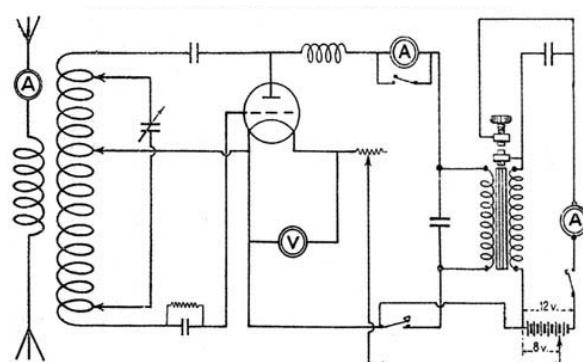


Università degli Studi di Firenze
Dipartimento di Fisica

Marcello Carlà

Appunti di Elettronica per Fisici

A.A. 2020-2021



Copyright © 2005-2023 Marcello Carlà

Ogni riproduzione completa o parziale di questo documento per uso privato o collettivo è assolutamente consentita per qualunque numero di copie.

Ogni copia dovrà riportare il testo di questo Copyright o, per copie di una piccola parte del documento, almeno l'indirizzo dove è disponibile la versione completa: <http://studenti.fisica.unifi.it/~carla/appunti> oppure <http://bambatron.duckdns.org>.

Quest'opera è stata rilasciata sotto la licenza "Creative Commons Attribution-Non-Commercial-ShareAlike 2.5". Per leggere il testo completo della licenza visita il sito web <http://creativecommons.org/licenses/by-nc-sa/2.5/> o invia una richiesta a Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

L'autore sarà grato a chi vorrà inviare segnalazioni di errori e/o suggerimenti.

e-mail: carla@fi.infn.it

In copertina: schema elettrico del trasmettitore radio utilizzato nel maggio-giugno 1928 dal marconista Biagi per inviare i segnali di richiesta di soccorso dalla banchisa polare.

Indice

1 Circuiti	1-1
1.1 Segnali elettrici	1-1
1.2 Leggi di Kirchhoff	1-3
1.2.1 I Legge di Kirchhoff - Legge dei nodi	1-3
1.2.2 II Legge di Kirchhoff - Legge delle maglie	1-3
1.3 Reti lineari - principio di sovrapposizione	1-4
1.4 Equazioni delle reti lineari	1-5
1.4.1 Legge di Ohm	1-5
1.4.2 Legge di Ohm complessa	1-5
1.4.3 Teorema di Thévenin	1-6
1.4.4 Teorema di Norton	1-7
1.5 Massimo trasferimento di potenza	1-8
1.6 Circuiti RC	1-9
1.6.1 Filtri passa-alto e passa-basso - Il <i>decibel</i>	1-9
1.6.2 Risposta all'onda quadra	1-11
1.7 Circuiti risonanti LRC	1-12
1.8 Filtri LRC	1-14
1.8.1 Circuito risonante parallelo - circuiti duali	1-17
1.8.2 Circuiti risonanti reali	1-18
2 Linee di trasmissione	2-1
2.1 Circuiti a costanti concentrate	2-1
2.2 Linee di trasmissione	2-1
2.3 Circuito equivalente di una linea di trasmissione	2-1
2.3.1 Cavo coassiale	2-2
2.4 Correnti e tensioni lungo la linea	2-4
2.5 Linea senza perdite	2-5
2.6 Lunghezza d'onda, lunghezza elettrica	2-6
2.7 Linea con perdite. Condizioni di Heaviside	2-6
2.8 Impedenza caratteristica della linea	2-7
2.9 Linea di trasmissione chiusa sulla sua resistenza caratteristica	2-8
2.10 Coefficiente di riflessione.	2-9
2.11 Onde stazionarie.	2-9
2.11.1 Trasformazione di impedenza.	2-10
2.11.2 Linea di lunghezza variabile	2-12
2.11.3 Tensioni alle estremità della linea	2-13

2.12 Segnali a banda larga e a banda stretta	2-14
2.13 Riflessione di un segnale a banda larga	2-15
2.14 Resistenza interna del generatore	2-17
2.15 Riflessione di un segnale a gradino	2-18
2.16 Riflessioni multiple	2-19
2.16.1 Riflessioni multiple in un sistema con perdite	2-20
2.17 Velocità di fase e di gruppo	2-23
3 Semiconduttori	3-1
3.1 Resistività	3-1
3.2 Germanio e Silicio	3-2
3.3 Elettroni e Lacune	3-2
3.4 Modello a bande di energia	3-4
3.5 Semiconduttore estrinseco	3-5
3.6 Legge di Azione di Massa	3-6
3.7 Corrente di diffusione	3-6
4 Diodi	4-1
4.1 Giunzione <i>P-N</i>	4-1
4.2 Regione di Transizione	4-2
4.3 Potenziale di giunzione	4-2
4.4 Diodo a giunzione	4-4
4.5 Polarizzazione diretta ed inversa	4-4
4.6 Correnti in un diodo polarizzato	4-6
4.6.1 Polarizzazione diretta	4-6
4.6.2 Polarizzazione inversa	4-8
4.7 Equazione della giunzione	4-9
4.7.1 Diodo in conduzione	4-10
4.7.2 Diodo in interdizione	4-10
4.8 Contatti ohmici	4-12
4.9 Coefficiente di temperatura del diodo a giunzione	4-12
4.10 Una applicazione dei diodi: il raddrizzatore	4-13
4.11 Tensione inversa di rottura (<i>Reverse Breakdown Voltage</i>)	4-14
4.12 Effetti termici nel diodo	4-15
4.13 Circuiti non lineari	4-16
4.14 Circuito <i>linearizzato</i>	4-17
4.15 Conduttanza dinamica del diodo	4-19
4.15.1 Diodo direttamente polarizzato	4-19
4.15.2 Diodo non polarizzato	4-20
4.15.3 Diodo contropolarizzato	4-20
4.16 Diodi stabilizzatori di tensione	4-20
4.17 Interruttore a diodi per segnali analogici	4-21
4.18 Capacità della giunzione	4-23
4.18.1 Diodo contropolarizzato - Capacità di transizione	4-23
4.18.2 Diodo in conduzione - Capacità di diffusione	4-25
4.18.3 Confronto tra capacità di transizione e capacità di diffusione	4-27
4.19 Spessore della regione di transizione	4-27

4.20	Tempi di commutazione di un diodo	4-28
4.21	Dispositivi Optoelettronici	4-30
4.21.1	Fotodiodi	4-30
4.21.2	Relazione tensione-corrente in un fotodiodo	4-31
4.21.3	Rendimento dei generatori fotovoltaici	4-32
4.21.4	Fotorivelatori	4-33
4.21.5	Diodi <i>led</i> (<i>light emitting diode</i> - diodi emettitori di luce)	4-34
4.21.6	Diodi <i>laser</i>	4-34
5	Misure su diodi	5-1
5.1	Tempi di commutazione	5-1
5.2	Capacità di transizione	5-4
5.2.1	Fit dei dati sperimentali con il programma <i>Gnuplot</i>	5-6
6	Reti a due porte	6-1
6.1	Circuiti attivi	6-1
6.2	Amplificatore ideale	6-1
6.3	Variabili indipendenti e dipendenti	6-3
6.4	Modelli lineari per i dispositivi a due porte	6-3
6.4.1	Trasformazioni di parametri	6-5
6.5	Circuito amplificatore	6-6
6.6	Amplificatori ideali	6-9
6.7	Reti passive	6-10
7	Il transistor <i>bjt</i>	7-1
7.1	Equazioni di Ebers-Moll	7-1
7.2	Regioni operative del <i>bjt</i>	7-3
7.2.1	Regione di interdizione	7-3
7.2.2	Regione attiva	7-3
7.2.3	Regione di saturazione	7-4
7.3	Configurazioni circuitali	7-4
7.3.1	base comune (<i>common base</i>)	7-4
7.3.2	emettitore comune (<i>common emitter</i>)	7-5
7.3.3	collettore comune (<i>common collector</i>)	7-5
7.4	Curve caratteristiche	7-5
7.5	Effetto Early	7-6
7.6	Amplificatore ad emettitore comune	7-7
7.6.1	Modello lineare a emettitore comune	7-8
7.6.2	Valori dei parametri <i>g</i> ed <i>h</i> in corrente continua e bassa frequenza	7-9
7.6.3	Amplificazione di tensione e di corrente	7-11
7.7	Amplificatore a collettore comune e base comune	7-13
7.7.1	<i>Emitter follower</i>	7-13
7.7.2	Reazione di emettitore	7-14
7.7.3	Amplificatore a base comune	7-15
7.7.4	Parametri a base comune e collettore comune	7-16
7.8	La polarizzazione del <i>bjt</i>	7-17
7.8.1	Polarizzazione con singola resistenza di base	7-18

7.8.2	Polarizzazione a quattro resistenze	7-19
7.8.3	Condensatori di accoppiamento e disaccoppiamento	7-20
7.9	Amplificatore differenziale	7-22
7.9.1	Amplificazione di modo comune	7-23
7.9.2	Amplificazione di modo differenziale	7-24
7.9.3	Un esempio di amplificatore differenziale	7-25
7.9.4	Reiezione di modo comune	7-26
7.10	Circuito <i>current mirror</i>	7-27
7.11	Circuito <i>Darlington</i>	7-28
7.12	Circuiti a simmetria complementare	7-29
8	Amplificatori a transistor e misure	8-1
8.1	Amplificazione	8-1
8.2	Banda passante	8-2
8.3	Risposta all'onda quadra in un amplificatore a larga banda	8-3
8.4	Frequenze di taglio multiple	8-6
8.5	Amplificatore con transistor bjt	8-8
8.6	Misure sui circuiti	8-12
8.6.1	Sonda compensata	8-12
8.6.2	Partitore compensato	8-12
8.7	Misura dell'amplificazione	8-15
8.8	Impedenza di ingresso e di uscita	8-16
8.8.1	Impedenza di uscita	8-16
8.8.2	Impedenza di ingresso	8-17
8.9	Distorsione	8-17
9	I transistor <i>mosfet</i> e <i>jfet</i>	9-1
9.1	Struttura del transistor <i>mosfet</i>	9-1
9.2	Canale di conduzione	9-2
9.3	Conduttanza del canale	9-5
9.4	Modulazione del canale	9-6
9.5	<i>Pinch-off</i>	9-7
9.6	Curve caratteristiche per il transistor <i>nmos</i>	9-8
9.7	Il transistor <i>pmos - mosfet complementari</i>	9-9
9.8	Transistor <i>mosfet</i> di tipo <i>enhancement</i> e <i>depletion</i>	9-10
9.9	Struttura del transistor <i>jfet</i>	9-11
9.10	Modulazione della corrente di canale nel transistor <i>jfet</i>	9-12
9.11	Varietà di transistor ad effetto di campo	9-13
9.12	Circuiti con transistor ad effetto di campo	9-13
9.12.1	Amplificatore <i>common source</i>	9-13
9.12.2	Modello per piccoli segnali per i transistor ad effetto di campo	9-14
10	Elettronica digitale	10-1
10.1	Algebra di Boole	10-3
10.1.1	Operazioni logiche elementari	10-3
10.1.2	Circuiti logici elementari (<i>porte</i>)	10-4
10.1.3	Le leggi di De Morgan	10-4

10.1.4 <i>NAND e NOR</i>	10-6
10.2 Il transistor come interruttore	10-6
10.3 Le famiglie logiche	10-7
10.3.1 Famiglia <i>CMOS</i>	10-10
10.3.2 Porta <i>NAND e NOR</i> in tecnologia CMOS	10-11
10.3.3 Famiglia <i>TTL</i>	10-12
10.4 Le commutazioni nei circuiti digitali	10-15
10.5 La serie “74”	10-16
10.6 Logica combinatoria	10-18
10.6.1 Exclusive- <i>OR</i> (<i>XOR</i>)	10-18
10.6.2 Semiaddizionatore (<i>Half adder</i>)	10-19
10.6.3 Addizionatore (<i>Full adder</i>)	10-20
10.6.4 Addizionatore a 4 bit	10-20
10.6.5 Convertitori di codice	10-21
10.7 Tempi di commutazione	10-21
10.8 Logica sequenziale	10-22
10.9 I <i>flip-flop</i>	10-23
10.9.1 <i>Set-Reset (SR)</i>	10-23
10.9.2 <i>Clocked-SR</i>	10-25
10.9.3 <i>D-type</i>	10-25
10.10 I flip-flop di tipo <i>Master-Slave</i>	10-26
10.10.1 <i>Shift register</i> (Registro a scorrimento)	10-27
10.10.2 Flip-flop <i>JK</i>	10-28
10.10.3 Contatore binario	10-29
10.10.4 Contatore sincrono	10-30
10.11 Contatore modulo N	10-31
10.12 Orologio digitale	10-32
10.13 Elaboratore digitale	10-33

11 Amplificazione e Reazione	11-1
11.1 Amplificazione	11-1
11.2 Anello di reazione (<i>feedback loop</i>)	11-2
11.3 Guadagno ad anello aperto e ad anello chiuso	11-3
11.4 Reazione positiva	11-4
11.5 Reazione negativa	11-4
11.6 L’Amplificatore Operazionale	11-5
11.6.1 L’amplificatore operazionale nei circuiti	11-6
11.6.2 Amplificatori Operazionali reali	11-6
11.6.3 Altri limiti degli amplificatori operazionali reali	11-9
11.7 Amplificatore di tensione <i>non invertente</i>	11-10
11.7.1 <i>Voltage follower</i>	11-11
11.8 Effetti della reazione negativa	11-11
11.8.1 Stabilizzazione del guadagno	11-12
11.8.2 Riduzione della distorsione	11-12
11.8.3 Aumento della banda passante	11-12
11.8.4 Rete di reazione non lineare	11-14
11.8.5 Effetto della reazione negativa sul rumore	11-15

11.9 Resistenze di ingresso e di uscita	11-17
11.9.1 Resistenza di uscita	11-19
11.9.2 Effetto di r_i ed r_o su G	11-20
11.10 Amplificatore <i>invertente</i>	11-21
11.10.1 Resistenza di uscita	11-22
11.10.2 Resistenza di ingresso	11-22
11.10.3 Teorema di Miller	11-23
12 Applicazioni dell'amplificatore operazionale	12-1
12.1 Convertitore <i>corrente-tensione</i>	12-1
12.1.1 Una applicazione del convertitore corrente-tensione: amplificatore per fotorivelatori	12-2
12.2 Circuito sommatore - Nodo di somma	12-3
12.3 Amplificatore <i>vero differenziale</i>	12-3
12.4 <i>Instrumentation Amplifier (IA)</i>	12-4
12.4.1 Ricevitore differenziale	12-5
12.5 Filtri <i>RC</i>	12-6
12.6 Integratore di Miller	12-7
12.7 Differenziatore	12-9
12.8 NIC (<i>Negative impedance converter</i>)	12-9
12.9 Generatore di corrente (Pompa di Howland)	12-10
12.10 Integratore non invertente	12-10
12.11 Oscillatore a resistenza negativa	12-11
12.12 Filtri attivi	12-12
12.12.1 Filtri Sallen-Key	12-12
12.12.2 Filtro KHN	12-14
12.13 Il filtro <i>passa tutto</i>	12-15
12.14 Comparatore	12-16
12.14.1 Comparatore con isteresi (<i>trigger di Schmitt</i>)	12-17
12.15 Generatore di funzioni	12-18
13 Elaborazione analogica	13-1
13.1 Il Calcolatore Analogico	13-1
13.2 Moltiplicatore a quattro quadranti	13-3
13.2.1 Funzione <i>quadrato</i>	13-5
13.3 Oscillatore armonico	13-6
13.4 Oscillatore <i>quasi armonico</i>	13-8
13.5 Potenziale non armonico. Oscillatore di <i>Duffing</i>	13-10
14 Le configurazioni della reazione	14-1
14.1 Le quattro configurazioni della reazione	14-1
14.2 Condizioni all'ingresso ed all'uscita	14-4
14.3 Quattro modelli per l'amplificatore operazionale	14-7
14.4 Quattro esempi di rete di reazione	14-9
14.5 Amplificazione	14-12
14.6 Guadagno d'anello	14-12
14.7 Resistenza e conduttanza di ingresso e di uscita	14-13

15 Stabilità delle reti lineari	15-1
15.1 Risposta stazionaria di una rete lineare	15-3
15.2 Comportamento durante un transitorio	15-4
15.3 Condizione di stabilità di un anello di reazione	15-6
15.4 Anelli del primo e del secondo ordine	15-7
15.4.1 Anello del secondo ordine	15-8
15.5 Criterio di Nyquist	15-11
15.5.1 Anello di reazione stabile	15-12
15.5.2 Anello di reazione instabile	15-13
15.6 Il criterio di Nyquist nei diagrammi di Bode	15-14
15.7 Margine di fase e di guadagno	15-15
15.8 Compensazione degli anelli di reazione	15-16
15.8.1 Polo dominante	15-18
15.8.2 Polo-zero	15-19
15.9 Stabilità condizionata	15-21
16 I transistor in alta frequenza	16-1
16.1 Transistor <i>bjt</i>	16-2
16.2 Amplificazione di corrente	16-2
16.3 Amplificazione di tensione - Effetto Miller	16-3
16.4 Transistor <i>fet</i>	16-4
16.5 Amplificatore a larga banda	16-5
16.6 Amplificatori a banda stretta	16-7
16.7 Adattamento di impedenza	16-10
16.7.1 Trasformatore di impedenza ad <i>L</i>	16-13
16.8 Circuito <i>cascode</i>	16-14
17 Generazione di segnali	17-1
17.1 Anello con reazione positiva. Il <i>moltiplicatore di Q</i>	17-2
17.2 Oscillatore Colpitts	17-5
17.2.1 Stabilità della frequenza	17-7
17.2.2 Oscillatore <i>a quarzo</i>	17-8
17.3 Rumore di fase	17-11
17.4 Sintesi di frequenza	17-14
17.5 Generatori <i>DDS</i> (<i>Direct Digital Synthesis</i>)	17-14
17.5.1 Rumore nei generatori DDS	17-16
17.6 Generatori <i>PLL</i> (<i>Phase Lock Loop</i>)	17-19
17.7 Stabilità degli anelli <i>PLL</i>	17-20
17.7.1 Guadagno d'anello in un <i>PLL</i>	17-21
17.7.2 Discriminatore di fase	17-23
17.7.3 Discriminatore di fase-frequenza	17-24
17.7.4 Aggancio di fase (<i>lock</i>)	17-26
18 Analisi numerica delle reti	18-1
18.1 Analisi per nodi di una rete	18-1
18.2 Composizione del <i>file</i> di descrizione del circuito	18-2
18.3 Nodi	18-2

18.4 Valori numerici	18-3
18.5 Componenti passivi semplici	18-3
18.5.1 Resistenze, condensatori ed induttanze	18-3
18.5.2 Mutua induttanza	18-4
18.6 Generatori indipendenti	18-4
18.6.1 Generatori di tensioni e correnti continue	18-4
18.6.2 Generatori AC per analisi lineare	18-4
18.6.3 Generatori per analisi di transitorio	18-4
18.7 Circuiti indefiniti o impossibili	18-5
18.8 Analisi dei circuiti	18-6
18.8.1 Determinazione del punto di lavoro	18-6
18.8.2 Funzione di trasferimento in continua	18-7
18.8.3 Funzione di trasferimento in alternata	18-7
18.8.4 Analisi di transitorio	18-7
18.9 Studio di un circuito risonante	18-8
18.10 Presentazione dei risultati	18-10
18.11 Semiconduttori	18-11
18.11.1 Modelli	18-11
18.11.2 Diodi a giunzione	18-12
18.11.3 Transistor <i>bjt</i>	18-13
18.11.4 Transistor <i>fet</i>	18-13
18.12 Esempi di circuiti con semiconduttori	18-14
18.12.1 Curve caratteristiche di un transistor <i>bjt</i>	18-14
18.12.2 Tempi di commutazione di un diodo	18-15
18.12.3 Circuito <i>current mirror</i>	18-15
18.13 Generatori controllati	18-16
18.13.1 Generatori controllati lineari	18-17
18.13.2 Generatori controllati non lineari	18-18
18.14 Sottocircuiti	18-18
18.14.1 Il comando <code>.include</code>	18-19
18.14.2 Modelli lineari per l'amplificatore operazionale	18-19
18.15 Anello di reazione con operazionale	18-20
19 Modulazione e conversione di frequenza	19-1
19.1 Modulazione di ampiezza	19-1
19.2 Identità trigonometriche	19-3
19.2.1 somme e differenze di angoli	19-3
19.2.2 formule di Werner (o di prostaferesi)	19-3
19.3 Bande laterali	19-3
19.4 Conversione di frequenza	19-4
19.5 Ricevitore <i>Supereterodina</i>	19-5
19.5.1 Vantaggi del circuito supereterodina	19-7
19.5.2 Svantaggi del circuito supereterodina	19-8
19.6 L'evoluzione della supereterodina: la SDR	19-8
19.6.1 La modulazione <i>SSB</i>	19-9
19.6.2 La trasformata di Hilbert	19-11
19.6.3 Il segnale analitico	19-12

19.6.4 Conversione di frequenza con il segnale analitico	19-12
19.7 La conversione <i>I-Q</i>	19-13
19.8 I filtri nella supereterodina SDR	19-14
20 Circuiti di potenza	20-1
20.1 Regolatori lineari	20-2
20.1.1 Vantaggi e svantaggi dei regolatori lineari	20-3
20.2 Regolatori <i>switching</i> (a commutazione)	20-4
20.2.1 Tensioni e correnti in un regolatore <i>switching</i>	20-5
20.2.2 Problemi di stabilità e velocità	20-7
20.2.3 Il ruolo del transistor Q_2	20-7
20.2.4 Efficienza energetica	20-8
20.2.5 Altre configurazioni dei regolatori <i>switching</i>	20-9
20.3 Trasformatori DC-DC	20-10
20.4 Amplificatori di potenza	20-11
20.4.1 Classe A	20-11
20.4.2 Classe B	20-13
20.4.3 Classe C	20-14
20.4.4 Classe D	20-15
A Valori dei componenti	A-1
B	B-1

Capitolo 1

Circuiti

Questo capitolo contiene un richiamo sintetico ai punti principali dello studio dei circuiti elettrici, per lo più come elenco senza dimostrazioni o spiegazioni approfondate, che invece possono essere trovate in “G. Poggi - Esperimenti di Elettricità e magnetismo” (http://studenti.fisica.unifi.it/~a_perego/Esperimentazioni_II-Lab_Fisica_II/Dispense/esper_elettr_magn.pdf).

1.1 Segnali elettrici

Un segnale elettrico è una quantità fisica, come una tensione o una corrente in un circuito elettrico, le cui variazioni in funzione del tempo contengono *informazione*.

Ad ogni segnale è associato il suo *spettro di Fourier* determinato dallo sviluppo in serie per i segnali periodici e dalla trasformata per i non periodici.

Se $v(t)$ è un segnale periodico con periodo T , frequenza $f = 1/T$ e frequenza angolare $\omega = 2\pi f = 2\pi/T$, valgono le relazioni:¹

$$v(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos(n\omega t) + B_n \sin(n\omega t)$$

oppure

$$v(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} S_n \cos(n\omega t + \varphi_n) \quad (1.1)$$

con

$$A_n = \frac{2}{T} \int_0^T v(t) \cos(n\omega t) dt \quad B_n = \frac{2}{T} \int_0^T v(t) \sin(n\omega t) dt$$

$$S_n = \sqrt{A_n^2 + B_n^2} \quad \varphi_n = \arctan(B_n/A_n) \quad (1.2)$$

¹Con le opportune condizioni matematiche, cui soddisfano praticamente tutti i segnali di interesse fisico.

o, in forma complessa:²

$$v(t) = \sum_{n=-\infty}^{\infty} V_n e^{jn\omega t} \quad (1.3)$$

$$V_n = \frac{1}{T} \int_0^T v(t) e^{-jn\omega t} dt \quad (1.4)$$

Poichè la funzione $v(t)$ è reale, vale la condizione $V_{-n} = V_n^*$ e di conseguenza $A_n = V_n + V_{-n}$, $B_n = V_n - V_{-n}$. La componente a frequenza angolare ω è la *fondamentale*, le componenti a frequenza $n\omega$ sono le *armoniche*. La quantità $A_0/2$ ($= V_0$) è la *componente continua*, cioè il valore medio del segnale.

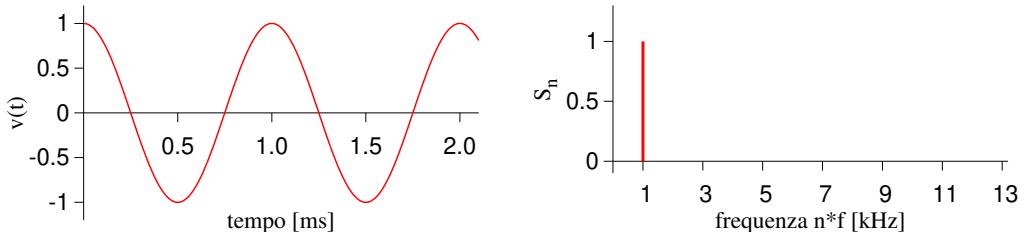


fig. 1.1a: segnale $\cos(2\pi ft)$ con $f = 1$ kHz.

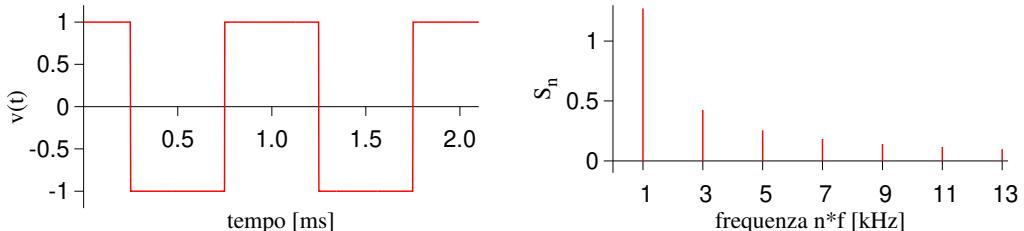


fig. 1.1b: onda quadra con frequenza $f = 1$ kHz.

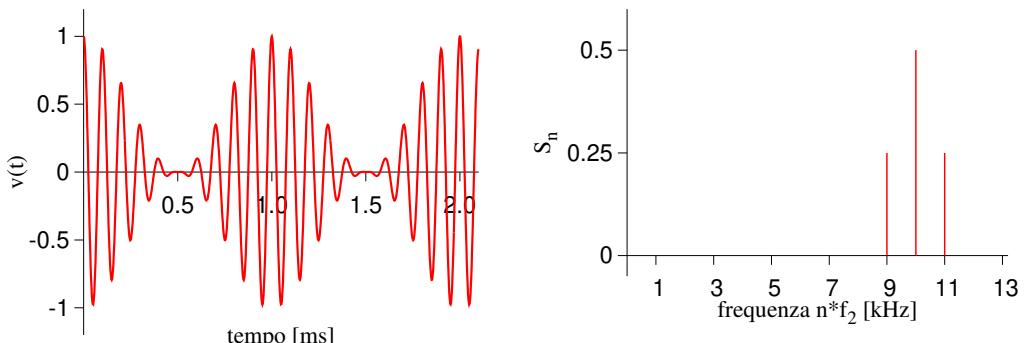


fig. 1.1c: modulazione di ampiezza: segnale $0.5 \cos(2\pi f_1 t) \cdot (1 + \cos(2\pi f_2 t))$ con $f_1 = 10$ kHz ed $f_2 = 1$ kHz.

Figura 1.1: Alcune funzioni $v(t)$ e le loro componenti spettrali S_n . In tutti e tre gli esempi l'origine dei tempi è stata scelta in modo di avere tutte le fasi φ_n nulle.

²In elettrotecnica ed in elettronica si usa di solito la lettera j come unità immaginaria, in quanto la i è utilizzata di preferenza per indicare la corrente.

Convenzionalmente si utilizzano le lettere minuscole con un eventuale indice minuscuso per rappresentare i segnali nel dominio del tempo e le lettere maiuscole, sempre con un eventuale indice minuscolo, per rappresentare le componenti degli spettri degli stessi segnali nel dominio della frequenza. Quindi, se $v_a(t)$ è una funzione che descrive l'andamento temporale della tensione sul nodo a di un circuito, $V_a(f)$ è il suo spettro di Fourier in funzione della frequenza.

Le due funzioni $v_a(t)$ e $V_a(f)$, legate tra loro dalle relazioni di Fourier, sono equivalenti nel rappresentare il segnale e contengono la stessa informazione.

Nella fig. 1.1 sono riportati tre esempi di segnali $v(t)$ funzioni del tempo ed i loro relativi spettri $S(f)$.

Il primo è un segnale sinusoidale puro con periodo di 1 ms ed il suo spettro contiene di conseguenza la sola componente armonica S_1 di frequenza $f = 1 \text{ kHz}$ con ampiezza corrispondente all'ampiezza della sinusode.

Il secondo è un'onda quadra con periodo 1 ms, quindi ancora con frequenza 1 kHz. Il suo spettro di Fourier contiene tutte le infinite armoniche dispari con ampiezza decrescente: $S_n = 4/(n\pi)$ con $n = \pm 1, 3, 5, \dots$. Solo le prime 7 sono riportate nel grafico.

Il terzo è un segnale sinusoidale di frequenza $f_1 = 10 \text{ kHz}$ modulato in ampiezza da un altro segnale sinusoidale di frequenza $f_2 = 1 \text{ kHz}$:

$$\begin{aligned} v(t) &= 0.5 \cdot \cos(\omega_1 t) \cdot (1 + \cos(\omega_2 t)) \\ &= 0.5 \cos(\omega_1 t) + 0.25 \cos((\omega_1 + \omega_2)t) + 0.25 \cos((\omega_1 - \omega_2)t) \end{aligned}$$

con $\omega_{1,2} = 2\pi f_{1,2}$. Lo spettro contiene ancora il segnale a frequenza f_1 (la *portante*), con ampiezza dimezzata, e le due *bande laterali* a frequenza $f_1 - f_2$ e $f_1 + f_2$.

In questi tre esempi l'origine dei tempi è stata scelta in modo tale che le fasi di tutte le componenti armoniche sono nulle e di conseguenza non sono state riportate nei grafici. Normalmente uno spettro di Fourier completo contiene oltre alle ampiezze delle componenti armoniche anche le rispettive fasi.

1.2 Leggi di Kirchhoff

Le leggi di Kirchhoff riguardano tutte le reti, sia lineari che non lineari, perchè sono l'applicazione ai circuiti elettrici di due leggi fondamentali della fisica: la conservazione della carica elettrica e la conservazione dell'energia. Il significato dei termini *rete*, *ramo* e *nodo* che compaiono nell'enunciato delle due leggi si suppone già noto.

1.2.1 I Legge di Kirchhoff - Legge dei nodi

La somma algebrica delle correnti nei rami che fanno capo ad uno stesso nodo di una rete è nulla (conservazione della carica elettrica).

1.2.2 II Legge di Kirchhoff - Legge delle maglie

La somma algebrica delle differenze di potenziale ai capi di tutti i rami di una stessa maglia di una rete è nulla (conservazione dell'energia).

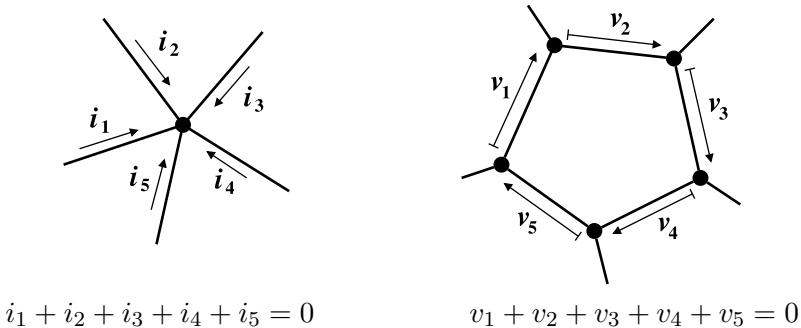


Figura 1.2: Leggi di Kirchhoff.

In base alle leggi di Kirchhoff è possibile scrivere un sistema di equazioni, ad esempio una per ogni nodo oppure una per ogni maglia del circuito, in cui compaiono le correnti o le tensioni come variabili incognite. La soluzione del sistema di equazioni fornisce la descrizione completa del funzionamento del circuito.

Per le reti lineari è sufficiente un sistema di equazioni algebriche, reali per i circuiti in corrente continua, complesse per i circuiti in corrente alternata. Nel caso più complicato, quello dei circuiti non lineari con tensioni e correnti variabili nel tempo, si avrà un sistema di equazioni differenziali non lineari da integrare in funzione del tempo. Questo compito oggi è parecchio facilitato dalla disponibilità di programmi di calcolo numerico come *Spice*, descritto nel cap. 18.

1.3 Reti lineari - principio di sovrapposizione

Un circuito elettrico si dice *rete lineare* quando i suoi rami contengono solamente componenti lineari, cioè componenti in cui le relazioni tra tensioni e correnti sono descritte da equazioni algebriche o differenziali lineari. Sono componenti lineari resistenze, condensatori, induttanze, mutue induttanze e generatori lineari, indipendenti e controllati.

Per le reti lineari (e solo per le reti lineari) vale il *principio di sovrapposizione*: la risposta di una rete lineare ad un segnale composto dalla somma di due segnali è data dalla somma delle singole risposte della rete ai segnali componenti considerati singolarmente uno per volta: se $v_o(t)$ e $u_o(t)$ sono le risposte di una rete ai singoli segnali $v_i(t)$ e $u_i(t)$, la risposta al segnale $v_i(t) + u_i(t)$ è il segnale $v_o(t) + u_o(t)$.

In modo del tutto equivalente, nel dominio delle frequenze: se $V_o(f)$ e $U_o(f)$ sono le risposte di una rete ai singoli segnali $V_i(f)$ e $U_i(f)$, allora la risposta al segnale $V_i(f) + U_i(f)$ è il segnale $V_o(f) + U_o(f)$.

La generalizzazione a segnali contenenti N componenti invece di 2 è immediata.

Grazie al principio di sovrapposizione la risposta di una rete lineare ad un segnale $v_i(t)$ funzione del tempo risultante dalla somma di tutte le sue componenti di Fourier può essere ottenuta conoscendo la risposta della rete ad ogni componente armonica in funzione della frequenza e sommando tutte le risposte alle singole componenti

armoniche $V_i(f)$.³ Se

$$v_i(t) = \sum_n V_i(f_n) e^{j2\pi f_n t} \quad (1.5)$$

allora

$$v_o(t) = \sum_n V_o(f_n) e^{j2\pi f_n t} = \sum_n H(f_n) \cdot V_i(f_n) e^{j2\pi f_n t} \quad (1.6)$$

dove $H(f) = V_o(f)/V_i(f)$ è la risposta della rete alla componente armonica pura di frequenza f .

1.4 Equazioni delle reti lineari

1.4.1 Legge di Ohm

La *legge di Ohm* non è una legge fisica, ma la definizione di una classe di conduttori, i conduttori *ohmici*, per i quali la relazione tra tensione e corrente è lineare:

$$v = i \cdot R \quad i = v \cdot G \quad (1.7)$$

dove v ed i sono tensione e corrente funzioni del tempo ed R (resistenza) e $G = 1/R$ (conduttanza) sono costanti.

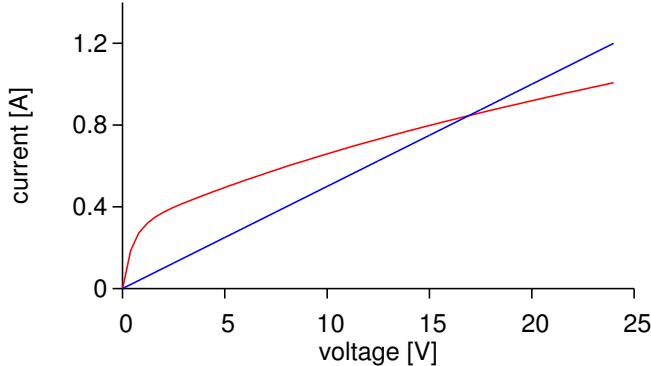


Figura 1.3: Relazione tensione-corrente per un conduttore ohmico e non ohmico. In blu: resistenza da 20Ω , in rosso filamento di tungsteno di una lampadina da $24 V, 25 W$.

Nella fig. 1.3 è riportata la relazione tensione-corrente per un conduttore ohmico (in blu, resistenza da 20Ω) ed un conduttore non ohmico (in rosso, filamento di tungsteno di una lampadina ad incandescenza da $24 V, 25 W$).

1.4.2 Legge di Ohm complessa

Le stesse relazioni 1.7 valgono anche tra le componenti armoniche $V(f)$ ed $I(f)$ dei segnali $v(t)$ ed $i(t)$:

$$V(f) = I(f) \cdot R \quad I(f) = V(f) \cdot G \quad (1.8)$$

³Con le limitazioni che si vedranno nel cap. 15 riguardo ai comportamenti di transitorio.

con V ed I quantità complesse contenenti ampiezza e fase di ciascuna componente. Nella stessa forma possono essere scritte le relazioni tra tensione e corrente per induttanze e capacità (condensatori). Per questi due componenti valgono le relazioni lineari⁴

$$v(t) = L \cdot \frac{di}{dt} \quad v(t) = \frac{1}{C} \cdot \int i \, dt \quad (1.9)$$

che diventano

$$V(\omega) = j\omega L I(\omega) \quad V(\omega) = \frac{1}{j\omega C} I(\omega) \quad (1.10)$$

tra i valori di tensioni e correnti delle singole componenti armoniche.

Le quantità $j\omega L$ e $1/(j\omega C)$ sono le *reattanze* rispettivamente dell'induttanza L e del condensatore C alla frequenza angolare ω e sono quantità immaginarie pure. Corrente e tensione nell'induttanza e nel condensatore sono *in quadratura*, con un angolo di fase di $+90^\circ$ per l'induttanza e di -90° per il condensatore.

La somma $Z = R + jX$ di una resistenza R reale e di una reattanza jX immaginaria è una *impedenza*. Analogamente si definisce una *ammettenza* $Y = G + jB$, somma di una conduttanza G e di una *suscettanza* B .

In analogia con le 1.7 si ha quindi la *legge di Ohm complessa*

$$V = I \cdot Z \quad I = V \cdot Y \quad (1.11)$$

dove V , I , Z ed Y sono tutte quantità complesse funzioni di ω .

1.4.3 Teorema di Thévenin

È un teorema fondamentale nello studio delle reti lineari, perchè consente di riassumere il comportamento di reti anche molto complesse in soli due parametri.

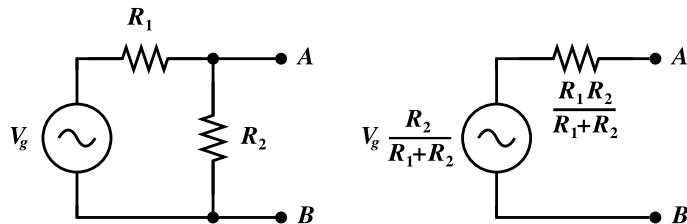


Figura 1.4: Partitore di tensione (a sinistra) e suo equivalente di Thévenin (a destra).

Consideriamo una rete comunque complessa (purchè lineare⁵) e due nodi a piacere della rete, A e B . Collegando i due nodi A e B ad altri circuiti, tutta la rete si comporta come un unico generatore di tensione V con in serie una resistenza R . Il valore di V è quello che esiste tra i due nodi con la rete in funzione a *circuito aperto*, cioè senza prelevare corrente dai nodi. La resistenza R è quella che si misura tra i

⁴La *differenza di potenziale* ai capi dell'induttanza ha segno opposto a quello della *forza elettromotrice* generata per autoinduzione e questo cancella il segno – della legge di Faraday-Lenz.

⁵O che *appaia* lineare ai terminali A e B : esistono reti al loro interno fortemente non lineari, che complessivamente si comportano come dispositivi quasi perfettamente lineari. Ad esempio, gli amplificatori in classe B descritti nel par. 7.12.

due nodi quando tutti i generatori indipendenti contenuti nella rete sono azzerati e sostituiti dalla loro resistenza interna.

Nella fig. 1.4 è riportato un esempio di applicazione del teorema di Thévenin. La tensione a vuoto all'uscita del partitore a sinistra è $V'_g = V_g \cdot R_2 / (R_1 + R_2)$; la resistenza che si misura tra A e B con il generatore V_g azzerato è data da $R_1 // R_2$ (R_1 ed R_2 in parallelo). Il circuito a destra, con i valori riportati per generatore e resistenza, è *completamente equivalente* a quello di sinistra. *Completamente equivalente* significa che se i due circuiti fossero racchiusi in una scatola nera,⁶ nessuna misura elettrica ai terminali A e B permetterebbe di distinguere il circuito originale dal suo equivalente di Thévenin.

1.4.4 Teorema di Norton

Ogni generatore (ideale) di tensione V con una resistenza R in serie equivale ad un generatore (ideale) di corrente I con una conduttanza G in parallelo, con I e G date da:

$$I = \frac{V}{R} \quad G = \frac{1}{R} \quad (1.12)$$

Come è possibile sostituire una rete complessa, collegata tramite due dei suoi nodi con altri circuiti, con il suo equivalente di Thévenin, così la stessa rete può essere sostituita con il suo *equivalente di Norton*: un generatore di corrente con una conduttanza in parallelo. È immediato verificare che i due generatori a destra ed a sinistra in fig. 1.5 danno gli stessi valori di tensione e di corrente in condizione di circuito aperto e di cortocircuito. Avendo stabilito che si tratta di dispositivi o reti lineari, tensioni e correnti coincideranno anche in tutte le altre condizioni.

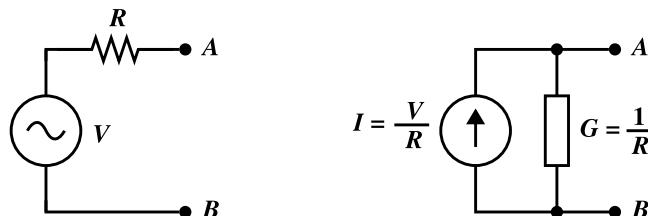


Figura 1.5: Generatore di tensione di Thévenin e suo equivalente di Norton.

È esclusivamente una questione di convenienza rappresentare una rete complessa con il suo generatore equivalente di Thévenin, con resistenza in serie, o con il suo generatore equivalente di Norton, con conduttanza in parallelo. Alcuni circuiti tendono a comportarsi come generatori di tensione, altri come generatori di corrente. Sarà più conveniente utilizzare il generatore equivalente di Thévenin nel primo caso, con una resistenza in serie che in prima approssimazione è zero o comunque di valore piccolo ed il generatore equivalente di Norton nel secondo caso, con una conduttanza in

⁶Si dice che un circuito è racchiuso in una *scatola nera* quando non si può avere nessuna altra informazione sulle sue caratteristiche oltre quelle che si ottengono dalle misure elettriche ai suoi terminali.

parallelo nulla in prima approssimazione, o piccola. Facendo le scelte opposte, nelle equazioni dei circuiti in prima approssimazione compaiono fastidiosi valori *infinito*.

1.5 Massimo trasferimento di potenza

Quando si ha a che fare con segnali molto deboli è importante riuscire a trasferire il massimo segnale possibile dalla sorgente ai circuiti successivi, destinati per lo più all'amplificazione.

Un esempio è dato dal collegamento dell'antenna di un radiotelescopio al suo ricevitore. Come ogni misura fisica è soggetta ad un *errore di misura*, così nei circuiti elettronici ogni segnale è soggetto al *rumore* di cui una parte è sempre generata all'interno dei circuiti stessi ed è inevitabile. Ogni perdita, anche piccola, del segnale disponibile riduce il *rappporto segnale/rumore*, cioè aumenta l'*errore* nella rivelazione del segnale e riduce di conseguenza il campo di esplorazione possibile.

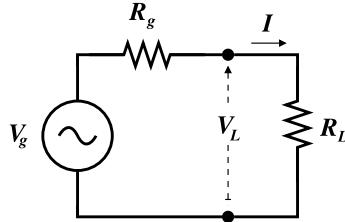


Figura 1.6: Accoppiamento di una sorgente di segnale (V_g , R_g) ad un circuito utilizzatore (R_L).

Il problema del massimo trasferimento di potenza è schematizzato nella figura 1.6, dove la sorgente del segnale composta dal generatore di tensione V_g con resistenza interna R_g è collegata ad un circuito utilizzatore schematizzato dalla resistenza R_L .

La potenza del segnale trasferito dal generatore all'utilizzatore è

$$P_L = |I|^2 \cdot R_L = |V_g|^2 \cdot \frac{R_L}{(R_g + R_L)^2} \quad (1.13)$$

ed ha un massimo, come ben noto, quando $R_g = R_L$.

Nel caso di circuiti in corrente alternata si può avere a che fare con un generatore con impedenza interna complessa $Z_g = R_g + jX_g$. Sostituendo nella eq. 1.13 Z_g per R_g ed aggiungendo anche ad R_L una parte immaginaria jX_L , si può determinare la nuova condizione per il *massimo trasferimento di potenza*.

Si trova di nuovo $R_L = R_g$ e per la parte immaginaria $X_L = -X_g$.

Per il massimo trasferimento di potenza l'impedenza dell'utilizzatore del segnale deve avere il valore complesso coniugato dell'impedenza della sorgente: $Z_L = Z_g^$.*

Questa condizione si chiama *adattamento di impedenza (impedance matching)*.

Rendimento

Una sorgente di segnale fornisce energia ai circuiti utilizzatori e quindi deve essere dotata di una propria fonte di energia. Il rendimento è il rapporto tra la potenza erogata all'utilizzatore del segnale e quella prelevata dalla sorgente di energia.

Nel caso che la sorgente di segnale sia *fisicamente* composta come indicato in fig. 1.6 da un generatore V_g con in serie una resistenza R_g , il rendimento η è

$$\eta = \frac{i^2 R_L}{i^2 R_L + i^2 R_g} = \frac{R_L}{R_L + R_g} \quad (1.14)$$

In condizioni di adattamento di impedenza ($R_L = R_g$) si ha $\eta = 50\%$; per $R_L < R_g$ il rendimento scende sotto il 50% e tende verso lo zero insieme ad R_L ; per $R_L > R_g$ il rendimento cresce e tende al 100% per $R_L \rightarrow \infty$. Quindi il massimo rendimento energetico si può avere solo con il massimo disadattamento di impedenza.

Ad esempio, se si tentasse di far lavorare un generatore di tensione quasi ideale come una batteria di avviamento d'automobile in condizioni di adattamento di impedenza si otterrebbero risultati disastrosi. Tipicamente un accumulatore al piombo per auto fornisce una tensione di 12.6 V e presenta una resistenza interna dell'ordine di pochi $m\Omega$. Ad esempio, con 3 $m\Omega$ per R_g ed altrettanti per R_L , si avrebbe una corrente di 2100 A, con una erogazione di potenza di oltre 13 kW verso il carico ed una dissipazione uguale all'interno della batteria. Che probabilmente esploderebbe dopo pochi secondi.

In generale, però, si deve tenere presente che la schematizzazione di Thévenin di fig. 1.6 *** non è *** valida ai fini del calcolo del rendimento. L'equivalente di Thévenin di una sorgente di segnale consente solo di calcolare la potenza scambiata tra due reti, non permette di dire nulla sui bilanci energetici interni di una rete: per poter analizzare il rendimento di una rete è necessaria la conoscenza completa di tutta la rete.

L'esempio precedente con la batteria è corretto solo perché la resistenza R_g è *fisicamente* presente all'interno della batteria come resistenza ohmica degli elettrodi, della soluzione eletrolitica ecc.

1.6 Circuiti RC

I due circuiti *RC* più semplici, la cella *passa alto* e la cella *passa basso* in fig. 1.7, alla pari del partitore di tensione già visto in fig. 1.4, compaiono talmente spesso nei circuiti elettronici che conviene richiamarne le caratteristiche principali.

1.6.1 Filtri passa-alto e passa-basso - Il *decibel*

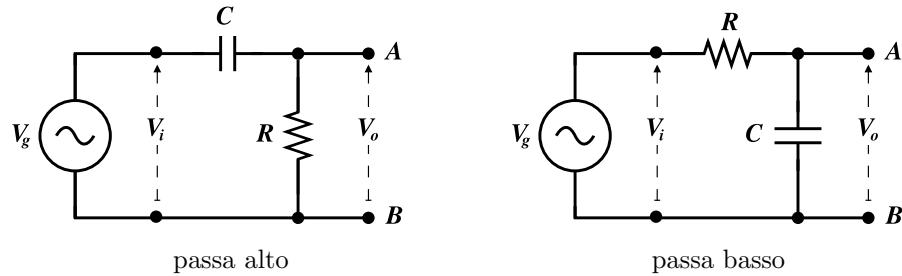


Figura 1.7: Filtro *RC* passa alto e passa basso.

Un partitore di tensione come nella fig. 1.7 in cui uno dei due rami è una resistenza e l'altro è un condensatore è un *filtro* in quanto il suo rapporto di attenuazione dipende dalla frequenza e parti diverse dello spettro delle frequenze saranno attenuate in modo diverso. Quando il ramo inferiore del partitore è la resistenza, si ha il *filtro passa-alto*; nel caso opposto, si ha il *filtro passa-basso*.

Applicando all'ingresso dei due filtri di fig. 1.7 una tensione alternata sinusoidale V_i di frequenza f , tra ingresso e uscita si avranno rispettivamente per il passa-alto ed il passa-basso le attenuazioni

$$A_a = \frac{V_o}{V_i} = \frac{1}{1 - j f_a/f} \quad A_b = \frac{V_o}{V_i} = \frac{1}{1 + j f/f_b} \quad (1.15)$$

dove il rapporto V_o/V_i è un numero complesso ed f_a ed f_b sono le *frequenze di taglio* determinate dalla costante di tempo RC del circuito: $f_{a,b} = 1/(2\pi RC)$ (a per alto, b per basso).

Per riportare in grafico $A_a(f)$, $A_b(f)$ sono particolarmente convenienti i *diagrammi di Bode* (fig. 1.8).

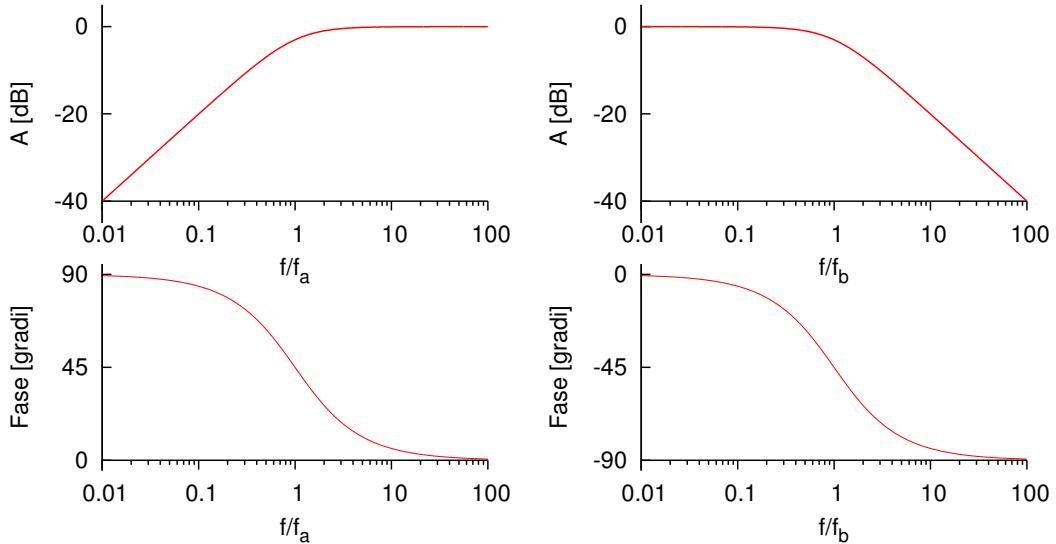


Figura 1.8: Diagrammi di Bode: modulo e fase della attenuazione in funzione della frequenza per il filtro passa-alto (a sinistra) e passa-basso (a destra).

L'attenuazione, e come si vedrà anche l'amplificazione, è un rapporto che può spaziare su molti ordini di grandezza e quindi è conveniente utilizzare una scala logaritmica. L'unità di misura logaritmica dell'attenuazione e dell'amplificazione è il *bel*:

$$A(\text{bel}) = \log_{10} |A|$$

Il *bel* non è praticamente mai entrato nell'uso comune, mentre è utilizzato universalmente il suo sottomultiplo *decibel* (*dB*): $1 \text{ dB} = 0.1 \text{ bel}$. Quindi una attenuazione/amplificazione di potenza espressa in *decibel* è

$$A_p(\text{dB}) = 10 \cdot \log_{10} |A_p|$$

A causa della relazione quadratica che esiste tra tensione e potenza, l'attenuazione/amplificazione di tensione in *decibel* è per definizione

$$A_v(dB) = 20 \cdot \log_{10} |A_v|$$

e analogamente per la corrente:

$$A_i(dB) = 20 \cdot \log_{10} |A_i|$$

Nei grafici di Bode vengono riportati il modulo dell'amplificazione, espresso in dB , e la fase, entrambi in funzione del logaritmo della frequenza.

Dai grafici di fig. 1.8 e dalle eq. 1.15 si vede come per il filtro passa-basso (a destra) l'attenuazione sia nulla alla corrente continua e nel limite delle basse frequenze; alla frequenza di taglio $f = f_b$ l'attenuazione ha il valore di $-3 dB$ ⁷ e la fase ha subito una rotazione di 45° . Al crescere ulteriore della frequenza l'attenuazione varia di $-20 dB$ per ogni decade di frequenza (pari a $-6 dB$ per ogni ottava⁸). Per il filtro passa alto, a sinistra nel grafico, si ha il comportamento speculare: l'attenuazione va a $-\infty$ con la pendenza di $6 dB$ per ottava al decrescere della frequenza verso lo zero (che su scala logaritmica si trova a $-\infty$), ha il valore $-3 dB$ a $f = f_a$ e tende asintoticamente a zero quando la frequenza cresce oltre la frequenza di taglio f_a .

1.6.2 Risposta all'onda quadra

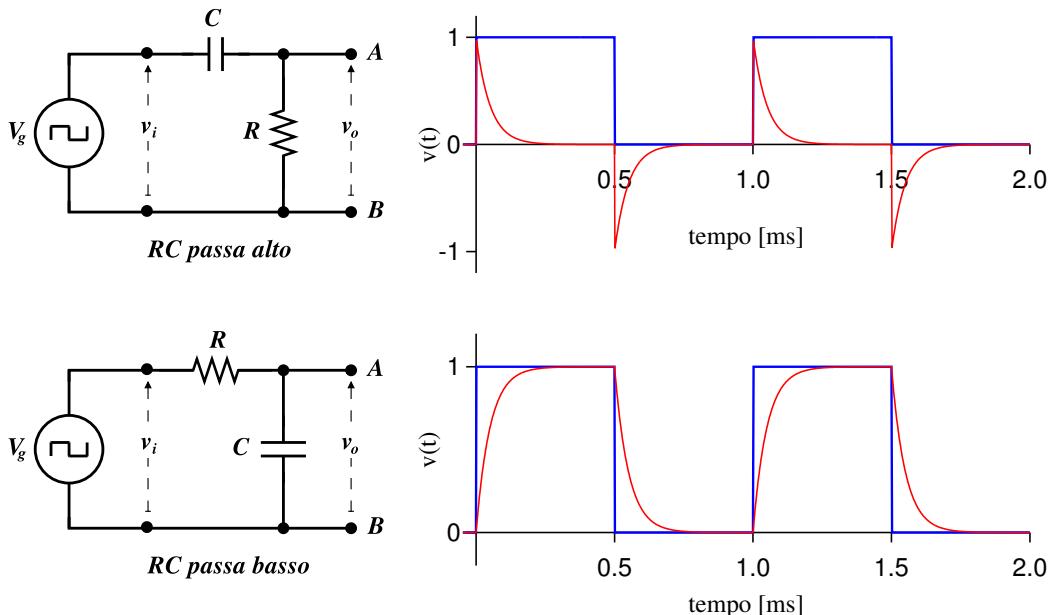


Figura 1.9: Cella RC passa alto e passa basso. La costante di tempo è $\tau = RC = 0.1 ms$.

⁷Più esattamente $-10 \log_{10}(2) = -3.0103....$

⁸Più esattamente $-20 \log_{10}(2) = -6.0206....$

Il comportamento dei due circuiti RC passa alto e passa basso in funzione del tempo si può vedere molto bene applicando all'ingresso un'onda quadra, come descritto nella fig. 1.9. Ad ogni fronte di salita o di discesa si ha nel circuito un transitorio di corrente $i(t)$ che decade esponenzialmente verso zero con una costante di tempo $\tau = RC$:

$$i(t) = i_0 \cdot e^{-t/\tau} \quad (1.16)$$

(i due circuiti di fig. 1.9 sono identici per quanto riguarda la corrente; differiscono per i nodi dove viene prelevata la tensione di uscita). Se il periodo T dell'onda quadra è sufficientemente lungo da consentire ai transitori di decadere completamente tra un fronte ed il successivo, come nei grafici, in cui $T = 1 \text{ ms}$ e $\tau = 0.1 \text{ ms}$, il valore iniziale della corrente ad ogni transitorio è

$$i_0 = \frac{\Delta V}{R} \quad (1.17)$$

dove ΔV è l'ampiezza del fronte dell'onda quadra ed R la resistenza presente nel circuito (è da notare che il valore della corrente nell'istante iniziale del transitorio non dipende dal valore della capacità C). Conoscendo la corrente $i(t)$ si possono scrivere le espressioni per $v_o(t)$:

$$\begin{aligned} \text{passa alto: } v_o(t) &= i(t) \cdot R = \Delta V e^{-t/\tau} \\ \text{passa basso: } v_o(t) &= \Delta V - i(t) \cdot R = \Delta V \left(1 - e^{-t/\tau}\right) \end{aligned}$$

1.7 Circuiti risonanti LRC

Il circuito risonante di fig. 1.10, composto da una capacità C , una induttanza L ed una resistenza R , è l'equivalente elettronico dell'oscillatore armonico meccanico.

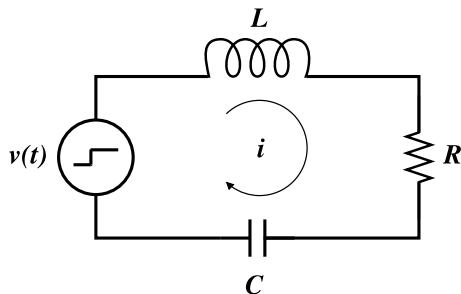


Figura 1.10: Circuito risonante composto da induttanza L , resistenza R e capacità C in serie.

Il suo funzionamento è descritto dall'equazione integro-differenziale

$$L \frac{di}{dt} + Ri + \frac{1}{C} \int idt = v(t) \quad (1.18)$$

equivalente a

$$L \ddot{q} + R \dot{q} + \frac{1}{C} q = v(t) \quad (1.19)$$

dove q è la carica accumulata sulle armature del condensatore e $i = \dot{q}$.

Limitandosi al caso $v(t) = \text{costante}$ per $t > 0$, ad esempio dopo un fronte di salita della tensione v da $v(t) = 0$ per $t < 0$ a $v(t) = v_1$ per $t > 0$, si ha la soluzione statica asintotica

$$q = v_1 C \quad i = 0 \quad (1.20)$$

A questa va aggiunta la soluzione dell'equazione omogenea associata

$$L\ddot{q} + R\dot{q} + \frac{1}{C}q = 0 \quad (1.21)$$

Sostituendo in quest'ultima, come indicato dall'analisi, la soluzione di tipo armonico/esponenziale

$$q(t) = Ae^{\lambda t} \quad (1.22)$$

si ottiene l'*equazione caratteristica*

$$L\lambda^2 + R\lambda + \frac{1}{C} = 0 \quad (1.23)$$

le cui radici sono

$$\lambda_{1,2} = -\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} \quad (1.24)$$

Introducendo come di consueto la frequenza di risonanza ω_0 ed il fattore di merito Q_0 definiti da

$$\omega_0 = \frac{1}{\sqrt{LC}} \quad Q_0 = \frac{\omega_0 L}{R} = \frac{1}{\omega_0 R C} = \frac{1}{R} \sqrt{\frac{L}{C}} \quad (1.25)$$

la eq. 1.24 diventa

$$\lambda_{1,2} = -\frac{\omega_0}{2Q_0} \pm \omega_0 \sqrt{\frac{1}{4Q_0^2} - 1} \quad (1.26)$$

A seconda che si abbia $Q_0 > 1/2$ oppure $Q_0 < 1/2$ le due soluzioni $\lambda_{1,2}$ sono complesse coniugate con parte reale ed immaginaria

$$\alpha = -\frac{\omega_0}{2Q_0} = -\frac{1}{2} \frac{R}{L} \quad \beta = \pm j\omega_0 \sqrt{1 - \frac{1}{4Q_0^2}} \quad (1.27)$$

oppure entrambe reali e negative

$$\alpha_{1,2} = -\frac{\omega_0}{2Q_0} \left(1 \pm \sqrt{1 - 4Q_0^2} \right) \quad (1.28)$$

Nel primo caso si hanno le oscillazioni armoniche smorzate

$$q(t) = e^{\alpha t} \cdot (Ae^{j\beta t} + Be^{-j\beta t}) \quad (1.29)$$

e nel secondo caso un rilassamento senza oscillazioni verso la condizione di equilibrio

$$q(t) = A e^{\alpha_1 t} + B e^{\alpha_2 t} \quad (1.30)$$

Il caso limite $Q_0 = 1/2$ porta ad $\alpha_1 = \alpha_2 = \omega_0/(2Q_0)$ e corrisponde alla condizione di *smorzamento critico*, cioè al più veloce rilassamento privo di oscillazioni. In queste condizioni anche la funzione $te^{\alpha t}$ è soluzione della eq. 1.21 per cui si ha

$$q(t) = e^{\alpha t} \cdot (A + Bt) \quad (1.31)$$

Le due costanti A e B sono determinate dalle condizioni iniziali e cioè dal valore della carica q_0 sul condensatore e della corrente i_0 nell'induttanza nell'istante $t = 0$.

1.8 Filtri LRC

Nel dominio delle frequenze l'impedenza complessa della serie *LRC* è

$$Z_s(\omega) = R + j \left(\omega L - \frac{1}{\omega C} \right) = R \left[1 + j Q_0 \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right) \right] \quad (1.32)$$

dove ω_0 e Q_0 hanno lo stesso significato del paragrafo precedente. L'andamento in modulo e fase della ammettenza $Y_s(\omega) = 1/Z_s(\omega)$ in funzione della frequenza è riportata in fig. 1.11.

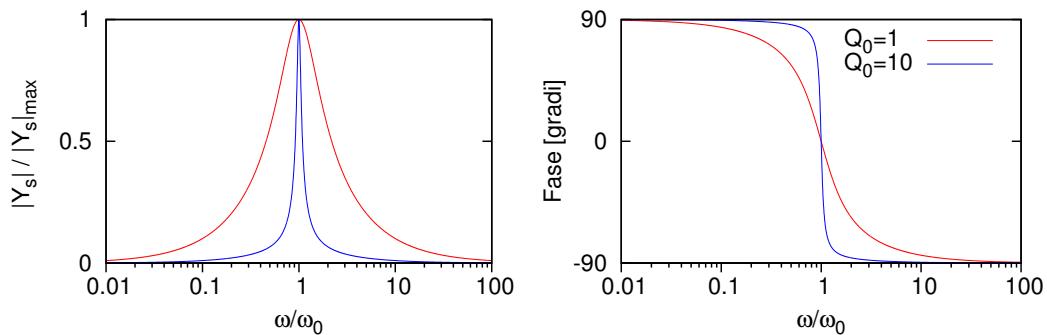


Figura 1.11: Picco di risonanza nel circuito *LRC* serie.

L'andamento della ammettenza corrisponde all'andamento della corrente nel circuito quando viene alimentato da un generatore ideale di tensione. Si osserva nel grafico il tipico picco di un sistema risonante.

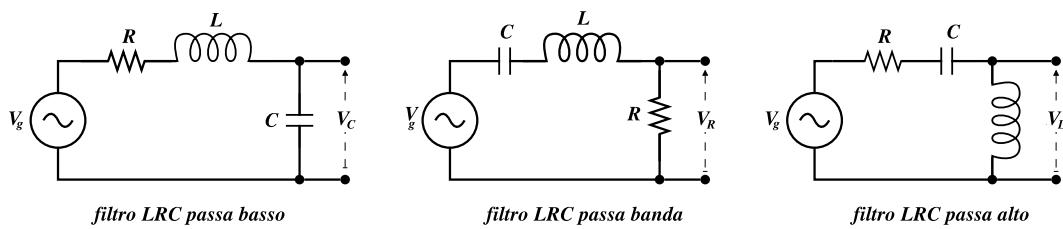


Figura 1.12: Filtri *LRC* passa basso, passa banda e passa alto.

Alimentando il circuito risonante con un segnale sinusoidale puro con frequenza angolare ω e prelevando come segnale di uscita la tensione che si ottiene ai capi del condensatore, della resistenza o della induttanza come indicato in fig. 1.12 si hanno le tre risposte in frequenza

$$H_l(\omega) = \frac{V_C}{V_G} = \frac{1}{j\omega CZ_s} = \frac{-jQ_0 \frac{\omega_0}{\omega}}{1 + jQ_0 \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right)} = \frac{1}{1 - \frac{\omega^2}{\omega_0^2} + \frac{j}{Q_0} \frac{\omega}{\omega_0}} \quad (1.33)$$

$$H_b(\omega) = \frac{V_R}{V_G} = \frac{R}{Z_s} = \frac{1}{1 + jQ_0 \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right)} = \frac{\frac{j}{Q_0} \frac{\omega}{\omega_0}}{1 - \frac{\omega^2}{\omega_0^2} + \frac{j}{Q_0} \frac{\omega}{\omega_0}} \quad (1.34)$$

$$H_h(\omega) = \frac{V_L}{V_G} = \frac{j\omega L}{Z_s} = \frac{jQ_0 \frac{\omega}{\omega_0}}{1 + jQ_0 \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right)} = \frac{-\frac{\omega^2}{\omega_0^2}}{1 - \frac{\omega^2}{\omega_0^2} + \frac{j}{Q_0} \frac{\omega}{\omega_0}} \quad (1.35)$$

che corrispondono a tre filtri di tipo rispettivamente passa basso (*low pass*), passa banda (*band pass*) e passa alto (*high pass*). Le variabili ω_0 e Q_0 hanno lo stesso significato del paragrafo precedente.

Passa basso

Nel primo caso, per $\omega \rightarrow 0$ e nel limite delle basse frequenze, la funzione H_l assume asintoticamente il valore costante 1, mentre nel limite $\omega \rightarrow \infty$ va asintoticamente a zero, con un andamento $|H_l| \propto 1/\omega^2$. Infatti, nel limite delle alte frequenze, nell'espressione di $Z_s(\omega)$ (eq. 1.32) è il termine ωL a prevalere e quindi la corrente circolante nella serie *RLC* diminuisce come $1/\omega$. D'altra parte la tensione viene prelevata ai capi del condensatore, la cui reattanza a sua volta diminuisce come $1/\omega$. I due effetti combinati producono l'andamento proporzionale a $1/\omega^2$. Nei diagrammi di Bode questo equivale ad una pendenza di -40 dB per decade (fig. 1.13), il doppio rispetto ai -20 dB per decade del circuito passa basso *RC*. Questi sono filtri del *secondo ordine*, a differenza dei precedenti filtri *RC*, detti del *primo ordine*.

Alla frequenza di risonanza $\omega = \omega_0$ si ha $H_l = -jQ_0$ e la funzione $|H_l|$ può presentare un picco, come si vede nella fig. 1.13.

Dal grafico e dalle equazioni si vede che in corrispondenza del picco la tensione ai capi del condensatore (e similmente si vedrà anche ai capi dell'induttanza) risulta Q_0 volte maggiore della tensione con cui il circuito è alimentato. Questa è una cosa di cui si deve sempre tenere conto nel valutare le sollecitazioni a cui un componente è sottoposto, come ci è stato insegnato da tempo dagli effetti delle risonanze nei sistemi meccanici.

In un certo senso nell'intorno della risonanza il circuito si comporta come un trasformatore in salita, con rapporto di trasformazione pari a Q_0 . Questa proprietà si vedrà che può essere utilizzata per ottenere un adattamento di impedenza in una piccola banda di frequenze (vedi par. 16.7).

In presenza del picco il modulo di H_l deve presentare un massimo, ovvero il suo inverso un minimo, e quindi

$$\frac{d}{d\omega} \left[\left(1 - \frac{\omega^2}{\omega_0^2} \right)^2 + \frac{\omega^2}{\omega_0^2 Q_0^2} \right] = 0 \quad (1.36)$$

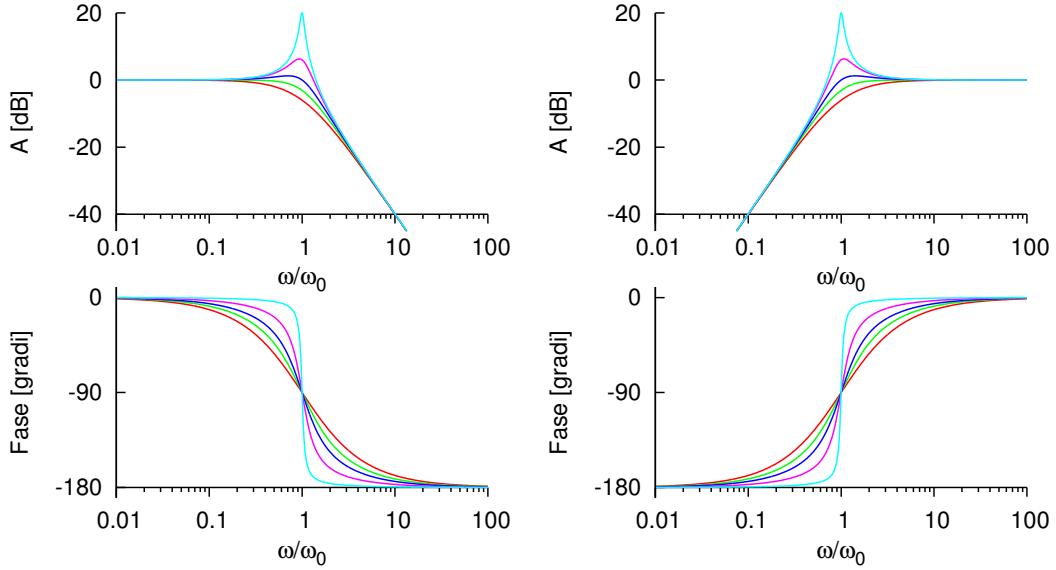


Figura 1.13: Risposte in frequenza per il filtro LRC passa basso (a sinistra) e passa alto (a destra) per i valori di Q : 0.5, 0.7, 1, 2, 10, in ordine dalla curva rossa alla celeste.

Questa equazione ha due soluzioni:

$$\omega_1 = 0 \quad \omega_2 = \omega_0 \sqrt{1 - \frac{1}{2Q_0^2}} \quad (1.37)$$

La prima soluzione, $\omega_1 = 0$, corrisponde alla risposta costante eguale ad 1 nel limite $\omega \rightarrow 0$; la seconda, ω_2 , corrisponde alla posizione del picco, che può esistere solo per $Q_0 \geq 1/\sqrt{2}$. Nel caso limite $Q_0 = 1/\sqrt{2}$ si ha $\omega_2 = 0$ ed il filtro presenta la risposta *massimamente piatta* in funzione della frequenza (curva verde nel grafico).

Il valore di Q_0 per la risposta massimamente piatta nel dominio delle frequenze non coincide con quello della condizione di smorzamento critico visto nel paragrafo precedente, che si ha per $Q_0 = 1/2$ (curva rossa nel grafico).

Passa alto

Non è necessario svolgere altri calcoli per studiare il comportamento del filtro passa alto. Osservando le eq. 1.33 e 1.35 si vede che vale la relazione

$$H_h\left(\frac{\omega_0}{\omega}\right) = H_l^*\left(\frac{\omega}{\omega_0}\right) \quad (1.38)$$

cioè la risposta del filtro passa alto è la complessa coniugata del passa basso con l'asse delle frequenze nel diagramma di Bode rovesciato, come si vede in fig. 1.13. Si possono quindi riutilizzare tutti i risultati già ottenuti e le osservazioni già fatte, compresa la condizione di risposta massimamente piatta.

Passa banda

La risposta del filtro passa banda è quella del filtro selettivo risonante riportata nella fig. 1.14 e presenta un picco per $\omega = \omega_0$, con una attenuazione sui fianchi per $\omega > \omega_0$ oppure $\omega < \omega_0$ tanto più ripida quanto più il valore di Q_0 è elevato. Si può

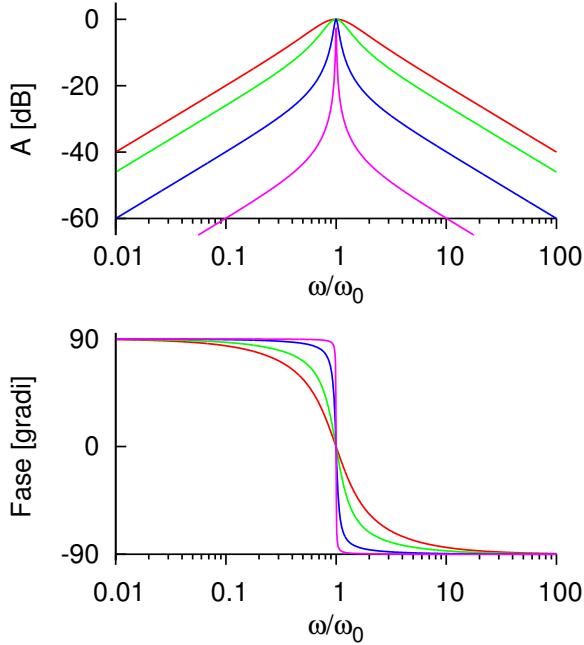


Figura 1.14: Risposte in frequenza per il filtro *LRC* passa banda per i valori di Q : 1, 2, 10, 100, in ordine dalla curva rossa alla viola. Questi grafici a parità di Q rappresentano gli stessi andamenti di quelli in fig. 1.11. L'apparente differenza è dovuta alla diversa scala, lineare in questo caso, logaritmica nell'altro.

verificare facilmente per sostituzione diretta nella eq. 1.34 che alle due frequenze

$$\omega_{1,2} = \omega_0 \left(\sqrt{\frac{1}{4Q_0^2} + 1} \pm \frac{1}{2Q_0} \right) \quad (1.39)$$

il modulo di H_b si è ridotto di un fattore $\sqrt{2}$, cioè di 3 dB, e la fase presenta una rotazione di $\pm 45^\circ$. La *banda passante a 3 dB* del filtro è quindi

$$B_w = \omega_1 - \omega_2 = \frac{\omega_0}{Q_0} \quad (1.40)$$

Le due frequenze ω_1 ed ω_2 sono simmetriche rispetto ad ω_0 sulla scala logaritmica del diagramma di Bode, cioè $\omega_1/\omega_0 = \omega_0/\omega_2$.

1.8.1 Circuito risonante parallelo - circuiti duali

In analogia con il circuito risonante di fig. 1.10 formato dai tre elementi L , R , e C in serie, esiste il circuito risonante *parallelo* di fig. 1.15 il cui funzionamento è descritto dall'equazione integro-differenziale

$$C \frac{dv}{dt} + Gv + \frac{1}{L} \int v dt = i(t) \quad (1.41)$$

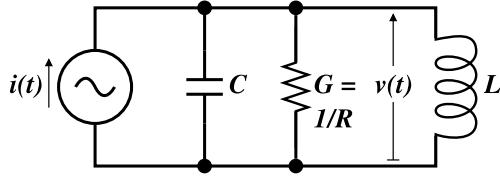


Figura 1.15: Circuito LRC parallelo, duale del circuito di fig. 1.10 e 1.12.

Questa equazione è formalmente identica alla 1.18 una volta che si siano scambiati i ruoli di nodi e maglie, serie e parallelo, di L con C e C con L , di R con la sua inversa $G = 1/R$, della tensione v con la corrente i e viceversa. Il generatore di tensione di fig. 1.10 diventa un generatore di corrente ed il collegamento di tutti gli elementi in serie si trasforma in un parallelo. Questo è un esempio di *rete duale*: per ogni rete elettrica esiste una sua duale che si ottiene con le sostituzioni elencate prima.

Poichè l'equazione che descrive il funzionamento di una rete e della sua duale è formalmente la stessa, tutte le soluzioni trovate per una sono valide anche per l'altra, con le opportune sostituzioni di variabili. In particolare per la rete di fig. 1.15 si avrà ancora $\omega_0 = 1/\sqrt{LC}$ mentre le espressioni per Q_0 ed α , analoghe alle eq. 1.25 e 1.27 diventano

$$Q_0 = \frac{\omega_0 C}{G} = \frac{1}{\omega_0 GL} = \frac{1}{G} \sqrt{\frac{C}{L}} \quad \alpha = -\frac{\omega_0}{2Q_0} = -\frac{1}{2RC} \quad (1.42)$$

L'espressione 1.32 sarà l'ammettenza del circuito, con G in sostituzione di R ed il grafico 1.11 indicherà invece l'andamento dell'impedenza.

Le tre risposte in frequenza delle eq. 1.33-1.35 diventano

$$H_l(\omega) = \frac{I_L}{I_G} \quad H_b(\omega) = \frac{I_R}{I_G} \quad H_h(\omega) = \frac{I_C}{I_G} \quad (1.43)$$

dove I_G è la corrente erogata dal generatore e I_L , I_R ed I_C le tre correnti attraverso rispettivamente induttanza, resistenza e condensatore.

Le espressioni per le tre funzioni H nelle 1.43 in funzione di ω_0 e Q_0 sono le stesse delle eq. 1.33-1.35.

1.8.2 Circuiti risonanti reali

Nella pratica si ha spesso a che fare con il circuito risonante di fig. 1.16a o con il suo duale di fig. 1.16b, poichè capacità ed induttanze reali portano entrambe un contributo dissipativo.

Per il circuito 1.16a si può scrivere per la ammettenza:

$$Y(\omega) = j\omega C + G + \frac{1}{R + j\omega L} = j\omega C + G + \frac{R - j\omega L}{\omega^2 L^2 \cdot \left(1 + \frac{R^2}{\omega^2 L^2}\right)} \quad (1.44)$$

Raccogliendo parte reale ed immaginaria e sostituendo $Q = \omega L/R$, si ottiene

$$Y(\omega) = j \left[\omega C - \frac{1}{\omega L (1 + 1/Q^2)} \right] + G + \frac{R}{\omega^2 L^2 (1 + 1/Q^2)} \quad (1.45)$$

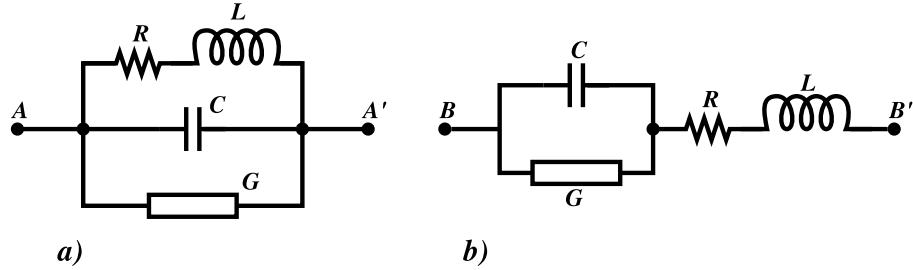


Figura 1.16: Circuiti risonanti parallelo (a) e serie (b) con contributo dissipativo sia da parte della resistenza R in serie all'induttanza L che della conduttanza G in parallelo alla capacità C .

Indicando con $\omega_0 = 1/\sqrt{LC}$ la frequenza di risonanza in assenza di perdite (cioè, con $R = 0, G = 0$), per valori di ω vicini ad ω_0 si ha $Q \simeq Q_0 = \omega_0 L/R$.

Sostituendo Q_0 per Q , dall'azzeramento della parte immaginaria si ricava la frequenza di risonanza in presenza delle perdite dovute ad R e G .

Con l'approssimazione $Q_0 \gg 1$, si ottiene

$$\omega'_0 = \frac{\omega_0}{\sqrt{1 + 1/Q^2}} \simeq \omega_0 \left(1 - \frac{1}{2Q_0^2}\right) \quad (1.46)$$

Dalla parte reale si ricava la conduttanza equivalente presente tra i terminali A-A' alla frequenza di risonanza

$$Y(\omega'_0) = G + \frac{R}{\omega^2 R^2 (1 + 1/Q^2)} \simeq G + \frac{1}{R} \cdot \frac{1}{Q_0^2 + 1} \simeq G + \frac{1}{Q_0^2 R} \quad (1.47)$$

Anche con $G = 0$, il circuito di fig. 1.16a presenta una conduttanza $Y(\omega'_0) = G'$ equivalente ad una resistenza $R' = R \cdot Q_0^2$.

Analogamente, per il principio di dualità, il circuito di fig. 1.16b presenta, con $R = 0$, una resistenza serie R' equivalente ad una conduttanza $G' = G \cdot Q_0^2$ (in questo caso con $Q_0 = \omega_0 C/G$).

Poichè il fattore di merito Q_0 compare nelle equazioni alla seconda potenza, già con $Q_0 = 10$ le approssimazioni introdotte hanno una influenza sui risultati finali dell'ordine dell'1%.

Capitolo 2

Linee di trasmissione

2.1 Circuiti a costanti concentrate

Nei circuiti a *costanti concentrate* le proprietà elettriche del circuito (resistenza, induttanza, capacità ecc.) si considerano tutte contenute nei relativi componenti circolari; i collegamenti tra un componente e l'altro si suppongono fatti con conduttori ideali, cioè privi di resistenza, induttanza o capacità.

Questa schematizzazione è valida fino a che le dimensioni geometriche dei componenti e le lunghezze dei collegamenti rimangono piccole rispetto alle lunghezze d'onda associate ai segnali presenti nel circuito, come definito più avanti.

Oppure, in termini equivalenti, fino a che il tempo di propagazione dei segnali tra due punti del circuito è piccolo rispetto ai tempi di variazione dei segnali stessi.

2.2 Linee di trasmissione

Una linea di trasmissione è invece un sistema di conduttori idealmente paralleli con sezione e caratteristiche costanti su tutta la lunghezza, le cui proprietà elettriche rilevanti sono proprio quelle distribuite sulla lunghezza.

Due esempi di linee di trasmissione molto diffuse e utilizzate sono la linea bifilare ed il cavo coassiale:

1) linea bifilare (o *piattina*): due conduttori rettilinei, di solito a sezione cilindrica, tenuti separati e paralleli da un supporto isolante (fig. 2.1).

2) cavo coassiale: un conduttore centrale cilindrico che corre coassiale ad un secondo conduttore tubolare esterno, detto *schermo*; lo spazio tra i due conduttori è riempito da un materiale isolante (tipicamente polietilene o teflon) (fig. 2.2).

2.3 Circuito equivalente di una linea di trasmissione

Ogni linea di trasmissione può essere considerata composta da infiniti elementi di lunghezza infinitesima Δx , ciascuno equivalente ad un circuito contenente una induttanza, capacità, resistenza e conduttanza infinitesime, come illustrato in fig. 2.3. I simboli L , C , R , G , Z e Y indicano in questa figura (ed in questo capitolo) rispettivamente induttanza, capacità, resistenza, conduttanza, impedenza ed ammettenza per unità di lunghezza del sistema di conduttori che costituiscono la linea.



Figura 2.1: Linea bifilare.



Figura 2.2: Cavo coassiale.

In alternativa, l'impedenza $Z\Delta x$ propria di ognuno degli elementi infinitesimi della linea può essere associata per metà ($Z\Delta x/2$) al conduttore superiore e per l'altra metà al conduttore inferiore (fig. 2.4). I calcoli svolti nei paragrafi seguenti sono identici e portano agli stessi risultati in entrambi i casi. La situazione riportata in fig. 2.3 corrisponde meglio al caso del cavo coassiale, almeno nel caso del cavo ideale, in cui i campi elettrici e magnetici sono confinati completamente all'interno della cavità cilindrica del conduttore di schermo: questo conduttore rimane esternamente una superficie equipotenziale ed è possibile collegarne insieme le due estremità senza influire in alcun modo sul comportamento dei segnali. La descrizione alternativa rispecchia invece la simmetria della linea di trasmissione bifilare e fa risaltare il fatto che il campo elettrico è distribuito sulla lunghezza di entrambi i conduttori. Non è possibile collegare insieme le due estremità di uno dei conduttori di una linea bifilare senza alterarne il funzionamento.

2.3.1 Cavo coassiale

Nel caso del cavo coassiale L e C possono essere calcolati con relativa facilità:¹

¹A rigore l'espressione per L vale solo per il caso di due conduttori tubolari coassiali di spessore molto sottile.

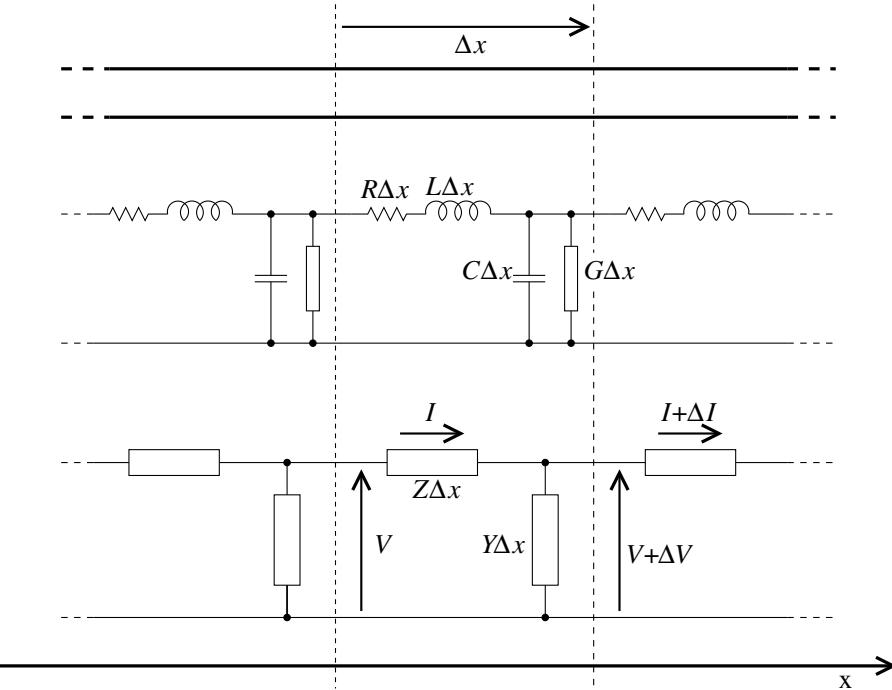


Figura 2.3: Circuito equivalente di una linea di trasmissione. L , C , R , G sono induttanza, capacità, resistenza e conduttanza per unità di lunghezza. $Z = R + j\omega L$ e $Y = G + j\omega C$ sono impedenza e ammettenza per unità di lunghezza. La corrente I è positiva quando scorre verso destra nel conduttore superiore; la tensione V è quella del conduttore superiore rispetto a quello inferiore.

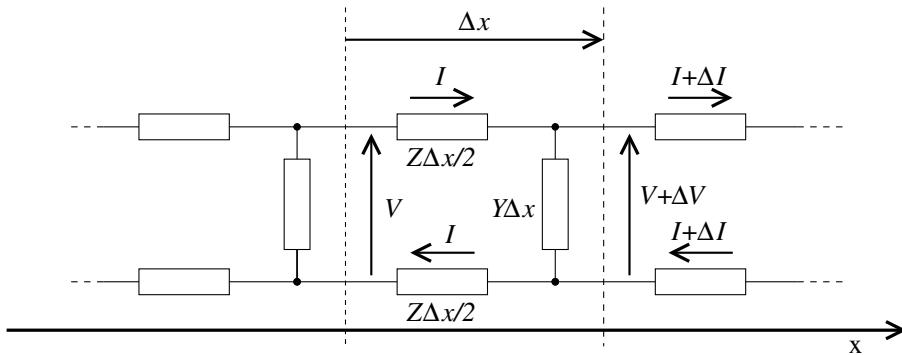


Figura 2.4: Descrizione alternativa delle costanti distribuite della linea di fig. 2.3, più aderente alla geometria della linea bifilare. Le equazioni che descrivono la propagazione dei segnali lungo la linea sono le stesse in entrambi i casi.

$$L = \mu \cdot \frac{\log(D_2/D_1)}{2\pi} \quad C = \epsilon \cdot \frac{2\pi}{\log(D_2/D_1)} \quad (2.1)$$

Le quantità D_1 e D_2 rappresentano rispettivamente il raggio esterno del conduttore interno ed il raggio interno del conduttore esterno; ϵ e μ sono la permeabilità dielettrica e magnetica del materiale isolante che riempie lo spazio tra i due conduttori.

Nel caso di una linea con geometria diversa da quella del cavo coassiale, lo schema di figura 2.3 continuerà ad essere valido, ma i valori dei parametri L e C saranno dati da espressioni diverse dalle eq. 2.1.

2.4 Correnti e tensioni lungo la linea

Per scrivere le equazioni per le tensioni e le correnti lungo la linea si utilizzano le convenzioni riportate in fig. 2.3: la coordinata x individua la posizione lungo la linea ed ha direzione positiva verso destra; una corrente è positiva quando scorre verso destra nel conduttore superiore (e verso sinistra in quello inferiore); le tensioni indicano il potenziale del conduttore superiore rispetto a quello inferiore.

Con queste convenzioni, applicando la legge di Ohm generalizzata ai circuiti in corrente alternata, si possono scrivere per ciascuna cella infinitesima che compone la linea due equazioni:

$$\frac{\partial V}{\partial x} = -ZI \quad \frac{\partial I}{\partial x} = -YV \quad (2.2)$$

dove $Z = R + j\omega L$ e $Y = G + j\omega C$ ². V ed I sono due quantità complesse che rappresentano ampiezza e fase di tensione e corrente lungo la linea per una componente alternata di frequenza f e frequenza angolare $\omega = 2\pi f$.

Derivando entrambe le equazioni rispetto alla x e sostituendo si ottiene

$$\frac{\partial^2 V}{\partial x^2} = -Z \frac{\partial I}{\partial x} \quad \frac{\partial^2 I}{\partial x^2} = -Y \frac{\partial V}{\partial x} \quad (2.3)$$

$$\frac{\partial^2 V}{\partial x^2} - ZYV = 0 \quad \frac{\partial^2 I}{\partial x^2} - ZYI = 0 \quad (2.4)$$

Le (2.4) sono equazioni differenziali lineari del secondo ordine ed hanno soluzioni:

$$V(x) = V_1 e^{+\gamma x} + V_2 e^{-\gamma x} \quad (2.5)$$

$$I(x) = I_1 e^{+\gamma x} + I_2 e^{-\gamma x} \quad (2.6)$$

La quantità complessa

$$\gamma = \alpha + j\beta = \sqrt{ZY} = \sqrt{(R + j\omega L)(G + j\omega C)} \quad (2.7)$$

è detta *costante di propagazione* o *coefficiente di propagazione*; la parte reale α (*neper/metro*) è la *costante di attenuazione* e misura la perdita di segnale lungo la linea a causa degli elementi dissipativi R e G ; la parte immaginaria β (*radianti/metro*) è la costante di fase e misura la differenza di fase tra punti diversi della linea ad un medesimo istante di tempo; $V_{1,2}, I_{1,2}$ sono costanti determinate dalle condizioni alle estremità della linea.

²Come già notato, in elettrotecnica ed in elettronica l'unità immaginaria è indicata di solito con la lettera j , in quanto la i è usata di preferenza per indicare la corrente.

2.5 Linea senza perdite

Nel caso ideale di una linea senza perdite si avrà $R = 0$ e $G = 0$ e quindi

$$\gamma = j\beta = j\omega\sqrt{LC} \quad (2.8)$$

L'espressione completa della tensione lungo la linea, inserendo esplicitamente anche la dipendenza dal tempo, è:

$$V(x, t) = V_1 e^{j(\omega t + \beta x)} + V_2 e^{j(\omega t - \beta x)} \quad (2.9)$$

Analogamente per la corrente:

$$I(x, t) = I_1 e^{j(\omega t + \beta x)} + I_2 e^{j(\omega t - \beta x)} \quad (2.10)$$

Queste due espressioni rappresentano la tensione e la corrente di due segnali con andamento sinusoidale in funzione del tempo e dello spazio, che si propagano lungo la linea. Dalla condizione $\omega t \pm \beta x = \text{costante}$, si ottiene la velocità

$$u = \frac{dx}{dt} = \mp \frac{\omega}{\beta} = \mp \frac{1}{\sqrt{LC}} \quad (2.11)$$

Il primo termine (V_1, I_1), con velocità negativa, prende il nome di onda regressiva, o riflessa, e si propaga da destra verso sinistra; il secondo termine (V_2, I_2), con velocità positiva, prende il nome di onda progressiva, o diretta e si propaga da sinistra verso destra.

Sostituendo nell'eq. 2.11 il valore di L e C per il cavo coassiale (eq. 2.1) si ottiene:

$$\beta = \omega \sqrt{\mu\epsilon} \quad (2.12)$$

da cui per una linea ideale senza perdite

$$|u| = \frac{\omega}{\beta} = \frac{1}{\sqrt{\mu\epsilon}} \quad (2.13)$$

cioè la stessa velocità che avrebbe un'onda elettromagnetica che si propagasse nel mezzo dielettrico che riempie lo spazio tra i due conduttori.

In un dielettrico ideale μ ed ϵ sono costanti, in particolare non dipendono da ω . In questo caso, in un segnale composto dalla sovrapposizione di più componenti di frequenze diverse, ciascuna componente avrà la stessa velocità u , dando luogo ad una propagazione senza *dispersione*. In queste condizioni *velocità di fase* e *velocità di gruppo* (vedi par. 2.17) coincidono e la velocità data dalla (2.13) sarà valida per qualsiasi tipo di segnale.

Nel caso che il materiale isolante sia l'aria o il vuoto:

$$\beta = \omega \sqrt{\mu_0 \epsilon_0} = \frac{\omega}{c} \quad (2.14)$$

dove μ_0 , ϵ_0 e c sono la permeabilità dielettrica e magnetica e la velocità della luce nel vuoto. È possibile realizzare un tale cavo inserendo tra i due conduttori coassiali una striscia elicoidale di materiale isolante di spessore molto sottile, in modo che il dielettrico sia formato praticamente quasi solo da aria (fig. 2.5).



Figura 2.5: Cavo coassiale per grosse potenze con aria come dielettrico. La striscia elicoidale di polietilene mantiene coassiali i due cilindri conduttori in rame, occupando solo una minima parte del volume. Alla frequenza di 2 MHz il cavo può trasportare una potenza di circa 300 kW. Diametro del conduttore interno: 32 mm; diametro del conduttore esterno: 76 mm. La moneta appoggiata sul bordo per evidenziare le dimensioni ha un diametro di 26 mm. Resistenza caratteristica: 50 Ω.

2.6 Lunghezza d'onda, lunghezza elettrica

Un segnale con andamento sinusoidale si ripete lungo la linea con periodicità $\beta x = 2\pi$. La quantità $\lambda = 2\pi/\beta = 2\pi u/\omega$ prende il nome di lunghezza d'onda.

Per un elemento circuitale di lunghezza $l \ll \lambda$ si potranno considerare le tensioni e le correnti uniformi su tutto l'elemento e ignorare le variazioni tra un punto ed un altro. Si potrà cioè considerare l'elemento come *puntiforme* e supporre le proprietà elettriche *concentrate* tutte nel punto. La condizione $l \ll \lambda$ equivale alla condizione $t \ll \tau$, dove $\tau = 2\pi/\omega$ è il periodo del segnale e $t = l/u$ è il tempo necessario a percorrere la lunghezza l : un circuito può essere considerato a *costanti concentrate* fino a che il tempo di propagazione dei segnali da un punto ad un altro è trascurabile rispetto ai periodi propri delle componenti di frequenza più alta nei segnali stessi.

Con il termine *lunghezza elettrica* di una linea si intende la lunghezza di una linea avente il vuoto come dielettrico e con lo stesso tempo di propagazione dei segnali della linea data: $\text{lunghezza elettrica} = c \cdot \text{lunghezza geometrica}/u$. Spesso, con lo stesso termine *lunghezza elettrica* si indica direttamente il tempo di propagazione $t = \text{lunghezza geometrica}/u$ di un segnale sulla lunghezza della linea. Il contesto (e le unità di misura) evitano qualunque possibilità di fraintendimento.

2.7 Linea con perdite. Condizioni di Heaviside

Nel caso di una linea con perdite ($R, G \neq 0$) si ha:

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} = \alpha + j\beta \quad (2.15)$$

In generale questa condizione comporta che la velocità di propagazione sia funzione della frequenza e quindi che si abbia dispersione. Nel caso particolare che le due quantità complesse $R + j\omega L$ e $G + j\omega C$ abbiano lo stesso argomento, cioè si abbia

$L/R = C/G$ (condizioni di Heaviside), si ottiene $\alpha = \sqrt{RG}$ e $\beta = \omega\sqrt{LC}$ e quindi ancora velocità costante e propagazione non dispersiva, anche se si tratta di una linea non ideale.

Il termine α descrive l'attenuazione del segnale lungo la linea per effetto delle perdite:

$$V(x, t) = V_1 e^{\alpha x} e^{j(\omega t + \beta x)} + V_2 e^{-\alpha x} e^{j(\omega t - \beta x)} \quad (2.16)$$

Il primo termine rappresenta un segnale che si propaga da destra verso sinistra e che si attenua quindi procedendo verso la direzione negativa delle x . Indicando con V_0 l'ampiezza del segnale nell'origine, si ha rispettivamente, per le componenti di onda regressiva e progressiva:

$$\left| \frac{V(x)}{V_0} \right| = e^{\pm \alpha x} \quad (2.17)$$

Volendo esprimere l'attenuazione A in decibel:

$$A = 20 \log_{10} (e^{\pm \alpha x}) = \pm 20 \alpha x \log_{10}(e) \simeq \pm 8.7 \alpha x \quad (2.18)$$

Se x ed $1/\alpha$ sono espressi in metri, la quantità 8.7α rappresenta l'attenuazione della linea di trasmissione in dB/m (decibel per metro).

2.8 Impedenza caratteristica della linea

Dalle eq. (2.5) e (2.6) si può calcolare il rapporto tra tensione e corrente in ogni punto della linea:

$$\frac{V(x)}{I(x)} = \frac{V_1 e^{\gamma x} + V_2 e^{-\gamma x}}{I_1 e^{\gamma x} + I_2 e^{-\gamma x}} \quad (2.19)$$

I valori di I_1 e I_2 si ottengono dalla prima delle eq. (2.2):

$$I(x) = -\frac{1}{Z} \frac{\partial V(x)}{\partial x} = -\frac{\gamma}{Z} V_1 e^{+\gamma x} + \frac{\gamma}{Z} V_2 e^{-\gamma x} \quad (2.20)$$

e confrontando questo risultato con l'eq. 2.6 (tenendo anche conto della 2.7):

$$I_1 = -\sqrt{\frac{Y}{Z}} V_1 \quad I_2 = \sqrt{\frac{Y}{Z}} V_2 \quad (2.21)$$

Sostituendo questi coefficienti nella (2.19):

$$\frac{V(x)}{I(x)} = \sqrt{\frac{Z}{Y}} \cdot \frac{V_1 e^{\gamma x} + V_2 e^{-\gamma x}}{-V_1 e^{\gamma x} + V_2 e^{-\gamma x}} \quad (2.22)$$

Oppure, nel caso di una linea senza perdite,

$$\frac{V(x)}{I(x)} = \sqrt{\frac{Z}{Y}} \cdot \frac{V_1 e^{j\beta x} + V_2 e^{-j\beta x}}{-V_1 e^{j\beta x} + V_2 e^{-j\beta x}} \quad (2.23)$$

Nel caso sulla linea sia presente una sola onda, ad esempio l'onda diretta, il rapporto tra tensione e corrente è:

$$\frac{V_2(x)}{I_2(x)} = \sqrt{\frac{Z}{Y}} = Z_0 \quad (2.24)$$

Questo rapporto è costante su tutta la linea, ha le dimensioni di una resistenza e prende il nome di *resistenza* o *impedenza caratteristica* della linea. Nel caso di linea ideale senza perdite, o di linea con perdite ma nelle condizioni di Heaviside, Z_0 è reale, cioè è una resistenza pura senza componente reattiva e non dipende dalla frequenza:

$$Z_0 = \sqrt{\frac{L}{C}} = R_0 \quad (2.25)$$

Facendo il rapporto tra tensione e corrente per un'onda riflessa, si ottiene:

$$\frac{V_1(x)}{I_1(x)} = -\sqrt{\frac{Z}{Y}} = -Z_0 \quad (2.26)$$

Il segno meno in questo risultato indica che ad uno stesso valore di tensione in un generico punto lungo la linea è associata, nel caso di un'onda riflessa, una corrente di valore opposto rispetto a quella di un'onda diretta.

Per il cavo coassiale

$$R_0 = \sqrt{\frac{L}{C}} = \frac{1}{2\pi} \sqrt{\frac{\mu_0 \mu_R}{\epsilon_0 \epsilon_R}} \cdot \log \frac{D_2}{D_1} \simeq \sqrt{\frac{\mu_R}{\epsilon_R}} \cdot \log \frac{D_2}{D_1} \cdot 60\Omega \quad (2.27)$$

Per i cavi coassiali comunemente in uso nella strumentazione di laboratorio ed in radiotecnica $\log(D_2/D_1)$ è dell'ordine dell'unità, $\mu_R = 1$ ed $\epsilon_R \simeq 2$ (isolante polietilene o teflon). I valori di R_0 che si ottengono sono dell'ordine di $50 \dots 100 \Omega$. Per ottenere valori di R_0 di un ordine di grandezza diverso, a causa della dipendenza logaritmica che compare nella eq. 2.27, occorrerebbero valori di D_2/D_1 di difficile realizzazione pratica.

2.9 Linea di trasmissione chiusa sulla sua resistenza caratteristica

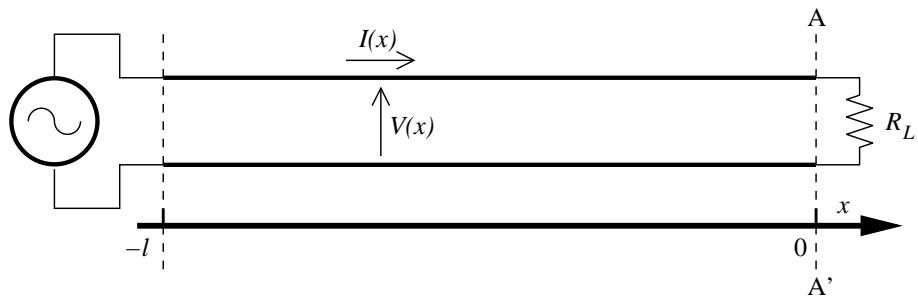


Figura 2.6: Linea terminata sulla sua resistenza caratteristica.

Nella fig. 2.6 è riportato un tratto di linea di trasmissione alimentata a sinistra da un generatore e collegata a destra ad una resistenza (*terminazione*) di valore $R_L = R_0$. Durante la propagazione del segnale inviato dal generatore, considerando la sola onda diretta, si ha in ogni punto della linea $V(x)/I(x) = R_0$; questo vale anche alla terminazione $A - A'$ della linea, dove però deve valere anche l'ulteriore

condizione $V/I = R_L$ (che però è ancora eguale ad R_0). Si vede quindi che le condizioni al contorno imposte dalla resistenza di terminazione sono soddisfatte da una soluzione composta dalla sola onda diretta.

2.10 Coefficiente di riflessione.

Riferendosi sempre alla fig. 2.6, supponiamo di avere $R_L \neq R_0$; in presenza della sola onda diretta il rapporto V/I vale R_0 in tutti i punti della linea a sinistra e fino ad $A - A'$, mentre deve valere R_L in $A - A'$. La presenza della sola onda diretta non è quindi in grado in questo caso di soddisfare alle condizioni al contorno imposte dalla resistenza di terminazione $R_L \neq R_0$; il segnale che si propaga da sinistra verso destra non può essere interamente assorbito dalla resistenza R_L , ma viene in parte *riflesso*. In presenza di entrambe le onde, diretta e riflessa, il rapporto V/I è quello dato dalla eq. 2.23:

$$\frac{V(x)}{I(x)} = R_0 \frac{V_1 e^{j\beta x} + V_2 e^{-j\beta x}}{-V_1 e^{j\beta x} + V_2 e^{-j\beta x}} \quad (2.28)$$

Conviene porre l'origine della coordinata x in $A - A'$; il generatore verrà ad avere coordinata $x = -l$, con l lunghezza della linea. Quindi, in $A - A'$:

$$\frac{V(0)}{I(0)} = R_0 \frac{V_1 + V_2}{-V_1 + V_2} = R_L \quad (2.29)$$

Da questa

$$\frac{V_1}{V_2} = \frac{R_L - R_0}{R_L + R_0} = \rho_v \quad (2.30)$$

La quantità ρ_v prende il nome di coefficiente di riflessione di tensione ed esprime il rapporto tra l'ampiezza dell'onda riflessa da una terminazione e l'ampiezza dell'onda diretta.

Viceversa, dalla conoscenza del coefficiente di riflessione ρ_v è possibile determinare il rapporto

$$\frac{R_L}{R_0} = \frac{1 + \rho_v}{1 - \rho_v} \quad (2.31)$$

Nel caso più generale che la terminazione sia una impedenza complessa Z_L invece di una resistenza pura R_L , anche il coefficiente di riflessione ρ_v sarà una quantità complessa.

2.11 Onde stazionarie.

La sovrapposizione lungo la linea delle onde diretta e riflessa da origine ad un fenomeno di interferenza. Supponendo che la linea abbia perdite nulle, o comunque trascurabili ($\alpha \simeq 0$):

$$V(x, t) = V_1 e^{j(\omega t + \beta x)} + V_2 e^{j(\omega t - \beta x)} \quad (2.32)$$

Spostandosi lungo la linea, ad esempio verso destra, l'onda diretta arriva con un ritardo di fase via via maggiore, mentre la fase dell'onda riflessa risulta anticipata: si attraversano una serie di regioni in cui le due onde si sovrappongono

dando luogo alternativamente ad interferenza costruttiva e distruttiva. Per vedere quantitativamente questo effetto, calcoliamo

$$|V(x, t)|^2 = |e^{j\omega t}|^2 \cdot (V_1 e^{j\beta x} + V_2 e^{-j\beta x}) \cdot (V_1 e^{j\beta x} + V_2 e^{-j\beta x})^* \quad (2.33)$$

Scrivendo V_1 e V_2 in notazione polare

$$V_1 = |V_1| e^{j\varphi}, \quad V_2 = |V_2| e^{j\psi} \quad (2.34)$$

si ottiene:

$$\begin{aligned} |V(x, t)|^2 &= \left(|V_1| e^{j(\beta x + \varphi)} + |V_2| e^{-j(\beta x + \psi)} \right) \\ &\quad \cdot \left(|V_1| e^{-j(\beta x + \varphi)} + |V_2| e^{j(\beta x + \psi)} \right) \\ &= |V_1|^2 + |V_2|^2 + 2|V_1| \cdot |V_2| \cdot \cos(2\beta x + \varphi + \psi) \end{aligned} \quad (2.35)$$

Si vede che il modulo di $V(x, t)$ oscilla tra un massimo $V_{max} = |V_2| + |V_1|$ ed un minimo $V_{min} = |V_2| - |V_1|$, corrispondenti rispettivamente ad interferenza massimamente costruttiva e distruttiva. Nel caso di riflessione completa $V_{max} = 2|V_2|$ e $V_{min} = 0$. La quantità

$$SWR = \frac{|V_2| + |V_1|}{|V_2| - |V_1|} \quad (2.36)$$

prende il nome di *rapporto di onde stazionarie (Standing Wave Ratio)*.

La distanza tra due massimi e due minimi consecutivi di $|V|$ si ha quando la differenza di fase tra onda diretta e riflessa compie una variazione complessiva di 2π , cioè quando $\Delta x = \pi/\beta = \lambda/2$.

2.11.1 Trasformazione di impedenza.

Per effetto dell'interferenza tra onda diretta ed onda riflessa descritta al par. 2.11 il rapporto tra tensione e corrente varia da punto a punto lungo la linea. Dalla eq. 2.23 si ha:

$$Z(x) = \frac{V(x)}{I(x)} = R_0 \frac{V_1 e^{j\beta x} + V_2 e^{-j\beta x}}{-V_1 e^{j\beta x} + V_2 e^{-j\beta x}} \quad (2.37)$$

e poichè $V_1 = V_2(Z_L - R_0)/(Z_L + R_0)$

$$\begin{aligned} Z(x) &= R_0 \frac{Z_L(e^{j\beta x} + e^{-j\beta x}) - R_0(e^{j\beta x} - e^{-j\beta x})}{-Z_L(e^{j\beta x} - e^{-j\beta x}) + R_0(e^{j\beta x} + e^{-j\beta x})} \\ &= R_0 \frac{Z_L \cos(\beta x) - jR_0 \sin(\beta x)}{-jZ_L \sin(\beta x) + R_0 \cos(\beta x)} \\ &= R_0 \frac{Z_L - jR_0 \tan(\beta x)}{R_0 - jZ_L \tan(\beta x)} \end{aligned} \quad (2.38)$$

Questo risultato ci dice che un generatore, collegato ad un tratto di linea di trasmissione di resistenza caratteristica R_0 e chiuso su una impedenza Z_L , vede una

impedenza $Z_l = Z(-l)$ che dipende dalla lunghezza l della linea. Nel caso di linee con perdite, l'eq. 2.38 diventa

$$Z(x) = R_0 \frac{Z_L(e^{\gamma x} + e^{-\gamma x}) - R_0(e^{\gamma x} - e^{-\gamma x})}{-Z_L(e^{\gamma x} - e^{-\gamma x}) + R_0(e^{\gamma x} + e^{-\gamma x})} = R_0 \frac{Z_L - R_0 \tanh(\gamma x)}{R_0 - Z_L \tanh(\gamma x)} \quad (2.39)$$

Alcuni esempi:

1. Linea chiusa in cortocircuito ($Z_L = 0$): $Z_l = -jR_0 \tan(\beta x) = jR_0 \tan(\beta l)$ in assenza di perdite. L'impedenza è una reattanza pura, che per valori di βl (cioè di l/λ) piccoli è di tipo induttivo; per $\beta l \rightarrow \pi/2$ si ha $Z_l \rightarrow \infty$: il generatore vede attraverso un pezzo di linea di lunghezza $l = \pi/(2\beta) = \lambda/4$ il cortocircuito trasformato in un circuito aperto. Per $l > \lambda/4$ la reattanza diventa negativa, cioè di tipo capacitivo, fino a $l = \lambda/2$, ecc. Questo andamento è riportato in fig. 2.7 (curve in blu), per una linea di lunghezza elettrica $l = 1.5 \text{ m}$, per frequenze da 0 a 300 MHz . La frequenza è riportata sull'asse x come f/f_λ , dove $f_\lambda = 200 \text{ MHz}$ è la frequenza la cui lunghezza d'onda è eguale alla lunghezza l della linea. In rosso è riportata l'impedenza $Z_l = R_0 \tanh(\gamma l)$ della stessa linea, in presenza di una perdita di 0.2 dB/m . In questo caso la parte reale non è più identicamente nulla ed in corrispondenza di $l = \lambda/4, 3\lambda/4$ ecc. l'impedenza non diverge più a infinito, ma diventa reale e finita.
2. Linea aperta ($Z_L = \infty$): $Z_l = -jR_0 / \tan(\beta l) = -jR_0 \cot(\beta l)$ in assenza di perdite e $Z_l = -R_0 \coth(\gamma l)$ nel caso di linea con perdite. A causa della complementarietà tra le funzioni $\tan()$ e $\cot()$, si hanno gli stessi comportamenti del caso precedente, con una traslazione di $1/4$ sull'asse x . Nelle stesse condizioni dell'esempio precedente risulta ancora valido il grafico di fig. 2.7, con i valori dell'asse x riportati in alto. Quindi per $f \ll f_\lambda$ la linea si comporta come un condensatore; per $f = f_\lambda/4$ la terminazione a circuito aperto viene trasformata in un cortocircuito ecc.
3. Linea senza perdite lunga esattamente $l = \lambda/2$ o un suo multiplo intero e terminata su una impedenza qualsiasi Z_x : $Z_l = Z_x$.

Quando $|\gamma l| \ll 1$, nella eq. 2.39 si può sostituire a $\tanh(\gamma x)$ il suo sviluppo in serie di Taylor arrestato al primo ordine. Si ottiene:

$$\begin{aligned} 1. \quad Z_l &= R_0 \gamma l = \sqrt{\frac{(R + j\omega L)}{(G + j\omega C)}} \cdot \sqrt{(R + j\omega L)(G + j\omega C)} \cdot l = (R + j\omega L) \cdot l \\ 2. \quad Z_l &= \frac{R_0}{\gamma l} = \sqrt{\frac{(R + j\omega L)}{(G + j\omega C)}} \sqrt{\frac{(R + j\omega L)(G + j\omega C)}{l}} = \frac{1}{(G + j\omega C) \cdot l} \end{aligned}$$

Quindi un tratto di linea molto corta equivale ad elementi circuituali *concentrati*: se la linea è aperta si comporta come un condensatore (con una conduttanza in parallelo), se è chiusa in cortocircuito come una induttanza (con una resistenza in serie), con i valori degli elementi propri del tratto di linea considerato.

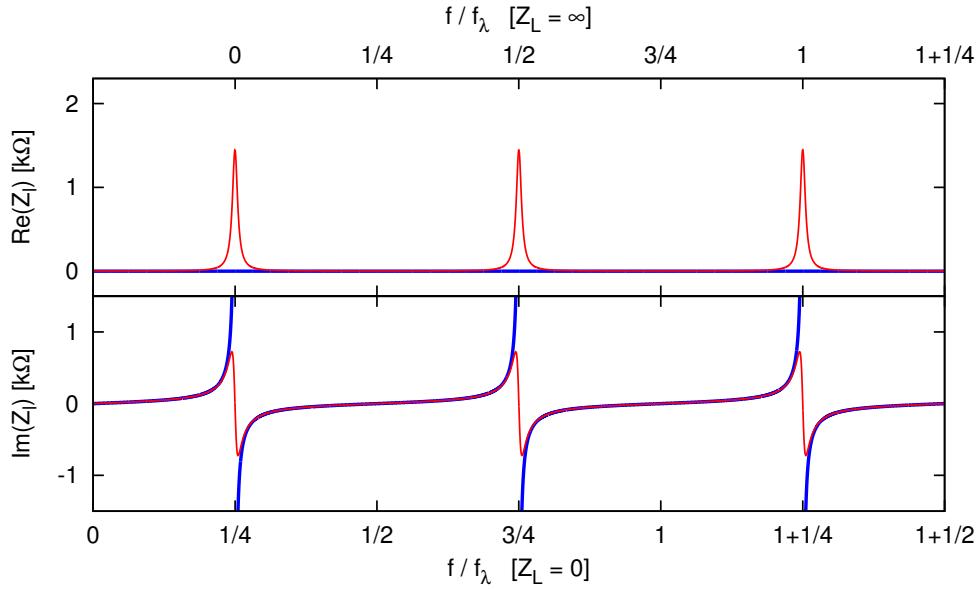


Figura 2.7: Parte reale ed immaginaria dell'impedenza Z_l in funzione della frequenza f per una linea di lunghezza elettrica $l = 1.5 \text{ m}$, con $R_0 = 50 \Omega$ e chiusa in cortocircuito ($Z_L = 0$) oppure terminata a circuito aperto ($Z_L = \infty$). Le curve in blu corrispondono ad una linea senza perdite; le curve in rosso ad una linea con una perdita di $0.2 \text{ dB}/\text{m}$ ($\alpha = 0.023 \text{ m}^{-1}$). La frequenza f_λ è quella frequenza per cui la lunghezza d'onda λ è eguale alla lunghezza l della linea. Le curve per i due casi $Z_L = 0$ e $Z_L = \infty$ differiscono solo per una traslazione di $1/4$ sull'asse x .

2.11.2 Linea di lunghezza variabile

I grafici in fig. 2.7 sono stati ottenuti considerando una linea di lunghezza pre-

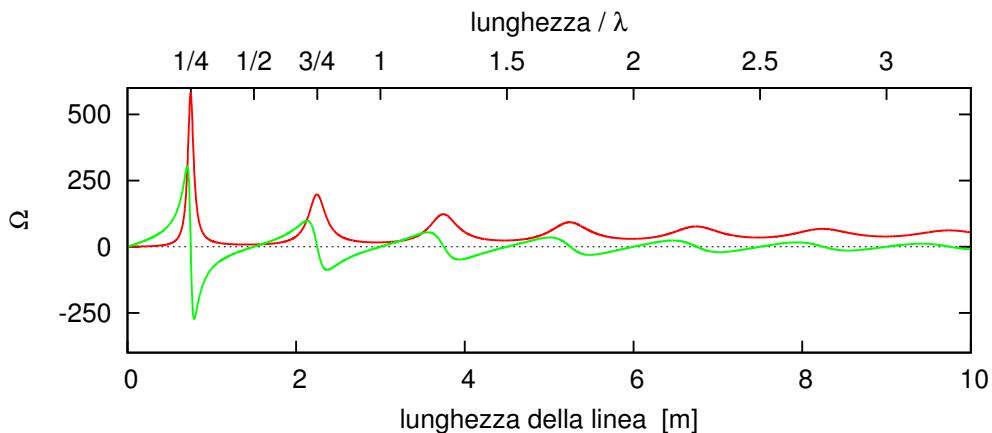


Figura 2.8: Impedenza equivalente di una linea chiusa in cortocircuito, di lunghezza variabile $0 \dots 10 \text{ m}$, con resistenza caratteristica $R_0 = 50 \Omega$ e perdite di $1 \text{ dB}/\text{m}$, alla frequenza $f = 100 \text{ MHz}$. Rosso: parte reale; verde: parte immaginaria.

fissata (1.5 m) ed una frequenza variabile ($0 \dots 300\text{ MHz}$). In fig. 2.8 è riportato il grafico della eq. 2.39 calcolato alla frequenza fissa $f = 100\text{ MHz}$, per una linea di lunghezza variabile $0 \dots 10\text{ m}$ chiusa in cortocircuito e con un coefficiente di perdita di 1.0 dB/m . Anche in questo caso compaiono i picchi in corrispondenza di $f/f_\lambda = 1/4, 3/4 \dots$ ecc. Ma al crescere della lunghezza della linea i picchi si fanno sempre più piccoli a causa dell'aumento delle perdite, fino a che, per $l \rightarrow \infty$ si ha $Z_l \rightarrow R_0$. Un grafico analogo si otterrebbe per la stessa linea aperta, con le differenze già viste per fig. 2.7.

È da notare che nel mondo reale una misura effettuata nelle condizioni di fig. 2.7 produrrebbe un grafico con un andamento analogo a quello di fig. 2.8, perché in una linea reale le perdite tendono sempre ad aumentare al crescere della frequenza.

2.11.3 Tensioni alle estremità della linea

Il calcolo della tensione presente al termine di una linea di trasmissione, conoscendo la tensione applicata all'ingresso e le caratteristiche e la terminazione della linea, richiede di tenere conto del fenomeno delle onde stazionarie.

Le tensioni V_i e V_t , presenti rispettivamente all'inizio ed al termine di una linea di lunghezza l e coefficiente di propagazione γ , risultano dalla somma dell'onda diretta V_2 e dell'onda riflessa V_1 :

$$V_i = V_1(0) + V_2(0) \quad V_t = V_1(l) + V_2(l) \quad (2.40)$$

dove la coordinata 0 corrisponde all'inizio della linea e la coordinata l all'altra estremità. Per effetto della riflessione alla terminazione, secondo la eq. 2.30, vale la relazione

$$V_1(l) = \rho_v V_2(l) \quad (2.41)$$

mentre per effetto della propagazione tra i valori di V_1 e V_2 alle due estremità valgono le relazioni

$$V_2(l) = V_2(0)e^{-\gamma l} \quad V_1(0) = V_1(l)e^{-\gamma l} \quad (2.42)$$

Sostituendo le espressioni 2.41 e 2.42 nelle 2.40 si ottiene

$$\begin{aligned} V_i &= V_2(0) \left(1 + \rho_v e^{-2\gamma l} \right) \\ V_t &= V_2(0) \left(e^{-\gamma l} + \rho_v e^{-\gamma l} \right) \end{aligned} \quad (2.43)$$

Dividendo tra loro le 2.43 e semplificando:

$$\frac{V_t}{V_i} = \frac{1 + \rho_v}{e^{\gamma l} + \rho_v e^{-\gamma l}} \quad (2.44)$$

Per $\rho_v = 0$, cioè per una linea correttamente terminata, si ha:

$$\frac{V_t}{V_i} = e^{-\gamma l} \quad (2.45)$$

cioè la normale propagazione senza riflessione. Per una linea aperta ($\rho_v = 1$):

$$\frac{V_t}{V_i} = \frac{2}{e^{\gamma l} + e^{-\gamma l}} = \frac{2}{\cos(\beta l) (e^{\alpha l} + e^{-\alpha l}) + j \sin(\beta l) (e^{\alpha l} - e^{-\alpha l})} \quad (2.46)$$

Quando βl diventa $\pi/2$, $3\pi/2$, ecc., il termine in coseno a denominatore si annulla e per valori di α piccoli rimane

$$\frac{V_t}{V_i} = \pm \frac{1}{j\alpha l} \quad (2.47)$$

Alimentando una linea aperta con un generatore a bassa resistenza interna, in grado di mantenere una tensione V_i anche su una impedenza che tende ad un cortocircuito (vedi esempio 2 del paragrafo precedente), la tensione alla terminazione della linea può assumere valori molto elevati. Per una linea chiusa in cortocircuito ($\rho_v = -1$)

$$\frac{V_t}{V_i} = 0 \quad (2.48)$$

come è ovvio che sia.

2.12 Segnali a banda larga e a banda stretta.

Quanto detto nei paragrafi precedenti si riferisce a segnali di tipo monocromatico, cioè segnali il cui spettro di Fourier contiene una singola componente di frequenza f (o frequenza angolare $\omega = 2\pi f$). Molti segnali di interesse fisico sono invece caratterizzati da uno spettro di Fourier ricco di componenti aventi frequenze diverse. L'intervallo nello spettro di un segnale compreso tra la componente di frequenza più bassa (f_{min}) e quella di frequenza più alta (f_{max}) costituisce la larghezza di banda (o occupazione di banda) del segnale.

In figura (2.9) sono riportate le bande di frequenza caratteristiche di vari tipi di segnale:

1. Segnale audio: $20\text{Hz} \div 20\text{kHz}$. Larghezza di banda $\simeq 20\text{kHz}$; rapporto $f_{max}/f_{min} = 10^3$
2. Impulsi generati da un fotomoltiplicatore: $20\text{kHz} \div 500\text{MHz}$. Larghezza di banda $\simeq 500\text{MHz}$; rapporto $f_{max}/f_{min} = 2.5 \cdot 10^4$
3. Segnale televisivo in banda UHF, canale 61: $790 \div 798\text{MHz}$. Larghezza di banda $= 8\text{MHz}$; rapporto $f_{max}/f_{min} = 1.01$

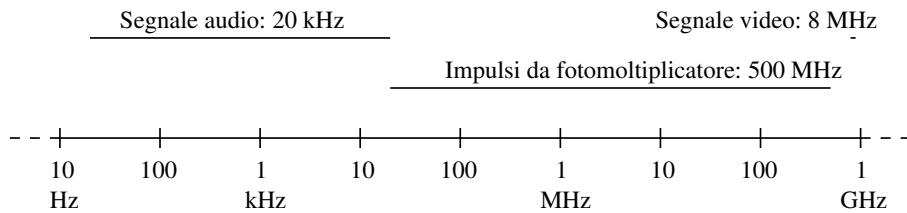


Figura 2.9: Occupazione di banda per vari tipi di segnali.

Il primo ed il secondo di questi esempi rappresentano segnali detti *a banda larga*; il terzo rappresenta un segnale *a banda stretta*. In questa classificazione l'aspetto importante non è la larghezza di banda in Hz o in kHz , ma il rapporto tra frequenza

massima e minima dello spettro del segnale. Un segnale in cui questo rapporto è molto vicino ad uno, come nel terzo esempio, può essere considerato come un segnale per molti aspetti monocromatico. Le equazioni del par. 2.11 si possono applicare al segnale del terzo esempio, a condizione che βx vari solo di poco nella banda di frequenze interessate. Prendendo come valore di x tutta la lunghezza l della linea e ricordando che $\beta = \omega/u$ (eq. 2.13), la condizione richiesta diventa $l\Delta\omega/u \ll 2\pi$, cioè $\Delta\omega \ll 2\pi u/l = \omega\lambda/l$ e quindi

$$\frac{\Delta\omega}{\omega} = \frac{\Delta f}{f} \ll \frac{\lambda}{l} \quad (2.49)$$

Se si considera una linea lunga anche solo poche lunghezze d'onda, si vede che solo il segnale riportato nel terzo esempio può soddisfare alla condizione richiesta. Le equazioni del par. 2.11 applicate ai segnali degli esempi (1) e (2) danno risultati che sono funzione di ω , cioè diversi per ogni componente dello spettro e in genere di scarsa utilità.

2.13 Riflessione di un segnale a banda larga

Consideriamo come segnale a banda larga un breve impulso (fig. 2.10). Se l'im-



Figura 2.10: Impulso di durata τ

pulso si ripete nel tempo con cadenza f il suo spettro di Fourier contiene la frequenza fondamentale f e le armoniche $2f, 3f$ ecc., di intensità via via decrescente. Nel caso limite di un impulso infinitamente stretto, cioè di una δ di Dirac, tutte le frequenze armoniche hanno la stessa ampiezza. Riducendo via via la frequenza f le componenti armoniche dello spettro diventano sempre più fitte e nel caso limite di una singola δ di Dirac, corrispondente a $f = 0$, si ha uno spettro continuo che contiene tutte le frequenze, tutte con la stessa ampiezza³.

Supponendo che la linea sia ideale o almeno soddisfi alle condizioni di Heaviside, la propagazione avverrà senza dispersione e quindi tutte le componenti armoniche si propagheranno con la stessa velocità u ; anche l'impulso quindi si propagherà lungo la linea con velocità u senza deformarsi.

³Più correttamente: con la stessa *densità spettrale*, perchè dopo il passaggio al limite non si può più parlare dell'ampiezza di una singola componente $a(f)$, ma solo di una densità spettrale da/df .

Al termine della linea ognuna delle componenti armoniche subirà una riflessione, come previsto dall'equazione 2.30. Se la resistenza R_L di terminazione della linea è una resistenza pura, non dipendente dalla frequenza, tutte le armoniche verranno riflesse nella stessa misura e la loro combinazione nell'onda riflessa darà un impulso di forma eguale a quello originario, ma scalato in ampiezza per il coefficiente ρ_v .

Negli esempi che seguono si considerano alcuni casi di riflessione di impulsi in una linea che abbia perdite trascurabili e sia terminata su una resistenza pura. Si suppone anche che gli impulsi si susseguano con una cadenza lenta abbastanza perché tutti i fenomeni prodotti da un impulso lungo la linea siano scomparsi prima dell'inizio dell'impulso successivo.

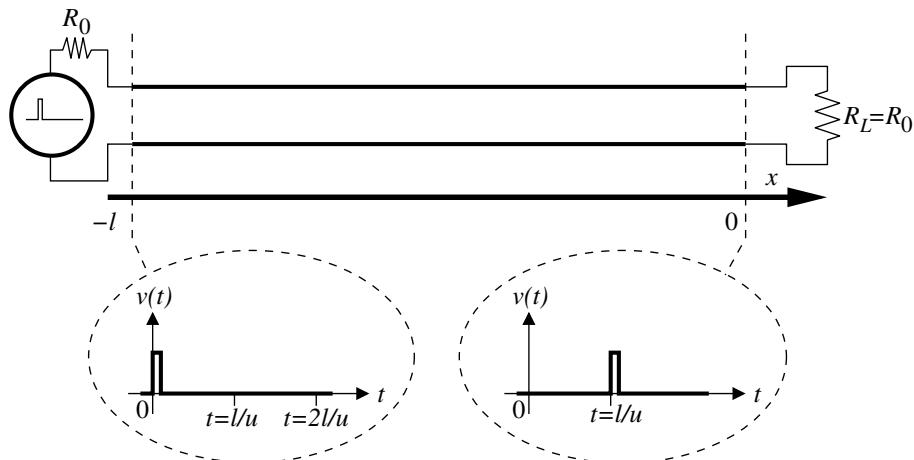


Figura 2.11: Propagazione di un impulso in una linea terminata sulla sua resistenza caratteristica. L'impulso parte dall'estremità sinistra al tempo $t = 0$ e raggiunge l'estremità destra dopo un tempo $t = l/u$.

1) $R_L = R_0$; $\rho_v = 0$: non si ha onda riflessa. L'impulso viene emesso dal generatore al tempo $t = 0$; viaggia lungo la linea da sinistra verso destra e dopo un tempo $t = l/u$ raggiunge la resistenza di terminazione R_L dove la sua energia viene interamente assorbita. (fig. 2.11).

2) $R_L > R_0$; $\rho_v > 0$: quando l'impulso arriva alla resistenza di terminazione viene in parte assorbito ed in parte riflesso, generando un impulso che torna indietro verso il generatore. L'impulso riflesso ha la stessa forma ed una ampiezza pari alla frazione ρ_v dell'impulso originale. Nel caso limite $R_L = \infty$, si ha $\rho_v = 1$ e l'impulso viene interamente riflesso, riportando indietro tutta l'energia verso il generatore (una resistenza di valore infinito non può assorbire alcuna energia dalla linea) (fig. 2.12).

3) $R_L < R_0$; $\rho_v < 0$: valgono le stesse considerazioni del punto 2, con la sola differenza che l'impulso riflesso ha polarità opposta a quella dell'impulso originale. Nel caso limite $R_L = 0$, si ha $\rho_v = -1$ e quindi ancora riflessione totale, ma con l'inversione della polarità dell'impulso (anche una resistenza di valore zero non può assorbire alcuna energia) (fig. 2.13).

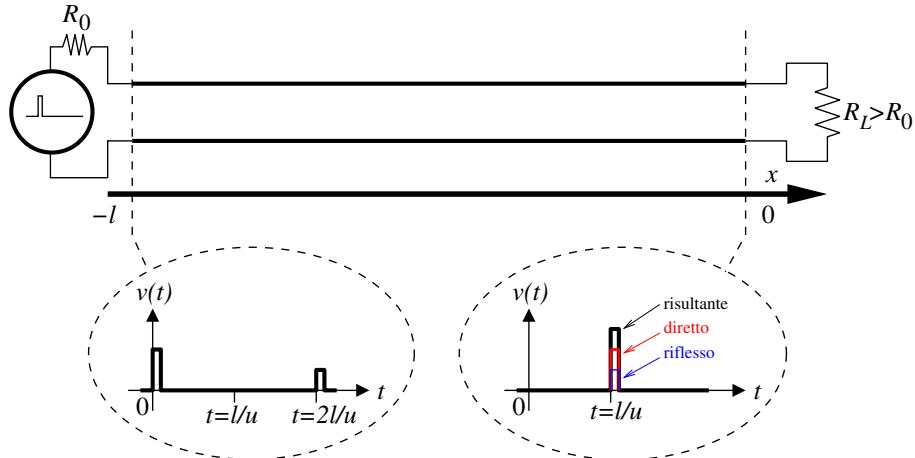


Figura 2.12: Propagazione di un impulso in una linea terminata su una resistenza $R_L > R_0$.

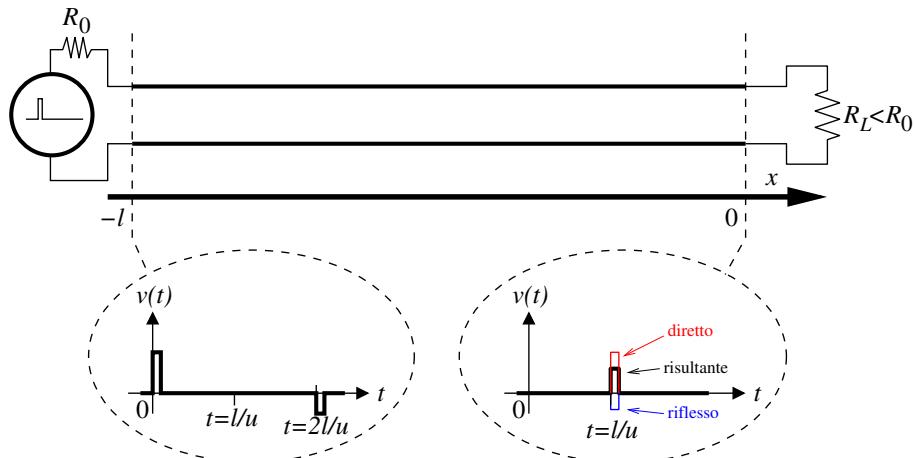


Figura 2.13: Propagazione di un impulso in una linea terminata su una resistenza $R_L < R_0$.

2.14 Resistenza interna del generatore

Nelle fig. 2.11–2.13 è stato indicato un generatore di tensione ideale, con in serie una resistenza interna pari alla resistenza caratteristica della linea; questo accorgimento è importante, anche nella pratica di laboratorio, se si vuole evitare che un impulso riflesso, una volta tornato all'inizio della linea di trasmissione, venga riflesso nuovamente, dando luogo al fenomeno delle riflessioni multiple.

In tutti e tre i casi se il generatore invia un impulso di ampiezza V_G , si ha una ampiezza $V_I = V_G/2$ all'estremità sinistra della linea. Infatti in presenza della sola onda diretta la linea si presenta come una resistenza di valore R_0 , che forma un partitore con rapporto 1/2 con la resistenza R_0 del generatore. All'estremità destra,

al termine della linea, il segnale ha ampiezza

$$V_E = V_I \cdot (1 + \rho_v) = \frac{V_G}{2} \cdot \left(1 + \frac{R_L - R_0}{R_L + R_0} \right) = V_G \cdot \frac{R_L}{R_L + R_0} \quad (2.50)$$

secondo la legge del normale partitore resistivo. Quindi *un generatore con impedenza interna R_0 collegato ad una linea di trasmissione con impedenza caratteristica R_0 appare al suo carico come se la linea non ci fosse*, a parte il ritardo nella propagazione del segnale ed eventuali attenuazioni nel caso di linea con perdite.

Questa è la ragione per cui nella strumentazione di laboratorio le impedenze interne degli apparecchi sono per lo più normalizzate al valore 50Ω e si usano per i collegamenti cavi coassiali con impedenze caratteristiche di 50Ω .

2.15 Riflessione di un segnale a gradino

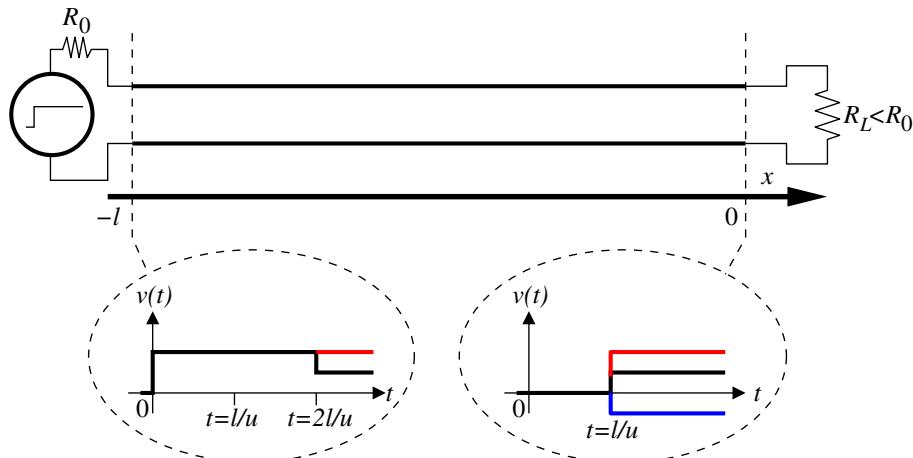


Figura 2.14: Riflessione di un segnale a gradino. I due grafici in tratto nero continuo indicano $V(t)$ all'inizio e alla fine della linea. I grafici in rosso rappresentano l'andamento di $V(t)$ se non ci fosse riflessione; il grafico in blu rappresenta il segnale riflesso.

È interessante riesaminare le situazioni descritte nel paragrafo (2.13) nel caso che il generatore invii un segnale “a gradino”: $v(t) = 0$ per $t < 0$, $v(t) = V^+ = V_G/2$ per $t \geq 0$ (Fig. 2.14). All'istante $t = 0$ e negli istanti immediatamente successivi il rapporto tra tensione e corrente all'inizio della linea è $v/i = R_0$ e sulla linea è presente la sola onda diretta. La corrente erogata dal generatore non può dipendere da quello che è collegato all'altra estremità della linea prima che sia trascorso un tempo $t = 2l/u$.

Il fronte del segnale a gradino viaggia lungo la linea per il tempo l/u , giunge all'altra estremità, viene riflesso a seconda delle condizioni della terminazione e dopo un altro tempo l/u ritorna al generatore.

A questo punto all'inizio della linea si ha la sovrapposizione dei due segnali diretto e riflesso:

$$\frac{v(t)}{i(t)} = \frac{V^+ + \rho_v V^+}{V^+/R_0 - \rho_v V^+/R_0} = R_0 \frac{1 + \rho_v}{1 - \rho_v} = R_L \quad (2.51)$$

Quindi una resistenza R_L collegata alla fine di un tratto di linea di trasmissione ideale viene vista dal generatore che invia un segnale a gradino all'inizio della linea come una resistenza di valore R_L (ovviamente) ma solo dopo il tempo $t = 2l/u$. Prima di questo tempo il generatore *vede* la resistenza R_0 .

In questo esempio l'analisi della sovrapposizione dei segnali diretto e riflesso risulta particolarmente semplice perché i segnali sono costanti nel tempo, a parte la variazione all'istante $t = 0$ da 0 a V^+ ; in pratica, per $t \geq 0$ si tratta di una corrente continua.

2.16 Riflessioni multiple

Nel caso che una linea di trasmissione sia terminata con una impedenza diversa dalla propria impedenza caratteristica a tutte e due le estremità si ha il fenomeno delle *riflessioni multiple*. In fig. 2.15 è riportato il caso di una linea senza perdite

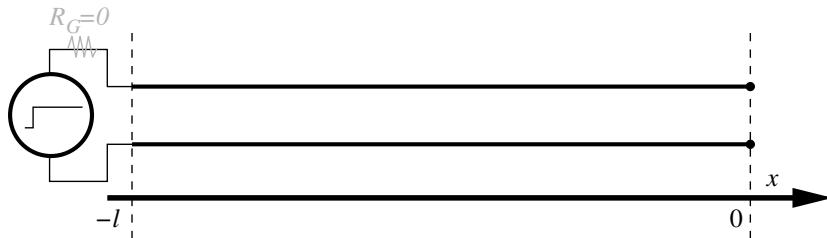


Figura 2.15: Linea di trasmissione senza perdite alimentata ad una estremità da un generatore di tensione a gradino con resistenza interna nulla. All'altra estremità la linea è aperta ($R_L = \infty$). Il sistema ideale non contiene alcun elemento dissipativo.

collegata ad una estremità ad un generatore ideale di tensione (quindi con resistenza interna $R_G = 0$) e lasciata aperta all'altra estremità ($R_L = \infty$).

Nel grafico 2.16 sono riportate in funzione del tempo la tensione $v(t)$ e la corrente $i(t)$ all'estremità sinistra della linea. Il generatore fornisce all'istante $t = 0$ un segnale di tensione a gradino come nel caso di fig. 2.14. Poiché il generatore ha resistenza interna nulla, la tensione $v(t)$ ai suoi capi (ed all'inizio della linea) è quella nominale: al tempo $t = 0$ si ha il salto da 0 a V_0 , valore che poi viene mantenuto stabile per un tempo indefinito. Contemporaneamente anche la corrente passa da 0 al valore $I_0 = V_0/R_0$.

Nel tempo $T = l/u$ il segnale giunge all'estremità opposta della linea, dove viene interamente riflesso con coefficiente di riflessione $\rho_v = 1$. L'ampiezza del segnale riflesso è ancora V_0 per quanto riguarda la tensione, ma $-I_0$ per la corrente, in quanto ora la propagazione sta avvenendo da destra verso sinistra. Al tempo $t = 2T$ il segnale riflesso giunge all'inizio della linea e si combina con quello originario. Contemporaneamente, però, il segnale riflesso trova la linea chiusa su un cortocircuito (la resistenza interna nulla del generatore) e viene quindi nuovamente riflesso da sinistra verso destra, stavolta con coefficiente di riflessione $\rho_v = -1$, quindi con

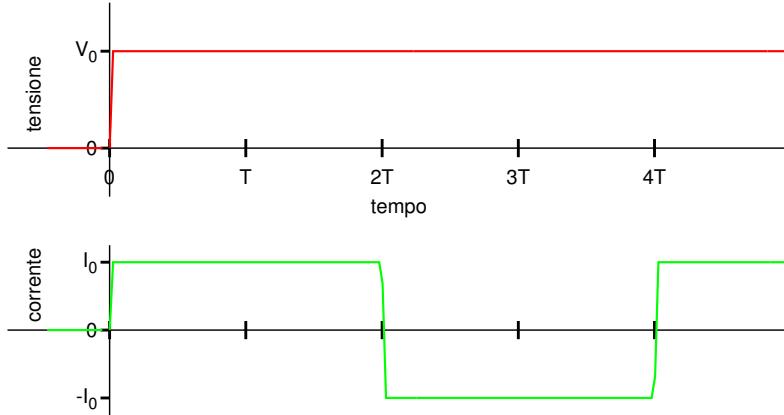


Figura 2.16: Tensione e corrente della linea di fig. 2.15 nel punto di alimentazione. La tensione passa da $0 V$ a V_0 all'istante $t = 0$ e rimane a questo valore indefinitamente; la corrente oscilla tra i valori $I_0 = V_0/R_0$ e $-I_0$ con periodo $2T = 2 \cdot l/u$ (l = lunghezza della linea, u = velocità di propagazione del segnale lungo la linea, R_0 = resistenza caratteristica della linea).

tensione $-V_0$ e corrente $-I_0$. In definitiva all'istante $t = 2T$ all'estremità sinistra della linea si ha

$$v(2T) = V_0 + V_0 - V_0 = V_0 \quad (2.52)$$

$$i(2T) = I_0 - I_0 - I_0 = -I_0 \quad (2.53)$$

Nel successivo intervallo di tempo $2T$ il fenomeno si ripete uguale, con l'unica differenza del cambiamento di segno nella corrente, dando all'istante $t = 4T$:

$$v(4T) = V_0 \quad (2.54)$$

$$i(4T) = I_0 \quad (2.55)$$

e così via all'infinito, come illustrato nel grafico 2.16.

Il fatto che la linea di trasmissione diventi (nel caso ideale) sede di una corrente oscillante in modo stazionario non deve stupire. Il fenomeno è analogo a quello che si verifica quando il segnale a gradino di un generatore di tensione ideale è applicato ad un circuito risonante ideale (fig. 2.17).

2.16.1 Riflessioni multiple in un sistema con perdite

Il caso più interessante in pratica è quello in cui si ha $R_0 \neq R_G \neq 0$. Nella figura 2.18 sono riportati in alto i grafici di corrente (rosso) e tensione (verde) all'inizio di una linea con $R_0 = 50 \Omega$, $R_G = 500 \Omega$, $T = 20 \text{ ns}$ ed $R_L = \infty$, con una ampiezza della tensione del generatore $V_G = 1 V$. In basso gli stessi grafici per $R_L = 500 \Omega$.

L'andamento è quello di una curva a gradini con l'altezza dei gradini che si riduce nel tempo con decadimento esponenziale. Nel limite in cui il tempo T di propagazione del segnale da un capo all'altro della linea è molto minore delle costanti di tempo formate dalla capacità totale della linea, lC ,⁴ e dalle resistenze che si trovano in

⁴Come detto all'inizio del capitolo, C è la capacità per unità di lunghezza.

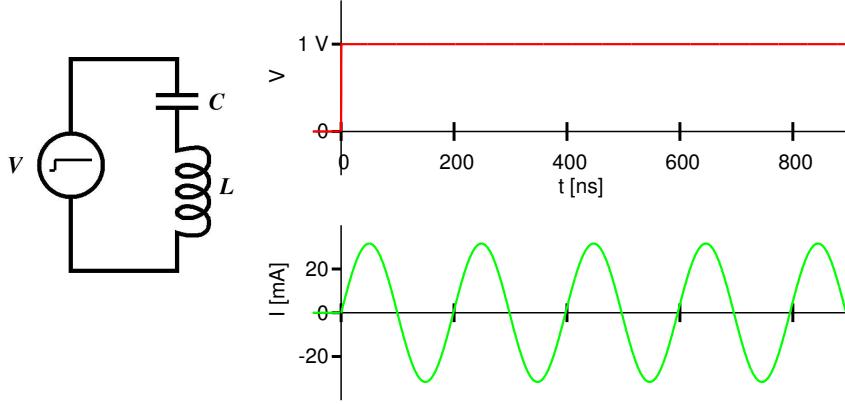


Figura 2.17: Circuito risonante con smorzamento nullo eccitato da una tensione a gradino. In alto, in rosso: tensione ai capi del circuito risonante; in basso, in verde, corrente attraverso condensatore ed induttanza ($L = 1 \mu H$, $C = 1 nF$).

parallelo alla linea, R_L e R_G , cioè quando i gradini diventano sempre più stretti e più bassi, la curva diventa l'esponenziale di carica del condensatore di un circuito RC a costanti concentrate.

Il gradino iniziale V_0 all'inizio della linea è determinato dal rapporto di partizione tra le resistenze interne del generatore, R_G , e la resistenza caratteristica della linea, R_0 :

$$V_0 = V_G \cdot \frac{R_0}{R_0 + R_G} \quad (2.56)$$

Dopo un intervallo di tempo $2T$ il gradino è giunto all'estremità destra della linea, è stato riflesso con coefficiente di riflessione

$$\rho_L = \frac{R_L - R_0}{R_L + R_0} \quad (2.57)$$

ed è tornato all'inizio della linea, dove si somma al gradino iniziale V_0 e viene nuovamente riflesso verso destra, questa volta con coefficiente

$$\rho_G = \frac{R_G - R_0}{R_G + R_0} \quad (2.58)$$

Quindi all'inizio della linea dopo un tempo $2T$ al gradino iniziale V_0 si sommano i due contributi $V_0\rho_L$ e $V_0\rho_L\rho_G$:

$$V(2T) = V_0 + V_0\rho_L + V_0\rho_L\rho_G \quad (2.59)$$

Dopo un ulteriore tempo $2T$ si hanno i due nuovi contributi $V_0\rho_L^2\rho_G$ e $V_0\rho_L^2\rho_G^2$ e così via. Sommando i contributi successivi si ha:

$$V(2nT) = V_0 \left[1 + \rho_L \left(1 + \rho_G \right) \sum_{i=0}^{n-1} \left(\rho_L \rho_G \right)^i \right] \quad (2.60)$$

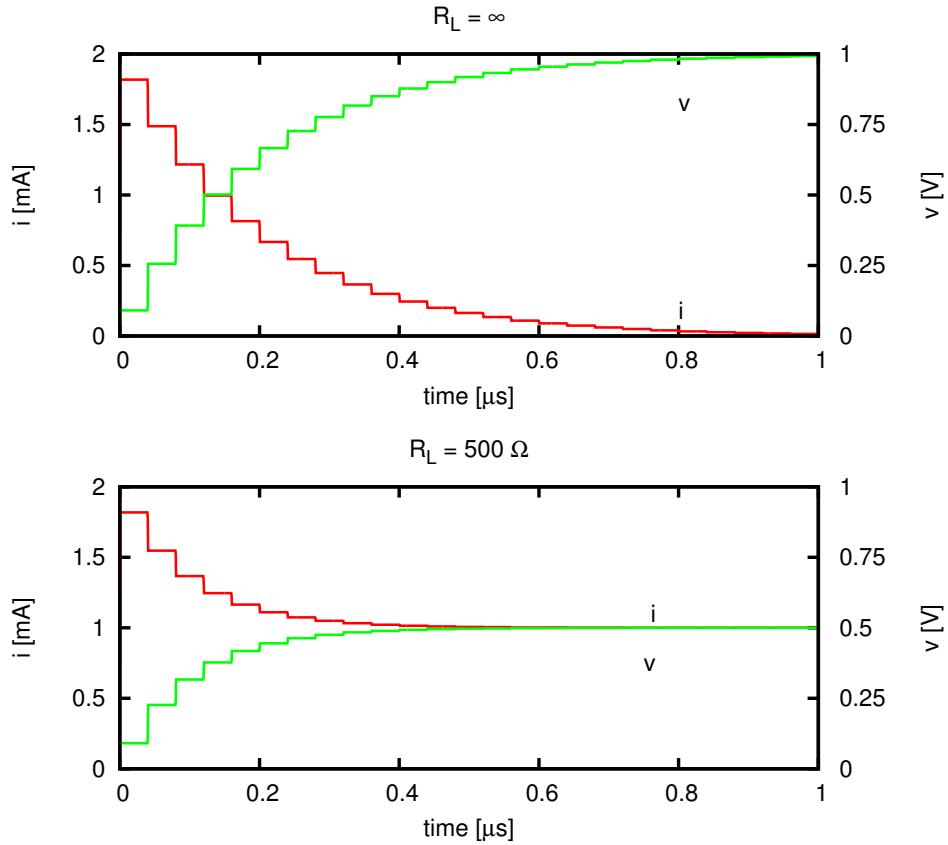


Figura 2.18: Corrente (in rosso) e tensione (in verde) nel punto di alimentazione di una linea con lunghezza elettrica $T = 20 \text{ ns}$ e resistenza caratteristica $R_0 = 50 \Omega$ da parte di un generatore di tensione a gradino con resistenza interna $R_G = 500 \Omega$. Il grafico in basso si riferisce ad una linea terminata con una resistenza $R_L = 500 \Omega$, quello in alto ad una linea aperta ($R_L = \infty$).

con $n = 0, 1, 2, \dots$. Ricordando che

$$\sum_{i=0}^{n-1} x^n = \frac{1 - x^n}{1 - x}$$

si ottiene

$$V(t) = V_0 \left[1 + \rho_L \left(1 + \rho_G \right) \cdot \frac{1 - \left(\rho_L \rho_G \right)^{\frac{t}{2T}}}{1 - \rho_L \rho_G} \right] \quad (2.61)$$

con $t = 0, 2T, 4T, 6T, \dots$. Sostituendo i valori di ρ_L , ρ_G e V_0 , dopo una certa manipolazione algebrica, si arriva a

$$V(t) = V_G \left[\frac{R_L}{R_G + R_L} - \left(\frac{R_L}{R_L + R_G} - \frac{R_0}{R_0 + R_G} \right) \cdot e^{\frac{t}{2T} \log(\rho_L \rho_G)} \right] \quad (2.62)$$

Dalle due equazioni 2.25 e 2.11 si ottengono le relazioni

$$L = R_0^2 C \quad (2.63)$$

$$T = \frac{l}{v} = l\sqrt{LC} = R_0 l C \quad (2.64)$$

Sostituendo $R_0 = T/(lC)$ nelle espressioni di ρ_L e di ρ_G , se

$$\frac{R_0}{R_L} = \frac{T}{lCR_L} \ll 1 \quad \text{e} \quad \frac{R_0}{R_G} = \frac{T}{lCR_G} \ll 1$$

si può sviluppare in serie di Taylor al primo ordine ottenendo

$$\log(\rho_L \cdot \rho_G) = -\frac{2T}{lC} \left(\frac{1}{R_L} + \frac{1}{R_G} \right) \quad (2.65)$$

Il risultato finale è quindi la legge esponenziale di carica di un condensatore dal valore di tensione iniziale $V_G R_0 / (R_0 + R_G)$ al valore finale $V_G R_L / (R_G + R_L)$ con costante di tempo $lCR_L R_G / (R_L + R_G)$.

2.17 Velocità di fase e di gruppo

Come visto nel par. 2.5, dalla equazione

$$V(x, t) = V^\pm e^{j(\omega t \pm \beta x)} \quad (2.66)$$

che descrive la propagazione di un'onda con frequenza angolare ω e lunghezza d'onda $\lambda = 2\pi\beta$, si può ricavare la velocità con cui procede nello spazio la fase, ponendo

$$\omega t \pm \beta x = \text{cost.} \quad (2.67)$$

Differenziando questa espressione, si ottiene la *velocità di fase*

$$u = \frac{dx}{dt} = \mp \frac{\omega}{\beta} \quad (2.68)$$

Nel caso ideale, cioè quando valgono le condizioni di Heaviside e R, L, C e G non dipendono da ω , si ha $\beta \propto \omega$ e quindi $u = \text{cost.}$

Per una linea reale questo vale solo in forma approssimata. Nei mezzi materiali la velocità di fase dipende sempre in qualche misura da ω .

La sovrapposizione di due segnali di frequenza ω e costante di propagazione β leggermente diverse ($\omega \pm \Delta\omega$ e $\beta \pm \Delta\beta$, con $\Delta\omega \ll \omega$ e $\Delta\beta \ll \beta$) possono essere considerati il risultato della modulazione di una *portante* ω, β con un segnale $\Delta\omega, \Delta\beta$:

$$\begin{aligned} v(x, t) &= v_0 \cdot \{\cos[(\omega + \Delta\omega)t - (\beta + \Delta\beta)x] \\ &\quad + \cos[(\omega - \Delta\omega)t - (\beta - \Delta\beta)x]\} \\ &= 2v_0 \cdot \cos(\omega t - \beta x) \cdot \cos(\Delta\omega t - \Delta\beta x) \end{aligned} \quad (2.69)$$

La portante si propaga come si è visto con velocità di fase $u = \omega/\beta$, mentre il segnale modulante (l'*informazione*) si propaga con la propria velocità di fase $u' = \Delta\omega/\Delta\beta$ e per $\Delta\omega, \Delta\beta \rightarrow 0$ si ha

$$u' = \frac{d\omega}{d\beta} \quad (2.70)$$

La velocità di fase della modulazione, velocità con cui si propaga l'informazione, è detta *velocità di gruppo*.

Capitolo 3

Semiconduttori

Con il termine *semiconduttori* si indicano alcuni elementi delle colonne III, IV e V della tavola periodica, caratterizzati da una resistività elettrica ρ intermedia tra quella molto bassa dei conduttori e quella estremamente elevata degli isolanti (fig. 3.1). I

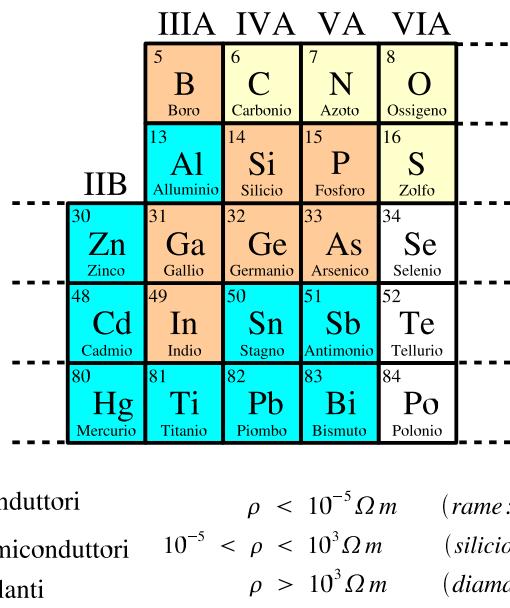


Figura 3.1: Semiconduttori.

semiconduttori possono essere composti da una singola specie atomica della quarta colonna (semiconduttori *elementali*: *Si*, *Ge*) oppure possono essere combinazioni degli elementi delle colonne III e V oppure II e VI (semiconduttori composti).

3.1 Resistività

La grandissima differenza di resistività che si osserva tra conduttori, semiconduttori e isolanti è dovuta principalmente alla differente concentrazione in volume dei

portatori di carica. Infatti la densità di corrente di conduzione \mathbf{J} , dovuta al moto di deriva dei portatori in presenza di un campo elettrico \mathbf{E} , per i conduttori ohmici, è:

$$\mathbf{J} = \text{densità di corrente } (A \text{ } m^{-2})$$

$$n = \text{Concentrazione dei portatori } (m^{-3})$$

$$\mathbf{J} = n q \mu \mathbf{E} \quad (3.1)$$

$$n q \mu = \sigma = \rho^{-1} \quad (3.2)$$

$$q = \text{Carica dei portatori } (C)$$

$$\mu = \text{Mobilità dei portatori } (m^2 \text{ } V^{-1} \text{ } s^{-1})$$

$$\sigma = \text{Conducibilità del materiale } (\Omega^{-1} \text{ } m^{-1})$$

$$\rho = \text{Resistività del materiale } (\Omega \text{ } m)$$

La mobilità μ differisce relativamente poco tra materiali diversi, mentre la concentrazione dei portatori di carica tra isolanti e conduttori spazia su oltre 20 ordini di grandezza ($10^1 \dots 10^{22} \text{ cm}^{-3}$). Nei semiconduttori si colloca, su scala logaritmica, a metà di questo intervallo (fig. 3.2).

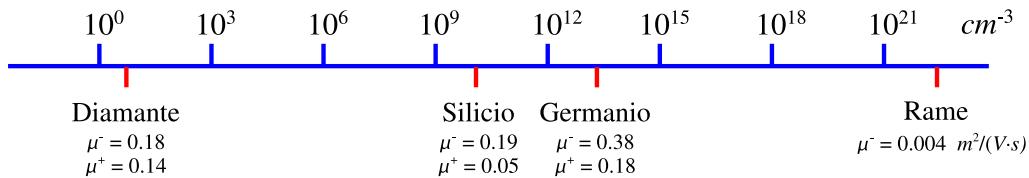


Figura 3.2: Concentrazione dei portatori di carica per alcuni materiali. Sotto i nomi sono riportate le mobilità di elettroni (μ^-) e lacune (μ^+).

3.2 Germanio e Silicio

Il Germanio è stato utilizzato per la produzione di dispositivi elettronici (diodi e transistor) soprattutto negli anni '50 e '60 del secolo scorso; successivamente è stato sostituito dal Silicio. Attualmente con il Silicio è realizzata la grandissima maggioranza dei dispositivi a semiconduttore (fig. 3.3).

Germanio e Silicio appartengono alla IV colonna della tavola periodica, sono tetravalenti e formano cristalli con reticolo tetraedrico. In figura 3.4 è riportata una proiezione in piano del reticolo cristallino tridimensionale, in cui ogni atomo forma quattro legami covalenti con i quattro atomi primi vicini.

3.3 Elettroni e Lacune

Nella figura 3.4a è schematizzata la situazione del reticolo cristallino alla temperatura dello zero assoluto: tutti gli elettroni di valenza risultano stabilmente impegnati nei rispettivi legami; non esistono elettroni liberi all'interno del reticolo e quindi il cristallo è un isolante perfetto. La situazione cambia per temperature diverse dallo zero assoluto: ogni elettrone ha una probabilità piccola ma finita di uscire dalla buca



Figura 3.3: Dispositivi a semiconduttore al Germanio ed al Silicio. Da sinistra: tre transistor al germanio; un diodo al silicio per piccoli segnali, uno di media potenza ed uno di potenza; due transistor al silicio per piccoli segnali ed uno di potenza; un circuito integrato digitale ed un microprocessore.

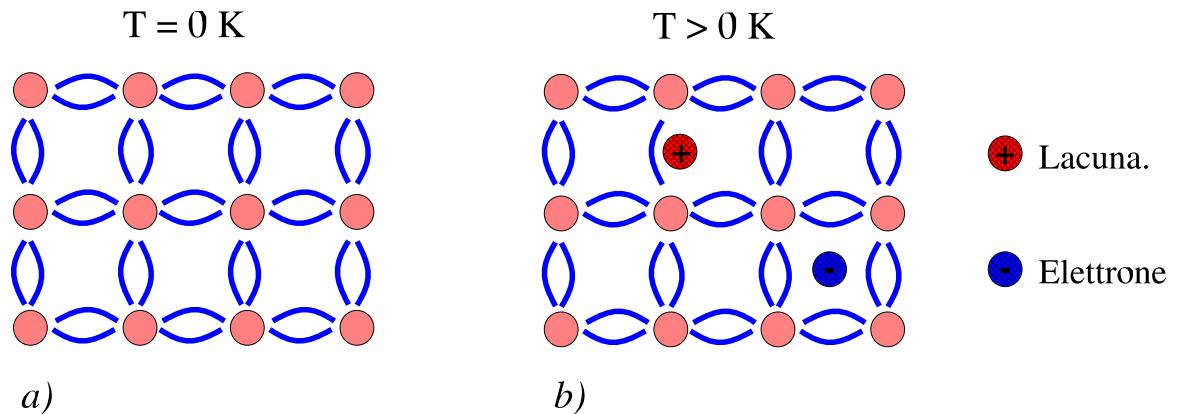


Figura 3.4: Semiconduttore intrinseco. *a)* Allo zero assoluto non ci sono nel reticolo portatori di carica disponibili; *b)* a temperature maggiori dello zero assoluto alcuni elettroni si rendono disponibili come elettroni di conduzione.

di potenziale corrispondente al legame covalente e trovarsi nel reticolo, disponibile come elettrone di conduzione (figura 3.4*b*).

Per ogni elettrone che abbandona il proprio legame covalente si formano due cariche libere: una carica negativa (l'elettrone) ed una positiva (la *lacuna*, cioè la carica positiva rimasta scoperta nel reticolo a seguito dell'allontanamento dell'elettrone). Le lacune sono portatori di carica a tutti gli effetti e danno un contributo al trasporto di carica indipendente da quello degli elettroni di conduzione. In fig. 3.5 (a sinistra) una lacuna si trova in *A*; un elettrone di valenza salta da *B* in *A*; la lacuna in *A* scompare e riappare in *B*: apparentemente la lacuna si è spostata da *A* a *B*. In realtà sono sempre gli elettroni che si muovono, ma poiché si tratta di elettroni di valenza questo meccanismo fornisce un movimento di carica indipendente da quello degli elettroni di conduzione.

