Lezioni di
Laboratorio di Segnali e Sistemi

A.NIGRO

Dipartimento di Fisica, Università La Sapienza di Roma

(Revisione dicembre 2007)

Introduzione

Queste dispense costituiscono una rielaborazione delle lezioni da me tenute per il corso di Esperimentazione di Fisica III (Laboratorio di Fisica I, nel vecchio ordinamento) negli ultimi anni. Nella nostra Università questo corso ha sempre avuto la funzione di insegnare agli studenti l'Elettronica; di fatto è l'unico corso che essi ricevono in questo campo (naturalmente fatta eccezione per coloro che seguono l'indirizzo Elettronico). L'obiettivo è di dare agli studenti una certa dimestichezza con questa materia, che li metta in grado, qualunque sia il loro mestiere futuro, di affrontare e risolvere semplici problemi applicativi, ovvero li metta in grado di interagire efficacemente con degli specialisti, per i problemi piu' complessi. Oggi la cultura elettronica è fondamentale in qualunque campo della Scienza e della Tecnica ed è quindi fondamentale, per un laureato in Fisica, capire, o almeno intravedere, cosa si puo', e cosa non si puo' fare, con l'Elettronica.

L'unico modo per conoscere seriamente questa materia è attraverso la pratica, quindi in questo corso hanno un ruolo centrale, come in passato, le esercitazioni che gli studenti svolgono in laboratorio parallelamente al corso in aula, imparando a progettare e a costruire circuiti via via più complicati[1].

Nel corso degli ultimi anni l'Elettronica ha fatto progressi giganteschi che stanno rivoluzionando il mondo in cui viviamo. Alcuni potrebbero quindi trovare strano che ancora si parli di circuiti RC e si perda tempo con antiquati transistors, mentre si potrebbe meglio impiegare il tempo con circuiti ultra-moderni capaci di prestazioni esaltanti. Io, come molti altri, sono in totale disaccordo con questa filosofia: è la filosofia delle scatole nere, che gli studenti imparano ad usare (leggendo i fogli illustrativi), senza nessuna possibilita' di capire cosa succede. Invece, nella nostra filosofia, lo studente deve capire cosa succede, e padroneggiare completamente i circuiti che costruisce ed usa. In altre parole, se si conoscono i fondamentali si è in grado, studiando, di aggiornarsi e di arrivare, in modo consapevole, ad usare scatole che ora non saranno piu' completamente nere, ma, almeno, semitrasparenti.

[1] La Guida alle Esercitazioni costituisce oggetto di un fascicolo separato, reperibile sul sito http://www-zeus.roma1.infn.it/nigro/.

Indice

Capitolo 1

Richiami di teoria delle reti lineari

1.1 Introduzione

In questo corso presupporremo noti alcuni concetti base già studiati nei corsi precedenti. In particolare daremo per acquisite le nozioni di differenza di potenziale elettrica, di corrente elettrica, di resistenza, di capacità e di induttanza. Riteniamo inoltre che gli studenti abbiano già avuto modo di studiare il comportamento di semplici circuiti elettrici, in cui compaiono generatori di tensione costante, ovvero generatori di tensione sinusoidale. Nella

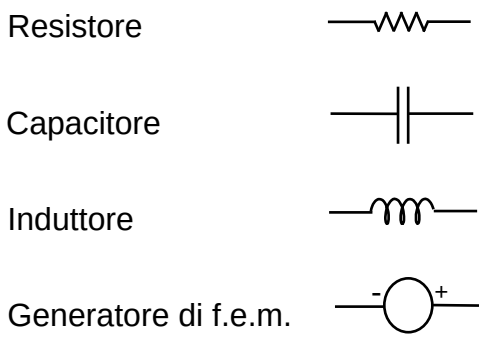

Figura 1.1: *Principali simboli utilizzati*

Fig. 1.1 sono rappresentati i simboli per gli elementi di circuito che utilizzeremo. Inoltre, indicheremo con lettere maiuscole grandezze fisiche indipendenti dal tempo (p.es, V, per simboleggiare una tensione costante), e con lettere minuscole grandezze fisiche dipendenti dal tempo (p.es. $i(t)$, per indicare una corrente variabile nel tempo).

1.2 Teoria delle reti

Una rete lineare é un circuito che da luogo a equazioni o sistemi lineari. Ad esempio, il circuito in Fig. 1.2a da luogo all'equazione

$$V = (R_1 + R_2)I$$

Supponendo noti i valori degli elementi che compongono il circuito, cioe' generatore e resistenze, puo' essere ricavato molto facilmente il valore della corre nte I, che circola nelle resistenze.

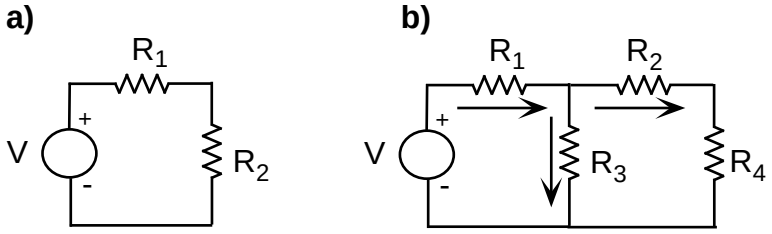

Figura 1.2: *a) Un semplice circuito elettrico; b) Un circuito più complesso*

In Fig. 1.2b abbiamo un altro esempio, questa volta più complicato, di circuito. Ora le incognite (le correnti che attraversano i vari elementi di circuito) sono più di una. Per poter risolvere il circuito dobbiamo quindi poter scrivere più equazioni (in numero almeno pari a quello delle incognite).

Prima di affrontare questo problema conviene approfondire le nostre nozioni sugli - elementi che costituiscono un circuito. I piú semplici elementi sono quelli bipolari (vedi Fig. 1.1), nei quali esiste una relazione funzionale tra la corrente $i(t)$ che circola nelÍ elemento e la differenza di potenziale $v(t)$ tra i suoi estremi. Si ha quindi:

$$v(t) = f[i(t)]$$

$$i(t) = g[v(t)]$$

Dove, evidentemente

$$g = f^{-1}$$

Le suddette relazioni funzionali tra tensione e corrente non sono necessariamente di tipo algebrico, ma possono essere più in generale di tipo analitico, cioe' espresse attraverso operazioni di derivazione o integrazione. Dalla forma della funzione f, gli elementi si possono distinguere tra lineari e non lineari, dove la linearità va intesa in senso esteso, cioe' anche analitica.

Un' altra importante distinzione va operata tra gli elementi attivi e passivi, cioé tra quelli che rispettivamente contengono e non contengono sorgenti interne di energia.

Gli elementi passivi che conosciamo sono:

- Resistore
- Induttore
- Condensatore (o capacitore)

Possiamo in modo del tutto generale scrivere:

$$v(t) = \{z\}i(t)$$

$$i(t) = \{y\}v(t)$$

dove l'operatore $\{z\}$ si chiama impedenza, mentre l'operatore inverso $\{y\}$ si chiama ammettenza. Si ha, per i tre elementi suddetti:

Resistore	$\{z\} \equiv R$	$\{y\} \equiv G$
Induttore	$\{z\} \equiv L\frac{d}{dt}$	$\{y\} \equiv \frac{1}{L}\int dt$
Condensatore	$\{z\} \equiv \frac{1}{C}\int dt$	$\{y\} \equiv C\frac{d}{dt}$

Come si vede questi tre elementi passivi sono di tipo lineare. Vedremo in seguito esempi di elementi non lineari.

Se diversi elementi sono connessi in serie l'impedenza complessiva del sistema e' data dalla somma delle impedenze dei singoli elementi:

$$\{z\} = \sum_i \{z_i\}$$

Se invece sono connessi in parallelo l'ammettenza complessiva e' data dalla somma delle ammettenze dei singoli elementi:

$$\{y\} = \sum_i \{y_i\}$$

Gli elementi attivi sono invece:

- Generatore di tensione ideale
- Generatore di corrente ideale

Nella Fisica Generale e' stato introdotto il concetto di generatore di forza elettromotrice: un dispositivo in grado di mantenere una differenza di potenziale tra i suoi morsetti e di erogare una corrente elettrica. In Elettronica e' conveniente schematizzare un generatore in due diversi modi, cioe' attraverso la definizione di due possibili modelli ideali di esso.

Il generatore di tensione ideale e' un dispositivo in grado di mantenere una differenza di potenziale tra i suoi morsetti, indipendentemente dal carico ad esso connesso, cioe' indipendentemente dalla corrente erogata. Invece il generatore di corrente ideale e' un dispositivo in grado di erogare una ben definita corrente, indipendentemente dal carico ad esso connesso.

I generatori reali hanno ovviamente un comportamento diverso da entrambi questi modelli; esso verrà chiarito meglio nel par. 1.5.

E' bene infine ricordare che se in un circuito ci sono piú induttori si ha tra essi una mutua induzione: salvo casi particolari, noi trascureremo questo effetto.

1.3 Elementi di una rete

Una rete puó essere pensata come composta da tanti elementi dipolari opportunamente connessi. Gli elementi topologici di una rete sono i nodi, i rami e le maglie (Fig. 1.3).

Chiamiamo *nodo* ogni punto della rete in cui confluiscono più di due elementi. I rami sono costituiti dagli elementi che uniscono tra loro due nodi. Chiamiamo invece *maglia* qualunque percorso chiuso possiamo effettuare partendo da un nodo per tornare al nodo medesimo. Data una rete con n nodi, ci sará un certo numero m di maglie, non tutte indipendenti tra loro. I rami possono essere distinti in:

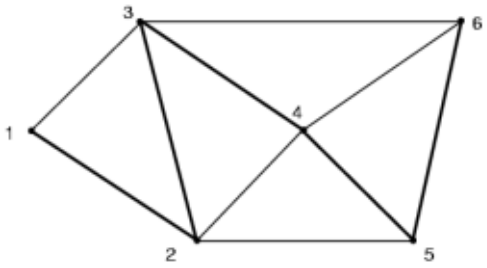

Figura 1.3: *Esempio di rete. in grassetto è mostrata una possibile scelta dei rami scheletro*

- rami *scheletro*, essenziali per connettere i nodi;
- rami *anello*, non essenziali per connettere i nodi, ma necessari per formare le maglie.

Vi sono ovviamente vari modi di scegliere i rami scheletro, ma il loro numero non cambia, essendo dato da:

$$r_s = n - 1$$

dove n e' il numero di nodi. Prendendo un nodo come riferimento di tensione, r_s rappresenta il numero di coppie di nodi indipendenti. Il numero di maglie, m,coincide con il numero di rami anello, r_a:

$$m = r_a = r - r_s = r - n + 1$$

dove r e' il numero totale di rami.

1.4 Analisi delle reti

Fare l'analisi di una rete consiste nel calcolare le risposte di una rete note le eccitazioni. Ad esempio, nel circuito in Fig. 1.2b abbiamo le tre incognite i_1, i_2, i_3, e quindi bisogna scrivere tre equazioni, che leghino le incognite al valore degli elementi del circuito. Per fare questo possiamo utilizzare i cosidetti *Principi di Kirchoff*:

- **Principio di Kirchoff delle maglie**

  La somma delle cadute di potenziale lungo un percorso chiuso e' uguale a zero, ovvero, se nel percorso scelto sono presenti generatori di tensione, essa e' uguale alla d.d.p. erogata dal (o dai generatori). Questo principio deriva dalla conservazione dell'energia.

- **Principio di Kirchoff dei nodi**

  La somma (algebrica) delle correnti che confluiscono in un nodo e' sempre pari a zero. Questo principio deriva dalla conservazione della carica elettrica; quindi le correnti che entrano in un nodo devono essere bilanciate dalle correnti che escono dal nodo stesso.

Utilizzando i principi di Kirchoff per il circuito in Fig. 1.2b si possono scrivere 3 equazioni delle maglie e 2 equazioni dei nodi, cioe' un sistema di 5 equazioni, in 3 incognite:

$$\begin{cases} v = R_1 i_1 + R_3 i_3 \\ v = R_1 i_1 + (R_2 + R_4) i_3 \\ 0 = -R_3 i_3 + (R_2 + R_4) i_2 \\ 0 = i_1 - i_2 - i_3 \\ 0 = i_2 + i_3 - i_1 \end{cases}$$

Si vede subito che le ultime due equazioni sono identiche tra loro: possiamo quindi eliminare l'ultima e rimanere con un sistema di 4 equazioni, *non linearmente indipendenti*. Scegliendo 3 di queste 4 equazioni e risolvendole, troveremo le nostre incognite. Naturalmente dovremo fare attenzione a scegliere un sotto-insieme di equazioni *linearmente indipendenti*.
Questa situazione e' del tutto generale: i principi di Kirchoff danno luogo ad un numero di equazioni *sempre* superiore al numero di incognite del problema. Dovremo quindi provvedere a ridurre il sistema ad un numero di relazioni *indipendenti* pari al numero di incognite. Anziche' fare questa riduzione in modo empirico, esistono due metodi classici per avere *immediatamente* il numero giusto e minimo indispensabile di relazioni. Essi vanno sotto i nomi di *metodo delle tensioni ai nodi* e *metodo delle correnti di maglia*.

- **Metodo delle tensioni ai nodi:**

Consideriamo il circuito in Fig. 1.4, in cui abbiamo le resistenze R_A,R_B,...R_E. Vi sono 3 nodi: prendendo il nodo 3 come nodo di riferimento possiamo scrivere un sistema di 2 equazioni nelle due incognite v_1 e v_2, cioe' le differenze di potenziale dei 2 nodi rispetto al nodo di riferimento:

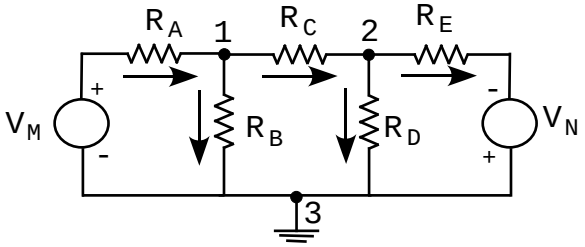

Figura 1.4: *Circuito con 3 nodi. Le frecce indicano i versi delle correnti, scelti arbitrariamente.*

$$\begin{cases} -\frac{v_1 - v_M}{R_A} + \frac{v_1}{R_B} + \frac{v_1 - v_2}{R_C} = 0 \\ \frac{v_2 - v_1}{R_C} + \frac{v_2}{R_D} + \frac{v_2 + v_N}{R_E} = 0 \end{cases}$$

Riordinando i termini si ottiene

$$\begin{cases} \frac{1}{R_A} v_M = (\frac{1}{R_A} + \frac{1}{R_B} + \frac{1}{R_C}) v_1 - \frac{1}{R_C} v_2 \\ -\frac{1}{R_E} v_N = -\frac{1}{R_C} v_1 + (\frac{1}{R_C} + \frac{1}{R_D} + \frac{1}{R_E}) v_2 \end{cases}$$

e, in termini di conduttanze, si ha

$$\begin{cases} G_A V_M &= (G_A + G_B + G_C)v_1 - G_C v_2 \\ -G_E V_M &= -G_C v_1 + (G_C + G_D + G_E)v_2 \end{cases}$$

Risolvendo, si trovano facilmente i valori delle tensioni incognite v_1 e v_2, e quindi di tutte le correnti.

- **Metodo delle correnti di maglia:**

Lo stesso circuito (Fig. 1.5) puo' essere studiato utilizzando un diverso approccio. Introduciamo 3 variabili ausiliarie, le cosidette correnti di maglia i_1, i_2 e i_3, scegliendone arbitrariamente il verso. Esse sono definite dalle seguenti eguaglianze:

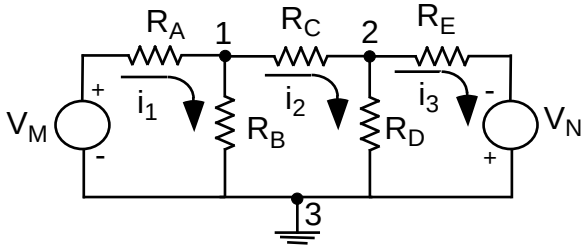

Figura 1.5: *Lo stesso circuito, studiato introducendo le correnti di maglia.*

$$\begin{cases} i_A = i_1 \\ i_B = i_1 - i_2 \\ i_C = i_2 \\ i_D = i_2 - i_3 \\ i_E = i_3 \end{cases}$$

In altre parole, se un elemento appartiene ad una sola maglia, la corrente che vi circola coincide con la corrente di maglia. Se invece un elemento e' comune a due maglie, la corrente che vi circola e' data dalla somma (o differenza) delle due correnti di maglia pertinenti. Possiamo ora scrivere il sistema di 3 equazioni in 3 incognite:

$$\begin{cases} V_M &= i_1 R_a + (i_1 - i_2)R_b \\ 0 &= i_2 R_c + (i_2 - i_3)R_D + (i_2 - i_1)R_B \\ V_N &= i_3 R_E + (i_3 + i_2)R_D \end{cases}$$

Una volta risolto il sistema, le correnti che effettivamente circolano nelle varie resistenze possono facilmente essere ricavate.

Come si vede il metodo delle tensioni ai nodi era più conveniente, poiche' dava luogo ad un minor numero di equazioni, in quanto il numero di coppie di nodi indipendenti

era minore del numero delle maglie. Cio' e' generalmente vero: il numero di coppie di nodi indipendenti e' in genere minore (o al più uguale) del numero di maglie indipendenti. Quindi il metodo dei nodi e' spesso più conveniente di quello delle maglie.

Vi sono pero' dei casi in cui il metodo dei nodi e' inapplicabile. Consideriamo ad esempio il circuito in Fig. 1.6. Per i nodi 1 e 2 si puo' scrivere:

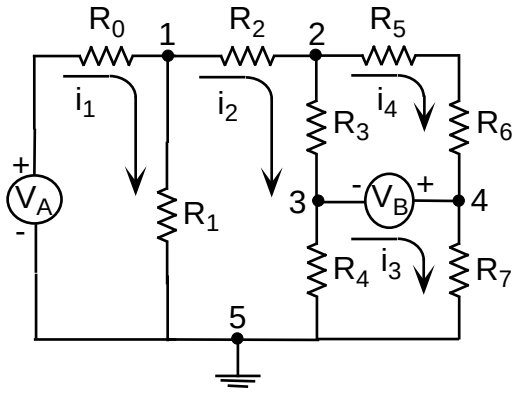

Figura 1.6: *Un circuito particolare*

$$\begin{cases} \frac{V_A - v_1}{R_0} - \frac{V_1}{R_1} - \frac{v_1 - v_2}{R_2} & = & 0 \\ \frac{v_1 - v_2}{R_2} - \frac{v_2 - v_3}{R_3} - \frac{v_2 - v_4}{R_5 + R_6} & = & 0 \end{cases}$$

Ma per i nodi 3 e 4 si ha un problema: come esprimere la corrente che circola attraverso V_B? Si capisce che in questa situazione un generatore di tensione ideale crea delle divergenze nell'equazione dei nodi connessi.

Un problema analogo nascerebbe quando si volesse usare il metodo delle maglie in un circuito in cui sono presenti generatori ideali di corrente: e' infatti impossibile esprimere la caduta di potenziale ai capi di essi. Queste apparenti incongruenze derivano dall'aver introdotto elementi di circuito non realistici: i generatori reali (vedi il prossimo paragrafo) non danno luogo a queste divergenze.

Finora abbiamo visto circuiti con sole resistenze. Consideriamo ora invece il circuito in Fig. 1.7. Esso ha 3 coppie di nodi indipendenti e 3 maglie indipendenti. Scrivendo le equazioni delle maglie si ha:

$$\begin{cases} f(t) & = & \{R_1 + L_1 \frac{d}{dt} + \frac{1}{C_1} \int dt\} i_1 - \{R_1\} i_2 \\ 0 & = & -\{R_1\} i_1 + \{(R_1 + R_2) + (L_2 + L_3) \frac{d}{dt}\} i_2 - \{R_2\} i_3 \\ 0 & = & -\{R_2\} i_2 + \{(R_2 + R_3) + L_4 \frac{d}{dt} + \frac{1}{C_2} \int dt\} i_3 \end{cases}$$

Abbiamo quindi non più un sistema algebrico, bensì un sistema di equazioni integro-differenziali. Si comprende quindi come in presenza di elementi reattivi (induttanze e capacità) la soluzione di un circuito possa rapidamente divenire un problema proibitivo. E' necessario quindi utilizzare altri metodi, che verranno illustrati nei successivi paragrafi.

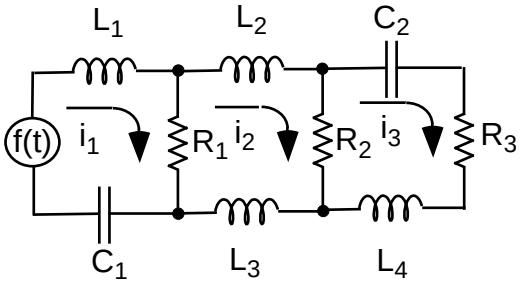

Figura 1.7: *Circuito con reattanze*

1.5 Ulteriori proprietà delle reti

Ricordiamo alcune importanti proprietà di cui godono le reti lineari; esse ci saranno utili nel seguito.

Teorema di sovrapposizione

La risposta di un circuito lineare si ottiene considerando separatamente le singole sorgenti e sommando le relative risposte. Ció segue direttamente dalla linearitá delle equazioni che governano il circuito (in ultima analisi, dalle equazioni di Maxwell).

Teorema di Thevenin

Ogni rete, vista da due terminali, puó essere sostituita da un generatore di tensione ideale in serie ad una opportuna impedenza (Fig. 1.8a). v_{eq} é ovviamente la tensione che si misura fra A e B (quando non sono connessi tra loro esternamente). Z_{eq} é l'impedenza che la rete presenta fra i terminali A e B.

Teorema di Norton:

Ogni rete, vista da due terminali, puó essere sostituita da un generatore di corrente ideale, con una opportuna impedenza in parallelo (Fig. 1.8b). i_{eq} e' uguale alla corrente che circola tra i due terminali quando essi sono esternamente cortocircuitati, con in parallelo l'impedenza che il circuito presenta fra i terminali stessi.

Poiché per ogni rete possiamo applicare il teorema di Thevenin e quello di Norton, segue che

$$i_{eq} = \frac{v_{eq}}{Z_{eq}}$$

E' quindi sempre possibile trasformare un circuito sostituendo un generatore di tensione ad uno di corrente, e viceversa. Ovviamente se la rete e' composta da sole resistenze l'impedenza equivalente sarà puramente resistiva.
Questi due teoremi ci fanno comprendere quindi che un generatore di forza elettromotrice reale, qualunque sia il suo principio fisico di funzionamento, e' sempre schematizzabile come un generatore ideale (di tensione o corrente) con una opportuna impedenza (in serie o in parallelo). In genere questa impedenza e' puramente resistiva e prende il nome di resistenza d'uscita del generatore.

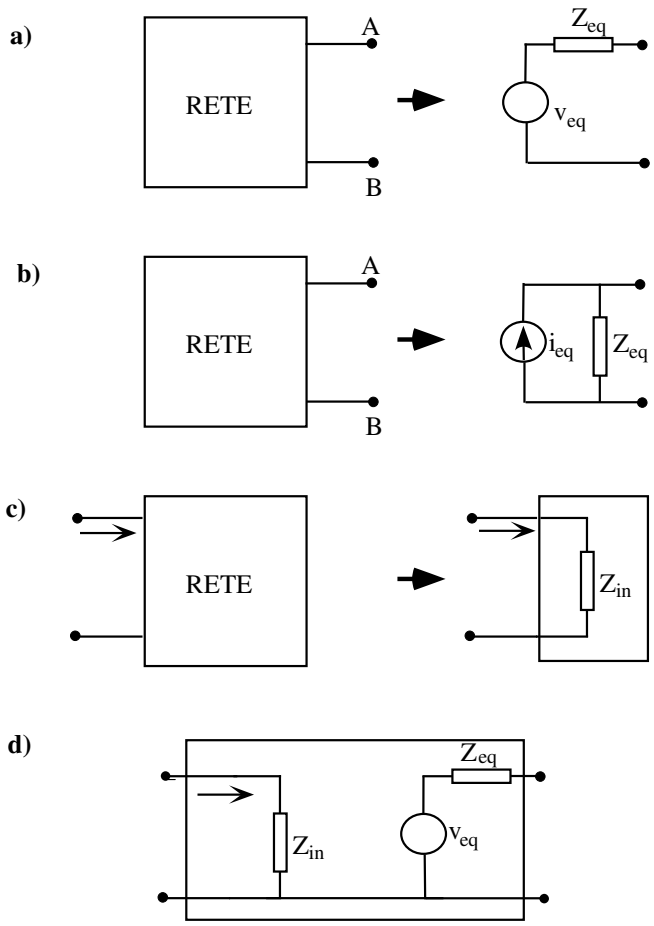

Figura 1.8: *a) Teorema di Thevenin; b) Teorema di Norton; c) Impedenza di ingresso; d) Schematizzazione completa di una rete quadrupolare, utilizzando il teorema di Thevenin.*

Quadrupoli

Molto spesso studieremo reti di tipo quadrupolare: in questo caso non siamo interessati a conoscere tutte le correnti che circolano nei vari rami (o le tensioni di tutti i nodi); ci interessa solo conoscere la *risposta* del circuito, intesa come tensione del nodo di uscita, in funzione della sollecitazione che applichiamo al nodo di ingresso.
E' chiaro quindi che i teoremi di Thevenin e Norton sono particolarmente utili in questo caso, perche' possono aiutarci a semplificare il problema, dal punto di vista dell'uscita.
In modo speculare, è possibile semplificare il problema dal punto di vista dell'ingresso, introducendo il concetto di *impedenza d'ingresso* del quadrupolo,Z_{in}, definita come il rapporto tra la tensione applicata al nodo d'ingresso e la corrente che entra nel quadrupolo stesso. (Fig. 1.8c).
In definitiva la rete quadrupolare può essere rappresentata come in (Fig. 1.8d), o in modo analogo utilizzando per l'uscita il teorema di Norton.

1.6 Trasformazioni di Fourier e Laplace

Come abbiamo visto, lo studio dei circuiti diventa rapidamente molto complicato, poiche' e' legato alla soluzione di sistemi di equazioni differenziali. In realtà e' possibile semplificare molto il problema utilizzando adeguate tecniche matematiche, che verranno brevemente, e non rigorosamente, illustrate in questo paragrafo.

Ricordiamo anzitutto che una funzione periodica $f(t)$ puó essere sviluppata in serie di Fourier, cioe'

$$f(t) = \frac{1}{2}a_0 + \sum_n \{a_n \cos n\omega t + b_n \sin n\omega t\}$$

dove

$$a_n = \frac{1}{T}\int_0^T f(t)\cos n\omega t dt \qquad b_n = \frac{1}{T}\int_0^T f(t)\sin n\omega t dt$$

$T = 2\pi/\omega$ é il periodo della funzione. Lo sviluppo di Fourier puo' anche essere espresso in forma complessa:

$$f(t) = \sum_{-\infty}^{\infty} c_n e^{jnt\omega}$$

dove

$$c_n = \frac{1}{T}\int_{-T/2}^{T/2} f(t)e^{-jn\omega t}dt$$

Si vede subito che $c_{-n} = c_n^*$ ed inoltre

$$f(t) = 2Re[\sum_0^{\infty} nc_n e^{jn\omega t}]$$

Le funzioni non periodiche non sono sviluppabili in serie di Fourier; tuttavia una funzione non periodica puo' essere considerata come una funzione periodica con periodo infinito. Quindi, passando al limite per $T \to \infty$ la sommatoria dello sviluppo di Fourier e' sostituita da un integrale e si ha:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(\omega)e^{j\omega t}d\omega$$

dove

$$g(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt$$

La funzione $g(\omega)$ e' detta la trasformata di Fourier della funzione $f(t)$.

Condizioni sufficienti affinché una funzione $f(t)$ sia trasformabile sono:

- che $f(t)$ sia continua, almeno a tratti;
- che $\int_{-\infty}^{\infty} |f(t)|dt$ sia convergente.

Vediamo ora alcuni esempi di trasformate di Fourier:
1) Funzione rettangolare:

$$f(t) = 1 \quad \text{per } |t| \leq \tau/2$$

$$f(t) = 0 \quad \text{per } |t| \geq \tau/2$$

$$g(\omega) = \int_{-\tau/2}^{\tau/2} e^{-j\omega t}dt = 2/\omega \sin \omega\tau/2$$

2) gaussiana

$$f(t) = e^{-\frac{t^2}{2\tau^2}}$$

$$g(\omega) = \sqrt{2\pi}\tau e^{-\frac{\tau^2\omega^2}{2}}$$

3) delta di Dirac

Come e' noto e' definita dalle condizioni

$$\delta(x) = 0 \quad \text{se } x \neq 0$$

$$\int_{-\infty}^{\infty} \delta(x)dx = 1$$

Inoltre, per ogni funzione f

$$\int f(y)\delta(x-y)dy = f(x)$$

Si ha quindi

$$g(\omega) = 1/2\pi \int \delta(t)e^{j\omega t}dt = 1$$

cioé

$$\delta(t) = 1/2\pi \int e^{j\omega t}d\omega$$

Un risultato analogo si ottiene se si fa tendere $\tau \to 0$ nell' impulso rettangolare o gaussiano.

Una importante proprietà delle trasformazioni di Fourier e' data dal teorema di Parseval:

$$\int |f(t)|^2 dt = 1/2\pi \int |g(\omega)|^2 d\omega$$

Infatti

$$\begin{aligned}
\int |f(t)|^2 dt &= \int dt(\frac{1}{2\pi}\int g^*(\omega)e^{-j\omega t}d\omega) \\
&= (\frac{1}{2\pi}\int g(\omega')e^{j\omega' t}d\omega' \\
&= \frac{1}{2\pi}\int d\omega g^*(\omega)\int d\omega' g(\omega')\frac{1}{2\pi}\int e^{i(\omega'-\omega)t}dt \\
&= \frac{1}{2\pi}\int d\omega g^*(\omega)\int d\omega' g(\omega')\delta(\omega'-\omega) \\
&= \frac{1}{2\pi}\int d\omega g^*(\omega)g(\omega) \\
&= \frac{1}{2\pi}\int |g(\omega)|^2 d\omega
\end{aligned}$$

Se la f(t) é una funzione pari

$$\begin{aligned}
g(\omega) &= \int_0^\infty f(t)e^{-j\omega t}dt + \int_{-\infty}^0 f(t)e^{-j\omega t}dt \\
&= \int_0^\infty f(t)(e^{j\omega t}+e^{-j\omega t})dt \\
&= 2\int_0^\infty f(t)\cos\omega t dt
\end{aligned}$$

Da notare che anche la $g(\omega)$ é pari. E quindi

$$f(t) = 1/\pi \int g(\omega)\cos\omega t d\omega$$

Se invece f(t) é una funzione dispari, si ha

$$\begin{aligned}
f(t) &= 1/\pi\int_0^\infty g(\omega)\sin\omega t d\omega \\
&= 2\int_0^\infty f(t)\sin\omega t dt
\end{aligned}$$

Ovviamente esistono funzioni non trasformabili secondo Fourier. Ad esempio, la funzione a gradino $u(t)$, definita come:

$$u(t) = 1 \quad \text{per } t > 0$$

$$u(t) = 0 \quad \text{per } t < 0$$

non e' trasformabile, perche'

$$\int_{-\infty}^{\infty} |u(t)|dt$$

non converge.

Si può però ricorrere ad una trasformazione piú generale, cioe' la trasformazione di Laplace.

Consideriamo la variabile complessa $s = \sigma + j\omega$, e sia $s_0 = \sigma_0 + j\omega_0$ un punto del piano complesso: se

$$\lim_{a\to 0}\lim_{b\to\infty}\int_a^b e^{-s_0 t}f(t)dt$$

converge diremo che f(t) é trasformabile secondo Laplace. Inoltre

$$\int_0^\infty e^{-st}f(t)dt$$

sarà convergente per ogni s tale che $\sigma > \sigma_0$. La trasformata di Laplace é

$$L[f(t)] \equiv F(s) = \int_0^\infty e^{-st} f(t) dt$$

mentre la anti-trasformata ci riporta alla funzione di partenza:

$$L^{-1}[F(s)] = f(t)$$

La trasformazione di Laplace gode delle seguenti proprietà:

a) $L[k] = \frac{k}{s}$

b) $L[\sum_k^n a_k f_k(t)] = \sum_k^n a_k F_k(s)$

c) $L[\frac{df}{dt}] = sF(s) - \lim t \to 0^+ f(t)$

d) $L[\int_0^t f(t')dt'] = \frac{1}{s} F(s)$

e) $\lim_{s\to 0} sF(s) = \lim_{t\to\infty} f(t)$

f) $\lim_{s\to\infty} sF(s) = \lim_{t\to 0} f(t)$

Dimostriamo ad esempio la proprietà c):

$$\begin{aligned} L[\frac{df}{dt}] &= \lim_{a\to 0}\lim_{b\to\infty} \int_a^b e^{-st}\dot{f}(t)dt \\ &= \lim_{a\to 0}\lim_{b\to\infty} \{f(t)e^{-st}|_a^b + s\int_a^b e^{-st} f(t)dt\} \\ &= \lim_{a\to 0}\lim_{b\to\infty} \{f(b)e^{-sb} - f(a)e^{-sa} + s\int_a^b e^{-st} f(t)dt\} \\ &= -f(0^+) + sF(s) \end{aligned}$$

Dove $f(0^+)$ ci ricorda che, a stretto rigore, dobbiamo considerare il valore che la funzione assume quando il tempo tende a 0 dal semiasse positivo. In modo del tutto analogo si ricava la proprietà d).
Dimostriamo invece la proprietà e). Poiche'

$$\int_0^\infty \dot{f}(t)e^{-st}dt = sF(s) - f(0^+)$$

applicando il limite ad ambo i membri si ha

$$\lim_{s\to 0}\int_0^\infty \dot{f}(t)e^{-st}dt = \lim_{s\to 0}(sF(s) - f(0^+))$$

D'altra parte

$$\lim_{s\to 0}\int_0^\infty \dot{f}(t)e^{-st}dt = f(\infty) - f(0^+)$$

quindi

$$f(\infty) - f(0^+) = \lim_{s\to 0} sF(s) - f(0^+)$$

cioé

$$\lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s)$$

come volevasi dimostrare.
Esempi di trasformate di Laplace sono riportati nella Tab. 1.1.
Si noti che la trasformata di Laplace ignora la funzione f(t) per $t < 0$, perché il fattore $e^{-\sigma t}$ divergerebbe per $t \to -\infty$. In pratica quindi la trasformata di Laplace si usa quando si é interessati alla funzione solo per $t \geq 0$, il che nel nostro caso non costituisce certo un problema.

$f(t)$	$F(s)$
k	$\frac{k}{s}$
kt	$\frac{k}{s^2}$
kt^n	$\frac{kn!}{s^{n+1}}$
e^{-at}	$\frac{1}{s+a}$
te^{-at}	$\frac{1}{(s+a)^2}$
$\sin at$	$\frac{a}{s^2+a^2}$
$\cos at$	$\frac{s}{s^2+a^2}$
$e^{-at}\sin \omega t$	$\frac{\omega}{(s+a)^2+\omega^2)}$
$e^{-at}\cos \omega t$	$\frac{s+a}{(s+a)^2+\omega^2}$
$\sinh at$	$\frac{a}{s^2-a^2}$
$\cosh at$	$\frac{s}{s^2-a^2}$
$f(t-t_1)$	$e^{-t_1 s}F(s)$

Tabella 1.1: Trasformate di Laplace di alcune funzioni

1.7 Applicazione alle reti della trasformazione di Laplace

Con la tecnica delle trasformazioni, le equazioni differenziali delle reti si trasformano in equazioni algebriche. Consideriamo ad esempio il circuito RLC serie (vedi Fig 1.10); si trova facilmente l'equazione della maglia

$$v(t) = \{R + \frac{1}{C}\int dt + L\frac{d}{dt}\}i(t) \tag{1.1}$$

Applicando ad ambo i membri la trasformazione di Laplace si ottiene

$$V(s) = RI(s) + \frac{I(s)}{sC} + sLI(s)$$

da cui si ricava

$$I(s) = \frac{V(s)}{R + 1/sC + sL}$$

A questo punto, trovando l'antitrasformata di $I(s)$ si ottiene la funzione incognita $i(t)$. Ponendo $s = j\omega$

$$I(\omega) = \frac{V(\omega)}{R + \frac{1}{j\omega C} + j\omega L}$$

otteniamo un'equazione formalmente identica a quella che si otteneva con il cosidetto metodo simbolico. Infatti quest'ultimo non e' altro che l'applicazione della tecnica della trasformazione di Laplace nel caso particolare di funzioni sinusoidali; nel prossimo paragrafo giustificheremo rigorosamente questa equivalenza.

Figura 1.9: *Rete quadrupolare*

In generale, data una rete quadrupolare (Fig. 1.9) si arriva ad una equazione del tipo

$$V_u(s) = A(s)V_i(s)$$

ovvero

$$A(s) = \frac{V_u(s)}{V_i(s)}$$

La funzione $A(s)$ e' detta funzione di trasferimento della rete. E' interessante notare che, per $V_i(s) = 1$, $V_u(s)$ coincide con $A(s)$. In altre parole, $A(s)$ rappresenta la trasformata della risposta di un circuito per una eccitazione $v_i(t) = \delta(t)$.

La conoscenza della funzione $A(s)$ ci permette di prevedere la risposta del circuito per qualunque eccitazione. In generale essa e' una funzione complessa, per cui dovremo conoscerne il modulo $|A(s)|$ e la fase $\phi(s)$, ovvero $|A(\omega)|$ e $\phi(\omega)$. Di norma si preferisce usare scale logaritmiche; più precisamente si riporta la quantità $20\log_{10}|A|$ in funzione di $\log_{10}(\omega/\omega_0)$, dove ω_0 e' una pulsazione di riferimento arbitraria (spesso e' una pulsazione caratteristica del particolare circuito, ma al limite può anche essere semplicemente la pulsazione unitaria). Il diagramma che ne risulta prende il nome di diagramma di Bode. La grandezza $20\log_{10}|A|$ rappresenta la misura in decibel della funzione di trasferimento. Invece, per quanto riguarda l'angolo di fase, si riporta ϕ (in gradi) in funzione di $\log_{10}(\omega/\omega_0)$.

Vediamo ora una importante conseguenza di due già citate proprietà delle trasformazioni di Laplace, cioe':

$$\lim_{s\to\infty} sF(s) = \lim_{t\to 0} f(t)$$

$$\lim_{s\to 0} sF(s) = \lim_{t\to\infty} f(t)$$

Consideriamo un quadrupolo soggetto ad una sollecitazione a gradino, ovvero

$$v_i(t) = u(t)$$

Applicando la trasformazione di Laplace si ha quindi

$$V_i(s) = \frac{1}{s}$$

e la funzione di trasferimento diviene

$$A(s) = sV_u(s)$$

Applicando le due suddette proprietà ne deriva

$$\lim_{s\to 0} A(s) = \lim_{s\to 0} sV_u(s) = \lim_{t\to\infty} v_u(t)$$

e

$$\lim_{s\to\infty} A(s) = \lim_{s\to\infty} sV_u(s) = \lim_{t\to 0} v_u(t)$$

In sostanza, se $s = j\omega$,

$$A(0) = \lim_{t\to\infty} v_u(t)$$

$$A(\infty) = \lim_{t\to 0^+} v_u(t)$$

Ciò vuol dire che il comportamento asintotico della funzione di trasferimento é legato alla risposta asintotica (a $t \to 0$ e a $t \to \infty$) ad una sollecitazione a gradino. Queste proprietà matematiche hanno un significato fisico, che si può immediatamente comprendere, in modo qualitativo e non rigoroso come segue: lo sviluppo in frequenza di una funzione a gradino contiene un grande addensamento di onde ad altissima frequenza attorno a $t = 0$, mentre per $t \to \infty$ si ha un prevalente contenuto di onde a bassissima frequenza. Quindi la risposta del circuito per $t \to 0$ e' quella tipica delle alte frequenze, mentre per $t \to \infty$ e' quella tipica delle bassissime frequenze. La conoscenza di un circuito può allora essere acquisita in due modi equivalenti: o studiandone la risposta per sollecitazioni sinusoidali, ovvero per sollecitazioni impulsive. Vi sarà una corrispondenza fra alte frequenze e tempi brevi ovvero basse frequenze e tempi lunghi.

Nel prossimo capitolo studieremo la risposta di alcuni circuiti semplici, ma fondamentali, in entrambi i regimi sinusoidale ed impulsivo.

1.8 Anti-trasformazione di Laplace

Abbiamo visto che, in generale, la risposta di un circuito é data da un'espressione del tipo

$$V_u(s) = A(s)V_i(s) \tag{1.2}$$

Per ottenere l'andamento temporale della risposta, cioé la funzione $v_u(t)$, dobbiamo anti-trasformare questa espressione. In generale la A(s) sará una funzione del tipo

$$A(s) = \frac{G(s)}{H(s)} = \frac{a_1 s + a_2 s^2 + \ldots a_m s^m}{b_1 s + b_2 s^2 + \ldots b_n s^n} \tag{1.3}$$

cioé un rapporto tra polinomi, in cui $n \geq m$. Il polinomio $H(s)$ può essere scritto come

$$H(s) = (s - s_1)^{r_1}(s - s_2)^{r_2} \dots (s - s_n)^{r_n} \tag{1.4}$$

dove ovviamente $r_1 + r_2 + \dots r_n = n$. Le costanti $s_1 \dots s_n$ sono sono le radici di $H(s)$ e costituiscono i poli della funzione di trasferimento.
Nel caso di n poli semplici

$$H(s) = (s - s_1)(s - s_2) \dots (s - s_n) \tag{1.5}$$

Si può dimostrare che la $A(s)$ può essere posta nella forma

$$A(s) = \sum_{j=1}^{n} \frac{k_j}{s - s_j} \tag{1.6}$$

Vediamo come si determinano le costanti k_j. Moltiplicando ambo i membri per $(s - s_i)$ si ha

$$(s - s_i)A(s) = (s - s_i) \sum_{j=1}^{n} \frac{k_j}{s - s_j} \tag{1.7}$$

e, passando al limite per $s \to s_i$

$$\lim_{s \to s_i} (s - s_i)A(s) = \lim_{s \to s_i} (s - s_i) \sum_{j=1}^{n} \frac{k_j}{s - s_j} \tag{1.8}$$

Per effetto del limite tutti i termini a secondo membro si annullano, tranne il termine per $j = i$, che vale proprio k_i. Quindi le costanti k_i sono date da

$$k_i = \lim_{s \to s_i} (s - s_i)A(s) \tag{1.9}$$

Una volta che la funzione di trasferimento sia stata posta nella forma 1.6 é facile dimostrare che la sua anti-trasformata é semplicemente data da

$$f(t) = \sum_{j=1}^{n} k_j e^{s_j t} \tag{1.10}$$

Quindi, se un polo é sull'asse reale, l'esponenziale é reale (crescente o decrescente); se il polo é complesso, anche il suo complesso coniugato é un polo della funzione, e si hanno onde sinusoidali moltiplicate per esponenziali reali (se il polo ha una parte reale diversa da zero).

Nel caso di poli multipli il ragionamento é un po' piú complicato, ma analogo, e si ottiene

$$f(t) = \sum_{i=1}^{n} \sum_{g=1}^{m_i} \frac{k_{ig}}{(m_i - g)!} t^{m_i - g} e^{s_i t} \tag{1.11}$$

l'indice g va da 1 a m_i, cioé la molteplicitá del polo i-esimo. I coefficienti k_{ig} si calcolano con la

$$k_{ig} = \lim_{s \to s_i} \frac{1}{(g-1)!} [\frac{d^{g-1}}{ds^{g-1}} (s - s_i)^{m_i} A(s)] \tag{1.12}$$

La collocazione dei poli nel piano complesso caratterizza completamente l'andamento della anti-trasformata $f(t)$. É chiaro infatti che ad un polo puramente immaginario corrisponde un termine sinusoidale, mentre le parti reali danno luogo ad esponenziali reali. Il polinomio a numeratore ,$G(s)$, influisce *solo* sull'ampiezza relativa dei contributi dei vari poli.

Naturalmente lo studio della $A(s)$ non risolve completamente il problema, perché dobbiamo, in generale, anti-trasformare $V_u(s)$, cioé il prodotto $A(s)V_i(s)$.
Ricordiamo però che se l'eccitazione é una delta di Dirac, $\delta(t)$, la sua trasformata é pari ad 1: pertanto la $A(s)$ rappresenta la risposta del circuito per una eccitazione di tal genere.

Sollecitazione impulsiva

É interessante il caso in cui la tensione d'ingresso sia la funzione a gradino $u(t)$. Allora, come abbiamo visto in precedenza, $V_i(s) = 1/s$ e quindi

$$V_u(s) = \frac{A(s)}{s} \tag{1.13}$$

Ricordando la 1.3, possiamo scrivere

$$V_u(s) = \frac{G(s)}{sH(s)} \tag{1.14}$$

questo significa, al massimo, aggiungere un polo per $s = 0$ (che corrisponde a un termine costante nell'antitrasformata).

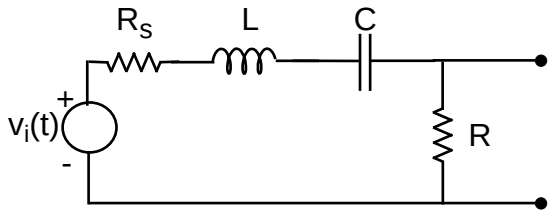

Figura 1.10: *Circuito RLC in serie*

Possiamo prendere, a titolo di esempio, il circuito RLC in serie (Fig. 1.10). Trascuriamo per semplicità la resistenza R_s della sorgente: otteniamo quindi, dopo qualche semplice passaggio, la funzione di trasferimento

$$A(s) = \frac{sRC}{s^2LC + sRC + 1} \tag{1.15}$$

e, per una sollecitazione a gradino,

$$V_u(s) = \frac{RC}{s^2RC + sLC + 1} \tag{1.16}$$

Quindi i poli della funzione possono essere trovati risolvendo l'equazione algebrica

$$s^2RC + sLC + 1 = 0 \tag{1.17}$$

ed é immediato trovare anche le costanti k_i.

Sollecitazione sinusoidale

L'altro caso importante é quello della sollecitazione sinusoidale, che abbiamo finora trattato utilizzando il metodo simbolico. Possiamo a questo punto dedurne in modo rigoroso la validità, partendo appunto dalle trasformazioni di Laplace.
Consideriamo infatti un circuito generico con funzione di trasferimento $A(s)$, soggetto ad una sollecitazione sinusoidale con pulsazione ω_0. La sollecitazione può essere scritta come

$$v_i = |v_i| e^{j\omega_0 t} \tag{1.18}$$

la cui trasformata é data da

$$L[v_i] = \frac{|v_i|}{s - j\omega_0} \tag{1.19}$$

La risposta del circuito sarà quindi

$$V_u(s) = A(s) \frac{|v_i|}{s - j\omega_0}$$

Il secondo membro può essere sviluppato, come abbiamo visto, in una somma di termini tipo

$$\frac{k_j}{(s - s_j)}$$

I poli della $V_u(s)$ saranno dati da tutti i poli della $A(s)$ più il polo per $s = j\omega$. I poli della $A(s)$ devono trovarsi tutti nel semipiano complesso avente parte reale negativa (altrimenti avremmo delle divergenze a tempi lunghi), quindi danno luogo a termini esponenziali tipo e^{-at}, che non interessano quando si voglia conoscere la soluzione stazionaria. Resta il termine corrispondente al polo $s = j\omega_0$ e per la soluzione stazionaria si ha quindi

$$V_u(s) = \frac{k_\omega}{s - j\omega_0} \tag{1.20}$$

Il coefficiente k_ω si determina con il metodo descritto in precedenza: si ottiene

$$k_\omega = A(\omega_0)|v_i| \tag{1.21}$$

e infine

$$V_u(s) = A(\omega_0) \frac{|v_i|}{s - j\omega_0} \tag{1.22}$$

Antitrasformando si ottiene la funzione di risposta

$$v_u(t) = A(\omega_0)|v_i| e^{j\omega_0 t} = |A(\omega_0)||v_i| e^{j\omega_0 t} e^{\phi} \tag{1.23}$$

Ritroviamo il risultato noto, cioé che la risposta ha un andamento sinusoidale, con pulsazione ω_0, ampiezza pari al prodotto di $|v_i|$ per il modulo di $A(\omega_0)$ e sfasamento dato dalla fase di $A(\omega_0)$, cioe' (ϕ). Quindi tutta l'informazione necessaria per conoscere la risposta quando la sollecitazione é sinusoidale, con frequenza ω_0, é data dalla funzione di trasferimento $A(s)$, calcolata nel punto $s = j\omega_0$.

In conclusione, é raramente necessario dover antitrasformare esplicitamente una funzione. Infatti, come abbiamo visto in questo paragrafo, la risposta del circuito per sollecitazioni sinusoidali (ovvero la risposta per un'eccitazione a gradino), ci fornisce tutte le informazioni necessarie per valutare la risposta per segnali d'ingresso di forma qualunque.

Capitolo 2

Circuiti passivi

2.1 Introduzione

In questo capitolo cominceremo a prendere familiarita' con i circuiti studiando approfonditamente alcuni esempi di quadrupoli passivi, esaminandone la risposta per sollecitazioni di tipo sinusoidale ed impulsivo.

2.2 Circuito RC passa-alto

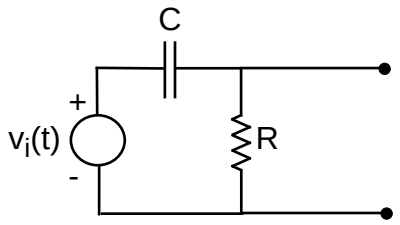

Figura 2.1: *Circuito RC passa-alto*

Consideriamo il circuito di Fig. 2.1; scrivendo l'equazione della maglia si ha:

$$v_i = \frac{1}{C}\int_0^t i dt' + Ri \tag{2.1}$$

avendo fatto l'ipotesi che il condensatore sia inizialmente scarico. Derivando ambo i membri:

$$\frac{dv_i}{dt} = \frac{1}{C}i + R\frac{di}{dt} \tag{2.2}$$

e, ovviamente

$$v_u = Ri \tag{2.3}$$

Risposta in regime sinusoidale

Supponiamo ora che la sollecitazione sia di tipo sinusoidale; possiamo allora utilizzare il metodo delle trasformate cioe', ponendo $s = j\omega$, il metodo simbolico; le equazioni 2.2 e 2.3

divengono:

$$j\omega V_i(\omega) = \frac{1}{C} I(\omega) + j\omega R I(\omega) \tag{2.4}$$

$$V_u(\omega) = R I(\omega) \tag{2.5}$$

e, combinando le due equazioni si ottiene

$$V_u(\omega) = \frac{j\omega R}{\frac{1}{C} + J\omega R} V_i(\omega) \tag{2.6}$$

La funzione di trasferimento del circuito e' quindi data da

$$A(\omega) = \frac{1}{1 + \frac{1}{j\omega\tau}} \tag{2.7}$$

dove abbiamo posto $RC = \tau$. Il fattore τ ha le dimensioni di un tempo, ed e' comunemente indicato come *costante di tempo* del circuito.
Il modulo e la fase della funzione di trasferimento sono quindi

$$|A(\omega)| = \frac{1}{\sqrt{1 + \frac{1}{\omega^2\tau^2}}} \tag{2.8}$$

$$\phi = \arctan \frac{1}{\omega\tau} \tag{2.9}$$

Osserviamo che per $\omega \to 0$ $A(\omega) \to 0$, mentre per $\omega \to \infty$ $|A(\omega)| \to 1$; inoltre se $1/(\omega^2\tau^2) \gg 1$ (cioe' se $\omega \gg 1/\tau$)

$$|A(\omega)| \approx \frac{1}{\sqrt{1/\omega^2\tau^2}} = \omega\tau \tag{2.10}$$

cioé $|A(\omega)|$ tende a zero linearmente con ω. La risposta complessiva del circuito e' quindi rappresentata dal diagramma di Bode in Fig. 2.2. Come si vede il circuito attenua i segnali a bassa frequenza, mentre trasmette le alte frequenze, da cui il nome del circuito stesso. La fase invece tende asintoticamente a 0 per $\omega \to \infty$, e a 90^o per $\omega \to 0$.

La frequenza di taglio, f_T, e' definita come la frequenza a cui la risposta del circuito e' attenuata di $3dB$, cioe'

$$20 \log |A(f_T)| = -3 \tag{2.11}$$

da cui segue che

$$|A(f_T)| \simeq \frac{1}{\sqrt{2}} \tag{2.12}$$

e quindi

$$f_T \simeq \frac{1}{2\pi\tau} \tag{2.13}$$

E' spesso conveniente scrivere la funzione di trasferimento come

$$A(f) = \frac{1}{1 - j\frac{f_T}{f}} \tag{2.14}$$

Possiamo allora riassumere il comportamento di questo circuito: per frequenze molto inferiori alla frequenza di taglio la funzione di trasferimento e' una retta con pendenza di 20 $dB/decade$, mentre al di sopra di essa e' una retta orizzontale.

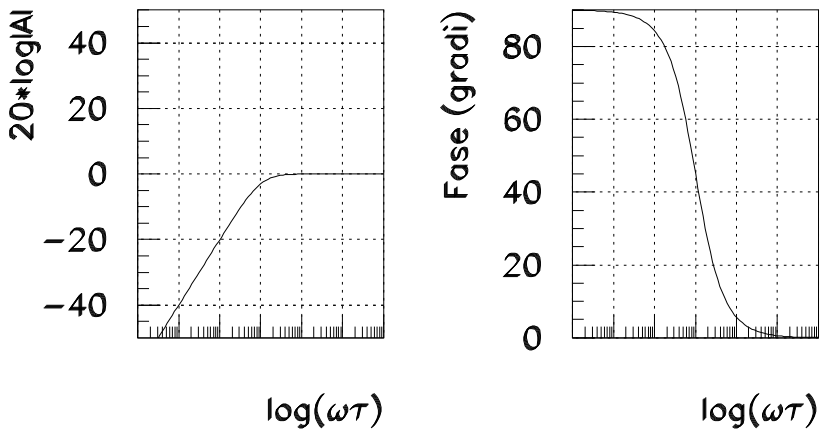

Figura 2.2: *Diagramma di Bode del circuito RC passa-alto*

Risposta in regime impulsivo

Per studiare la risposta del circuito in regime impulsivo possiamo porre come segnale d'ingresso una funzione a gradino $u(t)$ e risolvere direttamente l'equazione differenziale della maglia

$$\frac{dv_i}{dt} = \frac{1}{C}i + R\frac{di}{dt} \tag{2.15}$$

Per $t > 0$ il primo membro si annulla e si ha dunque un'equazione differenziale omogenea

$$i + RC\frac{di}{dt} = 0 \tag{2.16}$$

La cui soluzione e' una funzione del tipo

$$i(t) = i_0 e^{-\frac{t}{\tau}} \tag{2.17}$$

e quindi

$$v_u(t) = v_0 e^{-\frac{t}{\tau}} \tag{2.18}$$

La costante v_0 deve essere determinata dalle condizioni iniziali. Sfruttiamo in questo caso la relazione

$$v_u(0) = \lim_{\omega\to\infty} A(\omega) = 1 \tag{2.19}$$

da cui si ottiene immediatamente

$$v_u(t) = e^{-\frac{t}{\tau}} \tag{2.20}$$

In generale se $v_i(t) = v_r u(t)$, si avra'

$$v_u(t) = v_r e^{-\frac{t}{\tau}} \tag{2.21}$$

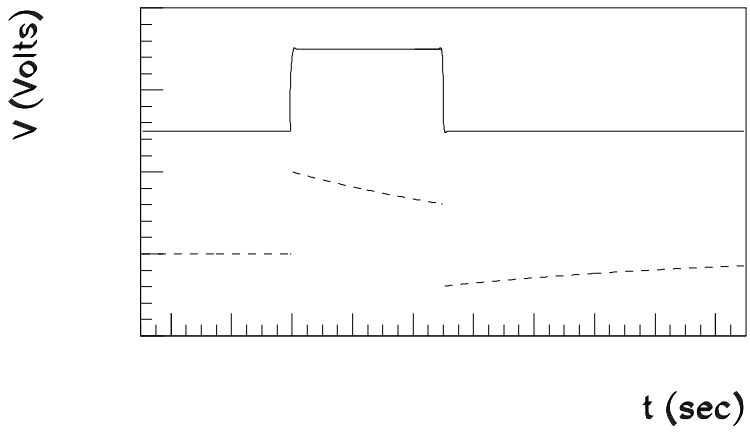

Figura 2.3: *Risposta del circuito passa-alto in regime impulsivo*

Un segnale d'ingresso rettangolare, di periodo T, puo' essere opportunamente descritto come combinazione di 2 funzioni a gradino

$$v_i(t) = v_r[u(t) - u(t-T)] \tag{2.22}$$

La risposta sara' data da (vedi Fig. 2.3)

$$v_u(t) = v_r e^{-\frac{t}{\tau}} \quad per \quad 0 < t < T \tag{2.23}$$

e

$$v_u(t) = v_r[e^{-\frac{t}{\tau}} - e^{-\frac{t-T}{\tau}}] \quad per \quad t > T \tag{2.24}$$

La risposta per una sollecitazione rettangolare periodica puo' essere facilmente dedotta a partire da quella relativa al singolo impulso.

2.3 Circuito RC Passa-Basso

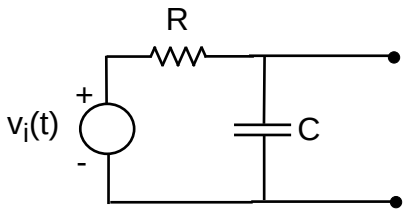

Figura 2.4: *Circuito RC passa-basso*

Consideriamo ora il circuito in Fig. 2.4. Nel caso di sollecitazione sinusoidale, con il

metodo delle trasformate si ha:

$$V_i = (R + \frac{1}{j\omega C})I \tag{2.25}$$

$$V_u = \frac{I}{j\omega C} \tag{2.26}$$

da cui si ricava

$$V_u = \frac{\frac{1}{j\omega C}}{R + \frac{1}{j\omega C}} V_i \tag{2.27}$$

La funzione di trasferimento del circuito e' quindi

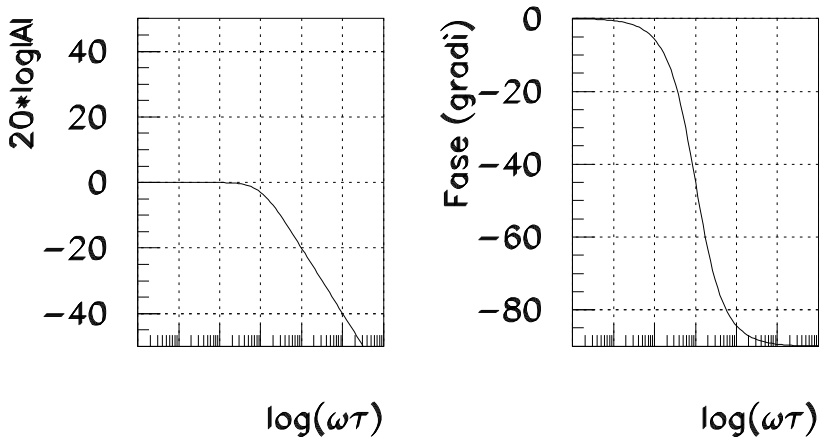

Figura 2.5: *Diagramma di Bode del circuito RC passa-basso*

$$A(\omega) = \frac{1}{1 + j\omega\tau} \tag{2.28}$$

avendo anche qui posto $RC = \tau$. Il modulo e la fase della funzione di trasferimento sono allora

$$|A(\omega)| = \frac{1}{\sqrt{1 + \omega^2\tau^2}} \tag{2.29}$$

$$\phi(\omega) = -\arctan\omega\tau \tag{2.30}$$

I valori asintotici della funzione di trasferimento sono

$$|A(0)| = 1 \quad |A(\infty)| = 0 \tag{2.31}$$

in questo caso, se $\omega^2\tau^2 \gg 1$, cioé se $\omega \gg 1/\tau$

$$|A(\omega)| \approx \frac{1}{\omega\tau} \tag{2.32}$$

La risposta complessiva del circuito e' quindi rappresentata dal diagramma di Bode in Fig. 2.5. Come si vede, il circuito ora attenua le frequenze alte, lasciando imperturbati i segnali di bassa frequenza. Anche in questo caso si definisce la frequenza di taglio, f_T, corrispondente ad una attenuazione di 3 dB. Si ha chiaramente

$$f_T = \frac{1}{2\pi\tau} \tag{2.33}$$

Per frequenze molto maggiori della frequenza di taglio la curva di risposta ha una pendenza di -20 $dB/decade$.

Risposta in regime impulsivo

Possiamo, come nel caso precedente, risolvere direttamente l'equazione differenziale della maglia, avendo come segnale d'ingresso una sollecitazione a gradino $v_r u(t)$; si trova molto facilmente che

$$v_u = v_r(1 - e^{-\frac{t}{\tau}}) \tag{2.34}$$

cioé il circuito in questo caso distorce il segnale d'ingresso ai tempi brevi (vedi Fig. 2.6). Un modo per parametrizzare questa distorsione é attraverso l'introduzione del *tempo di salita*, t_s, definito come il tempo impiegato dalla tensione di uscita per passare dal 10% al 90% del suo valore massimo. Chiamando t_1 e t_2 gli estremi di questo intervallo ($t_s = t_2 - t_1$) si ha

$$v_r(1 - e^{-\frac{t_1}{\tau}}) = .1v_r \tag{2.35}$$

$$v_r(1 - e^{-\frac{t_2}{\tau}}) = .9v_r \tag{2.36}$$

Dividendo tra loro le due equazioni si ottiene

$$e^{-\frac{t_2 - t_1}{\tau}} = 1/9 \tag{2.37}$$

da cui

$$t_s = \tau \ln 9 \approx 2.2\tau \tag{2.38}$$

Possiamo poi porre in relazione tempo di salita e frequenza di taglio: dai risultati precedenti si ottiene

$$t_s \approx .35 f_T \tag{2.39}$$

2.4 Derivatore e integratore

Consideriamo di nuovo il circuito passa-alto. Dall'equazione gia' vista si ha che

$$\frac{dv_i}{dt} = \frac{v_u}{RC} + \frac{dv_u}{dt} \tag{2.40}$$

cioé

$$\frac{d(v_i - v_u)}{dt} = \frac{v_u}{RC} \tag{2.41}$$

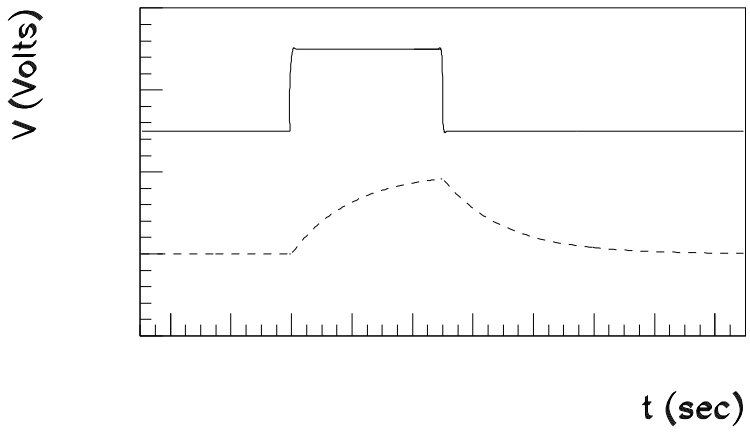

Figura 2.6: *Risposta impulsiva del passa-basso*

ora, se $v_u \ll v_i$

$$v_u \approx RC\frac{dv_i}{dt} \tag{2.42}$$

quindi v_u é circa proporzionale alla derivata del segnale d' ingresso[1].
Consideriamo invece il circuito passa-basso. Possiamo riscrivere l'equazione della maglia

$$v_i = Ri + v_u \tag{2.43}$$

ovvero

$$v_i - v_u = Ri \tag{2.44}$$

integrando ambo i membri

$$\int (v_i - v_u)dt = R\int idt \tag{2.45}$$

se $v_i \gg v_u$

$$\int v_i dt \simeq R\int idt \tag{2.46}$$

$$\simeq RCv_u \tag{2.47}$$

cioé il circuito si comporta da integratore, se la caduta di tensione su C é piccola rispetto a quella su R.

Possiamo vedere la cosa da un punto di vista diverso. Un derivatore *ideale* dovrebbe essere un circuito in cui la tensione di uscita e' proporzionale alla derivata della tensione d'ingresso, ovvero

$$v_u = k\frac{dv_i}{dt} \tag{2.48}$$

[1]La condizione $v_u \ll v_i$ equivale a dire che la caduta di tensione ai capi di R deve essere molto piccola rispetto a quella ai capi di C.

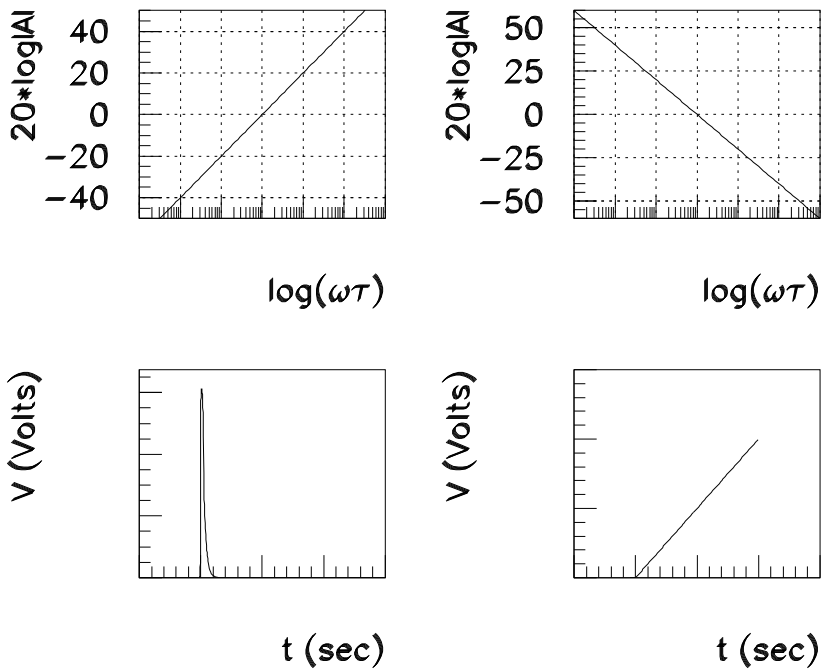

Figura 2.7: *a) Diagramma di Bode del derivatore ideale;*
b) Diagramma di Bode dell'integratore ideale;
c) Risposta impulsiva del derivatore ideale;
d) Risposta impulsiva dell'integratore ideale;

in termini di trasformate si dovrebbe quindi avere

$$V_u(\omega) = k\omega V_i(\omega) \tag{2.49}$$

In sostanza, la funzione di trasferimento del circuito dovrebbe essere

$$A(\omega) = k\omega \tag{2.50}$$

Il corrispondente diagramma di Bode e' mostrato il Fig. 2.7a. E' chiaro allora che il circuito RC si approssima al derivatore *ideale* per $\omega\tau \ll 1$, ovvero per frequenze basse, dove l' impedenza di C é alta.

Viceversa l'integratore *ideale* ha una funzione di trasferimento

$$A(\omega) = \frac{1}{\omega k} \tag{2.51}$$

(Fig. 2.7b). Quindi la funzione di trasferimento del RC passa basso approssima bene l'integratore ideale per $\omega\tau \gg 1$, nella regione delle alte frequenze, dove l'impedenza di C é bassa.

In termini di forme d'onda, per un segnale d'ingresso a gradino, si avrebbero in uscita le forme d'onda mostrate in Fig 2.7c e 2.7d.

Queste proprieta' dei circuiti RC sono spesso utilizzate nella elaborazione dei segnali impulsivi. Infatti un segnale impulsivo porta con se' generalmente due informazioni: una legata al suo tempo di arrivo, l'altra legata alla sua ampiezza (ovvero spesso all'integrale nel tempo dell'ampiezza). Poiche' i segnali reali non sono mai esattamente rettangolari, ma hanno sempre un tempo di salita diverso da zero (seppur piccolo) il tempo di arrivo del segnale puo' risultare piu' o meno mal definito e mal misurabile: facendo passare il segnale attraverso un derivatore, il tempo di salita viene diminuito drasticamente e il tempo di arrivo del segnale risulta quindi piu' facilmente misurabile. Viceversa, se si e' interessati all'area sotto il segnale, l'integratore e' il circuito adatto per fornire meglio questa informazione.

2.5 Attenuatore compensato

Gli effetti distorcenti dovuti alle capacita' non si manifestano solo quando si costruiscono dei circuiti RC, ma anche, spesso, per la presenza di capacita' parassite. L'esempio tipico e' quello dell'oscillografo: noi preleviamo un segnale tra due nodi di un circuito (tipicamentre tra un nodo e la massa) e lo portiamo all'ingresso dell'oscillografo. Quest'ultimo e' caratterizzato da una resistenza d'ingresso, con in parallelo una capacita', quindi l'operazione di *trasferimento* del segnale avviene attraverso una rete del tipo esemplificato in Fig. 2.8a, in cui R_1 rappresenta la resistenza d'uscita del nodo sotto esame, mentre R_2 e C_2 rappresentano l'impedenza d'ingresso dell'oscillografo. Questo circuito si comporta come un passa-basso (per capirlo basta fare l'equivalente di Thevenin del circuito *a monte* di C_2), con una costante di tempo

$$\tau = C_2 \frac{R_1 R_2}{R_1 + R_2} \tag{2.52}$$

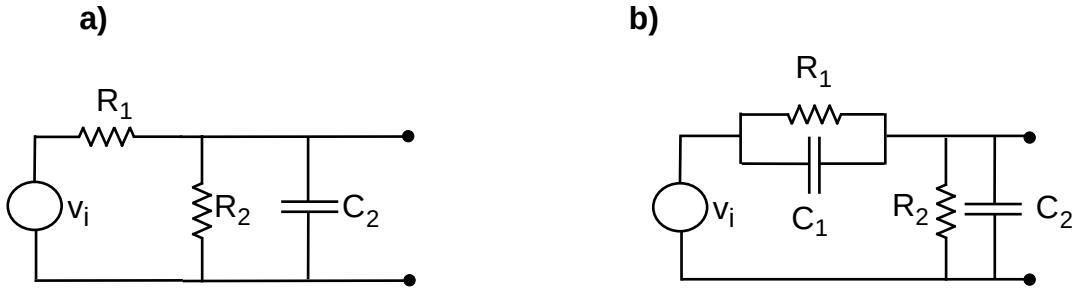

Figura 2.8: *a) Partitore resistivo con capacita' parassita; b) Partitore compensato*

Una situazione del genere si incontra di frequente, ovvero tutte le volte che si trasferisce un segnale da un quadrupolo ad un altro: alla partizione resistiva si aggiunge una distorsione dovuta alla (ineliminabile) capacita' parassita d'ingresso.
Se questa distorsione non e' accettabile, conviene aggiungere una capacita' C_1 in parallelo ad R_1 Fig. 2.8b. Possiamo allora scrivere le due equazioni:

$$V_i = \frac{R_1}{1+j\omega R_1 C_1} I + \frac{R_2}{1+j\omega R_2 C_2} I \tag{2.53}$$

$$V_u = \frac{R_2}{1+j\omega R_2 C_2} I \tag{2.54}$$

da cui ricaviamo la funzione di trasferimento

$$A(\omega) = \frac{\frac{R_2}{1+j\omega R_2 C_2}}{\frac{R_1}{1+j\omega R_1 C_1} + \frac{R_2}{1+j\omega R_2 C_2}} \tag{2.55}$$

$$= \frac{1}{1+\frac{R_1}{R_2}\frac{1+j\omega R_2 C_2}{1+j\omega R_1 C_1}} \tag{2.56}$$

Si vede subito che, se $R_1 C_1 = R_2 C_2$ il partitore é perfettamente *compensato*, cioé non distorce il segnale. Infatti, in questo caso si ha:

$$A(\omega) = \frac{1}{1+\frac{R_1}{R_2}} \tag{2.57}$$

E' tuttavia interessante notare che, se $R_2 C_2 \neq R_1 C_1$, alle basse frequenze, cioé per $\omega R_2 C_2 \ll 1$ e $\omega R_1 C_1 \ll 1$ si ha ancora

$$A(\omega) \approx \frac{1}{1+\frac{R_1}{R_2}} \tag{2.58}$$

Alle alte frequenze, cioé per $\omega R_2 C_2 \gg 1$ e $\omega R_1 C_1 \gg 1$ si ha invece

$$A(\omega) = \frac{1}{1+\frac{C_1}{C_2}} = \frac{C_1}{C_1+C_2} \tag{2.59}$$

cioé l'attenuazione dipende solo dal rapporto delle capacitá.

Possiamo comprendere in modo piu' intuitivo questo risultato, ridisegnando il circuito (Fig. 2.9a). Chiaramente, se

$$\frac{R_1}{R_2} = \frac{C_2}{C_1} \tag{2.60}$$

tra i punti A e B del circuito non vi e' differenza di potenziale, quindi il ramo che li connette puo' essere omesso e si vede che l'uscita e' data solo dalla partizione R_1, R_2. Nella realta'

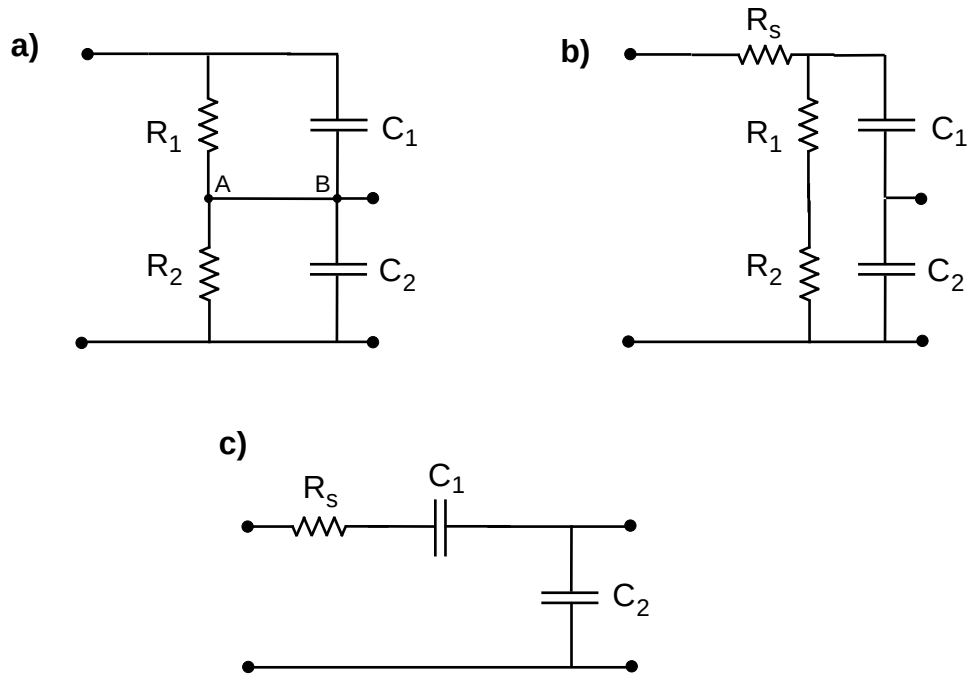

Figura 2.9: *a) Partitore compensato ridisegnato; b) Un caso piu' realistico; c) Equivalente di Thevenin approssimato*

non sempre e' possibile mettere un capacitore C_1 in parallelo ad R_1: spesso quest'ultima non e' altro che la resistenza d'uscita del primo stadio (quindi non accessibile). Il caso piu' realistico e' quindi descritto in Fig. 2.9b, in cui preleviamo l'uscita attraverso un partitore compensato. Ora pero', se $R_s \ll R_1 + R_2$, attraverso il teorema di Thevenin, arriviamo al circuito mostrato in Fig. 2.9c. Il tempo di salita é dato da

$$t'_s = 2.2\, R_s \frac{C_1 C_2}{C_1 + C_2} \tag{2.61}$$

Se non ci fossero C_1 ed R_1, si avrebbe un tempo di salita

$$t_s = 2.2\, R_s C_2 \tag{2.62}$$

ed il rapporto tra i due

$$\frac{t_s}{t'_s} = \frac{C_1 + C_2}{C_1} \tag{2.63}$$

Quindi, aggiungendo il condensatore C_1 possiamo ridurre il tempo di salita. Il prezzo che paghiamo e' nella maggiore attenuazione del segnale: infatti se vogliamo guadagnare un fattore 10 nel tempo di salita, dobbiamo attenuare dello stesso fattore il segnale.

