_R G_S} = \frac{1}{1 + \frac{k_1 k_p (T_I s + 1)}{T_I s(T_1 s + 1)(T_2 s + 1)}} = \frac{T_I s(T_1 s + 1)(T_2 s + 1)}{T_I s(T_1 s + 1)(T_2 s + 1) + k_1 k_p (T_I s + 1)}$$

Pro skok polohy

$$e_w(\infty) = \lim_{s\to 0} sG_{we}(s)W(s) = \lim_{s\to 0}\left[s\frac{T_I s(T_1 s + 1)(T_2 s + 1)}{T_I s(T_1 s + 1)(T_2 s + 1) + k_1 k_p (T_I s + 1)}\frac{W_0}{s}\right] = 0$$

Pro skok rychlosti

$$e_w(\infty) = \lim_{s\to 0} sG_{we}(s)W(s) = \lim_{s\to 0}\left[s\frac{T_I s(T_1 s + 1)(T_2 s + 1)}{T_I s(T_1 s + 1)(T_2 s + 1) + k_1 k_p (T_I s + 1)}\frac{W_1}{s^2}\right] = \frac{T_I W_1}{k_1 k_r}$$

Pro skok zrychlení

$$e_w(\infty) = \lim_{s\to 0} sG_{we}(s)W(s) = \lim_{s\to 0}\left[s\frac{T_I s(T_1 s + 1)(T_2 s + 1)}{T_I s(T_1 s + 1)(T_2 s + 1) + k_1 k_p (T_I s + 1)}\frac{2W_2}{s^3}\right] = \infty$$

Shrnutí pojmů 2.9.

Kapitola se věnuje popisu kvality regulačního pochodu v časové oblasti a to jak pomocí průběhu odezvy regulačního pochodu, tak pomocí integrálních kritérií.

Otázky 2.9.

1. Vznikne trvalá regulační odchylka v regulačním obvodu bez integrátoru a se skokem polohy žádané hodnoty?
2. K čemu slouží integrální kritéria?
3. Jak se určí doba regulace?

Úlohy k řešení 2.4.

1. Vypočítejte trvalé regulační odchylky řízení pro skok polohy.
   $G_R(s) = k_p(1 + T_D s + \frac{1}{T_I s})$ $G_S(s) = \frac{3k_1}{s(T_1 s + 1)^2}$
2. Vypočítejte trvalé regulační odchylky poruchy pro skok polohy.
   $G_R(s) = k_p(1 + T_D s)$ $G_S(s) = 3(T_1 s + 4) + T_1 T_2 + 8$

2.10. Syntéza regulačních obvodů

Čas ke studiu: 2 hodin

Cíl Po prostudování tohoto odstavce budete umět

- definovat pojem syntéza regulačního obvodu.
- popsat čtyři metody syntézy regulačních obvodů (Zieglerovy – Nicholsovy metody, metodu požadovaného modelu).
- popsat způsoby identifikace naměřených přechodových charakteristik.
- seřídit regulátor metodou kritických parametrů.

Výklad

Syntézou regulačních obvodů rozumíme návrh struktury regulačního obvodu a jeho parametrů tak, aby obvod splňoval požadavky kladené na regulační pochod. Do syntézy zahrnujeme pouze teoretický návrh regulačního obvodu a nikoliv i následnou volbu technické realizace návrhu. Při návrhu regulačního obvodu však vycházíme z provozních podmínek (např. rozměry a hmotnost zařízení, pracovní prostředí, režim provozu apod.), kladených na regulační obvod.

Můžeme se setkat s těmito výchozími předpoklady pro syntézu:

1. Můžeme libovolně měnit strukturu i parametry regulačního obvodu a jsme omezeni jen splněním podmínek fyzikální realizovatelnosti.
2. Je zadaná část struktury i část parametrů regulačního obvodu.
3. Struktura obvodu je plně zadána a jsou zadány i některé parametry regulačního obvodu.

První případ se v praxi vyskytuje pouze ojediněle, většina regulačních obvodů patří pod bod 3. Jsou to obvody, které lze rozdělit na regulovanou soustavu a regulátor. Úloha syntézy se zde redukuje pouze na určení stavitelných parametrů regulátoru. Vhodnost správné volby typu regulátoru a jeho parametrů můžeme ověřit jednak simulací na matematickém modelu navrženého regulačního obvodu a jednak provozními zkouškami regulovaného objektu přímo v provozu.

Pod 2. bod patří regulační obvody, u kterých nelze provést rozdělení na regulovanou soustavu a regulátor. Jsou to např. servomechanismy, regulační obvody sloužící k vlečné regulaci polohy a rychlosti. U těchto regulačních obvodů navrhujeme částečně jak jejich strukturu, tak i jejich parametry.

Při syntéze regulačního obvodu potřebujeme znát:

1. dynamické vlastnosti regulované soustavy (přenos, diferenciální rovnici přechodovou charakteristiku),
2. předpokládaný průběh řídící veličiny,
3. předpokládané průběhy poruchových veličin a místo jejich vstupu do regulované soustavy,
4. omezení akčních veličin (výstupní hodnota z regulátoru nemůže nabývat libovolně velké hodnoty),
5. požadavky na kvalitu regulace.

Pro zjednodušení výpočtu uvažujeme u bodu 2 a 3 jako vstupní veličiny funkce, jejichž matematické vyjádření je snadné a z odezvy regulačního na tyto funkce můžeme posoudit přesnost a kvalitu regulace. Nejčastěji používané funkce jsou jednotkový skok, jednotkový impuls a skok rychlosti a zrychlení vstupního průběhu.

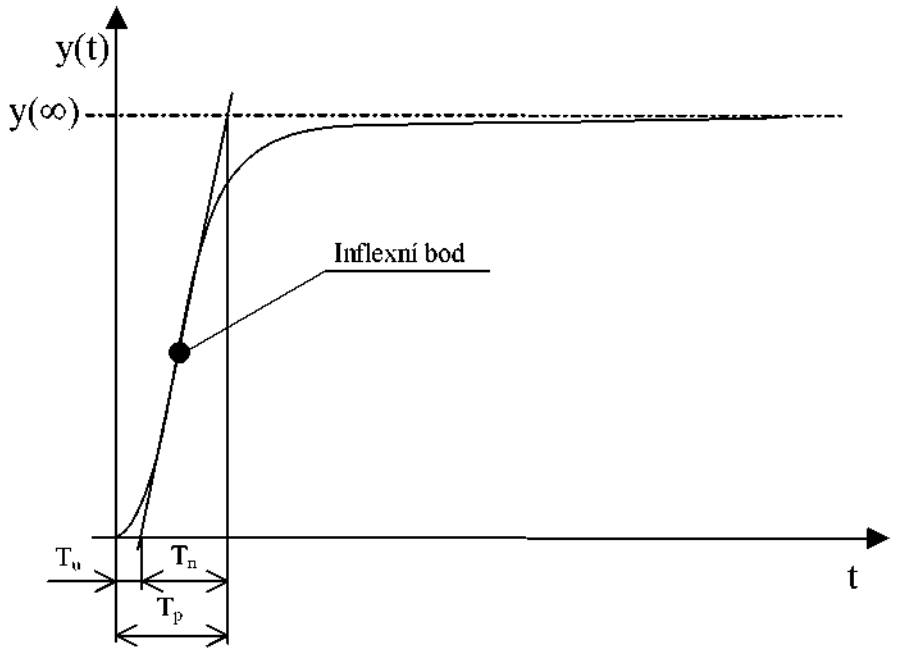
Ziegler-Nicholsova metoda přechodové charakteristiky (1.modifikace)

Tato metoda je použitelná pro lineární spojité i diskrétní regulační obvody. Předpokladem je aperiodická regulovaná soustava (pro regulované soustavy, které nemají přechodovou charakteristiku s překmitem). Vycházíme z odměřené přechodové charakteristiky regulované soustavy, kde určíme tyto parametry:

k_1 – zesílení přenosu proporcionální soustavy 1. řádu

T_u – doba průtahu ($T_u = T_d$ – dopravní zpoždění)

T_n – doba náběhu ($T_n = T_1$ – časová konstanta regulované soustavy)



Obrázek 2.93 - Určení parametrů přechodové charakteristiky regulované soustavy

Následující tabulka nám umožní určit optimální hodnoty stavitelných parametrů.

Tabulka 2.5 - Určení optimálních hodnot stavitelných parametrů regulátoru

Typ regulátoru	k_p^*	T_I^*	T_D^*
P	$\frac{T_n}{T_u \cdot k_1}$		
PI	$0{,}9\frac{T_n}{T_u \cdot k_1}$	$3{,}5\ T_u$	
PD	$1{,}2\ \frac{T_n}{T_u \cdot k_1}$		$0{,}25\ T_u$
PID	$1{,}25\frac{T_n}{T_u \cdot k_1}$	$2\ T_u$	$0{,}05\ T_u$

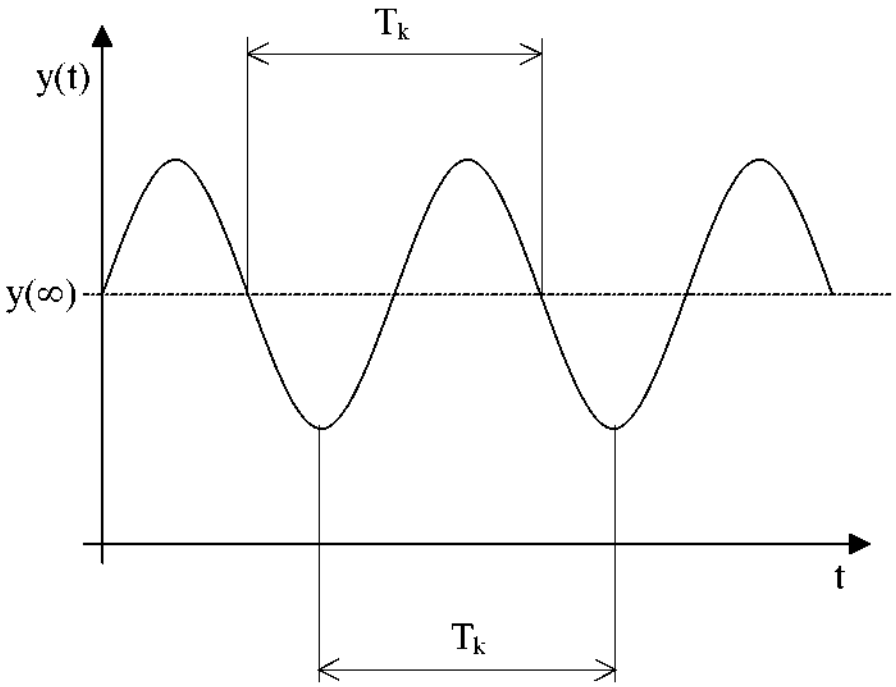
Ziegler-Nicholsova metoda kritických parametrů (2.modifikace)

Tato metoda je použitelná rovněž pro lineární spojité i diskrétní regulační obvody. Princip této metody spočívá v tom, že přivedeme regulační obvod do tzv. kritického stavu, tj. na kmitavou mez stability, přičemž regulátor pracuje pouze s proporcionální složkou a tedy integrační a derivační jsou vyřazeny nastavením $T_D = 0, T_I = \infty$.

a.) experimentální řešení

Postup pro seřizování regulátoru zapojeného do regulačního obvodu:

1. Vyřadíme integrační a derivační složku regulátoru. To provedeme nastavením integrační a derivační konstanty ($T_D \to 0, T_I \to \infty$).
2. Postupně zvyšujeme zesílení regulátoru k_p, až obvod začne kmitat s konstantní amplitudou (dosáhne kmitavé meze stability). Hodnotu zesílení regulátoru k_p, při kterém regulační obvod dosáhne meze stability, označíme jako kritické zesílení k_{pk}. Periodu kmitů, při které k tomu došlo, pak nazýváme periodu kritických kmitů $T = T_k$.



Obrázek 2.94 - Určení kritické periody

3. Konstanty regulátoru nastavíme s využitím hodnot k_{pk} a T_k viz tabulka 2.6.

Tabulka 2.6 - Optimální hodnoty stavitelných parametrů spojitého regulátoru

Typ regulátoru	k_p^*	T_I^*	T_D^*
P	0,5 k_{pk}		
PI	0,45 k_{pk}	0,83 T_k	
PD	0,4 k_{pk}		0,05 T_k
PID	0,6 k_{pk}	0,5 T_k	0,125 T_k

V případě, že se jedná o čistě integrační regulátor, přivedeme regulační obvod do kritického stavu tím, že zmenšujeme integrační časovou konstantu T_I, až dosáhneme kmitavé

meze stability. Potom je nejvhodnější nastavení integrační časové konstanty $T_I = 2T_K$ pro periodický tlumený průběh regulačního pochodu a $T_I = 4T_K$ pro aperiodický průběh.

b) analytické řešení

Při analytickém řešení musíme vypočítat k_{pk} a T_k, které charakterizují kmitavou mez stability. Využijeme Nyquistovo kritérium stability. Vyřadíme integrační a derivační složky regulátoru ($T_D \to 0, T_I \to \infty$) a určíme přenos otevřeného regulačního obvodu $G_O(s)$

$$G_o(s) = G_s(s).G_R(s) = \frac{M_o(s)}{N_o(s)}. \qquad (2.58)$$

Určíme reálnou a imaginární část kmitočtového přenosu otevřeného regulačního obvodu. Imaginární část položíme rovnu 0 a určíme kritický kmitočet

$$\mathrm{Im}\{G_0(j\omega)\} = 0 \qquad \Rightarrow \qquad \omega_k. \qquad (2.59)$$

Kritický kmitočet dosadíme do reálné části a ta musí být rovna hodnotě -1 a určíme kritické zesílení.

$$\mathrm{Re}\{G_0(j\omega_k)\} = -1 \qquad \Rightarrow \qquad k_{pk}. \qquad (2.60)$$

Periodu kritických kmitů T_k určíme z kritického kmitočtu dle vztahu

$$\omega_k = \frac{2\pi}{T_k}. \qquad (2.61)$$

Pro výpočet doporučeného seřízení příslušného regulátoru použijeme vztahy, které obsahuje tabulka 2.6.

Příklad 2.22. *Seřízení regulátorů.*

Pro regulovanou soustavu popsanou přenosem $G_S(s)$ seřiďte PID regulátor, použijte metodu kritických parametrů.

$$G_S(s) = \frac{2}{(s+1)^3}.$$

Řešení:

Nejdříve musíme vyřadit integrační a derivační složku regulátoru

$$T_I \to \infty, T_D = 0 \qquad \Rightarrow \qquad G_R(s) = k_p.$$

Teď určíme přenos otevřeného regulačního obvodu

$$G_0(s) = G_R G_S = \frac{2k_p}{(s+1)^3} = \frac{2k_p}{s^3 + 3s^2 + 3s + 1}$$

a vypočítáme kmitočtový přenos otevřeného regulačního obvodu a určíme jeho reálnou a imaginární část

$$G_0(j\omega) = G_0(s)\Big|_{s=j\omega} = \frac{2k_p}{-j\omega^3 - 3\omega^2 + 3j\omega + 1} = \frac{2k_p(1-3\omega^2) + 2k_p j\omega(\omega^2 - 3)}{(1-3\omega^2)^2 + \omega^2(\omega^2-3)^2},$$

$$\mathrm{Re}\{G_0(j\omega)\} = \frac{2k_p(1-3\omega^2)}{(1-3\omega^2)^2 + \omega^2(\omega^2-3)^2}, \mathrm{Im}\{G_0(j\omega)\} = \frac{2k_p\omega(\omega^2-3)}{(1-3\omega^2)^2 + \omega^2(\omega^2-3)^2}.$$

Teď určíme kritický kmitočet, periodu kritických kmitů a kritické zesílení:

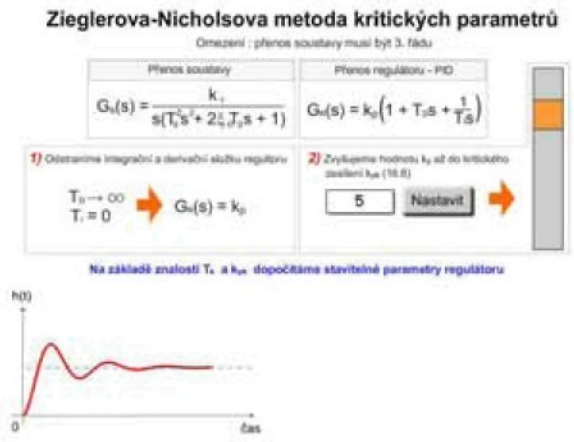
$$\mathrm{Im}\{G_0(j\omega)\} = 0 \quad \Rightarrow \quad 2k_p\omega(\omega^2-3) = 0 \quad \Rightarrow \quad \omega_k = \sqrt{3},$$

$$\mathrm{Re}\{G_0(j\omega_k)\} = -1 \quad \Rightarrow \quad \frac{2k_{pk}(1-3.3)}{(1-3.3)^2 + 3(3-3)^2} = -1 \quad \Rightarrow \quad k_{pk} = 4,$$

$$T_k = \frac{2\pi}{\omega_k} = \frac{2\pi\sqrt{3}}{3}.$$

Na základě tabulky 2.6 můžeme určit hodnoty stavitelných parametrů regulátoru:

$$k_p^* = 0{,}6k_{pk} = 2{,}4$$
$$T_I^* = 0{,}5T_k = 1{,}81$$
$$T_D^* = 0{,}125T_k = 0{,}43$$



Obrázek 2.95 – Animace seřízení RO pomocí metody Ziegler-Nichols

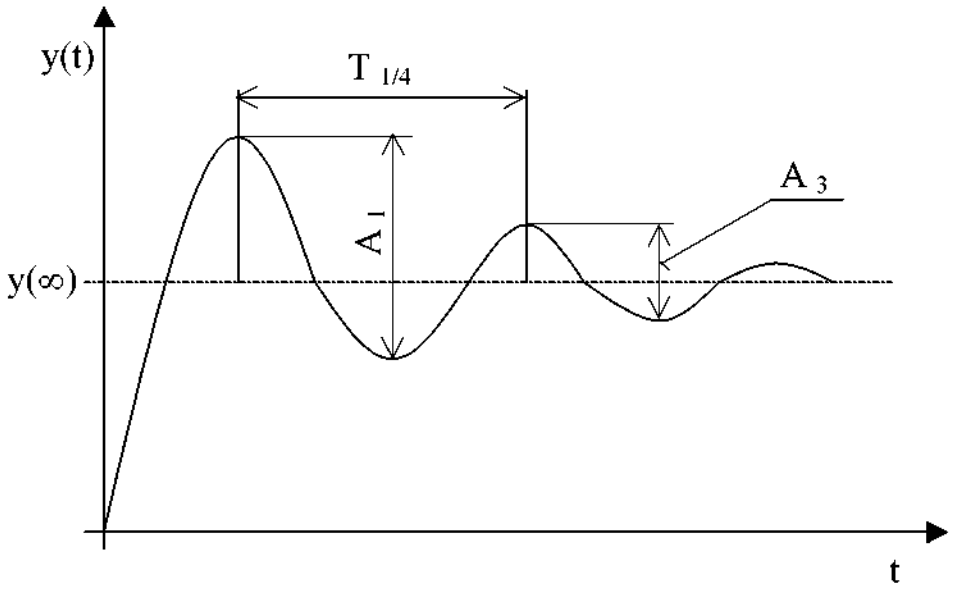
Ziegler-Nicholsova metoda čtvrtinového tlumení (3.modifikace)

Tato metoda je použitelná pro lineární spojité i diskrétní regulační obvody. Řeší pouze experimentálně, analyticky by se dala jen stěží využít. Používá se v případě, že nelze použít rozkmitání na kmitavou mez stability. Opět odstraníme integrační i derivační složky regulátoru ($T_D \to 0, T_I \to \infty$) a hodnotu k_p zvyšujeme tak dlouho, až průběh výstupní (regulované) veličiny bude ve tvaru

$$\frac{A_3}{A_1} = \frac{1}{4}, \tag{2.62}$$

kde A_1, A_3 -amplituda,

čili dojde ke čtvrtinovému tlumení, tzn. dvě po sobě jdoucí amplitudy v poměru 1:4.



Obrázek 2.96 - Určení periody tlumených kmitů $T_{1/4}$

Z průběhu regulované veličiny se stanoví hodnota čtvrtinové periody $T_{1/4}$, viz obrázek 2.96, a ze stupnice regulátoru čtvrtinové zesílení $k_{p1/4}$. Pro zvolený typ regulátoru se určí optimální hodnoty jeho stavitelných parametrů, viz . tabulka 2.7.

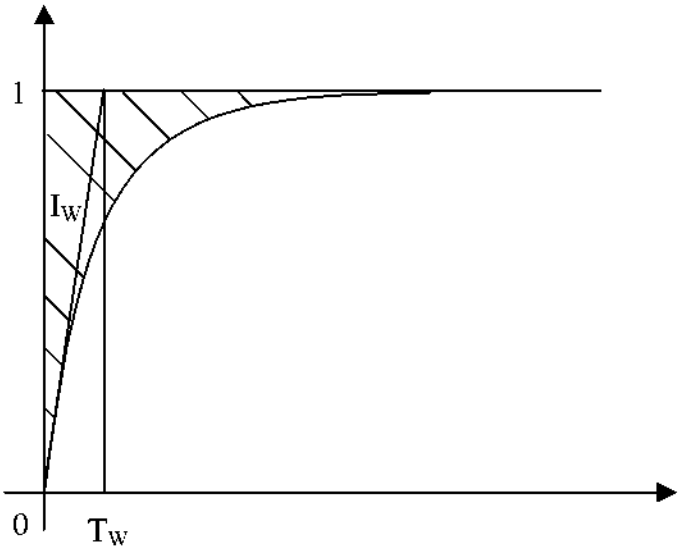
Tabulka 2.7 - Optimální hodnoty stavitelných parametrů regulátoru.

Typ regulátoru	k_p^*	T_I^*	T_D^*
P	$k_{p1/4}$		
PI	$0{,}9\ k_{p1/4}$	$T_{1/4}$	
PID	$1{,}2\ k_{p1/4}$	$T_{1/4}$	$0{,}25\ T_{1/4}$

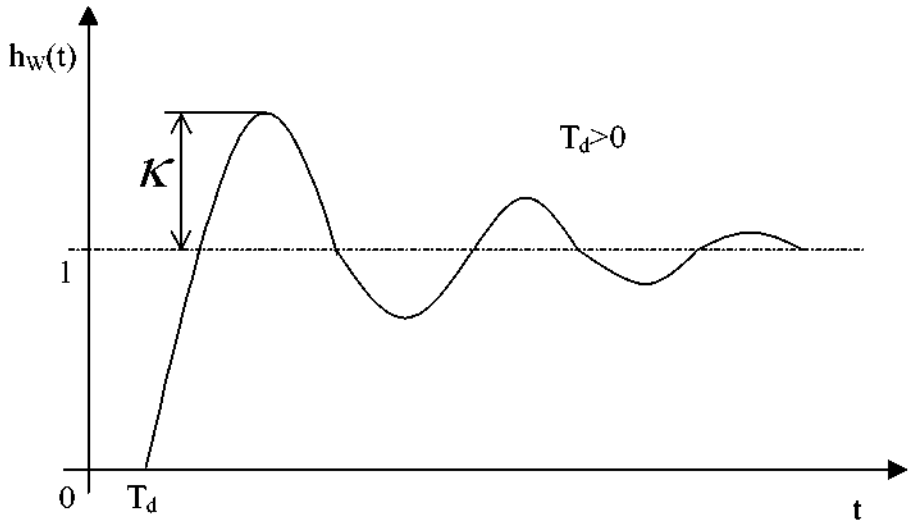
Metoda požadovaného modelu

Tato metoda je použitelná pro lineární spojité i diskrétní regulační obvody, umožňuje snadné a rychlé seřízení standardních typů analogových i číslových regulátorů pro základní typy regulovaných soustav i s dopravním zpožděním. Typ regulátoru je doporučen z hlediska vlastností regulované soustavy a požadavku na nulovou trvalou regulační odchylku

způsobenou skokovou změnou polohy žádané veličiny, resp. poruchy působící na výstupu regulované soustavy. [ŠULC&VÍTEČKOVÁ 2004], [VÍTEČKOVÁ&VÍTEČEK 2006]



Obrázek 2.97 - Předpokládaný průběh přechodové charakteristiky RO pro regulované soustavy bez dopravního zpoždění



Obrázek 2.98 - Předpokládaný průběh přechodové charakteristiky RO pro regulované soustavy s dopravního zpoždění

Pro regulované soustavy bez dopravního zpoždění $(T_d = 0)$ se předpokládá požadovaný aperiodický průběh přechodové charakteristiky $h_W(t)$ uzavřeného regulačního obvodu, viz. obrázek 2.97. Časová konstanta T_W musí být zvolena s ohledem na omezení akční veličiny a maximální nastavitelnou hodnotou zesílení regulátoru $k_{p\max}$ a u regulačního obvodu s číslicovým regulátorem musí platit $T_W > (2 \div 3)T$.

Pro regulované soustavy s dopravním zpožděním ($T_d > 0$) zvolíme požadovaný překmit $\kappa (0 \le \kappa \le 0{,}5)$ přechodové charakteristiky $h_W(t)$ uzavřeného regulačního obvodu podle hodnoty překmitu, viz tabulka 2.9, a určíme parametr a.

Podle tvaru přenosu regulované soustavy se určí doporučený typ regulátoru a odpovídající hodnoty jeho stavitelných parametrů, viz tabulka 2.8.

Tabulka 2.8– Parametry pro seřízení regulátorů

	Regulovaná soustava	Typ	k_p^*		T_I^*	T_D^*
			$T_d = 0$	$T_d > 0$		
1	$\frac{k_1}{s} e^{-T_d s}$	P	$\frac{2}{k_1(2T_W + T)}$	$\frac{a}{k_1}$	-	-
2	$\frac{k_1}{(T_1 s + 1)} e^{-T_d s}$	PI	$\frac{2T_I^*}{k_1(2T_W + T)}$	$\frac{aT_I^*}{k_1}$	$T_1 - \frac{T}{2}$	-
3	$\frac{k_1}{s(T_1 s + 1)} e^{-T_d s}$	PD	$\frac{2}{k_1(2T_W + T)}$	$\frac{a}{k_1}$	-	$T_1 - \frac{T}{2}$
4	$\frac{k_1}{(T_1 s + 1)(T_2 s + 1)} e^{-T_d s}$ $T_1 \ge T_2$	PID	$\frac{2T_I^*}{k_1(2T_W + T)}$	$\frac{aT_I^*}{k_1}$	$T_1 + T_2 - T$	$\frac{T_1 T_2}{T_1 + T_2} - \frac{T}{4}$
5	$\frac{k_1}{T_0^2 s^2 + 2\xi_0 s + 1} e^{-T_d s}$ $0 < \xi_0 \le 1$	PID	$\frac{2T_I^*}{k_1(2T_W + T)}$	$\frac{aT_I^*}{k_1}$	$2\xi_0 T_0 - T$	$\frac{T_0}{2\xi_0} - \frac{T}{4}$

Pozn. V případě analogových regulátorů T=0.

Tabulka 2.9 – Hodnoty koeficientů

κ	0	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
β	2,718	1,944	1,720	1,561	1,437	1,337	1,248	1,172	1,104	1,045	0,992
a	0,368	0,514	0,581	0,641	0,696	0,748	0,801	0,853	0,906	0,957	1,008

$$a = \frac{1}{\beta T_d + \alpha T}. \qquad (2.63)$$

Pro ideální proporcionální soustavu s dopravním zpožděním, jejíž obrazový přenos má tvar $\frac{k_1}{s} e^{-T_d s}$, je doporučený P-regulátor, kde

$$\text{pro} \quad T_d = 0 \qquad \rightarrow k_p^* = \frac{2}{k_1(2T_W + T)}, \qquad (2.64)$$

pro $T_d > 0$ $\rightarrow k_p^* = \frac{a}{k_1}$. (2.65)

❑ Jednoduché aproximační metody

Vzhledem k tomu, že pomocí metody požadovaného modelu lze seřídit regulátory pro regulované soustavy s obrazovým přenosem ve standardním tvaru, uvedeme si jednoduché metody pro aproximaci regulovaných soustav náhradními přenosy. Nejčastěji se používají dále uvedené postupy.

Způsob výpočtu

Z přechodové charakteristiky proporcionální regulované soustavy se určí poměrné hodnoty ustáleného stavu $y_{0,7} = 0,7y_\infty$, $y_{0,33} = 0,33y_\infty$ a odpovídající časy $t_{0,7}$ a $t_{0,33}$., viz obrázek 2.96. Dílčí náhradní zpoždění T_{d1}, T_{d2} a náhradní časová konstanta T_1 (popř. T_2) se vypočtou z některého z následujících vztahů.

1) Pro přenos $G(s) = \frac{k}{(T_1 s + 1)} e^{-T_{d1} s}$,

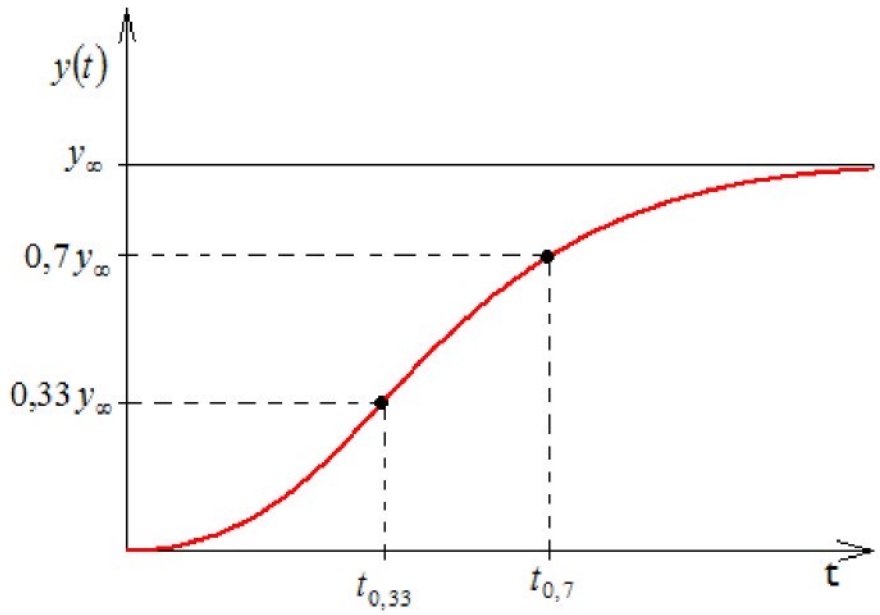
a) $T_{d1} = 1,498 t_{0,33} - 0,498 t_{0,73}$,

$T_1 = 1,245(t_{0,7} - t_{0,33})$,

2) Pro přenos $G(s) = \frac{k}{(T_1 s + 1)^2} e^{-T_{d2} s}$,

a) $T_{d2} = 1,937 t_{0,33} - 0,937 t_{0,7}$,

$T_1 = 0,794(t_{0,7} - t_{0,33})$,



Obrázek 2.96 - Aproximace pomocí dvou bodů odezvy

Pro regulované soustavy s obrazovým přenosem ve tvaru $\frac{1}{(T_n s+1)^n} e^{-T_{dn} s}$ je možno vypočítat přímo základní tvary podle převodní tabulky 2.10.

Tabulka 2.10 – Převodní tabulka

$\frac{1}{(T_n s+1)} e^{-T_{dn} s}$	n	1	2	3	4	5	6
$\frac{1}{T_1 s+1} e^{-T_{d1} s}$	$\frac{T_1}{T_n}$	1	1,568	1,980	2,320	2,615	2,881
	$\frac{T_{d1}-T_{dn}}{T_n}$	0	0,552	1,232	1,969	2,741	3,537
$\frac{1}{T_2 s+1} e^{-T_{d2} s}$	$\frac{T_2}{T_n}$	0,638	1	1,263	1,480	1,668	1,838
	$\frac{T_{d2}-T_{dn}}{T_n}$	- 0,352*	0	0,535	1,153	1,821	2,523

* - $T_{d1} \geq 0{,}352 T_1$

Další způsoby aproximace (např. náhrada malých časových konstant) jsou uvedeny v [VÍTEČKOVÁ&VÍTEČEK 2006.].

Shrnutí pojmů 2.10.

Kapitola je věnována syntéze regulačních obvodů a jsou zde popsány čtyři metody seřízení regulátorů a to tři modifikace Zieglerových-Nicholsových metod a metoda požadovaného modelu. Na závěr jsou popsány možnosti zjednodušení přenosů regulovaných soustav.

Otázky 2.10.

1. Jak se jmenuje metoda seřízení regulátoru, která byla odvozena na katedře ATŘ - 352?
2. Jaké označení má experimentální postup seřizování regulačních obvodů?
3. Co je to kritické zesílení?
4. Jak vypočteme vzorkovací periodu?

Úlohy k řešení 2.5.

1. Určete stavitelné parametry zadaného regulátoru.

$$G_S(s) = \frac{1}{(10s+1)(s+1)}$$ PID regulátor

2. Určete stavitelné parametry zadaného regulátoru.

$$G_S(s) = \frac{k_1}{s(T_0^2 s^2 + 2T_0\xi_0 s + 1)}$$ PID regulátor

3. Určete stavitelné parametry zadaného regulátoru.

$$G_S(s) = \frac{3}{(8s+1)^3}$$ PID regulátor

4. Určete stavitelné parametry zadaného regulátoru.

$$G_S(s) = \frac{1}{(10s+1)(s+1)(0{,}1s+1)}$$ PI regulátor

Další zdroje

[1] BALÁTĚ, J. 2003. *Automatické řízení*. Praha: Nakladatelství BEN, 2003, 654 s. ISBN 80-7300-020-2.

[2] BOLTON, W. 1992. *Control engineering*. New York: Longman scientific & Technical, 1992. 327s. ISBN 0-582-09729-0.

[3] DORF, R. C. & BISHOP, R. H. 1998: *Modern Control Systems*. Addison-Wesley : Harlow England 1998. ISNB 0-201-30864-9.

[4] FARANA, R. aj. 1996. *Programová podpora simulace dynamických systémů.*

Sbírka řešených příkladů. 1 vyd. Ostrava: KAKI 1996, 114 s. ISBN 80-02-01129-5.

[5] KATEDRA ATŘ. *Webový informační systém katedry*. Ostrava : Katedra ATŘ, VŠB-TU Ostrava. Dostupný z webu: <URL: http://www.352.vsb.cz>.

[6] KATEDRA MECHATRONIKY. *Webový informační systém katedry*. Liberec : Katedra MECHATRONIKY, TU Liberec. Dostupný z webu: <URL: http://www.fm.vslib.cz/>.

[7] ŠULC, B.&VÍTEČKOVÁ, M. 2004. *Teorie a praxe návrhu regulačních obvodů*. 1. vyd. Praha : Vydavatelství ČVUT, Praha, 2004. 333 s. ISBN 80-01-03007-5.

[8] ŠVARC, I. 1992. *Teorie automatického řízení I.* Brno: FS VUT, 1992. 210 s.

[9] ŠVARC, I. 2002. *Automatizace/Automatické řízení.* Brno: Nakladatelství CERM, 2002, ISBN 80-214-2087-1.

[10] VÍTEČKOVÁ, M. 2005. *Slovníky L- a Z- transformace s řešenými příklady*. Ostrava, KAKI 2005. 71 s. ISBN 80-248-0851-X.

[11] VÍTEČKOVÁ, M. 1989. POLOKOVÁ , J. *Logaritmické kmitočtové charakteristiky*. Ostrava: VŠB-TUO, 1989. 54s. Doplňkový učební text.

[12] VÍTEČKOVÁ, M.&VÍTEČEK, A. 2006. *Základy automatické regulace*. Ostrava: VŠB-TU Ostrava 2006. 198 s. ISBN 80-248-1068-9.

3 TECHNICKÉ PROSTŘEDKY ŘÍZENÍ

3.1. Základní pojmy

Čas ke studiu: 0,5 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- popsat přístrojovou skladbu regulačního obvodu – analogového, diskrétního.

Výklad

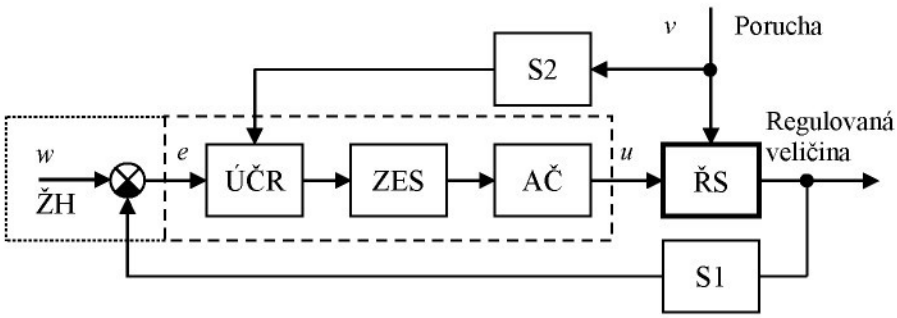
Tato kapitola se zabývá technickými prostředky pro logické, analogové a číslicové řízení. Doplňuje kapitoly pojednávající o teoretických základech vyjmenovaných způsobů řízení a regulaci.

Logické řízení a ovládání

Kapitola o logickém řízení obsahovala stručný popis základních elektronických prvků a složitých technických prostředků, jako jsou programové automaty. V této kapitole bude doplněna o prvky pneumatické a hydraulické, které lze také použit k logickému řízení. Jsou to například systémy elektropneumatického řízení pro různé výrobní a montážní linky.

Regulační obvod

Abstraktní rozdělení regulačního obvodu do dvou bloků představujících regulátor R a řízenou soustavu S lze z hlediska technického řešení dále podrobněji specifikovat. Podrobnější členění obvodu než je zmíněný matematický model, je znázorněno na obrázku 3.1. Toto schéma obsahuje prvky používané v analogovém regulačním obvodu. Na rozdíl od modelu jsou rovněž vyznačeny snímače regulované veličiny a poruchy.



Obrázek 3.1 – Regulační obvod se zpětnou vazbou

Analogový regulační obvod obsahuje mimo řízenou (regulovanou) soustavu ŘS tyto komponenty

- Regulátor
- S1 - snímač regulované veličiny y, který mění tuto veličinu na obvykle elektrický signál
- S2 - snímač poruchy v, který má stejnou funkci jako snímač S1

- Rozdílový člen pro výpočet regulační odchylky $e = w - y$
- ŽH – nastavení žádané hodnoty

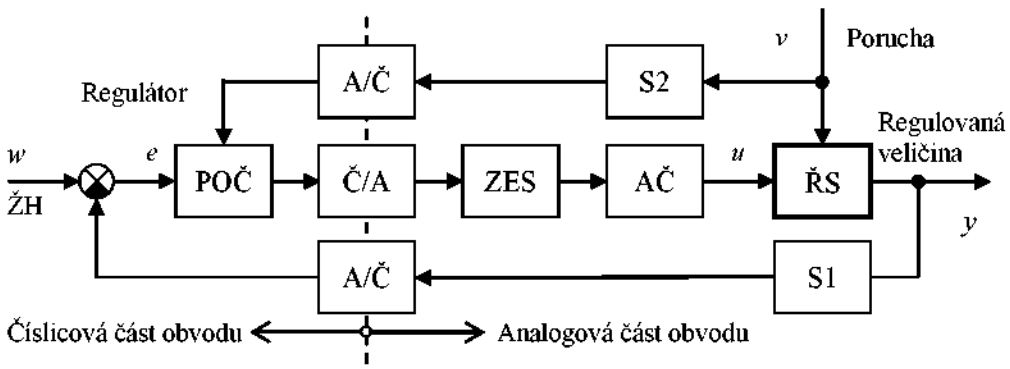
  Analogový regulátor se skládá z dílčích komponent
- ÚČR – ústřední člen regulátoru, který obsahuje obvody dynamických korekcí regulační odchylky
- ZES – zesilovač upravující výkonově výstup ÚČR tak, aby mohl ovládat akční člen
- AČ – akční člen určující akční veličinu u, která ovlivňuje řízenou soustavu.

Je třeba poznamenat, že rozdílový člen a zadání žádané hodnoty jsou součástí analogového regulátoru, na kterém jsou svorky pro signál regulované veličiny a poruchy a na jeho čelním panelu je nastavovací prvek (např. potenciometr, atd.). Obecně nemusí akční člen náležet k regulátoru, ale může být považován jako součást řízené soustavy. Pro některou variantu technického vybavení může splývat také zesilovač buď s akčním členem nebo ústředním členem regulátoru. Varianty tedy jsou

ÚČR + ZES = **R** a AČ + ŘS = **S** nebo ÚČR + ZES + AČ = **R** a ŘS = **S**.

Na příkladu analogového regulačního obvodu průtoku vzduchu bez měření poruchy si lze význam jednotlivých bloků vysvětlit. Průtok je snímán například diferenčním tlakem na cloně v potrubí s korekcemi na teplotu, tlak. Signál průtoku pro regulátor, který je výstupem snímače S1, je připojen k regulátoru. Jak již bylo řečeno, regulátor obsahuje rozdílový člen a zadávací prvek pro průtok. Odchylka je vstupem do dynamického korekčního členu, který tvoří nízkovýkonové elektronické prvky. Výstup tohoto členu je třeba zesílit tak, aby tento signál mohl ovládat servomotor pro natáčení clony v potrubí. Komplet servomotoru, převodu na páku spojenou se clonou lze považovat za akční člen.

Diskrétní regulační obvody obsahují navíc oproti analogovému obvodu převodníky, které zajišťují převody analogových veličin na číslicové a zpět číslicové veličiny na analogové.



Obrázek 3.2 – Diskrétní regulační obvod se zpětnou vazbou

Blokové schéma diskrétního regulačního obvodu je znázorněno na obrázku 3.2. Funkci ústředního členu regulátoru plní číslicový počítač POČ. Tento prvek může být rozdílných vlastností nebo výpočetního výkonu a kapacity paměti. Řešením pro nejjednodušší úlohy je mikropočítač třídy PIC. Složitější úlohy jsou řešitelné pomocí PLC, které obvykle podporují také zpětnovazební řízení pro malý počet regulačních obvodů. Nejsložitější úlohy lze řešit pomocí PC v průmyslovém provedení.

Funkce převodníků A/Č a Č/A byla popsána v částí skript popisujících programové logické automaty (PLC). Převodník Č/A plní současně funkci tvarovače nultého řádu, tj. paměti analogové veličiny mezi okamžiky výpočtu velikosti akční veličiny.

Ostatní prvky číslicového regulačního obvodu plní stejné funkce jako u analogového obvodu.

Obecné rozdělení technických prostředků zpětnovazebního řízení

Technické prostředky budou proto rozděleny na

- snímače získávající z vnějšího prostředí řídicího systému informace o velikosti různých veličin
- regulátory transformující získané informace na akční zásahy podle určité strategie
- akční členy vykonávající akční zásahy vně řídicího systému.

Regulační technika, zvláště snímače a akční členy, se vyvíjela podle aplikačních oborů. Hrubé dělení lze provést na tyto obory

- Tepelná technika – regulace teploty, dodávky tepla, průtoků, atd.
- Chemie – stejné regulační obvody jako u tepelné techniky, navíc regulace složení
- Regulace toku materiálu – kontinuální výrobní procesy v hutích a ve výrobě stavebních hmot
- Polohová regulace a regulace otáček – regulace pro obráběcí stroje, válcovací tratě.

Nově se vyvíjí aplikace v dopravních prostředcích (automobily, lodě, letadla).

Otázky 3.1.

1. Z čeho se skládá analogový regulační obvod?
2. Čím se liší analogový regulační obvod od diskrétního?

3.2. Pneumatické a hydraulické systémy pro řízení a ovládání

Čas ke studiu: 1 hodina

Cíl Po prostudování tohoto odstavce budete umět

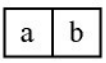
- popsat rozvaděč a jeho části.
- popsat funkci rozvaděče.
- popsat způsoby ovládání pneumatický a hydraulických prvků.

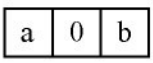
Výklad

V této kapitole budou popsány pneumatické systémy řízení umožňující změnit polohu ovládaného zařízení nebo zařízení spustit a zastavit. V těchto systémech nejde o zpětnovazební řízení, ale o logické řízení. Například je přestavován hydraulický válec do dvou krajních poloh. Popis pneumatických a hydraulických regulátorů pro zpětnovazební řízení je zařazen do kapitoly o analogových regulátorech. Tato kapitola je považována za úvod k problematice pneumatických a hydraulických systémů.

Hydraulické a pneumatické obvody

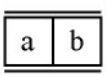
Prvky, které ovládají proudění a přívod stlačeného vzduchu nebo kapaliny k pracovním strojům, se nazývají ventily. Ventily umožňující jednoduché řízení směru pohybu hydromotoru nebo obecně řízení směru průtoku, případně jeho přehrazení, se nazývají rozváděče. Rozváděče mohou mít více poloh (vícepolohové ventily) nebo jinak řečeno stavů, například klidový stav (polohu) nebo aktivovaný stav (polohu). Jednotlivé polohy (stavy) jsou znázorněny symbolem okénka (políčka) a jsou uspořádány vedle. Platí, že počet poloh je shodný s počtem políček.

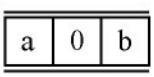




Výchozí poloha rozvaděče je stav po zapnutí systému, tj. po přivedení tlaku. Klidová poloha je u rozváděčů s vratnou pružinou taková poloha, která není aktivována. Ve schématech jsou přívody tlakového media kresleny k okénku, které odpovídá nulové nebo výchozí poloze. Aktivovaná poloha je opakem ke klidové poloze. Další hledisko k označování poloh rozvaděče je podle jeho průchodnosti. Průchozí poloha odpovídá stavu, kdy ventilem protéká pracovní medium k místu využití. Naproti tomu uzavřená poloha je neprůchozí pro médium k místu jeho pracovního využití. Stavy mohou být značeny písmeny uvnitř okénka. Přívody jsou kresleny k okénku odpovídajícímu výchozímu (klidovému) stavu. Aktivování rozváděče pomyslně přesouvá okénka z výchozí polohy tak, aby přívody mířily k příslušnému okénku, který je funkční v aktivovaném stavu.

Pro úplnost je označení rozvaděče nebo ventilu bez stabilní polohy, tj. se spojitým přesunem bloků doplněno následujícím způsobem

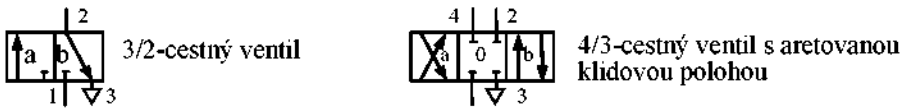




Tyto prvky jsou označovány jako proporcionální ventily a budou popsány později.

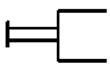
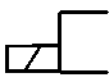
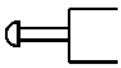
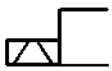
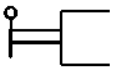
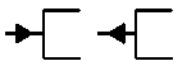
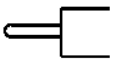
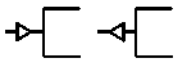
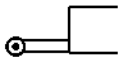
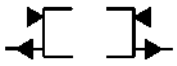
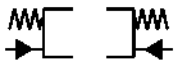
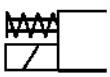
Označení druhu pracovního média a kreslení výstupů je znázorněno na následujících obrázcích.

Ve zkráceném označování rozváděčů je uveden počet řízených propojovaných přívodů (cest z okolí) a počet jeho poloh (stavů). Například označení 3/2-cestný ventil znamená, že jde o ventil třícestný a dvoupolohový. Na následujícím obrázku je příklad 3/2-cestného a 4/3-cestného ventilu (rozváděče) s očíslovanými přívody a vývody. Vývody jsou také nakresleny u bloku, který odpovídá klidovému stavu.



Polohy (stavy) ventilů (rozváděčů) se ovládají různými způsoby. Přehled schématického značení některých způsobů ovládání polohy ventilu je v tabulce 3.1. Popis k jednotlivým značkám dává stručnou představu, jak ke změně polohy ventilu dochází.

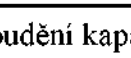
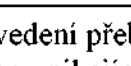
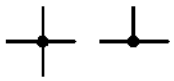
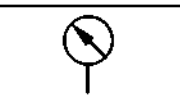
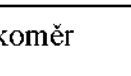
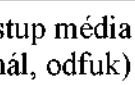
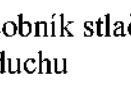
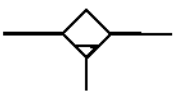
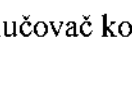
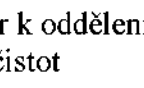
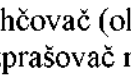
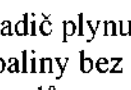
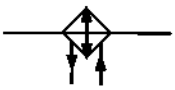
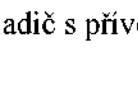
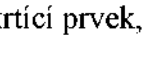
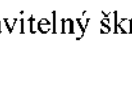
Tabulka 3.1 – Některé značky pro způsoby ovládání pneumatických a hydraulických prvků

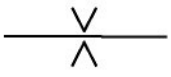
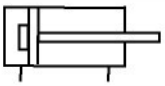
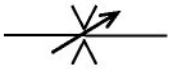
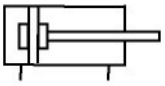
Značka	Popis způsobu ovládání	Značka	Popis způsobu ovládání
	obecné manuální ovládání		elektrické ovládání elektromagnetem s jedním vinutím
	manuální ovládání tlačítkem		elektrické ovládání elektromagnetem se 2 vinutími
	manuální ovládání pákou		přímé ovládání hydraulické přivedením a uvolněním tlaku
	mechanické ovládání páčkou nebo tlačítkem		přímé ovládání pneumatické přivedením a uvolněním tlaku
	mechanické ovládání kladkou		tlakové centrování
	mechanické ovládání pružinou		pružinové centrování
	elektromagnetem nebo manuálně se zpětnou pružinou		západka aretuje ventil v určitých polohách

K nakreslení pneumatického nebo hydraulického obvodu je třeba znát pravidla a značky základních prvků těchto obvodů. Přehledově jsou některé vybrané značky uvedeny v tabulce 3.2. Výběr byl proveden tak, aby bylo možné rozumět aspoň základním obvodům.

Jak je zřejmé, značky jsou společné pro hydraulické a pneumatické prvky. Odlišení obvodu se vzduchem a kapalinou je ve značce na vedení nebo uvnitř kompresoru nebo motoru. Plná šipka (hrot tvaru trojúhelníku) označuje kapalinu a šipka s prázdným vnitřkem vzduch. I když jsou značky kresleny pro vzduch, záměnou zmíněného označení platí také pro kapalinu.

Tabulka 3.2 – Některé značky používané ve schématech pneumatických a hydraulických systémů

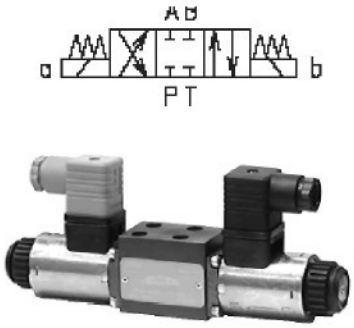
Značka	Funkce		
	proudění kapaliny		proudění vzduchu
	provozní, zpětné a napájecí vedení, přívod energie		řídicí vedení k přenosu tlakových řídicích signálů
	odvedení přebytečného nebo unikajícího media		zdroj tlaku, přípojka tlakového vzduchu
	spoje		zpětný ventil
	tlakoměr		zpětný ventil s pružinou
	výstup média (odlehčovací kanál, odfuk)		elektromotor
	zásobník stlačeného vzduchu		kompresor
	odlučovač kondenzátu		pneumatický motor pro jeden směr pohybu
	filtr k oddělení pevných nečistot		pneumatický motor pro oba směry pohybu
	zvlhčovač (olejovač), rozprašovač maziva		kyvný motor s omezeným úhlem otočení
	chladič plynu nebo kapaliny bez označení přívodů		jednočinný pneumatický válec - zpětný pohyb vnější silou
	chladič s přívody		jednočinný pneumatický válec - zpětný pohyb vratnou pružinou
	škrtící prvek, clona, dýza		dvojčinný pneumatický válec s jednostrannou pístní tyčí (pístnicí)
	Stavitelný škrticí ventil		dvojčinný pneumatický válec s oboustrannou pístní tyčí (pístnicí)

Značka	Funkce		
	clona		válec s vnitřním tlumením na dně (podrobná značka)
	stavitelná clona		válec s oboustranným vnitřním tlumením (podrobná značka)

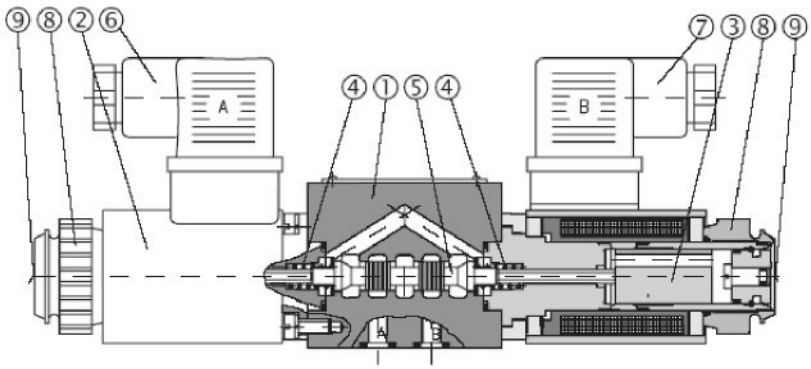
Příklad provedení rozvaděče

Podle konstrukce se rozváděče dělí na sedlové nebo šoupátkové. U sedlového ventilu je uzavíracím tělesem kulička, kužel, kotouč nebo talíř. U šoupátkových ventilů jsou vývody ovládány pomocí pístového šoupátka. Ventily s plochým šoupátkem mají uzavírací plochy rovinné.