Lacune ed elettroni si generano continuamente a coppie in seguito alla rottura di legami covalenti per effetto della agitazione termica e scompaiono a coppie, quando

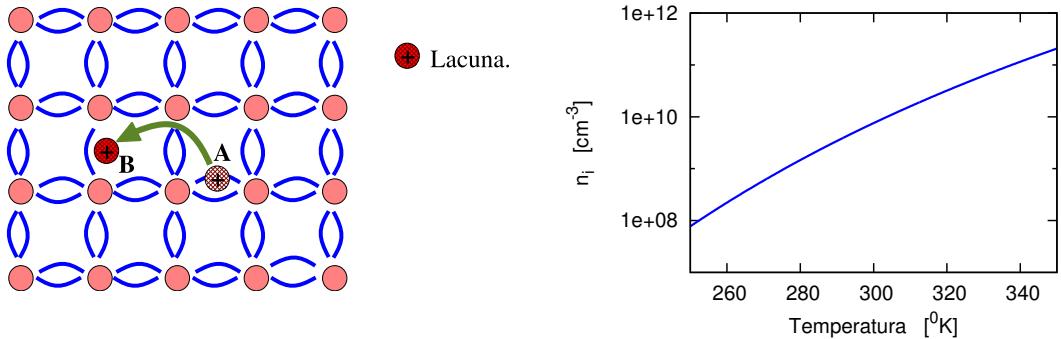


Figura 3.5: A sinistra: movimento delle lacune: una lacuna si trova in *A*; un elettrone di valenza salta da *B* in *A*; la lacuna in *A* scompare e riappare in *B*. A destra: concentrazione intrinseca n_i dei portatori in funzione della temperatura.

una lacuna ed un elettrone si ricombinano a riformare un legame covalente. Di conseguenza nel semiconduttore *intrinseco* (cioè non *drogato*, come si vedrà più avanti) la concentrazione delle lacune (p) sarà sempre eguale a quella degli elettroni (n). Il valore n_i di p ed n , detto *concentrazione intrinseca*, è determinato dall'equilibrio dinamico tra generazione e ricombinazione e dipende dalle caratteristiche del materiale semiconduttore e dalla temperatura T :

$$n_i^2 = n^2 = p^2 = B T^3 e^{-E_G/kT} \quad (3.3)$$

dove E_G è l'energia di rottura di un legame covalente, k la costante di Boltzmann e B una costante caratteristica del materiale. La concentrazione intrinseca n_i aumenta rapidamente con la temperatura (fig. 3.5 a destra) e questo aumento è la causa della forte diminuzione della resistività dei semiconduttori con il riscaldamento (a differenza dei metalli, la cui resistività cresce con la temperatura).

3.4 Modello a bande di energia

Una descrizione accurata dei meccanismi di trasporto nei semiconduttori si può avere solo facendo uso della meccanica quantistica, tramite la teoria delle bande (banda di conduzione e di valenza) ed esula dallo scopo di queste note. Qui è sufficiente utilizzare i risultati che descrivono fenomenologicamente tali meccanismi. Il risultato più importante è costituito proprio dal fatto che in un semiconduttore si ha un doppio meccanismo di conduzione, legato alla presenza di portatori di carica negativi (gli elettroni di conduzione) e positivi (le lacune), che portano ciascuno una unità di carica elettronica q di segno opposto. In un metallo invece gli unici portatori di carica sono gli elettroni. In presenza di un campo elettrico si stabilisce quindi una corrente di deriva (*drift current*) con densità \mathbf{J} :

Metalli**Semiconduttori**

$$\mathbf{J} = q n \mu_n \mathbf{E}$$

$$\mathbf{J} = q (n \mu_n + p \mu_p) \mathbf{E} \quad (3.4)$$

$$\sigma = q n \mu_n$$

$$\sigma = q (n \mu_n + p \mu_p) \quad (3.5)$$

Le caratteristiche chimico-fisiche principali alla base delle proprietà di conduzione di Silicio e Germanio intrinseco sono riportate in tabella 3.1.

		Ge	Si	
numero atomico		32	14	
peso atomico (relativo)		72.6	28.1	
densità		5.32	2.33	g/cm^3
densità atomica		$4.4 \cdot 10^{22}$	$5 \cdot 10^{22}$	$atomi/cm^3$
costante dielettrica relativa	ϵ_r	16	12	
Energia di legame (<i>bandgap</i>)	E_G	0.72	1.1	eV
concentrazione intrinseca	n_i	$2.5 \cdot 10^{13}$	$1.5 \cdot 10^{10}$	cm^{-3}
$d(\log n_i)/dT$		16	10	$\%/\circ K$
frazione portatori/legami		$0.14 \cdot 10^{-9}$	$0.75 \cdot 10^{-13}$	
resistività	ρ	0.45	2300	$\Omega \cdot m$
mobilità degli elettroni	μ_n	0.38	0.19	$m^2/(V \cdot s)$
mobilità delle lacune	μ_p	0.18	0.05	$m^2/(V \cdot s)$
coeff. di diffusione degli elettroni	D_n	$9.9 \cdot 10^{-3}$	$3.4 \cdot 10^{-3}$	m^2/s
coeff. di diffusione delle lacune	D_p	$4.7 \cdot 10^{-3}$	$1.3 \cdot 10^{-3}$	m^2/s

Tabella 3.1: Principali caratteristiche chimico-fisiche di Germanio e Silicio intrinseco a temperatura ambiente ($300^\circ K$).

3.5 Semiconduttore estrinseco

Si ha un semiconduttore *estrinseco* o *drogato* quando piccole quantità di atomi di elementi trivalenti (es.: boro, gallio, indio) o pentavalenti (es.: antimonio, fosforo, arsenico) vengono sostituite nel reticolo cristallino agli atomi originali del semiconduttore, Germanio o Silicio (fig. 3.6). La frazione di atomi sostituiti tipicamente è compresa nell'intervallo $10^{-3} \dots 10^{-9}$, per cui le proprietà chimico-fisiche del reticolo cristallino rimangono sostanzialmente quasi tutte inalterate. Gli atomi estranei si sostituiscono a quelli del semiconduttore originale e portano ciascuno nel reticolo un eccesso di un elettrone, nel caso degli elementi pentavalenti (atomi *donatori*, droggaggio di tipo *N*), o un difetto di un elettrone, cioè una lacuna nel caso degli elementi trivalenti (atomi *accettori*, droggaggio di tipo *P*).

In entrambi i casi ci troviamo ad avere inserito portatori di carica liberi nel cristallo: gli elettroni in eccesso portati dal droggaggio *N* hanno una energia di legame con il proprio atomo sufficientemente bassa da trovarsi già a temperatura ambiente tutti disponibili come elettroni di conduzione, e quindi come portatori di carica negativa; analogamente le lacune portate dal droggaggio di tipo *P* costituiscono altrettanti portatori di carica positivi.

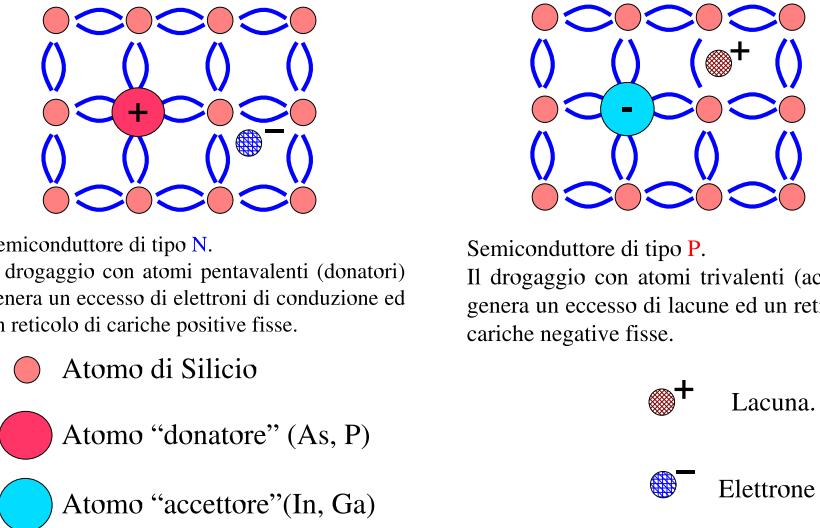


Figura 3.6: Modello di semiconduttore estrinseco

3.6 Legge di Azione di Massa

L'aumento della concentrazione di uno dei due tipi di portatori n o p nel reticolo in seguito al drogaggio porta conseguentemente una diminuzione della concentrazione dell'altro portatore. Infatti la probabilità di ricombinazione delle coppie lacuna-elettrone è legata al prodotto $n \cdot p$ delle due concentrazioni, mentre la probabilità di generazione di nuove coppie per effetto della agitazione termica è legata alla concentrazione di legami non ionizzati e quindi rimane sostanzialmente costante anche in seguito al drogaggio. Questo equilibrio è riassunto nella *legge di azione di massa*:

$$n \cdot p = n_i^2(T) \quad (3.6)$$

Nell'intervallo di drogaggi usati abitualmente si ha, indicando con N_A ed N_D rispettivamente la concentrazione di atomi accettori e donatori:

$$\begin{aligned} \text{semiconduttore } P: \quad & N_A \gg n_i \quad p_p \simeq N_A \quad n_p = n_i^2/N_A \\ \text{semiconduttore } N: \quad & N_D \gg n_i \quad n_n \simeq N_D \quad p_n = n_i^2/N_D \end{aligned}$$

Con p_p e n_n sono state indicate le concentrazioni di lacune ed elettroni nelle regioni rispettivamente di tipo P ed N ; questi sono i *portatori maggioritari*, in quanto presenti nella regione in quantità preponderante. I simboli n_p e p_n indicano invece le concentrazioni dei *portatori minoritari*: l'altra specie di portatori, la cui concentrazione risulta fortemente depressa a causa della abbondante presenza dei maggioritari (fig. 3.7).

3.7 Corrente di diffusione

Nei semiconduttori il trasporto di carica avviene come moto di deriva dei portatori per effetto del campo elettrico (corrente di conduzione), analogamente a quanto

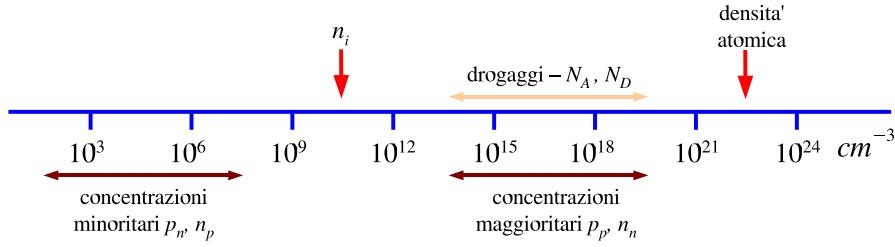


Figura 3.7: Livelli di droggaggio nei semiconduttori al silicio e concentrazioni dei portatori a temperatura ambiente.

avviene nei metalli. Oltre a questo meccanismo ve ne è un secondo, legato alla diffusione, che invece è irrilevante nel caso dei metalli: un gradiente di concentrazione dei portatori di carica dà origine ad un moto di diffusione e quindi ad una densità di corrente elettrica di lacune (\mathbf{J}_p) e di elettroni (\mathbf{J}_n) anche in assenza di un campo elettrico:

$$\mathbf{J}_p = -qD_p \nabla p \quad \mathbf{J}_n = qD_n \nabla n \quad (3.7)$$

La densità di corrente totale \mathbf{J} è data dalla somma dei contributi di conduzione e diffusione per entrambi i tipi di portatori:

$$\mathbf{J}_p = q \mu_p p \mathbf{E} - q D_p \nabla p \quad \mathbf{J}_n = q \mu_n n \mathbf{E} + q D_n \nabla n \quad (3.8)$$

$$\mathbf{J} = \mathbf{J}_p + \mathbf{J}_n \quad (3.9)$$

I coefficienti di diffusione D_p , D_n e le mobilità μ_p , μ_n sono legati dalla relazione di Einstein:

$$\frac{D_p}{\mu_p} = \frac{D_n}{\mu_n} = V_T = \frac{kT}{q}. \quad (3.10)$$

A temperatura ambiente $V_T \simeq 26 \text{ mV}$.

Capitolo 4

Diodi

4.1 Giunzione $P-N$

Una giunzione $P-N$ è formata dal contatto di due regioni con drogaggio diverso, una di tipo P e l'altra di tipo N . Indipendentemente dai procedimenti tecnologici che portano alla realizzazione di una giunzione di questo tipo, quello che è importante è che si realizzi una sottile zona in cui si passa rapidamente dalla regione P drogata con atomi accettori, e quindi ricca di portatori di carica positivi (lacune), alla regione N , drogata con donatori e ricca di portatori di carica negativi (elettroni). Nella giunzione si vengono quindi ad avere fortissimi gradienti di concentrazione (fig. 4.1).

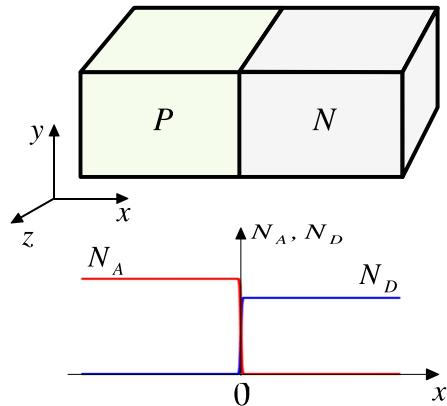


Figura 4.1: Profili di concentrazione del drogaggio in una giunzione $P-N$.

Per semplicità si considererà solo una giunzione dotata di simmetria piana: le concentrazioni quindi varieranno solo lungo la direzione x , mentre nel piano perpendicolare a questa direzione il sistema sarà omogeneo. Nella pratica, anche quando questa condizione non è rispettata, se il raggio di curvatura della giunzione è grande rispetto allo spessore, il sistema può essere considerato con buona approssimazione dotato di simmetria piana. Lo spessore geometrico della giunzione, cioè la zona di disomogeneità chimica tra le due zone omogenee P ed N , è compreso tra una frazione di nm e qualche decina di nm .

4.2 Regione di Transizione

I profili di concentrazione di donatori ed accettori riportati in fig. 4.1 non possono corrispondere, in condizioni di equilibrio, ai profili di concentrazione dei portatori di carica, elettroni e lacune. Immaginiamo idealmente di creare la giunzione portando a contatto le due regioni P ed N , ciascuna con la propria concentrazione di portatori N_A ed N_D , uniforme in tutto il volume. Nella regione di contatto si viene ad avere un gradiente di concentrazione di portatori elevatissimo, che dà immediatamente luogo a correnti di diffusione di lacune dalla regione P alla regione N e di elettroni nella direzione opposta, fino a raggiungere una nuova condizione di equilibrio (fig. 4.2).

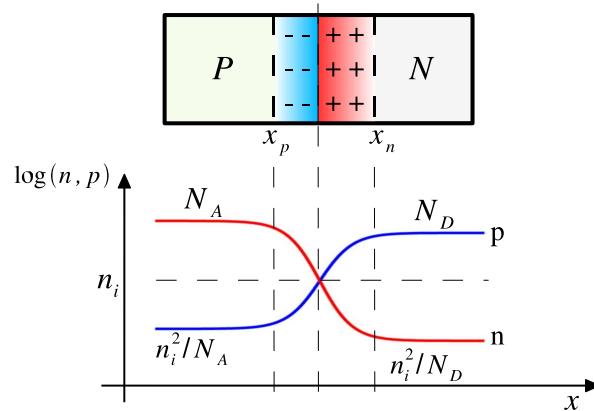


Figura 4.2: Regione di transizione e profili di concentrazione dei portatori di carica in una giunzione P - N . Linea rossa: concentrazione delle lacune; linea blu: concentrazione degli elettroni.

Il movimento di diffusione dei portatori da una regione all'altra crea una zona di carica elettrica in prossimità della giunzione, dovuta alla carica dei nuclei degli atomi del drogaggio immobilizzati nel reticolo cristallino e non più in equilibrio di carica con i propri elettroni. Si ha una carica positiva nella regione di tipo N , dove gli atomi pentavalenti rimangono con una unità di carica elettronica positiva in eccesso, non più neutralizzata dai corrispondenti elettroni di conduzione ed una carica negativa nella regione di tipo P , a causa dell'eccesso di elettroni di legame rispetto alle cariche dei nuclei trivalenti accettori. Si forma quindi una regione di doppio strato di carica, detta regione di *transizione* o di *carica spaziale* o di *svuotamento*, il cui campo elettrico ha direzione tale da opporsi al moto di diffusione dei portatori. Per evidenti ragioni termodinamiche il moto di diffusione non potrà continuare all'infinito. La giunzione troverà la sua condizione di equilibrio quando in ogni punto il movimento di diffusione delle cariche, generato dai gradienti di concentrazione, sarà bilanciato da un movimento opposto, di conduzione, dovuto al campo elettrico.

4.3 Potenziale di giunzione

Senza entrare nei dettagli dei profili di concentrazione dei portatori, illustrati qualitativamente in fig. 4.2, si può osservare che in generale all'equilibrio si dovrà

avere in ogni punto della regione di transizione, in media, un movimento nullo di cariche. In particolare, nella direzione x :

$$J_n = q \mu_n n E_x + q D_n \frac{dn}{dx} = 0 \quad (4.1)$$

$$J_p = q \mu_p p E_x - q D_p \frac{dp}{dx} = 0 \quad (4.2)$$

dove E_x è la componente x del campo elettrico e J_n e J_p sono i contributi dovuti rispettivamente ad elettroni e lacune della componente x del vettore densità di corrente. Le altre componenti sono nulle a causa della simmetria della giunzione ipotizzata al par. 4.1. Dalla equazione 4.1 si ottiene:

$$E_x = -\frac{1}{n} \cdot \frac{D_n}{\mu_n} \cdot \frac{dn}{dx} \quad (4.3)$$

Il campo elettrico è diverso da zero solo nella regione di transizione e nullo nelle parti omogenee delle due regioni P ed N (cioè per $x < x_p$ e $x > x_n$ in fig. 4.2) in quanto la giunzione nel suo complesso deve essere elettricamente neutra. La differenza di potenziale tra le due regioni è data dall'integrale di E_x dalla zona omogenea P alla zona omogenea N :

$$V_j = V_N - V_P = - \int_{x_p}^{x_n} E_x dx = \frac{D_n}{\mu_n} \int_{n_p}^{n_n} \frac{dn}{n} = V_T \cdot \log \left(\frac{n_n}{n_p} \right) \quad (4.4)$$

avendo utilizzato la relazione di Einstein (eq. 3.10):

$$\frac{D_n}{\mu_n} = \frac{kT}{q} = V_T \quad (4.5)$$

A temperatura ambiente $T = 300^\circ K$ si ha:

$$V_j \simeq 26 mV \cdot \log_e \left(\frac{n_n}{n_p} \right) \quad (4.6)$$

oppure

$$V_j \simeq 60 mV \cdot \log_{10} \left(\frac{n_n}{n_p} \right) \quad (4.7)$$

Il procedimento dalla equazione 4.1 alla 4.4 può essere ripetuto per la corrente J_p , arrivando al risultato:

$$V_j = V_T \cdot \log \left(\frac{p_p}{p_n} \right) \quad (4.8)$$

A causa della legge di azione di massa e ricordando che nelle zone omogenee $n_n \simeq N_D$ e $p_p \simeq N_A$, si ha:

$$\frac{n_n}{n_p} = \frac{N_D}{n_i^2/N_A} = \frac{N_D N_A}{n_i^2} = \frac{N_A}{n_i^2/N_D} = \frac{p_p}{p_n} \quad (4.9)$$

e quindi il valore di V_j calcolato con la 4.4 o la 4.8 è lo stesso, come è ovvio che sia. Analogamente è possibile verificare che il valore del campo elettrico che in ogni punto verifica la 4.1 soddisfa anche la 4.2.

La eq. 4.4 o l'equivalente 4.6 è una espressione della *distribuzione di Boltzmann*:

$$\frac{n_n}{n_p} = e^{qV_j/kT}. \quad (4.10)$$

Nella forma 4.7 è più nota come *legge di Nernst*.

Partendo dalla relazione di Boltzmann 4.10 e ripercorrendo il ragionamento a ritroso si sarebbe arrivati ad ottenere la relazione di Einstein 3.10.

I valori tipici dei drogaggi per il Silicio danno luogo, secondo la eq. 4.7 ad una *barriera di potenziale* o *potenziale di giunzione* V_j (in inglese, *built-in potential*) compreso tra 300 mV e 800 mV.

4.4 Diodo a giunzione

Dotando di un collegamento elettrico la regione P e la regione N di una giunzione $P-N$ si ottiene un dispositivo elettronico a due terminali, detto diodo a giunzione, o diodo semiconduttore, la cui caratteristica principale è di presentare una relazione corrente-tensione fortemente non lineare.

In condizioni di equilibrio termodinamico, cioè con il diodo a temperatura uniforme e non collegato a fonti di energia, i due terminali, supposti dello stesso metallo (ad es. rame), si troveranno allo stesso potenziale elettrico (fig. 4.3). Questo si può

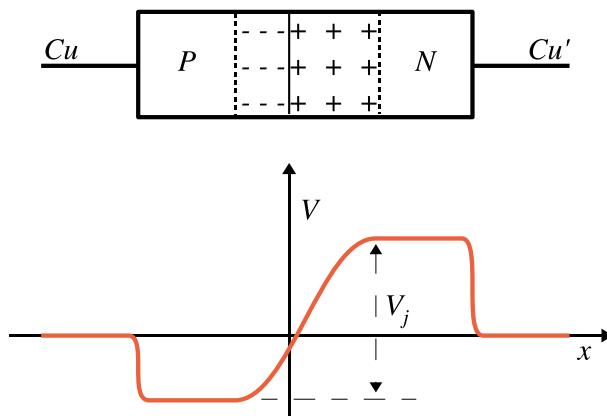


Figura 4.3: Potenziale elettrico attraverso la struttura di un diodo a giunzione.

immediatamente verificare se si pensa che le tre superfici di contatto $Cu||P$, $P||N$ e $N||Cu'$ sono tutte *reversibili*, cioè permeabili rispetto agli elettroni. In condizioni di equilibrio il potenziale elettrochimico degli elettroni dovrà essere lo stesso in tutte e quattro le regioni ed in particolare nei terminali Cu e Cu' . Essendo il metallo, e quindi il potenziale chimico, lo stesso, anche il potenziale elettrico in Cu e Cu' dovrà essere lo stesso.

4.5 Polarizzazione diretta ed inversa

Collegando il diodo a giunzione ad un generatore di d.d.p. V_D , si avranno gli effetti illustrati in fig. 4.4a e 4.4b. In figura 4.4a il generatore forza la barriera di potenziale

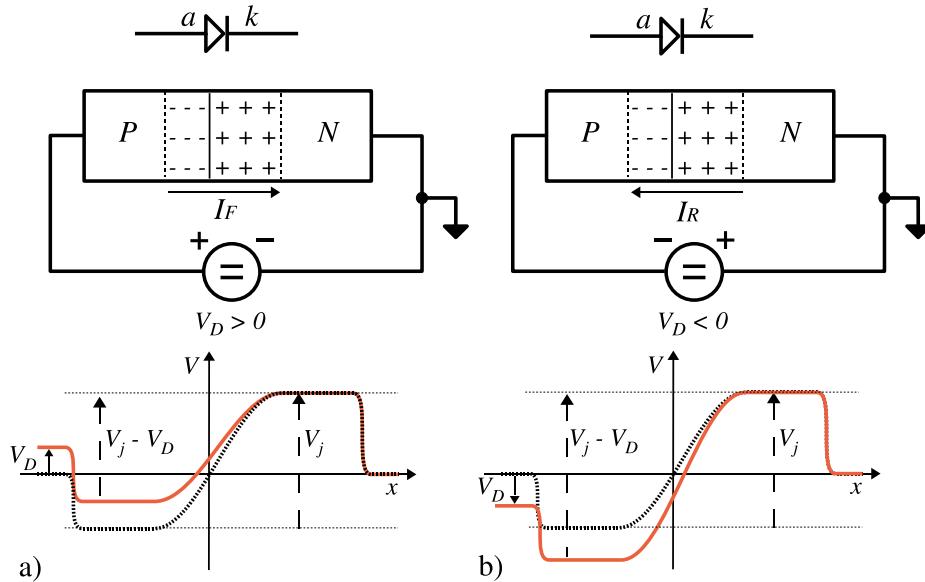


Figura 4.4: Polarizzazione diretta e inversa di una giunzione $P-N$. Il collegamento di massa indica il punto convenzionale rispetto a cui sono misurati i potenziali elettrici.

tra regione P ed N ad avere un valore più basso di quello corrispondente al potenziale V_j con l'effetto di sbilanciare l'equilibrio previsto dalle eq. 4.1 e 4.2 in favore del moto di diffusione delle lacune dalla regione P alla regione N attraverso la regione di transizione e di elettroni nel senso inverso. Entrambi questi flussi corrispondono ad una corrente I_F nella direzione indicata dalla freccia. Come si vedrà nel paragrafo successivo, la corrente I_F cresce in maniera esponenziale al crescere di V_D . In questa condizione il diodo si trova in *polarizzazione diretta*, o *conduzione* ($I_F = I_{Forward} = \text{corrente diretta}$).

Nel caso di fig. 4.4b la barriera di potenziale viene forzata ad un valore più alto di quello di equilibrio con l'effetto di indurre un ulteriore allontanamento di elettroni e lacune dalla regione di transizione ed un allargamento della medesima. Questa è la condizione di *polarizzazione inversa*, o *contropolarizzazione*, o *interdizione* del diodo. Si ha la circolazione di una debole corrente I_R che è da sei a dieci ordini di grandezza inferiore alla corrente I_F che si ha in condizioni operative normali ($I_R = I_{Reverse} = \text{corrente inversa}$).

Nella figura è riportato anche il simbolo con cui il diodo viene rappresentato negli schemi elettrici. Il terminale indicato con a è detto *anodo*; quello con k è detto *catodo*. Si ha conduzione quando l'anodo è positivo rispetto al catodo. La forma del simbolo, che ricorda una freccia, indica la direzione della corrente quando il diodo è in conduzione.

4.6 Correnti in un diodo polarizzato

4.6.1 Polarizzazione diretta

In un diodo in condizioni di polarizzazione diretta le lacune diffondono dalla regione P , dove sono i portatori maggioritari e sono quindi presenti in larghissima quantità, verso la regione N dove sono minoritari e quindi presenti in quantità ridottissima; analogamente accade per gli elettroni, che diffondono nella direzione opposta. Una

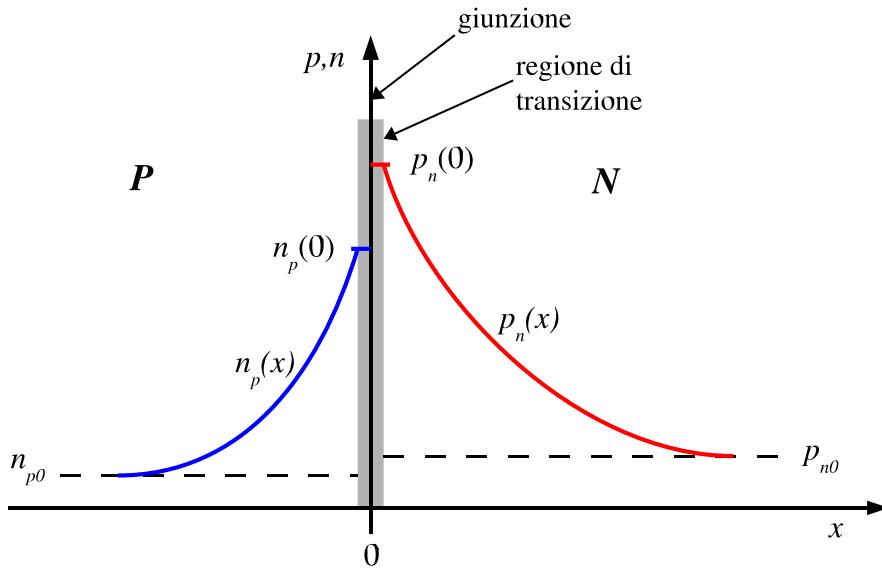


Figura 4.5: Concentrazione dei portatori minoritari in prossimità di una giunzione $P-N$ per un diodo direttamente polarizzato.

volta attraversata la giunzione e diventati quindi minoritari, i portatori continuano nel loro moto di diffusione allontanandosi sempre di più dalla regione di transizione. In prossimità della giunzione il sistema si viene a trovare in condizioni di non equilibrio: a causa della consistente iniezione di portatori minoritari si viene ad avere $p_n \cdot n_n \gg n_i^2$ nella regione N e $n_p \cdot p_p \gg n_i^2$ nella regione P . Ha luogo quindi un processo di ricombinazione tra maggioritari e minoritari fino a raggiungere nuovamente, allontanandosi a sufficienza dalla giunzione, i valori di equilibrio. La sovrapposizione dei due processi di diffusione e ricombinazione dà luogo ad uno stato stazionario con le distribuzioni di portatori raffigurate in fig. 4.5, mantenute dal generatore esterno V_D .

Le due linee tratteggiate orizzontali indicano i valori all'equilibrio delle concentrazioni degli elettroni nella regione P (n_{p0}) e delle lacune nella regione N (p_{n0}). Le due curve $n_p(x)$ e $p_n(x)$ indicano i profili di concentrazione in funzione della distanza dalla regione di transizione.

Le concentrazioni n_n e p_p dei portatori maggioritari, praticamente costanti al di fuori della regione di transizione e pari rispettivamente a N_D ed N_A , non sono indicate nella figura.

Partendo dalla equazione della diffusione (eq. 3.7), si può dimostrare che in condizioni stazionarie e per regioni P ed N sufficientemente estese le concentrazioni dei portatori minoritari in eccesso decadono esponenzialmente con la distanza dalla giunzione:

$$n_p(x) - n_{p0} = [n_p(0) - n_{p0}] e^{+x/L_n} \quad (4.11)$$

$$p_n(x) - p_{n0} = [p_n(0) - p_{n0}] e^{-x/L_p} \quad (4.12)$$

dove

$$L_n = \sqrt{D_n \cdot \tau_n} \quad L_p = \sqrt{D_p \cdot \tau_p} \quad (4.13)$$

sono le *lunghezze di diffusione* e τ_n , τ_p le *vite medie* di elettroni e lacune prima della ricombinazione, rispettivamente nella regione P ed N .

Introducendo alcune ipotesi semplificative è possibile determinare la relazione tra la corrente I_D attraverso il diodo e la tensione di polarizzazione V_D applicata:

- Lo spessore della regione di transizione intorno alla giunzione è trascurabile rispetto alle lunghezze L_p ed L_n : in questo paragrafo, *giunzione* e *regione di transizione* diventano quindi sinonimi.
- Il campo elettrico al di fuori della regione di transizione è trascurabile anche in condizioni di passaggio di corrente; questo vuol dire che il potenziale V_D applicato ai terminali del diodo viene a localizzarsi interamente come variazione del potenziale V_j attraverso la giunzione/regione di transizione.
- L'equazione di Boltzmann, che lega il potenziale di giunzione V_j alle concentrazioni dei portatori ai due lati della regione di transizione e che a rigore sarebbe valida solo in condizioni di equilibrio termodinamico, viene considerata valida con approssimazione sufficiente anche in condizione di stato stazionario vicino all'equilibrio.

La seconda e la terza ipotesi combinate insieme ci permettono di scrivere, in alternativa alla eq. 4.10:

$$\frac{n_n}{n_p(0)} = e^{q(V_j - V_D)/kT} \quad (4.14)$$

Dividendo la 4.10 per la 4.14 e tenendo conto che n_p della eq. 4.10 corrisponde nella figura 4.5 a n_{p0} si ha

$$n_p(0) = n_{p0} e^{V_D/V_T} \quad (4.15)$$

Analogamente, per le lacune

$$p_n(0) = p_{n0} e^{V_D/V_T} \quad (4.16)$$

Sostituendo le equazioni 4.15 e 4.16 nelle 4.11 e 4.12, derivando rispetto ad x ed utilizzando le 3.7, si ottiene la densità di corrente di diffusione associata con l'eccesso di elettroni e lacune iniettati attraverso la giunzione:

$$J_n \Big|_{x=0} = +q D_n \frac{dn_p}{dx} \Big|_{x=0} = \frac{q D_n n_{p0}}{L_n} \left[e^{V_D/V_T} - 1 \right] \quad (4.17)$$

$$J_p \Big|_{x=0} = -q D_p \frac{dp_n}{dx} \Big|_{x=0} = \frac{q D_p p_{n0}}{L_p} \left[e^{V_D/V_T} - 1 \right] \quad (4.18)$$

La corrente, interamente di diffusione di portatori minoritari a ridosso della giunzione, si trasforma gradualmente per effetto della ricombinazione, allontanandosi, in corrente di deriva di portatori maggioritari (supponendo, però, che questa deriva avvenga sotto l'azione di un campo elettrico trascurabile); il valore totale della corrente attraverso ogni sezione del diodo a diverse distanze dalla giunzione, somma dei contributi di diffusione e deriva, è costante.

La corrente totale I_D che scorre attraverso il diodo è data quindi dalla somma dei due contributi J_n e J_p dei due diversi tipi di portatori di carica:

$$I_D = A \cdot (J_n + J_p) = I_S \left(e^{V_D/V_T} - 1 \right) \quad (4.19)$$

dove A è l'area della giunzione e

$$I_S = qA \left(\frac{p_{n0}}{L_p} D_p + \frac{n_{p0}}{L_n} D_n \right) = qA n_i^2 \left(\frac{D_p}{N_D L_p} + \frac{D_n}{N_A L_n} \right) \quad (4.20)$$

4.6.2 Polarizzazione inversa

Tutte le considerazioni ed i passaggi matematici relativi alla condizione di polarizzazione diretta di un diodo possono essere ripetuti nel caso di polarizzazione inversa. In questo caso le distribuzioni dei portatori minoritari saranno quelle raffigurate in

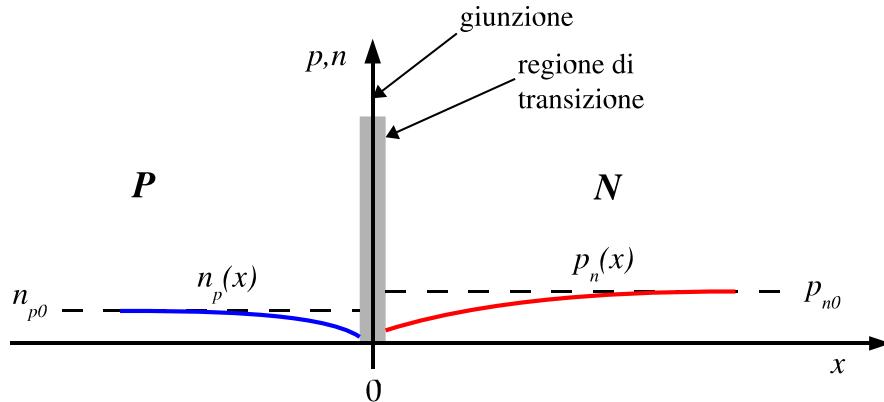


Figura 4.6: Concentrazione dei portatori minoritari per un diodo inversamente polarizzato.

fig. 4.6, con un impoverimento delle concentrazioni in prossimità della giunzione. In fondo comunque si arriva sempre alle equazioni 4.19 e 4.20, che quindi descrivono idealmente, cioè nel campo di validità delle approssimazioni fatte, il comportamento di un diodo a giunzione in ogni condizione di polarizzazione.

4.7 Equazione della giunzione

L'eq. 4.19 descrive in modo abbastanza accurato la relazione $I - V$ di un diodo a giunzione reale a patto di introdurre alcune correzioni per compensare, almeno in parte, le approssimazioni fatte durante la derivazione. La correzione più importante consiste nella introduzione di un fattore η , di valore compreso tra 1 e 2, a denominatore dell'esponente, per cui l'eq. 4.19 diventa:

$$I_D = I_S \left(e^{V_D/\eta V_T} - 1 \right) \quad (4.21)$$

In fig. 4.7 è riportato il grafico della *equazione della giunzione* 4.21 per $I_S = 1 \text{ nA}$, $\eta = 1.5$ e $V_T = 26 \text{ mV}$. V_D è la tensione di polarizzazione applicata al diodo

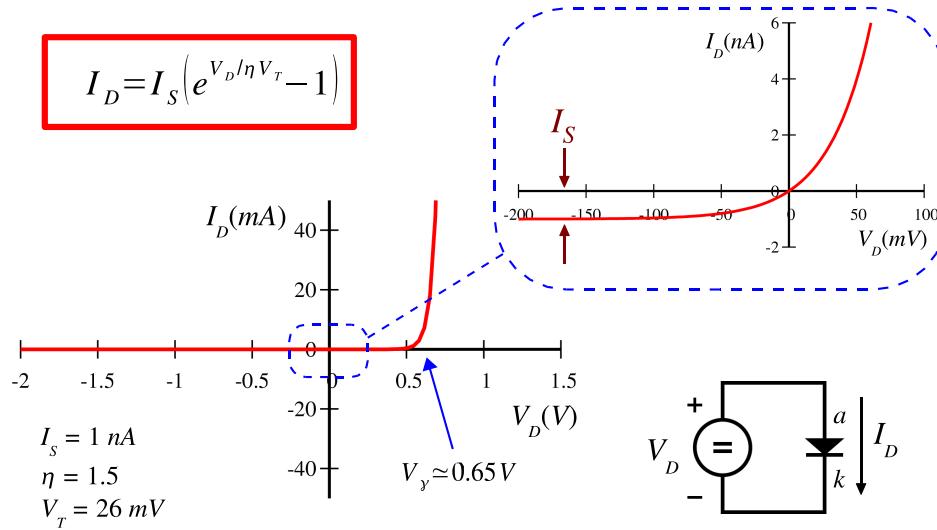


Figura 4.7: Relazione $I - V$ per una giunzione $P-N$ con $I_S = 1 \text{ nA}$, $\eta = 1.5$, $V_T = 26 \text{ mV}$.

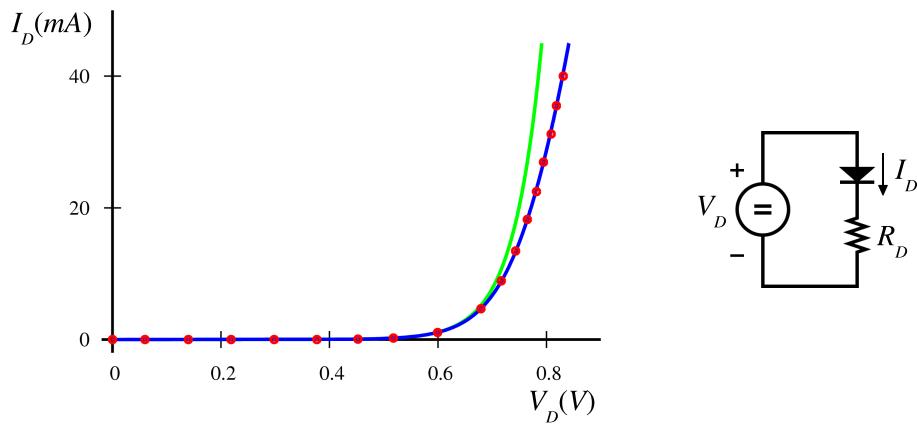
(tensione dell'anodo rispetto al catodo) e I_D è la corrente che scorre nel diodo da anodo a catodo. I_S è la *corrente di saturazione inversa* (eq. 4.20); il nome deriva dal fatto che in condizione di contropolarizzazione la corrente inversa, già con tensioni V_D di poche centinaia di mV , *satura* al valore asintotico I_S , come si può vedere nell'ingrandimento della regione intorno allo zero.

Nel grafico è indicata la tensione $V_\gamma \approx 0.65 \text{ V}$ in corrispondenza del punto in cui la corrente I_D comincia a crescere in maniera significativa. Se si osservano le curve $I_D(V_D)$ per altre giunzioni $P-N$, ad esempio per vari diodi al Silicio di modello diverso, ciascuna sulla scala di correnti per cui il diodo è stato realizzato, ci si rende conto che presentano tutte il ginocchio nell'intervallo $0.6 \dots 0.8 \text{ V}$. La ragione di questa uniformità va ricercata nel fatto che il rapporto I_D/I_S tende ad essere costante e tipico del materiale semiconduttore, quando i valori di I_D siano sulla scala dei valori massimi di corrente che la giunzione può sostenere: diodi in grado di sostenere correnti più elevate presentano valori di I_S proporzionalmente più elevati, principalmente perché hanno superfici di giunzione maggiori. Inoltre la dipendenza esponenziale di I_D da V_D dà luogo ad una variazione della tensione di soli 300 mV (con $\eta = 1$) per una

escursione della corrente di cinque ordini di grandezza. Il valore di V_γ per un diodo a giunzione al germanio è di circa 200 mV.

4.7.1 Diodo in conduzione

Nella fig. 4.8 è riportato l'andamento previsto dalla equazione della giunzione 4.21 in condizioni di polarizzazione diretta a confronto con il comportamento di un diodo reale (diodo per piccoli segnali per uso generale tipo 1N4151). I punti in rosso



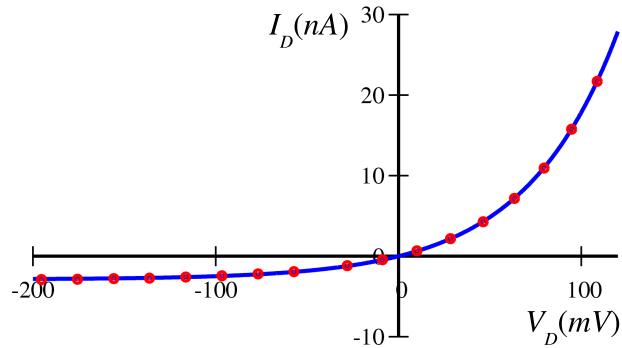


Figura 4.9: Confronto tra eq. 4.21 e diodo reale per valori piccoli di polarizzazione. I punti rossi indicano i valori misurati sullo stesso diodo di fig. 4.8; la curva blu corrisponde ai parametri $I_S = 2.9 \text{ nA}$ ed $\eta = 1.95$.

valore asintotico I_S per polarizzazioni di poche centinaia di mV e poi rimanga rigorosamente costante, nel diodo reale si vede una certa dipendenza di I_R dalla tensione di contropolarizzazione.

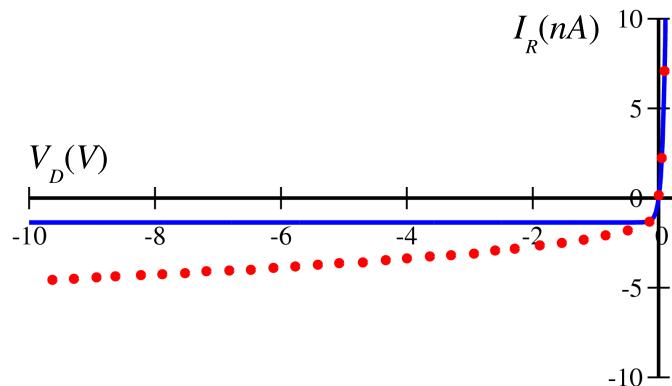


Figura 4.10: Corrente inversa I_R in un diodo reale in condizioni di contropolarizzazione. Mentre l'eq. 4.21 prevede $I_R = -I_S$ per valori di polarizzazione $V_D < -200...300 \text{ mV}$ (curva blu), il diodo reale mostra una dipendenza dell'ordine di 0.2 nA/V dalla tensione di contropolarizzazione (punti rossi).

La discrepanza è dovuta al fatto di avere trascurato la variazione dello spessore della regione di transizione al variare della tensione di contropolarizzazione (vedi par. 4.18.1). Aumentando lo spessore e quindi il volume della regione di transizione aumenta il contributo alla corrente inversa I_R dovuto alla generazione di coppie lacuna-elettrone che avviene continuamente a ritmo costante in tutto il semiconduttore per effetto della agitazione termica.

Nelle applicazioni pratiche del diodo, questa discrepanza di solito non è molto importante; nel grafico di fig. 4.8 per valori di V_D inferiori a V_γ la corrente I_D

scende rapidamente a valori talmente piccoli che non è possibile neanche apprezzarla. Rispetto ai mA della polarizzazione diretta non fa differenza qualche nA in più o in meno in contropolarizzazione.

In alcune situazioni, però, la corrente di contropolarizzazione di una giunzione $P-N$ o la sua dipendenza dalla tensione è il parametro più importante, come ad esempio, nel caso della giunzione *gate-source* di un *j-fet* in un amplificatore differenziale (vedi par. 11.9): da questa caratteristica dipende in massima parte il valore dell'impedenza di ingresso dell'amplificatore.

4.8 Contatti ohmici

Nel circuito di fig. 4.4 sono presenti due ulteriori giunzioni, tra i fili metallici e le regioni P ed N del semiconduttore. Queste regioni devono presentare un comportamento il più possibile *ohmico* ed una resistenza di contatto bassa. In realtà anche alla giunzione tra metallo e semiconduttore si forma una regione di svuotamento ed una barriera di potenziale, analogamente a quanto succede con la giunzione $P-N$. Se però si ha l'accortezza di aumentare opportunamente il livello di droggaggio del semiconduttore in prossimità del metallo si ottengono regioni di svuotamento talmente sottili da risultare in definitiva inefficaci ad impedire il passaggio dei portatori di carica. Una giunzione di questo tipo risulta *non polarizzabile*: solo minime variazioni sono richieste alla altezza della barriera di potenziale per consentire il passaggio di una corrente. Quindi, come indicato in figura 4.4, un potenziale di polarizzazione esterno applicato ad un diodo va a localizzarsi quasi tutto sulla giunzione $P-N$ e solo in misura trascurabile sulle due giunzioni semiconduttore-metallo.

4.9 Coefficiente di temperatura del diodo a giunzione

Invertendo l'equazione 4.21 si ottiene la tensione V_D ai capi del diodo in funzione della corrente I_D che lo percorre. In condizione di polarizzazione diretta:

$$V_D = \eta V_T \log \left(\frac{I_D}{I_S} + 1 \right) \simeq \eta V_T \log \frac{I_D}{I_S} \quad (4.22)$$

Per un dato valore costante di I_D , V_D dipende dalla temperatura tramite V_T (eq. 3.10) e I_S (eq. 4.20):

$$\frac{\partial V_D}{\partial T} \Big|_{I_D} = \frac{\eta V_T}{T} \log \frac{I_D}{I_S} - \eta V_T \frac{1}{I_S} \frac{dI_S}{dT} \quad (4.23)$$

Il primo termine a secondo membro è eguale a V_D/T ($\simeq +2.2 \text{ mV}/^\circ\text{K}$ a temperatura ambiente) ed è positivo. Il secondo termine è più complesso da calcolare perché I_S dipende dalla temperatura sia tramite n_i^2 (eq. 3.3) sia tramite i coefficienti di diffusione D_n e D_p (il coefficiente di diffusione diminuisce all'aumentare della temperatura). L'effetto prevalente è dato dall'aumento molto veloce di n_i^2 con la temperatura, come previsto nella eq. 3.3, per cui a temperatura ambiente il secondo termine è negativo e di valore circa doppio del primo. In totale, con i dati caratteristici del Silicio, si ha:

$$\frac{\partial V_D}{\partial T} \Big|_{I_D} \simeq -2.3 \text{ mV}/^\circ\text{K} \quad (4.24)$$

Questo risultato vale per tutte le giunzioni *P-N* al Silicio direttamente polarizzate, compresa la giunzione base-emettitore del transistor bjt, come si vedrà nel cap. 7.

4.10 Una applicazione dei diodi: il raddrizzatore

Dal grafico tensione - corrente in fig. 4.7 si vede che in condizioni di polarizzazione inversa la corrente che scorre nel diodo è praticamente nulla, mentre in condizioni di polarizzazione diretta si ha una caduta di potenziale ai capi del diodo dell'ordine di 0.65 V, quasi indipendente dalla corrente. Si sfrutta questa caratteristica nei circuiti *raddrizzatori*, con lo scopo di convertire la corrente alternata, ad es. quella proveniente dalla rete di distribuzione dell'energia elettrica, in corrente continua, necessaria ad es. per la alimentazione dei circuiti elettronici.

Il principio di funzionamento è illustrato nella fig. 4.11. Nel circuito in fig. 4.11a

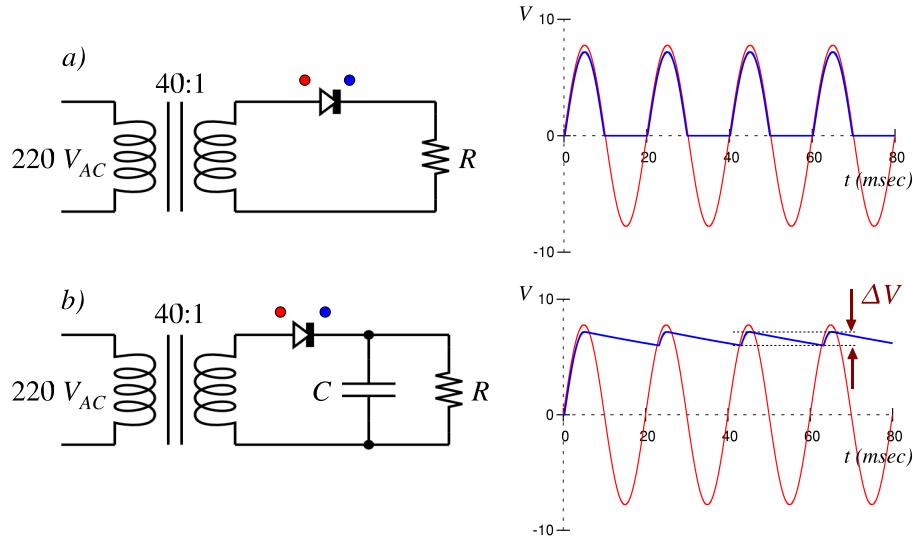


Figura 4.11: Circuito raddrizzatore ad una semionda. Il trasformatore riduce la tensione della rete di distribuzione elettrica ($220 \text{ V}_{\text{RMS}}$) al valore richiesto, in questo caso 5.5 V. Il diodo consente il passaggio della corrente solo durante la semionda positiva. I grafici riportano la tensione a monte del diodo, sull'anodo, in rosso, e a valle del diodo, sul catodo, in blu. Durante la conduzione le due tensioni differiscono di circa 0.65 V ($R = 100\Omega$, $C = 1000\mu\text{F}$; i diodi hanno le stesse caratteristiche di fig. 4.7).

si ha passaggio di corrente nella resistenza R solo durante la semionda positiva: si ottiene una corrente pulsante con valore medio diverso da zero. Nel circuito in fig. 4.11b l'aggiunta del condensatore C in parallelo ad R trasforma la corrente pulsante in una corrente continua cui è sovrapposta una certa ondulazione residua (*ripple*). Nell'ipotesi, abbastanza comune nella pratica, che la costante di tempo RC sia molto maggiore del periodo Δt della tensione alternata e che il tempo di conduzione del diodo sia invece trascurabile rispetto a Δt , il livello di ondulazione residua ΔV può essere stimato calcolando la scarica del condensatore tra un picco ed il successivo:

$$\Delta V \simeq \frac{V}{RC} \Delta t \quad (4.25)$$

I grafici in fig. 4.11 riportano in rosso la tensione misurata sull'anodo del diodo ed in blu quella sul catodo. Durante la conduzione le due tensioni differiscono di $V_\gamma \simeq 0.65 V$.

In figura 4.12 è riportato un circuito che permette di sfruttare tutte e due le semionde della corrente alternata, riducendo a metà il tempo di scarica del condensatore e di conseguenza anche l'ondulazione residua (a parità di R e di C).

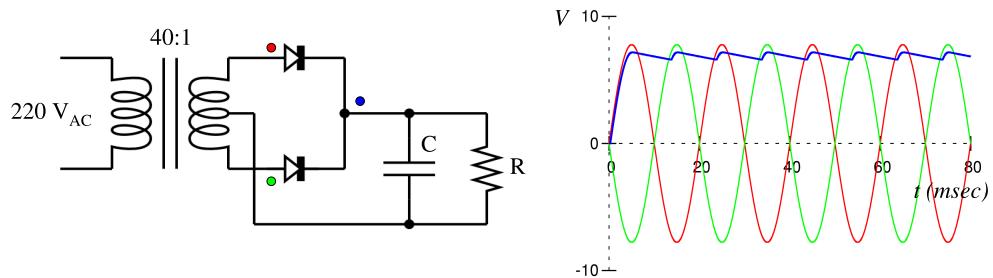


Figura 4.12: Circuito raddrizzatore a due semionde.

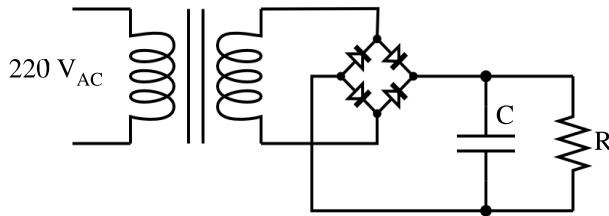


Figura 4.13: Circuito raddrizzatore a ponte.

Il circuito in fig. 4.13, detto *raddrizzatore a ponte*, consente egualmente di sfruttare entrambe le semionde senza dover ricorrere ad un trasformatore con avvolgimento secondario dotato di presa centrale. Si osservi però che in questo caso la caduta di potenziale sui diodi sarà pari a $2V_\gamma$.

4.11 Tensione inversa di rottura (*Reverse Breakdown Voltage*)

Aumentando il valore della tensione di polarizzazione inversa di un diodo, si raggiungerà un punto in cui la corrente inversa I_R comincerà ad aumentare in maniera estremamente rapida (fig. 4.14). Il valore V_{BR} di tensione a cui questo avviene prende il nome di *tensione inversa di rottura* (*reverse breakdown voltage*) e varia da pochi Volt fino a centinaia o migliaia di Volt, a seconda delle caratteristiche del diodo. I meccanismi fisici di questo comportamento sono due:

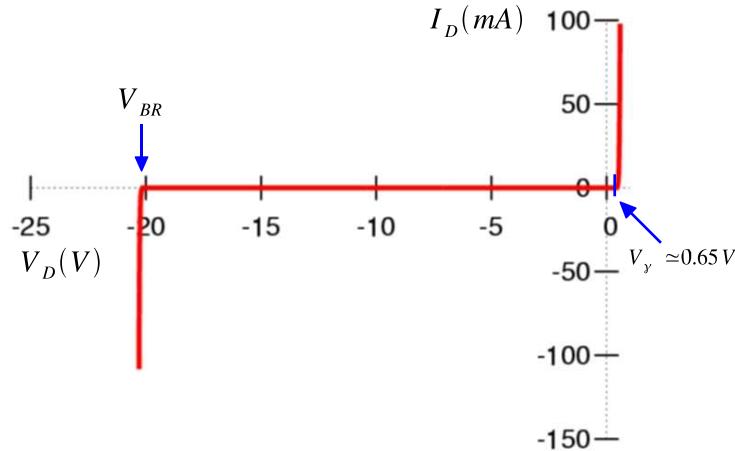


Figura 4.14: Reverse breakdown di una giunzione P-N.

a) Moltiplicazione a valanga dei portatori: il campo elettrico nella regione di transizione diventa sufficientemente intenso da fornire ai portatori tra un urto ed il successivo abbastanza energia da poter provocare la rottura di un legame covalente; ogni nuovo portatore prodotto in questo modo partecipa a sua volta al meccanismo con la produzione di altri portatori in un processo a valanga.

b) Effetto Zener: il campo elettrico nella regione di transizione diventa sufficientemente intenso da provocare direttamente la rottura di legami covalenti e quindi la generazione di una grande quantità di coppie elettrone - lacuna.

4.12 Effetti termici nel diodo

Contrariamente a quanto può fare pensare il termine “tensione inversa di rottura”, l’instaurarsi del regime *a*) oppure *b*) del paragrafo precedente non implica di per sé la distruzione del dispositivo: una volta che la tensione di polarizzazione sia tornata a valori (in modulo) più piccoli della tensione V_{BR} il diodo torna nello stato di interdizione e la corrente torna a valori vicini a I_S . La distruzione è però possibile, anzi in genere abbastanza probabile, specialmente con i diodi raddrizzatori, a causa degli effetti termici. Per il diodo descritto in fig. 4.14, in condizioni di polarizzazione diretta si ha, ad esempio, con una corrente $I_D = 100$ mA una produzione di calore $W_F = 0.7$ V · 0.1 A = 70 mW; supponiamo che il dispositivo sia stato dimensionato per riuscire a dissipare questo calore. Se in qualche circostanza la tensione V_D inversa dovesse arrivare a superare, in modulo, il valore V_{BR} , portando la corrente a -100 mA, avremmo una produzione di calore $W_R = 20$ V · 0.1 A = 2 W, quasi 30 volte superiore a W_F . Se questa non è una condizione operativa prevista, difficilmente il dispositivo sarà in grado di dissipare anche questa produzione di calore. Si avrà come conseguenza un surriscaldamento della giunzione ed un degrado delle caratteristiche o la distruzione completa del dispositivo. Le temperature limite di giunzione che non devono essere superate (ma preferibilmente nemmeno raggiunte) sono per il Germanio $\simeq 70 \dots 90^\circ C$ e per il Silicio $\simeq 180 \dots 200^\circ C$.

A regime la temperatura T_j della giunzione è, secondo la legge di Fourier per la conduzione del calore,

$$T_j = T_c + P_D \cdot R_\theta. \quad (4.26)$$

dove T_c è la temperatura del contenitore del dispositivo, P_D è la potenza dissipata e R_θ la resistenza termica tra giunzione e contenitore. A sua volta il contenitore dovrà dissipare la stessa potenza P_D verso l'ambiente. La resistenza termica tra contenitore ed ambiente dipende fortemente dalle condizioni ambientali (velocità di circolazione dell'aria, presenza di scambiatori di calore ecc.). Nella tabella 4.1 sono riportate le caratteristiche termiche di alcuni dei contenitori per diodi e transistor che compaiono nella fig. 3.3.

	$R_{\theta ja}$ (°C/W)	$R_{\theta jc}$ (°C/W)	P (W)	P_∞ (W)
a) diodo per piccoli segnali	350		0.5	
b) diodo raddrizzatore	175		1.0	
c) transistor di piccola potenza	580	150	0.3	1.2
d) transistor di piccola potenza	220	60	0.8	3
e) transistor di potenza			1.5	115

Tabella 4.1: Caratteristiche termiche dei contenitori di alcuni dei dispositivi di fig. 3.3. Nella prima colonna è riportato il riferimento alla figura; $R_{\theta ja}$ e $R_{\theta jc}$ sono rispettivamente la resistenza termica tra giunzione e ambiente e tra giunzione e contenitore, in °C/W. P è la potenza (in Watt) che può essere dissipata dal dispositivo in aria libera e P_∞ con un radiatore di dimensioni infinite, cioè con resistenza termica nulla tra contenitore e ambiente. I dati si riferiscono ad una temperatura ambiente di 25°C.

4.13 Circuiti non lineari

Un circuito contenente diodi semiconduttori, le cui relazioni tensione-corrente sono fortemente non lineari, sarà descritto da equazioni o sistemi di equazioni non lineari.

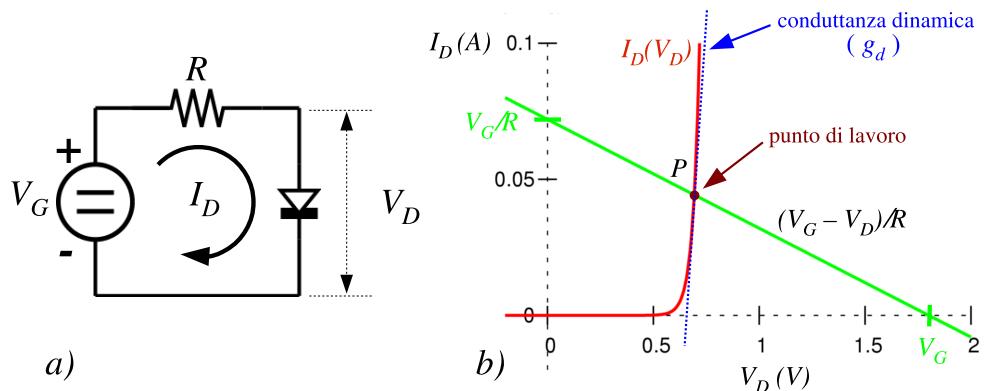


Figura 4.15: Circuito ad una sola maglia contenente un diodo semiconduttore.

Anche per un circuito semplice come quello in fig. 4.15a, l'equazione della maglia

$$I_D \cdot R + V_D(I_D) = V_G \quad (4.27)$$

non ha soluzione analitica (la relazione $V_D(I_D)$ è la 4.22 che si ottiene invertendo l'equazione della giunzione 4.21).

La figura 4.15b riporta l'equazione 4.27 in forma grafica. La curva in rosso rappresenta la relazione tensione-corrente del diodo $I_D(V_D)$, la retta in verde, detta *retta di carico*, corrisponde alla equazione lineare $I_D = (V_G - V_D)/R$. Il punto di intersezione P corrisponde ai valori di I_D e V_D che soddisfano la relazione $I_D = (V_G - V_D)/R = I_D(V_D)$, cioè la 4.27.

In passato il metodo grafico è stato utilizzato ampiamente per trovare le soluzioni statiche dei circuiti non lineari. Ad es., nel caso del raddrizzatore di fig. 4.11a il grafico della tensione ai capi della resistenza R corrisponde alla soluzione della eq. 4.27 per i valori di tensione V_G forniti istante per istante dal trasformatore (considerando il trasformatore come un generatore di tensione ideale e trascurando le capacità del diodo, descritte al par. 4.18). In pratica, traslando nella figura 4.15b la retta di carico parallelamente a se stessa si ottiene rapidamente la posizione del punto P in funzione dei valori di V_G .

Oggi il metodo grafico è stato convenientemente sostituito da programmi di calcolo numerico per computer. Uno di questi (Spice) verrà descritto nel capitolo 18.

Quando un circuito contiene elementi capacitivi od induttivi il suo funzionamento non può essere descritto solo mediante equazioni algebriche. Ad esempio, il raddrizzatore di fig. 4.11b, contenente il condensatore C , è descritto da un sistema di equazioni differenziali non lineari in cui compare la carica Q del condensatore e la sua derivata Q' rispetto al tempo:

$$Q/C + V_D(I_D) = V_G \quad (4.28)$$

$$Q' + Q/RC = I_D \quad (4.29)$$

Anche questi sistemi di equazioni possono essere risolti mediante integrazione numerica da un programma di calcolo come Spice.

4.14 Circuito linearizzato

In moltissimi casi il funzionamento di un circuito contenente elementi non lineari è limitato a *piccoli segnali*, cioè a piccole escursioni delle tensioni e delle correnti intorno a valori costanti di riferimento. Questo *non* è il caso dei circuiti raddrizzatori delle figure 4.11, 4.12 e 4.13. In questo caso i diodi passano continuamente dallo stato di interdizione completa fino alla conduzione piena e viceversa: sono circuiti che devono trattare *grossi segnali*. Il circuito in fig. 4.15 può invece essere considerato un circuito per piccoli segnali nel caso in cui la tensione V_G sia una tensione costante con sovrapposte variazioni nel tempo piccole rispetto a V_G stessa.

Si adotta la seguente convenzione, ampiamente utilizzata in tutta la letteratura:

- Le lettere maiuscole con indice maiuscolo indicano valori costanti: V_G , V_D , I_D sono tensioni e correnti continue.

- Le lettere minuscole con indice minuscolo indicano valori variabili nel tempo: i_d , v_g , v_d . Di solito queste tensioni e correnti sono molto più piccole delle corrispondenti costanti V_G , V_D , I_D .
- Le lettere minuscole con indice maiuscolo indicano la somma di un termine costante ed uno dipendente dal tempo:

$$v_G = V_G + v_g$$

$$v_D = V_D + v_d$$

$$i_D = I_D + i_d$$

- Le lettere maiuscole con indice minuscolo indicano il valore efficace della componente variabile nel tempo oppure modulo e fase di una componente armonica:
 V_g è il valore efficace di v_G
 I_d è il valore efficace di i_D
 $V_g(\omega)$ è un numero complesso che rappresenta ampiezza e fase della componente di frequenza angolare ω della tensione $v_g(t)$.

Con questa notazione l'eq. 4.27 diventa

$$i_D \cdot R + v_D(i_D) = v_G \quad (4.30)$$

Se si suppone $v_g \ll V_G$ (e di conseguenza $v_d \ll V_D$ e $i_d \ll I_D$), l'eq. 4.30 può essere riscritta sviluppando $v_D(i_D)$ in serie di Taylor ed arrestando lo sviluppo al primo ordine. Si ottiene l'equazione lineare:

$$(I_D + i_d) \cdot R + V_D + r_d \cdot i_d = V_G + v_g \quad (4.31)$$

La quantità $r_d = dv_D/di_D$ ($= v_d/i_d$ secondo la convenzione adottata) prende il nome di resistenza dinamica o resistenza incrementale; la sua inversa $g_d = i_d/v_d$ è la conduttanza dinamica (retta blu tratteggiata in fig. 4.15).

Indicando con I_P e V_P i valori di I_D e V_D corrispondenti al punto P di fig. 4.15 e sottraendo dalla eq. 4.31 la relazione tra i termini costanti data dalla eq. 4.27

$$I_P \cdot R + V_P(I_P) = V_G \quad (4.32)$$

rimane l'equazione

$$i_d \cdot R + i_d \cdot r_d = v_g \quad (4.33)$$

Quindi il circuito di fig. 4.16a può essere studiato separando il comportamento in corrente continua al punto $P = (I_P, V_P)$ descritto dall'eq. 4.32 dal comportamento per le variazioni (piccole) di tensioni e correnti intorno al punto P , descritto dall'eq. 4.33 e corrispondente al circuito di fig. 4.16b.

Il punto P viene indicato indifferentemente come *punto di lavoro (operating point)* oppure *punto di riposo (quiescent point)*. Questa grossolana ambiguità sindacale deriva dal fatto che il punto P corrisponde ai valori di tensione e di corrente che si hanno quando il circuito è in funzione (cioè non è spento), e quindi è un *punto di lavoro*; d'altra parte se tensione e corrente corrispondono esattamente ai valori del

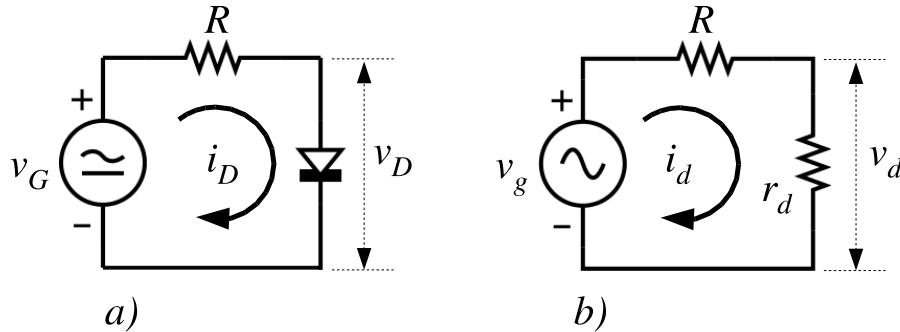


Figura 4.16: Linearizzazione di un circuito: a) circuito contenente un elemento non lineare (diodo); b) circuito linearizzato, per i soli segnali.

punto *P* nessun segnale è presente nel circuito oltre le polarizzazioni ($v_g = 0$) e allora il circuito non sta veramente *lavorando* sui segnali, e quindi *punto di riposo*.

L'operazione di linearizzazione di un circuito non è solo una questione di convenienza matematica: in moltissime situazioni si ha interesse ad un comportamento il più possibile lineare da parte di un circuito contenente elementi non lineari e molti circuiti di questo tipo hanno effettivamente un comportamento discretamente lineare se si sceglie in maniera opportuna il punto di lavoro e l'ampiezza dei segnali.

4.15 Conduttanza dinamica del diodo

Il valore della conduttanza dinamica del diodo può essere calcolato derivando l'equazione 4.21:

$$g_d = \frac{dI_D}{dV_D} = \frac{I_S}{\eta V_T} \cdot e^{V_D/\eta V_T} \quad (4.34)$$

Ricavando l'espressione per $e^{V_D/\eta V_T}$ dalla 4.21 stessa e sostituendola nella 4.34 si ottiene

$$g_d = \frac{I_D + I_S}{\eta V_T} \quad (4.35)$$

I valori di g_d che si ottengono sono profondamente diversi a seconda dello stato di polarizzazione del diodo.

4.15.1 Diodo direttamente polarizzato

$$I_D \gg I_S \quad g_d \simeq \frac{I_D}{\eta V_T} \quad (4.36)$$

In questa condizione la corrente I_S è assolutamente trascurabile rispetto a I_D . La conduttanza dinamica g_d risulta proporzionale alla corrente I_D di polarizzazione ed il fattore di proporzionalità è

$$\frac{1}{\eta V_T} \simeq 30 \frac{mA/V}{mA} \quad (4.37)$$

Le dimensioni della costante sono state indicate come $mA/V/mA$ (ovviamente eguale a V^{-1}) per sottolineare il fatto che g_d ammonta a circa 30 mA/V per ogni mA di corrente I_D . Si noti come il valore 30 della costante prescinda da qualsiasi considerazione sul modello di diodo e dipenda solo dalla temperatura. In altre parole, la costante è la stessa per ogni diodo. Il fattore η introduce un campo di variazione dell'ordine di 2; d'altra parte il campo di validità della eq. 4.36 copre circa 10 ordini di grandezza di I_D .

4.15.2 Diodo non polarizzato

$$V_D = 0 \quad I_D = 0 \quad g_d = \frac{I_S}{\eta V_T} \quad (4.38)$$

La conduttanza g_d dipende dal valore di I_S , ma assume comunque valori molto piccoli, dell'ordine di $10^{-8} \dots 10^{-12} A/V$.

4.15.3 Diodo contropolarizzato

Per valori di tensione di contropolarizzazione sufficientemente elevati, I_D tende a $-I_S$; $I_D + I_S$ tende a 0 e quindi anche g_d tende a 0. La resistenza dinamica $r_d = 1/g_d$ tende a infinito. Nella pratica si riscontrano valori molto più bassi di quelli teorici, compresi comunque tra le migliaia ed i milioni di $M\Omega$. Infatti, come visto al par. 4.7.2, la corrente nel diodo contropolarizzato invece di saturare a $-I_S$ mostra una certa dipendenza da V_D . Nell'esempio di fig. 4.10 intorno a $V_D = -5 V$ si ha $dI_D/dV_D \simeq 0.23 nA/V$ e quindi $r_d \simeq 4.3 G\Omega$ invece dei 10^{15} teorici $G\Omega$ previsti dalla equazione 4.34.

Quando la tensione di contropolarizzazione raggiunge il valore di *reverse breakdown* V_{BR} la conduttanza dinamica g_d passa rapidamente da valori praticamente nulli a valori molto elevati, come indicato dalla ripidità della curva $I_D(V_D)$ in fig. 4.14. Questa caratteristica è sfruttata nei diodi stabilizzatori di tensione, descritti al paragrafo seguente.

4.16 Diodi stabilizzatori di tensione

Molto spesso nei circuiti elettronici è necessario poter disporre di tensioni di valore noto e stabili nel tempo. Un diodo in condizione di *reverse breakdown* può essere utilizzato come stabilizzatore di tensione. Diodi costruiti per questo scopo sono chiamati genericamente *diodi zener*, anche quando il loro meccanismo di funzionamento non corrisponde strettamente all'effetto Zener.

Nel circuito in fig. 4.17a il generatore fornisce una tensione v_G continua, ma soggetta a variazioni nel tempo; le variazioni si suppongono tali che v_G risulti sempre opportunamente maggiore (in modulo) di V_{BR} . Il diodo D_Z (il simbolo, modificato in modo da richiamare una Z , indica che si tratta di un diodo zener) si troverà contropolarizzato ed in condizione di breakdown. Il punto di lavoro (o di riposo) $Q = (V_Z, I_Z)$ del circuito è determinato graficamente in fig. 4.17b. Quando la tensione v_G varia, anche la tensione v_Z varia, come indicato dalla retta tratteggiata parallela alla retta di carico. È immediatamente evidente però che le variazioni di v_G avranno

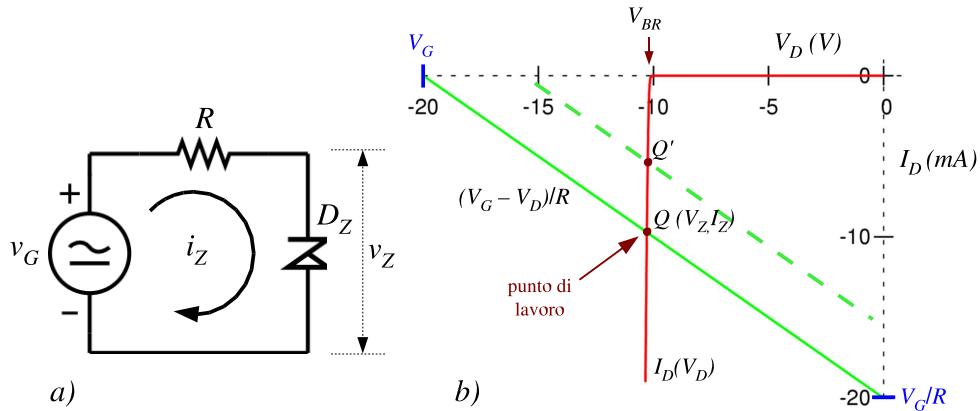


Figura 4.17: Diodo zener utilizzato come stabilizzatore di tensione.

poco effetto su v_Z , tanto meno quanto più la curva $I_D(V_D)$ sarà ripida a sinistra di V_{BR} . In fig. 4.18a è riportato il circuito linearizzato equivalente a quello di fig.

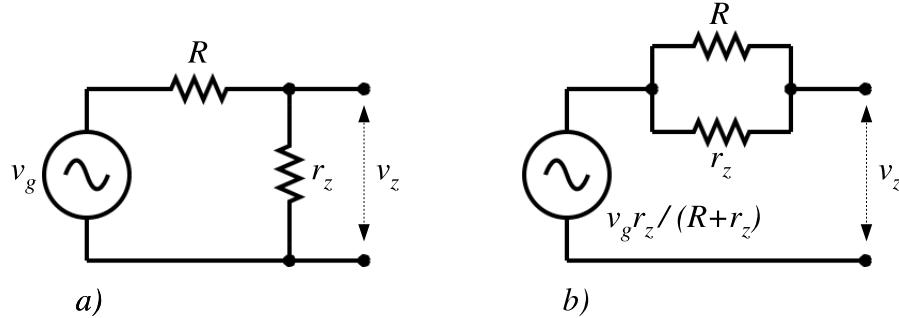


Figura 4.18: Circuito linearizzato dello stabilizzatore di tensione con diodo zener (a) e suo equivalente di Thèvenin (b).

4.17a, ed in figura 4.18b il suo equivalente di Thèvenin. La resistenza r_z è la resistenza dinamica del diodo zener al punto di lavoro Q .

Il grafico di fig. 4.17b corrisponde ai seguenti valori: $V_G = 20V$, $V_{BR} = 10V$, $R = 1000\Omega$, $r_z = 5\Omega$. La tensione che si ottiene all'uscita è circa $-10.6V$, la metà della tensione di ingresso V_G ; ma le variazioni v_z sovrapposte a V_Z sono ridotte di un fattore circa 200 rispetto a quelle originarie su V_G .

È superfluo osservare che il modello linearizzato del circuito è valido solo fino a che il punto Q si trova nel tratto ripido a sinistra di V_{BR} . Quando V_G raggiunge V_{BR} il modello non è più valido ed il suo utilizzo porta solo a conclusioni assurde.

4.17 Interruttore a diodi per segnali analogici

La grande differenza di resistenza dinamica di un diodo nei due stati di polarizzazione diretta ed inversa può essere utilizzata per realizzare un interruttore per segnali.

Nel circuito in fig. 4.19 il segnale proveniente dal generatore v_g viene trasferito o non trasferito alla resistenza di carico R_L a seconda che i diodi $D_1 \dots D_4$, controllati dal generatore I_C , si trovino nello stato di bassa o alta resistenza dinamica. Il comportamento è analogo a quello del circuito riportato sulla destra: quando l'interruttore S_1 è chiuso, tra generatore e carico viene inserita una resistenza nominalmente nulla ed il segnale giunge dal generatore al carico; quando S_1 è aperto, viene inserita una resistenza nominalmente infinita e tra generatore e carico si ha isolamento completo. I quattro diodi svolgono la funzione di interruttore comandato elettricamente dal segnale di controllo I_C . I diodi $D_1 \dots D_4$ sono direttamente polarizzati quando il

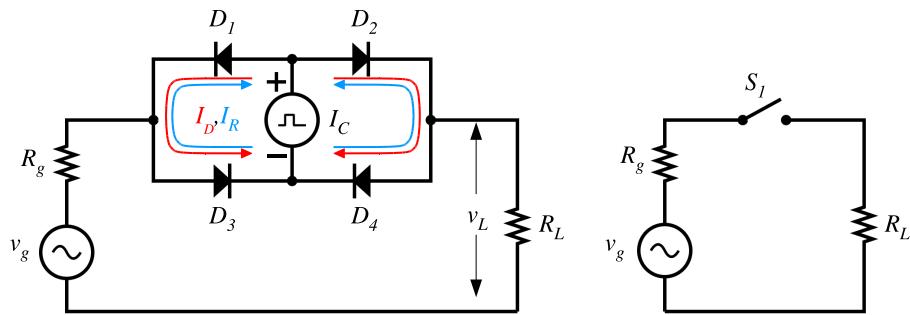


Figura 4.19: Interruttore elettronico per segnali.

generatore I_C fa scorrere una corrente di polarizzazione I_D come indicato dalle frecce rosse nella figura. Anche con correnti di polarizzazione di pochi mA, la resistenza dinamica scende a pochi ohm: con $I_D = 10mA$ si ha $r_d \simeq 3\Omega$ (eq. 4.36).

Con una polarizzazione di polarità opposta invece nei diodi scorre la corrente I_R , corrispondente praticamente alla corrente di saturazione inversa (frecce blu) ed i diodi sono contropolarizzati. La resistenza dinamica in questo caso diventa molto elevata, comunque superiore alle centinaia o migliaia di $M\Omega$ (vedi par. 4.15.3).

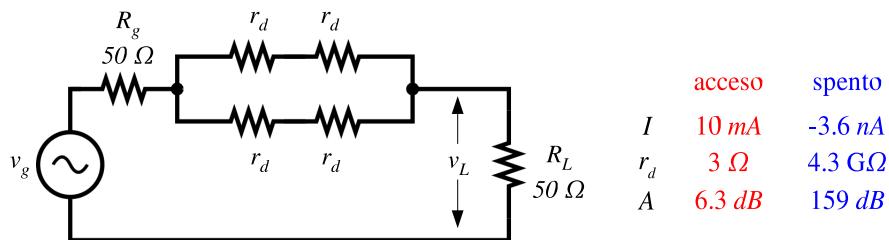


Figura 4.20: Circuito linearizzato equivalente per l'interruttore elettronico a diodi di fig. 4.19.

In fig. 4.20 è riportato il circuito linearizzato equivalente dell'interruttore di fig. 4.19, con un esempio di condizioni operative. Nello stato *acceso* l'attenuazione A è di soli 0.3 dB superiore a quella ideale (6 dB) che si avrebbe con resistenza dinamica nulla per effetto del rapporto di partizione tra resistenza del generatore e resistenza di carico. Nello stato *spento* teoricamente si dovrebbe avere una attenuazione elevatissima ($A = 159$ dB). Nella pratica difficilmente si arriva a superare i 120...140 dB,

a causa di numerosi piccoli effetti parassiti. Di questi, il più importante è la capacità dei diodi, descritta nei paragrafi seguenti, che si viene a trovare in parallelo alla resistenza dinamica e degrada l'attenuazione tanto più quanto più è alta la frequenza del segnale.

Come già visto nei paragrafi precedenti, anche qui la corrente di segnale che scorre attraverso le resistenze r_d dei diodi direttamente polarizzati non è altro che una (piccola) variazione della corrente di polarizzazione impostata dal generatore I_C . È necessario che la corrente del segnale rimanga piccola rispetto a quella di polarizzazione non solo per evitare di modificare lo stato dei diodi da accesi a spenti, ma anche per mantenere sufficientemente bassa la distorsione dovuta ai termini non lineari ignorati nello sviluppo in serie 4.31.

In linea di principio, a causa della simmetria del circuito, il segnale di controllo del generatore I_C non dovrebbe comparire all'uscita ai capi della resistenza R_L . Questo è tanto più vero quanto più i diodi $D_1 \dots D_4$ hanno caratteristiche identiche. I risultati migliori si ottengono utilizzando diodi *monolitici*, realizzati cioè contemporaneamente su uno stesso substrato di Silicio con lo stesso ciclo di fabbricazione.

4.18 Capacità della giunzione

Tutte le considerazioni svolte fino ad ora hanno riguardato solo il comportamento statico dei diodi. Anche quando si sono considerate tensioni e correnti dipendenti dal tempo, come nel circuito di fig. 4.11a, la risposta del diodo è stata considerata comunque istantanea. In realtà si è visto nei paragrafi precedenti che lo stato di polarizzazione determina variazioni nelle distribuzioni di carica in prossimità della giunzione, con trasferimenti che non possono essere istantanei, in quanto richiederebbero correnti di intensità infinita. Poichè la carica accumulata in prossimità della giunzione dipende dal potenziale di polarizzazione, ci dovremo aspettare un comportamento simile a quello di un condensatore.

Anche a questo riguardo, come si è già visto per la conduttanza dinamica, il comportamento del diodo è profondamente diverso a seconda che si trovi in condizione di polarizzazione diretta o inversa.

4.18.1 Diodo contropolarizzato - Capacità di transizione

Nel caso di un diodo contropolarizzato (fig. 4.4b) è possibile, con opportune semplificazioni, scrivere una relazione tra la carica presente nella regione di transizione e la differenza di potenziale tra le regioni P ed N . La distribuzione della carica è stata riportata qualitativamente in fig. 4.2. Si tenga presente però che in questo grafico per le concentrazioni è stata utilizzata una scala logaritmica. Se lo stesso grafico viene riportato con una scala lineare le transizioni tra le zone omogenee delle regioni P ed N e la zona di carica spaziale diventano molto più ripide, al punto che si può considerare con buona approssimazione di avere due distribuzioni uniformi con densità $\rho = -qN_A$ e $\rho = +qN_D$ rispettivamente di spessore x_A nella regione P e x_D nella regione N (fig. 4.21). Integrando l'equazione di Poisson unidimensionale

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\epsilon} \quad (4.39)$$

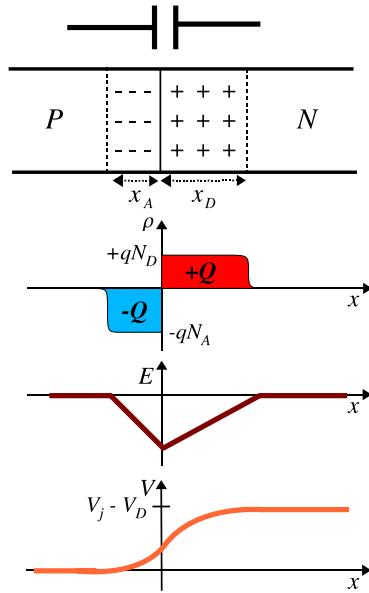


Figura 4.21: Distribuzioni di carica in una giunzione P - N ed andamento del campo elettrico e del potenziale.

sugli spessori x_A e x_D si ottiene per l'altezza totale della barriera di potenziale:

$$V_j - V_D = \frac{q}{2\epsilon} (N_A x_A^2 + N_D x_D^2) \quad (4.40)$$

Poichè la giunzione deve essere nel suo complesso elettricamente neutra, la carica totale Q per unità di superficie contenuta nello spessore x_D deve essere eguale ed opposta a quella contenuta nello spessore x_A :

$$q N_D x_D = q N_A x_A = Q \quad (4.41)$$

Ricavando dalla eq. 4.41 $x_A = Q/(N_A \cdot q)$ e $x_D = Q/(N_D \cdot q)$ e sostituendo nella 4.40 si ottiene:

$$V_j - V_D = \frac{Q^2}{2\epsilon q} \left(\frac{1}{N_D} + \frac{1}{N_A} \right) \quad (4.42)$$

La relazione 4.42 non è lineare, quindi non è possibile definire una capacità del diodo in base al rapporto Q/V_D nel modo abituale. Come nel caso della resistenza dinamica del diodo, anche in questo caso è possibile, anzi spesso utile, definire dalla 4.42 una capacità dinamica o capacità differenziale del diodo contropolarizzato, detta capacità di transizione C_T :

$$C_T = A \frac{dQ}{dV_D} = \frac{C_0}{\sqrt{1 - V_D/V_j}} \quad (4.43)$$

dove A è la superficie della giunzione e

$$C_0 = A \sqrt{\frac{\epsilon q}{2V_j} / \left(\frac{1}{N_D} + \frac{1}{N_A} \right)} \quad (4.44)$$

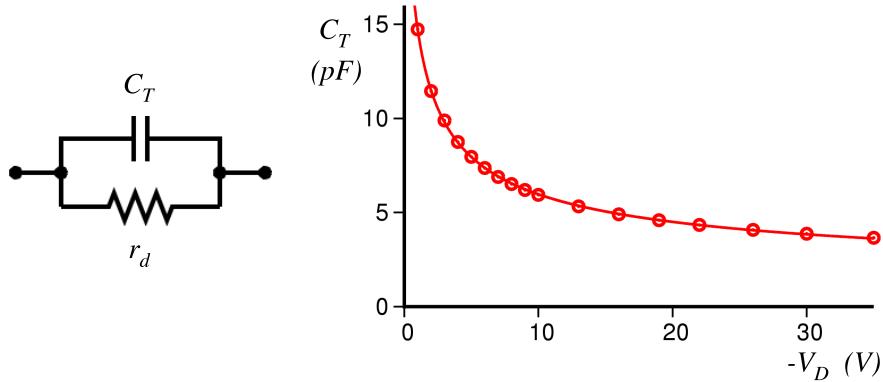


Figura 4.22: A sinistra, circuito lineare equivalente per un diodo contropolarizzato; a destra, capacità di transizione di un diodo di piccola potenza in funzione della tensione di polarizzazione inversa.

Quindi un diodo contropolarizzato equivale, per piccoli segnali intorno al punto di polarizzazione, al circuito lineare in fig. 4.22.

La capacità C_T per un diodo per piccoli segnali è compresa tra una frazione di pF e qualche pF ; per i diodi di potenza si arriva a centinaia o migliaia di pF . In fig. 4.22 è riportata la capacità di transizione di un diodo di piccola potenza: i punti corrispondono ai valori misurati; la linea continua riporta la funzione $C_T = C_0 / \sqrt{1 - V_D/V_J} + C_c$, con $C_0 = 25.5 \text{ } pF$, $V_j = 0.42 \text{ } V$ e $C_c = 0.84 \text{ } pF$. Il termine C_c , che non compare nella eq. 4.44, è la somma di tutte le capacità parassite (terminali, parti delle regioni P ed N non interessate dalla giunzione ecc.).



Figura 4.23: Da sinistra a destra: condensatore variabile meccanico e simboli nei circuiti per il condensatore variabile e per il diodo varicap.

Mentre nel funzionamento del diodo come raddrizzatore o come interruttore la capacità di transizione è un effetto *parassita* che ne limita le prestazioni, nei diodi varicap lo stesso effetto è utilizzato a vantaggio per inserire in un circuito elementi con capacità variabile in grado di sostituire i condensatori variabili meccanici (fig. 4.23).

4.18.2 Diodo in conduzione - Capacità di diffusione

Come visto al par. 4.6.2, in un diodo in conduzione la corrente I_D è una corrente di diffusione legata ad un eccesso di concentrazione dei portatori minoritari iniettati

attraverso la giunzione (fig. 4.5 ed eq. 4.15 ... 4.20). Ad ogni profilo di concentra-

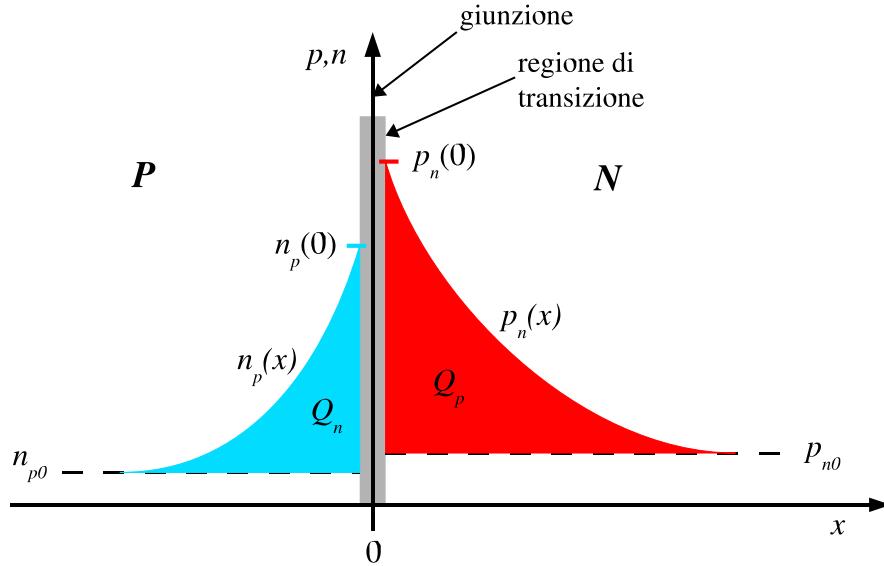


Figura 4.24: Eccesso di portatori minoritari in prossimità di una giunzione $P-N$.

zione $n_p(x)$ e $p_n(x)$ di fig. 4.5 corrisponde un eccesso di portatori di carica ai due lati della giunzione, positivi e negativi, come raffigurato dalle due aree rosse e blu in fig. 4.24:

$$Q_p = \int_0^\infty A q [p_n(x) - p_{n0}] dx \quad (4.45)$$

$$Q_n = \int_{-\infty}^0 A q [n_p(x) - n_{p0}] dx \quad (4.46)$$

La situazione *non* è analoga a quella di un condensatore a facce piane e parallele: la carica associata agli eccessi dei due tipi di portatori da una parte e dall'altra non si bilancia. Anzi, è abbastanza comune in un diodo avere drogaggi delle due regioni P ed N fortemente asimmetrici, per cui il contributo di uno dei due tipi di portatori di carica è prevalente e l'altro trascurabile. Ad esempio, se $N_A \gg N_D$, i portatori maggioritari nella regione P (lacune) avranno una concentrazione molto più alta dei maggioritari nella regione N (elettroni) e contribuiranno in maniera proporzionalmente maggiore alla corrente I_D , secondo le equazioni del par. 4.6.1. Anche gli eccessi di carica Q_p e Q_n associati ai portatori ai due lati della giunzione non saranno eguali, ma piuttosto in proporzione ai singoli contributi alla corrente I_D . Ognuna singolarmente delle due regioni P ed N sarà in prima approssimazione elettricamente neutra. Al par. 4.6.1 abbiamo considerato che i campi elettrici siano trascurabili nelle zone di diffusione; la carica dovuta all'eccesso dei portatori minoritari sarà piuttosto compensata da una piccola ridistribuzione dei maggioritari.

Considerando per semplicità il caso della giunzione asimmetrica in cui il contributo degli elettroni alla corrente I_D è trascurabile, si ha dall'integrale 4.45:

$$Q_p = A q L_p [p_n(0) - p_{n0}] \quad (4.47)$$

Utilizzando la eq. 4.16, la 4.18 e la 4.13:

$$Q_p = A q L_p p_{n0} \left(e^{V_D/qkT} - 1 \right) \quad (4.48)$$

$$= A J_n \frac{L_p^2}{D_p} \quad (4.49)$$

$$= I_D \tau_p \quad (4.50)$$

Ogni variazione del potenziale di polarizzazione del diodo comporta una variazione della corrente I_D e della carica Q_p ; si ha quindi anche qui un comportamento di tipo capacitivo. Anche in questo caso è utile definire una capacità differenziale, o incrementale, detta *capacità di diffusione*:

$$C_D = \frac{dQ_p}{dV_D} = \tau_p \frac{dI_D}{dV_D} = \tau_p g_d \quad (4.51)$$

dove g_d è la conduttanza dinamica del diodo in conduzione vista al par. 4.15.

4.18.3 Confronto tra capacità di transizione e capacità di diffusione

Come si è visto al paragrafo 4.18.1, la capacità di transizione C_T ha un valore dell'ordine dei pF per diodi di uso generale per segnali; dipende dal potenziale di polarizzazione inverso, arrivando ad un rapporto tra valore massimo e minimo generalmente tra 2 e 10; si trova in parallelo alla resistenza dinamica del diodo contropolarizzato, che ha valori molto elevati ($10^9 \dots 10^{12} \Omega$) e può generalmente essere ignorata. La capacità C_T si comporta quindi per i piccoli segnali come un vero condensatore, il cui valore dipende dal potenziale di contropolarizzazione del diodo. Tanto è vero che vi sono diodi costruiti appositamente per essere utilizzati come condensatori variabili in circuiti LC risonanti per alta frequenza (*diodi varicap*). La capacità di diffusione C_D ha caratteristiche completamente diverse. Dalla eq. 4.51 si vede che il valore di C_D dipende fortemente dal punto di lavoro. Nel caso di un diodo per segnali con $\tau_p = 1 \text{ nsec}$, si ha, per $I_D = 100 \mu A$, $C_D = 4 \text{ pF}$ mentre per $I_D = 100 \text{ mA}$ si arriva a $C_D = 4000 \text{ pF}$ (τ_p è caratteristico del semiconduttore e sostanzialmente indipendente dal punto di lavoro). Inoltre C_D si trova in parallelo alla resistenza dinamica r_d con cui forma una costante di tempo τ_p , indipendente dalla polarizzazione.

4.19 Spessore della regione di transizione

Utilizzando la condizione di elettroneutralità della giunzione (4.41)

$$N_A \cdot x_A = N_D \cdot x_D$$

l'eq. 4.40 può essere riscritta come

$$(V_j - V_D) \cdot \frac{2\epsilon}{qN_A} = x_A (x_A + x_D) \quad (4.52)$$

oppure

$$(V_j - V_D) \cdot \frac{2\epsilon}{qN_D} = x_D (x_A + x_D) \quad (4.53)$$

Sommando membro a membro le eq. 4.52 e 4.53 si ottiene

$$(V_j - V_D) \frac{2\epsilon}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right) = (x_A + x_D)^2 \quad (4.54)$$

da cui si ricava lo spessore della regione di transizione

$$w_j = x_A + x_D = \sqrt{\frac{2\epsilon(V_j - V_D)}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)} \quad (4.55)$$

Ad un basso livello di drogaggio ($N_A = N_D \simeq 10^{15} \text{ atomi} \cdot \text{cm}^{-3}$) si ottiene

$$V_j = 580 \text{ mV}$$

$$w_j = 0.36 \mu\text{m}$$

Ad un livello di drogaggio più alto ($N_A = N_D \simeq 10^{18} \text{ atomi} \cdot \text{cm}^{-3}$)

$$V_j = 940 \text{ mV}$$

$$w_j = 0.014 \mu\text{m}$$

4.20 Tempi di commutazione di un diodo

Gli effetti capacitivi in un diodo sono particolarmente importanti durante le *commutazioni*, cioè quando il diodo deve passare rapidamente dallo stato di conduzione a quello di interdizione o viceversa. In fig. 4.25 è riportato come esempio un circui-

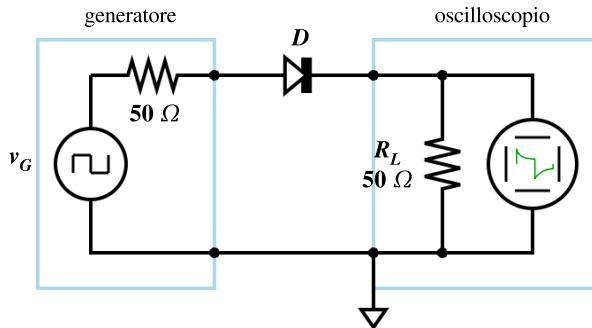


Figura 4.25: Circuito di misura dei tempi di commutazione di un diodo.

to per misurare il comportamento di un diodo durante una commutazione veloce. Il generatore V_G fornisce un'onda quadra di ampiezza $\pm 5 \text{ V}$ (fig. 4.26a). Durante la semionda positiva il diodo si trova in conduzione ed è percorso da una corrente $I_D \simeq (5 \text{ V} - V_\gamma)/100 \Omega \simeq 43 \text{ mA}$. Durante la semionda negativa il diodo è in interdizione ed è percorso dalla corrente di saturazione inversa. Collegando un oscilloscopio in parallelo alla resistenza R_L è possibile misurare la corrente che attraversa il diodo istante per istante.

Nel caso di un diodo ideale senza capacità le commutazioni istantanee e la corrente salterebbe di colpo tra i valori $\simeq 0$ e $\simeq 43 \text{ mA}$ (fig. 4.26b). Nel caso di un

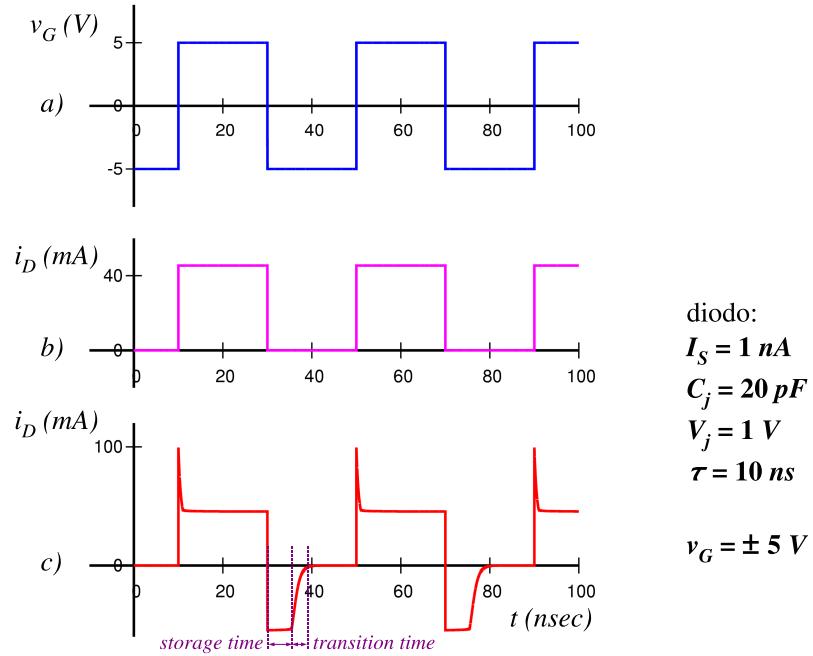


Figura 4.26: Tempi di commutazione di un diodo.

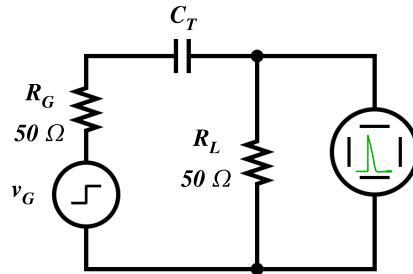


Figura 4.27: Circuito equivalente durante la commutazione del diodo da interdizione a conduzione.

diodo reale, caratterizzato dai parametri riportati in fig. 4.26, l'andamento temporale della corrente è quello che si vede in fig. 4.26c. Al tempo $t = 0$ il diodo è in interdizione ($v_g = -5 \text{ V}$); a $t = 10 \text{ ns}$ la tensione v_g passa da -5 a $+5 \text{ V}$. La corrente, prima di stabilizzarsi al valore di conduzione I_{on} , presenta un brevissimo picco di ampiezza circa $2 \cdot I_{on}$. Il circuito di fig. 4.27 ne spiega la ragione: il diodo contropolarizzato si comporta come un condensatore carico al potenziale $v_g = -5 \text{ V}$. Immediatamente dopo la commutazione, questa tensione si somma a quella del generatore, dando in totale una tensione quasi doppia, chiusa sulla resistenza $R = R_G + R_L$. Quindi la capacità C_T si scarica, con una costante di tempo $R \cdot C_T$, e la corrente si stabilizza al valore di conduzione I_{on} . Poiché la capacità C_T non è costante al variare della tensione di polarizzazione (non siamo nell'approssimazione dei piccoli segnali), il picco di corrente avrà solo approssimativamente la forma di una esponenziale e $R \cdot C_T$ sarà

solo una stima del tempo di scarica.

Al tempo $t = 30 \text{ ns}$ il generatore compie la transizione inversa, da $+5 \text{ V}$ a -5 V ; ci si aspetta che il diodo torni in condizioni di contropolarizzazione e quindi la corrente torni praticamente a zero. Invece si osserva dal grafico che la corrente si inverte di segno e continua a circolare per un tempo di circa 8 ns . Questo comportamento è dovuto all'accumulo dei portatori minoritari nella zona della giunzione durante la conduzione. Quando il generatore inverte la tensione, l'eccesso di portatori non scompare istantaneamente: deve trascorrere un certo tempo, detto *storage time* (*tempo di immagazzinamento*), perché avvenga la ricombinazione o perché, per effetto della corrente inversa, i portatori tornino indietro attraversando nuovamente la giunzione. Fino a che l'eccesso dei portatori non è scomparso il diodo rimane in conduzione e il valore della corrente è determinato praticamente solo dalla resistenza R : $I_R = (-v_g - V_\gamma)/R$.

Lo *storage time* t_s può essere calcolato partendo dalla eq. 4.50. Con il diodo in conduzione con una corrente I_F , si ha una carica in eccesso dei portatori minoritari prevalenti:

$$Q_0 = \tau \cdot I_F$$

Dopo che al tempo t_0 la corrente è passata bruscamente dal valore I_F al valore I_R , la carica in eccesso decade con velocità:

$$\frac{dQ}{dt} = -\frac{Q}{\tau} + I_R$$

in cui il termine Q/τ è il contributo della ricombinazione ed il termine I_R è dovuto al riatraversamento della giunzione (I_R è negativa). Integrando l'equazione con la condizione iniziale $Q(t_0) = \tau I_F$ e risolvendo per il tempo $t_0 + t_s$ a cui la carica Q si è ridotta a zero, si ottiene:

$$t_s = \tau \log \left(1 + \left| \frac{I_F}{I_R} \right| \right) \quad (4.56)$$

Nell'esempio di fig. 4.25-4.26 occorre un tempo di circa 8 ns perchè questo avvenga. Una volta terminato lo *storage time*, il diodo torna finalmente nello stato di interdizione con un rilassamento determinato dal caricamento della capacità di transizione (*transition time*). I grafici di fig. 4.26 sono stati calcolati integrando numericamente le equazioni differenziali non lineari del circuito con il programma Spice.

4.21 Dispositivi Optoelettronici

4.21.1 Fotodiodi

Un fotone che colpisce una giunzione $P-N$ di un diodo può essere assorbito e cedere la sua energia nella creazione di una coppia lacuna-elettrone. Se i due portatori di carica così prodotti vengono a trovarsi nella regione di svuotamento della giunzione, il campo elettrico presente li separa rapidamente, impedendone la ricombinazione: la lacuna viene spinta verso la regione P , l'elettrone verso la regione N (fig. 4.28).

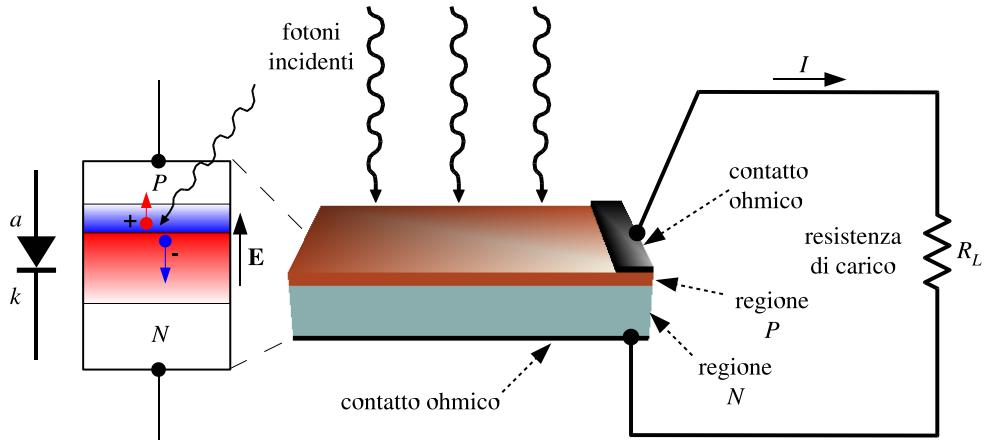


Figura 4.28: Principio di funzionamento del fotodiodo e della cella solare.

Questi eventi, ripetuti nel tempo, portano ad un accumulo di carica positiva nella regione P e negativa nella regione N , risultando in definitiva in una polarizzazione diretta della giunzione: si forma una differenza di potenziale tra i due terminali del diodo, in grado di far circolare una corrente in un eventuale circuito resistivo esterno. Il fotodiodo è quindi un dispositivo in grado di convertire l'energia trasportata da un fascio di luce in energia elettrica.

Fotodiodi costruiti appositamente per convertire impulsi di luce in segnali elettrici, detti anche *fotorivelatori*, sono largamente utilizzati nelle telecomunicazioni e nella strumentazione elettro-ottica. Dispositivi analoghi, ma con giunzioni di superficie molto grande, realizzati con lo scopo di raccogliere la radiazione luminosa solare e convertirla in energia elettrica, sono detti “celle solari” o “celle fotovoltaiche”.

4.21.2 Relazione tensione-corrente in un fotodiodo

Se non è presente un circuito di carico esterno che collega anodo e catodo di un fotodiodo, o se il valore della resistenza di carico è troppo elevato, l'accumulo delle cariche finisce con il portare la giunzione in conduzione e scaricare attraverso la giunzione stessa la corrente prodotta. Indicando con I_ϕ la corrente prodotta nella giunzione per effetto dei fotoni ad un dato livello di illuminazione e con V_D la tensione che si viene a stabilire ai capi del fotodiodo, la corrente I_L che circola nella resistenza di carico esterna è data dalla differenza tra I_ϕ e la corrente attraverso la giunzione per effetto della polarizzazione, secondo la eq. 4.21:

$$I_L = I_\phi - I_S \cdot \left(e^{V_D/\eta V_T} - 1 \right) \quad (4.57)$$

Un esempio di relazione tensione-corrente per una cella solare è riportata in fig. 4.29. Nel caso che il fotodiodo o la cella solare sia chiuso in cortocircuito si ha:

$$V_D = 0 \quad I_L = I_\phi$$

A circuito aperto invece, trascurando il termine -1 ,

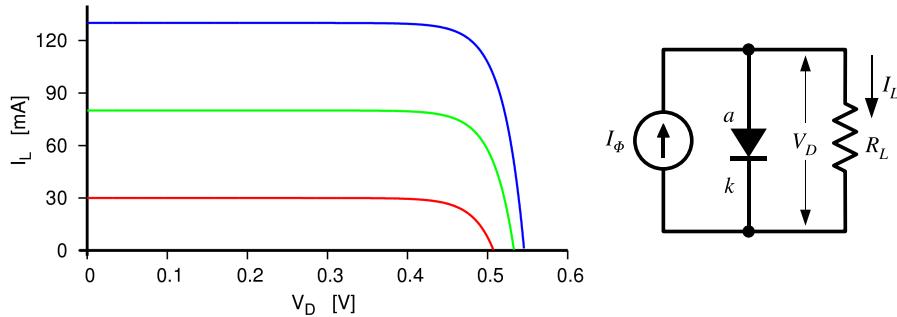


Figura 4.29: Relazione tensione-corrente per una cella solare per tre diversi livelli di illuminazione.

$$I_L = 0 \quad V_D = V_T \log \frac{I_\phi}{I_S}$$

Poichè la corrente I_ϕ è proporzionale al flusso luminoso incidente, anche la corrente di cortocircuito è proporzionale al flusso luminoso; la tensione a circuito aperto invece è proporzionale al logaritmo del flusso luminoso.

4.21.3 Rendimento dei generatori fotovoltaici

Nei due casi limite di circuito aperto e corto circuito si ha rispettivamente $R_L = \infty$, $I_L = 0$ e $R_L = 0$, $V_D = 0$. In entrambi i casi il trasferimento di energia tra cella solare e resistenza di carico è nullo. Esiste un valore intermedio di R_L (e quindi di V_D ed I_L) per cui la potenza $P = I_L \cdot V_D$ a cui avviene il trasferimento di energia è massima (fig. 4.30).

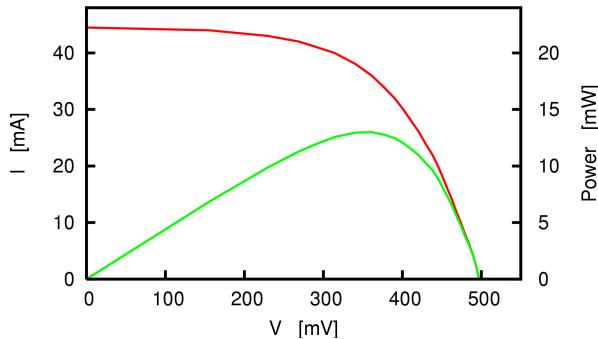


Figura 4.30: Rendimento di una cella solare di 1 cm^2 in condizioni di illuminazione artificiale costante. Linea rossa: relazione tensione-corrente; linea verde: potenza erogata al carico.

La derivata di P rispetto a V_D si annulla per

$$\frac{I_L}{V_D} = -\frac{dI_L}{dV_D} \quad (4.58)$$

Dalla eq. 4.57 e dalla 4.34 si ottiene

$$\frac{dI_L}{dV_D} = -g_d = -1/r_d \quad (4.59)$$

per cui la condizione di massima potenza di trasferimento è:

$$R_L = r_d \quad (4.60)$$

Questa condizione non è statica: r_d dipende dal punto di lavoro; per avere il massimo rendimento da una cella solare la resistenza di carico deve essere continuamente adattata alle condizioni di illuminazione.

4.21.4 Fotorivelatori

Quando un fotodiodo è utilizzato come rivelatore di segnali, come nelle trasmissioni su fibra ottica o nei lettori di *Compact Disc*, si ha interesse ad ottenere non solo la massima ampiezza di segnale possibile ma anche la massima velocità di risposta.

In fig. 4.31 è riportato il circuito elettrico equivalente allo schema di principio di fig. 4.28. L'impulso di corrente i_Φ prodotto dai fotoni nella giunzione del fotodiodo giunge al carico R_L come un segnale di tensione v_L deformato dalla presenza della capacità della giunzione C_j in parallelo alle resistenze r_d (resistenza dinamica della giunzione) ed R_L , come illustrato dalle forme d'onda nella figura a destra.

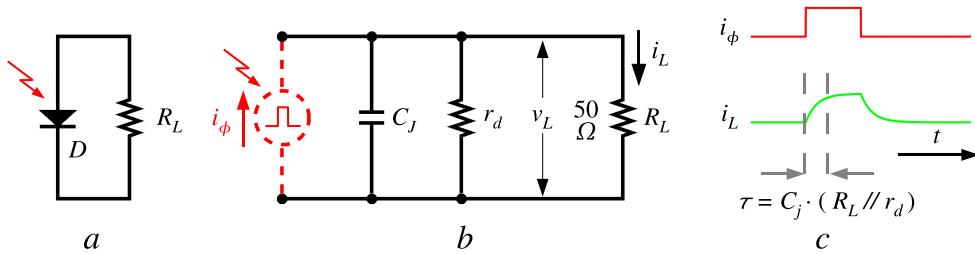


Figura 4.31: a) fotodiodo rivelatore collegato ad una resistenza di carico; b) circuito elettrico equivalente; c) risposta temporale del fotodiodo (in verde) per un impulso rettangolare di luce (in rosso).

Secondo quanto si è visto ai par. 4.18 e 4.19, tenendo il diodo fotorivelatore in condizione di contropolarizzazione si ottengono i seguenti vantaggi:

- la capacità C_j è costituita dalla capacità di transizione C_T , tanto più piccola quanto maggiore è la contropolarizzazione;
- lo spessore della regione di transizione in cui i fotoni generano la photocorrente aumenta con la contropolarizzazione, facendo crescere il segnale prodotto.

Come contropartita, la corrente di polarizzazione inversa attraverso la giunzione può essere causa di un errore rilevante in condizione di segnali molto piccoli.

In figura 4.32 sono riportati due esempi di circuiti per la misura della risposta di un fotodiodo contropolarizzato. In entrambi gli esempi il generatore V_P fornisce al fotodiodo D la tensione di contropolarizzazione. Il circuito in a presenta due problemi: 1) se il diodo dovesse andare in cortocircuito, la resistenza di terminazione R_L , magari *interna all'oscilloscopio*, brucerebbe; 2) non si può avere alcun terminale del diodo collegato a massa.

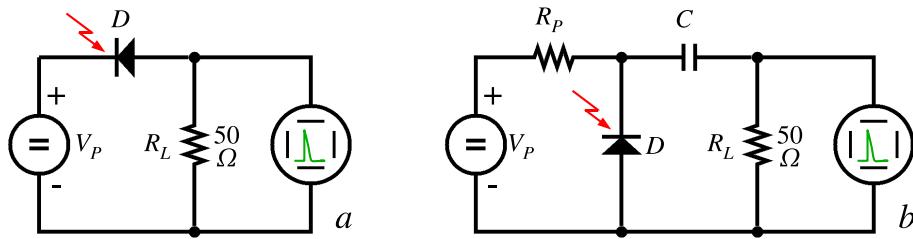


Figura 4.32: Esempi di circuiti di contropolarizzazione di un fotodiodo.

Nel circuito in *b*) la resistenza R_L è protetta dal condensatore C e dalla resistenza R_p ed è possibile avere un terminale del diodo collegato a massa. Si perde però la possibilità di osservare segnali in continua o in bassissima frequenza a causa della cella passa-alto formata da R_L e C .

4.21.5 Diodi led (*light emitting diode* - diodi emettitori di luce)

In condizioni di polarizzazione diretta i portatori di carica iniettati da una regione all'altra attraverso la giunzione si trovano ad essere minoritari in eccesso e quindi soggetti a ricombinarsi rapidamente con i portatori maggioritari. Nel processo di ricombinazione viene ceduta una certa quantità di energia. Questo può avvenire in maniera “non radiativa”, con l'energia ceduta sotto forma di calore al reticolo cristallino, oppure in maniera “radiativa”, con l'energia emessa sotto forma di un fotone. Nei diodi al Silicio la ricombinazione è sostanzialmente di tipo non radiativo. In giunzioni realizzate con semiconduttori composti, come il Fosfuro di Gallio (*GaP*) o l'Arseniuo di Gallio (*GaAs*), è possibile avere una frazione di ricombinazione radiativa sufficiente a realizzare emettitori di luce per uso come indicatori o schermi luminosi e, recentemente, anche per illuminazione.

4.21.6 Diodi laser

In un diodo *led* le emissioni dei fotoni per effetto della ricombinazione delle coppie lacuna-elettrone sono eventi di tipo casuale, non correlati tra di loro. La radiazione luminosa che viene generata è di tipo *incoerente* e lo spettro di emissione è piuttosto ampio. È possibile costruire diodi *led* in cui la regione della giunzione è realizzata come una cavità ottica risonante e quindi farli lavorare nelle condizioni necessarie per generare l'effetto *laser*. Le emissioni dei fotoni divengono allora correlate e la radiazione è di tipo *coerente*, cioè monocromatica e spazialmente ben collimata.

Capitolo 5

Misure su diodi

5.1 Tempi di commutazione

Il circuito reale per la misura dei tempi di commutazione di un diodo non è così semplice come appare nello schema di principio di fig. 4.25. È necessario utilizzare dei cavetti di collegamento per portare il segnale dal generatore al diodo e dal diodo all'oscilloscopio e questi cavetti non possono essere ignorati, come si fa normalmente nei circuiti in bassa frequenza quando i collegamenti sono di lunghezza trascurabile rispetto alle lunghezze d'onda corrispondenti alle frequenze dei segnali in gioco.

La scala dei tempi di commutazione di un diodo di uso generale è dell'ordine dei nanosecondi, mentre i tempi di propagazione dei segnali lungo i cavetti sono di 5 ns/m e quindi non trascurabili (nel dominio delle frequenze: lo spettro dei segnali si estende fino alle centinaia di MHz , con lunghezze d'onda associate dell'ordine del metro, cioè confrontabili con la lunghezza dei cavetti).

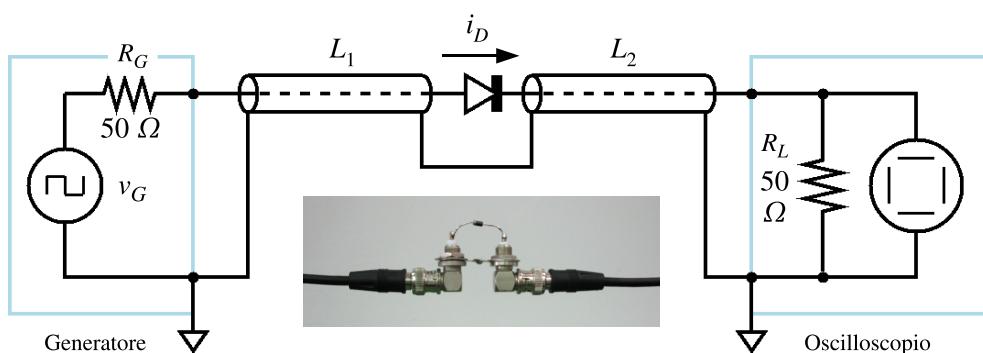


Figura 5.1: I collegamenti tra generatore e diodo in misura e tra diodo e oscilloscopio sono due linee di trasmissione. Nella foto al centro è raffigurato un metodo pratico per inserire il diodo in misura tra le due linee L_1 e L_2 .

Una rappresentazione del sistema di misura più aderente alla realtà è quindi quella di fig. 5.1, in cui i collegamenti tra generatore e diodo e tra diodo e oscilloscopio sono due tratti di linea di trasmissione, L_1 e L_2 , di cui non si può fare a meno di tenere conto.

È possibile però ricondursi al caso semplice di fig. 4.25 sfruttando le proprietà di una linea di trasmissione, che chiusa ad una estremità sulla sua impedenza caratteristica Z_0 presenta all'altra estremità la stessa impedenza Z_0 e permette di trasferire i segnali senza distorsione.

Si dovranno quindi rispettare le due condizioni seguenti:

- La linea L_1 dovrà avere impedenza caratteristica 50Ω in modo da essere adattata all'impedenza interna R_G del generatore. In questo modo il diodo vedrà ancora il segnale provenire da un generatore con impedenza interna 50Ω e la linea non avrà altro effetto se non di introdurre un ritardo τ_1 . Il fatto che la linea non sia adattata all'estremità verso il diodo non è un problema: il segnale, riflesso parzialmente dal disadattamento tra linea e diodo, tornerà verso il generatore dove sarà interamente assorbito dalla resistenza R_G .
- La linea L_2 dovrà avere impedenza caratteristica pari al valore della resistenza R_I posta alla sua estremità, all'ingresso dell'oscilloscopio. Normalmente l'impedenza dell'oscilloscopio è sufficientemente alta da avere un effetto trascurabile ($1 M\Omega$ in parallelo con $15 pF$, a meno che l'oscilloscopio non contenga già una resistenza da 50Ω per la terminazione della linea). Di nuovo, il diodo vedrà come carico la sola resistenza R_I ed il segnale arriverà all'oscilloscopio con un ritardo τ_2 pari al tempo di propagazione lungo la linea L_2 .

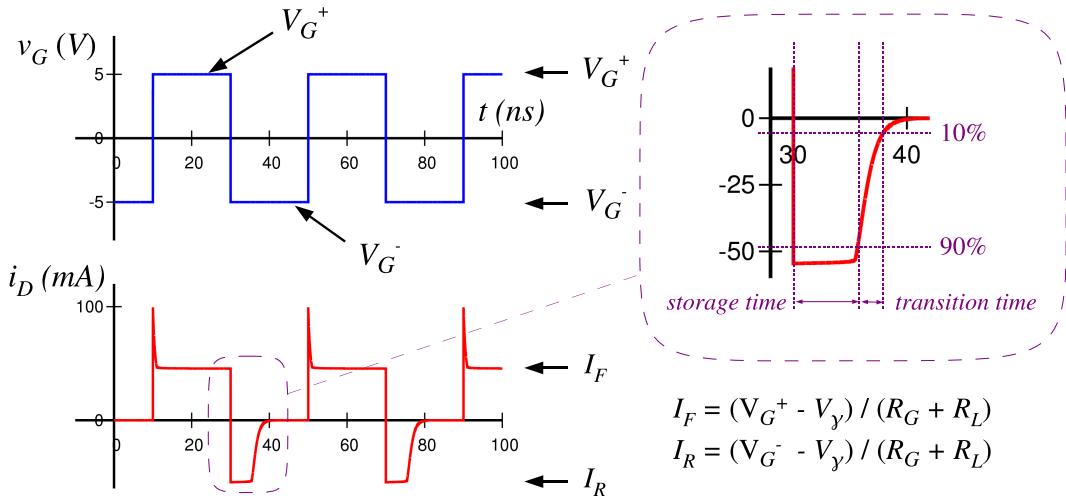


Figura 5.2: Corrente i_D in un diodo portato alternativamente dallo stato di acceso a spento e viceversa da una tensione ad onda quadra (par. 4.20). Il grafico di dettaglio in alto a destra fa vedere la corrente inversa I_R durante il *tempo di immagazzinamento* (*storage time*) ed il decadimento di I_R a zero durante il *transition time*, corrispondente all'accumulo di carica nella capacità di transizione.

I grafici in fig. 5.2 a sinistra riportano l'andamento temporale della tensione v_G (in alto) e della corrente i_D (in basso). I valori della tensione durante la semionda positiva (V_G^+) e negativa (V_G^-) sono regolabili indipendentemente impostando op-

portunamente l'ampiezza picco-picco del segnale a $V_P = V_G^+ - V_G^-$ e la componente continua (*offset*) a $V_O = (V_G^+ + V_G^-)/2$.

Durante la semionda positiva il diodo si trova in conduzione e la corrente ha un valore stazionario $I_F = (V_G^+ - V_\gamma)/(R_G + R_L)$, preceduto da un picco di ampiezza circa doppia, dovuto alla capacità di transizione C_T . Le caratteristiche di questo picco sono risolvibili solo con strumentazione di velocità adeguata.

Durante la semionda negativa il diodo non passa immediatamente in interdizione. La corrente passa al valore $I_R = (V_G^- - V_\gamma)/(R_G + R_I)$ e vi rimane per il tempo necessario per eliminare l'eccesso di portatori di carica presente in prossimità della giunzione (*storage time*, par. 4.20). Terminato lo storage time, la corrente può finalmente tornare a zero, con un rilassamento dovuto al caricamento della capacità di transizione (*transition time*).

Modificando opportunamente V_G^+ e V_G^- (tramite V_P e V_O) si possono fare assumere valori diversi a I_F e I_R , misurare il valore di corrispondente per lo *storage time* e confrontare i dati sperimentali con l'eq. 4.56.

Per ragioni che verranno approfondite meglio nel par. 8.3, è conveniente misurare i tempi tra i livelli 10% e 90% delle transizioni, come indicato nel dettaglio in alto a destra della fig. 5.2.

Non è possibile osservare durante la misura il segnale fornito dal generatore V_G . Il

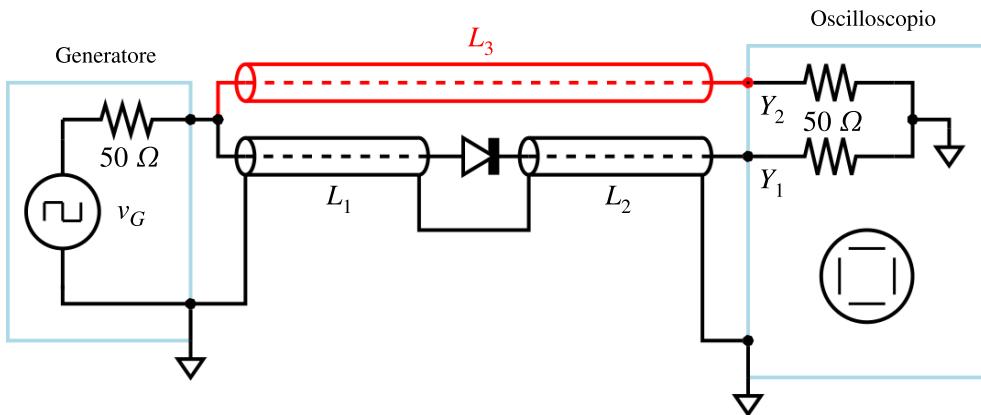


Figura 5.3: Il collegamento L_3 , disegnato in rosso, indica come *NON* si deve cercare di osservare il segnale del generatore V_G .

tentativo di prelevare questo segnale tramite una diramazione a T e portarlo, come in figura fig. 5.3, al secondo canale dell'oscilloscopio per osservarlo insieme alla corrente I_D è *sbagliato* per due motivi:

- La linea L_1 non si trova più ad essere terminata all'inizio sulla sua resistenza caratteristica di 50Ω , ma su una resistenza di 25Ω , data dal parallelo di R_G con l'impedenza equivalente di L_3 . Il segnale riflesso dalla terminazione non adattata tra L_1 ed il diodo, una volta tornato indietro verso il generatore, sarà solo parzialmente assorbito, mentre in parte sarà nuovamente riflesso verso il diodo. Questo segnale si andrà a sovrapporre a quello originale del generatore, deformandolo.

- Il segnale che si osserverà sul canale Y_2 dell'oscilloscopio non sarà quello originale v_G del generatore, ma conterrà sovrapposto anche il segnale riflesso dal disadattamento tra L_1 ed il diodo.

5.2 Capacità di transizione

La capacità di transizione di un diodo ha valori compresi tra la frazione di picoferad per un diodo veloce per piccoli segnali e le decine o centinaia di picoferad per diodi raddrizzatori di potenza o fotodiodi con aree di giunzione elevate e può essere misurata mediante uno strumento *Analizzatore di impedenza*, il cui schema di principio è riportato in fig. 5.4.

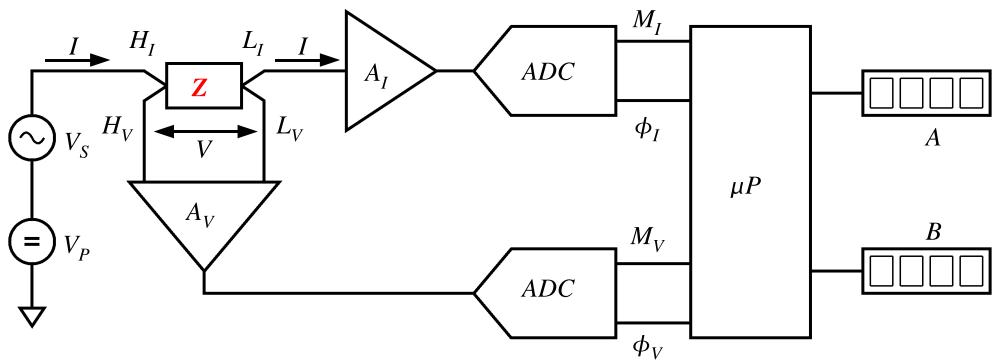


Figura 5.4: Schema di principio di un Analizzatore di Impedenza.

I due generatori V_P e V_S applicano al componente da misurare Z rispettivamente una tensione continua di polarizzazione ed un segnale di tensione alternata a frequenza f .

La corrente I attraverso l'impedenza Z e la tensione V ai suoi capi sono inviate tramite i due amplificatori A_I e A_V , di caratteristiche opportune, a due convertitori ADC (Analog to Digital Converter) che ne determinano ampiezza (M_I ed M_V) e fase Φ_I e Φ_V . Da questi dati l'elaboratore μP , che di solito fa parte dello strumento, calcola il valore dell'impedenza complessa e la riporta secondo la rappresentazione richiesta (modulo e fase, parte reale ed immaginaria, resistenza e capacità ecc.) sui due indicatori A e B .

La misura viene fatta con una connessione a quattro terminali per minimizzare l'effetto dei collegamenti.

In fig. 5.5 è riportato il pannello frontale di un analizzatore di impedenza mod. *HP4275A*. Il riquadro rosso sulla sinistra delimita i comandi che permettono di scegliere la modalità di rappresentazione dei dati. Questi comandi non influiscono sul modo di esecuzione della misura, ma esclusivamente sul modo con cui sono riportati i risultati in forma numerica sui due *display A* e *B*: lo stesso dato può essere riportato come impedenza o ammettenza, come parte reale ed immaginaria oppure modulo e fase, o anche con i valori della rete *RC/RL* serie/parallelo equivalente. L'informazione contenuta è sempre la stessa e si può passare numericamente dai valori di una rappresentazione a quelli di un'altra. Ad esempio, i dati ottenuti per il modello *RC*-serie



Figura 5.5: Pannello frontale dell’Analizzatore di impedenza complessa HP4275A.

possono essere convertiti nel loro equivalente RC -parallelo scrivendo l’equazione

$$Z = R_S + \frac{1}{j\omega C_S} = \frac{1}{j\omega C_P + 1/R_P} \quad (5.1)$$

in cui $\omega = 2\pi f$ ed f è la frequenza di misura. Eguagliando separatamente le parti reale ed immaginaria si ottiene:

$$C_P = \frac{C_S}{1 + \omega^2 R_S^2 C_S^2} \quad R_P = R_S + \frac{1}{\omega^2 R_S C_S^2} \quad (5.2)$$

Nel caso della misura della capacità di transizione, la scelta più semplice è ottenere i valori per una rete RC -parallelo.

La misura va ripetuta per valori di V_P che coprano l’intervallo di tensioni di polarizzazione che si intende studiare (la tensione V_P viene impostata mediante una scatola di comando esterna all’apparecchio). Si dovrà fare attenzione a rispettare la polarità e a non superare la tensione di *breakdown* del diodo in misura.

La misura va anche ripetuta, almeno per alcuni valori di polarizzazione V_P , per diversi valori di frequenza f ; la costanza dei valori di R e C in funzione della frequenza conferma la validità del modello adottato, mentre eventuali variazioni indicano la presenza di altri effetti di cui il modello non tiene conto. La frequenza si seleziona tramite i due pulsanti nel riquadro verde, nell’intervallo tra 10 kHz e 10 MHz , a passi $1 - 2 - 4 - 10$.

La scelta della tensione V_S , impostabile e verificabile tramite i comandi racchiusi nei riquadri arancione, è critica. Si dovrebbe rispettare sempre la condizione $V_S \ll V_P$, implicita nella definizione stessa di capacità di transizione (eq. 4.43). Nella pratica questo non è sempre possibile, specialmente per valori di V_P vicini a zero, perché i valori di V_S diventano talmente piccoli da far aumentare in modo inaccettabile l’errore di misura. Si tratta di trovare il compromesso migliore, osservando che in corrispondenza della riduzione di V_S l’apparecchio riduce il numero di cifre riportate sui due *display* alle sole che possono essere misurate in modo significativo,

mentre in corrispondenza di valori di V_S più elevati il risultato viene a dipendere dal valore di V_S a causa della non linearità del diodo.

La misura comunque è possibile anche per $V_P = 0$, purchè si abbia l'accortezza di utilizzare un valore per V_S sufficientemente piccolo da non portare il diodo in conduzione in modo significativo.

L'errore di misura dipende in modo piuttosto complicato da tutti i vari parametri. Come regola empirica, si può assumere un errore di ± 5 unità sulla cifra meno significativa riportata sul *display*.

I due pulsanti nel riquadro blu consentono di effettuare l'azzeramento dello strumento per impedenza zero (corto circuito) e conduttanza zero (circuito aperto).

5.2.1 Fit dei dati sperimentali con il programma *Gnuplot*

Gnuplot è un programma interattivo a linea di comando destinato prevalentemente alla grafica di tabelle di dati e funzioni matematiche, ma è dotato anche di risorse, tra cui il *fit* ai minimi quadrati, che consentono di effettuare elaborazioni di dati sperimentali. Il codice di *Gnuplot* fa parte del software distribuito liberamente e gratuitamente e può essere utilizzato praticamente su tutti o quasi tutti i sistemi operativi attuali.

Il grosso vantaggio di un programma a linea di comando utilizzato per l'elaborazione di dati per uso scientifico è dato dalla possibilità di scrivere l'intera procedura di elaborazione come sequenza di comandi (*script*). In questo modo si ha una traccia chiara e comprensibile, e soprattutto ripetibile, della procedura che è stata utilizzata. Una modifica all'elaborazione o la correzione di un errore richiede solo di modificare opportunamente la sequenza di comandi e poi ripetere l'esecuzione della procedura.

Di seguito viene riportato un esempio di procedura per riportare in grafico le misure di capacità di transizione in funzione del potenziale di polarizzazione come descritto nel paragrafo precedente, insieme ai risultati del fit ai minimi quadrati con l'eq. 4.43.

La procedura è composta di due *files*: in uno, denominato nell'esempio *ct.plt*, ci sarà la sequenza di comandi da sottoporre al programma *Gnuplot* per effettuare le operazioni desiderate; nell'altro, denominato *ct.dat*, saranno inseriti i dati sperimentali, cioè una coppia di valori tensione-capacità per ogni riga.

- File *ct.plt*

```
ct(v,cc,cz,vj) = cc + cz/sqrt(1+v/vj)
fit ct(x,cc,cz,vj) "ct.dat" via cc,cz,vj
plot "ct.dat", ct(x,cc,cz,vj)
```

- File *ct.dat*

#	-V	pF
1.00	14.73	
2.00	11.45	
3.00	9.895	
4.00	8.754	
5.00	7.957	

6.00	7.362
7.00	6.892
8.00	6.515
9.00	6.201
10.00	5.936
13.00	5.329
16.00	4.907
19.00	4.589
22.00	4.341
26.00	4.079
30.00	3.870
35.00	3.658

Il *file* dei comandi è composto di sole tre righe, contenenti un comando ciascuna. La prima riga contiene la definizione della funzione `ct()`, da utilizzare nei calcoli successivi; la funzione è l'espressione della capacità di transizione data dalla eq. 4.43, con $cz = C_0$, $v_j = V_j$ e l'aggiunta del parametro `cc` per tenere conto della capacità, piccola ma non trascurabile, dei terminali ed eventuali altre capacità parassite.

La seconda riga è il comando di *fit*. Contiene quattro elementi: la parola chiave `fit`, che identifica il comando; la funzione che deve essere utilizzata per effettuare il fit, `ct()`, definita nella riga precedente; il nome del *file* contenente i dati e infine la lista dei parametri della funzione, che devono essere aggiustati per il miglior accordo, preceduti dalla parola chiave `via`. La variabile indipendente della funzione è sempre identificata dalla lettera *x*.

Al termine del calcolo rimangono definite tre variabili, `cc`, `cz`, `vj`, con assegnati i valori ottenuti.

La terza riga è il comando *plot*, cioè il disegno del grafico, con due argomenti, il *file* contenente i dati sperimentali e la funzione utilizzata nel *fit*. Il risultato che si ottiene è riportato in fig. 5.6.

Le crocette rosse rappresentano i dati sperimentali, la linea verde continua è il grafico della funzione `ct()`.

Durante il *fit*, che viene eseguito con un procedimento iterativo, il programma *Gnuplot* riporta ad ogni passo informazioni relative alla convergenza; alla fine vengono riportati i valori trovati per i parametri, insieme agli errori statistici.

L'esempio seguente riporta la parte iniziale e finale dell'esecuzione della procedura `ct.plt`.

```
carla@lab22$ gnuplot
G N U P L O T
Version 4.2 patchlevel 2
last modified 31 Aug 2007
System: Linux 2.6.26-1-686

Copyright (C) 1986 - 1993, 1998, 2004, 2007
Thomas Williams, Colin Kelley and many others

Type 'help' to access the on-line reference manual.
The gnuplot FAQ is available from http://www.gnuplot.info/faq/
```

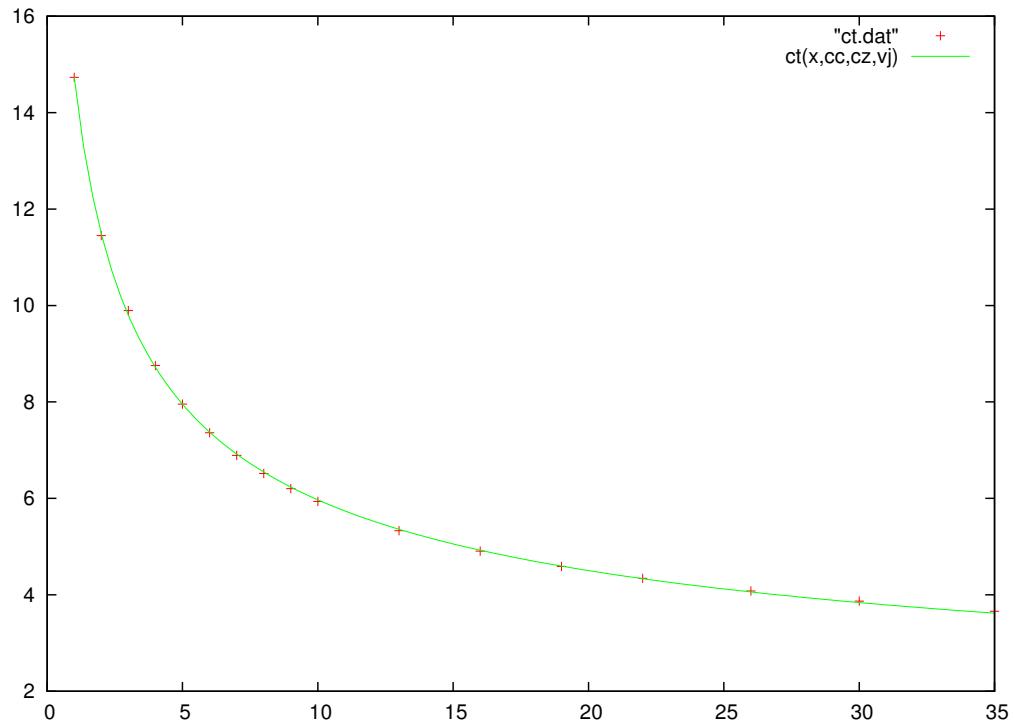


Figura 5.6: Capacità di transizione: grafico dei dati sperimentali e della funzione `ct()` con i valori dei parametri ottenuti dal *fit*.

Send bug reports and suggestions to <<http://sourceforge.net/projects/gnuplot>>

```
Terminal type set to 'wxt'
gnuplot> load "ct.plt"
```

```
Iteration 0
WSSR      : 645.33          delta(WSSR)/WSSR   : 0
delta(WSSR) : 0              limit for stopping : 1e-05
lambda    : 0.621325
```

```
initial set of free parameter values
```

```
cc      : 1
cz      : 1
vj      : 1
/
.....
.....
.....
```

```
After 47 iterations the fit converged.
final sum of squares of residuals : 0.0208152
rel. change during last iteration : -2.88667e-10
```

```

degrees of freedom      (FIT_NDF)          : 14
rms of residuals        (FIT_STDFIT) = sqrt(WSSR/ndf)    : 0.038559
variance of residuals   (reduced chisquare) = WSSR/ndf   : 0.0014868

Final set of parameters      Asymptotic Standard Error
=====                      =====
cc              = 0.835566      +/- 0.0379      (4.536%)
cz              = 25.5476       +/- 0.4354      (1.704%)
vj              = 0.420705      +/- 0.02054     (4.882%)
correlation matrix of the fit parameters:
cc      cz      vj
cc      1.000
cz      0.712  1.000
vj      -0.823 -0.982  1.000
gnuplot>

```

La prima linea contiene il comando di esecuzione di *Gnuplot*, che è semplicemente `gnuplot`. Dopo le presentazioni, il programma si predispone ad accettare i comandi scrivendo `gnuplot>`. L'unico comando da dare per eseguire la procedura `ct.plt` è `load "ct.plt"`. L'esecuzione parte e procede ad effettuare il *fit*. Al termine, in questo caso dopo 47 iterazioni, vengono scritti i valori trovati per i tre parametri, insieme al valore di χ^2 ed alla matrice di covarianza, quindi viene disegnato il grafico della funzione e dei dati sperimentali.

Il valore *rms* degli scarti (`rms of residuals`) risulta essere circa 0.04 *pF*, quasi un ordine di grandezza superiore all'errore stimato a priori sui dati sperimentali. Omettendo i primi due punti sperimentali, che risultano meno accurati degli altri a causa della difficoltà a rispettare la condizione $V_S \ll V_P$, e ripetendo il calcolo si ottiene invece uno scarto *rms* di 0.007 *pF*, in ottimo accordo con l'errore di misura stimato.

All'inizio della procedura vengono dichiarati i valori iniziali dei parametri che devono essere ottimizzati durante il fit. Non avendo inserito alcuna indicazione, viene utilizzato automaticamente il valore 1. A volte questo porta la procedura iterativa a non convergere, per cui può essere necessario specificare obbligatoriamente dei valori iniziali ragionevoli.

Complessivamente il programma *Gnuplot* è dotato di quasi duecento tra comandi e sottocomandi, che sono descritti a fondo nel manuale di istruzioni e con cui è possibile effettuare altre operazioni e definire in molti dettagli l'aspetto dei grafici.

Capitolo 6

Reti a due porte

6.1 Circuiti attivi

Gli elementi circuituali considerati sino ad ora, sia lineari (resistenze, capacità, induttanze e generatori indipendenti), sia non lineari (diodi), sono detti *passivi* in quanto, consentono di effettuare molte operazioni sui segnali elettrici, ma non una che è veramente fondamentale, l'*amplificazione*: ottenere da un segnale debole una copia fedele di potenza maggiore. I componenti attivi maggiormente in uso oggi da cui si può ottenere amplificazione sono i transistor. I tre capitoli seguenti sono dedicati alla descrizione del loro funzionamento.

L'amplificazione, combinata con la *reazione positiva* (*positive feedback* o *regenerative feedback*) consente di ottenere un'altra funzione fondamentale: la generazione dei segnali.

La combinazione di amplificazione e *reazione negativa* o *controreazione* (*negative feedback*)¹ consente invece di ottenere nelle prestazioni dei circuiti elettronici livelli di precisione impensabili sulla base della scarsa precisione intrinseca dei dispositivi a semiconduttore. Le applicazioni dell'amplificatore operazionale, descritte nei cap. 11 e 12 ne sono un esempio.

6.2 Amplificatore ideale

Il comportamento di un amplificatore ideale, ad esempio di tensione, è descritto dalla relazione lineare

$$V_{out} = A \cdot V_{in} \quad (6.1)$$

dove V_{in} è la tensione applicata all'ingresso dell'amplificatore, V_{out} la tensione ottenuta all'uscita e A è una costante reale, detta *coefficiente di amplificazione* o *guadagno*. Un tale dispositivo è schematizzato in fig. 6.1 come un *generatore controllato* ideale che fornisce una tensione V_{out} proporzionale alla tensione di controllo V_{in} .

Non esiste nessun singolo dispositivo fisico che consenta di ottenere direttamente le prestazioni descritte dalla eq. 6.1; nella pratica le relazioni tra le variabili di ingresso e di uscita di un dispositivo amplificatore sono più complicate e non lineari,

¹Insieme ad un uso accorto delle simmetrie nei circuiti, come nell'amplificatore differenziale (par. 7.9).

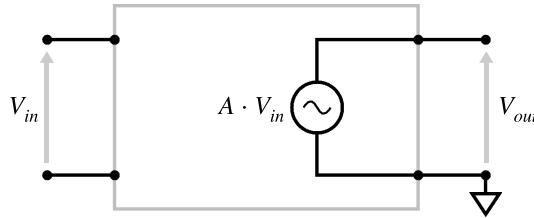


Figura 6.1: Amplificatore ideale di tensione: un generatore controllato fornisce una tensione V_{out} proporzionale alla tensione di ingresso V_{in} .

come illustrato ad esempio dalle famiglie di curve delle figure 7.7b e 9.9. È possibile però combinando opportunamente più dispositivi realizzare circuiti, ad esempio gli *amplificatori operazionali* descritti nel cap. 11, il cui comportamento si avvicina considerevolmente a quello ideale.

Tra i dispositivi fisici che consentono di ottenere amplificazione i più diffusi oggi sono di gran lunga i transistor: *bjt* (*bipolar junction transistor*), *jfet* (*junction field effect transistor*), *mosfet* (*metal oxide semiconductor field effect transistor*).

Una caratteristica comune a questi dispositivi attivi è di essere dotati di tre terminali, che formano due *porte*: porta 1 o *di ingresso* (*in*) e porta 2 o *di uscita* (*out*), con uno dei tre terminali in comune tra le due porte (fig. 6.2) ².

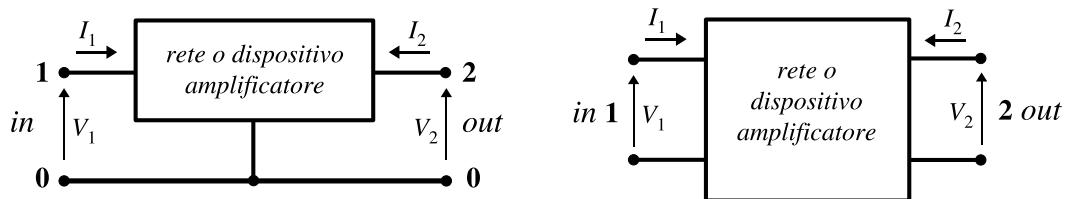


Figura 6.2: In una rete a due *porte* si ha convenzionalmente una porta di ingresso (*in*) ed una di uscita (*out*). In un amplificatore con transistor *bjt* o *fet* le due porte hanno un terminale in comune. Le tensioni V_1 e V_2 sono le tensioni dei nodi 1 e 2 misurate rispetto al nodo comune 0 e le correnti I_1 ed I_2 si considerano positive quando sono entranti nel dispositivo. Le due porte possono anche non avere un terminale comune, come nel caso dell'amplificatore differenziale (entro l'intervallo delle tensioni di ingresso di modo comune) o del trasformatore.

La possibilità di amplificare un segnale deriva dal fatto che, una volta che il dispositivo sia opportunamente *polarizzato*, le variabili elettriche della porta di ingresso, tensione o corrente, sono in grado di controllare e variare tensioni e correnti più intense nella porta di uscita.

Le relazioni matematiche tra le variabili di ingresso (I_1, V_1) e di uscita (I_2, V_2) sono determinate dai meccanismi fisici di funzionamento dei dispositivi amplificatori, esaminati nei capitoli seguenti. Alcune considerazioni però sono del tutto generali e prescindono dal particolare dispositivo utilizzato.

²È possibile ottenere amplificazione anche da componenti ad una sola porta, come ad esempio dalle resistenze negative fornite da particolari tipi di diodi (diodi tunnel, diodi Gunn).

6.3 Variabili indipendenti e dipendenti

La rete a due porte di fig. 6.2 è dotata di due gradi di libertà: delle quattro variabili I_1, V_1, I_2, V_2 se ne possono scegliere due come indipendenti e le altre due saranno funzioni delle prime. Nello studio dei circuiti elettronici risulta conveniente considerare come indipendenti una variabile scelta tra quelle di ingresso ed una scelta tra quelle di uscita, ad esempio V_1 e V_2 . Questo equivale concettualmente a collegare un generatore di tensione ideale alla porta di ingresso ed uno alla porta di uscita, impostare i valori delle tensioni ed avere le correnti determinate di conseguenza:

$$I_1 = I_1(V_1, V_2) \quad (6.2)$$

$$I_2 = I_2(V_1, V_2) \quad (6.3)$$

Analogamente è possibile scegliere I_1 e I_2 come indipendenti, collegando concettualmente due generatori di corrente ideali alle due porte, oppure una tensione ed una corrente. Sono possibili in tutto le quattro combinazioni:

$I_1(V_1, V_2)$	$I_1(V_1, I_2)$
$I_2(V_1, V_2)$	$V_2(V_1, I_2)$
(6.4)	
$V_1(I_1, V_2)$	$V_1(I_1, I_2)$
$I_2(I_1, V_2)$	$V_2(I_1, I_2)$

È possibile anche considerare come indipendenti entrambe le variabili di ingresso o di uscita ed avere quindi:

$V_2(V_1, I_1)$	$V_1(V_2, I_2)$
$I_2(V_1, I_1)$	$I_1(V_2, I_2)$

(6.5)

Questa scelta però risulta conveniente solo per lo studio di linee, reti di trasmissione e filtri e non viene normalmente utilizzata per lo studio del funzionamento dei dispositivi amplificatori.

Le varie combinazioni di variabili indipendenti e dipendenti forniscono descrizioni completamente equivalenti dal punto di vista matematico. Alcune combinazioni risulteranno nell'uso più convenienti di altre in relazione alle caratteristiche dei dispositivi o delle reti che si dovranno descrivere.

6.4 Modelli lineari per i dispositivi a due porte

Nel caso dei transistor le relazioni 6.4 sono praticamente sempre fortemente non lineari. Tuttavia di solito è possibile individuare un *punto di lavoro* intorno al quale le relazioni tra tensioni e correnti sono sufficientemente lineari da poter essere descritte da uno sviluppo in serie di Taylor arrestato al primo ordine (ed anche da consentire un funzionamento del dispositivo senza troppa *distorsione*).

Dalle equazioni 6.4, differenziando intorno al punto di lavoro rispetto a ciascuna delle due variabili indipendenti e sottraendo poi i valori costanti corrispondenti al

punto di lavoro, si ottiene:

$$\begin{aligned}
 i_1 &= \frac{\partial I_1}{\partial V_1} \Big|_{V_2} v_1 + \frac{\partial I_1}{\partial V_2} \Big|_{V_1} v_2 & i_1 &= \frac{\partial I_1}{\partial V_1} \Big|_{I_2} v_1 + \frac{\partial I_1}{\partial I_2} \Big|_{V_1} i_2 \\
 i_2 &= \frac{\partial I_2}{\partial V_1} \Big|_{V_2} v_1 + \frac{\partial I_2}{\partial V_2} \Big|_{V_1} v_2 & v_2 &= \frac{\partial V_2}{\partial V_1} \Big|_{I_2} v_1 + \frac{\partial V_2}{\partial I_2} \Big|_{V_1} i_2 \\
 v_1 &= \frac{\partial V_1}{\partial I_1} \Big|_{V_2} i_1 + \frac{\partial V_1}{\partial V_2} \Big|_{I_1} v_2 & v_1 &= \frac{\partial V_1}{\partial I_1} \Big|_{I_2} i_1 + \frac{\partial V_1}{\partial I_2} \Big|_{I_1} i_2 \\
 i_2 &= \frac{\partial I_2}{\partial I_1} \Big|_{V_2} i_1 + \frac{\partial I_2}{\partial V_2} \Big|_{I_1} v_2 & v_2 &= \frac{\partial V_2}{\partial I_1} \Big|_{I_2} i_1 + \frac{\partial V_2}{\partial I_2} \Big|_{I_1} i_2
 \end{aligned} \tag{6.6}$$

(le variabili in minuscole rappresentano i *segnali*, cioè le variazioni di tensioni e correnti rispetto ai valori del punto di lavoro rappresentati dalle maiuscole, secondo la notazione vista nel par. 4.14). Per ogni rete lineare (o linearizzata) a due porte sono perciò possibili quattro modelli equivalenti, ciascuno descritto da quattro costanti (o parametri), corrispondenti alle quattro derivate che compaiono in ciascuna coppia delle eq. 6.6. La scelta delle variabili indipendenti determina le dimensioni dei parametri del modello. Si possono avere quattro conduttanze ($\partial I / \partial V$), dette *parametri g*, quattro resistenze ($\partial V / \partial I$), dette *parametri r*, o quattro parametri misti, una conduttanza, una resistenza e due numeri puri ($\partial V / \partial V$, $\partial I / \partial I$), detti parametri *h* (*hybrid*) ed *m* (*mixed*). Ogni parametro è individuato da due indici di cui il primo indica la variabile dipendente (1:ingresso, 2:uscita) ed il secondo la variabile indipendente rispetto a cui è calcolata la derivata. Con queste notazioni, sostituendo $g_{11} = \partial I_1 / \partial V_1$, $g_{12} = \partial I_1 / \partial V_2$, ecc., le eq. 6.6 diventano

$$\begin{aligned}
 i_1 &= g_{11} v_1 + g_{12} v_2 & i_1 &= m_{11} v_1 + m_{12} i_2 \\
 i_2 &= g_{21} v_1 + g_{22} v_2 & v_2 &= m_{21} v_1 + m_{22} i_2 \\
 \\
 v_1 &= h_{11} i_1 + h_{12} v_2 & v_1 &= r_{11} i_1 + r_{12} i_2 \\
 i_2 &= h_{21} i_1 + h_{22} v_2 & v_2 &= r_{21} i_1 + r_{22} i_2
 \end{aligned} \tag{6.7}$$

(a volte viene utilizzata la lettera *y* per la famiglia di parametri qui indicata con la *g* e la lettera *g* per la famiglia di parametri misti *m*).

In ognuna delle quattro famiglie di parametri (*g*, *m*, *h*, *r*) compare un parametro di *ingresso* ed uno di *uscita*, individuati dagli indici ₁₁ e ₂₂ e due parametri di *trasferimento*. L'indice ₂₁ identifica il parametro di *trasferimento diretto* e descrive l'azione di un segnale, tensione o corrente, secondo i casi, applicato alla porta 1 (l'ingresso) sulle variabili di uscita alla porta 2; il parametro con indice ₁₂ indica invece il *trasferimento inverso*, cioè l'effetto di un segnale applicato al circuito di uscita sulle variabili di ingresso.

In base a questi significati i parametri sono identificati anche con gli indici _i, _f, _r, _o anzichè con gli indici numerici utilizzati fino ad ora. Ad esempio, nel caso dei parametri *g* ed *h*, che sono i più utilizzati nella descrizione dei transistor, si ha:

$$g_{11} \rightarrow g_i \text{ conduttanza di } \textit{ingresso} \text{ (input)}$$

$g_{12} \rightarrow g_r$ conduttanza di trasferimento *inverso* (*reverse*)

$g_{21} \rightarrow g_f$ conduttanza di trasferimento *diretto* (*forward*)

$g_{22} \rightarrow g_o$ conduttanza di *uscita* (*output*)

$h_{11} \rightarrow h_i$ resistenza di *ingresso* (*input*)

$h_{12} \rightarrow h_r$ trasferimento di tensione *inverso* (*reverse*)

$h_{21} \rightarrow h_f$ trasferimento di corrente *diretto* (*forward*)

$h_{22} \rightarrow h_o$ conduttanza di *uscita* (*output*)

Ai quattro gruppi di equazioni corrispondono i quattro diversi circuiti lineari riportati in fig. 6.3. Ad ognuna delle due porte è presente un generatore di tensione

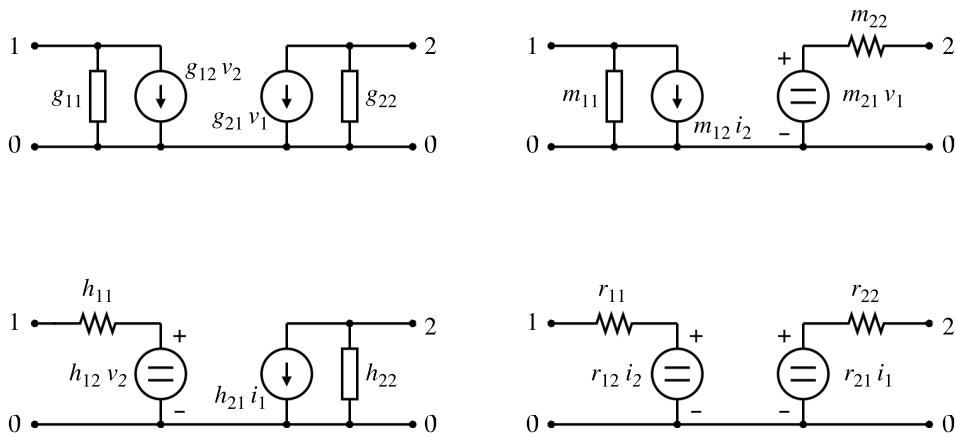


Figura 6.3: Modelli lineari di un dispositivo amplificatore corrispondenti alle diverse scelte di variabili indipendenti e dipendenti.

o corrente, seconda dei casi, controllato dalla variabile elettrica dell'altra porta; ad ogni generatore (e quindi ad ogni porta) è associata una conduttanza in parallelo, se di corrente, ed una resistenza in serie, se di tensione.

Un generatore controllato ideale è un generatore che eroga una tensione (o una corrente) il cui valore è proporzionale al valore di una variabile elettrica in un altro punto del circuito. È il generatore controllato che schematizza gli effetti fisici fondamentali alla base dell'amplificazione: un segnale elettrico (piccolo) all'ingresso di un dispositivo amplificatore controlla variabili elettriche nel circuito di uscita di potenza (auspicabilmente) molto maggiore.

6.4.1 Trasformazioni di parametri

Come già detto, i quattro sistemi di parametri forniscono descrizioni equivalenti del comportamento di una rete lineare ed è solo una questione di convenienza l'utilizzo di un sistema o di un altro.

In genere per una rete reale è sempre possibile passare da una famiglia di parametri ad un'altra (ma vedi il caso degli amplificatori ideali del par. 6.6 come esempio di

impossibilità). Ad esempio, nel sistema che utilizza i parametri g

$$\begin{aligned} i_1 &= g_{11} v_1 + g_{12} v_2 \\ i_2 &= g_{21} v_1 + g_{22} v_2 \end{aligned} \quad (6.8)$$

si può invertire la prima equazione e sostituire nella seconda, ottenendo:

$$v_1 = \frac{1}{g_{11}} i_1 - \frac{g_{12}}{g_{11}} v_2 \quad (6.9)$$

$$i_2 = \frac{g_{21}}{g_{11}} i_1 + \frac{g_{22} g_{11} - g_{12} g_{21}}{g_{11}} v_2 \quad (6.10)$$

cioè la descrizione della stessa rete espressa tramite i parametri h , con

$$\begin{aligned} h_{11} &= \frac{1}{g_{11}} & h_{12} &= -\frac{g_{12}}{g_{11}} \\ h_{21} &= \frac{g_{21}}{g_{11}} & h_{22} &= \frac{g_{22} g_{11} - g_{12} g_{21}}{g_{11}} \end{aligned} \quad (6.11)$$

6.5 Circuito amplificatore

Il circuito di fig. 6.4 è un esempio di *stadio* amplificatore in cui il dispositivo attivo è descritto mediante i parametri h . Il segnale proveniente dal generatore di tensione

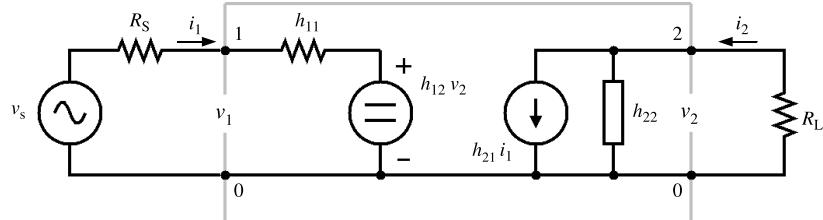


Figura 6.4: Nel riquadro grigio è raffigurato il circuito equivalente di un amplificatore descritto mediante i parametri h . A sinistra, la sorgente del segnale da amplificare (v_s , R_S), a destra i circuiti o apparecchi utilizzatori del segnale amplificato (R_L).

v_s , con resistenza interna R_S , è applicato alla porta di ingresso (1) del dispositivo amplificatore, ad esempio un transistor *bjt* o *fet* opportunamente polarizzato, come si vedrà nei cap. 7 e 9. Il segnale che si ottiene alla porta di uscita (2) viene inviato ad una resistenza R_L che rappresenta il circuito o apparecchio utilizzatore.

Per la maglia di ingresso, utilizzando la seconda legge di Kirchhoff, si ha

$$i_1 (R_S + h_{11}) = v_s - h_{12} v_2 \quad (6.12)$$

Per il nodo di uscita (2) con la prima legge di Kirchhoff si può scrivere

$$h_{22} v_2 + \frac{v_2}{R_L} + h_{21} i_1 = 0 \quad (6.13)$$

Da queste due equazioni si ottiene il sistema

$$\begin{aligned} (R_S + h_{11}) i_1 + h_{12} v_2 &= v_s \\ h_{21} i_1 + \left(h_{22} + \frac{1}{R_L} \right) v_2 &= 0 \end{aligned} \quad (6.14)$$

che ha i_1 e v_2 come incognite e v_s termine noto.

Risolvendo per v_2 e dividendo per v_s si ottiene l'*amplificazione* o *guadagno di tensione* dello stadio:

$$A_v = \frac{v_2}{v_s} = -\frac{-h_{21}}{(h_{11} + R_S) \left(h_{22} + \frac{1}{R_L} \right) - h_{21} h_{12}} \quad (6.15)$$

L'eq. 6.15 è la soluzione esatta del sistema 6.14; conviene però esaminarne una versione semplificata, in modo da mettere in evidenza il ruolo dei diversi termini.

Nel circuito di fig. 6.4 l'amplificazione è rappresentata dal generatore h_{21} , che fornisce nel circuito di uscita una corrente controllata dalla corrente nel circuito di ingresso. Il secondo generatore controllato, h_{12} , rappresenta spesso effetti secondari, corrispondenti ad accoppiamenti tra il circuito di uscita e quello di ingresso, che molte volte in prima approssimazione possono essere trascurati (a meno che non siano stati aggiunti al circuito per introdurre volutamente una *reazione*, come si vedrà nel cap. 11).

Considerando quindi nullo h_{12} , l'eq. 6.15 si semplifica in:

$$A_0 = \frac{v_2}{v_s} = \frac{-h_{21}}{(h_{11} + R_S) \left(h_{22} + \frac{1}{R_L} \right)} \quad (6.16)$$

In assenza del generatore $h_{12} v_2$ la maglia di ingresso è indipendente da quella di uscita, per cui l'eq. 6.16 potrebbe essere scritta direttamente, senza passare per il sistema 6.14: una tensione di segnale v_s genera nella maglia di ingresso una corrente $i_1 = v_s / (R_s + h_{11})$; la corrente i_1 , moltiplicata per il coefficiente di amplificazione h_{21} produce una tensione $-v_2$ all'uscita, circolando nella resistenza $1 / (h_{22} + 1 / R_L)$ formata da h_{22} ed R_L in parallelo. Il segno $-$ compare a causa delle convenzioni nei versi positivi di correnti e tensioni ed esprime il fatto che se l'amplificazione di corrente h_{21} è positiva, allora l'amplificazione di tensione è negativa.

Nel caso limite $R_S = 0$ e $1/R_L = 0$, cioè $R_L = \infty$, l'eq. 6.16 si riduce alla equazione

$$A'_0 = \frac{v_2}{v_s} = \frac{-h_{21}}{h_{11} h_{22}} \quad (6.17)$$

che descrive il funzionamento dell'amplificatore *di per sé* indipendentemente dagli altri circuiti cui è collegato, dando l'amplificazione di tensione tra il terminale di ingresso 1 e quello di uscita 2 quando all'uscita non è richiesto di erogare corrente.