La sonda dell'oscillografo

Le considerazioni precedenti sono abbastanza vere anche in assenza di perfetta compensazione, e sono alla base del funzionamento delle *sonde* che corredano gli oscillografi. Prendiamo ad esempio un'oscillografo con resistenza d'ingresso 1 $M\Omega$ e capacita' d'ingresso 10 pF: supponiamo di connettere, tramite un normale cavetto, l'oscillografo ad un particolare punto, o meglio, nodo di un circuito e che la resistenza d'uscita di questo nodo sia 100 k. Il tempo di salita del segnale diviene

$$t_s = 2.2 \times 100\ k \times 10\ pF = 2.2\ \mu s \tag{2.64}$$

cioe' piuttosto alto; utilizzando invece una sonda con attenuazione 10 il tempo di salita migliora dello stesso fattore.

2.6 Filtri in cascata

2.6.1 Doppio passa-basso

Consideriamo il circuito in Fig. 2.10 Applicando il metodo dei nodi, per il nodo 1 abbiamo

$$\frac{V_i - V_1}{R_1} = \frac{V_1}{Z_1} + \frac{V_1}{R_2 + Z_2} \tag{2.65}$$

mentre la tensione d'uscita e' data da

$$V_0 = \frac{Z_2}{R_2 + Z_2} V_1 \tag{2.66}$$

dove abbiamo indicato con Z_1 e Z_2 le impedenze dei due condensatori. Dalla prima relazione ricaviamo

$$\frac{V_1}{V_i} = \frac{1}{R_1(\frac{1}{R_1} + \frac{1}{Z_1} + \frac{1}{R_2 + Z_2})} \tag{2.67}$$

$$= \frac{1}{1 + \frac{R_1}{Z_1} + \frac{R_1}{R_2 + Z_2}} \tag{2.68}$$

per ottenere la funzione di trasferimento combiniamo la 2.66 con la 2.68. Infatti

$$A(\omega) = \frac{V_0}{V_1}\frac{V_1}{V_i} \tag{2.69}$$

$$= \frac{1}{1 + \frac{R_2}{Z_2}} \frac{1}{1 + \frac{R_1}{Z_1} + \frac{R_1}{R_2 + Z_2}} \tag{2.70}$$

$$= \frac{1}{1 + j\omega C_2 R_2} \frac{1}{1 + j\omega C_1 R_1 + \frac{R_1}{R_2 + 1/j\omega C_2}} \tag{2.71}$$

$$= \frac{1}{1 - \omega^2 C_1 R_1 C_2 R_2 + j\omega C_1 R_1 + j\omega C_2 (R_2 + R_1)} \tag{2.72}$$

$$\tag{2.73}$$

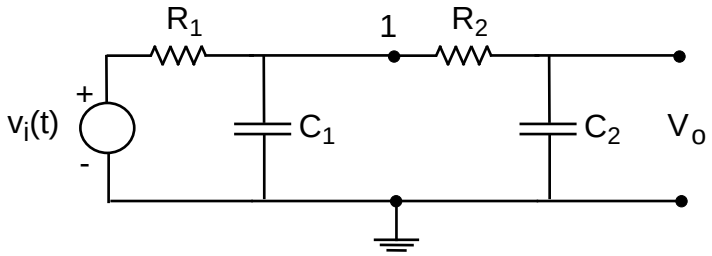

Figura 2.10: *Due passa-basso in cascata*

Ora, se $R_2 \gg R_1$:

$$A \approx \frac{1}{1-\omega^2 C_1 R_1 C_2 R_2 + j\omega C_1 R_1 + j\omega C_2 R_2} \tag{2.74}$$

$$= \frac{1}{(1+j\omega C_1 R_1)(1+j\omega C_2 R_2)} \tag{2.75}$$

$$\tag{2.76}$$

A coincide col prodotto delle funzioni di trasferimento dei 2 passa-basso in cascata. Questo esempio ci fa comprendere che, in generale, la funzione di trasferimento di due quadrupoli in cascata non e' uguale al prodotto delle due funzioni di trasferimento, perche' il comportamento del primo stadio e' perturbato dalla presenza del secondo. Quand'e' che questa perturbazione puo' essere trascurata? Quando l'impedenza d'ingresso del secondo stadio e' molto maggiore dell'impedenza d'uscita del primo. Infatti se ora consideriamo il doppio passa-basso alla luce di questa affermazione, osserviamo che l'impedenza di ingresso del secondo stadio e'data da:

$$Z_{i2} = R_2 + \frac{1}{j\omega C_2} \tag{2.77}$$

mentre l'impedenza d'uscita del primo stadio (applicando il teorema di Thevenin) e':

$$\frac{1}{Z_{o1}} = \frac{1}{R_1} + j\omega C_1 \tag{2.78}$$

Volendo avere la condizione di non perturbazione a tutte le frequenze, si deve imporre proprio che $R_2 \gg R_1$.
Possiamo sfruttare questa proprieta' per costruire filtri passa-basso con una selettivita' migliore. Infatti, se scegliamo i componenti in modo che $R_1 C_1 = R_2 C_2$ (mantenendo naturalmente $R_2 \gg R_1$), otteniamo un filtro con una discesa asintotica di 40 $dB/decade$

2.6.2 Doppio passa-alto

Naturalmente le stesse considerazioni si applicano al doppio passa-alto, cioe' il circuito in Fig. 2.11. Applicando anche qui il metodo dei nodi si ha:

$$\frac{V_i - V_1}{Z_1} = \frac{V_1}{R_1} + \frac{V_1}{R_2 + Z_2} \tag{2.79}$$

$$V_0 = \frac{R_2}{R_2 + Z_2} V_1 \tag{2.80}$$

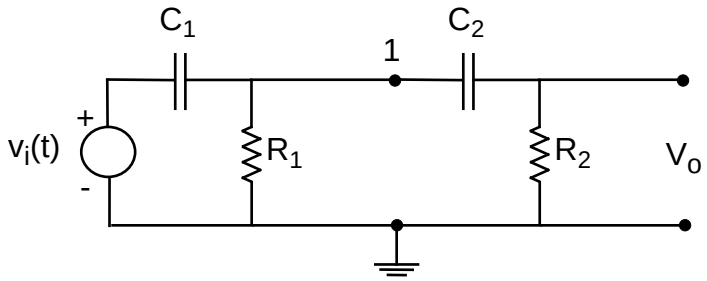

Figura 2.11: *Due passa-alto in cascata*

e, sviluppando il calcolo, analogamente al caso precedente, si arriva a

$$A = \frac{1}{1+\frac{1}{j\omega R_2 C_2}} \frac{1}{1+\frac{1}{j\omega R_1 C_1}+\frac{\frac{1}{j\omega C_1 R_2}}{1+\frac{1}{j\omega C_2 R_2}}} \tag{2.81}$$

$$= \frac{1}{1-\frac{1}{\omega^2 R_1 C_1 R_2 C_2}+\frac{1}{j\omega R_1 C_1}+\frac{1}{j\omega C_2 R_2}+\frac{1}{j\omega C_1 R_2}} \tag{2.82}$$

Anche in questo caso, se $R_2 \gg R_1$,

$$\frac{1}{j\omega C_1 R_1} \gg \frac{1}{j\omega C_1 R_2} \tag{2.83}$$

quindi l'ultimo addendo si trascura, e si ha

$$A \approx \frac{1}{(1+\frac{1}{j\omega R_1 C_1})(1+\frac{1}{j\omega R_2 C_2})} \tag{2.84}$$

cioé il prodotto di due passa-alto.

2.6.3 Passa-banda

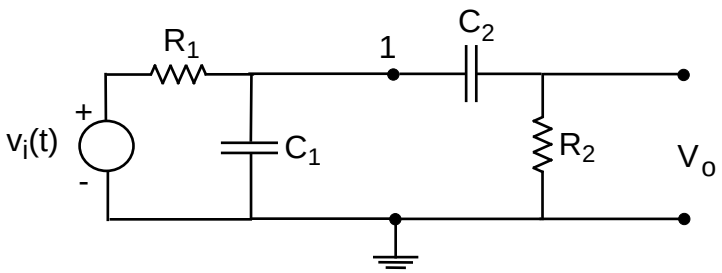

Figura 2.12: *Passa-banda*

E' possibile, utilizzando un passa-basso ed un passa-alto, ottenere un circuito passa-banda. Naturalmente dovremo scegliere i componenti in modo che la frequenza di taglio

del passa-basso, f_2, sia superiore alla frequenza di taglio del passa-alto, f_1. La larghezza di banda sara' allora data da

$$B = f_2 - f_1 \tag{2.85}$$

Abbiamo

$$\frac{V_i - V_1}{R_1} = \frac{V_1}{Z_1} + \frac{V_1}{R_2 + Z_2} \tag{2.86}$$

$$V_0 = \frac{R_2}{R_2 + Z_2} V_1 \tag{2.87}$$

Con il solito metodo si arriva a

$$A = \frac{1}{1 + \frac{1}{j\omega R_2 C_2}} \frac{1}{1 + j\omega R_1 C_1 + \frac{j\omega C_2 R_1}{1 + J\omega C_2 R_2}} \tag{2.88}$$

$$= \frac{1}{\frac{1 + j\omega R_2 C_2}{j\omega R_2 C_2}} \frac{1}{\frac{1 - \omega^2 R_1 C_1 R_2 C_2 + j\omega R_1 C_1 + j\omega R_2 C_2 + j\omega C_2 R_1}{1 + j\omega R_2 C_2}} \tag{2.89}$$

Se $R_2 \gg R_1$ l'ultimo addendo si trascura, e si ha

$$A \approx \frac{1}{\frac{1 - \omega^2 R_1 C_1 R_2 C_2 + j\omega R_1 C_1 + j\omega R_2 C_2}{j\omega R_2 C_2}} \tag{2.90}$$

che é uguale al prodotto di un passa-basso e di un passa-alto, cioé

$$A = \frac{1}{(1 + \frac{1}{j\omega R_2 C_2})} \frac{1}{(1 + j\omega R_1 C_1)} \tag{2.91}$$

2.7 Circuiti RLC

I filtri passa-banda sono molto importanti e necessari per numerosissime applicazioni. In alcuni casi e' sufficiente il circuito che abbiamo visto nel paragrafo precedente, ma in genere, se si ha bisogno di un circuito molto *selettivo*, si devono introdurre degli induttori, per costruire circuiti RLC.
Notiamo infatti che, se abbiamo un induttore ed un capacitore in serie, l'impedenza complessiva e'

$$Z_S = j\omega L + \frac{1}{j\omega C} \tag{2.92}$$

$$= j(\omega L - \frac{1}{\omega C}) \tag{2.93}$$

Si vede quindi che Z_S si annulla se

$$\omega = \frac{1}{\sqrt{LC}} \tag{2.94}$$

mentre diviene infinita per $\omega = \infty$ o per $\omega = 0$.

Se invece prendiamo il parallelo tra un induttore e un capacitore abbiamo un comportamento analogo per l'ammettenza complessiva

$$Y_P = j\omega C + \frac{1}{j\omega L} \tag{2.95}$$

$$= j(\omega C - \frac{1}{\omega L}) \tag{2.96}$$

In questo caso l'impedenza del parallelo, Z_P, diviene infinita quando

$$\omega = \frac{1}{\sqrt{LC}} \tag{2.97}$$

Questa frequenza critica del circuito si chiama, in entrambi i casi, *frequenza di risonanza* del sistema; i circuiti che ora vedremo sfruttano questi andamenti delle impedenze dei sistemi LC per ottenere una andamento della funzione di trasferimento selettivo in frequenza.

2.7.1 RLC in serie

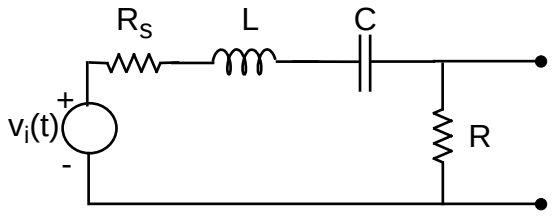

Figura 2.13: *Circuito RLC in serie*

Cosideriamo anzitutto il circuito RLC in serie (Fig. 2.13). Per eccitazioni sinusoidali avremo:

$$V_i = (R_s + R + j\omega L + \frac{1}{j\omega C})I \tag{2.98}$$

$$V_u = RI \tag{2.99}$$

Combinando le due equazioni si ottiene la funzione di trasferimento

$$A(\omega) = \frac{R}{R_s + R + j(\omega L - \frac{1}{\omega C})} \tag{2.100}$$

$$= \frac{1}{\frac{R_s + R}{R} + \frac{j}{R}(\omega L - \frac{1}{\omega C})} \tag{2.101}$$

Introducendo le variabili

$$\omega_0 = \frac{1}{\sqrt{LC}} \tag{2.102}$$

$$Q_0 = \frac{\omega_0 L}{R} \tag{2.103}$$

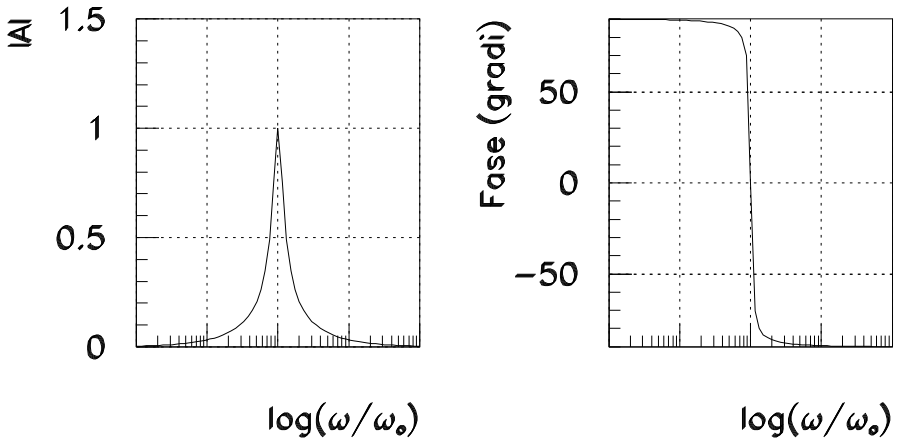

Figura 2.14: *Funzione di trasferimento di un circuito RLC serie, con* $Q_0 = 10$ *(si e' preso* $R >> R_s$*).*

la 2.101 puo' essere scritta

$$A(\omega) = \frac{1}{\frac{R_s + R}{R} + jQ_0(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega})} \tag{2.104}$$

Il modulo e la fase di A sono quindi date da

$$|A(\omega)| = \frac{1}{\sqrt{(\frac{R_s + R}{R})^2 + Q_0^2(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega})^2}} \tag{2.105}$$

$$\phi(\omega) = \arctan[-\frac{R}{R_s + R}Q_0(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega})] \tag{2.106}$$

Come si vede dalla Fig. 2.14 il modulo di A ha l'andamento richiesto. Alla frequenza di risonanza si ha

$$|A(\omega_0)| = \frac{R}{R + R_s} \tag{2.107}$$

mentre la larghezza del picco attorno al valore di risonanza diminuisce al crescere di Q_0, che, per questo motivo, prende il nome di *fattore di merito*. Per una data frequenza di risonanza Q_0 dipende da R ed L; si potrebbe pensare di aumentare il fattore di merito riducendo R. Pero' in questo modo diminuisce anche l'ampiezza sul picco (al limite, per $R = 0$ il fattore di merito e' infinito, ma il segnale in uscita avrebbe ampiezza nulla!). Inoltre R e' sostanziamente, nei casi reali, il carico rappresentato dallo stadio successivo del circuito, quello cioe' in cui la frequenza filtrata deve essere utilizzata ed e' quindi determinato anche da altre considerazioni. Un discorso analogo vale per R_s (resistenza d'uscita del generatore, ovvero dello stadio precedente): sarebbe comodo avere R_s molto piccolo, per poter ridurre conseguentemente R ed avere un alto fattore di merito, ma non e' detto che cio' sia fattibile in pratica. Si noti poi che l'induttore ha intrinsecamente una resistenza parassita, r, che e' in genere non trascurabile (puo' anche essere di decine di Ohm); le nostre formule restano invariate se si interpreta R_s come la somma di r e della

resistenza d'uscita del generatore, ma e' quindi chiaro che R_s non puo' essere mai annullata del tutto.
E' interessante notare che il fattore di merito é anche definibile in termini di energia come il rapporto tra il valore max dell'energia accumulata e l'en. dissipata in un periodo, moltiplicato per 2π

Il modulo della funzione di trasferimento gode di una particolare simmetria (simmetria geometrica): infatti, per ogni valore $\omega' < \omega_0$ esiste un valore $\omega'' > \omega_0$ per cui

$$|A(\omega')| = |A(\omega'')| \tag{2.108}$$

ed inoltre

$$\omega'\omega'' = \omega_0^2 \tag{2.109}$$

La dimostrazione e' molto facile: infatti la condizione 2.108 implica che

$$(\frac{\omega'}{\omega_0} - \frac{\omega_0}{\omega'}) = -(\frac{\omega''}{\omega_0} - \frac{\omega_0}{\omega''}) \tag{2.110}$$

da cui si ottiene subito la 2.109.
Possiamo infine introdurre la *larghezza di banda*, B, definita come la banda compresa tra le due frequenze, ω_1 e ω_2, dove il modulo di A diminuisce di 3 dB rispetto al valore massimo, cioe'

$$B = \frac{\omega_2 - \omega_1}{2\pi} \tag{2.111}$$

dove

$$|\frac{A(\omega_1)}{A(\omega_0)}| = |\frac{A(\omega_2)}{A(\omega_0)}| = \frac{1}{\sqrt{2}} \tag{2.112}$$

E' facile intuire che B e' correlata a Q_0; dalle definizioni (sfruttando la simmetria geometrica), si trova facilmente che

$$B = \frac{\omega_0}{2\pi Q} \tag{2.113}$$

e viceversa

$$Q = \frac{\omega_0}{2\pi B} = \frac{\omega_0}{\omega_2 - \omega_1} \tag{2.114}$$

Quindi il parametro Q_0 puo' essere valutato sperimentalmente misurando ω_0, ω_1 ed ω_2.

2.7.2 Circuito RLC parallelo

Possiamo ora studiare il circuito RLC in parallelo (Fig. 2.15a). Ci conviene trasformare il circuito utilizzando il teorema di Norton (Fig. 2.15b) e scrivere quindi l'equazione del nodo d'uscita

$$\frac{V_s}{R_s} = V_o(\frac{1}{R} + \frac{1}{R_s} + \frac{1}{j\omega L} + j\omega C) \tag{2.115}$$

Chiamando R_P il parallelo tra R ed R_s si ottiene la funzione di trasferimento

$$A(\omega) = \frac{1}{\frac{R_s}{R_P} + jR_s(\omega C - \frac{1}{\omega L})} \tag{2.116}$$

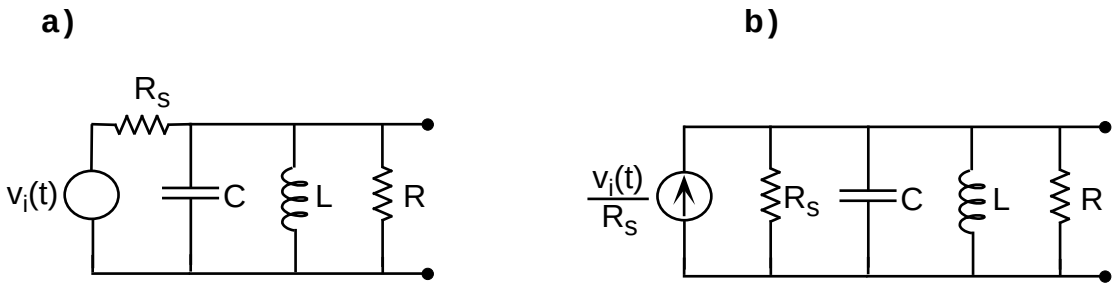

Figura 2.15: *a) Circuito RLC in parallelo; b) Circuito equivalente*

Definendo

$$\omega_0 = \frac{1}{\sqrt{LC}} \tag{2.117}$$

$$Q_0 = \frac{R_s}{\omega_0 L} \tag{2.118}$$

possiamo esprimere A nella consueta forma

$$A(\omega) = \frac{1}{\frac{R_s}{R_P} + jQ_0(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega})} \tag{2.119}$$

La funzione 2.119 e' del tutto analoga alla 2.104. Si noti che l'ampiezza sul picco e' data ora da

$$|A(\omega_0)| = \frac{R_P}{R_s} \tag{2.120}$$

quindi, se $R >> R_s$, $R_P \simeq R_s$ e l'ampiezza dell'uscita e' massima. D'altra parte ora il fattore di merito cresce al crescere di R_s; potremmo quindi incrementare R_s (aggiungendo per es. una resistenza in serie) per migliorare il fattore di merito, ma cio' penalizzerebbe l'ampiezza dell'uscita, a meno di non incrementare anche R (cioe' la resistenza d'ingresso dello stadio successivo). Di fatto, anche in questo caso, i fattori di merito realmente ottenibili sono abbastanza limitati. Vedremo nel Cap. 6 come, con l'uso di circuiti *attivi*, sia possibile ottenere fattori di merito molto piu' elevati.

Il comportamento di questo circuito e' pero' influenzato in modo complesso dalla presenza della resistenza parassita dell'induttore. Conviene quindi esaminare in dettaglio il circuito completo (Fig. 2.16a): riscriviamo l'equazione del nodo di uscita:

$$\frac{V_i}{R_s} = V_o(\frac{1}{R_P} + j\omega C + \frac{1}{Z_L}) \tag{2.121}$$

dove

$$Z_L = r + j\omega L \tag{2.122}$$

Possiamo osservare che

$$\frac{1}{Z_L} = \frac{1}{r + j\omega L} \tag{2.123}$$

$$= \frac{r - j\omega L}{r^2 + \omega^2 L^2} \tag{2.124}$$

$$= \frac{r}{r^2 + \omega^2 L^2} - \frac{j\omega L}{r^2 + \omega^2 L^2} \tag{2.125}$$

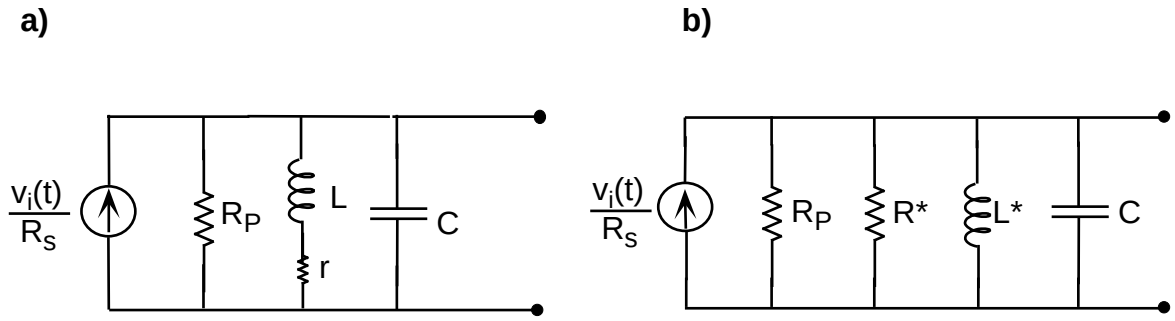

Figura 2.16: *a) Circuito RLC in parallelo; b) Circuito equivalente*

L'equazione 2.121 diviene allora:

$$\frac{V_i}{R_s} = V_u(\frac{1}{R_P} + \frac{r}{r^2+\omega^2L^2} + j\omega C - \frac{j\omega L}{r^2+\omega^2L^2}) \tag{2.126}$$

La 2.126 puo' essere interpretata come l'equazione del circuito in Fig. 2.16b, dove e' stata introdotta un resistore

$$R^* = \frac{r^2+\omega^2L^2}{r} = r + \frac{\omega^2L^2}{r} = r(1+\frac{\omega^2L^2}{r^2}) \tag{2.127}$$

ed un induttore

$$L^* = L(\frac{r^2}{\omega^2L^2}+1) \tag{2.128}$$

Non risolveremo esplicitamente l'equazione 2.126; si noti tuttavia che sia la frequenza di risonanza che il fattore di merito sono alterati dalla presenza di r. La situazione diviene piu' semplice nella regione di alta frequenza; infatti se ω é abbastanza grande, $(\omega L)/r \gg 1$, e quindi

$$R^* \approx \frac{\omega^2L^2}{r} \tag{2.129}$$

$$L^* \approx L \tag{2.130}$$

Risposta del circuito in regime impulsivo

Studiamo ora la risposta del circuito RLC parallelo in regime impulsivo, cioe' per una sollecitazione a gradino. Per semplicita' trascuriamo l'effetto della resistenza parassita r. L'equazione differenziale del nodo d'uscita e':

$$\frac{v_i}{R_s} = (\frac{1}{R_P} + \frac{1}{L}\int dt + C\frac{d}{dt})v_u \tag{2.131}$$

dove abbiamo indicato con R_P il parallelo tra R ed R_s. Derivando ambo i membri

$$\frac{1}{R_s}\frac{dv_i}{dt} = \frac{1}{R_P}\frac{dv_u}{dt} + \frac{v_u}{L} + C\frac{d^2v_u}{dt^2} \tag{2.132}$$

Se $v_i = u(t)$ il primo membro si annulla per $t > 0$, quindi si ha

$$C\frac{d^2v_u}{dt} + \frac{1}{R_P}\frac{dv_u}{dt} + \frac{v_u}{L} = 0 \tag{2.133}$$

Come e' noto le soluzioni sono funzioni del tipo e^{pt}: si ha quindi

$$Cp^2e^{pt} + \frac{pe^{pt}}{R} + \frac{e^{pt}}{L} = 0 \tag{2.134}$$

$$Cp^2 + \frac{1}{R_P}p + \frac{1}{L} = 0 \tag{2.135}$$

Risolvendo

$$p = \frac{-\frac{1}{R_P} \pm \sqrt{\frac{1}{R_P^2} - \frac{4C}{L}}}{2C} \tag{2.136}$$

$$= -\frac{1}{2R_PC} \pm \sqrt{\frac{1}{(2R_PC)^2} - \frac{1}{LC}} \tag{2.137}$$

Ponendo

$$\omega_0 = \frac{1}{\sqrt{LC}} \quad k = \frac{1}{2R_P}\sqrt{\frac{L}{C}} \tag{2.138}$$

si puó scrivere

$$p = -k\omega_0 \pm j\omega_0\sqrt{1-k^2} \tag{2.139}$$

come e' facile verificare direttamente. Si hanno allora varie possibilitá (Fig. 2.17):

1) $k = 0$: le soluzioni sono quindi $e^{j\omega_0 t}$ e $e^{-j\omega_0 t}$, cioé onde sinusoidali di frequenza ω_0. Questo caso corrisponde a R= ∞, cioé un caso non realistico;

2)$k = 1$: le soluzioni sono reali e coincidenti; quindi le soluzioni dell'equazione differenziale sono $e^{-\omega_0 t}$ e $ate^{-\omega_0 t}$. Tenendo conto delle condizioni iniziali si trova che

$$v_u = 2\omega_0 te^{-\omega_0 t} \tag{2.140}$$

(smorzamento critico);

3) $0 < k < 1$: la risposta é una sinusoide smorzata di frequenza $\omega_0\sqrt{1-k^2}$

4) $k > 1$: le soluzioni sono di nuovo esponenziali reali:

$$p = -k\omega_0 \pm \omega_0\sqrt{k^2-1} = -k\omega_0 \pm k\omega_0\sqrt{1-\frac{1}{k^2}} \tag{2.141}$$

Ovviamente per $t \to \infty$, $v_u \to 0$.Tuttavia, nel caso reale, dobbiamo considerare la resistenza parassita r; ai tempi lunghi il circuito equivale ad un partitore resistivo e quindi

$$v_u(\infty) \simeq \frac{r}{R_s + r}u(t) \tag{2.142}$$

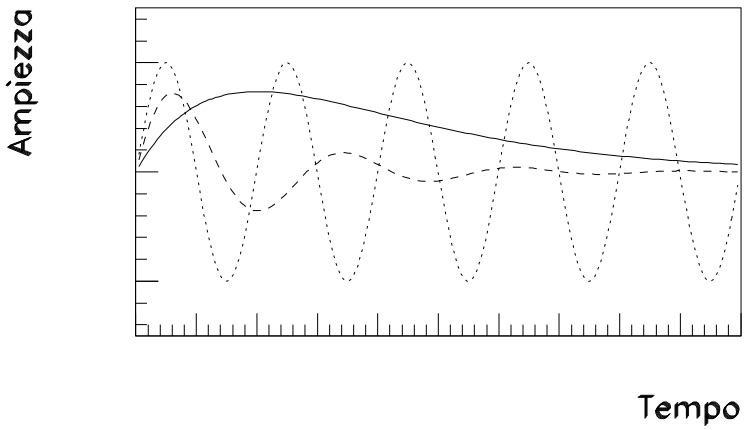

Figura 2.17: *Risposta impulsiva di un circuito RLC serie: $k = 0$ linea puntinata; $k = 0.5$ linea tratteggiata; $k = 1$ linea continua.*

poiche' possiamo ipotizzare che r sia molto minore di R.

Il comportamento in regime impulsivo del circuito RLC in serie e' assolutamente analogo: in questo caso dovremo scrivere l'equazione differenziale della maglia e discuterne le soluzioni sulla falsariga di quanto fatto per il circuito in parallelo.

2.8 Linee di trasmissione

Negli esempi visti finora abbiamo trascurato il fatto che la velocita' di propagazione dei segnali elettrici e' finita. In altre parole, abbiamo fatto l'ipotesi che le variazioni nel tempo della sollecitazione producano i loro effetti istantaneamente in tutti i punti del circuito. Questo comportamento non puo' essere esatto, visto che al massimo il segnale si propaghera' alla velocita' della luce c. Tuttavia il nostro trattamento costituisce una buona approssimazione quando le dimensioni fisiche del circuito sono piccole, tali cioe' per cui il tempo di propagazione del segnale e' piccolo rispetto alla scala dei tempi che ci interessa. Per segnali impulsivi questa scala dei tempi e' data dai tempi di salita che abbiamo, mentre per sollecitazioni sinusoidali e' data dal periodo T. Per esempio, se il nostro circuito ha dimensioni tipiche dell'ordine di $\sim 30\ cm$, i tempi di propagazione (nell'ipotesi che il segnale si propaghi a velocita' c) sono di $\sim 1\ ns$. Se i nostri impulsi hanno tempi di salita dell'ordine dei μs non ha senso preoccuparsi del tempo di propagazione. Lo stesso se il nostro generatore sinusoidale ha una frequenza di $1\ Mhz$ ($T = 1\ \mu s$). Invece, se la frequenza e' di $1\ Ghz$ ($T = 1\ ns$) non possiamo trascurare il tempo di propagazione, che e' dello stesso ordine. Inoltre e' evidente che non possiamo trascurare il tempo di propagazione quando siamo interessati a trasmettere un segnale, sia sinusoidale che impulsivo, su lunghe distanze (decine o centinaia di metri).

In questo paragrafo studieremo quindi il problema della trasmissione di un segnale, attraverso una coppia di conduttori, che prende il nome di linea di trasmissione. Questa coppia puo' essere composta da 2 fili paralleli tra loro (linea bifilare), da 2 conduttori

coassiali, ovvero da un filo e da un piano conduttore (situazione tipica, per es. di un segnale che si propaga su un circuito stampato). Se consideriamo un tratto infinitesimo di linea, esso puo' essere schematizzato, nel modo piu' generale possibile, come in Fig. 2.18. Si noti che R,L,G e C sono intese *per unita' di lunghezza*. Possiamo quindi definire

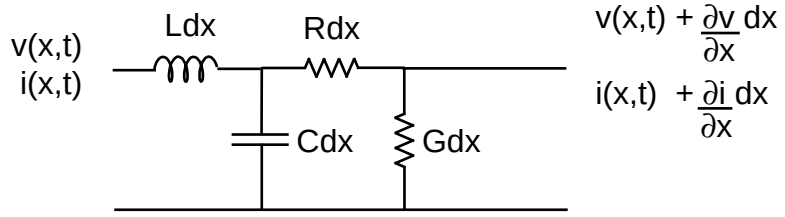

Figura 2.18: *Tratto infinitesimo di linea*

$$Z = R + j\omega L$$

$$Y = G + j\omega C$$

Applicando i principi di Kirchoff si hanno le due equazioni

$$-\frac{\partial V}{\partial x}dx = ZIdx$$

$$-\frac{\partial I}{\partial x}dx = YVdx$$

cioe'

$$\frac{\partial V}{\partial x} = -ZI$$

$$\frac{\partial I}{\partial x} = -YV$$

Derivando ancora rispetto ad x, si ottiene

$$\frac{\partial^2 V}{\partial x^2} - \gamma_0^2 V = 0$$

$$\frac{\partial^2 I}{\partial x^2} - \gamma_0^2 I = 0$$

dove

$$\gamma_0 = -\sqrt{ZY} = -\sqrt{(R + j\omega L)(G + j\omega C)}$$

che si puó scrivere anche come

$$\gamma_0 = -\tau_0\sqrt{-\omega^2 + j\omega(\frac{R}{L} + \frac{G}{C}) + \frac{RG}{LC}}$$

dove $\tau_0 = \sqrt{LC}$ ha le dimensioni di un tempo per unitá di lunghezza e prende il nome di ritardo per unita' di lunghezza. Il rapporto

$$Z_0 = \sqrt{\frac{Z}{Y}} = \sqrt{\frac{R + j\omega L}{G + j\omega C}}$$

prende il nome di impedenza caratteristica.

Se si verifica la condizione

$$\frac{R}{L} = \frac{G}{C}$$

allora

$$Z_0 = \sqrt{\frac{L}{C}} = R_0$$

cioé l'impedenza caratteristica é reale ed ha le dimensioni di una resistenza. Si ha allora

$$\gamma_0 = -\tau_0\sqrt{(\frac{R}{L} + j\omega)^2} = -\tau_0(\frac{R}{L} + j\omega)$$

cioé la linea si dice non distorcente.

Vediamo ora un caso ancora piú semplice, in cui R e G sono trascurabili (linea non dissipativa). Si ha allora

$$\frac{\partial v}{\partial x} = -L\frac{\partial i}{\partial t}$$

$$\frac{\partial i}{\partial x} = -C\frac{\partial v}{\partial t}$$

Derivando la prima equazione rispetto ad x, e la seconda rispetto a t

$$\frac{\partial^2 v}{\partial x^2} = -L\frac{\partial^2 i}{\partial t \partial x}$$

$$\frac{\partial^2 i}{\partial t \partial x} = -C\frac{\partial^2 v}{\partial t^2}$$

cioé

$$\frac{\partial^2 v}{\partial x^2} - LC\frac{\partial^2 v}{\partial t^2} = 0$$

ovvero un'equazione delle onde con velocitá di propagazione

$$u = \frac{1}{\sqrt{LC}} = \frac{1}{\tau_0}$$

Le soluzioni sono quindi del tipo

$$v(x,t) = f_1(x - ut) + f_2(x + ut)$$

$$i(x,t) = \frac{1}{R_0}[f_1(x - ut) - f_2(x + ut)]$$

Supponiamo ora di avere una linea di lunghezza infinita alimentata da un generatore di segnale con resistenza di uscita R_g (Fig. 2.19a). Si ha

$$v(x,t) = \frac{R_0}{R_0 + R_g} v_g(t - \tau_0 x)$$

$$i(x,t) = \frac{1}{R_0 + R_g} v_g(t - \tau_0 x)$$

cioé, dal punto di vista dell'ingresso, la linea si comporta come un carico resistivo R_0.

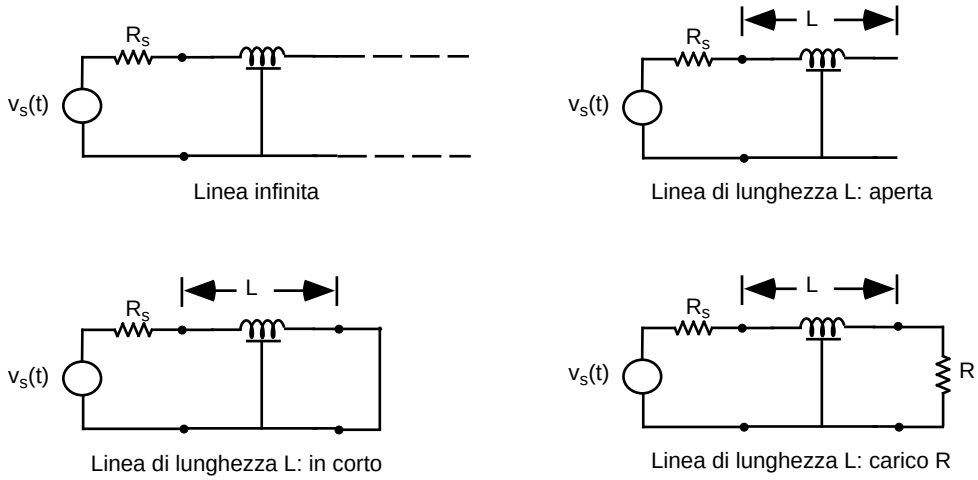

Figura 2.19: *a) Linea di lunghezza infinita; b) Linea di lunghezza finita,L; c) Linea chiusa in corto circuito; d) Linea chiusa su un carico R*

Se ora consideriamo una linea finita e terminata con una resistenza R_0, le equazioni della linea restano le stesse, ed anche la soluzione particolare é la stessa, cioé non ci sono onde retrograde. Dal punto di vista del segnale, quindi, una linea chiusa su una resistenza R_0 e' equivalente ad una linea infinita (linea *adattata*)

Supponiamo ora che la linea sia aperta e di lunghezza finita: $\tau = \tau_0 L$ e' il tempo di propagazione del segnale lungo la linea. Nel punto L la corrente deve essere 0, quindi ci deve essere un'onda retrograda, cioé

$$i(x,t) = \frac{1}{R_0 + R_g}[v_g(t - \tau_0 x) - v_g(t - \tau_0(2L - x))]$$

Questo é vero solo per $t \ll 2\tau_0$, per tempi successivi l'onda retrograda, arrivata all'origine,viene nuovamente riflessa e si somma alle altre. Ció non é vero se $R_g = R_0$; in tal caso la linea é *adattata* sul lato del generatore e non si hanno ulteriori riflessioni.

Se la linea e' invece chiusa in corto circuito (Fig. 2.19c) Si deve imporre la condizione

$$v(L,t) = 0$$

quindi si deve avere un'onda riflessa di tensione uguale cambiata di segno.

$$v(x,t) = \frac{R_0}{R_0 + R_g}[v_g(t - \tau x) - v_g(t - \tau(2L - x))]$$

$$i(x,t) = \frac{1}{R_0 + R_g}[v_g(t - \tau x) + v_g(t - \tau(2L - x))]$$

In generale, considerando la linea chiusa su una resistenza generica R (Fig. 2.19d) si avrá nel punto L una riflessione parziale; potremo definire un coefficiente di riflessione Γ_2

$$\Gamma_2 = \frac{R - R_0}{R + R_0}$$

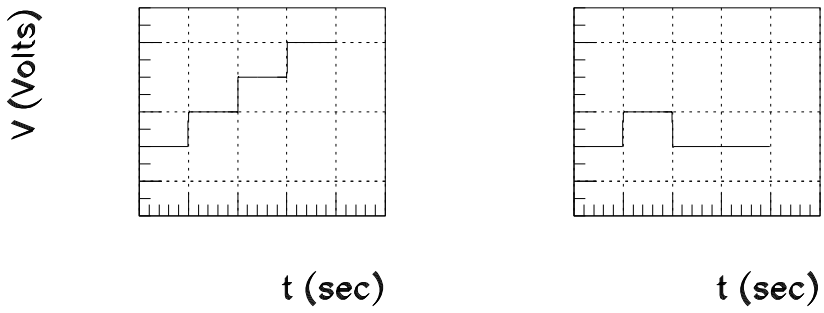

Figura 2.20: *Risposte della linea per sollecitazioni a gradino:*
a) Linea aperta;
b) Linea chiusa in corto.

Nei vari casi avremo infatti

$$R = R_0 \quad \Gamma_2 = 0$$
$$R = 0 \quad \Gamma_2 = -1$$
$$R = \infty \quad \Gamma_2 = 1$$

Si avrá quindi

$$v(x,t) = \frac{R_0}{R_0 + R_g}[v_g(t - \tau x) + \Gamma_2 v_g(t - \tau(2L - x))]$$
$$i(x,t) = \frac{1}{R_0 + R_g}[v_g(t - \tau x) - \Gamma_2 v_g(t - \tau(2L - x))]$$

Se $R_g \neq R_0$ si puó analogamente definire un coefficiente di riflessione all'inizio della linea

$$\Gamma_1 = \frac{R_g - R_0}{R_g + R_0}$$

e le soluzioni conterranno ulteriori termini dovuti a questa riflessione. Nella Fig. 2.20 sono riportati alcuni esempi di risposte della linea soggetta ad una sollecitazione a gradino, $u(t)$, per vari valori di Γ_1 e Γ_2.

2.8.1 Linee reali

Abbiamo visto che i parametri fisicamente rilevanti di una linea non dissipativa sono la velocita' di propagazione,v, e la resistenza caratteristica, R_0; essi sono dati da:

$$v = \frac{1}{\sqrt{LC}} \quad R_0 = \sqrt{\frac{L}{C}}$$

L'induttanza per unita' di lunghezza,L, e la capacita' per unita' di lunghezza, C, dipendono dalla geometria della linea e dalle caratteristiche elettro-magnetiche del mezzo interposto

tra i conduttori. Nel caso di linea coassiale (Fig. 2.21a) sono date da:

$$C = \frac{2\pi\epsilon}{\log D/d} \quad L = \frac{\mu}{2\pi}\log D/d$$

$$v = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{\mu\epsilon}}$$

In genere la permeabilita' magnetica del mezzo e' uguale a quella del vuoto, μ_0, per cui, ricordando la definizione di velocita' della luce nel vuoto, c:

$$v \approx \frac{c}{\sqrt{\epsilon_r}}$$

dove ϵ_r e' la costante dielettrica relativa del mezzo ($\sim$ 2 per i dielettrici normalmente utilizzati nella costruzione dei cavi).

Si ha poi

$$R_0 \equiv \sqrt{\frac{L}{C}} = (\frac{\mu}{4\pi^2\epsilon}\log^2 D/d)^{1/2} = \frac{1}{2\pi}(\log\frac{D}{d})\sqrt{\frac{\mu}{\epsilon}}$$

Si noti che R_0 dipende pochissimo dalla geometria del cavo (solo attraverso il logaritmo del rapporto delle dimensioni geometriche).

Valori tipici per un cavo coassiale utilizzato comunemente in ambiente elettronico (cavo RG58) sono:

$$C \simeq 100pF/m \quad L \simeq 250nH/m$$

$$v \simeq 0.7c \quad R_0 \simeq 50\Omega$$

Il cavo coassiale utilizzato per le antenne televisive ha invece una resistenza caratteristica di 75 Ω, valore standard in questo ambiente.
Invece, nel caso di linea bifilare (Fig. 2.21b):

$$C = \frac{\pi\epsilon}{\log(2D/d)} \quad L = \frac{\mu}{\pi}\log(2D/d)$$

che danno luogo ad una espressione della velocita' identica a quella del cavo coassiale e ad una espressione della resistenza caratteristica assai simile.

Ulteriori possibili geometrie sono rappresentate nella Fig. 2.21. Tutte le situazioni raffigurate sono caratterizzate da una dipendenza logaritmica dalle costanti geometriche e da una velocita' di propagazione dipendente solo da ϵ_r. Quindi tutte le geometrie portano a resistenze caratteristiche comprese tra le decine e le centinaia di Ohm.

Finora abbiamo fatto l'ipotesi, poco realistica, di trascurare gli effetti ohmici della linea, cioe'R e G. Nel caso di cavi reali e' in genere lecito trascurare G, ma non R, in genere piccolo ma non trascurabile completamente, se stiamo utilizzando una linea di considerevole lunghezza. La condizione di *non-distorsione* non e' piu' esattemente rispettata; inoltre il segnale verra' attenuato in modo esponenziale. Supponendo di avere una linea infinita (ovvero adattata all'estremita' lontana), l'ampiezza del segnale lungo la linea sara' data da:

$$v(x,t) = \frac{R_0}{R_0 + R_g} v_g(t - \tau x)e^{-\alpha x}$$

dove $\alpha = R/R_0$ é detto coefficiente di attenuazione della linea.

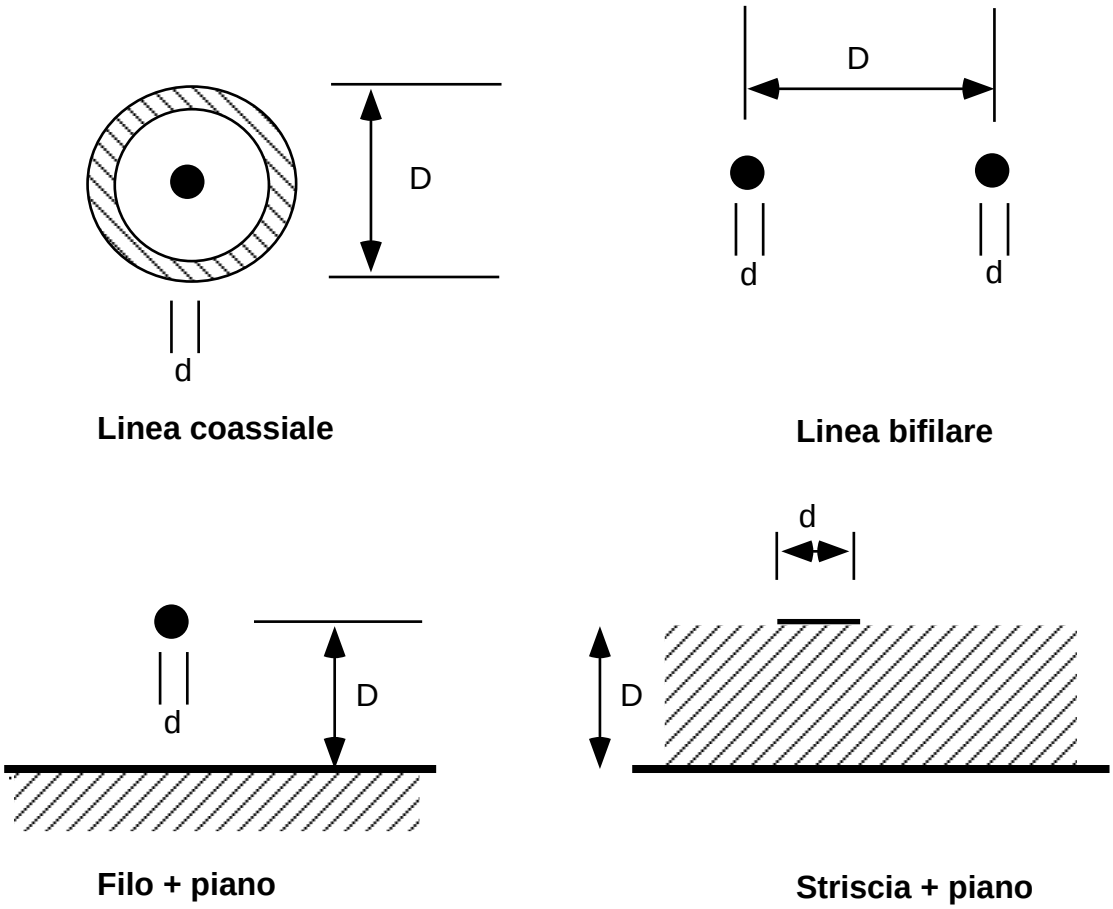

Figura 2.21: *a) Linea coassiale; b) Linea bifilare; c) Filo + piano d) Strisce*

Per i cavi commerciali R e' dell'ordine di $10^{-2}\ \Omega/m$.

Dobbiamo infine tenere presente che una linea va anche studiata per quelli che sono i suoi effetti *complessivi* sul trasporto del segnale. Consideriamo infatti una linea di lunghezza L (supposta adattata dal lato del generatore): vista dall'estremo lontano essa e' comunque equivalente ad un passa-basso con parametri complessivi $R_t = RL$ e $C_t = CL$. Essa introduce quindi una frequenza di taglio $f_2 = 1/(2\pi R_t C_t)$, ovvero, per segnali impulsivi, un tempo di salita $t_s = 2.2 R_t C_t$. E' chiaro quindi che la linea ha una banda passante finita, dipendente quadraticamente dalla lunghezza.
Inoltre possono essere non trascurabili gli effetti della induttanza complessiva della linea.

2.8.2 Ulteriori applicazioni

Una linea puo' essere utilizzata anche per modificare la forma di un segnale. Consideriamo ad esempio una linea di lunghezza L chiusa in corto, adattata dal lato del generatore.

Se il segnale fornito dal generatore e' rettangolare, di lunghezza T, con $T > 2\tau L$, il segnale prelevato all'inizio della linea sara' composto da due segnali rettangolari, uno positivo ed uno negativo, di lunghezza $2\tau L$ (Fig. 2.22). Tralasciando il segnale negativo (che puo' essere opportunamente eliminato, come vedremo in seguito), il segnale d'uscita

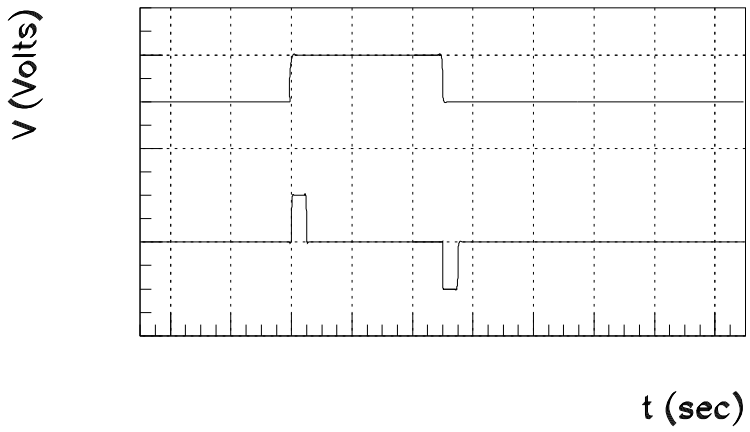

Figura 2.22: *Forma d'onda per una linea chiusa in corto*

puo' quindi essere formato a piacimento, variando L. Peraltro, per L molto corto, l'uscita e' approssimativamente proporzionale alla derivata del segnale d'ingresso, e questo puo' essere verificato anche per forme d'onda d'ingresso non rettangolari.

Quanto studiato finora ci fa anche comprendere che sono necessarie delle cautele quando si connettono tra loro cavi con impedenza diversa, ovvero quando si vuole suddividere un segnale su varie linee (Fig. 2.23).
a) Connessione di due cavi con impedenza diversa, con $Z_1 < Z_2$: per evitare che il segnale veda nel punto di connessione un aumento di impedenza, occorre aggiungere una resistenza verso massa, in modo che il parallelo di R e Z_2 eguagli Z_1, cioe':

$$R = \frac{Z_1 Z_2}{Z_2 - Z_1}$$

b) $Z_1 > Z_2$: il segnale vedrebbe una diminuzione di impedenza, quindi cio' puo' essere evitato mettendo in serie un resistore $R = Z_1 - Z_2$.
Le stesse considerazioni si applicano per un carico R_L che debba essere adattato a una linea Z. Se $R_L > Z$ occorre aggiungere un resistore in parallelo; mentre Se $R_L < Z$ occorre aggiungere un resistore in serie.
c): se si vuole suddividere un segnale su due o piu' linee, si ha inevitabilmente un disadattamento, che deve essere rimosso aggiungendo in serie dei resistori R. Il valore di R si trova imponendo che

$$\frac{R + Z}{2} = Z$$

cioé R=Z.

Se lo splitting é in n rami:

$$\frac{R + Z}{n} = Z \Rightarrow R = (n-1)Z$$

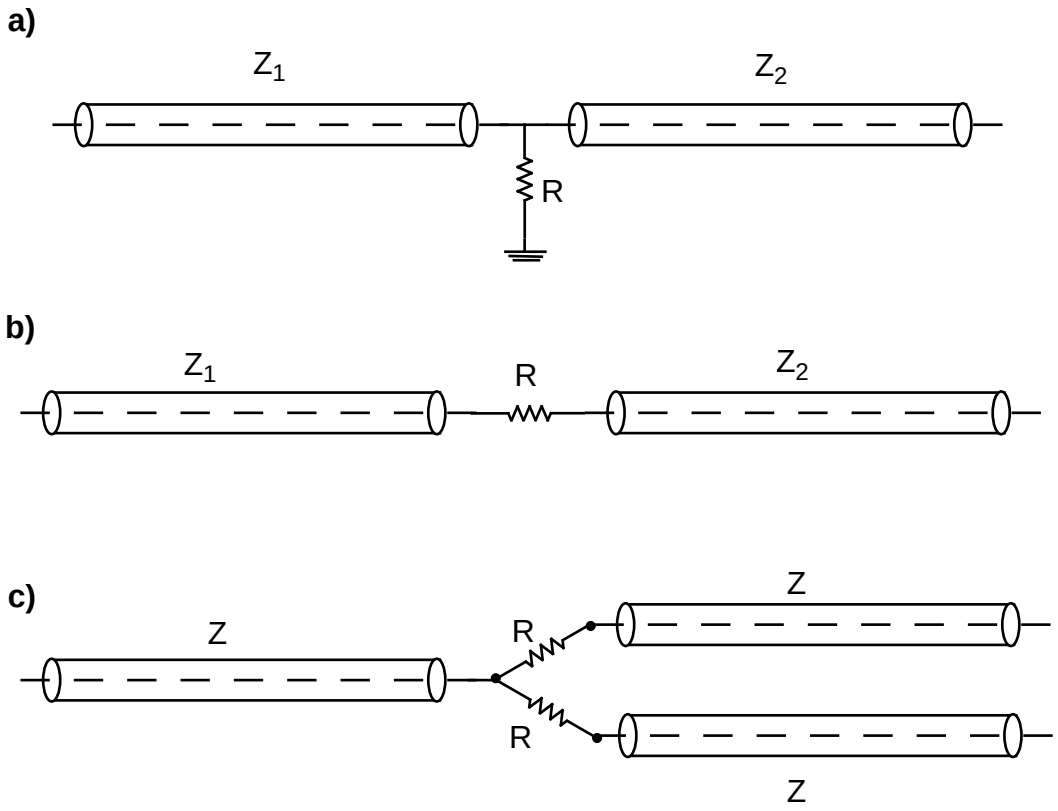

Figura 2.23: *a) Connessione di cavi a diversa impedenza ($Z_1 < Z_2$); b) Connessione di cavi a diversa impedenza ($Z_1 > Z_2$); c) Splitting di un segnale*

2.9 Cenni sul trasformatore

Il trasformatore e' un dispositivo comunemente noto per le sue applicazioni in elettrotecnica. Vogliamo qui studiarne brevemente il comportamento per sollecitazioni impulsive.

Il trasformatore puo' essere schematizzato (Fig. 2.24b) come due induttori accoppiati tra loro da un coefficiente di mutua induzione M.
Le equazioni del circuito sono:

$$\begin{cases} v_i &= L_p \dfrac{di_p}{dt} - M\dfrac{di_s}{dt} \\ 0 &= -M\dfrac{di_p}{dt} + L_s\dfrac{di_s}{dt} + i_s R_L \end{cases}$$

Si definisce

$$k \equiv \frac{M}{\sqrt{L_p L_s}}$$

Per un trasformatore ideale $k = 1$ e $L_p \to 0$ (in realtá $k < 1$ per qualche percento) e si ha

$$\frac{v_u}{v_i} = \frac{i_p}{i_s} = \sqrt{\frac{L_s}{L_p}} = \frac{N_s}{N_p} = n$$

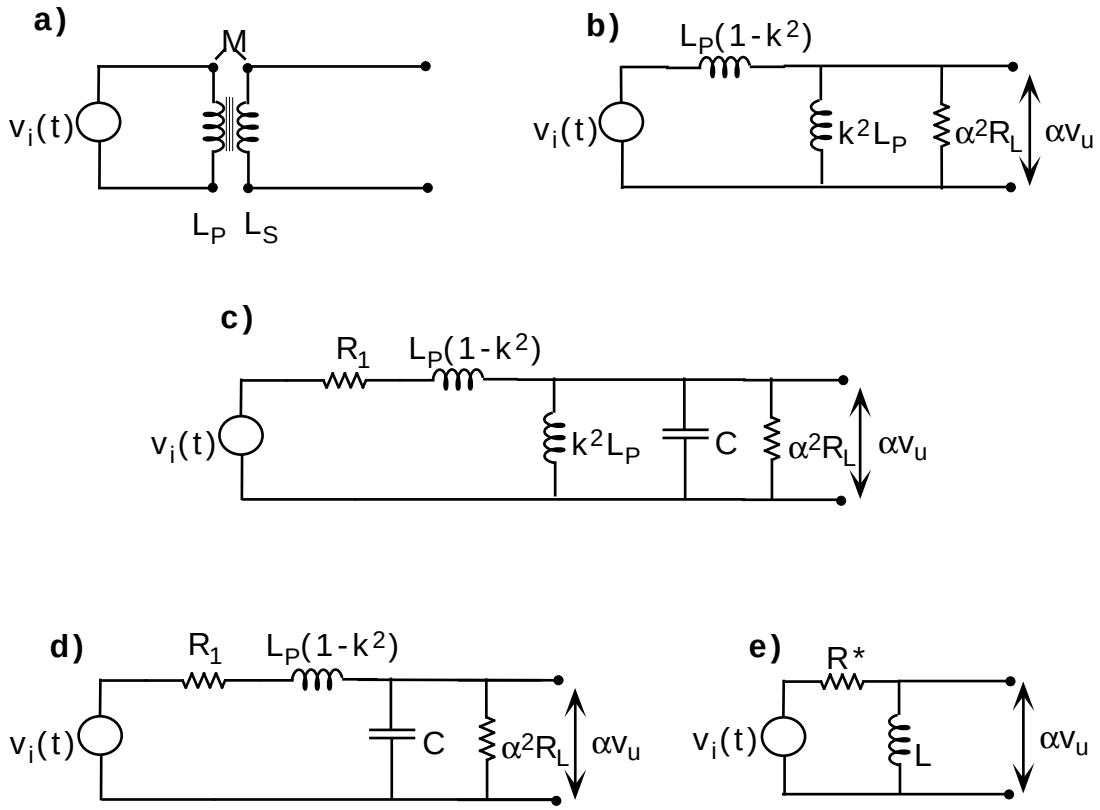

Figura 2.24: *a) Trasformatore; b) Circuito equivalente; c) Circuito equivalente completo; d) Circuito semplificato ai tempi brevi; e) Circuito semplificato ai tempi lunghi*

E' conveniente schematizzare il trasformatore come in Fig. 2.24b dove il parametro α e' dato da

$$\alpha = k\sqrt{\frac{L_p}{L_s}} = \frac{k}{n}$$

Il circuito e'governato dalle equazioni

$$\begin{cases} v_i &= L_p\dfrac{di_p}{dt} - \dfrac{k^2}{\alpha}L_p\dfrac{di_s}{dt} \\ 0 &= -k^2L_p\dfrac{di_p}{dt} + \dfrac{k^2}{\alpha}L_p\dfrac{di_s}{dt} + \dfrac{\alpha^2R_Li_s}{\alpha} \end{cases}$$

Dividendo l'ultima equazione per α si verifica che effettivamente i due circuiti sono equivalenti. In realtá il circuito equivalente completo é dato dalla Fig. 2.24c dove $R_1 = R_g + R_p$ rappresenta la resistenza del generatore piu' la resistenza del primario, mentre

$$\frac{1}{R_2} = \frac{1}{R'} + \frac{1}{\alpha^2(R_L + R_2')}$$

dove R_2' é la resistenza del secondario e R' é la resistenza in parallelo che viene dalle correnti di Foucault,ecc; C é il contributo complessivo delle capacitá.

Ai tempi brevi il circuito si semplifica come in Fig. 2.24d. Per una sollecitazione a gradino $v_i = u(t)$, si risolve l'equazione differenziale del circuito: le soluzioni sono del tipo

e^{pt}, con

$$p = -\omega_0 \xi \pm j\omega_0 \sqrt{1-\xi^2}$$

dove

$$\omega_0 = \frac{1}{\sqrt{\sigma C a}} \quad a = \frac{R_2}{R_1 + R_2}$$

$$\xi = (\frac{R_1}{\sigma} + \frac{1}{R_2 C})\frac{1}{2\omega_0}$$

Si ha quindi una situazione analoga a quella del circuito RLC, e la forma d'onda d'uscita e' legata al valore del parametro ξ.

Ai tempi lunghi il circuito si semplifica come in Fig. 2.24e; applicando il teorema di Thevenin,esso si semplifica ulteriormente come in Fig. 2.24f. La risposta é quindi un esponenziale reale del tipo $av_i e^{-t/\tau}$ dove $\tau = L/R$.

2.10 Sorgenti di segnale

Come abbiamo detto nell'Introduzione il nostro obiettivo e' lo studio dei segnali e dei circuiti destinati all'elaborazione dei segnali, in senso lato. Ma il lettore potrebbe a questo punto chiedersi come sono fatte in realta' le *sorgenti* dei segnali che trattiamo. I teoremi di Thevenin e Norton ci hanno fatto capire che, indipendentemente dalla sua natura fisica, una sorgente di segnale e' sempre vista dal circuito come un generatore ideale di tensione con una impedenza in serie (o di corrente con una impedenza in parallelo). Quindi non e' necessario per i nostri studi conoscere nulla di piu'; tuttavia puo' essere interessante ed istruttivo dare qualche cenno sulle sorgenti.

Ne vedremo qui di tre tipi: le antenne, i trasduttori e le sorgenti *a ionizzazione*.

Antenne

Si puo' definire con il termine *antenna* qualunque dispositivo che permette la trasmissione dell'energia elettromagnetica da un generatore allo spazio libero. Analogamente e' anche un'antenna ogni dispositivo che convoglia l'energia associata ad un'onda elettromagnetica dallo spazio libero ad un apparato limitato nello spazio (in genere un circuito).
L'antenna e' generalmente connessa al trasmettitore, o al ricevitore, attraverso una linea di trasmissione: possiamo quindi dire che le antenne sono in sostanza degli *adattatori* di impedenza tra la linea di trasmissione e lo spazio. Infatti l'adattamento di impedenza e' la condizione che si richiede per ottenere il massimo trasferimento di energia.
Lo studio delle antenne esula completamente dagli scopi di questo testo, quindi ci limiteremo solo a delle semplici considerazioni generali. Una linea di trasmissione (ad esempio bifilare), alimentata ad un estremo da un generatore e adattata all'altro estremo sulla propria impedenza caratteristica, e' essa stessa un'antenna trasmittente: infatti la propagazione di un'onda di tensione e di corrente lungo la linea e' associata alla propagazione di un campo elettromagnetico irradiato nello spazio circostante. Dal punto di vista circuitale la perdita di energia per irraggiamento e' del tutto equivalente ad una perdita ohmica lungo la linea stessa. Possiamo quindi in generale trattare l'antenna come una linea dissipativa; idealmente sarebbe desiderabile che tutte le perdite fossero associate all'irraggiamento, mentre in realta' ogni antenna avra' anche una dissipazione ohmica vera e propria. E' chiaro

quindi che anche per un'antenna trasmittente possiamo, in stretta analogia a quanto fatto per le linee, definire un'impedenza d'ingresso, la cui parte reale contiene un contributo R_r, *resistenza di radiazione*, che tiene conto proprio delle perdite di energia per irraggiamento. La stessa linea di trasmissione costituisce anche un'antenna ricevente: il campo elettrico

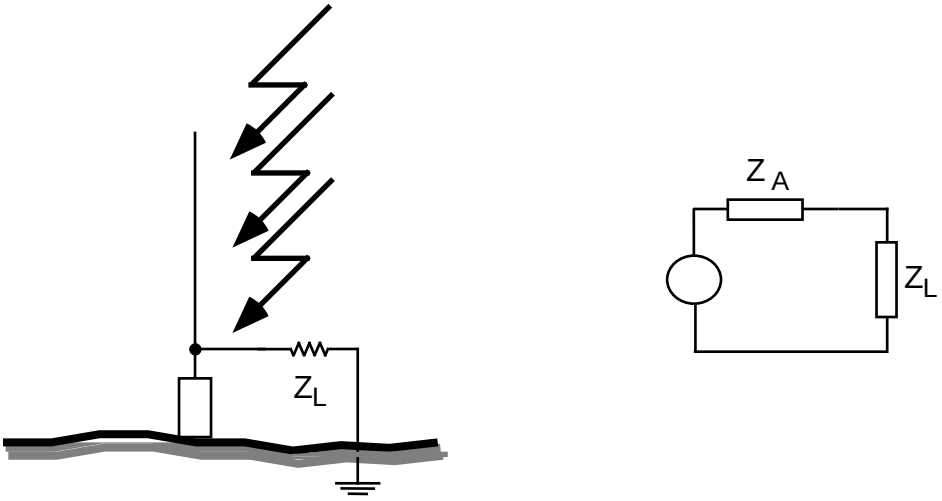

Figura 2.25: *Antenna ricevente e suo circuito equivalente*

associato all'onda che investe il conduttore induce una corrente, che puo' essere trasferita ad un ricevitore. Da un punto di vista circuitale si ha quindi la situazione presentata in Fig. 2.25, dove il generatore equivalente V_E e' chiaramente legato al campo elettrico E dell'onda[2]. Si puo' scrivere

$$V_E = Eh \tag{2.143}$$

Il parametro h, avente dimensioni di una lunghezza, e' detto *altezza equivalente dell'antenna*. Peraltro il principio di reciprocita'[3] ci consente di dire che l'impedenza caratteristica dell'antenna e' la stessa in trasmissione ed in ricezione. L'impedenza equivalente dell'antenna, Z_A e' naturalmente legata alla geometria dell'antenna stessa e conterra' una componente reattiva; la componente resistiva sara' data da

$$R_A = R_r + R_d \tag{2.144}$$

dove R_d e' legato alle perdite ohmiche dall'antenna e R_r e' la resistenza di radiazione dell'antenna. E' chiaro infatti che anche un'antenna ricevente irradia (proprio per il sopra citato principio di reciprocita'), quindi parte dell'energia captata viene reirradiata all'esterno. La potenza complessiva captata dall'antenna e' data da

$$W = \frac{(Eh)^2}{R_L + R_r + R_d} \tag{2.145}$$

[2]Si noti che l'antenna, in quanto linea, e' costituita dal conduttore (verticale nel caso in figura) e dal piano di terra.

[3]Questo principio, valido in generale per le reti lineari, si puo' enunciare come segue: se un generatore ideale di tensione V, posto tra i nodi A e B di un circuito, provoca tra i nodi A' e B' (tra loro corto-circuitati) una corrente I, allora ponendo lo stesso generatore tra i nodi A' e B', si avra' tra i nodi A e B (corto-circuitati tra loro) la stessa corrente I.

La frazione $R_L/(R_L + R_r + R_d)$ viene assorbita dal carico; la frazione $R_d/(R_L + R_r + R_d)$ e' dissipata, mentre la parte rimanente e' reirradiata. La massima potenza che un'antenna puo' estrarre dall'onda che la investe si ha quando $(R_d + R_L)$ eguagliano la resistenza di radiazione, e la parte reattiva del carico sia uguale ed opposta alla parte reattiva di Z_A.

Trasduttori

Un trasduttore e' un dispositivo che genera una tensione (o una corrente) legata da una relazione funzionale con una variabile fisica: in generale quindi abbiamo

$$v(t) = F[y(t)] \tag{2.146}$$

dove $y(t)$ e' la grandezza fisica in oggetto. E' evidente l'interesse di dispositivi di questo generale, ed e' chiaro che in generale si desidera una relazione *lineare* tra v e y.

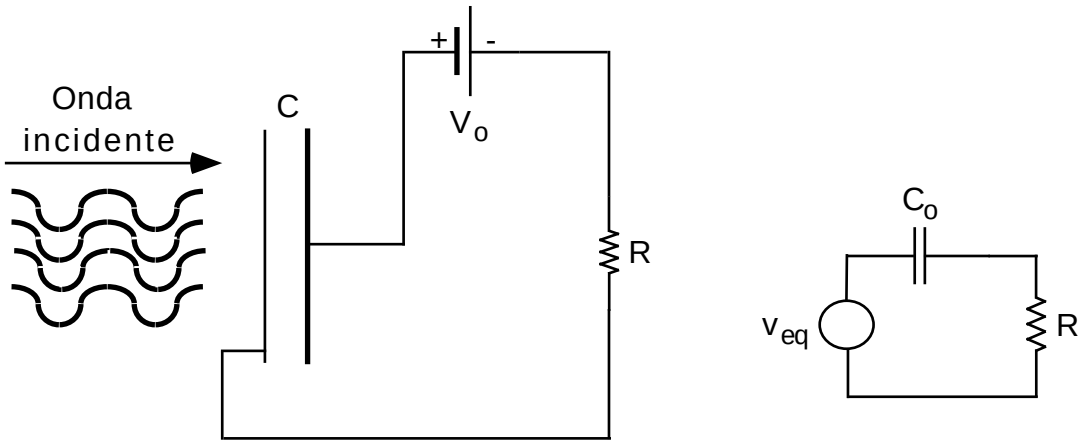

Figura 2.26: *a) Schema di principio di un microfono capacitivo; b) generatore di segnale equivalente*

Discuteremo qui brevemente un solo esempio, quello del microfono a capacita' (Fig. 2.26), cioe' un dispositivo che traduce l'onda di pressione associata al suono in un segnale elettrico. Il condensatore C ha un'armatura fissa, mentre l'altra e' libera di vibrare quando e' investita dal suono. Se ω e' la pulsazione dell'onda, la capacita' varia con una legge

$$C = C_0 + C_1 \sin \omega t \tag{2.147}$$

L'equazione della maglia e'

$$\frac{Q}{C(t)} - V_0 + iR = 0 \tag{2.148}$$

che possiamo quindi scrivere

$$\int i\, dt' - CV_0 + RC(t)i = 0 \tag{2.149}$$

Derivando rispetto a t e riordinando opportunamente i termini si arriva a

$$R(C_0 + C_1 \sin \omega t)\frac{di}{dt} + (1 + RC_1\omega \cos \omega t)i = V_0\omega C_1 \cos \omega t \tag{2.150}$$

Ora, se $C_1 << C_0$ e $\omega R C_1 << 1$, l'equazione si semplifica in

$$RC_0 \frac{di}{dt} + i = V_0 \omega C_1 \cos \omega t \tag{2.151}$$

la cui soluzione e' data da

$$i = \frac{C_1}{C0} \frac{V_0 \cos(\omega t - \phi)}{\sqrt{R^2 + \frac{1}{\omega^2 C_0^2}}} \tag{2.152}$$

dove

$$\tan \phi = \frac{1}{\omega R C_0} \tag{2.153}$$

E' interessante notare che, se le condizioni anzidette non sono soddisfatte, la soluzione e' piu' complessa, ma soprattutto contiene termini oscillanti in ω, 2ω, 4ω, ... Naturalmente un buon microfono deve soddisfare quelle condizioni; si vede comunque dalla 2.152 che il trasduttore e' equivalente (Fig. 2.26b) ad un generatore di tensione

$$v_e q = \frac{C_1}{C_0} V_0 \cos(\omega t - \phi) \tag{2.154}$$

con una impedenza in serie data dalla capacita' C_0 (non includiamo R, che rappresenta in sostanza il carico su cui il segnale viene riversato).
Come si vede questo e' un esempio di generatore con una impedenza d'uscita tipicamente reattiva, anche se e' sicuramente presente una parte resistiva che qui abbiamo trascurato (si pensi, per esempio alla resistenza del generatore costante V_0).

Sorgenti a ionizzazione

L'ultimo caso che esamineremo e' quello di sorgenti in cui il segnale e' generato attraverso meccanismi di ionizzazione. Esistono svariati esempi di dispositivi di questo tipo, molti dei quali utilizzati proprio per la *rivelazione* di particelle ionizzanti. In linea di principio il meccanismo di funzionamento e' quello illustrato nella Fig. 2.27. Due elettrodi racchiudono una porzione di spazio riempita per es. da un gas. Il passaggio di una particella carica attraverso questo spazio provoca la ionizzazione di un certo numero di molecole e la formazione di coppie elettrone-ione. Sotto l'effetto del campo elettrico applicato ai due elettrodi gli ioni migrano verso il polo negativo e gli elettroni verso quello positivo: si ha quindi un *segnale* che viene rivelato all'uscita, sovrapposto alla tensione costante V.

E' chiaro quindi che, anche in questo caso, la sorgente di segnale ha un'impedenza d'uscita con una parte reattiva importante, data proprio dalla capacita' degli elettrodi di raccolta. Naturalmente questa e' una sorgente di tipo impulsivo (si avra' un *fiotto* di carica collegato temporalmente al passaggio della particella); la forma dell'impulso, oltre ad essere legata alla dinamica del fenomeno in se', e' anche determinata dai valori di R e C, che nel loro insieme formano un'integratore.

2.11 Circuiti reali

Nei precedenti paragrafi abbiamo approfonditamente studiato il comportamento di circuiti con reattanze. Vedremo nel seguito che, in particolare, i circuiti RC sono di estrema importanza e molto frequentemente usati. Vale anche la pena di sottolineare che, spesso, il

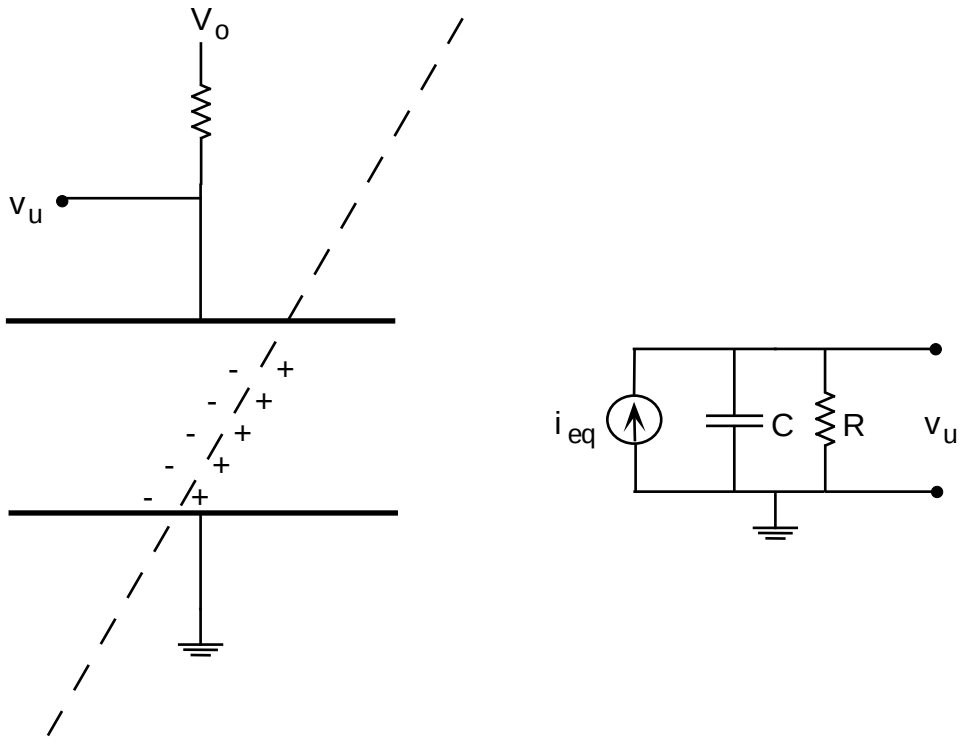

Figura 2.27: *a) Schema di principio di un rivelatore a ionizzazione; b) generatore di segnale equivalente*

comportamento di circuiti reali e' influenzato anche da reattanze parassite. In particolare, capacita' parassite sono inevitabilmente presenti in ogni circuito, generando accoppiamenti RC di tipo passa basso; nascono quindi frequenze di taglio superiori, che limitano la banda passante. Circuiti destinati ad essere utilizzati per altissime frequenze devono quindi essere progettati con particolare cura (con inevitabile aumento del costo); inoltre queste considerazioni si applicano anche ai regimi impulsivi; si comprende come sia concettualmente impossibile avere impulsi con tempo di salita zero, mentre tempi di salita molto corti implicano notevoli aumenti nel costo. Naturalmente ci aspettiamo che siano presenti anche induttanze parassite, e quindi in linea di principio ogni circuito puo' dare luogo a frequenze di risonanza indesiderate. Queste frequenze in genere sono molto alte (poiche' sono legate al prodotto LC), ma il loro effetto puo' a volte essere osservato, per es. in regime impulsivo, per l'insorgere di oscillazioni.