Příkladem elektromagneticky ovládaného rozváděče je řada výrobků firmy ARGO-HYTOS a.s. s označením RPE2-04. Jsou to šoupátkové rozváděče 4/3, 4/2 a 3/2 pro tlak 32 MPa a průtok 20 dm^3/min. Rozváděč je znázorněn na obrázku 3.3 a jeho řez na obrázku 3.4.



Obrázek 3.3 – Rozváděč HYTOS RPE2-04



Obrázek 3.4 – Řez rozváděčem HYTOS RPE2-04

Rozváděče se skládají z litinového tělesa (1), válcového šoupátka (5), vratných pružin (4) a ovládacích elektromagnetů (2, 3). Třípolohové rozváděče mají vždy dva ovládací elektromagnety a dvě vratné pružiny. Jestliže některou cívkou protéká proud, přesune se rozváděč z nulové do jedné z krajních poloh. Dvoupolohové rozváděče mají jednu vratnou pružinu a jeden ovládací elektromagnet, nebo dva ovládací elektromagnety a aretaci polohy válcového šoupátka. Ovládací elektromagnety jsou napájeny stejnosměrným proudem přes konektorové nástrčky A, B (6, 7) bez usměrňovače, nebo střídavým proudem přes konektorové nástrčky A, B (6, 7) s usměrňovačem. Po povolení upevňovací matice (8) lze elektromagnety (2, 3) natáčet kolem osy v rozsahu 360°. Do výše tlaku 2,5 MPa v kanálu „T“ lze rozváděče ovládat nouzovým ručním ovládáním (9)..

Otázky 3.2.

1. K čemu se používají rozvaděče?
2. Jaká je značka pro zásobník stlačeného vzduchu?

3. Jaká je značka pro jednočinný pneumatický válec?

3.3. Analogové regulátory pro zpětnovazební řízení

Čas ke studiu: 1 hodina

Cíl Po prostudování tohoto odstavce budete umět

- definovat analogové regulátory a jejich realizaci.
- popsat příklady jednoduchých mechanických regulátorů - Wattův rozněžním, splachovadlo, tlaková nádoba.

Výklad

Analogové regulátory lze rozdělit také podle druhu měronosné veličiny, která předává informaci do regulátoru, a způsobu zacházení s touto informací uvnitř regulátoru. Typickou měronosnou veličinou je elektrické napětí, které se přenáší párem vodičů. Jinou měronosnou veličinou je poloha například páky. Další měronosnou veličinou je tlak vzduchu nebo oleje. Tlakový signál se přenáší tzv. impulsním potrubím (pro vzduch typicky průměr 6 mm). Typy regulátorů jsou

- elektrické (zpracovávají elektrické napětí, výstup rovněž elektrický)
- mechanické (přenášejí pohyb na jiný pohyb)
- pneumatické (vstup je tlak vzduchu nebo mechanický pohyb a nebo elektrický signál, výstup tlak vzduchu)
- hydraulické (vstup je tlak kapaliny nebo mechanický pohyb a nebo elektrický signál, výstup tlak kapaliny)

Regulátory lze dělit také podle zdroje energie, kterou ke své činnosti potřebují. Jsou to

- direktní regulátory, které nepotřebují vnější přívod energie (např. splachovadlo)
- indirektní regulátory s vnějším přívodem energie (všechna elektrická zařízení)

Posledním hlediskem bude matematický popis chování regulátoru. Regulátory jsou rozděleny na

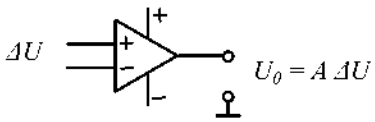
- lineární – matematický popis představuje lineární algebraická nebo diferenciální a nebo diferenční rovnice, většinou to jsou regulátory typu PID
- nelineární – modelem je nelineární rovnice, často s nespojitostí závislosti výstupu na vstupu, tj. regulační odchylce.

Elektronické regulátory

Nejrozšířenějším typem elektronického regulátoru je regulátor PID. Jedná se o lineární regulátor, jehož výstup obsahuje součet složek proporcionální, derivační, a složku integrační

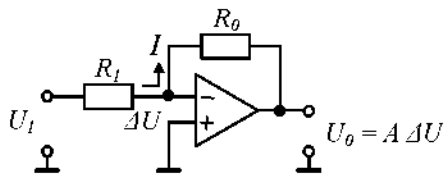
reagující na rychlost změny regulační odchylky úměrnou časovému integrálu regulační odchylky. Výstup tohoto regulátoru působí v čase spojitě.

Základním stavebním prvkem elektronického regulátoru je operační zesilovač. Obvodová značka a označení napětí je na obrázku 3.5. Vstup tohoto zesilovače je symetrický - diferenciální, tj. reaguje na rozdíl dvou napěťových signálů ΔU přivedených na dvě vstupní svorky. Výstup tohoto zesilovače U_0 je nesymetrický, tj. napěťový proti společné zemi (s nulovým elektrickým potenciálem).



Obrázek 3.5 – Operační zesilovač

Napájecí napětí operačních zesilovačů je ±12 V ss nebo ±15 V ss pro rozsah pracovních napětí ±10V. Operační zesilovač zesiluje i stejnosměrná napětí narozdíl od zesilovačů známých z audiotechniky. Tyto zesilovače jsou charakteristické vysokou vstupní impedancí a vysokým zesílením A, řádově například 10^5. Toto zesílení není časově stálé a proto se tento zesilovač doplňuje zpětnou vazbou tak, aby se zesílení omezilo na hodnoty o mnoho řádů méně, například 100.



Obrázek 3.6 – Invertující zesilovač

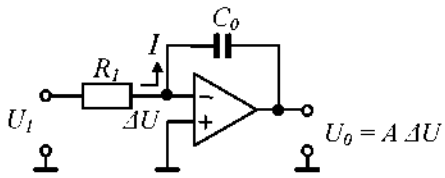
K tomuto účelu jsou využity elektrické odpory R_1 a R_0 v zapojení podle obrázku 3.6. První odpor je připojen na invertující vstup zesilovače a druhý odpor spojuje tento vstup s výstupem a tvoří zpětnou vazbu. Neinvertující vstup zesilovače je uzemněn. Ve skutečném zapojení je do tohoto spojení vložen odpor shodný s odporem R_1. Napěťový rozsah výstupního napětí tohoto obvodu je ±10 V.

Vzhledem k obrovskému zesílení odpovídá na vstupu zesilovače diferenční napětí maximálně desetiny milivoltu a lze jej zanedbat. Invertující vstup zesilovače je tedy na virtuální nule. Podle předpokladu má zesilovač vysokou vstupní impedanci, proto proud ze vstupu je shodný s proudem I zpětnovazebním odporem. Platí

$$I = \frac{U_1}{R_1} = -\frac{U_0}{R_0} \quad \Rightarrow \quad \frac{U_0}{U_1} = -\frac{R_0}{R_1} \tag{3.1}$$

Zesílení zesilovače je dáno pouze poměrem odporů. Znaménko minus znamená, že zesilovač invertuje, tj. obrací fázi střídavého signálu. Zesilovač se používá k proporcionální složce regulátoru PID. Poměr odporů se volí rak, aby nepřesáhl 100.

Integrátor je další variantou zapojení zpětnovazebního obvodu, viz obrázek 3.7. Místo odporu ve zpětné vazbě je použit elektrický kondenzátor s kapacitou C_0 . Jestliže za impedanci kondenzátoru se dosadí operátorový tvar $1/C_0 s$, pak se rovnice (3.1) změní na Laplaceův přenos

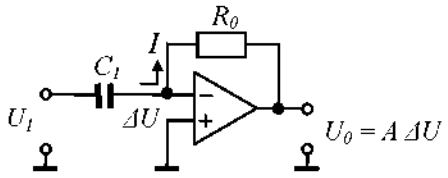


Obrázek 3.7 – Integrátor

$$I = \frac{U_1}{R_1} = -\frac{U_0}{1/C_0 s} \quad \Rightarrow \quad \frac{U_0}{U_1} = -\frac{1}{R_1 C_0 s} \tag{3.2}$$

Přenos vstupního napětí na výstup má tvar integračního článku s časovou konstantou $T_I = R_1 C_0$. Vysvětlit chování tohoto obvodu lze na příkladu jednotkového skoku vstupního napětí. Konstantní proud odporem R_1 nabíjí lineárně v čase kondenzátor, což se projeví vzrůstem výstupního napětí, viz obrázek 3.9 A. Kladný pól kondenzátoru je připojen ke vstupu operačního zesilovače, který má téměř nulový potenciál, proto napětí na výstupu roste do záporných hodnot.

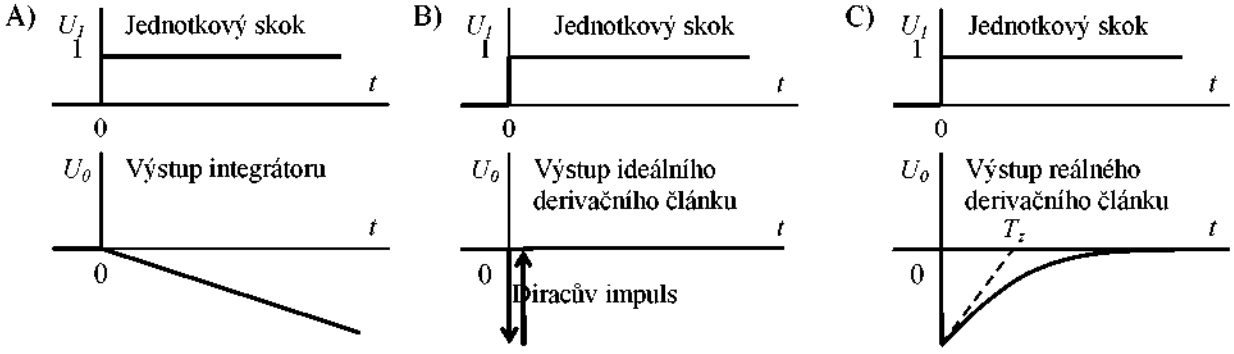
Derivační článek má na rozdíl od integrátoru je elektrický kondenzátor připojen mezi vstup obvodu a invertující vstup operačního zesilovače, viz obrázek 3.8. Kapacita tohoto kondenzátoru je označena C_1. Vzhledem k operátorovému tvaru impedance kondenzátoru se rovnice (3.1) se změní na tvar, který odpovídá Laplaceovu přenosu ideálního derivačního článku.



Obrázek 3.8 – Derivační článek

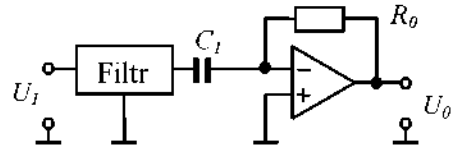
$$I = \frac{U_1}{1/C_0 s} = -\frac{U_0}{R_0} \quad \Rightarrow \quad \frac{U_0}{U_1} = -R_0 C_1 s\,, \tag{3.3}$$

Teoretický průběh výstupního napětí ve tvaru Diracova impulsu na obrázku 3.9 B, jako odezva na skokovou změnu vstupního napětí, není pro omezené napájecí napětí možný. Ideální derivační článek není proto fyzikálně realizovatelný. Nelze připustit skokové změny vstupního napětí, proto je tomuto ideálnímu článku sériově předřazen setrvačný článek (proporcionální soustava prvního řádu), viz obrázek 3.10. Odezva obvodu na skok vstupního napětí se změní na omezený prodloužený impuls, jak je zřejmé z obrázku 3.9 C.



Obrázek 3.9 – Přechodové charakteristiky kanálů I a D elektronického regulátoru

Časová konstanta exponenciálního doznívání odezvy reálného integračního článku T_z je shodná s časovou konstantou zmíněné proporcionální soustavy prvního řádu ve funkci dolnopropustného filtru. Pro přenos napětí platí



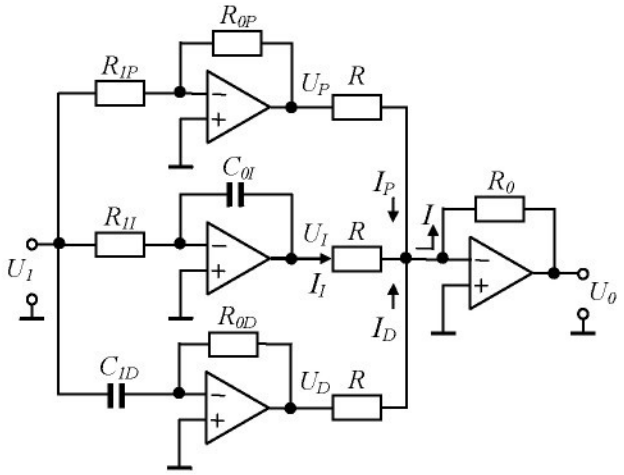
Obrázek 3.10 – Reálný derivační článek

$$\frac{U_0}{U_1} = -\frac{R_0 C_1 s}{1 + T_z s} \tag{3.4}$$

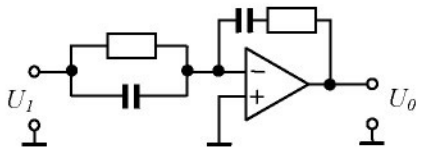
V textu této podkapitoly byly podrobně analyzovány všechny tři typy větví regulátoru typu PID. Zbývá pouze sčítací obvod. Schéma sčítacího obvodu jako součást úplného

regulátoru PID je znázorněno na obrázku 3.11. Do vstupu operačního zesilovače sumačního obvodu se sčítají tři proudy, které se musí rovnat proudu zpětnovazebním odporem. Platí

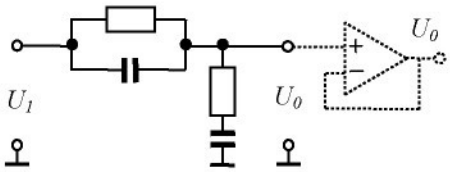
$$I = I_P + I_I + I_D = \frac{U_P}{R} + \frac{U_I}{R} + \frac{U_D}{R} = -\frac{U_0}{R_0} \quad \Rightarrow \quad U_0 = -\left(U_P + U_I + U_D\right) \qquad (3.5)$$



Obrázek 3.11 – Regulátor PID bez interakce nastavení parametrů



Obrázek 3.12 – Regulátor PID s interakcí nastavení parametrů



Obrázek 3.13 – Pasivní zapojení regulátoru PID

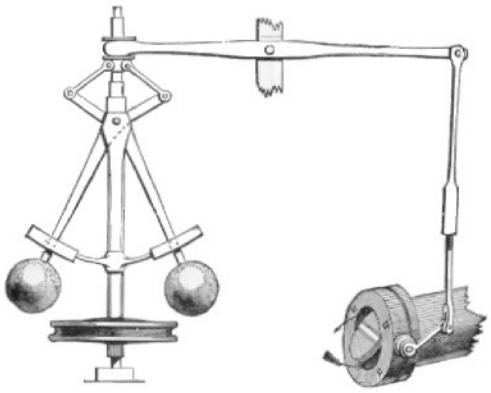
Sériové spojení jednotlivých větví zvlášť a sčítacího obvodu ruší dvojité minus v přenosech napětí. Výsledný obvod má přenos odpovídající regulátoru PID. Velikost kapacity a odporu ovlivňují vždy jeden parametr regulátoru, tj. některou z časových konstant nebo zesílení regulátoru. Tomuto zapojení se říká bez interakce. Naproti tomu zapojení na obrázku 3.12 je sice zdánlivě zjednodušeno na jeden operační zesilovač (druhý invertující je třeba ke správné polaritě výstupního napětí), ale nastavení velikosti kapacity a odporu ovlivňuje všechny parametry přenosu, proto toto zapojení má označení s interakcí.

Zapojení regulátorů na obrázcích 3.11 a 3.12 vyžaduje operační zesilovače, proto je označováno jako zapojení s aktivními prvky. Pasivní zapojení představuje filtr (čtyřpól) na obrázku 3.13. Vlastnosti jednotlivých prvků rovněž ovlivňují všechny parametry tohoto regulátoru. Zatímco výstup všech obvodů s operačními zesilovači bylo možné zatížit odběrem proudu dalšími obvody, pasivní zapojení vyžaduje impedanční oddělovač, tj. zesilovač s velmi vysokým vstupním odporem a zesílením 1. Zapojení tohoto obvodu je naznačeno také v obrázku 3.13.

❑ Mechanické regulátory

Nejstarší regulátory byly sestaveny na principu mechanickém. Od nejstarších dob dávali lidé umělým systémům schopnost regulovat samy sebe (autoregulace). Co jiného je přepad u nádržky s vodou než dvoustavový regulátor? V oblasti mechaniky byl zájem regulovat rychlost otáčení různých strojů nebo chod hodin a hracích strojků, které byly poháněny závažím nebo pružinou. První zmínka o krokovém mechanismu s kyvadlem pro hodiny je z roku 1250. Krokový

mechanismus kompenzoval ztrátu energie třením u hodinového kyvadla nebo později setrvačníku (nepokoje) v hodinkách a zároveň tyto jednohmotové mechanické systémy svou schopností kmitat pevnou frekvencí omezovaly jejich rychlost otáčení.

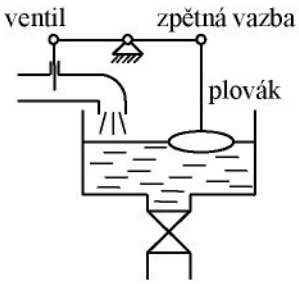


Obrázek 3.14 – Wattův roztěžník

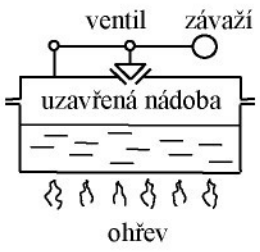
Nejpozoruhodnější inovaci konstrukce mechanického stroje, která spočívala v zavedení zpětnovazebního řízení, vytvořil skotský mechanik James Watt (1736 – 1819). Tento zdatný technik doplnil parní stroj nejen o setrvačník a dvoucestný ventil k rozvodu páry, ale z našeho hlediska hlavně také o odstředivý regulátor (jinak Wattův roztěžník), který zprostředkovává automatickou zpětnou vazbu, která reguluje otáčky parního stroje. Tento proporcionální regulátor navrhl Watt v roce 1788. Jeho kresba, která byla převzata z encyklopedie Wikipedia, je na obrázku 3.14. Roztěžník je poháněn regulovaným strojem. Dvě závaží se vychylují odstředivou silou a posouvají objímkou, která vychyluje pákový převod ovládající natočení klapky v přívodním potrubí vodní páry pro stroj.

Princip silového působení dvou rotujících hmot zůstal uplatněn až dodnes v mechanickém vstřikovacím čerpadle pro dieselový motor s funkcí omezovat dodávku paliva a tím i otáčky po překročení zadané meze, při které překoná odstředivá síla, působící na rotující hmoty, přítlak pružin a pohyb se přenese na posun regulační tyče. Tento posun způsobí pootočení ventilů ve vstřikovacím čerpadlu tak, že pohybem vačky se nasaje a pak vstřikne do válců menší množství paliva.

Jinou zajímavou aplikací je omezovač rychlosti pohybu výtahu, který zájemce může vidět v muzeu u krajského úřadu ve Zlíně. Tento stroj měl kontrolovat pohyb kanceláře ve výtahu Baťova mrakodrapu a při překročení určité rychlosti měla být uvedena do činnosti brzda.



Obrázek 3.15 – Splachovadlo



Obrázek 3.16 – Tlaková nádoba

Oblíbeným příkladem mechanického regulátoru je splachovadlo (obrázek 3.15). Pohyb plováku ovládá prostřednictvím páky ventil přítoku paliva. Změny polohy a plováku jsou lineární. Skutečné soudobé splachovadlo připomíná spíše klopný obvod. Průtok je po celou dobu udržován na maximální velikosti.

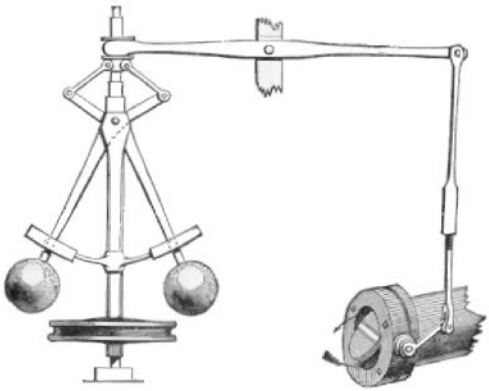
Po dosažení určité hladiny se pákový převod rychle překlopí a přívod vody uzavře (často nespolehlivě, což je důvodem napětí v mnohých domácnostech).

Posledním příkladem je regulace maximálního tlaku v nádobě s odpařující se tekutinou (obrázek 3.16). Tlak par v nádobě překoná přítlak ventilu a přebytečné páry odpustí.

Mechanické doplňky dávají v uvedených příkladech systému samoregulační (autoregulační) schopnosti, čímž nedovolí, aby se dostaly za vymezené hranice. Samoregulace je běžná i v jiných systémech a není umělého původu jako v příkladu mechanických systémů.

Otázky 3.3.

1. Jaké zesílení má operační zesilovač? Velké, střední, malé.
2. Kde se používá operační zesilovač?
3. Jak se jmenuje regulátor na obrázku?



3.4. Pneumatické regulátory

Čas ke studiu: 0,5 hodiny

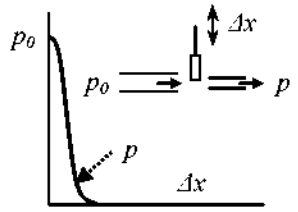
Cíl Po prostudování tohoto odstavce budete umět

- popsat výhody pneumatických regulátorů.
- popsat způsoby řízení proudu vzduchu.

Výklad

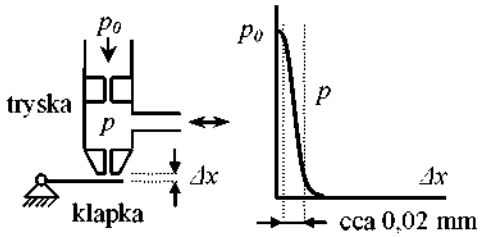
Důvodem vzniku pneumatických regulátorů je jejich jiskrová bezpečnost pro provozy s nebezpečím výbuchu v některých odvětvích chemického průmyslu. Tento konstrukční systém zahrnuje nejen regulátory, ale i snímače s tlakovým výstupem (měronosnou veličinou je tlak vzduchu) a akční členy ovládané tlakovým vzduchem. Také pneumatické regulátory využívají operační zesilovače, ovšem na principu přeměny malého mechanického pohybu na tlakový signál. Je to soustava tryska klapka, dvoukuličkový rozvod, šoupátko (šoupátkový rozvod) a proudová záchytná tryska.

Proudová záchytná tryska je znázorněna na obrázku 3.17. Změnou polohy clonky se dosáhne změna tlaku p na výstupu. Zajímavé je, že největší účinek se dosáhne při menším průměru záchytné trysky než je průměr proudové trysky.



Obrázek 3.17 Proudová záchytná tryska

Vzhledem k praktické realizaci budou popsány pouze dva systémy a to systém s klapkou-tryskou a šoupátky. Pokroky v elektronice však i první zmíněný systém postupně vytlačují. V již zmíněné první skupině regulátorů existují všechny varianty regulátorů odvozených od typu PID včetně zesilovačů výkonu. Jejich funkce je založena na ekvivalentu operačního zesilovače, kterým je soustava tryska klapka, viz obrázek 3.18. Tryska se napájí konstantním tlakem p_0 (obvykle 140 kPa).

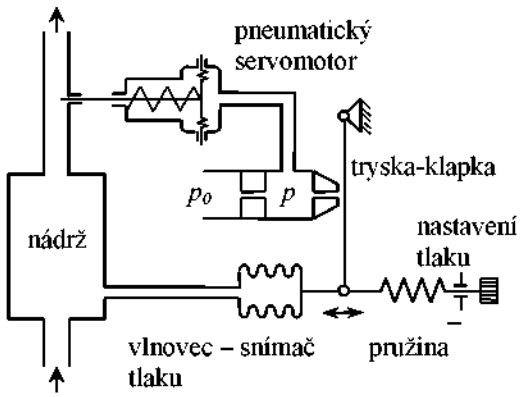


Obrázek 3.18 – Tryska-klapka

Průměr napájecího otvoru je 0,25 mm a průměr trysky 0,55 mm, takže jsou to kapiláry. Odběr tlaku je značen p. Vzduch prochází dvěma kapilárami, přičemž výstup je škrcen klapkou. Závislost tlaku na poloze klapky je znázorněna na grafu v tomto obrázku. Jak je zřejmé, velmi malé pohyby trysky v rozsahu 0,02 mm, způsobí velkou změnu tlaku p na výstupu z 20 na 100 kPa. Tento operační zesilovač má také slušnou linearitu.

Pneumatický operační zesilovač je schopen změnou tlaku vyvolat silovou rovnováhu pákové soustavy a udržet ji v rovnováze v potřebném pracovním bodě.

Funkci proporcionálního pneumatického regulátoru (D a I složka chybí) lze demonstrovat na příkladu na obrázku 3.19. Úkolem regulace je udržet v nádobě stálý tlak regulovaným odpouštěním náplně. Snímač tlaku představuje vlnovec. Proti vychýlení vlnovce působí silou pružina, jejímž předpětím se zadává žádaná hodnota tlaku.



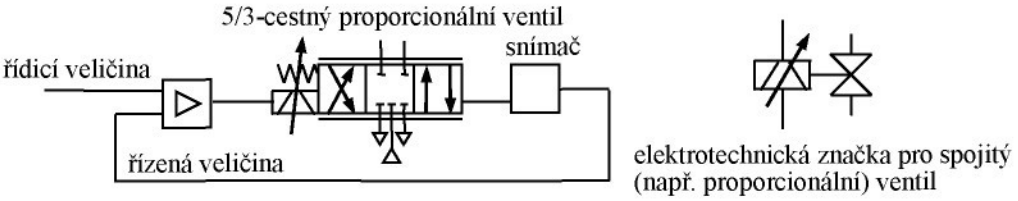
Obrázek 3.19 – Regulace tlaku proporcionální pneumatickým regulátorem

Pohyblivý konec vlnovce je spojen s klapkou systému tryska-klapka. Tryska je napájená konstantním tlakem p_0. Výstupní tlak p je přiveden do pneumatického servomotoru s vratnou pružinou. Jestliže se tlak p před pohyblivou membránou servomotoru zvětšuje, pak se zmenšuje průtočný průřez a odpouštění tlaku v nádobě se zmenšuje.

Nechť se zvětší tlak v nádobě nad stanovenou hodnotu. Důsledkem tohoto jevu je vychýlení rovnovážného stavu vlnovce takovým způsobem, že se zvětší mezera mezi tryskou a klapkou a v trysce poklesne tlak. V servomotoru vratná pružina zatáhne šoupátko dovnitř a

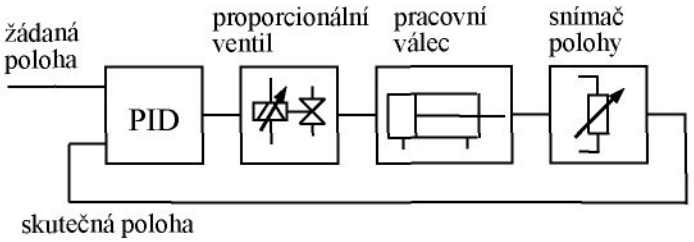
průtočný průřez ventilu se zvětší, což povede k poklesu tlaku. Při poklesu tlaku v nádobě je účinek opačný, průtočný průřez ventilu se zmenší.

Třetí způsob řízení proudu vzduchu relativně malým pohybem ovládacího prvku je založen na principu šoupátkového rozvodu. Podrobně bude uspořádání šoupátkového rozvodu popsáno u hydraulických regulátorů. Protože šoupátko nemá uvnitř ventilu jako běžný rozváděč dvě nebo tři stabilní polohy, ale posouvá se spojitě, je tento typ ventilu blíže pojmenován jako proporcionální. Pohyb šoupátka ventilu je často ovládán elektromagnetem, jehož silové působení na šoupátko je úměrné velikosti proudu elektromagnetickou cívkou. Tento ventil může být součástí uzavřené smyčky podle obrázku 3.20.



Obrázek 3.20 – 5/3-cestný regulovaný ventil – schématické značky

Polohový regulátor pracovního válce s proporcionálním ventilem (rozváděčem) je znázorněn na obrázku 3.21. V regulační smyčce je zařazen elektronický regulátor PID. Výstupní proud ovládá proporcionální ventil, který přivádí tlak do dvojčinného pracovního válce. Poloha pístní tyče (pístnice) je snímána snímačem s elektrickým výstupem. Signál polohy je zápornou zpětnou vazbou regulačního obvodu.



Obrázek 3.21 – Polohový regulátor s proporcionálním ventilem

Hraniční hodnota přesnosti pneumatického regulačního obvodu polohy je ±0,1 mm při rychlostech do 3 m/s. Lineární pohony mají délku do 2 m. Limitujícím faktorem je tření pístu pohyblivých prvků. Přesto z pneumatických prvků lze sestavit přesné a rychlé systémy řízení polohy.

Otázky 3.4.

1. Základní princip pneumatických regulátorů? Tryska + ?
2. Používají se operační zesilovače u pneumatických regulátorů?

3.5. Hydraulické zesilovací prvky

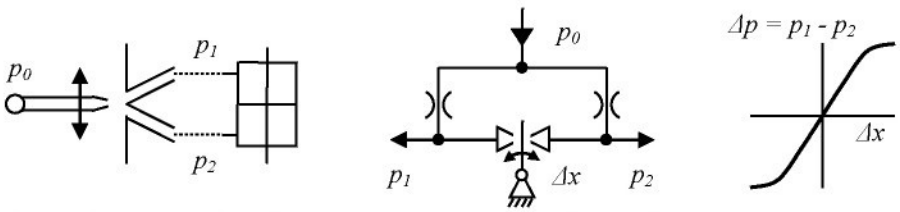
Čas ke studiu: 1 hodina

Cíl Po prostudování tohoto odstavce budete umět

- definovat princip hydraulických zesilovacích prvků.
- popsat funkci šoupátkového rozvodu.
- popsat provedení proporcionálního ventilu.

Výklad

Princip funkce těchto prvků je založen také na hydraulické verzi operačního zesilovače, tj. prvku, u kterého malá změna vstupu vyvolá velkou změnu odezvy na výstupu. Lze uvést tři principy řešení: tryskový rozvod (Askania), dvoutryskový rozvod, který se nazývá také systém klapka-tryska, a šoupátkový rozvod. První dvě řešení jsou znázorněny na obrázku 3.22. Šoupátkový rozvod bude popsán podrobněji. V úplném uspořádání působí na otočnou trysku a klapku pružný člen, který oba tyto pohyblivé prvky vrací do výchozí polohy.



A) Tryskový rozvod (Askania) B) Dvoutryskový rozvod

Obrázek 3.22 – Principy hydraulických operačních zesilovačů

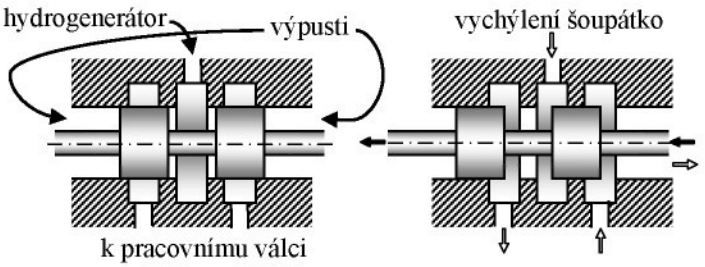
Vlastností operačních zesilovačů je způsobit malou změnou polohy velkou změnu tlaku na jejich výstupu. Princip tryskového regulátoru s názvem podle prvního výrobce (Askania) je uveden na obrázku 3.22 A. Jedná se o trysku, která je úměrně regulační odchylce vychylována směrem ke dvěma otvorům a tím je dosaženo tlakové diference mezi připojenými potrubími. Principem funkce je transformace průtoku na tlak podle Bernouliho zákona. Dvoutryskový obvod na obrázku 3.22 B je podobný systému tryska-klapka. Tlak v obou větvích určuje poloha otočné klapky (flapper).

Funkci šoupátkového rozvodu lze popsat podle obrázku 3.23. Protože se šoupátko přestavuje spojitě, je tento typ ventilů také označen proporcionální. Nákres vlevo znázorňuje šoupátko v centrované (základní) poloze. Na obrázku vpravo je šoupátko vychýleno a umožňuje průtok hydraulické kapaliny vnitřními kanálky k jednomu přívodu směrem k pracovnímu válci, zatímco druhý přívod je otevřen k výpusti.

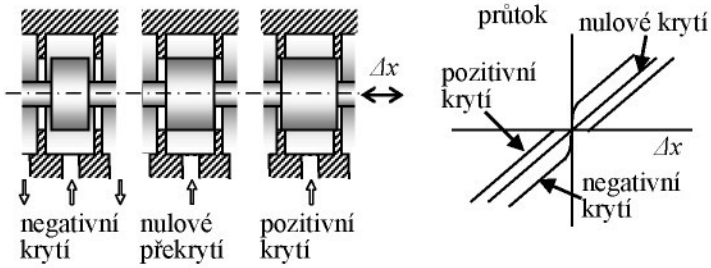
Charakteristiku šoupátkového rozvodu, tj. závislost průtoku na poloze šoupátka, ovlivňuje krytí šoupátka a kvalita pracovních hran uvnitř rozvodu. Tři různé druhy krytí (negativní, nulové a pozitivní) jsou znázorněny na obrázku 3.24. Tvar povrchu ventilu, například zápichy (prstencové drážky) nebo zploštění, může také podstatně změnit průtokovou charakteristiku ventilu. Vyrábějí se tak ventily s lomenou nebo parabolickou charakteristikou. Snížení opotřebení hran proporcionálního ventilu přispívá výřez v povrchu

válcového šoupátka, který se otvírá pro průchod kapaliny dříve něž hrany válcové části šoupátka.

Efektem negativního krytí (není přesah hran) na obrázku vlevo je stále otevřený průchod hydraulické kapaliny. Při pozitivním krytí je průchod přesahem hran uzavřen. V charakteristice šoupátkového rozvodu se objeví pásmo necitlivosti, zatímco charakteristika rozvodu s negativním krytím je v okolí střední polohy lineární. Linearita bezpochyby přispívá k lepší funkci regulačního obvodu. Napájecí tlak tohoto obvodu je řádově desítky MPa.



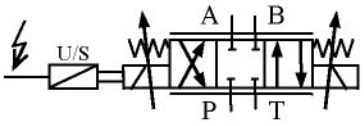
Obrázek 3.23 – Princip šoupátkového rozvodu



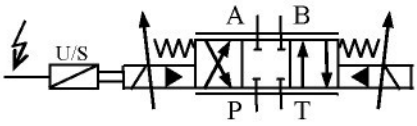
Obrázek 3.24 – Pozitivní a negativní krytí šoupátka

Obvodové značky pro proporcionální ventily jsou na obrázku 3.25. Jedná se o přímo a nepřímo řízený proporcionální ventil. Oba ventily mají čtyři vývody. K pracovnímu válci se připojují vývody A a B.

Napájecí tlaková kapalina je připojena k vývodu P a vývod T je výpusť. Šoupátko obou ventilů je centrováno dvěma pružinami. Poloha šoupátka je řízena proudem elektromagnetu. Ke zvýšení přesnosti nastavení je ventil vybaven indukčním snímačem polohy šoupátka. Regulační obvod pak může polohu šoupátka nastavovat velmi přesně. Dalším zpřesněním funkce ventilu je modulace ovládacího proudu slabým střídavým signálem (Dithersignal), který vyvolá jemné vibrace, zabraňující nedokončení přestavení polohy z důvodu tření.



Přímo řízený proporcionální ventil



Nepřímo řízený proporcionální ventil

Obrázek 3.25 – Značky pro proporcionální ventily

Nepřímo řízený proporcionální ventil má předřazeny dva tlakové omezovací ventily, které jsou ovládány elektromagnetem. Přivření nebo rozevření mezery změní odtok do výpusti a tím i tlak kapaliny, který působí na šoupátko hlavního ventilu. Toto šoupátko hlavního ventilu je drženo v centrální poloze také pružinami jako u přímo řízeného ventilu. Porucha rovnováhy sil od pružin a tlaku od předřazených omezovacích ventilů způsobí vychýlení šoupátka hlavního ventilu do nového rovnovážného stavu. Poloha tohoto šoupátka je rovněž snímána indukčním snímačem.

Proud k ovládání ventilů je v rozmezí ±20 mA. Ventil je ovládán dvěma cívkami elektromagnetu. Jejich zapojení je paralelní nebo sériové a nebo proti sobě (push-pull), což znamená, že zvýšení proudu v jedné cívce doprovází snížení proudu v cívce druhé.

❑ Příklad provedení proporcionálního ventilu (rozváděče)

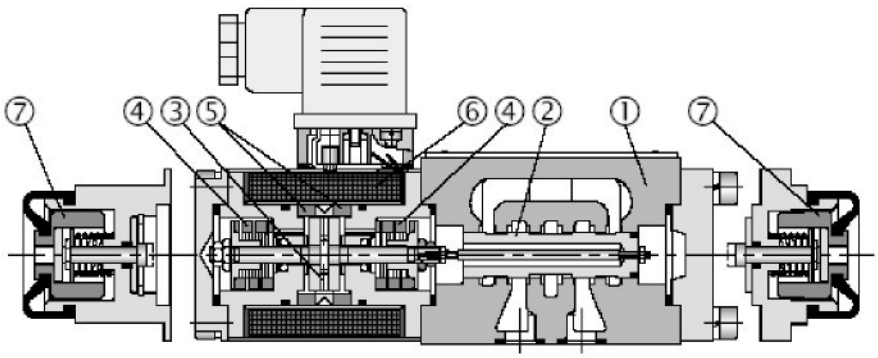
Příklad rychlého proporcionálního ventilu také od firmy ARGO-HYTOS a.s. představuje ventil typu PRL1-06 pro tlak 25 MPa a průtok 32 dm^3/min, který obsahuje lineární motor. Rozváděč je znázorněn na obrázku 3.26 a jeho řez na obrázku 3.27.

Proporcionální rozváděč PRL1 je určen ke spojitému dálkovému řízení hydromotorů a válců, převážně v mobilních aplikacích. Rozváděč sestává ze dvou částí. Hydraulickou část tvoří litinové těleso (1),do něhož je nalícováno šoupátko (2) uspořádané pro zajištění požadované funkce. Ovládací část tvoří lineární motor. Kotva (3) lineárního motoru je středěna pružinami (4) a pracovní mezery jsou protisměrně předmagnetizovány trvalými magnety (5) ze vzácných zemin.

Při vybuzení cívky (6) elektrickým proudem se posouvají kotva a s ní spojené šoupátko ze střední polohy. Výchylka kotvy je přitom úměrná řídicímu proudu a smysl výchylky závisí na směru průtoku proudu. Předností lineárního motoru je, že při výpadku napájení nebo přerušení kabelu se přesouvá kotva motoru a šoupátko do střední polohy. Provedení s nouzovým ovládáním umožňuje plynulé přestavení šoupátka do určité polohy, které se provádí zašroubováním šroubu (7).

Obrázek 3.26 – Proporcionální rozváděč HYTOS PRL1-06

Pro řízení lineárního motoru jsou k dispozici elektronické řídicí jednotky. Přestože jsou proporcionální rozváděče PRL1 určeny především k řízení velikosti a směru průtoku,mohou být použity jako řídicí stupně proporcionálních rozváděčů větších světlostí (v tomto případě jako prvky řídicí tlak). Dynamické vlastnosti proporcionálních rozváděčů PRL1 umožňují jejich použití v uzavřených regulačních obvodech.



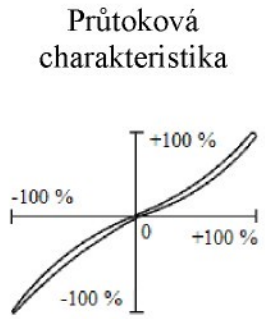
Obrázek 3.27 – Řez proporcionálním rozváděčem HYTOS PRL1-06

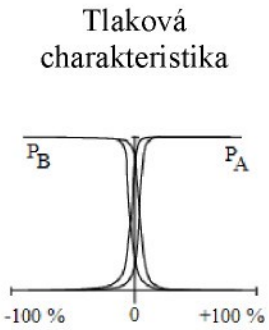
Statické charakteristiky ventilu jsou znázorněny na obrázku 3.28 a 3.29. V levé části obrázku jsou průtokové charakteristiky a v pravé části obrázku jsou tlakové charakteristiky.

Vodorovná osa všech grafů představuje relativní velikost řídicího proudu. Průtoková charakteristika odpovídá konstantnímu vstupnímu tlaku. Charakteristiky závisí na směru budícího proudu, tvoří hysterézní smyčku. Je to způsobeno ztrátami při magnetizaci a mechanickým třením. Obrázek 3.28 přísluší tzv. krytí 0 a obrázek 3.29 přísluší krytí 1. Obě průtokové charakteristiky se liší v ve svém průběhu pro malé řídicí proudy, tj. blízko nulovému proudu. Provedení s krytím 1 má výraznější pásmo necitlivosti oproti provedení s krytím 0.

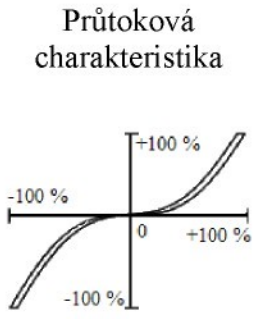
Na obrázku 3.30 je znázorněna frekvenční charakteristika proporcionálního rozvaděče. Jedná se o závislost poměru výstupního harmonického signálu (sinusovky) k amplitudě vstupního signálu. Frekvenční charakteristika je pro linearizovaný model, což znamená, že vstupní a výstupní signál má stejnou frekvenci. Druhá křivka znázorňuje velikost fázového posunu v úhlových stupních mezi vstupním a výstupním signálem. Amplitudová frekvenční charakteristika je v decibelech, tj. zmíněný poměr amplitud signálu na výstupu a vstupu je logaritmován a vynásoben 20.

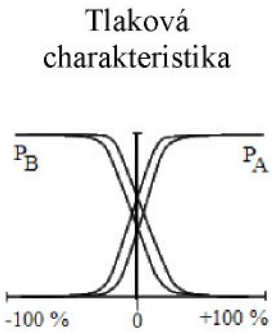
Frekvenční charakteristika vykazuje mezi 150 a 200 Hz slabou rezonanci.



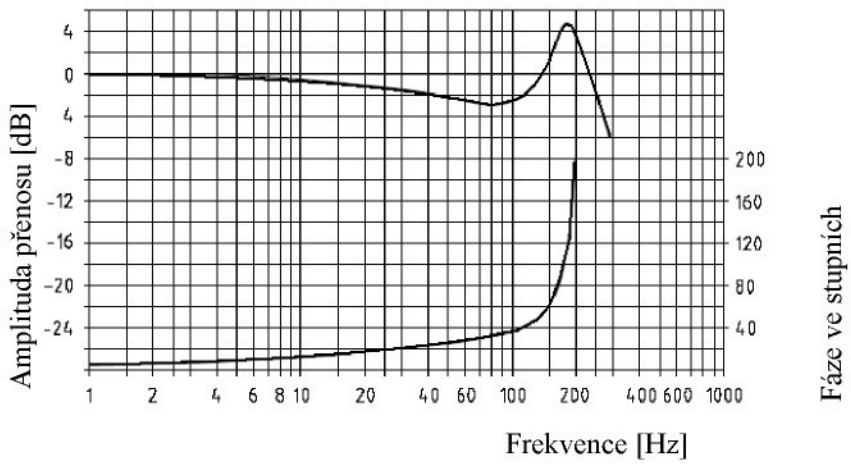


Obrázek 3.28 – Statické charakteristiky proporcionálního rozvaděče HYTOS PRL1-06 s krytím 0





Obrázek 3.29 – Statické charakteristiky proporcionálního rozvaděče HYTOS PRL1-06 s krytím 1

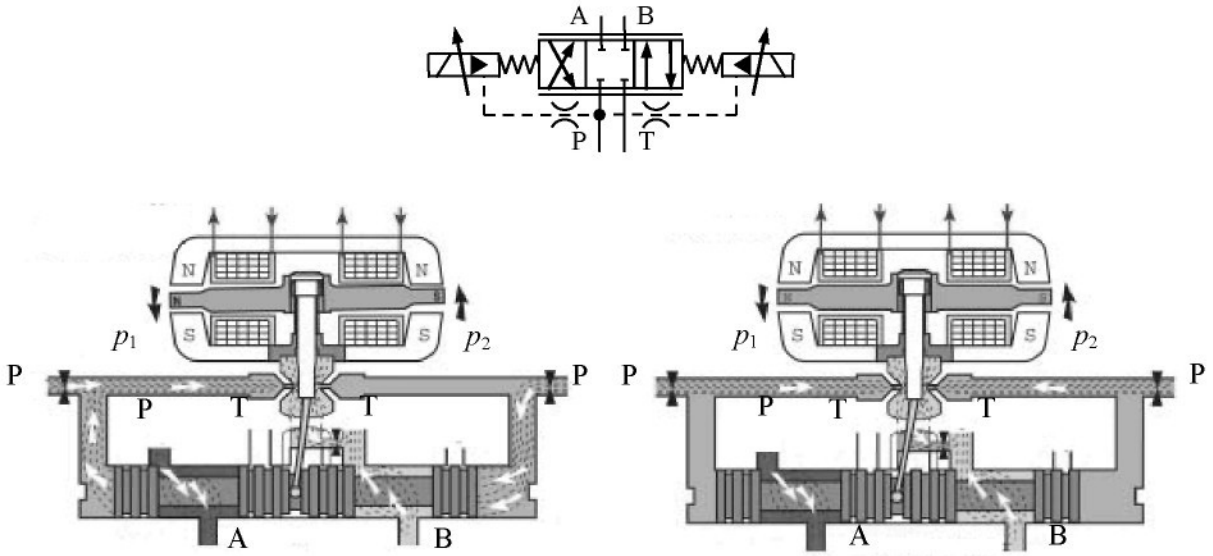


Obrázek 3.30 –Frekvenční charakteristika proporcionálního rozváděče HYTOS PRL1-06

❑ Servoventily

K hydraulických zesilovacím prvkům patří také servoventily, viz obrázek 3.31. Tyto ventily reagují rychleji než proporcionální ventily, avšak jsou mnohem citlivější na čistotu

oleje a jsou dražší. První stupeň odpovídá řešení podle obrázku 3.22B. Nulová poloha klapky je centrována pružinami. Na klapku působí také krouticí moment otočně uložené kotvy elektromagnetu s nastavitelným proudem. Výstupní tlaky p_1 a p_2 dvoutryskového rozvodu určují polohu šoupátka druhého stupně ventilu. Mezi polohou šoupátka a klapky působí mechanická zpětná vazba, která je zprostředkována pérem.



Ventil v okamžiku změny proudu cívkami

Ventil v ustáleném stavu po změně proudu

Obrázek 3.31 – Dvojstupňový servoventil

Chování ventilu lze rozfázovat. V okamžiku změny proudu cívkami dojde k vychýlení klapky a ovlivnění rozdílu tlaků v obou tryskách. Šoupátko druhého stupně ventilu se uvede do pohybu a začne unášet konec péra tak, že se snaží obnovit neutrální polohu klapky mezi tryskami. To vede k vyrovnání tlaků v tryskách a zastavení pohybu šoupátka druhého stupně.

Pro svoje vynikající dynamické vlastnosti mají servoventily použití v polohových a rychlostních regulačních obvodech.

3.6. Snímače

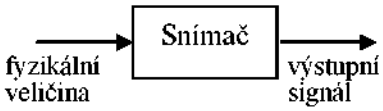
Čas ke studiu: 3 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat funkci snímače v regulačním obvodu, principy měření elektrických a neelektrických veličin.
- popsat snímače pro měření základních veličin – teploty, průtoků plynů a kapalin, síly a kroutícího momentu, tlaku, hladiny, polohyprůtoků sypkých hmot.

Výklad

V regulačním obvodu plní snímače funkci členů pro získávání informace o velikosti regulovaných a poruchových veličin. Obecně jsou snímače systémy, na jejichž vstupu je fyzikální veličina a na výstupu měronosný signál, nejčastěji elektrický, ale také tlakový nebo poloha.



Obrázek 3.32 – Snímač jako dynamický systém

V případě elektrických signálů lze provést další dělení výstupních signálů a to na signály dvouhodnotové (logická nula nebo jednička, kontakt sepnut nebo rozepnut, atd.)

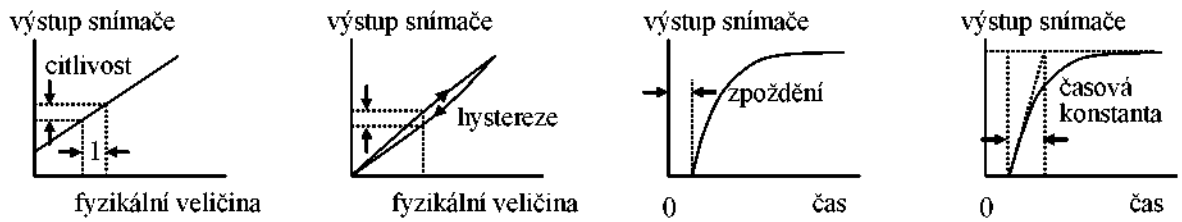
- vícehodnotové
- spojité (analogové) v unipolárním rozsahu, např. 0 až +5V, 0 až+ 10V, nebo v bipolárním rozsahu, např. . ±5V, ±10V, a nebo 4 až 20 mA
- číslicové (binárně dekadický kód, sběrnice)

Snímač jako dynamický systém mají následující vlastnosti, viz obrázek 3.32

- statickou charakteristiku
- citlivost v případě lineární statické charakteristiky
- hysterezi v případě nelineární statické charakteristiky
- přechodovou charakteristiku
- zpoždění
- časová konstanta

Další skupina parametrů snímačů obsahuje

- přesnost měření v procentech, reprodukovatelnost měření
- vstupní rozsah (měřené fyzikální veličiny), rozsah výstupního signálu (elektrické napětí nebo proud)
- přetížitelnost na vstupu, zátěž výstupu (pro elektrický signál zatěžovací proud)
- klimatické podmínky pro správnou funkci (teplota, vlhkost, a jiné), odolnost proti vibracím nebo záření
- rozměry, přípojné místo, napojení (konektory), napájecí napětí
- atd.



Obrázek 3.33 – Statické a dynamické charakteristiky snímačů

Některé obecné principy měření fyzikálních veličin

Souhrnně lze operaci měření s analogovým výstupem znázornit schématem

Fyzikální veličina → elektrický signál (napětí nebo proud)

Elektrické napětí o standardizovaném rozsahu je nejběžnější výstup snímačů. Některé snímače mají jako výstupní signál proud. V tomto případě je omezen odpor proudové smyčky. Nula rozsahu odpovídá často nenulovému proudu (např. 4 mA) pro možnost indikace přerušení proudové smyčky.

Variantou prvního schématu je mezistupeň s malým napětím

Fyzikální veličina → malé napětí → elektrický signál (napětí nebo proud)

Malé napětí v řádu milivoltů je třeba zesílit stejnosměrným elektronickým zesilovačem. V minulosti se pro napětí kolem jednoho mV používal magnetický zesilovač. Aby se konstruktéři vyhnuli problémům se stejnosměrným driftem (pomalá nahodilá změna výstupního napětí) zesilovačů, vstupní stejnosměrné napětí zesilovače bylo změněno na střídavé a výstupní napětí bylo pak synchronně usměrněno. Praktická realizace je například u citlivých inframetrů pro měření složení plynů použitím rotující clonky, přes kterou se střídavě ohřívají detekční komory, mezi kterými se měří rozdíl teplot.