Al momento che si collega all'ingresso un generatore v_s che presenta una resistenza interna R_S non nulla, questa crea con h_{11} un partitore che introduce una attenuazione $h_{11} / (R_s + h_{11})$. Analogamente all'uscita, al momento che si collega una resistenza R_L , si introduce un effetto di partizione con la conduttanza h_{22} pari a $h_{22} / (h_{22} + 1 / R_L)$.

Moltiplicando il guadagno A'_0 della eq. 6.17 per questi due coefficienti di attenuazione si ritorna ad avere l'eq. 6.16.

Tornando all'eq. 6.15, dividendo numeratore e denominatore per $(h_{11} + R_S)(h_{22} + 1/R_L)$ e sostituendo l'espressione di A_0 data dalla eq. 6.16, si ottiene:

$$A = \frac{v_o}{v_s} = \frac{A_0}{1 + A_0 h_{12}} \quad (6.18)$$

Il generatore $h_{12}v_2$ introduce nel circuito una reazione: la frazione h_{12} del segnale di uscita v_2 viene riportata nella maglia di ingresso e sottratta o sommata al segnale v_s , secondo che h_{12} sia positivo o negativo. L'effetto della reazione è di modificare il guadagno dividendolo per il fattore $(1 + A_0 h_{12})$. Questo è un risultato molto generale, che verrà analizzato più a fondo nel cap. 11 dedicato alla reazione.

Mentre con $h_{12} = 0$ il generatore v_s vede una resistenza di ingresso dell'amplificatore $R_i = h_{11}$ e la resistenza di carico R_L vede una conduttanza di uscita $G_o = 1/R_o = h_{22}$, la presenza di $h_{12} \neq 0$ complica le cose.

Supponendo di applicare una tensione $v_s = v_1$ all'ingresso 1 mediante un generatore a resistenza interna R_S nulla, si può calcolare dal sistema 6.14 il valore di i_1 e da questo il rapporto $R_i = v_1/i_1$, che è la resistenza di ingresso che l'amplificatore presenta realmente. Per la corrente i_1 si ha:

$$i_1 = v_1 \cdot \frac{\left(h_{22} + \frac{1}{R_L} \right)}{h_{11} \left(h_{22} + \frac{1}{R_L} \right) - h_{21}h_{12}} \quad (6.19)$$

e da questa

$$R_i = \frac{v_1}{i_1} = h_{11} \left[1 - \frac{h_{21}}{h_{11} \left(h_{22} + \frac{1}{R_L} \right)} \cdot h_{12} \right] \quad (6.20)$$

Sostituendo nella eq. 6.20 l'espressione di A_0 dell'eq. 6.16 (tenendo presente che si è posto $R_S = 0$) si ha come risultato che anche la resistenza di ingresso h_{11} risulta modificata del fattore $(1 + A_0 h_{12})$, questa volta a moltiplicare:

$$R_i = h_{11}(1 + A_0 h_{12}) \quad (6.21)$$

Per il calcolo di $G_o = 1/R_o$ si procede in maniera analoga, supponendo di porre $v_s = 0$ e di applicare una tensione v_2 ai terminali di uscita.

$$\begin{aligned} \frac{i_2}{v_2} &= G_o = h_{22} + h_{21} \frac{i_1}{v_2} = h_{22} - h_{21} \frac{h_{12}}{R_S + h_{11}} = \\ &= h_{22} \left[1 - \frac{h_{21}}{h_{22}(R_S + h_{11})} \cdot h_{12} \right] \end{aligned} \quad (6.22)$$

Anche questa volta il coefficiente del termine h_{12} è l'amplificazione data dall'eq. 6.16, avendo posto $1/R_L = 0$, per cui

$$G_o = \frac{1}{R_o} = h_{22} (1 + A_0 h_{12}) \quad (6.23)$$

Anche la conduttanza di uscita h_{22} viene quindi moltiplicata per il fattore $(1 + A_0 h_{12})$.

È da notare che l'espressione di A_0 della eq. 6.21 contiene il valore di R_L (e non di R_S), mentre l'espressione di A_0 della eq. 6.23 contiene il valore di R_S (e non di R_L): la resistenza di carico R_L influisce sul valore di R_i , la resistenza R_S del generatore influisce sul valore di R_o . In un circuito ben progettato queste influenze devono essere piccole. Il valore di R_S ovviamente non ha influenza su R_i , come quello di R_L non ha influenza su R_o .

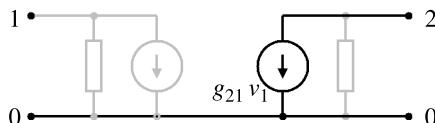
Combinando le equazioni 6.15 e 6.20 si ottiene l'amplificazione di corrente

$$A_i = \frac{i_2}{i_1} = \frac{v_2}{R_L} / \frac{v_1}{R_i} = -A_v \cdot \frac{R_i}{R_L} = \frac{h_{21}}{1 + h_{22}R_L} \quad (6.24)$$

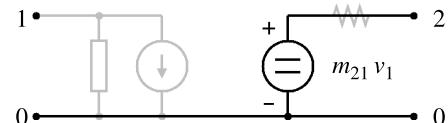
Questo risultato si sarebbe potuto ottenere anche da una analisi diretta del circuito, tenendo conto che $1/(1 + h_{22}R_L)$ è il rapporto di partizione tra R_L e h_{22} .

6.6 Amplificatori ideali

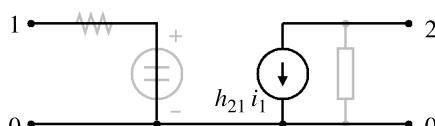
L'amplificatore di fig. 6.1 non è l'unico tipo di amplificatore ideale possibile. Ciascuno dei quattro modelli di fig. 6.3 può essere semplificato azzerando tutti i parametri eccetto il termine di trasferimento diretto g_{21} o m_{21} che è alla base della descrizione dell'amplificazione, cioè del controllo della variabile di uscita da parte della varabile di ingresso. Si ottengono i quattro amplificatori ideali di fig. 6.5.



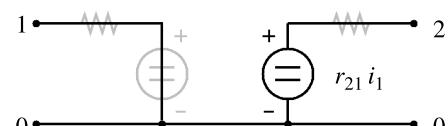
Amplificatore a transconduttanza



Amplificatore di tensione



Amplificatore di corrente



Amplificatore a transresistenza

Figura 6.5: Modelli lineari di amplificatore ideali.

Per ognuno è riportato il nome, che ne riassume le caratteristiche. L'amplificatore di tensione (in alto a destra) è probabilmente il più comune ed il più noto. All'uscita

si comporta come un generatore di tensione ideale, all'ingresso presenta conduttanza nulla (resistenza infinita). In basso a destra c'è l'amplificatore a transresistenza. Ha le stesse caratteristiche di uscita ma resistenza di ingresso nulla, in quanto la variabile di ingresso è una corrente.

La coppia di amplificatori nella colonna di sinistra presenta le stesse caratteristiche di ingresso dei corrispondenti sulla destra, ma all'uscita si comporta come un generatore ideale di corrente. Abbiamo quindi un amplificatore di corrente (in basso) ed un amplificatore a transconduttanza (in alto).

Le matrici che descrivono questi quattro modelli contengono tutte tre parametri nulli e quindi non è possibile effettuare le trasformazioni del par. 6.4.1 da una famiglia di parametri ad un'altra: è impossibile rappresentare un generatore di tensione ideale mediante un generatore di corrente o una resistenza nulla con una conduttanza e viceversa.

6.7 Reti passive

Eccetto alcune limitazioni come quelle viste alla fine del paragrafo precedente, qualunque rete lineare a due porte, sia attiva che passiva, può essere descritta con i modelli di fig. 6.3. A volte è utile usare questo formalismo anche per circuiti molto semplici, che potrebbero senza difficoltà essere risolti direttamente. Il circuito in fig. 6.6b è la descrizione mediante i parametri h del partitore resistivo in fig. 6.6a.

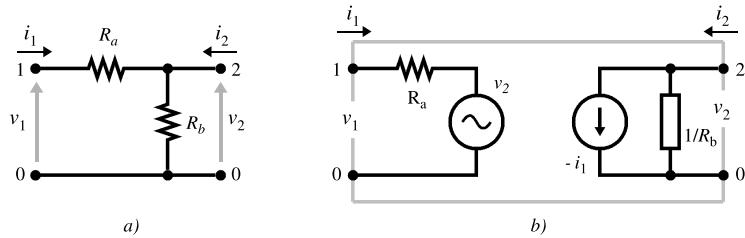


Figura 6.6: Partitore resistivo (a) e suo circuito equivalente a parametri h (b).

I valori dei parametri si determinano direttamente applicando le eq. 6.7:

$$h_i = \left. \frac{v_1}{i_1} \right|_{v_2=0} = R_a \quad h_f = \left. \frac{i_2}{i_1} \right|_{v_2=0} = -1$$

$$h_r = \left. \frac{v_1}{v_2} \right|_{i_1=0} = 1 \quad h_o = \left. \frac{i_2}{v_2} \right|_{i_1=0} = \frac{1}{R_b}$$

La condizione $v_2 = 0$ equivale a: *cortocircuitando i due terminali di uscita*; la condizione $i_1 = 0$ equivale a: *lasciando aperti i due terminali di ingresso*. Il segno $-$ in h_f deriva dal fatto che la corrente che scorre in R_a ha versi opposti se vista al nodo 1 o al nodo 2; per convenzione le correnti si considerano positive quando *entrano* nel nodo.

Descrizioni di questo tipo torneranno utili nel cap. 14 in uno studio più generale dei circuiti con reazione.

Capitolo 7

Il transistor *bjt*

La sigla *bjt* è un acronimo per *bipolar junction transistor* (transistor bipolare a giunzione). L'aggettivo *bipolare* è dovuto al fatto che nel funzionamento di questo dispositivo sono coinvolti entrambi i tipi di portatori di carica, lacune ed elettroni. Per contro, in altri dispositivi come i *mosfet* ed i *jfet* (cap. 9) solo uno dei due tipi di portatori contribuisce al funzionamento, o le lacune o gli elettroni.

La struttura di un transistor *bjt* è riportata schematicamente in fig. 7.1. Un transistor *bjt* è composto da tre regioni di semiconduttore drogato alternativamente *NPN* oppure *PNP*, dette *emettitore* (*E*), *base* (*B*) e *collettore* (*C*). Le tre regioni formano due giunzioni *P-N*, con la regione centrale, la base, in comune.

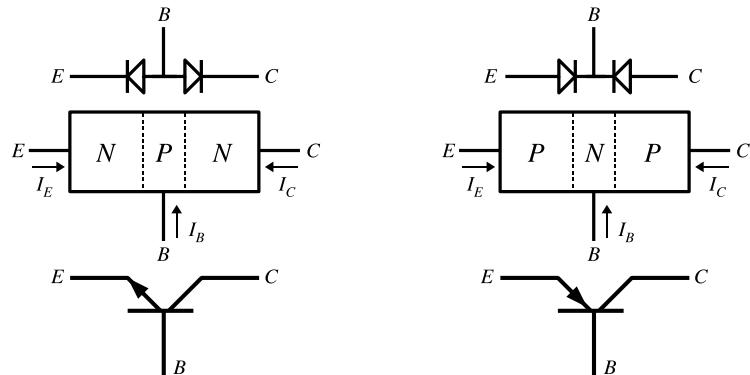


Figura 7.1: Transistor *bjt* di tipo *NPN* (a sinistra) e di tipo *PNP* (a destra). Il transistor *bjt* è formato da due giunzioni *PN* affacciate, con una regione in comune. In basso: i simboli utilizzati per indicare il transistor nei circuiti.

Costruttivamente in un transistor *bjt* emettitore e collettore non sono simmetrici: il collettore è di dimensioni maggiori e circonda ed avvolge l'emettitore (fig. 7.2). La regione di base, che separa emettitore e collettore, è molto sottile. Il suo spessore è compreso tra 0.1 e 100 μm .

7.1 Equazioni di Ebers-Moll

In figura 7.3 è descritto il funzionamento di un transistor *NPN*. Si considera per

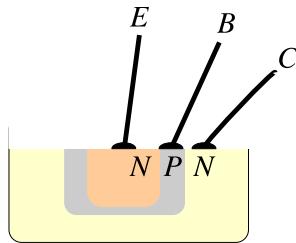
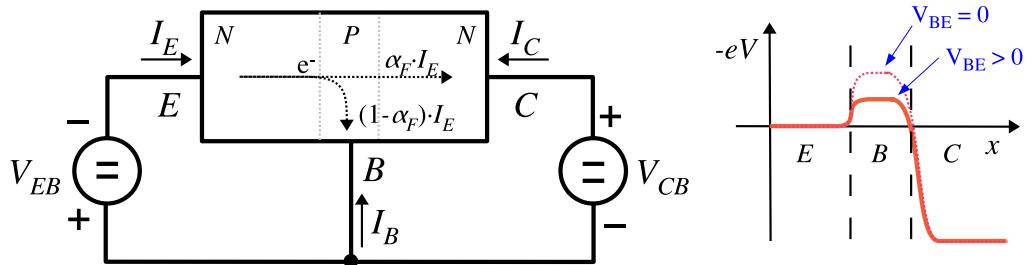
Figura 7.2: Geometria di principio di un transistor *bjt* di tipo *NPN*.

Figura 7.3: *Effetto transistor*: la frazione α_F degli elettroni iniettati dall'emettitore nella base (che costituiscono la quasi totalità della corrente della giunzione base-emettitore) attraversa, per diffusione, il sottile spessore della base e raggiunge la regione di collettore. Solo la frazione $(1 - \alpha_F)$ costituisce la corrente di base I_B . Il grafico a destra riporta l'andamento dell'energia potenziale degli elettroni nelle tre regioni. La tensione di polarizzazione V_{BE} riduce l'altezza della barriera di potenziale tra base ed emettitore. Gli elettroni che entrano dall'emettitore nella base e raggiungono per diffusione la giunzione base-collettore vengono catturati dal campo elettrico nella giunzione e cadono nella buca di potenziale della regione di collettore.

convenzione positivo il verso di una corrente *entrante* nel dispositivo. I due generatori V_{EB} e V_{CB} forniscono le tensioni di polarizzazione applicate rispettivamente al diodo base-emettitore ed al diodo base-collettore. Le correnti I_E ed I_C sono descritte dalle due equazioni:

$$I_E = -I_{ES} \left(e^{V_{BE}/\eta V_T} - 1 \right) + \alpha_R I_{CS} \left(e^{V_{BC}/\eta V_T} - 1 \right) \quad (7.1)$$

$$I_C = -I_{CS} \left(e^{V_{BC}/\eta V_T} - 1 \right) + \alpha_F I_{ES} \left(e^{V_{BE}/\eta V_T} - 1 \right) \quad (7.2)$$

note come *equazioni di Ebers-Moll*.

I primi due termini (in nero) nelle eq. di Ebers-Moll derivano dalla applicazione della *equazione della giunzione* 4.21: I_{ES} è la corrente di saturazione inversa della giunzione base-emettitore, I_{CS} l'analogo per la giunzione base-collettore. I segni meno sono dovuti alla convenzione adottata per il verso delle correnti. I due termini di destra (in rosso) non comparirebbero nelle equazioni se il transistor fosse realmente fatto solo di due diodi indipendenti; compaiono invece a causa dell'*effetto transistor*. La giunzione base-emettitore viene realizzata con un drogaggio asimmetrico, cioè con una concentrazione di donatori nell'emettitore molto maggiore di quella degli accettori nella base. A causa di questo, quando la giunzione base-emettitore è polarizzata

direttamente, la corrente è costituita per la quasi totalità da elettroni che vengono iniettati dalla regione di emettitore in quella di base, mentre il contributo dovuto alle lacune dalla regione di base all'emettitore risulta trascurabile. A causa della sottigliezza della regione di base e della sua geometria, una frazione α_F dei portatori (elettroni) iniettati nella base raggiunge per diffusione la giunzione con il collettore, viene catturata dal campo elettrico presente in questa regione e va a contribuire alla corrente di collettore I_C , dando origine al secondo termine della eq. 7.2. Il valore di α_F è tipicamente molto vicino all'unità, tra 0.95 e 0.999.

Analogamente, quando viene polarizzata direttamente la giunzione collettore-base, una frazione α_R dei portatori iniettati dal collettore nella base raggiunge l'emettitore e va a contribuire alla corrente I_E con il secondo termine della eq. 7.1; il valore di α_R in genere è sensibilmente minore di 1, tipicamente dell'ordine di 0.5. Questa asimmetria tra α_F ed α_R è dovuta alle caratteristiche costruttive delle due giunzioni.

7.2 Regioni operative del bjt

7.2.1 Regione di interdizione

Il transistor *bjt* si dice in interdizione o spento (*cutoff*) quando entrambe le giunzioni sono contropolarizzate e quindi le uniche correnti circolanti sono le deboli correnti di saturazione inversa I_{ES} e I_{CS} delle due giunzioni.

7.2.2 Regione attiva

Un transistor *bjt* si dice che lavora nella *regione attiva* quando la giunzione base-emettitore è polarizzata direttamente e quella base-collettore è contropolarizzata. In queste condizioni il primo termine della eq. 7.1 ed il secondo della 7.2 risultano preponderanti e le due equazioni, trascurando anche i due termini -1 rispetto agli esponenziali, si semplificano in:

$$I_E = -I_{ES} e^{V_{BE}/\eta V_T} \quad (7.3)$$

$$I_C = \alpha_F I_{ES} e^{V_{BE}/\eta V_T} = -\alpha_F I_E \quad (7.4)$$

Utilizzando la legge dei nodi (I legge di Kirchhoff) si ricava la corrente di base:

$$I_B = -I_E - I_C = \frac{I_C}{\alpha_F} - I_C = I_C \frac{1 - \alpha_F}{\alpha_F} \quad (7.5)$$

da cui

$$I_C = \beta_F I_B \quad (7.6)$$

$$I_E = -(\beta_F + 1) I_B \quad (7.7)$$

con

$$\beta_F = \frac{\alpha_F}{1 - \alpha_F} \quad (7.8)$$

Poiché α_F ha valori molto vicini ad uno, il valore di β_F risulta molto maggiore di uno, tipicamente tra 20 e 1000.

7.2.3 Regione di saturazione

Portando a zero la tensione di polarizzazione V_{CB} e successivamente invertendone la polarità la giunzione collettore-base passerà dalla interdizione alla conduzione, per cui il primo termine della eq. 7.2 non sarà più trascurabile, ma darà un contributo via via più significativo, di segno opposto a quello del secondo termine. A causa della dipendenza esponenziale delle correnti dalle tensioni di polarizzazione, per un valore di V_{CB} molto vicino a quello di V_{EB} , cioè per $V_{CE} \simeq 0$ i due termini avranno lo stesso valore e si elideranno dando in totale $I_C = 0$.

Un transistor bjt utilizzato come amplificatore normalmente è *polarizzato* per lavorare nella regione attiva.

Le due regioni di interdizione e saturazione sono le regioni tipiche di lavoro del bjt utilizzato come interruttore nei circuiti digitali.

7.3 Configurazioni circuitali

Il transistor è un dispositivo a tre terminali, come il dispositivo modello di figura 6.2. Di conseguenza nell'impiego come amplificatore uno dei tre terminali dovrà essere in comune ad entrambe le porte, di ingresso e di uscita, cioè collegato a massa. Sono possibili tre diverse *configurazioni circuitali*: *common base* (base comune o a massa), *common emitter* (emettitore comune o a massa) e *common collector* (collettore comune o a massa). In ogni caso, come si vede dagli schemi di fig. 7.4, la base fa sempre parte del circuito di ingresso ed il collettore di quello di uscita. A seconda della configurazione adottata si ottengono circuiti amplificatori con caratteristiche sensibilmente diverse.

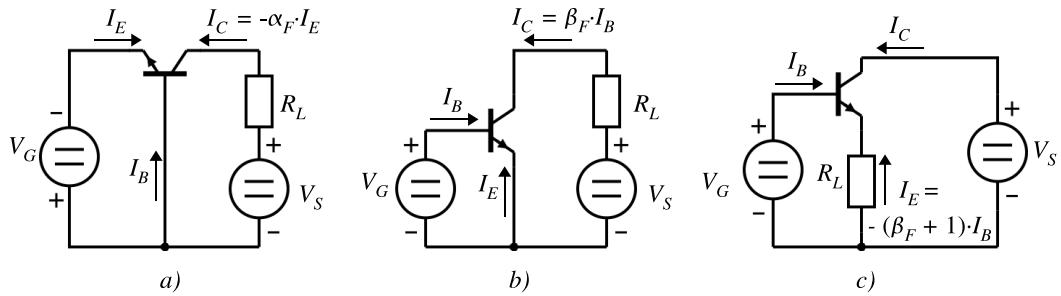


Figura 7.4: Configurazioni di utilizzo del transistor bjt: a) base comune; b) emettitore comune; c) collettore comune.

7.3.1 base comune (*common base*)

Il terminale di ingresso è l'emettitore, quello di uscita il collettore (fig. 7.4a). La corrente di ingresso I_E , inviata dal generatore V_G , viene *trasferita* dal transistor (*transistor*) nel circuito di uscita come corrente I_C , con una amplificazione di corrente $-\alpha_F$ leggermente minore di uno in modulo (eq. 7.4); in prima approssimazione tale corrente non dipende dalla tensione V_{CB} . La tensione che si genera ai capi

della *resistenza di carico* R_L , se il valore di R_L è scelto correttamente, è molto maggiore della tensione di ingresso all'emettitore. Un transistor con base a massa è quindi un amplificatore di tensione. Questa è stata storicamente la prima configurazione di utilizzo del *bjt*.

7.3.2 emettitore comune (*common emitter*)

Il terminale di ingresso è la base, quello di uscita il collettore (fig. 7.4b). Il generatore V_G polarizza la giunzione base-emettitore esattamente come nel caso della configurazione con base a massa, ma questa volta è chiamato a erogare solo la debole corrente di base I_B , che è β_F volte più piccola della corrente I_C (eq. 7.6). Un transistor bjt con emettitore a massa è quindi un amplificatore di corrente con fattore di amplificazione β_F , oltre ad essere un amplificatore di tensione analogamente al circuito con base a massa. Per queste ragioni la configurazione con emettitore comune è quella di utilizzo più frequente.

7.3.3 collettore comune (*common collector*)

Il terminale di ingresso è la base, quello di uscita l'emettitore (fig. 7.4c). La corrente che circola nella resistenza di carico R_L è $(\beta_F + 1)$ volte quella della base (eq. 7.7): il transistor bjt con collettore a massa è un amplificatore di corrente. L'amplificazione di tensione invece risulta leggermente minore di uno, in quanto la tensione che si trova all'uscita sul carico R_L è quella del generatore di ingresso V_G diminuita della frazione necessaria per polarizzare la giunzione base-emettitore. Il circuito con collettore comune è detto anche *inseguitore di tensione* (*voltage follower* o *emitter follower*).

7.4 Curve caratteristiche

Nella fig. 7.5 sono riportate in forma grafica le relazioni tra correnti e tensioni di ingresso e di uscita per un *bjt NPN* ad emettitore comune, secondo quanto previsto dalle equazioni di Ebers-Moll.

Le curve in fig. 7.5a descrivono la corrente di collettore I_C in funzione della tensione collettore-emettitore V_{CE} e della corrente di base I_B . La relazione $I_C(I_B)$ con V_{CE} costante è la *caratteristica di trasferimento diretto*, la relazione $I_C(V_{CE})$ con I_B costante è la *caratteristica di uscita*.

In fig. 7.5b è riportata la relazione tra corrente di base I_B e tensione base-emettitore V_{BE} . Anche questo grafico esprime due relazioni: la *caratteristica di ingresso* $I_B(V_{BE})$ e la *caratteristica di trasferimento inverso* $I_B(V_{CE})$, rispettivamente a V_{CE} e V_{BE} costanti. Il grafico contiene una sola curva perché, secondo le equazioni di Ebers-Moll, la relazione $I_B(V_{BE})$ è praticamente indipendente da V_{CE} in tutta la regione attiva. Si osserva la curva tipica di un diodo direttamente polarizzato, a parte il fatto che la corrente I_B è solo la frazione $1/(\beta_F + 1)$ della corrente totale attraverso la giunzione emettitore-base (eq. 7.7).

La curva blu nella figura 7.5a separa la regione operativa attiva, a destra, dalla regione di saturazione, a sinistra: i punti di intersezione con le curve corrispondenti ai diversi valori di I_B sono quelli per cui $V_{BE} = V_{CE}$ e quindi $V_{CB} = 0$.

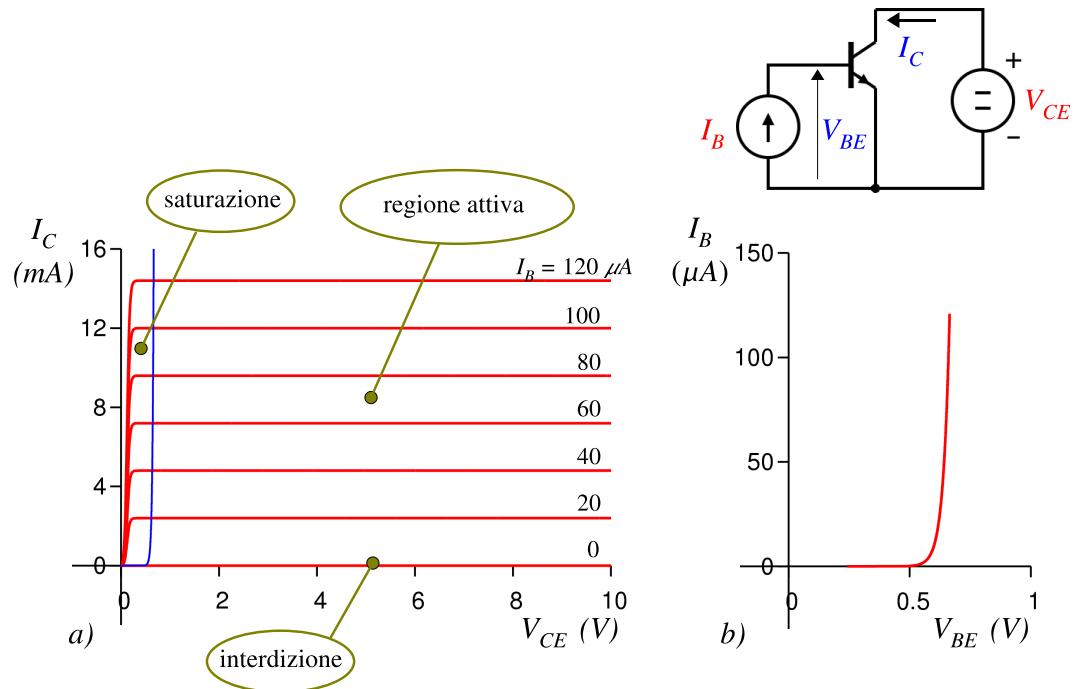


Figura 7.5: Curve caratteristiche di un transistor *NPN* bjt nella configurazione con emettitore a massa. *a*: caratteristiche di trasferimento diretto e di uscita; *b*: caratteristiche di trasferimento inverso e di ingresso. Le curve si riferiscono ad un transistor con $\beta_F = 120$. In alto: il circuito di misura. In rosso sono indicate le due variabili indipendenti, in blu le due dipendenti.

La curva più bassa, con $I_B = 0$, corrisponde al transistor spento, o in interdizione. Dalla curva di fig. 7.5*b* si vede che questa condizione si verifica praticamente per $V_{BE} < 0.5 \div 0.6V$, cioè quando il diodo base-emettitore non è sufficientemente polarizzato.

7.5 Effetto Early

Le equazioni di Ebers-Moll prevedono che nella regione attiva la corrente I_C dipenda solo dalla tensione V_{BE} (eq. 7.4) o dalla corrente I_B (eq. 7.6) e questo è evidente nel grafico in fig. 7.5*a*, dove le curve a destra della linea blu diventano rettilinee e parallele all'asse orizzontale. Una descrizione più accurata del comportamento di un transistor bjt è riportata in fig. 7.6: nella regione attiva la corrente I_C ha una dipendenza in prima approssimazione lineare da V_{CE} , descritta dalla equazione

$$I_C = \beta_F I_B \left(1 + \frac{V_{CE}}{V_A} \right) \quad (7.9)$$

Questa dipendenza è dovuta all'*effetto Early*: lo spessore della regione di transizione della giunzione base-collettore varia al variare della polarizzazione e questo fa variare il parametro α_F (e quindi β_F).

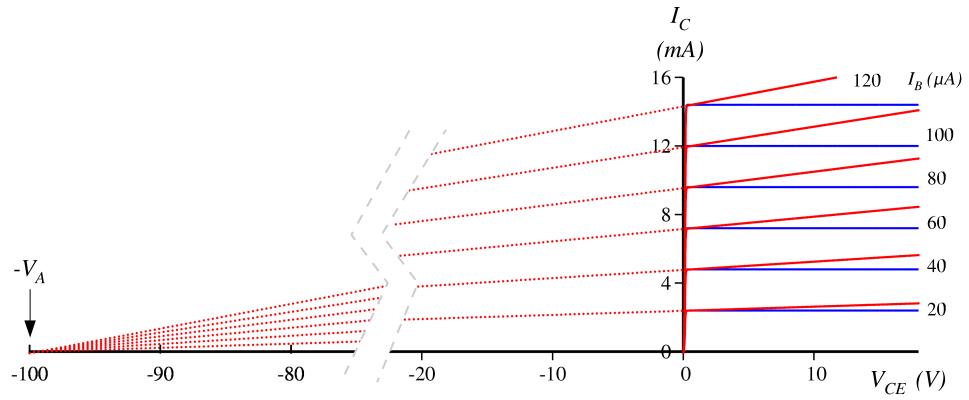


Figura 7.6: Effetto Early. Curve in blu: corrente di collettore come previsto dalle equazioni di Ebers-Moll. Curve in rosso: corrente di collettore tenendo conto dell'effetto Early. Le rette a sinistra (punteggiate) sono il prolungamento della parte rettilinea delle curve caratteristiche nella regione attiva: si incontrano sull'asse x nel punto di coordinata $-V_A$, la *tensione di Early*.

La tensione $-V_A$ corrisponde al punto dove si incontrano le rette ottenute dal prolungamento a sinistra delle curve caratteristiche nella regione attiva ed è detta *tensione di Early*.

7.6 Amplificatore ad emettitore comune

La configurazione ad emettitore comune, riportata nello schema di principio di figura 7.7a, è la più utilizzata in quanto offre contemporaneamente amplificazione di tensione e di corrente. I due generatori di corrente collegati alla base rappresentano i

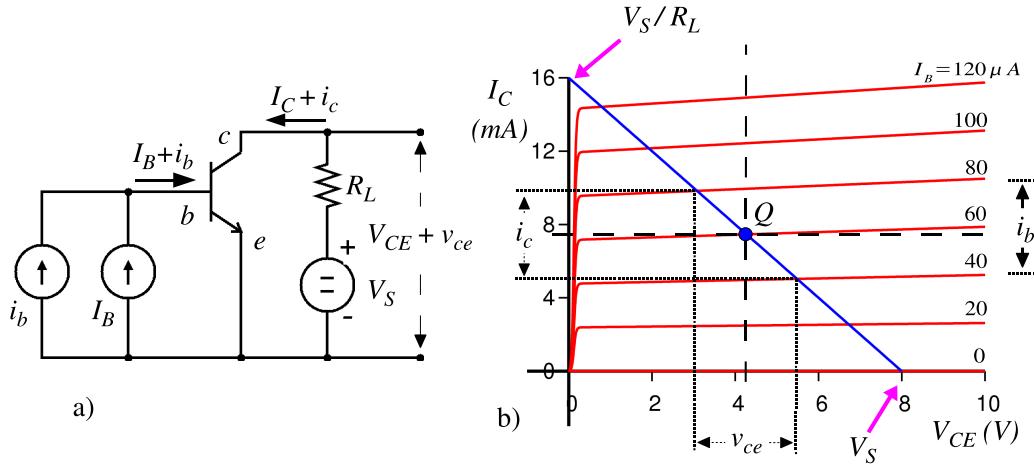


Figura 7.7: Schema di principio di un amplificatore con bjt in configurazione ad emettitore comune.

circuiti che forniscono la corrente di polarizzazione I_B che determina il punto di lavoro

e la corrente di segnale i_b , che corrisponde alle variazioni intorno al punto di lavoro. Il funzionamento è illustrato in forma grafica nella figura 7.7b, dove sono riportate le curve caratteristiche $I_C(I_B, V_{CE})$ e la retta di carico corrispondente all'equazione $I_C R_L + V_{CE} = V_S$. L'intersezione tra la retta di carico e la curva corrispondente al valore di I_B prescelto (nell'esempio $I_B = 60 \mu A$) individua il punto di lavoro Q del circuito, e quindi i valori di I_C e V_{CE} di riposo.

Il segnale di corrente i_b sovrapposto alla corrente di polarizzazione I_B nel circuito di ingresso genera nel circuito di uscita le variazioni i_e e v_{ce} indicate nel grafico. La tensione di ingresso v_{be} corrispondente alla corrente i_b può essere determinata tramite la curva di ingresso (fig. 7.5b).

L'utilizzo diretto mediante calcolo numerico (o in forma grafica) delle equazioni di Ebers-Moll 7.1 e 7.2 e della 7.9 è conveniente solo quando si è in presenza di escursioni di segnale sufficientemente grandi da portare il circuito in condizioni di non linearità significative, ad esempio nello studio dei circuiti di potenza. Negli altri casi, avendo a che fare con piccoli segnali, risulta particolarmente conveniente utilizzare i modelli lineari per i dispositivi a due porte descritti nel capitolo precedente.

7.6.1 Modello lineare a emettitore comune

Una volta stabilito il punto di lavoro e determinati i valori delle correnti e tensioni continue di polarizzazione, il circuito di fig. 7.7 può essere *linearizzato* e quindi ridisegnato in modo da contenere solo gli elementi relativi ai *segnali*, secondo uno dei modelli lineari visti al paragrafo 6.4. Le equazioni di Ebers-Moll suggeriscono di scegliere la corrente di collettore I_C come variabile dipendente di uscita e di conseguenza la tensione V_{CE} come variabile indipendente. Per quanto riguarda l'ingresso risulta altrettanto conveniente scegliere come variabile indipendente sia la tensione (e utilizzare quindi i parametri g), sia la corrente, come nell'esempio di fig. 7.7 (e utilizzare i parametri h). I circuiti relativi sono riportati in fig. 7.8. Ad ogni parametro dei modelli è stato aggiunto l'indice e per specificare che ci si riferisce ad un transistor utilizzato con emettitore comune. A seconda della convenienza la sorgente del segnale può essere rappresentata da un generatore equivalente di tensione con una resistenza interna in serie (generatore di Thèvenin v_g , R_g), o da un generatore equivalente di corrente con una conduttanza interna in parallelo (generatore di Norton i_g , G_g). La resistenza R_L è l'equivalente del *carico*, cioè del circuito utilizzatore collegato all'uscita (L da *Load*); si suppone che nel circuito di carico non siano presenti sorgenti di segnale.

I due elementi più importanti del modello lineare sono il parametro di trasferimento diretto (g_{fe} , h_{fe}) ed il parametro di ingresso (g_{ie} , h_{ie}); gli altri due elementi, raffigurati in colore chiaro nello schema, in prima approssimazione possono essere considerati nulli ed ignorati. Il fatto che la conduttanza di uscita (g_{oe} , h_{oe}) sia nulla in prima approssimazione chiarisce il motivo per cui i parametri g ed h siano i più convenienti nella descrizione del *bjt*. I modelli che utilizzano i parametri m ed r contengono nel circuito di uscita un generatore di tensione. Quindi, in prima approssimazione, avremmo dovuto utilizzare un generatore di tensione infinita in serie ad una resistenza infinita, in modo da simulare il comportamento del circuito di collettore, che è quello di un generatore di corrente. Saremmo finiti in una situazione decisamente scomoda da trattare matematicamente.

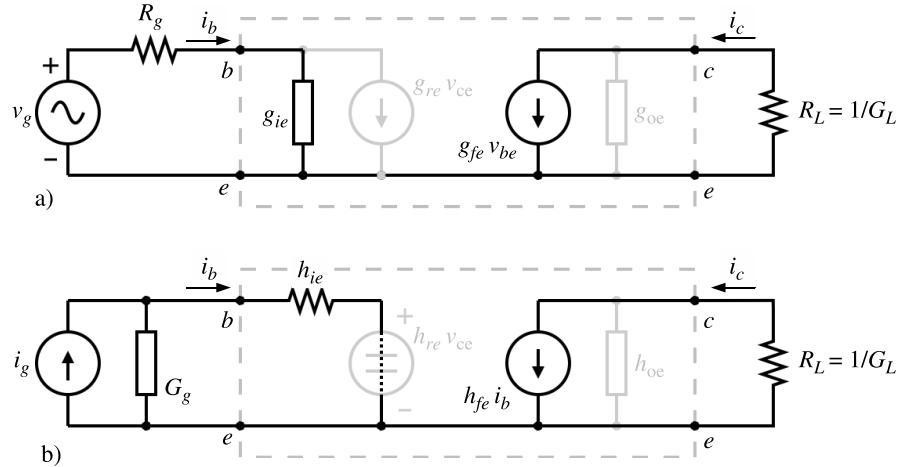


Figura 7.8: Modello lineare a parametri g (in alto) e a parametri h (in basso) per un amplificatore con transistor bjt ad emettitore comune. La sorgente del segnale può essere descritta sia come generatore di tensione che come generatore di corrente ($R_g = 1/G_g$, $v_g = i_g R_g$).

7.6.2 Valori dei parametri g ed h in corrente continua e bassa frequenza

I valori dei parametri dei modelli di fig. 7.8 si ricavano dalle equazioni che descrivono il funzionamento fisico del dispositivo, viste nei paragrafi precedenti. Dalla eq. 7.4 si ottiene il valore di g_{fe} , che coincide, a meno del coefficiente $\alpha_F \simeq 1$, con la conduttanza dinamica del diodo base-emettitore direttamente polarizzato:

$$g_{fe} = \left. \frac{i_c}{v_{be}} \right|_{v_{ce}=0} = \left. \frac{\partial I_C}{\partial V_{BE}} \right|_{V_{CE}} = \alpha_F \frac{I_{ES}}{\eta V_T} e^{V_{BE}/\eta V_T} \simeq \frac{|I_E|}{\eta V_T} \quad (7.10)$$

Come è stato già osservato relativamente alla eq. 4.36 per il diodo direttamente polarizzato, questo valore prescinde da ogni ipotesi sul modello di diodo e dipende solo dalla temperatura e dalla corrente del punto di lavoro. Questo è un risultato notevole, perché così uno dei parametri fondamentali nella descrizione del *bjt* come amplificatore risulta già noto *a priori*, indipendentemente dalla scelta del transistor.

Per quanto riguarda il secondo parametro del modello di fig. 7.8a, si ha:

$$g_{ie} = \left. \frac{i_b}{v_{be}} \right|_{v_{ce}=0} = \left. \frac{\partial I_B}{\partial V_{BE}} \right|_{V_{CE}} = \left. \frac{\partial I_B}{\partial I_C} \right|_{V_{CE}} \cdot \left. \frac{\partial I_C}{\partial V_{BE}} \right|_{V_{CE}} = \frac{g_{fe}}{\beta_f} \quad (7.11)$$

Il parametro $\beta_f = 1/(\partial I_B / \partial I_C)$ è in stretta connessione con il parametro β_F definito dalle equazioni 7.6 e 7.8, anzi coinciderebbe con questo se il valore di β_F fosse veramente una costante. In realtà β_F dipende in parte dalla corrente di emettitore I_E e dalla tensione collettore-emettitore V_{CE} per cui, in accordo con la notazione adottata per indicare i segnali ed i valori in corrente continua relativi alla polarizzazione, si ha:

$$\beta_F = \frac{I_C}{I_B} \quad \beta_f = \left. \frac{\partial I_C}{\partial I_B} \right|_{V_{CE}} = \left. \frac{i_c}{i_b} \right|_{v_{ce}=0} \quad (7.12)$$

Il valore di β_f , è una caratteristica costruttiva del singolo transistor e può spaziare in un intervallo di valori tra 20 e 1000, analogamente a β_F .

I parametri h corrispondenti a g_{fe} e g_{ie} sono il guadagno diretto di corrente

$$h_{fe} = \left. \frac{i_c}{i_b} \right|_{v_{ce}=0} = \beta_f = \frac{g_{fe}}{g_{ie}} \quad (7.13)$$

e la resistenza di ingresso

$$h_{ie} = \left. \frac{v_{be}}{i_b} \right|_{v_{ce}=0} = \frac{1}{g_{ie}} \quad (7.14)$$

in accordo con le regole di trasformazione viste nel par. 6.4.1.

Parametri di uscita e trasferimento inverso

Le equazioni di Ebers-Moll non prevedono alcuna dipendenza di I_C da V_{CE} (a parte la minuscola corrente di saturazione inversa del diodo base-collettore) né di I_B da V_{CE} . In questa approssimazione sono eguali a zero g_{re} , g_{oe} , h_{re} e h_{oe} e si ottengono i modelli semplificati a due soli parametri di fig. 7.8, in cui il funzionamento del transistor (in corrente continua ed in bassa frequenza) è determinato completamente quando siano noti il valore di β_f , che dipende dalle caratteristiche costruttive del singolo dispositivo, ed il valore della corrente di emettitore I_E , che determina il valore di g_{fe} .

In un transistor reale i parametri h_{oe} , h_{re} , g_{oe} e g_{re} risultano piccoli, ma non nulli, a causa degli effetti che sono stati trascurati nella prima approssimazione. Ad esempio, l'eq. 7.9 introduce una correzione al modello di Ebers-Moll, tenendo conto dell'effetto Early, cioè della dipendenza di I_C da V_{CE} . Da questa equazione possiamo calcolare

$$h_{oe} = \left. \frac{i_c}{v_{ce}} \right|_{i_b=0} = \left. \frac{\partial I_C}{\partial V_{CE}} \right|_{I_B} = \frac{\beta_F I_B}{V_A} = \frac{I_C}{V_A + V_{CE}} \quad (7.15)$$

Il valore della conduttanza di uscita h_{oe} definita dalla eq. 7.15 differisce dall'analoga g_{oe} , che invece è data dalla equazione

$$g_{oe} = \left. \frac{i_c}{v_{ce}} \right|_{v_{be}=0} = \left. \frac{\partial I_C}{\partial V_{CE}} \right|_{V_{BE}} \quad (7.16)$$

La differenza tra le due definizioni sta nelle condizioni di misura: g_{oe} è la conduttanza di uscita misurata con $v_{be} = 0$, cioè con la base cortocircuitata all'emettitore (per i segnali) mentre h_{oe} è la conduttanza di uscita con $i_b = 0$, cioè con il circuito di base aperto per quanto riguarda i segnali. Questa differenza è legata al fatto che anche i parametri h_{re} e g_{re} nel transistor reale sono non nulli. A causa dell'effetto Early, una tensione v_{ce} applicata all'uscita tra collettore ed emettitore induce una piccola corrente $v_{ce}g_{re}$ o tensione $v_{ce}h_{re}$ nel circuito di base. Il singolo stadio amplificatore con transistor *bjt* non è rigorosamente unilaterale.

Nella tabella 7.1 sono riportati i valori tipici dei parametri h e g in diverse condizioni di polarizzazione per un transistor *NPN* di uso generale per bassa frequenza modello *BC107C*.

I_E	0.1	1	10	mA
h_{ie}	130	16	2.1	$k\Omega$
h_{fe}	360	570	660	A/A
h_{re}	2.6	0.48	0.23	mV/V
h_{oe}	20	39	300	$\mu A/V$
g_{ie}	7.7	63	480	$\mu A/V$
g_{fe}	2.8	36	320	mA/V
g_{re}	-0.2	-0.3	-1.1	$\mu A/V$
g_{oe}	13	22	230	$\mu A/V$

Tabella 7.1: Valori tipici dei parametri h e g per un transistor $BC107C$ per $V_{CE} = 10$ V e diversi valori della corrente di emettitore.

La dispersione nel valore dei parametri tra due esemplari dello stesso modello di transistor è piuttosto alta. I valori di h_{fe} , h_{ie} e g_{ie} possono arrivare a differire di un fattore 2 dal valore nominale; per gli altri parametri le differenze possono arrivare ad essere dell'ordine di $\pm 10 \div 20\%$. È bene però ricordare che queste dispersioni nei valori non possono essere considerate errori indipendenti, in quanto i parametri sono legati tra di loro dalle eq. 7.11 e 7.13-7.14.

7.6.3 Amplificazione di tensione e di corrente

Se si considerano solo i parametri f_e e i_e i circuiti di fig. 7.8 sono sufficientemente semplici da poter essere risolti utilizzando direttamente le equazioni 6.7. Per il circuito in figura 7.8a si ottengono le due equazioni:

$$i_b = g_{ie} v_{be} \quad (7.17)$$

$$i_c = g_{fe} v_{be} \quad (7.18)$$

da cui è possibile calcolare il coefficiente di amplificazione (o *guadagno*) di corrente e di tensione

$$A_i = \frac{i_c}{i_b} = \frac{g_{fe}}{g_{ie}} \quad (7.19)$$

$$A_v = \frac{v_{ce}}{v_{be}} = \frac{-i_c R_L}{v_{be}} = -g_{fe} R_L \quad (7.20)$$

Ripetendo il calcolo per il circuito di fig. 7.8b (o semplicemente sostituendo a g_{fe} e g_{ie} le espressioni date dalle eq. 7.13 e 7.14) si ottiene

$$A_i = h_{fe} \quad (7.21)$$

$$A_v = -\frac{h_{fe}}{h_{ie}} R_L \quad (7.22)$$

Moltiplicando i due coefficienti di amplificazione si ottiene l'*amplificazione di potenza* A_p :

$$A_p = |A_i \cdot A_v| = \frac{g_{fe}^2}{g_{ie}} \cdot R_L = \frac{h_{fe}^2}{h_{ie}} \cdot R_L \quad (7.23)$$

Il calcolo completo per un circuito come quello di fig. 7.8b tenendo conto di tutti i parametri è stato già svolto nel par. 6.5. Dalle eq. 6.24 e 6.15 si hanno amplificazione di corrente e di tensione

$$A_i = \frac{i_c}{i_b} = \frac{h_{fe}}{1 + h_{oe}R_L} \quad (7.24)$$

$$A_v = \frac{v_{ce}}{v_g} = \frac{-h_{fe}}{(h_{ie} + R_g) \left(h_{oe} + \frac{1}{R_L} \right) - h_{fe}h_{re}} \quad (7.25)$$

Per valutare la differenza tra il risultato dell'eq. 7.24 e quello della 7.21 si deve tener conto del fatto che la resistenza R_L fa parte anche dei circuiti di polarizzazione (fig. 7.7a) e praticamente si trova quasi sempre ad avere un valore che è almeno un ordine di grandezza inferiore ad $1/h_{oe}$. Quindi $h_{oe}R_L \simeq 0.1$ e

$$A_i \simeq 0.9 \cdot h_{fe} \quad (7.26)$$

Per quanto riguarda l'effetto sul guadagno di tensione A_v dato dalla eq. 7.25, si deve porre $R_G = 0$, in quanto quello che si deve calcolare realmente è il rapporto tra la tensione di uscita (v_{ce}) e quella effettivamente presente ai terminali di ingresso (v_{be}), indipendentemente dal rapporto di partizione aggiuntivo dovuto alla resistenza R_G , che può avere qualunque valore. Per le considerazioni precedenti, però, si deve continuare a porre $h_{oe}R_L \simeq 0.1$. Con i valori dei parametri riportati nella tab. 7.1 per la corrente $I_E = 1 \text{ mA}$, si ottiene

$$A_v \simeq -0.95 \cdot \frac{h_{fe}}{h_{ie}} R_L \quad (7.27)$$

Dalle eq. 6.20 e 6.22 si possono calcolare la resistenza di ingresso e la conduttanza di uscita

$$R_i = h_{ie} \left[1 - \frac{h_{fe}h_{re}}{(h_{oe} + 1/R_L)h_{ie}} \right] \quad (7.28)$$

$$G_o = h_{oe} \left[1 - \frac{h_{fe}h_{re}}{(h_{ie} + R_g)h_{ie}} \right] \quad (7.29)$$

Utilizzando le stesse condizioni che hanno portato ai risultati 7.26 e 7.27 si ottiene

$$R_i = 0.90 \cdot h_{ie} \quad (7.30)$$

$$G_o = 0.57 \cdot h_{oe} \quad (7.31)$$

L'ultimo risultato potrebbe far pensare ad una correzione drammaticamente importante sulle caratteristiche di uscita dell'amplificatore. In realtà non è così. Si è già visto che h_{oe} rappresenta un effetto del secondo ordine, che contribuisce solo per circa il 10% alla conduttanza di uscita, determinata prevalentemente da R_L . L'eq. 7.31 indica che questa correzione può arrivare ad essere dimezzata nel caso limite $R_G = 0$, con un effetto quindi dell'ordine del 5%.

Confrontando i risultati 7.26, 7.27, 7.30 e 7.31 con quelli delle equazioni precedenti si vede come l'effetto prodotto da h_{re} ed h_{oe} non sia mai superiore al 10%, minore dell'effetto dovuto alla dispersione sui valori stessi dei parametri. Questo giustifica

il frequente utilizzo delle eq. 7.21-7.22 invece delle più complesse eq. 7.24-7.25 ed il fatto che molto spesso l'unico parametro caratteristico fornito dai produttori per il comportamento in bassa frequenza sia h_{fe} .

Si tenga presente però che l'analisi del circuito nella sua forma più semplice perde completamente di valore quando si considera il funzionamento del dispositivo in alta frequenza. Nel cap. 16 si vedrà che al crescere della frequenza il contributo del parametro di trasferimento inverso diventa sempre più significativo, al punto da non poter essere mai ignorato.

In alcuni circuiti inoltre il parametro h_{oe} risulta il termine principale nel determinare le caratteristiche del circuito di uscita, come ad esempio nel *current mirror* descritto nel par. 7.10.

7.7 Amplificatore a collettore comune e base comune

7.7.1 Emitter follower

Nel circuito in figura 7.9 la resistenza di carico R_L è stata inserita sul terminale di emettitore anziché sul collettore, realizzando così la configurazione a *collettore comune* del par. 7.3. Infatti, nel modello linearizzato per piccoli segnali (fig. 7.9b) il generatore di alimentazione V_S viene sostituito dalla sua resistenza interna, cioè da un corto circuito verso massa.

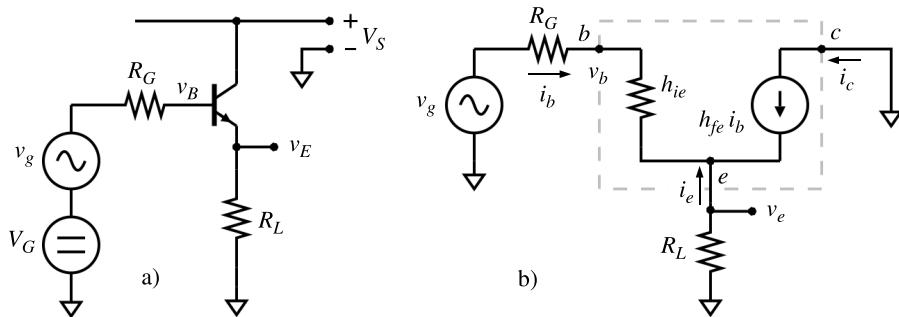


Figura 7.9: Emitter follower.

Nel circuito linearizzato di fig. 7.9b si considerano solo i due termini principali h_{ie} ed h_{fe} e si trascurano gli effetti di h_{oe} ed h_{re} . Nella resistenza R_L scorre la corrente

$$i_e = -(1 + h_{fe})i_b \quad (7.32)$$

per cui il potenziale dell'emettitore è

$$v_e = (1 + h_{fe}) R_L i_b \quad (7.33)$$

e quello della base

$$v_b = h_{ie} i_b + (1 + h_{fe}) R_L i_b \quad (7.34)$$

Rispetto al generatore v_g è come se in serie alla resistenza di ingresso h_{ie} si trovasse una resistenza di valore $(1 + h_{fe}) R_L$. Allora la tensione sull'emettitore è

$$v_e = v_b \cdot \frac{(1 + h_{fe}) R_L}{h_{ie} + (1 + h_{fe}) R_L} \quad (7.35)$$

Considerando che di solito R_L ed h_{ie} sono dello stesso ordine di grandezza mentre h_{fe} è dell'ordine di 100, si ha $h_{ie} \ll (1 + h_{fe})R_L$ e $v_e \simeq v_b$ con una differenza in meno dell'ordine del per cento (o di pochi per cento). Questa è la ragione del nome dato a questo circuito: *emitter follower, inseguitore di emettitore*. La tensione di uscita (sull'emettitore) *segue* strettamente la tensione all'ingresso, rimanendone sempre appena sotto.

Supponendo di porre $v_g = 0$ e di applicare dall'esterno la tensione v_e all'emettitore, si avrà una corrente

$$i_e = -i_b \cdot (h_{fe} + 1) = \frac{v_e}{h_{ie} + R_g} \cdot (h_{fe} + 1) \quad (7.36)$$

da cui si ottiene la resistenza di uscita R_o all'emettitore del transistor:

$$R_o = \frac{v_e}{i_e} = \frac{h_{ie} + R_g}{h_{fe} + 1} \quad (7.37)$$

La resistenza R_L , esterna al dispositivo, si viene a trovare in parallelo a R_o .

In definitiva, il comportamento dell'emitter follower è descritto dalle equazioni:

$$A_v = \frac{v_e}{v_b} = \frac{(1 + h_{fe}) R_L}{h_{ie} + (1 + h_{fe}) R_L} \simeq 1 \quad (7.38)$$

$$A_i = \frac{i_e}{i_b} = -(h_{fe} + 1) \quad (7.39)$$

$$R_i = \frac{v_b}{i_b} = h_{ie} + (1 + h_{fe}) R_L \quad (7.40)$$

$$R_o = \frac{v_e}{i_e} = \frac{h_{ie} + R_g}{h_{fe} + 1} \quad (7.41)$$

Per queste sue caratteristiche, l'applicazione principale dell'*emitter follower* è come amplificatore di corrente ed adattatore di impedenza.

7.7.2 Reazione di emettitore

Osservando il circuito di fig. 7.9b si può osservare che l'inserzione di una ulteriore resistenza tra il collettore e massa non ne modifica il comportamento, almeno fino a che si assume il modello semplificato con i due soli parametri h_{ie} ed h_{fe} , cioè fino a che il generatore di corrente $h_{fe}i_b$ viene considerato ideale e quindi con resistenza interna infinita. Questa configurazione è riportata nel circuito in fig. 7.10a con il suo circuito linearizzato equivalente in fig. 7.10b.

Dall'equazione 7.40 si può ricavare

$$i_b = \frac{v_b}{R_i} = \frac{v_b}{h_{ie} + (1 + h_{fe}) R_E} \quad (7.42)$$

$$v_c = -i_c R_C = -i_b h_{fe} R_C \quad (7.43)$$

$$= -\frac{v_b h_{fe} R_C}{h_{ie} + (1 + h_{fe}) R_E} \simeq -\frac{R_C}{R_E} v_b \quad (7.44)$$

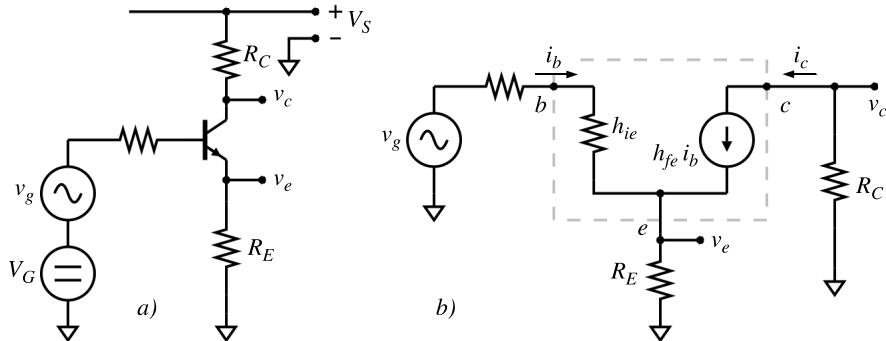


Figura 7.10: Circuito con resistenza di reazione negativa in serie all'emettitore e resistenza di carico sul collettore.

L'ultimo passaggio della eq. 7.44 è valido quando h_{ie} è trascurabile rispetto a $(1 + h_{fe})R_E$ e h_{fe} è molto maggiore di uno. Mentre la seconda ipotesi è praticamente sempre vera, la prima è valida solo per valori di R_E non troppo piccoli. È chiaro che nel caso limite di $R_E = 0$ l'eq. 7.44 si riduce alla 7.22 e l'ultimo passaggio smette di avere senso.

Inserire una resistenza R_E in serie all'emettitore di un transistor ha quindi l'effetto di ridurre l'amplificazione del circuito ad emettitore comune fino al valore limite R_C/R_E . Questo può apparire un controsenso: si costruisce un circuito per *amplificare* un segnale, poi si introduce nel circuito un elemento in più per *deteriorare* questa prestazione. In realtà, una analisi attenta del comportamento del circuito fa vedere che a fronte di questa perdita nell'amplificazione, si ottengono miglioramenti in altre prestazioni. Ad esempio, la resistenza di ingresso vista dal generatore v_g passa da h_{ie} in assenza della resistenza R_E al valore $h_{ie} + (1 + h_{fe})R_E$, e in molte situazioni una resistenza di ingresso elevata è una caratteristica preziosa. Un altro vantaggio portato dalla resistenza R_E lo si vedrà nello studio dei circuiti di polarizzazione. Ma la caratteristica certamente più importante è data dal fatto che introducendo R_E nel circuito, via via che se ne aumenta il valore, il coefficiente di amplificazione viene a dipendere sempre meno dalle caratteristiche del transistor, come risulta dalla 7.22, e tende invece al valore stabile R_C/R_E .

7.7.3 Amplificatore a base comune

Nell'amplificatore a base comune il segnale di ingresso viene applicato all'emettitore ed il segnale di uscita è prelevato dal collettore (fig. 7.11).

Osservando il circuito, si vede facilmente che l'amplificazione di tensione, considerando solo h_{ie} ed h_{fe} , è la stessa che per il circuito ad emettitore comune, con una inversione di segno (eq. 7.22):

$$A_v = \frac{h_{fe}}{h_{ie}} R_L \quad (7.45)$$

Per l'amplificazione di corrente, applicando la prima legge di Kirchhoff al nodo di emettitore si ottiene:

$$A_i = \frac{i_c}{i_e} = \frac{h_{fe}i_b}{-h_{fe}i_b - i_b} = \frac{-h_{fe}}{h_{fe} + 1} \simeq -1 \quad (7.46)$$

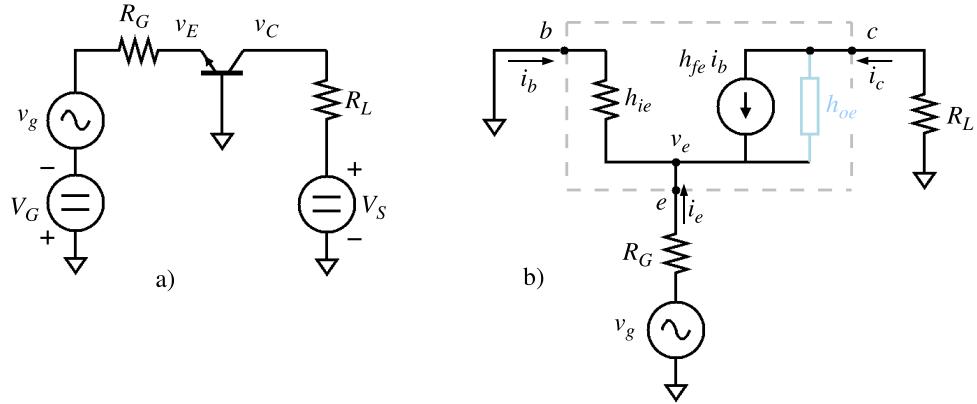


Figura 7.11: Amplificatore con base comune o a massa.

Utilizzando le stesse considerazioni delle eq. 7.36 e 7.37 si può calcolare la resistenza di ingresso

$$R_i = \frac{h_{ie}}{h_{fe} + 1} \quad (7.47)$$

che risulta quindi $h_{fe} + 1$ volte più piccola di quanto trovato nella configurazione ad emettitore comune.

La resistenza di uscita del transistor in fig. 7.11 risulta ovviamente infinita, se si ignora h_{oe} . Tenendo conto anche del contributo di h_{oe} la resistenza di uscita dipende fortemente dal valore della resistenza R_G . Nel caso che il generatore di segnale sia un generatore ideale di tensione, si ha $R_G = 0$ e quindi $R_o = 1/h_{oe}$, come nel circuito ad emettitore comune. Nel caso opposto, con un generatore di segnale che si comporta come un generatore ideale di corrente, si ha $R_G = \infty$, $i_b = -i_c$ e

$$v_o = i_c h_{ie} + \frac{1}{h_{oe}} (i_c + h_{fe} i_c) \quad (7.48)$$

da cui

$$R_o = \frac{v_o}{i_c} = h_{ie} + \frac{1 + h_{fe}}{h_{oe}} \simeq \frac{h_{fe}}{h_{oe}} \quad (7.49)$$

In queste condizioni il *bjt* con base a massa presenta una resistenza di uscita di uno o due ordini di grandezza maggiore di quella con emettitore a massa.

Il calcolo completo che tenga conto anche del coefficiente h_{re} può portare ad un risultato sensibilmente diverso, senza però alterare gli ordini di grandezza.

7.7.4 Parametri a base comune e collettore comune

Nei paragrafi precedenti i circuiti amplificatori a base comune e a collettore comune sono stati analizzati utilizzando il modello lineare del *bjt* ad emettitore comune e risolvendo le reti lineari risultanti. È ugualmente possibile introdurre reti lineari analoghe a quelle di fig. 7.8, con i terminali di base ed emettitore o quelli di base e collettore scambiati. Le reti saranno formalmente identiche e conterranno ognuna ancora quattro parametri, identificati dagli indici b e c invece di e : h_{ib} , h_{ic} , h_{fb} , ecc. Mediante trasformazioni algebriche si può passare dai valori dei parametri per una

data configurazione a quelli per un'altra. Ma alla fine i calcoli sono equivalenti a quelli visti nei paragrafi precedenti.

7.8 La polarizzazione del bjt

Nei paragrafi precedenti è stato analizzato il funzionamento di alcuni circuiti amplificatori supponendo che il punto di lavoro fosse già stato opportunamente stabilito ed impostato. Ad esempio, in fig. 7.7 compare il generatore I_B , che schematizza i circuiti di polarizzazione necessari. Nella figura 7.12 sono riportati due esempi di amplificatore ad emettitore comune completi di tutti i circuiti di polarizzazione. Il primo circuito è molto semplice, ma si vedrà che ha problemi tali da essere in pratica inutilizzabile; il secondo circuito invece descrive un sistema di polarizzazione molto diffuso nella realizzazione di amplificatori di segnale in corrente alternata.

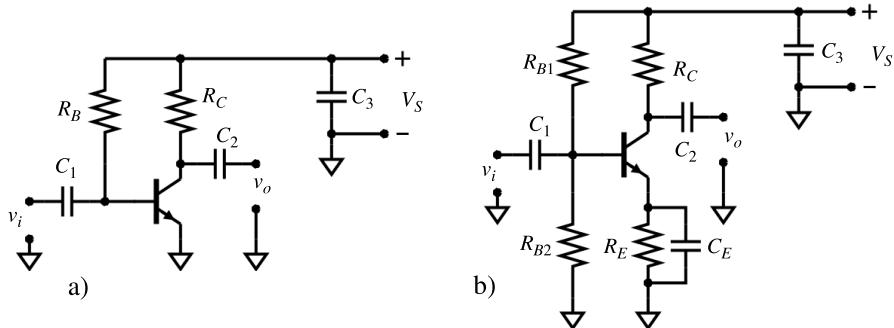


Figura 7.12: Due esempi di amplificatori ad emettitore comune completi dei circuiti di polarizzazione.

Nel calcolo dei valori dei componenti che costituiscono un circuito è necessario tener presente che alcuni di questi non possono essere determinati che da condizioni esterne al circuito stesso. Ad esempio la resistenza di carico R_C è determinata principalmente dalla *resistenza di uscita* che si vuole il circuito presenti verso lo stadio successivo cui il segnale di uscita è destinato; la tensione di alimentazione V_S deve essere scelta tra le tensioni disponibili, compatibilmente con l'ampiezza dei segnali che devono essere trattati; il transistor, in genere, è scelto cercando un accordo tra costi e prestazioni, ecc. È abbastanza normale che le varie esigenze risultino in conflitto tra di loro e che la progettazione richieda una serie di scelte di compromesso, anche tra diverse possibili soluzioni circuitali.

Supponiamo quindi di avere individuato i valori $V_S = 12 V$ ed $R_C = 10 k\Omega$ come idonei alle ipotetiche condizioni di utilizzo. Supponiamo anche di avere scelto un transistor con $\beta_F = 120$ e $V_A = 100 V$. Anche la posizione del punto di lavoro deve essere stabilita in base alle applicazioni del circuito. Frequentemente si sceglie un punto che sia grosso modo al centro della regione attiva, in modo da poter trattare in maniera simmetrica segnali di entrambe le polarità. A volte invece altre soluzioni sono preferibili: se un circuito è destinato ad amplificare segnali essenzialmente unipolari, ad esempio impulsi positivi, risulterà più conveniente posizionare il punto di lavoro in prossimità della regione di interdizione, per poter disporre della più ampia *dinamica di uscita* possibile. Come si vede dalle curve caratteristiche di fig. 7.5 e 7.7 la regione

attiva si estende fino a valori di V_{CE} di pochi decimi di volt (formalmente il confine con la regione di saturazione si trova a $V_{CE} = V_{BE}$, quindi a circa 0.65 V, ma fino a che la giunzione base-collettore non raggiunge una polarizzazione sufficiente, 0.4...0.5 V, corrispondenti a 0.2...0.1 V per V_{CE} , non si ha praticamente alcun effetto). Poniamo, nel nostro caso, $V_Q = 6$ V (e di conseguenza $I_Q = 0.6$ mA).

7.8.1 Polarizzazione con singola resistenza di base

Per completare la polarizzazione, nel circuito in fig. 7.12a rimane solo da determinare il valore di R_B che dia la giusta corrente di base:

$$R_B = \frac{V_S - V_{BE}}{I_B} = \frac{V_S - V_{BE}}{I_C} \cdot \beta_F \cdot \left(1 + \frac{V_{CE}}{V_A}\right) \quad (7.50)$$

$$\simeq \frac{12\text{ V} - 0.65\text{ V}}{0.6\text{ mA}} \cdot 120 \cdot 1.06 = 2.41\text{ M}\Omega \simeq 2.2\text{ M}\Omega \quad (7.51)$$

Nel calcolo sono state introdotte due approssimazioni: il valore di V_{BE} è stato posto a 0.65 V, con una approssimazione di circa ± 0.1 V nel numeratore della frazione e quindi un errore dell'1%; il valore ottenuto per la resistenza (2.41 MΩ) è stato arrotondato al valore disponibile più vicino (2.2 MΩ), introducendo un ulteriore errore di circa il 10% (per una discussione sui valori *disponibili* dei componenti, si veda l'appendice A). In definitiva, ci si potrebbe aspettare una corrente di base (e quindi di collettore) superiore di circa il 10% rispetto al valore preventivato, e quindi una tensione di lavoro V_Q di 5.4 V invece dei 6 V desiderati, e questo sarebbe nella maggior parte dei casi un risultato accettabile. La fonte di errore veramente grossa, che non è stata sino ad ora considerata e che rende invece il circuito in pratica inutilizzabile, è costituita dall'indeterminazione sul valore di β_F . In figura 7.13 è riportata la distribuzione dei valori di β_F per un campione di 50 transistor dello stesso modello (2N2222A) provenienti da uno stesso acquisto. È evidente come i valori siano

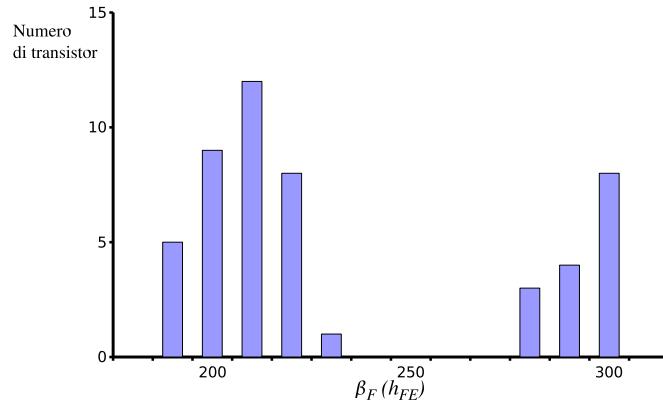


Figura 7.13: Distribuzione dei valori di β_F in un campione scelto casualmente di 50 transistor modello 2N2222A provenienti da uno stesso acquisto.

raggruppati intorno a 210 e 300: il campione era chiaramente costituito da esemplari provenienti da due cicli di lavorazione diversi. Mentre semiconduttori ottenuti in uno stesso ciclo di lavorazione su un unico *wafer* di Silicio tendono ad avere valori dei

parametri molto simili (e questa caratteristica viene sfruttata con enorme vantaggio in circuiti come *l'amplificatore differenziale* o il *current mirror* descritti nei prossimi paragrafi), si ha invece una grande dispersione di valori quando si confrontano componenti che provengono da lavorazioni diverse. Tanto è che parametri come β_F spesso sono specificati dai produttori come valore nominale con una possibilità di variazione di un fattore 2 in più o in meno. In queste condizioni, il punto di lavoro del nostro

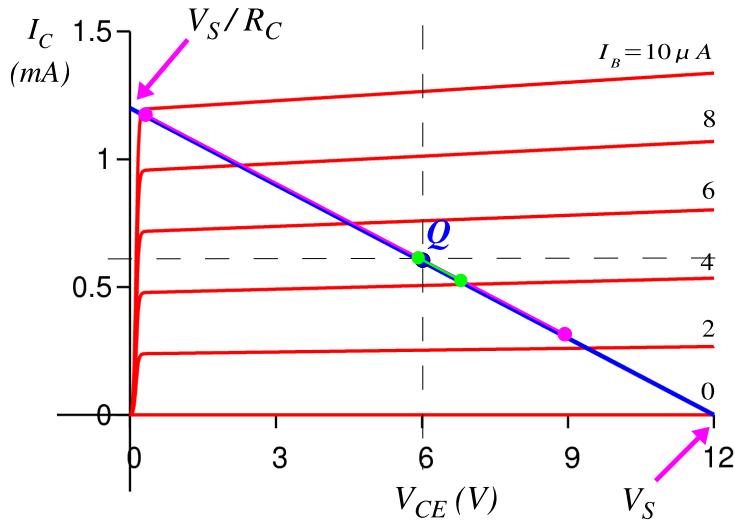


Figura 7.14: Variabilità della posizione del punto di lavoro per i due circuiti di polarizzazione di fig. 7.12 per valori di β_F tra 60 e 240. Il segmento in magenta, sovrapposto alla retta di carico in blu, si riferisce al circuito 7.12a, il segmento in verde al circuito 7.12b.

ipotetico circuito si potrebbe venire a trovare in tutto l'intervallo tra $I_C = 0.3 \text{ mA}$, $V_{CE} = 9 \text{ V}$ e $I_C = 1.2 \text{ mA}$, $V_{CE} \approx 0 \text{ V}$. In figura 7.14, dove sono riportate curve caratteristiche e retta di carico relative ai due circuiti di fig. 7.12, questo intervallo è indicato in color magenta sovrapposto alla retta di carico (in blu); ad un estremo si ha una notevole riduzione della escursione di potenziale possibile, all'altro estremo si ha addirittura il transistor in saturazione. Per poter utilizzare veramente il circuito di fig. 7.12a sarebbe necessario aggiustare il valore di R_B singolarmente su ciascun esemplare di circuito ed eventualmente riaggiustarlo in caso di sostituzione del transistor. Si tratta evidentemente di una soluzione poco pratica (ed anche costosa nel caso di una produzione su larga scala), per cui il circuito di fig. 7.12a ha valore poco più che illustrativo.

7.8.2 Polarizzazione a quattro resistenze

Il circuito di fig. 7.12b, noto come *polarizzazione a quattro resistenze*, presenta due sostanziali differenze rispetto a quello di fig. 7.12a: la base viene polarizzata tramite il partitore R_{B1}, R_{B2} e viene inserita una resistenza R_E in serie all'emettitore. La prima modifica cambia il tipo di pilotaggio della base. In fig. 7.12a si ha un pilotaggio della base *in corrente*: un generatore di tensione di valore elevato (12 V) ed una resistenza in serie di valore elevato ($2.2 \text{ M}\Omega$) stabiliscono il valore della corrente di base I_B , e

lo mantengono costante. La corrente di collettore $I_C = \beta_F I_B$ è influenzata da tutte le possibili variazioni di β_F . In fig. 7.12b invece, se le resistenze R_{B1} ed R_{B2} sono di valore sufficientemente basso, si ha un pilotaggio della base *in tensione*, come si vede dalla fig. 7.15 dove è riportato il circuito equivalente di Thèvenin della rete di alimentazione della base. La seconda modifica, l'inserimento della resistenza R_E ,

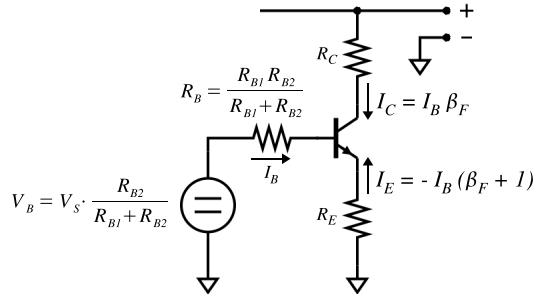


Figura 7.15: Circuito di polarizzazione a quattro resistenze con il partitore di base sostituito con il suo circuito equivalente di Thèvenin.

introduce nel circuito una *reazione di emettitore* in serie alla tensione di polarizzazione (come visto al paragrafo 7.7.2) e questo lega il punto di lavoro principalmente al valore delle resistenze, rendendolo meno dipendente dal β_F del transistor. Infatti, applicando la legge della maglia al circuito di base, si ha

$$V_B = I_B R_B + V_{BE} + I_B (\beta_F + 1) R_E \quad (7.52)$$

da cui

$$I_C = I_B \beta_F = \frac{(V_B - V_{BE}) \beta_F}{R_B + (\beta_F + 1) R_E} \quad (7.53)$$

Nella misura in cui $\beta_F \gg 1$, $(\beta_F + 1) \cdot R_E \gg R_B$ e $V_{BE} \simeq 0.65 V$, la corrente I_C è

$$I_C = \frac{V_B - 0.65 V}{R_E} \quad (7.54)$$

Nella figura 7.14 è riportato in verde, sovrapposto alla retta di carico, l'intervallo di valori in cui si viene a trovare il punto di lavoro con $R_{B1} = 100 k\Omega$, $R_{B2} = 18 k\Omega$ ed $R_E = 1.8 k\Omega$. Per lo stesso intervallo di variabilità di β_F considerato per l'altro circuito ($60 \dots 240$), la posizione del punto di lavoro rimane confinata tra 5.9 e 6.8 V. Naturalmente, questo miglioramento nella stabilità del circuito ha un prezzo: oltre alla maggiore complessità ed al maggior numero di componenti necessari, si ha una riduzione della escursione della tensione di uscita, che passa da quasi 12 V a circa 10.5 V a causa della caduta di potenziale sulla resistenza R_E ; inoltre le due resistenze del partitore di base R_{B1} ed R_{B2} si vengono a trovare, per quanto riguarda i segnali, tra base e massa, e quindi in parallelo alla resistenza di ingresso.

7.8.3 Condensatori di accoppiamento e disaccoppiamento

In entrambi i circuiti di fig. 7.12 la base del transistor non si trova al potenziale di massa: nel circuito 7.12a si trova a circa 0.6 V, mentre in 7.12b si trova a circa

1.6 V. In entrambi i casi è necessario introdurre il condensatore C_1 per isolare i circuiti di polarizzazione della base dal generatore del segnale (o dagli stadi precedenti). In mancanza di tale condensatore la base si troverebbe collegata a massa tramite la resistenza interna del generatore; se questa fosse sufficientemente piccola il transistor potrebbe anche finire per trovarsi in interdizione. Se invece la resistenza del generatore fosse molto grande, il condensatore C_1 sarebbe ancora necessario, perché il generatore potrebbe non gradire di ricevere la tensione continua presente sulla base. Il condensatore C_1 insieme con la resistenza interna del generatore e la resistenza di ingresso dell'amplificatore costituisce un filtro passa alto e introduce quindi una attenuazione delle frequenze al di sotto della sua *frequenza di taglio* (fig. 7.16). In-

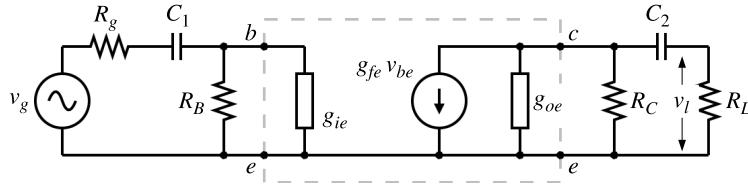


Figura 7.16: Circuiti passa alto formati dai condensatori di accoppiamento di ingresso e di uscita C_1 e C_2 .

dicando con r_i la resistenza di ingresso dell'amplificatore, formata da g_{ie} in parallelo alla resistenze di polarizzazione R_B (o R_{B1} ed R_{B2}), il rapporto tra la tensione v_{be} effettivamente presente all'ingresso e quella della sorgente v_g è

$$\frac{v_b}{v_g} = \frac{r_i}{r_i + R_g + 1/j\omega C_1} = \frac{r_i}{R_g + r_i} \cdot \frac{j\omega C_1(R_g + r_i)}{1 + j\omega C_1(R_g + r_i)} \quad (7.55)$$

L'espressione a destra nell'equazione 7.55 si compone di due parti: $r_i/(R_g + r_i)$ è l'attenuazione del segnale per frequenze molto alte ($\omega \rightarrow \infty$); l'altro termine è la risposta di un filtro passa alto con costante di tempo $\tau = C_1(R_g + r_i)$ e frequenza di taglio $f_0 = 1/(2\pi\tau)$. Un problema analogo si presenta all'uscita: il collettore si trova alla tensione continua del punto di lavoro (~ 6 V) ed è necessario isolarlo dal circuito dello stadio successivo. Il condensatore C_2 , introdotto a questo scopo, forma a sua volta un circuito passa alto con una sua propria frequenza di taglio. Indicando con r_o la resistenza di uscita dell'amplificatore, formata da g_{oe} e da R_C in parallelo, e con v_o la tensione di uscita a vuoto, cioè quando non vi è alcun carico collegato, la tensione v_l che si ottiene quando è presente il carico R_L , è:

$$\frac{v_l}{v_o} = \frac{R_L}{R_L + r_o + 1/j\omega C_2} = \frac{R_L}{R_L + r_o} \cdot \frac{j\omega C_2(R_L + r_o)}{1 + j\omega C_2(R_L + r_o)} \quad (7.56)$$

La presenza dei due condensatori C_1 e C_2 , detti di *accoppiamento* perché hanno lo scopo di accoppiare per quanto riguarda il segnale ingresso ed uscita del circuito agli altri stadi, impedisce definitivamente l'utilizzo dei circuiti di fig. 7.12 per l'amplificazione di segnali in corrente continua.

Il condensatore C_E in fig. 7.12b ha lo scopo di cortocircuitare la resistenza R_E e collegare a massa l'emettitore nel campo di frequenza dei segnali utili in modo da riportare il circuito alla configurazione originaria di *emettitore a massa*; altrimenti

l'effetto di riduzione del guadagno, utile in corrente continua per stabilizzare il punto di lavoro, si estenderebbe inevitabilmente anche alle frequenze di lavoro del circuito.

Un terzo condensatore (C_3) compare nei circuiti in parallelo ai terminali di alimentazione. Il suo scopo è di cortocircuitare i terminali di alimentazione, per quanto riguarda le frequenze dei segnali in gioco, in modo da far apparire il generatore V_S il più possibile un generatore di tensione ideale, cioè con resistenza interna nulla. In caso contrario, oltre ad avere una ulteriore resistenza su un ramo in comune tra resistenza di carico e resistenza di polarizzazione della base, con complicazioni di calcolo e di funzionamento del circuito, si avrebbe un trasferimento di segnale tra due stadi facenti capo ad uno stesso generatore di alimentazione, con effetti potenzialmente disastrosi sul funzionamento dei circuiti. Per questa sua funzione, questo condensatore è detto di *disaccoppiamento*

7.9 Amplificatore differenziale

Non è possibile produrre transistor di precisione: si è visto al paragrafo precedente che parametri come β_F sono soggetti ad una grande dispersione tra un esemplare e l'altro di uno stesso modello. Anche selezionando accuratamente i transistor in base ai valori dei parametri, rimarrebbe comunque il problema delle inevitabili dipendenze dalla temperatura. Ad esempio, la tensione V_{BE} di una giunzione base-emettitore a corrente di base costante presenta una deriva con la temperatura $\partial V_{BE} / \partial T \simeq -2.3 \text{ mV}/^\circ\text{K}$, tipica di tutte le giunzioni *PN* (par. 4.9). Questo rende vano ogni tentativo di costruire con un singolo transistor bjt un amplificatore di precisione per segnali in corrente continua che non siano già di ampiezza superiore a parecchi *mV*.

È però possibile produrre transistor con caratteristiche con ottima precisione identiche, sia per quanto riguarda i valori dei parametri che per le loro derive. Questo risultato lo si raggiunge con dispositivi realizzati contemporaneamente, sullo stesso *wafer* di Silicio, a distanza di pochi μm l'uno dall'altro, con uno stesso ciclo di lavorazione.

L'*amplificatore differenziale*, il cui schema di principio è riportato in fig. 7.17, è un circuito che impiegando due transistor dalle caratteristiche il più possibile identiche riesce ad ottenere, per quanto riguarda la stabilità, prestazioni fino a tre o quattro ordini di grandezza migliori di quelle ottenibili con il singolo transistor. Il principio che sta alla base del disegno è di duplicare nel circuito ogni fonte di errore con un'altra il più possibile identica, e combinarle in modo che i due effetti si elidano.

Come si vede nella fig. 7.17, il circuito è dotato di due ingressi, corrispondenti alle basi b_1 e b_2 dei due transistor, a cui sono collegati i due segnali v_1 e v_2 . I due transistor, che come si è detto hanno caratteristiche il più possibile eguali, hanno sul collettore due resistenze R_C , anche queste il più possibile eguali; gli emettitori sono collegati insieme ad un'unica resistenza R_E . L'alimentazione al circuito è fornita da due generatori: una tensione $+V_S$ alimenta le due resistenze di collettore R_C , una tensione $-V_S$ alimenta la resistenza di emettitore R_E . I due generatori hanno un terminale in comune, che costituisce il *circuito di massa* a cui sono riferite tutte le tensioni. Per analizzare il funzionamento del circuito, conviene scomporre idealmente i due segnali indipendenti v_1 e v_2 in un *segnale di modo comune* $v_c = (v_1 + v_2)/2$ ed un *segnale di modo differenziale* $v_d = (v_1 - v_2)$, come indicato in fig. 7.18. In questo

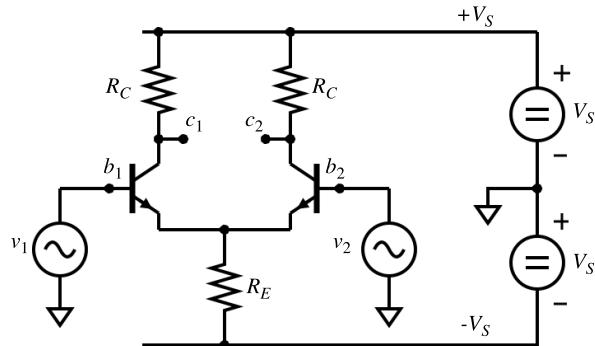


Figura 7.17: Circuito elettrico di principio di un amplificatore differenziale con due transistor bjt.

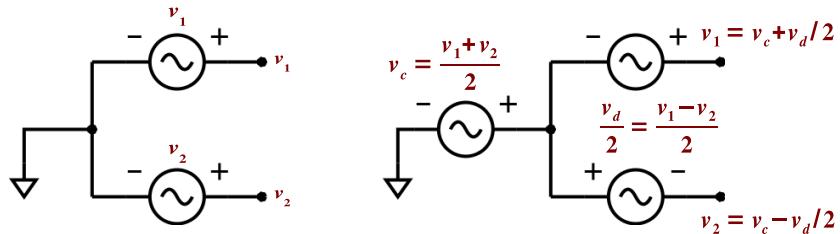


Figura 7.18: Scomposizione di due segnali indipendenti v_1 e v_2 (a sinistra) in un segnale di *modo comune* $v_c = (v_1 + v_2)/2$ ed in un segnale di *modo differenziale* $v_d = (v_1 - v_2)$ (a destra).

modo, qualsiasi coppia di valori v_1, v_2 applicata agli ingressi può essere ottenuta come somma e differenza dei due segnali v_c e v_d : $v_1 = v_c + v_d/2$, $v_2 = v_c - v_d/2$. In condizioni di comportamento lineare da parte del circuito (e vedremo che questo è vero, con le dovute approssimazioni) i segnali di uscita saranno la somma delle risposte ai due segnali, di modo comune e di modo differenziale, applicati indipendentemente.

7.9.1 Amplificazione di modo comune

Nella figura 7.19a viene esaminato il funzionamento del circuito in presenza del solo segnale di modo comune v_c , che viene applicato contemporaneamente ad entrambe le basi b_1 e b_2 . La resistenza di emettitore R_E è stata considerata idealmente composta da due resistenze identiche, di valore $2R_E$, in parallelo, associate ciascuna ad uno dei due transistor. In queste condizioni il circuito è composto da una parte destra ed una parte sinistra completamente simmetriche rispetto alla linea blu verticale; di conseguenza nel collegamento indicato in grigio tra i due emettitori e_1, e_2 la corrente è necessariamente nulla per ogni valore di v_c ed il circuito si comporta come se fosse composto da due circuiti identici ed indipendenti, ciascuno dei quali è un amplificatore con reazione di emettitore, già studiato al par. 7.7.2. Utilizzando l'equazione 7.44 si può calcolare l'*amplificazione di modo comune* A_c di ciascuna delle

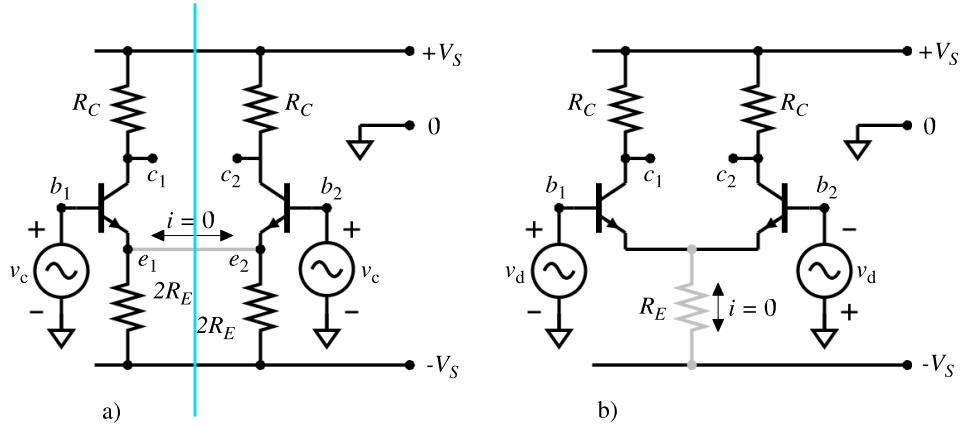


Figura 7.19: Analisi del funzionamento dell'amplificatore differenziale in presenza del solo segnale di modo comune (a) e del solo segnale di modo differenziale (b).

due metà:

$$A_c = \frac{v_{c1}}{v_{b1}} = \frac{v_{c2}}{v_{b2}} = -\frac{R_C}{2R_E} \quad (7.57)$$

Il segnale di uscita presente sui collettori c_1, c_2 è anch'esso un segnale di modo comune: $v_{c1} = v_{c2}$.

7.9.2 Amplificazione di modo differenziale

Nel circuito in figura 7.19b alle due basi viene applicato il segnale di modo differenziale, cioè uno stesso segnale v_d con polarità opposta su una base rispetto all'altra. In queste condizioni, nell'ipotesi che il circuito si comporti in modo lineare, la corrente di segnale nella resistenza R_E risulta nulla, per cui gli emettitori possono essere considerati collegati a massa per quanto riguarda i segnali. Per verificare questa

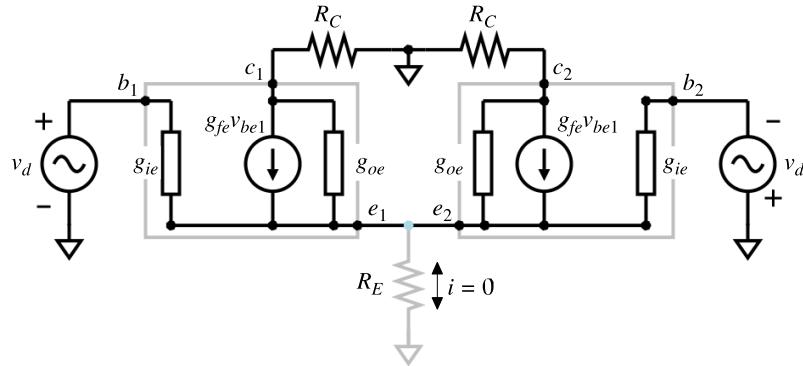


Figura 7.20: Modello lineare equivalente al circuito di fig. 7.17 in presenza del solo segnale di ingresso in modo differenziale v_d .

affermazione è necessario esaminare il modello lineare equivalente dell'amplificatore, riportato in fig. 7.20. Anche questo circuito è composto di una metà destra ed una

sinistra eguali. Poichè il circuito è lineare (condizione che nello studio del funzionamento in modo comune non era necessaria) e le tensioni applicate alle due basi b_1 e b_2 sono eguali e di segno opposto, le correnti nei due terminali di emettitore e_1 ed e_2 saranno anch'esse eguali ed opposte. Di conseguenza la corrente i nella resistenza R_E sarà nulla ed i due emettitori si troveranno sempre a potenziale di massa (relativamente ai segnali). Anche in questo caso quindi il circuito si comporta come se fosse costituito da due mezzi circuiti completamente indipendenti, ciascuno dei quali equivale, come modello lineare, all'amplificatore con emettitore comune descritto nel par. 7.6. Il coefficiente di amplificazione di modo differenziale A_d può quindi essere ottenuto scrivendo direttamente per ciascuna delle due metà del circuito 7.20

$$\frac{v_{c1}}{v_{b1}} = \frac{v_{c2}}{v_{b2}} = -\frac{g_{fe}R_C}{1 + g_{oe}R_C} \quad (7.58)$$

e quindi

$$A_d = \frac{v_{c1}}{v_{b1} - v_{b2}} = \frac{-v_{c2}}{v_{b1} - v_{b2}} = -\frac{1}{2} \cdot \frac{g_{fe}R_C}{1 + g_{oe}R_C} \quad (7.59)$$

Poichè i segnali di ingresso v_d applicati alle basi sono di polarità opposta, anche i segnali di uscita sui collettori c_1 e c_2 saranno di polarità opposta: $v_{c1} = -v_{c2}$. Il segnale presente sulle uscite è anch'esso un segnale di modo differenziale.

7.9.3 Un esempio di amplificatore differenziale

In figura 7.21 è riportato un esempio di amplificatore differenziale. La condizione

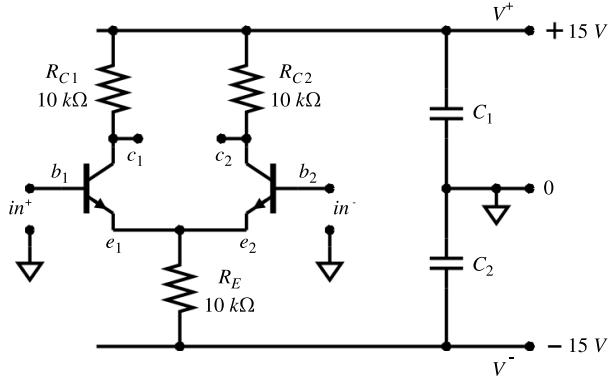


Figura 7.21: Esempio di stadio amplificatore differenziale a due transistor bjt.

di assenza di segnale sugli ingressi corrisponde ad avere entrambe le basi a potenziale 0, cioè a massa. Gli emettitori si troveranno a circa $-0.65 V$ e la corrente attraverso la resistenza R_E sarà con ottima precisione $(15 V - 0.65 V)/10k\Omega = 1.44 mA$. Per quanto si è detto sulla simmetria del circuito, la corrente si divide a metà tra i due transistor e quindi si hanno circa $0.72 mA$ in ciascuna delle due resistenze di carico ed una tensione di $7.8 V$ sui collettori. Questo è un punto di lavoro ragionevole, perchè la tensione di collettore si viene a trovare circa a metà della tensione di alimentazione.

I due condensatori C_1 e C_2 hanno la stessa funzione di disaccoppiamento rispetto alla alimentazione del condensatore C_3 visto nei circuiti di fig. 7.12.

Con i dati sopra indicati è possibile calcolare i coefficienti di amplificazione del circuito:

$$A_c = -\frac{R_C}{2R_E} = -0.5 \quad (7.60)$$

$$A_d = -\frac{1}{2} \cdot \frac{g_{fe}R_C}{1 + g_{oe}R_c} \simeq -\frac{1}{2} \cdot \frac{I_E}{\eta V_T} R_C \simeq -125 \quad (7.61)$$

Il guadagno A_d così calcolato risulterà leggermente sovrastimato (dell'ordine del 5 ... 10%), poiché si è ignorata la presenza di g_{oe} .

Sostituendo nell'eq. 7.61 per la corrente di emettitore il valore

$$I_E = \frac{1}{2} \cdot \frac{V^- - V_\gamma}{R_E} \quad (7.62)$$

e ponendo in base alle considerazioni precedenti $R_E = R_C$, si ottiene:

$$A_d = \frac{1}{4} \cdot \frac{V - v_\gamma}{\eta V_T} \cdot \frac{1}{1 + g_{oe}R_c} \quad (7.63)$$

A parte il termine correttivo $(1 + g_{oe}R_c)$ a denominatore, il guadagno A_d dipende solo dalla tensione di alimentazione e non dagli altri parametri del circuito. Il valore delle resistenze, tutte eguali, può essere scelto liberamente in base alle impedenze di ingresso/uscita che si vogliono ottenere e ad altre esigenze, ad esempio il consumo o la banda passante.

È interessante notare come dall'equazione 7.63 si ottenga il valore di A_d con una precisione dell'ordine del 10% senza fare uso di alcuna conoscenza delle caratteristiche del transistor.

7.9.4 Reiezione di modo comune

Confrontando le due espressioni 7.57 e 7.59 relative alla amplificazione dei segnali di modo comune A_c e di modo differenziale A_d ed i risultati ottenuti nel calcolo relativo all'esempio del paragrafo precedente si può vedere che $A_d \gg A_c$. In un circuito come quello di fig. 7.21 il rapporto A_d/A_c , detto *reiezione di modo comune* (*Common Mode Rejection Ratio, CMRR*), è solitamente dell'ordine di 10^2 ; nei paragrafi successivi verranno esaminate tecniche per aumentare ulteriormente di diversi ordini di grandezza questo rapporto.

L'aspetto notevole del circuito dell'amplificatore differenziale è che le più importanti fonti di errore sono segnali di modo comune mentre i segnali utili possono essere applicati agli ingressi come segnali di modo differenziale. Rendendo molto elevata la reiezione di modo comune è possibile amplificare i soli segnali utili e non amplificare o anche ridurre i segnali di errore.

I due problemi più grossi incontrati con l'amplificatore ad emettitore comune del par. 7.6 erano la presenza del gradino di tensione di 0.65 V tra base ed emettitore, che rende impossibile l'accoppiamento in corrente continua alla sorgente del segnale, e la deriva di $-2.3\text{ mV/}^\circ\text{K}$ di questa stessa tensione, che ne rende impossibile l'eliminazione mediante reti di polarizzazione per quanto elaborate. Nel circuito in fig. 7.17 e 7.21 il gradino di 0.65 V è comune ad entrambi i transistor ed eguale con notevole precisione: coppie di transistor *monolitiche*, cioè realizzate sullo stesso

substrato di Silicio, arrivano a presentare facilmente tra le due V_{BE} differenze inferiori al mV , riducibili con opportune tecniche di taratura fino al livello del μV . La tensione di ingresso di $0 V$ su entrambe le basi diventa quindi quella del punto di lavoro, ed è eliminato il problema dell'accoppiamento in corrente continua con la sorgente del segnale. La deriva delle due V_{BE} con la temperatura è un segnale di modo comune che viene quindi soppresso se si ha una sufficiente reiezione di modo comune. Poiché anche nelle migliori coppie monolitiche i transistor non saranno mai *esattamente* identici, rimarrà come fonte di errore di modo differenziale la differenza tra i coefficienti di deriva delle V_{BE} dei due transistor, compresa tra $10\mu V/^{\circ}K$ per i dispositivi più economici fino a $0.1\mu V/^{\circ}K$ per i migliori dispositivi di precisione.

Si noti che il problema dell'accoppiamento in corrente continua è stato risolto per quanto riguarda il circuito di ingresso; rimane, per ora, nel circuito di uscita, come si vede nell'esempio del paragrafo precedente, dove la tensione nominale dei collettori al punto di lavoro è di $7.8 V$.

7.10 Circuito *current mirror*

Aumentando la resistenza di emettitore R_E nell'amplificatore differenziale di fig. 7.17 o 7.21 l'amplificazione di modo comune A_c diminuisce, secondo l'equazione 7.57, mentre l'amplificazione di modo differenziale A_d rimane invariata, secondo l'equazione 7.59. Il risultato è un aumento della reiezione di modo comune.

Un modo efficiente per aumentare R_E e mantenere costanti le correnti, senza dover aumentare la tensione di alimentazione negativa, è di utilizzare la resistenza di uscita dinamica $1/g_{oe}$ di un transistor. In figura 7.22 sono riportati due schemi.

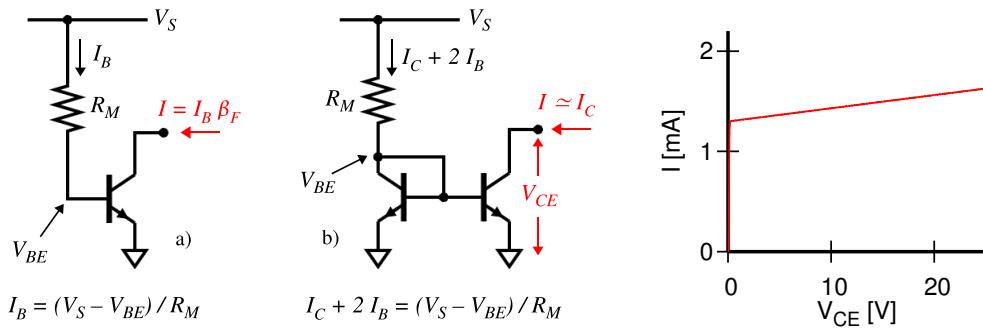


Figura 7.22: Due circuiti *generatori di corrente* che fanno uso della bassa conduttanza di uscita g_{oe} di un transistor. Il grafico a destra descrive il comportamento di un *current mirror* realizzato con una coppia monolitica CA 3046, con $R_M = 22 k\Omega$ e $V_S = 30 V$.

Nel circuito in a) il transistor è polarizzato da una corrente di base costante e stabile (se V_S è una tensione stabile e sufficientemente grande da poter trascurare le derive su V_{BE}). La corrente di collettore è $\beta_F I_B$ e soffre di tutte le instabilità legate a β_F . Nel circuito in b) il collettore e la base del transistor di sinistra di una coppia monolitica di transistor identici sono collegati insieme. La tensione tra base e massa sarà ancora $V_{BE} \simeq 0.65 V$, ma la corrente nella resistenza R_M , se il transistor ha

un β_F sufficientemente elevato, coincide praticamente con la corrente di collettore. Il secondo transistor si trova ad avere la stessa V_{BE} del primo e, formando una coppia di transistor identici, anche la stessa I_B e lo stesso β_F . Quindi la corrente nel secondo collettore sarà eguale a quella del primo e pari alla corrente nella resistenza R_M , a meno delle due piccole correnti di base I_B e delle variazioni dovute all'effetto Early (tutto questo, naturalmente, sino a che il secondo transistor si trova nella regione attiva e non nella regione di saturazione). Il nome del circuito, *current mirror*, deriva da questo modo di funzionare: la corrente nel secondo collettore ripete, *rispecchia*, la corrente nel primo.

Il grafico nella figura mostra come varia la corrente I nel collettore del transistor di destra al variare della tensione tra collettore ed emettitore. La resistenza dinamica equivalente è data dalla derivata $dV_{CE}/dI \simeq 76 \text{ k}\Omega$ a $V_{CE} = 15 \text{ V}$ e $I_C = 1,5 \text{ mA}$. La resistenza statica corrispondente è $V/I = 10 \text{ k}\Omega$.

7.11 Circuito *Darlington*

L'amplificatore *Darlington* utilizza in cascata due transistor di cui il primo in configurazione *emitter follower*.

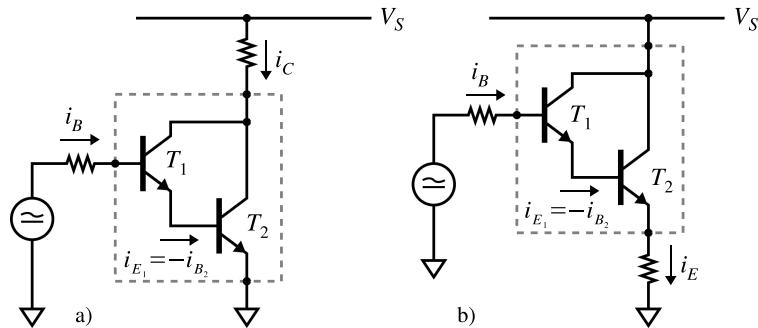


Figura 7.23: Due esempi di circuito amplificatore *Darlington*.

Nei due esempi in fig. 7.23 la corrente di emettitore del primo transistor diventa la corrente di base del secondo. Valgono quindi le relazioni:

$$i_{B_2} = -i_{E_1} = (1 + h_{FE_1}) \cdot i_{B_1} \quad (7.64)$$

$$i_C = i_{C_1} + i_{C_2} = h_{FE_1} i_{B_1} + h_{FE_2} i_{B_2} \quad (7.65)$$

$$= (h_{FE_1} + h_{FE_2} + h_{FE_1} h_{FE_2}) i_{B_1} \simeq h_{FE_1} h_{FE_2} i_{B_1} \quad (7.66)$$

In prima approssimazione il circuito Darlington si comporta come un singolo transistor con h_{FE} pari al prodotto degli h_{FE} dei due transistor costituenti la coppia. Nel circuito in fig. 7.23a il *transistor Darlington* è utilizzato ad emettitore comune; in fig. 7.23b invece è utilizzato in un circuito *emitter follower*.

Le due differenze più importanti nel funzionamento del *transistor Darlington* rispetto al normale bjt sono:

- la tensione di polarizzazione tra base di ingresso b_1 ed emettitore e_2 è $\simeq 2V_\gamma$ invece di V_γ , poiché le due giunzioni B-E sono in serie;

- la tensione di *saturazione* tra collettore (c_1, c_2) ed emettitore (e_2) è data dalla somma della tensione di saturazione del primo transistor e della V_γ del secondo e quindi non può mai essere meno di $0.8 \dots 0.9$ V invece di $0.15 \dots 0.25$ V.

7.12 Circuiti a simmetria complementare

Il circuito emitter follower riportato in fig. 7.24 soffre di un problema di asimmetria.

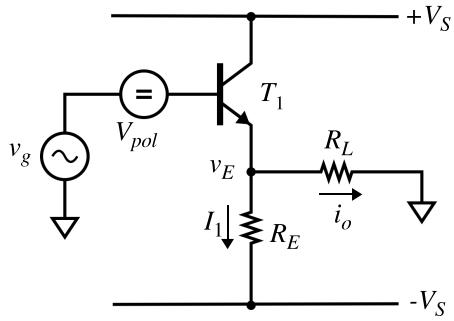


Figura 7.24: Circuito emitter follower con transistor *NPN*.

Supponiamo che per effetto del generatore V_{pol} , opportunamente scelto, si abbia a riposo $V_E = 0$ e quindi $I_1 = -I_E = V_S/R_E$. Durante la semionda positiva del segnale di ingresso v_g il transistor T_1 svolge la sua funzione di amplificatore di corrente, fornendo alle resistenze R_L ed R_E la corrente necessaria per portare l'uscita alla tensione che compete, fino al valore massimo possibile $v_E = V_S - V_{CEsat}$ e $i_o = I_{max}^+ = (V_S - V_{CEsat})/R_L$. Durante le semionde negative invece la corrente attraverso la resistenza R_L è fornita dalla resistenza R_E ; il transistor T_1 infatti può solo ridurre la propria corrente dal valore di riposo V_S/R_E fino a zero. Quindi $|I_{max}^-| = |-V_S/(R_E + R_L)| \ll I_{max}^+$. Il valore di R_E infatti non può essere scelto piccolo, perchè la dissipazione a riposo in R_E e T_1 è V_S^2/R_E , con problemi di riscaldamento dei componenti e consumo di energia tanto più grandi quanto più R_E è piccola.

Il circuito in fig. 7.25a è invece un emitter follower *a simmetria complementare*: un secondo transistor T_2 , di polarità opposta al primo (T_1 : *NPN*; T_2 : *PNP*) duplica per le semionde negative il ruolo svolto da T_1 per le positive.

Esaminando i grafici in fig. 7.25c si vede che i due transistor sono, a riposo, entrambi spenti. Solo quando v_g raggiunge il valore V_γ il transistor T_1 comincia a condurre, e analogamente T_2 quando v_g scende sotto $-V_\gamma$.

Il circuito in questa forma introduce una forte distorsione nei segnali. Il circuito riportato in fig. 7.25d elimina questo inconveniente tramite la rete di polarizzazione $R_1 - D_1 - D_2 - R_2$ che mantiene le basi di T_1 e T_2 in prossimità di $\pm V_\gamma$. Ulteriori accorgimenti sono necessari nell'uso pratico di questo circuito per stabilizzare il punto di riposo rispetto alle variazioni di temperatura, ad esempio si possono accoppiare termicamente D_1 con T_1 e D_2 con T_2 in modo che le variazioni dei potenziali V_γ delle giunzioni si compensino.

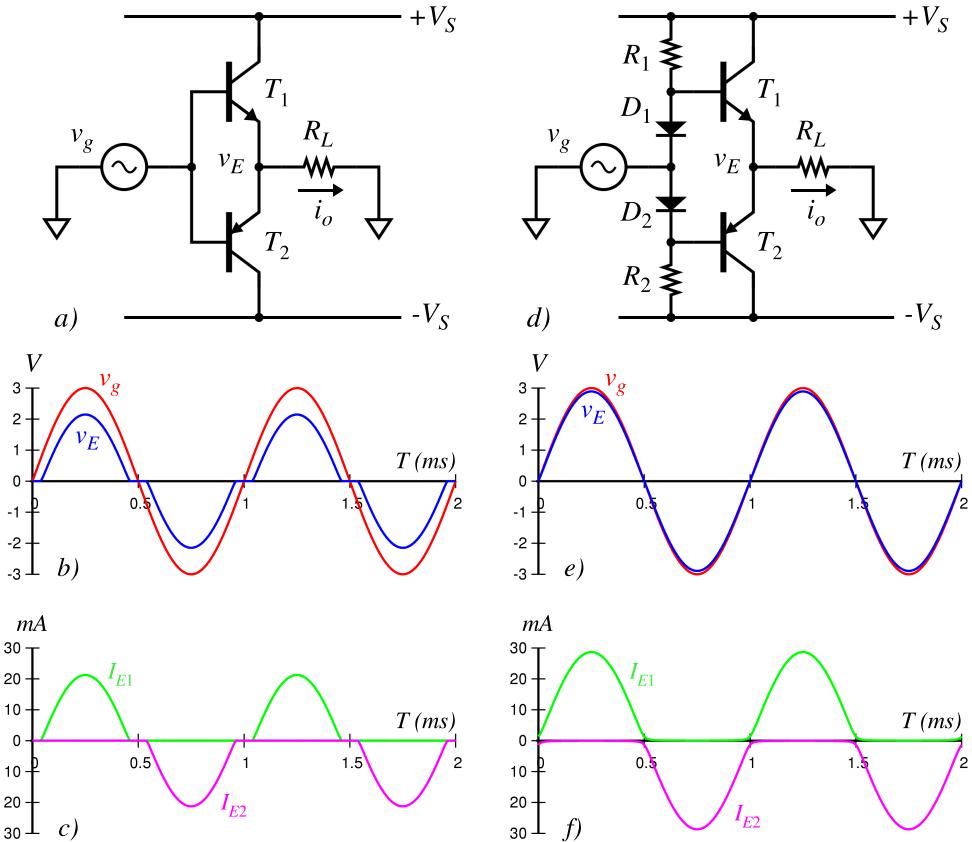


Figura 7.25: Circuito emitter follower a simmetria complementare. a) schema di principio; b) nel segnale di uscita v_E è presente una forte distorsione nell'intorno dello zero (distorsione di incrocio); in c) sono riportate le correnti nei due transistori: T_1 conduce durante la semionda positiva, T_2 durante quella negativa; intorno allo zero, quando $|v_g| < V_\gamma$, entrambi i transistori sono spenti. Nel circuito in d) i due transistori sono polarizzati alla soglia dell'accensione e la distorsione di incrocio è fortemente ridotta (e); nell'intorno dello zero i due transistori sono per breve tempo entrambi in conduzione (f).

Un circuito di questo tipo, in cui due dispositivi amplificatori lavorano alternativamente ciascuno su metà del ciclo del segnale, si dice *amplificatore in classe B*. I circuiti di fig. 7.12 o fig. 7.24, in cui un unico dispositivo è sempre attivo per l'intero ciclo del segnale, sono *amplificatori in classe A*.

Capitolo 8

Amplificatori a transistor e misure

8.1 Amplificazione

Come già visto nei capitoli precedenti, si definiscono una amplificazione di tensione ed una di corrente

$$A_v = \frac{V_o}{V_i} \quad A_i = \frac{I_o}{I_i}$$

o, nel caso dell'amplificatore differenziale (di tensione),

$$A_v = \frac{V_o}{V_+ - V_-}$$

ed una amplificazione di potenza

$$A_p = |A_v| \cdot |A_i|$$

Si ha vera amplificazione solo quando l'amplificazione di potenza A_p è maggiore di 1, indipendentemente dai singoli valori di A_v ed A_i . Si può avere $A_p > 1$ anche con $|A_v| < 1$ oppure con $|A_i| < 1$ (ma ovviamente non entrambi).

Se R_i è la resistenza che l'amplificatore presenta all'ingresso ed R_L la resistenza presentata all'uscita dall'utilizzatore del segnale (L per *Load*, cioè *carico*), l'amplificazione di potenza è

$$A_p = \frac{P_L}{P_i} = \frac{|V_o|^2}{R_L} / \frac{|V_i|^2}{R_i} = |A_v|^2 \cdot \frac{R_i}{R_L} \quad (8.1)$$

Come già visto nell'introduzione, è conveniente misurare l'amplificazione in decibel (dB):

$$A_p(\text{dB}) = 10 \cdot \log_{10} |A_p|$$

A causa della relazione quadratica nella eq. 8.1 l'amplificazione di tensione in decibel è per definizione

$$A_v(\text{dB}) = 20 \cdot \log_{10} |A_v|$$

Analogamente, si definisce l'amplificazione di corrente in decibel:

$$A_i(\text{dB}) = 20 \cdot \log_{10} |A_i|$$

In questo modo se R_i ed R_L sono eguali $A_p(dB)$, $A_v(dB)$ e $A_i(dB)$ coincidono, altrimenti differiscono per una costante.

8.2 Banda passante

Il comportamento dinamico di un amplificatore è descritto nel dominio delle frequenze dalla sua amplificazione, che è sempre funzione della frequenza; se ne può determinare la dipendenza applicando all'ingresso un segnale sinusoidale puro di ampiezza V_i e misurando la corrispondente tensione di uscita V_o per diversi valori di frequenza f e costruendo le due funzioni

$$A(f) = 20 \cdot \log_{10} \left| \frac{V_o(f)}{V_i(f)} \right| \quad \Phi(f) = \arg \left[\frac{V_o(f)}{V_i(f)} \right]$$

Le due quantità V_i e V_o sono complesse. $A(f)$ è il modulo dell'amplificazione espresso in decibel; $\Phi(f)$ è lo sfasamento tra il segnale di uscita e quello di ingresso. Nei *diagrammi di Bode* queste due quantità sono riportate in funzione del logaritmo della frequenza (fig. 8.1).

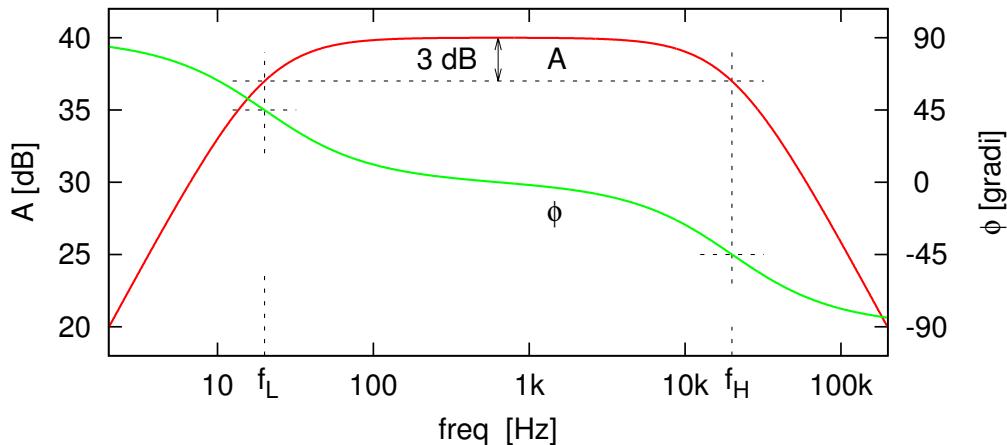


Figura 8.1: Diagrammi di Bode per un amplificatore con guadagno di 40 dB e banda passante 20 – 20000 Hz. La curva rossa rappresenta l'amplificazione $A(f)$, la verde lo sfasamento $\Phi(f)$.

Convenzionalmente si identificano nella funzione $A(f)$ una frequenza di taglio superiore, f_H , ed una frequenza di taglio inferiore, f_L , corrispondenti a quelle frequenze a cui il valore di $A(f)$ si è ridotto, all'aumentare o al diminuire della frequenza, di 3 dB rispetto al suo massimo. Si indica come *banda passante* dell'amplificatore l'intervallo di frequenze tra f_L ed f_H .

Un amplificatore si dice *a larga banda* quando $f_H/f_L \gg 1$, indipendentemente dal valore in Hz della banda passante $B_w = f_H - f_L$.

Negli amplificatori a larga banda la funzione $A(f)$ presenta spesso un andamento piatto su un'ampia gamma di frequenze, per poi decrescere di ampiezza al disopra del limite superiore f_H e al disotto del limite inferiore f_L .

8.3. RISPOSTA ALL'ONDA QUADRA IN UN AMPLIFICATORE A LARGA BANDA 8-3

Mentre la frequenza di taglio f_H è sempre presente in qualsiasi sistema fisico, è possibile realizzare amplificatori in grado di lavorare con segnali fino alla corrente continua, quindi con amplificazione $A(f)$ costante fino al limite $f_L = 0$. Tutti gli amplificatori che arrivano alla corrente continua sono da considerare a larga banda, anche se la loro banda passante è solo di pochi Hz .

Il diagramma di Bode riportato in fig. 8.1 si riferisce ad un amplificatore con guadagno di 40 dB , che copre lo spettro delle frequenze audio da 20 a 20000 Hz . Il comportamento è quello di un amplificatore ideale con banda passante infinita, associato ad una cella RC passa-alto che determina la frequenza f_L ed una cella passa-basso che determina la frequenza f_H (fig. 8.2). Grazie al valore elevato del rapporto tra le due frequenze f_H ed f_L gli angoli di fase e le attenuazioni corrispondenti sono quasi esattamente quelli delle due celle prese singolarmente e cioè +45 e -45 gradi e $1/\sqrt{2}$.

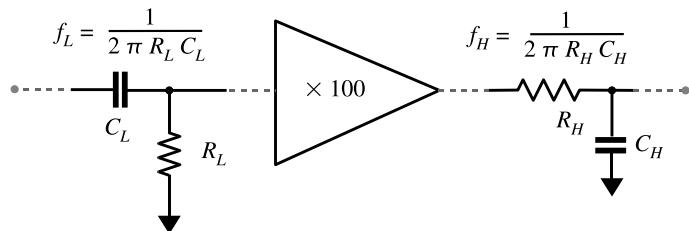


Figura 8.2: Un amplificatore ideale con amplificazione 100 ($\equiv 40 dB$) associato ad una singola cella passa alto ed una passa basso presenta la risposta in frequenza di fig. 8.1.

8.3 Risposta all'onda quadra in un amplificatore a larga banda

Nel dominio del tempo il comportamento dinamico di un amplificatore può essere descritto dalla sua risposta quando il segnale di ingresso è un'onda quadra. In fig. 8.3 è riportata in verde come esempio l'uscita dell'amplificatore di fig. 8.1 e 8.2,

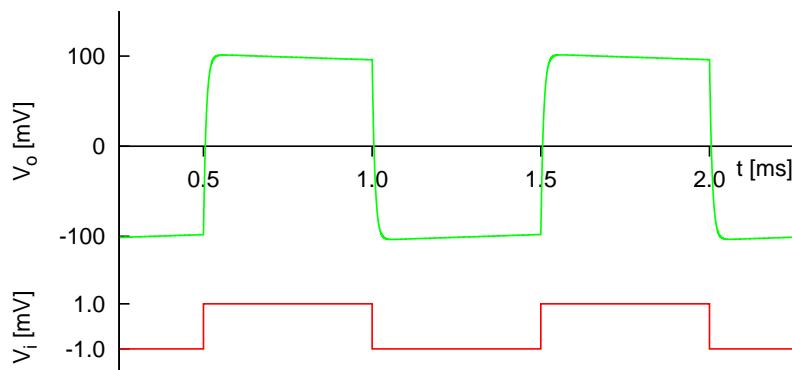


Figura 8.3: Risposta all'onda quadra dell'amplificatore di fig. 8.2. Il segnale di ingresso è in rosso, il segnale di uscita in verde.

quando all'ingresso viene applicato un segnale ad onda quadra di frequenza 1000 Hz e ampiezza 1 mV (linea rossa).

La forma d'onda all'uscita non riproduce in modo perfettamente fedele quella di ingresso: i due fronti di salita e di discesa sono leggermente rallentati rispetto alle due transizioni ideali all'ingresso ed i due livelli stabili tra una transizione e l'altra mostrano una lieve deriva verso lo zero. Il primo effetto è dovuto alla presenza della frequenza di taglio superiore f_H introdotta dalla cella passa-basso $R_H C_H$, il secondo alla frequenza di taglio inferiore f_L introdotta dalla cella passa-alto $R_L C_L$.

tempo di salita

Nella figura fig. 8.4 il fronte di salita è stato messo in evidenza espandendo la scala dei tempi. Si osserva l'andamento tipico della tensione ai capi di un condensatore all'uscita di una cella RC di tipo passa-basso:

$$v(t) = V_2 + (V_1 - V_2) \cdot e^{-t/\tau} \quad (8.2)$$

dove V_1 è il valore iniziale, V_2 è il valore finale e $\tau = RC$ è la costante di tempo della cella, legata alla frequenza di taglio:

$$f_H = \frac{1}{2\pi R_H C_H}$$

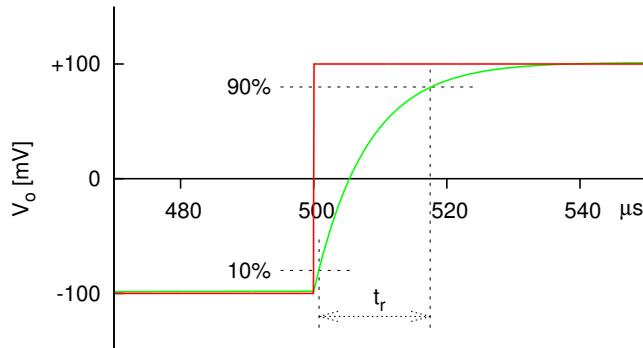


Figura 8.4: Ingrandimento del fronte di salita del segnale all'uscita dell'amplificatore di fig. 8.2 (linea verde), a confronto con il segnale che si avrebbe nel caso ideale di un amplificatore con banda passante infinita (linea rossa).

Convenzionalmente si definisce *tempo di salita* t_r di un segnale (r per *rise time*) il tempo richiesto per passare dal 10% al 90% dell'escursione completa. Dalla eq. 8.2

$$t_r = t_{90\%} - t_{10\%} = \tau \cdot [\log(0.9) - \log(0.1)] = \tau \cdot \log(9) = 2.2 \cdot \tau$$

mentre in un tempo pari alla costante di tempo τ il segnale passa dallo 0% al 63% (cioè da 0 a $1 - 1/e$).

Il tempo di salita t_r è più conveniente della costante di tempo τ per la caratterizzazione di un segnale. Infatti la costante di tempo può essere determinata solo per un fronte di salita (o di discesa) che presenti un andamento esponenziale puro, mentre il tempo di salita definito sopra è una quantità che è sempre ben determinabile e può essere misurata qualunque sia l'andamento della transizione, purchè siano chiaramente visibili i livelli iniziale e finale. Inoltre all'oscilloscopio è più facile misurare il tempo di salita t_r che non la costante di tempo τ .

8.3. RISPOSTA ALL'ONDA QUADRA IN UN AMPLIFICATORE A LARGA BANDA8-5

Nel caso della singola cella passa basso vale la relazione

$$t_r = 2.2 \cdot \tau = 2.2 \cdot \frac{1}{2\pi f_H} = \frac{0.35}{f_H}$$

Questa stessa relazione continua ad essere valida, in forma approssimata, anche quando la frequenza di taglio f_H è determinata dall'effetto combinato di più di una singola cella passa-basso. Ad esempio, un oscilloscopio con una frequenza di taglio superiore di 100 MHz, determinata dalla combinazione delle molteplici frequenze di taglio presenti nel canale di amplificazione, avrà comunque un tempo di salita t_r di circa 3.5 ns come risposta ad un segnale a gradino istantaneo.

Analogamente a quanto detto sopra si definisce un tempo di discesa t_f (f per *fall time*). Nel caso ideale di una rete perfettamente lineare t_f e t_r coincidono; in un amplificatore reale, specialmente nel caso di segnali molto veloci, ci possono essere fenomeni di saturazione (non lineari) che danno origine a tempi di salita e di discesa anche sensibilmente diversi.

tilt

Nella figura 8.5 è riportato in evidenza l'andamento temporale del segnale di uscita durante la semionda positiva dell'onda quadra, mentre il segnale di ingresso rimane stabile ad un valore costante.

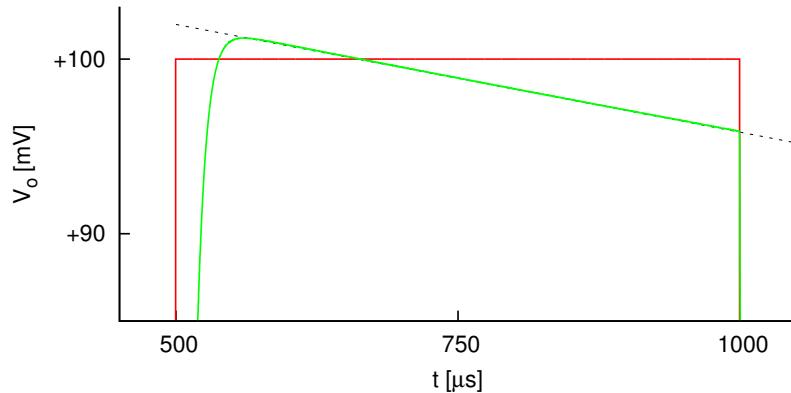


Figura 8.5: Effetto della cella passa-alto nella risposta all'onda quadra dell'amplificatore di fig. 8.2: il segnale all'uscita (in verde) non rimane costante nell'intervallo di tempo tra le due transizioni, come si avrebbe nel caso ideale (in rosso), ma presenta una deriva verso lo zero.

Si vede come il valore di tensione raggiunto al termine del fronte di salita non viene mantenuto nel tempo, ma si stabilisce subito una deriva lenta ma continua verso lo zero (*tilt*). Si tratta dell'inizio del decadimento esponenziale caratteristico della corrente (e della tensione) nella resistenza all'uscita di una cella *RC* di tipo passa-alto:

$$v(t) = V_1 \cdot e^{-t/\tau_L}$$

dove V_1 è la tensione iniziale e $\tau_L = R_L C_L$ è la costante di tempo della cella, legata alla frequenza di taglio:

$$f_L = \frac{1}{2\pi\tau_L} = \frac{1}{2\pi R_L C_L}$$

Quando, come in questo caso, la durata del semiperiodo dell'onda quadra è molto piccola rispetto alla costante di tempo τ_L , la deriva ha una pendenza praticamente costante e pari a

$$\frac{dv(t)}{dt} = -\frac{V_1}{\tau_L} \cdot e^{-t/\tau_L} \Big|_{t=0} = -\frac{V_1}{\tau_L}$$

Data la brevità del tempo di osservazione la costante τ_L non è direttamente determinabile, ma dalla misura della deriva dv/dt si può calcolare

$$\tau_L = -\frac{V_1}{dv/dt}$$

Quando il rapporto f_H/f_L negli amplificatori a larga banda diminuisce, l'esame della risposta all'onda quadra fornisce ancora una indicazione qualitativa della banda passante, ma non permette di determinare facilmente valori numerici accurati delle costanti di tempo e delle frequenze di taglio. Questo è visibile nella fig. 8.6 dove sono

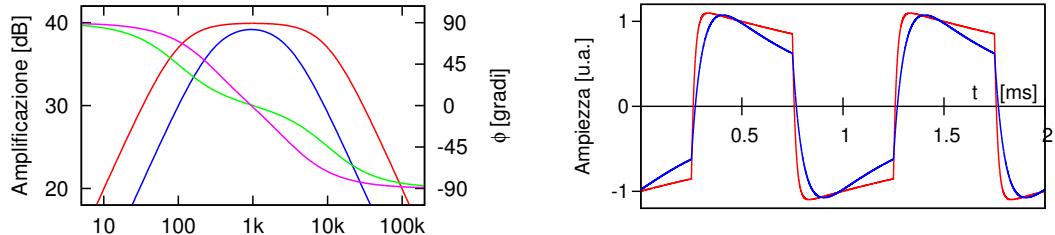


Figura 8.6: Diagrammi di Bode e risposta all'onda quadra per due amplificatori con $f_H/f_L = 10 \text{ kHz}/100 \text{ Hz}$ e $3 \text{ kHz}/300 \text{ Hz}$. Il modulo del guadagno e le risposte in funzione del tempo sono riportate in rosso ed in blu, le fasi in verde e violetto, rispettivamente per il primo ed il secondo amplificatore.

riportati i diagrammi di Bode di due amplificatori con f_H/f_L pari a $10 \text{ kHz}/100 \text{ Hz}$ e $3 \text{ kHz}/300 \text{ Hz}$. Al ridursi del rapporto tra f_H ed f_L i vari effetti si sovrappongono e si confondono l'un l'altro.

8.4 Frequenze di taglio multiple

Una rete complessa come un amplificatore può contenere più di uno stadio con comportamento di tipo passa-basso, ciascuno con la sua propria costante di tempo t_H e frequenza di taglio f_H . Nel caso che tra queste frequenze ve ne sia una *dominante*, perché sensibilmente più bassa delle altre, la frequenza di taglio dell'intero amplificatore sarà determinata da questa.

La curva blu in fig. 8.7 corrisponde a questo caso e si riferisce ad un amplificatore con una frequenza di taglio dominante $f_{H1} = 20 \text{ kHz}$, seguita da una seconda frequenza

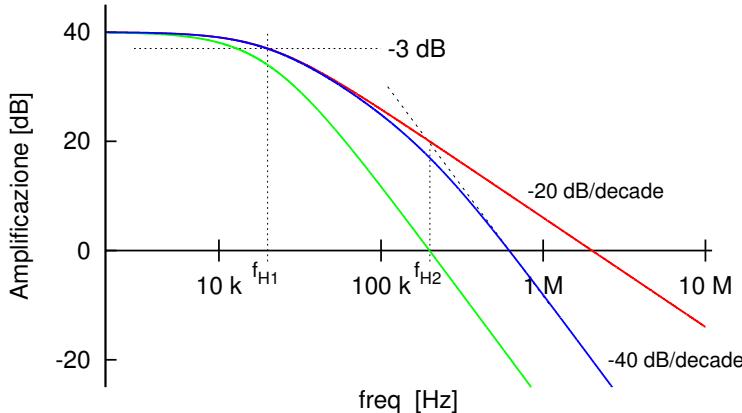


Figura 8.7: Banda passante di un amplificatore con due frequenze di taglio f_H . La curva verde corrisponde al caso di due frequenze identiche $f_{H1} = f_{H2} = 20 \text{ kHz}$; la curva blu corrisponde al caso in cui una delle due frequenze è dominante, perché sensibilmente minore dell'altra: $f_{H1} = 20 \text{ kHz}, f_{H2} = 200 \text{ kHz}$.

$f_{H2} = 200 \text{ kHz}$. Per confronto è riportata in rosso la risposta dell'amplificatore di fig. 8.1, che presenta la sola frequenza di taglio $f_{H1} = 20 \text{ kHz}$. Si può osservare nel grafico che la frequenza di taglio a -3 dB dell'amplificatore nel suo complesso è alterata molto poco dalla presenza della f_{H2} a 200 kHz ; infatti si passa da 20 kHz a 19.8 kHz . Solo quando si arriva in prossimità della seconda frequenza di taglio se ne può osservare l'effetto nella curva di risposta $A(f)$. In particolare, si può vedere che si passa dalla pendenza originaria di 20 dB per decade di frequenza alla pendenza doppia di 40 dB per decade.

La curva verde in fig. 8.7 si riferisce al caso opposto: un amplificatore con due frequenze di taglio coincidenti $f_{H1} = f_{H2} = 20 \text{ kHz}$. In questo caso la frequenza di taglio complessiva risulta sensibilmente più bassa (12.9 kHz) e la funzione $A(f)$ prende immediatamente la pendenza di 40 dB per decade.

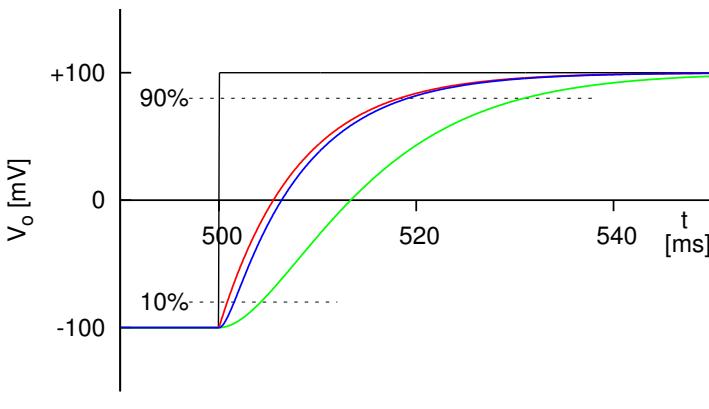


Figura 8.8: Risposta all'onda quadra per un amplificatore con frequenze di taglio come in fig. 8.7. In rosso: $f_H = 20 \text{ kHz}, t_H = 7.96 \mu\text{s}, t_r = 17.5 \mu\text{s}$; in blu: $f_{H1} = 20 \text{ kHz}, f_{H2} = 200 \text{ kHz}, t_{H2} = 0.796 \mu\text{s}, t_r = 17.6 \mu\text{s}$; in verde: $f_{H1} = 20 \text{ kHz}, f_{H2} = 20 \text{ kHz}, t_r = 27 \mu\text{s} (\approx 0.35/12.9 \text{ kHz})$.

In fig. 8.8 sono riportati i corrispondenti fronti di salita ottenuti dalla risposta all'onda quadra. In nero è riportato l'andamento nel caso ideale di banda passante infinita. I colori delle altre curve corrispondono ai casi della figura precedente. Anche qui si vede che in presenza di una frequenza di taglio dominante (curva blu) l'effetto della seconda frequenza è molto piccolo rispetto al caso della singola frequenza di taglio (curva rossa); infatti si passa da $17.5 \mu\text{s}$ a $17.6 \mu\text{s}$. Nel caso invece che le due

frequenze di taglio coincidano (curva verde) l'effetto sul tempo di salita è notevole: $t_r = 26.7 \mu s$. In quest'ultimo caso è notevole anche l'effetto sulla parte iniziale del fronte di salita: la transizione inizia con pendenza nulla e questo rende difficile individuare in maniera univoca l'istante iniziale. La misura dei due tempi ai livelli del 10% e del 90% risulta invece egualmente agevole in tutti e tre i casi.

Con una buona approssimazione per gli usi pratici si può assumere che i tempi di salita si sommino quadraticamente.

8.5 Amplificatore con transistor bjt

Il circuito in fig. 8.9 è un esempio di amplificatore a più stadi. Nel circuito di ingresso è stata utilizzata una coppia differenziale (T_1, T_2), con un circuito *current mirror* (T_4, T_5) che fornisce la corrente agli emettitori; lo stadio di uscita è un *emitter follower* di tipo *Darlington* con una coppia a simmetria complementare come secondo transistor ($T_3, T_6 - T_7$); l'accoppiamento tra T_3 e T_6 è fatto tramite un diodo zener per avere una tensione di uscita nominale a riposo di $0 V$ e rendere quindi il circuito utilizzabile in corrente continua.

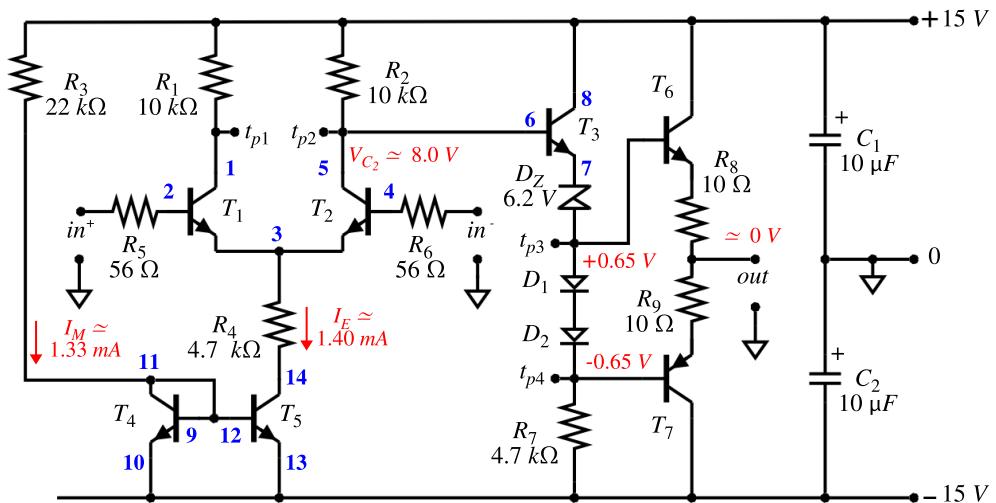


Figura 8.9: Circuito amplificatore a più stadi con transistor bjt.

Sia l'amplificatore differenziale di ingresso sia il circuito *current mirror* possono dare risultati soddisfacenti solo se vengono realizzati con transistor monolitici. È stato quindi utilizzato un *chip* che contiene cinque transistor di tipo *NPN* realizzati sullo stesso substrato di Silicio (fig. 8.10). I due transistor dello stadio di uscita a simmetria complementare (T_6 e T_7) sono invece dispositivi *discreti NPN* (2N2219) e *PNP* (2N2905). Le caratteristiche principali di questi componenti sono riassunte nella tab. 8.1.

La corrente I_M nel *current mirror* è determinata dalla tensione di alimentazione e dal valore di R_3 : $I_M = (30 V - 0.65 V) / 22 k\Omega = 1.33 mA$. La corrente negli emettitori della coppia differenziale avrebbe questo stesso valore (a meno delle correnti delle basi) nel caso di un circuito ideale. Risulta invece leggermente maggiore a causa della

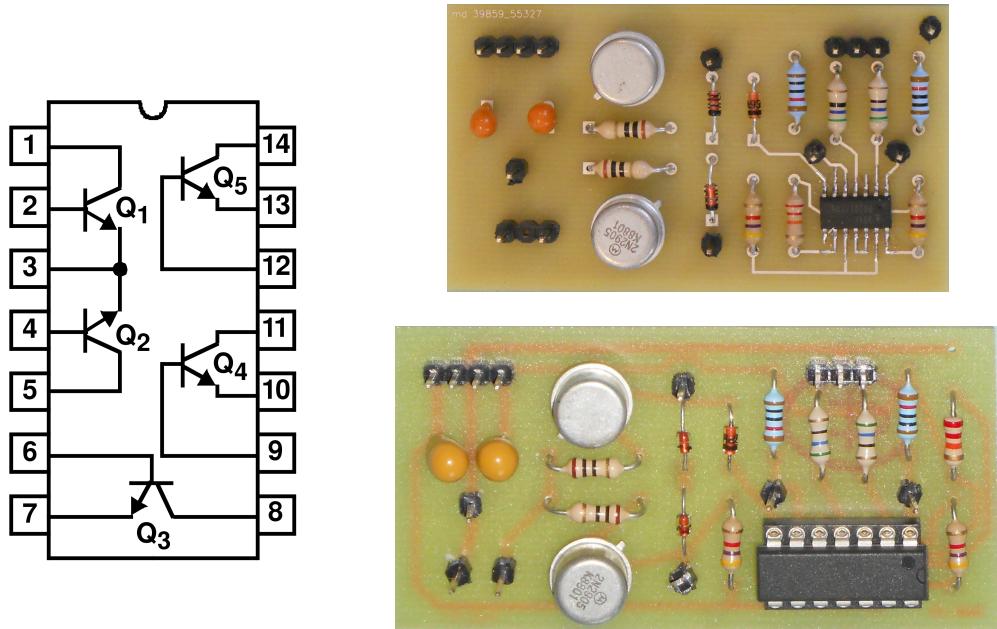


Figura 8.10: A sinistra: chip CA3046 contenente 5 transistor NPN monolitici. A destra: in basso, montaggio su circuito stampato dell'amplificatore di fig. 8.9 con chip di tipo DIP (dual in line package); in alto, con chip di tipo SMD (surface mount device).

conduttanza di uscita non nulla del transistor T_5 : $I_E = I_M + g_{oe} \cdot (V_{CE_5} - V_{CE_4})$. Con $g_{oe} = 13 \mu A/V$ e $V_{CE_5} - V_{CE_4} \approx 8 V$, si ha $I_{C5} \approx 1.41 mA$; in ciascuno dei due transistor T_1, T_2 si ha quindi $I_C \approx I_E = 0.70 mA$ e $V_C = 15 V - I_C \cdot R_{1,2} \approx 8.0 V$.

Queste condizioni, riassunte nella tabella 8.1, costituiscono un buon punto di lavoro, perché la tensione di collettore di T_1 (e di T_2) si trova molto vicino al valore intermedio tra i $+15 V$ che si hanno quando il transistor arriva a spingersi e gli $0 V$ al confine con la regione di saturazione (supponendo che la componente di modo comune del segnale sulle basi sia nulla).

	T_1, T_2	T_4	T_5	T_3	T_6, T_7	
I_E	0.70	1.32	1.41	3.1	7	mA
h_{FE}	113	108	116	121	200 \div 250	
g_{fe}	26	49	53	112		mA/V
g_{ie}	0.23		0.46	0.92		mA/V
g_{oe}	6.5		13	27		$\mu A/V$

Tabella 8.1: Principali caratteristiche dei transistor utilizzati nel circuito di fig. 8.9.

Ovviamente nella progettazione reale di un circuito è più facile che il calcolo precedente debba essere percorso in senso inverso, partendo cioè dalle esigenze del punto di lavoro e determinando (magari per approssimazioni successive) i valori dei componenti necessari per soddisfarle.

Il guadagno di tensione di modo differenziale del primo stadio sarà (eq. 7.59):

$$A_v = -g_{fe}R_L \cdot \frac{1}{1 + g_{oe}R_L} \quad (8.3)$$

$$= -26 \text{ mA/V} \cdot 10 \text{ k}\Omega \cdot \frac{1}{1 + 6.5 \mu\text{A/V} \cdot 10 \text{ k}\Omega} \simeq 240 \quad (8.4)$$

pari a 48 dB , e con buona approssimazione questo sarà anche il guadagno del circuito completo, in quanto lo stadio di uscita ha guadagno molto vicino a uno.

Il guadagno di modo comune sarà $\simeq -0.065$ (eq. 7.60 con $R_C = 10 \text{ k}\Omega$ e $R_E = 1/g_{oe4} = 77 \text{ k}\Omega$); la reiezione di modo comune è $CMRR = A_d/A_c = 3700$, pari a 71 dB .

In queste condizioni, l'effetto della deriva con la temperatura della tensione V_{BE} della coppia T_1-T_2 ($-2.3 \text{ mV/}^\circ\text{K}$) si ridurrebbe, in condizioni ideali, a $-0.62 \mu\text{V/}^\circ\text{K}$ differenziali *riferiti all'ingresso*. Nella pratica è abbastanza comune ottenere risultati di almeno un ordine di grandezza peggiori. La frase *riferiti all'ingresso* indica quale è l'ampiezza del segnale utile equivalente che applicato all'ingresso produce all'uscita lo stesso effetto della fonte di errore considerata, cioè con un errore del 100%.

Se si prendesse il segnale di uscita direttamente dal collettore di T_2 si avrebbero numerosi inconvenienti:

- con segnale di ingresso 0 V si avrebbe una tensione di uscita continua di 8 V , invece di zero;
- la resistenza di uscita del circuito sarebbe di circa $9 \text{ k}\Omega$ (dal parallelo di R_1 oppure R_2 e della corrispondente g_{oe}); il comportamento sarebbe quindi ben diverso da quello di un generatore di tensione, come ci si aspetta di trovare all'uscita di un amplificatore di tensione;
- la disponibilità di corrente all'uscita sarebbe assai limitata: massimo 0.7 mA di segnale con l'uscita chiusa su resistenza molto piccola verso massa.

Lo stadio di uscita costituito dai transistor T_5 , T_6 , T_7 e dai diodi D_Z , D_1 e D_2 ha lo scopo di eliminare questi problemi.

Il diodo zener D_Z viene scelto in maniera da avere $V_Z + V_{BE3} + V_{BE6} - R_8 I_{E6} = V_{C2}$ (la corrente I_{E6} , per le convenzioni adottate, è negativa); in questo modo la tensione di uscita nominale risulta 0 V in assenza di segnale. Idealmente $V_Z = 8.0 \text{ V} - 0.65 \text{ V} - 0.65 \text{ V} = 6.7 \text{ V}$ (la tensione ai capi di R_8 è di pochi mV e quindi è trascurabile). I diodi zener disponibili in commercio hanno, intorno a questo valore, tensioni intervallate di circa 0.6 V tra un modello e l'altro. Quindi nel peggiore dei casi la tensione di uscita sarà all'interno dell'intervallo $V_{out} = \pm 300 \text{ mV}$. Questo equivale ad una *tensione di offset* riferita all'ingresso di $\pm 300 \text{ mV}/A_v = \pm 1.4 \text{ mV}$.

La resistenza di uscita di $9 \text{ k}\Omega$ della coppia differenziale *vista* attraverso l'mitter follower Darlington risulta $h_{FE3} \cdot h_{FE6,7}$ volte più piccola e quindi dell'ordine di grandezza degli ohm (eq. 7.41). Molto più difficile è calcolare il contributo dovuto ai transistor T_6 e T_7 del termine $h_{ie}/(h_{fe} + 1)$ nella eq. 7.41. Infatti, il circuito di polarizzazione dei transistor T_6 e T_7 è piuttosto semplice e non permette di stabilire con grande precisione il punto di lavoro. La corrente nei due transistor a riposo

può variare dalla frazione di mA fino a qualche mA e di conseguenza il termine trascurato nell'impedenza di uscita può variare da un minimo di una frazione di ohm ad un massimo di qualche ohm. Si tratta comunque di valori sufficientemente piccoli da permetterci di considerare il circuito di uscita come equivalente ad un discreto generatore di tensione, per cui il contributo principale all'impedenza di uscita in bassa frequenza è dato dalle due resistenze R_8 ed R_9 .

Infine, la resistenza R_4 ha funzione di protezione della coppia differenziale da eventuali cortocircuiti accidentali e le resistenze R_5 ed R_6 hanno funzione di protezione da eventuali autooscillazioni del circuito a causa delle capacità parassite.

Le caratteristiche principali di funzionamento del circuito (amplificazione ed impedenze di ingresso e di uscita), calcolate con il programma *Spice* in funzione della frequenza, sono riportate in fig. 8.11 e 8.12. Il calcolo ha tenuto conto delle capacità di transizione e diffusione presenti nelle giunzioni di diodi e transistor (vedi cap. 16), ma non delle capacità parassite tra i componenti del circuito e degli effetti termici.

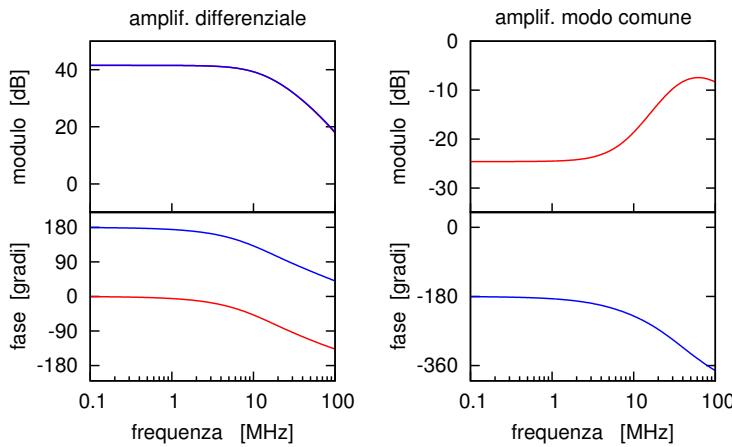


Figura 8.11: Sinistra: amplificazione differenziale con segnale all'ingresso invertente (curva blu) e non invertente (curva rossa). Il valore (41.5 dB) è 6 dB più basso di quello che si avrebbe applicando un vero segnale differenziale ad entrambi gli ingressi. Destra: amplificazione di modo comune (-25 dB in bassa frequenza).

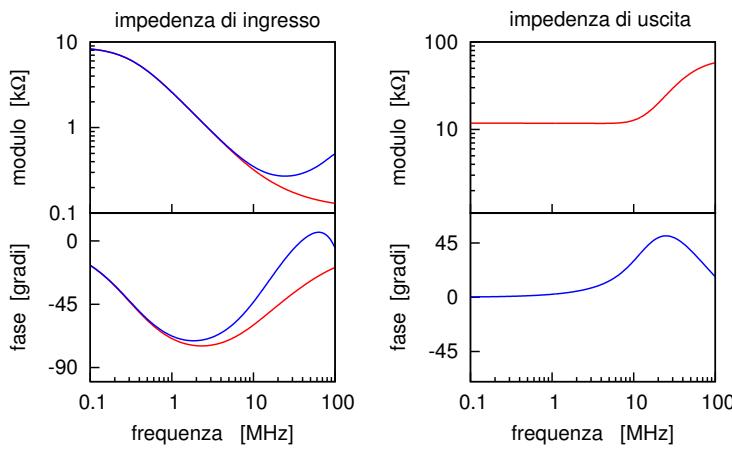


Figura 8.12: Sinistra: impedenza all'ingresso invertente (curva blu) e non invertente (curva rossa). Il valore in bassa frequenza è di $8.7 \text{ k}\Omega$ e si riduce di un fattore $\sqrt{2}$ a 320 kHz . L'impedenza di uscita (a destra) è 11.8Ω fino a circa 10 MHz ; di questi, 10Ω sono dovuti alle resistenze R_8 ed R_9 .

8.6 Misure sui circuiti

Il collegamento della strumentazione di misura ai punti interni di un circuito per esaminarne il funzionamento può portare a delle alterazioni anche molto pesanti delle condizioni operative. Nella fig. 8.13 un oscilloscopio è stato collegato all'uscita del

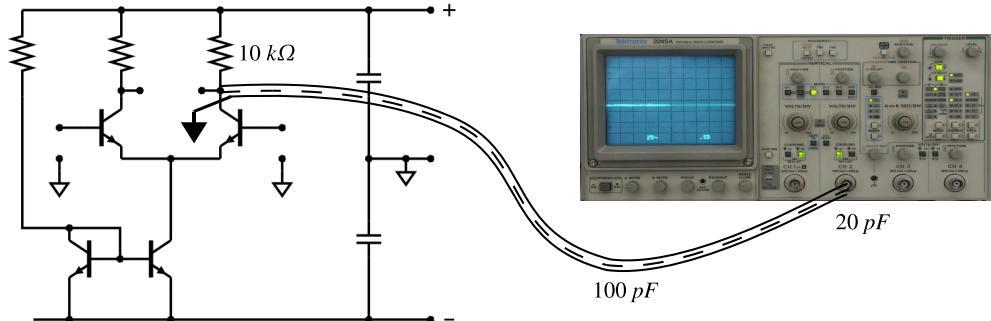


Figura 8.13: Il collegamento di un oscilloscopio ad un punto ad alta impedenza di un circuito tramite un cavetto coassiale introduce una costante di tempo ed una frequenza di taglio che possono limitare pesantemente la banda passante del circuito.

primo stadio dell'amplificatore descritto nel paragrafo precedente. Per il collegamento si è supposto di utilizzare un cavetto coassiale, come si fa di solito, per schermare il segnale da analizzare da eventuali disturbi e interferenze esterne. Un tale cavo presenta tipicamente una capacità dell'ordine di 100 pF/m . Con una lunghezza di 1 m ed una capacità all'ingresso dell'oscilloscopio di 20 pF , ci si trova ad inserire nel circuito una capacità di 120 pF in parallelo alla resistenza di uscita dello stadio, di circa $9 \text{ k}\Omega$. Anche trascurando le capacità già presenti nel circuito ed associate al comportamento in alta frequenza del transistor, come si vedrà nel cap. 16, la cella passa-basso che si forma può condizionare la banda passante dell'amplificatore. In questo caso si ha una costante di tempo di $1.1 \mu\text{s}$, corrispondente ad una frequenza di taglio di 150 kHz , che risulta essere la frequenza di taglio dominante.

8.6.1 Sonda compensata

Questo problema viene in parte mitigato dall'uso di una *sonda compensata* per prelevare il segnale, cioè di una sonda che contiene un partitore, tipicamente con un fattore di attenuazione $1 : 10$, come descritto nella fig. 8.14.

La perdita di sensibilità che si ha usando la sonda compensata porta come vantaggio la riduzione di un fattore analogo della capacità esterna che viene ad interferire con il circuito. Nell'esempio citato prima, da 120 pF si passa a circa 12 pF . La costante di tempo del circuito di uscita scende da $1.1 \mu\text{s}$ a circa $0.11 \mu\text{s}$ e la frequenza di taglio passa da 150 kHz a circa 1.5 MHz . L'interferenza dello strumento di misura con il funzionamento del circuito rimane, ma è sensibilmente ridotta.

8.6.2 Partitore compensato

Il partitore utilizzato nella sonda non è un normale partitore resistivo, ma è un partitore tra due impedenze ciascuna delle quali è composta da una resistenza

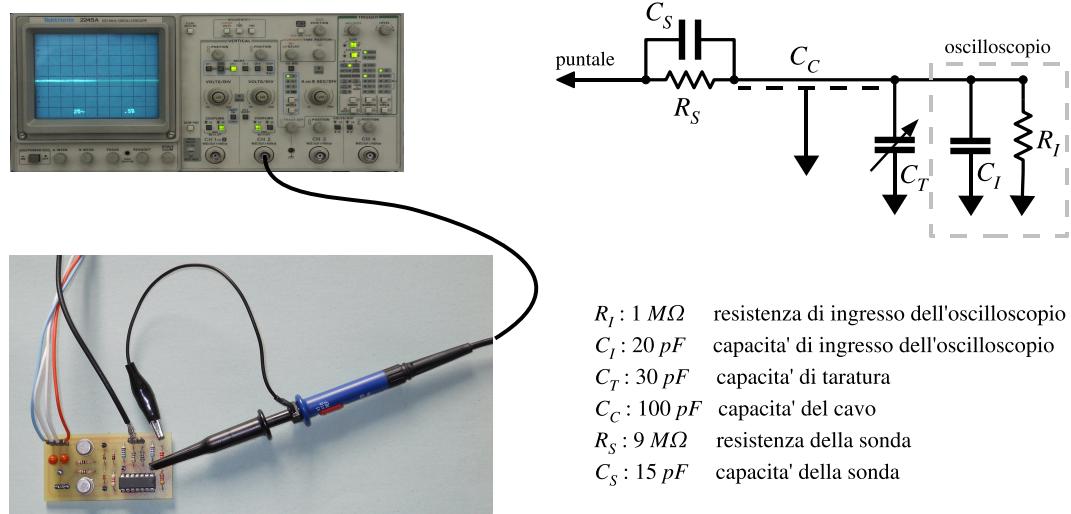


Figura 8.14: La *sonda compensata* riduce l'effetto dello strumento di misura sul circuito a spese di una riduzione dell'ampiezza del segnale.

ed una capacità in parallelo e prende il nome di *partitore compensato*. Questo è indispensabile se si vuole mantenere un coefficiente di attenuazione costante con la frequenza e quindi portare all'oscilloscopio una copia ridotta, ma fedele, del segnale. Infatti il ramo inferiore del partitore contiene inevitabilmente la capacità del cavo in

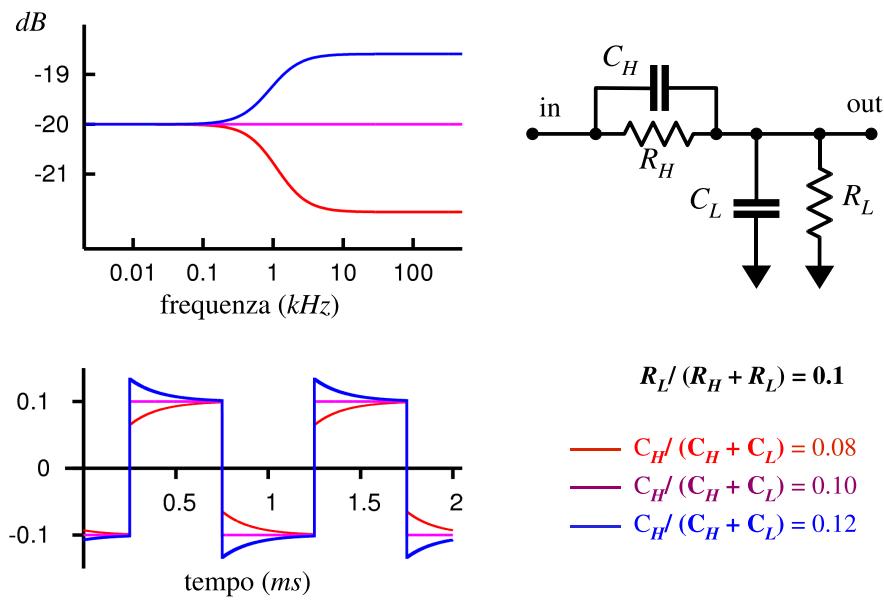


Figura 8.15: Nel partitore compensato le due impedenze $Z_H = R_H//C_H$ e $Z_L = R_L//C_L$ hanno lo stesso rapporto tra parte reale ed immaginaria, in modo da avere un fattore di attenuazione indipendente dalla frequenza.

parallelo alla resistenza ed alla capacità di ingresso dell'oscilloscopio. Solo utilizzando per il ramo superiore del partitore una impedenza Z_H proporzionale a quella del ramo inferiore (Z_L), quindi con lo stesso rapporto tra parte immaginaria e parte reale, si può ottenere che il rapporto $Z_L/(Z_L + Z_H)$ sia indipendente dalla frequenza.

Nella fig. 8.15 sono riportate la banda passante e la risposta all'onda quadra per i tre casi di un partitore esattamente compensato (curve in violetto), di un partitore sovracompensato (curve in blu) e sottocompensato (curve in rosso). Nel primo caso il rapporto di partizione $C_H/(C_H + C_L)$ tra i condensatori è esattamente eguale a quello tra le resistenze $R_L/(R_L + R_H)$; nel secondo caso, partitore sovracompensato, la risposta in alta frequenza, dominata dal partitore capacitivo, è maggiore di quella in bassa frequenza e la risposta all'onda quadra presenta un *overshoot*; nel terzo caso, partitore sottocompensato, all'opposto si ha una risposta in alta frequenza minore che in bassa ed un *undershoot* sui fronti dell'onda quadra.

Poichè non è possibile costruire sonde che siano già esattamente compensate, anche perchè gli oscilloscopi non hanno tutti esattamente la stessa capacità di ingresso, si aggiunge nella sonda una piccola capacità di regolazione (C_T in fig. 8.14), da regolare di volta in volta per ottenere l'esatta compensazione.

La costante di tempo dei due rilassamenti esponenziali in fig. 8.15 è pari a $(C_H//C_L) \cdot (R_H//R_L) \simeq 100\mu s$, quindi ben visibile con un'onda quadra a 1 kHz, come riportato nella figura. Un modo semplice per calcolare questa costante di tempo è descritto nella fig. 8.16.

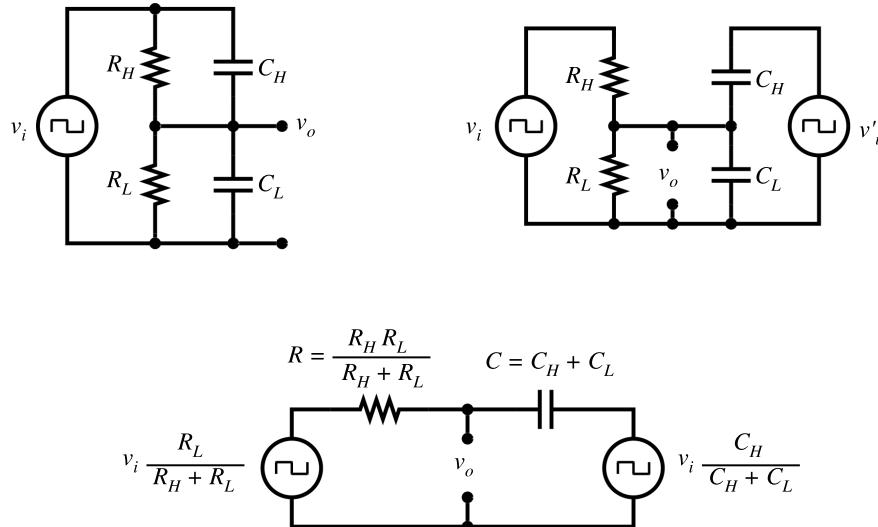


Figura 8.16: I due circuiti in alto sono quivalenti, con il generatore di tensione v_i sostituito da due generatori identici v_i e v'_i . Il circuito è stato ridisegnato in basso, sostituendo a ciascuno dei due partitori, R_H, R_L e C_H, C_L , il loro equivalente di Thévenin. È immediato verificare che al momento che la tensione di ingresso compie la sua transizione da 0 al valore V la tensione di uscita v_o passa da 0 a $V \cdot C_H / (C_H + C_L)$, per tendere poi asintoticamente a $V \cdot R_L / (R_H + R_L)$ con rilassamento esponenziale e costante di tempo $\tau = (C_H + C_L) \cdot R_H R_L / (R_H + R_L)$. Nel calcolo si è supposto un fronte di salita isolato, oppure un'onda quadra con periodo sufficientemente lungo da consentire un rilassamento completo dei transitori tra un fronte ed il successivo.

Per effettuare la taratura delle sonde gli oscilloscopi generalmente sono dotati di un generatore di onda quadra a 1 kHz, di ampiezza 0.5 V oppure 1 V, collegato ad una linguetta che compare sul pannello frontale e dove è possibile collegare il gancetto a molla di cui è dotata la punta della sonda. Quindi si può provvedere alla regolazione con un piccolo cacciavite, possibilmente di plastica, ruotando la vite che compare sul corpo della sonda oppure sul connettore che si collega all'oscilloscopio. Quando l'onda quadra appare sullo schermo priva di *overshoot* o *undershoot* la regolazione è completa.

8.7 Misura dell'amplificazione

Utilizzando generatore di segnali ed oscilloscopio, l'amplificazione può essere misurata come descritto nella fig. 8.17.

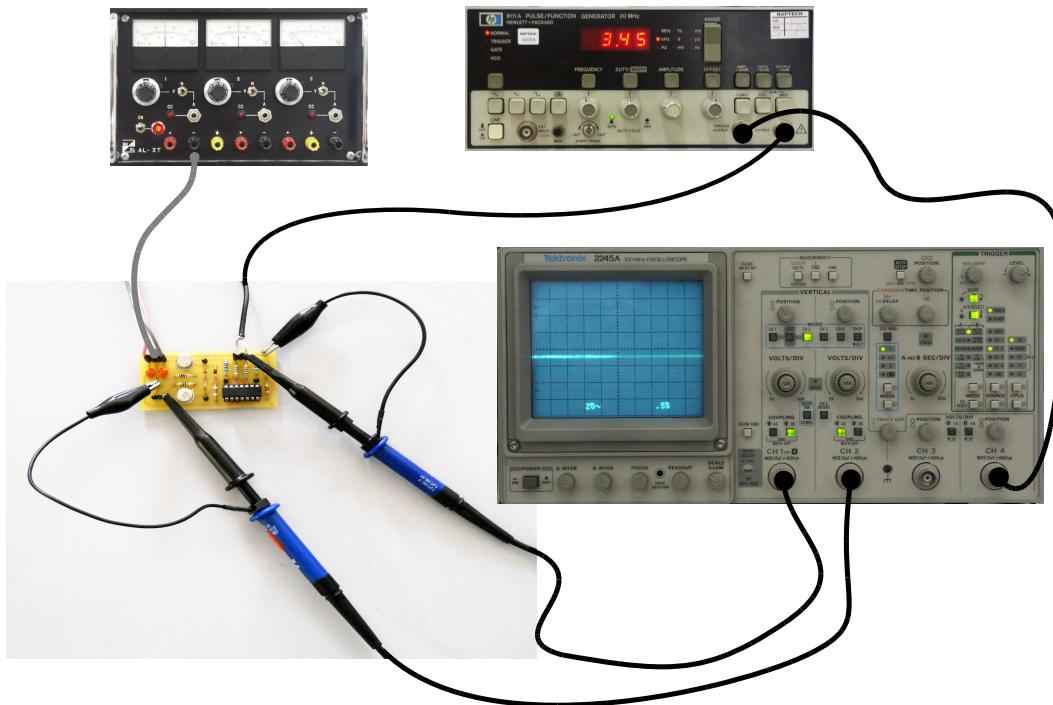


Figura 8.17: .