Infine va ricordato che gli stessi componenti elementari presentano caratteristiche - parassite. Abbiamo gia' visto il caso degli induttori; ma anche resistori e condensatori presentano, in genere in misura molto minore, lo stesso problema. In linea di principio ogni componente costituisce un piccolo circuito RLC : tuttavia le caratteristiche indesiderate divengono influenti solo a frequenze altissime, lontane in genere dalla regione che ci interessa. Per questo motivo trascureremo, nel seguito, questo aspetto.

Capitolo 3

Dispositivi a semiconduttore

3.1 Introduzione

I dispositivi a semiconduttore costituiscono i mattoni fondamentali dell'elettronica odierna. I meccanismi fisici che sono alla base dei semiconduttori sono abbastanza complessi e vengono approfonditamente studiati nei corsi di Struttura della Materia. Qui ci limiteremo ad una breve e non rigorosa sintesi che utilizzeremo come punto di partenza per lo studio dei dispositivi a semiconduttore, cioe' diodi e transistors.
Il transistor è un componente essenziale di ogni circuito elettronico, dal più semplice amplificatore al più complesso computer. I circuiti integrati, che oggi hanno largamente sostituito i circuiti a componenti discreti, sono essi stessi costituiti principalmente da reti di transistors.

Una buona comprensione del funzionamento di questo componente è perciò importante anche comprendere appieno il funzionamento e le proprieta' di ingresso e uscita dei circuiti integrati. E, comunque, ci sono ancora situazioni in cui il transistor e' insostituibile.
In questo capitolo studieremo i componenti a giunzione, mentre i transistors ad effetto di campo (FET) formeranno oggetto di un altro capitolo.

3.2 Cenni sulla fisica dei semiconduttori

3.2.1 Struttura elettronica degli elementi

Ricordiamo che gli atomi sono composti da un nucleo di protoni e neutroni e da una nuvola di elettroni orbitanti attorno al nucleo. Gli elettroni atomici sono caratterizzati da 4 numeri quantici n, l, m_l, ed s. Essi possono assumere i seguenti valori:

$$\begin{array}{ll} n & 1, 2, 3 \ldots \\ l & 0, 1, 2, \ldots (n-1) \\ m_l & -l, -(l+1), \ldots, (l-1), l \\ s & \pm\frac{1}{2} \end{array}$$

Il principio di esclusione di Pauli asserisce che in un atomo non ci possono essere due elettroni con gli stessi numeri quantici, quindi ogni elettrone e' caratterizzato univocamente

dal set di numeri quantici che esso possiede. In un atomo con piu' elettroni questi occupano i numeri quantici partendo da quelli cui compete la minore energia; si definiscono quindi delle shells, K,L,M,N ..., corrispondenti al valore del numero quantico principale n, e delle sub-shells $s, p, d, f, \ldots$, corrispondenti al numero quantico orbitale l.

shell	K	L	M	N
n	1	2	3	4
subshells	s	sp	spd	spdf
tot.el.	2	$2+6$	$2+6+10$	$2+6+10+14$

Tabella 3.1: *Orbite elettroniche nell'atomo.*

A questo punto, per ogni numero atomico Z, è molto semplice prevedere la configurazione elettronica. Prendiamo ad esempio $Z = 10$ (Neon): si ha $1s^2 2s^2 2p^6$, dove con questa scrittura sintetica si indica che vi sono 2 elettroni nella shell K ($n = 1$) ed 8 elettroni nella shell L ($n = 2$), suddivisi nelle sub-shells s e p. A noi interessano in particolare il Germanio (Z=32, $\ldots 4s^2 4p^2$) ed il Silicio (Z=14, $\ldots 3s^2 3p^2$).

3.2.2 Bande di energia in un solido

Nei solidi, cioe' aggregati di atomi organizzati in modo ordinato (cristalli) i livelli energetici piu' esterni dei singoli atomi costituenti interagiscono tra loro e si ha la formazione di bande di livelli energetici, eventualmente separate da bande di energie proibite. Le proprieta' di conduzione elettrica del materiale sono legate alla struttura di queste bande (vedi Fig. 3.1).

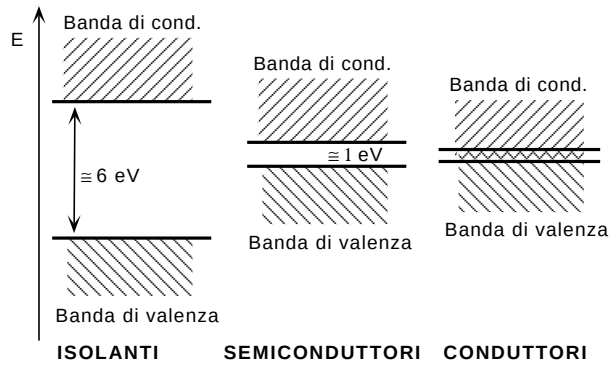

Figura 3.1: *Bande di energia*

I livelli energetici occupati dagli elettroni di valenza degli atomi formano la cosidetta banda di valenza: se questa banda e' completamente piena e, al di sopra di essa vi e' una gap di energia proibita, la conduzione elettrica e' impedita, in quanto gli elettroni non possono essere accelerati dalla applicazione di un campo elettrico esterno (gli elettroni dovrebbero aumentare la loro energia, ed occupare livelli superiori). In altre parole la conduzione e' legata alla posizione dei livelli energetici liberi piu' vicini a quelli della banda di valenza. Se questi livelli sono distanti qualche eV la conduzione e' impossibile, a meno di applicare

al cristallo campi elettrici giganteschi. Questi materiali sono detti isolanti. Se invece vi sono livelli liberi adiacenti alla banda di valenza, si puo' avere conduzione elettrica (metalli). Esistono alcune situazioni intermedie, in cui la gap proibita e' dell'ordine di 1 eV: in questo caso l'energia termica media degli elettroni (ricordiamo che $kT \simeq 0.026\ eV$ a temperatura ambiente) e' sufficiente a far si che una piccola frazione di elettroni, nella coda della distribuzione di energia, possa superare la gap proibita e raggiungere i livelli permessi superiori (banda di conduzione), rendendosi disponibili alla conduzione elettrica. A questo punto la banda di valenza non e' piu' completamente piena, ma vi e' un (limitato) numero di livelli liberi; questo rende possibile accelerare anche gli elettroni di questa banda, che quindi possono contribuire alla conduzione elettrica. Se immaginiamo di applicare un campo elettrico in una certa direzione, gli elettroni della banda di valenza tenderanno a muoversi nella direzione opposta; tutto appare come se i livelli liberi si spostassero nella stessa direzione del campo esterno, cioe' come se delle cariche positive si spostassero nella direzione del campo. E' quindi conveniente pensare alla conduzione in un semiconduttore come legata al moto di elettroni (quelli presenti nella banda di conduzione) e di lacune positive (cioe' i livelli liberi nella banda di valenza) che si muovono in verso opposto.

3.2.3 Conduzione nei metalli e nei semiconduttori

Cerchiamo ora di descrivere il fenomeno in modo piu' quantitativo. Sappiamo che, applicando un campo elettrico E in un metallo, gli elettroni sono accelerati nel verso opposto. A causa degli urti con il reticolo, essi mantengono tuttavia una velocità media finita, data in modulo da

$$v_d = \mu E$$

detta velocità di drift, dove μ è la mobilità degli elettroni. Se consideriamo un conduttore di lunghezza L ed area A, che contenga N elettroni, sotto l'influenza di un campo elettrico E, la corrente risultante è data da

$$I = \frac{qNv_d}{L}$$

e la densità di corrente è data da

$$J = \frac{qNv_d}{LA}$$

cioè

$$J = qv_d$$

dove n è la densità degli elettroni. Combinando questa eq. con la prima, si ha

$$J = qn\mu E = \sigma E$$

dove $\sigma = qn\mu$ è la conducibilità del materiale. Questa equazione è equivalente alla legge di Ohm.

In un semiconduttore intrinseco anche le buche contribuiscono alla conduzione, quindi

$$J = q(n\mu_n + p\mu_p)E = \sigma_i E$$

ovviamente, $p = n = n_i$, per cui

$$\sigma_i = qn_i(\mu_n + \mu_p)$$

mentre invece μ_n e μ_p sono diversi($\mu_p < \mu_n$). Mentre in un metallo $n \approx 10^{22}\ cm^{-3}$, in un semiconduttore $n_i \approx 10^{10}\ cm^{-3}$ a temperatura ambiente. E' chiaro che n_i dipende fortemente dalla temperatura; esso ha infatti un andamento del tipo

$$n_i^2 = A_0 T^3 e^{-\frac{E_G}{kT}}$$

dove E_G e' l'ampiezza della gap proibita ed A_0 una costante dipendente dal materiale.

3.2.4 Semiconduttori drogati

E' possibile realizzare cristalli non completamente puri, mediante l'aggiunta di piccole percentuali di impurita'. Aggiungendo ad esempio ad un semiconduttore atomi con 5 elettroni di valenza (fosforo, arsenico,antimonio) si ha un aumento degli elettroni di conduzione (drogaggio di tipo **n**). Infatti l'elettrone in eccesso risulta debolmente legato e puo' quindi facilmente liberarsi (per effetto dell'energia termica) e raggiungere la banda di conduzione. Cosa succede se aggiungiamo invece atomi con 3 elettroni di valenza (boro, gallio,indio)? In questo caso l'atomo di impurezza funziona come trappola per gli elettroni nella banda di valenza: l'elettrone catturato lascia libero un posto nella banda di valenza, e si ha un aumento delle lacune (drogaggio di tipo **p**). Da un punto di vista energetico si puo' schematizzare la presenza di impurezze come la presenza di livelli localizzati, vicini al bordo inferiore della banda di conduzione (impurezze pentavalenti) o al bordo superiore della banda di valenza (impurezze trivalenti) (vedi Fig. 3.2)

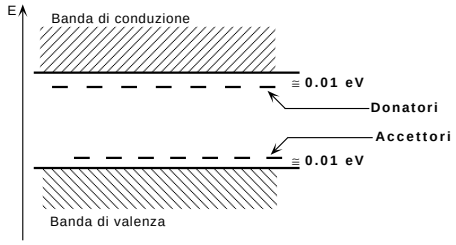

Figura 3.2: *Livelli energetici in semiconduttori drogati*

Con considerazioni di tipo statistico si può dimostrare che

$$np = n_i^2$$

(legge di azione di massa), cioè che il prodotto delle concentrazioni di elettroni e lacune resta costante ($=n_i^2$). Poichè il materiale è elettricamente neutro, si ha che

$$N_D + p = N_A + n$$

dove N_D e N_A sono le concentrazioni di impurezze, supposte tutte ionizzate. Per esempio, in un materiale di tipo n ($N_A = 0$), $n \gg p$, cioè

$$n \approx N_D$$

Quindi, dall'equazione $np = n_i^2$ segue che

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

In un semiconduttore il trasporto di carica può avvenire per diffusione in presenza di un gradiente di concentrazione. Per es., per le lacune si avrà

$$J_p = -qD_p\frac{dp}{dx}$$

dove D_p è il coefficiente di diffusione. $D_p, D_n, \mu_p e \mu_n$ sono legati dalla relazione di Einstein

$$\frac{D_p}{\mu_p} = \frac{D_n}{\mu_n} = \frac{kT}{q}$$

La corrente totale (drift + diffusione) è data da

$$J_p = q\mu_p pE - qD_p\frac{dp}{dx}$$

e

$$J_n = q\mu_n nE + qD_n\frac{dn}{dx}$$

3.3 Diodi a giunzione

Possiamo ora costruire una giunzione immaginando di saldare tra loro un cristallo di tipo p ed uno di tipo n (vedi Fig. 3.3). In realta' cio' viene fatto producendo un unico cristallo con una parte drogata con impurezze di tipo trivalente ed una parte drogata con impurezze di tipo p.

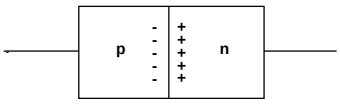

Figura 3.3: *Giunzione pn*

Per effetto di fenomeni di diffusione si avra' una migrazione di elettroni dal lato n al lato p ed una migrazione opposta delle lacune. In questo modo pero' si ha un eccesso di cariche negative nel lato p ed un eccesso di cariche positive nel lato n e di conseguenza la nascita di un campo elettrico interno che tende ad opporsi alla diffusione. Si raggiunge quindi una situazione di equilibrio (vedi Fig. 3.4).
L'andamento delle grandezze elettriche in funzione di x e' dato da:

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\epsilon}$$

$$E = -\frac{dV}{dx} = \int \frac{\rho}{\epsilon}dx$$

$$V = -\int Edx$$

Possiamo ora applicare un campo elettrico esterno, che puo' avere un verso concorde (Fig. 3.5a) o discorde (Fig. 3.5b) rispetto a quello interno.

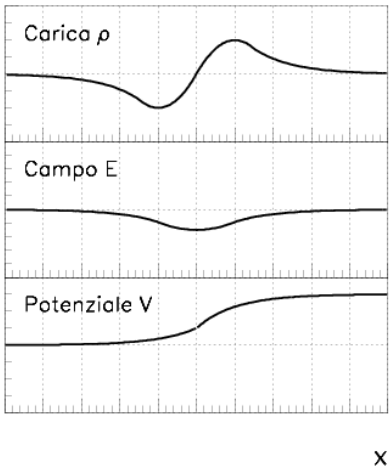

Figura 3.4: *Densita' di carica, campo elettrico e potenziale, in funzione di x*

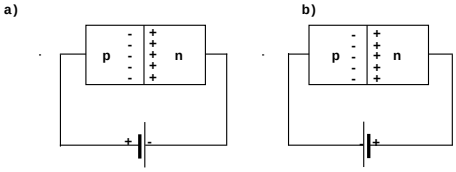

Figura 3.5: *a) Polarizzazione diretta; b) Polarizzazione inversa*

Nel primo caso (detto di polarizzazione inversa) non si ha passaggio di corrente: la differenza di potenziale indotta dall'esterno si somma alla barriera di potenziale V_0. Si ha comunque una corrente I_0 (corrente di saturazione inversa del diodo) dovuta alla generazione di coppie lacuna-elettrone (le lacune generate nella regione n passano alla p, e viceversa per gli elettroni). Si ha cioè una corrente molto piccola dovuta ai portatori minoritari.
Invece, nel secondo caso (polarizzazione diretta) la barriera di potenziale si abbassa e quindi è possibile il passaggio di lacune dalla regione p alla regione n e di elettroni da n a p.

In termini quantitativi la relazione tra corrente e tensione applicata dall'esterno e' data da:

$$I = I_0(e^{\frac{V}{\eta V_T}} - 1)$$

dove il parametro η e' un piccolo fattore correttivo, dipendente dal tipo di materiale e dal processo di fabbricazione[1]. Il termine V_T indica il cosidetto equivalente in volt della temperatura (spesso chiamato anche *tensione termica*), cioe'

$$V_T = \frac{kT}{q}$$

a temperatura ordinaria ($T = 300\ ^oK$) $V_T \simeq 0.026\ V$.

Come si vede (Fig. 3.6), per polarizzazione diretta ($V > 0$) la corrente sale bruscamente quando V supera $0.6\ V$ (tensione di ginocchio).

[1] η e' spesso omesso perche' approssimato a 1 (anche se potrebbe arrivare a circa 2 in alcuni casi).

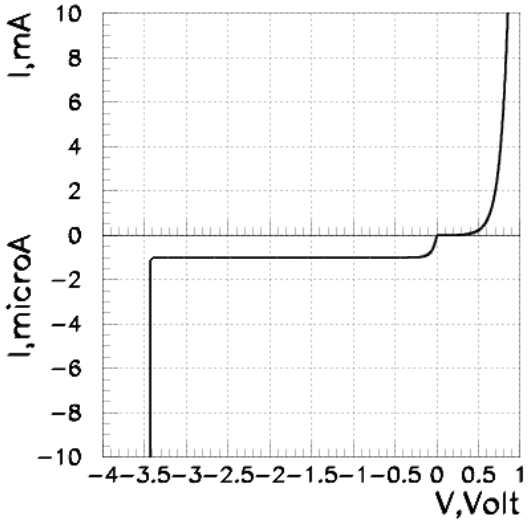

Figura 3.6: *Caratteristica corrente tensione in un diodo al silicio. Si noti la diversa scala di corrente nei due semipiani.*

Per polarizzazione inversa, si ha una piccola corrente, I_0, pressoche' costante (si noti la diversa scala verticale!), fino a quando la tensione applicata arriva ad un certo valore, in cui si ha un brusco cambiamento di regime: la caratteristica diviene quasi verticale. Questo significa che la resistenza del diodo diviene praticamente nulla. Il valore di tensione a cui avviene questo fenomeno (dovuto all'effetto tunnel) dipende dalle caratteristiche costruttive del diodo; questo fenomeno puo' essere sfruttato per costruire speciali diodi (diodi Zener) che si prestano, come vedremo piu' avanti, per realizzare dei regolatori di tensione.
Anche I_0 dipende dalla temperatura: si ha una legge del tipo

$$I_0 = AT^m e^{-\frac{V_0}{\eta V_T}}$$

dove m dipende di nuovo dal materiale (circa 1.5 nel silicio).

Il diodo e' un elemento non lineare, quindi la resistenza, intesa come rapporto tensione-corrente, non ha molto significato. Si puo' tuttavia definire una resistenza dinamica, $r = dV/dI$, che varia al variare di I; si ha quindi:

$$\frac{1}{r} \equiv g = \frac{dI}{dV} = \frac{I_0 e^{\frac{V}{\eta V_T}}}{\eta V_T} = \frac{I + I_0}{\eta V_T}$$

Quando $I \gg I_0$

$$g \approx \frac{I}{V_T} \tag{3.1}$$

In questa regione è possibile approssimare la caratteristica del diodo come una retta con pendenza $1/R_f$, dove R_f e'in genere dell'ordine di qualche Ohm.

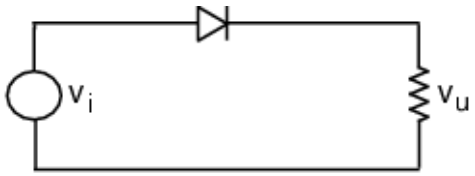

Figura 3.7: *Semplice circuito con diodo*

3.4 Circuiti con diodi

Consideriamo il circuito in Fig. 3.7. Scrivendo l'equazione della maglia si ha:

$$v = v_i - iR_L$$

Per risolvere questa equazione occorre conoscere la caratteristica corrente tensione del diodo, che, come abbiamo detto e' di tipo non lineare; la soluzione del problema diviene quindi complicata. Si puo' pero' ricorrere ad un metodo grafico. Nella Fig. 3.8a e' riportata la caratteristica del diodo e la retta la cui equazione e' data dall'espressione precedente (retta di carico). L'intersezione tra le due curve fornisce la relazione corrente-tensione di cui abbiamo bisogno per risolvere il circuito. In generale v_i e' funzione del tempo, quindi la retta di carico si muove: e' allora conveniente costruire la cosidetta caratteristica dinamica, cioe' la curva che lega i e v_i (Fig. 3.8b).
Poiche' $v_u = iR_L$ dalla caratteristica dinamica si puo' ricavare la relazione tra v_u e v_i (transcaratteristica), mostrata in Fig. 3.8c. Grazie alla transcaratteristica e' possibile, conoscendo l'andamento temporale della v_i ricavare in modo grafico l'andamento temporale della v_u
Il metodo predetto e' molto macchinoso. Per molte applicazioni e' sufficiente approssimare il diodo come una resistenza (normalmente piccola) R_f, quando la tensione ai suoi capi supera la tensione di ginocchio V_γ, e come una resistenza R_r molto grande (al limite infinita) altrove (Fig. 3.9a).
Considerando per esempio il circuito di Fig. 3.7, se

$$v_i = V_m \sin \omega t$$

si ha

$$\begin{aligned} i &= 0 && \text{per} \quad v_i < v_\gamma \\ i &= \frac{v_m \sin \alpha - v_\gamma}{R_L + R_f} && \text{per} \quad v_i > v_\gamma \end{aligned}$$

Il predetto circuito si chiama raddrizzatore (a singola semionda), infatti, se

$$v_i = V_m \sin \omega t$$

si ha

$$\begin{aligned} i &= I_m \sin \omega t && \text{per} \quad 0 \leq \omega t \leq \pi \\ i &= 0 && \text{per} \quad \pi \leq \omega t \leq 2\pi \end{aligned}$$

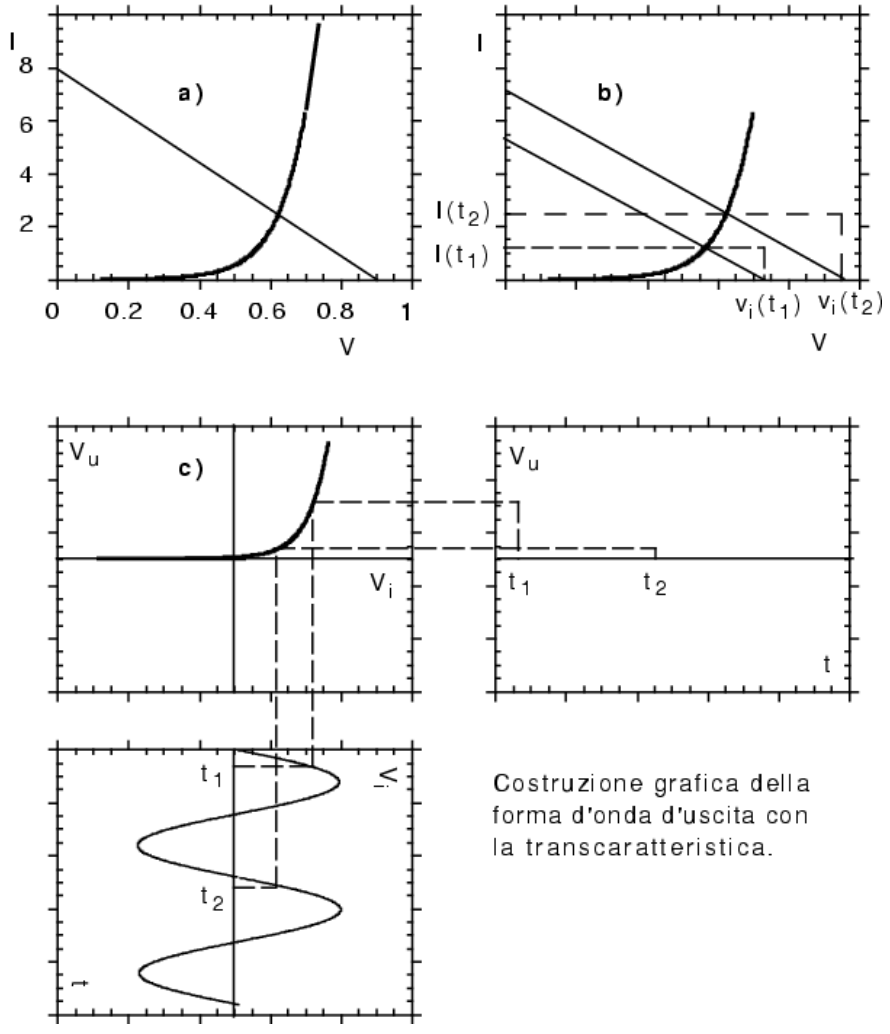

Figura 3.8: *a) Caratteristica del diodo; b) Caratteristica dinamica; c) Costruzione grafica della forma d'onda d'uscita a partire dalla forma d'onda d'ingresso*

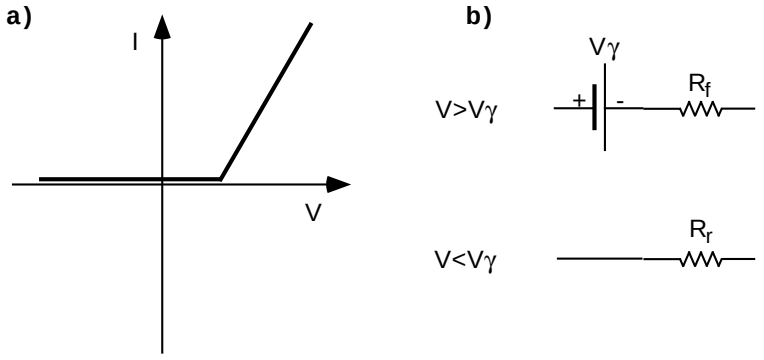

Figura 3.9: *a) Modello lineare del diodo; b) Circuito equivalente dell'esempio di Fig. 3.7*

dove

$$I_m = \frac{V_m}{R_f + R_L}$$

e quindi

$$v_u = R_L I_m \sin \omega t \quad \text{per} \quad 0 \leq \omega t \leq \pi$$
$$v_u = 0 \quad \text{per} \quad \pi \leq \omega t \leq 2\pi$$

Il valore medio di v_u è dato da

$$< v_u >= \frac{1}{2\pi} \int_0^{2\pi} R_L I_m \sin \omega t \, d\omega t = \frac{R_L I_m}{\pi}$$

In Fig. 3.10 sono mostrate le forme d'onda d'ingresso e d'uscita.

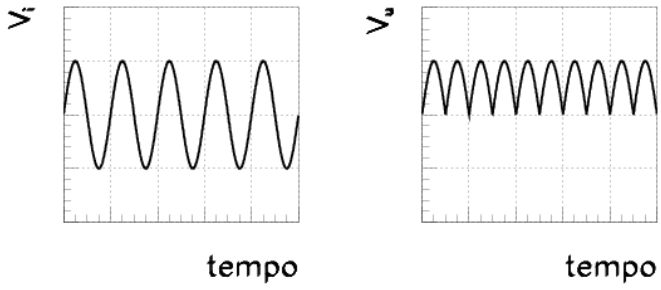

Figura 3.10: *Forme d'onda d'ingresso e d'uscita per il raddrizzatore*

I raddrizzatori sono in genere utilizzati per ottenere una tensione continua partendo da una tensione sinusoidale. Da questo punto di vista il circuito precedente ha delle prestazioni molto modeste: la tensione d'uscita ha un valor medio positivo, ma e' ben lungi dall'essere costante. Un miglioramento notevole puo' essere ottenuto aggiungendo un opportuno condensatore al circuito (Fig. 3.11a).
Se la costante di tempo RC e' grande rispetto al periodo dell'onda di ingresso, la forma d'onda d'uscita e' quella riportata in Fig. 3.11b.

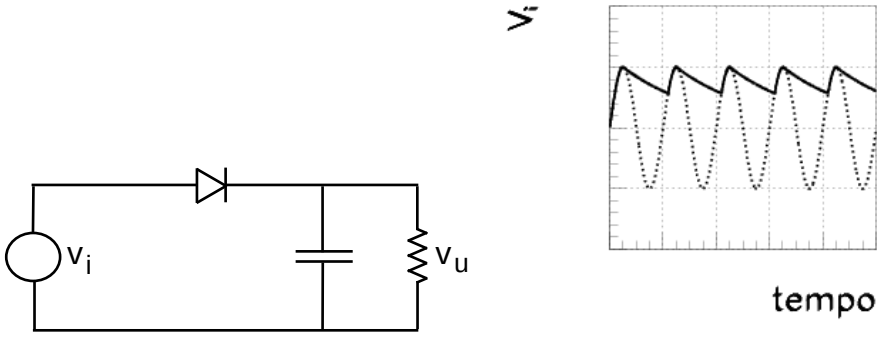

Figura 3.11: *a) Raddrizzatore con filtro capacitivo; b) Forma d'onda d'uscita*

Si possono poi avere prestazioni ancora migliori con il circuito raddrizzatore a doppia semionda (raddrizzatore a ponte) mostrato in Fig. 3.12a. La forma d'onda d'uscita e' quella riportata in Fig. 3.12b. L'aggiunta di un opportuno condensatore in parallelo al carico R_L puo', anche in questo caso, migliorare ulteriormente la forma d'onda d'uscita.

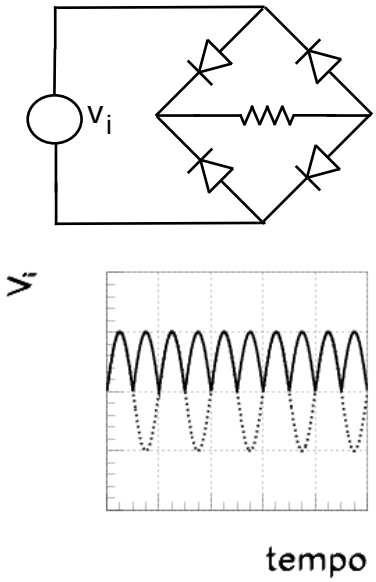

Figura 3.12: *a) Raddrizzatore a ponte; b) Forma d'onda d'uscita*

Nella Fig. 3.13 sono mostrati ulteriori esempi di circuiti con diodi, e le relative forme d'onda.

3.5 Approfondimento sulla fisica dei semiconduttori

Abbiamo visto che esistono due fenomeni che producono la corrente nei semiconduttori:
a) la deriva dei portatori, in presenza di un campo elettrico;
b) la diffusione dei portatori quando esiste un gradiente di concentrazione.

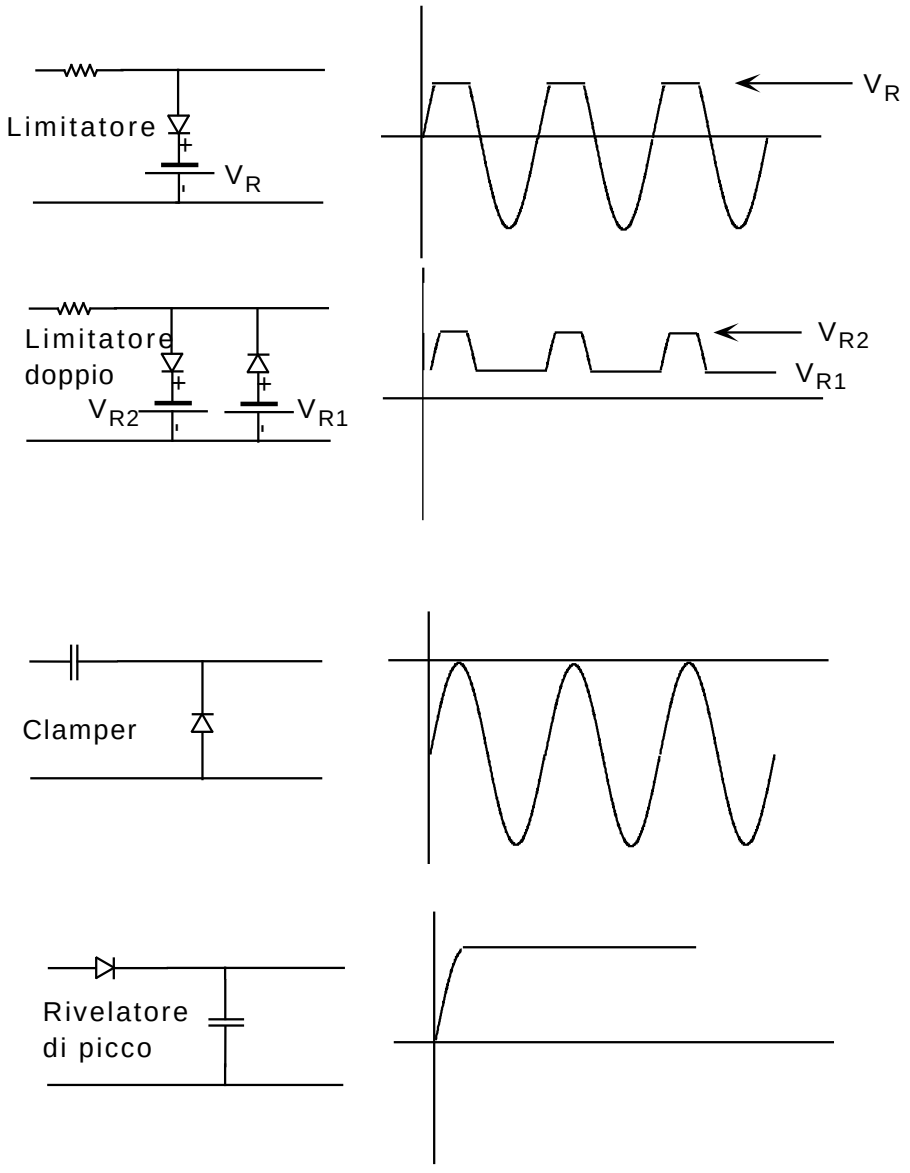

Figura 3.13: *Altri esempi di circuiti con diodi e forme d'onda d'uscita per sollecitazione sinusoidale*

La densità di corrente di deriva è data da

$$J = (n\mu_n + p\mu_p)qE = \sigma E$$

mentre la densità della corrente di diffusione è data da

$$J_p = -qD_p\frac{dp}{dx}$$

Inoltre la relazione di Einstein collega D_p e D_n:

$$\frac{D_p}{\mu_p} = \frac{D_n}{\mu_n} = V_T$$

Consideriamo ora un elemento di volume di area A e lunghezza dx, con una concentrazione di lacune p. Si abbia una corrente entrante in x e una corrente uscente $I_p + dI_p$ nel punto $x + dx$. Allora dI_p/q rappresenta la diminuzione (per secondo) della concentrazione di lacune nel volume Adx. Quindi la diminuzione per unità di tempo per unità di volume

$$\frac{dI_p}{qAdx} = \frac{dJ_p}{qdx}$$

Esiste poi un generatore di lacune di natura termica, che produce un incremento per unità di tempo

$$g = \frac{p_0}{\tau_p}$$

(dove p_0 è la concentrazione all'equilibrio e τ_p la vita media), e una diminuzione p/τ_p per opera della ricombinazione. Poichè la carica si conserva,

$$\frac{dp}{dt} = \frac{p_0 - p}{\tau_p} - \frac{dJ_p}{qdx}$$

che è l'equazione di continuità.

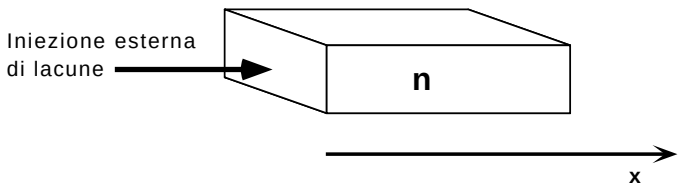

Figura 3.14: *Iniezione esterna di lacune*

Si consideri il caso in cui si ha una iniezione esterna di lacune in un materiale di tipo n (Fig. 3.14); si vuole analizzare come varia in funzione di x la concentrazione p in condizioni di regime. Ora

$$p = p' + p_0$$

dove p' rappresenta la concentrazione in eccesso. Si fa l'ipotesi che $p' \ll n$ e si trascura la corrente di deriva delle lacune (non quella degli elettroni). Si hanno le relazioni

$$J_p = -qD_p\frac{dp}{dx}$$

$$\frac{dp}{dt} = \frac{p_0 - p}{\tau_p} - \frac{dJ_p}{qdx}$$

Imponendo (a regime),

$$\frac{dp}{dt} = 0$$

si trova

$$\frac{d^2p}{dx^2} = \frac{p - p_0}{D_p\tau_p}$$

Posto $L_p = (D_p\tau_p)^{1/2}$ (lunghezza di diffusione), si arriva all' equazione per p':

$$\frac{d^2p'}{dx^2} = \frac{p'}{L_p^2}$$

la cui soluzione è

$$p'(x) = k_1 e^{-\frac{x}{L_p}} + k_2 e^{\frac{x}{L_p}}$$

Poiche' per $x \to \infty$ $p' \to 0$, si ha $k_2 = 0$. Quindi posto che $p'(0)$ e' la concentrazione per $x = 0$, si ha

$$p'(x) = p'(0)e^{-\frac{x}{L_p}} = p(x) - p_0$$

La corrente di diffusione è allora data da

$$I_p(x) = \frac{A_q D_p p'(0)}{L_p} e^{-\frac{x}{L_p}} = \frac{A_q D_p}{L_p}[p(0) - p_0]e^{\frac{x}{L_p}}$$

La corrente di diffusione degli elettroni è data da

$$A_q D_n \frac{dn}{dx}$$

Supponendo sempre che il materiale sia neutro, $n' = p'$, cioè

$$n - n_0 = p - p_0$$

poichè n_0 e p_0 non dipendono da x si ha quindi

$$\frac{dn}{dx} = \frac{dp}{dx}$$

Quindi

$$A_q D_n \frac{dn}{dx} = A_q D_n \frac{dp}{dx} = -\frac{D_n}{D_p} I_p$$

(con $D_n/D_p \approx 2$ per il Ge e ≈ 3 per il Si).

Se il semiconduttore è inserito in un circuito aperto, la corrente totale è nulla; deve quindi esistere una corrente di deriva dei portatori maggioritari I_{nd} tale che

$$I_p + (I_{nd} - \frac{D_n}{D_p} I_p) = 0$$

ovvero

$$I_{nd} = (\frac{D_n}{D_p} - 1)I_p$$

che quindi diminuisce esponenzialmente con la distanza. E' necessaria cioe' la presenza di un campo elettrico E che mantenga questa corrente. Questo campo è generato dai portatori iniettati

$$E = \frac{1}{Aqn\mu_n}(\frac{D_n}{D_p} - 1)I_p$$

nell'ipotesi che I_{pd} sia nulla (o trascurabile, il che è vero se $p \ll n$)

Quando si applica una polarizzazione diretta ad un diodo, le lacune vengono iniettate nel lato n e gli elettroni nel lato p. La corrente di diffusione delle lacune nel lato n è dato, per $x = 0$, da

$$I_{pn}(0) = \frac{A_q D_p}{L_p}(p_n(0) - p_{n0})$$

$p_n(0)$ é funzione del potenziale esterno V. Si trova che

$$p_n(0) = p_{n0}e^{V/V_t}$$

(legge delle giunzioni). Si ha quindi

$$I_{pn}(0) = \frac{A_q D_p p_{n0}}{L_p}(e^{V/V_T} - 1)$$

e analoga espressione per $I_{np}(0)$. La corrente totale in $x = 0$ è data da

$$I = I_{pn}(0) + I_{np}(0) = I_0(e^{V/V_T} - 1)$$

dove I_0 dipende dai parametri fisici del diodo. Poichè I è la stessa in tutto il diodo, essa non dipende più da x. I ragionamenti precedenti valgono anche per $V < 0$ (polarizzazione inversa).

Poichè I non dipende da x, si dovranno avere delle correnti maggioritarie I_{pp} e I_{nn} fortemente dipendenti da x, legate dalla relazione

$$I_{nn} = I - I_{pn}(x)$$

I_{nn} è composta da una parte di diffusione e una parte di deriva, e lo stesso vale per I_{pp}.

In questo ragionamento si è fatta l'ipotesi di poter trascurare i fenomeni di generazione e ricombinazione di coppie. Ciò è vero per il Ge ma non per il Si (da cui il fattore η diverso da 0).

3.5.1 Capacità della giunzione

In polarizzazione inversa la capacità è data dalla regione di carica spaziale (che dipende da V). Si definisce quindi una capacità della regione di transizione

$$C_T = |\frac{dQ}{dV}|$$

Si può dimostrare che, per una giunzione a gradino,

$$C_T = \frac{\epsilon A}{w}$$

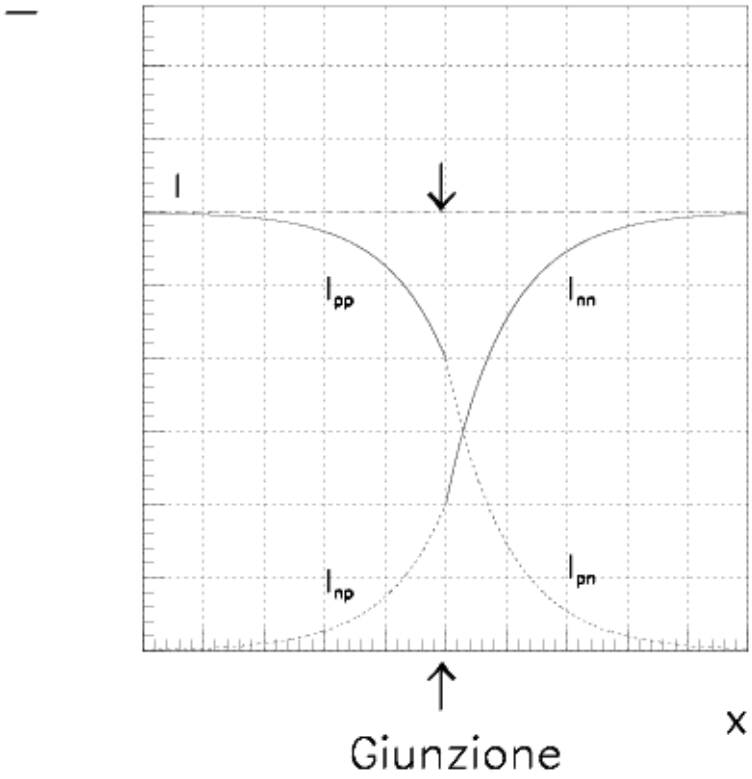

Figura 3.15: *Componenti della corrente in un diodo*

dove A è la superficie e w lo spessore della regione di transizione.

In polarizzazione diretta entrano in gioco le cariche iniettate, che danno luogo a una capacità assai più grande di C_T e che si chiama capacità di diffusione C_D. In condizioni statiche

$$C_D = \frac{dQ}{dV} = \tau \frac{dI}{dV} = \tau g_d = \frac{\tau}{r_d}$$

dove τ è la vita media. Da cui

$$C_D = \frac{\tau I}{\eta V_T}$$

Naturalmente se la corrente è dovuta a entrambi i tipi di portatori si avrà una capacita'

$$C_D = C_{D_p} + C_{D_n}$$

Comunque, in polarizzazione inversa, poichè g è molto piccolo, anche C_D è molto piccolo. Per polarizzazione diretta invece $C_D \gg C_T$.

3.6 Il transistor a giunzione

Il transistor e' un dispositivo realizzato con un cristallo semiconduttore composto da tre regioni di diverso drogaggio: due di tipo p separate da una sottile regione di tipo n (in

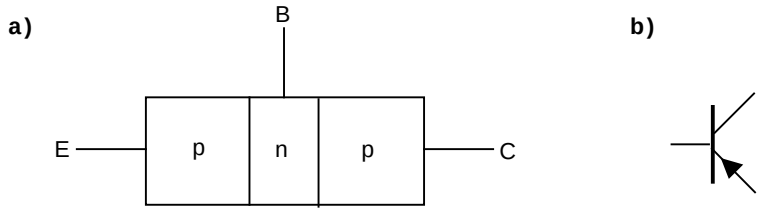

Figura 3.16: *a) Transistor pnp; b) Simbolo circuitale*

questo caso si parla di transistor pnp), ovvero due di tipo n separate da una sottile regione di tipo p (transistor npn). Nella Fig. 3.16a e' mostrato lo schema di un transistor pnp: le tre regioni (ciascuna dotata di una opportuna connessione metallica verso l'esterno) prendono il nome di Emettitore, Base e Collettore. Nella Fig. 3.16b e' mostrato il simbolo circuitale del transistor (l simbolo del tipo npn e' identico, cambia solo il verso della freccia, uscente anziche' entrante).

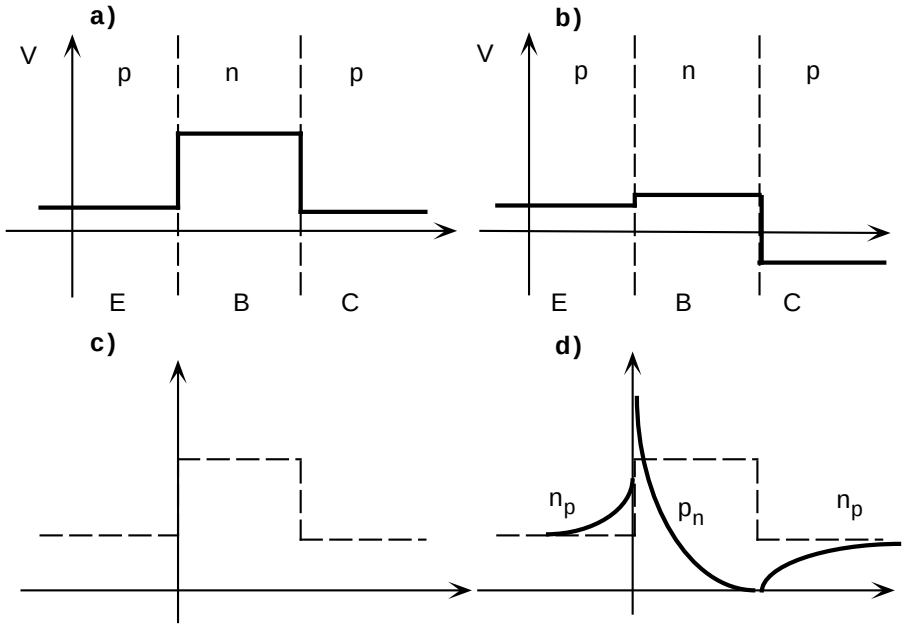

Figura 3.17: *a)Andamento del potenziale V per transistor non connesso; b)idem quando il transistor e' nella regione attiva; c)Concentrazione dei portatori minoritari per transistor non connesso; d)idem quando il transistor e' nella regione attiva*

Nella Fig. 3.17a sono mostrati l'andamento del potenziale nelle tre regioni e la concentrazione dei portatori minoritari, quando il transistor non e' connesso con l'esterno. Se ora polarizziamo direttamente la giunzione base-emettitore e inversamente quella base-collettore(Fig. 3.18), la situazione del potenziale diviene quella in Fig. 3.17b: si dice che il transistor e' polarizzato nella regione attiva. Per comprendere cio' che avviene dobbiamo esaminare accuratamente le varie componenti della corrente (Fig. 3.19): I_{pE} e I_{nE} sono le correnti di diffusione dei portatori minoritari, quindi

$$I_E = I_{pE} + I_{nE}$$

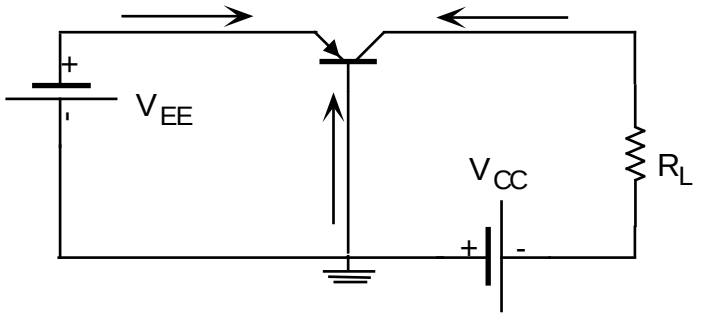

Figura 3.18: *Transistor in regione attiva*

In genere, $I_{nE} \ll I_{pE}$, perchè la base è drogata molto poco rispetto all'emettitore.

Gran parte della corrente I_{pE} raggiunge il collettore, perchè la regione di base è molto stretta; $(I_E - I_{pC_1})$ è ciò che si perde nella base per effetto della ricombinazione. La I_{C0}, composta da lacune ed elettroni, è la corrente di saturazione inversa del diodo BC. Si ha quindi

$$I_c = I_{CO} - I_{pC_1} = I_{C0} - \alpha I_E$$

Il fattore α e' molto vicino ad 1 ($\alpha \approx .9 \div .995$) e dipende da I_E, da V_{CB} e dalla temperatura assoluta T.

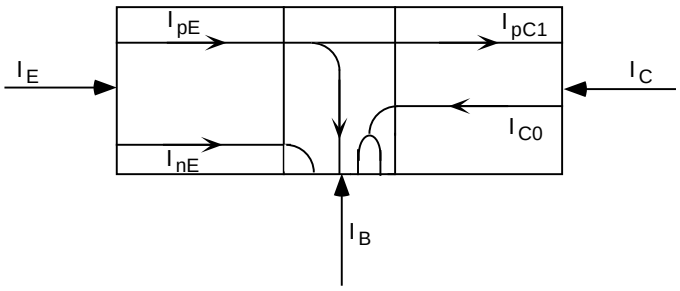

Figura 3.19: *Componenti della corrente in un transistor*

3.6.1 Rappresentazione di Ebers-Moll del transistor

Il transistor può essere descritto, anche quantitativamente, come un sistema di due diodi accoppiati (Fig. 3.20).
Usando le equazioni dei diodi si ha

$$I_E = I_{ED} - \alpha_R I_{CD} = I_{ES}(e^{\frac{V_{EB}}{V_T}} - 1) - \alpha_R I_{CS}(e^{\frac{V_C B}{V_T}} - 1)$$

$$I_C = -\alpha_F I_{ED} + I_{CD} = -\alpha_F I_{ES}(e^{\frac{V_{EB}}{V_T}} - 1) + I_{CS}(e^{\frac{V_{CB}}{V_T}} - 1)$$

Le quantità I_{CS} e I_{ES} sono le correnti di saturazione inversa dei due diodi. I_{CS}, I_{ES}, α_R, α_F sono funzione delle densità di impurezze e della geometria. In sostanza sono

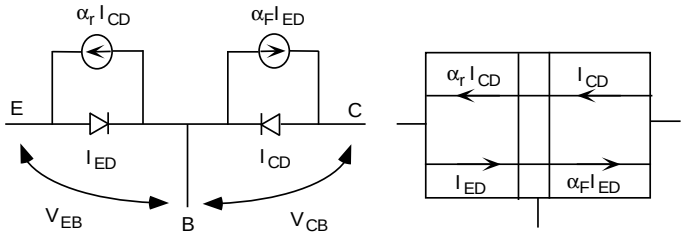

Figura 3.20: *Modello di Ebers-Moll*

parametri costruttivi e possono essere determinati solo sperimentalmente. Si può comunque dimostrare che essi sono legati tra loro dalla condizione di reciprocità

$$\alpha_F I_{ES} = \alpha_R I_{CS}$$

Come gia' abbiamo detto $\alpha_F \approx .9 \div .995$; invece $\alpha_R \approx .4 \div .8$. Le costanti I_{ES} e I_{CS} sono dell'ordine di $10^{-15}A$. Ovviamente

$$I_B = -(I_E + I_C)$$

Naturalmente si può trattare nello stesso modo un transistor npn, scegliendo opportunamente i versi dei diodi.

Ponendo $V_{CB} = 0$ si ottiene

$$I_E = I_{ES}(e^{\frac{V_{EB}}{V_T}} - 1)$$
$$I_C = -\alpha_F I_{ES}(e^{\frac{V_{EB}}{V_T}} - 1)$$

cioè $I_C = -\alpha_F I_E$.
Quindi

$$\alpha_F \equiv -\frac{I_C}{I_E}\bigg|_{V_{CB}=0}$$

Analogamente, per $V_{EB} = 0$

$$\alpha_R \equiv -\frac{I_E}{I_C}\bigg|_{V_{EB}=0}$$

Per $V_{CB} = 0$ posso scrivere

$$I_B = -(I_C + I_E) = -(1 - \alpha_F)I_E$$

e anche

$$I_C = \frac{\alpha_F}{1 - \alpha_F} I_B = \beta_F I_B$$
$$I_E = \frac{-I_B}{1 - \alpha_F} = -(1 + \beta_F)I_B$$

dove il parametro β_F[2] é definito come

$$\beta_F = \frac{\alpha_F}{1 - \alpha_F}$$

[2]questo parametro é spesso denominato h_{FE}, per motivi che diverranno chiari più avanti

Generalmente

$$\beta_F = 50 \div 400$$

ma esistono in commercio transistor con valori di β_F molto piú elevati.
Analogamente, si puó definire il parametro β_R come

$$\beta_R = \frac{\alpha_R}{1 - \alpha_R}$$

In genere si ha

$$\beta_R = 1 \div 5$$

L'enorme differenza tra β_F e β_R é ovviamente il risultato di scelte costruttive.
Da un punto di vista pratico e' utile introdurre le correnti I_{C0} e I_{E0}, definite rispettivamente come la corrente che circola nel collettore ad emettitore aperto, e come la corrente che circola nell'emettitore a collettore aperto. E' possibile ricavare la relazione che lega questi parametri a I_{CS} e I_{ES}; si trova

$$I_{C0} = (1 - \alpha_F \alpha_R) I_{CS}$$

$$I_{E0} = (1 - \alpha_F \alpha_R) I_{ES}$$

3.6.2 Modi di operazione del transistor

Poiché vi sono 2 giunzioni, sono possibili 4 modi di operazione in relazione alle possibili polarizzazioni delle due giunzioni:

EB	CB	Modo
diretto	inverso	Attivo
inverso	inverso	Cut-Off
diretto	diretto	Saturazione
inverso	diretto	Attivo-inverso

Il modo attivo-inverso è scarsamente usato. Il modo attivo (quello che abbiamo iniziato a studiare in precedenza) puo' essere ottenuto in diverse configurazioni.

Configurazione a base comune
Le equazioni delle due maglie sono

$$-V_{CC} = V_{CB} + I_C R_L$$

$$V_{EE} = R_I I_E + V_{EB}$$

Inoltre:

$$I_B = -(I_C + I_E)$$

Fra le 3 correnti e le 2 tensioni si hanno delle relazioni funzionali(vedi ad esempio il modello di Ebers-Moll). In genere si considerano

$$V_{BE} = \Phi_1(V_{CB}, I_E)$$

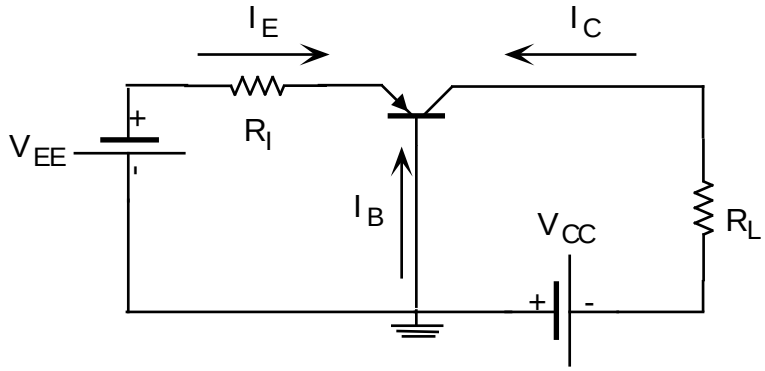

Figura 3.21: *Configurazione a base comune*

$$I_C = \Phi_2(V_{CB}, I_E)$$

cioè si guardano corrente d'uscita e tensioni d'ingresso come funzioni di tensione d'uscita e corrente d'ingresso. I_E e I_C vengono espresse in termini di I_{E0} e I_{C0} ricavando quindi

$$I_E = I_{E0}(e^{\frac{V_{EB}}{V_T}} - 1) - \alpha_R I_C$$

$$I_C = -\alpha_F I_E + I_{C0}(e^{\frac{V_{CB}}{V_T}} - 1)$$

Nel Fig. 3.22a e' riportata I_C in funzione di V_{CB}, per vari valori di I_E (caratteristiche d'uscita). Si distinguono la regione attiva e la regione di saturazione ($V_{CB} < 0$), mentre la regione di cut-off e' quella al di sotto della curva $I_E = 0$.

Nella Fig. 3.22b sono riportate le caratteristiche di ingresso, cioe' V_{EB} in funzione di I_E usando come parametro V_{CB}.

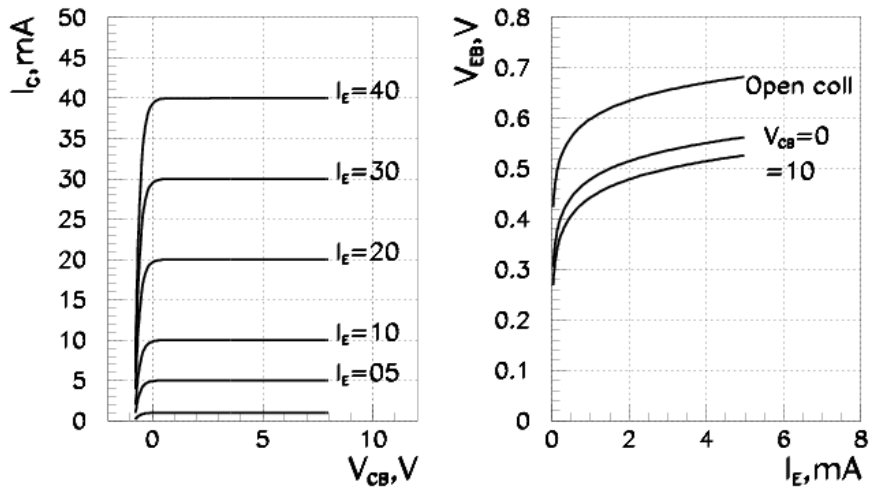

Figura 3.22: *Configurazione a base comune:a) Caratteristiche d'uscita; b) Caratteristiche d'ingresso*

La dipendenza delle caratteristiche d'ingresso da V_{CB} è dovuta al cosiddetto effetto Early (modulazione dello spessore di base): lo spessore effettivo della base diminuisce, quando V_{CB} cresce, quindi α_F cresce. Quindi, fissato V_{EB}, I_E cresce al crescere di V_{CB}.

Le due equazioni delle maglie individuano due rette di carico (d'uscita e di ingresso):

$$V_{CB} = -V_{CC} - I_C R_L$$

$$V_{EB} = V_{EE} - I_E R_I$$

Riportando le due rette rispettivamente sulle caratteristiche d'ingresso e di uscita e' quindi possibile trovare il punto di lavoro del transistor, cioe' i valori di I_C,V_{CB},V_{EB} ed I_E che soddisfano a tutte le condizioni imposte.
Questa configurazione e' spesso indicata con la sigla CB (Common Base in inglese).

Configurazione a emettitore comune

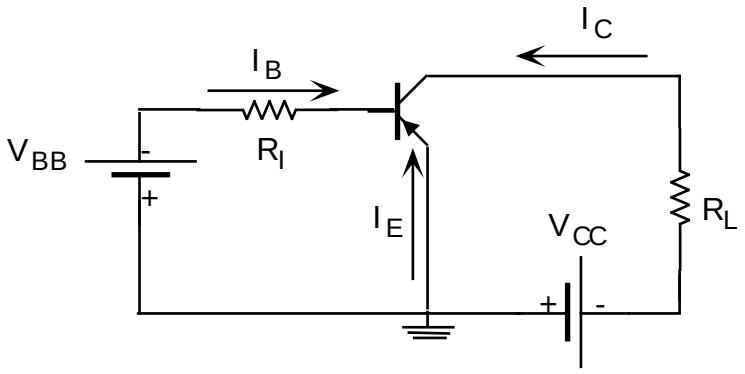

Figura 3.23: *Configurazione ad emettitore comune*

Nella configurazione ad emettitore comune si usano come variabili dipendenti I_C e V_{BE}, cioe'

$$V_{BE} = f_1(V_{CE}, I_B)$$

$$I_C = f_2(V_{CE}, I_B)$$

Possiamo scrivere

$$I_C = -\alpha_F I_E + I_{C0}$$

$$I_C = \frac{\alpha_F I_B}{1-\alpha_F} + \frac{I_{C0}}{1-\alpha_F}$$

$$I_C = \beta_F I_B + (1+\beta_F) I_{C0} \approx \beta_F I_B$$

poichè $I_B \gg I_{C0}$
Nella Fig. 3.24 sono mostrate le caratteristiche d'uscita e di ingresso in questa configurazione, spesso indicata con la sigla CE (Common Emitter). La dipendenza di I_C da V_{CE} è dovuta, di nuovo, all'effetto Early, cioè alla variazione di α_F con V_{CE}. Nella configurazione CB il cut-off si ha per $I_E = 0$ e quindi $I_C = -I_B = I_{C0}$; nella configurazione CE, quando $I_B = 0$ si ha $I_E = -I_C$

$$I_C = \frac{I_{C0}}{1-\alpha_F} \equiv I_{CE0}$$

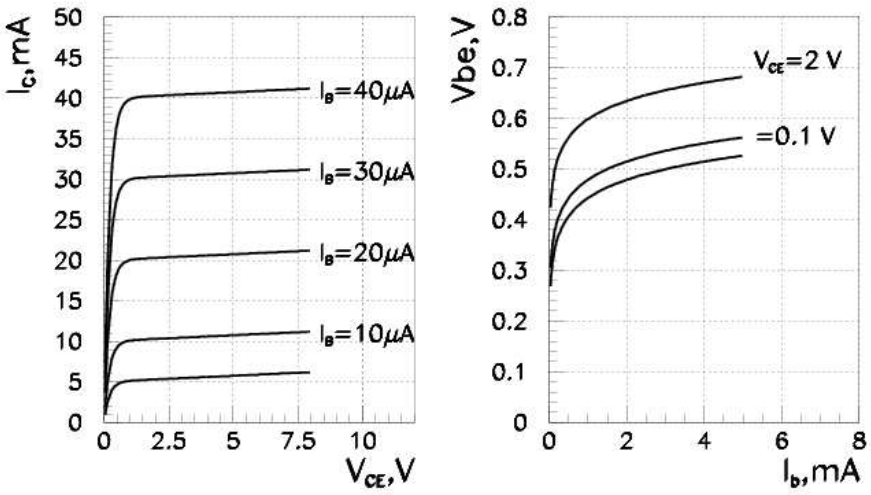

Figura 3.24: *Configurazione ad emettitore comune: a) Caratteristiche d'uscita; b) Caratteristiche d'ingresso*

Quando $I_C \approx I_{C0}$, α_F è praticamente zero, cioè prevale la ricombinazione nella base. Quindi

$$I_{C0} \approx I_{CE0}$$

Si vede anche che $V_{BE} \approx 0$.
Quando V_{CE} si avvicina a zero anche il diodo Collettore - Base comincia a condurre e si entra nella regione di saturazione: non vi è più una relazione lineare tra I_C e I_B e le curve si addensano. Piccole variazioni di V_{CE} provocano enormi variazioni di corrente. Questo avviene per valori di $V_{CE} \simeq 0.2\ V$ (per transistor al silicio). Ai fini pratici questo numero può essere considerato praticamente costante ed é comunemente denominato V_{CEsat}.

Configurazione a collettore comune
Il comportamento è sostanzialmente simile a quello in configurazione a emettitore comune. Nel prossimo Capitolo avremo modo di approfondirlo.

Modello del transistor in continua
Studiamo un po' più in dettaglio il comportamento del transistor in configurazione CE, usando alcune ipotesi semplificative. Supponiamo di avere il circuito di Fig 3.25a: se il transistor e' nella regione attiva, il circuito puo' essere considerato approssimativamente equivalente allo schema di Fig. 3.25b. Se invece e' in saturazione lo schema equivalente diviene quello di Fig. 3.25c
Purtroppo a priori non si sa se il transistor è nella regione attiva o no: bisogna quindi fare l'ipotesi che lo sia e verificare a posteriori se i risultati ottenuti sono coerenti con l'ipotesi. Supponiamo che i parametri del circuito in Fig 3.25a siano:

$$V_{CC} = 10\ V; \quad R_C = 2\ K\Omega; \quad R_B = 300\ K\Omega \quad \beta_F = 100$$

Possiamo scrivere l'equazione di Kirchoff della maglia che contiene R_B:

$$-V_{cc} + I_B R_B + V_{BE} = 0$$

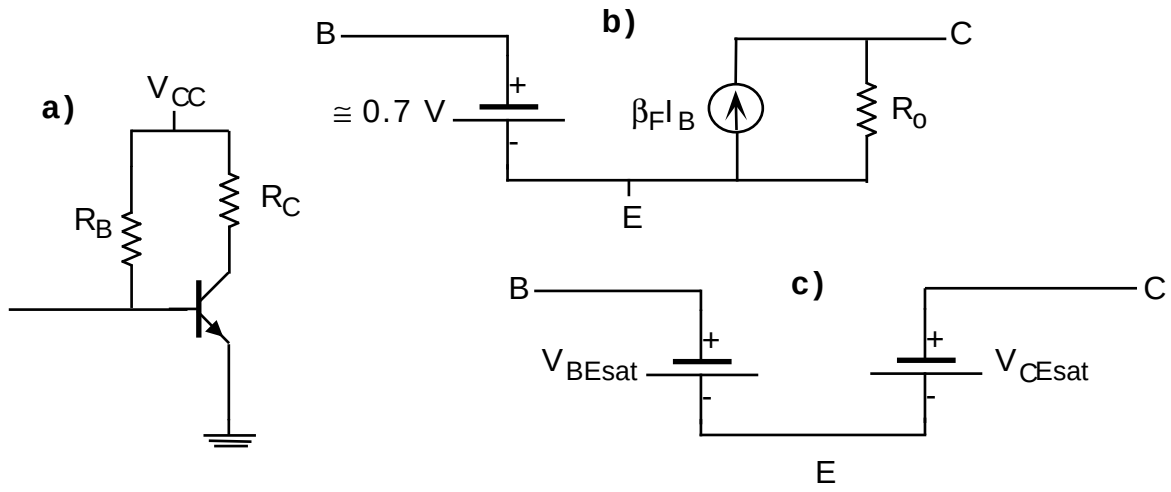

Figura 3.25: *a) Transistor ad emettitore comune; b) Schema equivalente nella regione attiva; c) Schema equivalente nella regione di saturazione.*

Da cui si ricava

$$I_B = \frac{V_{cc} - V_{BE}}{R_B}$$

Se il transistor é nella regione attiva, possiamo risolvere questa approssimativamente questa equazione assumendo $V_{BE} \simeq 0.7\ V$ e ricavare

$$I_B = 31\mu A$$

$$I_c = \beta_F I_B \simeq 3.1mA$$

Scriviamo ora l'equazione di Kirchoff della maglia che contiene R_C:

$$-V_{cc} + I_c R_c + V_{CE} = 0$$

Risolvendo troviamo

$$V_{CE} = 3.8\ V$$

Troviamo quindi che $V_{CE} > V_{CE_{sat}}$, quindi la nostra ipotesi é verificata.
Proviamo invece a verificare lo stato del circuito con un diverso valore di R_B, per esempio

$$R_B = 150\ K\Omega$$

Si troverebbe allora

$$I_B = 62\ \mu A$$

$$I_C = 6.2mA$$

e

$$V_{CE} = V_{CC} - I_c R_c < 0$$

cioe' un risultato fisicamente assurdo. Questo vuol dire che il transistor non é nella regione attiva, bensì in saturazione. Si ha quindi

$$I_B = \frac{V_{CC} - V_{BE_{sat}}}{R_B} = 61\ \mu A$$

$$I_C = \frac{V_{CC} - V_{CE_{sat}}}{R_c} = 4.90\ mA$$

Da questo esempio dovrebbe quindi essere chiaro che lo stato del transistor dipende in modo articolato dal complesso dei parametri circuitali. Se si desidera che il transistor sia nella regione attiva occorre assicurarsi che

$$V_{CE} = V_{CC} - \beta_F(V_{CC} - V_{BE})\frac{R_C}{R_B} >> V_{CE_{sat}}$$

3.7 Utilizzazione del transistor

Il transistor puo' essere utilizzato per costruire amplificatori, cioe' dispositivi lineari in cui la funzione di trasferimento e' maggiore di uno. Consideriamo infatti il circuito di Fig. 3.26a: la tensione e la corrente d'ingresso (base) risulteranno dalla sovrapposizione di un termine costante e di un termine variabile dovuto al segnale, v_s, che si vuole amplificare. In conseguenza la tensione d'uscita sul collettore, v_o, risultera' dalla sovrapposizione di un termine costante e di un termine variabile, che ripete la forma del segnale d'ingresso con ampiezza maggiore. Lo studio degli amplificatori formera' l'oggetto del prossimo Capitolo. Il transistor puo' anche essere utilizzato per realizzare interruttori pilotati (switch).

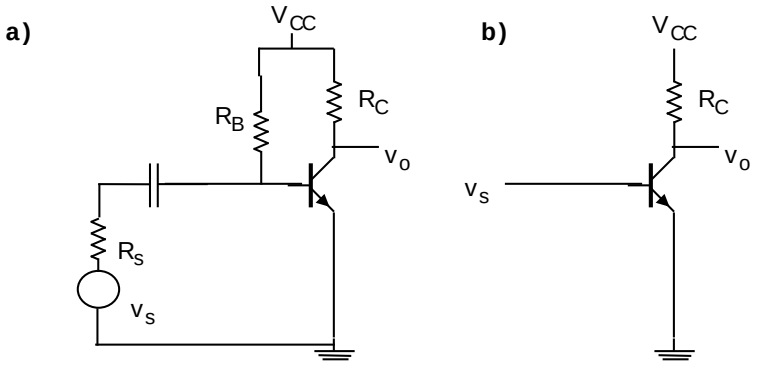

Figura 3.26: *a) Un semplice amplificatore; b) Un interruttore a transistor*

Prendiamo ad esempio il circuito in Fig 3.26b: se la tensione v_s e' negativa il transistor e' in cut-off e $v_o = V_{CC}$, mentre se v_s e' positiva (maggiore di 0.6 V) il transistor e' in saturazione e $v_o \simeq 0.2\ V$. Torneremo su questo dispositivo in uno dei prossimi Capitoli, quando studieremo i circuiti logici.

Capitolo 4

Amplificatori

4.1 Introduzione

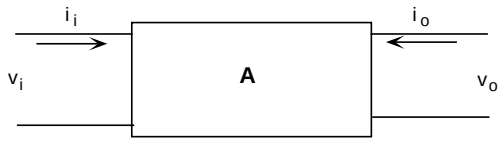

Figura 4.1: *Amplificatore*

Gli amplificatori sono dispositivi di uso corrente in un gran numero di applicazioni; essi costituiranno l'oggetto del presente capitolo. Un amplificatore ideale puo' essere schematizzato come il quadrupolo in Fig. 4.1, dove

$$v_u(t) = A_v v_i(t)$$

$$i_u(t) = A_i i_i(t)$$

almeno uno tra i due parametri A_v ed A_i e' maggiore di uno (ma spesso lo sono entrambi). Poiche' la potenza elettrica e' data dal prodotto vi si ha che la potenza della maglia d'uscita e' maggiore di quella immessa in entrata. E' chiaro quindi che un amplificatore non puo' essere realizzato con soli elementi passivi, ma richiede sorgenti interne di energia. Idealmente A_v ed A_i dovrebbero essere costanti (cioe' indipendenti dalla frequenza): in tal caso il segnale d'uscita ripete fedelmente la forma del segnale d'ingresso. In realta' vedremo che cio' e' irrealizzabile; negli amplificatori reali A_v e A_i sono costanti solo in un intervallo di frequenza finito.

Nel Capitolo precedente abbiamo brevemente visto che si puo' realizzare un amplificatore utilizzando opportunamente un transistor: il segnale da amplificare veniva sovrapposto ad una tensione costante e ne risultava un segnale d'uscita anch'esso sovrapposto ad un livello di tensione diverso da zero.

E' quindi importante usare una notazione corretta per distinguere valori istantanei, valori medi e variazioni attorno ai valori medi, utilizzando opportunamente simboli maiuscoli o minuscoli. Avremo ad esempio

i_A	valore istantaneo della corrente
I_A	valore in quiete della corrente
$i_a = i_A - I_A$	variazione della corrente

L'analisi completa dell'amplificatore a transistor é quindi abbastanza complessa ed é spesso utile separarla in due fasi: ricerca del punto di riposo, e studio delle variazioni delle grandezze intorno al punto di riposo. Per il secondo punto é spesso possibile usare una approssimazione lineare (modello per piccoli segnali).

4.2 Amplificatore ad emettitore comune

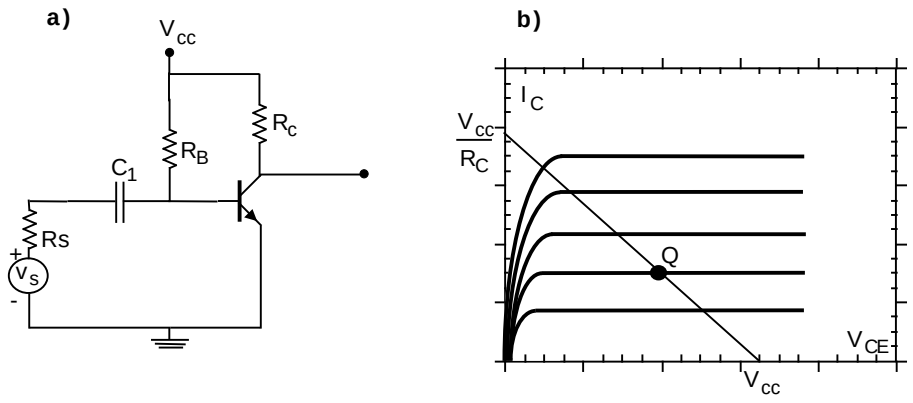

Figura 4.2: *a) Amplificatore ad emettitore comune; b) Caratteristiche di uscita*

Consideriamo il circuito in Fig. 4.2. Il segnale che vogliamo amplificare, v_s, e' applicato alla base del transistor attraverso un capacitore C_1. In questo modo abbiamo *disaccoppiato* in continua l'amplificatore dal generatore di segnale e possiamo determinare il punto di lavoro statico del transistor attraverso le due equazioni:

$$V_{CC} = R_C I_C + V_{CE}$$

$$V_{CC} = R_B I_B + V_{BE}$$

La seconda equazione determina la corrente I_B

$$I_B = \frac{V_{CC} - V_{BE}}{R_B}$$

che puo' essere calcolata molto rapidamente assumendo che $V_{BE} \approx 0.7\ V$ a questo punto e' possibile ricavare I_C e quindi V_{BE} (vedi capitolo precedente), assumendo di conoscere il valore di β_F. Naturalmente avremo scelto in modo opportuno i valori di R_B ed R_C, assicurandoci che il transistor sia nella regione attiva.
Possiamo ora esaminare il contributo del segnale v_s (che immaginiamo essere un'onda sinusoidale con pulsazione ω): se la reattanza di C_1 e' trascurabile alla frequenza del

segnale[1] la sua presenza e' irrilevante ed in uscita avremo una oscillazione sinusoidale della corrente I_C e della tensione V_{CE} attorno al punto di riposo Q. Questa oscillazione ha la stessa frequenza ma un'ampiezza, sia in corrente che in tensione, maggiore di quella del segnale d'ingresso, realizzando quindi una *amplificazione*.

Questo tipo di circuito di polarizzazione, come vedremo in seguito, consente di ottenere una grande amplificazione di tensione e di corrente. Tuttavia esso e' poco adatto a garantire una sufficiente stabilita' al circuito e quindi realizzare un buon amplificatore[2].

Possiamo realizzare un circuito migliore, in termini di stabilita', utilizzando lo schema di Fig. 4.3a, comunemente chiamato amplificatore con rete autopolarizzante.

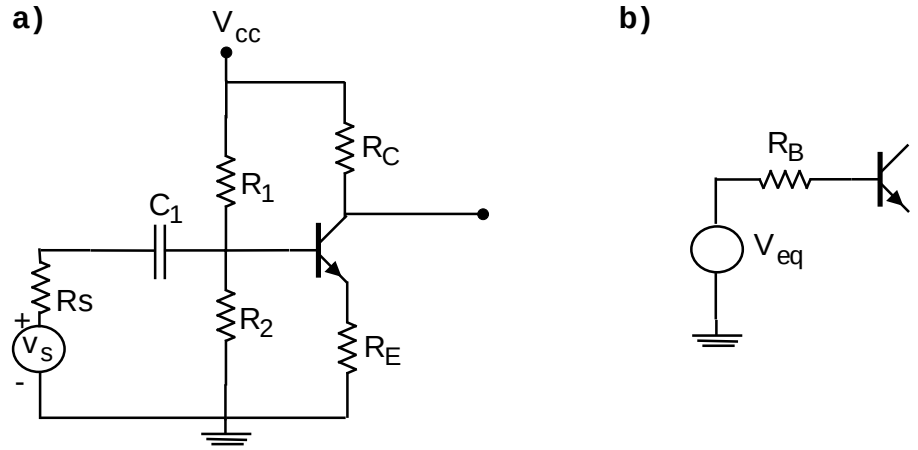

Figura 4.3: *a) Amplificatore con rete autopolarizzante; b) Equivalente di Thevenin della maglia di ingresso*

In questo caso l'equazione della maglia di uscita e':

$$\begin{aligned} V_{CC} &= R_C I_C + V_{CE} + (I_C + I_B) R_E \\ &= (R_C + R_E) I_C + V_{CE} \end{aligned} \tag{4.1}$$

avendo trascurato I_B rispetto ad I_C.
La corrente di base é determinata facendo l'equivalente di Thevenin della maglia d'ingresso (Fig. 4.3b),dove:

$$V_{eq} = \frac{R_2}{R_1 + R_2} V_{CC} \quad R_B = \frac{R_1 R_2}{R_1 + R_2}$$

e si ottiene

$$\begin{aligned} V_{eq} &= R_B I_B + V_{BE} + (I_B + I_C) R_E \\ &= R_B I_B + V_{BE} + I_C R_E \end{aligned} \tag{4.2}$$

[1]Come vedremo meglio in seguito il capacitore introduce un passa-alto nel circuito, cioe' un'attenuazione dei segnali a bassa frequenza. Ma, se lo eliminassimo, la tensione statica della base dipenderebbe anche dal valore di R_s; se R_s fosse molto minore di R_B la base andrebbe a tensione zero, portando in interdizione il transistor.