Univerzální značka pro elektronický zesilovač je znázorněna na obrázku 3.33.

Obrázek 3.34 – Značka pro zesilovač

U snímačů, u kterých je vztah mezi nimi dán integrací nebo derivací (například dráha, rychlost a zrychlení), je třeba si uvědomit, která z těchto veličin je primární a která je elektronicky odvozena derivací nebo integrací.

Konstrukce snímačů využívá různých fyzikálních zákonů. Jejich stručný přehled je následující

- Termoelektrický jev na kontaktu dvou různých vodivých materiálů – termočlánky
- Vztah mezi rychlosti pohybu vodiče (v) o délce (l), magnetickou indukcí (B) a generovaným napětím u – snímače proudění vody, tachodynama

$$u = B v l \tag{3.6}$$

- Vztah mezi proudem vodičem (J) o délce (l), magnetickou indukcí (B) a vznikající silou F – proudové váhy založené na kompenzací síly elektromagnetem

$$F = B J l \tag{3.7}$$

- Závislosti elektrického odporu na teplotě – odporové teploměry, měření tepla, měření rychlosti proudění (proud vzduchu ochlazuje procházejícím proudem ohřátý vodič)
- Závislosti změny elektrického odporu ΔR na protažení Δl vodiče – tenzometry, měření mechanického napětí, síly a krouticího momentu

$$\varepsilon = \Delta l / l, \quad \Delta R / R = k \varepsilon \tag{3.8}$$

- Dopplerův jev o změně frekvence (zvuku, ultrazvuku nebo světla) f při odrazu od předmětu, který se přibližuje rychlostí v při rychlosti šíření vlnění c – měření rychlosti pohybu

$$\Delta f = \frac{c}{c - v} f \tag{3.9}$$

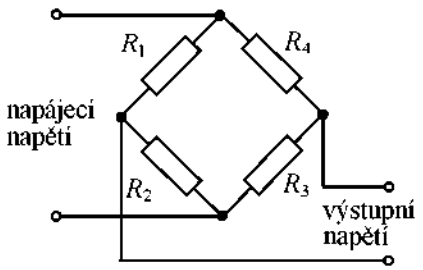
- Fotoelektrický jev – indikace přerušení světelného paprsku (lineární a rotační encodéry)
- Piezoelektrický jev – závislost generovaného náboje na působící síle na piezoelektrický materiál

Přehled fyzikálních zákonů není vyčerpávající, ale stačí k vysvětlení funkce mnoha čidel, které se používají v regulačních a řídicích obvodech.

Výše uvedené schéma snímačů lze v řadě případů popsat podrobněji. Mnohé principy měření jsou založeny na schématu

Fyzikální veličina → elektrický odpor → elektrický signál

Toto řešení má uplatnění například při měření teploty a mechanického napětí. Základní způsob přeměny změny odporu na elektrické napětí je odporový můstek na obrázku 3.34. Větve mezi napájecím napětím tvoří odporový dělič. Je zřejmé, že když obě větve budou mít stejný dělící poměr, tj.

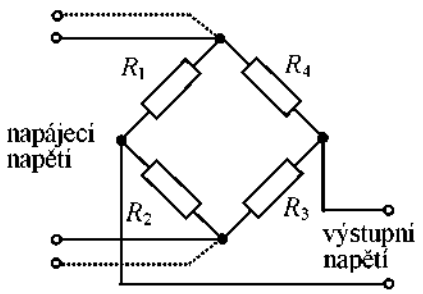


Obrázek 3.35 – Odporový můstek

$$\frac{R_1}{R_2} = \frac{R_4}{R_3} \tag{3.10}$$

pak bude můstek vyrovnán a na jeho výstupu bude nulové napětí. Změna odporů v můstku může způsobit jeho rozvážení, jen pokud nebude platit rovnice (3.10).Napájecí napětí můstku je stejnosměrné nebo také střídavé. Můstek může detekovat velmi malé změny odporů, výstupní napětí ve však velmi malé a je třeba zesilovače. Největší citlivosti měření se dosáhne, jestliže v protilehlých větvích můstku nastává změna odporu stejného znaménka a u sousedních odporů je změna znaménka opačná.

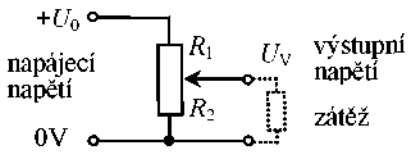
Vedení k můstku potřebuje 4 vodiče. Napájecí proud můstku způsobí nekontrolovatelný úbytek napětí a tím i změnu napětí na výstupu můstku. Aby bylo možné kontrolovat přesnou velikost napájecího napětí, přidávají se ke zmíněným 4 vodičům ještě dva další vodiče, které se zapojí do napájecích uzlů můstku, viz obrázek 3.35. Toto opatření je třeba při tenzometrických měřeních s vysokou přesností.



Obrázek 3.36 – Odporový můstek

Čtvrtinové můstkové zapojení znamená, že se vyhodnocují můstkovým zapojením změny jen jednoho odporu. V řadě případů se pracuje s polovičním můstkovým, tj. jen dvojící odporů, přičemž ve vyhodnocovací aparatuře je druhá polovina můstku.

Zvláštní řešení vyhodnocení změny odporů je v situaci, kdy zvětšení jednoho odporu je provázeno zmenšením druhého odporu o přesně stejnou velikost. Tento jev nastává u potenciometrů, viz obrázek 3.36. Poloha jeho jezdce je spřažena s posunem nebo otočením, což jsou hlavní měřené veličiny. Změny odporů se nevyhodnocují můstkovým zapojením, ale přímým vyhodnocením napětí mezi jezdcem a jedním uzlem, který je napájen napětím.



Obrázek 3.37 – Potenciometr

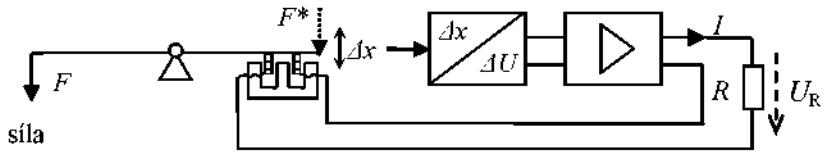
Jak již bylo zmíněno, vyhodnocuje se dráha. Jestliže je odporová dráha homogenní, pak je na výstupu odporového děliče lineárně závislé napětí na poloze jezdce. Při vyhodnocení výstupního napětí nesmí snímač napětí (voltmetr) zatěžovat potenciometr odběrem proudu. V elektrickém schématu zapojení potenciometru je tato zátěž vyznačena tečkovaně.

Další schéma převodu fyzikální veličiny na elektrický signál je následující

Tímto postupem se měří například tlak plynů a kapalin, tlaková diference, atd.. Tzv. proudové kompenzační váhy jsou založeny na principu

Fyzikální veličina → síla → elektrický proud v kompenzačním elektromagnetu → elektrický signál

Uspořádání proudových vah je znázorněno na obrázku 3.37. Měřená síly se vyvažuje sílou elektromagnetu F^*, která je přímo úměrná proudu, proto úbytek napětí U_R na odporu R je přímo úměrný měřené síle F. Velikost proudu ve smyčce je úměrná poloze ramene páky Δx. Obvod je zapojen tak (záporná zpětná vazba), že elektromagnet vrací páku do neutrální polohy. Toto zapojení má uplatnění například při měření malých tlaků a tlakových diferencí.



Obrázek 3.38 – Proudová váha

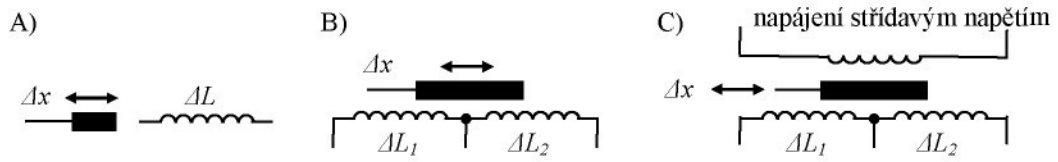
Velmi rozšířená jsou indukční snímače, které využívají tento postup měření

Fyzikální veličina → posun → indukčnost → elektrický signál

Vyhodnocování změn indukčnosti je podobné jako vyhodnocování změn odporu. Úbytek napětí U na cívce o indukčnosti L při průchodu proudu o úhlové frekvenci $\omega = 2\pi f$ je dán vzorcem

$$U = \omega L \qquad (3.11)$$

Změna indukčnosti indukuje změnu magnetické vodivosti jádra cívky, což lze dosáhnout jeho posunem. Vstupem indukčního snímače je posun jádra cívky a mezivýstupem změna indukčnosti. Uspořádání cívek je znázorněno na obrázku 3.38. Vyhodnocování změn umožňuje v případech ad A) a ad B) na uvedeném obrázku můstek doplněný odpory. Napájení odporů musí být střídavým proudem o obecně jiné frekvenci než je frekvence rozvodné sítě (50 Hz). Na části ad C) obrázku je znázorněn diferenciální transformátor, u kterého se vyhodnocuje rozdíl usměrněných napětí.



Obrázek 3.39 – Indukční snímače

Měření rychlosti pohybu, například otáčení bylo v minulosti založeno na schématu

Rychlost → otáčky tachodynama → elektrický signál

Nové systémy měření rychlosti využívají postup

Rychlost → frekvence impulsů → elektrický signál

nebo tento postup podrobněji

Rychlost → frekvence impulsů → délka časového intervalu → elektrický signál

K měření rychlosti pohybu plynů nebo kapalin se využívá Dopplerova posunu frekvencí

Rychlost → rozdíl frekvencí → elektrický signál

Tento princip se používá v rychlostních dopplerovských laserech pro bezdotykové měření vibrací. Jinou aplikací Dopplerova jevu je měření objemového proudu plynu nebo kapalin v potrubí.

Snímače složení plynů (inframetry) s nesymetrickou molekulou (CO, CO_2) využívají rozdílné absorpce infračerveného záření

Obsah plynu s nesymetrickou molekulou → elektrický odpor → elektrický signál

Obsah kyslíku v plynu se měří s využitím vzniku jeho proudu v nehomogenním magnetickém poli

Obsah kyslíku v plynu → elektrický odpor → elektrický signál.

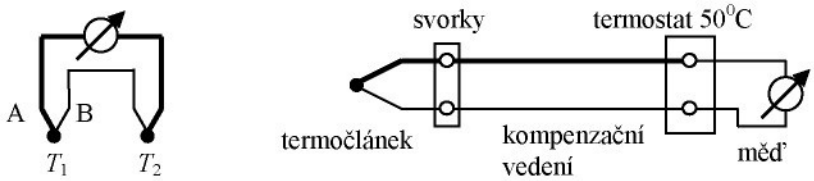
❑ Snímače teploty pro regulaci

Teplota se měří v průmyslových ohřívacích pecích, sušárnách, někdy také v topeništi, například v plynovém hořáku a v různých chemických reaktorech. Měření teploty představuje zvláštní specializaci a vědomosti. Snímače použité ke snímání teploty závisí na rozsahu měření.

Nejběžněji jsou pro měření do teploty asi 300 ^{0}C používány platinové odporové teploměry Pt 100. Tento odporový teploměr má odpor 100 Ω při 0^{0}C. Závislost odporu na teplotě je podrobně tabelována. Změny odporu jsou vyhodnocovány odporovým můstkem.

Do hodnoty kolem 1100 ^{0}C jsou běžné termočlánky, které využívají termoelektrický jev. Tento jev vzniká na spojích dvou různých kovů zapojených do série se dvěma spoji (kov A - spoj AB - kov B - spoj BA - kov A), viz obrázek 3.39. Mají-li spoje (teplý a studený) navzájem různou teplotu T_1 a T_2, vzniká na každém ze spojů odlišný elektrický potenciál, který je zdrojem proudu. Pro měření teploty je místo v obvodu s jedním spojem nahrazeno voltmetrem.

Napětí na termočlánku je dáno rozdílem teploty jeho teplého a studeného spoje. Aby naměřené napětí mohlo být spojeno s teplotou teplého konce, je třeba znát teplotu studeného konce termočlánku. Samozřejmě, že teplota studeného konce musí být stabilizována na zvolené hodnotě. Výhodná je volba teploty 50^{0}C, protože stačí pouze na této nevysoké teplotě studené konce ohřívat různě intenzivně v závislosti na teplotě okolí. K tomuto účelu se vodiče termočlánku přivádějí na svorky do termostatu, odkud pak pokračuje přenos signálu napětí k vyhodnocovacímu přístroji (ručkové měřidlo nebo A/Č převodník) měděnými vodiči. Protože je neúnosné vést dlouhá vedení z drahých kovů, ze kterých se mnohé termočlánky vyrábějí, je spojení mezi svorkami termostatu a termočlánkem z drahých kovů provedeno levnějším kompenzačním vedením. Mezi vodiči kompenzačního vedení a vodiči termočlánku při teplotě svorek je minimální termoelektrické napětí. Samotný termočlánek o délce kratší než jeden metr je umístěn v jímce a je kryt keramickými trubičkami (kapilárami). Na obrázku 3.39 je znázorněno napojení termočlánku na vyhodnocovací přístroj.

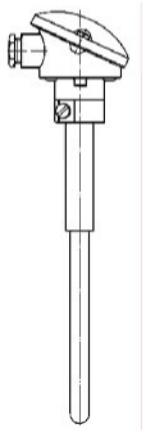


Obrázek 3.40 – Napojení termočlánku na ukazovací přístroj

Uvedený teplotní rozsah je pokryt termočlánky různých typů. V tabulce 3.3 jsou uvedeny typy termočlánků podle normy IEC 584 a rozsah maximálních teplot. V původním značení platinových termočlánků znamenají číslice procentní obsah druhého kovu ve slitině. U nás bylo hodně rozšířené používání norem DIN, avšak nyní se doporučují normy IEC.

Tabulka 3.3 – Označení podle IEC 584 a měřicí rozsah

IEC 584	Původní označení	Měřicí rozsah [° C]
T	Cu-CuNi, Cu-ko	- 200 až 350
J	Fe-CuNi	- 200 až 750
E	NiCr-CuNi, ch-ko	- 100 až 900
K	Ni-Cr-Ni, ch-a	- 200 až 1200
N	NiCrSi-NiSi	- 200 až 1200
S	PtRh10-Pt	0 až 1600
R	PtRh13-Pt	0 až 1600



Obrázek 3.41 – Provedení termočlánku

B	PtRh30-PtRh6	300 až 1700

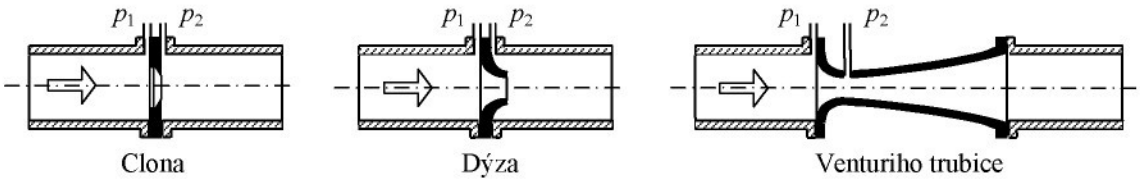
V normě nalezne zájemce tabulky termoelektrických napětí pro odstupňované teploty. Referenční hodnota je obvykle 0 ^{0}C, proto je třeba od napětí v mV odečíst napětí příslušné teplotě studeného konce termočlánku, kterou je teplota v termostatu. Nasazení termočlánku k měření znamená rovněž umění vybrat vhodnou jímku, aby byla odolná nejen teplotě, ale i korozi. Zvláštností jsou plášťové termočlánky,které jsou na rozdíl od drátových termočlánků sestaveny tak, že plášť, tepelná ochrana a samotný termočlánek tvoří integrovaný celek.

Kompenzační vedení má v ochranném obalu vpletenou niť. Například pro platinu je bílá a pro chrom-nikl je zelená

Nad teplotu 1200 ^{0}C lze použít prakticky jen bezkontaktní měření optickým pyrometrem. Tyto pyrometry jsou sestaveny z baterie termočlánků na jejichž teplý konec je optikou soustředěno tepelné záření. U zapalovacích hlav na aglomeračních spékacích pásech pro prachovou rudu je opticky pyrometr zaměřen na žáruvzdornou vyzdívku nebo do keramické jímky.

❑ Snímače průtoku plynů a kapalin pro regulaci

Pro kapaliny a plyn se používá nejčastěji k měření průtoku měřící clona nebo v případě větších průřezů Venturiho trubice. Jejich provedení je znázorněno na obrázku 3.41.



Obrázek 3.42 – Clona, dýza a Venturiho trubice

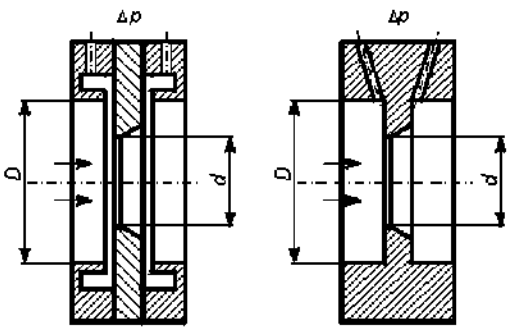
K měření průtoku se snímá tlakový spád (diference) na tomto umělém odporu proti proudění. Jestliže se v případě plynu mění značně vlivem teploty a tlaku měrná hmotnost plynu γ, pak se měří i obě tyto veličiny. Vzorce pro výpočet objemového průtoku jsou následující

$$Q = k_1\sqrt{\frac{\Delta p}{\gamma}}, \quad \left[m^3 s^{-1}; Pa, kgm^{-3}\right] \text{ nebo } Q = k_2\sqrt{\frac{\Delta p}{p}T}, \quad \left[m^3 s^{-1}; Pa, Pa, {}^0K\right] \quad (3.12)$$

Tlak a teplota jsou samozřejmě absolutní. Podstatné je, že průtok je dán odmocninou tlakové diference, jejíž snímač je pro měření průtoku nejdůležitější. Měření clonou nebo dýzou není přesné, uvádí se, že relativní nejistota je v pásmu 3 až 5% změřené hodnoty.

Návrh rozměrů clony, dýzy nebo Venturiho trubice pro daný vnitřní průměr potrubí a měřící rozsah je dán normou (ISO 5167-1). Tlaková diference se na měřící cloně snímá bodovými nebo komorovými odběry, viz obrázek 3.42.

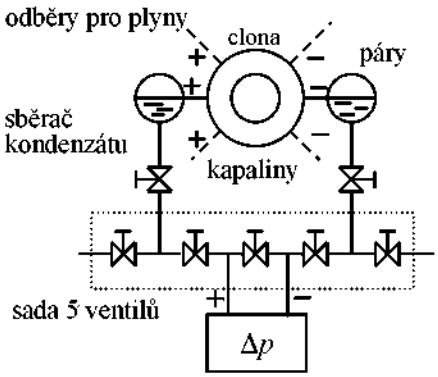
Bodový odběr u clon je vytvořen trubkami, které jsou přivařeny pod úhlem 45^0 na potrubí těsně před a za clonou mířící do osy potrubí a svírající úhel 90^0. Pro plyny je odběrné místo nahoře na osou potrubí a pro kapaliny dole. Páry se odebírají v rovině osy potrubí. Komorový odběr tlakové diference je tvořena komorou ve tvaru prstence o čtvercovém nebo obdélníkovém průřezu, která je spojena s prostředím uvnitř clony štěrbinou po celém obvodu potrubí. Obě varianty odběru diferenčního tlaku jsou znázorněny na obrázku 3.42. Clona a komorový odběr jsou vestavěny mezi příruby potrubí.



Obrázek 3.43 – Clony pro měření průtoku

Kromě clony nebo dýzy a nebo Venturiho trubice patří k vybavení impulsní potrubí, kterým se tlakový signál přivádí ke snímači tlaku nebo tlakové diference, také další příslušenství, viz obrázek 3.43. Jsou to různé kohouty a ventily, které odpojí přenos tlaku z prostředí uvnitř potrubí od tlaku na vstupu snímače. Při měření průtoku páry nebo vlhkého plynu vzniká v impulsním potrubí kondenzát, který je třeba občas ze sběrné nádoby odpustit. Sběrná nádoba pro kondenzát se umísťuje tak, aby kondenzát do ní mohl bezpečně stéci.

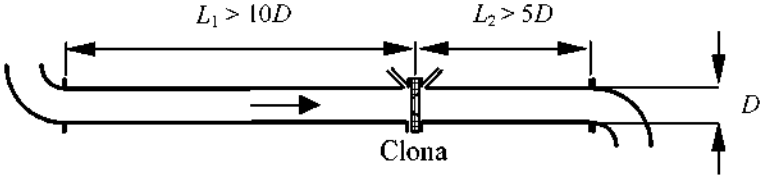
Snímač diferenčního tlaku se připojuje prostřednictvím sady pětinásobných ventilů. Aby uzavření jednoho impulsního potrubí z páru pro snímání tlakové diference nezpůsobilo přetížení snímače, je také v systému propojka mezi oběma impulsními potrubími těsně před snímačem, která se při připojování přívodu tlaku ke snímači zavře jako poslední a nebo při odpojování snímače se tato propojka otevře jako první.



Snímač diferenčního tlaku

Obrázek 3.44 – Odběr diferenčního tlaku pro snímač

Při projektování zástavby měřící clony je třeba dbát, aby potrubí bylo přímé v délce aspoň 10násobku jeho průměru (jmenovitá světlost) před clonou a 5násobku jeho průměru za clonou, jak je znázorněno na obrázku 3.44. Přímý úsek potrubí umožňuje ustavit laminární proudění, což je podmínkou přesného měření.



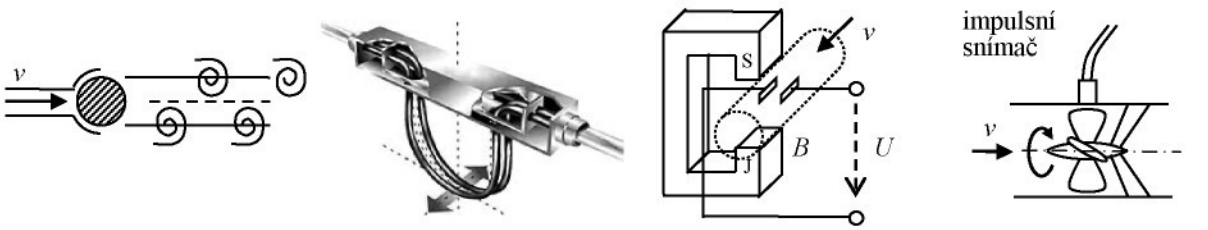
Obrázek 3.45 – Podmínky správné instalace clony do potrubí

Výjimečně lze měřit tlakovou diferenci na vnitřním a vnějším ohybu kolena.

Přehled použitelných snímačů průtoku lze nalézt na internetu [Orlíková, 2001]. Kromě snímačů se škrcením průtoku clonou, dýzou nebo Venturiho trubicí lze použít další principy, viz obrázek 3.45

- Vírové průtokoměry založené na principu von Karnamově efektu, kdy po stranách překážky neaerodynamického tvaru, která je obtékána tekutinou, se střídavě na obou stranách oddělují víry, které vytvářejí tzv. von Karmanovou stezku. Frekvence vzniku vírů je v určitém rozsahu hodnot Reynoldsova čísla nad 10^4 úměrná rychlosti proudění tekutiny, a proto i velikosti průtoku
- Průtokoměry založené na měření Coroliusovy síly, která vzniká ve vibrujících měřicích trubicích při průtoku hmotného média. Fázový posuv v rezonančním kmitání trubic, který vzniká jako důsledek působení Coriolisovy síly, je úměrný hmotnostnímu průtoku tekutiny trubicí a frekvence vlastních kmitů odpovídá hustotě tekutiny.
- Průtokoměry využívají principu již zmíněného Faradayova zákona elektromagnetické indukce. Pohybem vodiče (u měření průtoku pohybem tekutiny) v homogenním magnetickém poli se indukuje elektrické napětí.
- Turbínkové průtokoměry snímající objemový průtok. Jejich konstrukce je založena na volně otočném rotoru s lopatkami. Rotor se vlivem proudění tekutiny otáčí a jeho otáčky jsou úměrné rychlosti proudění tekutiny. Otáčky bývají snímány bezdotykovým indukčním snímačem, kdy výstupem jsou napěťové impulsy.

Vzájemné porovnání výše uvedených metod měření pro různé kapaliny a plyny je v tabulce 3.4.



Karmanův efekt | Coroliusovy síly | Faradayův zákon | Turbínkový snímač

Obrázek 3.46 – Principy dalších průtokoměrů

Tabulka 3.4 – Doporučení pro použití průtokoměrů

	Medium			
Průtokoměr	čisté plyny	pára	čisté kapaliny	kapaliny s pevnými částicemi
vírový	vhodný	vhodný	vhodný (pro kapaliny s malou viskozitou)	nevhodný
Coriolisův	vhodný	vhodný	vhodný (ne pro extrémně viskózní kapaliny)	vhodný (ne pro extrémně viskózní kapaliny)
magneticko-indukční	nepoužitelný	nepoužitelný	vhodný pro vodivé kapaliny	vhodný pro vodivé kapaliny
clona s měřením diferenčního tlaku	vhodný	vhodný	vhodný pro málo viskózní kapaliny	nevhodný
turbínový	vhodný za určitých podmínek	nevhodný	vhodný pro málo viskózní kapaliny	nevhodný

❑ Snímače síly a krouticího momentu pro regulaci

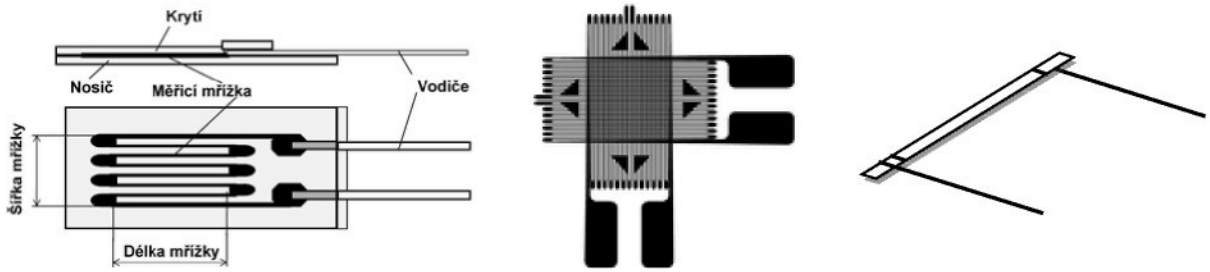
Nejběžněji se pro měření síly nebo krouticího momentu používají tenzometry. K dispozici jsou dva druhy a to kovové tenzometry a polovodičové tenzometry. Jejich princip funkce je odlišný. Kovové tenzometry využívají zúžení odporové dráhy při jejím protažení při neměnném měrném odporu materiálu, ze kterého se vyrábějí. Jedná se o slitinu konstantan (55% mědi a 45% cínu), jejíž měrný odpor téměř nezávisí na teplotě. Zatímco polovodičové tenzometry mění svůj měrný odpor v důsledku piezorezistentního jevu. Oba druhy tenzometrů mají rozdílné použití. Polovodičové tenzometry mají vysokou citlivost, naproti tomu kovové tenzometry jsou málo závislé na teplotě a mají vysokou stálost a také přesnost.

Kovové tenzometry se vyrábějí z jemného konstantanového drátku o průměru 0,01 až 0,02 mm nebo se leptají z tohoto materiálu, který je nanesen na fólii a jeho tloušťka je 0,001 mm. Elektrický odpor je 120 až 600 Ω. Součinitel deformační citlivosti, parametr k v rovnici (3.8), je roven 2. Přesnost měření pozitivně ovlivňuje nízký teplotní součinitel elektrického odporu konstantanu 5×10^{-6} °C 1/K. Polovodičové tenzometry mají součinitel deformační citlivosti roven 130.

Vedle tenzometrů pro měření povrchové deformace v tahu a tlaku v určeném směru jsou vyráběny tenzometry s mřížkami vzájemně pootočenými o 45°, které jsou určeny k

měření torzní povrchové deformace (např. hřídelů zatěžovaných krouticím momentem). Pro měření hlavního vektoru povrchové deformace v neznámé rovině napjatosti se používají tzv. rozety složené nejčastěji ze dvou tenzometrů vzájemně pootočených o 90°.

Aktivní délka tyčinek polovodičových tenzometrů mezi zlatými vývody je 2 až 10 mm, šířka 0,2 až 0,4 mm a tloušťka 0,01 až 0,03 mm. Ohmický odpor polovodičového tenzometru může být 60 až 1000 Ω. Polovodičové tenzometry jsou vyráběny pouze s přibližně stejnými metrologickými parametry a pro zapojení do měřicích obvodů jsou vybírány s těmito co nejpodobnějšími daty. Nároky na tmel spojující polovodičové tenzometry s měřeným objektem jsou vyšší, než tomu je u kovových tenzometrů.

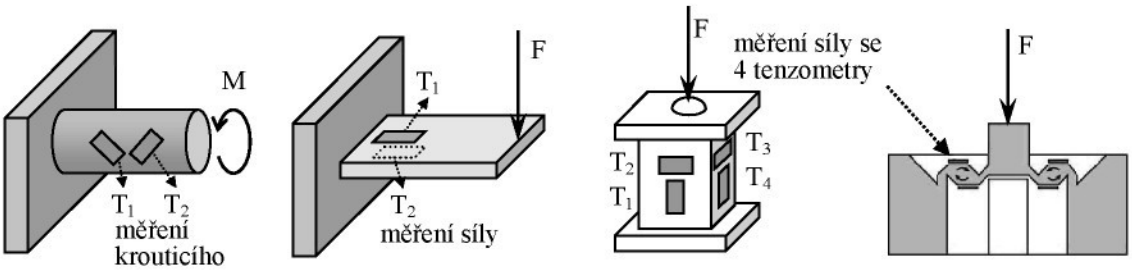


Drátový tenzometr | Fóliový tenzometr | Polovodičový tenzometr

Obrázek 3.47 – Provedení kovových a polovodičových tenzometrů

Statická charakteristika závislosti výstupního signálu na deformaci měřicího členu při konstantní teplotě může být zatížena chybou 0,1 až 1 %. Vliv teploty je podle dokonalosti výběru tenzometrů 0,05 až 0,5 % . 10 K^{-1}. Účinnou kompenzaci vlivu teploty lze zavést v intervalu 50 °C. Teoreticky lze polovodičové tenzometry využívat v intervalu teplot –70 °C až 300 °C.

Ukázka jejich provedení je na obrázku 3.46. Na obrázku 3.47 je znázorněno umístění tenzometrů pro měření síly a krouticího momentu. Pro měření síly na konci vetknutého nosníku a pro měření krouticího momentu jsou použity dva tenzometry. U vetknutého nosníku je tenzometr T_1 namáhán na tah a druhý na tlak. Při namáhání profilu prostým krutem je velikost hlavních napětí shodná s velikostí maxima smykových napětí v rovinách pootočených o 45^{0}, což určuje polohu tenzometrů T_1 a T_2, z nichž první je namáhán na tah a druhý na tlak. Oba tenzometry jsou shodně ovlivňovány teplotou, a proto se její účinek ruší, jestliže jsou ve větvi můstku, která je připojena na napájecí napětí. Při použití čtyř tenzometrů pro měření síly působící na deformační člen je jedna dvojice aktivní (největší citlivost na protažení) a druhá dvojice tenzometrů plní funkci pro kompenzaci teploty. Oba aktivní tenzometry jsou v můstku na jeho protějších stranách. Aktivní a kompenzační tenzometr jsou ve větvi můstku, která je připojena k napájecímu napětí. Vliv změn teplot na velikost odporu nemění dělicí poměr této větve. Tvar deformačního elementu může být volen tak, aby se dosáhlo co největší citlivosti měření. Zároveň je třeba jistit snímač proti přetížení, jak je ukázáno na nákresu deformačního členu v obrázku 3.47 vpravo.



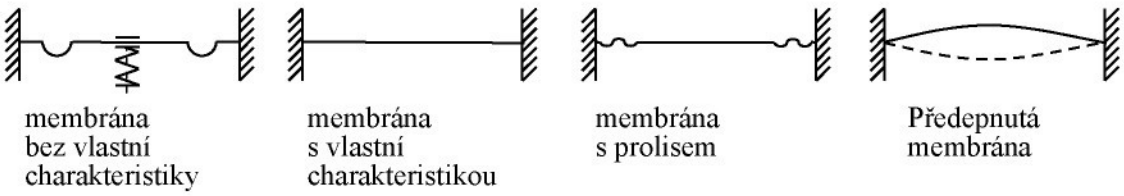
Obrázek 3.48 – Umístění tenzometrů

Snímače krouticího momentu jsou umísťovány na rotující hřídel, což způsobí problém přenosu na stacionární vyhodnocovací zařízení. Přenos se řeší rádiovým signálem nebo stěracími kroužky.

❑ Snímače tlaku a diferenčního tlaku pro regulaci

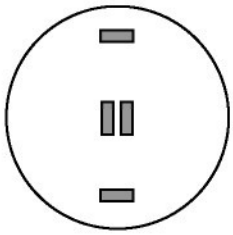
Tlakoměry lze rozdělit na kapalinové a deformační. Kapalinové tlakoměry vychylují kapalinu v U-trubici. V minulosti se pro diferenční tlak používala náplň se rtutí, která se působením tlaku vychýlila a natočila prstenec a s ním spojený ukazatel. Nyní se kapalinové tlakoměry používají jen pro samotné měření.

Deformační tlakoměry jsou membránové a nebo s Burdonovou pružinou, která je známá z mechanických tlakoměrů s ručičkou. Měřený tlak se převádí například některou variantou uspořádaní membrány na výchylku nebo sílu, která se měří dříve popsanými postupy. Varianty uspořádání membrány jsou znázorněny na obrázku 3.48. Charakteristika membrány je závislost prohnutí na tlaku. První uspořádání bez vlastní charakteristiky představuje spojení desky přes pružnou pryžovou manžetu ke stěně komory. Charakteristika membrány je dána pak připojenou pružinou. Ostatní způsoby řešení jsou zřejmé z popisu zmíněného obrázku.

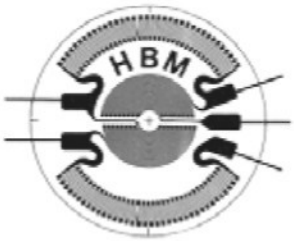


Obrázek 3.49 – Provedení membrán

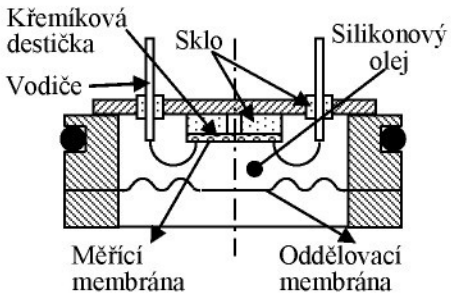
U membrány s vlastní charakteristikou lze snímat mechanické napětí tenzometrickou růžici nebo čtyřmi tenzometry, z nichž dva jsou aktivní (ve středu) a dva pasivní (u okraje) pro kompenzaci vlivu teploty. Zvláštní řešení je u membrán z polovodiče (monokrystalický křemík), které obsahují přímo vyleptané odpory ve funkci tenzometrů. V tomto případě je tlak převeden na sílu, která se měří tenzometry. Způsoby vyhodnocení mechanického napětí jsou zobrazeny na obrázku 3.49.



Umístění tenzometrů na membráně



Tenzometrická růžice firmy HBM



Polovodičový tlakoměr

Obrázek 3.50 – Umístění tenzometrů

Snímače tlakové diference jsou konstruovány pro malý rozdílový tlak. Nejlépe vyhoví membrána bez vlastní charakteristiky s převodem na sílu, kterou lze měřit například na kompenzačním principu podle obrázku 3.37.

Snímače hladiny pro regulaci

Hladinoměry rozdělujeme na

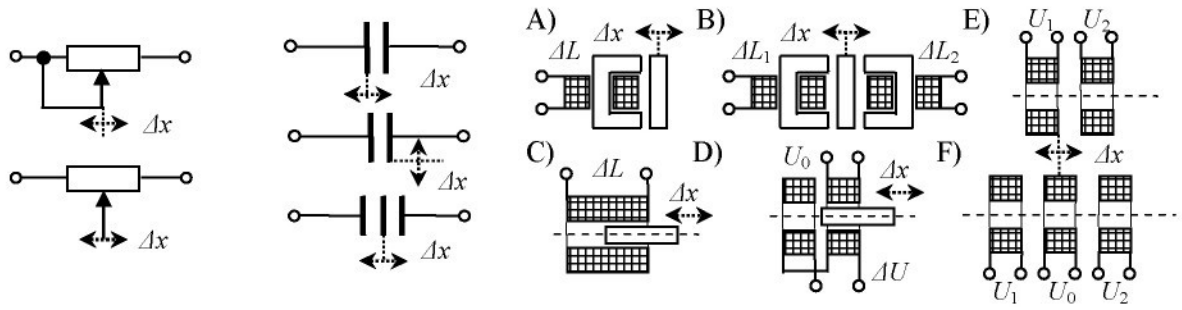
- Plovákové pro kapaliny s převodem na odměřování dráhy například potenciometrem.
- Kapacitní s převodem na změnu kapacity, která se vyhodnocuje elektronicky.
- Hydrostatické pro kapaliny s převodem na měření hydrostatického tlaku ($p = h\rho g$)
- Ultrazvukové, které jsou založeny na principu měření doby mezi vysláním a zpětným příjmem ultrazvukového signálu odraženého od hladiny materiálu.
- Radarové se stejným principem měření jako ultrazvukové, ale místo ultrazvuku jsou použity rádiové vlny.
- Ionizační se dvěma způsoby: první měří absorbované záření a druhý vyhodnocuje vzdálenost mezi zdrojem a detektorem záření.

Snímače polohy pro regulaci s analogovým výstupem

Snímačů polohy je velké množství druhů. Snímače využívají změny elektrického odporu, indukčnosti nebo kapacity, jak je znázorněno na obrázku 3.49. Odporové snímače mohou měnit velikost odporu ve vztahu k poloze nebo posunutí nebo představují dvojici odporů v uspořádání, které tvoří potenciometr. Oba odpory odděluje jezdec potenciometru. Mechanický pohyb je v tomto případě spřažen s tímto jezdcem. Vyhodnocovací metody jsou již popsané můstkové metody a tzv. výchylkové metody, u kterých odpor určuje proud v obvodu a jeho stupnice je ocejchována v posunutí.

Uspořádání snímačů založených na změně kapacity je rovněž uvedeno v obrázku 3.50. Jsou zde znázorněny pohyby elektrod ve dvou směrech a pohyb vložené elektrody u diferenciálního elektrického kondenzátoru. Další varianty uspořádání lze odvodit z pohybu dielektriky mezi elektrodami. Vyhodnocovací elektronické obvody používají již popsaný můstek nebo vlivu proměnné kapacity kondenzátoru ve zpětné vazbě zesilovače, ve kterém

jsou odpory nahrazeny elektrickými kondenzátory, a nebo rezonanční metody využívající změny rezonanční frekvence rezonančního obvodu.



Odporový snímač | Kapacitní snímače | Indukční snímače

Obrázek 3.51 – Metody měření polohy převodem na změnu elektrického odporu, kapacity a indukčnosti

Indukční snímače na obrázku 3.50 lze rozdělit na snímače

- s uzavřeným magnetickým obvodem (ad A) a ad B))
- s otevřeným magnetickým obvodem (ad C) a ad D))
- snímače bez feromagnetika (ad E) a ad F))
- snímače s potlačeným magnetickým polem (aplikace vířivých proudů).

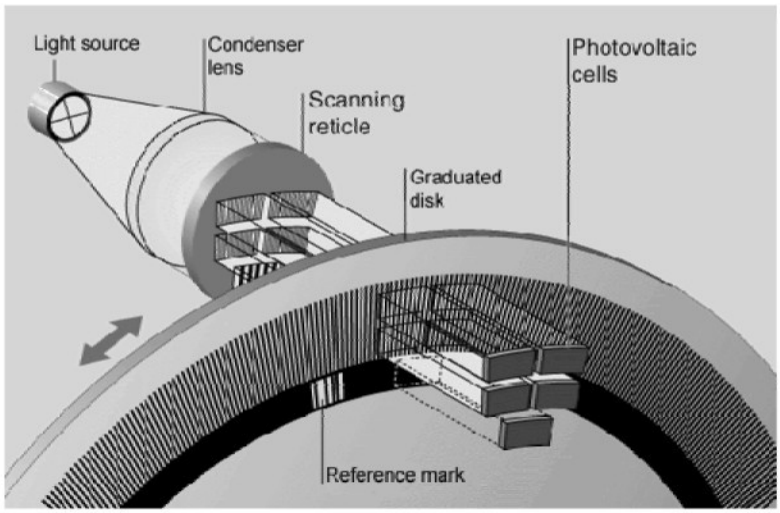
U indukčních snímačů s feromagnetikem v magnetickém obvodu je indikován posun části jádra cívka. Vzduchová mezera se může rozšiřovat nebo zužovat, ale také se může měnit překrytí průřezu jader. U snímačů bez feromagnetika se vzájemně posouvají samotné cívky, což ovlivňuje jejich indukční vazbu. Zapojení indukčních snímačů ad D) a ad F) představuje diferenciální transformátor (u zapojení ad F) je třeba spoje cívek s výstupem U_1 a U_2 doplnit.

Elektrické zapojení vyhodnocovacích obvodů se opírá principiálně o stejné zapojení jako v případě kondenzátorů. Podstatné je, že obvody pracují se střídavým proudem, zatímco odporový potenciometr dovoluje použít i stejnosměrné napájecí napětí.

❑ Snímače polohy pro regulaci s číslicovým výstupem

Analogové snímače nejsou vhodné pro velmi přesné měření jaké je požadováno například u obráběcích strojů. Základem pro velmi přesné odměřování polohy jsou optické snímače úhlového natočení a snímače lineárního pohybu, které využívají převod polohy přímo na číslicový údaj. Bez ohledu na konkrétní fyzikální princip lze tyto snímače rozdělit na snímače na

- inkrementální
- absolutní.

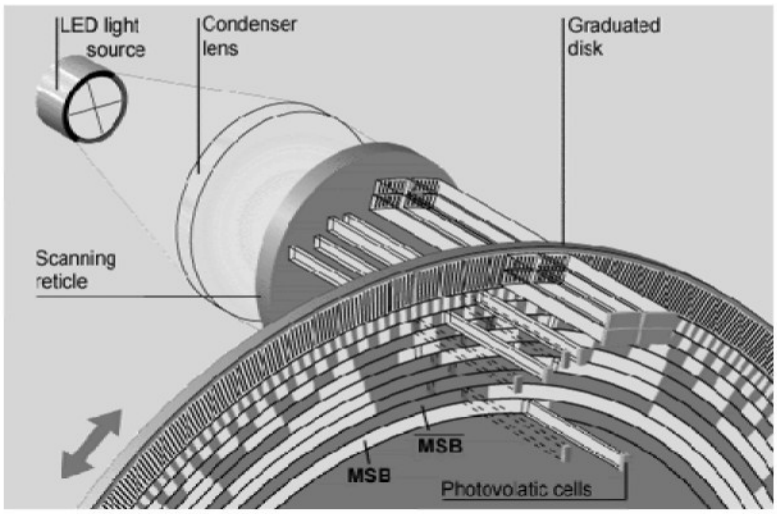


Obrázek 3.52 – Inkrementální optický snímač natočení

Inkrementální optické snímače například firmy Heidenhein obsahují zdroj světla (light source) jehož paprsky procházejí kondenzorovou (condenser lens) čočkou, přičemž vytvoří svazek rovnoběžných paprsků. Tyto paprsky procházejí optickou skenovací mřížkou (scanning reticle) a pak pohyblivým diskem opatřeným ryskami (graduated disk). Interferencí světla při otáčení disku vznikají stavy světla a tmy. Tyto změny osvětlení se vyhodnocují pomocí dvou párů fotobuněk (photovoltaic cells)) Optická skenovací mřížka je zhotovena tak, že vzniknou dva elektrické signály, které jsou fázově posunuty o čtvrť periody signálu. Vzájemný fázový posun určuje smysl otáčení kotouče. Elektronické zařízení dále dělí periodu signálu, takže rozlišení je od dva řády jemnější. Vzdálenost rysek na mřížce je řádově 10 až 20 mikronů, zatímco elektronickým dělením lze dosáhnout až 0,01 mikronu. K určení referenční polohy je na disku ještě zvláštní značka (reference mark). Funkční schéma úhlového inkrementálního snímače je na obrázku 3.51.

Úhlové snímače, které generují řadu impulsů, mohou být použity buď k měření uhlového natočení (počet přijatých impulsů je úměrný úhlu otočení) nebo jsou určeny k vyhodnocování otáček (frekvence impulsů je úměrná úhlové rychlosti). Snímače natočení se anglicky označují angle encoders, zatímco snímače otáček rotary encoders.

U inkrementálního snímače natočení je třeba po každém zapnutí stroje najet na referenční značku. Je také nebezpečí ztráty informace o poloze. Absolutní odměřování detekuje polohu takovou, jaká právě je, protože snímá informaci z velkého množství referenčních značek ve formě dalších stop na kódovacím kotouči. Princip funkce těchto snímačů je zřejmý z obrázku 3.52. Mezi těmito značkami je snímána poloha inkrementálně. Kódování polohy je pomocí Grayova kódu, jehož výhodou je to, že sousední kódové kombinace se liší v jednom bitu, což minimalizuje případnou chybu, která je nejvýše rovna otočení kódovacího kotouče o jednu rozeznatelnou polohu. Kdyby se při tomto otočení měnilo více bitů a změna by proběhla jen u některých bitů, pak by byla chyba mnohonásobkem základního rozlišení.



Obrázek 3.53 –Absolutní optický snímač natočení

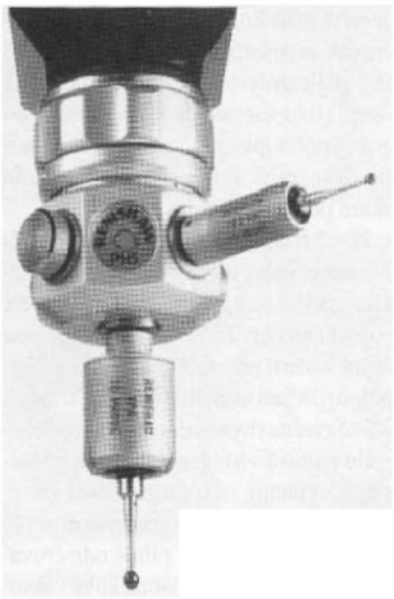
Odměřování polohy je rozdělováno na

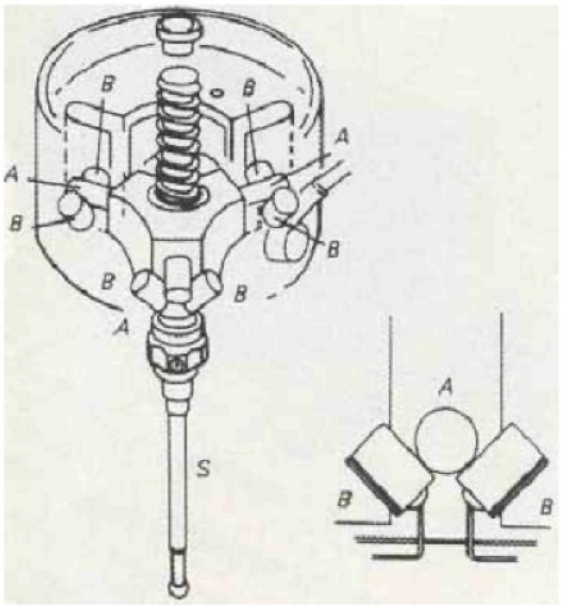
- přímé – snímač lineárního posunutí je nezávislý na pohonu, tj. je spojen s posunovanou součástí
- nepřímé – snímač úhlového natočení je na stejné ose jako servomotor pro přestavení polohy.

❑ Dotykové sondy

K přesnému odměřování polohy patří kontaktní sondy. Jsou to snímače s kontaktním výstupem. Příklad řešení takové sondy od firmy Reinishaw je na obrázku 3.53 vlevo a princip funkce je znázorněn vpravo. Kulový dotykový hrot je tvořen z otěruvzdorného materiálu. Dotykem s kontrolovaným obrobkem dojde k vychýlení hrotu a rozpojení elektrického obvodu. V tělese sondy jsou tři páry tangenciálně rozmístěných válečků B, mezi něž dosedají pod lehkým přítlakem centrální pružiny tři tangenciálně rozmístěné válečky A, které nesou vlastní dotyk S (Stylus). Všech šest válečků B je elektricky izolováno a elektricky propojeno po obvodu tak, že elektrické spojení je zprostředkováno válečky A. Obvodem prochází nepatrný proud. Po dotyku kulového konce se může sonda vychýlit o značný úhel, což zabrání jejímu mechanickému poškození, ovšem elektrický obvod se přeruší okamžitě po dotyku a nadzvednutí některého válečku A. K přerušení dojde i při axiálním dotyku sondy. Elektrické spojení se obnoví okamžitě po ztrátě dotyku s obrobkem. Sonda se tedy používá k indikaci dotyku kulového konce s povrchem obrobku při odměřování vzdáleností.

Dotykové sondy se používají v třísouřadnicových zařízení pro kontrolu rozměrů výrobků.





Provedení sondy fy Reinishaw Princip funkce kontaktní sondy

Obrázek 3.54 – Princip funkce kontaktní sondy

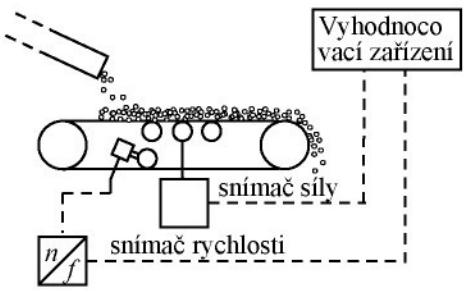
❑ Průtok sypkých hmot

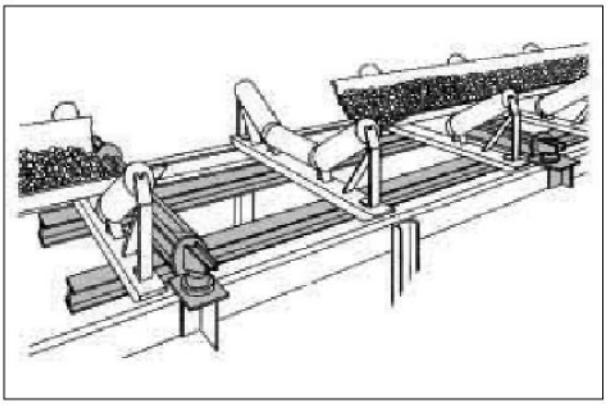
Regulace průtoku sypkých hmot a materiálů má široké uplatnění při vytváření směsí nebo dodávku materiálu do technologického zařízení. Příkladem může být zavážení paliva do kotle. Princip měření spočívá v umístění jednoho nebo skupiny více válečků dopravníku na snímače váhy. Součin této váhy a rychlosti pásu představuje průtočnou hmotu.

Pro zatížení, které je přepočteno pomocí gravitačního zrychlení g na hmotu o velikosti $q = F/g$ [kg] na úseku pásu délky l [m], který se pohybuje rychlostí v, je přepravované množství v [kg/s]

$$P = \frac{qv}{l} = \frac{ql}{lt} = \frac{q}{t}. \qquad (3.13)$$

Snímače síly jsou obvykle tenzometry uspořádané do tenzometrického můstku, který býval dříve napájen napětím z tachodynama. Nyní je toto zařízení řešeno číslicově, a proto se rychlost, která je stanovená impulsním tachodynamem, násobí se zatížením číslicově, protože zaručuje zvláště pro rychlost pohybu pásu přesnější měření. Příklad provedení jedno a víceválečkové pásové váhy je na obrázku 3.54.