Si avrà cura di prelevare i segnali all'ingresso ed all'uscita dell'amplificatore utilizzando due sonde eguali. Questo riduce gli errori di misura. L'amplificazione è un rapporto tra le ampiezze di due segnali; utilizzare due canali di misura eguali fa scomparire dal rapporto gli errori di scala comuni ai due canali. Al contrario, la cosa peggiore che si può fare dal punto di vista degli errori di misura è utilizzare il valore della tensione di ingresso fornito dal generatore e quello della tensione di uscita misurato dall'oscilloscopio.

Utilizzare per il sincronismo dell'oscilloscopio il segnale appositamente fornito dal generatore non è indispensabile in questa misura, ma offre il vantaggio di avere sullo schermo segnali stabili indipendentemente dalle variazioni nelle ampiezze.

La precisione che si può raggiungere con questa tecnica nella misura dell'amplificazione non è molto elevata, tipicamente dell'ordine del 5% ($\pm 0.4 \text{ dB}$). Si ha però l'enorme vantaggio di avere tramite l'oscilloscopio una visione panoramica del funzionamento del circuito, in modo da poter individuare eventuali difetti come la comparsa di distorsioni o di oscillazioni parassite che in una misura effettuata tramite un voltmetro, per quanto raffinato, in genere passano inosservate.

8.8 Impedenza di ingresso e di uscita

Senza dover disporre di ulteriore strumentazione, le impedenze di ingresso o di uscita di un amplificatore possono essere determinate aggiungendo in serie all'ingresso o in parallelo all'uscita una resistenza di valore opportuno, mantenendo tutte le altre condizioni costanti e misurando le variazioni indotte nel segnale di uscita.

8.8.1 Impedenza di uscita

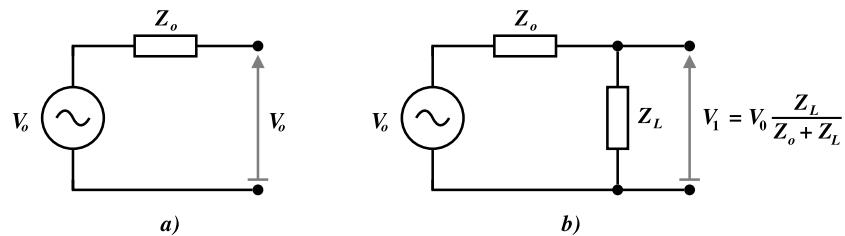


Figura 8.18: La doppia misura della tensione di uscita, *a vuoto* e con un carico Z_L , consente di calcolare il valore dell'impedenza di uscita $Z_o = Z_L \cdot (V_0 - V_1)/V_1$.

Come illustrato in fig. 8.18a, a circuito aperto (*a vuoto*) l'amplificatore presenta ai terminali la tensione V_0 ; quando invece è chiuso sulla impedenza di carico Z_L (fig. 8.18b) presenta la tensione $V_1 = V_0 \cdot Z_L / (Z_o + Z_L)$. Invertendo questa equazione si ottiene $Z_o = Z_L \cdot (V_0/V_1 - 1)$.

Le condizioni migliori per la sensibilità e la precisione della misura si hanno quando $Z_L \simeq Z_o$, e quindi $V_1 = V_0/2$, ma non sempre questa condizione può essere rispettata. Nel caso del circuito in fig. 8.9 ci si aspetta dal calcolo $Z_o \simeq 12 \Omega$ in corrente continua ed in bassa frequenza. Utilizzando per Z_L una resistenza di questo valore nei transistor finali T_6 e T_7 si avrebbe una corrente abbastanza elevata da provocarne la distruzione per surriscaldamento. In circuiti più raffinati in cui questo non avviene, perché dotati di protezioni che limitano la corrente di uscita entro valori di sicurezza, si avrebbe una perdita di linearità che renderebbe priva di senso la misura della impedenza di uscita. Anche questa è una condizione di cui ci si rende immediatamente conto se si osserva il segnale con un oscilloscopio, mentre un voltmetro continuerebbe a dare accurati valori numerici purtroppo privi di senso.

Per il circuito di fig. 8.9 un compromesso ragionevole è utilizzare per Z_L una resistenza da 50Ω .

In questa misura è indispensabile utilizzare il segnale di sincronismo tra generatore ed oscilloscopio perché V_0 e V_1 sono quantità complesse di cui va misurato il modulo e la fase (o meglio, la differenza di fase).

8.8.2 Impedenza di ingresso

Inserendo in serie all'ingresso dell'amplificatore una impedenza Z_S si viene a formare un partitore, come indicato in fig. 8.19, che riduce l'ampiezza del segnale di ingresso del fattore $(Z_i + R_G)/(Z_S + Z_i + R_G)$.

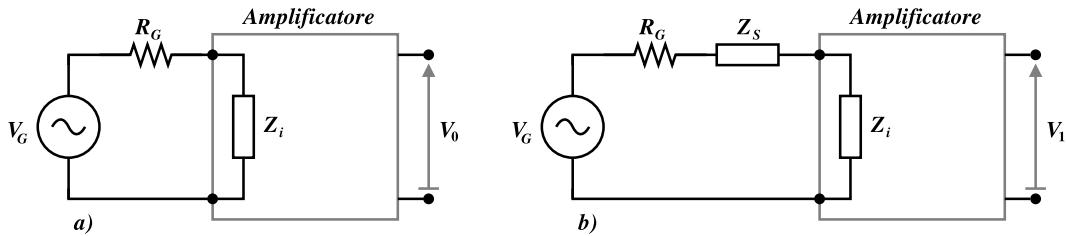


Figura 8.19: La doppia misura della tensione di uscita senza e con una impedenza Z_S in serie all'ingresso consente di calcolare il valore dell'impedenza di ingresso $Z_i = Z_S \cdot (V_0 - V_1)/V_1$ (trascurando R_G).

La tensione di uscita, che è inizialmente V_0 , diventa quindi a parità di tutte le altre condizioni $V_1 = V_0 \cdot (Z_i + R_G)/(Z_S + Z_i + R_G)$. Nel caso che R_G sia trascurabile rispetto a Z_i e Z_S , l'espressione si semplifica in $V_1 = V_0 \cdot Z_i/(Z_S + Z_i)$. Invertendo l'equazione come nel paragrafo precedente, si ottiene $Z_i = Z_S/(V_0/V_1 - 1)$.

Anche in questo caso è conveniente utilizzare per Z_S un valore simile a quello previsto per Z_i ; per il circuito di fig. 8.9, ad esempio, $Z_S = 10 \text{ k}\Omega$.

Anche in questo caso le tensioni V_0 e V_1 devono essere misurate in modulo e fase e quindi sarà necessario mantenere il segnale di sincronismo tra generatore ed oscilloscopio.

Si noti che nella misura di entrambe le impedenze, di ingresso e di uscita, l'unico segnale che si osserva con l'oscilloscopio è quello di uscita, anzi la sua variazione in conseguenza dell'inserimento dell'impedenza Z_L o Z_S . Questo è sufficiente se la misura viene eseguita mantenendo costanti tutte le altre condizioni, in particolare l'ampiezza del segnale fornito dal generatore.

8.9 Distorsione

Ogni amplificatore, anche il più lineare, presenta un certo grado di distorsione. Nel dominio del tempo la distorsione si manifesta come una deformazione del segnale da quella che è la sua forma originale. Per esempio, la fig. 8.20 mostra quello che si ottiene all'uscita di un amplificatore applicando all'ingresso un segnale sinusoidale puro, ma di ampiezza eccessiva, tale da spingere l'amplificatore alla *saturazione*: si ha il fenomeno del *clipping*, cioè della *tosatura* dei picchi della sinusoide quando la tensione di uscita supera i valori che l'amplificatore è in grado di fornire.

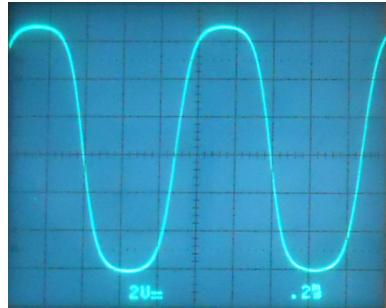


Figura 8.20: Segnale all'uscita di un amplificatore con un segnale di ingresso sinusoidale di ampiezza tale da portare l'amplificatore alla saturazione. Si vede chiaramente il fenomeno del *clipping*, la tosatuta dei picchi del segnale quando la tensione supera i valori che l'amplificatore è in grado di fornire.

Se la relazione tra tensione di ingresso e tensione di uscita di un amplificatore può essere espressa mediante una serie di Taylor (cosa normalmente possibile, a meno di valori di distorsione patologici), si ha:

$$v_o(t) = A_0 + A_1 \cdot v_i(t) + A_2 \cdot v_i^2(t) + A_3 \cdot v_i^3(t) \dots \quad (8.5)$$

I primi due termini, A_0 e A_1 , descrivono il comportamento lineare e rappresentano la tensione di *offset* e l'amplificazione. I termini successivi descrivono le non linearità. Applicando all'ingresso un segnale sinusoidale puro di frequenza f , dalla equazione si vede che all'uscita compariranno anche le frequenze armoniche $2f$, $3f$, $4f$ ecc., corrispondenti ai termini \sin^2 , \cos^2 , \sin^3 , \cos^3 ,

Nel caso che il segnale all'ingresso contenga più componenti a frequenze diverse, i termini non lineari daranno origine, oltre alle armoniche di ciascuna delle componenti, anche alle frequenze somma e differenza di ciascuna coppia di componenti, generate dai prodotti $\sin(2\pi f_i) \cdot \sin(2\pi f_j)$, che si originano nello sviluppo delle potenze del polinomio.

In genere, quando si riesce ad apprezzare ad occhio sull'oscilloscopio la distorsione di una sinusoide, allora si è già a livelli di contenuto armonico piuttosto elevati, dell'ordine del 5 – 10% dell'ampiezza del segnale fondamentale.

Misure accurate della distorsione di un amplificatore, anche a livelli molto più bassi di questi, possono essere fatte utilizzando un *Analizzatore di spettro* oppure un oscilloscopio digitale, che permette di campionare il segnale in funzione del tempo e poi di determinarne il contenuto armonico mediante una operazione *FFT* (*Fast Fourier Transform*, Trasformata di Fourier veloce) (fig. 8.21).

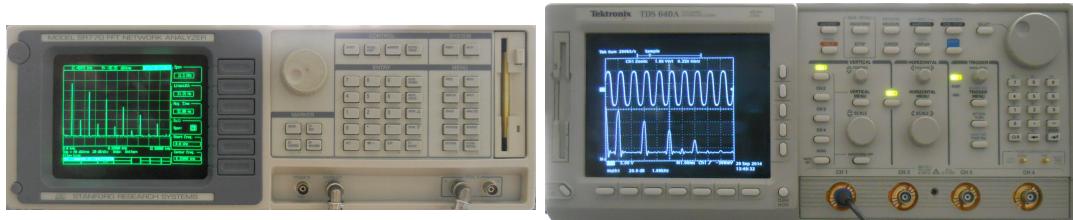


Figura 8.21: Analizzatore di spettro per bassa frequenza, da 0 a 100 kHz (a sinistra) ed oscilloscopio digitale con funzione FFT per frequenze fino a 500 MHz (a destra).

Il grafico in fig. 8.22 riporta il contenuto armonico all'uscita dell'amplificatore di fig. 8.9 con un segnale di ingresso sinusoidale alla frequenza di 1000 Hz e di

ampiezza variabile. Le curve rappresentano la differenza (in dB) all'uscita tra l'am-

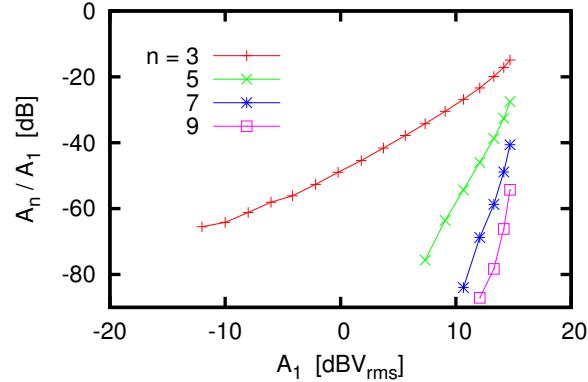


Figura 8.22: Armoniche dispari all'uscita dell'amplificatore di fig. 8.9 con un segnale di ingresso a frequenza di $1 kHz$ e ampiezza variabile. A_1 : ampiezza in dBV della componente fondamentale a $1000 Hz$ all'uscita. A_n/A_1 : rapporto tra le ampiezze dell'armonica n e della componente fondamentale.

piezza di ciascuna armonica e l'ampiezza della componente fondamentale, in funzione dell'ampiezza di quest'ultima. Le misure sono state fatte nell'intervallo di valori dell'ampiezza della fondamentale da $-12 dBV$ ($-12 dB$ rispetto ad una ampiezza di $1 V_{rms}$) a $+15 dBV$.

Sono riportate le sole armoniche dispari, dalla terza fino alla nona; per ragioni di simmetria, le armoniche dispari sono predominanti, quelle pari sono di ampiezza sensibilmente inferiore. Si vede che per una tensione di uscita fino a $3 dBV$ (circa $2 V$ di picco) si riesce a misurare la sola terza armonica, con un livello inferiore a $-40 dB$ (pari all' 1%) rispetto alla fondamentale. Oltre questo livello di tensione, l'amplificatore entra in saturazione, ed il livello delle armoniche, dalla quinta in poi, cresce molto rapidamente.

Capitolo 9

I transistor *mosfet* e *jfet*

9.1 Struttura del transistor *mosfet*

La sigla *mosfet* è un acronimo per *Metal-Oxide-Semiconductor Field-Effect-Transistor* (transistor ad effetto di campo di tipo metallo-ossido-semiconduttore). Il cuore del dispositivo è costituito dal condensatore *mos*, illustrato in fig. 9.1: una delle due armature del condensatore è formata da un *substrato* di semiconduttore drogato, nell'esempio Silicio di tipo *P*; l'altra armatura, detta *gate*, è formata da uno strato metallico, ad esempio Alluminio deposito per evaporazione sulla superficie del semiconduttore. Prima della deposizione del *gate* la superficie del substrato viene resa

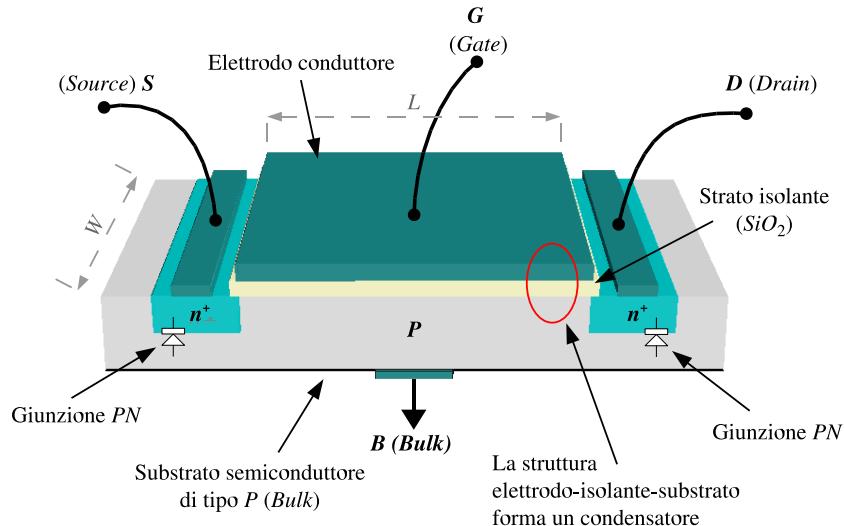


Figura 9.1: Struttura del transistor *mosfet*: il condensatore *mos*. L e W sono rispettivamente la lunghezza e la larghezza del condensatore. Le dimensioni del condensatore *mos* vanno dalla frazione di μm^2 per i dispositivi veloci per piccoli segnali fino al mm^2 per i dispositivi di potenza.

isolante mediante la formazione di uno strato di Ossido di Silicio (SiO_2), con uno spessore dell'ordine dei 100 nm. Su due lati di questa struttura vengono realizzate nel substrato due regioni, dette *source* e *drain* (*sorgente* e *drenaggio*), con droggaggio

opposto a quello del substrato (nell'esempio, N). Nella figura queste regioni sono contrassegnate come n^+ per indicare che si tratta di regioni con drogaggio molto forte.

Come nel caso del *bjt*, anche i transistor *mosfet* possono essere realizzati in due versioni con polarità simmetriche, a seconda che si parta da un substrato di tipo P o di tipo N . Si avranno quindi transistor complementari di tipo *nmos* e di tipo *pmos* (figura 9.2).

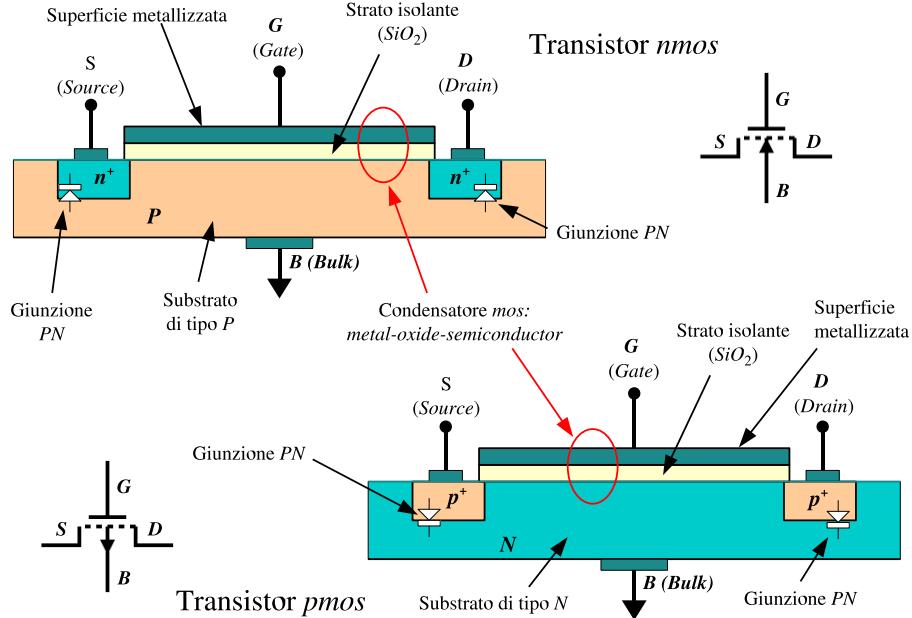


Figura 9.2: Transistor complementari *nmos* (in alto) e *pmos* (in basso) insieme con i simboli con cui sono rappresentati negli schemi elettrici.

In condizioni operative normali le due giunzioni $P-N$ presenti fra le due regioni di *source* e *drain* ed il substrato devono essere contropolarizzate: per un transistor *nmos* il substrato dovrà essere collegato ad un potenziale più negativo (o, quantomeno, non più positivo) di quello di *source* e *drain*, per un transistor *pmos* ad un potenziale più positivo. La presenza delle due giunzioni contropolarizzate contrapposte (rappresentate da diodi in fig. 9.2), rende impossibile il passaggio di carica tra *source* e *drain*. Applicando una tensione di polarizzazione di polarità e valore opportuni al *gate* è possibile arrivare alla formazione di un *canale* di conduzione costituito da una sottile lamina di cariche mobili in corrispondenza della superficie di separazione tra substrato e *gate*. Il canale di conduzione, le cui caratteristiche sono modulate dalla tensione di *gate*, rende possibile il fluire di una corrente tra *source* e *drain*.

9.2 Canale di conduzione

Applicando una differenza di potenziale tra *gate* e substrato di un transistor *mosfet* si modifica lo stato di carica del condensatore *mos*.

Nel caso di un transistor *nmos*, si hanno le situazioni schematizzate in fig. 9.3.

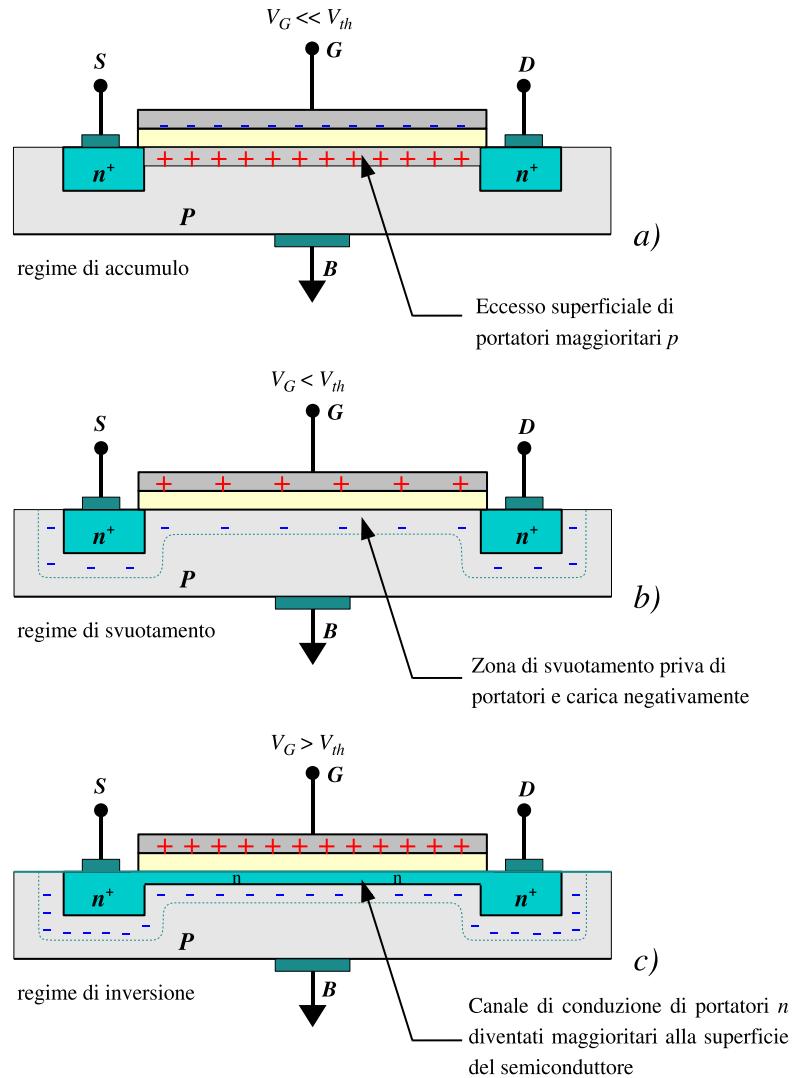


Figura 9.3: Formazione del canale di conduzione in un transistor *nmos*.

In 9.3a il gate si trova ad un potenziale più negativo del substrato, di tipo P . I portatori di carica positivi, maggioritari, vengono attratti verso il *gate* negativo e si accumulano alla superficie. Il condensatore *mos* si comporta come un normale condensatore a facce piane e parallele e presenta una capacità costante al variare del potenziale (fig. 9.4, regione tra -10 e -3 V circa). In questa condizione, detta di *accumulo*, i portatori di carica presenti tra *source* e *drain* non possono dare luogo ad alcuna conduzione, a causa della presenza delle due giunzioni contropolarizzate tra *source* e *drain* e substrato.

Portando il potenziale del *gate* verso valori più positivi, la carica del condensatore *mos* diminuisce, riducendosi a zero e successivamente invertendosi di segno, divenendo quindi positiva sulla superficie del *gate* e negativa sulla superficie del substrato (fig. 9.3b). La zona di carica negativa presente alla superficie del substrato in fig. 9.3b corrisponde ad una zona di svuotamento: i portatori positivi maggioritari del sub-

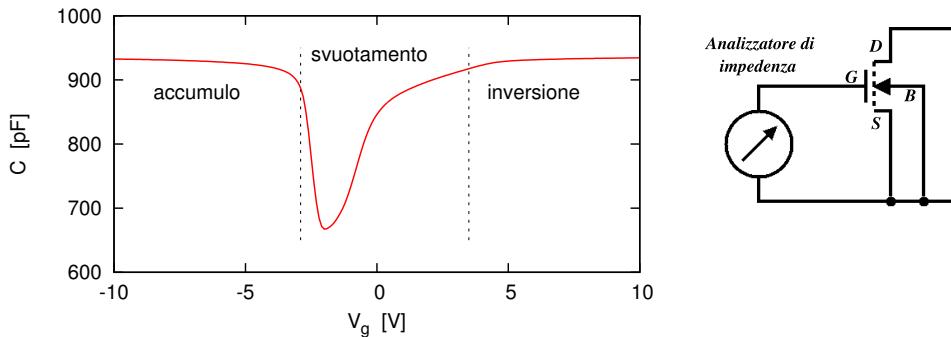


Figura 9.4: Capacità del condensatore mos in funzione della tensione di *gate* misurata su un dispositivo IRF820 (mosfet di potenza a canale N). A destra è riportato il circuito di misura. L'apparecchio analizzatore di impedenza provvede anche ad applicare la tensione di polarizzazione. Frequenza di misura: 1 MHz

strato, respinti dal campo elettrico, lasciano una zona scoperta di cariche negative nel reticolo cristallino. La capacità del condensatore *mos* diminuisce, in quanto la zona di svuotamento, analogamente a quanto succede nella giunzione PN del diodo, presenta una capacità di transizione che si riduce all'aumentare della contropolarizzazione e che si viene a trovare in serie con la capacità del condensatore *mos* che si aveva nel regime di accumulo (fig. 9.4, regione tra -3 e $+3.5$ V circa). Non essendoci portatori di carica nella zona di svuotamento, non vi può essere conduzione tra *source* e *drain*.

Portando il potenziale del *gate* verso valori ancora più positivi, la capacità passa per un minimo (intorno a -2 V in fig. 9.4), quindi torna ad aumentare. Infatti, all'aumentare del potenziale del *gate*, diminuisce sempre di più la concentrazione dei portatori positivi maggioritari nella zona di svuotamento, sino ad arrivare alla concentrazione del semiconduttore intrinseco. Passato questo punto, detto *inversione debole*, i portatori inizialmente minoritari divengono maggioritari in una zona laminare immediatamente sotto il *gate*, in cui si ha la inversione della polarità del semiconduttore da P ad N , e questo porta nuovamente alla formazione di uno strato di cariche libere, ora negative, alla superficie dello strato isolante del *gate*, ritornando quindi alla situazione del condensatore a facce piane del regime di accumulo.

Per un valore della tensione di *gate* V_G sufficientemente elevato ($V_G \geq V_{th}$, con V_{th} detta *tensione di soglia [threshold]*), la concentrazione dei portatori n , inizialmente minoritari in prossimità del *gate*, supera quella dei portatori p maggioritari nel substrato e dà luogo alla formazione di un *canale di conduzione* che interconnette *source* e *drain* e rende possibile il passaggio di cariche tra i due elettrodi (*inversione forte*, fig. 9.3c, regione oltre i 3.5 V in fig. 9.4).

È da notare che la condizione di carica zero del condensatore *mos* non coincide con la condizione di differenza di potenziale nulla tra i terminali di *gate* e substrato principalmente a causa della presenza dei potenziali di contatto tra i diversi materiali della intera catena elettrodica.

9.3 Conduttanza del canale

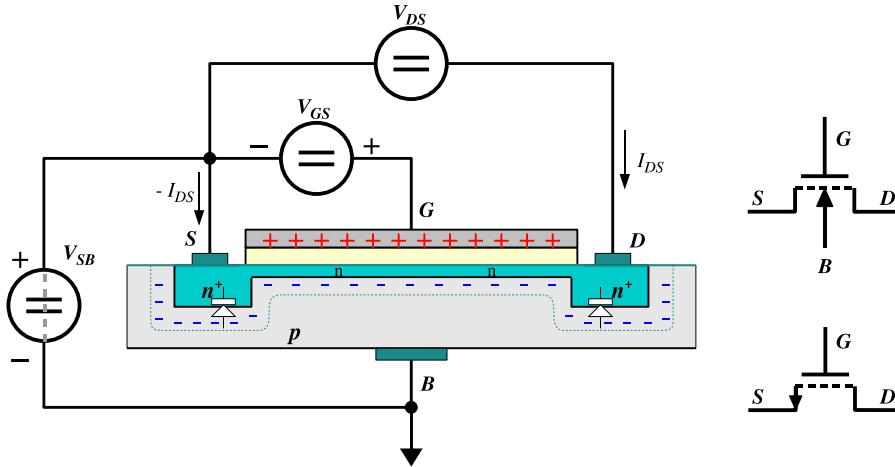


Figura 9.5: Polarizzazione di un transistor *nmos*. A destra sono riportati i simboli utilizzati per rappresentare il transistor *nmos* nei circuiti.

In fig. 9.5 sono riportate le condizioni di polarizzazione di un transistor *nmos*. Il generatore V_{SB} garantisce la condizione di contropolarizzazione della giunzione *source*-substrato. Il generatore V_{GS} polarizza positivamente il *gate*, in modo da garantire la formazione del canale; il generatore V_{DS} polarizza il *drain* rispetto al *source* e fornisce la corrente I_{DS} , la cui circolazione è resa possibile e modulata dalla presenza del canale. La polarità del generatore V_{DS} non è indicata nella figura. A differenza del *bjt*, dove emettitore e collettore sono due eletrodi con caratteristiche costruttive asimmetriche, il *drain* ed il *source* del *mosfet* sono simmetrici e di principio intercambiabili. Nella situazione riportata in fig. 9.5 il *drain* può essere sia positivo sia negativo rispetto al *source*, purchè sia rispettata la condizione $V_{SB} + V_{DS} > 0$ per garantire la contropolarizzazione della giunzione *drain*-substrato. Per valori di V_{DS} piccoli (secondo il criterio che si vedrà al paragrafo successivo) il canale ha un comportamento di tipo ohmico (fig. 9.6) con conduttanza

$$G_C = \mu_n Q_n W/L \quad (9.1)$$

dove μ_n è la mobilità dei portatori di carica (negativi) presenti nel canale, W ed L sono larghezza e lunghezza del canale e Q_n è la densità superficiale di carica dei portatori, funzione del potenziale di polarizzazione V_{GS} .

Indicando con V_{th} la tensione di soglia a cui il canale prende consistenza, corrispondente alla tensione di inversione forte, e considerando che il sistema ha la geometria di un condensatore piano con dielettrico di spessore d e costante dielettrica ϵ , si ha

$$Q_n = \frac{V_{GS} - V_{th}}{d} \epsilon \quad (9.2)$$

In pratica, non è strettamente necessario che le giunzioni *source*-substrato e *drain*-substrato siano realmente contropolarizzate: è sufficiente che non siano direttamente

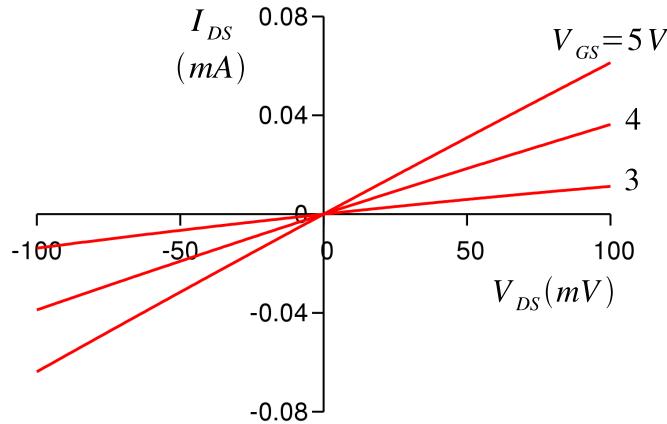


Figura 9.6: Relazione lineare tra tensione V_{DS} e corrente I_{DS} in un transistor *nmos* nell'intorno di $V_{DS} = 0$. La pendenza I_{DS}/V_{DS} è la conduttanza G_C del canale.

polarizzate. Si può quindi porre $V_{SB} = 0$, sostituendo il generatore tra *source* e substrato con il cortocircuito rappresentato a tratteggio in fig. 9.5. In questo caso però si perde la simmetria tra *source* e *drain*: si dovrà avere necessariamente $V_{DS} \geq 0$ e di conseguenza $I_{DS} \geq 0$. Dei due possibili quadranti operativi riportati in fig. 9.6 rimarrà accessibile solo quello superiore destro.

In fig. 9.5, in basso a destra, è riportata una variante del simbolo circuitale del transistor *nmos* utilizzata per rappresentare i dispositivi in cui *source* e substrato siano stati connessi internamente dal costruttore e non siano separatamente accessibili dall'esterno. Per il transistor *pmos* il verso della freccia del *source* è rovesciato.

9.4 Modulazione del canale

Applicando una tensione V_{DS} tra *drain* e *source* il potenziale lungo il canale non è più costante, ma viene a dipendere dalla posizione x tra *source* ($x = 0$) e *drain* ($x = x_D$). L'equazione 9.2 dovrà essere modificata per tenere conto di questo effetto:

$$Q_n(x) = C [V_{GS} - V_{th} - V(x)] \quad (9.3)$$

dove $V(x)$ è il potenziale elettrico lungo il canale, misurato rispetto al *source* e $C = \epsilon/d$ la capacità per unità di superficie tra *gate* e substrato (fig. 9.7). L'equazione 9.1 continua ad essere valida localmente, in ogni punto del canale. In presenza di un campo elettrico E_x diretto da *drain* a *source*, la densità superficiale di corrente è:

$$J_x = \mu_n Q_n(x) E_x \quad (9.4)$$

Poichè il canale è omogeneo lungo la direzione y ortogonale al piano del disegno in fig. 9.7, la corrente totale si può ottenere moltiplicando per la larghezza W del canale:

$$I_{DS} = -\mu_n Q_n(x) W E_x \quad (9.5)$$

(per le convenzioni adottate, la corrente I_{DS} ed il campo E_x hanno direzioni positive opposte).

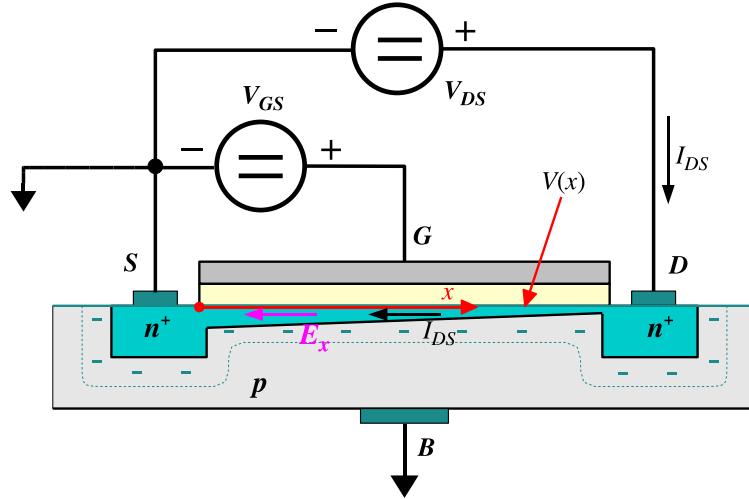


Figura 9.7: Modulazione della consistenza del canale in funzione della posizione tra *drain* e *source*.

Il valore di I_{DS} non dipende dalla coordinata x in quanto le cariche possono fluire lungo il canale solo tra *drain* e *source*: la corrente verso il *gate* è nulla a causa dello strato isolante di SiO_2 e la corrente verso il substrato consiste solo nella debolissima corrente di saturazione inversa delle giunzioni PN , di fatto trascurabile.

Sostituendo la eq. 9.3 nella 9.5

$$I_{DS} = -\mu_n C [V_{GS} - V_{th} - V(x)] W E_x \quad (9.6)$$

integrandolo sulla lunghezza del canale

$$\int_0^{x_D} I_{DS} dx = \mu_n C W \int_0^{x_D} [V_{GS} - V_{th} - V(x)] \frac{dV}{dx} dx \quad (9.7)$$

e considerando che I_{DS} è costante e $V(x_D) = V_{DS}$, si ottiene:

$$I_{DS} \cdot L = \mu_n C W \left(V_{GS} - V_{th} - \frac{V_{DS}}{2} \right) V_{DS} \quad (9.8)$$

da cui

$$I_{DS} = K_n \left(V_{GS} - V_{th} - \frac{V_{DS}}{2} \right) V_{DS} \quad (9.9)$$

con $K_n = \mu_n C W/L$.

9.5 Pinch-off

Aumentando la tensione V_{DS} fino al valore $V_{DS} = V_{GS} - V_{th}$ si ha, in base alla eq. 9.3,

$$Q_n(x_D) = C (V_{GS} - V_{th} - V_{DS}) = 0 \quad (9.10)$$

La consistenza del canale, data dalla densità superficiale $Q_n(x)$ dei portatori di carica, si assottiglia progressivamente muovendosi dal *source* verso il *drain*, fino ad annullarsi

del tutto in corrispondenza del *drain*. Questa è la condizione di *pinch-off* (*strozzamento*) del canale (fig. 9.8). Sostituendo nella eq. 9.9 la condizione di *pinch-off* data

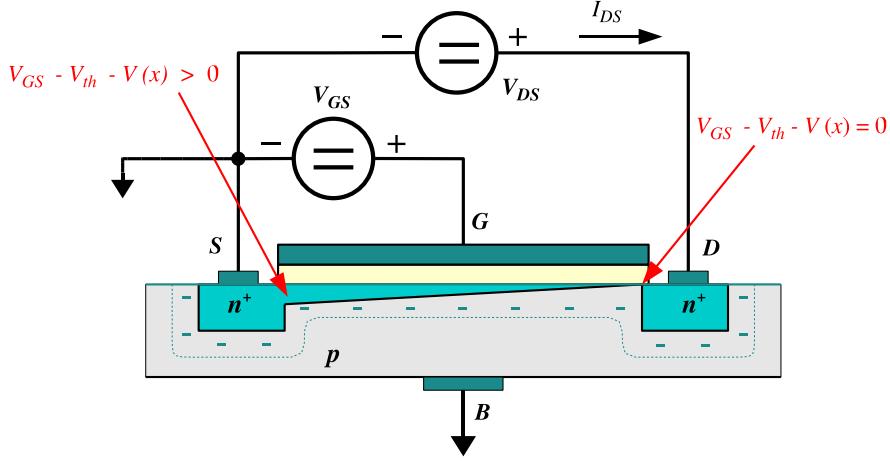


Figura 9.8: Condizione di *pinch-off*: strozzamento del canale di conduzione.

dalla 9.10 si ottiene per la corrente:

$$I_{DS} = \frac{K_n}{2} \cdot (V_{GS} - V_{th})^2 \quad (9.11)$$

Aumentando ulteriormente la tensione V_{DS} l'equazione 9.9 cessa di essere valida ed il transistor entra in regime di *saturazione*. La corrente I_{DS} tende ad aumentare solo molto lentamente, con una dipendenza approssimativamente lineare con la tensione:

$$I_{DS} = \frac{K_n}{2} \cdot (V_{GS} - V_{th})^2 (1 + \lambda V_{DS}) \quad (9.12)$$

Il coefficiente λ è l'analogo del *coefficiente di Early* ($1/V_A$) del *bjt*.

9.6 Curve caratteristiche per il transistor *nmos*

La famiglia di curve di I_{DS} in funzione di V_{DS} e V_{GS} riportata in fig. 9.9 riassume il funzionamento del transistor *nmos*.

La *regione lineare* corrisponde alla condizione $V_{GS} - V_{th} > V_{DS}$, descritta dalla equazione 9.9; per valori di tensione V_{DS} molto piccoli rispetto a $V_{GS} - V_{th}$ il termine $V_{DS}/2$ in parentesi risulta trascurabile ed il comportamento si avvicina strettamente a quello ohmico riportato in fig. 9.6. Nel transistor *bjt* la regione analoga alla regione lineare del *mosfet* è la *regione di saturazione*.

La *regione di saturazione* per il *mosfet*, la cui analogia per il *bjt* è la *regione attiva*, è descritta dalla eq. 9.11 o dalla 9.12. Il confine tra le due regioni è rappresentato dalla parabola $I_{DS} = V_{DS}^2 \cdot K_n/2$ (linea blu in figura).

La terza regione, di *interdizione* o *cutoff*, corrisponde a $V_{GS} < V_{th}$, quindi alla condizione di transistor spento ($I_{DS} = 0$). È di interesse principalmente in elettronica digitale, quando il transistor viene utilizzato come interruttore tra i due stati acceso (regione lineare o saturazione) / spento (interdizione).

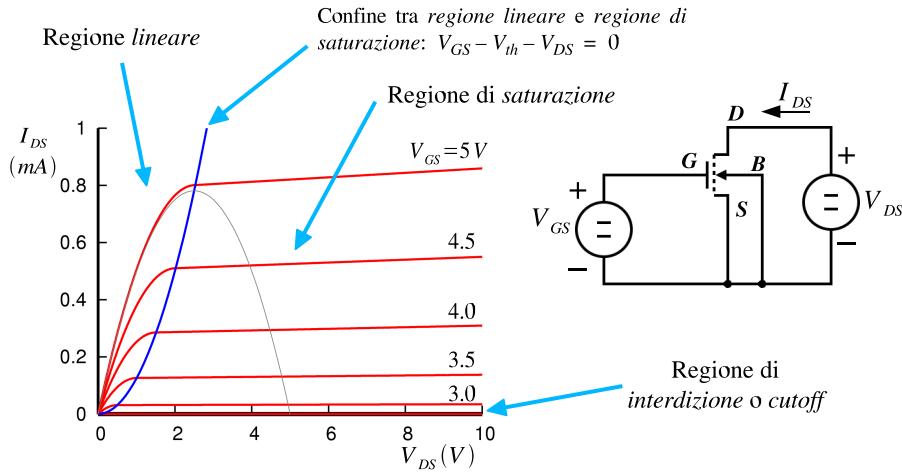


Figura 9.9: Famiglia di curve caratteristiche per un transistor *nmos* ($K_n = 0.25 \text{ mA/V}^2$, $V_{th} = 2.5 \text{ V}$, $\lambda = 0.01 \text{ V}^{-1}$).

Nella figura è riportato anche il circuito di misura per il rilievo delle curve caratteristiche. Il transistor è utilizzato con il *source* in comune tra il circuito di ingresso e quella di uscita (configurazione *common source*).

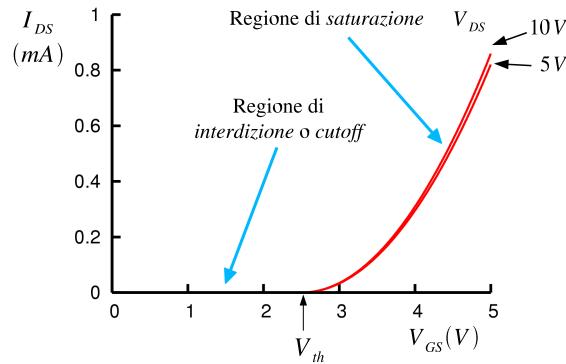


Figura 9.10: Curve $I_{DS}(V_{GS})$ per diversi valori di V_{DS} (stesso transistor di fig. 9.9).

In fig. 9.10 sono riportate le caratteristiche per lo stesso transistor di fig. 9.9, utilizzando V_{GS} come variabile indipendente e V_{DS} come parametro. Sono distinguibili la *regione di interdizione* ($V_{GS} < V_{th}$) e la *regione di saturazione* ($V_{GS} > V_{th}$), con la crescita parabolica della corrente in funzione di V_{GS} .

La regione lineare, corrispondente a valori molto piccoli di V_{DS} , non è riportata.

9.7 Il transistor *pmos* - *mosfet complementari*

Tutto quanto detto ai paragrafi precedenti per il transistor *mosfet* a canale *n* vale identicamente per il *mosfet* a canale *p*, invertendo le polarità di tutte le tensioni e

correnti (fig. 9.11). Transistor *mosfet* che hanno le medesime caratteristiche ($|V_{th}|$, K_n , $|\lambda|$) ma opposta polarità del canale sono denominati *complementari* o *cmos*.

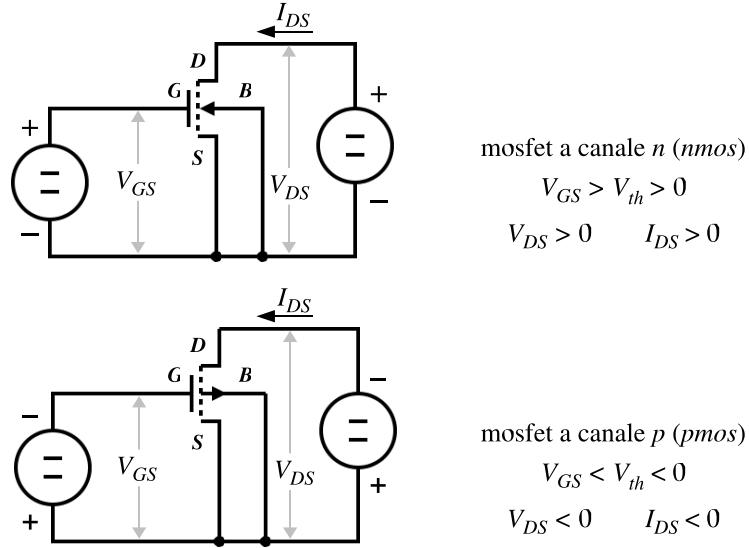


Figura 9.11: Transistor *mosfet* complementari.

9.8 Transistor *mosfet* di tipo *enhancement* e *depletion*

Nei transistor *mosfet* descritti nei paragrafi precedenti si ha la formazione del canale e quindi la conduzione tra *drain* e *source* solo in presenza di una tensione di polarizzazione tra *gate* e *source* sufficientemente elevata ($V_{GS} > V_{th}$ per l'*nmos* e $V_{GS} < V_{th}$ per il *pmos*). Questi *mosfet* sono detti di tipo *enhancement*: ad arricchimento. È possibile realizzare anche transistor *mosfet* in cui si *impanta* un canale di conduzione tra *source* e *drain* già durante il processo di fabbricazione, mediante la formazione con drogaggi opportuni di lamine di portatori di carica alla superficie tra substrato e *gate*. In questi transistor l'effetto della tensione di polarizzazione del *gate* è di aumentare la conduzione già presente nel canale, oppure di diminuirla fino ad azzerarla completamente, a seconda della polarità della tensione applicata. Questi dispositivi sono detti di tipo *depletion*: ad impoverimento.

In fig. 9.12 sono riportate a confronto le curve $I_{DS}(V_{GS})$ per un mosfet a canale *n* di tipo *enhancement* (curve rosse) ed uno di tipo *depletion* (curve blu). Le relazioni funzionali per i due tipi di mosfet sono esattamente le stesse; unica differenza è che le curve del *mosfet* di tipo *depletion* sono traslate a sinistra sull'asse delle tensioni V_{GS} rispetto alle curve del tipo *enhancement*. In un caso a $V_{GS} = 0$ si ha corrente I_{DS} nulla; nell'altro caso si ha una corrente finita che può essere variata in aumento o in diminuzione dalla tensione del *gate*.

Negli schemi elettrici i transistor di tipo *depletion* sono rappresentati con un simbolo simile a quello dei transistor *enhancement*. La sola differenza è il segmento che connette *source* e *drain* e rappresenta il canale: nel primo caso è continuo, nel secondo è tratteggiato (fig. 9.14).

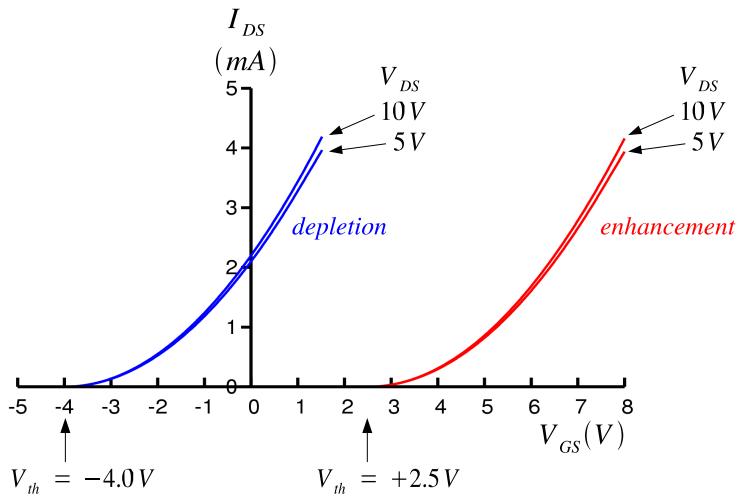


Figura 9.12: Curve $I_{DS}(V_{GS})$ per un *mosfet* di tipo *enhancement* (curve rosse) e di tipo *depletion* (curve blu) aventi eguali valori di K_n e $|\lambda|$.

9.9 Struttura del transistor *jfet*

La sigla *jfet* è un acronimo per *junction field effect transistor* (transistor a giunzione ad effetto di campo).

Il transistor jfet mostra forti analogie con il transistor mosfet per quanto riguarda le caratteristiche elettriche, anche se il principio di funzionamento è alquanto diverso. Anche nel jfet abbiamo una tensione di polarizzazione ad un elettrodo *gate* (*G*) che modula la conduzione di un *canale* che connette un elettrodo *source* (*S*) ad un *drain* (*D*) (fig. 9.13). Il gate però non è elettricamente isolato dal canale come nel mosfet, ma forma con questo una giunzione *P-N*. In condizioni operative normali questa giunzione è contropolarizzata per cui anche qui si ha $I_G = 0$ (a meno della corrente di saturazione inversa della giunzione).

A differenza di quanto avviene nel mosfet, la conduzione non viene modificata mediante un arricchimento od un impoverimento della concentrazione dei portatori nel canale, ma piuttosto variandone la sezione: una variazione della tensione di contropolarizzazione della giunzione gate-canale fa variare le dimensioni della regione di ssvuotamento. Normalmente il droggaggio della regione di gate (p^+ in figura) è molto più forte di quello del canale (n), per cui una estensione o riduzione della zona di ssvuotamento avviene principalmente a spese od a vantaggio di quest'ultima regione.

Lo schema di principio del transistor jfet è riportato in fig. 9.13. La regione del canale è drogata *N*; si tratta quindi di un jfet *di tipo n* o *a canale n*. Simmetricamente si possono realizzare jfet a canale *p*, invertendo la polarità dei drogaggi.

Anche nel jfet come nel mosfet la consistenza del canale varia in funzione della posizione tra drain e source per effetto della distribuzione della tensione V_{DS} lungo il canale. Anche in questo caso si ha il fenomeno della strozzatura del canale (pinch-off) per valori di V_{GS} sufficientemente negativi o per valori di V_{DS} sufficientemente positivi (per il jfet di tipo *n*).

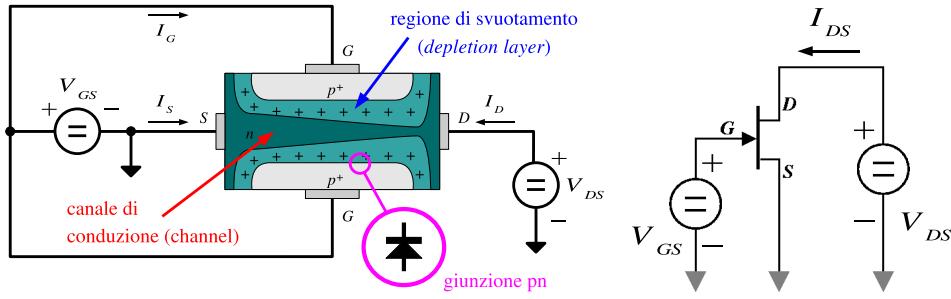


Figura 9.13: Schema di principio (in sezione) di un transistor jfet *di tipo n*. Il canale è costituito da semiconduttore *N*; il gate da una o due regioni ad alto drogaggio di semiconduttore *P*, che forma giunzioni *P-N* con il canale. A destra è riportato il simbolo per il jfet a canale *n* con l'indicazione delle polarizzazioni nelle condizioni operative normali. Per il jfet a canale *p* tutte le tensioni hanno polarità opposta ed il verso della freccia in corrispondenza del gate è rovesciato.

9.10 Modulazione della corrente di canale nel transistor jfet

L'analisi del trasporto della carica attraverso il canale di un jfet porta alle stesse equazioni che sono state trovate nel caso del mosfet (eq. 9.9, 9.11 e 9.12, par. 9.4 e 9.5), con alcune differenze nei simboli, dovute alla tradizione.

La tensione di gate per la quale la corrente del canale si riduce a zero (strozzamento del canale) prende il nome di tensione di *pinch-off* (V_P) ed è l'analogo della tensione di soglia V_{th} del mosfet.

Facendo riferimento ad un jfet a canale *n* come quello riportato in fig. 9.13, con $V_{GS} - V_P \geq V_{DS} \geq 0$ il dispositivo si trova nella *regione lineare* e vale la relazione:

$$I_{DS} = \frac{2 I_{DSS}}{V_P^2} \left(V_{GS} - V_P - \frac{1}{2} V_{DS} \right) V_{DS} \quad (9.13)$$

Il termine $2 I_{DSS}/V_P^2$ è l'analogo del coefficiente K_n della eq. 9.9.

Con $V_{DS} > V_{GS} - V_P$ il dispositivo si trova nella *regione di saturazione* e vale la relazione:

$$I_{DS} = \frac{I_{DSS}}{V_P^2} (V_{GS} - V_P)^2 \quad (9.14)$$

analoga alla 9.11 oppure, tenendo conto anche dell'effetto Early,

$$I_{DS} = \frac{I_{DSS}}{V_P^2} (V_{GS} - V_P)^2 (1 + \lambda V_{DS}) \quad (9.15)$$

analoga alla 9.12.

Dalle ultime due relazioni si vede che la quantità I_{DSS} rappresenta la corrente I_{DS} nel canale in condizioni di saturazione con tensione di gate $V_{GS} = 0$, trascurando l'effetto Early.

Le curve riportate in fig. 9.9 per il transistor nmos sono quindi egualmente valide per il jfet a canale *n*, con gli opportuni valori per la tensione di gate V_{GS} . Infatti, delle due famiglie di curve in fig. 9.12, quelle appropriate al transistor jfet

sono quelle di sinistra, relative al mosfet di tipo depletion: nel jfet alla tensione di gate $V_{GS} = 0$ il canale ha il massimo della consistenza e la corrente I_{DS} ha il valore massimo. L'applicazione di una tensione V_{GS} via via più negativa riduce la dimensione del canale e quindi il valore di I_{DS} , fino ad arrivare ad azzerare la conduzione per $V_{GS} = V_P$. Il valore di V_P è sempre negativo per jfet a canale n e positivo per jfet a canale p . A differenza del mosfet, in condizioni operative normali la tensione V_{GS} dovrà essere sempre compresa tra V_P e 0 e non potrà andare oltre lo zero, dovendo mantenere la condizione di contropolarizzazione della giunzione tra gate e canale.

9.11 Varietà di transistor ad effetto di campo

Riassumendo, si possono avere transistor mosfet a canale n ed a canale p , ciascuno dei quali può essere di tipo *enhancement* o di tipo *depletion*. Ognuno di questi può avere il substrato collegato o non collegato costruttivamente al *source*.

Inoltre si possono avere transistor jfet a canale n ed a canale p .

In fig. 9.14 sono riportati i simboli con cui viene rappresentata negli schemi elettrici tutta questa varietà di dispositivi.

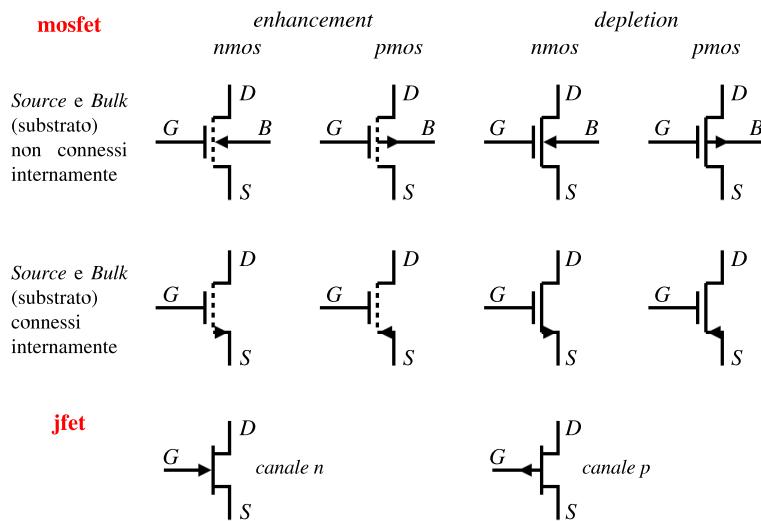


Figura 9.14: Simboli utilizzati per rappresentare negli schemi elettrici le diverse varietà di transistor ad effetto di campo.

9.12 Circuiti con transistor ad effetto di campo

9.12.1 Amplificatore *common source*

Due esempi di amplificatori *common source* realizzati con un transistor ad effetto di campo (*fet*) sono riportati in fig. 9.15. Il primo (fig. 9.15a) impiega un *nmos* di tipo *enhancement*; il secondo (fig. 9.15b) impiega un jfet a canale n o, equivalentemente, un *nmos* di tipo *depletion*. I circuiti di polarizzazione sono diversi nei due casi. In

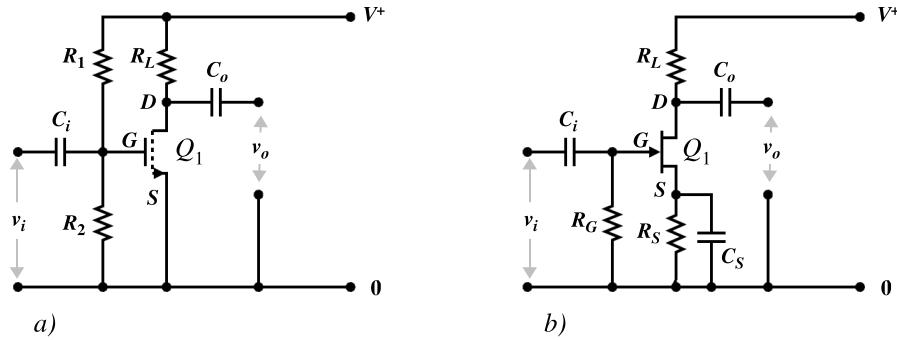


Figura 9.15: Due esempi di amplificatori *common source*. Il primo (a) utilizza un mosfet di tipo *enhancement*; il secondo (b) un jfet (o, in alternativa, un mosfet di tipo *depletion*).

a) si richiede una tensione di *gate* positiva rispetto al *source* e la si ottiene dalla alimentazione V^+ tramite il partitore R_1, R_2 :

$$V_{GS} = \frac{R_2}{R_1 + R_2} \cdot V^+ \quad (9.16)$$

Nel secondo caso la tensione V_{GS} deve essere negativa: la resistenza R_S porta il *source* alla tensione $V_S = I_{DS} \cdot R_S$, mentre la resistenza R_G mantiene il *gate* al potenziale $V_G = 0$. Di conseguenza $V_{GS} = -I_{DS} \cdot R_S$. Questo schema di polarizzazione è identico a quello che è stato in uso per buona parte del secolo scorso per la polarizzazione dei tubi a vuoto (triodi, pentodi ecc.) utilizzati come amplificatori.

Anche in questi circuiti, come già visto per il transistor *bjt*, i condensatori C_i e C_o isolano lo stadio amplificatore dagli altri circuiti per quanto riguarda le tensioni continue di polarizzazione, consentendo il solo passaggio dei segnali in corrente alternata. Il condensatore C_S in 9.15b cortocircuita la resistenza R_S per quanto riguarda i segnali, eliminando la reazione negativa che verrebbe altrimenti introdotta.

Come per gli altri casi, il funzionamento del circuito può essere studiato riportando sul grafico delle curve caratteristiche la retta di carico di equazione $I_{DS} = (V^+ - V_{DS})/R_L$ ed esaminando lo spostamento del punto di lavoro al variare di $v_{GS} = V_{GS} + v_{gs}$ e le conseguenti variazioni di v_{DS} (fig. 9.16) oppure risolvendo per via numerica le equazioni del circuito.

Nel caso del circuito di fig. 9.15b si ha una retta di carico *statica*, corrispondente al comportamento del circuito in corrente continua, cioè in pratica ignorando la presenza di C_S (retta di carico in magenta in fig. 9.16b), ed una retta *dinamica* (in blu) per i segnali con frequenze per cui C_S corrisponde ad un corto circuito: per segnali veloci v_S rimane fermo al valore V_S per effetto della presenza del condensatore C_S e v_D può variare nell'intervallo tra V_S e V^+ ; di conseguenza v_{DS} varia tra 0 e $V^+ - V_S$.

9.12.2 Modello per piccoli segnali per i transistor ad effetto di campo

Nel caso di *piccoli segnali* ($v_{gs} \ll V_{GS}$ e di conseguenza $v_{ds} \ll V_{DS}$ e $i_{ds} \ll I_{DS}$) è possibile utilizzare un modello lineare del transistor mosfet o jfet ed effettuare i calcoli sul circuito lineare equivalente.

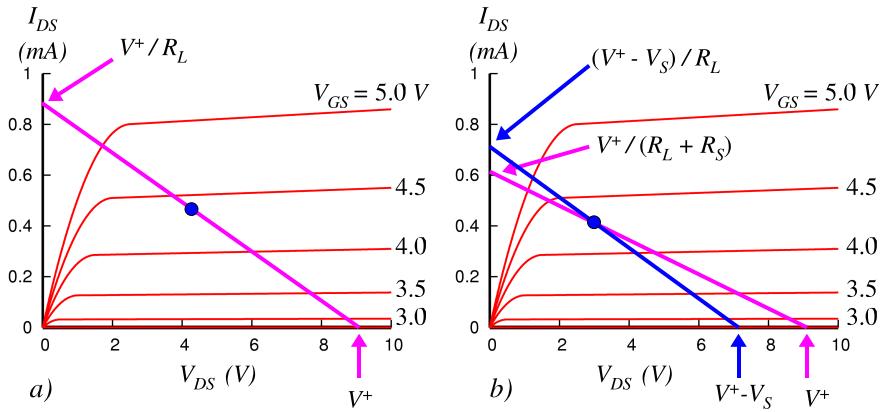


Figura 9.16: Curve caratteristiche e retta di carico per i due circuiti di fig. 9.15. Per il circuito *b* si hanno due rette di carico: una statica (in colore magenta), corrispondente al comportamento in corrente continua, ed una dinamica (in blu) che descrive il comportamento del circuito per quanto riguarda i segnali.

Delle quattro famiglie di parametri riportate al par. 6.4 per la descrizione dei dispositivi lineari a due porte, solo due sono utilizzabili nel caso dei transistor fet nella configurazione *common source*. Poiché la corrente del *gate* è praticamente nulla, almeno in corrente continua ed in bassissima frequenza, non ha molto senso scegliere i_1 come variabile indipendente: in questa configurazione il fet è un dispositivo in cui il parametro di ingresso è essenzialmente una tensione. Si utilizzeranno quindi i parametri g o m , a cui corrispondono i due circuiti lineari in fig. 9.17. In entrambi i

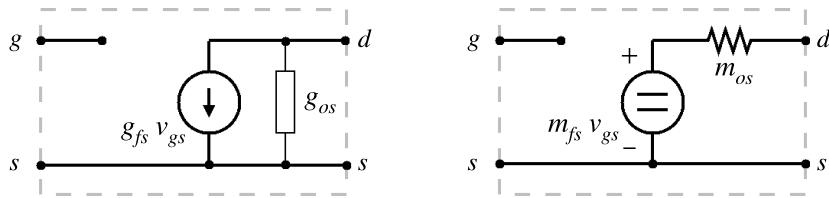


Figura 9.17: Circuiti lineari equivalenti per piccoli segnali per i transistor fet nella configurazione *common source*.

casi i parametri i ed r sono stati omessi in quanto di fatto nulli in corrente continua. Nella regione di saturazione, dove viene solitamente posizionato il punto di lavoro del *fet* come amplificatore, il valore di g_{os} è molto piccolo, in prima approssimazione zero se si trascura l'effetto della *tensione di Early* e quindi $m_{os} = \infty$. Di conseguenza i parametri g sono i più convenienti e i più comunemente usati nel modello lineare dei *fet*.

Con i transistor ad effetto di campo è possibile realizzare tutte le configurazioni circuitali viste nel caso del *bjt*, quindi in particolare l'amplificatore differenziale, il *current mirror* ed il *source follower*, che è l'equivalente dell'*emitter follower*.

La differenza più appariscente tra le prestazioni di transistor *bjt* e *fet* è costituita ovviamente dal parametro di ingresso: i transistor ad effetto di campo hanno impedenza di ingresso praticamente infinita, mentre il *bjt* presenta tipicamente impedenze di ingresso nel campo dai $k\Omega$ ai $M\Omega$. Questo è il motivo principale per cui oggi gli amplificatori operazionali hanno quasi invariabilmente uno stadio di ingresso che utilizza una coppia differenziale di *fet*, di solito *jfet*.

D'altra parte i transistor *bjt* sono superiori rispetto ai *fet* per un'altra caratteristica: a parità di corrente offrono un parametro g_f , e quindi un guadagno di tensione, maggiore. Inoltre, allo stato attuale dell'arte, il massimo di stabilità ed il minimo di rumore di tensione si riesce ad ottenerli con i transistor *bjt*.

In conclusione i transistor *fet* saranno da preferire in tutte le applicazioni in cui si richiede il massimo di sensibilità in corrente, mentre i *bjt* daranno risultati migliori nei circuiti in cui la sensibilità in tensione è l'obiettivo principale.

Capitolo 10

Elettronica digitale

In *elettronica analogica* le variabili elettriche in un circuito, tensioni e correnti, sono funzioni continue, non necessariamente lineari, delle grandezze fisiche che rappresentano; nei circuiti elettronici digitali invece le variabili elettriche assumono solo due valori discreti, indicati convenzionalmente con una coppia di simboli come 0/1 oppure *Falso/Vero* (*False/True*), *Basso/Alto* (*Low/High*), *Spento/Acceso* (*Off/On*), ecc. (fig. 10.1). Una variabile digitale elementare, con il suo valore 0/1, prende il nome di variabile logica o variabile binaria o *bit* (*bit = binary digit*).

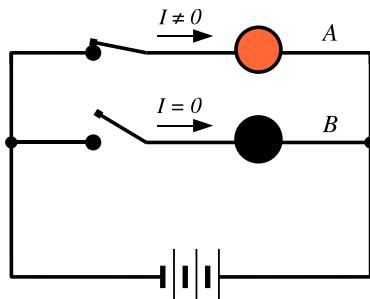


Figura 10.1: Circuiti digitali: il circuito *A* è *acceso* (circola corrente), il circuito *B* è *spento* (non circola corrente).

In elettronica digitale un segnale è rappresentato da un certo numero di bit che ne esprimono il valore secondo un opportuno codice. Nella tabella 10.1 sono riportati i primi 16 elementi del codice di numerazione binaria mediante il quale è possibile rappresentare, utilizzando N bit, 2^N valori diversi, ad esempio i primi 2^N numeri interi. Nella tabella 10.2 è riportata una parte del codice ASCII, in cui vengono utilizzate sequenze di 8 bit per rappresentare i caratteri dell'alfabeto ed i segni di interpunkzione.

Al confine tra il mondo dell'elettronica analogica e quello digitale vi sono i *Convertitori Analogico-Digitale (ADC)* e *Digitale-Analogico (DAC)*. Gli *ADC* hanno come ingresso un singolo segnale analogico e come uscita un certo numero di bit che, dopo la conversione, esprimono il valore numerico del segnale di ingresso; i convertitori *DAC* svolgono la funzione inversa: impostando sui bit che costituiscono le variabili

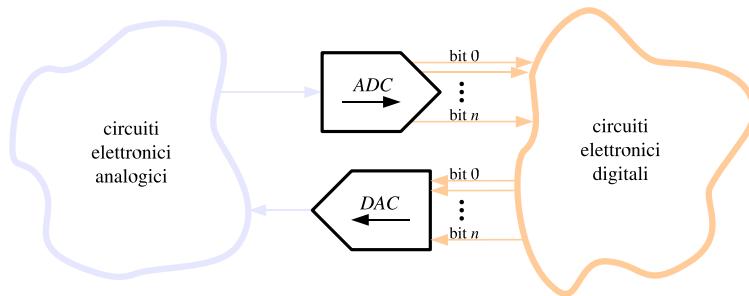
valore decimale	codice binario				valore decimale	codice binario			
	2^3	2^2	2^1	2^0		2^3	2^2	2^1	2^0
0	0	0	0	0	8	1	0	0	0
1	0	0	0	1	9	1	0	0	1
2	0	0	1	0	10	1	0	1	0
3	0	0	1	1	11	1	0	1	1
4	0	1	0	0	12	1	1	0	0
5	0	1	0	1	13	1	1	0	1
6	0	1	1	0	14	1	1	1	0
7	0	1	1	1	15	1	1	1	1

Tabella 10.1: Numeri interi da 0 a 15 e corrispondente codice binario.

	000	001	010	011	100	101	110	111
00100	"!"	".."	#	\$	%	&	'	
00101	()	*	+	,	-	.	/
00110	0	1	2	3	4	5	6	7
00111	8	9	:	;	<	=	>	?
01000	@	A	B	C	D	E	F	G
01001	H	I	J	K	L	M	N	O
01010	P	Q	R	S	T	U	V	W
01011	X	Y	Z	[\]	” ”	” ”	
01100	'	a	b	c	d	e	f	g
01101	h	i	j	k	l	m	n	o
01110	p	q	r	s	t	u	v	w
01111	x	y	z	{ }		{ }	{ }	del

Tabella 10.2: Parte del codice ASCII. I caratteri dell’alfabeto, i segni di interpunzione, le cifre ed altri segni aritmetici sono rappresentati da sequenze di 8 bit. La prima colonna a sinistra riporta i 5 bit più significativi (7...3); la prima riga in alto i tre bit meno significativi (2...0).

digitali di ingresso un codice numerico, il *DAC* sintetizza la tensione (o corrente) di valore corrispondente (fig. 10.2).

Figura 10.2: I convertitori da *Analogico a Digitale* (*ADC*) e da *Digitale ad Analogico* (*DAC*) costituiscono il legame tra le variabili continue delle reti analogiche ed i valori *acceso/spento* dei bit delle reti digitali.

La conversione da analogico a digitale è il destino finale di ogni misura di una

grandezza fisica analogica. Questo può avvenire tardi, alla fine del processo di misura, quando un operatore legge sulla scala numerata di un quadrante un valore e lo riporta in forma numerica su un quaderno di appunti, oppure può avvenire precocemente all'interno degli stessi strumenti di misura, ad opera di un *ADC* (fig. 10.3).



Figura 10.3: Strumenti di misura con lettura analogica e digitale.

Una volta che un segnale è disponibile in forma digitale è possibile elaborarlo tramite circuiti elettronici digitali (reti logiche) in grado di eseguire algoritmi logico-matematici.

Gli elaboratori elettronici (*computer*) sono oggi le più complesse e versatili reti di elaborazione digitale che siano state sviluppate.

10.1 Algebra di Boole

L'Algebra di Boole definisce tre operazioni *logiche* elementari che possono essere eseguite sulle variabili binarie. Tramite queste tre operazioni (in realtà si vedrà che una è addirittura ridondante) è possibile realizzare algoritmi logico-matematici come la somma, il confronto tra due numeri binari, ecc. Combinando in maniera opportuna molte operazioni elementari è possibile arrivare allo sviluppo di algoritmi sempre più complessi. In fin dei conti, il calcolo numerico insegna che la moltiplicazione si può ridurre ad una serie di addizioni, che un integrale oppure una funzione analitica possono essere calcolati con una serie di somme e moltiplicazioni sino al grado voluto di precisione, ecc.

Quindi se si è in grado di costruire circuiti elettronici in grado di effettuare le operazioni logiche elementari, allora è possibile costruire reti digitali contenenti tanti di questi circuiti elementari (spesso veramente tanti!) in grado di elaborare la soluzione di qualsiasi problema che possa essere espresso tramite un algoritmo logico-matematico.

10.1.1 Operazioni logiche elementari

Le tre operazioni logiche elementari sono convenzionalmente identificate dai termini inglesi *AND*, *OR* e *NOT* ed indicate nelle espressioni logiche con i simboli,

rispettivamente, \wedge, \vee, \sim . Le operazioni *AND* e *OR* agiscono su due variabili elementari (due bit) come operandi, dando un terzo bit come risultato, secondo le definizioni seguenti:

$C = A \wedge B$: la variabile C è 1 se A e B sono entrambe 1.

$C = A \vee B$: la variabile C è 1 se A oppure B è 1.

La terza operazione, *NOT*, agisce su un solo bit come operando e dà un bit come risultato:

$C = \sim A$: la variabile C è 1 se A è 0; la variabile C è zero se A è 1.

Un altro modo di descrivere le operazioni *AND*, *OR*, *NOT* è tramite le *tavole di verità*, riportate in fig. 10.4. Sotto ogni operazione è riportato anche il simbolo utilizzato negli schemi delle reti digitali per indicare la *porta*, cioè il circuito logico elementare, che effettua l'operazione.

A	B	$A \wedge B$	A	B	$A \vee B$	A		$\sim A$
0	0	0	0	0	0	0		1
0	1	0	0	1	1	1		0
1	0	0	1	0	1			
1	1	1	1	1	1			

The diagram shows three logic gate symbols below their respective truth tables:
AND: A symbol with two inputs (A and B) and one output (C). It consists of a rectangle with a diagonal line from top-left to bottom-right.
OR: A symbol with two inputs (A and B) and one output (C). It consists of a rectangle with a horizontal line from top-left to bottom-right.
NOT: A symbol with one input (A) and one output (C). It consists of a rectangle with a circle at the output end (C).

Figura 10.4: Descrizione delle operazioni logiche elementari tramite le *tavole di verità* e simboli con cui le operazioni vengono rappresentate negli schemi delle reti digitali.

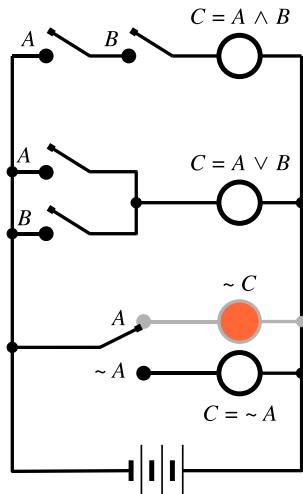
10.1.2 Circuiti logici elementari (*porte*)

I circuiti logici elementari possono essere realizzati utilizzando diverse *tecnologie*. Per esempio, in fig. 10.5 le operazioni *AND*, *OR*, *NOT* sono state realizzate mediante interruttori meccanici.

Nella descrizione delle operazioni contenuta nella figura è stata implicitamente utilizzata la corrispondenza $0 = \text{spento}$ e $1 = \text{acceso}$, indicata convenzionalmente come *logica positiva*. Anche se siamo portati istintivamente ad utilizzare questa associazione, la *logica negativa* opposta ($0 = \text{acceso}$ e $1 = \text{spento}$) è altrettanto legittima.

10.1.3 Le leggi di De Morgan

Le tavole di verità delle operazioni *AND* e *OR* interpretate secondo le due logiche sono riportate nella tabella 10.3.



AND: La lampadina C è accesa se l'interruttore A è chiuso **e** interruttore B è chiuso.

OR: La lampadina C è accesa se l'interruttore A è chiuso **oppure** interruttore B è chiuso.

NOT: La lampadina C è accesa se l'interruttore A **non** è chiuso.

Figura 10.5: Le operazioni logiche elementari effettuate tramite interruttori elettrici.

Operazione AND									
stato del circuito			logica positiva			logica negativa			
A	B	C	A	B	C	A	B	C	
spento	spento	spento	0	0	0	1	1	1	
spento	acceso	spento	0	1	0	1	0	1	
acceso	spento	spento	1	0	0	0	1	1	
acceso	acceso	acceso	1	1	1	0	0	0	

Operazione OR									
stato del circuito			logica positiva			logica negativa			
A	B	C	A	B	C	A	B	C	
spento	spento	spento	0	0	0	1	1	1	
spento	acceso	acceso	0	1	1	1	0	0	
acceso	spento	acceso	1	0	1	0	1	0	
acceso	acceso	acceso	1	1	1	0	0	0	

Tabella 10.3: Tavole di verità delle operazioni *AND* e *OR* interpretate secondo la logica positiva e negativa.

Si vede che una operazione *AND* secondo la logica positiva corrisponde ad una operazione *OR* quando i valori alle variabili vengono interpretati secondo la logica negativa e viceversa.

Queste equivalenze costituiscono le *Leggi di De Morgan*:

$$\sim A \wedge \sim B = \sim (A \vee B) \quad (10.1)$$

$$\sim A \vee \sim B = \sim (A \wedge B) \quad (10.2)$$

oppure

$$A \wedge B = \sim (\sim A \vee \sim B) \quad (10.3)$$

$$A \vee B = \sim (\sim A \wedge \sim B) \quad (10.4)$$

Quindi le due operazioni *AND* e *OR* non sono indipendenti. Qualunque funzione logica potrebbe essere ottenuta combinando soltanto *AND* e *NOT* oppure *OR* e *NOT* (fig. 10.6).

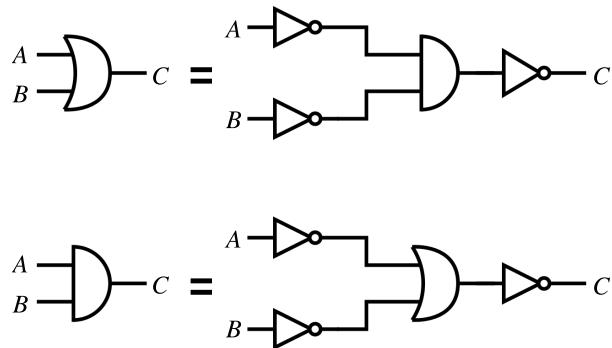


Figura 10.6: Leggi di De Morgan: l'operazione *OR* può essere ottenuta combinando *AND* e *NOT* e viceversa.

10.1.4 *NAND* e *NOR*

Le combinazioni delle due operazioni *AND-NOT* e *OR-NOT* che compaiono in fig. 10.6 sulla destra sono, come vedremo, le vere funzioni elementari delle reti digitali elettroniche, tanto da avere un proprio nome (*NAND* e *NOR*) ed un simbolo (fig. 10.7). In genere l'operazione *NOT* associata ad un'altra operazione viene riportata



Figura 10.7: Le operazioni elementari realizzate mediante circuiti elettronici digitali sono *NAND* e *NOR*.

negli schemi con il solo cerchietto invece che con il simbolo completo.

10.2 Il transistor come interruttore

Nei circuiti elettronici digitali sono i transistor, sia bjt che fet, che svolgono la funzione degli interruttori di fig. 10.5.

In fig. 10.8 è riportato uno schema di principio di un circuito *NOT* realizzato con un transistor bjt; a destra è riportata la funzione di trasferimento $V_{out}(V_{in})$.

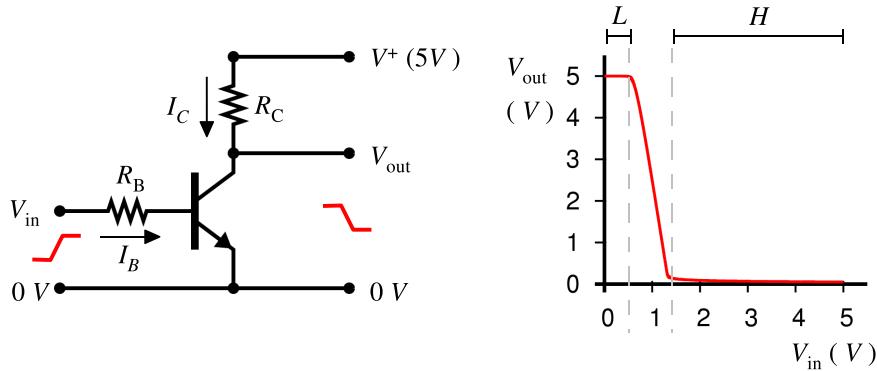


Figura 10.8: Circuito *NOT* realizzato con un transistor bjt.

Sino a che la tensione di ingresso V_{in} rimane nell'intervallo di valori indicato con L (*Low*, $0 \dots V_\gamma$), si ha corrente I_B (e quindi I_C) praticamente nulla; la tensione di uscita V_{out} rimane al valore V^+ ed il transistor è spento. Quando V_{in} supera V_γ , comincia a circolare una corrente di base $I_B = (V_{in} - V_\gamma)/R_B$ e una corrente di collettore $I_C = I_B \cdot h_{FE}$; la tensione $V_{out} = V^+ - I_C \cdot R_C$ scende rapidamente, il transistor attraversa la regione attiva ed infine entra in saturazione. Per tutti i valori di V_{in} compresi nell'intervallo H (*High*), V_{out} rimane al di sotto di 200 mV.

In definitiva il circuito *NOT* non è altro che un amplificatore fatto lavorare al di fuori della sua caratteristica lineare.

Dal grafico si vede bene come i due stati digitali 0 e 1 (L e H) non sono due valori di tensione ben definiti, ma piuttosto due intervalli, separati da una zona di transizione che viene solo attraversata rapidamente nel passaggio da uno stato logico all'altro.

È importante che tutti i valori possibili della tensione di uscita nei due stati logici siano riconosciuti in maniera non ambigua dal circuito di ingresso di un'altra porta logica con le stesse caratteristiche. Questo permette di realizzare *famiglie* di dispositivi logici che possono essere direttamente interconnessi tra di loro per ottenere funzioni logiche più complesse.

Ogni famiglia è caratterizzata da un proprio insieme di valori delle caratteristiche elettriche corrispondenti ai due stati logici.

10.3 Le famiglie logiche

Nella figura 10.9 è riportata la foto di una scheda, risalente ai primi anni '60 del secolo scorso, contenente una parte dell'unità di calcolo di un computer e realizzata a *componenti discreti*, cioè singoli transistor, resistenze, condensatori e diodi.

La scheda contiene complessivamente qualche decina di porte logiche elementari.

Con l'introduzione dei circuiti integrati è stato possibile costruire all'interno dello stesso *chip* di Silicio intere porte logiche e poi reti digitali via via più complesse, fino ad arrivare alla densità attuale di $10^6 \dots 10^9$ porte e oltre, cioè un intero computer, in un singolo circuito integrato.

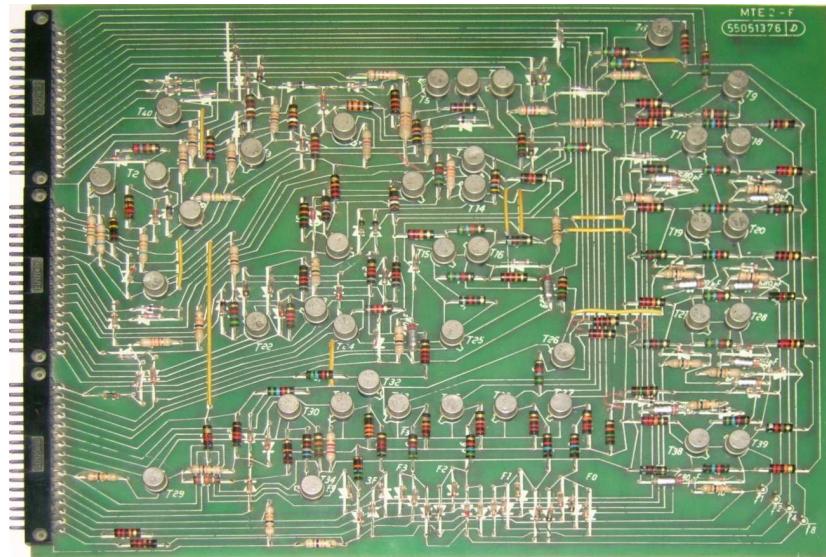


Figura 10.9: Parte di unità di calcolo realizzata a *componenti discreti*.

Nella tabella 10.4 è riportata l’evoluzione della complessità dei circuiti integrati digitali nel corso degli anni.

anno di riferimento	livello di integrazione		circuiti elementari per <i>chip</i>
1950		circuiti a componenti discreti	
1960	SSI	Small Scale integration	< 10
1966	MSI	Medium Scale integration	$10 \div 10^2$
1969	LSI	Large Scale integration	$10^2 \div 10^3$
1975	VLSI	Very Large Scale integration	$10^3 \div 10^6$
1990	ULSI	Ultra Large scale integration	$> 10^6$

Tabella 10.4: Evoluzione della complessità dei circuiti integrati digitali a partire dalla introduzione del transistor.

La foto in figura 10.10 riproduce l’intera unità di calcolo di un computer della serie PDP11, in produzione nei primi anni ’80, realizzata con circuiti integrati sia ad alta che a bassa integrazione.

L’evoluzione dei circuiti integrati ha portato nel corso degli anni allo sviluppo di diverse soluzioni circuituali e tecnologiche per la realizzazione delle porte logiche elementari *AND*, *OR*, *NOT* (o meglio *NAND*, *NOR*, *NOT*) ed una varietà di funzioni più o meno complesse ottenute dalla combinazione di più porte elementari.

Alcune delle principali *famiglie* di porte e funzioni logiche oggi in uso sono riportate nella tabella 10.5; si possono tutte raggruppare intorno a tre tecnologie principali, che fanno uso di transistor bipolari (famiglia *TTL* e sue discendenti e famiglia *ECL*), di transistor *mosfet* complementari (famiglia *CMOS* e sue discendenti) o misti bipolari e *mosfet* (famiglia *BiCmos* e discendenti).

Per ogni famiglia è riportato anche lo *stato evolutivo*, cioè la situazione riguardo allo sviluppo, affermazione nelle applicazioni e declino, riferiti all’anno 2010.

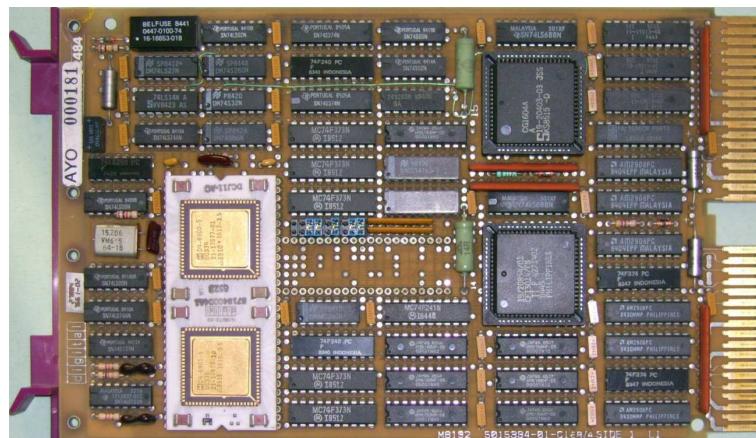


Figura 10.10: Unità di calcolo di un computer PDP11-23, realizzata a circuiti integrati sia a bassa che ad alta integrazione.

Famiglie logiche a transistor bipolar		
TTL	Transistor-Transistor Logic	obsoleta
S	Schottky Logic	estinta
LS	Low-Power Schottky Logic	in declino
AS	Advanced Schottky Logic	in declino
F	Fast Logic	in declino
ALS	Advanced Low-Power Schottky Logic	in declino
ECL	Emitter Coupled Logic	in declino
Famiglie logiche a transistor mosfet complementari		
C	CMOS Logic	estinta
CD4000	CMOS Logic	in declino
HC	High-Speed CMOS Logic	di uso corrente
HCT	High-Speed CMOS Logic	di uso corrente
AC	Advanced CMOS Logic	di uso corrente
ACT	Advanced CMOS Logic	di uso corrente
FCT	Fast CMOS Technology	di uso corrente
AHC	Advanced High-Speed CMOS	di uso corrente
AHCT	Advanced High-Speed CMOS	di uso corrente
ALVC	Advanced Low-Voltage CMOS Technology	in crescita
Famiglie logiche a tecnologia mista bipolari+mosfet		
BCT	BiCMOS Technology	di uso corrente
ABT	Advanced BiCMOS Technology	di uso corrente
LVT	Low-Voltage BiCMOS Technology	di uso corrente

Tabella 10.5: Alcune delle principali famiglie logiche e loro *stato evolutivo*, riferito all'anno 2010.

Tra una famiglia e l'altra di una stessa soluzione tecnologica vi sono differenze nelle prestazioni, come maggiore o minore velocità e consumi, differenti tensioni di alimentazione, diversa capacità di pilotaggio di altri circuiti, ecc. Ma le caratteristiche circuitali di base di tutte le famiglie di una stessa tecnologia sono essenzialmente le stesse (ad eccezione di *TTL* ed *ECL*).

Esistono altre famiglie non riportate nella tabella ma comunque in uso ed anche molte famiglie ormai completamente scomparse dalla produzione e dall'uso corrente e

di cui si può trovare traccia solo in apparecchiature e manuali sufficientemente vecchi.

10.3.1 Famiglia CMOS

Anche se storicamente la famiglia *CMOS* è nata diversi anni dopo l'affermazione della famiglia *TTL*, la sua enorme importanza attuale e la notevole semplicità di principio dei suoi circuiti suggeriscono di analizzarla per prima.

Il circuito base della tecnologia *CMOS* è l'*inverter* (porta *NOT*) riportato in fig. 10.11. Invece di avere un transistor ed una resistenza di carico, come nell'esempio

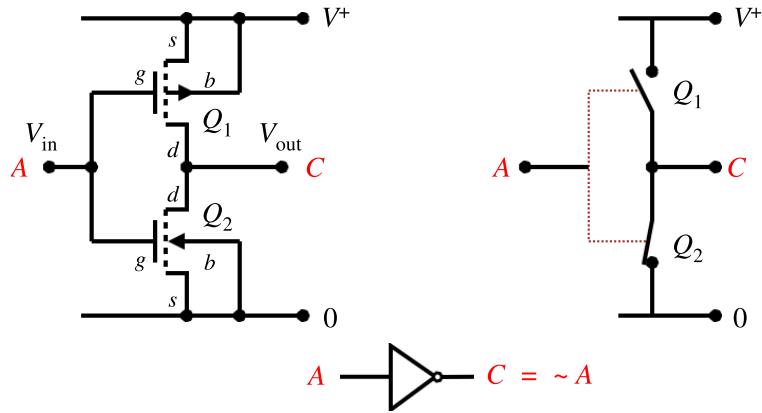


Figura 10.11: Porta *NOT* in tecnologia *CMOS*.

di fig. 10.8, il circuito è composto da due transistor *mosfet* complementari (Q_1 è un *pmos*, Q_2 è un *nmos*) che lavorano in maniera simmetrica, in modo che ciascuno faccia da carico all'altro. Da questo il nome dato alla famiglia: *CMOS* = *Complementary MOSfet*.

Normalmente i due transistor sono realizzati in modo da avere caratteristiche elettriche eguali, a parte gli opportuni segni. Quindi, ad esempio, $V_{th1} = -V_{th2}$ ecc. In queste condizioni è facile vedere che quando Q_1 è spento Q_2 è acceso e viceversa. I due transistor si comportano quindi come i due interruttori accoppiati e contrapposti in fig. 10.11 a destra, per cui l'uscita C si trova collegata alternativamente con la massa attraverso Q_2 , quando l'ingresso A è Alto ($V_{in} = V^+$) o con l'alimentazione V^+ attraverso Q_1 quando l'ingresso A è Basso ($V_{in} = 0$).

Questo circuito ha diversi vantaggi rispetto all'esempio di principio di fig. 10.8, che farebbe uso di un transistor e di una resistenza di carico:

- In condizioni statiche la dissipazione interna di energia è idealmente nulla, sia con l'uscita Alta che con l'uscita Bassa, in quanto uno dei due transistor è sempre spento, con resistenza nominalmente infinita, mentre l'altro è acceso, con resistenza molto bassa, nominalmente nulla. La dissipazione di energia si ha solo per la corrente che fluisce in un eventuale carico. Nel circuito di fig. 10.8 solo nello stato di uscita Alta (transistor spento) si ha dissipazione nulla; nell'altro stato viene dissipata una potenza $P \simeq V^{+2}/R_C$.
- L' ingresso di un circuito *CMOS* equivale ad un condensatore; in condizioni statiche non si ha passaggio di corrente e dissipazione di energia all'ingresso.

- I livelli elettrici relativi agli stati *Alto* e *Basso* sono simmetrici.

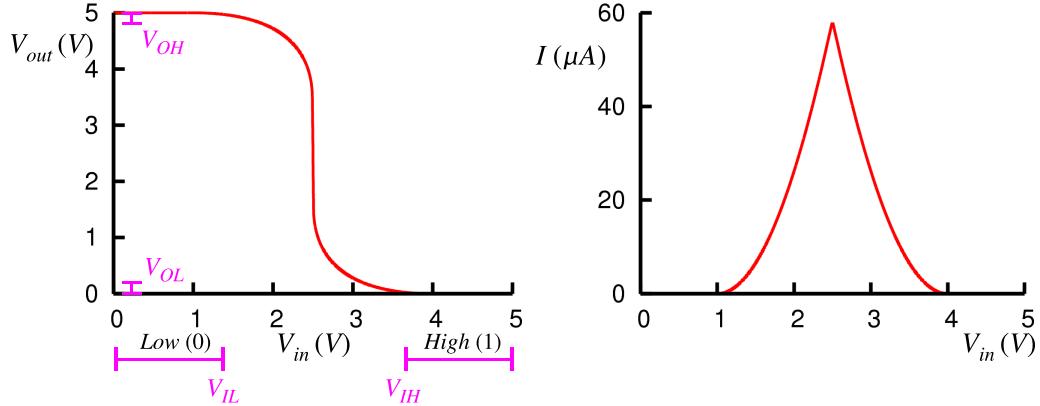


Figura 10.12: A sinistra: relazione tra tensione di ingresso V_{in} e tensione di uscita V_{out} nella porta *NOT* in tecnologia *CMOS*. A destra: corrente attraverso i transistor durante la commutazione.

La figura 10.12 riporta la relazione tra tensione V_{in} all'ingresso A e tensione V_{out} all'uscita C per una porta con due transistor complementari con $V_{th} = \pm 1.0\text{ V}$, $K_n = 50\text{ }\mu\text{A/V}^2$, $\lambda = 0.01\text{ V}^{-1}$ e tensione di alimentazione $V^+ = 5\text{ V}$.

I valori di tensione indicati in viola sono:

- | | |
|----------|--|
| V_{IL} | = massima tensione di ingresso considerata come 0 (<i>Low</i>) |
| V_{IH} | = minima tensione di ingresso considerata come 1 (<i>High</i>) |
| V_{OL} | = massima tensione di uscita quando l'uscita è 0 (<i>Low</i>) |
| V_{OH} | = minima tensione di uscita quando l'uscita è 1 (<i>High</i>) |

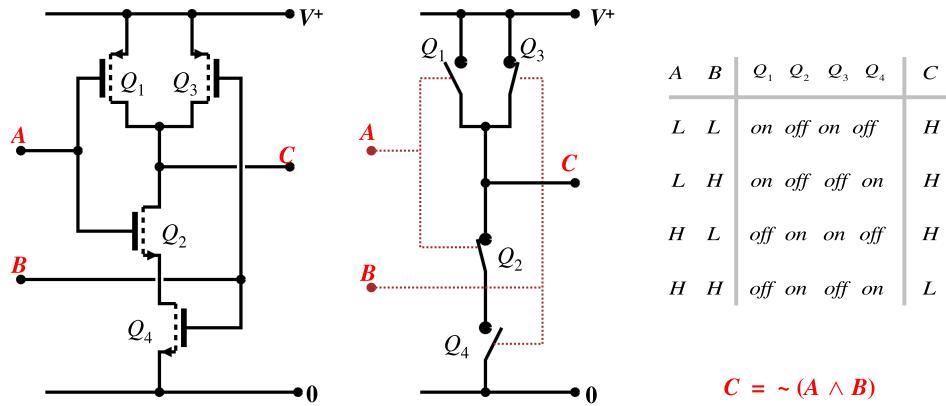
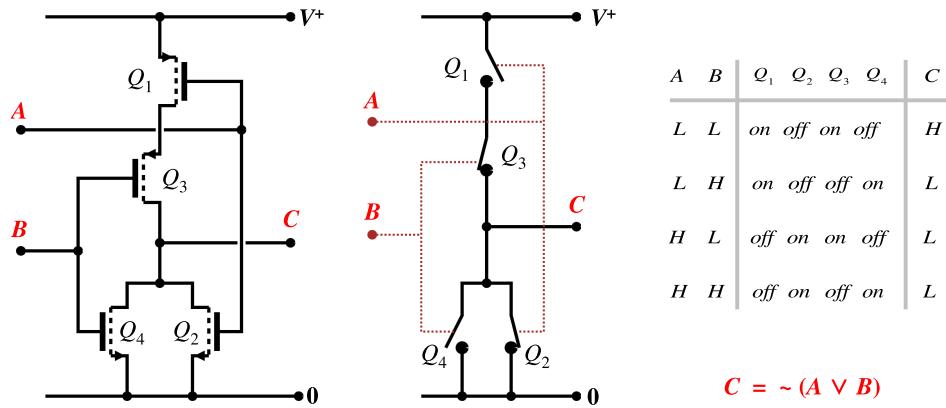
Gli intervalli di tensione $0 \dots V_{IL}$ e $V_{IH} \dots V^+$ corrispondono ai valori di ingresso riconosciuti rispettivamente come 0 e 1, mentre gli intervalli $0 \dots V_{OL}$ e $V_{OH} \dots V^+$ rappresentano i valori *garantiti* per la tensione di uscita negli stati 0 e 1.

Le differenze di tensione $V_{IL} - V_{OL}$ e $V_{IH} - V_{OH}$ sono il *margine di rumore*: l'ampiezza minima di un segnale esterno di disturbo necessaria per falsare il riconoscimento di un livello alto in basso o viceversa e generare un errore nelle operazioni della rete.

10.3.2 Porta *NAND* e *NOR* in tecnologia CMOS

Raddoppiando i transistor-interruttori della fig. 10.11 come riportato negli schemi di fig. 10.13 e 10.14 dalla porta *NOT* si ottiene una porta *NAND* oppure una porta *NOR*.

Le tabelle a fianco degli schemi elettrici descrivono il funzionamento dei circuiti, riportando oltre alla tavola di verità della porta anche lo stato acceso/spento (*on/off*) di tutti i transistor in corrispondenza degli stati alto/basso (*H/L* o *1/0*) degli ingressi. Per *alto* in questa tabella ed in tutti i paragrafi che seguono si intende un valore di tensione positivo (nominalmente V^+), per basso si intende il valore nominale 0 V.

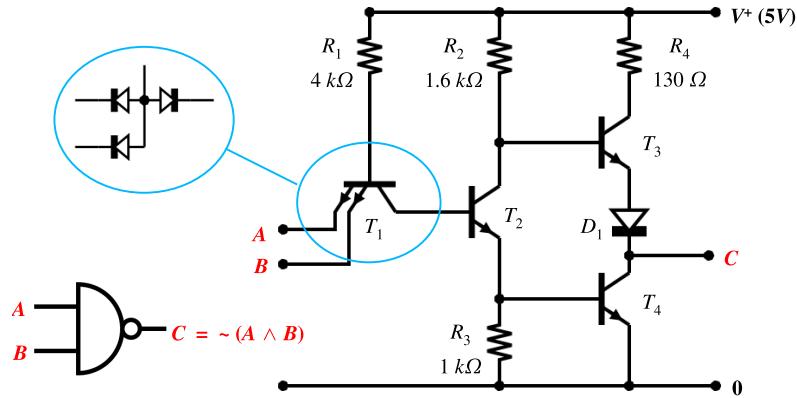
Figura 10.13: Porta *NAND* della famiglia CMOS.Figura 10.14: Porta *NOR* della famiglia CMOS.

10.3.3 Famiglia TTL

In figura 10.15 è riportato lo schema della porta *NAND*, elemento base dei dispositivi logici realizzati con la tecnologia *TTL*.

L'aspetto del circuito è reso intricato dalla presenza di un transistor con doppio emettitore sugli ingressi A e B , in una configurazione che apparentemente non corrisponde a nessuna di quelle viste nel cap. 7. Conviene inizialmente considerare questo transistor come una rete di tre diodi connessi come indicato in alto a sinistra. Il transistor bjt è formato da due diodi contrapposti, come visto nel par. 7.1; realizzando la regione di emettitore come due regioni separate a contatto con un'unica regione di base si ottiene la configurazione di fig. 10.15.

Supponiamo inizialmente che i due ingressi A e B si trovino entrambi nello stato alto ($V_A = V_B = +5 V$). Osservando il circuito è facile vedere che le due giunzioni base-emettitore di T_1 si troveranno contropolarizzate, mentre la giunzione base collettore di T_1 e le giunzioni base-emettitore di T_2 e T_4 si troveranno tutte e tre collegate in serie e direttamente polarizzate dalla tensione di alimentazione V^+ tramite la resistenza R_1 . La corrente attraverso R_1 , tra base e collettore di T_1 e nella base

Figura 10.15: Porta *NAND* in tecnologia *TTL*.

di T_2 è quindi $I_{R_1} = (V^+ - 3V_\gamma)/R_1$. Quindi i transistor T_2 e T_4 saranno entrambi accesi. I componenti del circuito, in particolare la resistenza R_2 , sono dimensionati in modo che in queste condizioni il transistor T_2 si trovi non solo acceso, ma anche in saturazione. Quindi si avrà:

$$V_{E_2} = V_{BE_4} = V_\gamma \simeq 0.65 \text{ V} \quad (10.5)$$

$$V_{C_2} = V_\gamma + V_{CE_{sat}} = 0.7 \dots 0.9 \text{ V} \quad (10.6)$$

Il transistor T_3 invece sarà spento. Infatti, a causa del diodo D_1 presente sul suo emettitore, per arrivare all'accensione è necessaria una tensione $V_{B_3} = V_{C_2}$ almeno superiore a $2V_\gamma \simeq 1.3 \text{ V}$, mentre si è visto che V_{C_2} non arriva ad 1 V. Il transistor T_3 spento si comporta come una resistenza di carico di valore infinito per T_4 , che risulta quindi in saturazione. In definitiva all'uscita si avrà una tensione $V_C = V_{CE_{SAT}} \leq 0.4 \text{ V}$.

Collegando a massa, o ad un livello di tensione sufficientemente basso, uno o tutti e due gli ingressi A e B , la giunzione base emettitore di T_1 si viene a trovare in polarizzazione diretta e quindi $V_{B_1} \simeq V_\gamma$. Questa tensione non è più sufficiente a tenere in conduzione le tre giunzioni in serie V_{BC_1} , V_{BE_2} e V_{BE_4} per cui T_2 e T_4 si troveranno spenti. In queste condizioni il transistor T_3 si trova in configurazione *emitter follower* con la base collegata a V^+ tramite R_2 . La tensione all'uscita sarà $V_C = V^+ - R_2 \cdot I_{B_3} - 2V_\gamma \simeq 3.7 \text{ V}$.

Nella figura 10.16 è riportata la relazione tra tensione di ingresso e tensione di uscita per una porta *NAND* della famiglia *TTL*, insieme ai valori di tensione in alcuni nodi interni del circuito. La curva in rosso in alto rappresenta la corrente di alimentazione fornita dal generatore V^+ durante la commutazione. Per un breve intervallo della tensione di ingresso V_{in} intorno a $\simeq 1.6 \text{ V}$ i due transistor T_3 e T_4 dello stadio di uscita si trovano entrambi in forte conduzione; la resistenza R_4 ha lo scopo di limitare la corrente durante questo transitorio.

Anche questo circuito è dotato di uno stadio di uscita contenente due transistor che lavorano in maniera complementare, come visto per le porte *CMOS*. In questo caso, però, i due transistor sono della stessa polarità, *npn* e non complementari. Que-

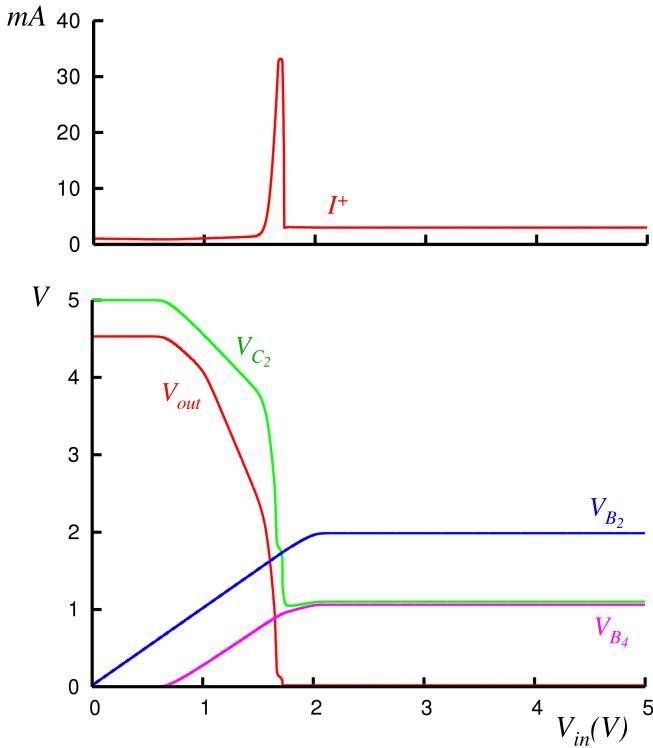


Figura 10.16: Grafico in basso: tensioni in alcuni punti del circuito della porta TTL di fig. 10.15; grafico in alto: corrente di alimentazione.

sta configurazione è denominata *totem pole*. Negli istanti immediatamente successivi alla commutazione degli ingressi A e B da 1 a 0, a causa della capacità di diffusione della base di T_2 che rimane carica per un breve istante ad una tensione positiva,

	<i>TTL</i>	<i>TTL-LS</i>	<i>ACT</i>	<i>AHC</i>		
V^+	5	5	5	5	V	tensione di alimentazione
V_{IL}	0.8	0.8	0.8	1.5	V	massima tensione di ingresso considerata <i>low</i>
V_{IH}	2.0	2.0	2.0	3.5	V	minima tensione di ingresso considerata <i>high</i>
V_{OL}	0.4	0.5	0.4	0.5	V	massima tensione di uscita nello stato <i>low</i>
V_{OH}	2.4	2.7	3.7	4.5	V	minima tensione di uscita nello stato <i>high</i>
I_{IL}	-1.6	-0.4		-0.001	mA	corrente di ingresso nello stato <i>low</i>
I_{IH}	0.04	0.02		0.001	mA	corrente di ingresso nello stato <i>high</i>
I_{OL}	16	8	24	8	mA	corrente di uscita massima nello stato <i>low</i>
I_{OH}	-0.4	-0.4	-24	-8	mA	corrente di uscita massima nello stato <i>high</i>
<i>fan-out</i>	10	20	40	40		ingressi equivalenti collegabili ad una uscita
NM_L	0.4	0.3	0.4	1	V	margine di rumore a livello <i>low</i>
NM_H	0.4	0.7	1.7	1	V	margine di rumore a livello <i>high</i>

Tabella 10.6: Principali caratteristiche elettriche di alcune famiglie CMOS e TTL. La famiglia ACT, di tipo CMOS, ha livelli elettrici compatibili con le famiglie TTL.

il transistor T_1 si viene a trovare in regione attiva e funziona effettivamente come transistor invece che semplicemente come rete di diodi. Questo rende più veloce la scarica della capacità di base di T_2 e quindi riduce i tempi di commutazione.

Nella tabella 10.6 sono riportati i valori caratteristici dei parametri elettrici per alcune famiglie di dispositivi *CMOS* e *TTL*.

Fan-out e fan-in

Il *fan-out* (ventaglio di uscita) è il numero di ingressi standard che possono essere comandati dall'uscita standard della stessa famiglia. Per le famiglie *TTL* e *TTL-LS* il *fan-out* è tipicamente 10 e 20 rispettivamente; per i dispositivi *CMOS* arriva a 40. In ogni famiglia esistono elementi in grado di fornire *fan-out* maggiori, come ad esempio i circuiti *buffer* (circuiti tampone, ripetitori del segnale, con l'unica funzione di aumentare la disponibilità di pilotaggio di una uscita).

Analogamente si definisce un *fan-in* (ventaglio di ingresso). Ogni porta logica standard di ogni famiglia ha *fan-in* pari a 1 per definizione. Circuiti complessi possono avere qualche ingresso che fornisce segnale a più di una porta logica. In questo caso si può avere *fan-in* maggiore di 1. Di solito, per *fan-in* superiori a 2 il costruttore inserisce all'ingresso interessato un *buffer*, in modo da ritornare al valore standard di 1 (con la penalizzazione di introdurre un ritardo nel segnale).

10.4 Le commutazioni nei circuiti digitali

Durante le commutazioni nei circuiti sia *CMOS* che *TTL* si hanno brevi impulsi di corrente attraverso i transistor, come si può vedere nel grafico a destra in fig. 10.12 ed in quello in alto in fig. 10.16. Ad ognuno di questi impulsi di corrente, caratteristici di ogni famiglia di dispositivi logici come forma e durata, corrisponde una dissipazione all'interno del dispositivo di una certa quantità di energia $E_i = QV^+$ data dal prodotto della carica totale Q dell'impulso (integrale della corrente nel tempo) per la tensione di alimentazione V^+ . Se f è la frequenza media di commutazioni di una porta, si avrà all'interno della porta una potenza media dissipata $\bar{P} = E_i \cdot f$ in aggiunta alla potenza comunque dissipata in condizioni statiche.

Nel caso dei dispositivi *CMOS* c'è da tenere conto di un ulteriore effetto: l'ingresso di una porta *CMOS* equivale elettricamente ad un condensatore che ad ogni transizione da basso ad alto si carica da 0 a V^+ per poi scaricarsi nella transizione da alto a basso.

Al termine della carica, l'energia immagazzinata nel condensatore di ingresso C_i è $E_C = QV^+/2$, mentre l'energia erogata complessivamente dal generatore per effettuare la commutazione è $E_+ = QV^+$, con $Q = C_iV^+$. La differenza tra queste due quantità, cioè $QV^+/2$, viene dissipata in calore all'interno della porta che pilota l'ingresso. Durante la successiva transizione dal livello alto al livello basso l'energia rimasta immagazzinata nel condensatore C_i viene a sua volta dissipata in calore all'interno della porta. In totale, indicando nuovamente con f la frequenza media dei cicli di commutazioni, si ha un contributo ulteriore di dissipazione $\bar{P}_C = C_iV^{+2}f$.

In definitiva, nei dispositivi *TTL* la generazione interna di calore aumenta, rispetto al valore che si ha in condizioni statiche, al crescere della frequenza media di commutazioni. Nei dispositivi *CMOS* invece la dissipazione è praticamente nulla

in condizioni statiche e da zero cresce linearmente con la frequenza f in condizioni operative.

Aumentando la frequenza di lavoro di un dispositivo oltre i limiti di sicurezza indicati dal costruttore si possono a volte ottenere prestazioni migliori, a scapito però di un più rapido deterioramento del dispositivo. Al limite, si può anche arrivare alla sua rapida distruzione per surriscaldamento.

C'e' poi un altro effetto importante prodotto dalle commutazioni veloci di corrente nei circuiti digitali. Un picco di corrente come quello di fig. 10.12 o quello di fig. 10.16, anche se riguarda correnti del valore dalla frazione di mA a qualche decina di mA , può contenere variazioni con velocità fino a $10^4 \div 10^6 A/s$. L'ordine di grandezza dell'induttanza di un conduttore rettilineo del tipo di quelli utilizzati nei montaggi dei circuiti digitali, o di una pista di circuito stampato, è di $1 nH/cm$.

Commutazioni simultanee di più porte collegate a conduttori di alimentazione lunghi molti cm può portare alla generazione per autoinduzione di *f.e.m.* dell'ordine della frazione di V , potenzialmente in grado di eccedere il margine di rumore dei dispositivi e generare errori nel riconoscimento del valore logico dei segnali.

Per questo motivo è obbligo tassativo nella realizzazione dei circuiti digitali disporre lungo i conduttori di alimentazione opportuni condensatori di *by-pass* (filtraggio) a distanze non superiori a pochi cm , con lo scopo di limitare la propagazione delle fluttuazioni veloci di corrente. La fig. 10.17 riporta un dettaglio della fig. 10.10, in cui si vede come accanto ad ogni *chip* si trovi un condensatore di *by-pass*.

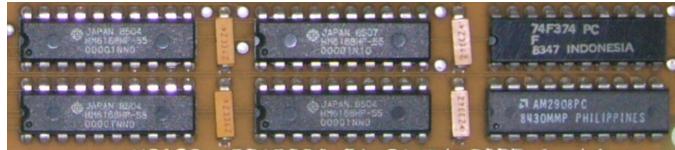


Figura 10.17: Dettaglio della scheda del computer PDP11-23 di fig. 10.10, in cui si vedono i condensatori di *by-pass* accanto ai circuiti integrati digitali.

10.5 La serie “74”

I circuiti integrati realizzati con la tecnologia *TTL* hanno avuto una larghissima diffusione tra la fine degli anni '60 e l'inizio degli anni '70 del secolo scorso, con l'introduzione di una famiglia identificata dalla sigla *SN74*. I componenti di questa famiglia erano circuiti integrati a bassa e a media integrazione (*SSI* e *MSI*); nelle fig. 10.18 - 10.20 sono riportati gli schemi logici e le connessioni elettriche di alcuni dei più utilizzati tra i componenti della famiglia.

I primi membri della serie contenevano soltanto un certo numero di porte logiche elementari da connettere esternamente per realizzare le funzioni desiderate (quattro porte *NAND* o *NOR*, sei porte *NOT* ecc.). La famiglia si è poi arricchita con il passare del tempo di *chip* contenenti funzioni sempre più complesse, alcune delle quali sono descritte nei paragrafi seguenti.

La serie 74 ha avuto una larghissima diffusione nella produzione di apparecchiature elettroniche di tipo commerciale, industriale e scientifico, diventando praticamente uno standard *de facto*.

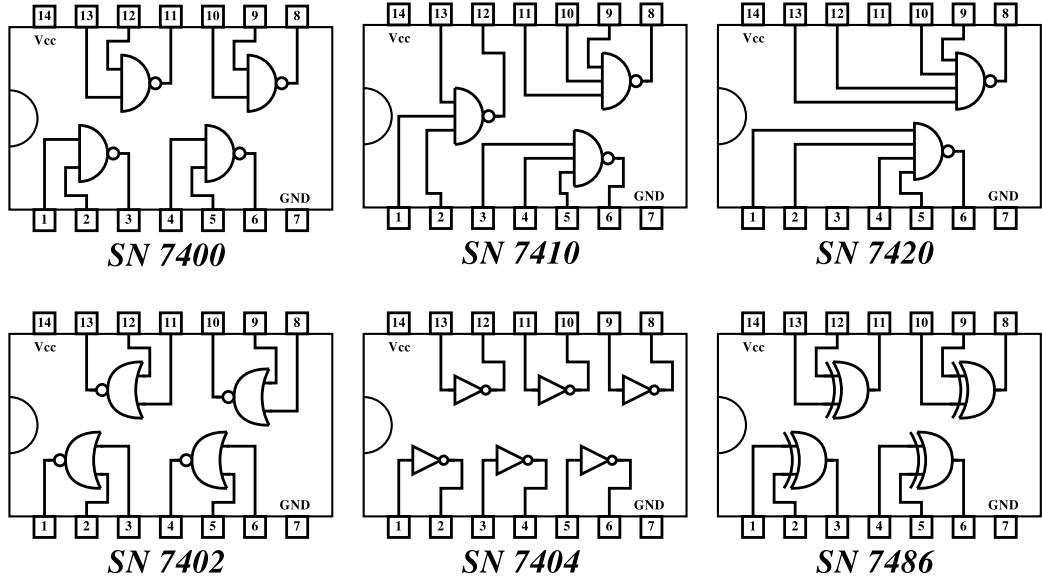


Figura 10.18: *Chips* della serie *SN74* contenenti porte logiche elementari: *NAND* a due, tre e quattro ingressi, *NOR*, *NOT* ed *Exclusive-OR* (par. 10.6.1).

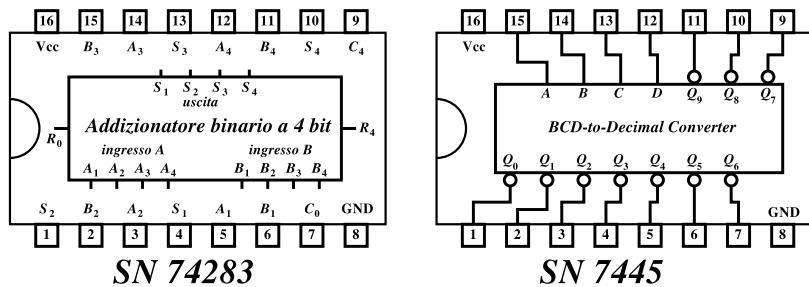


Figura 10.19: Esempi di *chips* della serie *SN74* contenenti funzioni combinatorie: addizionatore binario a 4 bit (par. 10.6.4) e convertitore BCD - decimale (par. 10.6.5).

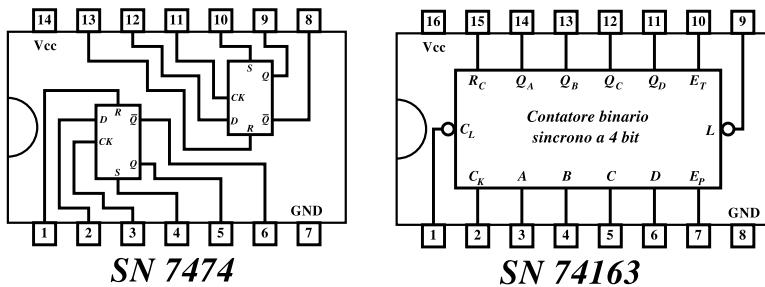


Figura 10.20: *Chips* della serie *SN74* contenenti funzioni sequenziali: flip-flop doppio di tipo *D* (par. 10.9.3) e contatore binario sincrono a 4 bit (par. 10.10.4).

Il successo è stato tale che le famiglie logiche che si sono sviluppate successivamente, alcune delle quali sono riportate nella tab. 10.5, hanno cercato di mantenere

il più possibile la compatibilità con la originaria serie *TTL 74*. Così, per esempio, accanto al *chip SN7400* contenente quattro porte *NAND* si è avuto il *chip SN74LS00, SN74C00, SN74AC00, SN74HC00, SN74ACT00* ecc., ognuno contenente le stesse funzioni logiche (quattro porte *NAND*) con lo stesso aspetto esterno, gli stessi collegamenti ai piedini ed offrendo quando possibile anche la compatibilità dei livelli elettrici.

Una rete logica disegnata per essere realizzata con dispositivi *SN74* poteva, con minimi o nessun aggiustamento, essere realizzata sostituendo tutti i dispositivi con quelli di un'altra famiglia, ad esempio la *SN74C* o la *SN74HCT*, sfruttando i vantaggi che l'evoluzione tecnologica aveva portato con la nuova famiglia.

Oggi la realizzazione di reti logiche complesse utilizzando i *chip* della serie *74* è decisamente obsoleta ed è stata quasi completamente sostituita dall'uso di *chip* di tipo *VLSI* e *ULSI* programmabili, che possono arrivare a contenere da milioni fino a centinaia di milioni di porte logiche elementari, con evidenti vantaggi di spazio, velocità,¹ consumo e costo.

Non è più necessario interconnettere tra loro mediante fili o piste di circuito stampato i vari elementi delle reti logiche, ma una volta impostato lo schema complessivo, in modo grafico o mediante opportuni linguaggi di programmazione, è possibile realizzare la rete trasferendo l'informazione all'interno del *chip* e creando così le connessioni necessarie.

Tuttavia il successo della serie *SN74* è stato tale che nella programmazione dei dispositivi *VLSI* e *ULSI* le funzioni di base sono ancora identificate con i nomi che avevano nella serie *SN74* e tra i microchip contenenti una singola porta logica ed utilizzati come *glue logic* (funzioni logiche di raccordo), una porta *NAND* è indicata come 1/4 di *74xx00*.

10.6 Logica combinatoria

Si definisce *logica combinatoria* quella realizzata con reti digitali in cui i valori delle variabili di uscita dipendono esclusivamente dai valori istantanei delle variabili di ingresso, a parte ovviamente i ritardi dovuti ai tempi di propagazione dei segnali ed ai tempi di risposta finiti dei circuiti. Per contro nella *logica sequenziale*, descritta nel paragrafo 10.8 e successivi, le variabili di uscita dipendono non solo dai valori istantanei delle variabili di ingresso, ma anche dalla storia precedente del circuito: la rete sequenziale è dotata di *memoria*.

10.6.1 Exclusive-OR (*XOR*)

La definizione di operazione *OR* stabilisce che l'uscita è 1 se uno dei due ingressi, o anche tutti e due, hanno valore 1. Esiste anche una definizione più stretta dell'operazione, che stabilisce che l'uscita è 1 quando uno o l'altro degli ingressi è 1, ma non tutti e due insieme. In questa versione l'operazione si chiama *exclusive-OR* (*OR esclusivo*) e viene indicata con il termine *XOR*. In fig. 10.21 è riportata una rete combinatoria che esegue l'operazione *XOR*, insieme con la tavola di verità ed il simbolo utilizzato negli schemi logici.

¹Il trasferimento di un segnale da un *chip* ad un altro è più lento e richiede più energia del trasferimento da una porta ad un'altra all'interno dello stesso *chip*.

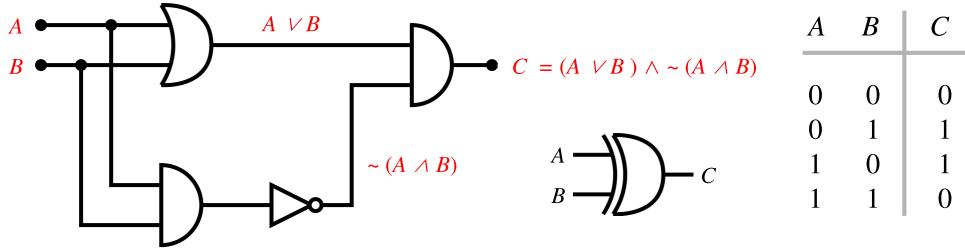


Figura 10.21: Operazione *exclusive-OR* (*XOR*). A sinistra: esempio di rete combinatoria per effettuare l'operazione; a destra: tavola di verità; al centro: simbolo con cui l'operazione viene rappresentata negli schemi.

Ogni funzione logica può essere ottenuta usando reti diverse, aventi lo scopo di ottimizzare, ad esempio, la velocità di risposta oppure il numero delle porte logiche utilizzate o la loro varietà, ecc. La fig. 10.22 riporta un circuito alternativo per la funzione *XOR* che fa uso solo di porte di tipo *NAND*.

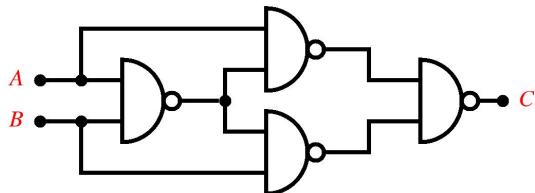


Figura 10.22: Funzione *XOR* realizzata utilizzando solo porte logiche di tipo *NAND*.

10.6.2 Semiaddizionatore (*Half adder*)

L'operazione *XOR*, con una piccola modifica, permette di realizzare un addizionatore ad un bit, cioè una rete combinatoria che ha due ingressi e due uscite e fornisce sulle uscite in codice binario il valore della somma aritmetica dei valori corrispondenti ai due bit di ingresso. In pratica esegue il calcolo $0 + 0 = 0$, $0 + 1 = 1$, $1 + 0 = 1$, $1 + 1 = 0$ con riporto di uno.

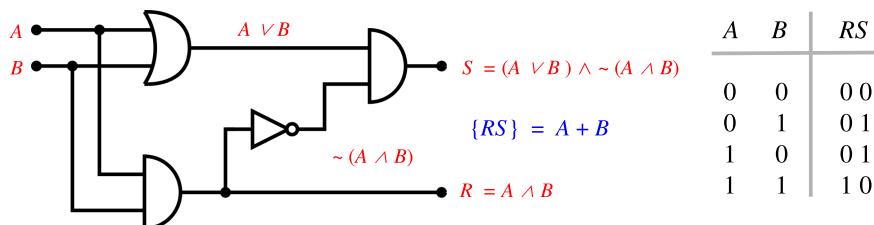


Figura 10.23: Circuito semi-addizionatore. Il codice binario formato dai due bit $\{RS\}$ rappresenta la somma aritmetica dei valori corrispondenti ai due bit A e B .

In questa forma il circuito prende il nome di semi-addizionatore (*half adder*). Infatti volendo effettuare la somma di numeri binari di lunghezza superiore ad un bit è necessario prevedere per ogni addizionatore un terzo ingresso per il riporto dall'addizionatore del bit immediatamente meno significativo.

10.6.3 Addizionatore (*Full adder*)

Un addizionatore completo può essere costruito mediante due semi-addizionatori ed una porta *OR* (fig 10.24). Ciascuno dei due rettangoli nella figura rappresenta un semi-addizionatore ($\frac{1}{2} \Sigma$) con i due ingressi A e B e le due uscite R e S come in fig. 10.23.

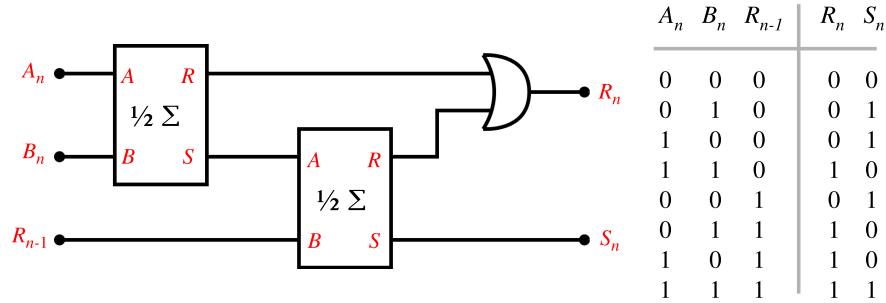


Figura 10.24: Circuito addizionatore (*full adder*). Con A_n e B_n si indicano i bit n -esimi di due numeri binari A e B ; R_{n-1} è il riporto che viene dal circuito addizionatore del bit immediatamente meno significativo; S_n è il bit n -esimo della somma e R_n il riporto verso il bit immediatamente più significativo. Sulla destra la tavola di verità.

Osservando la tavola di verità si può vedere come i due bit R_n ed S_n , se letti come numero binario, contengano la somma aritmetica degli 1 presenti sui tre ingressi A , B ed R_{n-1} .

10.6.4 Addizionatore a 4 bit

Combinando insieme N circuiti addizionatori come quello di fig. 10.24 è possibile realizzare una rete combinatoria in grado di effettuare l'operazione di *somma aritmetica* di due variabili numeriche di N bit. Nella figura 10.25 è riportato un esempio per $N = 4$.

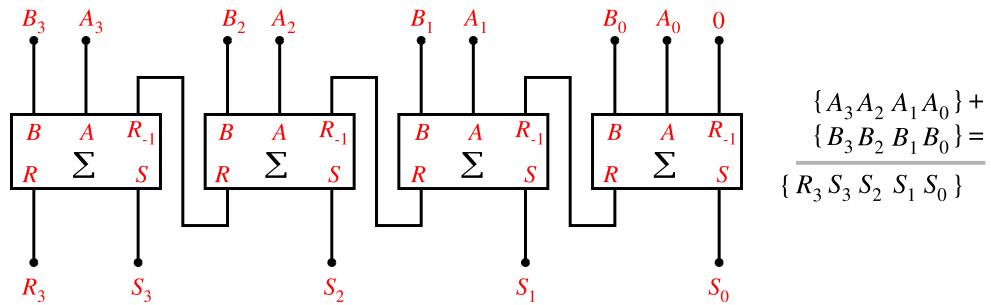


Figura 10.25: Circuito addizionatore a 4 bit.

Il circuito è composto da una catena di quattro addizionatori *full adder* (Σ) da un bit. Ogni singolo addizionatore opera su un bit della variabile $A = \{A_3 A_2 A_1 A_0\}$, sul bit corrispondente della variabile $B = \{B_3 B_2 B_1 B_0\}$ e sul riporto che proviene dal bit alla sua destra, cioè immediatamente meno significativo. Ogni sezione del

circuito genera un bit S_i della variabile somma $S = \{S_3\ S_2\ S_1\ S_0\}$ ed il riporto R_i per la sezione alla sua sinistra, cioè per il bit immediatamente più significativo.

L'ultimo riporto (R_3) costituisce il bit più significativo del risultato (in qualsiasi sistema di numerazione, la somma di due numeri di N cifre può dare un risultato fino a $N + 1$ cifre).

10.6.5 Convertitori di codice

Il circuito in fig. 10.26 è un convertitore da codice *BCD* a codice decimale. Il codice *BCD* (*Binary Coded Decimal*) è utilizzato per rappresentare in codice binario le cifre del sistema di numerazione decimale. Poiché con N bit è possibile avere 2^N combinazioni diverse, per codificare 10 elementi è necessario disporre di $N = \log_2 10 = 4$ bit (arrotondati per eccesso). Delle $2^4 = 16$ combinazioni possibili, dieci sono utilizzate per indicare le cifre da 0 a 9, le altre sei sono ridondanti.

Il circuito ha un ingresso a 4 bit ($B_3 \dots B_0$) e dieci uscite. Solo una delle dieci uscite si può trovare nello stato 1, come indicato dalla tavola a sinistra nella figura, selezionando quella delle dieci cifre indicata dal codice binario di ingresso.

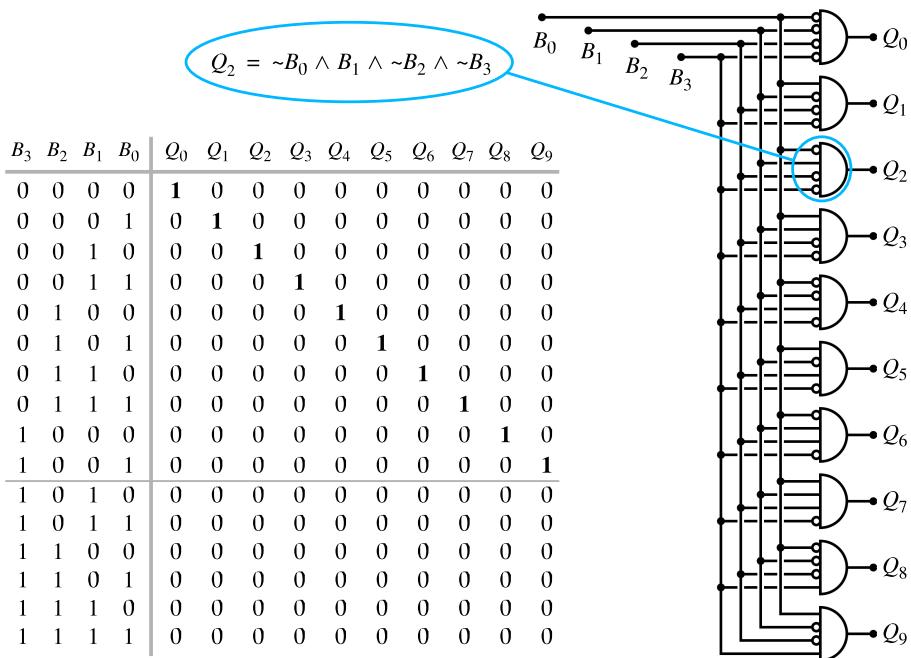


Figura 10.26: Convertitore da codice BCD a codice decimale.

10.7 Tempi di commutazione

La propagazione dei segnali tra ingressi e uscite di una rete combinatoria non è istantanea, ma richiede un tempo finito, dovuto principalmente alla velocità di risposta dei transistor delle porte che, per quanto elevata, è comunque sempre limitata. Per i dispositivi più veloci, anche il tempo di propagazione dei segnali lungo i collega-

menti può diventare significativo (un collegamento della lunghezza di 2 cm introduce un ritardo di circa 0.1 ns).

Ogni porta di ogni famiglia logica ha i suoi tempi caratteristici di commutazione. Nella figura 10.27 sono riportati i tempi caratteristici della porta *NAND* per alcune famiglie *TTL* e *CMOS*. Il tempo di commutazione non è necessariamente lo stesso nelle transizioni da livello alto a livello basso e viceversa. Nella figura sono definiti due tempi: un tempo t_{pHL} , corrispondente alla commutazione $H \rightarrow L$ dell'uscita in seguito al passaggio $L \rightarrow H$ dell'ingresso ed un tempo t_{pLH} corrispondente alla transizione opposta. Dalla tabella si vede che per alcune famiglie le differenze tra questi due tempi sono considerevoli.

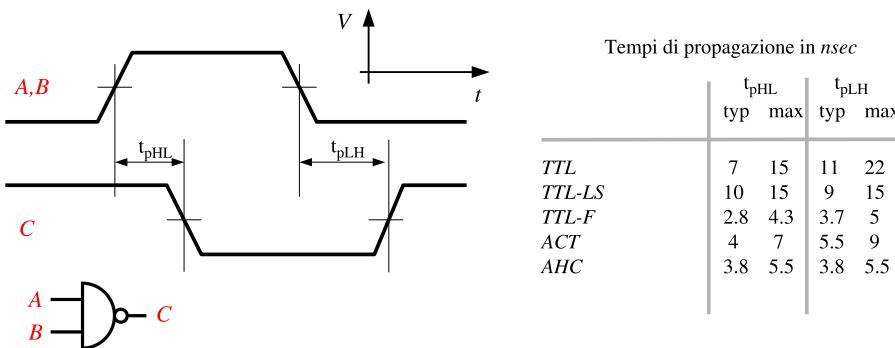


Figura 10.27: Definizione dei tempi di propagazione dei segnali da ingressi A e B ad uscita C attraverso una porta *NAND* e valori numerici per alcune famiglie *TTL* e *CMOS*.

10.8 Logica sequenziale

Anche nelle reti logiche è possibile introdurre la reazione (feedback), come nelle reti analogiche. La figura 10.28 riporta un esempio di reazione negativa (a sinistra) e di reazione positiva (a destra).

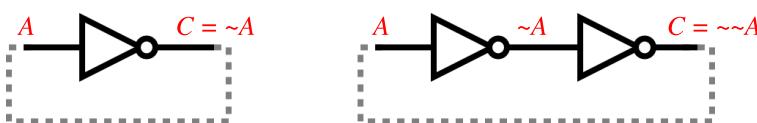


Figura 10.28: Due circuiti digitali con anello di reazione: negativa a sinistra, positiva a destra.

Da un punto di vista logico, chiudendo l'anello nel circuito a sinistra, si pone $A = C$; visto che per una porta *NOT* vale la relazione $C = \sim A$, si ottiene la relazione $A = \sim \sim A$, che è una contraddizione.

Il grafico in fig. 10.29 illustra quello che succede nel circuito. Nella figura è stata sovrapposta alla relazione di trasferimento tra ingresso e uscita della porta *NOT*, già vista in fig. 10.12, la relazione $V_{out} = V_{in}$ stabilita dal collegamento esterno. Il punto di incrocio tra le due curve corrisponde alla soluzione della equazione $V_{out}(V_{in}) = V_{in}$

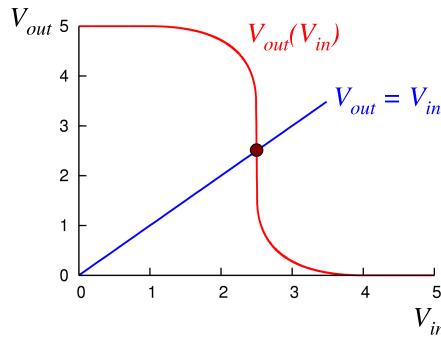


Figura 10.29: Una porta *NOT* con ingresso collegato all’uscita si comporta come un amplificatore con reazione negativa. Per la porta *NOT* CMOS si ha $V_{in} = V_{out} \simeq 2.5 \text{ V}$.

ed indica i livelli assunti dalle tensioni, che non corrispondono a nessuno dei valori previsti per gli stati logici 0/1. Il circuito smette di funzionare come porta digitale e ritorna piuttosto ad un comportamento da circuito analogico.

È da notare che la condizione che si viene a stabilire corrisponde ad un punto di elevata dissipazione all’interno del dispositivo (in una porta CMOS, ad esempio, i transistor sono entrambi accesi ed al massimo della corrente che li attraversa). Per alcune famiglie questa può essere una condizione operativa non accettabile, in quanto produce riscaldamento eccessivo.

Molto più interessante dal punto di vista delle reti digitali è il circuito nella fig. 10.28, a destra.

Una volta chiuso l’anello il funzionamento del circuito è descritto dalla relazione $C = \sim\sim A = A$, relazione perfettamente valida. L’aspetto interessante è che la relazione è verificata sia per $A = 0$ sia per $A = 1$. In altre parole, uno stato 1 all’ingresso A genera una uscita $C = 1$, che genera, coerentemente, un ingresso $A = 1$. Risultato analogo si ottiene per $A = 0$.

Quindi sia che il circuito si trovi nello stato $A = C = 1$ che nello stato $A = C = 0$, vi rimane indefinitamente, o almeno fino a quando lo stato non viene modificato per effetto di una azione esterna.

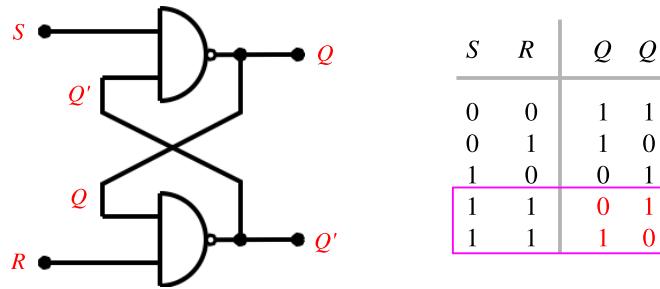
Il circuito è quindi dotato di *memoria* ed è l’elemento base della logica *sequenziale*: i valori delle variabili di uscita di una rete sequenziale dipendono non solo dai valori attuali delle variabili di ingresso, ma anche dalla storia precedente della rete.

Per la sua caratteristica di avere due possibili stati stabili, il circuito è detto *bistabile*. Il termine corrispondente in lingua inglese è *flip-flop*.

10.9 I *flip-flop*

10.9.1 *Set-Reset (SR)*

Il circuito bistabile descritto nell’introduzione precedente è solo di interesse illustrativo, in quanto non è provvisto di alcun terminale di ingresso tramite il quale sia possibile modificarne lo stato.

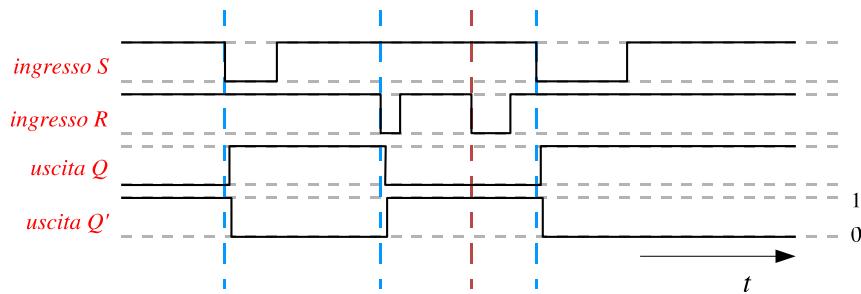
Figura 10.30: Il flip-flop di tipo *set-reset*.

In figura 10.30 è riportato il circuito di uno dei più semplici flip-flop di interesse applicativo, dotato di due terminali di ingresso contrassegnati con le lettere *S* (*Set*) ed *R* (*Reset*), da cui il nome *set-reset* o *SR*.

La tabella di valori a destra illustra il funzionamento del circuito. Quando gli ingressi *S* oppure *R* hanno valore 0, l'uscita corrispondente, *Q* o *Q'* è forzata al valore 1, indipendentemente dal valore dell'altro ingresso e dell'altra uscita. Nello stato *S* = 1 ed *R* = 1, le due porte *NAND* si comportano come due porte *NOT* per i due segnali che provengono dalle uscite *Q* e *Q'* ed il circuito si comporta come quello dell'esempio di fig. 10.28, a destra. I due stati $\{Q = 0, Q' = 1\}$ e $\{Q = 1, Q' = 0\}$ sono entrambi due stati possibili e stabili.

Tramite i due ingressi *S* ed *R* è possibile modificare lo stato del flip-flop: portando *S* a 0 (e lasciando *R* ad 1), le uscite assumono i valori $\{Q = 1, Q' = 0\}$ e questi valori sono mantenuti anche quando *S* torna a 1. Alternativamente, un valore 0 applicato a *R* (lasciando *S* ad 1) porta le uscite ai valori $\{Q = 0, Q' = 1\}$.

Il circuito *ricorda* quale è stato l'ultimo dei due ingressi, *S* o *R*, a passare da 0 a 1.

Figura 10.31: Diagrammi temporali per il flip-flop *SR*. Le uscite cambiano di stato in corrispondenza dei livelli 0 presenti sugli ingressi. In corrispondenza del terzo evento (linea verticale rossa) le due uscite sono già nello stato che dovrebbero raggiungere.

La figura 10.31 illustra il funzionamento del circuito tramite i diagrammi temporali. In corrispondenza delle transizioni da 1 a 0 degli ingressi *S* ed *R*, indicate dalle linee verticali tratteggiate, le uscite cambiano il loro stato di conseguenza, dopo un piccolo ritardo dovuto ai tempi di risposta finiti delle porte.

Si può ipotizzare che i due ingressi *S* ed *R* compiano entrambi la transizione $0 \rightarrow 1$ nello stesso istante. È molto poco probabile che la transizione avvenga *esattamente*

nello stesso istante; una differenza di tempo tra i due segnali, sia pur minima, c'è praticamente sempre. Tuttavia, se anche la transizione avvenisse realmente nello stesso istante, lo stato $S = 1$, $R = 1$, $Q = 1$, $Q' = 1$ non è uno stato stabile; in un tempo brevissimo il circuito ricadrebbe in uno dei due stati stabili, selezionato magari in maniera casuale da qualche minima interferenza sempre presente nei circuiti, o anche solo dall'agitazione termica delle cariche elettriche (rumore).

10.9.2 Clocked-SR

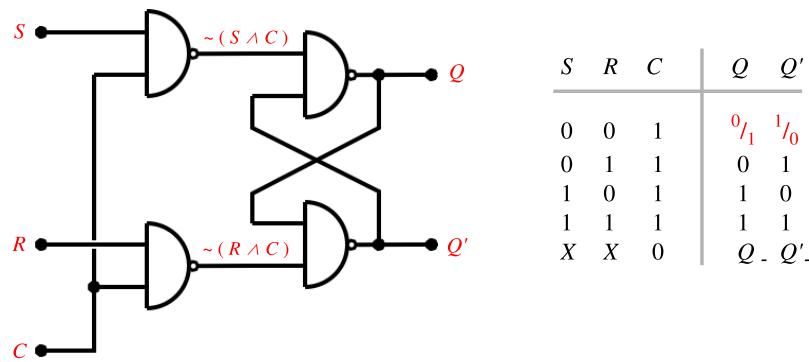


Figura 10.32: Il flip-flop di tipo *clocked set-reset*.

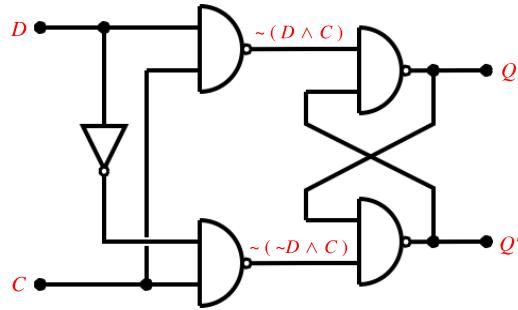
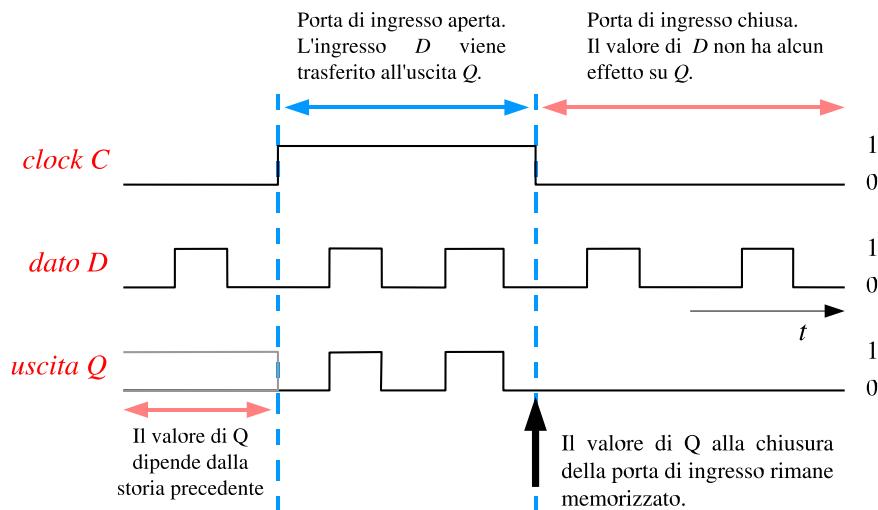
Una variante del flip-flop di tipo *SR* è illustrata in fig. 10.32: vengono introdotte due ulteriori porte di tipo *NAND*, tramite le quali i due segnali S ed R vengono applicati al flip-flop vero e proprio. Le due porte sono controllate da un terzo ingresso C (*Clock*). Quando l'ingresso C è a livello 1 le due porte di ingresso sono *aperte* e i segnali presenti su S ed R (invertiti, perché le porte sono *NAND*) sono applicati all'ingresso del flip-flop e modificano lo stato delle uscite. Quando le porte di ingresso si *chiudono* ($C = 0$), S ed R non possono più giungere agli ingressi del flip-flop e lo stato presente su Q e Q' al momento della chiusura delle porte rimane *congelato* fino a che le porte non vengono nuovamente aperte.

10.9.3 D-type

Nel flip-flop di tipo *D* (*Data*) si ha un ingresso D , che coincide con l'ingresso S del flip-flop *SR*; l'ingresso R viene ottenuto dalla negazione di D (fig. 10.33).

I quattro possibili stati dei due ingressi S ed R di fig. 10.32 si riducono a due $\{0, 1\}$ e $\{1, 0\}$.

Quando il *clock* C è alto (1), le due porte di ingresso sono aperte ed il dato D entra nel flip-flop e viene trasferito sull'uscita; quando il *clock* va basso (0), le porte si chiudono e, come nel caso precedente, rimane memorizzato il valore presente sulle uscite al momento della chiusura delle porte (fig. 10.34).

Figura 10.33: Il flip-flop di tipo *D*.Figura 10.34: Il dato presente all’ingresso *D* viene trasferito all’uscita *Q* (e *Q'*) fino a che il *clock C* è alto; quando *C* diventa basso il valore delle uscite viene congelato ed ogni ulteriore variazione su *D* non ha più alcun effetto.

10.10 I flip-flop di tipo *Master-Slave*

Due flip-flop, ad esempio di tipo *D* o di tipo *SR*, possono essere combinati insieme in una disposizione detta *master-slave* (fig. 10.35). L’uscita del primo flip-flop, il *master*, costituisce l’ingresso del secondo.

Un segnale di comando, il *clock*, è applicato sia al *master* che, invertito tramite una porta *NOT*, allo *slave*. Come nel flip-flop di tipo *D* del paragrafo precedente quando il *clock* è alto l’ingresso *D* e le sue variazioni entrano nel primo flip-flop della catena, il *master*, e sono presenti sulle uscite *Q_M* e *Q'_M*. L’ingresso al secondo flip-flop però è chiuso e quindi i segnali non possono giungere alle uscite *Q* e *Q'*. Solo quando il *clock* torna basso, si verifica l’apertura delle porte dello *slave*, con il trasferimento dei segnali dall’uscita del *master* all’uscita dello *slave*. Contestualmente, però, la porta del *master* si è chiusa, per cui le ulteriori variazioni presenti sull’ingresso *D* non possono più produrre alcun effetto. Al momento che il *clock* torna alto i valori di *D* possono tornare ad entrare nel *master*, ma non vengono più trasferiti nello *slave*.

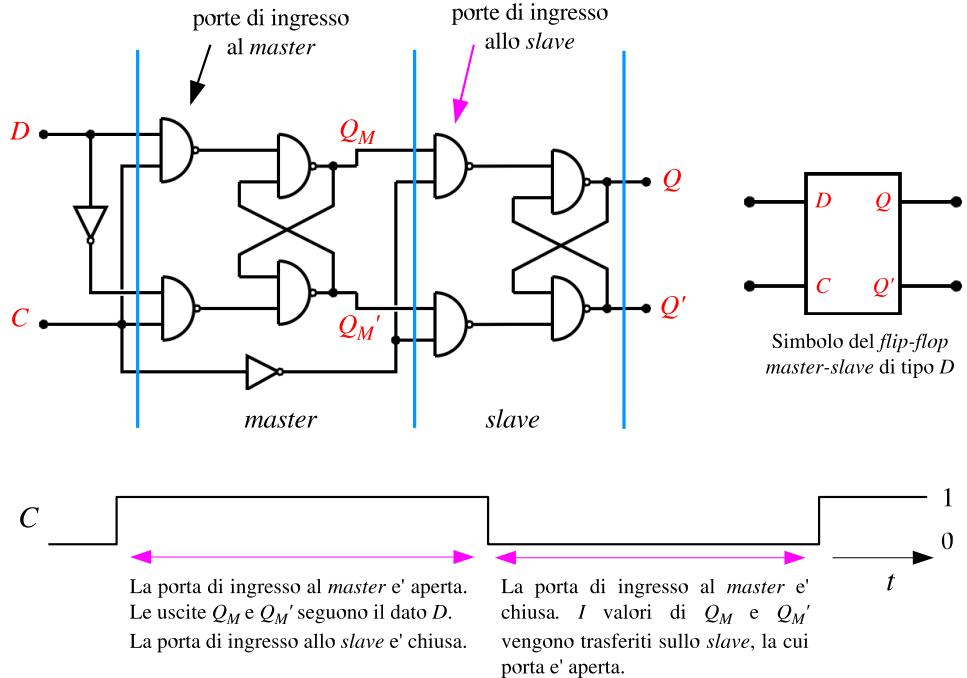


Figura 10.35: Flip-flop di tipo *master-slave* formato dalla combinazione di un flip-flop *D* e di un *clocked-SR*.

a causa della chiusura della porta di ingresso di quest'ultimo.

In conclusione, mentre nel flip-flop del paragrafo precedente l'uscita seguiva l'ingresso per tutto il tempo che il segnale *clock* rimaneva alto, ora l'uscita cambia di stato solo ed esclusivamente in corrispondenza della transizione da alto a basso dell'ingresso *clock*: il dato presente all'uscita del master al momento della transizione viene rimane congelato e trasferito sullo slave.

10.10.1 *Shift register* (Registro a scorrimento)

Una catena di flip-flop di tipo *master-slave* collegati come in fig. 10.36 costituisce un *registro a scorrimento* (*shift register*).

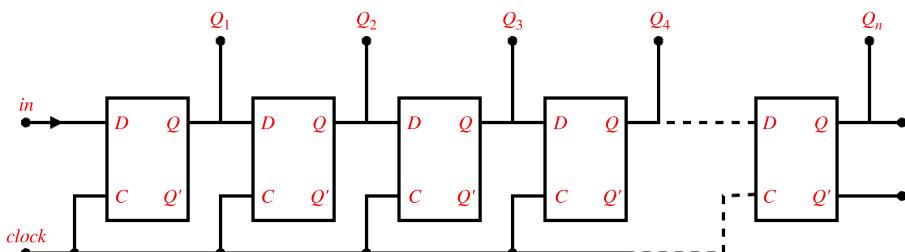


Figura 10.36: *Shift register*: il bit presente all'uscita di ogni flip-flop della catena viene inviato all'ingresso del flip-flop successivo; il segnale *clock* viene applicato contemporaneamente in parallelo a tutti i flip-flop.

Ogni rettangolo nello schema rappresenta un intero flip-flop del tipo descritto al paragrafo precedente, con i suoi ingressi D e C e le uscite Q e Q' . Dal secondo in poi, l'ingresso D di ogni flip-flop è collegato all'uscita Q del precedente. Il segnale $clock$ è applicato in parallelo contemporaneamente a tutti gli elementi della catena.

L'ingresso del primo flip-flop è collegato al segnale in che il circuito riceve dall'esterno.

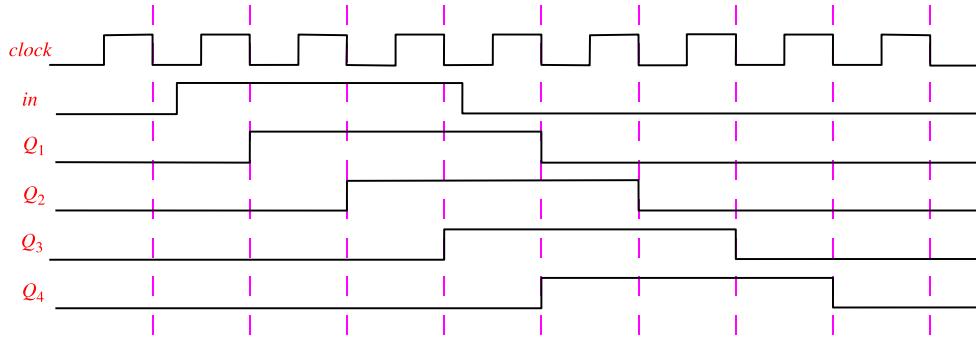


Figura 10.37: Ad ogni transizione da alto a basso del segnale $clock$ (linee tratteggiate viola) il bit presente all'ingresso D di ogni flip-flop della catena viene memorizzato e trasferito alla corrispondente uscita Q : ad ogni ciclo del $clock$ la sequenza di valori dei bit $\{Q_1, Q_2, \dots, Q_n\}$ scorre a destra di una posizione.

Ad ogni transizione del $clock$ da alto a basso ogni flip-flop memorizza il dato che trova presente al proprio ingresso D e lo trasferisce alla propria uscita Q . Il primo flip-flop memorizza il dato presente all'ingresso in di tutta la catena. In questo modo, ad ogni ciclo del segnale $clock$, il valore della linea in viene campionato e fatto scorrere di una posizione nella catena dei flip-flop, come indicato dal diagramma temporale dei segnali in fig. 10.37.

10.10.2 Flip-flop JK

Uno dei flip-flop più versatili è quello indicato come JK (fig. 10.38).

Si tratta di un flip-flop di tipo *master-slave*, formato da due flip-flop *clocked-SR*, con la differenza che l'ingresso al *master* è controllato da due porte *NAND* a tre ingressi. Sul terzo ingresso di ciascuna porta è presente uno dei segnali di uscita, Q o Q' , come indicato nella figura.

Come negli altri *master-slave* le uscite cambiano in corrispondenza della transizione $1 \rightarrow 0$ del $clock$ C , in base al valore degli ingressi J e K , come descritto nella tabella a destra nella figura.

La condizione $J = K = 1$ è particolarmente interessante: quando entrambe le porte di ingresso sono aperte, ad ogni ciclo del segnale $clock$ viene prelevato e trasferito all'uscita del *master* l'inverso dello stato presente all'uscita dello *slave*. Quindi l'uscita Q (e Q') cambia di stato ad ogni ciclo del segnale $clock$ e compie un ciclo completo ogni due cicli di quest'ultimo: il circuito è un *divisore di frequenza*. Se all'ingresso C viene applicata un'onda quadra di frequenza f , l'uscita Q fornisce un'onda quadra di frequenza $f/2$. Un flip-flop costruito appositamente per funzionare

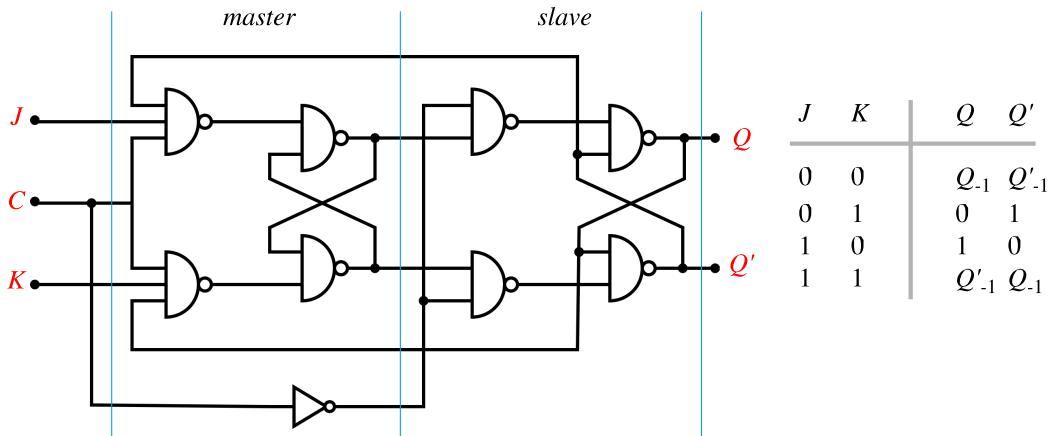


Figura 10.38: Schema del flip-flop di tipo JK . La tabella a destra indica i valori assunti dalle uscite in corrispondenza della transizione $1 \rightarrow 0$ dell'ingresso C . Con l'indice -1 si è indicato lo stato che le uscite avevano *prima* della transizione.

con questa modalità viene indicato come *Toggle* o *tipo T* (*toggle*: interruttore a due posizioni, fig. 10.39).

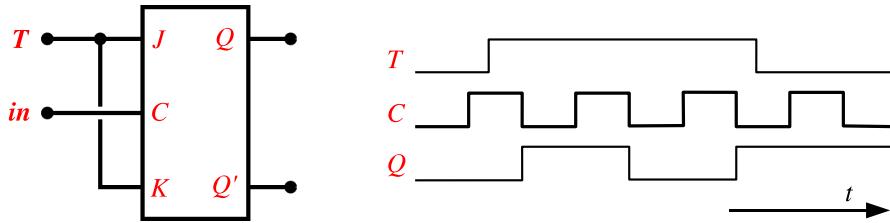


Figura 10.39: Il flip-flop JK con entrambi gli ingressi a 1 è un flip-flop di tipo T .

10.10.3 Contatore binario

In figura 10.40 è riportata una catena di flip-flop di tipo T .

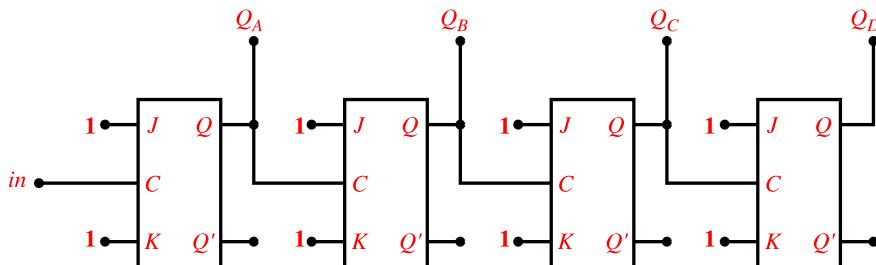


Figura 10.40: Un contatore binario può essere realizzato connettendo in cascata più flip-flop di tipo T (nello schema, flip-flop di tipo JK utilizzati in modalità T).

L'uscita Q_A cambia di stato ogni due cicli del segnale di ingresso in ; Q_B ogni 4; Q_C ogni 8 ecc., come indicato nel diagramma temporale in fig. 10.41.

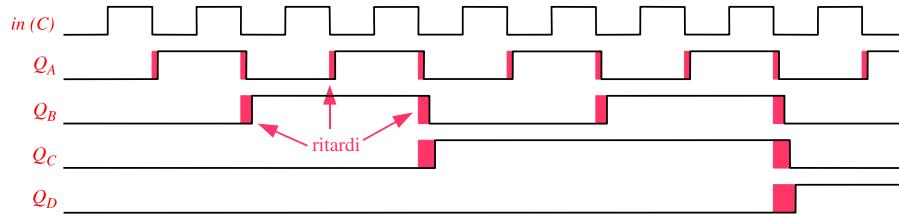


Figura 10.41: Diagramma temporale delle commutazioni in un contatore binario a 4 bit.

Dal grafico si vede facilmente come in ogni istante il valore dei bit $Q_D \dots Q_A$ indichi, in codice binario, il numero di cicli effettuati dal segnale applicato all'ingresso in della catena. Il bit Q_D è il più significativo, il bit Q_A il meno significativo.

Dal grafico temporale si vede anche come la commutazione delle uscite Q avvenga con un ritardo sempre crescente via via che ci si sposta verso destra lungo la catena. Il segnale nel percorso dall'ingresso in alle varie uscite deve attraversare un numero sempre maggiore di porte, ciascuna delle quali introduce il proprio ritardo.

Questo accumularsi dei ritardi pone dei limiti severi alla velocità massima di operazione di un contatore di questo tipo. Quando il ritardo accumulato dallo stadio n -esimo diventa eguale alla durata del ciclo del segnale di ingresso, non esiste più un intervallo di tempo durante il quale le uscite siano tutte quante valide e possano quindi essere lette correttamente.

10.10.4 Contatore sincrono

Il problema visto al paragrafo precedente dell'aumento del ritardo al crescere del numero degli stadi di un contatore binario può essere eliminato utilizzando una rete di conteggio *sincrona*.

Il principio base di ogni rete sincrona è di applicare il segnale che deve determinare la commutazione dei flip-flop contemporaneamente a tutti gli elementi della rete. Nel caso del contatore sincrono, il segnale in viene applicato a tutti gli ingressi C degli n flip-flop (fig. 10.42).

Circuiti logici ausiliari predispongono ciascun flip-flop nella condizione di effettuare o no la commutazione in corrispondenza di ogni fronte di discesa del *clock*. Nel caso del contatore, la logica ausiliaria deve predisporre ciascun ingresso T a 1 (e quindi predisporre per la commutazione al successivo passaggio $1 \rightarrow 0$ di C) se e solo se tutte le uscite Q dei flip-flop precedenti sono a 1, come richiesto dall'aritmetica del conteggio binario.

Nel circuito in figura 10.42 la frequenza massima di conteggio è limitata dal tempo necessario al segnale a giungere da ciascun ingresso C alla corrispondente uscita Q e propagarsi attraverso una singola porta *AND* fino all'ingresso T del flip-flop successivo.