[2]L'instabilita' e' legata alla variabilita' dei parametri del transistor da un esemplare all'altro, nonche' in funzione della temeratura. I parametri suscettibili di variazione sono β_F,V_{BE} e I_{CB0}.

Ponendo $I_C \simeq \beta_F I_B$ si ottiene

$$I_B = \frac{V_{eq} - V_{BE}}{R_B + \beta_F R_E} \tag{4.3}$$

A questo punto e' possibile ricavare V_{CE} dalla 4.1, sostituendo ancora $I_C \simeq \beta_F I_B$

$$V_{CE} = V_{CC} - (R_C + R_E)\beta_F I_B \tag{4.4}$$

Anche in questo caso, naturalmente, occorre verificare che il transistor sia nella regione attiva e non in saturazione.
Nella realta', come vedremo nel paragrafo 4.6, per progettare realmente un amplificatore questa procedura e' del tutto pleonastica e ci e' servita solo per comprendere il funzionamento del circuito. Questo circuito, come comprenderemo meglio in seguito, ha delle prestazioni molto piu' stabili[3].

Modello ibrido per piccoli segnali

Dobbiamo ora studiare quantitativamente la risposta di un amplificatore per piccoli segnali, cioe' cosa succede quando applichiamo all'ingresso una tensione sinusoidale, v_s, piccola rispetto al valore statico. E' ovvio che dovremo aspettarci di avere piccole variazioni di tutte le altre grandezze elettriche attorno al punto di lavoro statico determinato in precedenza.

Converra' fare per ora due ipotesi:
1) C_1 ha reattanza trascurabile alla frequenza di v_s;
2) Le reattanze interne al transistor, dovute alle capacita' delle giunzioni, sono anch'esse trascurabili.
Con queste ipotesi possiamo schematizzare il transistor come un quadrupolo come in Fig. 4.4a, al cui ingresso si applica una tensione V_I. $v_1 = v_{BE}$, $i_1 = i_B$, $v_2 = v_{CE}$, $i_2 = i_C$.

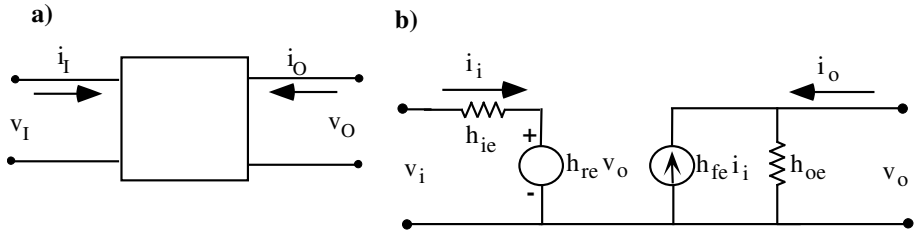

Figura 4.4: *a) Quadrupolo equivalente al transistor; b) Schema dettagliato*

Tensioni e correnti all'ingresso e all'uscita sono legate da relazioni funzionali; possiamo scegliere i_I e v_O come variabili indipendenti e scrivere quindi

$$\begin{aligned} v_I &= f_1(i_I, v_O) \\ i_O &= f_2(i_I, v_O) \end{aligned}$$

[3]Le considerazioni sulla *stabilita'* di un circuito possono essere fatte in modo rigoroso, introducendo i cosidetti fattori di stabilita', cioe' in sostanza studiando le derivate di I_C in funzione dei parametri del transistor.

Se i_I, i_O, v_I, v_O variano poco attorno al valore statico, posso sviluppare in serie di Taylor attorno ad esso, arrestando lo sviluppo al primo ordine:

$$\begin{aligned} \Delta v_I &= \left.\frac{\partial f_1}{\partial i_I}\right|_{v_O=cost} \Delta i_I + \left.\frac{\partial f_1}{\partial v_O}\right|_{i_I=cost} \Delta v_O \\ \Delta i_O &= \left.\frac{\partial f_2}{\partial i_I}\right|_{v_O=cost} \Delta i_I + \left.\frac{\partial f_2}{\partial v_O}\right|_{i_I=cost} \Delta v_O \end{aligned}$$

cioe', in sostanza

$$\begin{cases} v_i = h_{11} i_i + h_{12} v_o \\ i_o = h_{21} i_i + h_{22} v_o \end{cases}$$

dove abbiamo usato la notazione abbreviata per le variazioni ed indicato con h_{ij} le derivate parziali. Questi ultimi si chiamano parametri ibridi del transistor, perche' hanno dimensioni fisiche diverse. Nella configurazione ad emettitore comune essi vengono normalmente indicati con i nomi h_{ie} , h_{re} , h_{fe} , h_{oe}. Possiamo quindi scrivere:

$$v_i = h_{ie} i_i + h_{re} v_o \tag{4.5}$$
$$i_o = h_{fe} i_i + h_{oe} v_o \tag{4.6}$$

Queste sono le equazioni di un quadrupolo come quello in Fig. 4.4b (da cui si comprende il significato fisico dei 4 parametri), che rappresenta quindi lo schema equivalente del transistor per piccoli segnali. Quindi la conoscenza dei 4 parametri consente di studiare le prestazioni dell'amplificatore, come vedremo tra poco. In linea di principio il loro valore dipende dal punto di lavoro scelto, ma in realta' essi sono abbastanza costanti all'interno della regione attiva. I fabbricanti di transistors forniscono, tra le varie specifiche, il valore di questi parametri per ogni tipo che essi producono; tuttavia ci sono forti variazioni da un esemplare all'altro (ed anche variazioni con la temperatura). Per un transistor tipico i loro valori sono:

$$\begin{aligned} h_{fe} &\quad (50 \div 400) \\ h_{ie} &\quad (1 \div 10)\ Kohm \\ h_{re} &\quad \sim 10^{-4} \\ h_{oe} &\quad \sim 10^{-4}\ ohm^{-1} \end{aligned}$$

Si noti come i parametri h_{re} ed h_{oe} sono in genere molto piccoli (vedremo che in molti casi il loro effetto puo' essere trascurato): da un punto di vista fisico sono dovuti al cosidetto effetto Early (modulazione dello spessore effettivo della base). E' interessante notare che la nostra deduzione dei parametri h non si applica solo alla configurazione CE, ma e' del tutto generale: per ogni configurazione e' possibile definire 4 parametri con cui si schematizza il circuito equivalente per piccoli segnali. Si hanno quindi complessivamente 12 parametri:

h_{ie} h_{fe} h_{re} h_{oe} a emettitore comune

h_{ib} h_{fb} h_{rb} h_{ob} a base comune

h_{ic} h_{fc} h_{rc} h_{oc} a collettore comune

Naturalmente essi non sono indipendenti, e conoscendo i 4 parametri di una configurazione si possono ricavare gli altri.

Proviamo ad applicare questo modello al circuito di Fig. 4.2. Per lo studio dei piccoli segnali esso diviene equivalente allo schema di Fig. 4.5a.

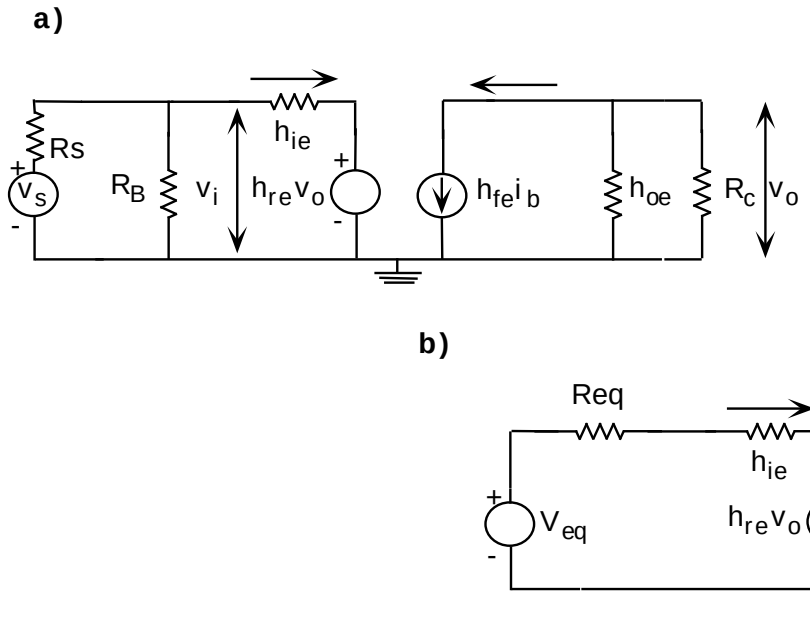

Figura 4.5: *a) Circuito equivalente dell'amplificatore CE; b) Circuito equivalente della maglia di ingressso*

Deve essere chiaro il significato di questo schema: esso rappresenta il circuito ai fini dello studio delle variazioni delle grandezze elettriche attorno al loro valore statico (cioe' attorno al valore che assumono in assenza del segnale d'ingresso). Percio' si comprende come i resistori R_C ed R_B (che nel circuito reale hanno un estremo collegato a V_{CC}) appaiano in questo schema collegate a massa: il punto a tensione V_{CC} e' infatti un punto a variazione zero.

Come abbiamo detto il termine h_{re} e' molto piccolo, percio' esso puo' spesso essere trascurato. Inoltre la resistenza d'uscita $1/H_{oe}$ e' in genere molto grande rispetto al resistore di carico R_C. Percio' e' generalmente sufficiente utilizzare un modello approssimato, il cosidetto modello ibrido semplificato (Fig. 4.6).

Allora lo schema equivalente del nostro amplificatore diviene quello della Fig. 4.7

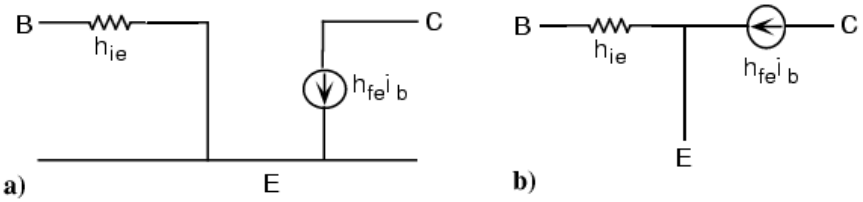

Figura 4.6: *a): modello ibrido semplificato; b): un modo piu' semplice di ridisegnare lo stesso circuito. Il lettore non dovrebbe avere difficolta' a convincersi che i due schemi sono totalmente equivalenti*

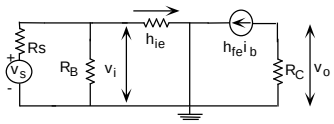

Figura 4.7: *Amplificatore CE con il modello semplificato*

Le prestazioni di un amplificatore sono completamente caratterizzate una volta note l'amplificazione di corrente A_i, l'amplificazione di tensione A_v, la resistenza d'ingresso R_i ed infine la resistenza d'uscita R_o. Per definizione

$$A_i \equiv \frac{i_o}{i_i}$$

dove i_o e' la corrente che circola nella corrente di uscita (convenzionalmente presa uscente), mentre i_i e' la corrente che entra nel transistor. Si ha allora:

$$A_i = -\frac{i_c}{i_b} = -h_{fe}$$

L'amplificazione di tensione é data da

$$A_v \equiv \frac{v_o}{v_i}$$

Poiche'

$$\begin{aligned} v_o &= -i_c R_C \\ &= -h_{fe} i_b R_C \\ v_i &= h_{ie} i_b \end{aligned}$$

si ha

$$A_v = \frac{-h_{fe} R_C}{h_{ie}}$$

In realtá si devono considerare

$$A_i' \equiv \frac{i_o}{i_s} \quad e \quad A_v' \equiv \frac{v_o}{v_s}$$

cioe' le amplificazioni rispetto alla tensione e corrente erogate dal generatore; esse si possono pero' ricavare facilmente da A_i e A_v. Facciamo infatti l'equivalente di Thevenin della maglia d'ingresso (Fig. 4.5b):

$$R_{eq} = R_S || R_B \quad v_{eq} = \frac{R_B}{R_S + R_B} v_s$$

$$A_v' = \frac{v_o}{v_s} = \frac{v_o}{v_{eq}} \frac{v_{eq}}{v_s}$$

ora, ripetendo la procedura giá vista per A_v

$$v_o = A_i i_i R_C$$

$$v_{eq} = (R_{eq} + h_{ie} + h_{re} R_C A_i) i_i$$

e quindi

$$A'_v = \frac{R_B}{R_S + R_B} \frac{A_i R_C}{R_{eq} + h_{ie} + h_{re} R_C A_i} \approx \frac{R_B}{R_S + R_B} \frac{A_i R_C}{R_{eq} + h_{ie}}$$

La resistenza d'ingresso e' definita come

$$R_i \equiv \frac{v_i}{i_i}$$

dove

$$v_i = h_{ie} i_b$$

e quindi

$$R_i = h_{ie}$$

Si verifica facilmente che

$$A_v = A_i \frac{R_C}{R_i}$$

questa relazione e'valida sempre per un quadrupolo: l'amplificazione di tensione e' data dall'amplificazione di corrente moltiplicata per il rapporto tra resistenza di carico e resistenza d'ingresso.
La resistenza d'ingresso e' un parametro importante di un amplificatore: un valore elevato significa che l'amplificatore assorbe poca corrente (e quindi poca potenza) dal generatore d'ingresso; questa e' in genere (anche se non sempre) una caratteristica positiva.
Sotto questo profilo R_B andrebbe incluso nella resistenza d'ingresso, che diventerebbe

$$R'_i = R_i || R_B$$

La resistenza d'uscita e' in sostanza la resistenza equivalente che si ricaverebbe applicando al circuito il teorema di Thevenin. Un metodo pratico di ricavarla consiste nel cortocircuitare (idealmente) tutti i generatori indipendenti presenti nel circuito, togliere la resistenza R_C ed immaginare di applicare una tensione v ai morsetti d'uscita. Il rapporto

$$\frac{1}{R_o} \equiv \frac{i'}{v}$$

dove i' e' la corrente risultante fornisce la conduttanza d'uscita.
Nel caso in esame, una volta enucleata la resistenza R_C, la maglia d'uscita del circuito e' costituita solamente da un generatore ideale di corrente. Pertanto la resistenza d'uscita (intesa nel senso di Norton) e' infinita [4].
In conclusione un amplificatore e' schematizzabile come un quadrupolo (Fig 4.8), con due possibili rappresentazioni della maglia d'uscita, chiaramente equivalenti tra loro.

[4] Come dovrebbe essere chiaro, la resistenza d'uscita e' data, con migliore approssimazione, dal termine $1/h_{oe}$, fin qui trascurato

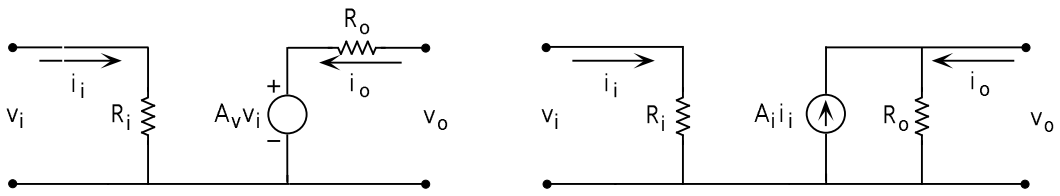

Figura 4.8: *Due possibili rappresentazioni di un amplificatore*

Configurazione ad Emettitore Comune con rete autopolarizzante

Studiamo ora le prestazioni dell'amplificatore schematizzato nella Fig. 4.3. Trascurando il capacitore di blocco C_1 ed utilizzando il modello ibrido semplificato otteniamo lo schema di Fig. 4.9.

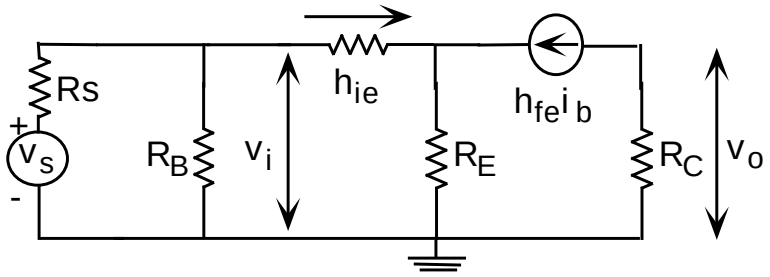

Figura 4.9: *Schema equivalente dell'amplificatore CE con rete autopolarizzante*

$$A_i = -h_{fe}$$

$$R_i = \frac{v_i}{i_b} = \frac{h_{ie} i_b + (-i_e R_E)}{i_b}$$

poiché

$$i_e = -(i_b + i_c) = -(i_b + h_{fe} i_b) = -(1 + h_{fe}) i_b$$

si ha

$$R_i = \frac{[h_{ie} + (1 + h_{fe}) R_E] i_b}{i_b} = h_{ie} + (1 + h_{fe}) R_E$$

cioé la resistenza d'ingresso é molto elevata.
(In realtá la vera resistenza d'ingresso deve tener conto del parallelo di R_1 e R_2).

$$A_v = A_i \frac{R_L}{R_i} = \frac{-h_{fe}}{h_{ie} + (1 + h_{fe}) R_E} R_L$$

poiche' $(1 + h_{fe}) R_E \gg h_{ie}$,

$$A_v \approx -\frac{R_L}{R_E}$$

$$R_o = \infty$$

Come si vede l'amplificazione di tensione é indipendente dai parametri del transistor, ma dipende solo dal rapporto tra due resistenze: e' quindi particolarmente stabile. Questo rapporto non puo' tuttavia essere troppo grande, per ragioni pratiche; quindi questo circuito non consente di ottenere grandi amplificazioni[5].

Il significato del modello ibrido semplificato

Con questo modello noi in sostenza rappresentiamo il transistor come un elemento a tre terminali (vedi Fig. 4.7b), in cui la relazione tra le correnti e' fissata:

$$i_c = h_{fe} i_b \tag{4.7}$$
$$i_e = (1 + h_{fe}) i_b \tag{4.8}$$
$$0 = i_b + i_c + i_e \tag{4.9}$$

Inoltre, *guardando* dal terminale di base, il transistor offre una resistenza h_{ie} verso l'emettitore. Non deve quindi stupire il fatto che questa schematizzazione possa essere lecitamente usata anche in configurazioni diverse da quella ad emettitore comune.
Inoltre, e' bene notare che h_{fe} rappresenta, per i piccoli segnali, cio' che il parametro β_F rappresentava per le correnti continue[6].

Il modello di Giacoletto

Un approccio piú fisico e meno formale suggerisce, tenendo conto del modello a 2 diodi, di schematizzare il transistor come in Fig. 4.10a.

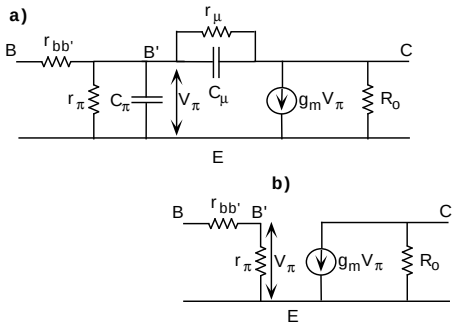

Figura 4.10: *a) Modello di Giacoletto; b) Modello semplificato a bassa frequenza*

Questo schema prende il nome di modello a π (o di Giacoletto) per la configurazione ad Emettitore Comune. Questo modello, come vedremo, ci consentira' di studiare il comportamento del transistor anche ad alte frequenze perche' tiene correttamente conto delle reattanze parassite delle giunzioni base-emettitore e base-collettore A media e bassa frequenza esse sono in genere trascurabili, quindi esso si riduce allo schema di Fig. 4.10, dove abbiamo trascurato anche r_μ che e' in genere molto grande.

[5] questo verra' compreso meglio quando progetteremo concretamente un amplificatore di questo tipo.
[6] Di fatto il parametro β_F e' generalmente indicato dai costruttori con il simbolo h_{FE}.

E' interessante confrontare questo schema con il modello ibrido semplificato. Si vede subito che

$$\frac{1}{R_o} = h_{oe}$$

$$r_{bb'} + r_\pi = h_{ie}$$

La resistenza $r_{bb'}$ viene introdotta per tener conto del fatto che il terminale di base presenta spesso una resistenza di natura puramente ohmica, non trascurabile, distinta dalla resistenza dinamica del diodo base-emettitore[7]. Tuttavia $r_{bb'}$ puo' essere trascurata, edallora si ha

$$r_\pi = h_{ie}$$

Inoltre

$$v_\pi = r_\pi i_b$$

da cui si ricava

$$g_m r_\pi = h_{fe}$$

Il parametro

$$g_m \equiv \frac{i_c}{v_\pi}$$

si chiama transconduttanza del transistor. Possiamo quindi scrivere

$$g_m = \frac{h_{fe}}{h_{ie}}$$

Se continuiamo a trascurare $r_{bb'}$ abbiamo che

$$g_m = \left.\frac{\Delta i_C}{\Delta v_{BE}}\right|_{v_{CE}=cost} = \left.\frac{\partial i_C}{\partial v_{BE}}\right|_{v_{ce}=0}$$

La condizione $v_{ce} = 0$ implica che R_o é cortocircuitata; allora $g_m v_\pi$ coincide con i_c.

D'altra parte, poiché $i_C = -\alpha_F i_E$,

$$g_m = -\alpha_F \left.\frac{\partial i_E}{\partial v_{BE}}\right|_{v_{ce}=0}$$

Cioe' abbiamo espresso g_m in termini delle variabili elettriche del diodo base-emettitore. Ma, in un diodo polarizzato direttamente

$$g_D = \frac{di_D}{dv_D} \sim \frac{I}{V_T}$$

(vedi la 3.1); quindi possiamo dedurre che

$$g_m = \frac{\alpha_F |I_E|}{V_T} = \frac{|I_C|}{V_T}$$

[7]Questo e' dovuto proprio alla geometria della base, generalmente molto piccola rispetto a emettitore e collettore; la resistenza ohmica complessiva dovuta al contatto tra il cristallo e il conduttore metallico puo' quindi essere non trascurabile.

dove abbiamo preso il modulo delle correnti per tener conto del fatto che per un transistor *npn* $V_{BE} = v_D$ e $I_E = -i_D$, mentre per un transistor *pnp* $V_{BE} = -v_D$ e $I_E = i_D$. A temperatura ordinaria si ha allora

$$g_m \approx \frac{|I_C|(mA)}{26(mV)} \tag{4.10}$$

Si noti che la temperatura del transistor e' in genere piu' alta della temperatura ambiente perche' in esso si ha una dissipazione di potenza dovuta al passaggio della corrente, quindi non e' facilmente misurabile. Tuttavia, entro una certa approssimazione, la relazione 4.10 ci da una informazione quantitativa importante, legando l'amplificazione di tensione alla corrente statica di collettore.
Prendiamo ad esempio l'amplificatore ad Emettitore Comune (non autopolarizzante): tenendo conto di quanto visto qui possiamo scrivere

$$A_v = -\frac{h_{fe}}{h_{ie}} R_C = -g_m R_C \simeq -\frac{|I_C|}{V_T} R_C \tag{4.11}$$

Quindi le prestazioni dinamiche possono essere in qualche misura predette sulla base del valore di I_C definito nel progetto.

In conclusione, come era lecito attendersi, i due modelli sono del tutto equivalenti se si trascurano le reattanze interne del transistor. Tuttavia lo schema di Giacoletto ci ha consentito di comprendere meglio il significato fisico dei parametri e le loro relazioni.

Amplificatore ad Emettitore Comune con capacita' di emettitore

Dal confronto dei due precedenti amplificatori si vede che la rete autopolarizzante produce una drastica riduzione dell'amplificazione di tensione, mentre garantisce una maggiore stabilita' del punto di lavoro. Possiamo modificare l'amplificatore con rete autopolarizzante in modo da conservarne la stabilita' ed aumentarne l'amplificazione, aggiungendo una capacita' opportuna sull'emettitore (Fig. 4.11a).

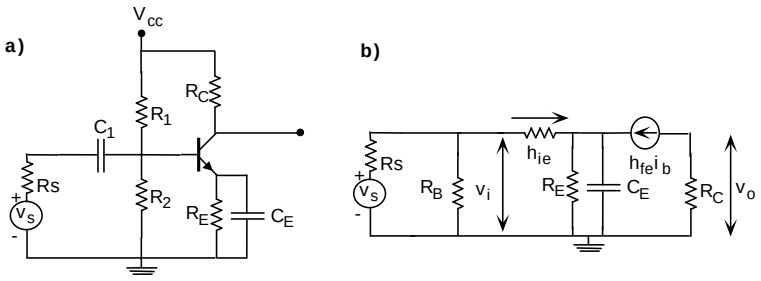

Figura 4.11: *a) Amplificatore con capacita' di emettitore; b) Schema equivalente per piccoli segnali*

Il condensatore C_E e' ininfluente ai fini della polarizzazione, quindi non modifica il punto di lavoro e la sua stabilita'. Invece a frequenze abbastanza alte la sua reattanza e' piccola rispetto ad R_E, per cui l'emettitore viene in pratica cortocircuitato verso la massa, e le prestazioni dell'amplificatore tornano ad essere quelle dell'amplificatore CE senza rete autopolarizzante. Questo schema e' comunemente utilizzato quando si voglia avere una alta amplificazione di tensione; C_E va scelto in modo che la sua reattanza sia trascurabile nell'intervallo di frequenza che vogliamo amplificare.

4.3 Amplificatore a collettore comune

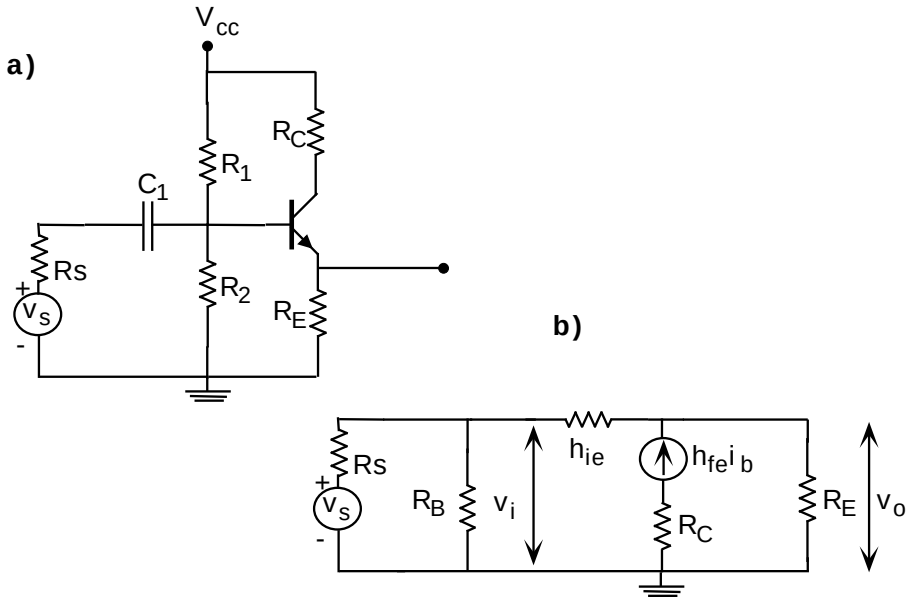

Figura 4.12: *a) Amplificatore a collettore comune; b) Schema equivalente per piccoli segnali*

Nella Fig. 4.12a e' mostrato un esempio di amplificatore a collettore comune. Esso e' comunemente chiamato *emitter follower*, poiche' come vedremo, la tensione di emettitore, cioe' l'uscita, *segue* la tensione d'ingresso. Possiamo studiare questo circuito utilizzando il modello ibrido semplificato, ottenendo quindi il circuito di Fig. 4.12b. Si ha

$$i_e = -(1 + h_{fe})i_b$$

$$A_i \equiv \frac{-i_e}{i_b} = 1 + h_{fe}$$

$$R_i \equiv \frac{v_i}{i_b} = \frac{h_{ie}i_b - R_E i_c}{i_b} = \frac{h_{ie}i_b + A_i R_E i_b}{i_b}$$

$$R_i = h_{ie} + A_i R_E = h_{ie} + (1 + h_{fe})R_E$$

$$A_v \equiv \frac{v_u}{v_i} = \frac{-i_e R_E}{h_{ie}i_b + A_i R_E i_b} = -\frac{i_e R_E}{i_b R_i} = A_i \frac{R_E}{R_i}$$

Ora,

$$R_E A_i = R_i - h_{ie}$$

quindi

$$A_v = \frac{R_i - h_{ie}}{R_i} \approx 1$$

L'impedenza di uscita si puó calcolare dal rapporto tra la tensione v idealmente applicata sui morsetti d'uscita e corrente i' che ne deriva (cortocircuitando il generatore d'ingresso). Si ha

$$v = -h_{ie}i_b - R'_S i_b R'_S = R_S || R_B$$

$$i' = -(1 + h_{fe})i_b$$

$$R_o = \frac{h_{ie} + R'_S}{1 + h_{fe}}$$

Naturalmente l'effettiva impedenza d'uscita comprende peró anche R_E, cioé

$$R'_o = R_o || R_E$$

Inoltre non abbiamo tenuto conto della presenza di R_B; bisognerebbe quindi calcolare A'_i,R'_i,A'_v.

Si noti che la resistenza R_C sul collettore e' assolutamente ininfluente ai fini delle prestazioni del circuito (entra invece nella determinazione del punto di lavoro). Si puo' quindi costruire il circuito senza di essa.

4.4 Amplificatore a base comune

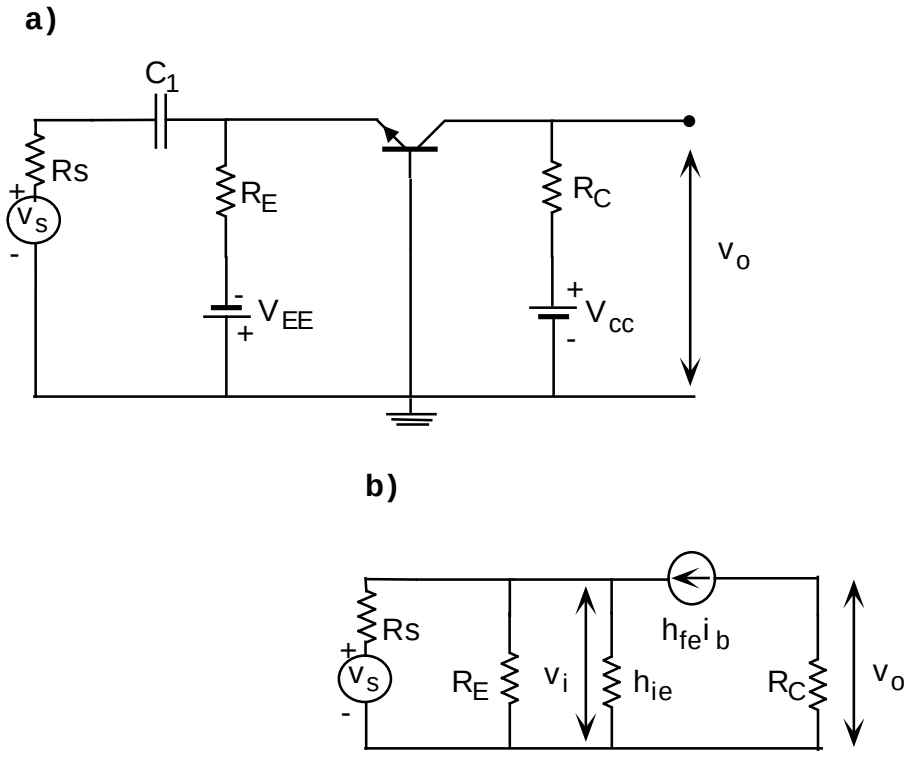

Figura 4.13: *a) Amplificatore a base comune; b) Schema equivalente per piccoli segnali*

Nella Fig. 4.13 e' mostrato un amplificatore a base comune ed il suo schema equivalente. Si ha

$$A_i = -\frac{i_c}{i_e} = \frac{h_{fe} i_b}{-(1 + h_{fe})i_b} = -\frac{h_{fe}}{1 + h_{fe}} \approx 1$$

$$R_i = \frac{v_i}{i_e} = \frac{-h_{ie} i_b}{-(1 + h_{fe})i_b} = \frac{h_{ie}}{1 + h_{fe}}$$

$$A_v = A_i \frac{R_C}{R_I} = \frac{1 + h_{fe}}{h_{ie}} R_C$$
$$R_o = \infty$$

4.5 Riepilogo

E' utile riepilogare le caratteristiche degli amplificatori fin qui visti.

	CE	CE^*	CC	CB
A_i	$-h_{fe}$	$-h_{fe}$	$1 + h_{fe}$	≈ 1
R_i	h_{ie}	$h_{ie} + (1 + h_{fe})R_E$	$h_{ie} + (1 + h_{fe})R_E$	$\frac{h_{ie}}{1 + h_{fe}}$
A_v	$-h_{fe}\frac{R_C}{h_{ie}}$	$-\frac{R_C}{R_E}$	≈ 1	$\frac{1 + h_{fe}}{h_{ie}} R_C$
R_o	∞	∞	$\frac{h_{ie} + R'_S}{1 + h_{fe}}$	∞

Con CE^* abbiamo indicato la configurazione ad emettitore comune con rete autopolarizzante.

Come si vede le varie configurazioni offrono al progettista la possibilita' di scegliere tra varie caratteristiche: alta o bassa impedenza d'ingresso; alta o bassa impedenza d'uscita; amplificazione solo di corrente, solo di tensione, o entrambe.

4.6 Progettare un amplificatore

Nei paragrafi precedenti abbiamo visto quali sono le prestazioni di alcuni amplificatori a transistor e come essi dovrebbero teoricamente essere progettati e costruiti. Nella pratica le cose sono un po' diverse quindi e' opportuno, attraverso degli esempi concreti, imparare a progettare un amplificatore. Infatti, come abbiamo gia' detto, i parametri fondamentali del transistor variano molto da un esemplare all'altro, e le curve caratteristiche fornite dal costruttore vanno intese come indicazioni di massima del comportamento di quel modello. Nel seguito vedremo due esempi di progetto, il primo di un amplificatore CE con rete autopolarizzante, il secondo di un inseguitore di tensione.

Amplificatore ad Emettitore Comune

Si vuole costruire un amplificatore utilizzando un transistor 2N2222A. Dai fogli illustrativi si vede che il costruttore indica un valore di h_{fe} compreso tra 50 e 375. Come si vede non ha nessun senso basare un progetto sull'ipotesi che il parametro h_{fe} abbia un valore noto e preciso. Ci baseremo invece solo sulla ipotesi (molto ragionevole) che la differenza di potenziale V_{BE} sia approssimativamente costante (e pari a 0.7 V) quando il transistor e' nella regione attiva. Questo significa che se fissiamo la tensione di base anche la tensione di emettitore e' fissa e, come vedremo, diverra' facilissimo scegliere il punto di lavoro del transistor.

Si richieda ad esempio che il dispositivo abbia una amplificazione $A_V = 10$ e che il segnale d'ingresso da amplificare non superi 200 mV.
Si deve anzitutto decidere la tensione di emettitore, V_E, a cui si vuole lavorare. Come? E' bene scegliere un valore basso, ma non troppo, perche' l'emettitore *segue* la base quando mandiamo un segnale e quindi la V_E deve poter variare attorno al suo valore statico di una quantita' pari al massimo segnale d'ingresso, senza avvicinarsi troppo allo zero. In caso contrario il segnale d'uscita verrebbe distorto. Possiamo ad esempio porre $V_E = 0.5\ V$.
Ora, poiche' l'amplificazione deve essere 10, la caduta di tensione sulla resistenza di collettore R_C e' 10 volte V_E. In sostanza, dall'equazione della maglia d'uscita si ha:

$$V_{CC} = V_E + 10\,V_E + V_{CE}$$

Possiamo ora decidere quanto deve essere V_{CE}. Come? Bisogna tener conto che V_{CE} deve poter variare attorno al suo valore statico, con una escursione ΔV_{CE} pari al massimo segnale d'ingresso moltiplicato per l'amplificazione, senza uscire dalla regione lineare. In sostanza V_{CE} deve restare sempre positiva e abbastanza lontana da zero. Nel nostro caso $\Delta V_{CE} = 2\ V$ e, per tenerci larghi sceglieremo $V_{CE} = 6\ V$; quindi dovremo porre $V_{CC} = 11.5$ (metteremo ovviamente $V_{CC} = 12\ V$!). A questo punto e' tutto ben determinato e restano da scegliere i valori effettivi delle resistenze. La scelta di R_E (e quindi di R_C) agisce *solo* sulla corrente I_C che scorre nel transistor: poniamo ad esempio $R_E = 100\ \Omega$: avremo di conseguenza $R_C = 1\ k\Omega$ ed una corrente di 5 mA, che costituisce un valore *ragionevole*.
Infine, dobbiamo progettare il partitore di base: vogliamo una tensione V_B pari a 1.2 V (0.7 piu' alta dell'emettitore), e vogliamo che essa resti fissa, indipendentemente dalla corrente di base. Questo si puo' ottenere avendo una corrente I_P del partitore molto maggiore della corrente che entra nella base. Nel caso piu' pessimistico ($h_{fe} = 50$) avremo:

$$I_B = \frac{I_C}{h_{fe}} = \frac{5\ mA}{50} = 100\ \mu A$$

Saremo quindi tranquilli se sceglieremo $I_P = 1\ mA$, da cui

$$R_1 + R_2 = \frac{V_{CC}}{I_P} = \frac{12\ V}{1\ mA} = 12\ k\Omega$$

D'altra parte si deve avere

$$\frac{R_2}{R_1 + R_2} V_{CC} = 1.2\ V$$

e questo ci porta a determinare

$$R_1 = 10800\ \Omega$$

$$R_2 = 1200\ \Omega$$

Naturalmente potremo variare leggermente questi valori e sceglierne di piu' comodi senza alterare in nulla la sostanza del circuito.
Controlliamo ora che tutte le dissipazioni siano al di sotto dei limiti:
il prodotto $I_C V_{CE}$ risulta di 30 mW (il limite per il 2N2222A e' 500 mW);
R_E dissipa 2.5 mW, mentre R_C ne dissipa 25 (non ci sono problemi usando resistori da 1/4 W); il partitore di base dissipa complessivamente 12 mW quindi possiamo approvare le scelte fatte.

L'ultimo controllo e' sul valore di R_E: siamo sicuri che il prodotto $h_{fe}R_E$ sia molto maggiore della resistenza h_{ie}? Questo e' importante per la *stabilita'* del circuito ma anche perche' altrimenti non e' piu' vero che l'amplificazione e' 10! Dai fogli illustrativi vediamo che h_{ie} (come del resto h_{fe}) dipende da I_C e purtroppo sono riportati solo i valori relativi a $I_C = 1\ mA$ (h_{ie} oscilla tra 2 e 8 $k\Omega$) e $I_C = 10\ mA$ ($0.25 \div 1.25\ k\Omega$). Interpolando rozzamente saremmo portati a dire che nel caso peggiore ($h_{fe} = 50$ e h_{ie} di qualche $k\Omega$) la nostra richiesta non e' rispettata; tuttavia occorre ricordare quanto abbiamo scoperto con il modello di Giacoletto: il rapporto h_{fe}/h_{ie} dipende solo da I_C e dalla temperatura. Cioe' i due parametri vanno di pari passo: se uno e' grande anche l'altro e' grande, e viceversa, quindi il caso pessimistico e' assolutamente irrealistico e possiamo essere ragionevolmente tranquilli della bonta' delle nostre scelte.

Ricapitoliamo allora la procedura da seguire:

- esaminare le richieste, cioe' amplificazione e dinamica del segnale d'ingresso;
- scegliere V_E, V_{CE} e quindi V_{CC};
- scegliere I_C e quindi R_E ed R_C;
- progettare il partitore di base in modo che $V_P = V_E + 0.7$ e I_P molto piu' grande di I_C/h_{fe} anche nel peggiore dei casi;
- verificare che tutte le dissipazioni siano entro i limiti e che il circuito soddisfi il requisito di stabilita'.

Il lettore e' invitato a *lavorare* con questo esempio, modificando le scelte fatto e cercando di capire le implicazioni conseguenti. Per esempio, che succede se raddoppiamo i valori di R_E ed R_C? Apparentemente miglioriamo in questo modo la *stabilita'*, ma, attenzione, in questo modo la I_C si dimezza e quindi diminuisce il rapporto h_{fe}/h_{ie}, il che va nel verso contrario. Che succede aumentando V_{CC}? Nulla; si puo' ovviamente avere una V_{CC} piu' grande, una V_{CE} piu' grande, e, magari, anche una I_C piu' grande, variando opportunamente tutti i parametri, per ottenere un circuito che funziona altrettanto bene. Ma, a questo punto, tutte le dissipazioni di potenza aumenteranno inutilmente! Inoltre dover avere grandi valori per le tensioni di alimentazione non e' comodo: il buon progettista cerchera' di usare valori adeguati alle reali necessita' del circuito. Come regola empirica si puo' dire che e' conveniente avere $V_{CE} \sim V_{CC}/2$: in questo modo si puo' avere la massima escursione possibile per il segnale d'uscita a parita' di alimentazione e, in assoluto, la scelta di V_{CE} dipende dalla dinamica d'uscita desiderata. Il lettore provi ad aumentare via via l'amplificazione, tenendo fissa V_{CC}, e vedra' che sara' costretto, per avere un segnale d'uscita non distorto, a diminuire il segnale d'ingresso.

Inseguitore di tensione

Con lo stesso transistor costruiamo un emitter follower. Questo circuito non serve ad amplificare in tensione, bensi' e' normalmente usato per fornire una grande corrente d'uscita, avendo una bassissima resistenza d'uscita. Dobbiamo quindi aspettarci un segnale d'ingresso gia' abbastanza ampio.

Immaginiamo di lavorare con segnali di ingresso con ampiezza fino a 2 V. E' chiaro che ora bisogna avere una tensione sulla base piu' alta di prima: per massimizzare l'escursione ci converra' porre

$$V_B = V_{CC}/2$$

e quindi V_E verra' automaticamente 0.7 V piu' bassa. Non mettiamo resistenza sul collettore, per cui:

$$V_{CC} = V_{CE} + V_E$$

$V_{CC} = 12\ V$ e' anche qui una buona scelta e, conseguentemente

$$\begin{aligned} V_B &= 6\ V \\ V_E &= 5.3\ V \\ V_{CE} &= 6.7\ V \end{aligned}$$

La scelta di R_E determinera' la corrente: per esempio, con $R_E = 1\ k\Omega$, si ha $I_C = 5.3\ mA$, valore del tutto ragionevole per questo transistor.
E' del tutto ovvio come progettare il partitore di ingresso: con questa scelta di I_C (che' e' uguale al caso dell'amplificatore CE visto in precedenza), si ha:

$$R_1 + R_2 = 12\ k\Omega$$

ma ora porremo $R_1 = R_2 = 6\ k\Omega$.
Lasciamo al lettore verificare che le dissipazioni sono entro i limiti e che il criterio di *stabilita'* e' rispettato ampiamente.

Amplificatore a due stadi

Possiamo accoppiare, come spesso si fa, l'emitter follower all'amplificatore CE, per realizzare un amplificatore a 2 stadi caratterizzato da una amplificazione complessiva pari a 10 e bassa resistenza d'uscita. Il modo piu' inelegante di farlo e' di interporre tra l'uscita del primo stadio e l'ingresso dell'emitter follower un condensatore di disaccoppiamento (Fig. 4.14a). Capiremo meglio nel prossimo paragrafo che i condensatori sono nocivi, in quanto introducono limitazioni nella risposta a basse frequenze. Possiamo chiederci se e' lecito farne a meno e accoppiare direttamente il collettore del primo stadio alla base del secondo (Fig. 4.14b), addirittura eliminando il partitore d'ingresso dell'emitter follower. Che succede in questo caso? Il collettore del transistor 1 e', nel nostro progetto a 5.5 V e questo e' il valore cui si porta la base del transistor 2. E' un valore ragionevole? Ovviamente si, perche' avremo una V_E di 4.8 V, quindi una corrente di 4.8 mA, poco diversa dal valore precedente, e saremo ancora ampiamente nei limiti della dinamica d'uscita richiesta. Dobbiamo solo verificare che il primo stadio non venga *perturbato* dal secondo. Ora, l'impedenza d'uscita del primo stadio, R_{o1} e', come sappiamo, sostanzialmente data da R_C. Invece l'impedenza d'ingresso dell'emitter follower, R_{i2} e' data da $h_{fe}R_E$. Nel nostro caso abbiamo quindi

$$R_{o1} = 1\ k\Omega$$

$$R_{i2} = 50\ k\Omega$$

Se prendiamo, pessimisticamente, $h_{fe} = 50$. Quindi possiamo tranquillamente accoppiare in continua i due amplificatori semplificando il circuito e migliorandone anzi le prestazioni.

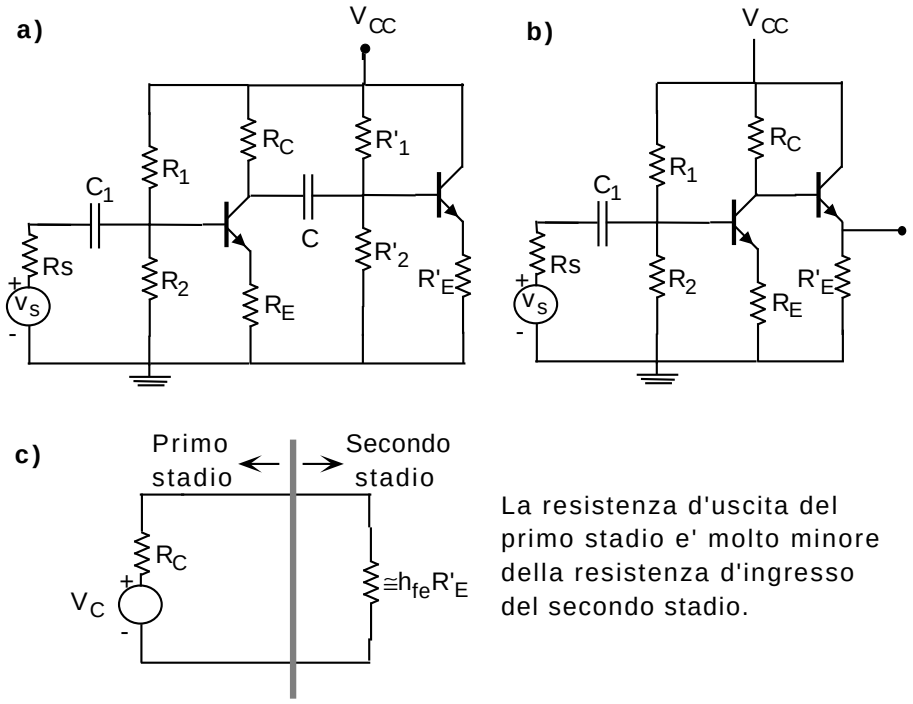

Figura 4.14: *a) Due stadi accoppiate tramite condensatore; b) Accoppiamento in continua; c) Il carico che il primo stadio vede alla sua uscita*

4.7 Amplificatori CE e CC con doppia alimentazione

Lo schema che abbiamo utilizzato finora per costruire l'amplificatore ad emettitore comune (Fig. 4.3) e l'emitter follower (Fig. 4.12) puo' non essere completamente soddisfacente in alcuni casi. Anzitutto obbliga a inserire un condensatore all'ingresso, che, come vedremo meglio in seguito, introduce una attenuazione a bassa frequenza. Per avere una soddisfacente risposta occorre mettere una grande capacita' con conseguente ingombro, aumento dei disturbi, ecc. Inoltre, il partitore che polarizza la base si traduce in una diminuzione della resistenza d'ingresso. In sostanza, una grossa frazione del segnale da amplificare viene dissipata inutilmente in quel partitore.

Si possono evitare entrambi questi inconvenienti con il montaggio mostrato nella Fig. 4.15, che richiede due alimentazioni. Si progetta il circuito in modo che l'emettitore sia a $-0.7\ V$ e la base a zero. In questo modo il condensatore d'ingresso puo' essere omesso e l'amplificatore risponde fino a frequenza zero. In realta' la base e' ad una tensione leggermente negativa, dovuta alla caduta di tensione $R_B I_B$. Il valore di R_B non e' particolarmente critico ai fini del funzionamento del circuito: essa serve solo a consentire il fluire della corrente di base, condizione necessaria per tenere il transistor nella regione attiva. Essa comunque riduce la resistenza d'ingresso del circuito, ma, volendo, puo' essere del tutto omessa nel momento in cui si connette il generatore d'ingresso, perche' attraverso di esso potra' fluire la corrente continua di base necessaria al transistor.

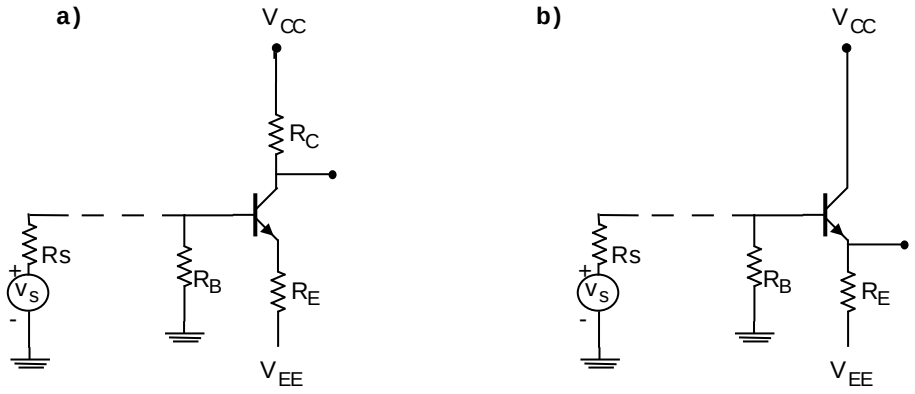

Figura 4.15: *Uso della doppia alimentazione. a) Amplificatore CE; b) Emitter follower. Con transistori npn V_{CC} e' positivo e V_{EE} negativo, l'opposto con transistori pnp.*

4.8 L'effetto Miller

Conviene a questo punto fare una breve parentesi e discutere una proprieta' generale delle reti, che va sotto il nome di teorema di Miller [8] (o effetto Miller); ci sara' molto utile nel seguito. Consideriamo una generica rete in cui tra due particolari nodi (nodo 1 e nodo 2) vi e' un'impedenza Z' (Fig. 4.16a).

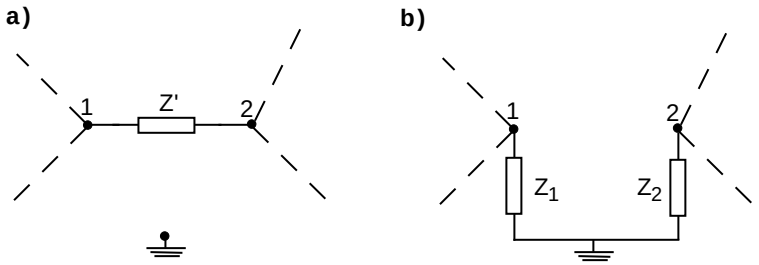

Figura 4.16: *Applicazione del teorema di Miller dei nodi.*

Nell'equazione del nodo 1 vi sará, tra gli altri, un termine

$$I_1 = \frac{V_1 - V_2}{Z'}$$

Potremo quindi manipolare questo termine nel modo seguente

$$I_1 = \frac{V_1 - V_2}{Z'} = \frac{V_1(1 - \frac{V_2}{V_1})}{Z'} = \frac{V_1}{Z'(1-k)}$$

dove si e' posto $k = V_2/V_1$. Analogamente, nell'equazione del nodo 2 vi sará, tra gli altri, un termine

$$I_2 = \frac{V_2 - V_1}{Z'}$$

[8] questo effetto prende il suo nome da quello di John Milton Miller, fisico americano, che lo mise in evidenza studiando gli amplificatori con valvole termoioniche, negli anni attorno al 1920.

che potremo trasformare come segue

$$I_2 = \frac{V_2 - V_1}{Z'} = \frac{V_2(1 - \frac{V_1}{V_2})}{Z'} = \frac{V_2(k-1)}{kZ'}$$

dove k é lo stesso fattore di prima.
In sostanza, nelle due equazioni, i due termini di corrente legati alla presenza di Z' possono essere scritti come

$$I_1 = \frac{V_1}{Z_1}$$

e

$$I_2 = \frac{V_2}{Z_2}$$

dove

$$Z_1 = \frac{Z'}{1-k}$$

e

$$Z_2 = \frac{Z'k}{k-1} = \frac{Z'}{1 - \frac{1}{k}}$$

Quindi l'effetto dell'impedenza Z' é equivalente, nelle due equazioni, alla presenza di un'impedenza Z_1 tra il nodo 1 e la massa ed un'impedenza Z_2 tra il nodo 2 e la massa (Fig. 4.16b).

Una formulazione duale del teorema di Miller puo' essere applicata alle maglie di un circuito. In Fig. 4.17b é mostrato un generico circuito, in cui l'elemento comune a due maglie adiacenti é l'impedenza Z'. Procedendo in analogia al ragionamento precedente é

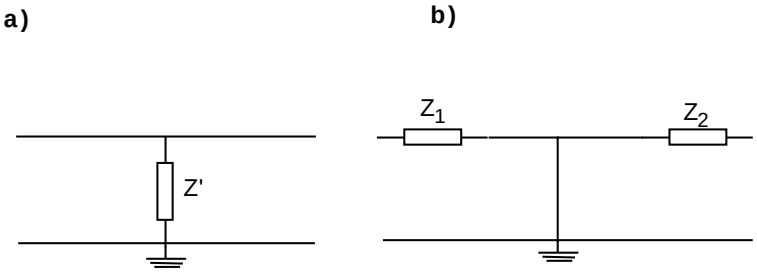

Figura 4.17: *Applicazione del teorema di Miller delle maglie.*

facile dimostrare che il circuito e' equivalente a quello di Fig. 4.17b in cui

$$Z_1 = Z'(1 - A_i) \quad Z_2 = \frac{A_i - 1}{A_i} Z'$$

e dove abbiamo posto

$$A_i = -I_2/I_1$$

Il teorema di Miller é uno strumento utile per calcolare la risposta di un circuito.

Consideriamo ad esempio il circuito amplificatore mostrato in Fig. 4.18a. In questo montaggio l'alimentazione per la base e' prelevata *a valle* di R_C. Applicando il teorema di Miller si arriva allo schema equivalente in Fig, 4.18b, dove

$$R_1 = \frac{R_B}{1 - A_v} \quad R_2 = \frac{R_B}{1 - 1/A_v}$$

In questo modo abbiamo disaccoppiato le maglie d'ingresso e di uscita. Se $|A_v| \gg 1$ $R_1 \approx -R_B/A_v$, mentre $R_2 \approx R_B$.

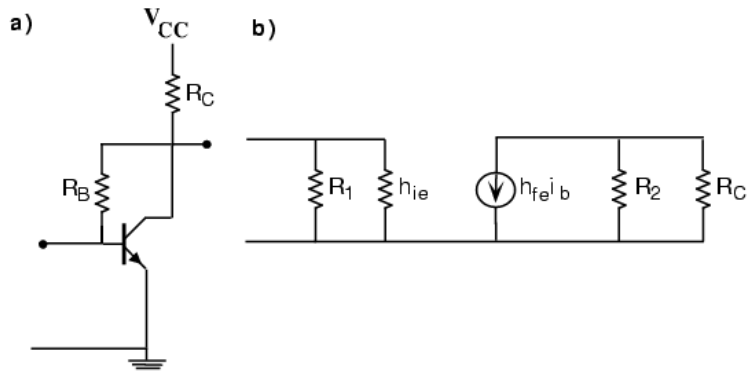

Figura 4.18: *a) Un nuovo amplificatore a emettitore comune; b) Schema equivalente per piccoli segnali dopo l'applicazione del teorema di Miller*

Un altro esempio e' dato dal gia' visto amplificatore con rete autopolarizzante, il cui schema equivalente (semplificato) e' riportato in Fig. 4.19a. Esso si trasforma nello schema in Fig. 4.19b. Ricordando che $A_i = h_{fe}$ e $h_{fe} \gg 1$ si possono ora ricalcolare le caratteristiche di questo amplificatore, ritrovando i gia' noti risultati.
Naturalmente il teorema di Miller si applica anche alle capacità, come avremo modo di verificare nei prossimi paragrafi.

4.9 Risposta in frequenza

4.9.1 Considerazioni generali

Finora ci siamo limitati a studiare i circuiti in una regione di frequenza in cui tutte le reattanze fossero trascurabili. Dobbiamo ora completare il nostro studio, cioe' esaminare l'andamento con ω delle amplificazioni di corrente e di tensione. Concentriamoci per ora sulla amplificazione di tensione: in generale avremo

$$A(\omega) = |A(\omega)|e^{i\phi(\omega)}$$

quindi dobbiamo tener conto sia del modulo che della fase. Consideriamo un segnale d'ingresso sinusoidale a frequenza ω'

$$v_i = v_m \sin \omega' t$$

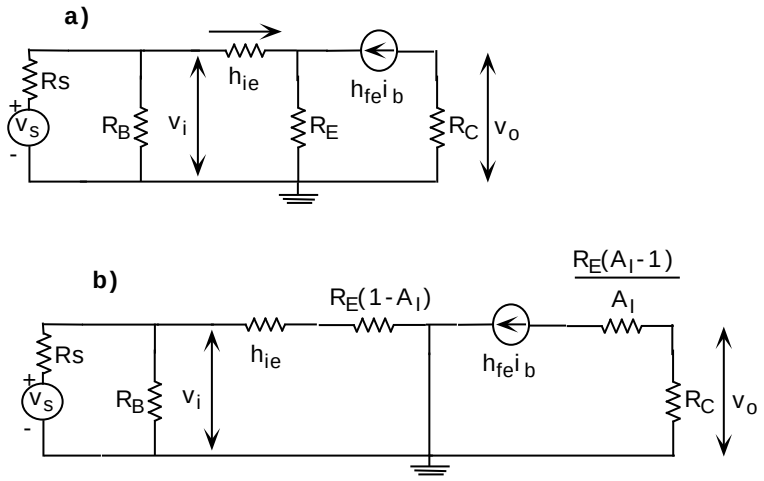

Figura 4.19: *a) L'amplificatore a emettitore comune con rete autopolarizzante; b) Dopo l'applicazione del teorema di Miller*

il segnale d'uscita sará dato da

$$\begin{aligned} v_o &= |A(\omega')| v_m \sin(\omega' t + \phi(\omega')) \\ &= |A(\omega')| v_m \sin[\omega'(t + \tfrac{\phi(\omega')}{\omega'}] \end{aligned}$$

Se prendiamo ora un segnale qualunque, cioé una sovrapposizione di onde di varia frequenza, la sua forma sará preservata se $|A|$ non dipende da ω e se ϕ é zero, oppure se ϕ dipende linearmente da ω. In quest'ultimo caso il segnale sará ritardato (o anticipato) nel tempo di un fattore: infatti, se $\phi(\omega) = k\omega$

$$\begin{aligned} v_o &= |A| v_m \sin[\omega'(t + \tfrac{k\omega'}{\omega'})] \\ &= |A| v_m \sin[\omega'(t + k)] \end{aligned}$$

Torniamo ora alle trasformate di Laplace. La funzione di trasferimento di un circuito é sempre del tipo

$$A(s) = \frac{P(s)}{Q(s)}$$

dove P é un polinomio di grado n, Q é un polinomio di grado m, con $m \geq n$. Le radici do P(s) sono gli zeri di A(s), mentre le radici di Q(s) sono i poli di A(s). Quindi

$$A(s) = k \frac{(s - z_1)(s - z_2)(s - z_3) \ldots (s - z_n)}{(s - p_1)(s - p_2) \ldots (s - p_m)}$$

ovvero anche

$$A(s) = \frac{k(1 - s/z_1)(1 - s/z_2) \ldots (1 - s/z_n)}{(1 - s/p_1)(1 - s/p_2) \ldots (1 - s/p_m)}$$

questo significa che A(s) é sempre pensabile come prodotto di termini

$$(1 + s/z) \quad e \quad \frac{1}{1 + s/p}$$

Il modulo di A(s) é dato dal prodotto dei moduli, e la fase dalla somma delle fasi dei vari termini. Se ora prendo il $\log|A(s)|$, esso é dato dalla somma dei logaritmi dei termini tipo

$$\log|1+s/z_i| \quad e \quad \log|\frac{1}{1+s/p_i}|$$

cioé, in termini di ω

$$\log(1+\frac{\omega^2}{\omega_i^2})^{1/2}$$

e

$$\log(1+\frac{\omega^2}{\omega_i^2})^{-1/2}$$

Prendiamo i termini del primo tipo. Asintoticamente essi valgono

$$\begin{array}{lll} \text{Per } \omega \ll \omega_i & \approx \log 1 = 0 & \text{retta orizzontale} \\ \text{Per } \omega \gg \omega_i & \approx \log \frac{\omega}{\omega_i} & \text{retta a 20db/decade} \end{array}$$

Analogamente nel caso dei termini del secondo tipo si ha

$$\begin{array}{lll} \text{Per } \omega \ll \omega_i & \approx \log 1 = 0 & \\ \text{Per } \omega \gg \omega_i & \approx \log \frac{\omega_i}{\omega} = -\log \frac{\omega}{\omega_i} & \text{retta a -20db/decade} \end{array}$$

Cioé, eccetto che nell'intorno degli zeri e dei poli, il $\log|A(s)|$ é dato dalla sovrapposizione di contributi orizzontali (a 0 db) e di contributi a ± 20 db/decade.

Consideriamo ad esempio il doppio passa-basso, che abbiamo gia' visto nel Cap. 2. La funzione di trasferimento e' data da

$$A(s) = \frac{1}{(1+s/s_1)(1+s/s_2)}$$

Quindi

$$\log|A| = \log\frac{1}{(1+\frac{\omega^2}{\omega_1^2})^{1/2}} + \log\frac{1}{(1+\frac{\omega^2}{\omega_2^2})^{1/2}}$$

Sia per es. $\omega_1 \ll \omega_2$; si avranno 3 regioni:

$$\begin{array}{lll} \omega \ll \omega_1\ ,\omega_2 & \log|A| \approx \log 1 + \log 1 = 0 & \text{orizzontale} \\ \omega_1 \ll \omega \ll \omega_2 & \log|A| \approx -\log\frac{\omega}{\omega_1} + \log 1 & \text{-20 db/decade} \\ \omega_1\ ,\omega_2 \ll \omega & \log|A| \approx -\log\frac{\omega}{\omega_1} - \log\frac{\omega}{\omega_2} & \\ & = -\log\frac{\omega^2}{\omega_1\omega_2} & \\ & = -2\log\frac{\omega}{\omega_1\omega_2} & \text{-40 db/decade} \end{array}$$

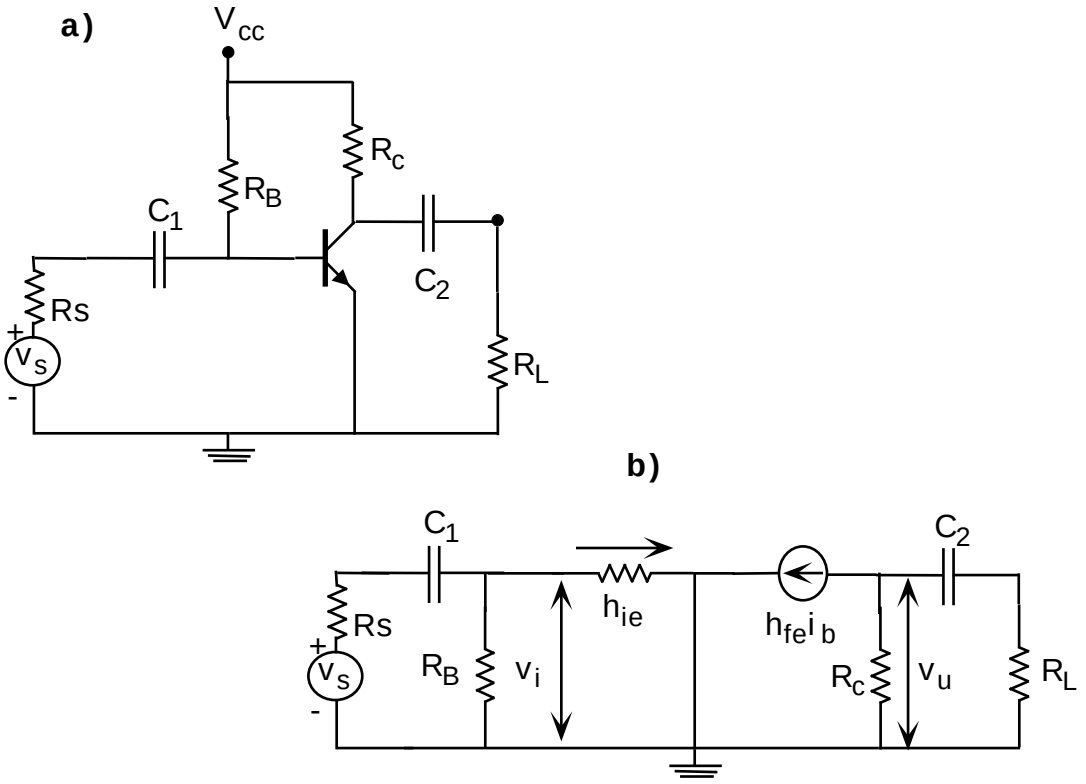

Figura 4.20: *a) Amplificatore a emettitore comune; b) Schema equivalente per piccoli segnali includendo le capacita' d'ingresso e d'uscita*

4.9.2 Risposta a bassa frequenza

Consideriamo il circuito in Fig. 4.20a ed il suo equivalente con il modello a parametri ibridi. Abbiamo anche aggiunto un carico esterno R_L, ed un capacitore di disaccoppiamento C_2. Con

$$A_v \equiv \frac{v_u}{v_i} = -h_{fe}\frac{R_C}{h_{ie}}$$

indichiamo l'amplificazione di tensione a media frequenza (trascuriamo per ora l'effetto del carico R_L, cioè consideriamo $R_L = \infty$). Ora,

$$A'_v = \frac{v_u}{v_s} = \frac{v_u}{v_i}\frac{v_i}{v_s} = A_v\frac{v_i}{v_s}$$

Passando nel dominio delle frequenze

$$A'_v(\omega) = A_v\frac{V_i}{V_s}$$

$$\frac{V_i}{V_s} = \frac{R'}{R' + R_s + \frac{1}{j\omega C_i}}$$

dove $R' = R_B||h_{ie}$

$$\frac{V_i}{V_s} = \frac{\frac{R'}{R'+R_s}}{\frac{R'+R_s}{R'+R_s} + \frac{1}{j\omega C_i(R'+R_S)}} = \frac{R'}{R'+R_S}\frac{1}{1+\frac{\omega_1}{j\omega}}$$

dove

$$\omega_1 = \frac{1}{C_1(R'+R_S)} = \frac{1}{\tau_1}$$

Quindi

$$A'_v = -h_{fe}\frac{R_C}{h_{ie}}\frac{R'}{R'+R_s}\frac{1}{1-j\frac{\omega_1}{\omega}}$$

Cioè si ha un passa-alto con frequenza di taglio

$$f_1 = \frac{\omega_1}{2\pi} = \frac{1}{2\pi C_1(R'+R_S)}$$

Nel caso del circuito con rete autopolarizzante si applica lo stesso calcolo dove ora però

$$A_v = -h_{fe}\frac{R_C}{R_E}$$

$$R_i \equiv h_{ie} + (1+h_{fe})R_C$$

Si avrà quindi

$$A'_v = -h_{fe}\frac{R_C}{R_E}\frac{R'}{R'+R_S}\frac{1}{1-j\frac{\omega_1}{\omega}}$$

dove

$$R' = R_B||R_i = R_B||(h_{ie} + (1+h_{fe})R_E)$$

e

$$\omega_1 = \frac{1}{C_1(R'+R_S)}$$

Lasciamo al lettore il completare questo studio con l'aggiunta del contributo dovuto al passa-alto sulla maglia di uscita.

4.9.3 Risposta ad alta frequenza

Possiamo ora comprendere qualitativamente il comportamento dell'amplificatore CE, utilizzando il modello di Giacoletto, che ci consente di schematizzare il circuito come in Fig. 4.21a. trascurando la resistenza r_μ. ed applicando il teorema di Miller, la capacita' C_μ puo' essere sostituita da due capacita', una in parallelo a C_π, l'altra sulla maglia di uscita (Fig. 4.21b)); si puo' poi ulteriormente semplificare il circuito nella maglia d'ingresso, con l'utilizzazione del teorema di Thevenin, e nella maglia di uscita, sostituendo le due resistenze con il loro parallelo (Fig. 4.21c)).
Grazie a queste semplificazioni e' evidente il comportamento dell'amplificatore ad alte frequenze: e' il comportamente di un doppio passa-basso. Le due frequenze critiche sono in

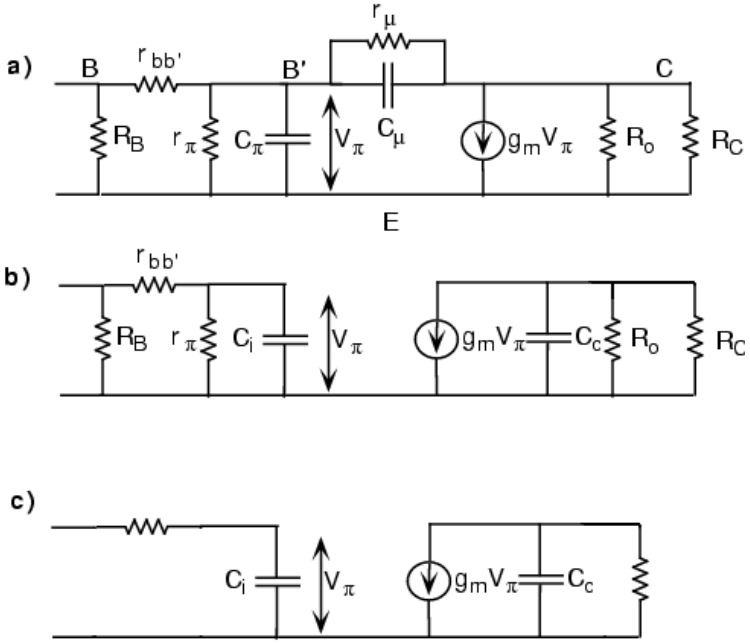

Figura 4.21: *a) Amplificatore a emettitore comune nel modello di Giacoletto; b) Dopo l'applicazione del teorema di Miller; c) Ulteriore semplificazione*

genere abbastanza distanziate tra loro: il passa-basso nella maglia d'ingresso ha una frequenza di taglio molto piu' bassa di quella dovuta alla maglia d'uscita[9], cio' significa che il diagramma di Bode presentera', in generale, un primo tratto con pendenza -20 dB/decade seguito da un secondo tratto a -40 dB/decade.

4.9.4 Risposta in frequenza del circuito amplificatore CE con capacità sull'emettitore

Conviene brevemente tornare su questo circuito per comprenderne, anche qui qualitativamente, il comportamente alle varie frequenze. Anche in questo caso, per semplicita', analizzeremo separatamente le varie regioni. Consideriamo anzitutto la frequenza bassa-media (in cui cioè il transistor è parametrizzabile con i parametri h semplificati). Mettiamoci inoltre in una regione in cui $\frac{1}{\omega C_e} \gg R_E$, cioè $Z_e = R_E||C_e \approx R_E$. Allora il comportamento del circuito e' sostanzialmente equivalente a quello di un normale amplificatore con resistenza sull'emettitore (amplificazione $A_v \simeq -R_C/R_E$). Viceversa a frequenze medio-alte (in cui ancora la capacita' parassite del transistor sono trascurabili) in cui l'impedenza del capacitore C_e diviene bassissima, l'emettitore e' sostanzialmente a massa, e il circuito si

[9]Questo e' dovuto proprio all'effetto Miller: la capacita' C_μ e' riportata sulla maglia d'ingresso moltiplicata per il fattore di amplificazione. Quindi la frequenza di taglio e' tanto piu' bassa quanto piu' e' grande l'amplificazione. Piu' in generale, si comprende che il comportamento ad alta frequenza del transistor non e' solo legato ai valori delle capacita' parassite, ma anche al valore dei resistori esterni e, come detto, al valore dell'amplificazione a bassa frequenza.

comporta come un amplificatore senza resistenza di emettitore, con

$$A_v = -\frac{h_{fe}R_C}{h_{ie}}$$

Aumentando ancora la frequenza si ricade nel comportamento visto nel paragrafo precedente. Ci aspettiamo quindi che il diagramma di Bode dell'amplificatore sia qualitativamente simile a quello schematizzato nella Fig. 4.22: a bassissima frequenza si ha un passa-alto dovuto al capacitore d'ingresso; segue una regione a bassa amplificazione da cui si transisce in una regione ad alta amplificazione, seguita infine da una discesa delle prestazioni dovuta all'intervento delle capacita' interne del transistor.

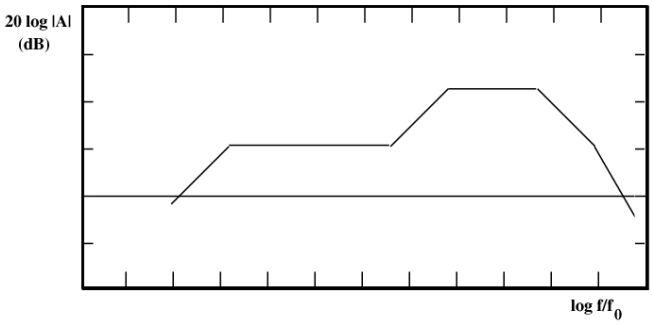

Figura 4.22: *Amplificatore CE con capacita' sull'emettitore: andamento qualitativo della risposta in funzione della frequenza.*

4.10 Amplificatore differenziale

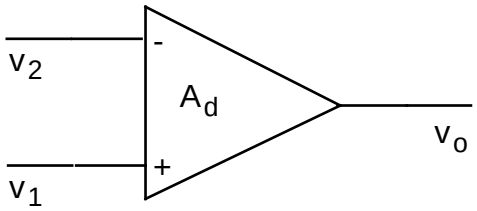

Figura 4.23: *a) L'amplificatore differenziale*

Studieremo ora l'amplificatore differenziale (Fig. 4.23), cioe' un dispositivo a due ingressi, in cui si richiede che

$$v_o = A_d(v_1 - v_2) \tag{4.12}$$

Possiamo comprendere come realizzare un dispositivo del genere osservando che, in generale, la tensione d'uscita v_o sara' funzione delle due tensioni d'ingresso; sviluppando in

serie e limitandosi al primo ordine si avra'

$$v_o = A_1 v_1 + A_2 v_2 \tag{4.13}$$

Introduciamo due nuove variabili

$$v_c = \frac{1}{2}(v_1 + v_2)$$

$$v_d = (v_1 - v_2)$$

Si ricava quindi che

$$v_1 = v_c + \frac{1}{2} v_d$$

$$v_1 = v_c - \frac{1}{2} v_d$$

e sostituendo nella 4.13 si ottiene

$$v_o = A_d v_d + A_c v_c \tag{4.14}$$

dove

$$A_d = \frac{1}{2}(A_1 - A_2) \tag{4.15}$$

$$A_c = (A_1 + A_2) \tag{4.16}$$

A_d e A_c prendono il nome di *amplificazione differenziale* e di *amplificazione di modo comune* rispettivamente. E' anche utile notare che A_d rappresenta l'amplificazione del circuito quando i segnali d'ingresso sono uguali ed opposti, mentre A_c rappresenta l'amplificazione quando i segnali d'ingresso sono uguali. Confrontando la 4.14 con la 4.12 si vede che possiamo ottenere il risultato voluto se costruiamo un dispositivo con $A_c = 0$ e $A_d \neq 0$, cioe' dobbiamo avere

$$A_1 = -A_2$$

Possiamo facilmente comprendere che in circuito reale questa condizione ben difficilmente puo' essere realizzata in modo esatto; quello che in realta' si puo' fare e' di avere A_c molto piccola rispetto ad A_d. E' logico quindi definire un fattore di merito dell'amplificatore differenziale come il rapporto

$$\rho = |\frac{A_d}{A_c}|$$

che prende il nome di Common Mode Rejection Ratio (CMRR);in un amplificatore differenziale ideale si ha quindi $\rho = \infty$.