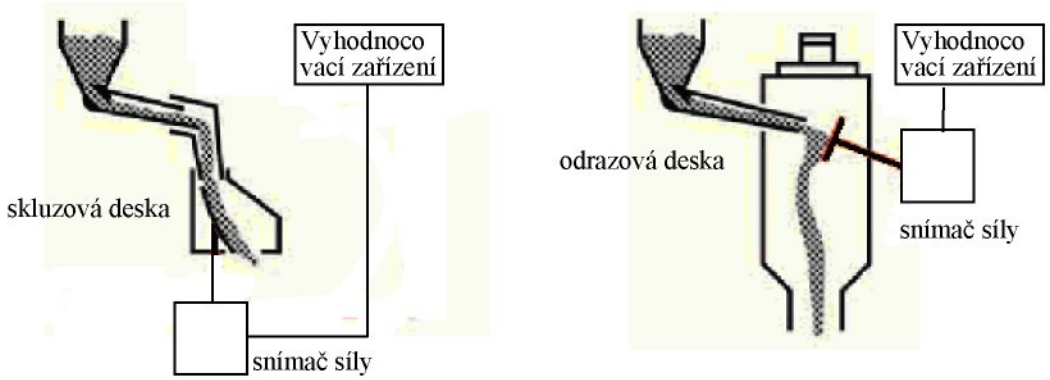


Jednoválečková váha Víceválečková váha Schenk

Obrázek 3.55 – Princip pásové váhy

Minimální provozní chyba je v rozmezí 0,5 až 1%. Maximální přepravní výkon je až 15000 t/h.

Pro případ, kdy se materiál nedopravuje materiál po páse se používají skluzové nebo odrazové desky, viz obrázek 3.55. Principem měření je úměra síly vyvozená změnou směru toku materiálu.



Obrázek 3.56 – Princip měření množství sypkých hmot skluzovou a odrazovou deskou

Otázky 3.5.

1. Kde se umístí na potrubí clona pro měření průtoku?
2. Lze umístit na potrubí clona pro měření průtoku do kolena?
3. Co se používá nejčastěji pro měření do teploty do 300 0C?
4. Při regulaci sypkých hmot a materiálů se používá odrazové desky. Jaký je její princip?

3.7. Akční členy

Čas ke studiu: 2 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- popsat akční členy, jejich rozdělení, převody pohybů, regulační orgány pro ovládání průtoku plynů, par a kapalin.
- popsat kontinuálního dávkování sypkých hmot.

Výklad

Akční členy jsou velmi závislé na aplikační oblasti řízení. Nejobecněji pojatou funkcí řízení a regulace je ovládání toku energie. Energie v podobě toku hmotných, plynných a

kapalných medií se ovládá mechanicky prostřednictvím škrticího orgánu, který mimochodem způsobuje ztráty, nebo řízením výkonu dávkovací váhy, čerpadla a nebo kompresoru. Jestliže tok energie nemá hmotnou podobu, jako je tomu například v případě elektrické energie, je ovládáno napětí nebo proud regulovaného zdroje. Také zde může být regulace toku ztrátová nebo bezeztrátová. Obecná charakteristika funkce řízení jako toku energie platí i v takovém případě, kdy jde o řízení polohy nebo síly, resp. tlaku. Systému je třeba dodat energii pro přestavení polohy nebo zvýšení (snížení) tlaku.

V učebním textu, který je určen studentům strojního inženýrství, je pozornost zaměřena na kinematiku (poloha) nebo dynamiku pohybu (síla). Může jít o pohyb akčních orgánů (ventily, klapky) nebo také regulované veličiny (poloha, rychlost, otáčky, síla, tlak). Do pohybu uvádějí různá mechanická zařízení pohony. Pohony lze rozdělit podle formy spotřebované energie takto

- elektrický
- hydraulický
- pneumatický
- piezoelektrický (místo pohon je používáno označení aktuátor).

Mechanické pohony mohou vykonávat dva druhy pohybu

- přímočarý
- otáčivý

Cílem řízení u otáčivého pohybu pohonů může být dosažení zadaných otáček nebo otočení. U přímočarých pohonů je cílem řízení dosažení žádoucího posunu nebo rychlosti pohybu. Někdy se také dělí řízení mechanických systémů na polohové nebo rychlostní.

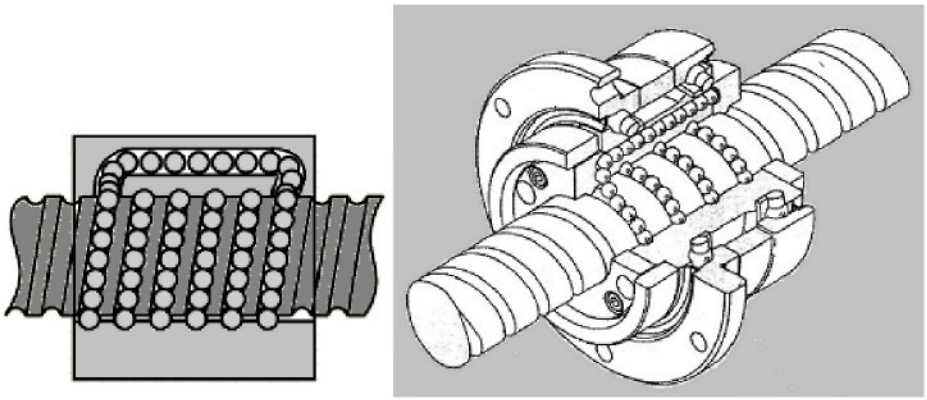
❑ Převody pohybů

Pohony pro automatizaci mají určité charakteristiky. Elektrické pohony lze konstruovat a dimenzovat snadno pro vysoké otáčky (výjimka jsou krokové motory). Často je žádoucí pohyb přímočarý. Pro mechanické systémy proto připadají v úvahu dva druhy převodů

- otáčivý pohyb na přímočarý
- otáčivý pohyb o vysokých otáčkách na otáčivý pohyb na nízkých otáčkách

Kuličkové šrouby

Tento strojní konstrukční prvek slouží k převedení rotačního pohybu na přímočarý. Plní funkci například posuvů obráběcích strojů. Je navržen tak, aby neměl prakticky žádnou vůli. Jeho konstrukce je zřejmá z obrázku 3.56. Kontakt mezi šroubem a otočnou maticí není třecí, ale odvalovací a je zprostředkován kuličkami.

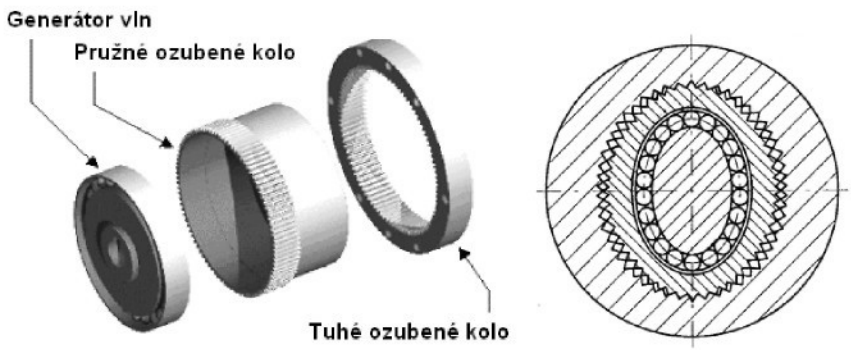


Obrázek 3.57 – Kuličkový šroub

Harmonické převodovky

K demonstraci funkce je vybrána harmonická převodovka firmy TG Drivers (www. TG-Drivers.cz), viz obrázek 3.57. Tyto převodovky typu jsou zvlášť vhodné pro aplikace vyžadující velké převodové poměry mezi motorem a poháněným zařízením prostřednictvím jednoho převodového stupně. Lze je použít k pohonu ramen robotů. Základními částmi převodovky jsou: pružné ozubené kolo, tuhé ozubené kolo a generátor vln (unášeč). Tuhé ozubené kolo má o dva nebo tři zuby více než pružné ozubené kolo. Pružné ozubené kolo s vnějším ozubením a počtem zubů například o dva menším než tuhé ozubené kolo s vnitřním ozubením je deformováno přes eliptický generátor vln a jeho zuby zapadají do zubových mezer v místě hlavní osy elipsy. Při otočení generátoru vln o jednu otáčku se posune pružné kolo o dva zuby vůči tuhému kolu v opačném směru. Pružné ozubené kolo má dno zesílené přírubou sloužící k upevnění výstupního hřídele převodovky. Tuhé ozubené kolo umožňuje montáž do skříně převodovky nebo na přírubu servomotoru. Generátor vln se montuje na hřídel servomotoru. Předností harmonických převodovek jsou následující

- velký převodový poměr (50 - 276)
- malé rozměry a nízká hmotnost, vysoká torzní tuhost
- vysoká opakovatelná přesnost, dlouhá životnost
- nízká vůle do 1 a 3 úhlových minut
- vysoká účinnost 75 až 85%.

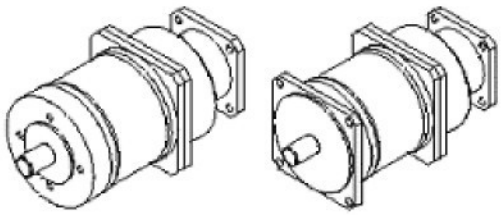


Obrázek 3.58 – Princip harmonické převodovky

Planetové převodovky

Planetové převodovky jsou konstruované pro použití v aplikacích pro řízení polohy. Je výhodné je využít při požadavku menších převodových poměrů (do $i = 50$). Redukují otáčky synchronního motoru, řízeného měničem frekvence. Přesným provedením převodovek je zaručena malá vůle, rovnoměrný chod a velká účinnost. Z hlediska vůlí se tyto převodovky vyrábějí v přesnostech od 1 do 20 úhlových minut. Uložení výstupní hřídele umožňuje radiální i axiální zatížení. Přesné a tvrdé provedení ozubení snižuje třecí momenty a umožňuje zmenšit rozměry ozubených kol, což má za následek snížení momentu setrvačnosti převodovky. Příklad jejich provedení je na obrázku 3.58.

Provedení čelní příruby



Obrázek 3.59 – Planetové převodovky firmy TG Drivers

- **Regulační orgány pro ovládání průtoku plynů par a kapalin**

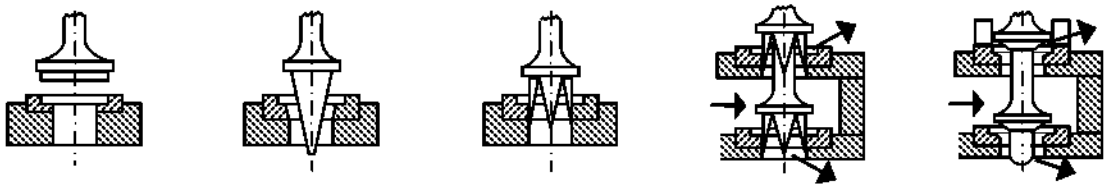
K ovládání průtoku uvedených médií se používají konstrukční prvky

- ventily (uzavírací, regulační, třícestné směšovací, pojistné zpětné a škrtící)
- kohouty (válcový, kuželový, kulový)
- šoupátka
- klapky (klapka jednoduchá a protisměrná, také křídlový ventil)
- žaluzie.

Ventil se skládá z tělesa, sedla, kuželky (podle tvaru talířové, jehlové, válcové a dvousedlové, dříku a těsnění). Charakteristickým znakem ventilu, že v něm dochází ke změně směru proudění a není tedy přes něj vidět a nelze jím prostrčit rovný drát. Ukázky konstrukce kuželek je na obrázku 3.59. Jehlové ventily slouží k jemné regulaci průtoku. Pracovním prvkem je zabroušená jehla, zasouvaná do kalibrovaného otvoru. Talířové ventily pracují s plochým talířem, zapadajícím do obrobeného osazení v tělese ventilu. Ventil je schopen udržet velmi vysoké tlaky, využíván je pro svou samotěsnící schopnost ve spalovacích motorech. Variantou talířových ventilů je pružně poddajná membrána. Používají se mimo jiné v konstrukcích dýchacích přístrojů a respirátorů či v automobilovém průmyslu [http://cs.wikipedia.org/wiki/Ventil].

Kuličkové ventily se používají k zamezení zpětného toku pracovního média. Pracují s kuličkou, unášenou v pracovním směru proudem média. Při obrácení toku je kulička médiem vtlačena do osazení a zastaví tak průtok nežádoucím směrem. Tento ventil je vhodný na kontrolu zapnutí/vypnutí.

Podle funkce lze ventily rozdělit na uzavírací, regulační, třícestné směšovací, pojistné zpětné a škrtící. Hlavní funkce uzavíracích ventilů je těsnit, například plyn, proti vyššímu tlaku. Do skupiny uzavíracích ventilů patří také solenoidové ventily ovládané elektrickým proudem. U regulačních ventilů je přípustná netěsnost po uzavření 1% u jednosedlového ventilu a 5% u dvousedlového ventilu.



kuželka talířová | kuželka jehlová | kuželka válcová | kuželka válcová dvousedlová | kuželka dvousedlová

Obrázek 3.60 – Konstrukce kuželek ventilů

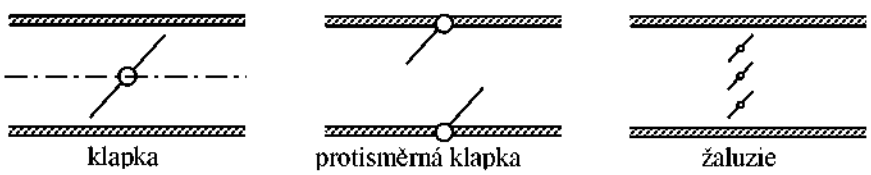
Třícestné ventily jsou rozdělovací a směšovací. Používají se v chemickém průmyslu a v tepelné technice pro regulaci teploty.

Zpětné ventily mají rozdílnou charakteristiku při změně směru proudu. Účel škrtících ventilů je obdobný elektrickému odporu.

Kohouty se uzavírají pootočením otočné části. Patří mezi uzavírací armatury, protože jsou velmi těsné. Podle tvaru otočného tělesa se dělí na válcové, kuželové, a kulové. Pracovní částí kohoutu je obrobený a provrtaný válec, koule nebo kužel, otáčející se v zalapovaném prostoru s přívodním a vývodovým potrubím umístěným v ose vrtání. Kohout může pracovat s vysokými tlaky při relativně malé potřebné ovládací síle. Z tohoto důvodu jsou často využívány u energetických zařízení. Tento ventil je vhodný na regulaci toku. Kohout lze snadno rozebrat a vyčistit, proto bývá využíván ve výrobních technologiích a chemických aparaturách.

Šoupátko je rovněž uzavírací armatura typu stavidla, které se pohybuje přímočaře.

Klapka je umístěná v průřezu potrubí a je otočná. Jsou vyráběny jako jednoduché klapky, protisměrné klapky a nebo žaluzie s více listy. Žaluzie a protisměrná klapka je vhodná v potrubích s obdélníkovým průřezem, které je používáno v klimatizaci. Tyto typy regulačních orgánů jsou znázorněny na obrázku 3.60.



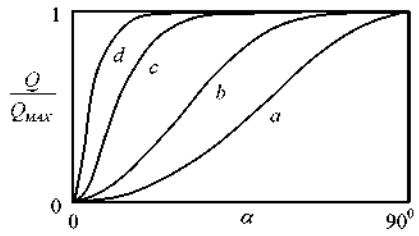
klapka | protisměrná klapka | žaluzie

Obrázek 3.61 – Regulační orgán pro průtok vzduchu a plynu

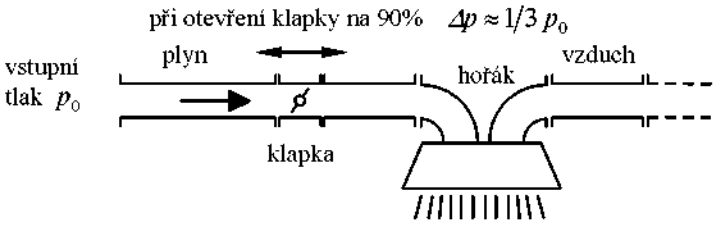
O funkci regulačního obvodu rozhoduje charakteristika regulačního orgánu, kterým je závislost průtočného množství na jeho poloze, nejčastěji úhlu natočení α. Průtočné množství se udává relativní vzhledem k průtočnému množství při plně otevřeném průtočném průřezu. Průtočné množství je přirozeně závislé na vstupním tlaku. Na obrázku 3.61 vlevo je znázorněn příklad této charakteristiky pro klapku v potrubí. Nejvýhodnější pro regulaci je

průtok přímo úměrný natočení klapky. Nejblíže lineární charakteristice je charakteristika *a*, zatímco charakteristika *d* využívá regulačního rozsahu natočení jen z části. Projektant regulačního obvodu musí správně dimenzovat škrtící orgán. Je vhodné, aby závislost velikosti průtoku na stupni otevření (reprezentováno posunem nebo natočením) byla v co největším rozsahu lineární. Například pro plyny platí, že při téměř otevřeném regulačním orgánu (např. na 90%) má být tlakový spád na tomto škrtícím orgánu asi třetinou vstupního tlaku jak je znázorněno na obrázku 3.61 vpravo. Špatně dimenzována škrtící klapka je například tehdy, jestliže při zavírání se do stupně otevření 10% průtok téměř nezmění a klapka začne regulovat průtok v posledních 10% procentech rozsahu nastavení úhlu otevření před uzavřením. Při projektování regulačního obvodu je třeba se pečlivě seznámit s doporučeními, které výrobce různých škrtících klapek nebo žaluzií doporučují pro jeho správnou funkci.

Jak již bylo zmíněno, směšovací trojcestné ventily slouží k regulaci teploty topné vody v systémech ústředního topení. Příklad je uveden na obrázku 3.62. Horká výstupní voda je směšována s vratnou ochlazenou vodou a do radiátorů proudí tedy voda se sníženou teplotou. Příklad škrtící klapky pro topnou vodu Honeywell (www.honeywell.cz) je na dalším obrázku 3.63.



Charakteristika regulační armatury



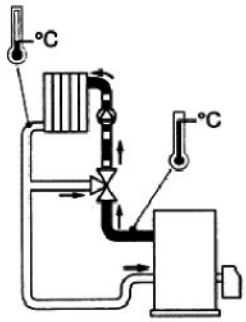
Podmínka správné funkce klaply nebo ventilu

Obrázek 3.62 – Volba vhodné regulační klapky

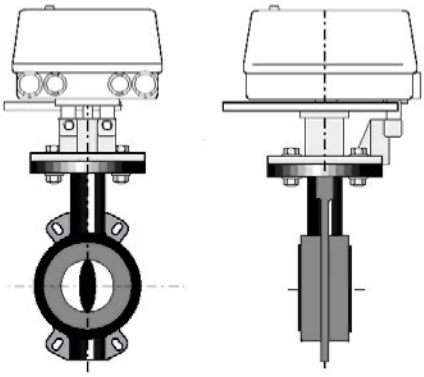
Průtok tekutiny ventilem lze popsat rovnicí (3.14), ve které k_V označuje průtokový součinitel ventilu. Průtokový součinitel je uváděn výrobci v katalogu a technickém popisu ventilu.

$$Q = k_V \sqrt{\frac{\Delta p}{\gamma}}, \quad \left[m^3 s^{-1}; Pa, kgm^{-3}\right] \tag{3.14}$$

Již výše bylo uvedeno, že podmínkou správné funkce klapky je vhodná tlaková ztráta při otevřené poloze. Diagram závislosti tlakové ztráty pro různý stupeň otevření a objemovém průtoku na rychlosti proudění lze najít k katalogu dodavatele těchto regulačních prvků. Z diagramu se lze dovědět, jaký lze očekávat průtok nebo rychlost proudění kapaliny pro danou tlakovou ztrátu a úhel otevření škrtící klapky.



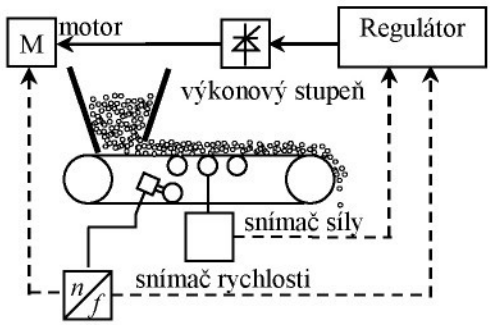
Obrázek 3.63 – Směšovací třícestný ventil u kotle ústředního topení



Obrázek 3.64 – Škrtící klapka pro topnou vodu, výrobek firmy Honeywell

❑ Kontinuální dávkování sypkých hmot

Zkrácené označení pro zařízení, které zajišťuje kontinuální tok sypkého materiálu ze zásobníku, je dávkovací váha. Toto zařízení je základem poměrové regulace průběžného dávkování směsi. Vynášení směsi ze zásobníku se řeší trojím způsobem, a to vibračními podávači, kde vynášený tok směsi je dán amplitudou vibrací, dále talířovými a pásovými vynášeči. U talířového podávače je pod zásobník umístěn rotující plochý talíř se svislou osou mimo osu otvoru zásobníku. Vynesená směs se pak shrnuje škrabkou.



Obrázek 3.65 – Princip pásové dávkovací váhy

Otáčky talířového podávače a rychlost vynášecího pásu spolu s velikostí štěrbiny mezi dolním okrajem zásobníku a pohybující se plochou určují vynášené množství komponenty směsi.

Poslední uvedený způsob dávkování, nyní již dominujícím, využívá krátkého podávacího pásu, jehož rychlost a případně velikost štěrbiny určuje dávkované množství. Princip funkce pásové dávkovací váhy je na obrázku 3.64.

Průběžná pásová dávkovací váha obsahuje jeden váleček, jehož svislé zatížení je snímáno a spolu s rychlostí pohybu pásu určuje vynášené množství dávkovaného sypkého materiálu. Vynášené množství je ovládáno změnou rychlosti pohybu pásu. Regulační smyčka pak změnami otáček pohonu vynášecího pásu udržuje zadaný průtok sypkého materiálu. Přesnost dávkování je shodná s přesností pásové váhy.

? Otázky 3.6.

1. K čemu je kuličkový šroub?
2. K čemu slouží harmonická převodovka?
3. K čemu slouží žaluzie?

4. Jaká je podmínka správné funkce klapky?

3.8. Elektrické pohony

Čas ke studiu: 3 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- popsat základní typy elektrických pohonů, pracovní oblasti, charakteristiky motoru, oblasti použití.
- charakterizovat základní prvky silnoproudé elektroniky sloužící k napájení pohonů a k jejich řízení.

Výklad

Tato kapitola obsahuje všeobecný popis vlastností elektrických pohonů včetně prvků silnoproudé elektroniky, které slouží k napájení pohonů a k jejich řízení.

Tranzistor ve spínacím režimu

Elektrický spínač uzavírá elektrický obvod pro průchod elektrického proudu. Je známo, že k sepnutí nebo k rozpojení spínače, například světelného obvodu v domácnosti, je třeba vnějšího impulsu. Pro větší proud budou na rozdíl od světelného vypínače kontakty spínače masivní a navíc může být výhodné, aby jeho sepnutí bylo dálkově ovladatelné. Řešení nabízí stykač, který byl již popsán v kapitole o logickém řízení.

Mechanické kontakty stykače se mohou opalovat obloukem, který vzniká ve fázi jejich přiblížení nebo oddálení před sepnutím nebo po rozepnutí zátěže indukčního charakteru. Jejich životnost není nekonečná. Je tedy výhodné používat bezkontaktní spínání polovodičovými prvky.

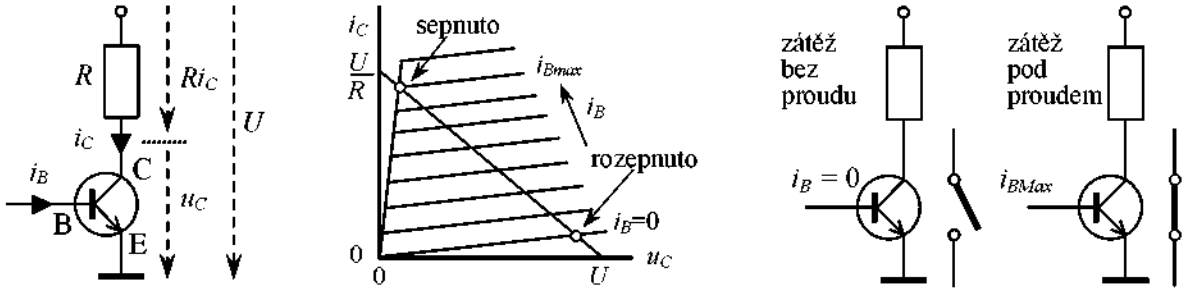
Nejjednodušší řešení poskytuje obyčejný tranzistor. Proud mezi kolektorem a emitorem u tohoto prvku je ovládán mnohem menším proudem v obvodu báze – emitor. Funkce bude vysvětlena na příkladu použití tranzistoru NPN, který je zapojen podle schématu na obrázku 3.65 vlevo. Charakteristika tranzistoru je funkce kolektorového proudu i_C na proudu bází i_B a napětí na kolektoru u_C. Platí

$$i_C = f(i_B, u_C). \tag{3.15}$$

Charakteristika tranzistoru je znázorněna v obrázku 3.66 uprostřed. Ve skutečnosti je to celá skupina křivek pro diskrétní hodnoty proudu bází. Linearizace poslední rovnice ukazuje, že mezi proudem báze a kolektorem platí přibližně $\Delta i_C = \beta \Delta i_B$, kde β je součinitel zesílení proudu. Mezi napájecí napětí o velikosti U a kolektor C je zapojen odpor o velikosti R. Napětí mezi kolektorem a emitorem (vývody tranzistoru C a E) je následující

$$u_C = U - R\, i_C. \tag{3.16}$$

V souřadnicích charakteristiky tranzistoru představuje poslední rovnice přímku, která protíná síť charakteristik tranzistoru a nazývá se zatěžovací. Velikost proudu bází je obvodu vnucena zvenčí, proto napětí mezi kolektorem a emitorem a proud kolektorem tvoří dvě neznámé veličiny soustavy dvou posledních rovnic. Jedna z těchto rovnic je lineární rovnice (3.16) a druhá je nelineární rovnice (3.15). Grafickým řešením této soustavy jsou průsečíky, tzv. pracovní body, charakteristik tranzistoru a zatěžovací přímky.



Tranzistor NPN | Charakteristika tranzistoru a zatěžovací přímka | Tranzistor jako spínací obvod

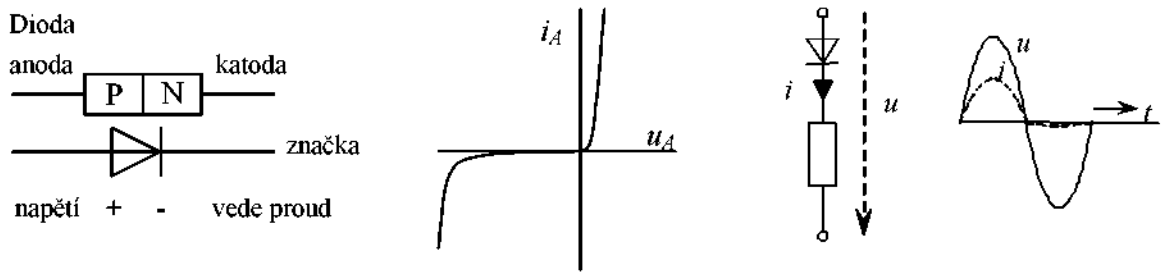
Obrázek 3.66 – Značka tranzistoru NPN, charakteristika a funkce ve spínacím režimu

Z množství pracovních bodů jsou ve spínacím režimu zajímavé pouze dva, a to pro nulový proud bází a pro proud bází, kterým jsou přechody tranzistoru nasyceny. Příslušné stavy jsou znázorněny na obrázku 3.65 vpravo. Při nulovém proudu bází prochází tranzistorem velmi malý proud a napětí na tranzistoru se blíží k napájecímu napětí. Zátěž není téměř pod proudem a chová se jako při rozepnutém ekvivalentním spínači. Druhý stav odpovídá proudu zátěží o velikosti téměř U/R a velmi malém úbytku napětí mezi kolektorem a emitorem. Tento stav odpovídá sepnutému ekvivalentnímu kontaktu. Teoreticky proud báze při sepnutí stačí takový, který odpovídá pracovnímu bodu sepnuto z obrázku 3.65. Pro bezpečnost funkce je vhodné volit proud báze poněkud větší. Pomůckou je parametr β tranzistoru. Příliš velký proud může prodloužit dobu rozepnutí u velmi rychlých obvodů.

Polovodičová dioda

Dioda je elektronický prvek, který za určitých podmínek vede nebo nevede elektrický proud. Princip funkce je založen na vlastnostech přechodu P-N, který je vytvořen z polykrystalického (selén) materiálu nebo monokrystalů (křemík nebo germanium). Tento materiál je slabě dotován nečistotami různého typu, které ve svém důsledku vyvolávají tzv. děrovou (P) nebo elektronovou vodivost (N). Základní vlastnosti polovodičové diody a obvodová značka jsou znázorněny na obrázku 3.68 vlevo. Volt-ampérová charakteristika diody, tj. závislost okamžitých hodnot proudu na napětí anoda-katoda, je na obrázku 3.66 uprostřed. Pro otevřenou diodu (kladné napětí na anodě) vzniká na diodě malý úbytek napětí a diodou protéká proud. Germaniové diody mají úbytek 0,4 V, zatímco křemíkové diody 0,6 až 0,8 V. Při uzavřené diodě (anoda záporná) nevede proud až do závěrného napětí, které je pro germaniové diody 100 V a křemíkové diody až 4000 V. Diody mohou být dimenzovány až na stovky ampér.

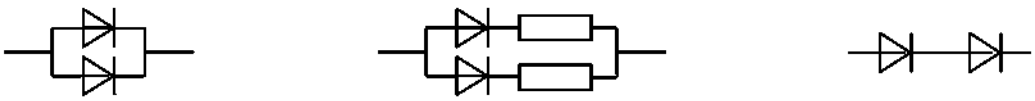
Dioda má usměrňovací efekt pro střídavý proud. Závislost časového průběhu napětí a proudu je znázorněna na obrázku 3.66 vpravo. Velikost proudu nelze ovládat.



Princip činnosti diody | Charakteristika diody | Usměrňovací efekt diody

Obrázek 3.67 – Značka diody, volt-ampérová charakteristika a usměrňovací efekt

Zvětšení proudu nebo závěrného napětí se řeší zapojeními, které jsou znázorněny na obrázcích 3.67. Pro přerozdělení proudu u diod lišících se svou charakteristikou se použijí přídavné sériové odpory.



Paralelní zapojení pro zvýšení proudu | Paralelní zapojení pro zvýšení proudu a přerozdělení úbytku napětí | Sériové zapojení pro zvýšení závěrného napětí

Obrázek 3.68 – Zapojení pro zvětšení proudu nebo závěrného napětí

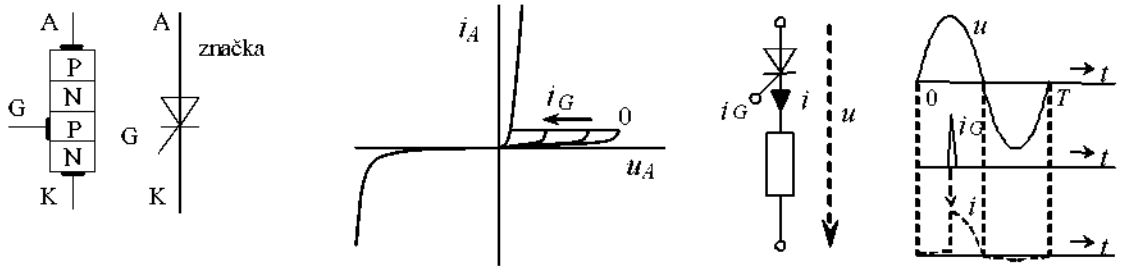
Tyristor

Tyristor je prvek, u kterého lze na rozdíl od diody ovládat propustnost proudu. Jde o sériové spojení dvou přechodů P-N, viz. obrázek 3.68 vlevo. K jednomu z přechodů je připojena elektroda G (Gate - hradlo). Volt-ampérová charakteristika tyristoru je na obrázku 3.68 uprostřed. Parametrem charakteristiky je zapalovací proud i_G. Impuls proudu přivedený na tuto elektrodu ve stavu s kladným napětím na anodě otevře tyristor k průchodu proudu, který potrvá až do okamžiku změny napětí na anodě na záporné. Tímto způsobem lze ovládat velikost proudu zátěží jak je znázorněno na obrázku 3.68 vpravo.

Velikost proudu se hodnotí nejčastěji jeho efektivním hodnotou I_{ef}, což je druhá odmocnina střední hodnoty kvadrátu proudu $i(t)$ za dobu jeho periody T. Efektivní hodnota proudu je dána integrováním druhé mocniny proudu podle času. Jen u střídavého proudu se sinusovým průběhem je efektivní hodnota vypočtena dělením jeho amplitudy odmocninou ze dvou.

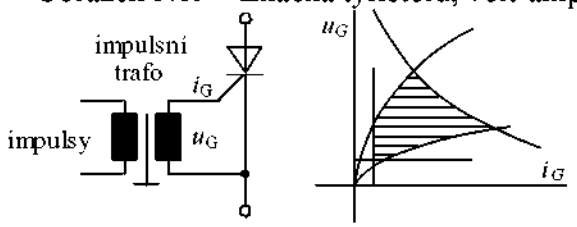
$$I_{ef} = \sqrt{\frac{1}{T}\int_0^T (i(t))^2 dt} \qquad (3.17)$$

Spuštění proudu tyristorem zajišťuje tzv. zapalovací obvod, který je znázorněn na obrázku 3.69. Z důvodu galvanickému oddělení se ke katodě a elektrodě G připojí sekundární vinutí impulsního transformátoru. Pro zjednodušení je obvodová značka pro vinutí transformátoru podobná odporu, ovšem odlišuje se vnitřním vybarvením. Spouštěcí napětí a proud musí vyhovovat určitému tolerančnímu poli v charakteristice i_G a u_G na tomtéž obrázku.



Princip činnosti tyristoru | Charakteristika tyristoru | Usměrňovací a řídicí efekt tyristoru

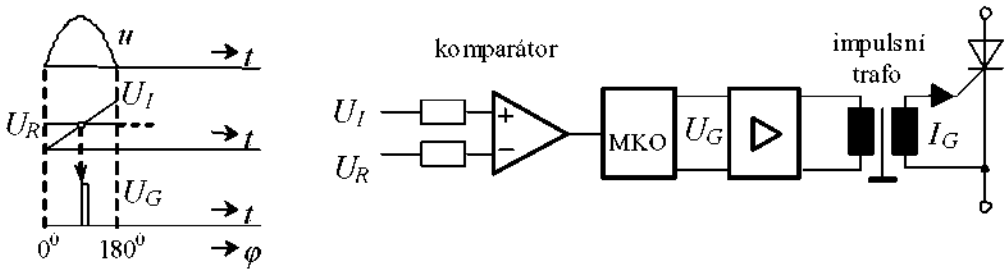
Obrázek 3.69 – Značka tyristoru, volt-ampérová charakteristika a řídicí účinek na proud



Obrázek 3.70 – Zapalovací obvod tyristoru

Úplný obvod pro generování spouštěcích impulsů je znázorněn na obrázku 3.70. Po dobu půlperiody napájecího napětí, kdy může tyristor vést proud, je generováno lineárně rostoucí napětí U_I, které je v komparátoru srovnáváno s řídicím napětím U_R. V okamžiku, kdy pilovité napětí přesáhne řídicí napětí, změní výstup komparátoru polaritu a monostabilní klopný obvod (MKO) vygeneruje napěťový impuls U_G o konstantní

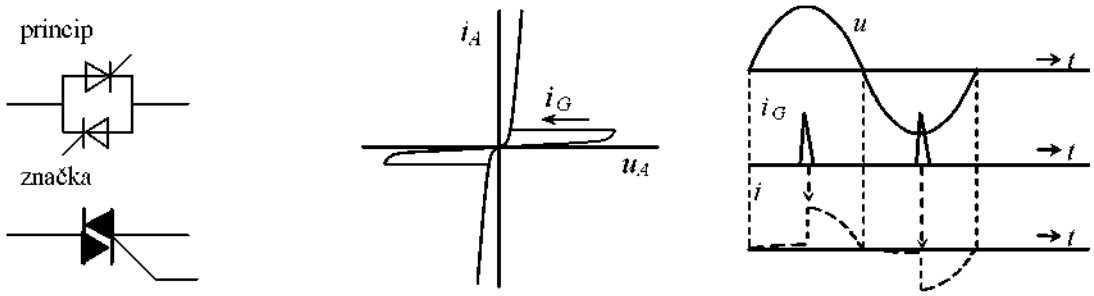
Šířce (délce časového intervalu). Napěťový impuls je zesílen a přiveden na vstup impulsního transformátoru, na jehož výstupu se objeví proudový spouštěcí impuls I_G.



Obrázek 3.71 – Obvod řízení otevření tyristoru

Triak

Funkci triaku odpovídá antiparalelní zapojení dvou tyristorů. Značka, volt-ampérová charakteristika a funkce při řízení proudu jsou znázorněny na obrázcích 3.71.



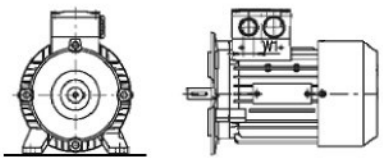
Princip a značka triaku | Charakteristika triaku | Funkce triaku

Obrázek 3.72 – Značka triaku, volt-ampérová charakteristika a řídicí účinek na proud

Přehled typů elektrických motorů

Elektrické motory mohou být děleny podle různých hledisek, například

- podle tvaru (viz obrázek 3.72)
  - patkové (montážní plocha – patky - je rovnoběžná s osou motoru)
  - přírubové (montážní plocha – příruba - je kolmá na osu motoru)
  - vestavné
- podle způsobu chlazení
  - bez chlazení (motory nejnižších výkonů)
  - vlastní ventilace pro otevřené motory (na hřídeli je ventilátor, který prohání vzduch motorem)
  - cizí ventilace (ventilátor s vlastním pohonem)
  - cizí chlazení (vodou)
- podle krytí (odolnost proti vlhkosti, prachu, apod.)
- podle napájecího napětí.
  - stejnosměrné motory (stejnosměrné napájení, ovlivňování otáček napájecím napětím a buzením)
  - střídavé motory (napájení střídavým proudem, otáčky jsou vázány na frekvenci otáčení magnetického pole a počet pólů)
  - asynchronní motory (mezi frekvenci napájecího napětí a otáčkami existuje skluz v závislosti na zatížení motoru)
  - komutátorové motory, které jsou používány například v ručních vrtačkách, bruskách, atd..
  - synchronní motory (otáčky a frekvence napájecího napětí jsou synchronní nezávisle na zatížení motoru)
  - krokové motory



patkové přírubové

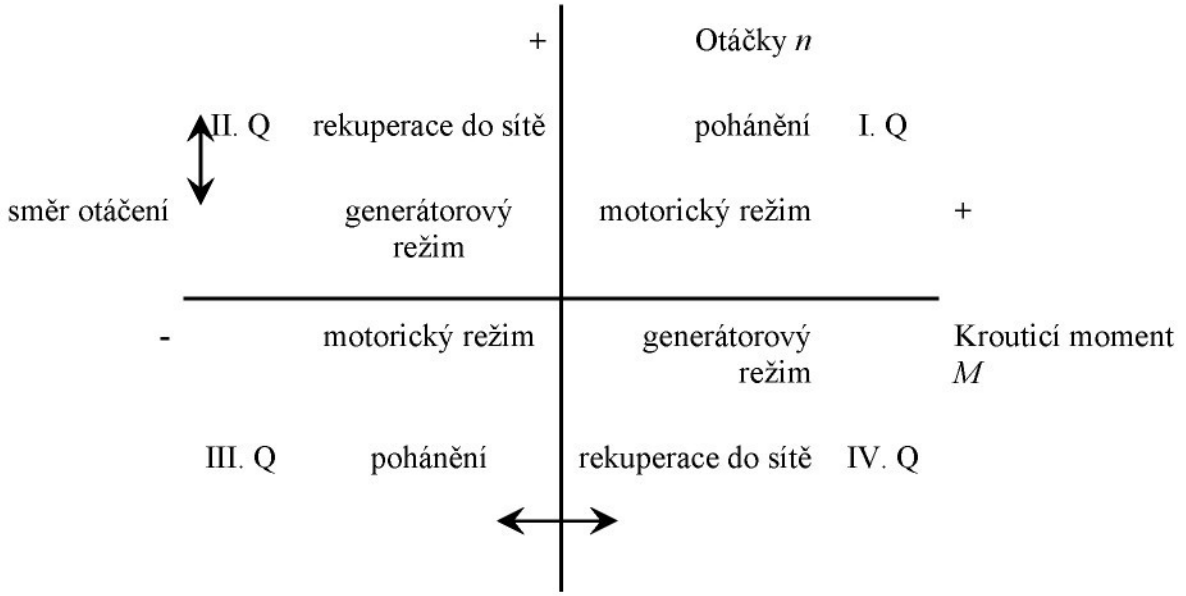
Obrázek 3.73 – Provedení elektromotorů

Napájecí napětí určuje konstrukci motoru a především jeho statické a dynamické vlastnosti. Každý elektrický pohon má své specifické použití. Vlastnosti pohonné jednotky je třeba posuzovat v komplexu napájecích obvodů a motoru.

Pracovní oblasti pohonů

Pohon může pracovat v motorickém režimu nebo v generátorovém režimu. V motorickém režimu pohání motor připojené zařízení (stroj) představující jeho zátěž, zatímco v rekuperačním režimu připojený stroj naopak pohání motor, který se jeví jako generátor, který odebírá ze stroje mechanickou energii. Tyto funkce může plnit pro jeden směr otáčení. Funkce poháněného stroje může vyžadovat reverzaci otáček, tj. chod motoru

dvěma směry. Všechny pracovní režimy motoru lze znázornit čtyřkvadrantovým diagramem. Vodorovná osa znázorňuje kroutící moment M motoru. Na svislé ose jsou otáčky motoru n. V diagramu je rozhodující znaménko zmíněných veličin.

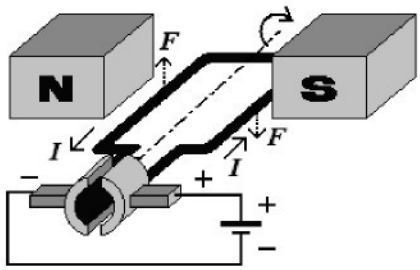


Jednokvadrantový pohon znamená, že připojený pohon má jen jeden směr otáčení a působí jako pohon (bez brzdění). Dvoukvadrantový pohon může pracovat buď v I. a IV. kvadrantu nebo v I. a II. kvadrantu. Čtyřkvadrantový pohon má pracovní režimy ze všech kvadrantů.

❑ Stejnosměrný motor

Stejnosměrný motor představuje nejjednodušší realizaci základního fyzikálního zákona o síle F působící na vodič, kterým protéká proud I a který je umístěn v magnetickém poli o indukci B. Jestliže je tento vodič o délce l orientován kolmo k siločárám pole, pak na něj působí síla $F = BIl$. Směr síly lze určit podle Flemingova pravidla levé ruky. Jestliže siločáry vstupují do dlaně a prsty ukazují směr proudu, pak palec určuje směr síly.

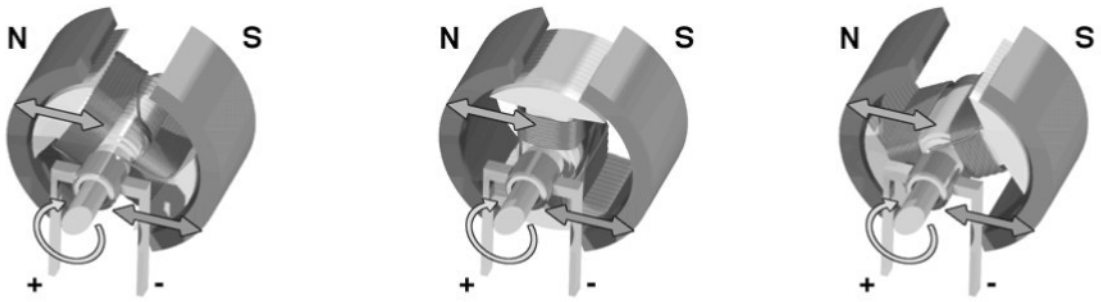
Vodič ve tvaru smyčky (součást rotující kotvy) s osou rovnoběžnou s osou rotace mění periodicky směr pohybu kolmý na siločáry, a proto je třeba ve vhodném okamžiku dvakrát za otáčku změnit směr proudu, což zajišťuje komutátor, viz obrázek 3.73. Konce proudové smyčky jsou připojeny ke dvěma shodným vodivým segmentům, ke kterým se přivádí přes kartáčky elektrický proud. V okamžiku, kdy je smyčka v rovině kolmé na magnetické siločáry, se vystřídá propojení segmentů a kartáčků, což způsobí změnu proudu v otáčející se smyčce a následnou změnu směru síly, která působí na vodiče smyčky v magnetickém poli.



Obrázek 3.74 – Funkce komutátoru

Klasický stejnosměrný motor má kotvu tvořenou elektromagnetem. Kroutící moment je vytvořen na základě vzájemného přitahování opačných magnetických pólů a odpuzování souhlasných pólů (Sever - North, Jih - South). Na obrázku 3.74 je znázorněn tento pohon ve

třech fázích natočení. Třetí fáze je kreslena těsně před okamžikem reverzace proudu komutátorem.



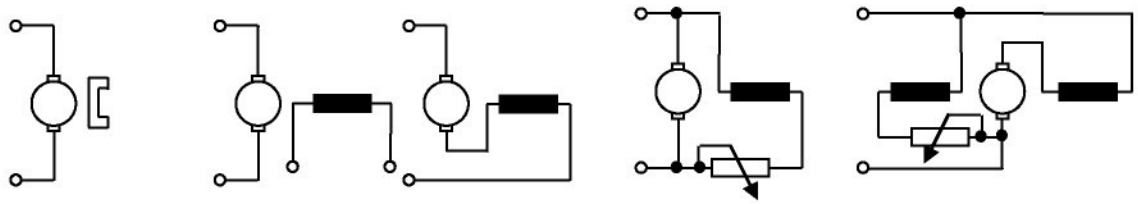
1. poloha 2. poloha 3. poloha

Obrázek 3.75 – Stejnosměrný motor s rotujícím elektromagnetem

Magnetické pole může být vybuzeno permanentním magnetem nebo cívkou napájenou stejnosměrným proudem ze zvláštního zdroje (cizí buzení) a nebo cívkou, kterou protéká proud napájející kotvu motoru. Schématicky je zapojení nakresleno na obrázku 3.75. Kromě sériového buzení může být motor buzen také paralelně (derivační buzení) nebo smíšeně (tzv. kompaudní motor). Podle budícího toku rozdělujeme stejnosměrné motory na

- motory s budícím tokem nezávislým na zatížení (buzení cizí, paralelní a permanentní magnety)
- motory s budícím tokem závislým na zátěži (sériové buzení)
- motory s buzením smíšeným (kompaundní buzení)

Permanentní magnet | Cizí buzení | Sériové buzení | Derivační buzení | Kompaudní buzení



Obrázek 3.76 – Rozdělení motorů podle buzení

Statické charakteristiky motoru popisují dvě rovnice, a to pro kroutící moment M v závislosti na magnetickém toku Φ a proudu kotvou I

$$M = k\Phi I \tag{3.18}$$

a pro elektromotorické napětí indukované na kotvě motoru v závislosti na magnetickém toku a úhlové rychlosti ω

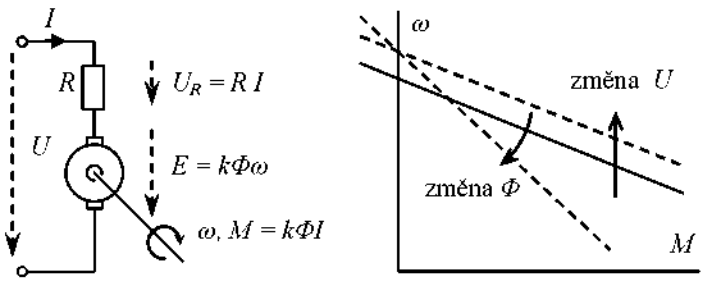
$$E = k\Phi\omega\,. \tag{3.19}$$

Při běhu bez zatížení pod napětím na kotvě způsobí náhlé odbuzení ($\Phi \to 0$) efekt, kdy se pro stálou velikost E roztočí motor ($\omega \to \infty$) až k hranici destrukce odstředivými silami (jak uvádí ve svých vzpomínkách doc. Chmelík z elektrofakulty).

Elektrické náhradní zapojení stejnosměrného motoru musí respektovat odpor kotvy R a indukčnost v obvodu kotvy. Magnetické pole vybuzené proudem, který protéká kotvou, se skládá s polem daným buzením v nezatíženém stavu motoru. Tento jev se nazývá reakce kotvy. Nepříznivý účinek na změnu směru magnetických siločar (a tím i komutaci) se kompenzuje kompenzačním vinutím, které je v sérii v obvodu kotvy. Tímto způsobem se ruší magnetický účinek proudu kotvou.

Jestliže bude motor v ustálením stavu, není třeba s indukčností počítat. Charakteristika motoru představuje závislost kroutícího momentu na otáčkách. Tuto charakteristiku lze odvodit z rovnice (3.20). Regulaci otáček motoru nebo krouticího momentu je možné změnami elektrických veličin, kterými jsou napětí na kotvě U a magnetický tok Φ. Magnetický tok u motoru s cizím buzení je dán proudem budící cívkou. U sériových motorů je budící proud totožný s proudem kotvy. Změna magnetického toku není u motorů s permanentním magnetem možná.

$$\omega = \frac{E}{k\Phi} = \frac{U - RI}{k\Phi} = \frac{U - RM/k\Phi}{k\Phi} = \frac{U}{k\Phi} - \frac{R}{(k\Phi)^2} M \,. \qquad (3.20)$$



Obrázek 3.77 – Náhradní zapojení stejnosměrného elektrického motoru a statická charakteristika motoru s cizím buzením nebo s paralelním buzením

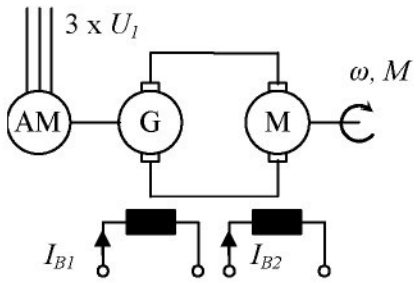
Chyba! Objekty nemohou být vytvořeny úpravami kódů polí.

Obrázek 3.78 – Charakteristika motoru se sériovým buzením

Charakteristika stejnosměrného motoru s cizím buzením nebo s paralelním buzením je přímka, jejíž sklon určuje magnetický tok, tj. proud budícím vinutí, viz obrázek 3.76. Přímka se posouvá změnou napětí na kotvě motoru. Změna buzení mění sklon charakteristiky. Otáčky motoru při tomto buzení jsou závislé na zatížení. Tento druh regulace je vhodný tam, kde se otáčky příliš nemění.

Charakteristika sériového motoru je na obrázku 3.77. Krouticí moment je nepřímo závislý na otáčkách. Její výhodou je vysoký záběrový moment pro nízké otáčky. Sériové motory mají využití u dopravních prostředků, u kterých je při rozjezdu vyvinout velký krouticí moment.

Dynamické charakteristiky elektrické části motoru musí respektovat indukčnost vinutí koty a indukčnost budícího vinutí. Poměr velikosti indukčnosti a odporu kotvy určuje elektrické časové konstanty. Mechanická časová konstanta souvisí s momentem setrvačnosti rotoru a připojeného zařízení.



Obrázek 3.79 –Ward-Leonardovo soustrojí

Regulace stejnosměrného motoru (M) vyžaduje zdroj regulovatelného napětí v minulosti se používalo dynamo (G), které bylo poháněno asynchronním motorem (AM), viz soustrojí Ward-Leonard (příjmení a jméno) na obrázku 3.78.

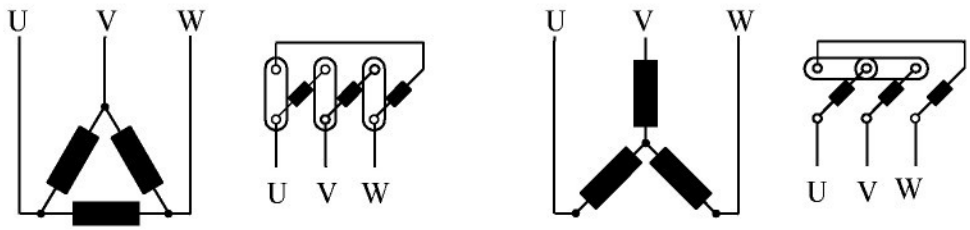
Protože otáčky asynchronního motoru jsou téměř neměnné, lze otáčky soustrojí regulovat změnou buzení jak u dynama, tak u motoru. Napětí a proud generátoru lze určit pomocí shodných rovnic jako v případě motoru. Moderní napájecí obvody obsahují jen polovodičové prvky.