La maggiore complessità circuitale, cioè l'aggiunta per ogni stadio del contatore di una porta *AND* con un numero di ingressi crescente con il numero di stadi, è il

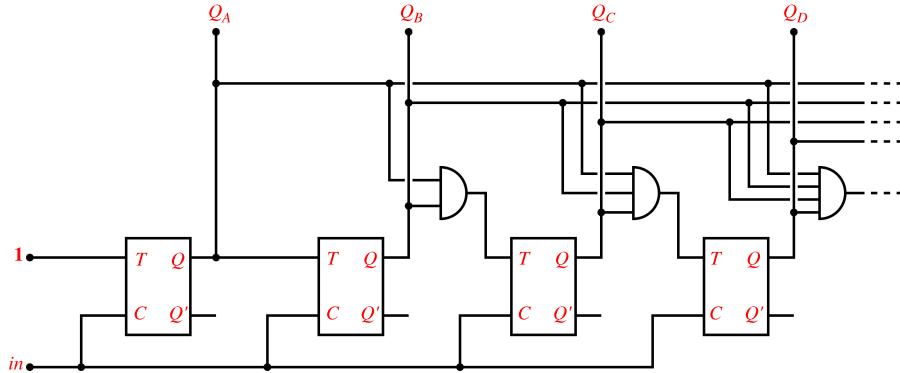


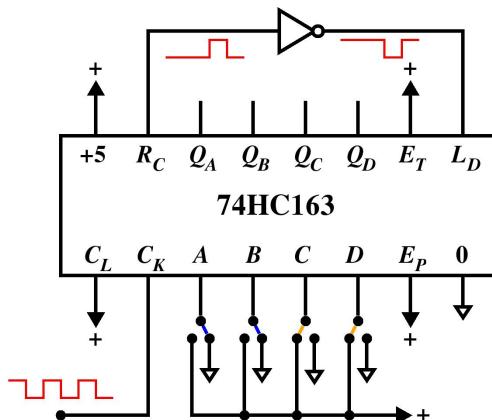
Figura 10.42: Schema di un contatore sincrono a 4bit.

prezzo da pagare per l'aumento di velocità che il contatore sincrono offre rispetto a quello *asincrono* di fig. 10.40.

10.11 Contatore modulo N

Un contatore binario, ad esempio a 4 bit, conta gli impulsi che riceve all'ingresso da 0 a 15 e poi ripete il ciclo ripartendo da zero. È quindi un contatore *modulo 16*. Utilizzando un numero di bit diverso si possono realizzare altri cicli di conteggio, più lunghi o più corti, ma sempre della lunghezza di un intero potenza di 2. Si possono realizzare anche contatori capaci di contare *modulo N* con N intero qualsiasi, non necessariamente potenza di 2.

Nella fig. 10.43 è riportato come esempio di circuito digitale sequenziale lo schema di un tale contatore a 4 bit, con N selezionabile da 1 a 16.

Figura 10.43: Contatore modulo N a 4 bit, con un esempio di ciclo di conteggio con $N = 4$.

Il *chip* 74HC163 è un contatore sincrono a 4 bit *programmabile*, cioè dotato di ingressi e circuiti logici di controllo che permettono di effettuare operazioni di impostazione oltre al conteggio. Un livello basso (0) sull'ingresso *clear* (C_L) oppure *load* (L_D) da luogo in corrispondenza del successivo fronte di salita del *clock* (C_K) rispettivamente all'azzeramento di tutti e quattro i flip-flop del contatore oppure al

caricamento dei valori presenti ai quattro ingressi A, B, C, D . L'uscita R_C (*ripple carry*, riporto) genera un livello alto (1) in corrispondenza dello stato 1 di tutte a quattro le uscite Q_A, Q_B, Q_C, Q_D . Questo segnale, insieme ai due segnali di *enable* E_T ed E_P può essere utilizzato per collegare in cascata altri *chip* identici e realizzare contatori a 8, 12, ... $n \cdot 4$ bit.

Nel circuito di fig. 10.43 il segnale R_C invertito è utilizzato per generare il segnale L_D : quando il conteggio arriva a 1111, al successivo fronte di salita del *clock* C_K le uscite non passano allo stato 0000, ma al valore presente sui quattro ingressi A, B, C, D , nello schema in figura impostati a 1100: D e C collegati a +5, stato logico 1, B e A collegati a massa, stato logico 0, in logica positiva. Il bit D è il più significativo, A il meno significativo (abitualmente i circuiti si disegnano con i segnali che si propagano da sinistra verso destra, quindi nei contatori il primo flip-flop, quello del bit meno significativo, compare per primo a sinistra).

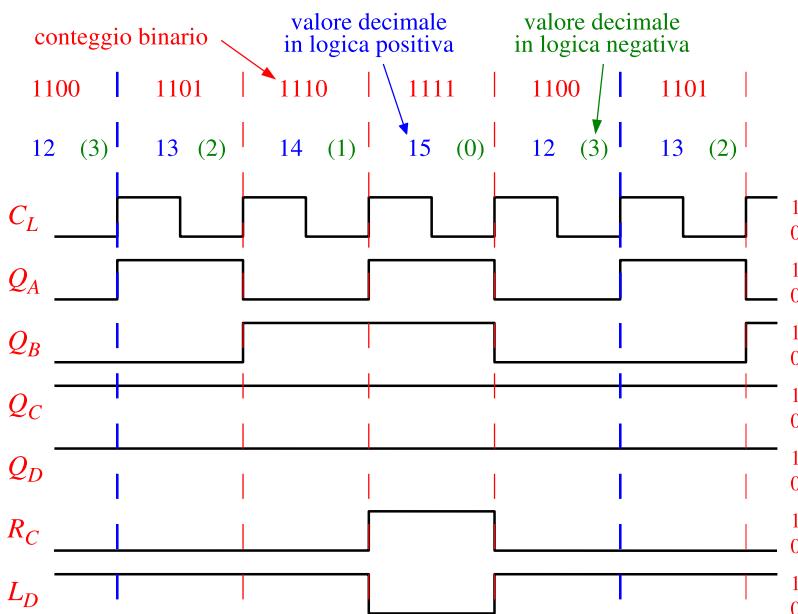


Figura 10.44: Quando le uscite Q sono tutte 1 diventa 1 anche l'uscita R_C (riporto) mentre l'ingresso L_D (load) diventa 0, predisponendo per il caricamento dei dati. Al successivo fronte di salita del *clock* C_K viene caricato il dato presente agli ingressi A, B, C, D (1100) ed R_C torna a zero.

La sequenza temporale degli eventi è illustrata nella fig. 10.44, per un ciclo di lunghezza 4, da 1100 (12_{10}) a 1111 (15_{10}). Le uscite Q_A, Q_B, Q_C, Q_D possono essere lette sia in logica positiva, con un conteggio in salita (*up*) da 12 a 15, sia in logica negativa, con un conteggio in discesa (*down*) da 3 a 0.

I contatori modulo N sono elementi essenziali di molti strumenti elettronici, per esempio degli orologi descritti nel paragrafo seguente e dei generatori di segnali di tipo *PLL*, descritti nel par. 17.6.

10.12 Orologio digitale

Il contatore binario, insieme ad un generatore di segnali di elevata precisione e stabilità, consente di raggiungere nella misura del tempo livelli di risoluzione e precisione che nessuna altra tecnica consente.

Nella fig. 10.45 è riportato lo schema di principio di un orologio digitale. Il segnale

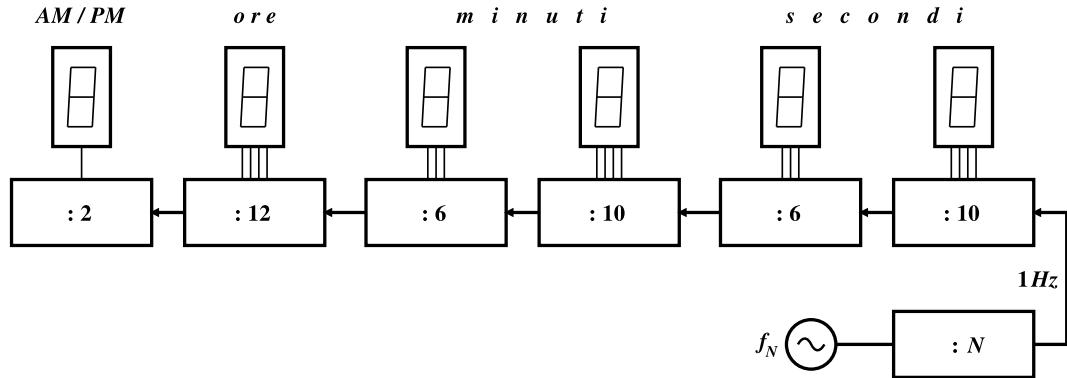


Figura 10.45: Schema a blocchi di un orologio digitale.

di frequenza f_N prodotto da un generatore molto accurato e stabile viene inviato ad un divisore modulo N , con N pari al valore numerico di f_N . Il segnale all'uscita del divisore avrà con altissima precisione la frequenza di 1 Hz .

Seguono, riportati da destra verso sinistra, una serie di contatori modulo 10, 6, 10, 6, 12, 2, che contano rispettivamente secondi, decine di secondi, minuti, decine di minuti, ore e l'indicazione AM/PM.

10.13 Elaboratore digitale

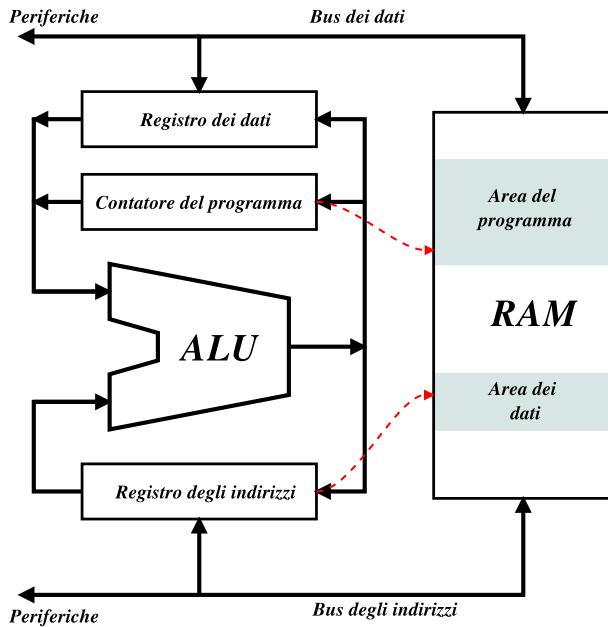


Figura 10.46: Schema a blocchi degli elementi principali di un computer. Il *contatore del programma* individua l'operazione che deve essere eseguita nella sequenza *programma*. Il *registro degli indirizzi* individua le varie logiche contenute nella memoria oggetto delle operazioni da parte della *ALU* (*Arithmetic Logic Unit*).

L'elaboratore digitale, o *computer*, è probabilmente la più complessa rete sequenziale che sia stata ideata. È caratterizzata dal fatto di avere una grande capacità di memoria, molto veloce (*RAM: Random Access Memory*), ed una unità di calcolo

(*ALU: Arithmetic-Logic Unit*) in grado di eseguire sia le operazioni logiche elementari sia operazioni più complesse come ad esempio l'addizione e la moltiplicazione aritmetica (fig. 10.46).

La *RAM* contiene oltre alle variabili oggetto delle operazioni anche le *istruzioni*, cioè le informazioni su quali operazioni debbano essere eseguite in sequenza (il *programma*) su quali variabili (i *dati*).

Capitolo 11

Amplificazione e Reazione

Come si è visto nei cap. 7-9, utilizzando i transistor *bjt* e *fet* è possibile costruire reti attive che consentono di amplificare un segnale, cioè di ottenere da un segnale debole una copia fedele di potenza maggiore. In effetti, l'elettronica nasce con il primo dispositivo amplificatore, il triodo termoionico.

Indipendentemente dalla tecnologia con cui viene ottenuta l'amplificazione, e cioè se si utilizzano transistor, tubi a vuoto o altri dispositivi, esistono alcune proprietà e principi che sono molto generali e comuni a tutti i sistemi capaci di amplificare un segnale. Tra questi sicuramente uno dei più importanti è l'*anello di reazione (feedback loop)*.

11.1 Amplificazione

Nella figura 11.1 sono riportati i due simboli utilizzati comunemente per rappresentare negli schemi elettrici un amplificatore (*a*) ed un amplificatore *differenziale* (*b*). In entrambi i casi le tensioni di ingresso (V_i , V_+ , V_-) e di uscita (V_o) sono

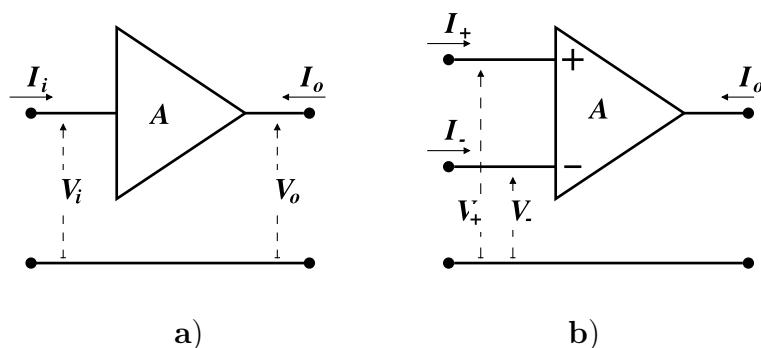


Figura 11.1: Simboli comunemente utilizzati per indicare negli schemi un amplificatore (a sinistra) ed un amplificatore differenziale (a destra).

misurate rispetto ad un nodo comune di riferimento, il nodo di *massa* o di *terra*. I circuiti interni dell'amplificatore sono collegati con il nodo di massa, o direttamente

o tramite le alimentazioni.¹ Spesso negli schemi questi collegamenti sono sottintesi e non indicati esplicitamente, come è stato fatto anche in fig. 11.1. Le correnti si considerano per convenzione positive quando sono *entranti* nel dispositivo.

11.2 Anello di reazione (*feedback loop*)

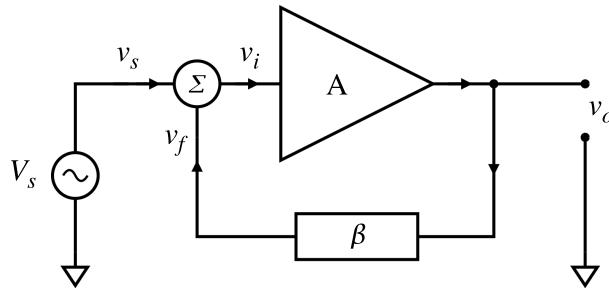


Figura 11.2: Anello di reazione.

In fig. 11.2 è riportato lo schema di un anello di reazione elettronico. Il blocco A rappresenta un amplificatore, di cui almeno in questo contesto non interessa il funzionamento interno, il blocco β una rete passiva, detta rete di reazione; l'elemento Σ schematizza una parte di circuito in cui viene effettuata la somma o la differenza tra il segnale $v_s(t)$ proveniente dal generatore V_s ed il segnale $v_f(t)$ proveniente dalla rete di reazione. Le frecce nel disegno indicano la direzione in cui i segnali si propagano lungo i circuiti.

In questo schema, per semplicità, entrambi gli elementi A e β sono considerati *unilaterali*: si suppone cioè che i segnali si propaghino solo nella direzione indicata dalle frecce, considerando nullo il trasferimento inverso, da destra verso sinistra attraverso l'amplificatore A e da sinistra verso destra attraverso la rete β . Si suppone, cioè, che valgano le relazioni lineari:

$$\begin{aligned} v_o &= A \cdot v_i \\ v_f &= \beta \cdot v_o \\ v_i &= v_s \pm v_f \end{aligned} \tag{11.1}$$

dove A e β sono due costanti. I limiti di questa ipotesi ed una analisi più completa saranno discussi nei paragrafi e nei capitoli seguenti.

Nel circuito si ha *reazione negativa* o *degenerativa* quando la frazione v_f del segnale di uscita riportato all'ingresso attraverso la rete β si contrappone al segnale di ingresso v_s ; quando invece il segnale v_f si somma in fase con il segnale v_s si ha *reazione positiva* o *rigenerativa*. Nell'esempio in figura 11.2, se $v_i = v_s - v_f$ (il circuito Σ effettua la differenza tra v_s e v_f) ed il prodotto $A\beta$ dei coefficienti di trasferimento dell'amplificatore e della rete di reazione è positivo, si ha reazione negativa; se in A

¹A parte il caso degli *amplificatori di isolamento*, costruiti in modo da avere i circuiti di ingresso e di uscita isolati galvanicamente tra di loro e quindi con masse indipendenti, tra le quali possono essere applicate differenze di potenziale anche di centinaia o migliaia di V.

oppure in β (ma non in entrambi) si ha una inversione di segno, la reazione diventa invece positiva.

Questo schema semplice vale fino a che A e β sono quantità costanti, in particolare indipendenti dalla frequenza dei segnali.

Una volta che si consideri che i valori di A e β dipendono dalla frequenza, cosa che succede *sempre* nei sistemi reali, allora sarà opportuno scrivere le equazioni 11.1 nel dominio delle frequenze:

$$\begin{aligned} V_o &= A \cdot V_i \\ V_f &= \beta \cdot V_o \\ V_i &= V_s \pm V_f \end{aligned}$$

e A e β dovranno essere considerate funzioni complesse della frequenza; la situazione diventerà molto più complicata e si potrà avere nello stesso circuito reazione negativa a certe frequenze e positiva ad altre (cap. 15).

11.3 Guadagno ad anello aperto e ad anello chiuso

Interrompendo idealmente l'anello di reazione nel punto P in fig. 11.3 si ha $v_f = 0$ e quindi $v_o/v_s = v_o/v_i = A$; il coefficiente A è quindi il *guadagno ad anello aperto*.

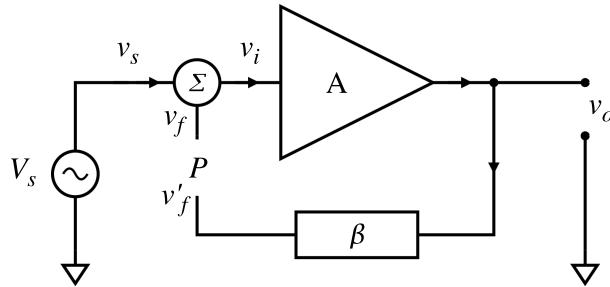


Figura 11.3: Apertura dell'anello di reazione.

Nell'operazione di *apertura dell'anello* si dovrà fare in modo che il comportamento delle due reti A e β rimanga invariato. Normalmente questo non succede: l'apertura dell'anello comporta un cambiamento delle impedenze su cui le due reti sono chiuse e questo modifica il valore dei due coefficienti A e β . Supponiamo di avere preso tutte le precauzioni del caso, ad esempio introducendo delle impedenze per terminare in modo opportuno le due connessioni rimaste aperte nel punto P .

Il prodotto $A\beta$ è il *guadagno d'anello*: il segnale v_i attraversa in successione le due reti A e β e alla fine del percorso, prima di ritornare all'ingresso ed essere sommato o sottratto a v_s , ha una ampiezza $v'_f = A\beta v_i$.

Con l'anello chiuso vale la relazione:

$$v_o = v_i \cdot A = (v_s \pm v_f) \cdot A = (v_s \pm \beta v_o) \cdot A \quad (11.2)$$

da cui si ottiene

$$v_o \cdot (1 \mp A\beta) = v_s \cdot A$$

ed il *guadagno ad anello chiuso*:

$$G = \frac{v_o}{v_s} = \frac{A}{1 \mp A\beta} \quad (11.3)$$

11.4 Reazione positiva

Nel caso che il segnale v_f sia sommato a v_s o, equivalentemente, che sia sottratto ma il prodotto $A\beta$ sia negativo, a denominatore della eq. 11.3 compare il segno $-$ e siamo in presenza di *reazione positiva*; per maggiore chiarezza conviene scrivere l'equazione 11.3 come:

$$G = \frac{A}{1 - |A\beta|} \quad (11.4)$$

Per valori di $|A\beta| < 1$ si vede che il guadagno G ad anello chiuso risulta maggiore del guadagno A ad anello aperto.

Storicamente la reazione nei circuiti elettronici è nata in radiotecnica come reazione positiva per aumentare l'amplificazione assai scarsa dei primi tubi a vuoto e per generare segnali persistenti². Infatti quando $|A\beta|$ tende ad 1, il guadagno G tende ad infinito ed il circuito diventa capace di generare un segnale di uscita anche con un segnale di ingresso nullo: diventa un *oscillatore*. La stessa cosa succede con $|A\beta| > 1$, anche se questo non risulta dalle equazioni precedenti: algebricamente sembra sempre possibile ottenere soluzioni che indicano un comportamento *normale* del circuito. Questo problema sarà affrontato nei capitoli 15 e 17, nell'ambito del problema più generale della stabilità delle reti, studiando la risposta dei circuiti non solo in condizioni statiche o stazionarie, ma considerando anche il comportamento durante i transitori.

11.5 Reazione negativa

Nel caso di reazione negativa a denominatore della eq. 11.3 compare il segno $+$ (con il prodotto $A\beta$ positivo) e quindi il guadagno ad anello chiuso risulta minore del guadagno ad anello aperto di un fattore $1 + A\beta$. A prima vista l'uso della reazione negativa sembra una cosa decisamente poco intelligente: costruire circuiti capaci di amplificare un segnale e poi introdurre ulteriori circuiti e componenti per degradarne questa capacità. Forse è stata proprio questa la prima risposta nel mondo dell'elettronica alla proposta di utilizzare la reazione negativa nei circuiti amplificatori.

In realtà la combinazione di *amplificazione* e *reazione negativa*³ è lo strumento più potente di cui si dispone oggi nella realizzazione dei sistemi elettronici. Infatti tecnologicamente è possibile costruire circuiti dotati di amplificazione, anche molto elevata, ma spesso non è possibile ottenere intrinsecamente da questi circuiti l'accuratezza, la stabilità o le caratteristiche di risposta richieste. D'altra parte la precisione,

²Lo schema elettrico riportato sulla copertina di questi appunti è un oscillatore a *triodo*, che fa uso della reazione positiva per generare segnali a radio frequenza

³Insieme ad un uso accorto delle simmetrie nei circuiti, come nell'amplificatore differenziale del par. 7.9.

l'accuratezza e la stabilità sono caratteristiche che è relativamente facile ottenere dai componenti passivi, soprattutto dalle resistenze e dai condensatori.

La reazione negativa permette di barattare in maniera conveniente le due caratteristiche: cedere parte del guadagno elevato ma poco accurato degli amplificatori in cambio della precisione elevata dei componenti passivi privi di amplificazione, ottenendo come risultato amplificazione accurata e stabile e funzioni di risposta ben definite.

11.6 L'Amplificatore Operazionale

Molto spesso in un anello di reazione l'amplificazione A viene fornita da un amplificatore differenziale di tensione detto *Amplificatore Operazionale (op-amp)*.

Un amplificatore differenziale è dotato di due ingressi denominati *invertente* (-) e *non invertente* (+) ed amplifica la tensione *di modo differenziale* $v_d = v_+ - v_-$ tra i due ingressi, indipendentemente dalla tensione *di modo comune* $v_c = (v_+ + v_-)/2$.

Nella figura 11.4 è riportato oltre al simbolo utilizzato per rappresentare l'amplificatore operazionale negli schemi elettrici, anche il suo modello elettrico equivalente.

Nella tab. 11.1 sono riportate le caratteristiche principali di un amplificatore operazionale ideale, il cui funzionamento è descritto dalla relazione

$$v_o = A \cdot (v_+ - v_-) \quad (11.5)$$

con A costante, indipendente dalla frequenza e tendente ad infinito.

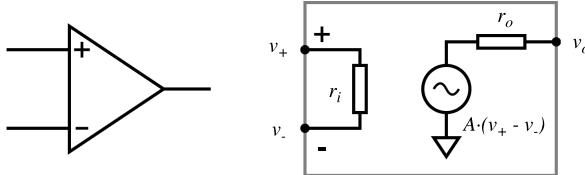


Figura 11.4: Simbolo utilizzato per rappresentare un amplificatore operazionale negli schemi elettrici (a sinistra) e circuito equivalente (a destra).

resistenza di ingresso	r_i	infinita
resistenza di uscita	r_o	zero
amplificazione in modo differenziale	A	infinita
amplificazione in modo comune	A_c	zero
banda passante	B_w	infinita

Tabella 11.1: Caratteristiche principali di un amplificatore operazionale ideale.

L'idea alla base di un anello di reazione negativa con un amplificatore con un guadagno idealmente infinito è che in un circuito come quello di fig. 11.2 il segnale di reazione v_f arrivi ad annullare quasi esattamente il segnale v_s in modo da lasciare

solo una piccola differenza, il segnale v_i , detto anche *segnale di errore*, come descritto dalle eq. 11.1-11.3. Tanto più elevato è il valore di A tanto più piccolo è il segnale residuo di errore v_i e quindi tanto migliore l'egualanza tra v_f e v_s .

Non esiste nessun singolo dispositivo fisico che consenta di ottenere direttamente le prestazioni descritte dalla eq. 11.5; nella pratica le relazioni tra le variabili di ingresso e di uscita di un dispositivo amplificatore sono più complicate e non lineari, come illustrato ad esempio dalle famiglie di curve caratteristiche dei transistor *bjt* e *fet* nelle figure 7.5 e 9.9. È possibile però combinando opportunamente più dispositivi realizzare circuiti il cui comportamento si avvicina considerevolmente a quello ideale.

11.6.1 L'amplificatore operazionale nei circuiti

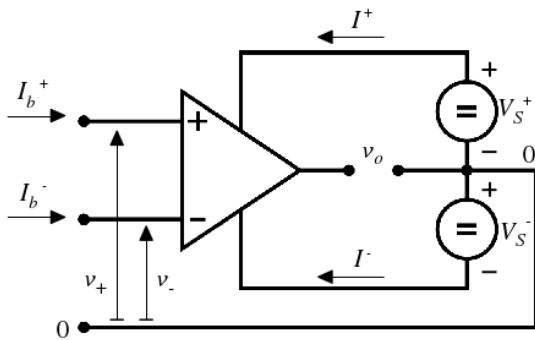


Figura 11.5: Circuiti di alimentazione di un amplificatore operazionale. Normalmente si utilizzano due generatori di tensione uguali, che forniscono una tensione di alimentazione positiva ed una negativa.

Come già visto nella fig. 11.1, nel simbolo di fig. 11.4 compaiono solo i terminali dell'amplificatore che riguardano i segnali e cioè i due terminali di ingresso v_+ e v_- ed il terminale di uscita v_o . In realtà, come tutti i circuiti attivi, anche l'amplificatore operazionale ha bisogno di una sorgente di energia per funzionare (*alimentazione*), anche se spesso negli schemi dei circuiti le alimentazioni vengono omesse per semplificare il disegno e si considerano sottointese.

Il più delle volte un amplificatore operazionale è alimentato da due generatori di tensione collegati come in fig. 11.5, in modo da fornire al circuito una tensione di alimentazione positiva (V_S^+) ed una negativa (V_S^-). I valori delle tensioni V_S^\pm possono andare da un minimo di meno di un Volt fino a diverse centinaia di Volt, ma comunemente sono compresi tra $\pm 5\text{ V}$ e $\pm 15\text{ V}$.

Il terminale in comune tra i due generatori costituisce il nodo di riferimento (*massa*) rispetto a cui sono misurate tutte le altre tensioni.

Nello schema nella figura sono indicate anche le correnti di *polarizzazione* (*bias*) degli ingressi del circuito, I_b^\pm .

11.6.2 Amplificatori Operazionali reali

L'amplificatore operazionale reale differisce da quello ideale praticamente in tutti gli aspetti. Tenendo conto delle principali fonti di errore, l'eq. 11.5 diventa:

$$v_o = A \cdot (v_+ - v_- + V_{os}) + A_c \cdot (v_+ + v_-) \quad (11.6)$$

dove V_{os} è una tensione di errore⁴, la tensione di *offset* tra i due ingressi, e A_c è l'*amplificazione di modo comune*. Inoltre le costanti A e A_c non sono realmente costanti, ma funzioni complesse della frequenza.

La tensione di *offset* (*sbilanciamento*) è dovuta principalmente alle imperfezioni anche minime nella simmetria dei componenti del circuito. In parte è una tensione costante che in linea di principio può essere compensata, in parte è affetta da numerose derive, le più importanti dovute al variare della temperatura e della tensione di alimentazione V_S^\pm ; in qualche misura V_{os} dipende anche dall'invecchiamento del dispositivo. L'entità di alcune di queste fonti di errore è riportata nella tab. 11.2.

Alla tensione di errore V_{os} , intrinseca al dispositivo amplificatore, nei circuiti si aggiunge un contributo $\Delta v_{ib} = I_b^+ R^+ - I_b^- R^-$ dovuto alle correnti di polarizzazione I_b^+ ed I_b^- dei due ingressi, che scorrono rispettivamente nelle resistenze R^+ ed R^- (fig. 11.6). Avere valori eguali per R^+ ed R^- tende a ridurre questo errore ma non

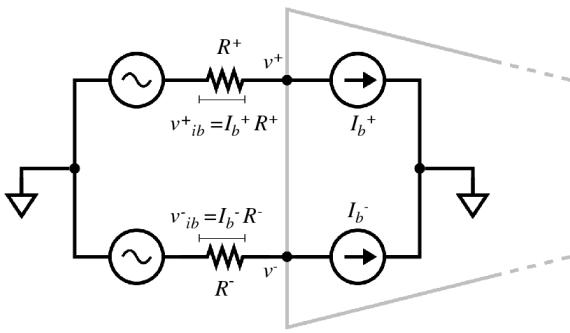


Figura 11.6: Errore Δv_{ib} prodotto dalle correnti di polarizzazione I_b^\pm . Le due cadute di potenziale $I_b^+ R^+$ ed $I_b^- R^-$ tendono a compensarsi, nella misura in cui i loro contributi sono eguali. La loro differenza è un contributo alla tensione di *offset*.

può eliminarlo del tutto, in quanto anche le correnti di polarizzazione non sono identiche, ma sono affette da uno sbilanciamento (I_{bo}). Inoltre correnti e sbilanciamento dipendono dalla temperatura.

Alcuni modelli significativi di amplificatori operazionali reali, con le loro caratteristiche principali, sono riportati nella tabella 11.2.

Il $\mu A709$, introdotto nel 1965, è stato storicamente il primo amplificatore operazionale ad alto guadagno *monolitico*, cioè realizzato interamente come circuito integrato su un unico substrato di silicio; in precedenza erano stati realizzati amplificatori operazionali a transistor e componenti discreti ed ancora prima con tubi a vuoto o anche con dispositivi elettromeccanici.

Il $\mu A741$ è una versione notevolmente migliorata del $\mu A709$, ed è stato per molti anni uno dei più utilizzati nei più svariati campi. Offre un guadagno A , una resistenza di ingresso r_i ed una robustezza maggiori. Il $\mu A741$ può tollerare un cortocircuito tra l'uscita e la massa o verso una delle due alimentazioni senza danneggiarsi, come invece era la regola con il $\mu A709$.

Il $TL081$ è il successore del $\mu A741$ e ne riprende e migliora molte caratteristiche, soprattutto l'impedenza di ingresso. Infatti il $TL081$ è equipaggiato con uno stadio di ingresso differenziale che fa uso di transistor *jfet* invece che *bjt* come nei dispositivi precedenti.

L' $AD549$ è un amplificatore *elettrometrico*, dotato di uno stadio di ingresso a *jfet*. È uno dei dispositivi più sensibili oggi disponibili per la misura di piccole correnti,

⁴Da non confondere con il segnale di errore v_i del par. 11.6.

Alcuni modelli di amplificatori operazionali

		$\mu A709$	$\mu A741$	$TL081$	$AD549$	$OP27$	$HA2540$
A	V/mV	45	200	200	1000	1800	15
B_w	MHz	1	1	3	1	8	400
r_i	$M\Omega$	0.25	2	10^6	10^7	0.1	0.01
r_o	Ω	150	200	200	500	70	30
V_{os}	mV	2	1	3	0.5	$10\mu V$	8
α_v	$\mu V/^\circ C$	10	6	18	10	0.4	20
I_b	pA	$300nA$	$80nA$	30	0.15	$15nA$	$5\mu A$
I_{bo}	pA		$20 nA$	5	0.05		
C_{MRR}	dB	90	90	86	90	120	72
P_{SRR}	dB		90	86		118	70
v_n	nV/\sqrt{Hz}		24	18	35	3.2	6
i_n	fA/\sqrt{Hz}		550	10	0.22	400	6

- A : coefficiente di amplificazione di tensione (amplificazione in modo differenziale);
 B_w : prodotto *guadagno* \times *larghezza di banda* (frequenza a cui l'amplificazione è scesa a 1, cioè a 0 dB);
 r_i : resistenza di ingresso;
 r_o : resistenza di uscita;
 V_{os} : tensione di *offset* riferita all'ingresso (sbilanciamento in tensione tra i due ingressi);
 α_v : coefficiente di deriva della tensione di *offset* con la temperatura;
 I_b : corrente di polarizzazione degli ingressi dell'amplificatore;
 I_{bo} : sbilanciamento tra le correnti di polarizzazione degli ingressi
 C_{MRR} : reiezione di modo comune (rapporto tra amplificazione in modo differenziale ed amplificazione in modo comune);
 P_{SRR} : *Power Supply Rejection Ratio*: rapporto tra variazione della tensione di alimentazione e variazione della tensione di *offset*;
 v_n : tensione equivalente di rumore;
 i_n : corrente equivalente di rumore;

Tabella 11.2: Caratteristiche principali di alcuni modelli di Amplificatori Operazionali di uso comune. I parametri A , r_i , r_o e C_{MRR} dipendono dalla frequenza. I valori riportati sono quelli limite per $f \rightarrow 0$ (valori *in continua*).

con una corrente di polarizzazione e di rumore estremamente bassa, dell'ordine dei fA .

Le caratteristiche di estrema sensibilità in corrente degli amplificatori con stadi di ingresso a *j fet* non si accoppiano bene con caratteristiche analoghe per quanto riguarda la tensione. Uno dei dispositivi migliori per quanto riguarda la tensione di *offset*, la sua stabilità ed il rumore di tensione è l'*OP27*, dotato di uno stadio di ingresso differenziale con transistor *bjt*.

Infine, l'*HA2540* è un amplificatore operazionale non molto brillante riguardo alle caratteristiche esaminate fino ad ora, ma caratterizzato da una larghezza di banda circa due ordini di grandezza superiore agli altri dispositivi.

La fig. 11.7 mostra alcuni esemplari di amplificatori operazionali.



Figura 11.7: Alcuni esemplari di amplificatori operazionali. Al centro, in alto, un $\mu A709$ ed in basso un $TL081$; a destra, un modello OPA544, in grado di erogare una potenza di uscita fino a 30 W. Il blocco a sinistra è un operazionale contemporaneo al $\mu A709$, non monolitico, ma realizzato a componenti discreti. Le sue prestazioni erano nettamente superiori a quelle del $\mu A709$, così come l'ingombro, il consumo ed il costo.

11.6.3 Altri limiti degli amplificatori operazionali reali

L'eq. 11.6 illustra le principali fonti di errore presenti nel funzionamento di un amplificatore operazionale reale, ma ha la caratteristica importante di essere ancora una equazione lineare. Gli amplificatori reali sono soggetti anche ad altri limiti fisici oltre i quali non si comportano più in modo lineare e di conseguenza non possono più essere descritti mediante sole equazioni algebriche. I più importanti di questi limiti sono:

- tensione di ingresso di modo comune (*common mode input range*): l'intervallo di valori delle tensioni ai due ingressi entro il quale sono garantite le caratteristiche di funzionamento del dispositivo. A causa del valore molto elevato dell'amplificazione di un operazionale, la tensione di ingresso in modo differenziale, cioè la differenza di tensione $v^+ - v^-$ tra i due ingressi in condizioni di funzionamento normale, è sempre molto piccola, dell'ordine dei μV . Questa tensione *differenziale* si può trovare sovrapposta ad una tensione di *modo comune* $(v^+ - v^-)/2$ che può essere ordini di grandezza maggiore e che normalmente può variare entro alcuni Volt dalle tensioni di alimentazione. Ad esempio, per il $TL081$ con tensioni di alimentazione di $\pm 15 V$ la massima tensione di ingresso di modo comune è garantita minimo $\pm 11 V$, fino ad un massimo di $\pm 15 V$.
- tensione massima di uscita: l'intervallo di tensione che il dispositivo può fornire all'uscita. Questo valore dipende dal *carico*, cioè dalla resistenza complessiva dei circuiti collegati all'uscita. Valori tipici sono $\pm 10 V$ con una resistenza di carico di $1 k\Omega$. Quando un amplificatore arriva al limite della tensione massima che può erogare si dice che *entra in saturazione* e la tensione di uscita non segue più in maniera lineare le variazioni delle tensioni di ingresso.
- corrente massima di uscita: quasi tutti gli amplificatori operazionali sono dotati di circuiti di controllo che limitano la corrente di uscita a valori di sicurezza

per evitare un riscaldamento eccessivo del dispositivo ed un possibile danneggiamento. Anche in questo caso quando viene raggiunto il valore massimo, tipicamente 10...20 mA, il dispositivo entra in saturazione e non si comporta più in modo lineare.

- velocità massima di variazione della tensione di uscita (*slew rate*): il limite nella corrente di uscita insieme alla capacità presente sull'uscita (e altri meccanismi analoghi presenti nei circuiti interni di un amplificatore operazionale) limitano la velocità con cui può variare la tensione di uscita da un valore ad un altro. In un segnale sinusoidale la tensione varia a velocità diverse: zero sul picco positivo e negativo, massima nel momento del passaggio per lo zero. Se viene superato il massimo consentito dallo *slew rate* del dispositivo, nel tratto interessato l'andamento del segnale non è più sinusoidale ma assume pendenza costante: tende a diventare un'onda triangolare. Ancora una volta il funzionamento del dispositivo perde di linearità.

11.7 Amplificatore di tensione *non invertente*

La prima e più semplice applicazione dell'amplificatore operazionale con reazione negativa è l'amplificazione di un segnale con un livello di precisione molto alto. Lo

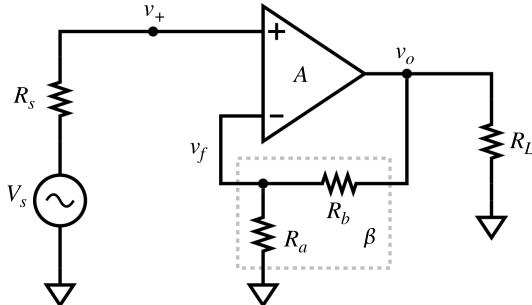


Figura 11.8: Amplificatore di tensione con guadagno $G = (R_a + R_b)/R_a$.

schema in fig. 11.8 è quello di fig. 11.2 con la rete β formata dal partitore R_a, R_b . Considerando inizialmente, per semplicità, di utilizzare un (ipotetico) amplificatore operazionale ideale, con $r_i = \infty$, $r_o = 0$ e $A \rightarrow \infty$, si ha:

$$\frac{v_f}{v_o} = \beta = \frac{R_a}{R_a + R_b} \quad (11.7)$$

La differenza tra il segnale di ingresso v_s ed il segnale di reazione v_f viene fatta direttamente dall'ingresso differenziale dell'amplificatore operazionale:

$$v_o = A \cdot (v^+ - v^-) = A \cdot (v_s - v_f) \quad (11.8)$$

Le resistenze R_s ed R_L non hanno alcun effetto sul funzionamento del circuito, la prima grazie alla resistenza di ingresso infinita e la seconda grazie alla resistenza di

uscita nulla dell'amplificatore A . Ad anello chiuso, secondo l'eq. 11.3 e nel limite $A \rightarrow \infty$, si ottiene il guadagno di tensione

$$G = \frac{A}{1 + A\beta} \simeq \frac{1}{\beta} = \frac{R_a + R_b}{R_a} \quad (11.9)$$

che dipende solo dalle caratteristiche della rete passiva β e non da quelle dell'amplificatore A .

11.7.1 *Voltage follower*

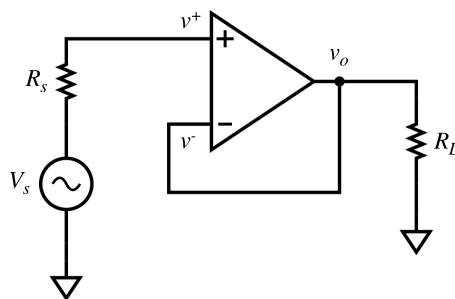


Figura 11.9: Circuito *inseguitore di tensione* (voltage follower)

Nel circuito precedente, nel caso limite $R_a = \infty$ ed $R_b = 0$, si ha $\beta = 1$ e quindi anche il guadagno risulta essere $G = 1$. In queste condizioni il circuito prende il nome di *voltage follower* (inseguitore di tensione): la tensione di uscita segue fedelmente quella di ingresso con elevatissima precisione. Il circuito non è più un amplificatore di tensione, ma solo di corrente. Viene utilizzato come adattatore di impedenza, interponendolo tra una sorgente di segnale con resistenza interna R_s elevata ed un carico R_L di valore troppo piccolo per essere collegato direttamente al generatore di segnale (fig. 11.9).

11.8 Effetti della reazione negativa

Il risultato del paragrafo precedente (eq. 11.9) è stato ottenuto con un ipotetico amplificatore operazionale quasi ideale. Nella pratica, come già visto, un amplificatore operazionale non ha amplificazione e resistenza di ingresso infinita, non ha resistenza di uscita zero, non è mai perfettamente lineare, anzi può presentare una distorsione apprezzabile, ed ha una banda passante limitata.

Inoltre i valori dei parametri riportati nella tabella 11.2 sono solo nominali. Gli amplificatori operazionali non sono dispositivi di precisione. Per quanto riguarda l'amplificazione o la resistenza di ingresso, esemplari diversi di uno stesso modello di amplificatore operazionale possono presentare valori che differiscono anche per un ordine di grandezza.

È solo l'accoppiamento della amplificazione elevata con la reazione negativa che consente di superare ognuno di questi limiti.

11.8.1 Stabilizzazione del guadagno

È relativamente facile costruire amplificatori operazionali con un guadagno A molto elevato. Valori di 10^6 sono oggi molto comuni in tutti i dispositivi di uso generale. È praticamente impossibile che questo guadagno sia accurato; si può arrivare ad una indeterminazione su questo parametro anche di un ordine di grandezza.

Nonostante questa scarsissima precisione, il guadagno G del circuito amplificatore di fig. 11.8 risulta molto preciso e stabile. Dalla eq. 11.9, dividendo numeratore e denominatore per $A\beta$ e tenendo conto del fatto che comunque $A\beta \gg 1$, si ottiene

$$G = \frac{A}{1 + A\beta} = \frac{1/\beta}{1 + 1/(A\beta)} \simeq \frac{1}{\beta} \cdot \left(1 - \frac{1}{A\beta}\right) = G_\infty \cdot \left(1 - \frac{1}{A\beta}\right) \quad (11.10)$$

La differenza frazionaria tra il guadagno G ad anello chiuso ed il guadagno ideale $G_\infty = 1/\beta$ che si avrebbe in condizioni ideali, cioè con $A \rightarrow \infty$, è $1/(A\beta)$, l'inverso del guadagno d'anello.

Una indeterminazione ΔA sul valore di A ha un effetto sul valore di G pari a

$$\frac{\Delta G}{G_\infty} = \frac{\Delta A}{A^2\beta} = \frac{1}{A\beta} \cdot \frac{\Delta A}{A}$$

L'indeterminazione relativa su A produce quindi una indeterminazione relativa su G più piccola di un fattore pari al guadagno d'anello $A\beta$.

In condizioni ordinarie A è dell'ordine di $10^5 \dots 10^6$ mentre β è dell'ordine di $0.01 \dots 1$. Di conseguenza il guadagno d'anello varia tra 10^3 nel peggiore dei casi e 10^6 nelle condizioni migliori.

11.8.2 Riduzione della distorsione

In fig. 11.10 è riportata la tensione di uscita V_O in funzione della tensione di ingresso V_I per un amplificatore che presenta una forte non linearità (curva rossa). L'amplificazione incrementale $a = v_o/v_i = dV_O/dV_I$ per i piccoli segnali, corrispondente alla pendenza della curva, è 1000 al centro del grafico, vicino allo zero, e diminuisce fino a circa 300 quando la tensione di uscita arriva a $\pm 5V$. La non linearità, o distorsione, può essere considerata come una variazione del guadagno incrementale a tra regioni diverse della curva.

Nella stessa figura è riportata anche la relazione $V_O(V_I)$ per lo stesso amplificatore utilizzando una rete di reazione negativa con $\beta = 0.1$ (curva blu). L'aumento della linearità e la riduzione della distorsione (a scapito dell'amplificazione) è immediatamente evidente.

Per effetto della reazione negativa, secondo l'eq. 11.9, il guadagno incrementale ad anello chiuso è $g = a/(1 + \beta a)$. Con i valori indicati, al centro della curva si ha $g = 1000/(1 + 0.1 \cdot 1000) \simeq 9.90$, mentre nell'intorno di $V_O = \pm 5V$ si ha $g = 300/(1 + 0.1 \cdot 300) \simeq 9.68$. La non linearità, espressa come variazione del guadagno incrementale, si è ridotta di un fattore dello stesso ordine di quello di cui si è ridotto il guadagno.

11.8.3 Aumento della banda passante

In questo paragrafo il funzionamento del circuito, che si suppone composto interamente da componenti lineari, è studiato nel dominio delle frequenze; nei paragrafi

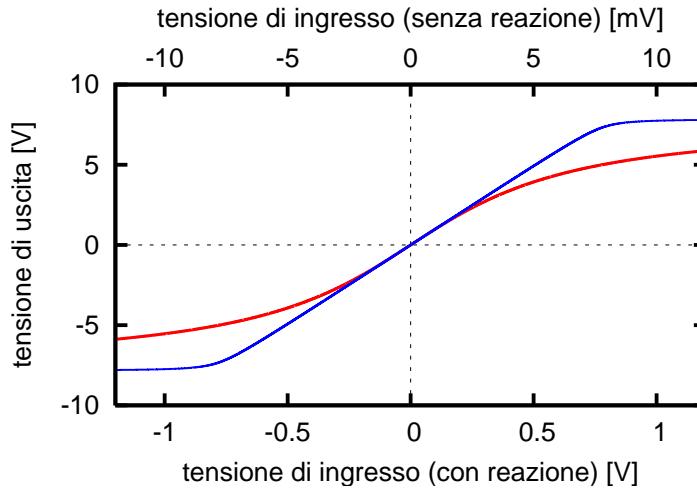


Figura 11.10: Tensione di uscita in funzione della tensione di ingresso per un amplificatore con forte distorsione. La curva rossa corrisponde al comportamento dell'amplificatore ad anello aperto, cioè senza reazione (tensione di ingresso, in mV , sul bordo superiore del grafico). La curva blu corrisponde al comportamento ad anello chiuso con un coefficiente di reazione negativa $\beta = 0.1$ (tensione di ingresso, in V , sul bordo inferiore del grafico.)

11.8.2 e 11.8.4, dove si considerano comportamenti non lineari, l'analisi è stata limitata per semplicità al caso di segnali abbastanza lenti da poter essere considerati statici. Il caso più generale di circuiti non lineari con risposte dipendenti anche dalle frequenze dei segnali deve essere affrontato necessariamente tramite l'integrazione, di solito numerica, del sistema di equazioni differenziali non lineari che descrive la rete (vedi par. 4.13 e cap. 18).

Nel par. 11.2 e nei successivi i coefficienti A e β sono stati considerati costanti, indipendenti dalla frequenza. Questa è una buona approssimazione per la rete β : normalmente è possibile realizzare una rete passiva costituita da resistenze e condensatori che dia una attenuazione sufficientemente costante sul campo di frequenze di interesse, fino a frequenze dell'ordine dei GHz . La stessa approssimazione non è praticamente mai utilizzabile per l'amplificatore, specie se dotato di guadagno elevato.

Per ragioni legate ai problemi di stabilità che verranno analizzati nel capitolo 15, spesso gli amplificatori operazionali presentano una dipendenza della amplificazione A dalla frequenza f del tipo

$$A(f) = \frac{V_o(f)}{V_i(f)} = A_0 \cdot \frac{1}{1 + j f / f_0}$$

corrispondente alla combinazione di un amplificatore ideale con banda passante infinita in serie ad una cella RC di tipo *passa-basso*, con frequenza di taglio $f_0 = 1/(2\pi RC)$ (fig. 11.11). Sostituendo questa espressione di A nella eq. 11.9, si ottiene:

$$G(f) = \frac{A(f)}{1 + A(f)\beta} = \frac{A_0/(1 + j f / f_0)}{1 + A_0\beta/(1 + j f / f_0)} = \frac{A_0}{1 + j f / f_0 + A_0\beta} \quad (11.11)$$

Conviene riscrivere questa equazione mettendo in evidenza il comportamento asintotico per $f \rightarrow 0$, separandolo dalla variazione in funzione della frequenza:

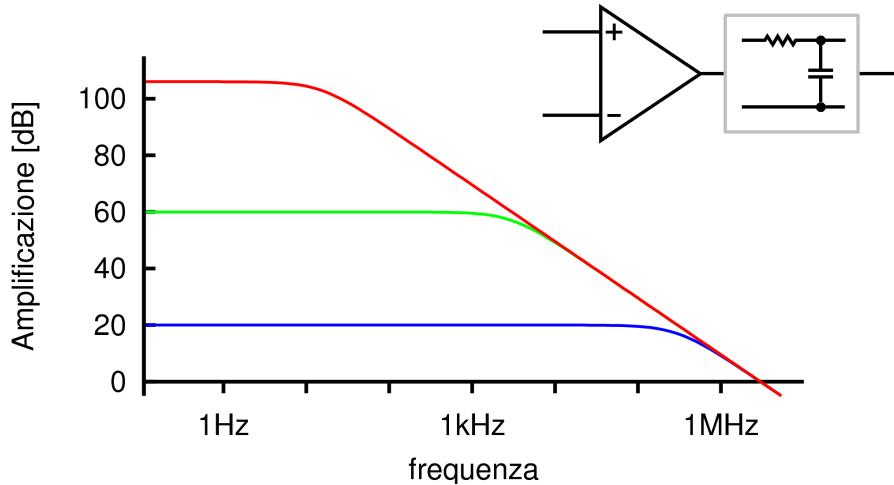


Figura 11.11: Risposta in frequenza dell'amplificatore operazionale TL081. La curva rossa è l'amplificazione ad anello aperto $A(f)$, espressa in dB; le due curve verdi e blu sono le curve del guadagno $G(f)$ in funzione della frequenza per $\beta = 0.001$ e $\beta = 0.1$ rispettivamente. In alto a destra è riportato il circuito lineare equivalente del TL081, composto da un amplificatore ideale con banda passante infinita, seguito da un filtro passa-basso con frequenza di taglio di 15 Hz.

$$G(f) = \frac{A_0}{1 + A_0\beta} \cdot \frac{1}{1 + jf/f'_0} \simeq \frac{1}{\beta} \cdot \frac{1}{1 + jf/f'_0} \quad (11.12)$$

con $f'_0 = f_0(1 + A_0\beta)$. La risposta in frequenza del circuito con reazione negativa è ancora quella di un amplificatore ideale (con guadagno $G = 1/\beta$), in serie con un filtro passa-basso con frequenza di taglio f'_0 che risulta, ad anello chiuso, maggiore di quella ad anello aperto dello stesso fattore $(1 + A_0\beta)$ per cui risulta diminuito il coefficiente di amplificazione. In conseguenza di questo, il prodotto ad anello chiuso tra frequenza di taglio ed amplificazione (prodotto Guadagno \times Banda passante) rimane costante per i diversi valori di reazione β e pari al valore $B_w = f_0 \cdot A_0$.

Nel grafico in fig. 11.11 è riportata in blu la curva $G(f)$ corrispondente a $\beta = 0.1$ e in verde a $\beta = 0.001$.

11.8.4 Rete di reazione non lineare

Si può immaginare di utilizzare per la rete β di fig. 11.2 una rete *non lineare*, in cui la relazione tra la tensione v_f riportata all'ingresso e la tensione v_o prelevata all'uscita sia

$$v_f = f_\beta(v_o) \quad (11.13)$$

dove la funzione f_β è non lineare.⁵

Le eq. 11.2-11.3, per il caso di reazione negativa, diventano:

$$v_o = [v_s - f_\beta(v_o)] \cdot A \quad (11.14)$$

⁵È necessario però che la funzione f_β sia monotona.

$$v_o + Af_\beta(v_o) = Av_s \quad (11.15)$$

Se A è sufficientemente grande da valere la relazione $Af_\beta(v_o) \gg v_o$, allora la eq. 11.15 diventa

$$\begin{aligned} Af_\beta(v_o) &= Av_s \\ f_\beta(v_o) &= v_s \\ v_o &= f_\beta^{-1}(v_s) \end{aligned} \quad (11.16)$$

dove con f_β^{-1} si è indicata la funzione inversa di f_β .

Quindi, se si dispone di una rete β in grado di trasformare un segnale secondo una data relazione non lineare f_β , è possibile ottenere la trasformazione inversa f_β^{-1} inserendo la rete β nell'anello di reazione di un amplificatore operazionale. Ad esempio, è possibile realizzare circuiti in grado di fornire una tensione di uscita v_o proporzionale al prodotto di due tensioni di ingresso v_1 e v_2 :

$$v_o = kv_1v_2 \quad (11.17)$$

e quindi anche al quadrato di una tensione di ingresso:

$$v_o = kv_1^2 \quad (11.18)$$

con k costante con dimensioni V^{-1} . Utilizzando un tale circuito come rete β in un anello di reazione è possibile ottenere la funzione inversa⁶

$$v_o = k\sqrt{v_s} \quad (11.19)$$

Il caso degli anelli PLL (*Phase Lock Loop*) è un esempio estremo del principio illustrato in questo paragrafo: è possibile realizzare circuiti divisorii digitali che consentono di ottenere da un segnale di frequenza f_1 un altro segnale di frequenza $f_n = f_1/n$ per qualsiasi valore di n intero. Non è possibile fare direttamente l'operazione inversa, cioè passare da un segnale con frequenza f_1 ad uno con frequenza $f_n = n \cdot f_1$ altro che per pochi e piccoli valori di n .

Un anello PLL consente invece di ottenere questo risultato: moltiplicare la frequenza di un segnale per un coefficiente n intero qualsiasi, utilizzando un divisore per n all'interno di un anello di reazione (par. 17.6).

11.8.5 Effetto della reazione negativa sul rumore

In un circuito complesso come un amplificatore operazionale sono presenti numerose sorgenti di rumore; praticamente ogni resistenza ed ogni giunzione partecipano in misura diversa al degradamento della qualità del segnale aggiungendovi il proprio contributo.

Nel circuito in fig. 11.12 all'interno dell'amplificatore sono state indicate come esempio due sorgenti di rumore: la prima, v_{n1} , si trova all'ingresso e quindi il suo contributo compare all'uscita moltiplicato per l'intera amplificazione $A = A_1 \cdot A_2$ del dispositivo; la seconda invece, v_{n2} , si trova in una posizione intermedia, più vicina

⁶Poichè la funzione parabolica v^2 non è monotona, è necessario limitare opportunamente l'intervallo di funzionamento del circuito ad uno solo dei due rami della parabola.

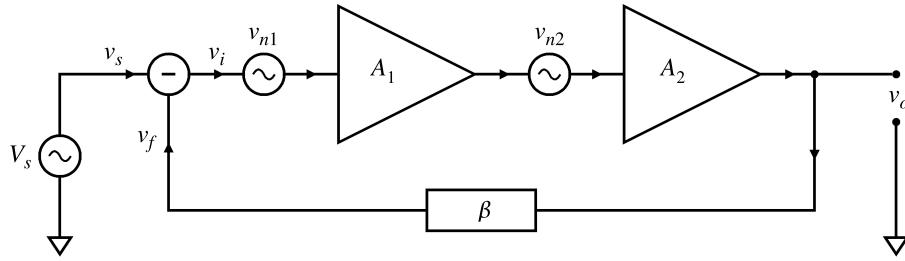


Figura 11.12: Circuito amplificatore con due sorgenti di rumore, una all'ingresso ed un'altra in una posizione intermedia verso l'uscita.

all'uscita, ed il suo contributo viene amplificato solo per il coefficiente A_2 . In una schematizzazione lineare è del tutto equivalente considerare che anche v_{n2} si trovi all'ingresso dell'amplificatore, ma che la sua ampiezza sia ridotta del fattore A_1 . Questa procedura permette di raccogliere in un unico generatore⁷ posto all'ingresso del dispositivo i contributi di tutte le fonti di rumore, come in fig. 11.13, dove si ha $v_n = v_{n1} + v_{n2}/A_1$.⁸

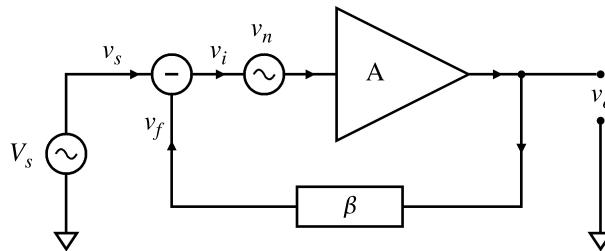


Figura 11.13: Tutte le fonti di rumore del dispositivo amplificatore sono state raccolte nell'unico generatore v_n , all'ingresso.

Le equazioni 11.2-11.3 diventano, nel caso di reazione negativa:

$$\begin{aligned} v_o &= (v_s - \beta v_o + v_n) \cdot A \\ v_o (1 + \beta A) &= (v_s + v_n) \cdot A \\ v_o &= (v_s + v_n) \cdot \frac{A}{1 + \beta A} \end{aligned} \quad (11.20)$$

Dall'ultima di queste equazioni si vede che la reazione negativa è totalmente ininfluente in quanto rumore e segnale vengono amplificati della stessa identica quantità.

Poichè la tensione di errore Δv compare nell'eq. 11.6 esattamente come la tensione di rumore nelle 11.20, la conclusione precedente vale anche per tutte le fonti di errore dell'amplificatore operazionale reale riassunte nel termine Δv . In particolare, la reazione negativa non ha nessun effetto per quanto riguarda la tensione di *offset*.

⁷In una descrizione completa normalmente si devono considerare un generatore di rumore di tensione ed uno di corrente, per tenere conto degli effetti della resistenza che l'amplificatore *vede* al suo ingresso.

⁸Poichè le due sorgenti di rumore v_{n1} e v_{n2} sono statisticamente indipendenti, i loro valori efficaci si sommano quadraticamente: $V_n^2 = V_{n1}^2 + V_{n2}^2/A_1^2$.

11.9 Resistenze di ingresso e di uscita

La resistenza di ingresso R_i di un circuito è data per definizione dal rapporto tra la tensione v^+ che un generatore esterno V_S applica all'ingresso e la corrente i^+ che fluisce nel terminale di ingresso (fig. 11.14). Nel caso dell'amplificatore ideale con

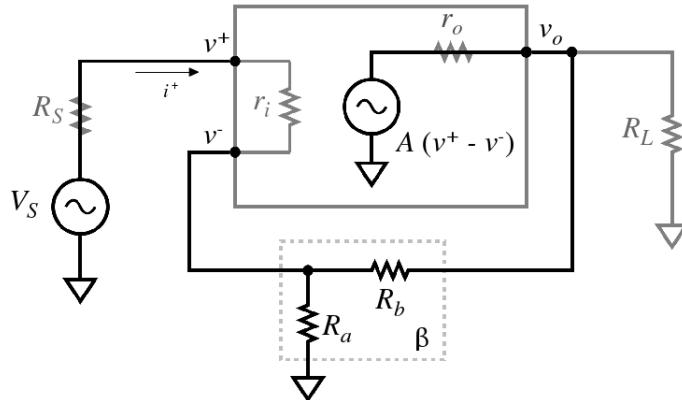


Figura 11.14: Circuito di ingresso dell'amplificatore non invertente. La resistenza di ingresso che il circuito presenta al generatore V_S è per definizione $R_i = v^+ / i^+$.

$r_i = \infty$ si ha ovviamente $i^+ = 0$ e $R_i = \infty$. Per valori di r_i finiti, invece, la corrente di ingresso è $i^+ = (v^+ - v^-) / r_i$ e quindi $R_i = r_i \cdot v^+ / (v^+ - v^-)$.

Se $r_i \gg R_a$, vale con buona approssimazione la relazione $v_- \simeq (v^+ - v^-) \cdot A\beta$; ricavando da questa v^- e sostituendo nella precedente si ottiene

$$R_i \simeq \frac{v^+}{i^+} = r_i \cdot (1 + A\beta) \quad (11.21)$$

L'effetto della reazione negativa è quindi di aumentare del fattore $1 + A\beta$, cioè normalmente di svariati ordini di grandezza, la resistenza di ingresso che un circuito amplificatore presenta complessivamente rispetto a quella che sarebbe la resistenza di ingresso intrinseca del solo dispositivo amplificatore.

Se si ha $r_i > R_a$, o anche soltanto $r_i \simeq R_a$, la relazione 11.21 continua ad essere valida, con un livello di approssimazione tanto peggior quanto più si riduce il fattore $A\beta$ per effetto della presenza di r_i .

Questo risultato non tiene conto del partitore di tensione formato dalla resistenza r_o con il parallelo tra R_L e la resistenza equivalente di tutta la rete R_b, R_a, r_i, R_S , che introduce nel prodotto $A\beta$ un ulteriore fattore di correzione minore di 1; è però facile verificare che anche questa correzione ha un effetto trascurabile sulla eq. 11.21.

Infatti la presenza della resistenza equivalente della rete R_b, R_a, r_i, R_S in parallelo con R_L introduce una penalizzazione nella ampiezza massima della tensione di uscita del circuito (*dinamica di uscita*). Un amplificatore operazionale senza alcun carico collegato può fornire all'uscita una tensione in un intervallo tipicamente di 1 o 2 V inferiore alle tensioni di alimentazione.⁹ Ad esempio, con alimentazioni ± 15 V, la tensione di uscita può essere compresa tra i valori ± 13 V.

⁹ Alcuni modelli più moderni, indicati con il termine *rail-to-rail*, arrivano a fornire un intervallo di tensioni di uscita pari a quello di alimentazione.

Indicando con R'_L la resistenza di carico complessiva presente all'uscita, si ha un fattore di attenuazione

$$a_o = \frac{R'_L}{r_o + R'_L} \quad (11.22)$$

Dovendo limitare per ragioni pratiche la perdita della dinamica di uscita ad esempio a non più del 10%, si dovrà avere $a_o \geq 0.9$ e quindi R'_L almeno un ordine di grandezza maggiore di r_o . Questa condizione pone un limite inferiore ai valori di R_L , R_a ed R_b .

Riassumendo le considerazioni precedenti, si dovrà quindi avere

$$r_o \ll R_L // (R_a + R_b) \quad R_a \ll r_i \quad (11.23)$$

Dalla tabella 11.2 si vede che il rapporto r_i/r_o va da un minimo di 330 (per l'*HA2540*) fino a $2 \cdot 10^{10}$ (per l'*AD549*). Non c'è quindi nessuna difficoltà a rispettare entrambe le condizioni indicate dalla eq. 11.23, con una riduzione del coefficiente $A\beta$ non superiore a qualche unità percentuale, correzione sicuramente meno significativa della stessa indeterminazione nel valore di A, che difficilmente è meglio di un fattore 2.

In conclusione, la resistenza di ingresso del circuito di fig. 11.14 viene comunque ad avere un valore elevatissimo. Anche con una r_i modesta, dell'ordine delle decine di $k\Omega$, grazie al valore molto elevato del guadagno d'anello $A\beta$, si può arrivare ad avere facilmente una R_i di centinaia o migliaia di $M\Omega$.

Bisogna però fare attenzione ai valori fantasticamente alti che si possono ottenere dal calcolo. Ad esempio, per un *AD549* utilizzato come *voltage follower*, quindi con guadagno d'anello pari a 10^6 , si otterebbe

$$R_i = 10^7 M\Omega \cdot 10^6 = 10^{19} \Omega$$

Questo valore è illusorio. Un modello più accurato del circuito di ingresso dell'amplificatore operazionale è riportato in fig. 11.15.

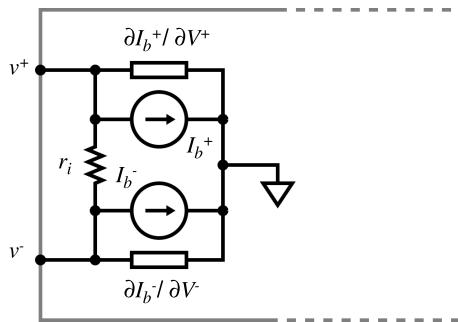


Figura 11.15: Circuito di ingresso dell'amplificatore operazionale, in cui si tiene conto degli effetti delle correnti di polarizzazione I_{bias} .

Ciascun ingresso ha la sua corrente di polarizzazione I_{bias} (schematizzata dai due generatori di corrente I_b^\pm) e questa corrente può essere anche ordini di grandezza superiore ai valori della corrente di ingresso che risulterebbe con i valori di R_i ottenuti dalla eq. 11.21. Per quanto le correnti I_b^\pm siano discretamente costanti al variare di V^+ e V^- , i due generatori non si comporteranno mai in modo completamente ideale e ad ognuno di essi andrà associata una conduttanza $\partial I_b^\pm / \partial V^\pm$.¹⁰ Questa quantità,

¹⁰Ad esempio, se lo stadio di ingresso dell'amplificatore utilizza transistor *jfet*, la corrente di polarizzazione è quella di una giunzione contropolarizzata, con un andamento analogo a quello riportato in fig. 4.10.

indicata come conduttanza o resistenza di ingresso di modo comune, viene riportata raramente nelle specifiche degli amplificatori operazionali ed è di interesse solo con i dispositivi più sensibili. La resistenza di ingresso di modo comune può limitare il valore effettivo della R_i anche di ordini di grandezza al di sotto del valore $r_i(1+A\beta)$. Ad esempio, l'AD549 ha una resistenza di ingresso di modo comune di $10^{15}\Omega$.

11.9.1 Resistenza di uscita

La resistenza di uscita R_o è definita in modo analogo alla resistenza di ingresso come il rapporto tra la tensione v_l e la corrente i_l di un ipotetico generatore collegato all'uscita al posto della resistenza R_L (fig. 11.16), avendo posto all'ingresso $v_s = 0$.

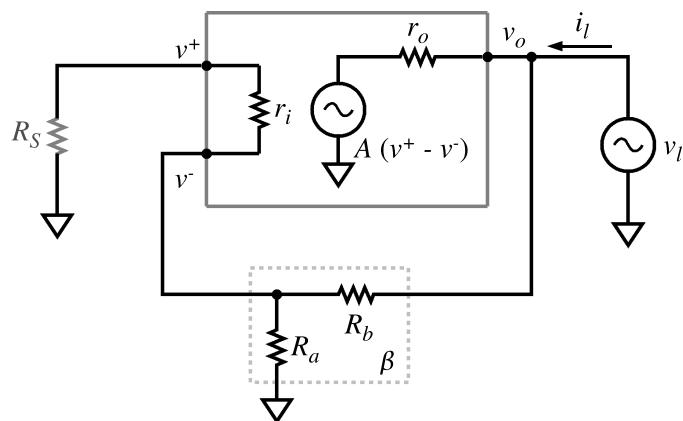


Figura 11.16: La resistenza di uscita di una rete è per definizione data dal rapporto v_l/i_l tra la tensione applicata da un generatore collegato al posto della resistenza di carico R_L e la corrente che fluisce verso il nodo di uscita v_o .

Dal punto di vista algebrico risulta più conveniente calcolare la conduttanza di uscita $G_o = 1/R_o = i_l/v_l$:

$$i_l = \frac{v_l}{R_\beta} + \frac{v_l - A(v^+ - v^-)}{r_o} \quad (11.24)$$

dove R_β è la resistenza equivalente della rete R_b, R_a, r_i (ed R_S) vista dal nodo v_o verso massa. La tensione $v_l (= v_o)$, attraverso il partitore β , genera una corrente attraverso r_i ed R_S e quindi una tensione tra i terminali di ingresso dell'amplificatore operazionale

$$v^- - v^+ = \beta \frac{r_i}{r_i + R_S + R'_\beta} v_l \quad (11.25)$$

dove $R'_\beta = R_a // R_b = R_a R_b / (R_a + R_b)$. Sostituendo nella eq. 11.24 si ottiene

$$G_o = \frac{1}{R_o} = \frac{i_l}{v_l} = \frac{1}{R_\beta} + \frac{1 + A\beta \frac{r_i}{r_i + R_S + R'_\beta}}{r_o} = \frac{1}{R_\beta} + \frac{1 + A'\beta}{r_o} \quad (11.26)$$

con $A' = Ar_i/(r_i + R_S + R'_\beta)$.

La resistenza di uscita è data quindi dalla combinazione in parallelo della resistenza R'_β e della resistenza $r_o/(1 + A'\beta)$.

In un circuito ben progettato si avrà cura di avere $r_i \gg R_S$, o almeno $r_i > R_S$, o se proprio questo non è possibile, un rapporto r_i/R_S non tanto piccolo da ridurre il guadagno d'anello $A'\beta$ a valori inaccettabili. Al limite, per $R_S \rightarrow \infty$ si avrebbe $A' \rightarrow 0$ e

$$G_o = \frac{1}{R'_\beta} + \frac{1}{r_o} \simeq \frac{1}{r_o} \quad (11.27)$$

Ma questo equivarrebbe ad avere scollegato il terminale di ingresso del circuito amplificatore da ogni altro circuito.

In definitiva, a meno di valori abnormemente elevati per R_S , si ha $A'\beta \approx A\beta$ e quindi, anche riutilizzando le relazioni 11.23,

$$R_o \approx \frac{r_o}{1 + A\beta} \quad (11.28)$$

Con i valori di A ed r_o della tab. 11.2 e con β tra 0.01 ed 1 l'eq. 11.28 dà un intervallo di valori di R_o compreso tra 0.2Ω nel peggiore dei casi e $0.5 m\Omega$ nelle condizioni migliori.

Anche in questo caso bisogna fare attenzione ai valori fantasticamente piccoli che si ottengono dal calcolo. Se si collega l'uscita dell'amplificatore ad altri circuiti utilizzando, ad esempio, un tratto di conduttore di rame di lunghezza 10 cm e diametro 0.3 mm, si aggiunge in serie al circuito una resistenza di $4 m\Omega$, che può anche essere di un ordine di grandezza superiore alla resistenza di uscita calcolata.

11.9.2 Effetto di r_i ed r_o su G

Nei due paragrafi precedenti si è visto che per effetto della reazione negativa anche nel caso di amplificatori operazionali con valori di r_i ed r_o non ideali, il comportamento del circuito tende ad avvicinarsi molto a quello ideale, con $R_i \rightarrow \infty$ ed $R_o \rightarrow 0$.

Il dubbio che viene lasciato non chiarito in quei paragrafi è quanto la presenza di r_i ed r_o possa modificare il risultato della eq. 11.9 nel par. 11.7, cioè quale può essere l'errore che viene introdotto nel valore di G .

Una risposta definitivamente convincente può venire solo dallo studio completo del circuito di fig. 11.14.

Ad esempio, scrivendo per i due nodi v^- e v_o le due equazioni di Kirchhoff:

$$\frac{v^-}{R_a} + \frac{v^- - v^+}{r_i} + \frac{v^- - v_o}{R_b} = 0 \quad (11.29)$$

$$\frac{v_o - v^-}{R_b} + \frac{v_o - A(v^+ - v^-)}{r_o} = 0 \quad (11.30)$$

con una certa manipolazione algebrica si arriva alla soluzione:

$$G = \frac{v_o}{v^+} = \frac{A + \beta \frac{r_o}{r_i}}{1 + A\beta + \beta \left(\frac{r_o}{r_i} + \frac{R_b}{r_i} + \frac{r_o}{R_a} \right)} \quad (11.31)$$

in cui, come prima, $\beta = R_a/(R_a + R_b)$, secondo la 11.7.

Si verifica molto facilmente che con le condizioni 11.23 tutti i termini correttivi che compaiono in più nella eq. 11.31 rispetto alla eq. 11.9 sono ininfluenti e non modificano la precisione del risultato già ottenuto.

Questo modo di affrontare il problema, però, è dispendioso perché richiederebbe di ripetere la verifica per ogni tipo o variante di circuito. Nel cap. 14 il problema verrà affrontato in maniera più formale, ma anche molto più generale, per arrivare ad una dimostrazione valida per tutti i tipi di anelli di reazione, mettendo in evidenza le due condizioni essenziali che garantiscono la validità dei risultati ottenuti:

- un guadagno d'anello βA sufficientemente elevato;
- la *unilateralità* dell'amplificatore A .

11.10 Amplificatore *invertente*

L'amplificatore di tensione di fig. 11.8 effettua sui segnali elettrici la prima e più semplice delle *operazioni* che è possibile realizzare con l'amplificatore operazionale:¹¹ la moltiplicazione dell'ampiezza di un segnale per una costante positiva e maggiore di uno. Per questo motivo il circuito di fig. 11.8 è indicato come *amplificatore non invertente*.

Il circuito di fig. 11.17 invece è in grado di fornire una amplificazione negativa, sia maggiore che minore di 1 in modulo ed è detto *amplificatore invertente*.

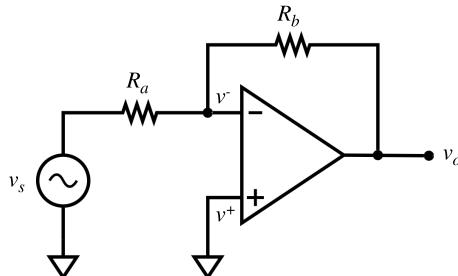


Figura 11.17: Amplificatore invertente. $G = v_o/v_s = -R_b/R_a$.

Poichè la rete è lineare, per il principio di sovrapposizione la tensione v^- si può scrivere come somma dei due contributi dovuti alla tensione v_s del generatore di segnale ed alla tensione $v_o = A \cdot (v^+ - v^-)$ del generatore di uscita dell'amplificatore operazionale:

$$v^- = v_s \cdot \frac{R_b}{R_a + R_b} + v_o \cdot \frac{R_a}{R_a + R_b} = v_s(1 - \beta) + v_o\beta \quad (11.32)$$

dove si è posto, come prima, $\beta = R_a/(R_a + R_b)$. Poichè $v^+ = 0$, si ha $v_o = -A \cdot v^-$ e sostituendo nella equazione 11.32 si ottiene

$$G = \frac{v_o}{v_s} = -(1 - \beta) \cdot \frac{A}{1 + A\beta} \quad (11.33)$$

¹¹Il nome *amplificatore operazionale* deriva proprio dal suo impiego in circuiti per l'esecuzione di operazioni matematiche sui segnali elettrici.

Nel limite $A \rightarrow \infty$:

$$G = -\frac{1-\beta}{\beta} = -\frac{R_b}{R_a} \quad (11.34)$$

L'espressione ottenuta per il guadagno G dell'amplificatore invertente (eq. 11.33) coincide con l'eq. 11.9 a meno del fattore costante $-(1-\beta)$, che non dipende da A . Di conseguenza, tutti i risultati e le considerazioni del par. 11.8 riguardo gli effetti della reazione negativa, basati sul comportamento della quantità $A/(1+A\beta)$, e cioè la stabilizzazione del guadagno, la riduzione della distorsione e l'aumento della banda passante, si applicano egualmente anche all'amplificatore invertente.

L'unica differenza è data dal fatto che il fattore $(1-\beta)$ introduce una certa penalizzazione, in quanto l'amplificazione viene ridotta del fattore $(1+A\beta)/(1-\beta)$ mentre gli effetti di errore solo del fattore $1+A\beta$.

11.10.1 Resistenza di uscita

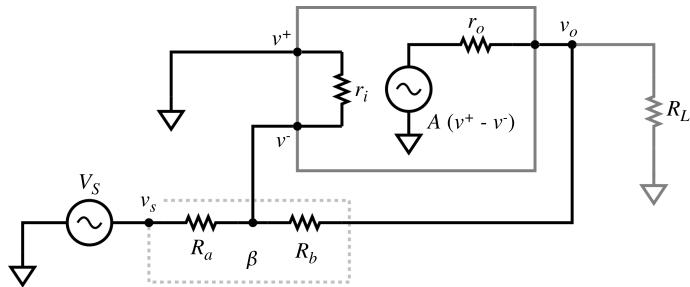


Figura 11.18: Circuito equivalente dell'amplificatore invertente, completo con le resistenze r_i ed r_o .

Sostituendo nel circuito equivalente dell'amplificatore invertente di fig. 11.18 il generatore V_S con un corto circuito e la resistenza di carico R_L con un generatore di tensione v_L per procedere a valutare la resistenza di uscita r_o come definita all'inizio del par. 11.9.1, si vede che si ottiene un circuito identico a quello di fig. 11.16, con $R_S = 0$: tutto quanto detto per l'amplificatore non invertente vale identicamente per l'amplificatore invertente, l'effetto della reazione negativa è di ridurre la resistenza di uscita dell'amplificatore da r_o a $r_o/(1+A\beta)$.

11.10.2 Resistenza di ingresso

Situazione completamente diversa si ha per quanto riguarda la resistenza di ingresso.

Il generatore V_S vede una resistenza di ingresso $R_i = R_a + R'_i$, dove R'_i è la resistenza di ingresso che l'amplificatore presenta direttamente al nodo v^- . Tralasciando per ora R_L , R'_i risulta composta dalla resistenza r_i in parallelo con la resistenza di ingresso di un amplificatore ideale con guadagno $-A$ e resistenza di reazione tra ingresso ed uscita pari a $r_o + R_b$ (fig. 11.19). Per il teorema di Miller, questo circuito presenta una resistenza di ingresso pari a $(r_o + R_b)/(A+1)$. Quindi

$$\frac{1}{R'_i} = \frac{1}{r_i} + \frac{A+1}{r_o + R_b} \quad (11.35)$$

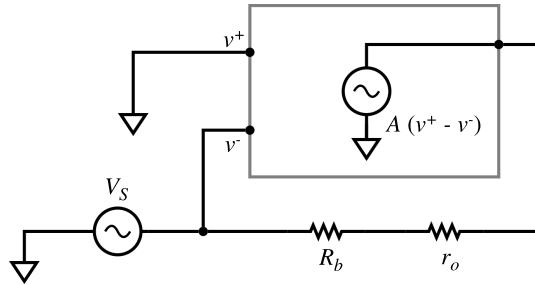


Figura 11.19: Per il teorema di Miller il circuito con amplificazione $-A$ e resistenza di reazione tra ingresso ed uscita $r_o + R_b$ presenta all'ingresso una resistenza equivalente $(r_o + R_b)/(A + 1)$.

Il primo termine è di diversi ordini di grandezza inferiore al secondo, per cui

$$R'_i \simeq \frac{r_o + R_b}{A + 1} \quad (11.36)$$

Con i valori già visti per r_o , R_b ed A , la resistenza R'_i viene ad assumere valori molto piccoli, dell'ordine da 1 fino a $10^{-3}\Omega$. Nel circuito di fig. 11.19 l'effetto della reazione negativa è stato di ridurre enormemente la resistenza di ingresso del circuito, fino a portarla praticamente a zero.

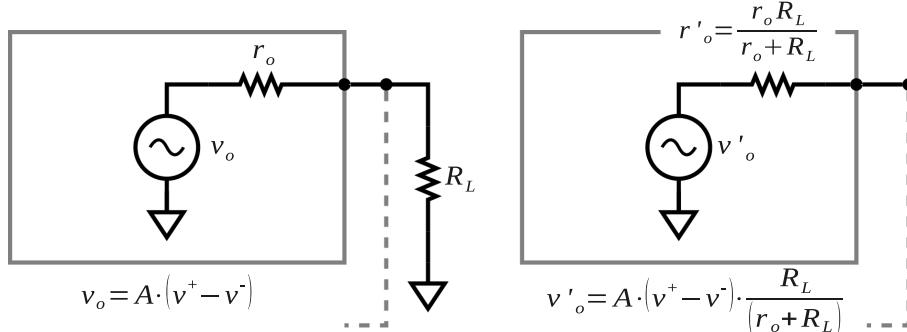


Figura 11.20: Circuito di uscita equivalente secondo Thévenin dell'amplificatore in presenza della resistenza di carico R_L . Fino a che $r_o < 0.1R_L$, l'effetto è solo di ridurre leggermente il guadagno A dell'amplificatore ed il valore effettivo di r_o .

Volendo considerare a questo punto l'effetto della R_L , si può sostituire il generatore ed il partitore formato da r_o ed R_L con il circuito equivalente secondo Thévenin (fig. 11.20). L'unico effetto sarà di ridurre lievemente il valore di r_o e di A , senza quindi modificare in maniera significativa il risultato 11.36.

Per il circuito completo della fig. 11.20 si avrà $R_i = R_a + R'_i \simeq R_a$.

11.10.3 Teorema di Miller

Nel circuito in fig. 11.21a tra l'ingresso e l'uscita di un amplificatore con amplificazione $-k$ è collegata una resistenza R . La corrente i_i che il generatore v_i deve

erogare è

$$i_i = \frac{v_i - (-kv_i)}{R} = v_i \cdot \frac{1+k}{R} \quad (11.37)$$

e quindi la resistenza che il generatore *vede* all'ingresso è

$$R_i = \frac{v_i}{i_i} = \frac{R}{1+k} \quad (11.38)$$

come indicato in fig. 11.21b. L'uscita dell'amplificatore invece eroga una corrente $i_o = (-kv_i - v_i)/R$ e quindi *vede* la resistenza

$$R_o = \frac{-kv_i}{(-kv_i - v_i)/R} = R \cdot \frac{k}{k+1} \quad (11.39)$$

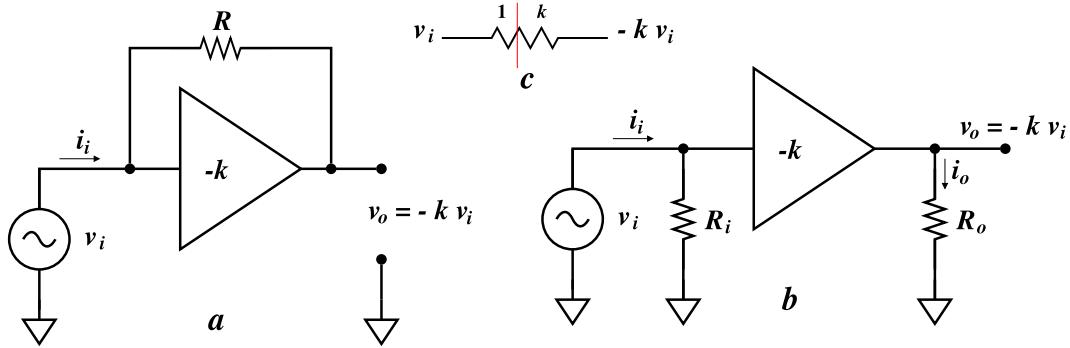


Figura 11.21: Teorema di Miller.

Il punto che divide la resistenza R in due parti in rapporto tra di loro come $1 : k$, individuato in fig. 11.21c da una barretta rossa, è una *massa virtuale* in quanto si trova sempre a potenziale 0. Nel limite $k \rightarrow \infty$ si ha $R_i \rightarrow 0$ e $R_o \rightarrow R$; l'ingresso che il circuito presenta verso il generatore v_i tende a diventare anch'esso una massa virtuale.

Una eventuale resistenza di ingresso r_i propria dell'amplificatore si viene a trovare in parallelo alla *resistenza di Miller* $R_i = R/(k+1)$ e non ha altri effetti sul circuito. Per la resistenza di uscita r_o dell'amplificatore, se ne può tenere conto come indicato nella fig. 11.19.

Capitolo 12

Applicazioni dell'amplificatore operazionale

Molte applicazioni dell'amplificatore operazionale si basano su circuiti che sono derivati da quello dell'amplificatore non invertente di fig. 11.8 o da quello dell'amplificatore invertente di fig. 11.17 e possono essere analizzati facilmente utilizzando i risultati già trovati nei paragrafi precedenti.

12.1 Convertitore *corrente-tensione*

Questo circuito ha un ingresso *in corrente* ed una uscita *in tensione*, da cui il nome di *convertitore corrente-tensione* o amplificatore a transresistenza (il coefficiente di amplificazione tra il segnale di uscita, tensione, e quello di ingresso, corrente, ha le dimensioni di una resistenza).

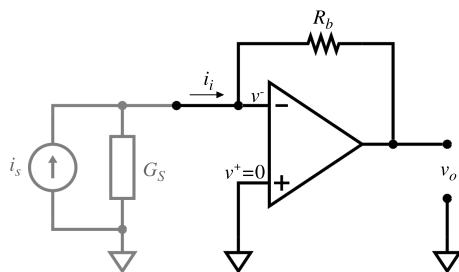


Figura 12.1: Convertitore corrente-tensione.

Lo schema, riportato in fig. 12.1, è lo stesso dell'amplificatore invertente di fig. 11.17, privato della resistenza R_a . Come dimostrato nel par. 11.10.2, la resistenza che il circuito presenta al suo terminale di ingresso è estremamente piccola, tra gli ohm ed i milliohm (eq. 11.36). Per il generatore di segnale esterno, in pratica, l'ingresso del circuito viene visto come un cortocircuito verso massa, da cui il nome di *massa virtuale* che viene assegnato al nodo v^- . In condizioni di funzionamento normale, la tensione di ingresso sulla massa virtuale è praticamente nulla. Infatti, se all'uscita è presente una tensione v_o dell'ordine dei Volt, in base alla relazione $v_o/v^- = -A$, essendo $v^+ = 0$, all'ingresso si avranno tensioni dell'ordine dei μV .

Se si considera la sorgente del segnale come un generatore di corrente i_s , la corrente di ingresso sarà $i_i = i_s$, indipendentemente dalla conduttanza G_S del generatore; se invece si descrive la sorgente come un generatore di tensione v_s con resistenza interna R_S , si avrà $i_i = v_s/R_S$. L'analisi svolta nel par. 11.10 dimostra che la corrente i_i fluisce praticamente per intero attraverso la resistenza R_b , con una frazione del tutto trascurabile attraverso la resistenza r_i di ingresso dell'amplificatore, dando quindi $v_o = -i_i R_b$.

Il convertitore corrente-tensione è utilizzato in tutte quelle situazioni in cui si deve campionare la corrente in un circuito senza introdurvi una resistenza e la relativa caduta di potenziale.

12.1.1 Una applicazione del convertitore corrente-tensione: amplificatore per fotorivelatori

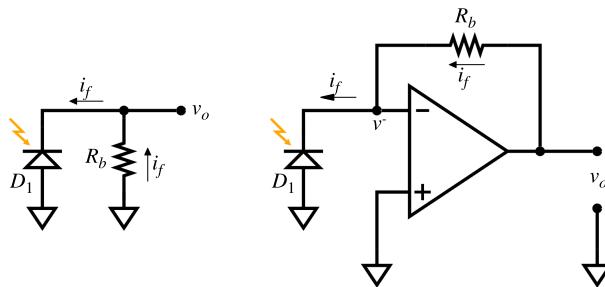


Figura 12.2: Il circuito convertitore corrente-tensione utilizzato per amplificare il segnale generato da un fotodiodo.

In fig. 12.2 il circuito di sinistra è lo schema di principio di un ricevitore di impulsi luminosi. Il fotodiodo D_1 sotto l'azione degli impulsi di luce genera una photocorrente i_f che, scorrendo nella resistenza R_b , produce una tensione $v_o = -i_f R_b$. Questo circuito non è una soluzione soddisfacente sotto due aspetti:

- La relazione tra intensità luminosa e corrente i_f (o tensione v_o) non è lineare perché la tensione v_o tende a portare il diodo in conduzione e a far fluire attraverso la giunzione stessa una parte della photocorrente (par. 4.21.2).
- La capacità del diodo forma con la resistenza R_b una costante di tempo che può rallentare sensibilmente la velocità di risposta del circuito.

Il circuito a destra nella figura risolve entrambi i problemi. La tensione di uscita è ancora $v_o = i_f R_b$ (con una inversione di segno), ma essendo il diodo chiuso su un cortocircuito, si ha $V_D = 0$ e la photocorrente scorre tutta nel circuito esterno al diodo (eq. 4.57); per lo stesso motivo la costante di tempo dovuta alla capacità del diodo viene ridotta di ordini di grandezza, aumentando sensibilmente la velocità di risposta.

12.2 Circuito sommatore - Nodo di somma

Quanto detto nel paragrafo precedente non cambia se sul nodo v^- confluiscono più rami provenienti da diversi generatori $v_1, v_2, v_3 \dots$ attraverso diverse resistenze $R_1, R_2, R_3 \dots$ (fig. 12.3).

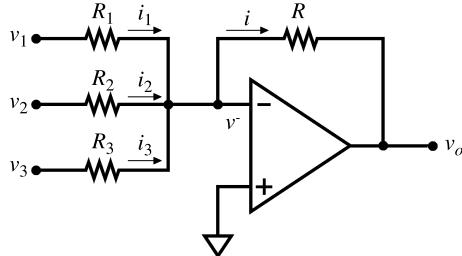


Figura 12.3: Circuito sommatore.

Per ogni ramo si avrà:

$$i_1 = v_1/R_1 \quad i_2 = v_2/R_2 \quad i_3 = v_3/R_3 \quad \dots$$

e, complessivamente:

$$i = i_1 + i_2 + i_3 \dots = \frac{v_1}{R_1} + \frac{v_2}{R_2} + \frac{v_3}{R_3} + \dots$$

$$v_o = -iR = -R \left(\frac{v_1}{R_1} + \frac{v_2}{R_2} + \frac{v_3}{R_3} + \dots \right)$$

Il circuito è una estensione dell'amplificatore invertente del par. 11.10 ed effettua la somma di più tensioni, ciascuna pesata secondo il fattore $-R/R_i$. Da questa applicazione il nodo v^- del circuito prende anche il nome di *nodo di somma*.

12.3 Amplificatore *vero differenziale*

L'amplificatore operazionale è un amplificatore differenziale, ma non può essere direttamente utilizzato per ottenere in modo accurato la differenza tra due segnali, perchè nell'espressione

$$v_o = (v^+ - v^-) \cdot A$$

la costante A è poco stabile, molto poco precisa e talmente elevata da avere scarsa utilità immediata.

Il circuito in fig. 12.4 consente invece di effettuare la differenza tra due segnali con la stessa precisione con cui il circuito di fig. 12.3 consente di effettuare la somma.

Anche questo circuito può essere analizzato sfruttando i risultati già ottenuti nel capitolo precedente. Poichè il circuito è lineare, vale il principio di sovrapposizione e si possono quindi considerare le risposte alle due tensioni v_1 e v_2 prese singolarmente e poi sommare i risultati ottenuti. Con $v_2 = 0$ il circuito equivale all'amplificatore invertente di fig. 11.17. Unica differenza è costituita dal fatto che l'ingresso + è collegato a massa tramite le due resistenze R_3 ed R_4 in parallelo. Normalmente la

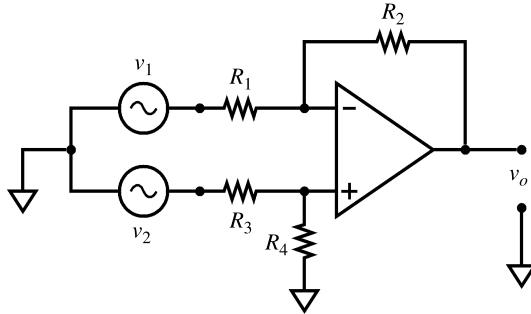


Figura 12.4: Circuito *vero differenziale* per effettuare la differenza tra due segnali: $v_o = k(v_1 - v_2)$.

caduta di potenziale dovuta alla corrente di polarizzazione del transistor di ingresso dell'amplificatore su queste due resistenze è trascurabile. Se così non fosse, basta avere l'accortezza di porre $R_1//R_2 = R_3//R_4$ perché l'effetto ad un ingresso si elida con l'altro.

Dalla eq. 11.34 si ha quindi

$$v_o = -v_1 \cdot \frac{R_2}{R_1} \quad (12.1)$$

Considerando ora di porre $v_1 = 0$ e di applicare la sola tensione v_2 , si ha l'amplificatore non invertente di fig. 11.8, con l'aggiunta del partitore R_3, R_4 all'ingresso.

Dalla eq. 11.9 si ha quindi

$$v_o = v_2 \cdot \frac{R_4}{R_3 + R_4} \cdot \frac{R_2 + R_1}{R_1} \quad (12.2)$$

Se si scelgono i valori delle resistenze in modo che $R_2/R_1 = R_4/R_3$, combinando i due contributi precedenti si ottiene:

$$v_o = v_2 \cdot \frac{R_4}{R_3 + R_4} \cdot \frac{R_2 + R_1}{R_1} - v_1 \cdot \frac{R_2}{R_1} = \frac{R_2}{R_1} \cdot (v_2 - v_1) \quad (12.3)$$

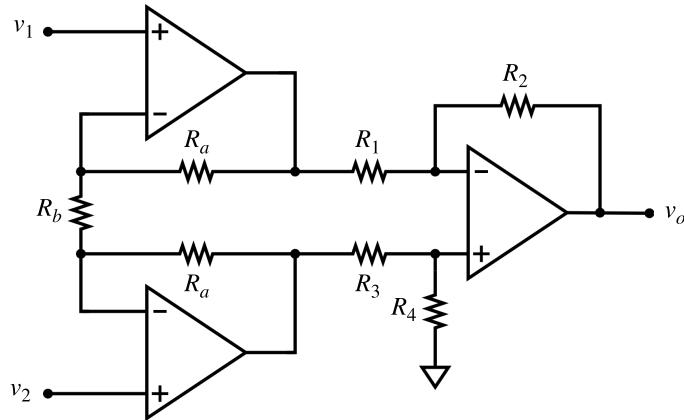
Il circuito quindi fornisce la differenza tra le due tensioni v_2 e v_1 , moltiplicata per il fattore di amplificazione R_2/R_1 . La precisione con cui è rispettata la condizione $R_2/R_1 = R_4/R_3$ determina la *Reiezione di Modo Comune* del circuito (*CMR - Common Mode Rejection*). Un segnale di modo comune $v_c = v_1 = v_2$ ai due ingressi è amplificato della quantità

$$A_c = \frac{v_o}{v_c} = \frac{R_4}{R_3 + R_4} \cdot \frac{R_2 + R_1}{R_1} - \frac{R_2}{R_1} \quad (12.4)$$

Si può aggiustare separatamente il rapporto R_2/R_1 per ottenere il valore desiderato per il guadagno ed il rapporto R_4/R_3 per azzerare l'amplificazione di modo comune.

12.4 *Instrumentation Amplifier (IA)*

L'amplificatore *vero differenziale* soffre di una notevole limitazione: l'analisi svolta nel paragrafo precedente è valida solo se le sorgenti di segnale v_1 e v_2 sono generatori di tensione con resistenza interna trascurabile. Altrimenti le resistenze interne

Figura 12.5: *Instrumentation Amplifier*

dei generatori si sommano alle resistenze R_1 ed R_3 , sbilanciando i rapporti su cui si basa il funzionamento del circuito. Il circuito in fig. 12.5 risolve questo problema, aggiungendo a ciascuno dei due ingressi un amplificatore non invertente, caratterizzato come si è visto da una elevatissima resistenza di ingresso ed una resistenza di uscita praticamente nulla. I due stadi di ingresso hanno in comune un ramo dei due partitori di reazione, la resistenza R_b . Nel funzionamento normale, per quanto visto nel par. 11.7, ai capi di R_b si viene a trovare la tensione $v_2 - v_1$. Di conseguenza la tensione che viene applicata in modo differenziale ai due ingressi del secondo stadio è $(v_2 - v_1)(2R_a + R_b)/R_b$. Il guadagno dell'intero amplificatore è quindi

$$G = \left(\frac{2R_a}{R_b} + 1 \right) \cdot \frac{R_2}{R_1} \quad (12.5)$$

ed è possibile variarlo modificando il valore di una sola resistenza, R_b .

Data la sua versatilità nei sistemi di misura, questo circuito ha preso il nome di *Instrumentation Amplifier (IA)*.

12.4.1 Ricevitore differenziale

L'amplificatore differenziale o l'*Instrumentation Amplifier* possono essere utilizzati in maniera conveniente per effettuare un trasferimento di segnali immune da interferenze, secondo il principio illustrato in fig. 12.6.

Nello schema in alto in fig. 12.6, la sorgente v_s è collegata al ricevitore A tramite una normale linea coassiale. Tra la massa della sorgente e la massa del ricevitore è presente una sorgente di rumore v_n . Il conduttore esterno del cavo coassiale, che fa da schermo di protezione per il segnale, collega anche le due masse di sorgente e ricevitore. Questo però non è sufficiente a garantire che i conduttori di massa siano veramente una rete equipotenziale a causa delle correnti che vi circolano, generate dalle sorgenti di rumore o di interferenza. Di conseguenza all'ingresso del ricevitore si ha la tensione $v_s + v_n$, somma di segnale ed interferenza, che compare eguale, eventualmente amplificata, anche all'uscita.

Nello schema in basso, invece, il ricevitore è un amplificatore differenziale e la sorgente v_s è collegata tramite una linea bifilare, eventualmente protetta da un terzo

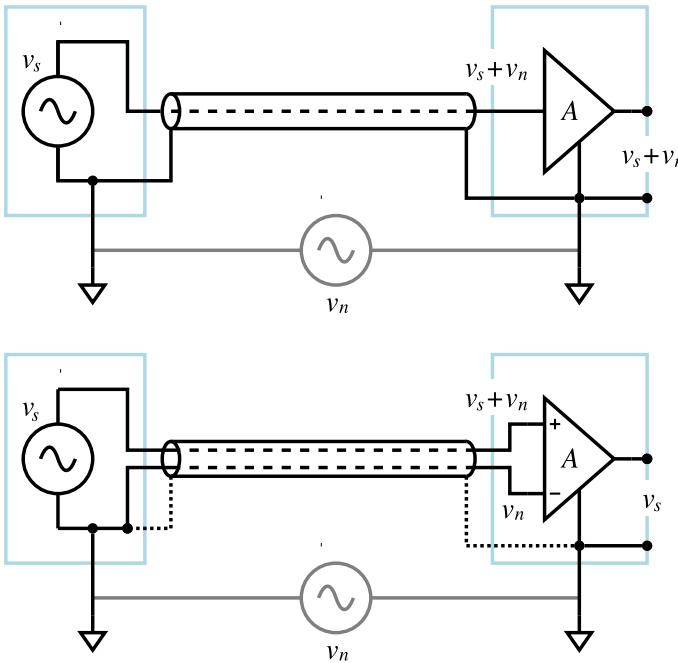


Figura 12.6:

In alto: trasmissione di un segnale mediante semplice cavo coassiale. Al ricevitore A giunge il segnale v_s sommato al segnale di rumore e interferenza v_n .

In basso: una linea bifilare porta ai due ingressi del ricevitore differenziale A sia il segnale v_s , sia la rispettiva massa, entrambi affetti nello stesso modo dal rumore v_n , che costituisce un segnale *di modo comune* e che si elide all'ingresso differenziale.

conduttore con funzione di schermo. Uno dei due conduttori della linea bifilare porta il segnale ad uno degli ingressi del ricevitore, l'altro conduttore porta la massa della sorgente all'altro ingresso. In questo modo l'amplificatore differenziale *vede* il segnale v_s riferito alla massa della sorgente, non alla propria, e può fornire all'uscita lo stesso segnale (eventualmente amplificato) e riferito alla propria massa.

Di solito il terzo conduttore, lo schermo, è collegato ad una sola estremità, o alla massa della sorgente del segnale o alla massa del ricevitore. Deve essere collegato ad entrambe le estremità ad entrambe le masse solo nel caso in cui i due apparecchi non abbiano nessuna altra connessione tra le loro masse.

Per proteggere meglio il segnale da interferenze dovute a campi elettromagnetici esterni, è possibile utilizzare una linea bifilare con i due conduttori attorcigliati tra loro (*twisted pair*), in modo che gli effetti siano eguali sui due conduttori e si elidano al momento della differenza all'ingresso del ricevitore.

Le due modalità di trasmissione hanno diversi nomi come, ad esempio, *linea bilanciata (balanced line)*, *trasmissione differenziale*, *twisted pair*, etc., la seconda, *linea asimmetrica*, *linea sbilanciata*, *single ended line*, etc., la prima.

12.5 Filtri RC

Come già visto nel cap. 1, un filtro è un circuito che amplifica in maniera selettiva i segnali contenuti nella banda di frequenze di interesse e riduce quello che è al di fuori della banda e che normalmente si considera *rumore* o *interferenza*.

Il circuito in fig. 12.7 è uno dei più semplici filtri di uso generale che si possono realizzare con gli amplificatori operazionali.

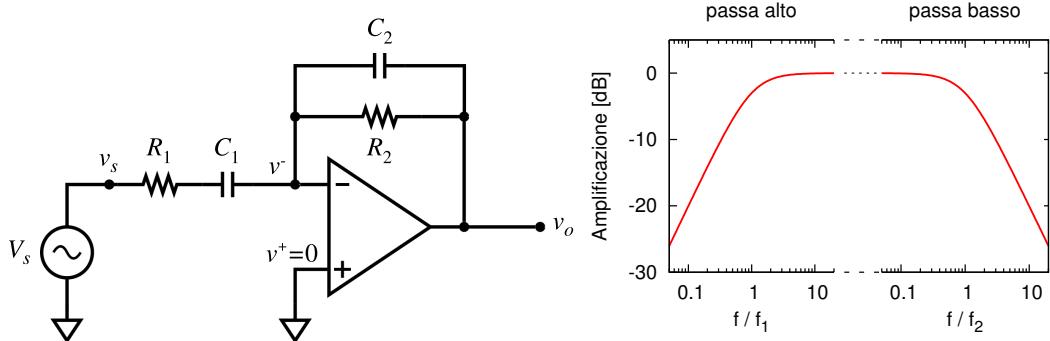


Figura 12.7: Filtro passa banda con celle RC nei rami di reazione dell'amplificatore invertente. Nel grafico a destra si vede separatamente la risposta delle due celle e quella complessiva di tutto il circuito. L'amplificazione è riferita al guadagno di *centro banda*, determinato dal rapporto R_2/R_1 (0 dB se $R_2 = R_1$).

Generalizzando ai circuiti in corrente alternata l'eq. 11.34 dell'amplificatore invertente come già visto nel par. 12.6, cioè sostituendo alla resistenza R_a l'impedenza $Z_1 = R_1 + 1/(j\omega C_1)$ ed alla resistenza R_b l'impedenza $Z_2 = R_2/(1 + j\omega R_2 C_2)$, si ottiene l'amplificazione, funzione della frequenza,

$$\frac{V_o(\omega)}{V_s(\omega)} = G(\omega) = -\frac{Z_2}{Z_1} = -\frac{1}{1 - jf_1/f} \cdot \frac{R_2}{R_1} \cdot \frac{1}{1 + jf/f_2} \quad (12.6)$$

Il primo termine nel prodotto a destra nell'eq. 12.6 rappresenta il comportamento del circuito RC *passa alto* del par. 1.6, con frequenza di taglio $f_1 = 1/(2\pi R_1 C_1)$; il terzo termine il comportamento del circuito RC di tipo *passa basso*, con frequenza di taglio $f_2 = 1/(2\pi R_2 C_2)$. Il comportamento complessivo, illustrato in fig. 12.7 è quello già visto in fig. 8.1, con $f_1 = f_L < f_2 = f_H$; il secondo termine nel prodotto (R_2/R_1) è il *guadagno a centro banda*.

12.6 Integratore di Miller

Nel circuito di fig. 12.7, nel caso limite $C_1 = \infty$, $R_2 = \infty$, il circuito diventa quello di fig. 12.8 e l'eq. 12.6 si riduce a

$$V_o(\omega) = G(\omega) \cdot V_s(\omega) = -\frac{1}{j\omega RC} \cdot V_s(\omega) \quad (12.7)$$

L'equazione 12.7 fornisce la relazione lineare tra la tensione di ingresso $V_s(\omega)$ e di uscita $V_o(\omega)$ nel dominio delle frequenze. Il risultato equivalente nel dominio del tempo si ottiene sostituendo alle funzioni $V_o(\omega)$ e $V_s(\omega)$ le loro trasformate di Fourier $v_o(t)$ e $v_s(t)$ e tenendo conto che la divisione per $j\omega$ nelle frequenze equivale alla integrazione nel tempo:

$$v_o(t) = -\frac{1}{RC} \int v_s(t) dt \quad (12.8)$$

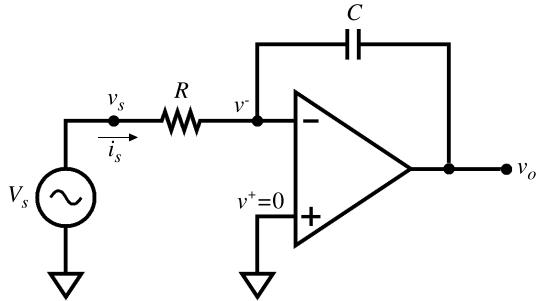


Figura 12.8: Circuito integratore di Miller.

Allo stesso risultato si può arrivare rapidamente utilizzando il concetto di massa virtuale per il nodo v^- . Poiché $v^- \simeq 0$, la corrente attraverso la resistenza R ed il condensatore C è $i_s = v_s/R = -C \cdot dv_o/dt$, da cui

$$\frac{dv_o}{dt} = -\frac{v_s}{RC} \quad (12.9)$$

equivalente alla eq. 12.8.

La tensione di uscita è proporzionale all'integrale nel tempo della tensione di ingresso, con costante di proporzionalità $1/RC$. Da questo la denominazione *integratore* che identifica il circuito. Le analogie con il circuito di fig. 11.21 sono invece all'origine della denominazione *di Miller*.

Le equazioni 12.7 e 12.8 contengono implicitamente le approssimazioni già viste e introdotte nel capitolo precedente e descrivono un comportamento ideale. Per esempio, l'eq. 12.7 prevede un guadagno che va a infinito quando ω tende a zero. Nel comportamento reale del circuito, riportato in fig. 12.9, si vede invece che il

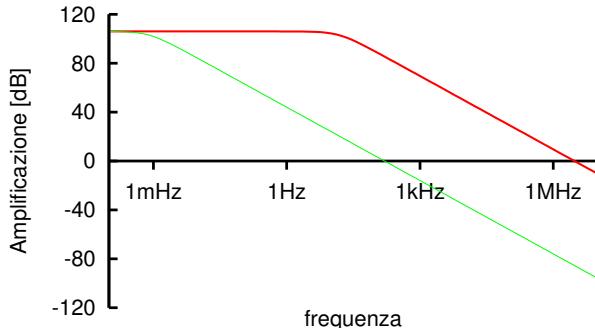


Figura 12.9: In verde: risposta in frequenza di un circuito integratore di Miller con un amplificatore operazionale TL081 e $RC = 1 \text{ ms}$. In rosso: risposta del solo amplificatore operazionale. L'integratore si comporta in maniera praticamente ideale da 1 mHz fino ad oltre 1 MHz . Al di sotto di 1 mHz l'amplificazione si ferma al valore per la corrente continua $A = 2 \cdot 10^5$ (106 dB).

guadagno cresce al diminuire di ω fino al valore A dell'amplificazione ad anello aperto e non oltre. L'andamento è quello tipico di un circuito passa-basso, con frequenza di taglio $f = [2\pi RC(A+1)]^{-1}$. Queste affermazioni possono essere verificate facilmente ripartendo dall'eq. 11.33 e ripercorrendo il cammino, con l'impedenza di C per R_b .

12.7 Differenziatore

In linea di principio dovrebbe essere possibile realizzare il circuito di fig. 12.10 per svolgere la funzione inversa di quello di fig. 12.8, cioè la derivata rispetto al tempo del segnale di ingresso. Sostituendo nella eq. 11.34 l'impedenza del condensatore C

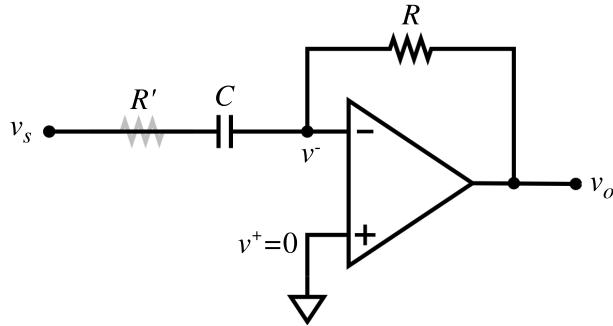


Figura 12.10: Schema di principio di un circuito *differenziatore*.

alla resistenza R_a si ottiene

$$V_o(\omega) = -j\omega RC \cdot V_s(\omega) \quad (12.10)$$

e quindi passando al dominio del tempo

$$v_o(t) = -RC \frac{d}{dt} v_s(t) \quad (12.11)$$

In pratica le cose non vanno così. A differenza dell'integratore, che è un circuito intrinsecamente stabile, il differenziatore si trova sempre ad essere instabile o al limite della instabilità ed è quindi inutilizzabile senza accorgimenti appropriati. L'aggiunta di una seconda resistenza in serie al condensatore (in grigio nella figura) riduce il problema dell'instabilità, ma anche l'efficienza del circuito.

Le ragioni di questo comportamento saranno chiarite dall'analisi generale del problema della stabilità degli anelli di reazione.

12.8 NIC (*Negative impedance converter*)

In questo circuito, come negli altri esempi che seguono, sono presenti contemporaneamente due canali di reazione: la reazione negativa, che determina il comportamento dell'amplificatore operazionale come già visto nei paragrafi precedenti, ed un secondo canale, di reazione positiva. Questa combinazione permette di ottenere comportamenti molto interessanti. In ogni caso, per mantenere la stabilità del circuito, il contributo di reazione positiva deve essere minore di quello di reazione negativa.

L'amplificatore A in fig. 12.11 insieme con le resistenze R_1 ed R_2 ed ignorando R_3 è un amplificatore non invertente per quanto riguarda il segnale applicato dal generatore v_s all'ingresso +. Quindi $v_o = (1 + R_1/R_2) \cdot v_s$.

A causa della presenza di R_3 , il generatore si trova ad erogare una corrente i_s data da

$$i_s = \frac{1}{R_3} \cdot (v_s - v_o) = \frac{1}{R_3} \left[v_s - \left(1 + \frac{R_1}{R_2} \right) \cdot v_s \right] = -\frac{v_s}{R_3} \cdot \frac{R_1}{R_2} \quad (12.12)$$

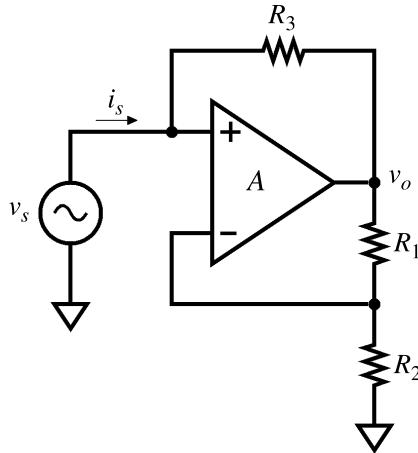


Figura 12.11: NIC: Negative Impedance Converter.

e di conseguenza la resistenza di ingresso R vista dal generatore v_s è

$$R = \frac{v_s}{i_s} = -R_3 \cdot \frac{R_2}{R_1} \quad (12.13)$$

cioè una resistenza negativa. Nel caso $R_1 = R_2$, si ha $R = -R_3$.

Il circuito è stabile fino a che il termine di reazione negativa $R_2/(R_1 + R_2)$ è maggiore di quello di reazione positiva $R_G/(R_G + R_3)$, dove R_G è la resistenza interna del generatore, che può essere quindi un generatore di tensione con resistenza interna $R_G < R_2 R_3 / R_1$, ma non un generatore di corrente ideale.

La reazione positiva può essere fatta diventare selettivamente superiore a quella negativa, come nel circuito del par. 12.11, ottenendo così un *generatore di segnali* (vedi anche cap. 15 e 17).

12.9 Generatore di corrente (Pompa di Howland)

Una applicazione del circuito NIC è la *pompa di corrente di Howland*, illustrata in fig. 12.12. Se si pone $R_4 = R_3 R_2 / R_1 = R$, il circuito equivale a quello a destra nella figura e per la corrente attraverso R_L si ha:

$$i_L = v_s \frac{\frac{R_L \cdot (-R)}{R_L - R}}{R + \frac{R_L \cdot (-R)}{R_L - R}} \cdot \frac{1}{R_L} = \frac{v_s}{R} \quad (12.14)$$

La corrente i_L dipende da v_s e da R , ma non da R_L . Il circuito si comporta come un generatore di corrente ideale. La stabilità è garantita purché $R_L \neq \infty$.

12.10 Integratore non invertente

Sostituendo nella *pompa di Howland* un condensatore C alla resistenza R_L si ottiene il circuito integratore di Deboo (fig. 12.13). A differenza dell'integratore di

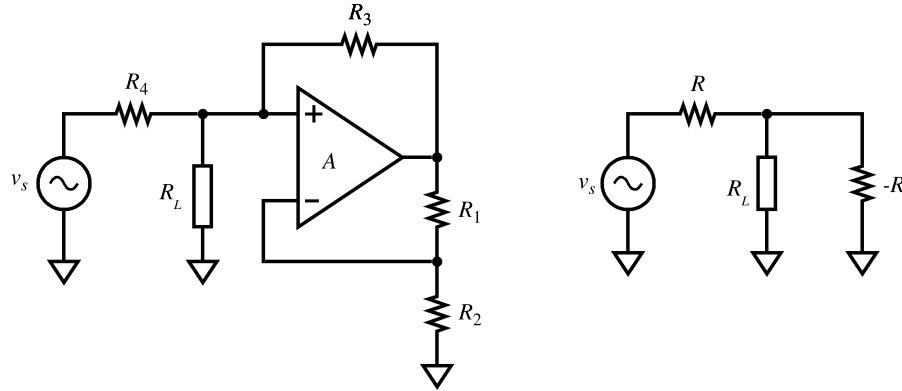


Figura 12.12: *Pompa di corrente di Howland*. La corrente attraverso R_L dipende da v_s e da $R_{1\ldots 4}$, ma non da R_L e quindi il circuito si comporta come un generatore di corrente. A destra: circuito equivalente, con il circuito NIC sostituito dalla sua resistenza negativa equivalente.

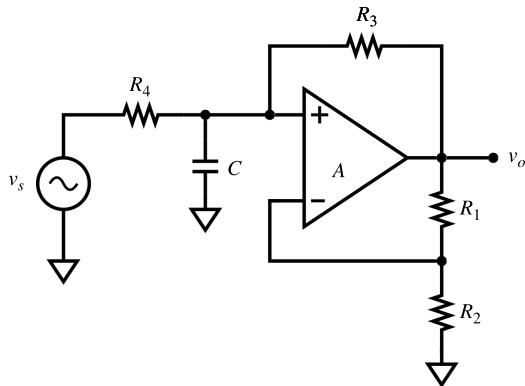


Figura 12.13: Integratore non invertente di Deboo.

Miller, in questo circuito non è presente l'inversione di polarità del segnale tra ingresso e uscita, per cui questo circuito è noto anche come *integratore non invertente*.

Combinando i risultati ottenuti dall'analisi dei circuiti precedenti con la condizione $R_1 = R_2 = R_3 = R_4 = R$ si ottiene facilmente la relazione

$$v_o = \frac{2}{RC} \int v_s dt$$

12.11 Oscillatore a resistenza negativa

Il circuito di fig. 12.14 utilizza un circuito NIC per produrre una resistenza negativa che, posta in parallelo ad un circuito risonante LCR_s , compensa gli effetti dissipativi della resistenza R_s e dà luogo ad una oscillazione di ampiezza crescente invece che smorzata.

Per un circuito risonante con un fattore di merito Q sufficientemente elevato, l'effetto dissipativo della resistenza R_s in serie con l'induttanza L è, alla risonanza, circa eguale a quello di una resistenza in parallelo $R_p = R_s \cdot Q^2$ (vedi par. 1.8.2).

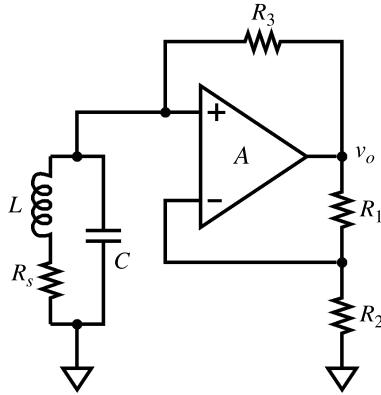


Figura 12.14: Circuito NIC utilizzato per la generazione di segnali persistenti.

Se $-|R|$ è la resistenza negativa presentata dal circuito NIC, la resistenza complessiva di smorzamento in parallelo al circuito risonante è:

$$R_d = \frac{|R| \cdot R_p}{|R| - R_p} \quad (12.15)$$

che diventa infinita per $|R| = R_p$ e negativa per $|R| < R_p$.

12.12 Filtri attivi

Le induttanze sono i componenti passivi più scomodi da realizzare nella pratica, specialmente nei circuiti in bassa frequenza, a causa di: ingombro, peso quando è richiesto un nucleo magnetico, non linearità introdotte dal nucleo stesso, difficoltà a tenere basse le perdite ed ottenere fattori di merito elevati, necessità di tenere il campo magnetico intrappolato all'interno del componente per evitare accoppiamenti con altre induttanze o altri circuiti in genere.

Utilizzando circuiti contenenti una componente di reazione positiva come visto nei paragrafi precedenti, è possibile realizzare circuiti di filtro, detti *filtri attivi*, in cui la funzione della induttanza è sostituita dalla presenza di condensatori, amplificazione e reazione.

12.12.1 Filtri Sallen-Key

In fig. 12.15 sono riportati tre esempi filtri attivi di tipo *Sallen-Key*, che sintetizzano ciascuno una delle tre funzioni di risposta del circuito risonante descritte dalle eq. 1.33-1.35: passa basso, passa alto e passa banda.

Per studiare il funzionamento di questi circuiti conviene scrivere la prima equazione di Kirchhoff per il nodo V_n . Ad esempio, per il circuito passa-basso (l_p), si ha

$$\frac{V_n - V_i}{R_1} + j\omega C_1 \cdot (V_n - V_o) + V_n \cdot \frac{j\omega C_2}{1 + j\omega C_2 R_2} = 0 \quad (12.16)$$

Tra V_n e V_o vale la relazione $V_o = GV_n/(1 + j\omega R_2 C_2)$ con $G = (R_3 + R_4)/R_4$. Combinando le espressioni e riarrangiando i termini, si ottiene

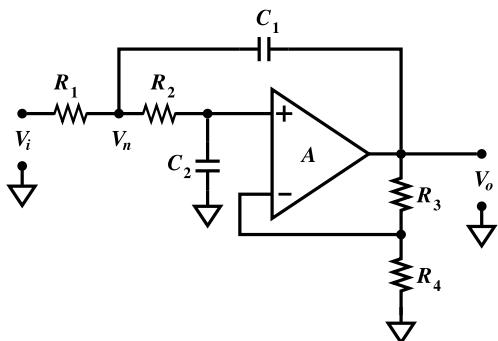
$$\left. \frac{V_o}{V_i} \right|_{l_p} = \frac{G}{1 - \omega^2 C_1 C_2 R_1 R_2 + j\omega [R_2 C_2 + R_1 C_2 + R_1 C_1 (1 - G)]} \quad (12.17)$$

Analogamente, per il passa alto (hp) ed il passa banda (bp):

$$\left. \frac{V_o}{V_i} \right|_{hp} = \frac{-G\omega^2 C_1 C_2 R_1 R_2}{1 - \omega^2 C_1 C_2 R_1 R_2 + j\omega [R_1 C_1 + R_1 C_2 + R_2 C_2 (1 - G)]} \quad (12.18)$$

$$\left. \frac{V_o}{V_i} \right|_{bp} = \frac{Gj\omega C_2 R_2}{1 + R_1/R_5 - \omega^2 C_1 C_2 R_1 R_2 + j\omega \{C_1 R_1 + C_2 R_1 + C_2 R_2 [1 + R_1/R_5 (1 - G)]\}} \quad (12.19)$$

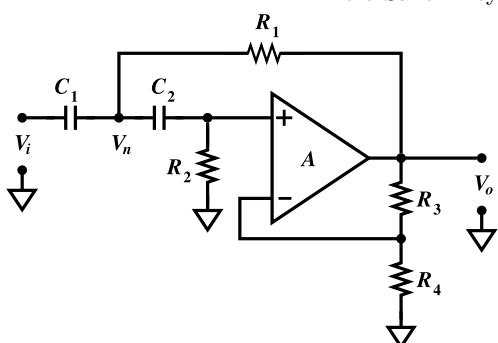
Queste equazioni, a meno di un coefficiente moltiplicativo, coincidono con le 1.33-1.35, con i valori di ω_0 e Q_0 riportati in fig. 12.15, con $\mathbb{R} = R_1/R_2$ e $\mathbb{C} = C_2/C_1$.



filtro Sallen-Key di tipo passa-basso

$$\omega_0 = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}}$$

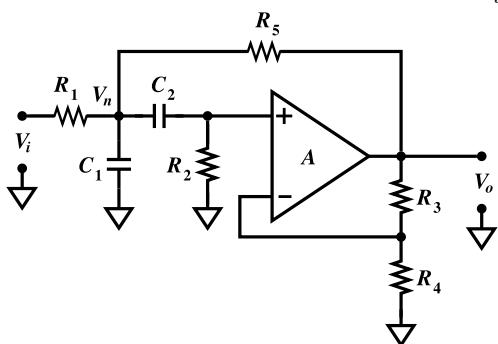
$$Q_0 = \frac{1}{(1 - G) \sqrt{\mathbb{R}/\mathbb{C}} + \sqrt{\mathbb{R}\mathbb{C}} + \sqrt{\mathbb{C}/\mathbb{R}}}$$



filtro Sallen-Key di tipo passa-alto

$$\omega_0 = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}}$$

$$Q_0 = \frac{1}{(1 - G) \sqrt{\mathbb{C}/\mathbb{R}} + \sqrt{\mathbb{R}\mathbb{C}} + \sqrt{\mathbb{R}/\mathbb{C}}}$$



filtro Sallen-Key di tipo passa-banda

$$\omega_0 = \frac{\sqrt{1 + R_1/R_5}}{\sqrt{R_1 R_2 C_1 C_2}}$$

$$Q_0 = \frac{\sqrt{1 + R_1/R_5}}{(1 - G') \sqrt{\mathbb{C}/\mathbb{R}} + \sqrt{\mathbb{R}\mathbb{C}} + \sqrt{\mathbb{R}/\mathbb{C}}}$$

$$G' = (R_1/R_5)(R_3/R_4)$$

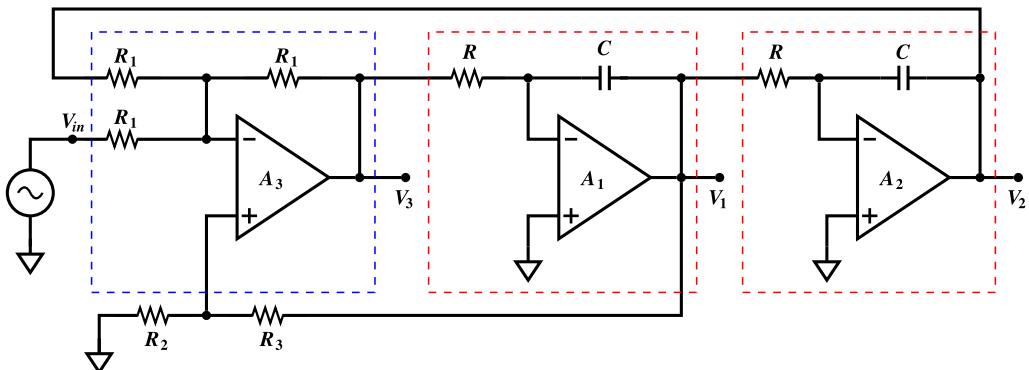
Figura 12.15: Filtri attivi di tipo *Sallen-Key* passa basso, passa alto e passa banda.

Nei primi due casi il valore di ω_0 può essere modificato indipendentemente da Q_0 variando contemporaneamente R_1 ed R_2 , mantenendone costante il rapporto, ed il valore di Q_0 può essere variato senza influire su ω_0 modificando il valore di G , variando contemporaneamente anche l'amplificazione del circuito. Ad esempio, con $\mathbb{R} = \mathbb{C} = 1$ (condizione usata di frequente per ragioni di semplicità nella scelta dei componenti) si ha $Q_0 = 1/(3 - G)$.

12.12.2 Filtro KHN

I filtri Sallen-Key utilizzano un numero minimo di componenti per ottenere la funzione di risposta desiderata, al prezzo di una certa rigidezza nella scelta dei valori.

Il filtro KHN (dalle iniziali degli inventori Kerwin, Huelsman, Newcomb), il cui schema è riportato in fig. 12.16, a fronte di una maggiore complessità, è in grado di fornire contemporaneamente tutte e tre le funzioni di risposta, passa basso (uscita V_3), passa banda (uscita V_1) e passa alto (uscita V_2), ed offre la caratteristica molto interessante di poter variare facilmente non solo la frequenza di risonanza ω_0 variando simultaneamente le due resistenze R (per esempio tramite un potenziometro doppio) ma anche la selettività, variando il fattore di merito Q_0 tramite il rapporto R_3/R_2 .



Utilizzando per semplicità la stessa costante di tempo $\tau = \tau_1 = \tau_2 = RC$ per i due integratori, il circuito di fig. 12.16 è descritto dal sistema di equazioni

$$\begin{aligned} V_1 &= -\frac{1}{j\omega\tau} \cdot V_3 \\ V_2 &= -\frac{1}{j\omega\tau} \cdot V_1 \\ V_3 &= -V_2 + \frac{V_1}{Q_0} - V_{in} \\ Q_0 &= \frac{1}{3} \left(1 + \frac{R_3}{R_2} \right) \end{aligned} \tag{12.20}$$

Con la sostituzione $\omega_0 = 1/\tau$, si ottengono le tre funzioni di trasferimento:

$$\begin{aligned}\frac{V_2}{V_{in}} &= \frac{-1}{1 - \frac{\omega^2}{\omega_0^2} + \frac{j}{Q_0} \frac{\omega}{\omega_0}} \\ \frac{V_1}{V_{in}} &= \frac{j \frac{\omega}{\omega_0}}{1 - \frac{\omega^2}{\omega_0^2} + \frac{j}{Q_0} \frac{\omega}{\omega_0}} \\ \frac{V_3}{V_{in}} &= \frac{\frac{\omega^2}{\omega_0^2}}{1 - \frac{\omega^2}{\omega_0^2} + \frac{j}{Q_0} \frac{\omega}{\omega_0}}\end{aligned}\quad (12.21)$$

che, a parte i segni ed un coefficiente Q_0 , coincidono con le eq. 1.33-1.35 del circuito risonante.

12.13 Il filtro *passa tutto*

Un esempio di filtro *passa tutto* è riportato nella fig. 12.17a. Il circuito è lo stesso

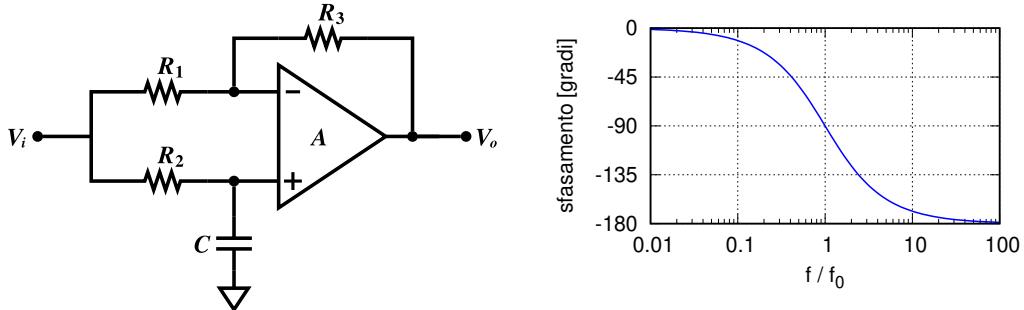


Figura 12.17: Schema elettrico (a sinistra) e sfasamento (a destra) di un filtro *passa tutto*. Il modulo della funzione di trasferimento è 1 per tutte le frequenze.

dell'amplificatore *vero differenziale* del par. 12.3, con il condensatore C al posto della resistenza R_4 e con lo stesso segnale V_i applicato contemporaneamente ad entrambi gli ingressi. Ponendo $R_1 = R_2 = R_3 = R$ ed $R_4 = (j\omega C)^{-1}$, dalla eq. 12.4 si ottiene la funzione di trasferimento

$$H(\omega) = \frac{V_o}{V_i} = \frac{1/(j\omega C)}{R + 1/(j\omega C)} \cdot 2 - 1 = \frac{1 - j\omega RC}{1 + j\omega RC} \quad (12.22)$$

È immediatamente evidente da questo risultato che $|H(\omega)| = 1$ per tutte le frequenze. Il filtro *passa tutto* non attenua né esalta alcuna frequenza; da questo il suo nome. La sua funzione è piuttosto di modificare la fase di un segnale in funzione della frequenza, lasciandone inalterata l'ampiezza. Infatti

$$\arg(H(\omega)) = \varphi = -2 \arctan(\omega RC) \quad (12.23)$$

Alla frequenza caratteristica $f_0 = (2\pi RC)^{-1}$ si ha $\varphi = -90^\circ$; per frequenze inferiori a f_0 la fase ruota e tende asintoticamente a 0° ; per frequenze superiori a f_0 tende asintoticamente a 180° (fig. 12.17).

Nel par. 19.6.1 si vedrà una applicazione di questi filtri, la modulazione SSB, in cui è necessario disporre di due copie di un segnale, che contengano le stesse componenti armoniche, della stessa ampiezza, ma con uno sfasamento tra una copia e l'altra di 90° per ciascuna componente.

Nella fig. 12.18 (a sinistra) sono riportati gli sfasamenti ottenuti con due celle passa tutto con frequenze caratteristiche $f_1 = 400 \text{ Hz}$ e $f_2 = 2500 \text{ Hz}$. Si vede dal grafico che in un piccolo intervallo di frequenza intorno a 1000 Hz (circa mezza decade) la differenza di fase tra le uscite delle due celle rimane intorno a 90° .

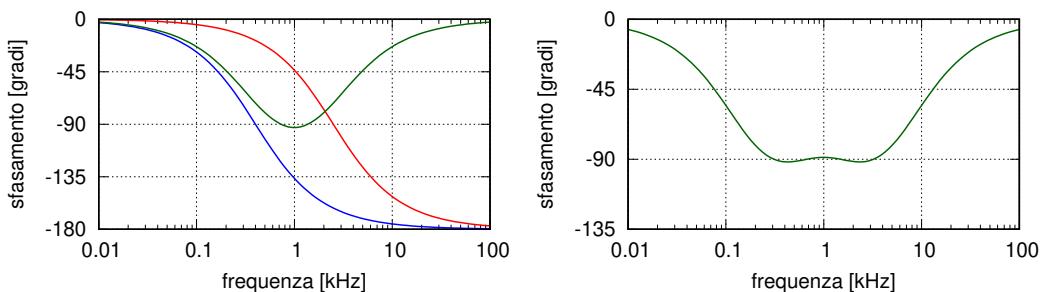


Figura 12.18: A sinistra sono riportati gli sfasamenti introdotti da due celle di tipo *passa-tutto* con $f_1 = 1/(2\pi RC) = 400 \text{ Hz}$ (curva blu) ed $f_2 = 2500 \text{ Hz}$ (curva rossa), insieme con la loro differenza (curva verde). A destra è riportata la differenza tra gli sfasamenti tra due serie di due celle, con frequenze ottimizzate per coprire circa una decade intorno ad 1 kHz , con la migliore approssimazione possibile.

Per ottenere lo stesso risultato su intervalli di frequenza più ampi e migliorare contemporaneamente la precisione con cui la differenza di fase approssima i 90° costanti è possibile combinare più coppie di filtri passa-tutto con frequenze caratteristiche opportune. Nella fig. 12.18 (a destra) è riportato come esempio la differenza di fase tra le uscite di due serie di due celle. Si ottiene una differenza di fase quasi costante nell'intervallo di frequenze $300 \dots 3000 \text{ Hz}$.

Prestazioni ancora migliori si ottengono aumentando il numero delle celle in ciascuna delle due serie.

12.14 Comparatore

Il comparatore è un circuito non lineare con un ingresso di tipo analogico ed una uscita di tipo digitale. Come dice il nome, il comparatore confronta la tensione di ingresso V_{in} con una tensione di riferimento V_{ref} e fornisce in uscita una variabile logica di un bit che indica tramite il suo valore 0/1 se V_{in} è maggiore o minore di V_{ref} . Il comparatore è quindi un convertitore Analogico → Digitale ad un bit.

Un amplificatore operazionale senza rete di reazione si comporta, in modo un po' rozzo, come un comparatore. Nel circuito di fig. 12.19 l'uscita dell'amplificatore rimane al valore della tensione di saturazione positiva (V_s^+ , valore logico 1) fino a che $V_{in} < V_{ref}$; nell'intorno di $V_{in} \simeq V_{ref}$, in un intervallo di pochi μV , si ha il passaggio

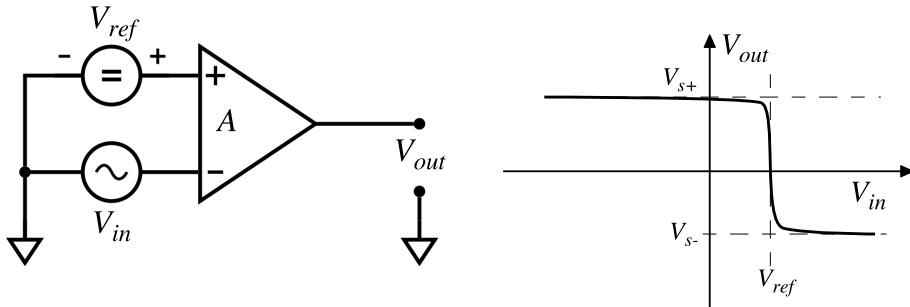


Figura 12.19: Amplificatore operazionale utilizzato come comparatore di tensione.

molto ripido di V_{out} da V_s^+ alla tensione di saturazione negativa V_s^- (valore logico 0), valore che si mantiene fino a che $V_{in} > V_{ref}$.

12.14.1 Comparatore con isteresi (*trigger di Schmitt*)

Nel comparatore di fig. 12.19 la transizione da uno stato all'altro dell'uscita avviene in corrispondenza di un intervallo di tensioni di ingresso piccolo, ma non nullo, introducendo un margine di incertezza nel valore esatto della tensione di ingresso V_{in} che viene letta come maggiore o minore di V_{ref} .

Questo non sarebbe un problema di per se: corrisponde semplicemente all'errore sperimentale con cui possono essere misurate V_{in} e V_{ref} .

Si può avere invece un problema anche serio se la tensione V_{in} , variando nel tempo in prossimità di V_{ref} si viene a trovare più volte sopra e sotto la soglia a causa del rumore sempre inevitabilmente associato ad ogni segnale, come si vede in fig. 12.20. Questo fa comparire nel segnale digitale di uscita un treno di commutazioni casuali invece del singolo cambio di stato da 0 a 1 o viceversa, che dovrebbe identificare l'evento.

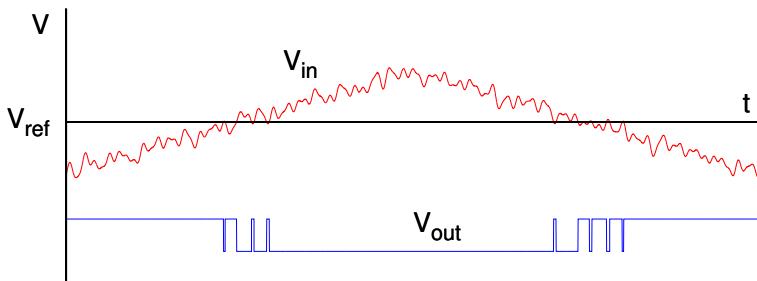


Figura 12.20: Un segnale che varia nel tempo può dare origine a commutazioni multiple a causa del rumore.

Il circuito di fig. 12.21 risolve questo problema utilizzando la reazione positiva per introdurre nel circuito comparatore un certo livello di isteresi.

Il partitore formato dalle due resistenze R_1 ed R_2 aggiunge alla tensione di soglia V_{ref} una frazione $R_1/(R_1 + R_2)$ della tensione di uscita V_{out} . Appena la tensione di ingresso V_{in} salendo raggiunge la tensione di soglia $V_{ref} + V_{hh}$ ed avviene la prima commutazione, la tensione V_{out} passa da V_{oh} a V_{ol} e di conseguenza la tensione di

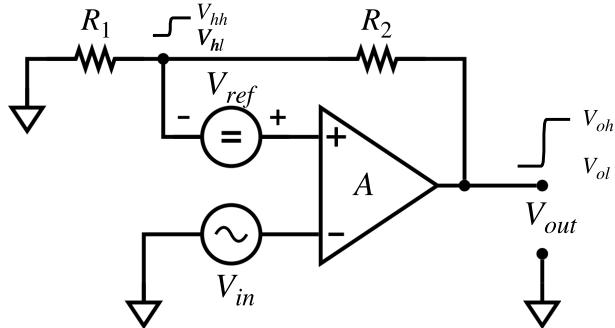


Figura 12.21: Circuito comparatore con isteresi (trigger di Schmitt).

soglia salta a $V_{ref} + V_{hl}$. La differenza tra V_{hh} e V_{hl} è una quantità piccola, ma più grande delle fluttuazioni dovute al rumore, ed ha polarità tale da confermare la commutazione e renderla irreversibile. Il processo è illustrato qualitativamente nella figura 12.22, insieme al processo complementare che si verifica quando la tensione V_{in} torna a scendere e raggiunge la nuova soglia $V_{ref} + V_{hl}$.

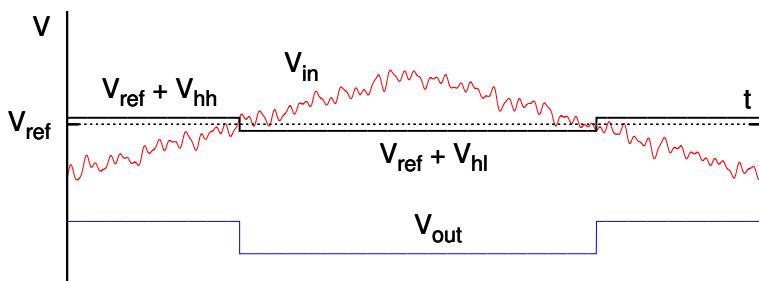


Figura 12.22:
L'aggiunta di una componente di isteresi al comparatore lo rende immune dalle commutazioni multiple dovute al rumore.

Con questo tipo di circuito il valore della tensione di soglia è diverso a seconda che la commutazione avvenga nella fase di salita o di discesa della tensione V_{in} : il circuito presenta isteresi.

Il primo comparatore elettronico con isteresi, realizzato con tubi a vuoto nel 1938, prese il nome dal suo inventore O.H. Schmitt. Il nome si è poi esteso a tutti i comparatori di tensione con isteresi, qualunque fosse la tecnologia utilizzata, tubi a vuoto, transistor o amplificatori operazionali.

12.15 Generatore di funzioni

Combinando insieme come indicato in fig. 12.23 un circuito integratore (ad es., l'integratore di Miller) con due comparatori ed una sorgente di tensione comandata si può realizzare un *generatore di funzioni*, circuito in grado di produrre onde quadre e triangolari (fig. 12.24). Da queste ultime è possibile ricavare anche segnali sinusoidali.

L'integratore genera alla sua uscita una rampa V_t a pendenza costante in salita o in discesa, a seconda del segno $-$ o $+$ della tensione costante V_i all'ingresso. Quando la rampa raggiunge il livello V_c^\pm di uno dei due comparatori, genera un impulso (C_1 o C_2 , a seconda della pendenza) che provoca il ribaltamento del flip-flop, il

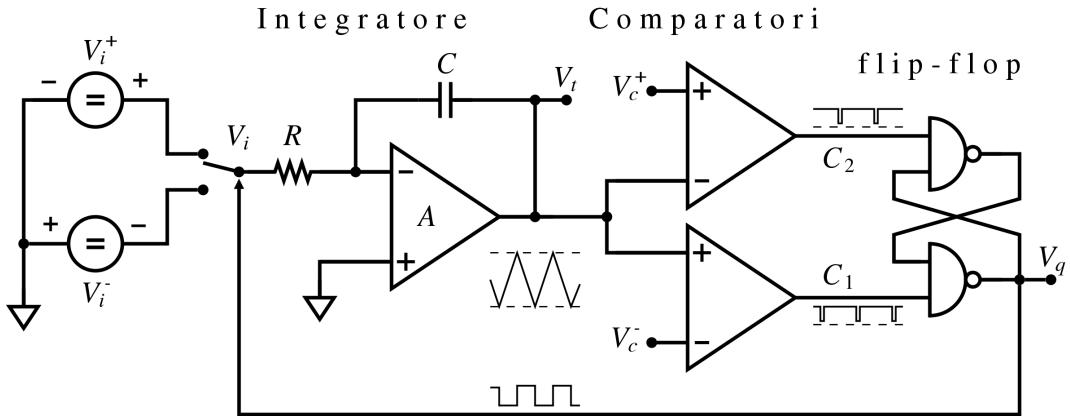


Figura 12.23: Schema a blocchi di un generatore di funzioni.

cambio della polarità di V_i e l'inversione della pendenza della rampa. Variando R e C è possibile cambiare con continuità la pendenza e quindi la frequenza dell'onda triangolare prodotta.

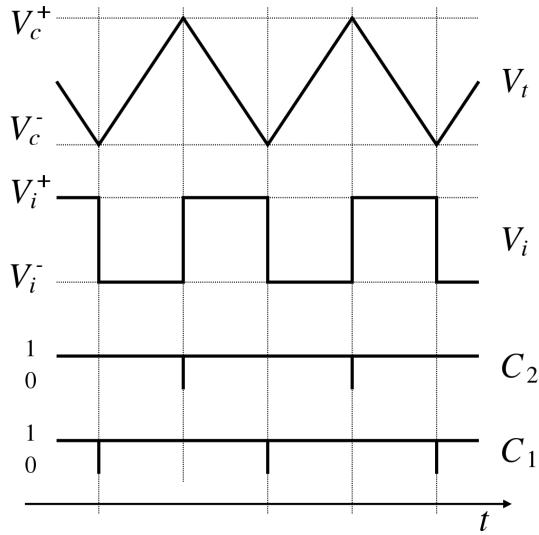


Figura 12.24: Il circuito integratore genera le rampe dell'onda triangolare V_t , in salita o in discesa, a seconda della polarità del generatore V_i selezionato. La tensione all'uscita dell'integratore continua a crescere o a diminuire fino a che non raggiunge il valore limite V_c^+ o V_c^- . A questo punto il comparatore corrispondente al limite raggiunto scatta generando un impulso (C_1 o C_2) e fa cambiare lo stato del flip-flop di tipo *set-reset*, che controlla la selezione del generatore di tensione V_i . La tensione applicata all'ingresso dell'integratore cambia di segno, e così la pendenza della rampa all'uscita, completando il ciclo che poi si ripete indefinitamente.

Segnali sinusoidali

Da un'onda triangolare di ampiezza $\pm V_t$ è possibile ottenere con discreta precisione un'onda sinusoidale di ampiezza $\pm V_s$ utilizzando una rete non lineare che abbia la funzione di trasferimento tra ingresso e uscita:

$$V_o = V_s \cdot \sin \left(\frac{\pi}{2} \frac{V_i}{V_t} \right)$$

Lo schema di principio in fig. 12.25 a sinistra mostra un esempio di come può essere realizzata una tale rete, il grafico a destra ne riporta la funzione di trasferimento (in

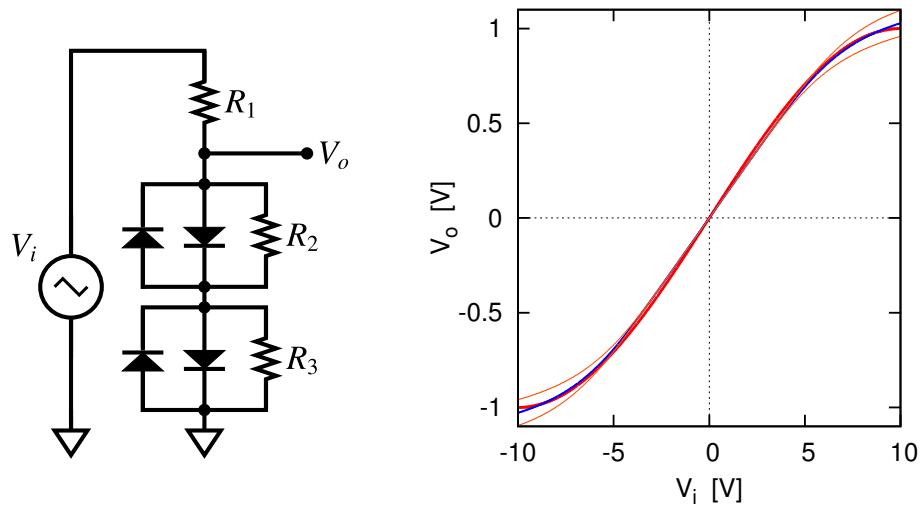


Figura 12.25: Circuito non-lineare per la trasformazione di una rampa lineare in una funzione seno. La funzione $V_o(V_i)$ è stata calcolata considerando quattro diodi uguali mod. 1N4148, con $R_1 = 123 \text{ k}\Omega$, $R_2 = 8.6 \text{ k}\Omega$ ed $R_3 = 11.9 \text{ k}\Omega$.

blu), a confronto con l'andamento della funzione ideale desiderata. Lo scarto massimo tra le due curve non supera il 3%. Nella pratica è necessario utilizzare reti più complesse, contenenti un numero maggiore di elementi non lineari, per ottenere livelli di precisione maggiori. Inoltre è necessario provvedere circuiti di compensazione per eliminare i problemi legati alle variazioni di temperatura. Le due curve a tratto sottile di colore arancione riportano la funzione di trasferimento a temperature differenti di $\pm 20^\circ\text{C}$ rispetto a quella nominale.

Capitolo 13

Elaborazione analogica

In questi appunti si è già visto più volte come una rete può essere descritta da un opportuno sistema di equazioni. Risolvendo le equazioni si ottiene la descrizione del funzionamento della rete stessa.

Questa considerazione può essere anche rovesciata: se si costruisce una rete che è descritta da un certo sistema di equazioni, l'andamento temporale di tensioni e correnti durante il funzionamento della rete fornisce la soluzione al sistema di equazioni. Questo è il principio alla base del funzionamento dei *calcolatori analogici*, macchine di calcolo che hanno avuto un certo ruolo intorno alla metà del secolo scorso.

Successivamente, l'enorme sviluppo del calcolatore digitale ha reso i calcolatori analogici decisamente poco convenienti. Il costo attuale, molto basso, delle risorse di calcolo numerico permette di ottenere tramite il calcolatore digitale le soluzioni che interessano, praticamente ad ogni livello di precisione desiderato.

13.1 Il Calcolatore Analogico

I due elementi principali che compongono un calcolatore analogico sono il circuito sommatore, descritto nel par. 12.2, e l'integratore di Miller, descritto nel par. 12.6.

Utilizzando questi due elementi è possibile per ogni sistema di equazioni differenziali lineari realizzare una rete che ne fornisce le soluzioni con il suo funzionamento.

Un esempio semplice di calcolatore analogico, composto da due integratori e da un circuito di somma, è riportato nella fig. 13.1. Tra le tensioni presenti nel circuito valgono le relazioni

$$v_1 = -\frac{1}{\tau_1} \int v_0 dt \quad v_2 = -\frac{1}{\tau_2} \int v_1 dt \quad (13.1)$$

$$v_3 = -(v_2 + v_a + v_{in}) \quad (13.2)$$

con $\tau_1 = R_1 C_1$ e $\tau_2 = R_2 C_2$.

Chiudendo il circuito con il conduttore tratteggiato in alto nella figura si aggiunge la condizione $v_0 = v_3$.

Trasformando le equazioni integrali in equazioni differenziali, combinando le tre espressioni e supponendo, per ora, che per la rete f_a valga la relazione lineare

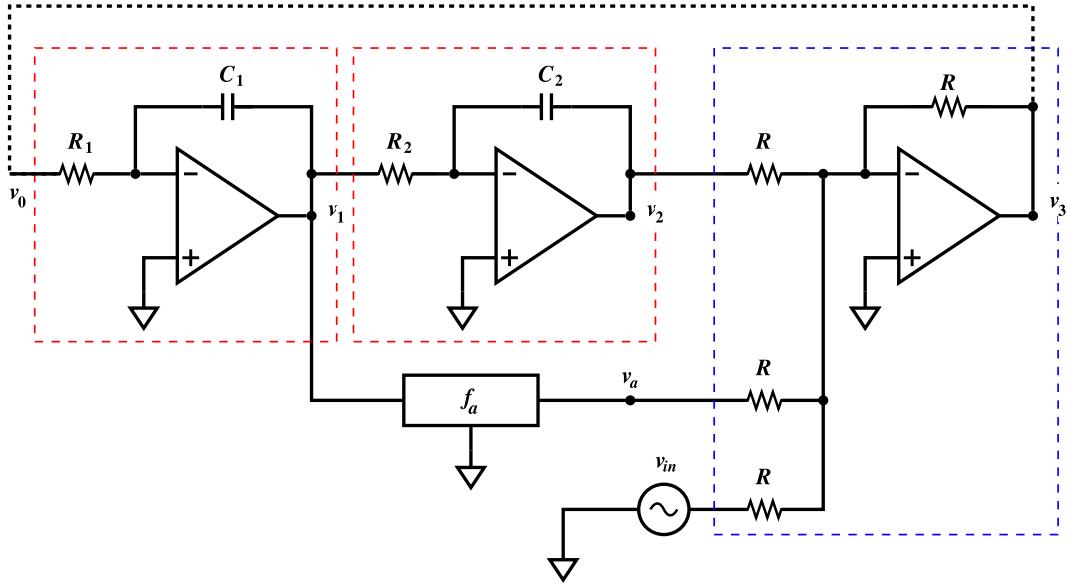


Figura 13.1: Semplice calcolatore analogico per la soluzione dell'equazione differenziale dell'oscillatore armonico. I due riquadri rossi racchiudono due integratori di Miller, il riquadro blu un circuito di somma.

$v_a = a \cdot v_1$, con a costante, si ottiene:

$$\tau_1 \tau_2 \ddot{v}_2 - a \tau_2 \dot{v}_2 + v_2 = -v_{in} \quad (13.3)$$

cioè l'equazione dell'oscillatore armonico smorzato, con uno smorzamento positivo o negativo dato dal coefficiente a .

Si possono impostare le *condizioni iniziali* al tempo $t = 0$ stabilendo la carica sui condensatori C_1 e C_2 e quindi il valore di v_1 e v_2 , ad esempio modificando il circuito come in fig. 13.2.

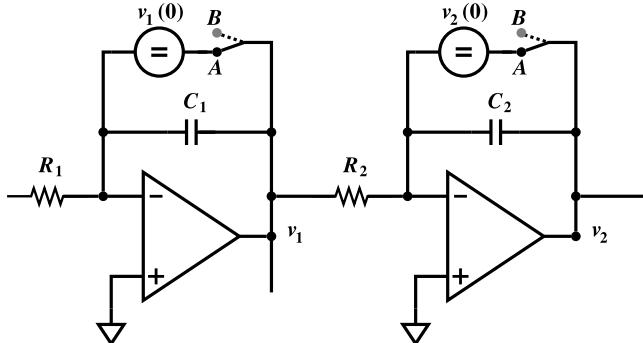


Figura 13.2: Impostazione delle condizioni iniziali. Con gli interruttori in posizione A le tensioni ai capi dei condensatori C_1 e C_2 vengono mantenute ai valori impostati dai due generatori. Portando gli interruttori in B , le tensioni evolvono nel tempo come descritto dalle quazioni del circuito.

Con i due interruttori in posizione A i valori di v_1 e v_2 vengono tenuti fermi a $v_1(0)$ e $v_2(0)$. Dal momento in cui entrambi gli interruttori simultaneamente si aprono, passando nella posizione B , le tensioni nel circuito cominciano a evolvere in funzione del tempo come descritto dalle equazioni 13.1-13.3 e la misura di queste tensioni fornisce le soluzioni cercate.

Si vede facilmente che uno dei limiti principali del calcolatore analogico è dato dalla difficoltà di ottenere risultati con precisioni che vadano oltre la terza o la quarta cifra decimale.

Nell'esempio visto l'uso del calcolatore analogico non risulta particolarmente conveniente, perché l'equazione considerata ha una soluzione analitica che può essere calcolata con facilità. Introducendo però nella rete anche elementi non lineari si possono rappresentare facilmente equazioni che sono intrattabili analiticamente e pesanti da risolvere numericamente in mancanza di strumenti adeguati di calcolo automatico. In questo contesto il calcolatore analogico aveva un suo ruolo.

Anche il filtro KHN, descritto nel par. 12.12.2, può essere considerato una variante del circuito di fig. 13.1 e quindi esempio di *elaborazione analogica*.

13.2 Moltiplicatore a quattro quadranti

Il *moltiplicatore a quattro quadranti* è uno dei dispositivi non lineari più interessanti. La sua funzione, come dice il nome, è di generare una tensione di uscita v_o proporzionale al prodotto di due tensioni di ingresso v_x e v_y :

$$v_o = \frac{v_x v_y}{K} \quad (13.4)$$

dove la costante K ha le dimensioni di una tensione.

Il cuore di un moderno moltiplicatore a quattro quadranti è la *cella di Gilbert*, disegnata in fig. 13.3.¹ La cella di Gilbert è composta da due coppie differenziali

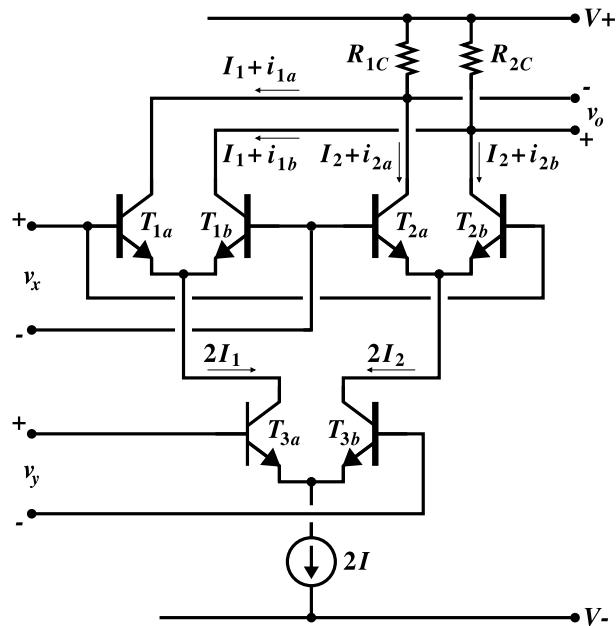


Figura 13.3: Cella di Gilbert.

¹Ovviamente questa tecnologia non era disponibile alla metà del secolo scorso e le soluzioni adottate all'epoca oggi sono decisamente obsolete e di interesse solamente storico.

identiche di transistor *bjt* con gli ingressi in comune invertiti e le uscite sovrapposte e con le correnti di emettitore generate da una terza coppia differenziale.

Per quanto visto al par. 7.9, per ciascuna delle due coppie differenziali si ha, per le correnti ci collettore,

$$i_{1a} - i_{1b} = g_{fe1} \cdot v_x \quad i_{2a} - i_{2b} = -g_{fe2} \cdot v_x \quad (13.5)$$

dove g_{fe1} e g_{fe2} sono le conduttanze di trasferimento diretto dei transistor rispettivamente della coppia T_1 e T_2 .

Poichè le correnti dei collettori si sommano nelle due resistenze R_{1C} ed R_{2C} , che si suppongono di valore eguale e pari ad R_C , si ha

$$v_o = (i_{1a} - i_{1b} + i_{2a} - i_{2b}) \cdot R_C = (g_{fe1} - g_{fe2}) \cdot R_C \cdot v_x \quad (2) \quad (13.6)$$

Se le correnti di polarizzazione di emettitore I_1 ed I_2 dei transistor delle due coppie differenziali sono uguali si ha, secondo l'eq. 7.10,

$$g_{fe1} = g_{fe2} = \frac{I_{1,2}}{\eta V_T}$$

e di conseguenza $v_o = 0$. In presenza di un segnale v_y all'ingresso della terza coppia differenziale le correnti I_1 ed I_2 si sbilanciano e si ha

$$I_1 - I_2 = v_y \cdot g_{fe3}$$

con g_{fe3} conduttanza di trasferimento diretto dei transistor della coppia T_3 .

Combinando i risultati precedenti e utilizzando nuovamente l'eq. 7.10

$$v_o = v_x R_C (g_{fe1} - g_{fe2}) = v_x R_C \frac{I_1 - I_2}{\eta V_T} = v_x v_y \cdot \frac{I R_C}{\eta^2 V_T^2} \quad (13.7)$$

dove I è la corrente di polarizzazione di emettitore dei transistor della coppia T_3 . L'eq. 13.7 è quanto si voleva: una tensione di uscita proporzionale al prodotto $v_x v_y$ delle due tensioni di ingresso. La costante di proporzionalità

$$K = \frac{\eta^2 V_T^2}{I R_C} \quad (13.8)$$

ha le dimensioni di una tensione e presenta una certa dipendenza dalla temperatura a causa del termine V_T^2 (circa 0.7%/°C), ma questo non è un grosso problema. È possibile compensare questa deriva introducendo nella catena di amplificazione successiva alla cella di Gilbert una dipendenza uguale ed opposta dell'amplificazione dalla temperatura.

Anche per il funzionamento della cella di Gilbert, come già si è visto per la coppia differenziale, è fondamentale la possibilità di realizzare con ottima precisione componenti con caratteristiche uguali tramite la tecnica dei circuiti monolitici.

Esistono in commercio diversi circuiti moltiplicatori di questo tipo che contengono oltre la cella di Gilbert anche i circuiti necessari per riportare intorno allo zero sia le tensioni dei due ingressi che quella di uscita. Le caratteristiche principali di alcuni di questi circuiti sono riportate nella tabella 13.1.

²Più precisamente, $R_C // g_{oe}$

modello	tensioni v_x (V)	tensioni v_y (V)	costante K (V)	errore su K	non linearità	frequenza massima (MHz)
AD633	± 10	± 10	10	1%	0.25%	1
AD734	± 10	± 10	10	1%	0.05%	10
AD835	± 1	± 1	1.05	5%	0.5%	250

Tabella 13.1: Caratteristiche principali di alcuni moltiplicatori analogici a quattro quadranti.

13.2.1 Funzione *quadrato*

Inviando ad entrambi gli ingressi di un moltiplicatore a quattro quadranti lo stesso segnale v_i , all'uscita si ha $v_o(t) = v_i^2(t)/K$. Se v_i è un segnale sinusoidale puro $A \sin(\omega t)$ di frequenza angolare ω (con fase nulla con una scelta opportuna dell'origine dei tempi) allora

$$v_o(t) = \frac{A^2}{K} \sin^2(\omega t) = \frac{A^2}{K} \cdot \frac{1 - \cos(2\omega t)}{2} \quad (13.9)$$

Questo risultato contiene due termini, il primo è la tensione costante $v'_o = A^2/(2K)$, il secondo è una oscillazione armonica pura

$$v''_o(t) = \frac{A_0^2}{2K} \cdot \cos(2\omega t)$$

con frequenza angolare 2ω , doppia di quella originale. Utilizzando un filtro passa-alto per eliminare il termine continuo v'_o , come indicato in fig. 13.4 a sinistra, si ottiene un circuito duplicatore di frequenza.

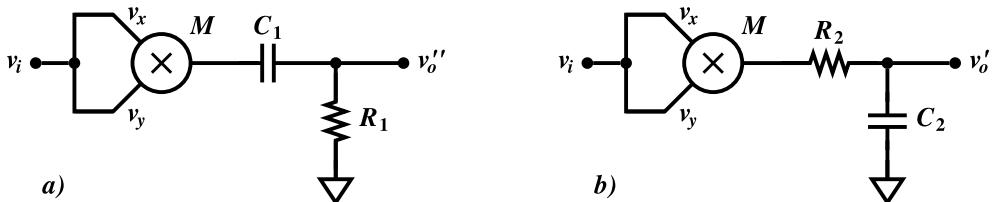


Figura 13.4: Utilizzo di un filtro passa-alto e di un filtro passa-basso per separare le due componenti v'_o e v''_o nel segnale $v_o(t)$ della eq. 13.9. In entrambi i casi si deve avere $RC \gg 1/\omega$.

Utilizzando invece un filtro passa-basso come nel circuito di fig. 13.4 a destra, si ottiene il risultato opposto: a regime la componente continua v'_o giunge inalterata all'uscita del filtro, mentre le componenti v''_o viene soppressa di un fattore $2\omega RC$. Il valore di v'_o coincide, a meno della costante K , con il valore *quadrato medio*³ di $v_i(t)$. Questo risultato vale anche se la tensione $v_i(t)$ non è un segnale sinusoidale puro, ma contiene più componenti armoniche.

³Con il termine *valore quadratico medio* si intende di solito il valore *rms* (*root mean square*): radice quadrata della media del quadrato. Qui il termine *quadrato medio* ha il significato letterale di *media del quadrato*, senza la radice quadrata.

13.3 Oscillatore armonico

Seguendo il procedimento del par. 1.7, l'eq. 13.3 dell'oscillatore armonico analoga alla 1.19 ha, con $v_{in} = 0$, soluzioni

$$v_2(t) = e^{\alpha t} \cdot (Ae^{j\omega t} + Be^{-j\omega t}) \quad (13.10)$$

con

$$\alpha = \frac{a}{2\tau_1} \quad \omega = \sqrt{\frac{a^2}{4\tau_1^2} - \frac{1}{\tau_1\tau_2}} \quad (13.11)$$

e $B = A^*$ per avere soluzioni reali.

Le equazioni 13.10 e 13.11 suggeriscono l'idea che se si pone $a = 0$, ad esempio con un circuito come quello in fig. 13.5, si possa ottenere il comportamento dell'oscillatore armonico ideale, con smorzamento nullo. Date le caratteristiche del circuito si avrebbe la generazione di due segnali sinusoidali puri, v_1 e v_2 , di frequenza $\omega_0 = 1/\sqrt{\tau_1\tau_2}$ ed in quadratura tra di loro.

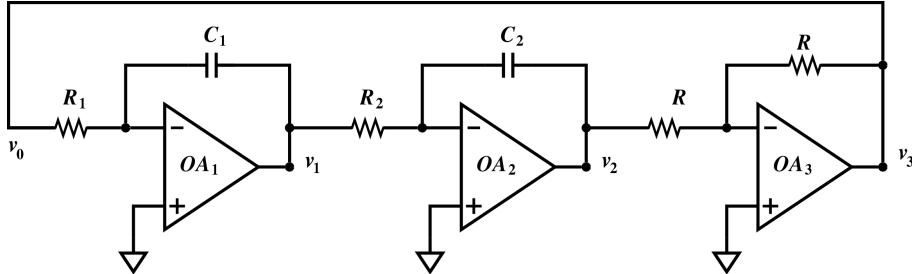


Figura 13.5: Oscillatore armonico con smorzamento teoricamente nullo.