Gli amplificatori differenziali sono dispositivi molto importanti e comunemente usati. Per comprenderne l'utilita' possiamo confrontare le due situazioni in Fig. 4.24, in cui un segnale (contenente una certa informazione) esce da una sorgente e viene trasferito all'ingresso di un amplificatore, mescolato a disturbi provenienti dall'esterno. Nel caso a) viene usato un amplificatore convenzionale, alla cui uscita il segnale ed il disturbo vengono ugualmente amplificati; nel caso b) il disturbo, essendo presente in egual misura su entrambi gli ingressi, viene fortemente soppresso. Si comprende quindi la convenienza ad utilizzare

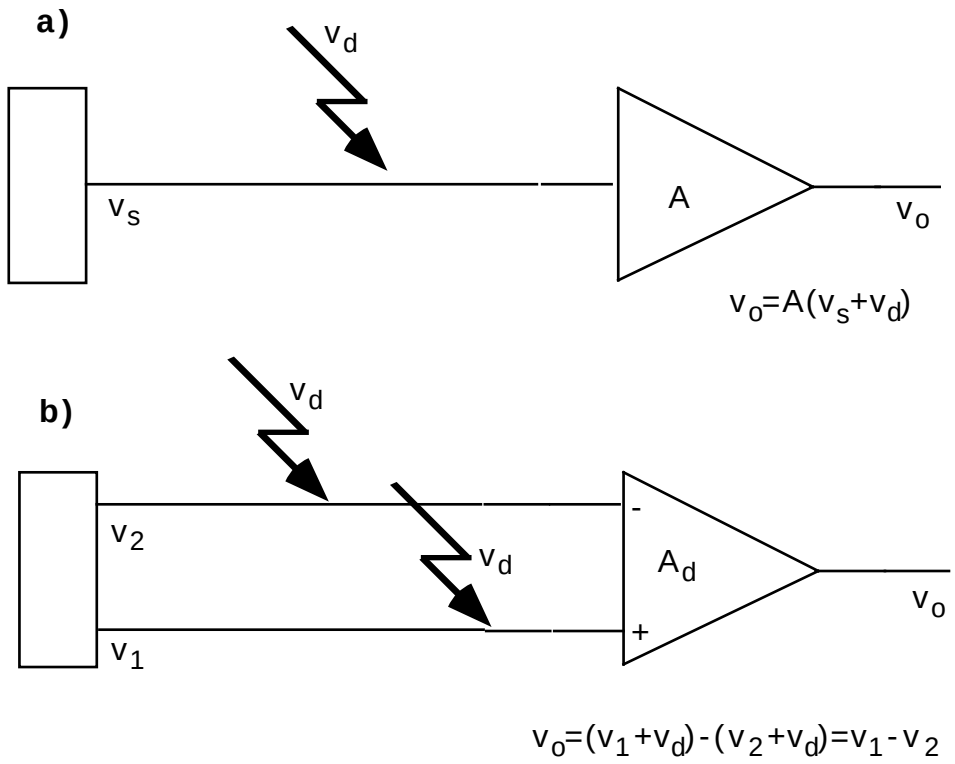

Figura 4.24: *Amplificazione di un segnale soggetto a disturbi: a) con amplificatore semplice; b) con amplificatore differenziale*

il modo differenziale per trasmettere ed elaborare segnali quando disturbi, sia provenienti dall'esterno ma anche interni ai circuiti (per esempio, il ripple dell'alimentazione in continua), debbano essere soppressi.

Un esempio di realizzazione di un amplificatore differenziale e' riportato in Fig. 4.25a, dove l'uscita puo' essere prelevata su uno qualunque dei due collettori[10]. L'uso della doppia alimentazione consente di evitare capacitori ai due ingressi ed avere quindi una buona risposta fino a frequenza zero.

Supporremo, per semplicita' di calcolo, che il circuito sia esattamente simmetrico e che quindi i due transistor siano assolutamente identici[11]. Dopo aver polarizzato i due transistor attraverso un'opportuna scelta dei valori delle resistenze, possiamo studiare il circuito utilizzando lo schema equivalente per piccoli segnali di Fig. 4.25b; da cui cercheremo di ricavare A_d ed A_c. Per fare cio' possiamo metterci in due casi limite:

a) Segnali uguali sui due ingressi.

Poiche' $v_2 = v_1$, dalle relazioni precedenti avremo

$$v_o = A_c v_1 \tag{4.17}$$

che ci consente di ricavare A_c.

[10] E' anche possibile, naturalmente, prelevare entrambe le uscite: si avra' allora un amplificatore differenziale con uscita differenziale.

[11] Il lettore potrebbe obiettare che questa condizione non e' praticamente realizzabile. In realta', il calcolo esatto delle prestazioni di questo circuito dimostra che questa condizione non e' affatto necessaria.

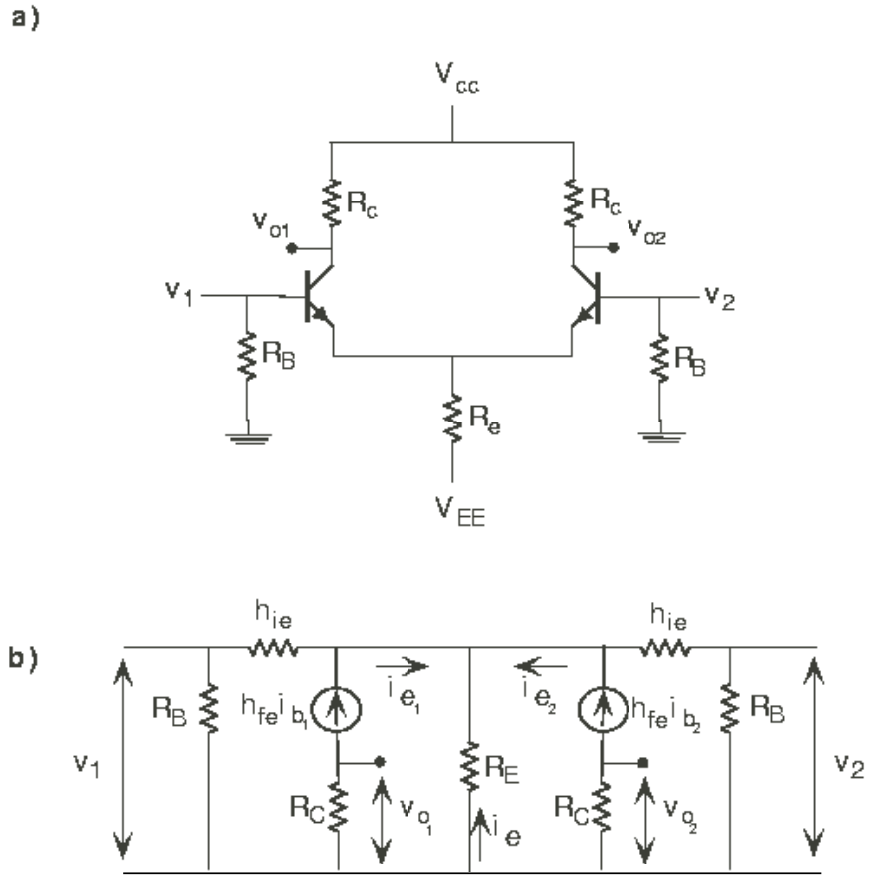

Figura 4.25: *a) Amplificatore differenziale; b) circuito equivalente*

In questo caso anche $i_{b_1} = i_{b_2}$, e potremo quindi scrivere

$$\begin{aligned} v_o &= -h_{fe} R_C i_{b_1} \\ v_1 &= h_{ie} i_{b_1} - i_e R_E \\ &= h_{ie} i_{b_1} - (i_{e_1} + i_{e_2}) R_E \\ &= h_{ie} i_{b_1} + (1 + h_{fe}) R_E i_{b_1} + (1 + h_{fe}) R_E i_{b_2} \\ &= h_{ie} i_{b_1} + 2(1 + h_{fe}) R_E i_{b_1} \end{aligned}$$

Si ha quindi

$$A_c = \frac{v_o}{v_1} = \frac{-h_{fe} R_C}{2(1 + h_{fe}) R_E} \tag{4.18}$$

b) Segnali uguali ed opposti sui due ingressi
In questo caso $v_2 = -v_1$ e $i_{b_2} = -i_{b_1}$ Quindi, sempre dalle relazioni precedenti

$$v_o = 2 A_d v_1 \tag{4.19}$$

che ci consente di ricavare A_d.

$$\begin{aligned} v_o &= -h_{fe}R_C i_{b_1} \\ v_1 &= h_{ie} i_{b_1} \end{aligned}$$

Posso ricavare

$$A_d = \frac{v_o}{2v_1} = \frac{-h_{fe}R_C}{2h_{ie}} \tag{4.20}$$

Utilizzando le relazioni 4.18 e 4.20 si ricava il CMRR di questo circuito:

$$\frac{A_d}{A_c} = \frac{h_{ie} + 2(1+h_{fe})R_E}{2h_{ie}} \simeq \frac{(1+h_{fe})R_E}{h_{ie}} \simeq g_m R_E \tag{4.21}$$

Gli stessi risultati potevano essere ottenuti valutando invece A_1 ed A_2 (cioe' le amplificazioni per segnale singolo) e ricavando poi A_c ed A_d dalle relazioni 4.15 e 4.16. Come si vede il CMRR del circuito dipende essenzialmente da R_E, e migliora al crescere di essa. Cio' si comprende osservando che l'amplificazione differenziale e' sostanzialmente quella di un normale amplificatore CE (a parte un fattore 2), mentre l'amplificazione di modo comune e' quella di un amplificatore CE con rete autopolarizzante (sempre a meno di un fattore 2). E' chiaro quindi che le prestazioni del circuito possono essere migliorate aumentando R_E: se $R_E \to \infty$ $A_c \to 0$. Tuttavia cio' ha dei limiti in quanto un aumento di R_E comporta una diminuzione delle correnti di collettore, I_C, dei due transistor, con conseguente diminuzione di g_m.

Amplificatore differenziale con generatore di corrente

Si possono ottenere migliori prestazioni sostituendo il resistore R_E con un transistor come in Fig. 4.26a.
Il transistor T_3 si comporta come un generatore di corrente (quasi ideale). Scrivendo l'equazione della sua maglia di base abbiamo

$$V_{EE}\frac{R_2}{R_1+R_2} = V_{BE_3} + I_3 R_3$$

e possiamo ricavare la corrente di collettore (e di emettitore)

$$I_0 \simeq I_3 = \frac{1}{R_3}\left(\frac{V_{EE}R_2}{R_1+R_2} - V_{BE_3}\right)$$

che é ovviamente indipendente dagli altri due transistor. Ora possiamo studiare il circuito utilizzando lo schema di Fig. 4.26b; si noti che il generatore di corrente che ha sostituito R_E corrisponde ad una variazione di corrente $i_e = 0$ (per definizione di generatore di corrente ideale) e quindi si deve avere

$$i_{e_1} = -i_{e_2}$$

Nel caso di segnali d'ingresso uguali si ha $v_1 = v_2$, $i_{b_1} = i_{b_2}$ e conseguentemente $i_{e_1} = i_{e_2}$. Le due correnti di emettitore devono essere percio' uguali, e contemporaneamente opposte: ne consegue che sono nulle ed e' nulla anche la tensione v_0.

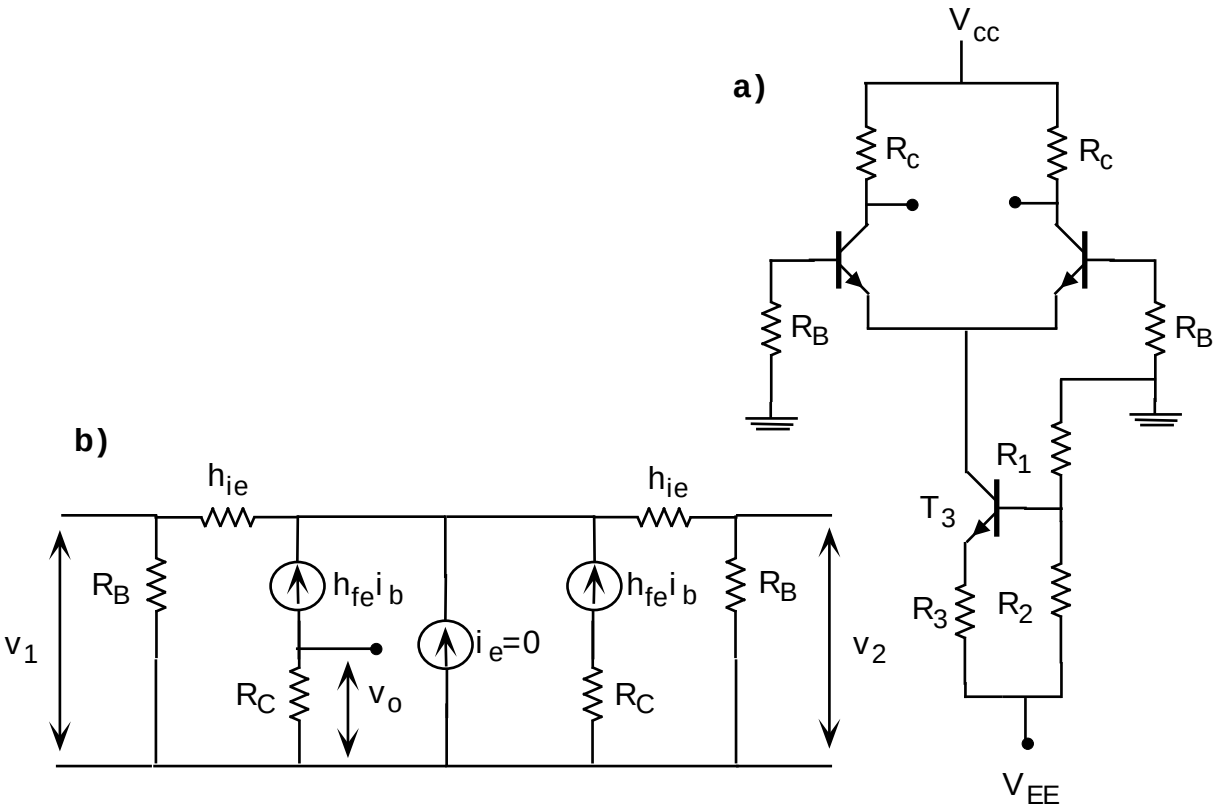

Figura 4.26: *a) Amplificatore differenziale con generatore di corrente; b) schema equivalente*

L'amplificazione di modo comune A_c e' allora nulla, mentre l'amplificazione differenziale resta la stessa del caso precedente, dando luogo ad un CMMR infinito, almeno in linea di principio.

Nella realta' il transistor T_3 non e' un generatore *ideale* di corrente; tuttavia lo approssima molto bene, poiché la sua resistenza d'uscita e' molto grande; avremo quindi un CMRR molto migliore rispetto al caso precedente.

4.11 Amplificatori con reazione negativa

La stabilita' e' un requisito essenziale degli amplificatori, in tutte le loro applicazioni. Si richiede cioe' che le prestazioni (amplificazione di corrente e di tensione) siano indipendenti da fattori esterni, p.es. la temperatura, e non siano legate ai valori individuali dei parametri dei transistors. In genere questi requisiti sono ottenuti introducendo effetti di reazione negativa (o controreazione). Abbiamo gia' utilizzato, senza saperlo, questi effetti; ora dobbiamo studiarli in modo esplicito.

Consideriamo la rete in Fig. 4.27, in cui il segnale X_o all'uscita dell'amplificatore A viene, in parte, rimiscelato all'ingresso, attraverso la rete passiva β. Con X_i, X_o, ecc.,

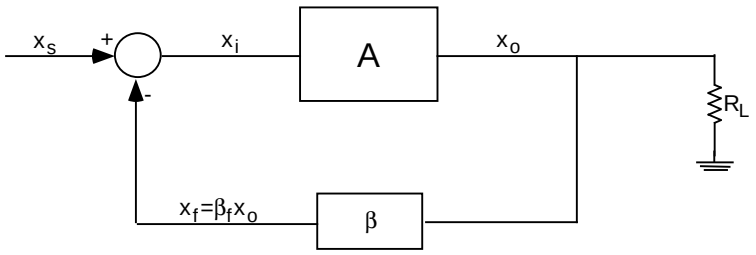

Figura 4.27: *Una rete reazionata*

abbiamo indicato una generica variabile elettrica, corrente o tensione. Abbiamo ora

$$\begin{aligned} X_i &= X_s - X_f \\ &= X_s - \beta X_o \end{aligned}$$

si noti che X_f e' *sottratto* al segnale d'ingresso X_s. Si ha quindi

$$A_f = \frac{X_o}{X_s} = \frac{X_0}{X_i + \beta X_o} \tag{4.22}$$

e, dividendo numeratore e denominatore per X_i, si ottiene

$$A_f = \frac{A}{1 + \beta A} \tag{4.23}$$

Chiaramente l'amplificazione complessiva (amplificazione *con reazione*) e' diminuita, tuttavia ne avremo guadagnato in stabilita': infatti, se $\beta A \gg 1$ si ha

$$A_f \simeq \frac{1}{\beta} \tag{4.24}$$

Questo e' molto importante perche' la funzione di trasferimento β e' in genere legata solo a componenti passive, quindi l'amplificazione con reazione non dipende piu' da parametri instabili, come ad esempio quelli dei transistors che costituiscono l'amplificatore A. Piu' in generale, differenziando la 4.23 rispetto ad A si ottiene:

$$\left| \frac{dA_f}{A_f} \right| = \frac{1}{|1 + \beta A|} \frac{dA}{A} \tag{4.25}$$

Cio' significa che le variazioni di A_f sono ridotte, rispetto a quelle di A di un fattore grande.

La reazione negativa ha inoltre effetto sulla larghezza di banda dell'amplificatore. Supponiamo che il nostro amplificatore sia approssimativamente esprimibile come un passa-basso:

$$A = \frac{A_o}{1 + j\frac{f}{f_H}} \tag{4.26}$$

Introducendo la reazione si ha

$$\begin{aligned} A_f &= \frac{\frac{A_o}{1+j\frac{f}{f_H}}}{1+\frac{\beta A_o}{1+j\frac{f}{f_H}}} \\ &= \frac{A_o}{1+\beta A_o + j\frac{f}{f_H}} \\ &= \frac{\frac{A_o}{1+\beta A_o}}{1+j\frac{f}{f_H(1+\beta A_o}} \\ &= \frac{A_{of}}{1+j\frac{f}{f'_H}} \end{aligned}$$

dove abbiamo indicato con A_{of} l'amplificazione a media frequenza con reazione. Si vede quindi che la nuova frequenza di taglio f'_H e' aumentata di un fattore $(1+\beta A_o)$, con conseguente incremento della larghezza di banda. E' interessante notare che il prodotto (amplificazione)x(larghezza di banda) e' costante: infatti

$$A_f f'_H = A f_H \tag{4.27}$$

Questa proprieta' deriva semplicemente dalla linearita' della discesa di amplificazione ad alta frequenza; non e' piu' valida se l'amplificazione scende con tratti di pendenza diversa.

Nella Fig. 4.28 si vedono le quattro possibilita' per introdurre la reazione negativa in un circuito. Infatti il segnale di reazione X_f puo' essere proporzionale alla tensione o alla corrente d'uscita, e puo' essere miscelato al segnale d'ingresso in serie o in parallelo. Si deve comprendere che la grandezza reazionata non e' necessariamente l'amplificazione di corrente o quella di tensione, ma puo' anche essere la transcoduttanza (rapporto tra corrente d'uscita e tensione d'ingresso) o la transresistenza (rapporto tra tensione d'uscita e corrente d'ingresso). Per ogni tipo di reazione e' stabilizzata una di queste quattro grandezze.

Prendiamo ad esempio il caso tensione-serie (Fig. 4.28a). Si ha

$$\begin{aligned} V_s &= V_i + \beta V_o \\ V_o &= A V_i \end{aligned}$$

Sostituendo la seconda nella prima si ricava

$$V_s = \frac{V_O}{A} + \beta V_o$$

da cui

$$A_{vf} = \frac{V_o}{V_s} = \frac{A}{1+\beta A}$$

In questo caso la reazione agisce sull'amplificazione di tensione.

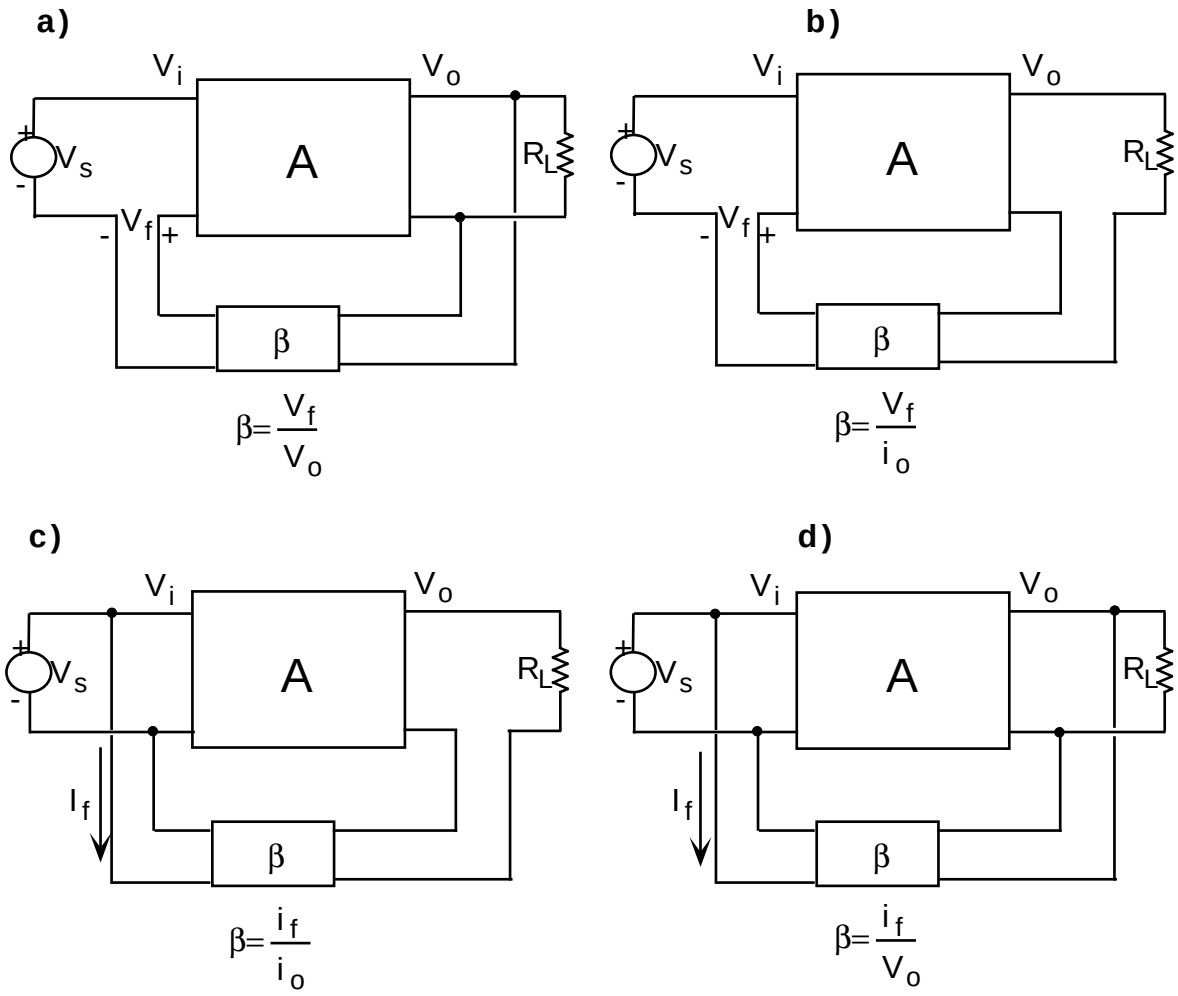

Figura 4.28: *Reazione negativa: a) di tensione in serie; b) di corrente in serie; c) di tensione in parallelo; d) di corrente in parallelo*

4.11.1 Esempi

Vediamo ora alcuni esempi di circuiti reali in cui si hanno i vari tipi di reazione. Si noti che normalmente il circuito non e' formato da un amplificatore cui si aggiunge una rete di reazione; al contrario spesso la reazione e' insita nel circuito stesso. Consideriamo l'emitter follower, schematicamente riportato in Fig. 4.29a. In questo circuito si ha un esempio di reazione tensione-serie; infatti, nella maglia d'ingresso viene riportata una tensione (in serie al generatore v_s) uguale alla tensione d'uscita v_o. In questo caso quindi il β e' uguale ad 1 e si ha

$$A_{vf} \simeq \frac{1}{\beta} = 1$$

Prendiamo invece l'amplificatore CE di Fig. 4.29b. In questo caso si riporta in ingresso una tensione proporzionale alla corrente d'uscita I_C. Si ha allora

$$\beta = \frac{V_f}{I_C} = \frac{-I_C R_E}{I_C} = -R_E$$

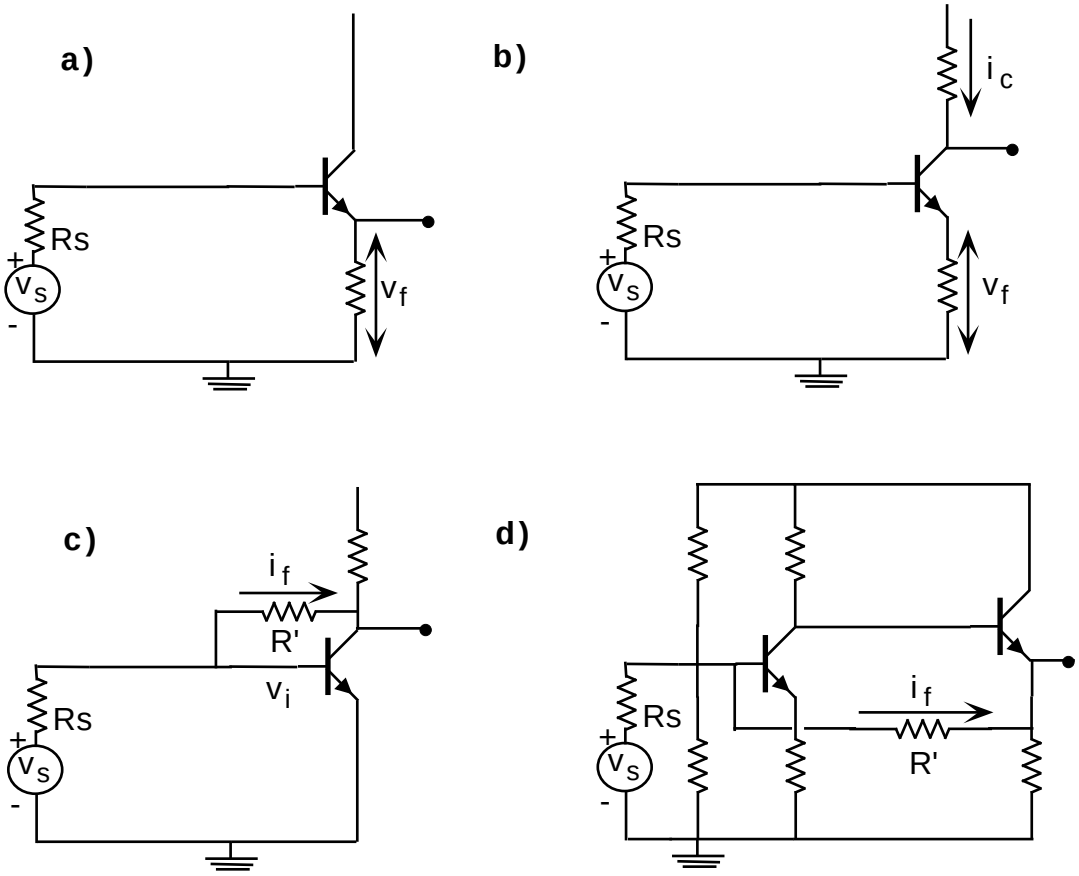

Figura 4.29: *Reazione negativa: a)esempio di tensione in serie; b) di corrente in serie; c) di corrente in parallelo; d) di tensione in parallelo*

In questo caso la grandezza reazionata e' la transconduttanza G_m cioe'

$$G_{mf} \simeq \frac{1}{\beta} = -\frac{1}{R_E}$$

Il circuito di Fig. 4.29c e' un esempio di reazione tensione-parallelo. Infatti la corrente che scorre in R' e' data da

$$I_f = \frac{V_i - V_o}{R'} \simeq -\frac{V_o}{R'}$$

La grandezza reazionata e' la transresistenza R_m, cioe'

$$R_{mf} \simeq \frac{1}{\beta} = -R'$$

Infine il circuito in Fig. 4.29d e' un esempio di reazione corrente-parallelo. Si tratta di un amplificatore a due stadi (CE+CC) in cui si riporta in parallelo all'ingresso del primo stadio un segnale proporzionale alla corrente d'uscita. Si ha infatti

$$I_f = \frac{V_{B_1} - V_{E_2}}{R'} \simeq \frac{-V_{E_2}}{R'} = \frac{(I_o - I_f)R_E}{R'}$$

Da cui si puo' ricavare I_f

$$I_f = \frac{R_E}{R' + R_E} I_o$$

La grandezza reazionata e' l'amplificazione di corrente ed il fattore β e' dato da

$$\beta = \frac{R_E}{R' + R_E}$$

Quindi l'amplificazione di corrente e' data da

$$A_{if} \simeq \frac{1}{\beta} = \frac{R' + R_E}{R_E}$$

La reazione negativa ha anche un effetto su resistenza d'ingresso e resistenza d'uscita. In particolare se la reazione e' in serie la resistenza d'ingresso aumenta, mentre diminuisce se e' in parallelo. Per quanto riguarda invece la resistenza d'uscita essa diminuisce se la reazione e' di tensione, aumenta se e' di corrente.

Conviene riepilogare in una tabella le caratteristiche dei vari tipi di reazione:

	Tensione serie	Corrente serie	Corrente parallelo	Tensione parallelo
R_o	Diminuisce	Aumenta	Aumenta	Diminuisce
R_i	Aumenta	Aumenta	Diminuisce	Aumenta
Stabilizza	A_v	G_m	A_i	R_m

In tutti i casi si ha una diminuzione della grandezza stabilizzata, compensata da un allargamento della banda passante.

Capitolo 5

Transistors ad effetto di campo

5.1 Introduzione

I transistors ad effetto di campo sono caratterizzati da una serie di proprieta' che li rendono preferibili, in molte applicazioni, ai transistors a giunzione. Essi possono anzitutto avere dimensioni molto ridotte, il che apre grandi possibilità di integrazione su larga scala; inoltre possono essere usati per realizzare resistori e capacitori e quindi si possono costruire circuiti integrati senza ricorrere a componenti discreti. Sono dispositivi a 'portatori di maggioranza', quindi meno sensibili alla temperatura e possono avere una resistività di ingresso molto alta. Esistono due tipi di transistors ad effetto di campo:
JFET (Junction Field Effect Transistor);
IGFET (Insulated Gate Field Effect Transistors), anche noti come MOSFET (Metal Oxide Semiconductor FET).

Noi studieremo principalmente i JFET, mentre daremo solo qualche cenno sui MOSFET.

5.2 Il transistor JFET

Il JFET a canale n (Fig. 5.1a) e' costituito da una barretta di materiale di tipo n con due inserzioni di materiale di tipo p fortemente drogato (che viene percio' indicato con p^+); le due inserzioni sono elettricamente collegate tra loro e formano il terminale di Gate (indicato con G), mentre i due estremi della barretta costituiscono (attraverso opportuni contatti metallici) i terminali di Drain e Source. Il JFET viene polarizzato come in Fig 5.1c in modo che la giunzione pn gate-canale sia polarizzata inversamente. Si crea una regione di svuotamento come in figura e, poiche' la zona p e' molto piu' drogata, la regione di svuotamento è tutta nel canale. Ricordiamo che nella regione di svuotamento la conducibilità è zero perchè non vi sono virtualmente cariche libere. Quindi l'effettiva larghezza del canale dipende dal voltaggio V_{gs} applicato: a un certo valore $V_{GS} = V_P$ la larghezza del canale si riduce a zero (tensione di pinch-off). Questo dispositivo si chiama FET a canale n. E' naturalmente possibile realizzare un dispositivo analogo usando materiale di tipo p, con inserzioni di tipo n^+ (FET a canale p).

In realta' il JFET viene utilizzato applicando una differenza di potenziale anche tra drain e source; si ha quindi che la corrente circolante tra questi due terminali, I_D, e'

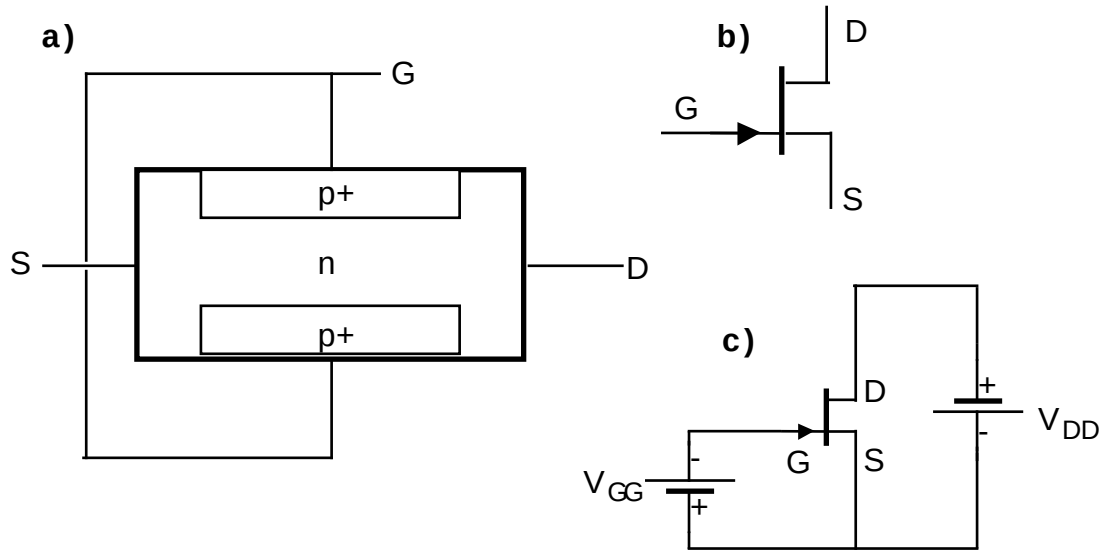

Figura 5.1: *a) JFET a canale n; b) Simbolo circuitale; c) Circuito di polarizzazione*

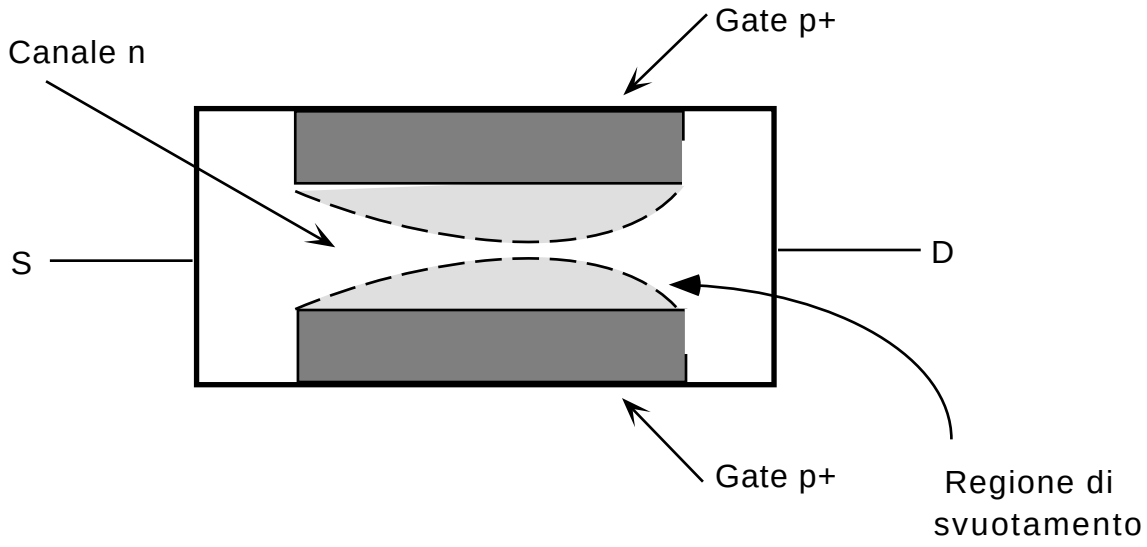

Figura 5.2: *La regione di svuotamento*

funzione di V_{GS} e di V_{DS}. Le curve caratteristiche che si ottengono sono mostrate in Fig. 5.3 Possiamo distinguere una regione *ohmica*, una regione di *saturazione*, una di *break-down*, ed infine una regione di *interdizione*.

Regione ohmica
Se la tensione V_{DS} e' abbastanza piccola, ci aspettiamo, per un dato valore di V_{GS} una relazione del tipo:

$$I_D = AqN_D\mu_n\epsilon = 2bWqN_D\mu_n\frac{V_{DS}}{L} = 2bqN_D\mu_n\frac{W}{L}V_D \tag{5.1}$$

dove 2b, W, L sono rispettivamente la larghezza, lo spessore, e la lunghezza del canale. Naturalmente b dipende da V_{GS}. Si ha quindi una relazione lineare e si puo' definire la resistenza del FET come

$$r_{DS(ON)} = \frac{1}{2aqN_D\mu_n}(\frac{L}{W}) \tag{5.2}$$

dove a e' la larghezza che il canale assume per $V_{GS} = 0$. Tipicamente $r_{ds} = 10 - 10^2$ ohm.

E' chiaro quindi che al variare di V_{GS}, poiche' varia la larghezza del canale, varia anche la resistenza del transistor e quindi la pendenza delle curve nell'origine. Poichè $\mu_p \ll \mu_\mu$ con i FET a canale p si realizzano resistenze molto più grandi.

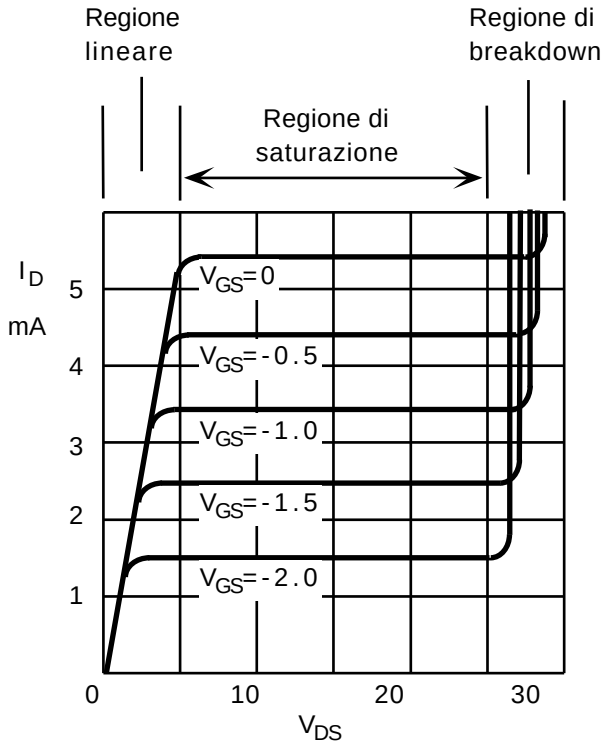

Figura 5.3: *Caratteristiche*

Regione di saturazione

Quando V_{DS} e' diversa da zero nasce lungo x un campo elettrico E non trascurabile. L'estremita' del gate rivolta verso il drain risulta polarizzata inversamente in misura maggiore dell'estremita' rivolta verso il source; pertanto i contorni della regione di svuotamento non sono paralleli all'asse longitudinale del canale, ma assumono l'andamento mostrato nella Fig. 5.2. Al crescere di V_{DS} le grandezze E e I_D aumentano, mentre b diminuisce poiche'

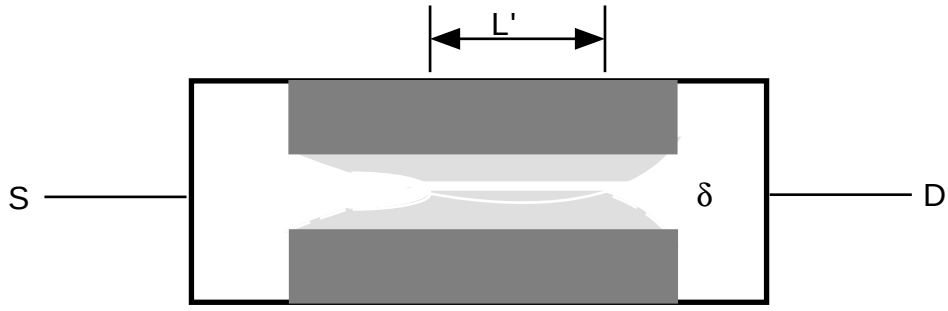

Figura 5.4: *Superata la tensione di restringimento, al crescere di V_{DS}, L' aumenta mentre δ e I_D rimangono costanti.*

il canale si restringe; pertanto la densita' di corrente J deve aumentare. Tuttavia essa non puo' aumentare oltre certi limiti: pertanto si osserva sperimentalmente che la mobilita' degli elettroni diminuisce in modo inversamente proporzionale ad E, per cui la velocita' di deriva v degli elettroni ($v = \mu E$) resta costante e la legge di Ohm non e' piu valida. In conseguenza la corrente I_D comincia a saturare; quando V_{DS} cresce oltre il valore di pinch-off il profilo della regione di svuotamento assume l'andamento mostrato in Fig. 5.4, in cui si osserva un'allungamento della regione L' dove la velocita' di deriva e' limitata.

Regione di Breakdown

Se la differenza di tensione tra drain e source cresce ulteriormente si puo' provocare un fenomeno di breakdown per valanga. Dalla Fig. 5.3 si vede che questo avviene a tensioni V_{DS} inferiori man mano che V_{GS} diminuisce. Questo e' dovuto al fatto che la tensione che polarizza inversamente la giunzione si aggiunge alla tensione di drain e quindi aumenta la differenza di potenziale effettiva ai capi del canale.

Normalmente i costruttori riportano la tensione BV_{DSS}, cioe' la tensione di breakdown tra drain e source, quando gate e source sono cortocircuitati tra loro (cioe' $V_{GS} = 0$). Questa tensione varia tipicamente tra i 20 ed i 50 V

Regione di Interdizione

Se $|V_{GS}| > |V_P|$ non si ha in linea di principio passaggio di corrente. In pratica e' presente una corrente, $I_{D,OFF}$, tipicamente dell'ordine dei nanoampere. Dello stesso ordine e' anche la corrente di gate all'interdizione, I_{GSS}, che rappresenta la corrente tra gate e source, con drain e source in corto circuito, quando $|V_{GS}| > |V_P|$.

Poiche' la corrente di gate e' trascurabile non ha praticamente senso esaminare le caratteristiche d'ingresso, diversamente da cio' che avviene nel transistor a giunzione. E' invece utile esaminare la transcaratteristica, cioe' la relazione che lega la corrente di drain alla

tensione V_{GS}. Si ha una relazione del tipo

$$I_D = I_{DSS}(1 - \frac{V_{GS}}{V_P})^2(1 + \lambda V_{DS}) \tag{5.3}$$

Le curve risultanti sono riportate nella Fig 5.5. Poiche' il parametro λ e' piccolo (dell'ordine di $10^{-2}\ V^{-1}$) la dipendenza da V_{DS} e' piccola. Cio' corrisponde all'andamento orizzontale delle curve nella Fig. 5.3. In pratica si puo' quindi considerare:

$$I_D \approx I_{DSS}(1 - \frac{V_{GS}}{V_P})^2 \tag{5.4}$$

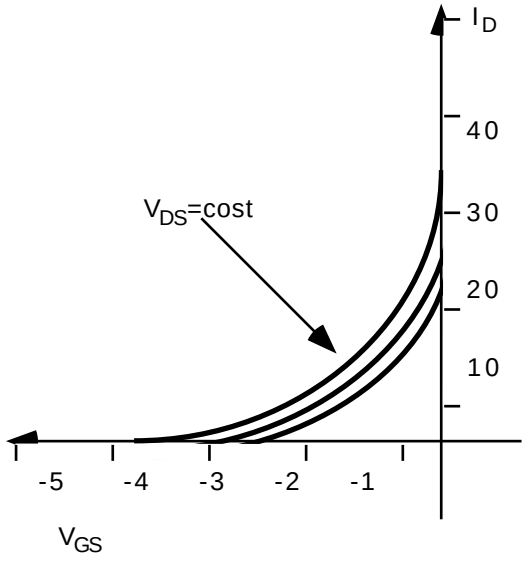

Figura 5.5: *Transcaratteristica*

5.3 Amplificatori con JFET

Il JFET puo' essere utilizzato per costruire amplificatori e, naturalmente, esistono due possibili configurazioni, a source comune (corrispondente all'amplificatore ad emettitore comune), e a drain comune (corrispondente al circuito a collettore comune).

Anche in questo caso conviene prima studiare la polarizzazione del transistor, e poi le prestazioni dinamiche del circuito utilizzando l'approssimazione per piccoli segnali, cioe' un modello lineare del FET.

5.3.1 Punto di lavoro di un JFET

Consideriamo il circuito in Fig 5.6a. Poichè $I_G = 0$, non vi è caduta di tensione ai capi di R_G, quindi, ricordando che $I_D = I_S$, abbiamo:

$$V_{GS} + I_D R_S = 0 \tag{5.5}$$

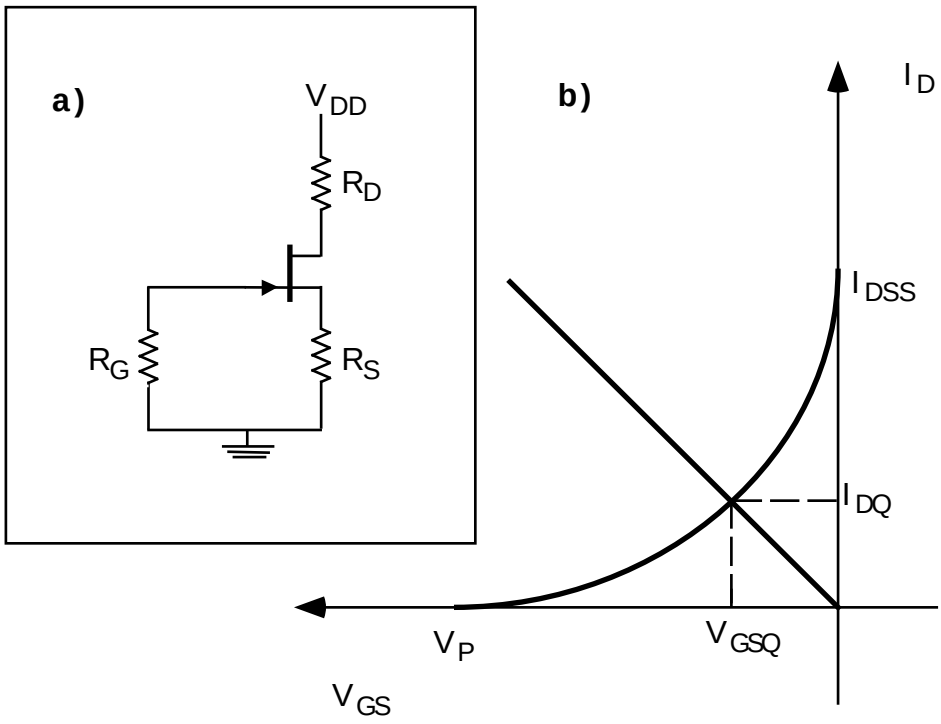

Figura 5.6: *a) Rete di polarizzazione; b) Retta di carico*

$$I_D = \frac{-V_{GS}}{R_S} \tag{5.6}$$

L'equazione 5.6 corrisponde a una retta di pendenza $-1/R_S$ nel piano delle transcaratteristiche e consente di individuare il punto di lavoro I_{DQ},V_{GSQ}. Possiamo inoltre scrivere

$$-V_{DD} + I_D R_D + V_{DS} + I_D R_s = 0 \tag{5.7}$$

Sostituendo I_{DQ} nella 5.7 si trova quindi la V_{DSQ}. Possiamo anche determinare V_{DSQ} tracciando la retta 5.7 sulle caratteristiche d'uscita (Fig. 5.3) e trovare il punto di lavoro come l'intersezione con la curva $V_{GS} = V_{GSQ}$.

5.3.2 Il modello per piccoli segnali

Consideriamo il circuito in Fig.5.7a. Se il transistor e' nella regione di saturazione il suo equivalente per piccoli segnali (a bassa frequenza) e' chiaramente il circuito di Fig.5.7b.

Si hanno quindi due parametri, g_m e r_{ds}. Ora

$$g_m \equiv \frac{di_D}{dv_{GS}}|_{v_{DS}=V_{DSQ}} = \frac{i_d}{v_{gs}}|_{v_{ds}=0} \tag{5.8}$$

utilizzando l'equazione 5.4 si ottiene

$$i_D = I_{DSS}(1 - \frac{v_{gs}}{V_P})^2 \tag{5.9}$$

e quindi

$$g_m = \frac{-2I_{DSS}}{V_P}(1 - \frac{V_{GSQ}}{V_P}) \tag{5.10}$$

Ma d'altra parte

$$(1 - \frac{V_{GSQ}}{V_p})^2 = \frac{I_{DQ}}{I_{DSS}} \tag{5.11}$$

per cui si ottiene in definitiva

$$g_m = \pm\frac{2}{V_p}\sqrt{I_{DQ}I_{DSS}} \tag{5.12}$$

ma g_m deve essere positivo, quindi si ha un'unica soluzione valida

$$g_m = -\frac{2I_{DSS}}{V_P}\sqrt{\frac{I_{DQ}}{I_{DSS}}} \equiv g_{m0}\sqrt{\frac{I_{DQ}}{I_{DSS}}} \tag{5.13}$$

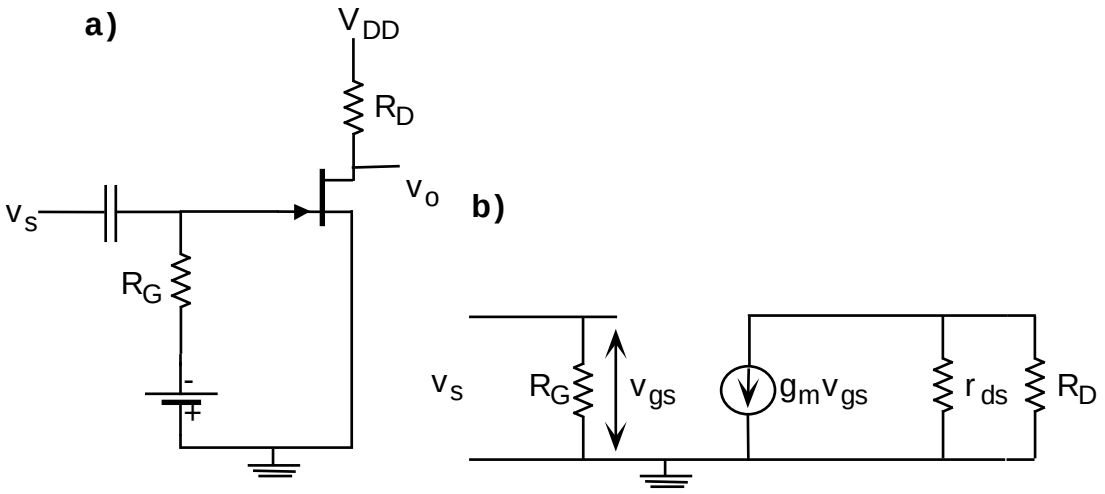

Figura 5.7: *a) Circuito reale; b) Equivalente per piccoli segnali*

Il parametro r_{ds} è invece definito come

$$\frac{1}{r_{ds}} \equiv g_{ds} = \frac{di_D}{dv_{DS}}|_{v_{gs}=V_{GSQ}} = \frac{i_d}{v_{ds}}|_{v_{gs}=0} \tag{5.14}$$

Chiaramente, a livello di approssimazione dell'equazione 5.4, il parametro r_{ds} viene infinito. Dobbiamo partire quindi dalla relazione non approssimata 5.3, con cui si ottiene

$$r_{ds} = \frac{1 + \lambda V_{DSQ}}{\lambda I_{DQ}} \approx \frac{1}{\lambda I_{DQ}} \tag{5.15}$$

Siamo quindi in una situazione abbastanza simile a quella del transistor a giunzione: la resistenza d'uscita r_{ds} e' molto grande (come lo era $1/h_{oe}$) e, ai fini di calcoli approssimati, puo' essere posta ad infinito.

Questo modello e' naturalmente valido quando sia possibile trascurare le capacita' parassite tra gate e drain e tra gate e source, cioe' a bassa-media frequenza. Non studieremo il comportamento ad alta frequenza, ma e' chiaro che esso sara' caratterizzato da un andamento tipo passa-basso.

5.3.3 Analisi dell'amplificatore common-source

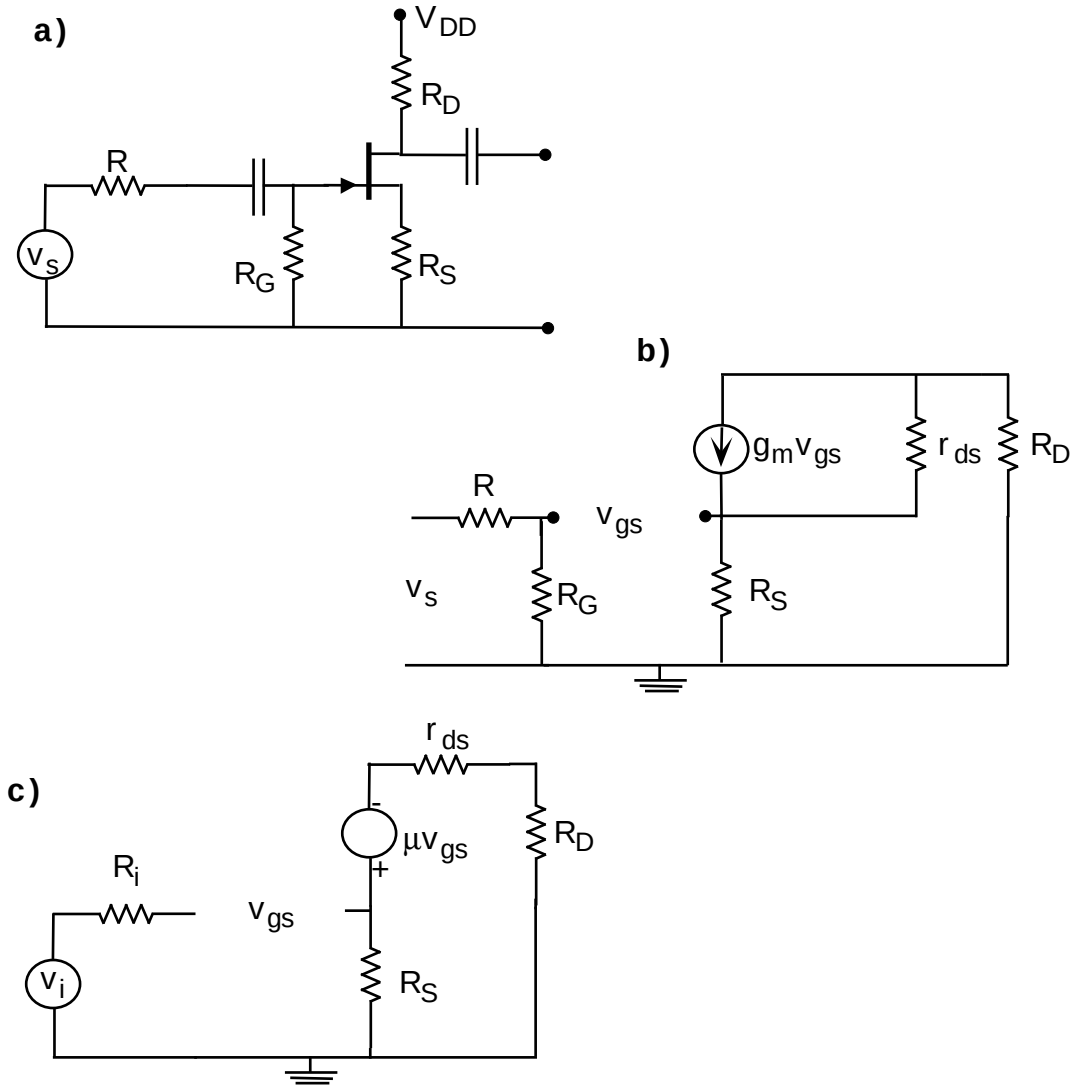

Figura 5.8: *a) Amplificatore common-source; b) Equivalente per piccoli segnali c) Dopo l'applicazione del teorema di Thevenin*

Siamo ora in condizione di studiare le prestazioni dell'amplificatore a source comune, nel caso piu' generale in cui si abbia una resistenza sia sul drain che sul source (Fig. 5.8a); il circuito equivalente per piccoli segnali e' rappresentato in Fig. 5.8b. Applicando il teorema di Thevenin sia alla maglia d'ingresso che alla maglia d'uscita arriviamo alla configurazione in Fig. 5.8c, dove:

$$\mu = g_m r_{ds}$$

$$v_i = \frac{R_G}{R_G + R} v_s$$

$$R_i = \frac{R_G R}{R_G + R}$$

Si ha allora, per la maglia d'uscita:

$$i_d R_D + i_d r_{ds} - \mu v_{gs} + i_d R_s = 0$$

e, per la maglia di ingresso:

$$v_i = v_{gs} + i_d R_s$$

cioè

$$v_{gs} = v_i - i_d R_s$$

Combinando le due equazioni si ricava

$$i_d = \frac{\mu}{r_{ds} + R_D + (1+\mu)R_S} v_i$$

Poichè $v_0 = -i_d R_D$ si ha:

$$A_V = \frac{-\mu R_D}{r_{ds} + R_D + (1+\mu)R_s} \tag{5.16}$$

La resistenza d'uscita R_0' (inclusa R_D) si puo' calcolare dal rapporto tra tensione d'uscita v_0 e la corrente di corto circuito i_{sc}:

$$v_0 = -i_d R_D = \frac{-\mu R_D}{r_{ds} + R_D + (1+\mu)R_s} v_i$$

$$i_{sc} = \frac{\mu}{r_{ds} + (1+\mu)R_S} V_i$$

Quindi

$$R_0' = \frac{R_D[r_{ds} + (1+\mu)R_s]}{R_D + R_{ds} + R_s(1+\mu)}$$

cioè è il parallelo tra R_D e la resistenza R_0, dove

$$R_0 = r_{ds} + (1+\mu)R_S \tag{5.17}$$

che rappresenta quindi la resistenza d'uscita del circuito a monte di R_D.

Naturalmente la resistenza d'ingresso e' infinita e di conseguenza non ha senso parlare di amplificazione di corrente.

Le prestazioni di un circuito senza resistenza R_S possono essere ricavate da qui semplicemente mandando a zero R_S: si ha quindi

$$A_v = \frac{-\mu R_D}{r_{ds} + R_D} = \frac{-g_m R_D}{1 + \frac{R_D}{r_{ds}}}$$

$$R_0 = r_{ds}$$

Questo risultato si applica naturalmente anche per un amplificatore con un condensatore C_S in parallelo a R_S, per frequenze

$$\nu \gg \frac{1}{2\pi R_S C_S}$$

5.3.4 Amplificatore common-drain (source-follower)

Si pone $R_D = 0$ e si preleva l'uscita sul source

$$i_d = \frac{\mu}{r_{ds} + (1+\mu)R_s} v_i$$

$$v_0 = i_d R_s = \frac{\mu R_s}{r_{ds} + (1+\mu)R_s} v_i$$

Quindi

$$A_V = \frac{v_0}{v_i} = \frac{\mu R_s}{r_{ds} + (1+\mu)R_s}$$

cioè, se $(1+\mu)R_S \gg r_{ds}$,

$$A_V \approx 1$$

La resistenza d'uscita, R_o, è data dal rapporto v_0/i_{sc}, dove

$$i_{sc} = \frac{\mu v_i}{r_{ds}}$$

Quindi

$$R'_o = \frac{R_s r_{ds}}{r_{ds} + (1+\mu)R_S} = \frac{R_S \frac{r_{ds}}{1+\mu}}{\frac{r_{ds}}{1+\mu} + R_S}$$

cioè è il parallelo tra R_S e

$$R_o = \frac{r_{ds}}{1+\mu} \approx \frac{1}{g_m}$$

5.4 Cenni sui transistors MOSFET

Il principio di funzionamento dei transistors MOSFET e' simile a quello dei JFET, con una fondamentale differenza: l'elettrodo di controllo (gate) funziona sul principio dell'induzione elettrostatica, quindi esso e' elettricamente isolato dal canale. Quindi la corrente di gate e' zero, e la resistenza d'ingresso infinita. Esistono due tipi di MOSFET: il cosidetto tipo *enhancement*, ed il tipo *depletion*.

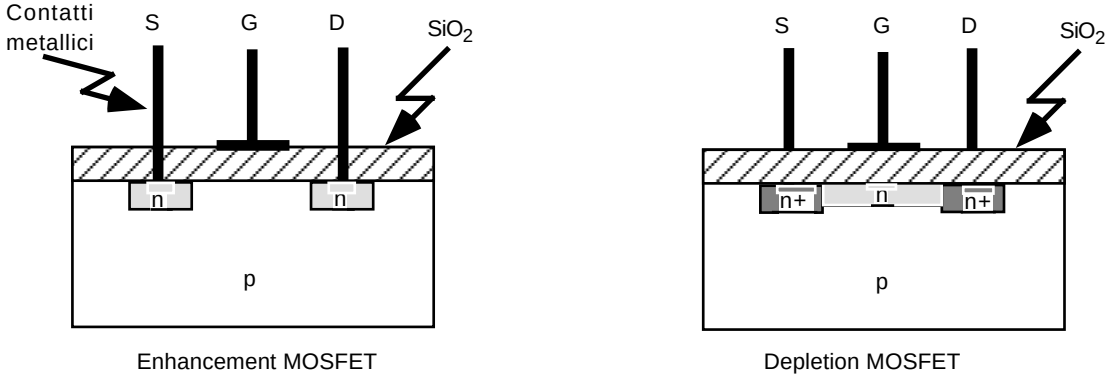

Figura 5.9: *a) Enhancement MOSFET; b) depletion MOSFET*

Enhancement MOSFET. Consideriamo il dispositivo in Fig. 5.9a: su un substrato di tipo p si hanno due inserzioni di tipo n che costituiscono gli elettrodi di source e drain, mentre il gate e' formato da uno strato metallico isolato dal substrato da uno strato isolante (tipicamente ossido di silicio). Supponiamo ora di polarizzare il gate positivamente rispetto al drain, al source ed al substrato (tutti messi a massa). Nel substrato tra D e S si formerà uno strato di cariche negative indotte, che costituirà un canale conduttivo. Si puo' facilmente comprendere che cio' avviene quando V_{GS} supera una certa tensione minima (tensione di soglia), cioe' si deve avere $V_{GS} > V_T$. Applichiamo ora una tensione tra drain e source. Finchè $V_{DS} < (V_{GS} - V_T)$ si ha un comportamento ohmico. Quando V_{DS} cresce, la caduta di tensione tra G e canale diminuisce, specie vicino al Drain, dove $V_{GD} = (V_{DS} - V_{GS})$. Quindi il canale si riduce e si arriva ad una situazione di saturazione.

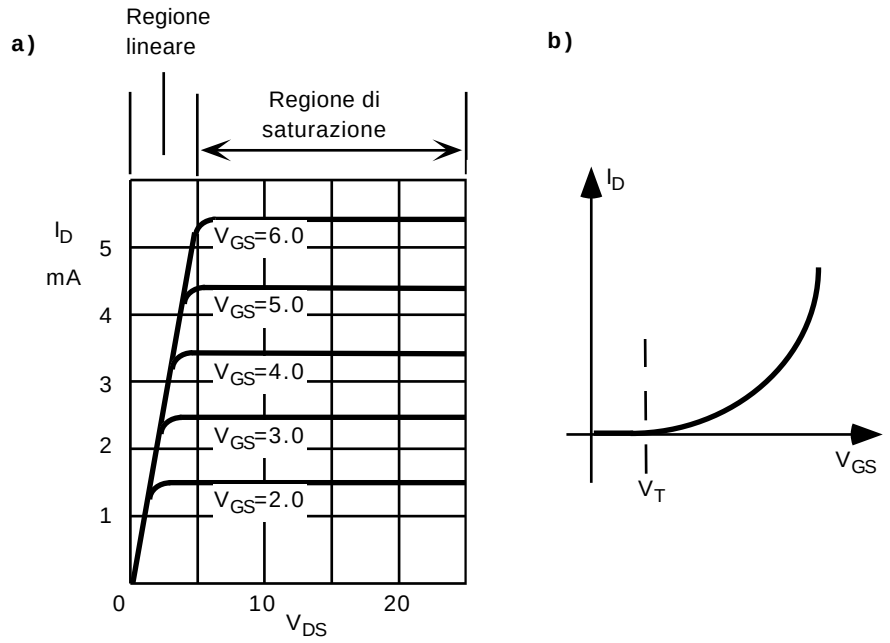

Figura 5.10: *a) Caratteristiche d'uscita dell'enhancement MOSFET; b) Transcaratteristica*

Le caratteristiche d'uscita sono quindi del tipo mostrato in Fig. 5.10a; la transcaratteristica e' invece data da

$$I_D = k\frac{W}{L}(V_{GS} - V_T)^2(1 + \lambda V_{DS})$$

dove L è la lunghezza del canale, W lo spessore, $k = \mu_n C_0/2$ e C_0 la capacità per unità di superficie; λ esprime la dipendenza di I da V_{DS} (effetto Early) ed è molto piccolo ($10^{-2}V^{-1}$) (Fig, 5.10. Poiche' il canale conduttivo e' di tipo n questo transistor prende il nome di NMOS; si può ovviamente costruire un dispositivo con canale di tipo p (PMOS). Ovviamente i due dispositivi avranno prestazioni diverse perche' le mobilita' di elettroni e lacune sono diverse.

Depletion MOSFET. Questo principio di funzionamento puo' essere modificato costruendo un dispositivo come quello schematizzato in Fig. 5.9b. Ora, quando $V_{GS} = 0$ si ha

gia' un canale conduttivo, quindi se applichiamo tra drain e source una tensione positiva ci aspettiamo dapprima un comportamento ohmico e, al crescere di V_{DS}, un fenomeno di saturazione. Se ora applichiamo anche al gate una tensione, il canale si arricchisce o si impoverisce di cariche, a seconda del segno di V_{GS}, e quindi la curva della I_{DS} in funzione di V_{DS} si sposta verso l'alto o verso il basso (Fig. 5.11).

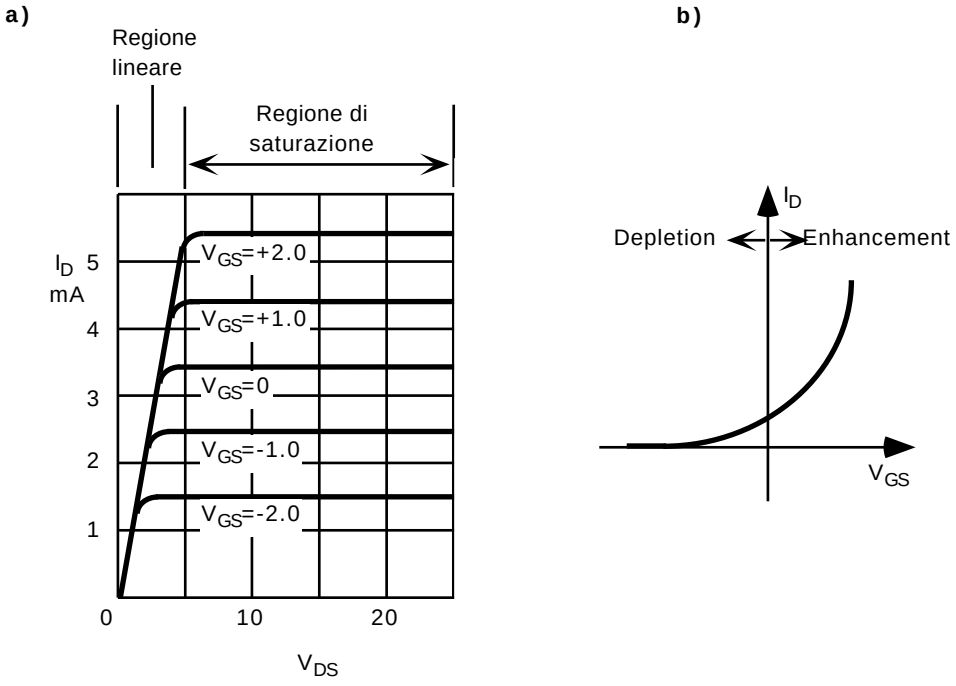

Figura 5.11: *Caratteristiche del depletion MOSFET*

Questi componenti offrono i vantaggi cui abbiamo gia' accennato all'inizio: altissima resistenza d'ingresso, cioe' assenza di corrente di gate. Si prestano, come i JFET, alla realizzazione di resistenze (e di capacita'), e, piu' in generale, alla realizzazione di dispositivi ad alto livello di integrazione

Conviene accennare anche al fatto che essi sono molto delicati. Infatti il gate del Mosfet costituisce una capacità priva di resistenza attraverso cui si possa scaricare l'elettricità statica eventualmente accumulata. Questo significa che possono facilmente danneggiarsi se maneggiati senza precauzioni. Ricordiamo infatti che, da un punto di vista elettrico, l'uomo può essere schematizzato come una capacita' dell'ordine di 100 pF con in serie una resistenza di qualche $kohm$. A causa della elettricita' statica presente nell'atmosfera questa capacita' puo' caricarsi ad una tensione di molti $kVolts$. Quindi, toccando il transistor, potremmo portare questa differenza di potenziale tra gate e substrato, provocando un'immediata scarica attraverso il dielettrico, che, come abbiamo detto, e' uno strato sottile di ossido di silicio, incapace di sopportare una differenza di potenziale cosi' alta. Questo e' il motivo per cui prima di manipolare integrati di tipo MOS l'operatore deve scaricare a terra, attraverso un buon conduttore, l'elettricita' statica eventualmente accumulata sul proprio corpo.

Capitolo 6

Amplificatori operazionali

6.1 Introduzione

Gli amplificatori integrati (comunemente chiamati amplificatori operazionali) sono divenuti ormai, grazie ai progressi nel campo dell'integrazione, uno dei componenti essenziali dell'elettronica. Essi sono degli amplificatori a molti stadi, accoppiati in continua, quasi sempre con ingresso differenziale, caratterizzati da un'altissima amplificazione di tensione ($10^5 \div 10^6$), alta resistenza d'ingresso e bassa resistenza d'uscita, realizzati su un unico circuito integrato. I costi di produzione sono ormai bassissimi e questo componente, che si presta, come vedremo, a svariati usi, sostituisce in moltissime applicazioni gli amplificatori convenzionali a transistors.

6.2 Caratteristiche generali

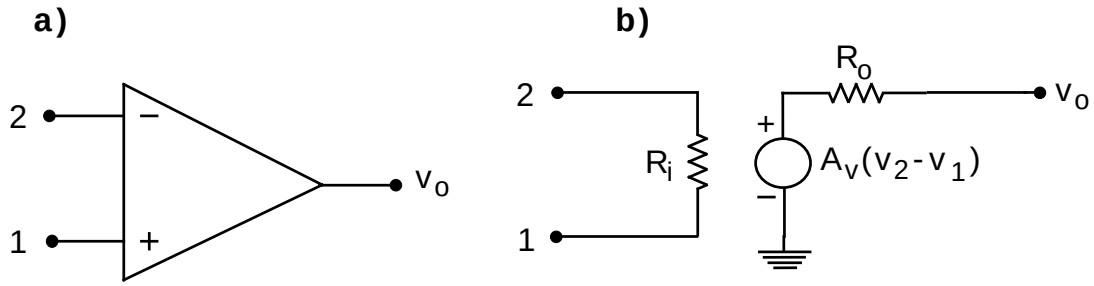

Figura 6.1: *a) Simbolo dell'amplificatore operazionale; b) Circuito equivalente*

Nella Fig. 6.1a e' riportato il simbolo circuitale usato per gli amplificatori operazionali con ingresso differenziale; la Fig. 6.1b mostra il circuito equivalente. Come abbiamo detto A_v ed R_i sono molto grandi, mentre R_o e' molto piccolo; possiamo quindi partire da un'ipotesi semplificativa (modello ideale), in cui:

$$\begin{aligned} A_v &= -\infty \\ R_i &= \infty \\ R_o &= 0 \end{aligned}$$

E' anche implicito nel modello ideale che il Rapporto di Reiezione di Modo Comune (CMMR) sia infinito [1].

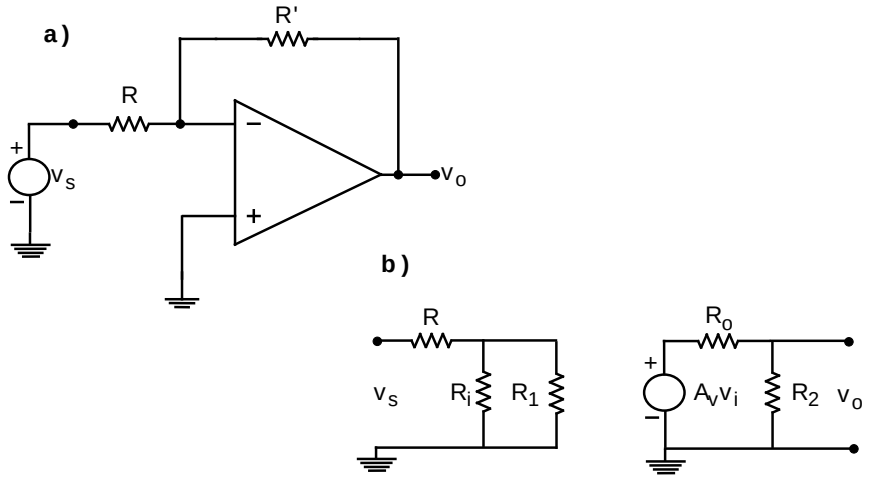

Figura 6.2: *a) Amplificatore invertente; b) Circuito equivalente con il teorema di Miller*

Possiamo immediatamente comprendere l'utilita' di questo componente attraverso alcuni semplici esempi. Consideriamo anzitutto il circuito in Fig. 6.2a, in cui abbiamo posto a massa l'ingresso non invertente, mentre l'altro ingresso e' collegato all'uscita attraverso una resistenza di reazione R'. Possiamo studiare il circuito utilizzando il teorema di Miller (Fig. 6.2b), dove

$$R_1 = \frac{R'}{1 - A_v} \quad R_2 = \frac{R'}{1 - \frac{1}{A_v}} \simeq R' \tag{6.1}$$

Poiche' $R' \gg R_o$ il suo effetto sulla maglia d'uscita e' trascurabile, mentre R_i e' trascurabile rispetto ad R_1, per cui abbiamo:

$$\begin{aligned} v_o &\simeq A_v(v_2 - v_1) = A_v v_2 \\ v_2 &= \frac{R_1}{R + R_1} v_s \end{aligned}$$

Possiamo quindi ricavare l'amplificazione con reazione

$$\begin{aligned} A_{vf} &= \frac{v_o}{v_s} \\ &= A_v \frac{R_1}{R + R_1} \\ &= A_v \frac{\frac{R'}{1-A_v}}{R + \frac{R'}{1-A_v}} \\ &\simeq -\frac{R'}{R} \end{aligned}$$

[1] Si noti che stiamo utilizzando, qui e nel seguito, una notazione per cui A_v è intrinsecamente negativo, come in molti manuali. Questo è a volte fonte di confusione negli studenti. Si può evitare ogni confusione ricordando sempre che l'amplificatore operazione inverte il segno di ciò che entra nel morsetto negativo.

avendo sfruttato il fatto che $A_v \to -\infty$.
Come si vede l'amplificazione e' negativa, da cui il nome di *amplificatore invertente*, dipende solo dal valore delle due resistenze esterne, e non piu' dai parametri dell'amplificatore. E' anche importante notare che la tensione v_2 e' praticamente zero, come la tensione v_1: la differenza di potenziale tra i due ingressi e' forzata a zero proprio dall'effetto Miller, che introduce tra essi una resistenza bassissima. Tuttavia la resistenza d'ingresso vista dal segnale e'

$$\frac{v_s}{i_s} = R + R_1 \simeq R \tag{6.2}$$

L'ingresso 2 e' quindi quello che si dice una *massa virtuale*, cioe' un nodo la cui tensione e' sostanzialmente zero, pur essendo connesso alla massa attraverso una resistenza molto elevata (R_i). In sostanza la presenza della resistenza di reazione R' *obbliga* i due ingressi dell'operazionale a portarsi allo stesso potenziale: qualunque tentativo noi facciamo di variare in un verso o nell'altro la tensione all'ingresso 2 provoca una reazione opposta che, attraverso la resistenza R', riporta il terminale d'ingresso 2 allo stato di partenza.
E' possibile ricalcolare l'amplificazione del circuito della Fig. 6.2a usando a priori questa proprieta'. Diciamo allora che:

1. la tensione v_2 (ingresso (-)) e' uguale alla tensione v_1 (ingresso (+)), cioe', nel caso in esame, pari a zero;

2. non entra corrente nell'operazionale (perche' la resistenza d'ingresso e' molto elevata), quindi la corrente i_s che circola in R e' la stessa che circola in R'.