❑ Střídavý asynchronní motor

Střídavý třífázový asynchronní motor je nejrozšířenějším typem pohonné jednotky. V regulaci měla dříve nevýhodu obtížná měnitelnosti otáček. Polovodičové frekvenční měniče tuto nevýhodu nyní odstraňují. Typický asynchronní motor má statorové vinutí pro třífázové napájení složené ze tří statorových cívek, které vytvářejí kruhové točivé magnetické pole. Motorek s jednou fázi by se teoreticky nemohl samostatně rozběhnout. K jeho rozběhu slouží pomocné vinutí, které je připojeno přes elektrický kondenzátor krátkodobě k fázi. Kondenzátor způsobí fázový posun, točivé magnetické pole je eliptické. Na trhu jsou také asynchronní motorky s trvale zapojeným pomocným vinutím přes kondenzátor.

Statorové cívky jsou zapojeny do hvězdy nebo do trojúhelníku. Pro určité napětí má zapojení do hvězdy 1,7krát (odmocnina ze tří) menší odběr proudu než do trojúhelníka, což se využívá všeobecně při rozběhu motoru s výkonem nad 4 kW. Nejprve se při rozběhu motor zapojí do hvězdy a po dosažení provozních otáček se přepojí do trojúhelníka.

Cívky statorového vinutí jsou zapojeny ke třem dvojicím svorek, které lze plochými propojkami trvale snadno přeskupit na oba zmíněné způsoby zapojení cívek, viz obrázek 3.79. Běžně se však používá spouštěcí otočný přepínač s polohami 0 (stop), Y (hvězda) a pak Δ (trojúhelník).



Zapojení do trojúhelníka Zapojení do hvězdy

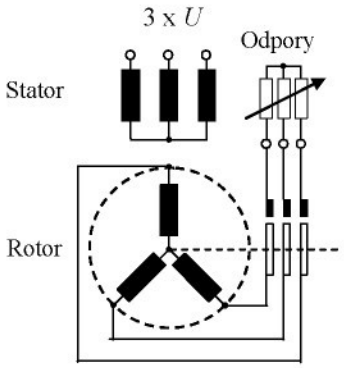
Obrázek 3.80 – Zapojení třífázového asynchronního motoru do trojúhelníka a do hvězdy

Rotor může mít dvojí provedení:

- kotva nakrátko
- vinutý, přičemž vinutí je vyvedeno z rotoru pomocí kroužků a kartáčů

Vinutý rotor má obdobně jako stator třífázové vinutí, které je založeno v drážkách. Toto vinutí se vždy spojuje do hvězdy a vzniknuvší uzel se nevyvádí ven, ale je vytvořen přímo na rotoru.

Počet pólů rotorového vinutí musí být shodný s počtem pólů statoru. Vývody z pohybujícího se rotorového vinutí lze realizovat prostřednictvím kluzných kontaktů, které jsou vytvořeny ze tří kroužků a uhlíkových kartáčů, jak je znázorněno na obrázku 3.80. Vinutí rotoru se pospojuje do uzlu přes vnější odpory, jejichž velikost lze měnit. Velikost těchto odporů lze měnit a takto „změkčit" charakteristiku motoru za rozběhu. Kluzný kontakt je relativně náročný na údržbu, proto se mnohem častěji užívají stroje s kotvou nakrátko, která tento kluzný kontakt nemá.



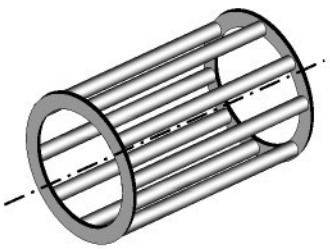
Obrázek 3.81 – Kroužkový asynchronní motor

Tento způsob regulace se používal například u pohonu pouťových kolotočů. Odpory představovala vodou naplněná nádobka, do které obsluha spouštěla tři segmenty napojené na rotorové vinutí pohonu. Rotor s klecí má v drážkách hliníkové (nebo měděné) tyče bez jakékoliv izolace a na koncích jsou tyče spojeny (rovněž hliníkovými) kruhy nakrátko, jak je znázorněno na obrázku 3.81. Toto provedení je nenáročné na údržbu.

Napájecí napětí ve statorových cívkách o frekvenci f_1 o počtu pólů p vytvoří točivé magnetické pole o úhlové rychlosti

$$\omega_0 = 2\pi \frac{f_1}{p}. \qquad (3.21)$$

Toto točivé magnetické pole indukuje ve vinutí rotoru nebo v kleci proud, který vyvolá sílu působící kolmo na vodiče rotoru. Dvojice sil na protilehlých stranách pak vyvolá krouticí moment.



Obrázek 3.82 – Klec nakrátko asynchronního motoru

Proud se v kleci indukuje jen za podmínky, že úhlová rychlost kotvy ω se liší od úhlové rychlosti ω_0 točivého elektrického pole. Tento rozdíl otáček se označuje jako skluzové otáčky $\omega_0 - \omega$. Skluz se určí ze vztahu

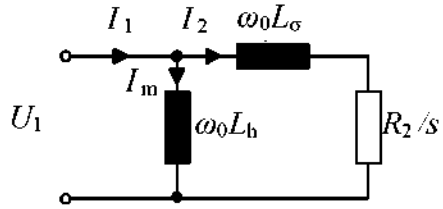
$$s = 1 - \frac{\omega}{\omega_0}. \qquad (3.22)$$

Frekvence proudu v kotvě lze vyjádřit vzorcem

$$f_2 = s\, f_1 .$$

(3.23)

Zjednodušené náhradní schéma asynchronního motoru je znázorněno na obrázku 3.82.



Obrázek 3.83 – Zjednodušené náhradní schéma asynchronního motoru

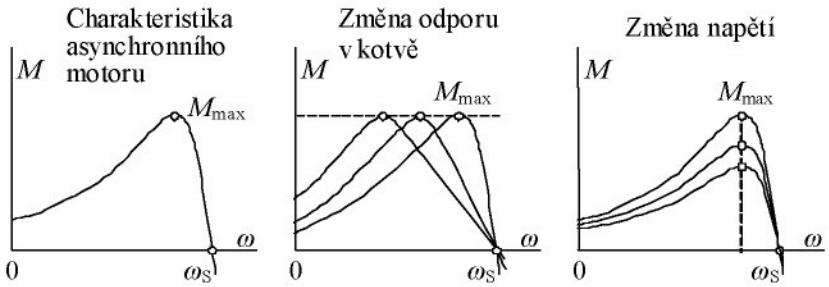
Zanedbány jsou ztráty ve statorovém vinutí a v železe v důsledku jeho magnetizace s hysterezí. Odpor vinutí rotoru je označen R_2. Napájecí napětí U_1 je předpokládáno konstantní. Skluz s určuje proud vinutí kotvy. Kroutící moment se odvodí z rovnosti elektrického a mechanického výkonu ($U_1 I_2 = M\omega$). Platí

$$M = \frac{\left|U_1^2\right|}{\omega_0 (1-s)\sqrt{(R_2/s)^2 + (\omega_0 L_\sigma)^2}} = \frac{\left|U_1^2\right|}{\omega_0 \sqrt{(R_2 (1-s)/s)^2 + ((1-s)\omega_0 L_\sigma)^2}} . \qquad (3.24)$$

Charakteristika asynchronního motoru je znázorněna v diagramu na obrázku 3.80. Při synchronních otáčkách (s = 0) je kroutící moment motoru nulový. Postupný vzrůst zatížení motoru krouticím momentem vyvolá pokles jeho otáček a tím i zvětšení skluzu. Motor se autoreguluje a zůstává ve stabilním stavu nebo jinak řečeno pracuje ve stabilním pracovním bodu, ve kterém se vyrovná hnací a zátěžný kroutící moment. V okamžiku, kdy je dosaženo maxima kroutícího momentu M_{max}, tak přestane autoregulace působit. Motor nemůže vyrovnat kroutící moment zátěže, a proto jeho otáčky poklesnou a hnací moment také. Pracovní bod již není stabilní, což způsobí překotný pokles otáček obvykle k nule, jestliže zatížení zmizí. V opačném případě se může změnit i směr otáčení. Jestliže se motor zastaví, jeho kroutící moment je menší než moment maximální a odběr proudu se zvětší. Tento jev je znám těm, kdo mají zkušenost s řezáním dříví na cirkulárce a občasným "přepálením pojistek" při zablokování otáčení pily, když se přitlačí více „na pilu".

Popsané chování asynchronního motoru se týkalo tzv. motorického režimu. Jestliže na motor působí místo zátěže naopak hnací moment (jízda výtahem dolů vlivem hmoty vyšší než je protizávaží), dojde k překročení synchronních otáček, motor se dostane do generátorového režimu a směr proudu se obrátí. Pokud je motor připojen přímo na rozvodnou síť, je zpětně do rozvodné sítě rekuperován elektrický výkon.

Na obrázku 3.83 je demonstrován rovněž vliv změny odporu ve vinutí kotvy u kroužkového motoru. Zvětšení odporu rozšíří stabilní oblast charakteristiky tím, že se připustí větší skluz. Změna napájecího napětí motoru naproti tomu má vliv jen na velikost maxima kroutícího momentu při neměnné hodnotě skluzu. Zmíněné možnosti ovládat otáčky a kroutící moment asynchronního motoru nejsou vhodné k regulaci jeho otáček. Pro regulaci otáček se využívá změny frekvence napájecího napětí.



Obrázek 3.84 – Charakteristiky asynchronního motoru

Obrázek 3.85 – Nízkonapěťový asynchronní motor

Otáčky asynchronního stroje lze řídit plynule změnou frekvence napájecího napětí užitím frekvenčních měničů, skokově přepínáním počtu pólů a řízením skluzu (odporem v rotoru u kroužkových motorů, napájením rotoru a impulsním napájením).

❑ Jednofázový komutátorový motor

Jde o stejnosměrný motor připojený na jednofázovou síť. Pro trvalý chod musí být obvod statoru vyroben z izolovaných plechů a velkého počtu lamel (segmentů) na komutátoru. Komutace se u tohoto typu motoru zhoršuje, což nutí ke snížení magnetického toku a snížení kmitočtu. Oproti synchronním a asynchronním motorům lze u těchto motorů měnit otáčky v širokých mezích a také překročit hranici 3000 otáček za minutu, což je u dříve zmíněných motorů pro síť 50 Hz nemožné. Nevýhodou je jiskření na kartáčích, což způsobuje jejich opotřebení.

Komutátorové motory velkých výkonů se používají jako trakční motory pro dráhové systémy s kmitočtem 16 ⅔ Hz, což je třetina 50 Hz (dráhy v Německu, Rakousku a Švýcarsku), aby se změna frekvence dala snadno realizovat šestipólovými rotačními konvertory frekvence a střídavé napětí bylo možné transformovat. Z důvodu zamezení rezonančních jevů a možnost použití polovodičových prvků byla frekvence změněna na 16,7 Hz. Komutátorové motory pro malé výkony lze použít také pro síť s 50 Hz. Jedná se o domácí spotřebiče (mixery, fény, mlýnky, vysavače) a elektrické nářadí (ruční vrtačky a brusky apod.).

❑ Elektronicky komutovaný motor (EC motor)

Motor s komutátorem má nevýhodu v nutnosti údržby komutátoru, což u EC motorů odpadá. U elektronicky komutovaného motoru (EC) je vyměněna funkce statoru a rotoru. Také se jim někdy říká stejnosměrné bezkartáčové motory. Rotor tvoří permanentní magnet a do vinutí statoru je přiváděno elektronicky komutované napětí s fází závislou na poloze rotoru.

Obrázek 3.86 – Řada AC motorů firmy Berger Lahr

Příkladem může být třífázový elektronicky komutovaný synchronní motor firmy Berger Lahr, jehož vzorky jsou uvedeny na obrázku 3.85. Při malých rozměrech dosahují tyto motory relativně velkého výkonu a dynamiky, mají také tichý a plynulý chod. Rotor je tvořen permanentním magnetem ze vzácných zemin. Pro snímání polohy rotoru se používají Hallovy sondy, které se připojují spolu s fázemi motoru k servozesilovači. Motory dosahují velkých otáček při malém krouticím momentu. Pro dosažení vyššího krouticího momentu se používají čelní převodovky (do 12 Nm) nebo planetové převodovky (do 50 Nm).

❑ Synchronní střídavé motory

Synchronní motor pracuje pouze na synchronních otáčkách bez skluzu. Aby rotor sledoval otáčky točivého magnetického pole, musí obsahovat permanentní magnet nebo cívku napájenou stejnosměrným proudem. Synchronní motory se rozbíhají jako asynchronní s pomocným vinutím, ve kterém zanikne při synchronních otáčkách proud, a nebo postupným zvětšováním kmitočtu.

Velké synchronní motory mají zvláštní využití v energetice podniku, protože je lze použít ke kompenzaci indukční jalové složky proudu. Tato složka proudu vzniká v provozech, kde je velké množství například asynchronních pohonů. Tento motor může například pohánět centrální kompresor pro rozvod stlačeného vzduchu.

V automatizaci mají význam malé motory buzené permanentními magnety na rotoru. Tyto motory jsou nejvíce používány pro polohové řízení výrobních (obráběcích) strojů, protože jsou vybaveny precizním snímačem polohy a otáček s vysokým rozlišením. Svou konstrukcí jsou vhodné pro dynamicky náročné úlohy. Doplněním vhodnou planetovou převodovkou je možno optimalizovat potřebný moment k otáčkám pohonu.

Tyto střídavé motory jsou bezkartáčové synchronní motory s permanentními magnety na rotoru a třífázovým vinutím ve statoru. Optimalizovaná koncepce motoru s použitím nových magnetických materiálů (neodym-železo-bór) dovoluje až 5-násobné momentové přetížení a motory vychází konstrukčně kratší oproti jiným typům. Optický snímač polohy je s vysokým rozlišením až 16 384 impulsů na otáčku.

❑ Krokové motory

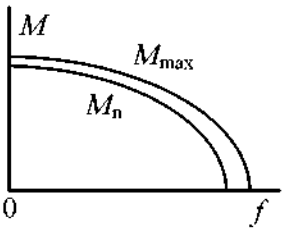
Podle konstrukčního provedení rotoru se krokové motory rozdělují do tří skupin

- krokové motory s pasivním rotorem mající vyjádřené póly (zuby) na statoru a rotoru, jsou označovány jako reakční nebo reluktanční využívající proměnlivou magnetickou vodivost (Variable Reluktance), podmínkou funkce je rozdílný počet pólů na statoru a rotoru

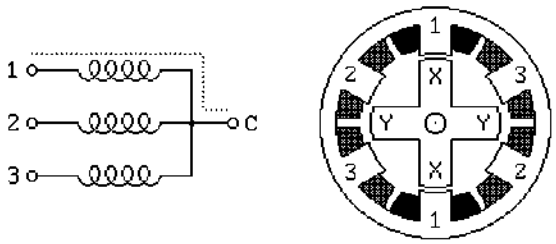
- krokové motory s aktivním rotorem, který obsahuje magneticky aktivní část, tj. budící vinutí nebo permanentní magnet, počet pólů rotoru a statoru je rozdílný; motory s permanentním magnetem udržují klidovou polohu i když statorové vinutí je bez proudu
- skupina slučující konstrukční řešení obou předchozích typů.

Statická charakteristika krokových motorů na obrázku 3.86, kterou poskytuje výrobce, obsahuje dvě křivky. První představuje závislost jmenovitého výkonu na frekvenci a druhá určuje maximální výkon, po jehož překročení krokový motor vypadne ze synchronismu.

Konstrukční princip reluktančního motoru je znázorněn na obrázku 3.87. Stator obsahuje tři cívky na pólových nástavcích 1-1, 2-2 a 3-3. Rotor má dvě dvojice zubů X-X a Y-Y. Na obrázku je nakreslen stav, kdy proud protéká cívkou 1-1 a rotor je svými výstupy (zuby) X-X natočen stabilně tak, aby uzavíral magnetický obvod. Přepnutí proudu do cívky 2-2 se pootočí rotor o 30^0 do nového stabilního stavu, ve kterém je uzavřen magnetický obvod mezi pólovými nástavci cívky 2-2 a výstupy rotoru Y-Y, atd.

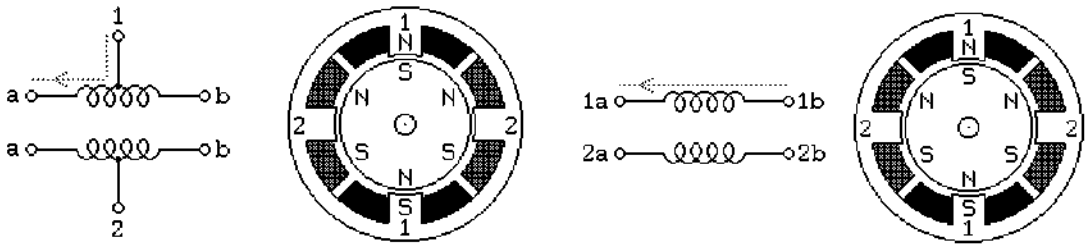


Obrázek 3.87 – Charakteristiky krokového motoru



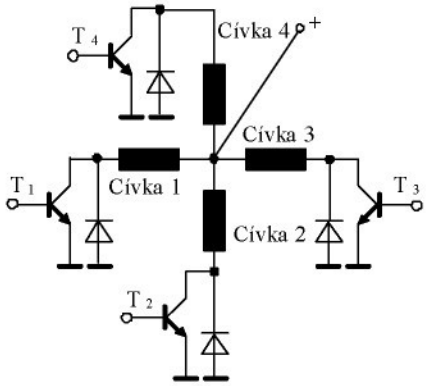
Obrázek 3.88 – Reluktanční krokový motor

Rotor se třemi permanentními magnety a dvěmi cívkami statorového vinutí je znázorněn ve dvou variantách zapojení cívek na obrázku 3.88. Postupným přepínáním proudů do jednotlivých sekcí vinutí statorových cívek se dosáhne stejného efektu jako v předcházejícím případě.



Obrázek 3.89 – Dvě varianty zapojení cívek krokového motoru s permanentními magnety

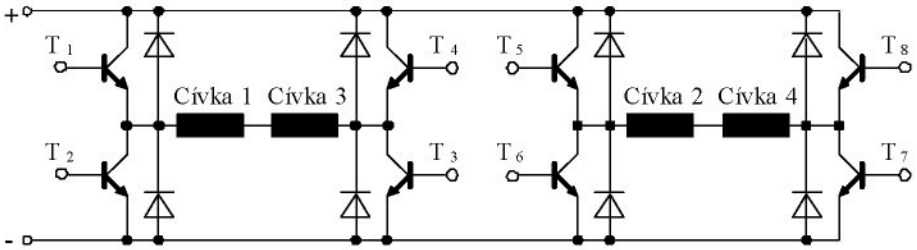
Řízení krokových motorů rozeznáváme unipolární a bipolární. Při unipolárním řízení krokového motoru se čtyřmi statorovými cívkami prochází v každém okamžiku proud jen jednou cívkou (viz obrázek 3.89), což je zajištěno sepnutím jen jednoho z tranzistorů T_1 až T_4. Sepnutí tranzistoru je zabezpečeno proudem do jeho báze po požadovanou dobu. Motor má menší odběr proudu a také nejmenší kroutící moment oproti bipolárnímu řízení. Při bipolárním řízení prochází proud dvěma protilehlými cívkami (viz obrázek 3.90). Tyto cívky jsou zapojeny tak, že mají navzájem opačně orientované magnetické pole. Motor v tomto režimu vyvozuje větší kroutící moment za cenu vyšší spotřeby proudu oproti zapojení unipolárnímu.



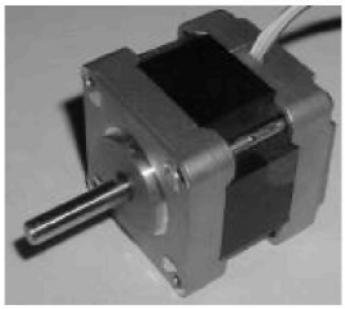
Obrázek 3.90 – Schéma unipolárního řízení

Krokové motory se řídí jednofázově nebo dvoufázově. Při jednofázovém řízení generuje magnetické pole pouze jedna cívka (případně dvojice cívek při bipolárním buzení). Při dvoufázovém řízení generují shodně orientované magnetické pole dvě sousední cívky. Počet poloh rotoru se zdvojnásobí, protože lze rotor natočit také do mezipoloh daných jednofázovým řízením.

Řízení lze provádět s plným nebo polovičním krokem. Řízení s plným krokem odpovídá počtu pootočení za otáčku rovném počtu zubů statoru daného motoru. Řízením s polovičním krokem dosáhneme dvojnásobné přesnosti.

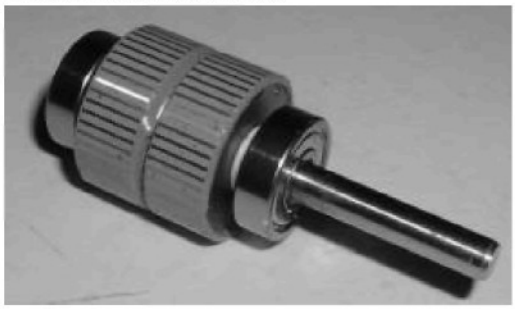


Obrázek 3.91 – Schéma bipolárního řízení čtveřice cívek



Celkový pohled

Stator



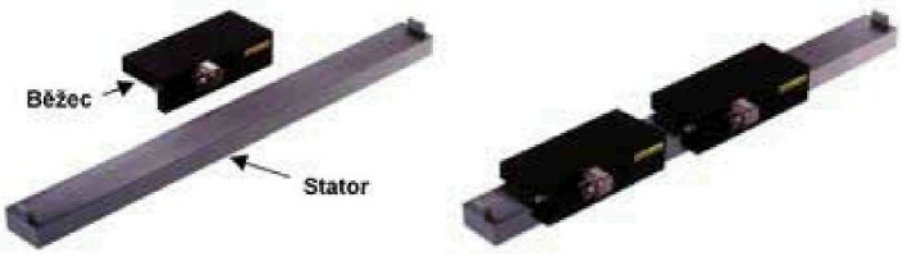
Rotor

Obrázek 3.92 – příklad krokového motoru s 200 kroky na otáčku (převzato z http://robotika.cz/articles/steppers/cs)

Na obrázku 3.91 je příklad krokového motoru s 200 kroky na otáčku (tj. 1.8 stupně na krok). Stator krokového motoru tvoří sada 4 dvojic cívek. Pólové nástavce statoru jsou vroubkovány se stejnou roztečí jako je rozteč magnetů na rotoru. Rotor tvoří hřídel usazená na kuličkových ložiskách a prstencem permanentních magnetů. Proud procházející cívkou statoru vytvoří magnetické pole, které přitáhne opačný pól magnetu rotoru. Vhodným zapojováním cívek dosáhneme vytvoření rotujícího magnetického pole, které otáčí rotorem.

Krokové motory nejsou jen pro otáčivý pohyb, ale také pro lineární posunutí. Příklad takového lineárního motoru je na obrázku 3.92 od firmy Berger Lahr. Lineární krokový motor využívá principu permanentně buzeného reluktančního krokového motoru**.** Stator je tvořen hranolem z nerezové oceli, ve kterém je vytvořena velmi přesná struktura zubů a mezer vyplněných polymerem. Povrch statoru je zabroušen. Stator tak vytváří současně vedení lineární osy a žádné další přídavné vedení není nutné. Běžec se pohybuje nad statorem na

silně stlačeném tenkém vzduchovém polštáři (asi 15 μm), který vzniká rozvedením tlakového vzduchu do trysek po vnitřní ploše běžce a působí proti magnetické síle přitahující běžec ke statoru. Vypnutím přívodu vzduchu je běžec magneticky pevně přitažen ke statoru a motor tedy nepotřebuje žádnou přídavnou brzdu. V běžci jsou kromě rozvodu tlakového vzduchu zabudovány tři fáze vinutí motoru, které jsou vyvedeny na kovový konektor.

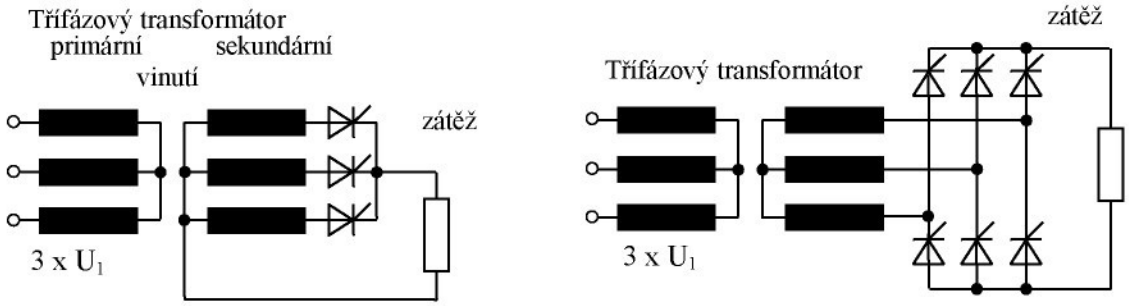


Obrázek 3.93 – Schéma bipolárního řízení čtveřice cívek

❑ Napájení elektrických motorů řízeným usměrňovačem

Řízený usměrňovač je zdrojem regulovatelného napájecího napětí pro stejnosměrné elektrické motory. Jejich funkce je založena na tyristorech. Výkonové řízené usměrňovače používají pro napájení třífázové napětí, které je transformováno třífázovým transformátorem. Řízené usměrňovače jsou buď uzlové (obrázek 3.93 vlevo) nebo můstkové (obrázek 3.93 vpravo). Tyristory mohou vést proud za podmínky, že na jejich anodě je vyšší napětí než katodě. Tento stav pro každý tyristor trvá po dobu jedné třetiny periody napětí. Uzlové zapojení vyžaduje vyvedení a hlavně zatížení nulového vodiče transformátoru, což není často dovoleno. Naproti tomu můstkové zapojení se obejde dokonce bez transformátoru.

Obvody obou usměrňovačů na zmíněném obrázku neobsahují pro jednoduchost zdroje řídicích impulsů pro tyristory. Tyto signály je třeba k oběma usměrňovačům připojit.

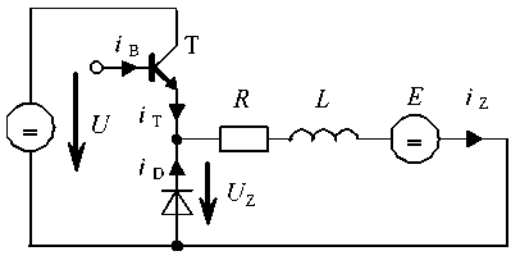


Uzlové zapojení řízeného usměrňovače Můstkové zapojení řízeného usměrňovače

Obrázek 3.94 – Řízené usměrňovače

Pohony představují pro zdroje napětí odporovou a indukční zátěž. V obvodech pro jejich řízení jsou používané rychlé polovodičové spínače. Jevy, které v těchto obvodech vznikají, lze demonstrovat na jednoduchém obvodu, který je znázorněn na obrázku 3.94. V tomto obvodu je spínán tranzistorem proud ze stejnosměrného napěťového zdroje o napětí U pro stejnosměrný motor, jehož kotva vykazuje odpor o velikosti R a indukčnosti L. Na kotvě motoru se indukuje elektromotorická síla o velikosti $E = k\Phi\omega$. Lze předpokládat, že setrvačnost motoru včetně poháněné soustavy je tak velká, že v intervalu spínání tranzistoru T se otáčky motoru téměř nemění a elektromotorická síla E je proto téměř konstantní.

Nechť proud báze i_B se impulsně mění podle časového průběhu na obrázku 3.91. Po dobu průchodu proudu bází tranzistoru se vybudí stav nasycení tranzistoru (ekvivalentní sepnutí fiktivního spínače mezi kolektorem a emitorem) a na jeho emitoru se objeví napětí blízké napětí zdroje, tj. U. Indukčnost vinutí kotvy motoru účinkuje tak, že se proud tranzistorem a vinutím začne exponenciálně zvětšovat s časovou konstantou $\tau = L/R$ z počátečního proudu, který bude označen i_{Z0}.



Obrázek 3.95 - Spínání proudu pro stejnosměrný motor

$$i_Z = i_{Z0} + \left(\frac{U}{R} - i_{Z0}\right)\left(1 - \exp\left(-t/\tau\right)\right). \tag{3.25}$$

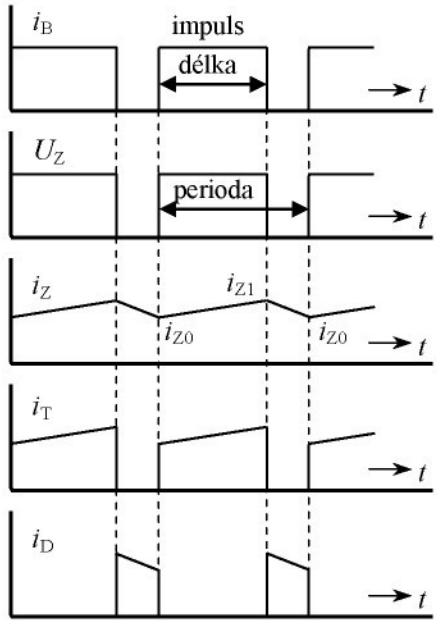
Protože proudové impulsy do báze se předpokládají mnohem kratší než je zmíněna elektrická časová konstanta motoru, jeví se náběh proudu vinutím motoru lineární s časem. Při přerušení proudu kolektorem a emitorem tranzistoru začne proud klesat z dosažené hodnoty i_{Z1} až k hodnotě i_{Z0}, kdy se vše opakuje znovu.

$$i_Z = i_{Z1} \exp\left(-t/\tau\right). \tag{3.26}$$

Úbytek napětí na cívce je dán výrazem $L di_z/dt$. Změna znaménka derivace proudu způsobí, že na anodě diody se objeví kladné napětí proti její katodě a dioda se stane pro proud vodivá. Proud indukčností motoru střídavě prochází tranzistorem a diodou, jak je naznačeno v časových průbězích na obrázku 3.95.

Kdyby v obvodu dioda chyběla, pak se na indukčnosti bude v okamžiku přerušení proudu indukovat vysoké napětí, které má udržet proud v neuzavřeném obvodu. Tato napěťová špička by prorazila přechody v tranzistoru, což by znamenalo jeho zničení. Dioda tedy plní také ochrannou funkci před napěťovým přetížením tranzistoru.

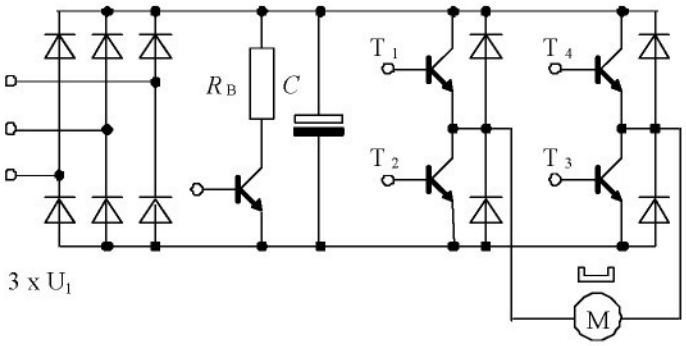
Kompletní zapojení pro toto řízení stejnosměrného motoru je na obrázku 3.96. Obvod umožňuje reverzování otáček motoru. Vzhledem k tomu, že diodový můstkový usměrňovač neumožňuje



Obrázek 3.96 – Časový průběh napětí a proudu při spínání

rekuperaci proudu do sítě, je při brzdění, např. během doběhu, sepnut tranzistor s odporem R_B, který zatěžuje motor ve funkci dynama tím, že zvýší jeho proudový odběr. Elektrolytický kondenzátor vyhlazuje napětí po usměrnění diodovým můstkem.

Tranzistory jsou v obvodu na obrázku 3.96 kresleny bez připojené báze. Předpokládá se, že na tyto vstupy bude přiveden impulsní řídicí signál podobně jako pro obvod na obrázku 3.95.

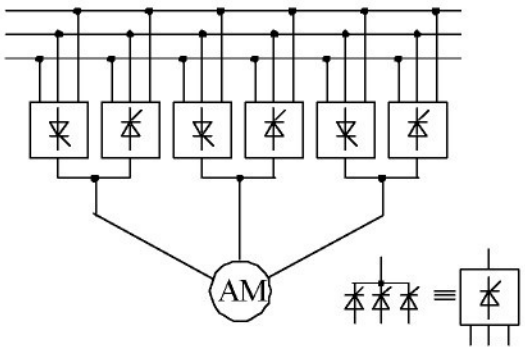


Obrázek 3.97 – Tranzistorový pulsní měnič

Měniče frekvence

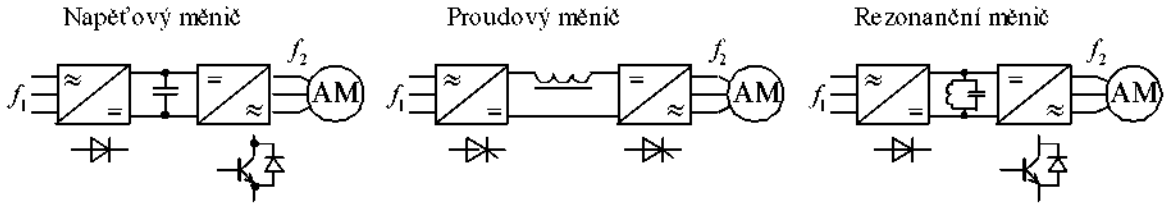
Elektrické pohony střídavé (synchronní a asynchronní) mají otáčky vázané na frekvenci napájecího proudu. Otáčky synchronních pohonů jsou přímo úměrné frekvenci. U asynchronních motorů ovlivňuje tuto závislost z menší části také proměnný skluz. Pro tyto střídavé motory je tedy zapotřebí proud o řiditelné frekvenci. V rozvodné síti je k dispozici proud o stálé frekvenci 50 Hz (60 Hz v USA). Proto je třeba změnit frekvenci proudu uměle speciálním zařízením, které se nazývá frekvenční měnič.

Frekvenční měniče se rozdělují na měniče přímé a nepřímé. U přímých měničů se napětí jedné frekvence (50 Hz) mění přímo na napětí proměnlivé frekvence. K přímým měničům se řadí cyklokonvertory a maticové měniče. Nepřímé měniče nejprve usměrňují elektrický proud na stejnosměrný, který pak střídáním mění na střídavý o libovolné (z určitého rozsahu) frekvenci. Princip cyklokonvertoru je znázorněn na obrázku 3.97. Je realizován třemi reverzačními tyristorovými usměrňovači s fázovým řízením. Každý tento usměrňovač napájí jednu fázi třífázového asynchronního motoru. Výstupní frekvence může dosáhnout jen asi 25% vstupní frekvence a proto má význam jen k řízení pomaluběžných motorů.

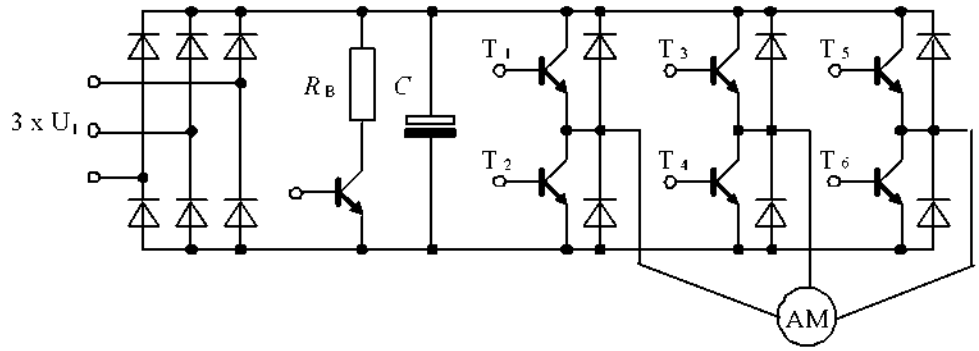


Obrázek 3.98 – Princip cyklokonvertoru

Nepřímé měniče nejprve střídavé napětí o frekvenci f_1 usměrní na stejnosměrný proud, který je spínacími prvky přerušován tak, že vznikne napětí o frekvenci f_2. Blokové schéma zapojení tří typů měničů je znázorněno na obrázku 3.98. Napěťový měnič má ve své stejnosměrné části vyhlazovací elektrický kondenzátor, který se chová jako napěťový zdroj s velmi malou vnitřní impedancí. Proudový měnič využívá k vyhlazení zvlnění proudu indukčnost (tlumivka). Měniče malých výkonů mají usměrňovací obvod s diodami, zatímco pro větší výkony se využívá tyristorů. Usměrňovač a střídač je vybaven tyristory, což umožňuje rekuperaci energie do sítě. To znamená, že energie pro zbrzdění pohonů se nemaří na teplo, ale je dodávána zpět do elektrorozvodné sítě. Frekvenční měnič využívá měkkého spínání tranzistorů, tj. spínání a rozpínání obvodu dochází v okamžiku, kdy je na kolektoru tranzistoru nulové napětí. Rezonanční obvod kmitá na vysoké frekvenci, např. 100 kHz. Harmonického průběhu proudu se dosahuje vynecháváním některých pulsů.



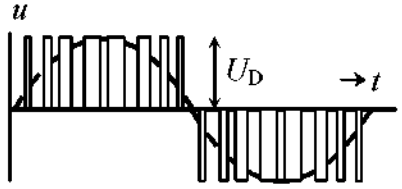
Obrázek 3.99 – Principy nepřímých měničů



Obrázek 3.100 – Napěťový měnič frekvence s diodovým napáječem

Třífázový napěťový měnič frekvence s diodovým napáječem je znázorněn na obrázku 3.99. Obsahuje diodový můstkový usměrňovač, elektrolytický vyhlazovací kondenzátor ve stejnosměrné části obvodu a tranzistorový třífázový střídač. Diodový usměrňovač podobně jako v obrázku 3.92 neumožňuje rekuperaci, proto se tranzistorem spíná do obvodu brzdný odpor. Spínacími tranzistory se napájí současně sinovým napětím všechna tři vinutí asynchronního motoru. Funkce diod v obvodu byla vysvětlena na zjednodušeném obvodu na obrázku 3.94 a 3.95.

K vytváření sinusového průběhu napětí se používá pulsně šířkové modulace impulsním signálem. V průběhu jedné periody se připojuje a odpojuje stejnosměrné napájecí napětí s šířkou impulsu úměrnou okamžité hodnotě sinusového napětí. Princip této modulace je znázorněn na obrázku 3.100. Impulsy o různé šířce jsou filtrační schopností indukčnosti elektromotoru vyhlazeny tak, že odpovídají sinusovce.



Obrázek 3.101 – Princip pulsní šířkové modulace

Otázky 3.7.

1. Které motory se zapojují do hvězdy a trojúhelníka?
2. Co je to asynchronní motor?
3. K čemu se používá střídač nebo měnič frekvence?
4. Co je to krokový motor a kde se použije?
5. Jak se liší řízený usměrňovač od střídače?
6. Jakým zdrojem elektrického proudu se napájí stejnosměrný elektrický motor?
7. Co je to synchronní motor?
8. Které elektrické motory se otáčejí stejně rychle jako točivé magnetické pole, které vybudí statorové cívky?

3.9. Hydraulické a pneumatické pohony

Čas ke studiu: 1 hodina

Cíl Po prostudování tohoto odstavce budete umět

- popsat pomocná zařízení obsahující pneumatické a hydraulické obvody (zdroje tlakového vzduchu nebo kapaliny, tlakové omezovací ventily, průtokové regulační ventily.
- popsat výhody a nevýhody pneumatických a hydraulických systémů.

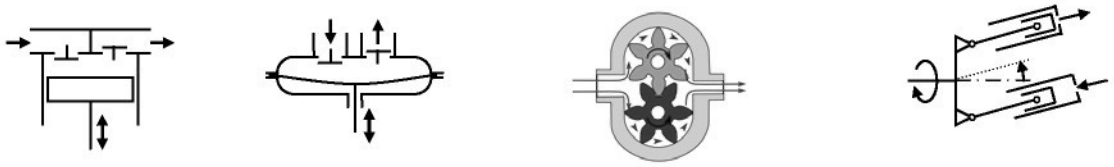
Výklad

K pneumatickému a hydraulickému obvodu patří také zdroje tlakového vzduchu nebo kapaliny. Tlakový vzduch dodávají kompresory. Tato zařízení lze rozdělit na několik typů: objemové kompresory (pístové, membránové), rotační kompresory (vícekomorový rotační, šroubové, zubové), proudové kompresory (axiální, radiální).

Tlakovou kapalinu (olej) dodávají hydrogenerátory: zubové, šroubové, lamelové, lopatkové a pístové (axiální nebo radiální pístové čerpadlo). Pro hydrogenerátor se používá také běžně označení čerpadlo.

Ukázky kompresorů a čerpadel jsou na obrázcích 3.101. Jedná se o kompresor pístový, kompresor membránový, zubové čerpadlo a axiální pístové čerpadlo.

Charakteristika čerpadla je obecně závislost průtoku pracovní kapaliny čerpadlem na jeho výstupním tlaku. Naklápěním pístů axiálního čerpadla oproti ose rotace lze řídit průtok tohoto čerpadla.

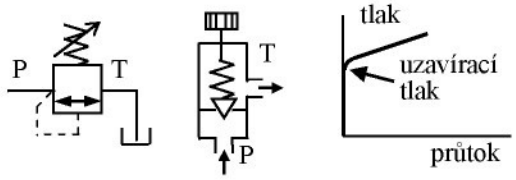


Pístový kompresor | Membránový kompresor | Zubové čerpadlo | Axiální pístové čerpadlo

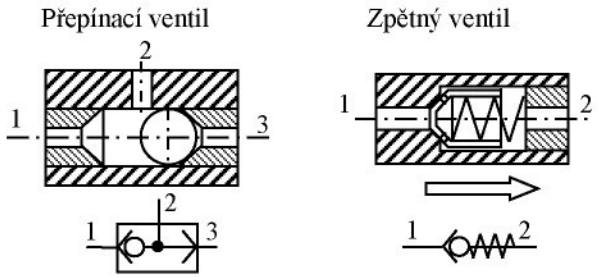
Obrázek 3.102 – Princip funkce kompresoru a hydrogenerátoru (čerpadla)

Pneumatické a hydraulické obvody vyžadují další pomocná zařízení. Některé z nich jsou součástí tabulky 3.2 s obvodovými značkami. Tlakový omezovací ventil je znázorněn na obrázku 3.102. Jeho úlohou je chránit obvod před přetlakem. Tlakový ventil se otevře, jestliže síla působící na kuželový ventil je větší než přítlačná síla pružiny. Předpětí pružiny lze nastavit ručně. Pohyb ventilového kužele může zpomalovat tlumící píst. Na charakteristice tohoto prvku je zřejmý uzavírací tlak, po jehož překročení začne hydraulická kapalina protékat. Jak je na obrázku zřejmé, uzavírací tlak je stavitelný změnou předpětí pružiny naznačeným šroubem.

Na obrázku 3.103 je přepínací a zpětný ventil. Přepínací ventil pracuje tak, že se na vývod 2 dostane vyšší tlak z přívodu 1 a 3, který přetlačí kuličku. Zpětný ventil umožňuje proudění jen jedním směrem.

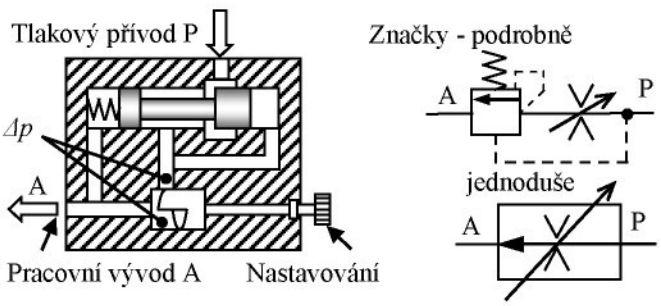


Obrázek 3.103 – Tlakový omezovací ventil



Obrázek 3.104 – Přepínací a zpětný ventil

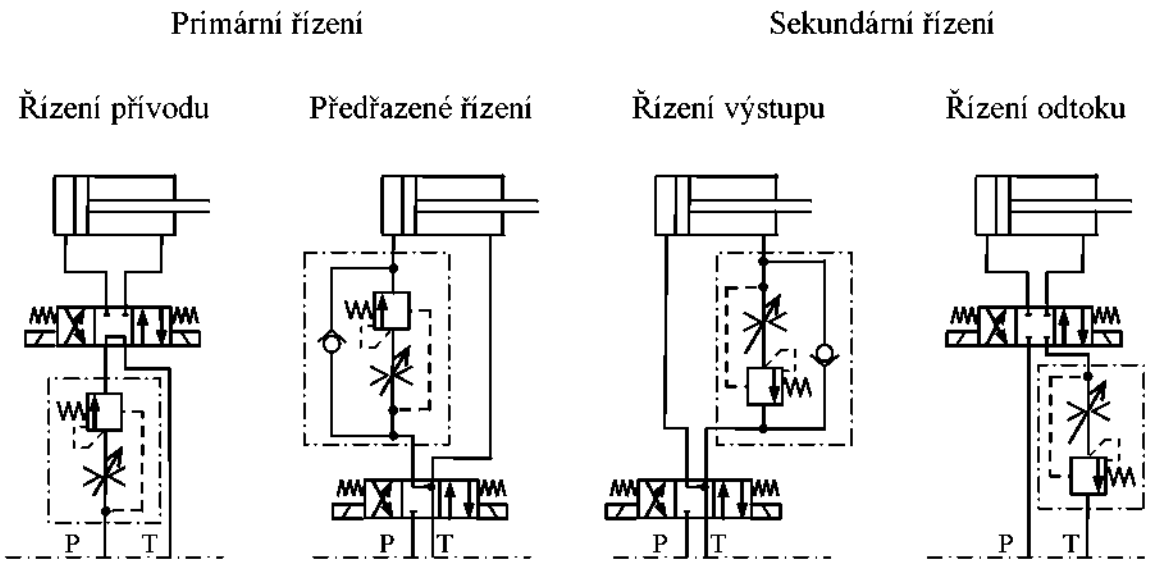
Na obrázku 3.104 je znázorněn průtokový regulační ventil, který se nazývá také tlaková váha. Tento ventil obsahuje nastavitelnou clonu, na které vzniká průtokem určitá tlaková diference Δp. Řídicí píst tlakové váhy má dvě části. Tlak oleje na vnitřní stěny obou pístů, které spojuje pístní tyč, nemá na polohu řídicího pístu vliv. O jeho poloze rozhoduje rovnováha přítlaku pružiny a síla daná velikostí tlakové diference a plochy pístu.



Obrázek 3.105 – Průtokový regulační ventil

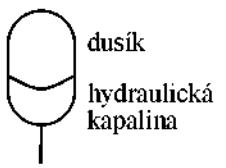
Jestliže se clonou zvětší průtok, například zvýšením tlaku na přívodu P, pak diference tlaků přetlačí vnitřní pružinu směrem doleva a přivře průtok oleje na tlakovém přívodu. Zpětná vazba působí tak, aby se obnovila nastavena velikost tlakové diference na cloně. Účelem průtokového regulačního ventilu je tedy udržovat průtok na nastavené hodnotě. Průtok oleje určuje rychlost pohybu pístu v hydraulickém válci. Konstantní průtok oleje zabezpečí také konstantní rychlost pohybu a omezí účinek kolísání tlaku a zátěže.

Průtokový regulační ventil lze využít k řízení rychlosti pohybu dvojčinného hydraulického válce. Podle umístění tohoto ventilu po směru toku oleje lze rozeznat primární a sekundární způsob řízení. Pro primární řízení jsou varianty řízení přívodu a předřazené řízení. Pro sekundární řízení je to řízení výstupu a řízení odtoku. Všechny varianty jsou znázorněny na obrázku 3.105. Čerchovaná čára odděluje funkční celky zapojení. Přívod tlakového oleje je označen písmenem P a výpusť T. Primární řízení šetří těsnění válců, namáháno je jen vedení mezi čerpadlem a ventilem. Chybějící protitlak může vést při střídavém zatížení ke zpětnému pohybu pístu. Při sekundárním řízení je brzděn odtok kapaliny z hydraulického válce. Regulační ventil tak vytváří protitlak pracovnímu tlaku a píst je držen mezi dvěma sloupci tlakového oleje. Nevýhodou řízení pomocí škrtících ventilů jsou energetické ztráty.



Obrázek 3.106 – Řízení rychlosti

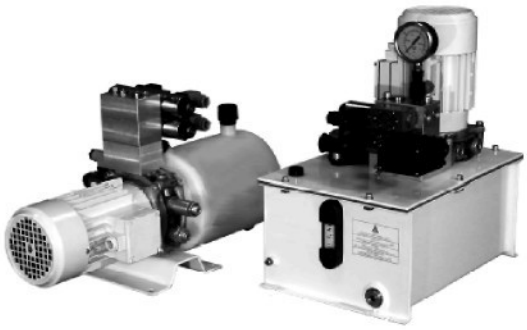
Akumulátor tlakové energie je znázorněn na obrázku 3.106. Jedná se o zásobník rozdělený pružnou membránou, kterou kapalina tlačí proti uzavřenému prostoru naplněnému dusíkem. Vysoká stlačitelnost plynné náplně dusíkem umožní pohlcovat tlakové rázy v potrubí. Hydraulická kapalina je oproti plynu téměř nestlačitelná a má vysokou tuhost.



Obrázek 3.107 – Akumulátor tlaku

Hydraulické obvody jsou otevřené a uzavřené. V otevřeném obvodu se olej shromažďuje v nádobě, odkud se zase čerpá zpět. Uzavřené obvody mají uzavřený oběh, doplňují se jen shromažďované ztráty z vložené nádrže.

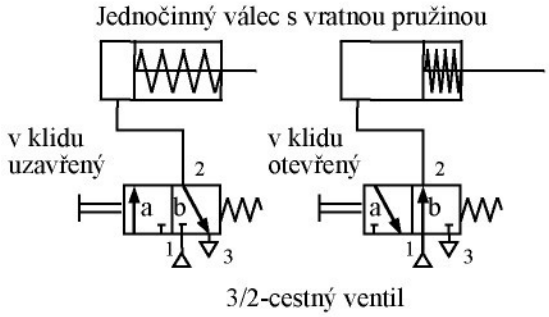
Čerpadlo, filtry, pojistné ventily a případně i akumulátor tlaku tvoří hydraulický agregát, jehož vzorové provedení je na obrázku 3.107. Je to výrobek firmy ARGO-HYTOS a.s., představující malý kompaktní agregát SMA 04 s nádrží na olej od 1,5 do 40 dm^3 pro maximální tlak 25 MPa a maximální průtok 17 dm^3/min. Výkon poháněcího elektrického motoru je 0,12 až 3 kW.



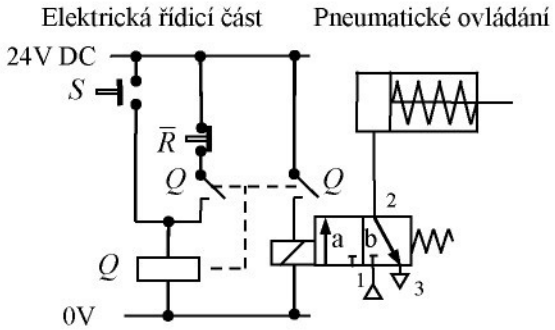
Obrázek 3.108 – Malé kompaktní agregáty SMA 04 od firmy ARGO-HYTOS

Po základních informacích o hydraulických a pneumatických obvodech následují ukázky jejich zapojení. Na obrázku 3.108 je třícestný dvoustavový ventil s vratnou pružinou ve spojení s jednočinným pneumatickým válcem se dvěma variantami klidové polohy ventilu, které se liší připojením vývodů k jednomu ze dvou bloků. Variantě obvodu s ventilem, který je v klidu uzavřen, odpovídá poloha pístu válce s roztaženou vratnou pružinou a naproti tomu u varianty s jiným ventilem, který je v klidu otevřen, je píst v poloze se stlačenou vratnou pružinou.

Ventily na obrázku 3.108 mají obecné ruční ovládání, zatímco na obrázku 3.109 je ovládání elektromagnetické. Přesun do obou krajních poloh je po stisku tlačítek, z nichž jedno je spínací a druhé rozpínací. Elektrická řídicí část obsahuje kromě tlačítek také stykač. Zapojení stykače odpovídá klopnému obvodu typu RS. Zapojení na obrázku představuje elektropneumatický systém řízení. Na rozdíl od obrázku 3.109 je ve schématech elektropneumatického řízení silová pneumatická část kreslena odděleně od elektrické řídicí části. Řídicí část se kreslí podle stejných pravidel, které platí pro schémata logického řízení.