Purtroppo eliminare la rete f_a non è un modo valido per porre $a = 0$ nella eq. 13.3 ed $\alpha = 0$ nella 13.10. Infatti le eq. 13.1 e 13.2 descrivono il funzionamento degli integratori di Miller e del nodo di somma con ottima precisione, ma non sono equazioni esatte. Le eq. 13.1 derivano dalla 12.8, che corrisponde, nel dominio delle frequenze, alla 12.7. In quest'ultima si vede chiaramente come il segnale V_o all'uscita dell'integratore sia in quadratura con il segnale V_s all'ingresso, ma questo è vero solo nel limite ideale ($A \rightarrow \infty, B_w \rightarrow \infty$ ecc.). In realtà l'angolo tra V_o e V_s differisce da $\pi/2$ di una quantità dell'ordine di $1/(\omega R C A)$ e quindi V_o presenta sempre una minuscola componente in fase con V_s . Effetti simili sono prodotti anche dagli altri componenti. Un condensatore non si comporta mai perfettamente come un condensatore, ma presenta sempre anche una piccola componente resistiva; una resistenza, come anche i semplici conduttori dei collegamenti, presenta sempre qualche capacità od induttanza parassita, ecc.

Di conseguenza nella eq. 13.3 compare sempre inevitabilmente un piccolo termine $\epsilon_2 v_2$, frutto delle non idealità del circuito, con ϵ anche molto piccolo, ma mai nullo.

Il comportamento reale del circuito di fig. 13.5 è mostrato in fig. 13.6, utilizzando i valori dei componenti indicati nella didascalia e campionando il segnale all'uscita v_2 con una cadenza di $10 \mu s$ per un tempo di 2 s.

Le condizioni iniziali $v_1 = v_2 = 0$ sono state impostate con un circuito simile a quello di fig. 13.2. Una volta rimosso il blocco le tensioni non sono rimaste ferme al

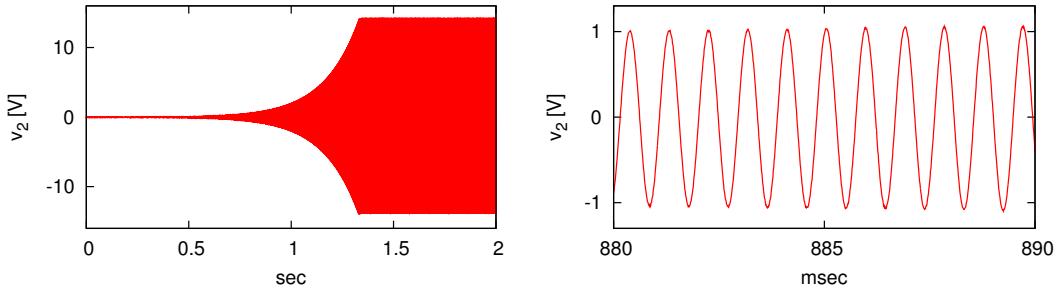


Figura 13.6: Comportamento dell'oscillatore di fig. 13.5 con i valori dei componenti: $C_1 = C_2 = 10.0 \text{ nF}$, $R_1 = R_2 = 14.75 \text{ k}\Omega$, $R = 10.0 \text{ k}\Omega$. Precisione: $\pm 1\%$ per le resistenze, 0.1% per i condensatori. $OA_{1,2,3} = TL081$, con $A = 2 \cdot 10^5$ e $B_w = 3 \text{ MHz}$. Nel grafico a sinistra si vede che il circuito ha in realtà un piccolo smorzamento negativo che fa crescere esponenzialmente l'oscillazione fino alla saturazione degli amplificatori. Il grafico a destra riporta un dettaglio dell'oscillazione, tra 880 e 890 ms

valore iniziale, ma partendo dai livelli delle tensioni di rumore sempre presenti in un circuito reale hanno cominciato ad evolvere nel tempo, come indicato dai grafici. Il grafico a sinistra riporta l'andamento di v_2 per la durata di tutti i 2 s della misura, il grafico a destra riporta un dettaglio della durata di 10 ms, per i valori della ampiezza intorno ad 1 V. Anche se l'effetto su ciascun ciclo è molto piccolo, sui tempi lunghi si vede chiaramente la crescita esponenziale dell'oscillazione, che si arresta solo quando si arriva ai valori di circa $\pm 14 \text{ V}$, corrispondenti alla saturazione degli amplificatori operazionali.

Dai valori dei componenti del circuito si dovrebbe avere $f_0 = \omega_0/(2\pi) = 1.079 \text{ kHz}$ ed $\alpha = 0$. Dal risultato del campionamento risultano i valori $f_0 = 1.072 \text{ kHz}$ ed $\alpha = 5.84 \text{ s}^{-1}$ ($\epsilon = 1.7 \cdot 10^{-3}$), corrispondenti ad una costante di tempo della crescita esponenziale di circa 170 ms. Si vede che il circuito in realtà è abbastanza preciso nel generare la sua funzione. La frequenza differisce solo di 0.7% dal valore atteso e l'errore introdotto dalla presenza del coefficiente ϵ diverso da zero nell'arco di tempo di un ciclo è dell'ordine di 0.5%. Inoltre va detto che nel realizzare questo circuito si sono utilizzati appositamente amplificatori operazionali di uso generale e di tipo economico, per ottenere un effetto più vistoso. Sostituendo nello stesso circuito ai *TL081* tre amplificatori *OPA604*, di qualità e costo superiori,⁴ si ottiene un valore di ϵ , e quindi una variazione di ampiezza tra un ciclo ed il successivo, di oltre un ordine di grandezza migliore.

Il comportamento *qualitativo* del circuito però è drasticamente diverso da quanto ci si aspettava: è quello di un oscillatore armonico con un termine di smorzamento (negativo) piccolo ma non nullo e quindi non può dare oscillazioni di ampiezza stabile.

Si può aggiungere allora esplicitamente la rete f_a e regolare a in modo da avere $\epsilon + a = 0$ ed ottenere finalmente il comportamento ideale dell'oscillatore non smorzato. Anche questa soluzione non può funzionare. Sia ϵ che a sono quantità fisiche soggette ad errore. Non si può ottenere e mantenere stabile nel tempo $\epsilon + a = 0$ esattamente, nel senso matematico del termine.

⁴ $A = 10^5$, $B_w = 20 \text{ MHz}$.

Un sistema fisico *perfettamente* lineare non si può trovare stabilmente in uno stato stazionario diverso da quello nullo $v_1(t) = v_2(t) = v_3(t) = 0$. Le cose cambiano se accettiamo di introdurre nel nostro sistema elementi non lineari.

13.4 Oscillatore *quasi armonico*

La prima e più semplice non linearità è quella presentata dagli amplificatori operazionali quando arrivano alla saturazione. Questo effettivamente stabilizza l'ampiezza dell'oscillazione, come si vede nella fig. 13.6, ed è una soluzione ampiamente adottata soprattutto nei generatori dei segnali in alta frequenza, come si vedrà nell'esempio al par. 17.2. Ma nel campo della bassa frequenza è un metodo brutale che degrada la qualità dei segnali in modo poco controllabile.

Un metodo più raffinato è di usare una rete per f_a che abbia una risposta del tipo $v_a = + (a - b \cdot f(|v_1|)) \cdot v_1$. La funzione $f(|v_1|)$ deve essere una funzione monotona crescente, non necessariamente lineare, che dipenda dall'ampiezza dell'oscillazione. Ad esempio, si potrebbe utilizzare $f(v_1) = \langle v_1^2 \rangle$, ottenendo il *quadrato* di v_1 tramite il *moltiplicatore a quattro quadranti* descritto nel paragrafo 13.2. In queste condizioni il secondo termine dell'eq. 13.3 è

$$-(\epsilon + a - b \langle \dot{v}_2^2 \rangle) \cdot \dot{v}_2 \quad (13.12)$$

Per valori molto piccoli dell'oscillazione il termine $-b \langle \dot{v}_2^2 \rangle$ è trascurabile e, se si è scelto un valore di a sufficientemente grande da avere $a + \epsilon > 0$ indipendentemente dal valore e dal segno di ϵ , il sistema si troverà in condizioni di smorzamento negativo e l'ampiezza dell'oscillazione crescerà nel tempo, tendendo asintoticamente al valore per cui

$$\epsilon + a - b \langle \dot{v}_2^2 \rangle = 0 \quad (13.13)$$

Una volta raggiunta questa condizione ogni aumento o diminuzione dell'ampiezza dell'oscillazione farà comparire di nuovo uno smorzamento, positivo o negativo, che riporterà il sistema verso lo stato stazionario.

Il circuito in fig. 13.7 è un esempio di come può essere realizzata la rete f_a utilizzando i moltiplicatori del par. 13.2.

I risultati ottenuti nel funzionamento dei due circuiti di fig. 13.6 e 13.7 sono riportati in fig. 13.8. In alto si vedono le forme d'onda come appaiono all'oscilloscopio, in basso il contenuto armonico dei segnali generati. I valori delle varie componenti armoniche sono riportati in *dB* rispetto alla componente fondamentale ($n = 1$: 0 dB).

Per il circuito di fig. 13.7 (grafici a sinistra) la rete di controllo f_a limita l'ampiezza dell'oscillazione a ± 10 V di picco, con un livello di distorsione molto basso. Nello spettro compaiono solo le armoniche dispari; la più intensa è la terza, 78.5 dB sotto la fondamentale, quindi con una ampiezza in tensione pari a circa lo 0.012% della fondamentale. Tutte le armoniche pari hanno un livello più basso della sensibilità dell'analizzatore di spettro utilizzato per la misura.

Per il circuito di fig. 13.5 (grafici a destra) l'ampiezza dell'oscillazione è limitata a circa ± 14 V solo dalla saturazione degli amplificatori operazionali, che è visibile nella lieve deformazione della cresta delle semionde positive del segnale. La composizione spettrale è molto più ricca, con la seconda e la quinta armonica al livello di circa 55 ÷ 56 dB sotto la fondamentale (circa 0.2% di ampiezza di tensione) e le altre di

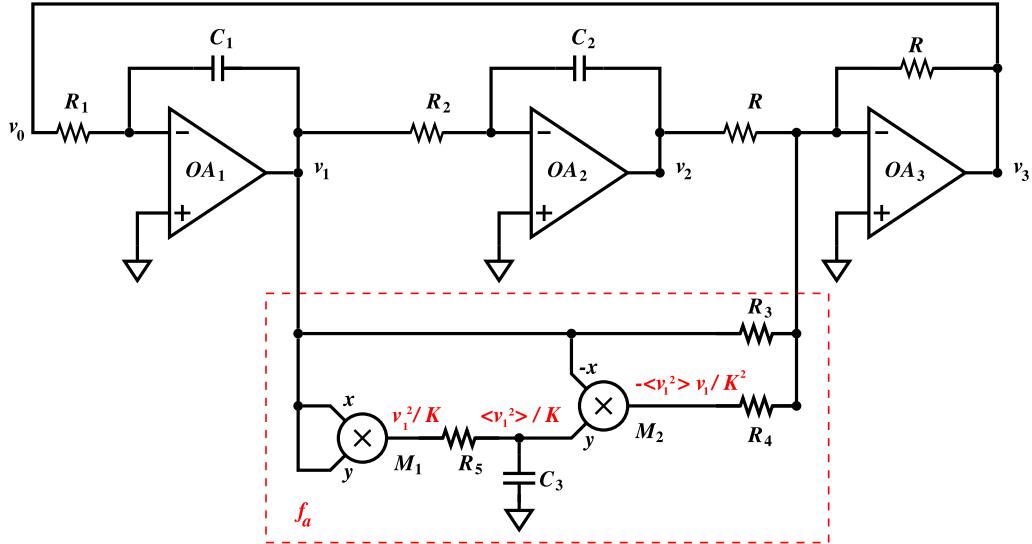


Figura 13.7: Oscillatore armonico composto dal circuito di fig. 13.6 con aggiunta una rete f_a per rendere stazionaria l'oscillazione facendo uso di due moltiplicatori a quattro quadranti. I valori dei componenti sono gli stessi di fig. 13.6, con in più $R_3 = 200 \text{ k}\Omega$, $R_4 = 100 \text{ k}\Omega$, $R_5 = 10 \text{ k}\Omega$, $C_3 = 0.5 \mu\text{F}$ e $M_1 = M_2 = \text{AD633}$.

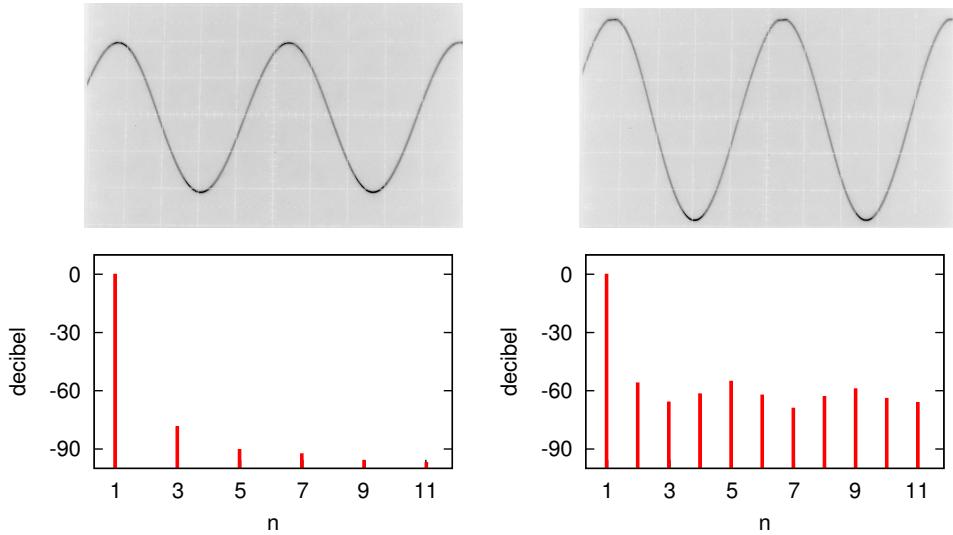


Figura 13.8: Segnale all'oscilloscopio e contenuto spettrale delle prime 11 componenti armoniche per i segnali generati dai circuiti di fig. 13.7 (a sinistra) e 13.6 (a destra).

poco inferiori. A differenza del grafico di sinistra, sono presenti anche tutte le armoeniche pari e l'intensità decresce molto lentamente con l'aumentare della frequenza. Questo comportamento è facilmente comprensibile anche intuitivamente, se si pensa che la distorsione che si osserva nel segnale all'oscilloscopio equivale a sovrapporre ad una sinusoida praticamente ideale una sequenza di impulsi molto brevi con la stessa frequenza della sinusoida: l'ampiezza delle componenti armoniche di una tale

sequenza di impulsi è praticamente costante fino a che la durata dell'impulso rimane piccola rispetto al periodo dell'armonica considerata.

La presenza delle armoniche pari invece è legata alla simmetria della distorsione: se la deformazione fosse simmetrica sui picchi positivi e negativi, anche nello spettro a destra scomparirebbero o sarebbero fortemente ridotte tutte le armoniche pari.

13.5 Potenziale non armonico. Oscillatore di *Duffing*

Utilizzando gli stessi componenti di fig. 13.7, con la modifica circuitale riportata in fig. 13.9, il circuito diventa un oscillatore con un potenziale non-armonico ed è

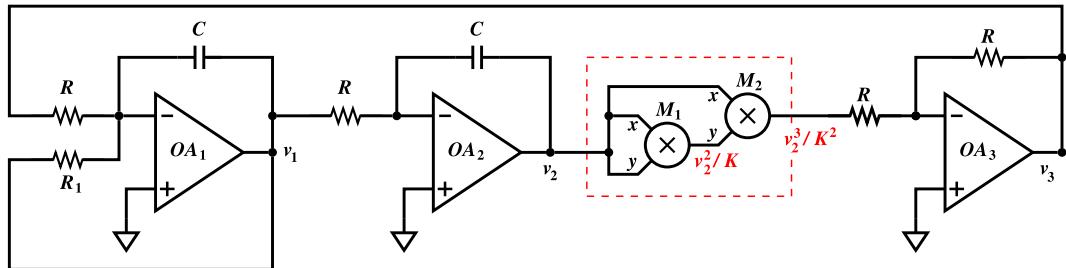


Figura 13.9: Oscillatore non armonico.

descritto dall'equazione

$$\tau^2 \ddot{v}_2 + a\tau \dot{v}_2 + \frac{v_2^3}{K^2} = 0 \quad (13.14)$$

con $a = R/R_1$ e $\tau = RC$.

Questo è uno dei più semplici tra gli oscillatori non armonici ed è noto come *oscillatore di Duffing*. La figura 13.10 riporta il comportamento del circuito partendo dalle condizioni iniziali $v_1 = 0 V$, $v_2 = 10 V$.

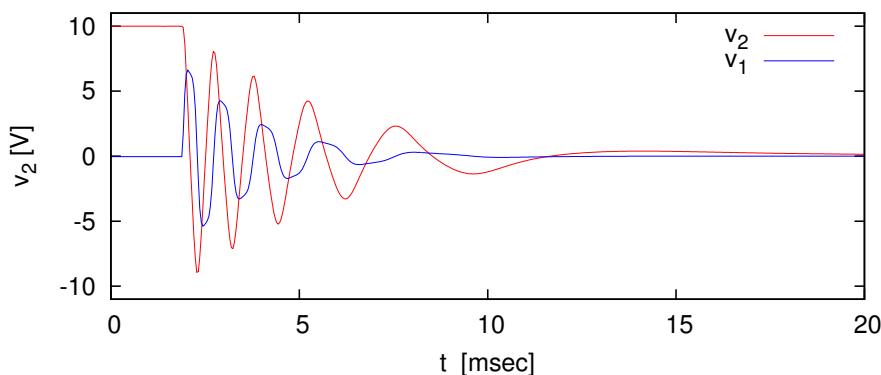


Figura 13.10: Oscillazione smorzata nell'oscillatore di fig. 13.9 con $R = 10 k\Omega$, $R_1 = 120 k\Omega$, $C = 10 nF$ e $K = 10 V$.

Si vede immediatamente dal grafico che la frequenza di oscillazione non è costante, ma diminuisce con il decrescere dell'ampiezza. Solo l'oscillatore armonico è isocrono.

Capitolo 14

Le configurazioni della reazione

Nel capitolo cap. 11 si è visto che la reazione ha effetto diametralmente opposto tra l'amplificatore non invertente (par. 11.7) e quello invertente (par. 11.10) per quanto riguarda la resistenza di ingresso, mentre per le altre caratteristiche (guadagno e resistenza di uscita) i due circuiti si comportano in maniera identica. Questo è dovuto al fatto che la connessione della rete di reazione sul lato dell'ingresso è diversa tra i due circuiti, come si può vedere in fig. 14.1, mentre è la stessa sul lato dell'uscita.

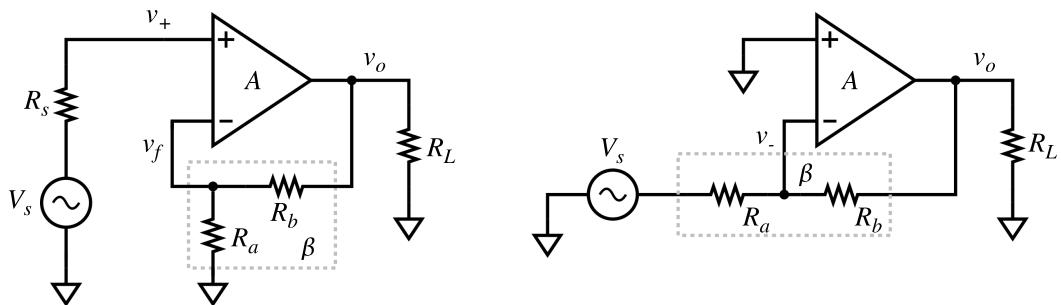


Figura 14.1: Circuiti dell'amplificatore non invertente (a sinistra) ed invertente (a destra).

Nell'amplificatore non invertente l'ingresso differenziale dell'amplificatore operazionale effettua la differenza tra la tensione v_+ proveniente dal generatore di segnale V_s e la tensione v_f riportata all'ingresso dalla rete di reazione β : segnale di ingresso e segnale di reazione si trovano connessi *in serie*.

Nell'amplificatore invertente viene effettuata invece la differenza tra due correnti convergenti *in parallelo* nel nodo v_- , quella proveniente dal generatore di segnale e quella proveniente dalla rete di reazione.

Sul lato dell'uscita, invece, la rete di reazione preleva in entrambi i circuiti la tensione di uscita v_o .

14.1 Le quattro configurazioni della reazione

È possibile costruire anche circuiti in cui la rete di reazione preleva all'uscita la corrente i_o invece che la tensione. Si possono quindi avere in tutto quattro di-

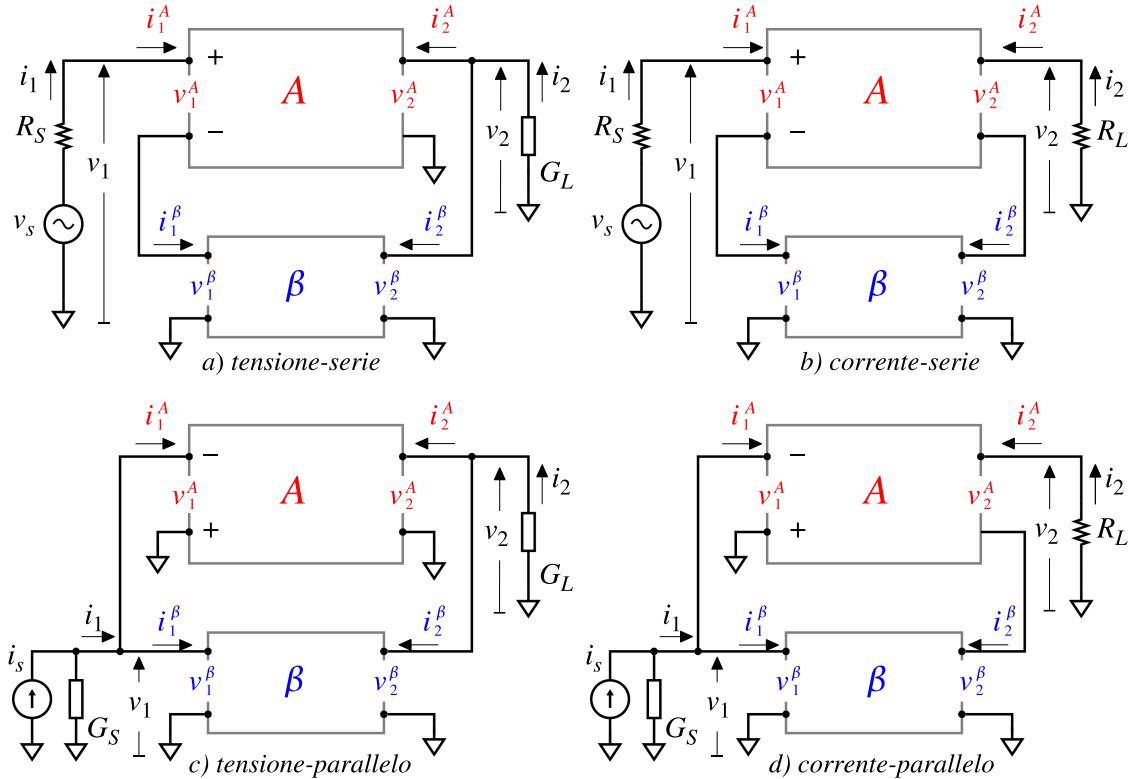


Figura 14.2: Le quattro configurazioni possibili per una anello di reazione. Il primo termine (tensione/corrente) si riferisce alla variabile di uscita che viene prelevata dalla rete β , il secondo termine (serie/parallelo) indica se il segnale generato dalla rete β viene combinato con il segnale di ingresso in serie, sommando le tensioni, oppure in parallelo, sommando le correnti.

verse configurazioni di reazione, indicate tradizionalmente come *tensione-serie*(a), *corrente-serie*(b), *tensione-parallelo*(c) e *corrente-parallelo*(d). Il primo termine (tensione/corrente) si riferisce alla variabile di uscita che viene prelevata dalla rete β , il secondo termine (serie/parallelo) indica se il segnale proveniente dalla rete β viene combinato con il segnale di ingresso in serie, quindi sommando le tensioni, oppure in parallelo, sommando le correnti (fig. 14.2).

Nei casi a e b la tensione v_1^β , riportata in serie alla tensione v_s e v_1^A , è proporzionale rispettivamente alla tensione di uscita v_2 ai capi di G_L oppure alla corrente di uscita i_2 attraverso R_L , secondo la costante di proporzionalità β , che ha le dimensioni rispettivamente di un numero puro (rapporto tra tensioni) o di una resistenza.

Nei casi c e d, invece, è la corrente i_1^β , ancora proporzionale a v_2 oppure i_2 , che viene riportata all'ingresso, in parallelo alla corrente i_1 del generatore. In questi due casi β ha le dimensioni rispettivamente di una conduttanza e di un numero puro (rapporto tra correnti).

Per ragioni di opportunità algebrica che saranno chiare nel seguito, la sorgente del segnale è stata rappresentata come un generatore di tensione (v_s, R_S) nei casi a e b (generatore equivalente di Thévenin) e come un generatore di corrente (i_s, G_s) nei

casi *c* e *d* (generatore equivalente di Norton).

Per ragioni analoghe, gli ulteriori circuiti o apparecchi utilizzatori del segnale presenti all'uscita sono stati rappresentati con una resistenza di carico R_L nei casi *b* e *d* e con una conduttanza $G_L = 1/R_L$ nei casi *a* e *c*.

La descrizione matematica completa delle quattro configurazioni diventa relativamente semplice se ciascuno dei due blocchi A e β è trattato come una rete lineare a due porte, utilizzando per ciascuna configurazione i parametri più appropriati tra quelli descritti nel par. 6.4, come riportato nella tab. 14.1.

<i>a) tensione-serie - parametri h</i> $v_1 = h_{11} i_1 + h_{12} v_2$ $i_2 = h_{21} i_1 + h_{22} v_2$	<i>b) corrente-serie - parametri r</i> $v_1 = r_{11} i_1 + r_{12} i_2$ $v_2 = r_{21} i_1 + r_{22} i_2$
<i>c) tensione-parallelo - parametri g</i> $i_1 = g_{11} v_1 + g_{12} v_2$ $i_2 = g_{21} v_1 + g_{22} v_2$	<i>d) corrente-parallelo - parametri m</i> $i_1 = m_{11} v_1 + m_{12} i_2$ $v_2 = m_{21} v_1 + m_{22} i_2$

Tabella 14.1: I quattro sistemi di equazioni che descrivono una rete lineare a due porte utilizzando i quattro sistemi di parametri h , r , g , m .

In realtà, le quattro descrizioni sarebbero equivalenti e normalmente sarebbe possibile passare da una descrizione all'altra con trasformazioni algebriche, come descritto nel par. 6.4.1, con le limitazioni indicate nel par. 6.6. È solo una questione di opportunità algebrica utilizzare lo schema della tab. 14.1. Infatti, indicando con l'indice A i termini relativi alla rete A e con l'indice β i termini relativi alla rete β , i quattro sistemi di equazioni 14.1 si possono scrivere sinteticamente, per ciascuna delle due reti:

$$\begin{aligned} y_1^A &= p_{11}^A x_1^A + p_{12}^A x_2^A & y_1^\beta &= p_{11}^\beta x_1^\beta + p_{12}^\beta x_2^\beta \\ y_2^A &= p_{21}^A x_1^A + p_{22}^A x_2^A & y_2^\beta &= p_{21}^\beta x_1^\beta + p_{22}^\beta x_2^\beta \end{aligned} \quad (14.1)$$

Le variabili x_1 e x_2 rappresentano le due variabili del sistema, rispettivamente una di ingresso (v_1 oppure i_1) e una di uscita (v_2 oppure i_2); y_1 ed y_2 sono i termini noti; i parametri p sono quelli coerenti con le variabili scelte (tab. 14.2);

	p	x_1	x_2	y_1	y_2	y_s	P_S	P_L
<i>a) tensione serie</i>	h	i_1	v_2	v_1	i_2	v_s	R_S	G_L
<i>b) corrente-serie</i>	r	i_1	i_2	v_1	v_2	v_s	R_S	R_L
<i>c) tensione parallelo</i>	g	v_1	v_2	i_1	i_2	i_s	G_S	G_L
<i>d) corrente parallelo</i>	m	v_1	i_2	i_1	v_2	i_s	G_S	R_L

Tabella 14.2: Corrispondenza tra i simboli delle eq. 14.1 e della fig. 14.2

Le altre variabili che compaiono nella tab. 14.2 sono il segnale del generatore, y_s , e P_S e P_L che rappresentano in maniera coerente R_S oppure G_S ed R_L oppure G_L ¹.

Confrontando la tabella con gli schemi di fig. 14.2 si vede che, grazie alle scelte fatte, la variabile x_1 è sempre in comune tra gli ingressi delle due reti A e β : i_1 se i due ingressi sono in serie, v_1 se sono in parallelo; lo stesso vale per la variabile x_2 per le due uscite.

Si ha quindi $x_1 = x_1^A = x_1^\beta$ e $x_2 = x_2^A = x_2^\beta$ ed è possibile sommare i due sistemi di equazioni 14.1, ottenendo il sistema

$$\begin{aligned} y_1 &= p_{11} x_1 + p_{12} x_2 \\ y_2 &= p_{21} x_1 + p_{22} x_2 \end{aligned} \quad (14.2)$$

dove i parametri p_{ij} rappresentano la somma dei due contributi delle reti A e β :

$$\begin{aligned} p_{11} &= p_{11}^A + p_{11}^\beta & p_{12} &= p_{12}^A + p_{12}^\beta \\ p_{21} &= p_{21}^A + p_{21}^\beta & p_{22} &= p_{22}^A + p_{22}^\beta \end{aligned} \quad (14.3)$$

e le due variabili y_1 ed y_2 sono la somma delle variabili corrispondenti di A e β :

$$\begin{aligned} y_1 &= y_1^A + y_1^\beta \\ y_2 &= y_2^A + y_2^\beta \end{aligned} \quad (14.4)$$

Se si considera la rete complessiva formata dalle due sottoreti A e β , si vede che anche questa è una rete lineare a due porte e che è descritta nel suo insieme dalle eq. 14.2. Quindi **il sistema 14.2 descrive ogni anello di reazione**, qualunque sia la sua configurazione, semplicemente assegnando i significati opportuni alle variabili x , y ed ai parametri p .

Le condizioni di funzionamento ideali corrispondono sempre a $P_S = 0$ e $P_L = 0$: in a e b la sorgente del segnale è un generatore di tensione e $P_S = 0$ significa $R_S = 0$, mentre in c e d la sorgente del segnale è un generatore di corrente e $P_S = 0$ significa $G_S = 0$; all'uscita, in a e c , $P_L = 0$ significa $G_L = 0$ e quindi nessuna corrente viene prelevata dall'uscita del circuito; in b e d invece $P_L = 0$ significa $R_L = 0$ e quindi il carico esterno non introduce nessuna ulteriore caduta di potenziale nel circuito di uscita formato dalle due porte $\frac{A}{2}$ e $\frac{\beta}{2}$ in serie.

14.2 Condizioni all'ingresso ed all'uscita

Le eq. 14.2 descrivono solo il circuito formato dall'amplificatore A e dalla rete di reazione β senza tenere conto della resistenza o conduttanza interna del generatore o del carico. Esaminando nuovamente gli schemi di fig. 14.2, si vede che per le condizioni all'ingresso ed all'uscita si possono scrivere le relazioni in tab. 14.3.

Sostituendo le variabili secondo le corrispondenze della tab. 14.2, le quattro coppie di equazioni vengono a coincidere:

$$\begin{aligned} y_1 &= y_s - x_1 P_S \\ y_2 &= -x_2 P_L \end{aligned} \quad (14.5)$$

¹Un termine usato per indicare queste variabili è *immettenza*, contrazione di *impedenza* e *ammettenza*.

a) tensione-serie - parametri h	b) corrente-serie - parametri r
$v_1 = v_s - i_1 R_S$ $i_2 = -v_2 G_L$	$v_1 = v_s - i_1 R_S$ $v_2 = -i_2 R_L$
c) tensione-parallelo - parametri g	d) corrente-parallelo - parametri m
$i_1 = i_s - v_1 G_S$ $i_2 = -v_2 G_L$	$i_1 = i_s - v_1 G_S$ $v_2 = -i_2 R_L$

Tabella 14.3: Condizioni all'ingresso ed all'uscita delle reti di fig. 14.2.

Amplificazione

Aggiungendo le condizioni determinate dalle eq. 14.5 al sistema 14.2 si ottiene:

$$\begin{aligned} (p_{11} + P_S) x_1 + p_{12} x_2 &= y_s \\ p_{21} x_1 + (p_{22} + P_L) x_2 &= 0 \end{aligned} \quad (14.6)$$

Queste due equazioni sono formalmente identiche alle 6.14. Questo significa che tutte le soluzioni già trovate nel par. 6.5 sono valide per ciascuna delle reti di fig. 14.2. Risolvendo per x_2 e dividendo per y_s si ottiene, analogamente a 6.15,

$$G = \frac{x_2}{y_s} = \frac{-p_{21}}{(p_{11} + P_S)(p_{22} + P_L) - p_{12}p_{21}} \quad (14.7)$$

La quantità G , rapporto tra variabile di uscita e variabile di ingresso determinata dal generatore ideale (v_s o i_s), è detta genericamente *guadagno* o *amplificazione*, ed ha il significato di:

- a) $G = v_2/v_1$ amplificazione di tensione
- b) $G = i_2/v_1$ transconduttanza
- c) $G = v_2/i_1$ transresistenza
- d) $G = i_2/i_1$ amplificazione di corrente

Se si pone a zero il coefficiente di trasferimento inverso p_{12} , come nella 6.16, l'eq. 14.7 diventa

$$G_0 = \frac{x_2}{y_s} = \frac{-p_{21}}{(p_{11} + P_S)(p_{22} + P_L)} \quad (14.8)$$

Dividendo numeratore e denominatore della eq. 14.7 per $(p_{11} + P_S)(p_{22} + P_L)$ ed effettuando le sostituzioni

$$G_0 = \frac{-p_{21}}{(p_{11} + P_S)(p_{22} + P_L)} \quad \beta = p_{12} \quad (14.9)$$

si ottiene il risultato

$$G = \frac{G_0}{1 + G_0\beta} \quad (14.10)$$

che coincide con l'eq. 6.18 e la 11.3.

A questo punto è abbastanza naturale identificare p_{12} , il parametro di trasferimento inverso da uscita a ingresso, con il coefficiente di reazione β e l'azzeramento di p_{12} lasciando tutte le altre condizioni inalterate come equivalente ad aprire l'anello di reazione come visto nel par. 11.3; di conseguenza G_0 è l'amplificazione ad anello aperto e $G_0\beta$ è il guadagno d'anello.

Se il guadagno d'anello è sufficientemente elevato ($G_0\beta \gg 1$) si ha come nel par. 11.7:

$$G = \frac{G_0}{1 + G_0\beta} \simeq \frac{G_0}{G_0\beta} = \frac{1}{\beta} \quad (14.11)$$

Il guadagno G_0 , e di conseguenza il guadagno d'anello $G_0\beta$, dipende chiaramente dalla presenza di P_S e P_L ed assume il valore massimo $G_0 = -p_{21}/(p_{11}p_{22})$ con $P_S = 0$ e $P_L = 0$.

Con $P_L \neq 0$ si introduce una attenuazione nel circuito di uscita a causa del partitore formato da p_{22} e P_L . Stessa situazione si verifica nel circuito di ingresso con $P_S \neq 0$ a causa del partitore formato da P_S e p_{11} .

In condizioni *normali* queste attenuazioni non dovrebbero alterare la condizione $G_0\beta \gg 1$.

Resistenza e conduttanza di ingresso e di uscita

Procedendo ancora come già visto al par. 6.5, è possibile calcolare i parametri (resistenza o conduttanza) di ingresso e di uscita (fig. 14.3).

Vista l'identità formale tra i sistemi 14.6 e 6.14 si possono scrivere direttamente i risultati

$$\begin{aligned} P_i = \frac{y_1}{x_1} &= p_{11} - \frac{p_{21}p_{12}}{p_{22} + P_L} \\ &= p_{11} \cdot \left[1 - \frac{p_{21}p_{12}}{p_{11}(p_{22} + P_L)} \right] \\ &= p_{11} \cdot (1 + G_0\beta) \end{aligned} \quad (14.12)$$

e

$$\begin{aligned} P_o = \frac{y_2}{x_2} &= p_{22} - \frac{p_{21}p_{12}}{p_{11} + P_S} \\ &= p_{22} \cdot \left[1 - \frac{p_{21}p_{12}}{(p_{11} + P_S)p_{22}} \right] \\ &= p_{22} \cdot (1 + G_0\beta) \end{aligned} \quad (14.13)$$

dove G_0 è ancora l'amplificazione ad anello aperto data dalla eq. 14.8, con $P_S = 0$ per la 14.12 e $P_L = 0$ per la 14.13.

L'effetto della reazione negativa è sempre e comunque quello di moltiplicare la caratteristica di ingresso p_{11} e la caratteristica di uscita p_{22} della rete per il coefficiente $(1 + G_0\beta)$, sia che si tratti di una resistenza, che di una conduttanza.

Poichè $1 + G_0\beta$ può essere anche molto grande, dell'ordine di $10^5 \div 10^6$, le variazioni di resistenza e conduttanza di ingresso e di uscita possono essere anche di svariati ordini di grandezza in più o in meno.

All'ingresso, se p_{11} è una resistenza, il suo valore aumenta e al limite tende a comportarsi come un circuito aperto; se p_{11} è una conduttanza, il suo valore, come resistenza equivalente, diminuisce e al limite tende a comportarsi come un corto circuito.

All'uscita, a seconda che p_{22} sia una resistenza o una conduttanza, se $G_0\beta$ è molto elevato il circuito tenderà a comportarsi come un generatore di corrente nei casi *c* e *d* ($P_o = R_o \rightarrow \infty$) e di tensione nei casi *a* e *b* ($P_o = G_o \rightarrow \infty$).

In sintesi:

All'ingresso, la reazione negativa:

serie: moltiplica per $(1 + G_0\beta)$ la resistenza di ingresso;

parallelo: divide per $(1 + G_0\beta)$ la resistenza di ingresso.

All'uscita, la reazione negativa:

di tensione: divide per $(1 + G_0\beta)$ la resistenza di uscita;

di corrente: moltiplica per $(1 + G_0\beta)$ la resistenza di uscita.

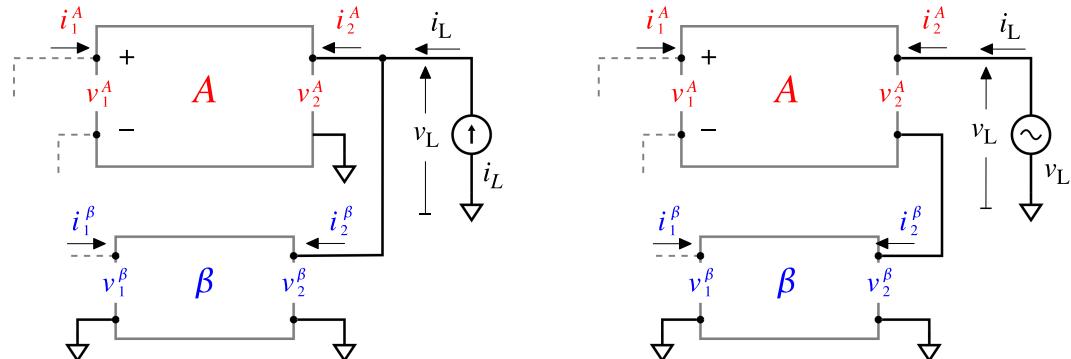


Figura 14.3: A sinistra, calcolo della conduttanza di uscita con reazione di tensione: la conduttanza G_L è idealmente sostituita dal generatore di corrente i_L e viene calcolata la tensione v_L . A destra, calcolo della resistenza di uscita con reazione di corrente: la resistenza R_L è sostituita dal generatore di tensione v_L e viene calcolata la corrente i_L .

14.3 Quattro modelli per l'amplificatore operazionale

Per poter utilizzare le equazioni del paragrafo precedente è necessario disporre delle quattro descrizioni dell'amplificatore operazionale di fig. 11.4 tramite i quattro gruppi di parametri h, g, r, m .

Le definizioni dei parametri come risultano dal cap. 6 sono:

$$\begin{aligned}
 h_{11} &= \frac{v_1}{i_1} \Big|_{v_2=0} & h_{12} &= \frac{v_1}{v_2} \Big|_{i_1=0} & r_{11} &= \frac{v_1}{i_1} \Big|_{i_2=0} & r_{12} &= \frac{v_1}{i_2} \Big|_{i_1=0} \\
 h_{21} &= \frac{i_2}{i_1} \Big|_{v_2=0} & h_{22} &= \frac{i_2}{v_2} \Big|_{i_1=0} & r_{21} &= \frac{v_2}{i_1} \Big|_{i_2=0} & r_{22} &= \frac{v_2}{i_2} \Big|_{i_1=0} \\
 g_{11} &= \frac{i_1}{v_1} \Big|_{v_2=0} & g_{12} &= \frac{i_1}{v_2} \Big|_{v_1=0} & m_{11} &= \frac{i_1}{v_1} \Big|_{i_2=0} & m_{12} &= \frac{i_1}{i_2} \Big|_{v_1=0} \\
 g_{21} &= \frac{i_2}{v_1} \Big|_{v_2=0} & g_{22} &= \frac{i_2}{v_2} \Big|_{v_1=0} & m_{21} &= \frac{v_2}{v_1} \Big|_{i_2=0} & m_{22} &= \frac{v_2}{i_2} \Big|_{v_1=0}
 \end{aligned} \tag{14.14}$$

e corrispondono ai circuiti modello di fig. 6.3, ripetuti per comodità in fig. 14.4.

Nelle definizioni la condizione $v_1, v_2 = 0$ significa che si deve imporre differenza di potenziale nulla tra i terminali della porta indicata, che equivale a dire che i due terminali devono essere cortocircuitati; la condizione $i_1, i_2 = 0$ significa invece che non deve circolare alcuna corrente, quindi la porta in questione deve essere lasciata a circuito aperto.

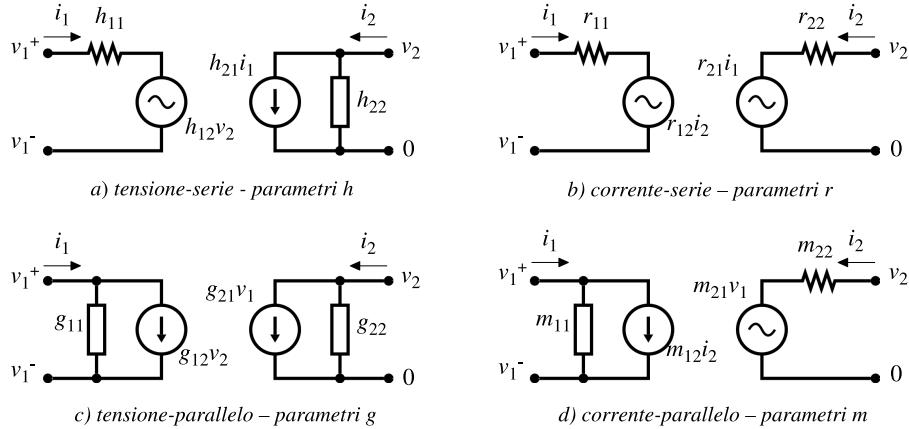


Figura 14.4: I quattro modelli di descrizione dell'amplificatore operazionale come rete lineare a due porte nelle quattro configurazioni della reazione.

Grazie all'ingresso differenziale dell'amplificatore operazionale, entrambi i nodi della porta 1 sono svincolati da massa; all'uscita invece uno dei nodi è necessariamente collegato a massa e di questo si dovrà tenere conto nella realizzazione dei circuiti di fig. 14.2 b e d. Applicando le definizioni al circuito di fig. 11.4, si ottengono le quattro matrici di parametri riportate nella tabella 14.4.

Su queste matrici ci sono alcune osservazioni da fare:

- Confrontando il modello dell'amplificatore operazionale di fig. 11.4 con i quattro sistemi di parametri, ci si rende conto che i parametri m sono i più adatti per

$$\mathbf{h}^{\mathbf{A}} = \begin{pmatrix} r_i & 0 \\ -A\frac{r_i}{r_o} & \frac{1}{r_o} \end{pmatrix} \quad \mathbf{r}^{\mathbf{A}} = \begin{pmatrix} r_i & 0 \\ Ar_i & r_o \end{pmatrix}$$

$$\mathbf{g}^{\mathbf{A}} = \begin{pmatrix} \frac{1}{r_i} & 0 \\ -\frac{A}{r_o} & \frac{1}{r_o} \end{pmatrix} \quad \mathbf{m}^{\mathbf{A}} = \begin{pmatrix} \frac{1}{r_i} & 0 \\ A & r_o \end{pmatrix}$$

Tabella 14.4: Parametri dei quattro modelli di fig. 14.4 per l'amplificatore operazionale di fig. 11.4.

descrivere un amplificatore di tensione, anche se tradizionalmente il parametro di ingresso m_{11} , che dovrebbe essere una conduttanza, viene indicato tramite il suo inverso $r_i = 1/m_{11}$. Nel caso dell'amplificatore ideale si ha $m_{21} = A$, mentre tutti gli altri termini sono zero.

- I parametri h , g ed r non possono essere utilizzati nel caso di un amplificatore ideale di tensione perché almeno uno tra i termini p_{11} , p_{21} , p_{22} ha valore infinito. Riguardo gli altri tipi di amplificatori (di corrente, a transconduttanza, a transresistenza), ognuno viene meglio descritto da uno in particolare tra i sistemi di parametri.
- Nel caso dell'amplificatore operazionale reale, si possono utilizzare in maniera equivalente tutti e quattro i sistemi di parametri, anche se le schematizzazioni possono apparire a prima vista innaturali (ad es., con i parametri h si è costretti a descrivere come generatore di corrente l'uscita di un amplificatore operazionale, anche se la resistenza di uscita è solo di pochi ohm).
- In tutti e quattro i casi il parametro p_{12} è zero. Questo è un punto molto importante, come si vedrà nei paragrafi seguenti, ed esprime il fatto che per un aspetto particolare l'amplificatore operazionale è un componente realmente ideale: è un dispositivo a due porte *unilaterale*: il trasferimento di segnale avviene solo dall'ingresso all'uscita (parametro p_{21}), mentre nella direzione opposta dall'uscita all'ingresso (parametro p_{12}) è sufficientemente piccolo da poter essere considerato nullo a tutti gli effetti.

14.4 Quattro esempi di rete di reazione

La quattro reti di reazione β della fig. 14.2 possono essere ovviamente realizzate in infiniti modi diversi. Nella fig. 14.5 sono riportati quattro esempi molto semplici di reti resistive, una per ciascuna delle quattro configurazioni. Lo schema di fig. 14.5a è quello dell'amplificatore non invertente di fig. 11.8; quello di fig. 14.5c è quello dell'amplificatore invertente di fig. 11.17, con la differenza che il ruolo della resistenza

R_a è svolto dalla resistenza R_S del generatore v_s . Questi sono i due circuiti di gran lunga più comuni.

I circuiti in fig. 14.5b e 14.5d sono di uso meno comune; si comportano all'uscita come un generatore di corrente, il circuito in b come amplificatore a transconduttanza, quello in d come amplificatore di corrente. Hanno l'inconveniente di richiedere entrambi i terminali della resistenza di carico svincolati da massa, per cui nell'uso pratico si preferiscono altri circuiti che non hanno questa limitazione. Infatti, con-

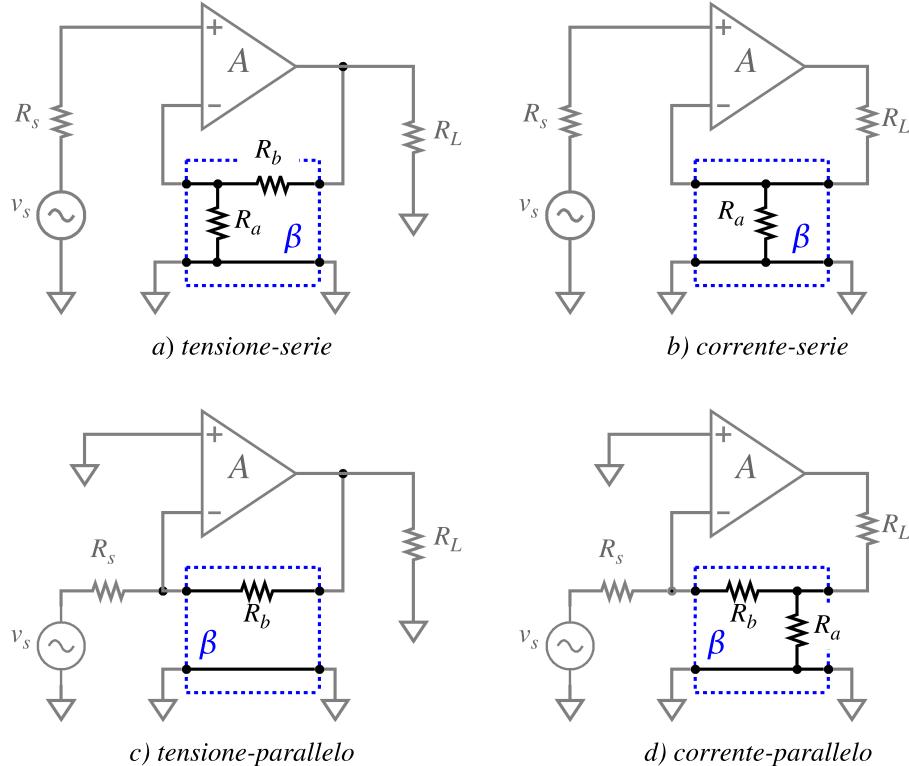


Figura 14.5: Esempi di reti resistive di reazione per le quattro configurazioni di fig. 14.2.

frontando gli schemi di fig. 14.5 con quelli di principio di fig. 14.2 si vede che nei due circuiti in b e d la connessione delle due reti e della resistenza di carico R_L è stata leggermente modificata poiché è richiesto che entrambe le porte 2^A e 2^β abbiano il nodo inferiore vincolato a massa.

Questa modifica porta al cambiamento di segno di i_2 e v_2 nel calcolo dei parametri \mathbf{r} ed \mathbf{m} , come indicato nella fig. 14.6, e di conseguenza al cambiamento di segno dei parametri p_{12} e p_{21} nelle due matrici nella tabella 14.5.

Ognuna delle quattro reti di reazione della fig. 14.5 può essere descritta mediante ognuno dei quattro sistemi di parametri, applicando nuovamente le definizioni date dalle equazioni 14.14. Ma quello che serve realmente è di utilizzare per ciascuna configurazione soltanto il sistema di parametri più idoneo, coerente con il modello di amplificatore operazionale riportato nel paragrafo precedente, e cioè le quattro matrici della tabella 14.5.

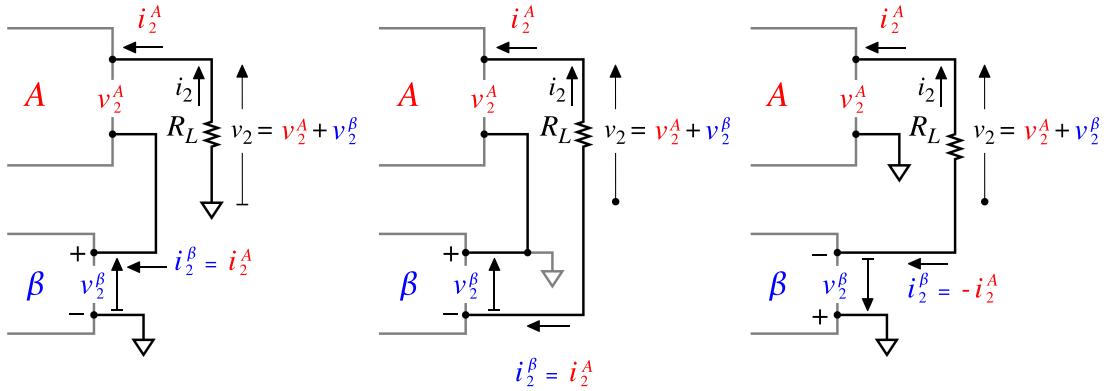


Figura 14.6: A sinistra: connessione all'uscita tra le reti A e β come riportato negli schemi di principio di fig. 14.2. Il collegamento a massa indicato è puramente convenzionale, perché ancora non sono state definite le caratteristiche delle reti; può essere rimosso e sostituito dal collegamento a massa del nodo comune tra v_2^A e v_2^β , come indicato nello schema al centro. Questo schema è poi ridisegnato a destra come appare nella fig. 14.5.

$$\begin{aligned}
 \mathbf{h}^\beta &= \begin{pmatrix} \beta R_b & \beta \\ -\beta & \beta/R_a \end{pmatrix} & \mathbf{r}^\beta &= \begin{pmatrix} -\beta & \beta \\ \beta & -\beta \end{pmatrix} \\
 \beta &= R_a/(R_a + R_b) & \beta &= -R_a \\
 \text{a) tensione-serie} & & \text{b) corrente-serie} & \\
 \mathbf{g}^\beta &= \begin{pmatrix} -\beta & \beta \\ \beta & -\beta \end{pmatrix} & \mathbf{m}^\beta &= \begin{pmatrix} \beta/R_a & \beta \\ -\beta & \beta R_b \end{pmatrix} \\
 \beta &= G_b = -1/R_b & \beta &= R_a/(R_a + R_b) \\
 \text{c) tensione-parallelo} & & \text{d) corrente-parallelo} &
 \end{aligned} \tag{14.15}$$

Tabella 14.5: Matrici di descrizione delle quattro reti di reazione dei circuiti in fig. 14.5. Per ogni rete è stato utilizzato il sistema di parametri più idoneo, in coerenza con la fig. 14.4 e la tabella 14.4. Il parametro p_{12} è stato sempre indicato con β e la sua espressione riportata per ogni matrice.

In queste matrici il parametro p_{12} , che è l'elemento più importante della rete di reazione, è stato sempre indicato con β , e gli altri parametri sono stati espressi di conseguenza. Per ogni sistema di parametri è riportata l'espressione di β relativa.

Nella tabella si è già tenuto conto dei cambiamenti di segno di i_2 e v_2 descritti nella fig. 14.6. Altrimenti le matrici \mathbf{r}^β e \mathbf{m}^β sarebbero state:

$$\mathbf{r}^\beta = \begin{pmatrix} \beta & \beta \\ \beta & \beta \end{pmatrix} \quad \mathbf{m}^\beta = \begin{pmatrix} -\beta/R_a & \beta \\ -\beta & -\beta R_b \end{pmatrix}$$

$$\beta = R_a \quad \beta = -R_a/(R_a + R_b)$$

14.5 Amplificazione

Sommando le matrici della tab. 14.4 relative all'amplificatore operazionale con le corrispondenti della tab. 14.5 delle reti di reazione, si ottengono i quattro sistemi di coefficienti delle eq. 14.2, riportati nella tab. 14.6, che descrivono il funzionamento dei quattro diversi circuiti di fig. 14.5.

$$\mathbf{h} = \begin{pmatrix} r_i + \beta R_b & \beta \\ -A \frac{r_i}{r_o} - \beta & \frac{1}{r_o} + \frac{\beta}{R_a} \end{pmatrix} \quad \mathbf{r} = \begin{pmatrix} r_i - \beta & \beta \\ Ar_i + \beta & r_o - \beta \end{pmatrix}$$

$$\beta = R_a/(R_a + R_b) \quad \beta = -R_a$$

$$\mathbf{g} = \begin{pmatrix} \frac{1}{r_i} - \beta & \beta \\ \frac{A}{r_o} + \beta & \frac{1}{r_o} - \beta \end{pmatrix} \quad \mathbf{m} = \begin{pmatrix} \frac{1}{r_i} + \frac{\beta}{R_a} & \beta \\ -A - \beta & r_o + \beta R_b \end{pmatrix}$$

$$\beta = G_b = -1/R_b \quad \beta = R_a/(R_a + R_b)$$

Tabella 14.6: Matrici dei parametri dei quattro circuiti di fig. 14.5, ottenute sommando termine a termine i parametri della rete A ed i parametri della rete β .

La cosa veramente notevole in queste quattro matrici è il fatto che il termine p_{12} risulta sempre dovuto solo al contributo della rete di reazione β , grazie alla *unilateralità* dell'amplificatore A , come già notato nel par. 14.3. Quando si va a calcolare il guadagno ad anello chiuso tramite l'eq. 14.10 del par. 14.2 si ottiene sempre $G \simeq 1/\beta$, nel limite di guadagno d'anello $G_0\beta$ sufficientemente alto.

14.6 Guadagno d'anello

Quello che si rimane da verificare è che il guadagno d'anello sia sufficientemente elevato e che rimanga tale anche tenendo conto delle caratteristiche della sorgente del segnale e del carico, cioè di P_S e P_L .

Dalle eq. 14.9 si ha per il guadagno d'anello, con $P_S = 0$ e $P_L = 0$:

$$G_0\beta = -\frac{p_{21}p_{12}}{p_{11}p_{22}} \quad (14.16)$$

ed utilizzando i coefficienti della tab. 14.6, si ottiene per $G_0\beta$ per ognuna delle quattro configurazioni:

$$\frac{A\beta + \beta^2 \frac{r_o}{r_i}}{1 + \beta \frac{R_b}{r_i} + \beta \frac{r_o}{R_a} + \beta^2 \frac{R_b}{r_i} \frac{r_o}{R_a}} \quad \frac{A - \frac{R_a}{r_i}}{1 + \frac{r_o}{r_i} + \frac{r_o}{R_a} + \frac{R_a}{r_i}}$$

a) tensione-serie

b) corrente-serie

$$G_0\beta =$$

$$\frac{A - \frac{r_o}{R_b}}{1 + \frac{r_o}{r_i} + \frac{R_b}{r_i} + \frac{r_o}{R_b}} \quad \frac{A + \beta}{1 - \beta + \frac{r_o}{\beta r_i} + \frac{r_o}{R_a} + \frac{R_b}{r_i}}$$

c) tensione-parallelo

d) corrente-parallelo

Il coefficiente β è stato mantenuto solo nei casi *a* e *d*, dove esprime il rapporto di partizione (di tensione o di corrente) $R_a/(R_a + R_b)$; nei casi *b* e *c*, dove invece ha le dimensioni di una resistenza e di una conduttanza, è stato sostituito dal valore corrispondente secondo le espressioni 14.15.

Esaminando le quattro espressioni, si verifica facilmente che le condizioni richieste per avere $G_0\beta \gg 1$ sono sempre le stesse: tutti e quattro i denominatori differiscono poco dall'unità se $r_o < \{R_a, R_b\} < r_i$ e β non è eccessivamente piccolo.

Con la condizione $G_0\beta \gg 1$, l'espressione del guadagno 14.10 diventa, nei quattro diversi casi:

$$G = \frac{v_2}{v_1} = \frac{R_a + R_b}{R_a} \quad G = \frac{i_2}{v_1} = -\frac{1}{R_a}$$

a) tensione-serie

b) corrente-serie

$$G = \frac{v_2}{i_1} = -R_b$$

$$G = \frac{i_2}{i_1} = \frac{R_a + R_b}{R_a}$$

c) tensione-parallelo

d) corrente-parallelo

(14.17)

14.7 Resistenza e conduttanza di ingresso e di uscita

Nei casi *a* e *b* (tensione-serie e corrente-serie), p_{11} è praticamente la resistenza di ingresso dell'amplificatore operazionale; moltiplicando questa quantità, già di suo molto elevata, per $(1 + G_0\beta)$ si ottengono valori di resistenza di ingresso R_i per il

circuito completo estremamente alti, come già visto nel par. 11.9 per il caso specifico della configurazione tensione-serie.

Nei casi *c* e *d* (tensione-parallelo e corrente-parallelo) p_{11} è la conduttanza di ingresso del circuito, determinata praticamente dalle resistenze della rete di reazione β , in quanto il contributo $1/r_i$ dell'amplificatore operazionale è del tutto trascurabile. In questi casi, quindi, è la conduttanza p_{11} ad essere moltiplicata per $(1+G_0\beta)$, dando come risultato una conduttanza di ingresso G_i per il circuito completo elevatissima, cioè una resistenza di ingresso $R_i = 1/G_i$ praticamente nulla. Considerando come ordini di grandezza tipici $1/p_{11} \simeq 10^2 \dots 10^4 \Omega$ e $1 + G_0\beta \simeq 10^4 \dots 10^6$, si trova $R_i \simeq 10^{-4} \dots 1 \Omega$. A causa di questo comportamento, il nodo di ingresso dei due circuiti in fig. 14.2 e 14.5 *c* e *d* prende il nome di *massa virtuale*: a tutti gli effetti è come se questo nodo fosse collegato a massa. Quindi, ad esempio, la corrente di ingresso è semplicemente $i_1 = v_s/R_S$.

Tutte queste considerazioni risultano valide fino a che vale la condizione

$$\frac{p_{22}}{p_{22} + P_L} \simeq 1$$

cioè $P_L < p_{22}$. Esaminando le quattro matrici di parametri della tab. 14.6 si vede che nei casi *a* e *c* (tensione-serie e corrente-serie) questo equivale a dire $R_L > r_o$. L'interpretazione di questa condizione è evidente: se si cortocircuita l'uscita dell'operazionale a massa, collegando una R_L troppo piccola (P_L grande), il circuito smette di funzionare.

Se un amplificatore operazionale è in grado di fornire una tensione di uscita nell'intervallo di valori $\pm V_M$ in condizioni ideali, cioè in assenza di carico, questo intervallo si riduce a $V_M \cdot R_L / (r_o + R_L)$ in presenza di un carico R_L . Per non penalizzare eccessivamente la dinamica di uscita, conviene sempre rispettare la condizione $R_L \geq 10 \cdot r_o$.

Nei casi *b* e *d* (corrente-serie e corrente parallelo) la condizione da rispettare è invece quella opposta, cioè che R_L non sia troppo grande. Se, al limite, si pone $R_L = \infty$, si è già visto che si apre l'anello di reazione ed il circuito smette di funzionare. In questi due casi, però, non sempre è conveniente rispettare la condizione $P_L < p_{22}$, perché si avrebbe di nuovo una forte penalizzazione nella dinamica di uscita. In questi circuiti, se il guadagno d'anello $G_0\beta$ è sufficientemente alto, è più comune avere $P_L/p_{22} = 0 \dots 10$ o anche $P_L/p_{22} = 0 \dots 100$, cioè avere resistenze di carico R_L fino a dieci o cento volte maggiori della resistenza R_a della rete di reazione.

Nei casi *a* e *c* (tensione-serie e tensione-parallelo) il parametro di uscita P_o è una conduttanza ed è praticamente la conduttanza di uscita $1/r_o$ dell'amplificatore operazionale, moltiplicata per $(1+G_0\beta)$. Si arriva facilmente a valori di resistenza di uscita di centesimi o di millesimi di ohm. Il circuito si comporta come un generatore di tensione ideale.

Nei casi *b* e *d* (corrente-serie e corrente-parallelo) p_{22} è una resistenza, data prevalentemente dalle resistenze R_a ed R_b della rete di reazione; il valore di questa resistenza viene moltiplicato per $(1+G_0\beta)$, dando per la resistenza di uscita R_o valori che possono arrivare alle decine, centinaia o anche migliaia di megaohm. Il circuito si comporta verso il carico R_L come un generatore di corrente praticamente ideale.

Anche per i valori del parametro P_S valgono considerazioni analoghe a quelle fatte per P_L : Si deve rispettare la condizione $P_S < p_{11}$. Nei casi *a* e *b* (tensione-serie e

corrente-serie) P_S è una resistenza e deve essere minore (o almeno non troppo più grande) di r_i ; al limite, se R_S va a infinito, si apre l'anello di reazione ed il circuito smette di funzionare.

Nei casi *c* e *d* (tensione-parallelo e corrente-parallelo), P_S è una conduttanza e deve essere più piccola (o almeno non troppo più grande) di quella dovuta alle resistenze R_a ed R_b della rete di reazione β . Una conduttanza P_S molto grande, al limite infinita, è una resistenza $R_S = 0$ che cortocircuita l'ingresso dell'amplificatore A impedendogli di ricevere il segnale di reazione dalla rete β .

Capitolo 15

Stabilità delle reti lineari

I circuiti descritti nei cap. 11-14 contenenti anelli di reazione possono essere *stabili* oppure *instabili*. Il comportamento si dice *stabile* quando applicando all'ingresso un segnale che decade nel tempo fino ad annullarsi si ottiene dal circuito (o *rete*) una risposta che analogamente decade e si annulla (fig. 15.1). Il comportamento si dice

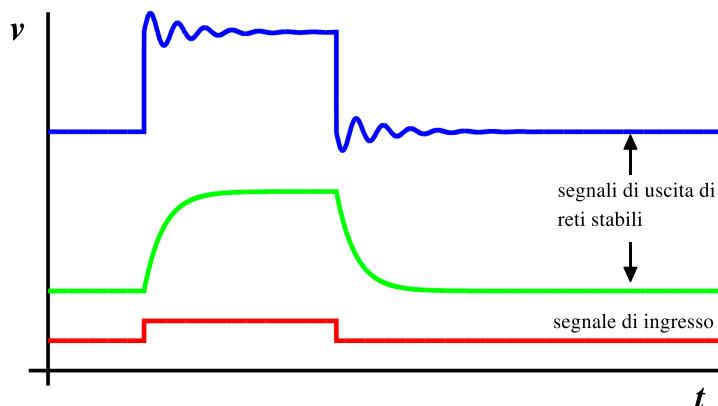


Figura 15.1: Risposta nel tempo di una rete stabile. L'impulso di ingresso (in basso, in rosso) appare all'uscita amplificato ed eventualmente modificato nella sua composizione spettrale: il segnale in verde corrisponde ad un amplificatore che deprime le componenti ad alta frequenza dello spettro, mentre il segnale in blu si ottiene con una esaltazione delle componenti ad alta frequenza.

invece *instabile* quando un segnale di ingresso che decade e scompare produce una risposta che dura indefinitamente nel tempo rimanendo stazionaria oppure crescendo in ampiezza, fino ad essere limitata dalle non linearità inevitabilmente presenti in qualunque rete reale (fig. 15.2). Il circuito a resistenza negativa descritto nel par. 12.11 ed il generatore del par. 13.3 e 13.4 sono esempi di reti instabili.

Reti stabili e instabili sono altrettanto importanti, ciascuna nel suo campo di applicazione: amplificatori e filtri sono esempi di reti stabili, i generatori di segnali sono applicazioni di reti instabili. In quest'ultimo caso va precisato che da un buon generatore ci si aspetta un segnale continuo e stabile all'uscita, senza alcun segnale di

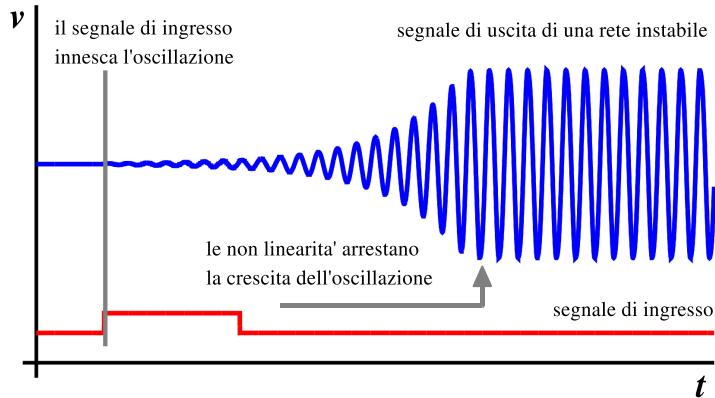


Figura 15.2: Risposta nel tempo di una rete instabile. L’impulso di ingresso (in rosso) innesca nella rete una oscillazione che continua a crescere in ampiezza anche dopo che l’impulso è terminato, fino ad essere limitata dalle non linearità dei componenti.

ingresso. Il comportamento della rete dovrà essere quindi instabile dal punto di vista della definizione data sopra, ma con caratteristiche *stazionarie* in modo da garantire la generazione di *segnali stabili*.¹

Nel capitolo 11 si è visto che con la reazione positiva quando il guadagno ad anello aperto $|A\beta|$ tende al valore 1 il guadagno ad anello chiuso G tende ad infinito, indicando che il circuito diventa in grado di fornire un segnale di uscita anche con un segnale di ingresso che scompare (eq. 11.4).

In realtà l’eq. 11.4 sembra fornire un risultato apparentemente plausibile anche nel caso $|A\beta| > 1$: il circuito sembra comportarsi come un amplificatore con un valore di G ben determinato, anche se negativo. L’esperienza invece mostra che in questi casi il circuito è per lo più completamente instabile e preda di oscillazioni spontanee. Inoltre, nello scrivere le eq. 11.3 e 11.4 si è supposto che A e β siano due costanti, in particolare indipendenti dalla frequenza. Questo non è mai vero nei circuiti reali. In qualsiasi amplificatore l’amplificazione A dipende dalla frequenza f (o da $\omega = 2\pi f$) e per $\omega \rightarrow \infty$ si ha sempre $|A| \rightarrow 0$.² Le stesse considerazioni valgono anche per il circuito di reazione β , anche se di solito è abbastanza facile realizzare circuiti puramente passivi il cui comportamento non dipende dalla frequenza fino a frequenze molto superiori a quelle a cui l’amplificatore A ha già smesso di funzionare. In questi casi, con le dovute accortezze, può essere consentito supporre che il circuito β si comporti in modo ideale. In generale, comunque, il guadagno d’anello T è funzione della frequenza:³

¹Nel par. 13.3 si è visto che quest’ultimo risultato non può essere realizzato con una rete perfettamente lineare; la stazionarietà dei modi di oscillazione può essere raggiunta solo al prezzo di introdurre nella rete almeno un elemento non lineare.

²In alcuni circuiti può sembrare che per $\omega \rightarrow \infty$ si abbia $|A| \rightarrow \text{cost.} > 0$. Questo vuol dire solo che alle frequenze più alte il circuito si comporta come un partitore tra ingresso ed uscita, di solito a causa di qualche elemento parassita sempre presente nei dispositivi reali.

³La presenza del segno – dipende esclusivamente dalla convenzione che si adotta nella definizione.

$$T(\omega) = -A(\omega)\beta(\omega) \quad (15.1)$$

Ad una dipendenza di T dalla frequenza è *sempre* associato uno sfasamento. Di conseguenza, in un circuito la reazione può essere negativa per progetto a qualche frequenza e poi diventare positiva quando lo sfasamento complessivo dovuto al variare di A e β con la frequenza raggiunge e magari supera il valore di 180° . Le equazioni 11.3 e 11.4 non sono in grado di dirci che cosa succede in questi casi.

Il problema è che le eq. 11.3 e 11.4 sono state ottenute considerando soltanto la componente stazionaria della risposta di un circuito ad un segnale di ingresso. Sono stati cioè ignorati completamente tutti i comportamenti in condizioni di transitorio.

La stabilità o instabilità di una rete lineare invece è legata ai suoi *modi propri di oscillazione*, cioè al comportamento in regime di transitorio.

15.1 Risposta stazionaria di una rete lineare

La risposta stazionaria di una rete lineare è quella che si ottiene eccitando la rete con generatori stazionari di tensione o corrente a frequenza angolare ω e determinando i valori di correnti e tensioni che si hanno su ogni nodo o maglia della rete dopo che tutti i comportamenti di transitorio si sono estinti.

La risposta stazionaria viene calcolata scrivendo un sistema di equazioni algebriche complesse, una per ogni nodo⁴ oppure una per ogni maglia indipendente⁵ della rete, applicando rispettivamente la prima o la seconda equazione di Kirchhoff:

$$\sum_{p=1}^n I_p = 0 \quad \sum_{q=1}^n V_q = 0 \quad (15.2)$$

Nella analisi per nodi si ottiene il sistema

$$\begin{pmatrix} Y_{11} & -Y_{12} & \dots & -Y_{1n} \\ -Y_{21} & Y_{22} & \dots & -Y_{2n} \\ \vdots & & \ddots & \vdots \\ -Y_{n1} & -Y_{n2} & \dots & +Y_{nn} \end{pmatrix} \cdot \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{pmatrix} = (I_1^G, I_2^G, \dots, I_n^G) \quad (15.3)$$

dove i termini (complessi) Y_{pq} , V_k e I_l^G rappresentano rispettivamente la conduttanza

$$Y_{pq} = \frac{1}{j\omega L_{pq}} + G_{pq} + j\omega C_{pq} \quad (15.4)$$

tra i nodi p e q (fig. 15.3a), il potenziale del nodo k e la corrente entrante nel nodo l per effetto dei generatori indipendenti. I termini Y_{pp} sulla diagonale della matrice \mathbf{Y} sono la somma di tutte le conduttanze Y_{pq} convergenti sul nodo p .

⁴Escluso il *nodo di massa*, cioè il nodo preso come riferimento a cui si assegna convenzionalmente il potenziale 0 e rispetto a cui si misurano le tensioni di tutti gli altri nodi.

⁵Mentre l'identificazione degli n nodi di una rete è univoca, esistono in genere più modi di identificare le maglie. Ognuno dei possibili percorsi chiusi lungo i rami di una rete è una maglia, ma di queste solo n sono descritte da equazioni linearmente indipendenti.

Nella analisi per maglie invece si ottiene il sistema

$$\begin{pmatrix} Z_{11} & -Z_{12} & \dots & -Z_{1n} \\ -Z_{21} & Z_{22} & \dots & -Z_{2n} \\ \vdots & & & \vdots \\ -Z_{n1} & -Z_{n2} & \dots & +Z_{nn} \end{pmatrix} \cdot \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{pmatrix} = (V_1^G, V_2^G, \dots, V_n^G) \quad (15.5)$$

dove Z_{pq} è l'impedenza complessa

$$Z_{pq} = \frac{1}{j\omega C_{pq}} + R_{pq} + j\omega L_{pq} \quad (15.6)$$

del ramo in comune tra le maglie p e q (fig. 15.3b), le I_k sono le *correnti di maglia* e le V_l^G le *f.e.m.* dei generatori indipendenti presenti nelle maglie. Le Z_{pp} sono la somma di tutte le impedenze Z_{pq} della maglia p .

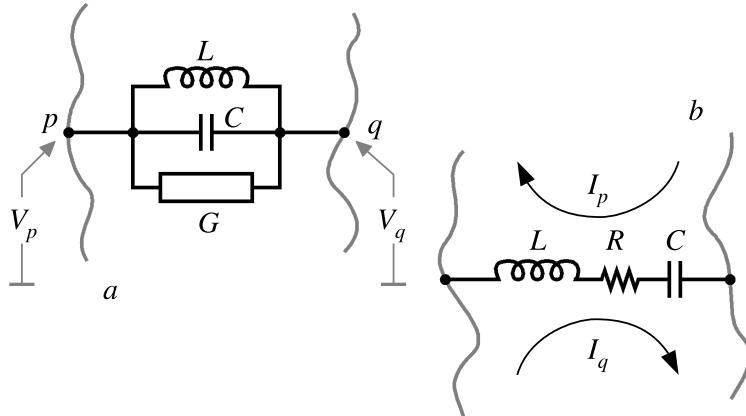


Figura 15.3: a) conduttanza tra due nodi di una rete: $Y_{pq} = j\omega C + G + 1/(j\omega L)$; b) impedenza di un ramo di una rete: $Z_{pq} = j\omega L + R + 1/(j\omega C)$.

La soluzione del sistema 15.3 fornisce il vettore delle variabili incognite $V_k(\omega)$ in funzione dei termini noti $I_l^G(\omega)$ e analogamente per il sistema 15.5 con le incognite $I_k(\omega)$ ed i termini noti $V_l^G(\omega)$.⁶

15.2 Comportamento durante un transitorio

Come già detto, le soluzioni algebriche dei due sistemi 15.3 e 15.5 descrivono soltanto il comportamento stazionario della rete, in quanto si considerano solo le soluzioni del tipo $e^{j\omega t}$, con ω reale, in presenza di soli termini forzanti stazionari ($I_1^G \dots I_n^G$) oppure ($V_1^G \dots V_n^G$)

⁶In realtà, nella pratica del calcolo si utilizza una *analisi per nodi modificata* per poter introdurre nella eq. 15.3 anche generatori di potenziale indipendenti V_m^G tra i termini noti a destra; alcuni dei termini del vettore (V_k) a sinistra nella 15.3 saranno quindi correnti ed alcuni dei termini della matrice \mathbf{Y} avranno dimensioni di impedenze invece che di ammettenze. Il programma *Spice* utilizza questa tecnica di calcolo. Per quanto riguarda le considerazioni di questi paragrafi sulla stabilità delle reti, l'utilizzo del sistema 15.3, 15.5 o della analisi per nodi modificata è completamente equivalente.

Il sistema algebrico 15.3 deriva dal sistema di equazioni differenziali lineari

$$\begin{pmatrix} y_{11} - y_{12} - & \dots & -y_{1n} \\ -y_{21} + y_{22} - & \dots & -y_{2n} \\ \vdots & & \vdots \\ -y_{n1} - y_{n2} - & \dots & +y_{nn} \end{pmatrix} \cdot \begin{pmatrix} v_1(t) \\ v_2(t) \\ \vdots \\ v_n(t) \end{pmatrix} = (i_1^G(t), i_2^G(t), \dots, i_n^G(t)) \quad (15.7)$$

in cui le i e v sono funzioni del tempo ed i termini y_{pq} sono gli operatori lineari

$$y_{pq} = \frac{1}{L_{pq}} \int dt + G_{pq} + C_{pq} \frac{d}{dt} \quad (15.8)$$

Come noto dall'Analisi, le soluzioni più generali di questo sistema di equazioni sono combinazioni lineari delle risposte stazionarie e dei comportamenti di transitorio del sistema nella condizione in cui tutti i termini forzanti sono nulli. Le soluzioni che descrivono il regime di transitorio sono funzioni del tipo e^{st} , dove la variabile $s = \alpha + j\omega$ è denominata *frequenza complessa*.

Inserendo queste soluzioni nel sistema 15.7, ponendo a zero i termini forzanti (i_1^G, \dots, i_n^G) , sostituendo la derivata di e^{st} con se^{st} e l'integrale con e^{st}/s , si arriva nuovamente ad un sistema di equazioni algebriche analogo al sistema 15.3 (o al 15.5) con la semplice sostituzione $j\omega \rightarrow s$:

$$\mathbf{Y}(s) \cdot \mathbf{V}(s) = 0 \quad (15.9)$$

Se la matrice \mathbf{Y} è non singolare, il sistema ha come unica soluzione

$$V_1 = V_2 = \dots = V_n = 0 \quad (15.10)$$

Solo nel caso che la matrice \mathbf{Y} abbia determinante nullo, allora possono esistere soluzioni con $V_k \neq 0$.

Nel dominio delle frequenze complesse $s = \alpha + j\omega$ i termini della matrice \mathbf{Y} sono espressioni del tipo:

$$Y_{pq} = \frac{1}{sL_{pq}} + G_{pq} + sC_{pq} = \frac{1}{s} \left(C_{pq}s^2 + G_{pq}s + \frac{1}{L_{pq}} \right) \quad (15.11)$$

per cui il determinante di \mathbf{Y} è sempre un polinomio a coefficienti reali della variabile s diviso per una qualche potenza di s :

$$\det \mathbf{Y} = \frac{P(s)}{s^m} \quad (15.12)$$

In base al teorema fondamentale dell'algebra il polinomio $P(s)$ avrà radici reali o complesse coniugate in numero pari al suo grado. Per ogni radice $s_k = \alpha_k + j\omega_k$ di $P(s)$ il determinante di \mathbf{Y} si annulla ed il sistema 15.3 ha una soluzione

$$v_k(t) = A_k e^{s_k t} = A_k e^{\alpha_k t} \cdot e^{\pm j\omega_k t} \quad (15.13)$$

che costituisce un *modo naturale* di oscillazione del sistema.

Le soluzioni con s_k reale corrispondono a rilassamenti ($\alpha < 0$) o derive ($\alpha > 0$) di tipo esponenziale

$$v(t) = a_k e^{\alpha_k t} \quad (15.14)$$

Le soluzioni con s_k complesso sono sempre a coppie complesse coniugate:

$$v(t) = a_k e^{\alpha_k t} \left[e^{j(\omega_k t + \varphi)} + e^{-j(\omega_k t + \varphi)} \right] \quad (15.15)$$

e corrispondono quindi ad oscillazioni smorzate, con $\alpha < 0$, o crescenti, con $\alpha > 0$. Nel caso molto particolare di $\alpha = 0$ si avrà una oscillazione persistente di ampiezza costante.

15.3 Condizione di stabilità di un anello di reazione

La definizione di sistema stabile implica che tutte le risposte di transitorio debbano decadere a zero nel tempo. Quindi perchè una rete sia stabile si deve avere $\alpha_k < 0$ per ogni valore s_k nelle soluzioni 15.13. Questo equivale a dire che tutti gli zeri del determinante della matrice \mathbf{Y} si devono trovare nella metà sinistra del piano delle frequenze complesse $s = \alpha + j\omega$ (fig. 15.4).

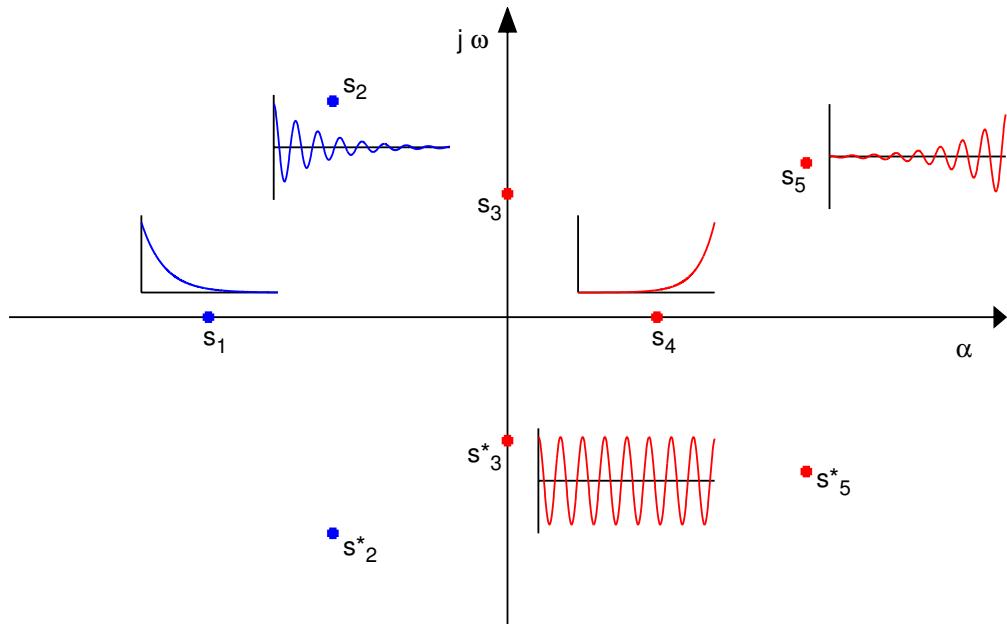


Figura 15.4: Possibili posizioni degli zeri del determinante della matrice \mathbf{Y} sul piano delle frequenze complesse $\alpha + j\omega$. In blu i comportamenti *stabili*, in rosso i comportamenti *instabili*. s_1 : rilassamento esponenziale $e^{-\alpha t}$; s_2 : oscillazione smorzata $e^{-\alpha t}(e^{j\omega t} + e^{-j\omega t})$; s_3 : oscillazione persistente stazionaria $e^{j\omega t} + e^{-j\omega t}$ ($\alpha = 0$); s_4 : deriva esponenziale $e^{\alpha t}$; s_5 : oscillazione crescente in modo esponenziale $e^{\alpha t}(e^{j\omega t} + e^{-j\omega t})$.

Nel caso dell'anello di reazione del cap. 11 (fig. 11.2 - 11.3) è molto importante poter stabilire se con una rete A stabile ed una rete β stabile al momento che si

chiude l'anello la rete risultante è stabile o instabile. Si è già visto che la stabilità delle reti A e β prese singolarmente non è affatto garanzia della stabilità dell'anello, anzi nel par. 13.3 reti stabili sono state combinate appositamente per ottenere un comportamento complessivo instabile.

Riscrivendo le equazioni 11.2-11.3 sotto forma di matrici

$$\begin{pmatrix} A & -1 \\ 1 & \beta \end{pmatrix} \cdot \begin{pmatrix} v_i \\ v_o \end{pmatrix} = (0 \quad v_s) \quad (15.16)$$

valgono tutte le considerazioni già fatte nei par. 15.1-15.2 e quindi la stabilità dell'anello è legata ai *modi naturali* della rete, determinati di nuovo dall'azzerarsi del determinante della matrice del sistema, che questa volta è $1 + A\beta$. Se gli zeri di $1 + A\beta$ si trovano tutti nel semipiano sinistro delle frequenze complesse $s = \alpha + j\omega$ la rete è stabile, altrimenti è instabile.

15.4 Anelli del primo e del secondo ordine

La determinazione della condizione di stabilità è abbastanza agevole nei casi semplici di anelli di reazione del primo e del secondo ordine, cioè che contengono rispettivamente solo una o due frequenze di taglio date da celle di tipo passa-basso.

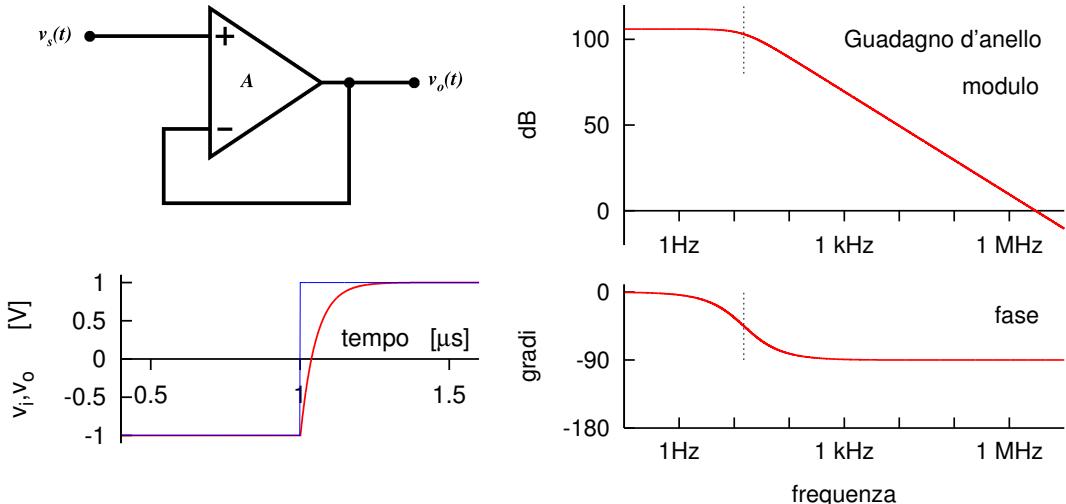


Figura 15.5: Anello del primo ordine: amplificatore operazionale con una singola frequenza di taglio $f_0 = 15$ Hz ($\tau = 10.6$ ms) ed amplificazione $A = 2 \cdot 10^5$, utilizzato come *voltage follower* ($\beta = 1$). A destra sono riportati i diagrammi di Bode del guadagno d'anello (modulo e fase). In basso a sinistra, sotto lo schema elettrico, è riportata la risposta in funzione del tempo del circuito ad anello chiuso. All'ingresso è applicato un segnale a gradino (curva blu). All'uscita (curva rossa) si ha la risposta esponenziale caratteristica di un filtro passa basso con costante di tempo $\tau' = \tau/(A\beta + 1) = 53$ ns.

Ad esempio, combinando un amplificatore caratterizzato da una sola frequenza di taglio con una rete di reazione puramente resistiva si ottiene

$$1 + A\beta = 1 + \frac{A_0}{1 + s\tau} \beta = \frac{1 + s\tau + A_0\beta}{1 + s\tau} \quad (15.17)$$

dove τ è la costante di tempo RC associata alla cella passa-basso. Ponendo $1 + A\beta = 0$ si ottiene una sola soluzione, reale e negativa:

$$s_1 = -\frac{1 + A_0\beta}{\tau} \quad (15.18)$$

che corrisponde ad un comportamento stabile, caratterizzato da un rilassamento esponenziale puro con costante di tempo $\tau' = \tau/(1 + A_0\beta)$. In fig. 15.5 è riportato un esempio di un anello del primo ordine.

In un anello del primo ordine si può avere una condizione di instabilità esclusivamente invertendo il segno della reazione: se nel circuito di fig. 15.5 si scambiano i due ingressi $+e-$ dell'amplificatore, la soluzione s_1 della eq. 15.18 cambia di segno e la risposta del circuito è una rapidissima deriva esponenziale della tensione di uscita fino a raggiungere il valore di saturazione positiva o negativa dell'amplificatore, a seconda delle condizioni iniziali.

15.4.1 Anello del secondo ordine

L'anello del secondo ordine presenta due frequenze di taglio, f_1 ed f_2 , determinate da due celle passa basso passive⁷ con costanti di tempo τ_1 e τ_2 , come nell'esempio di fig. 15.6.

L'equazione 15.17 diventa

$$1 + A\beta = \frac{(1 + s\tau_1)(1 + s\tau_2) + A_0\beta}{(1 + s\tau_1)(1 + s\tau_2)} = \frac{s^2\tau_1\tau_2 + s(\tau_1 + \tau_2) + 1 + A_0\beta}{(1 + s\tau_1)(1 + s\tau_2)} \quad (15.19)$$

Gli zeri del determinante $1 + A\beta$ sono dati ora dai valori di s

$$s_{1,2} = -\frac{1}{2} \left(\frac{1}{\tau_1} + \frac{1}{\tau_2} \right) \pm \frac{1}{2} \sqrt{\frac{1}{\tau_1\tau_2} \cdot \left(\frac{\tau_1}{\tau_2} + \frac{\tau_2}{\tau_1} - 2 - 4A_0\beta \right)} \quad (15.20)$$

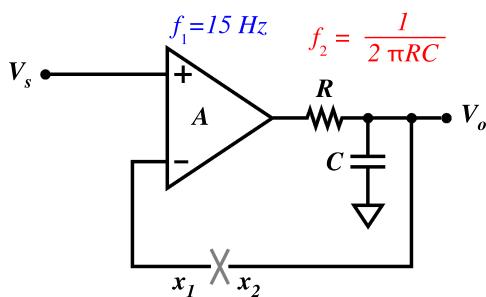


Figura 15.6: Anello del secondo ordine con due frequenze di taglio con comportamento di tipo *passa basso*. La prima frequenza, $f_1 = 15 \text{ Hz}$, è data dalla risposta dell'amplificatore operazionale, la seconda, f_2 , dal gruppo RC .

⁷Escludiamo il caso che nelle celle passa basso possano comparire elementi come una resistenza o una capacità negativa, generate da un circuito NIC *attivo*.

Quando l'espressione sotto radice è negativa si hanno due soluzioni complesse e coniugate, quando è positiva due soluzioni reali. Nel primo caso la risposta è una oscillazione armonica smorzata (la parte reale di $s_{1,2}$ è sempre negativa), nel secondo caso un rilassamento semplice, senza oscillazione.

Quando l'espressione sotto radice si annulla si ha la condizione di *smorzamento critico*, che corrisponde alla più rapida risposta senza oscillazione.

Nell'ipotesi $A_0\beta \gg 1$, come è la regola quando l'amplificatore A è un operazionale, nell'equazione

$$\frac{\tau_1}{\tau_2} + \frac{\tau_2}{\tau_1} - 2 - 4A_0\beta = 0 \quad (15.21)$$

il termine 2 risulta trascurabile e così uno dei due tra τ_1/τ_2 e τ_2/τ_1 , che sono uno l'inverso dell'altro. Supponendo ad esempio $\tau_1 > \tau_2$, si ha lo smorzamento critico per

$$\frac{\tau_1}{\tau_2} \simeq 4A_0\beta \quad \text{e} \quad s \simeq \frac{1}{2\tau_2} \quad (15.22)$$

Nella figura 15.7 sono riportati i diagrammi di Bode e la risposta in funzione del tempo del circuito 15.6 per diversi valori della frequenza f_2 , per $A_0 = 2 \cdot 10^5$. A

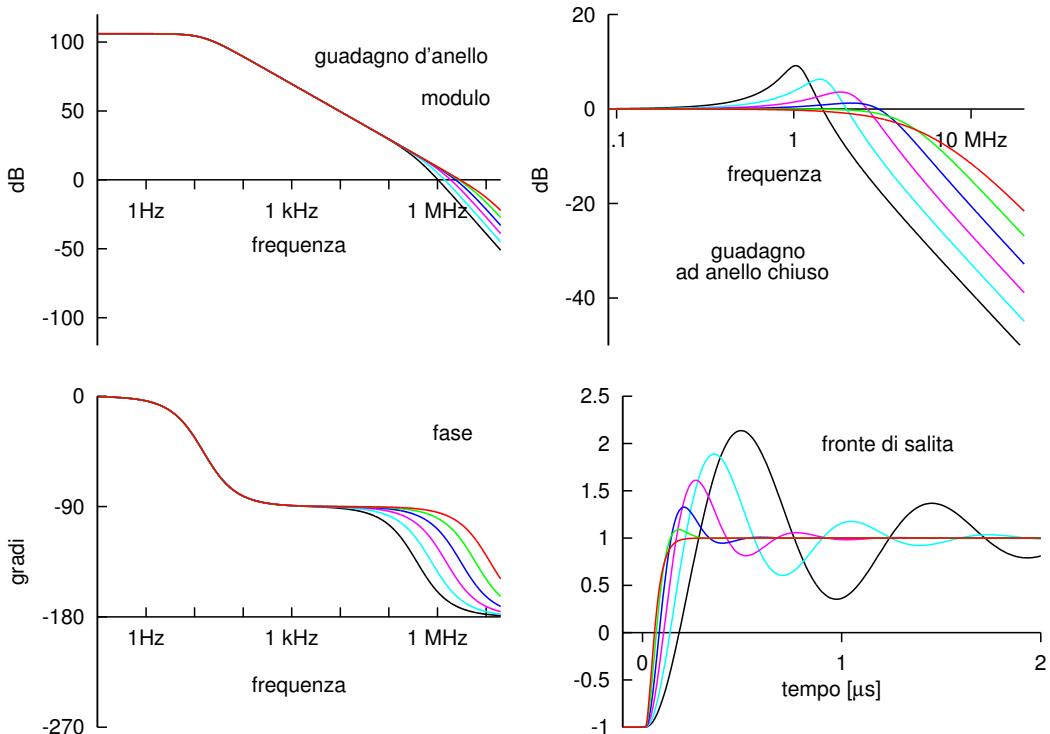


Figura 15.7: Risposta in frequenza e nel tempo per l'anello del secondo ordine. A sinistra: diagrammi di Bode del guadagno ad anello aperto; a destra in alto diagramma di Bode del modulo del guadagno ad anello chiuso; a destra in basso, risposta nel tempo ad anello chiuso ad un segnale a gradino. $f_1 = 15 \text{ Hz}$, $f_2 = 12, 6, 3, 1.5, 0.75 \text{ e } 0.375 \text{ MHz}$ per le curve, nell'ordine, rosse, verdi, blu, viola, celesti e nere.

sinistra è riportato il guadagno dell'anello aperto in funzione della frequenza, in alto il modulo, in basso la fase. A destra è riportata la risposta ad anello chiuso, in alto il modulo del guadagno in funzione della frequenza, in basso la risposta in funzione del tempo ad un segnale di ingresso a gradino.

Le curve in rosso corrispondono alla condizione di smorzamento critico, che si ha per la frequenza $f_2 = f_1 \cdot 4A_0\beta$. Con i valori indicati, $f_2 = 12 \text{ MHz}$. Le altre curve invece si riferiscono a valori di f_2 progressivamente più bassi, con una riduzione di un fattore due tra una curva e la successiva, quindi 6, 3, 1.5, 0.75 e 0.375 MHz rispettivamente per le curve verdi, blu, viola, azzurre e nere.

Quando ci si allontana dalla condizione di smorzamento critico nel verso che la frequenza f_2 diminuisce, entrambi i termini τ_1/τ_2 e τ_2/τ_1 diventano rapidamente piccoli rispetto a $4A_0\beta$. In queste condizioni, la frequenza dell'oscillazione smorzata è

$$f = \frac{\omega}{2\pi} \simeq \frac{1}{2\pi} \sqrt{\frac{A_0\beta}{\tau_1\tau_2}} \quad (15.23)$$

e la costante di tempo dello smorzamento

$$\alpha = \frac{1}{2} \left(\frac{1}{\tau_1} + \frac{1}{\tau_2} \right) \simeq \frac{1}{2\tau_2} \quad (15.24)$$

Nel caso della curva nera ($f_2 = 0.375 \text{ MHz}$, $\tau_2 = 0.42 \mu\text{s}$) si ha $f \simeq 1.06 \text{ MHz}$ e $1/\alpha \simeq 0.84 \mu\text{s}$.

Come si vede dai grafici, il comportamento del sistema rimane sempre *stabile* secondo la definizione, perché l'oscillazione transitoria (risonanza o *ringing*) provocata dai fronti di salita o di discesa del segnale di ingresso decade a zero e si estingue. Ma la deformazione del segnale che si ha per valori di f_2 via via più piccoli è evidente ed in molte applicazioni inaccettabile.

Di particolare interesse sono le prime tre curve dei grafici, rosse, verdi e blu, corrispondenti come si è detto ai valori di $f_2 = 12$, 6, e 3 MHz . La prima, la rossa, corrisponde, come già descritto, alla condizione di smorzamento critico.

La seconda, la verde, corrisponde ad una condizione di smorzamento lievemente *sottocritico*: non si ha una vera e propria coda di oscillazione smorzata, ma solo un leggero *overshoot* nella risposta in funzione del tempo, cioè un picco che va oltre il valore che deve essere raggiunto, seguito da un veloce rilassamento. La corrispondente risposta in frequenza presenta la caratteristica di essere *massimamente piatta*: questa quindi è la configurazione per un amplificatore che fornisce la amplificazione più uniforme possibile in funzione della frequenza. Qualitativamente questo si vede nel grafico in alto a destra in fig. 15.7 confrontando la curva verde con la rossa e la blu. Analiticamente si ottiene dalla espressione del guadagno ad anello chiuso

$$G(\omega) = \frac{A_0}{1 + A_0\beta + j\omega(\tau_1 + \tau_2) - \omega^2\tau_1\tau_2} \quad (15.25)$$

Derivando il modulo di $G(\omega)$ rispetto alla frequenza angolare ω si vede che questo presenta un massimo (o un minimo) per $\omega = 0$ ed un altro massimo o minimo ad una frequenza che dipende da τ_2 e quindi da f_2 . La risposta *massimamente piatta* si ha quando anche questo massimo o minimo si trova ad $\omega = 0$. Con la solita ipotesi $A_0\beta \gg 1$, si ottiene $f_2 = 2A_0\beta f_1$, cioè la condizione con cui è stata disegnata la curva verde ($f_1 = 15 \text{ Hz}$, $f_2 = 6 \text{ MHz}$, $A_0 = 2 \cdot 10^5$, $\beta = 1$).

La curva blu corrisponde al caso particolare $f_2 = A_0\beta f_1$. Anche qui siamo in condizioni di smorzamento *sottocritico*, con una lieve oscillazione fortemente smorzata.

Nella pratica la scelta tra l'una o l'altra di queste soluzioni è dettata dall'applicazione a cui il circuito è destinato e dalle preferenze. Con $f_2 = 4A_0\beta f_1$ si ha la risposta più veloce nel tempo, ma non la più uniforme nelle frequenze; con $f_2 = 2A_0\beta f_1$ si ha la risposta massimamente piatta in frequenza, ma un leggero *overshoot* nel tempo sui fronti di salita veloci; la condizione $f_2 = A_0\beta f_1$ spesso è già un compromesso accettabile.

Nel caso opposto rispetto a quelli considerati fino ad ora, cioè quando τ_2 diminuisce rispetto al valore che ha allo smorzamento critico e tende a zero, il sistema rimane con la sola costante di tempo τ_1 e si riduce al sistema del primo ordine visto all'inizio del paragrafo.

15.5 Criterio di Nyquist

Nei circuiti più complessi, caratterizzati da più di due costanti di tempo, la determinazione della condizione di stabilità o instabilità tramite il calcolo della posizione degli zeri del determinante non è sempre agevole.

A volte di una rete, o di un sistema, è nota solo la risposta (stazionaria) in funzione della frequenza. Questa comunque è quanto può essere determinato da misure di laboratorio; da questa si dovrebbe ricavare un circuito equivalente con le costanti di tempo opportune, per poi effettuare l'analisi dei par. 15.2-15.3.

In realtà allo scopo di determinare se un anello di reazione, composto da una rete stabile A e da una rete stabile β , è stabile oppure no, non è necessario sapere *dove* si trovino esattamente gli zeri del determinante nel piano complesso, ma unicamente *se* ve ne siano nel semipiano destro (o magari esattamente sull'asse $j\omega$).

A questa domanda si può rispondere in maniera abbastanza semplice con la sola conoscenza delle funzioni di risposta $A(\omega)$ e $\beta(\omega)$ delle due reti, utilizzando il criterio di Nyquist.

Dall'Analisi è noto che tramite l'integrale di Cauchy è possibile determinare la presenza degli zeri e dei poli di una funzione complessa (di caratteristiche opportune) all'interno di una regione del piano complesso mediante un integrale limitato al solo bordo della regione. Le funzioni considerate in questo capitolo per la descrizione delle reti lineari sono soltanto polinomi e quindi rispondono ai requisiti dell'Analisi.

Il criterio di Nyquist stabilisce che per un anello di reazione la presenza nel semipiano destro di zeri del determinante $1 + A\beta$ (del sistema 15.16) è rivelata dal comportamento del guadagno d'anello $T(\omega) = -A(\omega)\beta(\omega)$ sull'asse delle frequenze reali $j\omega$.⁸

L'enunciato del criterio afferma che, riportando sul piano complesso il tracciato della funzione $T(\omega)$ per ω che va da $-\infty$ a $+\infty$, questo può racchiudere al suo interno il punto del piano di coordinate $\{-1, 0\}$, oppure no. Nel primo caso il determinante del sistema che descrive la rete ha zeri nella metà destra del piano complesso e quindi la rete è instabile; nel secondo caso gli zeri sono tutti confinati nel semipiano sinistro e la rete è stabile. Il numero di volte che il tracciato racchiude il punto $\{-1, 0\}$

⁸L'asse $j\omega$ del piano complesso s è indicata spesso come *asse delle frequenze reali*.

corrisponde al numero di zeri nel semipiano destro. Ogni volta che il tracciato passa esattamente per il punto $\{-1, 0\}$, uno zero si trova sull'asse $j\omega$.

Date le caratteristiche delle funzioni $A(\omega)$ e $\beta(\omega)$ per le reti realizzate con componenti reali, vale sempre la relazione $A(-\omega) = A^*(\omega)$ e analogamente $\beta(-\omega) = \beta^*(\omega)$ e il limite per $\omega \rightarrow \pm\infty$ è sempre zero oppure un valore reale, per cui il tracciato di $A(\omega)\beta(\omega)$ sul piano complesso è sempre una curva chiusa.

Se il punto $\{-1, 0\}$ del piano complesso rimane sulla destra nel percorrere il tracciato per ω che va da $-\infty$ a $+\infty$, allora si trova all'interno; se rimane sulla sinistra si trova all'esterno.

15.5.1 Anello di reazione stabile

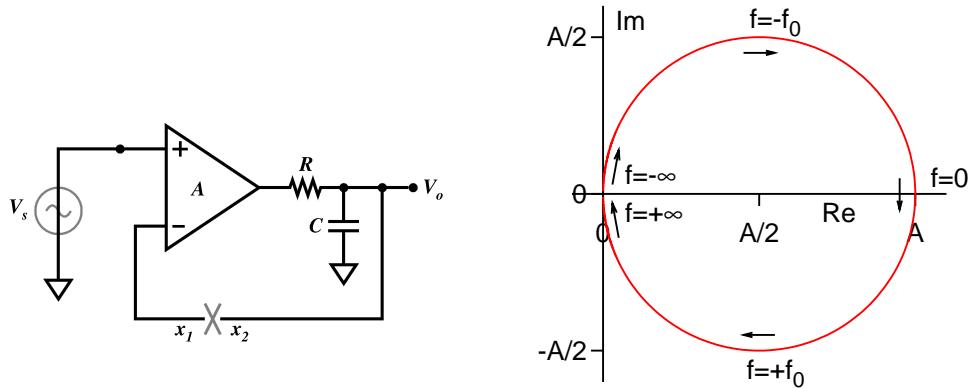


Figura 15.8: Anello di reazione stabile.

Nella figura 15.8 è riportato come esempio l'anello stabile del primo ordine del par. 15.4. Il circuito è stato disegnato come un amplificatore differenziale ideale con guadagno A e banda passante infinita ed una cella passa basso con frequenza di taglio $\omega_0 = (RC)^{-1}$. Chiudendo l'anello di reazione nel punto contrassegnato con la X si ha il circuito del *voltage follower* del par. 11.7.1. Ad anello aperto, sostituendo il generatore V_s con un cortocircuito, si ha un guadagno d'anello⁹ dal terminale x_1 al terminale x_2

$$T(\omega) = \frac{A}{1 + j\omega/\omega_0} \quad (15.26)$$

Nella figura 15.8, a destra, è riportato il relativo diagramma di Nyquist. Per $\omega = 0$ il guadagno T è reale e pari ad A ; per $\omega = \pm\omega_0$ si hanno i due punti $(1 \mp j)A/2$ e per $\omega \rightarrow \pm\infty$ il guadagno va a zero. Il tracciato rimane tutto nel semipiano destro e viene percorso nel verso indicato dalle frecce per $\omega = 2\pi f$ da $-\infty$ a $+\infty$. Qualunque sia il valore di A , il punto $\{-1, 0\}$ rimarrà comunque al di fuori del tracciato. Il circuito in fig. 15.8 è *intrinsecamente stabile*.

⁹Il calcolo va sempre fatto lungo il percorso effettivo dell'anello di reazione. In questo caso: $T_{x_1 \rightarrow x_2} = -(-A) \cdot \beta = A\beta$. Il primo segno – deriva dalla definizione di $T(\omega)$, il secondo dal segno negativo del guadagno per un segnale applicato all'ingresso – dell'amplificatore.

15.5.2 Anello di reazione instabile

Nella figura 15.9 l'anello di reazione di fig. 15.8 è stato modificato sostituendo la singola cella RC con una successione di tre celle RC identiche. La funzione di

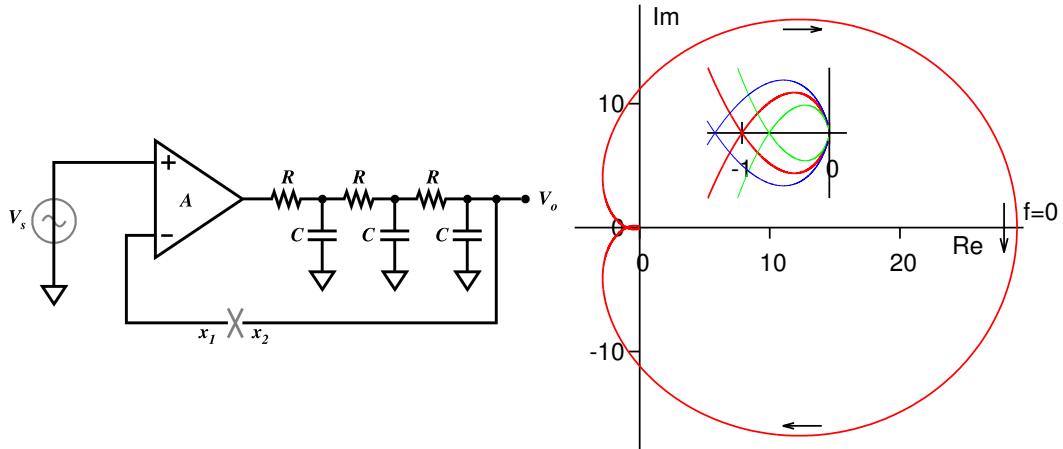


Figura 15.9: Esempio di anello di reazione instabile. Il valore di A è 29 per la curva rossa, 20 per la verde e 38 per la blu.

trasferimento di tale rete è

$$H(f) = \frac{1}{1 + j6\left(\frac{f}{f_o}\right) - 5\left(\frac{f}{f_o}\right)^2 - j\left(\frac{f}{f_o}\right)^3} \quad (15.27)$$

con $f_o = 1/(2\pi RC)$.

Nella figura, a destra, è riportato il diagramma di Nyquist del guadagno d'anello $A \cdot H(f)$ per $A = 29$. Il tracciato passa esattamente per il punto $\{-1, 0\}$ in corrispondenza della frequenza $f = \sqrt{6}f_o$. Infatti, sostituendo questo valore nella eq. 15.27 la parte immaginaria a denominatore si annulla, lasciando $H(\sqrt{6}f_o) = -1/29$. Nel grafico in piccolo, all'interno a destra, è riportata ingrandita la parte del piano complesso tra $\{-1, 0\}$ e $\{0, 0\}$ per tre valori di A : 29 (curva in rosso), 20 (curva in verde) e 38 (curva in blu).

Per ogni valore di A maggiore di 29 il tracciato di Nyquist racchiude il punto $\{-1, 0\}$ (due volte) e la rete è instabile. Per $A < 29$ il punto $\{-1, 0\}$ rimane esterno al tracciato e la rete è stabile. Il caso particolare $A = 29$ corrisponde ad avere una coppia di soluzioni complesse e coniugate esattamente sull'asse $j\omega$, come le s_3 nella figura 15.4, e quindi un comportamento della rete instabile, ma stazionario. Difatti, il circuito di fig. 15.9 è utilizzato come generatore di segnali sinusoidali a frequenze audio ed è denominato *oscillatore a rotazione di fase* (*phase shift oscillator*).¹⁰

¹⁰Poichè la situazione “ A esattamente eguale a 29” non è fisicamente realizzabile (e poi anche la rete RC non sarà né esatta né perfettamente stabile) il circuito reale dovrà contenere qualche meccanismo di regolazione che provveda a *ritoccare* il valore di A in modo da compensare ogni errore e deriva. Nella pratica, un metodo rozzo ma efficace, come già visto, è fare A leggermente maggiore

15.6 Il criterio di Nyquist nei diagrammi di Bode

Quando si ha a che fare con circuiti con guadagno molto elevato, dell'ordine di $A = 10^5 \dots 10^6$, come è tipico per un amplificatore operazionale, diventa impossibile stabilire dal diagramma di Nyquist se il punto $\{-1, 0\}$ si trovi dentro o fuori del tracciato. Già nell'esempio di fig. 15.9 con un guadagno tra 20 e 38 è stato necessario ripetere il grafico a scala maggiore nell'intorno dell'origine per poter individuare il comportamento del circuito. Per questo motivo nella pratica l'utilizzo del grafico di Nyquist è poco conveniente.

I diagrammi di Bode risultano più convenienti e di facile applicazione, sia per determinare le condizioni di stabilità o instabilità di un circuito sia per individuare le eventuali modifiche necessarie per portare un circuito nelle condizioni di stabilità volute. Nei diagrammi di Bode viene riportata l'amplificazione di una rete, espressa come modulo (in dB) e fase, entrambi in funzione del logaritmo della frequenza. Il fatto che il modulo sia espresso in dB , quindi su scala logaritmica, permette di esaminare con lo stesso dettaglio sia la regione intorno ad esempio a 10^5 o 10^6 sia quella intorno a 10^0 , eliminando quindi l'inconveniente visto prima per i grafici di Nyquist.

Il punto $\{-1, 0\}$ nel piano complesso di Nyquist corrisponde nei diagrammi di Bode ai punti in cui il guadagno è 0 dB , cioè ha modulo 1, e la fase ha raggiunto il valore di -180° .

Nel caso semplice in cui modulo e fase del guadagno d'anello sono funzioni monotone della frequenza la condizione che il tracciato di Nyquist racchiuda o meno il punto di coordinate $\{-1, 0\}$ equivale nel diagramma di Bode al fatto che il modulo del guadagno arrivi a 0 dB dopo o prima che la fase abbia oltrepassato il valore -180° .

I casi più complessi, in cui modulo e fase non sono funzioni monotone, verranno considerati in seguito.

Come esempio, nella fig. 15.10 sono riportate le funzioni di guadagno d'anello del circuito di fig. 15.9 per i tre valori di A già considerati e per una rete RC con $f_0 = 1 \text{ kHz}$.

La rotazione di fase è descritta dalla curva rossa in basso e non dipende da A , ma solo dalla funzione $H(f)$ della rete RC ; il valore di -180° viene raggiunto, come si è visto, alla frequenza $f_0 \cdot \sqrt{6} = 2449 \text{ Hz}$. La curva rossa in alto, corrispondente ad $A = 29$, passa per il valore 0 dB esattamente alla frequenza $f_0 \cdot \sqrt{6}$. La curva verde, corrispondente al guadagno 20, raggiunge il valore 0 dB prima che la fase abbia raggiunto i -180° e corrisponde quindi al comportamento stabile; all'opposto, la curva blu raggiunge il valore 0 dB quando la fase ha già superato i -180° e corrisponde al comportamento instabile.

I comportamenti in funzione del tempo per i tre casi sono riportati nel grafico fig. 15.11.

di 29 e lasciare che l'oscillazione cresca fino alla saturazione dell'amplificatore. Si ottiene la stabilità dell'ampiezza dell'oscillazione (stato stazionario) al prezzo di introdurre una lieve distorsione nel segnale. Non si avrà più una sinusoide pura, ma comparirà una piccola quantità di armoniche superiori, a frequenze $n \cdot \sqrt{6}f_0$, principalmente con n dispari per ragioni di simmetria. Altre tecniche più raffinate (e più complesse) consentono di ottenere lo stesso risultato per quanto riguarda la stabilizzazione dell'ampiezza, mantenendo livelli di distorsione nel segnale molto bassi.

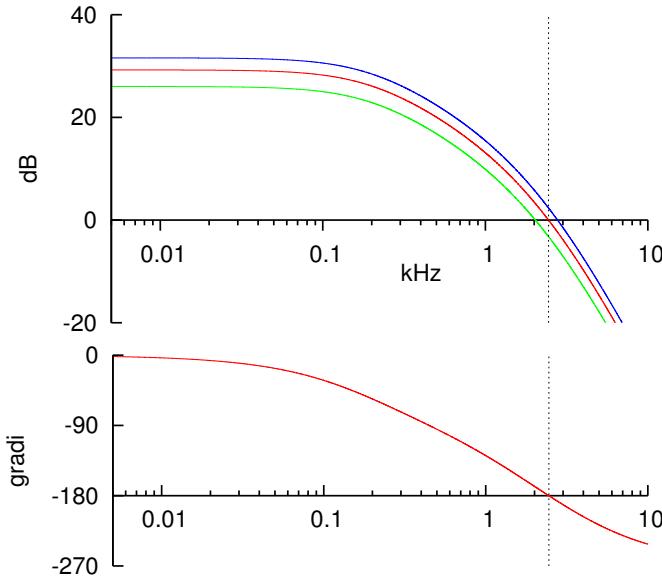


Figura 15.10: Diagrammi di Bode per il guadagno d'anello del circuito in fig. 15.9. Le tre curve in alto corrispondono ai valori di guadagno $A = 20$ (verde), 29 (rosso) e 38 (blu). La curva in basso mostra la variazione della fase, che dipende solo dalla funzione $H(f)$ e quindi è indipendente dal valore di A .

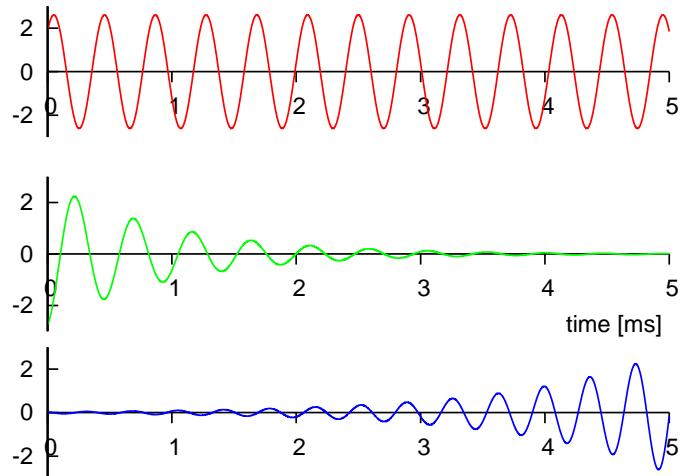


Figura 15.11: Comportamento ad anello chiuso in funzione del tempo della rete di fig. 15.9 per $A = 29$ (rosso), $A = 20$ (verde) e $A = 38$ (blu).

15.7 Margine di fase e di guadagno

L'anello di reazione del secondo ordine con due frequenze di taglio f_1 ed f_2 , esaminato nel par. 15.4.1, dovrebbe essere in teoria altrettanto intrinsecamente stabile quanto l'anello del primo ordine. La rotazione di fase del guadagno d'anello

$$T(\omega) = \frac{A_0}{(1 + j\omega/f_1)(1 + j\omega/f_2)} \quad (15.28)$$

tende asintoticamente a -180° senza mai arrivarcì; non si dovrebbe quindi mai arrivare a soddisfare la condizione di instabilità di Nyquist. Questo però equivale a dire che l'anello di reazione del secondo ordine può arrivare vicino quanto si vuole alla condizione di instabilità, cioè avere un comportamento di transitorio $s = \alpha + j\omega$ con α piccolo a piacere e quindi con tempi di smorzamento del transitorio lunghissimi.

Questo si vede chiaramente nei grafici 15.7 e 15.10–15.11: quanto più la fase arriva vicina al valore di -180° quando il guadagno dell'anello arriva 0 dB, tanto più il circuito si avvicina al limite della instabilità e presenta risonanze sempre più lunghe nella risposta temporale.

Si definisce *margine di fase* la differenza tra il valore della fase alla frequenza a cui il guadagno è arrivato a 0 dB e il valore -180° .

Per i sei casi di fig. 15.7 i margini di fase sono rispettivamente 76° , 66° , 52° , 39° , 28° e 20° , dal più stabile al meno stabile. Per il circuito di fig. 15.9 il margine di fase è 11° con guadagno di 20 e 0° con guadagno 29. Con guadagno 38 il margine di fase è negativo ed il circuito è in condizioni di instabilità.

Come regola di massima, il margine di fase dovrebbe essere non inferiore ai 45° per avere nelle risposte *overshoot* e *ringing* accettabili.

Per i circuiti in cui la rotazione di fase può andare oltre i -180° si può definire analogamente un *margine di guadagno*, corrispondente a quanto il guadagno è sceso sotto il valore di 0 dB quando la rotazione di fase raggiunge i -180° . Ad esempio, per il circuito di fig. 15.9 con $A = 20$, il margine di guadagno è 3.2 dB.

15.8 Compensazione degli anelli di reazione

Il circuito in fig. 15.12 è un anello del secondo ordine come quello del par. 15.4 e schematizza un sistema di controllo automatico. Il filtro passa basso con frequenza di taglio di 10 kHz all'uscita dell'amplificatore rappresenta il sistema fisico controllato, con una sua costante di tempo caratteristica di $16\mu s$. L'altra frequenza di taglio, $f_1 = 15\text{ Hz}$, è data dall'amplificatore operazionale.

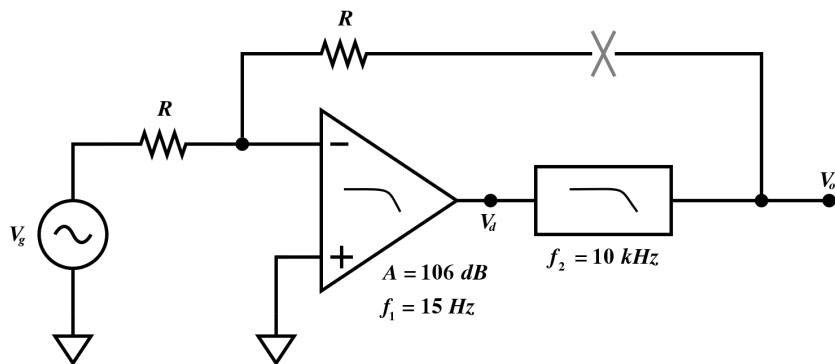


Figura 15.12: Anello di controllo del secondo ordine. La prima frequenza di taglio è data dall'amplificatore (15 Hz); la seconda, 10 kHz, è data dal sistema fisico controllato, che si suppone abbia un risposta di tipo *passa basso* con costante di tempo di $16\mu s$.

I grafici in fig. 15.13 riportano il guadagno d'anello e la risposta temporale ad anello chiuso per un'onda quadra di ampiezza unitaria e periodo 0.8 ms.

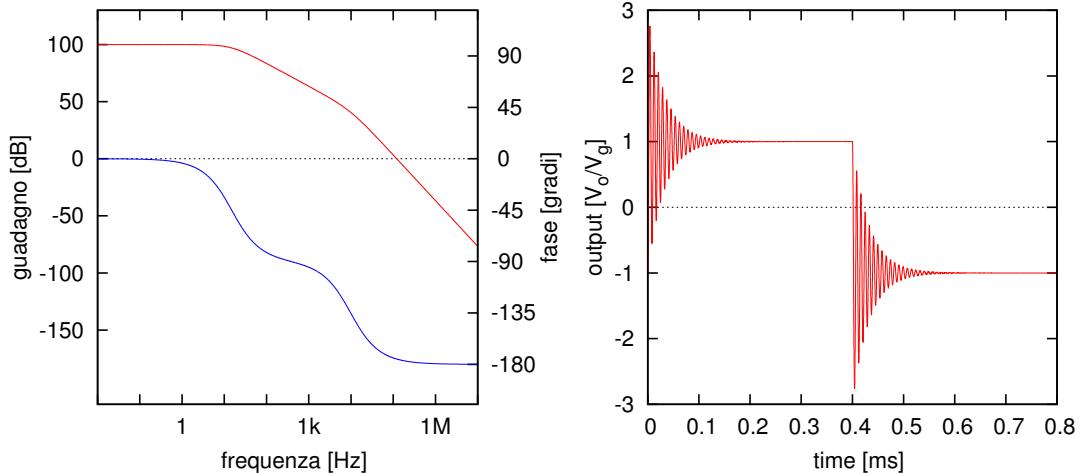


Figura 15.13: Comportamento del circuito di fig. 15.12. A sinistra: guadagno d'anello, modulo in rosso e fase in blu; a destra: risposta temporale ad anello chiuso ad un'onda quadra di ampiezza unitaria e periodo 0.8 ms. Il guadagno d'anello è 100 dB invece di 106 per via del partitore formato dalle due resistenze R , che introducono una attenuazione di un fattore 2.

Poichè in corrispondenza della seconda frequenza di taglio il guadagno è ancora circa 40 dB, occorre un'altra decade di frequenza per arrivare fino a 0 dB e questo dà spazio alla fase per ruotare fino quasi a 180°. Il margine di fase risulta infatti essere un misero 4.6° e la risposta in funzione del tempo al fronte di salita e di discesa dell'onda quadra dà origine ad una oscillazione che si smorza piuttosto lentamente nel tempo.

È evidente che una risposta di questo tipo è inaccettabile. È necessario modificare il guadagno d'anello in modo da portare il margine di fase ad un valore adeguato, ad esempio ai 66° o ai 76° corrispondenti alle curve rossa e verde di fig. 15.7. Questa procedura prende il nome di *compensazione*.

Una prima soluzione molto semplice è quella di ridurre il guadagno d'anello, ad esempio introducendo un partitore resistivo nel ramo di reazione β , aumentando quindi il guadagno ad anello chiuso. Questo ha l'effetto di traslare verso il basso la curva del modulo in fig. 15.13 lasciando inalterata la curva della fase e quindi di spostare verso sinistra verso frequenze più basse il punto di raggiungimento degli 0 dB, producendo un effettivo aumento del margine di fase.

In genere conviene ricorrere a questa soluzione solo quando non esiste realmente alcuna altra alternativa, perchè in questo modo si riduce la precisione del circuito anche alle frequenze basse, fino alla continua.

Le due tecniche di compensazione più utilizzate, dette di *polo dominante* e di *pole-zero*, prevedono di lasciare inalterato il valore asintotico del guadagno d'anello nel limite delle basse frequenze e della corrente continua, ma di arretrare la posizione della prima frequenza di taglio (detta *pole* dalla terminologia delle funzioni complesse)

in modo da raggiungere il valore di 0 dB in corrispondenza di minori rotazioni della fase.

Si suppone ovviamente che le caratteristiche della seconda frequenza di taglio siano legate al sistema fisico che si deve controllare e che non siano modificabili a piacere.

15.8.1 Polo dominante

La compensazione di polo dominante si ottiene modificando il circuito di fig. 15.12 secondo lo schema di fig. 15.14. La configurazione dell'amplificatore operazionale è

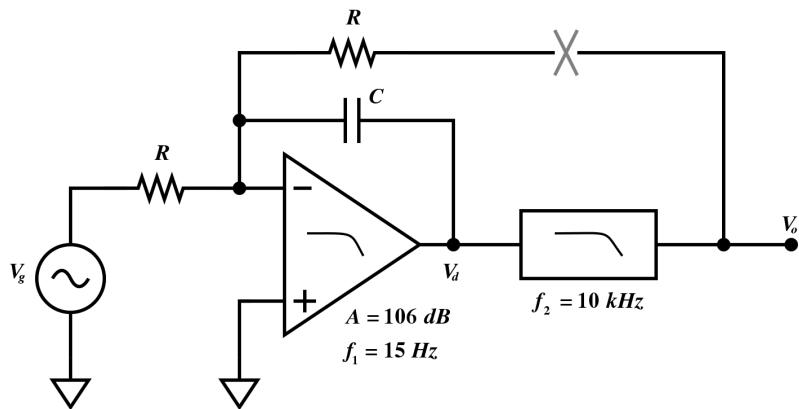


Figura 15.14: Circuito di compensazione a *polo-dominante*. Con i valori $R = 10 \text{ k}\Omega$ e $C = 1.6 \text{ nF}$ l'amplificatore ha guadagno unitario alla frequenza $f_0 = f(0 \text{ dB}) = 1/(2\pi RC) = 10 \text{ kHz}$ e frequenza di taglio $f'_1 = f_0/A = 100 \text{ mHz}$ ($A = 100 \text{ dB}$ a causa del partitore $R - R$ come in fig. 15.12 e 15.13).

quella dell'integratore di Miller, con la risposta in funzione della frequenza illustrata nel par. 12.6. Quello che si ottiene, con i valori di fig. 15.14, è l'arretramento della prima frequenza di taglio, dai 15 Hz caratteristici dell'operazionale fino a 100 mHz. A partire da questa frequenza il guadagno dell'amplificatore scende in modo regolare con pendenza di 6 dB per ottava e la fase ruota asintoticamente fino al valore di 90°.

Alla frequenza di 10 kHz, quando si incontra la seconda frequenza di taglio, il guadagno dell'amplificatore è 0 dB ed il guadagno d'anello -3 dB , con fase -135° . In corrispondenza del guadagno d'anello eguale a 0 dB la fase è a -128° e quindi il margine di fase è 52° , come nel terzo caso di fig. 15.7.

Le curve a sinistra in fig. 15.15 sono il guadagno d'anello della rete non compensata (in rosso) e con la compensazione di fig. 15.14 (in blu). A destra è riportata in blu la risposta all'onda quadra ad anello chiuso. È immediatamente evidente l'enorme differenza di comportamento rispetto al caso senza compensazione. Le altre due curve nei grafici sono state ottenute con una compensazione con costante di tempo RC doppia (violetto) e quadrupla (azzurro) rispetto al valore di fig. 15.14 e corrispondono al secondo e primo caso rispettivamente di fig. 15.7, con margine di fase di 66° e 76° . Come già visto, la curva azzurra corrisponde al caso di smorzamento critico, quella violetta alla risposta massimamente piatta in frequenza.

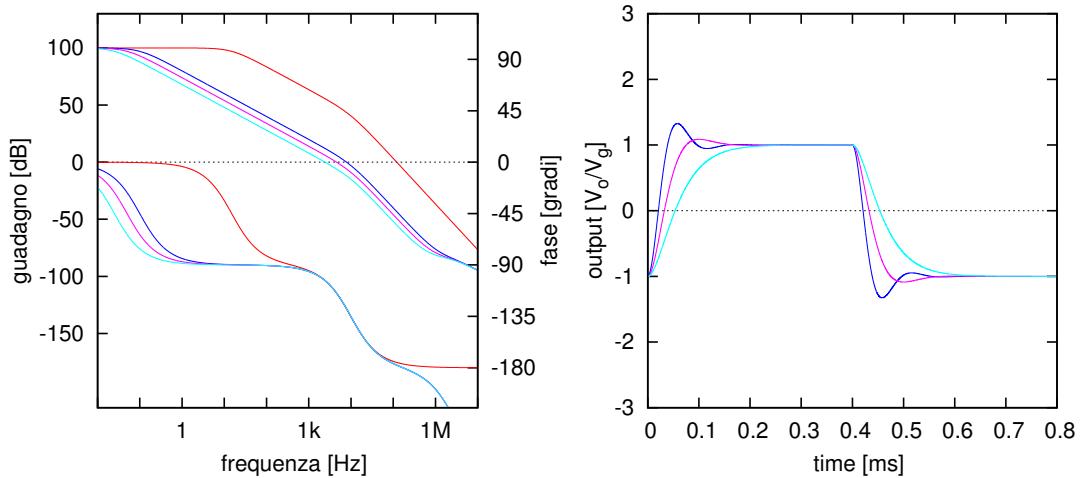


Figura 15.15: Compensazione di *polo dominante*. Il grafico a sinistra riporta il guadagno ad anello aperto senza compensazione (in rosso) e con la compensazione indicata in fig. 15.14 (in blu). A destra è riportata la risposta all'onda quadra della rete compensata ad anello chiuso (in blu). Le altre due curve corrispondono ad una compensazione con la costante RC di valore doppio (in violetto) e quadruplo (in azzurro) rispetto ai valori di fig. 15.14.

15.8.2 Polo-zero

Le risposte all'onda quadra che si vedono nella fig. 15.15 sono già molto buone rispetto a quello che si vedeva in fig. 15.13. Tuttavia si può fare di meglio utilizzando la compensazione di *polo-zero* come indicato nel circuito in fig. 15.16.

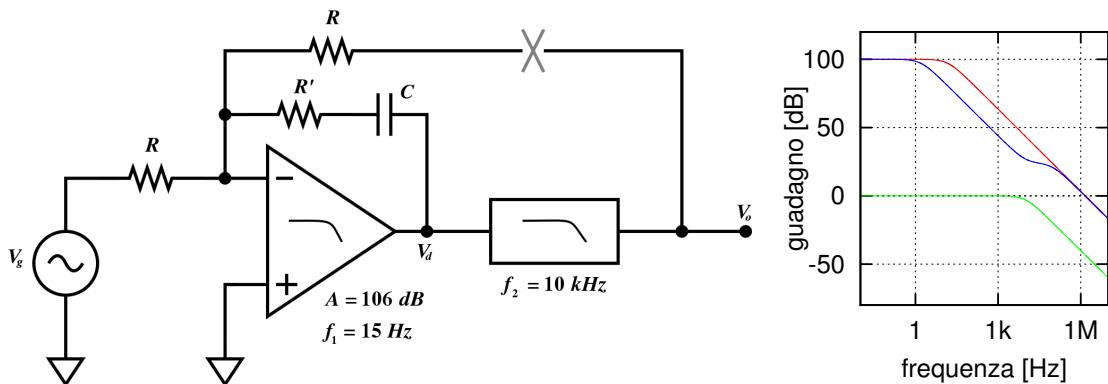


Figura 15.16: Circuito di compensazione di *polo-zero* e curve di risposta. Curva rossa: amplificatore operazionale senza compensazione; curva blu: amplificatore con compensazione di *polo-zero*; curva verde: circuito passa basso con frequenza di taglio di 10 kHz .

La differenza rispetto alla compensazione di polo dominante consiste nella introduzione della resistenza R' in serie al condensatore C dell'integratore di Miller.

La risposta del circuito amplificatore che ne risulta è quella indicata dalla linea blu nel grafico: dopo la prima frequenza di taglio, di tipo passa basso, determinata da R e da C , il guadagno scende con pendenza 6 dB per ottava, fino ad incontrare la seconda frequenza di taglio, di tipo passa-alto (uno *zero* nella funzione di trasferimento nella terminologia delle funzioni complesse), determinata da C e da R' .

A questa seconda frequenza il guadagno smette di diminuire e si stabilizza al valore R'/R , costante, mantenuto fino a che non viene raggiunto e superato (in discesa) dal guadagno dell'amplificatore non compensato nella sua diminuzione con l'aumentare della frequenza (curva rossa).

Se la costante di tempo $R'C$ è eguale a $1/(2\pi f_2)$ lo *zero* di $R'C$ ed il polo di f_2 (curva verde) si elideranno e l'andamento complessivo del guadagno d'anello sarà quello della curva verde nel diagramma di Bode in fig. 15.17.

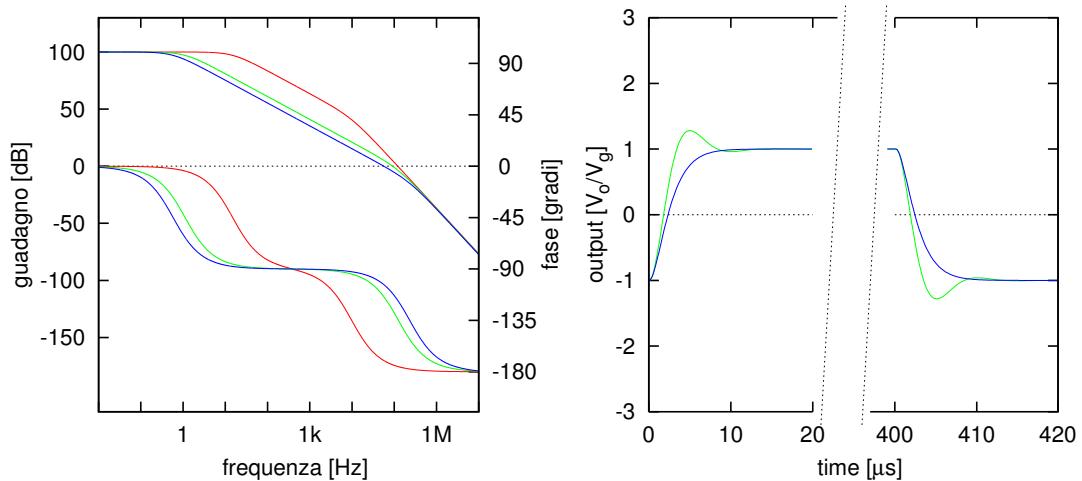


Figura 15.17: Compensazione di *polo-zero*. Curva rossa: guadagno ad anello aperto senza compensazione; curva verde: compensazione con $RC = 0.9 \mu\text{s}$ e $R'C = 16 \mu\text{s}$; curva blu: compensazione con $RC = 1.8 \mu\text{s}$ e $R'C = 16 \mu\text{s}$.

In pratica, si calcola la costante di tempo RC stabilendo che l'amplificatore compensato deve raggiungere il guadagno 0 dB , partendo dal valore iniziale di bassa frequenza e scendendo con pendenza costante di 6 dB per ottava, alla stessa frequenza a cui questo valore viene raggiunto dal guadagno d'anello senza compensazione, frequenza che nell'esempio di fig. 15.12 è 121 kHz :

$$RC = \frac{1}{2\pi f_{(0 \text{ dB})}} = 1.3 \mu\text{s} \quad (15.29)$$

Quindi si calcola R' dalla relazione

$$R'C = \frac{1}{2\pi f_2} = 16 \mu\text{s} \quad (15.30)$$

Queste condizioni corrispondono alle curve verdi nei grafici di fig. 15.17. La risposta all'onda quadra è nettamente più veloce che nel caso della compensazione di polo dominante, ma presenta un picco di *overshoot*. Riducendo di un fattore 2

la frequenza f_1 , cioè raddoppiando C (e dimezzando R' in modo da lasciare $R'C$ invariato) la risposta diventa quella delle curve blu, che corrisponde alla condizione di smorzamento critico.

La maggior velocità di risposta che si può ottenere con la compensazione *polo-zero* ha però un prezzo che non sempre è possibile pagare. Nella fig. 15.18 è riportata la tensione all'uscita dell'amplificatore che va a pilotare il sistema sotto controllo. Più è veloce la risposta complessiva dell'anello di reazione, più il segnale di pilotaggio diventa un picco stretto, veloce e alto. Al punto di poter facilmente arrivare a saturare i circuiti di uscita dell'amplificatore. E con la saturazione dell'amplificatore si introducono, oltre a non linearità che non sono sempre accettabili, anche ritardi che possono fare perdere tutti i vantaggi che si è pensato di poter ottenere.

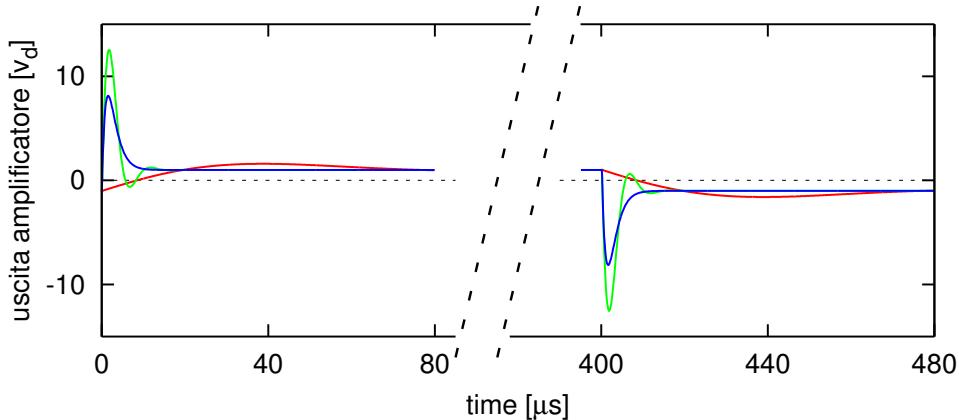


Figura 15.18: Ampiezza dei segnali all'uscita v_d dell'amplificatore per un segnale di ingresso v_g di ampiezza ± 1 V, onda quadra. Curva in rosso: compensazione di polo dominante di fig. 15.15; curve in verde ed in blu: compensazione di polo-zero di fig. 15.17.

15.9 Stabilità condizionata

Sono possibili circuiti in cui il modulo e/o la fase del guadagno d'anello non sono funzioni monotone della frequenza f . Un esempio è riportato in fig. 15.19.

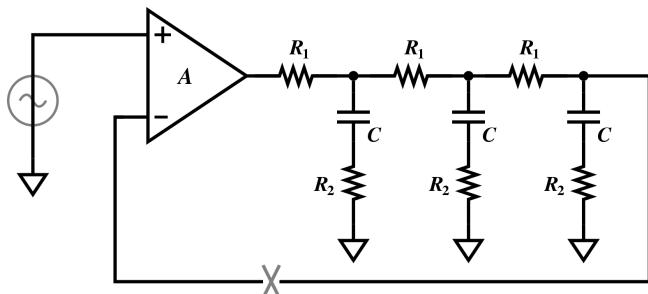


Figura 15.19: Esempio di circuito in cui la fase ha andamento non monotono in funzione della frequenza. $R_1 = 10 \text{ k}\Omega$, $R_2 = 400 \Omega$, $C = 160 \text{ nF}$, A = amplificatore con guadagno variabile.

Si tratta dello stesso circuito di fig. 15.9, con l'aggiunta di una resistenza in serie ad ogni condensatore. Questo introduce nel guadagno d'anello tre nuove frequenze di

taglio, tre *zeri*, dando origine ad una risposta complessiva in cui inizialmente il ritardo di fase aumenta con la frequenza per la presenza dei tre *poli* R_1C ed il guadagno diminuisce; successivamente i tre *zeri* arrestano la diminuzione del guadagno, ed il ritardo di fase diminuisce, tornando verso lo zero. Nel grafico 15.20, a sinistra, sono riportati i diagrammi di Bode per tre valori diversi di A : 25 (curva rossa), 250 (curva verde) e 2500 (curva blu). Il comportamento della fase (curva viola) non dipende dal valore di A ed è lo stesso nei tre casi.

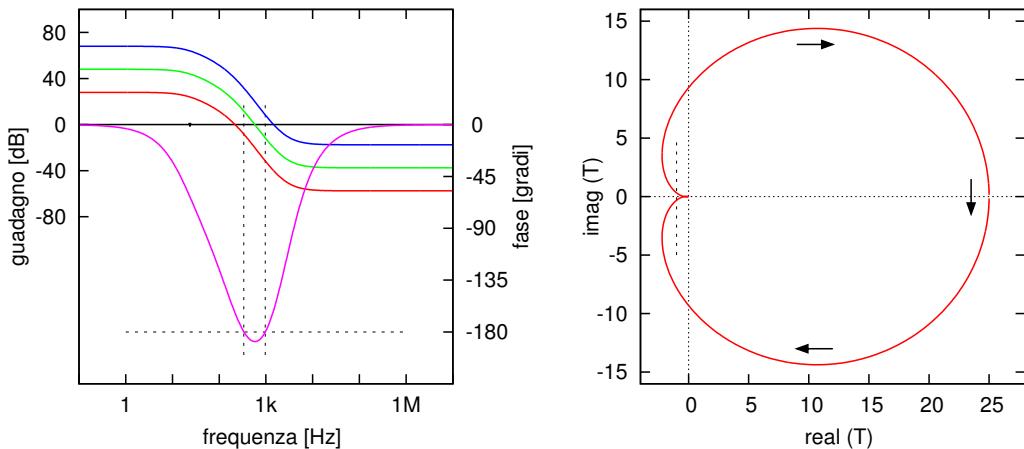


Figura 15.20: A sinistra: diagrammi di Bode del guadagno d'anello del circuito 15.19 per $A = 25$ (linea rossa), $A = 250$ (linea verde) e $A = 2500$ (linea blu). La fase (linea viola) è la stessa in tutti e tre i casi. A destra: diagramma di Nyquist per $A = 25$.

C'è un intervallo di frequenze, da circa 335 a 965 Hz, in cui la fase si trova oltre -180° . Come si vede dal grafico, a seconda del valore di A , la frequenza a cui il guadagno d'anello arriva al valore 0 dB si può trovare all'interno di questo intervallo (curva verde) oppure all'esterno a sinistra (curva rossa) o a destra (curva blu). Nel primo caso il punto $\{-1, 0\}$ del piano complesso si trova all'interno del tracciato di Nyquist e l'anello è instabile; negli altri due casi si trova all'esterno e l'anello è stabile.

Ci si trova nella situazione un poco controintuitiva, denominata da Nyquist *stabilità condizionata*, in cui riducendo il valore del guadagno d'anello A (in questo caso da maggiore a minore di circa 60 dB) si passa da un comportamento stabile ad uno instabile, per poi tornare scendendo sotto i 36 dB ad un comportamento nuovamente stabile.

Il diagramma di Nyquist a destra nel grafico 15.20 si riferisce al guadagno 25 (28 dB). Si vede, anche se con una certa difficoltà, che il punto $\{-1, 0\}$ si trova all'esterno del tracciato.

Negli altri due casi il diagramma di Nyquist è completamente illeggibile: quando il valore di A passa da 25 a 250 o a 2500, la forma del tracciato rimane la stessa, ma vengono moltiplicati per 10 o per 100 i valori sulle scale, rendendo impossibile individuare il punto $\{-1, 0\}$. Nella fig. 15.21 sono riportati due ingrandimenti della regione intorno all'origine, per rendere evidente il comportamento in questi due casi.

Normalmente avere un circuito in condizioni di *stabilità condizionata* è una situazione scomoda. Se il guadagno d'anello dovesse trovarsi a diminuire, il circuito

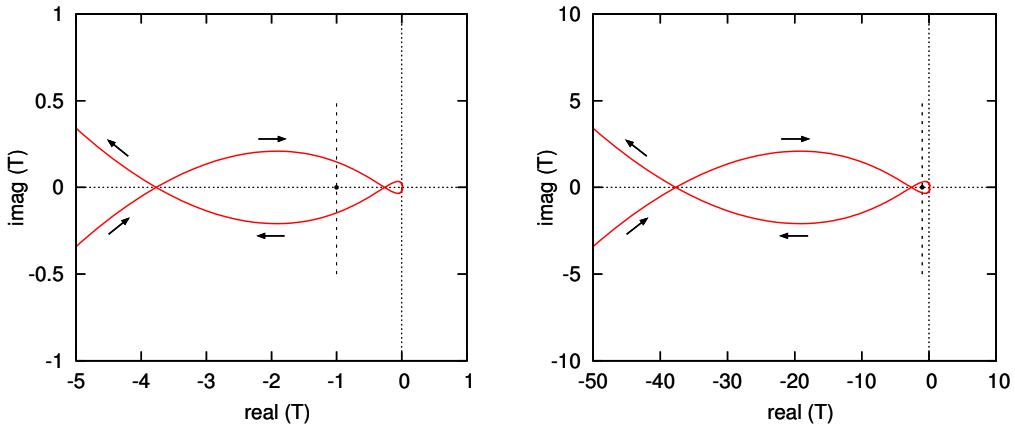


Figura 15.21: Tracciato di Nyquist per il guadagno d'anello del circuito 15.19 con $A = 250$ (a sinistra) ed $A = 2500$ (a destra). I due grafici sono uguali, cambiano solo le scale. Nel tracciato di destra, al diminuire del valore di A se si diminuisce in proporzione la scala si continua a vedere lo stesso disegno e si vede il punto $\{-1, 0\}$ spostarsi verso sinistra, fino ad uscire dal cappio più piccolo ed entrare in quello più grande. Nel percorrere il cappio piccolo nel grafico a destra il punto $\{-1, 0\}$ rimane sulla sinistra della linea e quindi è esterno al percorso; all'opposto, a sinistra, dove il punto $\{-1, 0\}$ si trova all'interno del cappio grande, rimane sulla destra della linea e quindi è all'interno del percorso.

diventerebbe instabile. Questo poteva succedere ad esempio con gli amplificatori che utilizzavano tubi a vuoto, il cui guadagno diminuiva con l'invecchiamento.

Nei moderni amplificatori a semiconduttori questo non si verifica. Ma in una situazione di transitorio un amplificatore può trovarsi ad avere guadagno ridotto, ad esempio per effetto della saturazione. Diventerebbe quindi instabile, entrando in oscillazione e facendo diventare permanente la condizione transitoria di saturazione.

Riportato al corretto funzionamento, ad esempio attraverso un ciclo di spegnimento-riaccensione, tornerebbe ad essere stabile. In attesa del transitorio successivo.

Capitolo 16

I transistor in alta frequenza

I modelli lineari per i dispositivi a due porte descritti al par. 6.4 sono astrazioni matematiche, analoghe ai teoremi di Thèvenin e Norton. Questi modelli sono stati utilizzati nei capitoli cap. 7–9 per descrivere il comportamento dei transistor *bjt* e *fet* nel regime dei piccoli segnali, scegliendo tra tutti quelli teoricamente possibili i più aderenti ai meccanismi fisici di funzionamento dei dispositivi (fig. 7.8 e 9.17). In questi modelli non è contenuta nessuna dipendenza esplicita del comportamento del transistor dalla frequenza dei segnali.

D'altra parte, come visto nel paragrafo 4.18, ad una giunzione *PN* direttamente polarizzata è associata una capacità di diffusione C_D e ad una giunzione inversamente polarizzata una capacità di transizione C_T . Nel *bjt* funzionante nella regione attiva queste due capacità sono entrambe presenti, la prima tra base ed emettitore in parallelo alla conduttanza di ingresso g_{ie} , essendo la giunzione base-emettitore direttamente polarizzata, la seconda in parallelo alla giunzione (inversamente polarizzata) tra base e collettore. Analogamente, nel caso dei transistor *fet* si avranno capacità tra *gate*, *source* e *drain*. A causa di queste capacità il comportamento dei transistor viene a dipendere dalla frequenza.

Volendo rimanere aderenti allo schema dei modelli lineari di fig. 7.8 e 9.17, la dipendenza dalla frequenza può essere introdotta considerando i quattro parametri g_{ie} , g_{fe} , g_{re} , g_{oe} (e analogamente h_{ie} , h_{fe} , h_{re} , h_{oe} ecc.) non più costanti reali, ma funzioni complesse della frequenza: $\hat{g}_{ie}(\omega)$, $\hat{g}_{fe}(\omega)$

In genere all'aumentare della frequenza risulterà anche che non saranno più trascurabili i parametri che potevano essere ignorati in prima approssimazione in corrente continua ed in bassa frequenza (g_{re} , h_{re} ...).

In alternativa, si può lasciare inalterato il modello per piccoli segnali valido per la corrente continua ed aggiungere esplicitamente al circuito altri elementi, corrispondenti alle capacità delle giunzioni. Il circuito diventa più complesso, ma ha il vantaggio di continuare ad essere composto solo da elementi il cui valore è indipendente dalla frequenza e che hanno una relazione diretta con gli aspetti fisici del dispositivo. Nei paragrafi seguenti verrà utilizzata questa seconda soluzione.

16.1 Transistor *bjt*

Nel caso del transistor *bjt*, aggiungendo al circuito la capacità di diffusione C_D della giunzione base-emettitore direttamente polarizzata e la capacità C_T della giunzione base-collettore inversamente polarizzata, si ottiene il circuito di fig. 16.1. Nei transistor che fanno parte di circuiti monolitici la presenza di un substrato comune a più dispositivi può dar luogo ad una capacità significativa tra collettore e substrato e tra emettitore e substrato (capacità C_{cs} e C_{es} a tratteggio nella figura). Se substrato ed emettitore sono collegati tra di loro, come succede comunemente nei circuiti con emettitore a massa, il condensatore C_{es} risulta cortocircuitato e rimane il solo C_{cs} tra collettore ed emettitore/massa.

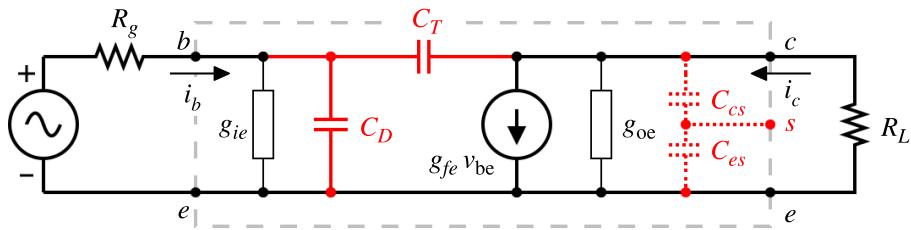


Figura 16.1: Modello lineare del *bjt* nella regione attiva completo della capacità di diffusione C_D della giunzione base-emettitore e della capacità di transizione C_T della giunzione base-collettore.

16.2 Amplificazione di corrente

Il guadagno di corrente del transistor, indipendentemente dal circuito in cui viene impiegato, si può studiare ponendo $R_L = 0$, cioè cortocircuitando collettore ed emettitore. La capacità C_T di solito è molto minore di C_D ed in queste condizioni dà un contributo trascurabile; il guadagno di tensione A_v si riduce a zero e il guadagno di corrente diventa

$$A_i = \frac{I_c}{I_b} = \frac{g_{fe}}{g_{ie} + j\omega C_D} = \frac{g_{fe}/g_{ie}}{1 + j\omega C_D/g_{ie}} \quad (16.1)$$

In corrente continua ed in bassissima frequenza il guadagno in corrente è costante e pari a $g_{fe}/g_{ie} = h_{fe}$ (eq. 6.11); al crescere della frequenza si ha un andamento di tipo *passa basso* (fig. 16.2), con frequenza di taglio

$$f_\beta = \frac{g_{ie}}{2\pi C_D} \quad (16.2)$$

Per frequenze molto maggiori di f_β il termine 1 a denominatore della 16.1 diventa trascurabile e si può scrivere

$$|A_i| = \frac{g_{fe}}{\omega C_D} \quad (16.3)$$

La frequenza f_T a cui si ha $|A_i| = 1$ prende il nome di *frequenza di transizione*:

$$f_T = \frac{g_{fe}}{2\pi C_D} \quad (16.4)$$

Il rapporto fra f_T e f_β è:

$$\frac{f_T}{f_\beta} = \frac{g_{fe}/2\pi C_D}{g_{ie}/2\pi C_D} = \frac{g_{fe}}{g_{ie}} = h_{fe} \quad (16.5)$$

La frequenza f_T ed il coefficiente di amplificazione di corrente h_{fe} sono due dei parametri fondamentali che caratterizzano un transistor bjt e sono normalmente specificati dal costruttore.

Invece della frequenza f_T a volte si trova specificato il *tempo di transizione* (*transition time*) $t_T = 1/(2\pi f_T)$.

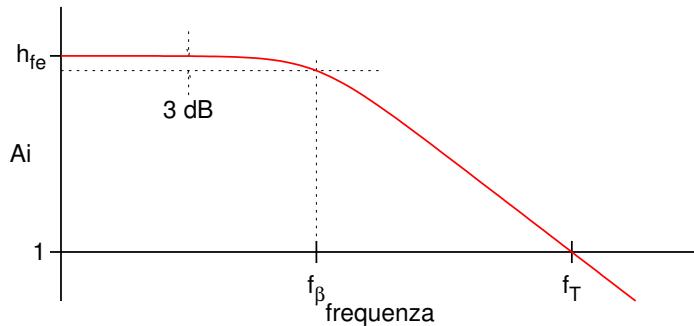


Figura 16.2: Amplificazione di corrente in funzione della frequenza per un transistor bjt. Entrambe le scale sono logaritmiche.

16.3 Amplificazione di tensione - Effetto Miller

Il calcolo completo di amplificazione di tensione e corrente del circuito di fig. 16.1, considerando anche la resistenza di carico R_L , deve essere fatto scrivendo e risolvendo il sistema di equazioni lineari complesse per le tre maglie di cui il circuito è composto.

Un calcolo approssimato, ma sufficiente per identificare il ruolo dei vari elementi e valutare costanti di tempo e frequenze di taglio, può essere fatto modificando il circuito come in fig. 16.3.

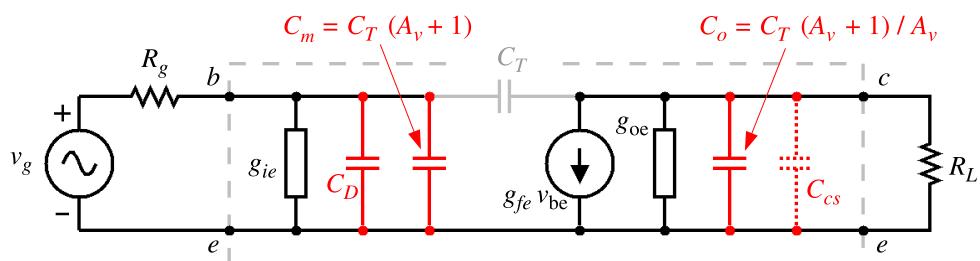


Figura 16.3: Modello del bjt nella regione attiva con la capacità di transizione C_T della giunzione base-collettore sostituita dalle due capacità equivalenti C_m e C_o secondo l'effetto Miller.

La capacità C_T si trova tra l'uscita e l'ingresso dell'amplificatore, che ha guadagno di tensione $-A_V$. Nell'ipotesi che l'uscita dell'amplificatore si comporti come un buon generatore di tensione, cioè abbia impedenza equivalente di uscita piccola rispetto alla reattanza del condensatore C_T , il circuito si comporta come se fosse presente una capacità $C_m = C_T \cdot (1 + A_V)$ tra l'ingresso dell'amplificatore e massa ed una capacità $C_o = C_T \cdot (1 + A_V) / A_V$ tra l'uscita e massa (effetto Miller, fig. 16.4 - vedi par. 11.10.3). La capacità C_m all'ingresso si viene a trovare in parallelo e quindi si somma a C_D e

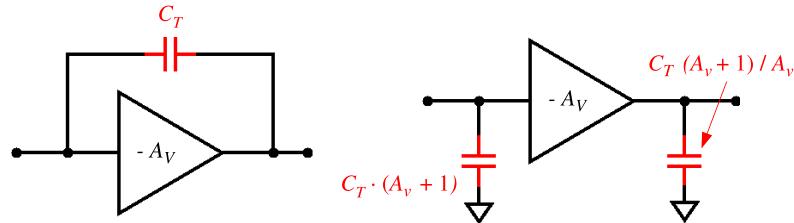
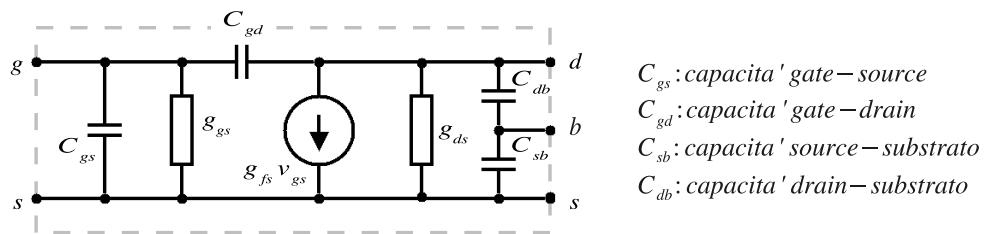


Figura 16.4: Effetto Miller.

contribuisce a ridurre la frequenza di taglio superiore dell'amplificatore. Per valori di A_V elevati, come è tipico per i *bjt* con emettitore a massa ($\sim 10^2$), la capacità *Miller* C_m spesso è l'elemento determinante della risposta in frequenza.

16.4 Transistor fet

Il *gate* di un transistor *mosfet* è a tutti gli effetti l'armatura di un condensatore la cui capacità può andare dalla frazione di pF per un dispositivo per piccoli segnali ed alta frequenza alle centinaia o migliaia di pF per i dispositivi di potenza. Il circuito lineare di fig. 9.17 dovrà essere modificato, come già è stato fatto per il *bjt*, per tenere conto di questi effetti, tanto più importanti quanto più alta è la frequenza di lavoro. Le due capacità C_{gs} e C_{gd} in fig. 16.5 rappresentano l'accoppiamento tra

Figura 16.5: Circuito lineare equivalente per i transistor *fet* in alta frequenza.

il *gate* e gli altri due elettrodi; le capacità C_{db} e C_{sb} corrispondono alle capacità di transizione delle giunzioni contropolarizzate tra *drain* e *source* e *substrato*. Nel caso che il *substrato* sia collegato al *drain*, la capacità C_{sb} risulta cortocircuitata e quindi ininfluente sul funzionamento del circuito.

Il circuito equivalente per il *jfet* in alta frequenza coincide con quello del *mosfet*. Le capacità C_{gs} e C_{gd} corrispondono alla capacità di transizione della giunzione *gate-canale*, rispettivamente dalla parte del *source* e del *drain*.

In questo schema è stato inserito anche il parametro di ingresso g_{gs} , corrispondente nel caso del *jfet* alla conduttanza dinamica della giunzione contropolarizzata del *gate*, che risulta comunque trascurabile, eccetto che nelle applicazioni dove le impedenze in gioco sono estremamente elevate.

Anche nel caso dei transistor *fet*, a causa dell'effetto Miller, la capacità C_{gd} diventa spesso l'elemento predominante nel determinare il comportamento in alta frequenza in presenza di amplificazioni di tensione elevate.

16.5 Amplificatore a larga banda

Come esempio, consideriamo lo stadio amplificatore di fig. 16.6, corrispondente al funzionamento in modo differenziale di ciascuno dei due transistor del circuito di fig. 7.21. Utilizzando transistor del *chip* monolitico *CA3046*, i parametri hanno

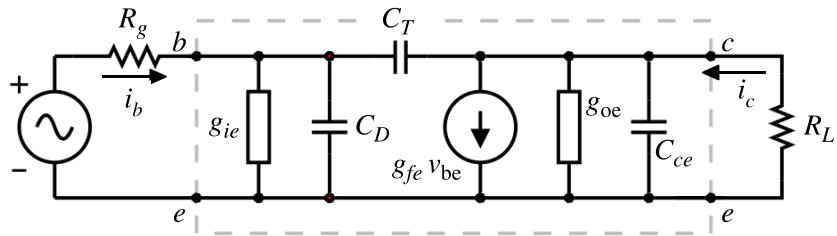


Figura 16.6: Modello lineare di un amplificatore a larga banda con transistor *CA3046*. Condizioni operative:

I_E	0.72 mA	V_{CE}	8.5 V	R_L	10 k Ω	R_g	50 Ω
g_{fe}	27 mA/V	g_{ie}	0.24 mA/V	g_{oe}	6.6 μ A/V	A_I	106
C_D	9.1 pF	C_T	0.44 pF	C_{ss}	0.5 pF	A_V	250

i valori riportati nella didascalia. Nelle fig. 16.7 sono riportati, in funzione della frequenza, i guadagni di corrente e di tensione, in modulo e fase, ed il guadagno di potenza, calcolati risolvendo numericamente le equazioni complete del circuito mediante il programma *Spice*. Nella fig. 16.8 è riportata l'impedenza di ingresso, in modulo e fase.

Dal calcolo, il guadagno di corrente in bassa frequenza risulta essere $A_i = 106$ (40.5 dB) e quello di tensione $A_v = 251$ (48 dB), in ottimo accordo con i valori previsti dalle equazioni del cap. 7 in cui si è considerato $h_{re} = 0$. La frequenza di taglio per il guadagno di corrente A_i in funzione della frequenza si trova a $f_{A_i} = 0.31$ MHz e quella per il guadagno di tensione a $f_{A_v} = 10.5$ MHz. Dall'eq. 16.2 la frequenza di taglio per A_i risulterebbe essere $f_\beta = 4.2$ MHz. Questa vistosa differenza è frutto dell'effetto Miller. Considerando, secondo quanto visto nel par. 16.3 una capacità di ingresso effettiva

$$C_i = C_D + C_T \cdot (1 + A_V) = 9.1 + 0.44 \cdot 252 = 120 \text{ pF} \quad (16.6)$$

si ottiene il valore $f_\beta = 0.32$ MHz, in ottimo accordo con il valore ottenuto con il calcolo completo.

