

Massimiliano Pieraccini

FONDAMENTI DI ELETTRONICA

Pearson Learning Solution

CODICE STUDENTE

ISBN 9788865188088B



Validità codice 31/12/2023

L.1742

Codice di accesso a **eText**

FONDAMENTI DI ELETTRONICA

Massimiliano Pieraccini

FONDAMENTI DI ELETTRONICA

Le informazioni contenute in questo libro sono state verificate e documentate con la massima cura possibile. Nessuna responsabilità derivante dal loro utilizzo potrà venire imputata agli Autori, a Pearson Italia S.p.A. o a ogni persona e società coinvolta nella creazione, produzione e distribuzione di questo libro.

Per i passi antologici, per le citazioni, per le riproduzioni grafiche, cartografiche e fotografiche appartenenti alla proprietà di terzi, inseriti in quest'opera, l'editore è a disposizione degli aventi diritto non potuti reperire nonché per eventuali non volute omissioni e/o errori di attribuzione nei riferimenti.

Le fotocopie per uso personale del lettore possono essere effettuate nei limiti del 15% di ciascun volume/fascicolo di periodico dietro pagamento alla SIAE del compenso previsto dall'art. 68, commi 4 e 5, della legge 22 aprile 1941 n. 633.

Le fotocopie effettuate per finalità di carattere professionale, economico o commerciale o comunque per uso diverso da quello personale possono essere effettuate a seguito di specifica autorizzazione rilasciata da CLEARedi, Centro Licenze e Autorizzazioni per le Riproduzioni Editoriali, Corso di Porta Romana 108, 20122 Milano, e-mail autorizzazioni@clearedi.org e sito web www.clearedi.org.

Realizzazione editoriale: Andrea Pessini

Revisione tecnica: Nicola Sommella

Grafica di copertina: Maurizio Garofalo

Stampa: TIP.LE.CO. - S. Bonico (PC)

Tutti i marchi citati nel testo sono di proprietà dei loro detentori.

9788865188088

Printed in Italy

1^a edizione: settembre 2014

Ristampa

00 01 02 03 04

Anno

14 15 16 17 18

Sommario

Prefazione	1
Capitolo 1 Fisica dei semiconduttori	5
1.1 La carica elettrica	6
1.2 Tensione	7
1.3 Corrente	10
1.4 Legge di Ohm	11
1.5 Isolanti e conduttori	17
1.6 Semiconduttori	20
1.7 Elettroni nei semiconduttori	23
1.8 Lacune	24
1.9 Effetto tunnel	24
1.10 Dualità onda-corpuscolo	26
1.11 Silicio intrinseco	29
1.12 Silicio drogato con impurità di tipo n	30
1.13 Silicio drogato con impurità di tipo p	32
1.14 Legge di azione di massa	32
1.15 Correnti di diffusione	35
1.16 Potenziale di built-in	37
1.17 Energia di Fermi	42
1.18 Tecnologia del silicio monocristallino	47
1.19 Storia di un materiale dalle proprietà insolite e poco riproducibili	48
1.20 Riepilogo	49
1.21 Esercizi svolti	51
1.22 Domande di verifica a scelta multipla	55
1.23 Domande di verifica a risposta aperta	60
1.24 Risposte alle domande a scelta multipla	61

Capitolo 2	La giunzione pn	63
2.1	Dalla galena alla giunzione pn	64
2.2	Realizzazione di una giunzione pn	66
2.3	Il diodo	66
2.4	Retta di carico, punto di lavoro e polarizzazione del diodo	68
2.5	Regione di svuotamento	69
2.6	Modello a bande	73
2.7	Giunzione metallo-semiconduttore	75
2.8	Il diodo pn completo	77
2.9	Principio di funzionamento del diodo	79
2.10	Concentrazione dei portatori al limite della regione di svuotamento	80
*2.11	Iniezione di cariche e ricombinazione	82
*2.12	Derivazione della caratteristica del diodo	85
*2.13	Correnti di diffusione e correnti di deriva nel diodo	86
*2.14	Limiti della caratteristica del diodo ideale	87
*2.15	Trappole in banda proibita	88
2.16	Il diodo Zener	91
2.17	Riepilogo	93
2.18	Esercizi svolti	94
2.19	Domande di verifica a scelta multipla	108
2.20	Domande di verifica a risposta aperta	112
2.21	Risposte alle domande a scelta multipla	112
Capitolo 3	Fotodiodi e LED	113
3.1	Interazione luce-semiconduttore	114
3.2	Il fotodiodo	115
3.3	Sensori a matrice per fotocamere	117
3.4	Efficienza di conversione di una cella fotovoltaica	118
3.5	Semiconduttori diretti e indiretti	120
3.6	LED	123
3.7	Dall'effetto fotoelettrico all'illuminazione a LED	125
3.8	Riepilogo	126

3.9	Esercizi svolti	126
3.10	Domande di verifica a scelta multipla	128
3.11	Domande di verifica a risposta aperta	130
3.12	Risposte alle domande a scelta multipla	130

Capitolo 4 Il transistor a effetto campo 131

4.1	Il transistor	132
4.2	Tipologia dei transistor a effetto campo	133
4.3	Il capacitore MOS	134
*4.4	Calcolo della tensione di soglia	137
4.5	MOSFET a canale n ad arricchimento	140
4.6	MOSFET a canale n ad arricchimento per grandi tensioni	145
*4.7	Strozzamento del MOSFET	150
4.8	Modulazione del canale	154
4.9	Realizzazione fisica del MOSFET	155
4.10	Recenti sviluppi della tecnologia dei MOSFET	162
4.11	Legge di Moore e miniaturizzazione dei MOSFET	164
4.12	MOSFET a canale n ad arricchimento con carico resistivo	165
4.13	MOSFET come amplificatore di tensione	167
4.14	Polarizzazione del MOSFET a canale n ad arricchimento	169
4.15	Rete di polarizzazione con resistenza di source	170
4.16	Polarizzazione con generatore di corrente	172
4.17	Specchio di corrente	173
4.18	Rete di polarizzazione con resistenza di gate	174
4.19	MOSFET in saturazione come carico attivo	175
4.20	MOSFET a svuotamento	178
4.21	Polarizzazione del MOSFET a svuotamento	179
4.22	MOSFET a svuotamento come carico attivo	181
4.23	MOSFET a canale p ad arricchimento	183
4.24	Prospetto riassuntivo dei MOSFET a tre terminali	184
4.25	MOSFET a quattro terminali	186
*4.26	Effetto body	188

4.27	JFET	190
*4.28	Derivazione della caratteristica del JFET	194
4.29	La travagliata nascita del transistor a effetto campo	197
4.30	Riepilogo	199
4.31	Esercizi svolti	200
4.32	Domande di verifica a scelta multipla	222
4.33	Domande di verifica a risposta aperta	225
4.34	Risposte alle domande a scelta multipla	225

Capitolo 5	Il transistor bipolare	227
-------------------	-------------------------------	------------

5.1	Il BJT	228
5.2	Realizzazione fisica del BJT	228
5.3	Regioni di funzionamento	233
5.4	Regione attiva diretta	234
5.5	Regione attiva inversa	238
5.6	Interdizione	238
5.7	Saturazione	239
5.8	Caratteristica di uscita del BJT	239
5.9	Effetto Early	240
5.10	Il BJT pnp	242
5.11	Il BJT come interruttore	243
5.12	Il BJT come amplificatore di tensione	246
5.13	Polarizzazione del BJT	249
5.14	Polarizzazione del BJT mediante rete a quattro resistori	250
5.15	Polarizzazione del BJT mediante generatore di corrente	253
5.16	Specchio di corrente a BJT	253
5.17	BJT e FET	254
5.18	La nascita dell'industria dei semiconduttori	255
5.19	Riepilogo	256
5.20	Esercizi svolti	257
5.21	Domande di verifica a scelta multipla	282
5.22	Domande di verifica a risposta aperta	285
5.23	Risposte alle domande a scelta multipla	285

Capitolo 6 Famiglie logiche e memorie 287

6.1	Al largo dell'isola greca di Anticitera...	288
6.2	Architettura di base dei sistemi digitali	288
6.3	Famiglie logiche	290
6.4	Interruttore ideale e resistore	297
6.5	Interruttori complementari ideali	305
6.6	La logica CMOS	308
6.7	Bi-CMOS	320
6.8	Classificazione delle memorie	322
6.9	RAM statica	323
6.10	RAM dinamica	329
6.11	FLASH	331
6.12	Riepilogo	336
6.13	Esercizi svolti	337
6.14	Domande di verifica a scelta multipla	343
6.15	Domande di verifica a risposta aperta	346
6.16	Risposte alle domande a scelta multipla	347

Capitolo 7 Amplificatori lineari 349

7.1	Grida e falò...	350
7.2	Gli amplificatori	351
7.3	Amplificatori lineari	353
7.4	Principio di sovrapposizione e non linearità	356
7.5	Modello linearizzato del FET a tre terminali	359
*7.6	Modello linearizzato del FET a quattro terminali	365
7.7	Resistenze viste ai terminali del MOSFET	366
7.8	Le tre configurazioni del FET	371
7.9	Modello linearizzato del BJT	377
7.10	Resistenze viste ai tre terminali del BJT	385
7.11	Le tre configurazioni del BJT	386
7.12	Analisi e progetto	388
7.13	Analisi di un amplificatore di tensione a singolo MOSFET in configurazione CS	388

7.14	Progetto di un amplificatore di tensione a singolo MOSFET in configurazione CS	392
7.15	Progetto di un buffer di tensione a singolo MOSFET in configurazione CD	394
7.16	Amplificatori reazionati	397
7.17	Progetto di un amplificatore di tensione reazionato CE	400
7.18	Teorema di Miller	403
7.19	Amplificatore a MOSFET con resistore tra gate e drain	405
7.20	Amplificatori a carico attivo	406
7.21	Riepilogo	408
7.22	Esercizi svolti	412
7.23	Domande di verifica a scelta multipla	441
7.24	Domande di verifica a risposta aperta	443
7.25	Risposte alle domande a scelta multipla	443

Capitolo 8 Risposta in frequenza degli amplificatori lineari 445

8.1	Risposta in frequenza degli amplificatori	446
8.2	Risposta in frequenza delle reti lineari	446
8.3	Risposta in frequenza degli amplificatori lineari	448
8.4	Risposta di un amplificatore di tensione CS alle basse frequenze	449
8.5	Risposta di un amplificatore CE alle basse frequenze	451
8.6	Regola generale per la stima della frequenza di taglio inferiore	454
8.7	Amplificatori alle alte frequenze	454
8.8	Capacità tra gate e source nel MOSFET in saturazione	455
8.9	Risposta dell'amplificatore CS ad alta frequenza	455
8.10	Modello del MOSFET ad alta frequenza	457
8.11	Capacità parassite del BJT	458
8.12	Modello del BJT ad alta frequenza	462
8.13	Risposta dell'amplificatore CE ad alta frequenza	464
8.14	Regola generale per la stima della frequenza di taglio superiore	467
8.15	Cascode	467

8.16	Riepilogo	471
8.17	Esercizi svolti	472
8.18	Domande di verifica a scelta multipla	478
8.19	Domande di verifica a risposta aperta	478
8.20	Risposte alle domande a scelta multipla	478

Appendici

A	Circuiti con capacitori	479
B	Circuiti lineari	483
C	Analisi di circuiti lineari mediante numeri complessi	487

Indice analitico	495
-------------------------	------------

Prefazione

Come ogni anno, all'inizio del corso di elettronica per il triennio di ingegneria, che ormai tengo da oltre dieci anni, mi pongo il problema del testo da adottare. A prima vista si direbbe un problema di facile soluzione: l'elettronica è una materia di base, ci sono forse un centinaio di insegnamenti attivi, ci saranno decine di buoni libri di testo tagliati per la nostra realtà universitaria. Invece no. La solita rapida indagine sui programmi in rete dei miei colleghi rileva che anche quest'anno i libri adottati sono i soliti tre o quattro, quasi tutti traduzioni dall'inglese. Non sono brutti libri, sono ben curati e ben scritti, ma proprio non riesco a usarli come libri di testo. Il problema è che alcuni di questi sono monografie specialistiche, sicuramente utili per la consultazione, ma ben poco didattiche; gli altri tendono a presentare i dispositivi elettronici come una collezione di scatolette che fanno qualcosa, ma non si sa bene come. Vorrei qualcosa che possa soddisfare la curiosità intellettuale dei miei studenti senza che debbano necessariamente diventare degli specialisti nella progettazione dei dispositivi. Insomma ci vorrebbe un libro.

Un libro rivolto agli studenti universitari che seguono il loro primo insegnamento di elettronica dovrebbe avere un chiaro filo conduttore in cui gli argomenti non appaiano giustapposti ma vengano affrontati in sequenza, quando se ne vede la necessità logica. I concetti rimangono e si radicano solo se strettamente interconnessi. La nostra memoria non è come quella di un computer. Non basta indirizzare una cella per ottenere l'informazione conservata, ma funziona piuttosto come un filo che si dipana. Trovato il bandolo si sbroglia la matassa ed è questo l'approccio che ho voluto seguire, sia nell'architettura generale del testo, sia nella trattazione del singolo dispositivo.

Un limite che spesso avverto nei libri di testo disponibili in commercio è proprio questo: gli argomenti affrontati appaiono più giustapposti che sequenziali e logicamente connessi. Per questo ho dedicato particolare cura a quegli argomenti che potrei definire "connettivi". Faccio un esempio concreto. Di solito dopo aver introdotto il transistor si parla direttamente delle tre configurazioni nelle quali si usa (CS, CG, CD nel caso del MOSFET; CE, CB, CC nel caso del BJT), ma lo studente rimane sempre con un dubbio: perché solo queste tre? Le possibili configurazioni sono sei, perché le altre non si considerano? La questione non è di quelle fondamentali, ma la spiegazione è semplice e aiuta ad avere un quadro completo. Un altro esempio di argomento "connettivo" sono i potenziali che si creano tra metallo e semiconduttore e che compensano il potenziale di giunzione. Di solito le trattazioni della giunzione pn si limitano a descrivere la formazione

del potenziale di built-in, ma lo studente si chiede: perché se applico un multimetro ai capi di una giunzione non misuro alcun potenziale?

Un libro rivolto agli studenti universitari non può limitarsi a descrivere i dispositivi senza parlare di come funzionano. D'altra parte la spiegazione completa della fisica dei dispositivi va ben al di là di un corso introduttivo di elettronica. La soluzione che ho scelto è puntare direttamente sul principio di funzionamento, nel modo più semplice possibile. Quando ho potuto, sono partito da un'immagine metaforica per catturare l'idea e, soprattutto, l'attenzione dello studente. Ad esempio per parlare delle correnti di diffusione immagino un campo di calcio in cui gli atomi sono i giocatori e gli elettroni i palloni. Una volta fornita l'idea, passo poi alla spiegazione rigorosa e matematica. Ho sperimentato in classe questo approccio per anni e l'ho sempre trovato molto efficace. Ha il grande pregio di evitare che lo studente, concentrato nei dettagli del calcolo, perda di vista lo scopo e la sostanza di ciò che sta facendo.

Uno dei primi problemi che un docente si pone nel pianificare un corso di elettronica di base è quanto spazio dare al transistor bipolare (BJT). Il BJT è stato alla base dell'elettronica per decenni, dai primi anni sessanta agli anni novanta, ma oggi è decisamente in declino. Alcuni testi fanno la scelta di non trattarlo più. Io la trovo un po' troppo radicale. Il BJT è sicuramente scomparso nelle applicazioni digitali ma rimane essenziale negli amplificatori ad alto guadagno a bassa cifra di rumore e negli amplificatori di potenza. D'altra parte una trattazione convenzionale del BJT richiede molte ore di lezione. Un po' paradossalmente, il funzionamento del BJT è più complesso di quello del MOSFET. La soluzione che ho affinato con il tempo è una trattazione semplificata del BJT, che anche senza introdurre il modello di Ebers-Moll coglie gli aspetti essenziali mantenendo coerenza e rigore.

Un'altra importante questione preliminare è come affrontare la tecnologia planare. La soluzione più semplice potrebbe essere un capitolo dedicato, ma la mia impressione è che questo approccio sia didatticamente poco efficace. Lo studente non coglie il collegamento profondo tra la microelettronica e la sua tecnologia realizzativa. Ho quindi preferito, di norma, descrivere prima come si realizza un dato dispositivo e poi come funziona. Lo scopo è dare subito allo studente un'idea concreta di ciò di cui si sta parlando, prima che si possa perdere in calcoli e modelli.

Alcuni argomenti trattati in questo libro possono risultare un po' troppo avanzati per un corso introduttivo, ad esempio la derivazione della caratteristica statica del diodo e il calcolo della tensione di soglia del MOSFET. Si tratta, a mio parere, di argomenti importanti e formativi, ma possono essere omessi senza che ciò pregiudichi la comprensione delle parti che seguono. Questi argomenti "opzionali" sono evidenziati con un asterisco davanti al titolo del paragrafo.

Le materie scientifiche sono trattate senza mai fare riferimento all'evoluzione storica che ha portato a certe invenzioni o scoperte. Quest'approccio pragmatico è piuttosto efficace e rispecchia un'importante assunzione implicita nella scienza: scoperte e invenzioni valgono in sé, non per l'autorità o il prestigio di chi le ha proposte. Tuttavia l'esclusione totale di ogni riferimento storico rende la materia un po' sospesa in un mondo senza tempo. Per questo, ritengo importante, all'inizio o al termine di ogni capitolo, collocare i concetti nel loro contesto storico. La scienza e la tecnologia non sono solo la descrizione di fatti obiettivi (qualunque sia il significato che vogliamo dare al termine "obiettivo"), ma il prodotto di un'evoluzione storica. Un futuro ingegnere dovrebbe essere consapevole di ciò. Inoltre, più prosaicamente, una manciata di aneddoti e qualche commento storico aiutano a rendere meno pesanti certe lezioni, specialmente se collocate nel primo pomeriggio.

Talvolta lo studente può trovare difficoltà a distinguere i concetti importanti da quelli, per così dire, accessori. Per questo alla fine di ogni capitolo ho inserito un riepilogo che riporta per punti i concetti chiave.

Le materie tecniche si comprendono solo quando si è in grado di svolgere un esercizio. Non mi stancherò mai di ripetere questo concetto. È attraverso la personale riflessione su un problema non banale che i concetti appresi si consolidano. Per questa ragione alla fine di ogni capitolo ho inserito una nutrita collezione di esercizi svolti. È importante che tali esercizi non siano pure applicazioni di formule o di algoritmi definiti. Lo studente deve aver modo di riflettere autonomamente e proporre una soluzione originale. A questo scopo ho evitato al massimo esercizi con calcoli lunghi e complessi, privilegiando problemi con soluzioni semplici ed eleganti. Contrassegnati con un asterisco, ho inserito anche alcuni esercizi che la maggior parte degli studenti troverà piuttosto difficile.

Si può molto discutere sui limiti delle domande a scelta multipla ma sono uno strumento senza dubbio veloce e obiettivo per valutare la preparazione di uno studente. Il problema è semmai formulare le domande in modo che non siano nozionistiche, ma richiedano riflessione. A questo proposito grande cura e attenzione è stata posta nella loro elaborazione. Tra l'altro quasi tutte sono state testate dai miei studenti.

Il grande limite delle domande a scelta multipla è che non sono in grado di valutare la capacità dello studente di articolare un'argomentazione. Per questo ho provveduto ogni capitolo anche di alcune domande a risposta aperta. Ho cercato di fare in modo che non si sovrapponevano per argomenti alle domande a scelta multipla. La maggior parte delle domande a risposta aperta, infatti, si riferisce a dimostrazioni o concetti non strettamente quantitativi.

Il docente userà come meglio crede il materiale di questo libro, tuttavia mi permetto di dare qualche linea guida. Gli otto capitoli del libro possono essere svolti interamente in un corso di 9-10 crediti formativi (80-90 ore) adatto per tutti i corsi di laurea dell'area dell'informazione. Per un corso più breve si dovranno fare delle scelte.

I primi 5 capitoli con i capitoli 7 e 8 costituiscono un corso di 8-9 crediti formativi (70-80 ore) adatto agli studenti di Ingegneria Elettronica e Ingegneria delle Telecomunicazioni che hanno nel piano di studi un successivo corso di elettronica digitale. Per un corso più breve (7-8 crediti, 60-70 ore) si potranno omettere i paragrafi con l'asterisco e anche il capitolo 3 (Fotodiodi e LED). Per un corso minimale (5-6 crediti, 40-50 ore) si può scegliere di non trattare il transistor bipolare (capitolo 5) avendo poi cura di non svolgere i paragrafi e gli esercizi dei capitoli 7 e 8 che si riferiscono al BJT. Ciò non pregiudica il filo logico della trattazione in quanto gli ultimi due capitoli sono strutturati in modo da rimanere comunque autoconsistenti.

I primi 6 capitoli costituiscono, invece, un corso di 6 crediti (50-60 ore) adatto agli studenti di Ingegneria Informatica. Per un corso minimale (30-40 ore), si potranno omettere i paragrafi con l'asterisco, il capitolo 3 (Fotodiodi e LED) e il capitolo 5 (Il transistor bipolare). Il capitolo successivo (Famiglie logiche e memorie) cita il BJT in un solo paragrafo che può essere saltato senza pregiudicare la comprensione del capitolo stesso.

Pearson Learning Solution

Il codice di registrazione che trovate sulla copertina di questo libro consente l'accesso per diciotto mesi al Pearson eText, l'edizione digitale del libro arricchita da funzionalità che permettono di personalizzarne la lettura, evidenziare il testo, inserire segnalibri e annotazioni, studiare e condividere note anche su tablet con l'app Pearson eText.

1

Fisica dei semiconduttori

In questo capitolo si forniscono i concetti di Fisica dello stato solido necessari per la comprensione dei dispositivi a semiconduttore. L'approccio alla materia privilegia le idee fisiche piuttosto che il formalismo matematico, quando possibile anche con immagini vivide che possano stimolare l'immaginazione dello studente. Con questo spirito sono trattati il concetto di lacuna, la legge di azione di massa, l'energia di Fermi e il potenziale di built-in. Conclude il capitolo un paragrafo storico sui semiconduttori.

1.1 La carica elettrica

Gli scrittori medievali erano soliti iniziare con brevi cenni sull'universo. Noi inizieremo parlando di cariche elettriche: l'universo è fatto di atomi e in un atomo ci sono corpuscoli molto piccoli chiamati elettroni, che per convenzione si dicono avere carica negativa.

Quando un elettrone per qualche ragione si allontana dall'atomo, quest'ultimo resta carente di una carica negativa, ovvero acquisisce una carica positiva. Sarebbe stato più logico chiamare positiva la carica dell'elettrone e negativa quella dell'atomo senza un elettrone, ma quando William Gilbert, medico di corte di sua maestà britannica Elisabetta I nell'Anno Domini 1600, scoprì l'elettricità senza saper nulla di atomi, aveva il cinquanta per cento di possibilità di fare la scelta giusta.

Purtroppo non siamo stati fortunati. Da allora cariche e correnti elettriche sembrano avere sempre il segno sbagliato.

Circa un secolo dopo la pubblicazione del trattato *De Magnete, Magneticisque Corporibus, et de Magno Magnete Tellure Physiologia Nova* di William Gilbert, l'ingegnere del re (così si chiamavano gli ingegneri militari prima della rivoluzione francese) Charles Augustin de Coulomb scoprì che le cariche dello stesso segno si respingono e le cariche di segno opposto si attraggono. Scoprì inoltre che la forza F tra due cariche q_1 e q_2 è data da:

$$F = k \frac{q_1 q_2}{R^2} \quad (1.1)$$

Nella formula la carica elettrica si esprime in Coulomb (C). Purtroppo la carica di un singolo elettrone è un valore piccolissimo e molto scomodo da usare: $1.602 \times 10^{-19} \text{C}$, diciannove ordini di grandezza più piccolo dell'unità di misura.

D'altronde, Charles Augustin de Coulomb, così come William Gilbert, nulla poteva sapere di atomi e di elettroni, né tantomeno immaginare quanto piccoli potessero essere. Per questo a distanza di tre secoli continuiamo a trascinarci questa ingombrante e scomoda unità di misura.

La costante k che compare nella formula è detta costante di Coulomb e vale $8.99 \times 10^9 \text{ Nm}^2\text{C}^{-2}$. R è la distanza tra le cariche. Il fatto che la legge di Coulomb sia formalmente identica alla legge di gravitazione universale (eccetto per il fatto che le masse sono sempre positive e le cariche hanno un segno) ha affascinato gli scienziati per secoli e ha a che fare con la struttura profonda dello spazio. Un altro mistero che ha sempre affascinato i fisici è che la carica positiva dell'universo è esattamente uguale alla carica negativa, ovvero l'universo è neutro. Non è così ovvio come sembra.

1.2 Tensione

Le cariche elettriche, abbiamo già detto, si attraggono o si respingono. Di solito sono confinate all'interno dell'atomo e quindi non ci accorgiamo delle forze tra le cariche. Tuttavia, per qualche ragione, può capitare che ci sia un accumulo di cariche in una regione di spazio e quindi un accumulo di cariche opposte in un'altra regione. Gli antichi si erano già accorti di questo fenomeno. Strusciando un vello di pecora su un pezzetto di ambra si genera appunto un accumulo di cariche negative sull'ambra (ambra in greco si dice *ηλεκτρον* [*élektron*], da cui il termine elettricità). Tra una regione di accumulo di cariche negative e una regione di accumulo di cariche positive, per la legge di Coulomb, ci sono delle forze. Prendiamo una carica elettrica (ad esempio un atomo privato di un elettrone, ovvero uno ione positivo) e facciamole percorrere un qualunque cammino tra due punti A e B nello spazio tra l'accumulo di carica negativa e l'accumulo di carica positiva (vedi Figura 1.1).

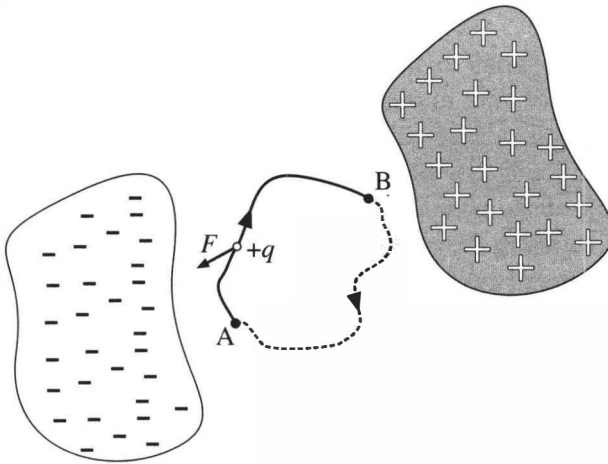


Figura 1.1 Cammino chiuso tra due punti nello spazio

Per fare ciò dovremo fornirle un'energia E pari all'integrale sul cammino della forza a cui è soggetta la carica.

$$E = \int_A^B \vec{F} \cdot d\vec{l} \quad (1.2)$$

Se facciamo percorrere alla stessa carica il cammino inverso (o un qualunque altro cammino che la riporti dal punto B al punto A) la stessa energia è restituita dalla carica sotto forma di energia cinetica. Si può dimostrare che questa energia è indipendente dal particolare percorso che compie la carica

elettrica e dipende solo dalle posizioni dei punti A e B nello spazio. Quindi, fissato un qualunque punto di riferimento nello spazio (ad esempio A), possiamo definire una funzione che dipende solo dalla generica posizione P nello spazio e che esprime l'energia tra il punto A e il punto P. Poiché la forza elettrica, data dalla legge di Coulomb, è proporzionale alla carica elettrica, anche l'energia fornita o restituita sotto forma di energia cinetica sarà proporzionale alla carica.

Per definire una grandezza che non dipenda dal valore della carica mobile, ma solo dalla configurazione spaziale delle altre cariche che determinano le forze elettriche, dividiamo l'energia per la carica e definiamo questa nuova grandezza, dipendente solo dalla posizione nello spazio, **potenziale elettrico**: $\varphi(x,y,z)$.

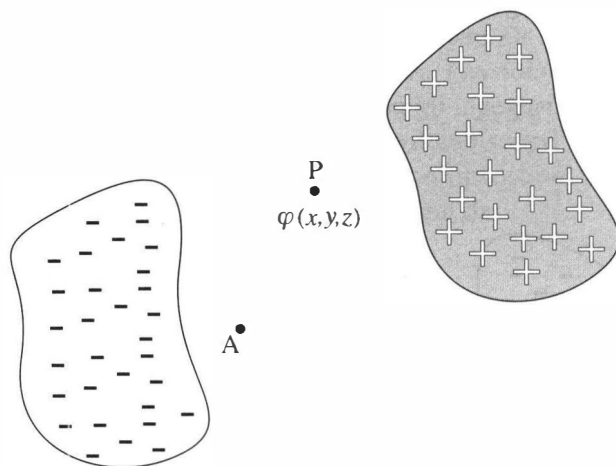


Figura 1.2 Potenziale elettrico

Il potenziale è dimensionalmente un'energia diviso una carica ovvero Joule per unità di carica, J/C. Questa unità di misura prende il nome di **Volt (V)**. Per come è stato definito, il potenziale aumenta quando muovendo una carica positiva dobbiamo fornire un'energia, diminuisce quando la carica positiva restituisce energia in forma di energia cinetica. Per una carica negativa vale l'opposto: dobbiamo fornirle energia quando il potenziale diminuisce, restituisce energia quando il potenziale aumenta.

Per quanto abbiamo detto sopra, il valore del potenziale dipende dal punto scelto come riferimento, ma dati due punti nello spazio la differenza di potenziale è indipendente dal riferimento.

Questa differenza di potenziale è detta **tensione** ed è una delle grandezze fondamentali dell'elettronica.

Si può dimostrare che il potenziale elettrico è legato alla forza elettrica (a cui è soggetta una generica carica elettrica q nello spazio) mediante la seguente relazione:

$$\vec{F}(x, y, z) = \left(-q \frac{\partial \varphi(x, y, z)}{\partial x}, -q \frac{\partial \varphi(x, y, z)}{\partial y}, -q \frac{\partial \varphi(x, y, z)}{\partial z} \right) = -q \nabla \varphi \quad (1.3)$$

La forza elettrica, normalizzata rispetto alla carica, è detta **campo elettrico**:

$$\vec{E} = \frac{\vec{F}}{q} \quad (1.4)$$

Pertanto, vale in generale:

$$\vec{E} = -\nabla \varphi \quad (1.5)$$

In definitiva possiamo dire che una distribuzione di cariche nello spazio determina un potenziale e un campo elettrico definito in tutto lo spazio. Si noti, tuttavia, che a rigore anche la carica elettrica mobile che abbiamo usato per definire il potenziale interagisce con le altre cariche provocando una modifica della distribuzione di carica. Il problema si risolve immaginando una carica molto piccola che non influenzi la distribuzione complessiva di carica. Se poi volessimo essere veramente pignoli, potremmo domandarci se ha senso ipotizzare una carica molto piccola, sapendo che tutti gli elettroni sono uguali e portatori della stessa carica elementare. Ma qui si sconfina nella filosofia.

Nella maggior parte dei casi pratici che affronteremo nel resto di questo libro, potenziale e campo elettrico sono definiti in uno spazio monodimensionale. In tal caso, valgono le seguenti semplici relazioni:

$$\varphi(x) = - \int_0^x E dx \quad (1.6)$$

$$E = - \frac{d\varphi}{dx} \quad (1.7)$$

Purtroppo si è soliti indicare energia e campo elettrico con la stessa lettera (E maiuscola). Il problema della notazione è comune a chiunque si accinga a scrivere un testo scientifico. Una notazione rigorosa e assolutamente non ambigua risulta immancabilmente artificiosa e complessa. In generale, seguiremo le notazioni più comuni, discostandocene solo quando ne risulterebbe un'ambiguità non facilmente risolvibile con un po' di buon senso.

1.3 Corrente

Uno sbilanciamento di cariche, ovvero un accumulo di cariche positive in una regione e negative in un'altra, dà origine a un potenziale non costante, ovvero a un potenziale con un gradiente diverso da zero. Per la (1.3), o nel caso monodimensionale per la (1.7), un gradiente significa una forza applicata alle cariche elettriche. Se una forza è applicata a una carica mobile questa tenderà a muoversi: nel senso opposto al gradiente se si tratta di una carica positiva, nel senso del gradiente se si tratta di una carica negativa. In altre parole, una tensione, ovvero una differenza di potenziale, in presenza di cariche mobili positive o negative, dà origine a una corrente. Per fissare le idee, supponiamo ad esempio che ci sia uno sbilanciamento di cariche tra due regioni: A e B. Alcuni elettroni degli atomi della regione A sono stati strappati dai rispettivi atomi e spostati nella regione B (ad esempio sfregando un vello di pecora con una bacchetta d'ambra, o più semplicemente un maglione di lana con una penna di plastica).

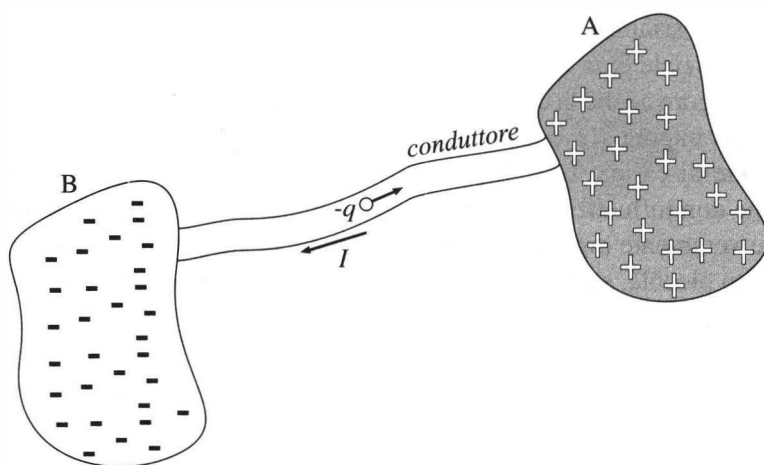


Figura 1.3 Corrente elettrica

Collegiamo ora la regione A e la regione B con un filo conduttore (ancora non abbiamo detto cosa sia un materiale conduttore, ma lo diremo presto, per ora basta sapere che è una sorta di “tubo” dove possono passare gli elettroni). Spinti dalla tensione elettrica, gli elettroni strappati dagli atomi in A e quindi mobili, fluiranno da B a A. Poiché le cariche mobili sono negative, diremo che c'è una corrente negativa $-I$ da B ad A, ovvero una corrente positiva I da A a B. Se William Gilbert, nell'anno 1600, avesse fatto la scelta opposta nei segni delle cariche elettriche, non avremmo avuto bisogno di doverci ricordare continuamente che la corrente scorre nella direzione opposta agli elettroni. In realtà, non siamo così sfortunati, perché in elettronica si fa uso anche di portatori di carica

positiva, questi sono atomi mobili privi di un elettrone (ioni), oppure un altro strano tipo di portatore, detto «lacuna». Correnti di ioni positivi o di lacune hanno lo stesso verso dell'effettivo flusso dei portatori.

Una tensione, applicata a una regione con cariche libere, dà origine a una corrente. Tale corrente esprimerà quanta carica elettrica fluisce nell'unità di tempo, ovvero C/s , che di solito si indica con il termine **Ampere (A)** in onore del fisico francese André-Marie Ampère. Il senso di tale corrente è tale da ridurre lo squilibrio di cariche. Nel caso visto sopra, ad esempio, è tale da riportare gli elettroni strappati ai propri atomi. Quindi man mano che la corrente fluisce la tensione cala. Quando tutti gli elettroni sono tornati nella zona A, la tensione elettrica è nulla. Affinché la tensione possa permanere, così come la corrente, è necessario che qualche meccanismo esterno continui a strappare elettroni dagli atomi in A e a portarli in B. Questo meccanismo, detto genericamente forza elettromotrice, può essere una pila elettrica oppure un alimentatore collegato alla rete elettrica.

1.4 Legge di Ohm

Alla sua pubblicazione nel 1827, il trattato *Die galvanische Kette, mathematisch bearbeitet*, di un certo Georg Simon Ohm, all'epoca professore di fisica in un istituto gesuita tedesco, fu accolto piuttosto freddamente: troppa matematica. Dovette passare una quindicina d'anni prima che la Royal Society britannica, la più prestigiosa istituzione scientifica del tempo, riconoscesse in quel trattato una delle più importanti scoperte della nascente scienza dell'elettricità: la relazione lineare tra tensione (V) e corrente (I) valida per la gran parte dei materiali in grado di condurre una corrente elettrica. Tale relazione oggi è nota come **legge di Ohm** e si è soliti scriverla nella forma:

$$V = RI \quad (1.8)$$

La costante di proporzionalità R è detta **resistenza** ed è dimensionalmente V/A , tale grandezza fisica è detta **Ohm (Ω)**.

Vediamo in dettaglio l'origine fisica della legge di Ohm. Prendiamo in considerazione un cilindro di un qualche materiale solido, in cui ci siano degli elettroni liberi, ovvero non legati ai rispettivi atomi; quale sia il meccanismo fisico per cui ci possano essere tali elettroni lo vedremo nei paragrafi successivi. Per ciascun elettrone libero, ci sarà un atomo privo di un elettrone e quindi caricato positivamente (ione), tali atomi sono fissi nel reticolo cristallino e quindi non contribuiscono in alcun modo all'eventuale corrente. Il materiale è nel suo complesso neutro e quindi gli elettroni non sono soggetti a forze elettriche, ma si muoveranno solo per agitazione termica. Possiamo visualizzare ciò che avviene all'interno del materiale pensando ai giocatori di una squadra di calcio (gli

atomi) che si allenano eseguendo tra di loro passaggi con un certo numero di palloni (gli elettroni). Ciascun giocatore quando riceve un pallone lo stoppa e lo lancia in una direzione casuale. Il tempo che il pallone impiega per andare da un giocatore all'altro è ovviamente molto variabile, ma se i giocatori sono distribuiti piuttosto uniformemente e il campo di calcio è molto grande, possiamo definire un tempo medio τ che dipenderà dall'energia media che i giocatori trasferiscono ai palloni, ovvero fuor di metafora, dall'agitazione termica. All'aumentare della temperatura, ovvero dell'energia media trasferita a ogni calcio, τ diminuisce.

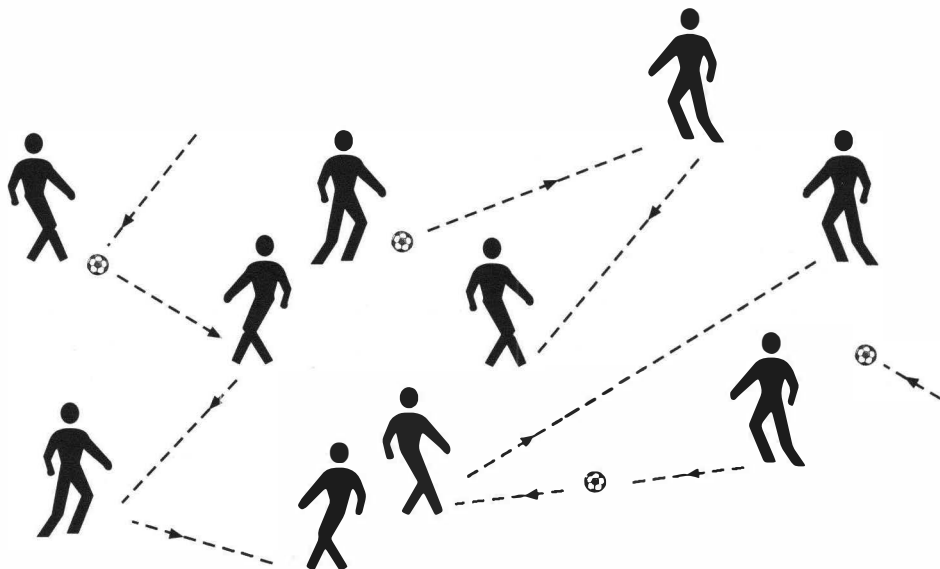


Figura 1.4 Legge di Ohm

Applichiamo ora una tensione (V) ai due estremi del cilindro. All'interno del cilindro ci sarà ora un gradiente di potenziale che tenderà a spingere gli elettroni verso il lato a potenziale maggiore. Tornando all'analogia del campo di calcio, è come se i giocatori si allenassero in un campo in discesa. I palloni saranno ancora lanciati a caso in tutte le direzioni, ma in media tenderanno a spostarsi verso il lato più basso del campo. Questo moto, sovrapposto a quello casuale dovuto all'agitazione termica, è detto «deriva». Si noti che il moto termico rimane preponderante rispetto alla deriva e solo nel lungo termine si noterà questa tendenza dei palloni ad andare verso una specifica direzione.

Vediamo ora come quantificare quanto detto sopra. Prendiamo in considerazione una piccola zona del cilindro in cui a causa del gradiente di potenziale c'è un

campo elettrico E : se la zona è piccola il campo è costante. Un elettrone che si trova in tale piccola zona subirà una forza

$$F = -qE \quad (1.9)$$

con q carica elementare dell'elettrone. A tale forza corrisponde un'accelerazione:

$$a = -\frac{qE}{m} \quad (1.10)$$

con m massa dell'elettrone. Supponiamo che all'istante $t = 0$ l'elettrone subisca un urto con un atomo (il calciatore calcia il pallone); se all'istante t lo stesso elettrone colpisce un altro atomo, a causa del campo elettrico esso avrà acquisito una velocità nella direzione del campo (supplementare a quella che gli ha trasferito il primo urto) pari a

$$v = -\frac{qE}{m} t \quad (1.11)$$

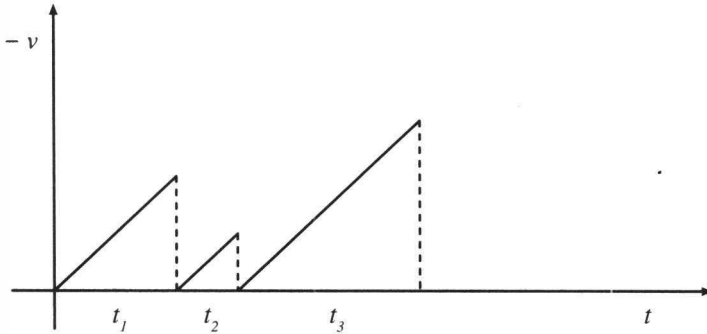


Figura 1.5 Velocità di deriva

Sia t_m l'intervallo di tempo tra l'urto m e l'urto $m + 1$. La velocità media supplementare indotta dal campo elettrico durante il percorso tra i due urti è pari a

$$v_m = -\frac{1}{2} \frac{qE}{m} t_m \quad (1.12)$$

Se ora consideriamo il tempo medio $\langle t_m \rangle = \tau$ tra un urto e l'altro, la velocità media di deriva provocata dal campo elettrico sarà

$$v_d = \langle v_m \rangle = -\frac{q}{2m} E \langle t_m \rangle = -\frac{q}{2m} E \tau \quad (1.13)$$

Definiamo ora mobilità

$$\mu = \frac{q\tau}{2m} \quad (1.14)$$

Segue la relazione lineare tra velocità di deriva e campo elettrico:

$$v_d = -\mu E \quad (1.15)$$

Si noti che la mobilità dipende dalle caratteristiche del portatore (in particolare il rapporto tra carica e massa dell'elettrone) e dalla temperatura. All'aumentare della temperatura, la mobilità diminuisce. Vogliamo, tuttavia, precisare che la (1.14) è stata ottenuta con un modello un po' troppo semplificato della realtà fisica. Alla scala degli elettroni la fisica classica può fornire al più un'indicazione qualitativa. In effetti la mobilità calcolata mediante la (1.14) per il rame, ad esempio, è sei volte minore di quella misurata sperimentalmente. Anche la dipendenza dalla temperatura data dalla (1.14) è corretta solo qualitativamente. Infatti il tempo medio tra un urto e il successivo (τ) è inversamente proporzionale alla velocità media degli elettroni liberi $\langle v \rangle$. In generale la statistica di Boltzmann, prevede che per una particella in grado di muoversi con tre gradi di libertà vale

$$\frac{1}{2} m \langle v^2 \rangle = \frac{3}{2} kT \quad (1.16)$$

Pertanto sostituendo nella (1.14) si ottiene

$$\mu \propto \frac{1}{\sqrt{T}} \quad (1.17)$$

Ma sperimentalmente la mobilità è circa proporzionale a $1/T$, piuttosto che alla sua radice. Queste discrepanze si risolvono con una trattazione quantistica del moto degli elettroni tra gli atomi del reticolo cristallino.

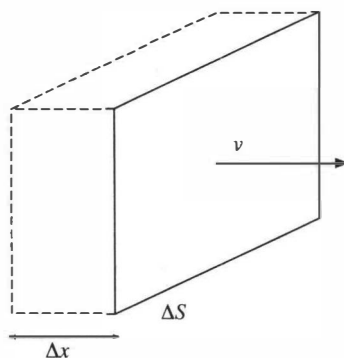


Figura 1.6 Densità di corrente

Consideriamo ora una superficie infinitesima ΔS attraversata da un flusso di cariche q di velocità v . La carica che attraversa la superficie ΔS nel tempo infinitesimo Δt è tutta la carica nel volume adiacente alla superficie di spessore Δx , con $\Delta x = v \Delta t$

$$\Delta Q = qn\Delta S\Delta x \quad (1.18)$$

con n cariche q per unità di volume. La densità di corrente j è la carica che attraversa una superficie nell'unità di tempo, ovvero

$$j = \frac{\Delta Q}{\Delta S\Delta t} \quad (1.19)$$

Pertanto

$$j = qnv \quad (1.20)$$

La (1.20) è valida anche nel caso in cui le cariche che fluiscono siano negative (ad esempio elettroni). Infatti la carica del singolo portatore è negativa e la velocità di deriva ha direzione opposta al campo elettrico, pertanto

$$j = (-q)n(-\mu E) = qn\mu E \quad (1.21)$$

Definiamo la conducibilità

$$\sigma = qn\mu \quad (1.22)$$

La densità di corrente può essere quindi espressa come segue

$$j = \sigma E \quad (1.23)$$

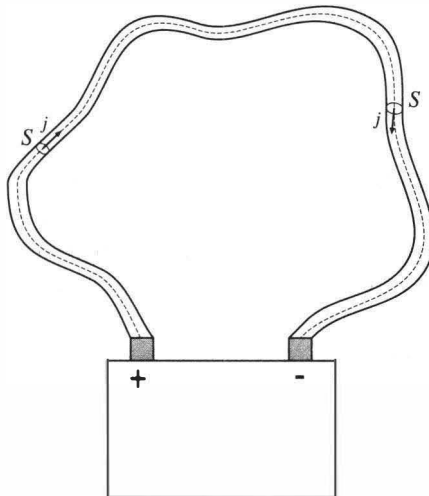


Figura 1.7 Corrente in un conduttore

Consideriamo ora un cilindro di sezione costante (ma anche non rettilineo purché di curvature non tali da produrre locali strozzamenti della sezione) in cui la corrente può fluire all'interno, ma non uscire dalle pareti laterali. Alle due estremità è applicata una tensione V . La corrente I che attraversa una generica sezione S del cilindro è pari a

$$I = \int_S \vec{j} \cdot \vec{n} dS \quad (1.24)$$

con $\vec{n}$ versore ortogonale alla superficie di integrazione.

Facciamo ora un'ipotesi piuttosto stringente, ma ragionevole: supponiamo la densità di corrente sempre parallela all'asse del cilindro e di modulo costante su ciascuna sezione ortogonale all'asse del cilindro. Si tratta di una di quelle ipotesi che a priori non sono del tutto giustificate, ma semplificano molto la trattazione e danno un primo modello che produce risultati in ottimo accordo con le verifiche sperimentali. Se poi volessimo approfondire la questione dovremmo considerare un piccolo gradiente di concentrazione degli elettroni tra centro e periferia del cilindro, in quanto gli elettroni tendono a respingersi e quindi ad addensarsi più sulla superficie. Questo fenomeno è detto "effetto pelle", ma nel resto della trattazione lo trascureremo.

Con questa ipotesi per ciascuna sezione vale

$$I = jS \quad (1.25)$$

Poiché S e I sono costanti lungo il cilindro, segue che anche j è costante (in modulo). Ma se j è costante, per la (1.23) anche il campo elettrico è costante (in modulo) e, come la densità di corrente, ha direzione sempre parallela all'asse del cilindro. Pertanto:

$$V = - \int_l \vec{E} d\vec{l} = -lE \quad (1.26)$$

A questo punto il lettore potrebbe domandarsi cosa distorce il campo elettrico tra i due poli della batteria, tanto da fargli seguire l'asse del conduttore per quante volute possa fare. La risposta è che la corrente scorre seguendo il campo elettrico, ma, non potendo uscire dal cilindro, quando incontra le pareti dà luogo a un addensamento di cariche che produce un campo elettrico tale da mantenere il campo elettrico totale in asse con il cilindro. Questi addensamenti di cariche statiche non producono lavoro, ma si limitano a modificare la direzione del campo elettrico guidando le cariche nel conduttore.

Dalla (1.26) segue che

$$I = \sigma ES = -\sigma \frac{S}{l} V \quad (1.27)$$

che può essere riscritta come

$$V = -RI \quad (1.28)$$

avendo definito

$$R = \frac{1}{\sigma} \frac{L}{S} \quad (1.29)$$

La (1.28) è, appunto, la legge di Ohm. Il senso della (1.29) è molto intuitivo: a parità di tensione applicata, fili elettrici lunghi e stretti conducono meno corrente di fili corti e larghi.

Il segno meno nella (1.28) sta ad indicare che la corrente scorre nel verso opposto al verso positivo della tensione (ovvero dal positivo al negativo e non viceversa). Scegliendo positivo il verso della corrente opposto al verso della tensione, la legge di Ohm si scrive nella consueta forma:

$$V = RI \quad (1.30)$$

1.5 Isolanti e conduttori

In elettronica si usano molti differenti materiali che tuttavia rientrano sempre in tre grandi categorie: isolanti, conduttori e semiconduttori. Gli isolanti hanno conducibilità tipicamente compresa tra 10^{-20} e $10^{-8} \Omega^{-1} \text{cm}^{-1}$. I conduttori hanno conducibilità compresa tra 10^4 e $10^8 \Omega^{-1} \text{cm}^{-1}$. I semiconduttori hanno conducibilità intermedia tra gli isolanti e i conduttori. Ad esempio la conducibilità del silicio puro, che è appunto un semiconduttore, è dell'ordine di $10^{-6} \Omega^{-1} \text{cm}^{-1}$. Si noti bene che tra la conducibilità dei conduttori e la conducibilità degli isolanti ci sono almeno 12 ordini di grandezza: un'enormità neppure immaginabile.

Sia per gli isolanti sia per i conduttori vale la legge di Ohm, tuttavia applicando una tensione (di valore ragionevole) a un isolante la corrente indotta sarà così piccola che nella gran parte dei casi pratici potrà essere del tutto trascurata. Ad esempio applicando una tensione di un Volt a un cilindretto di isolante di conducibilità $10^{-8} \Omega^{-1} \text{m}^{-1}$ lungo 1cm e di raggio 1mm, la corrente che scorre è di circa 3pA. Per ogni scopo pratico due punti collegati da un isolante, sono appunto isolati. Un circuito elettrico interrotto da un materiale isolante è quindi un circuito aperto, ovvero un circuito in cui non scorre corrente.

Anche per i conduttori vale la legge di Ohm, tuttavia la conducibilità è così grande che anche una piccola tensione produce una corrente enorme. Ad esempio applicando una tensione di un Volt a un cilindretto di materiale conduttore di conducibilità $10^6 \Omega^{-1} \text{m}^{-1}$ lungo 1cm e di raggio 1mm, la corrente che scorre è di circa 300A, che è un valore enorme per un dispositivo elettronico, infatti 300A su una differenza di tensione di 1V dissipa sul materiale attraversato 300W: una potenza in grado di danneggiare la maggior parte dei

dispositivi elettronici di cui ci occuperemo in questo libro. Nella pratica, nei fili conduttori o nelle piste di materiale conduttore in un circuito elettrico scorrono tipicamente correnti di qualche mA, che nell'esempio di cui sopra per la legge di Ohm daranno luogo a differenze di tensione dell'ordine di circa $3\mu\text{V}$, valore del tutto trascurabile nella gran parte dei casi pratici. In definitiva possiamo affermare che due punti collegati da un conduttore saranno allo stesso potenziale indipendentemente dalla corrente che scorre. Questa assunzione sarà fatta frequentemente nel corso di questo libro, senza ulteriori specificazioni.

Isolanti e conduttori hanno conducibilità molto diverse. Vediamo ora in dettaglio la fisica che sta dietro tale diversità.

In un materiale isolante solido gli atomi sono legati da legami covalenti e formano un reticolo cristallino. Gli elettroni sono confinati negli orbitali interni degli atomi o nei legami covalenti, che sono una sorta di orbitali più estesi che due atomi condividono. In un legame covalente, così come in un orbitale, possono stare al massimo due elettroni. Gli elettroni nei legami covalenti possono essere strappati via più facilmente degli elettroni degli orbitali interni, ovvero hanno una minore energia di estrazione. A una data temperatura, diversa dallo zero assoluto, gli atomi vibrano e si muovono con una certa energia che si scambiano l'uno con l'altro: un po' come se si dessero reciprocamente dei piccoli calci. La maggior parte degli scambi di energia avviene a valori bassi, ma per quanto bassa sia la temperatura la probabilità di un calcetto con energia qualsivoglia grande non è nulla. Pertanto a qualsiasi temperatura ci saranno degli scambi di energia tra gli atomi tali da superare l'energia di estrazione degli elettroni nei legami covalenti, ovvero ci saranno elettroni mobili in grado di condurre una corrente. Nel caso degli isolanti l'energia per liberare gli elettroni dei legami covalenti è piuttosto alta e quindi il numero di elettroni liberi è estremamente basso: per questa ragione un isolante ha conducibilità così bassa. All'aumentare della temperatura, aumenta il numero di elettroni liberi, molto più di quanto diminuisce la mobilità e quindi nel complesso la conducibilità aumenta.

Un altro possibile legame, diverso dal legame covalente, ma comunque in grado di tenere insieme un reticolo cristallino, è il cosiddetto "legame metallico".

Certi tipi di atomi (i metalli, appunto) possono legarsi costituendo una sorta di grande nube elettronica comune a tutti gli atomi del reticolo. In questo caso, ciascun atomo mette in compartecipazione uno o due elettroni che diventano liberi di muoversi all'interno del reticolo. Questo tipo di legame è possibile quando l'energia di ionizzazione degli elettroni più esterni è sufficientemente bassa e pertanto l'agitazione termica a temperatura ambiente è in grado di liberare praticamente tutti gli elettroni più esterni di ciascun atomo.

Ad esempio l'atomo di rame (Cu) ha 29 elettroni. In generale gli elettroni si dispongono intorno al nucleo su orbitali. Un orbitale è un livello energetico

e ci possono stare al massimo due elettroni. Gli orbitali di un atomo, sebbene di energia diversa, si dispongono in cortecce, ovvero in una sorta di strati concentrici intorno al nucleo. La prima corteccia è costituita da 2 elettroni (un singolo orbitale), la seconda da 8 (quattro orbitali), la terza da 18 (9 orbitali) e la quarta, la più esterna, da un solo elettrone. In generale, il completamento o il completo svuotamento della corteccia più esterna è una condizione di minimo energetico per l'atomo, ciò significa che un atomo con uno o due elettroni nella corteccia più esterna tenderà a cedere anche quei pochi formando una nube elettronica, che riempie l'intero reticolo. Gli elettroni di tale nube rimangono comunque legati al reticolo dal fatto che gli atomi sono ioni positivi, la cui carica è schermata proprio dalla nube negativa. Se un elettrone si allontana, il materiale perde la sua neutralità diventando positivo e quindi attirando a sé l'elettrone.

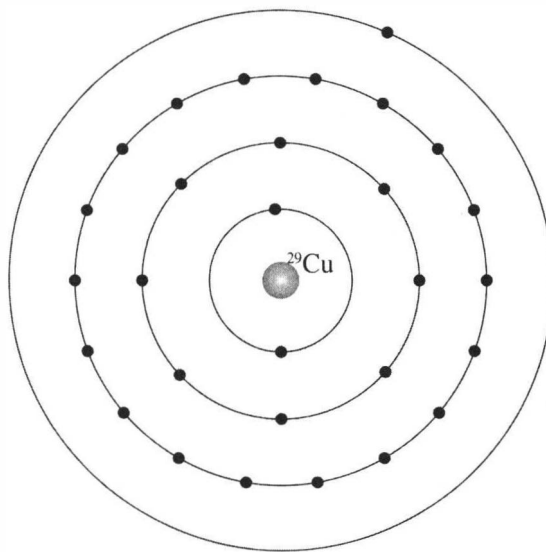


Figura 1.8 Rappresentazione schematica dell'atomo di rame

Come è facilmente immaginabile in un metallo il numero di elettroni liberi è enormemente più grande rispetto al caso degli isolanti e varia poco con la temperatura. Questa è la ragione per cui la conducibilità dei conduttori è così elevata. Tra l'altro la mobilità nei metalli è tipicamente piuttosto bassa (ad esempio più bassa che per il silicio o il germanio), ma ciò nonostante la più elevata concentrazione di portatori fa sì che la conducibilità dei metalli sia molto più grande che nei semiconduttori e negli isolanti. Si noti, inoltre, che il numero degli elet-

troni mobili nei metalli è circa costante con la temperatura, mentre la mobilità diminuisce con l'aumentare della temperatura, pertanto la conducibilità dei conduttori diminuisce con l'aumentare della temperatura.

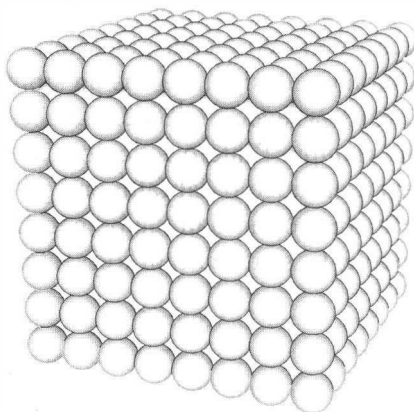


Figura 1.9 Reticolo cristallino

1.6 Semiconduttori

I semiconduttori hanno conducibilità molto più alta degli isolanti, ma minore dei conduttori. Gli atomi sono legati da legami covalenti come per gli isolanti, con i quali condividono anche la dipendenza della conducibilità dalla temperatura. In effetti, Faraday nel 1833 scoprì il primo semiconduttore, proprio osservando uno strano materiale, il solfuro d'argento, che conduceva quasi come un metallo, ma la cui conducibilità aumentava con la temperatura al contrario dei metalli conduttori.

Il più comune semiconduttore è il silicio. Anche il germanio è un semiconduttore. Silicio e germanio sono gli unici semiconduttori non composti, ovvero costituiti da atomi di un solo tipo. Sono semiconduttori i seguenti composti binari: GaAs, GaP, GaSb, AlSb, InP, InSb, InAs. Proviamo a localizzare gli atomi di ciascuna coppia nella tavola periodica. Noteremo che sono tutti a cavallo della IV colonna: quella del carbonio, del silicio e del germanio. Evidentemente le particolari proprietà dei semiconduttori hanno a che fare con questa colonna.

Si possono realizzare semiconduttori anche prendendo due atomi a sinistra della IV colonna e uno a destra formando un composto ternario: $(\text{Ga}_x\text{In}_{1-x})\text{As}$, $(\text{Al}_x\text{In}_{1-x})\text{As}$. La x sta a indicare la frazione di una specie atomica. Tali composti, infatti, non sono stechiometrici, ovvero non hanno una frazione fissa di atomi di un tipo e atomi di un altro tipo, ma sono piuttosto delle leghe, la cui composizione può essere variata con una certa continuità. Con la stessa logica si possono realizzare composti quaternari come ad esempio: $(\text{In}_{1-x}\text{Ga}_x)(\text{As}_{1-y}\text{P}_y)$.

	III	IV	V	VI	VII
	B 5	C 6	N 7	O 8	
	Al 13	Si 14	P 15		
Zn 30	Ga 31	Ge 32	As 33		
Cd 48	In 49	Sn 50			

Figura 1.10 Tavola periodica degli elementi

Malgrado tanta varietà, praticamente tutti i dispositivi elettronici sono realizzati in silicio, con le sole eccezioni dei LED (Light Emitting Diode) e dei LASER a semiconduttore, per una limitazione fisica fondamentale del silicio di cui parleremo nel capitolo dedicato alle applicazioni optoelettroniche. Nel seguito, quando non diversamente specificato, ci riferiremo al silicio.

Il silicio ha numero atomico 14, ovvero ha 14 elettroni disposti in orbitali intorno al nucleo. Due elettroni sono sistemati nell'orbitale più interno (che è anche la corteccia più interna), otto elettroni nei quattro orbitali molto vicini tra loro che costituiscono la seconda corteccia. I restanti quattro elettroni sono sistemati nella terza corteccia, anch'essa potenzialmente costituita da quattro orbitali in grado di contenere otto elettroni.

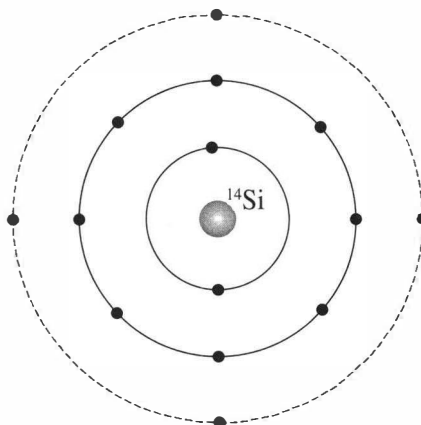


Figura 1.11 Rappresentazione schematica dell'atomo di silicio

Tutti gli atomi della IV colonna, non solo il silicio, hanno la corteccia più esterna riempita a metà ed è proprio questa la caratteristica che li rende speciali. Infatti, tali elementi si trovano nella condizione di non avere preferenze nel liberare o acquistare elettroni. Questa proprietà rende questi atomi in grado di costituire facilmente legami covalenti, ma anche di dare luogo a una strana forma di conduzione elettrica mediante portatori “virtuali”, detti lacune, di cui parleremo nel seguito.

I quattro elettroni della corteccia più esterna dell'atomo di silicio, per quanto detto sopra, andranno a costituire quattro legami covalenti con altrettanti atomi vicini di silicio. La struttura spaziale che ne risulta è quella in Figura 1.12.

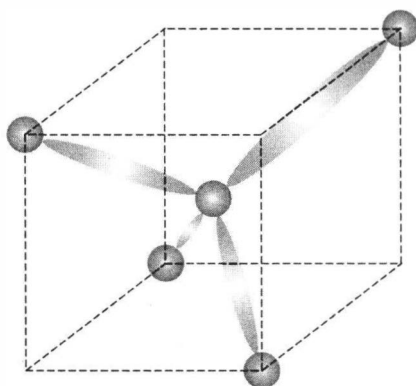


Figura 1.12 Disposizione degli atomi nel reticolo di silicio

Nel seguito, per comodità grafica disegneremo il reticolo cristallino del silicio come se fosse planare (vedi Figura 1.13).

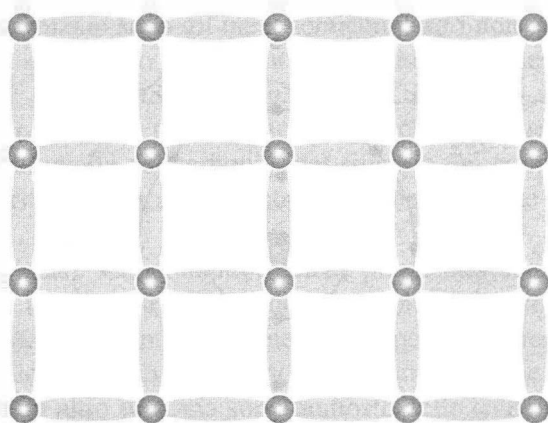


Figura 1.13 Rappresentazione planare del reticolo cristallino del silicio

1.7 Elettroni nei semiconduttori

Nel silicio gli elettroni più debolmente legati al reticolo cristallino sono gli elettroni dei legami covalenti. A causa dell'agitazione termica nei semiconduttori, come per gli isolanti, ci sarà una certa concentrazione di elettroni mobili.

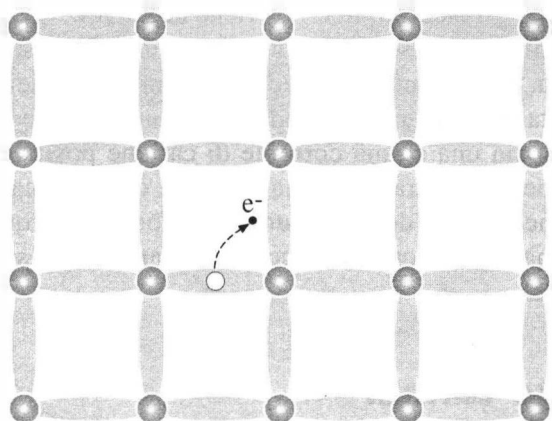


Figura 1.14 Elettrone di conduzione

Nel silicio a temperatura ambiente (300 K) la concentrazione di elettroni liberi è circa 10^{10} cm^{-3} . In un cristallo di silicio la distanza tra un atomo e l'altro è $5.4 \times 10^{-10} \text{ m}$, e quindi la densità degli atomi è circa $(5.4 \times 10^{-10} \text{ m})^{-3} \approx 6.4 \times 10^{21} \text{ cm}^{-3}$. Ogni atomo forma quattro legami covalenti, nei quali partecipano quattro elettroni. Due di questi possono diventare elettroni di conduzione. Quindi la concentrazione complessiva di elettroni potenzialmente in grado di diventare elettroni liberi è $2 \times 6.4 \times 10^{21} \text{ cm}^{-3} \approx 10^{22} \text{ cm}^{-3}$. Poiché la densità di elettroni liberi è circa 10^{10} cm^{-3} , risulta che a temperatura ambiente solo un legame ogni mille miliardi (10^{12}) è spezzato dall'agitazione termica e dà luogo a un elettrone mobile. Vale la pena di soffermarsi sul fatto che le correnti elettriche, di cui tanto ci occuperemo in questo libro, dal punto di vista fisico sono un effetto quasi marginale.

La stragrande maggioranza degli atomi di silicio non vi contribuisce in alcun modo. Se potessimo farci piccoli fino alle dimensioni degli atomi (come nel film di fantascienza *Viaggio allucinante*), vedere passare nelle vicinanze un elettrone mobile sarebbe un evento rarissimo, molto più raro di fare una passeggiata per il bosco vicino casa e trovare un orecchino d'oro. Eppure tutta l'elettronica si basa su questi pochi elettroni.

1.8 Lacune

Nei conduttori gli atomi che hanno liberato gli elettroni più esterni rimangono impacchettati strettamente l'uno all'altro legati dalla nube di elettroni mobili che li avvolge.

Negli isolanti, quando un legame covalente si spezza liberando un elettrone, l'atomo rimane legato agli atomi vicini mediante altri legami. Quindi, sia nei conduttori, sia negli isolanti, quando viene applicata una tensione le uniche cariche a muoversi sono gli elettroni liberi (cariche negative).

Gli ioni positivi, per quanto soggetti a un campo elettrico esterno dovuto alla tensione applicata, rimangono fissi. Vedremo tuttavia che nei semiconduttori quando è applicata una tensione si produce sia una corrente di cariche negative (gli elettroni) sia una strana corrente di cariche positive. E la cosa più sorprendente è che non si tratta di una corrente di ioni, ma di una sorta di corrente di cariche virtuali dette «lacune». Per parlare di questi portatori è necessario fare una digressione di meccanica quantistica e spiegare il cosiddetto effetto tunnel.

1.9 Effetto tunnel

Un elettrone legato a un atomo può essere visto come una particella dentro una buca di potenziale. La carica positiva dell'atomo genera un campo elettrico ($\propto 1/R^2$) e quindi un potenziale ($\propto 1/R$).

Un elettrone libero trova questa buca di potenziale e vi cade dentro, nel precipitare perde la propria energia potenziale sotto forma di calore e quindi non è più in grado di risalire, esattamente come il pallino cade in una buca nel gioco del biliardo. In fisica classica non c'è molto altro da dire: se qualcuno non va a riprendere il pallino questo rimarrà in buca per sempre. Invece in meccanica quantistica, ovvero nel mondo subatomico, possono succedere delle cose veramente strane.

A questo proposito il romanzo di avventura *Il Conte di Montecristo* di Alexandre Dumas ci può venire in aiuto. Si tratta della storia di Edmond Dantès, un giovane che ebbe la sfortuna di essere imprigionato nella terribile prigione del castello d'If nell'omonima isoletta davanti a Marsiglia, ma che riuscì a evadere mediante un tunnel. Proviamo a prendere ispirazione (con qualche licenza). Dumas descrive le celle sotterranee del castello come dei pozzi senza finestre e senza scale. Un carceriere passeggia tra un pozzo e l'altro e di tanto in tanto guarda dentro.

Se tutte le celle sono occupate, ogni volta che guarda dentro troverà un prigioniero. Ma supponiamo che non tutte siano occupate e che, ad esempio, la cella numero 28 sia vuota. Ebbene, una mattina durante il consueto giro di ispezione il carceriere trova la cella numero 27 vuota, subito in allarme controlla tutte le altre celle e scopre che sono tutte occupate: anche la 28 che il giorno prima era vuota.

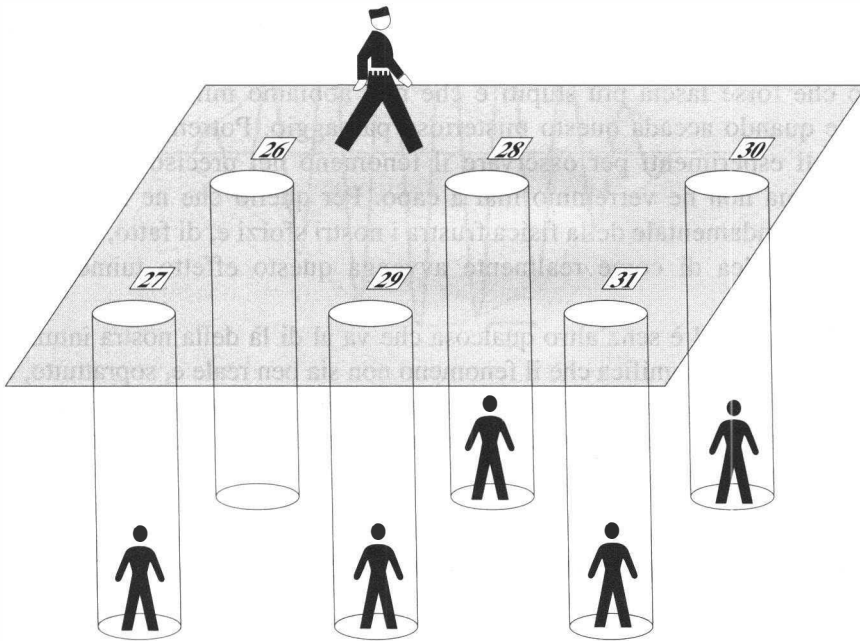


Figura 1.15 Il castello d'If

A questo punto si tranquillizza: non ci sono evasioni. Per scrupolo decide di scendere nelle celle e controllare con attenzione se tra la cella 28 e 27 ci sia un tunnel. Ma le pareti dei pozzi sono assolutamente intatte. Perplesso decide di soprassedere. Passa la notte. La mattina successiva trova che la cella 27 è vuota. Questa volta prima di allarmarsi controlla le celle vicine e, guarda caso, scopre che la cella 26 che il giorno prima era vuota ora è occupata. Ancora una volta scende nelle celle per cercare tunnel, ma non ne trova. Nei giorni successivi sempre la stessa storia. Se una cella è vuota, i prigionieri delle celle vicine di tanto in tanto passano in quella vuota. In sé il fenomeno non è molto grave, nessun prigioniero evade mai, ma il carceriere è un tipo pignolo e ligio alle regole, decide quindi di intensificare la sorveglianza: fa ronde ogni pochi minuti, passa la notte in piedi, ispezioni a sorpresa, ma niente da fare: i passaggi tra una cella e l'altra avvengono sempre quando il carceriere sta guardando da un'altra parte.

Quello che abbiamo raccontato può sembrare solo una stramba storia di fantasia, ma in realtà è ciò che accade continuamente in un semiconduttore.

Quando un elettrone si libera da un legame covalente, si crea una buca di potenziale vuota e gli elettroni dei legami covalenti vicini possono andarla ad occupare lasciando a loro volta una buca di potenziale vuota. Il fenomeno è del tutto analogo, e ugualmente misterioso, a quello descritto sopra. Gli elet-

troni dentro le buche di potenziale non hanno abbastanza energia per saltare fuori e cadere in un'altra, eppure passano da una buca all'altra, come se ci fosse un tunnel (per questo si chiama effetto tunnel). Ma *non* c'è alcun tunnel. E ciò che forse lascia più stupiti è che non abbiamo minimamente idea di come e quando accada questo misterioso passaggio. Potremmo ideare i più sofisticati esperimenti per osservare il fenomeno nel preciso istante in cui avviene, ma non ne verremmo mai a capo. Per quello che ne sappiamo un principio fondamentale della fisica frustra i nostri sforzi e, di fatto, non abbiamo alcuna idea di come realmente avvenga questo effetto tunnel "senza tunnel".

L'effetto tunnel è senz'altro qualcosa che va al di là della nostra intuizione fisica, ma ciò non significa che il fenomeno non sia ben reale e, soprattutto, ben inquadrato in una teoria in grado di dare previsioni quantitative e risultati certi (non si tratta di magia!). Il concetto fondamentale alla base dell'effetto tunnel è la dualità onda-corpuscolo.

1.10 Dualità onda-corpuscolo

Un modo semplice per immaginare un elettrone in moto da un punto all'altro nello spazio è pensare a una pallina molto piccola che segue una certa traiettoria. Questo modo di vedere l'elettrone, così come qualunque altra particella o corpuscolo, non è sbagliato, è solo una descrizione a larga scala, è come dire che una città è un pallino sulla carta geografica: è vero finché non ci avviciniamo troppo.

Se ingrandiamo sull'elettrone ci accorgiamo di due cose per certi aspetti contraddittorie: la prima è che l'elettrone a qualunque scala è sempre un punto, non vediamo mai la struttura interna; la seconda è che man mano che ci avviciniamo la posizione dell'elettrone è sempre meno definita, talmente poco definita che siamo costretti a parlare di un'onda di probabilità: un elettrone può essere visto come un'onda che si propaga. Anzi le cose sono anche più strane: ciò che si propaga, seguendo una particolare equazione d'onda, è una grandezza un po' fantomatica detta φ , funzione dello spazio e del tempo. Questa grandezza è un numero complesso, ovvero provvisto di parte reale e parte immaginaria. Il quadrato del modulo (ovvero la somma dei quadrati della parte reale e della parte immaginaria) di φ è la probabilità di trovare l'elettrone in una certa posizione a un dato istante. Sembra tutto molto strano e in effetti lo è, ma la natura è fatta così e non possiamo farci molto.

Abbiamo detto che $\varphi(x, t)$ è un'onda in quanto rispetta una particolare equazione che ha per soluzioni funzioni dello spazio e del tempo che sembrano propagarsi come siamo abituati a vedere propagarsi le onde. Per inciso, si tratta dell'equazione di Schrödinger, una delle tre o quattro equazioni alla base della fisica così come la conosciamo.

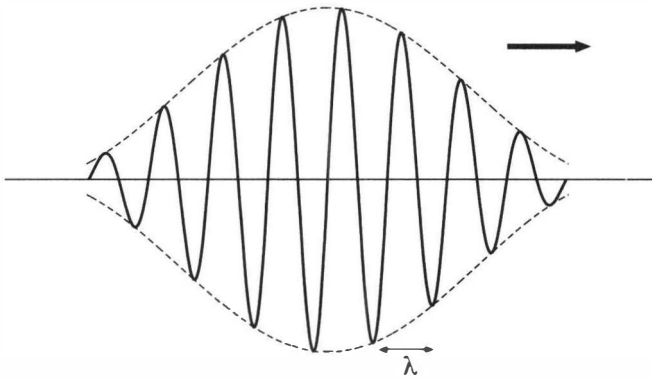


Figura 1.16 Pacchetto d'onda associato a un elettrone in moto

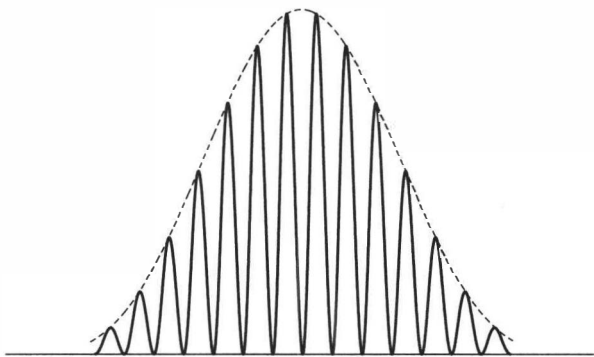


Figura 1.17 Distribuzione della probabilità di trovare l'elettrone nello spazio

Un elettrone che si muove nello spazio può, quindi, essere visto come un pacchetto d'onde con una sua propria lunghezza d'onda. La dimensione del pacchetto potrà essere più o meno grande, ma non potrà mai scendere sotto la lunghezza d'onda, questa è una proprietà piuttosto generale delle onde: se il pacchetto non contiene almeno una lunghezza d'onda, il concetto stesso di lunghezza d'onda perde di significato. In altre parole ciò significa che è non possibile confinare un elettrone in uno spazio più piccolo della sua lunghezza d'onda.

D'altra parte la probabilità di trovare un elettrone in una data posizione è proporzionale al quadrato del modulo della funzione d'onda, ma in generale il quadrato dell'ampiezza di un'onda è un'onda a frequenza doppia, quindi la distanza entro la quale l'elettrone è confinato è la semilunghezza d'onda della sua funzione d'onda.

De Broglie, nel 1924, scoprì una relazione fondamentale tra velocità dell'elettrone e lunghezza d'onda della funzione d'onda:

$$\lambda = \frac{h}{mv} \quad (1.31)$$

h è la costante di Planck (6.63×10^{-34} J s), m è la massa dell'elettrone (9.1×10^{-31} kg).

In realtà, la (1.31) vale per qualunque oggetto, particella subatomica, atomo o anche corpi macroscopici, ma h è molto piccolo e per m più grande di pochi atomi, la lunghezza d'onda, anche solo con la velocità impressa dall'agitazione termica, diventa così incredibilmente piccola che può essere considerata nulla ad ogni effetto pratico.

Proviamo a stimare la lunghezza d'onda di un elettrone in una buca di potenziale. A causa dell'agitazione termica avrà un'energia cinetica E_k pari a

$$E_k = \frac{1}{2}mv^2 \quad (1.32)$$

con v velocità media. Sostituendo la (1.32) nella (1.31):

$$\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2mE_k}} \quad (1.33)$$

D'altra parte l'energia cinetica media lungo una qualsiasi direzione di un elettrone a temperatura ambiente è:

$$E_k = \frac{1}{2}kT \approx 13\text{meV} \quad (1.34)$$

con k costante di Boltzmann e T temperatura ambiente. Sostituendo nella (1.33) si ottiene $\lambda = 10\text{nm}$, ovvero una dimensione minima della regione di confinamento dell'elettrone pari a 5nm ($\lambda/2$). Ma la distanza tra gli atomi di silicio è $5.4 \times 10^{-10}\text{m}$ (10 volte più piccola della dimensione minima di confinamento), in altre parole la funzione d'onda dell'elettrone nella buca di potenziale è più grande della buca di potenziale! Il significato fisico è che l'elettrone ha una probabilità non nulla di "materializzarsi" in un punto anche piuttosto lontano dalla buca di potenziale dove è confinato (anche senza avere abbastanza energia per saltare fuori dalla buca di potenziale!) e se in questo punto casualmente si trova una buca di potenziale vuota, ecco che avviene una sorta di teletrasporto: l'elettrone si sposta. Come avviene nei dettagli questo spostamento? Per quello che ne sappiamo esattamente come avviene in Star Trek: *"Energia! Scotty!" e il capitano Kirk scompare dalla plancia dell'Enterprise per comparire dopo pochi istanti sul pianeta dei Klingon.*

1.11 Silicio intrinseco

Facciamo ora il punto di ciò che abbiamo detto fin qui. Il silicio forma un reticolo cristallino di legami covalenti. Se la presenza di eventuali atomi diversi dal silicio è trascurabile si parla di silicio intrinseco. L'agitazione termica spezza alcuni legami e fa sì che nel materiale ci sia una certa concentrazione n_i di elettroni liberi che si muovono casualmente all'interno del reticolo. Quando è applicato un campo elettrico, a questo moto casuale si sovrappone una piccola velocità di deriva v_n , proporzionale al campo elettrico:

$$v_n = -\mu_n E \quad (1.35)$$

la costante μ_n è detta mobilità degli elettroni: nel silicio è circa $1350 \text{ cm}^2/(\text{Vs})$.

Per ogni elettrone che si libera dal proprio legame covalente, si crea una buca di potenziale vuota. La particolarità dei semiconduttori è che, poiché la funzione d'onda degli elettroni dei legami covalenti vicini è più grande delle rispettive buche di potenziale, c'è una probabilità tutt'altro che trascurabile che un elettrone di un legame vicino si "materializzi" nella buca di potenziale vuota, lasciando a sua volta una buca di potenziale, e così via anche per questa buca di potenziale. Insomma, è come se le buche di potenziale vuote migrassero di moto casuale all'interno del silicio (queste buche migranti sono dette «lacune»). Poiché a queste buche di potenziale vuote è associata una carica positiva (quella dell'atomo privo di un elettrone associato a ciascun legame covalente spezzato) possiamo immaginarle come cariche positive mobili. In realtà nessuna carica positiva si muove realmente, sono gli elettroni che passano da una buca all'altra per effetto tunnel, ma a tutti gli effetti pratici è come se ci fossero cariche positive mobili. E, in effetti, l'analogia tra lacune e cariche positive è così forte che quanto detto a proposito delle cariche mobili e della legge di Ohm vale anche per le lacune (che, ripetiamo, non sono particelle, ioni o altri corpuscoli mobili). Anche per le lacune si può parlare di velocità di deriva v_p , proporzionale al campo elettrico applicato:

$$v_p = \mu_p E \quad (1.36)$$

Il segno positivo (contrariamente al caso degli elettroni) indica che le lacune si muovono nella direzione positiva del campo elettrico. Si noti, tuttavia, che la mobilità delle lacune μ_p non è uguale a quella degli elettroni ma sensibilmente minore: $\mu_p = 550 \text{ cm}^2/(\text{Vs})$.

Dal punto di vista fisico quando si applica una tensione si creano due flussi di elettroni (vedi Figura 1.18):

- 1) un flusso di elettroni liberi nel senso opposto al campo elettrico;

- 2) un flusso di elettroni “sotterraneo” determinato dal fatto che gli elettroni, pur non avendo energia sufficiente per uscire dalle buche di potenziale dei legami covalenti, passano per effetto tunnel da un legame occupato a un legame privo di un elettrone (lacuna). Il moto degli elettroni è nel senso opposto al campo elettrico, ma il fenomeno può essere visto anche come un moto di lacune (legami covalenti privi di un elettrone) nel senso del campo elettrico.

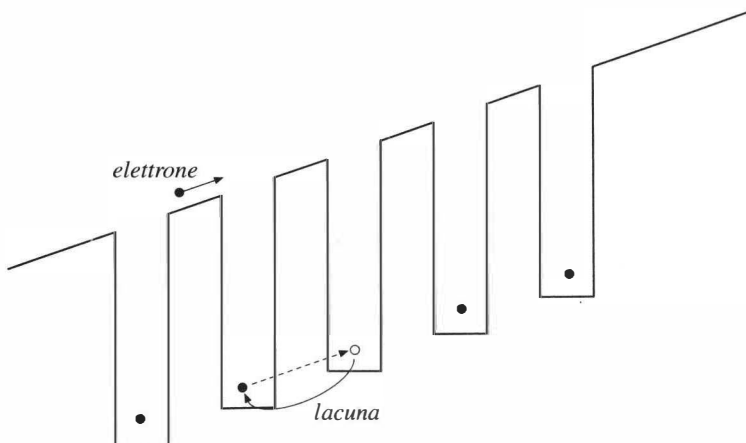


Figura 1.18 Elettroni e lacune

Quando è applicato un campo elettrico, elettroni e lacune contribuiscono entrambi alla densità di corrente locale j :

$$j = j_n + j_p = -qn_i v_n + qn_i v_p = qn_i \mu_n E + qn_i \mu_p E \quad (1.37)$$

Si noti che elettroni e lacune contribuiscono con lo stesso segno alla corrente che quindi è in valore assoluto maggiore rispetto al caso in cui ci fossero solo elettroni. Tuttavia, sebbene il numero di elettroni e lacune sia lo stesso in un semiconduttore intrinseco, il contributo della corrente di elettroni è circa 2.5 volte maggiore del contributo della corrente di lacune, a causa della loro maggior mobilità.

1.12 Silicio drogato con impurità di tipo n

Nel silicio intrinseco elettroni e lacune sono in ugual numero e la loro concentrazione, a temperatura ambiente, è pari a circa 10^{10}cm^{-3} . Vedremo ora come l'introduzione nel reticolo di opportune impurità permette di controllare la concentrazione dell'uno o dell'altro portatore.

Il silicio è un atomo in IV colonna che quindi forma 4 legami covalenti. Durante il processo di formazione del reticolo si possono aggiungere alcuni atomi in V

colonna, ad esempio il fosforo (P). Questi atomi sono detti donatori, oppure impurità di tipo n, il semiconduttore con tali impurità è detto di tipo n. Un semiconduttore che contiene impurità di tipo n (o tipo p come vedremo in seguito) è detto drogato, le impurità sono dette anche droganti.

In un semiconduttore drogato la concentrazione delle impurità deve essere tale che gli atomi di silicio rimangano fortemente maggioritari. In tal caso il reticolo cristallino si formerà come se ci fossero solo atomi di silicio, con l'unica differenza che, di tanto in tanto, un atomo di silicio sarà sostituito da un atomo di fosforo. Questo si legherà, come il silicio, ai quattro atomi di silicio vicini mediante legami covalenti, ma rimarrà con il quinto elettrone nella corteccia più esterna senza la possibilità di legarsi.

Questo elettrone spaiato si troverà in una condizione di bassa energia (una buca di potenziale meno profonda delle buche di potenziale dove sono confinati gli elettroni dei legami covalenti) per cui l'agitazione termica lo porterà facilmente in conduzione.

Di fatto a temperatura ambiente tutte le impurità di tipo n risulteranno ionizzate. In altre parole tutte le impurità libereranno un elettrone mobile e andranno a costituire ioni positivi fissi, ben legati al reticolo cristallino. Gli elettroni mobili nel reticolo saranno quindi la somma degli elettroni prodotti dagli atomi di fosforo e dagli elettroni liberati dai legami covalenti. Questi ultimi al più saranno pari alla concentrazione degli elettroni nel silicio intrinseco (10^{10}cm^{-3}), ma se la concentrazione delle impurità di fosforo è molto maggiore, possiamo trascurare il contributo degli elettroni liberati dai legami covalenti e supporre il numero di elettroni (n) pari al numero di impurità donatrici (N_D)

$$n \approx N_D \quad (1.38)$$

È importante notare che la concentrazione di impurità deve essere molto minore della densità degli atomi di silicio, ma molto maggiore della concentrazione degli elettroni liberi nel silicio intrinseco, ovvero

$$10^{10}\text{cm}^{-3} \ll N_D \ll 10^{22}\text{cm}^{-3} \quad (1.39)$$

Fortunatamente ci sono ben 12 ordini di grandezza!

La presenza di droganti nel semiconduttore, oltre a modificare la concentrazione dei portatori, modifica la mobilità degli elettroni. Infatti la presenza di impurità produce deformazioni del reticolo cristallino che ostacolano il movimento degli elettroni. E anche la sola presenza di ioni fissi aumenta la probabilità di interazione degli elettroni con il reticolo, riducendone la velocità. La mobilità di un semiconduttore molto drogato può essere anche dieci volte inferiore al semiconduttore intrinseco.

1.13 Silicio drogato con impurità di tipo p

In alternativa ad impurità in V colonna, durante il processo di produzione di un cristallo di silicio, si possono introdurre impurità in III colonna, come ad esempio il boro (B). Questi sono detti atomi accettori, oppure impurità di tipo p. Il semiconduttore drogato con accettori è detto di tipo p.

Un atomo in III colonna ha tre elettroni nella corteccia più esterna e quindi può costituire tre legami covalenti. Se un atomo di boro sostituisce un atomo di silicio nel reticolo, uno dei quattro atomi vicini non può legarsi al boro. Questo atomo di silicio si troverà quindi con solo tre legami covalenti, pur potendone formare quattro. Possiamo pensare questa condizione come se in corrispondenza del quarto elettrone ci fosse un posto vuoto, ovvero una buca di potenziale. Questa buca è più profonda rispetto alle buche di potenziale relative agli elettroni dei legami vicini. Per effetto tunnel un elettrone dei legami covalenti vicini può “materializzarsi” in questa buca di potenziale, dalla quale non potrà più uscire, essendo questa buca più profonda delle altre. Il risultato è: da un lato un elettrone imprigionato, ovvero uno ione negativo fisso, dall’altro una lacuna mobile. In altre parole per ogni impurità accettore si avrà uno ione fisso negativo e una lacuna positiva. Se la concentrazione N_A delle impurità di tipo p è molto maggiore della concentrazione delle lacune nel silicio intrinseco, la concentrazione delle lacune sarà circa uguale a quella delle impurità

$$p \approx N_A \quad (1.40)$$

Come nel caso delle impurità di tipo n, anche per le impurità di tipo p dovrà valere la condizione

$$10^{10} \text{ cm}^{-3} \ll N_A \ll 10^{22} \text{ cm}^{-3} \quad (1.41)$$

Come per il caso dei semiconduttori n, la presenza di droganti nel semiconduttore, oltre a modificare la concentrazione dei portatori, riduce la mobilità dei portatori (lacune, in questo caso).

1.14 Legge di azione di massa

L’introduzione di impurità di tipo n nel reticolo cristallino del silicio permette di controllare il numero di elettroni liberi mediante la (1.38); analogamente l’introduzione di impurità di tipo p permette di controllare il numero di lacune mediante la (1.40). Cosa succede se sono presenti impurità dei due tipi? Possiamo controllare indipendentemente l’una dall’altra le concentrazioni di elettroni e di lacune? La risposta è no. Donatori e accettori si annullano reciprocamente, ciò significa che se in un semiconduttore c’è una concentrazione N_D ,

di atomi donatori e una concentrazione N_A di atomi accettori, se $N_D > N_A$ la situazione è equivalente ad avere una concentrazione di donatori pari a $N_D - N_A$. Viceversa se $N_A > N_D$, la situazione è equivalente ad avere una concentrazione di accettori pari a $N_A - N_D$. In altre parole, è possibile controllare, mediante impurità, solo una delle due concentrazioni: elettroni oppure lacune, non entrambi. La domanda è quindi: fissata la concentrazione degli elettroni mediante impurità di tipo n, qual è la concentrazione delle lacune? E viceversa: fissata la concentrazione delle lacune mediante impurità di tipo p, qual è la concentrazione degli elettroni? La risposta è fornita dalla cosiddetta **legge di azione di massa**.

Elettroni e lacune si muovono casualmente all'interno di un semiconduttore. Quando si trovano molto vicini l'uno all'altra l'elettrone è intrappolato nella lacuna, il legame covalente è ricostituito ed entrambi cessano di esistere come portatori mobili di carica. Questo fenomeno è detto ricombinazione.

Prendiamo in considerazione una scacchiera di N caselle dove posizionare N_n pedine bianche e N_p pedine nere.

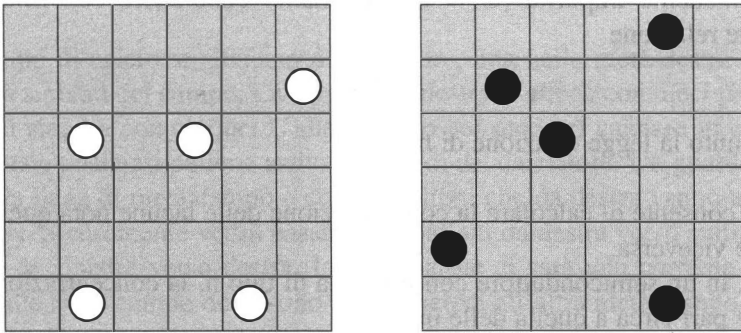


Figura 1.19 Legge di azione di massa

La probabilità che una pedina bianca si trovi in una particolare casella sarà N_n/N , analogamente la probabilità che una pedina nera si trovi in una particolare casella sarà N_p/N . La probabilità che entrambe le pedine si trovino nella stessa casella sarà il prodotto delle due probabilità. Per ottenere il numero totale di caselle in cui ci sono entrambe le pedine dovremo moltiplicare questa probabilità per il numero di caselle ottenendo $(N_n/N)(N_p/N)N$.

Se ora torniamo ai semiconduttori, possiamo ragionevolmente pensare che il tasso di ricombinazione per unità di volume R , ovvero il numero di coppie elettrone-lacuna che si ricombinano nell'unità di tempo cessando di essere portatori mobili, sarà proporzionale al prodotto delle concentrazioni degli elettroni e delle lacune:

$$R = rnp \quad (1.42)$$

con r opportuna costante. Abbiamo già detto che l'agitazione termica provoca la rottura dei legami covalenti e la generazione di coppie elettrone-lacuna. Fissata una temperatura, ci sarà quindi un tasso di generazione G . Questo tasso è circa lo stesso sia per il silicio intrinseco sia per il silicio drogato. Infatti, la concentrazione delle impurità è sempre molto minore della densità dei legami covalenti e, quindi, qualunque effetto possano avere è trascurabile rispetto alla generazione delle coppie elettrone-lacuna del solo silicio.

D'altra parte, il tasso di generazione deve essere uguale al tasso di ricombinazione. Infatti se $G > R$, si creano più elettroni e lacune di quanti se ne ricombinano, quindi n e p aumentano. Segue che il prodotto np aumenta, ovvero aumenta $R = np$ finché $R = G$. Analogamente se $G < R$, R diminuisce finché $R = G$. Pertanto:

$$G = R = np \quad (1.43)$$

Per il silicio intrinseco potremmo scrivere

$$G_i = R = np = n_i^2 \quad (1.44)$$

Ma poiché il tasso di generazione è lo stesso per il silicio intrinseco e per il silicio contenente impurità ($G = G_i$), possiamo concludere con la seguente importante relazione

$$np = n_i^2 \quad (1.45)$$

che è appunto la legge di azione di massa.

La (1.45) consente di calcolare la concentrazione delle lacune nota quella degli elettroni e viceversa.

Infatti, in un semiconduttore con impurità di tipo n, la concentrazione degli elettroni è pari circa a quella delle impurità

$$n \approx N_D \quad (1.46)$$

pertanto

$$p \approx \frac{n_i^2}{N_D} \quad (1.47)$$

Poiché deve valere la condizione $10^{10} \text{ cm}^{-3} \ll N_D \ll 10^{22} \text{ cm}^{-3}$, prendiamo ad esempio $N_D = 10^{13} \text{ cm}^{-3}$, d'altra parte $n_i = 10^{10} \text{ cm}^{-3}$, segue che $p = 10^7 \text{ cm}^{-3}$. Riflettiamo su questi numeri: in un semiconduttore di tipo n, ovvero drogato con impurità di tipo n, la concentrazione degli elettroni è enormemente superiore alla concentrazione delle lacune (nell'esempio 6 ordini di grandezza, ma potrebbe arrivare anche a 12 ordini di grandezza!).

Il caso del silicio drogato con atomi accettori è del tutto analogo. Poiché

$$p \approx N_A \quad (1.48)$$

la concentrazione degli elettroni sarà:

$$n \approx \frac{n_i^2}{N_A} \quad (1.49)$$

Anche per impurità di tipo p deve valere la condizione $10^{10} \text{ cm}^{-3} \ll N_A \ll 10^{22} \text{ cm}^{-3}$, segue che in un semiconduttore di tipo p la concentrazione delle lacune è enormemente superiore alla concentrazione degli elettroni.

1.15 Correnti di diffusione

In un materiale in cui possono muoversi portatori di carica, l'applicazione di una tensione (e quindi localmente di un campo elettrico) produce una corrente, detta corrente di deriva. Tale corrente segue la legge di Ohm ed è proporzionale alla tensione applicata. Tuttavia, in generale, può formarsi una corrente anche in assenza di un campo elettrico, tale corrente è detta corrente di diffusione. Le correnti di diffusione, come vedremo, svolgono un ruolo fondamentale nel funzionamento dei diodi e dei transistor bipolari. Per comprenderne il meccanismo ricorremo a un esempio.

In un campo di calcio si allenano due squadre, una nella metà destra e l'altra nella metà sinistra del campo. La squadra di destra si allena con dieci palloni, la squadra di sinistra con cinque. L'allenamento consiste nel calciare in una direzione casuale qualsiasi pallone arrivi tra i piedi del calciatore. Un guardalinee si mette sulla linea di metà campo e conta i palloni che da destra vanno a sinistra e viceversa. Sicuramente vedrà passare più palloni da destra verso sinistra piuttosto che da sinistra verso destra. In altre parole ci sarà una corrente netta di palloni dalla metà campo dove sono più numerosi all'altra metà campo.

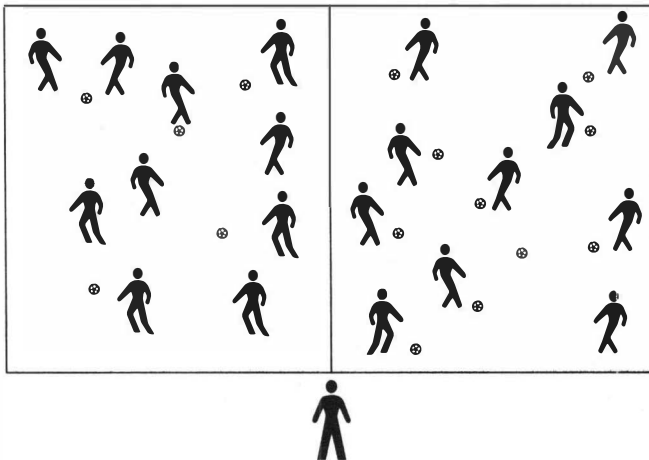


Figura 1.20 Corrente di diffusione

Si noti bene che non c'è nessuna forza esterna che spinge i palloni in una direzione anziché in un'altra, si tratta solo di una questione di probabilità: è più probabile che un pallone vada dalla metà campo dove sono più numerosi a quella dove sono meno numerosi piuttosto che viceversa. Questa è una corrente di diffusione. Se il campo di calcio avesse una pendenza (un campo elettrico applicato) alla corrente di diffusione si sommerebbe una corrente di deriva, in quanto i palloni tra un calcio e l'altro tenderebbero ad acquistare una piccola velocità nella direzione della pendenza.

Se ora consideriamo una distribuzione continua di concentrazione dei palloni (non semplicemente un certo numero in una metà campo e un numero differente nell'altra) è ragionevole pensare che la densità di corrente locale j_p sia proporzionale al gradiente della concentrazione $p(x)$ con verso positivo nella direzione in cui il gradiente diminuisce, quindi nel caso monodimensionale

$$j_p = -qD_p \frac{dp(x)}{dx} \quad (1.50)$$

q è la carica elettrica portata da ciascun pallone, la costante D_p è detta coefficiente di diffusione. La (1.50), fuor di metafora, è l'equazione della corrente di diffusione delle lacune (palloni con carica positiva). Per gli elettroni, portatori di carica negativa, la corrente elettrica sarà di segno opposto al flusso degli elettroni e quindi

$$j_n = qD_n \frac{dn(x)}{dx} \quad (1.51)$$

D_n è il coefficiente di diffusione degli elettroni.

Riassumendo, in un semiconduttore possono sussistere quattro differenti correnti: corrente di deriva degli elettroni, corrente di diffusione degli elettroni, corrente di deriva delle lacune, corrente di diffusione delle lacune. La densità di corrente locale j è la somma di questi quattro contributi:

$$j = qn\mu_n E + qD_n \frac{dn(x)}{dx} + qp\mu_p E - qD_p \frac{dp(x)}{dx} \quad (1.52)$$

μ_n e μ_p sono rispettivamente la mobilità degli elettroni e delle lacune.

Nella (1.52) le correnti di diffusione e di deriva dipendono da due costanti diverse: il coefficiente di diffusione (D) e la mobilità (μ). Tuttavia, per quanto diverse, una relazione profonda e per molti aspetti sorprendente le unisce:

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{kT}{q} \quad (1.53)$$

dove k è la costante di Boltzmann, T la temperatura e q la carica elementare dell'elettrone. La (1.53) è detta **relazione di Einstein** e, come vedremo, ha un ruolo fondamentale nel principio di funzionamento dei dispositivi a semiconduttore.

kT/q è dimensionalmente una tensione ed è detto potenziale termico V_T

$$V_T = \frac{kT}{q} \quad (1.54)$$

A temperatura ambiente (300 K) il potenziale termico è 26mV.

1.16 Potenziale di built-in

Prendiamo ora in considerazione un semiconduttore (ad esempio p) in cui il drogaggio non sia costante lungo un asse ma, giusto per fissare le idee, sia decrescente come in Figura 1.21.

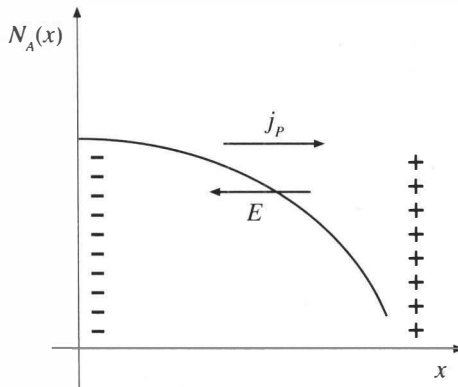


Figura 1.21 Gradiente nella concentrazione delle impurità p

Nella zona in cui il drogaggio è maggiore ci saranno più lacune. Queste muovendosi del tutto casualmente tenderanno a diffondersi verso la zona a minore concentrazione. Ma essendo cariche positive, produrranno uno sbilanciamento di cariche e quindi un campo elettrico che si oppone al moto di diffusione. In definitiva la corrente di diffusione in un verso produce un campo elettrico nel verso opposto che a sua volta crea una corrente di deriva di segno opposto rispetto alla corrente di diffusione. Queste due correnti in condizioni di equilibrio si devono compensare, pertanto deve valere

$$j_p = qp\mu_p E - qD_p \frac{dp(x)}{dx} = 0 \quad (1.55)$$

ovvero, ricordando la relazione di Einstein:

$$pE = V_T \frac{dp(x)}{dx} \quad (1.56)$$

Questo punto è fondamentale per capire la fisica dei semiconduttori. Cerchiamo di visualizzarlo con un'immagine intuitiva.

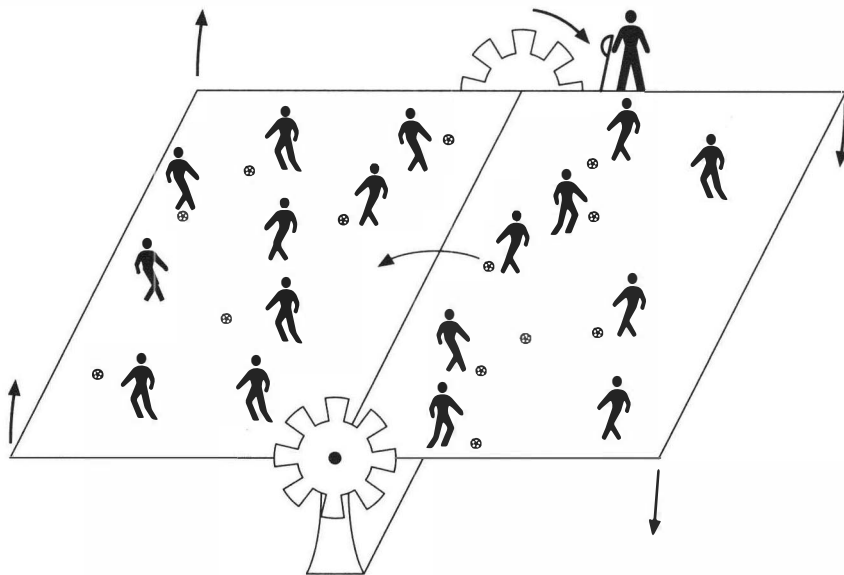


Figura 1.22 Campo di calcio basculante

Torniamo al campo di calcio della Figura 1.20. I giocatori continuano ad allenarsi con un gran numero di palloni. Inizialmente sono molti di più in una metà campo che nell'altra. Introduciamo qualche modifica tecnica. Ora il campo è in equilibrio su un meccanismo basculante in modo da poterne controllare l'inclinazione. In questo modo simuliamo la forza che determina la corrente di deriva. L'arbitro sulla metà campo ogni volta che vede una palla andare da destra verso sinistra con un apposita leva inclina il campo verso destra di un dente dell'ingranaggio che ruota il piano. E viceversa: quando una palla va verso destra, ruota il campo di un dente nell'altro verso. In questo modo si crea un equilibrio. Il campo risulterà un po' inclinato verso il lato con maggiore densità di palle, queste tenderanno ad essere calciate più verso la zona a bassa concentrazione, ovvero verso la metà del campo più in alto, ma la traiettoria un po' incurvata verso il basso compensa l'effetto. Questo bilanciamento fa sì che il numero di palle in una metà campo rimanga praticamente inalterato rispetto alla condizione iniziale.

Ma vediamo ora un caso un po' più realistico di un "campo di calcio basculante". Prendiamo in considerazione una barretta di semiconduttore lunga w con una distribuzione di impurità p esponenzialmente decrescente (Figura 1.23).

$$p(x) \approx N_A(x) = N_0 e^{-\frac{x}{L}} \quad (1.57)$$

Una distribuzione del genere si può fisicamente realizzare mediante diffusione di un gas di impurità.

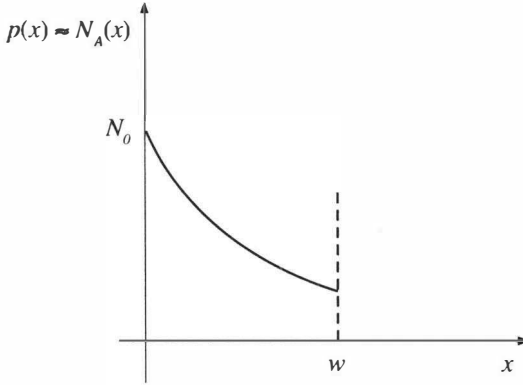


Figura 1.23 Distribuzione esponenziale di impurità p

Sostituendo la (1.57) nella (1.56) si ottiene che all'interno del materiale è presente un campo elettrico negativo costante pari a

$$E = -\frac{V_T}{L} \quad (1.58)$$

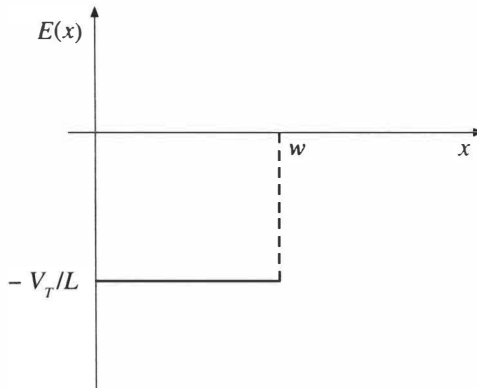


Figura 1.24 Campo elettrico nel semiconduttore

Un campo elettrico costante significa, per la legge di Poisson, che non ci sono cariche non compensate all'interno del materiale. Tuttavia, poiché il campo è

nullo esternamente ci deve essere una distribuzione di cariche superficiali non compensate (Figura 1.25) che si deve essere formata durante il transitorio prima dell'equilibrio.

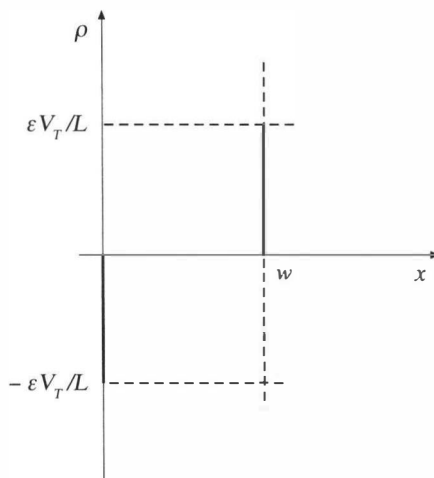


Figura 1.25 Distribuzione di cariche non compensate

La densità di questa carica superficiale si calcola con la legge di Gauss:

$$\int_s \vec{E} \cdot \vec{n} dS = \frac{Q}{\epsilon} \quad (1.59)$$

dove $\vec{E}$ è il campo elettrico, S è una qualunque superficie chiusa, $\vec{n}$ è il versore sulla superficie, Q è la carica in essa racchiusa. Prendendo in considerazione un parallelepipedo attraverso il quale passa una delle due facce del materiale si ottiene

$$E \Delta S = \frac{Q}{\epsilon} \quad (1.60)$$

quindi

$$E = \frac{\rho}{\epsilon} \quad (1.61)$$

con ρ densità di carica per unità di superficie. Sostituendo la (1.58) nella (1.61) si ottiene la densità di carica sulle facce del semiconduttore

$$\rho = \epsilon \frac{V_T}{L} \quad (1.62)$$

Un campo elettrico costante come in Figura 1.24 produce un potenziale crescente come in Figura 1.26, che prende il nome di **potenziale di built-in**. Tale potenziale non è direttamente misurabile perché sempre compensato dalle cariche di superficie ma, come vedremo nei capitoli seguenti, è alla base del funzionamento del diodo e del transistor bipolare.

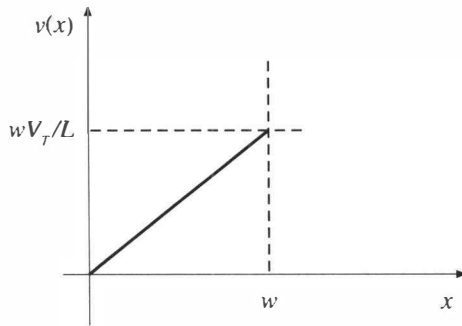


Figura 1.26 Potenziale elettrico nel semiconduttore

Vogliamo ora trovare un'espressione generale che leghi la distribuzione di portatori al potenziale di built-in. La (1.56) può essere riscritta come segue

$$Edx = V_T \frac{dp}{p} \quad (1.63)$$

Integriamo entrambi i membri

$$\int_{x=0}^x Edx = V_T \int_{p(0)}^{p(x)} \frac{dp}{p} \quad (1.64)$$

da cui

$$v(x) = -V_T \ln \left(\frac{p(x)}{p(0)} \right) \quad (1.65)$$

ovvero

$$p(x) = p(0) e^{-\frac{v(x)}{V_T}} \quad (1.66)$$

che è l'espressione che lega le concentrazioni di portatori p in due generici punti ($x = 0$ e x) al potenziale elettrico tra gli stessi due punti. Quanto detto sopra vale ovviamente anche per i portatori n , con l'unica differenza del segno nell'argomento dell'esponenziale:

$$n(x) = n(0) e^{\frac{v(x)}{V_T}} \quad (1.67)$$

1.17 Energia di Fermi

In un conduttore gli atomi mettono in compartecipazione l'elettrone (o gli elettroni) più esterno. Questi vanno a costituire una sorta di nube elettronica che avvolge l'intero reticolo cristallino.

L'applicazione dell'equazione di Schrödinger a un elettrone di conduzione di un metallo consente di calcolare i suoi possibili stati quantici in funzione dell'energia. Risulta che tali stati sono discreti, ma molto vicini l'uno all'altro, tanto da poter essere trattati come un continuo, si può pertanto parlare di densità di stati per unità di energia. Inoltre tali stati aumentano all'aumentare delle dimensioni fisiche del semiconduttore ed è quindi ragionevole definire una densità di stati $\rho(E)$ per unità di energia e unità volume.

Gli elettroni hanno una proprietà quantistica detta spin che può assumere due valori $1/2$ e $-1/2$. Il **principio di esclusione di Pauli** prescrive che due atomi di stesso spin non possono occupare lo stesso stato quantico e pertanto in un conduttore non più di due elettroni possono avere la stessa energia. Si noti che si tratta di un risultato importante e sorprendente. Nel mondo classico nulla impedisce alle auto di una strada di procedere alla stessa velocità l'una dietro all'altra. Nel mondo quantistico questo è vietato; solo due elettroni – e solo se di spin diversi – possono muoversi alla stessa velocità, ovvero avere la stessa energia.

Gli urti causati dall'agitazione termica tra gli atomi del reticolo e gli elettroni fanno sì che questi ultimi saltino continuamente da uno stato all'altro. Questi salti seguono una particolare statistica (detta **statistica di Fermi-Dirac**) che, al contrario di quella classica derivata da Boltzmann, è compatibile con il principio di esclusione di Pauli. Tale statistica prevede che la probabilità di occupazione $p(E)$ di uno stato quantico di energia E sia pari a

$$p(E) = \frac{1}{1 + e^{\frac{E - E_F}{kT}}} \quad (1.68)$$

dove E_F è una costante detta energia di Fermi. Si noti che la probabilità di occupazione di un livello energetico all'energia di Fermi è pari al 50%.

L'energia di Fermi è un parametro fondamentale per le considerazioni che faremo nel seguito ed è quindi importante averne anche un'idea intuitiva, oltre che una definizione. Immaginiamo un materiale i cui stati energetici siano uniformemente distribuiti al variare dell'energia (questo materiale non esiste, ma è un'utile astrazione). Se non ci fosse agitazione termica tutti gli elettroni mobili si sistemerebbero su questi stati due per ogni stato partendo da quello a energia più bassa. Il livello energetico che gli elettroni raggiungono, se impilati come abbiamo detto, è l'energia di Fermi. Per visualizzarla, possiamo pensare che i livelli energetici siano come un bicchiere, l'acqua sono gli elettroni, il livello che l'acqua raggiunge quando è versata nel bicchiere è l'energia di Fermi.

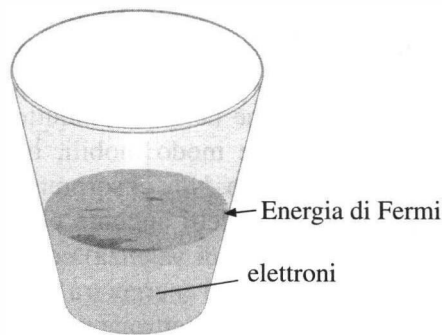


Figura 1.27 Energia di Fermi

A temperatura diversa dallo zero assoluto, gli urti tra reticolo ed elettroni fanno sì che questi saltino da un livello all'altro e che non ci sia più un confine netto tra stati occupati e stati vuoti, ma una probabilità di occupazione che segue la (1.68). Come mostrato in Figura 1.28, al diminuire della temperatura la distribuzione è sempre più ripida, fino a diventare uno scalino allo zero assoluto. L'energia di Fermi non varia al variare della temperatura e dipende essenzialmente dal numero di elettroni liberi per unità di volume. Questo numero può essere alterato inserendo nel metallo degli atomi trappola, in grado di legare un elettrone. All'aumentare della densità degli atomi trappola, l'energia di Fermi cala.

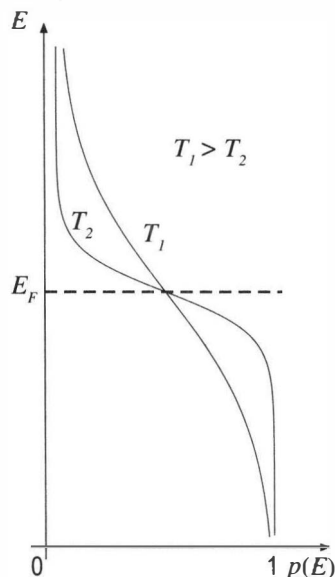


Figura 1.28 Distribuzione di Fermi

Prendiamo ora in considerazione un semiconduttore. Per gli elettroni liberi vale quanto detto per i metalli: a partire da un valore minimo detto energia di conduzione E_C , si può definire una fascia molto densa, praticamente continua, di stati energetici associati a elettroni mobili (elettroni di conduzione). La differenza rispetto ai metalli è che nei semiconduttori anche gli elettroni dei legami covalenti sono in qualche modo mobili, infatti possono saltare per effetto tunnel da un legame all'altro. Non ci sorprenderà, dunque, se anche per gli elettroni dei legami covalenti, detti appunto elettroni di valenza, si potrà definire una fascia di livelli (banda di valenza) con un valore massimo, detto energia di valenza E_V . La differenza di energia tra E_C e E_V sarà proprio l'energia minima che è necessario fornire a un elettrone per strapparli dal legame covalente, ovvero per passare dalla valenza alla conduzione.

In un semiconduttore intrinseco l'energia di Fermi è esattamente al centro della banda proibita. Supponiamo costante la densità dei livelli in banda di valenza e banda di conduzione (in realtà non lo è, ma non è essenziale ai fini di ciò che stiamo dicendo). Il prodotto densità di stati e probabilità di occupazione ci fornisce la densità di elettroni. Possiamo vedere che in conduzione il numero massimo di elettroni è sul fondo della banda, poi decresce esponenzialmente. Anche in banda di valenza gli elettroni si addensano in basso lasciando degli stati liberi (dei posti vuoti) nella parte superiore della banda: questi posti vuoti sono le lacune. In un semiconduttore intrinseco il numero di elettroni in conduzione è uguale al numero di lacune e ciò è possibile, in questo modello, solo quando l'energia di Fermi è al centro della banda proibita.

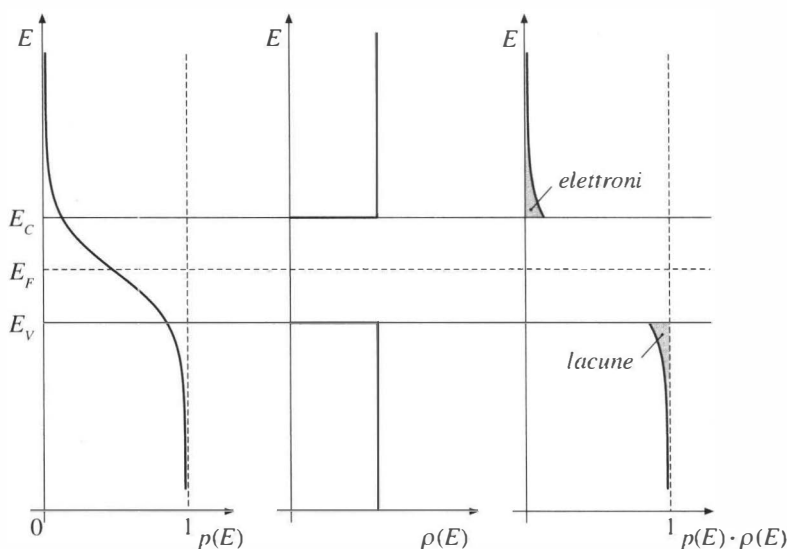


Figura 1.29 Distribuzione degli elettroni nel semiconduttore intrinseco

Un semiconduttore drogato con impurità di tipo n ha un numero di elettroni in conduzione molto maggiore di un semiconduttore intrinseco, ciò significa che l'energia di Fermi è più alta. Riprendendo la metafora del bicchiere: aggiungendo acqua il livello sale.

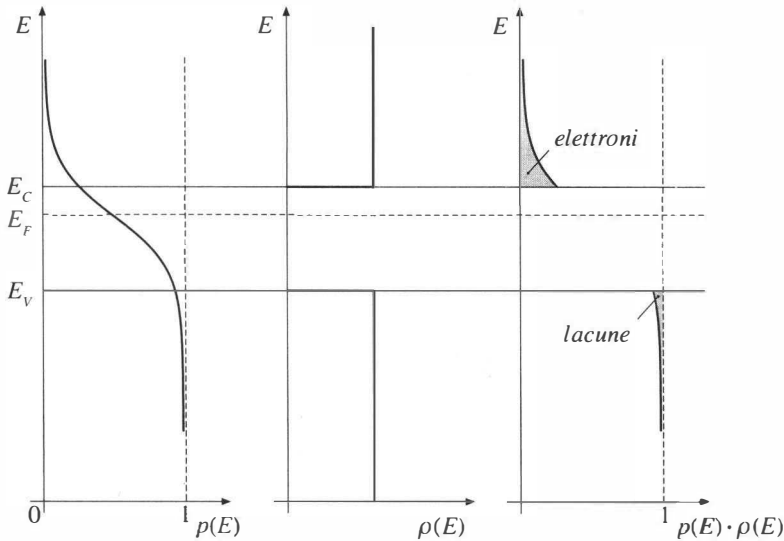


Figura 1.30 Distribuzione degli elettroni nel semiconduttore n

Viceversa drogare un semiconduttore con impurità di tipo p significa ridurre il numero di elettroni (per la legge di azione di massa) e quindi abbassare il livello dell'energia di Fermi.

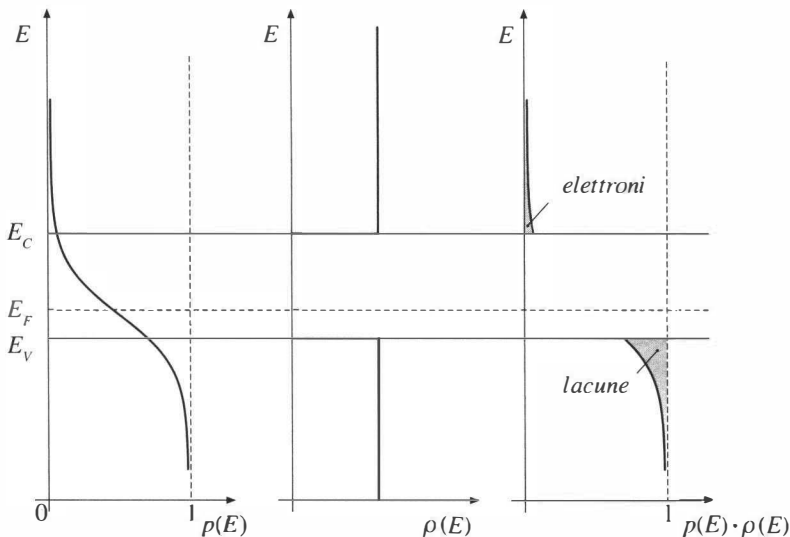


Figura 1.31 Distribuzione degli elettroni nel semiconduttore p

Si noti bene che il livello di Fermi non corrisponde a un effettivo livello fisico come la metafora dell'acqua in un bicchiere potrebbe far pensare. Sarebbe così solo per il caso (non reale) di una distribuzione uniforme degli stati. Nella realtà tale distribuzione forma delle bande ed è tutt'altro che costante anche all'interno di ciascuna banda.

Si noti che la distribuzione di Fermi per $E - E_F \gg kT$ si può scrivere

$$p(E) \approx e^{-\frac{E - E_F}{kT}} \quad (1.69)$$

che è molto simile alla distribuzione di Boltzmann valida per i sistemi classici:

$$p_{\text{Boltz}}(E) \propto e^{-\frac{E}{kT}} \quad (1.70)$$

L'unica differenza è che la distribuzione è spostata in alto di un valore pari all'energia di Fermi. Cioè è una diretta conseguenza del principio di esclusione di Pauli. Si consideri il caso di un gas perfetto trattato classicamente, niente impedisce agli atomi di assumere la stessa energia, e in effetti a $T = 0$ tutti gli atomi sono fermi, ovvero nello stato $E = 0$. Ma il caso degli elettroni in un semiconduttore è differente: non più di 2 elettroni possono stare nello stesso stato. Pertanto è come se gli elettroni si impilassero a coppie fino a raggiungere un livello massimo, che è appunto l'energia di Fermi.

Il numero di elettroni in conduzione per unità di volume è l'integrale del prodotto della densità di stati per la distribuzione di Fermi sopra il livello E_C .

$$n = \int_{E_C}^{\infty} p(E) \rho(E) dE \quad (1.71)$$

Trascurando la dipendenza di ρ dall'energia possiamo scrivere

$$n \propto \int_{E_C}^{\infty} e^{-\frac{E - E_F}{kT}} dE \propto e^{-\frac{E_C - E_F}{kT}} \quad (1.72)$$

La (1.72) mostra come all'aumentare dell'energia di Fermi, aumenti il numero di elettroni in conduzione. Nel caso del silicio intrinseco $E_C - E_V = E_G = 1.12 \text{ eV}$ e $E_C - E_F = E_G/2$, quindi

$$n_i \propto e^{-\frac{E_G}{2kT}} \quad (1.73)$$

1.18 Tecnologia del silicio monocristallino

L'elettronica dei dispositivi a semiconduttore si basa su materiali puri e monocristallini. Vediamo i passi principali del processo di produzione del silicio monocristallino, il semiconduttore di gran lunga più usato. Il silicio si trova in natura sotto forma di quarzite da cui per fusione si ottiene silicio metallurgico al 98%. Mediante reazione chimica con acido cloridrico (HCl) gassoso si produce triclorosilicato liquido (SiHCl_3) che è purificato per distillazione e poi fatto reagire con idrogeno gassoso per tornare al silicio in forma solida elementale. In questo modo si ottiene silicio policristallino di elevata purezza (circa una parte su miliardo). Il passo successivo è la realizzazione del lingotto monocristallino. Il silicio è liquefatto in un crogiolo. Un seme di silicio monocristallino è montato su una punta rotante che viene immersa nel silicio fuso (di solito leggermente drogato con impurità n o p). Estraendo molto lentamente la punta rotante, gli atomi di silicio fuso si fissano al seme cristallino seguendone i piani cristallini fino a realizzare un unico grande cristallo di forma cilindrica tipicamente di diametro 10 cm e lungo 50 cm. Il cilindro viene tagliato in fette di spessore 0.5 mm (dette *wafer*) che costituiranno la base delle lavorazioni successive. Si noti che questo spessore è determinato solo dalla necessità di mantenere una certa rigidità meccanica, lo strato nel quale vengono realizzati i circuiti elettronici ha spessore di pochi μm .

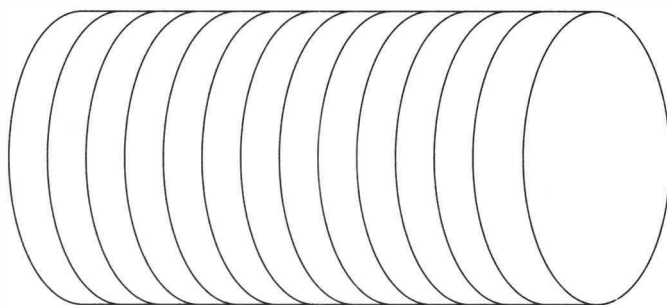


Figura 1.32 Lingotto monocristallino a fette

Sulla superficie del wafer si possono realizzare delle stratificazioni di drogaggi diversi mediante un processo detto **crescita epitassiale**, che consiste nella deposizione di atomi di silicio sul wafer a partire da un composto gassoso (o liquido). Tale processo avviene in un forno a circa 1000°C , in cui è possibile controllare l'erogazione di gas (o liquidi). Questi reagiscono liberando atomi di silicio che si depositano sulla superficie del wafer seguendone i piani cristallini. Durante la crescita epitassiale si possono aggiungere in modo controllato impurità n o p.

1.19 Storia di un materiale dalle proprietà insolite e poco riproducibili

Nella seconda metà del 1700, Alessandro Volta mette le basi di quella che sarà la scienza dell'elettricità: definisce la tensione (che non a caso si misura in Volt), la capacità elettrica, inventa la pila e propone (senza realizzarlo) il generatore elettrico. In un suo scritto del 1782 usa per la prima volta la parola "semiconduttore". Volta osserva che toccando un elettrometro con un metallo questo si scarica istantaneamente; toccandolo con altri materiali, detti isolanti, non si scarica e con altri ancora, che lui definisce semiconduttori, l'elettrometro si scarica molto lentamente. Un'interessante osservazione ma, poiché non fornisce ulteriori dettagli, non possiamo affermare con certezza che Volta sia stato lo scopritore dei semiconduttori.

Il vero scopritore dei semiconduttori fu Michael Faraday nel 1833 (ben cinquant'anni dopo). All'epoca si sapeva che la conducibilità dei metalli (ottimi conduttori) diminuiva all'aumentare della temperatura. Faraday scoprì un'intera classe di composti (tra cui il solfuro di argento, Ag_2S) che, pur conducendo come i metalli, avevano conducibilità che aumentava con la temperatura. Ci sarebbe voluto un secolo per capirne la fisica. La comprensione della conduzione nei semiconduttori necessita, infatti, della meccanica quantistica. Solo nel 1931 l'inglese Alan Watson, durante un periodo di studio a Lipsia presso il gruppo del premio Nobel Werner Heisenberg, elaborò un modello soddisfacente. Watson comprese perfettamente il meccanismo della conduzione nei semiconduttori, ma fu Werner Heisenberg ad avere il coraggio di parlare di lacuna come una sorta di fantomatica particella positiva.

Tuttavia, la strada verso lo sfruttamento industriale dei semiconduttori sarebbe stata ancora molto lunga. Il principale problema con questi materiali era la riproducibilità delle loro caratteristiche. Nello stesso anno in cui Watson elabora la teoria della conduzione nei semiconduttori, il Premio Nobel Wolfgang Pauli, sempre molto caustico, smorza l'entusiasmo dell'amico Rudolf Peierls scrivendogli: "Sui semiconduttori proprio non si può lavorare, è un pasticcio, chissà poi se esistono veramente!".

1.20 Riepilogo

- La legge di Ohm esprime una relazione lineare tra tensione (V) e corrente (I)

$$V = RI$$

con R resistenza.

- La resistenza di un materiale di lunghezza L e sezione S è dato da

$$R = \frac{1}{\sigma} \frac{L}{S}$$

con σ conducibilità del materiale.

- La conducibilità è una proprietà del materiale ed è data da

$$\sigma = qn\mu$$

con q carica del portatore, n concentrazione dei portatori, μ mobilità dei portatori.

- I materiali dell'elettronica si distinguono in conduttori, isolanti e semiconduttori.
- I conduttori hanno alta conducibilità che diminuisce all'aumentare della temperatura.
- Gli isolanti hanno conducibilità molto bassa che aumenta con la temperatura.
- I semiconduttori hanno conducibilità intermedia tra conduttori e isolanti, che aumenta all'aumentare della temperatura.
- Nei semiconduttori la corrente è portata da due tipi di portatori di carica: gli elettroni (particelle di carica negativa pari a $q = 1.6 \times 10^{-19}$ C) e le lacune (particelle "virtuali" positive di carica pari a q).
- Il semiconduttore più comune è il silicio.
- Nel silicio intrinseco (ovvero senza impurità) il numero di elettroni per unità di volume n è pari al numero di lacune per unità di volume p

$$n = p = n_i \approx 10^{10} \text{ cm}^{-3}$$

- Nel silicio drogato con impurità di tipo n

$$n \approx N_D$$

$$p \approx \frac{n_i^2}{N_D}$$

con N_D concentrazione delle impurità.

- Nel silicio drogato con impurità di tipo p

$$p \approx N_A$$

$$n \approx \frac{n_i^2}{N_A}$$

con N_A concentrazione delle impurità.

- La densità di corrente j in un semiconduttore è, in generale, data da

$$j = qn\mu_n E + qp\mu_p E + qD_n \frac{dn}{dx} - qD_p \frac{dp}{dx}$$

con

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{kT}{q} = V_T \approx 26\text{mV}$$

- In un semiconduttore all'equilibrio

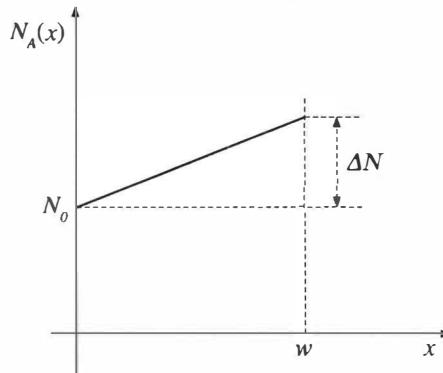
$$p(x) = p(0)e^{-\frac{x}{V_T}}$$

$$n(x) = n(0)e^{\frac{x}{V_T}}$$

1.21 Esercizi svolti

*1) Calcolare la distribuzione di campo elettrico e la distribuzione delle cariche non compensate in un semiconduttore p di spessore w all'equilibrio con distribuzione di impurità data da

$$N_A(x) = N_0 + \frac{\Delta N}{w}x$$



All'equilibrio

$$p(x) \approx N_A(x)$$

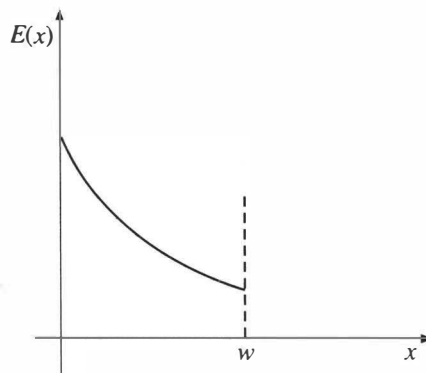
Sostituendo nella relazione

$$pE(x) = V_T \frac{dp}{dx}$$

si ottiene

$$E(x) = \frac{V_T \Delta N}{w} \frac{1}{N_0 + \frac{\Delta N}{w}x}$$

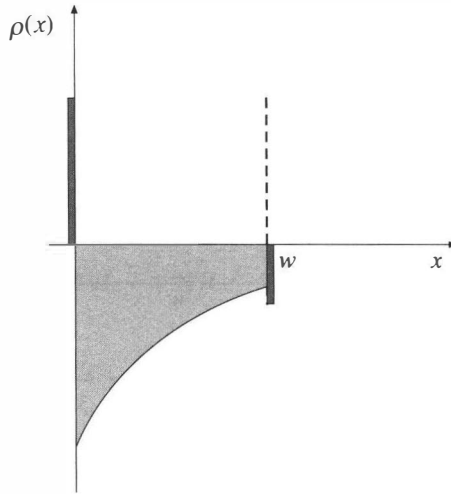
il cui andamento è mostrato in figura



Per calcolare la distribuzione di carica non compensata si usa la relazione di Poisson

$$\rho(x) = \varepsilon \frac{dE}{dx} = -\varepsilon V_T \frac{1}{\left(\frac{N_0}{\Delta N} w + x \right)^2}$$

il cui andamento è mostrato in figura



Tale distribuzione deve essere compensata da un'adeguata concentrazione di cariche superficiali. Per il teorema di Gauss, tali cariche sono pari al campo elettrico superficiale per la costante dielettrica quindi

$$\rho'(0) = \varepsilon \frac{V_T \Delta N}{w} \frac{1}{N_0}$$

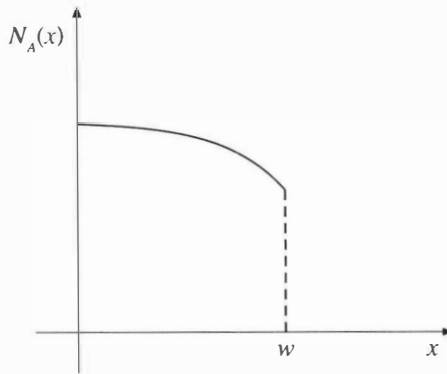
$$\rho'(w) = -\varepsilon \frac{V_T \Delta N}{w} \frac{1}{N_0 + \Delta N}$$

Si noti che ρ' essendo una carica superficiale è dimensionalmente C/m^2 , mentre ρ è una carica di volume quindi C/m^3 .

***2) Calcolare la distribuzione di campo elettrico e la distribuzione delle cariche non compensate in un semiconduttore p di spessore w all'equilibrio con distribuzione di impurità data da**

$$N_A(x) = N_0 \cos\left(\frac{x}{L}\right)$$

con $0 < x < w$ e $w < \pi L/4$



All'equilibrio

$$p(x) \approx N_A(x)$$

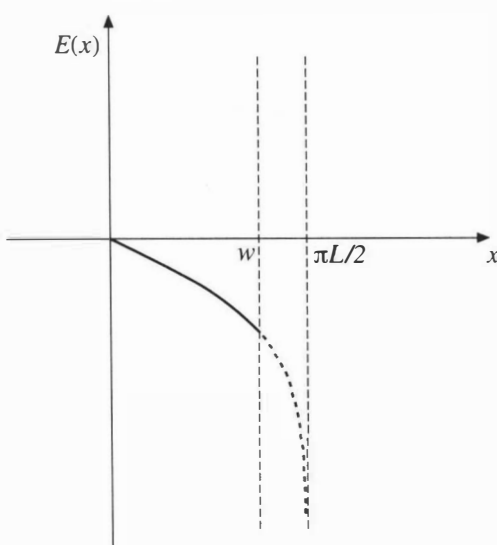
Sostituendo nella relazione

$$pE(x) = V_\tau \frac{dp}{dx}$$

si ottiene

$$E(x) = -\frac{V_\tau}{L} \tan\left(\frac{x}{L}\right)$$

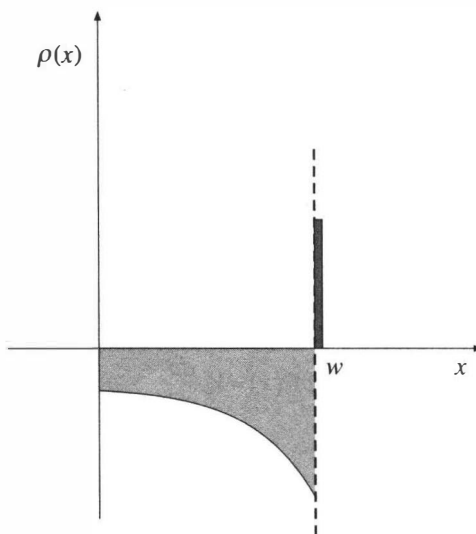
il cui andamento è mostrato in figura:



Per calcolare la distribuzione di carica non compensata si usa la relazione di Poisson

$$\rho(x) = \epsilon \frac{dE}{dx} = -\epsilon \frac{V_T}{L^2} \frac{1}{\cos^2\left(\frac{x}{L}\right)}$$

il cui andamento è mostrato in figura:



Tale distribuzione deve essere compensata da un'adeguata concentrazione di cariche superficiali. Per il teorema di Gauss, tali cariche sono pari al campo elettrico superficiale per la costante dielettrica, quindi

$$\begin{aligned} \rho'(0) &= 0 \\ \rho'(w) &= \epsilon \frac{V_T}{L} \tan\left(\frac{w}{L}\right) \end{aligned}$$

Si noti che ρ' essendo una carica superficiale è dimensionalmente C/m^2 , mentre ρ è una carica di volume, quindi C/m^3 .