Possiamo scrivere l'equazione del nodo di ingresso (-), assegnando arbitrariamente un verso alla corrente i_s

$$\begin{aligned} \frac{(v_s - 0)}{R} &= \frac{(0 - v_0)}{R'} \\ \frac{v_s}{R} &= \frac{-v_0}{R'} \end{aligned}$$

Da cui si ricava immediatamente

$$\frac{v_o}{v_s} = -\frac{R'}{R}$$

cioe' ritroviamo in modo immediato il risultato gia' noto.
Queste due ipotesi semplificative (uguaglianza di tensione tra i morsetti d'ingresso e resistenza d'ingresso infinita) ci consentiranno di calcolare in modo molto semplice il comportamento di moltissimi circuiti basati sull'amplificatore operazionale. E' bene pero' sempre ricordare che si tratta di una *approssimazione*: vedremo in seguito dei casi in cui questa approssimazione ci porterebbe a risultati grossolanamente errati. Ci conviene quindi calcolare anche l'amplificazione *esatta*, nel caso piu' generale, in cui al posto di R ed R' poniamo due generiche impedenze Z e Z' (Fig.6.3).

Applicando il metodo dei nodi all'ingresso 2 ed all'uscita abbiamo

$$\frac{V_s - V_i}{Z} = \frac{V_i}{R_i} + \frac{V_i - V_o}{Z'} \tag{6.3}$$

$$0 = \frac{A_V V_i - V_o}{R_o} + \frac{V_i - V_o}{Z'} \tag{6.4}$$

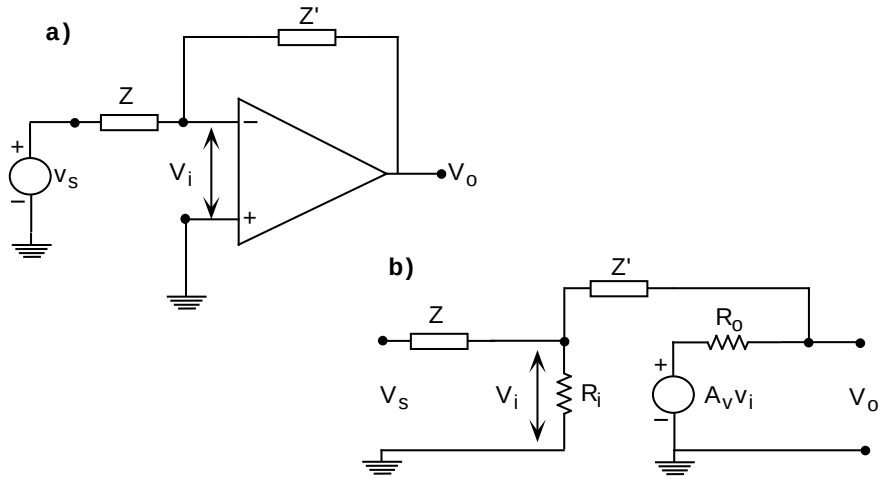

Figura 6.3: *a) Il caso di generiche impedenze; b) Circuito equivalente*

dove naturalmente stiamo lavorando nel dominio complesso. Eliminando V_i tra le due equazioni, si ottiene con alcuni passaggi

$$\frac{V_o}{V_s} = \frac{Y}{-Y' + \frac{Y_o + Y'}{A_V Y_o + Y'}(Y + Y_i + Y')} \tag{6.5}$$

dove, per semplicita' di scrittura, abbiamo sostituito ogni impedenza con la corrispondente ammettenza. E' facile verificare che, per $A_V \to \infty$, la 6.5 si riduce a

$$\frac{V_o}{V_s} = -\frac{Y}{Y'} = -\frac{Z'}{Z} \tag{6.6}$$

cioe' la generalizzazione del risultato ottenuto nel caso di impedenze puramente resisistive. Il risultato che abbiamo ottenuto con la 6.5 ci sara' utile in seguito.

Vediamo ora il circuito di Fig. 6.4a. Anche qui, abbiamo una rete di reazione sull'ingresso invertente, ma ora il segnale che vogliamo amplificare e' connesso al morsetto non invertente. Lo schema equivalente e' mostrato in Fig. 6.4b; abbiamo

$$v_i = v_2 - v_1 = v_2 - v_s \tag{6.7}$$

L'equazione del nodo 2 e' data da

$$\frac{v_2}{R} = \frac{v_o - v_2}{R'} + \frac{v_s - v_2}{R_i} \tag{6.8}$$

da cui si ricava

$$v_2\left(\frac{1}{R} + \frac{1}{R'} + \frac{1}{R_i}\right) = \frac{v_o}{R'} + \frac{v_s}{R_i} \tag{6.9}$$

trascurando $1/R_i$ a primo membro e trascurando v_s/R_i al secondo membro si ricava

$$v_2 = \frac{R}{R + R'} v_o \tag{6.10}$$

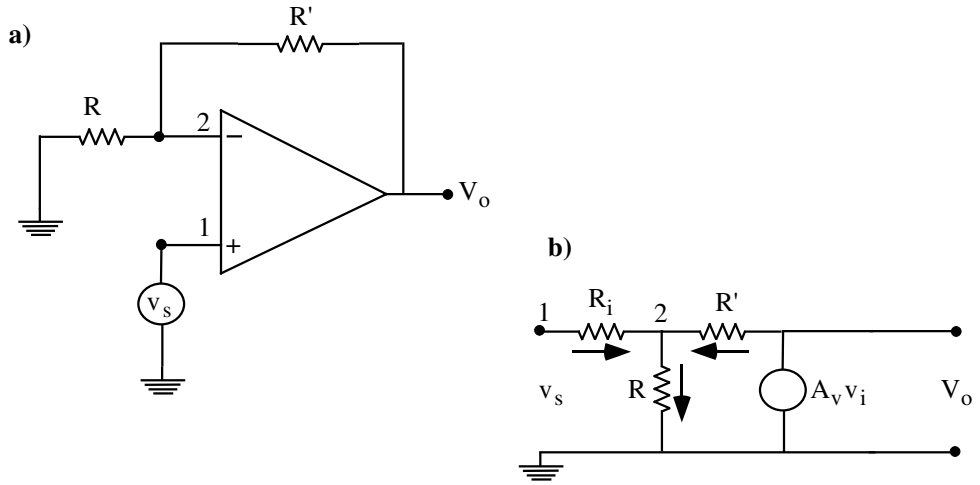

Figura 6.4: *a) Amplificatore non-invertente; b) Circuito equivalente; le frecce indicano i versi delle correnti scelti arbitrariamente*

Abbiamo ora

$$\begin{aligned} v_o &= A_v v_i \\ &= A_v v_2 - A_v v_s \\ &= \frac{R}{R+R'} A_v v_o - A_v V_s \end{aligned}$$

da cui possiamo finalmente ricavare

$$A_{vf} = \frac{v_o}{v_s} = \frac{A_v}{\frac{R}{R+R'} A_v - 1} \tag{6.11}$$

Poiché $A_v \to -\infty$

$$A_{vf} \simeq \frac{R+R'}{R} \tag{6.12}$$

Cioe', anche in questo caso, abbiamo un'amplificazione che dipende solo dai valori delle due resistenze esterne.

Si noti che, anche in questo caso, i due ingressi dell'operazionale sono allo stesso potenziale; infatti

$$\frac{v_2}{v_s} = \frac{v_2}{v_o}\frac{v_o}{v_s} = \frac{R}{R+R'}\frac{R+R'}{R} = 1 \tag{6.13}$$

Di nuovo, cio' e' dovuto alla reazione negativa presente nel circuito. Anche in questo caso quindi sarebbe stato possibile calcolare le prestazioni del circuito assumendo a priori l'uguaglianza di tensione tra i due ingressi, trovando immediatamente

$$\frac{v_o}{v_s} = \frac{v_o}{v_2} = \frac{R+R'}{R} \tag{6.14}$$

A differenza del circuito precedente, la resistenza d'ingresso vista dal segnale e' in questo caso proprio R_i, cioe' la resistenza d'ingresso dell'amplificatore operazionale. Questo schema prende il nome di *amplificatore non-invertente*

Ponendo $R \gg R'$ si ha $A_{vf} = 1$; si realizza in modo immediato un ottimo inseguitore di tensione, con bassa resistenza d'uscita ed alta resistenza d'ingresso. Spesso cio' e' fatto ponendo semplicemente $R' = 0,\ R = \infty$, cioe' semplicemente collegando l'ingresso invertente con l'uscita.

Da questi due esempi abbiamo quindi compreso che l'amplificatore operazionale ha grandissime potenzialita', semplicita' d'uso e consente di realizzare amplificatori molto stabili. Vedremo nei prossimi paragrafi molte applicazioni in cui sfrutteremo la reazione negativa; potremo quindi calcolarne facilmente le prestazioni utilizzando le ipotesi semplificative, la cui validita' abbiamo qui constatato.

6.3 Amplificatori operazionali reali

Un amplificatore operazionale e' in genere costituito da vari stadi: dopo uno stadio d'ingresso di tipo differenziale, si hanno alcuni stadi di amplificazione in cascata, l'ultimo dei quali e' un inseguitore di tensione che fornisce una bassa resistenza d'uscita. Sono spesso necessarie due tensioni di alimentazione (tipicamente $\pm 10 \div 15\ V$) che delimitano la massima escursione della tensione d'uscita[2]. E' chiaro quindi che le prestazioni non sono esattamente quelle ideali finora considerate, e si deve tener conto di una serie di fattori che possono avere influenza sul comportamento reale del circuito.

Offset di tensione e di corrente

Nei due terminali d'ingresso scorrono delle correnti, che prendono il nome di correnti di ingresso. Nel caso di amplificatori con ingresso BJT esse corrispondono naturalmente alla corrente di base dei due transistors d'ingresso, essenziale per il funzionamento dei medesimi; analogamente, se i transistors d'ingresso sono di tipo FET, le correnti di ingresso non sono altro che le correnti di gate. Nel primo caso esse sono dell'ordine di decine di nA, mentre nel secondo sono dell'ordine delle decine di pA (cioe' 1000 volte piu' piccole).

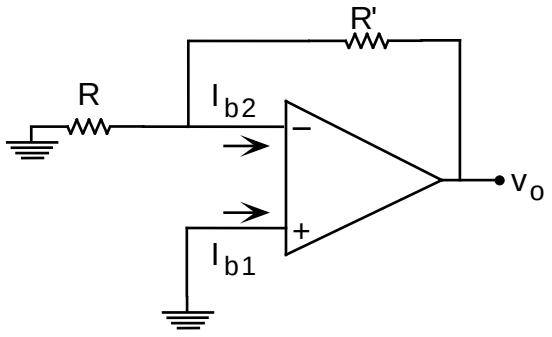

Figura 6.5: *Correnti d'ingresso: il verso reale dipende dal tipo di giunzione (pn o np)*

Il significato di queste correnti e' che esse producono delle cadute di tensione sulle resistenze che compongono la rete di reazione, o sulla resistenza d'uscita del generatore di segnale collegato all'ingresso. Prendiamo ad esempio il circuito di Fig. 6.5: la corrente I_{b_2} scorre

[2] In realta' la tensione d'uscita e' in genere costretta in un intervallo un po' piu' ristretto

sulle resistenze R ed R', quindi il morsetto 2 non e' piu' alla stessa tensione del morsetto 1 e ovviamente questa differenza si presenta amplificata sull'uscita. Se la tensione v_o e' comunque piccola possiamo approssimativamente considerare R ed R' in parallelo, e quindi il morsetto 2 si porta alla tensione

$$v_2 \simeq \frac{RR'}{R+R'} I_{b_2} \tag{6.15}$$

Il valore assoluto di v_2 e' legato quindi a R ed R'; dipende naturalmente dall'utilizzo che dobbiamo fare del circuito stabilire se questo produca o meno un effetto *intollerabile* sull'uscita. Tipicamente esso e' del tutto trascurabile nel caso di amplificatori con ingresso FET, mentre puo' non esserlo per ingressi BJT. Possiamo minimizzare l'effetto scegliendo valori di R ed R' non troppo alti, ma possiamo anche , se necessario, curare il problema alla radice bilanciando il carico sui due ingressi: nel circuito di Fig. 6.6 la resistenza di $9.1k$, attraversata dalla corrente I_{b_1}, riporta, in assenza di segnale esterno, entrambi gli ingressi alla stessa tensione.

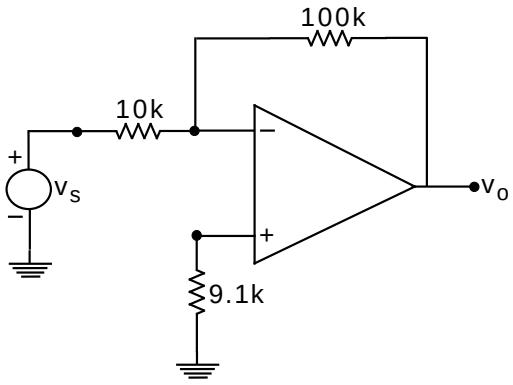

Figura 6.6:

Tuttavia e' impensabile poter realizzare due ingressi esattamente identici: piccole differenze costruttive sono inevitabili e quindi le due correnti, I_{b_1} e I_{b_2}, non saranno mai uguali. Cio' significa che, pur in presenza di carichi assolutamente bilanciati, si avra' una tensione d'uscita diversa da zero anche in assenza di un segnale d'ingresso. Si definisce quindi la corrente di *bias* come la media di I_{b_1} e I_{b_2}, mentre il valore assoluto della differenza

$$I_{os} = |I_{b_1} - I_{b_2}| \tag{6.16}$$

prende il nome di *corrente di offset*: tipicamente e' tra il 10 ed il 50% della corrente di bias.

Piu' in generale le due giunzioni base-emettitore dei due transistors di ingresso possono avere caratteristiche diverse, ma operativamente cio' si traduce sempre in un *offset* di tensione (continua) presente all'uscita anche in assenza di segnale d'ingresso: si definisce allora la tensione di offset d'ingresso (V_{os}), come la tensione che occorre porre tra i due ingressi, attraverso carichi uguali, per azzerare perfettamente l'uscita. In genere V_{os} e' dell'ordine di pochi mV; questo significa che ponendo entrambi gli ingressi a massa, in assenza di reazione, la tensione d'uscita andrebbe a decine o centinaia di Volts! Questo in

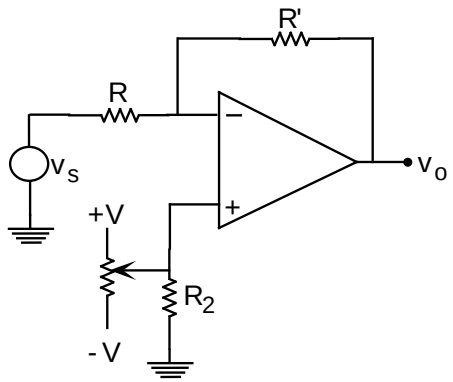

Figura 6.7:

realta' non puo' avvenire, ma l'uscita si porta alla massima (o minima) tensione consentita dall'alimentazione. Per misurare in pratica la tensione di offset d'ingresso e' piu' conveniente ridurre l'amplificazione introducendo la reazione negativa: si puo' quindi utilizzare un circuito come quello della Fig. 6.5: si avrebbe

$$V_{os} \simeq -v_o \frac{R}{R'} \tag{6.17}$$

questo purche' sia possibile trascurare l'effetto dello sbilanciamento dei due ingressi (in sostanza, se $RI_{b2} << V_{os}$). Per una misura piu' precisa e' opportuno bilanciare i carichi, come in Fig. 6.6.
Se necessario, possiamo neutralizzare la V_{os} con un opportuno partitore posto all'ingresso non utilizzato, come per esempio in Fig. 6.7: prima di collegare il segnale v_s si regola il potenziometro in modo che l'uscita sia esattamente zero.
Peraltro molti operazionali in commercio hanno ingressi appositi per azzerare l'offset

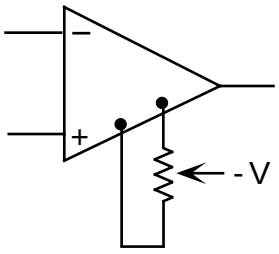

Figura 6.8:

(Fig. 6.8), senza impegnare gli ingressi ordinari. Il costruttore fornisce, tra le varie specifiche, anche indicazioni sul valore piu' adatto del partitore di correzione. Da quanto detto finora dovrebbe essere chiaro che e' opportuno eseguire questa operazione con le stesse condizioni di carico con cui poi si utilizzera' l'amplificatore.
Se e' vero quindi che e' sempre possibile riportare l'operazionale ad una condizione di perfetto bilanciamento diviene importante la stabilita' temporale di questa condizione, soprattutto in funzione di variazioni di temperatura, altrimenti saremmo costretti ad inseguire queste derive agendo continuamente sul partitore di azzeramento. Tipicamente la

deriva termica della tensione di offset e' di alcuni μV per grado centigrado (e di alcuni pA per grado relativamente alla I_{os}); per usi normali non si hanno quindi particolari problemi in questo senso. Le correnti di ingresso possono essere misurate , utilizzando alternativa-

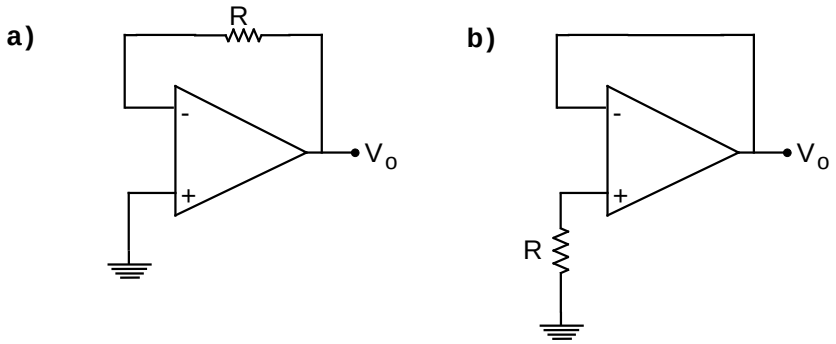

Figura 6.9: *a) Misura della corrente I_{b2}; b) Misura della corrente I_{b1}*

mente i due schemi in Fig. 6.9a e 6.9b, purche' si possa trascurare l'effetto della V_{os}. Nel primo caso, se $RI_{b2} >> V_{os}$, si ha

$$v_o \simeq I_{b2}R \tag{6.18}$$

mentre nel secondo caso, se $RI_{b1} >> V_{os}$,

$$v_o = -I_{b1}R \tag{6.19}$$

Resistenza d'ingresso e di uscita

La resistenza d'ingresso di operazionali realizzati con BJT e' dell'ordine di qualche $M\Omega$. Come abbiamo visto nel paragrafo precedente essa e' completamente mascherata dall'effetto Miller quando utilizziamo l'ingresso invertente, quindi per misurarla si deve utilizzare la configurazione non-invertente, ovvero, molto semplicemente, un inseguitore di tensione.
Tipiche resistenze d'ingresso di 10^{12} Ω si possono invece ottenere in operazionali con stadio d'ingresso a JFET.
Poiche', come abbiamio gia' detto, lo stadio d'uscita di un operazionale e' in genere un inseguitore di tensine, la resistenza d'uscita e' quella tipica di tali circuiti, cioe' dell'ordine dei 10 Ω. Questo non significa tuttavia che la corrente d'uscita possa divenire arbitrariamente grande: i normali operazionali possono erogare al massimo qualche decina di mA (questo valore dipende anche dalla tensione d'uscita), oltre i quali l'amplificatore puo' degradare le sue prestazioni o danneggiarsi.

Risposta in frequenza

Un amplificatore operazionale reale non puo' che avere una banda passante finita. Gli accoppiamenti in continua consentono effettivamente l'amplificazione fino a frequenza zero, ma, poiche' esso e' composto da molti stadi l'andamento ad alta frequenza dell'amplificazione e' dominato dalla presenza di numerosi poli nella funzione di trasferimento (dovuti alle capacita' parassite dei transistors presenti nel circuito), e il diagramma di Bode avra'

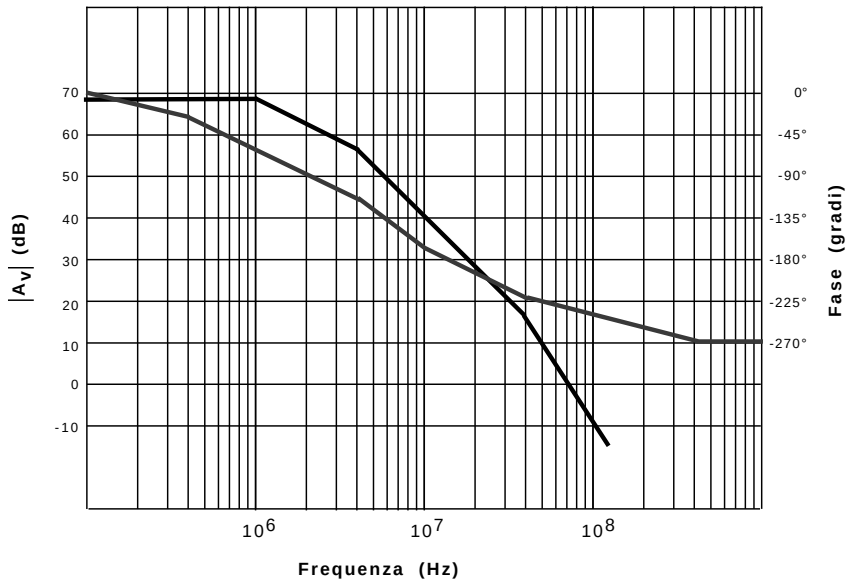

Figura 6.10: *Amplificazione di tensione ad anello aperto e sfasamento di un amplificatore operazionale*

quindi un andamento simile a quello mostrato in Fig. 6.10. Questo puo' creare grossi problemi di instabilita' nel circuito nel momento in cui utilizziamo l'operazionale con una rete di reazione negativa, cioe' abbiamo una funzione di trasferimento del tipo

$$A_{vf} = \frac{A}{1 + \beta A} \tag{6.20}$$

Infatti, a causa della rotazione di fase, al crescere della frequenza la reazione diviene *positiva* (quando lo sfasamento raggiunge 180^0), cioe' β diviene negativo. Se esiste una frequenza per cui $\beta A = -1$ l'amplificazione diverge, cioe' l'amplificatore puo' entrare in oscillazione. Di fatto quindi, gli inevitabili disturbi presenti su tutto lo spettro di frequenza sono sufficienti a trasformare il nostro amplificatore in un oscillatore.
Possiamo studiare quantitativamente questo fenomeno prendendo ad esempio il circuito di Fig. 6.4, cioe' l'amplificatore non invertente. In questo caso la reazione e' di tipo *tensione-serie*, quindi la grandezza stabilizzata e' proprio l'amplificazione di tensione. Il fattore di reazione β e'

$$\beta = \frac{R}{R + R'} \tag{6.21}$$

e si ottiene quindi

$$A_{Vf} = \frac{A_V}{1 + \frac{R}{R + R'} A_V} \tag{6.22}$$

Nel caso mostrato in Figura 6.10 si vede che la fase raggiunge -180^o a circa 12 Mhz: a questa frequenza il modulo di A_V vale circa 36 dB. Quindi, se a quella frequenza

$$\frac{R}{R + R'} |A_V| = 1 \tag{6.23}$$

l'amplificatore non e' stabile e puo' oscillare. Si verifica immediatamente che cio' corrisponde a

$$\frac{R+R'}{R} = 63 \tag{6.24}$$

In definitiva, se il suddetto rapporto e' inferiore a 63 la 6.22 puo' divergere a qualche frequenza superiore a 12 Mhz. In realta', se si vuole avere un margine di sicurezza, si deve richiedere che l'amplificazione con reazione sia ben piu' alta: usualmente si richiede un margine di fase di 45^o, e questo corrisponde nel caso in esame ad un limite inferiore di 316 nell'amplificazione.
Limiti analoghi si possono ricavare nella configurazione invertente, che corrisponde ad una reazione di tipo *tensione-parallelo*.

Questa e' chiaramente una limitazione molto severa nell'uso dell'amplificatore operazionale: in molti casi e' quindi necessario aggiungere al circuito una rete RC per modificare in modo opportuno l'andamento del modulo e della fase di A_V. Questa operazione e' detta *compensazione* e puo' essere sostanzialmente fatta in 3 modi:

1. *Compensazione con polo dominante*

   Si inserisce un ulteriore polo nella funzione di trasferimento ad una frequenza molto piu' bassa di quelli esistenti.

2. *Compensazione con polo e zero*

   Si aggiunge sia un polo che uno zero nella funzione di trasferimento: la frequenza dello zero e' scelta in modo da cancellare il primo polo, cioe' $f_z = f_1$.

3. *Compensazione per anticipo di fase*

   Si aggiunge solo uno zero alla funzione di trasferimento: si ottiene cosi' un anticipo della fase e conseguentemente aumenta la frequenza per cui si ha la rotazione di -180^o.

Molti amplificatori operazionali in commercio sono gia' dotati internamente di una rete di compensazione, in genere con il metodo del polo dominante: in sostanza si aggiunge una capacita' tra l'ingresso e l'uscita di uno degli stadi di amplificazione. L'effetto Miller provvede a moltiplicare questa capacita' per un grosso fattore, introducendo quindi un polo a frequenza molto bassa (in molti casi circa 10 Hz): l'effetto e' quello mostrato in Fig 6.11. La rotazione di fase arriva ora a -180^o quando il modulo di A_V e' molto inferiore ad uno, eliminando quindi ogni pericolo di instabilita'.
Apparentemente la banda passante e' ora molto modesta: in realta' essa e' molto piu' grande in presenza di reazione (si ricordi infatti che il prodotto *banda-guadagno* e' costante). Quindi, se ad esempio l'amplificazione ad anello aperto e' 10^5 e la banda passante 10 Hz, si ha un prodotto *banda-guadagno* di 10^6: un amplificatore che amplifichi 10 ha quindi una banda di 100 kHz.
Esistono comunque in commercio amplificatori con polo dominante a frequenza piu' elevata (per esempio l'integrato LF157 della National ha una frequenza di taglio ad anello aperto attorno ai 100 Hz).

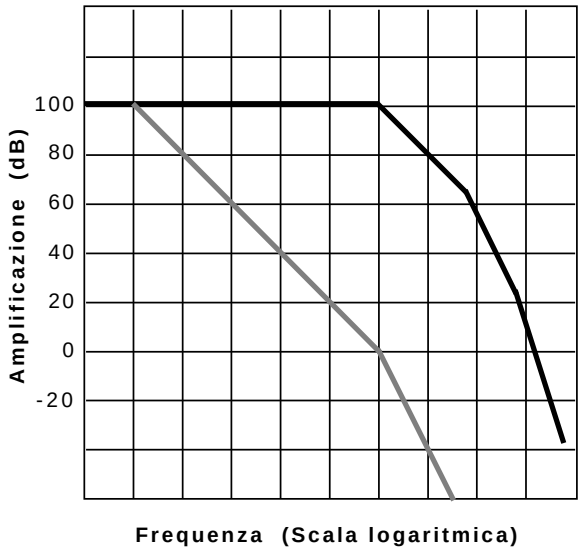

Figura 6.11: *Effetto della compensazione*

Slew rate

Usualmente le prestazioni in frequenza dell'amplificatore vengono valutate anche in base alla *slew rate*, cioe' alla massima velocita' con cui la tensione d'uscita puo' variare nel tempo, per ampi segnali di ingresso; in altri termini

$$S = \left| \frac{dv_o}{dt} \right|_{max} \tag{6.25}$$

Chiaramente la slew rate e' legato alla frequenza di taglio del circuito; essa puo' essere misurata direttamente inviando all'ingresso un'onda rettangolare ed osservando la distorsione nel segnale d'uscita.
Il 741 ha una slew rate di circa 1 $V/\mu s$; ma lo LF157 arriva a 50 $V/\mu s$ ed esistono in commercio tipi ancor piu' *veloci*.

Disponibilita'

La varieta' di operazionali disponibili in commercio e' enorme e tutti i principali produttori di integrati offrono una vasta scelta.
Velocita', precisione (cioe' minimi valori di offset di ingresso), consumo e, naturalmente, costo, sono i parametri principali di cui il progettista deve tener conto per scegliere il tipo piu' adatto alle esigenze di una particolare applicazione. Ma vi sono molte altre caratteristiche che possono essere rilevanti, come ad esempio:

Alimentazione: in generale sono necessarie due alimentazioni, ma alcuni tipi possono essere utilizzati con una sola (la seconda e' sostituita dalla massa);

Massima escursione del segnale d'uscita: spesso e' limitata in un intervallo assai piu' ristretto rispetto alle alimentazioni;

Tensioni d'ingresso (modo comune):per assicurare un corretto funzionamento del circuito i due ingressi non possono uscire da un certo intervallo di tensione; spesso questo e' piu' ristretto rispetto alle alimentazioni;

Massima tensione differenziale d'ingresso: la stessa cosa, ma riferita alla differenza tra i due segnali d'ingresso.

Nella Figura 6.12 sono sintetizzate alcune delle principali caratteristiche di tre operazionali in commercio che abbiamo gia' citato.

Sigla		$\mu A741$			$LM358$			$LF157$		
Costruttore		Fairchild			National			National		
Parametro	Unita'	Min	Typ	Max	Min	Typ	Max	Min	Typ	Max
Condizioni di misura		$V_{CC} \pm 15\ V$			$V^+ = 5\ V$			$V_{CC} \pm 15\ V$		
Tensione di offset d'ingresso	(mV)		1.0	5.0		2.0	7.0		1.0	2.0
Corrente di offset d'ingresso	(nA)		20	200		5	50		3 (pA)	10 (pA)
Corrente di bias d'ingresso	(nA)		80	500		45	250		30 (pA)	50 (pA)
Impedenza d'ingresso	($M\Omega$)	0.3	2.0		Non riportata				10^6	
Alimentazione	(V)			± 22	± 1.5		± 16			± 22
Alimentazione singola	(V)	Non prevista			3		32	Non prevista		
Limite tensioni d'ingresso	(V)	± 12	± 13		da -0.3 a $V^+ - 1.5$			± 11	$+15/-12$	
Limite tensione diff. d'ingresso	(V)		± 30		Come alimentazione				± 40	
Guadagno di tensione	(V/mV)	50	200		25	100		50	200	
Common Mode Rejection	(dB)	70			65	85		85	100	
Escursione d'uscita	(V)	± 12			da 0 a $V^+ - 1.5$			± 12	± 13	
Banda	(Mhz)		1.0			1.0		15	20	
Slew Rate	($V/\mu s$)		0.5			0.5		40	50	

Figura 6.12: *Confronto tra le caratteristiche di tre operazionali in commercio*

Il 741 e' un operazionale di uso generale, tra i piu' noti in commercio. E' fabbricato (con sigle lievemente diverse) da moltissimi costruttori, a volte con prestazioni lievemente diverse per quanto riguarda i limiti di precisione (offset di tensione, di corrente, ecc.). Comunque tutti hanno ingressi separati per il circuito di correzione dell'offset.

L'integrato LM358, 2 operazionali nello stesso contenitore, ha prestazioni simili al 741, ma non ha gli ingressi separati per il circuito di correzione dell'offset. Puo' funzionare anche con una sola alimentazione (almeno $3\ V$), ed ha un assorbimento di potenza sensibilmente piu' ridotto.

Infine l'amplificatore LF157 ha ingressi JFET (quindi resistenza d'ingresso elevatissima e correnti d'ingresso trascurabili). E' molto piu' veloce avendo una rete di compensazione con polo dominante attorno ai $100\ Hz$.

6.4 Applicazioni

L'amplificatore operazionale ha un elevatissimo numero di usi ed applicazioni. Abbiamo gia' visto come realizzare semplici amplificatori (invertenti e non); vedremo ora alcuni altri esempi, che risolveremo utilizzando quasi sempre il modello di operazionale *ideale*.

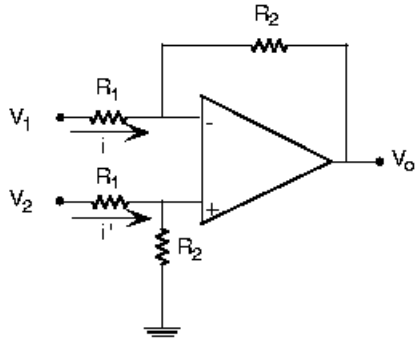

Figura 6.13: *Amplificatore differenziale*

Amplificatore differenziale

Per costruire un amplificatore differenziale si puo' utilizzare lo schema mostrato nella Fig. 6.13. In base a quanto ormai abbiamo imparato possiamo assumere che nell'operazionale non entra corrente e che i terminali d'ingresso sono alla stessa tensione, v. Dal morsetto invertente ricaviamo

$$v_o = v - iR_2 \tag{6.26}$$

dove

$$i = \frac{v_1 - v}{R_1} \tag{6.27}$$

quindi

$$v_o = v - \frac{R_2}{R_1}(v_1 - v) \tag{6.28}$$

ovvero

$$v_o = v(1 + \frac{R_2}{R_1}) - \frac{R_2}{R_1}v_1 \tag{6.29}$$

Dal morsetto non invertente possiamo ricavare v, cioe'

$$v = v_2\frac{R_2}{R_1 + R_2} \tag{6.30}$$

e combinando le due equazioni si ottiene

$$v_o = \frac{R_2}{R_1}(v_2 - v_1) \tag{6.31}$$

Naturalmente questo risultato presuppone un amplificatore operazionale ideale e fornisce quindi un CMRR infinito. In realta' si avra' una reiezione di modo comune finita anche se elevata; si noti inoltre che abbiamo presupposto l'assoluta eguaglianza delle due resistenze R_1 e delle due resistenze R_2.

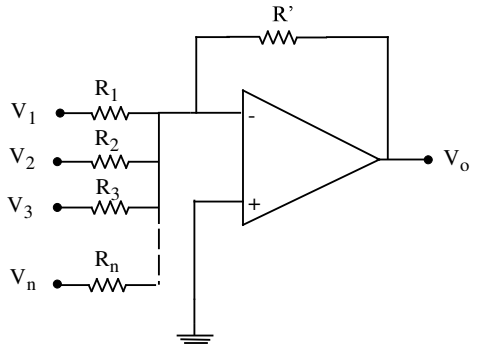

Figura 6.14: *Sommatore analogico*

Sommatore analogico

Consideriamo il circuito in Fig 6.14; i due morsetti d'ingresso sono ora a tensione zero, quindi la corrente i e' data da

$$i = \frac{v_1}{R_1} + \frac{v_2}{R_2} + \cdots + \frac{v_n}{R_n} \tag{6.32}$$

d'altra parte

$$v_o = -R'i \tag{6.33}$$

e quindi

$$v_o = -R'(\frac{v_1}{R_1} + \frac{v_2}{R_2} + \cdots + \frac{v_n}{R_n}) \tag{6.34}$$

Ponendo

$$R_1 = R_2 = \cdots = R_n \tag{6.35}$$

la tensione d'uscita sara' data da

$$v_o = -\frac{R'}{R}(v_1 + v_2 + \cdots + v_n) \tag{6.36}$$

cioe' sara' proporzionale alla somma delle tensioni d'ingresso.

Generatore ideale di corrente

L'amplificatore operazionale puo' essere utilizzato come sorgente ideale di corrente. Infatti, se consideriamo il circuito in Fig. 6.15, la corrente che circola nel carico R_L e' data da

$$i = \frac{v_s}{R} \tag{6.37}$$

cioe' non dipende da R_L.

Derivatore e integratore

Il circuito di Fig. 6.16 si comporta come un derivatore. Infatti, applicando la relazione (6.6) otteniamo

$$\frac{V_o}{V_s} = -j\omega RC \tag{6.38}$$

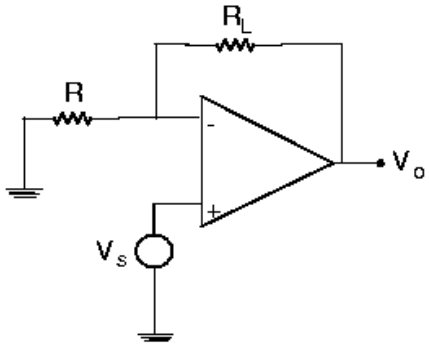

Figura 6.15: *Generatore di corrente*

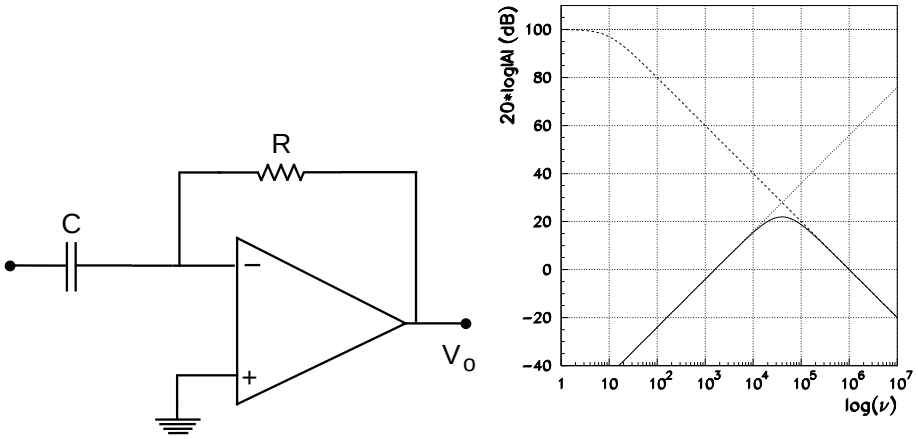

Figura 6.16: *Derivatore; il diagramma di Bode ideale e' la linea puntinata, mentre quello reale (per* $RC = 10^{-4}$ *s) e' rappresentato dalla linea continua, quando si tenga conto della vera funzione di trasferimento dell'operazionale (linea tratteggiata)*

cioe' proprio la funzione di trasferimento del derivatore ideale. Tuttavia questo calcolo e' puramente teorico, come si comprende osservando che la funzione di trasferimento divergerebbe a bassa frequenza. Il modello ideale di amplificatore non e' applicabile e occorre tener conto in modo piu' realistico della funzione di trasferimento vera. Prendiamo ad esempio un operazionale *compensato* internamente: possiamo descrivere la sua amplificazione ad anello aperto con l'espressione ad un solo polo

$$A_V = \frac{A_{Vo}}{1 + j\frac{f}{f_1}} \tag{6.39}$$

dove f_1 e' la frequenza del polo dominante. Utilizziamo quindi la (6.5) dove ora A_V non e' costante, ma e' espresso dalla (6.39). Si otterra' allora la funzione di trasferimento piu' realistica riportata in Fig 6.16

Analogamente, il circuito di Fig. 6.17 si comporta come un integratore: la sua funzione di trasferimento e' infatti

$$\frac{V_o}{V_s} = -\frac{1}{j\omega RC} \tag{6.40}$$

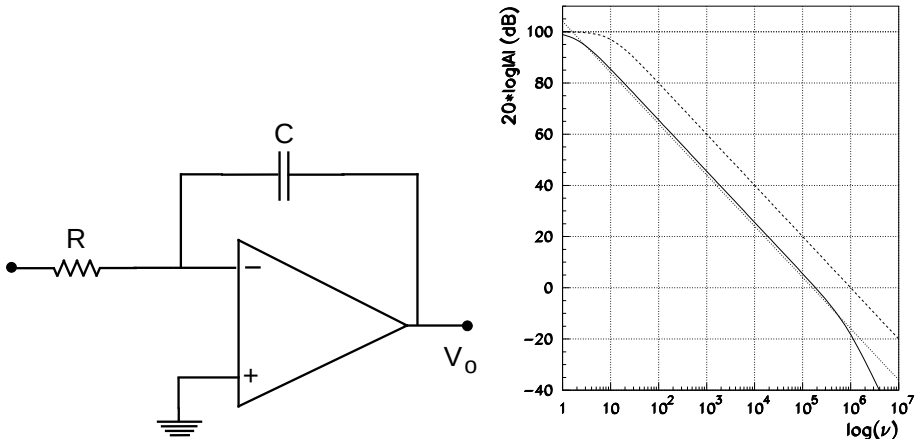

Figura 6.17: *Integratore; il diagramma di Bode ideale e' la linea puntinata, mentre quello reale (per* $RC = 10^{-6}$ *s) e' rappresentato dalla linea continua, quando si tenga conto della vera funzione di trasferimento dell'operazionale (linea tratteggiata)*

Anche qui il calcolo e' puramente ideale; utilizzando di nuovo il modello piu' realistico otteniamo la curva continua della Fig. 6.17. L'integratore non puo' concretamente fun-

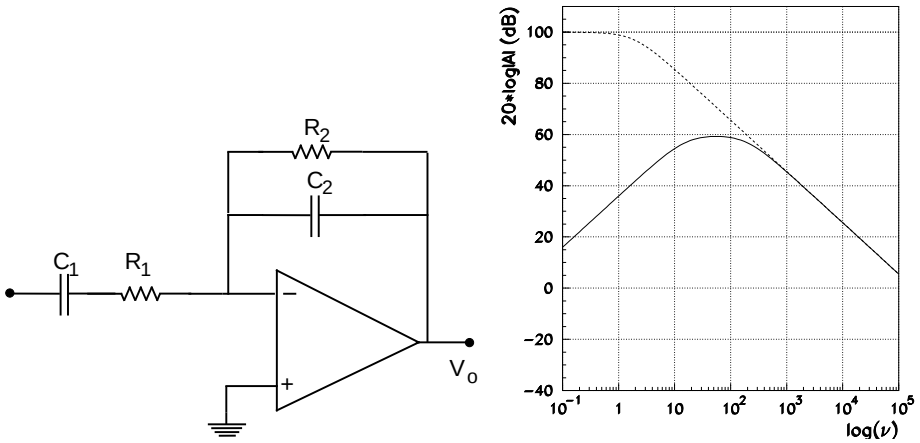

Figura 6.18: *a) Integratore modificato; b) Diagramma di Bode (la linea tratteggiata e' relativa al nuovo circuito;*

zionare; infatti l'amplificazione a bassa frequenza (fino alla tensione continua) e' troppo alta. Quindi una piccola componente di bassa frequenza (o una livello di tensione continua) presente nel segnale d'ingresso porterebbe immediatamente in saturazione l'uscita. E' quindi necessario introdurre una capacita' di blocco C_1 (Fig. 6.18a); a questo punto pero' non ci sarebbe piu' un percorso per la corrente di ingresso I_{b_2}; si deve allora aggiungere una resistenza R_2 in parallelo a C_2. Il circuito e' ora simultaneamente un derivatore ed un integratore: occorre che l'effetto derivante sia trascurabile, per cui si deve avere $R_2 \gg R_1$,

$C_1 \gg C_2$ e $R_1 C_2 \gg T$, dove T e' il periodo del segnale che si vuole integrare. In Fig. 6.18b vediamo l'effetto di una capacita' $C_1 = 1.\ \mu F$ e di una resistenza $R_2 = 10\ M\Omega$.

Raddrizzatore

Consideriamo il circuito di Fig. 6.19a): se v_i e' negativo il diodo non conduce, non vi e' quindi contro reazione e l'uscita v_o e' nulla. Quando

$$v_i > \frac{V_\gamma}{A_v} \tag{6.41}$$

(dove V_γ e' la tensione di ginocchio del diodo) il diodo entra in conduzione e v_o ricopia fedelmente la forma dell'ingresso. Il vantaggio rispetto ad un normale raddrizzatore e' che la tensione di ginocchio del diodo e' divisa per un fattore elevatissimo, quindi di fatto il diodo commuta esattamente alla tensione zero.

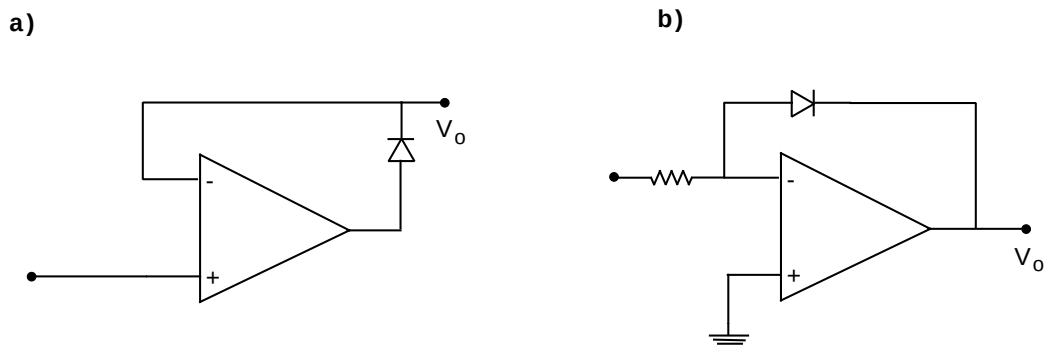

Figura 6.19: *a) Raddrizzatore; b) Amplificatore logaritmico*

Amplificatore logaritmico

Un amplificatore logaritmico puo' essere ottenuto con lo schema in Fig. 6.19b: la corrente che circola nel diodo e' data da

$$I_f = I_o(e^{\frac{V_f}{\eta V_T}} - 1) \simeq I_o\, e^{\frac{V_f}{\eta V_T}} \tag{6.42}$$

Mentre la corrente che circola nella resistenza e' data da

$$I_s = \frac{V_s}{R} \tag{6.43}$$

Poiche' le due correnti sono uguali e $V_o = -V_f$ si ricava

$$V_o = -\eta V_T(\ln \frac{V_s}{R} - \ln I_o) \tag{6.44}$$

Quindi l'uscita e' proporzionale al logaritmo dell'ingresso. In realta' questo circuito non da prestazioni molto soddisfacenti (tra l'altro e' molto instabile per variazioni di temperatura);

si puo' pero' notevolmente migliorare, utilizzando un transistor al posto del diodo, ed anche minimizzare gli effetti termici.
E' anche possibile costruire amplificatori anti-logaritmici, ovvero esponenziali, sempre sfruttando opportunamente il termine esponenziale presente nell'equazione di una giunzione.

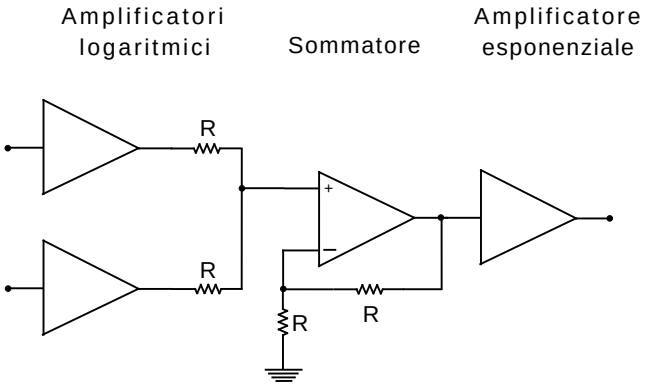

Figura 6.20:

Moltiplicatori analogici

La moltiplicazione analogica di due segnali, cioe' un'uscita proporzionale al prodotto di due ingressi, puo' essere ottenuta sfruttando le note proprieta' dei logaritmi. Lo schema in Fig. 6.20 e' un esempio di come si possa raggiungere lo scopo. Si noti tuttavia che esistono delle limitazioni nell'uso di questo circuito: i segnali d'ingresso devono essere positivi (il logaritmo di un numero negativo non e' definito).
Un moltiplicatore a 4 quadranti, cioe' capace di moltiplicare ingressi di qualunque segno e

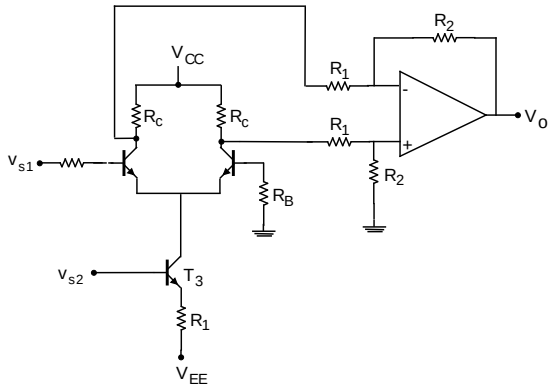

Figura 6.21:

di fornire il segno giusto all'uscita, puo' essere realizzato in modo molto piu' elegante, con il cosidetto amplificatore di transconduttanza (Fig. 6.21). Qui si sfrutta uno stadio d'ingresso di tipo differenziale: la sua risposta e' proporzionale (oltreche' al segnale d'ingesso v_{s1}) alla transconduttanza dei due transistors T_1 e T_2, cioe' alla loro corrente statica I_C: questa e'

a sua volta determinata dal generatore di corrente T_3, il cui punto di lavoro puo' essere determinato dal secondo segnale d'ingresso v_{s2}.
Esistono in commercio integrati basati su questo schema di principio: per esempio, lo MC1495 (Motorola) e' un moltiplicatore a quattro quadranti con due ingressi differenziali ed uscita differenziale. Si ha quindi

$$v_{o1} - v_{o2} = K(v_{x1} - v_{x2})(v_{y1} - v_{y2}) \tag{6.45}$$

dove il fattore K puo' essere scelto dall'utilizzatore mediante una rete resistiva esterna.

6.5 Filtri attivi

Con gli amplificatori operazionali si possono realizzare dei filtri, cioe' circuiti selettivi in frequenza, in modo molto piu' efficiente di quanto sia possibile fare utilizzando solo componenti passivi. Inoltre e' possibile costruire filtri *passa-banda* senza utilizzare induttori, che, come abbiamo visto nel Cap. 2, sono componenti assai scomodi da usare: ingombranti, costosi, non trascurabile resistenza parassita, e molti altri inconvenienti. Per comprendere come cio' sia possibile, consideriamo il circuito in Fig. 6.22a, e calcoliamone l'impedenza di ingresso. Poiche' i due ingressi dell'operazionale sono alla stessa tensione, la corrente

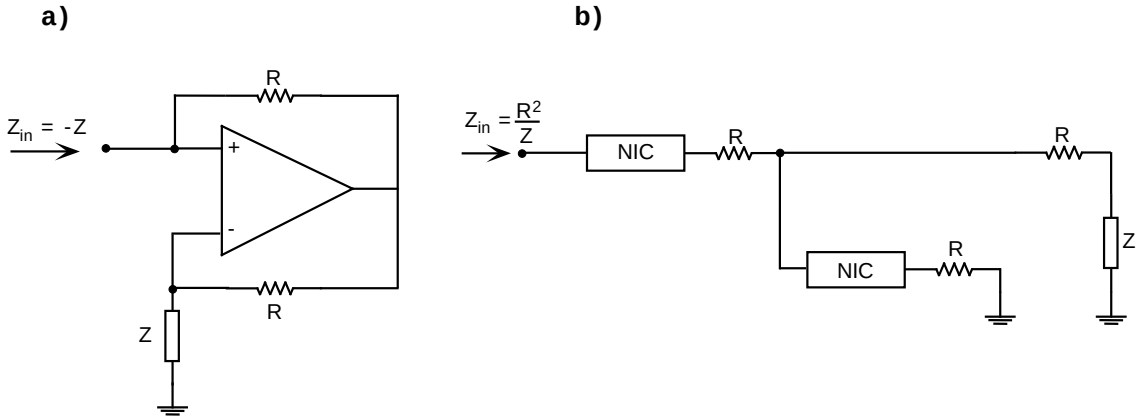

Figura 6.22: *a) Convertitore di impedenza; b) Giratore*

che scorre nelle due resistenze e' identica, ed e' quindi uguale alla corrente che scorre nell'impedenza Z. Si ha quindi

$$V_1 = V_2 = -ZI \tag{6.46}$$

Quindi l'impedenza d'ingresso e' proprio $-Z$. Questo tipo di circuiti prende il nome di *negative-impedance converters*, spesso abbreviato in NIC.
Se ora calcoliamo l'impedenza di ingresso del circuito in Fig. 6.22b, otteniamo

$$Z_i = \frac{R^2}{Z} \tag{6.47}$$

Quindi, se

$$Z = \frac{1}{j\omega C} \;\rightarrow\; Z_i = j\omega C R^2 \tag{6.48}$$

cioe' il circuito *simula* un induttore di induttanza $L = CR^2$. Questo circuito si chiama *giratore* e dimostra la possibilita' di costruire filtri in cui l'induttore e' sostituito da un opportuno giratore.

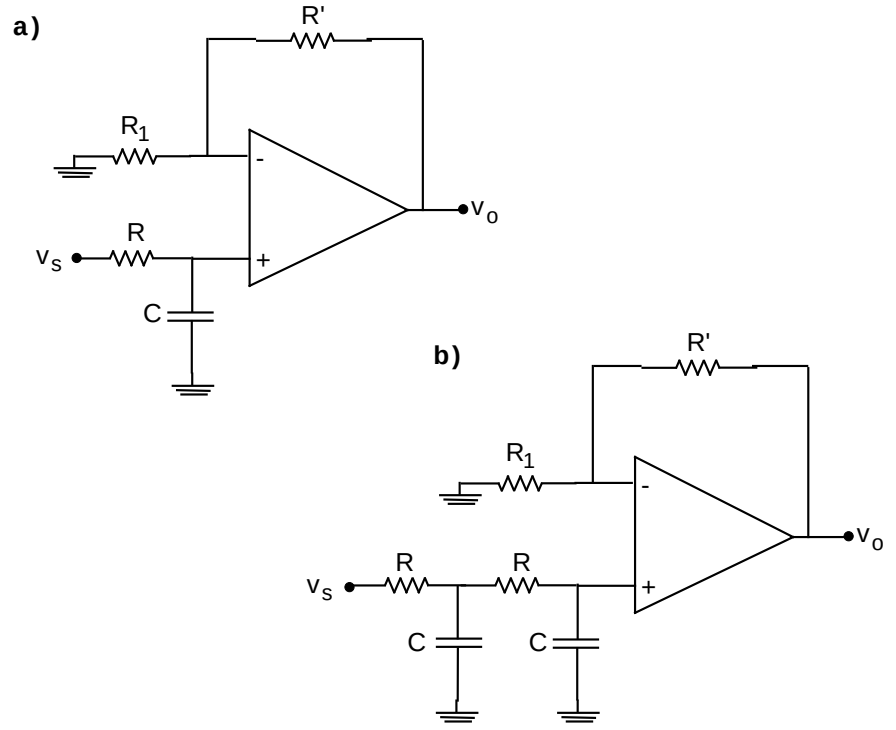

Figura 6.23: *a) Filtro passa-basso del primo ordine; b) Filtro passa-basso del secondo ordine*

6.5.1 Filtri passa-basso e passa-alto

La Fig. 6.23a mostra un esempio di filtro passa-basso. E' immediato verificare che la sua funzione di trasferimento e' quella desiderata. La pulsazione di taglio e' data da

$$\omega_o = \frac{1}{RC} \tag{6.49}$$

e si ha la usuale discesa a $-20\, dB/decade$. A basse frequenze si ha un'amplificazione

$$A = \frac{R_1 + R'}{R_1} \tag{6.50}$$

E' invece piu' interessante studiare il circuito di Fig. 6.23b. Si ha

$$\begin{aligned} \frac{V_o}{V_i} &= \frac{R_1 + R'}{R_1} \\ I &= sCV_i \\ V_1 &= I(R + \frac{1}{sC}) \\ &= sCV_i(R + \frac{1}{sC}) \\ &= \frac{V_o}{A_o}(1 + sRC) \end{aligned}$$

Ora scrivendo l'equazione del nodo 1 ed utilizzando le relazioni trovate prima si ottiene

$$A(s) = \frac{A_o}{(sRC)^2 + (3 - A_o)(sRC) + 1} \tag{6.51}$$

cioe' una funzione del tipo

$$A(s) = \frac{A_o}{(s/\omega_o)^2 + 2K(s/\omega_o) + 1} \tag{6.52}$$

La pulsazione di taglio e' data da

$$\omega_o = \frac{1}{RC} \tag{6.53}$$

ma la pendenza e' ora doppia, cioe' $-40\ dB/decade$.
Questo circuito si chiama filtro del secondo ordine. Chiaramente possiamo costruire filtri di ordine superiore (cioe' con pendenze ancora piu' ripide) mettendo in serie tra loro filtri del primo o del secondo ordine.

Filtri Sallen-Key

Possiamo costruire un filtro passa-basso del secondo ordine lievemente diverso, modificando il circuito come nella Fig. 6.24a: il primo capacitore e' connesso all'uscita, anziche' alla massa. Questo circuito e' comunemente noto come filtro Sallen-Key, dai nomi dei suoi inventori.
Ora i morsetti d'ingresso sono (entrambi) alla tensione V_o, quindi la corrente I che scorre nel condensatore verso massa e' data da

$$I = sCV_o \tag{6.54}$$

mentre la tensione del nodo 1 e' evidentemente

$$V_1 = I(R + \frac{1}{sC}) \tag{6.55}$$

Combinando le due relazioni si ha

$$V_1 = sCV_o(R + \frac{1}{sC}) \tag{6.56}$$

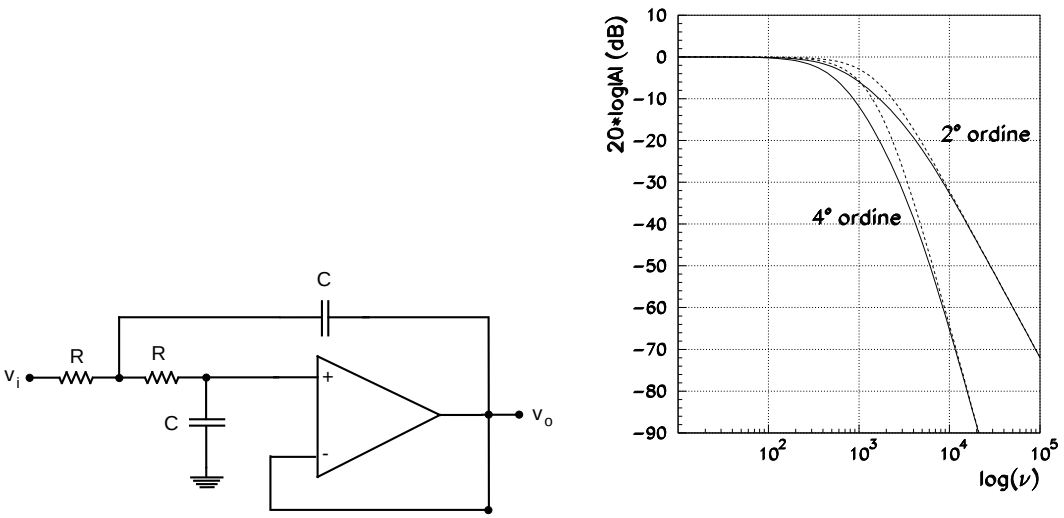

Figura 6.24: *a) Filtro Sallen-Key passa-basso; b) Le linee continue rappresentano le funzioni di trasferimento di filtri normali, quelle tratteggiate di filtri tipo Sallen-Key*

L'equazione del nodo 1 e'

$$\frac{V_i - V_1}{R} = sC(V_1 - V_o) + \frac{V_1}{R + \frac{1}{sC}} \tag{6.57}$$

e con pochi passaggi si arriva a

$$A(s) = \frac{1}{(sRC)^2 + 2s(RC) + 1} \tag{6.58}$$

che, in termini di pulsazione possiamo anche scrivere

$$A(\omega) = \frac{1}{(1 + j\omega RC)(1 + j\omega RC)} \tag{6.59}$$

Possiamo verificare il differente comportamento di questo filtro rispetto al precedente, confrontando i diagrammi di Bode per il 2^o ed il 4^o ordine (Fig. 6.24b): il comportamento asintotico e' lo stesso (discesa a 40 dB/decade e ad 80 dB/decade), ma la regione di transizione e' diversa. I Sallen-Key hanno un ginocchio piu' netto, ed arrivano piu' rapidamente alla pendenza asintotica.
Ovviamente scambiando i condensatori con i resistori possiamo realizzare un filtro Sallen-Key passa-alto.

Filtri VCVS

I filtri VCVS (Voltage Controlled Voltage Source) sono una variante dei circuiti tipo Sallen-Key. L'inseguitore di tensione e' sostituito da un amplificatore non invertente con guadagno superiore ad uno.

La forma della risposta di un filtro VCVS dipende, nella regione di transizione, dal valore del fattore di amplificazione K, come si puo' osservare dalla Fig. 6.26, dove sono

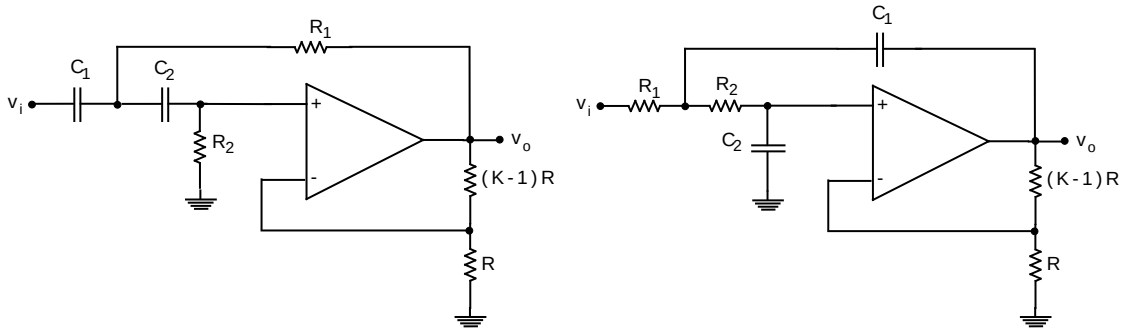

Figura 6.25: *Filtri VCVS passa-alto e passa basso*

riportati gli andamenti del modulo di A e della fase per un passa basso del 2^o ordine. Per valori di K superiori a ~ 1.6 il modulo di A mostra un *overshooting* attorno alla frequenza critica, mentre la fase mostra, al crescere di K, una transizione sempre piu' netta.

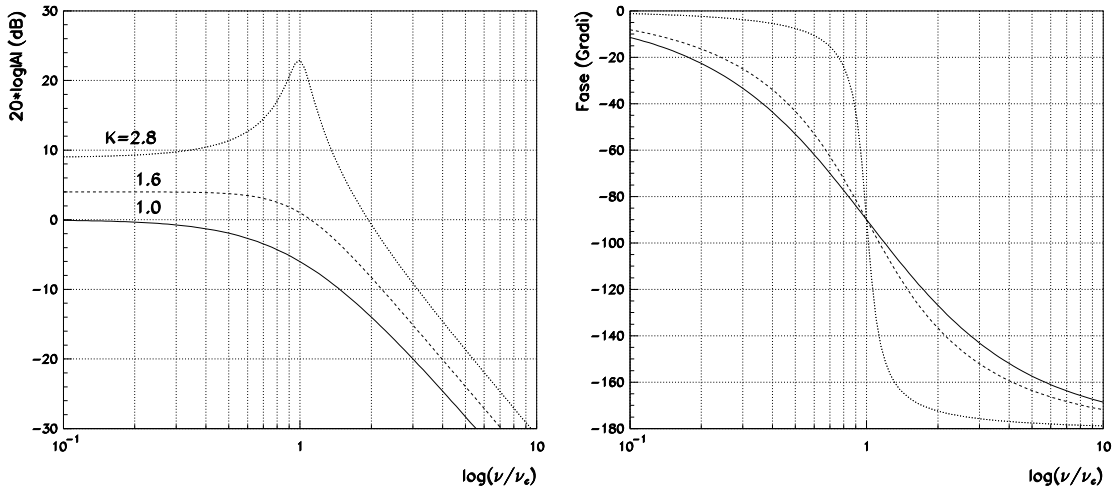

Figura 6.26: *Ampiezza e fase di un filtro VCVS passa-basso del 2^o ordine, per 3 valori di K. La scala orizzontale e' normalizzata alla frequenza critica $\nu_c = 1/(2\pi RC)$.*

6.5.2 Filtri passa-banda

Un filtro passa-banda puo' essere ottenuto mettendo in serie un filtro passa-basso ed uno passa-alto (con opportune frequenze di taglio). Si ottiene in questo modo una banda passante *piatta* tra le due discese a $20n$ db/decade, dove n e' l'ordine dei filtri utilizzati. Questo puo' essere fatto sia con filtri normali, sia con filtri VCVS. Un filtro VCVS passa-banda del 2^o ordine puo' anche essere piu' semplicemente costruito partendo da uno dei circuiti di Fig 6.25: basta scambiare C_1 con R_1, lasciando invariato il resto.
Sono pero' piu' interessanti i filtri di tipo *risonante*, cioe' con una funzione di trasferimento

simile a quella del classico RLC che abbiamo visto nel Cap. 2: il circuito di Fig. 6.27 ne e' un esempio.
La reazione negativa e' riportata due volte verso l'ingresso, sia attraverso il resistore R_3 che attraverso i due capacitori C_1 e C_2; percio' questi circuiti sono noti come filtri a *reazione multipla*.
Possiamo studiare il comportamento di questo circuito utilizzando il modello di operazionale ideale. Per semplicita' consideriamo il caso in cui i due capacitori sono uguali ($C_1 = C_2 = C$. Poiche' l'ingresso non invertente dell'operazionale e' a massa anche la tensione di quello invertente e' nulla; quindi possiamo scrivere l'equazione del nodo in cui confluiscono R_1, R_2 e i due capacitori

$$\frac{V_s - V_1}{R_1} + (V_o - V_1)sC - V_1 sC = \frac{V_1}{R_2} \tag{6.60}$$

cioe'

$$\frac{V_s}{R_1} - V_1(\frac{1}{R_1} + \frac{1}{R_2} + 2sC) + sCV_o = 0 \tag{6.61}$$

dove con V_1 abbiamo indicato la tensione del nodo stesso.

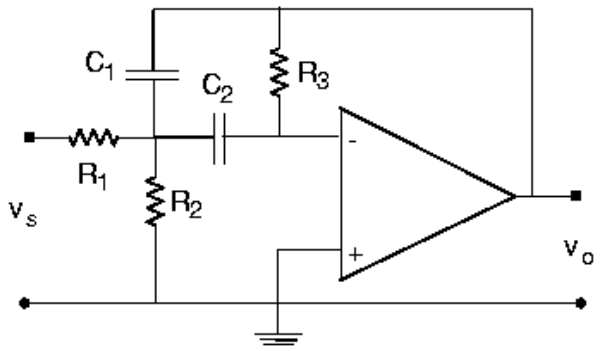

Figura 6.27: *Filtro passa-banda*

La relazione tra V_1 e V_o puo' d'altra parte trovarsi subito osservando che la corrente che fluisce in R_3 e' data da

$$I_3 = \frac{V_o}{R_3} \tag{6.62}$$

e la tensione V_1 e' data da

$$V_1 = -\frac{I_3}{sC} \tag{6.63}$$

Per cui

$$V_1 = -\frac{V_o}{sCR_3} \tag{6.64}$$

Sostituendo nella 6.61 si ottiene

$$\frac{V_s}{R_1} + V_o(\frac{1}{sCR_3R'} + \frac{2}{R_3} + sC) = 0 \tag{6.65}$$

dove

$$R' = \frac{R_1 R_2}{R_1 + R_2} \tag{6.66}$$

cioe' il parallelo tra R_1 ed R_2. Possiamo quindi scrivere la funzione di trasferimento

$$A(\omega) = \frac{V_o}{V_s} = \frac{-\frac{R_3}{2R_1}}{(1 + j(\frac{R_3 C}{2}\omega - \frac{1}{2CR'\omega}))} \tag{6.67}$$

dove abbiamo come al solito sostituito s con $j\omega$. Si vede che la 6.67 rappresenta una funzione risonante con frequenza di risonanza

$$\omega_o^2 = \frac{1}{R' R_3 C^2} \tag{6.68}$$

e con ampiezza massima (sulla risonanza)

$$A_o = -\frac{R_3}{2R_1} \tag{6.69}$$

Introducendo il fattore di merito, Q

$$Q = \omega_o \frac{R_3 C}{2} \tag{6.70}$$

la 6.67 puo' essere scritta

$$A = \frac{A_o}{(1 + jQ(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}))} \tag{6.71}$$

Ovvero

$$|A| = \frac{|A_o|}{\sqrt{1 + Q^2(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega})^2}} \tag{6.72}$$

Le equazioni 6.71 e 6.72 son le stesse (a parte il fattore A_o) di quelle del circuito RLC, ed ovviamente Q rappresenta il fattore di merito del circuito. Come si vede, possiamo scegliere A_o, ω_o e Q variando i parametri del circuito C, R_1, R_3 e R'. Poiche' abbiamo 4 parametri uno di essi puo' essere fissato arbitrariamente e gli altri 3 vengono fissati in conseguenza ai valori di progetto desiderati.

Ad esempio, supponiamo di voler progettare un filtro con

$$A_o = 5 \quad \nu_o = 10\ kHz \quad Q = 50 \tag{6.73}$$

Poniamo

$$C = 3\ nF \tag{6.74}$$

a questo punto R_3 e' dato da

$$R_3 = \frac{2Q}{2\pi\nu_o C} \simeq 500\ k\Omega \tag{6.75}$$

quindi possiamo scegliere R' che e' dato da

$$\frac{1}{4\pi^2 \nu_o^2 C^2 R_3} \simeq 50\ \Omega \tag{6.76}$$

ed infine

$$R_1 = \frac{R_3}{2A_o} = 50\ k\Omega \tag{6.77}$$

Poiche R' e' il parallelo tra R_1 (gia' fissato) ed R_2, si trova immediatamente che

$$R_2 \simeq 50\ \Omega \tag{6.78}$$

Si vede quindi che abbiamo la possibilita' di costruire filtri passa-banda molto differenziati (grande o piccola selettivita', grande o piccola amplificazione) senza avere utilizzato induttori che sono componenti assai piu' scomodi da usare.

6.5.3 Filtri a variabili di stato

Il circuito in Fig. 6.28 e' assai piu' complesso, richiedendo 3 operazionali. Consente pero' di scegliere piu' agevolmente frequenza di risonanza, fattore di merito e guadagno di picco [3].

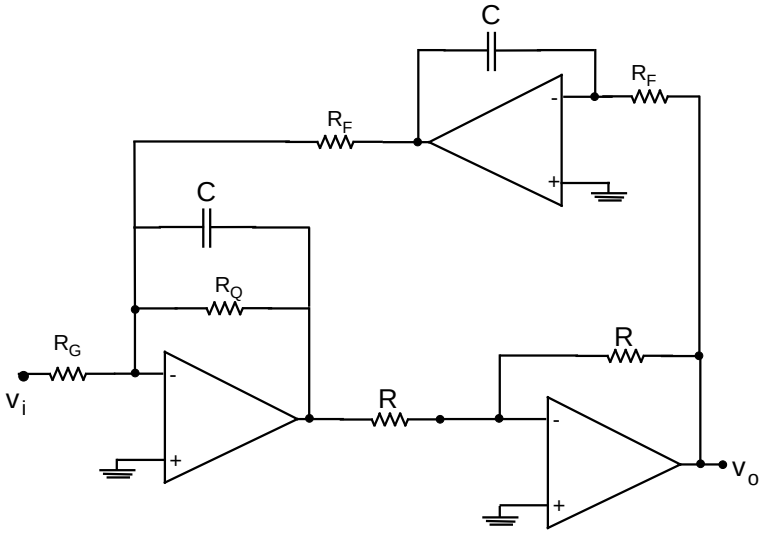

Figura 6.28: *Filtro*

Si possono infatti facilmente trovare le seguenti relazioni:

$$f_o = \frac{1}{2\pi R_F C}$$
$$Q = \frac{R_B}{R_F}$$
$$G = \frac{R_B}{R_G}$$

Volendo quindi progettare un filtro conviene procedere nel seguente modo:

[3] E' bene comunque prendere con cautela i risultati ottenuti, che sono validi solo a frequenze raginevolmente basse. Calcoli più attendibili devono essere fatti tenendo conto della funzione di trasferimento reale di un operazionale.

1. scegliere un valore ragionevole per il capacitore, in base alla frequenza f_o desiderata

$$C = \frac{10}{f_o} \, (\mu F)$$

2. calcolare il valore di R_F necessario per soddisfare la relazione data prima;
3. scegliere R_B in base al Q desiderato;
4. calcolare il valore di R_G necessario per aggiustare il guadagno al valore voluto.

Capitolo 7

Circuiti digitali

Finora abbiamo studiato dispositivi e circuiti di tipo *lineare*, ovvero delle situazioni in cui l'andamento temporale delle variabili elettriche d'uscita (tensione e/o corrente) riproduceva quello delle variabili d'ingresso. Invece i circuiti digitali (o circuiti logici) sono sistemi in cui lo stato dell'uscita, o delle uscite, e' legato in modo fortemente non lineare a quello dell'ingresso o degli ingressi. Piu' specificamente i circuiti logici sono caratterizzati dal fatto che le variabili elettriche d'uscita, tipicamente la tensione, assume due soli valori possibili (per es. $0\ V$ e $5\ V$) in conseguenza al valore di tensione applicato all'ingresso. Circuiti di questo tipo sono alla base, come vedremo, dei sistemi di elaborazione numerica e rivestono quindi una importanza fondamentale. Abbiamo gia' visto esempi molto rudimentali di circuiti del genere quando abbiamo parlato del transistor utilizzato come interruttore: in quel caso la tensione V_C del collettore puo' essere fatta variare tra un valore V_{CC} ed un valore V_{CEsat} (molto prossima a zero) in dipendenza della tensione applicata alla base.
Per comprendere meglio le applicazioni di questo tipo di circuiti e' necessario fare alcune premesse di natura matematica.

7.1 Numerazione binaria ed algebra di Boole

Il sistema numerico che noi utilizziamo e' di tipo posizionale, basato sulle potenze di 10. Infatti la scrittura di un qualunque numero non e' altro che l'espressione sintetica di una combinazione lineare di potenze di 10. Per esempio, il numero 327 non e' che un'abbreviazione per l'espressione

$$7 \cdot 10^0 + 2 \cdot 10^1 + 3 \cdot 10^2$$

Le 4 operazioni aritmetiche vengono di conseguenza eseguite tenendo conto del significato posizionale delle cifre che compongono il numero ed utilizzando, se necessario, il riporto da una posizione a quella, piu' significativa, alla sinistra della precedente.
In realta' la base 10 non ha nulla di fondamentale e potrebbe essere tranquillamente sostituita da qualunque altra base. La base piu' semplice puo' essere considerata quella fondata sul 2 (non ha senso infatti la base 1!); possiamo quindi costruire un sistema numerico in base 2, utilizzando le potenze di questo numero per esprimerne uno qualunque. Per fare cio' abbiamo bisogno di due soli simboli numerici, 0 ed 1 (nel sistema in base 10 avevamo 10 simboli, da 0 a 9). La nostra numerazione sara' allora: Le 4 operazioni aritmetiche pos-

Num. decimale	Num. binaria
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
...	...
16	10000
...	...
32	100000
...	...