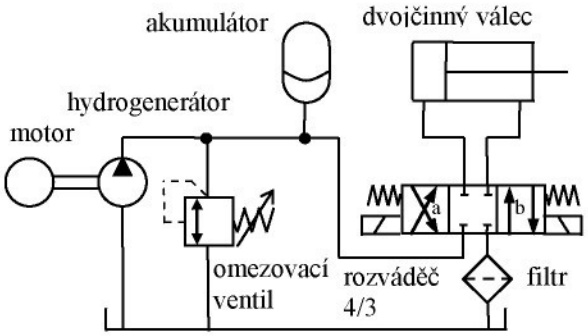
Obrázek 3.109 – Varianty funkce třícestného dvoustavového ventilu



Obrázek 3.110 – Elektropneumatické řízení

Druhá ukázka na obrázku 3.110 se týká hydraulického obvodu se čtyřcestným třístavovým ventilem, který je ovládán dvěma cívkami působícími v opačných směrech, a centrován v klidové poloze pružinami. V klidové poloze je přívod kapaliny do hydraulického válce zablokován. Ve dvou zbývajících polohách je dodávána tlaková kapalina na opačné strany pístu. Dalšími důležitými prvky, kterými je tento hydraulický obvod vybaven jsou filtr, čerpadlo (hydrogenerátor), akumulátor tlaku, pojistný tlakový ventil a samozřejmě dvojčinný hydraulický válec.

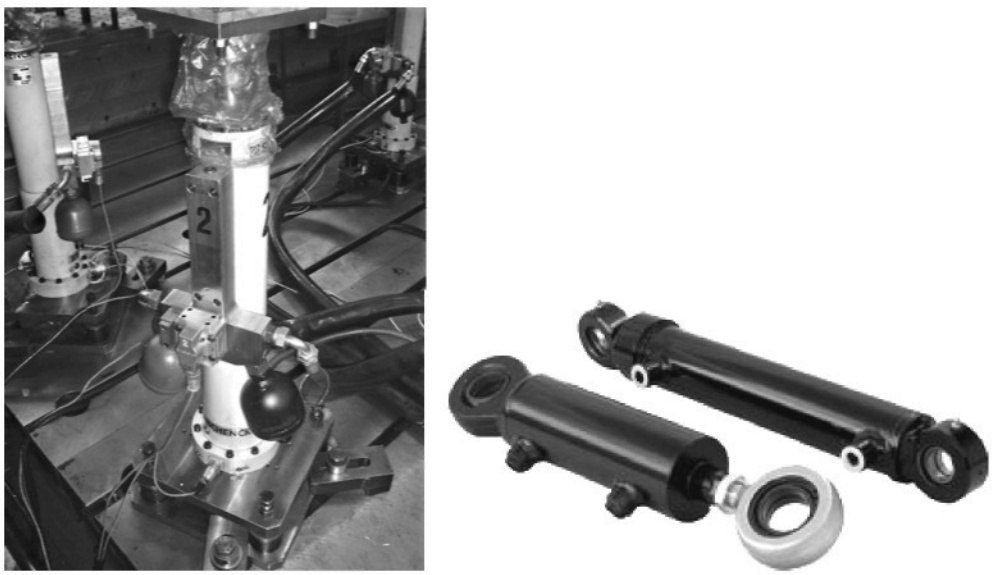
Všímavý student zjistí, že při centrované poloze ventilu prochází z čerpadla celé dodávané množství přes pojistný ventil zpět bez užitku, tj. ztrácí se výkon daný součinem tlakové ztráty (Pa) a objemového průtoku (m^3/s). Takto je mařen výkon motoru čerpadla. Moderním hydraulické pohony jsou vybaveny čerpadlem s regulací na konstantní tlak, konstantní výkon nebo s tzv. load-sensing regulací.



Obrázek 3.111 – Příklad hydraulického obvodu

Přehled hydraulických a pneumatických pohonů byl zmíněn při definici jejich obvodových značek. Hydromotory mají v některých případech stejnou konstrukci jako generátory. Od generátorů se liší tím, že kapalina se do motoru přivádí pod tlakem, a tak je schopna zaplňovat pracovní prostor. Hydromotory mohou konat rotační, přímočarý a kývavý pohybe. Rotační pohyb může vykonávat například zubové čerpadlo, jestliže se napojí na zdroj tlakového oleje. Hydromotory pro přímočarý pohyb (válce) jsou jednočinné nebo dvojčinné. U jednočinných působí tlak kapaliny jen na jednu stranu pístu, tzn. jen jeden zdvih je pracovní. Zpětný (nepracovní) pohyb pístu musí být zajištěn vnější silou (pružinou hmotností pístu a pracovního nástroje nebo jiným hydromotorem). Dvojčinné hydraulické válce mají přívod tlakové kapaliny střídavě na obě strany pístu. Tyt hydromotory mou být s jednostrannou nebo dvoustrannou pístnici, což znamená, že plocha pístu ze strany pístnice je menší než na straně bez pístnice. K hydromotorům s přímočarým pohybem patří také plunžr a teleskopický válec. Hydromotory s kývavým pohybem jsou výrobně náročné a používají se jen ve speciálních případech. Jedno z možných řešení používá píst s hřebenovým ozubením, se kterým zabírá pastorek.

Ukázka hydraulického válce, který je určen pro dopravní stroj je na obrázku 3.111.



Obrázek 3.113 – Ukázka provedení hydraulických válců

Porovnání hydraulických a pneumatických obvodů ovládání a řízení

Pneumatické systém mají následující přednosti a nevýhody:

- vyžadují centrální rozvod stlačeného vzduchu nebo aspoň kompresor, který je velmi hlučný
- nutnost rozprášení oleje do tohoto vzduchu pro bezchybnou funkci zařízení
- nelze dosáhnout velkých sil na pístech pneumatických válců, protože se používá tlak jen do 1 MPA
- dynamika řízení je značně závislá na zatížení pneumatických pohonů
- nelze dosáhnout pomalých pohybů
- možnost práce v prostředí s nebezpečím výbuchu nebo pod vodou.

Hydraulické systémy lze charakterizovat těmito výhodami a nevýhodami:

- prostorově malé jednotky přenášející velké síly a výkony
- rychle, jemně a plynule přestavitelné rychlosti pohybů válců a motorů
- jednoduché zabránění přetížení omezením tlaku
- viskozita hydraulického oleje závisí na teplotě
- vznikají ztráty prosakováním
- náchylnost ke kmitání a hluku.

Otázky 3.8.

1. Co je charakteristika čerpadla?
2. Jak se liší uzavřený a otevřený hydraulické obvody?
3. Mohou pneumatické pohony pracovat v prostředí s nebezpečím výbuchu?
4. Které pohony přenášejí vyšší výkony, pneumatické nebo hydraulické?

3.10. Piezoelektrické aktuátory

Čas ke studiu: 0,5 hodin

Cíl Po prostudování tohoto odstavce budete umět

- definovat piezoelektrický aktuátor oblasti jeho aplikace.
- popsat provedení piezoelektrických aktuátorů (vrstvené, laminární,).

Výklad

Piezoelektrické aktuátory se staly velmi důležitým prvkem pro mikromanipulaci (micropositioning), kde mají nezastupitelnou úlohu. Uplatňují se také v systémech aktivního potlačování hluku a vibrací. Tyto aktuátory využívají inverzní piezoelektrický efekt. Přímý piezoelektrický efekt objevili bratří Pierre a Jacques Curie v předminulém století, kteří zjistili, že některé krystalické materiály (keramika) mají schopnost generovat elektrický náboj úměrně vnější síle, která na ně působí. Tento jev je využíván ve snímačích síly, především dynamické. Inverzní jev spočívá v tom, že elektrické pole, které je orientováno paralelně ke směru polarizace, vybudí mechanické roztažení nebo smrštění materiálu. Piezoelektrický efekt je anizotropní, tj. působí jen v určitém směru.

Nejvíce populárním piezoelektrickým materiálem je keramika PZT (Lead-Zirconate-Titanate) nebo polymer PVDF (Polyvinilidene fluoride).

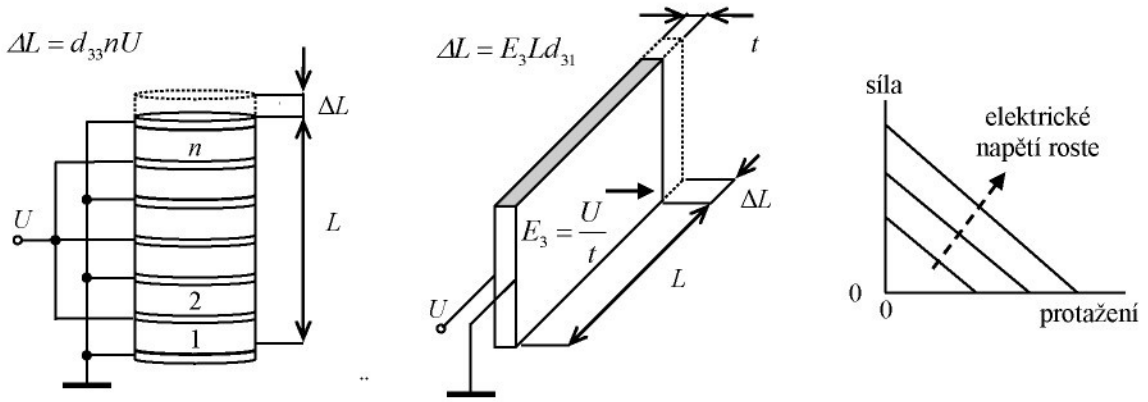
Lineární piezoelektrický aktuátor je vrstven z tenkých keramických destiček (tloušťka mezi 0,1 až 1 mm), které jsou odděleny elektrodami, jak je znázorněno na obrázku 3.112 vlevo. Potřebné napětí pro jednu destičku tvaru disku tloušťky 0,1 mm je asi 100 V, což je označováno jako nízkonapěťová LVPZ (Low voltage piezo). Destička o tloušťce 1 mm vyžaduje 1000 V a je proto označována jako vysokonapěťová HVZP. Maximální protažení je mezi 0,1 až 0,13 %, což znamená, že aktuátor o délce 100 mm má maximální protažení 100 μm. Tyto prvky vykazují hysterezí až 15 %.

Laminární uspořádání je znázorněno na obrázku 3.112 uprostřed.

Napětí, které je přivedeno na elektrody, je označeno U. K výpočtu protažení podle vzorců uvedených v obrázcích je třeba znát kromě vyznačených délek také piezoelektrické konstanty d_{33} a d_{31}. K dimenzování je navít třeba znát také maximální mechanické napětí a intenzitu elektrického pole. Tyto údaje jsou uvedeny v tabulce 3.5.

Tabulka 3.5 – Vlastnosti piezoelektrických materiálů

Materiál	PZT	PVDF
Piezoelektrické konstanty d_{33} (10^{-12} C/N nebo m/V) d_{31} (10^{-12} C/N nebo m/V)	 300 -150	 -25 jednoosý d_{31}=15 d_{32}=3 dvouosý d_{31}=d_{32}==3
Maximální napětí (MPa) tlak tah	 600 80	 200 200
Max. intenzita elektrického pole (V/mm)	2000	5.10^5
Hustota (kg/m^3)	7600	1800



Vrstvené provedení aktuátoru

Laminární provedení aktuátoru

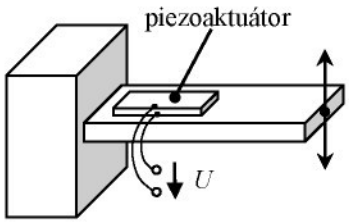
Charakteristika aktuátoru

Obrázek 3.113 – Provedení piezoelektrických aktuátorů a jejich charakteristika

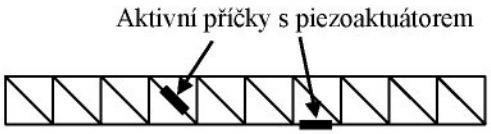
Charakteristika aktuátoru je znázorněna na obrázku 3.112 vpravo. Jedná se o závislost síly a protažení s parametrem elektrické napětí.

Firma CEDRAT dodává vrstvené piezoaktuátory v provedení označeném PI 840.30 pro protažení 45 μm a sílu 1500 N, dále APA 50 S pro protažení 55 μm a sílu 45 N (APA 100

N, 110 m a 117 N). Tloušťka vrstvy u laminárních aktuátorů je typicky 250 μm a je potažena stříbrnými elektrodami. Materiál PVDF je ohebnější ve srovnání s PZT, avšak tyto aktuátory mají 300krát menší citlivost.



Obrázek 3.114 – Vetknutý nosník s piezoaktuátorem



Obrázek 3.115 – Příhradová konstrukce s aktivními příčkami

Jak již bylo uvedeno, piezoaktuátory mají uplatnění v systémech aktivního tlumení vibrací a zvuku (obrázek 3.113 a 3.114). Jestliže budou umístěny blízko vetknutí nosníku, pak jeho volný konec lze rozkmitat nebo v případě kmitání utlumit. Jiným aplikačním příkladem je příhradová konstrukce. Jestliže je třeba její kmitání utlumit, pak se místo některých prvků umístí piezoelektrické aktuátory, které kmitání vhodným způsobem řízení utlumí. Tento princip byl studován pro tlumení anténních systémů na oběžné dráze kolem země.

Další aplikace lze nalézt v reproduktorech, pohybových ústrojích zrcadel laserů, frekvenční modulace laserového paprsku užitím Dopplerova jevu, vstřikování paliva do válců systémem common rail, inkoustové tiskárny, mikroskopy. Dokonce byl zkonstruován rotační krokový motor, které je piezoaktuátory poháněn.

? Otázky 3.9.

1. Jaký je nejčastěji používaný piezoelektrický materiál?
2. Jaké veličiny tvoří charakteristiku aktuátoru?
3. Uveďte tři příklady aplikace piezoaktuátorů.

3.11. Počítačové sítě

Čas ke studiu: 1 hodina

Cíl Po prostudování tohoto odstavce budete umět

- definovat komunikační protokol, IP adresy.
- popsat rozdělení sítí z hlediska aplikací, územní rozlehlosti nebo topologie.
- vysvětlit symbolické adresy, vybrané služby www.

Výklad

Předávání informací je důležitá součást naší civilizace. První drátový přenos informací začala zajišťovat telefonní síť již více než před 100 lety. Každý ví, že telefonovat lze kamkoliv do světa. Počítačová síť je vymoženost konce 20. století a také se s ní lze setkat kdekoliv v civilizovaných místech na zeměkouli. Sítě vznikly s cílem sdílet

- počítačové programy a dat
- technické vybavení (například společná tiskárna)

Z hlediska aplikací lze sítě rozdělit na

- jednoúčelové sítě
  - rezervační (letenky)
  - bankovní
  - pro důležité organizace (armáda)
  - pro řízení výroby
  - atd.
- veřejné sítě

Z hlediska územní rozlehlosti lze jsou sítě s tímto označením

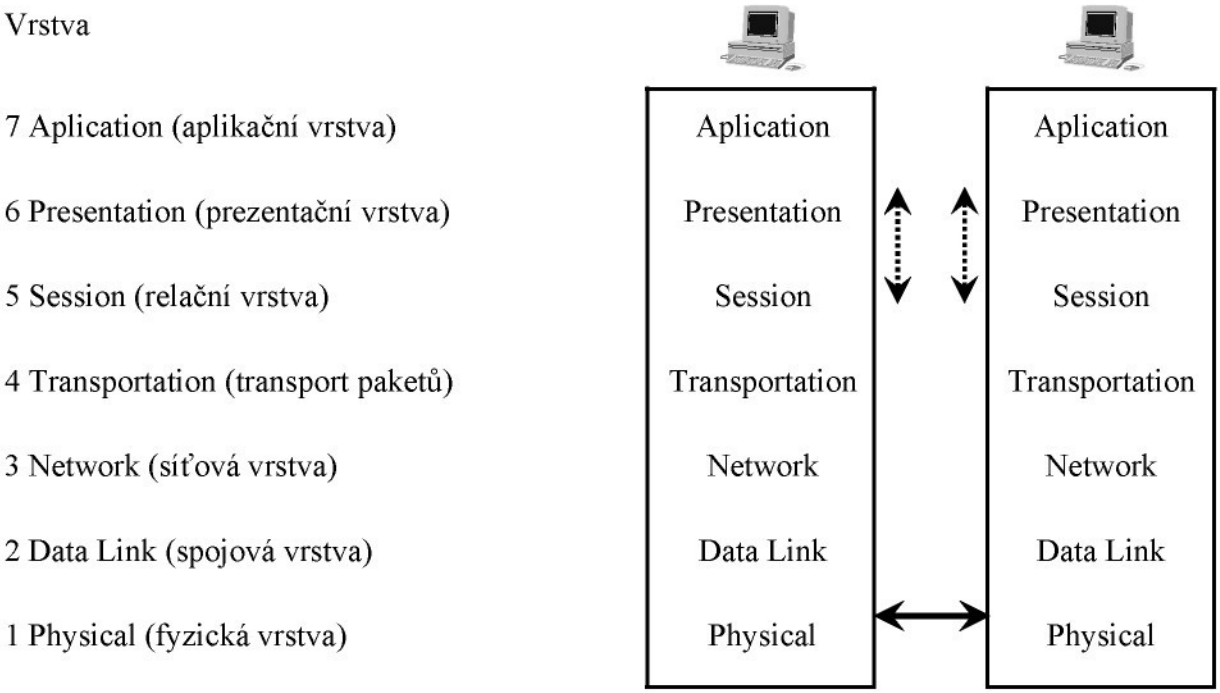
- LAN (Local Area Network) pro jednu instituci, například univerzita
- MAN (Metropolitan Area Network) pro městskou aglomeraci
- WAN (Wide Area Network) přesahující města, například síť Internet

Síť může být z hlediska topologie, tj. vzájemného pospojování počítačů

- sběrnicová (na jednu sběrnici jsou připojeny všechny počítače bez priority)
- kruhová (sběrnice je uzavřena do kruhu)
- hvězdicová (jeden centrální počítač – centrální uzel, ke kterému jsou ostatní připojení)
- typ stromu (síť se rozvětvuje v rozbočovačích)
- typ polygonu (každý s každým)

Velmi často jsou sítě hybridní s částmi tvořenými různou topologií. Sítě jsou standardizovány v normě IEEE 802 zobecněním konkrétních výrobků jako například sítě arcnet, token-ring a ethernet. Arcnet je historicky nejstarší standard (1982). Token-ring je síťovou implementací firmy IBM

Pravidla komunikace po sítích určují komunikační protokoly. Komunikace má řadu vrstev (problémových oblastí) od řešení fyzického propojení (vlastnosti signálů elektrických a optických) až po aplikační programy. Standard ISO OSI (Open System Interconection) rozeznává 7 vrstev.



Obrázek 3.116 – Komunikační protokol ISO OSI

Označení vrstva je zvoleno, protože mezi nimi dochází k předávání informací. Mezi dvěma počítači se samozřejmě bezprostřední přenos informací odehrává na nejspodnější fyzické vrstvě.

Technické prostředky obsahují

- opakovač – repeater, tvaruje a zesiluje přijatý zkreslený signál (pracuje na fyzické vrstvě)
- převodník, nejen zesiluje, ale i převádí z jednoho kabelu na jiný, například elektrický signál na optický (pracuje na fyzické vrstvě)
- přepínač – switch, zajišťuje větvení signálu například v sítích se stromovou topologií (pracuje na fyzické vrstvě)
- most – bridge, spojuje dva segmenty kabelové sítě a zajišťuje přenos paketů (nese data jedné aplikace – jednoho spojení). Toto zařízení vykonává některé inteligentní funkce tím, že reaguje na obsah paketů a rozeznává část jejich adresy a rozhoduje, který paket má být poslán vně svého segmentu, kde vznikl (pracuje na úrovni spojové vrstvy). Most spojuje také sítě, které se liší na úrovni fyzické vrstvy.

- směrovač – router, plní obdobné funkce jako most, ovšem na vyšší úrovni lití LAN (pracuje na úrovni síťové vrstvy). Směřovač je schopen určit každému paketu jeho další cestu, tato činnost se nazývá směrování paketů.
- brána – gate, pracuje na nejvyšší úrovni a slouží k propojení LAN na cizí prostředí.

Síťové operační systémy představují nástavbu operačního systému počítače, která má zajistit komunikaci s ostatními účastníky sítě. Síťové programové vybavení zajišťuje sdílení prostředků (společné využití například tiskáren), synchronizaci počítačů,ochranu prostředků sítě, obsluhu tisku a organizování front, zpracování chyb a administrativní úkony (statistika spojení).

Každý síťový operační systém umožňuje vytvořit server a pracovní stanici. Servery jsou počítače, ke kterým jsou připojeny sdílená zařízení (například tiskárny) a které obsahují společná data. Servery poskytují služby ostatním účastníkům sítě.

Servery jsou různých typů: diskový, souborový, tiskový, komunikační a zálohovací. Pro programovou obsluhu sítě typu Lan se používají dva typy operačních systémů, síť server-klient a pear-to-pear (já pán, ty pán).

❑ Síť Internet

Síť internet se řídí protokolem označeným zkratkou TCP/IP (Transmision Control Protocol/Internet Protocol). Má schopnost maximálně pružně směrovat zprávy v rámci rozlehlé sítě. Síťovým protokolem se označují i konkrétní softwarové prostředky, které slouží ke komunikaci v počítačové síti. V našem případě jsou to prostředky nezávislé na platformě (počítač, operační systém) a na síťovém hardware (síťové karty, kabeláž).

Internet protokol odpovídá síťové vrstvě a přenáší mez vzdálenými počítači IP-datagramy, které ve svém záhlaví obsahují adresu příjemce. Síť může přenášet každý datagram samostatně. Lze uvést, že datagramy mohou ke příjemci dorazit i v jiném pořadí, než byly vyslány. Každé síťové rozhranní (síťová karta) má v rámci celosvětové sítě jednoznačnou IP adresu, kterou nesmí používat žádné jiné síťové rozhranní. Internet je tvořen jednotlivými sítěmi, které jsou propojeny pomocí směrovačů (router).

Protokol TCP odpovídá transportní vrstvě. Jeho úkolem je transportovat datové segmenty jednotlivým aplikacím. Protokol TCP zajišťuje spojení mezi aplikacemi (programy), které běží na vzdálených počítačích. Zabezpečovat však také může komunikaci mezi procesy jednoho počítače. Adresou pro doručování datových segmentů je port.

Aplikační protokoly odpovídají několika vrstvám OSI: relační, prezentační a aplikační, které jsou spojeny do jediné aplikační vrstvy TCP/IP. Aplikačních protokolů je velké množství, a proto z hlediska uživatele jsou uvedeny jen tyto

- HTTP (Hypertext Transfer Protokol)
- TELNET (vzdálený terminál počítače)
- FTP (File Transfer Protocol) pro přenos souborů
- POM3, SMTP a IMAP pro elektronickou poštu (E-mail)

❑ IP adresy

IP-adresa je 32bitové číslo, které identifikuje síťové rozhranní. Existují dva nejpoužívanější způsoby, jak toto číslo uvádět:

- Dekadický zápis, například 10.240.5.100
- Binární (bitový) zápis, například 00001010.11110000.00000101.01100100

V obou případech se jedná o posloupnost čtyř čísel oddělených tečkou. Každé číslo může nabývat dekadických hodnot 0 až 255 (omezeno 8 bity).

IP adresa má dvě části, tzv. síťovou a hostitelskou (Host Number). Obě části se rozeznají podle síťové masky. Síťovou masku můžeme zadávat ve třech způsobech zápisu:

- Dekadický zápis, například 255.255.255.0
- Binární (bitový) zápis, například 11111111.11111111.11111111.00000000
- Zkrácený zápis, například /24

Nejpřehlednější je pravděpodobně druhý zápis, nejúspornější naopak třetí. Bity s logickými jedničkami odpovídají síťové adrese. Princip získání jednotlivých částí IP adresy je velmi jednoduchý. Tam, kde má síťová maska jedničky (v binárním zápisu), tam je v IP adrese síťová část, a tam, kde jsou nuly, je host část. Síťová část adresy se dostane logickým vynásobením adresy hodnotou masky a host část logickým vynásobením adresy negovanou hodnotou masky .

IP sítě se dělí do tří tříd podle velikosti. Úplná IP adresa obsahuje síťovou část a hostitelskou část (adresu zařízení v rámci sítě). Síťová a hostitelská část mají v různých třídách sítí různou délku, jak ukazuje tabulka 3.6.

Sítím připojeným k Internetu jsou přiřazeny typy tříd, které určují maximální počet možných hostitelů v síti. Předchozí tabulka š.6 ukazuje, jak se v jednotlivých třídách liší síťová a hostitelská část IP adresy. Třída A je přiřazena síti, která má více než 65 535 hostitelů; třída B je pro sítě, které mají 256 až 65 534 hostitelů; třída C je pro sítě s méně než 256 hostiteli. Všechny síťové adresy mimo tyto rozsahy pro třídy D a E jsou rezervovány nebo používány pro experimentální sítě nebo skupinové vysílání.

Tabulka 3.6 – Rozdělení IP-adresy pro jednotlivé třídy

Bity	0					8								16								24							
Třída A	0	síťová adresa								host																		
Třída B	1	0	síťová adresa															host										
Třída C	1	1	0	síťová adresa																					host			
Třída D	1	1	1	0	adresa pro skupinové (vícesměrné) vysílání (multicast)																							
Třída E	1	1	1	1	0	rezervováno pro budoucí použití																						

Pokud hostitelská část IP adresy obsahuje samé nuly, identifikuje síť a ne hostitele. Takovou adresu nelze přiřadit žádnému fyzickému zařízení. Síťová část adresy musí začínat hodnotou od 1 do 126 nebo od 128 do 223. Všechny ostatní hodnoty v síťové části mohou být od 0 do 255 s výjimkou, že ve třídě B jsou rezervovány síťové adresy 128.0.0.0 a 191.255.0.0 a ve třídě C jsou rezervovány síťové adresy 192.0.0.0 a 223.255.255.0.

Hodnoty v hostitelské části IP adresy fyzického zařízení mohou být v rozmezí 0 až 255, tato část však nesmí obsahovat samé 0 nebo samé 255.

Síťová část musí být stejná pro všechna IP zařízení v samostatné fyzické síti (např. jedné ethernetové LAN nebo připojení WAN). Hostitelská část musí být různá pro každé IP zařízení (nebo přesněji, pro každý IP port nebo rozhraní), které je přímo připojeno do této sítě. Síťová část IP adresy se v tomto návodu označuje jako číslo sítě; hostitelská část jako číslo hostitele (stanice).

Pro připojení k Internetu nebo libovolné privátní IP síti, která používá Internetem přiřazené číslo sítě, se musí získat registrované číslo IP sítě od internetového autorizačního síťového informačního centra.

Symbolické adresy

IP adresa je výhodná pro počítač, avšak pro uživatele je nepohodlná. Naštěstí existuje síťová služba DNS (Domain Name System), která zajišťuje používání jmenných názvů uzlů místo jejich IP adres, například symbolickou adresu www.vsb.cz na místo IP adresy 158.196.149.74. Princip DNS je velmi jednoduchý, základem je databáze se dvěma položkami a to jmenným názvem počítače a IP adresou.

Databáze DNS má hierarchické uspořádání, což znamená, že jednotlivé jmenné názvy jsou rozděleny do vzájemně nadřazených skupin, které se nazývají domény a subdomény.

Jména v jedné doméně k sobě logicky patří, ať už v rámci jedné organizace či celého státu. V rámci jednotlivých domén lze rozlišovat podskupiny jmen, tzv. subdomény, ty mohou mít další vnořené subdomény atd. Každé doméně a subdoméně je přiřazeno nějaké jméno. Název síťového uzlu je pak tvořen z posloupnosti řetězců, které jsou odděleny tečkami

jméno_uzlu.subdoménaN. subdoména2.subdoména1.doména

Doména je nejobecnější část z celého jména. V našem případě to bude označení státu České republiky (cz).

Databáze DNS je distribuovaná, což znamená, že je rozmístěna na více počítačích po celém světě. Těmto počítačům se říká *name* servery. Pro každou doménu existují dva *name* servery: primární a sekundární. Primární jmenný server udržuje platné informace a sekundární tyto informace v pravidelných časových úsecích zálohuje. Tyto *name* servery jsou organizovány distribuovaně. Pro každou úroveň DNS existuje jedna skupina serverů. Jinými slovy servery, které se starající o nejvyšší úroveň DNS, znají servery, které mají na starost nižší úroveň.

Některé služby internetu

Nejvíce využívanou službou internetu je služba WWW (World Wide Web), která zprostředkuje přístup k internetovým (webovým) stránkám, které se vyvinuly z prostého textu, přes jednoduché obrázky až k dnešním multimediálním prezentacím.

Přenos webových stránek zajišťuje protokol HTTP (Hyper Text Transfer Protocol). Platí pro něj platformní nezávislost (PC nebo Apple Macintosh), a tak lze surfovat po Internetu pod jakýmkoliv současným operačním systémem. K prohlížení internetových stránek je zapotřebí tzv. internetový prohlížeč. Na protější straně komunikace je spuštěn WWW server, který přijímá požadavky od klientů (prohlížečů) a jako odpověď posílá zdrojový kód stránek. Ten je v prohlížeči interpretován a zobrazen. Internetové stránky jsou psány v jazyce HTML (Hypertext Markup Language). Zdrojový kód stránky obsahuje

formátovací příkazy. Kromě tohoto jazyka se užívají skriptovací nástroje typu JavaScript či VBScript, popřípadě plnohodnotné programovací jazyky (Java, C#, VB.NET).

Přenos souborů zajišťuje služba FTP a protokol FTP (File Transfer Protocol), kterého tato služba využívá. Jde o nástroj pro vzdálenou práci se soubory a složkami. Dovoluje kopírovat soubory mezi vzdálenými počítači, popřípadě zakládat a rušit složky.

Další oblíbenou službou internetu je poštovní služba. Umožňuje posílat pomocí Internetu elektronickou poštu - emaily. Poštovní služby zajišťují poštovní servery, které shromažďují došlou korespondenci a tuto rozesílají. V Internetu tedy komunikují poštovní servery mezi sebou a teprve v okamžiku připojení k serveru elektronickou poštu předají adresátovi. Poštu zajišťují tři nejrozšířenější poštovní protokoly POP3, moderní IMAP a SMTP. Některé protokoly jsou určené pro odesílání pošty, jiné pro přijímání. Protokol POP3 (Post Office Protocol) poštu přijímá. Také protokol IMAP (Internet Mail Access Protocol) elektronickou poštu dovede přijmout, avšak na rozdíl oproti POP3 poskytuje možnost práce s dopisy přímo na poštovním serveru. Tímto odpadá kopírování na lokální počítač jako u protokolu POP3. Protokol SMTP (Simple Mail Transfer Protocol) je určen k odesílání pošty z lokálního počítače na poštovní server.

Otázky 3.10.

1. Co je to IP adresa na internetu?
2. Co je to LAN?
3. Vyjmenujte některé služby internetu.
4. K čemu se používá přepínač – switch?

3.12. Vizualizace a řízení z operátorského pracoviště

Čas ke studiu: 2 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat pojem vizualizace, operátorské řízení, alarm, trend, runtime, vývojové prostředí SCADA/HMI, tenký a tlustý klient
- popsat hlavní funkce vizualizace v průmyslové automatizaci
- vysvětlit jednotlivé úrovně řízení a jejich uspořádání do systémové architektury

Výklad

Vizualizace

Vizualizace znamená zviditelnění. Umožňuje zobrazit vhodným způsobem to, co není vidět nebo co je našimi smysly nepostřehnutelné, a tím srozumitelnou formou přibližuje nějaký děj. Takovýmto dějem může být právě získávání, zpracování a grafické zobrazení dat z řízeného procesu, či zviditelnění rozpracované výroby v reálném čase. Vizualizační systém má za úkol:

- grafické znázornění řízeného děje,
- sledování – jinými slovy monitorování - a archivace zvláštních a chybových stavů - alarmů,
- sledování průběhů a archivace vývoje vybraných parametrů v čase – časové charakteristiky – neboli tzv. trend,
- možnost zpětného vyvolání průběhu děje z archívu k analýzám a k dalšímu zpracování potřebných parametrů,
- vytváření a archivace reportů o průběhu řízeného děje.

Pro vizualizace je využívána celá řada programových a komunikačních prostředků. Byly vytvořeny proto, aby byly zviditelněny vazby mezi výrobním zařízením a hardware a software prostředky pro jejich řízení.

Obrázek 3.118 - Jednoduchá vizualizace pro signalizaci stavu zapnuto/vypnuto a signalizaci poruchy nebo krajního stavu (blikání žárovky, diody) a moderní software prostředí pro operátorské řízení (rozhranní člověk-stroj)

❑ Systémy řízení z operátorského pracoviště

Vizualizace nám tedy umožňuje na monitoru počítače nebo řídicím panelu zobrazit probíhající technologický proces, aktuální měřené hodnoty v podobě čísel nebo grafu, být informováni o vzniklých poruchách popř. možnostech jejich řešení. Je nám rovněž dovoleno nastavovat žádané hodnoty nebo manuálně zasahovat do automatického procesu, apod. V takovém případě mluvíme o vizualizaci procesu v reálném čase, někdy nazývaném jako runtime systému.

Systémy pro operátorské řízení jsou často označovány jako SCADA/HMI (z anglického *Supervisory Control And Data Aquisition/ Human Machine Interface*), tedy systémy pro supervizí řízení a získávání dat/ s rozhranním člověk – stroj). Jsou využívány také jako vývojová prostředí již při projektování systémů řízení ještě před jejich nasazením do provozu. Představují tak programovou podporu při projektování systémů řízení. Jsou rovněž důležité pro aktualizace a inovace systémů řízení již zprovozněných.

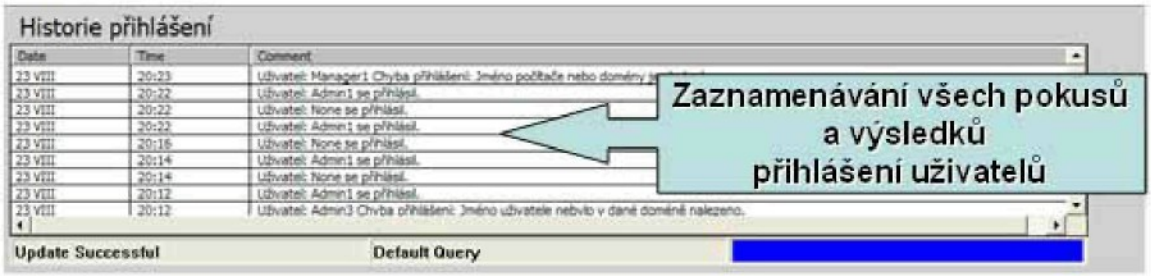
Klíčové funkce SCADA/HMI systémů jsou:

- Sběr a ukládání dat 7/24
- Zpracování dat do vhodných a námi definovaných formátů
- Prezentace výsledků:
  - včas
  - na správné místo
  - kompetentním osobám

Úloha sběru a prezentace dat na monitor operátorského pracoviště – vizualizace – tak musí být rozšířena o další systémovou funkci, a sice zabezpečení aplikace – aby k ní měli přístup pouze kompetentní osoby.

O bezpečný přístup k datům se v aplikaci starají konfigurace:

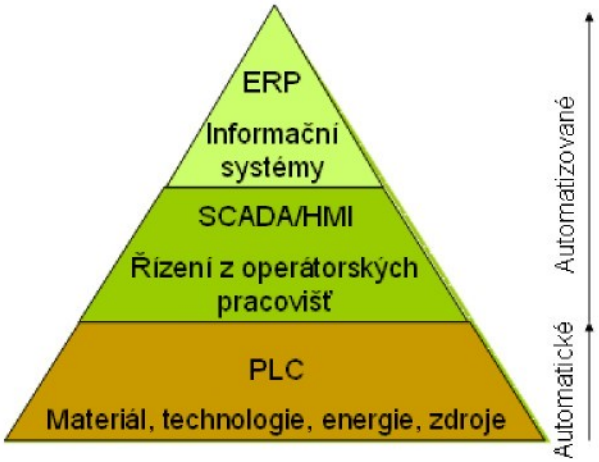
- Hesla a jména uživatele systému
- Přístupová práva k aplikaci
- Viditelnosti objektů aplikace až po přihlášení do aplikace



Obrázek 3.119 - Vizualizace zabezpečení přístupu k datům aplikace

Architektura vizualizačního systému

Programové a komunikační prostředky pracují na platformách specifických pro danou oblast průmyslu a síťovém prostředí. Každý průmyslový podnik, který chce nasadit systém vizualizace a řízení z operátorského pracoviště, potřebuje nejdříve vybudovat síťovou architekturu pro komunikaci mezi jednotlivými zařízeními a vlastním řídicím systémem postaveným na algoritmech naprogramovaných pro PLC. Úroveň řízení postavená na PLC je často nazývaná jako úroveň řízení 1 (automatické řízení). Řízení z operátorských pracovišť pak tvoří úroveň nadřazenou, tedy úroveň 2 (automatizované řízení). Dvou a více úrovňové architektury řízení jsou tak hierarchické. Rozhranním mezi jednotlivými úrovněmi řízení se označuje jako interface. Rozhraní mezi úrovní 1 a 2 je často nazýváno jako rozhranní člověk – stroj (HMI z anglického human machine interface), protože na rozdíl od úrovně 1, kde řízení probíhá automaticky podle programů napsaných z algoritmů, se od úrovně 2 již počítá s aktivní úlohou člověka – operátora - a jeho zasahováním do probíhajícího řízení z operátorského pracoviště.



Obrázek 3.120 - Hierarchie architektury systémů řízení

Příkladem software architektury je tzv. Service Oriented Architecture, zkráceně SOA. Využívá služby - software services – které jsou nezávislé na vyvíjecích platformách, programovacím jazyce a technologiích, např. Java, .NET apod.

Služby psané v C# běží na .Net platformě, služby psané v Java běží na Java EE platformě. Oba typy umí poskytovat služby stejným SOA systémům a aplikacím. Aplikace na kterékoli platformě umí používat službu té druhé jako webové služby (viz. kapitola 3.10). Definice rozhraní (interface) skrývá způsob provedení služby.

❑ Vizualizace dat v Internet/Intranet sítích

Přenos dat z řízených technologických procesů v síti Internet je příkladem více-úrovňové architektury řízení. Data získávaná z technologie a předávaná z PLC do operátorských pracovišť jsou navíc prostřednictvím služby www (World Wide Web) poskytovaná do sítě Internet. Tímto způsobem jsou data z technologie dále předávána mezi Internetovými klienty a servery. Internetový klient je počítač žádající o zasílání dat-dokumentů. Internetový server je počítač, na kterém běží software schopný odpovídat na požadavky. Komunikace mezi klientem a serverem probíhá pomocí protokolu HTTP(HyperText Transfer Protokol).

Internetový server může data předávat ve standardním formátu, např. ve formátu HTML (HyperText Markup Language) nebo XML (eXtended Markup Language). Dokument v tomto formátu lze prohlížet v běžném prohlížeči Internetu (Internet Explorer). Tato vizualizace se nazývá tenký klient. Tenkým klientem pak může být i počítač, který je v hierarchii řízení součástí nejvyšší úrovně řízení označované jako úroveň 3. Tímto způsobem jsou data sdílena v informačních systémech zodpovědných za komplexní plánování výrobních zdrojů podniku (ERP z anglického Enterprise and Ressource Planning). Mezi výrobní zdroje patří data o lidských zdrojích, tedy zaměstnancích, materiále a energiích, výrobním zařízení atd.

Internetový server může také data předávat v jiném než standardizovaném formátu. V takovém případě je nutné mít na straně klienta nainstalovaný vizualizační software, který umí získaná data zpracovat a zobrazit (InTouch a Internet Explorer). V tomto případě se jedná o tlustého klienta.

Alarmy

Alarm je definován jako specifický či výjimečný stav nebo hodnota monitorované veličiny systému procesu řízení. Úkolem tohoto generovaného stavu je signalizovat změny definované v procesu jako důležité (start, ukončení pracovní činnosti) nebo závažné (např. překročení limitní hodnoty).

Alarmy lze dělit podle typu sledované veličiny:

- Analogový alarm- je definován na analogové hodnoty, veličiny definované v určitém rozsahu.
- Diskrétní alarm– vždy jeden ze dvou stavů je před-definovaný jako alarm.
- Odchylkový alarm- je aktivován tehdy, pokud dojde k procentuální odchylce od určené meze. Procentuální odchylka je vypočítána v reálném čase.
- Alarm na rychlost změn - se vyvolá vždy, pokud je rychlost změny sledované veličiny větší než nadefinovaná mez.
- Speciální alarm– definuje ostatní změny statických vlastností.

Alarmy mohou být organizovány ve skupinách. Tyto skupiny alarmů jsou obvykle seřazeny **hierarchicky** – jako nadřazené a podřazené. Podřazené skupiny jsou navzájem propojené a posílají se jako alarmové hlášení jednomu nadřazenému systému, který je sjednocuje a vyhodnocuje – tzv. **systém s distribuovanými alarmy**. Pokud je v dané kompetenci tohoto systému vyrozumět operátora na dané úrovni řízení, tak to udělá. Tímto způsobem je možné rozdělit kompetence pro obsluhu systému a zabránit tomu, aby na jednu obrazovku operátora nechodila hlášení, která s řízením procesu z této stanice nesouvisí.

Priorita alarmu označuje závažnost alarmu a určuje, jak kritický nebo důležitý je sledovaný stav. V případě překročení limitu hodnoty ve sledovaném prostředí může být alarm zařazen do kategorie "velmi závažný" a vyžadovat okamžitou reakci a jednání za účelem záchrany života nebo majetku. Naopak míra závažnosti a tedy i priorita alarmu nastaveném pro zakončení pravidelné pracovní činnosti bude minimální.

Sledovaná veličina může mít hned několik definovaných alarmů. V takovém případě se jedná o **více-stavový alarm**. Například může mít alarmový stav několik mezních hodnot-limitů:

- Vysoký- (označen Hi, High)
- Velmi vysoký (HiHi, High High)
- Nízký – (Lo, Low)
- Velmi nízký – (LoLo, Low Low).

Hlavním účelem alarmů je upozornit obsluhu systému na výjimečný stav, úkolem obsluhy je pak tyto stavy vzít na vědomí („acknowledge"), tím potvrdí, že jsou se situací obeznámeni a umí patřičně zasáhnou. Toto potvrzení nemusí automaticky znamenat nápravu, ani vracení sledované veličiny do normálního stavu.

Trendy

Trend je definován jako průběh sledované veličiny v čase. Existují dva základní způsoby, jak zobrazit trend:

1. **Reálný trend** představuje průběh sledované veličiny, který je zobrazen na obrazovku operátora v reálném čase. Důležité je si však uvědomit, že je zobrazován pouze časový

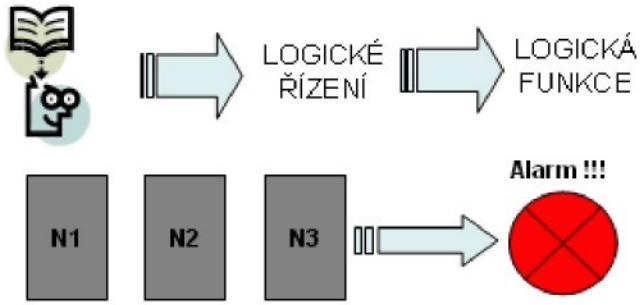
úsek definovaný rozsahem kreslícího nástroje. Data sledované veličiny nejsou nikam ukládána, není tedy potřeba tento nástroj zvlášť konfigurovat pro záznam historie dat.

2. **Historický trend** představuje nástroj pro záznam a ukládání dat s možností jejich dodatečného prohlížení. Data jsou ukládána do souboru a pak teprve posílána do kreslícího nástroje, je tedy nutné pohled na trend aktualizovat například tlačítkem pro aktualizaci („update"). Tento typ nástroje umožňuje operátorovi vyhledávat v minulosti uložené průběhy, k čemuž má k dispozici konfigurovatelná tlačítka a lišty pro pohyb a přiblížení v časové oblasti (zoom).

Příklad řešení vizualizace laboratorní úlohy

Zadání: předpokládejme známou úlohu dílny, kde jsou umístěny tři nádrže s technologickými kapalinami. V každém zásobníku je umístěno čidlo indikující minimální množství kapaliny nutné pro provoz dílny. V případě, že alespoň ve dvou zásobnících je již méně kapaliny, než je minimální množství, je třeba kontrolkou informovat operátora, aby doplnil kapalinu.

Obrázek 3.121 ukazuje návrh řešení úlohy vizualizace



Obrázek 3.121 - úloha pro vizualizaci

Řešení úlohy provedeme podle dříve popsaných metod řešení, viz. schéma na obr. xxx

Blokové schéma

Pravdivostní tabulka

N1	N2	N3	ALMO
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1		0	1
1	1	1	1

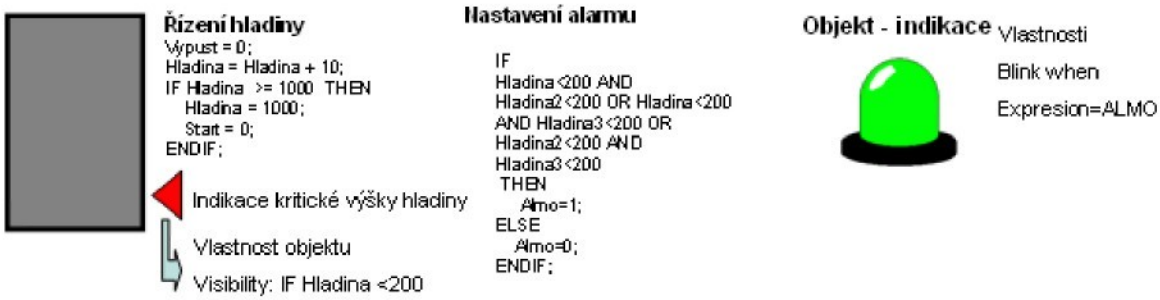
Algebraická rovnice

$$ALMO = N1.N2 + N1.N3 + N2.N3$$

Dalším krokem bude:

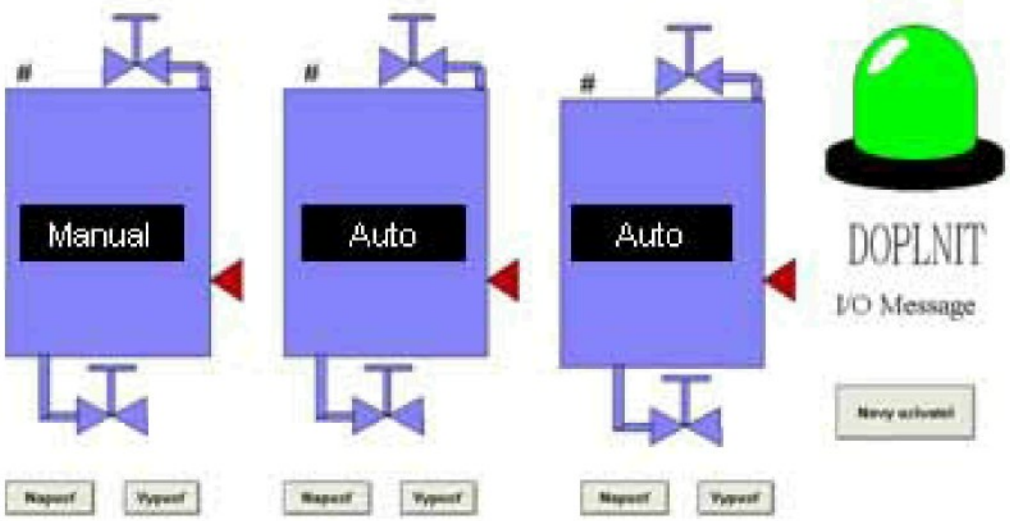
- Navrhnout vzhledu aplikace ve vizualizačním prostředí
- Definice proměnných pro aplikaci vizualizace
- Spuštění aplikace v provozním režimu (runtime), viz. schéma na obrázek 3.122.

Proměnné	Typy	Popis
Hladina	analog	Změna výšky hladiny
VENTIL	discrete	Indikace stavu ventilu
ALARM	discrete	Výstup logické funkce ALMO



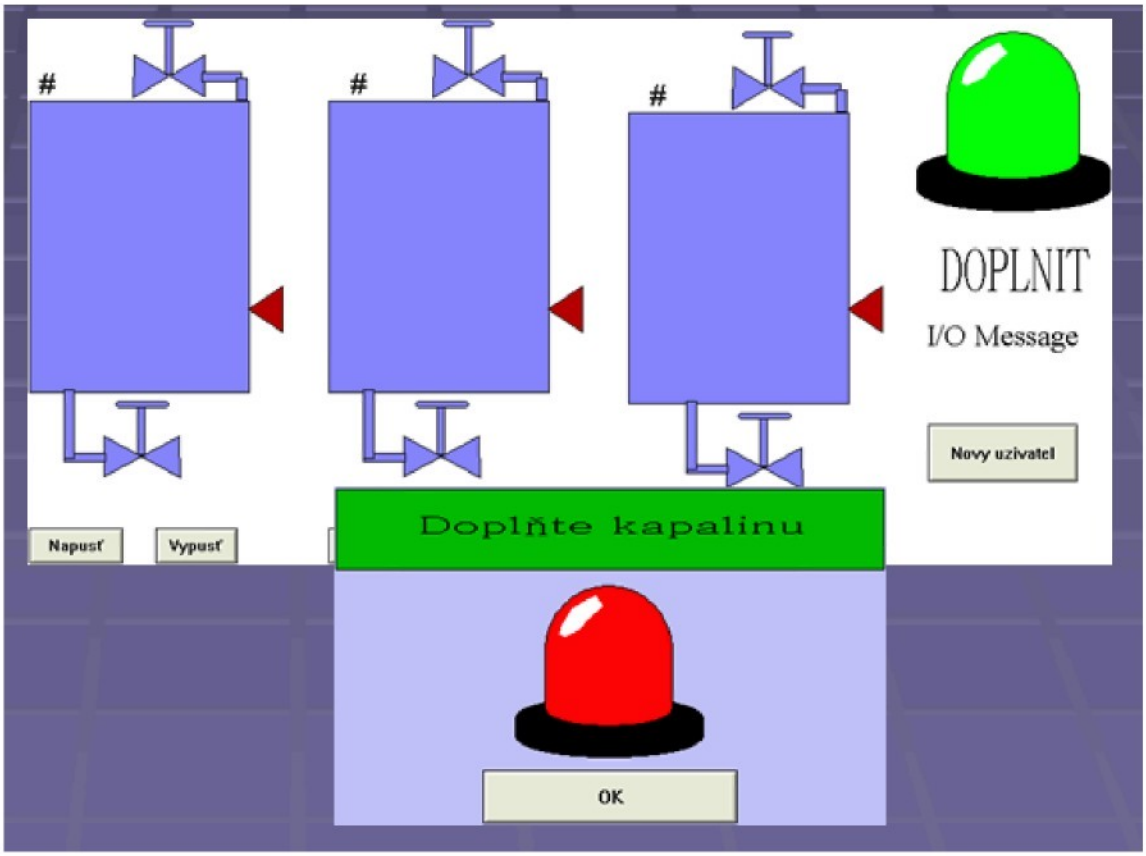
Obrázek 3.122 - Návrh aplikace

Vzhled navržené vizualizace je na obrázku 3.123.