1.22 Domande di verifica a scelta multipla

1. Negli isolanti a temperatura diversa da zero
 - a) ci sono cariche positive mobili
 - b) non ci sono cariche mobili
 - c) gli elettroni si possono muovere solo per effetto tunnel
 - d) ci sono elettroni mobili
2. La conducibilità degli isolanti all'aumentare della temperatura
 - a) aumenta perché aumentano gli elettroni liberi
 - b) diminuisce perché aumentano gli elettroni liberi
 - c) diminuisce perché diminuisce il tempo medio tra un urto e il successivo
 - d) diminuisce perché aumenta il tempo medio tra un urto e il successivo
3. La conducibilità dei metalli all'aumentare della temperatura
 - a) aumenta leggermente perché aumentano gli elettroni liberi
 - b) diminuisce perché diminuisce il tempo medio tra un urto e il successivo
 - c) diminuisce perché aumenta il tempo medio tra un urto e il successivo
 - d) aumenta leggermente perché diminuiscono gli elettroni liberi
4. Una sola delle seguenti affermazioni è vera.
 - a) Gli elettroni in un conduttore si possono muovere solo nella direzione del campo elettrico
 - b) Gli elettroni in un conduttore soggetto a un campo elettrico si muovono in modo del tutto casuale
 - c) Gli elettroni in un conduttore si muovono per effetto tunnel.
 - d) Gli elettroni in un conduttore non soggetto a campi elettrici si muovono in modo del tutto casuale
5. Una sola delle seguenti affermazioni è vera.
 - a) Nei conduttori ci sono cariche positive mobili
 - b) Nei conduttori ci sono sia elettroni sia cariche positive mobili
 - c) Nei conduttori ci sono solo elettroni mobili
 - d) Nei conduttori ci sono cariche positive che si muovono per effetto tunnel
6. La velocità di deriva è proporzionale al campo elettrico applicato
 - a) perché gli urti con gli atomi fanno dissipare l'energia cinetica acquisita dalle cariche
 - b) perché in generale la velocità è proporzionale alla forza applicata
 - c) perché nei casi in cui è necessario applicare la meccanica quantistica la velocità risulta proporzionale alla forza
 - d) perché la velocità di deriva è generata dalla diffusione dei portatori che è proporzionale al campo elettrico
7. In un conduttore
 - a) la velocità di deriva degli elettroni è preponderante rispetto alla velocità media indotta dall'agitazione termica
 - b) la velocità di deriva aumenta all'aumentare del tempo di applicazione della tensione
 - c) la velocità di deriva dipende dalla temperatura
 - d) la velocità di deriva è inversamente proporzionale alla mobilità

8. Nel silicio intrinseco
 - a) la concentrazione degli elettroni liberi è molto maggiore di quella degli atomi, in quanto un singolo atomo può liberare fino a 4 elettroni
 - b) la concentrazione delle lacune è sempre uguale a quella degli atomi, in quanto un atomo può generare una sola lacuna
 - c) la concentrazione degli elettroni liberi è molto minore della concentrazione degli atomi
 - d) la concentrazione degli elettroni liberi è molto maggiore di quella delle lacune
9. Una sola delle seguenti affermazioni è vera.
 - a) Il numero di impurità in un semiconduttore drogato deve essere circa uguale al numero di lacune nello stesso semiconduttore prima del drogaggio
 - b) Il numero di impurità di tipo n deve essere sempre molto maggiore delle eventuali impurità di tipo p già presenti nel semiconduttore
 - c) Il numero di impurità in un semiconduttore deve essere molto minore del numero di atomi
 - d) Il numero di impurità di tipo p in un semiconduttore deve essere molto minore della differenza tra il numero di impurità di tipo n e il numero di atomi
10. Un elettrone può passare per effetto tunnel da una buca di potenziale all'altra
 - a) quando, per agitazione termica, ha sufficiente energia per saltare fuori dalla buca
 - b) quando le buche di potenziale sono sufficientemente vicine e poco profonde
 - c) quando le buche di potenziale sono sufficientemente distanziate in modo che le rispettive funzioni d'onda non possano interferire tra loro
 - d) quando la lunghezza d'onda associata all'elettrone è molto più piccola della larghezza della buca di potenziale
11. Una corrente che scorre in un semiconduttore
 - a) non produce dissipazione termica, in quanto la velocità di deriva è molto minore della velocità media dovuta all'agitazione termica
 - b) produce una dissipazione termica proporzionale al quadrato della corrente
 - c) produce una dissipazione termica inversamente proporzionale alla tensione applicata
 - d) a parità di corrente, produce una dissipazione termica proporzionale alla conducibilità
12. Una sola delle seguenti affermazioni è vera.
 - a) Le lacune sono ioni positivi mobili
 - b) Le lacune si muovono casualmente quando non è applicato un campo elettrico
 - c) Le lacune sono protoni liberati dagli atomi
 - d) Le lacune sono elettroni che liberati da un'impurità rimangono per un breve periodo in conduzione prima di essere catturati da un'altra impurità
13. Un semiconduttore drogato con impurità di tipo p
 - a) ha un tasso di generazione di coppie elettrone-lacuna maggiore di un semiconduttore intrinseco
 - b) ha un numero minore di elettroni liberi rispetto a un semiconduttore intrinseco
 - c) ha un tasso di generazione di coppie elettrone-lacuna minore di un semiconduttore intrinseco
 - d) ha un tasso di ricombinazione maggiore di un semiconduttore intrinseco

14. All'aumentare del numero di portatori n
- il tasso di ricombinazione rimane invariato
 - il tasso di ricombinazione diminuisce
 - il tasso di generazione diminuisce
 - il tasso di generazione aumenta
15. Un semiconduttore drogato con impurità di tipo n
- ha circa la stessa conducibilità dello stesso semiconduttore intrinseco
 - ha conducibilità minore dello stesso semiconduttore intrinseco
 - ha conducibilità maggiore dello stesso semiconduttore intrinseco
 - dipende dalla concentrazione delle impurità, può avere conducibilità maggiore o minore dello stesso semiconduttore intrinseco
16. Una sola delle seguenti affermazioni è vera.
- Una corrente di lacune in un senso è sempre accompagnata da una corrente di elettroni nell'altro senso di pari entità
 - Quando lacune ed elettroni si muovono in senso opposto, entrambi contribuiscono ad aumentare la corrente risultante che quindi è maggiore rispetto al caso in cui sia presente solo corrente di lacune o solo corrente di elettroni
 - Quando lacune ed elettroni si muovono in senso opposto la corrente risultante è la differenza tra le due correnti ed è quindi minore rispetto al caso in cui sia presente solo corrente di lacune o solo corrente di elettroni
 - La somma algebrica delle correnti di lacune e di elettroni è sempre zero
17. Una sola delle seguenti affermazioni è vera.
- La somma algebrica di corrente di diffusione e corrente di deriva in un semiconduttore è sempre zero
 - Correnti di diffusione non possono sussistere quando ci sono tensioni applicate
 - All'aumentare della temperatura e a parità di altre condizioni la corrente di diffusione aumenta
 - La somma algebrica di corrente di diffusione e corrente di deriva in un semiconduttore è zero, quando sono presenti sia elettroni sia lacune
18. Il movimento delle lacune
- è dovuto alla presenza di elettroni liberi che tendono a riempire le buche di potenziale vuote
 - è dovuto all'effetto tunnel tra una buca di potenziale e l'altra
 - è dovuto all'effetto Zener che produce il passaggio di carica per ionizzazione degli atomi di silicio
 - è dovuto all'effetto valanga che strappa elettroni dalle buche di potenziale, producendo un movimento di cariche positive in senso opposto
19. Una sola delle seguenti affermazioni è vera.
- La corrente di diffusione è generata dal campo elettrico prodotto dal movimento dei portatori
 - La corrente di diffusione è maggiore dove la concentrazione dei portatori è maggiore
 - La corrente di diffusione aumenta all'aumentare della temperatura
 - La corrente di diffusione è sempre di segno opposto rispetto alla corrente di deriva

20. Una sola delle seguenti affermazioni è vera.
- a) L'energia di Fermi è una costante del materiale semiconduttore e non dipende dalla presenza di impurità
 - b) L'energia di Fermi è il valor medio tra l'energia di valenza e l'energia di conduzione
 - c) L'energia di Fermi aumenta all'aumentare della concentrazione di impurità di tipo n
 - d) L'energia di Fermi diminuisce all'aumentare della temperatura
21. Una sola delle seguenti affermazioni è vera.
- a) L'energia di Fermi è il valore medio tra l'energia di valenza e l'energia di conduzione
 - b) L'energia di Fermi aumenta con la temperatura
 - c) L'energia di Fermi diminuisce con la temperatura
 - d) L'energia di Fermi ha lo stesso valore per materiali in contatto elettrico
22. In un semiconduttore di tipo p, l'energia di Fermi
- a) è circa uguale all'energia di gap
 - b) è trascurabile
 - c) coincide con il potenziale di estrazione
 - d) è circa uguale all'energia di valenza
23. L'energia di Fermi in un semiconduttore
- a) aumenta con la temperatura, in quanto gli elettroni vanno a occupare stati a maggior energia
 - b) aumenta drogando il materiale con impurità di tipo n, in quanto aumentano gli elettroni liberi
 - c) aumenta drogando il materiale con impurità di tipo p, in quanto aumentano le lacune
 - d) è una costante ed è pari al valor medio tra energia di conduzione ed energia di valenza
24. In un metallo
- a) la corrente scorre nello stesso senso del movimento dei portatori
 - b) non può scorrere una corrente altrimenti la tensione ai suoi capi diverrebbe infinita
 - c) la corrente è proporzionale alla tensione applicata
 - d) la corrente scorre per effetto tunnel tra le buche di potenziale determinate dagli atomi che mettono in compartecipazione l'elettrone più esterno
25. In un metallo la densità di corrente
- a) è proporzionale alla conducibilità
 - b) non è una costante del materiale in quanto dipende dal numero di elettroni della corrente che scorre
 - c) si esprime dimensionalmente in Ampere
 - d) è inversamente proporzionale al numero di elettroni nell'unità di volume
26. Il potenziale termico
- a) dipende dal materiale, i metalli hanno tipicamente potenziali termici maggiori degli isolanti
 - b) è il potenziale ai capi di una giunzione pn polarizzata direttamente
 - c) è circa uguale alla tensione di soglia nei MOSFET a canale n
 - d) è un valore che non varia per materiali diversi alla stessa temperatura

27. Sono semiconduttori
- tutti gli elementi in quarta colonna
 - solo gli elementi in quarta colonna
 - combinazioni di elementi della quarta e quinta colonna
 - alcuni elementi della quarta colonna
28. Solo una delle seguenti affermazioni è corretta.
- un elettrone quando è in quiete ha lunghezza d'onda nulla
 - la lunghezza d'onda è una proprietà dei campi elettromagnetici, non delle particelle come gli elettroni
 - all'aumentare della sua velocità, un elettrone ha lunghezza d'onda minore
 - un elettrone ha una lunghezza d'onda propria che non dipende dalla sua velocità
29. Solo una delle seguenti affermazioni è corretta.
- All'aumentare della temperatura la mobilità degli elettroni diminuisce nei metalli e aumenta negli isolanti
 - L'integrale della velocità di deriva lungo la direzione della corrente è proporzionale alla differenza di tensione
 - Conducibilità e mobilità sono sinonimi
 - La conducibilità è una caratteristica intrinseca del materiale, pertanto non dipende dalla temperatura
30. In un semiconduttore
- il numero di elettroni liberi è pari al numero di donatori e il numero di lacune è pari al numero di accettori
 - il numero di elettroni liberi è sempre uguale a $n_i = 10^{10} \text{ cm}^{-3}$, perché la concentrazione dei droganti è sempre molto minore rispetto alla densità degli atomi
 - se sono presenti donatori in maggior numero degli accettori, la concentrazione degli elettroni liberi è pari alla differenza delle loro concentrazioni
 - il numero delle lacune non può essere maggiore del numero degli elettroni liberi, perché ogni elettrone che si libera lascia nel reticolo una lacuna
31. Nel silicio intrinseco
- gli elettroni liberi sono 4 volte il numero di atomi, in quanto un atomo di silicio ha 4 elettroni nella corteccia più esterna
 - il numero di lacune è circa pari al numero di atomi, perché ogni atomo di silicio può intrappolare un solo elettrone libero
 - c'è un elettrone libero ogni mille miliardi di atomi
 - non ci sono elettroni liberi
32. Nel silicio intrinseco, quando è applicata una tensione
- scorre una corrente di soli elettroni che aumenta esponenzialmente con la tensione
 - scorre una corrente di elettroni e di lacune che aumenta esponenzialmente con la tensione
 - scorre una corrente di coppie elettrone-lacuna costante al variare della tensione, se la sezione è costante
 - scorre una corrente proporzionale alla tensione applicata

33. La conducibilità di una piastrina rettangolare di silicio
 - a) aumenta all'aumentare della sezione della piastrina trasversale alla corrente
 - b) aumenta all'aumentare della lunghezza della piastrina lungo la direzione della corrente
 - c) aumenta con la temperatura
 - d) diminuisce quando il silicio è illuminato da luce nello spettro visibile
34. Solo una delle seguenti affermazioni è vera.
 - a) il silicio è l'unico materiale semiconduttore
 - b) sono semiconduttori tutti gli elementi in IV colonna
 - c) esistono semiconduttori costituiti da quattro elementi diversi
 - d) il gallio è un semiconduttore
35. Il rapporto tra coefficiente di diffusione e mobilità
 - a) nei metalli è molto più grande che nei semiconduttori
 - b) è unitario
 - c) aumenta all'aumentare della temperatura
 - d) è una costante universale

1.23 Domande di verifica a risposta aperta

1. Dimostrare la legge di Ohm.
2. Spiegare perché un elettrone a temperatura ambiente non può essere confinato in una regione (ad esempio una buca di potenziale) più piccola di circa 5nm.
3. Dimostrare la legge di azione di massa.
4. Scrivere l'espressione completa della densità di corrente in un semiconduttore spiegando il significato fisico di ciascun termine.
5. Spiegare il modello a bande del semiconduttore con particolare riferimento al significato dell'energia di Fermi.

1.24 Risposte alle domande a scelta multipla

2 a	14 a	26 d
3 b	15 c	27 d
4 d	16 b	28 c
5 c	17 c	29 b
6 a	18 b	30 c
7 c	19 c	31 c
8 c	20 c	32 d
9 c	21 d	33 c
10 b	22 d	34 c
11 b	23 b	35 c
12 b	24 c	

La giunzione pn

Il capitolo inizia con un breve excursus storico sugli sviluppi tecnologici che hanno portato al diodo a stato solido. La giunzione pn è introdotta a partire dalla sua caratteristica non lineare e dal suo impiego circuitale. Si discute, quindi, la formazione della regione di svuotamento e la sua interpretazione nell'ambito del modello a bande. Prima di passare alla descrizione del diodo completo, si trattano brevemente le giunzioni metallo-semiconduttore. Infine, il principio di funzionamento del diodo è spiegato prima intuitivamente e poi matematicamente derivando esplicitamente la caratteristica del diodo dal processo di iniezione e ricombinazione dei portatori minoritari.

2.1 Dalla galena alla giunzione pn

Nel 1874 Carl Ferdinand Braun (premio Nobel nel 1909 con Guglielmo Marconi) scoprì che per alcune giunzioni tra un metallo e un altro materiale (in particolare la galena, PbS) la curva tensione-corrente non era lineare (ohmica) come per la maggior parte dei materiali, ma aveva uno strano comportamento: la corrente sembrava scorrere in una sola direzione, invertendo la tensione la corrente quasi si annullava. Questi contatti vennero chiamati “rettificanti” in quanto una tensione alternata positiva e negativa veniva raddrizzata, cioè diveniva solo positiva passando attraverso il contatto (Figura 2.1).

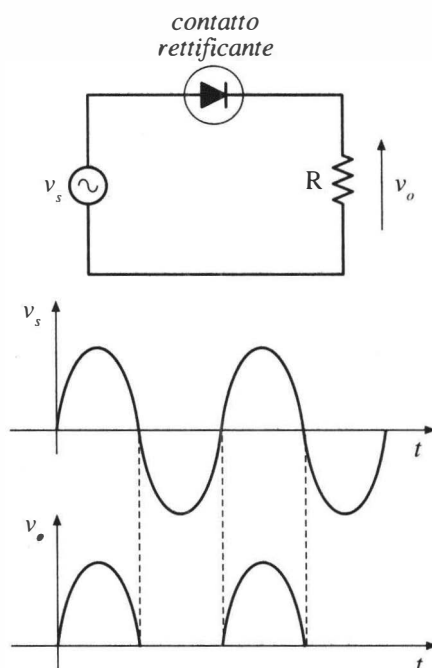


Figura 2.1 Contatto rettificante

Un'altra stranezza di questo dispositivo era che se i due materiali erano premuti fortemente l'uno contro l'altro il contatto diveniva ohmico, ma se erano premuti troppo debolmente la resistenza in conduzione diveniva troppo alta. Vennero pertanto sviluppati dei dispositivi molto delicati, detti *catwhisker*, in cui un sottile filo elettrico (come un baffo di gatto) era delicatamente premuto su un altro materiale.

I contatti rettificanti avrebbero potuto essere solo una curiosità scientifica, ma i primi del Novecento furono gli anni del “telegrafo senza fili”. All'epoca l'informazione era codificata come una variazione di ampiezza della portante (AM: *amplitude modulation*), e mediante un dispositivo rettificante era possibile realizzare un demodulatore AM semplice ed efficace (vedi esercizio 11 di questo capi-

to). In effetti, già nel 1904 l'indiano Jagadish Chandra Bose brevettò il primo rivelatore AM a galena. Due anni dopo, l'americano Greenleaf Whittier Pickard brevettò anche il contatto metallo-silicio e per la prima volta tracciò il simbolo del diodo, il cui originario significato era quello di rappresentare il contatto tra un filo e un materiale. La freccia indicava solo il contatto puntiforme, indipendentemente dal verso della corrente. Ma i primi contatti rettificanti avevano il problema della scarsa riproducibilità (a causa della poca purezza dei materiali impiegati) e della criticità del contatto "a baffo di gatto". Per questo motivo, nei primi anni del Novecento i rivelatori più usati rimasero i tubi a vuoto.

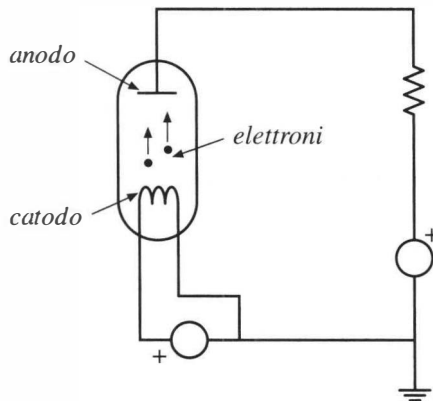


Figura 2.2 Diodo a tubo

In un tubo a vuoto un filamento riscaldato libera elettroni in prossimità del filamento stesso. La differenza di tensione tra il filamento e una placca metallica, detta anodo, produce una corrente di elettroni in direzione di quest'ultima. Per come è realizzato, questo dispositivo può condurre solo in un senso: gli elettroni possono andare dal catodo all'anodo e non viceversa. La presenza di aria tra catodo e anodo costituisce un impedimento al moto degli elettroni, per questo deve essere realizzato il vuoto pneumatico. I tubi a vuoto erano una tecnologia affidabile (basata su quella ben consolidata delle lampade a incandescenza), ma ingombrante e dispendiosa. I contatti rettificanti a stato solido (ovvero non a tubo) sembravano poter offrire una vantaggiosa alternativa. Un decisivo passo avanti fu compiuto nel 1939 presso i laboratori della Bell, la compagnia telefonica statunitense. Russell Ohl scoprì la "barriera pn", ovvero scoprì che nei semiconduttori tra una zona drogata n e una zona drogata p si crea una sorta di contatto rettificante. Tuttavia la cosiddetta "legge del diodo" (che esprime la corrente in funzione della tensione) per una giunzione pn fu derivata da William Shockley, un altro scienziato della Bell, lo stesso che in seguito avrebbe vinto il premio Nobel come coinventore del transistor bipolare.

Oggi l'importanza della giunzione pn va ben oltre il suo impiego come contatto rettificante. In qualunque dispositivo a semiconduttore ci sono interfacce tra

zone n e p, pertanto ci sono giunzioni pn: il transistor bipolare può essere visto come due giunzioni pn con una regione in comune; il principio di funzionamento del transistor JFET si basa sulla contropolarizzazione di una giunzione pn; nel transistor MOSFET ci sono due giunzioni pn in corrispondenza del drain e del source. Praticamente tutti i dispositivi microelettronici contengono giunzioni pn.

2.2 Realizzazione di una giunzione pn

Un semiconduttore è detto di tipo n o p se nel reticolo cristallino sono presenti impurità di tipo n o p. Prendiamo, ad esempio, un semiconduttore di tipo p. Tale semiconduttore può essere realizzato direttamente facendo solidificare una soluzione contenente silicio e impurità p. Un lato del solido così ottenuto è successivamente bombardato con un fascio di ioni di impurità di tipo n. Queste penetrano nel semiconduttore fino a una certa profondità e, se sono in concentrazione sufficientemente elevata, prima compensano le impurità p e poi trasformano il semiconduttore da p a n. Il risultato è un materiale con due zone nettamente distinte, una zona n e una p, senza alterazioni della struttura cristallina all'interfaccia tra le due zone, ovvero senza rotture dei piani reticolari o impurità difficilmente controllabili. L'interfaccia tra le due zone è detta giunzione pn. Nel seguito, per semplicità, assumeremo tale interfaccia una superficie netta, anche se nella realtà ci sarà inevitabilmente un certo gradiente nella concentrazione delle impurità per passare dalla zona n alla zona p. In ogni caso tale gradiente non è essenziale per la comprensione del principio di funzionamento dei dispositivi a semiconduttore.

2.3 Il diodo

Il diodo è il dispositivo più semplice che fa uso di una giunzione pn. Le due zone p e n sono munite ciascuna di un contatto elettrico (reoforo) in modo da poter applicare una tensione. Per un motivo che vedremo nel seguito la zona n in contatto con il metallo del reoforo deve essere molto drogata.

Il reoforo in contatto con la zona p è detto anodo (A), l'altro catodo (K). La tensione applicata è positiva se l'anodo è positivo rispetto al catodo.

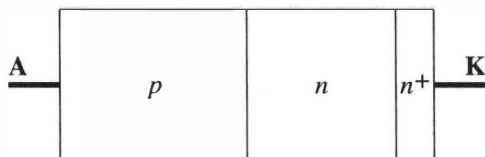


Figura 2.3 Diodo

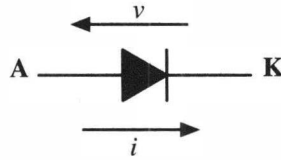


Figura 2.4 Simbolo circuitale del diodo

Applicando una tensione v ai due reofori scorre una corrente che, entro certi limiti che vedremo nel seguito, è data dalla seguente espressione (**caratteristica del diodo ideale**):

$$i = I_s \left(e^{\frac{v}{V_T}} - 1 \right) \quad (2.1)$$

V_T è la tensione termica (26mV), I_s è una costante detta corrente di saturazione inversa. Soffermiamoci un attimo su questa espressione. Per $v = 0$ la corrente è nulla, come è ragionevole aspettarci. Per v minore di zero (contropolarizzazione), la corrente tende rapidamente a $-I_s$, che per questo si chiama corrente di saturazione inversa. I_s è tipicamente un valore molto piccolo, nA o pA, e quindi in molti casi pratici possiamo assumere nulla la corrente che scorre nel diodo quando la tensione è negativa. Per v positiva (polarizzazione diretta), la corrente cresce in modo esponenziale e per $v \gg 2V_T \approx 50\text{mV}$ possiamo assumere

$$i \approx I_s e^{\frac{v}{V_T}} \quad (2.2)$$

L'esponenziale ha un tipico andamento: all'aumentare dell'argomento, assume prima valori molto piccoli per poi crescere repentinamente. Il valore di tensione per cui la corrente comincia ad aumentare molto rapidamente per i diodi al silicio è circa 0.6V e in generale dipende dal tipo di semiconduttore.

Superata la tensione di soglia, il diodo comincia a condurre una corrente significativa. Ci domandiamo ora per quale valore della tensione applicata la corrente che scorre nel diodo diventerà 1000 volte tanto, ovvero scorrerà una corrente molto più grande.

Ebbene, è sufficiente un aumento di tensione 0.18V per produrre un aumento di corrente di un fattore mille. In altre parole, superata la tensione di soglia, in condizioni operative la tensione varia poco (non più di 0.2V). Nel seguito, quando non diversamente specificato, per un diodo al silicio assumeremo 0.6V come tensione di soglia e 0.8V come tensione massima. Si tratta, ovviamente, di assunzioni piuttosto drastiche, ma molto utili per una rapida comprensione del funzionamento di molti circuiti.

2.4 Retta di carico, punto di lavoro e polarizzazione del diodo

Polarizzare un diodo (così come un qualsiasi altro dispositivo elettrico attivo o passivo) significa applicare ad esso una tensione o una corrente costanti in modo da mantenerlo in un particolare punto della sua caratteristica tensione-corrente, detto punto di lavoro. Il modo più semplice per polarizzare un diodo è applicare ad esso una tensione mediante un generatore di tensione costante e una resistenza come mostrato in Figura 2.5. Applicando la legge di Ohm a tale resistenza (detta “resistenza di carico”) si ottiene

$$V^+ - v_D = Ri_D \quad (2.3)$$

ovvero l'espressione della cosiddetta “retta di carico”:

$$i_D = \frac{V^+}{R} - \frac{v_D}{R} \quad (2.4)$$

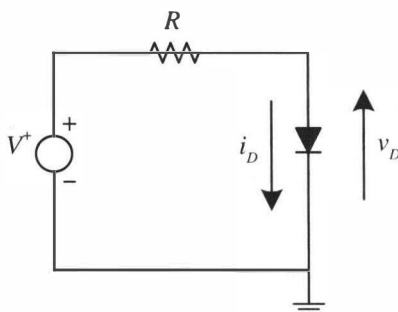


Figura 2.5 Polarizzazione del diodo

La retta di carico esprime i_D in funzione di v_D , così come la caratteristica del diodo, pertanto possono essere entrambi rappresentati in uno stesso grafico come mostrato in Figura 2.6. L'intersezione delle due curve è detto **punto di lavoro** o **punto di polarizzazione** e le sue coordinate esprimono l'effettiva tensione ai capi del diodo e la corrente che vi scorre.

Il concetto di retta di carico è, come vedremo nel seguito, piuttosto importante nello studio dei circuiti non lineari (in particolare i transistor), tuttavia nel caso specifico del diodo spesso la corrente si può stimare più semplicemente supponendo la tensione costante e pari a 0.6V. Con questa ipotesi, la corrente i_D che scorre nel circuito di Figura 2.5 è circa:

$$i_D = \frac{V^+ - (0.6V)}{R} \quad (2.5)$$

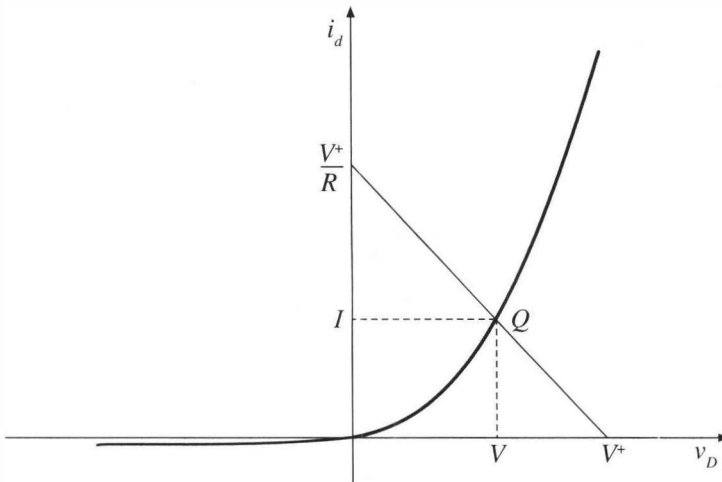


Figura 2.6 Retta di carico e caratteristica del diodo

2.5 Regione di svuotamento

Il funzionamento di un diodo è determinato da ciò che accade all'interfaccia tra la zona n e la zona p.

Nella zona p c'è un'alta concentrazione di lacune ($p \approx N_A$) e una bassa concentrazione di elettroni ($n \approx n_i^2/N_A$). Viceversa nella zona n: $n \approx N_D$, $p \approx n_i^2/N_D$. Le differenze di concentrazione daranno origine a due correnti di diffusione: una di lacune da p a n e una di elettroni da n a p. Gli elettroni migrano dalla zona n lasciando ioni fissi positivi non compensati. Gli stessi elettroni nella zona p si ricombinano con lacune con il risultato di lasciare degli ioni negativi fissi non compensati. Analogamente le lacune che migrano dalla parte p verso la parte n lasciano ioni negativi nella parte p e producono ioni positivi nella parte n. Ioni fissi positivi nella zona n e ioni fissi negativi nella zona p danno origine a un campo elettrico che ha il verso giusto per opporsi alla migrazione delle lacune e degli elettroni. In breve si raggiungerà una situazione di equilibrio tale che il campo elettrico arresti completamente l'ulteriore migrazione di cariche per diffusione.

La distribuzione delle cariche non compensate all'equilibrio dipende dal profilo di drogaggio della zona n e p, ma se la concentrazione dei droganti è costante nelle due zone (**giunzione brusca**) si crea una regione completamente svuotata di cariche mobili a destra (zona p) e a sinistra (zona n) della superficie di separazione tra la zona n e p, e questa regione avrà bordi netti nella zona n e p. Infatti all'equilibrio la somma delle correnti di diffusione e deriva deve essere nulla. Se le concentrazioni di impurità sono costanti non ci potranno essere correnti di diffusione e quindi neppure correnti di deriva. Ma se non ci sono correnti di

deriva in una regione in cui il campo è presente, significa che non ci sono portatori mobili, ovvero la regione tra zona n e p in cui è presente il campo elettrico deve essere completamente svuotata.

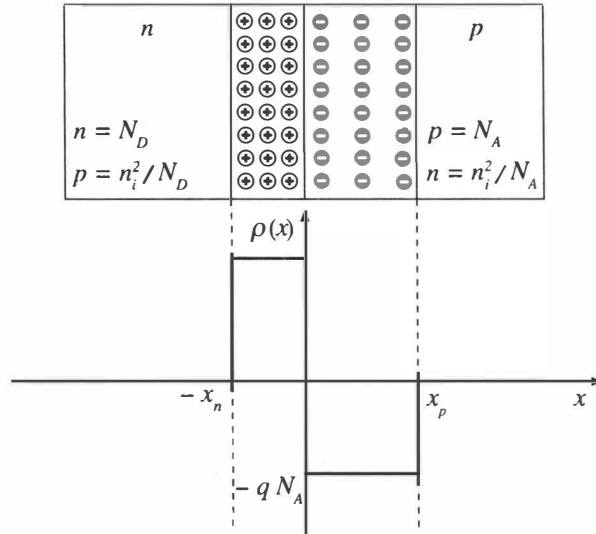


Figura 2.7 Ioni fissi nella regione di svuotamento

La Figura 2.7 mostra la distribuzione delle cariche fisse. L'origine dell'asse x è sulla superficie di separazione tra le zone n e p; $x = -x_n$ e $x = x_p$ sono i bordi della regione di svuotamento. Poiché il semiconduttore era neutro prima dello spostamento delle cariche e nessuna carica si è allontanata dal semiconduttore, la neutralità complessiva dovrà essere conservata, pertanto la carica accumulata nella parte n della regione di svuotamento Q_n deve essere uguale alla carica accumulata nella parte p della regione di svuotamento Q_p :

$$Q_n = Q_p \quad (2.6)$$

da cui:

$$qN_D S x_n = qN_A S x_p \quad (2.7)$$

dove S è la superficie di separazione tra le zone n e p. Da cui

$$N_D x_n = N_A x_p \quad (2.8)$$

Si noti che per la (2.8) la regione di svuotamento risulta più estesa nella regione meno drogata.

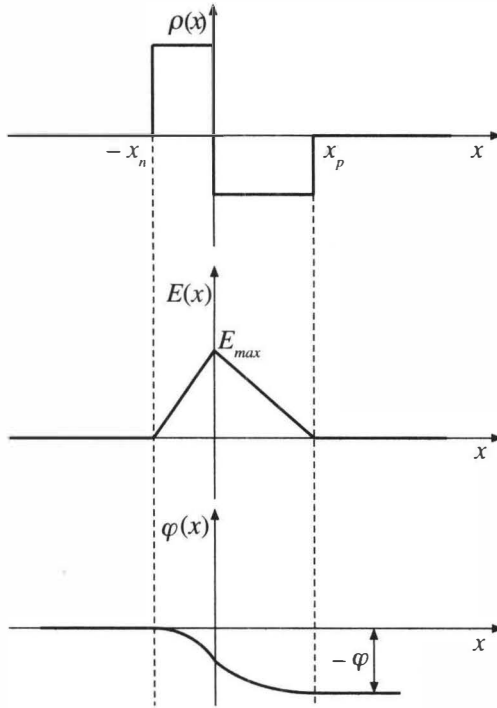


Figura 2.8 Carica, campo elettrico e potenziale nella regione di svuotamento

Campo elettrico e densità di carica (ρ) sono legati dall'equazione di Poisson:

$$\frac{\partial E}{\partial x} = \frac{\rho}{\varepsilon} \quad (2.9)$$

dove ε è la costante dielettrica del materiale (del silicio nello specifico). La (2.9) per il lato n e il lato p della regione di svuotamento sarà rispettivamente

$$\frac{\partial E_n}{\partial x} = \frac{qN_D}{\varepsilon} \quad (2.10)$$

$$\frac{\partial E_p}{\partial x} = -\frac{qN_A}{\varepsilon} \quad (2.11)$$

da cui:

$$E_n(x) = \frac{qN_D}{\varepsilon} x + C_n \quad (2.12)$$

$$E_p(x) = -\frac{qN_A}{\varepsilon} x + C_p \quad (2.13)$$

Le costanti si determinano imponendo che le due funzioni siano nulle rispettivamente per $x = -x_n$ e $x = x_p$, segue quindi che

$$E_n(x) = \frac{qN_D}{\varepsilon}x + \frac{qN_D}{\varepsilon}x_n \quad (2.14)$$

$$E_p(x) = -\frac{qN_A}{\varepsilon}x + \frac{qN_A}{\varepsilon}x_p \quad (2.15)$$

Il campo elettrico nella regione di svuotamento determina una differenza di potenziale tra la zona p e n. Questo potenziale $\varphi(x)$ può essere calcolato come

$$\varphi = \int_{-x_n}^{x_p} E(x) dx \quad (2.16)$$

Si faccia attenzione al segno. In generale il potenziale è l'integrale del campo elettrico cambiato di segno, ma in Figura 2.8 il potenziale è stato definito $-\varphi$ in modo che φ risulti positivo.

$$\varphi = \int_{-x_n}^0 E_n(x) dx + \int_0^{x_p} E_p(x) dx \quad (2.17)$$

Sostituendo le (2.12) e (2.13) nella (2.16) e calcolando gli integrali, si ottiene:

$$\varphi = \frac{q}{2\varepsilon} (N_D x_n^2 + N_A x_p^2) \quad (2.18)$$

Usando la relazione (2.8) che lega x_p e x_n , possiamo esplicitare x_p e x_n nella (2.18) ottenendo

$$x_n = \sqrt{\frac{2\varepsilon}{q} \varphi \frac{N_A}{N_D} \frac{1}{N_A + N_D}} \quad (2.19)$$

$$x_p = \sqrt{\frac{2\varepsilon}{q} \varphi \frac{N_D}{N_A} \frac{1}{N_A + N_D}} \quad (2.20)$$

e quindi:

$$w = x_p + x_n = \sqrt{\frac{2\varepsilon}{q} \varphi \left(\frac{1}{N_A} + \frac{1}{N_D} \right)} \quad (2.21)$$

La (2.21) permette di calcolare lo spessore della regione di svuotamento, noto il potenziale di giunzione.

D'altra parte per un semiconduttore all'equilibrio valgono

$$n(x) = n(0) e^{\frac{v(x)}{V_T}} \quad (2.22)$$

$$p(x) = p(0) e^{-\frac{v(x)}{V_T}} \quad (2.23)$$

Applicando una delle due agli estremi della regione di svuotamento, si ottiene

$$n(-x_n) = n(x_p) e^{\frac{\varphi}{V_T}} \quad (2.24)$$

La concentrazione degli elettroni per $x = -x_n$ è pari a N_D e la concentrazione degli elettroni per $x = x_p$ è n_i^2/N_A , quindi:

$$\varphi = V_T \ln \left(\frac{N_D N_A}{n_i^2} \right) \quad (2.25)$$

Proviamo a usare la (2.25) con valori numerici realistici. Supponiamo ad esempio $N_D = 10^{17} \text{ cm}^{-3}$, $N_A = 10^{14} \text{ cm}^{-3}$. Poiché $V_T = 26 \text{ mV}$, il potenziale di giunzione risulta 0.66 V .

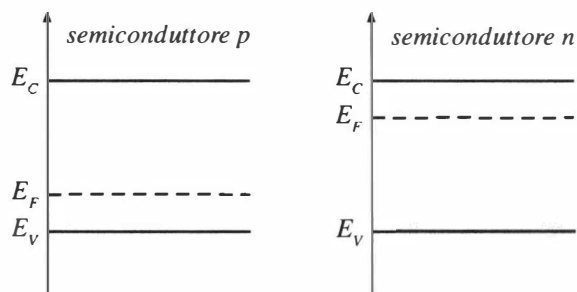
Noto il potenziale di giunzione possiamo usare la (2.21) per calcolare lo spessore della regione di svuotamento. A questo proposito si tenga presente che la costante dielettrica relativa del silicio ϵ_r è circa 12 e che la costante dielettrica del vuoto ϵ_0 è $8.85 \times 10^{-12} \text{ F/m}$.

$$\epsilon = \epsilon_r \epsilon_0 \quad (2.26)$$

Con questi valori lo spessore w della regione di svuotamento è circa $3 \mu\text{m}$. Rispetto alle dimensioni della zona n o p della maggior parte dei diodi, qualche μm è una distanza piccola, quindi in molti casi assumeremo trascurabile lo spessore della regione di svuotamento.

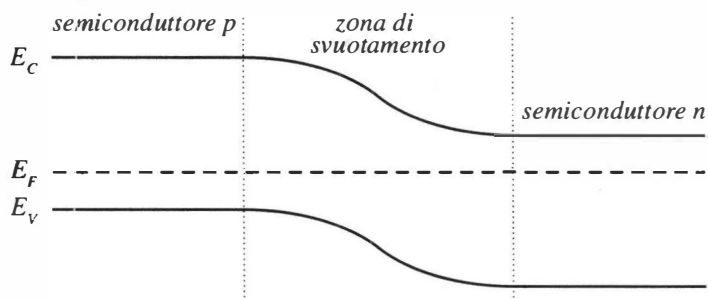
2.6 Modello a bande

Un modo alternativo per visualizzare ciò che accade all'interfaccia tra una zona p e una zona n è fare riferimento al modello a bande. Abbiamo già visto nel Capitolo 1 che l'energia di Fermi assume valori diversi per il silicio drogato n e drogato p.

**Figura 2.9** Modello a bande

In particolare, la presenza di impurità n aumenta il numero di elettroni e quindi alza il livello dell'energia di Fermi, viceversa nel caso di impurità di tipo p. D'altra parte, l'energia di Fermi può essere vista come il livello che raggiunge un liquido in un bicchiere (il liquido sono gli elettroni, l'altezza del bicchiere è l'energia dei livelli). Pertanto quando una zona n è in contatto con una zona p, elettroni tenderanno a fluire da n a p, come l'acqua tra due vasi comunicanti. Anche gli elettroni in banda di valenza tenderanno a fluire nella zona n dalla zona p, ma gli elettroni di valenza che si spostano in un senso possono essere visti come lacune che si spostano nel senso opposto, che quindi migrano da p a n.

Man mano che gli elettroni passano nella zona p e lacune nella zona n, rimarranno cariche fisse positive nella zona n e cariche negative fisse nella zona p. Tali cariche generano un campo elettrico da n a p. D'altra parte il campo elettrico è il gradiente dell'energia dei livelli energetici normalizzata rispetto alla carica elementare. Il significato fisico di quanto detto è che i livelli energetici si incurvano nella regione di svuotamento in modo da uguagliare l'energia di Fermi in tutto il semiconduttore.

**Figura 2.10** Modello a bande della giunzione

2.7 Giunzione metallo-semiconduttore

Il modello a bande della giunzione descritto nel paragrafo precedente offre un'immediata chiave di lettura per comprendere ciò che accade all'interfaccia tra un metallo e un semiconduttore. Dobbiamo distinguere tre casi.

1) Giunzione metallo-semiconduttore *n*

L'energia di Fermi di un metallo è più bassa dell'energia di Fermi del semiconduttore, che è circa uguale all'energia di conduzione. Pertanto se messi in contatto, un po' sorprendentemente, elettroni fluiranno dal semiconduttore al metallo, malgrado il fatto che la densità di elettroni nel metallo sia molto più alta della densità nel semiconduttore.

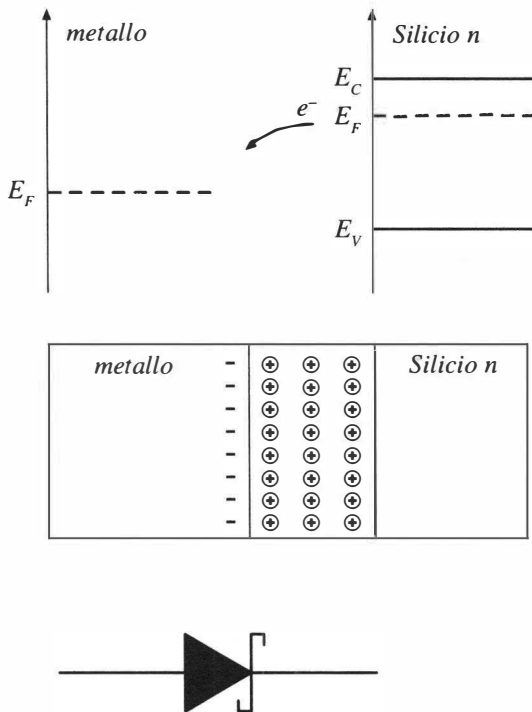


Figura 2.11 Metallo-silicio *n*

Man mano che gli elettroni fluiscono dal semiconduttore, si creerà una regione di svuotamento nel semiconduttore e quindi un potenziale di giunzione. In definitiva il semplice contatto tra metallo e semiconduttore dà luogo a una giunzione rettificante come nel caso pn.

In effetti ci sono diodi, detti Schottky, il cui simbolo è quello mostrato in basso in Figura 2.11, che si basano su questo principio. I diodi Schottky hanno tipicamente tensione di soglia più bassa dei diodi a giunzione pn e altre caratteristiche (minore capacità parassita) che li rendono particolarmente adatti per impieghi ad alta frequenza. Tuttavia la loro realizzazione non è semplice, in quanto impurità e rotture del reticolo sulla superficie rendono le caratteristiche di questo tipo di diodo difficili da controllare e poco riproducibili.

2) *Giunzione metallo-semiconduttore p, energia di Fermi del metallo maggiore dell'energia di valenza*

Prendiamo in considerazione i semiconduttori di tipo p. Se l'energia di Fermi del metallo è maggiore dell'energia di Fermi del semiconduttore (che è prossima all'energia di valenza), gli elettroni vanno dal metallo al semiconduttore compensando le cariche positive delle lacune e generando una zona di svuotamento di cariche negative fisse (ioni) nel semiconduttore in prossimità del metallo. Si crea pertanto una giunzione rettificante, ovvero una sorta di diodo con anodo dalla parte del semiconduttore e catodo dalla parte del metallo.

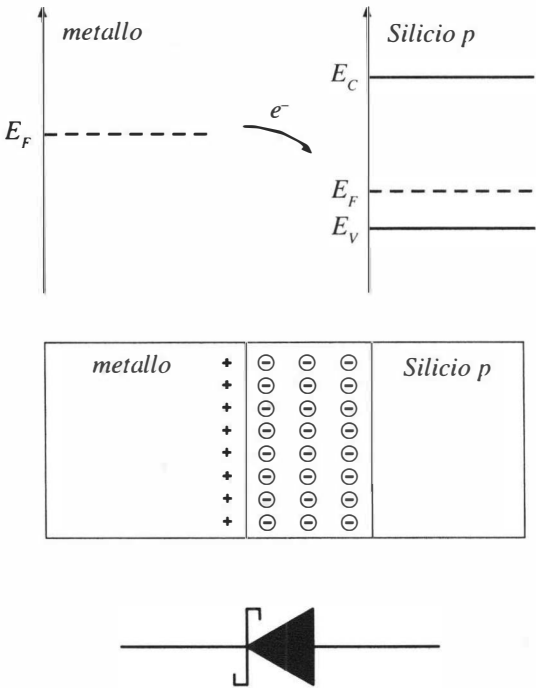


Figura 2.12 Metallo-silicio p, energia di Fermi del metallo superiore all'energia di Fermi del silicio drogato p

3) Giunzione metallo-semiconduttore p, energia di Fermi del metallo minore dell'energia di valenza

Se l'energia di Fermi del metallo è minore dell'energia di Fermi del semiconduttore di tipo p (che è prossima all'energia di valenza) gli elettroni vanno dal semiconduttore al metallo, ma senza produrre una regione di svuotamento nel semiconduttore. Quindi non si crea alcuna giunzione rettificante, ma la corrente può scorrere in entrambi i sensi come in una resistenza.

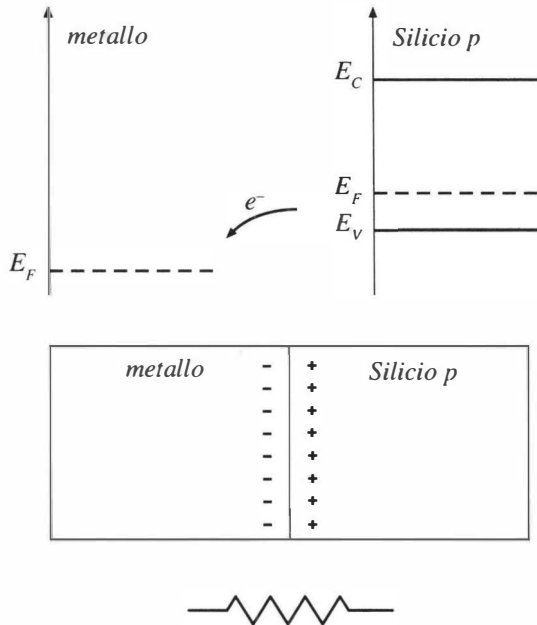


Figura 2.13 Metallo-silicio p, energia di Fermi del metallo inferiore all'energia di Fermi del silicio drogato p

2.8 Il diodo pn completo

Fin qui ci siamo occupati di ciò che succede all'interfaccia tra zona n e p di un diodo pn, ma a ben vedere un diodo con i suoi due refori è in realtà costituito da tre interfacce: metallo-zona n, zona n-zona p, zona p-metallo. È un po' come se dovessimo sempre considerare tre diodi in serie.

A questo proposito dobbiamo distinguere due casi a seconda del tipo di metallo; come mostrato in Figura 2.14 e Figura 2.15. Si noti che in ogni caso la giunzione rettificante tra semiconduttore n e metallo ha verso opposto alla giunzione tra p e n. D'altra parte in una giunzione rettificante può scorrere corrente solo in un verso, segue che due giunzioni contrapposte non consentono il passaggio di

corrente in nessun verso e quindi tutta la teoria fatta fin qui è inutile: il diodo non può funzionare! Per fortuna le cose non stanno esattamente così e la fisica quantistica ci viene in aiuto.

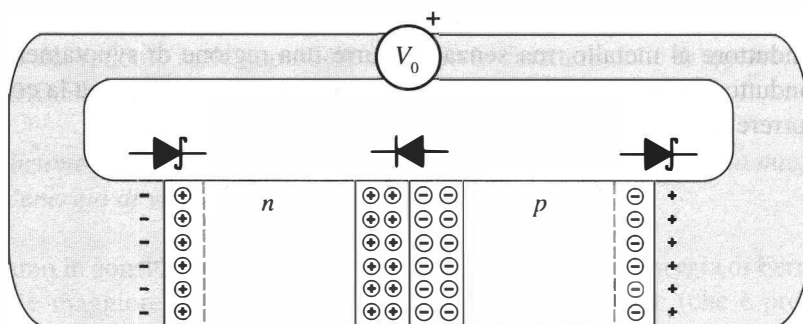


Figura 2.14 Diodo e reofori metallici, energia di Fermi del metallo superiore all'energia di valenza del semiconduttore

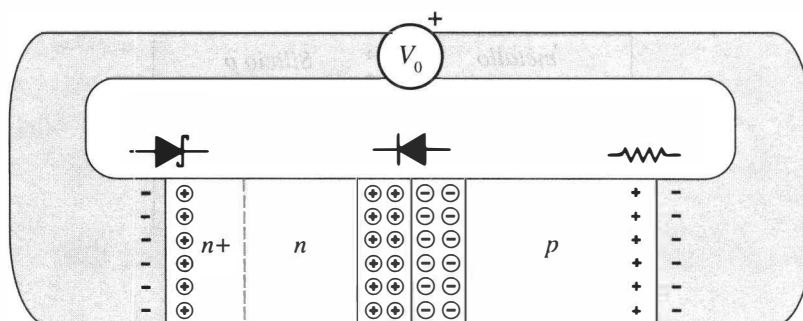


Figura 2.15 Diodo e reofori metallici, energia di Fermi del metallo inferiore all'energia di valenza del semiconduttore

Abbiamo già detto che a un elettrone è associata una lunghezza d'onda che a temperatura ambiente è circa 10nm e che una barriera più sottile della semilunghezza d'onda (5nm) può essere attraversata dall'elettrone mediante una sorta di "materializzazione" al di là della barriera. Quindi per far scorrere la corrente nel circuito contenente il diodo è necessario che il diodo contropolarizzato tra semiconduttore n e metallo abbia una barriera di potenziale sufficientemente sottile da poter essere superata dagli elettroni per effetto tunnel. D'altra parte lo spessore della regione di svuotamento (dove fisicamente è localizzata la barriera di potenziale) diminuisce all'aumentare della concentrazione delle impurità n. Questa è la ragione per cui il semiconduttore in contatto con il catodo in un diodo deve essere fortemente drogato. Ad esempio il silicio drogato n forma con

l'alluminio un contatto ohmico se la concentrazione di impurità è maggiore di 10^{16} cm^{-3} .

Per quanto riguarda la giunzione metallo-semiconduttore di tipo p, il suo comportamento dipende dal tipo di metallo. Se il metallo ha energia di Fermi superiore all'energia di valenza del semiconduttore si forma una regione di svuotamento e quindi una giunzione rettificante nel verso opposto a quello presente tra semiconduttore n e p, in tal caso è necessario realizzare una zona p molto drogata per ridurre la dimensione della zona di svuotamento a un valore sufficientemente piccolo. Se il metallo ha energia di Fermi inferiore all'energia di valenza del semiconduttore non si ha formazione di una regione di svuotamento e la corrente può scorrere liberamente nei due sensi.

Può essere interessante notare che la ragione per cui quando si misura con un multimetro la tensione sui reofori di un diodo a circuito aperto non si misura il potenziale di giunzione, ma una tensione nulla, è proprio il fatto che i potenziali di giunzione sono tre disposti in serie e che in condizioni di equilibrio questi si compensano esattamente.

2.9 Principio di funzionamento del diodo

Nei paragrafi precedenti abbiamo visto come tra le zone n e p di un semiconduttore si creino una regione di svuotamento e un potenziale elettrico che impedisce agli elettroni della zona n di migrare nella zona p e alle lacune della zona p di migrare nella zona n.

Se il polo positivo del generatore di tensione è applicato alla zona p, il verso della tensione tende a diminuire il potenziale sulla giunzione, ovvero a ridurre la barriera tra p e n. Se tale tensione esterna ha un valore tale da annullare la barriera (circa 0.6V) i portatori n potranno liberamente fluire nella zona p e viceversa, dando luogo a una corrente di diffusione. È importante notare che non c'è alcun campo elettrico esterno che spinge le cariche in una giunzione pn in conduzione, la tensione esterna ha il solo compito di annullare la barriera di potenziale: le cariche si muovono per diffusione.

Può essere utile avere un'idea intuitiva del funzionamento della giunzione. Torniamo alla metafora del campo di calcio che abbiamo già usato nel Capitolo 1 per chiarire il concetto di corrente di diffusione. Nella metà sinistra un gran numero di giocatori (gli atomi di silicio) si allenano calciando molti palloni bianchi (elettroni). Nella metà destra un gran numero di calciatori si allenano con alcuni palloni neri (lacune). Nella metà campo una rete impedisce ai palloni bianchi e neri di passare nell'altra metà. Al fischio dell'arbitro la rete è abbassata e, siccome i calciatori calciano casualmente i palloni, i palloni bianchi cominceranno a fluire verso la zona dove non ce ne sono e viceversa. La forza che spinge i palloni non è una tensione esterna (nel nostro esempio potrebbe essere un'e-

ventuale inclinazione del campo) né l'intenzionalità dei calciatori, che non stanno disputando una partita ma solo allenandosi. I palloni migrano semplicemente perché da un lato sono più concentrati che dall'altro.

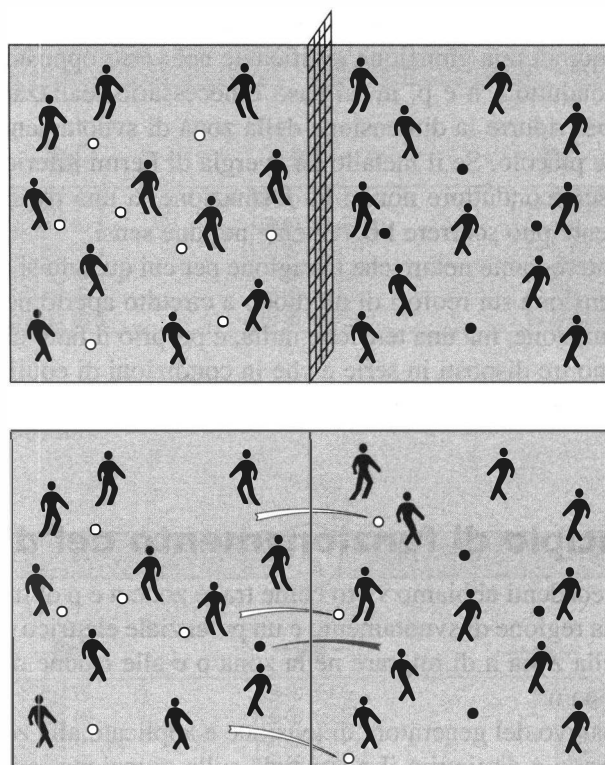


Figura 2.16 Principio di funzionamento del diodo

2.10 Concentrazione dei portatori al limite della regione di svuotamento

Abbiamo già detto che una tensione v_D applicata ai capi di un diodo (positiva dal lato p della giunzione) abbassa il potenziale sulla giunzione. Pertanto i portatori n tenderanno a diffondersi dalla zona n alla zona p e al limite della regione di svuotamento nella zona p ci sarà una concentrazione $n(x_p)$ diversa dalla concentrazione all'equilibrio $n = n_i^2/N_A$.

Analogamente ci sarà una densità di portatori $p(x_n)$ diversa dalla concentrazione di equilibrio $p = n_i^2/N_D$. Vogliamo ora calcolare questa concentrazione dipendente dalla tensione v_D .

Facciamo un'ipotesi a prima vista piuttosto azzardata, ma molto utile: la distribuzione degli elettroni nel semiconduttore in presenza di una corrente è circa

uguale alla distribuzione all'equilibrio. Questo è sicuramente vero per valori di v_D sotto la soglia di conduzione e lo sarà quindi anche per valori di poco superiori.

Se vale questa ipotesi, con riferimento ai portatori n , possiamo scrivere:

$$n(x) = n(0)e^{\frac{v(x)}{V_T}} \quad (2.27)$$

che applicata ai due estremi della regione di svuotamento dà

$$v = V_T \ln \left(\frac{N_D}{n(x_p)} \right) \quad (2.28)$$

Facciamo ora un'ulteriore ipotesi (anche questa piuttosto azzardata): la tensione esterna v_D si applica interamente alla giunzione e va ad abbassare il potenziale di built-in. È evidente che ciò, a rigore, non può essere del tutto vero perché parte della tensione esterna andrà a cadere sui potenziali di contatto tra semiconduttore e metallo. Inoltre quando comincerà a scorrere una corrente ci sarà anche una caduta ohmica nel semiconduttore. Ma se i contatti metallo semiconduttore sono realizzati come abbiamo visto nel Paragrafo 2.8 e se la corrente che scorre non è troppo grande da produrre una significativa caduta di tensione nel semiconduttore, possiamo effettivamente approssimare il potenziale di built-in v come $\varphi - v_D$, con φ potenziale di built-in all'equilibrio. Pertanto

$$\varphi - v_D = V_T \ln \left(\frac{N_D}{n(x_p)} \right) \quad (2.29)$$

Ricordando l'espressione già calcolata per φ (vedi (2.24)), riarrangiando i termini troviamo l'espressione per la concentrazione dei portatori n al limite della regione di svuotamento nella zona p

$$n(x_p) = \left(\frac{n_i^2}{N_A} \right) e^{\frac{v}{V_T}} \quad (2.30)$$

Riflettiamo sul significato dell'equazione trovata. Per $v = 0$, la concentrazione è pari a quella all'equilibrio $n(x_p) = n_i^2 / N_A$. Quando $v > 0$, la concentrazione aumenta esponenzialmente. L'abbassamento del potenziale di giunzione produce una sorta di iniezione di elettroni dalla zona n alla zona p . Ed è proprio questo eccesso di elettroni, rispetto al valore di equilibrio nel resto del semiconduttore p , a produrre la corrente di diffusione alla base del principio di funzionamento del diodo. Nel prossimo paragrafo vedremo in dettaglio l'effetto di tale iniezione di cariche.

*2.11 Iniezione di cariche e ricombinazione

In questo paragrafo calcoleremo la distribuzione di portatori n in un semiconduttore di tipo p , supponendo che in un lato del semiconduttore ci sia una concentrazione nota $n(0)$, ovvero una iniezione di elettroni.

Prima di procedere è necessario scrivere la cosiddetta equazione di continuità. Prendiamo in considerazione un generico volumetto e supponiamo che vi scorra una corrente di elettroni nella direzione x . Se all'interno del volumetto non avvengono fenomeni di ricombinazione, la corrente entrante dovrà essere uguale a quella uscente:

$$j_n(x + \Delta x)\Delta y\Delta z = j_n(x)\Delta y\Delta z \quad (2.31)$$

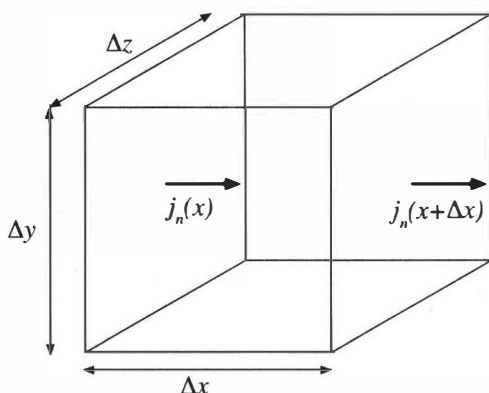


Figura 2.17 Volumetto infinitesimo

In realtà, a causa della ricombinazione elettrone-lacuna ci sarà una diminuzione del numero di elettroni, ovvero una diminuzione di cariche negative. Scriveremo quindi:

$$j_n(x + \Delta x)\Delta y\Delta z - j_n(x)\Delta y\Delta z = (-q)\frac{dn}{dt}\Delta x\Delta y\Delta z \quad (2.32)$$

che in forma differenziale diviene

$$\frac{dj_n}{dx} + q\frac{dn}{dt} = 0 \quad (2.33)$$

A questo punto è necessario avere un'espressione per il tasso di diminuzione degli elettroni, ovvero il tasso di ricombinazione. Il tasso di ricombinazione dipende dal prodotto tra la concentrazione degli elettroni n e la concentrazione delle lacune p :

$$\frac{dn}{dt} = -rnp \quad (2.34)$$

con r opportuna costante. Per il silicio: $r = 5 \times 10^{-18} \text{ cm}^{-3}/\text{s}$.

A rigore un aumento di n , dovuto all'iniezione di portatori, produce una riduzione di p , ma se la concentrazione delle lacune è molto grande ($p \approx N_A$) questa riduzione sarà piccola e quindi possiamo scrivere

$$\frac{dn}{dt} = -\frac{n}{\tau} \quad (2.35)$$

con τ costante definita come:

$$\tau = \frac{1}{rN_A} \quad (2.36)$$

Oltre alla ricombinazione è necessario tener conto della generazione per agitazione termica degli elettroni nella zona p.

$$\frac{dn}{dt} = rn_i^2 = rN_A \frac{n_i^2}{N_A} = \frac{n_{p0}}{\tau} \quad (2.37)$$

avendo definito

$$n_{p0} = \frac{n_i^2}{N_A} \quad (2.38)$$

Tenendo conto sia della ricombinazione sia della generazione, ovvero sommando algebricamente la (2.35) e la (2.37), scriviamo

$$\frac{dn}{dt} = -\frac{n - n_{p0}}{\tau} \quad (2.39)$$

Sostituendo la (2.39) nella (2.33) si ottiene

$$\frac{1}{q} \frac{dj_n}{dx} - \frac{n - n_{p0}}{\tau} = 0 \quad (2.40)$$

Tenendo presente che la corrente di diffusione è legata al gradiente della concentrazione dalla seguente espressione

$$j_n = qD_n \frac{dn}{dx} \quad (2.41)$$

la (2.40) diviene

$$\frac{d^2n}{dx^2} = \frac{n - n_{p0}}{\tau D_n} \quad (2.42)$$

che è, appunto, l'equazione di continuità dei portatori n nella zona p.

Definiamo, per comodità, l'eccesso di elettroni:

$$n' = n - n_{po} \quad (2.43)$$

e la lunghezza di diffusione degli elettroni

$$L_n = \sqrt{\tau D_n} \quad (2.44)$$

La (2.42) può essere scritta più semplicemente come

$$\frac{d^2 n'}{dx^2} = \frac{n'}{L_n^2} \quad (2.45)$$

che ha soluzione

$$n'(x) = n'(0) e^{-\frac{x}{L_n}} \quad (2.46)$$

La (2.46) è stata trovata per iniezione di cariche nel punto $x = 0$. Per iniezione nel punto $x = x_p$ (il limite della regione di svuotamento nella zona p) la (2.46) diviene

$$n'(x) = n'(x_p) e^{-\frac{(x-x_p)}{L_n}} \quad (2.47)$$

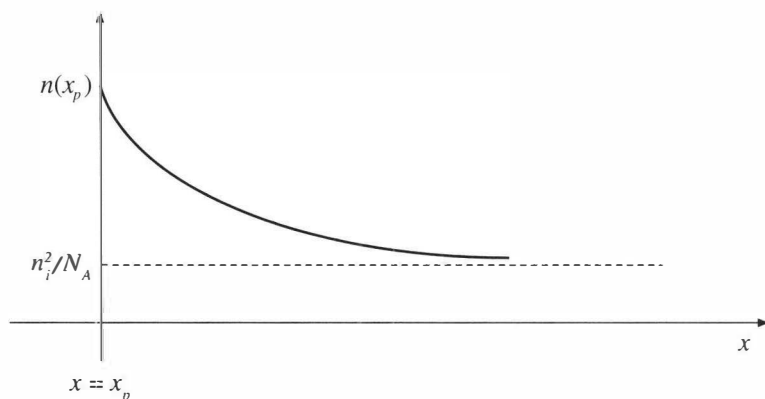


Figura 2.18 Concentrazione di portatori minoritari

Il significato fisico della (2.46) è che l'iniezione di un eccesso di elettroni (rispetto al valore di equilibrio) su una faccia di un semiconduttore p determina due effetti: la diffusione degli elettroni e la loro ricombinazione. Il risultato combinato di diffusione e ricombinazione è una riduzione esponenziale della concentrazione degli elettroni con la distanza dalla superficie dove sono stati iniettati, fino al raggiungimento del valore di equilibrio n_i^2/N_A .

*2.12 Derivazione della caratteristica del diodo

Possiamo ora derivare la caratteristica del diodo a giunzione pn. Di solito una delle due zone è molto più drogata dell'altra, ad esempio la zona n. In tal caso possiamo fare l'ipotesi che la corrente che scorre attraverso la regione di svuotamento sia principalmente dovuta ai portatori n. A causa della tensione esterna v_D al confine tra regione di svuotamento e zona p ($x = x_p$) ci sarà una iniezione di portatori n che per diffusione migreranno verso la zona p. In generale la corrente di diffusione è proporzionale al gradiente della concentrazione. Ciò significa che la corrente di diffusione degli elettroni è massima per $x = x_p$ e poi va a diminuire fino a zero quando la concentrazione degli elettroni raggiunge l'equilibrio (Figura 2.18). Ovviamente la corrente che scorre nel diodo deve essere costante lungo il diodo stesso, quindi se la corrente di diffusione diminuisce, deve essere compensata da una corrente di deriva. In particolare, la corrente sarà interamente di diffusione per $x = x_p$, mentre sarà interamente di deriva lontano dalla giunzione. Possiamo sfruttare questo fatto per calcolare la corrente che scorre sul diodo, semplicemente calcolando la corrente di diffusione per $x = x_p$.

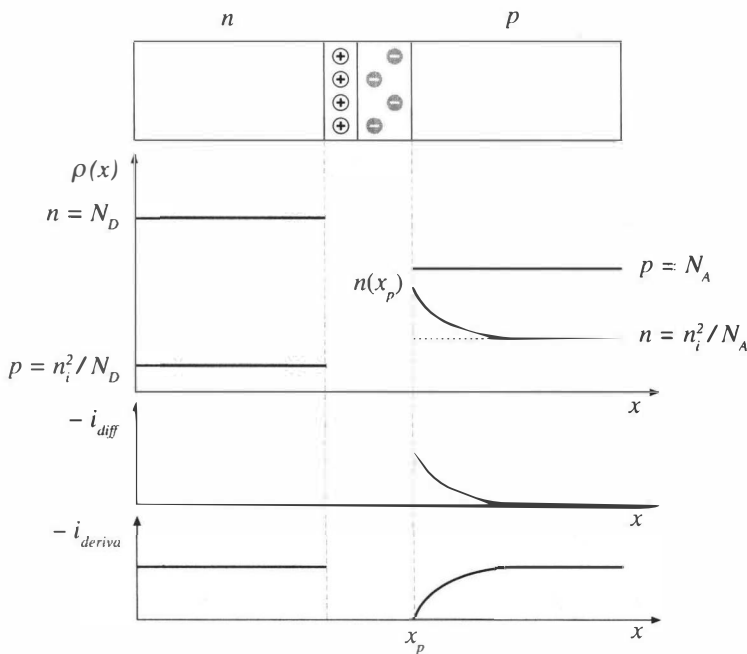


Figura 2.19 Concentrazione dei portatori, corrente di diffusione, corrente di deriva

$$i = -Sj_n^{(diff)} = -qSD_n \frac{dn}{dx} \Big|_{x=x_p} = -qSD_n \frac{dn}{dx} \Big|_{x=x_p} \quad (2.48)$$

con S sezione del diodo. Il segno meno è dovuto al fatto che con riferimento alla geometria di Figura 2.19 scegliamo come verso positivo della corrente quello da destra verso sinistra (che è l'effettivo verso della corrente). Sostituendo la (2.47) nella (2.48) si ottiene

$$i = \frac{qSD_n}{L_n} \left(n(x_p) - \frac{n_i^2}{N_A} \right) \quad (2.49)$$

Infine, sostituendo la (2.30) nella (2.49) si ottiene l'espressione completa della corrente in funzione della tensione applicata, la cosiddetta **caratteristica del diodo ideale**.

$$i = \frac{qSD_n}{L_n} \frac{n_i^2}{N_A} \left(e^{\frac{v}{V_T}} - 1 \right) \quad (2.50)$$

che può essere anche scritta come

$$i = qS \sqrt{\frac{rD_n}{N_A}} n_i^2 \left(e^{\frac{v}{V_T}} - 1 \right) \quad (2.51)$$

*2.13 Correnti di diffusione e correnti di deriva nel diodo

Nel paragrafo precedente abbiamo ragionato sulla distribuzione dei portatori n e p lungo la giunzione.

Poiché eravamo interessati solo all'andamento dei portatori n in prossimità del confine tra la zona di svuotamento e zona p , non ci siamo soffermati sull'andamento effettivo lungo tutto il diodo, da anodo a catodo. Ciò può aver generato qualche perplessità nel lettore: qual è l'andamento delle concentrazioni nella regione di svuotamento e in prossimità dei reofori? La Figura 2.20 mostra l'effettivo andamento delle concentrazioni.

Si noti che quando il diodo è polarizzato direttamente, in prossimità dell'anodo il generatore esterno estrae elettroni e quindi la densità dei portatori n è prossima a zero, in prossimità del catodo inietta elettroni e quindi la densità degli elettroni aumenta rispetto al livello di equilibrio. Gli elettroni immessi poi si ricombinano generando un gradiente di concentrazione. Inoltre per la legge di azione di massa quando l'iniezione degli elettroni diminuisce la concentrazione delle lacune aumenta e viceversa. In definitiva si creano dei gradienti di concentrazioni in prossimità dei reofori.

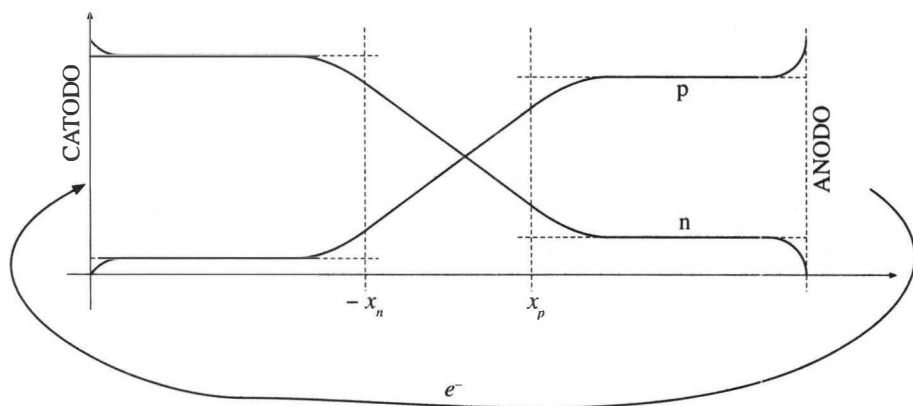


Figura 2.20 Concentrazione dei portatori lungo il diodo da catodo ad anodo

In generale le correnti di diffusione (n e p) sono proporzionali al gradiente della concentrazione, quindi sono presenti nella regione di svuotamento, nelle regioni n e p in prossimità della regione di svuotamento e, per quanto osservato sopra, nelle regioni n e p in prossimità dei reofori come mostrato in Figura 2.21. Ovviamente la somma di corrente di diffusione e corrente di deriva deve essere costante lungo il diodo, pertanto la corrente di deriva sarà minima (praticamente nulla) ai contatti con i reofori e nella regione di svuotamento.

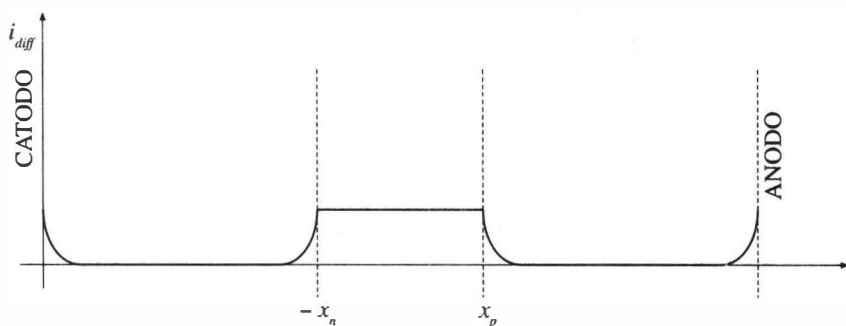


Figura 2.21 Corrente di diffusione nel diodo

*2.14 Limiti della caratteristica del diodo ideale

La caratteristica del diodo ricavata fin qui è detta "ideale". Rispetto all'effettiva caratteristica del diodo di silicio ha delle importanti discrepanze. La più evidente è relativa al valore della soglia.

Nel silicio intrinseco il numero di elettroni liberi è dato da

$$n_i \propto e^{-\frac{E_G}{2kT}} \quad (2.52)$$

con $E_G = E_C - E_V = 1.12\text{eV}$. Sostituendo nell'equazione del diodo ideale si ottiene

$$i \approx I_{s0} e^{-\frac{E_G}{kT}} e^{\frac{v}{V_T}} = I_{s0} e^{-\frac{E_G}{kT}} e^{\frac{qv}{kT}} = I_{s0} e^{\frac{1}{kT}(qv - E_G)} \quad (2.53)$$

con I_{s0} opportuna costante. In generale un esponenziale ha valori molto bassi, prossimi a zero, quando l'argomento è minore di 1 e valori che aumentano repentinamente quando l'argomento è maggiore 1. Possiamo quindi considerare la soglia tra non conduzione (quasi come un isolante) e conduzione (quasi come un conduttore) il valore della tensione per cui

$$v > \frac{E_G}{q} + \frac{kT}{q} \quad (2.54)$$

E_G/q è il potenziale necessario a strappare un elettrone, che per isolanti e semiconduttori è dell'ordine del Volt. kT/q è il potenziale termico che a temperatura ambiente è circa 26 mV. Possiamo quindi trascurare il secondo termine e concludere che la tensione di soglia V_γ è circa pari a

$$V_\gamma = \frac{E_G}{q} \quad (2.55)$$

Per il silicio $E_G = 1.12\text{eV}$ e pertanto la tensione di soglia nei diodi al silicio dovrebbe essere circa 1.12V. Purtroppo è facile verificare sperimentalmente che un diodo al silicio inizia a condurre intorno a 0.6V. Non è una discrepanza trascurabile.

Un'altra notevole discrepanza tra la caratteristica del diodo ideale e la curva sperimentale del diodo di silicio è che la corrente di saturazione inversa del diodo reale è molto più grande (circa un fattore 1000) di quella prevista dall'equazione del diodo ideale e per di più non dipende da n_i^2 , ma da n_i . Quest'ultima affermazione può essere facilmente verificata sperimentalmente misurando la variazione della corrente di saturazione con la temperatura, il numero di elettroni nel silicio intrinseco dipende dalla temperatura mediante la (2.52).

Da quanto detto, è evidente che la teoria fin qui esposta spiega solo una parte di ciò che accade nella giunzione pn. In effetti abbiamo totalmente trascurato la presenza delle cosiddette trappole in banda proibita.

*2.15 Trappole in banda proibita

La caratteristica del diodo derivata nei paragrafi precedenti sarebbe corretta se il silicio cosiddetto intrinseco fosse un cristallo perfetto del tutto puro. In realtà la struttura cristallina del silicio ha sempre qualche imperfezione. Inoltre, per quanto purificato, nel silicio rimangono sempre delle impurità non del tutto tra-

scurabili. Il risultato è che nella cosiddetta banda proibita ci sono in realtà dei livelli energetici detti **stati trappola**, che possono accogliere un elettrone di conduzione e immobilizzarlo. In generale, in relazione a tali livelli, potranno verificarsi quattro fenomeni:

- a) **cattura di un elettrone**, definiamo tasso di cattura R_a
- b) **emissione di un elettrone**, definiamo tasso di cattura R_b
- c) **cattura di una lacuna**, definiamo tasso di cattura R_c
- d) **emissione di una lacuna**, definiamo tasso di cattura R_d

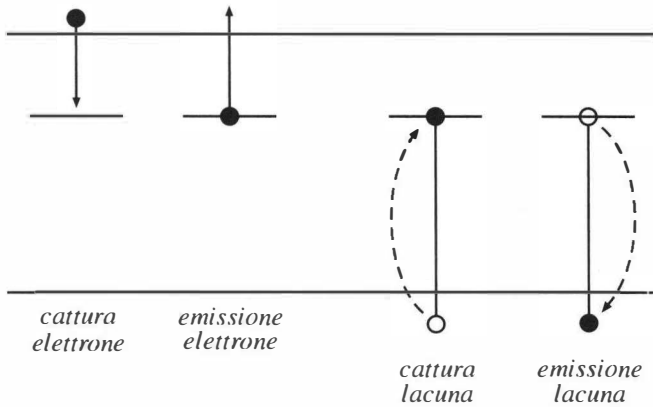


Figura 2.22 Stati trappola

La probabilità di un salto di un elettrone diminuisce esponenzialmente con la differenza di energia tra i livelli. Ma poiché il numero netto di elettroni catturati ($R_a - R_b$) deve essere uguale al numero netto di lacune catturate ($R_c - R_d$), la probabilità che entrambi avvengano è limitata dal salto meno probabile. La conclusione di questo ragionamento è che i processi di ricombinazione elettrone-lacuna attraverso una trappola sono dominati dalle trappole a centro banda, le cosiddette **trappole profonde**.

Consideriamo ora la giunzione pn. Tra le due zone scorre una corrente di elettroni in un verso e di lacune nel verso opposto. Parte dei portatori che fluiscono (sia elettroni, sia lacune) vengono catturati e si ricombinano attraverso le trappole profonde. È un po' come ci fosse un pozzo nella regione di svuotamento nel quale cadono elettroni e lacune. Poiché provengono da direzioni opposte, l'effetto cumulativo è una corrente positiva nel verso del moto delle lacune. Si può dimostrare che vale la seguente relazione

$$i_R \propto e^{-\frac{E_t/q}{2V_T}} \left(e^{\frac{V}{2V_T}} - 1 \right) \quad (2.56)$$

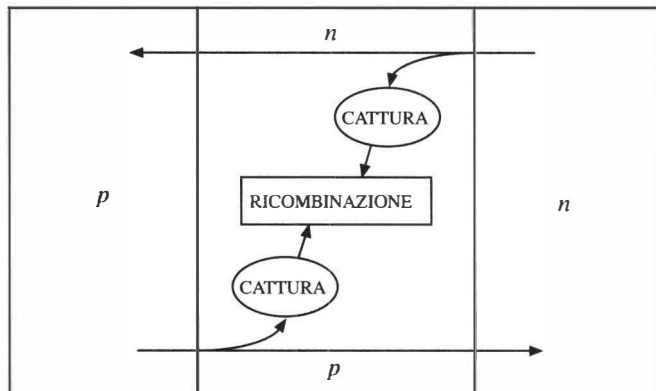


Figura 2.23 Ricombinazione nella regione di svuotamento

La corrente del diodo è quindi data dalla somma della corrente di ricombinazione e la corrente di diffusione:

$$i = I_R \left(e^{\frac{v}{2V_T}} - 1 \right) + I_S \left(e^{\frac{v}{V_T}} - 1 \right) \quad (2.57)$$

con I_R e I_S opportune costanti dipendenti dalla temperatura, ma non dalla tensione applicata. Quando $v \gg V_T = 26 \text{ mV}$ il termine di diffusione ha sempre il sopravvento perché il coefficiente nell'argomento dell'esponenziale della corrente di diffusione è due volte quello della corrente di ricombinazione. Pertanto la caratteristica del diodo ideale è sicuramente corretta per valori grandi di v , non lo è però per il calcolo della soglia. A tal proposito scriviamo la (2.57) esplicitando la dipendenza dalla temperatura dei termini a moltiplicare

$$i = I_{R0} e^{-\frac{E_g/q}{2V_T}} \left(e^{\frac{v}{2V_T}} - 1 \right) + I_{S0} e^{-\frac{E_g/q}{V_T}} \left(e^{\frac{v}{V_T}} - 1 \right) \quad (2.58)$$

con I_{R0} e I_{S0} opportune costanti. Definiamo tensione di soglia il valore per cui $i/I_{S0} = 1$.

$$\frac{i}{I_{S0}} = \frac{I_{R0}}{I_{S0}} e^{-\frac{E_g/q}{2V_T}} \left(e^{\frac{v}{2V_T}} - 1 \right) + e^{-\frac{E_g/q}{V_T}} \left(e^{\frac{v}{V_T}} - 1 \right) \quad (2.59)$$

La (2.59) può essere risolta numericamente. Si noti che il valore della soglia dipende solo dall'energia di gap ($E_g = 1.12 \text{ eV}$ per il silicio) e dal rapporto I_{R0}/I_{S0} . Se I_{R0}/I_{S0} è zero la soglia è 1.12 V , man mano che aumenta la soglia si abbassa. Tipicamente per il silicio I_{R0}/I_{S0} è dell'ordine di 10^7 . Per tale valore si ottiene una tensione di soglia di circa 0.6 V .

2.16 Il diodo Zener

Un diodo ideale continua a far scorrere una piccola corrente di saturazione inversa qualunque tensione di contropolarizzazione venga applicata. Un diodo reale, invece, a una particolare tensione di contropolarizzazione V_z , detta *breakdown*, cessa di essere un (quasi) circuito aperto per diventare un quasi cortocircuito, ovvero inizia a condurre repentinamente. Alcuni diodi, detti Zener, sono progettati per lavorare proprio in questa condizione. L'impiego tipico è come stabilizzatore di tensione.

Le cause di questa sorta di rottura dell'isolamento (breakdown) sono due, che possono operare separatamente oppure concorrere: l'effetto Zener e l'effetto valanga.

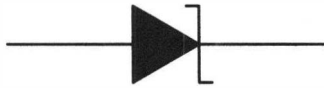


Figura 2.24 Diodo Zener

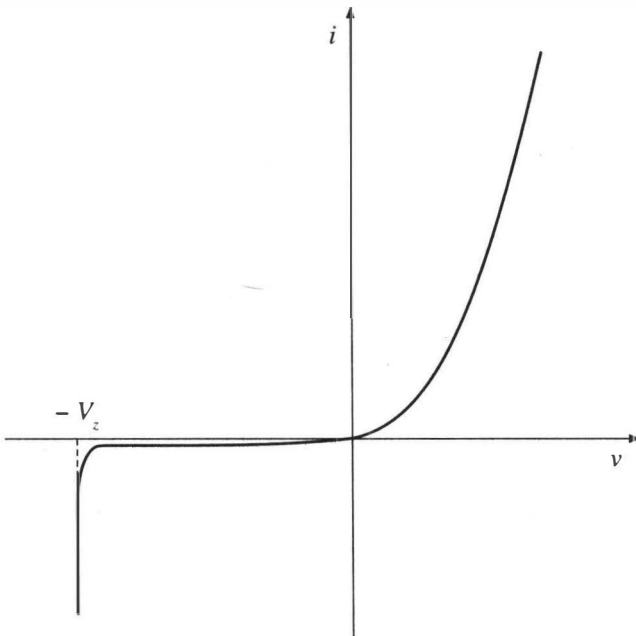


Figura 2.25 Caratteristica del diodo Zener

Come mostrato in Figura 2.26, la tensione di contropolarizzazione v_R ha l'effetto di alzare il potenziale di giunzione, allargare la regione di svuotamento e alzare il valore massimo del campo. Per grandi tensioni di contropolarizzazione il campo elettrico può diventare così intenso ($> 1\text{MV/cm}$) da strappare via gli

elettroni di valenza del silicio generando una scarica, analogamente a quello che accade quando si applica un forte campo elettrico a un isolante. Questo meccanismo è detto “effetto Zener”. All’aumentare della temperatura, diventa più facile strappare elettroni dagli atomi, quindi un aumento di temperatura abbassa la tensione di breakdown dovuta all’effetto Zener.

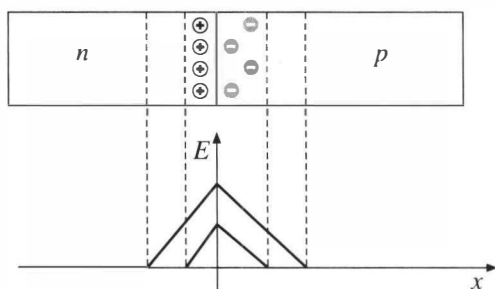


Figura 2.26 Allargamento della regione di svuotamento e aumento del campo elettrico

In alternativa un forte campo elettrico può indurre il cosiddetto “effetto valanga”. Per agitazione termica nella regione di svuotamento si creano elettroni e lacune. Gli elettroni, accelerati dall’intenso campo elettrico, possono raggiungere una velocità tale che, quando urtano con un atomo, hanno abbastanza energia cinetica da strappare un elettrone e creare un’ulteriore coppia elettrone lacuna. Questo meccanismo si autosostiene in quanto gli elettroni così strappati a loro volta sono accelerati dal campo elettrico fino a raggiungere una energia cinetica in grado di strappare altri elettroni. Per questa ragione è detto “effetto valanga”. All’aumentare della temperatura la mobilità degli elettroni diminuisce, quindi è necessario un campo elettrico maggiore per innescare la valanga, pertanto un aumento di temperatura alza la tensione di breakdown dovuta all’effetto valanga.

I diodi Zener sono realizzati con drogaggi tali che entrambi i meccanismi intervengano. Poiché al variare della temperatura i due effetti hanno comportamento opposto (l’uno diminuisce, l’altro aumenta la tensione di breakdown) il risultato è che la tensione di breakdown risulta estremamente costante non solo al variare della corrente che vi scorre, ma anche al variare della temperatura. Per questa ragione gli Zener sono utilizzati per fornire tensioni stabili di riferimento come mostrato in Figura 2.27.

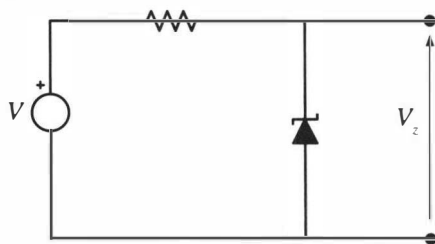


Figura 2.27 Tensione stabilizzata con Zener

Si noti la presenza della resistenza in serie con lo Zener che ha lo scopo di limitare la corrente su diodo e produrre la caduta di tensione opportuna tra la tensione di stabilizzare (V^*) e la tensione del diodo. Senza tale resistenza la corrente sullo Zener aumenterebbe indefinitamente fino a danneggiare il diodo.

2.17 Riepilogo

- Una giunzione pn è costituita da due zone adiacenti drogate rispettivamente p e n.
- In una giunzione pn la corrente che scorre è legata alla tensione (positiva nel senso dalla zona p alla zona n) dalla seguente relazione

$$i = I_s \left(e^{\frac{v}{V_T}} - 1 \right)$$

dove I_s è una costante e V_T è la tensione termica (26mV).

- La caratteristica tensione corrente di una giunzione pn di silicio può essere spesso approssimata come segue: $i = 0$ per $v < 0.6V$; quando $i > 0$, $v \approx 0.6V$.
- Tra la zona p e n di una giunzione pn si crea una regione di svuotamento (ovvero senza cariche mobili) di larghezza w pari a:

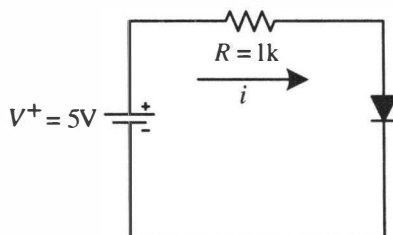
$$w = \sqrt{\frac{2\varepsilon}{q}(\varphi - v) \left(\frac{1}{N_A} + \frac{1}{N_D} \right)}$$

con ε costante dielettrica del silicio, q carica elementare dell'elettrone, $\varphi \approx 0.6V$, v tensione ai capi della giunzione pn (positiva nel verso da p a n), N_A concentrazione delle impurità nella zona p, N_D concentrazione delle impurità nella zona n.

- Una giunzione metallo semiconduttore ha una curva caratteristica simile alla giunzione pn (la tensione di soglia è però tipicamente dell'ordine di 0.1V).
- Alcune giunzioni pn (dette "diodi Zener") sono progettate per lavorare in polarizzazione inversa in una modalità detta di breakdown in cui la tensione è praticamente costante al variare della corrente. Questi diodi sono utilizzati come stabilizzatori di tensione.

2.18 Esercizi svolti

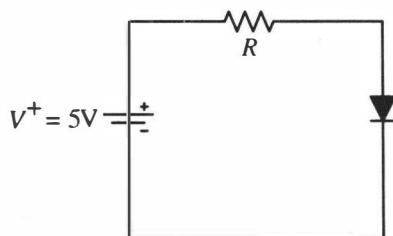
1) Con riferimento al circuito in figura si calcoli la corrente che eroga il generatore di tensione.



Poiché la tensione sul diodo in conduzione può essere considerata pari a 0.6V indipendentemente dalla corrente che vi scorre, possiamo calcolare la corrente i come segue:

$$i = \frac{V^+ - 0.6}{R} = 4.4 \text{ mA}$$

2) Con riferimento al circuito in figura si dimensioni la resistenza tenendo presente che il diodo non è in grado di dissipare più di 125mW.

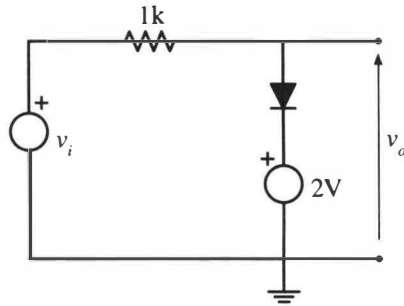


La potenza dissipata sul diodo è data da

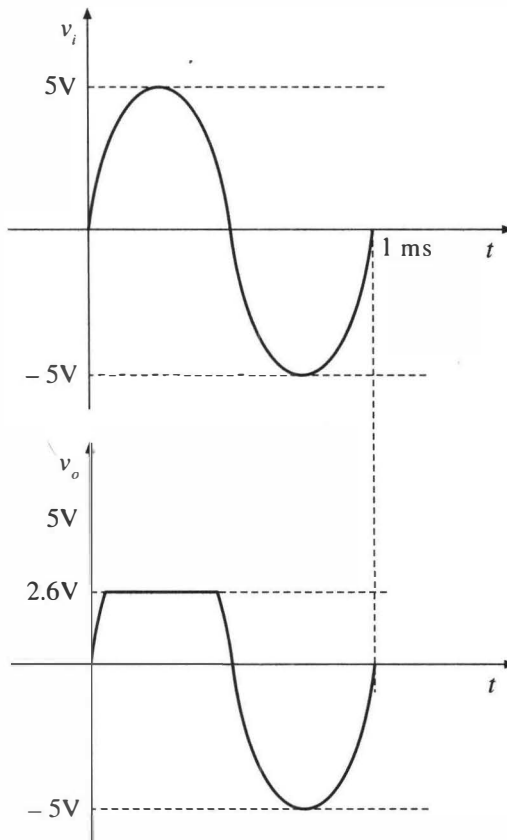
$$P = (0.6\text{V})i = (0.6\text{V}) \frac{4.4\text{V}}{R}$$

Ponendo la potenza pari a 125mW si ottiene $R > 21\Omega$.

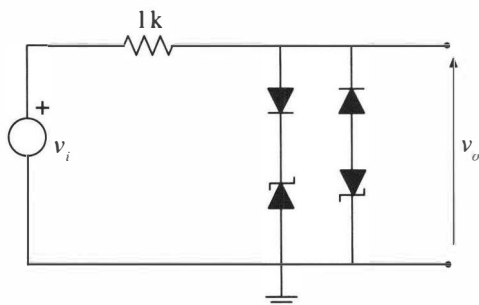
3) Con riferimento al circuito in figura, il generatore di tensione fornisce un segnale sinusoidale $v_i = V \sin(2\pi ft)$ con $V = 5\text{V}$, $f = 1\text{kHz}$. Si tracci la forma d'onda v_o in uscita.



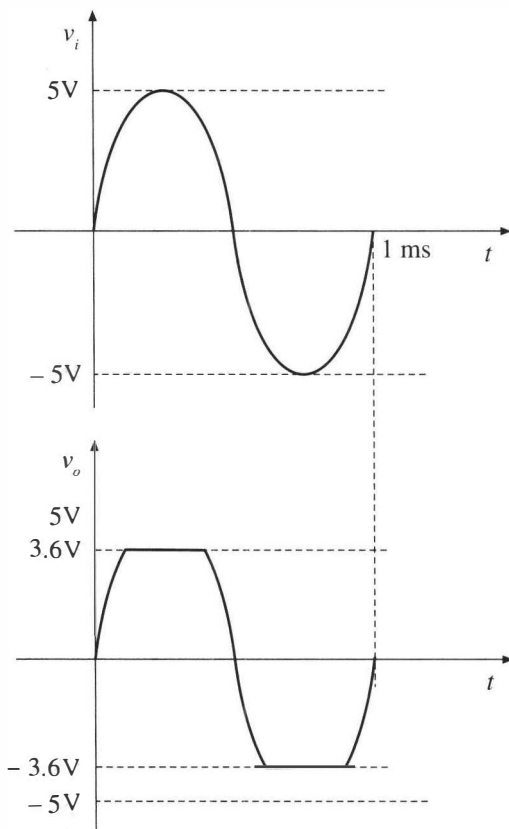
Il diodo è in conduzione solo quando $v_i > 2.6\text{V}$. Quando il diodo è in conduzione $v_o = 2.6\text{V}$ pertanto la forma d'onda in uscita è quella mostrata in figura.



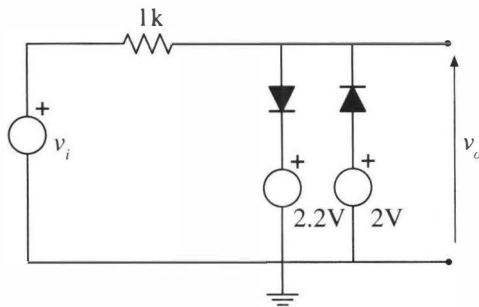
4) Con riferimento al circuito in figura, il generatore di tensione fornisce un segnale sinusoidale $v_i = V \sin(2\pi f t)$ con $V = 5\text{V}$, $f = 1\text{kHz}$. I diodi Zener hanno tensione di breakdown pari a 3V . Si tracci la forma d'onda v_o in uscita.



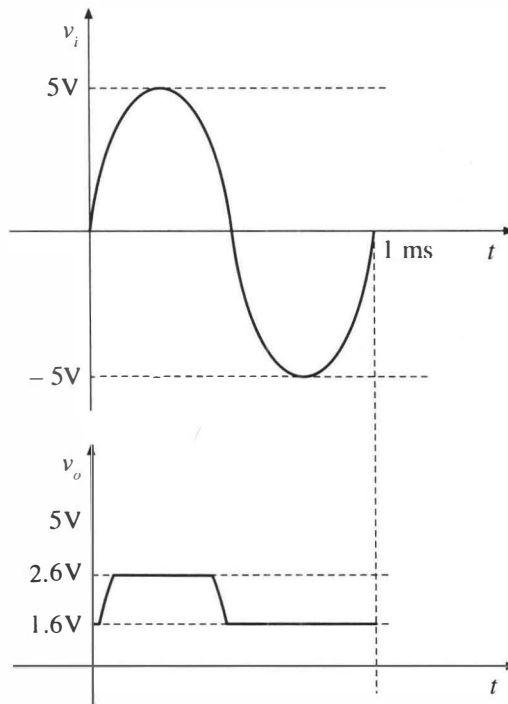
Il primo ramo è in conduzione solo quando $v_i > 3.6\text{V}$. Quando il ramo è in conduzione $v_o = 3.6\text{V}$. Il secondo ramo è in conduzione solo quando $v_i < -3.6\text{V}$. Quando il ramo è in conduzione $v_o = -3.6\text{V}$. Pertanto la forma d'onda in uscita è quella mostrata in figura.



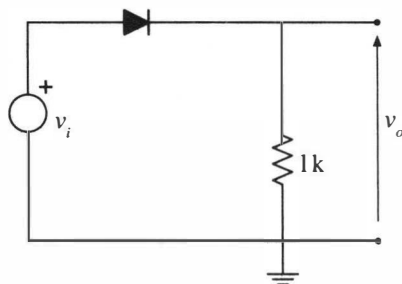
5) Con riferimento al circuito in figura, il generatore di tensione fornisce un segnale sinusoidale $v_i = V \sin(2\pi f t)$ con $V = 5\text{V}$, $f = 1\text{kHz}$. Si tracci la forma d'onda v_o in uscita.



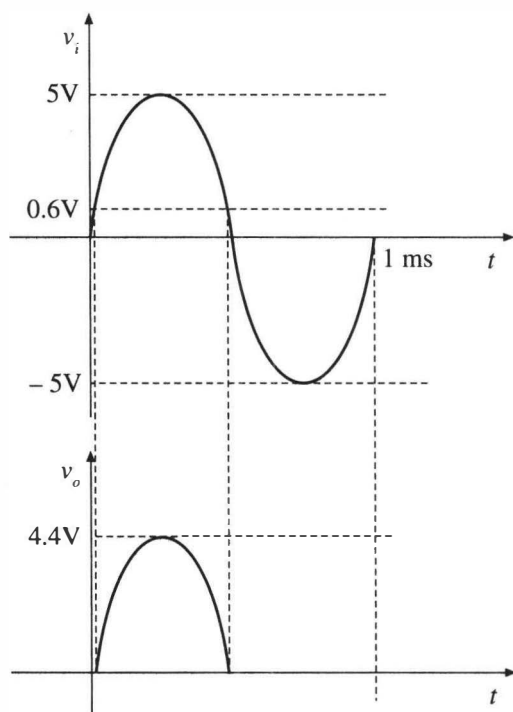
Il primo ramo è in conduzione solo quando $v_i > 1.6\text{V}$. Quando il ramo è in conduzione $v_o = 1.6\text{V}$. Il secondo ramo è in conduzione solo quando $v_i < 2.6\text{V}$. Quando il ramo è in conduzione $v_o = 2.6\text{V}$. Pertanto la forma d'onda in uscita è quella mostrata in figura.



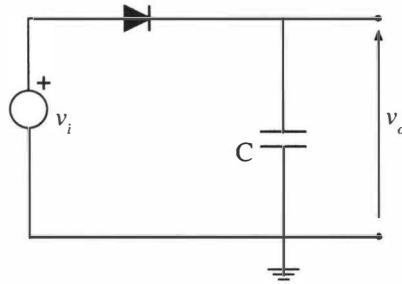
6) Con riferimento al circuito in figura, il generatore di tensione fornisce un segnale sinusoidale $v_i = V \sin(2\pi f t)$ con $V = 5\text{V}$, $f = 1\text{kHz}$. Si tracci la forma d'onda v_o in uscita.



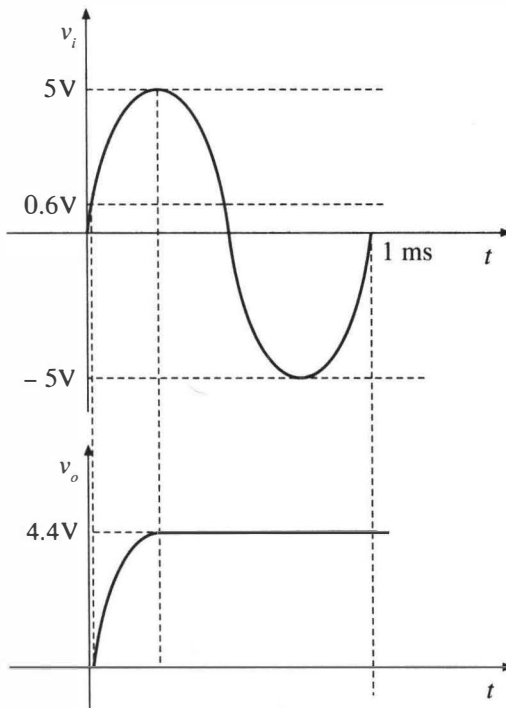
Il diodo è in conduzione solo quando $v_i > 0.6\text{V}$. Quando il diodo è in conduzione $v_o = v_i - 0.6\text{V}$ pertanto la forma d'onda in uscita è quella mostrata in figura.



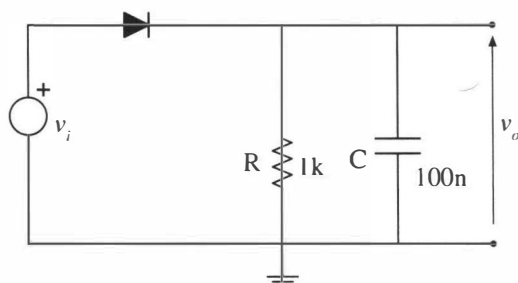
7) Con riferimento al circuito in figura, il generatore di tensione fornisce un segnale sinusoidale $v_i = V\sin(2\pi ft)$ con $V = 5\text{V}$, $f = 1\text{kHz}$. Si tracci la forma d'onda v_o in uscita.



Il condensatore è caricato da v_i mediante il diodo. Finché v_i è crescente e maggiore di 0.6V il diodo è in conduzione e $v_o = v_i - 0.6\text{V}$. Ma quando v_i , raggiunto il picco, inizia a diminuire, il diodo è OFF e il condensatore non può più scaricarsi. L'andamento della forma d'onda è in figura.

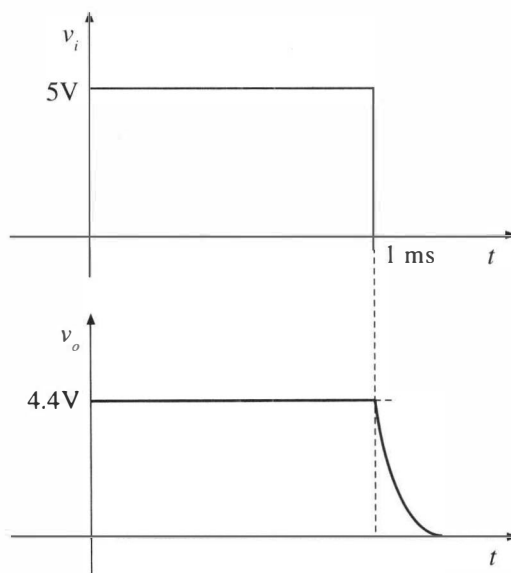


8) Con riferimento al circuito in figura, il generatore di tensione fornisce un impulso positivo di ampiezza 5V durata 1ms con fianchi molto ripidi. Si tracci la forma d'onda v_o in uscita.

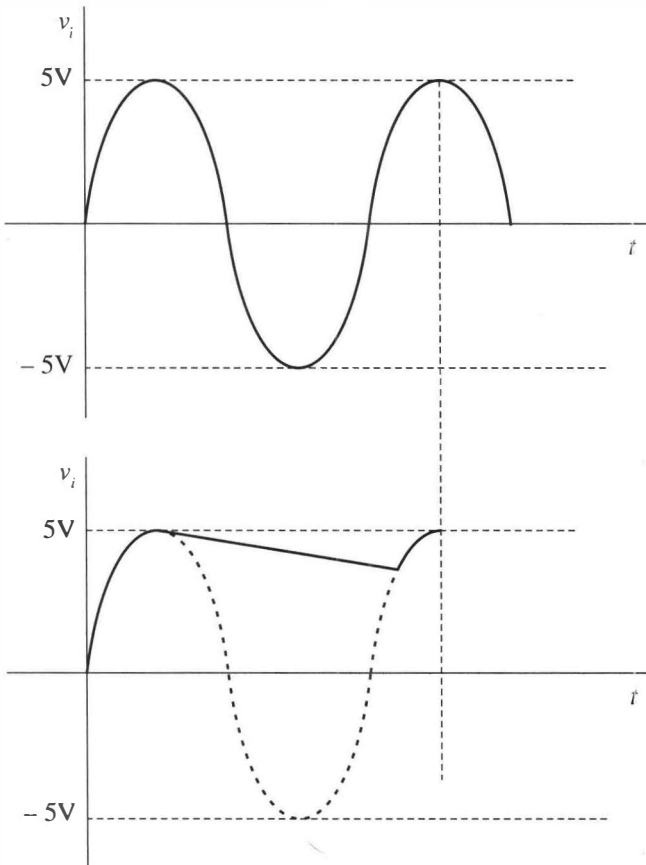
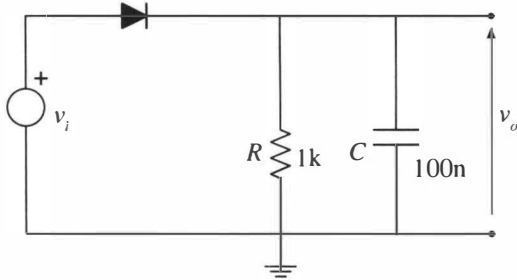


Il capacitore è caricato da v_i mediante il diodo. Quando v_i , raggiunto il picco, inizia a diminuire, il diodo è OFF e il capacitore si scarica attraverso la resistenza $R = 1\text{ k}\Omega$. La scarica ha andamento nel tempo (vedi Appendice A):

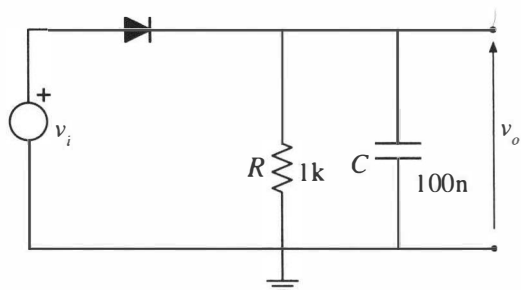
$$v(t) = Ve^{\frac{t-t_i}{RC}} = 4.4e^{\frac{t-1\text{ms}}{0.1\text{ms}}}$$



9) Con riferimento al circuito in figura, il generatore di tensione fornisce un segnale sinusoidale $v_i = V \sin(2\pi f t)$ con $V = 5\text{V}$, $f = 10\text{kHz}$. Si tracci la forma d'onda v_o in uscita. Per semplicità si assuma che la tensione di soglia del diodo sia $V_\gamma = 0$.



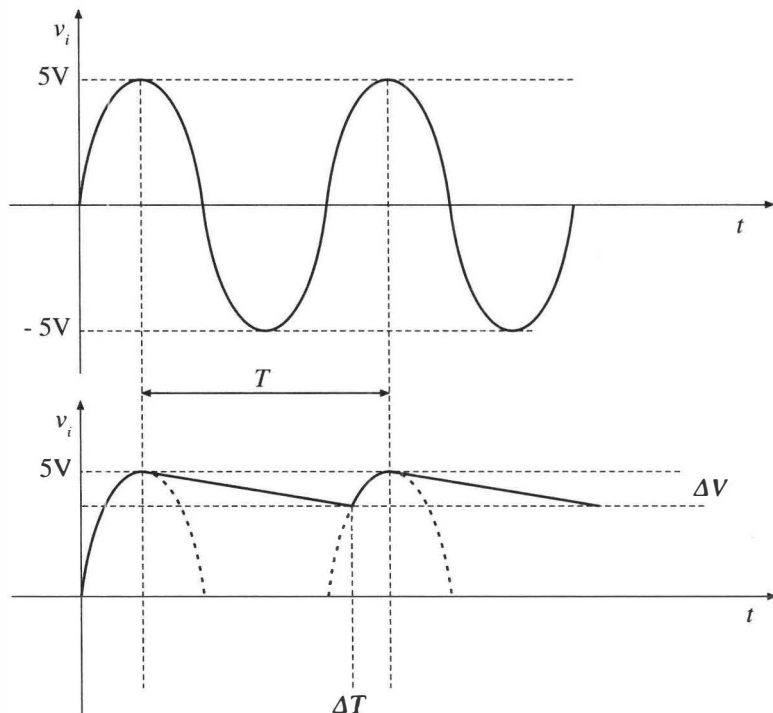
***10)** Con riferimento al circuito in figura, il generatore di tensione fornisce un segnale sinusoidale $v_i = V\sin(2\pi ft)$ con $V = 5\text{V}$, $f = 100\text{kHz}$. Si tracci la forma d'onda v_o in uscita e si calcoli la massima escursione a regime del segnale in uscita. Per semplicità si assuma che la tensione di soglia del diodo sia $V_\gamma = 0$.



L'esercizio è simile al 9, l'unica differenza è che il tempo di scarica del capacitore è molto più lungo. Poiché $RC = 100\mu\text{s} \gg T = 10\mu\text{s}$, possiamo approssimare:

$$v(t) = Ve^{-\frac{t}{RC}} \approx V\left(1 - \frac{t}{RC}\right)$$

L'andamento è quello mostrato in figura.

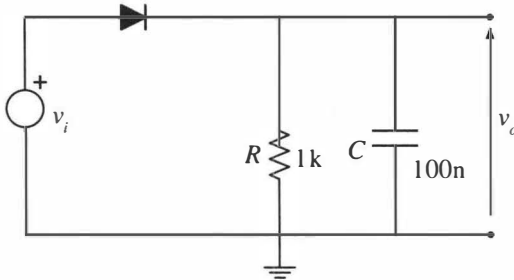


Poiché la scarica è lineare, la massima escursione a regime è pari a

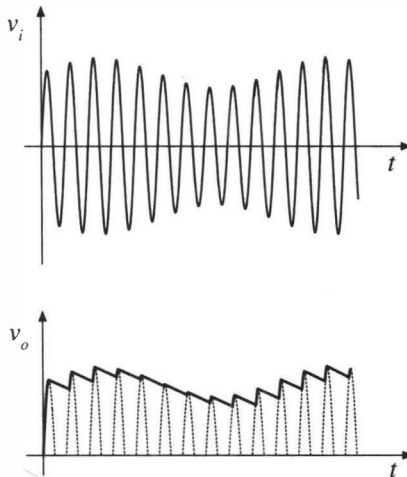
$$\Delta V = \frac{V}{RC}(T - \Delta t) \approx \frac{VT}{RC} = 0.5V$$

*Nota: il circuito oggetto di questo esercizio è un **raddrizzatore di tensione**, che consente di ottenere un segnale in continua (DC: direct current) circa pari all'ampiezza del segnale in alternata (AC: alternate current).*

***11)** Con riferimento al circuito in figura, il generatore di tensione fornisce un segnale sinusoidale $v_i = (V + v_m)\sin(2\pi f t)$ con $v_m = A\sin(2\pi f_m t)$ ($V = 5V, f = 100kHz, f_m = 1kHz, A = 0.5V$) Si tracci la forma d'onda v_o . Per semplicità si assuma che la tensione di soglia del diodo sia $V_\gamma = 0$.

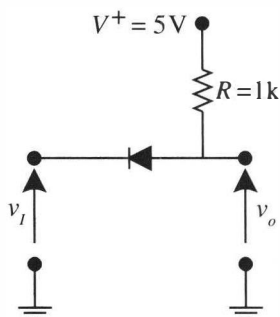


Poiché il segnale v_m è molto più lento di $V\sin(2\pi f t)$ possiamo procedere come nell'esercizio 10 ottenendo l'andamento mostrato in figura

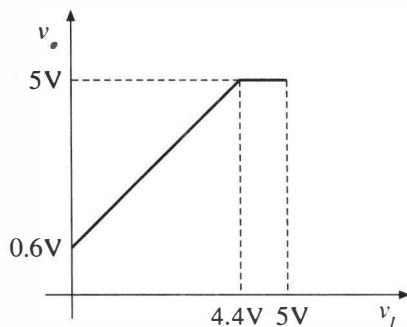


*Nota: il circuito oggetto di questo esercizio è un **demodulatore AM a diodo**, che consente di estrarre l'informazione (il segnale v_m) da una portante a radiofrequenza modulata in ampiezza. Questo circuito, oggi non più usato con questo scopo, è stato fondamentale per la nascita delle radiocomunicazioni nei primi anni del Novecento.*

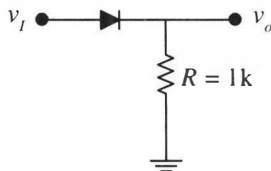
12) Con riferimento al circuito in figura, si tracci la funzione di trasferimento tra 0 e 5V, ovvero il grafico della tensione in uscita v_o in funzione di v_i .



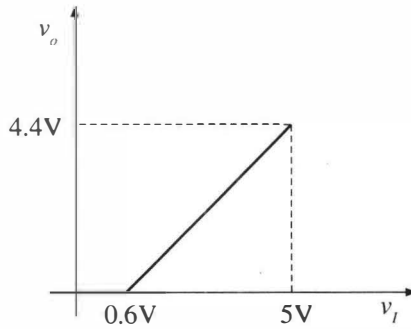
Se $v_i < 5V - 0.6V = 4.4V$, il diodo è in conduzione e la tensione in uscita è pari alla tensione in ingresso più la caduta di 0.6V sul diodo. Per $v_i > 4.4V$ il diodo è interdetto e la tensione in uscita è tenuta a 5V dalla resistenza (detta di pull-up).



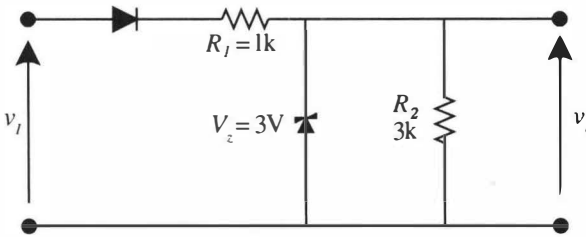
13) Con riferimento al circuito in figura, si tracci la funzione di trasferimento tra 0 e 5V, ovvero il grafico della tensione in uscita v_o in funzione di v_i .



Il diodo entra in conduzione per $v_i > 0.6\text{V}$. Per valori minori la resistenza forza l'uscita a massa, per valori maggiori l'uscita è uguale all'ingresso meno la caduta di 0.6V sul diodo.



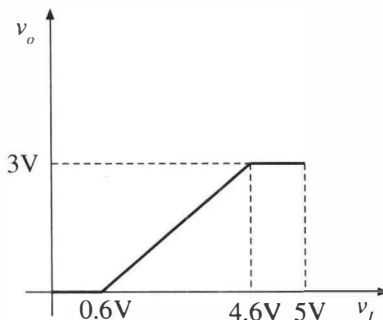
14) Con riferimento al circuito in figura, si tracci la funzione di trasferimento tra 0 e 5V , ovvero il grafico della tensione in uscita v_o in funzione di v_i .



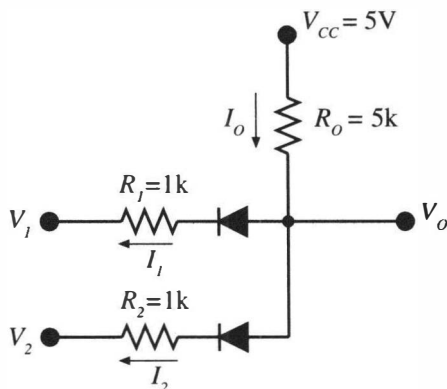
Il diodo entra in conduzione per $v_i > 0.6\text{V}$. Per valori minori la resistenza R_2 forza l'uscita a massa, per valori maggiori il diodo entra in conduzione e finché lo Zener è interdetto l'uscita è uguale a

$$v_o = \frac{v_i - 0.6\text{V}}{4\text{k}\Omega} 3\text{k}\Omega$$

Quando la tensione in uscita raggiunge la tensione di breakdown di 3V il diodo Zener cessa di essere un circuito aperto e mantiene la tensione in uscita a 3V .



15) Con riferimento al circuito si calcoli la tensione in uscita considerando che gli ingressi possono avere valori pari a 0V e 5V.



Quando entrambi gli ingressi sono a 0V

$$V_{CC} - (0.6\text{V}) = I_1 R_1 + I_o R_o$$

Poiché $I_o = 2I_1$

$$V_{CC} - (0.6\text{V}) = I_1 R_1 + 2I_1 R_o$$

$$I_1 = \frac{V_{CC} - (0.6\text{V})}{R_1 + 2R_o} = 0.4\text{mA}$$

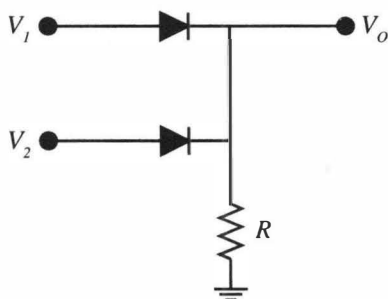
Quindi

$$V_o = V_{CC} - R_o I_o = V_{CC} - 2R_o I_1 = 1\text{V}$$

Gli altri tre casi si risolvono in maniera analoga.

V_1	V_2	V_o
0V	0V	1V
0V	5V	1.33V
5V	0V	1.33V
5V	5V	5V

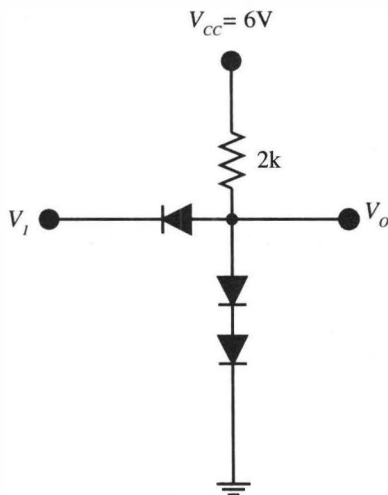
16) Con riferimento al circuito si calcoli la tensione in uscita considerando che gli ingressi possono avere valori pari a 0V e 5V. $R = 1\text{k}\Omega$.



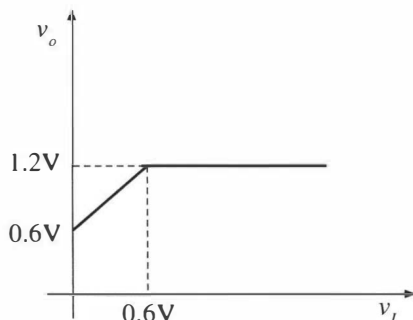
Se i due ingressi sono a massa (0V) non scorre corrente né attraverso diodi né attraverso la resistenza. La caduta di tensione sulla resistenza è quindi nulla e $V_o = 0$. Se uno dei due ingressi è a 5V, scorre corrente su un diodo e quindi anche sulla resistenza. La caduta di tensione sul diodo è 0.6V, quindi l'uscita è 4.4V.

V_1	V_2	V_o
0V	0V	0V
0V	5V	4.4V
5V	0V	4.4V
5V	5V	4.4V

17) Si tracci la funzione di trasferimento del circuito in figura.



Quando $V_i = 0$, la corrente scorre solo sul diodo nel ramo di ingresso, perché i due diodi in serie necessitano di una tensione pari a $0.6V + 0.6 = 1.2V$ e il diodo nel ramo di ingresso forza V_o a $0.6V$. Quando la tensione di ingresso è pari a $0.6V$ il ramo con i due diodi entra in conduzione e forza l'uscita a $1.2V$. Se la tensione di ingresso supera $1.2V$, il diodo nel ramo di ingresso è OFF e i due diodi in serie mantengono l'uscita a $1.2V$.



2.19 Domande di verifica a scelta multipla

1. La tensione di soglia di un diodo
 - a) aumenta all'aumentare della sezione del diodo
 - b) aumenta all'aumentare della temperatura
 - c) per i diodi al silicio è circa $0.6V$
 - d) ha lo stesso valore della tensione termica
2. Tra una zona p e una zona n si crea una regione di svuotamento
 - a) perché portatori p e n essendo di segno opposto si respingono creando una zona senza portatori
 - b) perché i portatori n e p muovendosi a caso tendono a fluire verso zone dove minore è la loro densità
 - c) perché il campo elettrico allontana le cariche nella zona di contatto tra la zona p e n
 - d) perché il potenziale di giunzione tende a muovere i portatori p e n in modo da creare una zona senza cariche elettriche
3. Una sola delle seguenti affermazioni è vera.
 - a) Il potenziale di giunzione dipende dalla temperatura
 - b) Il potenziale di giunzione è lo stesso per tutti i diodi al silicio
 - c) Il potenziale di giunzione è la tensione che si legge con un multimetro applicato ad anodo e catodo di un diodo quando non scorre corrente
 - d) Il potenziale di giunzione è determinato dall'effetto tunnel tra il semiconduttore e i reofori metallici
4. Una sola delle seguenti affermazioni è vera.
 - a) In un diodo scorrono solo correnti di diffusione

- b) La corrente di diffusione è massima allontanandosi dalla regione di svuotamento
 c) Man mano che i portatori minoritari si ricombinano aumenta la corrente di deriva
 d) In un diodo possono scorrere solo correnti di deriva
5. Una sola delle seguenti affermazioni è vera.
 a) Si crea una giunzione Schottky ogni volta che si mettono in contatto materiali con diversa energia di Fermi
 b) La giunzione Schottky è detta anche contatto ohmico
 c) La corrente scorre in una giunzione Schottky per effetto breakdown
 d) Nella regione di svuotamento di una giunzione Schottky ci sono ioni fissi di un solo tipo: o tutti positivi o tutti negativi
6. All'aumentare della contropolarizzazione di una giunzione pn reale la corrente di contropolarizzazione
 a) aumenta sensibilmente a causa dell'effetto Early
 b) aumenta esponenzialmente come quando il diodo è polarizzato direttamente
 c) aumenta leggermente a causa della ricombinazione dei portatori minoritari nella zona n
 d) aumenta leggermente a causa dell'aumento della generazione di coppie elettrone-lacuna nella regione di svuotamento che aumenta all'aumentare della tensione di contropolarizzazione
7. In un diodo polarizzato direttamente, quando scorre una grande corrente
 a) la corrente tende a saturare a un valore massimo determinato dallo strozzamento del canale
 b) la corrente tende ad aumentare linearmente piuttosto che esponenzialmente
 c) entra in breakdown e la tensione si stabilizza a un valore di circa 3.3V
 d) entra in breakdown e la corrente si stabilizza per effetto tunnel
8. Una giunzione metallo-semiconduttore
 a) è sempre un contatto ohmico
 b) si comporta sempre come un circuito aperto e per questa ragione permette di realizzare i gate dei MOSFET
 c) può comportarsi come una giunzione pn
 d) è sempre un circuito chiuso e per questa ragione permette di realizzare i reofori dei dispositivi a semiconduttore
9. Nella regione di svuotamento tra una zona p e una zona n
 a) c'è un campo elettrico che è la derivata del potenziale termico tra i due materiali
 b) c'è un campo elettrico generato dagli ioni fissi nella regione di svuotamento
 c) c'è un campo elettrico che è il gradiente della buca di potenziale che genera l'effetto tunnel tra lacune vicine
 d) non c'è alcun campo elettrico
10. Si consideri una giunzione pn: $x = 0$ è il confine tra zona p e zona n. $x = -x_n$ è il confine tra zona n e regione di svuotamento, $x = x_p$ è il confine tra zona p e regione di svuotamento, E è il campo elettrico, φ è il potenziale di giunzione. Una sola delle seguenti relazioni è vera.

a)
$$\int_{-\infty}^{+\infty} E(x) dx = \varphi$$

$$b) \int_0^{+\infty} E(x) dx = \varphi$$

$$c) \int_0^{x_p} E(x) dx = \varphi$$

$$d) \int_{-x_n}^{x_p} E(x) dx = w$$

11. Si consideri una giunzione pn: $x = 0$ è il confine tra zona n e zona p, $x = -x_n$ è il confine tra zona n e regione di svuotamento, $x = x_p$ è il confine tra zona p e regione di svuotamento, N_D è la concentrazione delle impurità nella zona n. Il campo elettrico nella zona $-x_n < x < 0$ è dato da:

$$a) E(x) = \frac{qN_D}{\epsilon} x + \frac{qN_D}{\epsilon} x_n$$

$$b) E(x) = \frac{qN_D}{\epsilon} x^2 + \frac{qN_D}{\epsilon} x_n^2$$

$$c) E(x) = \frac{qN_D}{w\epsilon} x^2 - \frac{qN_D}{w\epsilon} x_n^2$$

$$d) E(x) = -\frac{qN_D}{\epsilon} x + \frac{qN_D}{\epsilon} x_n$$

- *12. Una sola delle seguenti affermazioni, relative a un diodo in conduzione, è vera.

- Scorrono solo correnti di diffusione
- La corrente di deriva è massima nella zona di contatto tra n e p in quanto nella regione di svuotamento c'è un campo elettrico che la produce
- Le correnti di diffusione sono localizzate alle interfacce tra zona n e p, tra n e reoforo e tra p e reoforo
- La somma delle correnti di deriva e di diffusione è nulla su ogni sezione del diodo

13. La regione di svuotamento in una giunzione pn

- aumenta all'aumentare della corrente che scorre nella giunzione
- ha spessore nullo quando la tensione applicata ai capi della giunzione è nulla
- si estende principalmente nella zona meno drogata
- Si estende allo stesso modo nelle due zone, perché la carica elettrica nelle due zone (n e p) è la stessa

14. In una giunzione pn

- l'energia di Fermi è circa l'energia di valenza nella zona p e circa l'energia di conduzione nella zona n
- per la legge di azione di massa, l'energia di Fermi è nulla sia nella zona n sia nella zona p
- l'energia di conduzione ha lo stesso valore nella zona p e nella zona n
- l'energia di Fermi ha lo stesso valore nella zona p e nella zona n

15. Il diodo Zener

- si usa come riferimento di tensione
- è una giunzione metallo-semiconduttore
- si usa nei ricevitori AM per demodulare il segnale
- si usa come stabilizzatore di corrente nei circuiti di potenza

16. Un diodo Zener
- a) quando è applicata una tensione di contropolarizzazione non conduce
 - b) quando è applicata una tensione di polarizzazione diretta non conduce
 - c) la corrente può scorrere solo dal catodo verso l'anodo
 - d) la corrente può scorrere sia dal catodo verso l'anodo, sia viceversa
- *17. La tensione di soglia in un diodo
- a) è sempre pari all'energia di gap diviso la carica dell'elettrone
 - b) nel silicio è circa pari all'energia di gap diviso due volte la carica dell'elettrone
 - c) è un valore che dipende dalla lunghezza del diodo, più è lungo minore è la soglia
 - d) è determinata dalla concentrazione delle impurità nella zona più drogata
- *18. La caratteristica reale del diodo
- a) tende a divenire una parabola per grandi correnti
 - b) è esponenziale per tensioni tra 0V e 0.6V
 - c) quando la tensione supera 0.8V la corrente ha un massimo e poi diminuisce all'aumentare della tensione
 - d) la corrente diventa repentinamente negativa quando la tensione applicata supera i 3V
19. Nella regione di svuotamento di una giunzione pn, il campo elettrico
- a) è nullo, infatti non ci sono cariche mobili
 - b) ha verso positivo da p a n
 - c) è nullo, ma è presente una differenza di potenziale alle sue estremità
 - d) si oppone al moto degli elettroni dalla zona n alla zona p
20. In una giunzione pn
- a) quando la tensione applicata è nulla scorre una piccola corrente detta corrente di saturazione
 - b) quando è applicata una tensione di contropolarizzazione la corrente è circa nulla fino a circa 0.6V, poi aumenta repentinamente
 - c) quando è polarizzata direttamente, la corrente aumenta di una decade (un fattore 10) ogni 60mV
 - d) la caratteristica tensione-corrente è simmetrica e lineare
21. In una giunzione pn con zona n molto più drogata della zona p, lo spessore della zona di svuotamento
- a) aumenta all'aumentare del drogaggio della zona n
 - b) diminuisce all'aumentare del drogaggio della zona n, ma aumenta all'aumentare del drogaggio della zona p
 - c) dipende solo dal tipo di semiconduttore (silicio, germanio, arseniuro di gallio)
 - d) diminuisce all'aumentare del drogaggio della zona p
22. In un volumetto infinitesimo di silicio p con le facce parallele alla corrente di elettroni
- a) la corrente di elettroni entrante è sempre uguale alla corrente di elettroni uscente
 - b) il numero di elettroni è costante
 - c) il tasso di ricombinazione degli elettroni è proporzionale alla loro concentrazione
 - d) la corrente di elettroni in un senso deve essere sempre uguale alla corrente di lacune nell'altro senso

2.20 Domande di verifica a risposta aperta

1. Disegnare la caratteristica del diodo ideale.
2. Calcolare lo spessore della regione di svuotamento in funzione del potenziale di giunzione.
3. Calcolare la concentrazione dei portatori minoritari al confine della regione di svuotamento.
- *4. Calcolare l'andamento della concentrazione dei portatori n quando iniettati in un semiconduttore p .
- *5. Derivare la caratteristica del diodo, nota l'espressione al punto precedente.
- *6. Descrivere come la caratteristica reale del diodo si discosta dall'andamento ideale.
7. Si tracci la caratteristica del diodo Zener e se ne illustri il principio di funzionamento.

2.21 Risposte alle domande a scelta multipla

1 c	10 a	19 d
2 b	11 a	20 c
3 a	12 c	21 d
4 c	13 c	22 c
5 d	14 d	
6 d	15 a	
7 b	16 d	
8 c	17 b	
9 b	18 b	

3

Fotodiodi e LED

In questo capitolo si forniscono le basi fisiche del funzionamento dei fotodiodi e dei LED. Si richiama il concetto di fotone e di interazione luce-semiconduttore per poi mostrare come una giunzione pn non polarizzata o contropolarizzata generi una corrente quando illuminata. Dopo aver introdotto la differenza tra semiconduttori diretti e indiretti, si mostra come una giunzione polarizzata direttamente possa generare un flusso di fotoni per ricombinazione radiativa. Il capitolo termina con un breve paragrafo di inquadramento storico.

3.1 Interazione luce-semiconduttore

Nell'anno 1900, il professore di fisica Max Planck, sebbene inizialmente con poca convinzione, fece un'ipotesi che avrebbe rivoluzionato la fisica: in natura gli scambi energetici non avvengono in maniera continua, ma attraverso pacchetti. In altre parole due sistemi fisici non possono scambiarsi una quantità qualsiasi di energia, ma solo un multiplo di una quantità minima di energia detta *quanto*. Se lo scambio energetico avviene mediante radiazione elettromagnetica l'energia E del quanto è data dalla seguente formula:

$$E = \frac{hc}{\lambda} \quad (3.1)$$

con h costante di Planck, c velocità della luce e λ lunghezza d'onda.

L'interazione tra onde elettromagnetiche e semiconduttori non fa eccezione.

I livelli energetici in un semiconduttore presentano la tipica distribuzione a bande: banda di conduzione per $E > E_C$, banda di valenza per $E < E_V$ e banda proibita, ovvero una banda priva di livelli energetici, tra la valenza e la conduzione. La larghezza della banda proibita è detta energy gap e per il silicio vale 1.12eV

$$E_g = E_C - E_V = 1.12\text{eV} \quad (3.2)$$

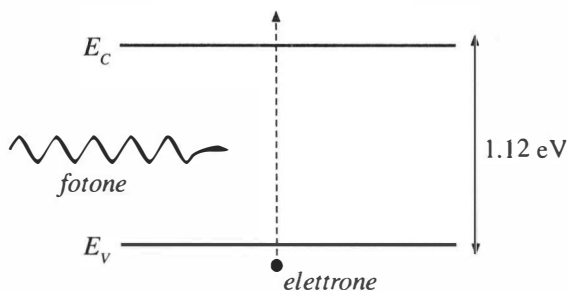


Figura 3.1 Interazione luce-semiconduttore

Illuminando un semiconduttore con un fascio di onde elettromagnetiche, se queste hanno lunghezza d'onda abbastanza piccola, i relativi fotoni saranno in grado di far saltare gli elettroni dalla banda di valenza alla banda di conduzione, questo fenomeno è detto effetto fotoelettrico e per la sua scoperta Albert Einstein ebbe il premio Nobel.

Uguagliando la (3.1) con la (3.2) si ottiene che la lunghezza d'onda massima è 1.1μm. La luce visibile è compresa tra 400nm e 700nm, pertanto quando un semiconduttore viene illuminato si generano in esso coppie elettrone-lacuna.

3.2 Il fotodiodo

Il fotodiodo è un dispositivo in grado di fornire una corrente elettrica sfruttando l'effetto fotoelettrico.

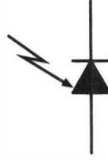


Figura 3.2 Simbolo circuitale del fotodiodo

Un diodo non polarizzato (nessuna tensione applicata) o contropolarizzato (applicando una tensione positiva al catodo) ha una regione di svuotamento tanto più ampia quanto più il diodo è contropolarizzato. In generale un fotone di lunghezza d'onda minore di $1.1\mu\text{m}$ che incide nel silicio dà origine a una coppia elettrone-lacuna.

Se tale generazione avviene nella regione di svuotamento, il campo elettrico (che è presente anche quando nessuna tensione esterna è applicata) separa i due portatori, che quindi danno origine a una corrente di deriva.

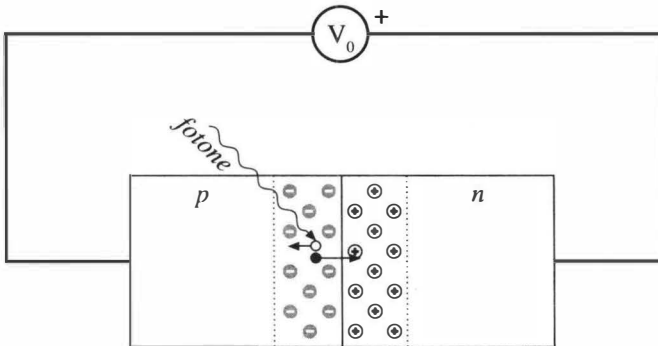


Figura 3.3 Generazione coppia elettrone-lacuna per assorbimento di un fotone

Con riferimento alla Figura 3.3, l'elettrone percorre il tratto dal punto in cui incide il fotone fino alla regione n, la lacuna percorre il tratto complementare dal punto in cui incide il fotone alla regione p.

Poiché elettroni e lacune portano la stessa carica in valore assoluto (la carica elementare dell'elettrone), complessivamente è come se una singola carica elementare percorresse l'intera zona di svuotamento. In altre parole l'assorbimento di un fotone produce un passaggio di carica nel circuito pari alla carica di un elettrone.

Se, quindi, indichiamo con ϕ il flusso di fotoni (fotoni al secondo) che incide nella regione di svuotamento, la corrente i che scorre nel circuito è pari a:

$$i = q\phi \quad (3.3)$$

con q carica elementare dell'elettrone. D'altra parte la potenza P_{opt} della radiazione luminosa è pari al flusso di fotoni per l'energia del singolo fotone:

$$P_{opt} = \phi \frac{hc}{\lambda} \quad (3.4)$$

Il rapporto i/P (A/W), detto responsività ($\mathcal{R}$), è quindi dato da:

$$\mathcal{R} = \frac{i}{P_{opt}} = \frac{q}{hc} \lambda \quad (3.5)$$

che è una relazione veramente notevole, perché ci dice che la corrente, a parità di potenza ottica, dipende linearmente dalla lunghezza d'onda mediante un fattore che contiene solo costanti fondamentali della fisica. Esprimendo $\mathcal{R}$ in A/W e la lunghezza d'onda in μm , la (3.5) può essere scritta in maniera più comoda come:

$$\mathcal{R} [\text{A/W}] = \frac{\lambda [\mu\text{m}]}{1.24} \quad (3.6)$$

La luce visibile è compresa tra $0.4\mu\text{m}$ e $0.7\mu\text{m}$, il valore medio è $0.55\mu\text{m}$ (verde). Dalla (3.6) segue che la luce visibile produce in un fotodiode circa mezzo Ampere per ogni Watt di potenza ottica.

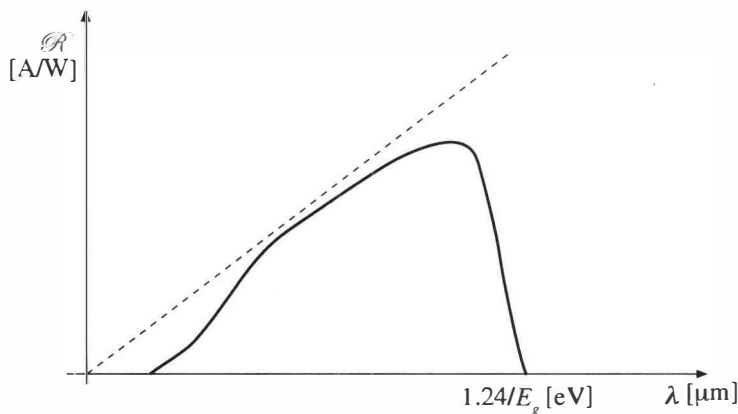


Figura 3.4 Responsività del fotodiode

La (3.6) vale ovviamente solo nel caso ideale di un fotodiodo con efficienza unitaria, più in generale vale

$$\mathcal{R}[\text{A/W}] = \eta \frac{\lambda[\mu\text{m}]}{1.24} \quad (3.7)$$

con η efficienza quantica. Si noti che l'efficienza quantica si riduce notevolmente per lunghezze d'onda prossime a 200-300nm (banda ultravioletto, UV). La ragione è che la profondità di penetrazione delle onde elettromagnetiche si riduce al diminuire della lunghezza d'onda e nella banda UV è già una frazione di millimetro. Pertanto in questa banda i fotoni vengono assorbiti prima di poter entrare nella regione di svuotamento e dare origine a una coppia elettrone-lacuna.

3.3 Sensori a matrice per fotocamere

Una delle più diffuse applicazioni dei fotodiodi è in forma di matrice come sensore d'immagine per le fotocamere. Il singolo pixel, nella sua forma basica, è un fotodiodo con un condensatore in parallelo (nella realtà non è un vero e proprio condensatore fisico, ma la capacità parassita del gate di un MOSFET).

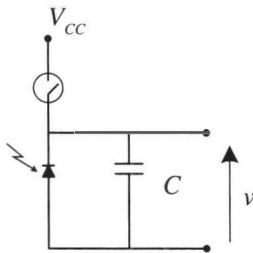


Figura 3.5 Pixel di una fotocamera

Quando la fotocamera è accesa l'interruttore è ON e i condensatori di ogni singolo pixel sono carichi alla tensione di alimentazione. Lo scatto (o l'acquisizione del frame per un filmato) consiste nello scollegare l'alimentazione alla coppia fotodiodo-capacitore. Se il pixel non è illuminato, il fotodiodo, contropolarizzato dal condensatore carico, è praticamente un circuito aperto e il condensatore non si scarica. Se, al contrario, il pixel è illuminato, il fotodiodo genera una corrente proporzionale alla potenza ottica ricevuta e scarica il condensatore.