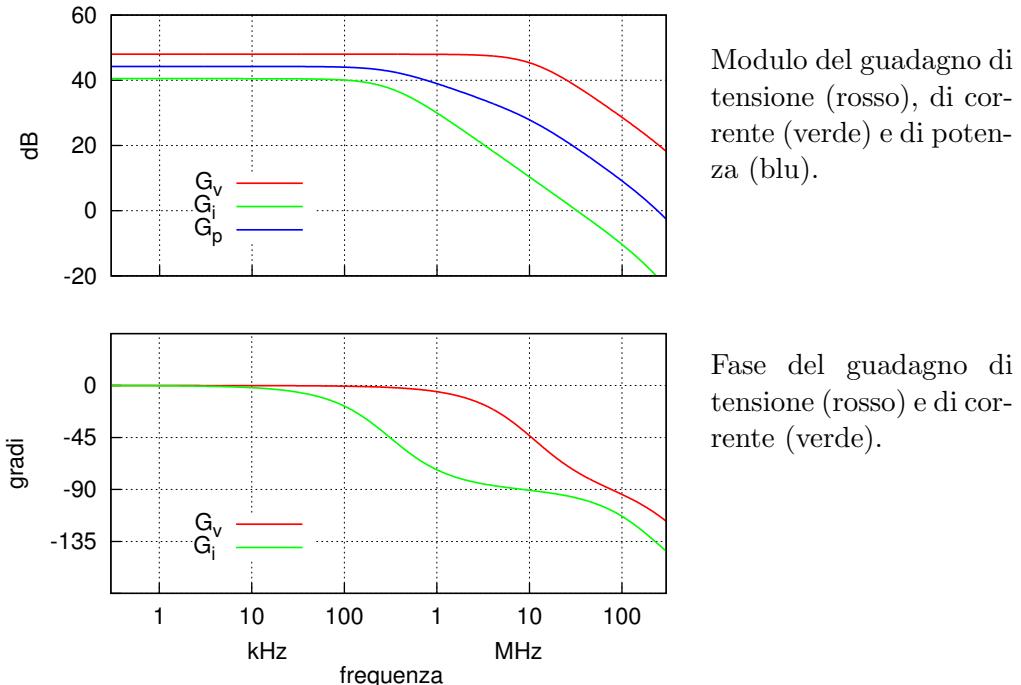


Figura 16.7: Guadagno di tensione, di corrente e di potenza dell'amplificatore in fig. 16.6

Nonostante la capacità di transizione C_T sia molto più piccola della capacità di diffusione C_D (0.44 contro 9.1 μF), a causa dell'effetto Miller il suo contributo alla capacità totale di ingresso (120 μF) è largamente predominante.

La frequenza di taglio per il guadagno di tensione A_v risulta di oltre un ordine di grandezza superiore a quella per A_i . Questo è dovuto alle condizioni di funzionamento del circuito, in cui si è ipotizzato di utilizzare un generatore con resistenza interna piccola ($R_g = 50 \Omega$). Infatti, mentre la frequenza di taglio per l'amplificazione di corrente A_i è determinata dalla costante di tempo di C_i e g_{ie} in parallelo, nel calcolo della frequenza di taglio di A_v bisogna considerare la costante di tempo di C_i , g_{ie} ed anche R_g in parallelo.

Questo si vede chiaramente nel grafico in fig. 16.8 che riporta l'impedenza di ingresso Z_i dell'amplificatore, come viene *vista* dal generatore di segnale, in funzione della frequenza. In corrente continua ed in bassa frequenza si ha $Z_i = 4.2 k\Omega$; alla frequenza di taglio di 320 kHz il modulo di Z_i si è ridotto di un fattore $\sqrt{2}$ e la fase è arrivata a 45°. Questa frequenza di taglio coincide con quella del guadagno di corrente A_i . A frequenze più alte di questa, la corrente di ingresso aumenta rapidamente, mentre la tensione rimane sostanzialmente costante grazie al fatto che la resistenza R_g del generatore (50 Ω) è molto minore di $|Z_i|$: il generatore sta pilotando la base del transistor in tensione, erogando una corrente sempre maggiore, proporzionalmente al crescere della frequenza.

Andando verso frequenze ancora più alte, si incontra una seconda frequenza di taglio, determinata dalla costante di tempo τ_o del circuito di uscita formato dalla

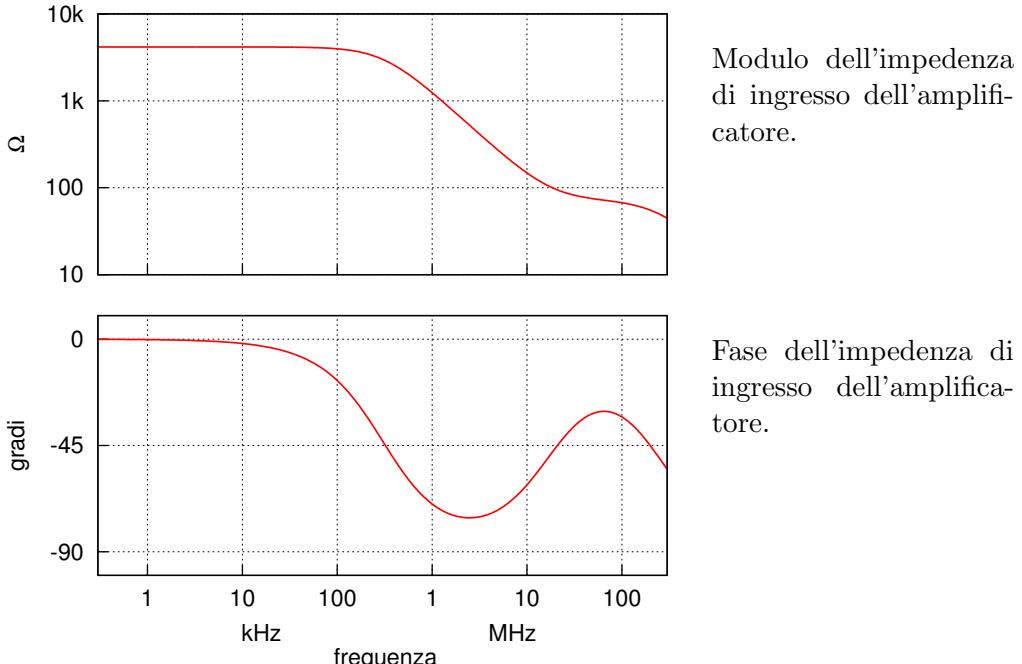


Figura 16.8: Impedenza di ingresso dell'amplificatore descritto in fig. 16.6.

combinazione in parallelo di R_L , g_{oe} , C_o e C_{cs} ($\tau_o \simeq 8.8 \text{ ns}$). Questa seconda frequenza di taglio si trova a circa 18 MHz . Contemporaneamente, però, il modulo di Z_i , diminuendo costantemente all'aumentare della frequenza, viene ad avere valori confrontabili con R_g . Questi due effetti combinati insieme determinano la frequenza di taglio $f_{Av} = 10.5 \text{ MHz}$ dell'amplificazione di tensione.

16.6 Amplificatori a banda stretta

Non sempre si ha interesse a realizzare amplificatori con banda passante la più ampia possibile. In molti casi, specialmente nei circuiti a radiofrequenza, si ha l'interesse opposto: amplificare i segnali contenuti all'interno di una determinata banda di frequenza, sopprimendo il più possibile quelli che si trovano al di fuori, che costituiscono *rumore* o *interferenza*.

Nel circuito in fig. 16.9a questo scopo viene raggiunto aggiungendo in parallelo alla resistenza di carico R un circuito risonante parallelo LC . L'impedenza Z risultante è funzione della frequenza e presenta il suo massimo R (in modulo) in corrispondenza della frequenza di risonanza $f_0 = 1/(2\pi\sqrt{LC})$, per poi diminuire quando ci si allontana dalla risonanza verso frequenze più alte o più basse:

$$Z = \frac{R}{1 + jQ_0 \left(\frac{f}{f_0} - \frac{f_0}{f} \right)} \quad (16.7)$$

dove $Q_0 = R\sqrt{C/L}$ è il *fattore di merito* del circuito alla frequenza di risonanza (eq. 1.42).

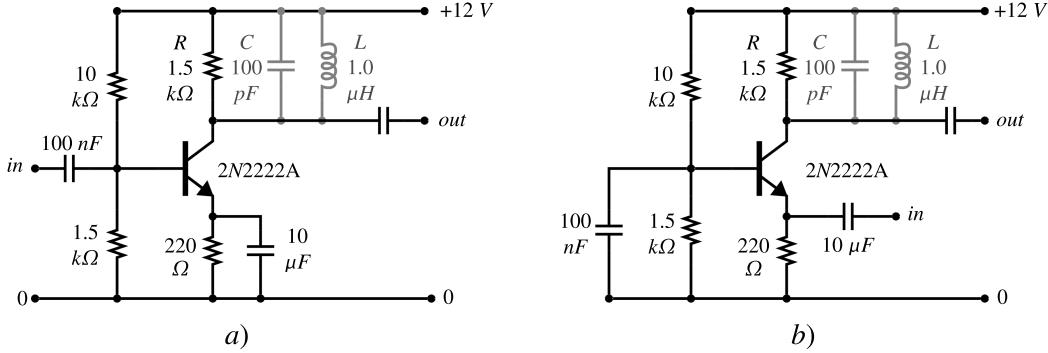


Figura 16.9: Amplificatore con transistor *bjt* con emettitore a massa (a sinistra) e base a massa (a destra). Quando il carico è formato dalla sola resistenza R si ha un amplificatore a *banda larga*; aggiungendo in parallelo ad R anche il circuito risonante LC si ottiene un amplificatore a *banda stretta*.

In prima approssimazione l'amplificazione di tensione è $A_v = V_{out}/V_{in} = g_{fe}Z$ e quindi segue l'andamento dell'impedenza di carico, come riportato in fig. 16.10.

La curva rossa corrisponde alla presenza della sola resistenza R e mostra l'andamento già visto nel par. 16.5: l'amplificazione rimane costante fino alla frequenza di circa 29 MHz , determinata dalla costante di tempo di resistenze e capacità presenti sul collettore del transistor, quindi comincia a diminuire con un andamento di tipo *passa-basso*.¹

La curva blu è stata ottenuta invece considerando la presenza del circuito risonante LC in parallelo alla resistenza R ; l'amplificazione di tensione quindi segue l'andamento dell'impedenza di carico Z data dalla eq. 16.7. Si ha un picco a $f_0 = 15.7\text{ MHz}$ ed una *larghezza di banda* $BW \simeq 1.1\text{ MHz}$ (a -3 dB), corrispondente ad un fattore di merito $Q_0 \simeq 14$ (con i valori dei componenti riportati nello schema si dovrebbe avere $f_0 = 15.9\text{ MHz}$ e $Q_0 = 14.3$; il calcolo effettuato con il programma *Spice* tiene conto anche della capacità $C_T = 2.9\text{ pF}$, interna al transistor; la capacità effettiva del circuito risonante passa quindi da 100 a quasi 103 pF).

In fig. 16.9b è riportato lo stesso circuito di fig. 16.9a, nelle stesse condizioni operative, ma in configurazione *base a massa* invece che *emettitore a massa*. Per quanto riguarda l'amplificazione di tensione il comportamento dei due circuiti è praticamente identico; i grafici relativi al circuito con base a massa sono riportati in colore verde e violetto e non sono distinguibili dagli altri.

Il comportamento dei due circuiti, emettitore a massa o base a massa, differisce invece profondamente per quanto riguarda l'impedenza di ingresso in funzione della frequenza, riportata in fig. 16.11 per entrambi i circuiti ed entrambe le condizioni operative, banda larga e stretta.

¹Porre $A_v = V_{out}/V_{in}$ equivale ad ignorare la resistenza interna del generatore di segnale oppure supporre un generatore ideale di tensione, e quindi ad eliminare gli effetti della costante di tempo all'ingresso, effetti che si ritrovano poi nell'amplificazione di corrente, come visto nel par. 16.5.

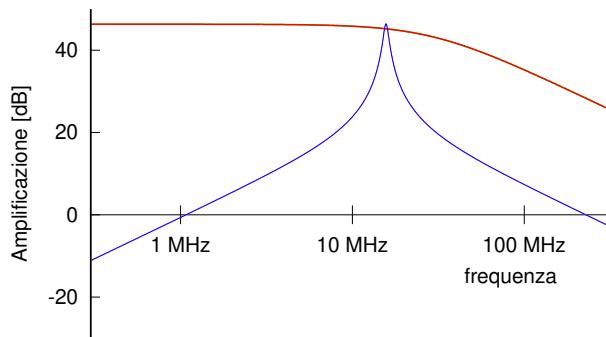
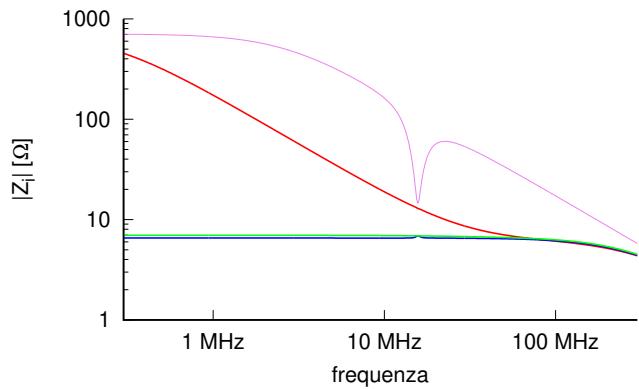


Figura 16.11: Modulo dell'impedenza di ingresso dei due circuiti di fig. 16.9 per le due condizioni operative. Rosso: emettitore comune, banda larga; violetto: emettitore comune, banda stretta; verde: base comune, banda larga; blu: base comune, banda stretta.



Nel circuito con base a massa l'impedenza di ingresso è di valore molto basso (circa 7Ω) e si mantiene molto stabile fino ad oltre 100 MHz , quando incomincia a decrescere (curve blu e verde). Nel circuito con emettitore a massa, invece, si osserva il comportamento già visto nel par. 16.5: in bassa frequenza il valore dell'impedenza di ingresso è relativamente alto (circa 700Ω), ma comincia presto a diminuire al crescere della frequenza, più precocemente nel circuito a larga banda (rosso), a frequenze leggermente più alte nell'altro caso (violetto) in quanto la ridotta amplificazione di tensione fuori della risonanza limita il contributo dell'effetto Miller. Nell'intorno della frequenza di risonanza, che è l'intervallo di frequenze di maggior interesse nel caso del circuito a banda stretta, a causa dell'effetto Miller l'impedenza di ingresso ad emettitore comune presenta un brusco minimo,

La cosa peggiore però la si vede nel grafico di fig. 16.12, dove vengono riportate le fasi per ciascuna delle impedenze di ingresso di fig. 16.11.

In entrambi i circuiti con base a massa la fase rimane a zero fino a circa 10 MHz , poi comincia a scendere lentamente verso i -45° , che verranno raggiunti alla frequenza di transizione $f_T = 350 \text{ MHz}$, per poi proseguire verso i -90° a frequenze ancora più alte (curve verde e blu).

Nei circuiti con emettitore a massa già a frequenze minori di 1 MHz la fase ha una forte componente capacitiva; nel circuito a larga banda (curva rossa) questa si riduce leggermente al crescere della frequenza, tra i 3 ed i 100 MHz ; nel circuito a banda stretta la fase continua a crescere, supera i -90° ed arriva fino a quasi

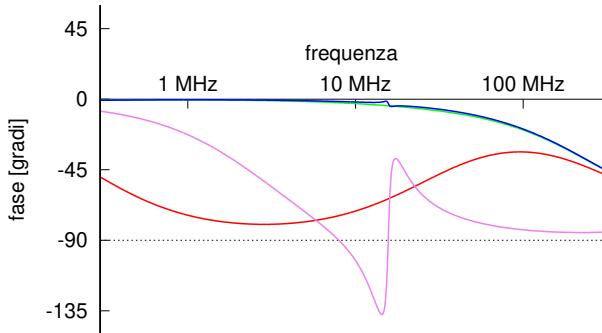


Figura 16.12: Fasi delle impedenze di ingresso riportate in fig. 16.11. I colori delle curve sono gli stessi di fig. 16.11.

-140° . Questo corrisponde nel piano complesso ad una impedenza che possiede una componente reale, cioè resistiva, negativa. In determinate condizioni, cioè in presenza di una componente induttiva nel circuito di ingresso, si può arrivare ad avere un circuito risonante con smorzamento negativo, l'amplificatore può diventare instabile e da amplificatore trasformarsi in generatore di segnali propri. Questo comportamento è dovuto al forte accoppiamento presente tra uscita ed ingresso a causa della capacità di transizione C_T . Problemi simili si hanno utilizzando altri dispositivi amplificatori, come i transistor *mosfet* o *jfet* nella configurazione *common source*².

Questo è uno dei motivi principali per cui quando si vuole far lavorare un transistor a frequenze alte, e soprattutto quando nel circuito sono presenti elementi reattivi (condensatori, induttanze), è conveniente utilizzare le configurazioni *common base* per i *bjt* e *common gate* per i *fet*.

16.7 Adattamento di impedenza

Nell'uso dell'amplificatore a banda stretta del paragrafo precedente normalmente la resistenza R non è fisicamente presente, ma rappresenta la resistenza di ingresso del circuito utilizzatore del segnale (ed in parte anche le perdite proprie del circuito risonante LC), come disegnato in fig. 16.13a.

Se però il collegamento deve essere fatto ad un carico che presenta un valore di impedenza diverso dai 1500Ω previsti, ad esempio ad un apparecchio che presenta il valore di impedenza di ingresso normalizzato standard di 50Ω , è necessario provvedere ad un *adattamento di impedenza*, modificando il circuito come nell'esempio in fig. 16.13b.

Una analisi semplice ed intuitiva, anche se non completamente rigorosa, del meccanismo della *trasformazione di impedenza* tramite i circuiti risonanti si può fare partendo dal circuito di fig. 16.14.

Il circuito ha una sua frequenza di risonanza $f_0 = 1/(2\pi\sqrt{LC})$, con la capacità $C = C_1C_2/(C_1 + C_2)$ risultante dalla serie di C_1 e C_2 . Poiché C_1 e C_2 formano un partitore capacitivo, si ha anche $V'/V = C_1/(C_1 + C_2)$.

Inserendo la resistenza R' in parallelo a C_2 tutte le caratteristiche del circuito vengono alterate. Se però si suppone che valga la condizione $R \gg 1/(\omega C_2)$, allora le

²Oppure i triodi dei tempi antichi nei circuiti con *catodo a massa*.

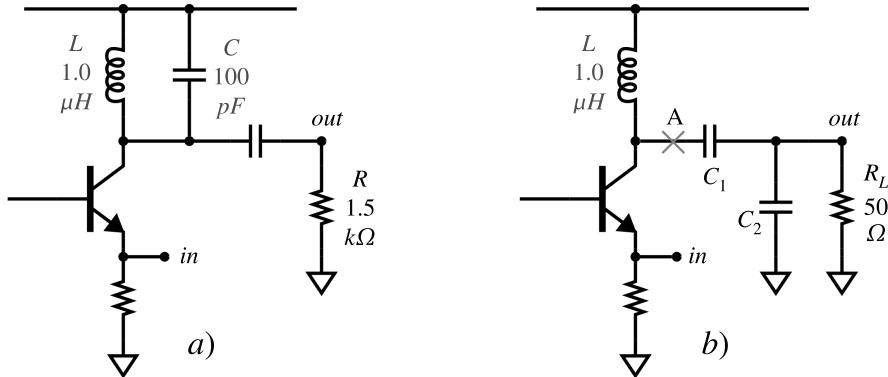


Figura 16.13: Adattamento di impedenza. Il circuito di sinistra è del tutto equivalente all'amplificatore a banda stretta di fig. 16.9. Il carico esterno R ($1.5 k\Omega$) è collegato tramite un condensatore che isola il circuito per quanto riguarda le polarizzazioni ma di valore tale da essere un cortocircuito per quanto riguarda i segnali. Nel circuito a destra il carico esterno ($R_L = 50 \Omega$) ha un valore incompatibile con l'uscita dell'amplificatore e deve essere *adattato* tramite il trasformatore di impedenza formato dal partitore capacitivo $C_1 - C_2$ in parallelo all'induttanza L .

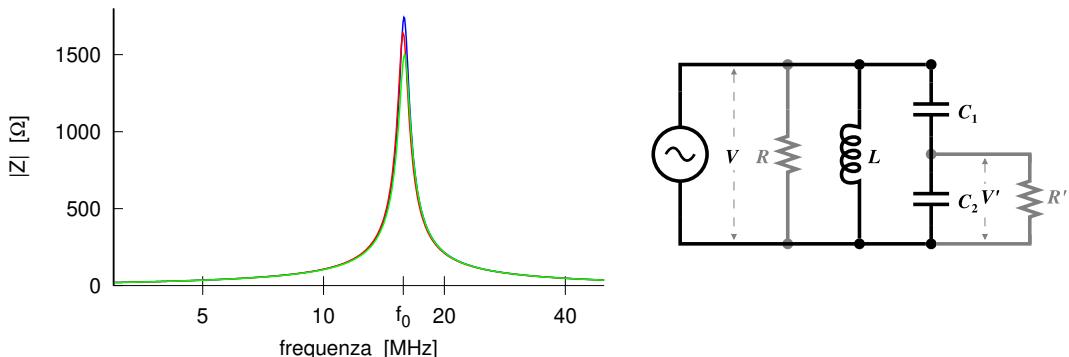


Figura 16.14: Adattamento di impedenza tramite circuito risonante. La resistenza R' collegata all'uscita del partitore capacitivo viene *vista* dal generatore come una resistenza $R \simeq R' \cdot (1 + C_2/C_1)^2$ in parallelo al circuito risonante. Il grafico riporta in verde la curva di risonanza senza R' e con $R = 1.5 k\Omega$ ed in rosso la curva senza R e con $R' = 50 \Omega$ ($C_1 = 122 pF$; $C_2 = 547 pF$; $L = 1 \mu H$). La curva in blu è quella che si ottiene arrotondando i valori dei componenti ottenuti dal calcolo a quelli effettivamente disponibili e cioè $120 pF$ per C_1 e $560 pF$ per C_2 . La frequenza di risonanza è ancora $15.8 MHz$, ma l'altezza del picco è diventata $1.74 k\Omega$.

alterazioni saranno piccole: si avrà un certo spostamento della frequenza di risonanza e la divisione di tensione del partitore $C_1 - C_2$ continuerà a valere anche se solo in forma approssimata.

Alla risonanza il circuito si presenterà al generatore come una resistenza reale pura R e quindi assorberà dal generatore una potenza V^2/R . Questa potenza sarà quella ceduta alla resistenza R' , poiché nel circuito non sono presenti altri elementi dissipativi (si stanno naturalmente considerando induttanza e condensatori come

componenti ideali). Si avrà quindi

$$\frac{V^2}{R} = \frac{V'^2}{R'} \simeq \frac{V^2}{R'} \left(\frac{C_1}{C_1 + C_2} \right)^2 \quad (16.8)$$

da cui

$$\frac{R}{R'} = \left(\frac{C_1 + C_2}{C_1} \right)^2 \quad \frac{C_2}{C_1} = \sqrt{\frac{R}{R'}} - 1 \quad (16.9)$$

Con i valori dell'esempio di fig. 16.13 si dovrà avere:

$$\frac{C_1 C_2}{C_1 + C_2} = 100 \text{ pF} \quad \frac{C_2}{C_1} \simeq \sqrt{\frac{1500}{50}} - 1 = 4.48 \quad (16.10)$$

da cui

$$C_1 = 122 \text{ pF} \quad C_2 = 547 \text{ pF} \quad (16.11)$$

Il risultato si vede nel grafico di fig. 16.14. La frequenza di risonanza è di 15.8 MHz invece di 15.9 MHz , mentre la resistenza sul picco della risonanza è $1.64 \text{ k}\Omega$ invece di $1.5 \text{ k}\Omega$.

Per analizzare in modo formalmente più accurato il funzionamento del circuito, conviene partire dalla conduttanza Y_A del partitore $C_1 C_2 R_L$ di fig. 16.13b tra il punto A e massa:

$$\begin{aligned} Y_A &= \left[\frac{1}{j\omega C_1} + \frac{R_L}{1 + j\omega R_L C_2} \right]^{-1} = \frac{j\omega C_1 (1 + j\omega R_L C_2)}{1 + j\omega R_L (C_1 + C_2)} \\ &= \frac{j\omega C_1 (1 + j\omega R_L C_2) [1 - j\omega R_L (C_1 + C_2)]}{1 + \omega^2 R_L^2 (C_1 + C_2)^2} \end{aligned} \quad (16.12)$$

La parte reale ed immaginaria di Y_A sono l'inverso della resistenza e della reattanza che si vengono a trovare in parallelo all'induttanza L ; nell'intorno della risonanza, quindi per $\omega \simeq \omega_0 = 2\pi f_0$, vogliamo siano eguali rispettivamente a $1/R$ e $j\omega_0 C$ di fig. 16.13a:

$$\Re(Y_A) = \frac{\omega_0^2 R_L C_1^2}{1 + \omega_0^2 R_L^2 (C_1 + C_2)^2} = \frac{1}{R} \quad (16.13)$$

$$\Im(Y_A) = \frac{\omega_0 C_1 [1 + \omega_0^2 R_L^2 C_2 (C_1 + C_2)]}{1 + \omega_0^2 R_L^2 (C_1 + C_2)^2} = \omega_0 C \quad (16.14)$$

Dalla prima di queste due espressioni si ricava

$$\frac{(C_1 + C_2)^2}{C_1^2} = \frac{R_L}{R} - \frac{1}{\omega_0^2 R_L^2 C_1^2} \quad (16.15)$$

Supponendo di poter trascurare il secondo termine a destra si ottiene nuovamente l'eq. 16.9:

$$\frac{(C_1 + C_2)^2}{C_1^2} = \frac{R}{R_L} \quad \frac{C_2}{C_1} = \sqrt{\frac{R}{R_L}} - 1 \quad (16.16)$$

Dalla seconda espressione si ottiene:

$$\frac{\frac{C_1}{\omega_0^2 R_L^2 (C_1 + C_2)} + C_1 C_2}{\frac{1}{\omega_0^2 R_L^2 (C_1 + C_2)} + (C_1 + C_2)} = C \quad (16.17)$$

Supponendo nuovamente di poter trascurare i due termini frazionari a numeratore e denominatore, l'espressione si riduce alla prima delle eq. 16.10, cioè alla capacità equivalente della serie di C_1 e C_2 :

$$\frac{C_1 C_2}{C_1 + C_2} = C \quad (16.18)$$

I termini trascurati sono quindi

$$\frac{1}{\omega_0^2 R_L^2 C_1^2} = \left(\frac{1}{2\pi \cdot 15.8 \cdot 10^6 \cdot 50 \cdot 122 \cdot 10^{-12}} \right)^2 = 2.7 \quad (16.19)$$

rispetto a $R/R_L = 30$;

$$\frac{1}{\omega_0^2 R_L^2 (C_1 + C_2)} = \left(\frac{1}{2\pi \cdot 15.8 \cdot 10^6 \cdot 50} \right)^2 \cdot \frac{1}{669 \cdot 10^{-12}} = 61 \text{ pF} \quad (16.20)$$

rispetto a $C_2 \simeq 547 \text{ pF}$ e $C_1 + C_2 \simeq 669 \text{ pF}$.

I termini trascurati hanno quindi un peso confrontabile con le tolleranze tipiche dei componenti (5 – 10%).

La condizione alla base delle approssimazioni fatte equivale a supporre che il fattore di merito Q_0 del circuito risonante sia adeguatamente elevato, tipicamente almeno 10, come già anticipato nella analisi meno rigorosa.

16.7.1 Trasformatore di impedenza ad L

Il circuito trasformatore di impedenza di fig. 16.14 è composto di tre elementi, L , C_1 e C_2 e quindi possiede tre gradi di libertà. Infatti è possibile, data la frequenza di lavoro f_0 e la resistenza equivalente R desiderata, scegliere il valore di L in base al fattore di merito (e quindi alla banda passante) che si vuole ottenere. I valori di C_1 e C_2 possono quindi essere scelti per avere il rapporto R/R' e la frequenza di risonanza che si desiderano.

A volte può essere conveniente rinunciare ad un grado di libertà, di solito la possibilità di scegliere Q_0 , ed utilizzare una rete di trasformazione con due soli componenti, una induttanza ed un condensatore.

Ad esempio, nel grafico 16.11 si vede che l'impedenza di ingresso dell'amplificatore con base a massa è molto bassa, 6.9Ω . Volendo portare anche questa al valore normalizzato standard di 50Ω si può utilizzare il circuito adattatore di fig. 16.15 composto soltanto dalla induttanza L e dal condensatore C .

Per analizzarne il funzionamento si può ripartire dalle eq. 16.13 e 16.14, ponendo $C_2 = 0$. Il generatore V vede una ammettenza complessa le cui parti reali ed immaginarie

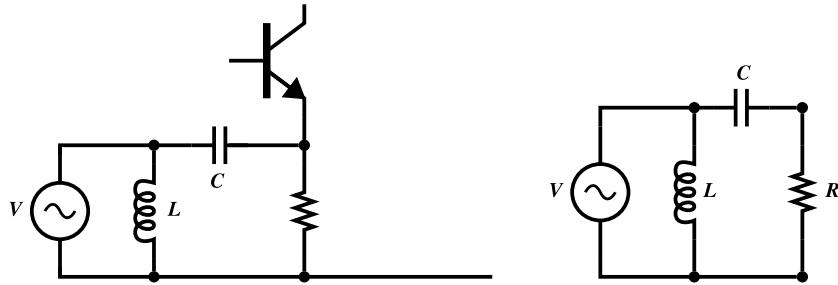


Figura 16.15: Circuito trasformatore di impedenza a due elementi, L e C , detto *circuito ad L* perchè nello schema induttanza e condensatore descrivono la forma di una L rovesciata. La resistenza R nel circuito di destra rappresenta la resistenza di ingresso dell'amplificatore.

$1/R'$ e $\omega_0 C'$ sono:

$$\Re(Y) = \frac{\omega_0^2 C^2 R}{1 + \omega_0^2 C^2 R^2} = \frac{1}{R'} \quad (16.21)$$

$$\Im(Y) = \frac{\omega_0 C}{1 + \omega_0^2 C^2 R^2} = \omega_0 C' \quad (16.22)$$

La quantità $\omega_0 R C$ è l'inverso del fattore di merito Q_0 del circuito risonante RCL per cui dalle 16.21 e 16.22 si ottengono

$$\frac{R'}{R} = Q_0^2 + 1 \quad \frac{C'}{C} = \frac{Q_0^2}{Q_0^2 + 1} \quad (16.23)$$

Con i valori del circuito 16.9b, $R = 6.9 \Omega$, $R' = 50 \Omega$ e $f_0 = 15.9 MHz$, si dovrà avere:

$$Q_0 = \sqrt{\frac{R'}{R} - 1} = 2.5 \quad L = \frac{R'}{\omega_0 Q_0} = 200 nH \quad (16.24)$$

$$C' = \frac{1}{\omega_0^2 L} = 500 pF \quad C = \frac{Q_0^2 + 1}{Q_0^2} \cdot C' = 580 pF \quad (16.25)$$

Il valore di Q_0 non è stato scelto, ma è stato dettato dal rapporto di trasformazione richiesto. Si tratta però di un valore perfettamente accettabile, a meno che non si abbiano motivi per voler introdurre un fattore di selettività più stretto nel circuito di ingresso.

16.8 Circuito cascode

Come visto nel par. 16.3, la limitazione principale alla banda passante di un amplificatore ad emettitore comune è data dalla capacità di transizione C_T della giunzione collettore-base, moltiplicata dall'effetto Miller.

Il circuito in fig. 16.16, detto amplificatore *cascode*, elimina questo problema, al costo di una maggiore complessità, impiegando due transistor invece di uno.

Il primo transistor, T_1 , è utilizzato in configurazione *emettitore a massa*, in condizioni operative analoghe a quelle del circuito in fig. 16.9a.

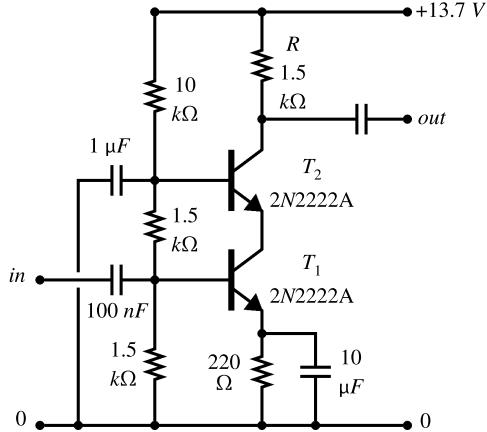
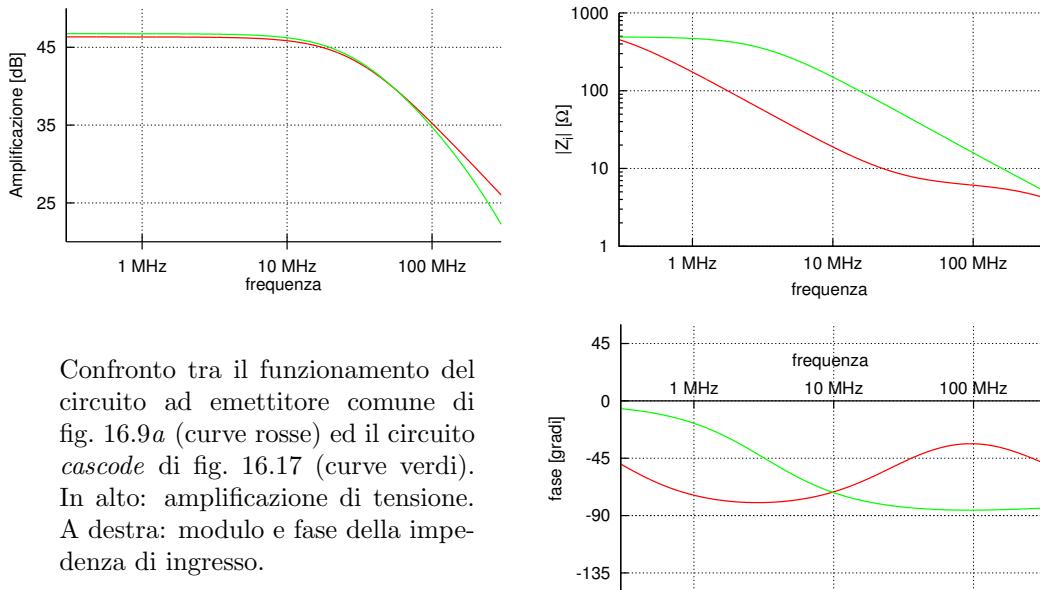


Figura 16.16: Schema elettrico del circuito *cascode*

L'impedenza di carico di questo transistor, però, non è la resistenza R , ma l'impedenza di ingresso all'emettitore del transistor T_2 , che è utilizzato nella configurazione *base a massa*. Questa coincide con l'impedenza di uscita del circuito *emitter follower* e, per quanto visto nel par. 7.7.1, è pari a $h_{ie}/(h_{fe} + 1) \simeq g_{fe}$ (eq. 7.41).

Poichè i due transistor T_1 e T_2 hanno praticamente la stessa corrente di emettitore (a meno della piccola correzione dovuta alla I_B di T_2), hanno anche la stessa transcondutanza $g_{fe1} \simeq g_{fe2} \simeq I_E/V_T$. Di conseguenza, l'amplificazione di tensione per il transistor T_1 è $A_{V1} = g_{fe1}/g_{fe2} = 1$, riducendo drasticamente il contributo dell'effetto Miller.

In pratica, il transistor T_1 fornisce solo l'amplificazione di corrente, il transistor T_2 solo l'amplificazione di tensione.



Confronto tra il funzionamento del circuito ad emettitore comune di fig. 16.9a (curve rosse) ed il circuito *cascode* di fig. 16.17 (curve verdi). In alto: amplificazione di tensione. A destra: modulo e fase della impedenza di ingresso.

Figura 16.17: Confronto tra circuito *emitter follower* e *cascode*

I risultati che si ottengono con questo circuito sono riportati in fig. 16.17, a confronto con quelli che si ottengono dal transistor singolo di fig. 16.9a, nelle stesse condizioni operative.

L'amplificazione di tensione è praticamente identica fino a circa 50 MHz ; dopo questa frequenza, l'attenuazione è più rapida per il circuito *cascode*, perchè si ha la combinazione delle frequenze di taglio introdotte da entrambi i transistor.

La frequenza di taglio relativa alla impedenza di ingresso, e quindi alla amplificazione di corrente, si trova invece spostata ad un valore più alto di circa un ordine di grandezza (da 0.3 a 3 MHz). L'andamento della fase della impedenza di ingresso nel circuito *cascode* è quella usuale di una cella con resistenza e condensatore in parallelo, senza gli effetti ulteriori che si vedono per il circuito ad emettitore comune: la configurazione *cascode* effettivamente riduce fortemente l'accoppiamento tra il circuito di uscita e quello di ingresso.

Capitolo 17

Generazione di segnali

Nei capitoli precedenti sono stati già descritti circuiti capaci di generare segnali persistenti sinusoidali (o quasi), come l'oscillatore a resistenza negativa (par. 12.11), l'oscillatore armonico (par. 13.3) e l'oscillatore a rotazione di fase (par. 15.5.2).

Le caratteristiche principali che determinano la qualità di un segnale (e del suo generatore) sono:

- precisione e stabilità della frequenza;
- precisione e stabilità della ampiezza;
- regolabilità della frequenza;
- purezza spettrale.

Di queste caratteristiche va subito detto che la seconda è la meno critica. In genere è sempre possibile con circuiti a valle del generatore correggere opportunamente l'ampiezza di un segnale.

Le caratteristiche di gran lunga più importanti sono invece la possibilità di ottenere da un generatore la frequenza voluta (regolabilità) in maniera accurata e stabile nel tempo. Come è facile da immaginare, queste caratteristiche sono di coesistenza difficile: ogni punto di regolazione presente in un circuito è anche un punto di *sregolazione*. Cioè, se nel circuito di un generatore compaiono componenti che ne determinano la frequenza ed il valore di questi componenti può essere variato per ottenere la frequenza desiderata, sarà inevitabile che questi componenti siano anche soggetti a derive tanto maggiori quanto più è grande il campo di variabilità del loro valore.

In definitiva, si possono realizzare generatori molto stabili, ma a costo di fare uso di elementi molto rigidi nella determinazione della frequenza (ad es., il cristallo piezoelettrico) e quindi con limitatissime possibilità di regolazione. Si possono realizzare generatori con una ottima possibilità di regolare la frequenza su un intervallo di valori anche molto ampio e con una grande risoluzione, ma con una stabilità inevitabilmente limitata.

Uno dei più importanti successi dell'elettronica di questi ultimi decenni è stato lo sviluppo di soluzioni circuituali che combinando opportunamente elettronica analogica e digitale, anche insieme alla reazione negativa, hanno permesso di realizzare generatori che offrono ottima regolabilità, precisione e stabilità di frequenza.

Un discorso a parte merita l'ultima caratteristica, la purezza spettrale. Un segnale ideale ha uno spettro *monocromatico*, contiene cioè una sola componente, con la frequenza desiderata.

Il segnale prodotto da un generatore reale ha uno spettro più complesso in cui compaiono:

- Armoniche superiori della frequenza fondamentale: il segnale è *distorto*.
- Segnali spurii a frequenze non correlate con la frequenza fondamentale: problema che si presenta in special modo nei generatori che fanno uso di elettronica digitale, le cui frequenze di funzionamento possono generare interferenze.
- Rumore, che va distinto tra un fondo di rumore continuo su tutte le frequenze, come è presente in tutti i circuiti reali, ed un rumore concentrato in una stretta banda di frequenze intorno alla frequenza fondamentale, per cui lo spettro del segnale non è più composto da una sola *riga* nel senso geometrico del termine, ma piuttosto da una distribuzione, normalmente a campana, intorno alla frequenza fondamentale. Questo è il *rumore di fase* (par. 17.3).

Nei tre esempi di generatori richiamati all'inizio del capitolo compaiono tre elementi base:

- un anello di reazione contenente amplificazione e reazione *positiva*;
- una rete il cui comportamento in funzione della frequenza determina la frequenza di oscillazione;
- un meccanismo di regolazione che garantisce l'avvio e la crescita dell'oscillazione e poi il raggiungimento e mantenimento di uno stato stazionario in cui l'ampiezza dell'oscillazione rimane costante.

I circuiti che fanno uso di reti *RC* per la determinazione della frequenza, come l'oscillatore armonico del par. 13.3 o l'oscillatore a rotazione di fase del par. 15.5.2 sono adatti soprattutto per segnali di bassa frequenza. Infatti presentano il vantaggio di non avere bisogno di induttanze, che nel campo delle basse frequenze sono componenti con moltissimi problemi.

Il generatore del par. 12.11, invece, per determinare la frequenza usa un circuito risonante *LC*, accoppiato ad un circuito amplificatore con reazione positiva per compensare le perdite nell'induttanza e nel condensatore e generare oscillazioni persistenti.

Questa soluzione è quella di gran lunga più usata per la generazione di segnali ad alta frequenza.

17.1 Anello con reazione positiva. Il *moltiplicatore di Q*

Mentre la reazione negativa, come visto al cap. 11, può essere usata per ridurre la dipendenza dalla frequenza del guadagno di un amplificatore, la reazione positiva può essere utilizzata per produrre l'effetto opposto, ad esempio per aumentare la selettività di un circuito risonante.

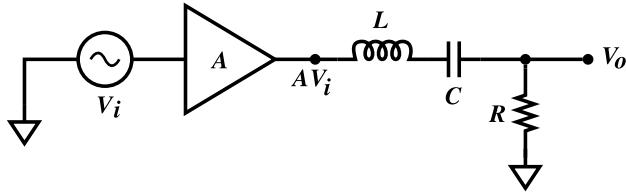


Figura 17.1: Amplificatore selettivo composto da un amplificatore A a larga banda seguito da un filtro passa-banda formato da un circuito risonante LRC .

Nella fig. 17.1 l'amplificatore A è seguito da un filtro passa-banda composto da un circuito risonante serie LC che forma con la resistenza R un partitore selettivo. Se si suppone A indipendente dalla frequenza, almeno nell'intorno della frequenza di risonanza del circuito LC , il guadagno complessivo della rete sarà

$$G_0 = \frac{V_o}{V_i} = A \cdot \frac{R}{R + j \left(2\pi f L - \frac{1}{2\pi f C} \right)} = \frac{A}{1 + j Q_0 \left(\frac{f}{f_0} - \frac{f_0}{f} \right)} \quad (17.1)$$

dove $f_0 = 1/(2\pi\sqrt{LC})$ è la frequenza di risonanza del circuito LC e $Q_0 = 2\pi f_0 L / R$ è il fattore di merito alla risonanza in presenza della resistenza R (considerando trascurabili la resistenza dovuta alle perdite di L e C e la resistenza di uscita dell'amplificatore A). Per $\Delta f = f - f_0 \ll f_0$ l'espressione tra parentesi può essere approssimata da

$$\left(\frac{f}{f_0} - \frac{f_0}{f} \right) = \frac{f^2 - f_0^2}{f_0 f} = \frac{(f - f_0)(f + f_0)}{f_0 f} \simeq \frac{2\Delta f}{f_0} \quad (17.2)$$

e quindi

$$G_0 = \frac{V_o}{V_i} = \frac{A}{1 + 2jQ_0\Delta f/f_0} \quad (17.3)$$

In fig. 17.2 è riportato l'andamento di $|G_0|^2 = |V_o/V_i|^2$ per $Q_0 = 10$ e $A = 1$, calcolato con l'approssimazione 17.2-17.3 (curva verde) e con l'espressione esatta 17.1 (curva rossa).

In fig. 17.3 il circuito è stato modificato realizzando un anello a reazione positiva che riporta il segnale V_o all'ingresso e lo somma a V_i .

Secondo quanto visto nel cap. 11, fino a che $A < 1$, l'eq. 11.3 fornisce correttamente la risposta del circuito ad un segnale sinusoidale stazionario:

$$G = \frac{G_0}{1 - \beta G_0} = \frac{\frac{A}{1 + 2jQ_0\Delta f/f_0}}{1 - \frac{A}{1 + 2jQ_0\Delta f/f_0}} = \frac{A}{1 - A + 2jQ_0\Delta f/f_0} \quad (17.4)$$

Riscrivendo questo risultato nella forma

$$G = \frac{A}{1 - A} \cdot \frac{1}{1 + 2j\frac{Q_0}{1 - A} \cdot \frac{\Delta f}{f_0}} \quad (17.5)$$

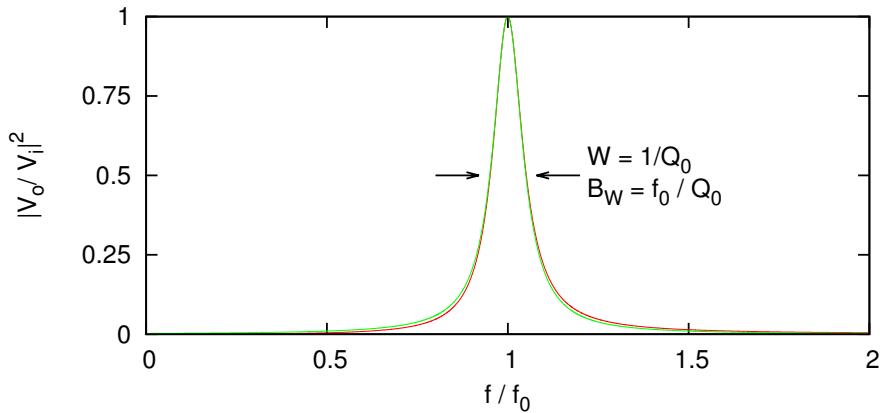


Figura 17.2: Curva di selettività di un circuito risonante con $Q_0 = 10$. La curva in verde è stata calcolata con l'espressione approssimata 17.3, quella in rosso con l'espressione esatta 17.1. La larghezza della curva a metà altezza è pari a $W = 1/Q_0$, corrispondente alla larghezza di banda $B_W = f_0/Q_0$.

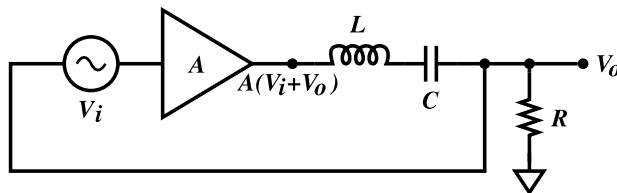


Figura 17.3: L'amplificatore di fig. 17.1 con il segnale V_o all'uscita, filtrato dal circuito risonante LRC , sommato al segnale di ingresso in modo da formare un anello con reazione positiva.

si vede che il circuito con reazione positiva presenta un guadagno ed un fattore di merito moltiplicati entrambi per il fattore $1/(1 - A)$ rispetto a quelli che avrebbe ad anello aperto (questo naturalmente sarà vero anche per tutte le altre caratteristiche del circuito che venivano attenuate dalla reazione negativa e che invece verranno esaltate: distorsione, dipendenza dai valori dei parametri etc).

Nel circuito di fig. 17.3 ad anello chiuso, per A che tende ad 1, il guadagno G ed il fattore di merito Q_0 vanno entrambi ad infinito. In questa condizione idealmente il circuito, avendo guadagno infinito, può fornire una tensione di uscita V_o anche in presenza di una tensione di ingresso V_i nulla e quindi diventare un generatore di segnali. Infatti, nel piano complesso del diagramma di Nyquist il tracciato del guadagno d'anello che compare nella eq. 17.4 per $A = 1$ passa due volte esattamente per il punto $\{-1,0\}$ (fig. 17.4). Questo indica la presenza di due zeri coniugati sull'asse $j\omega$, che difatti si trovano in corrispondenza dei punti $\pm 2j\pi f_0$, come la coppia S_3, S_3^* in fig. 15.4.

Questa condizione è riassunta nel *criterio di Barkhausen*: un anello con reazione positiva è sede di oscillazioni persistenti stazionarie quando il guadagno d'anello è esattamente $\{1,0\}$.

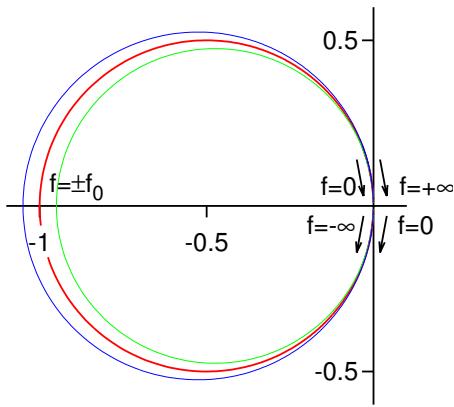


Figura 17.4: Diagramma di Nyquist per il circuito di fig. 17.3. La curva in rosso, ottenuta con $A = 1$, passa due volte esattamente per il punto $\{-1, 0\}$: il percorso della curva viene seguito una prima volta per f che va da $-\infty$ a 0 , con il punto $\{-1, 0\}$ in corrispondenza della frequenza $-f_0$ ed una seconda volta per f che va da 0 a $+\infty$, con il punto $\{-1, 0\}$ in corrispondenza della frequenza $+f_0$. Le altre due curve sono state ottenute con i valori $A = 0.95$ (curva verde) e $A = 1.05$ (curva blu). La curva verde non racchiude il punto $\{-1, 0\}$ e quindi corrisponde ad un comportamento di transitorio con oscillazione smorzata; la curva blu racchiude il punto $\{-1, 0\}$ e corrisponde ad un transitorio che cresce esponenzialmente.

In base a questa analisi la tensione all'uscita del circuito dovrebbe essere un'oscillazione perfetta con andamento $V_0 \cos(2\pi f_o t + \varphi)$, con ampiezza di picco V_0 costante. Nella realtà le cose sono più complicate.

Come si è già visto nel par. 13.3, *esattamente* non è una condizione che si possa ottenere e mantenere stabilmente in una rete reale. Qualunque minima deriva del valore di A porterà i due zeri da $\pm 2j\pi f_o$ a $\pm\epsilon \pm 2j\pi f_o$, con la comparsa quindi di un termine di smorzamento positivo o negativo $e^{\pm\epsilon t}$ che farà crescere o decadere l'oscillazione. Questo problema può essere risolto dotando il circuito di un sistema attivo di regolazione come quello in fig. 13.7 che, misurando l'ampiezza del segnale V_o all'uscita, intervenga aumentando o diminuendo leggermente il valore di A non appena V_o accenna a diminuire o ad aumentare. Nella pratica molto spesso questo risultato viene raggiunto impostando $A > 1$ per livelli di segnale piccoli e lasciando poi crescere l'ampiezza della tensione V_o dell'oscillazione fino a che la saturazione dell'amplificatore non riduce il valore effettivo di A ad 1. Ogni deriva successiva del valore di A verrà compensata da una leggera variazione di V_o .

Un esempio di circuito di questo tipo è l'*oscillatore Colpitts*, descritto nel paragrafo seguente.

17.2 Oscillatore Colpitts

L'amplificatore accordato descritto nella fig. 16.13b può essere facilmente trasformato in un generatore di segnali, collegando l'uscita (*out*) con l'ingresso (*in*) realizzando così, secondo lo schema di principio di fig. 17.3, un circuito oscillatore noto come *oscillatore Colpitts* (fig. 17.5).

Il rapporto tra i due condensatori C_1 e C_2 deve essere scelto in modo da avere un guadagno d'anello complessivo maggiore di uno. Normalmente, anzi, si fa in modo che il guadagno d'anello sia decisamente maggiore di uno, lasciando poi alla saturazione del transistor il compito di limitare e stabilizzare l'ampiezza dell'oscillazione. La distorsione del segnale che ne consegue non è un grosso problema in questo tipo di

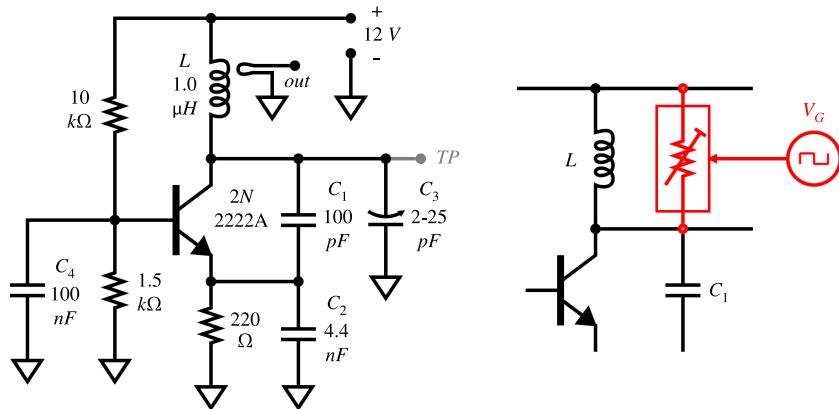


Figura 17.5: Schema elettrico di un oscillatore *Colpitts*. Con i valori dei componenti indicati la frequenza di oscillazione è di $13 \dots 15 \text{ MHz}$, regolabile tramite il condensatore variabile C_3 .

circuiti, perchè la presenza del circuito risonante LC garantisce una buona soppressione delle armoniche, almeno fino a che il suo fattore di merito Q_0 è sufficientemente elevato ed il tasso di distorsione non raggiunge livelli eccessivi.

In fig. 17.6 è riportata la registrazione, effettuata tramite un oscilloscopio digitale, della nascita, crescita e stabilizzazione dell'oscillazione, misurata sul *test point* TP .

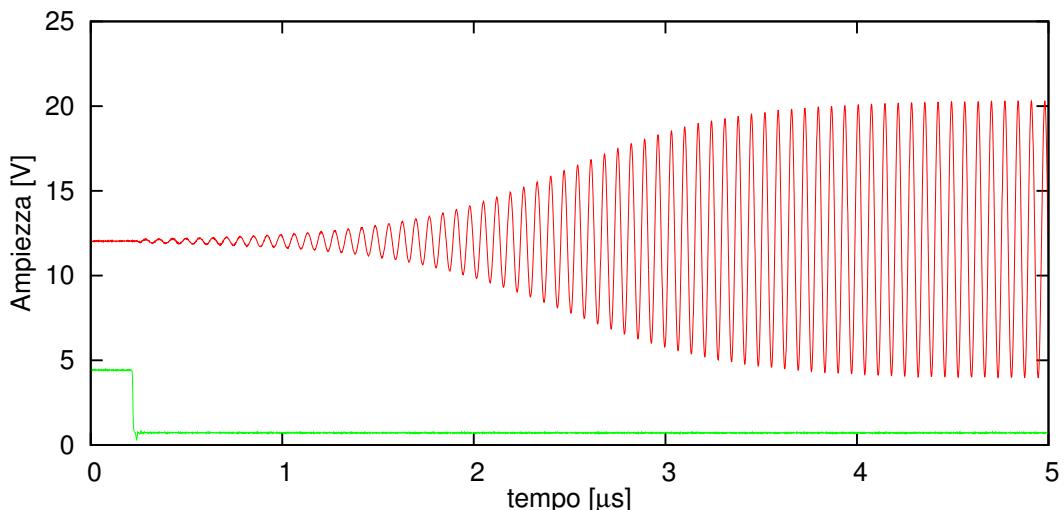


Figura 17.6: Crescita e stabilizzazione dell'ampiezza del segnale nell'oscillatore *Colpitts* di fig. 17.5.

Per effettuare questa misura è stato utilizzato il circuito modificato come indicato nella fig. 17.5, a destra, in rosso: in parallelo al circuito risonante è stato aggiunto un transistor *JFET* in funzione di resistenza variabile, comandato dal generatore V_G , programmato per inviare un'onda quadra. I valori del circuito sono stati aggiustati in modo da avere, con il transistor *JFET* acceso, un guadagno d'anello minore di uno e

spengere l'oscillazione; viceversa, con il transisitor spento, il guadagno d'anello torna a valori maggiori di uno e l'oscillazione può nascere e crescere.

Il segnale a gradino in verde nella figura è il fronte di discesa dell'onda quadra, che comanda lo spegnimento del *fet*. L'oscillazione nasce dalla piccola perturbazione indotta nel circuito risonante dall'onda quadra attraverso la capacità *gate-drain* del *fet*. In assenza di questa, l'oscillazione avrebbe origine comunque da ogni minima fluttuazione dovuta al rumore di agitazione termica sempre presente nei circuiti elettrici. Nel grafico si vede chiaramente come l'oscillazione cresce di ampiezza con andamento esponenziale per circa $2 \mu s$ e poi, in corrispondenza dell'ampiezza di $\pm 5 V$, comincia a crescere sempre più lentamente fino a che non si stabilizza asintoticamente intorno a circa $\pm 8 V$.

Nella figura 17.7 è riportato lo spettro del segnale generato una volta raggiunto il livello stazionario. L'intensità del segnale alla frequenza fondamentale $f_0 = 15.2 \text{ MHz}$ è di -2.1 dBm ($\text{dBm} = \text{decibel}$ rispetto ad 1 mW). La seconda armonica, a frequenza $2f_1$, si trova a -40 dBm , quindi a 38 dB al di sotto della fondamentale: circa 10000 volte più piccola. La quinta armonica, di intensità -83 dBm , si trova circa 3 dB sopra il fondo di rumore dell'analizzatore di spettro.

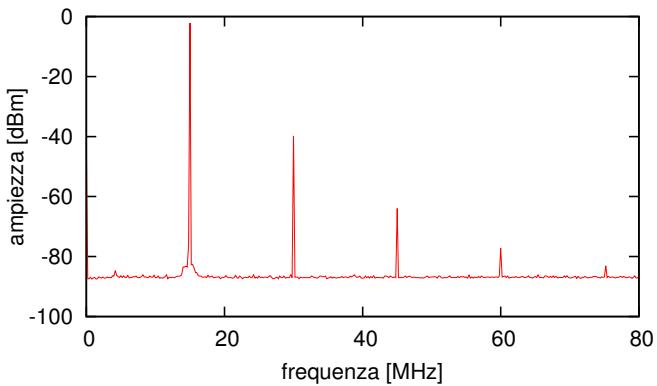


Figura 17.7: Spettro del segnale all'uscita del generatore di fig. 17.5. La frequenza fondamentale è $f_1 = 15.2 \text{ MHz}$, con una ampiezza di -2.1 dBm . Si vedono le prime quattro armoniche, oltre la fondamentale, con intensità via via decrescente: -38 , -62 , -75 e -81 dB rispetto alla fondamentale.

17.2.1 Stabilità della frequenza

Una delle prestazioni più importanti richieste ad un generatore di segnali, sia nelle applicazioni in telecomunicazioni sia nel campo della strumentazione scientifica, è la precisione e la stabilità della frequenza generata. Le due cause principali di deriva della frequenza sono le variazioni della tensione di alimentazione e le variazioni di temperatura. Per un circuito simile a quello di fig. 17.5 queste derive sono riportate nella fig. 17.8 (curve in rosso, scale dei valori a sinistra).

Nel grafico in alto è riportata la deriva al variare della tensione di alimentazione di 2 V in più o in meno rispetto al valore nominale di 12 V . Nel grafico in basso la deriva al variare della temperatura tra 26 e 42°C . I coefficienti di deriva sono rispettivamente -21 kHz/V (a 12 V) e $2.3 \text{ kHz/}^\circ\text{C}$ (a 34°C) intorno alla frequenza nominale di 12.7 MHz . In condizioni ambientali ordinarie si può avere facilmente una escursione di temperatura di 10°C nell'arco di una giornata ed anche di 40°C tra una stagione e l'altra. Questo porta a fluttuazioni della frequenza generata fino a

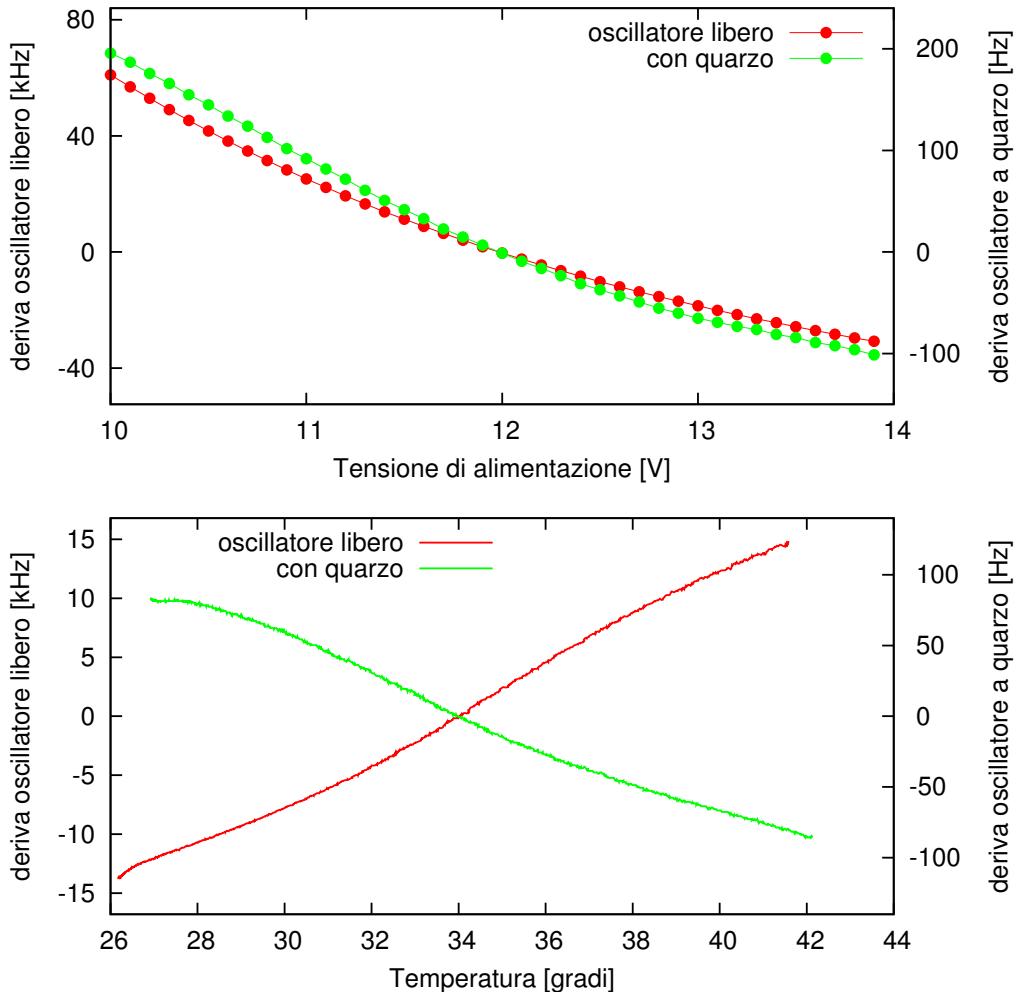


Figura 17.8: Deriva della frequenza in funzione della tensione di alimentazione (in alto) e della temperatura (in basso). Le curve rosse (con scala dei valori a sinistra) si riferiscono al generatore controllato dal solo circuito risonante LC ; le curve verdi (scala dei valori a destra) al generatore controllato da un cristallo di quarzo. Attenzione: la scala a sinistra è in $k\text{Hz}$, quella a destra in Hz .

$\pm 46 \text{ kHz}$, pari a $\pm 0.36\%$ della frequenza nominale. Un orologio controllato da questo generatore potrebbe accumulare nell'arco di una giornata un errore di diversi minuti; sarebbe un orologio di qualità decisamente scadente.

Lasciando il controllo della frequenza a componenti come induttanze e capacità (e quindi anche agli effetti delle capacità parassite, come ad esempio la capacità collettore-base del transistor) non è facile ottenere prestazioni molto migliori.

17.2.2 Oscillatore a quarzo

Per ottenere una migliore stabilità di frequenza da un oscillatore è necessario accoppiare al circuito un risuonatore che abbia una frequenza di risonanza il più possibile insensibile alle condizioni ambientali ed un fattore di merito Q molto elevato,

in modo che sia questo l'elemento di gran lunga preponderante nello stabilire la frequenza di oscillazione del circuito.

Una soluzione molto semplice e molto usata in elettronica è l'utilizzo di un risuonatore meccanico formato da una lamina di cristallo di quarzo. Le caratteristiche meccaniche del cristallo di quarzo sono molto stabili sia al variare della temperatura, sia per quanto riguarda l'invecchiamento con il passare del tempo. Il fatto che il quarzo sia un materiale piezoelettrico permette di accoppiare facilmente il risuonatore meccanico al circuito elettronico semplicemente utilizzando la lamina come dielettrico di un condensatore.

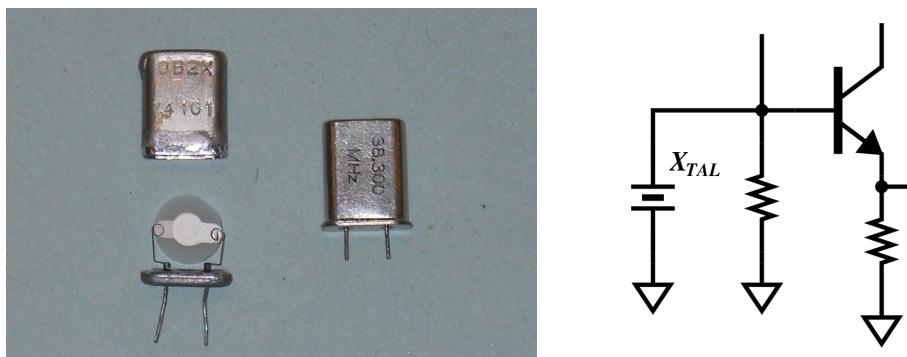


Figura 17.9: Nella foto, a sinistra: lamina di quarzo estratta dal contenitore, con gli elettrodi metallici depositati sulle due facce; a destra: il cristallo incapsulato nel suo contenitore, pronto per l'uso. Lo schema elettrico a destra indica come il cristallo di quarzo può essere inserito nel circuito oscillatore di fig. 17.5, prendendo il posto del condensatore C_4 .

Nella fig. 17.9 si vede, a sinistra nella foto, come è realizzato in pratica il risuonatore: una sottile lamina (la frequenza di risonanza dipende dallo spessore), di forma circolare (ma può essere anche quadrata o rettangolare), con due elettrodi metallici depositati sulle due facce e collegati ai terminali. Una volta incapsulata nel suo contenitore, si presenta come si vede a destra nella foto. Nello schema elettrico, a destra nella figura, è riportato il simbolo circuitale del risuonatore al quarzo (X_{TAL}), inserito nel circuito dell'oscillatore di fig. 17.5 in sostituzione del condensatore C_4 .

Elettricamente il risuonatore meccanico al quarzo equivale ad un circuito risonante RLC serie; in parallelo a questo si viene a trovare la capacità degli elettrodi, per cui il circuito equivalente complessivo è quello riportato in fig. 17.10 in alto; in basso è riportato l'andamento di modulo e fase dell'impedenza in funzione della frequenza nell'intorno della risonanza. Si osservano due risonanze: la prima, ad una frequenza più bassa (f_s), è la *risonanza serie* del circuito RLC_s (leggermente corretta dalla presenza di C_p) ed è caratterizzata dal minimo dell'impedenza (circa 10Ω per l'esemplare di cristallo utilizzato nella misura); la seconda, ad una frequenza f_p più alta di circa 25 kHz rispetto ad f_s , è la *risonanza parallelo*, caratterizzata da un massimo dell'impedenza (oltre $100 \text{ k}\Omega$), e generata dal comportamento induttivo di RLC_s in parallelo a C_p per frequenze $f > f_s$.

Poichè il circuito amplificatore su cui è costruito l'oscillatore è di tipo *base a massa*, è la frequenza f_s quella che realizza la condizione di massima amplificazione

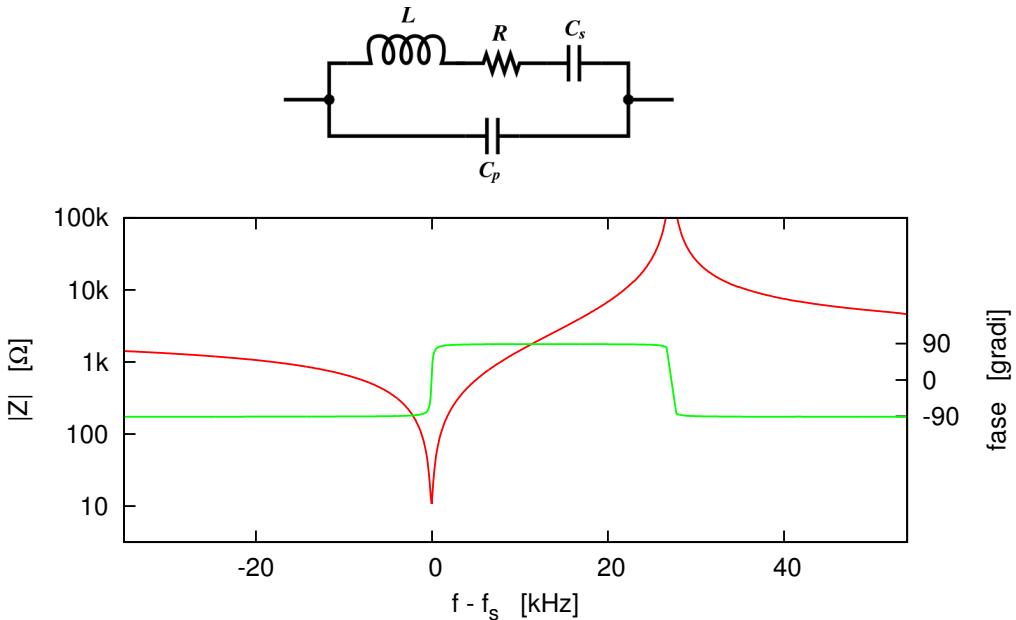


Figura 17.10: Circuito elettrico equivalente di un risuonatore al quarzo e sua impedenza complessa in funzione della frequenza, in modulo (curva rossa) e fase (curva verde). Il circuito RLC_s serie corrisponde all'elemento piezoelettrico; il condensatore C_p corrisponde alla capacità tra gli elettrodi. La frequenza f_s per il dispositivo utilizzato per questa misura è 12762.4 kHz .

e quindi corrisponde alla frequenza di oscillazione. Il fatto che la fase vari molto rapidamente da -90° a $+90^\circ$ intorno a f_s fa sì che uno spostamento lievissimo di f da f_s possa compensare eventuali errori o fluttuazioni nella fase del guadagno d'anello dagli 0° esatti necessari per avere l'oscillazione stabile.

Le derive misurate sul circuito modificato secondo lo schema di fig. 17.9 sono riportate in verde nella fig. 17.8. Nelle stesse condizioni operative utilizzate per l'oscillatore libero, si hanno rispettivamente -79 Hz/V e $-15 \text{ Hz/}^\circ\text{C}$ al variare della tensione di alimentazione e della temperatura. Il miglioramento nella stabilità è di oltre due ordini di grandezza per entrambe le derive.

Come si può facilmente intuire, il miglioramento nella stabilità del segnale ha un prezzo come contropartita: nel circuito di fig. 17.5 è possibile variare facilmente la frequenza del segnale generato, anche di alcuni MHz , ad esempio utilizzando per C_3 un condensatore variabile meccanico oppure un diodo *varicap*, la cui capacità di transizione dipende dalla tensione di contropolarizzazione applicata; nel circuito modificato, con la frequenza controllata dal cristallo di quarzo, le variazioni possibili sono solo di poche centinaia di Hz o al massimo qualche kHz .

Quanto più si chiede ad un generatore precisione e stabilità, tanto più diventa difficile avere anche la possibilità di variare a piacere la frequenza del segnale generato.

Solo con l'utilizzo di circuiti *sintetizzatori di frequenza* è possibile arrivare ad ottenere da un generatore la precisione e la stabilità dell'oscillatore a quarzo insieme alla possibilità di regolazione della frequenza tipica dell'oscillatore libero.

17.3 Rumore di fase

Come anticipato nell'introduzione al capitolo, una delle caratteristiche più importanti dei segnali prodotti da un generatore è la loro *purezza spettrale*.

Ogni circuito reale presenta un certo livello di non linearità, che si traduce in distorsione dei segnali e generazione di armoniche; anzi, il principio di funzionamento descritto nei par. 17.1 e 17.2 ha messo in evidenza come un certo tasso di distorsione, sia pure minimo, è fisiologico al mantenimento dello stato stazionario nel funzionamento di un oscillatore.

Un basso contenuto armonico quindi è sicuramente un segno di buona qualità per un generatore. Tuttavia, le armoniche sono segnali con frequenze che sono multipli interi della frequenza fondamentale, e quindi facilmente separabili o eliminabili tramite filtri abbastanza semplici.

Un'altra fonte di segnali spurii che limita la purezza spettrale di un generatore è data dal fatto che nel circuito di fig. 17.3 non si può porre semplicemente $V_i = 0$.

Qualunque circuito fisico, per quanto perfetto, contiene sempre numerose sorgenti di rumore, prima di tutto il rumore Johnson legato a tutte le componenti resistive del circuito e poi le varie sorgenti di rumore legate ai meccanismi fisici di funzionamento dei dispositivi attivi, diodi e transistor. Si può tenere conto dell'effetto complessivo di tutte queste sorgenti di rumore inserendo all'ingresso dell'amplificatore A , al posto del generatore V_i o in serie a questo, un *generatore equivalente di rumore* U_n che riproduca all'uscita l'effetto complessivo di tutti i generatori di rumore presenti. La presenza di questo rumore farà sì che il segnale all'uscita non possa più essere considerato un segnale *monocromatico puro*, cioè una sinusoide perfetta, ma abbia una composizione spettrale più complessa. In ogni istante il valore effettivo della tensione $v_o(t)$ differirà in più o in meno dal valore teorico $V_0 \cos(2\pi ft + \varphi)$, come se l'oscillazione stesse leggermente ritardando o anticipando: la fase del segnale non sarà più esattamente $2\pi ft + \varphi$, ma $2\pi ft + \varphi + \delta$, con δ variabile aleatoria. Da qui il termine *rumore di fase* per questo effetto.¹

Nell'analisi che segue sarà necessario supporre che la rete sia lineare, quindi si supporrà di avere un sistema di regolazione attivo, come quello di fig. 13.7, anche se poi nella pratica si utilizzerà il metodo di far leggermente saturare il dispositivo amplificatore. Normalmente la distorsione così introdotta si mantiene a livelli sufficientemente bassi da non alterare sensibilmente i risultati.

In presenza di un generatore di rumore U_n al posto di V_i , l'eq. 17.4 diventa

$$U_o = \frac{A}{1 - A + j2Q_0\Delta f/f_0} \cdot U_n \quad (17.6)$$

In questa equazione U_n non è l'ampiezza di una tensione nel senso abituale, ma una *densità spettrale di tensione* e sarà quindi espressa in V/\sqrt{Hz} . Di conseguenza anche U_o sarà adesso una densità spettrale di tensione, espressa nelle stesse unità; ad ogni intervallo di frequenza df corrisponde un contributo quadratico alla tensione efficace di rumore $dV_n^2 = U_n^2 df$.

La potenza P_o presente in uscita al circuito sulla resistenza R sarà data dal contributo di tutte le componenti spettrali presenti nel segnale U_o . Dalla eq. 17.6:

¹Anche la quantità V_0 viene modulata nel tempo dal rumore, ma poiché i due effetti si sommano in modo inseparabile, vengono cumulativamente indicati con il termine *rumore di fase*.

$$P_o = \int_0^\infty \frac{U_o^2(f)}{R} df = \frac{1}{R} \int_0^\infty \frac{U_n^2 A^2}{(1-A)^2 + 4Q_0^2(\Delta f/f_0)^2} df \quad (17.7)$$

Utilizzando la relazione

$$\int \frac{dx}{a^2 + b^2 x^2} = \frac{1}{ab} \arctan\left(\frac{b}{a} \cdot x\right)$$

con $x = f - f_0 = \Delta f$, $dx = df$, $a = (1 - A)$ e $b = 2Q_0/f_0$ si ottiene²

$$\begin{aligned} P_o &= \frac{U_n^2 A^2}{R} \cdot \frac{1}{(1-A)2Q_0/f_0} \cdot \left[\arctan\left(\frac{2Q_0}{1-A} \cdot \frac{f-f_0}{f_0}\right) \right]_0^{+\infty} \\ &= \frac{U_n^2 A^2}{R} \cdot \frac{f_0}{2(1-A)Q_0} \cdot \pi \end{aligned} \quad (17.8)$$

In generale A sarà veramente molto vicina ad 1, per cui si può sostituire nell'ultima equazione $A^2 \simeq 1$ e poi risolvere per $(1 - A)$:

$$(1 - A) = \frac{\pi}{2} \frac{f_0}{Q_0} \frac{U_n^2}{RP_o} \quad (17.9)$$

Il rapporto tra la densità spettrale U_n^2 del rumore all'ingresso dell'amplificatore e la potenza P_o del segnale all'uscita che viene riportato all'ingresso come reazione positiva, insieme alle caratteristiche f_0 e Q_0 del circuito risonante che fa da filtro, determinano quanto piccolo sia il valore di $(1 - A)$.

Sostituendo l'espressione 17.9 nella 17.6, e ponendo a numeratore $A \simeq 1$, si ottiene la densità spettrale di potenza del segnale:

$$P(f) = \frac{U_0^2(f)}{R} = \frac{U_n^2/R}{(1-A)^2 + (2Q_0\Delta f/f_0)^2} = \frac{P_o}{\pi f_2} \cdot \frac{1}{1 + (\Delta f/f_2)^2} \quad (17.10)$$

con

$$f_2 = \frac{\pi}{4} \frac{f_0^2}{Q_0^2} \frac{U_n^2}{RP_o} = \frac{\pi}{4} \frac{f_0^2}{Q_0^2} \frac{4kT}{P_0} \quad (17.11)$$

L'equazione 17.10 descrive una curva a campana con una semilarghezza a metà altezza pari a f_2 . Supponendo di avere, per esempio, un circuito con $f_0 = 100 \text{ MHz}$, $Q_0 = 10$, $R = 100 \Omega$, $P_o = 10 \text{ mW}$ (pari a $V_o = 1 \text{ V}_{RMS}$) e che la resistenza R sia l'unica fonte di rumore presente nel circuito, si ha

$$\begin{aligned} U_n^2 &= 4kTR = 4 \times 1.38 \times 10^{-23} \times 300 \times 100 = 1.66 \times 10^{-18} \text{ V}^2/\text{Hz} \\ f_2 &= \frac{\pi}{4} \cdot \frac{10^{16}}{10^2} \cdot \frac{1.66 \times 10^{-18}}{10^2 \times 10^{-2}} = 1.3 \times 10^{-4} \text{ Hz} \end{aligned} \quad (17.12)$$

In una situazione reale la potenza del segnale riportato all'ingresso è solo una frazione relativamente piccola del segnale totale presente all'uscita dell'amplificatore

²Poichè $1 - A$ è veramente molto piccolo, si ha $\arctan\left(\frac{-2Q_0}{1-A}\right) \simeq -\frac{\pi}{2}$

ed il rumore effettivo può essere notevolmente superiore a quello della sola resistenza di ingresso. Questo può portare a valori di f_2 più grandi anche di alcuni ordini di grandezza. Ad esempio, supponendo che il rumore effettivo sia di 10 dB superiore al puro rumore Johnson della resistenza, si ha

$$f_2 = 1.3 \times 10^{-3} \text{ Hz}$$

In grafico si otterebbe lo stesso andamento di fig. 17.2, ma con un picco veramente molto molto stretto. È quindi ragionevole associare con ottima precisione tutta la potenza del segnale alla frequenza centrale f_0 , la *portante*, di ampiezza V_o : $P_o = V_o^2/R$. La potenza associata a frequenze che distano da f_0 anche solo di pochi Hz è quasi evanescente.

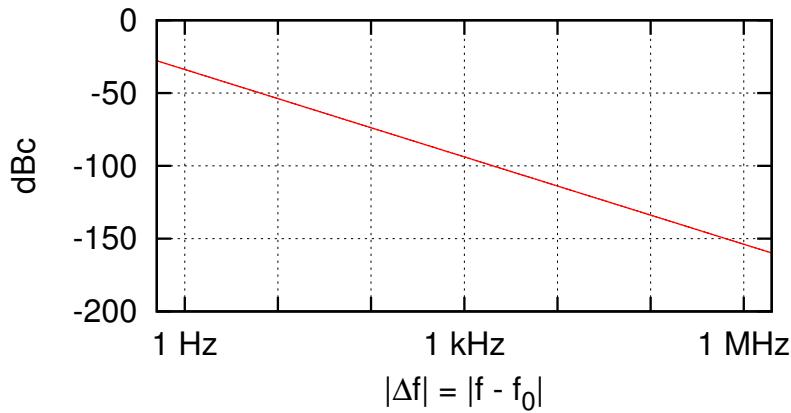


Figura 17.11: Spettro del rumore di fase del generatore di segnale corrispondente all'esempio di eq. 17.12.

Le cose però appaiono molto diverse se gli stessi dati sono osservati su una scala logaritmica. In fig. 17.11 viene riportata la funzione 17.10 espressa in unità dB_c , cioè come rapporto tra la potenza associata ad una larghezza di banda convenzionale di 1 Hz a distanza Δf dalla portante e la potenza della portante (che praticamente coincide con la potenza totale). La lettera *c* in dB_c sta per *carrier* (portante). Quindi $dB_c = decibel\ rispetto\ alla\ portante$.

Con questa definizione, con f_2 in Hz, si ha

$$dB_c [P(f)] = 10 \cdot \log_{10} \left(\frac{U_o^2(f)}{RP_o} \right) = 10 \cdot \log_{10} \left[\frac{1}{\pi f_2} \cdot \frac{1}{1 + (\Delta f/f_2)^2} \right] \quad (17.13)$$

Come si vede, le bande laterali di rumore si riducono di intensità al ritmo di 20 dB per decade di frequenza. Ma ancora ad una distanza di 50 kHz dalla portante hanno una intensità di $-128\ dB_c$. Un eventuale altro segnale presente a questa distanza dalla portante, che occupi ad esempio una banda di 2 kHz, sarebbe disturbato da un rumore con una potenza pari a $-128 + 33 = -95\ dB$ rispetto alla portante. Questo è un effetto per niente trascurabile in molte situazioni in cui si ha a che fare con strumenti molto sensibili come un ricevitore radio, un analizzatore di spettro o un radiotelescopio, in grado di elaborare segnali che si estendono su scale che possono arrivare a 120 dB di dinamica.

17.4 Sintesi di frequenza

Un *sintetizzatore di frequenza* è un apparecchio con cui si può ottenere dal segnale di un generatore di buona qualità a frequenza fissa, come ad esempio quello di un oscillatore a quarzo, altri segnali che mantengano il più possibile la qualità originaria, ma con valori di frequenza selezionabili a piacere.

I due tipi di sintetizzatori di frequenza di gran lunga più diffusi oggi sono il DDS (*Direct Digital Synthesis, Sintesi digitale diretta*) ed il PLL (*Phase Lock Loop, Anello ad aggancio di fase*).

17.5 Generatori DDS (*Direct Digital Synthesis*)

In modo molto schematico, un generatore *DDS* è una rete di calcolo digitale che fornisce ad alta velocità e con cadenza costante il valore numerico, funzione del tempo, del segnale che si vuole generare. La sequenza di valori numerici viene inviata ad un convertitore Digitale → Analogico, alla cui uscita si ottiene il segnale analogico desiderato.

Lo schema a blocchi che descrive questo principio di funzionamento è riportato in fig. 17.12. Il funzionamento della rete è regolato dall'*orologio*, che genera un segnale

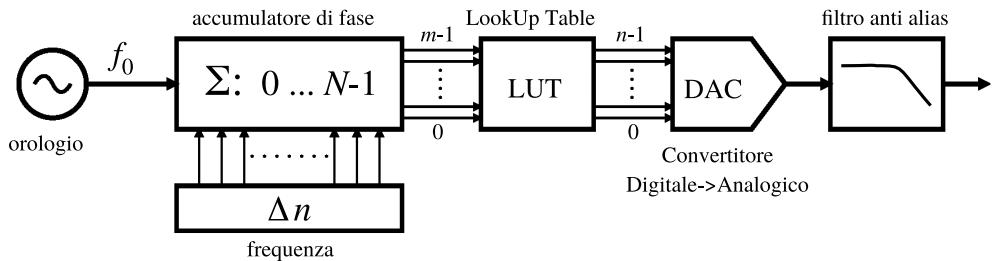


Figura 17.12: Schema a blocchi di un generatore DDS (Direct Digital Synthesis).

con frequenza f_0 , molto accurata e stabile. L'accumulatore di fase Σ è un registro digitale in grado di contenere un valore numerico n da 0 ad $N - 1$.

Ad ogni ciclo dell'orologio, quindi con cadenza f_0 , il valore n nel registro Σ viene incrementato della quantità Δn contenuta nel registro *frequenza*.

Supponendo di partire dal valore iniziale $n = 0$, l'accumulatore Σ conterrà in sequenza i valori $0, \Delta n, 2\Delta n, 3\Delta n$, ecc. Una volta raggiunto e superato N , il conteggio proseguirà ciclicamente *modulo N*, come illustrato nell'esempio della fig. 17.13.

Un ciclo completo di conteggio dell'accumulatore Σ da 0 a $N - 1$ si fa corrispondere per definizione ad un ciclo del segnale che si vuole generare, cioè ad un avanzamento della fase di 2π radianti. Il valore del bit meno significativo di Σ è di $1/N$ cicli, pari a $2\pi/N$ radianti e gli incrementi successivi Δn corrispondono ad una crescita della fase con velocità (frequenza angolare)

$$\omega = \frac{d\varphi}{dt} = \frac{2\pi}{N} \cdot \frac{\Delta n}{\Delta t} = 2\pi \cdot \frac{\Delta n}{N} \cdot f_0 \quad (17.14)$$

corrispondente ad una frequenza $f = f_0 \Delta n / N$, selezionabile con una risoluzione $\Delta f = f_0 / N$.

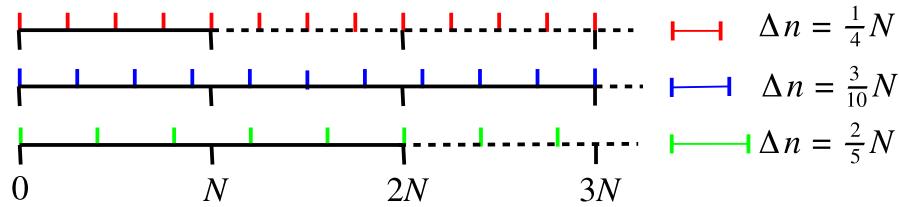


Figura 17.13: Cicli di conteggio con $\Delta n/N = 1/4, 3/10$ e $2/5$ (dall'alto in basso).

La linea nera continua identifica il periodo di ripetizione del conteggio. Nel primo caso, ogni quattro cicli di f_0 si ha un ciclo completo di Σ ed il conteggio riparte da 0 al quinto ciclo. Nel secondo caso la periodicità è di dieci cicli di f_0 per ogni tre cicli di Σ , nel terzo caso cinque cicli di f_0 ogni due cicli di Σ . Nel caso limite in cui N e Δn siano primi tra loro, il conteggio si ripete eguale solo ogni N cicli di orologio e Δn cicli dell'accumulatore Σ .

Il valore di N può essere qualunque numero intero, ma nella pratica si ha quasi sempre $N = 2^p$ oppure $N = 10^q$. Nel primo caso, l'accumulatore Σ è un contatore binario e p è compreso tipicamente tra 24 e 48 bit. Nel secondo caso, Σ è un contatore decimale e q di solito è tra 8 e 14 cifre decimali. Il contatore binario offre il massimo delle prestazioni per quanto riguarda velocità e semplicità circuitale, il contatore decimale permette di impostare codici di conteggio (e quindi frequenze generate) con valori più facilmente corrispondenti a quelli di uso comune.

Il blocco a destra dell'accumulatore Σ , cioè la *Look Up Table* (LUT), ha la funzione di ottenere dal valore $n(t)$, corrispondente a $\varphi(t) = 2\pi n(t)/N$ il valore di $\sin[\varphi(t)]$, oppure di $\cos[\varphi(t)]$, o di qualunque altra funzione periodica che si intende generare. Il successivo convertitore DAC trasformerà questo valore numerico nella corrispondente tensione analogica, funzione del tempo.

Il modo di gran lunga più veloce per ottenere il valore di una funzione trascendente come $\sin()$ e $\cos()$, che richiede uno sviluppo in una serie polinomiale contenente numerosi termini, è di utilizzare una tabella di risultati precalcolati: la LUT è una memoria digitale contenente M valori della funzione desiderata, ad esempio $\sin(\varphi)$, organizzati in maniera tale da coprire esattamente un ciclo. Il valore m -esimo della tabella, sarà $\sin(2\pi m/M)$.

Il numero M di valori della tabella è collegato alla risoluzione del convertitore DAC, corrispondente al numero R di valori diversi di tensione che possono essere generati, con $R = 2^r$ per un convertitore binario ad r bit, oppure $R = 10^r$, per un convertitore a base 10 con r cifre (ma questa soluzione nella pratica è molto meno comune della prima).

Poichè normalmente $R \ll N$, non è necessario che la LUT sia una tabella con N valori, tanti quanti sono i valori possibili nell'accumulatore Σ ; si avrebbe una inutile ripetizione di valori eguali. Normalmente si avrà $M > R$ di un fattore corrispondente ad una differenza di lunghezza delle due quantità di alcuni bit.

Ad esempio, nel caso di un convertitore binario a 12 bit, M potrà essere un numero di lunghezza da 14 a 16 bit. Allora, i 14...16 bit più significativi dell'accumulatore Σ , contenenti il valore m e corrispondenti alla fase $\varphi_m = 2\pi m/M$, inviati alla LUT, sono utilizzati per selezionare il valore m -esimo nella tabella, che sarà il valore della funzione desiderata, ad esempio $\sin(\varphi_m)$, con la precisione di r bit.

Non deve sorprendere il fatto che di solito N sia sensibilmente maggiore di M e di R . Il valore di M dipende dalla risoluzione di tensione R del DAC, che è limitata da vincoli fisici e tecnologici. Il valore N corrisponde invece alla risoluzione di frequenza che si vuole ottenere, che è limitata da vincoli molto meno stringenti, legati solo alla velocità delle porte logiche della rete digitale.

Utilizzando, ad esempio, un accumulatore Σ di 32 bit, con $N = 2^{32}$, corrispondente ad un numero decimale da 0 a circa $4.3 \cdot 10^9$, con $f_0 = 100 \text{ MHz}$, si potrebbero ottenere (in teoria) tutte le frequenze

$$f = \frac{\omega}{2\pi} = \frac{\Delta n}{N} \cdot f_0 = 0 \dots 50 \text{ MHz} \quad (17.15)$$

con una risoluzione di $100 \text{ MHz}/2^{32} = 0.023 \text{ Hz}$.

In base al *teorema del campionamento* (Teorema di Shannon-Nyquist) la massima frequenza che può essere sintetizzata in questo modo è $f < f_0/2 = 50 \text{ MHz}$, corrispondente a $\Delta n = N/2$. Nella pratica un limite più realistico è $f \leq f_0/3$.

17.5.1 Rumore nei generatori DDS

Il segnale che si ottiene all'uscita del DAC nel circuito di fig. 17.12 non è esattamente quello desiderato. Nel caso della funzione $\sin(\omega t)$, ad esempio, non si avrà un segnale sinusoidale, funzione continua del tempo, ma una *scaletta* di valori discreti, a causa del numero finito di campionamenti nel tempo e del numero finito di livelli di tensione del convertitore.

Anche la dimensione M della LUT è una fonte di errore, perché introduce un arrotondamento del valore della fase generata dall'accumulatore ed inviata al DAC, prima ancora dell'analogico effetto prodotto dal DAC sui valori della tensione. Questa è la ragione per cui normalmente si adotta $M > R$ di un fattore sufficiente a ridurre questo errore a livelli trascurabili.

Nell'analisi che segue verranno quindi considerati solo gli effetti delle prime due fonti di errore: il numero finito di campionamenti nel tempo ed il numero finito di livelli di tensione del convertitore.

Nella fig. 17.14 sono riportati due esempi per un generatore con $N = 4 \cdot 10^9$, $R = 4000$ ed $f_0 = 100 \text{ MHz}$. Questi valori sono molto vicini a valori binari di uso comune ($N = 2^{32} = 4\ 294\ 967\ 296$, $R = 2^{12} = 4\ 096$), ma sono arrotondati a valori decimali per maggiore semplicità di presentazione.

Nel grafico a sinistra, un segnale con frequenza $f = 400 \text{ Hz}$ all'uscita del DAC non è praticamente distinguibile da quello desiderato. Nel grafico al centro è riportato un ingrandimento del grafico di sinistra, per un intervallo di ampiezza pari allo 0.6% del fondo scala, in prossimità dello zero. Si vede chiaramente l'effetto dei valori discreti della tensione, che varia a *scaletta*, con gradini di ampiezza pari alla risoluzione del DAC. La traccia verde riporta le differenze intorno allo zero tra il segnale generato dal DAC ed il segnale teorico.

Nel grafico a destra, calcolato per una frequenza di lavoro $f = 12.5 \text{ MHz}$, si vede che il segnale all'uscita del DAC è solo una copia molto rozza del segnale che si desidera. In questo caso il fattore limitante è la velocità con cui vengono generati i valori all'uscita del DAC: ogni ciclo del segnale è composto da solo 8 campionamenti diversi.

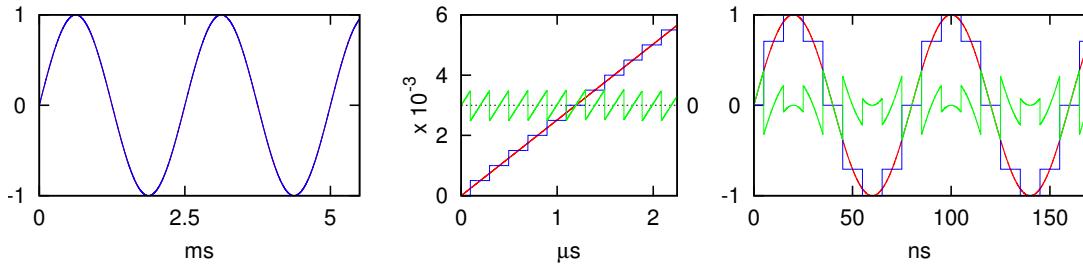


Figura 17.14: Confronto tra i valori teorici della funzione $\sin(\omega t)$ (in rosso) ed i valori ottenuti nelle condizioni operative indicate nel testo (in blu). La scala verticale riporta l'ampiezza relativa al fondo scala. A sinistra: segnale di frequenza 400 Hz. Il valore teorico e quello reale non sono distinguibili nel grafico. Al centro: lo stesso segnale, nel primo μs . I salti di livello hanno l'ampiezza di $1/2000 = 0.05\%$ del fondo scala. A destra: segnale di frequenza $f = 12.5 MHz$. In verde sono riportate le differenze tra il segnale sinusoidale desiderato e quello presente all'uscita del DAC.

In ogni caso, il segnale generato (curve in blu nei grafici) è composto dal segnale ideale desiderato (curve in rosso nei grafici), più il segnale differenza tra questo ed il segnale reale (curve in verde). Le componenti spettrali del segnale differenza sono il *rumore* che il generatore produce associato al segnale. Il filtro presente a destra del DAC ha lo scopo di eliminare il più possibile tale rumore.

Pur essendo il segnale differenza calcolabile matematicamente, come si vede dagli esempi, il calcolo degli spettri di rumore non è agevole. Infatti, le sequenze di valori generate dal DAC si ripetono periodicamente solo al raggiungimento di un valore cumulativo di conteggio dell'accumulatore Σ pari al minimo comune multiplo tra N e Δn , quindi con un periodo che salta tra valori molto diversi anche per frequenze contigue, per esempio frequenze il cui Δn differisce di una sola unità.

In generale, dal grafico centrale di fig. 17.14 si può osservare che per segnali di bassa frequenza il rumore ha una ampiezza picco-picco complessiva di ± 1 LSB = 0.05%. Lo spettro di Fourier corrispondente a questo rumore si può vedere in fig. 17.15, fino alla frequenza di 20 kHz. Il segnale principale, a 400 Hz, è riportato con l'ampiezza

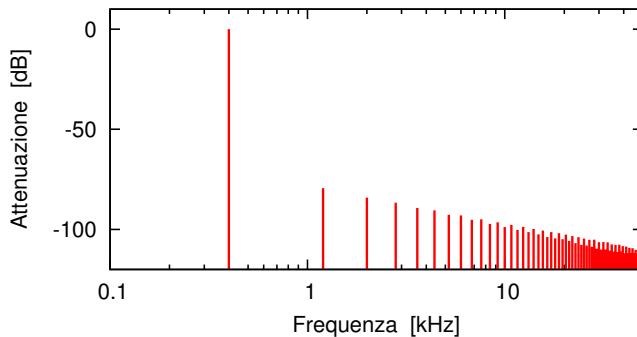


Figura 17.15: Spettro fino a 20 MHz del segnale a 400 Hz di fig. 17.14, con la fascia di rumore dovuta al campionamento.

di riferimento di 0 dB. Oltre a questo, compaiono le armoniche dispari, di ampiezza via via decrescente. La più significativa è la terza, alla frequenza di 1200 Hz, con ampiezza di circa 80 dB inferiore al segnale principale (in tensione, $1/10^4$). La som-

ma di tutte le armoniche che costituiscono lo spettro del rumore porta all'ampiezza totale di $5/10^4$ vista prima.

In questo caso il calcolo dello spettro è stato agevolato dal fatto che un ciclo del segnale corrisponde esattamente ad un ciclo di conteggio da 0 ad N , perché Δn è sottomultiplo di N .

Diversamente, lo spettro va calcolato su un periodo corrispondente a più cicli, come visto prima. Questo fa comparire componenti subarmoniche del segnale che si vuole generare. Ma l'ampiezza cumulativa di tutte le componenti di rumore dello spettro rimane comunque pari a ± 1 LSB. Di conseguenza è molto difficile veder comparire nel rumore componenti a livello superiore a -80 dB.

È intuitivo che il miglioramento del rapporto Segnale/Rumore in queste condizioni operative può essere ottenuto principalmente aumentando la risoluzione del DAC. L'utilizzo di filtri all'uscita del DAC, invece, è di efficacia scarsa, soprattutto se si vuole mantenere la flessibilità che l'apparecchio offre di poter variare liberamente e facilmente su una banda molto ampia la frequenza che si intende generare.

La situazione riportata nell'esempio a destra della fig. 17.14, per la generazione di un segnale a 12.5 MHz, è invece completamente diversa. In questo caso il problema più grosso è dato dal numero limitato di campionamenti contenuti all'interno di un ciclo, per cui l'ampiezza del rumore (differenza tra segnale generato e desiderato) arriva come valore di picco oltre il 30% del fondo scala. Lo spettro corrispondente è riportato nella figura fig. 17.16.

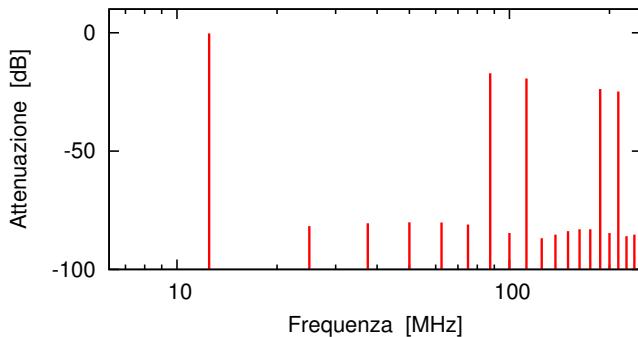


Figura 17.16: Spettro del segnale a 12.5 MHz di fig. 17.14, con la fascia di rumore dovuta al campionamento.

Anche qui Δn è un sottomultiplo di N ed un ciclo del segnale corrisponde esattamente ad un conteggio da 0 a N . Lo spettro contiene quindi la frequenza fondamentale a 12.5 MHz e le armoniche a $25, 37.5, 50 \dots$ MHz, tutte con ampiezza di circa 80 dB inferiore alla fondamentale.

Alle frequenze di 87.5 e 112.5 MHz compaiono invece i due segnali *immagine* del segnale a 12.5 MHz, differenza e somma di f_0 con f , con ampiezza di solo 18 dB inferiore al segnale fondamentale. Questi segnali spurii possono essere fortemente ridotti dal filtro anti-alias (anti-immagine) presente all'uscita del DAC, filtro passabasso che nelle condizioni operative dell'esempio dovrà avere una frequenza di taglio tra 30 e 50 MHz. Nei generatori di buona qualità il filtro anti-alias sarà composto da una sequenza di celle passa-basso come quella di fig. 1.12 (per gli apparecchi per alta frequenza) o come quella di fig. 12.15 (per gli apparecchi per bassa frequenza), con valori di frequenza e selettività combinati in maniera opportuna per generare una

curva di risposta il più possibile piatta nella regione di frequenze da trasmettere ed il più possibile ripida una volta oltrepassata la frequenza di taglio.

Infine, nella fig. 17.17 è riportato lo spettro nel caso di una frequenza $f = 15 \text{ MHz}$. Il ciclo di conteggio complessivo risulta essere lungo $3N$ e contiene 3 cicli della frequen-

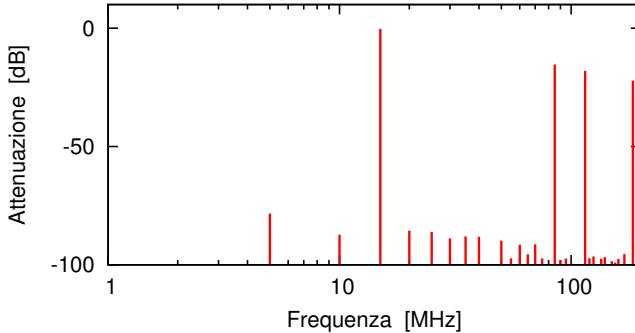


Figura 17.17: Spettro di un segnale a 15 MHz .

za f . La frequenza di ripetizione fondamentale del conteggio è quindi $f/3 = 5 \text{ MHz}$. Difatti, nello spettro compaiono nuovamente armoniche e subarmoniche tutte multiple di 5 MHz , oltre alle due frequenze immagine a 85 e 115 MHz . L'ampiezza massima tra armoniche e subarmoniche è -78 dB , per la componente a 5 MHz .

Anche qui il filtro anti-alias può eliminare solo le due frequenze-immagine e la parte del fondo con frequenze superiori alla frequenza di taglio (30 MHz).

Il fondo di segnali spuri con composizione fortemente variabile in funzione del rapporto f/f_0 e presenza di subarmoniche è una caratteristica tipica ed ineliminabile dei generatori DDS.

La qualità del segnale generato può essere migliorata solo tenendo basso il rapporto f/f_0 ed utilizzando convertitori DAC ad alta risoluzione. Entrambe queste caratteristiche sono in conflitto con la generazione di frequenze f molto elevate.

17.6 Generatori ***PLL (Phase Lock Loop)***

Mentre è relativamente facile, come visto nel par. 10.11, da un segnale a frequenza fissa e stabile f_0 generare tutte le frequenze sottomultiples f_0/N , con N intero qualsiasi, la cosa realmente utile è la funzione opposta: da una frequenza di riferimento f_0 , generare tutti i multipli $N \cdot f_0$. Ma il circuito *moltiplicatore di frequenza*, a differenza del divisore, non è possibile altro che per pochi valori di N , piccoli, principalmente 2 e 3.

Anche in questo caso, come già visto in altri casi, è possibile ottenere la funzione non realizzabile direttamente (la moltiplicazione di frequenza) inserendo nella rete di reazione di un anello con reazione negativa la funzione inversa (la divisione di frequenza), che invece è facilmente realizzabile.

L'anello di reazione che svolge questa funzione, il cui schema di principio è riportato in fig. 17.18, prende il nome di *Phase Lock Loop (PLL - Anello ad aggancio di fase)*.

Un oscillatore a frequenza variabile controllato da una tensione, ad esempio variando la frequenza di risonanza di un circuito *LC* mediante un diodo varicap (*VCO*)

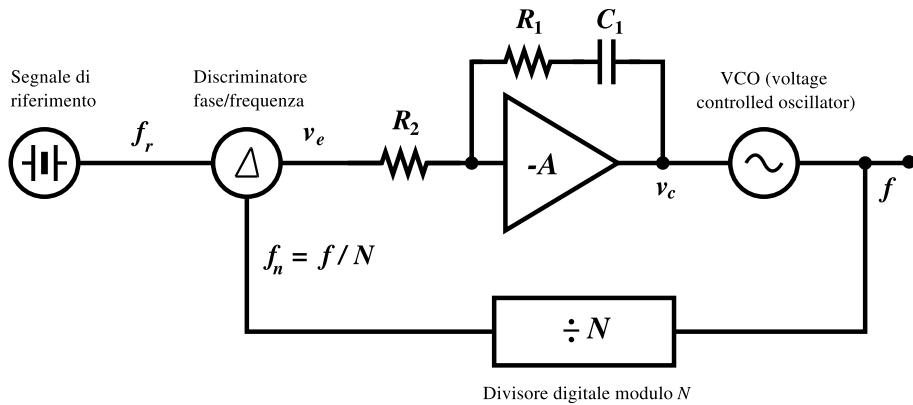


Figura 17.18: Schema a blocchi di un generatore PLL (Phase Lock Loop - Anello ad aggancio di fase).

- *Voltage Controlled Oscillator*), genera un segnale a frequenza f . Un *Divisore digitale* programmabile, del tipo di quello descritto nel par. 10.11, trasforma il segnale a frequenza f in un nuovo segnale a frequenza f/N . Un segnale di riferimento con frequenza f_r viene generato da un oscillatore a frequenza fissa, il più possibile accurato e stabile. Il circuito *Discriminatore di fase-frequenza* confronta il segnale f/N con il segnale f_r , generando un segnale di errore v_e che attraverso un circuito amplificatore/integratore va a controllare la frequenza f generata dal VCO. Se il segnale f/N risulta avere una frequenza maggiore di f_r , o se la sua fase anticipa quella di f_r , il segnale v_e agisce nel senso di far *rallentare* il VCO. Viceversa, se il segnale f/N risulta in ritardo rispetto a f_r , il VCO viene fatto *accelerare*. Se il circuito è ben progettato, raggiunge rapidamente una condizione stazionaria in cui $v_e = 0$ ed i due segnali f/N ed f_r sono in coincidenza sia in frequenza che in fase. In questa condizione all'uscita del circuito è disponibile un segnale a frequenza $f = f_r \times N$. Variando N possono essere ottenuti tutti i multipli desiderati della frequenza f_r .

17.7 Stabilità degli anelli *PLL*

Come tutti gli anelli con reazione negativa, anche il PLL è sottoposto alle condizioni di stabilità ed instabilità oggetto del criterio di Nyquist e che sono state descritte nel cap. 15. L'analisi del funzionamento dell'anello di fig. 17.18 si può fare in modo analogo agli anelli già visti, ma con una complicazione. La variabile *frequenza* compare due volte nelle equazioni dell'anello nel dominio delle frequenze: come variabile indipendente nelle funzioni di trasferimento dei vari blocchi che formano l'anello; come *valore* del segnale f del VCO, del divisore digitale, del discriminatore fase-frequenza. Le funzioni di trasferimento di questi blocchi risultano quindi essere del tipo *frequenza(frequenza)*, in cui i due termini *frequenza* significano cose diverse, generando una certa dose di confusione mentale. Una volta avvertiti di questo problema, ed usando le dovute attenzioni, il modo di procedere non è diverso da quanto visto nel cap. 15.

17.7.1 Guadagno d'anello in un PLL

Nelle espressioni che seguono si indicherà con f la frequenza del segnale generato dal VCO e con f_r quella del segnale di riferimento, come riportato nella fig. 17.18. Si utilizzerà invece la lettera F per indicare la variabile frequenza nei diagrammi di Bode e nelle funzioni di trasferimento dei vari blocchi dell'anello. Quindi, la funzione di trasferimento del blocco amplificatore sarà

$$H_a = \frac{V_c}{V_e} = -\frac{R_1}{R_2} \left(1 - \frac{j}{2\pi F R_1 C_1} \right) \quad (17.16)$$

dove, seguendo la convenzione abituale, V_c e V_e sono le trasformate nel dominio delle frequenze delle funzioni del tempo v_c e v_e .

A differenza di H_a , che è adimensionale, la funzione di trasferimento del VCO avrà dimensioni Hz/V , in quanto esprime il rapporto tra la frequenza f del segnale generato e la tensione di controllo v_c inviata al VCO. Di solito questa quantità può essere considerata una costante, non dipendente da F (attenzione, F è la frequenza con cui varia la tensione di controllo e con cui varia quindi l'altra frequenza, f , quella all'uscita del generatore). In altre parole, si suppone che la risposta del generatore sia *pronta*, f segue v_c senza ritardo. Si avrà quindi $df/dv_c = k_v = \text{costante}$ e $H_v = k_v$. Se la risposta del VCO non fosse *pronta*, cioè non fosse sufficientemente veloce ma fosse caratterizzata da un comportamento di tipo *passa basso* con costante di tempo τ_v , allora si dovrebbe scrivere:

$$H_v = \frac{k_v}{1 + j2\pi F \tau_v} \quad (17.17)$$

Si può anche avere il caso in cui il VCO non sia perfettamente lineare, cioè che la quantità $k_v = df/dv_c$ dipenda dal valore di v_c e dalla frequenza generata f . Anzi, si può dire che questo sia il caso più comune. Ci sono applicazioni in cui questa non linearità può essere un problema, ma non per quanto riguarda la stabilità dell'anello. Almeno fino a che le variazioni non vadano ad interessare l'ordine di grandezza di k_v .

Il blocco successivo dell'anello è il divisore di frequenza modulo N . La sua funzione di trasferimento è

$$H_d = \frac{\text{frequenza di uscita } f_n}{\text{frequenza di ingresso } f} = \frac{1}{N} \quad (17.18)$$

Il discriminatore fase-frequenza che si incontra lungo il percorso del segnale dopo il divisore è l'oggetto più complesso di tutto l'anello. Quando viene raggiunto lo stato stazionario (il *lock*, l'aggancio della fase), i due segnali f_r e $f_n = f/N$ sono sincroni, cioè hanno la stessa frequenza e la stessa fase. Fluttuazioni nel funzionamento, dovute a rumore, derive, interferenze, si traducono in una differenza di fase tra f_r e f_n che viene trasformata dal discriminatore in una tensione di errore v_e . Questa, amplificata ed integrata dal blocco amplificatore, modifica la marcia del VCO fino a che l'errore non viene azzerato. In questo modo di funzionamento il discriminatore sarà solo un rivelatore di fase, caratterizzato da un coefficiente di sensibilità k_f con dimensioni, ad esempio, V/ciclo . Poiché però il segnale di ingresso al discriminatore è una frequenza e la fase è l'integrale della frequenza, la funzione di trasferimento del blocco sarà:

$$H_f = \frac{k_f}{j2\pi F} \quad (17.19)$$

Il discriminatore fase-frequenza di solito comprende anche un filtro passa-basso, con costante di tempo τ_d (si vedrà in seguito perchè) e quindi la funzione di trasferimento diventa

$$H_f = \frac{k_f}{j2\pi F} \cdot \frac{1}{1 + j2\pi F\tau_d} \quad (17.20)$$

Combinando insieme i quattro blocchi si ha il guadagno d'anello

$$\begin{aligned} H(F) &= H_a \cdot H_v \cdot H_d \cdot H_f \\ &= -\frac{R_1}{R_2} \left(1 - \frac{j}{2\pi F R_1 C_1} \right) \cdot k_v \cdot \frac{1}{N} \cdot \frac{k_f}{j2\pi F} \cdot \frac{1}{1 + j2\pi F\tau_d} \end{aligned} \quad (17.21)$$

Le due costanti k_v e k_f hanno dimensioni rispettivamente $(cicli/sec)/V$ e $V/ciclo$, in cui il *ciclo* è quello del segnale f_n proveniente dal VCO. Quindi il prodotto $k_v \cdot k_f$ ha dimensioni s^{-1} , come la variabile F a denominatore dell'ultimo termine, ed il guadagno d'anello risulta adimensionale come deve essere.

Il diagramma di Bode della funzione $H(F)$ è riportato nella fig. 17.19.

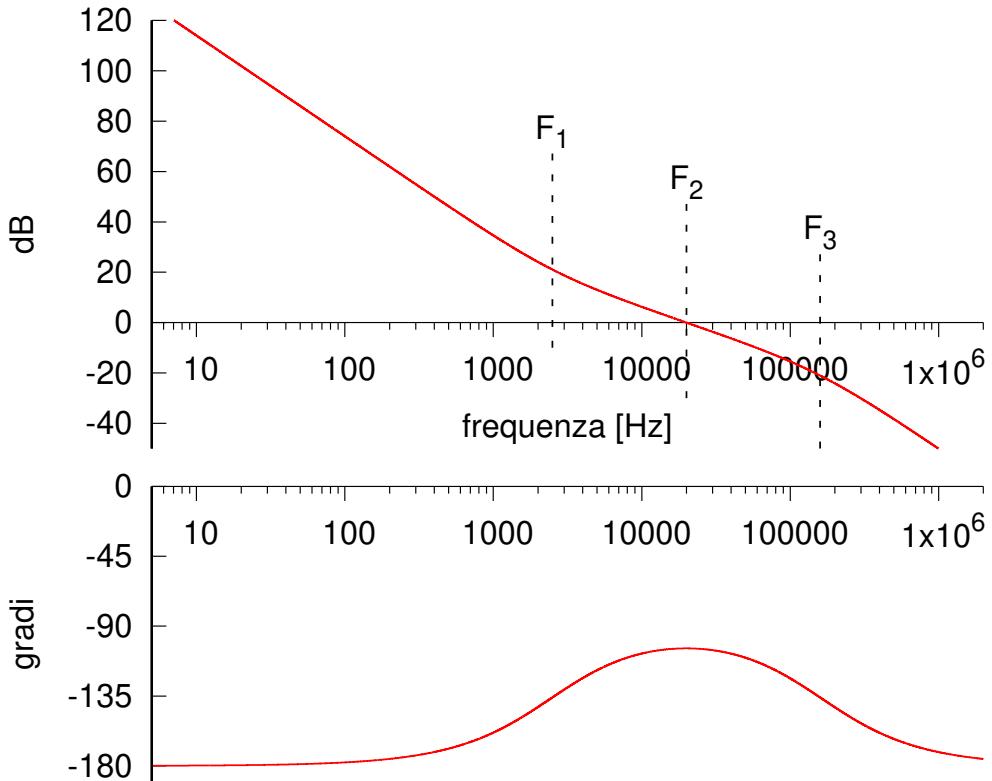


Figura 17.19:
Diagramma di Bode del guadagno d'anello del circuito di fig. 17.18 (eq. 17.21).

Nel limite delle frequenze molto basse si ha $H(F) \propto 1/F^2$ per effetto delle due integrazioni date dalla relazione tra frequenza e fase e dal blocco amplificatore. Nel grafico di Bode questo corrisponde al tratto che si trova a sinistra della frequenza

F_1 ed è caratterizzato da una pendenza di 40 dB/decade e da una fase che arriva a -180° .

La presenza della resistenza R_1 introduce nella funzione di trasferimento la frequenza di taglio F_1 , oltre la quale il guadagno continua a scendere, ma con pendenza di 20 dB/decade , mentre la fase ruota tornando verso i -90° . Questo è indispensabile per arrivare ad avere un margine di fase adeguato quando $|H(F)| = 1$ (frequenza F_2).

La terza frequenza di taglio F_3 , data dalla costante di tempo τ_d del discriminatore fase-frequenza, introduce un ulteriore polo nella funzione $H(F)$, che riporta la pendenza a 40 dB/decade e fa nuovamente ruotare la fase verso i -180° . Quindi F_3 si deve trovare sufficientemente distante da F_2 , quando il modulo di $H(F)$ è sceso sufficientemente al disotto di 0 dB e l'effetto sul margine di fase intorno a F_2 è limitato.

Nella equazione 17.21 i termini k_v e k_f dipendono dalle caratteristiche del VCO e del discriminatore e sono quindi prestabiliti; il fattore N è determinato dal rapporto f/f_r ; la costante di tempo τ_d è determinata sulla base di altre considerazioni, discusse più avanti.

Una volta nota $F_3 = 1/(2\pi\tau_d)$, si dovranno scegliere i valori di F_1 ed F_2 in modo da avere un rapporto conveniente per F_3/F_2 ed F_2/F_1 per il miglior compromesso tra velocità di risposta e stabilità.

Nell'intervallo di frequenze tra F_1 ed F_3 il guadagno d'anello è descritto dalla relazione

$$|H(F)| = \frac{R_1}{R_2} \cdot k_v \cdot \frac{1}{N} \cdot \frac{k_f}{2\pi F} \quad (17.22)$$

da cui, ponendo $|H(F)| = 1$ si ottiene

$$\frac{R_1}{R_2} = \frac{2\pi F_2 N}{k_v k_f} \quad (17.23)$$

Questo determinerà il rapporto R_1/R_2 , mentre il prodotto $R_1 C_1$ sarà dato da F_1 . Considerazioni di opportunità determineranno i singoli valori di R_1 , R_2 e C : ordine di grandezza più conveniente per le resistenze, valori di C disponibili, ecc.

Nel grafico di fig. 17.19, si ha $F_3 = 160 \text{ kHz}$ e si è scelto un rapporto $F_3/F_2 = F_2/F_1 = 8$ come compromesso tra velocità e stabilità.

17.7.2 Discriminatore di fase

La differenza di fase φ tra due segnali sinusoidali $s_1 = A \cos(\omega_1 t + \varphi)$ e $s_2 = B \cos(\omega t)$ alla stessa frequenza³ ω può essere ottenuta usando un moltiplicatore analogico, come ad esempio quello descritto nel par. 13.2.

Dalle formule di prostaferesi (vedi par. 19.1) si ha

$$A \cos(\omega_1 t + \varphi) \cdot B \cos(\omega t) = \frac{AB}{2} [\cos(2\omega t + \varphi) + \cos \varphi] \quad (17.24)$$

Un filtro passa basso all'uscita del moltiplicatore sopprime la componente a frequenza 2ω e lascia inalterato il termine $\cos \varphi$, come nel circuito in fig. 13.4b.

³Ovviamente angolare; la relazione $\omega = 2\pi f$ viene data per nota e scontata.

Perchè questo circuito funzioni correttamente è necessario che i due segnali siano in quadratura, cioè φ sia nell'intorno di 90° : in questa condizione il segnale di errore v_e è proporzionale all'errore di fase $\Delta\varphi = \varphi - 90^\circ$:

$$v_e = \frac{AB}{2K} \sin \Delta\varphi \approx \frac{AB}{2K} \Delta\varphi \quad (17.25)$$

dove K è il fattore di scala del moltiplicatore (eq. 13.4).

Nel caso che i due segnali s_1 e s_2 siano in forma di onda quadra, una porta XOR svolge la stessa funzione del moltiplicatore analogico (fig. 17.20).

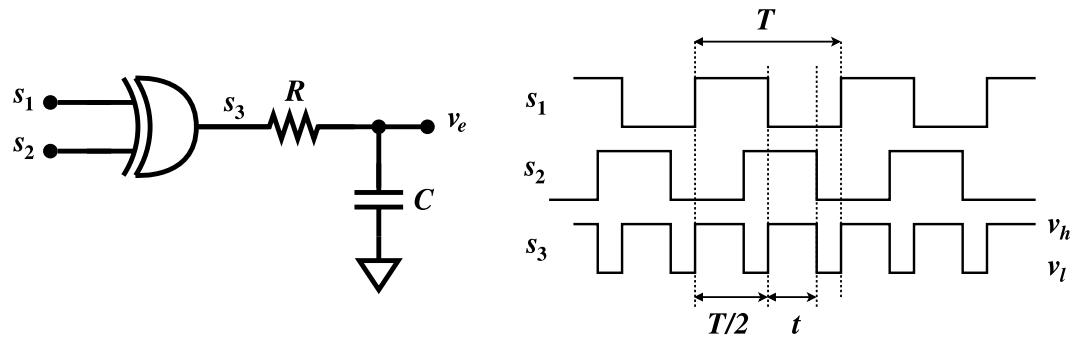


Figura 17.20: Discriminatore di fase tra due segnali ad onda quadra di eguale frequenza. Se $RC \gg T$, all'uscita v_e sarà presente solo il valore medio di s_3 : $\langle s_3 \rangle = (v_h - v_l) \cdot 2t/T + v_l$.

17.7.3 Discriminatore di fase-frequenza

I due circuiti visti nel paragrafo precedente possono essere utili solo quando i due segnali da confrontare hanno la stessa frequenza, o frequenze molto vicine.

Infatti, se le frequenze non coincidono, ma hanno valori ω_1 e ω_2 , entrambi i circuiti daranno all'uscita due segnali con frequenza $\omega_1 + \omega_2$ ed $\omega_1 - \omega_2$. Il segnale di errore v_e avrà media nulla e questo renderà molto difficile, se non impossibile, l'aggancio della fase.⁴

Per questo motivo negli anelli PLL in cui è richiesto che l'aggancio di fase sia possibile su un intervallo ampio di frequenze generate dal VCO, si utilizzano quasi esclusivamente comparatori fase-frequenza.

In fig. 17.21 è riportato un esempio di circuito molto comune per questo scopo.

I due segnali s_1 ed s_2 comandano ciascuno l'ingresso *clock* di un flip-flop di tipo D. Sul fronte di salita del segnale il flip-flop commuta portando l'uscita Q al livello alto. Quando entrambe le uscite Q_1 e Q_2 sono al livello alto, la porta NAND P_3 genera un livello basso che tramite P_4 e P_5 azzerà entrambi i flip-flop P_1 e P_2 , iniziando di nuovo il ciclo.

Se il segnale s_1 è in anticipo rispetto a s_2 si avranno una serie di impulsi di livello alto (v_h) all'uscita Q_1 della porta P_1 ; viceversa, se è il segnale s_2 ad essere in anticipo, sarà l'uscita Q'_2 a generare una serie di impulsi, questa volta di livello basso (v_l).

⁴Almeno in prima approssimazione. Effetti del secondo ordine possono consentire l'aggancio per valori di $\omega_1 - \omega_2$ sufficientemente piccoli.

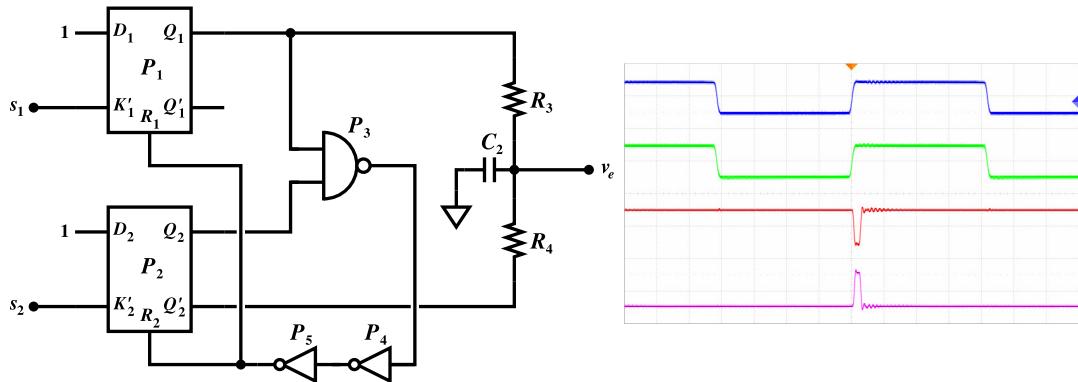


Figura 17.21: Schema elettrico di un discriminatore fase-frequenza. A destra: oscillogramma dei segnali s_1 (blu), s_2 (verde), Q'_2 (rosso), Q_1 (violetto). I due segnali s_1 ed s_2 sono isocroni ed in fase, con frequenza 1.2 MHz . Scala orizzontale $1 \mu\text{s}/\text{divisione}$, scala verticale $5 \text{ V}/\text{divisione}$.

Quando i flip-flop sono entrambi a riposo, l'uscita Q_1 è bassa, l'uscita Q'_2 è alta, ed il valore di v_e è al livello $(v_h + v_l)/2$.

Se i segnali s_1 e s_2 non hanno la stessa frequenza, viene generata all'uscita Q_1 oppure Q'_2 una serie di impulsi, positivi o negativi a seconda di quale delle due frequenze sia maggiore. Questi impulsi non sono tutti eguali, ma variano in durata con un ciclo corrispondente alla frequenza di battimento tra s_1 ed s_2 , come illustrato in fig. 17.22a. Nella fig. 17.22b invece viene riportata la tensione generata sul nodo v_e da questo treno di impulsi attraverso il circuito passa-basso $R_{3,4}C_2$.

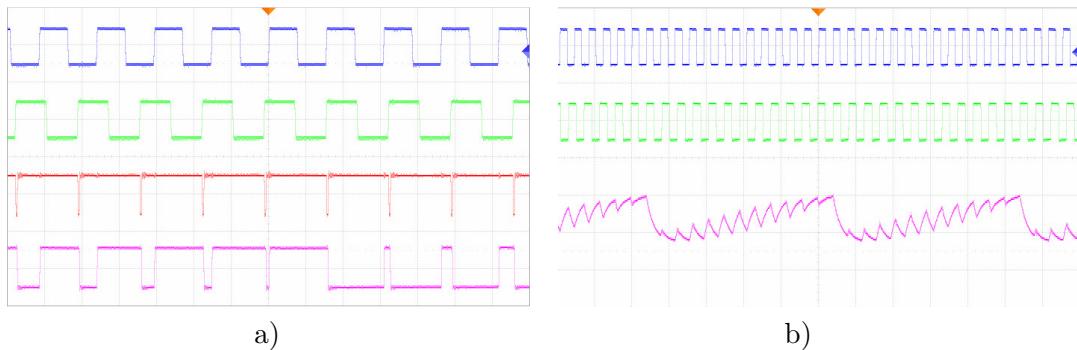


Figura 17.22: Uscite del discriminatore fase-frequenza con segnali s_1 (blu) ed s_2 (verde) con frequenze rispettivamente di 1.3 e 1.2 MHz . A sinistra, in violetto, l'uscita Q_1 , con una serie di impulsi positivi di durata via via crescente, fino a diventare pari all'intero periodo e ripartire da zero. In rosso l'uscita Q'_2 . A destra, in b), in violetto è riportata la tensione sul nodo v_e . La scala verticale è di $5 \text{ V}/\text{divisione}$ per tutti i segnali escluso il violetto in b), la cui scala è di $2 \text{ V}/\text{divisione}$. La scala orizzontale è di $0.5 \mu\text{s}/\text{divisione}$ in a) e di $2 \mu\text{s}/\text{divisione}$ in b).

Se i due segnali sono esattamente sincroni come fase e come frequenza, i due flip-flop commutano contemporaneamente e vengono immediatamente azzerati. Poiché *esattamente* ed *immediatamente* sono parole che non fanno parte del mondo fisico,

le due porte P_4 e P_5 ritardano leggermente il segnale generato da P_3 in modo che in condizione di sincronismo di fase e frequenza sia P_1 che P_2 generino un breve impulso che si sovrappone e si elide sul nodo v_e (vedi fig. 17.21 a destra). Senza questo accorgimento ci potrebbe essere il rischio di avere una zona morta per valori molto piccoli nella differenza di fase tra s_1 ed s_2 , in cui nessuno dei due impulsi, positivo e negativo, viene generato. Il nodo v_e verrebbe quindi abbandonato a se stesso fino a che una deriva nella fase non porti il circuito ad uscire dalla zona morta.

Nell'esempio descritto questo rischio è quasi solo teorico, ma sarebbe assai più reale se dopo le uscite Q_1 e Q'_2 ci fossero altri circuiti prima di arrivare al nodo v_e . La soluzione di fig. 17.21 è sicuramente semplice e veloce, ma è anche leggermente *sporca*, perché le correnti attraverso le resistenze R_3 ed R_4 , che generano il segnale di errore v_e , sono prodotte da tensioni di circuiti digitali, quindi facilmente affette da disturbi. Una soluzione più pulita ma anche più complessa e più lenta sarebbe di utilizzare i segnali digitali per comandare due circuiti di commutazione analogica.

In ogni caso è necessaria la presenza del condensatore C_2 per filtrare le componenti veloci degli impulsi di corrente che arrivano attraverso R_3 ed R_4 . La costante di tempo data da C_2 insieme con R_3 ed R_4 (eventualmente in parallelo ad R_2) è quella che determina la frequenza di taglio F_3 ed in definitiva la risposta complessiva dell'anello. Anche qui è necessario un compromesso tra velocità di risposta che si desidera più alta possibile e oscillazioni residue nel segnale v_e che, attraverso il circuito amplificatore-integratore, vanno a modulare la frequenza del VCO e generano rumore di fase.

17.7.4 Aggancio di fase (*lock*)

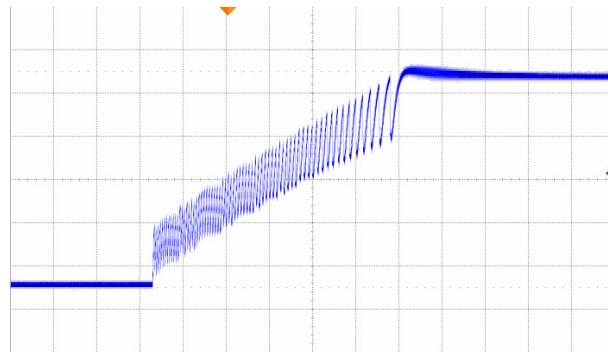


Figura 17.23: Tensione v_c di pilotaggio del VCO a seguito del passaggio della frequenza f_r da 1.5 a 2.0 MHz. La scala orizzontale è di 50 μs /divisione. Il tempo complessivo necessario al PLL per effettuare l'*aggancio* alla nuova frequenza è di circa 0.5 ms.

Nella fig. 17.23 è riportato l'andamento della tensione v_c all'uscita dell'amplificatore-integratore di un circuito PLL con il guadagno d'anello della fig. 17.18 e sollecitato ad effettuare un salto di frequenza f facendo variare da 1.5 MHz a 2 MHz la frequenza f_r .

In una prima fase, subito dopo la variazione di f_r , f_r ed f_n non coincidono più ed all'ingresso dell'amplificatore-integratore è presente un segnale del tipo di quello in fig. 17.22-b (curva in violetto). Questo produce la rampa oscillante nel passaggio

da un livello di tensione di pilotaggio del VCO ad un altro. Una volta che f_n ha raggiunto nuovamente la frequenza f_r , l'oscillazione cessa e si può vedere una leggera gobba (*overshoot*) in v_c , corrispondente all'aggiustamento della fase. Dopo di che il PLL è agganciato (*locked*) e la tensione di pilotaggio del VCO rimane costante.

Capitolo 18

Analisi numerica delle reti

Nei capitoli precedenti sono stati presentati più volte grafici per illustrare l'andamento di tensioni e correnti nei circuiti. Questi dati sono stati ottenuti o da misure di laboratorio oppure risolvendo le equazioni, algebriche e differenziali, lineari e non lineari, derivate applicando ai circuiti le leggi di Kirchhoff e le equazioni caratteristiche dei componenti.

Il metodo di gran lunga più efficace per svolgere questi calcoli oggi è l'utilizzo di un calcolatore elettronico ed un programma di calcolo idoneo. Il più noto e diffuso di questi è il programma *Spice*, sviluppato presso la *University of California*, a Berkeley negli anni tra il 1970 ed il 2000. Di questo programma ne sono state sviluppate numerose versioni, sia commerciali che distribuite con licenza *open source*, come ad esempio *spice3f5s* (ultima versione sviluppata a Berkeley) e *ngspice* (parte del progetto *gEDA: GPL Electronic Design Automation*).

Quelle che seguono sono soltanto le istruzioni principali per l'utilizzo dei programmi *spice3f5s* e *ngspice*, derivate per lo più dalla traduzione del manuale originale dello *Spice* di Berkeley. Per una descrizione più approfondita si rimanda ai manuali completi.

18.1 Analisi per nodi di una rete

Spice utilizza l'analisi per nodi modificata di una rete, citata nel par. 15.1, che risulta particolarmente conveniente per descrivere un circuito elettronico in una forma facilmente utilizzabile da un programma di calcolo.

Una volta numerati i nodi del circuito, la descrizione consiste nel costruire un *file* contenente l'elenco dei componenti (resistenze, condensatori, induttanze, diodi, transistor ecc.) con le loro caratteristiche e l'indicazione dei due (o più) nodi a cui ciascun componente è connesso. Da queste informazioni il programma di calcolo costruisce le matrici descritte nel cap. 15, algebriche per il calcolo lineare, di operatori integro-differenziali per il calcolo non lineare, e provvede ad effettuare i calcoli necessari. Il *file* di descrizione del circuito può essere costruito in due modi:

- Utilizzando un *editor* di testo (ad esempio, *emacs* o *gedit*) per comporre manualmente il *file* con tutte le informazioni necessarie.

- Utilizzando un programma di interfaccia grafica per disegnare uno schema del circuito con le indicazioni delle caratteristiche dei componenti, lasciando poi allo stesso programma il compito di generare il *file* di testo con la descrizione del circuito.

L'utilizzo di uno schema grafico è sicuramente vantaggioso, se non indispensabile, quando si ha a che fare con circuiti complessi; nel caso di circuiti semplici, come quasi tutti quelli riportati in questi appunti, la composizione diretta del file di testo è di gran lunga più veloce e conveniente.

Il programma *Spice* è stato rilasciato fin dall'inizio con una licenza di uso che equivaleva in buona parte a quello che ora chiamiamo *open source*. Questo ha consentito una ampia adozione e diffusione del programma e dei suoi formati da parte di utenti sia scientifici che industriali. Il formato dei *file* di descrizione dei circuiti è rimasto praticamente lo stesso per tutti i programmi derivati da *Spice*, per cui un *file* può essere portato da una versione all'altra di *Spice* con minime o anche nessuna variazione.

18.2 Composizione del *file* di descrizione del circuito

Il *file* è composto di *righe*; ogni riga contiene la descrizione di un componente e la sua collocazione nel circuito, oppure un comando relativo ai calcoli da effettuare.

La prima riga è sempre il *titolo* ed il suo contenuto è ignorato dal programma. Se la prima riga viene utilizzata per la descrizione di un componente, questo componente non comparirà nel circuito, generando errori a volte apparentemente incomprensibili.

Ogni riga contenente un '*' come primo carattere è considerata un commento ed il suo contenuto viene ignorato (è destinato alla lettura da parte degli *umani* e non al programma di calcolo).

Le righe vuote sono ignorate e possono essere inserite liberamente per migliorare la leggibilità del *file*.

Ogni riga contenente un '+' come primo carattere viene considerata la continuazione della riga precedente. In questo modo un comando molto lungo può essere spezzato in più linee di testo.

Ogni componente del circuito è descritto da una riga che contiene il nome del componente, i nodi a cui è connesso ed i valori dei parametri che ne determinano le caratteristiche elettriche. Il nome di un componente deve essere una stringa lunga da due a otto caratteri alfanumerici. Il primo carattere del nome indica il tipo di componente.

I nomi dei componenti, dei nodi e dei parametri ed i valori numerici non possono contenere spazi. Uno o più spazi invece separano tra loro i vari elementi della riga.

18.3 Nodi

I nodi del circuito sono identificati da stringhe alfanumeriche arbitrarie. Ad esempio possono essere numeri (1, 2, 3, 4, 5, ...), oppure nomi (`in`, `out`, `comando`, `sig1`, `sig2`, ...).

Degli N nodi che costituiscono la rete ne deve essere scelto uno, arbitrariamente, sulla base della convenienza, come riferimento. Questo sarà il nodo di massa e sarà

sempre indicato come nodo 0; convenzionalmente avrà potenziale elettrico 0 V. Il potenziale di ognuno degli altri $N - 1$ nodi sarà misurato rispetto al potenziale del nodo 0.

Poichè i nodi sono identificati da stringhe e non da numeri, le stringhe `zero`, `00`, `000`, ecc. indicheranno tutte nodi diversi tra loro e diversi dal nodo di riferimento 0.

18.4 Valori numerici

Ogni valore numerico che compare nella descrizione dei componenti del circuito viene implicitamente considerato espresso nelle unità opportune del sistema MKSA: tensioni in volt, correnti in ampere, resistenze in ohm ecc. Un valore numerico può essere un numero intero (12, -135) oppure un numero decimale (3.1415), eventualmente seguiti da un esponente intero (`1e-12`, `4.7e3`).

È possibile anche indicare multipli e sottomultipli utilizzando i suffissi:

<code>m</code>	10^{-3}	milli	<code>k</code>	10^3	kilo
<code>u</code>	10^{-6}	micro	<code>Meg</code>	10^6	mega
<code>n</code>	10^{-9}	nano	<code>G</code>	10^9	giga
<code>p</code>	10^{-12}	pico	<code>T</code>	10^{12}	tera
<code>f</code>	10^{-15}	femto			

ed il suffisso fuori sistema

$$\text{mil } 25.4 \cdot 10^{-6}$$

che equivale, come lunghezza, ad un millesimo di pollice. Tra il valore numerico ed il suffisso non devono essere lasciati spazi. Altri caratteri che seguono il suffisso sono ignorati.

18.5 Componenti passivi semplici

18.5.1 Resistenze, condensatori ed induttanze

I nomi di questi componenti hanno come primo carattere `R` per la resistenza, `C` per il condensatore ed `L` per l'induttanza (`L` da *Linkage*):

```
Rxxxxxx N1 N2 val
Cxxxxxx N1 N2 val [ic=volt]
Lxxxxxx N1 N2 val [ic=ampere]
```

`N1` ed `N2` sono i due nodi tra cui il componente è connesso; `val` è il valore del componente (le opzioni `[ic=volt]` e `[ic=ampere]` sono descritte nel par. 18.8.4). Ad esempio,

```
R1 1 2 120
Cin input 8 1u
L12 3 7 1m
```

indicano rispettivamente una resistenza da 120Ω collegata tra i nodi 1 e 2, un condensatore da $1 \mu F$ tra i nodi `input` ed 8 ed una induttanza da $1 mH$ tra i nodi 3 e 7. Poichè resistenze, condensatori ed induttanze sono dispositivi simmetrici, l'ordine dei due nodi `N1` ed `N2` è irrilevante.

18.5.2 Mutua induttanza

Tra due induttanze può essere indicato un coefficiente di accoppiamento:

Kxxxxxx L1 L2 val

L1 ed L2 sono le due induttanze; **val** è il valore del coefficiente di accoppiamento magnetico. Se L_1 ed L_2 sono i valori delle due induttanze, il coefficiente di accoppiamento k è definito dalla relazione $M = k\sqrt{L_1 L_2}$ dove M è la mutua induttanza tra L1 ed L2.

18.6 Generatori indipendenti

La definizione dei generatori è spesso causa di disorientamento tra gli utenti di Spice. Infatti il ruolo di un generatore cambia a seconda che il circuito venga analizzato come rete lineare in funzione della frequenza oppure integrando le equazioni integro-differenziali, sia lineari che non lineari, in funzione del tempo.

18.6.1 Generatori di tensioni e correnti continue

Vxxxxxx N+ N- [[DC] val]¹
Ixxxxxx N+ N- [[DC] val]

Il primo è un generatore ideale di tensione continua, il secondo un generatore ideale di corrente continua. Il primo nodo **N+** è il polo positivo del generatore, il secondo, **N-**, il polo negativo.

L'indicazione **DC** è opzionale, come anche il valore **val**; in assenza di indicazione diversa **val** è considerato zero. Un generatore di tensione nulla, come si vedrà, è un *amperometro*. Il verso positivo per una corrente che fluisce attraverso un generatore di tensione è quando la corrente entra nel nodo **N+** del generatore e ne esce dal nodo **N-**.

18.6.2 Generatori AC per analisi lineare

Aggiungendo optionalmente alla definizione di un generatore l'indicazione
AC [ampiezza [fase]]

si specifica che il generatore in questione fornisce, durante l'analisi lineare in funzione della frequenza, un segnale dell'ampiezza di 1 V, oppure del valore indicato, con fase 0, oppure con il valore di fase indicato.

La frequenza non è specificata, perché non è una caratteristica propria del generatore, ma è la variabile indipendente dell'analisi.

18.6.3 Generatori per analisi di transitorio

Come si fa nella pratica di laboratorio, è possibile analizzare la risposta di un circuito ad un segnale che sia una certa funzione del tempo (ad esempio, onda sinusoidale, onda quadra, triangolare, segnale a gradino, impulso ecc.). Le forme d'onda

¹Le parti di un comando racchiuse tra parentesi [...] sono opzionali.

usate più di frequente nell'analisi di transitorio con *Spice* sono l'impulso ed il segnale sinusoidale, che possono essere specificati aggiungendo alla definizione di un generatore V o I le caratteristiche:

Impulso

PULSE (V1 V2 td tr tf Pw per)

Il segnale generato è raffigurato in fig. 18.1.

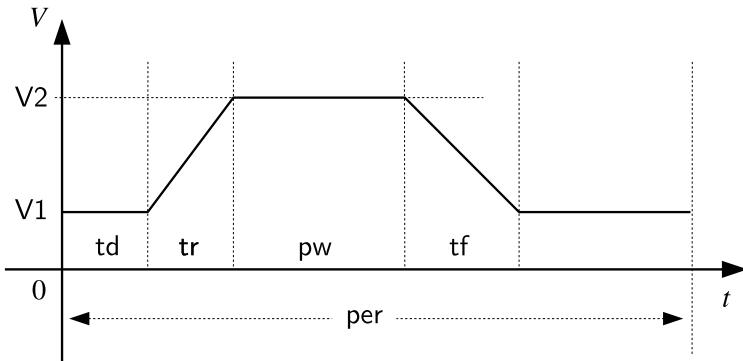


Figura 18.1: Impulso.

Quindi, ad esempio, un'onda quadra con periodo 1 ms, *duty cycle* 50% e livelli 0 e 1 V sarà descritta da

PULSE (0 1 0.5m 0 0 0.5m 1m)

Segnale sinusoidale

SIN (Vo Va f td theta)

Il segnale generato è descritto dall'equazione:

$$\begin{aligned} t < t_d : \quad v(t) &= V_0 \\ t > t_d : \quad v(t) &= V_0 + V_a e^{-\frac{(t-t_d)}{\theta}} \cdot \sin [2\pi f(t + t_d)] \end{aligned}$$

e corrisponde quindi ad una oscillazione sinusoidale smorzata ($\theta > 0$), crescente ($\theta < 0$) oppure stabile ($\theta = 0$).

Le caratteristiche DC, AC, PULSE() e SIN() possono convivere tutte contemporaneamente nella definizione di uno stesso generatore. Verranno utilizzate od ignorate dal programma in base al contesto del calcolo effettuato.

18.7 Circuiti indefiniti o impossibili

Ci sono alcuni circuiti che noi trattiamo abitualmente assumendo implicitamente ipotesi che in realtà sono arbitrarie. Ad esempio, con due condensatori in serie conoscendo il potenziale dei nodi agli estremi della serie non è possibile determinare

il potenziale del nodo centrale a meno di non assumere una ipotesi sulle condizioni iniziali, ad esempio che si parta con i condensatori entrambi scarichi. *Spice* non utilizza automaticamente ipotesi di questo tipo. Ogni nodo di un circuito deve avere un percorso elettrico in corrente continua che lo colleghi a massa oppure il potenziale di quel nodo deve essere definito esplicitamente.

Un problema simile è dato da due induttanze ideali in parallelo: avendo entrambe resistenza nulla, il calcolo della ripartizione della corrente continua darebbe luogo ad una divisone 0/0.

Altre situazioni inaccettabili per *Spice* sono due generatori di tensione (ideali) in parallelo, due di corrente in serie, oppure il calcolo del punto di lavoro con una induttanza in parallelo ad un generatore di tensione oppure un condensatore in serie ad un generatore di corrente.

Tutte le volte che una di queste situazioni si verifica, *Spice* interrompe l'esecuzione del calcolo segnalando un errore, per esempio lamentando che la matrice del sistema di equazioni che descrive il circuito è singolare e quindi non può essere invertita.

18.8 Analisi dei circuiti

Il file contenente la descrizione del circuito deve essere completato con le istruzioni relative ai calcoli da effettuare ed ai risultati da presentare.

Queste istruzioni possono essere inserite in sequenze di controllo in qualunque punto del *file*, prima del circuito, dopo, o anche a metà, ma devono essere separate logicamente dalla descrizione racchiudendole tra due righe che segnalano l'inizio e la fine della sequenza di controllo:

```
.control
.
.
.
.endc
```

All'interno della sequenza di controllo non ci possono essere righe di descrizione del circuito. Le principali analisi che è possibile effettuare con *Spice* sono:

op	determinazione del punto di lavoro
dc	funzione di trasferimento in continua
ac	funzione di trasferimento in alternata
tran	analisi di transitorio

18.8.1 Determinazione del punto di lavoro

op

Viene determinato il punto di lavoro del circuito in corrente continua; ogni induttanza è considerata un cortocircuito, ogni condensatore un circuito aperto. Se il circuito contiene elementi non lineari, come diodi o transistor, il sistema di equazioni viene risolto con procedimento iterativo per approssimazioni successive.

Una analisi **op** viene eseguita automaticamente prima di una analisi di tipo **tran** per determinare le condizioni iniziali da cui far partire l'integrazione delle equazioni differenziali e prima di una analisi di tipo **ac** per sostituire agli elementi non lineari

presenti nel circuito il loro equivalente *linearizzato* nell'intorno del punto di lavoro, analogamente a quanto visto nei par. 4.13 e 6.4.

18.8.2 Funzione di trasferimento in continua

```
dc gen start stop incr [gen2 start2 stop2 incr2]
```

Questa analisi equivale a ripetere l'operazione **op** del paragrafo precedente assegnando successivamente al generatore di tensione o di corrente **gen** i valori da **start** a **stop** a passi **incr**. Opzionalmente può essere indicato un secondo generatore **gen2**, la cui tensione o corrente viene fatta variare da **start2** a **stop2** a passi **incr2**. La scansione del primo generatore **gen** viene ripetuta per ognuno dei valori del secondo generatore **gen2**.

18.8.3 Funzione di trasferimento in alternata

```
ac (dec | oct | lin)2 Np fstart fstop
```

La funzione di trasferimento in corrente alternata del circuito, lineare o linearizzato da una operazione **op** implicita, può essere determinata facendo una scansione in frequenza su scala logaritmica per decadi (**dec**), per ottave (**oct**), o su scala lineare (**lin**). I due parametri **fstart** ed **fstop** indicano sempre la frequenza di inizio e fine della scansione; nel caso **lin** il parametro **Np** indica il numero totale di punti della scansione; nei casi **dec** e **oct** indica il numero di punti rispettivamente per ogni decade o per ogni ottava di frequenza.

18.8.4 Analisi di transitorio

```
tran tstep tstop [tstart [tmax]] [uic]
```

Le equazioni differenziali lineari o non lineari del circuito vengono integrate numericamente nel tempo da 0 fino al tempo **tstop** ed i risultati vengono riportati in funzione del tempo con un intervallo **tstep** tra ogni punto ed il successivo. La quantità **tstep** viene utilizzata come passo di integrazione, a meno che l'algoritmo numerico sulla base dell'andamento delle funzioni non valuti l'opportunità di utilizzare un passo di integrazione più piccolo (ma non più grande). I risultati dell'integrazione saranno riportati a partire dal tempo 0 fino a **tstop**, oppure a partire da **tstart** nel caso che opzionalmente questo parametro sia specificato. L'ulteriore parametro opzionale **tmax** permette di stabilire un limite superiore per il passo di integrazione, minore di **tstep**.

Come condizioni iniziali per l'integrazione, al tempo zero, viene preso automaticamente il punto di lavoro, a meno che non sia specificata l'opzione **uic**. In questo caso per ogni condensatore e per ogni induttanza il valore della tensione a $t = 0$ sarà quello (opzionalmente) indicato con **uic=volt** o con **uic=ampere** nella definizione del componente.

Al termine di una delle analisi descritte i risultati possono essere stampati o presentati in forma grafica mediante i due comandi **print** e **plot**.

²La barra verticale ' | ' indica la presenza di più alternative possibili in un comando.

18.9 Studio di un circuito risonante

Nella fig. 18.2 è riportato lo schema elettrico di un circuito risonante *RLC* serie, alimentato da un generatore di tensione; a destra è riportato il *file* contenente la descrizione del circuito e le istruzioni per l'analisi con *Spice* sia nel dominio del tempo che nel dominio della frequenza.

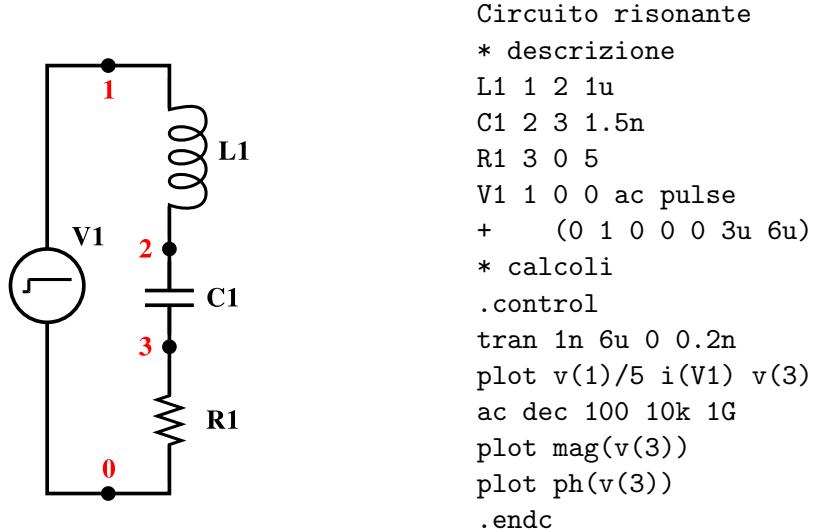


Figura 18.2: Circuito risonante serie e *file* di descrizione ed elaborazione in formato *Spice* per il calcolo della risposta in funzione della frequenza e del tempo. I quattro numeri in rosso identificano i nodi del circuito.

La prima riga del file (**Circuito risonante**), come visto, è il titolo; la seconda un commento; seguono poi quattro righe, di cui l'ultima divisa in due, che descrivono i quattro componenti del circuito (induttanza, capacità, resistenza e generatore).

Il generatore **V1** appare con una doppia personalità: come generatore **ac**, con i valori predefiniti di ampiezza e fase (1 *V*, 0°), per l'analisi lineare in funzione della frequenza, e come generatore di onda quadra con periodo 6 μs (frequenza 167 *kH*z) e livelli di tensione 0 e 1 *V* per l'analisi di transitorio.

Dopo un'altra riga di commento (*** calcoli**), le due righe **.control** e **.endc** racchiudono i comandi di elaborazione e presentazione dei risultati.

Il primo comando (**tran 1n 6u 0 0.2n**) richiede una analisi di transitorio, cioè l'integrazione in funzione del tempo delle equazioni differenziali del circuito, in questo caso lineari. L'integrazione sarà effettuata tra il tempo 0 ed il tempo 6 μs , con un passo di integrazione non superiore a 0.2 *ns* ed i risultati saranno generati ad intervalli di 1 *ns*, quindi per un totale di 6001 punti. I risultati saranno vettori contenenti i valori reali delle tensioni per ogni nodo del circuito e delle correnti attraverso ogni generatore, in funzione del tempo. Il successivo comando **plot v(1)/5 i(V1) v(3)** richiede al programma di produrre un grafico con tre funzioni. La funzione **v()** ha come argomento l'indicazione di un nodo, quindi **v(1)** produrrà il grafico della

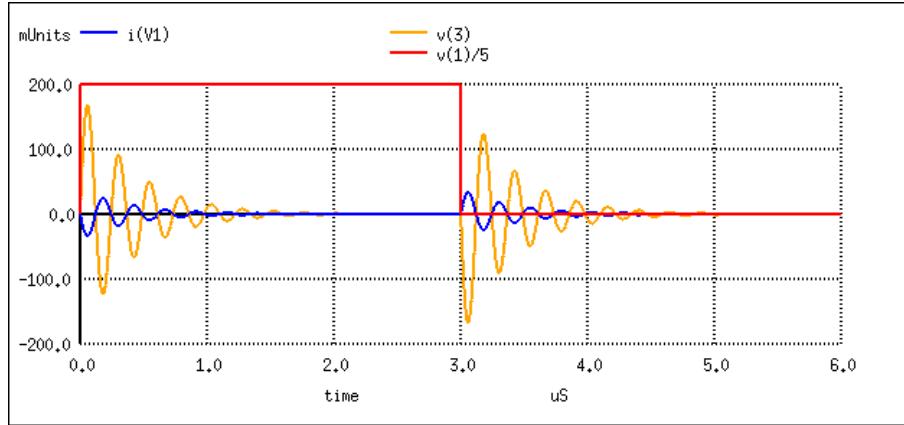


Figura 18.3: Transitorio nel circuito risonante di fig. 18.2. Curva rossa: onda quadra prodotta dal generatore V1; curva blu: corrente attraverso il generatore V1 (di segno opposto alla corrente della maglia); curva gialla: tensione sul nodo V3.

tensione sul nodo 1, $v(3)$ la tensione sul nodo 3, in funzione del tempo. La funzione $i()$ ha per argomento il nome di un generatore di tensione. Quindi $i(V1)$ produrrà il grafico della corrente attraverso il generatore V1 (ed attraverso la serie RLC). In questo senso, i generatori in *Spice* sono anche gli *amperometri*: un generatore di tensione nulla può essere inserito in un ramo per conoscere la corrente che vi circola, senza alterare il circuito.

Il grafico con le tre funzioni è riprodotto in fig. 18.3. Per ragioni di opportunità di scala, la tensione del generatore è stata divisa nel grafico per un fattore 5: da 1 V a 200 mV. Sull'asse x il programma riporta le unità di misura utilizzate: μs . Sull'asse y devono essere riportati i valori sia di tensioni che di correnti, quindi viene indicato solo il valore numerico ed il fattore di scala (`mUnits`); le unità di misura saranno quelle coerenti con le grandezze, quindi mV per le tensioni, mA per le correnti.

Nel grafico si può osservare come la tensione sul nodo 3 (curva gialla), proporzionale alla corrente che scorre in $R1$, appare con polarità opposta rispetto alla corrente $i(V1)$ (curva blu). Questo è dovuto alla convenzione vista nel par. 18.6.1: le correnti nei generatori di tensione sono positive quando scorrono attraverso il generatore dal nodo N+ al nodo N-, quindi con segno opposto rispetto alla corrente della maglia.

Dopo l'analisi di transitorio viene richiesta una analisi lineare di risposta in corrente alternata, calcolando tensioni e correnti (complesse) nel circuito, in funzione della frequenza, da 10 kHz a 1 GHz prendendo per ogni decade di frequenza 100 punti spaziati in maniera logaritmica (`ac dec 100 10k 1G`).

Poichè in questo calcolo tensioni e correnti sono quantità complesse della frequenza, per conoscere la tensione sul nodo 3 è necessario richiedere due grafici, uno per il modulo, tramite la funzione `mag()`, l'altro per la fase, tramite la funzione `ph()` (fig. 18.4). In alternativa si sarebbe potuto utilizzare di nuovo un solo grafico, come fatto nella figura precedente, scalando una delle due quantità per una costante opportuna.

L'esempio riportato in questo paragrafo ha solo uno scopo dimostrativo del meccanismo di funzionamento e di utilizzo di *Spice*, perchè tutti i risultati si sarebbero

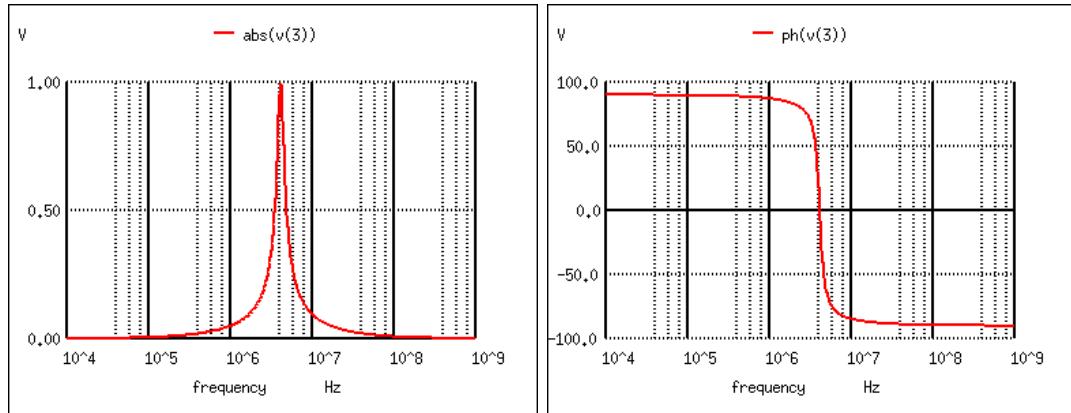


Figura 18.4: Modulo e fase in funzione della frequenza della tensione ai capi della resistenza nel circuito risonante di fig. 18.2. Nel grafico a destra la fase è riportata in gradi, ma viene indicata (erroneamente) l'unità di misura della grandezza di origine, cioè di una tensione. Questo non può comunque essere fonte di confusione, perché in *Spice* ogni grandezza è sempre espressa nelle sue unità naturali.

potuti ottenere direttamente per via analitica.

Il maggior interesse per *Spice* deriva invece dalla possibilità di risolvere circuiti ben più complessi dell'esempio di fig. 18.2 e soprattutto di trattare circuiti contenenti dispositivi a semiconduttore (diodi e transistor) intrinsecamente non lineari e quindi affrontabili praticamente solo per via numerica.

18.10 Presentazione dei risultati

Come visto nell'esempio precedente, il comando `plot` permette di produrre un grafico dei vettori di dati ottenuti tramite una delle analisi `dc`, `ac` e `tran`.

Nella forma più semplice il comando

```
plot expr1 expr2 expr3 ... exprN
```

genera un grafico delle N espressioni elencate. Queste espressioni possono essere, come si è visto nell'esempio, la semplice indicazione dei vettori di dati [come `v(3)`] o loro funzioni [come `v(1)/5` e `mag(v(3))`] applicate a ciascun elemento del vettore.

La variabile sull'asse x del grafico è automaticamente la stessa su cui è stata fatta la scansione, la frequenza nel caso di analisi `ac`, il tempo nel caso di analisi `tran`. I valori di queste variabili sono contenuti nei due vettori `frequency` e `time` rispettivamente.

Oltre alle funzioni già note `mag()` e `ph()`, nelle espressioni è possibile inserire numerose altre funzioni matematiche, di cui le più comuni sono:

<code>real()</code>	parte reale di una quantità complessa
<code>imag()</code>	parte immaginaria di una quantità complessa
<code>j()</code>	$j(a)$ equivale a $j * a$, dove j è l'unità immaginaria
<code>db()</code>	$db(a) = 20 \log_{10} a $
<code>log()</code>	logaritmo (base 10)
<code>ln()</code>	logaritmo (base e)
<code>exp()</code>	esponenziale
<code>abs()</code>	valore assoluto
<code>sqrt()</code>	radice quadrata
<code>sin()</code>	seno
<code>cos()</code>	coseno
<code>tan()</code>	tangente
<code>atan()</code>	arcotangente
<code>deriv()</code>	derivata numerica dei dati contenuti nel vettore

I risultati dei calcoli possono anche essere scritti su terminale, o meglio in un *file*, per altre elaborazioni. A questo scopo si può utilizzare il comando

```
print expr1 expr2 expr3 ... exprN [> file]
```

L'uscita dei dati avviene sul terminale a meno che non sia presente il reindirizzamento su *file*. Nei manuali si può trovare ogni altro dettaglio su questi due comandi.

18.11 Semiconduttori

I semiconduttori sono dispositivi troppo complessi per poter essere riassunti in una sola riga, come si può fare con una resistenza od un condensatore. Nei capitoli precedenti si sono visti i meccanismi fisici di funzionamento di questi dispositivi ed alcune delle equazioni che li descrivono, ad esempio l'equazione della giunzione per i diodi (eq. 4.7) e le equazioni di Ebers-Moll per i transistor *bjt* (eq. 7.1-7.2). Quello che differenzia un dispositivo da un altro sono i valori dei parametri che compaiono nelle equazioni. Per una descrizione sommaria che tenga conto soltanto delle caratteristiche principali bastano pochi parametri; una descrizione più completa e raffinata, invece, che tenga conto anche degli effetti secondari può avere fino a 15 parametri per un diodo e fino a 41 per un transistor *bjt*.

Inoltre un certo diodo od un certo transistor possono comparire più volte in un circuito, e di certo non è pratico ripetere tante volte la stessa (lunga) descrizione.

Questi componenti sono descritti quindi in due fasi: una riga, analoga a quelle già viste per i componenti semplici, indica il nome del dispositivo, i nodi a cui è connesso ed il *modello*. Un secondo comando, spesso spezzato in più linee a causa della lunghezza, contiene la descrizione del *modello*, cioè i parametri che compaiono nelle equazioni del dispositivo.

18.11.1 Modelli

La descrizione di un modello è fatta tramite una riga `.model`, che può comparire in qualunque punto del file:

```
.model nome tipo (par1=val1 par2=val2 par3=val3 ... )
```

Il nome è una stringa alfanumerica a piacere, che spesso per comodità coincide direttamente con la sigla commerciale del dispositivo che si vuole descrivere. Il **tipo** è un codice che indica appunto che tipo di dispositivo è descritto dal modello. I codici dei tipi più comuni sono:

D	diodo a giunzione
NPN	transistor <i>bjt</i> di tipo <i>npn</i>
PNP	transistor <i>bjt</i> di tipo <i>pnp</i>
NJF	transistor <i>jfet</i> a canale <i>n</i>
PJF	transistor <i>jfet</i> a canale <i>p</i>
NMOS	transistor <i>mosfet</i> a canale <i>n</i>
PMOS	transistor <i>mosfet</i> a canale <i>p</i>

La lista di **par=val** racchiusa tra parentesi contiene il nome di ciascun parametro (**par**) a cui si vuole assegnare un valore (**val**). Per ogni parametro esiste comunque un valore predefinito, utilizzato ogni volta che non c'è un valore assegnato esplicitamente. Quindi si può anche non assegnare alcun valore e si otterrà così il diodo *standard* od il transistor *standard*.

18.11.2 Diodi a giunzione

Dxxxxxx Na Nk modello

La lettera **D** all'inizio del nome individua il componente *diodo*. Il primo nodo (**Na**) corrisponde all'anodo del diodo, il secondo (**Nk**) al catodo. Il diodo è polarizzato direttamente e quindi è in conduzione quando l'anodo è a potenziale maggiore del catodo.

Il **modello** indica il nome del modello che descrive il diodo. I seguenti modelli descrivono due diodi di uso generale, molto comuni:

```
.model 1n4148 D (is=0.1n rs=16 cjo=2p tt=12n bv=100)
.model 1n4007 D (is=1.1e-8 n=1.78 rs=0.041 cjo=28.1p vj=0.51
+ m=0.32 tt=9.85e-6 bv=1100 ibv=0.1)
```

Per il primo diodo vengono specificati solo i cinque parametri più significativi; per il secondo diodo ne vengono indicati nove, per una descrizione più raffinata. Per ogni parametro non indicato, come già detto, *Spice* utilizza un valore predefinito.