Obrázek 3.123 - Vzhled vizualizace aplikace

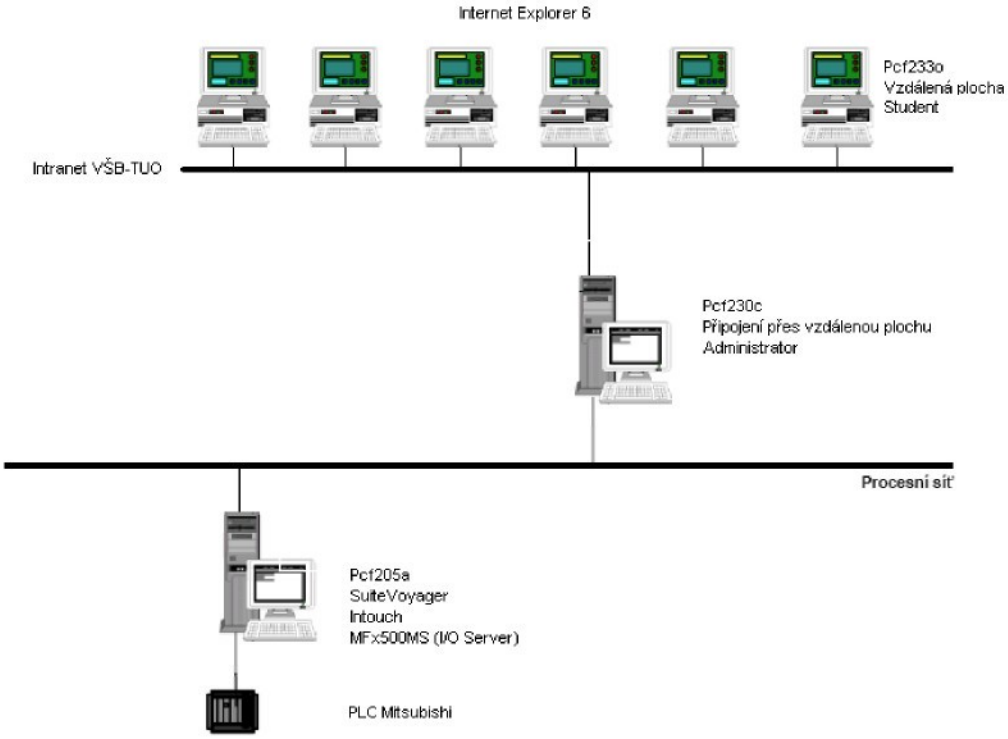
V případě dosažení minimální hladiny u více než dvou nádrží je operátor vyzván k zásahu funkcí alarmu, viz. obr.:



Obrázek 3.124 - Alarmové hlášení v aplikaci vizualizace

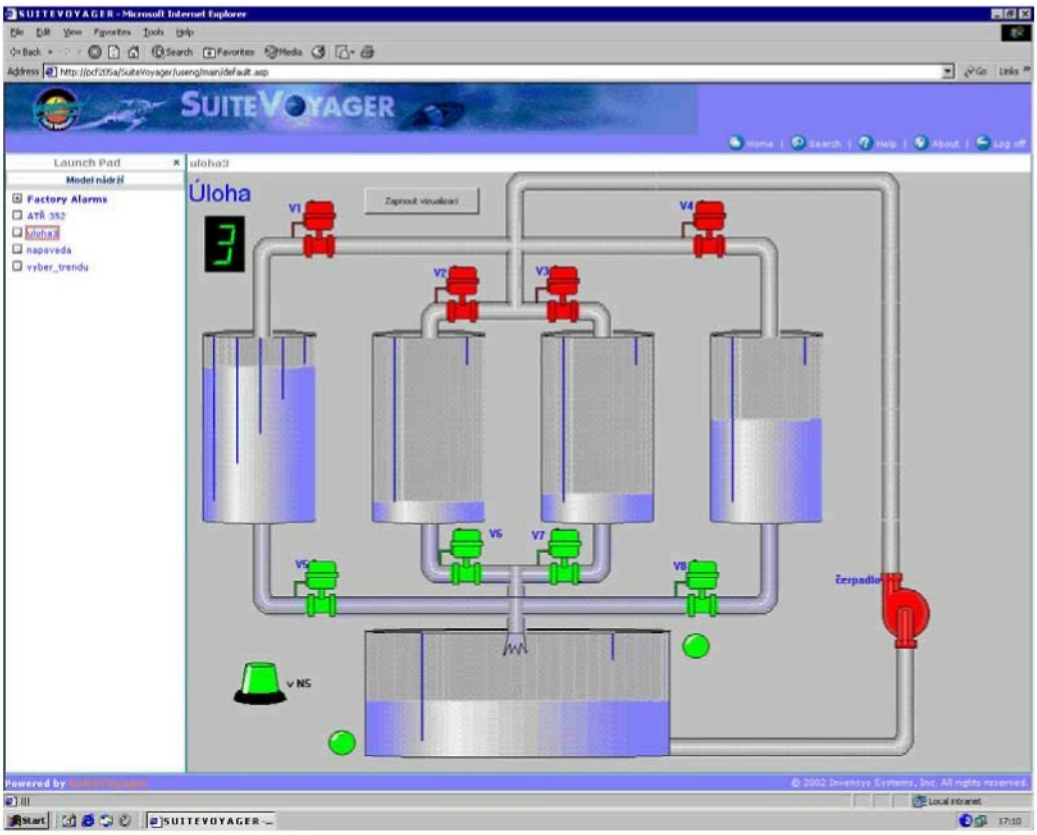
- **Řešení vizualizace dat z laboratorní úlohy do prostředí internetu**

Výrobní informační portál je prostředí, kam bývají publikována výrobní data z různých zdrojů, tj. aplikací běžících na vzdálených počítačích. Tato data pak portál vhodně organizuje, zabezpečuje a poskytuje definované skupině oprávněných klientů ze sítí Internet/intranet. Klientskou aplikací pro přístup k informacím publikovaným na portál je Microsoft Internet Explorer od verze 5.5. Klienti (uživatelé) nemusí mít na svých klientských počítačích instalován žádný speciální software (jsou to tzv. softwarově tencí klienti). Architektura systému pro pokročilé úlohy vizualizace je znázorněna na obrázku 3.125.



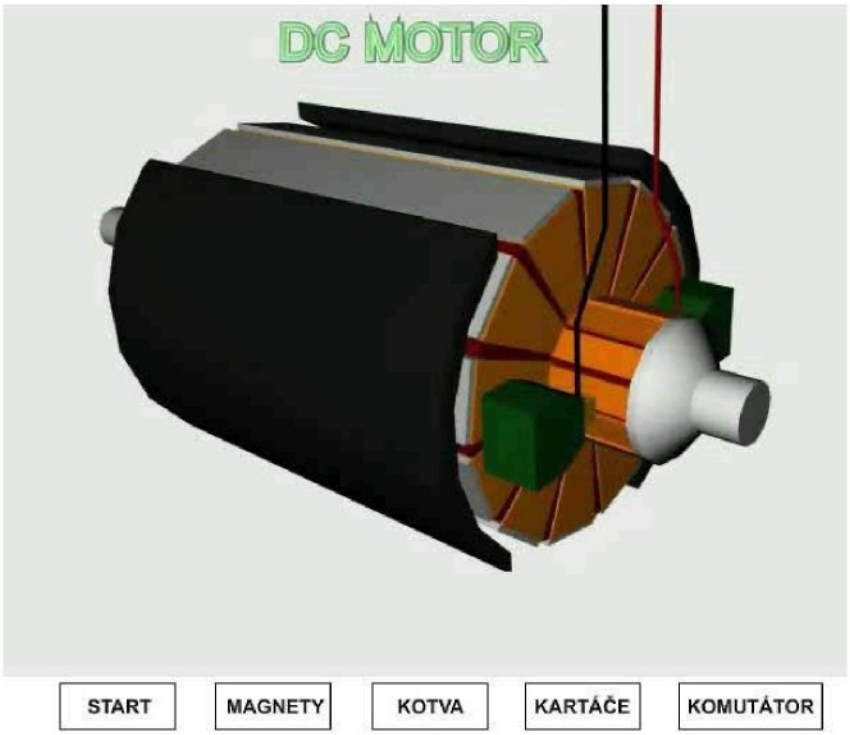
Obrázek 3.125 - Architektura systému

Vzhled aplikace v prohlížeči oprávněného klienta je vidět na obr. 3.126.

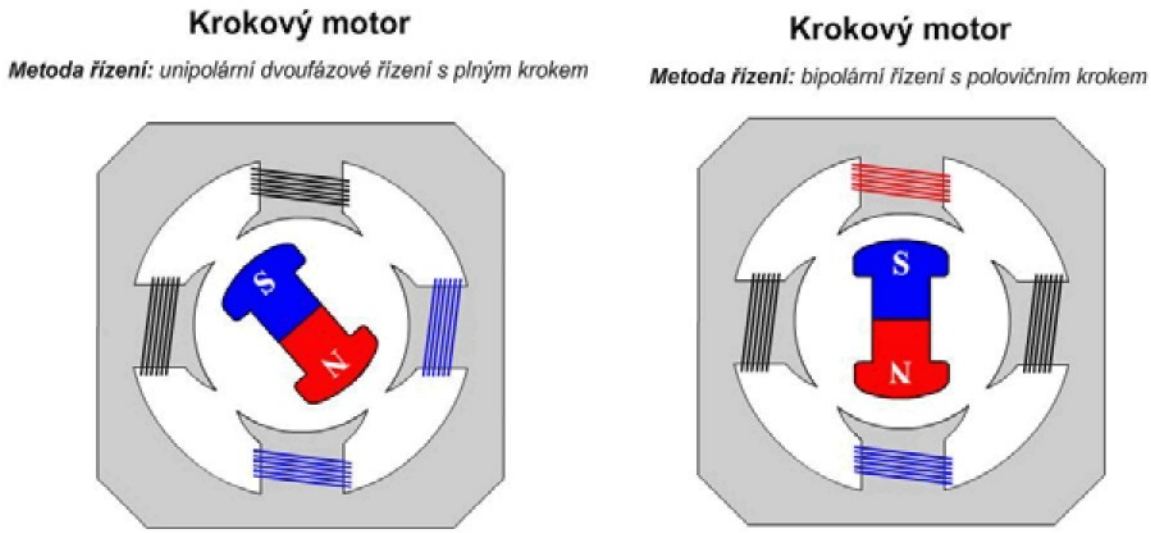


Obrázek 3.126 - Vizualizace úlohy v prohlížeči klienta

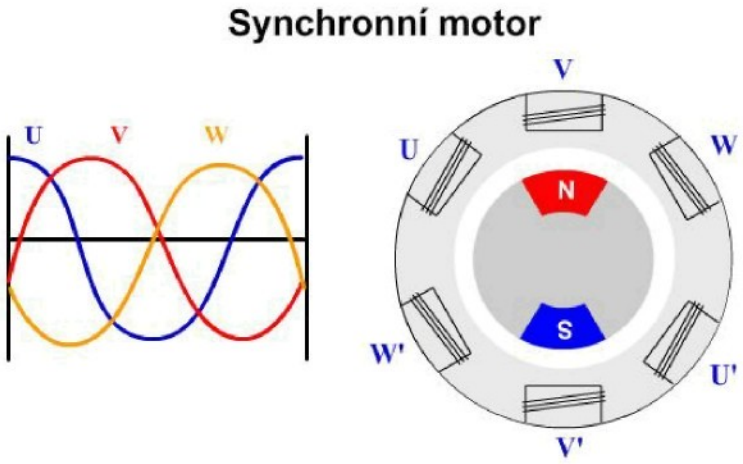
Kliknutím na následující obrázky lze spustit vytvořené animace, které vysvětlují funkci vybraných regulačních obvodů nebo akčních členů.



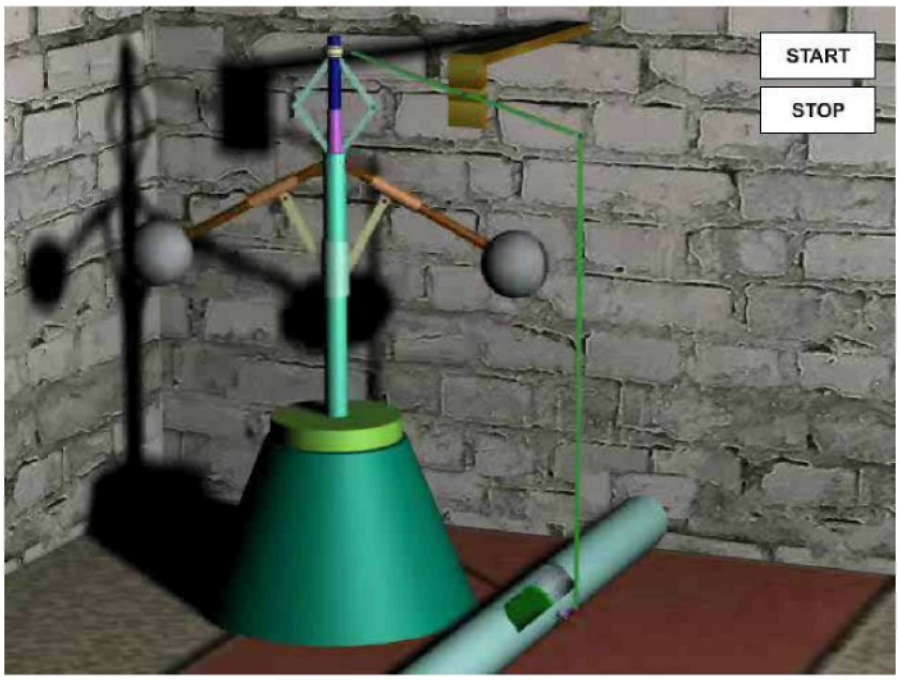
Obrázek 3.127 – Animace stejnosměrného motoru



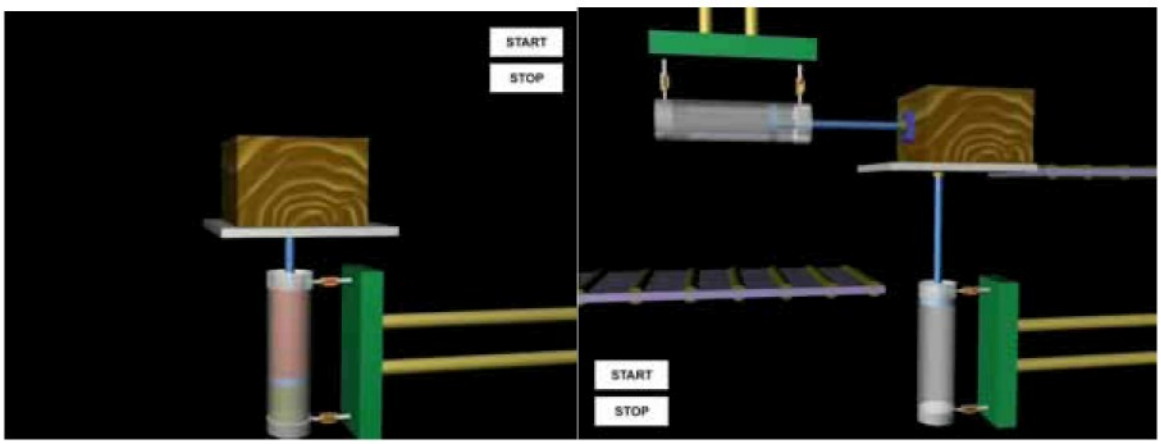
Obrázek 3.128 – Animace krokových motorů



Obrázek 3.129 – Animace synchronního motoru



Obrázek 3.130 – Animace funkce Wattova regulátoru



Obrázek 3.131 – Animace pneumatického (vlevo) a hydraulického (vpravo) polohování

Otázky 3.11.

1. Co je na nejnižším stupni hierarchické pyramidy řízení?
2. Co je v systému SCADA/HMI alarm?
3. Co jsou to historické trendy ve SCADA/HMI?

Další zdroje

[1] ORLÍKOVÁ, S. Měření průtoku tekutin - principy průtokoměrů. *Elektrorevue*. 2001/49. Dostupný z www: <URL:http://www.elektrorevue.cz/clanky/01049/>

[2] http://www-dt.fme.vutbr.cz/measure/Senzor/Cl-TerCl.htm

[3] http://uprt.vscht.cz/ucebnice/mrt/F4/F4k45-prut.htm

[4] VOLÁČEK, A, *Fóliové tenzometry - princip, provedení, použití, historie*. Dostupný z www: <URL:http://automatizace.hw.cz/view.php?cisloclanku=2005111201>

[5] BENEŠ, P., CHLEBNÝ, J., LANGER, J., MARTINÁSKOVÁ, J. & VORÁČEK, R *Automatizace a automatizační technika, Prostředky automatizační techniky*. Praha : Computer Press, 2000, 204 s. ISBN 80-7226-248-3.

4 APLIKACE AUTOMATIZAČNÍ TECHNIKY V TECHNICKÉ PRAXI

Čas ke studiu: 3 hodiny

Cíl Po prostudování tohoto odstavce budete umět

- definovat kaskádovou regulaci, integrační smyčkovou regulaci.
- popsat způsoby regulace teploty, dávkování směsí sypkých hmot, regulaci zásoby, řízení polohy.
- popsat řídicí jednotky automobilu a sběrnici CAN, řízení vstřikování a zapalování paliva zážehových motorů, servořízení.
- popsat použití GSM automatu pro zabezpečení „inteligentního domu".

Výklad

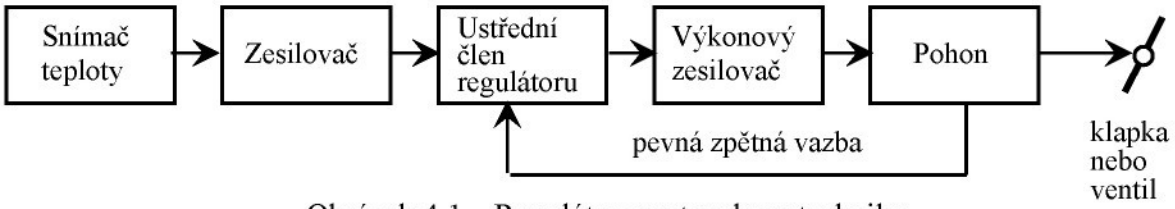
4.1. Regulace v tepelné technice

Regulace teploty

Teplota se reguluje v průmyslových ohřívacích pecích, sušárnách, někdy také v topeništi, například v plynovém hořáku a v různých chemických reaktorech. Podle akčních zásahů lze regulační obvody teploty rozdělit na dvě skupiny, a to akční zásahy do průtoku topného media respektive teplonosné kapaliny a regulační zásahy změnou velikosti elektrického proudu v topném vinutí.

První skupina ovládá průtok páry nebo horké vody a nebo plynu pro hořák. Používaným prvkům se říká regulační armatury. Těmito regulačními orgány je třeba posouvat nebo jimi otáčet. Slovo regulační znamená, že zvenčí se ovládá odpor škrticího orgánu proti proudění, nejčastěji velikosti průtočného průřezu. K pohonu slouží elektrické (nebo pneumatické) servomotory buď pákové nebo přímočaré. Součástí pohonu je snímač polohy regulačního orgánu, který se může pohybovat v rozsahu 0 až 100% otevření. Elektronické obvody ústředního členu regulátoru nemohou ovládat servomotor přímo, a proto je do obvodu zařazen výkonový zesilovač s tyristory a triaky. Zesilovač musí zajistit otáčení servomotoru dvěma směry (otvírání a zavírání).

Řetězec složený z ústředního členu regulátoru, zesilovače, servomotoru a regulačního orgánu (nejčastěji klapky nebo ventilu) je uzavřen tzv. pevnou zpětnou vazbou (tradiční označení), která přenáší do ústředního členu regulátoru informaci o poloze škrtícího orgánu. Tepelné procesy jsou obvykle z hlediska rychlosti odezvy na akční zásah velmi pomalé. Regulátor je vybaven možnosti chodu v automatickém nebo ručním (manuálním) režimu. Ruční režim znamená, že operátor velínu pohybuje svými příkazy regulačním orgánem a je současně informován o jeho poloze na ukazateli, který je součástí regulátoru. Blokové schéma regulátoru pro tepelnou techniku je znázorněno na obrázku 4.1.



Obrázek 4.1 – Regulátor pro tepelnou techniku

Zvláštní problematiku regulace teploty představují kotle pro výrobu obvykle přehřáté páry. Při ohřevu plynem jde o výše popsanou regulaci průtoku. Jinou možností je foukání práškového uhlí a nebo jeho dávkování na rošt. Zde se uplatňují regulovatelné podávače nebo přímo dávkovací váhy.

Přesná regulace teploty přehřáté páry se realizuje také vstřikem studené vody do vody před přehřívákem jako průtočným kotlem, což je znázorněno na schématu v obrázku 4.2. Teplota páry T_1 a horké páry T_2 se měří termočlánky. Je zřejmé, že odezva teploty páry v místě před přehřívákem na změnu skokovou změnu přídavku studené vody bude rychlejší než odezva teploty přehřáté páry na tuto akční veličinu regulačního obvodu. Ve zmíněném obrázku je různost odezev demonstrována na přechodových charakteristikách. Regulace teploty se realizuje kaskádově dvěma regulátory. První regulátor má žádanou hodnotu rovnou teplotě přehřáté páry. Výstup tohoto regulátoru představuje žádanou hodnotu teploty páry po vstřiku studené vody pro druhý regulátor, který průtok této vody reguluje. S ohledem na vyloučení trvalé regulační odchylky je první regulátor typu PI, tj. s integrační složkou, která tuto funkci zajišťuje. Druhý regulátor je typu PD. Derivační složka přispívá ke stabilite soustavy.

Průmyslové pece ohřívané elektrickým proudem mají topné vinutí z odporového materiálu (např. konstanta). Malé uhlíkové retorty jsou ohřívány přímo průchodem elektrického proudu. Při regulaci teploty se používají podle funkce dva druhy snímačů, často termočlánků nebo odporových teploměrů. Jeden, tzv. regulační snímač teploty bývá umístěn tak, že má rychlou odezvu na akční zásah. U elektricky ohřívaných pecí je situován co nejblíže topného vinutí. Druhý snímač se nazývá měřící a je umístěn do prostředí, jehož teplota se má regulovat. Záznam teploty z tohoto snímače je obvykle rozhodující pro správnou funkci regulačního obvodu.

Pro velkou setrvačnost regulované soustavy je regulace vykonávána dvouhodnotovým regulátorem (zapnuto-vypnuto). Malé proudy lze regulovat plynule. Proud pro laboratorní pece se v dávné minulosti plynule měnil magnetickými zesilovači, nyní jsou to již výhradně polovodičové prvky.

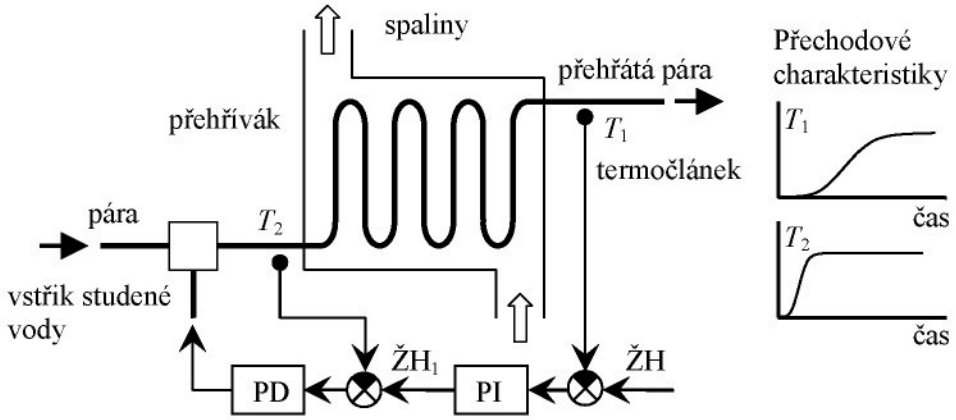
❑ Kaskádová regulace v tepelné technice

V tepelné technice je běžné, že regulovaná soustava má pomalou odezvu, tj. relativně velké dopravní zpoždění a velkou časovou konstantu, nebo řád jejího přenosu je také vyšší než je obvyklé. Po změně akční veličiny následuje velmi pomalá odezva v časovém průběhu regulované teploty. Těmto pro regulaci nevýhodným podmínkám se čelí tzv. kaskádovou regulací. V regulované soustavě se vyhledá teplota, jejíž odezva je podstatně rychlejší než odezva regulované teploty.

Příklad kaskádové regulace je ilustrován na přehříváku tlakové páry na obrázku 4.2. Znázorněný přehřívák je průtočný kotel. Teplota páry pro například turbínu je požadována velmi stálá. Kolísání teploty páry a tím i turbíny může urychlit únavu materiálu, ze kterého je vyrobena. Teplota výstupní páry se reguluje změnami průtoku studené vody, která je

vstřikována do páry před přehřívákem. Odezva na vstřik studené vody ve změně teploty páry před ohřívákem je téměř okamžitá. Součástí obrázku 4.2 jsou přechodové charakteristiky demonstrující rozdílnost odezev u obou zmíněných teplot páry na skok průtoku studené vody.

Kaskádové zapojení obsahuje dva regulátory. První regulátor reguluje teplotu páry na výstupu přehříváku. Akční veličina tohoto regulátoru je žádanou hodnotou teploty páry na vstupu do přehříváku pro druhý regulátor. Tento druhý regulátor pracuje s regulovanou veličinou, kterou je teplota páry na vstupu do přehříváku, a s akční veličinou, kterou je průtok studené vody. První regulátor musí pracovat bez trvalé regulační odchylky a proto obsahuje integrační člen. Druhý regulátor musí pracovat s co nejrychlejší odezvou a proto obsahuje derivační složku.

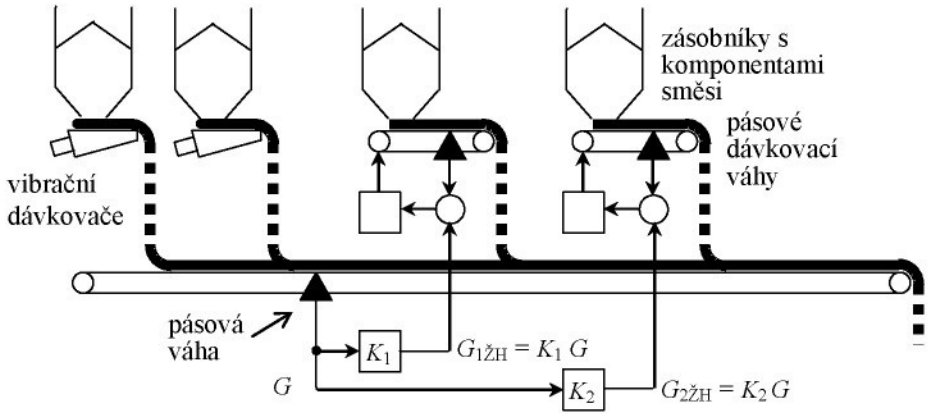


Obrázek 4.2 – Kaskádová regulace

4.2. Řízení dávkování směsi a řízení zásoby sypkých hmot

❑ Dávkování směsi sypkých hmot

Regulace průtoku sypkých hmot a materiálů má široké uplatnění při vytváření směsí nebo dodávku materiálu do technologického zařízení. Příkladem může být zavážení paliva do kotle nebo spékání aglomerátu v hutích a nebo výroba stavebních hmot. Směs lze připravovat obecně dvojím způsobem. První je jednoduché odvážení potřebné dávky, které se následně nějakým způsobem jednorázově promíchají. Toto dávkování nelze označit slovem kontinuální regulace. Druhý způsob vytváření směsi je nepřetržité a má svoje uplatnění v technologických procesech, které probíhají kontinuálně. Dávkuje se obvykle na dopravní pás, přičemž směs je dále promíchána v průtočném míchacím zařízení, například v otáčejícím se bubnu s mírně skloněnou osou, a pokračuje k dalšímu zpracování. Kontinuální dávkování vyžaduje průběžné měření průtoku směsi po dopravníku.



Obrázek 4.3 – Dávkování směsi

Uspořádání systému dávkování směsi může kombinovat použití levnějších postupů vynášení základní komponenty směsi, přičemž je měřen její celkový průtok, a k ní se regulovatelnou pásovou dávkovací váhou přidávají zbylé komponenty. Uspořádání systému řízení dávkování směsi je znázorněno na obrázku 4.3. Základní komponenta směsi se vynáší ze zásobníků na sběrná dopravní pás vibračními podávači. Hmotový průtok této komponenty G je použit k odvození žádané hodnoty pro obě pásové dávkovací váhy.

❑ Regulace zásoby

Obrázek 4.4 – Provedení snímačů zatížení

Kontinuální technologické procesy probíhají obvykle v několika stupních tak, že materiál postupuje od jednoho stupně k druhému. Mezi těmito stupni se vytváří zásoba, aby bylo možné vykrýt neplánované prostoje a kontinuita procesu v daném stupni nebyla porušena. Při výrobě aglomerátu (kovonosná vsázka pro vysoké pece po zkusovění prachových rud spékáním) je technologie výroby rozdělena do dvou stupňů a to přípravy směsi a pak spékání na spékacích pásech. Před spékacími pásy jsou mezizásobníky pro hodinu provozu spékacích pásů. Velikosti této zásoby je třeba přizpůsobit dodávku směsi ze stupně přípravy směsi.

Kriterium regulace není přesně dodržovaná hladina zaplnění, ale pouze předejití situaci, kdy se zásobníky vyprázdní nebo přeplní. Zároveň nejsou žádoucí prudké výkyvy v dodávce směsi do zmíněných mezizásobníků.

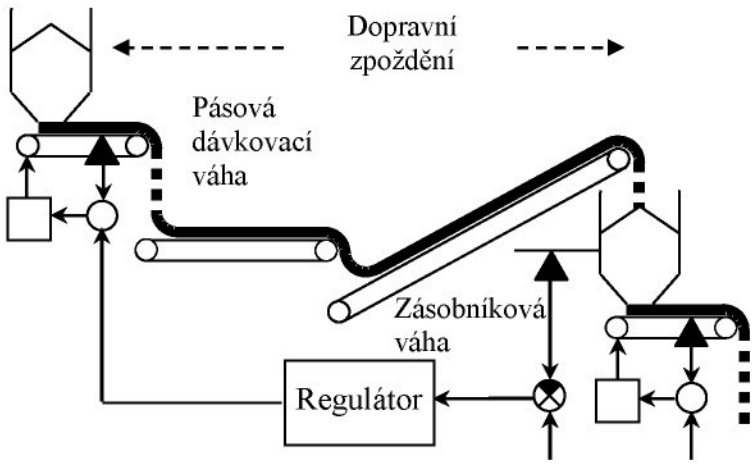
Velikost zaplnění (zásobu) je tedy třeba regulovat. K tomuto účelu je třeba průběrně znát velikost zaplnění a mít možnost ovládat výkon dávkovacího zařízení. Velikost zásoby se snímá různými snímači. Nejoblíbenější jsou tenzometrické váhy (tenzometrické krabice) s rozsahem desítek tun, na které se konstrukce zásobníku prostě umístí. Staticky určité je umístění zásobníku na tři tenzometrické krabice. S únosnou přesností lze uložit zásobník také na čtyři snímače. Měřící rozsah snímačů Schenck typu RTN je od 1 do 470 t. Při konstrukci je třeba dbát na uchycení horizontálních sil a zabránění klopení zásobníku. Ukázka snímačů a umístění tenzometrických krabic je na obrázku 4.4 a 4.5.

Dalším typem měření velikosti zásoby v zásobnících jsou kapacitní sondy založené na principu měření kapacity mezi materiálem a proměnlivě ponořenou elektrodou sondy. Jsou také čidla založena na odměřování výšky hladiny, například mechanicky (příkladem je vysoká

pec s periodicky spouštěným závažím), dobou odrazu ultrazvuku a nebo radioizotopy s čidlem sledujícím rozhraní mezi stínem a volným průchodem záření.

Hladina kapalin se snímá podobně, zvláštní předností je plochost hladiny. U kapaliny lze využít ke zjišťování výšky hladiny hydrostatickým tlakem. Podobně množství plynu v uzavřené nádobě je dáno jeho tlakem.

Obrázek 4.5 – Zásobníková váha se třemi snímači zatížení



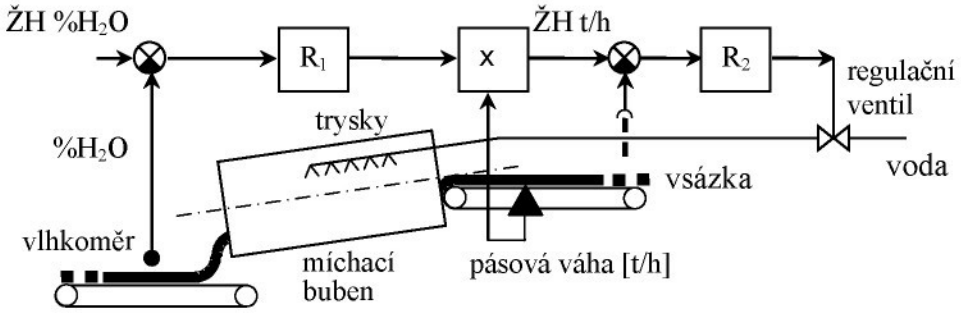
Obrázek 4.6 – Regulační obvod zásoby

Účelem regulace zásoby je vyrovnání produkce jednoho výrobního úseku s nerovnoměrností spotřeby v následujícím výrobním úseku. Charakteristickou vlastností regulované soustavy je dopravní zpoždění mezi zásobníkem a prvky ovlivňujícími velikost průtoku materiálu do zásobníku, jak je znázorněno na obrázku 4.6. Odůvodněním zpoždění jsou dopravní cesty z pásových dopravníků, které se pohybují stálou (neproměnlivou) rychlostí. Regulace rychlosti pohybu dopravníků se může jevit pro nákladnost plynule regulovaných pohonů nepřijatelná.

❑ Kontinuální regulace vlhkosti sypkých hmot

Některé technologické procesy vyžadují, aby tok vstupní suroviny obsahoval zadané procento vlhkosti. Takovou surovinou může být například aglomerační vsázka z prachové rudy, paliva a přísad. Optimální prodyšnost této směsi na spékacím pásu zajistí určité procento vlhkosti. Spékací pás vyrábí aglomerát pro vysoké pece. Shodný problém může být při výrobě cementu nebo jiných stavebních hmot. Směs je přirozeně málo vlhká a navíc její vlhkost se nahodile mění. Proto se dovlhčuje v míchacím (peletizačním) bubnu, kde se při otáčení převaluje a sbaluje do pelet (kuliček). Regulační obvod dávkování vody do rozprašovacích trysek sestává ze dvou regulátorů jak je znázorněno na obrázku 4.7. První regulátor R_1 určuje akční zásah na základě rozdílu žádané hodnoty a skutečné výsledné vlhkosti. Vlhkost se měří kontinuálně například na principu pohlcování infračerveného záření nebo pohlcování neutronů. Výstup regulátoru R_1 je ve významu dodané vlhkosti v procentech. Násobením tohoto procenta hmotnostním průtokem vsázky se dostane hmotnostní průtok vody do trysek uvnitř míchacího bubnu. Regulátor R_2 tento požadovaný průtok vody reguluje změnami otevření ventilu.

Vzájemné propojení regulačních obvodů je označováno jako vlečná regulace. Ve srovnání s kaskádovou regulací se jedná o shodné zapojení. Z důvodu zvyklostí se používá název vlečná regulace.



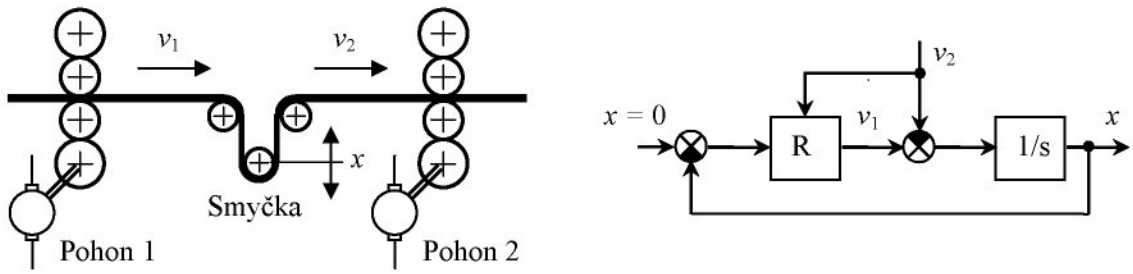
Obrázek 4.7 – Vlečná regulace

❑ Integrační smyčková regulace

Tento způsob uspořádání regulačního obvodu je používán v případech, kdy je zapotřebí synchronizovat otáčky dvou různých zařízení. Příkladem může být pohon navíjecího bubnu přičemž drát nebo pás je podáván svou vlastní rychlosti. Je evidentní, že průměr cívky se při navíjení mění, což vyžaduje přizpůsobení otáček navíjecího bubnu. Jiným příkladem je papírenský stroj, u kterého se předává pás papíru od jedné dvojice válců k jiné dvojici válců, přičemž každá dvojice válců má vlastní pohon. Nedokonalým řešením regulace pohonu by bylo společné zadání otáček oběma regulátorům nebo odvození žádané hodnoty otáček druhého regulátoru od skutečných otáček prvního regulátoru. I sebemenší chyba v měření otáček by vedla k integrování jejich rozdílu a posléze k roztržení pásu papíru nebo naopak k jeho havarijnímu nakupení mezi válci.

Tomuto jevů se čelí vložením smyčky, jejíž délka se reguluje změnou rychlosti pohybu jednoho z pohonů. Na obrázku 4.8 je uspořádání obvodu integrální smyčkové regulace. Vlevo je znázorněn regulovaný objekt s válci, dvěma pohony a vloženou smyčkou. Pohon 2 je řízen nezávisle na pohonu 1. Tento pohon určuje rychlost pohybu pásu v_2. Délka smyčky x je výsledkem integrace (symbolicky 1/s) rozdílu rychlostí v_1 a v_2. Regulátor R určuje otáčky pohonu 1 a tím i rychlost v_1. Rychlost v_2 je měřenou poruchou a představuje pro regulátor R přímou vazbu. Regulační odchylkou je velikost prodloužení nebo zkrácení smyčky x od nulové polohy.

.

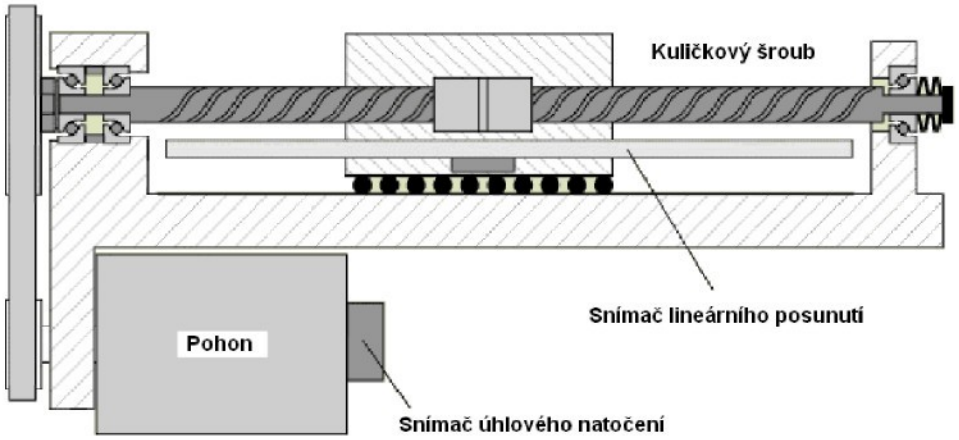


Obrázek 4.8 – Integrační smyčková regulace

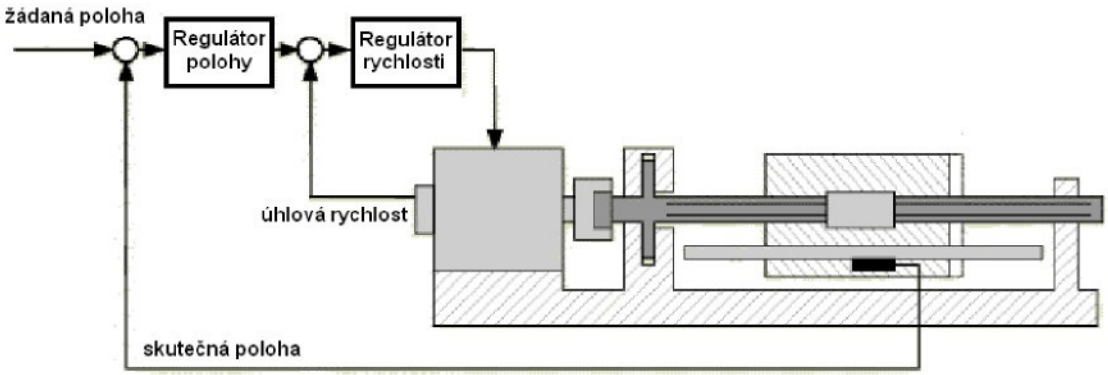
4.3. Řízení polohy

Nejpřesnější systémy řízení polohy mají uplatnění pro řízení posuvů obráběcích strojů. Dalším požadavkem je velký rozsah rychlostí pohybů v závislosti na tom, zda se obráběcí nástroj jen přesouvá nebo je v záběru. Tyto požadavky plní kombinací krokového motoru a kuličkového šroubu (posuvy) nebo harmonické převodovky (otočení). Jak již bylo zmíněno, odměřování

polohy je dvojí: přímé a nepřímé. Příklad tohoto uspořádání je na obrázku 4.9. Snímač úhlového natočení na pohonu odměřuje polohu stolu s obrobkem nepřímo a lineární snímač odměřuje tuto polohu přímo.



Obrázek 4.9 – Typické uspořádání pohonu pro posuvný pohyb obráběcího stroje

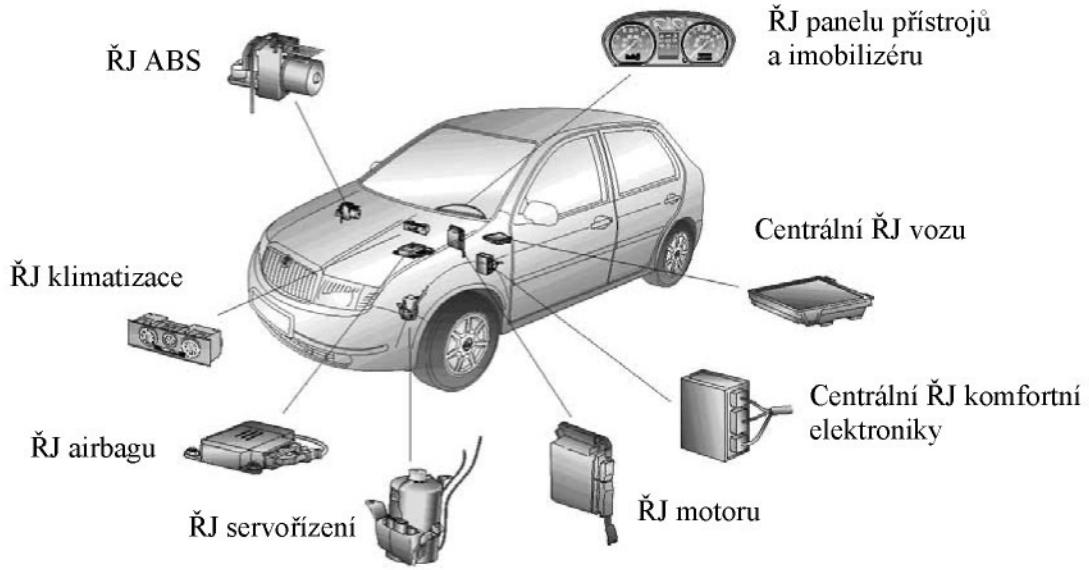


Obrázek 4.10 – Zapojení regulačního obvodu posunu s vnitřní rychlostní zpětnou vazbou

Blokové schéma systému řízení polohy u obráběcího stroje je znázorněno na obrázku 4.10. Obvod obsahuje zpětnou vazbu od rychlosti pohybu k regulátoru rychlosti a zpětnou vazbu od polohy s vlastním polohovým regulátorem. Akční veličina polohového regulátoru je žádanou hodnotou rychlostního regulátoru. Okamžitá rychlost pohybu je odvozena od frekvence impulsního signálu snímajícího úhel natočení poháněcího motoru.

4.4. Soubor řídicích jednotek automobilu a sběrnice CAN

Následující příklady jsou vybrány tak, aby byly blízké profesi strojního inženýra. Například moderní osobní automobil obsahuje řadu řídicích jednotek s mikropočítači, které plní koordinovaně různé funkce. Mikropočítače nepředstavují izolované uzly, ale vzájemně komunikují prostřednictvím sběrnice CAN (Control Network Area), která byla navržena firmou Bosch. Ukázka vybavení automobilů je na obrázku 4.11.



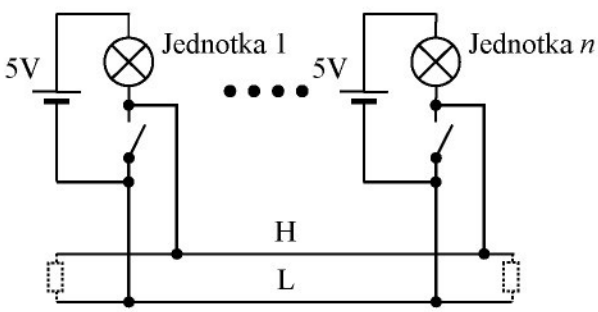
Obrázek 4.11 – Řídicí jednotky (ŘJ) u vozidel ŠKODA

Sběrnice CAN představují dva zkroucené vodiče, které mají označení L a H. Protože jsou po sběrnici posílány vysokofrekvenční impulsy, je vedení zakončeno odporem o velikosti jeho charakteristického odporu, aby nevznikaly odrazy vysokofrekvenčního vlnění a tak nebyly zkreslovány impulsní signály. Princip sběrnice CAN lze demonstrovat na schématu na obrázku 4.12, ve kterém jsou použity jen spínače a žárovky. Pro jednu řídicí jednotku je to obvod s jednou žárovkou, zdrojem a spínačem. Tyto spínače od všech jednotek jsou prostřednictvím sběrnice spojeny paralelně. Ve skutečnosti je spínání bezkontaktní a stav žárovky představuje logickou proměnnou řídicí jednotky, avšak vedení sběrnice je shodné. Jestliže jsou všechny spínače rozepnuty (zakončovací odpory pro jednoduchost zanedbáme), pak je na sběrnici +5V, tj. plné napětí a žádná žárovka nesvítí. Jestliže jedna jednotka svůj spínač sepne, pak se napětí na sběrnici sníží na nulu a hlavně, všechny ostatní jednotky se o tom dovědí, protože jejich žárovky se rozsvítí. Aby na sběrnici nevznikl chaos, jsou stanovena pravidla, tzv. datový protokol, podle kterých se jednotlivé jednotky řídí.

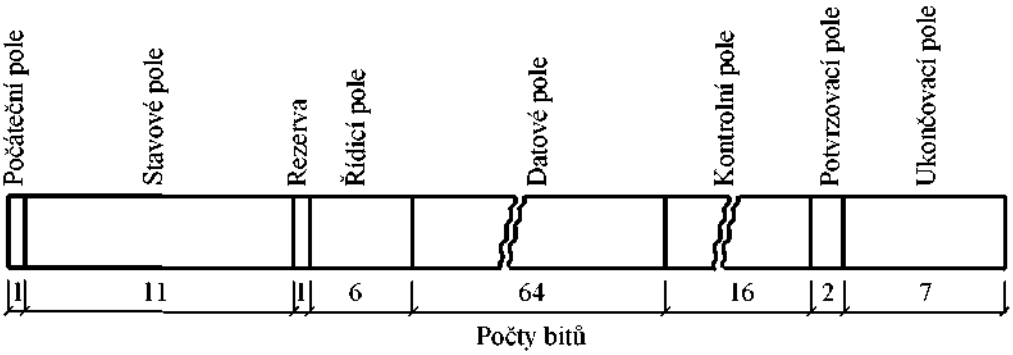
Sepnutí a rozepnutí kontaktu představuje informaci jednoho bitu. Samozřejmě, že je třeba přenést bohatší informaci o více bitech, proto jsou bity sdruženy do zprávy, tzv. datového rámce (Data Frame) s určitým uspořádáním. Jejich délka není stejná a závisí na délce jedné částí datového rámce, která obsahuje tzv. datové pole. Obsah datového rámce je na obrázku 4.13.

Datový rámec se skládá ze sedmi polí:

- počáteční pole (1 bit)
- stavové pole (11 bitů)
- řídicí pole (6 bitů)
- datové pole (maximálně 64 bitů)
- kontrolní pole (16 bitů)
- potvrzovací pole (2 bity)
- ukončovací pole (7 bitů)



Obrázek 4.12 – Princip funkce sběrnice CAN



Obrázek 4.13 – Datový rámec (Data Frame)

Počáteční pole - Start of Frame označuje počátek datového přenosu.

Stavové pole - Arbitration Field obsahuje informaci o prioritě datového přenosu, tj. stanovuje, která jednotka má při vysílání zpráva přednost. Zároveň je ve stavovém poli informace o obsahu zprávy, například, že se jedná o otáčky motoru.

Řídicí pole - Control Field obsahuje jako kód počet informací, které jsou obsaženy v datovém poli.

Datové pole - Data Field přenáší informace důležité pro ostatní jednotky, které zprávu přijmou. Délka zprávy je až 64 bitů a je členěna po bajtech sdružujících 8 bitů.

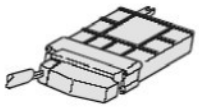
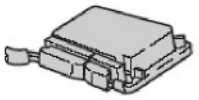
Kontrolní pole - CRC Field (CRC = Cyclical Redundancy Check) slouží ke zjišťování chyb v přenosu. Kontrola se provádí metodou cyklického výpočtu kontrolního kódu.

Potvrzovací pole - ACK Field (ACK = Acknowledgement). V tomto poli potvrzuje příjemce zprávy jejímu odesilateli, že zpráva byla přijata bez chyb. Jestliže se však zjistí chyba, pak je to odesilateli zprávy sděleno a dochází k opakování vysílání zprávy.

Ukončovací pole (End of Frame). V tomto poli kontroluje vysílač svůj datový protokol a potvrdí objektu, který zprávu vyslal, zda je v pořádku. Jestliže není, dojde okamžitě k přerušení a opakovanému zahájení přenosu. Tím je datový přenos protokolu ukončen.

Jestliže chce více jednotek vysílat a sběrnice je volná, je třeba rozhodnout, která z nich má přednost, tj. vyšší prioritu než jiná jednotka. Z důvodu bezpečnosti bude jistě potřebné dát přednost zprávě z jednotky ABS nebo ABS s EDS před zprávou z automatické převodovky. Rozhodnutí o prioritě je na základě srovnání bitů ze stavového pole. Každá jednotka ví, zda pomyslný spínač sepnula (logická nula na sběrnici) a nebo nesepnula, a také ví, v jaké poloze má svůj spínač. Ty jednotky, které spínač samy sepnuly mají vyšší prioritu než jednotka, která spínač nesepnula a takto by umožnila nechat na sběrnici logickou jedničku. Jednotka vysílající bit jednotkový, tj. bez priority, na některém z 11 bitů stavového poli, přičemž na sběrnici je logická nula, přestane okamžitě vysílat a začne pasivně přijímat bity, které vysílají jednotky s vyšší prioritou. Po srovnání všech bitů stavového zbude jednotka s nejvyšší prioritou. Diagram dokumentující výše popsaný postup je v tabulce 4.1.

Tabulka 4.1 – Seznam čidel na motoru

		Start of Frame	Arbitration Field = 11 bit				
			Bit 1	Bit 2	Bit 3		
Řídicí jednotka automatické převodovky		0	1				...
Řídicí jednotka ABS nebo ABS a EDS		0	0	0	1	0	...
Řídicí jednotka motoru Motronic		0	0	1			...
Vodiče datové sběrnice		0	0	0	1	0	...

4.5. Řízení vstřikování a zapalování paliva zážehových motorů

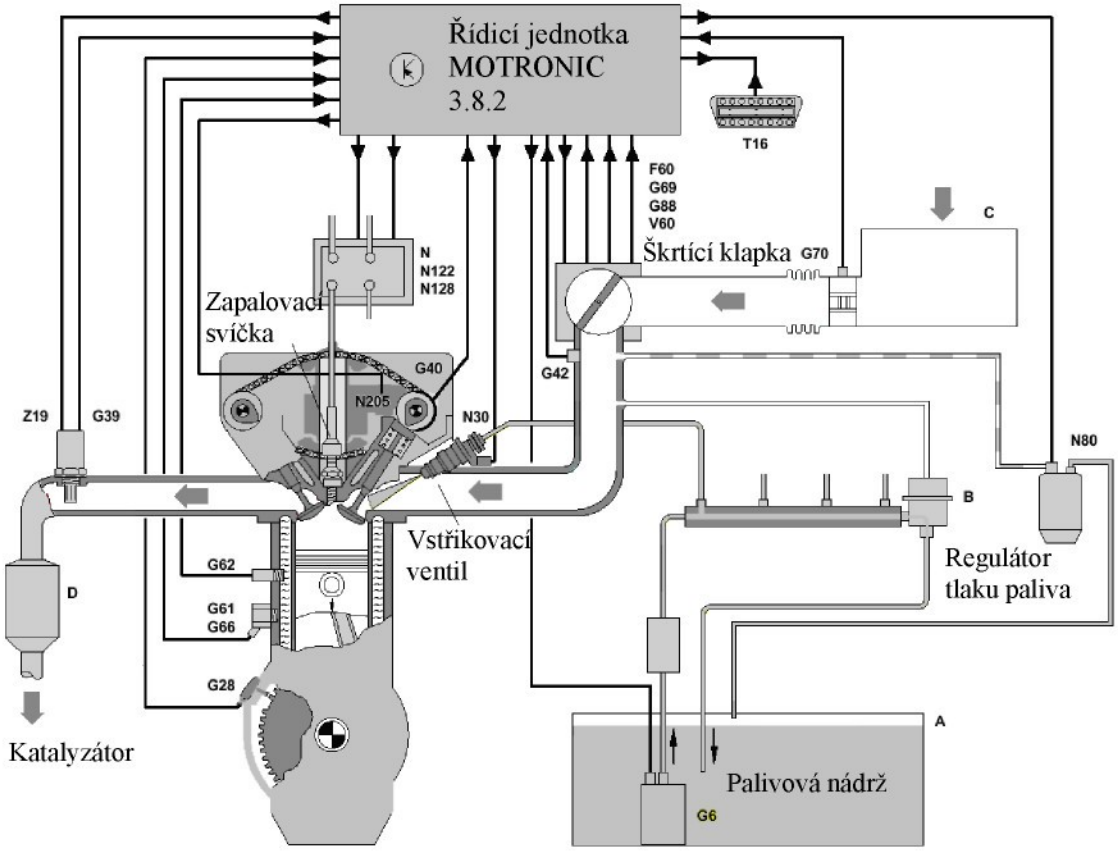
Jako příklad aplikace řídicí jednotky motoru je vybrán benzinový motor vozidla ŠKODA OCTAVIA 1,8 l s označením AGN o výkonu 92 kW. Jedná se o řadový čtyřválec s pěti ventily na válec, jeho dvě vačkové hřídele jsou umístěny v hlavě válce. U tohoto motoru je použita řídicí jednotka Bosch Motronic 3.8.2 se sekvenčním vstřikováním a datovým polem řízeným zapalování s regulací klepání každého válce. Účinnost spalování je řízena podle údajů lambda sondy. Způsob řízení podle datového pole se rozumí nastavování hodnot parametrů podle tabulky v závislosti na jednom nebo více parametrech, obvykle měřených, které určují okamžitý stav motoru (teplota, otáčky). Schéma zapojení je na obrázku 4.14.