La carica Q accumulata da un condensatore è proporzionale alla tensione ai suoi capi:

$$Q = Cv \quad (3.8)$$

Quindi

$$i = \frac{dQ}{dt} = C \frac{dv}{dt} \quad (3.9)$$

Quando il fotodiodo scarica il condensatore con una corrente

$$i = \mathcal{R} P_{opt} \quad (3.10)$$

Quindi la tensione sul condensatore dopo un tempo di esposizione Δt è pari a

$$v(\Delta t) = V_{CC} - \frac{\mathcal{R}}{C} P_{opt} \Delta t \quad (3.11)$$

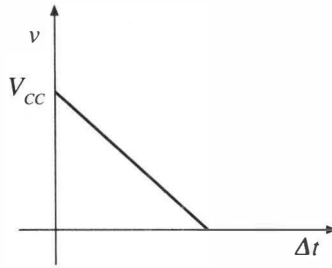


Figura 3.6 Tensione sul condensatore del pixel di una fotocamera durante il tempo di esposizione

Ovviamente il tempo di esposizione dovrà essere scelto abbastanza breve da non scaricare del tutto il condensatore, così come nella fotografia analogica il tempo di esposizione andava scelto in modo da non far saturare la pellicola.

3.4 Efficienza di conversione di una cella fotovoltaica

Il fotodiodo è l'elemento base dei pannelli solari per la produzione di energia elettrica. A questo proposito si faccia riferimento al circuito in Figura 3.7. Tracciamo le caratteristiche del diodo al variare della potenza ottica incidente. L'illuminazione del fotodiodo ha l'effetto di traslare verso il basso la caratteristica. Per la resistenza di carico vale

$$v_D = R i_R \quad (3.12)$$

ovvero

$$v_D = -R i_D \quad (3.13)$$

che è detta retta di carico. Si tratta di una retta a pendenza negativa, in generale ciò sta a significare che il circuito fornisce potenza al carico. Si noti che pur variando la pendenza della retta di carico (ovvero R) la tensione rimane intorno alla tensione di soglia del diodo: circa 0.6V.

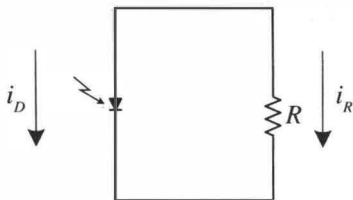


Figura 3.7 Fotodiodo e resistenza di carico

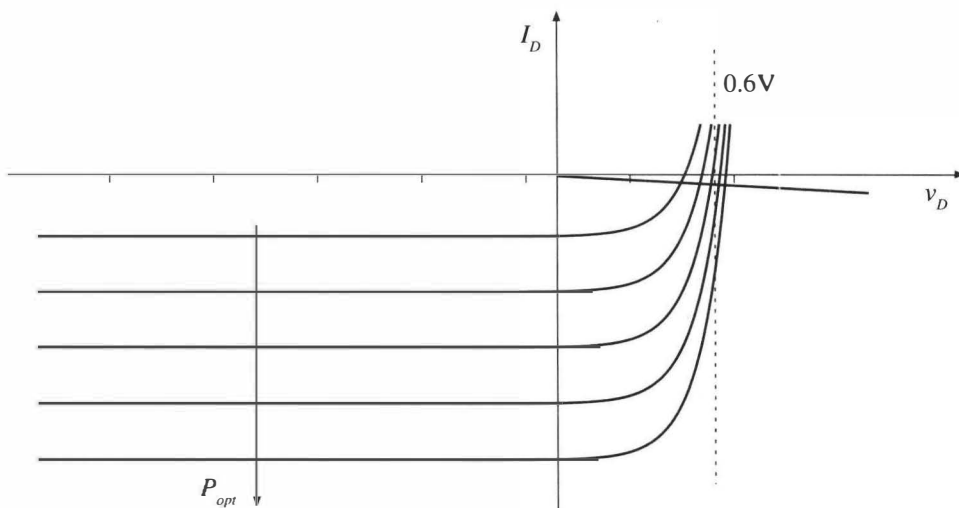


Figura 3.8 Caratteristica del fotodiodo e retta di carico

In altre parole una cella fotovoltaica è come una pila elettrica di tensione pari al potenziale di giunzione che è circa uguale alla tensione di soglia del diodo. Ciò è in accordo con l'idea intuitiva che la corrente scorre attraverso la regione di svuotamento e che la tensione è pari al potenziale di giunzione $\varphi \approx 0.6V$. La potenza massima teorica erogata da una cella al silicio è quindi

$$P_{elet} = \varphi i = (0.6V)(0.5A/W)P_{opt} \quad (3.14)$$

e l'efficienza

$$\eta = \frac{P_{elet}}{P_{opt}} \approx 0.3 \quad (3.15)$$

Nella pratica non tutti i fotoni riusciranno a raggiungere la regione di svuotamento e non tutti gli elettroni generati potranno contribuire alla corrente. I pannelli commercialmente disponibili hanno efficienza tra il 10% e il 15%.

3.5 Semiconduttori diretti e indiretti

Un fotone incidente su un semiconduttore, se ha abbastanza energia, fa saltare un elettrone dalla banda di valenza alla banda di conduzione generando una coppia elettrone-lacuna. Potremmo ragionevolmente domandarci se si verifica il processo inverso, ovvero se un elettrone quando si combina con una lacuna genera un fotone pari all'energia di gap tra banda di conduzione e banda di valenza. La risposta a questa domanda è «talvolta, ma non sempre». Purtroppo il silicio fa parte dei semiconduttori che non producono fotoni per ricombinazione elettrone-lacuna.

Per capirne il meccanismo è necessario fare una premessa. Un elettrone mobile nello spazio libero è caratterizzato da due grandezze: energia cinetica E_k e quantità di moto p :

$$E_k = \frac{1}{2}mv^2 \quad (3.16)$$

$$p = mv \quad (3.17)$$

che sono legate dalla seguente relazione

$$E_k = \frac{p^2}{2m} \quad (3.18)$$

È ragionevole supporre che l'energia di un elettrone in conduzione sia, in prima approssimazione, la somma della sua energia cinetica con l'energia potenziale E_c minima della banda di conduzione

$$E = \frac{p^2}{2m} + E_c \quad (3.19)$$

La (3.19) può essere applicata ai livelli energetici degli elettroni in conduzione in un semiconduttore. Il caso delle lacune è un po' più complesso, per il fatto che una lacuna non è una particella e non ha una massa. Tuttavia anche nel loro caso si può definire una sorta di massa virtuale m^* e scrivere per l'energia dei livelli delle lacune

$$E = -\frac{p^2}{2m^*} + E_v \quad (3.20)$$

Il segno negativo deriva dal fatto che le lacune sono positive e quindi il loro asse energetico è come se fosse capovolto rispetto al caso dell'elettrone. Rappresentando la (3.19) e la (3.20) sullo stesso grafico in funzione della quantità di moto, osserviamo che i livelli di conduzione più vicini al vertice della parabola sono i più bassi e quindi hanno maggiore probabilità di occupazione.

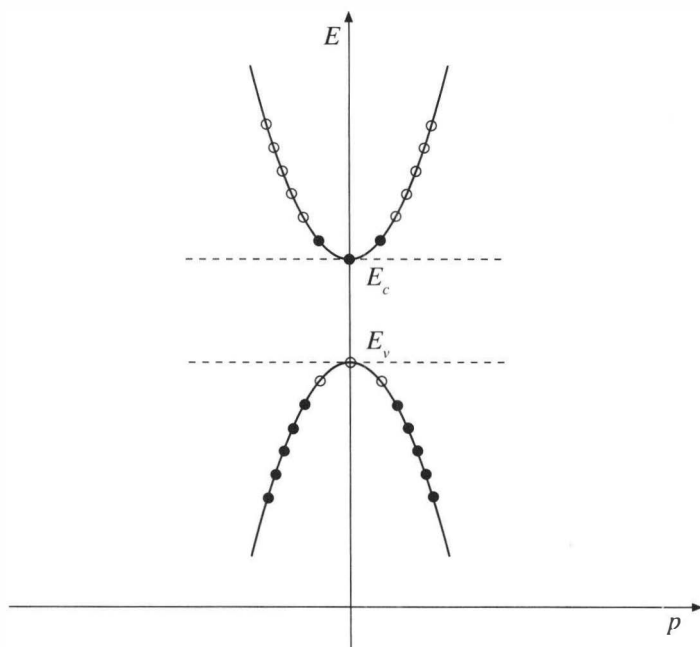


Figura 3.9 Semiconduttore *direct gap*

Anche per i livelli di valenza, i livelli bassi hanno maggiore probabilità di essere occupati, ma ciò significa che le lacune saranno nei livelli più alti, ovvero sulla sommità della parabola. In definitiva gli elettroni tenderanno a stare in un minimo relativo dei livelli energetici di conduzione in funzione della quantità di moto e le lacune in un massimo relativo dei livelli di valenza.

Putroppo in un semiconduttore i livelli energetici in funzione della quantità di moto non sono semplicemente delle parabole (d'altra parte stiamo parlando di elettroni all'interno di un reticolo cristallino e sarebbe quantomeno sorprendente che potessero essere assimilati del tutto a elettroni liberi), ma hanno forme complesse con più massimi e minimi. L'andamento parabolico è una buona approssimazione solo nei punti stazionari. Il punto $p = 0$ è sempre un punto stazionario di minimo locale per la conduzione e di massimo locale per la valenza, ma non è detto che sia di minimo assoluto per la conduzione. E in effetti nel silicio non lo è. In tal caso si parla di *indirect gap*.

Per quello che riguarda la generazione di una coppia elettrone-lacuna, non ci sono complicazioni per i semiconduttori *indirect gap*. Un elettrone eccitato da un fotone salta in conduzione a un livello di pari quantità di moto e poi ridiscende al livello più basso disponibile, perdendo o acquisendo, mediante urti con il reticolo, la quantità di moto necessaria per collocarsi nel primo stato libero in conduzione.

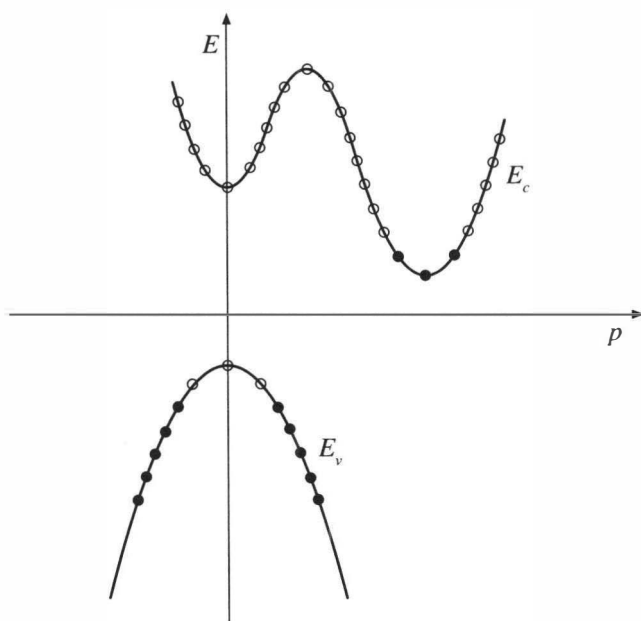


Figura 3.10 Semiconduttore *indirect gap*

Il caso della ricombinazione è invece più complesso. Gli elettroni si trovano, infatti, in maggioranza in uno stato energetico con quantità di moto diversa da zero, mentre le lacune sono per lo più in uno stato con quantità di moto nulla. Quindi una ricombinazione diretta elettrone-lacuna non è possibile, semplicemente perché l'uno è in moto e l'altra è ferma. Nella realtà elettrone e lacuna si ricombinano comunque, ma non direttamente e senza l'emissione di un singolo fotone. Si parla, infatti, di **ricombinazione non radiativa**. Tale processo di ricombinazione implica piccoli salti in banda proibita, scambi di quantità di moto con il reticolo cristallino, scambi con i livelli più profondi degli atomi (effetto Auger). L'effetto cumulativo è comunque il riscaldamento del semiconduttore. Per questo motivo il silicio, pur essendo un buon rivelatore di fotoni nel visibile, non può essere utilizzato come sorgente di fotoni nel visibile. Per emettere luce è necessario utilizzare altri semiconduttori, come ad esempio l'arseniuro di gallio (GaAs).

3.6 LED

LED sta per Light Emitting Diode ed è sostanzialmente un diodo realizzato in un semiconduttore *direct gap* in una capsula trasparente. I semiconduttori comunemente usati per i LED sono: GaAs (arseniuro di gallio), GaP (fosfuro di gallio), GaAsP (fosfuro arseniuro di gallio), SiC (carburo di silicio) e GaInN (nitruro di gallio e indio).



Figura 3.11 Simbolo circuitale del LED

Polarizzando direttamente un LED, ovvero applicando una tensione positiva all'anodo, si riduce la barriera di potenziale tra zona n e zona p. Gli elettroni dalla zona n si diffondono nella zona p, così come le lacune dalla zona p si diffondono nella zona n. Elettroni e lacune ricombinandosi generano fotoni di energia pari al gap tra banda di conduzione e banda di valenza.

Si noti che la ricombinazione e quindi la generazione di fotoni non avviene nella regione di svuotamento, ma nella zona n o p. D'altra parte una delle due è tipicamente molto più drogata dell'altra. Nel caso in cui la zona n è molto più drogata della zona p, la corrente risulta principalmente dovuta agli elettroni che si ricombinano nella zona p.

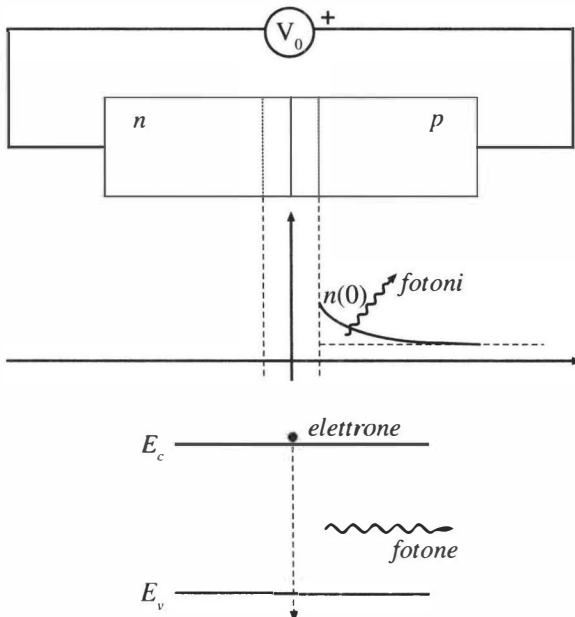


Figura 3.12 Emissione di fotoni nel LED

Idealmente, per ogni carica elementare che scorre nel circuito di polarizzazione del diodo, si genera un fotone

$$\phi = \frac{i}{q} \quad (3.21)$$

In realtà, come abbiamo già visto, solo una frazione delle ricombinazioni è di natura radiativa. Nei semiconduttori *direct gap* questa frazione può essere abbastanza significativa (ad esempio per il GaAs è circa 0.5), nel silicio è minore di 10^{-5} . La (3.21) deve essere quindi completata con un termine a moltiplicare minore di 1 detto efficienza intrinseca (η_i).

$$\phi = \eta_i \frac{i}{q} \quad (3.22)$$

D'altra parte la potenza P della radiazione luminosa è pari al flusso di fotoni per l'energia del singolo fotone:

$$P = \phi \frac{hc}{\lambda} \quad (3.23)$$

Il rapporto P/i (W/A), detto responsività ($\mathcal{R}$), risulta quindi pari a

$$\mathcal{R} = \frac{P}{i} = \eta_i \frac{hc}{q} \frac{1}{\lambda} \quad (3.24)$$

La lunghezza d'onda emessa dal LED dipende dall'*energy gap* del particolare semiconduttore impiegato

$$\lambda = \frac{hc}{E_g} \quad (3.25)$$

ovvero:

$$\lambda[\mu\text{m}] = \frac{1.24}{E_g[\text{eV}]} \quad (3.26)$$

Un LED è una sorgente a spettro stretto, ma non è del tutto monocromatica a causa dell'agitazione termica che produce un allargamento spettrale pari a

$$\Delta E = kT \quad (3.27)$$

con k costante di Boltzmann. Tenendo conto della relazione di Plank, si ottiene la variazione di lunghezza d'onda. Ad esempio a temperatura ambiente ($kT = 26\text{meV}$), per un LED a $0.6\mu\text{m}$, si ottiene 13nm .

La corrente tipica che è necessario far scorrere in un LED affinché si illumini è $10\text{-}20\text{mA}$. La Figura 3.13 mostra il circuito di polarizzazione di un LED, la resistenza in serie ha lo scopo di limitare la corrente al valore di progetto.

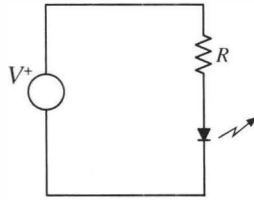


Figura 3.13 Circuito di polarizzazione di un LED

Infatti

$$I = \frac{V^+ - V_F}{R} \quad (3.28)$$

con V_F tensione del LED in polarizzazione diretta. V_F , come per il diodo al silicio, è circa costante al variare della corrente e dipendente solo dal tipo di semiconduttore. In linea di massima al diminuire della lunghezza d'onda (e quindi all'aumentare di E_g) la tensione in polarizzazione diretta aumenta, ma la relazione non è così diretta perché il valore di soglia dipende da come si sommano le correnti di diffusione e di ricombinazione nella regione di svuotamento dello specifico semiconduttore.

3.7 Dall'effetto fotoelettrico all'illuminazione a LED

La prima osservazione del cosiddetto effetto “fotoelettrico” è del 1880, a opera di Heinrich Rudolf Hertz. Lo scienziato tedesco notò che la scarica tra due elettrodi metallici era favorita con luce ultravioletta. Studi sperimentali più sistematici dimostrarono che in generale la luce ultravioletta provoca l'emissione di elettroni nei metalli. Nell'ambito dell'elettromagnetismo classico questo fenomeno era del tutto inspiegabile. Solo dopo la scoperta della quantizzazione dell'energia, Albert Einstein trovò una spiegazione soddisfacente del fatto che un'onda elettromagnetica anche di bassa intensità, se di lunghezza d'onda opportuna, poteva produrre elettroni liberi nel materiale in cui incideva. Fu proprio per questa scoperta che Einstein ottenne il Premio Nobel.

Il fotodiodo al silicio fu inventato nel 1939 da Russell Ohl nei laboratori della Bell. Ohl, quando scoprì l'effetto rettificante della giunzione pn, notò anche che essa era in grado di produrre una corrente inversa se illuminata.

L'invenzione del LED ha invece richiesto molto più tempo. Solo nel 1963 Nick Holonyak Jr., presso la General Electric, realizza la prima giunzione pn in grado di emettere luce visibile, dalla quale derivano gli schermi a LED e oggi anche l'illuminazione a LED.

3.8 Riepilogo

- Un fotodiodo è una giunzione pn (di solito di silicio) non polarizzata o polarizzata inversamente in cui la luce può arrivare (attraverso un'apposita finestra) nella zona di svuotamento.
- Un fotodiodo (ideale) illuminato da una potenza ottica P (in Watt) produce una corrente (inversa, ovvero che va dalla zona n alla zona p) I (in Ampere) data da

$$I = \frac{\lambda [\mu\text{m}]}{1.24} P$$

dove λ è la lunghezza d'onda espressa in μm della luce incidente.

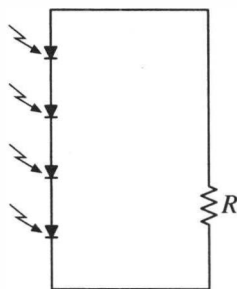
- Un LED è una giunzione pn (non di silicio perché è un semiconduttore *indirect gap*, ma di solito di GaAs che è un semiconduttore *direct gap*) polarizzata direttamente.
- Un LED (ideale) in cui scorre una corrente I (in Ampere) emette una potenza ottica (in Watt) data da

$$P = \frac{1.24}{\lambda [\mu\text{m}]} I$$

dove λ è la lunghezza d'onda espressa in μm della luce emessa.

3.9 Esercizi svolti

1) Quattro celle fotovoltaiche sono collegate in serie come in figura. La superficie di ciascuna cella è 25cm^2 . L'irradiazione è 300W/m^2 . La responsivity è 0.25A/W . Calcolare la corrente generata. Calcolare la tensione e la potenza sul carico. Cosa succede se uno dei 4 fotodiodi è in ombra?



Poiché sono collegati in serie la corrente complessiva è pari alla corrente generata dalla singola cella.

$$I = (0.25 \text{ A/W})(300 \text{ W/m}^2)(25 \times 10^{-4} \text{ m}^2) = 0.19 \text{ A}$$

La tensione sul carico è la somma delle 4 tensioni delle celle:

$$V = 4 \times 0.6 = 2.4 \text{ V}$$

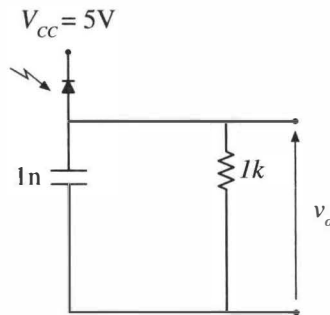
La potenza è quindi pari a 0.46W.

Se una delle quattro celle è in ombra, ovvero non produce corrente, l'intera serie non produce corrente. La corrente della serie è sempre pari a quella della cella meno efficiente o meno irradiata.

2) Il consumo medio su base annua di una famiglia è 300W. La potenza media dell'irradiazione del sole alle nostre latitudini su base annua è 300 W/m^2 , si calcoli la dimensione minima di superficie a pannelli solari per soddisfare il consumo di una famiglia.

Considerando un'efficienza del 10%, la potenza elettrica media per unità di superficie è 30 W/m^2 . Quindi per soddisfare un fabbisogno medio di 300W è necessaria una superficie coperta di pannelli solari pari a 10 m^2 .

***3) Con riferimento alla figura per $t < 0$ il condensatore è scarico e il fotodiodo è al buio. All'istante $t = 0$ si apre un otturatore il fotodiodo è illuminato con $20 \mu\text{W}$ ottici. Dopo 10ms l'otturatore è chiuso e il fotodiodo è al buio. Si tracci l'andamento nel tempo della tensione v_o .**



La responsività di un fotodiodo è 0.5 A/W , quindi quando illuminato fornisce una corrente $I = 10 \mu\text{A}$ che si ripartisce tra corrente attraverso il condensatore i_C e corrente attraverso la resistenza i_R :

$$I = i_C + i_R$$

ovvero

$$I = C \frac{dv_o}{dt} + \frac{v_o}{R}$$

che può essere riscritta come

$$\tau \frac{dv_o}{dt} + v_o = IR$$

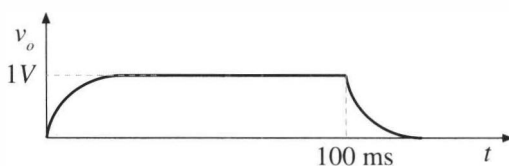
con $\tau = RC$.

Questa è un'equazione differenziale lineare che ha soluzione

$$v_o = IR \left(1 - e^{-\frac{t}{\tau}} \right) = (1V) \left(1 - e^{-\frac{t}{1\mu s}} \right)$$

Dopo qualche centinaio di μs , $v_o \approx 1V$. Dopo 10ms il capacitore il fotodiodo non fornisce più corrente e il capacitore si scarica attraverso la resistenza R . L'andamento della scarica è (vedi appendice A):

$$v_o = e^{-\frac{t - 10ms}{1\mu s}}$$



3.10 Domande di verifica a scelta multipla

- Per ogni fotone incidente nella zona attiva di un fotodiodo
 - si creano due cariche mobili
 - si creano un elettrone e uno ione fisso
 - si creano una lacuna e uno ione fisso
 - nel circuito scorre una carica pari a due volte la carica elementare dell'elettrone
- A parità di potenza ottica, all'aumentare della lunghezza d'onda della luce incidente
 - scorre una corrente maggiore, in quanto i singoli fotoni hanno maggiore energia
 - scorre una corrente maggiore, in quanto i fotoni sono più numerosi
 - scorre una corrente minore, in quanto i singoli fotoni hanno minore energia
 - scorre la stessa corrente, in quanto la corrente dipende solo dalla potenza ottica
- In un fotodiodo
 - i fotoni generano coppie elettrone-lacuna solo nella zona esterna alla regione di svuotamento
 - i fotoni possono generare coppie elettrone-lacuna solo nella regione di svuotamento dove ci sono ioni negativi
 - i fotoni generano coppie elettrone-lacuna solo nella regione di svuotamento
 - i fotoni generano coppie elettrone-lacuna ovunque incidano nel semiconduttore

4. Un fotodiodo di silicio non è sensibile a fotoni di lunghezze d'onda maggiori di 1100 nm
 - a) perché non penetrano abbastanza in profondità nel semiconduttore
 - b) perché sono riflessi dalla superficie del semiconduttore
 - c) perché non hanno abbastanza energia per creare una coppia elettrone-lacuna
 - d) perché la responsivity è inversamente proporzionale alla lunghezza d'onda
5. Un fotodiodo di silicio non è sensibile a fotoni di lunghezze d'onda minori di 200 nm
 - a) perché non penetrano abbastanza in profondità nel semiconduttore
 - b) perché non hanno abbastanza energia per creare una coppia elettrone-lacuna
 - c) perché sono riflessi dalla superficie del semiconduttore
 - d) perché la responsivity è direttamente proporzionale alla lunghezza d'onda
6. In un fotodiodo di silicio la responsivity all'aumentare della lunghezza d'onda
 - a) aumenta, perché la stessa potenza ottica genera una corrente elettrica minore
 - b) aumenta, perché a parità di potenza ottica il numero di fotoni è maggiore
 - c) diminuisce, perché i fotoni diminuiscono la loro energia all'aumentare della lunghezza d'onda
 - d) diminuisce, perché i fotoni aumentano la loro energia all'aumentare della lunghezza d'onda
7. In un LED
 - a) i fotoni si generano per effetto tunnel nella regione di svuotamento
 - b) i fotoni si generano per ricombinazione nella zona esterna alla regione di svuotamento
 - c) la zona dove si generano fotoni diminuisce all'aumentare del coefficiente di diffusione dei portatori
 - d) i fotoni si generano per ricombinazione nella regione di svuotamento
8. Una sola delle seguenti affermazioni è vera.
 - a) L'efficienza di un pannello solare è prossima al 100%, in quanto ogni fotone incidente produce una corrente corrispondente alla carica di un elettrone
 - b) L'efficienza di un pannello solare dipende dall'insolazione e dalla latitudine
 - c) Una centrale a gas per la produzione di energia elettrica ha efficienza molto minore di un campo solare
 - d) L'efficienza di un pannello solare è tra il 10% e il 20%
9. In un fotodiodo, a parità di potenza ottica incidente
 - a) la corrente è proporzionale all'inverso della frequenza della radiazione ottica incidente
 - b) la corrente è proporzionale al quadrato della lunghezza d'onda della radiazione incidente
 - c) la corrente dipende solo dalla potenza della radiazione incidente e non dalla sua lunghezza d'onda
 - d) la corrente è massima quando la radiazione incidente ha la massima frequenza compatibile con le caratteristiche di riflettanza del materiale
10. Un fotodiodo di silicio
 - a) è in grado di rivelare fotoni di lunghezza d'onda pari a 10mm
 - b) è in grado di rivelare fotoni di lunghezza d'onda pari a 0.5mm
 - c) non può rivelare fotoni, in quanto è semiconduttore di tipo *indirect gap*
 - d) è in grado di rivelare qualunque tipo di fotone

3.11 Domande di verifica a risposta aperta

1. Ricavare la formula della responsività di un fotodiodo.
2. Ricavare la formula della responsività di un LED.
3. Illustrare la differenza tra semiconduttori diretti e indiretti.

3.12 Risposte alle domande a scelta multipla

1 a	5 a	9 a
2 b	6 b	10 b
3 d	7 b	
4 c	8 d	

4

Il transistor a effetto campo

In questo capitolo si tratta il transistor a effetto campo (FET). Nell'affrontare gli aspetti più legati ai principi fisici si fornisce prima un'idea intuitiva e poi una trattazione rigorosa, lasciando al docente la scelta del livello di approfondimento che vuole dare alla materia. Con questo approccio sono trattati il concetto di soglia, la formazione del canale, lo strozzamento, l'effetto body. Il processo di fabbricazione del MOSFET è dettagliato passo-passo con riferimento alla tecnologia 45nm. Si descrivono i differenti tipi di MOSFET: ad arricchimento, a svuotamento, a tre e a quattro terminali. Infine, si discutono le applicazioni circuitali: interruttori, reti di polarizzazione, carichi attivi e specchi di corrente. Il JFET è trattato a parte seguendo lo stesso approccio del MOSFET. Il capitolo si conclude con un paragrafo storico sull'invenzione del transistor.

4.1 Il transistor

L'origine del nome transistor è, nella migliore tradizione filologica, oggetto di dispute. Secondo alcuni deriva dalla fantasiosa contrazione dei termini TRANS-conduttance varISTOR: letteralmente variatore di transconduttanza. Secondo altri è una contrazione di TRANSfer resISTOR: trasferitore di resistenza. Qualunque sia la versione corretta è un nome che dice poco o nulla anche a un ingegnere elettronico, ma negli anni cinquanta suonava moderno e futuribile ed ebbe subito uno straordinario successo.

Probabilmente sarebbe stato più appropriato il termine “valvola”, già in uso in Inghilterra (e anche in Italia) per indicare il progenitore del transistor: il tubo a vuoto o valvola termoionica.

In effetti, il modo migliore per visualizzare il principio di funzionamento di un transistor è proprio pensare a una valvola idraulica. Un transistor è un dispositivo in grado di regolare la corrente che scorre in una maglia (maglia di uscita), mediante la tensione applicata a un'altra maglia (maglia di ingresso).

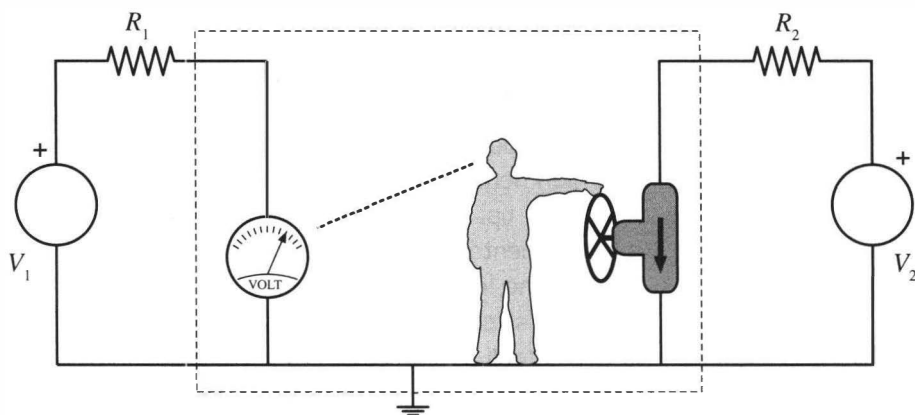


Figura 4.1 Il transistor

Il transistor, così come una valvola idraulica, può operare in due modi: 1) come interruttore; 2) come regolatore di flusso.

Un transistor usato come interruttore è un dispositivo logico elementare a due stati: ON e OFF. Su questa base si possono realizzare famiglie logiche, memorie e l'intera elettronica digitale. Un transistor usato come modulatore di corrente è a tutti gli effetti un amplificatore. In particolare, potendo controllare con una tensione della maglia di ingresso la corrente che scorre su una resistenza nella maglia di uscita, è possibile ottenere su tale resistenza una caduta di tensione molto più grande della tensione di controllo. In tal caso si ha un amplificatore tensione-tensione.

4.2 Tipologia dei transistor a effetto campo

Il transistor a effetto campo (FET: *Field Effect Transistor*) è oggi il transistor più largamente impiegato. Ne esistono di molti tipi diversi e ciò rende il FET particolarmente versatile nelle applicazioni.

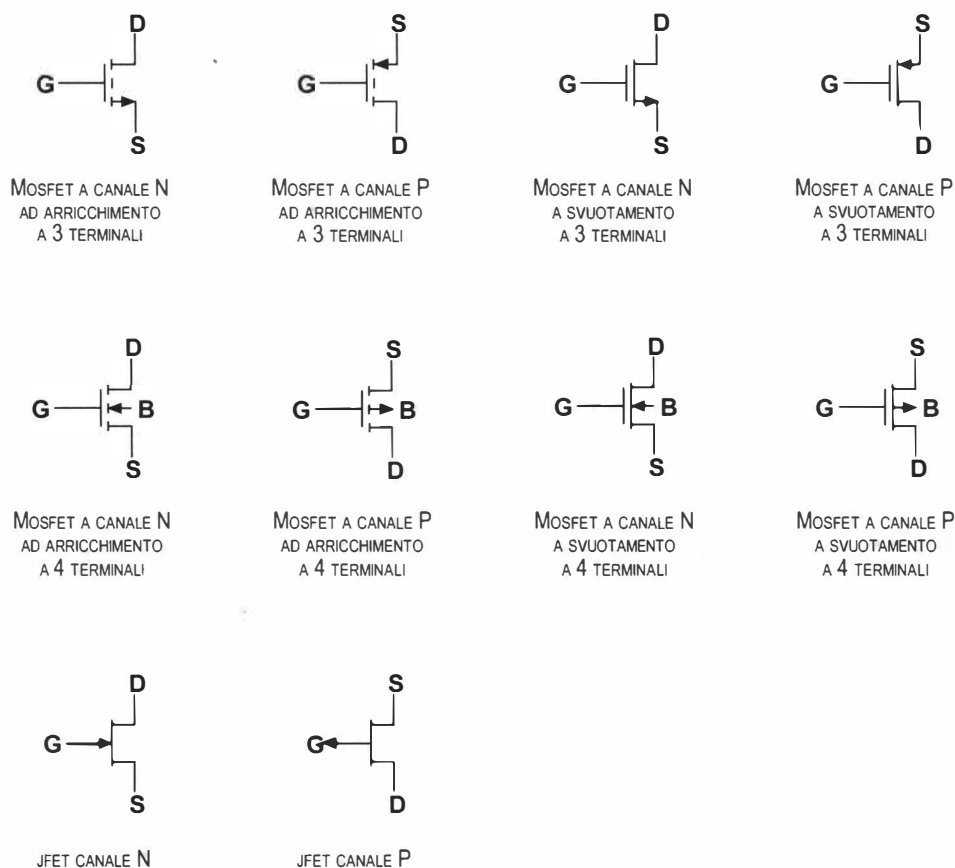


Figura 4.2 Simboli circuitali dei FET

Con riferimento ai simboli in Figura 4.2, la corrente scorre dall'alto verso il basso e la tensione di controllo è applicata tra i terminali G e S.

Innanzitutto si distingue tra MOSFET (Metal Oxide Silicon FET) e JFET (Junction FET). I MOSFET possono essere a canale n oppure a canale p, inoltre possono essere distinti in MOSFET ad arricchimento (*enhancement*) e a svuotamento (*depletion*). Infine i MOSFET possono essere a 3 o 4 terminali. Poiché tutte le combinazioni sono possibili, i tipi diversi di MOSFET sono otto. I JFET possono essere a canale n e canale p e quindi ne esistono solo due tipi differenti.

4.3 Il capacitore MOS

I MOSFET si basano essenzialmente sulle proprietà del capacitore MOS (*Metal Oxide Silicon*). Il MOS è un sandwich di tre strati sovrapposti: metallo, ossido di silicio (SiO_2) e silicio di tipo p o di tipo n. L'ossido è un isolante che non permette il passaggio di cariche elettriche tra metallo e semiconduttore. Consideriamo il caso in cui il silicio sia di tipo p.

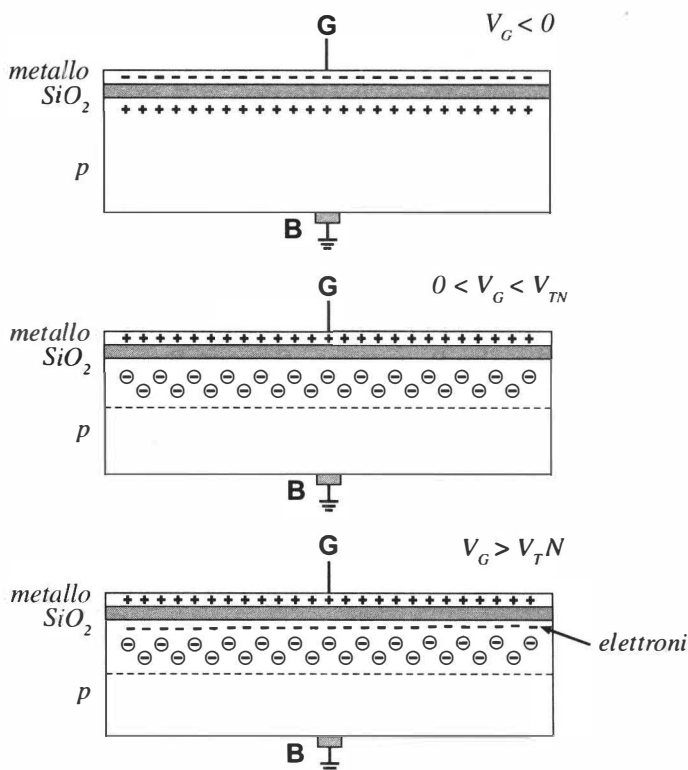


Figura 4.3 Capacitore MOS

Sul metallo è applicato un reoforo detto *gate* (G). Il semiconduttore è collegato a un secondo reoforo detto *body* (B). Applicando una tensione negativa tra gate e body, il gate si carica negativamente e attrae lacune mobili dal semiconduttore verso l'ossido. Queste non potendo attraversare lo strato di isolante semplicemente si accumulano in prossimità dell'ossido.

Quando invece sul gate è applicata una tensione positiva, le cariche positive sul metallo allontanano le lacune dalla zona in prossimità dell'ossido, creando una regione di svuotamento di cariche fisse negative.

Fin qui, nulla di particolarmente interessante. Quando però la tensione (positiva) sul gate supera una data tensione di soglia (V_{TN}), un po' sorprendentemente appaiono delle cariche mobili negative (elettroni) in prossimità dell'ossido. Poiché queste cariche sono mobili si dice che sotto l'ossido si è creato un canale nel quale può scorrere una corrente.

Vedremo nel seguito che questo canale ha un ruolo fondamentale nel funzionamento del MOSFET. Prima però dobbiamo soffermarci sulla ragione fisica per cui si crea il canale.

Sicuramente, quando eravate bambini, avrete tutti giocato sulla spiaggia con la paletta e vi sarete accorti che facendo una buca nella sabbia a un certo punto questa si riempie d'acqua. Ecco, questo è il principio alla base della creazione del canale: il livello della sabbia è l'energia di conduzione, il livello del mare è l'energia di Fermi.

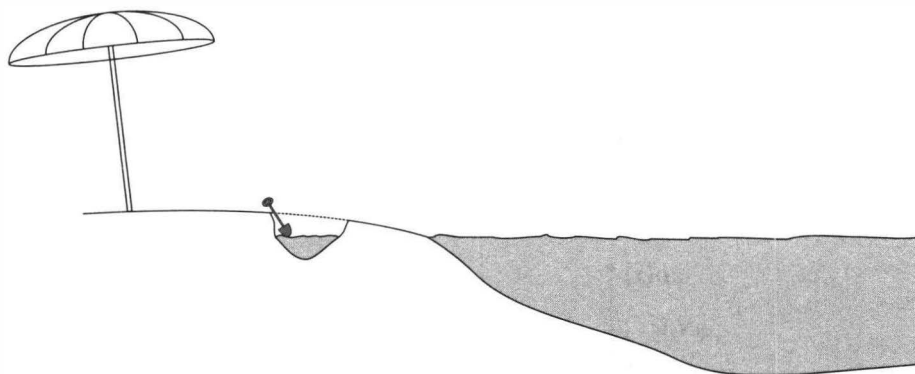


Figura 4.4 Tensione di soglia

Prendiamo in considerazione il modello a bande del silicio mostrato in Figura 4.5. Un accumulo di cariche positive sul metallo (isolato dal semiconduttore) produce un campo elettrico positivo nel verso dal metallo al semiconduttore. I livelli energetici degli elettroni nel silicio sono, di fatto, dei potenziali elettrici moltiplicati per la carica (negativa) dell'elettrone e il gradiente di un potenziale elettrico è il campo elettrico. Ne consegue che il campo elettrico prodotto dalle cariche incurva i livelli energetici degli elettroni, che pertanto in prossimità dell'ossido saranno più bassi. All'aumentare del campo elettrico i livelli energetici più bassi della conduzione si incurveranno così tanto da scendere sotto il livello di Fermi. L'energia di Fermi può essere vista come il livello dell'acqua: ciò che sta sotto è sommerso, ovvero i livelli energetici sotto il livello di Fermi sono occupati dagli elettroni. La ragione fisica per cui, applicando un potenziale positivo sufficientemente elevato, in prossimità dell'ossido appaiono elettroni liberi è che le cariche positive accumulate sul metallo con le corrispondenti

cariche negative nella regione di svuotamento abbassano così tanto i livelli della conduzione che questi sono sommersi dagli elettroni del semiconduttore.

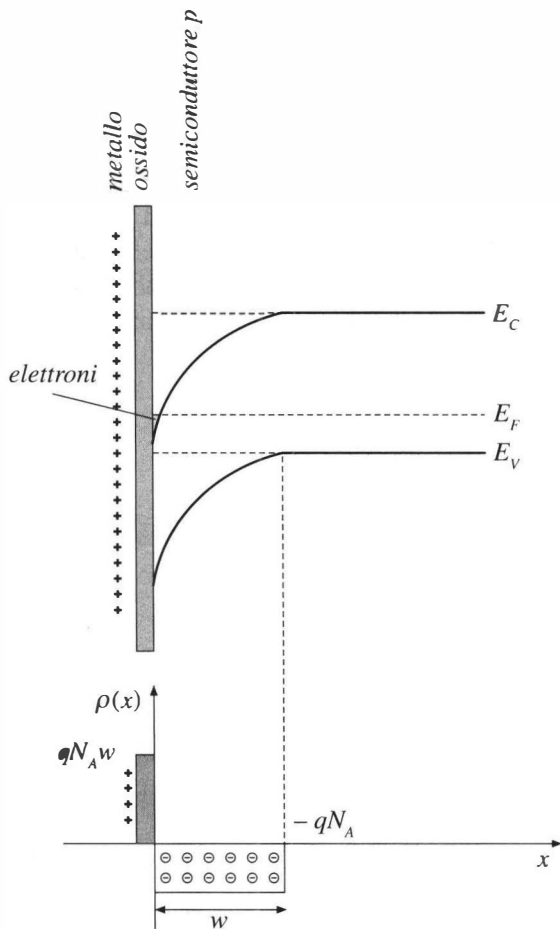


Figura 4.5 Tensione di soglia

Ci si potrebbe forse domandare da dove vengano questi elettroni. La risposta è: dall'agitazione termica. Ovunque nel semiconduttore per agitazione termica si creano continuamente coppie elettrone-lacuna. In prossimità dell'ossido gli elettroni rimangono perché attratti dal campo elettrico, le lacune sono spinte via dallo stesso campo elettrico verso la zona p.

Si noti, inoltre, che la soglia non è così netta come si potrebbe pensare, ma piuttosto sfumata: un po' come per la soglia di un diodo. La presenza di elettroni nel canale aumenta esponenzialmente (come per il diodo) all'aumentare della tensione, finché questa non raggiunge un valore tale da portare il livello della conduzione prossimo all'energia di Fermi. Tra l'altro, molti moderni dispositivi a bassa tensione per ridurre i consumi sfruttano proprio la piccola conduzione del canale anche in condizione sottosoglia.

*4.4 Calcolo della tensione di soglia

Come abbiamo già detto, il valore della soglia non è così inequivocabilmente identificabile, è pertanto necessario definirlo rigorosamente prima di procedere al calcolo.

Definiamo, quindi, **tensione di soglia** il valore per cui si ha la cosiddetta “inversione”, ovvero il valore della tensione alla quale al di sotto dell’ossido un semiconduttore p con una data concentrazione di portatori p si comporta come un semiconduttore n di pari concentrazione di portatori.

Definiamo φ_F la differenza di potenziale tra l’energia di Fermi del semiconduttore p e l’energia di Fermi del semiconduttore intrinseco. Se tutti i livelli si abbassano di una quantità pari a $2\varphi_F$ il semiconduttore da p diventa n, nel senso che se il semiconduttore p era caratterizzato da una data concentrazione p di lacune, quando i livelli si abbassano di $2\varphi_F$ il semiconduttore sarà caratterizzato da una concentrazione n di elettroni pari a p.

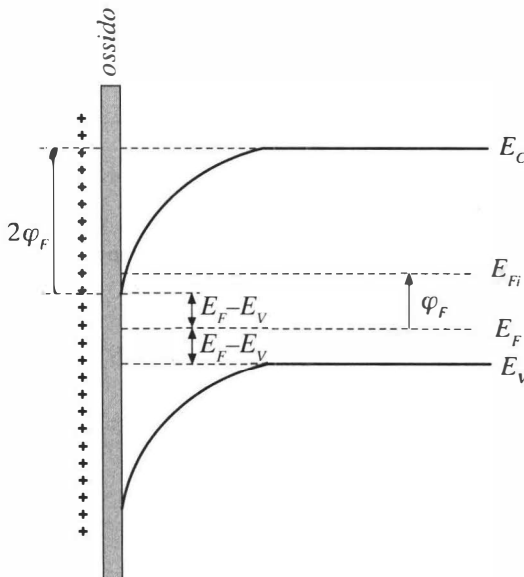


Figura 4.6 Inversione

Prendiamo ora in considerazione un MOS a cui sia applicata una tensione maggiore di zero, ma minore della soglia. La densità di carica, il campo elettrico e il potenziale sono quelli mostrati in Figura 4.7.

Nella distribuzione delle cariche si noti che, per la neutralità del MOS, la carica a sinistra dell’ossido, tutta concentrata in uno strato infinitesimo, è pari alla carica spaziale nella regione di svuotamento.

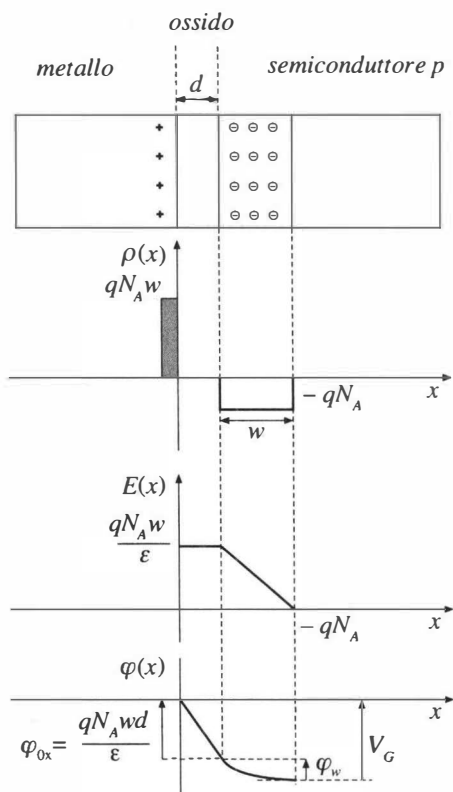


Figura 4.7 Zona di svuotamento, distribuzione di cariche, campo elettrico e potenziale in un condensatore MOS

Per il calcolo del campo elettrico si applica la legge di Gauss al parallelepipedo in Figura 4.8.

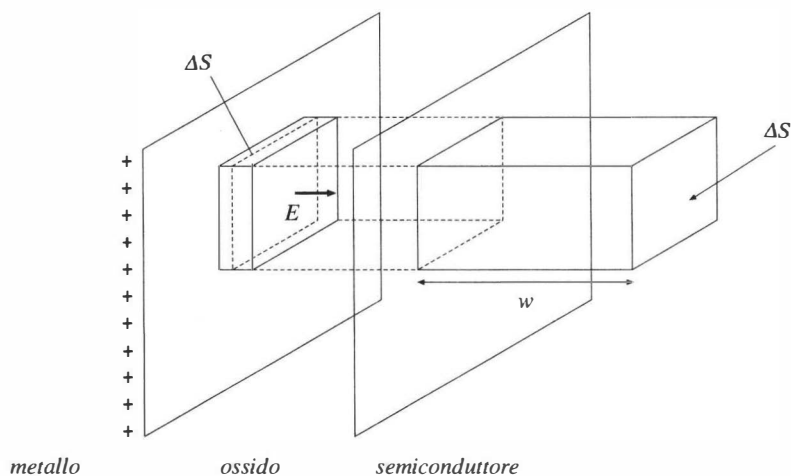


Figura 4.8 Calcolo del campo elettrico nell'ossido

Scriviamo la legge di Gauss:

$$\int_s \vec{E} \cdot \vec{n} dS = \frac{Q}{\epsilon} \quad (4.1)$$

dove $\vec{E}$ è il campo elettrico, $\vec{n}$ è il versore normale alla superficie, Q è la carica elettrica all'interno della superficie S e ϵ è la costante dielettrica. Il campo è ortogonale alla faccia interna del parallelepipedo e le cariche positive sono concentrate sull'interfaccia metallo-ossido, quindi

$$E \Delta S = \frac{Q}{\epsilon} \quad (4.2)$$

ΔS è l'area della faccia del parallelepipedo come indicato in Figura 4.8. D'altra parte la carica positiva sul metallo deve essere equivalente alla carica negativa nella regione di svuotamento corrispondente, quindi

$$Q = N_A \Delta S w \quad (4.3)$$

Sostituendo nella precedente equazione si ottiene

$$E = \frac{q N_A w}{\epsilon} \quad (4.4)$$

Poiché il campo elettrico attraverso l'ossido è costante, φ_{ox} è pari a:

$$\varphi_{ox} = \frac{q N_A w}{\epsilon} d \quad (4.5)$$

con d spessore dell'ossido.

La tensione di soglia V_{TN} sarà data dalla somma di quanto i livelli energetici devono essere incurvati per avere l'inversione ($2\varphi_F$) e il potenziale attraverso l'ossido (φ_{ox})

$$V_{TN} = 2\varphi_F + \varphi_{ox} \quad (4.6)$$

con

$$\varphi_{ox} = \frac{q N_A w}{\epsilon} d \quad (4.7)$$

D'altra parte lo spessore della zona di svuotamento è dato da

$$w = \sqrt{\frac{2\epsilon}{q} \varphi_w \frac{1}{N_A}} \approx \sqrt{\frac{2\epsilon}{q N_A} (2\varphi_F)} \quad (4.8)$$

Sostituendo la (4.8) nella (4.7) e questa nella (4.6) si ottiene

$$V_{TN} = (2\varphi_F) + d \sqrt{\frac{2qN_A}{\epsilon}} (2\varphi_F) \quad (4.9)$$

Questa equazione ci dice che la tensione di soglia può essere controllata in sede di progetto del MOS mediante lo spessore dell'ossido e la concentrazione dei droganti. In generale meno il silicio è drogato e minore è lo spessore dell'ossido, più la soglia è bassa.

Di solito una soglia bassa è un vantaggio soprattutto nei dispositivi digitali, perché consente di utilizzare tensioni di alimentazione minori e quindi avere consumi più bassi. Questa è la ragione per cui i drogaggi dei MOSFET sono piuttosto leggeri rispetto ai BJT.

4.5 MOSFET a canale n ad arricchimento

Possiamo ora analizzare la struttura completa di un MOSFET a canale n ad arricchimento. Un capacitore MOS è realizzato su un substrato p. In realtà, solo nei primi MOSFET c'era un vero capacitore MOS metallo-ossido-silicio: ormai da decenni il metallo è stato sostituito con il polisilicio fortemente drogato n, ovvero da silicio amorfo (con cristalli molto piccoli orientati casualmente). Questo materiale ha caratteristiche molto simili a un metallo, pur mantenendo la compatibilità con i processi di microlitografia del silicio. Se volessimo essere precisi dovremmo chiamare questo transistor "POSFET", ma come in molti altri casi i termini, una volta entrati nell'uso, difficilmente si modificano, anche se imprecisi e fuorvianti.

A lati opposti del capacitore MOS vengono realizzate due piazzole n molto drogate sulle quali sono fissati due reofori detti *source* (S) e *drain* (D). La struttura è perfettamente simmetrica e source e drain sono intercambiabili, tuttavia uno dei due terminali, nella configurazione a tre terminali, è collegato al body (B) del MOS e questo terminale diviene univocamente il source.

Fisicamente un MOSFET ha una struttura molto più larga che lunga, ovvero la distanza tra le piazzole source e drain è più corta della larghezza nell'altra direzione sul piano del semiconduttore. Come vedremo nel seguito il gate è il terminale di controllo, un po' come la manopola di una valvola idraulica. La corrente scorre dal drain verso il source.

Vediamo ora in dettaglio come funziona. Le piazzole n+ con il substrato p costituiscono delle giunzioni pn e quindi ci sarà una regione di svuotamento intorno alle piazzole che si estende principalmente nel semiconduttore p che è meno drogato della piazzola n+.

Applichiamo al gate una tensione negativa o uguale a zero. Sebbene ci siano cariche mobili sotto il gate, applicando una tensione tra drain e source non scorre alcuna corrente. Questo perché le due giunzioni pn in corrispondenza delle due piazzole n+ hanno verso opposto e, mentre quella tra substrato e source è polarizzata direttamente, la giunzione tra drain e substrato è contropolarizzata e quindi impedisce il flusso della corrente.

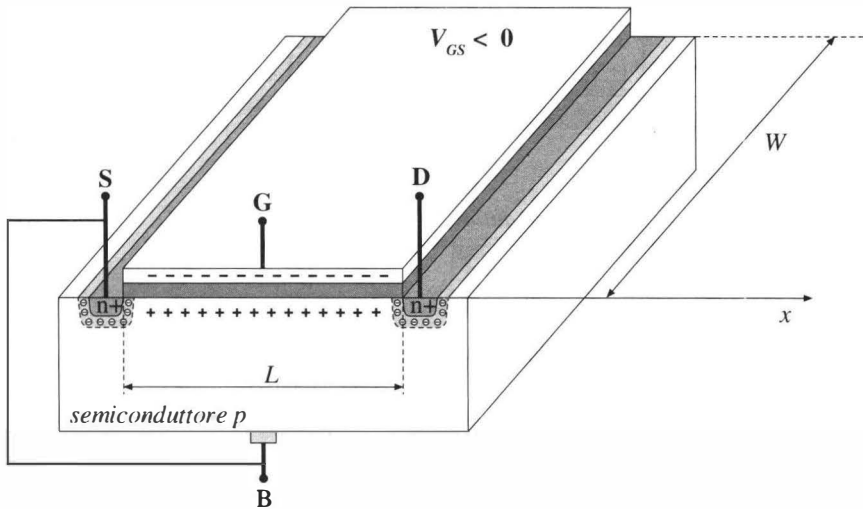


Figura 4.9 MOSFET a canale n

Applicando una tensione positiva ma minore della soglia tra gate e source si crea una regione di svuotamento sotto il gate e quindi a maggior ragione non potrà scorrere corrente tra drain e source.

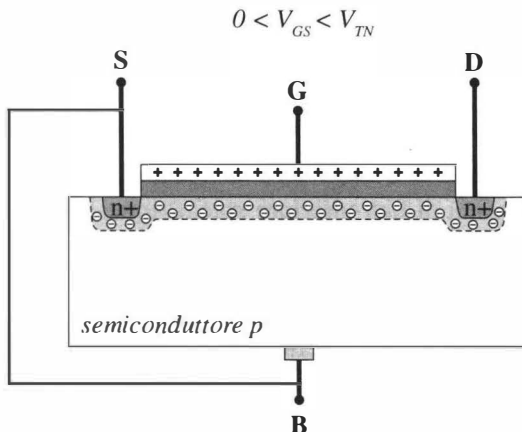


Figura 4.10 MOSFET a canale n

La cosa interessante succede quando la tensione tra gate e source supera la tensione di soglia. In questo caso si crea un canale di elettroni liberi lungo tutto l'ossido sotto il gate e se il gate è realizzato in modo da lambire la piazzola n+, viene a crearsi un unico canale di elettroni liberi da piazzola a piazzola e, quindi, applicando una tensione tra drain e source potrà scorrere una corrente.

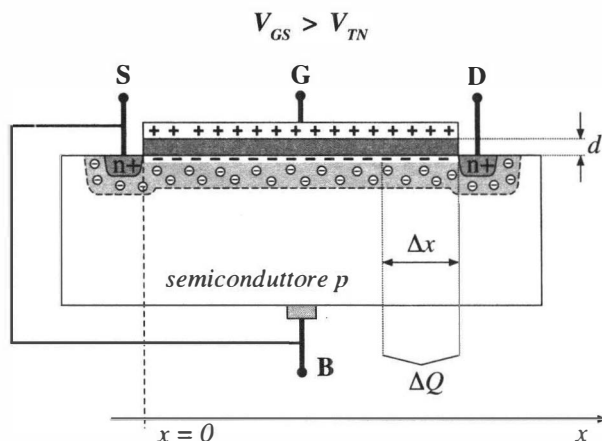


Figura 4.11 Formazione del canale

Vediamo come quantificare questa corrente.

Fintanto che la tensione sul gate è minore della soglia, un aumento della tensione di gate produce un allargamento della regione di svuotamento (e un conseguente incremento di ioni fissi non compensati). Ma appena la v_{GS} supera la soglia, sotto l'ossido si accumula una carica mobile Q (elettroni) pari a

$$Q = -C_{ox}(v_{GS} - V_{TN}) \quad (4.10)$$

dove

$$C_{ox} = \epsilon \frac{WL}{d} \quad (4.11)$$

è la capacità dell'ossido, con L lunghezza del canale e W larghezza. La (4.10) è ragionevole, ma necessita di una ulteriore spiegazione. La capacità di un MOS è pari a quella di un condensatore a facce piane e parallele di spessore d come dato dalla (4.11), solo se le cariche sono addensate in prossimità dell'ossido. Ciò è sicuramente vero per le cariche del gate, ma non per quelle della regione di svuotamento che ha un'estensione certo non trascurabile. In effetti la capacità differenziale $C = dQ/dv_{GS}$ del MOS varia al variare della v_{GS} . Se v_{GS} è minore di zero, il gate attrae portatori p (cariche mobili) verso l'ossido e quindi la carica positiva è quasi tutta concentrata in prossimità dell'ossido. In questa condizione

la capacità differenziale è pari a quella di un condensatore a facce piane e parallele. Quando v_{GS} è maggiore di zero la capacità differenziale diminuisce, in quanto man mano che la regione di svuotamento aumenta la carica che va ad aggiungersi è sempre più distante dall'ossido. La diminuzione prosegue fino alla tensione di soglia. Per tensioni maggiori compaiono cariche elettriche mobili (elettroni) che vanno ad addensarsi in prossimità dell'ossido e quindi la capacità differenziale torna quella di un condensatore a facce piane e parallele la cui espressione è data dalla (4.11). È pertanto corretto applicare la (4.10) con la (4.11) per la carica (mobile) incrementale rispetto alla carica (fissa) accumulata nella regione di svuotamento fino alla tensione di soglia. La Figura 4.12 mostra come varia la capacità differenziale del gate al variare della tensione v_{GS} .

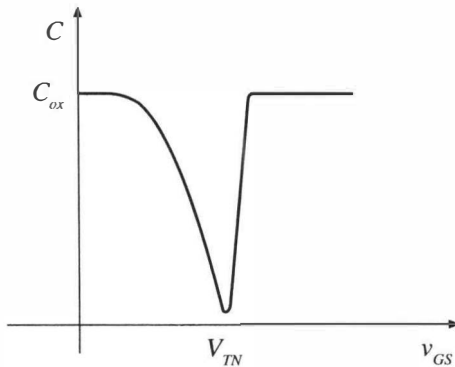


Figura 4.12 Capacità del MOS

Definiamo la capacità per unità di superficie dell'ossido come

$$C_{ox}'' = \frac{\epsilon}{d} \quad (4.12)$$

La (4.10) può essere riscritta come

$$Q = -C_{ox}'' WL(v_{GS} - V_{TN}) \quad (4.13)$$

Quando scorre una corrente, è come se questa carica distribuita sotto il gate scorresse rigidamente dal source verso il drain (non viceversa perché si tratta di una carica negativa). Con riferimento alla Figura 4.11 nel tempo Δt scorrerà verso il drain la carica ΔQ del tratto Δx , quindi:

$$i_{DS} = -\frac{\Delta Q}{\Delta t} = -\frac{\Delta Q}{\Delta x} \frac{\Delta x}{\Delta t} \quad (4.14)$$

Il segno meno sta a indicare che il verso positivo della corrente è quello da destra verso sinistra nell'asse delle x .

D'altra parte $\Delta Q/\Delta x$ è la densità di carica per unità di lunghezza, ovvero

$$\frac{\Delta Q}{\Delta x} = \frac{Q}{L} = -C_{ox}'' W (v_{GS} - V_{TN}) \quad (4.15)$$

$\Delta x/\Delta t$ è la velocità di deriva degli elettroni nel canale, ovvero

$$\frac{\Delta x}{\Delta t} = v_n \quad (4.16)$$

Poiché il canale è molto corto rispetto alla larghezza, il campo elettrico in esso è circa costante come in un condensatore a facce piane e parallele (le facce in questo caso sono le piazzole di drain e di source), quindi:

$$v_n = -\mu_n \bar{E} = -\mu_n \left(-\frac{v_{DS}}{L} \right) \quad (4.17)$$

con $\bar{E}$ campo elettrico. Sostituendo la (4.15) e (4.17) nella (4.14) si ottiene

$$i_{DS} = K_n (v_{GS} - V_{TN}) v_{DS} \quad (4.18)$$

avendo definito

$$K_n = \frac{C_{ox}'' \mu_n W}{L} \quad (4.19)$$

La (4.18) è, di fatto, l'espressione di una corrente che scorre in una resistenza variabile di valore pari a

$$R_{ON} = \frac{1}{K_n (v_{GS} - V_{TN})} \quad (4.20)$$

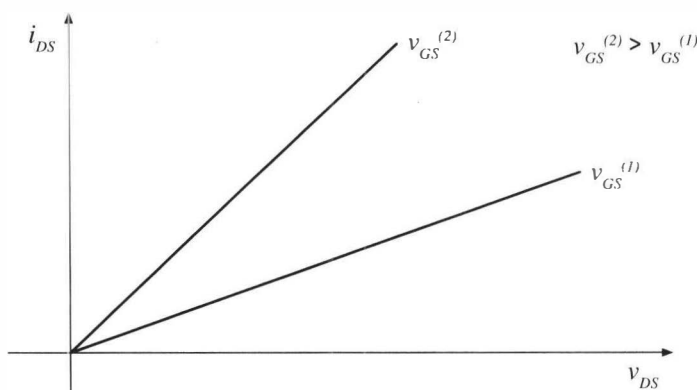


Figura 4.13 Corrente in funzione di v_{DS}

La tensione v_{GS} modifica la resistenza, quando v_{GS} cresce la resistenza diminuisce e quindi la corrente aumenta.

Vale la pena notare che la corrente scorre solo nel canale e questo è confinato sotto il gate. Si tratta di una caratteristica importante perché significa che si possono realizzare MOSFET molto vicini senza la necessità di zone di isolamento. Un altro aspetto notevole di questa tecnologia è proprio il fatto che il MOSFET si possa comportare come una resistenza. Un problema tipico dei circuiti integrati è che le resistenze tendono a essere fisicamente molto più grandi dei transistor stessi, finendo per essere il principale limite all'integrazione dei circuiti. La ragione è che, se le resistenze sono realizzate semplicemente come piste di silicio, per ragioni tecnologiche non possono essere troppo sottili, e quindi per ottenere resistenze grandi è necessario realizzare piste lunghe e ripiegate su se stesse a serpentina. Le zone di guardia tra una pista e l'altra sono un ulteriore problema nell'integrazione di tali resistenze. Questi problemi non sono presenti in una resistenza realizzata con un MOSFET. Come vedremo nel prossimo paragrafo questa "resistenza a MOSFET" tende a diventare non lineare quando la tensione ai capi cresce, ma questo può non essere un problema nella progettazione dei circuiti integrati.

4.6 MOSFET a canale n ad arricchimento per grandi tensioni

La trattazione del MOSFET fin qui svolta non tiene conto di un aspetto importante: la caduta di tensione per effetto Ohm lungo il canale. In effetti la (4.18) è un'approssimazione per v_{DS} piccola.

Il punto essenziale è che lo scorrere della corrente lungo il canale dal drain al source produce, a causa della resistività del canale, una caduta di tensione diversa da punto a punto, a seconda della distanza dal source. Chiamiamo questa tensione $v(x)$ con x posizione lungo il canale. Il punto $x = 0$ corrisponde al source, mentre per $x = L$ la tensione $v(x)$ è pari a v_{DS} , ovvero: $v(L) = v_{DS}$.

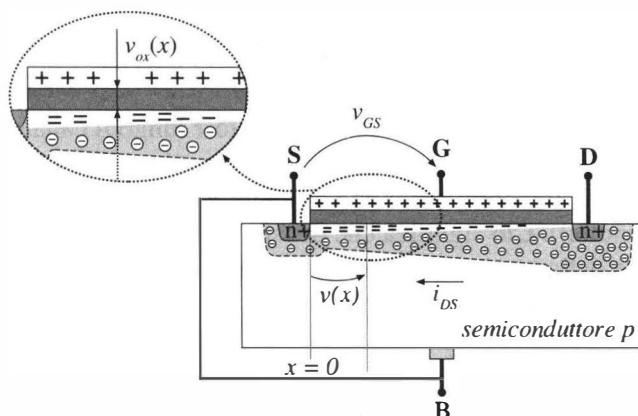


Figura 4.14 Canale per grandi correnti

La tensione ai capi dell'ossido, quando scorre una corrente, non è costante lungo il gate, ma sarà data dalla seguente espressione

$$v_{ox}(x) = v_{GS} - v(x) \quad (4.21)$$

La densità di carica mobile per unità di lunghezza del canale (Q') sarà proporzionale alla densità di capacità per unità di lunghezza del canale ($C_{ox}'' W$) e alla tensione sull'ossido meno la tensione di soglia

$$Q' = -C_{ox}'' W (v_{GS} - v(x) - V_{TN}) \quad (4.22)$$

Ragionando come nel paragrafo precedente, ovvero immaginando che la carica lungo il canale si muova traslando dal source al drain, la corrente i_{DS} in un generico punto del canale è data da

$$i_{DS} = -\frac{\Delta Q}{\Delta t} = -\frac{\Delta Q}{\Delta x} \frac{\Delta x}{\Delta t} = -Q' v_n \quad (4.23)$$

Quindi

$$i_{DS} = -C_{ox}'' W (v_{GS} - v(x) - V_{TN}) \mu_n \bar{E} \quad (4.24)$$

dove $\bar{E}$ è il campo elettrico. In generale un campo elettrico è il gradiente del potenziale e, poiché l'unico termine del potenziale nel canale dipendente da x è $v(x)$, dovrà valere

$$\bar{E} = -\frac{dv(x)}{dx} \quad (4.25)$$

Pertanto

$$i_{DS} = C_{ox}'' \mu_n W (v_{GS} - v(x) - V_{TN}) \frac{dv(x)}{dx} \quad (4.26)$$

La (4.26) esprime la corrente in un generico punto x lungo il canale. Sia la densità di carica, sia il gradiente del potenziale variano da punto a punto, tuttavia non è pensabile che la corrente non sia costante nel canale, (una corrente non può sparire nel nulla!). Ciò significa che densità di carica e velocità (proporzionale al gradiente della tensione) possono variare lungo il canale, ma sempre in modo da mantenere costante il loro prodotto: quando la carica diminuisce, la velocità aumenta e viceversa. Poiché quindi i_{DS} è costante lungo il canale possiamo scrivere

$$i_{DS} = \frac{1}{L} \int_{x=0}^{x=L} i_{DS} dx \quad (4.27)$$

Sostituendo la (4.26) nella (4.27)

$$i_{DS} = \frac{1}{L} \int_{x=0}^{x=L} C_{ox}'' \mu_n W (v_{GS} - v(x) - V_{TN}) \frac{dv(x)}{dx} dx \quad (4.28)$$

$$i_{DS} = \frac{1}{L} \int_{v=0}^{v=v_{DS}} C_{ox}'' \mu_n W (v_{GS} - v - V_{TN}) dv \quad (4.29)$$

Infine

$$i_{DS} = K_n \left((v_{GS} - V_{TN}) v_{DS} - \frac{1}{2} v_{DS}^2 \right) \quad (4.30)$$

con

$$K_n = \frac{C_{ox}'' \mu_n W}{L} \quad (4.31)$$

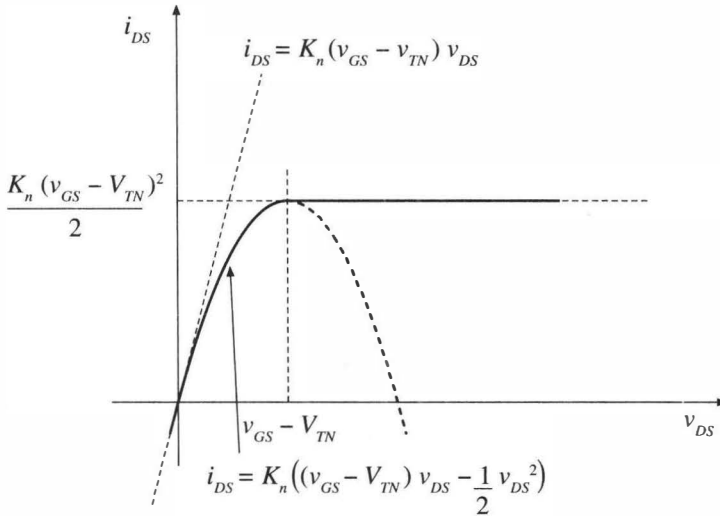


Figura 4.15 Caratteristica di uscita del MOSFET

Si noti che l'espressione della corrente nella (4.30) quando il termine quadratico è molto minore del termine lineare, ovvero per

$$v_{DS} \ll 2(v_{GS} - V_{TN}) \quad (4.32)$$

coincide con la (4.18) del paragrafo precedente. La (4.30) è una parabola passante per l'origine ($i_{DS} = 0$ per $v_{DS} = 0$) e avente vertice nel punto

$$v_{DS} = (v_{GS} - V_{TN}) \quad (4.33)$$

Il valore della corrente nel massimo è

$$i_{DS} = \frac{1}{2} K_n (v_{GS} - V_{TN})^2 \quad (4.34)$$

Matematicamente, la corrente espressa dalla parabola (4.30) superato il massimo dovrebbe scendere. Il fatto che all'aumentare della tensione la corrente scenda è quantomeno strano. Non impossibile, perché in realtà esiste un dispositivo con questa caratteristica (il diodo tunnel, un dispositivo oggi decisamente in disuso), ma un andamento del genere è almeno un po' sospetto. Vediamo in dettaglio cosa succede per $v_{DS} > v_{GS} - V_{TN}$.

La densità di carica mobile nel canale, come abbiamo visto, non è costante ma è massima per $x = 0$ ($v(0) = 0$) e minima per $x = L$ ($v(L) = v_{DS}$) e segue questa espressione:

$$Q'(x) = -C_{ox}'' W (v_{GS} - v(x) - V_{TN}) \quad (4.35)$$

Quando

$$v_{DS} = v_{GS} - V_{TN} \quad (4.36)$$

la (4.35) diviene

$$Q' = -C_{ox}'' W (v_{DS} - v(x)) \quad (4.37)$$

Poiché $v(x)$ aumenta all'aumentare di x e assume il valore massimo $v = v_{DS}$ per $x = L$, dalla (4.37) segue che per $x = L$, ovvero in prossimità del drain, la densità di carica è nulla, ovvero il canale non sussiste.

Se ora aumentiamo di Δv_{DS} la v_{DS} (quindi ora $v_{DS} > v_{GS} - V_{TN}$) il punto in cui la densità di carica è nulla ($v(x) = v_{GS} - V_{TN}$) cadrà lungo il tratto tra drain e source. Oltre questo punto la tensione ai capi dell'ossido è sotto la soglia, quindi la densità di carica è nulla e il canale è interrotto. Tale fenomeno è detto strozzamento del canale.

Questo fatto potrebbe far pensare semplicemente che la corrente cessi di scorrere (non essendoci più un canale in grado di supportarla), ma in realtà le cose sono un po' più complesse.

Innanzitutto osserviamo che la corrente non può interrompersi o diminuire perché, se così fosse, la caduta di tensione lungo il canale cesserebbe o diminuirebbe, comportando un aumento della tensione sotto l'ossido con relativo accrescimento della concentrazione delle cariche mobili sotto l'ossido e allargamento del

canale. Si deve allora raggiungere un qualche equilibrio tra la tendenza del canale a ridursi e la necessità che la corrente persista. Raggiunto tale equilibrio la corrente rimane costante al variare della v_{DS} e si dice che il MOSFET è passato dalla regione lineare alla saturazione.

L'andamento delle caratteristiche di uscita risulta quindi quello mostrato in Figura 4.16. Il confine tra regione lineare e saturazione, ovvero il luogo dei punti in cui la corrente diventa costante, è la parabola di espressione

$$i_{DS} = \frac{K_n}{2} v_{DS}^2 \quad (4.38)$$

Si noti che viene chiamata “regione lineare” tutta la zona a sinistra della parabola, anche se a rigore è lineare solo nel primo tratto quando la v_{DS} è piccola. Ma come al solito una nomenclatura quando è entrata nell'uso è difficile da cambiare anche se risulta palesemente fuorviante, come appunto in questo caso.

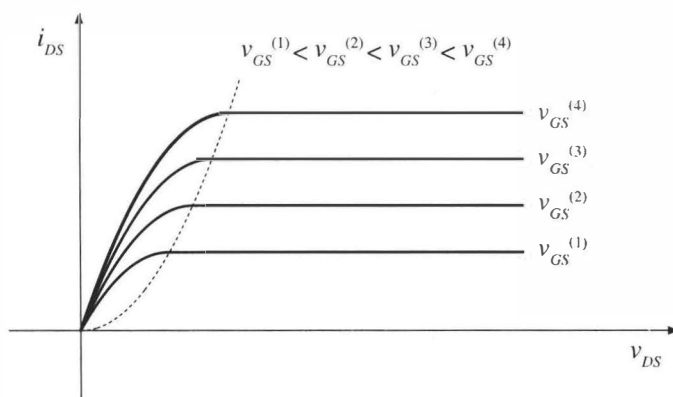


Figura 4.16 Caratteristiche di uscita del MOSFET

La corrente i_{DS} è nulla per $v_{GS} < V_{TN}$ (interdizione). Quando v_{GS} supera la soglia, se rimane verificato la condizione di saturazione

$$v_{DS} > v_{GS} - V_{TN} \quad (4.39)$$

la corrente i_{DS} è funzione solo di v_{GS} e ha andamento parabolico dato da

$$i_{DS} = \frac{1}{2} K_n (v_{GS} - V_{TN})^2 \quad (4.40)$$

come mostrato in Figura 4.17.

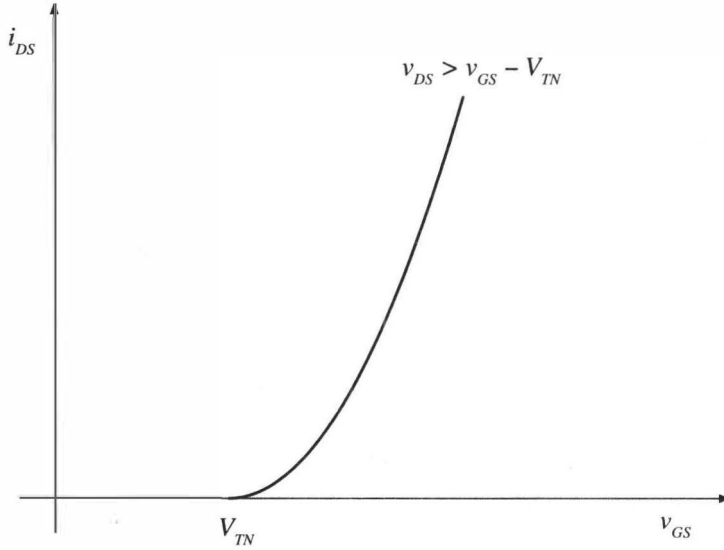


Figura 4.17 Caratteristica di trasferimento del MOSFET in saturazione

*4.7 Strozzamento del MOSFET

Vogliamo ora vedere con maggior dettaglio cosa succede in un MOSFET in saturazione. Abbiamo già notato che in regione lineare lungo il canale la densità di carica e la velocità degli elettroni non sono costanti, ma dal source al drain la densità diminuisce e la velocità aumenta mantenendo la corrente che scorre costante. La ragione di ciò è che la tensione $v(x)$ dal source al generico punto x lungo il canale non è costante, ma è crescente con x . Per calcolare l'andamento di $v(x)$ torniamo all'equazione (4.26), che per comodità riscriviamo

$$i_{DS} = C_{ox}'' \mu_n W (v_{GS} - v(x) - V_{TN}) \frac{dv(x)}{dx} \quad (4.41)$$

Integrando la variabile x in entrambi i membri dell'uguaglianza e tenendo presente che i_{DS} è una costante (il cui valore è dato dalla (4.30)) si ottiene

$$x = \left(\frac{K_n L}{i_{DS}} \right) \left(-\frac{1}{2} v^2 + (v_{GS} - V_{TN}) v \right) \quad (4.42)$$

che è una parabola passante per l'origine e avente vertice nel punto

$$v_{\max} = v_{GS} - V_{TN} \quad (4.43)$$

$$x_{\max} = \frac{(v_{GS} - V_{TN})^2}{2\left((v_{GS} - V_{TN})v_{DS} - \frac{1}{2}v_{DS}^2\right)} L \quad (4.44)$$

La parabola è graficata in Figura 4.18. Si noti che $x_{\max} > L$. Il ramo discendente non ha interesse.

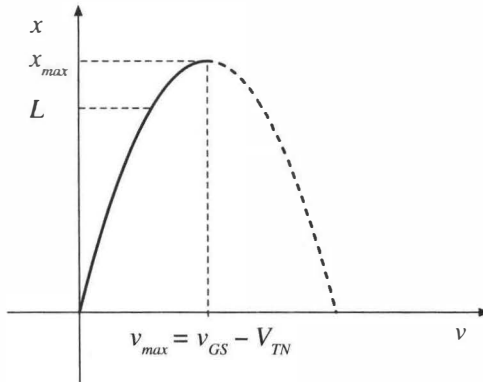


Figura 4.18 In funzione di v lungo il canale

La densità di carica (in valore assoluto) per unità di lunghezza nel canale è pari a

$$\rho(x) = C_{ox} W (v_{GS} - V_{TN} - v(x)) \quad (4.45)$$

La velocità di deriva degli elettroni è pari a

$$\bar{v}(x) = -\mu_n E = \mu_n \frac{dv(x)}{dx} \quad (4.46)$$

In Figura 4.19 sono mostrati i grafici della caduta di tensione $v(x)$, la densità di carica $\rho(x)$ e la velocità di deriva degli elettroni lungo il canale. $\rho(x)$ ha un profilo parabolico e diminuisce lungo il canale. Al contrario la velocità di deriva aumenta con un andamento a iperbole che va a infinito per $x = x_{\max}$.

In regione lineare $x_{\max} > L$. All'aumentare di v_{DS} , $x_{\max}$ diminuisce fino a diventare pari a L per $v_{DS} = v_{GS} - V_{TN}$. Tuttavia, quando $x_{\max} = L$ accade qualcosa di molto strano: la velocità di deriva in prossimità del drain ($x = L$) sembra schizzare a infinito e questo non è fisicamente possibile.

Nella realtà la velocità di deriva degli elettroni è proporzionale al campo elettrico solo per valori bassi del campo elettrico. Per grandi campi elettrici la mobilità degli elettroni diminuisce, finché la velocità di deriva raggiunge un valore massimo di saturazione come mostrato in Figura 4.20. Pertanto la velocità di deriva lungo il canale non va a infinito quando $v_{DS} \geq v_{GS} - V_{TN}$, ma raggiunge un valore massimo e poi rimane costante. Di conseguenza, poiché la corrente lungo il canale deve

rimanere costante, anche la densità delle cariche elettriche non tende a zero, ma raggiunge un valore minimo e, come mostrato in Figura 4.21, non diminuisce ulteriormente. Tale valore è comunque abbastanza piccolo da poter considerare il campo elettrico costante nel tratto dove il canale è strozzato. Sotto queste ipotesi la caduta di tensione aumenta linearmente fino al drain.

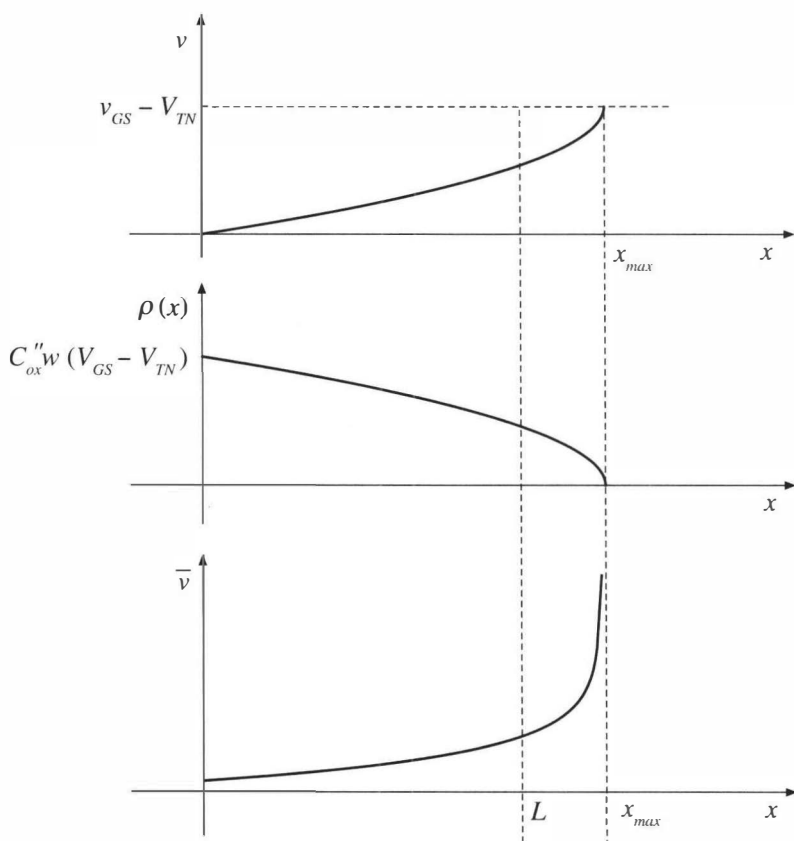


Figura 4.19 Caduta di tensione, densità di carica e velocità di deriva lungo il canale

Abbiamo detto che quando il canale è strozzato (per $v_{DS} \geq v_{GS} - V_{TN}$), la corrente i_{DS} non aumenta ulteriormente all'aumentare di v_{DS} . Tuttavia, all'aumentare di v_{DS} il punto in cui inizia lo strozzamento si allontana dal drain e si avvicina al source. In Figura 4.21 con L_0 si indica la distanza tra source e drain e con L la lunghezza del canale. L è minore di L_0 quando il canale è strozzato. È importante precisare, inoltre, che il valore di campo elettrico per cui la velocità di deriva non aumenta ulteriormente è molto alto, quindi il punto lungo il canale in cui inizia lo strozzamento è piuttosto vicino al punto in cui il campo andrebbe a infinito se non considerassimo il fatto che la velocità di deriva tende a saturare.

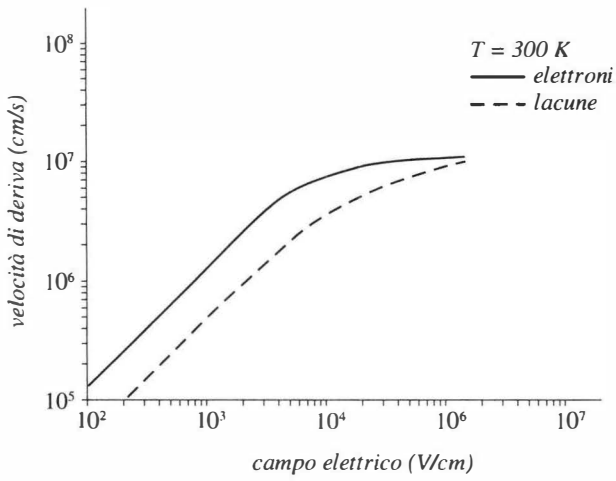


Figura 4.20 Velocità di deriva in funzione del campo elettrico

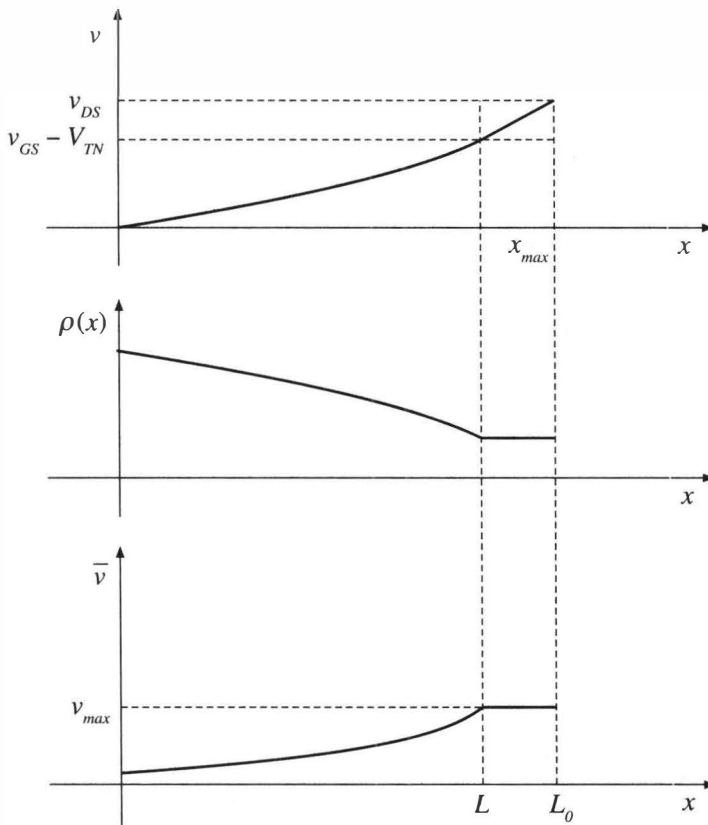


Figura 4.21 Caduta di tensione, densità di carica e velocità di deriva lungo il canale strozzato

4.8 Modulazione del canale

Proprietà fondamentale del MOSFET in saturazione è il fatto che la corrente i_{DS} , in prima approssimazione, non dipende dalla tensione tra drain e source, ma solo dalla tensione di controllo v_{GS} . In realtà, come è ragionevole aspettarsi, la corrente di drain non è del tutto indipendente dalla tensione tra drain e source, ma aumenta leggermente all'aumentare della tensione. Tale fenomeno è detto "modulazione del canale". Vediamone la ragione fisica.

Quando il MOSFET è in saturazione

$$i_{DS} = \frac{C_{ox}'' \mu_n W}{L} (v_{GS} - V_{TN})^2 \quad (4.47)$$

dove L è la lunghezza del canale, che in prima approssimazione abbiamo considerato pari alla distanza tra D e S. Tuttavia in saturazione l'effettiva lunghezza del canale è minore della distanza L_0 tra drain e source a causa dello strozzamento. Scriviamo quindi

$$i_{DS} = \frac{C_{ox}'' \mu_n W}{L_0 - \Delta L} (v_{GS} - V_{TN})^2 \quad (4.48)$$

con ΔL diminuzione del canale rispetto alla distanza L_0 tra S e D. La tensione tra source e strozzamento è pari a $(v_{GS} - V_{TN})$. Nel tratto ΔL tra la fine del canale e il drain il campo elettrico E sotto l'ossido è costante in quanto la concentrazione di cariche nel tratto strozzato può essere considerata nulla, pertanto

$$v_{DS} = (v_{GS} - V_{TN}) + E \Delta L \quad (4.49)$$

ovvero

$$v_{DS} = (v_{GS} - V_{TN}) + \frac{v_{max}}{\mu_n} \Delta L \quad (4.50)$$

Sostituendo la (4.50) nella (4.48) si ottiene

$$i_{DS} = \frac{C_{ox}'' \mu_n W}{L} (v_{GS} - V_{TN})^2 \frac{1}{1 - \frac{\mu_n}{v_{max}} L (v_{DS} - (v_{GS} - V_{TN}))} \quad (4.51)$$

Questa è l'espressione generale della corrente i_{DS} in saturazione. Tuttavia quando vale le seguente condizione

$$v_{GS} - V_{TN} \ll v_{DS} \ll \frac{v_{max} L}{\mu_n} + (v_{GS} - V_{TN}) \quad (4.52)$$

la (4.51) può essere approssimata come

$$i_{DS} = \frac{C_{ox}^n \mu_n W}{L} (v_{GS} - V_{TN})^2 (1 + \lambda v_{DS}) \quad (4.53)$$

con

$$\lambda = \frac{\mu_n}{v_{max} L} \quad (4.54)$$

La (4.53) è l'espressione che comunemente è utilizzata per tenere conto della modulazione del canale nei modelli circuitali del MOSFET. Si noti che $1/\lambda$ è fisicamente una tensione. Fissato $V_{DS} \gg V_{GS} - V_{TN}$ la tangente alla caratteristica intercetta l'asse v_{DS} nel punto $-1/\lambda$.

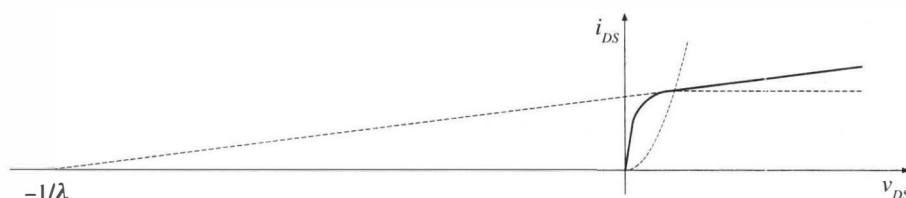


Figura 4.22 Modulazione del canale

4.9 Realizzazione fisica del MOSFET

L'aspetto forse più notevole del transistor è il fatto che è possibile realizzarne migliaia o milioni su una superficie di pochi millimetri interconnessi in modo da formare circuiti molto complessi detti IC (*Integrated Circuit*), in piccole piastrine rettangolari dette *chip*. Tali piastrine sono tagliate da un unico disco circolare di spessore 0.3 mm e diametro 10 cm, chiamato *wafer*, il cui processo produttivo è già stato descritto nel Capitolo 1.

Il MOSFET è il transistor oggi più largamente impiegato nei circuiti integrati. I dettagli tecnologici della loro fabbricazione sono in continua evoluzione e possono variare a seconda del costruttore. Il processo che descriveremo nel seguito va quindi considerato a titolo esemplificativo.

1) Deposizione di nitruro di silicio. La superficie del wafer di silicio monocristallino è interamente coperta con **uno strato di nitruro di silicio** (Si_3N_4) che è un composto del silicio molto resistente e praticamente impenetrabile ai droganti normalmente usati.

2) Mascheratura mediante fotolitografia. Sul nitruro di silicio si deposita uno strato di **fotoresist**, che è un polimero sensibile alla luce. Si appoggia,

quindi, una maschera di vetro con zone trasparenti e zone opache. Si illumina il wafer mascherato con un fascio di elettroni. In passato si usava luce ultravioletta (UV), ma oggi le dimensioni dei particolari della mascheratura sono più piccoli della lunghezza d'onda dell'UV (200-300nm). Le zone in ombra rimangono inalterate, le zone illuminate si induriscono (esistono dei fotoresist detti **positivi**, che invece si sciolgono quando illuminati da un fascio di elettroni, in tal caso le zone trasparenti e in ombra devono essere invertite). Il fotoresist viene lavato via con un solvente e rimane solo dove è stato indurito dalla luce. Il fotoresist rimasto viene ulteriormente indurito e stabilizzato con dei fissanti chimici.

3) **Rimozione del nitruro di silicio.** Il wafer è immerso in un bagno chimico aggressivo in grado di rimuovere il nitruro di silicio nelle zone non protette dal fotoresist.

4) **Rimozione del fotoresist.** Mediante opportuni reagenti il fotoresist è asportato e il wafer rimane con zone scoperte e zone protette dal nitruro di silicio. Le zone protette corrispondono a singoli MOSFET che si vorranno realizzare.

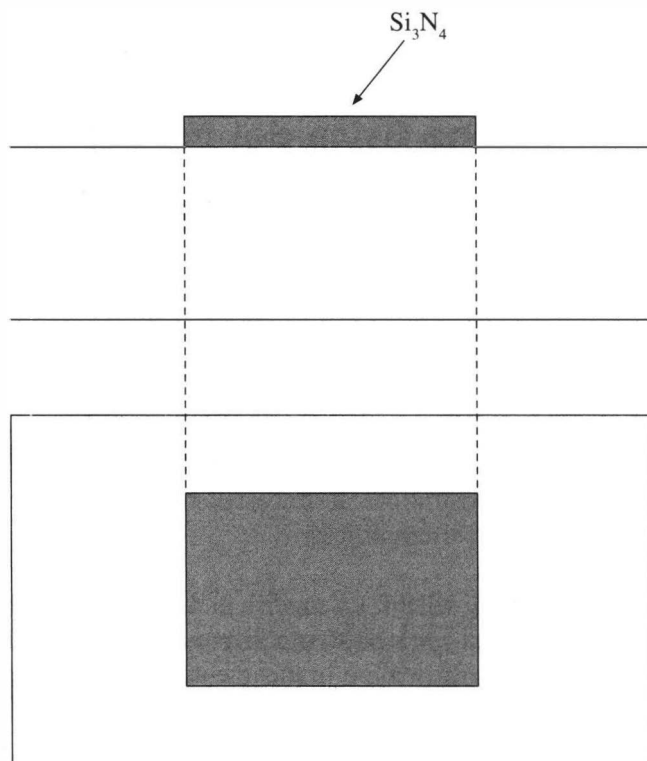


Figura 4.23 Nitruro di silicio asportato selettivamente mediante fotolitografia

5) **Impiantazione ionica.** Il wafer è bombardato con un fascio di atomi ionizzati che penetrano il silicio non protetto dal nitruro di silicio. In questo modo si realizza una sorta di cornice di impurità p+ che separa ciascun MOSFET.

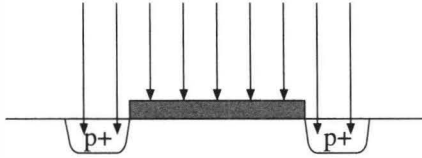


Figura 4.24 Impiantazione ionica

6) **Ossidazione.** Sul wafer viene fatto crescere uno strato piuttosto spesso ($1\mu\text{m}$) di ossido di silicio (SiO_2). L'ossido cresce solo dove il silicio è scoperto e viene detto **ossido di campo**.



Figura 4.25 Deposizione ossido di campo

7) **Rimozione del nitruro di silicio.** Il nitruro di silicio viene asportato con un attacco chimico.



Figura 4.26 Asportazione nitruro di silicio

8) **Seconda ossidazione.** Si deposita un sottile strato di ossido di silicio (dell'ordine della decina di nm) che andrà a costituire l'ossido sotto il gate.



Figura 4.27 Deposizione ossido di silicio

9) **Deposizione di polisilicio.** Si deposita uno strato di polisilicio, ovvero silicio policristallino molto drogato n.

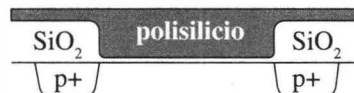


Figura 4.28 Deposizione polisilicio

10) **Seconda litografia.** Si esegue una seconda mascheratura litografica con la quale si proteggono le aree dove verranno realizzati i gate.

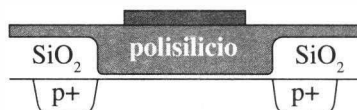


Figura 4.29 Mascheratura litografica

11) **Rimozione del polisilicio.** Il polisilicio è asportato chimicamente.

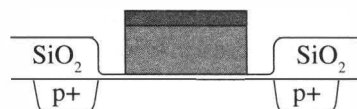


Figura 4.30 Asportazione polisilicio

12) **Rimozione del fotoresist.** Il fotoresist è dilavato con opportuni reagenti chimici.



Figura 4.31 Rimozione fotoresist

13) **Seconda impiantazione ionica.** A questo punto si esegue un'impiantazione ionica n+. È importante notare che si sfruttano come “maschera” lo spesso strato di ossido di campo e il polisilicio (che essendo già molto drogato n+ non viene sostanzialmente alterato dall'immissione di ulteriori impurità n). Il sottile strato di ossido tra gate e ossido di campo è invece attraversato dagli ioni e sotto di esso si formano il source e il drain. Questo processo, che usa il gate come maschera, consente di “autoallineare” le piazzole di source e drain al margine del gate. In questo modo si evita di creare capacità parassite tra il gate e le piazzole di drain e source. Queste capacità parassite sono un elemento critico che limita l'operatività del MOSFET ad alta frequenza (come vedremo nel Capitolo 8).

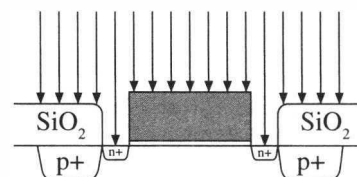


Figura 4.32 Realizzazione delle piazzole di source e drain mediante impiantazione ionica

14) **Terza ossidazione.** Il wafer è ora interamente ricoperto da uno strato di ossido di silicio.

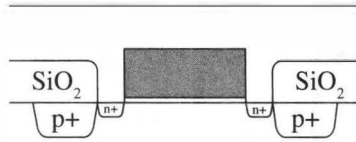


Figura 4.33 Deposizione di ossido di silicio

15) **Terza litografia.** Si esegue una terza litografia per selezionare i punti dove si vogliono inserire i contatti elettrici.

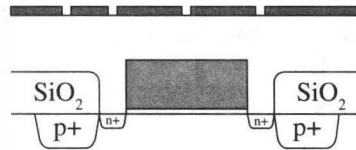


Figura 4.34 Mascheratura

16) **Attacco chimico.** Mediante attacco chimico si rimuove l'ossido dalle zone in cui si vogliono realizzare i contatti elettrici.

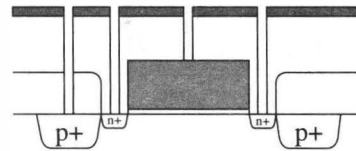


Figura 4.35 Attacco chimico selettivo

17) **Rimozione della mascheratura.** Mediante specifici reagenti chimici si rimuove la mascheratura.

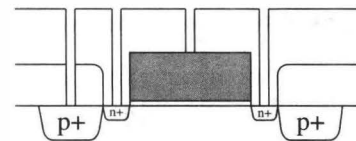


Figura 4.36 Rimozione mascheratura

18) **Metallizzazione.** Mediante vaporizzazione di metallo (alluminio e più recentemente anche rame) si riempiono i pozzi dei contatti elettrici e si copre tutto il wafer di una pellicola metallica.

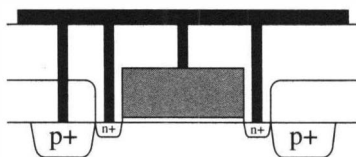


Figura 4.37 Vaporizzazione di metallo

19) **Quarta litografia.** Si esegue una quarta litografia per selezionare le piste dei contatti elettrici.

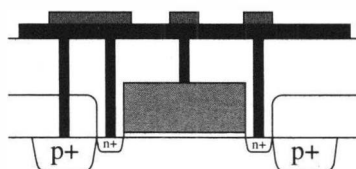


Figura 4.38 Quarta litografia

20) **Rimozione del metallo** mediante attacco chimico.

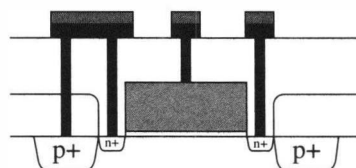


Figura 4.39 Asportazione metallo

21) **Rimozione della mascheratura.** Questa è l'ultima operazione per realizzare il singolo transistor. Per un circuito più complesso è necessario creare una rete di interconnessioni metalliche. Queste si realizzano mediante strati successivi di ossido e metallo con punti di contatto detti vie (*vias*).

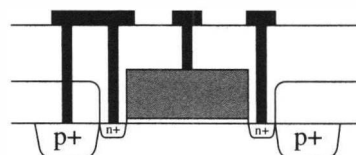


Figura 4.40 Asportazione mascheratura

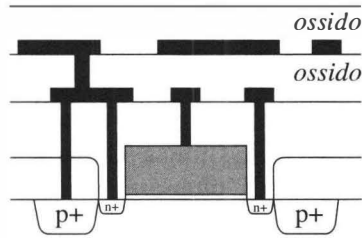


Figura 4.41 MOSFET completo con due strati di interconnessioni elettriche

Come si può capire, la fabbricazione del MOSFET è complessa e costosa. La fabbricazione delle lastre litografiche è un'operazione particolarmente lunga e costosa, per di più specifica per il particolare circuito che si vuole realizzare. Questi elevati costi sono ammortizzati dal numero di circuiti integrati che si possono produrre sullo stesso wafer. A titolo d'esempio si possono fabbricare 10.000 circuiti integrati ciascuno costituito da un milione di singoli transistor.

I MOSFET su uno stesso wafer sono realizzati con le stesse caratteristiche chimico-fisiche, se però sono necessari MOSFET con diverso K_n si può agire sul rapporto geometrico W/L . Nei circuiti ad alta integrazione L è pari alla dimensione minima che la tecnologia consente e si varia W a multipli di L . L'inconveniente di usare la dimensione minima è che questa è associata alla massima incertezza sul suo effettivo valore. In effetti nei circuiti analogici, dove questa indeterminazione è più critica, si usano transistor sensibilmente più grandi della dimensione minima.

Un circuito, ovviamente non è costituito da soli MOSFET. Per quello che riguarda i capacitori, essi si realizzano facilmente con la stessa tecnologia del MOSFET (la cui base è per l'appunto un condensatore MOS). I resistori sono elementi più critici perché, anche realizzandoli come strati stretti e sottili, tendono a occupare molto spazio a causa della bassa resistività del silicio (drogato) che si usa nella tecnologia planare. Quando è possibile si cerca di sostituire i resistori con MOSFET. Nel seguito vedremo degli esempi. Le induttanze, infine, non possono essere realizzate in tecnologia planare. Pertanto si studiano soluzioni circuitali alternative all'uso delle induttanze, oppure si prevede il montaggio di induttanze esterne al circuito integrato.

4.10 Recenti sviluppi della tecnologia dei MOSFET

La tecnologia descritta nel paragrafo precedente si riferisce al processo produttivo impiegato per realizzare MOSFET fino a 32nm di dimensione fisica del gate (che era lo stato dell'arte del 2009). L'ulteriore miniaturizzazione ha richiesto un sostanziale cambio di geometria.

L'obiettivo dei più recenti sviluppi della tecnologia dei MOSFET è aumentare la superficie di contatto tra gate e canale senza allargare la dimensione planare del canale. La Figura 4.42 mostra la soluzione adottata da Intel a partire dal 2012: il canale anziché essere planare ha la forma di una pinna che entra dentro il corpo del gate. In questo modo la superficie di contatto tra canale e gate è 2-3 volte la superficie planare del transistor. I processori i5 e i7 della Intel sono realizzati con questa tecnologia.

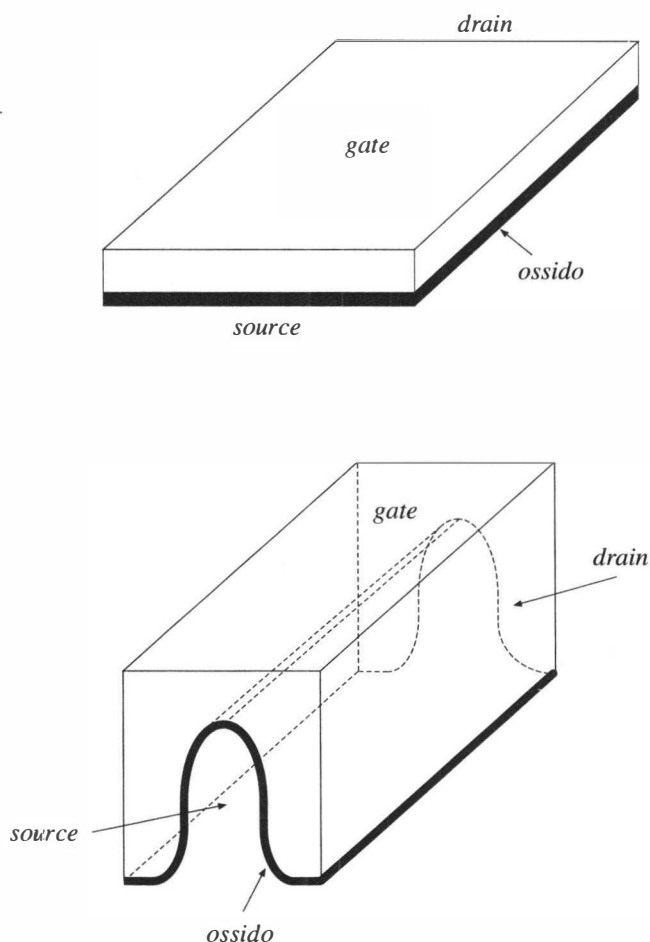


Figura 4.42 MOSFET a pinna

Sono allo studio tecnologie ancora più radicali basate su nanotubi completamente immersi nel gate. Tali strutture possono essere orizzontali (Figura 4.43) o verticali (Figura 4.44). Allo stato dell'arte (2014) sono però ancora dei prototipi.

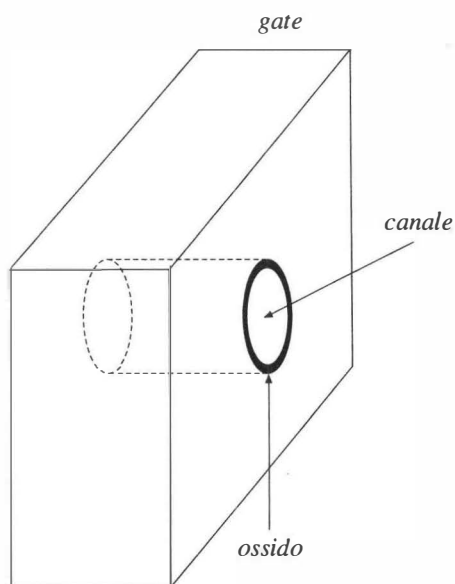


Figura 4.43 Nanotubo orizzontale

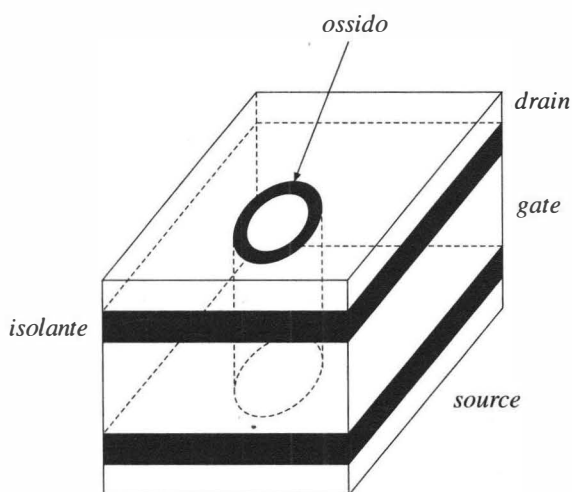


Figura 4.44 Nanotubo verticale

4.11 Legge di Moore e miniaturizzazione dei MOSFET

Gordon Moore, cofondatore della Intel, già nei primi anni sessanta osservò che il numero di transistor realizzabili su un singolo chip cresceva esponenziale nel tempo. Inizialmente Moore ipotizzò un tempo di raddoppio dei transistor per unità di superficie pari a un anno, successivamente pari a due, attualmente, sulla base dei dati storici dagli anni sessanta a oggi, il tempo di raddoppio risulta prossimo a 18 mesi.

Si tratta di un fatto piuttosto sorprendente, oltre che di enorme impatto economico e sociale: nei cinquant'anni di sviluppo dei semiconduttori le tecnologie sono cambiate più volte, ma la tendenza alla miniaturizzazione del singolo transistor è rimasta invariata. In effetti, lo straordinario successo del MOSFET è dovuto in gran parte al fatto che il suo processo produttivo consente un'elevata miniaturizzazione.

Nel 2003 lo stato dell'arte era una dimensione fisica del gate (*size*) di 90nm, nel 2005 di 65nm, nel 2007 di 45nm, nel 2009 di 32nm e nel 2011 appena di 22nm. Come si vede la legge di Moore è ben rispettata ancora oggi.

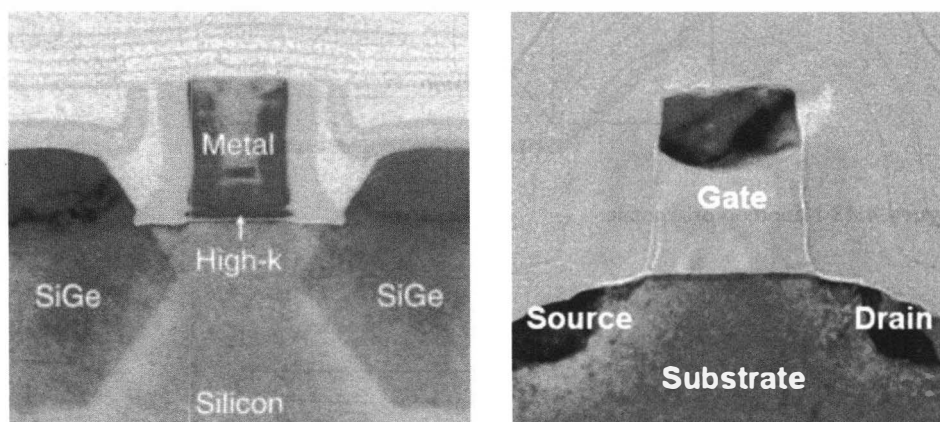


Figura 4.45 Immagine al microscopio elettronico di un MOSFET prodotto dalla Intel con gate di 45nm (a sinistra) e con gate di 32nm (a destra)

Una domanda fondamentale è fino a che punto potrà spingersi la miniaturizzazione del singolo transistor. L'attuale lunghezza dei gate è 22nm, altri 3 o 4 dimezzamenti, ovvero cinque o sei anni e l'intero MOSFET diverrà un oggetto del tutto quantistico, con proprietà e leggi molto diverse da quelle a cui siamo abituati. Dobbiamo allora chiederci se i "transistor quantistici" potranno continuare a essere affidabili ed economici come o più degli attuali, oppure se la corsa alla miniaturizzazione dovrà subire un arresto. Qualunque sia la risposta, lo sapremo molto presto.

4.12 MOSFET a canale n ad arricchimento con carico resistivo

Un MOSFET può essere impiegato per controllare la corrente che scorre su una resistenza di carico (carico resistivo) in modo da operare come interruttore controllato in tensione o come amplificatore. A questo proposito si consideri lo schema in Figura 4.46. Si definisce funzione di trasferimento la funzione che esprime la relazione tra la tensione di uscita (presa sul drain) e la tensione di ingresso sul gate.

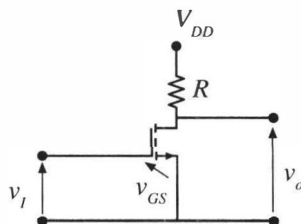


Figura 4.46 MOSFET con resistenza di carico

La tensione $v_{DS} = v_o$ è data da

$$v_{DS} = V_{DD} - i_{DS}R \quad (4.55)$$

che è la retta di carico nel piano delle caratteristiche di uscita del MOSFET come mostrato Figura 4.47.

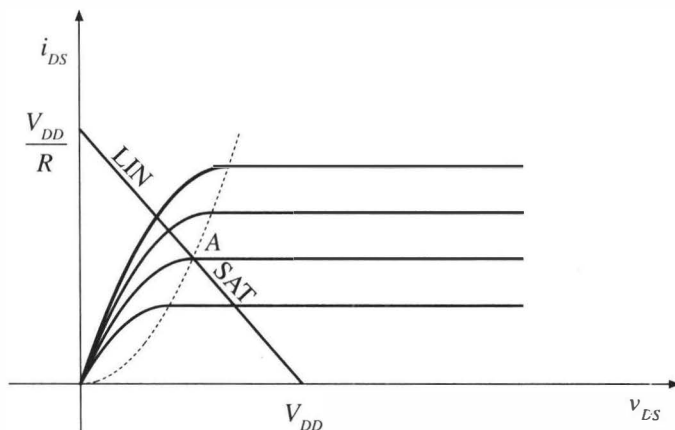


Figura 4.47 Caratteristiche di uscita e retta di carico

Il valore di $v_I = v_{GS}$ fissa la caratteristica. Il valore della corrente i_{DS} è dato dall'incrocio tra la caratteristica e la retta di carico. Quando v_{GS} è minore della soglia V_{TN} , il MOSFET è in interdizione. Poiché $i_{DS} = 0$, segue che $v_{DS} = V_{DD}$. Appena v_I raggiunge la tensione di soglia, il MOSFET è in conduzione, scorre della cor-

rente sulla resistenza R e quindi la tensione di uscita cala. Il punto di lavoro si sposta lungo la retta di carico restando in saturazione fino al punto A, poi passa in regione lineare. In generale il MOSFET è in saturazione se

$$v_{DS} > v_{GS} - V_{TN} \quad (4.56)$$

ovvero

$$v_O > v_I - V_{TN} \quad (4.57)$$

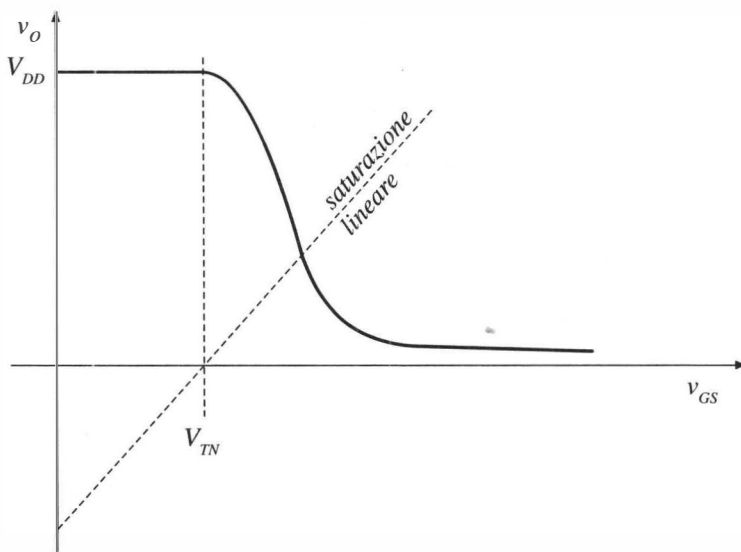


Figura 4.48 Funzione di trasferimento

Tracciamo ora il grafico di v_{DS} in funzione di v_{GS} . Per $v_{GS} < V_{TN}$, $v_{DS} = V_{DD}$, poi cala al suo valore minimo. Se in questo grafico tracciamo la retta

$$v_O = v_I - V_{TN} \quad (4.58)$$

questa separa il piano in due zone: nella zona superiore il MOSFET è in saturazione, nella zona inferiore è in regione lineare. In saturazione l'andamento della tensione è facilmente calcolabile come

$$v_O = V_{DD} - R \frac{K_n}{2} (v_I - V_{TN})^2 \quad (4.59)$$

ed è quindi il ramo discendente di una parabola rovesciata. L'andamento in regione lineare richiede di risolvere l'equazione di secondo grado

$$v_O = V_{DD} - RK_n \left((v_{GS} - V_{TN}) v_O - \frac{1}{2} v_O^2 \right) \quad (4.60)$$

La funzione di trasferimento è già una porta NOT elementare, infatti l'uscita ha un valore alto, quando l'ingresso è basso e viceversa.

4.13 MOSFET come amplificatore di tensione

Supponiamo che in ingresso al circuito in Figura 4.49 ci sia una rete lineare progettata per mantenere il MOSFET in regione di saturazione. Tale rete potrà sempre essere vista come il suo equivalente di Thévenin (vedi Appendice B), ovvero un generatore di tensione V_G e una resistenza R_G . Poiché il gate è un capacitore, per una tensione costante è assimilabile a un circuito aperto, ovvero una resistenza infinita. Pertanto $V_{GS} = V_G$.

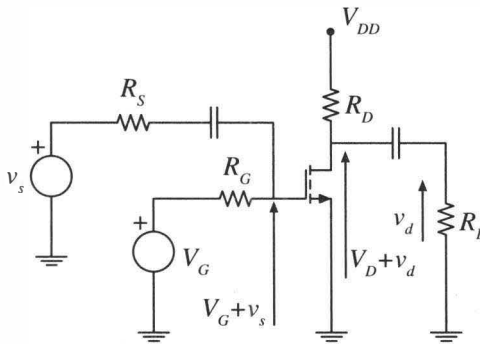


Figura 4.49 Amplificatore

Consideriamo ora un generatore di segnale, ovvero un generatore di tensione v_s variabile nel tempo con una sua resistenza di uscita $R_S \ll R_G$, collegato al gate mediante un capacitore di capacità molto grande (il dimensionamento di tale capacitore sarà trattato nel Capitolo 8). Tale capacitore detto **capacitore di accoppiamento** è un circuito aperto per la tensione costante V_G , ma un cortocircuito per il segnale v_s variabile nel tempo. Pertanto la componente in continua della tensione nel gate sarà pari a V_G , mentre la componente variabile nel tempo sarà pari a v_s (la caduta di tensione su R_S è trascurabile se $R_S \ll R_G$). Le due componenti sono indipendenti e si sommano, quindi

$$v_{GS} = V_G + v_s \quad (4.61)$$

Applicando la tensione data dalla (4.61) alla funzione di trasferimento, risulta che la tensione sul drain v_d sarà la somma di un valore costante V_D e un segnale v_d in controfase rispetto a v_s (in controfase significa che quando v_i aumenta v_d diminuisce).

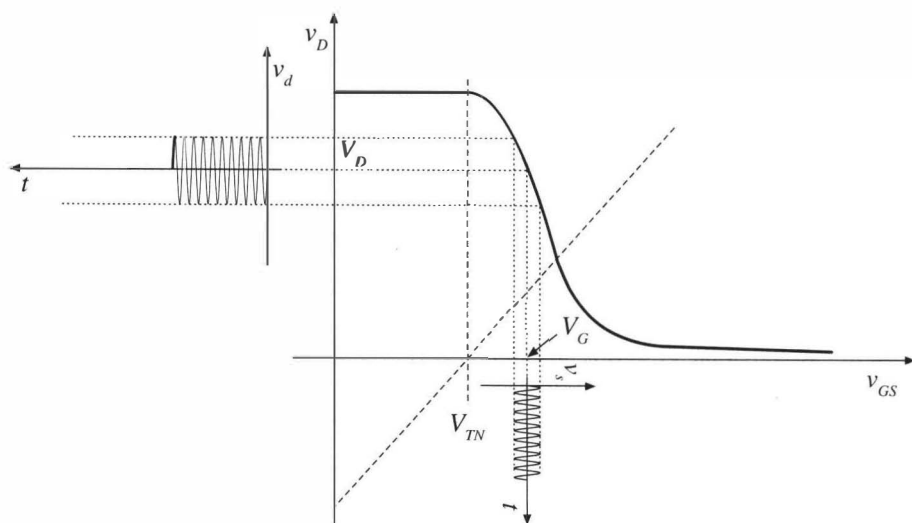


Figura 4.50 Amplificazione del segnale

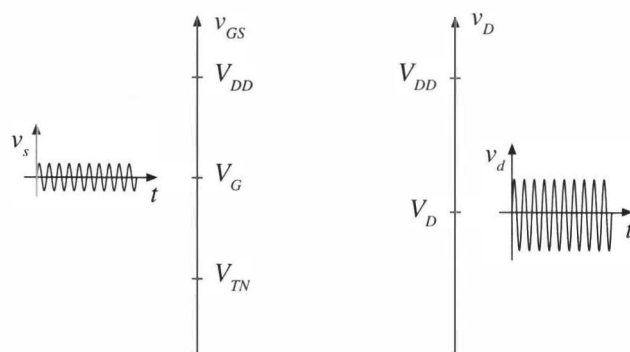


Figura 4.51 Tensioni in continua e variabili

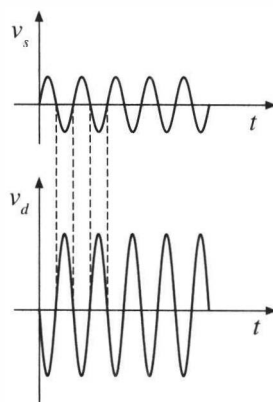


Figura 4.52 Segnale in ingresso e in uscita

Poiché la pendenza della funzione di trasferimento è maggiore di 1 (in valore assoluto), l'ampiezza di v_d è maggiore di quella del segnale in ingresso. Abbiamo pertanto realizzato un amplificatore di tensione.

Resta ora da vedere come trasferire il segnale su una resistenza di carico R_L . Ciò può essere fatto mediante un altro capacitore di capacità abbastanza grande tra il drain e il carico. Infatti per la componente in continua della tensione di drain il capacitore è un circuito aperto e pertanto essa non viene trasferita al carico.

Per la componente variabile il capacitore è un cortocircuito e pertanto essa è interamente trasferita al carico, ovvero sul carico c'è la stessa tensione variabile che c'è sul drain.

4.14 Polarizzazione del MOSFET a canale n ad arricchimento

Per utilizzare un MOSFET come amplificatore è necessario che sia polarizzato in saturazione. Il modo più semplice per farlo è con un partitore di tensione come mostrato in Figura 4.53

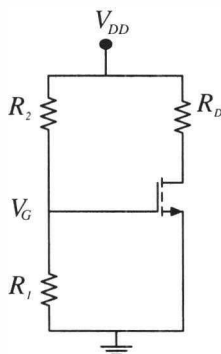


Figura 4.53 Circuito di polarizzazione del MOSFET

La corrente che scorre tra drain e source è pari a

$$i_{DS} = \frac{K_n}{2} (V_G - V_{TN})^2 \quad (4.62)$$

con

$$V_G = \frac{R_2}{R_1 + R_2} V_{DD} \quad (4.63)$$

La condizione da rispettare affinché il MOSFET sia in saturazione è

$$V_{DS} > V_G - V_{TN} \quad (4.64)$$

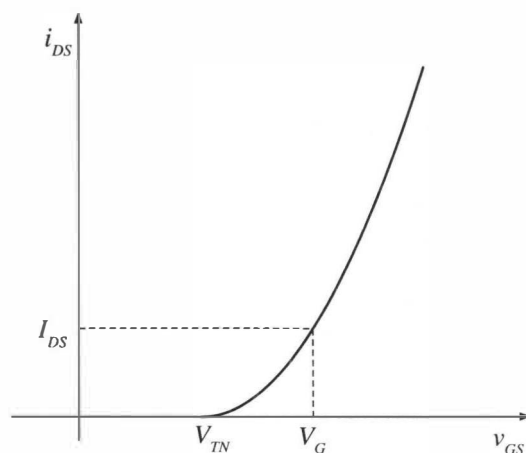


Figura 4.54 Corrente di polarizzazione

4.15 Rete di polarizzazione con resistenza di source

Le caratteristiche dei MOSFET (in particolare il K_n) sono piuttosto variabili anche per transistor nominalmente identici. Se nella progettazione della rete di polarizzazione si tiene conto di tale variabilità (Figura 4.55), il valore della corrente i_{DS} risulta poter variare in un range piuttosto ampio.

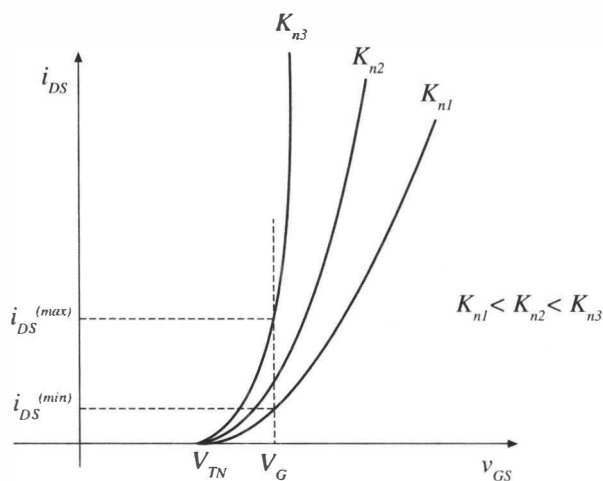


Figura 4.55 Variabilità delle caratteristiche

Una possibile soluzione per ridurre l'effetto della variabilità della caratteristica sulla corrente i_{DS} è l'inserimento di una resistenza R_s tra source e massa come in Figura 4.56.

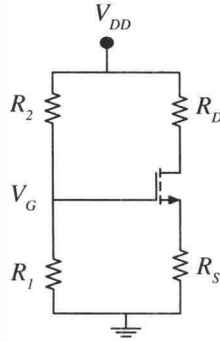


Figura 4.56 Polarizzazione del MOSFET con rete a quattro resistori

Per tale circuito deve valere

$$i_{DS} = \frac{V_G - V_{GS}}{R_S} \quad (4.65)$$

Pertanto tracciando la retta (4.65) sulla curva i_{DS} in funzione di v_{GS} del MOSFET in saturazione si ottiene il grafico di Figura 4.57, che mostra come il range di escursione di i_{DS} sia notevolmente minore rispetto al caso mostrato in Figura 4.55.

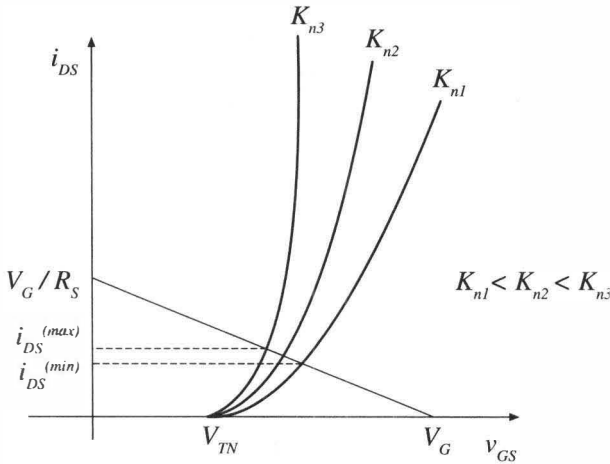


Figura 4.57 Variabilità delle caratteristiche

Il meccanismo mediante il quale la resistenza R_S stabilizza il punto di lavoro del MOSFET è una delle molte applicazioni della **reazione negativa** in elettronica. Rispetto al caso di polarizzazione con semplice partitore sul gate, la rete con resistenza di source è meno sensibile alle variazioni della curva caratteristica del MOSFET per il fatto che la caduta di tensione su R_S tende a ridurre v_{GS} . Pertanto se una diversa curva caratteristica del MOSFET determina un valore di i_{DS} più elevato, il conseguente aumento della caduta di tensione su R_S produce una riduzione di v_{GS} che tende ad attenuare la variazione di i_{DS} stessa.

4.16 Polarizzazione con generatore di corrente

In alternativa alla resistenza di source si può utilizzare un generatore di corrente come in Figura 4.58

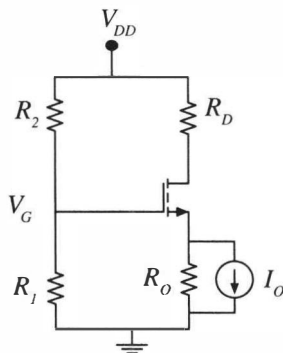


Figura 4.58 Rete di polarizzazione

Scriviamo

$$i_{DS} = \frac{V_G - V_{GS}}{R_0} + I_0 = \frac{V_G + I_0 R_0}{R_0} - \frac{V_{GS}}{R_0} \quad (4.66)$$

Il punto di polarizzazione è quindi l'intersezione delle curve nel grafico in Figura 4.59.

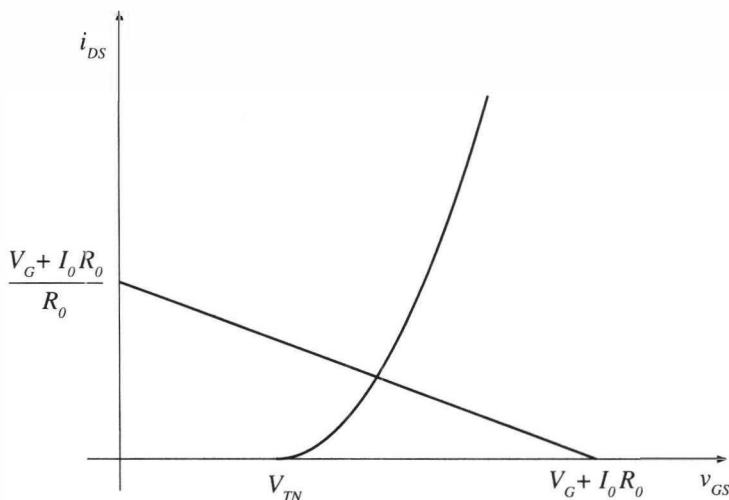


Figura 4.59 Punto di polarizzazione

Il grande vantaggio di questa configurazione è che la R_0 per un generatore di corrente può essere molto grande e quindi la retta praticamente piatta (Figura 4.60).

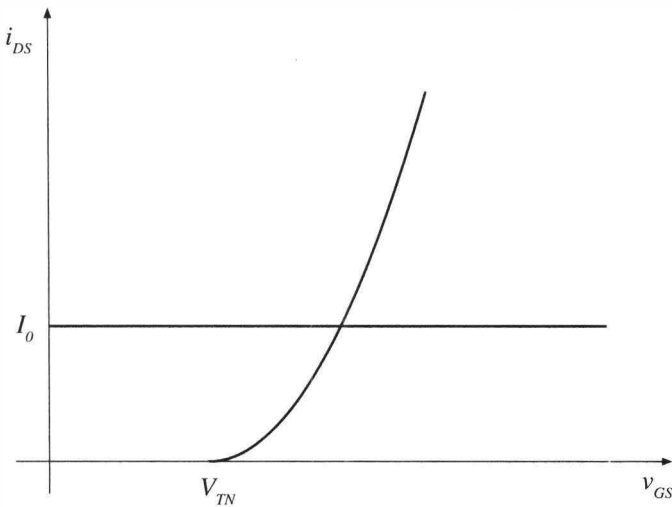


Figura 4.60 Punto di polarizzazione

4.17 Specchio di corrente

Il generatore di corrente citato nel paragrafo precedente è talvolta realizzato con uno **specchio di corrente**, ovvero una coppia di MOSFET identici collegati come in Figura 4.61.

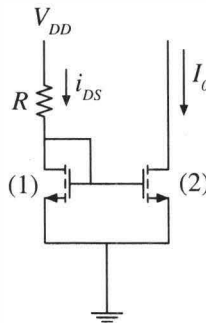


Figura 4.61 Specchio di corrente

Poiché i due MOSFET hanno stessa v_{GS} , tirano la stessa corrente (per questo il circuito è chiamato specchio di corrente). D'altra parte la corrente che scorre nel MOSFET (1) è data da

$$i_{DS} = \frac{V_{DD} - V_{GS}}{R} \quad (4.67)$$

Sovrapponendo la retta (4.67) alla caratteristica del MOSFET in saturazione si ottiene il grafico di Figura 4.62, analogo a quello di Figura 4.57.

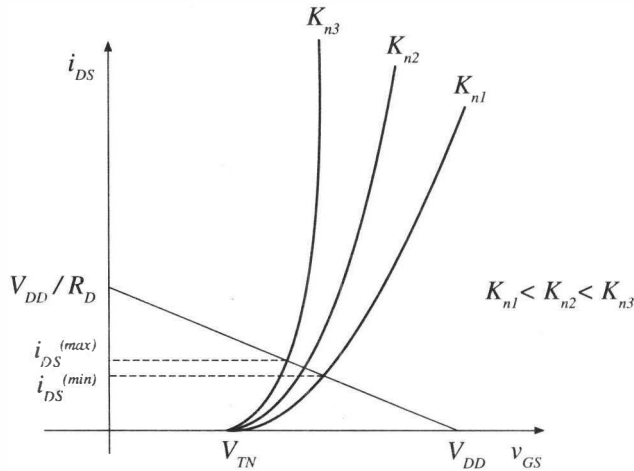


Figura 4.62 Punto di polarizzazione

Si noti che inserendo lo specchio di corrente nel circuito di Figura 4.58, la stabilità del punto di polarizzazione del MOSFET dipende dal punto di lavoro dello specchio di corrente, fissato come mostrato nel grafico di Figura 4.62, analogo al grafico di Figura 4.57. Tuttavia, la differenza non da poco è che nel caso del generatore di corrente la retta di carico può intercettare la caratteristica a valori maggiori di corrente o essere meno ripida. In altre parole è possibile far scorrere più corrente oppure, a parità di corrente, ottenere una migliore stabilità.

4.18 Rete di polarizzazione con resistenza di gate

Un'altra possibile rete di polarizzazione è mostrata in Figura 4.63.

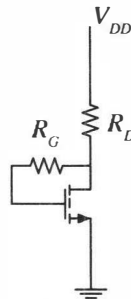


Figura 4.63 Rete di polarizzazione

Poiché la corrente che entra nel gate è nulla la caduta di tensione su R_G è nulla. Gate e drain sono alla stessa tensione e quindi

$$i_{DS} = \frac{V_{DD} - V_{GS}}{R_s} \quad (4.68)$$

Sovrapponendo la retta in (4.68) alla caratteristica del MOSFET in saturazione si ottiene il grafico in Figura 4.64, del tutto analogo di quello in Figura 4.57 e Figura 4.62.

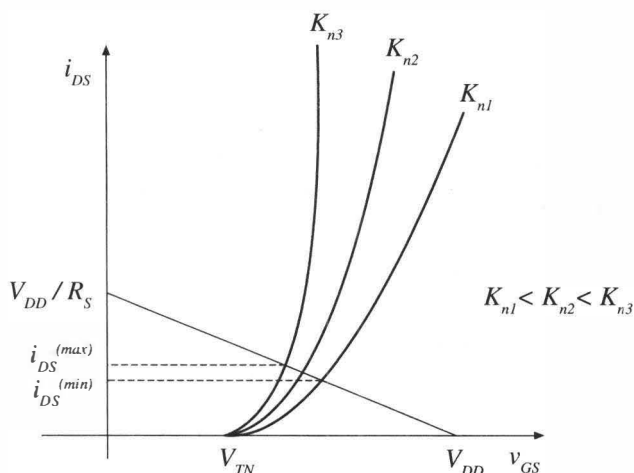


Figura 4.64 Punto di polarizzazione

Si noti che il punto di polarizzazione sarebbe esattamente lo stesso cortocircuitando il gate con il drain, ma la resistenza R_G è necessaria per separare il gate dal drain quando il segnale è immesso nel gate ed estratto dal drain. Se non ci fosse questa resistenza il segnale in ingresso sarebbe uguale a quello di uscita senza alcuna amplificazione.

4.19 MOSFET in saturazione come carico attivo

Nel paragrafo 4.12 abbiamo esaminato un MOSFET in configurazione detta a source comune con carico resistivo. In alternativa al resistore si può usare un secondo MOSFET in saturazione come carico attivo del MOSFET di controllo, come mostrato in Figura 4.65.

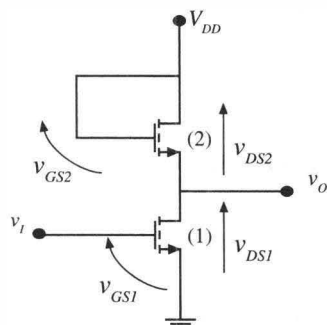


Figura 4.65 MOSFET con carico attivo

Si noti che il drain e il gate del MOSFET impiegato come carico sono cortocircuitati, in tal modo la condizione di saturazione

$$v_{DS2} > v_{GS2} - V_{TN2} \quad (4.69)$$

è sempre verificata.

Tracciamo le caratteristiche e la curva di carico. Per tracciare le caratteristiche di uscita del MOSFET di controllo (1), si osservi che a destra della curva di separazione tra regione lineare e regione di saturazione la corrente è pari a

$$i_{DS1} = \frac{K_{n1}}{2} (v_{GS1} - V_{TN1})^2 \quad (4.70)$$

dove K_{n1} è la costante relativa al MOSFET (1).

La curva di carico è la caratteristica del MOSFET (2) che essendo sempre in saturazione (in quanto D e G sono cortocircuitati) è data da

$$i_{DS2} = \frac{K_{n2}}{2} (V_{DD} - v_{DS1} - V_{TN2})^2 \quad (4.71)$$

dove K_{n2} è la costante relativa al MOSFET (2).

Poiché le due correnti devono essere uguali, possiamo tracciare le curve in Figura 4.66.

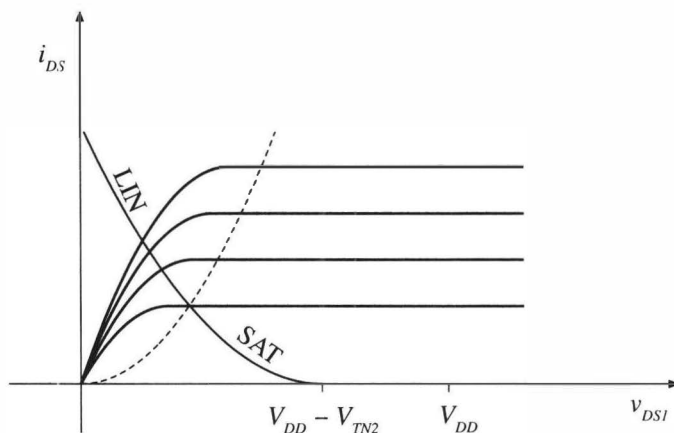


Figura 4.66 Curva di carico sulle caratteristiche di uscita

La curva di carico è un ramo di parabola con vertice nel punto $v_{DS1} = V_{DD} - V_{TN2}$. Al variare dei $v_i = v_{GS1}$ il punto di lavoro si sposta sulla curva di carico da destra verso sinistra e il MOSFET di controllo passa dalla regione di saturazione alla regione lineare.

Vediamo ora l'andamento della funzione di trasferimento. Quando $v_I < V_{TN1}$, il MOSFET di controllo è interdetto, non scorre corrente e $v_O = v_{DS1} = V_{DD}$. Quando v_I supera la soglia il MOSFET di controllo entra in saturazione. Uguagliando la (4.70) e la (4.71) e tenendo presente che $v_{GS1} = v_I$ e $v_{DS1} = v_O$ si ottiene

$$v_O = V_{DD} - V_{TN2} - \sqrt{\frac{K_{n1}}{K_{n2}}}(v_I - V_{TN1}) \quad (4.72)$$

ovvero nel tratto in cui entrambi i MOSFET sono in saturazione la funzione di trasferimento è lineare. L'andamento complessivo è dunque quello di Figura 4.67.