Tabella 7.1: *Numerazione binaria*

sono essere effettuate nel sistema numerico a base 2 (cosi' come in qualunque altro sistema numerico), con le stesse regole del sistema decimale; la sola differenza e' che il riporto deve essere effettuato quando si arriva a 2 e non quando si arriva a 10.
Possiamo fin da ora comprendere che il sistema binario, per il fatto di avere due soli numeri fondamentali, si presta particolarmente ad essere associato a sistemi elettronici in cui le variabili elettriche assumono due valori possibili, e quindi ci aiutera' a costruire macchine per l'elaborazione numerica. Dobbiamo pero' ancora approfondire gli aspetti matematici del sistema binario.
Possiamo infatti definire, oltre alle normali operazioni aritmetiche, delle operazioni *logiche*, che costituiscono nel loro complesso la cosidetta algebra di Boole; queste regole non si applicano solo all'insieme dei numeri binari, ma piu' in generale a tutti gli insiemi binari, cioe' composti da due elementi, che possiamo chiamare 0 ed 1, ma anche VERO e FALSO, o BIANCO e NERO, ecc.

Algebra di Boole. Definiamo tre operazioni logiche tra gli elementi dell'insieme binario:

1. Prodotto logico (detto anche AND): dati n elementi $x_1, x_2, \ldots, x_n$, il loro prodotto logico ha come risultato 1 se e solo se tutti gli elementi valgono 1; altrimenti il risultato e' 0. Cioe':

$$x_1 \cdot x_2 \cdots x_n = \begin{cases} 1 & \text{se ogni } x_j = 1 \\ 0 & \text{altrimenti} \end{cases}$$

Il punto viene spesso omesso, se cio' non da' luogo ad ambiguita', quindi $x_1 \cdot x_2$ puo' semplicemente essere scritto x_1x_2.

2. Somma logica (detta anche OR):dati n elementi $x_1, x_2, \ldots, x_n$, la loro somma logica ha come risultato 1 se almeno uno degli elementi e' uguale ad 1; altrimenti (cioe' se tutti gli elementi valgono 0) il risultato e' 0. Cioe'

$$x_1 + x_2 \cdots + x_n = \begin{cases} 0 & \text{se ogni } x_j = 0 \\ 1 & \text{altrimenti} \end{cases}$$

3. Negazione logica (detta anche NOT): dato un elemento x, l'elemento negato, che si indica con $\overline{x}$, vale 1 se x vale 0, mentre vale 0 se x vale 1. Cioe':

$$\overline{x} = 0 \quad \text{se} \quad x = 1$$
$$\overline{x} = 1 \quad \text{se} \quad x = 0$$

Naturalmente si ha $\overline{\overline{x}} = x$.

Le suddette operazioni ci consentono di definire funzioni di variabili logiche (o binarie); per esempio, date 3 variabili x_1, x_2, x_3, possiamo definire la funzione

$$F = x_1 \cdot x_2 \cdot \overline{x}_3 + x_1 \cdot \overline{x}_2 \cdot x_3 + \overline{x}_1 \cdot x_2 \cdot x_3$$

anche F potra' assumere solo i valori 0 o 1 in dipendenza dai valori assunti dalle variabili, cioe' il co-dominio della funzione e' ancora l'insieme binario. Questo modo di scrivere una funzione prende il nome di *forma canonica*: e' chiaro che ogni funzione puo' essere scritta in questo modo. Un modo diverso, spesso utile, per rappresentare una funzione e' quello di scrivere la cosidetta tavola della verita: e' semplicemente una tabella che lega il valore assunto dalla funzione ai valori assunti dalle variabili. Per esempio, la tavola della verità della funzione F vista prima e' data da:

x_3	x_2	x_1	F
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

Come si vede la tavola della verita' ha tante righe quante sono le combinazioni di tutti i valori possibili delle variabili (in questo caso 8) e per ogni riga la funzione assume un valore. Quindi, se abbiamo una funzione di n variabili, la sua tavola della verita' sara' composta da 2^n righe.
E' utile scrivere la tavola della verita' di alcune funzioni semplici, ma fondamentali: ad esempio, il prodotto di due variabili (AND), cioe' $Y = x_2 \cdot x_1$:

x_2	x_1	Y
0	0	0
0	1	0
1	0	0
1	1	1

Oppure della funzione somma (OR) di due variabili, cioe' $Y = x_2 + x_1$:

x_2	x_1	Y
0	0	0
0	1	1
1	0	1
1	1	1

Un'altra funzione importante e' il cosidetto OR ESCLUSIVO, a volte chiamato brevemente XOR, definita come:

$$x_1 \oplus x_2 \cdots \oplus x_n = \begin{cases} 1 & \text{se solo uno degli } x_j = 1 \\ 0 & \text{altrimenti} \end{cases}$$

La tavola della verita' per due variabili e'quindi data da:

x_2	x_1	Y
0	0	0
0	1	1
1	0	1
1	1	0

E' anche possibile scrivere una qualunque funzione in forma canonica partendo dalla tavola della verita'. Consideriamo ad esempio una funzione, G,di cui conosciamo la cui tavola della verita':

x_3	x_2	x_1	G
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	0

La sua forma canonica sara' allora data da:

$$G = \overline{x}_3\overline{x}_2x_1 + \overline{x}_3x_2x_1 + x_1\overline{x_2}x_3$$

cioe' avremo una somma di tanti addendi quante sono le righe per cui la funzione vale 1. In ogni addendo compare il prodotto delle variabili: se in quella riga una variabile vale 0 esso compare *negata*.
Quindi l'OR ESCLUSIVO visto in precedenza corrisponde ad una funzione la cui forma canonica e'

$$Y = x_1\overline{x}_2 + \overline{x}_1x_2$$

L'algebra di Boole gode di alcune proprieta' molto semplici. Indicando con $A, B, C, \ldots$ delle variabili binarie, si ha:

1. Proprietà commutativa della somma:

$$A + B = B + A$$

2. Proprietà commutativa del prodotto:

$$A \cdot B = B \cdot A$$

3. Proprietà associativa della somma:

$$(A + B) + C = A + (B + C)$$

4. Proprietà associativa del prodotto:

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

5. Prima proprietà distributiva:

$$A \cdot (B + C) = A \cdot B + A \cdot C$$

6. Seconda proprietà distributiva:

$$(A+B)\cdot(A+C) = A + B\cdot C$$

Alcune di queste proprieta' sono essenzialmente dei postulati, mentre le altre possono essere verificate immediatamente.
Un'ulteriore importante proprieta' e' data dal teorema di De Morgan, che lega tra loro le operazioni di somma e di prodotto logico. Si hanno due enunciazioni del teorema:

$$\overline{A} + \overline{B} = \overline{A\cdot B} \quad \text{ovvero} \quad A\cdot B = \overline{\overline{A} + \overline{B}}$$
$$\overline{A+B} = \overline{A}\cdot\overline{B} \quad \text{ovvero} \quad A + B = \overline{\overline{A}\cdot\overline{B}}$$

Queste relazioni sono naturalmente valide per qualunque numero di variabili. Omettiamo la dimostrazione di questo teorema, che puo' essere verificato in modo diretto attraverso le tavole della verita'. Esso e', come vedremo,molto importante, perche' consente di trasformare una funzione in un'altra equivalente, attraverso ripetute applicazioni del teorema. Inoltre, da un punto di vista concettuale, dimostra che le tre operazioni logiche non sono indipendenti: l'operazione di OR puo' essere costruita con AND e NOT, ovvero l'operazione di AND puo' essere costruita con OR e NOT.
Nella tabella 7.2 sono riassunte tutte le proprieta' ed i teoremi dell'algebra di Boole.

Identita'	$A+0=A$	$A\cdot 1 = A$
Dominanza	$A+1=1$	$A\cdot 0 = 0$
Idempotenza	$A+A=A$	$A\cdot A = A$
Complementazione	$A+\overline{A}=1$	$A\cdot\overline{A}=0$
Involuzione	$\overline{\overline{A}} = A$	
Commutativa	$A+B=B+A$	$A\cdot B = B\cdot A$
Associativa	$A+(B+C)=(A+B)+C$	$A\cdot(B\cdot C)=(A\cdot B)\cdot C$
Distributiva	$(A+B)\cdot(A+C)=A+B\cdot C$	$A\cdot(B+C)=A\cdot B+A\cdot C$
Assorbimento I	$A+A\cdot B=A$	$A\cdot(A+B)=A$
Assorbimento II	$A+\overline{A}\cdot B=A+B$	$A\cdot(\overline{A}+B)=A\cdot B$
Assorbimento III	$A\cdot B+B\cdot C+\overline{A}\cdot C=$	$(A+B)\cdot(B+C)\cdot(\overline{A}+C)=$
	$A\cdot B+\overline{A}\cdot C$	$(A+B)\cdot(\overline{A}+C)$
De Morgan	$\overline{A+B}=\overline{A}\cdot\overline{B}$	$\overline{A\cdot B}=\overline{A}+\overline{B}$
Shannon	$F(A,B,C)=$	$F(A,B,C)=$
	$A\cdot F(1,B,C)+\overline{A}\cdot F(0,B,C)$	$=[A+F(0,B,C)]\cdot[\overline{A}+F(1,B,C)]$

Tabella 7.2: *Proprieta' e teoremi dell'Algebra di Boole*

7.2 Circuiti logici

In sostanza, da quanto abbiamo visto nel precedente paragrafo, qualunque calcolo si voglia fare su numeri binari si riconduce a funzioni in cui compaiono le tre operazioni logiche AND,OR e NOT. Da un punto di vista circuitale questo significa che avendo a disposizione tre circuiti che svolgono queste funzioni elementari, ogni altra funzione puo' essere realizzata combinando in modo opportuno questi tre mattoni fondamentali.
Ma prima di tutto dobbiamo stabilire una associazione tra variabili logiche e variabili elettriche. Chiaramente questa associazione e' del tutto arbitraria, tuttavia e' chiara la necessita' di uniformare questa scelta in modo da rendere compatibili e intercambiabili circuiti realizzati da diversi produttori o comunque da diversi utenti. Esistono varie convenzioni, ma tutte perlopiu' considerano come variabili elettriche rilevanti le tensioni e quindi il circuito logico e' un sistema in cui, applicando certe tensioni agli ingressi, si ottiene una tensione d'uscita che corrisponde ad una ben definita funzione logica degli ingressi. Indipendentemente dai dettagli di realizzazione si usa uniformemente una simbologia come quella in Fig. 7.1, in cui sono mostrati i simboli dei circuiti (o porte logiche) basilari.

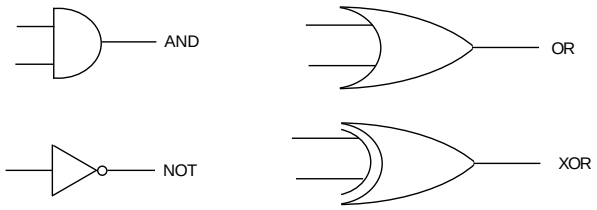

Figura 7.1: *Simboli delle porte logiche fondamentali*

E' importante notare che l'associazione di una tensione con un valore logico puo' essere fatta in 2 modi:

$$\begin{aligned} &0 \text{ logico} \quad \rightarrow \text{tensione bassa} \\ &1 \text{ logico} \quad \rightarrow \text{tensione alta} \end{aligned}$$

Ovvero:

$$\begin{aligned} &0 \text{ logico} \quad \rightarrow \text{tensione alta} \\ &1 \text{ logico} \quad \rightarrow \text{tensione bassa} \end{aligned}$$

Nel primo caso si ha quella che si chiama una logica *positiva*; nel secondo una logica *negativa*. C'e' una interessante relazione tra queste due scelte. Infatti il teorema di De Morgan ci dice che:

$$Y = AB = \overline{(\overline{A} + \overline{B})}$$

Il secondo membro puo' essere pensato come un OR in logica negativa: infatti negare una variabile e' del tutto equivalente a passare da logica positiva a logica negativa. La

conseguenza e' che un AND in logica positiva equivale ad un OR in logica negativa. Analogamente, poiche' si ha:

$$Y = A + B = \overline{\overline{A} \cdot \overline{B}}$$

possiamo dire che un OR in logica positiva e' equivalente ad un AND in logica negativa. Quindi lo stesso circuito svolge le funzioni di AND o di OR a seconda di quale associazione logica noi facciamo tra livelli di tensione e valori logici.
Un'ulteriore conseguenza del teorema di De Morgan e' che. come abbiamo gia' detto, AND e OR sono collegati tra loro, per cui in linea di principio non e' necessario costruire i circuiti che realizzano le tre funzioni fondamentali, ma ne bastano due, per esempio AND e NOT, oppure OR e NOT (si noti che il NOT e' comunque necessario). In realta' questo suggerisce di considerare come porta *fondamentale* il cosidetto NAND, ovvero la porta che realizza la funzione

$$Y = \overline{A \cdot B}$$

la cui tavola della verita' e' data da:

A	B	Y
0	0	1
0	1	1
1	0	1
1	1	0

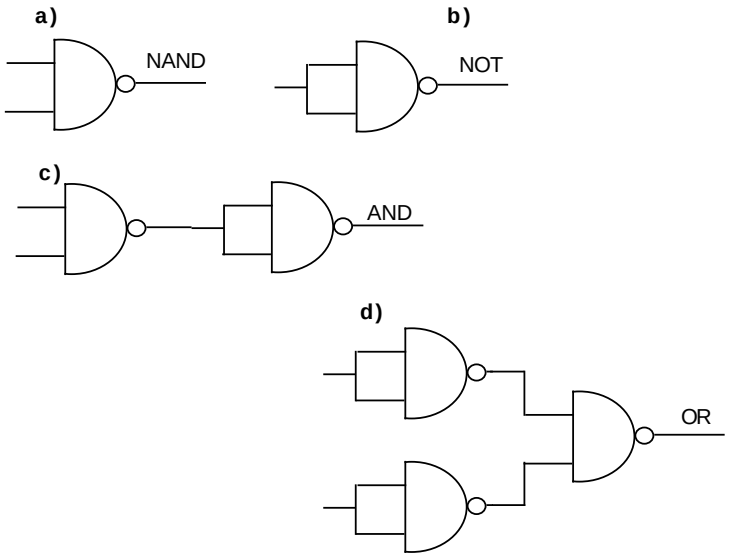

Figura 7.2: *a) Simbolo del NAND; b) Realizzazione del NOT c) Realizzazione dell' AND d) Realizzazione dell'OR*

Come si vede dalla Fig. 7.2 con una o piu' porte NAND e' possibile realizzare le tre operazioni elementari e, quindi, tutte le funzioni logiche.

Da quanto detto si comprende quindi che la costruzione di funzioni logiche piu' o meno complesse viene fatto costruendo delle reti formate da porte logiche; e' chiaro quindi che non e' sufficiente definire i livelli logici (cioe' la corrispondenza tra tensioni e variabili logiche), ma occorre anche che. per esempio l'uscita di una porta possa essere connessa all'ingresso di un'altra. Questo significa una piu' completa compatibilita' in termini di correnti, impedenze, ecc. Inoltre si vuole che l'uscita di una porta possa servire piu' ingressi di porte diverse, e questo pone ulteriori esigenze, come vedremo nei prossimi paragrafi.
In generale una certa funzione logica, grazie al teorema di De Morgan, puo' essere scritta in piu' modi; questo si traduce in diverse realizzazioni circuitali. Vediamo, a titolo di esempio, il caso dell'OR esclusivo. Si ha

$$Y = A\overline{B} + \overline{A}B$$

Ma, usando il teorema di De Morgan, si puo' anche scrivere:

$$Y = (A + B) \cdot (\overline{A \cdot B})$$

$$Y = \overline{A \cdot B + \overline{A} \cdot \overline{B}}$$

$$Y = (A + B) \cdot (\overline{A} + \overline{B})$$

A ciascuna di queste forme corrisponde ovviamente una realizzazione circuitale diversa (Fig. 7.3), di maggiore o minore complessita' e costo. Si sono quindi sviluppate tecniche, che noi non tratteremo, per ottimizzare la traduzione circuitale di una funzione logica, nel senso di minimizzare il numero di porte elementari necessarie a realizzarla.

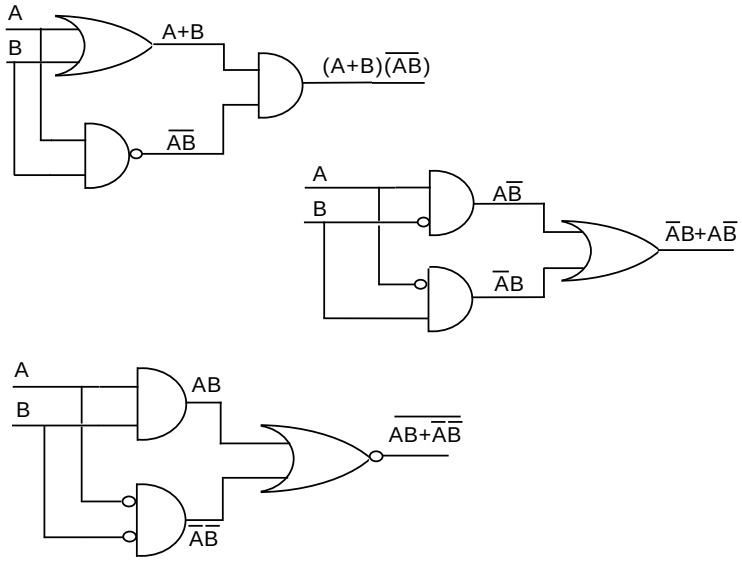

Figura 7.3: *Varie realizzazioni dell'OR ESCLUSIVO*

7.3 Famiglie logiche

Una famiglia logica e' definita da un insieme di regole e convenzioni, tali da rendere possibile la costruzione di circuiti complessi utilizzando circuiti logici elementari compatibili

pienamente tra loro. Nel corso degli anni si sono sviluppate varie famiglie logiche e sul mercato esiste una vastissima offerta i circuiti appartenenti ad esse. Le più importanti famiglie logiche in commercio sono

DTL	Diode-Transistor Logic
RTL	Resistor-Transistor Logic
TTL	Transistor-Transistor Logic
DCTL	Direct-Coupling Transistor Logic
ECL	Emitter-Coupled Logic
MOS	Basata su transistor a effetto di campo

Noi studieremo ed utilizzeremo in Laboratorio circuiti appartenenti alla famiglia TTL. Tuttavia, e' interessante ed istruttivo parlare brevemente della famiglia DTL, che oggi non e' piu' molto diffusa, ma ci consente comprendere alcune delle cose dette nel precedente paragrafo

7.3.1 Famiglia DTL

Le funzioni elementari AND e OR possono essere realizzate solo con diodi, mentre il NOT richiede l'uso di un transistor. I livelli logici sono (in logica positiva):

$$\begin{aligned} 0 \text{ logico} &\quad \rightarrow 0\ V \\ 1 \text{ logico} &\quad \rightarrow 12\ V \end{aligned}$$

Naturalmente la corrispondenza si inverte passando alla logica negativa. In Fig. 7.4 sono mostrate le realizzazioni di AND e OR e si vede come passando da logica positiva a logica negativa lo stesso circuito svolge entrambe le funzioni. Invece NOT e NAND mostrati nella stessa figura sono pensati per logica positiva; per passare a logica negativa occorre sostituire i transistors *npn* con transistors *pnp*.
Si capisce da questi esempi che i livelli logici devono essere definiti con certe tolleranze; infatti nel circuito NOT mostrato l'uscita a livello logico 0 corrisponde ad una tensione reale di circa 0.2 V (V_{CEsat} del transistor). In generale quindi i livelli logici corrispondono a due intervalli di tensione, ben separati tra loro, in modo che eventuali disturbi non alterino il significato logico delle tensioni ed il comportamento dei circuiti

7.3.2 Famiglia TTL

E' la famiglia logica piu' nota ed utilizzata, e costituisce uno standard industriale. I livelli logici nominali (in logica positiva) sono:

$$\begin{aligned} 0 \text{ logico} &\quad \rightarrow 0.2\ V \\ 1 \text{ logico} &\quad \rightarrow 5\ V \end{aligned}$$

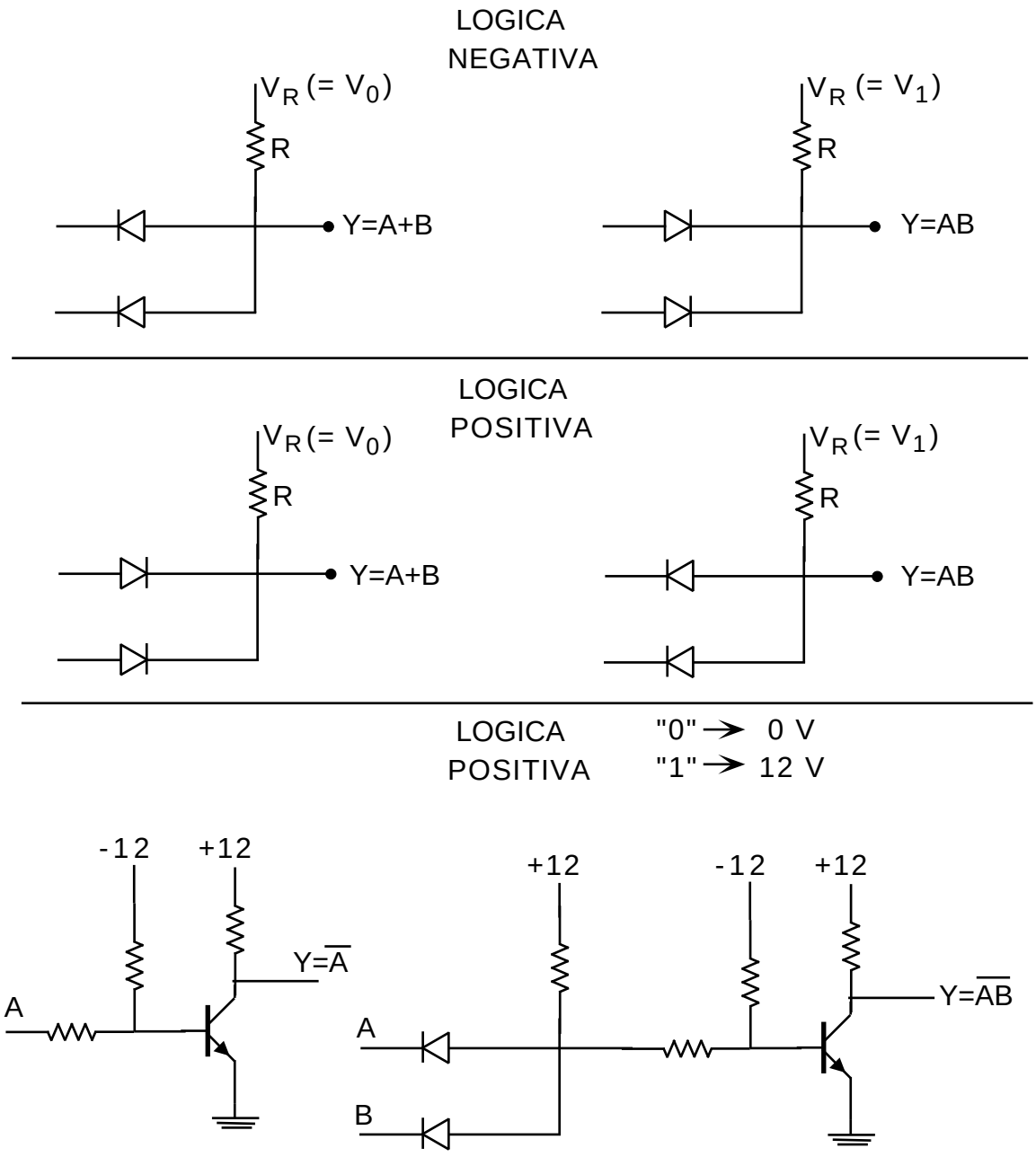

Figura 7.4: *AND, OR, NOT e NAND nella famiglia DTL*

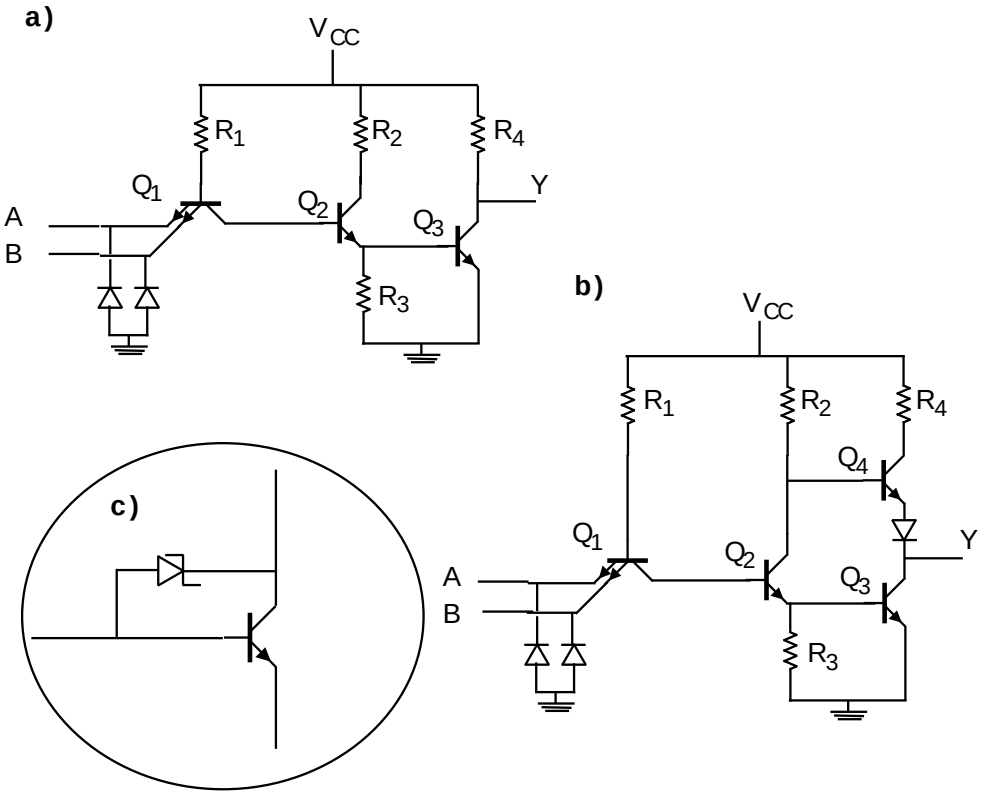

Figura 7.5: *NAND TTL. a) Schema di principio; b) Circuito migliorato; c) Dettaglio dell'uscita con diodo Schottky*

La porta logica fondamentale e' il circuito NAND, il cui schema di principio e' mostrato in Fig. 7.5a. Lo schema di principio non e' molto diverso dal NAND DTL; i diodi di ingresso sono sostituiti da un transistor ad emettitori multipli, facile da realizzare con la tecnologia dei circuiti integrati. Se almeno un ingresso è a $0.2\ V$ la base del transistor Q_1 è a $0.2 + 0.7 = 0.9\ V$. Per far condurre la giunzione di collettore di Q_1 e i transistor Q_2 e Q_3, occorrerebbe che la base di Q_1 fosse a $0.7 + 0.7 + 0.7 = 2.1\ V$. Quindi Q_2 e Q_3 sono interdetti e $Y = V_{CC}$. Se invece A e B sono a $5\ V$, la giunzione base-emettitore di Q_1 è polarizzata inversamente, la base di Q_1 sale di tensione, Q_2 e Q_3 entrano in saturazione. I diodi all'ingresso proteggono il circuito da eventuali ed indesiderate tensioni di ingresso negative.

La famiglia TTL appartiene alle cosidette logiche saturate, in quanto uno dei livelli logici corrisponde allo stato di saturazione del transistor d'uscita. Naturalmente cio' comporta una elevata dissipazione di potenza, quando l'uscita e' bassa, legata al valore della resistenza R_4. Da questo punto di vista converrebbe avere questa resistenza molto grande; tuttavia cio' creerebbe problemi con il tempo di commutazione del circuito (cioe' il tempo che l'uscita impiega a modificare il suo livello un relazione ad un cambiamento logico degli ingressi). Infatti questo tempo è dato da $\tau = R_4 C_L$, dove C_L è la capacità connessa con l'uscita (cioe' la capacita' d'ingresso dello stadio successivo.

Conviene quindi migliorare il circuito (Fig. 7.5b),con l'aggiunta di un ulteriore transistor; ora R_4 e' una resistenza piccola. Quando l'uscita e' alta ($Y = 1$) Q_2 è interdetto, Q_3 è

interdetto, Q_4 è in conduzione (cioè $Y = V_{CC} - V_{BE} - V_D \approx 3 \div 4\ V$). Quando l'uscita e' bassa ($Y = 0$), Q_2 e Q_3 sono in saturazione e Q_4 è interdetto, e ciò è ottenuto grazie al diodo D. Si noti quindi che ora non c'e' mai corrente che circola in R_4; quando l'uscita e' bassa il transistor Q_3 assorbe corrente dall'ingresso dello stadio successivo. Vediamo ora la commutazione dell'uscita da 0 a 1: Q_2 si interdice e Q_3 anche; Q_4 entra in saturazione e D conduce. In questo modo Q_4 fornisce la corrente per caricare C_L con una costante di tempo

$$\tau = (R_4 + R_S + R_D)C_L$$

dove R_S e' la resistenza di Q_4 in saturazione e R_D la resistenza diretta del diodo: l'uscita Y sale e si porta a $5\ V$. Per la commutazione da 1 a 0, Q_2 e Q_3 entrano in conduzione (e poi in saturazione), C_L si scarica attraverso Q_3; la tensione di Y scende e si porta in conduzione. Questo tipo di realizzazione dell'uscita e' comunemente noto come uscita a totem pole.

Un ulteriore miglioramento puo' essere ottenuto inserendo un diodo Schottky tra la base ed il collettore di Q_4 (Fig. 7.5c). Il diodo è di metallo-semiconduttore, con tempo di immagazzinamento trascurabile e soglia di circa $3\ V$. Esso impedisce che il transistor entri in saturazione e quindi rende più veloce la commutazione.

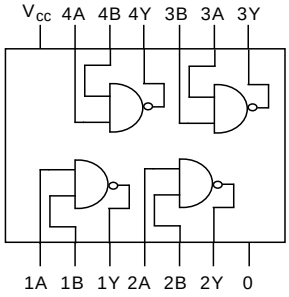

Figura 7.6: *Diagramma dell'integrato 7400, in contenitore Dual-in-line a 14 piedini*

Esistono in commercio circuiti integrati della famiglia TTL che svolgono varie funzioni; molto spesso lo stesso circuito integrato contiene piu' porte logiche identiche. La serie piu' diffusa e' la cosidetta serie 74, che comprende un grandissimo numero di componenti. Ad esempio, l'integrato 7400 racchiude in un contenitore a 14 piedini 4 porte NAND a due ingressi (Fig. 7.6); il 7410 contiene 3 porte NAND a tre ingressi. Esistono altre versioni della serie 74, contraddistinte da ulteriori caratteri nella sigla: per esempio la sigla 74LS00 si riferisce ad un circuito tipo Schottky, a bassa dissipazione (Low); invece la sola S indica circuiti tipo Schottky, caratterizzati da alta velocita' di commutazione, ma anche alta dissipazione. Le caratteristiche tipiche delle varie versioni sono riassunte nella Tabella 7.3. Abbiamo gia' detto che i livelli logici *effettivi* devono includere delle ampie tolleranze rispetto ai valori nominali. Nella Tabella 7.4 sono riportate, a titolo esemplificativo, le specifiche che un costruttore fornisce per i propri circuiti della serie 74.

Come si vede dalla tabella, il costruttore garantisce il corretto funzionamento del circuito se ai suoi ingressi vengono fornite tensioni non superiori a $0.8\ V$ come 0 logico e non inferiori a $2\ V$ come 1 logico; invece garantisce all'uscita una tensione non superiore a $0.4\ V$ come 0 logico (ma tipicamente del valore di 0.2) e non inferiore a $2\ V$ come 1 logico

Serie	Tempo di commutazione (nS)	Dissipazione (mW)
- standard	10	10
S Schottky	3	19
LS Low Schottky	9.5	2

Tabella 7.3: *Caratteristiche delle serie TTL*

Livello logico		Tensione (Volt)		
		minima	tipica	massima
Input	0			0.8
	1	2		
Output	0		0.2	0.4
	1	2.4	3.4	

Tabella 7.4: *Tolleranze di ingresso e uscita rispetto ai livelli logici TTL nominali.*

(tipicamente 3.4). E' chiaro quindi che ingressi compresi tra 0.8 e 2 V danno luogo ad un comportamento imprevedibile del circuito.

Un discorso analogo va fatto per quanto riguarda l'alimentazione. I circuiti TTL richiedono un'alimentazione nominale di 5 V; in realta' essi possono funzionare correttamente con tensioni di alimentazione fino a ~ 7 V. E' in genere conveniente quindi utilizzare una tensione di alimentazione un po' sovradimensionata rispetto al valore nominale, sia per avere una tensione d'uscita (a livello logico 1) un po' piu' alta; ma anche per compensare eventuali cadute di tensione dovute al carico complessivo dell'alimentatore.

7.4 Esempi di circuiti digitali

Vediamo ora alcuni esempi di circuiti piu' complessi che si possono realizzare partendo dalle porte logiche elementari.

7.4.1 Sommatori

Per comprendere come deve essere costruito un circuito capace di sommare due numeri binari possiamo anzitutto considerare la somma di due numeri a 1 bit (cioe' ad una sola cifra).

A	B	D	C	Somma
0	0	0	0	00
0	1	1	0	01
1	0	1	0	01
1	1	0	1	10

Vediamo dalla tabella che il risultato (colonna Somma) diviene di due cifre quando entrambi gli addendi sono 1; quindi l'operazione di somma produce sia un risultato (D) che un riporto (C) anch'essi mostrati nella tabella. Il riporto C, nel caso di numeri a piu' cifre, andrebbe aggiunto alla somma del bit immediatamente a sinistra. Come si vede D non e' altro che l'OR ESCLUSIVO degli ingressi, mentre C e' l'AND. Il circuito da costruire e' quindi quello mostrato in Fig. 7.7, e prende il nome di Semisommatore (Half Adder). Ora per

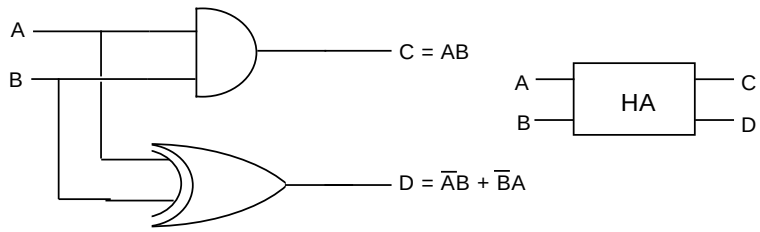

Figura 7.7: *Semisommatore; b) Simbolo complessivo*

realizzare la somma di numeri a piu' bits possiamo utilizzare lo schema di Fig. 7.8; come si vede per ogni cifra, tranne la meno significativa, sono necessari 2 Semisommatori, da cui il nome.

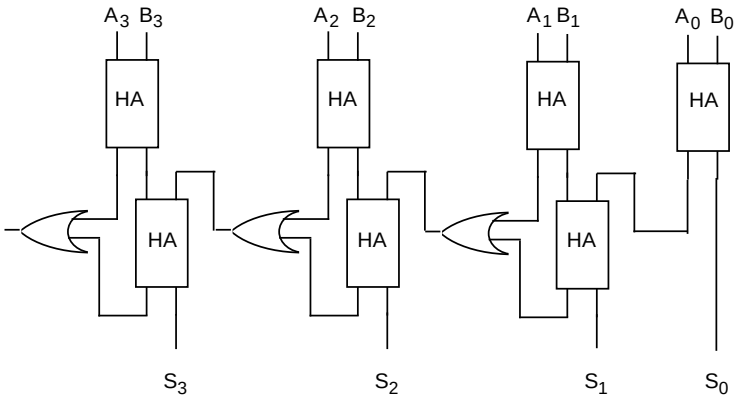

Figura 7.8: *Somma a molti bits*

La stessa funzione puo' essere realizzata in modo diverso. Supponiamo infatti di voler sommare due numeri, A e B, a molti bits; per il generico bit n noi dobbiamo fare una

somma a tre addendi, per tener conto dell'eventuale riporto C_{n-1} proveniente dalla cifra precedente, e fornire una somma, S_n, piu' un eventuale riporto, C_n. Le uscite logiche da fornire sono quindi:

$$\begin{aligned} S_n &= \overline{A}_n\overline{B}_nC_{n-1} + A_n\overline{B}_n\overline{C}_{n-1} + \overline{A}_nB_n\overline{C}_{n-1} + A_nB_nC_{n-1} \\ C_n &= \overline{A}_nB_nC_{n-1} + A_n\overline{B}_nC_{n-1} + A_nB_n\overline{C}_{n-1} + A_nB_nC_{n-1} \end{aligned}$$

come si puo' anche verificare dalla tavola della verita':

A_n	B_n	C_{n-1}	S_n	C_n
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

Il circuito completo diviene allora quello in Fig. 7.9 e prende il nome di Full Adder (o Sommatore completo). Circuiti sommatori sono naturalmente disponibili in forma integrata; per esempio, il 74LS283 e' un Full Adder a 4 bit; diversi esemplari possono essere collegati in cascata per realizzare somme a $8, 12, 16, \ldots$ bit.

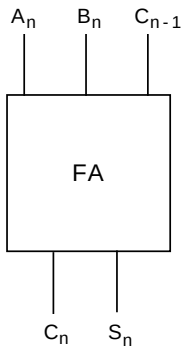

Figura 7.9: *Full adder*

E' interessante, a titolo di esercizio, vedere se e' possibile semplificare le espressioni di C_n ed S_n, sfruttando le proprieta' dell'algebra di Boole. Si ha

$$\begin{aligned} C_n &= \overline{A}_n B_n C_{n-1} + A_n \overline{B}_n C_{n-1} + A_n B_n \overline{C}_{n-1} + A_n B_n C_{n-1} \\ &= \overline{A}_n B_n C_{n-1} + A_n \overline{B}_n C_{n-1} + A_n B_n (\overline{C}_{n-1} + C_{n-1}) \\ &= \overline{A}_n B_n C_{n-1} + A_n \overline{B}_n C_{n-1} + A_n B_n \\ &= \overline{A}_n B_n C_{n-1} + A_n (\overline{B_n} C_{n-1} + B_n) \\ &= \overline{A}_n B_n C_{n-1} + A_n C_{n-1} + A_n B_n \\ &= C_{n-1}(\overline{A}_n B_n + A_n) + A_n B_n \\ &= C_{n-1}(A_n + B_n) + A_n B_n \\ &= A_n B_n + A_n C_{n-1} + C_{n-1} B_n \end{aligned}$$

Abbiamo utilizzato qui le proprieta' di complementazione e di assorbimento.
Con manipolazioni analoghe si puo' arrivare a semplificare anche l'espressione di S_n, con il risultato

$$S_n = C_{n-1}(\overline{A_n \oplus B_n}) + \overline{C_{n-1}}(A_n \oplus B_n)$$

Sottrazione. Vediamo come possiamo realizzare un circuito capace di effettuare una sottrazione tra due numeri binari, per esempio a 4 cifre.
In generale, dato un numero binario A, si definisce il suo complemento ad 1, $\overline{A}$, come il numero che si ottiene negando uno per uno tutti i bits. Si ha evidentemente che:

$$A + \overline{A} = 1111$$

se ora aggiungiamo 1

$$A + \overline{A} + 1 = 10000$$

Quindi possiamo anche dire che

$$A = 10000 - \overline{A} - 1$$

Adesso, volendo sottrarre tra loro i numeri B ed A si ha

$$B - A = (B + \overline{A} + 1) - 10000$$

Ad esempio, si voglia fare $12 - 9$. In binario

$$1100 - 1001 = (1100 + 0110 + 0001) - 10000 = 10011 - 10000 = 0011$$

cioè 3, come previsto. In pratica sottrarre 10000 significa ignorare il quinto bit, cioe' ignorare il bit fuori dalla dimensione degli operandi (bit di overflow).
L'operazione di sottrazione si fa quindi complementando il sottraendo ed eseguendo poi un'operazione di somma. In realtà tutto ciò funziona se $B > A$; se invece $B < A$ non c'e' riporto nel quinto bit, che rimane quindi zero. A questo punto per avere il risultato giusto (in valore assoluto), basta rifare il complemento a 2.
Si voglia ad esempio calcolare $13 - 15$. Si ha

$$1101 - 1111 = 1101 + 0001 = 01110$$

Il quinto bit e' zero, quindi facendo il complemento a 2 si ottiene 0010 cioe' il valore assoluto del risultato.
In pratica lo stato del bit di overflow ci fa capire il segno del risultato: se esso e' uguale ad 1 il risultato e' positivo, se uguale a 0 il risultato e' negativo (e va rieseguito il complemento a 2).

7.4.2 Moltiplicazione e divisione

Poiche' siamo in tema di operazioni numeriche conviene accennare brevemente alle operazioni di moltiplicazione e divisione, nel sistema binario. Moltiplicare un numero per due significa spostare tutti i bit di una posizione verso sinistra (cioè verso il bit più significativo). Moltiplicare per una potenza di due vuol dire ripetere più volte questa operazione; per moltiplicare un numero per un altro che non è potenza di due si scompone quest'ultimo in somme di potenze di due e si sommano i risultati. Si voglia fare ad esempio 27×33; possiamo scrivere

$$27 \times 33 = 27 \times (32 + 1) = 27 \times 2^5 + 27 \times 2^0$$

Quindi, poiche' il numero 27 in binario e' 11011, il risultato e' ottenuto dalla somma

$$1101100000 + 110110 = \ldots$$

dove abbiamo sommato il numero che si ottiene da 27 con 5 traslazioni verso sinistra con il numero ottenuto da 27 con 1 traslazione verso sinistra. Puo' sembrare complicato scomporre un numero in potenze di 2, ma e' invece semplicissimo in binario. Dato infatti un qualunque numero, per es. 1101100, esso si scompone in

$$\begin{array}{ll} 1000000 & + \\ 0100000 & + \\ 0001000 & + \\ 0000100 & = \\ \hline 1101100 & \end{array}$$

Cioe' basta prendere tutti i numeri che si ottengono azzerando via via tutti i bits tranne 1.
L'operazione di divisione si svolge in modo analogo, traslando i numeri verso destra, anziche' verso sinistra.

Vedremo piu' avanti come si possono realizzare traslazioni (verso destra o verso sinistra) di gruppi di bits.

7.4.3 Comparatore digitale

Abbiamo visto che e' spesso necessario confrontare tra loro due numeri binari per stabilire qual'e' il maggiore. Come al solito cominciamo ad affrontare il problema considerando per ora due numeri ad un solo bit. Siano A e B i nostri due numeri; possiamo costruire 3 funzioni logiche

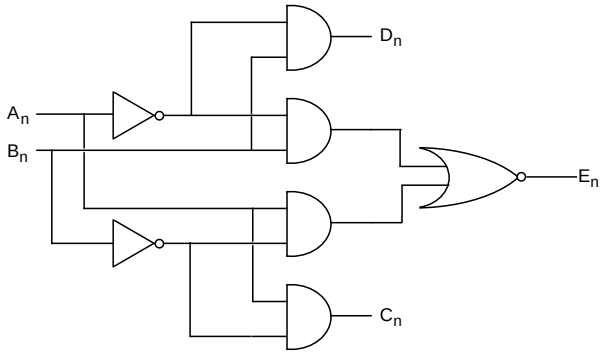

Figura 7.10: *Comparatore digitale ad 1 bit*

$$\begin{aligned} E_0 = \overline{A\overline{B} + \overline{A}B} &= 1 \quad \text{se } A = B \\ C_0 = A\overline{B} &= 1 \quad \text{se } A > B \\ D_0 = \overline{A}B &= 1 \quad \text{se } A < B \end{aligned}$$

Questi tre operatori ci consentono di effettuare il confronto. Chiaramente, per numeri a molti bits, per esempio 4, avremo:

$$E = E_3E_2E_1E_0 \begin{cases} = 1 & A = B \\ = 0 & A \neq B \end{cases}$$

D'altra parte la condizione $A > B$ e' vera se
$A_3 > B_3$
o $A_3 = B_3$ e $A_2 > B_2$
o $A_3 = B_3$ e $A_2 = B_2$ e $A_1 > B_1$
o $A_3 = B_3$ e $A_2 = B_2$ e $A_1 = B_1$ e $A_0 > B_0$
Cioè corrisponde all'operatore

$$C = A_3\overline{B_3} + E_3A_2\overline{B_2} + E_3E_2A_1\overline{B_1} + E_3E_2E_1A_0\overline{B_0}$$

Quindi

$$C = \begin{cases} 1 & \text{se } A > B \\ 0 & \text{altrimenti} \end{cases}$$

La condizione $A < B$ si ottiene in modo analogo, scambiando A con B. Nella Fig. 7.10 e' mostrato lo schema dell circuito che realizza la comparazione ad 1 bit.
Anche in questo caso sono reperibili in commercio circuiti che effettuano la comparazione: per esempio, il 74LS85 e' un comparatore a 4 bits, utilizzabile anche in cascata per comparare numeri a dimensione maggiore

Naturalmente l'industria dei circuiti integrati ha sviluppato anche circuiti capaci di compiere piu' operazioni aritmetiche o logiche. Questi circuiti sono comunemente noti come

Arithmetic and Logic Units (ALU) e costituiscono a loro volta uno degli elementi funzionali dei sistemi di elaborazione piu' complessi. Ad esempio, l'integrato 74LS381 e' una ALU a 4 bits, che puo' svolgere le seguenti operazioni:

Operazioni aritmetiche	Operazioni logiche
$B - A$	$A + B$
$A - B$	$A \oplus B$
$A + B$	$A \cdot B$

7.5 Circuiti sequenziali

I circuiti che abbiamo visto finora sono di tipo *combinatorio*, cioe' l'uscita (o le uscite) e' ad ogni istante una certa funzione logica degli ingressi. La variabile tempo non e' quindi concettualmente rilevante ai fini della funzionalita' di questi circuiti, ma lo e' solo da un punto di vista pratico, nel senso che inevitabilmente ogni circuito ha un tempo di commutazione finito e non puo' variare istantaneamente il suo stato.

Vi sono invece circuiti in cui la temporizzazione dei fenomeni e' intrinsecamente rilevante; essi prendono il nome di circuiti sequenziali. L'elemento fondamentale di partenza e' il circuito bistabile (Fig. 7.11a): la connessione tra le due porte NOT conduce al fatto che esistono due stati stabili ed equiprobabili del circuito

$$Q = 1 \quad \overline{Q} = 0$$
$$Q = 0 \quad \overline{Q} = 1$$

Al momento dell'accensione del circuito esso si porra' in uno dei due stati: in linea di principio non e' possibile prevedere quale, in realta', poiche' e' impossibile realizzare un circuito esattamente simmetrico, in termini di velocita' di commutazione delle porte, ci sara' uno stato preferito, in cui il circuito si pone al momento dell'accensione.

7.5.1 Flip-flops

Partendo dal bistabile possiamo costruire un circuito in cui possiamo assegnare e modificare lo stato delle uscite (Fig. 7.11b). Questo circuito prende il nome di flip-flop SET-RESET, ovvero **flip-flop S-R**. L'ingresso C_k (detto ingresso di clock) serve in sostanza ad abilitare il circuito: infatti, finche' C_k e' uguale a zero, l'uscita delle porte 1 e 2 e' sempre 1, indipendentemente dallo stato di S e di R; cio' significa che le porte 3 e 4 restano nello stato in cui si trovavano inizialmente. Per comprendere allora il funzionamento del circuito dobbiamo allora procedere nel seguente modo:
si applicano ad S e R due livelli logici;
si porta il clock ad 1 per un certo tempo;
si osserva lo stato delle uscite quando l'impulso di clock e' finito; la tavola della verita' del circuito e'allora

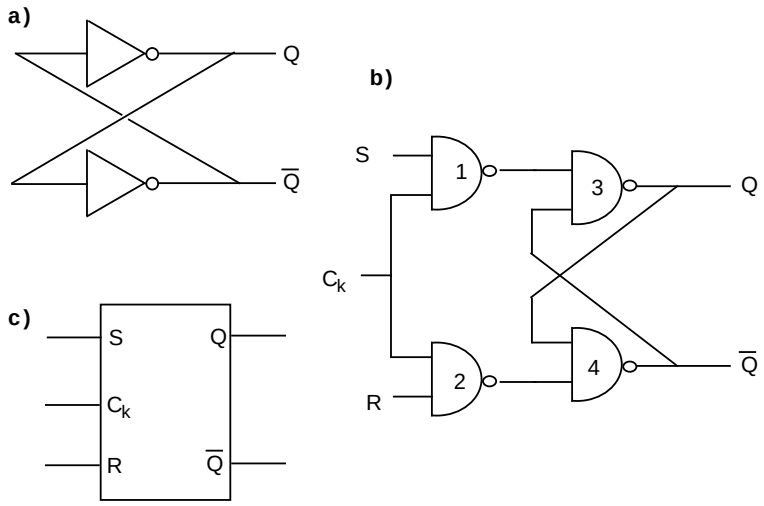

Figura 7.11: *a) Circuito bistabile; b) Flip-flop S-R c) Simbolo circuitale*

S_n	R_n	Q_{n+1}
0	0	Q_n
1	0	1
0	1	0
1	1	?

dove gli indici n ed $n+1$ stanno proprio a ricordare il significato sequenziale: S_n e R_n sono gli stati degli ingressi applicati prima dello n-esimo impulso di clock, Q_{n+1} e' lo stato dell'uscita dopo la fine di questo impulso. E' facile verificare il comportamento sopra descritto: se $S = 0$ e $R = 0$ l'uscita delle porte 1 e 2 resta uguale ad 1 prima, durante e dopo l'impulso di clock, quindi l'uscita non cambia e rimane uguale a Q_n. Se $S = 1$ e $R = 0$, l'uscita della porta 1 diviene 0 (durante l'impulso di clock) e questo forza ad 1 l'uscita della porta 3. Di conseguenza gli ingressi della porta 4 sono ora ad 1 e 1, percio' $\overline{Q}$ va a 0. Finito l'impulso di clock questo stato resta *congelato*. Il circuito si comporta ovviamente in modo analogo nel caso inverso: $S = 0$ e $R = 1$ porta $Q = 0$ e $\overline{Q} = 1$. Cosa succede invece se $S = 1$ e $R = 1$? Durante l'impulso di clock le uscite 1 e 2 vanno entrambe a 0 e quindi Q e $\overline{Q}$ vanno ad 1; questo stato persiste finche' e' presente il clock, ma quando esso finisce il bistabile formato dalle porte 3 e 4 deve mettersi in uno dei suoi due stati stabili. Non e' prevedibile teoricamente quale dei due, perche' chiaramente esso dipendera' da quale delle due porte NAND e' piu' veloce a commutare e quindi a forzare lo stato dell'altra.

Flip-Flop J-K. E' possibile costruire un flip-flop che dia sempre risposte definite, utilizzando lo schema in Fig. 7.12a. Come si vede le uscite sono ora riportate agli ingressi dello stadio di abilitazione; questo in sostanza equivale allo schema di Fig. 7.12b, che possiamo usare per comprendere il funzionamento. Tenendo presente che ora $S = J\overline{Q}$ e $R = KQ$ abbiamo la seguente tavola della verita':

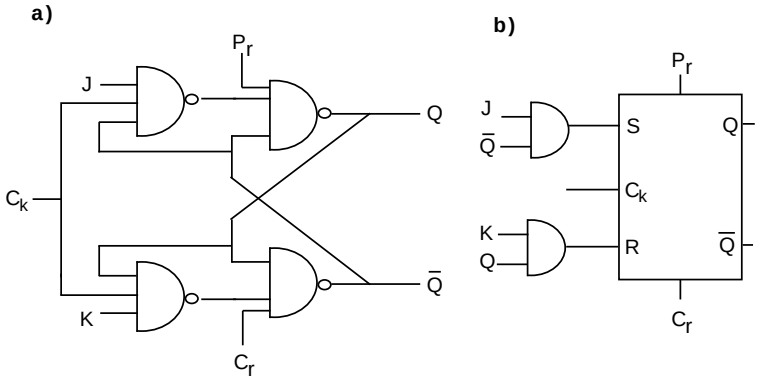

Figura 7.12: *a) Flip-flop J-K; b) Schema equivalente*

J_n	K_n	Q_n	$\overline{Q}_n$	S_n	R_n	Q_{n+1}
0	0	0	1	0	0	Q_n
0	0	1	0	0	0	Q_n
1	0	0	1	1	0	1
1	0	1	0	0	0	$Q_n = 1$
0	1	0	1	0	0	$Q_n = 0$
0	1	1	0	0	1	0
1	1	0	1	1	0	$\overline{Q}_n\ (= 1)$
1	1	1	0	0	1	$\overline{Q}_n\ (= 0)$

Le ultime due righe mostrano la differenza rispetto al S-R; ora quando entrambi gli ingressi sono ad 1 l'uscita commuta rispetto allo stato precedente. In definitiva, la tavola della verita' del flip-flop J-K e':

J_n	K_n	Q_{n+1}
0	0	Q_n
1	0	1
0	1	0
1	1	$\overline{Q}_n$

Gli ingressi P_r (pre-set) e C_r (clear) servono a fissare lo stato iniziale del circuito indipendentemente dal clock, cioe' in modo che potremmo definire asincrono. Infatti si ha:

C_r	P_r	Q	$\overline{Q}$
0	1	0	1
1	0	1	0

Quindi, dopo aver fissato lo stato desiderato, occorre mantenere C_r e P_r ad 1 per abilitare il flip-flop ($P_r = 0$, $C_r = 0$ non deve essere usato perchè porta a situazioni ambigue).

In realtà il flip-flop J-K ha degli inconvenienti, che inficiano il funzionamento che abbiamo descritto. Per esempio, supponiamo di essere nello stato $Q = 0$ e di porre J=1, K=1. Quando arriva l'impulso di clock l'uscita Q si porta ad 1, dopo un ritardo Δt legato ai naturali tempi di commutazione. Ora, se il clock e' ancora attivo, cio' provoca un'ulteriore commutazione (J e K infatti sono ancora ad 1) e l'uscita Q ritorna a 0. Cioè finche' e' presente l'impulso di clock l'uscita oscilla continuamente fra 0 e 1 e lo stato finale non e' chiaramente predicibile. Questo inconveniente potrebbe essere evitato solo calcolando accuratamente la durata dell'impulso di clock, in relazione ai tempi di commutazione del circuito, in modo da assicurarsi che l'uscita abbia tempo di commutare solo una volta. Evidentemente non e' una soluzione praticabile, se non in situazioni molto particolari.

E' invece possibile ovviare a questo inconveniente con il cosidetto schema *master-slave* (Fig. 7.13) Si hanno sostanzialmente due stadi successivi, ma il secondo stadio e' abilitato

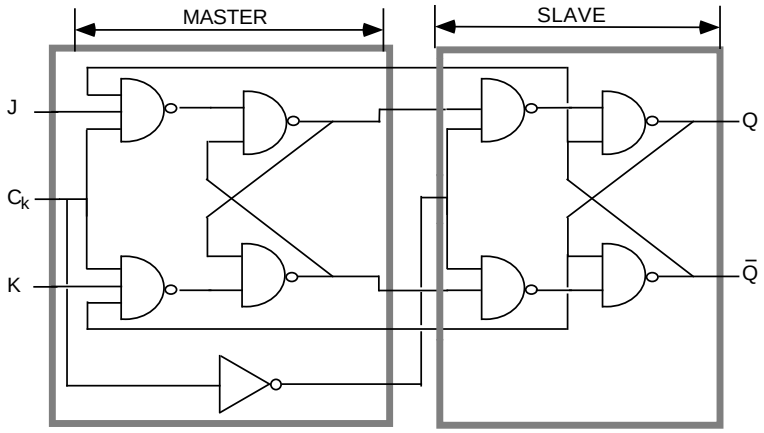

Figura 7.13: *a) Flip-flop J-K master-slave*

dal segnale di clock negato, mentre le uscite del bistabile finale sono riportate agli ingressi del primo stadio. Ora, durante l'impulso di clock il secondo stadio e' inibito, quindi le uscite non possono commutare. Invece, quando cessa l'impulso di clock, il primo stadio e' inibito, e lo stato delle uscite Q_M e $\overline{Q}_M$ e' trasferito su Q e $\overline{Q}$. Gli ingressi P_r, C_r servono per fissare lo stato iniziale, come nel caso precedente.

Flip-flop di tipo D e di tipo T Un flip-flop S-R (o J-K) in cui gli ingressi sono collegati come in Fig, 7.14a realizza il cosidetto tipo D (delay). Ora gli ingressi sono sempre opposti

tra loro, quindi la tavola della verita' e' sostanzialmente

D_n	Q_{n+1}
1	1
0	0

In sostanza questo circuito produce un ritardo di un ciclo di clock (da cui il nome). Esso rappresenta sostanzialmente una memoria a 1 bit: l'informazione viene scritta sul circuito (abilitando il clock) e permane su di esso, potendo essere riletta attraverso l'uscita Q. Questo circuito e' comunemente chiamato *transparent latch*, perche', nell'intervallo di tempo in cui il clock e' alto, l'uscita *vede* lo stato dell'ingresso come se il circuito fosse trasparente.

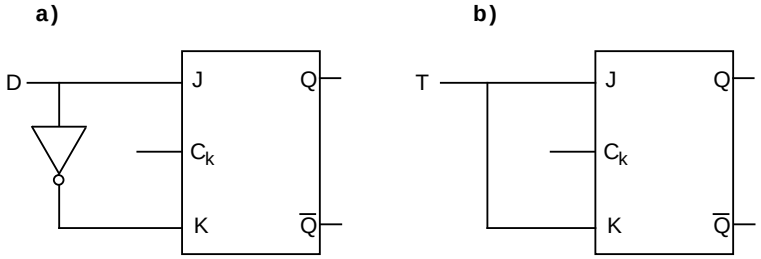

Figura 7.14: *a) Flip-flop D; b) Flip-flop T*

Invece, quando gli ingressi di un flip-flop J-K (master-slave) sono connessi direttamente tra loro (Fig, 7.14b) si ha il cosidetto tipo T (toggle). Questo circuito cambia il valore di Q ad ogni ciclo di clock; la sua tavola della verita' e' semplicemente

T_n	Q_{n+1}
1	$\overline{Q_n}$
0	Q_n

Edge triggered flip-flops. Finora abbiamo studiato dei flip-flops sensibili al livello del clock, cioe' abilitati a commutare, se necessario, solo durante l'intervallo di tempo in cui il segnale di clock e' alto. Esistono invece flip-flops sensibili solo al fronte del segnale di clock, comunemente noti come edge triggered. In questo caso il circuito e' sensibile allo stato degli ingressi solo durante un breve intervallo di tempo appena precedente (o, raramente appena successivo) al fronte di salita del clock. La Fig. 7.15 fa comprendere la differenza delle due situazioni. Il set-up time (tipicamente $15-20\ nS$) e' l'intervallo di tempo in cui gli ingressi devono essere stabili per consentire un corretto funzionamento del dispositivo, mentre esso e' indifferente allo stato degli ingressi al di fuori di questo intervallo. Il set-up time puo' essere tutto precedente il fronte del clock, ovvero parzialmente a cavallo: in tal caso la frazione successiva al fronte prende il nome di hold-time. Puo' sembrare paradossale il fatto che il flip-flop sia sensibile allo stato degli ingressi *prima* del clock; in realta' cio' e'

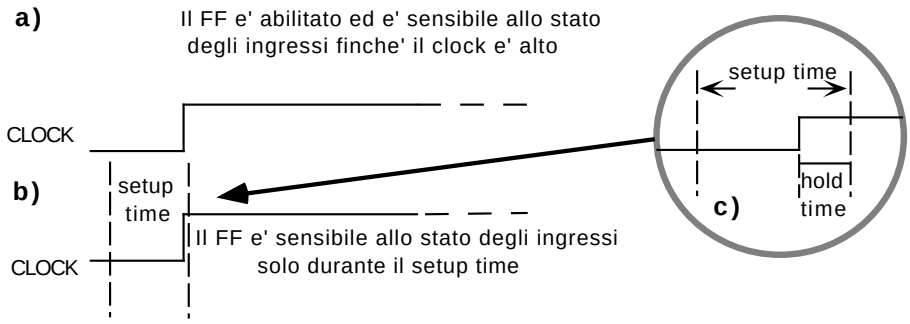

Figura 7.15: *a) Flip-flop sensibile al livello; b) Flip-flop sensibile al fronte del clock c) Dettaglio del set-up time e del hold-time*

ottenibile semplicemente giocando sui ritardi interni tra le linee degli ingressi e la linea del clock.
Chiaramente questo tipo di funzionamento elimina il problema della corsa critica, e quindi non c'e' la necessita' di costruire sistemi master-slave. Il simbolo circuitale dei flip-flops

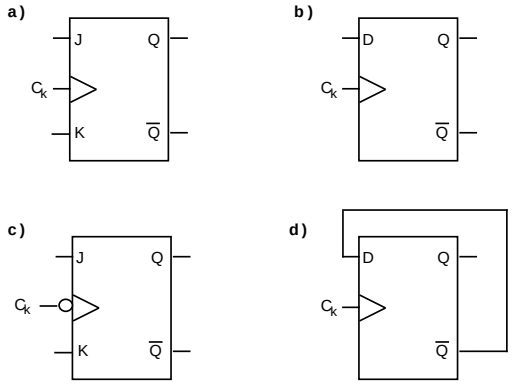

Figura 7.16: *a) Simbolo del J-K edge triggered; b) Simbolo del D edge triggered; c) FF sensibile al fronte di discesa d) Il tipo T ottenuto dal D*

edge triggered e' simile a quello dei circuiti sensibili al livello (Fig. 7.16; l'unica differenza e' un triangolino posto all'ingresso del clock. Esistono anche dispositivi che si attivano sul fronte di discesa del clock: sono distinguibili simbolicamente tramite un cerchietto sull'ingresso di clock.
Naturalmente esistono nella famiglia TTL dispositivi di questo tipo: il 7474 e' un integrato che contiene due flip-flops tipo D edge triggered, mentre il 74112 contiene 2 J-K.
E' interessante notare che, con un tipo D e' possibile ottenere un tipo T semplicemente riportando all'ingresso l'uscita $\overline{Q}$ (Fig. 7.16d). Ad ogni impulso di clock l'uscita Q commuta passando da 0 a 1 e viceversa.

7.5.2 Shift register

Lo shift register (in italiano registro a scorrimento) è formato da n flip-flop di tipo J-K (o S-R) master-slave. Consente di memorizzare n bits, che possono essere caricati sia in modo

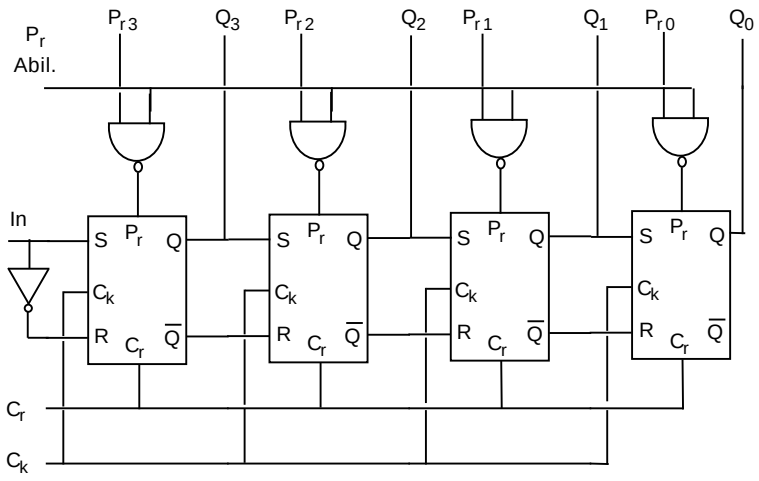

Figura 7.17: *Shift register a 4 bits*

seriale, attraverso l'ingresso di sinistra, sia in parallelo, attraverso gli ingressi di preset. Il contenuto puo' essere riletto sia in parallelo, che in serie. Vediamo ora in dettaglio le varie operazioni possibili, riferendoci all'esempio in Fig. 7.17:

1. Azzeramento: si pone $C_r = 0$, $P_{rAbil} = 0$;

2. Caricamento parallelo: $C_r = 1$, $P_{rAbil} = 1$ $P_{rj} = 1/0$: i vari flip-flops vengono caricati con il contenuto presentato agli ingressi P_{rj};

3. Lettura parallela: si effettua dalle uscite Q_i;

4. Caricamento seriale: si presenta il primo bit da caricare sull'ingresso seriale; fornendo un impulso di clock l'informazione viene caricata sul flip-flop F_4; successivamente si presenta il secondo bit e si fornisce un secondo impulso di clock: il primo bit passa al flip-flop F_3 e in F_4 entra il secondo bit; continuando si possono caricare tutti i flip-flops;

5. Lettura seriale: fornendo ulteriori impulsi di clock i bits memorizzati si presentano successivamente sull'uscita Q_0.

Si possono costruire registri in grado di effettuare scorrimenti verso destra e verso sinistra. Essi possono quindi essere utilizzati per effettuare moltiplicazioni e divisioni, che, come abbiami visto, richiedono questo tipo di manipolazione. In generale lo Shift register può essere usato come memoria, come convertitore serie-parallelo, parallelo-serie, ritardo digitale, ecc., e costituisce quindi un circuito molto importante in una grande varieta' di applicazioni.
L'integrato 74LS95 e' un esempio di shift register a 4 bit, capace di compiere sia scorrimenti verso destra che verso sinistra, con ingressi e uscite sia seriali che parallele. Un'esempio piu' sofisticato e' il 74LS293 che e' un moltiplicatore a 4 bit: il moltiplicando viene caricato in parallelo, mentre il moltiplicatore viene caricato serialmente fornendo opportuni impulsi di clock. Il risultato viene letto in parallelo.

7.5.3 Contatore asincrono

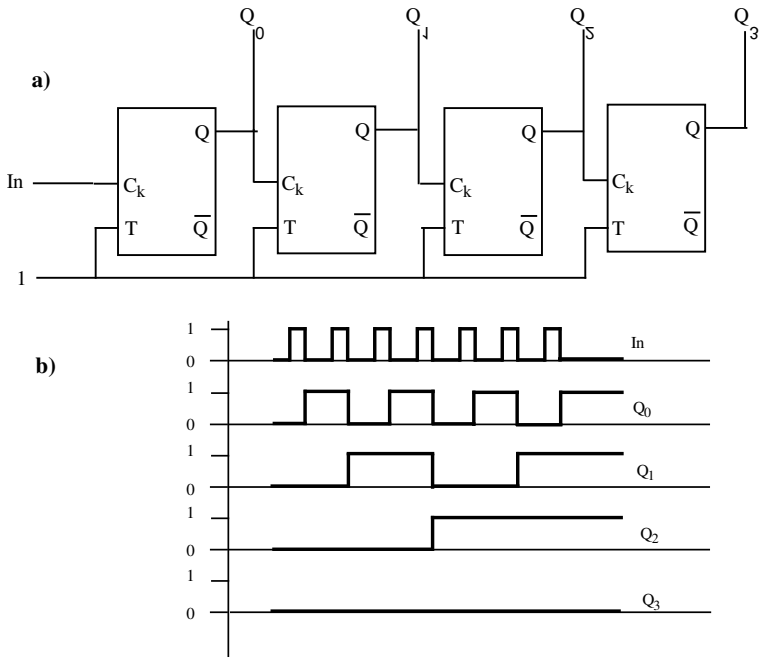

Figura 7.18: *a)Contatore asincrono a 4 bits; b)Diagramma temporale*

Si puo' costruire un contatore binario utilizzando flip-flops di tipo T (master-slave). Tutti gli ingressi vengono posti ad 1; gli impulsi da contare entrano nel clock del primo stadio, mentre ogni uscita fornisce il clock allo stadio successivo. Poiche' i flip-flops sono master-slave, ogni uscita commuta sul fronte di discesa del clock ricevuto all'ingresso (Fig. 7.18b). Supponendo quindi che inizialmente tutti i flip-flops siano a zero, lo stato delle uscite all'arrivare degli impulsi e' mostrato nella Tabella 7.5 Come si vede l'insieme delle uscite costituisce un numero a 4 bits (Q_0 e' il bit meno significativo) che conta il numero di impulsi arrivati In generale un contatore ad n stadi conta quindi fino a 2^n poi ricomincia da zero.

E' interessante notare che le varie uscite del contatore commutano a frequenze via via decrescenti: il bit 0 commuta alla stessa frequenza del segnale d'ingresso, il bit 1 a frequenza meta', il bit n a frequenza 2^{-n} volte piu' bassa. Percio' i contatori possono anche essere utilizzati come divisori di frequenza, per potenze di due.

Con opportuni aggiustamenti e' possibile realizzare divisori (e quindi contatori) per potenze diverse; esistono in commercio divisori per 5, 6, 10 e 12.

Contatore avanti-indietro. Un contatore che possa effettuare un conteggio in entrambe le direzioni e' detto contatore avanti-indietro (*up-down* in inglese). Per contare all'indietro occorre che l'ingresso di clock di ogni stadio sia collegato all'uscita $\overline{Q}$ dello stadio precedente, mentre il primo stadio resta invariato. In questo modo il primo stadio commuta ad ogni impulso, mentre gli stadi successivi commuteranno quando l'uscita $\overline{Q}$ proveniente dallo stadio precedente passa da 1 a 0. E' facile ora convincersi che, partendo da uno

Numero di impulsi	Q_3	Q_2	Q_1	Q_0
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	1
...				
15	1	1	1	1
16	0	0	0	0

Tabella 7.5: *Lo stato delle uscite in funzione degli impulsi in ingresso*

stato iniziale qualunque del contatore, l'arrivo di un impulso provoca un decremento del contenuto (si noti che, partendo dallo stato iniziale 0000, si transisce allo stato 1111).
Un contatore avanti-indietro puo' essere realizzato come nello schema di Fig. 7.19: il segnale di controllo, X, determina la direzione del conteggio. In un contatore asincrono

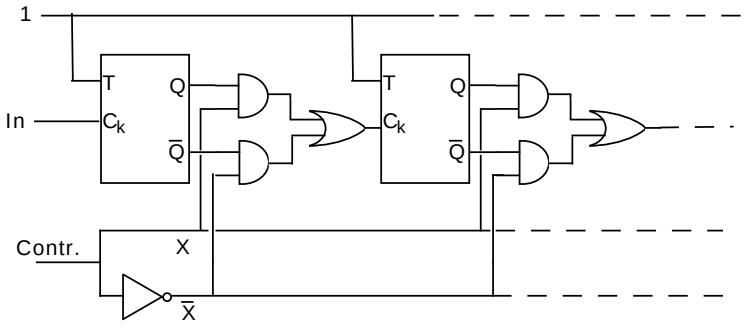

Figura 7.19: *Contatore asincrono avanti-indietro*

la frequenza massima di conteggio è limitata dal ritardo introdotto da ogni stadio, che deve dare il clock allo stadio successivo; ciò vuol dire che se arriva un impulso mentre il contatore non ha ancora commutato completamente, si perde il conteggio. Si può ovviare a questo inconveniente realizzando un contatore sincrono.

7.5.4 Contatore sincrono

Gli impulsi da contare vengono simultaneamente forniti a tutti i flip-flop, ma:

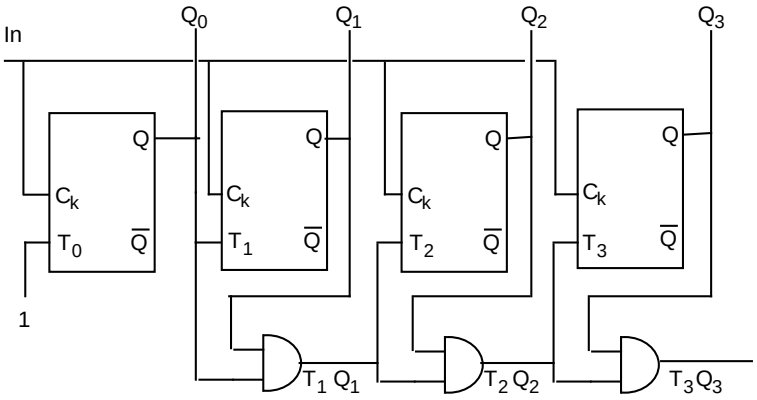

Figura 7.20: *Contatore sincrono a 4 bits con riporto in serie*

Q_0	commuta ad ogni impulso;	$T_0 = 1$
Q_1	solo se $Q_0 = 1$;	$T_1 = Q_0$
Q_2	solo se $Q_0 = Q_1 = 1$;	$T_2 = Q_0 Q_1$
Q_3	solo se $Q_0 = Q_1 = Q_2 = 1$;	$T_3 = Q_0 Q_1 Q_2$

Il tempo di commutazione non e' piu' dipendente dal numero di stadi, ed e' notevolmente piu' breve del caso precedente. Infatti esso e' ora dato da

$$T = T_F + (n-2)T_G$$

dove T_F e' il ritardo di propagazione di un flip-flop, mentre T_G e' il ritardo di propagazione di una porta AND.
E' possibile migliorare ulteriormente la velocita' utilizzando una tecnica di riporto in parallelo; questa richiede l'uso di AND a molti ingressi, chiaramente piu' scomodi, ma consente di arrivare ad un tempo complessivo di commutazione

$$T = T_F + T_G$$

Naturalmente esistono in commercio contatori sincroni up-down (ad esempio il 74LS191, contatore a 4 bits). Esistono anche contatori up-down (per esempio il 74LS193) con due ingressi di clock separati: uno per i segnali da contare in verso positivo, l'altro per i segnali da contare in verso negativo.

7.6 Conversione digitale-analogica

Il mondo dell'elettronica digitale non e' separato da quello dell'informazione analogica, ed esiste spesso la necessita' di passare dall'uno all'altro. La conversione digitale-analogica e' un'operazione attraverso la quale si costruisce una tensione V proporzionale ad un numero (binario), A, dato. Si ha quindi

$$V = K \sum_{0}^{n-1} 2^i a_i$$

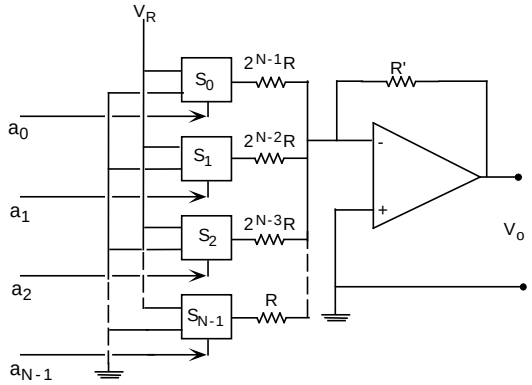

Figura 7.21: *Convertitore digitale-analogico a pesiera*

dove gli a_i sono i bits (0 o 1) che costituiscono il numero A (formato da n bits) e K e' una costante di proporzionalita' (dimensionalmente una tensione). Se il numero A e' uguale a 0 V e' anche uguale a 0, mentre se il numero e' formato da tutti 1 (cioe' ha la massima grandezza esprimibile con n bits) la tensione sara' data da

$$V_{max} = K(2^n - 1)$$

Quindi e' in generale piu' comodo scrivere

$$V = \frac{V_{max}}{2^n - 1} \sum_{0}^{n-1} 2^i a_i$$

I dispositivi che realizzano questa conversione si chiamano, in inglese, *Digital to Analog Converters* e percio' vengono brevemente chiamati DAC.