Jak každý motorista ví, řidič určuje výkon motoru stupněm sešlápnutí plynového pedálu. Tímto způsobem je uzavřen prostřednictvím rozhodovací funkce řidiče regulační obvod rychlosti vozidla. U starších automobilů byl pedál mechanicky spojen s klapkou v karburátoru. U dále popisovaného systému toto mechanické spojení chybí. Poloha pedálu je přenesena do řídicí jednotky motoru, která škrtící klapku nastaví do požadované polohy pomocí servomotoru samočinně. K přestavování je přirozeně využita zpětná vazba od skutečné polohy škrtící klapky. Poloha klapky určuje průtok vzduchu do motoru. Úměrně průtoku vzduchu je vstřikovačem dávkováno palivo. Množství vstřikovaného paliva je řiditelné dobou otevření vstřikovacích ventilů. Poměr průtoku paliva k průtoku vzduchu je ovlivňován dalšími parametry. Chladný motor například vyžaduje více paliva. Účinnost spalování je kontrolována lambda sondou, což je snímač přebytku nebo kyslíku ve výfukových plynech. Lambda-sonda porovnává koncentraci zbytkového kyslíku ve výfukových plynech s koncentrací kyslíku v referenčním vzduchu (výfukové plyny se porovnávají s okolním vzduchem, protože vnitřek lambda-sondy je otvorem spojen s okolní atmosférou). Lambda-sonda pracuje jako tzv. „dvoubodová“ sonda s reléovou charakteristikou. Ukazuje jen, zda je ve výfukových plynech směs bohatá ($\lambda < 1$), nebo chudá ($\lambda > 1$). Ke své funkci se musí sonda ohřát integrovaným ohřevem na teplotu aspoň 350 ^{0}C. Na rozdíl od dvoubodové lambda sondy lineární sonda ukazuje stupeň přebytku kyslíku spojitě.

V celém systému řízení motoru existuje jediná trvalá zpětná vazba a to od lambda-sondy. Řídicí jednotka však realizuje množství vazeb mezi veličinami, které se jen z částí

dříve dařilo vytvořit mechanickým spojením, jako bylo například ovlivňování předstihu zapálení od podtlaku v sacím potrubí.

Volitelně lze využít funkce pro řízení rychlosti vozidla, kdy další uzavřený regulační obvod přebírá funkci automatického ovládání plynového pedálu podle odchylky mezi požadovanou a skutečnou rychlosti vozidla (tempomat) obvykle jen pro nejvyšší převodový stupeň.



Obrázek 4.14 – Řízení vstřikování a zapalování paliva motoru 1,8 l AGN u vozidel ŠKODA OCTAVIA

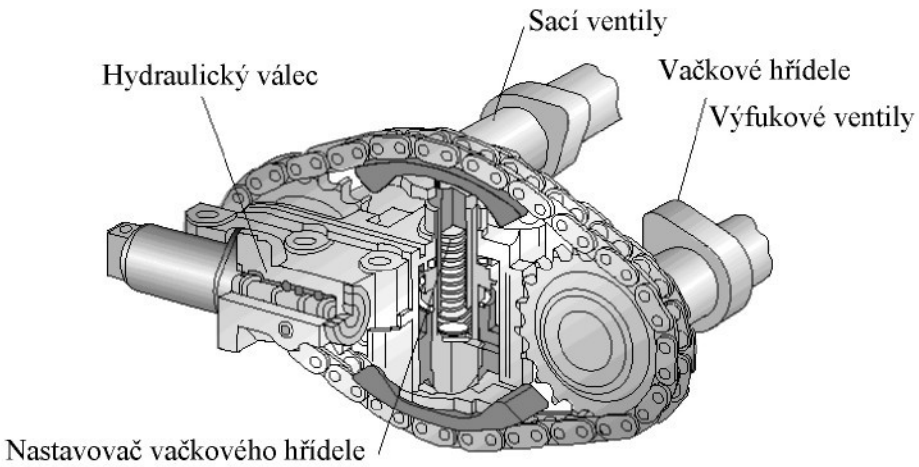
Mezi technické znaky systému sběru informací pro funkci řídicí jednotky patří

- měření polohy klikového hřídele na základě impulsů snímače G28, které jsou odvozeny z 60 rovnoměrně rozložených zubů na jednu otáčku, přičemž dva zuby jsou odfrézovány a tato mezera v řadě impulsů slouží k indikaci (s předstihem) dosažení horní úvratě válců
- rozpoznání fáze motoru v dvojotáčkovém cyklu podle údaje Hallového snímače G40 na hlavě válce motoru pro snímání polohy sacího vačkového hřídele (jeden impuls za otáčku).

Soupis ostatních čidel je uveden v tabulce 4.2. Pamětníci vědí, že jediným snímačem na motoru byl dříve snímač tlaku oleje a na karburátoru někdy pomocný elektrický ventil. Zapalovací soustava (přerušovač, rozdělovač a svíčky) byla dříve také velmi jednoduchá a umožňovala jen podtlakovou regulaci předstihu zapalování. Zmíněná tabulka dokládá pokrok techniky a potřebu interdisciplinární kvalifikace inženýra strojaře v oblasti řídicí techniky.

Tabulka 4.2 – Seznam čidel na motoru

Snímač	Popis funkce	Snímač	Popis funkce
F60	koncový spínač volnoběžných otáček	N	zapalovací cívka 1
G88	potenciometr škrtící klapky pro řízení volnoběhu	N80	elektromagnetický ventil nádobky s aktivním uhlím
G69	potenciometr škrtící klapky	N122	koncový výkonový stupeň I
G6	palivové čerpadlo	N128	zapalovací cívka 2
V60	nastavovač škrtící klapky	N192	koncový výkonový stupeň 2
G39	Lambda sonda	G28	snímač otáček motoru
G40	Hallův snímač otáček	G62	čidlo teploty chladicí kapaliny
G42	čidlo teploty nasávaného vzduchu	G70	snímač množství nasávaného vzduchu
G61	snímač klepání (I)	J338	jednotka ovládání škrticí klapky
G66	snímač klepání (II)	RTP	regulátor tlaku paliva
J220	řídicí jednotka motoru Motronic 3.8.2	N205	elektromagnetický ventil nastavování vačkového hřídele
Z19	Ohřev lambda sondy	T16	připojení diagnostiky
N30 až 33	Vstřikovací ventily		



Obrázek 4.15 – Zařízení ke změně časování ventilů

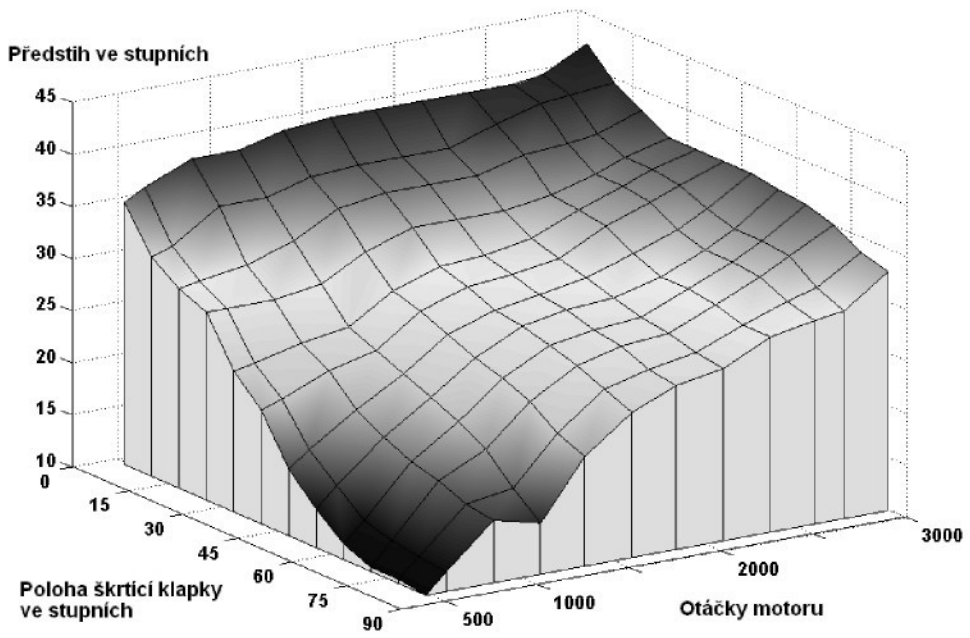
Charakteristika motoru (závislost kroutícího momentu na otáčkách) se ovlivňuje proměnlivým časováním rozvodu. Proto je motor opatřen nastavovacím zařízením (obrázek 4.15) pro sací vačkový hřídel, které je řízeno datovým polem, což umožní žádoucím způsobem ovlivňovat charakteristiku motoru. Vačkový hřídel sacích ventilů a vačkový hřídel

výfukových ventilů leží vedle sebe. Vačkový hřídel výfukových ventilů je poháněn ozubeným řemenem od klikového hřídele. Vačkový hřídel sacích ventilů je poháněn od vačkového hřídele výfukových ventilů pomocí řetězu. Řetěz je napínán nastavovačem vačkového hřídele. Ke změně časování dojde přesunutím bodu obratu řetězu ve svislém směru. Tím se zároveň natočí vačkový hřídel sacích ventilů. Nastavovač (aktuátor) je ovládán elektricky řízeným hydraulickým válcem, který je umístěn v nastavovači vačkového hřídele. Ventil je řízen podle zmíněného datového pole prostřednictvím elektromagnetického ventilu.

Řídicí jednotka je výkonný mikropočítač, který zpracovává údaje z několika čidel, které ovlivňují jak zapalování paliva tak jeho vstřikování. Pro funkci zapalování je v paměti mikropočítače uloženo datové pole určující okamžik zapálení paliva (úhel předstihu před horní úvratí pístu). Základní předstih je určen otáčkami motoru a některým parametrem charakterizujícím průtok vzduchu. Může to být otevření škrtící klapky, nebo sací podtlak a nebo přímo měřený průtok vzduchu. Ukázka datového pole pro určení základního předstihu je ve formě 3D grafu na obrázku 4.16. Velikost předstihu je dále korigována podle teploty motoru a teploty nasávaného vzduchu.

Výčet funkcí řídicí jednotky při řízení sekvenčního vstřikování

- základní nastavení pomocí datového pole
- řízení startu motoru
- obohacování směsi při opakovaném startu, při startu teplého motoru a při zrychlení vozidla
- odpojování vstřiku paliva při deceleraci vozidla
- omezování maximálních otáček
- lambda-regulace (adaptační dílčí systém)



Obrázek 4.16 – Závislost základního předstihu na otáčkách motoru a poloze škrtící klapky

Funkce při řízení zapalování

- základní nastavení pomocí datového pole
- řízení úhlu pro uzavírání sacích a výfukových ventilů
- korekce běhu teplého motoru
- řízení startu
- stabilizace volnoběhu
- regulace klepání každého válce zvlášť (adaptační dílčí systém).

Funkce pro odvětrávání palivové nádrže

- řízeno datovým polem,
- korigováno pomocí lambda-regulace

Vlastní diagnostika

- sledování snímačů, čidel a nastavovacích členů
- ukládání závad do paměti, čtení uložených závad
- diagnóza akčních členů a výstup naměřených hodnot
- nouzová funkce.

Efekt automatického řízení vstřikování paliva a jeho zapálení systémem Motronic 3.8.2 spočívá ve

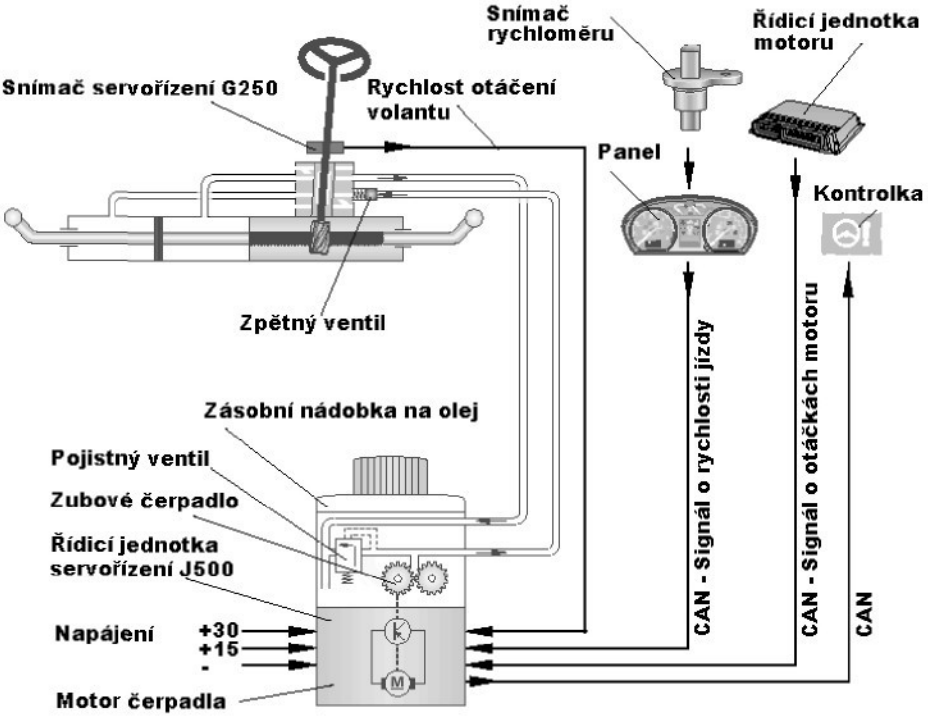
- snížení obsahu škodlivin ve výfukových plynech díky optimalizaci vstřikovaného množství paliva a okamžiku zapálení směsi a lambda-regulaci
- snížení spotřeby paliva
- neměnnosti charakteristiky zapalování v průběhu provozu
- zjednodušení údržby zapalovacího systému, která spočívá pouze ve výměně zapalovacích svíček po 60 000 km.

4.6. Elektrohydraulické servořízení

U starých automobilů se natočení volantu přenášelo na natočení kol přední nápravy automobilu prostřednictvím převodky řízení, která posouvala jedním ramenem lichoběžníku zajišťujícího správný úhel natočení kol při jízdě vozidla zatáčkou. K usnadnění řízení, zvláště automobilů s předním náhonem, byl zkonstruován posilovač řízení (Power Assisted Steering). V případě, kdy je použita hydraulika a elektronika, je systém označován jako EPHS (Electrically Powered Hydraulic Steering). Konstrukce posilovače musí umožnit řízení směru pohybu vozidla také při jeho poruše.

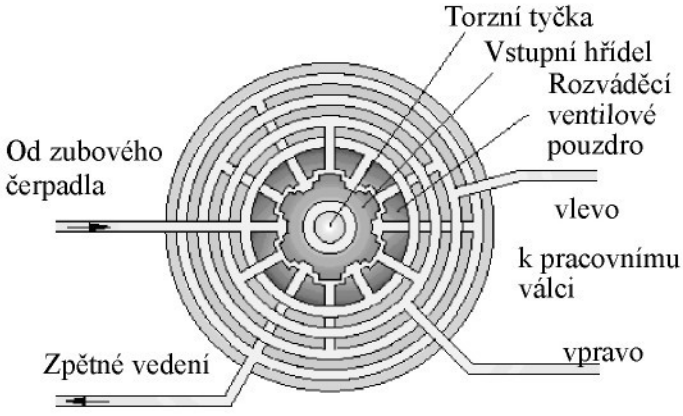
Nejnovější řešení posilovače řízení, jehož komplexní schéma je znázorněno na obrázku 4.17, pracuje se signálem rychlosti otáčení volantem a podle této rychlosti je upravován výkon elektrického motorku, který pohání zubové čerpadlo oleje nezávisle na otáčkách motoru. Výhodou tohoto řešení je to, že posilovací účinek je vysoký, i když otáčky motoru jsou jen volnoběžné. Čím rychleji se volantem otáčí, tím je k dispozici větší tlak oleje. Tato vazba slábne z důvodu bezpečnosti při zvyšování rychlosti vozidla, aby se znesnadnila rychlá změna směru jízdy při vysokých rychlostech.

Akční člen posilovače je přímočarý hydraulický motor, jehož pístnice je napojena na tyč s hřebenovým ozubením. Tlakový olej je do posilovače přiváděn přes rozváděcí ventilové pouzdro, které je znázorněno na obrázcích 4.18 a 4.19. Vstupní hřídel posilovače (od volantu) je s tímto rozváděcím pouzdrem a s pastorkem, který zabírá s hřebenovým ozubením, propojen přes torzní tyčku. V nehybné poloze volantu je tlak oleje na obou stranách pístu shodný, torzní tyčka a rozváděcí ventilové pouzdro jsou v neutrální poloze.

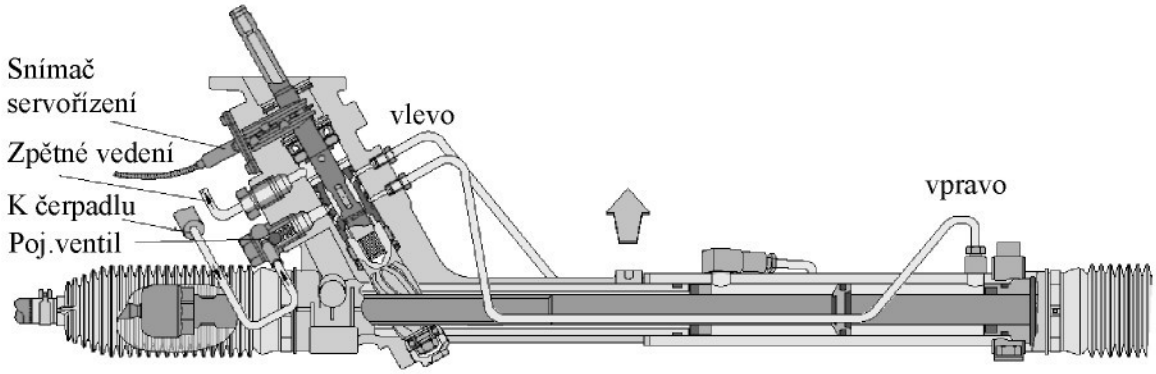


Obrázek 4.17 – Posilovač řízení – funkční schéma

Vychýlení volantu nebo vnější síla, působící na kola řízené nápravy, vyvolá natočení torzní tyčky úměrně velikosti krouticího momentu a také přerozdělení toku oleje mezi hranami rozváděcího ventilového pouzdra a vstupního hřídele tak, aby se vzájemné natočení torzní tyčky a rozváděcího ventilového pouzdra dostalo opět do neutrální polohy při daném natočení volantu. Hrany rozváděcího pouzdra a vstupního hřídele umožňují trvalou cirkulaci oleje.



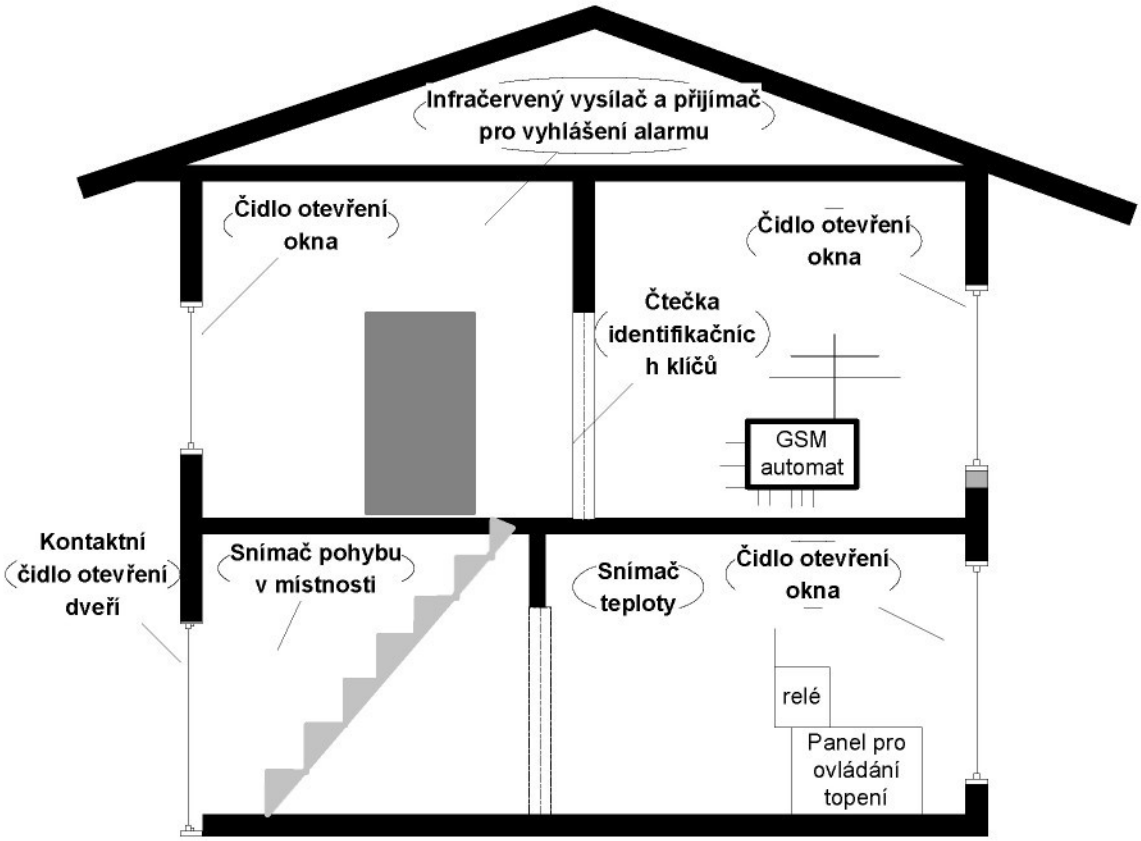
Obrázek 4.18 – Schéma hydraulického řídicího ventilu



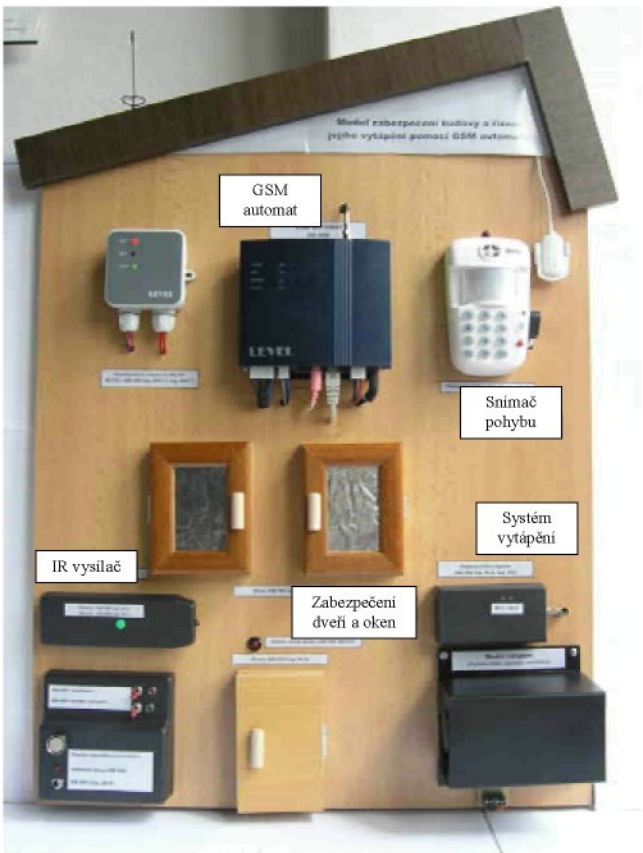
Obrázek 4.19 – Posilovač řízení – konstrukční provedení

4.7. Inteligentní dům

Pod tímto označením si jistě každý představí něco jiného. Opustíme-li však oblast literatury Scifi (Science fiction), označujeme jako „inteligentní dům“ stavbu vybavenou řízením distribuce všech energií, včetně vytápění, osvětlení, klimatizace, obvykle také doplněné o bezpečnostní systémy, případně také záznam příchodů a odchodů osob. Pro realizaci v rodinných domech se v současnosti velmi často doplňují systémy řízení také o možnost vzdáleného dohledu a ovládání využívající bezdrátové komunikační sítě, nejčastěji GSM (Global System for Mobile Communications, globální systém pro mobilní komunikaci). Jako vhodný nástroj může použít kompaktní GSM automat [2], který sdružuje komunikační systém s nástroji logického automatu včetně možnosti měření technologických veličin generování akčních zásahů k řízení těchto procesů.



Obrázek 4.20 – Rozmístění prvků zabezpečení „inteligentního domu“



Ke zjišťování stavu domu, nastavování požadovaných hodnot technologických veličin a naopak pro hlášení o narušení objektu apod. se využívají specializované zprávy SMS (Short Message Service, krátká textová zpráva), s jejich pomocí je například možné dálkově nastavit požadovanou teplotu vytápění.

Na obrázku 4.20 je zobrazen návrh rozmístění jednotlivých prvků systému v rodinném domě [1]. Sdružuje tyto hlavní části:

- systém řízení vytápění s možností dálkového nastavení požadované teploty,
- zabezpečení objektu proti vniknutí okny a dveřmi, doplněný o snímač pohybu ve vstupní hale a čtečku identifikačních klíčů pro zapínání a vypínání zabezpečení,
- infračervený vysílač a přijímač

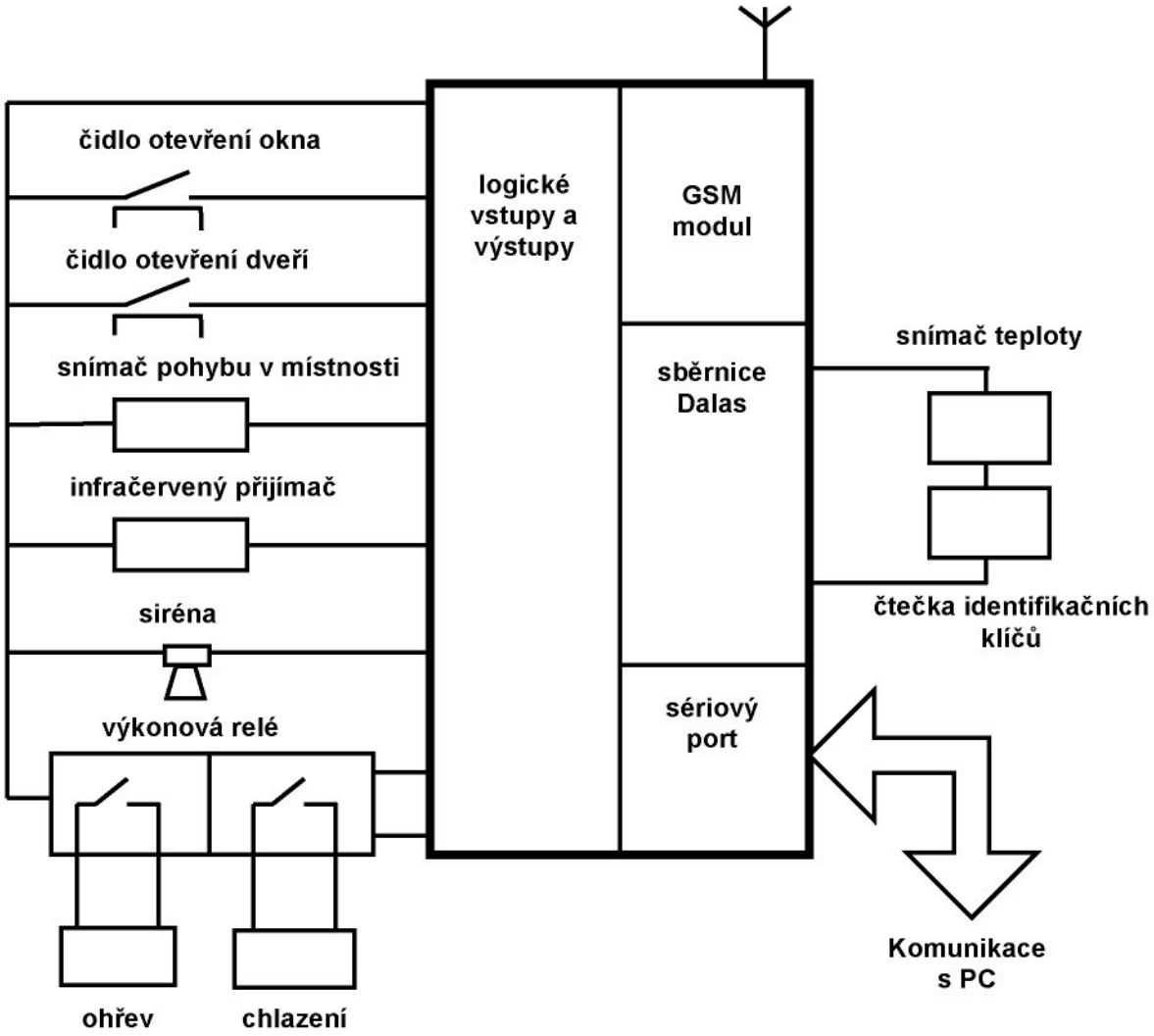
Obrázek 4.21 – Model „inteligentního domu“

pro vyhlášení alarmu, například pro použití handicapovanou osobou žijící v domě k přivolání pomoci.

Funkčnost navrženého systému je možné si ověřit na modelu v laboratořích katedry ATŘ, který obsahuje všechny popsané funkční prvky, viz obrázek 4.21.

Schéma zapojení jednotlivých prvků, napojení jednotlivých snímačů a akčních členů systému je ukázáno na obrázku 4.22.

K programování GSM automatu se používá moderní přístup založený na obsluze událostí jako je například změna hodnoty logického vstupu, příchod řídicí SMS, ale také činnosti odvozené od interního časovače, prováděné v nastavených okamžicích vzorkování. Například reakcí automatu na příchod signálu o otevření okna může být zahájení telefonního volání na určené číslo nebo odeslání SMS určeného textu, aktivace některé skupiny činností nebo aktivace některého výstupu automatu, například spuštění sirény. Skládáním a zřetězením těchto jednoduchých akcí je pak možno vytvořit celý řídicí algoritmus pro zajištění veškerých služeb systému.



Obrázek 4.22 – Schéma zapojení prvků systému

Otázky 4.1.

1. Co znamená kaskádová regulace?
2. Co je charakteristická vlastnost regulované soustavy v případě regulace zásob?
3. Kdy se používá smyčková regulace?
4. Co je akční člen posilovače řízení?
5. Lze při použití GSM automatu dálkově nastavit např. požadovanou teplotu vytápění?

Další zdroje

[1] BEMBENEK, P. 2004. *Řešení bezpečnostního systému pomocí GSM automatu.* Ostrava : VŠB-TU Ostrava, 2004. 78 s. Diplomová práce, vedoucí: Farana, R.

[2] LEVEL s.r.o. 2006 *Firemní informační systém*. [on-line]. HTML formát [cit. 2006-08-31]. Dostupný z www: <URL: http://www.levelna.cz>.

[3] Školící manuály firmy Škoda Auto a.s.

[4] VOBECKÝ, J. & ZÁHLAVA, V. *Elektronika. Součástky a obvody, principy a příklady*. (2 rozšířené vydání). Praha : Grada Publishing, 2001, 192 str. ISBN 80-7169-884-9.

Klíč k řešení

Odpovědi na otázky.

O 1.2 Odpověď na otázky 1.2.

O 1.1. 1. Co je to logické řízení?

Dvoustavové řízení, zapnuto – vypnuto.

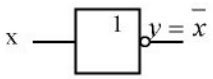
2. Jak je definována logická funkce?

Jako přiřazení dvouhodnotové proměnné y ke kombinacím hodnot nezávislých logických proměnných $x_1, x_2, \ldots x_n$.

3. Co je to NAND a NOR?

Negovaný součin a negovaný součet.

4. Jak vypadá značka negace NOT?



5. Na co se používá Boolova algebra?

K minimalizaci logické funkce.

6. Jak vypadají De Morganovy zákony?

$\overline{a+b} = \overline{a} \cdot \overline{b} \quad \overline{a \cdot b} = \overline{a} + \overline{b}$

7. Můžeme v Karanughově mapě sdružovat 3 políčka?

Ne.

8. Může obsahovat Karnaughova mapa nedefinované hodnoty?

Ano.

9. Muže být logická funkce zadána slovně?

Ano.

10. Pomocí jakých vyjádření můžeme vypsat logickou funkci z pravdivostní tabulky?

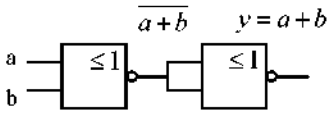
úplná disjunktní normální forma a úplná konjunktivní normální forma

O1.2 1. Jak se zapojuje logický součin pomocí kontaktních prvků?

2. Dá se kombinační logický obvod zapojit pomocí dvoustopého hradla NAND?

Ano.

3. Jak se zapojuje logický součet pomocí hradel NOR?



O1.3 1. Co je to sekvenční logická funkce?

Výstupní proměnné závisí nejen na momentální kombinaci vstupních proměnných, ale také na dosaženém stavu, který bývá určen kombinací vnitřních logických proměnných.

2. Co obsahují sekvenční logické obvody?

Paměťové prvky.

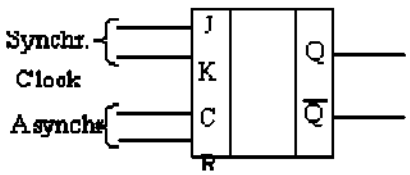
O1.4 1. Kolik mají stavů klopné obvody?

Dva stavy.

2. Kdy se klopí výstup KO typu D?

Jen když je na vstup D přivedena logická jednotka.

3. Jakého obvodu je značka na obrázku?



Klopný obvod J-K.

O1.5 1. Z čeho se skládá čítač?

Z několika klopných obvodů.

O1.6 1. Jak se označuje počáteční krok v sekvenčním diagramu?

Dvojitým orámováním bloku.

2. Jak se liší způsoby větvení AND a OR?

Při větvení AND přechází aktivní krok do všech větví diagramu, kdežto u větvení OR se vykonává posloupnost jen v jedné větvi.

O1.7 1. Co patří mezi stavební prvky logického řízení?

Relé, logické obvody, programovatelné logické automaty.

2. Co jsou to programovatelné logické automaty?

Jsou to programovatelné řídicí systémy specializované na úlohy převážně logického typu.

3. Co je to PLC v modulárním provedení?

Programovatelný logický automat, který se dá spojovat do větších a složitějších celků.

4. Jaké vstupy a výstupy obsahují PLC?

Binární a analogové.

5. Jaké jazyky používáme k programování PLC?

Jazyk mnemokódů, jazyk kontaktních schémat, jazyk logických schémat, jazyk strukturovaného textu.

O2.1 1. Co je to prvek?

Prvek je již dále nedělitelná část systému při dané rozlišovací schopnosti.

2. Co je to vazba?

Přenáší informace z prvku na prvek nebo mezi systémem a okolím systému.

3. Co je to nestacionární systém?

Systém, jehož tvar matematického popisu se v čase mění.

4. Co je to dynamický systém?

Systém, jehož chování se v čase mění.

O2.2 1. Jaká funkce je obraz funkce u Laplaceovy transformace?

Racionálně lomená funkce.

2. Jak zní vlastnost Laplaceovy transformace pro n-tou derivaci?

$$L\left\{f^{(n)}(t)\right\} = s^n F(s) - s^{n-1} f(0) - s^{n-2} f'(0) - \ldots - f^{n-1}(0)$$

3. Jaké typy rovnic se dají pomocí Laplaceovy transformace výhodně řešit?

Lineární diferenciální rovnice

O2.3 1. Definujte statickou charakteristiku.

Závislost výstupní veličiny na vstupní v ustáleném stavu (stavu, kdy hodnoty vstupu i výstupu jsou konstantní).

2. Kterými charakteristikami lze popsat dynamické vlastnosti systémů?

Přechová, impulsní a kmitočtová charakteristika.

3. Jak je definován přenos?

Podíl L obrazu výstupní veličiny k L obrazu vstupní veličiny při nulových počátečních podmínkách.

4. Co je to Diracův impuls?

$\delta(t)$ je vstupní signál, který se jeví jako nekonečně krátký impuls o nekonečně velké amplitudě, jehož plocha je rovna jedné a Laplaceův obraz se rovněž roven jedné.

5. Čím se nahradí komplexní proměnná s ve vzorci pro kmitočtový přenos?

$J\omega$

O2.4 1. V jakých hodnotách se počítá s reálným kmitočtem?

Interval $(0,\infty)$

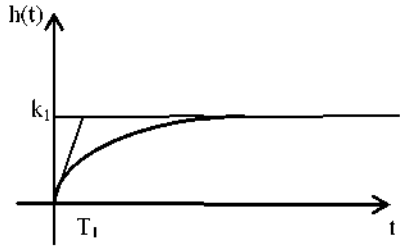
2. Kolik dB odpovídá zesílení 10?

20 dB

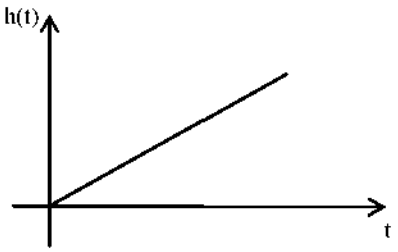
3. Jak se vykreslí LKCH pro systémy v sériovém zapojení?

Výsledné LKCH jsou dány součtem dílčích LKCH.

O.2.5 1. Jak vypadá přechodová charakteristika setrvačného článku s časovou konstantou T_1?



2. Nakreslete přechodovou charakteristiku integračního článku.



3. Uveďte příklad ideálního derivačního systému.

Neexistuje

4. Jaká je ustálená hodnota přechodové charakteristiky pro derivační systémy?

Nulová hodnota.

5. V jaké hodnotě začíná amplitudo-fázová kmitočtová charakteristika proporcionálního systému?

Na kladné reálné poloose.

6. Lze vytknout proměnnou s v čitateli přenosu integračních systémů?

Nelze.

O2.6 1. Z čeho se skládá regulační obvod?

Regulátor a regulovaná soustava, popř. měřicí člen.

2. Co je to regulovaná veličina?

Veličina, jejíž hodnota je výstupem z regulované soustavy a jež se regulací udržuje na požadované hodnotě.

3. Co je to akční veličina?

Je výstupní veličinou regulátoru, která je vstupní veličinou regulované soustavy.

4. Jak se spočítá náhradní přenos paralelně zapojených soustav?

$$G(s) = \frac{Y(s)}{U(s)} = G_1(s) + G_2(s)$$

6. Jak se vypočítá odchylkový přenos řízení?

$$G_{ew}(s) = \frac{E(s)}{W(s)} = \frac{1}{1 + G_O(s)}$$

5. Co obsahují číslicové regulační obvody navíc než analogové?

A/Č a Č/A převodníky.

O2.7 1. Co to je stabilita?

Vlastnost systémů definovaná jako schopnost dosáhnout rovnovážného stavu po odeznění vstupních sil, způsobujících nerovnováhu.

2. S čím je svázáno jméno Nyquist?

S kmitočtovým kritériem stability.

3. Která kritéria stability patří mezi algebraická?

Hurwitzovo, Rothovo-Shurovo.

4. Jak určíme charakteristický mnohočlen regulačního obvodu popsaného přenosem otevřené smyčky?

Sečteme jmenovatel ačitatel přenosu.

O2.8 1. Existují regulátory typu PD?

Ano.

2. Co je to pásmo proporcionality?

$$PP = \frac{1}{k_p} 100\%$$

3. Jaké jsou stavitelné parametry PID regulátoru?

Zesílení kp, integrační časová konstanta T_I a derivační časová konstanta T_D.

4. Napište zkratky pro regulátory odvozené od typu PID?

P, I, PI, PD, PIDi

5. Existuje regulátor typu D?

Ne.

6. Ustálí se regulovaná veličina v regulačním obvodu s nespojitým regulátorem?

Neustálí.

7. Jak se realizují analogové regulátory?

Pomocí operačních zesilovačů.

O2.9

1. Vznikne trvalá regulační odchylka v regulačním obvodu bez integrátoru a se skokem polohy žádané hodnoty?

Ano.

2. K čemu slouží integrální kritéria?

K vyčíslení regulační plochy.

3. Jak se určí doba regulace?

Doba, kdy se regulovaná veličina dostane do okolí 2δ žádané hodnoty a už ho neopustí

O2.10

1. Jak se jmenuje metoda seřízení regulátoru, která byla odvozena na katedře ATŘ - 352?

Metoda požadovaného modelu.

2. Jaké označení má experimentální postup seřizování regulačních obvodů?

Metody Ziegler-Nichols.

3. Co je to kritické zesílení?

Zesílení regulátoru, které přivede regulační obvod na kmitavou mez stability.

4. Jak vypočteme vzorkovací periodu?

$$T = \frac{2\pi}{\omega}$$

O3.1

1. Z čeho se skládá analogový regulační obvod?

Regulátor, regulovaná soustava, měřící člen.

2. Čím se liší analogový regulační obvod od diskrétního?

Navíc musí obsahovat A/Č a Č/A převodníky.

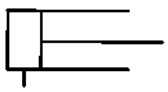
O3.2

1. K čemu se používají rozvaděče?

Ventily umožňující řízení směru průtoku, případně jeho přehrazení.

2. Jaká je značka pro zásobník stlačeného vzduchu?

3. Jaká je značka pro jednočinný pneumatický válec?



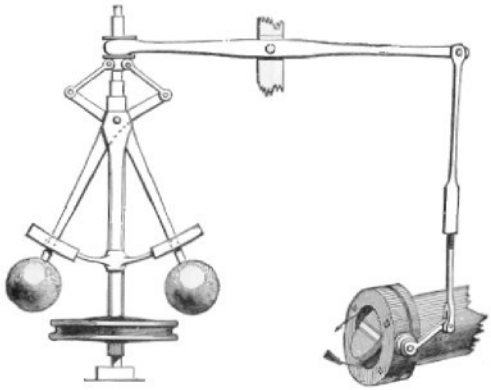
O3.3 1. Jaké zesílení má operační zesilovač? Velké, střední, malé.

Velké.

2. Kde se používá operační zesilovač?

K realizaci analogových PID regulátorů.

3. Jak se jmenuje regulátor na obrázku?



Wattův rozběžník.

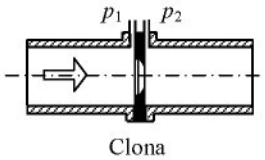
O3.4 1. Základní princip pneumatických regulátorů? Tryska + ?

Tryska+klapka

2. Používají se operační zesilovače u pneumatických regulátorů?

Ano.

O3.5 1. Kde se umístí na potrubí clona pro měření průtoku?



Clona

2. Lze umístit na potrubí clona pro měření průtoku do kolena?

Ne.

3. Co se používá nejčastěji pro měření do teploty do 300 °C?

Platinové odporové teploměry Pt 100

4. Při regulaci sypkých hmot a materiálů se používá odrazové desky. Jaký je její princip?

Principem měření je úměra síly vyvozená změnou směru toku materiálu.

O3.6 1. K čemu je kuličkový šroub?

Kuličkové šrouby slouží k převedení rotačního pohybu na přímočarý.

2. K čemu slouží harmonická převodovka?

Jsou zvlášť vhodné pro aplikace vyžadující velké převodové poměry mezi motorem a poháněným zařízením prostřednictvím jednoho převodového stupně.

3. K čemu slouží žaluzie?

K regulaci průtoku vzduchu a plynu.

4. Jaká je podmínka správné funkce klapky?

Vhodná tlaková ztráta při otevřené poloze.

O.3.7 1. Které motory se zapojují do hvězdy a trojúhelníka?

Střídavé třífázové asynchronní motory, stejnosměrné motory, krokové motory.

2. Co je to asynchronní motor?

Má statorové vinutí pro třífázové napájení složené ze tří statorových cívek.

3. K čemu se používá střídač nebo měnič frekvence?

Měnitelnost otáček u asynchronního třífázového motoru.

4. Co je to krokový motor a kde se použije?

Krokové motory jsou motory, u kterých můžeme měnit jeho natočení. Nejsou jen pro otáčivý pohyb, ale také pro lineární posunutí.

5. Jak se liší řízený usměrňovač od střídače?

Řízený usměrňovač mění střídavý proud na stejnosměrný a střídač mění stejnosměrný proud na střídavý.

6. Jakým zdrojem elektrického proudu se napájí stejnosměrný elektrický motor?

Řízeným usměrňovačem, který převádí střídavé napětí na stejnosměrné.

7. Co je to synchronní motor?

Otáčky a frekvence napájecího napětí jsou synchronní nezávisle na zatížení motoru.

8. Které elektrické motory se otáčejí stejně rychle jako točivé magnetické pole, které vybudí statorové cívky?

Synchronní.

O3.8 1. Co je charakteristika čerpadla?

Závislost průtoku pracovní kapaliny čerpadlem na jeho výstupním tlaku.

2. Jak se liší uzavřený a otevřený hydraulické obvody?

Otevřený obvod – olej se shromažďuje v nádobě, odkud se čerpá zpět. Uzavřený – mají uzavřený oběh.

3. Mohou pneumatické pohony pracovat v prostředí s nebezpečím výbuchu?

Ano.

4. Které pohony přenášejí vyšší výkony, pneumatické nebo hydraulické?

Hydraulické.

O3.9 1. Jaký je nejčastěji používaný piezoelektrický materiál?

Keramika PZT.

2. Jaké veličiny tvoří charakteristiku aktuátoru?

Závislost síly a protažení s parametrem elektrické napětí.

3. Uveďte tři příklady aplikace piezoaktuátorů.

Reproduktory, pohybové ústrojí zrcadel laserů, vstřikování paliva, inkoustové tiskárny.

O3.10 1. Co je to IP adresa na internetu?

Skládá se ze čtyř 8 bitových čísel. Musí být jedinečná v celé internetové síti.

2. Co je to LAN?

Počítačová síť, hlavně v budovách.

3. Vyjmenujte vrstvy komunikačního protokolu.

Fyzická, spojová, síťová, transportní, relační, prezentační, aplikační.

4. K čemu se používá rozbočovač-hub?

Zajišťuje větvení signálů.

O3.11 1. Co je na nejnižším stupni hierarchické pyramidy řízení?

Proces, stroj, technika, energie, materiál, technologie.

2. Co je v systému SCADA/HMI alarm?

Typ zprávy, kdy dojde k nějakému překročení hodnot.

3. Co jsou to historické trendy ve SCADA/HMI?

Pomáhají operátorovi v řízení procesu, hlavně při monitorování průběhu minulých a žádaných veličin.

O4.1 1. Co znamená kaskádová regulace?

Regulace, kdy máme více regulačních smyček.

2. Co je charakteristická vlastnost regulované soustavy v případě regulace zásob?

Dopravní zpoždění.

3. Kdy se používá smyčková regulace?

Při synchronizaci otáček dvou různých zařízení.

4. Co je akční člen posilovače řízení?

Přímočarý hydraulický motor.

5. Lze při použití GSM automatu dálkově nastavit např. požadovanou teplotu vytápění?

Ano, pomocí SMS zpráv.

Rejstřík

Akční členy, 205

Algebra blokových schémat, 125, 129

amplitudo fázová kmitočtová charakteris, 92

Booleova algebra, 9

Dávkování sypkých hmo, 211

derivační systém, 110

Derivační systém, 116

dopravní zpoždění, 121

Elektrické motory, 216

elektrické pohony, 212

hradla, 23

Hurwitzovo kriterium stability, 139

hydraulické systémy, 171, 232

Impulsní funkce, 90

integrační systém, 110

Integrační systém, 118

IP-adresa, 242

Karnaughova mapa, 18

Karnaughova mapa (K-mapa), 10

Kaskádová regulace, 259

klopný obvod, 44

Klopný obvod, 41

Kmitočtový přenos, 92

Kontaktní schéma, 22

Krokové motory, 224

Kvalita regulace, 152

Laplaceova transformace, 78

liniové schéma, 22

logaritmické kmitočtové charakteristiky, 93, 99

Logaritmické kmitočtové charakteristiky, 108

Logická funkce, 9

Logické řízení, 8

Logický součet, 11

Logický součin, 12

Matematické modely, 95, 96

Měniče frekvence, 229

NAND, 12

Negace, 11

NOR, 12

Nyquistovo kriterium stability, 144

Obrazový přenos, 87

Piezoelektrické aktuátory, 237

pneumatické systémy, 171

Pneumatické systémy, 232

Počítačové sítě, 240

Polovodičová dioda, 213

programovatelný logický automat, 53

Programovatelný logický automat, 26

proporcionální systém, 110

proporcionální systémy, 113

Přechodová funkce, 88

Převody pohybů, 206

Regulace průtoku sypkých hmot, 260

Regulace teploty, 258

Regulace vlhkosti, 262

Regulace zásoby, 261

Regulační obvod, 73

Regulační obvod, 74

Regulační orgány, 208

regulátor, 148; analogový, 148, 175; číslicový, 149; nespojitý, 150; pneumatický, 180

Routhovo-Schurovo kriterium stability, 141

Řešení diferenciálních rovnic, 82

Řízení polohy, 263

Sekvenční funkční diagram, 47

Sekvenční logický obvod, 38

snímače, 188; konstrukce, 189

Snímače: Dotykové sondy, 203; hladiny, 200; konstrukce, 189; polohy, 200; průtoku plynů a kapalin, 194; síly a krouticího momentu, 197; teploty, 192; tlaku, 199

Stabilita, 137

Statická charakteristika, 86

Stejnosměrný motor, 217

Střídavý asynchronní motor, 220

Synchronní střídavé motory, 224

Syntéza, 155

Syntéza sekvenčního logického řízení, 38

Systém, 70

Tranzistor, 212

Triak, 215

Třídění signálů, 71

Třídění systému, 73

Tyristor, 214

Vizualizace, 246; Alarmy, 250; Trendy, 250

Výlučný (exkluzivní) součet, 12

Zadání logické funkce, 10

Zákony Booleovy algebry, 14

Seznam animací

Obrázek 1.48 – Animace popisující základní logické operace 66
Obrázek 1.49 – Animace popisující převod mezi tabulkou hodnot a K-mapou 66
Obrázek 1.50 – Animace popisující převod mezi tabulkou hodnot a K-mapou 67
Obrázek 1.51 – Animace popisující základní logické obvody 67
Obrázek 2.71 – Animace chování proporcionálních systémů 123
Obrázek 2.72 – Animace chování integračních a derivačních systémů 124
Obrázek 2.73 – Animace chování systému hmota na tlumiči a pružině 124
Obrázek 2.74 – Animace chování kyvadla 124
Obrázek 2.83 – Animace kontroly stability pomocí Hurwitzova kritéria stability 141
Obrázek 2.84 – Animace kontroly stability pomocí Routhova-Schurova kritéria stability 144
Obrázek 2.95 – Animace seřízení RO pomocí metody Ziegler-Nichols 160
Obrázek 3.127 – Animace stejnosměrného motoru 255
Obrázek 3.128 – Animace krokových motorů 255
Obrázek 3.129 – Animace synchronního motoru 256
Obrázek 3.130 – Animace funkce Wattova regulátoru 256
Obrázek 3.131 – Animace pneumatického (vlevo) a hydraulického (vpravo) polohování 256