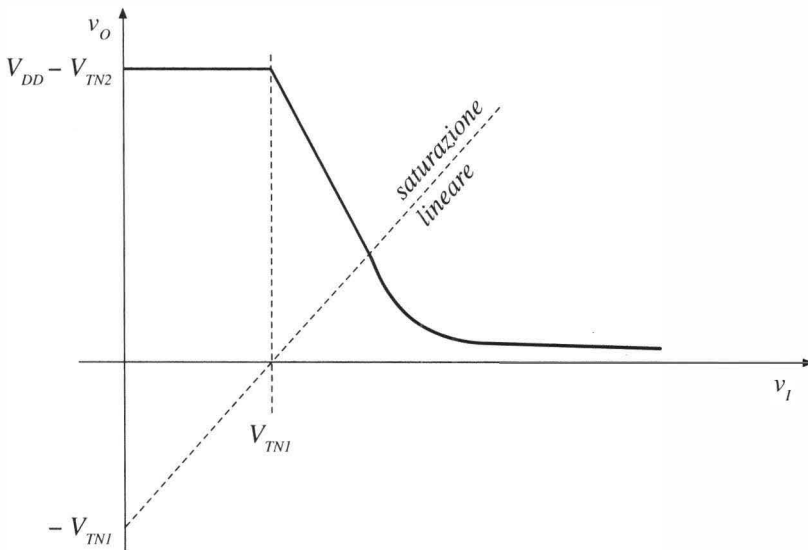


Figura 4.67 Funzione di trasferimento

Il fatto che la funzione di trasferimento in un tratto sia lineare è un aspetto importante di questa configurazione circuitale. In generale un amplificatore di tensione ideale dovrebbe avere funzione di trasferimento perfettamente lineare, ma nella maggior parte degli amplificatori ciò è (quasi) vero solo per piccoli segnali in ingresso. In questo specifico caso, invece, la linearità si estende in un ampio range del segnale.

Impiegando due MOSFET realizzati sullo stesso substrato, quindi con stesso C''_{ox} e μ_n , si ottiene che il guadagno in tensione è dato da

$$A_{vv} = -\sqrt{\frac{K_{n1}}{K_{n2}}} = -\sqrt{\frac{(W/L)_{n1}}{(W/L)_{n2}}} \quad (4.73)$$

Il rapporto W/L di un MOSFET è detto **parametro di forma**. Si noti che per avere un certo guadagno A_{vv} , il MOSFET di controllo deve essere A^2 più largo dell'altro (fissata la lunghezza L).

Questa è una forte limitazione nell'impiego di tale soluzione. Ad esempio per realizzare un amplificatore di guadagno 100 è necessario usare l'area equivalente di 10.000 transistor per realizzare il MOSFET di controllo. Tali considerazioni spiegano perché questa soluzione venga utilizzata solo laddove il guadagno richiesto sia non maggiore di 10-20.

4.20 MOSFET a svuotamento

I MOSFET finora analizzati sono detti ad arricchimento e sono caratterizzati dal fatto che per tensione nulla tra gate e source il transistor costituisce un interruttore aperto (OFF). Una configurazione alternativa prevede la realizzazione di un canale fisico (detto "canale sepolto") di semiconduttore n in substrato p (o viceversa per i MOSFET a canale p).

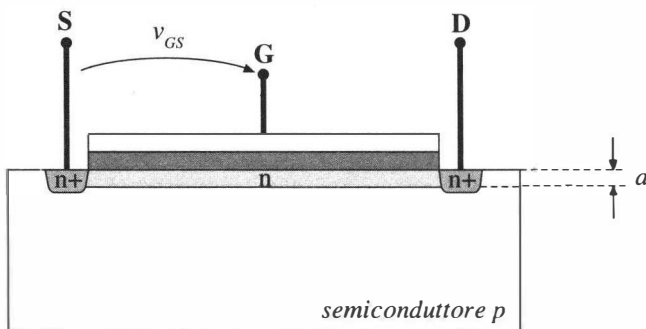


Figura 4.68 MOSFET con canale sepolto

La presenza di questo canale sepolto fa sì che anche per tensione di controllo nulla il transistor sia in conduzione. Per portarlo in interdizione è quindi necessario applicare una tensione negativa.

Consideriamo il caso del MOSFET a canale n: per piccole o nulle correnti, la carica (mobile) accumulata sotto l'ossido è data da

$$Q = -C_{ox}WL(v_{GS} - V_{TN}) - WLaN_D \quad (4.74)$$

con a spessore del canale sepolto e N_D concentrazione delle impurità n nel canale sepolto. La (4.74) può essere riscritta come

$$Q = -C_{ox}WL\left(v_{GS} - \left(V_{TN} - \frac{aN_D}{C_{ox}}\right)\right) \quad (4.75)$$

che mostra come l'effetto del canale sepolto sia spostare la soglia verso valori negativi. La tensione di soglia sarà quindi

$$V_{TN}^{(d)} = V_{TN} - \frac{aN_D}{C_{ox}} \quad (4.76)$$

che, ricordando l'espressione per la capacità dell'ossido per unità di superficie, può essere scritta come

$$V_{TN}^{(d)} = V_{TN} - \frac{N_D ad}{\epsilon} \quad (4.77)$$

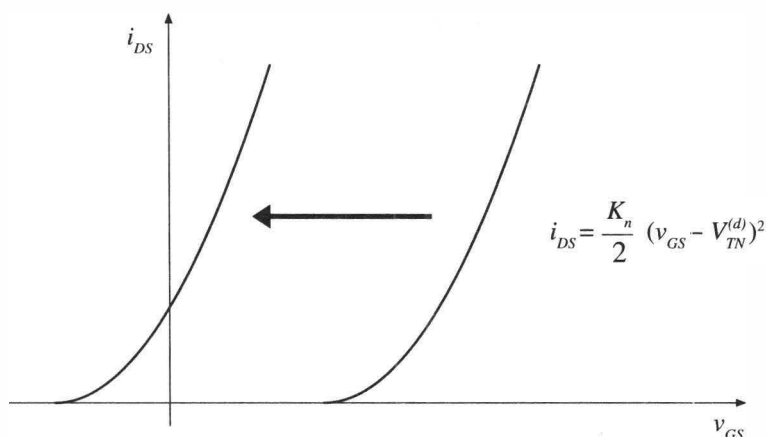


Figura 4.69 Caratteristica di trasferimento per MOSFET a svuotamento in saturazione

A rigore la traslazione della soglia verso sinistra può anche non essere tale da non portare a una soglia negativa. Tuttavia in tal caso, dal punto di vista esterno, il MOSFET a canale n a svuotamento è del tutto indistinguibile da quello ad arricchimento. Pertanto si è soliti riservare il termine “MOSFET a canale n a svuotamento” solo ai MOSFET con soglia negativa.

Per quello che riguarda la condizione che discrimina la regione lineare dalla regione di saturazione, essa è identica al caso del MOSFET n ad arricchimento con l'unica differenza che la soglia è negativa, quindi

$$v_{DS} > v_{GS} + V_{TN}^{(d)} \quad (4.78)$$

4.21 Polarizzazione del MOSFET a svuotamento

Per polarizzare un MOSFET a canale n a svuotamento è necessario fornire al gate una tensione negativa, questo può essere fatto con partitore di tensione purché sia disponibile una alimentazione duale.

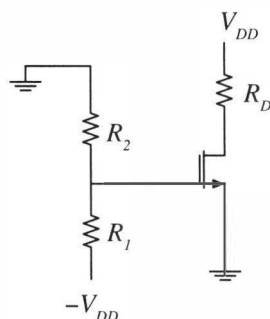


Figura 4.70 Polarizzazione con alimentazione duale del MOSFET a svuotamento

In alternativa può essere usato il circuito di Figura 4.71 per il quale

$$i_{DS} = -\frac{v_{GS}}{R_s} \quad (4.79)$$

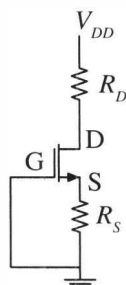


Figura 4.71 Polarizzazione del MOSFET a svuotamento

Tracciando la retta in (4.79) sulla caratteristica del MOSFET in saturazione si ottiene il grafico di Figura 4.72. L'intersezione tra le due curve è il punto di lavoro, che risulta piuttosto stabile al variare delle caratteristiche.

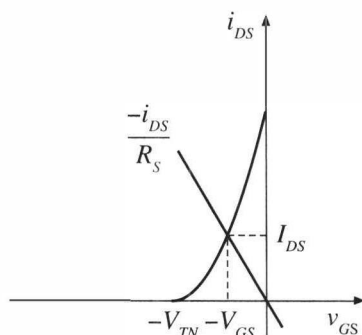


Figura 4.72 Punto di polarizzazione

4.22 MOSFET a svuotamento come carico attivo

Abbiamo già visto un MOSFET in configurazione a source comune che utilizza un ulteriore MOSFET in saturazione come carico attivo. Un'altra possibilità è impiegare come carico attivo un MOSFET a svuotamento come mostrato in Figura 4.73.

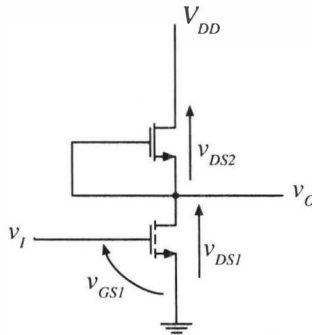


Figura 4.73 MOSFET a svuotamento come carico attivo

Il MOSFET di carico con gate e source cortocircuitati è sempre in conduzione, ed è anche in saturazione se

$$v_{DS2} > V_{TN2} \quad (4.80)$$

con V_{TN2} soglia del MOSFET di carico. Per soglia qui si intende il valore assoluto della v_{GS} quando il MOSFET entra in conduzione positiva.

Tracciamo la curva di carico relativa al secondo MOSFET sulle caratteristiche di uscita del MOSFET di controllo.

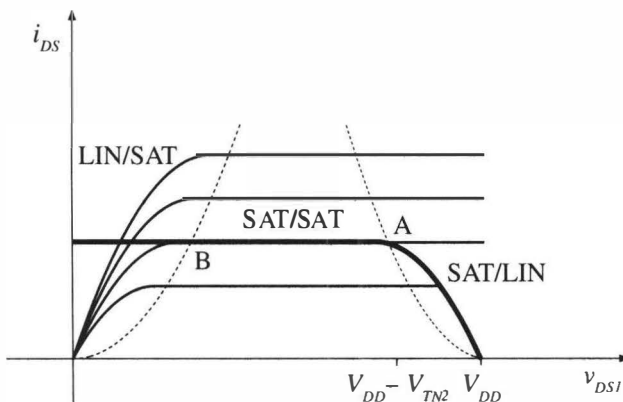


Figura 4.74 Curva di carico

Quando la $v_I = v_{GS1}$ supera la soglia V_{TN1} del MOSFET di controllo il punto di lavoro percorre la curva di carico da destra verso sinistra. Nel primo tratto (fino al punto A)

il MOSFET di controllo è in saturazione e quello di carico è in regione lineare. Nel tratto tra il punto A e B sono entrambi in saturazione e possiamo scrivere

$$i_{DS1} = \frac{K_{n1}}{2} (v_I - V_{TN1})^2 \quad (4.81)$$

$$i_{DS2} = \frac{K_{n2}}{2} (V_{TN2})^2 \quad (4.82)$$

Uguagliando le due correnti, si ottiene il valore di v_I corrispondente al tratto dove entrambi i MOSFET sono in saturazione

$$v_I = \sqrt{\frac{K_{n2}}{K_{n1}}} V_{TN2} + V_{TN1} \quad (4.83)$$

Continuando a seguire la curva di carico oltre il punto B in Figura 4.74, all'aumentare di v_I il MOSFET di controllo entra in regione lineare, mentre quello di carico rimane in saturazione. Il corrispondente andamento della funzione di trasferimento è mostrato Figura 4.75. Si noti il tratto verticale tra il punto A e B (corrispondenti agli analoghi punti A e B nella curva di carico). In generale, nella realizzazione di dispositivi logici una funzione di trasferimento molto ripida è un vantaggio importante. E in effetti il circuito in Figura 4.73 è stato la base di una famiglia logica (la NMOS) molto usata dalla metà degli anni settanta alla metà degli anni ottanta. Una funzione di trasferimento molto ripida è importante anche nelle applicazioni del transistor come amplificatore. In teoria una funzione di trasferimento verticale permetterebbe di realizzare un amplificatore di tensione di guadagno infinito, tuttavia nella realtà le caratteristiche di uscita del MOSFET non sono piatte a causa della modulazione del canale e quindi la funzione di trasferimento non è verticale.

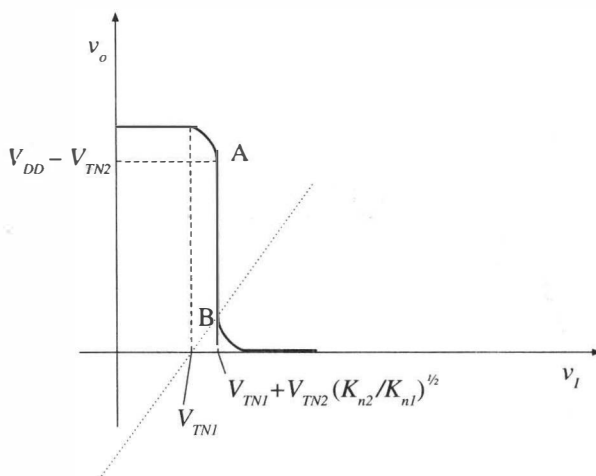


Figura 4.75 Funzione di trasferimento

4.23 MOSFET a canale p ad arricchimento

In alternativa al MOSFET a canale n si può realizzare un MOSFET a canale p su un substrato di semiconduttore n con piazzole di drain e di source drogate p+. In tal caso la corrente nel canale è portata da lacune anziché da elettroni.

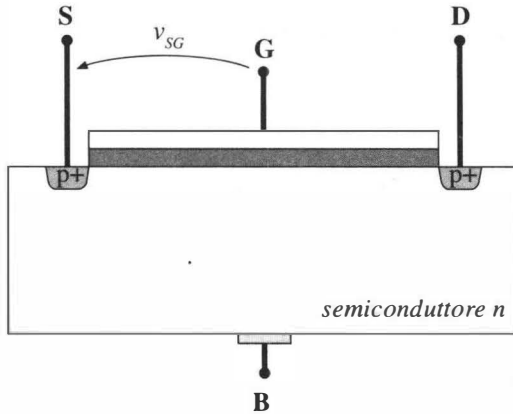


Figura 4.76 MOSFET a canale p

La tensione di controllo è positiva dal source al gate (anziché dal gate al source come nel MOSFET a canale n) e la corrente scorre dal source al drain. Per tutto il resto vale quanto già detto per il MOSFET a canale n. La corrente in regione lineare è data da

$$i_{SD} = K_p \left((v_{SG} - V_{TP})v_{SD} - \frac{1}{2}v_{SD}^2 \right) \quad (4.84)$$

con

$$K_p = \frac{C_{ox}^n \mu_p W}{L} \quad (4.85)$$

La corrente in saturazione è

$$i_{SD} = \frac{K_p}{2} (v_{SG} - V_{TP})^2 \quad (4.86)$$

La condizione per cui il MOSFET passa dalla regione lineare alla saturazione è

$$v_{SD} > v_{SG} - V_{TP} \quad (4.87)$$

4.24 Prospetto riassuntivo dei MOSFET a tre terminali

Poiché i MOSFET a tre terminali possono essere a canale n e a canale p, ad arricchimento e a svuotamento, le possibili configurazioni sono quattro. Ciò rende i transistor MOSFET particolarmente versatili. Per comodità del lettore riassumiamo in questo paragrafo le caratteristiche salienti delle quattro configurazioni.

1) MOSFET a canale n ad arricchimento

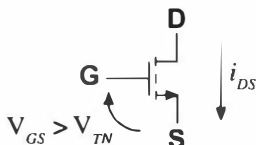


Figura 4.77 n-MOSFET ad arricchimento

La tensione di controllo è applicata dal gate al source e la corrente scorre dal drain al source. Il MOSFET è in conduzione quando la tensione di controllo supera un valore di soglia positivo (V_{TN}). La corrente in saturazione è data da

$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN})^2 \quad (4.88)$$

La condizione affinché il MOSFET sia in saturazione è

$$v_{DS} \geq v_{GS} - V_{TN} \quad (4.89)$$

La corrente in regione lineare è data da

$$i_{DS} = K_n \left((v_{GS} - V_{TN})v_{DS} - \frac{1}{2}v_{DS}^2 \right) \quad (4.90)$$

2) MOSFET a canale n a svuotamento

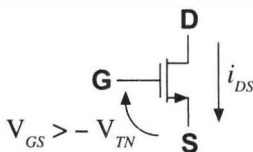


Figura 4.78 n-MOSFET a svuotamento

La tensione di controllo è applicata dal gate al source e la corrente scorre dal drain al source. Il MOSFET è in conduzione quando la tensione di controllo supera un valore di soglia negativo ($-V_{TN}$). La corrente in saturazione è data da

$$i_{DS} = \frac{K_n}{2} (v_{GS} + V_{TN})^2 \quad (4.91)$$

La condizione affinché il MOSFET sia in saturazione è

$$v_{DS} \geq v_{GS} + V_{TN} \quad (4.92)$$

La corrente in regione lineare è data da

$$i_{DS} = K_n \left((v_{GS} + V_{TN}) v_{DS} - \frac{1}{2} v_{DS}^2 \right) \quad (4.93)$$

3) MOSFET a canale p ad arricchimento

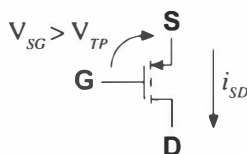


Figura 4.79 p-MOSFET ad arricchimento

La tensione di controllo è applicata dal source al gate e la corrente scorre dal source al drain. Il MOSFET è in conduzione quando la tensione di controllo supera un valore di soglia positivo (V_{TP}). La corrente in saturazione è data da

$$i_{SD} = \frac{K_p}{2} (v_{SG} - V_{TP})^2 \quad (4.94)$$

La condizione affinché il MOSFET sia in saturazione è

$$v_{SD} \geq v_{SG} - V_{TP} \quad (4.95)$$

La corrente in regione lineare è data da

$$i_{SD} = K_p \left((v_{SG} - V_{TP}) v_{SD} - \frac{1}{2} v_{SD}^2 \right) \quad (4.96)$$

4) MOSFET a canale p a svuotamento

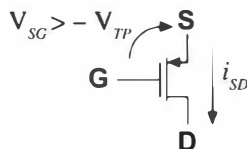


Figura 4.80 p-MOSFET a svuotamento

La tensione di controllo è applicata dal source al gate e la corrente scorre dal source al drain. Il MOSFET è in conduzione quando la tensione di controllo supera un valore di soglia negativo ($-V_{TP}$). La corrente in saturazione è data da

$$i_{SD} = \frac{K_p}{2} (v_{SG} + V_{TP})^2 \quad (4.97)$$

La condizione affinché il MOSFET sia in saturazione è

$$v_{SD} \geq v_{SG} + V_{TP} \quad (4.98)$$

La corrente in regione lineare è data da

$$i_{SD} = K_p \left((v_{SG} + V_{TP}) v_{SD} - \frac{1}{2} v_{SD}^2 \right) \quad (4.99)$$

4.25 MOSFET a quattro terminali

Finora abbiamo sempre supposto che uno dei due terminali attraverso i quali scorre la corrente fosse collegato al body e quindi distinto dall'altro. Questo terminale era detto source. Nel MOSFET a canale n, poiché la corrente scorre dal drain al source, il drain è sempre tenuto a tensione maggiore o uguale a quella del source, in questo modo si è sicuri che entrambe le giunzioni pn (tra source e body e tra drain e body) siano contropolarizzate. Se infatti invertiamo il drain con il source, come in Figura 4.81, tra body e drain è presente una giunzione polarizzata e la corrente vi scorre senza passare attraverso il canale. Il MOSFET diviene semplicemente un diodo.

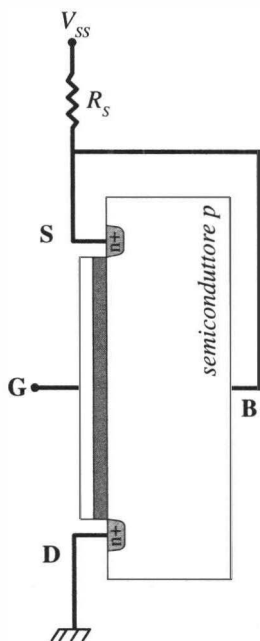


Figura 4.81 MOSFET montato al contrario

Tuttavia ci sono applicazioni per cui è necessario far scorrere la corrente sia da S a D sia viceversa, ad esempio in elettronica digitale per realizzare una connessione che può essere attivata o meno indipendentemente dal verso dell'eventuale corrente che vi può scorrere. In questi casi è necessario usare un MOSFET che abbia source e body indipendenti.

Poiché nell'impiego del MOSFET a 4 terminali non è scontato che entrambe le giunzioni con il body siano contropolarizzate è necessario porre sempre attenzione al fatto che tra drain e body, così come tra source e body, sia effettivamente presente una giunzione pn in contropolarizzazione.

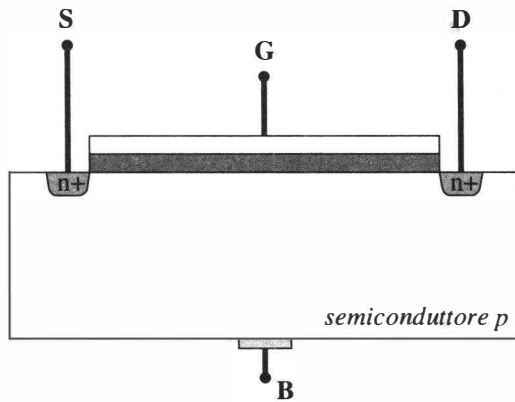


Figura 4.82 MOSFET a quattro terminali

Nel caso del MOSFET a canale n, il body deve essere mantenuto a una tensione più bassa di qualunque tensione applicata al drain o al source; nel caso del MOSFET a canale p il body deve essere mantenuto a una tensione più alta di qualunque tensione applicata al drain o al source. La freccia nel simbolo del MOSFET a 4 terminali ricorda appunto il senso di questa giunzione pn.

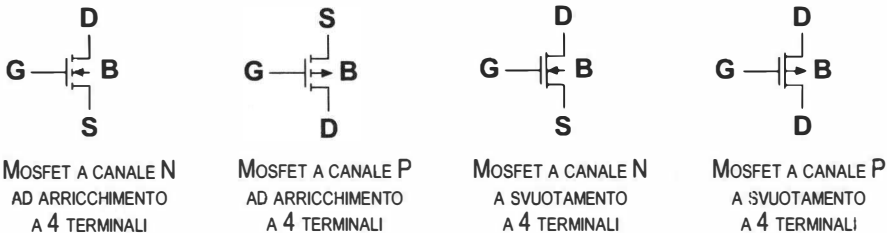


Figura 4.83 Simboli dei MOSFET a quattro terminali

*4.26 Effetto body

Nei MOSFET a 4 terminali la tensione di soglia non è del tutto fissa, ma ha una certa dipendenza dalla tensione tra canale e body di cui talvolta è necessario tenere conto. Questo fenomeno è detto **effetto body**. La tensione tra canale e body quando diversa da zero allarga la zona di svuotamento sotto il gate e ciò fa aumentare (in valore assoluto) la tensione di soglia.

Prima di procedere, però, vediamo in dettaglio cosa succede fisicamente nel canale di un MOSFET (per ora a 3 terminali) quando passa dall'interdizione alla conduzione. Fissata una tensione $v_{DS} > 0$ quando v_{GS} è sotto la soglia il MOSFET è in interdizione e il canale non sussiste. Appena v_{GS} supera la soglia, varrà sicuramente la condizione $v_{DS} > 2(v_{GS} - V_{TN})$ e quindi il MOSFET passa direttamente dall'interdizione alla saturazione. In pratica significa che si forma solo uno spezzone di canale in prossimità del source, mentre il resto del canale è strozzato. All'aumentare della v_{GS} il canale si allunga fino a toccare il drain, quando il MOSFET passa dalla saturazione alla regione lineare. Possiamo quindi dire in generale che in un MOSFET la tensione di controllo è quella tra il gate e il terminale a tensione più bassa (il source), dove inizia a formarsi il canale.

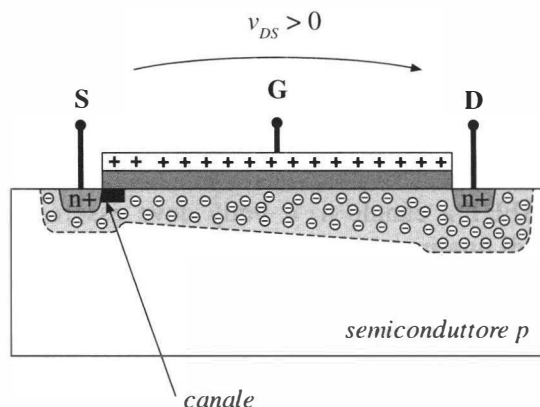


Figura 4.84 Formazione del canale

Premesso ciò, calcoliamo ora il valore della soglia in prossimità del source per un MOSFET a 4 terminali. Per quanto detto sopra, la tensione di controllo che determina la conduzione o l'interdizione è quella tra il gate e il terminale a tensione più bassa (che quindi definiamo univocamente come source). Nel paragrafo 4.3 abbiamo ricavato che la tensione di soglia è in generale data dalla somma di due termini:

$$V_{TN} = 2\varphi_F + \varphi_{ox} \quad (4.100)$$

In definitiva nel MOSFET a canale n a 4 terminali il source è il terminale a tensione più bassa, la tensione di controllo è quella tra gate e source e la corrente scorre dal drain al source. Nel MOSFET a canale p il source è il terminale a tensione più alta, la tensione di controllo è quella tra source e gate e la corrente scorre dal source al drain. Si noti però che source e drain non possono essere etichettati in modo permanente, ma è il senso della corrente a determinare quale dei due è il source e quale il drain.



Figura 4.86 MOSFET a 4 terminali a canale n e a canale p

4.27 JFET

Il principio di funzionamento del JFET (*Junction FET*) è sostanzialmente diverso da quello del MOSFET, anche se i due tipi di transistor hanno molte caratteristiche comuni. Prendiamo in considerazione un JFET a canale n. Su un substrato n vengono realizzate due zone p contrapposte collegate tra loro che costituiscono il gate. Ai due estremi della zona n si collegano i due reofori del drain e del source. Si noti che il JFET è perfettamente simmetrico e la distinzione tra drain e source dipende solo dal senso della corrente che vi scorre: nel JFET a canale n il source è il terminale dal quale esce la corrente (nel JFET a canale p è il contrario).

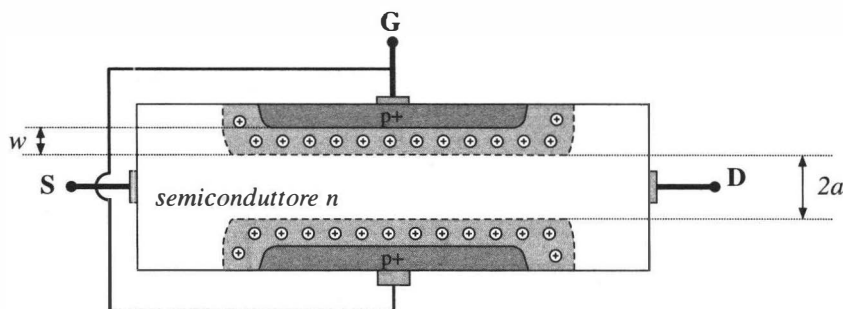


Figura 4.87 JFET

Il principio di funzionamento è del tutto analogo a quello di una valvola idraulica. Applicando una tensione tra drain e source scorre una corrente nel semiconduttore n portata dagli elettroni in conduzione. Quando le giunzioni pn vengono contropolarizzate, la regione di svuotamento, che si estende principalmente nella

zona n che è meno drogata, si allarga finché non occupa l'intera sezione del JFET impedendo la conduzione tra drain e source. Il valore di tensione per cui il JFET è interdetto (la valvola è chiusa) è detta **tensione di pinch-off** (V_p).

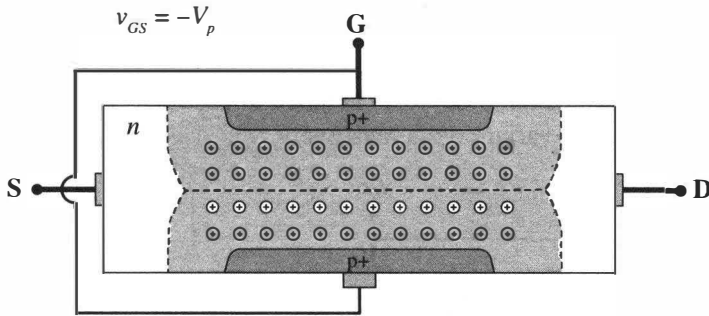


Figura 4.88 JFET in interdizione

Si noti che la tensione tra gate e source deve essere sempre minore di zero affinché la giunzione pn tra gate e substrato rimanga contropolarizzata, quindi per un JFET a canale n

$$-V_p < v_{GS} < 0 \quad (4.107)$$

Per piccole correnti, ovvero per piccole cadute di tensione tra drain e source, il JFET può essere visto semplicemente come un conduttore con sezione controllabile dalla tensione sul gate. Per la legge di Ohm

$$i_{DS} = q\mu_n N_D \frac{2ab}{L} v_{DS} \quad (4.108)$$

dove q è la carica dell'elettrone, μ_n la mobilità degli elettroni, N_D la concentrazione delle impurità n nel canale, $2a$ è lo spessore del canale, b la larghezza del JFET e L la sua lunghezza.

Per correnti non piccole, si verifica un fenomeno analogo allo strozzamento nel MOSFET.

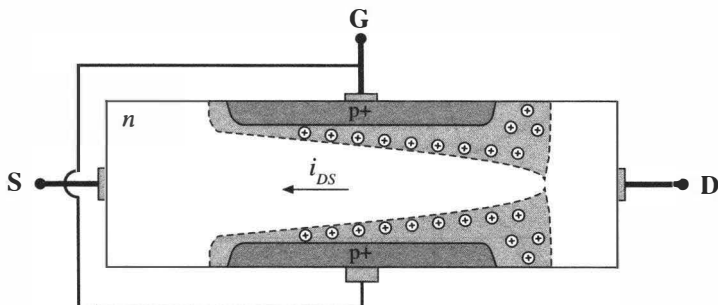


Figura 4.89 JFET per grandi correnti

La caduta di tensione lungo il canale, dal drain al source, causa una tensione diversa da punto a punto ai capi della regione di svuotamento. Questa risulterà minore in prossimità del source e maggiore in prossimità del drain, la larghezza del canale sarà quindi maggiore in prossimità del source e minore in prossimità del drain. All'aumentare della corrente (e della tensione) tra drain e source, in prossimità del drain la regione di svuotamento può arrivare a occludere il canale. A questo punto la situazione è del tutto analoga al caso del MOSFET e la corrente rimane costante al variare della tensione v_{DS} .

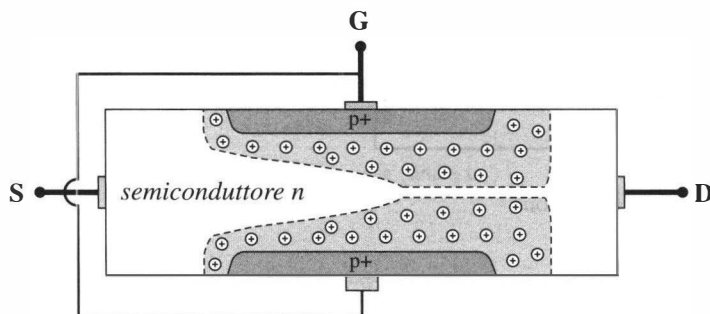


Figura 4.90 JFET in saturazione

Anche nel JFET si distingue quindi una zona lineare (dove la corrente ha andamento parabolico) e una zona di saturazione (dove la corrente è costante).

Per la corrente in saturazione si è soliti usare la seguente formula (che è matematicamente equivalente a quella del MOSFET, anche se espressa con parametri diversi):

$$i_{DS} = I_{DSS} \left(1 + \frac{v_{GS}}{V_p} \right)^2 \quad (4.109)$$

dove I_{DSS} è la massima corrente erogabile dal JFET che si ottiene per $v_{GS} = 0V$.

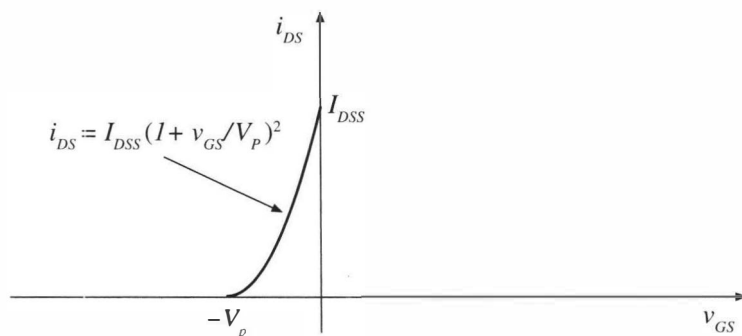


Figura 4.91 Caratteristica di trasferimento del JFET a canale n in saturazione

Per $v_{GS} < -V_p$ il JFET è in interdizione.

L'espressione in zona lineare è

$$i_{DS} = \frac{2I_{DSS}}{V_p^2} \left((v_{GS} + V_p)v_{DS} - \frac{1}{2}v_{DS}^2 \right) \quad (4.110)$$

e la condizione per cui il JFET passa dalla regione lineare alla saturazione è

$$v_{DS} > v_{GS} + V_p \quad (4.111)$$

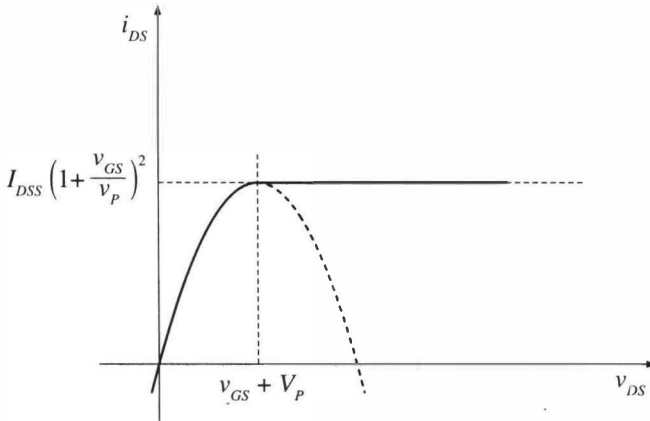


Figura 4.92 Caratteristica di uscita del JFET

Nel JFET a canale n il canale comincia a formarsi dal terminale a più bassa tensione che per questo viene chiamato source. Al contrario nel JFET a canale p il canale comincia a formarsi dal terminale a tensione più alta ed è questo che, pertanto, è chiamato source. Si noti che per il JFET, così come per il MOSFET a 4 terminali, non si può etichettare in maniera permanente un terminale come drain o source, perché dipende dal verso della corrente che vi scorre.



Figura 4.93 JFET a canale n e a canale p

Per un JFET a canale p deve sempre valere

$$-V_p < v_{GS} < 0 \quad (4.112)$$

In saturazione, ovvero per

$$v_{SD} > v_{SG} + V_{TP} \quad (4.113)$$

vale la seguente

$$i_{SD} = I_{SDS} \left(1 + \frac{v_{SG}}{V_p} \right)^2 \quad (4.114)$$

Anche per il JFET a canale p la corrente massima è erogata quando la tensione tra gate e source è nulla.

*4.28 Derivazione della caratteristica del JFET

Vediamo ora in dettaglio come si ricava la caratteristica del JFET. Consideriamo direttamente il caso generale in cui la caduta di tensione dovuta alla corrente tra drain e source non è trascurabile.

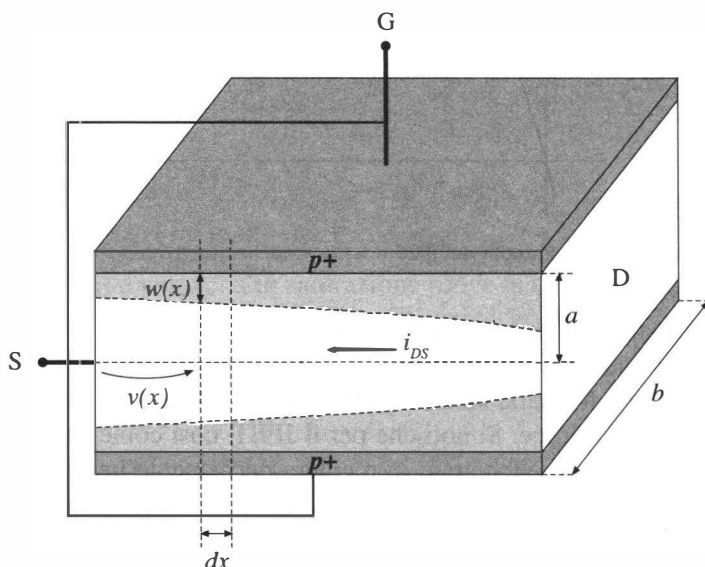


Figura 4.94 JFET

L'ampiezza della regione di svuotamento nella zona n varia lungo x ed è pari a

$$w \approx \sqrt{\frac{2\varepsilon}{q} (\varphi + v_{SG} + v(x)) \frac{1}{N_D}} \quad (4.115)$$

con ε costante dielettrica del silicio, q carica dell'elettrone, φ potenziale di giunzione.

Quando il MOSFET è in interdizione $v_{SG} = V_p$, $v(x) = 0$ e la regione di svuotamento occupa tutto lo spazio disponibile, quindi

$$a = \sqrt{\frac{2\varepsilon}{q}(\varphi + V_p) \frac{1}{N_D}} \quad (4.116)$$

dove a è la semilarghezza del canale.

Per $v_{SG} < V_{TP}$ La carica mobile dQ contenuta in una fetta verticale di spessore infinitesimo dx è pari a

$$dQ = -2(a - w(x))bN_D dx \quad (4.117)$$

con N_D concentrazione delle impurità n nel silicio del canale. Dividiamo entrambi i membri per il tempo infinitesimo dt nel quale le cariche percorrono lo spazio dx .

$$\frac{dQ}{dt} = -2(a - w(x))bN_D \frac{dx}{dt} \quad (4.118)$$

D'altra parte

$$\frac{dQ}{dt} = -i_{DS} \quad (4.119)$$

(il segno negativo è dovuto al fatto che la corrente è definita positiva quando va in senso opposto all'asse x) e

$$\frac{dx}{dt} = v_n = -\mu_n E(x) = \mu_n \frac{dv(x)}{dx} \quad (4.120)$$

con v_n velocità di deriva degli elettroni, μ_n mobilità degli elettroni, E campo elettrico. Sostituendo nella (4.118)

$$i_{DS} = 2N_D \mu_n b(a - w(x)) \frac{dv}{dx} \quad (4.121)$$

Da cui, procedendo integrando come già fatto per il MOSFET, si ottiene

$$i_{DS} = \frac{2N_D \mu_n b}{L} \int_{x=0}^{x=L} (a - w(x)) \frac{dv}{dx} dx \quad (4.122)$$

ovvero

$$i_{DS} = \frac{2N_D \mu_n b}{L} \int_{v=0}^{v=v_{DS}} (a - w(x)) dv \quad (4.123)$$

Calcoliamo ora esplicitamente l'argomento dell'integrale

$$a - w(x) = a - \sqrt{\frac{2\varepsilon}{qN_D}(\varphi + v_{SG} + v(x))} \quad (4.124)$$

da cui, tenendo presente la (4.116), si ottiene

$$a - w(x) = a \left(1 - \frac{\sqrt{\varphi + v_{SG} + v(x)}}{\sqrt{\varphi + V_p}} \right) \quad (4.125)$$

che è equivalente a

$$a - w(x) = a \left(1 - \sqrt{1 + \frac{v_{SG} + v(x) - V_p}{\varphi + V_p}} \right) \quad (4.126)$$

Supponiamo ora che il MOSFET lavori in prossimità del pinch-off, ovvero l'apertura del canale sia piccola rispetto all'intero canale. Questa condizione è verificata nella maggior parte dei casi pratici. In queste condizioni

$$v_{SG} + v(x) - V_p \ll V_p + \varphi \quad (4.127)$$

Quindi sviluppando con Taylor arrestato al primo ordine la radice nella (4.126), si ottiene

$$a - w(x) = \frac{a}{2} \left(\frac{V_p - v_{SG} - v(x)}{\varphi + V_p} \right) \quad (4.128)$$

Sostituendo nell'integrale (4.123)

$$i_{DS} = \frac{N_D \mu_n a b}{L(\varphi + V_p)} \int_{v=0}^{v=v_{DS}} ((V_p - v_{SG}) - v) dv \quad (4.129)$$

da cui

$$i_{DS} = \frac{N_D \mu_n a b}{L(\varphi + V_p)} \left((V_p - v_{SG}) v_{DS} - \frac{1}{2} v_{DS}^2 \right) \quad (4.130)$$

che è una parabola rovesciata avente vertice in

$$v_{DS} = V_p - v_{SG} \quad (4.131)$$

La corrente di saturazione si ottiene sostituendo la (4.131) nella (4.130) e risulta pari a

$$i_{DS} = I_{DSS} \left(1 + \frac{v_{GS}}{V_p} \right)^2 \quad (4.132)$$

dove

$$I_{DSS} = \frac{\mu_n N_D a b}{2L} \frac{V_p^2}{V_p + \varphi} \quad (4.133)$$

Come è ragionevole la corrente aumenta all'aumentare della sezione del JFET (ab) e diminuisce all'aumentare della sua lunghezza L .

4.29 La travagliata nascita del transistor a effetto campo

L'inizio del Novecento fu l'epoca delle lampadine a incandescenza. Nel 1904 l'inglese John Ambrose Fleming inventò il diodo a tubo a vuoto: esso era costituito da un filamento riscaldato che liberava elettroni per effetto termoionico e da un anodo che li attirava a sé mediante un campo elettrico. Gli elettroni potevano andare solo dal filamento (catodo) all'anodo attraversando una regione dove era fatto il vuoto pneumatico.

Il passo successivo lo compì nel 1906 l'americano Lee De Forest: inserì una griglia tra anodo e catodo in modo da modulare il flusso di elettroni e di fatto realizzò un dispositivo in grado di commutare e pure di amplificare. De Forest lo chiamò *audion*; solo successivamente prevalse il nome triodo per associazione con diodo. Tra l'altro, De Forest oltre che scienziato fu un noto regista e produttore cinematografico. È sua anche l'invenzione dei film sonori e della reazione negativa per la stabilizzazione degli amplificatori.

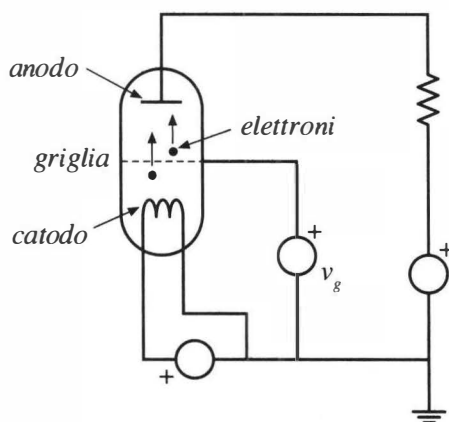


Figura 4.95 Triodo

Fino agli anni sessanta il triodo (noto anche come valvola termoionica) è stato il dispositivo fondamentale dell'elettronica. Ma i suoi limiti erano evidenti: ingombro, fragilità, elevato consumo elettrico e velocità di commutazione piuttosto limitata dalle sue stesse dimensioni fisiche che davano luogo a grandi capacità parassite.

Per questo molti inventori e centri di ricerca lavorarono fin da subito allo sviluppo di un dispositivo a stato solido. Già nel 1925 l'austriaco Julius Edgar Lilienfeld brevettò una versione a stato solido del triodo. Non realizzò mai la sua invenzione, ma l'idea base dell'odierno MOSFET cominciò a circolare. Nel 1936 Marvin Kelly, il direttore dei laboratori Bell, incaricò William Shockley di

studiare la possibilità di realizzare un interruttore elettronico a stato solido che sostituisse gli interruttori meccanici allora in uso. La guerra interruppe il progetto di ricerca prima di portare a qualche risultato (i laboratori Bell furono pesantemente impegnati nello sforzo bellico), ma subito dopo, nel 1945, Shockley fu messo a capo di un gruppo costituito da John Bardeen e Walter Brattain con lo stesso obiettivo. L'idea che i tre perseguivano era un transistor basato sul principio dell'effetto campo, come gli odierni JFET. Tuttavia non ebbero successo.

Shockley decise quindi di occuparsi di altro, lasciando i due colleghi a studiare la ragione del fallimento. Bardeen ebbe l'intuizione che la causa dell'insuccesso fossero gli stati trappola superficiali, ma ciò non aiutava a trovare una soluzione. Nel dicembre 1947 accadde un evento fortunato. In uno dei tanti esperimenti, Brattain, inavvertitamente, con il filo elettrico del reoforo ruppe la metallizzazione e penetrò lo strato di ossido sottostante.

Con la terminologia moderna, ruppe il gate e inserì la punta elettrica direttamente nel canale. I due non si accorsero del danno ed eseguirono l'esperimento. Inaspettatamente il transistor funzionò. Capirono subito che avevano inventato un dispositivo completamente diverso da quello che cercavano di realizzare, un dispositivo che funzionava per iniezione diretta di portatori, non per modulazione del canale.

Shockley rimase affranto dalla notizia: aveva perso l'occasione della sua vita, ma si mise a lavorare duro per migliorare il dispositivo inventato dai colleghi. Il transistor di Brattain e Bardeen sfruttava le giunzioni metallo-semiconduttore formate da sottili e delicati fili elettrici premuti sul semiconduttore. Nel 1949 Shockley realizzò il primo transistor bipolare come oggi lo intendiamo, ovvero basato sulle giunzioni pn.

Nel 1956 i tre scienziati, Bardeen, Brattain e Shockley, furono premiati con il Nobel in fisica come coinventori del transistor bipolare. In effetti, può sembrare curioso come un dispositivo dal principio di funzionamento piuttosto complesso, il BJT, sia stato realizzato prima di un dispositivo equivalente concettualmente più semplice (il FET), ma la ragione è, appunto, storica e tecnologica. La realizzazione del MOSFET fu inizialmente impedita dalla presenza degli stati trappola superficiali, l'errore sperimentale di Brattain aggravò il problema. L'immediato successo del BJT rallentò l'ulteriore sviluppo dell'idea del FET. Pertanto solo nel 1960 l'egiziano Martin "John" M. Atalla e il coreano Dawon Khang, presso i laboratori Bell, realizzarono il primo MOSFET funzionante.

Ma i laboratori Bell sottovalutarono le potenzialità di questo dispositivo, che appariva in ritardo di dieci anni rispetto al BJT e lo lasciarono sviluppare dalla RCA e dalla Fairchild. Ci vollero ancora un paio di decenni prima che il MOSFET si affermasse come il transistor più utilizzato.

4.30 Riepilogo

- I MOSFET possono essere a canale n o a canale p, a tre terminali o a quattro terminali.
- In un MOSFET a canale n a tre terminali, quando la tensione tra gate (G) e source (S) è maggiore di un valore di soglia (V_{TN}), si crea un canale di cariche negative mobili tra drain (D) e source (S) che permette alla corrente di scorrere. Al contrario, quando $v_{GS} < V_{TN}$ non scorre corrente e il transistor è detto in regione di interdizione.
- In un MOSFET a canale n a tre terminali, quando $v_{GS} > V_{TN}$ e $v_{DS} < v_{GS} - V_{TN}$ (regione lineare), la corrente i_{DS} è data da:

$$i_{DS} = K_n \left((v_{GS} - V_{TN})v_{DS} - \frac{1}{2}v_{DS}^2 \right)$$

con K_n costante del transistor.

- In un MOSFET a canale n a tre terminali, quando $v_{GS} > V_{TN}$ e $v_{DS} > v_{GS} - V_{TN}$ (saturazione), la corrente i_{DS} è data da:

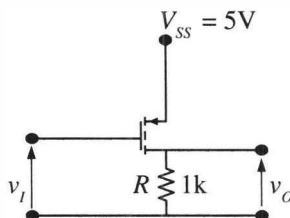
$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN})^2$$

con K_n costante del transistor.

- In un MOSFET a canale n in saturazione la corrente è solo approssimativamente costante, in realtà aumenta leggermente all'aumentare della tensione tra drain e source per effetto della "modulazione del canale".
- Un MOSFET che lavora come interruttore commuta tra la regione lineare e la regione di interdizione passando per la saturazione.
- Un MOSFET polarizzato in saturazione può operare come amplificatore.
- Un JFET a canale n è in interdizione quando $v_{GS} < -V_P$.
- Un JFET a canale n è in conduzione quando $-V_P < v_{GS} < 0$.

4.31 Esercizi svolti

1) Tracciare la funzione di trasferimento del circuito in figura. Il MOSFET a canale p ad arricchimento ha $V_{TP} = 1\text{V}$ e $K_p = 10\text{mA/V}^2$.



Per $v_I > V_{SS} - V_{TP} = 4\text{V}$ il MOSFET è interdetto e $v_O = 0\text{V}$.

Per $v_I < 4\text{V}$ il MOSFET è in conduzione, resta da vedere quando è in regione lineare e quando in saturazione. La condizione per la saturazione è

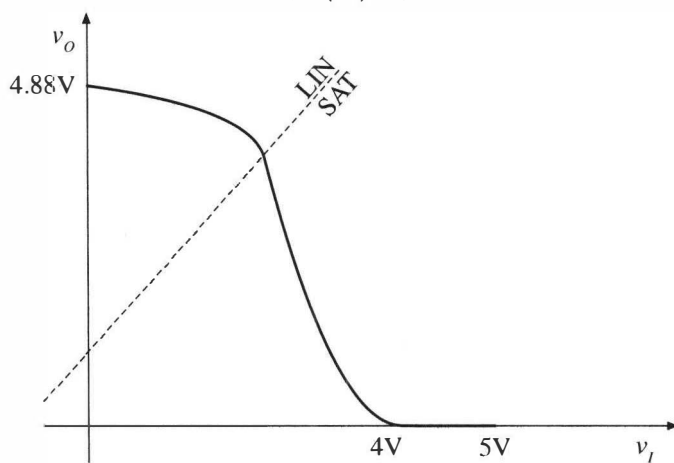
$$v_{SD} > v_{SG} - V_{TP}$$

ovvero

$$V_{SS} - v_O > V_{SS} - v_I - V_{TP}$$

$$v_O < v_I + V_{TP}$$

$$v_O < (1\text{V}) + v_I$$



Calcoliamo ora il valore di v_O per $v_I = 0\text{V}$. Supponiamo il MOSFET nella regione per cui $v_{SD} \ll 2(v_{SG} - V_{TP})$

$$R_{ON} = \frac{1}{K_p (v_{SG} - V_{TP})} = 25\Omega$$

quindi

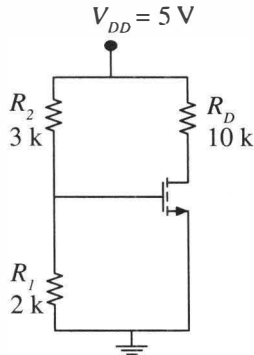
$$v_O = \frac{R}{R + R_{ON}} V_{SS} = 4.88\text{V}$$

Verifichiamo che per $v_i = 0V$

$$v_{SD} \ll 2(v_{SG} - V_{TP})$$

$$0.12V \ll 8V$$

2) Con riferimento alla figura, determinare la corrente di drain del transistor n-MOSFET. $V_{TN} = 1V$, $K_n = 10mA/V^2$.



La tensione sul gate è

$$V_G = \frac{R_1}{R_1 + R_2} V_{DD} = 2V$$

Supponiamo il MOSFET in saturazione

$$i_{DS} = \frac{K_n}{2} (V_G - V_{TN})^2 = 5mA$$

Verifichiamo che il MOSFET sia effettivamente in saturazione, la condizione è

$$V_{DS} > V_{GS} - V_{TN}$$

$$V_{DD} - i_{DS} R_D > V_G - V_{TN}$$

$$-45V > 1V$$

Purtroppo la condizione non è verificata e il MOSFET è in regione lineare. Dovremmo a questo punto considerare l'equazione della parabola, ma prima proviamo a supporre di essere nel tratto della caratteristica per cui

$$V_{DS} \ll 2(V_{GS} - V_{TN})$$

Se vale questa condizione, è come se tra drain e source ci fosse una resistenza R_{ON} pari a

$$R_{ON} = \frac{1}{K_n (V_{GS} - V_{TN})} = 100\Omega$$

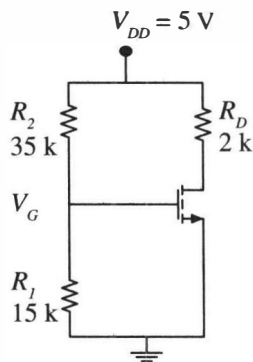
e quindi

$$V_{DS} = \frac{R_{ON}}{R_{ON} + R_D} V_{DD} = 50\text{mV} \ll 2(V_{GS} - V_{TN}) = 2\text{V}$$

La condizione è quindi ben verificata. Possiamo quindi calcolare i_{DS} come

$$i_{DS} = \frac{V_{DD}}{R_D + R_{ON}} \approx \frac{V_{DD}}{R_D} = 0.5\text{mA}$$

3) Con riferimento a circuito in figura si calcoli la corrente i_{DS} con $K_n = 10\text{mA/V}^2$ e $V_{TN} = 1\text{V}$. Si calcoli, inoltre, come varia la corrente i_{DS} al variare di K_n (ovvero si calcoli la derivata di i_{DS} rispetto a K_n nel punto di lavoro fissato nell'esercizio precedente.) Si calcoli inoltre come varia i_{DS} al variare di V_{TN} .



La tensione sul gate è

$$V_G = \frac{R_1}{R_1 + R_2} V_{DD} = 1.5\text{V}$$

Supponiamo il MOSFET in saturazione

$$i_{DS} = \frac{K_n}{2} (V_G - V_{TN})^2 = 1.25\text{mA}$$

Verifichiamo che il MOSFET sia effettivamente in saturazione, la condizione è

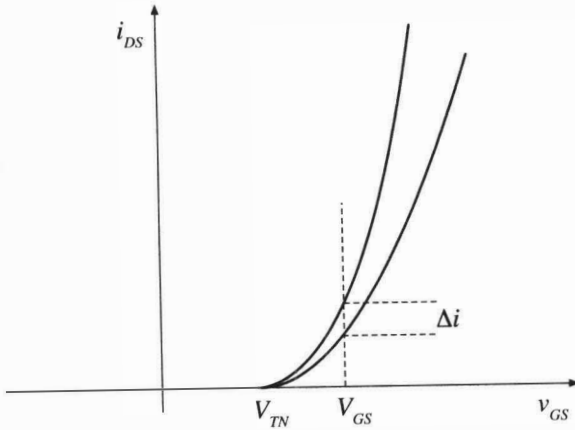
$$\begin{aligned} V_{DS} &> V_{GS} - V_{TN} \\ V_{DD} - i_{DS} R_D &> V_G - V_{TN} \\ 2.5\text{V} &> 0.5\text{V} \end{aligned}$$

Passiamo ora al calcolo delle variazioni di i_{DS} al variare di K_n e V_{TN} . Per il MOSFET in saturazione vale la seguente

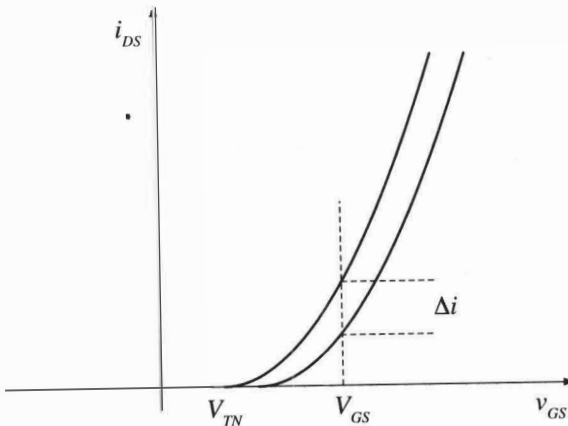
$$i_{DS} = \frac{K_n}{2} (V_G - V_{TN})^2$$

da cui

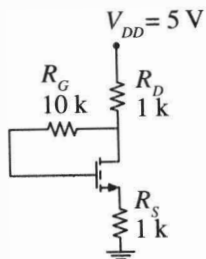
$$\frac{di_{DS}}{dK_n} = \frac{1}{2} (V_G - V_{TN})^2 = 0.125 \text{ V}^2$$



$$\frac{di_{DS}}{dV_{TN}} = -K_n (V_G - V_{TN}) = -5 \text{ mA/V}$$



4) Con riferimento alla circuito in figura si calcoli la I_{DS} . $K_n = 2 \text{ mA/V}^2$, $V_{TN} = 1 \text{ V}$.

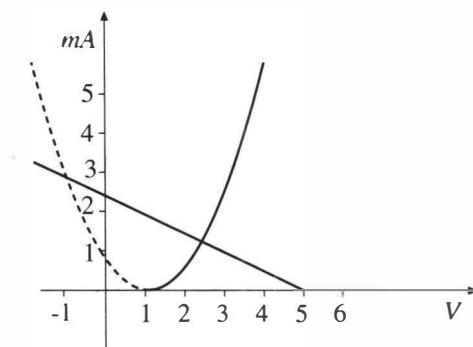


La corrente che scorre su R_G è nulla, quindi G e D sono allo stesso potenziale e il MOSFET è sempre in saturazione. Il punto di polarizzazione di questo circuito è l'intersezione delle curve

$$i_{DS} = \frac{V_{DD} - v_{GS}}{R_S + R_D}$$

$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN})^2$$

che può essere trovata approssimativamente per via grafica



Per il calcolo esatto è necessario risolvere l'equazione di secondo grado ricordando che

$$ax^2 + bx + c = 0$$

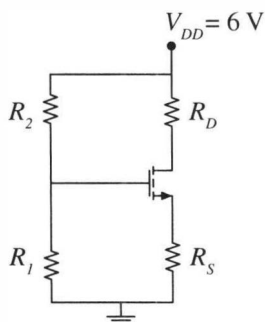
ha soluzioni date da

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Nel caso specifico le soluzioni sono: -0.69V e 2.19V . La prima soluzione è da scartare perché si riferisce al ramo sinistro della parabola, come si vede dal grafico tracciato sopra. Verifichiamo ora che effettivamente il MOSFET sia in saturazione, come supposto inizialmente

$$v_{DS} = V_{DD} - R_S I_{DS} - R_D I_{DS} = 2\text{V}$$

5) Si progetti la rete di polarizzazione in figura in modo che $I_{DS} = 2\text{mA}$. Le specifiche del MOSFET sono $K_n = 10\text{mA/V}^2$, $V_{TN} = 1\text{V}$.



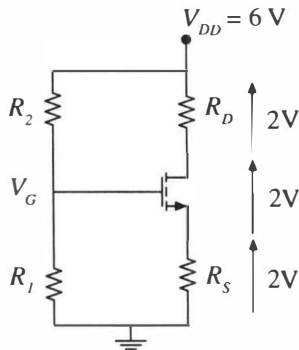
Poiché

$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN})^2$$

segue che

$$v_{GS} = V_{TN} + \sqrt{\frac{2I_{DS}}{K_n}} = 1.63\text{V}$$

per proseguire è necessario fare delle scelte sulle tensioni nella maglia di uscita, imponiamo le tensioni come nella figura che segue



Prima di procedere è necessario verificare che per le tensioni scelte il MOSFET sia in saturazione:

$$v_{DS} > v_{GS} - V_{TN}$$

$$2\text{V} > 0.63\text{V}$$

Passiamo quindi al dimensionamento delle resistenze:

$$R_S = R_D = \frac{2\text{V}}{2\text{mA}} = 1\text{k}\Omega$$

Per dimensionare il partitore $R_1 - R_2$ è necessario scegliere il valore della corrente che vi scorre. Scegliamo $I_p = 10\mu\text{A}$. Si ottiene

$$R_1 = \frac{V_G}{I_p} = \frac{2\text{V} + 1.63\text{V}}{10\mu\text{A}} = 363\text{k}\Omega$$

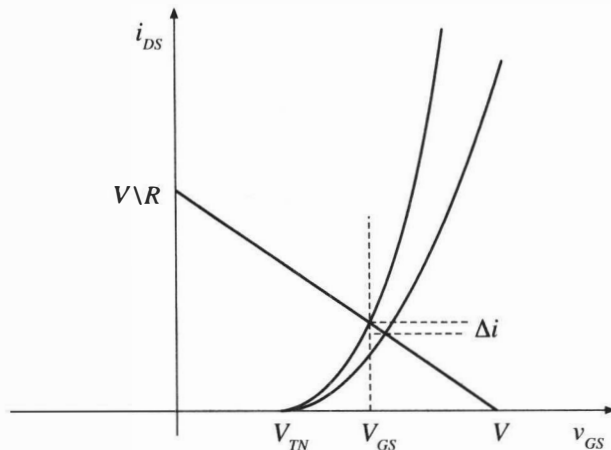
$$R_2 = \frac{V_{DD} - V_G}{I_p} = \frac{6\text{V} - 3.63\text{V}}{10\mu\text{A}} = 237\text{k}\Omega$$

***6)** Con riferimento all'esercizio precedente si calcoli come varia la corrente i_{DS} al variare di K_n (ovvero si calcoli la derivata di i_{DS} rispetto a K_n nel punto di lavoro fissato nell'esercizio precedente.) Si calcoli inoltre come varia i_{DS} al variare di V_{TN} .

La condizione da rispettare è

$$\frac{K_n}{2} (v_{GS} - V_{TN})^2 = \frac{V_G - v_{GS}}{R_S}$$

dove $V_G = 3.63\text{V}$ e $R_S = 1\text{k}\Omega$.



Deriviamo entrambi i membri rispetto a K_n .

$$\frac{1}{2} (v_{GS} - V_{TN})^2 + K_n (v_{GS} - V_{TN}) \frac{dv_{GS}}{dK_n} = -\frac{1}{R} \frac{dv_{GS}}{dK_n}$$

da cui

$$\frac{dv_{GS}}{dK_n} = -\frac{\frac{1}{2} (v_{GS} - V_{TN})^2}{K_n (v_{GS} - V_{TN}) + \frac{1}{R}}$$

D'altra parte

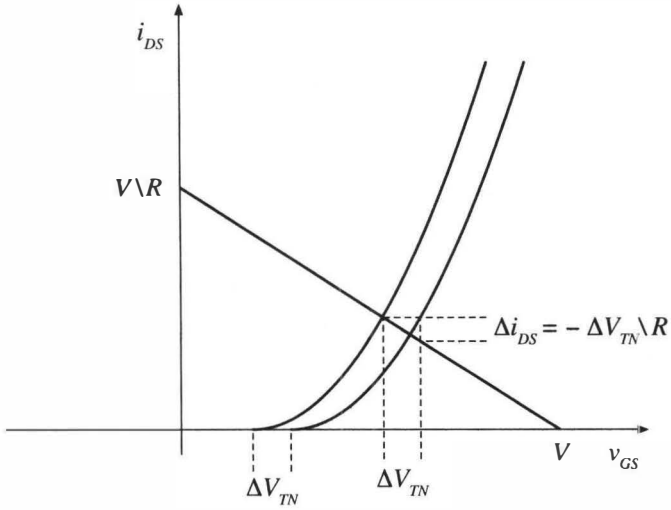
$$\frac{di_{DS}}{dK_n} = \frac{di_{DS}}{dv_{GS}} \frac{dv_{GS}}{dK_n} = -\frac{1}{R_S} \frac{dv_{GS}}{dK_n}$$

quindi

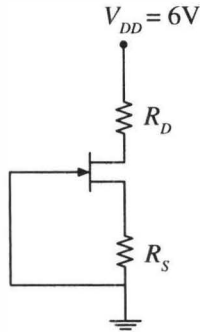
$$\frac{di_{DS}}{dv_{GS}} = \frac{\frac{1}{2} (v_{GS} - V_{TN})^2}{R_S K_n (v_{GS} - V_{TN}) + 1} = 0.027\text{V}^2$$

Per il calcolo delle sensibilità rispetto alle variazioni di V_{TN} con riferimento alla figura che segue possiamo direttamente scrivere

$$\frac{di_{DS}}{dv_{GS}} = \frac{\Delta i_{DS}}{\Delta V_{TN}} = -\frac{1}{R_S} = -1\text{mA/V}$$



7) Si progetti la rete di polarizzazione in figura in modo che $I_{DS} = 5\text{mA}$. Le specifiche del JFET sono $I_{DSS} = 30\text{mA}$, $V_P = 2\text{V}$.



Poiché

$$i_{DS} = I_{DSS} \left(1 + \frac{v_{GS}}{V_P} \right)^2$$

segue che

$$v_{GS} = V_P \left(\sqrt{\frac{I_{DS}}{I_{DSS}}} - 1 \right) = -1.18\text{V}$$

Quindi

$$R_S = \frac{v_{SG}}{I_{DS}} = 236\Omega$$

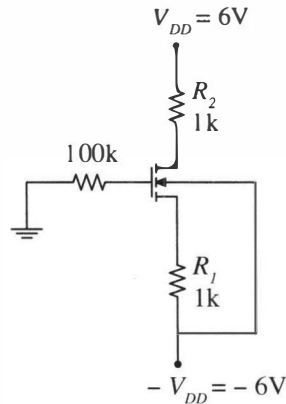
Per il dimensionamento di R_D , dobbiamo scegliere V_{DS} con il vincolo che il JFET sia in saturazione. La condizione da rispettare è

$$v_{DS} > V_P + v_{GS} = 0.82V$$

Scegliamo $v_{DS} = 3V$. Si ottiene

$$R_D = \frac{V_{DD} - 3V - 1.18V}{5mA} = 364\Omega$$

***8) Si calcoli la corrente i_{DS} del circuito in figura. Le caratteristiche del MOSFET sono $V_{TN0} = 1V$, $K_n = 4mA/V^2$, $\varphi_F = 0.2V$, $\gamma = 0.2V^{1/2}$.**



Si tratta di un MOSFET a 4 terminali quindi

$$V_{TN} = V_{TN0} + \gamma \left(\sqrt{2\varphi_F + v_{SB}} - \sqrt{2\varphi_F} \right)$$

Cominciamo supponendo $V_{TN} = V_{TN0}$, il punto di polarizzazione si calcola come intersezione delle curve

$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN0})^2$$

$$i_{DS} = \frac{V_{DD} - v_{GS}}{R_1}$$

Uguagliando le due espressioni e risolvendo l'equazione di secondo grado nell'incognita v_{GS} , si ottengono le due soluzioni: $-0.85V$ e $2.35V$; la prima delle due è da scartare perché si riferisce al ramo sinistro della parabola.

Per questo valore di v_{GS}

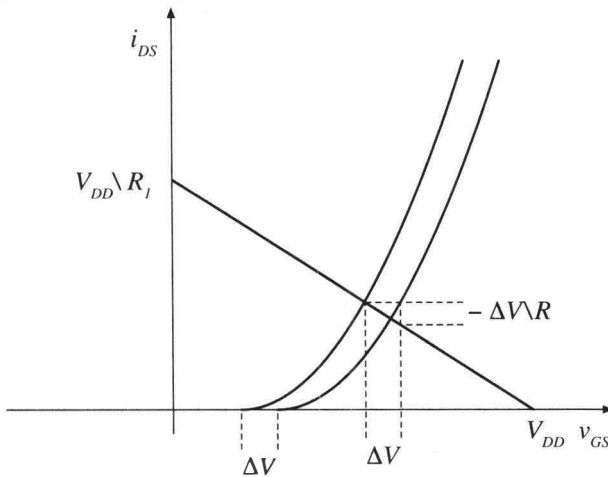
$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN0})^2 = 3.65mA$$

Possiamo ora calcolare la correzione alla soglia dovuta all'effetto body:

$$\Delta V = V_{TN} - V_{TN0} = \gamma \left(\sqrt{2\varphi_F + (V_{DD} - v_{GS})} - \sqrt{2\varphi_F} \right) = 0.28V$$

Se ora supponiamo piccola la correzione, possiamo approssimare linearmente come mostrato in figura e si ottiene

$$\Delta i = -\frac{\Delta V}{R_1} = -0.28mA$$

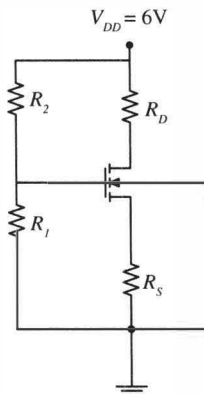


Quindi $i_{DS} = 3.65mA - 0.28mA = 3.37mA$.

Verifichiamo, infine, che il MOSFET sia effettivamente in saturazione, ovvero

$$\begin{aligned} v_{DS} &> v_{GS} - V_{TN} \\ 12 - (1k\Omega)(3.37mA) - (1k\Omega)(3.37mA) &> 2.35V \\ 5.26V &> 2.35V \end{aligned}$$

9) Si progetti la rete di polarizzazione in figura in modo che $i_{DS} = 1mA$. Le caratteristiche del MOSFET sono $V_{TN0} = 1V$, $K_n = 4mA/V^2$, $\varphi_F = 0.2V$, $\gamma = 0.5V^{1/2}$.



Per progettare il circuito dobbiamo fare delle scelte sulle tensioni della maglia di uscita. Poniamo la caduta di tensione su R_s pari a 1V da cui segue che $R_s = 1\text{k}\Omega$:

$$V_{TN} = V_{TN0} + \gamma \left(\sqrt{2\varphi_F + v_{SB}} - \sqrt{2\varphi_F} \right) = 1\text{V} + (0.5) \left(\sqrt{0.4 + 1} - \sqrt{0.4} \right) \text{V} = 1.28\text{V}$$

Possiamo quindi calcolare la v_{GS} come

$$v_{GS} = V_{TN} + \sqrt{\frac{2I_{DS}}{K_n}} = 1.99\text{V}$$

Quindi $V_G = 2.99\text{V}$. Fissiamo ora la corrente del partitore, ad esempio $10\mu\text{A}$. Segue che

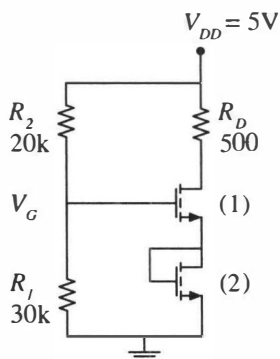
$$R_1 = \frac{2.99\text{V}}{10\mu\text{A}} = 299\text{k}\Omega$$

$$R_2 = \frac{6\text{V} - 2.99\text{V}}{10\mu\text{A}} = 301\text{k}\Omega$$

Fissiamo ora la caduta di tensione su R_D , ad esempio 2V. Segue che $R_D = 2\text{k}\Omega$. Resta ora da verificare che il MOSFET sia in saturazione, ovvero

$$\begin{aligned} v_{DS} &> v_{GS} - V_{TN} \\ 3\text{V} &> 0.71\text{V} \end{aligned}$$

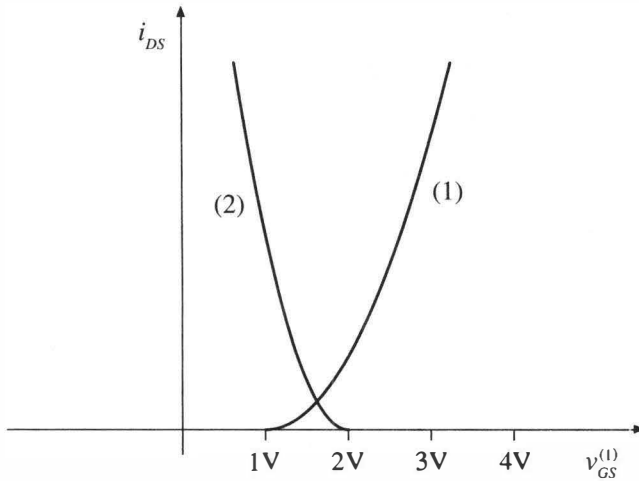
10) Si calcoli la I_{DS} del circuito in figura. $V_{TN1} = V_{TN2} = 1\text{V}$, $K_{n1} = 100\text{mA/V}^2$, $K_{n2} = 10\text{mA/V}^2$.



Se i due MOSFET sono in saturazione:

$$i_{DS} = \frac{K_n^{(1)}}{2} (v_{GS}^{(1)} - V_{TN}^{(1)})^2$$

$$i_{DS} = \frac{K_n^{(2)}}{2} (V_G - v_{GS}^{(1)} - V_{TN}^{(2)})^2$$

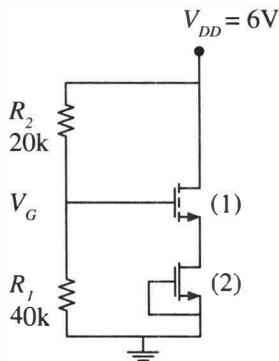


Uguagliando le due espressioni si ottiene $v_{GS}^{(1)} = 1.24\text{V}$, quindi $i_{DS} = 2.88\text{mA}$. Verifichiamo ora che il MOSFET 1 sia in saturazione (il MOSFET 2 lo è sicuramente avendo drain e gate cortocircuitati):

$$v_{DS}^{(1)} > v_{GS}^{(1)} - V_{TN}^{(1)}$$

$$1.8\text{V} > 0.24\text{V}$$

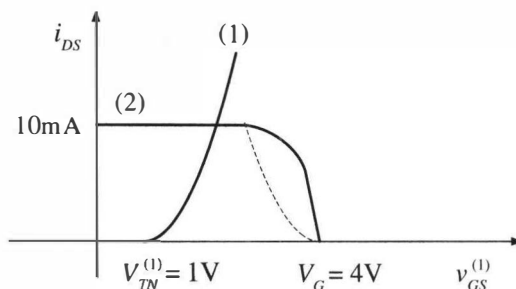
11) Si calcoli la I_{DS} del circuito in figura. $V_{TN1} = V_{TN2} = 1\text{V}$, $K_{n1} = 10\text{mA/V}^2$, $K_{n2} = 20\text{mA/V}^2$.



Se i due MOSFET sono in saturazione:

$$i_{DS} = \frac{K_n^{(1)}}{2} (v_{GS}^{(1)} - V_{TN}^{(1)})^2$$

$$i_{DS} = \frac{K_n^{(2)}}{2} (V_{TN}^{(2)})^2 = 10\text{mA}$$



Uguagliando le due espressioni si ottiene $v_{GS}^{(1)} = 2.4V$. Resta da controllare che per questo valore di $v_{GS}^{(1)}$ i due MOSFET siano in saturazione. Per il MOSFET (1)

$$v_{DS}^{(1)} > v_{GS}^{(1)} - V_{TN}^{(1)}$$

$$V_{DD} - (V_G - v_{GS}^{(1)}) > v_{GS}^{(1)} - V_{TN}^{(1)}$$

$$4.4V > 1.4V$$

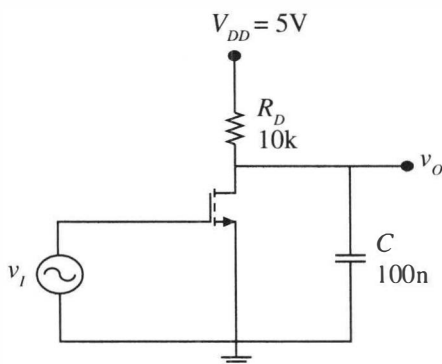
Per il MOSFET (2)

$$v_{DS}^{(2)} > V_{TN}^{(2)}$$

$$V_G - v_{GS}^{(1)} > V_{TN}^{(1)}$$

$$1.6V > 1V$$

12) Si consideri lo schema in figura. Si consideri in ingresso un fronte molto ripido tra 5V e 0V. Calcolare il tempo che impiega l'uscita a raggiungere il 90% del valore massimo. $V_{TN} = 1V$, $K_n = 10mA/V^2$.



Quando $v_i = 5V$ si verifica facilmente che il MOSFET è in regione lineare con $v_{DS} \ll 2(v_{GS} - V_{TN})$ e quindi la tensione v_o è circa zero.

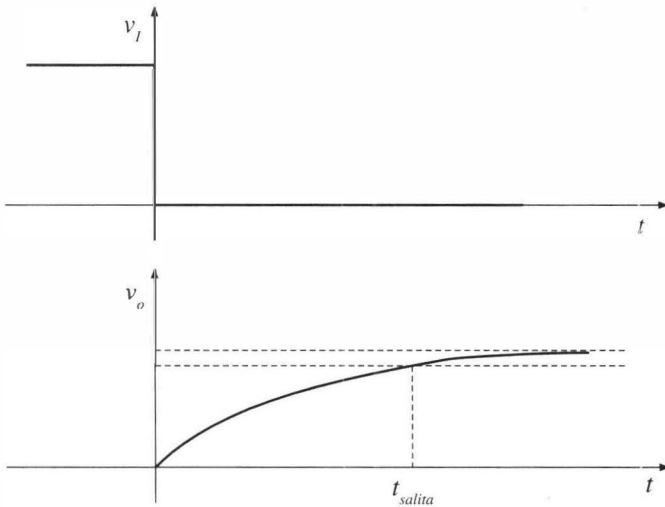
Quando l'ingresso repentinamente passa da 5V a 0V, il MOSFET entra in interdizione e il capacitore si carica attraverso la resistenza R_D secondo la formula

$$v_o = V_{DD} \left(1 - e^{-\frac{t}{R_D C}} \right)$$

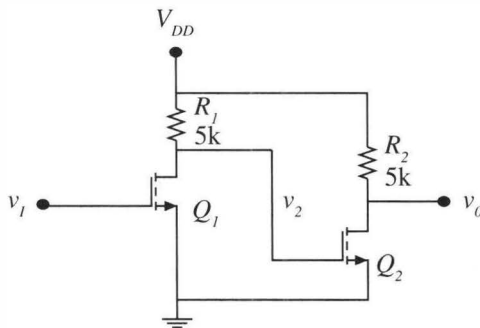
Il tempo di salita si calcola quindi come

$$\frac{9}{10} = 1 - e^{-\frac{t}{R_D C}}$$

$$t = -R_D C \ln(0.1) = 2.3 \text{ ms}$$



13) Si tracci la funzione di trasferimento del circuito in figura. $K_n = 1 \text{ mA/V}^2$; $V_{TN} = 2 \text{ V}$; $V_{DD} = 5 \text{ V}$.



Consideriamo la tensione v_2 .

Per $v_i < V_{TN} = 2 \text{ V}$ Q_1 è interdetto e $v_2 = 5 \text{ V}$.

Per $v_i > 2 \text{ V}$ è ON ed è in saturazione quando

$$v_{DS1} > v_{GS1} - V_{TN}$$

ovvero

$$v_2 > v_I - (2V)$$

In saturazione possiamo scrivere

$$v_2 = V_{DD} - R_1 i_{DS1} = V_{DD} - R_1 \frac{K_n}{2} (v_I - V_{TN})^2$$

Q_2 passa dalla conduzione all'interdizione per $v_2 = 2V$, quindi per $v_I = 3.1V$. Per tale valore Q_1 è ancora in saturazione in quanto

$$v_2 = v_{DS1} > (v_{GS1} - V_{TN}) = 1.1V$$

Quando Q_2 è in interdizione l'uscita è pari a 5V. Calcoliamo ora la tensione di uscita quando Q_2 è in conduzione. Facciamo l'ipotesi che sia nella regione lineare in cui la caratteristica del MOSFET è approssimabile con una retta:

$$R_{ON} = \frac{1}{K_n (V_{DD} - V_{TN})} = 330\Omega$$

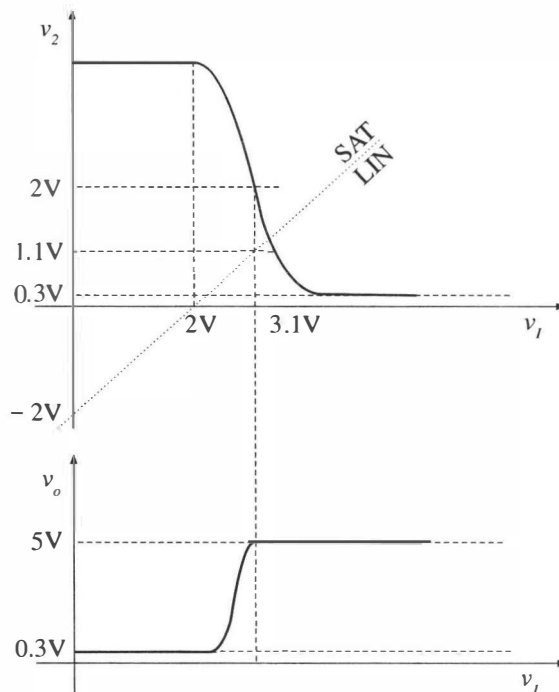
quindi

$$v_O = \frac{R_{ON}}{R_{ON} + R_2} V_{DD} = 0.3V$$

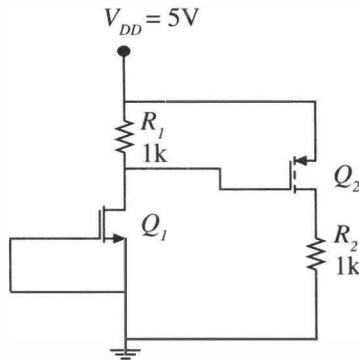
La condizione affinché l'ipotesi di linearità sia corretta è

$$v_O \ll 2(V_{GS} - V_{TN}) = 6V$$

che è ben rispettata.



14) Il MOSFET a svuotamento ha $V_{TN} = 2\text{V}$, $K_n = 1\text{mA/V}^2$. Il MOSFET ad arricchimento ha $V_{TP} = 1\text{V}$ e $K_p = 1\text{mA/V}^2$. Si calcoli la corrente i_{SD} del MOSFET ad arricchimento.



Supponiamo Q_1 in saturazione.

$$I_{DS} = \frac{K_n}{2} V_{TN}^2 = 2\text{mA}$$

La caduta di tensione su R_1 è 2V e

$$V_{DS1} = 3\text{V} > V_{TN} = 2\text{V}$$

pertanto Q_1 è effettivamente in saturazione.

$$V_{SG2} = V_{DD} - V_{DS1} = 2\text{V} > V_{TP} = 1\text{V}$$

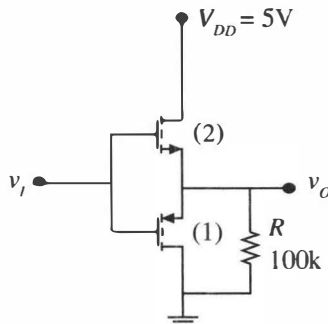
Pertanto Q_2 è in conduzione. Supponiamo che sia in saturazione.

$$I_{SD2} = \frac{K_p}{2} (V_{SG2} - V_{TP})^2 = 0.5\text{mA}$$

Verifichiamo che sia effettivamente in saturazione.

$$V_{SD2} = V_{DD} - R_2 I_{SD2} = 4.5\text{V} > V_{SG2} - V_{TP} = 1\text{V}$$

15) Si calcoli il valore della tensione in uscita quando $v_i = 0\text{V}$ e $v_i = 5\text{V}$. $V_{TN} = V_{TP} = 1\text{V}$, $K_n = K_p = 1\text{mA/V}^2$.



Quando $v_i = 0V$ il MOSFET 2 (in alto) è interdetto e la resistenza sull'uscita porta a massa l'uscita stessa, $v_{SG1} = 0$ e quindi anche il MOSFET 1 (in basso) è interdetto.

Quando $v_i = 5V$, $v_{GS2} = v_{DS2}$. Il MOSFET 2 è in saturazione, infatti la condizione per la saturazione è

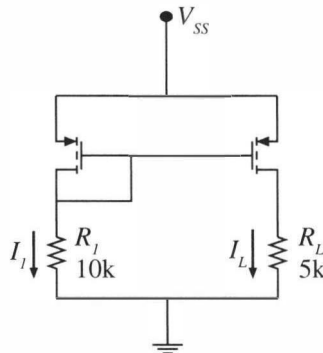
$$v_{DS2} > v_{GS2} - V_{TN}$$

che è sempre verificata. Scriviamo quindi

$$R \frac{K_n}{2} (v_{GS2} - V_{TN})^2 = V_{DD} - v_{GS2}$$

che è un'equazione di secondo grado nell'incognita v_{GS2} . Le soluzioni sono due: $v_{GS2} = 0.707V$ e $v_{GS2} = 1.27V$. La prima non è accettabile in quanto l'espressione per la corrente in saturazione vale solo per $v_{GS2} > V_{TN} = 1V$. Pertanto quando $v_i = 5V$, $v_o = 3.73V$.

16) Si calcoli la corrente I_L , $V_{SS} = 5V$, $K_p = 20mA/V^2$, $V_{TP} = 2V$.



Supponiamo che il punto di polarizzazione del MOSFET sia piuttosto vicino alla soglia

$$v_{SG} \approx V_{TP}$$

Possiamo quindi calcolare I_I come

$$I_I = \frac{V_{SS} - V_{TP}}{R_I} = 0.3mA$$

Poiché i due MOSFET sono uguali e hanno la stessa v_{SG} possiamo concludere che $I_L = I_I = 0.3mA$.

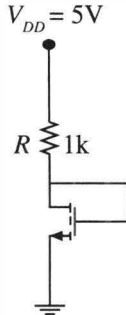
Vogliamo ora verificare l'ipotesi fatta inizialmente.

Se $V_{SG} \approx V_{TP}$, $(V_{SG} - V_{TP}) \approx 0 < V_{SD}$ quindi se l'ipotesi è vera il MOSFET è in saturazione e

$$I_{SD} = \frac{K_p}{2} (V_{SG} - V_{TP})^2 = 0.3mA$$

da cui $(V_{SG} - V_{TP}) = 0.18V$ che è molto minore di V_{TP} , quindi effettivamente $V_{SG} \approx V_{TP}$.

17) Si calcoli il valore della tensione V_{DS} , $V_{TN} = 1\text{V}$, $K_n = 2\text{mA/V}^2$.

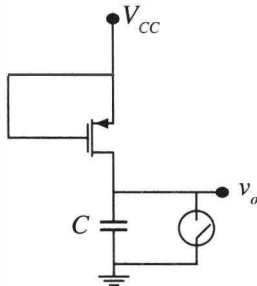


Il MOSFET è in saturazione e vale

$$\frac{K_n}{2} (V_{GS} - V_{TN})^2 = \frac{V_{DD} - v_{GS}}{R}$$

con $V_{DD} = 5\text{V}$ e $R = 1\text{k}\Omega$. L'equazione di secondo grado con incognita v_{GS} ha due soluzioni di cui accettabile solo $v_{GS} = 2.56\text{V}$.

18) All'istante $t = 0$ l'interruttore passa da ON a OFF. Calcolare l'istante al quale $v_o = 4\text{V}$, $V_{CC} = 5\text{V}$, $C = 100\text{nF}$, $V_{TP} = 1\text{V}$, $K_p = 10\text{mA/V}^2$.

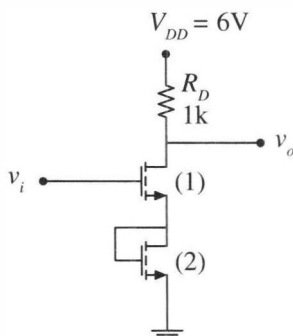


Il MOSFET è a canale p a svuotamento. Se $v_{SD} > V_{TP} = 1\text{V}$ è in saturazione e quindi fornisce una corrente costante ($i_{DS} = 5\text{mA}$) che carica linearmente il capacitore fino alla tensione $v_o = V_{CC} - V_{TP} = 4\text{V}$. Oltre questo valore il MOSFET non è più in saturazione. Il tempo che impiega il capacitore per raggiungere 4V è

$$\Delta t = \frac{CV_0}{i_{DS}} = 80\mu\text{s}$$

con $V_0 = 4\text{V}$ e $C = 100\text{nF}$.

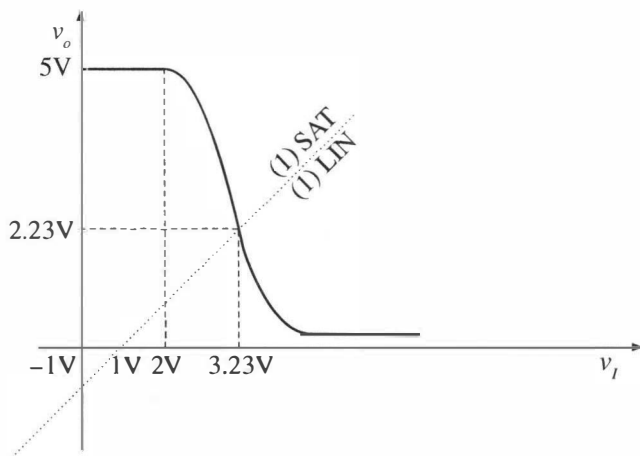
19) Tracciare la funzione di trasferimento per v_i compresa tra 0 e 6V. I due MOSFET sono identici. $K_n = 20\text{mA/V}^2$, $V_{TN} = 1\text{V}$. Calcolare il valore di v_i per cui il MOSFET (1) passa dalla saturazione alla regione lineare



Poiché i due MOSFET sono identici e conducono la stessa corrente $v_{GS1} = v_{GS2}$, quindi se $v_i < 2V_{TN} = 2\text{V}$ i due MOSFET sono interdetti e la tensione di uscita è pari a $V_{DD} = 6\text{V}$. Quando v_i supera 2V, i due MOSFET entrano simultaneamente in conduzione. Poiché il MOSFET (2) ha D e S cortocircuitati è sempre in saturazione. Il MOSFET (1) invece è in saturazione quando

$$\begin{aligned} v_{DS}^{(1)} &> v_{GS}^{(1)} - V_{TN} \\ v_o - \frac{v_i}{2} &> \frac{v_i}{2} - V_{TN} \\ v_o &> v_i - V_{TN} \end{aligned}$$

da cui il grafico in figura.

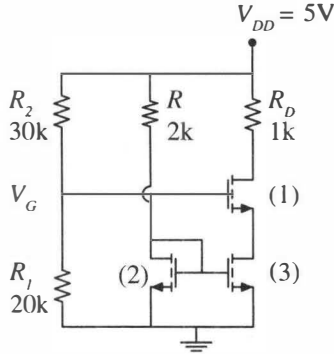


Per calcolare la tensione v_i per cui il MOSFET passa dalla saturazione alla regione lineare è necessario risolvere l'equazione

$$V_{DD} - R_D \frac{K_n}{2} \left(\frac{v_i}{2} - V_{TN} \right)^2 = v_i - V_{TN}$$

che ha soluzione $v_i = 3.23\text{V}$, da cui $v_o = 2.23$.

20) Calcolare la i_{DS} del MOSFET (1). I 3 MOSFET sono identici. $K_n = 20\text{mA/V}^2$, $V_{TN} = 1\text{V}$.



Nei MOSFET (1) e (3) scorre la stessa corrente perché sono in serie. Il MOSFET (2) è sicuramente in saturazione perché ha D e G cortocircuitati. Se anche il MOSFET (3) è in saturazione, nei tre transistor scorre la stessa corrente.

Per calcolare la corrente che scorre nel MOSFET (1) è necessario risolvere la seguente equazione di secondo grado con incognita v_{GS}

$$\frac{K_n}{2} (v_{GS} - V_{TN})^2 = \frac{V_{DD} - v_{GS}}{R}$$

L'equazione ha due soluzioni di cui accettabile solo $v_{GS} = 1.42\text{V}$, per la quale $i_{DS} = 1.79\text{mA}$.

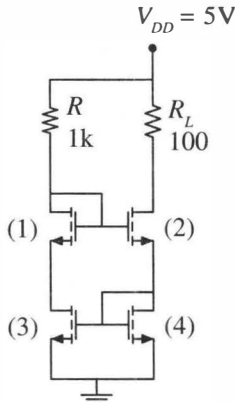
Resta da verificare che il MOSFET (3) sia in saturazione.

$$v_{DS}^{(3)} > v_{GS} - V_{TN}$$

$$\frac{R_1}{R_1 + R_2} V_{DD} - v_{GS} > v_{GS} - V_{TN}$$

$$0.58\text{V} > 0.42\text{V}$$

21) Calcolare la i_{DS} del MOSFET (2). I 4 MOSFET sono identici. $K_n = 2\text{mA/V}^2$, $V_{TN} = 1\text{V}$.



Nei MOSFET (1) e (3) scorre la stessa corrente perché sono in serie. Nei MOSFET (2) e (4) scorre la stessa corrente perché sono in serie. I MOSFET (1) e (4) sono sicuramente in saturazione perché hanno D e G cortocircuitati. Se anche il MOSFET (3) è in saturazione, nei quattro transistor scorre la stessa corrente.

Poiché i 4 MOSFET sono uguali e vi scorre la stessa corrente segue che hanno la stessa v_{GS} . Per calcolare corrente che scorre nel MOSFET (1) è necessario risolvere la seguente equazione di secondo grado con incognita v_{GS} :

$$\frac{K_n}{2} (v_{GS} - V_{TN})^2 = \frac{V_{DD} - 2v_{GS}}{R}$$

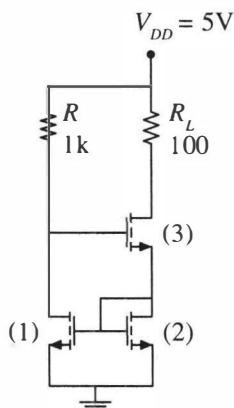
L'equazione ha due soluzioni di cui accettabile solo $v_{GS} = 2V$, per la quale $i_{DS} = 1\text{ mA}$.

Resta da verificare che il MOSFET (3) sia in saturazione.

$$\begin{aligned} v_{DS}^{(3)} &> v_{GS} - V_{TN} \\ i_{DS}R - v_{GS} &> v_{GS} - V_{TN} \\ 2V &> 0.42V \end{aligned}$$

Nota: La corrente che scorre nel carico R_L è fissata dalla resistenza R . Questo circuito è un generatore di corrente molto usato nei circuiti integrati.

22) Calcolare la i_{DS} del MOSFET (3). I 3 MOSFET sono identici. $K_n = 2\text{ mA/V}^2$, $V_{TN} = 1V$.



Se il MOSFET (1) è in saturazione nei 3 MOSFET scorre la stessa corrente e hanno la stessa v_{GS} . Per calcolare corrente che scorre nel MOSFET (1) è necessario risolvere la seguente equazione di secondo grado con incognita v_{GS} :

$$\frac{K_n}{2} (v_{GS} - V_{TN})^2 = \frac{V_{DD} - 2v_{GS}}{R}$$

L'equazione ha due soluzioni, di cui accettabile $v_{GS} = 2V$. Quindi

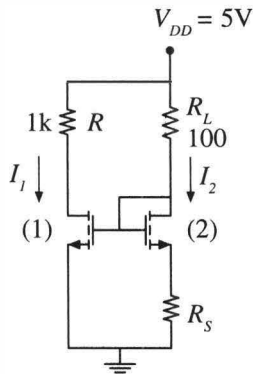
$$i_{DS} = \frac{V_{DD} - 2v_{GS}}{R} = 1mA$$

Verifichiamo ora che il MOSFET (1) sia in saturazione.

$$\begin{aligned} v_{DS}^{(1)} &= V_{DD} - Ri_{DS} = 4V \\ v_{DS}^{(1)} &> v_{GS} - V_{TN} \\ 4V &> 1V \end{aligned}$$

*Nota: La corrente che scorre nel carico R_L è fissata dalla resistenza R . Questo circuito, noto come **specchio di Wilson**, è un generatore di corrente molto usato nei circuiti integrati.*

***23) Dimensionare R_S in modo $I_1 = 1mA$ e $I_2 = 0.1mA$. I 2 MOSFET sono identici. $K_n = 2mA/V^2$, $V_{TN} = 1V$.**



Dall'analisi della maglia del circuito risulta

$$\begin{aligned} v_{GS}^{(1)} &= V_{DD} - RI_1 \\ v_{GS}^{(2)} &= v_{GS}^{(1)} - R_S I_2 \end{aligned}$$

Sostituendo la prima equazione nella seconda

$$v_{GS}^{(2)} = V_{DD} - RI_1 - R_S I_2$$

D'altra parte

$$I_2 = \frac{K_n}{2} (v_{GS}^{(2)} - V_{TN})^2 = \frac{K_n}{2} (V_{DD} - RI_1 - R_S I_2 - V_{TN})^2$$

da cui

$$\sqrt{\frac{2I_2}{K_n}} = V_{DD} - I_1 R - I_2 R_S - V_{TN}$$

$$R_S = \frac{V_{DD}}{I_2} - \frac{I_1 R}{I_2} - \frac{V_{TN}}{I_2} - \sqrt{\frac{2}{I_2 K_n}}$$

Sostituendo i valori numerici $R_S = 26.8 \text{ k}\Omega$.

*Nota: La corrente che scorre nel carico R_L è fissata dalle resistenze R e R_S . Questo circuito è un generatore di corrente (noto come **specchio di Widlar**) molto usato nei circuiti integrati.*

4.32 Domande di verifica a scelta multipla

1. Quando è applicata una tensione negativa a un capacitore MOS metallo-semiconduttore di tipo p
 - a) si crea una regione di svuotamento nel semiconduttore in prossimità dell'ossido
 - b) si crea una zona di cariche positive mobili nel semiconduttore in prossimità dell'ossido
 - c) si crea una zona di cariche negative mobili nel semiconduttore in prossimità dell'ossido
 - d) scorre una corrente proporzionale alla tensione applicata
2. Quando è applicata una tensione positiva a un capacitore MOS metallo-semiconduttore di tipo p
 - a) si crea una regione di svuotamento nel semiconduttore in prossimità dell'ossido
 - b) si crea un accumulo di cariche positive nel semiconduttore in prossimità dell'ossido
 - c) scorre una corrente proporzionale alla tensione applicata
 - d) si crea un canale di ioni negativi mobili
3. La tensione di soglia in un capacitore MOS è la tensione per cui
 - a) l'energia di conduzione è maggiore della tensione termica
 - b) ci sono cariche mobili nel semiconduttore in prossimità dell'ossido
 - c) la regione di svuotamento occupa l'intero semiconduttore
 - d) l'energia di gap coincide con l'energia di Fermi
4. Un MOSFET a canale n è in regione lineare
 - a) quando scorre una corrente tra drain e source
 - b) quando la tensione di drain è sufficientemente elevata
 - c) quando la corrente di drain è proporzionale alla tensione applicata al drain
 - d) quando la tensione di gate è sufficientemente elevata
5. Un MOSFET a canale n è in saturazione
 - a) quando la tensione di drain supera la tensione di alimentazione
 - b) quando il canale occupa l'intero semiconduttore
 - c) quando drain e source sono cortocircuitati
 - d) quando la tensione di gate è appena superiore alla soglia

6. A parità di altre condizioni un MOSFET conduce una corrente maggiore
 - a) se lo spessore dell'ossido è maggiore
 - b) se la lunghezza del canale è maggiore
 - c) se la larghezza del canale è minore
 - d) se il substrato è di tipo p invece che di tipo n
7. In un MOSFET in strozzamento
 - a) non scorre corrente nel canale
 - b) la corrente aumenta fino a danneggiare il transistor
 - c) la lunghezza del canale diminuisce
 - d) il canale non può più sussistere
8. In un p-MOSFET
 - a) il substrato è di tipo p
 - b) un flusso di elettroni scorre tra drain e source
 - c) la corrente fluisce dal source al drain
 - d) la corrente è controllata dalla tensione tra gate e drain
9. Un n-MOSFET a svuotamento
 - a) è sempre in interdizione quando la tensione tra gate e source è zero
 - b) è in conduzione quando la tensione tra gate e source è zero
 - c) la corrente scorre con verso positivo tra gate e drain
 - d) la massima corrente scorre quando la tensione tra gate e source è zero
10. In un JFET
 - a) la corrente di source è la somma della corrente di drain e la corrente di gate
 - b) la tensione tra gate e source è circa 0.6V
 - c) la corrente di source è uguale alla corrente di drain
 - d) non può scorrere corrente tra source e drain, ma solo tra drain e source
11. In un MOSFET a canale n ad arricchimento a tre terminali, quando gate e source sono cortocircuitati
 - a) può scorrere corrente tra drain e source
 - b) c'è una giunzione polarizzata direttamente tra source e body
 - c) potrebbe scorrere corrente tra source e drain se il source fosse posto a una tensione maggiore di almeno 0.6 V rispetto al drain
 - d) scorre una corrente di contropolarizzazione tra il drain e il gate
12. In un n-MOSFET ad arricchimento
 - a) la corrente di source è la somma della corrente di drain e della corrente di gate
 - b) la tensione tra gate e source è circa 0.6V
 - c) la corrente di source è uguale alla corrente di drain
 - d) può scorrere corrente se gate e source sono cortocircuitati
13. In un n-MOSFET ad arricchimento in saturazione
 - a) la velocità di deriva degli elettroni è costante lungo il canale
 - b) la tensione è costante lungo il canale
 - c) la densità di carica è costante lungo il canale
 - d) la corrente è costante lungo il canale

14. In un n-MOSFET ad arricchimento in saturazione, la corrente
 - a) aumenta all'aumentare della v_{DS} descrivendo una parabola rovesciata
 - b) diminuisce leggermente all'aumentare di v_{DS}
 - c) aumenta all'aumentare della v_{GS}
 - d) è nulla
15. Un JFET a canale n con $V_p = 1\text{ V}$ e $I_{DSS} = 10\text{ mA}$
 - a) per $v_{GS} = -0.5\text{ V}$ conduce
 - b) per $v_{GS} = 0$, $i_{DS} = 0$
 - c) per $v_{GS} > 0$ è interdetto
 - d) per $v_{GS} = -1.5\text{ V}$ conduce
16. In un MOSFET a canale n in conduzione, la concentrazione degli elettroni mobili
 - a) è maggiore in prossimità del source
 - b) è sempre costante lungo il canale, perché la corrente è costante
 - c) è sempre nulla in prossimità del drain
 - d) è nulla in prossimità del source
17. In un MOSFET a canale n ad arricchimento, quando tra D e S è applicata una tensione positiva
 - a) il canale tra D e S è strozzato
 - b) la tensione nel semiconduttore sotto l'ossido non è costante da D a S
 - c) D e G sono alla stessa tensione
 - d) scorre una corrente uscente dal G
18. In un MOSFET a canale n ad arricchimento in saturazione
 - a) la densità di cariche mobili nel canale è sempre costante, in quanto la corrente che scorre è la stessa lungo il canale
 - b) il rapporto tra densità di carica e velocità degli elettroni è costante lungo il canale
 - c) non ci sono cariche mobili nel canale
 - d) la velocità delle cariche mobili lungo il canale aumenta da S a D
19. L'autoallineamento delle piazzole S e D con il gate si realizza
 - a) mediante due maschere perfettamente collimate
 - b) tagliando il polisilicio del gate con un laser
 - c) mediante un'impiantazione ionica dopo l'asportazione selettiva del polisilicio
 - d) collimando con un fascio di elettroni il canale del MOSFET con il terminale di body
20. L'ossido di silicio
 - a) è permeabile alle impurità n
 - b) è un effetto dell'usura e finisce con il danneggiare il circuito integrato
 - c) è un acido impiegato per l'attacco chimico del silicio nei processi fotolitografici
 - d) è uno strato protettivo del silicio

4.33 Domande di verifica a risposta aperta

1. Si elenchi e si tracci il simbolo dei principali transistor a effetto campo.
2. Si ricavi l'espressione per la tensione di soglia di un MOSFET a canale n ad arricchimento.
3. Si ricavi l'andamento della caratteristica di uscita per un MOSFET a canale n ad arricchimento per piccole tensioni tra drain e source.
4. Si ricavi l'andamento della caratteristica di uscita per un MOSFET a canale n ad arricchimento in regione lineare e in saturazione.
5. Si illustri la causa fisica per cui la caratteristica di uscita di un MOSFET in saturazione non è piatta.
6. Si illustri qualitativamente il principio di funzionamento del JFET.

4.34 Risposte alle domande a scelta multipla

1 b	8 c	15 a
2 a	9 b	16 a
3 b	10 c	17 b
4 d	11 c	18 d
5 d	12 c	19 c
6 d	13 d	20 d
7 c	14 c	

5

Il transistor bipolare

In questo capitolo si tratta il transistor bipolare (BJT). Si definiscono le quattro regioni di funzionamento (attiva diretta, attiva inversa, interdizione e saturazione) per passare a illustrare il principio di funzionamento in regione attiva diretta, con riferimento alla corrente di trasporto e ai processi di ricombinazione in base. Si passa quindi alle applicazioni circuitali: interruttori, reti di polarizzazione e specchi di corrente. Un paragrafo è dedicato al confronto tra BJT e FET. Il capitolo si conclude con un cenno sulla nascita dell'industria dei semiconduttori.

5.1 Il BJT

Il transistor bipolare (BJT: *Bipolar Junction Transistor*) è stato il primo dispositivo a stato solido, commercializzato e largamente diffuso, in grado di sostituire le valvole termoioniche, ovvero in grado di funzionare sia come amplificatore sia come interruttore elettronico. Nei primi anni sessanta fu il protagonista della rivoluzione elettronica che ha interessato tutti gli aspetti della società e solo a partire dagli anni novanta ha cominciato a essere sostituito dai transistor a effetto campo. Oggi, pur ridimensionandosi nell'uso, rimane il transistor di elezione nelle applicazioni in cui è necessario grande guadagno e capacità di pilotare grandi correnti.

5.2 Realizzazione fisica del BJT

Innanzitutto è necessario specificare che il BJT è un dispositivo a tre terminali chiamati collettore, base ed emettitore. Ci sono due tipi di BJT: npn e pnp. Nell'npn la corrente, nelle condizioni operative più comuni, scorre dal collettore verso l'emettitore ed è controllata dalla tensione tra base ed emettitore. Nel pnp la corrente scorre dall'emettitore al collettore ed è controllata dalla tensione tra emettitore e base.

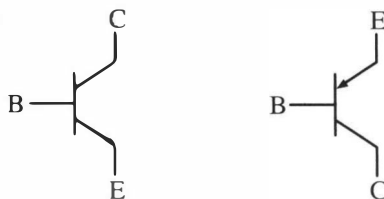


Figura 5.1 npn e pnp

La tecnologia del BJT è antecedente a quella del MOSFET e pertanto si basa su processi meno sofisticati. In particolare il drogaggio del semiconduttore avviene per diffusione anziché per impiantazione ionica. Per impressionare il fotoresist si usa luce ultravioletta (UV) anziché un fascio di elettroni. Questi processi limitano le dimensioni minime ottenibili del transistor, ma d'altra parte oggi difficilmente si usa il BJT in circuiti ad alta integrazione. In ogni caso i dettagli possono variare a seconda del produttore. Nel seguito ci riferiremo al processo produttivo più diffuso e consolidato.

Per fissare le idee prendiamo in considerazione il BJT npn. La fabbricazione di un BJT inizia dal wafer, ovvero da una fetta di silicio monocristallino drogato p di spessore 0.2mm. I passi successivi sono i seguenti:

1) **Ossidazione.** Il wafer è ricoperto di uno strato di ossido di silicio (SiO_2), che è un composto piuttosto resistente e impermeabile ai droganti normalmente usati.

Sull'ossido di silicio si deposita uno strato di **fotoresist**. Si appoggia una maschera di vetro con zone trasparenti e zone opache. Si illumina il wafer mascherato con luce UV. Le zone in ombra rimangono inalterate, le zone illuminate da luce si induriscono. Il fotoresist viene lavato via con un solvente e rimane solo dove è stato indurito dalla luce. Il fotoresist rimasto viene ulteriormente indurito e stabilizzato con dei fissanti chimici. Il wafer è immerso in un bagno chimico aggressivo in grado di rimuovere l'ossido di silicio nelle zone non protette dal fotoresist. Infine, mediante opportuni reagenti, il fotoresist è asportato e il wafer rimane con zone scoperte e zone protette dall'ossido di silicio. Le zone scoperte corrispondono a singoli BJT che si vorranno realizzare.

2) **Diffusione di impurità n.** La zona non mascherata è drogata fortemente con impurità n mediante diffusione di un composto gassoso contenente atomi di impurità che, ad alta temperatura, penetrano nel silicio. È un'operazione che richiede diverse ore. In questo modo si realizza lo **strato sepolto**.

3) **Rimozione dell'ossido.** L'ossido è rimosso mediante opportuni solventi.

4) **Crescita epitassiale.** La fase successiva prevede la crescita di uno strato di silicio sopra il wafer. Ciò è ottenuto mediante un composto del silicio in forma gassosa (o liquida) che in un forno ad alta temperatura si deposita sul wafer seguendone i piani cristallini. Durante la crescita si aggiungono impurità n in modo da realizzare del silicio debolmente drogato. Durante questa fase, a causa dell'elevata temperatura le impurità dello strato sepolto tendono a diffondersi nello strato epitassiale.

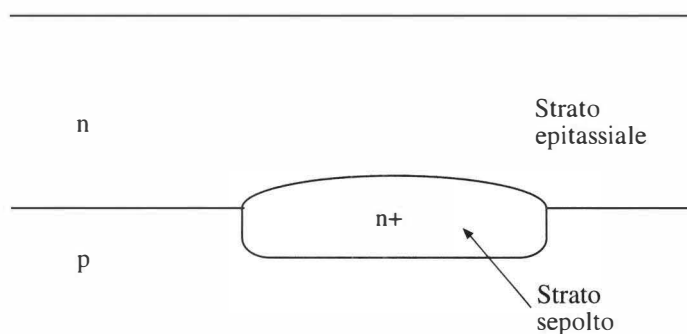


Figura 5.2 Strato sepolto e strato epitassiale

5) **Seconda ossidazione.** Le zone scoperte corrispondono alle regioni di separazione tra un BJT e l'altro.

6) **Diffusione di impurità p.** La zona non mascherata è drogata fortemente con impurità p mediante diffusione di un composto gassoso contenente atomi di impurità che, ad alta temperatura, penetrano nel silicio. Il tempo di diffusio-

ne deve essere sufficientemente lungo da far sì che le impurità arrivino fino al substrato p. Si noti che il processo di diffusione non consente di realizzare superfici di separazione nette e dritte, ma la zona drogata tende a essere più stretta nelle regioni più profonde. Questo è un limite importante della tecnologia BJT. I singoli transistor devono essere isolati l'uno dall'altro (al contrario dei MOSFET che sono autoisolati) da zone di guardia profonde fino al substrato che, quindi, devono essere piuttosto larghe in superficie.

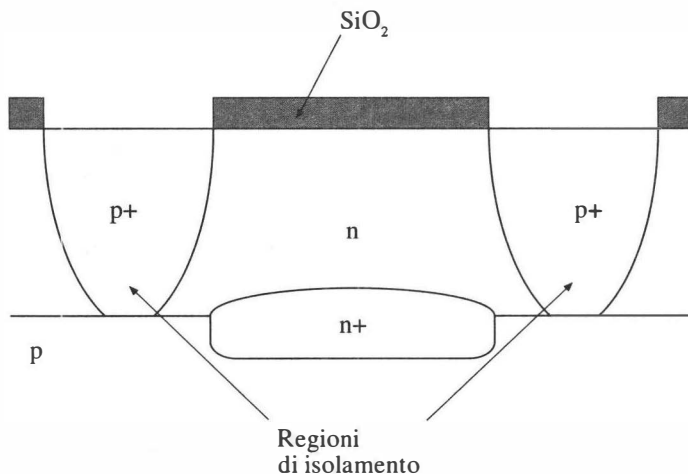


Figura 5.3 Regioni di isolamento

7) Rimozione dell'ossido di silicio.

8) **Terza ossidazione.** Le zone scoperte corrispondono alle regioni di base dei singoli BJT che si vogliono realizzare.

9) **Diffusione di impurità p.** La zona non mascherata è drogata con impurità p mediante diffusione.

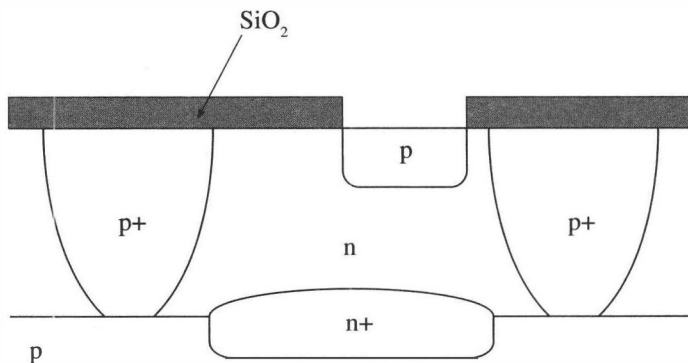


Figura 5.4 Realizzazione della base

10) Rimozione dell'ossido di silicio.

11) **Quarta ossidazione.** Le zone scoperte corrispondono alle zone di emettitore e alle piazzole di contatto dei collettori.

12) **Diffusione di impurità n.** La zona non mascherata è drogata fortemente con impurità n mediante diffusione. Il processo di diffusione va attentamente calibrato in modo che la zona di emettitore n+ dentro la regione di base arrivi molto vicino al bordo inferiore della base, senza oltrepassarlo.

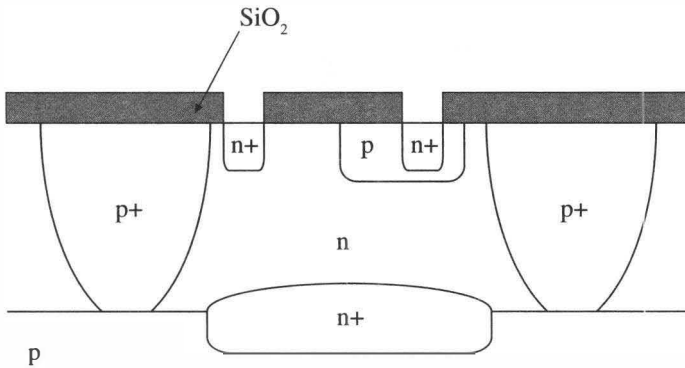


Figura 5.5 Realizzazione dell'emettitore

13) Rimozione dell'ossido di silicio.

14) **Quinta ossidazione.** Le zone scoperte corrispondono ai contatti elettrici che si vogliono realizzare.

15) **Metallizzazione.** In una camera a vuoto vapori di alluminio vengono fatti depositare sul wafer.

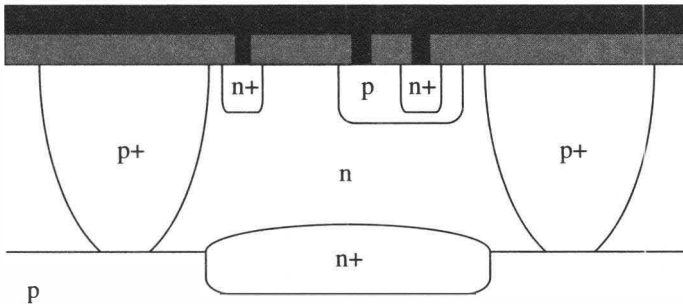


Figura 5.6 Metallizzazione

16) **Sesta ossidazione.** Le zone scoperte corrispondono alle zone da cui si vuole asportare il metallo.

17) **Attacco chimico del metallo.**

18) **Rimozione della mascheratura.** Questa è l'ultima operazione per realizzare il singolo transistor. Per un circuito più complesso è necessario creare una rete di interconnessioni metalliche. Queste si realizzano mediante strati successivi di ossido e metallo con punti passanti (*vias*) tra due strati.

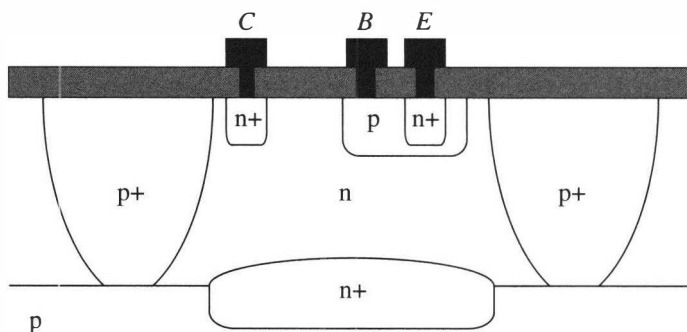


Figura 5.7 Rimozione della mascheratura

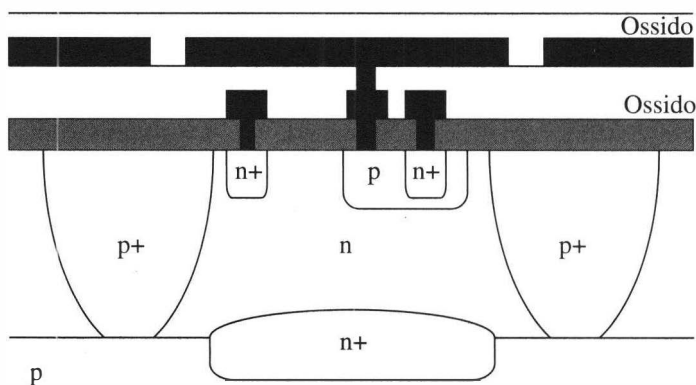


Figura 5.8 BJT completo con due strati di interconnessioni elettriche

Come vedremo in dettaglio in seguito, in un BJT in condizioni operative tipiche la corrente scorre dal collettore verso l'emettitore attraversando in senso verticale il sottile strato di base tra collettore ed emettitore. Tale corrente percorre un lungo tratto nella zona n, poco drogata, del collettore. La caduta di tensione in questo percorso può non essere trascurabile, soprattutto in considerazione del

fatto che il BJT è in grado di pilotare grandi correnti. Ed è appunto per questa ragione che sotto il collettore si realizza uno strato sepolto molto drogato in modo da offrire un percorso parallelo a bassa impedenza. Come noto, infatti, quando due resistenze sono in parallelo la resistenza complessiva è sicuramente minore della minore tra le due.

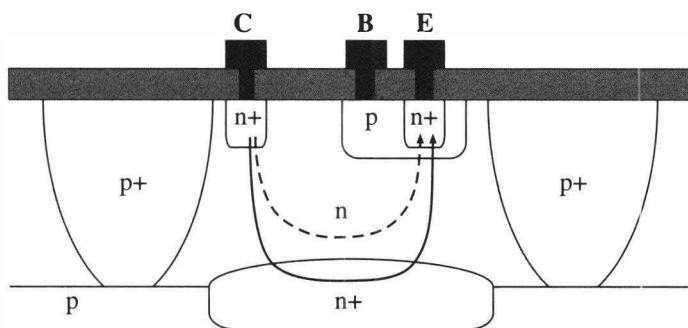


Figura 5.9 Corrente da collettore a emettitore

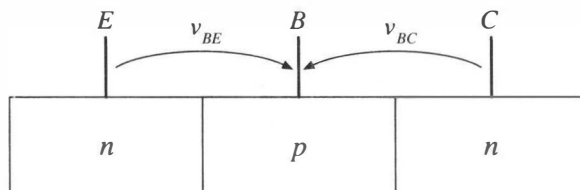
Abbiamo descritto sopra le fasi del processo di realizzazione di un BJT in tecnologia planare. Il grande vantaggio di tale tecnologia è il fatto di poter realizzare contemporaneamente migliaia o milioni di transistor variamente connessi. A questo proposito si noti che il processo produttivo non permette di variare drogaggi e spessori, ma solo le dimensioni planari. Come vedremo nel seguito uno dei parametri fondamentali di un BJT è il guadagno di corrente (β_F) che dipende dallo spessore della base e dal suo drogaggio. Ciò significa che il progettista di circuiti integrati deve trovare soluzioni circuitali che non richiedano BJT con β_F diversi.

Così come per la tecnologia MOSFET, nella realizzazione di circuiti integrati a BJT, i resistori tendono a occupare molto spazio e le induttanze non possono proprio essere realizzate. Riguardo ai capacitori, in tecnologia bipolare vengono realizzati, di solito, contropolarizzando una delle due giunzioni pn di cui è costituito un BJT integrato. Come vedremo nel seguito, una giunzione pn contropolarizzata è di fatto una capacità (il calcolo esatto di tale capacità sarà eseguito nel Capitolo 8, quando verranno discusse le capacità parassite del BJT).

5.3 Regioni di funzionamento

Un BJT npn è essenzialmente costituito da due giunzioni pn con la regione p in comune (la regione di base). Ciascuna giunzione pn può essere nello stato ON o OFF, ovvero condurre o non condurre, a seconda che la tensione applicata sia maggiore o minore della tensione di soglia (0.6V). I possibili stati, ovvero le regioni di funzionamento del BJT, sono quindi quattro:

- 1) regione attiva diretta: giunzione BE ON e giunzione BC OFF
- 2) regione attiva inversa: giunzione BE OFF e giunzione BC ON
- 3) saturazione: giunzione BE ON e giunzione BC ON
- 4) interdizione: giunzione BE OFF e giunzione BC OFF



	$v_{BC} < 0.6V$ - OFF	$v_{BC} > 0.6V$ - ON
$v_{BE} < 0.6V$ - OFF	INTERDIZIONE	REGIONE ATTIVA INVERSA
$v_{BE} > 0.6V$ - ON	REGIONE ATTIVA DIRETTA	SATURAZIONE

Figura 5.10 Regioni di funzionamento del BJT

Quando il BJT lavora come un interruttore passa dallo stato di saturazione (interruttore ON) allo stato di interdizione (interruttore OFF). La regione attiva diretta è lo stato in cui si trova il BJT quando opera come amplificatore per piccoli segnali. La regione attiva inversa non è usata, se non in un'unica importante applicazione che è lo stadio d'ingresso della famiglia logica TTL. Tale famiglia, pur avendo dominato il mercato per tutti gli anni ottanta, è oggi obsoleta.

5.4 Regione attiva diretta

In regione attiva diretta la giunzione tra base ed emettitore è polarizzata direttamente ($v_{BE} \geq 0.6V$) e la giunzione tra base e collettore è polarizzata inversamente ($v_{BC} < 0.6V$). Indichiamo positiva la corrente entrante nel collettore, la corrente entrante in base e la corrente uscente dall'emettitore, con questa convenzione nel normale funzionamento del BJT le correnti risulteranno tutte positive. Poiché la giunzione BC è polarizzata inversamente e la giunzione BE direttamente, la corrente entrante in base non potrà che scorrere verso l'emettitore. Inoltre ci potrà essere una corrente diretta, detta corrente di trasporto i_T , che fluisce tra collettore ed emettitore attraverso la base (Figura 5.11). La corrente di collettore

Gli elettroni iniettati dall'emettitore in base (portatori minoritari, in quanto in base sono in minore concentrazione rispetto alle lacune) trovano tra base e collettore una giunzione inversamente polarizzata, quindi un potenziale elettrico che impedisce alle lacune (portatori maggioritari) di diffondersi dalla base al collettore, ma tale da spingere gli elettroni dalla base verso il collettore. Pertanto la concentrazione degli elettroni $n(w)$ al confine tra la base e la regione di svuotamento tra base e collettore è circa nulla.

$$n(w) = 0 \quad (5.3)$$

A rigore il confine tra la base e la regione di svuotamento tra base e collettore non coincide esattamente con la separazione tra zona p di base e zona n di collettore. Tuttavia, poiché la base è molto più drogata del collettore, la parte di regione di svuotamento che si estende in base risulta abbastanza piccola da poter essere trascurata.

La concentrazione degli elettroni in base avrà un qualche andamento dipendente dalla lunghezza di diffusione degli elettroni in base, ma se la base è molto sottile, ovvero molto più corta della lunghezza di diffusione degli elettroni, questo andamento potrà essere approssimato come lineare.

In generale un gradiente di concentrazione produce una corrente di diffusione; la corrente di trasporto i_T è, appunto, la corrente data dal gradiente dei portatori minoritari in base, quindi

$$i_T = -qSD_n \frac{dn}{dx} \quad (5.4)$$

dove S è la sezione del BJT, D_n il coefficiente di diffusione degli elettroni. Il segno negativo è dovuto al fatto che il senso della corrente in Figura 5.11 è definito positivo nel verso opposto all'asse x . Poiché l'andamento della concentrazione degli elettroni in base è circa lineare possiamo scrivere

$$i_T \approx -qSD_n \frac{n(w) - n(0)}{w} \quad (5.5)$$

Sostituendo la (5.2) e la (5.3) nella (5.5), otteniamo l'espressione della corrente di trasporto

$$i_T = I_T e^{\frac{V_{BE}}{V_T}} \quad (5.6)$$

con

$$I_T = \frac{qSD_n}{w} \left(\frac{n_i^2}{N_B} \right) \quad (5.7)$$

Abbiamo visto che in un BJT in regione attiva diretta scorre una corrente di elettroni attraverso la base. D'altra parte la base è semiconduttore di tipo p e quindi almeno una frazione di questi elettroni si ricombinerà con lacune. Pertanto non tutti gli elettroni immessi dall'emettitore potranno attraversare la base e arrivare fino al collettore. Per ogni elettrone che si ricombina ci sarà una lacuna in meno in base, quindi uno ione negativo (fisso) elettricamente non compensato dalla rispettiva lacuna. In prossimità del reoforo di base si generano coppie elettrone-lacuna per agitazione termica, gli elettroni sono sospinti verso il reoforo dalla carica negativa della base, mentre le lacune sono attratte verso l'interno della stessa. L'effetto complessivo è una corrente positiva entrante dal reoforo che va a compensare la ricombinazione degli elettroni. Oppure, in maniera equivalente, possiamo dire che è come se parte degli elettroni della corrente di trasporto uscissero dal reoforo di base anziché arrivare fino al collettore. Se la base è sottile questa frazione è piccola e quindi la (5.6) rimane valida. Per il calcolo della corrente di base entrante è sufficiente calcolare il numero di ricombinazioni al secondo che avvengono nel volume della base e moltiplicarla per la carica unitaria dell'elettrone.

$$i_B = q \int_V r n p dV \quad (5.8)$$

D'altra parte il numero di lacune (portatori maggioritari) in base rimane sempre molto maggiore del numero di elettroni (portatori minoritari) e quindi

$$p \approx N_B \quad (5.9)$$

Pertanto

$$i_B = q r N_B S \int_{x=0}^{x=w} n(x) dx \quad (5.10)$$

con S sezione della base. Definendo

$$\tau = \frac{1}{r N_B} \quad (5.11)$$

possiamo scrivere

$$i_B = \frac{q S}{\tau} \int_{x=0}^{x=w} n(x) dx \quad (5.12)$$

D'altra parte, poiché la concentrazione n degli elettroni in base va linearmente da $n(0)$ a $n(w) = 0$

$$i_B = \frac{q S}{\tau} \frac{n(0)}{2} w = q S w \frac{1}{2 \tau} \left(\frac{n_i^2}{N_B} \right) e^{\frac{V_{BE}}{V_T}} \quad (5.13)$$

Dividiamo ora i_C (pari a i_T) per i_B

$$\frac{i_C}{i_B} = \frac{2D_n \tau}{w^2} \quad (5.14)$$

Ricordando l'espressione per la lunghezza di diffusione

$$L_n = \sqrt{\tau D_n} \quad (5.15)$$

si ottiene l'importante relazione lineare tra corrente di collettore e corrente di base in un BJT in regione attiva diretta

$$i_C = \beta_F i_B \quad (5.16)$$

$$\beta_F = 2 \left(\frac{L_n}{w} \right)^2 \quad (5.17)$$

La (5.17) lega il guadagno in corrente di un BJT alle sue caratteristiche costruttive: il guadagno aumenta al diminuire dello spessore della base e all'aumentare della lunghezza di diffusione degli elettroni in base, ovvero al diminuire della concentrazione delle impurità p in base.

5.5 Regione attiva inversa

In regione attiva inversa la giunzione tra base e collettore è polarizzata direttamente ($v_{BC} > 0.6V$) e la giunzione tra base ed emettitore è polarizzata inversamente ($v_{BE} < 0.6V$). Per la simmetria della struttura del BJT npn è ragionevole aspettarsi che il comportamento in regione attiva inversa sia analogo al caso della regione attiva diretta, con l'unica differenza che la corrente scorrerà dall'emettitore al collettore e il rapporto β_R tra corrente di emettitore e corrente di base sarà diverso da β_F , a causa dei diversi drogaggi del collettore e dell'emettitore (l'emettitore è molto più drogato del collettore). In effetti mentre in regione attiva diretta β_F è compreso tra 50 e 200, in regione attiva inversa β_R è compreso tra 0-20. Proprio per questo ridotto guadagno di corrente raramente un BJT è impiegato in regione attiva inversa.

5.6 Interdizione

In interdizione entrambe le giunzioni sono polarizzate inversamente ($v_{BE} < 0.6V$, $v_{BC} < 0.6V$), quindi nessuna delle due giunzioni inietta portatori minoritari in base, né può scorrere una corrente di base. Semplicemente tutte le correnti sono nulle.

5.7 Saturazione

In saturazione entrambe le giunzioni sono polarizzate direttamente ($v_{BE} > 0.6\text{V}$, $v_{BC} > 0.6\text{V}$). La tensione tra collettore ed emettitore in generale è data da

$$v_{CE} = v_{BE} - v_{BC} \quad (5.18)$$

D'altra parte, nel Capitolo 2 abbiamo già detto che difficilmente la tensione ai capi di un diodo supera 0.8V , quindi

$$|v_{CE}| \leq 0.2\text{V} \quad (5.19)$$

Nel seguito assumeremo spesso che se la v_{CE} cala fino a 0.2V , il BJT entra in saturazione. Ciò a rigore non è vero: per alcuni BJT la v_{CE} potrebbe scendere ulteriormente prima di entrare in saturazione. In ogni caso per una prima analisi dei circuiti non fa molta differenza e quindi di solito assumeremo che in saturazione $v_{CE} = 0.2\text{V}$. Se, però, in alcuni esercizi risulterà più comodo assumere in saturazione $v_{CE} = 0\text{V}$, lo faremo senz'altro.

5.8 Caratteristica di uscita del BJT

Consideriamo il circuito (ideale) in Figura 5.12. In questo circuito l'emettitore si dice "a comune" perché il reoforo dell'emettitore è comune sia alla maglia di ingresso sia alla maglia di uscita.

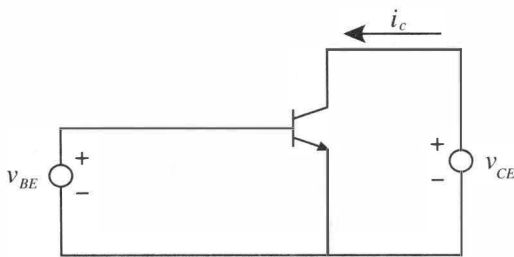


Figura 5.12 BJT in configurazione a emettitore comune

Fissata la v_{BE} , la corrente i_c in funzione di v_{CE} è la cosiddetta **caratteristica di uscita del BJT**. L'andamento è quello mostrato in Figura 5.13.

Dove le curve sono piatte il BJT è in regione attiva diretta, dove sono rapidamente crescenti è in saturazione. Il confine non è così netto e si sposta a valori più

elevati per le curve più alte, ma per una prima analisi dei circuiti possiamo assumere che per $v_{CE} \leq 0.2V$ il BJT sia in saturazione. In molti casi si potrà assumere anche che in saturazione $v_{CE} \approx 0V$.

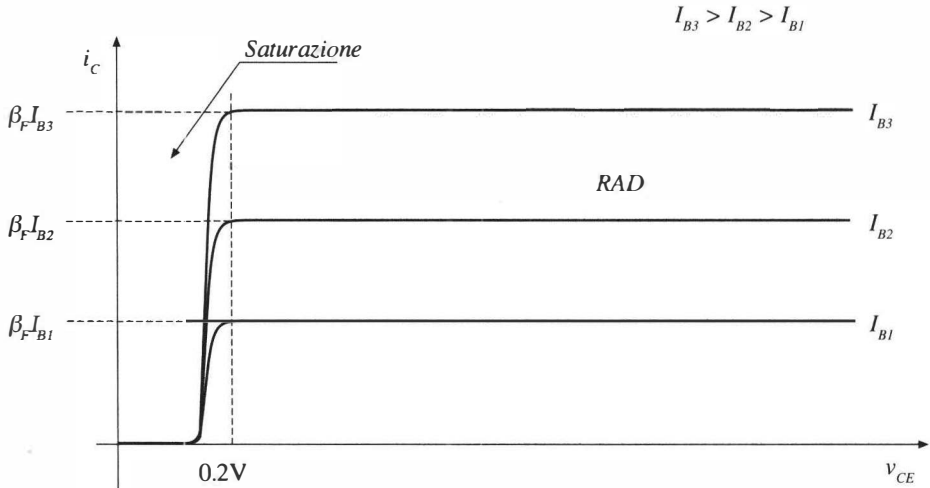


Figura 5.13 Caratteristiche di uscita del BJT

5.9 Effetto Early

Caratteristica fondamentale del BJT in regione attiva diretta è il fatto che la corrente di collettore non dipende dalla tensione tra collettore ed emettitore. In realtà, come è ragionevole aspettarsi, la corrente di collettore non è del tutto indipendente dalla tensione tra collettore ed emettitore: le curve caratteristiche non sono piatte, ma crescono leggermente all'aumentare della v_{CE} . Infatti, fissata la v_{BE} :

$$i_C = i_T = I_T e^{\frac{v_{BE}}{V_T}} \frac{qSD_n}{w} \left(\frac{n_i^2}{N_B} \right) e^{\frac{v_{BE}}{V_T}} \quad (5.20)$$

dove w è lo spessore della base. All'aumentare della v_{CE} , poiché la v_{BE} rimane circa costante, la v_{CB} aumenta. La giunzione CB è contropolarizzata e un aumento della v_{CB} produce un ulteriore allargamento della regione di svuotamento a scapito dello spessore della base che si riduce.

Indichiamo con w l'effettivo spessore della base che a rigore è un po' più piccolo dello spessore fisico (w_0) della base a causa delle regioni di svuotamento x_{BE} e x_{BC} .

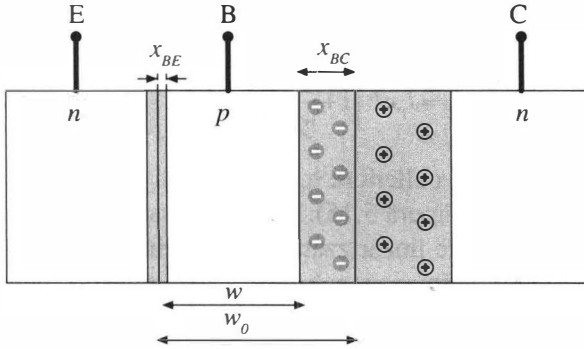


Figura 5.14 Effetto Early

In regione attiva diretta la v_{BE} rimane sempre intorno a 0.6 – 0.7V. La regione di svuotamento x_{BE} è circa costante e molto minore di w_0 , possiamo quindi trascurarla. Al contrario, poiché la giunzione BC è contropolarizzata, la x_{BC} può variare molto.

Lo spessore della regione di svuotamento nella zona di base è dato da:

$$x_{BC} = \sqrt{\frac{2\varepsilon}{q} (\varphi + v_{CB}) \frac{N_C}{N_B} \frac{1}{N_B + N_C}} \quad (5.21)$$

$$x_{BC} = \sqrt{\frac{2\varepsilon}{q} (\varphi + v_{CE} - v_{BE}) \frac{N_C}{N_B} \frac{1}{N_B + N_C}} \quad (5.22)$$

Possiamo ragionevolmente supporre $v_{BE} \approx \varphi$ e $N_B \gg N_C$, quindi

$$x_{BC} \approx \sqrt{\frac{2\varepsilon}{q} (v_{CE}) \frac{N_C}{N_B^2}} \quad (5.23)$$

Lo spessore effettivo della base è quindi

$$w = w_0 - x_{BC} \quad (5.24)$$

Pertanto sostituendo nella (5.20) si ottiene

$$i_C = I_{T0} e^{\frac{v_{BE}}{V_T}} \frac{1}{1 - \frac{1}{w_0} \sqrt{\frac{2\varepsilon}{q} \frac{N_C}{N_B^2} v_{CE}}} \quad (5.25)$$

con

$$I_{T0} = \frac{qSD_n}{w_0} \left(\frac{n_i^2}{N_B} \right) \quad (5.26)$$

Poiché $x_{BC} \ll w_0$ possiamo approssimare

$$i_C = I_T e^{\frac{v_{BE}}{V_T}} \left(1 + \frac{1}{w_0} \sqrt{\frac{2\varepsilon N_C}{q N_B^2}} v_{CE} \right) \quad (5.27)$$

La relazione tra corrente di collettore e v_{CE} è quindi un ramo di parabola (con asse lungo x nel grafico di Figura 5.15). Tuttavia, fissato un valore $v_{CE} = V_{CE}$ per piccole variazioni può essere linearizzata. Con riferimento alla Figura 5.15 per ogni $v_{CE} = V_{CE}$ possiamo definire un valore V_A (detta tensione di Early) per cui

$$\left. \frac{\partial i_C}{\partial v_{CE}} \right|_{v_{CE}=V_{CE}} \approx \frac{I_C}{V_{CE} + V_A} \quad (5.28)$$

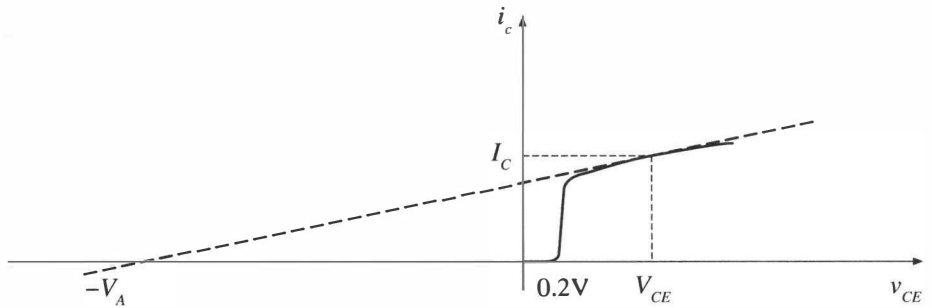


Figura 5.15 Caratteristica di uscita del BJT (effetto Early)

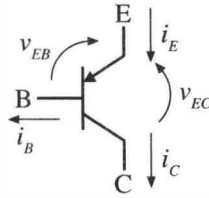
La tensione di Early V_A si calcola sostituendo la (5.26) nella (5.27)

$$V_A = w_0 \frac{N_B}{\sqrt{N_C}} \sqrt{\frac{2q}{\varepsilon}} V_{CE} + V_{CE} \quad (5.29)$$

Nella pratica, la tensione di Early ha tipicamente valori compresi tra 20 e 50V. Dalla (5.30) si deduce che i BJT ad alto guadagno, ovvero con w piccolo, hanno corrente di collettore meno costante al variare di v_{CE} (ovvero V_A non molto grande).