I parametri utilizzati dai modelli di *Spice* corrispondono in buona parte a quelli introdotti nel cap. 4 nella descrizione del funzionamento del diodo a giunzione:

- is** = corrente di saturazione inversa (I_S nelle eq. 4.6 e 4.7);
- rs** = resistenza ohmica in serie al diodo (R_D , par. 4.7);
- cjo** = capacità della giunzione a polarizzazione nulla (C_0 nelle eq. 4.43 e 4.44);
- tt** = *transit time* (corrisponde agli effetti di τ_p, τ_n del par. 4.6);
- bv** = tensione di *breakdown* (la V_{BR} del par. 4.11);
- ibv** = corrente inversa alla tensione di *breakdown*;
- n** = fattore di correzione η (par. 4.7);
- vj** = potenziale di giunzione (V_j , par. 4.3);
- m** = coefficiente del profilo di drogaggio della giunzione (0.5 per una giunzione a gradino, 0.33 per un profilo di drogaggio lineare).

Per gli altri i parametri non contenuti in questa tabella si rimanda al manuale.

18.11.3 Transistor *bjt*

Qxxxxxx Nc Nb Ne modello

Il transistor *bjt* è indicato dalla lettera **Q**. La polarità del transistor (*pnp* o *npn*) non compare in questa riga, ma è specificata nel modello dalla sigla **NPN** o **PNP** che compare dopo il nome. I tre terminali del transistor, collettore, base ed emettitore, sono connessi nell'ordine ai nodi **Nc**, **Nb** ed **Ne**. Un modello di transistor *npn* di uso generale è:

```
.model 2N2222A NPN (is=31f ise=7.5f isc=11p bf=220 br=4
+          ikf=0.52 ikr=0.24 ne=1.4 nc=1.4
+          vaf=104 var=28 rb=0.13 re=0.22 rc=0.12
+          cjc=9.12p vjc=0.41 mjc=0.35 cje=27p
+          tf=0.325n tr=100n)
```

Per descrivere il transistor *bjt*, *Spice* utilizza il modello di Gummel”-Poon, più completo (e complesso) del modello di Ebers-Moll visto nel par. 7.1, per cui alcuni parametri che compaiono nel **.model** precedente non hanno corrispondenza con quanto visto nel cap. 7. Nella tabella seguente i parametri definiti solo nel modello di Gummel”-Poon sono indicati con la sigla (GP).

is = valore di saturazione della corrente di trasporto (GP);
ise = corrente di saturazione della giunzione b-e (I_{ES} nelle eq. 7.1-7.2);
isc = corrente di saturazione della giunzione b-c (I_{CS} nelle eq. 7.1-7.2);
bf = β_F (eq. 7.8);
br = β_R (l'analogo di β_F con α_R invece di α_F);
ikf = parametro che esprime la dipendenza di β_F dalla corrente (GP);
ikr = parametro che esprime la dipendenza di β_R dalla corrente (GP);
ne = coefficiente η per la giunzione b-e (eq. 7.1-7.2);
nc = coefficiente η per la giunzione b-c (eq. 7.1-7.2);
vaf = tensione di Early (par. 7.5);
var = tensione di Early a collettore-emettitore invertiti;
rb = resistenza ohmica in serie alla base;
re = resistenza ohmica in serie all'emettitore;
rc = resistenza ohmica in serie al collettore;
cjc = capacità di transizione base-collettore a polarizzazione nulla (par. 16.1);
vjc = potenziale della giunzione base-collettore (V_j , eq. 4.4);
mjc = fattore esponenziale della giunzione base-collettore (GP);
cje = capacità di transizione base-emettitore a polarizzazione nulla (par. 16.1);
tf = $1/(2\pi f_T)$ (par. 16.4);
tr = analogo di **tf** a collettore-emettitore invertiti;

Omettendo nella descrizione del **2N2222A** i parametri contrassegnati con (GP), specifici del modello di Gummel”-Poon, si ottiene la descrizione meno accurata fornita dal modello di Ebers-Moll.

18.11.4 Transistor *fet*

In modo simile ai *bjt* sono definiti i transistor *jfet*:

Jxxxxxx Nd Ng Ns modello

e *mosfet*

Mxxxxxx Nd Ng Ns Nb modello

I nodi Nd, Ng ed Ns corrispondono rispettivamente a *drain*, *gate* e *source*. Nel caso del *mosfet* è necessario indicare anche il nodo Ns a cui è connesso il substrato. Un esempio di transistor *j fet* a canale n (NJF) è:

```
.model 2N3819 NJF (beta=1.57m lambda=4m vto=-1.95
+ cgd=1.5p cgs=1.5p is=170f)
```

con i seguenti significati per i parametri:

$\text{beta} = I_{DSS}/V_P^2$ (eq. 9.10);
 $\lambda = \lambda$ (eq. 9.15);
 $V_P = V_P$ (eq. 9.10);
 c_{gd} = capacità *gate-drain* (par. 16.4);
 c_{gs} = capacità *gate-source* (par. 16.4);
 i_s = corrente di saturazione inversa della giunzione *gate-canale*;

18.12 Esempi di circuiti con semiconduttori

Gli esempi seguenti riportano gli schemi elettrici, i file di descrizione di *Spice* ed i risultati dei calcoli per alcuni dei circuiti utilizzati in questi appunti.

18.12.1 Curve caratteristiche di un transistor *bjt*

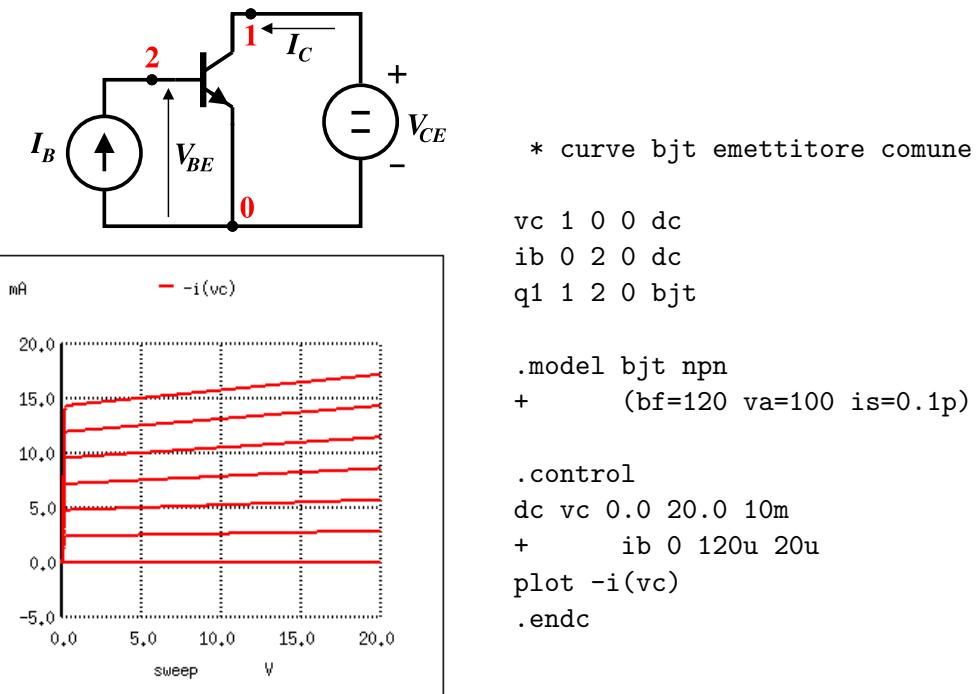


Figura 18.5: Determinazione della famiglia di curve caratteristiche per un transistor *bjt* ad emettitore comune.

Nell'esempio in fig. 18.5 viene fatta una doppia scansione in corrente continua: il generatore **vc** genera la tensione V_{CE} da 0 a 20 V a passi di 10 mV per ogni valore della corrente di base determinata dal generatore **ib** da 0 a $120\mu A$ a passi di $20\mu A$. Il circuito è quello utilizzato nel cap. 7 per disegnare le curve caratteristiche del transistor *bjt*. Per il transistor sono stati indicati soltanto alcuni dei parametri *statici*; questo modello non sarebbe assolutamente valido per lo studio in funzione del tempo o della frequenza.

18.12.2 Tempi di commutazione di un diodo

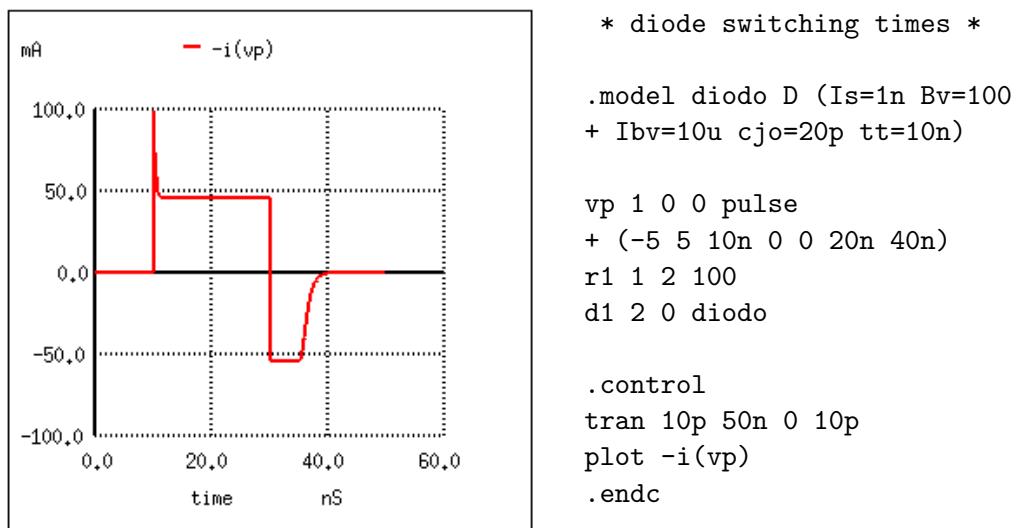


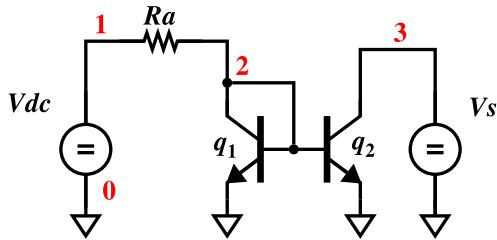
Figura 18.6: Correnti calcolate durante le commutazioni di un diodo e descrizione del circuito utilizzato.

Il circuito reale per la misura dei tempi di commutazione di un diodo è quello riportato in fig. 5.1. Se non si ha interesse a valutare i tempi di ritardo tra generatore ed oscilloscopio e le due linee di trasmissione L_1 ed L_2 hanno perdite trascurabili, allora possono essere omesse e si può tornare allo schema di principio di fig. 4.25. Avendo interesse a conoscere solo la corrente attraverso il diodo, le due resistenze da 50Ω possono essere raccolte in una sola da 100Ω , ottenendo il circuito molto semplice descritto dal file riportato in fig. 18.6 insieme al risultato del calcolo. Il modello del diodo deve avere ovviamente una capacità di transizione c_{jo} ed un tempo di transito tt diversi da zero, altrimenti si ha il diodo ideale a risposta istantanea di fig. 4.26b. Il segnale prodotto dal generatore **vp** è quello riportato in fig. 4.26a.

18.12.3 Circuito *current mirror*

Il circuito *current mirror* descritto nel par. 7.10 è stato studiato utilizzando il circuito riportato in fig. 18.7 ed il file di descrizione e comandi in fig. 18.8.

La colonna di sinistra del file contiene la descrizione del circuito e la sequenza di comandi per il calcolo, la colonna di destra contiene una accurata descrizione dei transistor utilizzati nell'esempio, una coppia dell'*array monolitico CA3046* contenente 7 transistor identici.

Figura 18.7: Circuito *current mirror*.

Nel circuito sono presenti due generatori: V_{dc} (30 V), che fornisce l'alimentazione; V_s (0 ... 25 V), che invece è il generatore con cui viene effettuata la scansione della tensione al collettore del secondo transistor e con cui è misurata la corrente.

```
**** Current mirror ****
Vdc 1 0 dc 30
Vs 3 0 dc 15 ac
Ra 1 2 22k
Q1 2 2 0 CA3046
Q2 3 2 0 CA3046
.control
dc vs 0.0 25 0.01
plot -i(vs) ylimit 0 2m
ac dec 200 1e4 0.9G
plot 1/abs(i(vs))
.endc
.model CA3046 NPN (
+ IS=10e-15 XTI=3 EG=1.11
+ VAF=1e2 VAR=1e2 BF=145.7
+ ISE=114.286e-15 NE=1.48 IKF=46.7e-3
+ XTB=0 BR=0.1 ISC=10.005e-15
+ NC=2 IKR=10e-3 RC=10
+ CJC=991.71e-15 MJC=0.333 VJC=0.75
+ FC=5e-1 CJE=1.02e-12 MJE=0.336
+ VJE=0.75 TR=10e-9 TF=277.01e-12
+ ITF=1.75 XTF=309.38 VTF=16.37
+ PTF=0 RE=0 RB=0
+ CJS=0.5e-12)
```

Figura 18.8: Descrizione del circuito e comandi (a sinistra) e modello dei transistor utilizzati (a destra) per lo studio del circuito *current mirror*.

Vengono effettuate due analisi: la prima (**dc**) calcola l'andamento della corrente continua nel collettore del transistor Q2 al variare della tensione collettore-emettitore da 0 a 25 V. I risultati ottenuti sono riportati nella fig. 18.9 a sinistra. La seconda analisi (**ac**) calcola l'impedenza dinamica *vista* dal generatore V_s . Per questa seconda analisi il generatore V_s svolge due funzioni: generatore di tensione continua come indicato nella definizione (15 V) e generatore di tensione alternata per l'analisi **ac** in funzione della frequenza. Poiché la tensione di questo generatore è quella implicita (1 V), l'inverso della corrente *alternata* che fluisce corrisponde numericamente all'impedenza complessa del circuito di collettore di Q2. Il grafico a destra in fig. 18.9 riporta il modulo di questa quantità. Il valore in bassa frequenza ($75 \text{ k}\Omega$) corrisponde alla pendenza della curva $i(v)$ a sinistra. All'aumentare della frequenza, arrivando a frequenze dell'ordine di 1 MHz, gli effetti capacitivi nel transistor cominciano a farsi sentire e rapidamente il modulo dell'impedenza scende verso lo zero.

18.13 Generatori controllati

Un generatore controllato è un generatore di tensione o di corrente comandato da tensioni e correnti presenti in altri punti del circuito. *Spice* riconosce quattro tipi di generatori controllati lineari e due tipi non lineari. Ognuno di questi è descritto

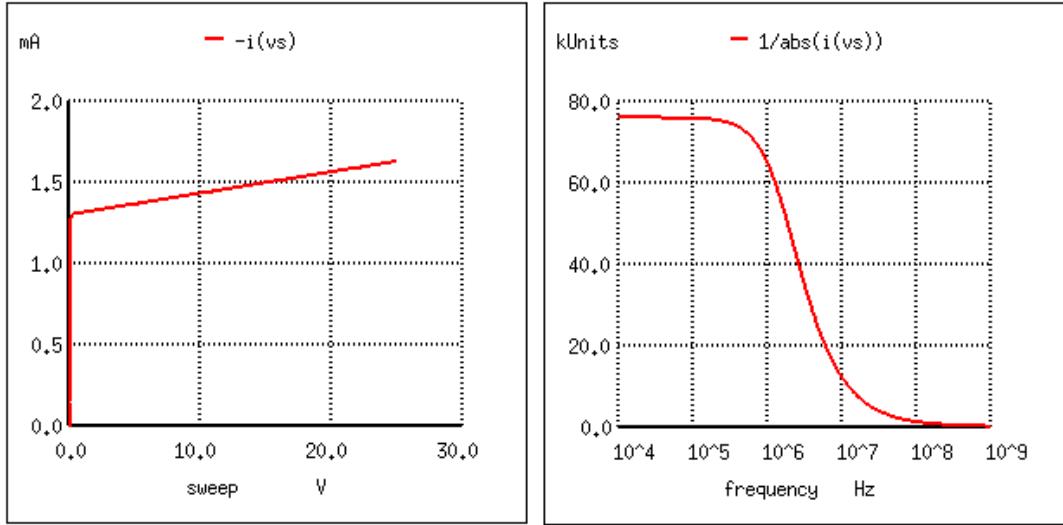


Figura 18.9: Corrente continua in funzione della tensione (a sinistra) e modulo dell'impedenza equivalente in funzione della frequenza (a destra) per il circuito *current mirror*.

da una riga simile a quelle già viste per gli altri componenti, composta dal nome il cui primo carattere individua il tipo di generatore e dai due nodi a cui il generatore è connesso. Seguono poi le caratteristiche del generatore, che sono diverse a seconda del tipo.

18.13.1 Generatori controllati lineari

Generatore di tensione controllato da una tensione

Exxxxxx N+ N- NC+ NC- val

La tensione generata tra i due nodi **N+** ed **N-** è pari a quella presente tra i due nodi di controllo **NC+** ed **NC-** moltiplicata per la costante **val**.

Generatore di corrente controllato da una tensione

Gxxxxxx N+ N- NC+ NC- val

La corrente generata e fatta circolare nel circuito dal nodo **N+** al nodo **N-** è pari alla differenza di potenziale tra i due nodi di controllo **NC+** ed **NC-** moltiplicata per la costante **val**, che ha quindi le dimensioni di una *transconduttanza*.

Generatore di tensione controllato da una corrente

Hxxxxxx N+ N- Vc val

La tensione generata tra i due nodi **N+** ed **N-** è pari alla corrente che circola nel generatore-amperometro di nome **Vc**, moltiplicata per la costante **val** che ha le dimensioni di una *transresistenza*. Il verso positivo della corrente nel generatore **Vc** è

dal nodo positivo al nodo negativo *attraverso il generatore*, quindi con segno opposto rispetto alla corrente che il generatore immette nel circuito

Generatore di corrente controllato da una corrente

`Fxxxxxx N+ N- Vc val`

La corrente generata è pari a quella che circola nel generatore-amperometro `Vc`, moltiplicata per la costante `val`. Per i segni valgono le stesse considerazioni viste per i due generatori precedenti.

18.13.2 Generatori controllati non lineari

Generatore non lineare di tensione

`Bxxxxxx N+ N- V=expr`

Generatore non lineare di corrente

`Bxxxxxx N+ N- I=expr`

I generatori non lineari sono tutti individuati dalla lettera `B` come iniziale del nome. La funzione, generatore di tensione o di corrente, è indicata dalla lettera `V` o `I` che compare dopo i nodi `N+` ed `N-`. Ovviamente solo una delle due funzioni può essere indicata. Il valore della tensione o corrente generata è dato dal valore della espressione `expr` che può essere una qualsiasi funzione reale delle tensioni presenti sui nodi del circuito e delle correnti attraverso i generatori-amperometri. L'espressione `expr` può contenere, oltre ai normali operatori aritmetici, anche numerose funzioni matematiche, tra cui ad esempio `abs()`, `sin()`, `cos()`, `log()`, `ln()`, `exp()`, `sqrt`, ecc.

Nel caso di una analisi `ac`, che calcola la funzione di trasferimento complessa, definita solo per le reti lineari, il generatore non lineare viene *linearizzato*, sostituendolo con uno o più generatori lineari controllati i cui parametri sono eguali alle derivate della espressione della sorgente, calcolate nel punto di lavoro `dc`.

18.14 Sottocircuiti

In circuiti di una certa complessità succede di frequente che una parte del circuito, un *sottocircuito*, compaia identico più volte. Non è conveniente ripetere la stessa descrizione di un sottocircuito, sia per l'inutile moltiplicazione di righe identiche, sia perché quando si dovesse presentare la necessità di modificare le caratteristiche del sottocircuito, bisognerebbe farlo in maniera identica in tutte le copie che compaiono. In questo caso diventa alta la probabilità di compiere errori.

Spice offre la possibilità di definire i sottocircuiti come moduli che possono essere inseriti una o più volte all'interno di uno stesso circuito, in maniera analoga a come sono definiti ed utilizzati i modelli di un componente, come si è visto ad esempio per il transistor `bjt`. Un sottocircuito è definito da una sequenza di istruzioni che inizia con il comando

```
.subckt <nome> <nodo> <nodo> ... <nodo>
e viene terminata dal comando
```

```
.ends
```

Le linee di codice tra l'inizio `.subckt` e la fine `.ends` contengono la descrizione del sottocircuito seguendo le regole già viste per i circuiti in genere.

Il `nome` che compare come primo parametro nel comando `.subckt` è una stringa alfanumerica a piacere che identifica in modo univoco il *modello* di sottocircuito. La lista che segue dopo il nome è l'elenco dei nodi del sottocircuito attraverso cui il sottocircuito si connette con il circuito principale. Un sottocircuito viene inserito nel circuito principale tramite il comando

```
Xxxxxxx <nodo> <nodo> ... <nodo> <nome>
```

La lettera `X` identifica il componente come un sottocircuito allo stesso modo che `R` e `C` identificano condensatori e resistenze e analogamente `xxxxxx` è una stringa alfanumerica a piacere. La lista che segue indica i nodi del circuito principale a cui devono essere connessi i nodi del sottocircuito. I nodi nelle due liste nei comandi `.subckt` e `Xxxxxxx` vengono fatti coincidere uno ad uno in base all'ordine con cui compaiono; non è necessario che nodi coincidenti abbiano lo stesso nome, anzi spesso questo non è nemmeno possibile.

Il `nome` che segue al termine della lista dei nodi nel comando `Xxxxxxx` identifica quale sottocircuito tra quelli definiti deve essere inserito come indicato.

18.14.1 Il comando `.include`

La descrizione di un sottocircuito o un modello di un dispositivo possono essere contenuti in un *file* separato da quello del circuito principale e richiamati tramite il comando

```
.include nome_file
```

Il contenuto del *file* richiamato viene inserito senza alcuna modifica nella procedura nel punto esatto in cui si trova il comando `.include`.

18.14.2 Modelli lineari per l'amplificatore operazionale

Un utilizzo tipico dei sottocircuiti è nella descrizione degli amplificatori operazionali. Molti costruttori forniscono una descrizione dei loro prodotti in un formato compatibile (o quasi) con numerose versioni del programma *Spice*.

Quando non si ha interesse ad una descrizione completa del dispositivo, in particolare quando interessa studiarne solo il comportamento in regime lineare, ignorando gli effetti non lineari dovuti ad esempio alla saturazione, è possibile utilizzare reti molto semplici che riproducono le caratteristiche essenziali dell'amplificatore ed in cui l'effetto dell'amplificazione è fornito da un generatore controllato.

Il sottocircuito riportato in fig. 18.10, ad esempio, descrive in maniera soddisfacente il comportamento in regime lineare dell'amplificatore operazionale *TL081*.

I nodi 1 e 2 sono rispettivamente l'ingresso non invertente ed invertente dell'amplificatore ed `Ri` la resistenza di ingresso (in modo differenziale).

Il generatore controllato `Eg` fornisce l'amplificazione, la rete passa basso `Rf` e `Cf` riproduce l'andamento dell'amplificazione con la frequenza.

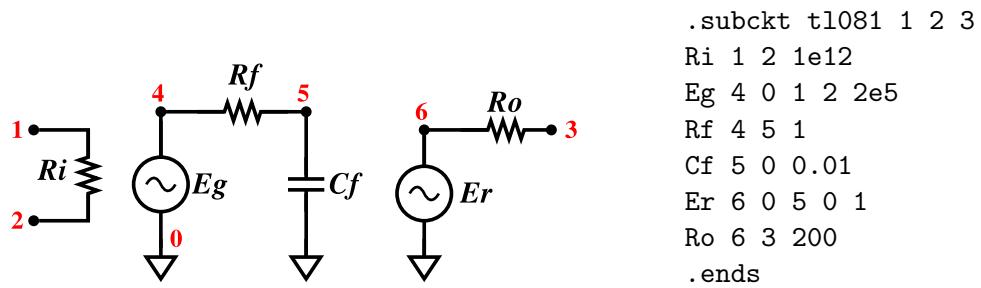


Figura 18.10: Schema elettrico del modello lineare di un amplificatore operazionale *TL081* e sua descrizione in un *sottocircuito*.

Un secondo generatore controllato, E_r , ripete la tensione presente sul nodo 5 all'uscita del circuito passa-basso e la fornisce all'uscita dell'amplificatore tramite la resistenza di uscita R_o .

Può sorprendere a prima vista il valore di 1Ω per la resistenza R_f e $0.01 F$ per il condensatore C_f . Questi due componenti, come anche il generatore E_g , non sono collegati elettricamente con nessuna altra parte del circuito. La tensione sul nodo 5, utilizzata dal secondo generatore controllato E_r , dipende solo da E_g e dalla costante di tempo $R_f \cdot C_f$, non dai singoli valori di R_f e C_f . In questi casi è pratica comune usare valori *unitari* per i componenti, quindi ad esempio 1Ω e $0.01 F$, oppure $1 F$ e 0.01Ω . Qualunque coppia di valori che abbia lo stesso prodotto è equivalente.³

18.15 Anello di reazione con operazionale

Il controllo dell'ampiezza dell'oscillazione ottenuto nel circuito di fig. 13.7 tramite i due moltiplicatori a quattro quadranti può essere realizzato in modo meno accurato ma molto più semplice tramite una coppia di diodi, come riportato nel circuito di fig. 18.11, descritto in linguaggio *Spice* dal file in fig. 18.12

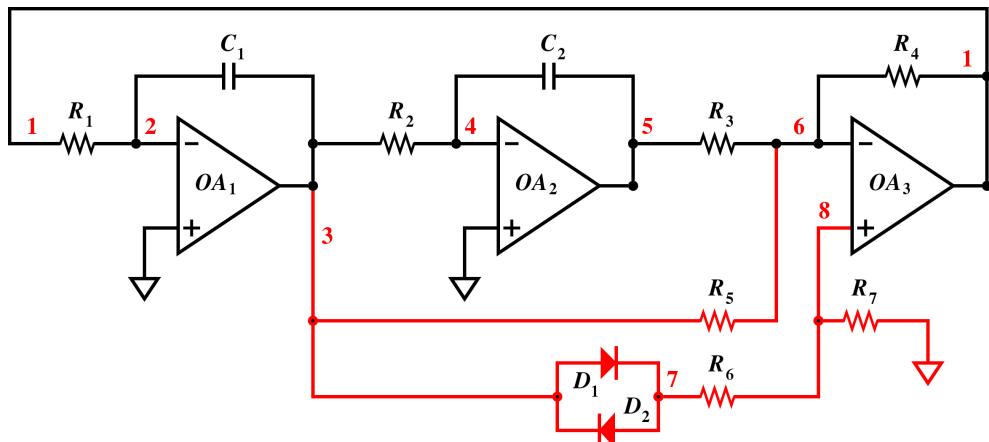


Figura 18.11: Oscillatore con ampiezza controllata dalla non linearità di due diodi.

³Nel limite in cui non si scelgono valori numerici *patologici* come $1e-64$, con il rischio di mandare fuori precisione l'aritmetica di *Spice*.

*** anello con operazionali ***

```
r1    1  2      10k          .model diodo d
c1    2  3      15n  ic=0.001 * (is=1e-10 rs=16 bv=100)
r2    3  4      10k
c2    4  5      15n  ic=0   .subckt tl081  1  2  3
r3    5  6      10k
r4    6  1      10k
xop1  8  6  1  tl081
xop2  0  2  3  tl081
xop3  0  4  5  tl081
r5    3  6      10k
d1    3  7      diodo
d2    7  3      diodo
r6    7  8      9.2k
r7    8  0      5k
                               .control
                               tran 1u 30m 0 1u uic
                               plot v(5)
                               .endc
```

Figura 18.12: Analisi del circuito di fig. 18.11. Nella colonna a sinistra è riportata la descrizione del circuito; nella colonna a destra è riportato il modello del diodo 1N4148 (limitato alla sola parte statica), il modello lineare dell'amplificatore operazionale *TL081* ed i comandi per l'analisi di transitorio ed il disegno dei risultati.

In prima approssimazione, con $R_3 = R_4 = R_5$, la rete disegnata in rosso introduce un contributo

$$v_{(1)} = -v_{(3)} + (v_{(3)} - V_\gamma) \cdot \frac{3R_7}{R_6 + R_7} = -\frac{1}{R_1 C_1} v_{(3)} \quad (18.1)$$

che si azzera quando l'ampiezza dell'oscillazione crescente arriva al valore $v_3 \simeq 10 \text{ V}$.

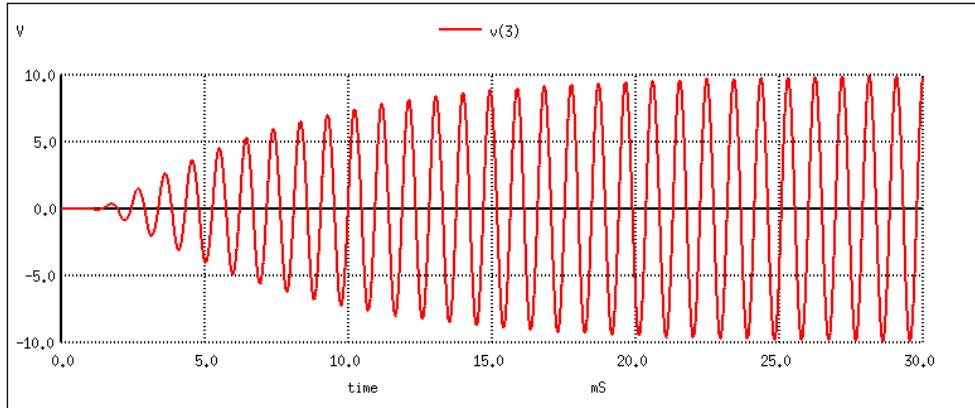


Figura 18.13: Nascita, crescita e stabilizzazione dell'oscillazione nel circuito di fig. 18.11.

La figura 18.13 riporta la nascita, la crescita e la stabilizzazione dell'oscillazione nei primi 30 ms di funzionamento del circuito.

Nella descrizione del circuito sono state impostate le condizioni iniziali sui due condensatori: 1 mV su C_1 , 0 su C_2 . Il calcolo di *Spice* non tiene conto dell'agitazione termica. Il valore di tensione iniziale di 0 V sia su C_1 che su C_2 corrisponde ad una condizione di equilibrio del circuito, sia pure instabile. In assenza di rumore il circuito non si allontana da questa posizione. Conviene introdurre una perturbazione iniziale che porti il circuito leggermente fuori dall'equilibrio per simulare l'effetto di *avvio* indotto dal rumore.⁴

⁴A volte si osserva nel calcolo che il circuito si allontana comunque dalla posizione di equilibrio ed evolve nel tempo anche senza l'accorgimento illustrato. Questo è dovuto agli arrotondamenti legati all'aritmetica del calcolo, che a volte (ma non sempre) agiscono introducendo una lieve perturbazione che porta il circuito fuori dalla condizione di equilibrio.

Capitolo 19

Modulazione e conversione di frequenza

L'elettronica nasce all'inizio del secolo scorso come *radiotecnica*, la tecnica di trasmissione a distanza dei segnali utilizzando onde elettromagnetiche generate, trasmesse e ricevute tramite dispositivi e circuiti *elettronici*.

Lo sviluppo successivo in questo campo è stato immenso ed è ancora in corso molto attivamente. Non è possibile in nessun modo in questi appunti pensare di poterne dare a nessun livello una descrizione adeguata. I paragrafi che seguono sono soltanto una selezione abbastanza arbitraria di alcuni argomenti che sono alla base del funzionamento non solo della radio, ma anche di moderna strumentazione di interesse scientifico.

Dopo l'epoca pionieristica della trasmissione con onde smorzate prodotte dai trasmettitori a scintilla, in cui il segnale irradiato copriva uno spettro molto ampio di frequenze, e dopo l'invenzione dei tubi a vuoto, si è rapidamente affermato l'utilizzo della trasmissione tramite un'*onda portante* (oscillazione monocromatica pura di frequenza sufficientemente alta), *modulata* dal segnale contenente l'informazione da trasmettere, caratterizzato da uno spettro con frequenze molto minori di quella della portante.

19.1 Modulazione di ampiezza

Una *portante*, segnale monocromatico stazionario $v(t) = A_p \cos(\omega_p t + \varphi)$,¹ non contiene alcuna informazione oltre al valore, costante, di A_p , ω_p e φ . Per trasmettere informazione è necessario aggiungere alla originale oscillazione pura una *modulazione*, cioè una variazione nel tempo di uno o più dei suoi parametri.

A seconda del parametro che si fa variare, si ha rispettivamente modulazione di ampiezza, di frequenza e di fase. Storicamente, la prima ad essere utilizzata è stata la modulazione di ampiezza (*AM*), ottenuta variando l'ampiezza della portante in

¹Come già fatto presente in altri capitoli di questi appunti, per indicare la frequenza di un segnale sinusoidale verrà utilizzato indifferentemente sia il simbolo f_x che $\omega_x (= 2\pi f_x)$, come più conveniente, senza indicare ogni volta l'equivalenza, per evitare di riempire il testo e le espressioni di fattori 2π .

funzione del tempo, proporzionalmente al valore del segnale da trasmettere,² secondo l'espressione:

$$s(t) = A_p[1 + a(t)] \cos(\omega_p t + \varphi) \quad (19.1)$$

Un esempio di portante modulata in ampiezza si vede nella fig. 19.1. L'andamento del segnale modulante è riprodotto nell'inviluppo della portante. Dalla stessa figura si vede facilmente che il valore di $a(t)$ deve essere limitato tra -1 e +1 per poter essere in relazione biunivoca con l'ampiezza del segnale modulato, che assume valori tra 0 e $2A_p$.

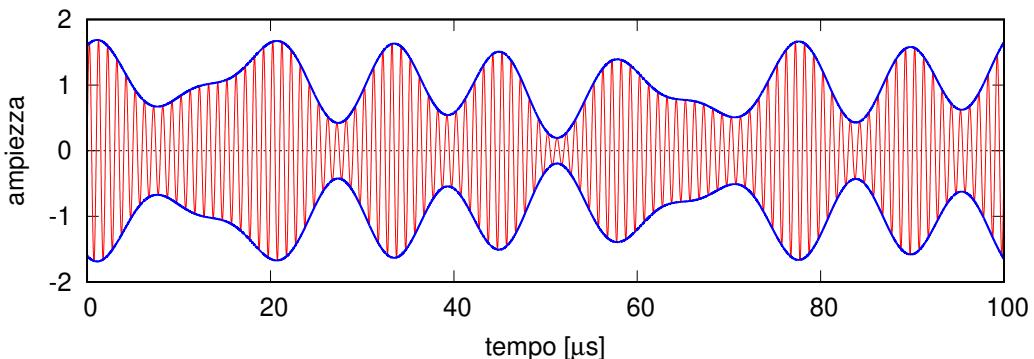


Figura 19.1: Modulazione di ampiezza: un segnale sinusoidale a frequenza $f_p = 1 \text{ MHz}$ (linea rossa) modulato in ampiezza da un segnale con variazione temporale più lenta (linea blu). Il modulo del valore di picco di a nella eq. 19.1, tra 0 ed 1 oppure tra 0 e 100%, è l'*indice o profondità della modulazione*.

Per fare un esempio, $a(t)$ può essere un segnale audio, con uno spettro di componenti armoniche che si estende da 20 Hz a 20 kHz (suono ad *alta fedeltà*), o da 300 a 3400 Hz , nel caso del segnale limitato alle principali componenti spettrali della voce umana (*segnale telefonico*).

Anche senza entrare nei dettagli dell'analisi di Fourier, si può comunque considerare la funzione $a(t)$ formata matematicamente dalla somma di componenti armoniche elementari:³

$$a(t) = \sum_m A_m \cos(\omega_m t + \theta_m) \quad (19.2)$$

Combinando le equazioni 19.1 e 19.2 si ottiene

$$s(t) = A_p \left\{ 1 + \sum_m A_m \cos(\omega_m t + \theta_m) \right\} \cdot \cos(\omega_p t + \varphi) \quad (19.3)$$

²La prima e più antica forma di modulazione è stata la trasmissione telegrafica in codice *Morse*, denominata *CW* (= *continuous wave*, trasmissione in onda continua), ottenuta inviando un'onda elettromagnetica ad intervalli lunghi e brevi opportunamente spaziati, in modo da riprodurre i punti e le linee del telegrafo elettrico. La modulazione *CW* è quindi una modulazione di ampiezza di tipo *tutto/niente*, segnale presente/assente.

³Oppure anche come integrale di una distribuzione continua. Ma i passaggi matematici successivi rimangono in sostanza gli stessi. Si suppone $a(t)$ a media nulla.

L'operazione alla base della modulazione di ampiezza è quindi il *prodotto* tra due segnali, operazione che può essere ottenuta utilizzando un *modulatore*, ad esempio il moltiplicatore a quattro quadranti descritto nel par. 13.2, o un altro circuito equivalente.

19.2 Identità trigonometriche

Prodotti e potenze di funzioni trigonometriche possono essere trasformate in somme di funzioni trigonometriche di somme e differenze di angoli. Nelle tabelle seguenti vengono riportate come promemoria due identità trigonometriche relative alle somme e differenze di angoli e le *formule di Werner* (o di *prostaferesi*) che ne derivano.

19.2.1 somme e differenze di angoli

$$\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta \quad (19.4)$$

$$\sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \cos \alpha \sin \beta \quad (19.5)$$

19.2.2 formule di Werner (o di prostaferesi)

$$\sin \alpha \cos \beta = \frac{1}{2} \{\sin(\alpha + \beta) + \sin(\alpha - \beta)\} \quad (19.6)$$

$$\cos \alpha \cos \beta = \frac{1}{2} \{\cos(\alpha + \beta) + \cos(\alpha - \beta)\} \quad (19.7)$$

$$\sin \alpha \sin \beta = \frac{1}{2} \{\cos(\alpha - \beta) - \cos(\alpha + \beta)\} \quad (19.8)$$

19.3 Bande laterali

Applicando la formula 19.7 a ciascun termine della somma nella eq. 19.3 e scegliendo per semplicità l'origine del tempo in modo da avere $\varphi = 0$, si ottiene:

$$s(t) = A_p \left\{ \cos(\omega_p t) + \frac{1}{2} \sum_m A_m \cos(\omega_p t + \omega_m t + \theta_m) + \frac{1}{2} \sum_m A_m \cos(\omega_p t - \omega_m t - \theta_m) \right\} \quad (19.9)$$

Nella figura 19.2 sono riportati a confronto in colore blu gli spettri delle frequenze del segnale in *banda base* [$a(t)$] e della portante [$v(t)$] all'ingresso del modulatore ed in colore rosso lo spettro del segnale modulato $s(t)$ come risulta dalla equazione 19.9.

Per ogni componente di frequenza (angolare) ω_m del segnale modulante si hanno nello spettro del segnale modulato due componenti di frequenza $\omega_p + \omega_m$ e $\omega_p - \omega_m$, che vanno a formare due *bande laterali* ai lati della portante: una banda laterale superiore (USB, *upper side band*) ed una banda laterale inferiore (LSB, *lower side band*), ciascuna con lo stesso andamento spettrale del segnale in banda base, ma con la LSB rovesciata specularmente rispetto a ω_p . Quindi il segnale modulato contiene in ciascuna banda laterale la stessa informazione che è presente nel segnale in banda base.

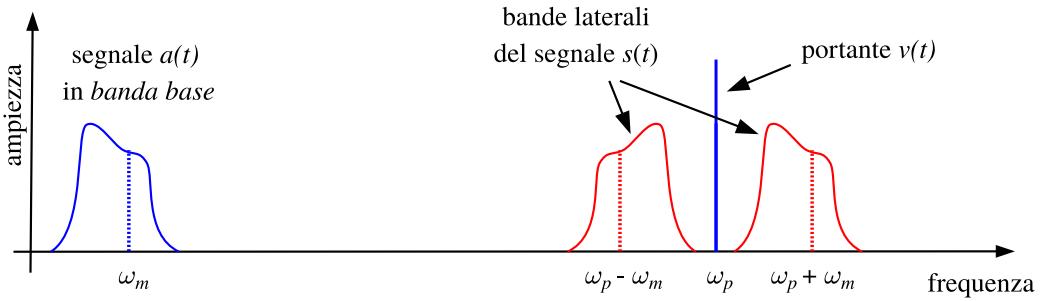


Figura 19.2: Modulazione di ampiezza: ogni componente spettrale di frequenza angolare ω_m del segnale contenente l'informazione (segnale in *banda base*), tramite il processo di modulazione, viene traslata nelle due frequenze $\omega_p \pm \omega_m$ nelle due *bande laterali* intorno alla frequenza portante.

Nella trasmissione radio il processo di modulazione (e successiva demodulazione alla ricezione per recuperare il segnale in banda base) è praticamente indispensabile. Infatti, il segnale in banda base non può essere trasmesso direttamente inviandolo ad una antenna. Si tratta quasi sempre di un segnale a banda larga, secondo le definizioni viste nel par. 2.12. Anche nel caso del segnale telefonico, la cui banda va da 300 a 3400 Hz , e quindi si estende solo per 3100 Hz , il rapporto tra frequenza massima e frequenza minima è 11.3. Non è possibile⁴ fare risuonare una antenna trasmittente in maniera efficiente su una banda di frequenze così ampia e con lunghezze d'onda che vanno dai cento ai mille km (per essere ragionevolmente efficiente, un'antenna deve avere dimensioni geometriche non più di un ordine di grandezza inferiori alla lunghezza d'onda del segnale).

Trasferendo l'informazione dalla banda base alle due bande laterali intorno ad una portante a frequenza sufficientemente alta, come illustrato nella figura, si ottiene un segnale a banda stretta. Ad esempio, con una portante a frequenza di 1 MHz , il rapporto tra frequenza massima e minima, per il segnale telefonico, diventa 1.007. In queste condizioni diventa molto facile realizzare antenne risonanti ad elevata efficienza, di dimensioni tra le decine ed il centinaio di metri. Con portanti di frequenza più alta, le dimensioni dell'antenna si riducono ancora: 30 m per una frequenza di 10 MHz , 3 m per 100 MHz e così via.

19.4 Conversione di frequenza

Lo stesso processo alla base della modulazione di ampiezza viene utilizzato per la *conversione di frequenza*, cioè per traslare un segnale da una regione dello spettro delle frequenze ad un'altra.

Moltiplicando nuovamente il segnale $s(t)$ della eq. 19.9 con un segnale a frequenza ω_q , sia la portante che tutte le componenti delle bande laterali vengono nuovamente duplicate in due nuove bande di frequenze $\omega_q \pm \omega_s$, dove le ω_s sono le frequenze delle componenti di $s(t)$.

⁴Quanto meno, è molto difficile.

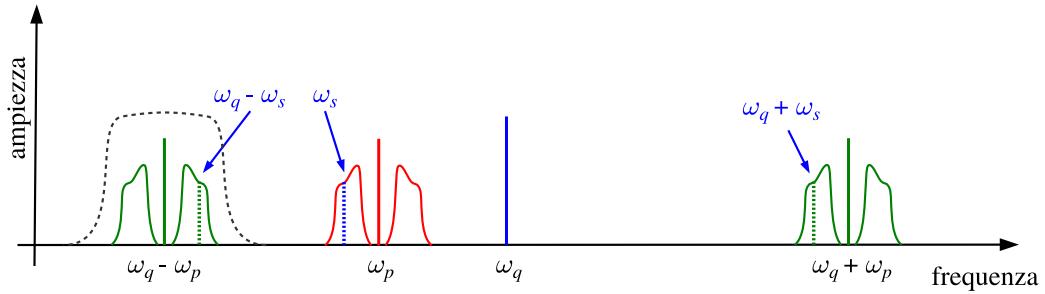


Figura 19.3: Conversione di frequenza: il segnale intorno alla portante ω_p (in rosso), moltiplicato per il segnale monocromatico a frequenza ω_q (in blu), origina due nuove copie intorno alle frequenze $\omega_q + \omega_p$ e $\omega_q - \omega_p$ (in verde). Un filtro passa banda (curva tratteggiata in grigio), centrato sulla frequenza $\omega_q - \omega_p$, seleziona solo una delle due copie.

Se ω_p non è troppo piccolo rispetto a ω_q , le due nuove bande contenenti i duplicati del segnale $s(t)$ si trovano in regioni dello spettro sufficientemente distanti da poter essere separate con un filtro (fig. 19.3).

La conversione di frequenza è fondamentale nel funzionamento dei ricevitori radio.

19.5 Ricevitore *Supereterodina*

Le funzioni di base di un ricevitore radio dall'antenna all'utilizzatore finale del segnale, per esempio un altoparlante o uno strumento di misura, sono:

- Amplificazione
- Selettività
- Rivelazione

Queste tre funzioni sono comuni sia ai normali apparecchi per l'ascolto dei programmi di radiodiffusione sia a strumentazione più raffinata per uso scientifico o tecnologico, come un radiotelescopio, un analizzatore di spettro, un apparecchio per misure di risonanza magnetica (nucleare o elettronica) o un radar per navigazione.

Amplificazione

Il limite ultimo di sensibilità di un ricevitore radio è dato dal rumore termico, che a temperatura ambiente ha una densità spettrale $dP/df = 4kT \simeq 1.66 \cdot 10^{-20} W/Hz$ con k costante di Boltzmann ($1.38 \cdot 10^{-23} J/^{\circ}K$) e T temperatura ambiente ($300^{\circ}K$). Per un canale telefonico con banda passante di $3100\ Hz$ questo corrisponde ad una potenza $P = 51.5 \cdot 10^{-18} W$, che equivale per una antenna con impedenza di 50Ω ad una tensione di $51\ nV_{rms}$. Per portare questi segnali ai livelli dei mV o dei V necessari per le applicazioni, si richiedono amplificazioni fino a $100 \div 150\ dB$.

Selettività

Un buon ricevitore deve essere in grado di selezionare in tutto lo spettro elettromagnetico, che si estende dai 100 kHz delle onde lunghe fino ai GHz delle microonde, esclusivamente l'intervallo di frequenze di interesse, che può essere anche molto stretto. Ad esempio, nel caso delle telecomunicazioni, un canale audio può essere, come si è visto, anche di soli 3100 Hz . Ricevitori per altre applicazioni possono richiedere prestazioni simili e a volte anche più stringenti.

Rivelazione

La rivelazione (o demodulazione) è il processo con cui si estrae dal segnale ricevuto, amplificato e filtrato con la dovuta selettività, il segnale originale in banda base. La demodulazione di un segnale AM come quello di fig. 19.1 è particolarmente semplice ed immediata: è sufficiente un raddrizzatore come quello di fig. 4.11, con un filtro RC passa basso con costante di tempo sufficientemente lunga da eliminare in maniera efficace l'ondulazione residua della portante, ma non tanto da alterare l'andamento temporale del segnale modulante. Ogni altro tipo di modulazione avrà, ovviamente, un suo demodulatore specifico.

A partire dal secondo decennio del 1900 si è affermato uno schema di principio per i ricevitori radio, detto *supereterodina*, basato sulla conversione di frequenza, per effettuare in modo molto efficiente le tre funzioni descritte, ciascuna nella banda di frequenza più conveniente (fig. 19.4).

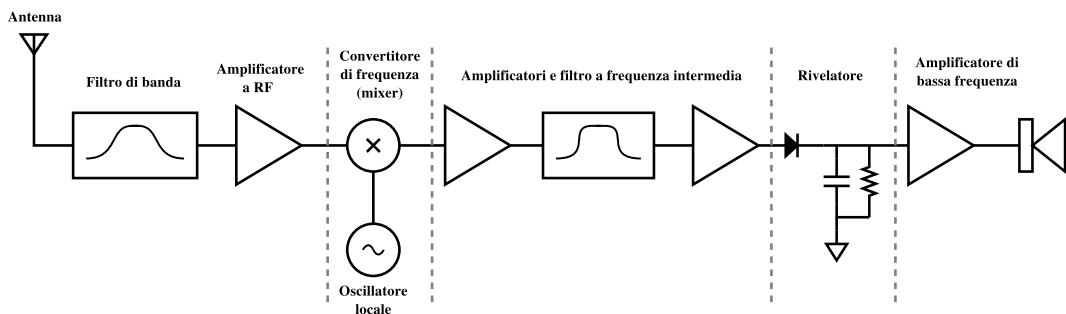


Figura 19.4: Schema a blocchi di un ricevitore supereterodina .

Il primo stadio amplifica il segnale a radiofrequenza (RF) proveniente dall'antenna, effettuando anche una prima selezione grossolana di una banda contenente le frequenze di interesse.

Il secondo stadio, il convertitore di frequenza, combina il segnale a radiofrequenza con un segnale sinusoidale puro prodotto dall'*oscillatore locale*. La frequenza dell'oscillatore locale è scelta in maniera tale da *convertire* l'intervallo di frequenza che si intende ricevere in modo da farlo coincidere con la banda passante dello stadio successivo, contenente amplificatori e filtro di media frequenza. Nella fig. 19.3 lo spettro del segnale di media frequenza è quello in verde a sinistra e la curva tratteggiata in grigio è la banda passante del filtro di media frequenza.

Dopo l'amplificazione ed il filtraggio a media frequenza il segnale viene *demodulato* (o *rivelato*). Nello schema è riportato come esempio il rivelatore AM.

Dopo la rivelazione segue una ulteriore amplificazione, detta di *bassa frequenza*, per portare il segnale al livello di potenza necessario per le applicazioni (nella figura, un altoparlante).

19.5.1 Vantaggi del circuito supereterodina

I due punti di forza principali del circuito supereterodina sono la distribuzione su più stadi, che lavorano a frequenze diverse, dell'amplificazione totale e della selezione del solo segnale desiderato con la soppressione di tutti gli altri.

Il primo punto è legato alla difficoltà di mantenere stabile un amplificatore con un guadagno molto elevato. Se una amplificazione elevata, ad esempio di 120 dB, fosse concentrata tutta nei circuiti a RF, sarebbe praticamente impossibile evitare il rientro di una piccola frazione di segnale dall'uscita verso l'ingresso, con il risultato di fare entrare i circuiti in oscillazione, a meno di non usare schermi adeguatamente efficienti, ma inevitabilmente scomodi, pesanti, ingombranti e costosi.

Il secondo punto invece è legato al fatto che si possono costruire filtri con una curva di selettività quasi ideale, ma che lavorano solo ad una frequenza fissa. Altrimenti, si possono realizzare filtri a frequenza variabile, ma con prestazioni di selettività nettamente inferiori. Il circuito supereterodina permette di avere selettività accurata su tutte le frequenze di interesse con un solo filtro a frequenza fissa, grazie al fatto che il segnale che si intende ricevere viene sempre convertito alla frequenza fissa del filtro a frequenza intermedia, variando opportunamente la frequenza dell'oscillatore locale.

Esempi di filtri

Nella fig. 19.5 sono riportate a confronto, le curve di selettività di alcuni tipi di filtri.

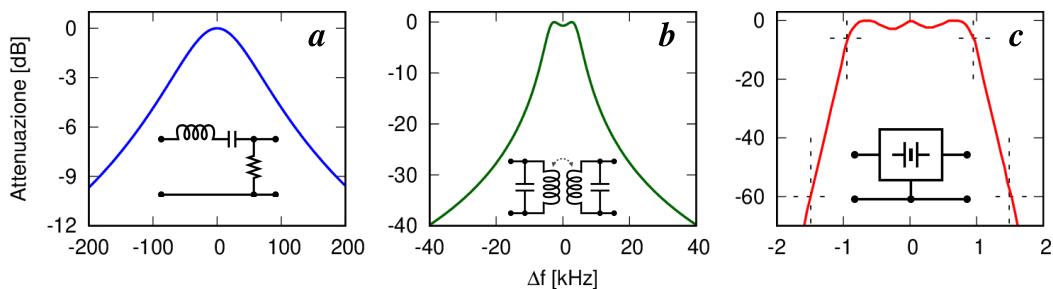


Figura 19.5: Curve di risposta di alcuni filtri. A sinistra, curva blu: circuito risonante LC con frequenza di risonanza $f_0 = 14 \text{ MHz}$ e fattore di merito $Q = 100$. Al centro, curva verde: filtro formato da due circuiti risonanti LC debolmente accoppiati, con frequenza di risonanza $f_0 = 467 \text{ kHz}$ (valore di media frequenza comunemente usato negli apparecchi per radiodiffusione in Europa) e $Q = 100$. A destra, curva rossa: risposta di un filtro commerciale di buona qualità realizzato con una rete di risuonatori a cristallo, con frequenza di centro banda $f_0 = 455 \text{ kHz}$ (valore di media frequenza comunemente usato negli Stati Uniti).

Dal grafico si vede come il primo filtro (*a*), formato da un circuito *LC* semplice, possa essere utilizzato come filtro di banda, per una preselezione del segnale; può anche essere *sintonizzato* tramite un condensatore variabile. Il secondo filtro (*b*) è più complesso, ma anche più efficace per selezionare i segnali di interesse nel canale di media frequenza. È realizzato con due circuiti risonanti debolmente accoppiati in modo da rendere più uniforme la risposta nell'intervallo di frequenza della banda passante. Sarebbe più difficile da sintonizzare, perché richiederebbe la variazione simultanea e sincrona di due condensatori. Il terzo filtro (*c*) è quello che si avvicina di più al comportamento del filtro ideale: attenuazione nulla o minima nella banda passante ed attenuazione molto elevata al di fuori di questa, con fianchi della curva molto ripidi. Essendo realizzato con una rete di risuonatori a cristallo è del tutto impossibile modificarne la frequenza.

Una misura di quanto un dato filtro si avvicina a quello ideale è data dal *fattore di forma*, definito come il rapporto tra la banda passante ad un valore di attenuazione molto elevato (ad esempio 60 dB) ed alla frequenza di taglio, convenzionalmente presa a -3 o a -6 dB. Per il filtro ideale, il fattore di forma è 1; per il filtro *c* (a destra) il fattore di forma tra 6 e 60 dB è 1.6; per il filtro *b* tra 3 e 30 dB è 5.7; per il filtro *a* è ancora peggio.

19.5.2 Svantaggi del circuito supereterodina

Il primo svantaggio è chiaramente la maggiore complessità del circuito rispetto al ricevitore ad *amplificazione diretta* che si avrebbe accoppiando direttamente l'uscita dell'amplificatore a RF con il rivelatore, eliminando oscillatore locale, convertitore e tutti i circuiti di media frequenza.

Il secondo svantaggio è dato dal problema della *frequenza immagine*.

Se f_l è la frequenza dell'oscillatore locale ed f_i il valore della media frequenza, ci saranno due frequenze, $f_1 = f_l - f_i$ ed $f_2 = f_l + f_i$, che saranno entrambe convertite alla frequenza f_i , sovrapponendosi nella banda del filtro di media frequenza. Se f_1 è la frequenza del segnale che intendiamo ricevere, l'eventuale segnale presente sull'antenna a frequenza f_2 (detta *frequenza immagine* o *alias*) sarà un segnale spurio e verrà ad interferire con la nostra ricezione.

Da questo si capisce l'importanza di avere una preselezione, sia pure rozza, dei segnali prima che giungano al convertitore. Non importa che il *filtro di banda* dell'amplificatore a RF abbia una banda passante molto stretta; basta che riesca ad attenuare in maniera sufficientemente efficace i segnali a frequenza f_2 che differiscono di $2f_i$ da f_1 .

19.6 L'evoluzione della supereterodina: la SDR

Per una cinquantina d'anni l'evoluzione dei circuiti della radio si è basata sul miglioramento delle prestazioni di ciascuna parte e sulla introduzione di varianti, come i ricevitori a *doppia conversione* di frequenza, dove lo stadio di media frequenza è seguito da un secondo convertitore ed un secondo filtro ed amplificatore di media frequenza (ovviamente a frequenza diversa dalla prima), prima del demodulatore.

Questa situazione ha cominciato a cambiare drasticamente con l'introduzione delle tecniche digitali. L'oscillatore locale, che inizialmente era un oscillatore *LC* a fre-

quenza variabile, è stato sostituito da un sintetizzatore (PLL, ma più recentemente anche DDS, vedi cap. 17), molto più accurato e stabile. Successivamente si è iniziato a sostituire una parte della catena di elaborazione del segnale con una elaborazione numerica mediante processori digitali.

L'idea di base è estremamente semplice: dato il segnale di ingresso, lo si digitalizza tramite un campionatore sufficientemente accurato e veloce e si svolgono per via numerica tutte le operazioni di amplificazione, filtraggio e demodulazione inizialmente affidate al circuito di fig. 19.4. Si ottiene in questo modo la *SDR*: *Software Defined Radio*, la radio realizzata via software.

Tanto l'idea di base è semplice nel principio, quanto è complessa la sua realizzazione pratica. Il limite principale è dato dal teorema di Shannon-Nyquist, che richiede una frequenza di campionamento pari almeno al doppio della banda di frequenze che si intende elaborare. Questa esigenza si scontra con la possibilità tecnologica di realizzare convertitori analogico-digitali con risoluzione sufficiente alla frequenza di campionamento richiesta.

Per avere un ordine di grandezza, un buon ricevitore o un buon analizzatore di spettro realizzati con la tecnica analogica tradizionale hanno una *dinamica*, cioè una capacità di risoluzione di due segnali vicini, uno molto forte ed uno molto debole, non inferiore a 80...90 dB. Per ottenere le stesse prestazioni, un convertitore ADC deve avere una risoluzione effettiva tra i 14 e i 16 bit. La frequenza massima a cui ogni convertitore può funzionare è limitata, anche se questo limite si sposta ogni giorno più in alto per effetto della evoluzione tecnologica. Il confine tra le frequenze che possono essere ricevute tramite una radio interamente digitale e quelle che richiedono almeno una prima conversione analogica verso frequenze più basse, oggi, si può immaginare a livello di qualche centinaio di MHz, con l'elettronica disponibile commercialmente. Nei laboratori di ricerca questo confine è già superato.

19.6.1 La modulazione SSB

La modulazione a Banda Laterale Unica o SSB (*Single Side Band*) è una variante della modulazione AM. È interessante esaminarla a questo punto per i suoi aspetti matematici, in comune con la matematica della SDR.

Osservando l'eq. 19.9 e lo spettro del segnale modulato in fig. 19.2 si può vedere che la portante non contiene in realtà alcuna informazione e le due bande laterali contengono esattamente la stessa informazione. Da qui nasce l'idea di trasmettere solo una delle due bande laterali. Trasmettendo solo la banda laterale superiore si ha la modulazione USB (*Upper Side Band*); trasmettendo quella inferiore si ha invece la modulazione LSB (*Lower Side Band*). Con la modulazione SSB si ha una minore occupazione dello spettro elettromagnetico ed il risparmio dell'energia associata alla portante.

La soppressione della portante è facile da ottenere. Se nella eq. 19.1 si sopprime il termine 1 nell'espressione tra le parentesi quadre, questo scompare anche nella eq. 19.3 e si porta via il termine $\cos(\omega_p t)$ nella eq. 19.9. Il termine 1 nella eq. 19.1 rappresenta una polarizzazione che è necessario aggiungere al segnale modulante per garantire, quando il segnale modulante è nullo, la presenza della portante al 50% del valore massimo che si raggiunge sui picchi della modulazione. Senza questa polarizzazione, il moltiplicatore a quattro quadranti utilizzato come modulatore fornisce

solo le due bande laterali senza la portante, e in questa configurazione prende il nome di *modulatore bilanciato*.

La soppressione di una delle due bande laterali è operazione più laboriosa. Un metodo è quello di usare un filtro che lasci passare una delle due bande ed elimini l'altra. L'idea è molto semplice concettualmente, ma presenta una difficoltà nella realizzazione pratica: le due bande laterali distano tra loro solo di 600 Hz (nel segnale audio), intorno ad una portante che può avere una frequenza anche di centinaia di kHz o qualche MHz . È quindi necessario utilizzare filtri con fattori di forma molto buoni, come ad esempio quello descritto dalla curva della fig. 19.5c.

Un secondo metodo per la generazione di un segnale SSB è noto come *modulazione a sfasamento*, ed è realizzato con un circuito come quello disegnato in fig. 19.6. Si compone di due *modulatori bilanciati* M_1 ed M_2 che combinano il segnale modu-

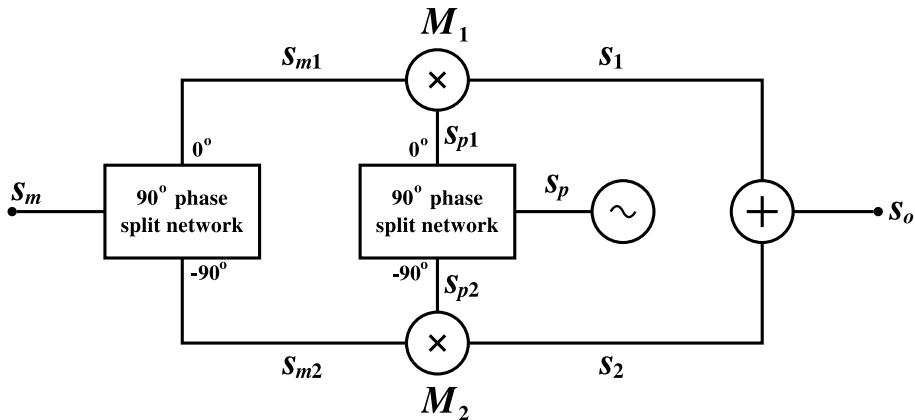


Figura 19.6: Modulatore a sfasamento per la generazione di un segnale SSB. Le due reti *phase split network* generano due copie dei segnali in cui ciascuna componente appare alle due uscite con una differenza di fase di 90° . A seconda che il circuito sommatore all'uscita effettui l'operazione $s_1 + s_2$ oppure $s_1 - s_2$ si avrà $s_o = \cos(\omega_p t - \omega_m t - \varphi)$ oppure $\cos(\omega_p t + \omega_m t + \varphi)$.

lante s_m con il segnale s_p della portante. Entrambi i segnali, prima di giungere ai modulatori passano attraverso una rete a sfasamento che ne genera due copie con una differenza di fase di 90° , nel senso che a ciascuna componente di tipo $\cos()$ all'uscita 0° corrisponde una componente della stessa ampiezza con un ritardo di 90° all'altra uscita, quindi di tipo $\sin()$, e viceversa a ciascuna componente di tipo $\sin()$ corrisponde una componente $-\cos()$. Nel par. 12.13 è stato descritto il *filtro passa tutto*, con cui è possibile realizzare simili reti.

Considerando una portante $s_p(t) = A_p \cos(\omega_p t)$ ed un segnale modulante $s_m = \sum_i A_i \cos(\omega_i t + \varphi_i)$, si ha:

$$\begin{aligned}
s_{m1} &= \sum_i A_i \cos(\omega_i t + \varphi_i) \\
s_{p1} &= A_p \cos(\omega_p t) \\
s_1 &= A_p \sum_i A_i \frac{\cos(\omega_p t + \omega_i t + \varphi_i) + \cos(\omega_p t - \omega_i t - \varphi_i)}{2} \\
s_{m2} &= \sum_i A_i \sin(\omega_i t + \varphi_i) \\
s_{p2} &= A_p \sin(\omega_p t) \\
s_2 &= A_p \sum_i A_i \frac{\cos(\omega_p t - \omega_i t - \varphi_i) - \cos(\omega_p t + \omega_i t + \varphi_i)}{2}
\end{aligned}$$

I due segnali s_1 ed s_2 vengono sommati (o sottratti) dal circuito sommatore in uscita, dando

$$s_{o^-} = A_p \sum_i A_i \cos(\omega_p t - \omega_i t - \varphi_i)$$

oppure

$$s_{o^+} = A_p \sum_i A_i \cos(\omega_p t + \omega_i t + \varphi_i)$$

cioè una sola delle due bande laterali.

19.6.2 La trasformata di Hilbert

La difficoltà grossa della modulazione SSB a sfasamento non sta nella indubbia complessità del circuito, ma nella realizzazione delle reti a sfasamento a 90° . Per il segnale della portante il problema è semplice: data la singola frequenza ω_p è facile costruire una rete contenente celle LRC o anche solo RC che forniscano un segnale sfasato di 90° , con la stessa ampiezza del segnale originale. Una rete invece che introduca uno sfasamento *esattamente* di 90° su *tutte* le componenti del segnale s_m , lasciandone inalterate le ampiezze, è impossibile da realizzare.

Dai testi di Analisi Matematica si può verificare che ruotare di 90° tutte le componenti armoniche di un segnale $s(t)$, lasciandone inalterate le ampiezze, equivale ad effettuare la trasformata di Hilbert del segnale:

$$H\{s(t)\} = \frac{1}{\pi} \mathbf{P} \int_{-\infty}^{+\infty} \frac{s(\tau)}{t - \tau} d\tau \quad (19.10)$$

dove il simbolo \mathbf{P} indica il *valore principale* di Cauchy.

L'integrale si estende dal tempo $-\infty$ a $+\infty$. Una rete elettronica che effettua questa operazione deve quindi conoscere $s(\tau)$ nel futuro.

Anche in questo caso, come accade spesso, l'operazione che non può essere effettuata *esattamente*, può essere approssimata con il livello di precisione desiderato, utilizzando ad esempio una rete a sfasamento di adeguata complessità, del tipo descritto nel par. 12.13.

È importante osservare che un problema analogo esiste anche nell'altra tecnica di modulazione SSB, mediante l'uso di un filtro molto selettivo. Il filtro ideale, detto *a*

scatola (in inglese: *brickwall filter*), che dovrebbe avere una curva di selettività con i fianchi verticali, passando in modo discontinuo da attenuazione nulla ad attenuazione infinita, è altrettanto impossibile da realizzare come la rete a sfasamento ideale. La sua risposta ad un impulso genererebbe una risposta nel passato, prima dell'inizio dell'impulso all'ingresso, anche in questo caso violando il principio di causalità. Ma, come visto in fig. 19.5, sono possibili ottime approssimazioni del filtro ideale.

19.6.3 Il segnale analitico

A questo punto si deve osservare che se si definiscono le due funzioni complesse

$$\begin{aligned} S_m(t) &= s_{m1}(t) + j s_{m2}(t) = A_i e^{j(\omega_i t + \varphi)} \\ S_p(t) &= s_{p1}(t) + j s_{p2}(t) = A_p e^{j(\omega_p t)} \end{aligned} \quad (19.11)$$

l'operazione svolta dal circuito di fig. 19.6 è descritta dalle due espressioni matematiche

$$\begin{aligned} s_{o^-} &= \Re\{S_p \cdot S_m\} \\ s_{o^+} &= \Re\{S_p \cdot S_m^*\} \end{aligned} \quad (19.12)$$

Le due funzioni $S_m(t)$ e $S_p(t)$ prendono il nome di *segnale analitico*.

Il segnale analitico può essere considerato come una estensione del formalismo delle quantità complesse utilizzato comunemente nello studio dei circuiti in corrente alternata.

Una tensione (o una corrente) alternata nel formalismo complesso si scrivono

$$V(t) = V_0 e^{j\omega t} = V_0 [\cos(\omega t) + j \sin(\omega t)]$$

con la condizione che il segnale fisico $v(t)$ corrisponde alla parte reale di $V(t)$. Quindi

$$V(t) = v(t) + j H\{v(t)\}$$

In pratica, il segnale analitico $V(t)$ è composto da un segnale fisico $v(t)$ a cui è aggiunta come componente immaginaria la sua trasformata di Hilbert.

Negli algoritmi della SDR il segnale analitico risulta particolarmente conveniente, come nell'esempio nel paragrafo seguente.

19.6.4 Conversione di frequenza con il segnale analitico

Il prodotto tra due segnali fisici a frequenze ω_1 e ω_2 , come già visto nei paragrafi precedenti ed in accordo con le formule di Werner, genera due segnali con frequenza somma e differenza, $\omega_1 \pm \omega_2$, e questo, come visto nel par. 19.5.2, è all'origine del problema della *frequenza immagine* nei ricevitori supereterodina.

Lo stesso problema non esiste con il segnale analitico. Presi due segnali

$$S_m(t) = A_m e^{j\omega_m t} \quad S_p(t) = A_p e^{j\omega_p t}$$

il loro prodotto (complesso) è

$$S_m(t) \cdot S_p(t) = A_m A_p e^{j(\omega_m + \omega_p)t}$$

cioè un segnale (analitico) a frequenza $\omega_m + \omega_p$.

Nella SDR la conversione di frequenza si ottiene semplicemente moltiplicando il segnale che si vuole convertire con la funzione $e^{j\Delta\omega t}$, dove $\Delta\omega$ è lo spostamento di frequenza che si vuole ottenere.

È utile osservare che $\Delta\omega$ può essere sia positivo che negativo, producendo così uno spostamento del segnale a scelta verso frequenze più alte oppure più basse.

Nel mondo dei segnali fisici una frequenza negativa non ha molto senso: $\cos(\omega t + \varphi)$ differisce da $\cos(-\omega t + \varphi)$ [= $\cos(\omega t - \varphi)$] solo per una fase, ma è sotto ogni altro aspetto indistinguibile.

Nel campo dei segnali analitici invece $e^{j\omega t} = \cos(\omega t) + j \sin(\omega t)$ e $e^{-j\omega t} = \cos(\omega t) - j \sin(\omega t)$ sono due funzioni distinte e non sovrapponibili: nel primo caso la componente immaginaria segue quella reale di un quarto di periodo, nel secondo caso invece la precede della stessa quantità e questo rende i due segnali indipendenti e distinguibili.

19.7 La conversione *I-Q*

Nella fig. 19.7 è riportato lo schema a blocchi di un circuito utilizzato spesso come stadio di ingresso (*front-end*) di un ricevitore SDR.

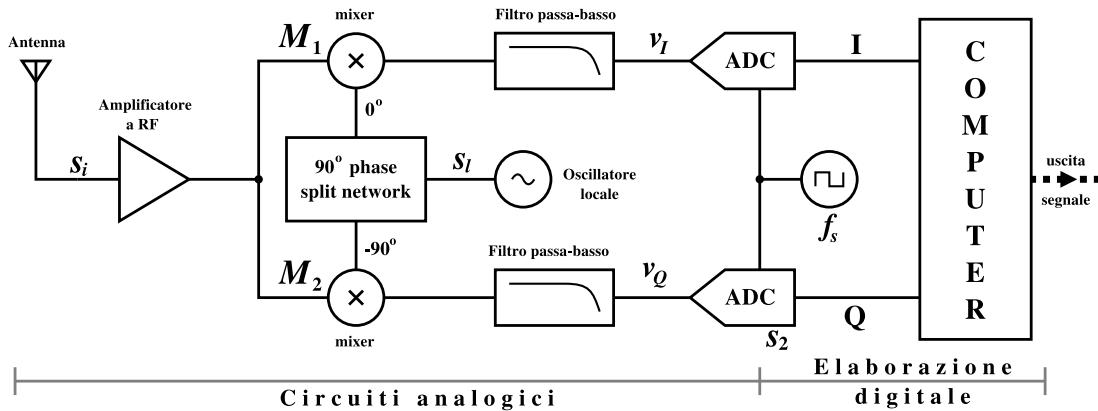


Figura 19.7: Schema a blocchi di un convertitore *I-Q* per ricevitore SDR. La parte a sinistra è il *front-end* analogico che trasforma il segnale fisico proveniente dall'antenna nelle due componenti di un segnale analitico. Gli ADC convertono le due componenti in un flusso di valori numerici, che viene elaborato dal computer a destra secondo gli algoritmi della SDR.

Lo scopo del circuito è di combinare il segnale fisico di ingresso $s_i(t)$ proveniente ad esempio da una antenna con il segnale fisico $s_l(t)$ generato da un oscillatore locale, in modo da ottenere le due componenti, reale ed immaginaria, del segnale analitico. Queste due componenti, indicate comunemente come *I* (*In phase*) e *Q* (*Quadrature*), una volta campionate con una cadenza adeguata f_s tramite due ADC, sono il flusso

di dati all'ingresso di un computer per la successiva elaborazione numerica secondo gli algoritmi della SDR.

All'uscita dei due *mixer* M_1 ed M_2 si hanno i due segnali

$$\begin{aligned} v_I(t) &= a_i \cos(\omega_i t + \varphi) \cdot \cos(\omega_l t) \\ v_Q(t) &= a_i \cos(\omega_i t + \varphi) \cdot \sin(\omega_l t) \end{aligned}$$

Per semplicità si è considerato eguale a 1 il rapporto tra l'ampiezza del segnale dell'oscillatore locale ed il fattore di scala dei *mixer*.

Sviluppando le due equazioni si ottiene:

$$v_I(t) = \frac{a_i}{2} \cos(\omega_i t + \varphi + \omega_l t) + \frac{a_i}{2} \cos(\omega_i t + \varphi - \omega_l t) \quad (19.13)$$

$$v_Q(t) = \frac{a_i}{2} \sin(\omega_i t + \varphi + \omega_l t) - \frac{a_i}{2} \sin(\omega_i t + \varphi - \omega_l t) \quad (19.14)$$

È abbastanza comune collocare la frequenza ω_l dell'oscillatore locale al centro della banda che si intende ricevere, come indicato nella fig. 19.8. Di conseguenza in ciascuna delle due eq. 19.13 e 19.14 si ha un termine con frequenze $\omega_l + \omega_i \simeq 2\omega_l$ ed uno con frequenze $\omega_l - \omega_i$ intorno a 0.

I due filtri passa-basso che seguono i mixer, con frequenza di taglio f_h , eliminano i due termini a frequenza più alta, lasciando soltanto

$$\begin{aligned} v_I(t) &= +\frac{a_i}{2} \cos(\omega_i t + \varphi - \omega_l t) \\ v_Q(t) &= -\frac{a_i}{2} \sin(\omega_i t + \varphi - \omega_l t) \end{aligned}$$

Questi due segnali, campionati dai due ADC in maniera sincrona a frequenza f_s , formano all'ingresso del computer, che provvede alla successiva elaborazione numerica, il segnale analitico

$$V(t) = v_I(t) + jv_Q(t) \quad (19.15)$$

Nella figura fig. 19.8 si vede come l'intervallo di frequenze ω_i da $\omega_l - \omega_h$ a $\omega_l + \omega_h$ all'ingresso viene traslato da ciascuno dei due convertitori nella banda da $-\omega_h$ a $+\omega_h$.

Per ciascuno dei due segnali fisici I e Q presi singolarmente all'uscita dei filtri, come si è già visto, la componente a frequenza positiva si sovrappone a quella a frequenza negativa, in maniera indistinguibile.

Per il segnale analitico $V(t)$ della eq. 19.15, invece, a ciascuna componente ω_i nell'intervallo di ingresso da $\omega_l - \omega_h$ a $\omega_l + \omega_h$ corrisponde una componente distinta con frequenza positiva oppure negativa tra $-\omega_h$ e $+\omega_h$.

19.8 I filtri nella supereterodina SDR

Il segnale analitico $V(t) = v_I(t) - jv_Q(t)$ all'uscita degli ADC del circuito di fig. 19.7 contiene tutte le componenti armoniche di tutti i segnali presenti nell'intervallo di frequenze da $\omega_l - \omega_h$ a $\omega_l + \omega_h$, traslate alle frequenze da $-\omega_h$ a $+\omega_h$. Per ottenere un ricevitore completo occorre prima filtrare da questo segnale le sole componenti che si trovano all'interno del canale di comunicazione che interessa, e quindi *demodulare* o *rivelare* il segnale filtrato per estrarne l'informazione.

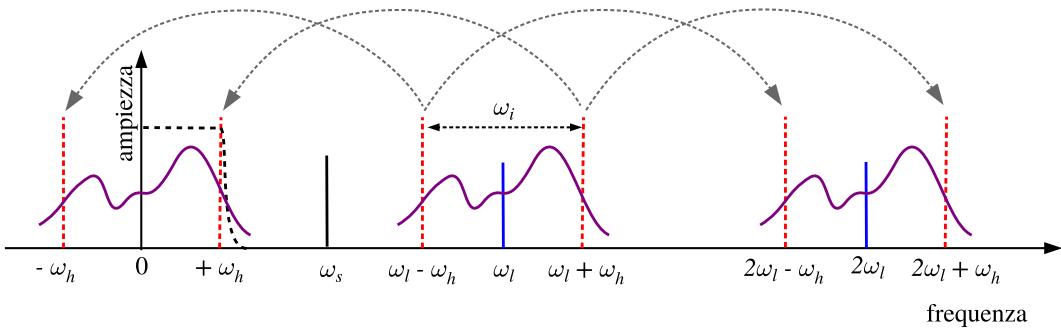


Figura 19.8: Conversione di frequenza I-Q: la frequenza ω_l dell'oscillatore locale si trova al centro della banda di frequenze ω_i che si vuole convertire. All'uscita di ciascuno dei due mixer nel circuito in fig. 19.7 si hanno due copie delle frequenze ω_i , una centrata intorno a $2\omega_l$ e l'altra intorno a zero. I due filtri passa-basso con banda passante da zero a ω_h (curva nera tratteggiata) lasciano passare verso gli ADC solo le componenti tra $-\omega_h$ e $+\omega_h$, che diventano i due segnali v_I e v_Q , parte reale ed immaginaria del segnale analitico che dopo il campionamento viene elaborato dal computer.

Tutte queste operazioni nella SDR sono svolte per via numerica da un computer o da un DSP (*Digital Signal Processor*, rete per elaborazione numerica di segnali) elaborando il flusso di numeri che proviene dagli ADC. La parte di SDR che si trova a valle degli ADC è quindi composta solo da algoritmi per elaborazione numerica e di conseguenza al di fuori del campo dell'*elettronica* in senso stretto.

Per scopo esclusivamente illustrativo di questo tipo di procedure, in questo paragrafo viene descritto un semplice esempio di filtro di media frequenza.

Nella figura fig. 19.9c è riportato lo spettro del segnale proveniente da un convertitore I-Q simile a quello descritto nel par. 19.7. Il convertitore è sintonizzato alla frequenza di 90 MHz, all'interno della banda di radiodiffusione a modulazione di frequenza (FM) e fornisce all'uscita un flusso dati con cadenza di $2 \cdot 10^6$ campioni I-Q al secondo, contenente nella banda di 2 MHz tra -1 e $+1$ MHz i segnali ricevuti nella porzione di spettro tra 89 e 91 MHz.

Si vede che sono presenti sei segnali distinti. Per selezionare uno solo di questi, ad esempio il terzo da sinistra, occorrerà un filtro con una banda passante come riportato in fig. 19.9b: i segnali tra $\omega_1 = -175$ kHz e $\omega_2 = +25$ kHz dovranno passare inalterati attraverso il filtro, quelli al di fuori di questa banda dovranno essere il più possibile eliminati.

Nella figura è riportato anche l'andamento della fase della funzione di trasferimento del filtro. È preferibile ed è stato scelto un andamento lineare, di modo che il ritardo introdotto dal filtro è costante per tutte le componenti spettrali del segnale.

L'operazione di filtraggio è quindi definita nel dominio delle frequenze: l'ampiezza di ogni componente dello spettro deve essere moltiplicata per una funzione *finestra* che idealmente ha valore 1 per ω compreso tra ω_1 e ω_2 e zero per tutte le altre ω .

Ma l'operazione deve essere effettuata nel dominio del tempo, sul segnale $V(t)$.

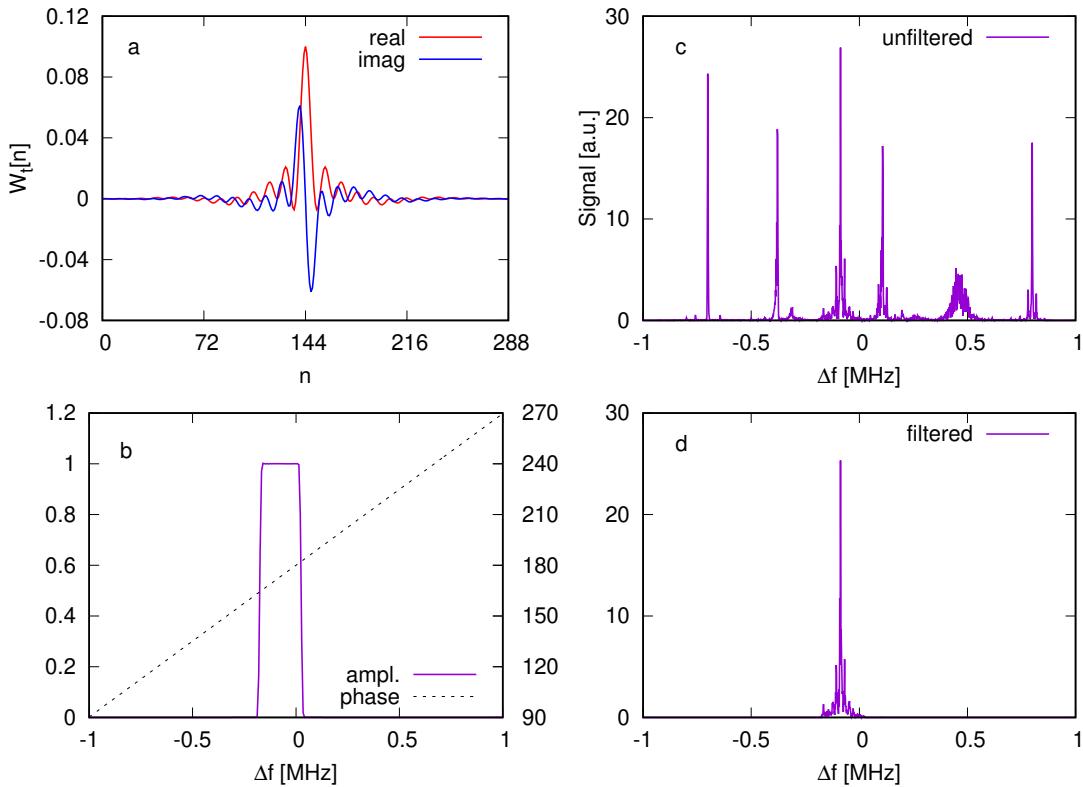


Figura 19.9: Esempio di filtro per un ricevitore SDR. *c)* banda di ampiezza 2 MHz contenente sei segnali FM; *b)* banda passante da -175 kHz a $+25$ kHz di un filtro disegnato per selezionare il terzo segnale presente nello spettro *c* (la linea viola descrive il modulo della funzione di trasferimento, la linea nera tratteggiata la fase); *a)* trasformata di Fourier (complessa) della banda passante *b* del filtro; *d)* segnali presenti all'uscita del filtro.

Dall'analisi di Fourier è noto che al prodotto tra due funzioni nel dominio delle frequenze corrisponde l'integrale di convoluzione tra le loro trasformate nel dominio del tempo.

Quindi, data la funzione $W_f(\omega)$ che descrive il comportamento del filtro nel dominio delle frequenze, se ne prende la sua trasformata (complessa) $W_t(t)$ e la si applica al segnale $V(t)$:

$$U(t) = \int_{-\infty}^t V(\tau)W_t(t - \tau)d\tau \quad (19.16)$$

Poichè il segnale $V(t)$ all'uscita degli ADC non è una funzione continua del tempo ma è composto da una sequenza di valori discreti che procedono alla velocità di campionamento f_s , si avrà una sommatoria piuttosto che un integrale:

$$U(n) = \sum_{i=n-M}^n V(i)W_t(n - i)\Delta t \quad (19.17)$$

dove n indica il campionamento corrispondente all'istante corrente t e $\Delta t = 1/f_s$ è l'intervallo tra due campionamenti.

La sommatoria non si può ovviamente estendere dal tempo $-\infty$, ma coprirà solo gli ultimi M campionamenti, dove M è la lunghezza del vettore $W_t(n)$ (nell'esempio riportato $M = 289$). Questo equivale a troncare la funzione $W_t(\tau)$, assegnandole valore 0 per $\tau > M\Delta t$.

La funzione (complessa) $W_t(\tau)$, o meglio il vettore dei valori $W_t(n)$ ottenuto tramite trasformata di Fourier della funzione *finestra* $W_f(\omega)$, è riportato nella fig. 19.9a.

I fianchi della curva che descrive la banda passante sono molto ripidi, ma non verticali.⁵ Maggiore è la pendenza della curva $W_f(\omega)$, più lentamente decade l'ampiezza dei valori di $W_t(n)$ e quindi maggiore deve essere M per evitare effetti di troncamento. Un compromesso è necessario tra qualità della risposta del filtro e volume di calcolo richiesto. Comunque, il fattore di forma del filtro descritto è circa 1.1, valore molto difficile da ottenere con un filtro analogico anche di ottima qualità.

⁵Nel confrontare le curve dei diversi filtri si tenga presente che quella in fig. 19.9b è su scala lineare, quelle in fig. 19.5 sono su scala logaritmica.

Capitolo 20

Circuiti di potenza

Ogni circuito elettronico contenente componenti attivi ha bisogno di una sorgente di energia per il suo funzionamento. Per un apparecchio portatile di piccola potenza questa può essere una batteria; per un apparecchio fisso può essere un alimentatore che preleva energia dalla rete di distribuzione elettrica sotto forma di corrente alternata (comunemente in Europa a $220\text{ V}, 50\text{ Hz}$) e la fornisce come corrente continua opportunamente livellata ed alla tensione richiesta.

Nel par. 4.10 si è visto come questo si può ottenere utilizzando un trasformatore per portare la tensione al valore desiderato, un raddrizzatore a diodi ed un condensatore di filtro.

Tuttavia una soluzione di questo tipo spesso non è adeguata perché la tensione generata non è accurata e non è stabile (le fluttuazioni di tensione sulle reti elettriche possono arrivare al $\pm 10\%$ del valore nominale) e contiene una ondulazione residua che viene solo parzialmente ridotta dalla presenza del condensatore di filtro.

Per poter alimentare un circuito elettronico con tensioni di valore preciso e stabile e con livelli di ondulazione sufficientemente bassi si utilizzano gli *alimentatori stabilizzati*, circuiti regolatori di tensione i cui principi di funzionamento sono descritti nei paragrafi seguenti.



Figura 20.1: Alimentatore da laboratorio contenente tre alimentatori stabilizzati completamente indipendenti e separati tra di loro, con tensione di uscita regolabile da 0 a 30 V per i primi due e da 0 a 8 V per il terzo. Ogni sezione è dotata di uno strumento che permette di misurare tensione o corrente erogata e di un circuito limitatore della corrente che permette di impostare il valore massimo di corrente erogabile.

Un alimentatore stabilizzato può essere un apparecchio a sé (ad esempio, un alimentatore da laboratorio di uso generale, come quello in fig. 20.1, oppure può essere un circuito contenuto all'interno dell'apparecchio che deve essere alimentato (ad esempio, un televisore, un computer, ecc.).

Attualmente gli alimentatori elettronici stabilizzati, ed i circuiti di potenza in generale, si dividono in due grandi categorie: regolatori lineari e regolatori di tipo *switching*. La tecnologia di ciascuno di questi due tipi di circuiti è ormai diventata molto varia e complessa ed argomento di trattati specialistici. Nei paragrafi seguenti, quindi, ci potrà essere solo una breve descrizione degli schemi e dei principi di base.

20.1 Regolatori lineari

Lo schema di principio di un alimentatore stabilizzato *lineare* è riportato in fig. 20.2.

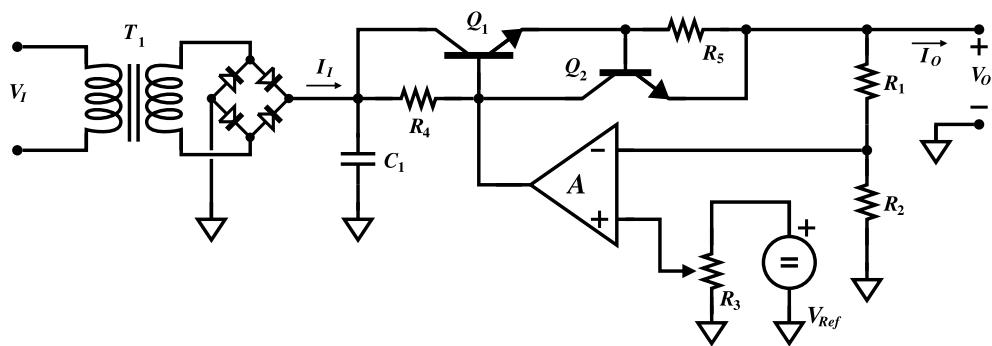


Figura 20.2: Schema di principio di un regolatore lineare. Il generatore V_{Ref} fornisce una tensione di riferimento accurata e stabile, ma non è in grado di erogare altro che una corrente minima. L'amplificatore A ed il transistor regolatore Q_1 moltiplicano la tensione V_{Ref} per una costante e forniscono la corrente e quindi la potenza necessarie.

Il blocco a sinistra, formato da trasformatore, raddrizzatore e condensatore di filtro, come si è visto fornisce una tensione continua non regolata, cioè non stabile e non perfettamente livellata. Queste prestazioni sono compito del circuito elettronico riportato sulla destra.

Il generatore V_{Ref} (tensione di riferimento) schematizza una sorgente di tensione continua, accurata, stabile e *pulita*, nel senso che contiene livelli di rumore ed ondulazione residua molto bassi. Normalmente il generatore V_{Ref} non è in grado di erogare altro che correnti molto piccole e si basa sull'utilizzo di un diodo Zener (o altro componente equivalente), come visto nell'esempio del par. 4.16.¹

La tensione V_{Ref} è applicata all'ingresso + di un amplificatore differenziale (ad esempio un operazionale), al cui ingresso - è applicata tramite il partitore R_1, R_2 una frazione della tensione V_O erogata all'uscita.

È facile riconoscere in questo circuito l'anello a reazione negativa dell'amplificatore non invertente di fig. 11.8, con il generatore V_{Ref} ed il potenziometro R_3 che forniscono il segnale di ingresso v_+ . L'unica differenza è la presenza del transistor

¹Quello del par. 4.16 è solo un circuito di principio; come sempre, il circuito reale può essere anche molto più complicato. Sono disponibili in commercio circuiti integrati monolitici che contengono il diodo Zener e l'elettronica ausiliaria per compensare le derive termiche e ridurre a valori estremamente bassi la resistenza dinamica. È anche disponibile in commercio una grande varietà di circuiti integrati che contengono l'intero regolatore lineare, compreso il transistor passante Q_1 .

Q_1 , utilizzato come *emitter follower*. La funzione di questo transistor, che può essere di tipo bjt o mosfet, è di fornire la potenza: un amplificatore operazionale tipicamente può erogare all'uscita solo pochi milliamper; il transistor Q_1 moltiplica questa corrente per uno o più ordini di grandezza, facendosi carico anche dei problemi di dissipazione termica collegati. Per fattori di moltiplicazione molto elevati (più di due ordini di grandezza) al singolo transistor Q_1 può essere sostituita una coppia in configurazione Darlington (par. 7.11).

A tutti gli effetti il transistor Q_1 , indicato comunemente come *regolatore passante*, può essere considerato come parte dell'amplificatore A e della sua presenza può essere tenuto conto nel valore del guadagno di A .² Di conseguenza la tensione all'uscita sarà

$$V_O = V_{Ref} \cdot K \cdot \left(\frac{R_1 + R_2}{R_2} \right) \quad (20.1)$$

dove con K si è indicato il rapporto di partizione, compreso tra 0 e 1, del potenziometro R_3 .

Il transistor Q_2 e la resistenza R_5 forniscono la protezione contro valori di corrente superiori ai limiti previsti. Quando la differenza di potenziale ai capi di R_5 raggiunge il valore V_γ del transistor Q_2 , questo entra in conduzione e *ruba* la corrente alla base di Q_1 , impedendogli di aumentare ancora la corrente erogata.

20.1.1 Vantaggi e svantaggi dei regolatori lineari

Il regolatore lineare si è affermato come la migliore soluzione per ottenere tensioni di alimentazione accurate, stabili e con livelli di rumore ed ondulazione residua estremamente bassi.

A fronte di questi vantaggi il regolatore lineare soffre di un grosso limite: la scarsa efficienza energetica, che ne sconsiglia l'uso, se appena è possibile, in tutte le applicazioni di potenza, vale a dire oltre uno o al massimo pochi watt.

Trascurando il consumo del generatore V_{Ref} e dell'amplificatore A , che è possibile tenere a valori molto bassi, l'efficienza energetica del circuito è

$$\eta = \frac{P_O}{P_I} = \frac{V_O \cdot I_O}{V_I \cdot I_I} \simeq \frac{V_O}{V_I} \quad (20.2)$$

poichè $I_O \simeq I_I$ a meno di contributi che possono essere trascurati.

La differenza tra V_I e V_O non può essere troppo piccola: V_I può oscillare in un intervallo che arriva a $\pm 15\%$ del valore nominale e $V_I - V_O$ non può essere meno del valore minimo necessario per il corretto funzionamento del transistor passante.

Tipicamente per apparecchi con tensione V_O fissa si può arrivare a valori di rendimento $\eta \simeq 70 \div 80\%$.

Le cose vanno molto peggio per gli alimentatori con tensione variabile, come ad esempio i generatori di laboratorio.

Quando si utilizza un apparecchio del genere per ottenere una tensione V_O di valore molto basso, l'efficienza η scivola inesorabilmente verso lo 0%.

²Poichè Q_1 è utilizzato come *emitter follower*, il suo effetto sul valore di A in bassa frequenza è trascurabile. La cosa più importante è che Q_1 introduce nell'anello di reazione una nuova frequenza di taglio di tipo passa-basso, di solito abbastanza bassa da creare problemi di stabilità all'anello di reazione, se non opportunamente compensata (vedi par. 15.8).

Questo può non essere un grosso problema per apparecchi di piccola potenza, o per un uso saltuario, come un esperimento in laboratorio. Diventa un fattore determinante quando le potenze in gioco sono grosse e l'uso è di tipo continuativo. Oltre al problema dello spreco energetico dovuto alla scarsa efficienza, c'è da tener presente che l'energia persa finisce in produzione di calore nel transistor regolatore passante. Questo calore deve essere rimosso e dissipato, creando la necessità di grossi radiatori di calore, sistemi di ventilazione, ecc.

Questo fatto, unito alla presenza del trasformatore tradizionale T_1 realizzato in ferro e rame, rende i regolatori lineari di potenza ingombranti, pesanti e costosi.

20.2 Regolatori *switching* (a commutazione)

Nella fig. 20.3 è riportato lo schema di principio di un regolatore di tipo *switching*, che svolge una funzione analoga al regolatore lineare visto nel paragrafo precedente: da una tensione di ingresso V_I non regolata ottenere una tensione di uscita $V_O < V_I$ stabile. Nello schema non viene riportata la parte trasformatore-raddrizzatore-condensatore di filtro, che può anche essere eguale a quella di fig. 20.2, ma che normalmente negli alimentatori switching viene realizzata con una tecnologia diversa.

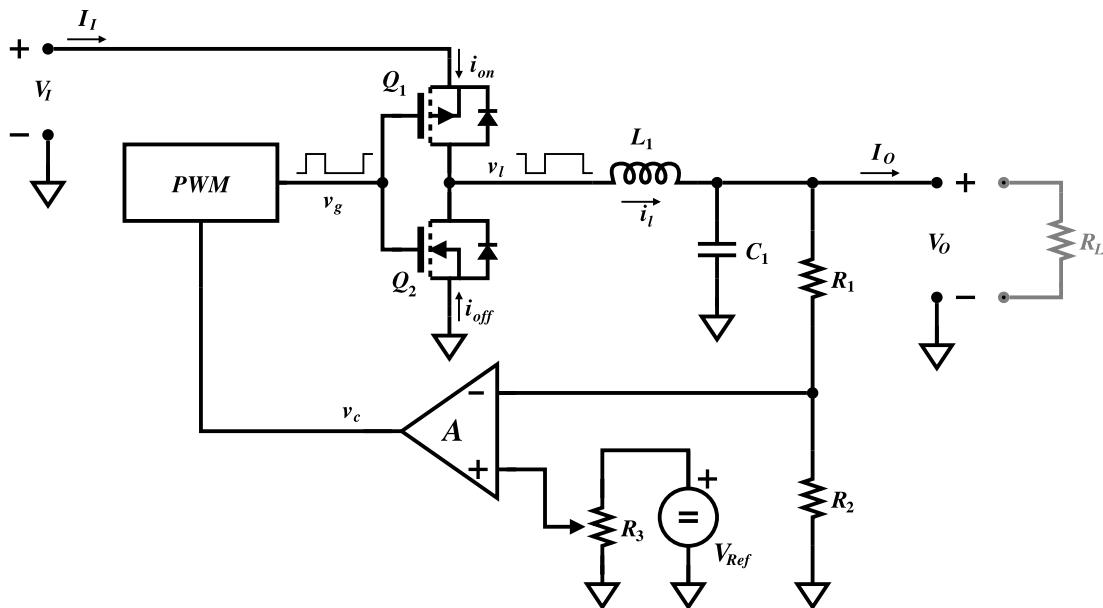


Figura 20.3: Schema di principio di un regolatore di tipo *switching*.

Il blocco indicato con PWM è un *Pulse Width Modulator* (modulatore di larghezza di impulso): un circuito che genera un'onda quadra (o meglio, rettangolare) con periodo T fisso e durata t_h dell'intervallo a livello alto che varia linearmente con la tensione di controllo v_c (fig. 20.4). Il rapporto $\delta = t_h/T$ è il *duty cycle* dell'onda quadra.

Il segnale generato dal PWM comanda i *gate* dei due mosfet complementari Q_1 e Q_2 , che lavorano in controfase: quando Q_1 è acceso, Q_2 è spento e viceversa.

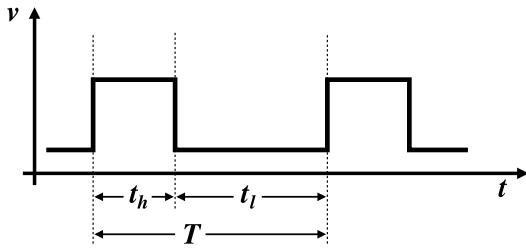


Figura 20.4: Segnale generato da un circuito PWM. Il periodo T è fisso, la parte t_h varia linearmente con la tensione di controllo v_c .

sa, analogamente a quanto si è visto per la porta *NOT* in tecnologia *CMOS* nel par. 10.3.1.

All'uscita della coppia di mosfet ed ingresso dell'induttanza L_1 è quindi presente un'onda rettangolare v_l che oscilla tra 0 e V_I , con gli stessi tempi del segnale del PWM (ma con polarità invertita).

In maniera brutale, ma di fatto corretta, si può dire che il valore medio di v_l è

$$\langle v_l \rangle = V_I \cdot \frac{t_h}{T} = V_I \cdot \delta \quad (20.3)$$

e sovrapposto a questo valore, che rappresenta una tensione continua, si trovano tutte le componenti armoniche che si ottengono dallo sviluppo in serie di Fourier del segnale di fig. 20.4.

L'induttanza L_1 , insieme al condensatore C_1 , forma un filtro passa basso che lascia passare inalterata la componente continua e, se opportunamente dimensionato, sopprime fino al livello desiderato le componenti oscillanti.

Questo passaggio dalla tensione di ingresso V_I alla tensione di uscita $V_O = \langle v_l \rangle$, qualunque sia il valore di V_O , avviene con efficienza molto alta. Nel caso limite di componenti ideali l'efficienza è del 100%: induttanza e condensatore sono componenti reattivi che non possono dissipare energia (a parte le eventuali perdite, che nel caso ideale sono nulle); lo stesso vale per i due transistor mosfet che funzionano come interruttori acceso/spento, quindi idealmente con resistenza 0 o ∞ e di conseguenza senza dissipazione. È chiaramente evidente la differenza rispetto al regolatore lineare, la cui efficienza nel caso limite ideale è $\eta = V_O/V_I$.

Nella pratica, ovviamente, i rendimenti non possono raggiungere il limite ideale, ma in alimentatori switching ben progettati e costruiti con componenti di qualità adeguata possono arrivare fino al 95 ÷ 98%.

20.2.1 Tensioni e correnti in un regolatore *switching*

Per una analisi dettagliata di tensioni e correnti in un regolatore *switching* si può partire dalla rete L_1 , C_1 , R_L , riportata senza i circuiti di controllo in fig. 20.5 a sinistra, schematizzando il PWM ed i transistor interruttori mosfet con il generatore di onde quadre v_l .

I grafici a destra, realizzati integrando numericamente le equazioni del circuito con il programma Spice,³ riportano, dall'alto in basso:

³Il circuito è un semplice LRC risonante, ed agli stessi risultati si può arrivare facilmente anche per via analitica.

- La tensione v_l fornita dal generatore; in questo esempio, un'onda quadra con periodo di $25 \mu s$, *duty cycle* 50%, livello alto $+12 V$ e livello basso $0 V$.
- La corrente i_l che percorre l'induttanza L_1 , per tre diversi valori di C_1 ($0, 5$ e $20 \mu F$).
- La tensione v_o presente ai capi del carico R_L , per gli stessi tre valori di C_1 .

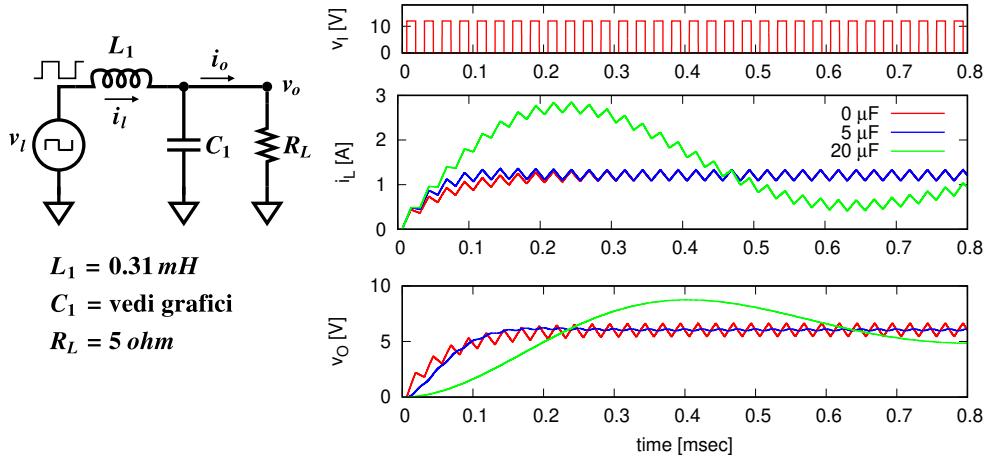


Figura 20.5: Tensioni e correnti nel gruppo LRC di un regolatore *switching*. Il generatore v_l fornisce un'onda quadra con frequenza $40 kHz$ e livelli di tensione $12 V$ (alto) e $0 V$ (basso).

Nel calcolo si è supposto di applicare l'onda quadra v_l al tempo $t = 0$, con i_l e v_o eguali a zero come condizioni iniziali.

Il segnale v_l può essere scomposto come somma di un segnale a gradino che al tempo $t = 0$ passa dal valore 0 a $V_I \cdot \delta$ e di un'onda quadra a media nulla che oscilla tra un valore positivo $V_H = V_I \cdot (1 - \delta)$ ed un valore negativo $V_L = -V_I \cdot \delta$.

Nei grafici della corrente i_l si vede chiaramente la risposta del circuito al segnale a gradino, con sovrapposta una oscillazione periodica dovuta alla componente ad onda quadra. La curva in rosso corrisponde ad avere $C_1 = 0$, quindi al solo rilassamento di un circuito RL , che nell'esempio ha costante di tempo $\tau = L/R = 62 \mu s$.

Nella curva in verde, corrispondente a $C_1 = 20 \mu F$ si vede invece l'inizio di una oscillazione smorzata, risposta del circuito risonante alla sollecitazione data dalla tensione a gradino.

La curva in blu, calcolata per $C = 5 \mu F$, corrisponde ad un comportamento vicino alla condizione di smorzamento critico.

L'oscillazione sovrapposta alla risposta al segnale a gradino ha la stessa frequenza dell'onda quadra v_l ed ampiezza sufficientemente piccola da consentirne il calcolo tramite uno sviluppo in serie di Taylor limitato al primo ordine. Durante l'intervallo t_h la corrente i_l cresce della quantità

$$\Delta i_l = t_h \cdot \frac{V_I - v_o}{L_1} \quad (20.4)$$

mentre durante l'intervallo t_l si ha una diminuzione

$$\Delta i_l = t_l \cdot \frac{v_o}{L_1} \quad (20.5)$$

A regime si raggiunge uno stato stazionario in cui questi due valori si egualano:

$$t_h \cdot \frac{V_I - v_o}{L_1} = t_l \cdot \frac{v_o}{L_1} \quad \Rightarrow \quad v_o = V_I \cdot \frac{t_h}{t_h + t_l} \quad (20.6)$$

in accordo con il risultato già visto con la eq. 20.3.

Dalla eq. 20.5 si può calcolare il valore dell'oscillazione della corrente a regime:

$$\Delta i_l = V_I \cdot \frac{T}{L_1} \cdot (1 - \delta) \delta \quad (20.7)$$

Quindi Δi_l è massima per $\delta = 0.5$, e per un dato δ diminuisce all'aumentare dell'induttanza o della frequenza dell'onda quadra.

In assenza del condensatore C_1 l'oscillazione nella corrente si ripercuote in una analoga oscillazione nella tensione di uscita v_o , come si può vedere nel grafico più basso della fig. 20.5, nella curva in rosso. Le altre due curve dello stesso grafico mostrano come la presenza del condensatore C_1 riduca l'oscillazione tanto più quanto più grande è la capacità.

20.2.2 Problemi di stabilità e velocità

Dalla equazione 20.5 e dalle considerazioni precedenti si vede che si può ridurre l'ondulazione residua su V_O aumentando i valori di L_1 e C_1 oppure riducendo il periodo T . Questo però ha un prezzo:

- Aumentando L_1 e C_1 Viene rallentata la velocità di risposta del circuito: il tempo necessario per adeguare la tensione V_O da un valore stazionario ad un altro aumenta con L_1 e C_1 .
- La funzione di trasferimento del filtro $L_1 - C_1 - R_L$ è un fattore del guadagno dell'anello di controllo del regolatore: è indispensabile tenerne conto nel determinare la condizione di stabilità della reazione negativa.
- La riduzione del periodo T , cioè l'aumento della frequenza dell'onda quadra, fa aumentare le perdite, principalmente nei transistor Q_1 e Q_2 .

Come sempre, la soluzione da cercare è il miglior compromesso tra prestazioni, efficienza e costo, operazione che nella progettazione dei regolatori *switching* è abbastanza complessa.

20.2.3 Il ruolo del transistor Q_2

Osservando attentamente lo schema di fig. 20.3 si può vedere che la freccia della corrente i_{off} che scorre nel transistor Q_2 ha direzione opposta a quella che uno si aspetterebbe nel funzionamento normale di un transistor NMOS.

Durante il tempo t_h la corrente nel transistor Q_1 e nella induttanza L_1 ha la direzione indicata dalle rispettive frecce, scorrendo dal polo positivo verso il polo negativo.

Quando la tensione v_l passa dal valore alto a quello basso del tempo t_l , la corrente i_l non può cambiare di colpo e tanto meno invertirsi: a causa della presenza dell'induttanza, dovrà continuare a scorrere attraverso il transistor Q_2 nel verso indicato nello schema. Questo è possibile perché Q_2 è un transistor di tipo mosfet, generosamente polarizzato per essere in regione lineare con una resistenza del canale molto bassa (vedi par. 9.3 e 9.6).

Fino a che la caduta ohmica prodotta dalla corrente attraverso il canale rimane sufficientemente piccola, cioè tale da non arrivare a polarizzare la giunzione *drain/substrato* in maniera apprezzabile, il dispositivo continua a lavorare in maniera bipolare come indicato nella fig. 9.6. Un transistor bjt nella stessa posizione non potrebbe funzionare.

È possibile anche sostituire il transistor Q_2 con un diodo. Nello schema, in parallelo ai due transistor Q_1 e Q_2 , sono presenti due diodi con funzione di protezione dei transistor dalle inversioni di corrente. Di solito questi diodi sono realizzati direttamente sullo stesso chip del transistor. Se si rimuove il transistor Q_2 e si lascia al suo posto un diodo, il circuito funziona egualmente, con un certo degrado delle prestazioni perché la caduta di potenziale sul diodo di solito è maggiore di quella attraverso il mosfet in conduzione. Utilizzando diodi di tipo Schottky,⁴ si hanno tensioni dell'ordine di 300 mV con il diodo in conduzione; si può avere la metà o meno di questo valore utilizzando invece un buon transistor mosfet.

L'utilizzo del diodo è la soluzione preferita negli apparecchi più economici, quando è accettabile un livello di efficienza energetica un po' minore.

20.2.4 Efficienza energetica

Come già visto nel paragrafo introduttivo 20.2, in condizioni ideali l'efficienza energetica dei regolatori a commutazione tende al 100%. Nel mondo reale i contributi di perdita più importanti sono:

- La dissipazione per effetto Joule della corrente attraverso la resistenza del canale dei due transistor e del filo della induttanza L_1 .
- Le perdite prodotte dalle componenti alternate di i_l nel materiale magnetico su cui è avvolta l'induttanza L_1 e nel dielettrico del condensatore C_1 .
- I tempi di commutazione finiti dei due transistor nel passaggio da acceso a spento e viceversa. Durante la commutazione la potenza istantanea $P(t) = i(t) \cdot v(t)$ dissipata sulla resistenza del canale passa per valori molto alti. Il valore medio, ottenuto integrando su ogni ciclo, può essere tenuto basso solo utilizzando transistor in grado di effettuare commutazioni molto veloci.
- Durante le commutazioni è fisiologico che ci sia una sia pur minima sovrapposizione tra gli stati dei due transistor, che si trovano ad essere entrambi parzialmente accesi, creando per un brevissimo istante un cortocircuito della tensione V_l . Anche questo contributo può essere ridotto da una elevata velocità di commutazione, oltre che da una accurata sincronizzazione degli eventi.

⁴Il diodo Schottky è costruito con una giunzione metallo-semiconduttore invece di una giunzione semiconduttore P-N. È più veloce ed ha una tensione di soglia V_γ più bassa del diodo P-N.

Gli ultimi due punti descrivono entrambi un contributo alle perdite che cresce con la frequenza di commutazione, ponendo dei limiti alla possibilità di affrontare per questa via i problemi di velocità di risposta considerati nel par. 20.2.2.

Nelle applicazioni pratiche le frequenze di commutazione utilizzate vanno dai 20 kHz fino al MHz. Frequenze più basse normalmente vengono evitate perché richiedono l'utilizzo di induttanze di valore più elevato, quindi con nuclei magnetici più grossi e pesanti e maggior quantità di rame negli avvolgimenti. Le frequenze sotto i 20 kHz cadono nello spettro acustico e possono produrre vibrazioni fastidiose nei nuclei magnetici delle induttanze.

20.2.5 Altre configurazioni dei regolatori *switching*

Una caratteristica importante dei regolatori *switching* è che possono essere costruiti in modo da fornire tensioni di uscita V_O più elevate della tensione di ingresso V_I o anche di polarità opposta alla tensione di ingresso, prestazioni in nessun modo possibili con i regolatori lineari.

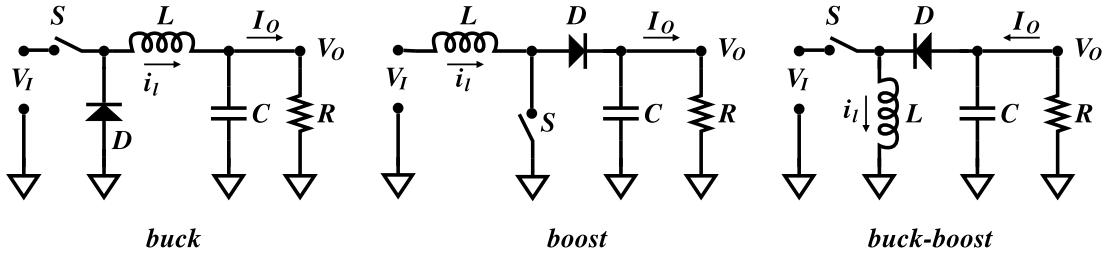


Figura 20.6: Schemi di base dei tre tipi principali di regolatori di tipo *switching*.

Nella fig. 20.6 sono riportati i tre schemi di base corrispondenti ai tre principali modi di operazione. Per semplicità, il transistor Q_1 è stato schematizzato con l'interruttore S , il transistor Q_2 è stato sostituito con il diodo D , come descritto nel par. 20.2.3. Da sinistra a destra:

- *buck converter*: il circuito discusso nei paragrafi precedenti, con

$$V_O = V_I \cdot \delta$$

- *boost converter*: circuito che fornisce una tensione di uscita maggiore di quella di ingresso; ripetendo il calcolo delle eq. 20.4–20.6 si ottiene

$$V_O = \frac{V_I}{1 - \delta}$$

- *buck-boost converter*: circuito che fornisce una tensione di uscita in modulo sia maggiore che minore di quella di ingresso, ma di polarità invertita:

$$V_O = -V_I \cdot \frac{\delta}{1 - \delta}$$

20.3 Trasformatori DC-DC

In teoria un regolatore *switching* potrebbe fare a meno del trasformatore della tensione di rete: il raddrizzatore a ponte della fig. 20.2 potrebbe essere collegato direttamente alla rete e la tensione continua così ottenuta (circa 300 V) potrebbe essere ridotta al valore desiderato, con alta efficienza energetica, utilizzando ad esempio un regolatore di tipo *buck*.

Questa soluzione non dovrebbe (leggi: non deve) essere mai adottata per ragioni di sicurezza: l'apparecchio alimentato si troverebbe collegato direttamente alla rete elettrica, con presenza quindi di tensioni pericolose per l'utente.

D'altra parte il trasformatore tradizionale in ferro e rame è pesante, ingombrante e costoso, e riduce notevolmente i vantaggi dei regolatori *switching*.

Utilizzando circuiti di commutazione simili a quelli visti in questi paragrafi è possibile realizzare trasformatori da corrente continua a corrente continua (trasformatori DC-DC) con efficienza pienamente confrontabile con i trasformatori tradizionali.

Il nocciolo del problema è che il trasformatore di rete lavora alla frequenza di 50 Hz; questo richiede valori di induttanza elevati negli avvolgimenti e nuclei di ferro di sezione grossa per evitare la saturazione magnetica. Convertendo la corrente continua in corrente alternata a frequenza relativamente alta (20 kHz ÷ 1 MHz), è possibile utilizzare nuclei ed avvolgimenti di dimensioni, peso e costo proporzionalmente minori. La fig. 20.7 mostra uno schema di massima di una applicazione di questo tipo.

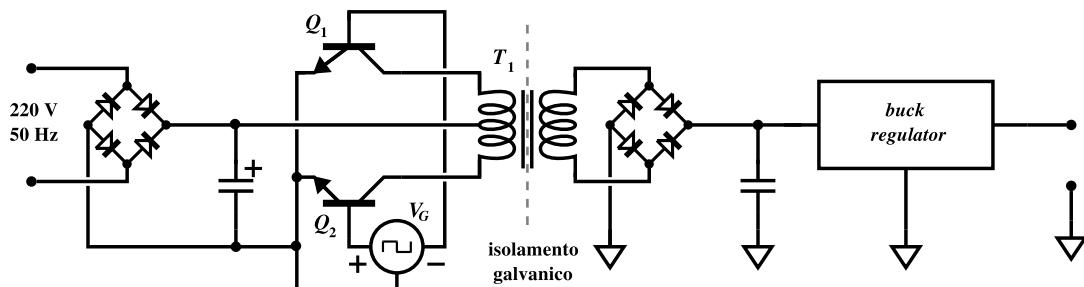


Figura 20.7: Alimentatore *switching* collegato alla rete tramite un convertitore DC-DC.

La tensione di rete viene raddrizzata e filtrata, quindi va ad alimentare i due transistor Q_1 e Q_2 che, lavorando in controfase, inviano un'onda quadra al trasformatore T_1 . I due transistor sono pilotati dal generatore V_G , un oscillatore che produce due onde quadre in opposizione (indicate con i segni + e -), in modo da accendere alternativamente uno dei due transistor e spegnere l'altro.

Il trasformatore T_1 fornisce la separazione galvanica tra la parte sinistra del circuito, che lavora collegata alla rete elettrica, e la parte a destra che alimenta gli apparecchi utilizzatori, raggiungendo lo stesso livello di isolamento che veniva offerto dal trasformatore tradizionale.

La guerra delle correnti

Verso la fine del 1800, con l'inizio della produzione e distribuzione dell'energia elettrica su scala sempre maggiore, si sono confrontate le due alternative, se usare per la distribuzione corrente continua o corrente alternata. Gli episodi di questo confronto sono passati alla storia con il nome di "guerra delle correnti".

A prevalere è stato l'utilizzo della corrente alternata, perché la disponibilità dei trasformatori permetteva di effettuare il trasporto di energia su lunghe distanze ad alta tensione riducendo le perdite, mantenendo la possibilità di riottenere all'arrivo la bassa tensione necessaria per le applicazioni domestiche. La corrente continua non poteva disporre di un apparecchio equivalente, che consentisse con la stessa praticità ed efficienza una trasformazione analoga.

Tale possibilità viene offerta oggi dai convertitori DC-DC funzionanti sul principio descritto nel paragrafo precedente, con lo stesso livello di efficienza dei trasformatori tradizionali.

Come conseguenza, si sta diffondendo l'utilizzo della corrente continua nel trasporto di energia elettrica in quelle situazioni in cui la corrente alternata è soggetta a perdite maggiori, come ad esempio nelle condotte sottomarine.

20.4 Amplificatori di potenza

Per un amplificatore di segnale il rendimento energetico, date le piccole potenze in gioco, non è il parametro più importante. A seconda dell'uso cui il circuito è destinato, possono prevalere altre esigenze, come il livello di rumore, l'assenza di distorsione, la banda passante, la precisione e stabilità dei parametri, ecc. Il consumo diventa importante solo nelle applicazioni con alimentazione a batteria, dove interessa una autonomia più lunga possibile.

Il discorso è completamente diverso per gli amplificatori di potenza. In questo caso il rendimento energetico diventa uno dei parametri più rilevanti, non solo per il maggior consumo di energia elettrica che deriva da un basso rendimento, ma anche perché, come visto per i regolatori lineari, l'energia persa finisce in riscaldamento dei transistor. Il calore prodotto deve essere rimosso tramite radiatori e sistemi di ventilazione, complicando la realizzazione degli apparecchi. Inoltre temperature di lavoro più elevate riducono la vita utile dei semiconduttori.

All'epoca dei tubi a vuoto gli amplificatori di potenza sono stati catalogati in tre classi di funzionamento: A, B, C. Con l'introduzione dei semiconduttori si è aggiunta la classe D e varie altre classi, non sempre ben definite.

20.4.1 Classe A

Un amplificatore in classe A utilizza un singolo dispositivo amplificatore che gestisce interamente il segnale nei cicli sia positivi che negativi.

Prima dell'introduzione dei transistor l'amplificazione di potenza in classe A era la regola negli apparecchi domestici per riproduzione acustica, come radio e fonovaligie, per potenze limitate a pochi watt.

Con il passaggio dai tubi a vuoto ai semiconduttori l'uso di amplificatori finali in classe A è scomparso completamente, sostituito dai circuiti in classe B, più convenienti sotto tutti i punti di vista.

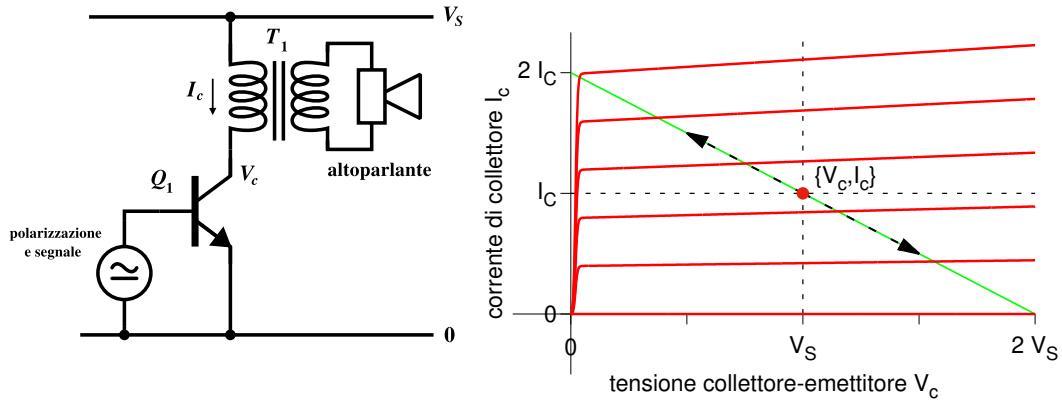


Figura 20.8: Ipotetico amplificatore finale in classe A con l'utilizzo di un transistor. Le curve caratteristiche nel grafico a destra descrivono il funzionamento elettrico del circuito.

Le parti essenziali di un amplificatore audio in classe A a transistor sono riportate nello schema in fig. 20.8.

È indispensabile utilizzare un trasformatore, componente costoso ed ingombrante, per accoppiare l'altoparlante al transistor, perché non è bene fare scorrere in un altoparlante la corrente continua del punto di lavoro del transistor. Già questo fatto rende il circuito poco pratico e poco conveniente.

Al punto di lavoro, indicato nel grafico delle curve caratteristiche del transistor dal punto rosso all'incrocio delle due linee tratteggiate, ed in assenza di segnale, la tensione tra collettore ed emettitore è pari alla tensione di alimentazione \$V_S\$. Con una corrente di collettore \$I_C\$ si ha una potenza assorbita (e dissipata nel transistor) \$P_a = I_C \cdot V_S\$, trascurando nel calcolo la resistenza ohmica dell'avvolgimento primario del trasformatore.

In presenza di segnale la corrente di collettore diventa \$I_c = I_C + i_c\$ e la tensione sul collettore diventa \$V_c = V_S + v_c\$.

Per il massimo dell'efficienza si deve rispettare la condizione \$R_L = V_S/I_C\$, dove \$R_L\$ è la resistenza dell'utilizzatore (altoparlante) trasferita dal trasformatore verso il transistor. In questo modo il punto \$\{V_c, I_c\}\$ si sposta nel piano secondo le due frecce lungo la linea verde, che rappresenta la retta di carico del circuito. La corrente \$I_c\$ può oscillare tra i valori 0 e \$2I_C\$, e la tensione \$V_c\$ tra 0 e \$2V_S\$.

Con un segnale sinusoidale di ampiezza \$V\$ ai capi del trasformatore la potenza media fornita è

$$\langle P_o \rangle = \frac{V_{eff}^2}{R_L} = \frac{V^2}{2R_L} \quad (20.8)$$

La potenza media assorbita invece rimane sempre la stessa:

$$\langle P_a \rangle = \langle I_C + i_c \rangle \cdot V_S = I_C \cdot V_S \quad (20.9)$$

poichè i_c è a media nulla, per cui il rendimento risulta

$$\eta = \frac{\langle P_o \rangle}{\langle P_a \rangle} = \frac{V^2}{2R_L} \cdot \frac{1}{I_C V_S} = \frac{V}{2V_S} \quad (20.10)$$

Quando V ha il valore massimo possibile ($= V_S$), allora $\eta = 50\%$. Nel caso invece che il segnale sia un'onda quadra, $V_{eff} = V$ e $P_o = V^2/R_L$, per cui η varia da 0 al 100%, secondo l'ampiezza del segnale.

Il segnale audio musicale si colloca in una posizione intermedia tra sinusoida ed onda quadra, e così anche il rendimento.

Ma la cosa più antipatica dell'amplificatore in classe A è il fatto che il consumo $P_a = I_C \cdot V_S$ rimane costante, anche in assenza di segnale.

20.4.2 Classe B

Negli amplificatori in classe B si hanno sempre due dispositivi che lavorano in modo complementare, uno sulla semionda positiva del segnale, l'altro sulla semionda negativa, come nel circuito emitter follower complementare di fig. 7.25. In questo modo, in assenza di segnale, il consumo è nullo o comunque molto ridotto.⁵

Supponendo nuovamente un segnale sinusoidale, di periodo T , la potenza media assorbita ad esempio durante la semionda positiva, indicando con V_S la tensione di alimentazione e con I_S il valore di picco della corrente, è:

$$\langle P_a^+ \rangle = \frac{2V_S I_S}{T} \int_0^{T/2} \sin(2\pi t/T) dt = \frac{2}{\pi} V_S I_S \quad (20.11)$$

Data la simmetria del circuito, lo stesso risultato si ottiene durante la semionda negativa quando lavora l'altro transistor. Quindi $\langle P_a \rangle = \langle P_a^+ \rangle$.

Nelle stesse condizioni la potenza media fornita alla resistenza di carico R_L è:

$$\langle P_L \rangle = \frac{R_L I_S^2}{T} \int_0^T \sin^2(2\pi t/T) dt = \frac{R_L I_S^2}{2} \quad (20.12)$$

ed il rendimento

$$\eta = \frac{\langle P_L \rangle}{\langle P_a \rangle} = \frac{R_L I_S^2}{2} \cdot \frac{\pi}{2V_S I_S} = \frac{\pi}{4} \frac{I_S}{I_M} \quad (20.13)$$

avendo definito $I_M = V_S/R_L$ il valore massimo possibile per la corrente I_S (anche qui, trascurando le tensioni di saturazione dei transistor).

Il rendimento massimo teorico per l'amplificatore in classe B con segnale sinusoidale è quindi $\eta = \pi/4 \simeq 78\%$. Anche qui, nel caso di un'onda quadra con la massima ampiezza possibile, si ha $\eta = 100\%$.

L'amplificatore in classe B ha quindi diversi vantaggi rispetto a quello in classe A:

- Consumo nullo o minimo in assenza di segnale.

⁵Per ovviare ai problemi di distorsione che si presentano all'*incrocio* quando un dispositivo si spegne e l'altro si accende sono stati sviluppati diversi schemi di polarizzazione che prevedono una parziale sovrapposizione nel funzionamento dei due dispositivi. Quindi si sono avuti, specialmente con i tubi a vuoto, amplificatori in classe AB, AB₁, AB₂ ... come tentativi di trovare il miglior compromesso tra efficienza energetica e basso livello di distorsione.

- Nessun bisogno di trasformatore.
- Rendimento più elevato (78% invece di 50% nelle condizioni ideali indicate nel calcolo).

Può sembrare che la differenza di rendimento indichi un vantaggio tutto sommato modesto. La cosa va considerata dal punto di vista dei transistori. Nelle condizioni migliori, per una certa potenza di uscita P il circuito in classe A dissipava la stessa potenza P nei transistori; quello in classe B, invece, dissipava nei transistori $P \cdot (4/\pi - 1) \approx 0.27P$, cioè poco più di un quarto di quello in classe A, con relativi minori problemi di riscaldamento.

Se però si ipotizza di avere un rendimento medio di quest'ordine per un impianto audio da 8 kW, valore non insolito in una manifestazione musicale, a fronte di un consumo medio di 10.2 kW, se ne avrebbero 2.2 dissipati in calore. È evidente in questi casi l'utilità di poter disporre di circuiti di potenza con efficienza maggiore.

20.4.3 Classe C

Gli amplificatori in classe C consentono di ottenere livelli di rendimento molto elevati, ma per il loro principio di funzionamento sono limitati esclusivamente ad applicazioni a radiofrequenza a banda stretta.

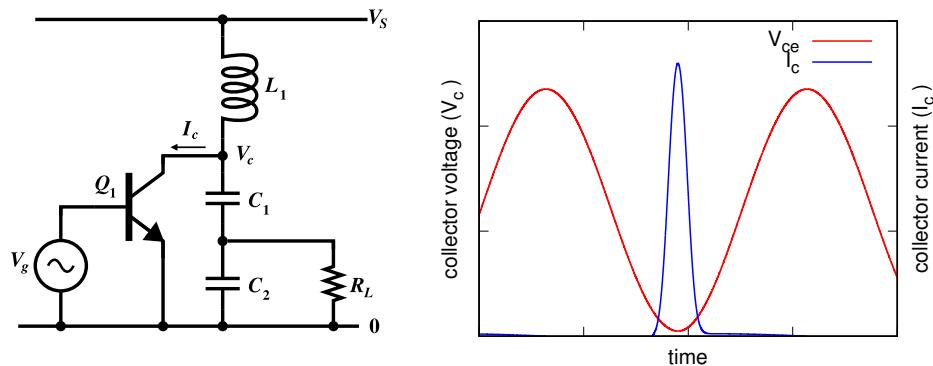


Figura 20.9: Schema di principio di un amplificatore in classe C e, a destra, curve di corrente e tensione di collettore durante il funzionamento.

In fig. 20.9 è riportato uno schema di amplificatore in classe C con gli elementi essenziali, insieme ai grafici di corrente e tensione del transistor durante il funzionamento.

La prima cosa da notare è che il transistor Q_1 non è polarizzato. Il generatore V_g fornisce solo una tensione alternata a media nulla, che porta il transistor in conduzione per una breve parte del ciclo durante i picchi positivi, quando $V_g > V_\gamma$. Questo lo si vede nel grafico della corrente I_c (curva blu): la corrente scorre soltanto a brevi impulsi. Nonostante questo, la tensione sul collettore del transistor (curva rossa) ha un andamento sinusoidale praticamente puro, grazie alla presenza del circuito risonante $L_1-C_1-C_2$.

Il funzionamento è molto simile a quello di un'altalena: l'oscillazione viene avviata e mantenuta mediante piccole spinte periodiche applicate in sincronia con il periodo dell'oscillazione libera.

Se il circuito risonante è correttamente sintonizzato, l'impulso di corrente viene a coincidere con il minimo della tensione tra collettore ed emettitore, minimizzando la potenza dissipata nel transistor, data dal prodotto $I_c \cdot V_c$ e massimizzando la potenza ceduta al circuito risonante data da $I_c \cdot (V_S - V_c)$.

Lo spettro di Fourier della corrente è ovviamente ricchissimo di armoniche, ma se il fattore di merito Q_0 del circuito risonante, detto *circuito volano* (in inglese: *tank circuit*), è sufficientemente elevato, queste vengono filtrate via e ridotte ad un livello accettabile, come si vede dal grafico della tensione V_c .

In questo schema, come già visto nei circuiti del par. 16.7, il circuito volano svolge anche la funzione di adattamento di impedenza tra la resistenza di carico R_L e le caratteristiche di uscita del transistor.

Il rendimento degli amplificatori in classe C può arrivare al 90% ed oltre, e questo li rende preziosi nella produzione di segnali di elevata potenza come i trasmettitori radio o i generatori di radiofrequenza negli acceleratori di particelle.

A fronte della loro apparente semplicità, gli amplificatori in classe C sono oggetti critici e delicati.

Una variazione significativa del carico dato dalla resistenza R_L (trasformata dal circuito volano) o la perdita di sintonia del circuito volano stesso può portare ad una rapida distruzione del dispositivo amplificatore: se il circuito va fuori sintonia, il picco di corrente può venire a coincidere con un valore elevato della tensione di collettore, portando ad un aumento della dissipazione fino ad un ordine di grandezza. Un amplificatore in classe C deve necessariamente essere dotato di circuiti e sistemi di protezione efficienti.

20.4.4 Classe D

Nelle applicazioni di media ed alta potenza a frequenze audio si sono affermati circuiti denominati amplificatori in classe D, funzionanti sullo stesso principio del regolatore *switching* del par. 20.2.

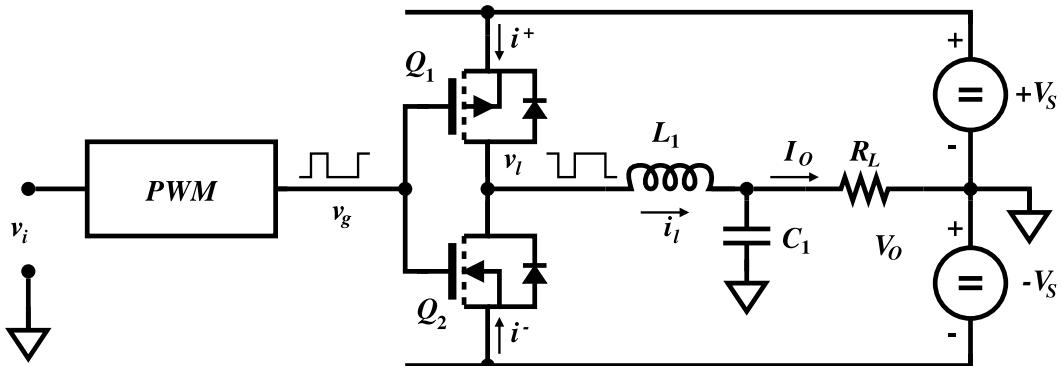


Figura 20.10: Schema di principio di un amplificatore in classe D.

Gli elementi essenziali dello stadio di potenza di questi amplificatori sono riportati in fig. 20.10. La differenza principale rispetto al circuito di fig. 20.3 è nella alimentazione, che in questo caso è simmetrica, con tensioni eguali ed opposte $\pm V_S$, con il carico R_L collegato tra l'uscita del filtro L_1C_1 ed il nodo di massa, in comune tra i due alimentatori.

In queste condizioni, con un ragionamento simile a quello utilizzato per scrivere l'eq. 20.3, si può vedere che si ha

$$\langle v_l \rangle = 2(\delta - 0.5) \cdot V_S \quad (20.14)$$

ed all'uscita del filtro $v_o = \langle v_l \rangle$, con una parte di ondulazione residua.

Utilizzando un circuito PWM con una dipendenza lineare del *duty cycle* dalla tensione v_i , ad esempio

$$\delta = 0.5 + \frac{v_i}{k} \quad (20.15)$$

si ha per il funzionamento di tutto il circuito la relazione lineare

$$v_o = v_i \cdot \frac{2V_S}{k} \quad (20.16)$$

Perchè tutto questo funzioni, però, è indispensabile che la frequenza f_q dell'onda quadra generata dal PWM sia molto maggiore della componente di frequenza più alta del segnale v_i .

Tipicamente,⁶ per la riproduzione di segnali audio che possono arrivare nominalmente fino a 20 kHz si utilizzano onde quadre con frequenza dell'ordine di $f_q = 300$ kHz. Questo permette di adottare per il filtro passa basso una frequenza di taglio che sia sufficientemente minore di f_q , tale ridurre in maniera efficace l'ondulazione residua, ma superiore a 20 kHz, per non filtrare via anche le componenti veloci di v_i .

Per il rendimento energetico di questi circuiti valgono le stesse considerazioni viste nel par. 20.2.4.

⁶A meno di non utilizzare circuiti PWM basati su algoritmi più complessi, come i convertitori delta-sigma.

Appendice A

Valori dei componenti

I valori di resistenze e condensatori utilizzati in un circuito elettronico spaziano su un campo molto ampio, dai centesimi di ohm ai terahom per le resistenze ($10^{-2} \div 10^{12} \Omega$), dal pF fino alle decine di F per i condensatori ($10^{-12} \div 10^2 F$), arrivando a coprire 14 ordini di grandezza.

In commercio sono disponibili resistenze e condensatori con serie di valori standard dipendenti dalla tolleranza richiesta. Ogni serie è identificata da una sigla composta dalla lettera E seguita da un numero che indica quanti sono i valori, spaziati logaritmicamente, per ogni decade. Le serie commerciali più comuni sono E6, E12, E24, E48, E96, E192.

La tabella che segue riporta i valori per le serie E12, E24 ed E48, che hanno passo tra due valori successivi rispettivamente di 1.21, 1.10 e 1.05.

E12	E24	E48
100	100	100 105
	110	110 115
120	120	121 127
	130	133 140
150	150	147 154
	160	162 169
180	180	178 187
	200	196 205

E12	E24	E48
220	220	215 226
	240	237 249
270	270	261 274
	300	287 301
330	330	316 332
	360	348 365
390	390	383 402
	430	422 442

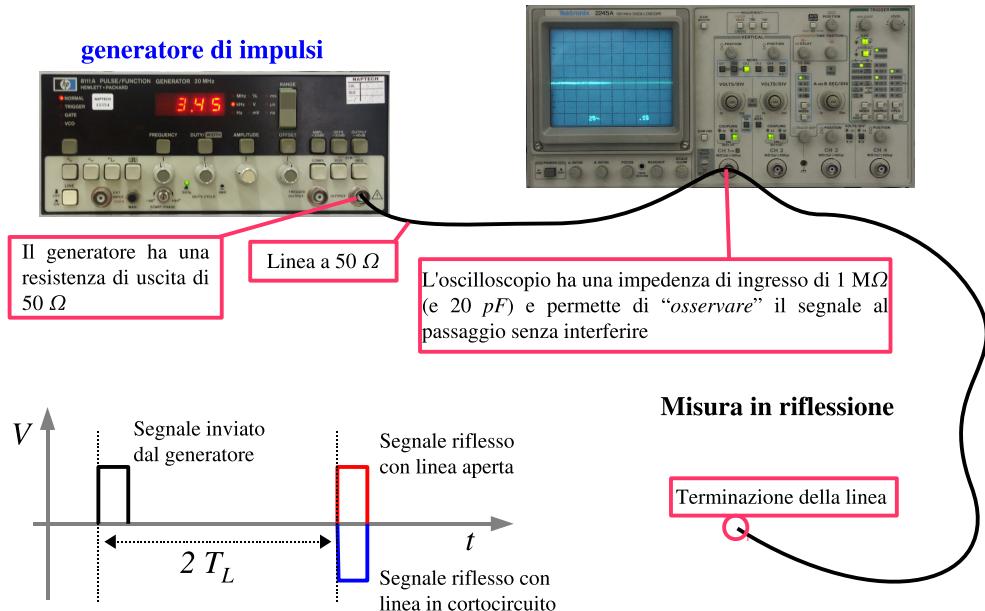
E12	E24	E48
470	470	464 487
	510	511 536
560	560	562 590
	620	619 649
680	680	681 715
	750	750 787
820	820	825 866
	910	909 953

A seconda della serie utilizzata è quindi possibile disporre di ogni valore entro una tolleranza che va da $10^{1/6} = \pm 24\%$ per la E6 a $10^{1/192} = \pm 0.6\%$ per la E192.

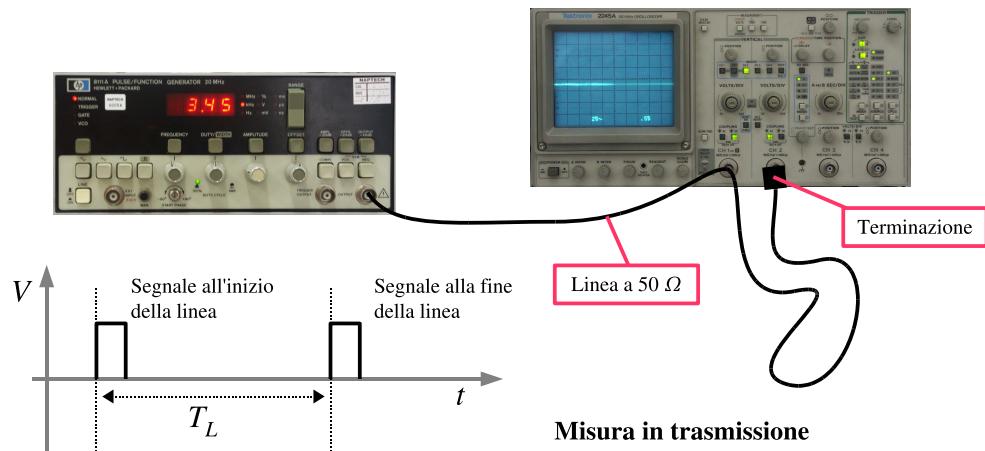
Appendice B

Misure su linee di trasmissione

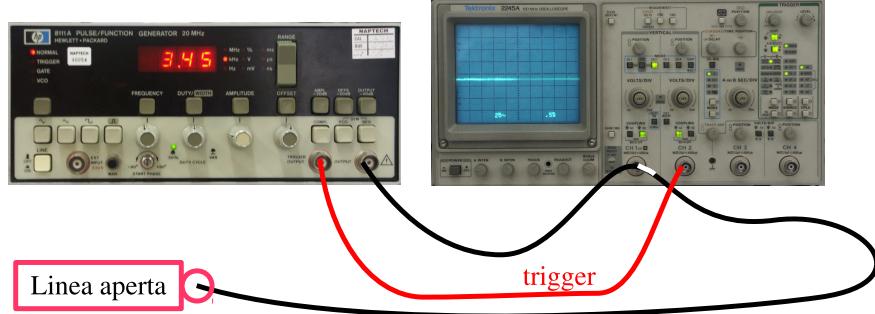
Misura della lunghezza elettrica di una linea di trasmissione oscilloscopio



- 1) Misurare, in trasmissione o in riflessione, la lunghezza elettrica T_L della linea.
- 2) Dal valore di T_L e dalla lunghezza geometrica calcolare la velocità di propagazione dei segnali lungo la linea e la costante dielettrica dell'isolante interno.
- 3) Terminare la linea con una resistenza variabile ($0 \dots 100\Omega$); dalla condizione di riflessione nulla verificare l'impedenza caratteristica.



Misura della trasformazione di impedenza con una linea di trasmissione

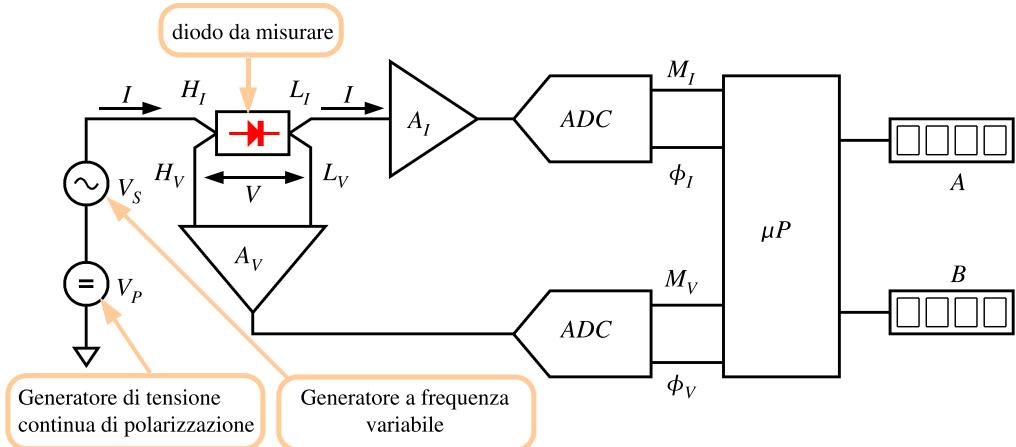


- 4) Misurare l'impedenza complessa della linea aperta per alcune frequenze, comprese tra 0 e $1/T_L$. Tra la tensione V_0 che si ha all'oscilloscopio con la linea non collegata e la tensione V_1 che si ha con la linea collegata vale la relazione:

$$V_1 = V_0 \cdot \frac{Z_L}{50\Omega + Z_L}$$

Dai valori complessi (modulo e fase) di V_0 e V_1 calcolare il valore (complesso) di Z_L (sincronizzare oscilloscopio e generatore tramite *trigger*).

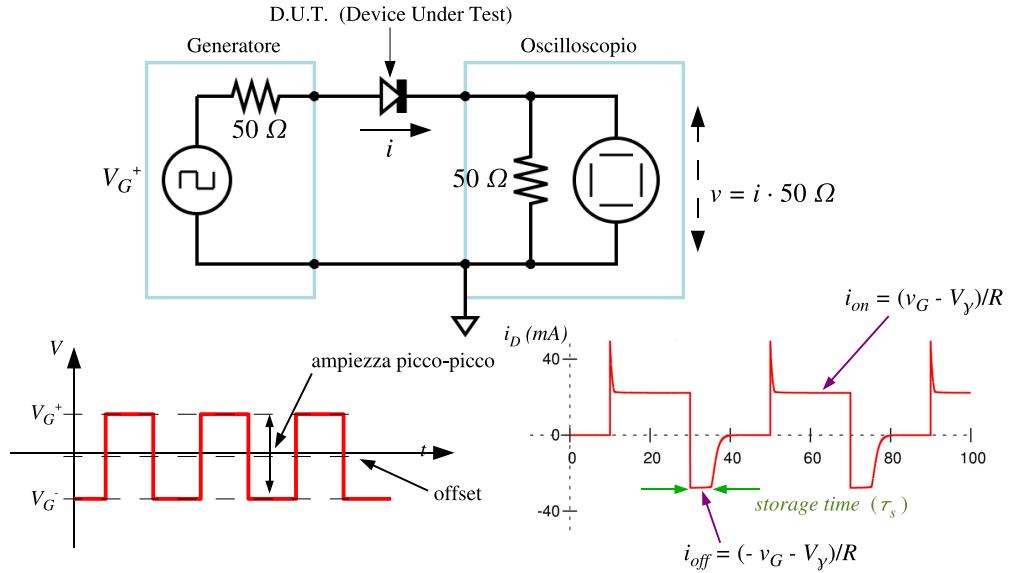
Misura della capacita' di transizione di un diodo



Misurare il valore di C_T per $V_P = -1 \dots -30 V$ per frequenze di $100 \text{ kHz} \dots 10 \text{ MHz}$. Confrontare i dati ottenuti con il modello teorico:

$$C_T = C_a + \frac{C_0}{\sqrt{1 - V_D/V_J}}$$

Misura dei tempi di commutazione di un diodo

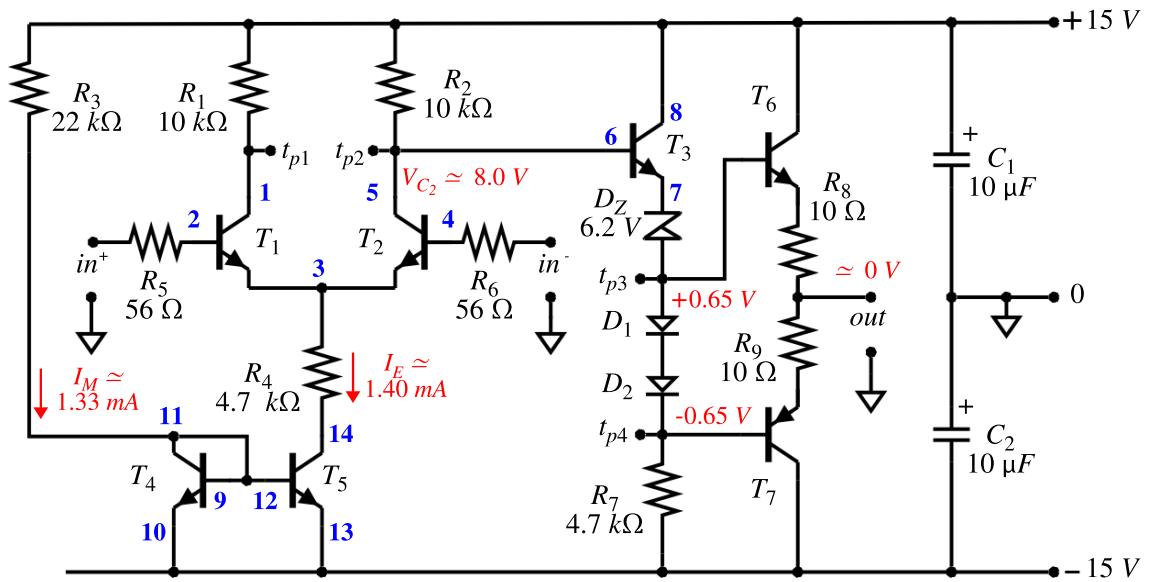


Misurare lo *storage time* t_s per diversi valori di corrente I_F (ad es. 50 ... 100mA) e diversi rapporti I_F/I_R (0.5 ... 1... 2).

Confrontare i dati ottenuti con il modello teorico

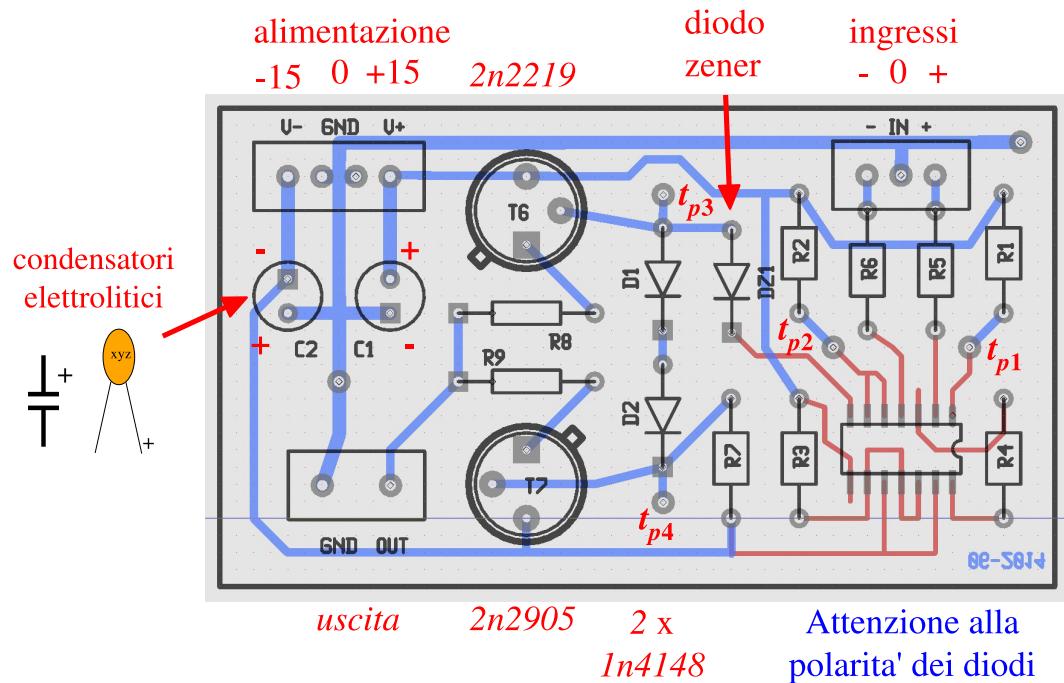
$$t_s = \tau \log \left(1 + \frac{I_F}{I_R} \right)$$

Amplificatore differenziale con stadio di uscita *emitter follower*



I transistor $T_1 \dots T_5$ sono contenuti nel chip CA3046.
 T_6 : 2n2219; T_7 : 2n2905; D_1, D_2 : 1N4148.

Circuito stampato visto dal lato componenti

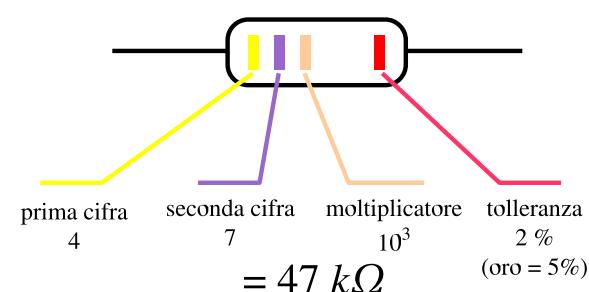


1. Montare il circuito, esclusi i transistor T_6 e T_7 .
2. Con il circuito non collegato, regolare due sezioni dell'alimentatore al valore 15 V, spegnere l'alimentatore e collegare il circuito.
3. Accendere l'alimentatore e verificare che tutte le tensioni siano corrette. In particolare le tensioni ai capi dei diodi D_1 e D_2 devono essere di circa 0.65 V.
4. Ad alimentazione spenta, completare il circuito montando i transistor T_6 e T_7 .
5. Misurare la tensione di *offset* all'uscita e verificare che sia compresa nei valori previsti.
6. Misurare l'amplificazione di modo comune e l'amplificazione di modo differenziale in bassa frequenza ($f \simeq 1\text{kHz}$).
7. Misurare l'amplificazione di modo comune e l'amplificazione di modo differenziale in funzione della frequenza.
8. Misurare l'ampiezza massima dei segnali di ingresso che non danno luogo a distorsione apprezzabile, sia in modo comune che in modo differenziale.
9. Misurare la risposta all'onda quadra dell'amplificatore. Confrontare le costanti di tempo ottenute con le caratteristiche della banda passante determinata in precedenza.
10. Misurare la risposta all'onda quadra del solo primo stadio dell'amplificatore, prelevando il segnale dal collettore dei transistor T_1 e T_2 con una sonda sia in modalità $\times 1$ sia $\times 10$. Confrontare i risultati ottenuti con quelli dell'amplificatore completo.
11. Misurare le impedenze di ingresso e di uscita agli estremi ed al centro della banda passante (chiudendo l'amplificatore su una resistenza di carico da 50Ω il segnale di uscita si riduce a ...). Inserendo una resistenza da $10k\Omega$ in serie all'ingresso il segnale di uscita si riduce a ...).
12. Misurare il contenuto armonico per diverse ampiezze del segnale di uscita ($1V_{pp}, 2V_{pp}, 4V_{pp}, \dots$).
13. Riportare in una tabella i risultati ottenuti, specificando le condizioni di misura.
14. Confrontare i risultati ottenuti con i valori previsti dai modelli.

Codice dei colori per i valori dei componenti elettronici

Valori standard
della serie E12

0	nero	
1	marrone	
2	rosso	
3	arancio	
4	giallo	
5	verde	
6	blu	
7	viola	
8	grigio	
9	bianco	



1.0

1.2

1.5

1.8

2.2

2.7

3.3

3.9

4.7

5.6

6.8

8.2

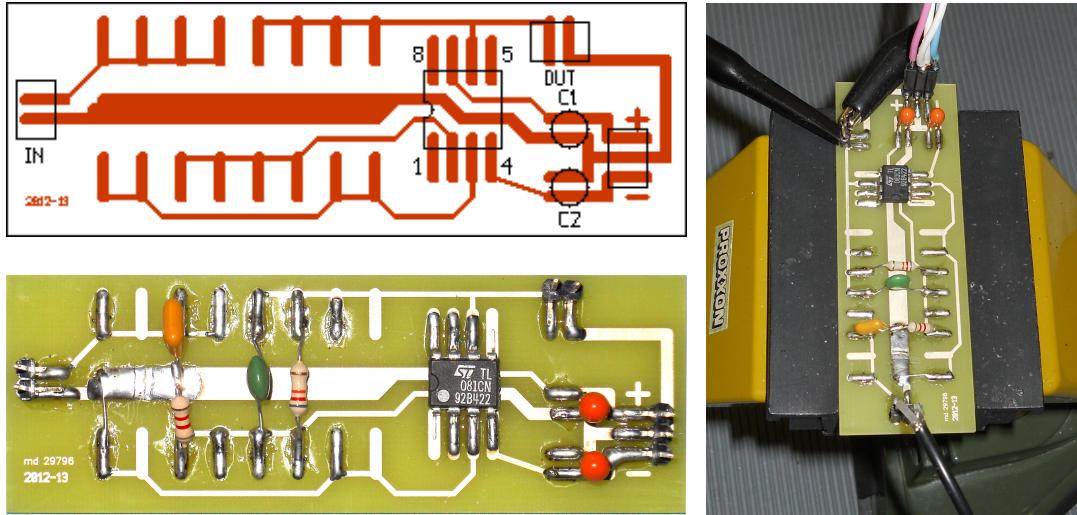


Figura B.1: Montaggio di amplificatore operazionale su circuito stampato.

Amplificatore operazionale

- 1) Montare sul circuito stampato l'amplificatore operazionale, i condensatori ed i connettori per alimentazione ed ingresso/uscita dei segnali. La chiave di riferimento sul contenitore dell'amplificatore deve essere allineata con quella nel disegno in fig. B.1 in alto. C_1 e C_2 sono due condensatori *elettrolitici al tantalio* e sono *polarizzati*. Il polo positivo corrisponde al terminale più lungo e si trova a destra quando il condensatore viene visto dal lato dove è scritto il valore tenendo i terminali rivolti verso il basso. I condensatori elettrolitici, se montati con polarità invertita, *possono esplodere*.
- 2) Completare il circuito secondo uno degli schemi in fig. B.2, utilizzando le piazzole come nell'esempio in fig. B.1 in basso. Usando l'oscilloscopio, misurare la deriva della tensione di uscita dv_o/dt e da questa ricavare la corrente di polarizzazione (*bias*) $I_b = C \cdot dv_o/dt$. $C = 100\ldots 1000 \text{ pF}$.

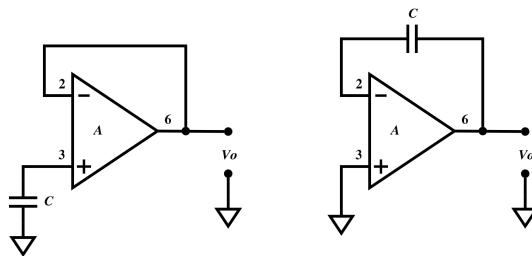


Figura B.2: Circuiti per la misura della corrente di polarizzazione (*bias*).

- 3) Montare il circuito amplificatore di fig. B.3 e misurare l'amplificazione di tensione e la frequenza di taglio a 3 dB per diversi valori di R_2/R_1 , tra 1 e 100. Per

ogni misura calcolare il prodotto *Guadagno* \times *Banda Passante* e confrontare i risultati ottenuti.

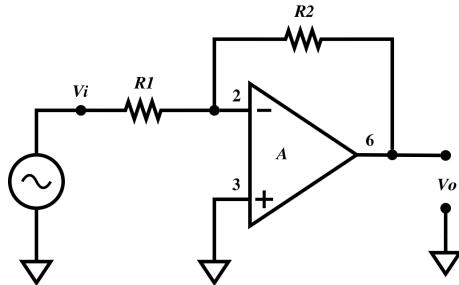


Figura B.3: Amplificatore invertente. $R_1 = 2.2\ldots47 \text{ k}\Omega$, $R_2 = 10 \text{ k}\Omega\ldots1 \text{ M}\Omega$.

- 4) Montare il circuito amplificatore di fig. B.4 e misurare l'amplificazione di tensione in funzione della frequenza tra 2 Hz e 200 kHz . Determinare le frequenze di taglio inferiore f_l e superiore f_h e gli andamenti del guadagno per $f < f_l$, $f_l < f < f_h$, $f > f_h$.

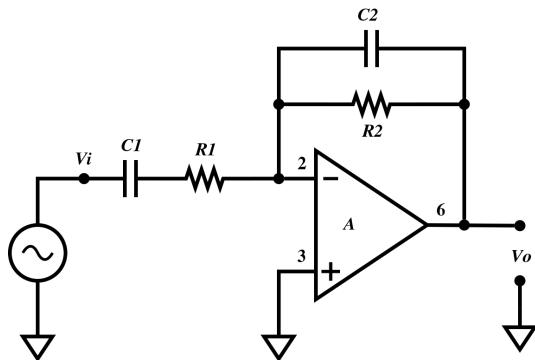
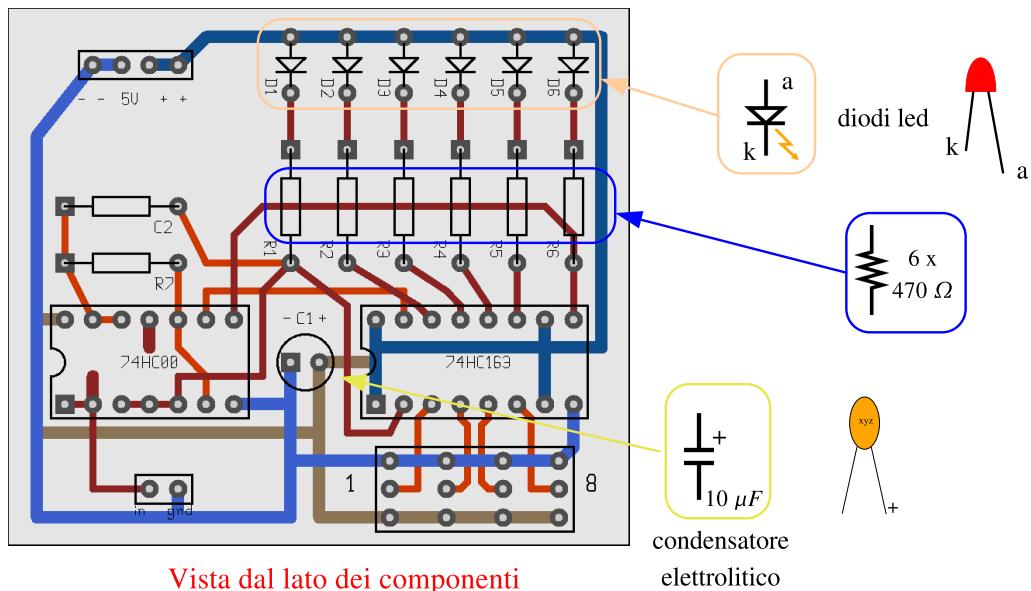
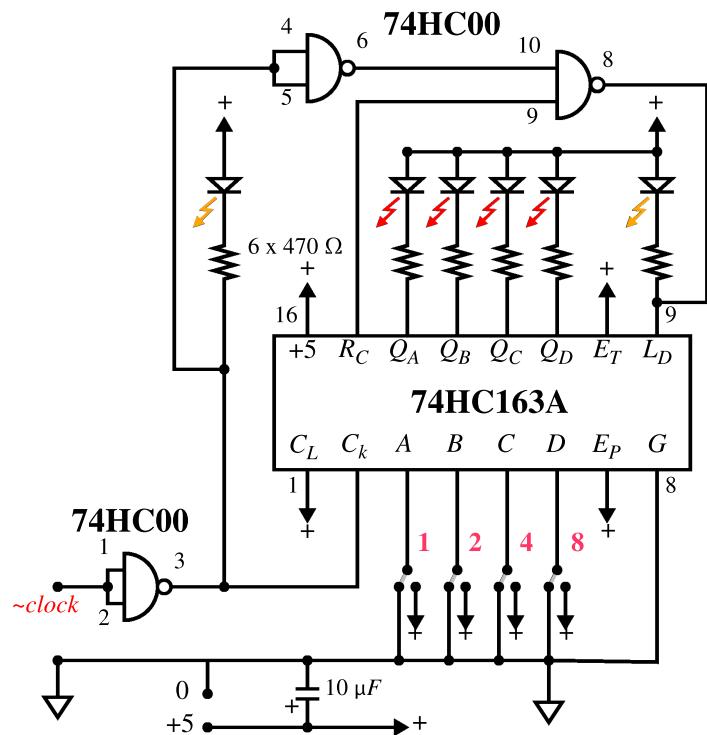


Figura B.4: Amplificatore selettivo. Valori indicativi per i componenti: $R_1 = 8.2 \text{ k}\Omega$, $R_2 = 82 \text{ k}\Omega$, $C_1 = 1 \mu\text{F}$, $C_2 = 100 \text{ pF}$.

Per il montaggio e per l'esecuzione delle misure tenere sempre fermo il circuito stampato utilizzando il morsetto da banco, come riportato ad esempio in fig. B.1 a destra.

Contatore programmabile modulo N ($1 \leq N \leq 16$)



Collegare un generatore all'ingresso $\sim\text{clock}$; regolare per una frequenza bassa ($< 10\text{Hz}$) e verificare che la sequenza di conteggio sia corretta, per diversi valori di N . Aumentare la frequenza di clock e misurare (con l'oscilloscopio) i ritardi tra l'ingresso C_k e le uscite Q ed RC ed il tempo di propagazione attraverso le porte $NAND$. Aumentare ancora la frequenza e determinare (se possibile) la frequenza massima di conteggio.