5.10 Il BJT pnp

La teoria del transistor pnp è del tutto analoga a quella dell'nnp. In un pnp in regione attiva diretta la corrente di trasporto scorre dall'emettitore al collettore e la tensione di controllo è la v_{EB} . Il senso positivo delle tensioni e delle correnti è quello mostrato in Figura 5.16.

**Figura 5.16** Transistor pnp

È importante notare una notevole differenza tra i due transistor. Per l'npn vale

$$\beta_{nF} = 2 \left(\frac{L_n}{w} \right)^2 = \frac{2}{w^2} \frac{D_n}{rN_B} \quad (5.30)$$

Ricordando la relazione di Einstein

$$\frac{D_n}{\mu_n} = \frac{kT}{q} \quad (5.31)$$

segue

$$\beta_{nF} = \frac{2}{qw^2} \frac{kT}{rN_B} \mu_n \quad (5.32)$$

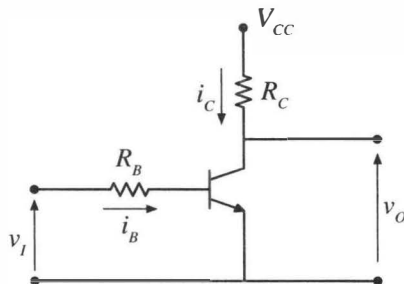
Ragionando in maniera analoga per il pnp si ottiene

$$\beta_{pF} = \frac{2}{qw^2} \frac{kT}{rN_B} \mu_p \quad (5.33)$$

Poiché la mobilità degli elettroni μ_n è circa 2.5 volte la mobilità delle lacune μ_p , a parità di altre condizioni il guadagno in corrente di un npn è 2.5 volte il guadagno di un pnp. Questa è la ragione per cui il transistor npn è più frequentemente utilizzato del pnp nella progettazione di dispositivi elettronici.

5.11 Il BJT come interruttore

Un transistor può essere impiegato come interruttore controllato in tensione. A questo proposito si consideri lo schema in figura.

**Figura 5.17** BJT come interruttore

Tracciamo le caratteristiche di uscita del BJT nel piano $i_C - v_{CE}$ come in Figura 5.18. La presenza della resistenza R_C impone che

$$i_C = \frac{V_{CC} - v_{CE}}{R_C} \quad (5.34)$$

che può essere vista come una retta (retta di carico) nel piano $i_C - v_{CE}$. L'effettivo valore della corrente i_C e della tensione v_{CE} , è dato dall'intersezione della retta di carico con la caratteristica di uscita relativa al valore di corrente in base.

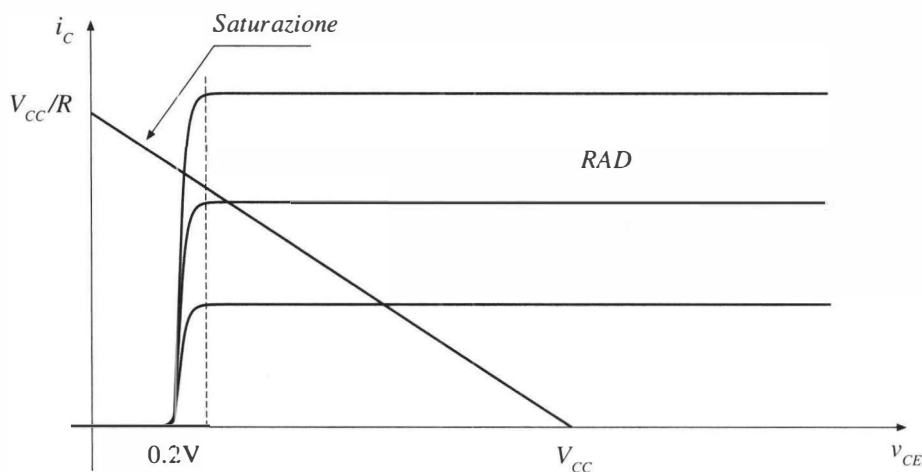


Figura 5.18 Retta di carico e caratteristiche di uscita

Definiamo ora funzione di trasferimento la funzione tra la tensione di uscita v_O (presa sul collettore e quindi pari a v_{CE}) e la tensione di ingresso v_I .

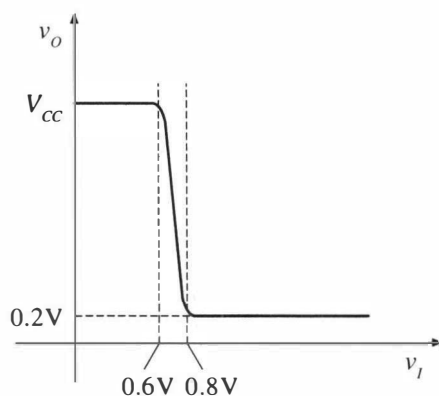


Figura 5.19 Funzione di trasferimento

Quando v_I è minore di 0.6 V la giunzione BE è OFF. La tensione v_{BC} è data da

$$v_{BC} = -v_{CB} = -(V_{CC} - R_C i_C - V_{BE}) = R_C i_C + V_{BE} - V_{CC} \quad (5.35)$$

D'altra parte la caduta di tensione su R_C non potrà essere certamente maggiore della tensione di alimentazione V_{CC} , quindi

$$v_{BC} = R_C i_C + V_{BE} - V_{CC} < V_{BE} < (0.6V) \quad (5.36)$$

ovvero la giunzione BC è OFF. Il BJT è in interdizione. Nel piano della retta di carico ci troviamo nel punto $v_{CE} = V_{CC}$, nella funzione di trasferimento nel tratto $v_O = V_{CC}$.

Appena v_I raggiunge la tensione di soglia, BE è ON mentre BC rimarrà inizialmente OFF. Il BJT è quindi in regione attiva diretta. Per $v_I > 0.6V$ la corrente che scorre nella maglia di ingresso è

$$i_B = \frac{v_I - (0.6V)}{R_B} \quad (5.37)$$

Al variare di i_B la caratteristica di uscita cambia e si sposta verso l'alto. Dalla Figura 5.18 si vede come l'intersezione tra retta di carico e caratteristiche di uscita si sposti verso valori più bassi di v_{CE} , fino ad arrivare alla zona di saturazione.

Finché il BJT è in regione attiva diretta, se in prima approssimazione trascuriamo l'effetto Early, possiamo scrivere:

$$i_C = \beta_F \frac{v_I - (0.6V)}{R_B} \quad (5.38)$$

quindi la caduta di tensione su R_C (pari a $R_C i_C$) aumenta con v_I e la tensione di uscita cala. Tuttavia quando la tensione tra collettore ed emettitore arriva a circa 0.2V il BJT entra in saturazione. In questo ragionamento abbiamo però trascurato il fatto che in realtà v_{BE} non rimane costante a 0.6V, ma aumenta leggermente fino a 0.8V quando il BJT entra in saturazione. Quindi il valore della tensione d'ingresso per cui il BJT entra in saturazione può essere stimato come segue:

$$v_O = v_{CE} = V_{CC} - R_C i_C = V_{CC} - \frac{\beta_F R_C}{R_B} (v_I - (0.8V)) = 0.2V \quad (5.39)$$

ovvero

$$v_I = \frac{R_B}{\beta_F R_C} (V_{CC} - (0.2V)) + (0.8V) \quad (5.40)$$

Se R_B non è molto più grande di R_C , poiché β_F è grande il primo addendo del secondo termine è piccolo, quindi la condizione su v_I per cui il BJT entra in saturazione è semplicemente

$$v_I \approx 0.8V \quad (5.41)$$

Si noti che la tensione v_{CE} non può diminuire ulteriormente quando il BJT entra in saturazione, ovvero si stabilizzerà alla soglia della saturazione. Si ricordi inoltre che non sempre la v_{CE} è 0.2V alla soglia della saturazione. Per BJT ad alto guadagno la v_{BE} rimane sempre comunque poco superiore a 0.6V e la v_{CE} alla soglia della saturazione può essere anche di pochi mV.

Tornando al circuito in esame, possiamo concludere che la funzione di trasferimento è alta quando l'ingresso è basso ($< 0.6V$) ed è bassa quando l'ingresso è alto, di fatto il circuito analizzato è già una possibile porta NOT se impiegato in un circuito digitale.

5.12 Il BJT come amplificatore di tensione

Supponiamo che un BJT ad emettitore comune come in Figura 5.20 sia mantenuto in regione attiva diretta mediante una rete lineare. Su come realizzare tale rete torneremo nei paragrafi successivi. Tale rete potrà sempre essere vista come il suo equivalente di Thévenin, ovvero un generatore di tensione V_B e una resistenza R_B (per un richiamo sui concetti di rete lineare ed equivalente di Thévenin vedi Appendice B). Poiché la corrente di base è piccola, supporremo la connessione tra base ed emettitore un circuito aperto.

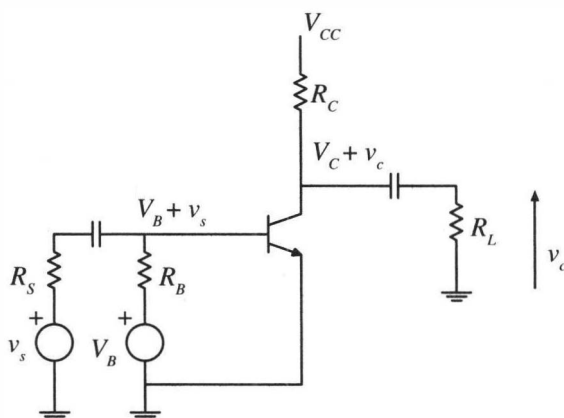


Figura 5.20 Amplificatore a BJT

Consideriamo ora un generatore di segnale, ovvero un generatore di tensione v_s variabile nel tempo con una sua resistenza di uscita R_s , collegato alla base mediante un capacitore di capacità molto grande (il dimensionamento di tale capacitore sarà trattato nel Capitolo 8). Tale capacitore, detto **capacitore di accoppiamento**, è un circuito aperto per la tensione costante V_B , ma è un cortocircuito per il segnale v_s variabile nel tempo. Pertanto la componente in continua della tensione nel gate sarà pari a V_B , mentre la componente variabile nel tempo sarà pari a v_s . Le due componenti sono indipendenti e si sommano, quindi

$$v_{BE} = V_B + v_s \quad (5.42)$$

Applicando la tensione data dalla (5.41) alla funzione di trasferimento, risulta che la tensione sul collettore v_c sarà la somma di un valore costante V_C e un segnale v_c in controfase rispetto a v_s (ovvero quando v_s aumenta v_c diminuisce).

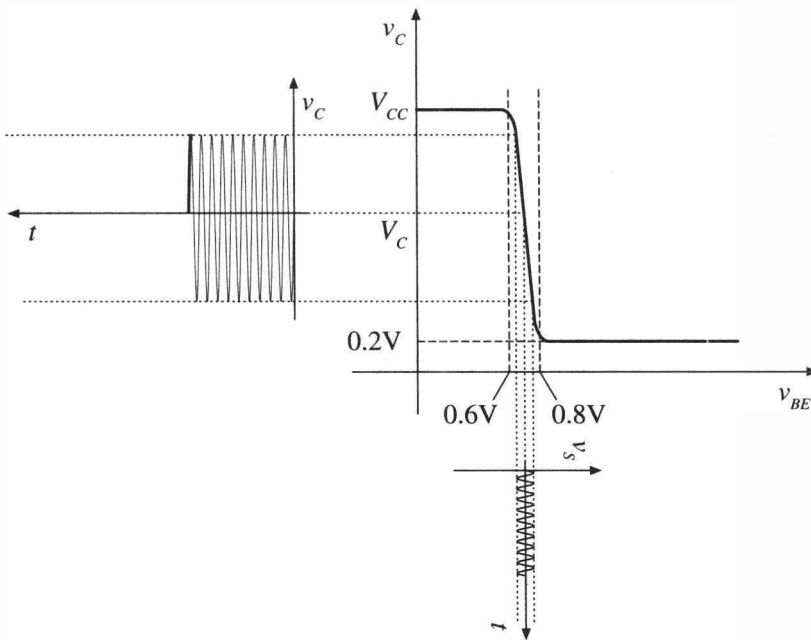


Figura 5.21 Amplificazione del segnale

Poiché la pendenza della funzione di trasferimento è maggiore di 1 (in valore assoluto), l'ampiezza di v_c è maggiore di quella del segnale in ingresso. Abbiamo pertanto realizzato un amplificatore di tensione.

Resta ora da vedere come estrarre il segnale dal collettore e trasferirlo a una resistenza di carico R_L . Ciò può essere fatto mediante un altro capacitore di capacità abbastanza grande posto tra il collettore e il carico. Infatti per la componente in continua della tensione di collettore il capacitore è un circuito aperto e pertanto essa non viene trasferita al carico. Per la componente variabile il capacitore è un cortocircuito e pertanto essa è interamente trasferita al carico, ovvero sul carico c'è la stessa tensione variabile che c'è sul drain.

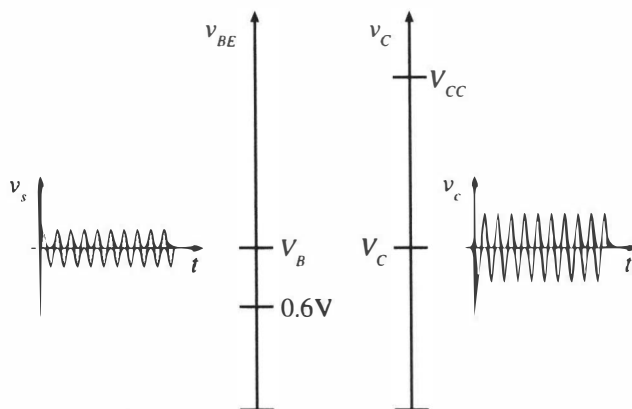


Figura 5.22 Tensioni in continua e variabili

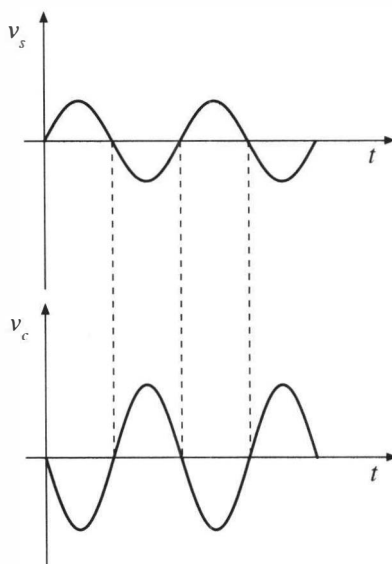


Figura 5.23 Segnale in ingresso e in uscita

5.13 Polarizzazione del BJT

Un modo molto semplice per polarizzare un BJT in regione attiva diretta è il circuito in Figura 5.24.

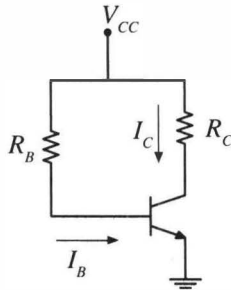


Figura 5.24 Polarizzazione del BJT

La corrente di base è data da

$$I_B = \frac{V_{CC} - V_{BE}}{R_B} \quad (5.43)$$

dove V_{BE} può essere considerata costante e pari a circa 0.6V. La corrente di collettore è quindi

$$I_C = \beta_F I_B \quad (5.44)$$

La condizione da rispettare affinché il BJT sia in regione attiva diretta è

$$R_C I_C < V_{CC} - 0.2V \quad (5.45)$$

Questa rete di polarizzazione è semplice ma ha il problema che il dimensionamento dei resistori dipende da β_F che è un parametro che ha una forte variabilità anche a parità di modello e tecnologia costruttiva.

In effetti il costruttore non fornisce mai il valore esatto di β_F , ma solo il valore minimo garantito.

5.14 Polarizzazione del BJT mediante rete a quattro resistori

Un soluzione molto efficace per polarizzazione un BJT è la rete a quattro resistori mostrata in Figura 5.25.

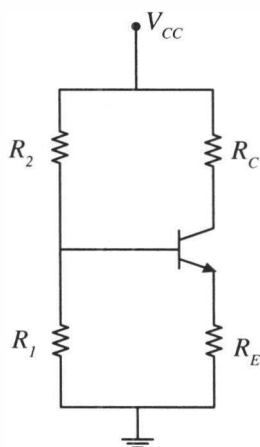


Figura 5.25 Rete di polarizzazione a quattro resistori

Il circuito può essere spezzato in due circuiti come in Figura 5.26.

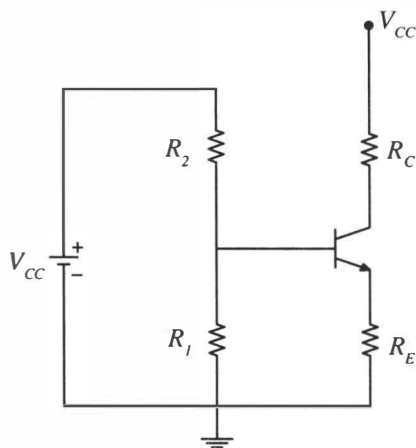


Figura 5.26 Polarizzazione del BJT

Il circuito in ingresso può essere sostituito con il suo equivalente di Thévenin:

$$V_{Th} = \frac{R_1}{R_1 + R_2} V_{CC} \quad (5.46)$$

$$R_{Th} = \frac{R_1 R_2}{R_1 + R_2} \quad (5.47)$$

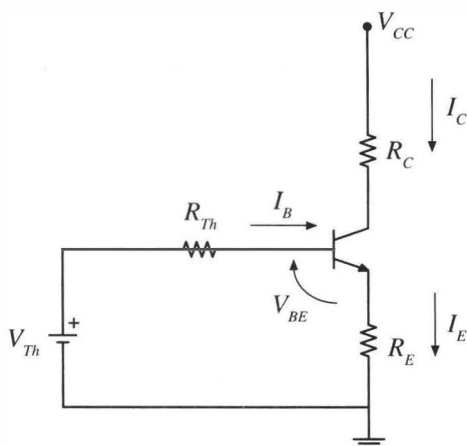


Figura 5.27 Polarizzazione del BJT

Considerando la maglia di ingresso possiamo scrivere:

$$V_{Th} - V_{BE} - R_{Th} I_B - R_E I_E = 0 \quad (5.48)$$

Poiché

$$I_E = I_B + I_C = (1 + \beta_F) I_B \approx \beta_F I_B = I_C \quad (5.49)$$

La (5.49) diviene

$$V_{Th} - V_{BE} = R_{Th} I_B + R_E I_C \quad (5.50)$$

Tuttavia

$$R_{Th} I_B \ll R_E I_C \quad (5.51)$$

Infatti questa condizione è equivalente a

$$\frac{I_C}{I_B} \gg \frac{R_{Th}}{R_E} \quad (5.52)$$

ovvero

$$\beta_F \gg \frac{R_{Th}}{R_E} \quad (5.53)$$

Poiché β_F è grande, la (5.53) di solito è ben verificata.

Quindi dalla (5.49) si ottiene:

$$I_C = \frac{V_{Th} - V_{BE}}{R_E} \quad (5.54)$$

D'altra parte V_{BE} in regione attiva diretta varia poco ed è circa 0.6V, quindi

$$I_C = \frac{V_{Th} - (0.6V)}{R_E} \quad (5.55)$$

Se V_{Th} e R_E sono scelti in modo che

$$V_{CE} > 0.2V \quad (5.56)$$

ovvero

$$V_{CC} - I_C(R_C + R_E) > 0.2V \quad (5.57)$$

il BJT è polarizzato in regione attiva diretta.

Le possibili variazioni di V_{Th} e V_{BE} , dovute a qualunque causa, da una variazione della tensione di alimentazione a una variazione termica che modifica la V_{BE} , per la (5.54) sono attenuate di un fattore pari alla resistenza di emettitore R_E .

La rete di polarizzazione a quattro resistori è una delle molteplici applicazioni della reazione negativa in elettronica. In generale la reazione negativa è un meccanismo che stabilizza una grandezza in uscita, usando la stessa uscita per controllare l'ingresso. Nello specifico la grandezza da stabilizzare è la corrente di collettore in modo da mantenere il BJT in regione attiva diretta.

La resistenza R_E fa da meccanismo di reazione. La corrente che scorre su R_E è la corrente di emettitore che è circa uguale alla corrente di collettore. Un aumento della corrente di collettore produce un aumento della tensione ai capi di R_E , ovvero un aumento della tensione di emettitore. D'altra parte la tensione di base è mantenuta costante dal partitore costituito da R_1 e R_2 . Quindi V_{BE} diminuisce all'aumentare della corrente di collettore. Ma, poiché la corrente di collettore dipende esponenzialmente dalla V_{BE} , essa diminuisce al diminuire di V_{BE} . In definitiva, un aumento della corrente produce un aumento della tensione su R_E che, a sua volta, produce una diminuzione della corrente. Possiamo quindi concludere che questo *loop* tra corrente di collettore e tensione tra base ed emettitore stabilizza entrambi.

5.15 Polarizzazione del BJT mediante generatore di corrente

In alternativa alla resistenza di emettitore si può utilizzare un generatore di corrente come in Figura 5.28.

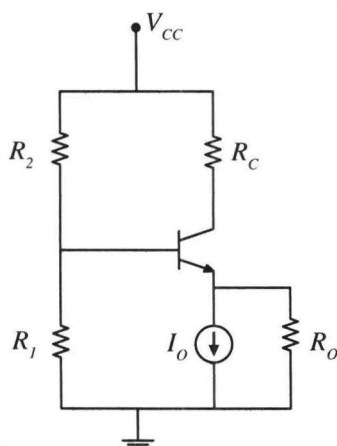


Figura 5.28 Polarizzazione con generatore di corrente

Per tale circuito, se R_o è grande, la corrente è fissata dal generatore di corrente

$$I_C \approx I_O \quad (5.58)$$

Si noti che le variazioni della corrente di collettore dovute a variazioni dell'alimentazione o della V_{BE} sono attenuate di un fattore pari alla resistenza R_o che, essendo la resistenza di uscita di un generatore di corrente, è tipicamente molto grande.

5.16 Specchio di corrente a BJT

Il generatore di corrente a cui ci siamo riferiti nel paragrafo precedente può essere realizzato mediante uno specchio di corrente a BJT, come in Figura 5.29.

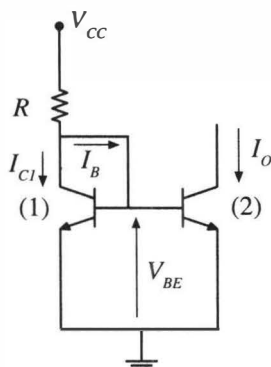


Figura 5.29 Specchio di corrente

I due BJT sono realizzati insieme sullo stesso substrato, quindi sono perfettamente identici. Il BJT (1) è sicuramente in regione attiva diretta in quanto ha collettore e base cortocircuitati. Se anche il BJT (2) è in regione attiva diretta, i due transistor hanno stessa corrente di collettore in quanto le due basi sono cortocircuitate e gli emettitori sono a comune (entrambi collegati a massa). D'altra parte poiché la corrente di base I_b è piccola rispetto alle correnti di collettore, la corrente di collettore del primo BJT è semplicemente data da

$$I_{C1} = \frac{V_{CC} - (0.7V)}{R} \quad (5.59)$$

E quindi

$$I_O = I_{C1} = \frac{V_{CC} - (0.7V)}{R} \quad (5.60)$$

In altre parole, forzando il primo BJT a mantenere una data corrente di collettore, questo a sua volta forza il BJT gemello a mantenere il valore dato.

La resistenza di uscita di questo generatore è la pendenza della caratteristica di uscita di un BJT in regione attiva diretta data dall'effetto Early ed è pertanto

$$R_O \approx \frac{V_A}{I_O} \quad (5.61)$$

che è un valore piuttosto alto (dell'ordine di 20-50Ω).

Si noti che inserendo lo specchio di corrente nel circuito di Figura 5.28, la stabilità del punto di polarizzazione del BJT dipende dalla stabilità del punto di lavoro dello specchio di corrente.

5.17 BJT e FET

BJT e FET hanno molte analogie e significative differenze. Per fissare le idee consideriamo un BJT npn e un MOSFET a canale n ad arricchimento.

Nel BJT la corrente tra collettore ed emettitore è controllata dalla tensione tra base ed emettitore, in maniera analoga nel FET la corrente tra drain e source è controllata dalla tensione tra gate e source.

La tensione di soglia per cui il transistor comincia a condurre per il BJT è sempre circa 0.6V, mentre per un MOSFET è un valore determinato dalle caratteristiche costruttive del transistor.

Il parametro fisico che rende conto del guadagno nel BJT è il β_F , nel MOSFET è il K_n . Dimensionalmente sono parametri diversi: il primo è un rapporto tra la corrente di collettore e la corrente di base ed è un numero puro; il secondo è una costante che, moltiplicata per il quadrato della differenza tra tensione di controllo (v_{GS}) e la soglia, dà la corrente di drain.

Emerge subito il fatto che mentre il BJT controlla la corrente di collettore essenzialmente mediante la corrente che entra in base, il MOSFET controlla la corrente di drain direttamente mediante una tensione. Ciò fa sì che nel MOSFET non sia necessario che il segnale da amplificare possa fornire una corrente, e questo è un decisivo vantaggio in molte applicazioni. Di contro il BJT è in grado sia di pilotare una corrente (di collettore) maggiore, sia di fornire un guadagno tensione-corrente molto più grande, ovvero piccole variazioni di tensione in ingresso (tra base ed emettitore) producono grandi variazioni nella corrente di collettore.

In un BJT le variazioni di tensione in ingresso sono sempre molto piccole, pertanto la v_{BE} è sempre prossima alla soglia di 0.6V (al più abbiamo detto arriva intorno a 0.8V). Questo non è vero per il MOSFET, la cui tensione di controllo può discostarsi significativamente dalla soglia. Il fatto che in un BJT la tensione tra base ed emettitore è sempre circa 0.6V è una significativa facilitazione in fase di progetto e di analisi perché consente di assumere praticamente costante la tensione di controllo.

5.18 La nascita dell'industria dei semiconduttori

Il BJT è stato il grande protagonista della nascita dell'industria dei semiconduttori. Nel 1952, cinque anni dopo la sua invenzione, la Bell organizzò un seminario che può essere considerato la data di nascita dell'industria dei semiconduttori. Quaranta aziende, pagando \$25,000 di anticipo sulle royalties del brevetto del BJT, assistettero a una full-immersion di 9 giorni sulla rivoluzionaria tecnologia dei semiconduttori. Parteciparono grandi aziende come GE e RCA, ma anche piccolissime aziende come la Texas Instruments. E fu proprio quest'ultima che due anni dopo stupì i concorrenti producendo il primo transistor al silicio. La tecnologia originale era infatti basata sul germanio.

Nel 1955, William Shockley, uno dei tre inventori del transistor, in contrasto con gli altri due fondò a Palo Alto in California la "Shockley Semiconductor Laboratory". Fu la prima azienda di semiconduttori in quella che qualche anno dopo verrà chiamata la "Silicon Valley". L'impresa commerciale fu un totale fallimento, Shockley si rivelò un manager dispotico e irragionevole, puntando tutto su un nuovo transistor a quattro strati pnpn che non ebbe successo. Tuttavia i brillanti ingegneri che aveva attratto nella sua impresa fondarono poi le aziende che sarebbero diventate le leader mondiali del settore: Fairchild, AMD, Intel.

Intanto, la Texas Instruments confermava la sua fama di azienda brillante e innovativa: nel 1958, Jack Kilby alla Texas realizzò il primo circuito integrato. L'enorme importanza scientifica, e non solo tecnologica, di questo fatto fu riconosciuta a distanza di quarant'anni con il conferimento a Kilby del Premio Nobel per la Fisica nel 2000.

Alla fine degli anni cinquanta le basi della grande industria dei semiconduttori erano già ben solide. Tutto era pronto per la grande rivoluzione elettronica che avrebbe caratterizzato i decenni successivi.

5.19 Riepilogo

- Il transistor bipolare (BJT) è un dispositivo a semiconduttore che può operare come interruttore elettronico oppure come amplificatore.
- I BJT sono di due tipi: npn e pnp. Il più comune è l'npn.
- Il transistor bipolare ha tre terminali: collettore, base ed emettitore. La corrente tra collettore e emettitore è controllata dalla corrente di base.
- Un BJT ha quattro regioni di funzionamento: regione attiva diretta, regione attiva inversa, saturazione, interdizione.
- In regione attiva diretta, la corrente di collettore i_C è legata linearmente alla corrente di base i_B dalla seguente relazione:

$$i_C = \beta_F i_B$$

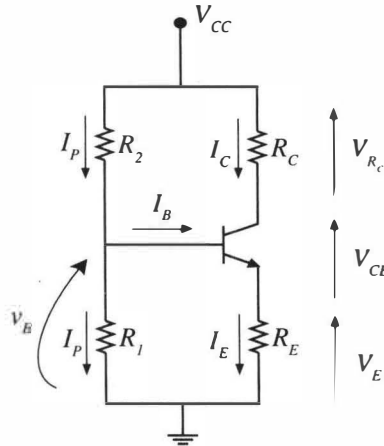
$$\beta_F = 2 \left(\frac{L_n}{w} \right)^2$$

con L_n lunghezza di diffusione dei portatori minoritari in base, w larghezza della base.

- In saturazione la tensione tra collettore ed emettitore è $\leq 0.2V$.
- In interdizione tra collettore ed emettitore non scorre corrente.
- Un BJT che lavora come interruttore commuta tra la regione di saturazione e la regione di interdizione passando per la regione attiva diretta.
- Un BJT polarizzato in regione attiva diretta può operare come amplificatore.
- In regione attiva diretta la corrente è solo approssimativamente costante al variare della tensione tra collettore ed emettitore, infatti aumenta leggermente all'aumentare della tensione per effetto Early.

5.20 Esercizi svolti

1) Polarizzare un BJT in regione attiva diretta con una rete di polarizzazione a quattro resistori in modo che la corrente di collettore sia pari a 1mA. La tensione di alimentazione è 5V.



Per progettare la rete di polarizzazione conviene cominciare fissando le tensioni: V_{R_C} , V_{CE} , V_E . Il vincolo è che la loro somma sia pari a $V_{CC} = 5V$ e che $V_{CE} > 0.2V$. Scegliamo, ad esempio: $V_{R_C} = 2V$, $V_{CE} = 2V$ e $V_E = 1V$. Poiché $I_C = 1mA$, segue che:

$$R_C = \frac{V_{R_C}}{I_C} = 2k\Omega$$

$$R_E = \frac{V_E}{I_C} = 1k\Omega$$

Dobbiamo ora dimensionare R_1 e R_2 . Innanzitutto facciamo in modo che le due resistenze costituiscano un partitore di tensione, ovvero che la corrente I_P che vi scorre sia molto maggiore della corrente I_B che entra in base.

$$I_B = \frac{I_C}{\beta_F}$$

β_F è nell'ordine di 50-200. Quindi al massimo la corrente di base sarà pari a $20\mu A$. Scegliamo quindi la corrente I_P almeno 10 volte tanto, quindi $I_P = 0.2mA$.

La tensione V_B è pari a

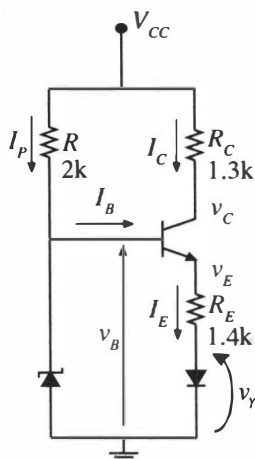
$$V_B = V_E + 0.6V = 1.6V$$

Quindi:

$$R_1 = \frac{V_B}{I_P} = 8k\Omega$$

$$R_2 = \frac{V_{CC} - V_B}{I_P} = 17k\Omega$$

2) Con riferimento alla figura, si calcoli la corrente di collettore. La tensione di alimentazione è 5V. Lo Zener ha tensione di breakdown pari a 3V.



La tensione in base è fissata a 3V dallo Zener. La tensione V_E è quindi pari a

$$V_E = V_B - 0.6V = 2.4V$$

La tensione sulla resistenza R_E è pari a

$$V_{R_E} = V_E - V_Y = V_E - 0.6V = 1.8V$$

Quindi

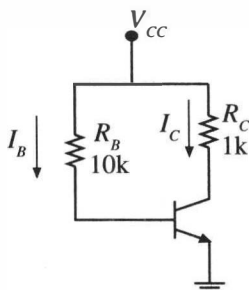
$$I_E \approx I_C = \frac{1.8V}{1.4k\Omega} = 1.29mA$$

Controlliamo che il BJT sia in regione attiva diretta

$$V_C = V_{CC} - I_C R_C = 3.32V$$

$$V_{CE} = 3.32V - 2.4V = 0.9V > 0.2V$$

3) Con riferimento alla figura si calcoli la corrente di polarizzazione. $V_{CC} = 5V$, $\beta_F = 200$. In che regione di funzionamento si trova il BJT?



Cominciamo con supporre che il BJT sia in regione attiva diretta. La corrente di base è pari a

$$I_B = \frac{V_{CC} - (0.6V)}{R_B} = 0.44mA$$

Quindi:

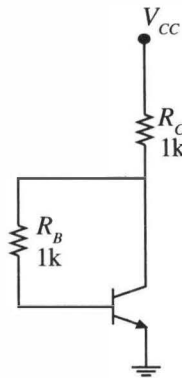
$$I_C = \beta_F I_B = 88mA$$

La caduta di tensione su R_C sarà quindi pari a

$$V_{R_C} = I_C R_C = 88V$$

Ma attenzione: 88V è maggiore della tensione di alimentazione! Evidentemente l'ipotesi che il BJT si trovasse in regione attiva diretta era sbagliata. Il BJT è quindi in saturazione: V_{CE} è pari a 0.2V e $V_{R_C} = 5V - 0.2V = 4.8V$.

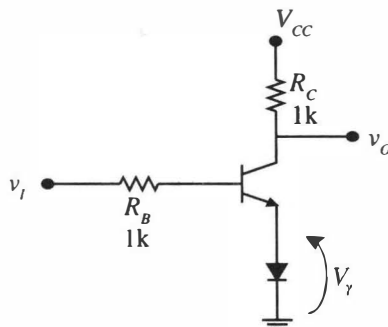
4) Con riferimento alla figura si calcoli la corrente di collettore.



La corrente di base è sicuramente piccola rispetto alla corrente di collettore. Se la supponiamo nulla è nulla anche la caduta di tensione sulla resistenza di base, quindi $v_{CB} = 0$ e $v_{CE} = v_{BE} = 0.6V$. Pertanto la corrente di collettore è pari a

$$I_C \approx \frac{V_{CC} - V_{BE}}{R_C} = 4.4mA$$

5) Con riferimento alla figura si calcoli la funzione di trasferimento tra tensione di ingresso e tensione di uscita. $V_{CC} = 5V$.

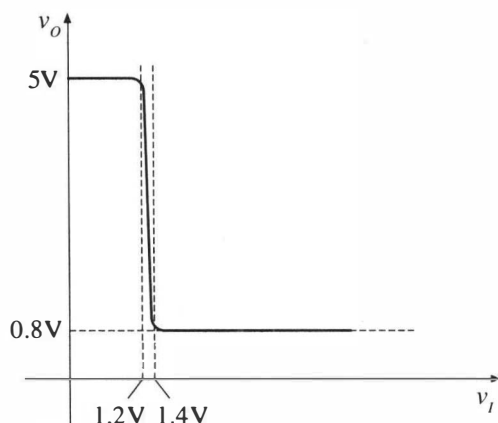


Cominciamo notando che la corrente di base, per un BJT in interdizione o in regione attiva diretta, è piccola e quindi la caduta di tensione su R_b è trascurabile. Questa ipotesi non è però vera quando il BJT è in saturazione. In effetti lo scopo della R_b è proprio limitare la corrente in base quando il BJT è in saturazione.

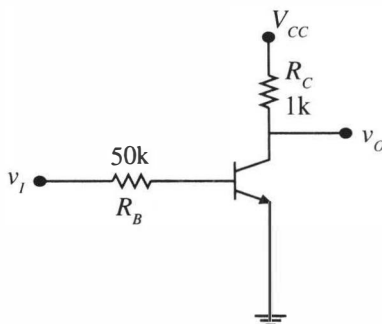
Per $v_i < 0.6V + V_\gamma = 1.2V$ non può scorrere corrente nella maglia di ingresso e quindi il BJT è in interdizione. In un BJT in interdizione la corrente di collettore è nulla e quindi nulla anche la caduta di tensione sulla resistenza R_c . Segue che la tensione di uscita è pari alla tensione di alimentazione.

Per $v_i > 1.2V$, scorre corrente di collettore, cade una tensione su R_c e la tensione di uscita cala, finché il BJT non va in saturazione per $v_i = 0.8V + 0.6V = 1.4V$. In saturazione la tensione di uscita sarà pari a $v_o = 0.2V + V_\gamma = 0.8V$.

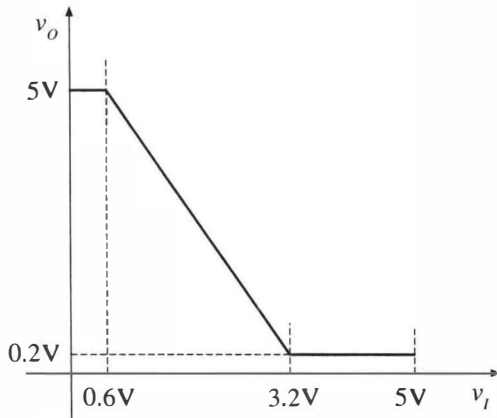
Un possibile dubbio nel risolvere questo esercizio potrebbe essere quale valore dare alla tensione del diodo quando il BJT è in saturazione. Si potrebbe pensare che entrambi siano a $0.8V$ in quanto vi scorre la stessa corrente e non sarebbe sbagliato. Tuttavia i diodi sono di solito progettati perché vi possa scorrere una discreta corrente senza che la tensione vari molto, quindi nella pratica è più corretto supporre $0.6V$ la tensione del diodo anche quando il BJT è in saturazione.



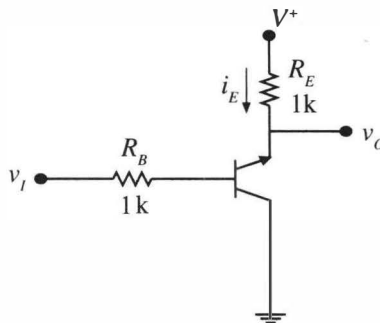
6) Con riferimento alla figura si calcoli la funzione di trasferimento tra tensione di ingresso e tensione di uscita. $\beta_F = 100$, $V_{cc} = 5V$.



Si noti che R_B è grande e quindi anche la piccola corrente di base è in grado di produrre una caduta di tensione non trascurabile, per di più R_C è piuttosto piccola quindi la corrente di collettore può diventare piuttosto grande e dar luogo a una corrente di base non trascurabile. Tenendo conto di questo l'andamento è quello in figura.



7) Con riferimento alla figura si calcoli la funzione di trasferimento. $\beta_F = 100$, $\beta_R = 10$, $V_{CC} = 5V$.



Si noti che il BJT è montato con il collettore a massa (ovvero è “capovolto” rispetto ai casi che abbiamo fin qui visto).

Quando $v_I < 0.6V$ il BJT è interdetto e quindi l'uscita è a 5V.

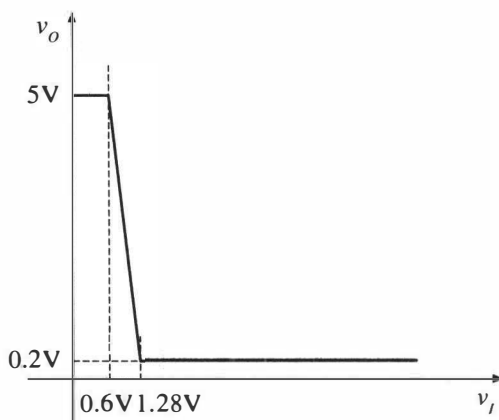
Quando $v_I > 0.6V$ il BJT è in regione attiva inversa.

$$i_E = \beta_R i_B$$

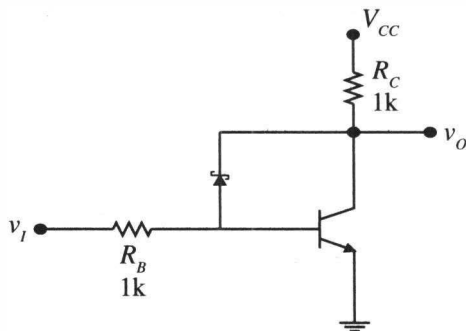
$$i_E = \beta_R \frac{v_I - (0.6V)}{R_B}$$

$$v_O = V_{CC} - R_E i_E = V_{CC} - \beta_R \frac{R_E}{R_B} (v_I - (0.6V))$$

IL BJT entra in saturazione quando $v_O = 0.2\text{V}$ e $V_{BE} = 0.8\text{V}$ quindi per $v_I = 1.28\text{V}$.

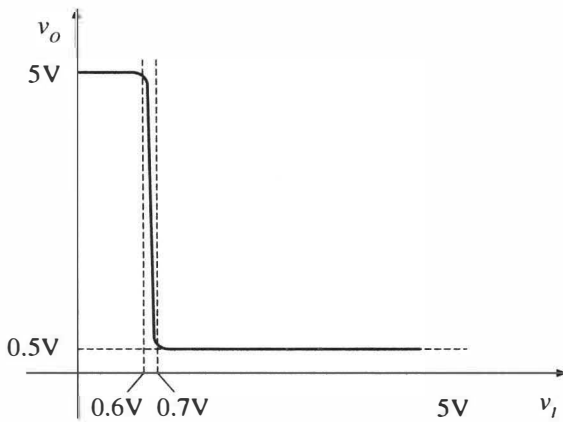


8) Con riferimento alla figura si calcoli la funzione di trasferimento tra tensione di ingresso e tensione di uscita. La tensione di polarizzazione del diodo Schottky è 0.2V . $V_{CC} = 5\text{V}$.

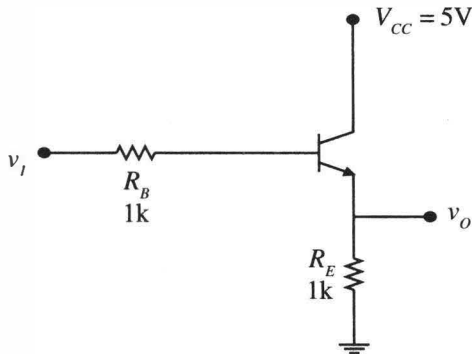


Per $v_I < 0.6\text{V}$ non può scorrere corrente nella maglia di ingresso (il diodo Schottky è contro-polarizzato) e quindi il BJT è in interdizione. In un BJT in interdizione la corrente di collettore è nulla e quindi nulla anche la caduta di tensione sulla resistenza R_C . Segue che la tensione di uscita è pari alla tensione di alimentazione.

Per $v_I > 0.6\text{V}$, scorre una corrente di collettore che produce una caduta su R_C . Tuttavia il BJT non va mai in saturazione perché quando $v_{BE} - v_O > 0.2\text{V}$ lo Schottky entra in conduzione. Quindi il valore minimo della tensione di uscita sarà pari a $v_{BE} - 0.2\text{V} = 0.5\text{V}$ che raggiunge quando $v_I \approx 0.7\text{V}$ (poco oltre la soglia).



9) Si calcoli la funzione di trasferimento (tensione di uscita in funzione della tensione in ingresso) del seguente circuito.

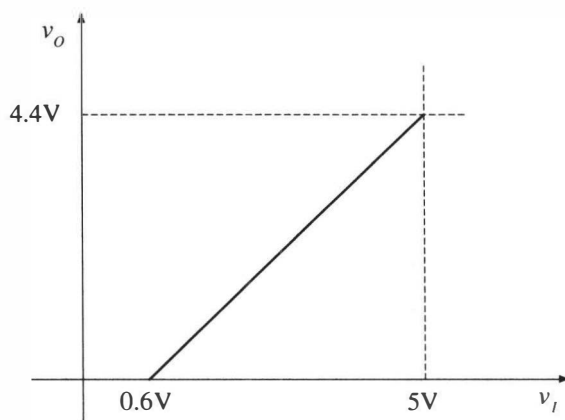


Per $v_i < 0.6\text{V}$ non può scorrere corrente nella maglia di ingresso e quindi il BJT è in interdizione. In un BJT in interdizione la corrente di collettore/emettitore è nulla e quindi nulla anche la caduta di tensione sulla resistenza R_E . Segue che la tensione di uscita è pari a zero.

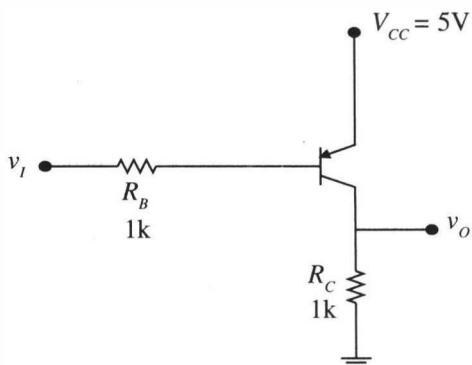
Per $v_i > 0.6\text{V}$, scorre una corrente di collettore che produce una caduta su R_E . Trascurando la caduta di tensione su R_B , la tensione di uscita sarà pari a

$$v_o = v_i - (0.6\text{V})$$

ovvero cresce linearmente con la tensione di ingresso. Quando $v_i = 5\text{V}$, $v_o = 4.4\text{V}$. Si noti che il BJT non entra mai in saturazione.

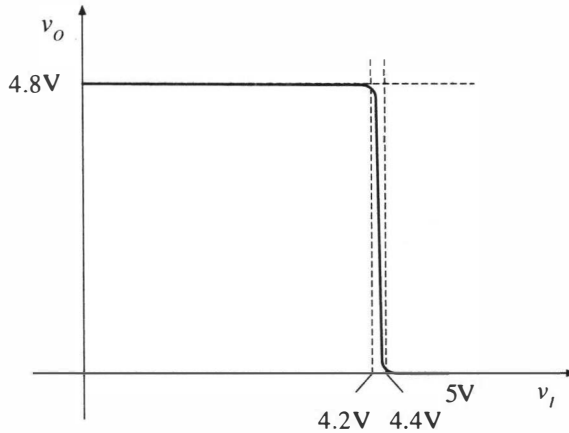


10) Si calcoli la funzione di trasferimento (tensione di uscita in funzione della tensione in ingresso) del seguente circuito, quando v_I varia tra 0 e 5V.

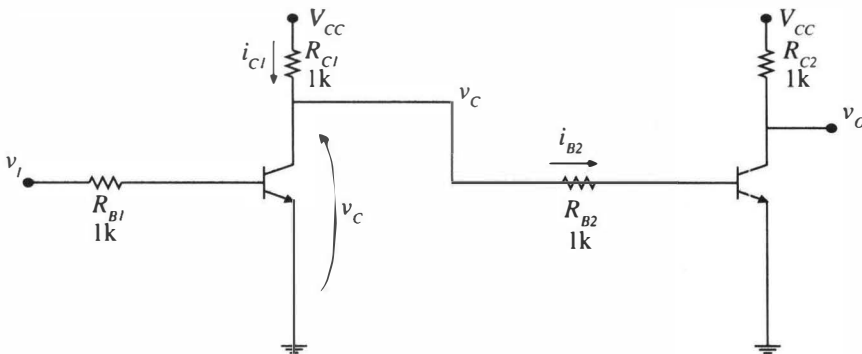


Per $v_I > 5\text{V} - 0.6\text{V} = 4.4\text{V}$ non può scorrere corrente tra emettitore e base e quindi il BJT è in interdizione. In un BJT in interdizione la corrente di collettore è nulla, così come la caduta di tensione su R_C . Quindi $v_O = 0\text{V}$.

Per $v_I < 4.4\text{V}$, scorre una corrente di collettore che produce una caduta su R_C finché il BJT non va in saturazione per $v_I = 5\text{V} - 0.8\text{V} = 4.2\text{V}$. In saturazione la tensione di uscita sarà pari a $v_O = 5\text{V} - 0.2\text{V} = 4.8\text{V}$.



11) Con riferimento allo schema in figura, si tracci la v_c e v_o in funzione di v_i . $V_{CC} = 5V$.



Per $v_i < 0.6V$ non può scorrere corrente nella maglia di ingresso e quindi il BJT è in interdizione ed è come se tra collettore ed emettitore ci fosse un circuito aperto, la corrente quindi scorre dall'alimentazione attraverso R_{C1} e R_{B2} fino alla giunzione BE del secondo BJT che pertanto verrà portato in saturazione.

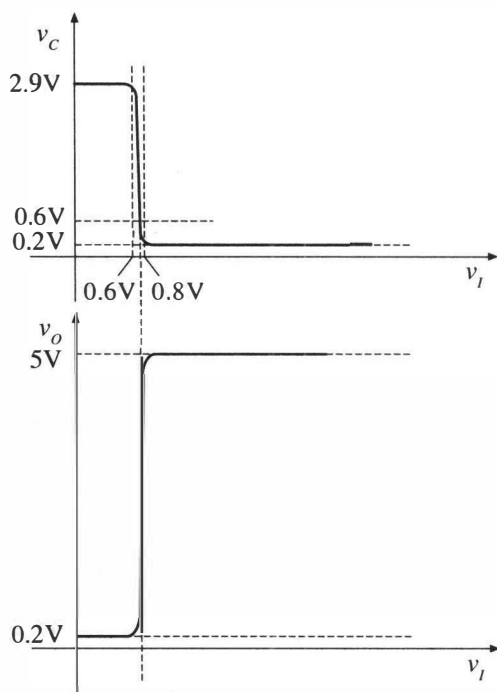
$$i_{C1} = i_{B2} = \frac{V_{CC} - V_{BE2}}{R_{C1} + R_{B2}} = 2.1mA$$

La tensione v_c è quindi

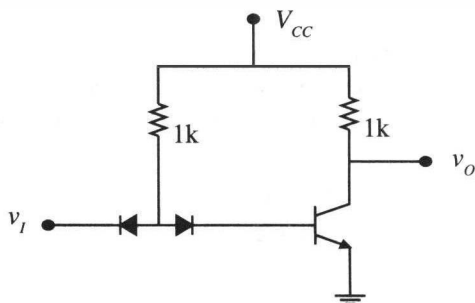
$$v_c = V_{CC} - i_{C1}R_{C1} = 2.9V$$

Per $v_i > 0.6V$, scorre una corrente di collettore che produce una caduta su R_{C1} e quindi v_c cala finché il primo BJT non va in saturazione per $v_i = 0.8V$.

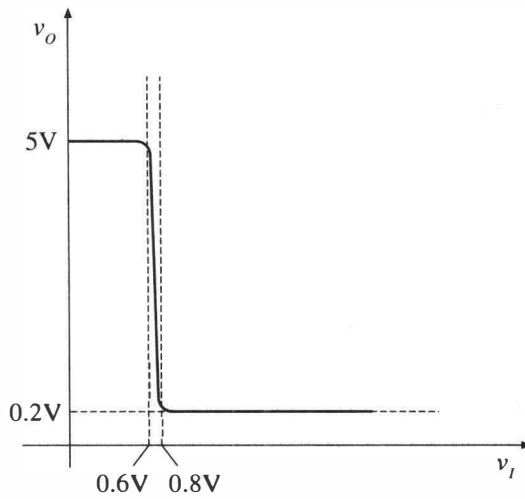
Per tracciare la tensione in uscita, si può procedere per via grafica tenendo presente che quando v_c è maggiore di 0.6V il secondo BJT è in conduzione e quindi l'uscita bassa, quando v_c è minore di 0.6V è in interdizione e quindi l'uscita è alta.



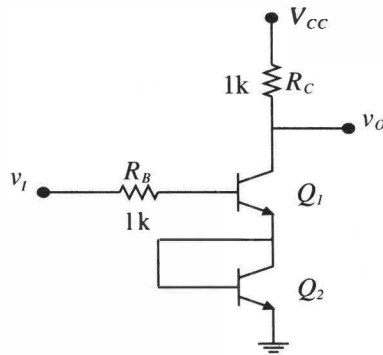
12) Si calcoli la funzione di trasferimento. $V_{cc} = 5V$.



Quando v_I è minore di 0.6V il secondo diodo e il BJT sono in interdizione, quindi $v_o = 5V$. Per $v_I > 0.6V$ il secondo diodo e il BJT sono in conduzione e v_o cala fino a 0.2V.



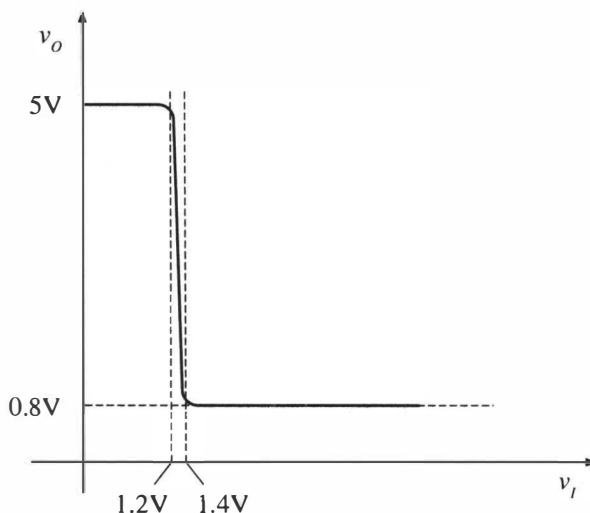
13) Con riferimento alla figura si tracci la funzione di trasferimento. I due BJT sono uguali. $V_{CC} = 5\text{V}$.



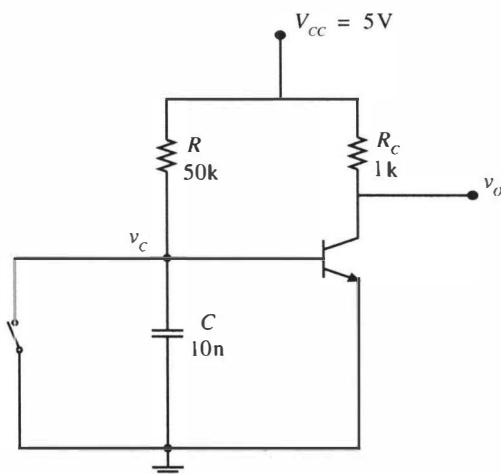
Per $v_i < v_{BE1} + v_{BE2} = 0.6\text{V} + 0.6\text{V} = 1.2\text{V}$ i due transistor sono interdetti.

Per $v_i > 1.2\text{V}$ sono in conduzione e Q_1 va in saturazione quando $v_i > 0.8\text{V} + 0.6\text{V} = 1.4\text{V}$.

La tensione di uscita quando Q_1 è in saturazione è $v_o = 0.2\text{V} + 0.6\text{V} = 0.8\text{V}$.



14) Il BJT in figura ha $\beta = 50$. Per $t < 0$ l'interruttore è ON. All'istante $t = 0$ l'interruttore passa da ON a OFF. Si tracci l'andamento nel tempo per $t > 0$ della tensione in uscita.



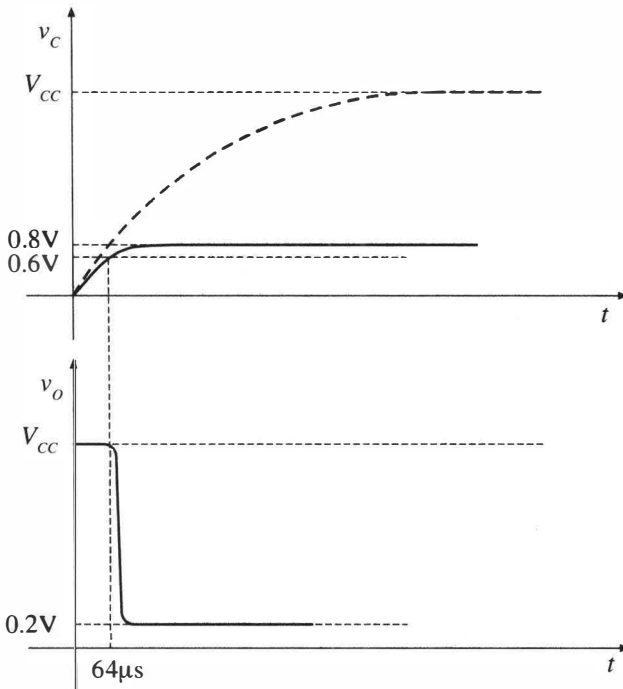
Quando l'interruttore è ON la base del BJT è forzata a massa e quindi il BJT è interdetto. L'uscita v_o è pari a 5V. Quando l'interruttore è OFF il condensatore C si carica attraverso la resistenza R_i seguendo la legge

$$v_c(t) = V_{CC} \left(1 - e^{-\frac{t}{R_i C}} \right)$$

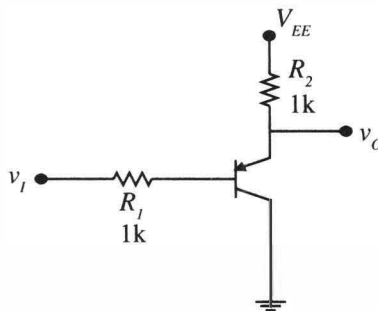
Il BJT passa dall'interdizione alla conduzione quando la tensione sul condensatore raggiunge 0.6V:

$$0.6 = 5 \left(1 - e^{-\frac{t}{0.5\text{ms}}} \right)$$

da cui: $t = 64\mu\text{s}$.

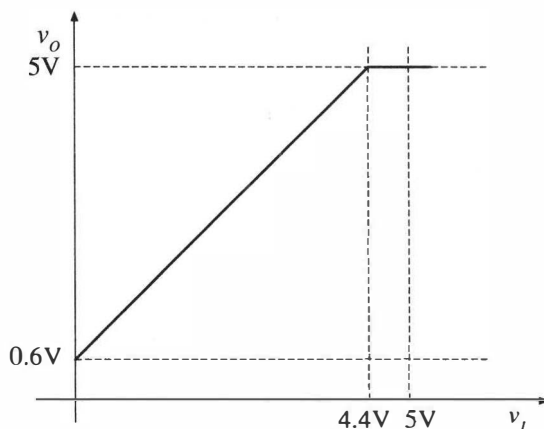


15) Tracciare la funzione di trasferimento. $V_{EE} = 5\text{V}$.

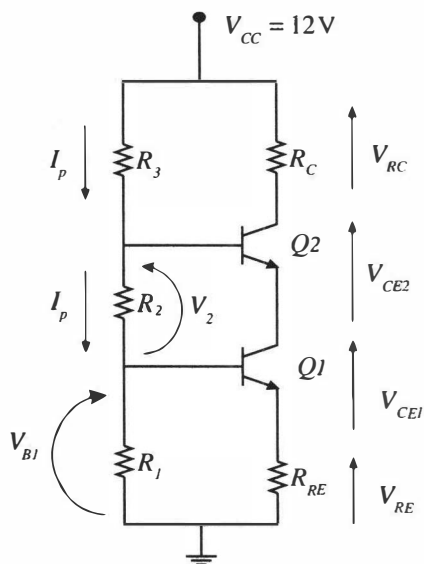


Quando $v_I > 5\text{V} - 0.6\text{V} = 4.4\text{V}$ il BJT pnp è in interdizione e l'uscita è pari a $V_{EE} = 5\text{V}$.

Quando $v_I < 4.4\text{V}$ il BJT è in regione attiva diretta e $v_O = v_I + 0.6\text{V}$.



16) Dimensionare i resistori in modo che entrambi i BJT siano in regione attiva diretta e che scorra su entrambi una corrente di collettore di 2mA.



Per risolvere il problema è necessario innanzitutto fissare le tensioni sulla maglia contenente i collettori dei due BJT, l'unico vincolo è che le tensioni tra collettore ed emettitore siano maggiori di 0.2V e che la loro somma sia 12V. Fissiamo, ad esempio:

$$V_E = V_{CE1} = V_{CE2} = V_{RC} = 3V$$

Poiché la corrente sulla stessa maglia è 2mA.

$$R_E = \frac{3V}{2mA} = 1.5k\Omega$$

$$R_C = \frac{3V}{2mA} = 1.5k\Omega$$

La tensione sulla base di Q_1 è

$$V_{B1} = V_E + (0.6V) = 3.6V$$

Dobbiamo fissare la corrente che scorre sulla maglia contenente R_1 , R_2 e R_3 . Questa dovrà essere molto minore delle correnti di base dei due BJT. Fissiamo ad esempio $I_p = 1mA$.

$$R_1 = \frac{V_{B1}}{I_p} = 3.6k\Omega$$

Per dimensionare R_2 è necessario calcolare la caduta di tensione sulla stessa resistenza.

$$V_2 = V_{CB1} + V_{BE2} = (V_{CE1} - V_{BE1}) + 0.6V = 3V - 0.6V + 0.6V = 3V$$

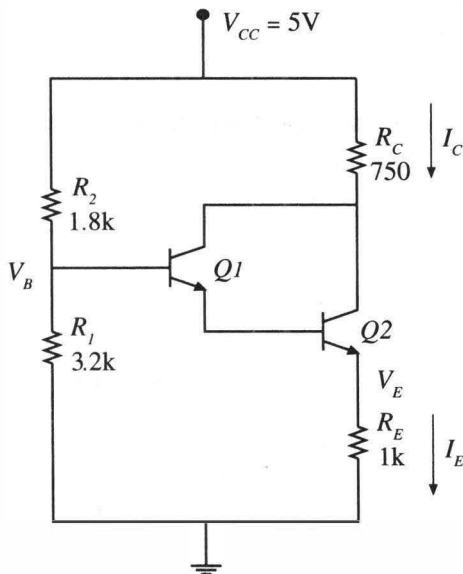
Quindi

$$R_2 = \frac{3V}{I_p} = 3k\Omega$$

Infine

$$R_3 = \frac{12V - 3V - 3.6V}{I_p} = 5.4k\Omega$$

17) Si calcoli la corrente che scorre nella resistenza R_C del circuito in figura.



Supponiamo entrambi i BJT in regione attiva diretta. Trascuriamo le correnti di base.

$$V_B = \frac{R_1}{R_1 + R_2} V_{CC} = 3.2\text{V}$$

Quindi

$$V_E = V_B - 0.6\text{V} - 0.6\text{V} = 2\text{V}$$

$$I_C = I_E = \frac{V_E}{R_E} = 2\text{mA}$$

Resta da verificare che i BJT siano in regione attiva diretta.

La caduta di tensione su R_C è

$$V_{R_C} = I_C R_C = 1.5\text{V}$$

quindi

$$V_{CE2} = V_{CC} - V_{R_C} - V_E = 1.5\text{V} > 0.2\text{V}$$

Poiché la tensione tra collettore ed emettitore di Q_2 è maggiore di 0.2V , Q_2 non è in saturazione.

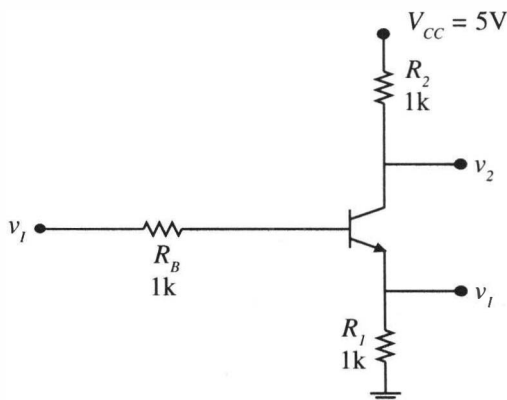
Inoltre:

$$V_{CE1} = V_{CC} - V_{R_C} - V_E - 0.6\text{V} = 0.9\text{V} > 0.2\text{V}$$

quindi anche Q_1 non è in saturazione.

*Nota: la configurazione a 2 BJT esaminata in questo esercizio è nota come coppia **Darlington** ed è molto usata quando è necessario un grande guadagno. L'unico inconveniente è la tensione di soglia piuttosto alta (1.2V).*

***18) Si traccino sullo stesso grafico le tensioni di uscita v_1 e v_2 in funzione della tensione di ingresso v_i .**



Per $v_I < 0.6\text{V}$ il BJT è in interdizione pertanto $v_I = 0\text{V}$, $v_2 = 5\text{V}$.

Per $v_I \geq 0.6\text{V}$ il BJT è in regione attiva diretta e

$$\begin{aligned}v_I &= v_I - (0.6\text{V}) \\i_C &= \frac{v_I - (0.6\text{V})}{R_I} \\v_2 &= V_{CC} - R_2 i_C = V_{CC} - \frac{R_2}{R_I} (v_I - 0.6\text{V}) = V_{CC} + 0.6\text{V} - v_I\end{aligned}$$

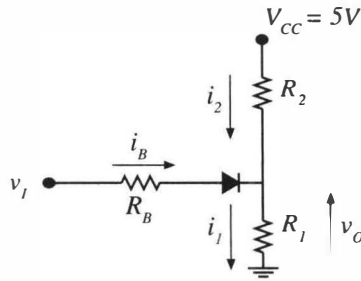
Poiché $v_{CE} = 0.2\text{V}$ quando

$$v_I = v_I - (0.6\text{V}) = \frac{V_{CC} - (0.2\text{V})}{2} = 2.4\text{V}$$

segue che il BJT entra in saturazione per

$$v_I = 3.2\text{V}$$

Vediamo ora cosa succede quando il BJT è in saturazione. Per semplificare l'analisi supponiamo entrambe le giunzioni polarizzate direttamente al valore di 0.7V (in realtà $v_{BE} = 0.8\text{V}$ e $v_{BC} = 0.6\text{V}$, ma non fa molta differenza). In queste condizioni possiamo ridisegnare il circuito come in figura



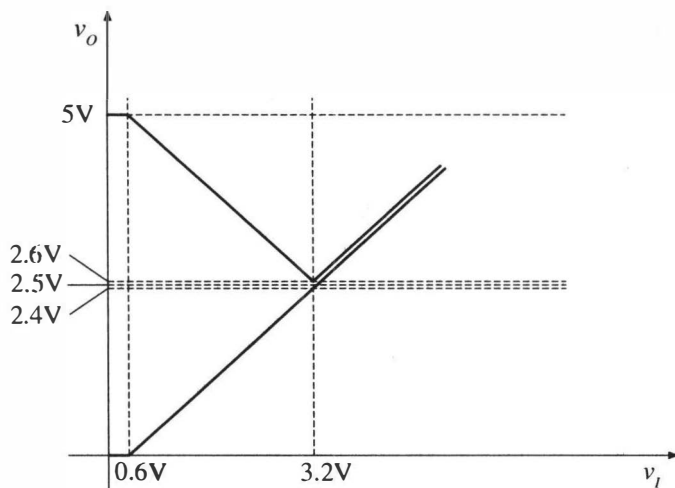
Dalle leggi di Kirchhoff si ottiene il seguente sistema lineare:

$$\begin{aligned}v_I - 0.7\text{V} &= i_B R_B + i_I R_I \\i_B + i_2 &= i_I \\i_I R_I + i_2 R_2 &= 5\text{V}\end{aligned}$$

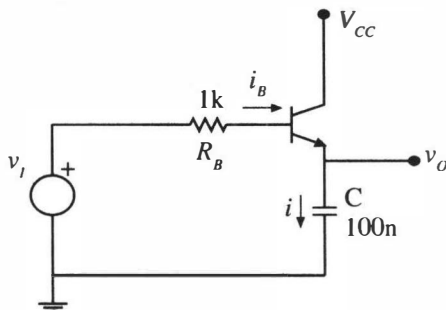
da cui

$$v_O = R_I i_I = \frac{v_I}{3} + 1.43\text{V}$$

Pertanto possiamo concludere che in saturazione le due tensioni in uscita al variare della tensione in ingresso aumentano linearmente con pendenza 0.33.



19) Il generatore di tensione fornisce un gradino da 0V a $V_{CC} = 5V$ molto ripido. Si calcoli il tempo in cui v_o arriva al 90% del suo valore massimo. $\beta_F = 100$.



La corrente di base per $t > 0$ è

$$i_B = \frac{V_{CC} - 0.6V - v_o}{R_B}$$

Quindi la corrente di collettore è data da

$$i = \beta_F \frac{V_{CC} - 0.6V - v_o}{R_B}$$

La corrente di emettitore carica il capacitore C e pertanto deve valere

$$i = C \frac{dv_o}{dt}$$

L'andamento nel tempo di v_o è quindi la soluzione della seguente equazione differenziale:

$$V_{CC} - 0.6V - v_o = \frac{R_B C}{\beta_F} \frac{dv_o}{dt}$$

Definiamo

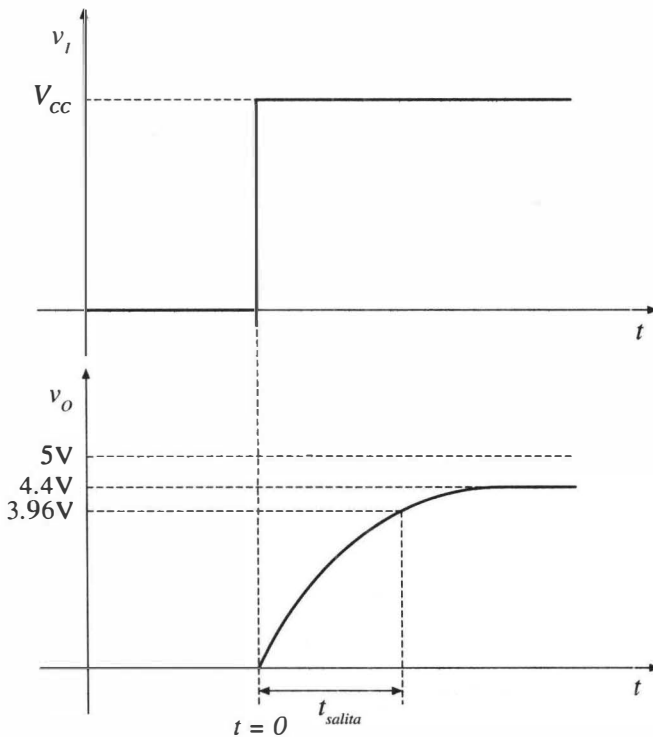
$$\tau = \frac{R_B C}{\beta_F} = 1\mu s$$

La soluzione è

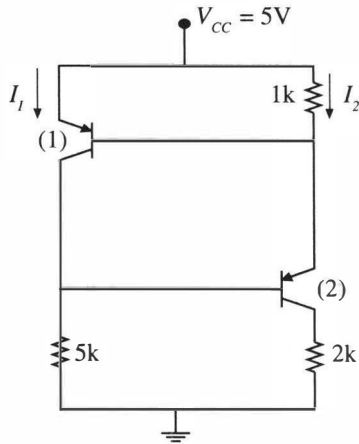
$$v_o(t) = (V_{CC} - 0.6V) \left(1 - e^{-\frac{t}{\tau}}\right)$$

Il tempo in cui arriva al 90% del suo valore massimo ($V_{CC} - 0.6V$) è dato da

$$t_{salita} = \tau \ln(10) = 2.3\mu s$$



20) Calcolare le correnti che scorrono nei due BJT.



Supponiamo i 2 BJT in regione attiva diretta. La corrente I_2 è data da

$$I_2 = \frac{V_{BE1}}{1k\Omega} = \frac{0.6V}{1k\Omega} = 0.6mA$$

La tensione V_{B2} su nodo di base del BJT(2) è

$$V_{B2} = V_{CC} - V_{BE1} - V_{BE2} = 3.8V$$

Quindi

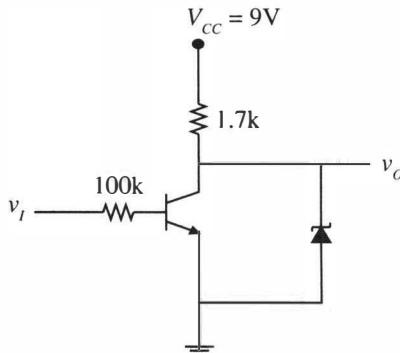
$$I_1 = \frac{V_{B2}}{5k\Omega} = 0.76mA$$

È necessario verificare che i 2 BJT non siano in saturazione

$$V_{CE1} = V_{CC} - I_1(5k\Omega) = 1.2V > 0.2V$$

$$V_{CE2} = V_{CC} - V_{BE1} - I_2(2k\Omega) = 3.2V > 0.2V$$

21) Tracciare la funzione di trasferimento. $V_i = 5.6V$, $\beta_F = 120$.



Per $v_i < 0.6V$ il BJT è interdetto e l'uscita è pari a 5.6V. Si noti che la resistenza R_b è piuttosto grande, quindi sarà necessario considerare la caduta di tensione su tale resistenza.

La tensione sullo Zener potrà scendere a un valore inferiore a 5.6V disattivando lo Zener stesso per

$$i_C = \frac{9V - 5.6V}{1.7k\Omega} = 2mA$$

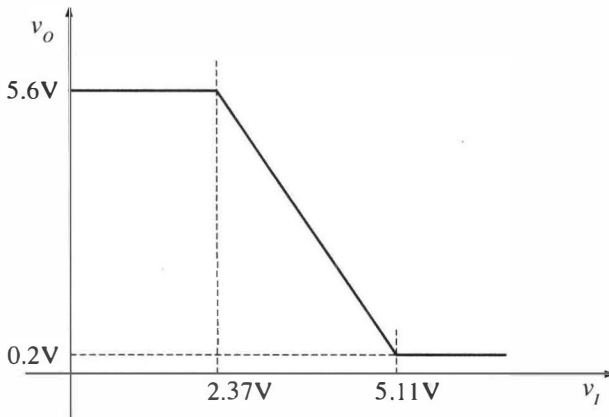
che corrisponde in regione attiva diretta a una corrente di base pari a

$$i_B = \frac{i_C}{\beta_F} = 16.7\mu A$$

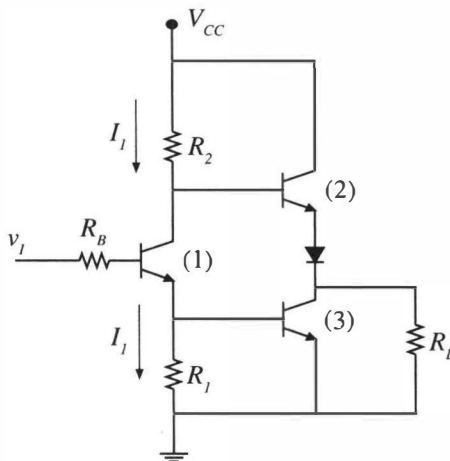
La tensione v_i si calcola tenendo presente che

$$\frac{v_i - 0.7V}{100k\Omega} = 16.7\mu A$$

Da cui $v_i = 2.37V$. In maniera analoga si calcola il valore di tensione in ingresso per cui il BJT entra in saturazione ($v_i = 5.11V$).



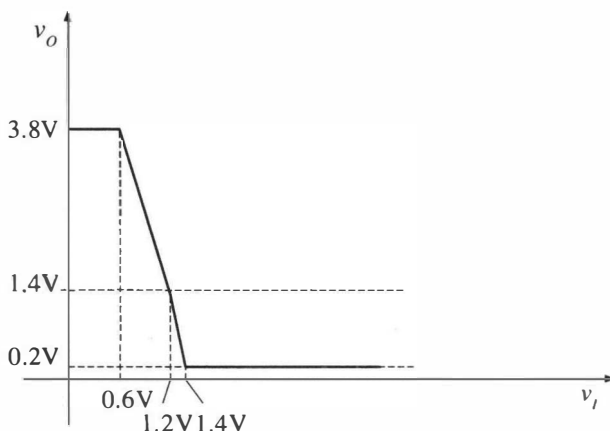
22) Tracciare la funzione di trasferimento per v_i tra 0 e 5V. $R_1 = 1k\Omega$, $R_2 = 4k\Omega$, $R_b = 1k\Omega$, $R_L = 10k\Omega$.



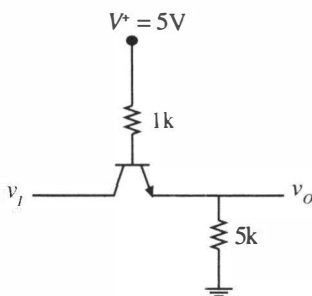
Per $v_i < 0.6V$ il BJT (1) è interdetto. Se BJT (1) è interdetto non scorre corrente su R_1 e quindi $V_{BE3} < 0.6V$. Segue che è interdetto anche il BJT (3). Se trascuriamo la corrente di base del BJT (2), la sua base è praticamente a V_{CC} e quindi l'uscita $v_o = V_{CC} - 0.6V - 0.6V = 3.8V$. Appena v_i supera la soglia scorre corrente su R_1 , quando la tensione su R_1 supera $0.6V$ entra in conduzione anche il BJT (3). Se consideriamo piccola la corrente di base del BJT (1), il valore v_i per cui il BJT (3) entra in conduzione è $v_i = 0.6V + 0.6V = 1.2V$. Calcoliamo il valore della tensione di uscita per $v_i = 1.2V$. $I_1 = (0.6V)/R_1 = 0.6mA$. Quindi

$$v_o = V_{CC} - I_1 R_2 - 0.6V - 0.6V = 1.4V$$

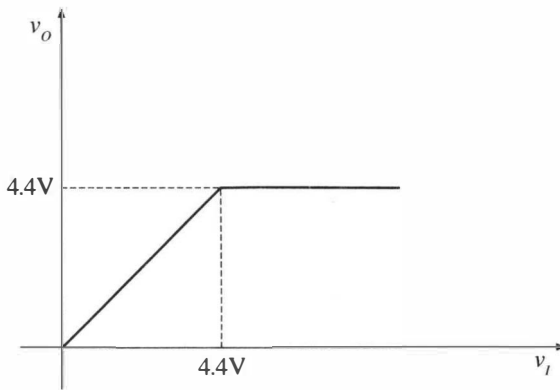
Se v_i aumenta ulteriormente la caduta di tensione su R_1 è tale da mandare il BJT (3) in saturazione ovvero $v_o = 0.2V$. Il valore di v_i per cui il BJT (3) è in saturazione è circa $v_i = 0.8V + 0.6V = 1.4V$



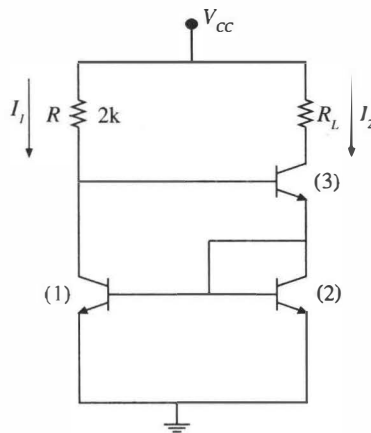
23) Tracciare la funzione di trasferimento per v_i tra 0 e 5V.



Poniamo $v_i = 5V$. Il BJT è in regione attiva diretta. Infatti, se trascuriamo la caduta di tensione sulla resistenza di base $v_o = V^* - 0.6V = 4.4V$ e $V_{CE} = v_i - v_o = 0.6V$. Supponiamo per semplicità che in saturazione $V_{CE} = 0$. Quando $v_i < 4.4V$, il BJT entra in saturazione. Al diminuire di v_i , la V_{BC} rimane fissa a $0.6V$ e quindi l'uscita è pari a $v_o = v_i + V_{BC} - V_{BE} = v_i + 0.6V - 0.6V = v_i$



24) Calcolare la corrente I_2 . $V_{CC} = 5V$.



Nei 3 BJT scorre la stessa corrente. Se il BJT (3) è in regione attiva diretta, allora

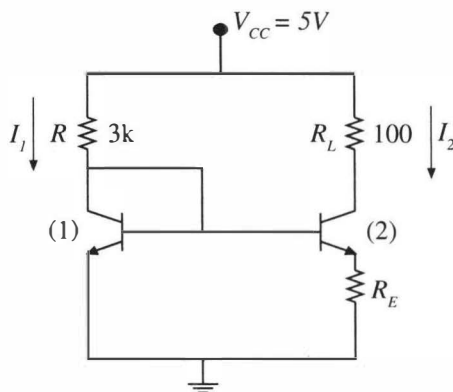
$$I_2 = I_1 = \frac{V_{CC} - V_{BE3} - V_{BE1}}{R} = 1.9\text{mA}$$

Verifichiamo ora che il BJT (3) sia in regione attiva diretta

$$v_{CE}^{(3)} = V_{CC} - I_2 R_L - 0.6V = 4.21V > 0.2V$$

*Nota: Questo circuito, noto come **specchio di Wilson**, è un generatore di corrente molto usato nei circuiti integrati.*

***25) Dimensionare R_E in modo che $I_2 = 0.1\text{mA}$.**



La corrente I_1 è pari a:

$$I_1 = \frac{V_{CC} - V_{BE1}}{R} = 1.47\text{mA}$$

Se i due BJT sono uguali, nell'ipotesi che il BJT (2) sia in regione attiva diretta, possiamo scrivere le seguenti espressioni per le due correnti, con I_T opportuna costante:

$$I_1 = I_T e^{\frac{V_{BE1}}{V_T}}$$

$$I_2 = I_T e^{\frac{V_{BE2}}{V_T}}$$

Da cui

$$\frac{I_1}{I_2} = e^{\frac{V_{BE1} - V_{BE2}}{V_T}} = e^{\frac{R_E I_2}{V_T}}$$

$$R_E = \frac{V_T}{I_2} \ln \left(\frac{I_1}{I_2} \right)$$

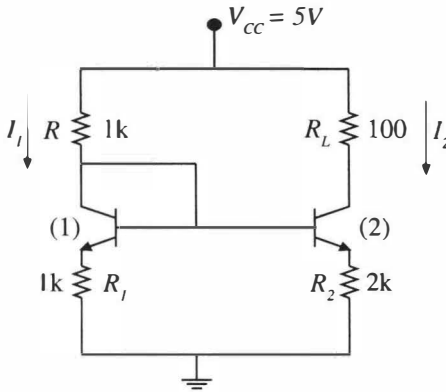
Sostituendo i valori numerici: $R_E = 672\Omega$.

Il BJT (1) è sicuramente in regione attiva diretta. Verifichiamo che lo sia anche il BJT (2):

$$V_{CE} = V_{CC} - I_2(R_E + R_L) = 4.92\text{V} > 0.2\text{V}$$

*Nota: Questo circuito, noto come **specchio di Widlar**, è un generatore di corrente molto usato nei circuiti integrati.*

26) Calcolare la corrente I_2 .



Nell'ipotesi che il BJT (2) sia in regione attiva diretta

$$I_1 = \frac{V_{CC} - V_{BE1}}{R_I + R} = 2.2\text{mA}$$

Scriviamo

$$I_1 R_I + V_{BE1} = I_2 R_2 + V_{BE2}$$

Poiché $V_{BE1} = V_{BE2} = 0.6\text{V}$, segue che

$$I_2 = \frac{R_I}{R_2} I_1 = 1.1\text{mA}$$

Verifichiamo che il BJT (2) sia in regione attiva diretta

$$V_{CE2} = V_{CC} - I_2 (R_L + R_2) = 2.69\text{V} > 0.2\text{V}$$

5.21 Domande di verifica a scelta multipla

1. In un BJT
 - a) la base è la zona più drogata
 - b) l'emettitore è drogato con lo stesso tipo di impurità del collettore
 - c) il collettore è la zona più drogata
 - d) base ed emettitore sono cortocircuitati
2. In un BJT in regione attiva diretta
 - a) scorre una corrente significativa tra collettore e base
 - b) tra collettore e base c'è una giunzione pn polarizzata direttamente
 - c) tra collettore ed emettitore la tensione è minore di 0.2V
 - d) tra base ed emettitore scorre corrente
3. In un BJT in regione attiva inversa
 - a) scorre corrente tra base e collettore
 - b) tra collettore e base c'è una giunzione pn polarizzata inversamente
 - c) tra collettore ed emettitore la tensione è minore di 0.2V
 - d) tra base ed emettitore scorre corrente
4. In un BJT in saturazione
 - a) tra base e collettore c'è un circuito aperto
 - b) tra collettore ed emettitore la tensione è maggiore di 0.2V
 - c) tra base e collettore c'è una giunzione pn polarizzata direttamente
 - d) tra base ed emettitore non scorre corrente
5. In un BJT in regione attiva diretta
 - a) la corrente di collettore aumenta all'aumentare della tensione tra collettore e base
 - b) la corrente di collettore diminuisce sensibilmente all'aumentare della tensione tra collettore ed emettitore
 - c) la corrente di emettitore non può variare in alcun modo
 - d) la corrente di emettitore diminuisce all'aumentare della corrente di base
6. In un BJT in regione attiva diretta
 - a) la corrente di emettitore è uguale alla corrente di base
 - b) la corrente di collettore è legata da una relazione esponenziale alla tensione tra base ed emettitore
 - c) la corrente di collettore è legata da una relazione esponenziale alla tensione tra collettore ed emettitore
 - d) la corrente di collettore è sempre maggiore della corrente di emettitore
7. In un BJT in regione attiva diretta la corrente di collettore aumenta leggermente al variare di v_{CE} perché
 - a) l'aumento di v_{CE} polarizza ulteriormente la giunzione pn tra base ed emettitore che quindi pilota una corrente maggiore
 - b) l'aumento di v_{CE} contropolarizza ulteriormente la giunzione tra collettore e base, riducendo lo spessore della base
 - c) l'aumento di v_{CE} polarizza ulteriormente la giunzione tra collettore e base, che quindi pilota una corrente maggiore
 - d) l'aumento di v_{CE} polarizza ulteriormente la giunzione tra collettore e base, riducendo lo spessore della zona di collettore

8. La caratteristica di uscita di un BJT a emettitore comune
 - a) dipende dalla corrente di base
 - b) mette in relazione la tensione tra base e collettore e la corrente di collettore
 - c) è circa lineare per piccole tensioni tra collettore e base
 - d) è una retta la cui pendenza dipende dalla resistenza tra alimentazione e collettore
9. In un BJT npn in regione attiva diretta
 - a) gli elettroni non possono fluire dalla base al collettore, in quanto tra base e collettore c'è una giunzione pn polarizzata inversamente
 - b) le lacune non possono fluire dalla base all'emettitore, in quanto sono portatori minoritari
 - c) la corrente tra i reofori di base ed emettitore è principalmente determinata dalla ricombinazione degli elettroni in base
 - d) tra collettore ed emettitore non può scorrere una corrente in quanto tra collettore e base c'è una giunzione contropolarizzata che impedisce il passaggio dei portatori
10. In un BJT pnp polarizzato direttamente
 - a) la corrente scorre dalla base al collettore
 - b) la corrente scorre dal collettore all'emettitore
 - c) la corrente in base è uscente
 - d) la corrente all'emettitore è uscente
11. In un BJT in regione attiva diretta
 - a) la corrente di collettore è legata linearmente alla tensione tra base ed emettitore
 - b) la corrente di collettore è legata linearmente alla tensione tra collettore ed emettitore
 - c) la corrente di collettore è molto minore della corrente di base
 - d) la corrente di collettore è circa proporzionale alla corrente di base
12. In un BJT in interdizione
 - a) la tensione di base può essere maggiore della tensione di emettitore
 - b) la tensione tra collettore ed emettitore è nulla
 - c) la tensione tra collettore e base deve essere sempre maggiore di 0.6V
 - d) la tensione tra base ed emettitore è maggiore di 0.6V
13. In un BJT in saturazione
 - a) la corrente di collettore è circa proporzionale alla corrente di base, per questo è impiegato come amplificatore
 - b) la corrente di collettore è circa uguale alla corrente di emettitore
 - c) la corrente di base è trascurabile
 - d) la corrente di collettore, in prima approssimazione, non dipende dalla corrente di base
14. In un BJT npn polarizzato direttamente
 - a) la corrente di base scorre tra base e collettore
 - b) la corrente di base è dovuta alla ricombinazione degli elettroni iniettati in base dall'emettitore
 - c) la corrente di base è nulla
 - d) la corrente di base è circa uguale alla corrente di trasporto

15. In un BJT npn in regione attiva diretta
- La v_{CE} è sempre compresa tra 0.6V e 0.8V
 - La v_{CE} è nulla
 - La corrente di collettore è data da v_{CE}/R_C dove R_C è la resistenza tra collettore e alimentazione
 - In prima approssimazione la corrente di collettore non dipende da v_{CE}
16. Collettore ed emettitore di un BJT npn vengono cortocircuitati, la base è collegata a una tensione di 1V mediante una resistenza R_B (1 k Ω), l'emettitore è collegato a massa. Il BJT
- è in interdizione
 - è in saturazione
 - è in regione attiva diretta
 - è in regione attiva inversa
17. In un BJT npn lo strato sepolto
- ha lo scopo di aumentare la resistenza di collettore
 - è una zona n molto drogata
 - corrisponde alla regione di base
 - non è presente perché è una caratteristica dei soli MOSFET
18. L'effetto Early
- è dovuto all'aumento della caduta di tensione nella regione di collettore
 - è dovuto alla riduzione dello spessore della base
 - è dovuto all'aumento della regione di svuotamento tra base ed emettitore
 - non è presente nei BJT, ma solo nei MOSFET
19. In un BJT in saturazione
- la v_{CE} può essere pari a 0V
 - la v_{CE} è pari a 0.8V
 - la v_{BE} è minore di 0.2V
 - la v_{BC} può essere pari a 0V
20. Per effetto Early, in un BJT in regione attiva diretta
- la v_{CE} è sempre pari a circa 0.2V
 - la corrente di collettore aumenta all'aumentare di v_{CE}
 - il rapporto tra corrente di collettore e corrente di base è costante
 - la v_{CE} ha un valore massimo detto tensione di breakdown
21. Una sola delle seguenti affermazioni è vera.
- La crescita epitassiale è un processo chimico per la rapida ossidazione del silicio
 - La base è interamente costituita da ossido di silicio
 - L'ossido di silicio separa le piste metalliche delle connessioni elettriche
 - Lo strato sepolto è una regione di silicio intrinseco

5.22 Domande di verifica a risposta aperta

1. Disegnare in sezione la struttura fisica di un transistor npn; qual è la funzione dello strato sepolto?
2. Dimostrare che in regione attiva diretta la corrente di collettore è proporzionale alla corrente di base.
3. Dimostrare che a parità di altre condizioni il transistor npn ha guadagno in corrente superiore rispetto al pnp di un fattore 2.5.
4. Spiegare la causa fisica dell'effetto Early.

5.23 Risposte alle domande a scelta multipla

1 b	8 a	15 d
2 d	9 c	16 b
3 a	10 c	17 b
4 c	11 d	18 b
5 a	12 a	19 a
6 b	13 d	20 b
7 b	14 b	21 c

6

Famiglie logiche e memorie

Le famiglie logiche e le memorie sono gli elementi di base dei sistemi elettronici digitali. Questo capitolo, sia per scelta degli argomenti sia per livello di trattazione, è pensato soprattutto per quegli studenti che non hanno nel loro piano di studio un successivo insegnamento di elettronica digitale. Lo scopo è, infatti, fornire il raccordo tra dispositivo fisico e dispositivo logico. Con questo spirito, dopo aver definito i parametri prestazionali di una generica famiglia logica, si descrive la famiglia a interruttori complementari CMOS. Si tratta, infine, il funzionamento delle tre memorie oggi più largamente usate: SRAM, DRAM e FLASH.

6.1 Al largo dell'isola greca di Anticitera...

Al largo dell'isola greca di Anticitera, nell'anno 1900, in un relitto risalente all'87 avanti Cristo, fu trovato un blocchetto metallico molto incrostato e corrosivo. Nessuno gli dette importanza, ma due anni dopo l'archeologo Valerios Stais scoprì che quella che sembrava una pietra era in realtà un meccanismo a ingranaggi estremamente sofisticato: un computer meccanico per il calcolo dei movimenti degli astri. La "macchina di Anticitera" è il più antico computer giunto fino a oggi. Probabilmente ne sono esistiti altri più antichi, forse anche realizzati da popoli diversi, ma certo è che già in epoca ellenistica esistevano sofisticate macchine di calcolo.

Indubbiamente l'informatica è nata molto prima dell'elettronica. In effetti le regole del calcolo automatico, la cosiddetta algebra di Boole, furono pubblicate per la prima volta nel 1854, ben prima della realizzazione di qualsiasi computer elettronico. In linea di principio l'informatica necessita solo di un qualche dispositivo che disponga di più stati (per memorizzare informazione) e possa modificarli (per elaborare l'informazione). Un ingranaggio è l'oggetto più semplice in grado di fare entrambe le operazioni: la posizione angolare marcata dai denti sulla cremagliera memorizza l'informazione, la rotazione permette di modificarla. Il problema delle macchine a ingranaggi non è nel principio di funzionamento, ma in due limitazioni pratiche: commutazioni troppo lente e poca informazione immagazzinabile.

Il primo passo verso la nascita dell'informatica moderna fu l'invenzione dell'interruttore elettronico, ovvero un dispositivo con solo due stati (acceso e spento) ma in grado di commutare molto velocemente non tramite parti meccaniche, ma mediante la variazione di una tensione di controllo. La valvola termionica faceva esattamente questo ma era fragile, ingombrante, e consumava tanta energia. Il transistor risolse ogni problema: veloce più delle valvole e piccolissimo. I sofisticati ingranaggi della macchina di Anticitera si sono trasformati in un elegante disegno di piste di silicio.

6.2 Architettura di base dei sistemi digitali

In linea di principio un qualunque sistema digitale per l'elaborazione di dati (da una rete di computer al piccolo controllore di un ascensore) è costituito da cinque elementi fondamentali: unità logiche, memorie, dispositivi di input/output, clock e un sistema di connessione.

Le **unità logiche** sono dispositivi in grado di implementare funzioni logiche, ovvero di fornire un'uscita alta o bassa in funzione dello stato alto o basso di un certo numero di ingressi.

Le **memorie** sono dispositivi in grado di memorizzare uno stato alto o basso. Le memorie propriamente dette possono memorizzare migliaia (kilo), milioni (Mega) o miliardi (Giga) di stati (bit), ma si fa largo uso anche di memorie più

piccole dette **registri** e di memorie particolari per tenere traccia del numero di fronti di salita (o di discesa) del segnale al suo ingresso (**contatori**). Le singole celle di memoria possono essere indirizzate direttamente, oppure mediante un qualche criterio di accesso, ad esempio: il primo dato che entra è il primo a uscire (**FIFO**: *First Input First Output*); o al contrario il primo dato che entra è l'ultimo a uscire (**LIFO**: *Last Input First Output*).

I **dispositivi di input e output (I/O)** sono le interfacce del sistema digitale. Le uscite nel caso più semplice possono essere luci, suoni o tensioni a certi terminali, in altri casi l'output può essere l'immagine sullo schermo o uno stampato. Gli ingressi possono essere i caratteri digitati in una tastiera, tensioni su certi terminali, segnali veicolati mediante onde elettromagnetiche.

Il **clock** è un oscillatore che, per così dire, dà il tempo e sincronizza le azioni di tutti i dispositivi. Tipicamente un sistema digitale ha un unico clock, per evitare problemi di sincronizzazione tra clock diversi, ma non è sempre così: una rete di computer è un sistema a più clock; un computer e la sua stampante sono un sistema digitale ad almeno due clock.

I dispositivi digitali devono essere in qualche modo collegati per poter interagire e costituire un sistema. Il **sistema di connessione** può essere semplicemente costituito da connessioni elettriche punto-punto tra un dispositivo e l'altro, oppure si può trattare di un sistema più complesso detto **bus**, che gestisce in maniera flessibile le comunicazioni tra molti dispositivi.

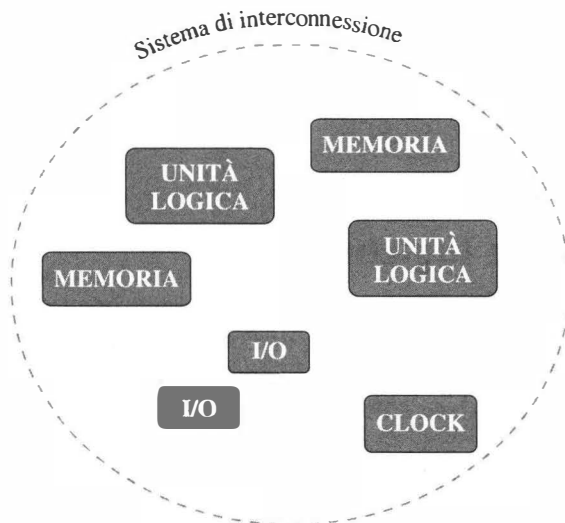


Figura 6.1 Architettura di un sistema digitale

L'architettura tipica di un computer è ancora oggi essenzialmente quella proposta per la prima volta da Von Neumann nei primi anni cinquanta: un'unità logica detta CPU (*Central Processing Unit*), una grande memoria le cui celle possono essere indirizzate direttamente (RAM: *Random Access Memory*) e un bus che connette unità logica, memoria e dispositivi I/O. Il successo che ha avuto questa architettura è dovuto essenzialmente alla sua semplicità e versatilità, che ricalca la struttura del computer ideale, la "macchina di Turing", costituito semplicemente da una grande memoria (immaginata come una lunghissima striscia di carta) e da una macchina a stati (immaginata come un carrello programmabile) in grado di leggere e scrivere su ciascuna cella di memoria. La diffusione di questa architettura non deve far pensare che sia l'unica possibile. Molti dispositivi digitali, soprattutto se progettati per specifiche applicazioni, possono essere piuttosto diversi, ad esempio avere più unità logiche e poca memoria o magari nessun bus, ma connessioni dirette punto-punto.

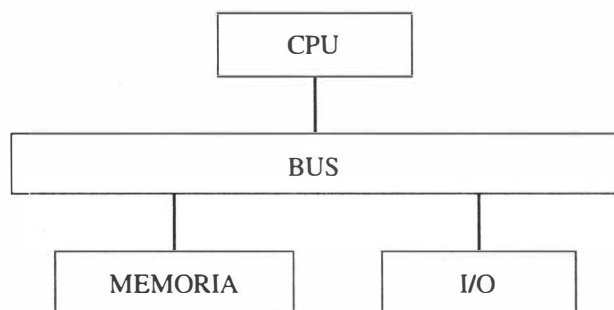


Figura 6.2 Architettura di Von Neumann

6.3 Famiglie logiche

Una **famiglia logica** è un insieme di dispositivi elettronici che possono essere connessi tra loro in modo da realizzare una qualsiasi funzione logica.

Per **funzione logica** si intende una funzione che per ogni possibile configurazione degli stati (alti o bassi) degli ingressi definisce lo stato (alto o basso) di un'uscita.

Una funzione logica con N uscite (come è tipico nei sistemi digitali) formalmente può sempre essere vista come costituita da N funzioni a singola uscita, pilotate dagli stessi ingressi.

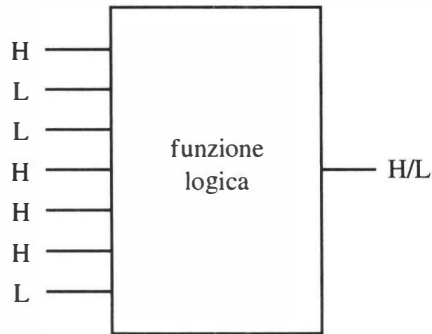


Figura 6.3 Funzione logica

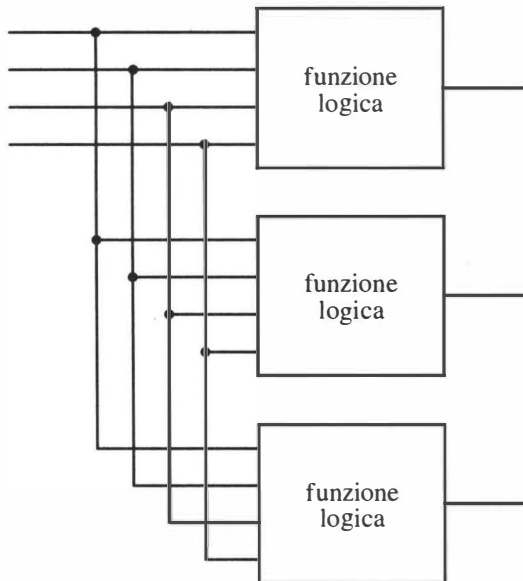
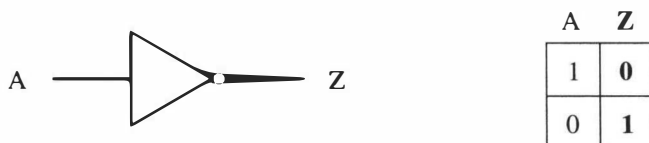


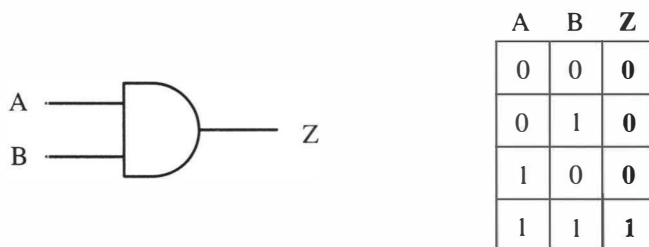
Figura 6.4 Funzione logica a più uscite

Fin dalla metà dell'Ottocento, ben prima di qualunque dispositivo elettronico, il matematico inglese George Boole aveva scoperto come qualunque funzione logica potesse essere realizzata connettendo in maniera opportuna la funzione NOT e una seconda funzione che poteva essere la AND o la OR.

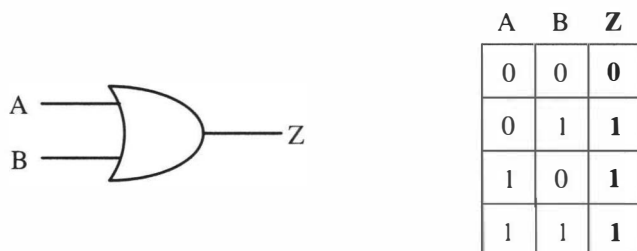
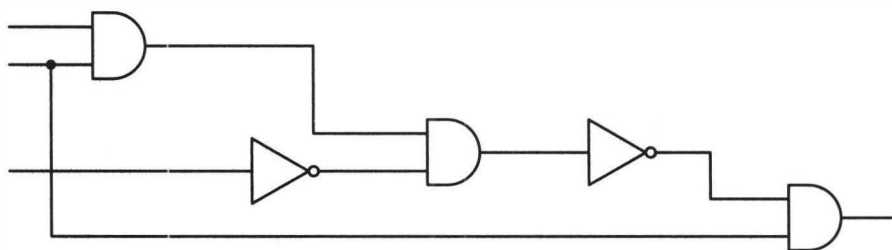
La funzione NOT ha un unico ingresso e un'unica uscita: quando l'ingresso è alto (1) l'uscita è bassa (0) e viceversa. Il simbolo e la tabella di verità sono mostrati in Figura 6.5.

**Figura 6.5** NOT

La funzione AND ha due ingressi e fornisce 1 solo quando entrambi sono 1.

**Figura 6.6** AND

La funzione OR ha due ingressi e fornisce 1 quando almeno uno dei due è 1.

**Figura 6.7** OR**Figura 6.8** Generica funzione logica

In linea di principio una famiglia logica è completa, cioè consente di realizzare qualunque funzione logica, se tra i suoi dispositivi è presente una funzione NOT e una AND o una OR. Al limite è sufficiente realizzare una NAND. Infatti, una NAND seguita da una NOT è equivalente a una AND. Inoltre una NAND è equivalente a una NOT se i due ingressi sono uniti.

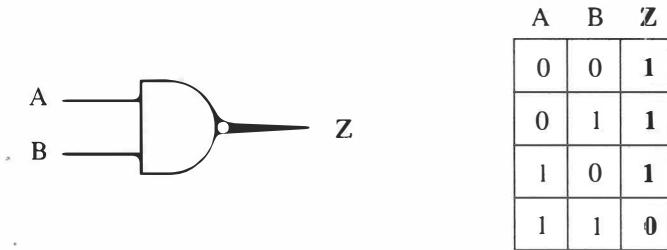


Figura 6.9 NAND

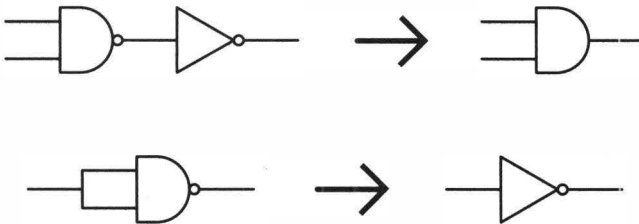


Figura 6.10 NAND

Nella pratica, tuttavia, le famiglie logiche comprendono dispositivi che realizzano, già integrate, una gran varietà di funzioni logiche, non solo le fondamentali.

In generale, una famiglia logica per essere tale deve soddisfare quattro requisiti:

- 1) riconoscere due stati in ingresso e fornire due possibili stati in uscita;
- 2) l'uscita di un dispositivo deve essere compatibile con l'ingresso di un dispositivo della stessa famiglia;
- 3) realizzare qualunque funzione logica;
- 4) rigenerare livelli e fronti.

Il significato dei primi tre requisiti è evidente, il quarto necessita qualche parola di commento.

In una data famiglia logica lo stato alto è codificato con una data tensione, che indichiamo con V_H , lo stato basso con una tensione pari a V_L . Idealmente in

ingresso a un dispositivo logico dovremmo trovare uno di questi due valori, nella realtà, a causa di possibili disturbi, anziché V_H potremmo trovare, ad esempio, $V_H + \Delta V$. Per questo un dispositivo logico dovrà essere realizzato in modo tale che se ΔV è minore di un dato valore (detto “margine di rumore”) il disturbo in uscita dovrà essere di entità minore rispetto al disturbo in ingresso. In altre parole il rumore non si deve propagare: questo si intende per rigenerazione dei livelli. Questa è una caratteristica chiave dell'elettronica digitale ed è anche la ragione per cui nelle telecomunicazioni i segnali digitali sono più robusti dei segnali analogici.

Vediamo ora cosa si intende per rigenerazione dei fronti. Le commutazioni alto-basso di un dispositivo digitale sono controllate da un clock, ovvero da un oscillatore che tipicamente produce un'onda quadra a frequenza costante. Le commutazioni possono avvenire sul fronte di salita o sul fronte di discesa del segnale di clock. Idealmente i fronti dovrebbero essere molto ripidi, nella realtà le capacità parassite delle linee necessitano di tempo per caricarsi e scaricarsi, e quindi i fronti risultano più morbidi. Se il segnale passando da una porta digitale all'altra continuasse a smussarsi, dopo un certo numero di passaggi lo stato dell'uscita risulterebbe indeterminato. Quindi analogamente ai livelli di tensione, i dispositivi digitali devono ricostruire i fronti, ovvero il fronte di salita (o di discesa) in uscita deve mantenersi ripido anche quando il segnale in ingresso è piuttosto smussato.

Dalla comparsa dei primi transistor molte famiglie logiche sono state proposte e sviluppate, alcune hanno avuto vita effimera, altre hanno resistito molti decenni (è il caso della TTL: *Transistor Transistor Logic*, oggi obsoleta). In generale, le prestazioni di famiglie logiche diverse si confrontano sulla base dei seguenti parametri.

1) Margini di rumore

Prendiamo in considerazione una generica porta NOT (la base di qualunque logica). La caratteristica di trasferimento ingresso-uscita dovrebbe essere idealmente una funzione a gradino: alta per valori bassi dell'ingresso e bassa per valori alti dell'ingresso.

Nella realtà la funzione di trasferimento avrà un andamento più smussato come ad esempio in Figura 6.12 da un valore massimo V_{OH} (lo stato alto) a un valore minimo V_{OL} (lo stato basso).

Su tale funzione si definiscono due valori di tensione in ingresso (V_{IL} e V_{IH}) dove la pendenza della funzione è pari a -1 . Un segnale in ingresso maggiore di V_{IH} darà un'uscita prossima a V_{OL} , quindi bassa; un segnale minore di V_{IL} darà un'uscita prossima a V_{OH} , quindi alta. Ma non solo. Se a un segnale minore di V_{IL} si sovrappone un disturbo ΔV , poiché la pendenza della funzione di trasferimento, in valore assoluto, è minore di 1, la variazione in uscita sarà



Figura 6.11 Funzione di trasferimento ideale di una porta NOT

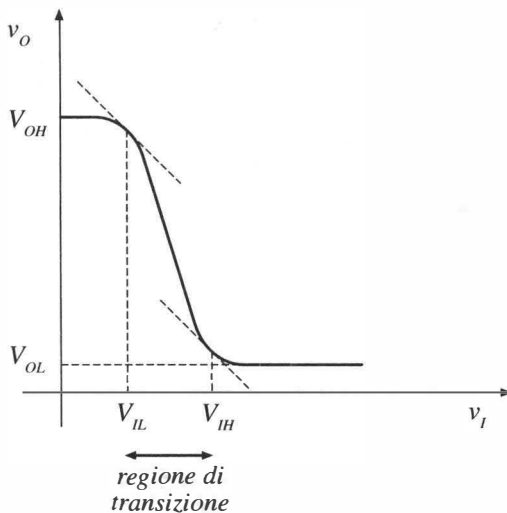


Figura 6.12 Funzione di trasferimento reale di una porta NOT

minore del disturbo in ingresso: in altre parole la porta logica tende a ricostruire i livelli.

Vale lo stesso ragionamento per i segnali maggiori di V_{IH} . In definitiva possiamo definire margine di rumore alto (NM_H) e margine di rumore basso (NM_L) come segue:

$$NM_H = V_{OH} - V_{IH} \quad (6.1)$$

$$NM_L = V_{IL} - V_{OL} \quad (6.2)$$

Più grandi sono i margini di rumore più la famiglia logica è in grado di ricostruire i livelli anche in presenza di disturbi.

Per quanto riguarda la ricostruzione dei fronti, il parametro importante è la regione di transizione tra V_{IL} e V_{IH} . Più stretta è la regione di transizione più ripidi saranno i fronti ricostruiti. In ultima analisi i quattro parametri V_{OH} , V_{OL} , V_{IL} e V_{IH} caratterizzano completamente le prestazioni della famiglia logica in termini di ricostruzione di livelli e fronti.

2) Tempo di propagazione

Supponiamo di dare in ingresso a una porta NOT un fronte di discesa molto ripido (vedi Figura 6.13). Il tempo di propagazione del segnale è l'intervallo di tempo tra il fronte in ingresso e l'istante al quale l'uscita è salita fino al 50% del livello alto. I tempi di propagazione sono due: t_{PLH} quando l'uscita va da basso a alto e t_{PHL} quando l'uscita va da alto a basso (vedi Figura 6.14).

3) Tempi di salita e di discesa

Supponiamo di dare in ingresso a una porta NOT un fronte di discesa molto ripido (vedi Figura 6.13). Il tempo di salita del segnale è l'intervallo di tempo per cui l'uscita passa dal 10% al 90% del livello alto. Il tempo di discesa è, invece, l'intervallo di tempo per cui l'uscita passa dal 90% al 10% del livello alto, quando in ingresso è presente un fronte di salita molto ripido (vedi Figura 6.14).

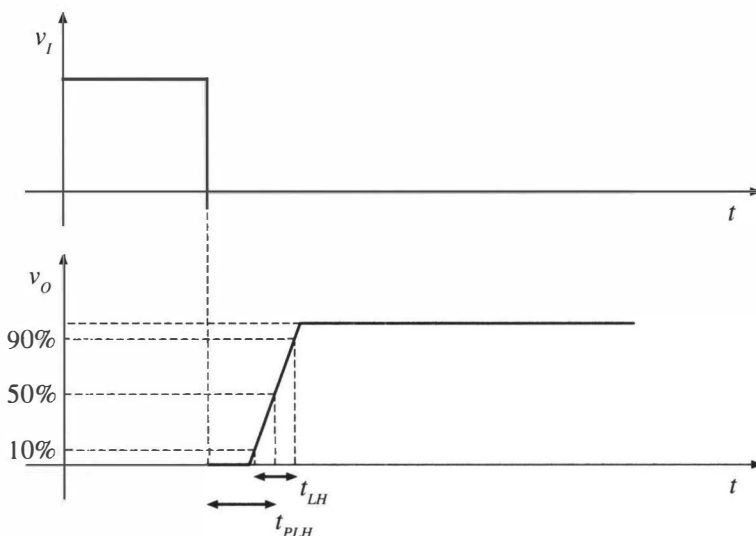


Figura 6.13 Commutazione basso-alto

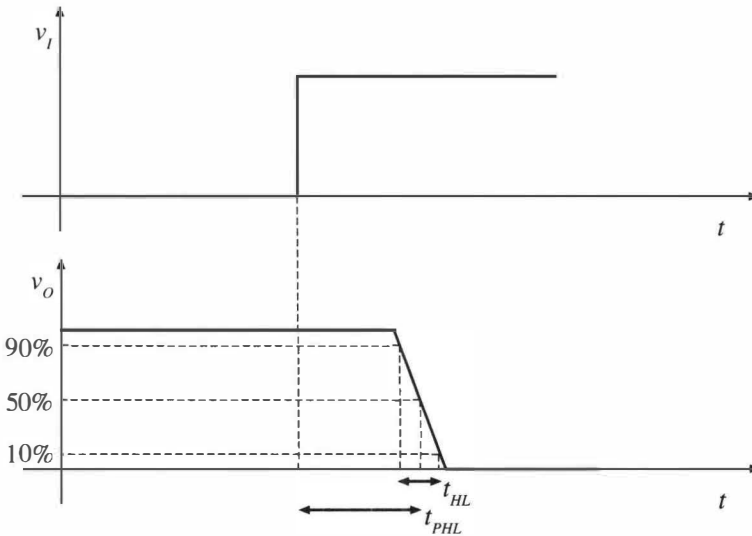


Figura 6.14 Commutazione alto-basso

4) Consumo statico

Il consumo statico è il valore medio dei consumi in condizione statica di una porta quando l'uscita è alta (P_{STAT-H}) e quando l'uscita è bassa (P_{STAT-L})

$$P_{STAT} = \frac{P_{STAT-H} + P_{STAT-L}}{2} \quad (6.3)$$

5) Consumo dinamico

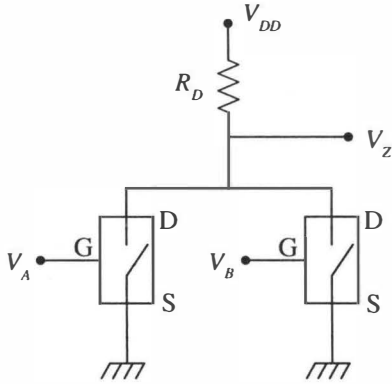
In un ciclo di commutazione alto-basso-alto, una porta logica dissipa una certa quantità di energia E (in Joule). Se la porta è pilotata con clock a frequenza f_{CLOCK} , avrà un consumo dinamico pari a

$$P_{DYN} = f_{CLOCK} E \quad (6.4)$$

6.4 Interruttore ideale e resistore

A scopo esemplificativo, prendiamo in considerazione una famiglia logica ideale costituita da una resistenza e da un interruttore ideale controllato in tensione disposti come in Figura 6.15.

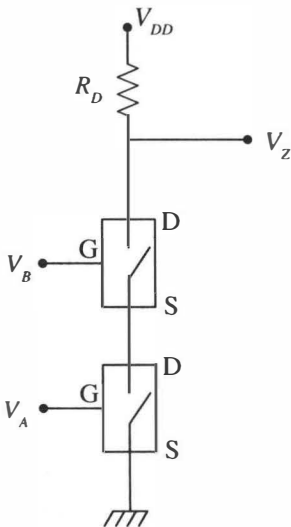
ingresso è maggiore della soglia, il relativo interruttore è in conduzione e forza l'uscita a massa. L'uscita può rimanere alla tensione di alimentazione solo se entrambi gli interruttori sono OFF.



A	B	Z
0	0	1
0	1	0
1	0	0
1	1	0

Figura 6.16 NOR

Una NAND può essere realizzata come in Figura 6.17. Anche in questo caso le tensioni sono riferite a massa. Se la tensione di ingresso dell'interruttore collegato a massa è minore della soglia, l'interruttore è OFF e non può scorrere corrente nella maglia, quindi l'uscita è alta indipendentemente dallo stato



A	B	Z
0	0	1
0	1	1
1	0	1
1	1	0

Figura 6.17 NAND

dell'altro interruttore. Vediamo ora il caso in cui la tensione di ingresso (riferita a massa) dell'interruttore non collegato a massa sia minore della soglia. Qualunque sia lo stato dell'altro interruttore, la tensione tra G e S sarà comunque minore della soglia. Quindi nella maglia non scorre corrente e la tensione di uscita sarà alta.

In generale una qualunque funzione logica può essere realizzata combinando variamente interruttori in serie e in parallelo (vedi Figura 6.18).

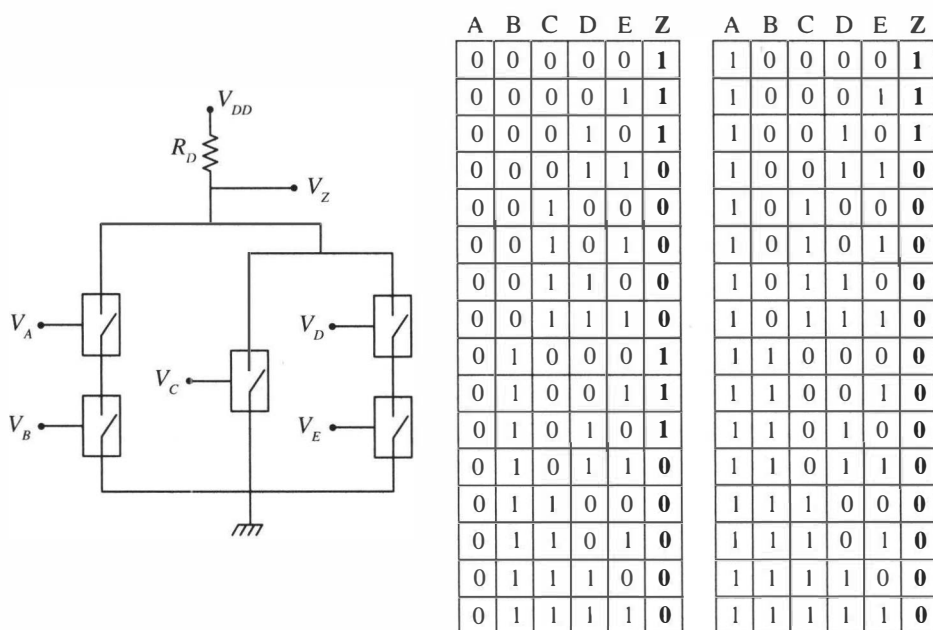


Figura 6.18 Generica funzione logica

4) Rigenerare livelli e fronti

Se l'interruttore scatta repentinamente alla soglia, i livelli possibili in uscita possono essere solo alto o basso, quindi i livelli sono ricostruiti.

Riguardo alla ricostruzione dei fronti, prendiamo in considerazione il caso di un segnale in ingresso a una porta NOT che non sia un fronte ripido di salita ma piuttosto la lenta carica di un capacitore (vedi Figura 6.19), se l'interruttore è ideale scatterà solo quando il segnale in ingresso raggiunge la soglia, e quindi il segnale in uscita avrà un fronte ripido (compatibilmente con un'eventuale capacità di carico all'uscita) del tutto indipendente dalla forma del segnale in ingresso.

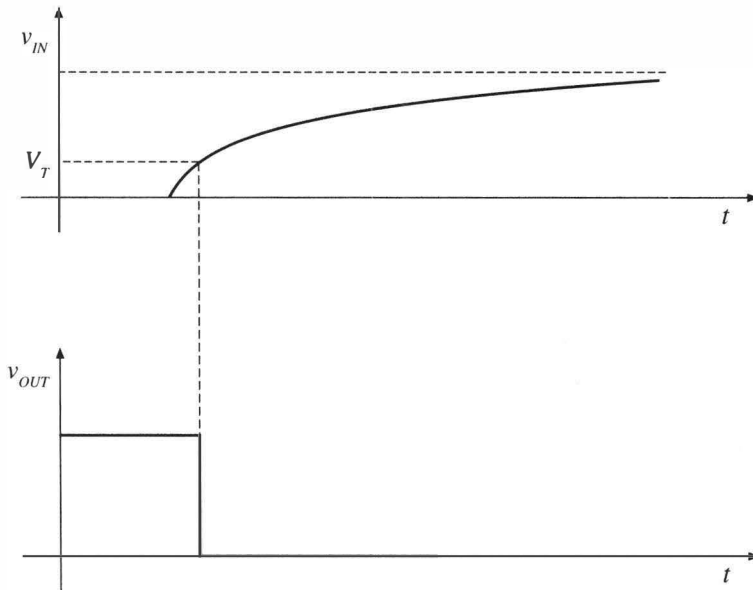


Figura 6.19 Rigenerazione dei fronti

Vediamo ora i parametri relativi alle prestazioni.

1) Margini di rumore

Poiché l'interruttore è sempre OFF per tensione di soglia minore dell'ingresso, il margine di rumore per il livello basso (NM_L) è pari alla tensione di soglia, mentre il margine di rumore per il livello alto (NM_H) è pari a $V_{DD} - V_T$.

2) Tempo di propagazione

Il tempo di propagazione da alto a basso è nullo ($t_{PHL} = 0$), mentre il tempo di propagazione da basso ad alto è determinato dal tempo di carica della capacità parassita (C) dello stadio successivo, ovvero dalla capacità tra G e S (C_{GS}). Se più ingressi sono collegati alla stessa uscita (il termine inglese è *fan-out*), la capacità che vede l'uscita è

$$C = NC_{GS} \quad (6.5)$$

con N numero di ingressi collegati alla stessa uscita. Il tempo di propagazione aumenta di conseguenza.

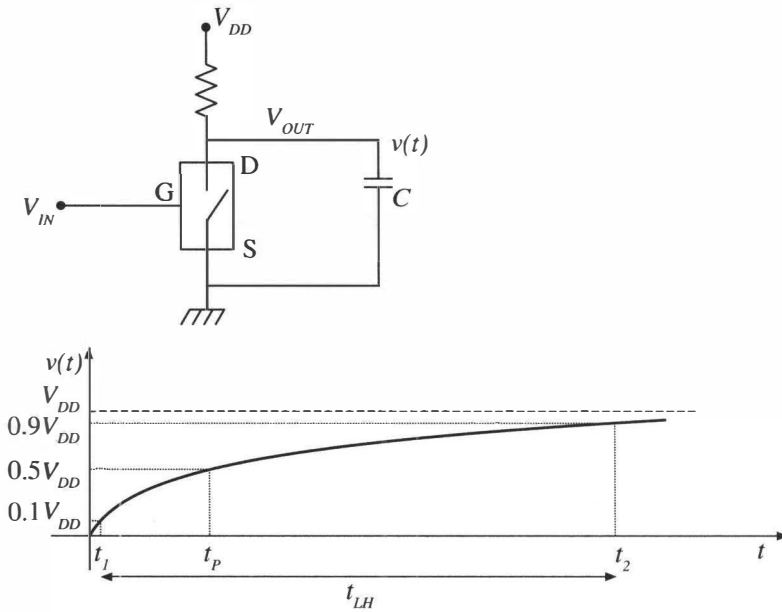


Figura 6.20 Commutazione dell'uscita da basso a alto

La tensione del capacitore nella fase di carica è data da:

$$v(t) = V_{DD} \left(1 - e^{-\frac{t}{RC}} \right) \quad (6.6)$$

Il tempo di propagazione da basso ad alto può, quindi, essere calcolato come segue:

$$\frac{1}{2} V_{DD} = V_{DD} \left(1 - e^{-\frac{t_{PLH}}{RC}} \right) \quad (6.7)$$

$$t_{PLH} = RC \ln(2) = 0.69RC \quad (6.8)$$

3) Tempi di salita e di discesa

Il tempo di discesa è nullo ($t_{HL} = 0$). Il tempo di salita può essere calcolato dalla (6.6) come $t_2 - t_l$ in Figura 6.20. Pertanto t_2 si calcola come segue:

$$\frac{9}{10} V_{DD} = V_{DD} \left(1 - e^{-\frac{t_2}{RC}} \right) \quad (6.9)$$

$$t_2 = RC \ln(10) = 2.3RC \quad (6.10)$$

Invece t_l si calcola come segue:

$$\frac{1}{10} V_{DD} = V_{DD} \left(1 - e^{-\frac{t_f}{RC}} \right) \quad (6.11)$$

$$t_f = RC \ln \left(\frac{10}{9} \right) = 0.1 RC \quad (6.12)$$

Quindi $t_{LH} = 2.2 RC$.

I tempi di salita e di discesa limitano la velocità massima alla quale può essere pilotata la famiglia logica in esame. Tali tempi dipendono dalla capacità parassita e dalla resistenza: minore è la resistenza, più veloce è la logica.

4) Consumo statico

In una porta NOT quando l'uscita è alta la corrente che scorre nella resistenza è nulla e quindi

$$P_{STAT-H} = 0 \quad (6.13)$$

Quando l'uscita è bassa la corrente che scorre sulla resistenza è V_{DD}/R e quindi la potenza dissipata è pari a

$$P_{STAT-L} = \frac{V_{DD}^2}{R} \quad (6.14)$$

La dissipazione statica (media) è

$$P_{STAT} = \frac{V_{DD}^2}{2R} \quad (6.15)$$

Minore è la tensione di alimentazione minore è il consumo, tuttavia minore è la tensione di alimentazione minori sono i margini di rumore e quindi la sensibilità ai disturbi.

A parità di tensione di alimentazione, maggiore è la resistenza minore è il consumo. Purtroppo maggiore è la resistenza maggiore è il tempo di commutazione.

5) Consumo dinamico

Oltre alla dissipazione statica è necessario considerare l'energia dissipata a ogni colpo di clock, che è pari all'energia che l'alimentazione deve fornire per caricare la capacità (parassita) dello stadio successivo. Poiché la carica Q e la tensione v in un condensatore sono legate dalla relazione

$$Q = Cv \quad (6.16)$$

segue che

$$i = C \frac{dv}{dt} \quad (6.17)$$

Per calcolare l'energia che l'alimentazione fornisce si integra la corrente nel tempo:

$$E = \int_{t=0}^{t=\infty} V_{DD} i dt = V_{DD} C \int_{t=0}^{t=\infty} \frac{dv}{dt} dt = V_{DD} C \int_{v=0}^{v=V_{DD}} dv = CV_{DD}^2 \quad (6.18)$$

Può essere interessante calcolare anche l'energia accumulata nel capacitore quando è portato alla tensione V_{DD} .

$$E_C = \int_{q=0}^{q=Q} v dq = \frac{1}{C} \int_{q=0}^{q=Q} q dq = \frac{1}{2C} Q^2 = \frac{1}{2} CV_{DD}^2 \quad (6.19)$$

Si noti che per caricare il capacitore l'alimentazione deve fornire un'energia pari a CV_{DD}^2 , ma solo la metà è effettivamente immagazzinata nel capacitore, l'altra metà è dissipata nella resistenza R . Quando poi l'uscita passa da alta a bassa, l'energia del capacitore si scarica a massa dissipandosi nella resistenza interna dell'interruttore (l'ipotesi di interruttore ideale con resistenza nulla, evidentemente, non è ragionevole per questo calcolo).

Dalla (6.4) segue che la dissipazione dinamica è pari a

$$P_{dyn} = CV_{DD}^2 f_{CLOCK} \quad (6.20)$$

con f_{CLOCK} frequenza del clock.

La soluzione interruttore e resistenza ha il pregio della semplicità, ma ha un problema critico: la dissipazione statica.

La dissipazione statica è un problema non tanto per il consumo di potenza in sé, che in fin dei conti non è una questione critica per computer alimentati dalla rete elettrica, ma per il fatto che limita il numero di porte logiche per unità di superficie sul chip. Infatti, il calore per unità di superficie che il silicio può smaltire ha un limite nella conducibilità termica del silicio. Già alla fine degli anni ottanta risultò evidente che la dissipazione statica era divenuta il principale problema per l'ulteriore miniaturizzazione delle porte logiche. La dissipazione dinamica, al contrario di quella statica, non costituisce un problema per la miniaturizzazione. La potenza dissipata è, infatti, proporzionale alla capacità parassita dell'ingresso della porta logica. Riducendo le dimensioni della porta, si riduce proporzionalmente la capacità, pertanto la potenza per unità di superficie rimane costante.

La ragione per cui la famiglia logica oggi più diffusa (la CMOS) non si basa sulla configurazione interruttore e resistenza, ma si basa su una diversa configurazione, detta a **interruttori complementari**, è proprio perché quest'ultima, come vedremo nel prossimo paragrafo, ha il fondamentale vantaggio di non presentare dissipazione statica.

6.5 Interruttori complementari ideali

Lo schema di una porta NOT a interruttori complementari è mostrata Figura 6.21. Entrambi gli interruttori sono ideali e sono pilotati dalla tensione tra G e S, ma l'interruttore indicato con n è ON quando $v_{GS} > V_{TC}$ (con V_{TC} valore positivo) mentre l'interruttore indicato con p è ON quando $v_{GS} < -V_{TC}$, ovvero $v_{SG} > V_{TC}$. Sebbene ideali, non trascureremo la capacità parassite tra G e S di ciascun interruttore.

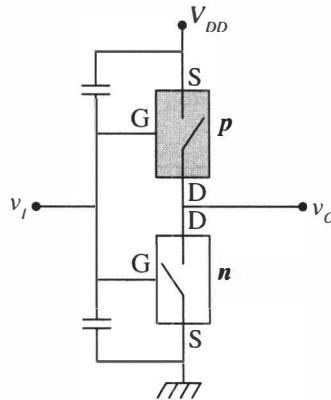


Figura 6.21 Porta NOT a interruttori complementari ideali

Per semplificare l'analisi supponiamo $V_{TC} = V_{DD}/2$. In tal caso per $v_I < V_{TC}$ l'interruttore n è OFF e l'interruttore p è ON: l'uscita è quindi cortocircuitata alla tensione di alimentazione. Per $v_I > V_{TC}$ l'interruttore n è ON e l'interruttore p è OFF: l'uscita è quindi cortocircuitata a massa.

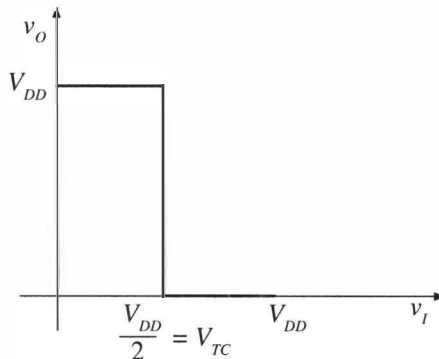


Figura 6.22 Funzione di trasferimento

È facile verificare i requisiti 1, 2, e 4 delle famiglie logiche. Il requisito 3 (realizzare qualunque funzione logica) richiede qualche precisazione. Le Figure 6.23 e 6.24 mostrano le porte NAND e NOR. In generale è agevole sintetizzare direttamente qualsiasi funzione logica, avendo cura che le due reti (una tra alimentazione e uscita e l'altra tra uscita e massa) siano complementari, nel senso che quando una è un circuito aperto l'altra è un cortocircuito e viceversa.

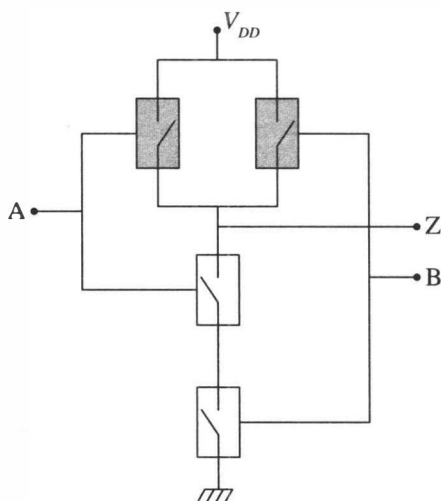


Figura 6.23 NAND

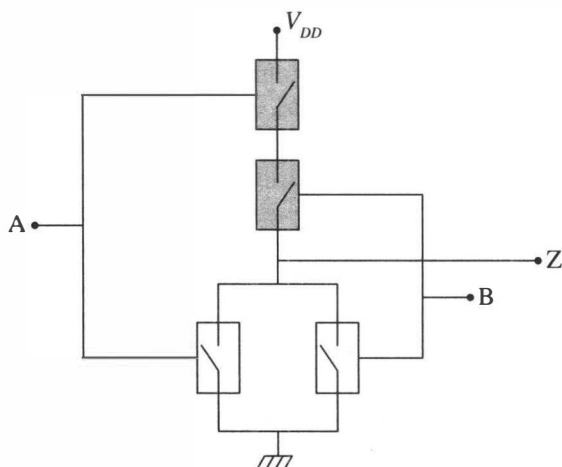


Figura 6.24 NOR

La configurazione a interruttori complementari ha due importanti vantaggi:

1) *Il consumo statico è sempre nullo*, in quanto per qualunque valore di tensione in ingresso non scorre corrente nella porta, essendo uno dei due interruttori OFF.

2) *I tempi di commutazione (salita, discesa e propagazione) sono idealmente nulli*. Infatti, se anche consideriamo in uscita un capacitore che rappresenta la capacità parassita dello stadio successivo, i tempi di carica e scarica sono nulli se la resistenza di ON dei due interruttori è nulla.

Si noti, tuttavia, che sebbene la dissipazione statica e i tempi di commutazione siano nulli, non lo è il consumo dinamico. Se infatti l'uscita della porta è collegata all'ingresso di un'altra porta le due capacità parassite degli interruttori devono caricarsi e scaricarsi durante la commutazione.

Calcoliamo, quindi, il consumo dinamico tenendo conto delle capacità parassite.

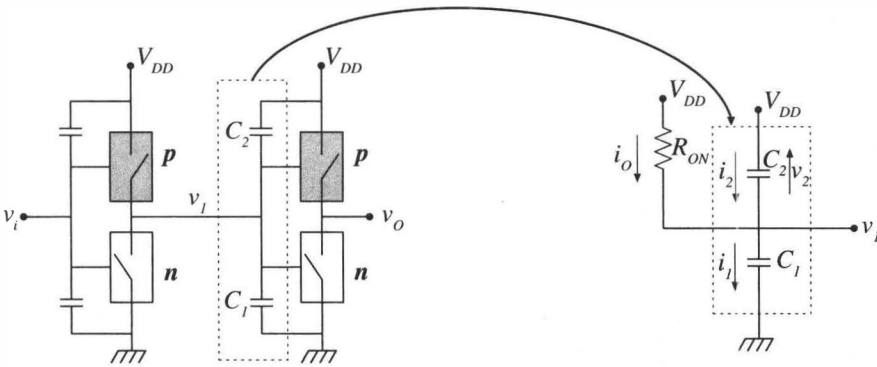


Figura 6.25 Capacità parassite

Con riferimento alla Figura 6.25, prendiamo in considerazione la variazione di carica delle due capacità parassite della seconda porta, mediante la corrente che fornisce la prima porta. Questa corrente scorre dall'alimentazione attraverso l'interruttore p . Chiamiamo R_{ON} la resistenza dell'interruttore quando è chiuso.

$$\frac{dv_1}{dt} = \frac{1}{C_1} i_1 \quad (6.21)$$

$$\frac{dv_2}{dt} = \frac{1}{C_2} i_2 \quad (6.22)$$

Poiché

$$\frac{dv_1}{dt} = -\frac{dv_2}{dt} \quad (6.23)$$

segue che

$$\frac{i_1}{i_2} = -\frac{C_1}{C_2} \quad (6.24)$$

Poiché

$$i_O = i_1 - i_2 \quad (6.25)$$

segue che

$$i_O = \frac{C_1 + C_2}{C_1} i_1 \quad (6.26)$$

Sostituendo la (6.26) nella (6.21):

$$\frac{dv_1}{dt} = \frac{1}{C_1 + C_2} i_O \quad (6.27)$$

il cui significato fisico è che la tensione in uscita v_1 aumenta come se il carico fosse costituito da due condensatori in parallelo tra uscita e massa. Possiamo quindi concludere che l'energia dissipata per ogni carica e scarica è pari a

$$E = (C_1 + C_2) V_{DD}^2 \quad (6.28)$$

e quindi

$$P_{dyn} = (C_1 + C_2) V_{DD}^2 f_{CLOCK} \quad (6.29)$$

La famiglia logica descritta in questo paragrafo è puramente ideale (in quanto basata su interruttori ideali), tuttavia su tale schema di principio si basa la logica oggi più largamente diffusa che verrà descritta nel paragrafo che segue.

6.6 La logica CMOS

Una porta NOT CMOS (*Complementary MOS*) è essenzialmente costituita da un MOSFET a canale n e un MOSFET a canale p disposti come in Figura 6.26.

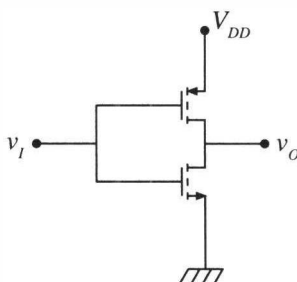


Figura 6.26 Porta NOT CMOS

Tracciamo ora la caratteristica ingresso-uscita della porta CMOS. I due MOSFET hanno stesso valore di soglia (V_{TC}), inoltre $V_{DD} > 2V_{TC}$. Quando $v_I < V_{TC}$ il MOSFET n è interdetto.

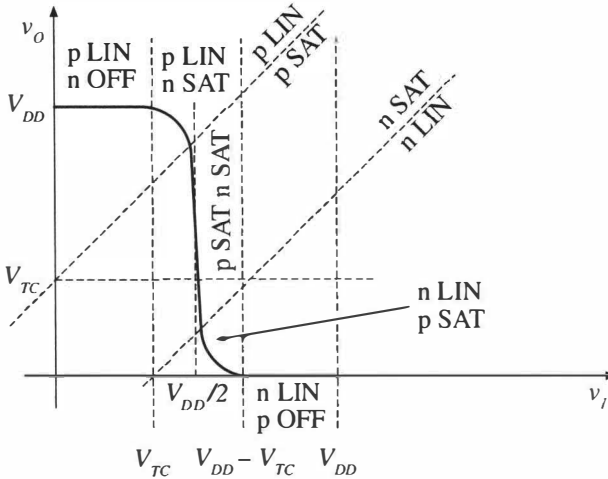


Figura 6.27 Funzione di trasferimento

La condizione affinché il MOSFET p sia ON è

$$v_{SG}^{(p)} > V_{TC} \quad (6.30)$$

$$V_{DD} - v_I > V_{TC} \quad (6.31)$$

$$v_I < V_{DD} - V_{TC} \quad (6.32)$$

Poiché $V_{DD} > 2V_{TC}$ segue che quando $v_I < V_{TC}$ il MOSFET p è ON.

Resta da verificare se il MOSFET p si trova in regione lineare o in saturazione.

La condizione per la saturazione è

$$v_{SD}^{(p)} > v_{SG}^{(p)} - V_{TC} \quad (6.33)$$

$$V_{DD} - v_O > V_{DD} - v_I - V_{TC} \quad (6.34)$$

$$v_O < v_I + V_{TC} \quad (6.35)$$

Poiché il MOSFET n è OFF e il carico in uscita è supposto una resistenza molto grande, la corrente che scorre sul MOSFET p è nulla e quindi la tensione di uscita è pari alla V_{DD} . Essendo $V_{DD} > 2V_{TC}$, la (6.35) non è verificata e quindi il MOSFET p è in regione lineare.

Quando v_I supera la soglia V_{TC} , il MOSFET n entra in conduzione, scorre corrente nel MOSFET p e quindi la tensione di uscita comincia a scendere.

Per verificare se il MOSFET n è in regione lineare o in saturazione, dobbiamo scrivere la sua condizione per la saturazione:

$$v_{DS}^{(n)} > v_{GS}^{(n)} - V_{TC} \quad (6.36)$$

$$v_O > v_I - V_{TC} \quad (6.37)$$

Quindi per v_I appena superiore alla soglia il MOSFET n è in saturazione, mentre il MOSFET p è in regione lineare. Al diminuire della tensione di uscita anche il MOSFET p entra in saturazione e si avrà quindi una zona in cui entrambi i MOSFET sono in saturazione. Aumentando ancora la tensione in ingresso, il MOSFET n uscirà dalla saturazione per entrare in regione lineare e, infine, per $v_I > V_{DD} - V_{TC}$ il MOSFET p sarà OFF e il MOSFET n in regione lineare.

Calcoliamo ora il valore di v_I per cui entrambi i MOSFET sono in saturazione.

$$i_{DS}^{(n)} = i_{SD}^{(p)} \quad (6.38)$$

$$\frac{K}{2} (v_I - V_{TC})^2 = \frac{K}{2} (V_{DD} - v_I - V_{TC})^2 \quad (6.39)$$

da cui

$$v_I = \frac{V_{DD}}{2} \quad (6.40)$$

Vediamo ora i parametri relativi alle prestazioni.

1) Margini di rumore

Per calcolare V_{IL} è necessario trovare il punto per cui la pendenza della funzione di trasferimento è pari a -1 . Tale punto si trova nella regione in cui il MOSFET n è in saturazione e il MOSFET p è in regione lineare. Poiché le correnti dei due MOSFET devono essere uguali (nell'ipotesi che la resistenza associata al carico applicata in uscita sia molto grande), possiamo scrivere:

$$\frac{K}{2} (v_{GS}^{(n)} - V_{TC})^2 = K \left((v_{SG}^{(p)} - V_{TC}) v_{SD}^{(p)} - \frac{1}{2} (v_{SD}^{(p)})^2 \right) \quad (6.41)$$

$$(v_I - V_{TC})^2 = 2(V_{DD} - v_I - V_{TC})(V_{DD} - v_O) - (V_{DD} - v_O)^2 \quad (6.42)$$

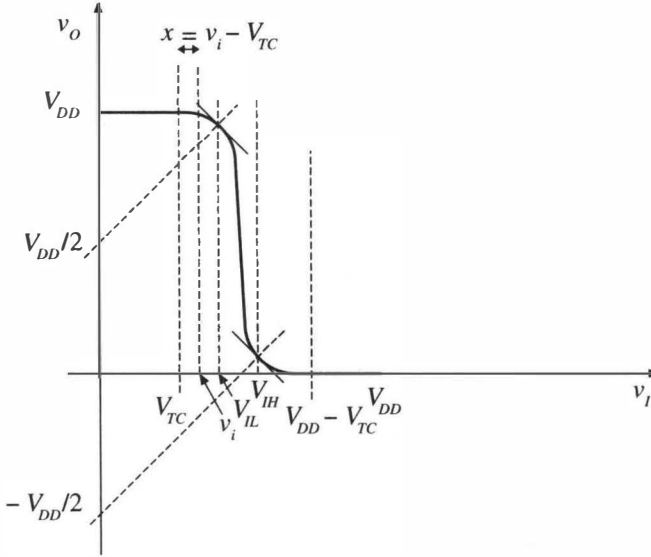


Figura 6.28 Calcolo dei margini di rumore

Calcoliamo la derivata rispetto a v_i :

$$\begin{aligned} 2(v_I - V_{TC}) &= -2(V_{DD} - v_O) + 2(V_{DD} - v_I - V_{TC}) \left(-\frac{dv_O}{dv_i} \right) + \\ &-2(V_{DD} - v_O) \left(-\frac{dv_O}{dv_i} \right) \end{aligned} \quad (6.43)$$

Ponendo nella (6.43)

$$\frac{dv_O}{dv_i} = -1 \quad (6.44)$$

si ottiene

$$v_O = v_I + \frac{V_{DD}}{2} \quad (6.45)$$

che nel piano $v_I - v_O$ è una retta di pendenza unitaria passante per il punto $(0, V_{DD}/2)$.

Per calcolare $v_I = V_{IL}$ si sostituisce la condizione (6.45) nella (6.42). Per comodità di calcolo conviene usare come incognita

$$x = v_I - V_{TC} \quad (6.46)$$

Dopo qualche semplice passaggio algebrico si ottiene

$$x = \frac{3}{4} \left(\frac{V_{DD}}{2} - V_{TC} \right) \quad (6.47)$$

V_{IL} è quindi dato da

$$V_{IL} = V_{TC} + x = \frac{3}{8} V_{DD} + \frac{1}{4} V_{TC} \quad (6.48)$$

Per la simmetria della funzione di trasferimento, V_{IH} può essere calcolato come

$$V_{IH} = V_{DD} - V_{TC} - x = \frac{5}{8} V_{DD} - \frac{1}{4} V_{TC} \quad (6.49)$$

La regione di indeterminazione è pari a

$$V_{IH} - V_{IL} = \frac{V_{DD} - 2V_{TC}}{4} \quad (6.50)$$

Proviamo a dare dei valori numerici. Per $V_{DD} = 5V$ e $V_{TC} = 1V$, risulta $V_{IL} = 2.1V$ e $V_{IH} = 2.9V$. La regione di indeterminazione è pari a $0.8V$.

2) Tempo di propagazione

Per il calcolo del tempo di propagazione, cominciamo con osservare che, per la simmetria della porta, $t_{PLH} = t_{PHL}$. Calcoliamo quindi t_{PHL} . Supponendo il fronte in ingresso molto ripido, quando l'ingresso passa da $v_I = 0$ a $v_I = V_{DD}$ il MOSFET p si interdice e la capacità (parassita) C si scarica attraverso il MOSFET n, il quale fino al tempo t_f sarà in saturazione poi passerà in regione lineare.

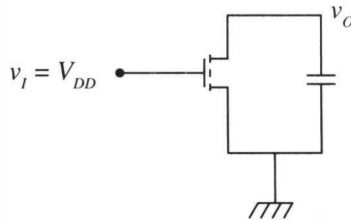


Figura 6.29 Schema circuitale relativo alla commutazione dell'uscita alto-basso

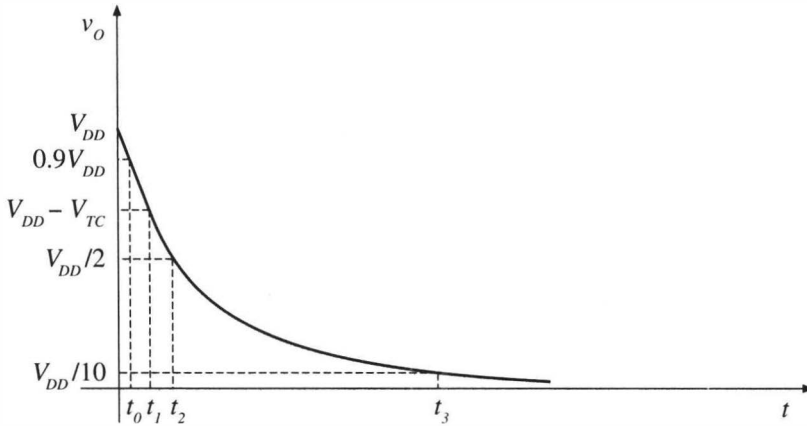


Figura 6.30 Commutazione alto-basso

Calcoliamo il tempo t_1 . La corrente i_{DS} che tira il MOSFET in saturazione è data da

$$i_{DS} = \frac{K}{2} (V_{DD} - V_{TC})^2 \quad (6.51)$$

In generale, per la corrente che fluisce da un capacitore vale

$$i = -C \frac{dv_O}{dt} \quad (6.52)$$

Quindi, il capacitore si scarica linearmente fino a che il MOSFET non entra in regione lineare per $v_{DS} = v_I - V_{TC}$, ovvero per $v_O = V_{DD} - V_{TC}$. Pertanto t_1 è dato da:

$$t_1 = \frac{2C}{K} \frac{V_{TC}}{(V_{DD} - V_{TC})^2} \quad (6.53)$$

Calcoliamo ora $\Delta t = t_2 - t_1$. In regione lineare deve valere

$$K \left((v_{GS} - V_{TC}) v_{DS} - \frac{1}{2} v_{DS}^2 \right) = -C \frac{dv_O}{dt} \quad (6.54)$$

$$K \left(2(V_{DD} - V_{TC}) v_O - v_O^2 \right) = -2C \frac{dv_O}{dt} \quad (6.55)$$

$$\frac{K}{2C} \int_{t_1}^{t_2} dt = - \int_{V_{DD}-V_{TC}}^{\frac{V_{DD}}{2}} \frac{dv_O}{2(V_{DD} - V_{TC}) v_O - v_O^2} \quad (6.56)$$

D'altra parte

$$\int \frac{dx}{(a-x)x} = -\frac{1}{a} \ln\left(\frac{x-a}{x}\right) \quad (6.57)$$

Quindi

$$\Delta t = t_2 - t_1 = \frac{C}{K} \frac{1}{(V_{DD} - V_{TC})} \ln\left(\frac{3V_{DD} - 4V_{TC}}{V_{DD}}\right) \quad (6.58)$$

Il tempo di propagazione è pari a

$$t_{PHL} = t_1 + \Delta t \quad (6.59)$$

Proviamo a quantificare i risultati ottenuti. Consideriamo il caso realistico $K = 2\mu A/V^2$, $V_{DD} = 5V$, $V_{TC} = 1V$, $C = 0.1pF$. Si ottiene $t_1 = 6.3ns$ e $\Delta t = 9.9ns$. Il tempo di propagazione è quindi $16.2ns$.

È importante notare che il tempo di propagazione (analogamente ai tempi di salita e di discesa, come vedremo nel paragrafo successivo) dipende linearmente dal rapporto C/K .

$$t_p \propto \frac{C}{K} \quad (6.60)$$

D'altra parte la capacità di carico di una porta logica è, come minimo, la capacità parassita all'ingresso della porta successiva. Quindi nel caso della logica CMOS

$$C = C_n + C_p = C_{ox}'' W_n L_n + C_{ox}'' W_p L_p \quad (6.61)$$

D'altra parte

$$K = \frac{C_{ox}'' \mu_n W_n}{L_n} = \frac{C_{ox}'' \mu_p W_p}{L_p} \quad (6.62)$$

da cui segue che

$$\frac{C}{K} = \frac{L_n^2}{\mu_n} + \frac{L_p^2}{\mu_p} \quad (6.63)$$

Si tratta di un risultato molto importante: poiché il tempo di propagazione è proporzionale a C/K , segue che dipende solo dal quadrato della lunghezza dei

canali: più sono corti più la porta è veloce. Aumentare la larghezza del canale non ha effetto sulla velocità di una logica.

3) Tempi di salita e di discesa

Per la simmetria della porta, $t_{HL} = t_{LH}$. Con riferimento alla Figura 6.30, il tempo di discesa può essere calcolato come

$$t_{HL} = t_I + (t_3 - t_I) - t_0 \quad (6.64)$$

$(t_3 - t_I)$ si calcola mediante l'integrale

$$\frac{K}{2C} \int_{t_I}^{t_3} dt = - \int_{V_{DD}-V_{TC}}^{\frac{V_{DD}}{10}} \frac{dv_O}{2(V_{DD} - V_{TC})v_O - v_O^2} \quad (6.65)$$

da cui

$$t_3 - t_I = \frac{C}{K} \frac{1}{(V_{DD} - V_{TC})} \ln \left(\frac{19V_{DD} - 20V_{TC}}{V_{DD}} \right) \quad (6.66)$$

Se $V_{TC} > V_{DD}/10$, t_0 è nella zona in cui la curva ha un andamento lineare e si calcola come

$$t_0 = \frac{2C}{K} \frac{V_{DD}}{10(V_{DD} - V_{TC})^2} \quad (6.67)$$

Quantificando come per il calcolo del tempo di propagazione, si ottiene $(t_3 - t_I) = 34\text{ns}$. Quindi $t_{HL} = 37\text{ns}$.

4) Consumo dinamico

Il consumo dinamico, dovuto all'energia necessaria per caricare e scaricare la capacità parassita $C = C_n + C_p$, è dato da

$$P_{dyn} = (C_n + C_p) V_{DD}^2 f_{CLOCK} \quad (6.68)$$

Tuttavia a tale valore va sommato un altro contributo dovuto al fatto che il consumo della porta (anche in condizioni statiche) non è nullo quando l'uscita è in uno stato intermedio tra alta e bassa. Infatti quando

$$V_{TC} < v_I < V_{DD} - V_{TC} \quad (6.69)$$

entrambi i MOSFET sono in conduzione, scorre una corrente i e pertanto l'alimentazione deve fornire una potenza istantanea

$$P = V_{DD}i \quad (6.70)$$

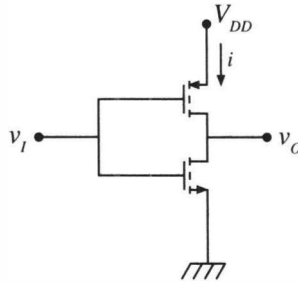


Figura 6.31 Corrente erogata dall'alimentazione

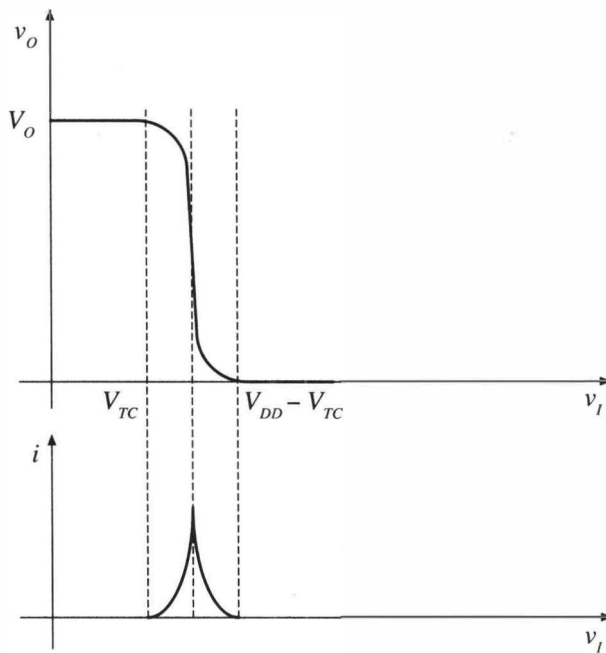


Figura 6.32 Funzione di trasferimento e corrente

La corrente è massima quando entrambi i MOSFET sono in saturazione. In tale condizione è pari a

$$I_{max} = \frac{K}{2} \left(\frac{V_{DD}}{2} - V_{TC} \right)^2 \quad (6.71)$$

Fortunatamente la porta ha questo consumo solo durante le commutazioni. Per una stima quantitativa della dissipazione dovuta a questa corrente, si consideri il caso di due porte in cascata. L'ingresso della prima è un fronte di discesa molto ripido. L'uscita della prima porta è collegata all'ingresso della seconda e la capacità di ingresso di quest'ultima è l'elemento limitante la velocità di commutazione della prima porta. All'uscita della seconda porta non c'è capacità di carico. Vogliamo stimare la dissipazione dinamica della seconda porta.

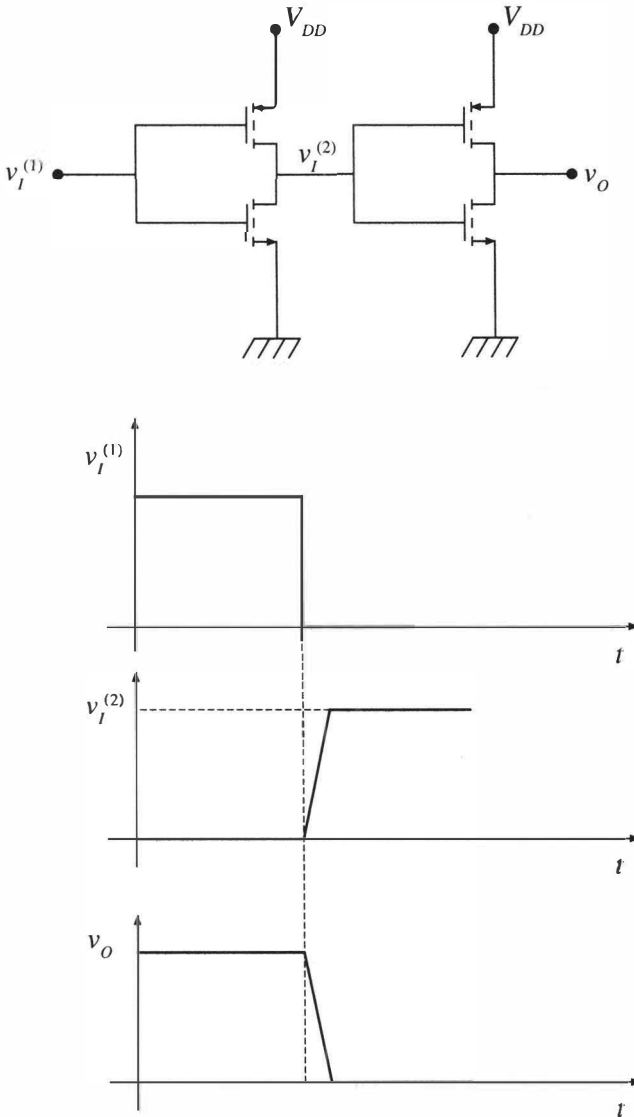


Figura 6.33 Calcolo della dissipazione dovuta alla corrente che scorre tra i MOSFET in conduzione

Se, molto grossolanamente, approssimiamo lineare la discesa della tensione e il picco di corrente nel tempo come un triangolo, l'energia dissipata per ogni commutazione è circa pari a

$$E \approx V_{DD} \frac{i_{DS}^{(max)} t_{HL}}{2} = V_{DD} \frac{K}{4} \left(\frac{V_{DD}}{2} - V_{TC} \right)^2 t_{HL} \quad (6.72)$$

In un periodo del clock ci sono due commutazioni, quindi la potenza dissipata è pari a

$$P = f_{CLOCK} V_{DD} \frac{K}{2} \left(\frac{V_{DD}}{2} - V_{TC} \right)^2 t_{HL} \quad (6.73)$$

con f_{CLOCK} frequenza del clock.

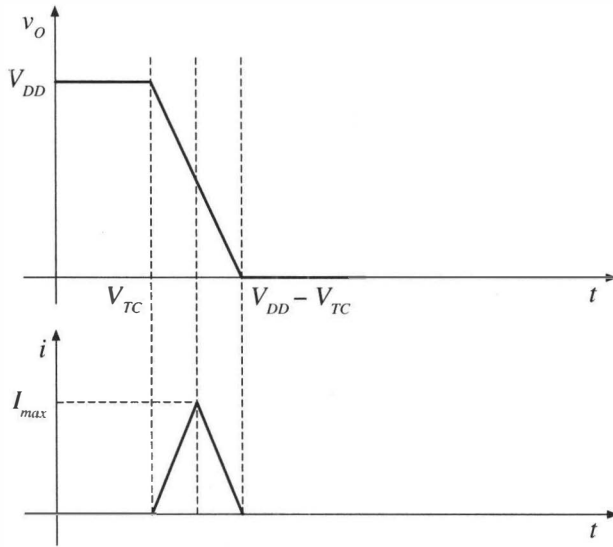


Figura 6.34 Funzione di trasferimento e corrente linearizzati

Si noti che

$$t_{HL} \propto \frac{C}{K} \quad (6.74)$$

Quindi la (6.73) diviene

$$P = f_{CLOCK} V_{DD} V' C \quad (6.75)$$

con V' opportuna costante (dimensionalmente una tensione). In definitiva il contributo alla dissipazione dinamica dovuto alla corrente di commutazione dipende linearmente dalla capacità parassita C e non dipende da K .

Confrontiamo i due contributi della dissipazione dinamica. Usando i valori numerici già usati, l'energia dissipata dalla carica-scarica del capacitore è 2.5pJ, mentre l'energia data dalla (6.69) è 0.21pJ.

Un aspetto molto importante dei parametri prestazionali della famiglia CMOS è che sia i tempi di commutazione sia la dissipazione dipendono linearmente dalla capacità parassita C del gate dei MOSFET. D'altra parte tale capacità è proporzionale alla superficie occupata dai MOSFET. Quindi più piccoli sono i transistor più la porta è veloce e minore è il consumo. Non si potrebbe volere niente di meglio! E, in effetti, le porte CMOS in linea di principio possono essere miniaturizzate con l'unico limite che le dimensioni dei particolari costruttivi dei transistor rimangano molto più grandi della dimensione degli atomi. Superato questo limite (e cominciamo ad essere vicini) sarà necessario prendere in considerazione il comportamento quantistico di atomi ed elettroni.

Un altro punto importante relativo alle prestazioni della CMOS è che la potenza dissipata dipende dal quadrato della tensione di alimentazione. Minore è la tensione di alimentazione minore è il consumo. E, in effetti, le logiche CMOS per i dispositivi portatili a basso consumo sono progettate per lavorare alle tensioni minime compatibili con la tensione di soglia. Anzi talvolta lavorano persino sotto soglia. Infatti, la soglia non è così netta come si potrebbe pensare. Il valore di soglia è definito come il valore di tensione per cui si ha "inversione", ovvero il valore per cui la concentrazione dei portatori n nel semiconduttore di tipo p in prossimità dell'ossido è pari alla concentrazione dei portatori p lontano dall'ossido, ma ciò non significa che al di sotto della soglia non ci sia del tutto conduzione: anche per valori di tensione minori della soglia c'è un certo numero di elettroni in conduzione nel canale. Questo fatto è utilizzato da alcuni dispositivi a bassissimo consumo che operano con tensioni minori della somma delle due soglie (del MOSFET p e del MOSFET n).

Purtroppo la CMOS non ha solo aspetti positivi: i tempi di salita e discesa non sono così rapidi come ci si potrebbe aspettare da una famiglia basata sul principio degli interruttori complementari. Il problema è nella scarica del capacitore attraverso il MOSFET: solo nella prima fase della discesa il MOSFET è in saturazione e quindi in grado di pilotare una discreta corrente, nella fase successiva è in regione lineare e la corrente che è in grado di pilotare è molto minore. In generale, i transistor bipolari sono in grado di pilotare correnti molto maggiori e rimangono in regione attiva diretta (la regione di funzionamento equivalente alla saturazione del MOSFET) fino a valori di tensione molto bassi (tipicamente 0.2V). Vedremo nel prossimo paragrafo come è possibile coniugare i vantaggi del MOSFET con quelli del BJT (a prezzo, tuttavia, di una maggior complessità e quindi anche di una maggiore dimensione fisica della porta elementare).

6.7 Bi-CMOS

Si faccia riferimento al circuito in Figura 6.35 (detto *totem-pole*).

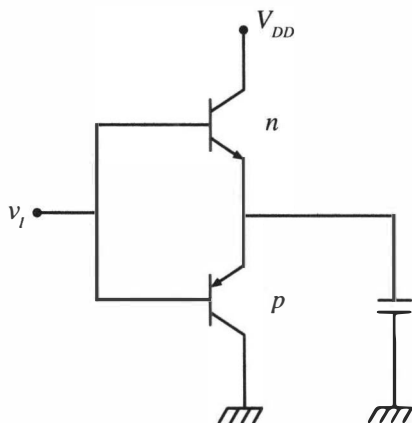


Figura 6.35 Totem-pole

Supponiamo che la capacità sia inizialmente scarica e che la tensione in ingresso passi repentinamente da 0V a V_{DD} (Figura 6.36). Il pnp è interdetto, l'nnp è in regione attiva diretta. La corrente che carica il capacitore è pari a

$$i_I = \beta_F i_{B1} \quad (6.76)$$

dove β_F è il guadagno in corrente del BJT (tipicamente tra 50 e 200) e i_{B1} la corrente che può fornire il generatore di segnale in ingresso.

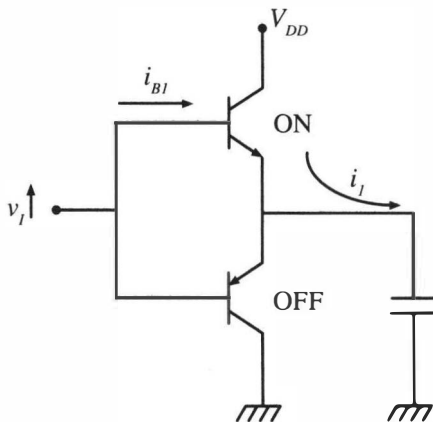


Figura 6.36 Carica del capacitore in uscita

Consideriamo ora il caso in cui la capacità C è carica (a tensione pari a V_{DD}) e in ingresso la tensione passa repentinamente da V_{DD} a 0V (Figura 6.37). L'nnp si interdice e il pnp passa in regione attiva diretta e quindi la capacità si scarica mediante una corrente pari a

$$i_2 = \beta_F i_{B2} \quad (6.77)$$

dove i_{B2} è la corrente che può pilotare il generatore di segnale in ingresso.

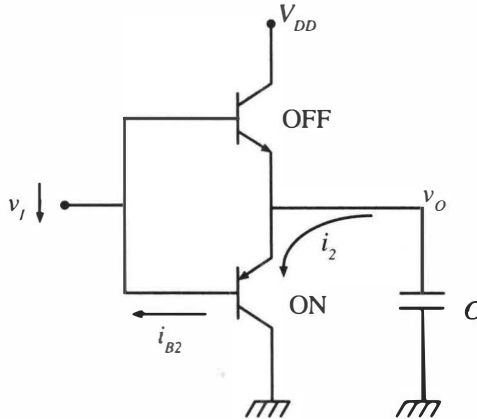


Figura 6.37 Carica del capacitore in uscita

L'idea alla base della logica Bi-CMOS è utilizzare un buffer a BJT come stadio di uscita di una porta CMOS, come mostrato in Figura 6.38. In questo modo la corrente in uscita al primo stadio (il CMOS) è amplificata dallo stadio successivo di un fattore pari a β_F e pertanto la carica (o la scarica) della capacità di carico è più rapida dello stesso fattore.

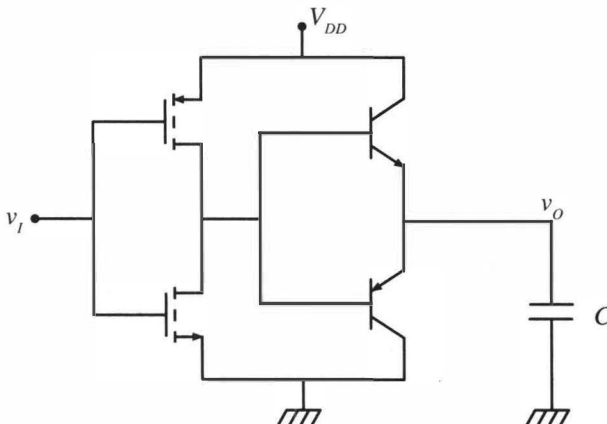


Figura 6.38 Bi-CMOS

Si noti che quando la tensione di uscita supera $V_{DD} - 0.6V$ la giunzione base emettitore del BJT in alto non è più polarizzata sopra soglia. La tensione di uscita continua a salire, ma molto più lentamente. La stessa cosa succede quando la tensione di uscita è inferiore a $0.6V$. Di fatto in condizioni operative, l'escursione del segnale in uscita è $V_{DD} - 1.2V$. Per poter usare l'intera escursione della tensione di alimentazione si può utilizzare una seconda porta CMOS collegata direttamente all'uscita in modo da forzare l'uscita a 0 o a V_{DD} .

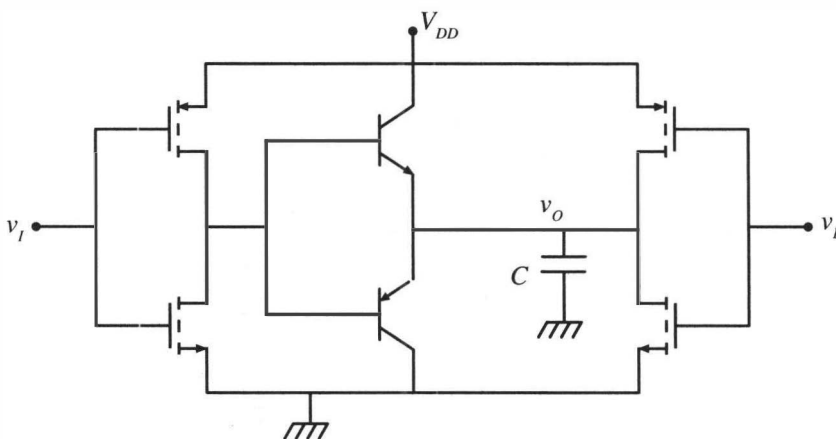


Figura 6.39 Bi-CMOS ad ampia escursione

6.8 Classificazione delle memorie

Gli elementi fondamentali di un sistema di elaborazione dati sono l'unità logica e la memoria. Per la precisione in un processore ci sono anche altri dispositivi: i **registri** e i **contatori**. In linea di principio sono anch'essi memorie, ma si è soliti usare il termine memoria per dispositivi in grado di immagazzinare grandi quantità di dati (almeno molte migliaia di byte).

Innanzitutto le memorie (propriamente dette) si classificano in volatili e non volatili. Le memorie volatili perdono l'informazione contenuta quando staccate dall'alimentazione, sono pertanto memorie di lavoro. Le memorie non volatili mantengono l'informazione indipendentemente dall'alimentazione e quindi sono usate per conservare l'informazione.

Sono memorie volatili i registri e i contatori. I primi sono piccole memorie di 8, 16, 32 bit utilizzate per immagazzinare i risultati parziali di un calcolo mentre l'unità logica lavora. I contatori tengono traccia del numero di colpi di clock e sono fondamentali per ogni processo sequenziale. Le memorie volatili propriamente dette sono essenzialmente di due tipi: SRAM (*Static Random Access Memory*) e DRAM (*Dynamic Random Access Memory*). In entrambi i casi a ogni

singola cella di memoria è associato un indirizzo e quindi è possibile accedere singolarmente a ciascuna cella senza un particolare ordine (da qui il termine “random”, come al solito un po’ fuorviante...). La SRAM, contrariamente a quanto potrebbe far supporre il termine “Static”, è molto più veloce della DRAM ed è usata per la cosiddetta memoria “cache” (termine francese che significa nascosta), ovvero la memoria più interna al processore, quella sulla quale scrive direttamente. La DRAM è la memoria che comunemente viene chiamata RAM. Di solito è una schedina separata che può essere sostituita piuttosto facilmente (anche dall’utente) per migliorare le prestazioni del computer: la dimensione della RAM è infatti uno degli elementi che maggiormente limitano le prestazioni computazionali di un computer (l’altro fattore è il clock, ma la sua sostituzione con uno più veloce, di solito, manda fuori specifica le porte logiche, con conseguenti malfunzionamenti).

Un’altra memoria volatile piuttosto usata è la FIFO (*First In First Out*). Si tratta di una memoria senza indirizzi ma con solo un ingresso e un’uscita. La sua funzione è fare da tampone tra due sistemi digitali con clock diversi. Un esempio tipico può essere il trasferimento dei dati dal computer alla stampante mediante la porta USB. Il trasferimento dei dati avviene dal computer all’ingresso di una FIFO della stampante collegata all’USB, la stampante preleva dall’uscita della FIFO con i tempi che gli sono propri.

Un’altra grande categoria di memorie è costituita da quelle non volatili. Ce ne sono di vari tipi: le ROM (*Read Only Memory*), che sono memorie in cui l’informazione è direttamente cablata nella disposizione dei collegamenti circuitali tra i transistor (a rigore non sono memorie, nel senso di dispositivi in grado di immagazzinare un dato e modificarlo, ma semplicemente delle grandi funzioni logiche che rigidamente collegano un’uscita a un indirizzo); le EPROM (*Erasable and Programmable ROM*) che sono delle ROM cancellabili e programmabili, ovvero scrivibili, sebbene con una procedura piuttosto lenta e complessa; infine, le FLASH. Queste ultime hanno raggiunto la capienza e la velocità dei dischi rigidi con il grande vantaggio di una memoria interamente a stato solido, senza parti meccaniche. Nel seguito descriveremo il principio di funzionamento delle memorie SRAM, DRAM e FLASH.

6.9 RAM statica

Una RAM è un dispositivo che nella sua configurazione più semplice ha N linee per l’indirizzamento, una linea di controllo per abilitare la scrittura o la lettura e M linee ingresso/uscita. In realtà, come vedremo, su una cella di memoria è possibile memorizzare un solo bit per ogni indirizzo, quindi ad esempio una memoria a 64 bit (8 byte) è un banco di 64 memorie da 1 bit indirizzate dalle stesse linee.

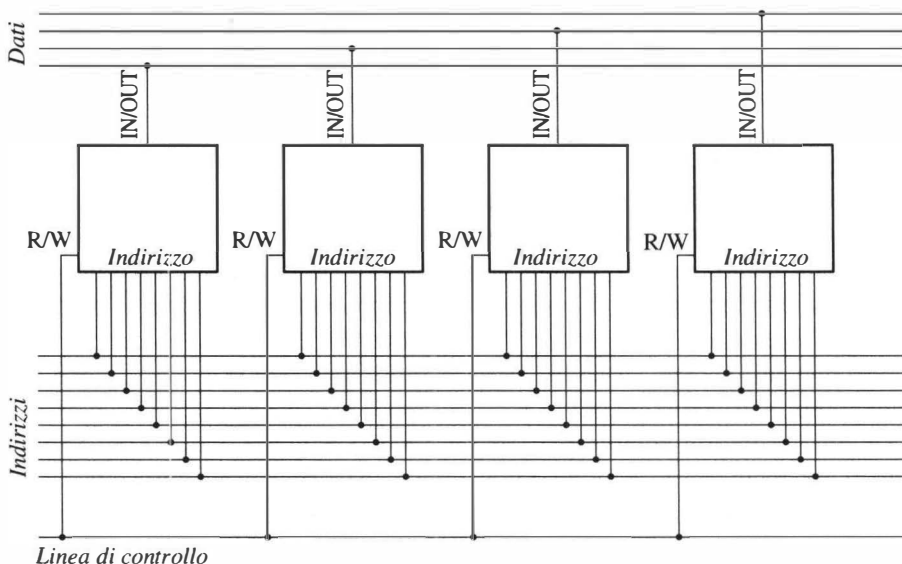


Figura 6.40 Quattro memorie da un bit indirizzate dalle stesse linee realizzano una memoria a 4 bit

Una memoria a 64 bit (64 memorie da 1 bit) con 26 linee di indirizzo (2^{26} celle) è in grado di memorizzare 512MB. Infatti, ricordando che 1byte = 8bit = 2^3 bit, 1kB è $1024 = 2^{10}$ byte, 1MB è $1024 \times 1024 = 2^{20}$ byte.

$$\frac{64 \times 2^{26}}{2^{20} \times 2^3} = 512\text{MB} \quad (6.78)$$

La singola cella di memoria statica è essenzialmente costituita da due porte NOT a loop. Questa configurazione è detta *latch* (catenaccio).

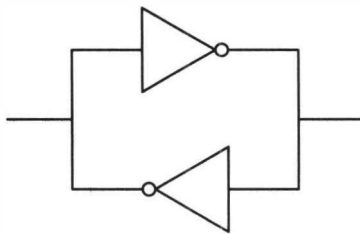


Figura 6.41 Latch

Per comprenderne il funzionamento si consideri la funzione di trasferimento di due NOT in cascata. La funzione di trasferimento complessiva fornirà un'uscita bassa per valori bassi dell'ingresso e alta per valori alti (vedi Figura 6.42).

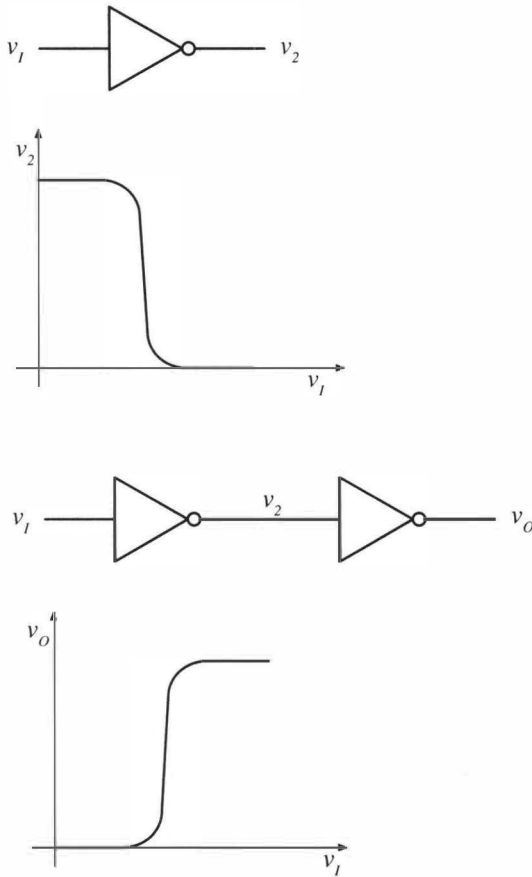


Figura 6.42 Funzione di trasferimento di due NOT in cascata

Collegiamo ora l'uscita con l'ingresso, ovvero chiudiamo il loop. Ciò equivale a tracciare la retta $v_O = v_I$ sul piano $v_O v_I$ come mostrato in Figura 6.43. I possibili stati in cui può trovarsi un latch sono le intersezioni della funzione di trasferimento con la retta di pendenza unitaria passante per l'origine: lo stato basso (A), lo stato alto (C) e uno stato intermedio (B).

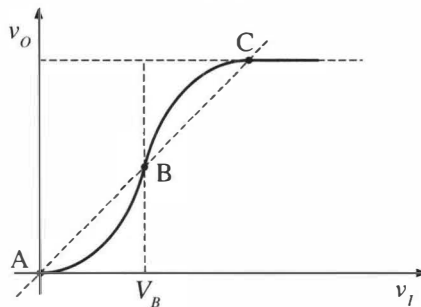


Figura 6.43 Punti di equilibrio di un latch

Gli stati alto e basso sono stabili, ovvero permangono anche in presenza di disturbi. Infatti, un disturbo in ingresso Δv_i , a causa della pendenza (in valore assoluto) minore di 1 della funzione di trasferimento, dà un disturbo in uscita $\Delta v_o < \Delta v_i$. Questo disturbo attenuato è riportato in ingresso e dopo un altro loop è ulteriormente attenuato e così via.

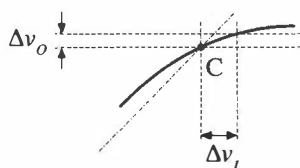


Figura 6.44 Stabilità del punto C

Di fatto le due configurazioni alto e basso sono stabili e robuste, adatte a memorizzare un bit di informazione. Il terzo possibile stato si trova in una regione a pendenza (in valore assoluto) maggiore di 1, ciò significa che qualsiasi disturbo, anche piccolissimo, verrebbe amplificato facendo precipitare il latch in uno dei due stati stabili: A (basso) o C (alto).

Nella pratica una SRAM è realizzata come una matrice quadrata di latch a cui corrisponde una griglia di linee elettriche. Una singola cella è indirizzata ponendo alte la linea e la colonna corrispondenti. Su due lati ortogonali della matrice ci sono due dispositivi logici detti decoder (che vedremo in dettaglio più avanti). Un decoder per ogni combinazione delle $N/2$ linee in ingresso pone alta una sola linea in uscita. Quindi $N/2$ linee in ingresso corrispondono a $2^{N/2}$ linee della matrice. In definitiva per ogni combinazione delle N linee complessive dei due decoder vengono asserite (poste alte) una riga e una colonna, quindi è identificata una cella.

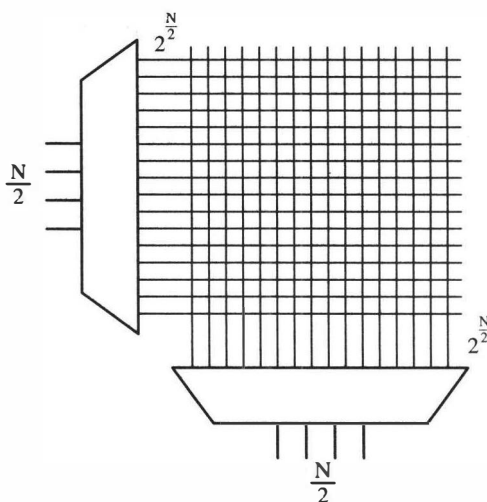


Figura 6.45 Indirizzamento

Vediamo in dettaglio il decoder. La Figura 6.46 mostra lo schema di principio con 2 linee in ingresso e 4 linee in uscita.

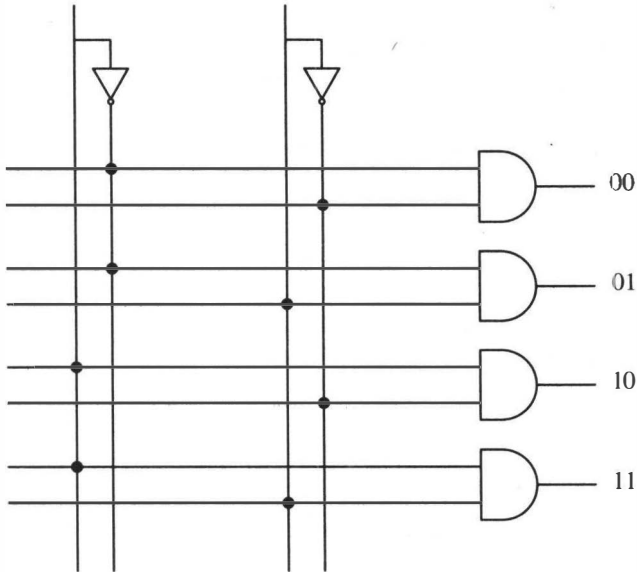


Figura 6.46 Decoder

La Figura 6.47 mostra una possibile implementazione a MOSFET. I MOSFET sono usati come interruttori: quando il gate è posto alla tensione dello stato alto il transistor è in conduzione.

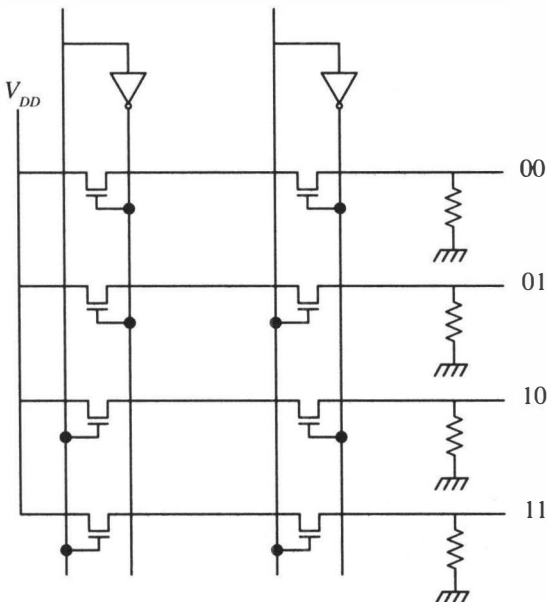


Figura 6.47 Implementazione di un decoder

6.10 RAM dinamica

Lo schema di principio di una DRAM è mostrato in Figura 6.49. Mentre l'indirizzamento è analogo a quanto visto per la RAM statica, l'aspetto notevole della RAM dinamica è che la cella di memoria è realizzata con un singolo MOSFET collegato a un piccolo capacitore integrato.

L'informazione della cella è contenuta nello stato di carica o scarica del capacitore.

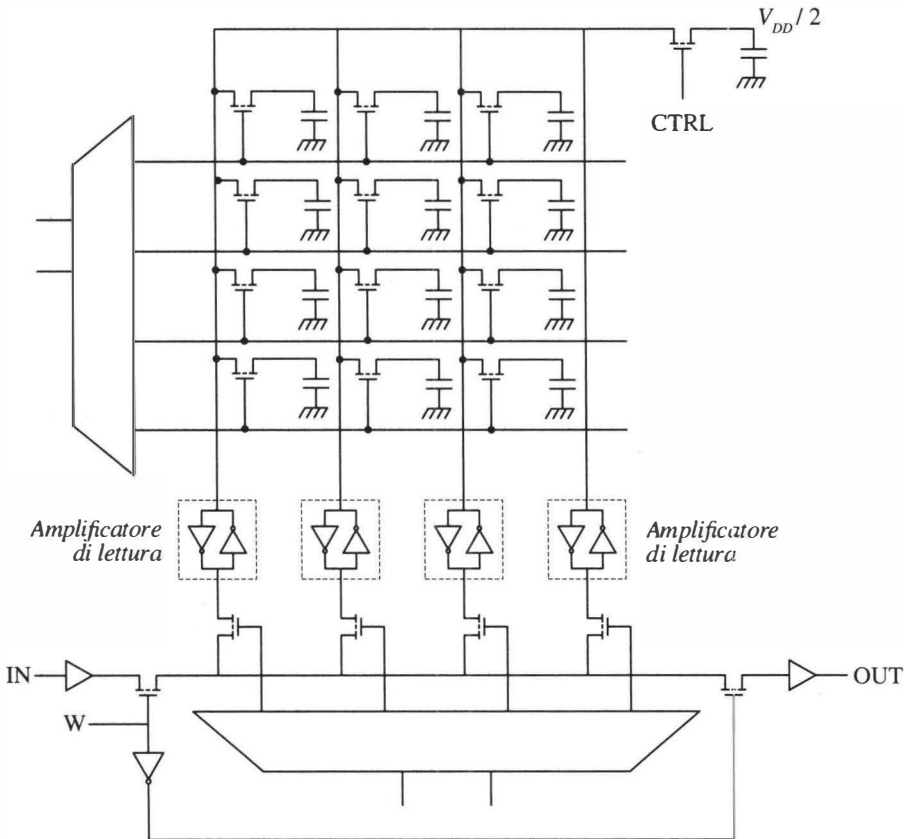


Figura 6.49 DRAM

Un aspetto molto delicato di questo tipo di memoria è relativo a ciò che accade quando il transistor di accesso al capacitore è attivato e la linea elettrica relativa alla colonna è in contatto elettrico con la capacità C della cella. Il problema è che la linea ha dimensioni fisiche molto maggiori del capacitore di memoria e quindi presenta una capacità parassita $C_L \gg C$.

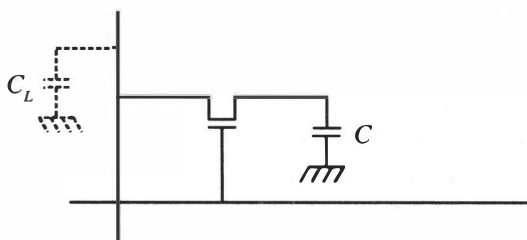


Figura 6.50 Cella di memoria e capacità parassita della linea

Per fissare le idee, supponiamo la cella di memoria nello stato alto ($V_H = V_{DD}$). La linea, prima di essere in contatto elettrico con il capacitore, è precaricata alla tensione $V_{DD}/2$ (vedremo nel seguito maggiori dettagli sull'operazione di pre-carica della linea che è essenziale per la lettura di una cella). La carica elettrica nel capacitore della cella è

$$Q_C = CV_{DD} \quad (6.79)$$

La carica elettrica accumulata nella capacità parassita della linea precaricata a tensione $V_{DD}/2$ è

$$Q_L = C_L \frac{V_{DD}}{2} \quad (6.80)$$

Quando le due capacità sono in contatto elettrico, i due capacitori sono equivalenti a un unico capacitore di capacità pari alla somma delle due capacità. La carica elettrica si ridistribuisce tra i due capacitori e la tensione alla quale si trova la linea (e il capacitore della cella) è pari a

$$v = \frac{Q_C + Q_L}{C + C_L} = \frac{C + \frac{1}{2}C_L}{C + C_L} V_{DD} \quad (6.81)$$

La variazione di tensione Δv della linea dopo che il MOSFET di accesso alla cella è stato attivato è

$$\Delta v = v - \frac{V_{DD}}{2} = \frac{1}{2} \frac{C}{C + C_L} V_{DD} \approx \frac{C}{C_L} \frac{V_{DD}}{2} \quad (6.82)$$

Poiché $C \ll C_L$, la variazione di tensione da rilevare è estremamente piccola e per questo si impiega un particolare amplificatore basato sul comportamento dinamico di un latch. Pochi istanti prima della lettura, il latch è intenzionalmen-

te messo nella condizione di equilibrio instabile (il punto B in Figura 6.43, che nell'esempio di DRAM che stiamo considerando supponiamo pari a $V_{DD}/2$) collegandolo con una tensione di riferimento pari al valore $V_B = V_{DD}/2$. Prima di attivare il collegamento con la cella da leggere, il latch è isolato e per qualche istante, a causa della capacità parassita della linea, rimane in equilibrio instabile. Quando è attivato il collegamento con la cella di memoria, il latch in equilibrio instabile precipita in uno dei due punti di equilibrio stabile, quale dei due dipende dalla piccola variazione di tensione indotta sulla linea dalla carica del capacitore contenente l'informazione. L'amplificatore a latch risolve quindi il problema dell'amplificazione di una piccolissima variazione di tensione fino al livello logico alto o basso, ma risolve anche un secondo importante problema a cui non abbiamo ancora accennato.

Il capacitore di cella C , quando mantiene l'informazione in esso immagazzinata, è isolato dal MOSFET di lettura. Tale isolamento tuttavia non è totale. Un MOSFET in interdizione consente infatti il passaggio di una piccola corrente (la corrente di saturazione inversa delle giunzioni pn contropolarizzate) e quindi i capacitori si scaricano piuttosto rapidamente (in tempi dell'ordine delle decine di millisecondi), perdendo l'informazione in essi immagazzinata. È pertanto necessario, di tanto in tanto, "rinfrescare" (*refreshing*) lo stato di carica (o scarica) dei capacitori di tutte le celle. A questo proposito si noti che l'operazione di lettura mediante l'amplificatore a latch fa esattamente questo: dopo aver letto la piccola variazione, forza l'ingresso (ovvero la linea e il capacitore della cella) al valore alto o basso a seconda del senso positivo o negativo della variazione letta. L'operazione di lettura mediante latch è rigenerativa. Pertanto una DRAM avrà una sua logica di controllo che periodicamente legge tutte le celle, anche senza una richiesta esterna. Questa operazione, ovviamente, ha un certo impatto sulla prontezza dell'accesso a una DRAM che, al momento della richiesta di una lettura, potrebbe essere impegnata in operazioni di *refreshing*.

6.11 FLASH

Le memorie FLASH sono memorie RAM (ogni cella è indirizzabile singolarmente in scrittura e in lettura) con il fondamentale vantaggio di essere in grado di mantenere l'informazione anche senza alimentazione. Il costo per bit delle memorie FLASH è ancora piuttosto alto rispetto agli hard disk convenzionali, ma la tecnologia sta progredendo molto rapidamente ed è prevedibile a breve la completa sostituzione degli hard disk con memorie FLASH, con evidenti vantaggi in termini di compattezza e robustezza alle sollecitazioni.

La cella di memoria FLASH si basa sul MOSFET a gate isolato (*floating gate*). La struttura è mostrata in Figura 6.51.

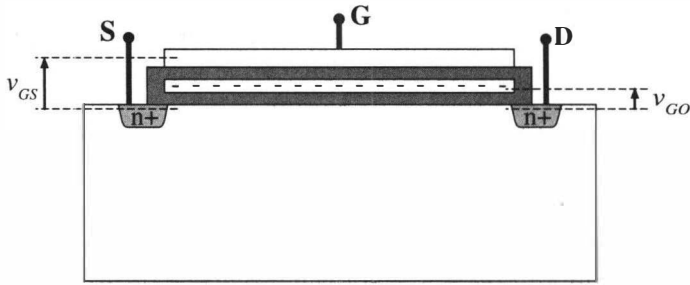


Figura 6.51 MOSFET con gate isolato

Nell'ossido tra gate e canale è realizzato un secondo gate completamente isolato dall'ossido stesso. Il gate accessibile e il gate isolato costituiscono un capacitore metallo-metallo, il gate isolato e il canale costituiscono un capacitore MOS. Supponiamo che il gate isolato sia caricato con una carica negativa $-Q$ (vedremo in seguito come), tale carica può permanere indefinitamente nel tempo. Se tra gate isolato e canale è presente una tensione v_{G0} (con il verso come indicato in Figura 6.51), nel gate isolato deve essere presente una carica pari a

$$Q_1 = C v_{G0} \quad (6.83)$$

Anche l'eventuale tensione tra gate e gate isolato fa sì che nel gate isolato ci sia una carica

$$Q_2 = -C(v_{GS} - v_{G0}) \quad (6.84)$$

La carica complessiva è la somma delle due cariche:

$$Q_1 + Q_2 = -Q \quad (6.85)$$

Pertanto

$$v_{G0} = \frac{1}{2} \left(v_{GS} - \frac{Q}{C} \right) \quad (6.86)$$

D'altra parte, nel MOS costituito dal gate isolato e il silicio, compaiono cariche libere, e quindi un canale, quando

$$v_{G0} > V_{TN} \quad (6.87)$$

ovvero sostituendo la (6.85) nella (6.86)

$$v_{GS} > 2V_{TN} + \frac{Q}{C} \quad (6.88)$$

In conclusione il MOSFET con gate isolato è come se avesse tensione di soglia (V_{TF}) pari a

$$V_{TF} = 2V_{TN} + \frac{Q}{C} \quad (6.89)$$

L'aspetto notevole della (6.89) è che la tensione di soglia può essere controllata da un'eventuale carica nel gate isolato. Al limite, immettendo una carica (negativa) sufficiente a portare la tensione di soglia a un valore maggiore dell'alimentazione, il MOSFET risulta sempre interdetto.

L'idea alla base della memoria FLASH è realizzare una matrice di MOSFET a gate isolato. L'informazione è associata allo stato di interdizione o meno del singolo MOSFET quando selezionato dalla matrice di indirizzamento. Nell'esempio in Figura 6.52 una cella di memoria con il MOSFET a gate isolato scarico (quindi attivo), quando indirizzata, dà un valore basso in uscita. Al contrario se il MOSFET a gate isolato è carico l'uscita rimane alta anche quando la cella è selezionata.

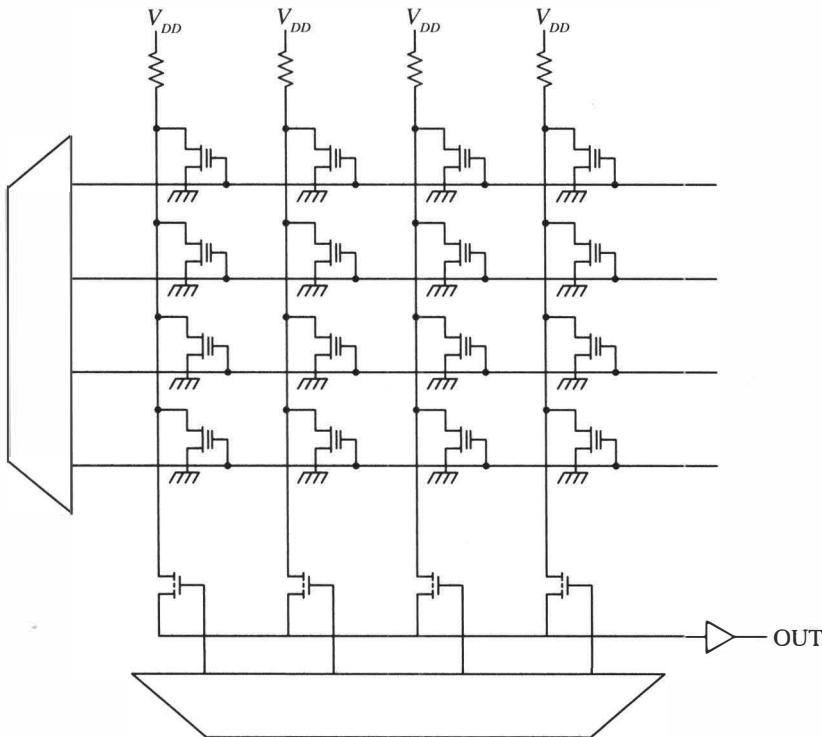


Figura 6.52 FLASH

Vediamo ora come è possibile immettere elettroni nel gate isolato e come scaricarlo. Applicando grandi tensioni positive v_{DS} e v_{GS} e rispettando la condizione

$$v_{DS} > v_{GS} - V_{TF} \quad (6.90)$$

si mette il MOSFET in condizione di forte saturazione. Gli elettroni vengono iniettati a grande velocità tra la fine del canale e il drain. Poiché il gate è posto a una elevata tensione positiva, il campo elettrico positivo dal canale verso il gate devia gli elettroni verso l'ossido. L'interfaccia silicio-ossido per gli elettroni costituisce una barriera di energia pari 3.2eV. Se l'energia degli elettroni deviati verso il gate è maggiore (i fisici parlano di “elettroni caldi” per la loro solita propensione a usare termini piuttosto fuorvianti), questi attraversano il sottile strato di ossido e caricano negativamente il gate isolato. Il processo è autolimitante perché man mano che il gate isolato si carica produce un campo elettrico che si oppone alla tensione positiva del gate.

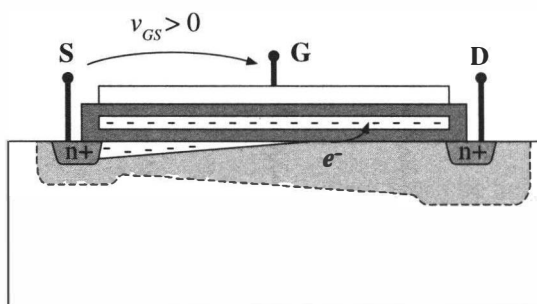


Figura 6.53 Carica del gate isolato mediante iniezione di elettroni caldi

Per scaricare il gate isolato si sfrutta l'effetto tunnel. Come mostrato in Figura 6.54 una parte del gate isolato si sovrappone alla piazzola del source. Facciamo riferimento a questa zona in cui source, gate isolato e gate si sovrappongono.

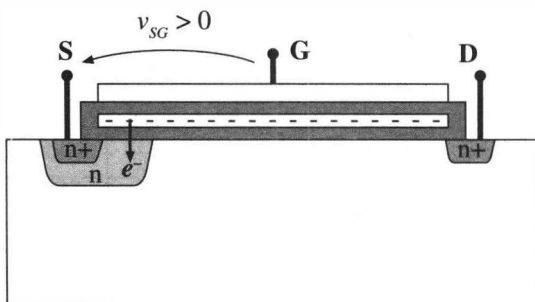


Figura 6.54 Scarica del gate isolato mediante effetto tunnel

In Figura 6.55 è mostrato l'andamento dei livelli energetici quando il semiconduttore e i due gate sono allo stesso potenziale.

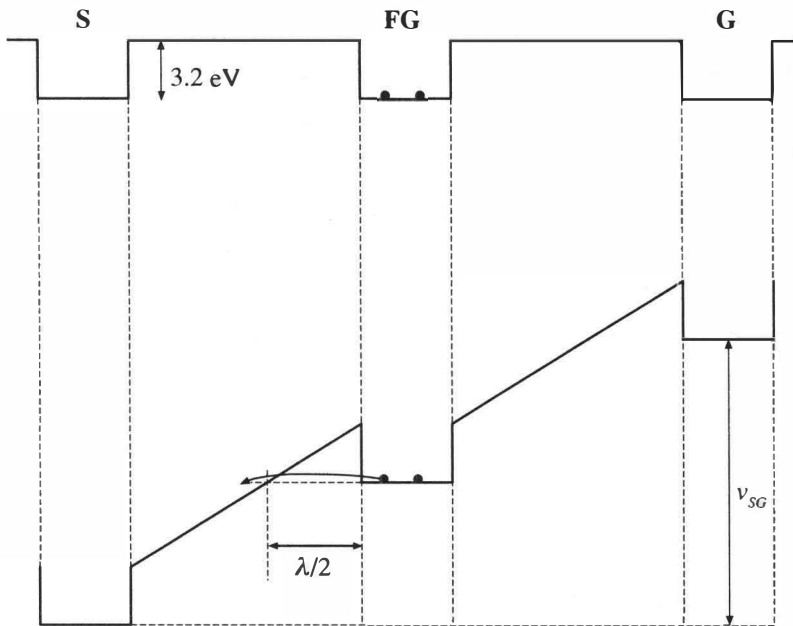


Figura 6.55 Effetto tunnel

L'ossido che separa il source dal gate isolato e il gate isolato dal gate può essere visto come una barriera alta 3.2eV molto più larga della lunghezza d'onda e quindi in grado di confinare gli elettroni. Se ora applichiamo una grande tensione $v_{SG} > 0$ il campo elettrico tra gate e source determina una pendenza dei livelli nell'ossido (ma non nel gate e nel semiconduttore del source, perché avendo cariche mobili queste si ridistribuiscono annullando il campo elettrico). All'aumentare della pendenza dei livelli energetici, lo spessore di ossido che gli elettroni devono attraversare prima di trovarsi in una regione a più bassa energia diminuisce. Abbiamo già visto nel Capitolo 1 che a un elettrone a temperatura ambiente è associata una lunghezza d'onda di circa 10nm e che esso può attraversare una barriera minore della sua semilunghezza d'onda (5nm). Pertanto quando lo spessore dell'ossido da attraversare risulta pari a 5nm gli elettroni possono fluire liberamente dal gate isolato al source.

6.12 Riepilogo

- Un sistema digitale per l'elaborazione di dati è costituito da: unità logiche, memorie, dispositivi di input e output, clock e un sistema di connessione.
- Una famiglia logica è un insieme di dispositivi elettronici che possono essere connessi tra loro e realizzare una qualsiasi funzione logica.
- Una famiglia logica è caratterizzata da: margini di rumore, tempo di propagazione, tempi di salita e discesa, consumo statico, consumo dinamico.
- La porta NOT-CMOS è costituita da due MOSFET giustapposti, rispettivamente a canale n (tra uscita e massa) e a canale p (tra alimentazione e uscita), controllati da uno stesso terminale (l'ingresso della porta).
- Le caratteristiche di una porta NOT-CMOS simmetrica sono le seguenti:

1. Margini di rumore

$$V_{IL} = \frac{3}{8}V_{DD} + \frac{1}{4}V_{TC}$$

$$V_{IH} = \frac{5}{8}V_{DD} - \frac{1}{4}V_{TC}$$

2. Tempo di propagazione

$$t_P = \frac{2C}{K} \frac{V_{TC}}{(V_{DD} - V_{TC})^2} + \frac{C}{K} \frac{1}{(V_{DD} - V_{TC})} \ln \left(\frac{3V_{DD} - 4V_{TC}}{V_{DD}} \right)$$

3. Tempi di salita e di discesa

$$t_{HL} = t_{LH} = \frac{2C}{10K} \frac{9V_{TC}}{(V_{DD} - V_{TC})^2} + \frac{C}{K} \frac{1}{(V_{DD} - V_{TC})} \ln \left(\frac{19V_{DD} - 20V_{TC}}{V_{DD}} \right)$$

4. Potenza statica

$$P_{stat} = 0$$

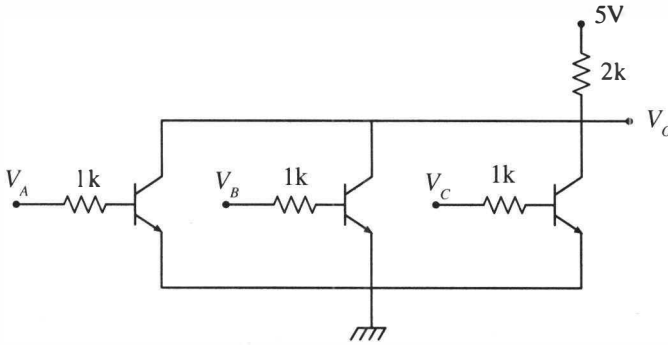
5. Potenza dinamica

$$P_{dyn} = f_{CLOCK} CV_{DD}^2 + f_{CLOCK} V_{DD} K \left(\frac{V_{DD}}{2} - V_{TC} \right)^2 t_{HL}$$

- La cella di memoria di una SRAM è costituita da due porte NOT collegate a *latch*.
- La cella di memoria di una DRAM è costituita da un condensatore collegato alla linea mediante un MOSFET che fa da interruttore di accesso.
- La cella di memoria di una FLASH è costituita da un MOSFET a gate isolato (*floating gate*).

6.13 Esercizi svolti

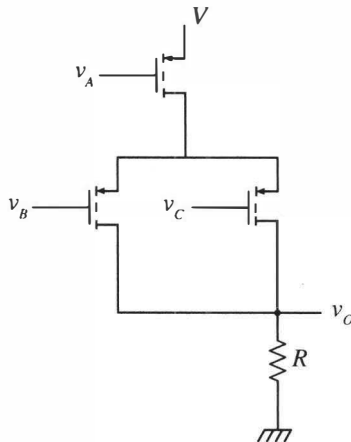
1) Calcolare la tabella di verità del circuito logico in figura ($1 = 5V$, $0 = 0V$).



Quando uno (o più di uno) degli ingressi è alto, il rispettivo BJT è in conduzione e quindi l'uscita è bassa. Segue il risultato in tabella:

V_C	V_B	V_A	V_o
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	0

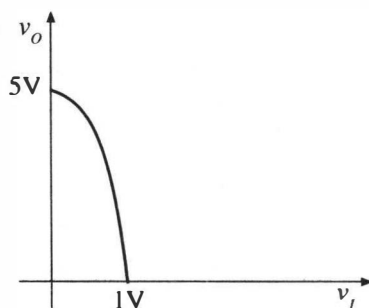
2) Calcolare la tabella di verità del circuito logico in figura ($1 = 5V$, $0 = 0V$). $V = 5V$, $K_p = 4mA/V^2$, $V_{TP} = 1V$, $R = 1k\Omega$.



Quando la tensione in ingresso è alta il rispettivo MOSFET è OFF. La tensione di uscita è alta solo quando c'è un collegamento tra il terminale di uscita e l'alimentazione, altrimenti la resistenza di pull-down porta l'uscita a massa. Il risultato è quello in tabella:

V_C	V_B	V_A	V_O
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	0

3) Il NOT di un'ipotetica famiglia logica ha funzione di trasferimento parabolica come mostrato in figura. Calcolare V_{OH} , V_{OL} , V_{IL} , V_{IH}



Dalla figura è evidente che $V_{OH} = 5V$, $V_{OL} = 0V$.

Per calcolare V_{IL} è necessario trovare il punto in cui la derivata è pari a -1 .

La parabola che passa per i punti $(0, 5V)$ e $(1V, 0)$ è

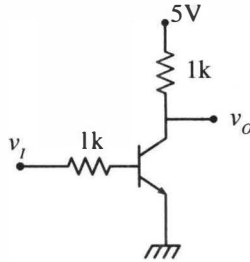
$$v_O = -5v_I^2 + 5$$

Poniamo la derivata pari a -1 :

$$\frac{dv_O}{dv_I} = -10v_I = -1$$

Si ottiene $V_{IL} = 0.1V$. Non c'è un secondo punto in cui la derivata è pari a -1 , ma si può ragionevolmente pensare che in una curva reale il punto $(1V, 0)$ sia un po' smussato e che quindi l'altro punto con derivata -1 sia proprio $V_{IH} = 1V$.

4) Con riferimento al circuito in figura relativo a una porta NOT, si calcoli il consumo statico ($1 = 5\text{V}$, $0 = 0\text{V}$).



Quando $v_i = 0$, il BJT è interdetto e l'uscita è alta. Poiché non scorre corrente: $P_L = 0$.

Quando $v_i = 5$, il BJT è in saturazione e

$$i_B = \frac{5\text{V} - 0.8\text{V}}{1\text{k}\Omega} = 4.2\text{mA}$$

$$i_C = \frac{5\text{V} - 0.2\text{V}}{1\text{k}\Omega} = 4.8\text{mA}$$

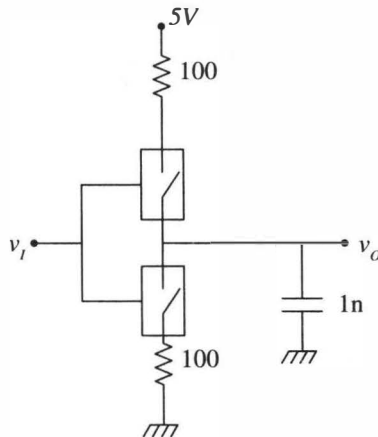
Quindi

$$P_L = (i_B + i_C)(5\text{V}) = 45\text{mW}$$

La potenza statica è

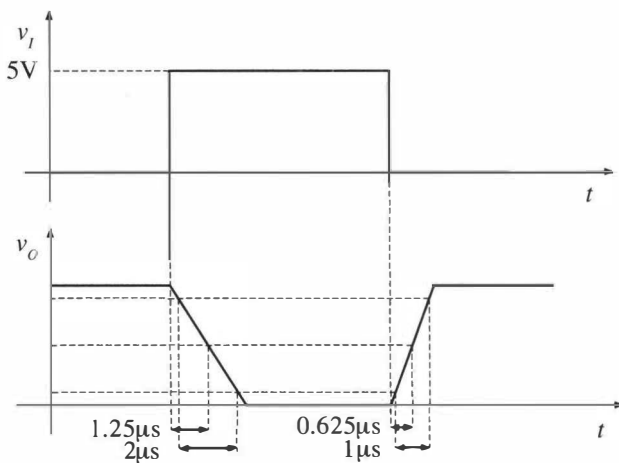
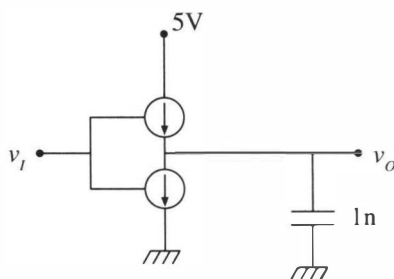
$$P_{STAT} = \frac{P_H + P_L}{2} = 22.5\text{mW}$$

5) Un'ipotetica famiglia logica è realizzata come in figura. I due interruttori sono complementari: quando $v_i < 1\text{V}$ l'interruttore 1 è ON e l'interruttore 2 è OFF; quando $v_i > 1\text{V}$ l'interruttore 1 è OFF e l'interruttore 2 è ON. Calcolare: tempo di salita, tempo di discesa, tempo di propagazione alto-basso, tempo di propagazione basso-alto.



Il capacitore si carica e si scarica attraverso una resistenza di 100Ω . Quindi ponendo $C = 1\text{nF}$ e $R = 100\Omega$ segue che $t_{PLH} = t_{PHL} = 0.69RC = 69\text{ns}$, $t_{HL} = t_{LH} = 2.2RC = 220\text{ns}$.

6) Un'ipotetica famiglia logica è realizzata con due generatori di corrente controllati dalla tensione v_I . Quando $v_I > 3V$ il generatore collegato all'alimentazione è spento e il generatore verso massa si aziona e tira 2mA. Quando $v_I < 3V$ il generatore verso massa è spento, ma il generatore collegato all'alimentazione è acceso ed eroga 4mA. Il carico è costituito da un capacità di 1nF. Calcolare: tempo di salita, tempo di discesa, tempo di propagazione alto-basso, tempo di propagazione basso-alto. Si considerino i fronti di salita e discesa in ingresso molto ripidi.



Il capacitore si scarica e si carica linearmente. La velocità di scarica è

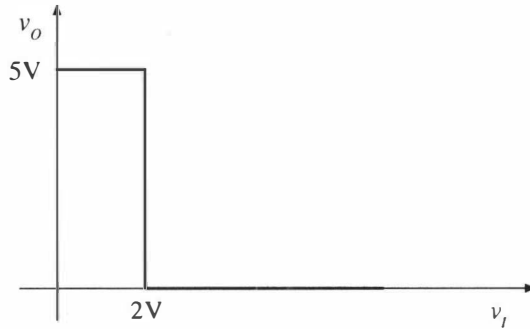
$$\frac{dv_O}{dt} = -\frac{2\text{mA}}{1\text{nF}} = -2 \times 10^6 \text{ V/s}$$

La velocità di carica è

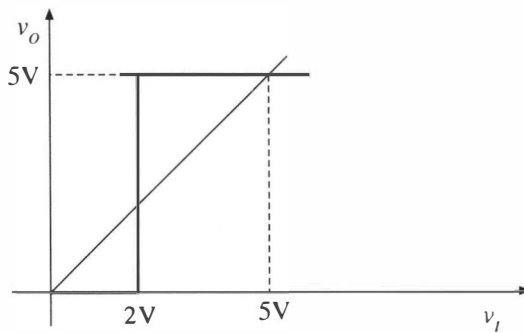
$$\frac{dv_O}{dt} = \frac{4\text{mA}}{1\text{nF}} = 4 \times 10^6 \text{ V/s}$$

Quindi: $t_{HL} = 2\mu\text{s}$, $t_{PHL} = 1.25\mu\text{s}$, $t_{LH} = 1\mu\text{s}$, $t_{PLH} = 0.625\mu\text{s}$.

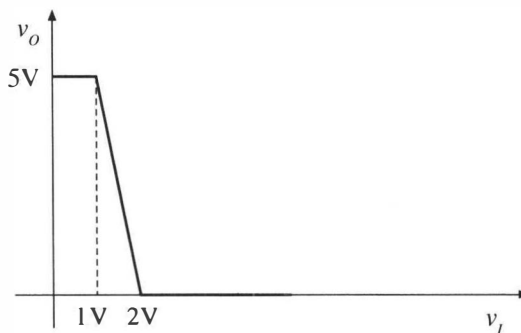
7) Il NOT di un'ipotetica famiglia logica ha funzione di trasferimento come mostrato in figura. Si calcolino i valori dei tre punti di equilibrio di un latch.



La funzione di trasferimento di due porte NOT in cascata è mostrata nella figura che segue. Tracciando la retta di pendenza unitaria si ottengono i tre punti di equilibrio: 0V, 2V, 5V. I punti stabili sono 0V e 5V.



8) Il NOT di un'ipotetica famiglia logica ha funzione di trasferimento con pendenza lineare come mostrato in figura. Si calcolino i valori dei tre punti di equilibrio di un latch.

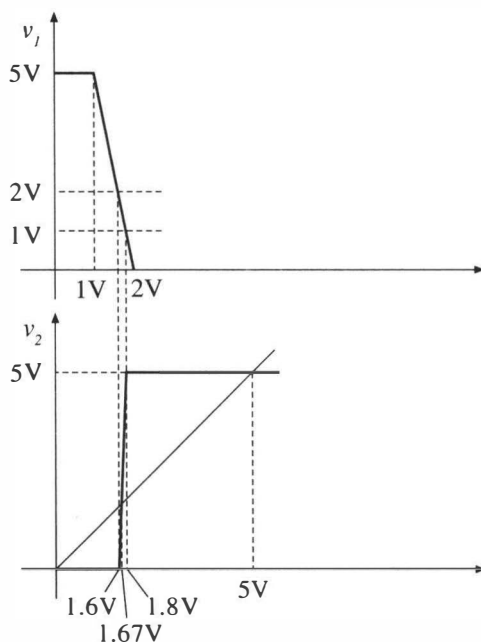


La funzione di trasferimento di due NOT in cascata è mostrata nella figura che segue, dove v_1 è l'uscita della prima porta e v_2 l'uscita della seconda.

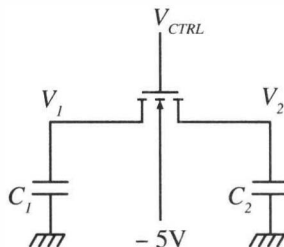
L'espressione di v_2 nella zona crescente è

$$v_2 = -40 + 25v_1$$

L'intersezione con la retta $v_2 = v_1$ è 1.67V. Pertanto i punti di equilibrio sono: 0V, 1.67V, 5V; di cui 0V e 5V stabili.



9) Con riferimento al circuito in figura, $C_1 = 100\text{nF}$, $C_2 = 50\text{nF}$. Supponiamo la tensione di soglia pari a 1V e non dipendente dalla tensione di body (ovvero trascuriamo la dipendenza della tensione di soglia da V_{SB}). Per $t < 0$, $V_1 = 5\text{V}$, $V_2 = 0\text{V}$, $V_{CTRL} = 0\text{V}$. All'istante $t = 0$, V_{CTRL} passa da 0V a 5V. Si calcolino i valori delle tensioni V_1 e V_2 dopo un tempo sufficientemente lungo da poter trascurare il transitorio.



La tensione finale è data da

$$V = \frac{C_1}{C_1 + C_2} V_1 = 3.33\text{V}$$

6.14 Domande di verifica a scelta multipla

1. Si consideri una ipotetica famiglia logica basata su resistenza e interruttore ideale (con resistenza d'ingresso infinita, ma capacità non nulla).
 - a) Il consumo statico dipende dalla capacità d'ingresso
 - b) Il consumo dinamico è proporzionale al quadrato della frequenza di clock
 - c) Il consumo statico è nullo
 - d) Il consumo dinamico diminuisce all'aumentare del tempo di clock
2. Si consideri una ipotetica famiglia logica basata su resistenza e interruttore ideale (con resistenza d'ingresso infinita, ma capacità non nulla).
 - a) Il tempo di commutazione è zero sia per i fronti di salita sia per i fronti di discesa
 - b) Il tempo di salita è più breve di quello di discesa
 - c) Il tempo di salita diminuisce al diminuire della resistenza
 - d) Il tempo di discesa aumenta al diminuire della capacità
3. Si consideri una ipotetica famiglia logica basata su due interruttori complementari ideali (con resistenza d'ingresso infinita, ma capacità non nulla).
 - a) Il tempo di commutazione è zero sia per i fronti di salita sia per i fronti di discesa
 - b) Il tempo di salita è più breve di quello di discesa
 - c) Il tempo di salita diminuisce all'aumentare della tensione di alimentazione
 - d) Il tempo di discesa diminuisce al diminuire della capacità
4. Un requisito fondamentale di ogni famiglia logica è
 - a) impedenza d'ingresso infinita
 - b) tensione di uscita alta pari alla tensione dell'alimentazione
 - c) deve essere in grado di realizzare una porta NOT
 - d) capacità parassita in ingresso trascurabile
5. Nei dispositivi digitali all'aumentare della tensione di alimentazione.
 - a) la velocità di commutazione aumenta
 - b) il consumo dinamico aumenta, ma il consumo statico diminuisce
 - c) le capacità parassite aumentano
 - d) il consumo dinamico e statico aumentano
6. Una sola delle seguenti affermazioni è vera.
 - a) In una porta CMOS non scorre corrente per nessun valore della tensione in ingresso
 - b) Il consumo dinamico di una porta CMOS dipende solo dal tipo di transistor e non dalla tensione di alimentazione
 - c) Il consumo dinamico di una porta CMOS aumenta all'aumentare della frequenza di clock
 - d) Una porta CMOS non ha consumo dinamico
7. In una porta CMOS simmetrica
 - a) V_{IH} e V_{IL} coincidono e sono pari a $V_{DD}/2$
 - b) i due transistor possono essere entrambi in saturazione
 - c) i due transistor sono entrambi in regione lineare quando $v_i = V_{DD}/2$
 - d) V_{IH} è nella retta di pendenza 1 passante per il punto $v_i = V_{DD}/2$ e $v_o = 0$

8. Il tempo di propagazione basso-alto
 - a) è sempre la metà del tempo di salita
 - b) è sempre minore del tempo di salita
 - c) può essere più lungo del tempo di salita
 - d) è sempre uguale al tempo di propagazione alto-basso.
9. Il tempo di propagazione della porta NOT-CMOS
 - a) aumenta all'aumentare della larghezza del gate
 - b) aumenta al diminuire dello spessore dell'ossido
 - c) aumenta linearmente con la tensione di alimentazione
 - d) aumenta con il quadrato della lunghezza del canale
10. In una porta CMOS simmetrica operante con alimentazione a 5V e soglia a 1V
 - a) per tensione d'ingresso nulla scorre corrente
 - b) per tensione di ingresso 5V scorre corrente
 - c) per qualunque valore della tensione d'ingresso non scorre corrente
 - d) per tensione pari a 2.5 V scorre corrente
11. Una sola delle seguenti affermazioni è vera.
 - a) Le memorie dinamiche sono più veloci delle memorie statiche
 - b) I banchi di RAM intercambiabili in un PC sono memorie statiche
 - c) La memoria cache di un PC è costituita da una memoria statica
 - d) La RAM e la cache sono memorie dello stesso tipo, la differenza è che la cache è all'interno del processore mentre la RAM è esterna ed è intercambiabile
12. Una sola delle seguenti affermazioni è vera.
 - a) La memoria statica mantiene i dati memorizzati anche dopo lo spegnimento dell'alimentazione
 - b) Sia la memoria statica sia la memoria dinamica perdono i dati memorizzati dopo lo spegnimento
 - c) La memoria dinamica conserva i dati indefinitamente nel tempo
 - d) La memoria statica necessita di un *refresh* periodico per il mantenimento dei dati nel tempo
13. In una memoria FLASH, il *gate isolato*
 - a) è sempre in contatto elettrico con il canale a causa dell'effetto tunnel
 - b) è scaricato sfruttando l'iniezione di elettroni caldi
 - c) è caricato applicando una elevata tensione positiva al gate
 - d) è scaricato facendo scorrere una elevata corrente tra drain e source
14. Un decoder
 - a) è un dispositivo logico sequenziale in quanto è in grado di memorizzare dati
 - b) ha $2N$ ingressi e N uscite
 - c) può avere più di una linea alta in ingresso
 - d) non può essere realizzato utilizzando porte logiche elementari
15. Una sola delle seguenti affermazioni è vera.
 - a) Una DRAM ben progettata ha capacità parassita delle linee molto più piccola della capacità di immagazzinamento del singolo bit
 - b) La variazione di tensione rivelata dall'amplificatore di lettura è molto più piccola della tensione nel capacitore quando immagazzina uno stato alto

- c) La lettura di una cella DRAM è sempre distruttiva
 - d) Una cella DRAM se non viene letta può conservare un bit indefinitamente
16. Una sola delle seguenti affermazioni è vera.
- a) La RAM di un PC è costituita da un banco di memoria FLASH
 - b) Le memorie statiche si usano per realizzare pen drive USB
 - c) Con le memorie dinamiche si realizzano gli hard-disk dei PC
 - d) Memorie statiche e dinamiche si indirizzano nello stesso modo
17. La cella di una memoria SRAM in tecnologia CMOS
- a) è costituita da 6 transistor
 - b) è costituita da un unico transistor
 - c) è costituita da 2 transistor e 2 resistenze
 - d) è costituita da 2 transistor e una resistenza
18. In una SRAM
- a) il singolo bit è memorizzato nello stato di carica di un capacitore
 - b) la singola cella non può essere più piccola di due porte logiche NOT
 - c) durante la lettura di uno stato basso la carica immagazzinata dalla capacità parassita delle linee di lettura fluisce nel capacitore di immagazzinamento del bit
 - d) il tempo di lettura può essere ritardato dalle operazioni di *refreshing*
19. La SRAM
- a) è una memoria di sola lettura interna al processore
 - b) è una memoria il cui funzionamento si basa sull'effetto tunnel tra canale e gate isolato
 - c) è una memoria in cui ciascuna singola cella richiede almeno 4 transistor
 - d) è una memoria non volatile
20. Un latch a CMOS è costituito
- a) da 2 MOSFET
 - b) da un MOSFET e una capacità di immagazzinamento dello stato logico
 - c) da almeno 6 MOSFET
 - d) da 4 MOSFET
21. Il tempo di salita dell'uscita di una porta CMOS simmetrica
- a) è nullo
 - b) è più lungo del tempo di discesa
 - c) è molto più breve del tempo di propagazione
 - d) diminuisce all'aumentare della costante K
22. Un latch
- a) ha due stati di equilibrio stabile
 - b) ha uno stato di equilibrio stabile e due stati di equilibrio instabile
 - c) non ha stati di equilibrio stabile, infatti oscilla continuamente tra due stati
 - d) ha quattro stati di equilibrio: due stabili e due instabili
23. Una cella di memoria DRAM
- a) conserva l'informazione in un latch
 - b) può essere letta mediante un latch
 - c) è indirizzata mediante una matrice tridimensionale di linee
 - d) è in grado di conservare l'informazione anche quando non è alimentata

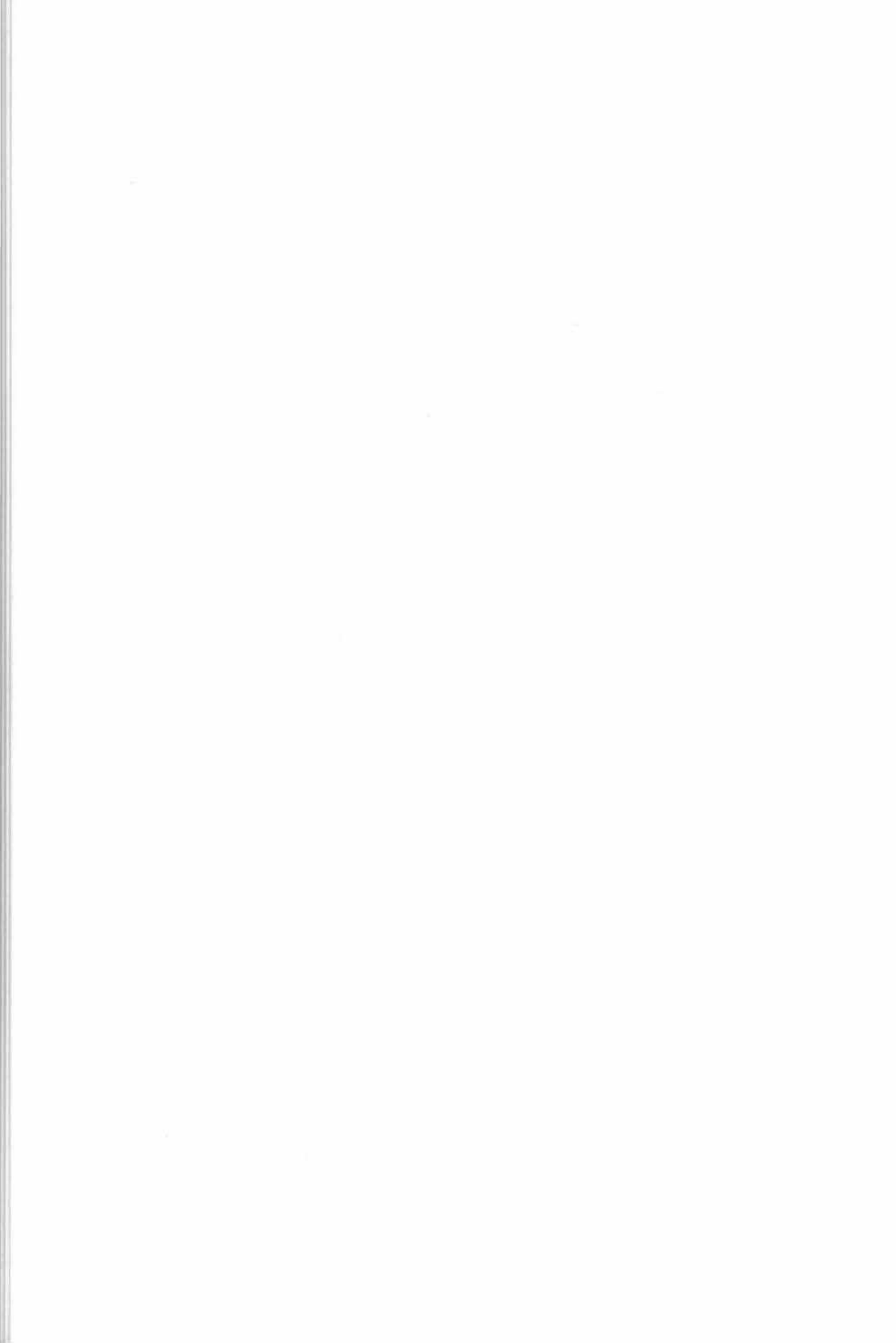
24. Un decoder con N ingressi
 - a) ha $2N$ uscite
 - b) ha 2^N uscite
 - c) ha N^2 uscite
 - d) ha una sola uscita
25. In una generica famiglia logica
 - a) V_{IL} è il valore più basso che l'ingresso può raggiungere
 - b) V_{IH} è il valore di ingresso per cui la pendenza della funzione di trasferimento è 0
 - c) V_{OL} è il valore di uscita per cui la pendenza della funzione di trasferimento è -1
 - d) V_{OH} è il valore massimo che l'uscita può raggiungere
26. In una memoria FLASH il singolo bit è memorizzato
 - a) mediante un latch
 - b) sfruttando la magnetoresistenza del silicio fortemente drogato
 - c) mediante lo stato di carica o scarica di un condensatore
 - d) come stato quantico di un singolo atomo di silicio
27. In una memoria DRAM
 - a) non esistono indirizzi, ma il primo bit scritto è anche l'ultimo a uscire
 - b) la lettura è sempre distruttiva del dato
 - c) l'indirizzamento avviene mediante una matrice di MOSFET gate isolato preprogrammata
 - d) la lettura avviene mediante una sequenza di passi, di cui il primo è applicare un'opportuna tensione a tutte le colonne

6.15 Domande di verifica a risposta aperta

1. Enunciare e spiegare i quattro requisiti di una famiglia logica.
2. Si ricavi la formula del consumo dinamico per una porta a interruttori complementari ideali.
3. Si tracci la funzione di trasferimento della porta NOT-CMOS giustificandone l'andamento.
4. Si calcolino i margini di rumore della famiglia CMOS.
5. Si calcoli il tempo di propagazione della porta NOT-CMOS.
6. Quali sono i vantaggi e i limiti della famiglia logica CMOS?
7. Si illustri il funzionamento della porta NOT-Bi-CMOS.
8. Come funziona la cella di memoria delle SRAM?
9. Si descriva il processo di lettura di una cella di memoria DRAM.
10. Come è realizzato l'amplificatore di lettura delle DRAM?
11. Come funziona una memoria FLASH?

6.16 Risposte alle domande a scelta multipla

1 d	10 d	19 c
2 c	11 c	20 d
3 a	12 b	21 d
4 c	13 c	22 a
5 d	14 c	23 b
6 c	15 b	24 b
7 b	16 d	25 d
8 c	17 a	26 c
9 d	18 b	27 d



7

Amplificatori lineari

In questo capitolo si trattano gli amplificatori lineari. Dopo aver richiamato il principio di sovrapposizione nei circuiti lineari, si introducono i modelli linearizzati dei FET. Si calcolano le resistenze viste ai tre terminali e si analizzano le tre configurazioni (CS, CD, CG). I BJT sono trattati in maniera analoga. Infine si riportano alcuni esempi di analisi e progetto di amplificatori in varie configurazioni.

7.1 Grida e falò...

I segnali si attenuano: è una ineluttabile legge di natura. Un grido a distanza diviene un impercettibile sussurro. Un falò nella notte, un puntino luminoso. L'uomo ha sempre sognato di acuire i propri sensi. All'inizio semplicemente cercando di concentrare il segnale raccolto su un'ampia superficie. Il telescopio funziona essenzialmente su questo principio. Durante la prima guerra mondiale furono usati degli "amplificatori" acustici passivi per sentire a grande distanza l'arrivo degli aerei. E forse proprio dalla loro forma deriva l'attuale simbolo elettrico dell'amplificatore. Ma il passo decisivo fu l'invenzione di un amplificatore attivo, ovvero un dispositivo che non si limitasse a "concentrare" la poca energia rivelata, ma fosse in grado di usare il segnale in arrivo per modulare un'altra sorgente di energia in modo da fornire un segnale "analogo" che poteva essere molto più intenso di quello originario.



Figura 7.1 "Amplificatore" acustico

Nel 1906 l'americano Lee De Forest fu il primo a brevettare un amplificatore attivo a valvole (l'*audion*). E questa può essere considerata la data di nascita del dispositivo chiave di ogni moderno sistema di comunicazione e di acquisizione dati da sensori.

7.2 Gli amplificatori

Un amplificatore è un dispositivo nel quale il segnale procede da un ingresso verso un'uscita, nel senso che il segnale in ingresso determina il segnale in uscita e non viceversa. Si noti che, contrariamente a quanto potrebbe suggerire il nome, non è il fatto di amplificare un segnale che rende un dispositivo un amplificatore. Un trasformatore, ad esempio, è in grado di dare in uscita una tensione più grande di quella in ingresso, ma non è un amplificatore. E d'altra parte ci sono amplificatori che danno in uscita una tensione (o una corrente) minore di quella in ingresso. La caratteristica peculiare degli amplificatori è la direzionalità, ovvero la capacità di controllare la corrente (e/o la tensione) in una maglia di un circuito, mediante la corrente (e/o la tensione) in un'altra maglia indipendente dalla prima.

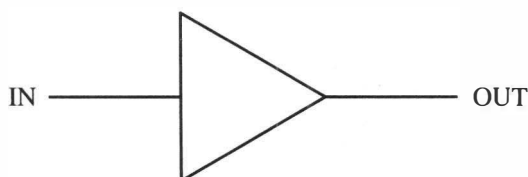


Figura 7.2 Simbolo dell'amplificatore

Un amplificatore è caratterizzato da una funzione di trasferimento che dà il valore del segnale in uscita S_{OUT} in funzione di quello in ingresso S_{IN} .

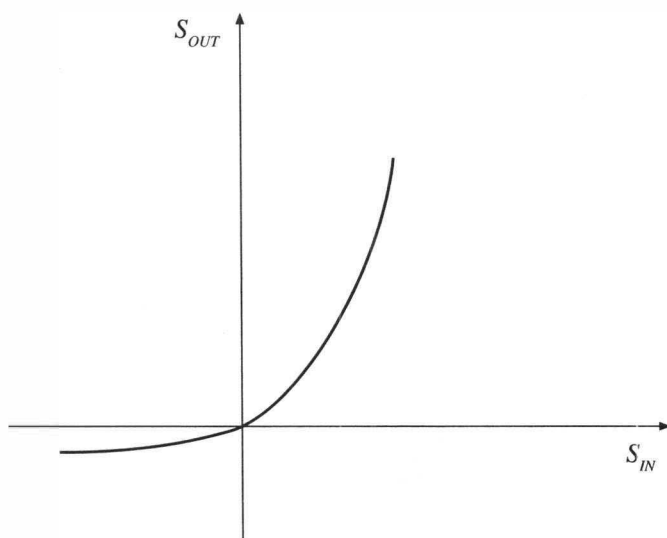


Figura 7.3 Funzione di trasferimento di un amplificatore

Se la funzione di trasferimento non è lineare, il segnale in uscita sarà in qualche modo distorto. Ad esempio per la funzione di trasferimento in Figura 7.3, una sinusoide in ingresso dà un segnale in uscita periodico con dei massimi a cuspide e dei minimi più allargati (Figura 7.4).

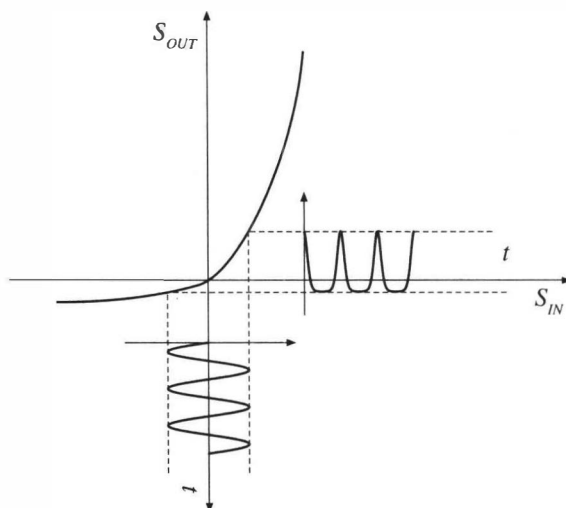


Figura 7.4 Distorsione

Solo una funzione di trasferimento lineare assicura che il segnale in uscita abbia la stessa forma del segnale in ingresso, ovvero che il segnale in uscita sia semplicemente il segnale in ingresso moltiplicato per un fattore (Figura 7.5).

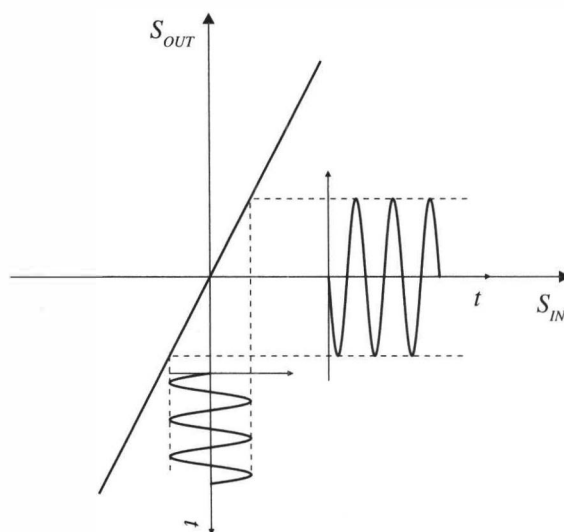


Figura 7.5 Amplificazione senza distorsione

7.3 Amplificatori lineari

Un circuito lineare contiene solo resistenze e generatori di tensione (o corrente) legati linearmente a tensioni (o correnti) su resistenze (non consideriamo per ora capacitori e induttori). Un circuito lineare ha una funzione di trasferimento lineare e un circuito con una risposta lineare può essere sempre visto come un circuito lineare (per un richiamo ai concetti di base dei circuiti lineari vedi Appendice B).

In generale, un amplificatore elettronico è un dispositivo a quattro terminali, due per l'ingresso e due per l'uscita.

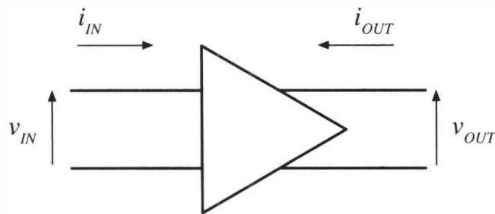


Figura 7.6 Amplificatore

Se i circuiti visti ai terminali di ingresso e di uscita sono lineari, possono essere idealmente sostituiti con il loro equivalente di Thévenin (un generatore di tensione con una resistenza in serie), oppure con il loro equivalente di Norton (un generatore di corrente con una resistenza in parallelo). Le due rappresentazioni sono equivalenti.

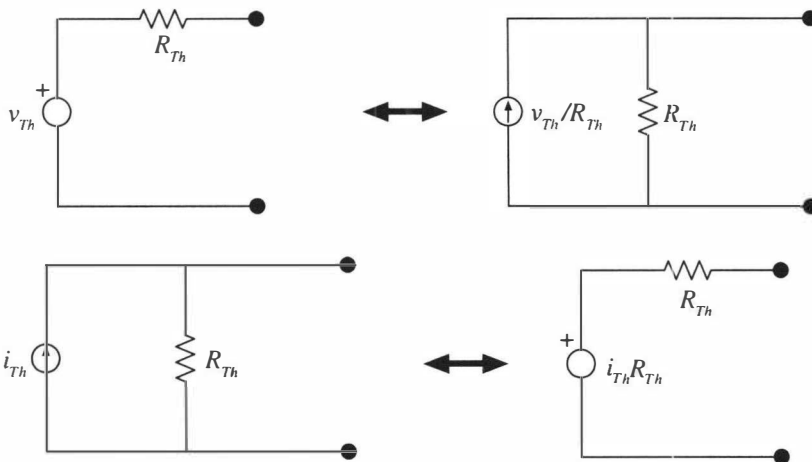


Figura 7.7 Generatori di tensione e di corrente

Per fissare le idee facciamo riferimento alla prima rappresentazione. La resistenza equivalente del circuito in ingresso è detta resistenza di ingresso e la resistenza equivalente del circuito in uscita è detta resistenza di uscita. La tensione ai capi dei due terminali in ingresso è detta tensione di ingresso e la corrente che vi scorre corrente di ingresso. Analogamente si definiscono la tensione di uscita e la corrente di uscita. Il generatore di tensione in uscita è controllato, mediante una relazione lineare, dalla tensione sulla resistenza di ingresso. Il generatore di tensione in ingresso è controllato, mediante una relazione lineare, dalla tensione sulla resistenza di uscita.

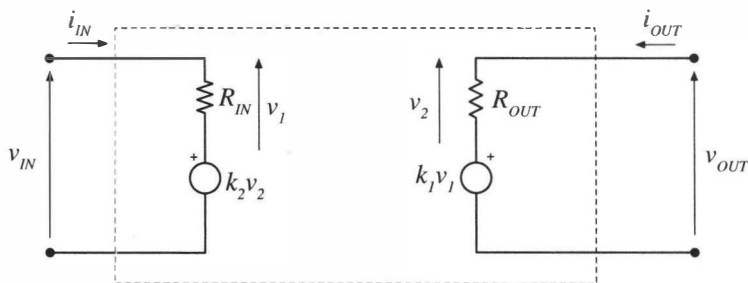


Figura 7.8 Amplificatore come doppio dipolo

Tuttavia, gli amplificatori sono progettati in modo che il segnale proceda in un senso e non nell'altro, quindi il generatore di tensione in ingresso controllato dall'uscita è quasi sempre nullo o trascurabile. Pertanto

$$K_2 \approx 0 \quad (7.1)$$

$$v_1 \approx v_{IN} \quad (7.2)$$

Supponiamo ora che sia lineare non solo l'amplificatore, ma anche i circuiti a monte e a valle dell'amplificatore. Consideriamo quindi un generatore di segnale riducibile al suo equivalente di Thévenin (un generatore di tensione e una resistenza). Anche in uscita dovremmo considerare l'equivalente di Thévenin del circuito, ma di solito l'uscita di un amplificatore è semplicemente un carico passivo (R_L). Tra l'altro abbiamo già visto che l'ingresso di un amplificatore è di fatto un carico passivo pari alla resistenza di ingresso, pertanto se in uscita a un amplificatore c'è un secondo amplificatore in cascata (cosa piuttosto frequente), quest'ultimo può essere visto come un carico passivo. In definitiva possiamo schematizzare un amplificatore lineare come in Figura 7.9.

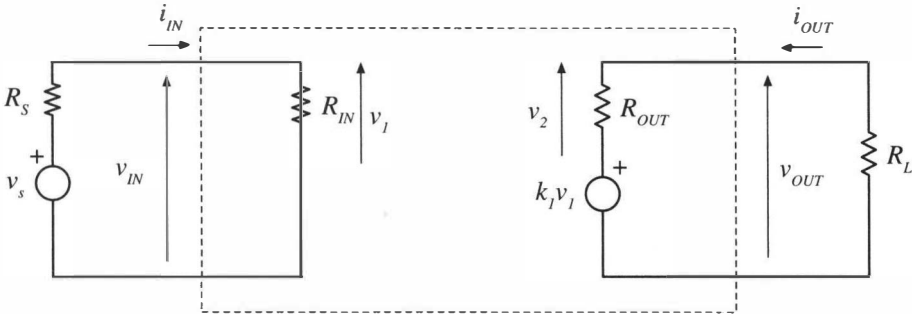


Figura 7.9 Amplificatore come doppio dipolo

Consideriamo il caso $R_s \ll R_{IN}$. Poiché R_s è piccola, la caduta di tensione su tale resistenza è trascurabile e $v_s = v_{IN}$. In altre parole tutta la tensione del generatore di segnale è applicata direttamente all'ingresso dell'amplificatore. In tal caso si parla di **trasferimento di tensione**.

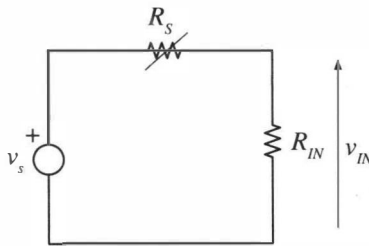


Figura 7.10 Trasferimento di tensione

Consideriamo ora il caso $R_s \gg R_{IN}$. Anziché l'equivalente di Thévenin (generatore di tensione con resistenza in serie) consideriamo l'equivalente di Norton (generatore di corrente con resistenza in parallelo). La resistenza R_s è ora in parallelo con R_{IN} ed essendo grande rispetto a R_{IN} può essere approssimata con un circuito aperto. Se ne deduce che tutta la corrente disponibile al generatore di segnale è entrante nell'amplificatore e, pertanto, in tal caso si parla di **trasferimento di corrente**.

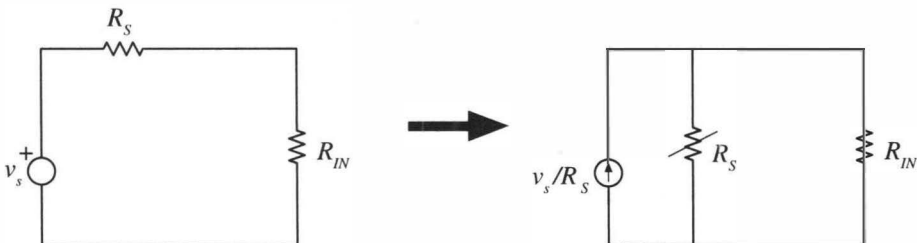


Figura 7.11 Trasferimento di corrente

Per il circuito di uscita si può fare un ragionamento analogo. Per $R_{OUT} \ll R_L$ si ha **trasferimento di tensione** e per $R_{OUT} \gg R_L$ **trasferimento di corrente**.

Sia nel caso di trasferimento di tensione sia nel caso di trasferimento di corrente, la potenza trasferita è molto piccola, idealmente nulla. Talvolta si preferisce trasferire potenza da uno stadio all'altro, piuttosto che tensione o corrente. La massimizzazione della potenza trasferita in un circuito lineare si ottiene quando la resistenza del generatore di segnale è uguale alla resistenza di ingresso dello stadio successivo. Quindi si ha trasferimento di potenza in ingresso quando $R_S = R_{IN}$ e trasferimento di potenza in uscita quando $R_{OUT} = R_L$.

Di solito si cerca di progettare gli stadi di ingresso e di uscita di un amplificatore in modo da massimizzare il trasferimento della tensione, della corrente o della potenza, anche se ovviamente si tratta solo di un criterio di buona progettazione e non di una regola assoluta.

7.4 Principio di sovrapposizione e non linearità

Per circuiti lineari vale il **principio di sovrapposizione** degli effetti. Se in un circuito lineare ci sono più di un generatore di tensione o corrente indipendenti si può analizzare l'effetto sul circuito di un generatore alla volta cortocircuitando tutti gli altri generatori di tensione e trasformando in circuiti aperti tutti gli altri generatori di corrente. Quindi, dopo aver calcolato tensioni e correnti nel circuito per ogni generatore indipendente, si possono semplicemente sommare tutte le tensioni e tutte le correnti in ciascuna maglia e nodo. Il principio di sovrapposizione vale anche per ogni singola frequenza di uno stesso generatore di segnale. Ciò significa che possiamo analizzare il circuito a una data frequenza indipendentemente dalle altre. Nella pratica possiamo analizzare un circuito lineare in due fasi.

- 1) **Analisi DC** (*direct current*). Si calcolano tensioni e correnti del circuito in condizioni di sola polarizzazione dei componenti (senza segnale). Poiché la frequenza è nulla, tutti i capacitori possono essere considerati dei circuiti aperti.
- 2) **Analisi AC** (*alternate current*). Si analizza il circuito alla frequenza del segnale con la certezza che il segnale non modifichi lo stato di polarizzazione dei componenti del circuito.

Si noti bene che il principio di sovrapposizione vale solo per i circuiti lineari. A questo proposito si consideri il circuito non lineare in Figura 7.12, che è un amplificatore a MOSFET come già visto nel Capitolo 3.

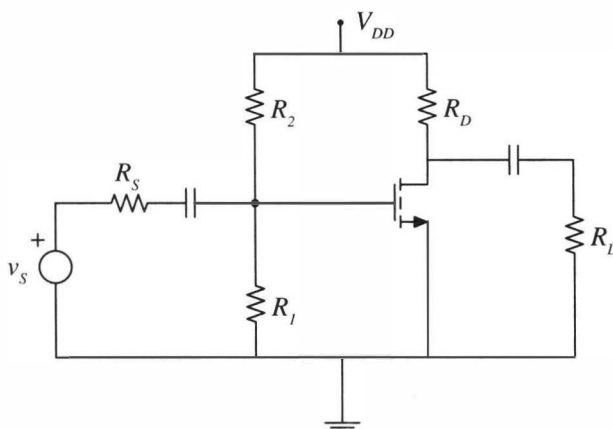


Figura 7.12 Amplificatore a MOSFET

La corrente i_{DS} è una **funzione non lineare** di v_{GS} data da

$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN})^2 \quad (7.3)$$

Il generatore di segnale fornisce una tensione sinusoidale con pulsazione ω .

$$v_s = V_s \cos(\omega t) \quad (7.4)$$

Se R_s è sufficientemente piccolo, v_{GS} sarà la somma della tensione costante di polarizzazione V_{GS} (fissata dal partitore R_1 e R_2) e il segnale v_s . Definiamo

$$x = \frac{V_s}{V_{GS} - V_{TN}} \quad (7.5)$$

si ottiene

$$i_{DS} = I_{DS} (1 + 2x \cos(\omega t) + x^2 \cos^2(\omega t)) \quad (7.6)$$

con

$$I_{DS} = \frac{K_n}{2} (V_{GS} - V_{TN})^2 \quad (7.7)$$

La (7.6) può essere riscritta come

$$i_{DS} = I_{DS} \left(1 + \frac{x^2}{2} + 2x \cos(\omega t) + \frac{1}{2} x^2 \cos(2\omega t) \right) \quad (7.8)$$

La corrente i_{DS} ha quindi tre termini: un termine costante (termine di ordine 0), uno alla frequenza di ingresso (termine di ordine 1) e uno alla frequenza doppia (termine di ordine 2). Il termine costante definisce il punto di polarizzazione del MOSFET. Si noti che tale termine dipende da x , che a sua volta dipende dall'ampiezza del segnale. Questo è un tipico effetto non lineare: il segnale ha effetti anche sulla componente continua e quindi sposta il punto di polarizzazione.

L'ultimo termine è a frequenza doppia rispetto al segnale. In generale una funzione di trasferimento non lineare produce componenti (dette **armoniche**) a frequenze multiple della frequenza del segnale. Il numero di armoniche e la loro ampiezza dipende dalla forma della funzione di trasferimento. Nel caso del MOSFET la funzione di trasferimento è del secondo ordine e l'unica armonica presente è quella a frequenza doppia. Se la funzione di trasferimento avesse avuto un termine del terzo ordine (ovvero se lo sviluppo di Taylor della funzione avesse avuto un termine del terzo ordine non nullo), ci sarebbe stata anche un'armonica alla frequenza tripla.

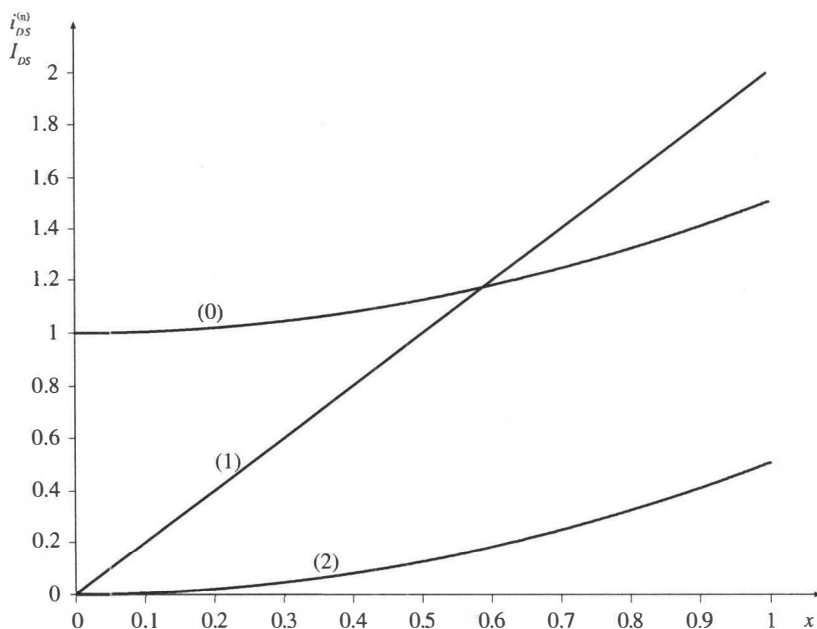


Figura 7.13 Ampiezza dei termini di ordine 0, 1 e 2

L'analisi di un circuito non lineare diventa anche più complessa se in ingresso ci sono contemporaneamente più frequenze. Se, ad esempio, il segnale è composto da due termini alle frequenze ω_1 e ω_2 , ragionando come sopra si può facilmente vedere che una funzione di trasferimento del secondo ordine genera sei frequenze: ω_1 , ω_2 , $2\omega_1$, $2\omega_2$, $\omega_1 + \omega_2$, $|\omega_1 - \omega_2|$, ovvero due armoniche e due **prodotti di intermodulazione**. Le stesse due frequenze applicate a una funzione di tra-

sferimento contenente un termine del terzo ordine producono le ulteriori frequenze: $3\omega_1$, $3\omega_2$, $2\omega_1 + \omega_2$, $\omega_1 + 2\omega_2$, $|2\omega_1 - \omega_2|$, $|2\omega_2 - \omega_1|$ (due armoniche e quattro prodotti di intermodulazione).

La non linearità non è certo una peculiarità dei soli FET. I BJT sono addirittura meno lineari in quanto la funzione di trasferimento esponenziale contiene armoniche a tutti gli ordini (ovviamente decrescenti in ampiezza all'aumentare dell'ordine):

$$i_C = I_T e^{\frac{v_{BE}}{V_T}} = I_T \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{v_{BE}}{V_T} \right)^n \quad (7.9)$$

Si tenga però presente che lo stesso BJT è praticamente un dispositivo lineare se pilotato in corrente, infatti

$$i_C = \beta_F i_B \quad (7.10)$$

Purtroppo, per un buon trasferimento di corrente è necessaria una piccola resistenza di ingresso, e il terminale di base del BJT è invece un ingresso ad alta resistenza. Questo fatto limita molto l'utilità della (7.10) ai fini di ottenere un amplificatore lineare a BJT.

7.5 Modello linearizzato del FET a tre terminali

Nel paragrafo precedente abbiamo visto come l'analisi di un circuito non lineare possa essere piuttosto complessa. Per questa ragione, quando possibile si utilizza il modello linearizzato del FET. In questo paragrafo ricaveremo tale modello e ne definiremo i limiti di validità. Prendiamo in considerazione un MOSFET a canale n ad arricchimento. Definiamo chiaramente la simbologia. Una grandezza indicata con la lettera maiuscola indica una tensione (o una corrente) costante, la minuscola indica una grandezza variabile. Tensioni e correnti costanti sono fissate dalla polarizzazione del transistor. Il pedice di una grandezza può essere maiuscolo o minuscolo. Il pedice minuscolo indica un piccolo segnale, ovvero abbastanza piccolo da poter approssimare la funzione di trasferimento del transistor con il suo sviluppo di Taylor arrestato al primo ordine. Sotto questa ipotesi una qualunque tensione o corrente sarà la somma di due termini: 1) un termine costante dipendente solo dal punto di polarizzazione; 2) un termine piccolo e variabile dipendente solo dal segnale. Con riferimento alle grandezze in Figura 7.14 possiamo scrivere:

$$i_{DS} = I_{DS} + i_{ds} \quad (7.11)$$

$$v_{DS} = V_{DS} + v_{ds} \quad (7.12)$$

$$v_{GS} = V_{GS} + v_{gs} \quad (7.13)$$

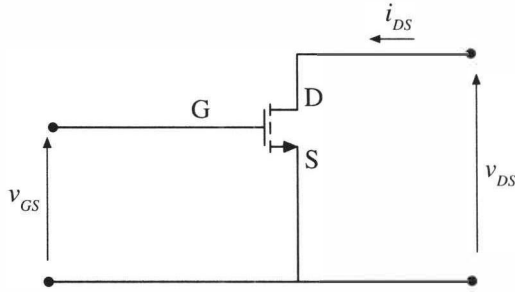


Figura 7.14 MOSFET

La corrente di drain è funzione di v_{GS} e di v_{DS}

$$i_{DS} = i_{DS}(v_{GS}, v_{DS}) \quad (7.14)$$

Sviluppiamo secondo Taylor e arrestiamo lo sviluppo al primo ordine

$$i_{DS} = i_{DS}(V_{GS}, V_{DS}) + \left. \frac{\partial i_{DS}}{\partial v_{GS}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS}}} v_{gs} + \left. \frac{\partial i_{DS}}{\partial v_{DS}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS}}} v_{ds} \quad (7.15)$$

ovvero

$$i_{ds} = \left. \frac{\partial i_{DS}}{\partial v_{GS}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS}}} v_{gs} + \left. \frac{\partial i_{DS}}{\partial v_{DS}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS}}} v_{ds} \quad (7.16)$$

Cominciamo con il calcolo della derivata di i_{DS} rispetto a v_{GS} . In saturazione la relazione tra i_{DS} e v_{GS} è parabolica:

$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN})^2 (1 + \lambda v_{DS}) \quad (7.17)$$

Quindi

$$\left. \frac{\partial i_{DS}}{\partial v_{GS}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS}}} = K_n (V_{GS} - V_{TN}) (1 + \lambda V_{DS}) \approx K_n (V_{GS} - V_{TN}) \quad (7.18)$$

Definiamo la **transconduttanza del FET** come segue:

$$g_m = K_n (V_{GS} - V_{TN}) \quad (7.19)$$

Quindi

$$\left. \frac{\partial i_{DS}}{\partial v_{GS}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS}}} = g_m \quad (7.20)$$

Passiamo ora al calcolo della derivata di i_{DS} rispetto a v_{DS} .

$$\left. \frac{\partial i_{DS}}{\partial v_{DS}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS}}} = \lambda \frac{K_n}{2} (V_{GS} - V_{TN})^2 = \lambda I_{DS} \quad (7.21)$$

Definiamo **resistenza di drain**

$$r_o = \frac{1}{\lambda I_{DS}} \quad (7.22)$$

Quindi

$$\left. \frac{\partial i_{DS}}{\partial v_{DS}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS}}} = \frac{1}{r_o} \quad (7.23)$$

La resistenza di drain dipende dalla corrente di polarizzazione tra drain e source; per valori di tale corrente dell'ordine del mA, r_o è dell'ordine della decina di k Ω .

Sostituendo la (7.20) e la (7.23) nella (7.16), si ottiene l'equazione del modello linearizzato del FET

$$i_{ds} = g_m v_{gs} + \frac{v_{ds}}{r_o} \quad (7.24)$$

Questa equazione è equivalente al circuito in Figura 7.15. Tra gate e source c'è un circuito aperto e quindi non può scorrere corrente. La corrente i_{ds} è determinata da due contributi, un generatore di corrente controllato dalla v_{gs} e un contributo proporzionale alla v_{ds} quindi equivalente a una resistenza in parallelo tra drain e source.

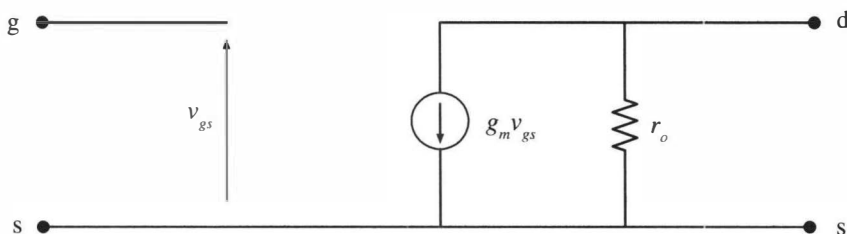


Figura 7.15 Modello linearizzato del MOSFET

Il modello linearizzato ci fornisce un modo immediato per stimare il massimo guadagno in tensione ottenibile con un FET. Supponiamo di poter applicare la tensione del segnale direttamente tra il terminale g e il terminale s . Supponiamo inoltre di prelevare la tensione di uscita tra il terminale d e il terminale s con un carico infinito. In queste condizioni ideali il guadagno in tensione è dato da

$$A_{vv} = \frac{v_{gs}}{v_{ds}} = -g_m r_o \quad (7.25)$$

Il prodotto $g_m r_o$ (talvolta indicato con la lettera μ) è il massimo guadagno di tensione teoricamente ottenibile dal transistor.

Affinché il modello linearizzato sia applicabile è necessario che lo sviluppo di Taylor della funzione di trasferimento del FET in saturazione sia arrestabile al primo ordine. Calcoliamo quando tale condizione è verificata.

La corrente i_{DS} ha un andamento parabolico al variare di v_{GS} il cui sviluppo è

$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN})^2 = \frac{K_n}{2} (V_{GS} + v_{gs} - V_{TN})^2 \quad (7.26)$$

ovvero

$$i_{DS} = \frac{K_n}{2} (V_{GS} - V_{TN})^2 \left(1 + \frac{v_{gs}}{V_{GS} - V_{TN}} \right)^2 \quad (7.27)$$

da cui

$$i_{DS} = \frac{K_n}{2} (V_{GS} - V_{TN})^2 \left(1 + 2 \left(\frac{v_{gs}}{V_{GS} - V_{TN}} \right) + \left(\frac{v_{gs}}{V_{GS} - V_{TN}} \right)^2 \right) \quad (7.28)$$

La condizione per applicare il modello lineare è che il termine di secondo ordine sia molto minore del termine di primo ordine:

$$\left(\frac{v_{gs}}{V_{GS} - V_{TN}} \right)^2 \ll 2 \left(\frac{v_{gs}}{V_{GS} - V_{TN}} \right) \quad (7.29)$$

da cui

$$v_{gs} \ll 2(V_{GS} - V_{TN}) \quad (7.30)$$

Si noti che tale condizione dipende da V_{GS} che è fissata dalla rete di polarizzazione e da V_{TN} che è la soglia del particolare MOSFET utilizzato. Per quantificare meglio la (7.30) se

$$v_{gs} < 2(V_{GS} - V_{TN})/10 \quad (7.31)$$

il termine di secondo ordine è, in potenza, 100 volte minore del primo. In scala logaritmica è -20dB ($= 20 \log_{10}(1/10)$). Questa è un'assunzione accettabile nella maggior parte dei casi.

In conclusione il modello linearizzato del MOSFET a canale n ad arricchimento può essere riassunto come in Figura 7.16.

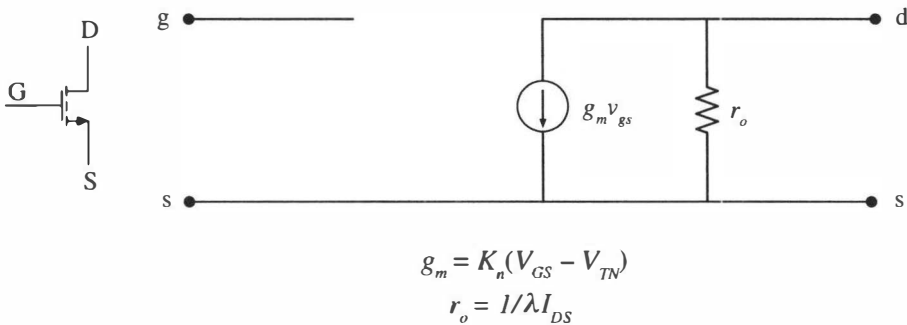


Figura 7.16 Modello linearizzato del MOSFET a canale n ad arricchimento

Se ora consideriamo il MOSFET a svuotamento, il fatto che la tensione V_{GS} di polarizzazione sia negativa non ha alcun effetto sul modello che si riferisce solo ai termini variabili. Pertanto il modello del MOSFET a svuotamento è del tutto identico a quello ad arricchimento.

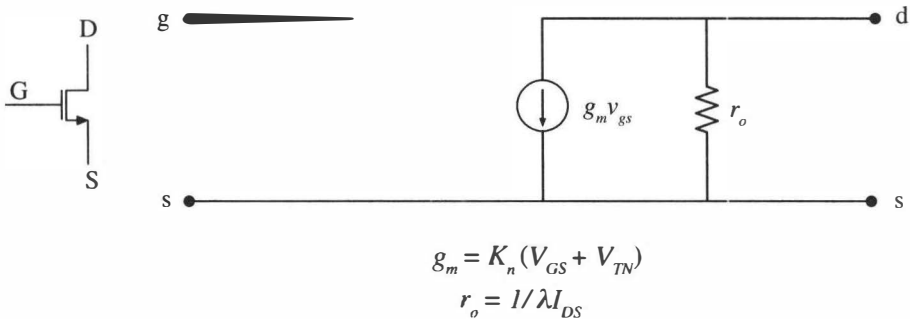


Figura 7.17 Modello linearizzato del MOSFET a canale n a svuotamento

Vediamo ora il caso del MOSFET ad arricchimento a canale p. Ragionando in analogia con quello a canale n, si ottiene il modello in Figura 7.18. Ma osser-

vandolo con attenzione, un po' sorprendentemente risulta del tutto identico al modello del MOSFET a canale n.

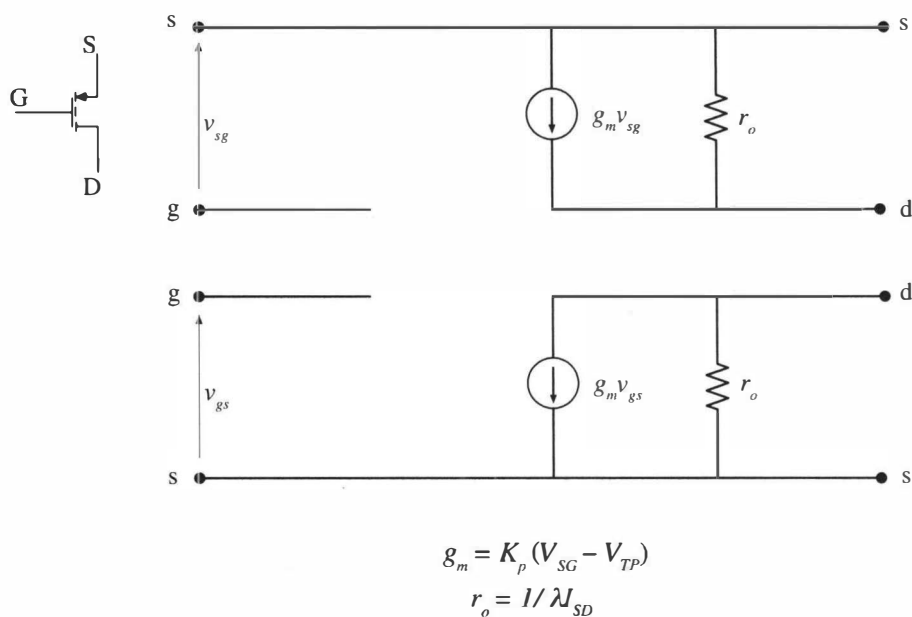


Figura 7.18 Modello linearizzato del MOSFET a canale p ad arricchimento

Per completezza, in Figura 7.19 è mostrato il modello linearizzato del MOSFET a canale p a svuotamento.

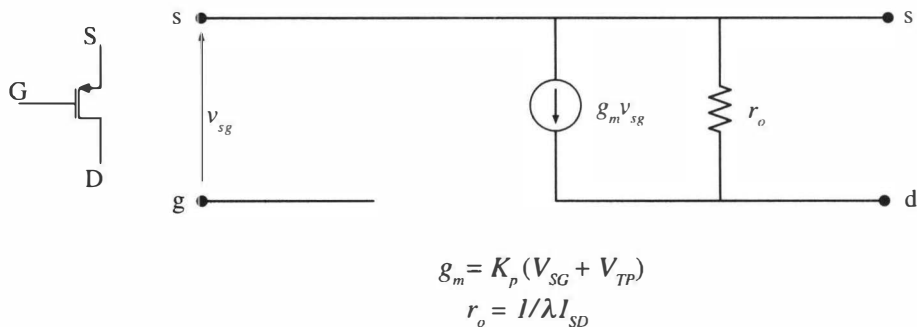


Figura 7.19 Modello linearizzato del MOSFET a canale p a svuotamento

*7.6 Modello linearizzato del FET a quattro terminali

Vogliamo ora ricavare il modello linearizzato del MOSFET a quattro terminali. Consideriamo il caso del MOSFET a canale n ad arricchimento. A causa dell'effetto body la corrente che scorre tra drain e source ora è funzione di tre variabili v_{GS} , v_{DS} e v_{SB} .

$$i_{DS} = i_{DS}(v_{GS}, v_{DS}, v_{SB}) \quad (7.32)$$

Le derivate rispetto alle prime due variabili sono le stesse già calcolate nel paragrafo precedente. Rimane da calcolare la derivata rispetto a v_{SB} . Le espressioni che ci servono sono

$$i_{DS} = \frac{K_n}{2} (v_{GS} - V_{TN})^2 \quad (7.33)$$

$$V_{TN} = V_{TN0} + \gamma \left(\sqrt{2\varphi_F + v_{SB}} - \sqrt{2\varphi_F} \right) \quad (7.34)$$

Quindi

$$\left. \frac{\partial i_{DS}}{\partial v_{SB}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS} \\ v_{SB} = V_{SB}}} = -K_n (v_{GS} - V_{TN}) \left. \frac{\partial V_{TN}}{\partial v_{SB}} \right|_{\substack{v_{GS} = V_{GS} \\ v_{DS} = V_{DS} \\ v_{SB} = V_{SB}}} = -g_m \frac{\gamma}{2\sqrt{2\varphi_F + V_{SB}}} \quad (7.35)$$

Il modello linearizzato che ne risulta sarà quindi quello in Figura 7.20.

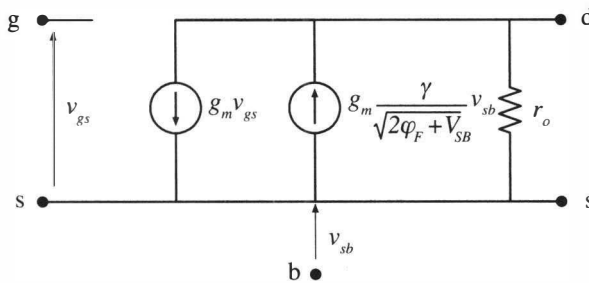


Figura 7.20 Modello linearizzato del MOSFET a canale n ad arricchimento a 4 terminali

In definitiva, l'effetto body si concretizza in un generatore di corrente che tende a ridurre la corrente erogata dal MOSFET di una piccola quantità proporzionale a v_{sb} .

Il modello linearizzato a canale p è mostrato in Figura 7.21, l'unica differenza è che nell'argomento della radice è presente la tensione V_{BS} invece della V_{SB} .

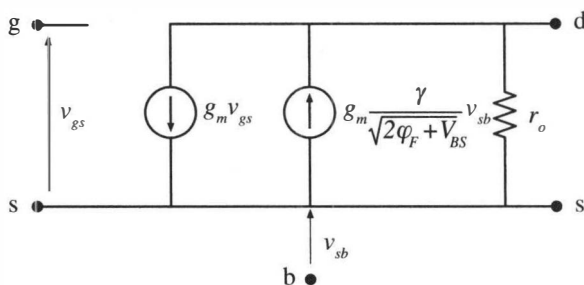


Figura 7.21 Modello linearizzato del MOSFET a canale p ad arricchimento a 4 terminali

Il modelli linearizzati dei MOSFET a svuotamento sono identici ai corrispondenti MOSFET ad arricchimento. La presenza dello strato sepolto sposta solo il punto di polarizzazione, ma non altera il comportamento dinamico.

7.7 Resistenze viste ai terminali del MOSFET

Prendiamo in considerazione un MOSFET polarizzato in saturazione. Per fissare le idee, possiamo riferirci alla rete a quattro resistori in Figura 7.22, ma tutte le considerazioni che faremo nel seguito sono generali e valgono per qualunque altra rete di polarizzazione.

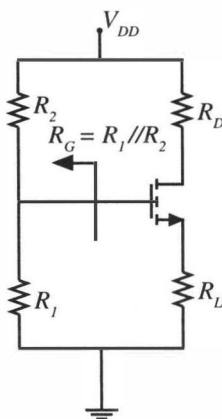


Figura 7.22 Rete di polarizzazione

Cominciamo calcolando la resistenza vista “guardando” dentro il gate. Poiché il gate è un circuito aperto, tale resistenza è infinita. Infatti, con riferimento alla Figura 7.23, un generatore di tensione v_x di prova eroga una corrente i_x nulla (vedi Appendice B per un richiamo al concetto di resistenza vista a due nodi di una rete lineare).

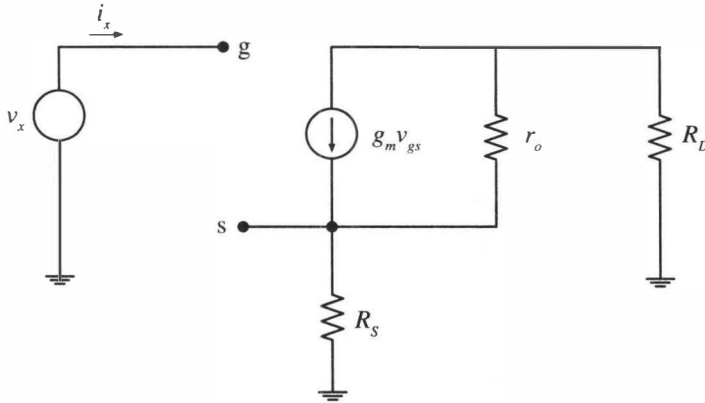


Figura 7.23 Calcolo della resistenza vista al terminale di gate

Le resistenze R_S e R_D che compaiono in Figura 7.23 dipendono dalla rete di polarizzazione.

Per calcolare la resistenza vista “guardando” dentro il drain facciamo riferimento alla Figura 7.24.

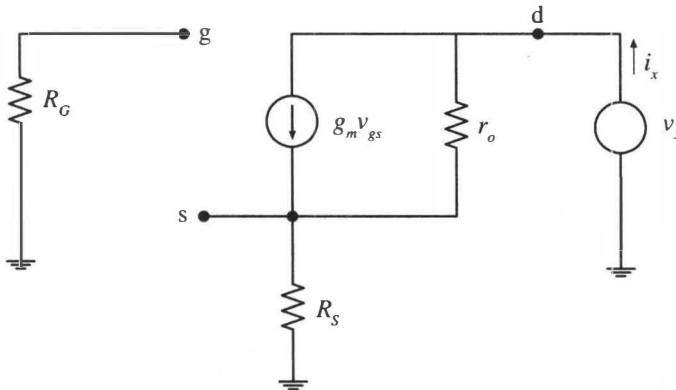


Figura 7.24 Calcolo della resistenza vista dal terminale di drain

Le resistenze R_G e R_S dipendono dalla rete di polarizzazione. Poiché su R_G non scorre corrente, il nodo g è a massa come in Figura 7.25.

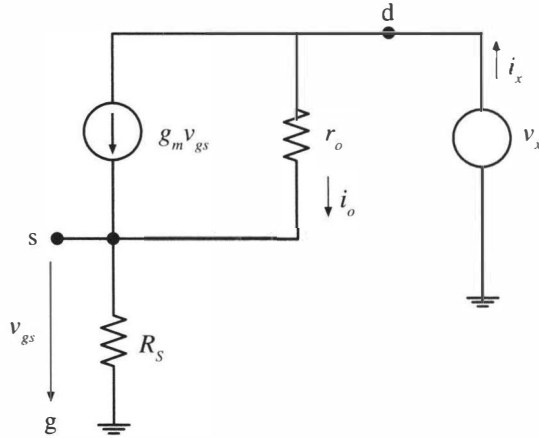


Figura 7.25 Calcolo della resistenza vista dal terminale di drain

Dalla circuitazione della maglia di uscita si ottiene

$$v_x = i_o r_o + i_s R_S \quad (7.36)$$

$$v_x = (i_s - g_m v_{gs}) r_o + i_s R_S \quad (7.37)$$

$$v_x = (i_s + g_m R_S i_s) r_o + i_s R_S \quad (7.38)$$

$$v_x = i_s \left((1 + g_m R_S) r_o + R_S \right) \quad (7.39)$$

D'altra parte $i_s = i_x$, quindi

$$R_x = \frac{v_x}{i_x} = (1 + g_m R_S) r_o + R_S \quad (7.40)$$

La resistenza vista “guardando” nel drain dipende dal guadagno del MOSFET e da R_S , ma è comunque un valore molto grande. Il valore minimo si ha per $R_S = 0$ ed è $R_x = r_o$, che è una resistenza dell'ordine di alcune decine di $k\Omega$.

Passiamo ora all'ultimo terminale: il source. Per calcolare la resistenza che si vede “guardando” nel source si consideri il circuito in Figura 7.26:

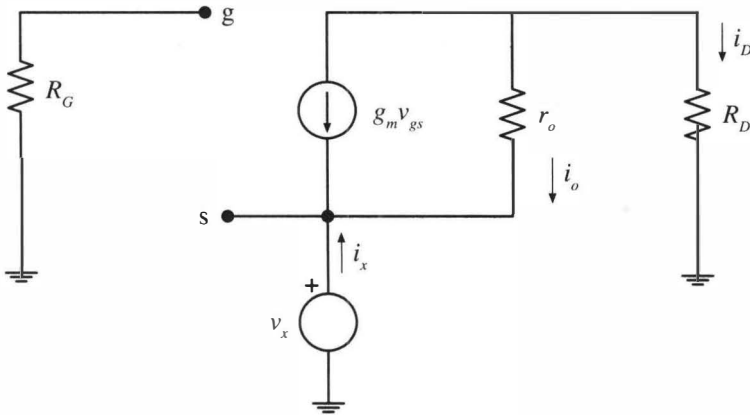


Figura 7.26 Calcolo della resistenza vista dal terminale di source

Poiché la corrente che scorre su R_G è nulla, il nodo di gate è a massa e il circuito può essere ridisegnato come in Figura 7.27:

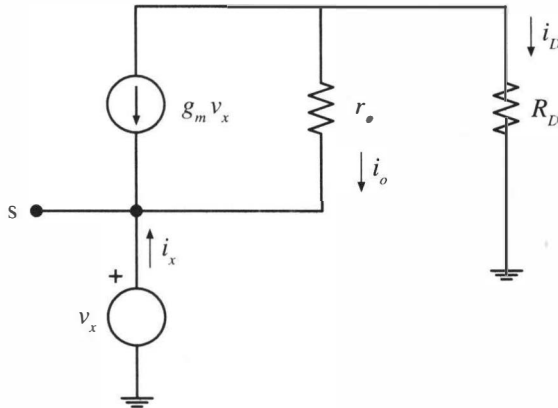


Figura 7.27 Calcolo della resistenza vista dal terminale di source

Dalla circuitazione della maglia di uscita si ottiene

$$v_x = -r_o i_o + R_D i_D \quad (7.41)$$

$$v_x = -r_o (i_x - g_m v_x) + R_D i_D \quad (7.42)$$

D'altra parte $i_D = i_x$, quindi

$$v_x = -r_o i_x + g_m v_x r_o + R_D i_x \quad (7.43)$$

da cui

$$R_x = \frac{v_x}{i_x} = \frac{R_D - r_o}{1 - g_m r_o} \approx \frac{1}{g_m} \quad (7.44)$$

Ma r_o è una resistenza piuttosto grande e di solito molto più grande di R_D . Inoltre $g_m r_o$ è il massimo guadagno in tensione ottenibile dal MOSFET ed è tipicamente molto più grande di 1. Pertanto

$$R_x \approx \frac{1}{g_m} \quad (7.45)$$

L'inverso della transconduttanza è tipicamente una resistenza piuttosto piccola (da pochi Ω a qualche centinaia di Ω). La Figura 7.28 riassume graficamente quanto trovato in questo paragrafo. Le resistenze viste al gate e al drain sono molto grandi, la resistenza vista al source è piccola e pari all'inverso della transconduttanza.

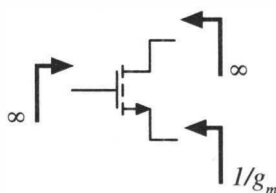


Figura 7.28 Resistenze viste ai tre terminali di un FET

Quanto ricavato sopra è valido per qualunque tipo di FET, non importa se MOSFET o JFET, a canale n o a canale p, ad arricchimento o a svuotamento. Per il MOSFET a quattro terminali, basta tenere presente che il terminale di body vede una giunzione pn contropolarizzata e quindi una grande resistenza.

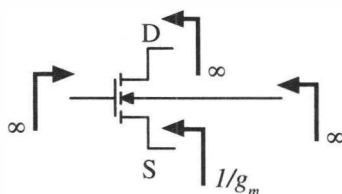


Figura 7.29 Resistenze viste ai terminali di un FET

7.8 Le tre configurazioni del FET

Un MOSFET ha tre terminali: gate, source e drain. Si può immettere o estrarre un segnale variabile (con frequenza minima sufficientemente alta) da ciascun terminale mediante un capacitore detto **capacitore di accoppiamento**. Il terminale da cui non si immette o estrae il segnale è collegato a massa (messo "a comune") con un capacitore detto **capacitore di by-pass**. Ovviamente non ha senso immettere ed estrarre il segnale dallo stesso terminale, quindi le possibili configurazioni sono sei.

	G	S	D
1	IN	C	OUT
2	OUT	C	IN
3	C	IN	OUT
4	C	OUT	IN
5	IN	OUT	C
6	OUT	IN	C

Tabella 7.1

Tuttavia non tutte le configurazioni hanno utilità pratica. Infatti ha poco senso immettere il segnale dal terminale di drain, in quanto in prima approssimazione la corrente di drain non dipende dalla tensione ad esso applicata. Un MOSFET ideale (senza modulazione del canale) è del tutto insensibile a un generatore di segnale applicato al drain. Ha ugualmente poco senso estrarre segnale dal gate, perché il gate non può fornire corrente e, inoltre, la variazione di tensione è piccola rispetto alle tensioni agli altri terminali. Delle sei configurazioni ne rimangono, quindi, solo tre: CS (source comune), CD (drain comune), CG (gate comune). La denominazione specifica solo quale dei tre terminali è a comune, ovvero non è usato per immettere o estrarre il segnale. Per identificare quale degli altri due è usato per immettere il segnale e quale per estrarlo basta ricordare quanto detto sopra, ovvero:

- 1) dal gate non si estrae il segnale
- 2) nel drain non si immette il segnale.

	G	S	D
CS	IN	C	OUT
CG	C	IN	OUT
CD	IN	OUT	C

Tabella 7.2

Vediamo quindi in dettaglio le tre configurazioni. La figura Figura 7.30 mostra un MOSFET in configurazione CS. Il source è a comune, l'ingresso è sul gate e l'uscita sul drain. Poiché nel paragrafo precedente abbiamo visto che la resistenza vista nel gate è molto grande, praticamente infinita, l'ingresso non può che essere in tensione. Anche la resistenza di drain è molto grande, ma il drain è un terminale di uscita e quindi l'uscita è in corrente.

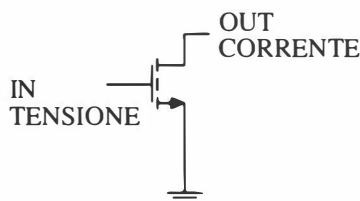


Figura 7.30 CS

Ragionando in maniera analoga con le altre configurazioni si ottiene che nella configurazione CD si ha tensione in ingresso e tensione in uscita, nella configurazione CG si ha corrente in ingresso e corrente in uscita (Figure 7.31 e 7.32).

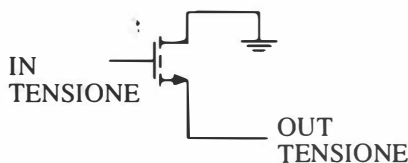


Figura 7.31 CD

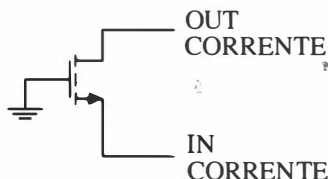


Figura 7.32 CG

Vogliamo ora calcolare i guadagni (ovvero i rapporti tra segnale di uscita e segnale di ingresso) nei tre casi.

In Figura 7.33 è mostrata la configurazione CS come in Figura 7.30, dove il transistor è stato sostituito con il suo modello linearizzato. La resistenza di uscita del generatore di segnale è omessa in quanto molto più piccola della resistenza tra g e s (che è praticamente infinita).

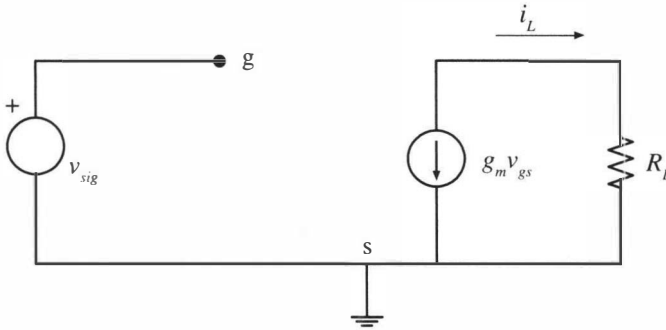


Figura 7.33 Guadagno in configurazione CS

Dal circuito risulta evidente che

$$A_{vt} = \frac{i_L}{v_{sig}} = \frac{-g_m v_{gs}}{v_{gs}} = -g_m \quad (7.46)$$

Vediamo ora la configurazione CD. La Figura 7.34 mostra il circuito da analizzare.

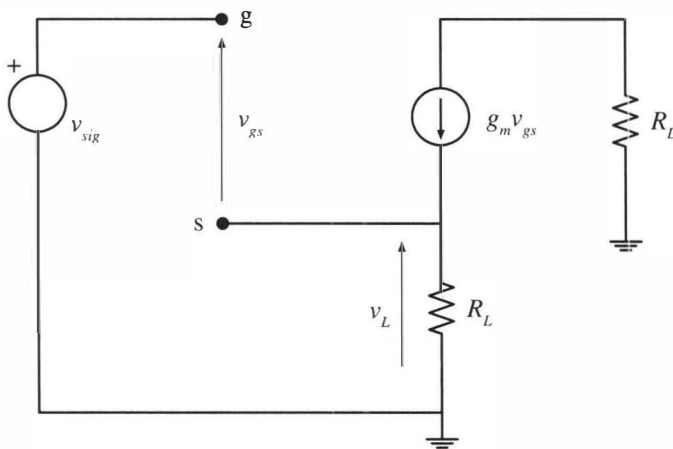


Figura 7.34 Guadagno in configurazione CD

Dal circuito si ottiene

$$v_L = R_L g_m v_{gs} = R_L g_m (v_{sig} - v_L) \quad (7.47)$$

da cui

$$A_{vv} = \frac{v_L}{v_{sig}} = \frac{R_L g_m}{1 + R_L g_m} \approx 1 \quad (7.48)$$

Il guadagno è quindi unitario. La configurazione CD non amplifica il segnale ma realizza un **buffer di tensione**, una sorta di stadio “cuscinetto” (buffer) tra due amplificatori in cascata con valori di resistenza di uscita e di ingresso non idonei a un trasferimento di tensione. Il buffer di tensione è uno degli amplificatori più usati.

Vogliamo chiarire questo concetto con un esempio. Supponiamo di voler trasferire tensione da uno stadio con resistenza di uscita $R_{sig} = 1\text{k}\Omega$ a un carico di resistenza $R_L = 500\Omega$. Il trasferimento di tensione è pari a

$$\frac{v_L}{v_{sig}} = \frac{R_L}{R_s + R_L} = 0.33 \quad (7.49)$$

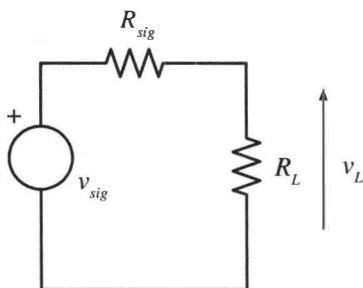


Figura 7.35 Trasferimento di segnale a un carico

Se ora interponiamo tra i due un buffer di tensione con resistenza di ingresso $R_{IN} = 10\text{k}\Omega$ e resistenza di uscita $R_{OUT} = 50\Omega$, il trasferimento di tensione è pari a

$$\frac{v_L}{v_{sig}} = \frac{v_L}{v_1} \frac{v_1}{v_{sig}} = \frac{R_{IN}}{R_{IN} + R_{sig}} \frac{R_L}{R_L + R_{OUT}} = 0.83 \quad (7.50)$$

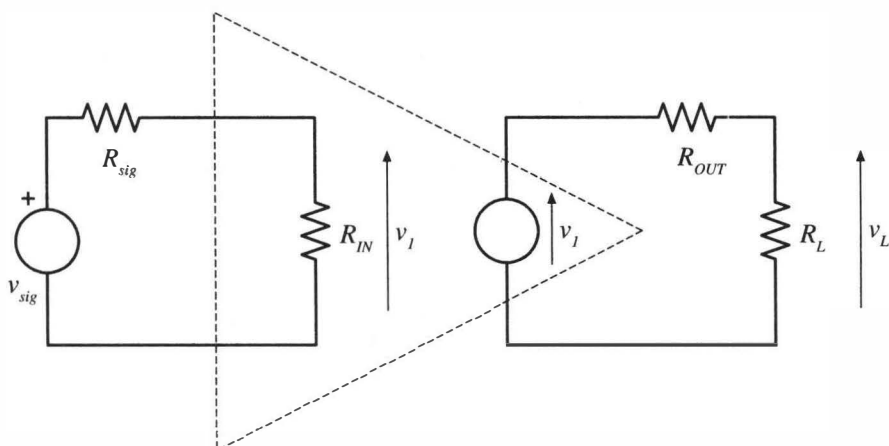


Figura 7.36 Trasferimento di tensione mediante buffer

Vediamo ora la configurazione CG. La Figura 7.37 mostra il circuito da analizzare, nel quale si considerano infinite la resistenza del generatore di segnale (in corrente) e la resistenza r_o del transistor.

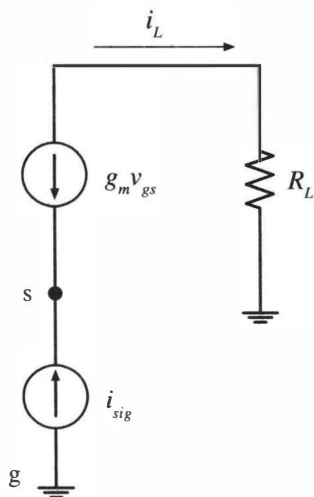


Figura 7.37 Guadagno in configurazione CG

Con queste approssimazioni, di fatto il circuito si riduce a una singola maglia e quindi

$$A_{ii} = \frac{i_L}{i_{sig}} = 1 \quad (7.51)$$

Si tratta, quindi, di un **buffer di corrente** di guadagno unitario.

Il transistor consente di realizzare direttamente un **amplificatore di transconduttanza** (ovvero tensione in ingresso e corrente in uscita), un buffer di tensione e un buffer di corrente. In numerose applicazioni però è più utile un amplificatore tensione-tensione. L'unico modo per realizzare un'amplificazione di tensione con un transistor è trasformare la corrente in uscita in una tensione dissipandola in una resistenza. La Figura 7.38 mostra un amplificatore CS con uscita in tensione.

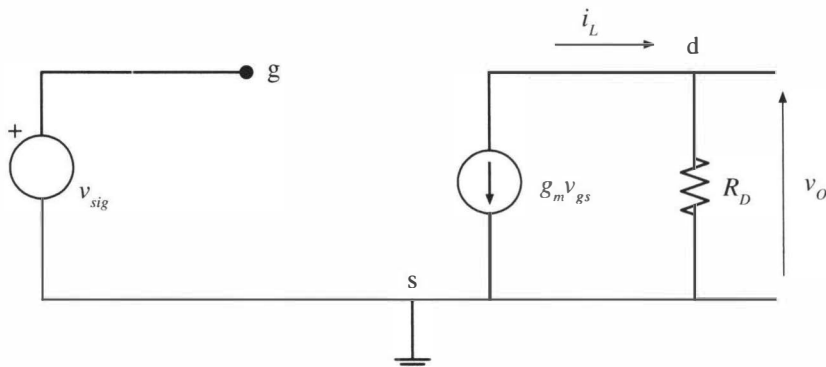


Figura 7.38 CS

Dal circuito risulta evidente che

$$A_{vv} = \frac{v_O}{v_{sig}} = -g_m R_D \quad (7.52)$$

Si noti che si tratta di un guadagno negativo, nel senso che quando la tensione di ingresso aumenta, quella di uscita diminuisce, ovvero il segnale in uscita è in **controfase** rispetto all'ingresso. La resistenza di uscita di questo amplificatore è pari a R_D . Un buon amplificatore con uscita in tensione dovrebbe avere resistenza di uscita piccola (per massimizzare il trasferimento di tensione) tuttavia R_D non può essere piccola se l'amplificatore deve avere un guadagno consistente. Questo problema si risolve con un amplificatore a due stadi. Un buffer di tensione in cascata allo stadio di amplificazione fa sì che la resistenza di uscita sia più adatta a un buon trasferimento di tensione.

Le figure successive riassumono graficamente il contenuto di questo paragrafo.

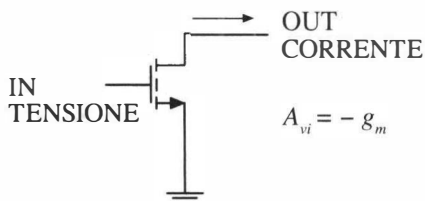


Figura 7.39 CS con uscita in corrente

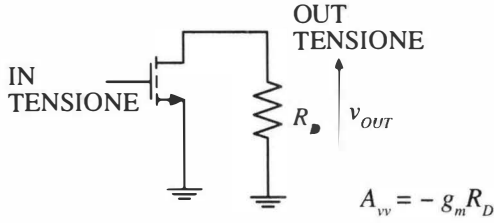


Figura 7.40 CS con uscita in tensione

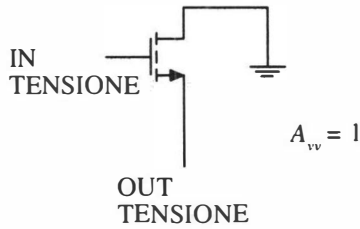


Figura 7.41 CD

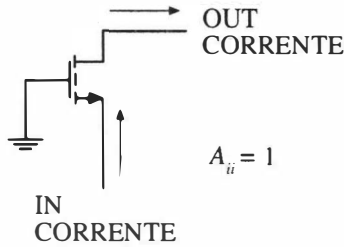


Figura 7.42 CG

7.9 Modello linearizzato del BJT

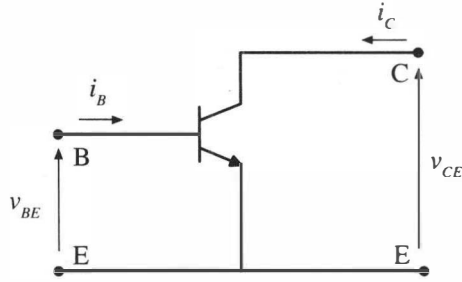
Per un BJT valgono considerazioni analoghe a quelle fatte sopra per il MOSFET. Cominciamo, quindi, con il modello linearizzato del BJT npn. Prima di procedere definiamo chiaramente i simboli impiegati secondo la convenzione già usata per il MOSFET. Con riferimento alle grandezze in Figura 7.43 possiamo scrivere:

$$i_C = I_C + i_c \quad (7.53)$$

$$i_B = I_B + i_b \quad (7.54)$$

$$v_{CE} = V_{CE} + v_{ce} \quad (7.55)$$

$$v_{BE} = V_{BE} + v_{be} \quad (7.56)$$

**Figura 7.43** BJT

Scriviamo:

$$i_C = i_C(v_{BE}, v_{CE}) \quad (7.57)$$

$$i_B = i_B(v_{BE}, v_{CE}) \quad (7.58)$$

Sviluppando secondo Taylor e arrestando lo sviluppo al primo ordine si ottiene:

$$i_C = \left. \frac{\partial i_C}{\partial v_{BE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} v_{be} + \left. \frac{\partial i_C}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} v_{ce} \quad (7.59)$$

$$i_b = \left. \frac{\partial i_B}{\partial v_{BE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} v_{be} + \left. \frac{\partial i_B}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} v_{ce} \quad (7.60)$$

La (7.59) e la (7.60) sono il modello linearizzato per piccoli segnali del BJT. Restano da calcolare le quattro derivate.

Cominciamo con il calcolo della derivata di i_C rispetto a v_{BE} . Nel Capitolo 5 abbiamo visto che in regione attiva diretta la relazione tra i_C e v_{BE} è essenzialmente esponenziale, ovvero:

$$i_C = I_T e^{\frac{v_{BE}}{V_T}} \quad (7.61)$$

con

$$I_T = \frac{qSD_n}{w} \left(\frac{n_i^2}{N_B} \right) \quad (7.62)$$

dove q è la carica elementare dell'elettrone, S la sezione della regione di base, D_n il coefficiente di diffusione degli elettroni, w lo spessore della base, n_i la concentrazione degli elettroni nel silicio intrinseco e N_B la concentrazione delle impurità in base. A rigore lo spessore della base, w , varia con v_{BE} , in quanto all'aumentare della v_{BE} si riduce la regione di svuotamento della giunzione base emettitore e quindi aumenta lo spessore della base. Tuttavia, essendo la giunzione base emettitore polarizzata direttamente, la regione di svuotamento è piccola e varia poco con la tensione. Quindi nel calcolo della derivata di i_C rispetto a v_{BE} tale dipendenza può essere trascurata.

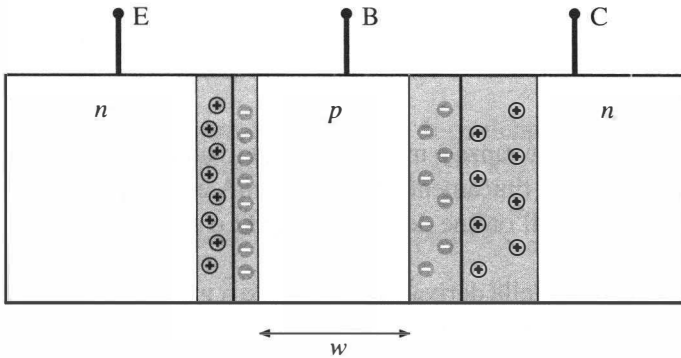


Figura 7.44 BJT npn

Procediamo al calcolo:

$$\left. \frac{\partial i_C}{\partial v_{BE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \left. \frac{\partial}{\partial v_{BE}} I_T e^{\frac{v_{BE}}{V_T}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{I_T e^{\frac{V_{BE}}{V_T}}}{V_T} = \frac{I_C}{V_T} \quad (7.63)$$

Definiamo la **transconduttanza del BJT** come segue:

$$g_m = \frac{I_C}{V_T} \quad (7.64)$$

Quindi

$$\left. \frac{\partial i_C}{\partial v_{BE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = g_m \quad (7.65)$$

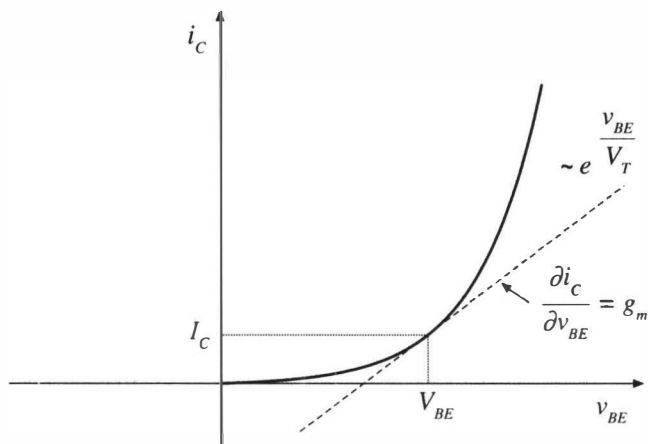


Figura 7.45 Transconduttanza

L'inverso della transconduttanza è una resistenza talvolta indicata come r_e . Poiché la corrente di collettore è tipicamente di qualche mA e V_T è 26 mV, segue che r_e è una resistenza piccola dell'ordine della decina di Ω o anche meno.

Passiamo ora al calcolo della derivata di i_C rispetto a v_{CE} . Nel Capitolo 3 abbiamo visto che in regione attiva diretta i_C dipende essenzialmente da v_{BE} piuttosto che da v_{CE} . Tuttavia, fissata v_{BE} , non è del tutto da trascurare una debole dipendenza circa lineare di i_C da v_{CE} a causa dell'effetto Early.

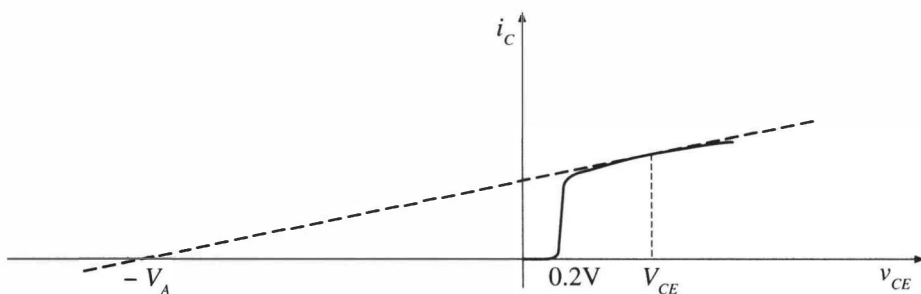


Figura 7.46 Effetto Early

Con riferimento alla Figura 7.46, la pendenza (ovvero la derivata) di tale dipendenza è

$$\left. \frac{\partial i_C}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{I_C}{V_{CE} + V_A} \quad (7.66)$$

Poiché V_A è dell'ordine di 50-100V e V_{CE} è di solito di qualche Volt, possiamo approssimare:

$$\left. \frac{\partial i_C}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{I_C}{V_A} \quad (7.67)$$

Definiamo **resistenza di collettore**

$$r_o = \frac{V_A}{I_C} \quad (7.68)$$

Quindi

$$\left. \frac{\partial i_C}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{1}{r_o} \quad (7.69)$$

Poiché la corrente di collettore è tipicamente di qualche mA, nei casi pratici la resistenza di collettore è dell'ordine delle decine di kΩ.

Passiamo ora al calcolo della derivata di i_B rispetto a v_{BE} . Ancora nel Capitolo 5 abbiamo visto che

$$i_B = I_{B0} e^{\frac{v_{BE}}{V_T}} \quad (7.70)$$

$$I_{B0} = \frac{qSw}{2\tau} \left(\frac{n_i^2}{N_B} \right) \quad (7.71)$$

con τ tempo di ricombinazione degli elettroni in base. Pertanto:

$$\left. \frac{\partial i_B}{\partial v_{BE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{\partial}{\partial v_{BE}} I_{B0} e^{\frac{v_{BE}}{V_T}} \bigg|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{I_{B0} e^{\frac{V_{BE}}{V_T}}}{V_T} = \frac{I_B}{V_T} = \frac{I_C}{\beta_F V_T} = \frac{g_m}{\beta_F} \quad (7.72)$$

Definiamo la **resistenza di base** come segue:

$$r_\pi = \frac{\beta_F}{g_m} \quad (7.73)$$

Quindi

$$\left. \frac{\partial i_B}{\partial v_{BE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{1}{r_\pi} \quad (7.74)$$

La resistenza di base può essere scritta anche come

$$r_{\pi} = \beta_F r_e \quad (7.75)$$

β_F è il guadagno di corrente ed è un numero puro tipicamente piuttosto grande (50-200), r_e è invece una resistenza piuttosto piccola dell'ordine della decina di Ω o anche meno. Pertanto r_{π} è una resistenza dell'ordine di qualche $k\Omega$.

Rimane ora da calcolare la derivata di i_B rispetto a v_{CE} . Dalla (7.71) risulta una dipendenza della corrente di base da v_{CE} , in quanto all'aumentare della v_{CE} aumenta la regione di svuotamento tra collettore e base e quindi diminuisce lo spessore utile della base w . Si tratta di una sorta di effetto Early relativo alla corrente di base invece che alla corrente di collettore. Nel calcolo della derivata di i_B rispetto a v_{CE} , l'unica grandezza che varia è w e possiamo scrivere

$$\left. \frac{\partial i_B}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{qS}{2\tau} \left(\frac{n_i^2}{N_B} \right) e^{\frac{V_{BE}}{V_T}} \left. \frac{\partial w}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} \quad (7.76)$$

D'altra parte per la corrente di collettore vale

$$i_C = \frac{qSD_n}{w} \left(\frac{n_i^2}{N_B} \right) e^{\frac{V_{BE}}{V_T}} \quad (7.77)$$

Quindi

$$\left. \frac{\partial i_C}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{qSD_n}{w^2} \left(\frac{n_i^2}{N_B} \right) e^{\frac{V_{BE}}{V_T}} \left. \frac{\partial w}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = -I_C \frac{1}{w_0} \left. \frac{\partial w}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} \quad (7.78)$$

dove w_0 è la larghezza della base nel punto di lavoro.

Ma avevamo già calcolato la derivata di i_C rispetto a v_{CE} come

$$\left. \frac{\partial i_C}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = \frac{I_C}{V_A} \quad (7.79)$$

Quindi

$$\left. \frac{\partial w}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = -\frac{w_0}{V_A} \quad (7.80)$$

che, sostituita nella (7.76), dà

$$\left. \frac{\partial i_B}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = -\frac{I_B}{V_A} \quad (7.81)$$

ovvero

$$\left. \frac{\partial i_B}{\partial v_{CE}} \right|_{\substack{v_{BE} = V_{BE} \\ v_{CE} = V_{CE}}} = - \frac{1}{\beta_F} \frac{I_C}{V_A} = - \frac{1}{\beta_F r_o} \quad (7.82)$$

β_F è il guadagno di corrente ed è un numero puro tipicamente piuttosto grande (50-200), r_o è la resistenza di collettore e ha un valore piuttosto grande dell'ordine della decina di k Ω . Quindi al denominatore c'è una resistenza dell'ordine del M Ω .

Sostituendo le (7.65), (7.69), (7.74) e (7.82) nelle (7.59) e (7.60), si ottengono le equazioni del modello per piccoli segnali del BJT

$$i_c = g_m v_{be} + \frac{v_{ce}}{r_o} \quad (7.83)$$

$$i_b = \frac{v_{be}}{r_\pi} - \frac{v_{ce}}{\beta_F r_o} \quad (7.84)$$

Il secondo termine della (7.84) può essere trascurato se

$$\frac{v_{ce}}{\beta_F r_o} \ll \frac{v_{be}}{r_\pi} \quad (7.85)$$

ovvero, facendo le opportune sostituzioni:

$$\frac{v_{ce}}{v_{be}} \ll \frac{V_A}{V_T} \quad (7.86)$$

Il primo termine di questa disuguaglianza corrisponde al rapporto tra la tensione del segnale in uscita e la tensione del segnale in ingresso: questo rapporto se il BJT è impiegato come amplificatore di tensione può non essere piccolo. Tuttavia il secondo termine è il rapporto tra la tensione di Early che è tipicamente 50-100V e la tensione termica (25mV), ovvero un valore superiore a 2000. La disuguaglianza (7.86) è quindi di solito ben verificata e la (7.84) può essere riscritta come

$$i_b = \frac{v_{be}}{r_\pi} \quad (7.87)$$

La (7.83) e la (7.87) descrivono un circuito come quello mostrato in Figura 7.47. Infatti, v_{be} e i_b sono legate dalla resistenza r_π tra base ed emettitore; la corrente di collettore è la somma di due contributi: un generatore di corrente controllato

dalla v_{be} e un contributo proporzionale alla v_{ce} quindi equivalente a una resistenza in parallelo tra collettore ed emettitore.

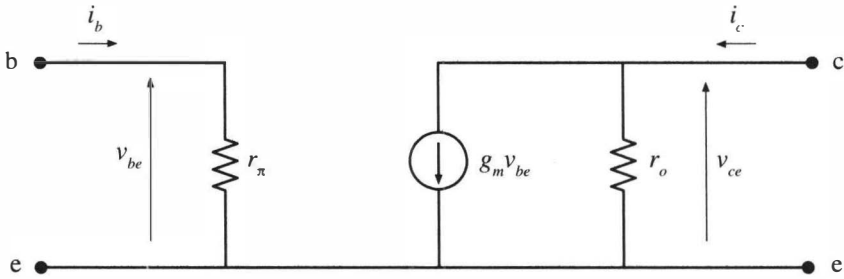


Figura 7.47 Modello linearizzato del BJT

Il generatore di corrente controllato dalla v_{be} può essere, in maniera equivalente, visto come un generatore controllato dalla corrente i_b , infatti:

$$g_m v_{be} = g_m r_\pi i_b = g_m \frac{\beta_F}{g_m} i_b = \beta_F i_b \quad (7.88)$$

Il modello linearizzato ci fornisce un modo immediato per stimare il massimo guadagno in tensione ottenibile con un BJT. Nel caso ideale la tensione del segnale si applica interamente tra il terminale b e il terminale e . La tensione di uscita si preleva tra il terminale c e il terminale e con un carico infinito.

Quindi

$$A_{vv} = \frac{v_{ce}}{v_{be}} = -g_m r_o = -\frac{V_A}{V_T} \quad (7.89)$$

che è un valore maggiore di 2000. Si noti però che per un guadagno di tensione così grande il secondo termine nella (7.84) non è più trascurabile, ovvero l'amplificatore non è del tutto direzionale dall'ingresso verso l'uscita.

Nel definire il modello per piccoli segnali, abbiamo inoltre assunto di poter arrestare lo sviluppo di Taylor della funzione di trasferimento del BJT in regione attiva diretta al primo ordine. Calcoliamo quando tale condizione è verificata.

La corrente di collettore ha un andamento esponenziale al variare di v_{BE} il cui sviluppo è

$$i_c = I_T e^{\frac{v_{BE}}{V_T}} = I_T e^{\frac{V_{BE} + v_{be}}{V_T}} = I_T e^{\frac{V_{BE}}{V_T}} \left(1 + \frac{v_{be}}{V_T} + \frac{1}{2} \left(\frac{v_{be}}{V_T} \right)^2 + \dots \right) \quad (7.90)$$

Per approssimare al primo ordine è necessario che il termine di secondo ordine (e quindi anche i successivi) sia molto minore del termine di primo ordine:

$$\frac{1}{2} \left(\frac{v_{be}}{V_T} \right)^2 \ll \frac{v_{be}}{V_T} \quad (7.91)$$

da cui

$$v_{be} \ll 2V_T \approx 50\text{mV} \quad (7.92)$$

Si noti che questa condizione è molto stringente. Significa che i segnali, affinché possano essere considerati piccoli, devono avere ampiezza di pochi mV. Sono di questo ordine i segnali provenienti da un'antenna o da un sensore, ma non i segnali interni a un circuito elettronico. Vedremo nel seguito come poter fare operare i BJT in regime di piccoli segnali anche con segnali decisamente più grandi.

7.10 Resistenze viste ai tre terminali del BJT

Il calcolo delle resistenze del BJT procede in modo del tutto analogo a quello per il FET, l'unica differenza è che la resistenza tra base ed emettitore non è infinita, ma pari a r_π . Il calcolo è solo leggermente più complesso. Il risultato è riassunto in Figura 7.48. R_E è la resistenza vista dal terminale di emettitore "guardando" verso l'esterno.

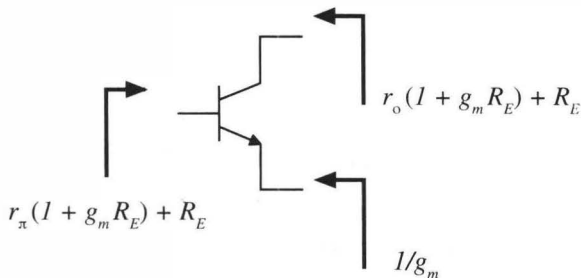


Figura 7.48 Resistenze viste dai terminali del BJT

Dal punto di vista pratico, utilizzando valori realistici per i parametri del BJT e del MOSFET, ci sono alcune significative differenze tra i due dispositivi di cui è importante tenere conto.

- 1) L'inverso della transconduttanza per un BJT è una resistenza molto piccola, tipicamente minore di 10Ω . Nel MOSFET tale valore è di solito sensibilmen-

- te più grande. Pertanto la resistenza vista “guardando” nell’emettitore del BJT è tipicamente molto più piccola di quella del FET.
- 2) La resistenza vista “guardando” nel gate è a tutti gli effetti grandissima, al contrario della resistenza in base, che non è certo piccola, ma neanche così grande da poter essere considerata un circuito aperto in ogni circostanza. In effetti la resistenza vista “guardando” in base dipende dalla presenza di una resistenza di emettitore R_E fissata dalla rete di polarizzazione. Il caso peggiore è quando $R_E = 0$. Per tale valore la resistenza in base è pari a r_π , che è una resistenza dell’ordine del $k\Omega$ e talvolta anche minore.
 - 3) Le resistenze viste “guardando” nel drain o nel collettore sono entrambe maggiori di qualche decina di $k\Omega$, quindi molto grandi.

Con le precisazioni che abbiamo detto sopra, anche per il BJT possiamo riassumere le resistenze viste ai tre terminali come in Figura 7.49. Sulla resistenza in base il segno “ $<$ ” sta a ricordare che è necessaria un po’ di cautela nel considerare questa resistenza infinita

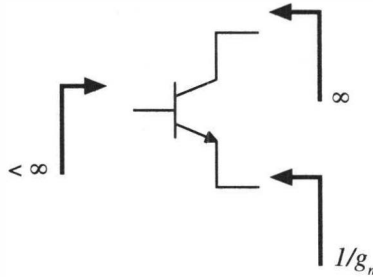


Figura 7.49 Resistenze viste ai terminali del BJT

7.11 Le tre configurazioni del BJT

Per il BJT valgono tutte le considerazioni già fatte per il FET. La Tabella 7.3 e i grafici successivi ne riassumono i risultati.

	B	E	C
CE	IN	C	OUT
CB	C	IN	OUT
CC	IN	OUT	C

Tabella 7.3

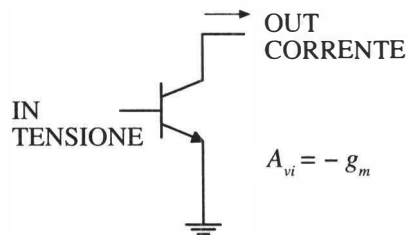


Figura 7.50 CE con uscita in corrente

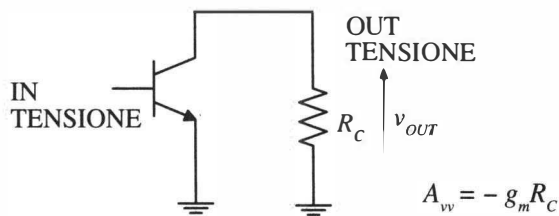


Figura 7.51 CE con uscita in tensione

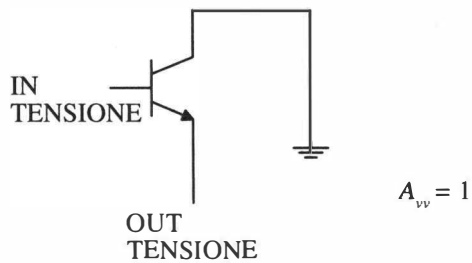


Figura 7.52 CC

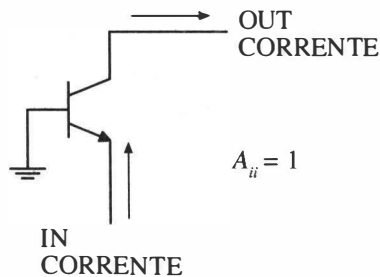


Figura 7.53 CB

7.12 Analisi e progetto

Le considerazioni teoriche svolte nei paragrafi precedenti sono preliminari all'analisi e alla progettazione degli amplificatori. Per **analisi** si intende il calcolo delle caratteristiche di un amplificatore il cui circuito è già stato progettato. Nello specifico in questo capitolo ci occuperemo di come calcolare la resistenza di ingresso, la resistenza di uscita, il guadagno e la dinamica. Nel capitolo successivo ci occuperemo delle caratteristiche relative alla risposta in frequenza. L'analisi di un circuito si esegue in due passi successivi: l'**analisi DC** (*direct current*) e l'**analisi AC** (*alternate current*). L'analisi DC è il calcolo delle correnti e delle tensioni in tutte le maglie, ovvero il calcolo del **punto di polarizzazione** o **punto di riposo**. La seconda fase è l'analisi AC, che consiste nel sostituire i transistor con i loro modelli linearizzati e calcolare, quindi, le componenti variabili delle correnti e delle tensioni nel circuito. Entrambe le analisi possono essere eseguite da un programma di simulazione. Il più noto è **SPICE**, di cui sono disponibili molte versioni anche scaricabili gratuitamente. Tuttavia, non è superfluo saper analizzare un circuito con "carta e penna" perché consente di capirne gli elementi critici e operare delle scelte consapevoli in fase di progetto. Nel seguito eseguiremo l'analisi "carta e penna" di numerosi circuiti.

Per progettazione si intende l'ideazione di un circuito sulla base di specifiche date. Un errore tipico del principiante è affrontare la progettazione come un esercizio di matematica che ammette una sola soluzione. Nella realtà le possibili soluzioni sono molteplici, talvolta ne esiste una ottimale, ma più spesso ci si accontenta di una soluzione solo soddisfacente. Non esistono delle **regole di progetto** di validità generale, ma ci sono criteri specifici per particolari configurazioni circuitali. Nel seguito mostreremo alcuni esempi di progettazione.

7.13 Analisi di un amplificatore di tensione a singolo MOSFET in configurazione CS

Come primo esempio di analisi consideriamo l'amplificatore a MOSFET in Figura 7.54. Il generatore di segnale in ingresso ha resistenza di uscita di 50Ω . Il carico in uscita è di $100k\Omega$. Le caratteristiche del MOSFET utilizzato sono $K_n = 20\text{mA/V}^2$ e $V_{TN} = 1\text{V}$. Vogliamo calcolare resistenze di ingresso e di uscita, guadagno e dinamica.

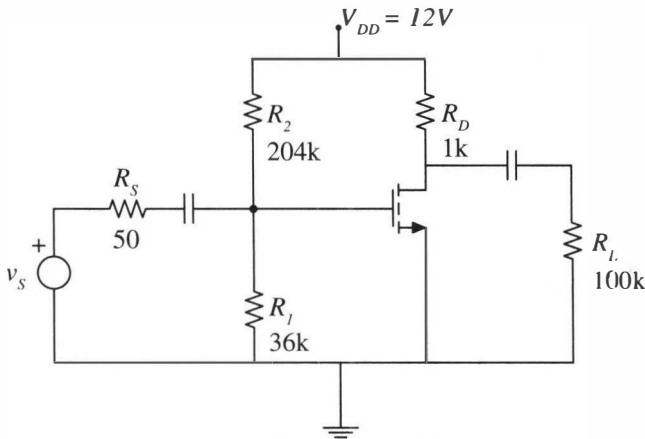


Figura 7.54 Amplificatore a MOSFET

Innanzitutto è necessario procedere all'analisi DC. Il partitore costituito dalle resistenze R_1 e R_2 fissa la tensione V_{GS} , che è quindi

$$V_{GS} = \frac{R_1}{R_1 + R_2} V_{DD} = 1.8 \text{ V} \quad (7.93)$$

La corrente i_{DS} è quindi pari a

$$i_{DS} = \frac{K_n}{2} (V_{GS} - V_{TN})^2 = 6.4 \text{ mA} \quad (7.94)$$

Calcoliamo ora la tensione al drain che risulta pari a

$$V_{DS} = V_{DD} - R_D i_{DS} = 5.6 \text{ V} \quad (7.95)$$

Verifichiamo che il MOSFET sia effettivamente in saturazione. La condizione è

$$V_{DS} > V_{GS} - V_{TN} \quad (7.96)$$

nello specifico

$$5.6 \text{ V} > 0.8 \text{ V} \quad (7.97)$$

che, quindi, è ampiamente rispettata.

Passiamo ora all'analisi AC. Il circuito per l'analisi AC è quello in Figura 7.55, dove i capacitori di accoppiamento sono sostituiti da cortocircuiti, e la tensione di alimentazione cortocircuitata a massa.

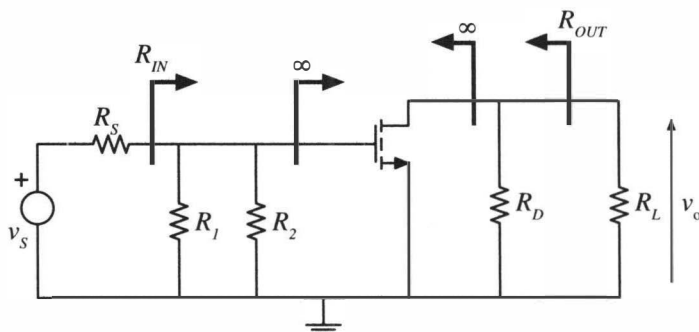


Figura 7.55 Circuito per l'analisi AC

Possiamo immediatamente ricavare che

$$R_{IN} = R_1 \parallel R_2 = 30.6 \text{ k}\Omega \quad (7.98)$$

$$R_{OUT} = R_D = 1 \text{ k}\Omega \quad (7.99)$$

Poiché

$$R_{IN} \gg R_s \quad (7.100)$$

$$R_{OUT} \ll R_L \quad (7.101)$$

l'ingresso è in tensione e l'uscita è in tensione. Il guadagno è quindi dato da

$$A_{vv} = -g_m R_D = -K_n (V_{GS} - V_{TN}) R_D = -16.0 \quad (7.102)$$

Calcoliamo ora la dinamica in ingresso. Affinché possa operare in regime lineare, l'ampiezza del segnale in ingresso v_s deve rispettare la seguente condizione:

$$v_s \ll 2(V_{GS} - V_{TN}) = 1.6 \text{ V} \quad (7.103)$$

Più quantitativamente un segnale di ampiezza 160mV (un decimo di 1.6V) produce termini del secondo ordine inferiori di -20dB.

Vediamo ora cosa limita la dinamica in uscita. Al variare del segnale in ingresso aumenta e diminuisce la corrente i_{DS} che scorre su R_D e ciò fa variare la v_{DS} . Ma la v_{DS} in alto è limitata dalla tensione di alimentazione ($V_{DD} = 12V$) e in basso dal punto in cui il MOSFET passa dalla saturazione alla regione lineare ($V_{GS} - V_{TN} = 0.8V$).

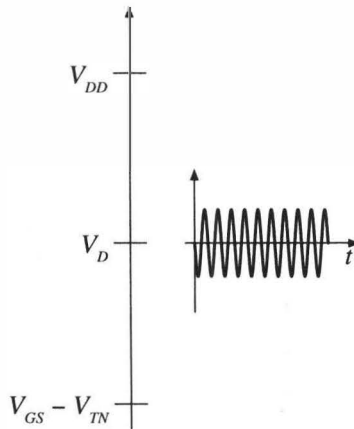


Figura 7.56 Dinamica in uscita

Possiamo ragionevolmente assumere che il segnale sia simmetrico, nel senso che la massima ampiezza positiva è pari alla massima ampiezza negativa, pertanto la dinamica di uscita è il valore minimo tra la massima escursione in alto e quella in basso. In questo caso è 4.8V. Questo valore di ampiezza in uscita corrisponde a 300mV (4.8V/16) in ingresso. Possiamo quindi concludere che la dinamica d'ingresso è 300mV, ma il valore del segnale per cui la distorsione armonica è inferiore a - 20dB è 160mV. La Tabella 7.4 riassume le caratteristiche dell'amplificatore.

R_{IN}	30.6k Ω
R_{OUT}	1 k Ω
A_{VV}	- 16.0
Dinamica	300mV
Distorsione - 20dB	160mV

Tabella 7.4

7.14 Progetto di un amplificatore di tensione a singolo MOSFET in configurazione CS

Vogliamo ora progettare un amplificatore di tensione a singolo MOSFET in configurazione CS sulla base delle seguenti specifiche: tensione di alimentazione $V_{DD} = 12\text{V}$; guadagno di tensione $A_{vv} = -10$; $R_{IN} > 10\text{k}\Omega$; $R_{OUT} < 1\text{k}\Omega$. Le caratteristiche del MOSFET da utilizzare sono $K_n = 20\text{mA/V}^2$ e $V_{TN} = 1\text{V}$.

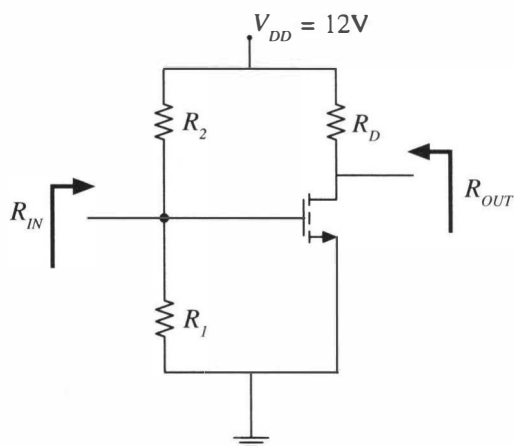


Figura 7.57 MOSFET in configurazione CS

La resistenza di uscita è

$$R_{OUT} = R_D \quad (7.104)$$

poiché R_{OUT} deve essere minore di $1\text{k}\Omega$, fissiamo $R_D = 500\Omega$.

Il guadagno è dato da

$$-g_m R_D = A_{vv} \quad (7.105)$$

$$K_n (V_{GS} - V_{TN}) R_D = |A_{vv}| \quad (7.106)$$

da cui il valore di $V_{GS} - V_{TN}$:

$$V_{GS} - V_{TN} = \frac{|A_{vv}|}{K_n R_D} = 1\text{V} \quad (7.107)$$

Per tale valore, i_{DS} risulta pari a

$$i_{DS} = \frac{K_n}{2} (V_{GS} - V_{TN})^2 = 10\text{mA} \quad (7.108)$$

I valori delle tensioni nelle maglie sono quindi quelli in Figura 7.58, che assicurano una dinamica di uscita di 5V. La dinamica non è tra le specifiche di progetto, ma quando è possibile è buona regola renderla più grande possibile.

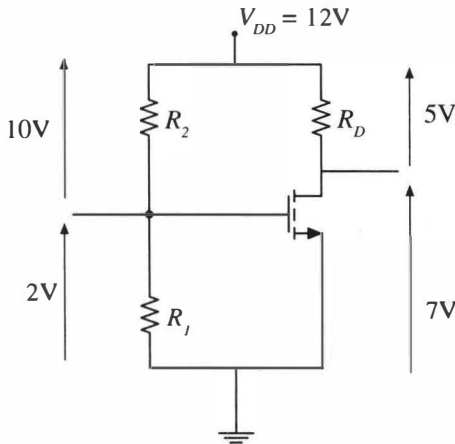


Figura 7.58 Tensioni alle maglie

Resta da dimensionare il partitore R_1 e R_2 in modo da fissare $V_{GS} = 2\text{V}$. Conviene scegliere un valore di corrente I_p che scorre nel partitore piuttosto piccolo, in modo da mantenere basso il consumo e tenere alta la resistenza di ingresso. Un valore ragionevole può essere $100\mu\text{A}$. Quindi

$$R_1 = \frac{V_{GS}}{I_p} = 20\text{k}\Omega \quad (7.109)$$

$$R_2 = \frac{V_{DD} - V_{GS}}{I_p} = 100\text{k}\Omega \quad (7.110)$$

La resistenza di ingresso è quindi

$$R_{IN} = R_1 \parallel R_2 = 16.6\text{k}\Omega > 10\text{k}\Omega \quad (7.111)$$

Il progetto ora è completo e la Figura 7.59 mostra i valori delle resistenze ottenuti:

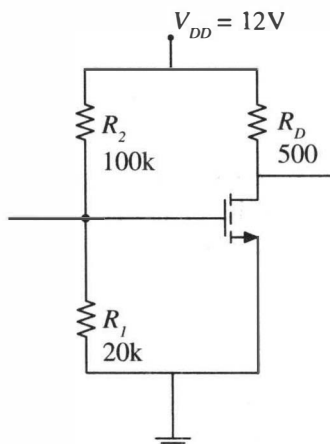


Figura 7.59 Dimensionamento delle resistenze

Si noti come non sia l'unica possibile soluzione: nel corso della progettazione abbiamo fatto alcune scelte, in certa misura arbitrarie, entro le specifiche imposte.

7.15 Progetto di un buffer di tensione a singolo MOSFET in configurazione CD

Vogliamo ora progettare un buffer di tensione a singolo MOSFET in configurazione CD sulla base delle seguenti specifiche: tensione di alimentazione $V_{DD} = 12\text{V}$; $R_{IN} > 10\text{k}\Omega$; $R_{OUT} < 100\Omega$. Le caratteristiche del MOSFET da utilizzare sono $K_n = 20\text{mA/V}^2$ e $V_{TN} = 1\text{V}$.

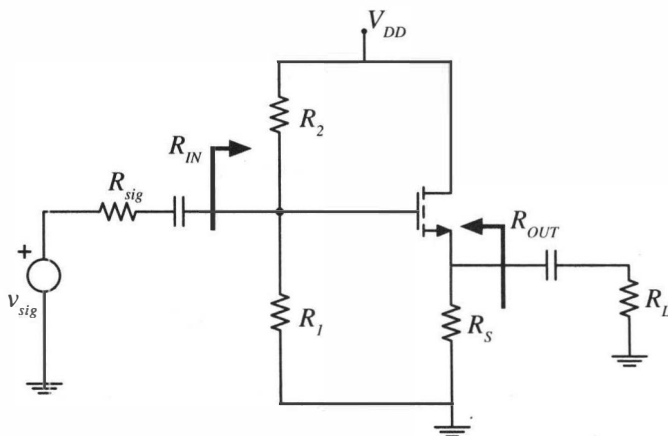


Figura 7.60 Buffer di tensione a MOSFET

La resistenza di uscita è pari a

$$R_{OUT} = \frac{1}{g_m} = \frac{1}{K_n (V_{GS} - V_{TN})} \quad (7.112)$$

Poniamo $R_{OUT} = 50\Omega$. Si ottiene $(V_{GS} - V_{TN}) = 1V$.

Dobbiamo ora fissare le tensioni nella maglia di uscita. Per massimizzare l'escursione del segnale in uscita conviene porre la tensione di source a riposo a circa metà tensione di alimentazione. Poniamo quindi $V_S = 6V$. Le tensioni sulle maglie risultano quindi quelle in Figura 7.61.

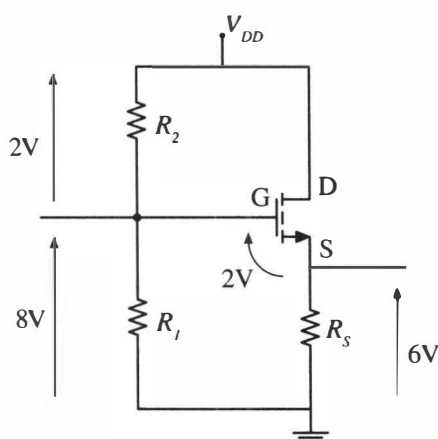


Figura 7.61 Tensioni alle maglie

Il progetto è completo, ma vogliamo ora calcolare la dinamica dell'amplificatore che abbiamo progettato.

La tensione al source può variare verso il basso di $V_S = 6V$ e verso l'alto di $V_{DD} - V_S - (V_{GS} - V_{TN}) = 5V$. La dinamica di uscita è il minore di questi due valori, quindi 5V, che, poiché il guadagno del buffer è unitario, corrisponde a circa 5V anche in ingresso.

Calcoliamo ora il limite sulla dinamica imposto dalla condizione di linearità. In generale per un MOSFET deve valere

$$v_{gs} \ll 2(V_{GS} - V_{TN}) \quad (7.113)$$

Dobbiamo quindi calcolare la relazione tra v_s (segnale in ingresso) e v_{gs} . A questo scopo si faccia riferimento al circuito AC in Figura 7.62:

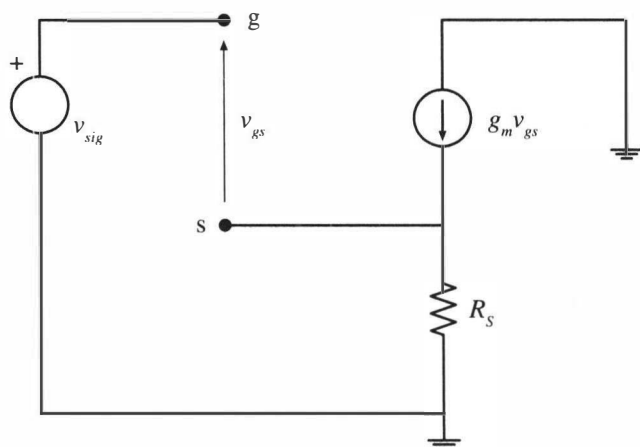


Figura 7.62 Calcolo della relazione tra v_{gs} e v_{sig}

Dalla circuitazione della maglia in ingresso si ottiene

$$v_{sig} = v_{gs} + R_S g_m v_{gs} \quad (7.114)$$

da cui

$$v_{gs} = \frac{v_{sig}}{1 + R_S g_m} \quad (7.115)$$

Sostituendo nella (7.113) si ottiene

$$v_{gs} \ll 2(V_{GS} - V_{TN})(1 + R_S g_m) \approx 2(V_{GS} - V_{TN})R_S g_m \quad (7.116)$$

$$v_{sig} \ll 4R_S I_{DS} = 4V_S = 24V \quad (7.117)$$

L'espressione appena trovata è molto utile e intuitiva: in presenza di R_S il limite di tensione da tenere come riferimento ai fini della linearità del transistor non è $2(V_{GS} - V_{TN})$, ma $4V_S$.

In conclusione il buffer progettato ha 5V di dinamica e 2.4V come punto di distorsione a -20dB.

7.16 Amplificatori reazionati

Nella pratica non sempre è necessario che un amplificatore abbia un elevato guadagno. Ad esempio in una catena di amplificazione è buona regola che solo il primo amplificatore sia ad alto guadagno; i successivi è bene che amplifichino molto meno per non saturare l'uscita, la cui escursione è sempre limitata dall'alimentazione. Un'ottima soluzione per ridurre il guadagno è l'impiego della reazione negativa. L'idea è prelevare parte del segnale in uscita e riportarlo in ingresso in controfase, in modo che all'aumentare dell'uscita diminuisca l'ingresso, riducendo quindi l'effettivo guadagno. Una resistenza R_S sul source agisce appunto come rete di reazione, nel senso che all'aumentare della corrente i_{DS} erogata dal transistor aumenta la caduta di tensione su R_S e quindi, a parità di v_{sig} , diminuisce la v_{gs} .

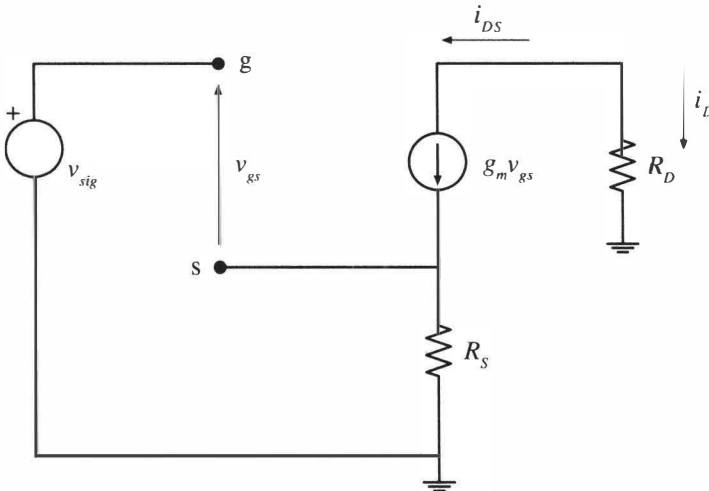


Figura 7.63 Circuito AC con reazione

Il guadagno tensione-corrente di un amplificatore reazionato si calcola mediante il circuito in Figura 7.63. Dalla circuitazione della maglia di ingresso:

$$v_{sig} = v_{gs} + R_S g_m v_{gs} \quad (7.118)$$

da cui

$$v_{gs} = \frac{v_{sig}}{1 + R_S g_m} \quad (7.119)$$

E quindi

$$i_{ds} = \frac{g_m v_{sig}}{1 + R_S g_m} \quad (7.120)$$

Se vale la condizione

$$R_S g_m \gg 1 \quad (7.121)$$

ovvero

$$2 R_S I_{DS} \gg V_{GS} - V_{TN} \quad (7.122)$$

allora il guadagno dell'amplificatore di transconduttanza è

$$A_{vi} = - \frac{1}{R_S} \quad (7.123)$$

che è un'espressione molto utile, perché non dipende dal particolare transistor utilizzato, ma solo dal valore di una resistenza. È quindi un valore ben determinato, che non risente della variabilità delle caratteristiche dei transistor.

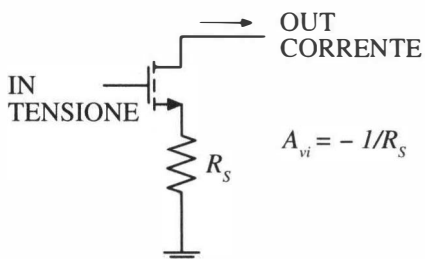


Figura 7.64 CS con reazione

Se l'amplificatore reazionato è utilizzato con uscita in tensione, il guadagno è dato da

$$A_{vv} = - \frac{R_D}{R_S} \quad (7.124)$$

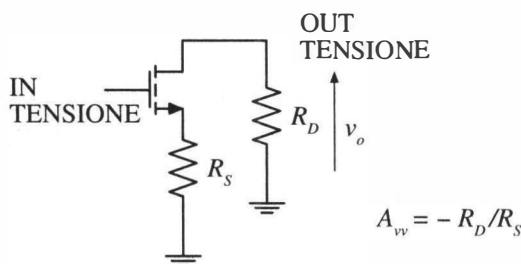


Figura 7.65 CS reazionato con uscita in tensione

Un vantaggio importante della reazione è che aumenta il valore del segnale in ingresso, per cui l'amplificatore può essere considerato lineare. La condizione di linearità è infatti

$$v_{gs} = \frac{v_{sig}}{1 + R_S g_m} \ll 2(V_{GS} - V_{TN}) \quad (7.125)$$

ovvero

$$v_{sig} \ll 2(1 + R_S g_m)(V_{GS} - V_{TN}) \approx 4R_S I_{DS} \quad (7.126)$$

Le considerazioni fatte sopra valgono anche per il BJT, per il quale

$$A_{vi} = -\frac{1}{R_E} \quad (7.127)$$

$$A_{vv} = -\frac{R_C}{R_E} \quad (7.128)$$

La condizione di linearità dell'amplificatore è

$$v_{sig} \ll (1 + R_E g_m) V_T \approx R_E I_C \quad (7.129)$$

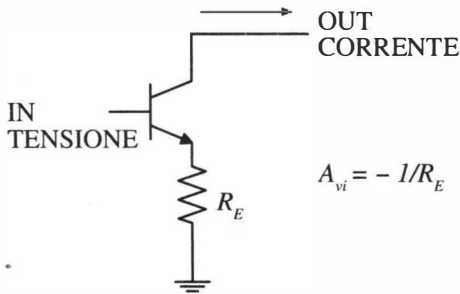


Figura 7.66 CE con reazione

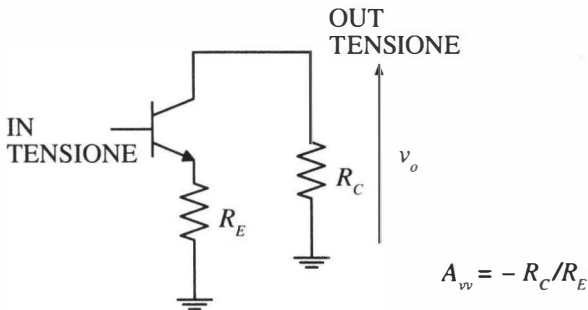


Figura 7.67 CE reazionato con uscita in tensione

La condizione affinché la reazione negativa possa operare efficacemente è

$$R_E g_m \gg 1 \quad (7.130)$$

che è equivalente alla (7.121). La (7.130) può essere riscritta anche come

$$I_C R_E \gg V_T \quad (7.131)$$

7.17 Progetto di un amplificatore di tensione reazionato CE

La reazione negativa è spesso utilizzata in combinazione con la configurazione CE. Infatti il BJT in configurazione CE permette di ottenere un grande guadagno ma ha alcuni importanti limiti:

- 1) Affinché possa operare in regime lineare, il segnale di ingresso deve essere molto minore di $2V_T = 50\text{mV}$, che è un valore piccolo.
- 2) La resistenza d'ingresso non è molto grande (tipicamente non è più di qualche $\text{k}\Omega$).
- 3) Il guadagno è proporzionale a $g_m = I_C/V_T$ con V_T tensione termica. Poiché la tensione termica dipende dalla temperatura, il guadagno non è costante. Per di più $g_m = I_C/V_T$ vale solo in prima approssimazione, più in generale è necessario considerare un fattore di non idealità a moltiplicare la tensione termica. Tale fattore dipende dalla corrente di collettore e dal particolare BJT. In definitiva il guadagno è grande, ma poco controllabile.

La reazione negativa mitiga questi problemi al prezzo di una riduzione del guadagno.

A titolo di esempio progettiamo un amplificatore CE reazionato con le seguenti specifiche: $V_{CC} = 12\text{V}$; $A_{vv} = -5$; $R_{OUT} = 1\text{k}\Omega$.

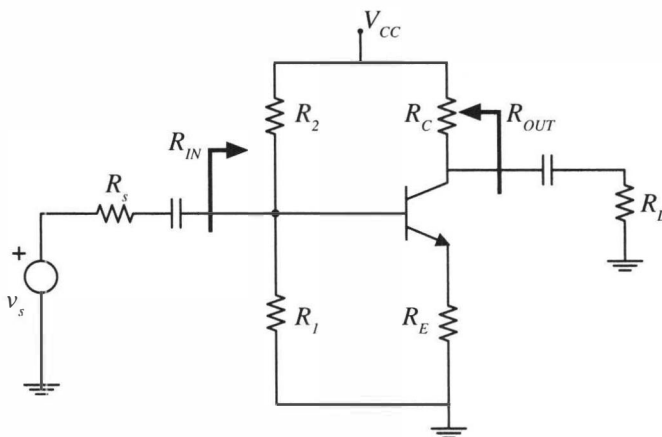


Figura 7.68 Amplificatore a BJT in configurazione CE reazionato

Poiché $R_{OUT} = R_C$, dalle specifiche segue che $R_C = 1\text{k}\Omega$.

Il guadagno in tensione è dato da

$$A_{vv} = -\frac{R_C}{R_E} = -5 \quad (7.132)$$

Quindi $R_E = 200\Omega$. La somma delle tensioni sulla maglia di uscita deve essere pari a V_{CC} , quindi

$$I_C R_C + V_{CE} + I_C R_E = V_{CC} \quad (7.133)$$

In presenza del segnale la tensione sul collettore varia in maniera simmetrica rispetto al valore di riposo. L'escursione massima in alto è pari alla caduta di tensione su R_C . Per avere circa la stessa escursione in basso è necessario impostare V_{CE} uguale alla caduta su R_C . Pertanto la (7.133) diviene

$$2I_C R_C + I_C R_E = V_{CC} \quad (7.134)$$

da cui

$$I_C = \frac{V_{CC}}{2R_C + R_E} = 5.45 \text{ mA} \quad (7.135)$$

Per questo valore di corrente

$$I_C R_E = 1.09 \text{ V} \quad (7.136)$$

Affinché la reazione negativa possa operare agevolmente deve essere rispettata la condizione

$$R_E I_C \gg V_T = 25 \text{ mV} \quad (7.137)$$

$R_E I_C = 1.09 \text{ V}$ è senz'altro molto maggiore di 25 mV .

Le tensioni ai nodi del circuito di polarizzazione sono quindi quelle in Figura 7.69, dove si è assunto $V_{BE} = 0.6 \text{ V}$.

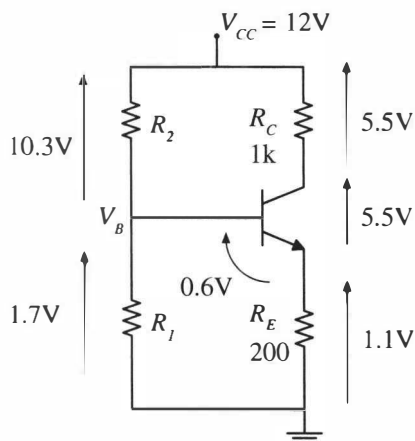


Figura 7.69 Tensioni ai nodi del circuito di polarizzazione

Resta ora da dimensionare le resistenze del partitore della maglia di ingresso. La corrente che vi scorre I_p deve essere molto maggiore della corrente di base:

$$I_p \gg I_B = \frac{I_C}{\beta_F} \quad (7.138)$$

Assumendo β_F maggiore di 100, segue che $I_p \gg 55\mu\text{A}$. Scegliamo $I_p = 0.5\text{mA}$.

Ne consegue che

$$R_1 = \frac{V_B}{I_p} = 3.4 \text{ k}\Omega \quad (7.139)$$

$$R_2 = \frac{V_{CC} - V_B}{I_p} = 20.6 \text{ k}\Omega \quad (7.140)$$

Il progetto del circuito sulla base delle specifiche date è concluso. Vogliamo però calcolare anche la resistenza di ingresso e la dinamica dell'amplificatore progettato.

La presenza della resistenza di emettitore R_E fa sì che la resistenza di ingresso in base sia

$$R_{IN}^{(B)} = r_\pi (1 + g_m R_E) + R_E \approx \beta_F R_E = 20 \text{ k}\Omega \quad (7.141)$$

avendo assunto $\beta_F = 100$.

La resistenza di ingresso è quindi

$$R_{IN} = R_1 \parallel R_2 \parallel \beta_F R_E = 2.55 \text{ k}\Omega \quad (7.142)$$

Calcoliamo ora la dinamica. Poiché l'escursione massima in uscita è 5.5V e il guadagno pari a -5, la massima ampiezza in ingresso per non saturare l'uscita è 1.1V. Vediamo ora la condizione di piccolo segnale.

$$v_s \ll (1 + R_E g_m) V_T \approx R_E I_C = 1.1 \text{ V} \quad (7.143)$$

Pertanto la dinamica è 1.1V e il punto a distorsione -20dB è 110mV. La Tabella 7.5 riassume le caratteristiche dell'amplificatore.

R_{IN}	2.55k Ω
R_{OUT}	1 k Ω
A_{VV}	- 5.0
Dinamica	1.1V
Distorsione - 20dB	110mV

Tabella 7.5

7.18 Teorema di Miller

Nel seguito troveremo molto utile il Teorema di Miller. Questo permette di semplificare l'analisi di un circuito lineare quando un'impedenza collega il morsetto di ingresso con quello di uscita ed è noto il rapporto tra le due tensioni.

Il teorema afferma che il circuito in Figura 7.70 è del tutto equivalente a quello di Figura 7.71 dove

$$A = \frac{v_o}{v_i} \quad (7.144)$$

L'equivalenza è dimostrata se l'impedenza vista in ingresso e in uscita è la stessa per i due circuiti.

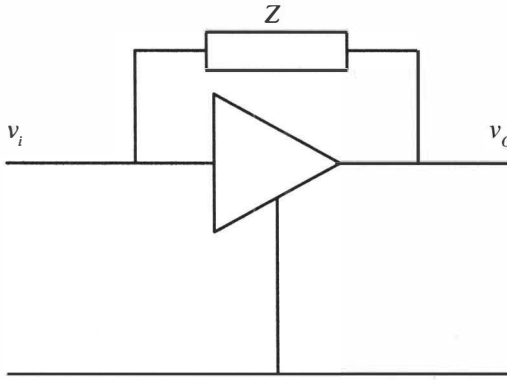


Figura 7.70 Teorema di Miller

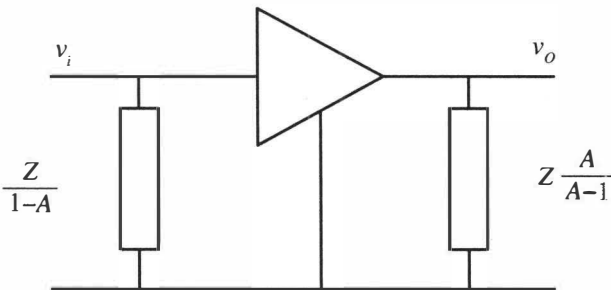


Figura 7.71 Teorema di Miller

Dimostriamo il teorema. Con riferimento alla Figura 7.72:

$$i = \frac{v_i - v_o}{Z} \quad (7.145)$$

da cui

$$i = \frac{1-A}{Z} v_i \quad (7.146)$$

Un generatore di tensione di prova tra l'ingresso e massa vedrà un'impedenza pari a

$$Z'_{IN} = \frac{v_i}{i} = \frac{Z}{1-A} \quad (7.147)$$

come mostrato in Figura 7.71.

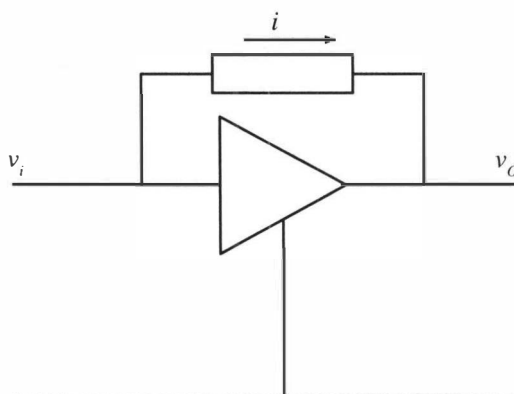


Figura 7.72 Dimostrazione del teorema di Miller

In maniera analoga si dimostra che un generatore di prova in uscita vede un'impedenza pari a $ZA/(A-1)$, come mostrato in Figura 7.71.

7.19 Amplificatore a MOSFET con resistore tra gate e drain

Si consideri il circuito in Figura 7.73. Supponiamo $R_G \gg R_D$. Si noti che il punto di polarizzazione è indipendente dal valore della resistenza R_G (che quindi può essere scelta grande a piacere). Se vale la condizione detta sopra il guadagno in tensione è $A_{vv} = -g_m R_D$.

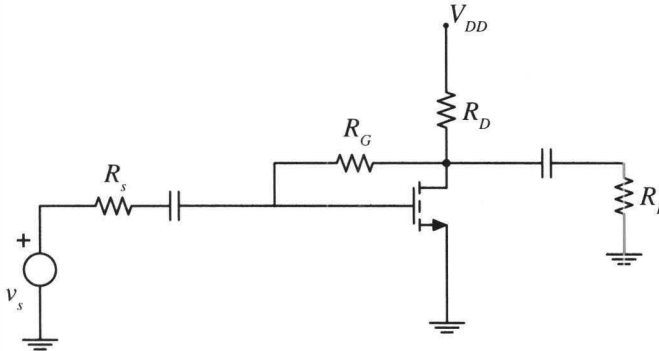


Figura 7.73 Amplificatore CS

Applichiamo ora il teorema di Miller.

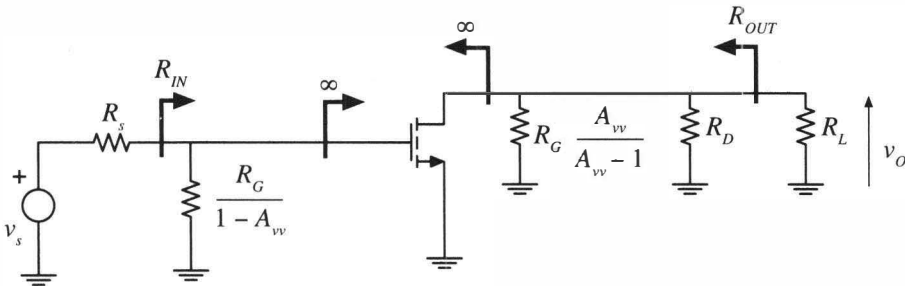


Figura 7.74 Circuito AC

Dal circuito in Figura 7.74 si ottiene che

$$R_{IN} = \frac{R_G}{1 - A_{vv}} \approx \frac{R_G}{g_m R_D} \quad (7.148)$$

$$R_{OUT} = R_G \left(\frac{A_{vv}}{1 - A_{vv}} \right) \parallel R_D \approx R_G R_D \quad (7.149)$$

7.20 Amplificatori a carico attivo

Nei paragrafi precedenti abbiamo sempre utilizzato il transistor come dispositivo di controllo della corrente erogata a una resistenza. Tuttavia, come abbiamo già visto nel Capitolo 5, un ulteriore transistor può essere utilizzato anche come carico attivo. In Figura 7.75 è mostrato un esempio.

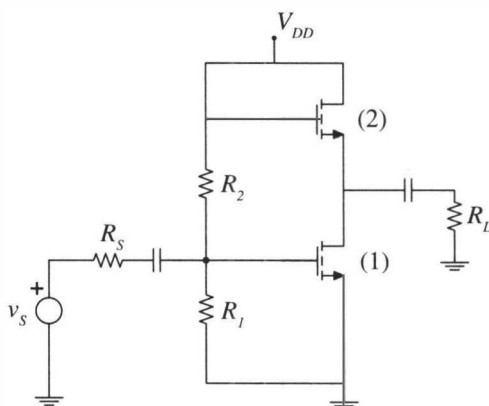


Figura 7.75 Amplificatore con carico attivo

Abbiamo già visto nel paragrafo 4.19 che la funzione di trasferimento del circuito in Figura 7.76 ha una regione di funzionamento lineare (vedi Figura 7.77) di pendenza pari a

$$p = -\sqrt{\frac{K_n^{(1)}}{K_n^{(2)}}} \quad (7.150)$$

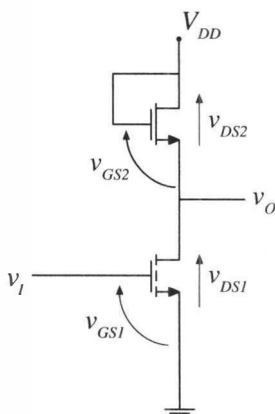


Figura 7.76 MOSFET come carico attivo

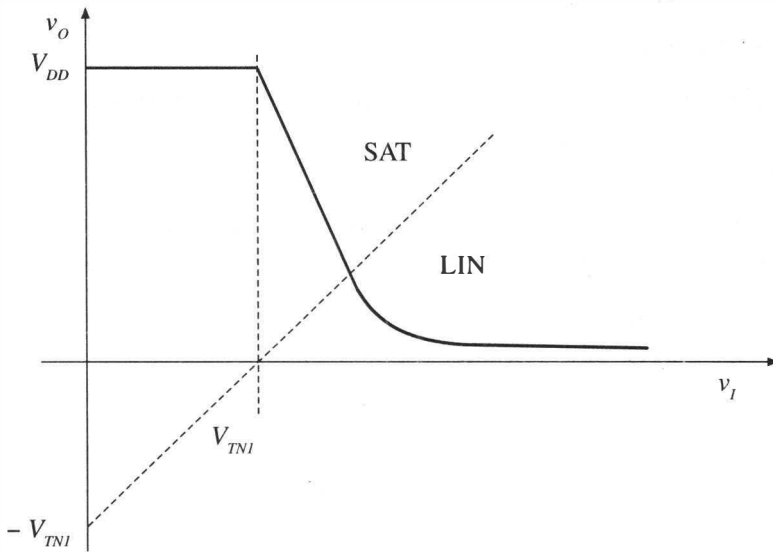


Figura 7.77 Funzione di trasferimento

Pertanto il guadagno in tensione è pari a

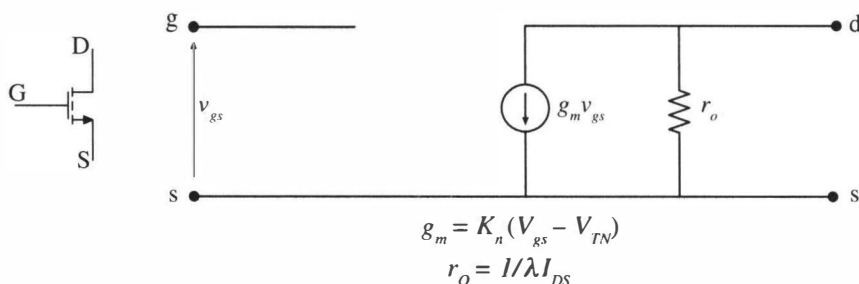
$$A_{vv} = -\sqrt{\frac{K_n^{(1)}}{K_n^{(2)}}} \quad (7.151)$$

Si noti che questo amplificatore è lineare fintanto che l'uscita rimane nel tratto lineare, indipendentemente dall'ampiezza del segnale in ingresso. Si tratta di una caratteristica molto importante di questa soluzione circuitale. Il carico attivo compensa la non linearità del MOSFET estendendo notevolmente il range di linearità, ben oltre le condizioni di piccolo segnale viste nei paragrafi precedenti. Per la stima della resistenza di uscita è immediato verificare che in uscita si ha il parallelo della resistenza vista al drain del MOSFET (1) (r_o , un valore piuttosto grande) con la resistenza vista al source che è pari a $1/g_m$. Quest'ultima è tipicamente piuttosto piccola e quindi in parallelo predominante sull'altra, pertanto

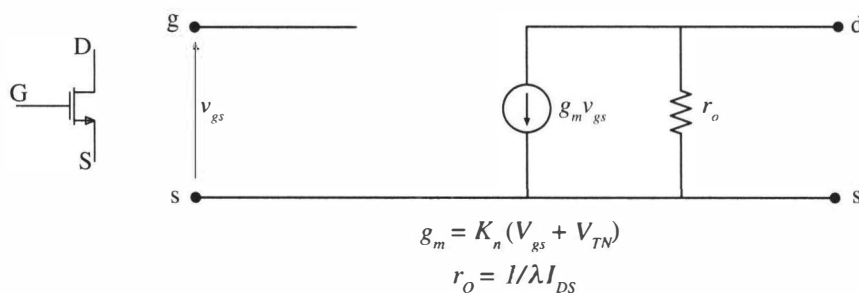
$$R_{OUT} = \frac{1}{g_m^{(2)}} \quad (7.152)$$

7.21 Riepilogo

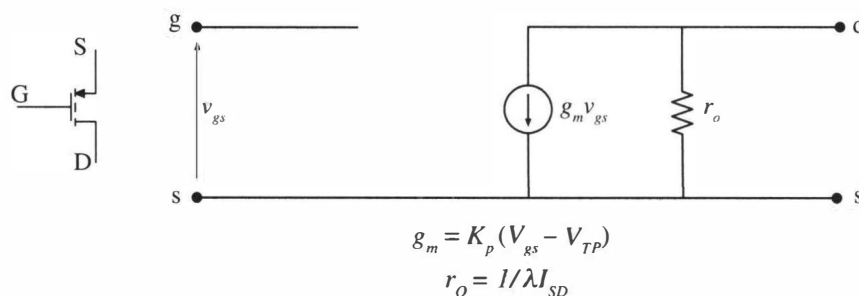
- Un amplificatore lineare non distorce il segnale.
- Il principio di sovrapposizione valido per i circuiti lineari consente di trattare separatamente la componente continua e ciascuna singola frequenza del segnale.
- Per piccoli segnali un **MOSFET a canale n ad arricchimento** polarizzato in saturazione può essere approssimato con il circuito che segue:



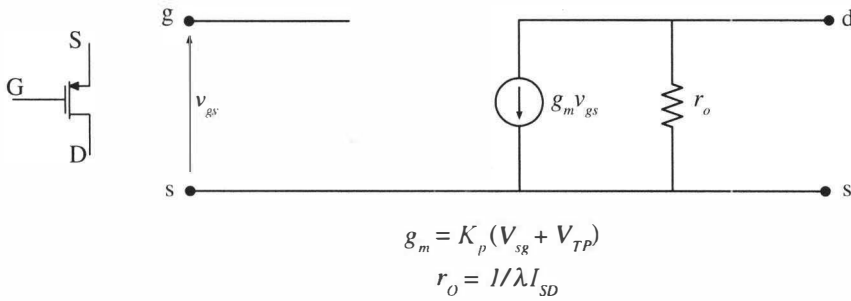
- Per piccoli segnali un **MOSFET a canale n a svuotamento** polarizzato in saturazione può essere approssimato con il circuito che segue:



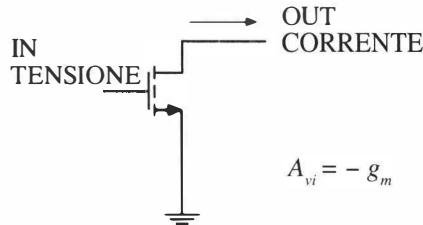
- Per piccoli segnali un **MOSFET a canale p ad arricchimento** polarizzato in saturazione può essere approssimato con il circuito che segue:



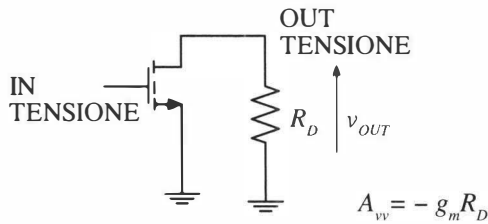
- Per piccoli segnali un **MOSFET a canale p a svuotamento** polarizzato in saturazione può essere approssimato con il circuito che segue:



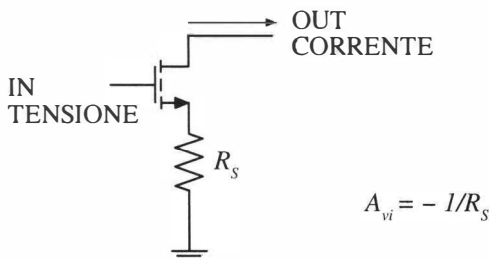
- Configurazione CS**



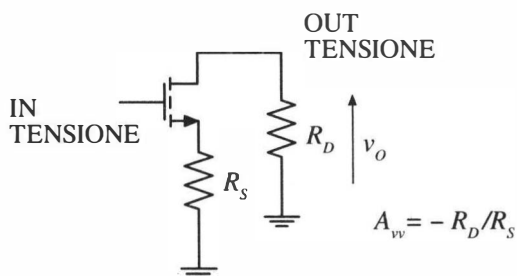
- Configurazione CS con uscita in tensione**



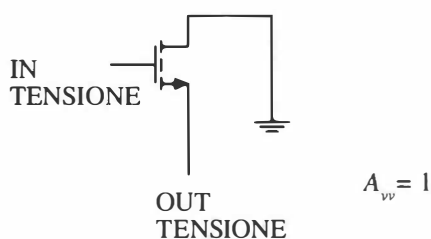
- Configurazione CS con reazione**



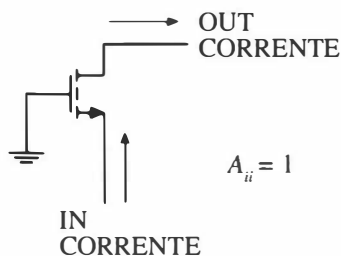
- **Configurazione CS con reazione e uscita in tensione**



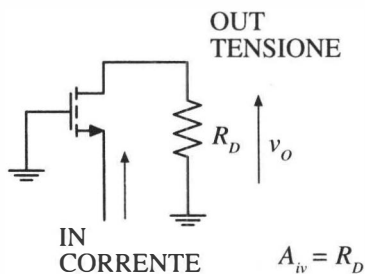
- **Configurazione CD**



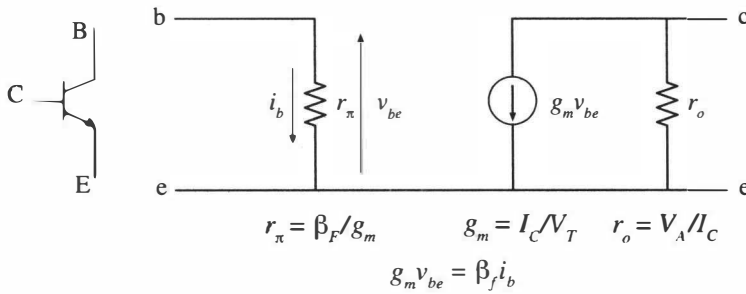
- **Configurazione CG**



- **Configurazione CG con uscita in tensione**

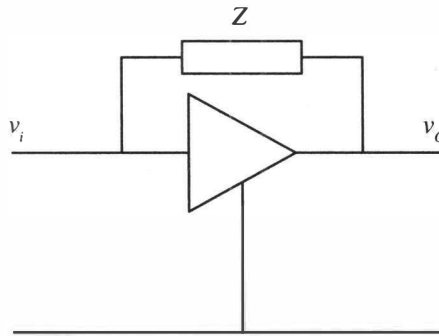


- Per piccoli segnali un **BJT npn** polarizzato in regione attiva diretta può essere approssimato con il circuito che segue:

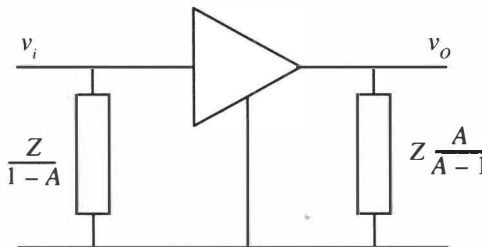


- Il modello linearizzato del **BJT pnp** è del tutto identico al modello del BJT npn.
- Le stesse configurazioni dei FET valgono per i BJT sostituendo nei reofori: G→B, D→C, S→E.
- Teorema di Miller**

Il circuito:



è equivalente a:

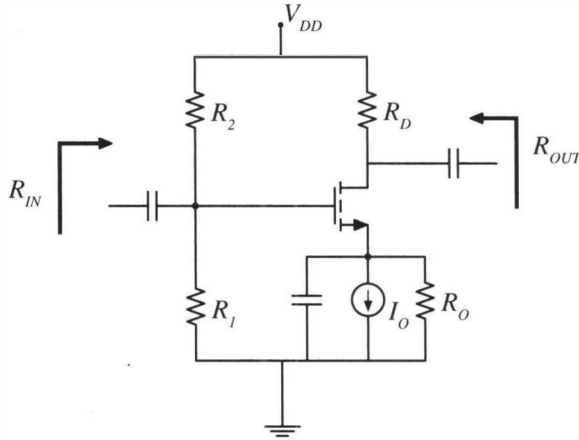


con $A = v_o / v_i$

7.22 Esercizi svolti

ANALISI DI AMPLIFICATORI A FET

1) Calcolare resistenza di ingresso, resistenza di uscita, guadagno, dinamica, punto di distorsione a -20dB , consumo dell'amplificatore tensione-tensione in figura. $R_I = 30\text{k}$, $R_2 = 30\text{k}$, $V_{DD} = 6\text{V}$, $R_D = 400$, $I_0 = 5\text{mA}$, $R_0 = 50\text{k}$, $K_n = 10\text{mA/V}^2$, $V_{TN} = 1\text{V}$.



La resistenza di ingresso è

$$R_{IN} = R_I \parallel R_2 = 15\text{k}\Omega$$

La resistenza di uscita è

$$R_{OUT} = R_D = 400\Omega$$

Supponiamo $R_0 \gg 1/g_m$, allora la corrente I_{DS} è circa uguale alla corrente I_0 e quindi

$$I_0 = \frac{K_n}{2} (V_{GS} - V_{TN})^2$$

da cui

$$V_{GS} = V_{TN} + \sqrt{\frac{2I_0}{K_n}} = 2\text{V}$$

Verifichiamo l'ipotesi fatta sopra:

$$g_m = K_n (V_{GS} - V_{TN}) = 0.01\Omega^{-1}$$

Quindi

$$\frac{1}{g_m} \ll R_o$$

$$100\Omega \ll 50k\Omega$$

Il guadagno è dato da

$$A_{vv} = -g_m R_D = -4$$

Per la dinamica è necessario calcolare

$$I_{DS} R_D = 2V$$

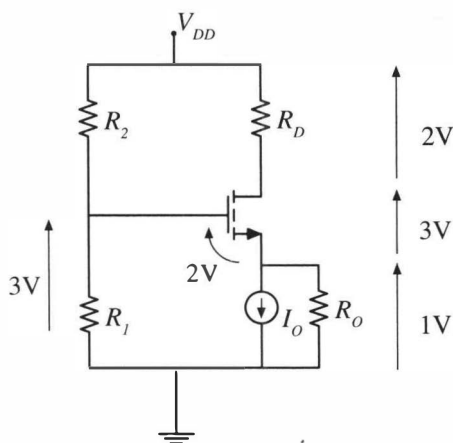
$$V_G = \frac{R_1}{R_1 + R_2} V_{DD} = 3V$$

$$V_S = V_G - V_{GS} = 1V$$

$$V_{DS} = V_{DD} - I_{DS} R_D - V_S = 3V$$

$$V_{DS} - (V_{GS} - V_{TN}) = 2V$$

L'escursione superiore in uscita è 2V, l'escursione inferiore è 2V. La dinamica è $2V/4 = 0.5V$.



Il punto di distorsione a $-20dB$ è

$$v_{sig} < 2(V_{GS} - V_{TN})/10 = 0.2V$$

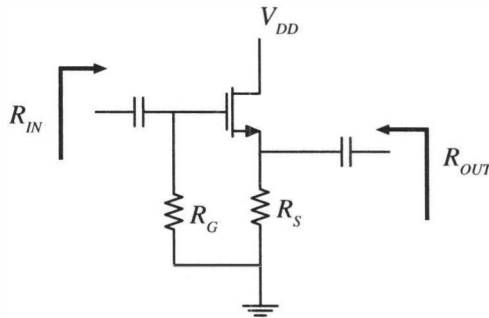
Il consumo è

$$P = V_{DD} \left(I_O + \frac{V_{DD}}{R_1 + R_2} \right) = 30.6mW$$

La tabella riassume i risultati dell'analisi effettuata.

R_{IN}	15k Ω
R_{OUT}	400 Ω
A_{VV}	- 4.0
Dinamica	0.5V
Distorsione - 20dB	0.2V
Consumo	30.6mW

2) Calcolare resistenza di ingresso, resistenza di uscita, guadagno, dinamica, punto di distorsione a - 20dB dell'amplificatore tensione-tensione in figura. $K_n = 20\text{mA/V}^2$, $V_{TN} = 1\text{V}$, $V_{DD} = 3\text{V}$, $R_S = 1\text{k}\Omega$, $R_G = 50\text{k}\Omega$.



Si tratta di un CD. La resistenza di ingresso è

$$R_{IN} = R_G = 50\text{k}\Omega$$

Per calcolare il punto di lavoro, scriviamo

$$\frac{K_n}{2} (V_{GS} + V_{TN})^2 = - \frac{V_{GS}}{R_S}$$

che è un'equazione di secondo grado nell'incognita V_{GS} . Le due soluzioni sono - 1.37V e - 0.73V. La prima è sotto la soglia ($V_{GS} < - 1\text{V}$) e quindi la soluzione accettabile è $V_{GS} = - 0.73$. Da cui

$$I_{DS} = - \frac{V_{GS}}{R_S} = 0.73\text{mA}$$

$$g_m = K_n (V_{GS} + V_{TN}) = 5.4 \times 10^{-3}\Omega^{-1}$$

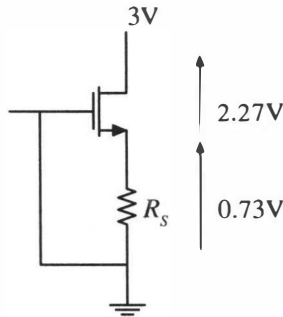
La resistenza vista “guardando” dentro il source è $1/g_m = 185\Omega$. La resistenza di uscita è

$$R_{OUT} = R_s \parallel \frac{1}{g_m} = 156\Omega$$

Essendo un buffer di tensione il guadagno nominale è 1, ma il guadagno effettivo dipende dal carico.

$$A_{vv} = \frac{g_m (R_s \parallel R_L)}{1 + g_m (R_s \parallel R_L)}$$

Nelle specifiche non viene dato il valore del carico di uscita, ma poiché si tratta di un buffer con uscita in tensione dovrà essere molto maggiore di $R_{OUT} = 156\Omega$. Scegliamo come valore significativo $R_L = 1.56k\Omega$, per il quale $A_{vv} = 0.77$.



Calcoliamo

$$I_{DS} R_s = 0.73V$$

$$V_{DD} - I_{DS} R_s - (V_{GS} + V_{TN}) = 2V$$

L'escursione in uscita è 0.73V, la dinamica in ingresso per il valore di guadagno calcolato sopra è pari a 0.95V.

Il punto di distorsione a -20dB è

$$v_{sig} < 2(V_{GS} + V_{TN}) \left(1 + g_m (R_s \parallel R_L) \right) / 10$$

Per $R_L = 1.56k\Omega$, si ottiene $v_{sig} < 0.232V$

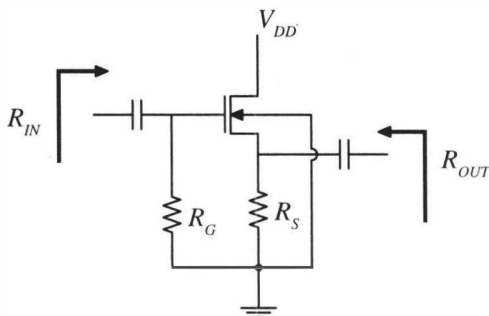
Il consumo è

$$P = V_{DD} I_{DS} = 2.2mW$$

La tabella riassume i risultati dell'analisi effettuata

R_{IN}	50k Ω
R_{OUT}	156 Ω
$A_{VV} @ R_L = 1.56k\Omega$	0.77
Dinamica @ $R_L = 1.56k\Omega$	0.95V
Distorsione - 20dB @ $R_L = 1.56k\Omega$	0.232V
Consumo	2.2mW

***3) Calcolare resistenza di ingresso, resistenza di uscita e guadagno dell'amplificatore tensione-tensione in figura. $K_n = 20\text{mA/V}^2$, $V_{TN0} = 1\text{V}$, $\gamma = 0.2\text{V}$, $2\phi_F = 0.6\text{V}$, $V_{DD} = 3\text{V}$, $R_S = 1k\Omega$, $R_G = 50k\Omega$.**



Si tratta dello stesso circuito dell'esercizio precedente con l'unica differenza che è necessario tenere conto dell'effetto body. L'effetto body sposta leggermente la soglia. Per calcolare tale variazione si può supporre che sia abbastanza piccola da poter calcolare il punto di polarizzazione trascurando l'effetto body. Usiamo quindi il risultato dell'esercizio precedente: $V_{SB} = V_{SG} = 0.73\text{V}$. La variazione della soglia può quindi essere stimata come

$$\Delta V_{TN} = V_{TN} - V_{TN0} = \gamma \left(\sqrt{2\phi_F + V_{SB}} - \sqrt{2\phi_F} \right) = 75.7\text{mV}$$

Il punto di lavoro è dato da

$$\frac{K_n}{2} (V_{GS} + V_{TN})^2 = -\frac{V_{GS}}{R_S}$$

Derivando entrambi i membri rispetto a V_{TN} si ottiene

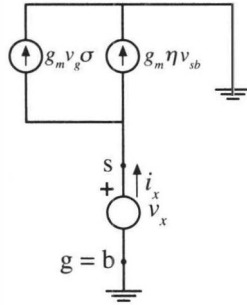
$$\frac{dV_{GS}}{dV_{TN}} = -\frac{K_n (V_{GS} + V_{TN})}{K_n (V_{GS} + V_{TN}) + 1} = -0.844$$

Pertanto la variazione del punto di lavoro è circa

$$\Delta V_{GS} \approx \frac{dV_{GS}}{dV_{TN}} \Delta V_{TN} = -63.9 \text{ mV}$$

Pertanto l'effettivo punto di polarizzazione è circa $V_{SB} = V_{SG} = 0.67 \text{ V}$.

Per calcolare la resistenza di uscita disegniamo il circuito AC in figura.



Dove

$$\eta = \frac{\gamma}{\sqrt{2\varphi_F + V_{SB}}} = 0.178$$

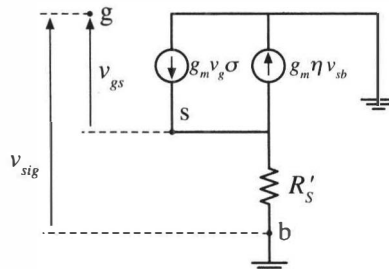
Dal circuito AC si ottiene

$$R_x = \frac{v_x}{i_x} = \frac{1}{g_m (1 + \eta)} = 157 \Omega$$

Possiamo quindi dire che, in questo specifico caso, l'effetto body migliora leggermente le prestazioni dell'amplificatore in termini di resistenza di uscita (nell'esercizio precedente la resistenza vista al source era 185Ω). La resistenza di uscita complessiva è

$$R_{OUT} = \frac{1}{g_m (1 + \eta)} \parallel R_S = 136 \Omega$$

Per calcolare il guadagno disegniamo il circuito AC in figura.



dove

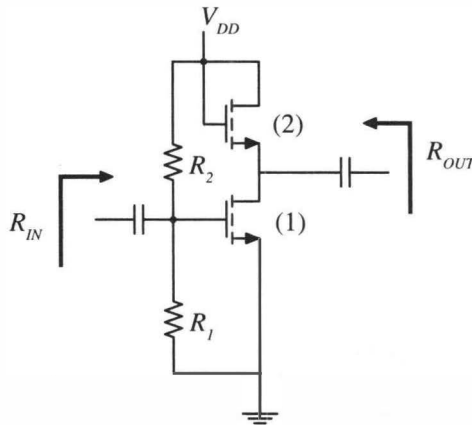
$$R'_S = R_S \parallel R_L$$

Si ottiene

$$A_{vv} = \frac{v_{sb}}{v_{sig}} = \frac{g_m R'_S}{1 + g_m R'_S (1 + \eta)}$$

per $R_L = 1.36\text{k}\Omega$ (dieci volte la resistenza di uscita), $A_{vv} = 0.69$. Possiamo quindi dire che, in questo specifico caso, l'effetto body peggiora leggermente le prestazioni dell'amplificatore in termini di guadagno (nell'esercizio precedente il guadagno era 0.77).

4) Calcolare resistenza di ingresso, resistenza di uscita, guadagno, dinamica, punto di compressione a -20dB dell'amplificatore tensione-tensione in figura. $K_{n1} = 10\text{mA/V}^2$, $K_{n2} = 1\text{mA/V}^2$, $V_{TN1} = V_{TN2} = 1\text{V}$, $R_1 = 20\text{k}\Omega$, $R_2 = 100\text{k}\Omega$. $V_{DD} = 12\text{V}$.



Si tratta di un CS. La resistenza di ingresso è

$$R_{IN} = R_1 \parallel R_2 = 16.7\text{k}\Omega$$

Il guadagno

$$A_{vv} = \sqrt{\frac{K_{n1}}{K_{n2}}} = 3.16$$

Il partitore fissa la V_{GS1} che è pari a

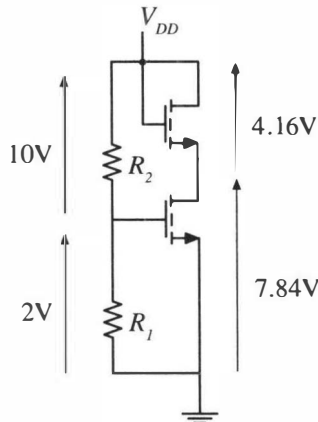
$$V_{GS1} = \frac{R_1}{R_1 + R_2} V_{DD} = 2\text{V}$$

$$I_{DS1} = \frac{K_{n1}}{2} (V_{GS1} - V_{TN1})^2 = 5\text{mA}$$

Per il calcolo di V_{GS2} possiamo usare l'uguaglianza

$$\frac{K_{n1}}{2} (V_{GS1} - V_{TN1})^2 = \frac{K_{n2}}{2} (V_{GS2} - V_{TN2})^2$$

da cui $V_{GS2} = 4.16\text{V}$.



La resistenza di uscita è pari a

$$R_{OUT} = \frac{1}{g_{m2}} = \frac{1}{K_{n2} (V_{GS2} - V_{TN2})} = 316\Omega$$

L'escursione superiore in uscita è pari a $V_{GS2} = 4.16\text{V}$. L'escursione inferiore in uscita è pari a $V_{DD} - V_{GS2} - (V_{GS1} - V_{TN1}) = 6.84\text{V}$. Quindi l'escursione in uscita è 4.16V e la dinamica $4.16\text{V}/3.16 = 1.32\text{V}$.

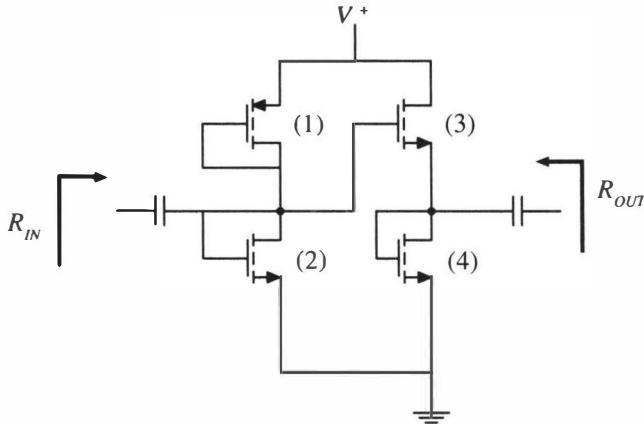
L'amplificatore è teoricamente lineare per qualunque ampiezza del segnale in ingresso. Il consumo è

$$P = V_{DD} \left(I_{DS1} + \frac{V_{DD}}{R_1 + R_2} \right) = 61.2\text{mW}$$

La tabella riassume i risultati dell'analisi effettuata.

R_{IN}	16.7k Ω
R_{OUT}	316 Ω
A_{VV}	3.16
Dinamica	1.32V
Distorsione - 20dB	∞
Consumo	61.2mW

5) Calcolare resistenza di ingresso, resistenza di uscita, guadagno, dinamica, punto di compressione a -20dB dell'amplificatore tensione-tensione in figura. $K_p^{(1)} = K_n^{(2)} = 0.1\text{mA/V}^2$, $K_n^{(3)} = K_n^{(4)} = 10\text{mA/V}^2$, $\lambda^{(1)} = \lambda^{(2)} = \lambda^{(3)} = \lambda^{(4)} = 0.01\text{V}^{-1}$, $V_{TP}^{(1)} = V_{TN}^{(2)} = V_{TN}^{(3)} = V_{TN}^{(4)} = 1\text{V}$, $V^+ = 5\text{V}$.



I MOSFET (1) e (2) fissano la tensione sulla base del MOSFET (3). Il MOSFET (4) sostituisce la resistenza di source del MOSFET (3). Si tratta quindi di un CD.

I MOSFET (1), (2) e (4) sono sicuramente in saturazione perché hanno G e D cortocircuitati.

I MOSFET (1) e (2) hanno stesso K, stessa soglia, e stesso λ . Poiché vi scorre la stessa corrente devono avere la stessa caduta di tensione. Quindi: $v_{DS1} = v_{DS2} = V^+/2 = 2.5\text{V}$. Inoltre

$$I_{SD}^{(1)} = I_{DS}^{(2)} = \frac{K_n^{(2)}}{2} (V_{GS2} - V_{TN2})^2 = 0.113\text{mA}$$

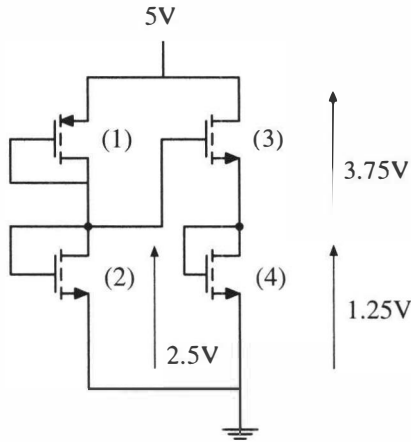
e

$$r_o^{(1)} = r_o^{(2)} = \frac{1}{\lambda^{(2)} I_{DS}^{(2)}} = 885\text{k}\Omega$$

La resistenza di ingresso è

$$R_{IN} = r_o^{(1)} \parallel r_o^{(2)} = 442\text{k}\Omega$$

I MOSFET (3) e (4) hanno stesso K, stessa soglia, e stesso λ . Se anche (3) è in saturazione (cosa da verificare nel seguito), poiché nei tre transistor scorre la stessa corrente devono avere la stessa v_{GS} . D'altra parte $v_{GS4} + v_{GS3} = v_{GS2}$. Segue che $v_{GS4} = v_{GS3} = 1.25\text{V}$.



Quindi

$$I_{DS}^{(3)} = I_{DS}^{(4)} = \frac{K_n^{(4)}}{2} (V_{GS4} - V_{TN4})^2 = 0.313 \text{ mA}$$

Verifichiamo che il MOSFET (3) sia in saturazione

$$V_{DS3} = V^+ - 1.25 \text{ V} = 3.75 \text{ V} > V_{GS3} - V_{TN3} = 0.25 \text{ V}$$

La resistenza di uscita è

$$R_{OUT} = \frac{1}{g_m^{(3)}} \parallel r_o^{(4)} \approx \frac{1}{g_m^{(3)}} = \frac{1}{K_n^{(3)} (V_{GS3} - V_{TN3})} = 400 \Omega$$

Calcoliamo il guadagno:

$$A_{vv} = \frac{(r_o^{(4)} \parallel R_L) g_m^{(3)}}{1 + (r_o^{(4)} \parallel R_L) g_m^{(3)}}$$

Poiché

$$r_o^{(4)} = \frac{1}{\lambda^{(4)} I_{DS}^{(4)}} = 319 \text{ k}\Omega$$

è un valore molto grande, l'effettivo guadagno dipende praticamente solo dal carico:

$$A_{vv} = \frac{R_L g_m^{(3)}}{1 + R_L g_m^{(3)}}$$

Nelle specifiche non viene dato il valore del carico di uscita, in ogni caso dovrà essere molto maggiore di $R_{OUT} = 400 \Omega$. Prendiamo $R_L = 4 \text{ k}\Omega$, per il quale $A_{vv} = 0.91$.

La massima escursione del segnale in uscita è $V_{GS4} = 1.25 \text{ V}$. La dinamica è quindi $1.25 \text{ V} / 0.91 = 1.37$.

Il punto di distorsione a -20dB è

$$v_{sig} < 2 \left(V_{GS}^{(3)} - V_{TN}^{(3)} \right) g_m^{(3)} \left(r_0^{(4)} \parallel R_L \right) / 10 \approx 2 \left(V_{GS}^{(3)} - V_{TN}^{(3)} \right) g_m^{(3)} R_L / 10$$

che dipende dal carico. Assumendo ancora $R_L = 4\text{k}\Omega$, il punto di distorsione a -20dB è 0.5V .
Il consumo è

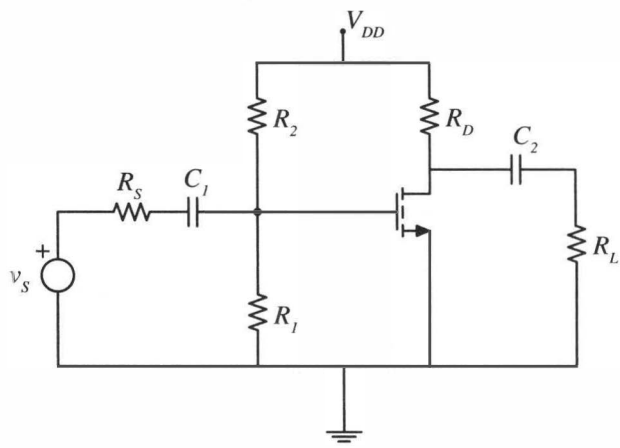
$$P = V_{DD} \left(I_{DS1} + I_{DS4} \right) = 2.13\text{mW}$$

La tabella riassume i risultati dell'analisi effettuata.

R_{IN}	442k Ω
R_{OUT}	400 Ω
$A_{VV} @ R_L = 4\text{k}\Omega$	0.91
Dinamica @ $R_L = 4\text{k}\Omega$	1.37V
Distorsione $-20\text{dB} @ R_L = 4\text{k}\Omega$	0.5V
Consumo	2.13mW

PROGETTO DI AMPLIFICATORI A FET

1) Progettare un amplificatore tensione-tensione a singolo MOSFET a canale n ad arricchimento in configurazione CS con le seguenti specifiche: $R_S = 1\text{k}\Omega$, $R_L = 100\text{k}\Omega$, $A_{vv} = -10$, $V_{DD} = 5\text{V}$, $K_n = 10\text{mA/V}^2$, $V_{TN} = 1\text{V}$.



La condizione sulla resistenza di uscita è

$$R_D \approx R_{OUT} \ll R_L = 100\text{k}\Omega$$

Fissiamo $R_D = 2\text{k}\Omega$.

La condizione sul guadagno è

$$A_{vv} = -10$$

ovvero

$$K_n (V_{GS} - V_{TN}) R_D = 10$$

da cui $V_{GS} = 1.5\text{V}$.

Controlliamo che il MOSFET sia in saturazione. Se è in saturazione:

$$i_{DS} = \frac{K_n}{2} (V_{GS} - V_{TN})^2 = 1.25\text{mA}$$

quindi

$$v_{DS} = V_{DD} - R_D i_D = 2.5\text{V}$$

La condizione da verificare è

$$v_{DS} > v_{GS} - V_{TN}$$

ovvero

$$2.5\text{V} > 0.5\text{V}$$

Dimensioniamo ora il partitore R_1 e R_2 in modo che la tensione sul gate sia 1.5V . Non ci sono vincoli particolari sulla corrente che vi scorre; scegliamo un valore piuttosto piccolo, compatibilmente con le resistenze disponibili, per ridurre la dissipazione. Scegliamo ad esempio $I_1 = I_2 = 10\mu\text{A}$, quindi

$$R_1 = \frac{V_{GS}}{I_1} = 150\text{k}\Omega$$

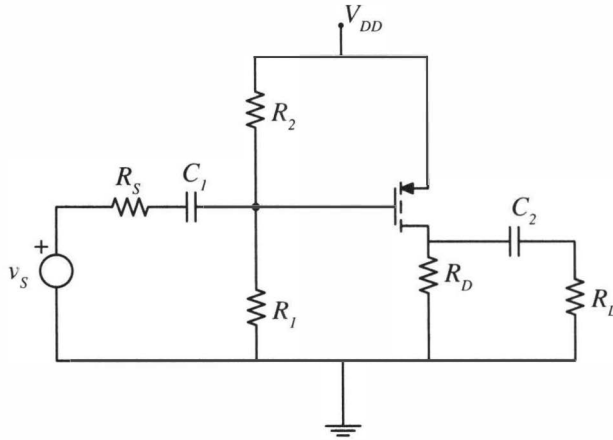
$$R_2 = \frac{V_{DD} - V_{GS}}{I_1} = 350\text{k}\Omega$$

Verifichiamo la condizione

$$R_{IN} \gg R_S = 50$$

$$R_{IN} = R_1 \parallel R_2 = 105\text{k}\Omega \gg 1\text{k}\Omega$$

2) Progettare un amplificatore tensione-corrente a singolo MOSFET a canale p ad arricchimento in configurazione CS con le seguenti specifiche: $R_s = 500\Omega$, $R_L = 50\Omega$, $A_{vi} = -5\text{mA/V}$, $V_{DD} = 5\text{V}$, $K_p = 10\text{mA/V}^2$, $V_{TP} = 1\text{V}$.



La condizione sul guadagno è

$$K_p (V_{SG} - V_{TP}) = 5\text{mA/V}^2$$

da cui $V_{SG} = 1.5\text{V}$.

Dimensioniamo ora il partitore R_1 e R_2 . Non ci sono vincoli particolari sulla corrente che vi scorre; scegliamo un valore piuttosto piccolo, compatibilmente con le resistenze disponibili, per ridurre la dissipazione. Scegliamo ad esempio $I_1 = I_2 = 10\mu\text{A}$, quindi

$$R_2 = \frac{V_{SG}}{I_1} = 150\text{k}\Omega$$

$$R_1 = \frac{V_{DD} - V_{SG}}{I_1} = 350\text{k}\Omega$$

Per dimensionare R_D è necessario fissare V_D . Non ci sono vincoli stringenti, purché il MOSFET sia in saturazione: fissiamo $V_D = 2.5\text{V}$, quindi

$$R_D = \frac{V_D}{I_{DS}} = \frac{V_D}{\frac{K_p}{2} (V_{SG} - V_{TP})^2} = \frac{2.5\text{V}}{1.25\text{mA}} = 2\text{k}\Omega$$

Verifichiamo la condizione

$$R_{IN} \gg R_s = 500\Omega$$

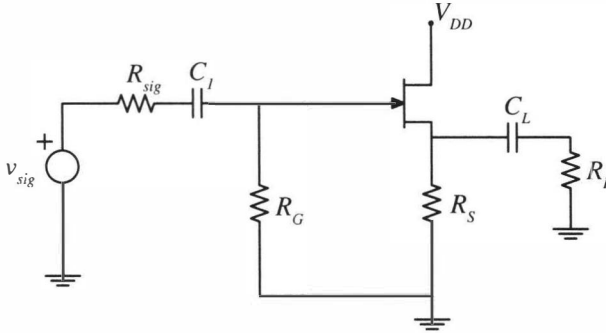
$$R_{IN} = R_1 \parallel R_2 = 105\text{k}\Omega$$

Verifichiamo la condizione

$$R_{OUT} \gg R_L = 50\Omega$$

$$R_{OUT} \approx R_D = 2k\Omega \gg 50\Omega$$

3) Progettare un buffer di tensione a singolo JFET a canale n in configurazione CD con le seguenti specifiche: $R_{sig} = 10k\Omega$, $R_L = 10k\Omega$, $V_{DD} = 5V$, $I_{DSS} = 10mA$, $V_p = 2V$. L'ampiezza massima del segnale in ingresso è 200mV.



Affinché un JFET sia in saturazione il gate deve essere a una tensione più bassa del source: una possibile soluzione è mantenere a massa il gate mediante una resistenza R_G (che può essere anche molto grande, in quanto non vi scorre corrente se non è applicato un segnale) e alzare la tensione sul source mediante una resistenza R_S nella quale scorre la corrente i_{DS} .

La corrente in saturazione è data da

$$i_{DS} = I_{DSS} \left(1 + \frac{V_{GS}}{V_p} \right)^2$$

Quindi

$$g_m = \left. \frac{\partial i_{DS}}{\partial V_{GS}} \right|_{V_{GS}=V_{GS}} = 2 \frac{I_{DSS}}{V_p} \left(1 + \frac{V_{GS}}{V_p} \right)$$

La condizione affinché lavori in regime di piccoli segnali è

$$V_{sig} \ll 2(V_p + V_{GS}) = 2V$$

poniamo quindi

$$V_p + V_{GS} = 1V$$

da cui $V_{GS} = -1V$.

La condizione di uscita in tensione impone che

$$R_{OUT} \approx \frac{1}{g_m} \ll 10\text{k}\Omega$$

$$\frac{1}{g_m} = \frac{1}{2 \frac{I_{DSS}}{V_P} \left(1 + \frac{V_{GS}}{V_P} \right)} = 200\Omega \ll 10\text{k}\Omega$$

Possiamo ora dimensionare R_S come segue

$$R_S = \frac{V_{GS}}{I_{DS}} = \frac{1\text{V}}{2.5\text{mA}} = 400\Omega$$

Per questo valore di R_S la resistenza di uscita non è propriamente pari a $1/g_m$, ma

$$R_{OUT} = \frac{1}{g_m} \parallel R_S = 133\Omega \ll 10\text{k}\Omega$$

Verifichiamo che il JFET sia effettivamente in saturazione, la condizione è

$$V_{DS} > V_P + V_{GS}$$

ovvero

$$V_{DD} - V_{SG} > 1\text{V}$$

$$4\text{V} > 1\text{V}$$

Per dimensionare R_G si tenga presente che

$$R_G = R_{IN} \gg R_{sig} = 10\text{k}\Omega$$

Fissiamo ad esempio $R_G = 100\text{k}\Omega$.

Prima di concludere è necessario fare una precisazione sul guadagno di questo buffer.

In generale vale

$$A_{vv} = \frac{R'_S g_m}{1 + R'_S g_m}$$

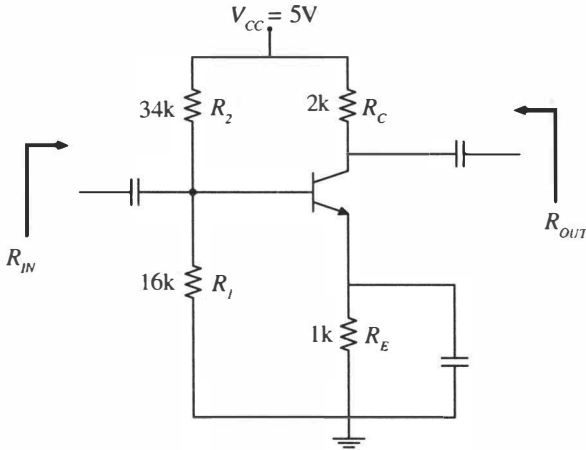
Se

$$R'_S g_m \gg 1$$

il guadagno è circa 1, tuttavia $R'_S g_m = 2$, quindi per questo buffer $A_{vv} = 0.67$.

ANALISI DI AMPLIFICATORI A BJT

1) Calcolare resistenza di ingresso, resistenza di uscita, guadagno, dinamica, punto di distorsione a -20dB , consumo. Si consideri $\beta_F = 200$.

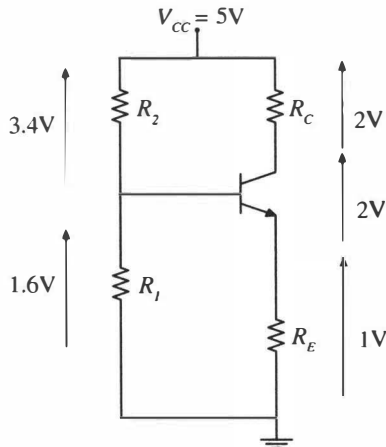


Si tratta di un CE. Se la corrente di base è trascurabile rispetto alla corrente che scorre in R_1 e R_2 :

$$V_B \approx \frac{R_1}{R_1 + R_2} V_{CC} = 1.6\text{V}$$

$$V_E = V_B - 0.6\text{V} = 1.0\text{V}$$

$$I_C \approx I_E = \frac{V_E}{R_E} = 1\text{mA}$$



Verifichiamo l'ipotesi sulla corrente di base:

$$I_B = \frac{I_C}{\beta_F} = 5\mu A \ll I_P = \frac{V_{CC}}{R_1 + R_2} = 100\mu A$$

Possiamo ora calcolare la resistenza di ingresso:

$$R_{IN} = R_1 \parallel R_2 \parallel r_\pi = (10.88k\Omega) \parallel r_\pi$$

$$r_\pi = \beta_F \frac{V_T}{I_C} = 5k\Omega$$

Quindi

$$R_{IN} = 3.43k\Omega$$

La resistenza di uscita è data da

$$R_{OUT} \approx R_C = 2k\Omega$$

Poiché $R_{IN} \gg R_s$ e $R_{OUT} \ll R_L$ si tratta di un amplificatore tensione-tensione il cui guadagno è quindi dato da

$$A_{vv} = - \frac{I_C R_C}{V_T} = -80$$

Calcoliamo ora la dinamica. L'ampiezza massima positiva in uscita è

$$V^+ = I_C R_C = 2V$$

L'ampiezza massima negativa è

$$V^- = V_{CE} - 0.2V = 1.8V$$

Tra le due è necessario considerare la minore. Quindi in ingresso

$$v_s < 22.5mV$$

La condizione affinché la distorsione sia $-20dB$ sotto il segnale è

$$v_s < 5mV$$

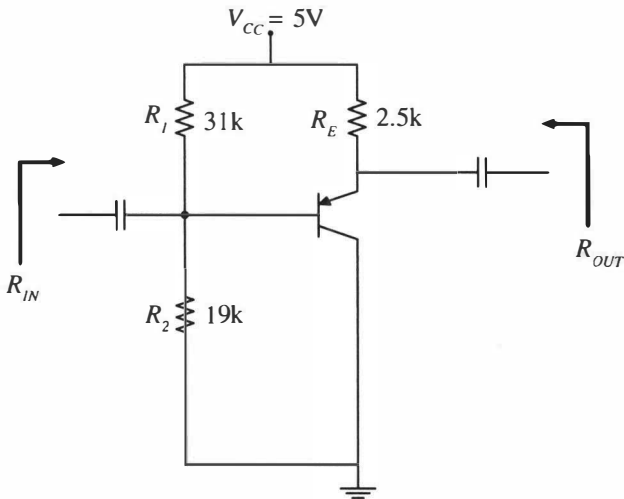
Calcoliamo ora il consumo.

$$P = V_{CC} (I_C + I_P) = 5.5mW$$

La tabella riassume i risultati dell'analisi effettuata.

R_{IN}	3.43k Ω
R_{OUT}	2k Ω
A_{VV}	- 80
Dinamica	22.5mV
Distorsione - 20dB	5mV
Consumo	5.5mW

2) Calcolare resistenza di ingresso, resistenza di uscita, guadagno, dinamica punto di distorsione a - 20dB, consumo. Si consideri $\beta_F = 100$.



Si tratta di un CC. Se la corrente di base è trascurabile rispetto alla corrente che scorre in R_1 e R_2 :

$$V_{CC} - V_B \approx \frac{R_1}{R_1 + R_2} V_{CC} = 3.1V$$

$$V_{CC} - V_E = V_{CC} - V_B + 0.6V = 2.5V$$

$$I_C \approx I_E = \frac{V_{CC} - V_E}{R_E} = 1mA$$

Verifichiamo l'ipotesi sulla corrente di base:

$$I_B = \frac{I_C}{\beta_F} = 10\mu A \ll I_P = \frac{V_{CC}}{R_1 + R_2} = 100\mu A$$

La resistenza di uscita è data da

$$R_{OUT} \approx \frac{V_T}{I} = 25\Omega$$

La resistenza di ingresso è data da

$$R_{IN} = R_I \parallel R_2 \parallel \left(r_\pi (1 + g_m (R_E \parallel R_L)) + R_E \parallel R_L \right)$$

Il valore dipende dal carico, che comunque dovrà essere molto più grande della resistenza di uscita. Come caso limite prendiamo: $R_L = 250\Omega$.

$$r_\pi (1 + g_m (R_E \parallel R_L)) + R_E \parallel R_L = 25.45k\Omega$$

Quindi $R_{IN} = 8.05k\Omega$.

Il guadagno di tensione (idealmente 1) dipende dal carico. Per $R_L = 250\Omega$.

$$A_{vv} = \frac{g_m (R_E \parallel R_L)}{1 + g_m (R_E \parallel R_L)} = 0.90$$

Calcoliamo ora la dinamica. L'ampiezza massima positiva in uscita è

$$V^+ = I_E R_C = 2.5V$$

L'ampiezza massima negativa è

$$V^- = V_{CE} - 0.2V = 2.3V$$

Tra le due è necessario considerare la minore. Quindi in ingresso $v_s < 2.3/0.90 = 2.56V$. Per il calcolo del punto di distorsione a $-20dB$

$$v_s < 2V_T (1 + g_m (R_E \parallel R_L)) / 10 = 50mV$$

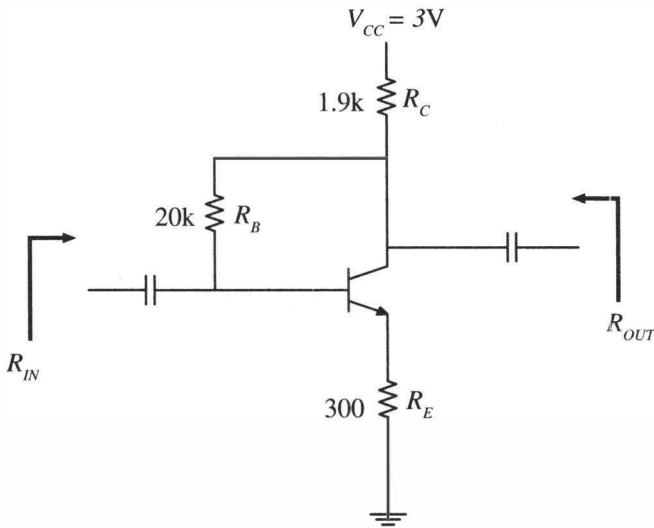
Calcoliamo ora il consumo.

$$P = V_{CC} (I_C + I_p) = 5.5mW$$

La tabella riassume i risultati dell'analisi effettuata.

$R_{IN} @ R_L = 250\Omega$	8.05k Ω
R_{OUT}	25 Ω
$A_{vv} @ R_L = 250\Omega$	0.90
Dinamica	2.56V
Distorsione - 20dB @ $R_L = 250\Omega$	50mV
Consumo	5.5mW

3) Calcolare resistenza di ingresso, resistenza di uscita, guadagno, dinamica, consumo.
 $\beta_F = 200$.



Si tratta di un CE reazionato. Se trascuriamo la caduta di tensione su R_B :

$$I_C = \frac{V_{CC} - V_{BE}}{R_C + R_B} = 1.1\text{mA}$$

Verifichiamo che la caduta su R_B sia piccola.

$$\Delta V = \frac{I_C}{\beta_F} R_B = 0.1\text{V} \ll (V_{CC} - 0.6\text{V}) = 2.4\text{V}$$

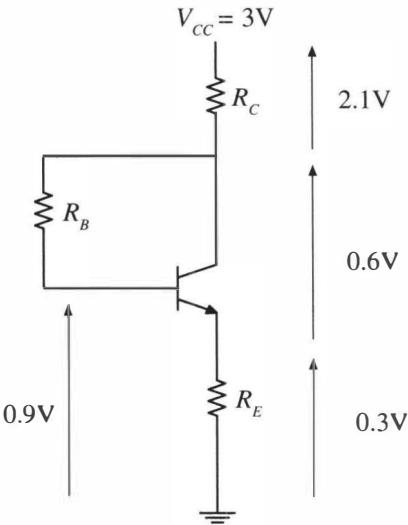
Il guadagno è

$$A_{vv} = - \frac{g_m R_C}{1 + g_m R_E} = -5.89$$

Per calcolare la resistenza di ingresso è necessario applicare il teorema di Miller.

$$R_{IN} = \frac{R_B}{1 - A_{vv}} \parallel (r_\pi (1 + g_m R_E) + R_E) = 2.90\text{k}\Omega \parallel 65\text{k}\Omega = 278\text{k}\Omega$$

La resistenza di uscita è circa pari a $R_C = 1.9\text{k}\Omega$.



L'escursione superiore del segnale è 2.1V, l'escursione inferiore 0.6V – 0.2V = 0.4V. Quindi la dinamica in ingresso è 75mV.

Il punto di distorsione a – 20dB è

$$v_{sig} \ll 2V_T (1 + g_m R_E) = 71\text{mV}$$

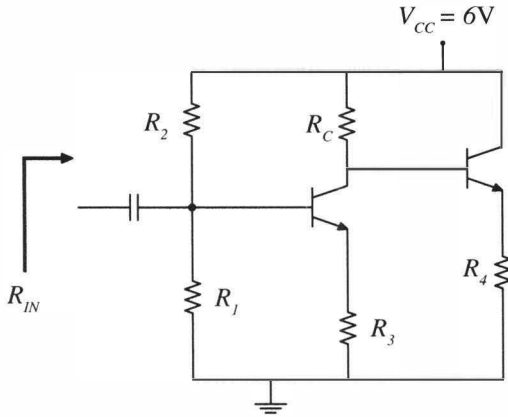
Il consumo è

$$P = V_{CC} I_C = 3.3\text{mW}$$

La tabella riassume i risultati dell'analisi effettuata.

R_{IN}	2.78k Ω
R_{OUT}	1.9k Ω
A_{VV}	– 5.89
Dinamica	68mV
Distorsione – 20dB	71mV
Consumo	3.3mW

4) Con riferimento all'amplificatore tensione-tensione in figura, si calcoli il guadagno, la resistenza di ingresso, la resistenza di uscita e la dinamica. $\beta_F = 200$ per entrambi i BJT.



Si tratta di un CE seguito da un CC.

Le resistenze R_1 e R_2 fissano il punto di lavoro:

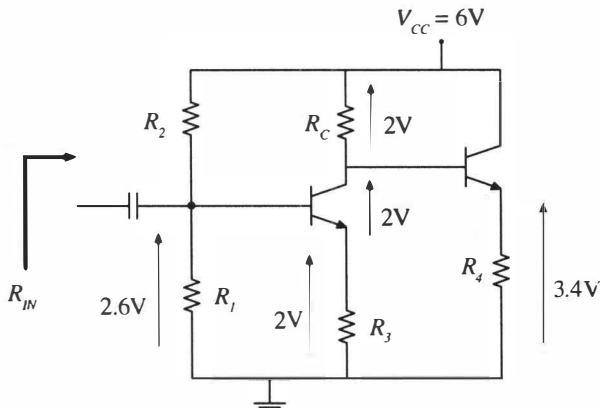
$$V_{B1} = \frac{R_2}{R_1 + R_2} V_{CC} = 2.6\text{V}$$

$$I_{C1} = \frac{V_{B1} - (0.6\text{V})}{R_3} = 2\text{mA}$$

Essendo piccola la corrente di base del secondo BJT:

$$V_{B2} = V_{CC} - I_{C1} R_C = 4\text{V}$$

$$I_{C2} = \frac{V_{B2} - (0.6\text{V})}{R_4} = 1.7\text{mA}$$



La resistenza di ingresso è

$$R_{IN} = R_1 \parallel R_2 \parallel \frac{\beta_F}{g_m} = 2082\Omega$$

La resistenza di uscita è

$$R_{OUT} = \frac{1}{g_{m2}} = 15\Omega$$

Il guadagno del CC è circa 1, quindi il guadagno complessivo è

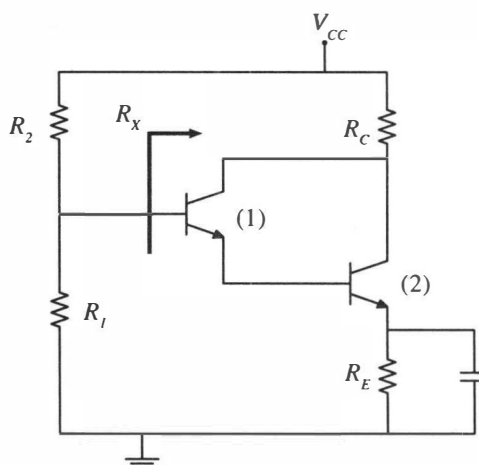
$$A_{vv} = -g_{m1}R_C = -160$$

L'escursione del primo stadio è 2V. L'escursione di ingresso del secondo stadio è pari a quella di uscita: 2.4V. Quindi è il primo stadio che limita la dinamica. La dinamica di ingresso è pertanto 12.5mV.

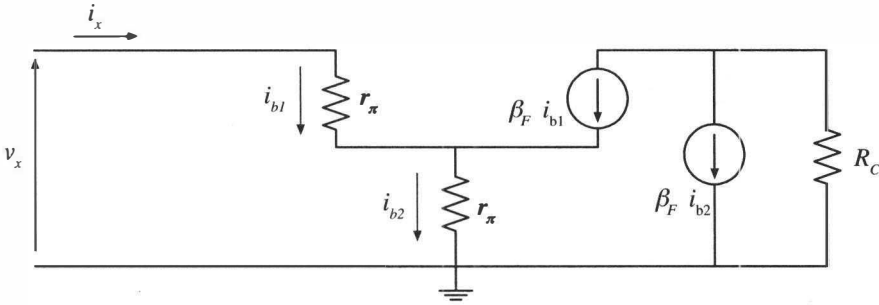
La tabella riassume i risultati dell'analisi effettuata.

R_{IN}	2.01k Ω
R_{OUT}	15 Ω
A_{VV}	- 160
Dinamica	12.5mV

***5) Calcolare la resistenza vista “guardando” nella base del BJT (1).**



Per calcolare R_x si consideri il circuito AC in figura.



Scriviamo

$$v_x = r_\pi i_{b1} + r_\pi i_{b2} = r_\pi i_{b1} + r_\pi (\beta_F + 1) i_{b1}$$

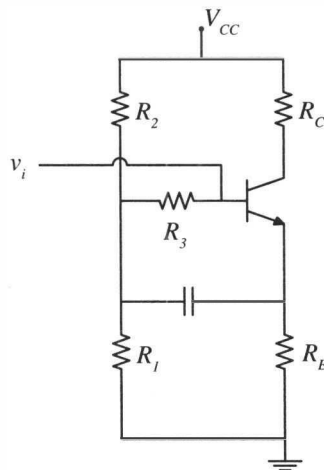
$$i_x = i_{b1}$$

Pertanto

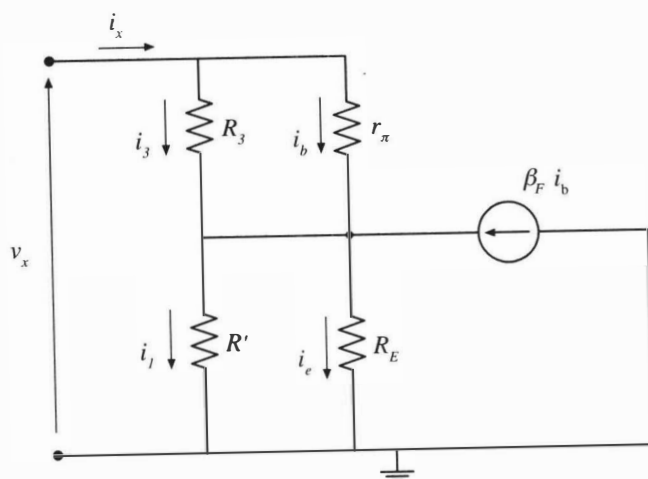
$$R'_{IN} = \frac{v_x}{i_x} = r_\pi + r_\pi (\beta_F + 1) \approx \beta_F r_\pi$$

*Nota: il circuito in figura è una configurazione nominata **Darlington**, che ha proprio il pregio di offrire una grande resistenza di base.*

***6) Calcolare la resistenza vista al terminale v_i .**



Il circuito AC da analizzare è mostrato in figura.



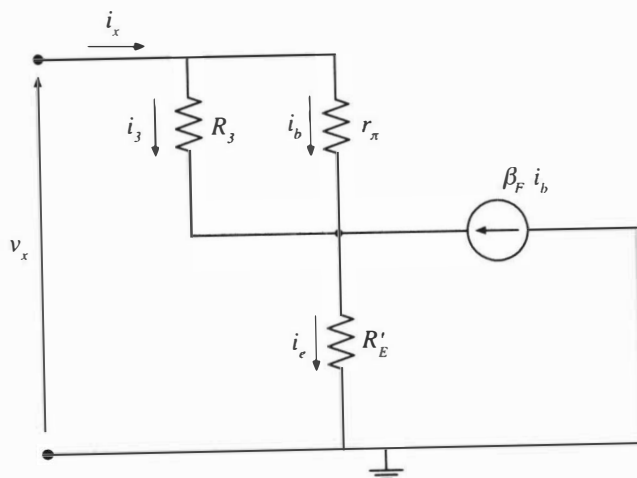
Dove

$$R' = R_1 \parallel R_2$$

Definiamo ora

$$R'_E = R_E \parallel R'$$

Il circuito può essere ora tracciato come nella figura che segue:



Calcoliamo ora

$$v_x = i_b r_\pi + i_2 R'_E$$

D'altra parte

$$i_2 = i_3 + i_b + \beta_F i_b$$

Inoltre circuitando la maglia che contiene R_3 e r_π si ottiene

$$i_3 R_3 = r_\pi i_b$$

ovvero

$$i_3 = \frac{r_\pi}{R_3} i_b$$

Sostituendo nell'espressione di v_x si ottiene

$$v_x = i_b r_\pi + \left(\frac{r_\pi}{R_3} + 1 + \beta_F \right) R'_E i_b$$

D'altra parte

$$i_x = i_3 + i_b = \left(\frac{r_\pi}{R_3} + 1 \right) i_b$$

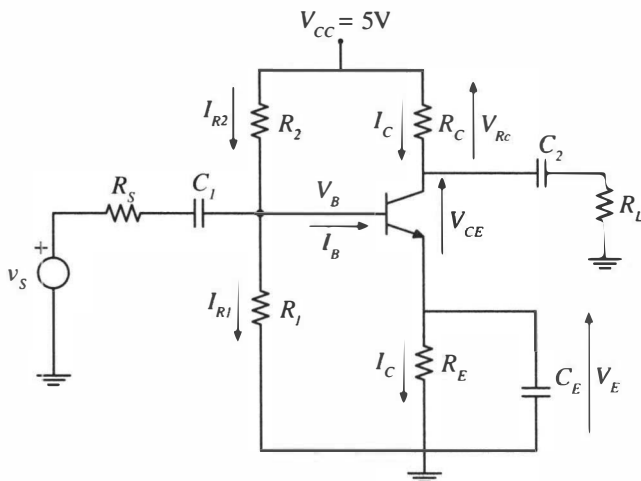
Pertanto

$$R_{IN} = \frac{v_x}{i_x} = \frac{r_\pi + R'_E \left(\frac{r_\pi}{R_3} + \beta_F + 1 \right)}{\frac{r_\pi}{R_3} + 1} \approx \beta_F R'_E$$

*Nota: il circuito in figura è una configurazione denominata **bootstrapping**, che ha il pregio di offrire una grande resistenza di ingresso.*

PROGETTO DI AMPLIFICATORI A BJT

1) Progettare un amplificatore tensione-tensione a singolo BJT in configurazione CE con le seguenti specifiche: $R_S = 50\Omega$, $R_L = 125k\Omega$, $A_{vv} = -50$, $V_{CC} = 5V$.



Consideriamo il circuito in figura. La condizione sul guadagno in tensione impone:

$$A_{vv} = -g_m R_C = -\frac{I_C R_C}{V_T} = -50$$

Quindi

$$V_{RC} = I_C R_C = 50 \cdot (25mV) = 1.25V$$

In generale conviene porre $V_{RC} = V_{CE}$ in modo che la tensione di collettore possa salire o scendere della stessa quantità in dipendenza del segnale in ingresso che ragionevolmente si suppone simmetrico in ampiezza (l'ampiezza massima positiva sarà circa uguale all'ampiezza massima negativa). Fissiamo quindi

$$V_{CE} = 1.25V$$

V_E si ottiene tenendo presente che la somma delle tre tensioni nella maglia di uscita deve essere pari all'alimentazione (5V):

$$V_E = 2.5V$$

Per procedere è necessario fissare la corrente di collettore. Il progettista ha una certa libertà in ciò: fissiamo ad esempio $I_C = 0.2mA$. Per questo valore di corrente otteniamo

$$R_C = \frac{V_{RC}}{I_C} = \frac{1.25\text{V}}{0.2\text{mA}} = 6.25\text{k}\Omega$$

$$R_E \approx \frac{V_E}{I_C} = \frac{2.5\text{V}}{0.2\text{mA}} = 12.5\text{k}\Omega$$

Se il BJT è in regione attiva diretta, la tensione in base è pari a

$$V_B = V_E + 0.6\text{V} = 3.1\text{V}$$

Resta ora da dimensionare R_1 e R_2 in modo che la tensione di base sia proprio 3.1V. Per far ciò conviene che la corrente che scorre in R_1 e R_2 sia molto maggiore della corrente entrante in base, in modo da poter considerare R_1 e R_2 semplicemente un partitore di tensione.

La corrente di base è

$$I_B = \frac{I_C}{\beta_F}$$

β_F non è tra i dati del problema, ma possiamo ragionevolmente supporre un valore > 50 . Supponiamo quindi $\beta_F = 50$. $I_B = 4\mu\text{A}$. Fissiamo quindi

$$I_{R2} = I_{R1} = 10I_B = 40\mu\text{A}$$

Pertanto

$$R_1 = \frac{V_B}{I_{R1}} = \frac{3.1\text{V}}{0.04\text{mA}} = 77.5\text{k}\Omega$$

$$R_2 = \frac{V_{CC} - V_B}{I_{R1}} = \frac{1.9\text{V}}{0.04\text{mA}} = 47.5\text{k}\Omega$$

Verifichiamo la condizione

$$R_{IN} \gg R_S = 50\Omega$$

$$R_{IN} = R_1 \parallel R_2 \parallel r_\pi = (29.45\text{k}\Omega) \parallel r_\pi$$

$$r_\pi = \beta_F \frac{V_T}{I} = \beta_F (125\Omega) = 6.25\text{k}\Omega$$

quindi

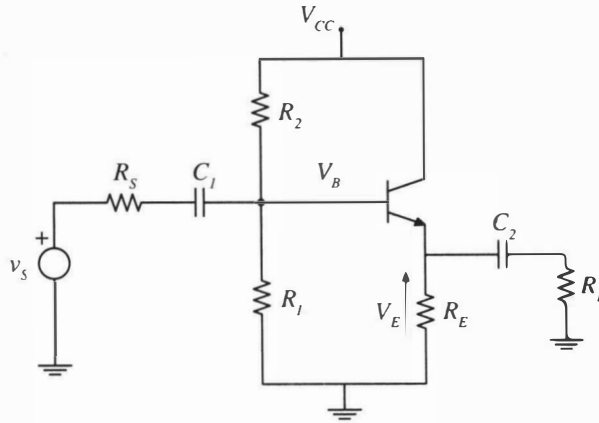
$$R_{IN} = 5.16\text{k}\Omega \gg 50\Omega$$

Verifichiamo la condizione

$$R_{OUT} \ll R_L = 125\text{k}\Omega$$

$$R_{OUT} \approx R_C = 6.25\text{k}\Omega \ll 125\text{k}\Omega$$

2) Progettare un buffer di tensione usando un BJT ($\beta_F = 100$) in configurazione CC con le seguenti specifiche: $V_S = 100\text{mV}$, $R_S = 3\text{k}\Omega$, $R_L = 10\text{k}\Omega$, $V_{CC} = 12\text{V}$.



Fissiamo $V_E = 6\text{V}$.

La condizione di uscita in tensione impone

$$\frac{1}{g_m} \ll R_L$$

$$I_C \gg \frac{V_T}{R_L} = 2.5\mu\text{A}$$

Fissiamo $I_C = 0.25\text{mA}$. Quindi

$$R_E = \frac{V_E}{I_C} = 24\text{k}\Omega$$

Dimensioniamo ora la maglia di ingresso:

$$V_B = V_E + 0.6\text{V} = 6.6\text{V}$$

Resta ora da dimensionare R_1 e R_2 in modo che la tensione di base sia proprio 6.6V . Per far ciò conviene che la corrente che scorre in R_1 e R_2 sia molto maggiore della corrente entrante in base, in modo da poter considerare R_1 e R_2 semplicemente un partitore di tensione.

La corrente di base è

$$I_B = \frac{I_C}{\beta_F}$$

$\beta_F = 100$. $I_B = 2.5\mu\text{A}$. Fissiamo quindi

$$I_{R2} = I_{R1} = 20I_B = 50\mu\text{A}$$

Pertanto

$$R_1 = 132 \text{ k}\Omega$$

$$R_2 = 108 \text{ k}\Omega$$

Verifichiamo la condizione

$$R_{IN} \gg R_s = 3 \text{ k}\Omega$$

$$R_{IN} = R_1 \parallel R_2 = 59.4 \text{ k}\Omega \gg 3 \text{ k}\Omega$$

Verifichiamo ora la condizione sull'ampiezza del segnale in ingresso

$$v_s \ll I_C (R_E \parallel R_L)$$

$$100 \text{ mV} \ll 1.76 \text{ V}$$

7.23 Domande di verifica a scelta multipla

1. Un amplificatore tensione-tensione a singolo BJT in configurazione CE con segnale d'ingresso molto minore di 50mV può essere considerato lineare
 - a) sempre
 - b) purché anche il segnale in uscita sia molto più piccolo di 50mV
 - c) purché il segnale di uscita non superi la dinamica di uscita
 - d) purché la corrente di base dovuta al segnale non diventi molto più piccola della corrente di base di polarizzazione
2. Un buffer di tensione a singolo BJT in configurazione CC può essere considerato lineare
 - a) quando il segnale in ingresso è molto minore di 50mV
 - b) quando il segnale in uscita è molto minore di 50mV
 - c) quando il segnale in ingresso è molto minore della dinamica di uscita
 - d) quando la dinamica di uscita è molto minore di 50mV
3. Una sola delle seguenti affermazioni è vera.
 - a) La resistenza di ingresso di un amplificatore CB è uguale alla resistenza di uscita di un amplificatore CC (a parità di punto di polarizzazione)
 - b) Un amplificatore CE ha resistenza di uscita molto grande
 - c) Un amplificatore CC può essere usato come buffer di corrente
 - d) La resistenza di ingresso di un amplificatore corrente-tensione CE non varia al variare del punto di polarizzazione
4. Una sola delle seguenti affermazioni è vera.
 - a) Il guadagno tensione-corrente di un amplificatore CE si esprime in $(\text{Ohm})^{-1}$
 - b) Il guadagno tensione-tensione di un amplificatore CC è proporzionale alla resistenza di collettore
 - c) Il guadagno corrente-tensione di un amplificatore CB è circa 1
 - d) Con i MOSFET non si possono realizzare buffer di corrente

5. Un amplificatore corrente-tensione
 - a) ha resistenza di ingresso molto più grande della resistenza di uscita del segnale
 - b) ha guadagno espresso da un numero puro
 - c) ha resistenza di uscita molto più piccola della resistenza di carico
 - d) ha guadagno espresso in Volt
6. Un amplificatore tensione-tensione è realizzato ponendo in cascata un amplificatore CE con un amplificatore CB. I due BJT sono polarizzati con due identiche reti a quattro resistori. Il guadagno complessivo è pari a
 - a) 1
 - b) $-g_m R_C/2$
 - c) $-g_m R_C$
 - d) $-R_C/R_E$
7. Un amplificatore a singolo BJT in configurazione CB
 - a) ha resistenza di ingresso molto grande, per questo è impiegato come buffer di tensione
 - b) ha resistenza di uscita molto piccola
 - c) la sua resistenza di ingresso diminuisce all'aumentare della corrente di collettore
 - d) non ha applicazioni pratiche
8. Nel modello per piccoli segnali del BJT
 - a) la resistenza tra base ed emettitore aumenta all'aumentare della corrente di polarizzazione del collettore
 - b) la resistenza di uscita aumenta all'aumentare della corrente di polarizzazione del collettore
 - c) quando aumenta la tensione (di piccolo segnale) tra collettore ed emettitore, la corrente di base aumenta
 - d) quando aumenta la tensione (di piccolo segnale) tra base ed emettitore, la tensione tra collettore ed emettitore diminuisce
9. In un amplificatore tensione-tensione
 - a) il trasferimento di potenza è massimizzato in ingresso e in uscita
 - b) il trasferimento di tensione è minimo in ingresso e in uscita
 - c) la resistenza di uscita è massimizzata
 - d) la resistenza di ingresso è massimizzata
10. Il modello per piccoli segnali del BJT
 - a) non è valido per grandi variazioni della tensione tra collettore ed emettitore
 - b) è valido solo per variazioni di corrente di collettore piccole rispetto alle variazioni della corrente di base
 - c) è valido solo per piccole variazioni della corrente di base
 - d) è valido solo quando la corrente di base è maggiore della corrente di collettore
11. Un amplificatore a singolo BJT in configurazione CC
 - a) ha resistenza di uscita molto piccola
 - b) può essere usato come buffer di corrente
 - c) ha resistenza di ingresso pari a V_T/I_C
 - d) può essere considerato lineare solo se il segnale di ingresso è molto più piccolo di 50mV

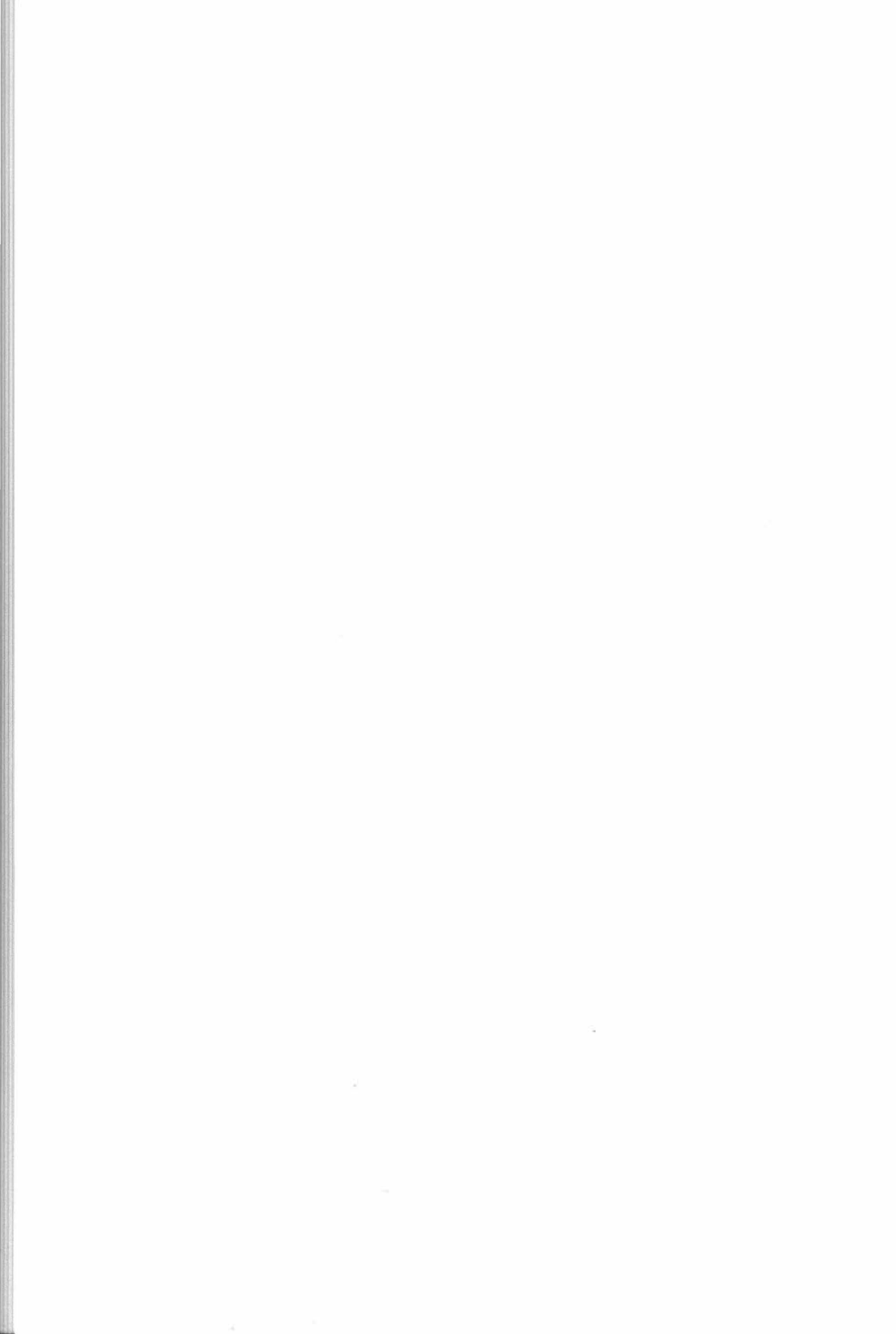
12. Il modello per piccoli segnali di un transistor
- a) contiene elementi non lineari quali generatori di corrente controllati in tensione
 - b) contiene solo resistenze e condensatori
 - c) dipende dal punto di polarizzazione del transistor
 - d) può essere utilizzato solo in circuiti CE (per i BJT) e CS (per i FET)

7.24 Domande di verifica a risposta aperta

1. Si ricavi il modello linearizzato del MOSFET a canale n ad arricchimento
2. Si ricavi il modello del BJT per piccoli segnali.
3. Si dimostri l'espressione del guadagno di un amplificatore a singolo BJT in configurazione CE reazionato.
4. Si dimostri che la resistenza di uscita di un amplificatore a singolo BJT in configurazione CC è pari a $1/g_m$.
5. Si dimostri che il guadagno tensione-tensione di un amplificatore a singolo MOSFET in configurazione CD è 1.
6. Si dimostri che il guadagno corrente-corrente di un amplificatore a singolo MOSFET in configurazione CG è 1.

7.25 Risposte alle domande a scelta multipla

1 c	5 c	9 d
2 c	6 c	10 c
3 a	7 c	11 a
4 a	8 d	12 c



8

Risposta in frequenza degli amplificatori lineari

In questo capitolo si tratta la risposta in frequenza degli amplificatori lineari, definendo il regime delle basse, medie e alte frequenze. Con riferimento ad esempi circuitali si fornisce la regola per la stima della frequenza di taglio inferiore. Si trattano, quindi, le capacità parassite dei FET e dei BJT. Infine, con riferimento ad esempi circuitali si fornisce la regola per la stima della frequenza di taglio superiore.

8.1 Risposta in frequenza degli amplificatori

Nel capitolo precedente ci siamo occupati di amplificatori senza considerare gli effetti di possibili elementi reattivi (capacità e induttanze). Di fatto si è assunto che eventuali capacità fossero o molto grandi da poter essere considerate dei cortocircuiti o molto piccole da poter essere considerate circuiti aperti. Con queste forti ipotesi semplificative gli amplificatori hanno la stessa risposta a qualsiasi frequenza e non hanno limiti né in bassa né in alta frequenza. Si tratta, ovviamente, di condizioni non realistiche. In questo capitolo considereremo l'effetto degli elementi reattivi e vedremo come essi influenzano la risposta degli amplificatori.

8.2 Risposta in frequenza delle reti lineari

Si consideri una generica rete lineare costituita da resistenze, capacitori, induttanze, generatori di tensione (per un richiamo ai concetti base delle reti lineari vedi Appendice C). Un segnale è immesso a due nodi ed estratto da altri due nodi.

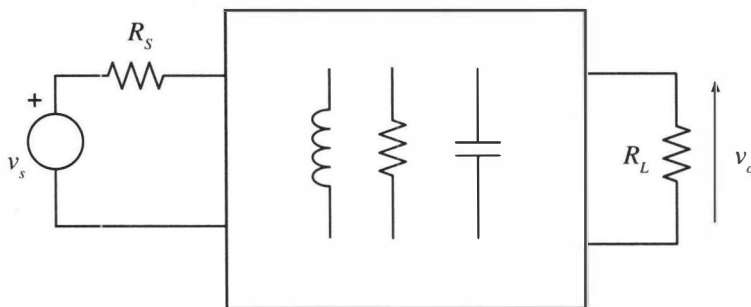


Figura 8.1 Rete lineare

Dalla linearità della rete segue che il segnale in ingresso ha la stessa frequenza di quello in uscita e che pertanto si può definire **risposta in frequenza** in tensione della rete il numero complesso

$$R(\omega) = \frac{v_o}{v_i} = A e^{j\varphi} \quad (8.1)$$

dove v_i e v_o sono le tensioni di ingresso e uscita in notazione complessa (i cosiddetti “fasori”). A è l'ampiezza della risposta e φ la fase. In maniera analoga si

può definire la risposta in frequenza in corrente, tensione-corrente, corrente-tensione. La risposta in frequenza di una rete si rappresenta spesso con i cosiddetti **diagrammi di Bode**, che mostrano separatamente l'andamento in funzione della frequenza dell'ampiezza e della fase.

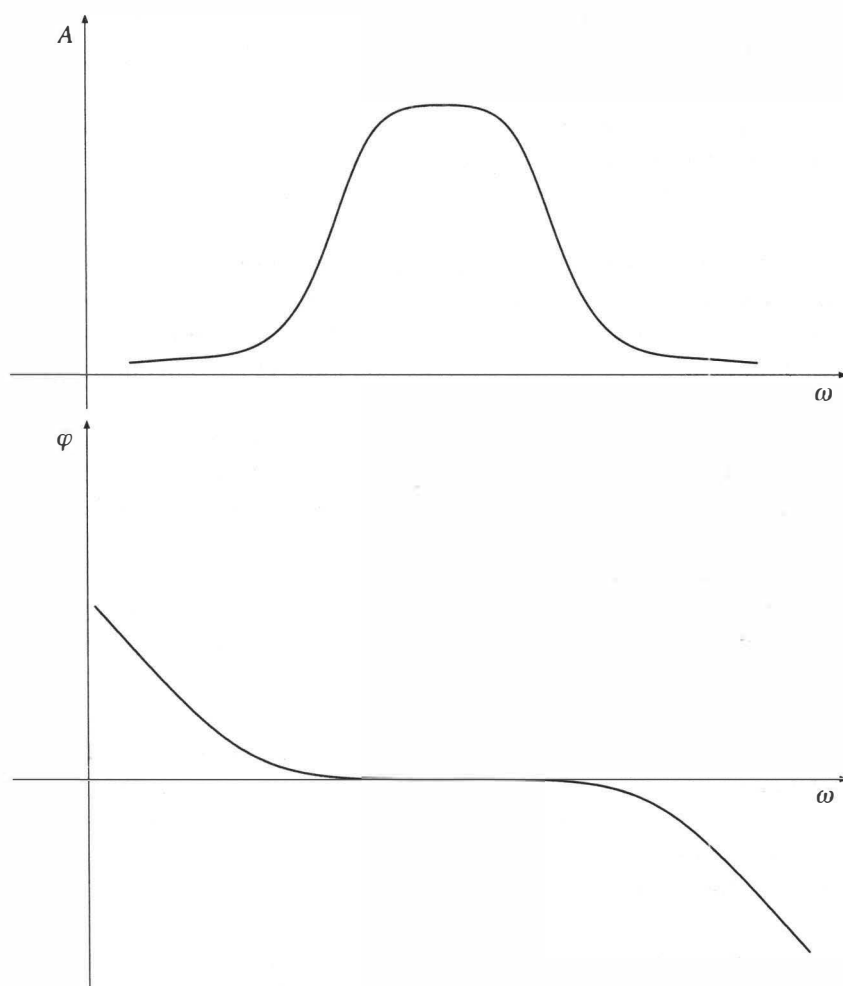


Figura 8.2 Diagrammi di Bode

Si noti che una rete puramente resistiva non può che avere risposta costante in frequenza. L'andamento variabile in frequenza è dato dalla presenza di capacità e induttanze.

8.3 Risposta in frequenza degli amplificatori lineari

Gli amplificatori lineari sono un sottoinsieme delle reti lineari e sono pertanto anch'essi caratterizzati da una risposta in frequenza. Tuttavia sugli amplificatori lineari si possono fare delle assunzioni che vincolano la forma della risposta e ne semplificano l'analisi. Innanzitutto in un amplificatore non ci sono induttanze. Inoltre le capacità possono essere distinte in due classi ben separate:

- 1) grandi capacità dovute ai capacitori di accoppiamento (per l'immissione e l'estrazione del segnale) e ai capacitori di by-pass;
- 2) piccole capacità parassite, per lo più localizzate tra i terminali dei transistor.

Con queste ipotesi possiamo distinguere tre regioni di funzionamento:

- 1) **Medie frequenze.** Le capacità di accoppiamento e di by-pass possono essere considerate dei cortocircuiti. Le capacità parassite possono essere considerate dei circuiti aperti. Di fatto le capacità non intervengono in questa regione e pertanto la risposta è costante in frequenza.
- 2) **Basse frequenze.** Le capacità parassite possono essere considerate dei circuiti aperti, ma le capacità di accoppiamento e di by-pass non possono essere considerate dei cortocircuiti e quindi vanno trattate con i metodi della teoria delle reti.
- 3) **Alte frequenze.** Le capacità di accoppiamento e di by-pass possono essere considerate dei cortocircuiti, ma le capacità parassite non possono essere considerate dei circuiti aperti e quindi vanno trattate con i metodi della teoria delle reti.

Fuori dalla regione delle medie frequenze (detta anche **banda operativa**) l'ampiezza del guadagno cala e lo sfasamento del segnale in uscita diventa dipendente dalla frequenza. La variazione dell'ampiezza rispetto all'ampiezza di centro banda è detta **distorsione di ampiezza**, la variazione della fase rispetto al centro banda è detta **distorsione di fase**.

Si dice **frequenza di taglio** (f_c) la frequenza per cui il quadrato dell'ampiezza della risposta in frequenza $A(f)$ si riduce della metà rispetto al valore massimo, ovvero

$$\left(\frac{A(f)}{A_0} \right)^2 = \frac{1}{2} \quad (8.2)$$

con A_0 ampiezza a centro banda, ovvero al centro della regione delle medie frequenze. La (8.2) è verificata quando l'ampiezza si riduce di un fattore pari a $1/\sqrt{2} = 0.71$, ovvero, in maniera equivalente, quando l'ampiezza si riduce di 3dB. Infatti, dalla (8.2) segue che

$$10\log \left(\left(\frac{A}{A_0} \right)^2 \right) = 10\log \left(\frac{1}{2} \right) \quad (8.3)$$

$$20\log \left(\frac{A}{A_0} \right) = -3\text{dB} \quad (8.4)$$

Poiché l'ampiezza del guadagno diminuisce sia alle alte sia alle basse frequenze, le frequenze di taglio sono due: la frequenza di taglio inferiore (f_L) e la frequenza di taglio superiore (f_H).

8.4 Risposta di un amplificatore di tensione CS alle basse frequenze

Vogliamo ora stimare la frequenza di taglio inferiore dell'amplificatore tensione-tensione in Figura 8.3.

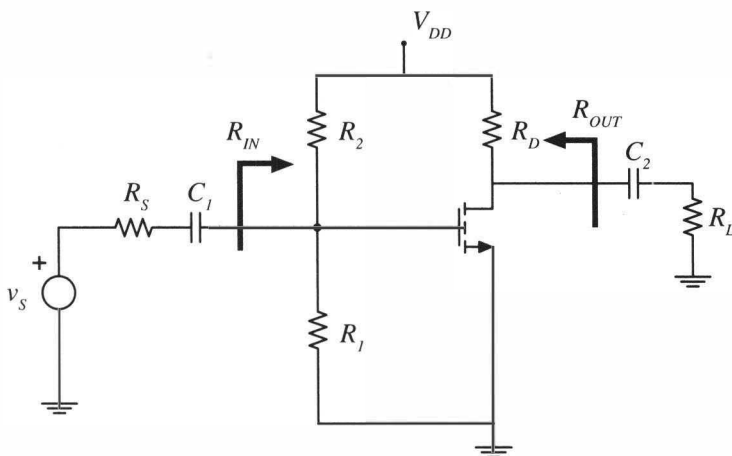


Figura 8.3 Amplificatore CS

Essendo un amplificatore progettato per operare con ingresso in tensione, la resistenza R_s sarà molto minore di R_{IN} e pertanto la maglia di ingresso può essere vista come

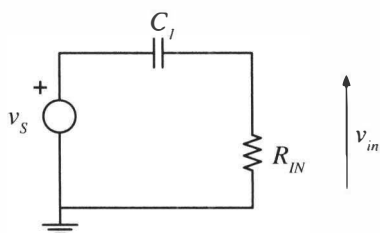


Figura 8.4 Maglia di ingresso

Il circuito in Figura 8.4 è un filtro **passa alto RC** la cui risposta in frequenza (v_{in}/v_s) è quella mostrata in Figura 8.5 (vedi Appendice C).

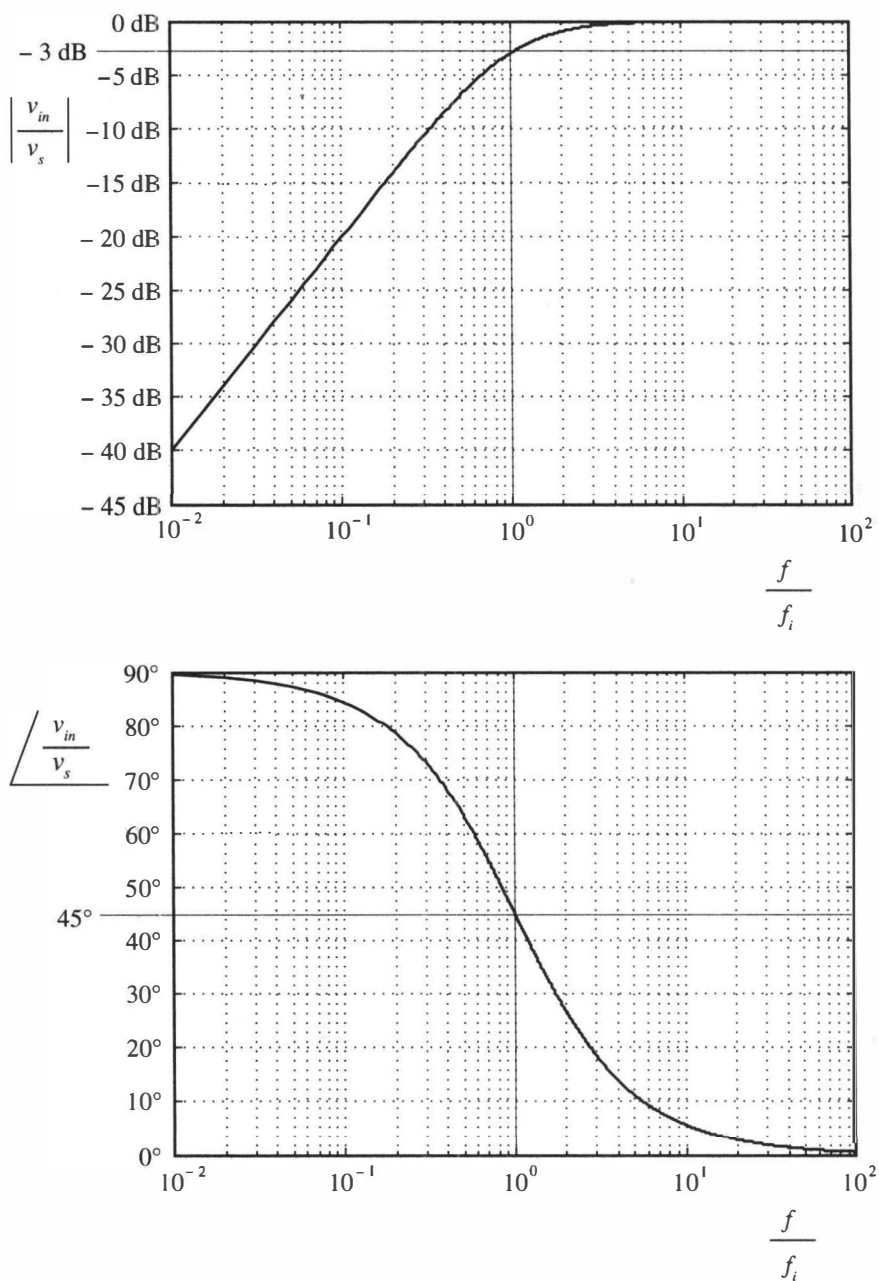


Figura 8.5 Passa alto

La frequenza di taglio per cui il segnale è attenuato di 3dB e sfasato di 45° è pari a

$$f_1 = \frac{1}{2\pi C_1 R_{IN}} \quad (8.5)$$

Anche la maglia di uscita può essere considerata un passa alto la cui frequenza di taglio è

$$f_2 = \frac{1}{2\pi C_2 R_L} \quad (8.6)$$

La frequenza più alta tra le due è già una prima stima della frequenza di taglio dell'amplificatore. Nel seguito vedremo un metodo più accurato, sebbene comunque approssimato per il calcolo della frequenza di taglio inferiore.

8.5 Risposta di un amplificatore CE alle basse frequenze

Vogliamo ora stimare la frequenza di taglio inferiore dell'amplificatore tensione-tensione a BJT in configurazione CE mostrato in Figura 8.6.

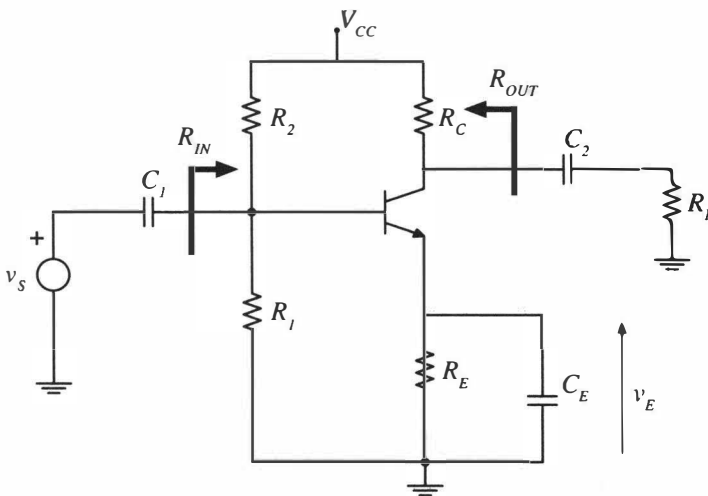


Figura 8.6 Amplificatore a singolo BJT in configurazione CE

I capacitori C_1 e C_2 possono trattati come nel paragrafo precedente e costituiscono dei passa alto con frequenze di taglio

$$f_1 = \frac{1}{2\pi C_1 R_{IN}} \quad (8.7)$$

$$f_2 = \frac{1}{2\pi C_2 R_L} \quad (8.8)$$

Un po' più complesso è l'effetto del capacitore C_E . Innanzitutto si osservi che la tensione v_E è la somma di un termine costante di polarizzazione V_E e un termine variabile v_e : solo quest'ultimo modifica il comportamento in frequenza dell'amplificatore. Il capacitore C_E alle alte frequenze è un cortocircuito e pertanto $v_e \approx 0$; man mano che la frequenza si abbassa, la v_e aumenta fino a raggiungere il suo valore massimo, che è pari a v_s , ovvero il segnale in ingresso. Si noti, infatti, che prendendo v_e come uscita il circuito può essere visto come un buffer di tensione (ingresso in base e uscita sul collettore: il fatto che il collettore non sia cortocircuitato a massa come è tipico dei buffer di tensione a BJT non è essenziale, in quanto un'eventuale resistenza vista "guardando" fuori dal collettore ha un effetto nullo sul comportamento del resto del circuito). L'uscita di tale buffer può essere vista come un generatore di tensione pari al segnale in ingresso e con una resistenza di uscita pari a $1/g_m$. Tale uscita è collegata al carico costituito da R_E in parallelo con C_E . Il circuito per il calcolo di v_e è quindi quello della Figura 8.7.

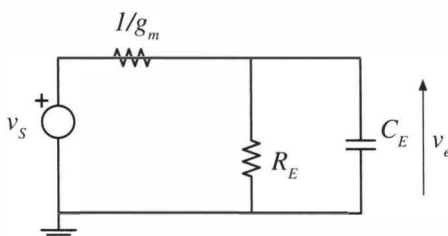


Figura 8.7 Maglia di uscita relativa a v_e

Poiché R_E è una resistenza grande rispetto a $1/g_m$, possiamo considerarla un circuito aperto. Il circuito che ne risulta è un **passa basso RC**, la cui risposta è mostrata in Figura 8.8, e la frequenza di taglio è data da (vedi appendice C):

$$f_E = \frac{1}{2\pi C_E (1/g_m)} \quad (8.9)$$

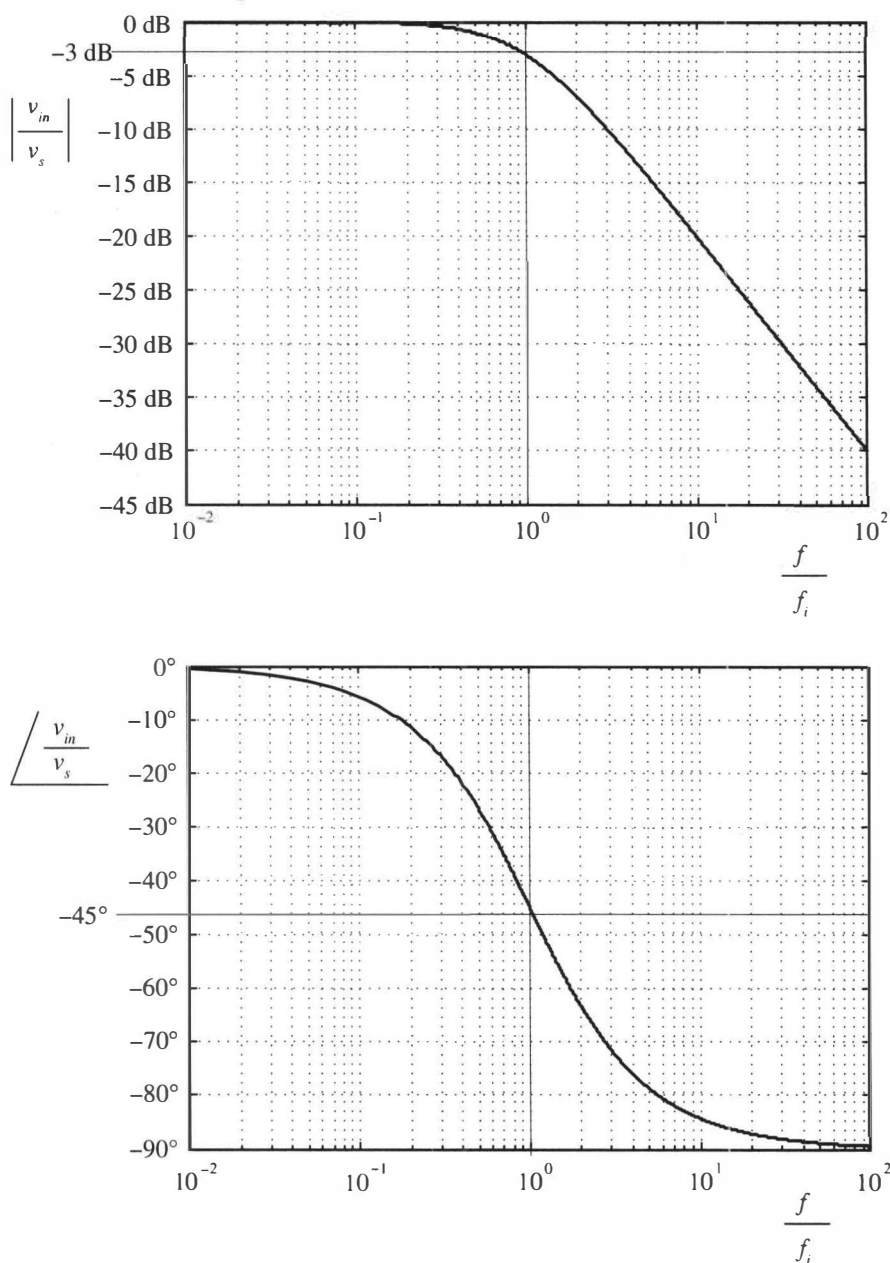


Figura 8.8 Passa basso

Quando v_e è nulla l'amplificatore è un CE e il guadagno in tensione pari a $-g_m R_C$, quando v_e aumenta opera come una reazione negativa che riduce il guadagno. A frequenze molto basse C_E è un circuito aperto e l'amplificatore diventa un CE

reazonato, quindi con guadagno in tensione pari a $-R_C/R_E$. Ne segue che il guadagno rimane alto (a $-g_m R_C$) solo se la frequenza è molto maggiore di f_E .

In definitiva in un CE si possono individuare tre frequenze di taglio. In prima approssimazione la frequenza di taglio dell'amplificatore può essere considerata la più alta delle tre. Nel prossimo paragrafo vedremo un metodo meno approssimato per combinare le tre frequenze.

8.6 Regola generale per la stima della frequenza di taglio inferiore

Nei paragrafi precedenti abbiamo analizzato la risposta in frequenza di due specifici circuiti. In generale le frequenze di taglio di un amplificatore lineare si possono calcolare con esattezza risolvendo il circuito con i metodi della teoria dei circuiti lineari. Si tratta di un calcolo piuttosto tedioso se svolto a mano, ma senza particolari difficoltà. Per questo, di norma, è eseguito utilizzando programmi di simulazione circuitali, come ad esempio SPICE. Tuttavia, per una stima più immediata, si può dimostrare (non del tutto in generale, ma per condizioni piuttosto ragionevoli) che la frequenza di taglio inferiore (f_L) è approssimativamente data da

$$f_L \approx \sum_i f_i \quad (8.10)$$

con

$$f_i = \frac{1}{2\pi C_i R_i} \quad (8.11)$$

dove R_i è la resistenza vista dal capacitore C_i **cortocircuitando tutte le altre capacità**. “Vedere una resistenza” in questo contesto significa sostituire il capacitore con un generatore di prova di tensione v_x , calcolare la corrente i_x che eroga e fare il rapporto v_x/i_x . Si fa presente che questa procedura si applica ai soli capacitori di accoppiamento e di by-pass, mentre le capacità parassite rimangono dei circuiti aperti e pertanto sono ignorate.

8.7 Amplificatori alle alte frequenze

I modelli di amplificatori analizzati fin qui non hanno elementi che limitano il guadagno ad alta frequenza, ovvero potrebbero operare a frequenza qualsivoglia alta. Nella realtà gli amplificatori hanno inevitabilmente un limite superiore in frequenza e questo è un parametro critico in molte applicazioni quali ad esempio i sistemi di trasmissione e ricezione radio. Tale limite è dovuto alle capacità parassite. Queste sono delle piccole capacità che a bassa frequenza si comporta-

no come grandi impedenze e quindi sono assimilabili a circuiti aperti. Ad alta frequenza, invece, l'impedenza si riduce fino a dar luogo a un cortocircuito.

Le capacità parassite possono localizzarsi nelle connessioni, nelle saldature e nei connettori, ma, in una scheda ben progettata, le più critiche sono quelle del transistor stesso.

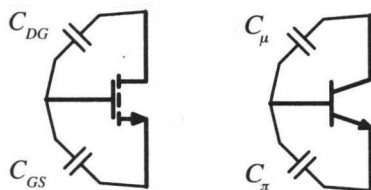


Figura 8.9 Passa basso

Tali capacità quando passano da circuiti aperti a cortocircuiti possono: cortocircuitare a massa l'ingresso o l'uscita, oppure cortocircuitare l'ingresso con l'uscita. Nei tre casi l'effetto è quello di impedire all'amplificatore di operare propriamente e quindi all'aumentare della frequenza ridurre il guadagno fino ad annullarlo.

8.8 Capacità tra gate e source nel MOSFET in saturazione

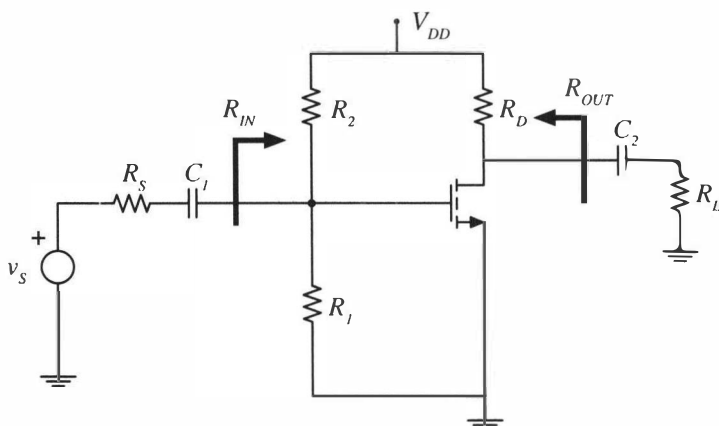
Un MOSFET è essenzialmente costituito da un capacitore MOS. La capacità del MOS è collegata da un lato al gate e dall'altro al canale. Quando il MOSFET è in saturazione è come se il canale fosse collegato solo con il source, e tra gate e source è presente una capacità parassita pari a circa la capacità del MOS:

$$C_{GS} \approx \frac{\epsilon}{d} WL \quad (8.12)$$

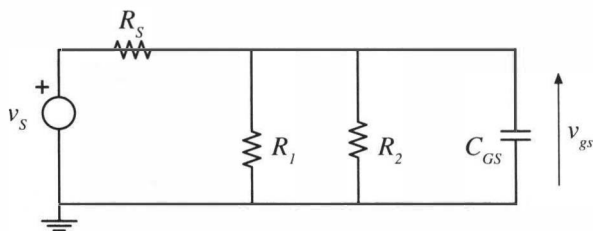
con W larghezza del canale, L lunghezza, d spessore dell'ossido, ϵ costante dielettrica dell'ossido.

8.9 Risposta dell'amplificatore CS ad alta frequenza

Stimiamo ora la frequenza di taglio superiore (f_H) di un amplificatore tensione-tensione a singolo MOSFET in configurazione CS.

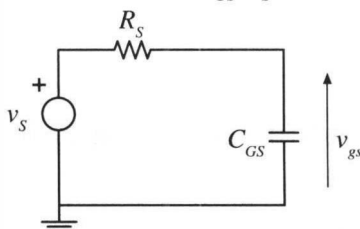
**Figura 8.10** Amplificatore CS

I capacitori di accoppiamento e di by-pass si considerano cortocircuiti, così come l'alimentazione si cortocircuita a massa. La maglia di ingresso è quindi il circuito in Figura 8.11.

**Figura 8.11** Maglia di ingresso

Se l'ingresso è in tensione il parallelo di R_1 e R_2 è molto più grande di R_S e quindi può essere trascurato. Il risultato è un passa basso RC. La cui frequenza di taglio a -3dB è

$$f_H = \frac{1}{2\pi C_{GS} R_S} \quad (8.13)$$

**Figura 8.12** Passa basso

In conclusione la frequenza di taglio superiore è data dalla capacità C_{GS} e dalla resistenza di uscita del generatore di segnale. In generale è sempre vantaggioso che la resistenza di uscita di un dispositivo con uscita in tensione (generatore di segnale, amplificatore, sensore...) sia molto bassa per due motivi: 1) perché massimizza il trasferimento di tensione; 2) perché aumenta la frequenza di taglio superiore dello stadio successivo e quindi allarga la banda operativa.

8.10 Modello del MOSFET ad alta frequenza

La capacità tra gate e source è, di solito, la principale causa della limitazione in frequenza dei MOSFET. Tuttavia ci sono altre capacità il cui effetto non è trascurabile ad alta frequenza. Con riferimento alla Figura 8.13, tra gate e le piazzole S o D c'è sempre una piccola capacità di sovrapposizione (*overlapping*). Inoltre le regioni di svuotamento delle giunzioni pn contropolarizzate costituiscono delle capacità tra S e B e tra D e B.

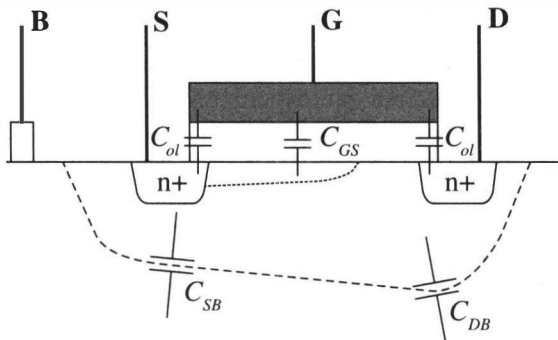


Figura 8.13 Capacità parassite del MOSFET in saturazione

Oltre a tali capacità, è necessario considerare delle resistenze parassite in serie ai tre terminali. Queste sono dovute a possibili cadute ohmiche nel gate, o nelle piazzole di drain e source. In particolare la resistenza di gate può essere molto critica perché è in serie con la resistenza del generatore di segnale e quindi contribuisce direttamente a diminuire la frequenza di taglio superiore. Si tratta comunque di una resistenza che può essere resa molto piccola aumentando il drogaggio del polisilicio che costituisce il gate e realizzando con cura i contatti elettrici. Il modello completo del MOSFET a tre terminali ad alta frequenza che tiene conto di tutte le possibili resistenze e capacità è mostrato in Figura 8.14.

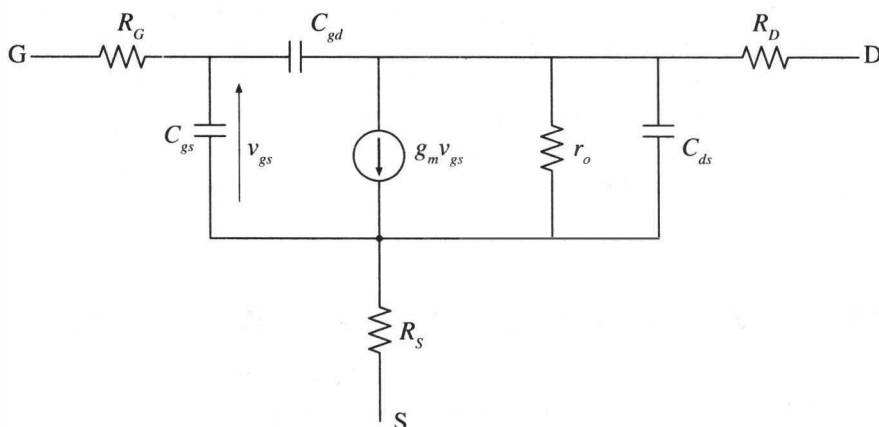


Figura 8.14 Modello ad alta frequenza del MOSFET a 3 terminali

Questo è il modello che viene implementato nei simulatori circuitali.

8.11 Capacità parassite del BJT

Vediamo ora le capacità parassite di un BJT. Un BJT in regione attiva diretta può essere visto come due giunzioni pn rispettivamente in polarizzazione diretta (BE) e in polarizzazione inversa (BC).

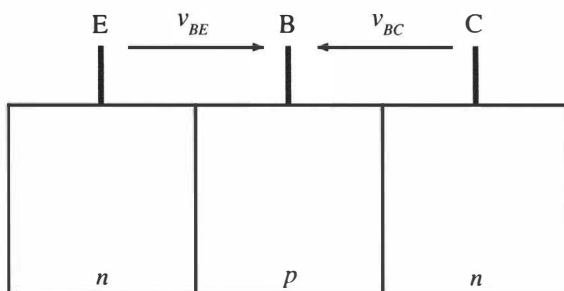


Figura 8.15 BJT npn

In generale una giunzione pn polarizzata inversamente dà luogo a una capacità dovuta alla carica positiva e negativa ai due lati della regione di svuotamento.

Nel Capitolo 2 abbiamo già calcolato lo spessore della regione di svuotamento nelle due zone n e p, in particolare:

$$x_n = \sqrt{\frac{2\varepsilon}{q}(\varphi + v_{CB}) \frac{N_A}{N_D} \frac{1}{N_A + N_D}} \quad (8.14)$$

dove ϵ è la costante dielettrica del silicio, q la carica dell'elettrone, φ il potenziale di giunzione, N_A concentrazione delle impurità p, N_D concentrazione delle impurità n.

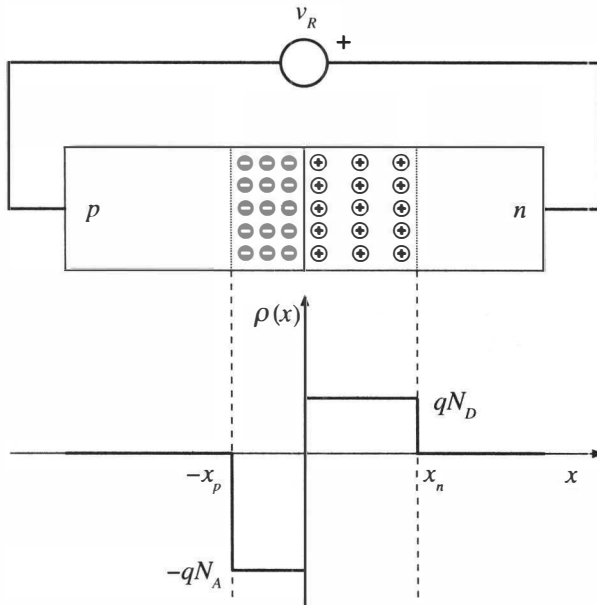


Figura 8.16 Regione di svuotamento di una giunzione pn contropolarizzata

La carica positiva accumulata nella zona n è pari a

$$Q = qSN_D x_n = qSN_D \sqrt{\frac{2\epsilon}{q}(\varphi + v_{CB})} \frac{N_A}{N_D} \frac{1}{N_A + N_D} \quad (8.15)$$

dove S è la sezione della giunzione.

La carica non dipende linearmente dalla tensione applicata (v_{CB}), pertanto a rigore la giunzione non è schematizzabile come un capacitore lineare, ma per piccole variazioni della tensione v_{CB} , ovvero per piccoli segnali, possiamo approssimare la capacità tra collettore e base (C_μ) con la derivata della carica Q rispetto alla tensione v_{CB} (capacità differenziale).

$$C_\mu \approx \left. \frac{dQ}{dv_{CB}} \right|_{v_{CB} = V_{CB}} \quad (8.16)$$

Sostituendo la (8.15) nella (8.16) e indicando con w_0 lo spessore della regione di svuotamento quando la tensione applicata è nulla:

$$w_0 = \sqrt{\frac{2\varepsilon}{q} \varphi \left(\frac{1}{N_A} + \frac{1}{N_D} \right)} \quad (8.17)$$

si ottiene la notevole equazione

$$C_\mu = \frac{C_0}{\sqrt{1 + \frac{V_{CB}}{\varphi}}} \quad (8.18)$$

con

$$C_0 = \varepsilon \frac{S}{w_0} \quad (8.19)$$

Il risultato trovato ha un immediato significato fisico: la capacità della regione di svuotamento in assenza di polarizzazione è equivalente alla capacità di un condensatore a facce piane e parallele la cui separazione è pari allo spessore della regione di svuotamento.

Tra base ed emettitore è presente un'altra capacità (C_π) parassita dovuta all'accumulo di carica in base. I portatori minoritari (provenienti dall'emettitore) in base hanno una carica complessiva pari a

$$Q = \frac{qSn(0)W}{2} \quad (8.20)$$

con W larghezza utile della base e

$$n(0) = \left(\frac{n_i^2}{N_B} \right) e^{\frac{V_{BE}}{V_T}} \quad (8.21)$$

q è la carica dell'elettrone, n_i è la concentrazione degli elettroni nel silicio intrinseco, N_B è la concentrazione delle impurità in base.

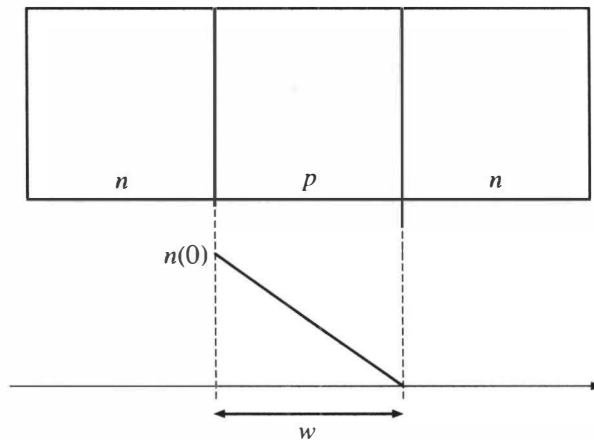


Figura 8.17 Portatori minoritari in base

Anche in questo caso la carica non dipende linearmente dalla tensione applicata (v_{BE}), pertanto a rigore la giunzione non è schematizzabile come un condensatore lineare, ma per piccole variazioni della tensione v_{BE} , ovvero per piccoli segnali, possiamo approssimare:

$$C_{\pi} \approx \left. \frac{dQ}{dv_{BE}} \right|_{v_{BE} = V_{BE}} \quad (8.22)$$

Sostituendo la (8.21) nella (8.22) e ricordando che

$$I_C = \frac{qSD_n}{w} \left(\frac{n_i^2}{N_B} \right) e^{\frac{v_{BE}}{V_T}} \quad (8.23)$$

dove D_n è il coefficiente di diffusione dei portatori n, si ottiene

$$C_{\pi} = \frac{1}{2} \frac{w^2}{D_n} \left(\frac{I_C}{V_T} \right) \quad (8.24)$$

Ricordando che

$$g_m = \frac{I_C}{V_T} \quad (8.25)$$

e definendo **tempo di transito**

$$\tau = \frac{1}{2} \frac{w^2}{D_n} \quad (8.26)$$

si ottiene l'espressione

$$C_{\pi} = \tau g_m \quad (8.27)$$

che mostra come la capacità parassita tra base ed emettitore sia linearmente dipendente da g_m .

8.12 Modello del BJT ad alta frequenza

Inseriamo le capacità parassite direttamente nel modello linearizzato del BJT come mostrato in Figura 8.18.

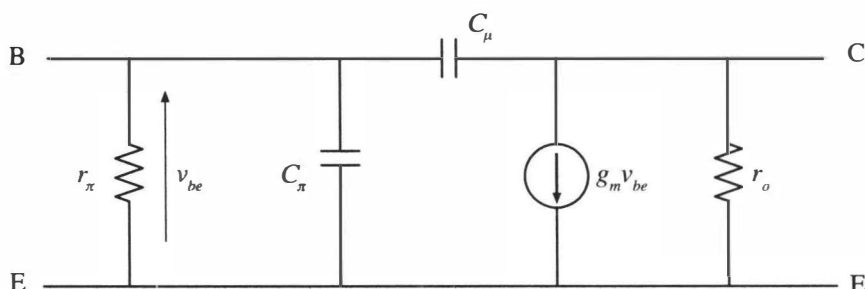


Figura 8.18 Modello del BJT ad alta frequenza

Questo modello non è ancora del tutto soddisfacente ad alta frequenza, perché quando le capacità parassite diventano dei cortocircuiti (o comunque cominciano a condurre una corrente significativa) è necessario tener conto della cosiddetta **resistenza di spreading**. Come mostrato in Figura 8.19, la corrente di base prima di entrare nella zona delle due giunzioni deve attraversare un percorso non breve. Ciò può essere visto come una resistenza in serie r_x tra il reoforo B e la giunzione come mostrato in Figura 8.20, che è appunto il modello per piccoli segnali ad alta frequenza che useremo nel seguito. r_x ha tipicamente un valore intorno ai 100 Ohm, è quindi trascurabile rispetto a r_π (che è dell'ordine di qualche k Ω) se tutta la corrente che scorre su r_x scorre anche su r_π . Ciò non è più vero quando le due capacità parassite entrano in conduzione.

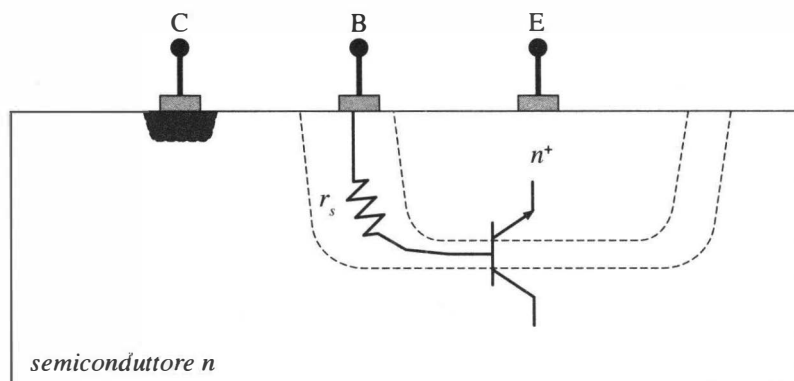


Figura 8.19 Resistenza di spreading

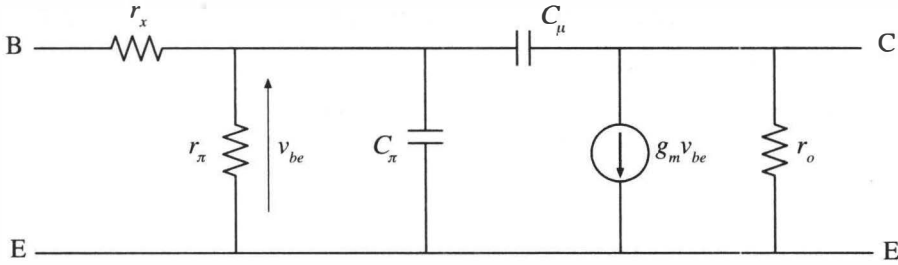


Figura 8.20 Modello AC del BJT ad alta frequenza

Il modello in Figura 8.20 è più che soddisfacente per la maggior parte delle applicazioni e nel seguito è quello a cui faremo riferimento, tuttavia il modello completo dovrebbe tener conto sia delle resistenze in serie parassite agli altri due terminali, sia della resistenza in parallelo a C_μ , dovuta alla corrente inversa di saturazione della giunzione contropolarizzata tra C e B.

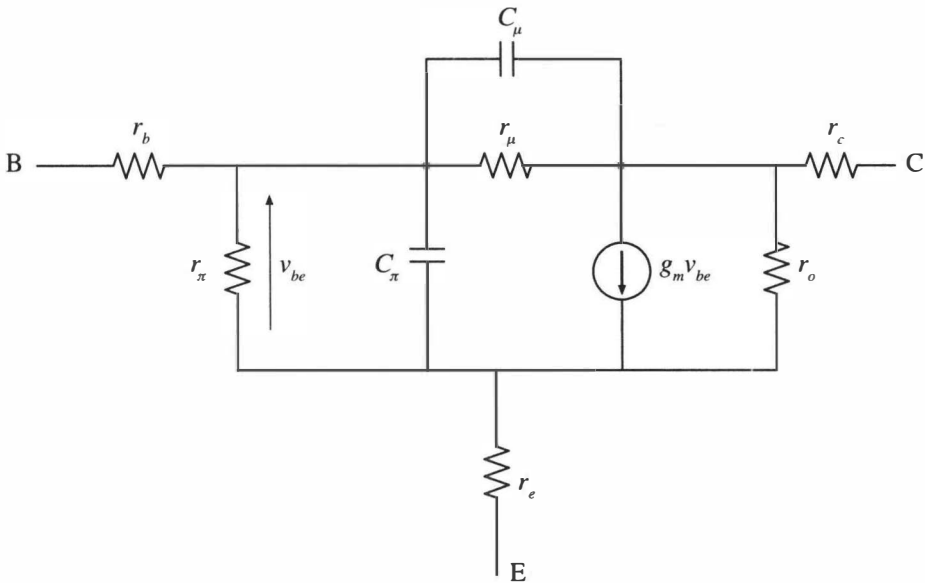


Figura 8.21 Modello AC completo

In Figura 8.21 è mostrato il modello completo usato nei simulatori circuitali.

8.13 Risposta dell'amplificatore CE ad alta frequenza

Stimiamo ora la frequenza di taglio superiore (f_H) di un amplificatore tensione-tensione a BJT in configurazione CE.

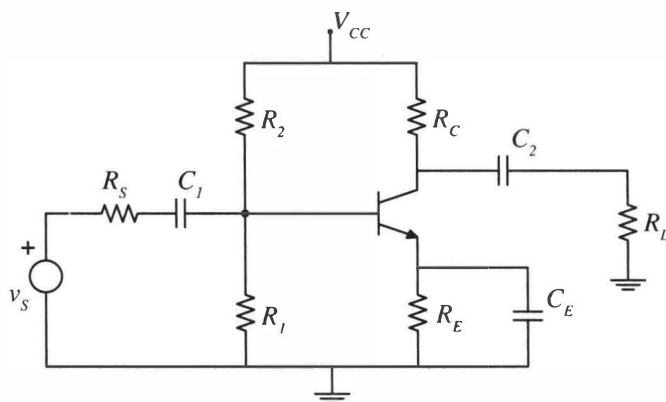


Figura 8.22 Amplificatore a singolo BJT in configurazione CE

Il circuito equivalente dopo aver cortocircuitato l'alimentazione di tutti i condensatori di accoppiamento e di by-pass è quello in Figura 8.23.

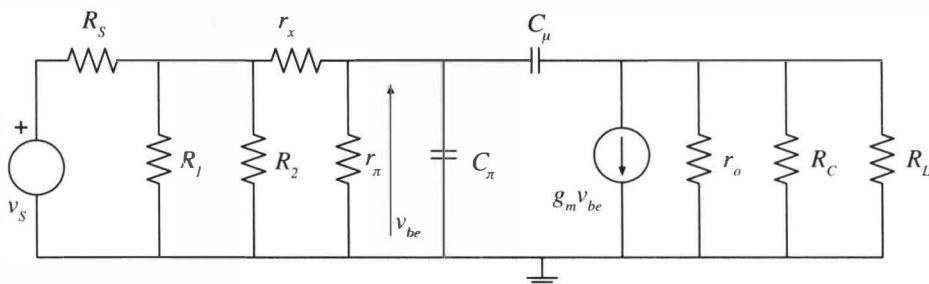


Figura 8.23 Modello linearizzato ad alta frequenza dell'amplificatore CE

Per l'analisi di questo circuito conviene applicare il teorema di Miller. Nel caso specifico il teorema afferma che il circuito di Figura 8.24 è equivalente al circuito di Figura 8.25.

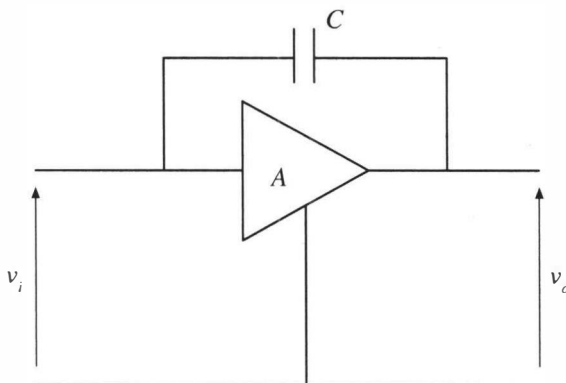


Figura 8.24 Teorema di Miller

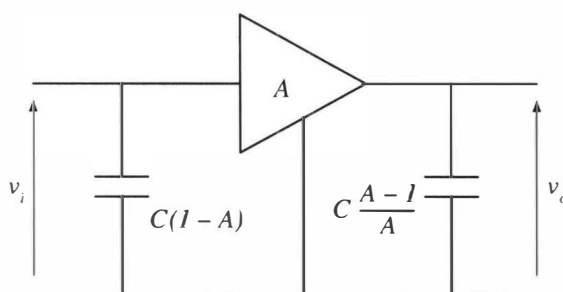


Figura 8.25 Teorema di Miller

Applicando quindi il teorema di Miller alla capacità C_μ e considerando $g_m R_C \parallel R_L \gg 1$ si ottiene il circuito di Figura 8.26. Si noti che la capacità C_μ è riportata nella maglia di ingresso moltiplicata per il guadagno in tensione (**effetto Miller**). Si tratta di un effetto importante che spesso è la principale causa che limita la banda operativa degli amplificatori di tensione ad alto guadagno.

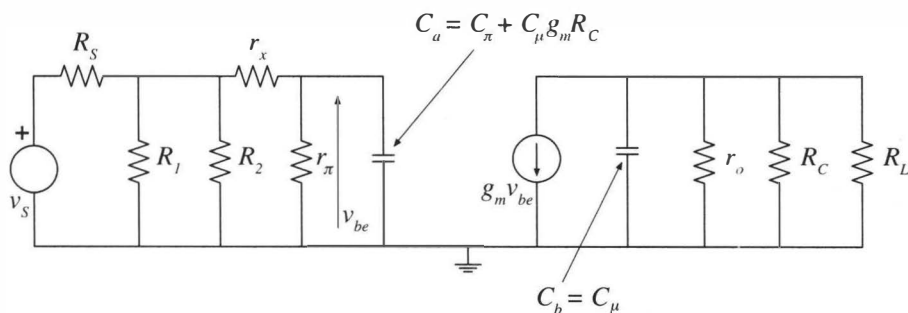


Figura 8.26 Modello linearizzato ad alta frequenza dell'amplificatore CE

Poiché R_1 , R_2 , r_π sono grandi rispetto a R_S e r_x , il capacitore C_a costituisce un passa basso RC con $R_S + r_x$ e la sua frequenza di taglio è

$$f_a = \frac{1}{2\pi C_a (R_S + r_x)} \quad (8.28)$$

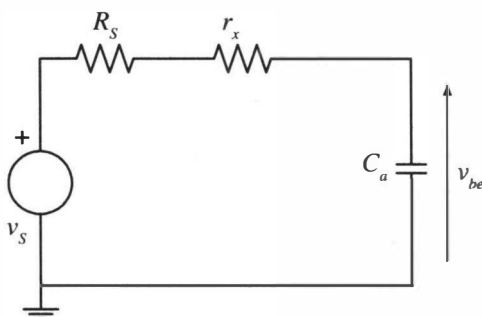


Figura 8.27 Maglia di ingresso

Se l'uscita è in tensione, la resistenza di uscita $R_{OUT} = R_C$ è molto più piccola di R_L , pertanto la maglia di uscita può essere vista come in Figura 8.28, che è un passa basso con frequenza di taglio

$$f_b = \frac{1}{2\pi C_b R_C} \quad (8.29)$$

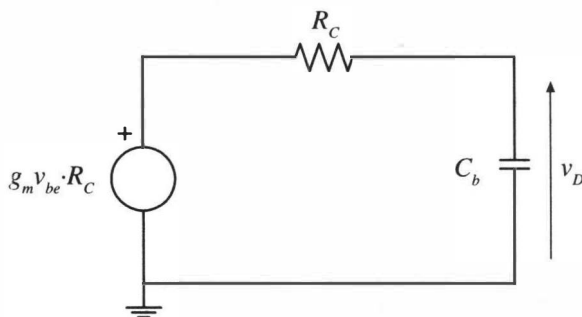


Figura 8.28 Maglia di uscita

Per una prima stima della frequenza di taglio dell'amplificatore nel suo complesso, possiamo prendere la minore tra le due frequenze. Nel paragrafo successivo verrà fornita una formula per una stima migliore.

8.14 Regola generale per la stima della frequenza di taglio superiore

Si può dimostrare (non del tutto in generale, ma per condizioni piuttosto ragionevoli) che la frequenza di taglio superiore (f_H) è approssimativamente data da

$$\frac{1}{f_H} \approx \sum_a \frac{1}{f_a} \quad (8.30)$$

con

$$f_a = \frac{1}{2\pi C_a R_a} \quad (8.31)$$

dove R_a è la resistenza vista dal capacitore C_a **trasformando in circuiti aperti tutte le altre capacità**. La (8.30) si applica alle sole capacità parassite, i capacitori di accoppiamento e di by-pass alle alte frequenze rimangono dei cortocircuiti e pertanto devono essere ignorati.

8.15 Cascode

Nei paragrafi precedenti abbiamo visto che l'effetto Miller può essere la principale causa che limita la risposta in frequenza di un amplificatore ad alto guadagno. Una possibile soluzione per estendere la banda di un amplificatore CE è il cosiddetto **cascode**, ovvero porre un BJT in configurazione CB in cascata al BJT in configurazione CE, come ad esempio nell'amplificatore tensione-tensione il cui circuito è mostrato in Figura 8.29.

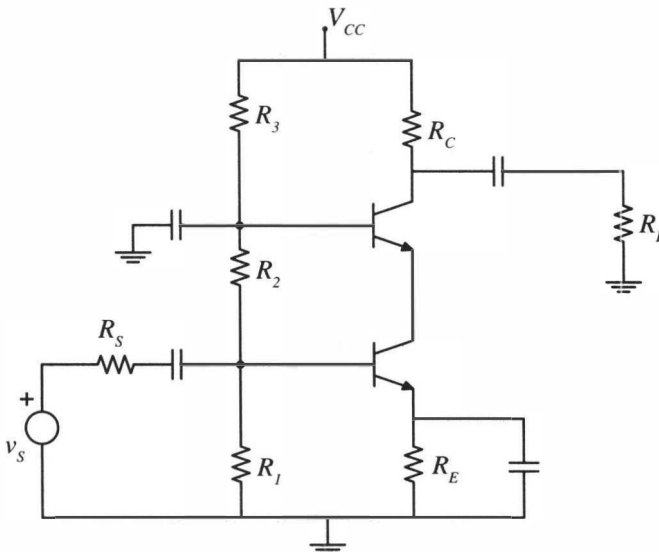


Figura 8.29 Cascode

Prima di analizzarne il comportamento in frequenza calcoliamo il guadagno. Se non consideriamo l'effetto della rete di polarizzazione sul guadagno (ovvero supponiamo la rete di polarizzazione ben progettata), il circuito in AC si riduce a quello mostrato in Figura 8.30.

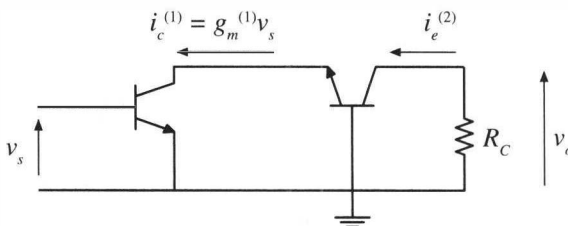


Figura 8.30 Cascode

La corrente che tira il primo BJT è $g_m v_s$. D'altra parte, nel secondo BJT la corrente di collettore deve essere uguale alla corrente di emettitore, pertanto

$$v_o = -R_C i_e^{(2)} = -R_C g_m^{(2)} v_s \quad (8.32)$$

Quindi

$$A_{vv} = \frac{v_o}{v_s} = -R_C g_m^{(2)} \quad (8.33)$$

Passiamo ora all'analisi in frequenza. Il circuito AC ad alta frequenza è mostrato in Figura 8.31.

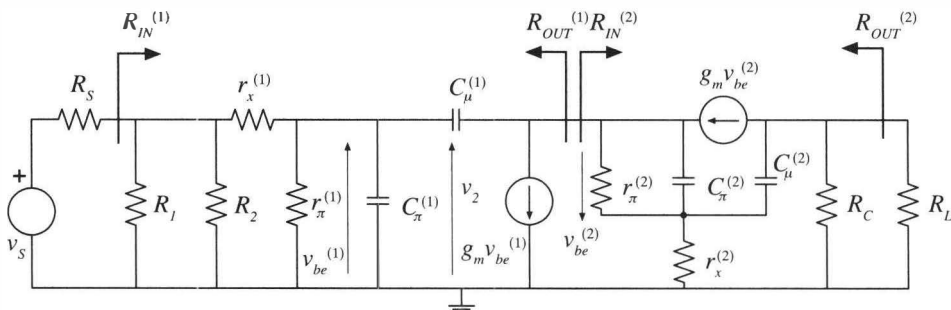


Figura 8.31 Cascode AC

L'amplificazione in tensione del primo stadio è data da

$$A_{vv}^{(1)} = \frac{v_2}{v_{be}^{(1)}} = -g_m^{(1)} R_{IN}^{(2)} \quad (8.34)$$

Il secondo stadio è un CC e quindi ha resistenza di ingresso pari a $1/g_m^{(2)}$. Pertanto:

$$A_{vv}^{(1)} = -\frac{g_m^{(1)}}{g_m^{(2)}} \approx 1 \quad (8.35)$$

Applichiamo ora il teorema di Miller al circuito in Figura 8.31, considerando anche $R_s \ll R_I \parallel R_2$ e $R_L \gg R_{OUT}^{(2)}$. Si ottiene il circuito in Figura 8.32.

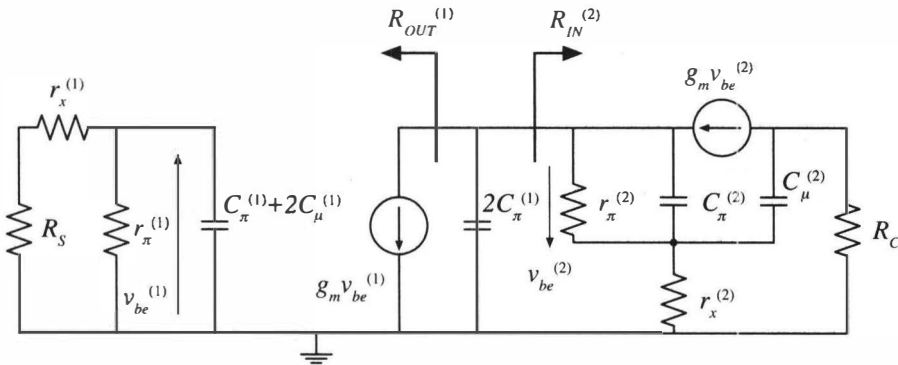


Figura 8.32 Circuito AC del cascode

Poiché il guadagno in tensione dello stadio CE è basso (circa unitario), l'effetto Miller non riporta in ingresso una capacità moltiplicata per un grande fattore, ma semplicemente C_μ moltiplicata per due.

Calcoliamo ora le frequenze associate a ciascuno dei quattro capacitori. La resistenza che vede ciascun capacitore si calcola considerando tutti gli altri dei circuiti aperti.

$$f_a = \frac{1}{2\pi (C_\pi^{(1)} + 2C_\mu^{(1)}) (r_\pi^{(1)} \parallel (r_x^{(1)} + R_s))} \approx \frac{1}{2\pi (C_\pi^{(1)} + 2C_\mu^{(1)}) (r_x^{(1)} + R_s)} \quad (8.36)$$

$$f_b = \frac{1}{2\pi(2C_\mu^{(1)})(R_{OUT} \parallel R_{IN}^{(2)})} \approx \frac{g_m^{(2)}}{2\pi(2C_\mu^{(1)})} \quad (8.37)$$

$$f_c = \frac{1}{2\pi C_\pi^{(2)} r_\pi^{(2)}} \quad (8.38)$$

$$f_d = \frac{1}{2\pi C_\mu^{(2)}(r_x^{(2)} + R_C)} \approx \frac{1}{2\pi C_\mu^{(2)} R_C} \quad (8.39)$$

Applicando la (8.30) si ottiene

$$\frac{1}{f_H} = \frac{1}{f_a} + \frac{1}{f_b} + \frac{1}{f_c} \quad (8.40)$$

Tipicamente le tre capacità sono dello stesso ordine di grandezza. Di solito fra le tre resistenze viste dai capacitori, R_C è decisamente più grande delle altre e quindi per una prima approssimazione possiamo scrivere:

$$f_H \approx f_c = \frac{1}{2\pi R_C C_\mu^{(2)}} \quad (8.41)$$

Si noti che senza lo stadio CB interposto tra l'amplificatore CE e la resistenza R_C la frequenza di taglio superiore sarebbe stata

$$f_H \approx \frac{1}{2\pi(R_x^{(1)} + R_s)C_\mu^{(2)}g_m^{(1)}R_C} \quad (8.42)$$

In definitiva la soluzione cascode permette di alzare la frequenza di taglio superiore di un fattore pari a $g_m^{(1)}(R_x^{(1)} + R_s)$.

8.16 Riepilogo

- La frequenza di taglio inferiore di un amplificatore è determinata dalla presenza dei capacitori di accoppiamento e di by-pass. Tale frequenza può essere stimata come

$$f_L \approx \sum_i f_i$$

dove

$$f_i = \frac{1}{2\pi C_i R_i}$$

con C_i i-esimo capacitore e R_i resistenza che vede l'i-esimo capacitore, considerando gli altri capacitori (attivi a bassa frequenza) dei cortocircuiti.

- La frequenza di taglio superiore di un amplificatore è determinata dalle capacità parassite dei transistor e può essere stimata come

$$\frac{1}{f_H} \approx \sum_{\alpha} \frac{1}{f_{\alpha}}$$

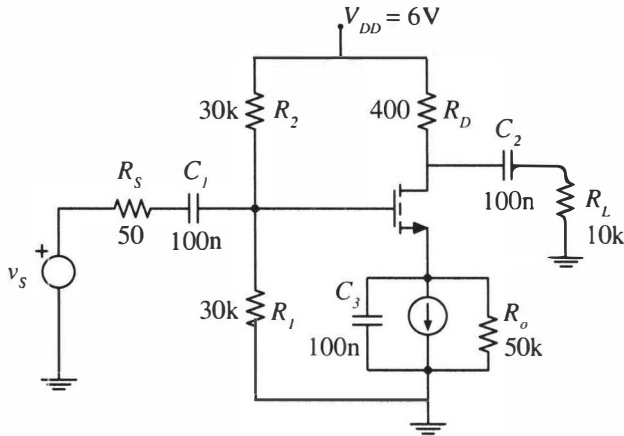
dove

$$f_{\alpha} = \frac{1}{2\pi C_{\alpha} R_{\alpha}}$$

con C_{α} α -esimo capacitore e R_{α} resistenza che vede l' α -esimo capacitore, considerando gli altri capacitori (attivi ad alta frequenza) dei circuiti aperti.

8.17 Esercizi svolti

1) Calcolare le frequenze di taglio superiore e inferiore dell'amplificatore tensione-tensione in figura. $I_0 = 5\text{mA}$, $K_n = 10\text{mA/V}^2$, $V_{TN} = 1\text{V}$, $\lambda = 0.01\text{V}^{-1}$, $C_{GS} = 60\text{pF}$.



La frequenza di taglio associata a C_1 è

$$f_1 = \frac{1}{2\pi C_1 (R_s + R_1 \parallel R_2)} \approx \frac{1}{2\pi C_1 (R_1 \parallel R_2)} = 106\text{Hz}$$

La frequenza di taglio associata a C_2 è

$$f_2 = \frac{1}{2\pi C_2 \left(R_L + R_D \parallel \left(\frac{1}{\lambda I_o} \right) \right)} \approx \frac{1}{2\pi C_2 R_L} = 159\text{Hz}$$

La frequenza di taglio associata a C_3 è

$$f_3 = \frac{1}{2\pi C_3 \left(R_o \parallel \left(\frac{1}{g_m} \right) \right)} \approx \frac{1}{2\pi C_3 \left(\frac{1}{g_m} \right)} = 16\text{kHz}$$

La frequenza di taglio inferiore dell'amplificatore è

$$f_L = f_1 + f_2 + f_3 \approx f_3 = 16\text{kHz}$$

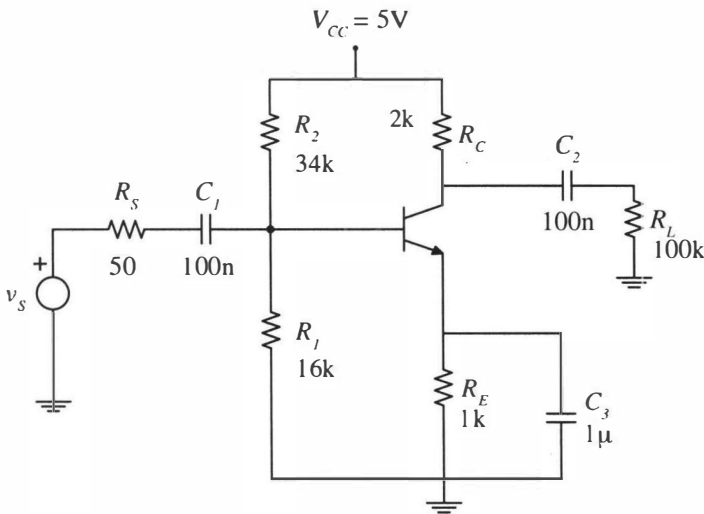
La frequenza di taglio associata a C_{GS} è

$$f_{GS} = \frac{1}{2\pi C_{GS} (R_1 \parallel R_2 \parallel R_s)} \approx \frac{1}{2\pi C_{GS} (R_s)} = 53\text{MHz}$$

La frequenza di taglio superiore dell'amplificatore è

$$f_H = f_{GS} = 53 \text{ MHz}$$

2) Calcolare le frequenze di taglio superiore e inferiore dell'amplificatore tensione-tensione in figura. $\beta_F = 200$, $V_A = 80\text{V}$, $C_{be} = 30\text{pF}$, $C_{cb} = 3\text{pF}$.



La frequenza di taglio associata a C_1 è

$$f_{f_1} = \frac{1}{2\pi C_1 (R_s + R_1 \parallel R_2 \parallel r_\pi)} \approx \frac{1}{2\pi C_1 r_\pi} = 306\text{Hz}$$

La frequenza di taglio associata a C_2 è

$$f_2 = \frac{1}{2\pi C_2 \left(R_L + R_C \parallel \left(\frac{V_A}{I_C} \right) \right)} \approx \frac{1}{2\pi C_2 R_L} = 16\text{Hz}$$

La frequenza di taglio associata a C_3 è

$$f_3 = \frac{1}{2\pi C_3 \left(R_E \parallel \left(\frac{1}{g_m} \right) \right)} \approx \frac{1}{2\pi C_3 \left(\frac{1}{g_m} \right)} = 6.4\text{kHz}$$

La frequenza di taglio inferiore dell'amplificatore è

$$f_L = f_1 + f_2 + f_3 \approx f_3 = 6.4\text{kHz}$$

Applicando il teorema di Miller si ottengono due frequenze di taglio rispettivamente nella maglia di ingresso e nella maglia di uscita. La frequenza di taglio associata alla maglia di ingresso è

$$f_a = \frac{1}{2\pi(C_{be} + (1 - A_{vv})C_{cb})(R_1 \parallel R_2 \parallel R_s)} \approx \frac{1}{2\pi(C_{be} + |A_{vv}|C_{cb})(R_s)} = 11.8\text{MHz}$$

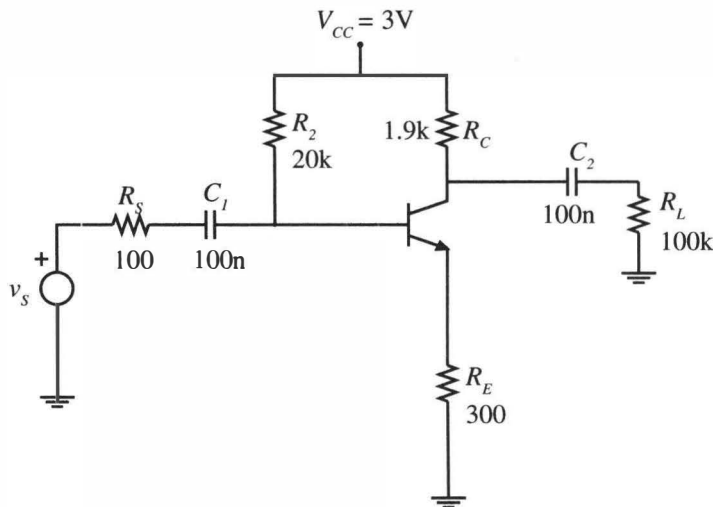
La frequenza di taglio associata alla maglia di uscita è

$$f_b = \frac{1}{2\pi\left(\frac{(1 - A_{vv})}{-A_{vv}}C_{cb}\right)(R_C \parallel R_L)} \approx \frac{1}{2\pi C_{cb}R_C} = 26.5\text{MHz}$$

Pertanto la frequenza di taglio superiore dell'amplificatore è

$$f_H = \frac{1}{\frac{1}{f_a} + \frac{1}{f_b}} = 8.16\text{MHz}$$

3) Calcolare le frequenze di taglio superiore e inferiore dell'amplificatore tensione-tensione in figura. $\beta_F = 200$, $V_A = 80\text{V}$, $C_{be} = 30\text{pF}$, $C_{cb} = 3\text{pF}$.



Applicando il teorema di Miller, la resistenza R_B è riportata in ingresso divisa per $(1 - A_{vv})$. Trattandosi di un CE reazionato:

$$A_{vv} \approx \frac{R_C}{R_E} = 6.33$$

La frequenza di taglio associata a C_1 è

$$f_1 = \frac{1}{2\pi C_1 \left(R_S + \frac{R_B}{1 - A_{vv}} \parallel \left(r_\pi (1 + g_m R_E) + R_E \right) \right)} \approx \frac{1}{2\pi C_1 \frac{R_B}{1 - A_{vv}}} = 583 \text{ Hz}$$

La frequenza di taglio associata a C_2 è

$$f_1 = \frac{1}{2\pi C_2 \left(R_L + R_C \parallel \left(\frac{V_A}{I_C} \right) \right)} \approx \frac{1}{2\pi C_2 R_L} = 838 \text{ Hz}$$

Quindi la frequenza di taglio inferiore dell'amplificatore è

$$f_L = f_1 + f_2 = 1.42 \text{ kHz}$$

Calcoliamo ora la frequenza di taglio superiore. Applicando il teorema di Miller si ottengono due frequenze di taglio rispettivamente nella maglia di ingresso e nella maglia di uscita. La frequenza di taglio associata alla maglia di ingresso è

$$f_a = \frac{1}{2\pi (C_{be} + (1 - A_{vv}) C_{cb}) (R_s \parallel (r_\pi (1 + g_m R_E) + R_E))} \approx \frac{1}{2\pi (C_{be} + (1 - A_{vv}) C_{cb}) R_s} = 30.6 \text{ MHz}$$

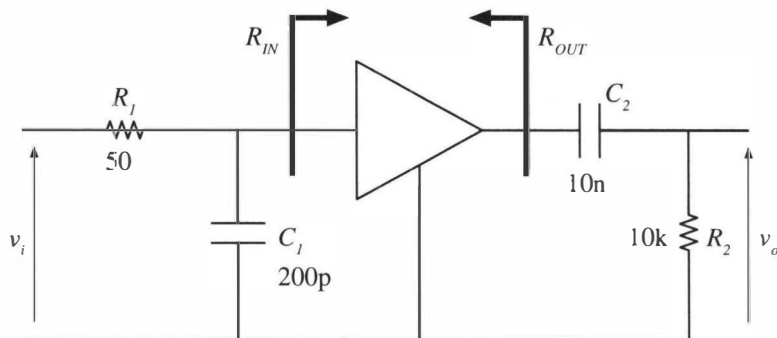
La frequenza di taglio associata alla maglia di uscita è

$$f_b = \frac{1}{2\pi \left(\frac{(1 - A_{vv})}{-A_{vv}} C_{cb} \right) (R_C \parallel R_L)} \approx \frac{1}{2\pi C_{cb} R_C} = 27.9 \text{ MHz}$$

La frequenza di taglio superiore dell'amplificatore è

$$f_H = \frac{1}{\frac{1}{f_a} + \frac{1}{f_b}} = \frac{f_a f_b}{f_a + f_b} = 14.6 \text{ MHz}$$

4) Calcolare le frequenze di taglio inferiore e superiore del circuito in figura. L'amplificatore è un tensione-tensione con resistenza d'ingresso $1\text{M}\Omega$, resistenza di uscita 50Ω , $A_{vv} = -50$, $f_L = 100\text{Hz}$, $f_H = 100\text{MHz}$.



Il capacitore C_i ha funzione di passa basso. La frequenza di taglio associata a C_i è

$$f_l = \frac{1}{2\pi C_i R_i \parallel R_{IN}} \approx \frac{1}{2\pi C_i R_i} = 15.9\text{MHz}$$

Pertanto la frequenza di taglio superiore è

$$f_H^{(0)} \approx \frac{1}{\frac{1}{f_l} + \frac{1}{f_H}} = 13.7\text{MHz}$$

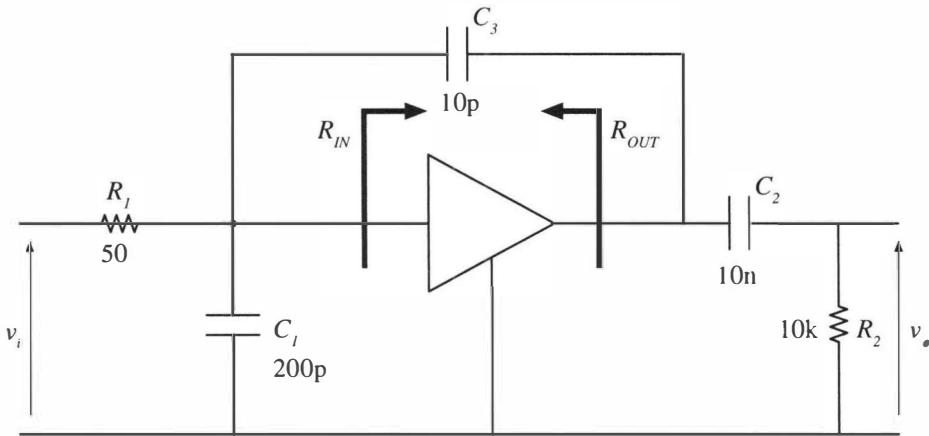
Il capacitore C_2 ha funzione di passa alto. La frequenza di taglio associata a C_2 è

$$f_2 = \frac{1}{2\pi C_2 (R_2 + R_{OUT})} \approx \frac{1}{2\pi C_2 R_2} = 1.6\text{kHz}$$

La frequenza di taglio inferiore è quindi

$$f_L^{(0)} \approx f_l + f_2 = 1.7\text{kHz}$$

5) Calcolare le frequenze di taglio inferiore e superiore del circuito in figura. L'amplificatore è un tensione-tensione con resistenza d'ingresso $1\text{M}\Omega$, resistenza di uscita 50Ω , $A_{vv} = -50$, $f_L = 100\text{Hz}$, $f_H = 100\text{MHz}$.



Applicando il teorema di Miller, C_3 è riportato in ingresso moltiplicato per il guadagno. La frequenza di taglio del passa basso in ingresso è

$$f_l = \frac{1}{2\pi(C_1 + (1 - A_{vv})C_3)R_i \parallel R_{IN}} \approx \frac{1}{2\pi(C_1 + |A_{vv}|C_3)R_i} = 4.55\text{MHz}$$

Il capacitore C_3 è riportato in parallelo all'uscita. Costituisce quindi un passa basso in uscita la cui frequenza di taglio è

$$f_3 = \frac{1}{2\pi C_3(R_{OUT} \parallel R_2)} \approx \frac{1}{2\pi C_3 R_{OUT}} = 318\text{MHz}$$

La frequenza di taglio superiore è quindi

$$f_H^{(0)} \approx \frac{1}{\frac{1}{f_l} + \frac{1}{f_3} + \frac{1}{f_H}} = 4.29\text{MHz}$$

Si tenga presente che la formula applicata è comunque un'approssimazione piuttosto grossolana la cui accuratezza non deve essere sopravvalutata. In questo caso, poiché una delle tre frequenze (4.55MHz) è molto più bassa delle altre, una stima ugualmente accettabile può essere $f_H^{(0)} = 4.55\text{MHz}$.

Il capacitore C_2 ha funzione di passa alto. La frequenza di taglio associata a C_2 è

$$f_2 = \frac{1}{2\pi C_2(R_2 + R_{OUT})} \approx \frac{1}{2\pi C_2 R_2} = 1.6\text{kHz}$$

La frequenza di taglio inferiore è quindi

$$f_L^{(0)} \approx f_i + f_L = 1.7\text{kHz}$$

8.18 Domande di verifica a scelta multipla

1. La frequenza massima alla quale un amplificatore a BJT può operare è determinata
 - a) dai condensatori di accoppiamento del segnale in ingresso e in uscita
 - b) dalle capacità parassite del BJT
 - c) dalle capacità parassite dei reofori
 - d) dalla capacità che si mette tra alimentazione e massa per ridurre i disturbi indotti dall'alimentazione
2. La frequenza minima alla quale un amplificatore può operare è determinata
 - a) dai condensatori di accoppiamento
 - b) dalle capacità parassite del transistor
 - c) dalle capacità parassite dei reofori
 - d) dalla capacità che si mette tra alimentazione e massa per ridurre i disturbi indotti dall'alimentazione
3. La capacità associata a una giunzione pn contropolarizzata
 - a) aumenta all'aumentare della contropolarizzazione, perché la regione di svuotamento è più grande
 - b) è pari alla capacità di un condensatore a facce piane e parallele di superficie pari alla sezione della giunzione e distanza tra le placche pari alla larghezza complessiva della regione di svuotamento
 - c) è un cortocircuito a bassa frequenza
 - d) è minore della capacità associata alla giunzione non contropolarizzata

8.19 Domande di verifica a risposta aperta

1. Si ricavi l'espressione della capacità associata a una giunzione pn contropolarizzata.
2. Si ricavi l'espressione per la capacità parassita tra base ed emettitore in un BJT.

8.20 Risposte alle domande a scelta multipla

1 b	2 a	3 d
-----	-----	-----

Circuiti con capacitori

A.1 Capacitori e capacità elettrica

Prendiamo in considerazione un condensatore costituito da due placche metalliche piane e parallele separate da un materiale isolante (eventualmente anche aria o il vuoto) come in Figura A.1.

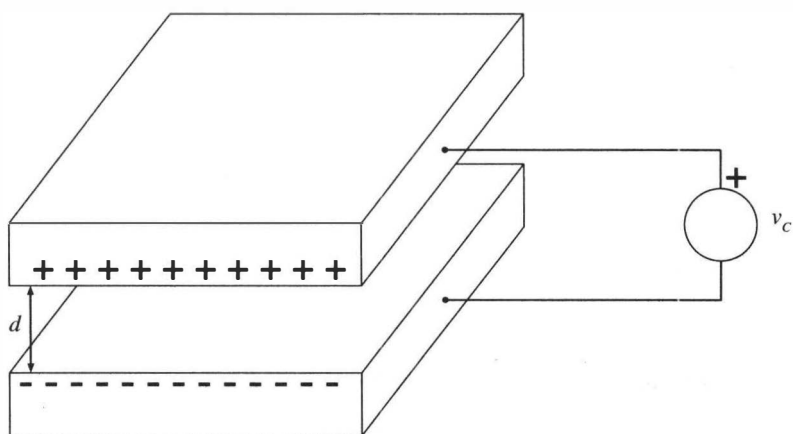


Figura A.1 Capacitore a facce piane e parallele

Applicando un generatore di tensione v_c alle due placche, le cariche elettriche nel metallo si muoveranno fino alla configurazione di equilibrio, che consiste in una distribuzione uniforme di cariche sulle due facce contrapposte: positive su una placca e negative sull'altra. La carica complessiva su ciascuna faccia è pari a

$$Q = C v_c \quad (\text{A.1})$$

dove C è la capacità elettrica del condensatore ed è data da

$$C = \epsilon \frac{S}{d} \quad (\text{A.2})$$

con ϵ costante dielettrica dell'isolante interposto ($\epsilon = \epsilon_0 = 8.8544 \times 10^{-12}$ C/m se si tratta di aria o vuoto), S superficie delle placche e d distanza tra le placche.

Dalla (A.1) si ottiene che

$$i_c = \frac{dQ}{dt} = C \frac{dv_c}{dt} \quad (\text{A.3})$$

che vale per un generico capacitore il cui simbolo circuitale è mostrato in Figura A.2.

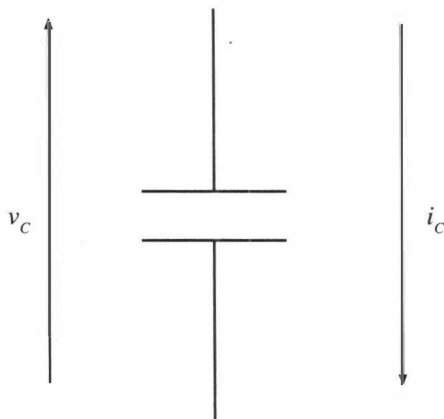


Figura A.2 Capacitore

A.2 Carica di un capacitore

Prendiamo in considerazione il circuito in Figura A.3.

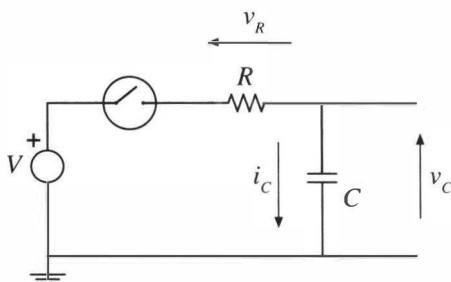


Figura A.3 Circuito per la carica di un capacitore

Per $t < 0$ il capacitore è scarico, ovvero la tensione ai capi delle placche è nulla. All'istante $t = 0$, l'interruttore passa da aperto a chiuso e il generatore di tensione carica il capacitore. Calcoliamo l'andamento nel tempo della tensione v_c sul capacitore. Con riferimento alle convenzioni adottate in figura per i segni delle correnti

$$i_R = i_c \quad (\text{A.4})$$

Applichiamo la Legge di Ohm al resistore di resistenza R e la (A.3) alla corrente del capacitore:

$$\frac{v_R}{R} = C \frac{dv_C}{dt} \quad (\text{A.5})$$

La somma delle tensioni sul capacitore e sul resistore è pari alla tensione V del generatore di tensione, quindi

$$\frac{V - v_C}{R} = C \frac{dv_C}{dt} \quad (\text{A.6})$$

che può essere riscritta come

$$\tau \frac{dv_C}{dt} + v_C = V \quad (\text{A.7})$$

con $\tau = RC$. Questa è un'equazione differenziale di primo grado la cui generica soluzione è

$$v_C = A + B e^{-\frac{t}{\tau}} \quad (\text{A.8})$$

Le costanti A e B si determinano con le condizioni al contorno. La prima condizione è che per $t = 0$, $v_C = 0$. Da questa segue che $A = -B$. La seconda condizione è che per $t \rightarrow \infty$, $v_C = V$. Sostituendo nella (A.8) si ottiene la soluzione che cercavamo:

$$v_C(t) = V \left(1 - e^{-\frac{t}{\tau}} \right) \quad (\text{A.9})$$

il cui andamento è mostrato in Figura A.4.

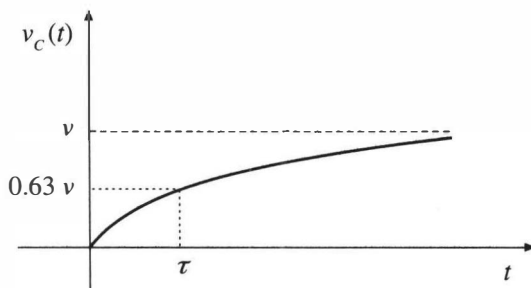


Figura A.4 Carica del capacitore attraverso un resistore

A.3 Scarica di un condensatore

Vediamo ora come il condensatore si scarica. Prendiamo in considerazione il circuito in Figura A.5.

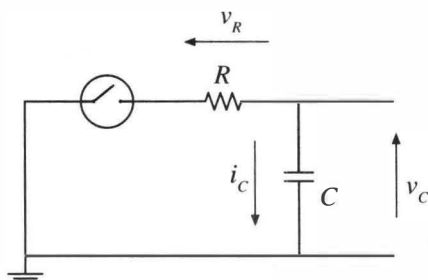


Figura A.5 Scarica del condensatore attraverso un resistore

Supponiamo che $t < 0$, il condensatore sia carico alla tensione V e l'interruttore sia OFF. All'istante $t = 0$ l'interruttore è ON e il condensatore si scarica attraverso la resistenza R . Ragionando come sopra si ottiene che l'espressione di v_C nel tempo è

$$v_C(t) = Ve^{-\frac{t}{\tau}} \quad (\text{A.10})$$

il cui andamento è mostrato in Figura A.6.

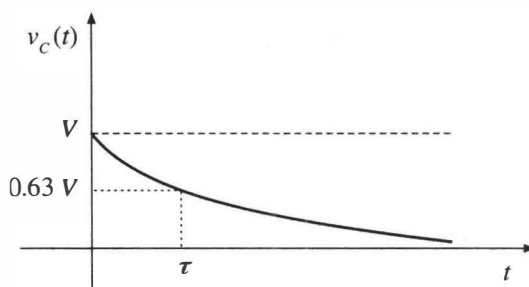


Figura A.6 Scarica del condensatore attraverso un resistore

Circuiti lineari

B.1 Circuiti lineari

Un circuito lineare è una rete in di componenti le cui equazioni costitutive sono lineari. Il caso più semplice è una rete di resistori, per i quali la tensione e la corrente sono legate dalla legge di Ohm. Tali resistori sono supposti essere collegati da fili elettrici di conducibilità infinita. Inoltre potranno essere presenti:

- 1) generatori di tensione indipendenti;
- 2) generatori di corrente indipendenti;
- 3) generatori di tensione controllati linearmente da tensioni o correnti su resistori;
- 4) generatori di corrente controllati linearmente da tensioni o correnti su resistori.

Nelle Figure B.1, B.2, B.3 sono mostrati i simboli circuitali del resistore, del generatore di tensione e del generatore di corrente.

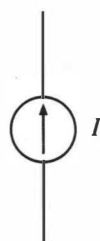
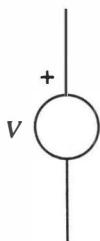
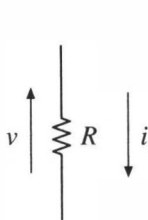


Figura B.1 Resistore **Figura B.2** Generatore di tensione **Figura B.3** Generatore di corrente

In Figura B.4 è mostrato un esempio di circuito lineare. Per tutti i resistori vale

$$v_k = R_k i_k \quad (\text{B.1})$$

Per **nodo** si intende una sezione equipotenziale del circuito: di fatto un filo elettrico di resistenza nulla che connette due o più resistenze è un nodo. I nodi sono indicati con le lettere A, B, C, D, E. Per **maglia** si intende, invece, un qualunque percorso chiuso della corrente, ovvero un percorso che inizia e termina nello stesso nodo.

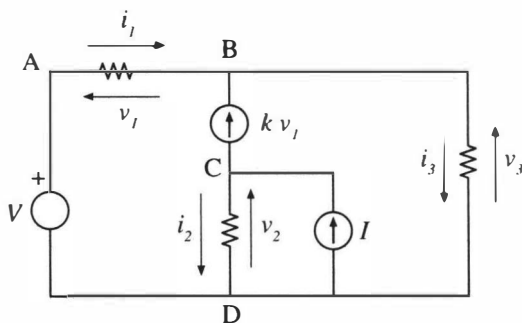


Figura B.4 Circuito lineare

B.2 Prima legge di Kirchhoff

La prima legge di Kirchhoff afferma che in un qualunque nodo di un circuito elettrico la somma algebrica delle correnti che vi confluiscono è nulla. Con riferimento al circuito in Figura B.4, se, ad esempio, applichiamo la legge al nodo B, sommando le correnti entranti e sottraendo quelle uscenti otteniamo

$$i_1 + k v_1 - i_3 = 0 \quad (\text{B.2})$$

B.3 Seconda legge di Kirchhoff

La seconda legge di Kirchhoff afferma che la somma delle tensioni lungo una qualsiasi maglia di una rete elettrica è nulla. Con riferimento al circuito in Figura B.4, se, ad esempio, applichiamo la legge alla maglia ABCE otteniamo

$$V - v_1 - v_3 = 0 \quad (\text{B.3})$$

B.4 Resistenza vista da due nodi

Per calcolare la resistenza vista da due nodi di un circuito lineare si procede come segue:

- 1) si cortocircuitano tutti i generatori di tensione indipendenti;
- 2) si aprono tutti i generatori di corrente indipendenti;
- 3) si applica un generatore di tensione v_x di prova ai due nodi e si calcola la corrente i_x che eroga;
- 4) la resistenza ai due nodi è $R_x = v_x / i_x$.

A titolo di esempio calcoliamo la resistenza vista dai nodi B e D del circuito in Figura B.4. Dopo aver cortocircuitato il generatore di tensione indipendente e aperto il generatore di corrente indipendente si ottiene il circuito in Figura B.5.

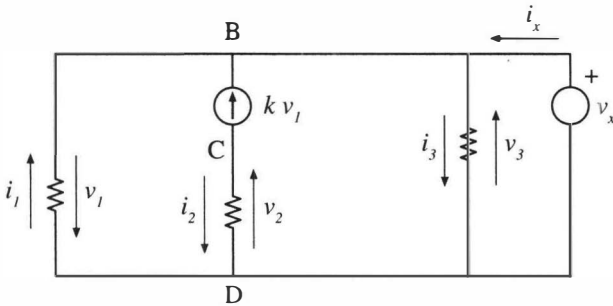


Figura B.5 Calcolo della resistenza R_x

Applichiamo la prima legge di Kirchhoff al nodo B.

$$i_x = -i_1 - K v_1 + i_3 \quad (\text{B.4})$$

da cui

$$i_x = -\frac{v_1}{R_1} - K v_1 + \frac{v_3}{R_3} \quad (\text{B.5})$$

D'altra parte

$$v_1 = -v_x \quad (\text{B.6})$$

$$v_3 = v_x \quad (\text{B.7})$$

Sostituendo nella (B.5)

$$i_x = \frac{v_x}{R_1} + K v_x + \frac{v_x}{R_3} \quad (\text{B.8})$$

si può calcolare i_x . Da cui

$$R_x = \frac{v_x}{i_x} = \frac{1}{\frac{1}{R_1} + K + \frac{1}{R_3}} \quad (\text{B.9})$$

B.5 Teorema di Thévenin

Il teorema di Thévenin afferma che un circuito lineare, visto da due nodi, è equivalente a un generatore di tensione in serie con un resistore (circuito equivalente di Thévenin).

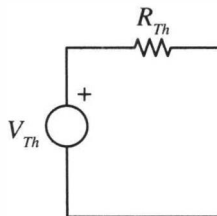


Figura B.6 Circuito equivalente di Thévenin

Per calcolare l'equivalente di Thévenin di un generico circuito si procede come segue:

- 1) si calcola la tensione ai due nodi e si ottiene V_{th} ;
- 2) si calcola la resistenza vista ai due nodi e si ottiene R_{th} .

A titolo di esempio calcoliamo l'equivalente del circuito in Figura B.4 visto dai nodi B e D.

Applicando le regole di Kirchoff si ottiene che la tensione V_{th} ai nodi C e E è pari a

$$V_{th} = \frac{k + \frac{1}{R_I}}{\frac{1}{R_I} + \frac{1}{R_3} + k} V \quad (\text{B.10})$$

La resistenza di Thévenin R_{th} è la resistenza R_x già calcolata nel paragrafo precedente.

B.6 Teorema di Norton

Il teorema di Norton afferma che un circuito lineare, visto da due nodi, è equivalente a un generatore di corrente in parallelo con un resistore (circuito equivalente di Norton).

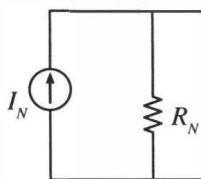


Figura B.7 Circuito equivalente di Norton

Per calcolare l'equivalente di Norton di un generico circuito si procede come segue:

- 1) si cortocircuitano i due nodi e si calcola la corrente tra i due nodi, tale corrente è I_N ;
- 2) si calcola la resistenza vista ai due nodi e si ottiene R_N .

Il circuito equivalente di Thévenin e quello di Norton sono legati dalle seguenti relazioni:

$$R_{th} = R_N \quad (\text{B.11})$$

$$I_N = \frac{V_{th}}{R_{th}} \quad (\text{B.12})$$

Analisi di circuiti lineari mediante numeri complessi

C.1 Rappresentazione di segnali mediante numeri complessi

In generale, un segnale sinusoidale a una data frequenza è caratterizzato da due parametri: l'ampiezza A e la fase φ . Infatti

$$u(t) = A \cos(\omega t - \varphi) \quad (\text{C.1})$$

con

$$\omega = \frac{2\pi}{T} \quad (\text{C.2})$$

$$\varphi = 2\pi \frac{\Delta t}{T} \quad (\text{C.3})$$

T è il periodo e Δt è l'anticipo temporale del segnale sinusoidale rispetto a un segnale di riferimento alla stessa frequenza ($f = 1/T$).

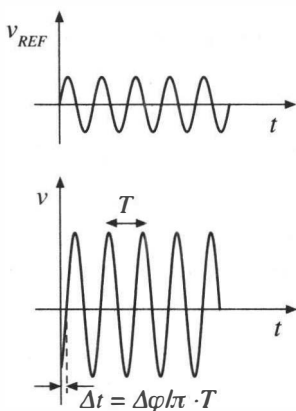


Figura C.1 Ampiezza e sfasamento

Questo fatto suggerisce la possibilità di utilizzare un singolo numero complesso per rappresentare un segnale sinusoidale a una data frequenza. A tal fine ricorriamo la seguente fondamentale relazione matematica (identità di Eulero):

$$e^{j\vartheta} = \cos\vartheta + j\sin\vartheta \quad (\text{C.4})$$

che può essere visualizzata nel piano complesso come mostrato in Figura C.2:

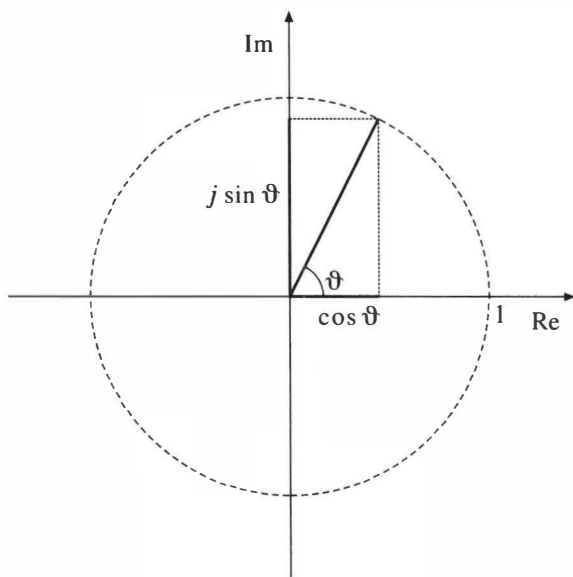


Figura C.2 Rappresentazione nel piano complesso

Dalla (C.4) segue che il segnale sinusoidale può essere visto come la parte reale di un numero complesso:

$$A\cos(\omega t + \varphi) = \text{Re}(Ae^{j(\omega t + \varphi)}) \quad (\text{C.5})$$

Il grande vantaggio di questa notazione è che utilizzando i numeri complessi si può estendere il concetto di linearità alla fase oltre che all'ampiezza di un segnale. Infatti, poiché

$$Ae^{j(\omega t + \varphi)} = Ae^{j\varphi} e^{j\omega t} \quad (\text{C.6})$$

possiamo dire che i segnali $Ae^{j(\omega t + \varphi)}$ e $e^{j\omega t}$ sono linearmente legati dal numero complesso $Ae^{j\varphi}$.

C.2 Resistenze, capacità e induttanze

Nel Capitolo 1 abbiamo discusso la relazione lineare tra tensione e corrente dovuta alla legge di Ohm.

$$v = Ri \quad (\text{C.7})$$

La legge di Ohm vale ovviamente solo per resistenze.

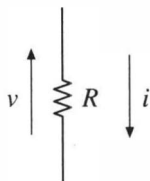


Figura C.3 Resistenza

Vedremo però che usando la notazione complessa anche per capacità e induttanze possiamo scrivere una relazione lineare tra tensione e corrente.

Infatti, la relazione tra tensione e corrente ai capi di un capacitore è

$$i = C \frac{dv}{dt} \quad (\text{C.8})$$

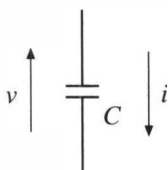


Figura C.4 Capacità

Consideriamo ora un segnale in tensione sinusoidale in notazione complessa:

$$v = Ve^{j\omega t + \varphi} \quad (\text{C.9})$$

Risulta immediato che

$$i = C \frac{dv}{dt} = j\omega Cv \quad (\text{C.10})$$

da cui

$$v = Z_C i \quad (\text{C.11})$$

con Z_C (impedenza del capacitore) data da

$$Z_C = \frac{1}{j\omega C} \quad (\text{C.12})$$

Una capacità quindi è caratterizzata da una relazione lineare tra tensione e corrente in cui il numero reale R è sostituito dal numero complesso Z_C . Si noti però che mentre R non dipende dalla frequenza, Z_C a frequenze molto basse aumenta (in modulo) fino a trasformarsi nell'impedenza di un circuito aperto ($Z = \infty$). A frequenze molto alte, al contrario, l'impedenza di una capacità diventa equivalente a quella di un cortocircuito ($Z = 0$).

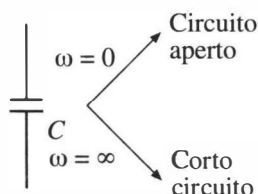


Figura C.5 Capacità a bassa e ad alta frequenza

Per un'induttanza vale

$$v = -L \frac{di}{dt} \quad (\text{C.13})$$

da cui

$$v = Z_L i \quad (\text{C.14})$$

dove

$$Z_L = j\omega L \quad (\text{C.15})$$

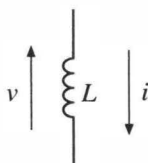


Figura C.6 Induttanza

Anche un'induttanza quindi è caratterizzata da una relazione lineare tra tensione e corrente. Si noti che Z_L a frequenze molto basse diminuisce (in modulo) fino a trasformarsi nell'impedenza di un cortocircuito ($Z = 0$). A frequenze molto alte, al contrario, l'impedenza di una induttanza diventa equivalente a quella di un circuito aperto ($Z = \infty$).

C.3 Risposta in frequenza di un filtro passa basso RC

Si consideri la rete R-C in Figura C.7. Si tratta di un filtro passa basso in quanto tende ad attenuare l'ampiezza delle componenti ad alta frequenza del segnale. Calcoliamo la tensione v_o .

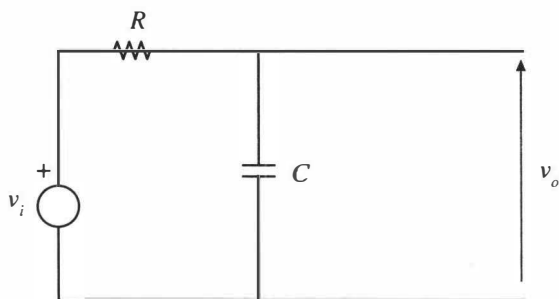


Figura C.7 Passa basso

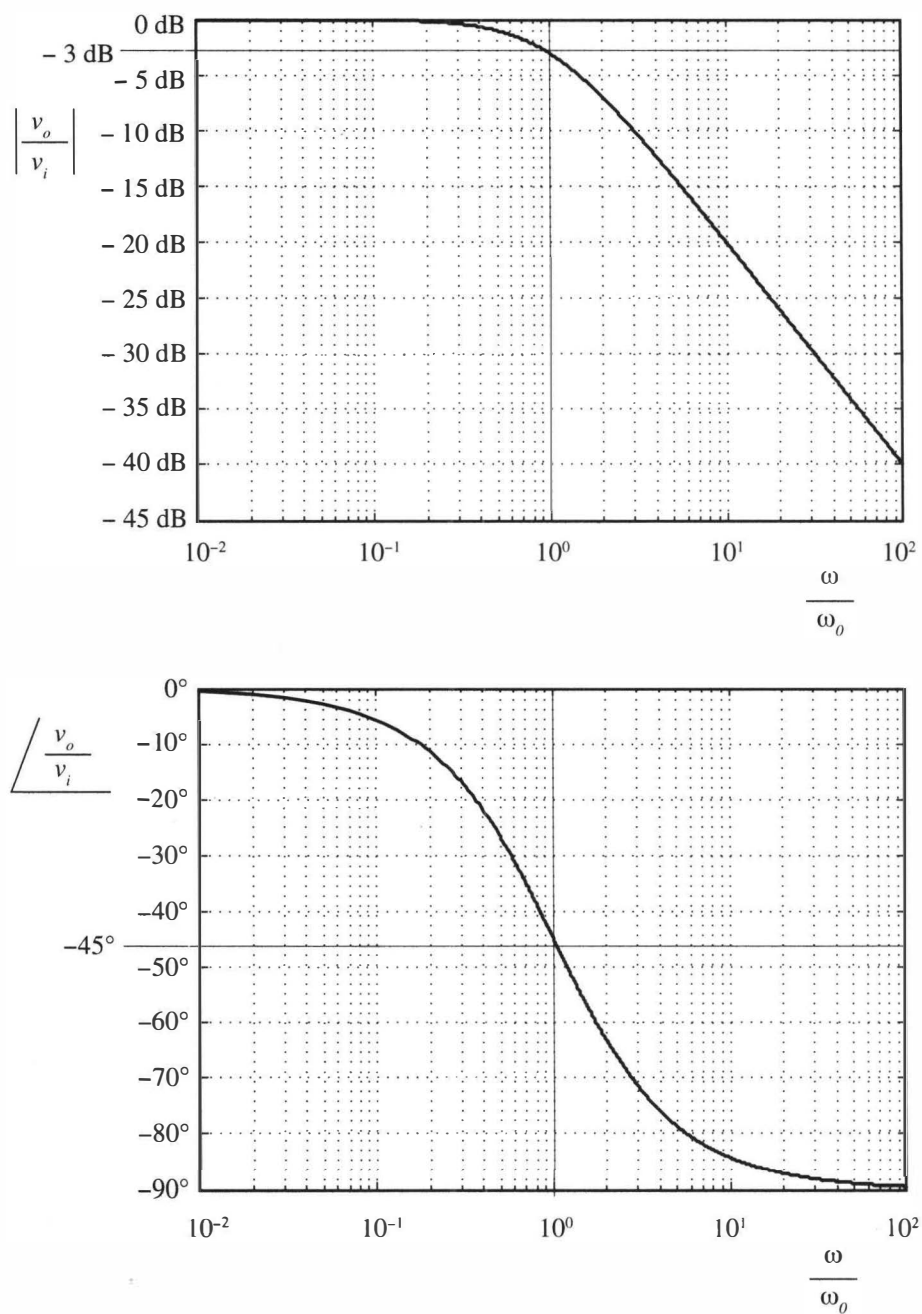
Dall'analisi del circuito risulta che

$$\frac{v_o}{v_i} = \frac{\frac{1}{j\omega C}}{\frac{1}{j\omega C} + R} = \frac{1}{1 + j\left(\frac{\omega}{\omega_0}\right)} \quad (8.16)$$

dove

$$\omega_0 = \frac{1}{CR} \quad (8.17)$$

La Figura C.4 mostra i grafici dell'ampiezza e della fase di v_o/v_i , da cui risulta che alle frequenze più alte il segnale è attenuato e progressivamente sfasato fino a 90° . Per $\omega \ll \omega_0$ il segnale attraversa inalterato il circuito.

**Figura C.8** Risposta in frequenza di un RC passa basso

C.4 Risposta in frequenza di un filtro passa alto RC

Si consideri la rete R-C in Figura C.9. Si tratta di un filtro passa alto in quanto tende ad attenuare l'ampiezza delle componenti a bassa frequenza del segnale.

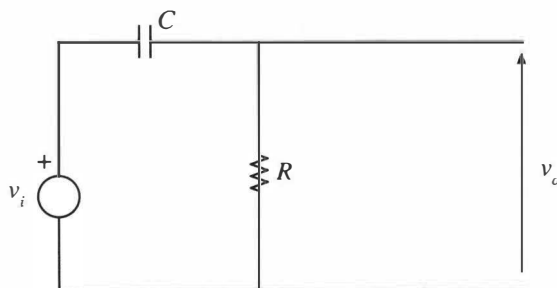


Figura C.9 Passa alto RC

Calcoliamo la tensione v_o . Dall'analisi del circuito risulta che

$$\frac{v_o}{v_i} = \frac{R}{\frac{1}{j\omega C} + R} = \frac{j\left(\frac{\omega}{\omega_0}\right)}{1 + j\left(\frac{\omega}{\omega_0}\right)} \quad (\text{C.18})$$

dove

$$\omega_0 = \frac{1}{CR} \quad (\text{C.19})$$

La Figura C.10 mostra i grafici dell'ampiezza e della fase di v_o/v_i , da cui risulta che alle frequenze più basse il segnale è attenuato e progressivamente sfasato fino a 90° . Per $\omega \gg \omega_0$ il segnale attraversa inalterato il circuito.

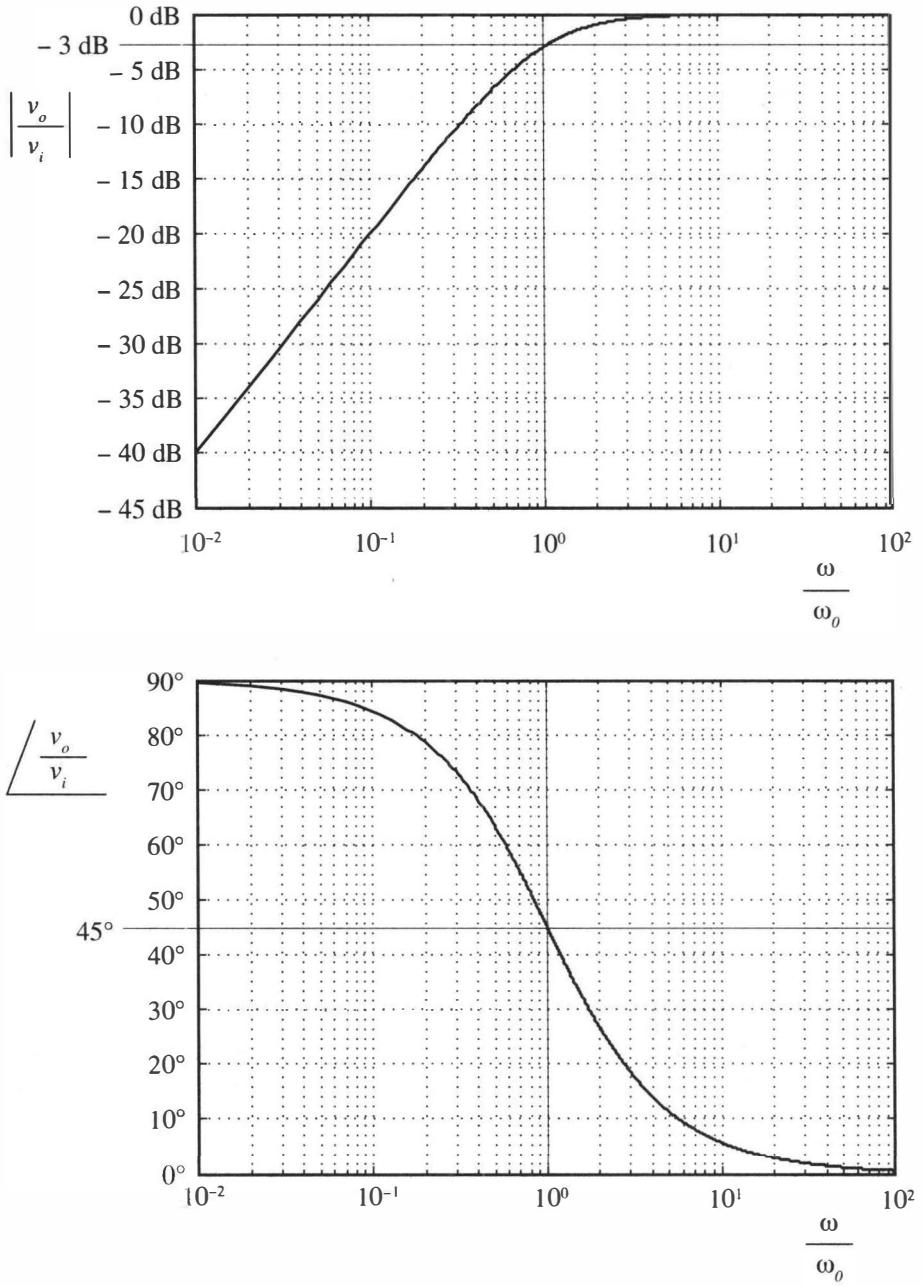


Figura C.10 Risposta in frequenza di un passa alto RC

Indice analitico

A

accettori 32
alimentazione duale 179
allargamento spettrale 124
alte frequenze 448
AMD 255
amplificatore di tensione 167, 246
amplificatore di transconduttanza 376
amplificatori 351
amplificatori alle alte frequenze 454
amplificatori lineari 353
amplificatori reazionati 397
analisi AC (alternate current) 356, 388
analisi DC (direct current) 356, 388
AND 292
anodo 65, 66
armoniche 358
arricchimento 140
arseniuro di gallio 122-123
Atalla, M.J.M. 198
audion 197, 350

B

banda di conduzione 44, 114
banda di valenza 44, 114
banda operativa 114, 448
banda proibita 44, 114
banda ultravioletto 117
Bardeen, J. 198
barriera pn 65
base 228
basse frequenze 448
Bell, laboratori 65, 125, 198, 255
Bi-CMOS 320
body 134, 140, 186
Boole, G. 291
bootstrapping 437
boro 32

Bose, J.C. 65
Brattain, W. 198
Braun, C.F. 64
breakdown 91-92
buffer di corrente 375
buffer di tensione 374, 394
bus 289

C

cache 323
campo elettrico 9
canale sepolto 178
capacità 479, 489
capacità differenziale 143, 459
capacità parassite 307, 455
capacità parassite del BJT 458
capacità parassite del MOSFET 457
capacitore 479
capacitore di accoppiamento 167, 247, 371, 448
capacitore di by-pass 371, 448
caratteristica del diodo ideale 67, 86, 87
caratteristica del JFET 194
caratteristica di trasferimento 150
caratteristica di uscita del BJT 239, 242
caratteristiche di uscita 149, 165, 176, 181, 244
carburo di silicio 123
carica di un condensatore 480
carica elettrica 6
carico attivo 175, 181, 406
cascode 467
catodo 65, 66
cattura 89
catwhisker 64
CB 387
CC 387
CD (drain comune) 371, 372, 373, 394
CE 387, 399, 400, 451, 464

cella fotovoltaica 118
 CG (gate comune) 371, 375
 chip 155
 circuiti integrati 161
 circuiti lineari 483
 circuito equivalente di Norton 486
 circuito equivalente di Thévenin 485
 clock 289
 CMOS 304, 308
 coefficiente di diffusione 36, 461
 collettore 228
 commutazioni 294
 composti binari 20
 composti quaternari 20
 composto ternario 20
 conducibilità 15
 conduttori 17
 consumo dinamico 297, 303, 315
 consumo statico 297, 303
 contatori 322
 contatto ohmico 79
 contatto rettificante 64
 controfase 167, 247, 376
 contropolarizzazione 67
 coppie elettrone-lacuna 114-115
 corrente 10
 corrente di collettore 234
 corrente di deriva 36, 115
 corrente di diffusione 35, 90, 236
 corrente di emettitore 235
 corrente di ricombinazione 90
 corrente di saturazione inversa 67, 88
 corrente di trasporto 234, 236
 cortecce 19
 costante di Boltzmann 28, 37, 124
 costante di Planck 28
 CPU 290
 crescita epitassiale 47, 229
 CS 371, 373, 388, 392, 398, 405, 455
 curva di carico 176, 181

D

Darlington 435
 De Broglie, L.-V. 28
 decoder 327

De Forest, L. 197, 350
 demodulatore AM a diodo 103
 densità di corrente 15
 nitruro di silicio 155
 polisilicio 157
 diagrammi di Bode 447
 diffusione di impurità 229
 dimensione minima 161
 dinamica in uscita 391, 395
 diodo 66
 diodo Zener 91
 direct gap 124
 distorsione di ampiezza 448
 distorsione di fase 448
 distribuzione di Boltzmann 46
 distribuzione di Fermi 43, 46
 donatori 31
 doppio dipolo 354
 drain 140
 DRAM 322, 329
 droganti 31
 dualità onda-corpuscolo 26

E

eccesso di elettroni 84
 effetto Auger 122
 effetto body 188, 365
 effetto Early 240, 254, 380
 effetto fotoelettrico 114
 effetto Miller 465
 effetto Ohm 145
 effetto tunnel 24, 334
 effetto valanga 92
 effetto Zener 92
 efficienza di conversione 118
 efficienza intrinseca 124
 efficienza quantica 117
 Einstein, A. 114, 125
 elettroni caldi 334
 elettroni di conduzione 23, 44, 89
 emettitore 228
 emissione 89, 123
 energia di Fermi 42, 74-77, 135-136
 energia di valenza 44
 energy gap 114, 124

EPROM 323
 equazione di continuità 83
 equazione di Einstein 37
 equazione di Poisson 71
 equazione di Schrödinger 26
 equivalente di Norton 353, 355
 equivalente di Thévenin 353, 355

F

FET 133
 Fairchild 198, 255
 famiglia logica 290
 fan-out 301
 Faraday, M. 48
 FIFO 289, 323
 FLASH 323, 331
 Fleming, J. A. 197
 floating gate 331
 fosforo 31
 fosfuro arseniuro di gallio 123
 fosfuro di gallio 123
 fotocamera 117
 fotodiodo 115, 125
 fotolitografia 155
 fotone 115
 fotoresist 155, 229
 frequenza di taglio 448
 frequenza di taglio inferiore 449, 454
 frequenza di taglio superiore 449, 467
 funzione di trasferimento 165, 177, 181, 244, 307
 funzione logica 290

G

galena 64
 gate 134
 gate isolato 331
 GE 255
 generatore di corrente 172, 483
 generatore di tensione 483
 generatore di Widlar 222
 generazione 34, 83
 germanio 20
 giunzione brusca 69

giunzione metallo-semiconduttore n 75
 giunzione metallo-semiconduttore p 76-77
 giunzione pn 66
 giunzione rettificante 75
 guadagno 372-377, 398-399
 guadagno di corrente 233, 238
 guadagno in tensione 400, 407

H

Heisenberg, W. 48
 Hertz, H.R. 125
 Holonyak, N. 125

I

identità di Eulero 487
 I/O 289, 290
 Impiantazione ionica 157
 impurità di tipo n 30
 impurità di tipo p 32
 indirect gap 121
 induttanza 490
 iniezione di cariche 82
 Integrated Circuit 155
 Intel 164, 255
 interdizione 149, 234, 238
 intermodulazione 358
 interruttori complementari 304
 interruttori complementari ideali 305
 ione negativo fisso 32
 ioni positivi 31
 isolanti 17

J

JFET 133, 190

K

Kelly, M. 197
 Khang, D. 198
 Kilby, J. 255

L

lacune 24
 LASER 21
 latch 324
 LED 21, 122, 125

legame metallico 18
 legami covalenti 18, 23
 legge del diodo 65
 legge di azione di massa 33
 legge di Gauss 40, 138
 legge di Moore 164
 legge di Ohm 17
 legge di Poisson 39
 LIFO 289
 Lilienfeld, J.E. 197
 luce visibile 116
 lunghezza d'onda 27, 116, 125
 lunghezza del canale 154
 lunghezza di diffusione 84, 238

M

macchina di Turing 290
 maglia di ingresso 132
 maglia di uscita 132
 margine di rumore 294, 301, 307
 margine di rumore alto 295
 margine di rumore basso 295
 mascheratura 155, 228
 medie frequenze 448
 memorie 288, 322
 metallizzazione 160, 231
 mobilità 14, 36
 mobilità degli elettroni 29, 243
 mobilità delle lacune 29, 243
 modello a bande 73
 modello del BJT ad alta frequenza 462
 modello linearizzato del BJT 377
 modello linearizzato del FET 359
 modulazione del canale 154
 MOS 134
 MOSFET 133, 140
 MOSFET a canale p ad arricchimento 183
 MOSFET a quattro terminali 186
 MOSFET a tre terminali 184

N

NAND 299, 306
 nanotubo 163

nitruro di gallio e indio 123
 NMOS 182
 non linearità 356
 NOR 298, 306
 NOT 291, 305, 307
 npn 228

O

Ohl, R. 65, 125
 Ohm 11
 OR 292
 orbitali 18
 ossidazione 157, 228
 ossido 134
 ossido di silicio 134, 228
 overlapping 457

P

parametro di forma 178
 passa alto RC 450, 493
 basso RC 452, 491
 Pickard, G.W. 65
 pinch-off 196
 pinna 162
 pixel 117
 Planck, M. 114
 pnp 228, 242
 polarizzazione 68
 polarizzazione con generatore di corrente 172
 polarizzazione del BJT 249
 polarizzazione del BJT mediante generatore di corrente 253
 polarizzazione del MOSFET 169, 178
 polarizzazione diretta 67
 polisilicio 140, 157
 portatori maggioritari 236
 portatori minoritari 236
 potenziale di built-in 37, 41, 81
 potenziale di giunzione 73, 194
 potenziale elettrico 8
 potenziale termico 37
 prima legge di Kirchhoff 484
 principio di esclusione di Pauli 42

principio di sovrapposizione 356
 punto di lavoro 68, 171, 181
 punto di polarizzazione 68, 388
 punto di riposo 388

Q

quanto 114
 quarzite 47

R

RAM 290, 323, 331
 raddrizzatore di tensione 103
 RAM dinamica 329
 RAM statica 323
 RCA 198, 255
 reazione negativa 171, 252, 397, 400
 refreshing 331
 regione attiva diretta 234, 244, 246
 regione attiva inversa 234, 238
 regione di svuotamento 69, 115, 134, 194, 459
 regione lineare 149
 registri 322
 regole di progetto 388
 relazione di Einstein 37, 38, 243
 reoforo 66, 134
 resistenza 11, 145, 489
 resistenza di base 381
 resistenza di carico 68, 165
 resistenza di collettore 381
 resistenza di drain 361
 resistenza di ingresso 354
 resistenza di spreading 462
 resistenza di uscita 354
 resistore 485
 responsività 116, 124
 rete a quattro resistori 171, 250, 366
 reti lineari 446
 reticolo cristallino 18
 retta di carico 68, 165, 244
 ricombinazione 33, 82, 89
 ricombinazione non radiativa 122
 rigenerazione dei fronti 294
 risposta in frequenza 446
 ROM 323

S

saturazione 149, 234, 244
 scarica di un condensatore 482
 seconda legge di Kirchhoff 484
 Shockley, W. 65, 197, 255
 silicio 20
 silicio intrinseco 29
 silicio metallurgico 47
 silicio monocristallino 47, 228
 silicio policristallino 47
 Silicon Valley 255
 size 164
 source 140
 specchio di corrente 173
 specchio di corrente a BJT 253
 specchio di Widlar 222, 280
 specchio di Wilson 221, 279
 SPICE 388, 454
 SRAM 322, 328
 stati trappola 89
 statistica di Boltzmann 14
 statistica di Fermi-Dirac 42
 strato sepolto 229, 233
 strozzamento 148, 150, 154, 191
 svuotamento 178

T

tecnologia planare 161
 tempi di salita e di discesa 296, 302, 315
 tempo di esposizione 118
 tempo di propagazione 296, 301, 307
 tempo di transito 461
 tensione 8
 tensione di Early 242, 383
 tensione di pinch-off 191
 tensione di soglia 67, 88, 90, 135-136, 178, 333
 teorema di Miller 403, 405, 464, 469
 teorema di Norton 486
 teorema di Thévenin 485
 Texas Instruments 255
 Totem-pole 320
 transconduttanza 360
 transconduttanza del BJT 379
 transistor 132

transistor bipolare 198, 228
trappole in banda proibita 88
trappole profonde 89
trasferimento di corrente 355-356
trasferimento di tensione 355-356
triclorosilicato 47
triode 197
TTL 294
tubi a vuoto 65

U

unità logiche 288
USB 323

UV 117, 156, 228

V

valvola 132
velocità di deriva 14, 29, 152, 194
vias 160, 232
Volt 8
Volta, A. 48
Von Neumann, J. 290

W

wafer 47, 155, 161, 228
Watson, A. 48

Fondamenti di elettronica

con eText

Indice

- Elenco dei capitoli sfogliabili

Note

- Raccolte di appunti inseriti dallo studente nel testo

Segnalibri

- Rimandi alle pagine, inseriti dallo studente



Per accedere alla piattaforma collegati al sito

www.pearson.it/place con:

- il tuo indirizzo e-mail
- il codice di accesso che trovi in copertina

Pearson Learning Solution

Pearson Learning Solution nasce dall'esperienza del principale editore educational nel mondo.

Ogni progetto didattico Pearson è composto da un manuale di riferimento e da una serie di supporti digitali - piattaforme di e-learning, eText e siti dedicati - che accrescono la qualità dello studio, riducono i tempi di apprendimento, aiutano a misurare il livello di preparazione raggiunto tramite test ed esercizi di autovalutazione.

L'esperienza nazionale e internazionale mostra che i prodotti didattici Pearson migliorano in maniera significativa i risultati raggiunti.

€ 35,00



9788865188088