Convertitore D/A a pesiera.
Il primo e piu' semplice esempio di convertitore e' quello in Fig. 7.21. Si tratta in sostanza di un sommatore analogico, in cui le tensioni da sommare sono pesate secondo le potenze di due, grazie alle resistenze R, $2R$, $4R$,...,$2^{N-1}R$. I blocchi $S_0 \ldots S_{N-1}$ sono interruttori a 2 vie che connettono ogni ingresso a massa, ovvero a V_R, in relazione al valore 0 o 1 dei bits che costituiscono il numero da convertire; in questo modo si realizza proprio la funzione desiderata. Naturalmente l'uscita del convertitore non varia in modo continuo, bensi' a gradini, passando dalla tensione 0 alla tensione V_{max}, ed e' data da

$$V_0 = -\frac{R'}{R} V_R (a_{N-1} + \frac{a_{N-2}}{2} + \cdots + \frac{a_0}{2^{N-1}})$$

La precisione e la linearita' di questo dispositivo sono legate alla precisione dei rapporti di valore tra le resistenze; si comprende come questo renda notevolmente critico il dispositivo, specie se si vogliono convertire numeri con molti bits.

Convertitore D/A con rete a scala.
In questa soluzione si hanno solo resistenze di valore R e $2R$. La tensione d'uscita V_0 vale

$$V_0 = \frac{R_1 + R'}{R_1} V_i$$

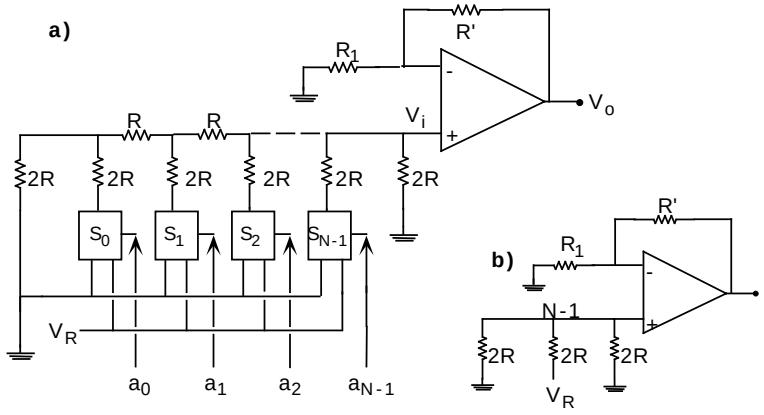

Figura 7.22: *a) Convertitore D/A con rete a scala; b) Circuito equivalente della rete di ingresso solamente il bit* $N-1$ *e' ad* 1

dove V_i è una funzione dei valori degli N bits del tipo richiesto. La verifica del comportamento di questo circuito puo' apparire molto complessa; conviene quindi prima esaminare alcuni casi semplici. Consideriamo anzitutto il caso in cui tutti i bits sono a zero: in questo caso V_i vale 0 e quindi anche l'uscita e' nulla. Se invece uno solo dei bits e' pari ad 1 il corrispondente nodo si porta alla tensione $V_R/3$. Infatti, in questo caso, qualunque sia la sua posizione, il nodo vede alla sua sinistra ed alla sua destra una resistenza $2R$ verso la massa (Fig. 7.22b). Ora, se si tratta del bit $N-1$ il nodo corrispondente coincide con l'ingresso dell'operazionale, quindi si ha

$$V_i = \frac{V_R}{3}$$

se invece si tratta del bit $N-2$, a causa della resistenza R presente tra il nodo $N-2$ e l'ingresso si ha

$$V_i = \frac{V_R}{3}\frac{1}{2}$$

cosi', nel caso del bit $N-3$ si ha

$$V = \frac{V_R}{3}\frac{1}{4}$$

Il comportamento non cambia se piu' bits sono pari ad uno, e si ha in generale

$$V = \frac{V_R}{3}(\frac{a_{N-1}}{2} + \frac{a_{N-2}}{4} + \cdots + \frac{a_0}{2^{N-1}})$$

Questo dispositivo e' chiaramente di piu' semplice realizzazione poiche' dobbiamo usare solo resistenze di valore R e $2R$.

Con opportuni miglioramenti e' possibile realizzare convertitori capaci di prestazioni piu' flessibili; si puo' infatti ottenere che l'uscita vari tra una tensione V_{min} ed una tensione V_{max} entrambe diverse da zero, oppure di segno diverso; la pendenza di conversione puo' essere positiva o negativa.
Alla stessa categoria appartiene il circuito mostrato in Fig. 7.23; in questo caso l'uscita del circuito e' una *corrente*, proporzionale all'ingresso digitale. Naturalmente si puo' ottenere

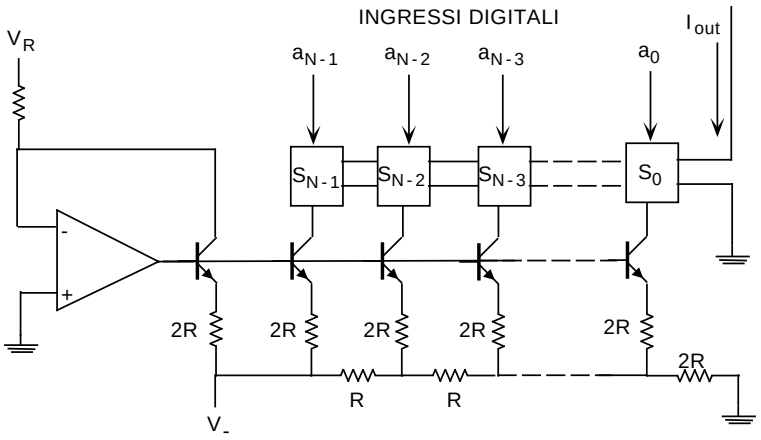

Figura 7.23: *a) DAC con uscita in corrente*

una tensione mettendo un opportuno resistore sull'uscita. Utilizzeremo in laboratorio proprio un convertitore di questo genere, ad 8 bits, per varie esperienze.

7.7 Conversione analogico-digitale

E' l'operazione contraria a quella vista in precedenza: l'uscita e' un numero binario di n bits, proporzionale ad una tensione d'ingresso. I dispositivi che realizzano questa funzione sono comunemente chiamati ADC *Analog to Digital Converter*, e possono essere realizzati in vari modi.

Convertitore A/D a conteggio. Consideriamo lo schema di Fig. 7.24a. La tensione da convertire, V_i e' presentata all'ingresso positivo del comparatore, ed il contatore binario e' inizialmente azzerato. L'uscita $V1$ del DAC e' nulla, quindi l'uscita del comparatore e' a livello logico 1. Si comincia ora ad inviare impulsi di clock: essi vanno ad incrementare il contatore, e quindi l'uscita del DAC sale formando dei gradini (Fig. 7.24b Quando $V1$

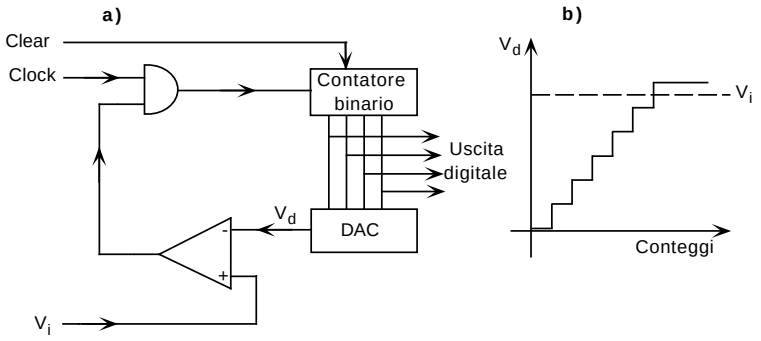

Figura 7.24: *a) Convertitore analogico-digitale a 4 bits; b) Andamento temporale della tensione* V_1

supera V_i l'uscita del comparatore scende a livello logico 0, l'ingresso di clock viene percio' inibito ed il contatore si ferma: il numero presentato all'uscita e' quindi proporzionale

al valore di V_i. Naturalmente il dispositivo funziona correttamente solo se la tensione da convertire e' compresa tra 0 e la tensione massima di uscita del DAC, V_{1max}, corrispondente ad avere tutti 1 nel contatore.
Il tempo di risposta di questo dispositivo, ovvero il tempo necessario affinche' l'uscita binaria arrivi al valore voluto e' chiaramente legto alla frequenza del clock, che deve essere adeguata ai tempi di risposta del contatore e del DAC. Esso inoltre non e' costante, ma cresce al crescere di V_i.

Tracking ADC. Una versione migliorata del convertitore a conteggio e' il *convertitore ad inseguimento*, o *tracking ADC* (Fig. 7.25). Non serve in questo caso un segnale di

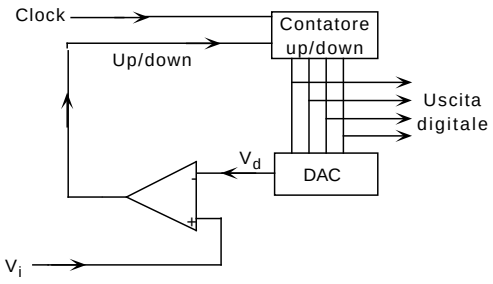

Figura 7.25: *Tracking ADC*

azzeramento. Infatti si supponga che inizialmente l'uscita abbia un valore qualunque e che, corrispondentemente, l'uscita del DAC sia inferiore alla tensione d'ingresso V_a: l'uscita del comparatore e' allora positiva e il contatore conta in avanti, finche' l'uscita del DAC supera la tensione d'ingresso. Allora il contatore inverte il verso del conteggio, facendo tornare l'uscita del DAC ad un valore inferiore a V_a. In sostanza l'uscita oscillera' avanti e indietro di 1 bit attorno al valore corretto.
Un dispositivo di questo genere e' quindi particolarmente adatto per convertire una tensione variabile nel tempo; il tempo di conversione e' piccolo se le variazioni del segnale analogico sono piccole.

ADC ad approssimazioni successive. Il tempo di conversione puo' essere in media molto ridotto usando una *strategia* di ricerca del valore corretto piu' intelligente. Sostituendo al contatore binario un programmatore (cioe' un registro piu' complesso) si puo realizzare una ricerca ad approssimazioni successive. Il programmatore pone inizialmente ad 1 il bit piu' significativo e a 0 tutti gli altri. Se la risultante tensione d'uscita del DAC e' maggiore della tensione d'ingresso, il bit viene rimesso a 0 e si prova con quello immediatamente meno significativo. In caso contrario il bit piu' significativo rimane ad 1, si mette ad 1 anche quello immediatamente meno significativo e si compara di nuovo. E' facile dimostrare che, per un sistema ad n bits il valore corretto e' trovato dopo n tentativi (cioe' impulsi di clock); invece in un ADC a conteggio sono necessari, nel caso peggiore 2^n impulsi.

ADC a doppia rampa. Consideriamo il circuito in Fig. 7.26a. V_i e' la tensione da convertire (positiva) e V_R e' una tensione negativa, con $|V_R| > V_i$. All'istante t_1 si connette

l'ingresso dell'integratore a V_i, per un tempo T_1 costante, pari a $n_1 T$, dove T e' il periodo del clock. All'istante t_2 si connette l'integratore a V_R e simultaneamente si fa partire il clock. La tensione V_0 scende (Fig. 7.26b) e all'istante t_3 torna a zero: l'uscita del comparatore blocca il clock e quindi il conteggio del contatore. Ora possiamo scrivere che

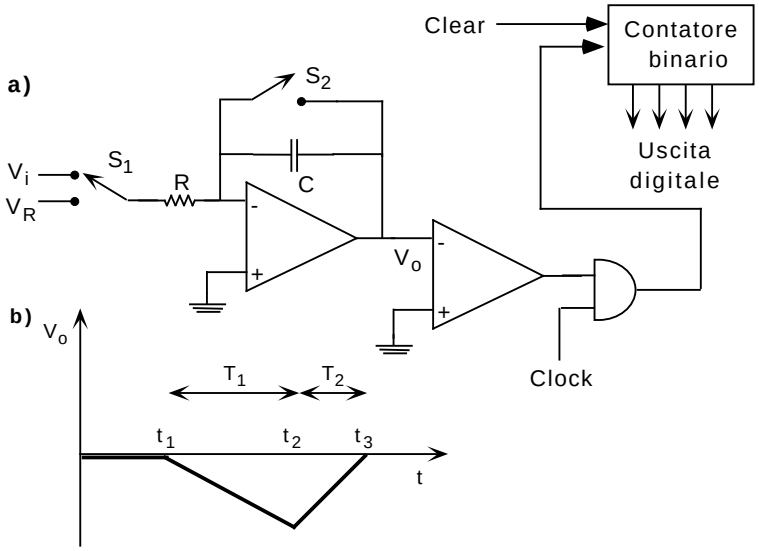

Figura 7.26: *a) Convertitore A/D a doppia rampa; b) Andamento temporale di* V_0

$$V(t_3) = -\frac{1}{RC}\int_{t_1}^{t_2} V_i dt - \frac{1}{RC}\int_{t_2}^{t_3} V_R dt = 0$$

$$V_i(t_2 - t_1) + V_R(t_3 - t_2) = 0$$

ovvero

$$V_i T_1 + V_R T_2 = 0$$

dove

$$T_2 = n_2 T$$

$$T_1 = n_1 T$$

quindi

$$V_i = \frac{n_2}{n_1}|V_R|$$

e finalmente

$$n_2 = n_1 \frac{V_i}{|V_R|}$$

cioè n_2 è proporzionale a V_i essendo n_1 e V_R delle costanti note. La sensibilita' di questo dispositivo e' legata alla frequenza del clock (limitata dalla velocita' del contatore), mentre la linearita' e' chiaramente legata alla linearita' dell'integratore; da questo punto di vista e' importante il fatto che la risposta non dipenda dai valori di R e C: possiamo quindi sceglierli in modo da sfruttare la parte lineare della rampa d'integrazione.

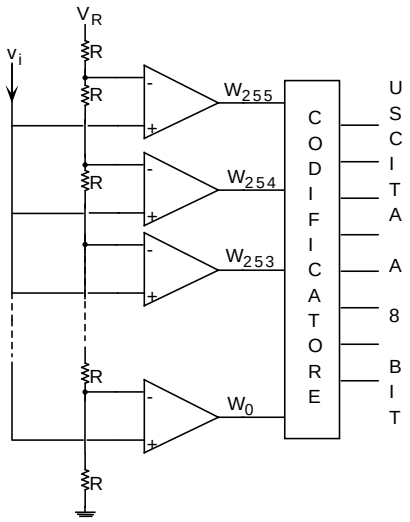

Figura 7.27: *a) Flash ADC*

Flash ADC. Questo e' il dispositivo concettualmente piu' semplice, ma anche il piu' difficile e costoso da realizzare (Fig. 7.27): il segnale da convertire viene inviato simultaneamente ad n comparatori con soglie equispaziate. Chiaramente la soglia dell'i-esimo comparatore e' data da:

$$V_i = V\frac{i}{n+1}$$

quindi tutti i comparatori la cui soglia e' inferiore a V_{in} scattano, mentre gli altri danno risposta zero. Tutte le uscite dei comparatori vengono poi inviate ad un codificatore che trasforma l'informazione in un numero binario. Il numero complessivo di comparatori e' enorme: per un convertitore ad 8 bits ne occorrono 255! Questa spiega perche' questi dispositivi sono estremamente costosi e solo i recenti progressi nel campo dell'integrazione su larga scala ne hanno reso possibile la realizzazione a costi accessibili. Questi dispositivi sono ovviamente i piu' veloci: il tempo di conversione e' infatti dovuto essenzialmente al tempo di risposta del codificatore. E' inutile dire che la precisione del dispositivo e' legata alla precisione delle n soglie; inoltre un problema e' anche rappresentato dalla enorme capacita' di ingresso che si forma, quando n e' molto grande. Di fatto i dispositivi in commercio non vanno oltre i 10 o 11 bits.

Capitolo 8

Il microprocessore Z80

8.1 Introduzione

Dal 1946, anno di costruzione del primo computer (il famoso ENIAC) ad oggi, si e' progressivamente consolidata un'architettura dei sistemi di calcolo basata su una unita' centrale, detta CPU (Central Processing Unit), che lavora in associazione con una memoria centrale, e con una serie di periferiche, come dischi, nastri, terminali videografici, ecc. La CPU e' in grado di eseguire una sequenza di istruzioni (programma) che siano state precedentemente immesse nella memoria; inoltre coordina il funzionamento delle periferiche, acquisendo dati dalle unita' d'ingresso e fornendone altri alle unita' di uscita. La stessa memoria utilizzata per contenere i programmi serve anche per contenere dati: essa e' quindi sostanzialmente una unita' periferica allo stesso livello delle altre.

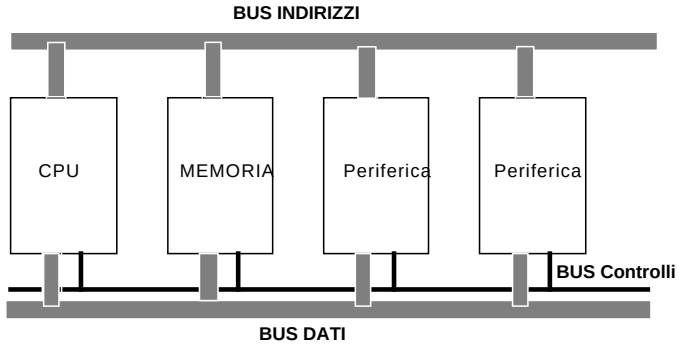

Figura 8.1: *Architettura di un computer*

Le comunicazioni tra la CPU e le periferiche avvengono attraverso un insieme di linee, che prendono il nome di *bus*; in realta' sono necessari piu' bus per il funzionamento del sistema (vedi Fig. 8.1): infatti e' necessario

1. definire da quale indirizzo deve essere prelevata un'informazione e a quale indirizzo deve essere destinata (Address Bus);
2. trasferire i dati dalla CPU alla periferica o viceversa (Data Bus);
3. inviare e ricevere segnali di controllo (Control Bus)

Sappiamo che l'unita' elementare di informazione e' il bit; e' chiaro pero' che qualunque operazione realistica richiede la capacita' di operare su numeri a molti bit. Una caratteristica importante di un computer e' quindi data dall'ampiezza dei numeri che esso e' in grado di gestire e manipolare; questa dimensione si chiama *parola*: quindi un computer con parole di 16 bit e' in grado di operare su numeri di questa ampiezza. Tipici valori sono 8, 16, 32 e oggi anche 64 bit. Naturalmente la grandezza della parola determina anche il numero di linee che formano il bus dei dati (anche se non necessariamente devono essere uguali).
Un'altra caratteristica importante e' il numero di linee di indirizzo che costituiscono il sistema. Infatti ogni periferica corrisponde ad uno o piu' indirizzi; nel caso delle memorie, ad esempio, ogni locazione (intesa come gruppo di bit) deve avere un indirizzo univoco; quindi il massimo numero di locazioni indirizzabili e' legato al numero di bit, e quindi di linee, che definiscono un indirizzo. Se ad esempio si hanno a disposizione 16 linee, si possono definire $2^{16} = 65536$ indirizzi diversi. Nella terminologia informatica si ha $1024 = 1\ k$, pertanto 2^{16} corrisponde a $64\ K$. In genere le memorie commerciali consentono di indirizzare gruppi di 8 bit, che cotituiscono 1 byte; percio' l'ampiezza delle memorie e' normalmente espressa in bytes
Il bus di controllo, il cui scopo comprenderemo meglio in seguito, richiede un numero abbastanza basso di linee; esso serve a definire il tipo di scambio che si vuole effettuare (lettura di un dato dalla memoria o scrittura di un dato sulla memoria), oppure a consentire alle periferiche di attirare l'attenzione della CPU. Infatti il controllo dei bus e', normalmente, compito della CPU, che coordina e gestisce ogni operazione; tuttavia le periferiche devono a volte potere intervenire in modo attivo nelle transazioni (si pensi ad esempio alla situazione in cui l'utente preme un tasto della tastiera per immettere un dato o un comando nel sistema).

In passato una CPU era costituita da un enorme volume di circuiti elettronici; oggi, i progressi fatti nel campo dell'integrazione hanno reso possibile racchiudere tutte le funzioni in un unico chip di silicio, in cui sono racchiuse migliaia o decine di migliaia di porte logiche elementari. Questi integrati prendono il nome di microprocessori; nella tabella 8.1 sono mostrate le principali caratteristiche di alcuni noti microprocessori.

Costruttore	Sigla	Parola (bit)	Indirizzi (bit)
Intel	8080	8	16
Motorola	6800	8	16
Zilog	Z80	8	16
Intel	8086	16	16
Motorola	68000	16	16
Intel	80386	32	32
Motorola	68020	32	32

Tabella 8.1: *Alcuni microprocessori in commercio*

E' importante notare che queste macchine costituiscono dei sistemi *sincroni*, cioe' svolgono le loro funzioni attraverso una sequenza di operazioni temporizzate da un segnale di clock (che spesso deve essere fornito dall'esterno); la massima frequenza di clock cui il sistema puo' operare e' chiaramente legata alla velocita' con cui i vari circuiti possono svolgere i loro compiti e costituisce quindi un fattore di merito del sistema. Noi studieremo

in dettaglio un particolare microprocessore, lo Zilog Z80; certamente esso e' superato in prestazioni da realizzazioni piu' recenti, tuttavia mantiene ancora oggi tutta la sua validita' come strumento didattico, sia per consentire agli studenti di apprendere attraverso un caso particolare una serie di nozioni valide in generale per tutti i microprocessori, ma soprattutto perche' consente di realizzare semplici ed istruttive applicazioni pratiche, che i molto piu' sofisticati microprocessori delle ultime generazioni rendono assai piu' problematiche.

8.1.1 Il sistema esadecimale

Nel Capitolo precedente abbiamo visto il sistema di numerazione binario e abbiamo capito che esso e' una chiave importante per tradurre circuitalmente problemi aritmetici (e logici). Cio' e' vero naturalmente anche per i microprocessori; tuttavia la notazione binaria e' estremamente pesante e scomoda da ricordare, non appena si ha a che fare con numeri di 8 o addirittura 16 bit. Si possono pero' esprimere numeri binari in modo piu' abbreviato usando la notazione esadecimale. Consideriamo infatti il sistema a base 16: esso ha bisogno di 15 simboli, quindi si utilizzano i numeri da 0 a 9 e poi le lettere A, B, C, D, E, F. Si ha quindi la numerazione come nella Tabella 8.2 Poiche' $16 = 2^4$ si ha una relazione

Num. decimale	Num. binaria	Num. esadecimale
0	0	0
1	1	1
2	10	2
3	11	3
4	100	4
5	101	5
6	110	6
7	111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F
16	10000	10
17	10001	11
...	...	...
255	11111111	FF
256	100000000	100
...	...	

Tabella 8.2: *Numerazione esadecimale*

molto semplice tra numerazione esadecimale e numerazione binaria: 1 cifra esadecimale corrisponde a 4 cifre binarie. Tradurre un numero binario in notazione esadecimale e' allora semplicissimo, perche' puo' essere fatto, partendo dalla cifra meno significativa, per gruppi di 4 bit per volta. Si abbia ad esempio il numero binario 10 1011 1100; esso si traduce immediatamente in $2\,B\,C$. Viceversa, se abbiamo il numero esadecimale $1AEF$, esso corrisponde a 1 1010 1110 1111.
La notazione esadecimale viene usata estensivamente nel campo informatico, proprio per la sua semplice connessione con la numerazione binaria, ed e' quindi bene che gli studenti

acquistino un minimo di dimestichezza con essa. Abbiamo per esempio visto che avendo a disposizione 16 linee di indirizzi, e' possibile indirizzare 65536 locazioni diverse; in notazione esadecimale e' immediato dire che queste locazioni vanno dall'indirizzo 0000 (corrispondente al numero binario 0000 0000 0000 0000) all'indirizzo $FFFF$ (corrispondente a 1111 1111 1111 1111).

8.1.2 Logica tri-state

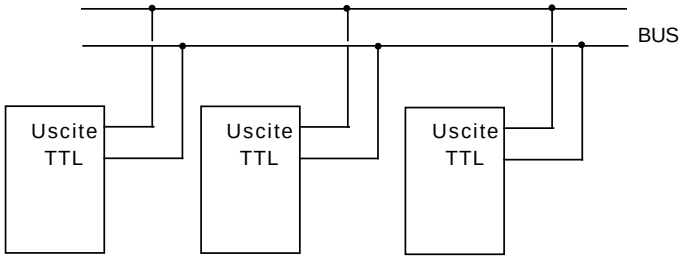

Figura 8.2: *3 sistemi logici connessi allo stesso bus*

Lo schema concettuale mostrato in Fig. 8.1 e' certamente molto funzionale ma ci pone dei problemi elettronici che dobbiamo comprendere e risolvere immediatamente. Infatti l'idea di un bus, cioe' un sistema di linee attraverso cui l'informazione si trasferisce da un blocco ad un altro (per esempio dalla memoria alla CPU, o viceversa) non e' realizzabile utilizzando i normali circuiti logici TTL che conosciamo. Prendiamo infatti la situazione schematizzata in Fig. 8.2; in cui 3 sistemi logici hanno le uscite connesse al bus (per semplicita' abbiamo disegnato solo 2 linee del bus). Se ricordiamo il funzionamento delle uscite TTL comprendiamo immediatamente che questo schema non puo' funzionare: infatti una eventuale uscita bassa forza a $0\ V$ la tensione della corrispondente linea, indipendentemente da cio' che fanno gli altri blocchi logici. E' chiaro quindi la connessione di piu' circuiti logici ad un bus non puo' essere fatta in questo modo. Uno dei modi comunemente usati per risolvere questo problema consiste nell'introdurre nei circuiti logici un terzo stato, il cosidetto stato di alta impedenza. Allora l'uscita del circuito puo' essere in 3 stati: 0 logico, 1 logico e alta impedenza, dove quest'ultimo significa che l'uscita presenta una alta impedenza verso la massa, cioe' e' come se fosse sconnessa dal circuito da cui proviene. Questo risolve il problema del bus: occorre fare in modo che uno e uno solo dei dispositivi connessi al bus presenti un'uscita logicamente valida; tutti gli altri devono essere nello stato di alta impedenza, H. Ogni circuito logico avra' percio' un ulteriore ingresso di abilitazione (Enable) che consente di introdurre questo terzo stato di uscita. Ad esempio, la tavola della verita' di un flip-flop D tipo tri-state sara'

Enable	D_n	Q_{n+1}
1	0	0
1	1	1
0	x	H

Cioe', se il segnale Enable e' basso, l'uscita e' su alta impedenza, qualunque sia l'ingresso.

8.2 Struttura dello Z80

Vediamo ora in dettaglio la struttura del microprocessore Z80 [1], per cominciare a comprenderne il funzionamento. Questo microprocessore si presenta come un contenitore Dual In Line da 40 piedini (Fig. 8.3). Di questi, 8 piedini corrispondono agli 8 bit del

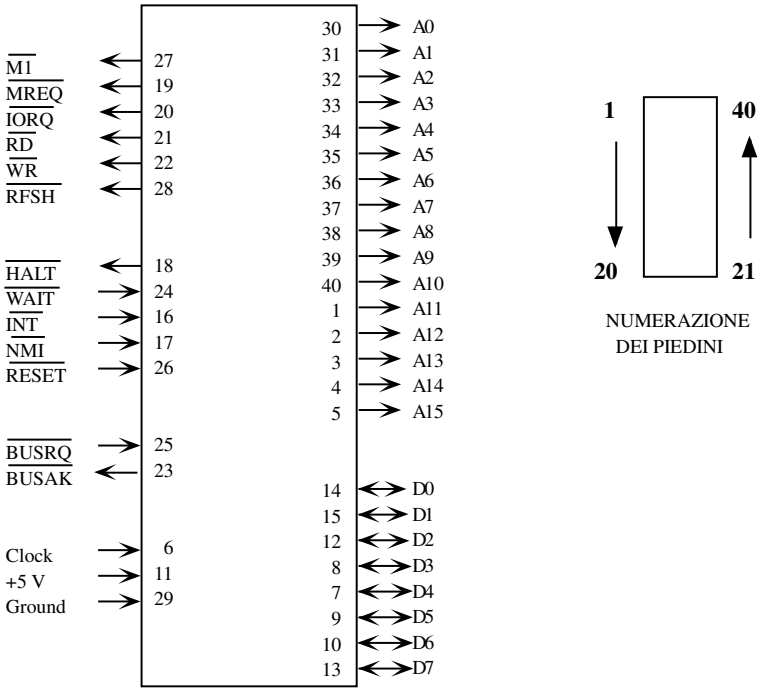

Figura 8.3: *I piedini dell'integrato Z80: le frecce indicano il verso in cui fluisce l'informazione.*

dato ($D_0 \ldots D_7$), 16 ai 16 bit degli indirizzi ($A_0 \ldots A_{15}$). Due piedini servono per l'alimentazione ($+5\ V$ e GND); uno e' l'ingresso per il clock, ϕ. I restanti 13 piedini sono utilizzati per i segnali di controllo. Ogni segnale e' indicato con una sigla mnemonica che ricorda la funzione svolta; poiche' e' il livello basso (0 logico) quello che attiva la funzione, le sigle sono soprasegnate. Alcuni segnali sono prodotti dalla CPU, e percio' in figura sono caratterizzati dalla freccia uscente; altri, invece, sono ricevuti dalla CPU e devono essere forniti dai dispositivi esterni. Vediamo ora in dettaglio questi segnali:

$\overline{BUSRQ}$ Richiesta del bus. Viene usato da un dispositivo esterno il controllo dei bus. Quando $\overline{BUSRQ}$ viene attivato, la CPU pone nello stato di alta impedenza le sue connessioni col bus dei dati, col bus degli indirizzi e con le uscite di controllo, non appena termina il ciclo macchina in corso.

[1] Lo Z80 è nato nel luglio del 1976 per opera di Federico Faggin che, lasciata la Intel dopo aver lavorato sull'8080, aveva fondato la Zilog. Era progettato per offrire compatibilità binaria con l'Intel 8080 in modo che il codice 8080 (in particolare il sistema operativo CP/M) potesse essere eseguito sullo Z80 senza modifiche.
Zilog ha concesso in licenza il core dello Z80 senza royalty a tutte le aziende che volessero costruire il chip. Questo ha fatto si che il prodotto guadagnasse consensi nel mercato mondiale, in quanto aziende di gran lunga più grandi iniziarono a produrre lo Z80.

$\overline{BUSAK}$ Riconoscimento del bus. Questo segnale viene attivato quando, a seguito di un $\overline{BUSRQ}$, i bus sono stati posti in uno stato di alta impedenza.

$\overline{RESET}$ Azzeramento. Questo segnale rimette a zero il contenuto del registro PC (Program Counter) ed esegue altre azioni di inizializzazione generale della CPU. A seguito di questo segnale la CPU comincia ad eseguire il programma, partendo dall'istruzione memorizzata nella locazione di memoria 0000.

$\overline{HALT}$ Stato di alt. Indica che la CPU ha eseguito una istruzione di HALT ed e' in attesa di comando per riprendere l'esecuzione del programma. Finche' perdura questo stato la CPU esegue istruzioni fittizie (NOP) per consentire comunque l'esecuzione di attivita' di refresh

$\overline{MREQ}$ Richiesta di memoria. Il segnale indica che il bus degli indirizzi contiene un indirizzo valido per un'operazione di lettura o di scrittura in memoria

$\overline{M1}$ Ciclo macchina 1. Il segnale indica che e' in atto un ciclo di prelievo (fetch) del codice operativo dalla memoria. L'inizio dell'esecuzione di ogni istruzione e' quindi caratterizzato da questo segnale. La presenza simultanea di $\overline{M1}$ e $\overline{IORQ}$ sta ad indicare un ciclo di riconoscimento di una interruzione

$\overline{IORQ}$ Richiesta di ingresso/uscita. Indica che il bus degli indirizzi contiene, negli 8 bit meno significativi, un indirizzo di un dispositivo di I/O valido per una oprazione di lettura o scrittura

$\overline{WAIT}$ Attesa. Indica alla CPU che la memoria, o altri dispositivi di I/O indirizzati non sono pronti per un trasferimento dati. Finche' questo segnale e' attivo la CPU continua l'attuazione di stati di attesa.

$\overline{RD}$ Lettura. La CPU vuole leggere dalla memoria o dal dispositivo di I/O indirizzato. La memoria o altro dispositivo usa questo segnale per porre sul bus dei dati l'informazione richiesta

$\overline{WR}$ Scrittura. Indica che il bus dei dati contiene dati validi da memorizzare nella memoria o nel dispositivo di I/O

$\overline{RFSH}$ Refresh. Indica che i 7 bit meno significativi del bus degli indirizzi contengono un indirizzo valido per una operazione di rinfresco della memoria e che il segnale $\overline{MREQ}$ attivo in quel momento e' usato per effettuare una lettura

$\overline{INT}$ Richiesta di interruzione (Interrupt). Deve essere generato da dispositivi di I/O; l'interruzione viene accettata dalla CPU, in particolari condizioni, alla fine dell'istruzione in corso.

$\overline{NMI}$ Richiesta di interruzione non mascherabile. Questo segnale e' attivo sul fronte di discesa. Viene sempre accettato, senza condizioni, alla fine dell'istruzione in corso. La CPU riprende l'esecuzione a partire dalla locazione di memoria 0066.

Dovremo man mano approfondire la descrizione precedente, che puo' aver disorientato il lettore, anche per l'introduzione di una serie di termini oscuri, come refresh, ciclo macchina,

ecc. Tuttavia alcuni punti dovrebbero cominciare ad essere chiari. Si capisce che la memoria deve contenere una serie di istruzioni, per esempio a partire dalla locazione 0000; dando al microprocessore un segnale di $\overline{RESET}$ esso iniziera' a prelevarle ed a eseguirle in sequenza. Una istruzione implica tipicamente prelevare dei dati dalla memoria, eseguire un'operazione su di essi e riscrivere in una certa locazione di memoria il risultato, che eventualmente costituisce l'operando di successive manipolazioni. Non poniamoci per ora il problema di come si possano introdurre nella memoria i codici delle istruzioni da eseguire; lo vedremo in seguito. Cominciamo invece ad esaminare la struttura funzionale del microprocessore (Fig. 8.4). Il cuore della CPU e' costituito da una serie di registri, da una unita' aritmetica

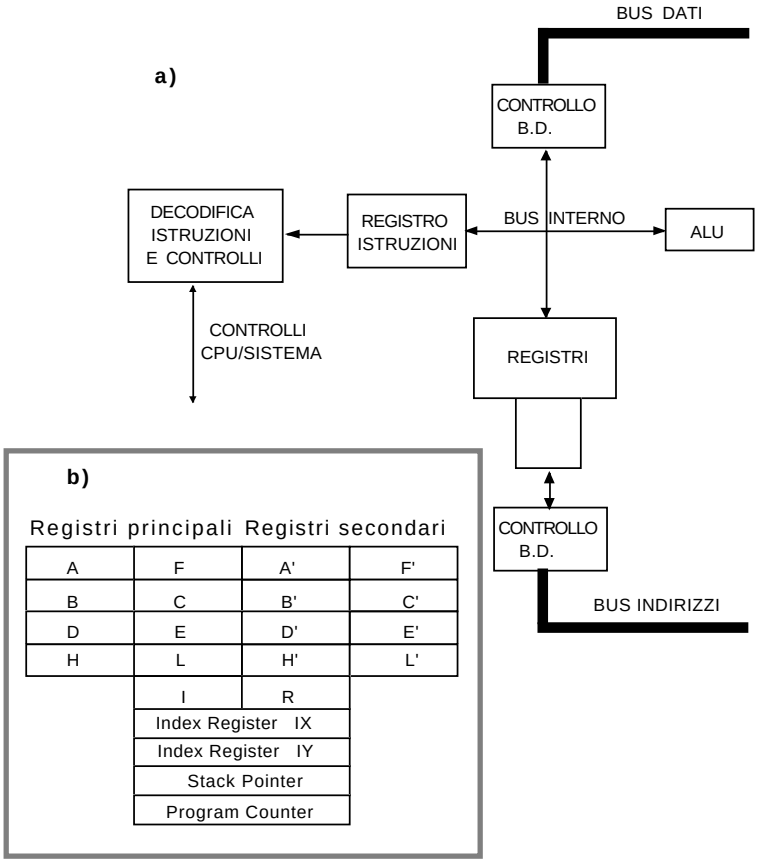

Figura 8.4: *a) Architettura interna dello Z80; b) Registri*

e logica (ALU) e da alcuni blocchi di controllo. Un bus interno (non accessibile cioe' da fuori) consente il trasferimento locale delle informazioni. Come si vede l'interazione verso l'esterno avviene attraverso i controllori dei vari bus: indirizzi, dati, e segnali di controllo. Il blocco dei registri contiene 18 registri a 8 bit e 4 registri a 16 bit.

I registri di uso generale sono A,F,B,C,D,E,H,L; accanto a questi, detti principali, ve ne sono altri 8, indicati con A',.....L', detti secondari, di cui per ora non ci occuperemo; la loro funzione verra' spiegata molto piu' avanti. I registri di uso generale, tranne il registro F, sono destinati come appoggio per gli operandi delle istruzioni ed i risultati dei calcoli eseguiti. Il piu' frequentemente usato e' il registro A, che prende, per ragioni storiche, il

nome di accumulatore. Spesso i registri vengono usati in coppie, per contenere informazioni a 16 bit (in particolare si usano le coppie BC, DE, HL).
Il registro F ha invece una funzione completamente diversa: ciascuno dei singoli bit che lo compongono da infatti informazioni su particolari caratteristiche dell'ultima operazione eseguita dal microprocessore. In particolare si ha:

Bit number	Simbolo	Significato
0	C	Carry
1	N	Add/Subtract
2	P/V	Parity/Overflow
3	X	Non usato
4	H	Half carry
5	X	Non usato
6	Z	Zero
7	S	Sign

Ad esempio, il bit 6 indica se l'ultima operazione eseguita ha avuto come risultato zero (in tal caso il bit viene posto ad 1); il bit 0 viene posto ad 1 se l'ultima operazione ha dato luogo ad un riporto. Tutte queste informazioni possono, come vedremo, essere utili all'utente per organizzare un programma.
Vediamo ora la funzione dei registri di uso speciale:

Program Counter, PC Questo registro da 16 bit contiene ad ogni istante l'indirizzo della locazione di memoria in cui sta l'istruzione da eseguire. Normalmente il microprocessore esegue le istruzioni in sequenza ed il registro viene quindi incrementato ogni volta di 1. A volte pero' il programma contiene istruzioni di salto per cui si richiede che l'istruzione successiva venga attinta da un indirizzo completamente diverso: in tal caso il nuovo indirizzo viene trascritto sul PC. Si capisce quindi che il bus degli indirizzi viene predisposto sulla base del contenuto del Program Counter.

Stack Pointer, SP La programmazione dello Z80 ammette l'uso di subroutines (cosi' come le conosciamo nei linguaggi evoluti come il FORTRAN). In tal caso, un salto dal programma principale ad una subroutine richiede che venga tenuta memoria del punto da cui si e' effettutato il salto, per potervi tornare una volta esaurita l'esecuzione della subroutine. Inoltre una subroutine puo' chiamarne un'altra, la quale a sua volta puo' chiamarne un'altra ancora e cosi' via; si devono quindi memorizzare tanti indirizzi concatenati, per ritrovare la strada inversa. Questo viene fatto, come vedremo meglio in seguito, con la tecnica dello stack (catasta), in cui tutti questi indirizzi vengono memorizzati in una zona della memoria: lo stack pointer contiene un indirizzo che ci consente di ritrovare all'indietro la strada percorsa in tutte le chiamate a subroutines concatenate.

Registri indice, IX, IY Sono impiegati quando si utilizza una particolare tecnica di indirizzamento, detta indicizzata.

Interrupt vector, I In questo registro viene immagazzinato un byte che contiene informazioni necessarie per gestire alcuni tipi di interruzione.

Memory refresh, R Lo Z80 incorpora la funzione di refresh automatico. E' un'operazione necessaria per il funzionamento delle memorie CMOS (memorie dinamiche); infatti in questo tipo di circuiti l'informazione e' memorizzata sotto forma di carica elettrica immagazzinata in una piccola capacita'. In teoria questa capacita' e' isolata, per cui dovrebbe mantenere indefinitamente il suo stato di carica, ma in realta' cio' non puo' avvenire, per cui essa tende a scaricarsi. E' necessario allora rinfrescare continuamente (con periodicita' dell'ordine delle decine di millisecondi) tutte le celle e questo viene semplicemente fatto rileggendone continuamente il contenuto. Il registro R aiuta a fare questa operazione, sfruttando intervalli di tempo in cui il microprocessore non deve accedere alla memoria per fare reali operazioni. Noi non useremo memorie di questo tipo, quindi non sfrutteremo questa opzione; ne vedremo pero' gli effetti quando studieremo in dettaglio la temporizzazione del microprocessore.

8.3 Programmazione dello Z80

Possiamo ora, finalmente, capire un po' meglio il funzionamento del sistema attraverso alcuni semplici esempi. Come abbiamo piu' volte detto il microprocessore esegue un compito definito leggendo dalla memoria una serie di istruzioni, che noi avremo preventivamente immagazzinato, che costituiscono il programma. Abbiamo visto che, fornendo un comando di $\overline{RESET}$, lo Z80 comincia ad eseguire il programma partendo dalla locazione 0000: possiamo quindi immaginare di scrivere il nostro programma partendo da quella locazione. Naturalmente le istruzioni devono essere date in linguaggio macchina, cioe' in codice binario, certamente assai difficile da apprendere e ricordare: e' quindi utile al programmatore fare ricorso ad una forma mnemonica per indicare ogni istruzione, che ci aiutera' nello scrivere un programma, ma anche nel capire cosa fa un programma, nel momento in cui andiamo a rileggerlo.
Non tutte le istruzioni dello Z80 si esauriscono in un byte; ci sono invece istruzioni piu' complesse che richiedono 2,3 o anche 4 bytes. Di conseguenza l'istruzione deve essere memorizzata in piu' locazioni di memoria consecutive. Ma vediamo ora un primo esempio, in cui useremo la notazione esadecimale, sia per scrivere gli indirizzi delle locazioni di memoria, sia le istruzioni:

Indirizzo	Istruzione	Mnemonico	Commento
0000	00	NOP	Non fa niente
0001	C3	JP 0000	Salta alla locazione 0000
0002	00		
0003	00		

Questo semplice programma di due istruzioni esegue un loop infinito: l'istruzione NOP non fa niente (corrisponde al CONTINUE del FORTRAN), l'istruzione JP 0000 salta alla

locazione 0000, cioe' ritorna al punto di partenza. Come si vede questa istruzione richiede 3 bytes e di conseguenza impegna 3 locazioni di memoria.
Vediamo ora un esempio piu' articolato:

Indirizzo	Istruzione	Mnemonico	Commento
0000	06	LDB, 64	Carica nel registro B il numero
0001	64		esadecimale 64
0002	05	DEC B	Decrementa il registro B di 1
0003	C2	JPNZ, 0002	Salta alla locazione 002 se
0004	02		l'ultima operazione non ha dato
0005	00		risultato zero
0006	76	HALT	Si ferma

Questo programma esegue un loop 64 volte (esadecimale) e poi si ferma. In esso abbiamo usato istruzioni lunghe 1,2 e 3 bytes. Abbiamo caricato in B un numero e successivamente lo abbiamo decrementato a passi di 1, eseguendo ogni volta un test per verificare il contenuto del registro. Questo tipo di istruzioni usa proprio il bit 6 del registro F, che viene posto ad 1 quando l'ultima operazione fatta (in questo caso DEC B) fornisce un risultato zero. Si noti poi il modo con cui abbiamo dato l'indirizzo nell'istruzione JPNZ, 0002: prima il byte meno significativo (02) e poi quello piu' significativo (00). Questa convenzione e' del tutto generale e deve essere seguita in tutte le istruzioni che contengono un indirizzo.
Il formato che stiamo usando per scrivere il nostro programma e' quindi chiaro: occorre in generale che il programmatore tenga conto delle locazioni di memoria in cui sono contenute le istruzioni, perche' le istruzioni contengono indirizzi ben precisi. In questa situazione il programma e' detto *non rilocabile* proprio' perche' esso funziona solo se e' caricato partendo da una precisa locazione di memoria. Possiamo anche apprezzare l'utilita' del codice mnemonico (che, ripetiamo, non serve al microprocessore, ma solo a noi). Da questo punto di vista e' utile anche assegnare dei nomi arbitrari, a scopo mnemonico, a locazioni di memoria. Chiameremo *labels* questi nomi; l'esempio precedente potrebbe allora essere riscritto:

Indirizzo	Label	Istruzione	Mnemonico	Commento
0000		06	LDB, 64	Carica nel reg. B
0001		64		il numero esadec. 64
0002	loop	05	DEC B	Decrementa il reg. B di 1
0003		C2	JPNZ, loop	Salta alla loc. 002 se
0004		02		l'ultima operazione non ha
0005		00		dato risultato zero
0006		76	HALT	Si ferma

Alla locazione 0002 (il bersaglio del salto) abbiamo dato il nome loop: questo rende piu' facile rileggere il programma e capirne il funzionamento. E' un artificio utile, che useremo spesso nei prossimi esempi.
Il lettore si porra' ora due domande: e' possibile eseguire un programma che non e' memorizzato a partire dalla locazione 0000? E' possibile scrivere un programma in modo *rilocabile*, cioe' in modo che esso funzioni qualunque sia la zona di memoria in cui esso e' stato immagazzinato? La risposta e' si ad entrambe le domande e possiamo vederlo con un ulteriore esempio. Questa volta scriviamo un programma che esegue la somma di due numeri ad 8 bit, e supponiamo di averlo memorizzato a partire dalla locazione 0100; i due addendi sono contenuti nelle locazioni di memoria 0200 e 0201, mentre il risultato viene salvato nella locazione 0202:

Indirizzo	Label	Istruzione	Mnemonico	Commento
0100		3A	LDA,(0200)	Carica in A
0101		00		il primo addendo
0102		02		
0103		2A	LDHL, 0201	Carica in HL
0104		01		l'indirizzo del
0105		02		secondo addendo
0106		86	ADDA,(HL)	Somma
0107		32	LD (0202),A	Scrive in memoria
0108		02		il risultato
0109		02		
010A		76	HALT	

Qui abbiamo usato varie tecniche di indirizzamento: abbiamo prelevato il primo addendo fornendo direttamente l'indirizzo; invece per il secondo addendo abbiamo scritto nella coppia di registri HL l'indirizzo e poi esguito una somma tra il contenuto di A e il contenuto della locazione di memoria il cui inirizzo e' fornito da HL: questo e' il cosidetto indirizzamento indiretto.
Come facciamo a mettere in esecuzione questo programma? Scriveremo semplicemente:

Indirizzo	Label	Istruzione	Mnemonico	Commento
0000		C3	JP, 0100	Salta a 0100
0001		00		
0002		01		

Possiamo quindi memorizzare in memoria molti programmi e scegliere quale eseguire modificando semplicemente l'indirizzo posto nelle locazioni 0001 e 0002. Partendo da questo semplice concetto potremmo poi sviluppare tecniche assai piu' sofisticate per selezionare il

programma da eseguire. Notiamo ora che l'esempio appena fatto non contiene nessun riferimento assoluto ad indirizzi di memoria (se non quelli degli operandi e del risultato): quel programma quindi funziona ovunque sia caricato, ed e' percio' un esempio di programma *rilocabile*. Tutti i compilatori evoluti (per esempio il FORTRAN) traducono il programma scritto dall'utente in un codice rilocabile; questo aspetto e' chiaramente essenziale per consentire al programma di funzionare in qualunque zona della memoria venga caricato.

8.3.1 Temporizzazione dello Z80

La sequenza delle operazioni che un microprocessore effettua e' rigorosamente scandita dagli impulsi di clock; nello Z80 il clock deve essere fornito dall'esterno, attraverso il piedino 6, e puo' avere una frequenza qualunque da praticamente zero fino ad un massimo di 4 Mhz. Questa caratteristica e' particolarmente utile per fini didattici perche' potremo, fornendo un clock a bassa frequenza, seguire la temporizzazione dei processi in corso con un normale oscillografo.
E' chiaro che l'esecuzione di un'istruzione, anche semplice, si compone da un punto di vista elettronico di piu' fasi; pertanto un ciclo di clock non e' mai sufficiente a completare un'istruzione. La durata minima e' di 4 cicli di clock; infatti durante i primi due cicli lo Z80 acquisisce il codice dell'istruzione, mentre durante i due successivi decodifica tale istruzione (e, simultaneamente, opera un refresh della memoria). E' evidente che questo e' il caso piu' semplice possibile, non esistendo istruzioni che facciano meno di questo. Un esempio e' l'istruzione DEC A, in cui si richiede di decrementare il contenuto dell'accumulatore. Se l'istruzione non prevede ulteriori accessi alla memoria, essa termina e si passa all'istruzione successiva, incrementando di 1 il Program Counter; altrimenti ogni ulteriore accesso in memoria richiede 3 cicli di clock aggiuntivi.
Possiamo quindi parlare di:

Ciclo di clock;

Ciclo di macchina;

Ciclo di istruzione;

Un ciclo di istruzione si compone quindi di uno o piu' cicli macchina, mentre un ciclo macchina si compone di 4 o di 3 cicli di clock, a seconda che sia o meno la prima fase dell'istruzione.
Vediamo ora la scala dei tempi in cui avvengono le operazioni descritte (Fig. 8.5).
Ogni istruzione inizia con il trasferimento del contenuto del Program Counter sul bus degli indirizzi: questo avviene all'inizio di T_1, primo periodo di clock. Nella seconda meta' di T_1 i segnali $\overline{MREQ}$ e $\overline{RD}$ divengono attivi (cioe' vanno a 0 logico). Nella restante parte di T_1 e nel periodo seguente T_2 la memoria scrive sul bus dei dati il contenuto della locazione indirizzata; questo dato viene letto dalla CPU il corrispondenza al fronte di salita di T_3. E' possibile che la memoria non sia in grado, in termini di velocita', di soddisfare la richiesta nel tempo previsto: in tal caso essa deve emettere un segnale sulla linea di $\overline{WAIT}$. La CPU controlla lo stato di questa linea in corrispondenza del fronte di discesa a meta' di T_2: se essa e' attiva inserisce uno o piu' cicli di clock di attesa, in modo da aspettare la risposta della memoria. Durante T_3 e T_4 si ha un nuovo segnale di $\overline{MREQ}$, questa volta senza $\overline{RD}$, ma con $\overline{RFSH}$; infatti ora la CPU e' occupata a decodificare ed eseguire

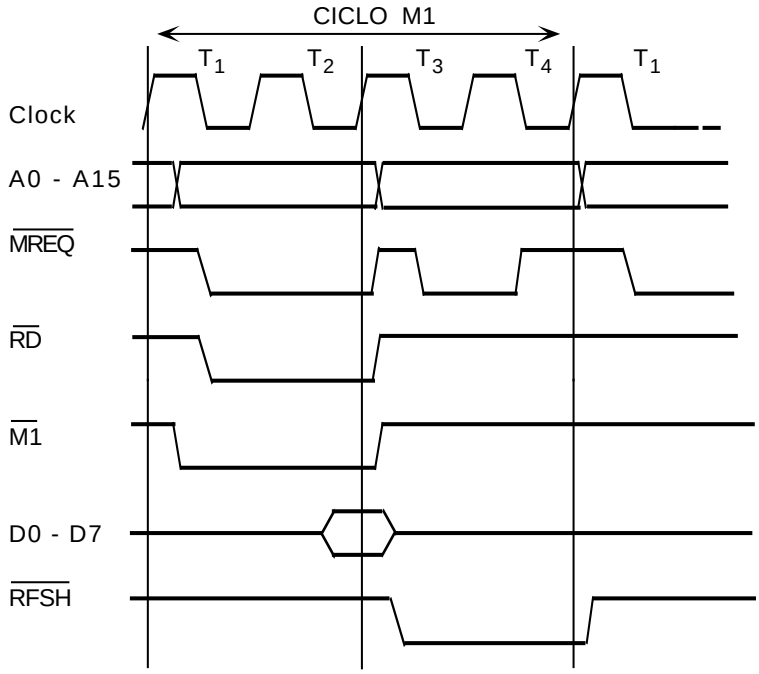

Figura 8.5: *Temporizzazioni dello Z80: fetch dell'istruzione*

l'istruzione e questo tempo viene usato per il refresh. Si noti che ora sul bus degli indirizzi viene presentato un indirizzo diverso, costruito partendo dal contenuto del registro R e non del PC.
Il segnale $\overline{M_1}$ diviene attivo simultaneamente a T_1 e resta tale fino alla fine di T_2: questo segnale contraddistingue quindi la fase di *fetch* di ogni istruzione, e permette tra l'altro, di individuarne l'inizio.
Abbiamo detto che un'istruzione puo' richiedere un ulteriore accesso, in lettura o scrittura, della memoria. Nel caso di lettura esso e' descritto dalla Fig. 8.6a. In questo caso lo Z80 legge il dato sul fronte di discesa di T_3.
L'operazione di scrittura e' abbastanza diversa (Fig. 8.6b); in questo caso la CPU presenta il dato sul bus a meta' di T_1, ed emette il segnale di $\overline{WR}$ a meta' di T_2: la memoria puo' quindi usare il fronte di discesa di questo segnale per leggere il dato. Naturalmente, sia in lettura che in scrittura, l'uso della linea di $\overline{WAIT}$ consente di sincronizzarsi con memorie piu' lente. Le operazioni di Input/Output con dispositivi diversi dalla memoria avvengono in modo analogo, con due differenze:
si usa la linea di $\overline{IORQ}$ invece di $\overline{MREQ}$;
viene aggiunto comunque un ciclo di wait tra T_2 e T_3; questo serve al dispositivo di I/O per avere il tempo di decodificare l'indirizzo ed attivare eventualmente la linea di $\overline{WAIT}$ (Fig. 8.6c).
Conviene infine mostrare cio' che avviene quando si ha, dall'esterno, un segnale di $\overline{BUSRQ}$ (Fig. 8.6d). Tale segnale viene campionato dalla CPU con il fronte di salita dell'ultimo periodi di clock di ogni ciclo macchina. Se esso e' attivo, la CPU pone in alta impedenza le proprie uscite sui bus e sui segnali di controllo, in coincidenza con il fronte di salita del successivo impulso di clock ed attiva la linea $\overline{BUSAK}$. Da questo momento, qualsiasi

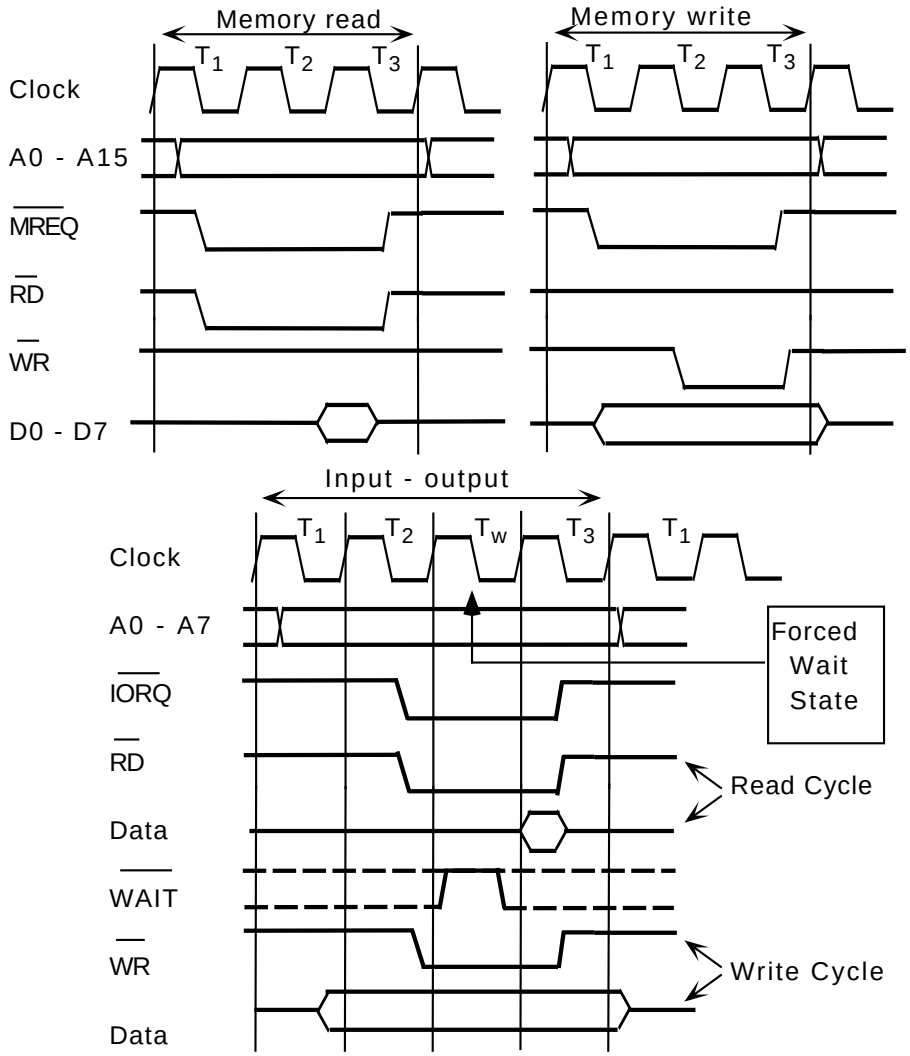

Figura 8.6: *Temporizzazioni dello Z80: a)ciclo di read; b)ciclo di write; c)input/output.*

dispositivo esterno puo' controllare il bus per effettuare trasferimenti tra la memoria ed i dispositivi di I/O. Sfrutteremo questa funzione per caricare in memoria i nostri programmi di prova.

8.3.2 Le istruzioni dello Z80

Possiamo ora analizzare in modo sistematico il set di istruzioni disponibile per lo Z80. Diciamo anzitutto che il progettista ha cercato di sfruttare al massimo le potenzialita' del sistema; infatti in linea di principio il set di istruzioni indispensabile per poter risolvere qualunque problema matematico o di manipolazione di dati e' in realta' relativamente esiguo. Tuttavia l'avere a disposizione istruzioni addizionali, e la possibilita' di usare molti diversi modi di indirizzamento puo' consentire di scrivere programmi molto piu' veloci ed efficienti.

Le istruzioni piu' importanti hanno, come abbiamo gia' visto, un codice a 1 byte; pertanto sono possibili 256 istruzioni diverse, che possiamo visualizzare sotto forma di matrice (Fig. 8.7), dove le righe sono associate al primo carattere esadecimale, e le colonne al secondo: ad esempio, il codice $0D$ corrisponde all'istruzione DEC C, mentre il codice 86 corrisponde all'istruzione ADD A,(HL), gia' vista in un precedente esempio.

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NOP	LD BC, **	LD (BC),A	INC BC	INC B	DEC B	LD B, *	RLCA	EX AF,AF'	ADD HL,BC	LD A,(BC)	DEC BC	INC C	DEC C	LD C, *	RRCA
1	DJNZ *	LD DE, **	LD (DE),A	INC DE	INC D	DEC D	LD D, *	RLA	JR *	ADD HL,DE	LD A,(DE)	DEC DE	INC E	DEC E	LD E, *	RRA
2	JRNZ, *	LD HL, **	LD (**),HL	INC HL	INC H	DEC H	LD H, *	DAA	JRZ, *	ADD HL,HL	LD HL,(**)	DEC HL	INC L	DEC L	LD L, *	CPL
3	JRNC, *	LD SP, **	LD (**),A	INC SP	INC (HL)	DEC (HL)	LD (HL),*	SCF	JRC, *	ADD HL,SP	LDA, **	DEC SP	INC A	DEC A	LD A, *	CCF
4	LD B,B	LD B,C	LD B,D	LD B,E	LD B,H	LD B,L	LD B,(HL)	LD B,A	LD C,B	LD C,C	LD C,D	LD C,E	LD C,H	LD C,L	LD C,(HL)	LD C,A
5	LD D,B	LD D,C	LD D,D	LD D,E	LD D,H	LD D,L	LD D,(HL)	LD D,A	LD E,B	LD E,C	LD E,D	LD E,E	LD E,H	LD E,L	LD E,(HL)	LD E,A
6	LD H,B	LD H,C	LD H,D	LD H,E	LD H,H	LD H,L	LD H,(HL)	LD H,A	LD L,B	LD L,C	LD L,D	LD L,E	LD L,H	LD L,L	LD L,(HL)	LD L,A
7	LD (HL),B	LD (HL),C	LD (HL),D	LD (HL),E	LD (HL),H	LD (HL),L	HALT	LD (HL),A	LD A,B	LD A,C	LD A,D	LD A,E	LD A,H	LD A,L	LD A,(HL)	LD A,A
8	ADD A,B	ADD A,C	ADD A,D	ADD A,E	ADD A,H	ADD A,L	ADD A,(HL)	ADD A,A	ADC A,B	ADC A,C	ADC A,D	ADC A,E	ADC A,H	ADC A,L	ADC A,(HL)	ADC A,A
9	SUB A,B	SUB A,C	SUB A,D	SUB A,E	SUB A,H	SUB A,L	SUB A,(HL)	SUB A,A	SBC A,B	SBC A,C	SBC A,D	SBC A,E	SBC A,H	SBC A,L	SBC A,(HL)	SBC A,A
A	AND B	AND C	AND D	AND E	AND H	AND L	AND (HL)	AND A	XOR B	XOR C	XOR D	XOR E	XOR H	XOR L	XOR (HL)	XOR A
B	OR B	OR C	OR D	OR E	OR H	OR L	OR (HL)	OR A	CP B	CP C	CP D	CP E	CP H	CP L	CP (HL)	CP A
C	RET NZ	POP BC	JPNZ, **	JP **	CALL NZ,**	PUSH BC	ADD A,*	RST 0H	RET Z	RET	JPZ, **	(1)	CALL Z,**	CALL **	ADC A,*	RST 8H
D	RET NC	POP DE	JPNC, **	OUT (*),A	CALL NC,**	PUSH DE	SUB *	RST 10H	RET C	EXX	JPC, **	IN A,(*)	CALL C,**	(1)	SBC A,*	RST 18H
E	RET PO	POP HL	JPPO, **	EX (SP),HL	CALL PO,**	PUSH DE	AND *	RST 20H	RET PE	JP (HL)	JPPE, **	EX DE,HL	CALL PE,**	(1)	XOR *	RST 28H
F	RET P	POP AF	JPP, **	DI	CALL P,**	PUSH AF	OR *	RST 30H	RET M	LD SP,HL	JPM, **	EI	CALL M,**	(1)	CP *	RST 38H

NOTE
(1): Primo byte di una istruzione con codice a piu' bytes;
***: L'istruzione si completa con un byte;**
****: L'istruzione si completa con due bytes.**

Figura 8.7: *Codici operativi dello Z80*

In questa matrice sono anche contenute le istruzioni a piu' bytes: ad esempio $C3$ (salta a ...) si completa con due bytes di indirizzo, mentre $3E$ (carica in A un numero a 8 bit) si completa con un byte contenente il dato da caricare nell'accumulatore. Quando l'istruzione viene decodificata la CPU capisce quindi se deve o meno aspettarsi ulteriori bytes di completamento.
Sarebbe troppo lungo qui descrivere dettagliatamente tutte le 256 istruzioni (chi fosse interessato puo' consultare il manuale dello Z80). Possiamo tuttavia tentare una descrizione generale osservando che esse possono essere divise in varie categorie:

Caricamento a 8 o 16 bit

Queste istruzioni muovono i dati tra i vari registri della CPU, ovvero tra i registri e la memoria. Questo gruppo comprende pure istruzioni di caricamento immediato del dato specificato nell'istruzione stessa in uno dei registri o in una qualsiasi posizione di memoria. Le istruzioni di scambio (Exchange) consentono poi di effettuare lo scambio tra il contenuto di due registri. In questo ambito rientrano anche istruzioni che, ad esempio, consentono di scambiare il contenuto dei registri principali con quello dei registri secondari.

Istruzioni aritmetiche e logiche
: Operano su dati posti nell'accumulatore e in un altro qualsiasi dei registri di uso generale, oppure su dati posti in accumulatore e una qualsiasi locazione di memoria. Sono anche possibili somme e sottrazioni a 16 bit.

Istruzioni di rotazione e scorrimento
: Permettono di far ruotare verso destra o verso sinistra il contenuto di qualunque registro o locazione di memoria, con o senza l'utilizzo del bit di Carry del registro F. Inoltre sono possibili rotazioni separate del gruppo meno significativo di 4 bit (nibble). La differenza tra rotazione e scorrimento e' mostrata in Fig. 8.8.

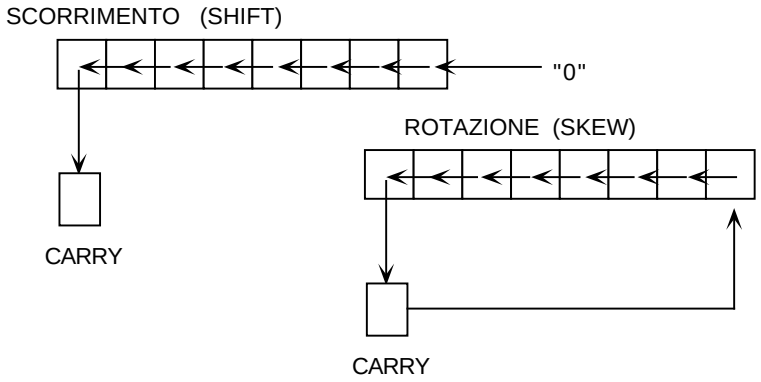

Figura 8.8: *Rotazioni e scorrimenti dei registri*

Manipolazioni di singoli bit
: E' possibile esaminare, porre a 1 (set), o porre a 0 (reset), singoli bit di ogni registro, o di qualunque locazione di memoria.

Istruzioni di salto, chiamata, o ritorno
: E' possibili fare salti (Jump), ovvero chiamate di subroutines (Call) con relativi ritorni. I salti possono essere condizionati al valore di specifici bits del registro F, in particolare il bit Z (zero) o il bit C (carry).

Istruzioni di input/output
: Esistono varie istruzioni per effettuare operazioni di I/O con dispositivi esterni, che vengono selezionati utilizzando gli 8 bit meno significativi del bus degli indirizzi. E' quindi possibile in linea di principio avere 256 dispositivi di I/O diversi. Torneremo piu' avanti su queste istruzioni.

Trasferimento e ricerca di blocchi di dati
: E' possibile trasferire un blocco di dati di qualsiasi dimensioni tra due diverse posizioni di memoria. Inoltre e' possibile analizzare un blocco di memoria per ricercare una particolare configurazione di 8 bit.

E' anche interessante osservare che esistono vari modi di indirizzamento della memoria, come abbiamo gia' imparato attraverso gli esempi visti; e' a questo punto utile esaminarli sistematicamente.

Indirizzamento immediato:
: Il dato (1 byte) viene fornito direttamente e caricato sul registro. Sono quindi istruzioni del tipo:
LDr, n: carica il byte n nel registro r
Il registro puo' essere A,B,C,D,E,H,L

Indirizzamento immediato esteso:
: LDdd, nn: carica i due bytes nn nella coppia di registri dd
La coppia di registri puo' essere BC,DE,HL oppure il registro SP.
Ovviamente i due bytes vanno dati nel consueto ordine (LSB e MSB)

Indirizzamento implicito:
: LDr, r': il contenuto del registro r' e' copiato in r
I registri utilizzabili sono A,B,C,D,E,H,L

Indirizzamento indiretto:
: LDr,(dd): carica nel registro r il contenuto della locazione di memoria il cui indirizzo e' nella coppia di registri dd.
Normalmente la coppia usata e' HL.

Indirizzamento indicizzato:
: LDr,(IX+d): carica nel registro r il contenuto della locazione di memoria indirizzata da IX piu' un offset d;
d e' un byte che viene fornito direttamente, ed e' interpretato come un numero di 7 bit con segno. Quindi l'offset possibile e' ± 127.

Indirizzamento relativo:
: Si applica solo nelle istruzioni di salto; consente un salto di ± 127 locazioni *relativamente* alla locazione corrente. E' un'istruzione del tipo:
JR e: salta dell'offset e (fornito direttamente)
ovvero
JRcc, e: salto condizionato (cc puo' essere Z,NZ,C,NC)

Modificato in pagina zero:
: Esiste solo per l'istruzione RST (restart): il programma salta ad alcune locazioni predefinite della pagina zero (cioe' della parte iniziale della memoria). Poiche' e' un'istruzione a un solo byte, la sua esecuzione e' piu' veloce del normale salto, e puo' quindi essere utile quandi si deve velocemente reagire ad un interrupt esterno.
L'istruzione quindi e' scritta come:
RSTgg, dove gg puo' essere $00, 08, 10, 18, 20, 28, 30, 38$.

8.3.3 Il concetto di catasta

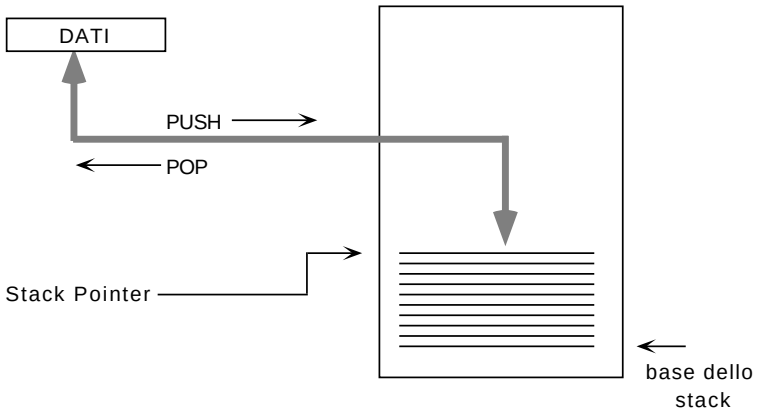

Figura 8.9: *Catasta*

Vediamo ora come viene gestito il problema degli indirizzi quando si utilizzano programmi strutturati in subroutines. Il programmatore deve inizialmente caricare nello Stack Pointer un indirizzo, corrispondente ad una zona di memoria libera, cioe' non utilizzata per altri scopi. Nel momento in cui, nel corso del programma, viene incontrata l'istruzione CALL nn
dove nn e' l'indirizzo di partenza della subroutine, lo Z80 effettua le seguenti operazioni:

il contenuto del Program Counter viene trasferito nelle locazioni di memoria immediatamente precedenti l'indirizzo contenuto nello Stack Pointer;

il contenuto dello Stack Pointer viene decrementato di 2;

nn viene trasferito nel Program Counter.

L'ultima istruzione della subroutine deve essere l'istruzione RET che viceversa provoca i seguenti effetti:

il contenuto della locazione di memoria puntata dallo Stack Pointer e di quella immediatamente successiva vengono trasferite nel Program Counter;

il contenuto dello Stack Pointer viene incrementato di 2.

Con questa tecnica e' possibile avere infinite chiamate a subroutines, una dentro l'altra, e ritrovare via via tutti gli indirizzi di ritorno man mano che ogni subroutines finisce. Si noti che lo stack (cioe' la catasta) parte da una base e cresce verso indirizzi decrescenti (Fig. 8.9).
E' anche possibile intervenire direttamente sulla catasta con le istruzioni PUSH e POP. Infatti l'istruzione:
PUSH pq (pq: coppia di registri)
trasferisce sulla catasta il contenuto dei due registri e dcrementa di 2 lo Stack Pointer. Invece

POP pq
copia nei registri il contenuto delle 2 locazioni piu' alte della catasta e incrementa di 2 il contenuto dello Stack Pointer.
Oltre che per gestire le subroutines, lo schema della catasta puo' essere usato per l'immagazzinamento temporaneo di dati, o per la gestione degli interrupts, di cui parleremo piu' avanti.

8.3.4 Operazioni di ingresso/uscita

In linea di principio un dispositivo periferico potrebbe essere considerato come una locazione nello spazio degli indirizzi, e quindi si potrebbero effettuare trasferimenti di dati attraverso semplici istruzioni di load. Molti microprocessori operano in questo modo; invece lo Z80 prevede istruzioni speciali per le operazioni di ingresso/uscita, che utilizzano solo gli 8 bit meno significativi del bus degli indirizzi e si distinguono dalle operazioni in memoria perche' attivano la linea $\overline{IORQ}$ anziche' la linea $\overline{MREQ}$. Uno dei vantaggi e' quello di avere istruzioni a due soli bytes, quindi piu' veloci.
Le istruzioni di input sono di due tipi:
IN A,(n) codice oggetto DB ..
il dispositivo periferico con indirizzo n e' letto ed il dato e' posto in A. n e' un numero a un byte fornito direttamente.
Oppure:
IN r,(C)
il dispositivo periferico il cui indirizzo e' contenuto in C e' letto ed il dato e' posto nel registro r. I registri utilizzabili sono A,B,C,D,E,H,L.
Analogamente, le operazioni di output sono gestite dalle istruzioni:
OUT (n),A codice oggetto D3 ..
ovvero
OUT (C),r
dove di nuovo r indica uno dei registri A,B,C,D,E,H,L.

8.3.5 Interruzioni

L'interazione del microprocessore con il modo esterno richiede spesso la possibilita' di intervenire in modo *asincrono*; in altre parole il microprocessore deve poter gestire operazioni di ingresso/uscita che vengono richieste dall'esterno, in un momento qualunque. Normalmente il microprocessore sta eseguendo un programma e deve a un certo istante servire la richiesta esterna: questo significa interrompere il programma in corso, andare ad eseguire una subroutine di servizio, e poi riprendere il normale lavoro da dove lo si e' interrotto. Naturalmente per poter riprendere il lavoro e' necessario aver salvato il contenuto dei registri quale era prima dell'interruzione. Questo e' molto facilitato dalla esistenza dei registri secondari; infatti il programmatore puo', all'inizio della subroutine di servizio, spostare il contenuto dei registri principali in quelli secondari, eseguire cio' che deve essere fatto per servire l'interruzione e, alla fine, ripristinare il contenuto dei registri principali. Lo Z80 ha a disposizione 3 linee per la gestione degli interrupts:

$\overline{INT}$

Questo segnale puo' abilitato con l'istruzione EI, o puo' essere mascherato, cioe'

disabilitato, con l'istruzione DI. La reazione puo' essere di 3 tipi diversi, selezionati con le istruzioni IM0,IM1 e IM2.

interrupt di modo 0
: All'arrivo del segnale $\overline{INT}$ lo Z80 genera un segnale $\overline{M1}$ e $\overline{IORQ}$ e si aspetta di ricevere sul bus dei dati, entro il ciclo di clock successivo, un byte che viene interpretato come un'istruzione da eseguire. Tipicamente la periferica interessata porra' sul bus dei dati un'istruzione RSTgg, che provoca il salto alla locazione gg; qui il programmatore avra' posto una routine che serve l'interruzione. E' chiaro che in questo modo si possono prevedere fino a sei routines diverse, ciascuna destinata ad una specifica periferica

interrupt di modo 1
: In questo caso lo Z80 salta direttamente alla locazione 38 (cioe' come se avesse ricevuto un comando RST 38); chiaramente e' una modalita' molto semplice e veloce ma poco flessibile.

interrupt di modo 2
: E' il modo piu' sofisticato e flessibile: come nel modo 0 lo Z80 legge il dato presente sul bus, ma lo interpreta come la parte meno significativa di un indirizzo. La parte piu' significativa viene presa dal registro I, e l'indirizzo completo viene caricato sul Program Counter; in questo modo possono essere messe in esecuzione molte routines diverse per servire in modo adeguato svariati dispositivi di I/O

$\overline{NMI}$
: Non puo' essere disabilitato da programma (la sigla significa infatti Interrupt Non Mascherabile). Alla fine dell'istruzione in corso lo Z80 salva nello stack il contenuto del Program Counter carica su di esso l'indirizzo contenuto nelle locazioni 66 e 66+1. Quindi il programmatore deve mettere in queste locazioni l'indirizzo di partenza della routine di servizio.

$\overline{BUSRQ}$
: E' il meccanismo a piu' alta priorita', e viene normalmente usato per il DMA (Direct Memory Acess), cioe' trasferimenti diretti da memoria a periferica o viceversa.

Le prestazioni offerte dallo Z80 possono sembrare fin troppo complicate; tuttavia esse possono essere necessarie per gestire sistemi complessi, con molte periferiche. In linea di principio, il programmatore deve prevedere anche la possibilita' che un interrupt arrivi mentre lo Z80 sta servendo un interrupt arrivato in precedenza da un'altra periferica; si capisce come il problema diventi rapidamente molto delicato.

8.4 La scheda didattica Z80

Data la complessita' dell'hardware e del software necessari per trasformare un microprocessore in un sistema funzionale, esistono i cosidetti sistemi di sviluppo (microcomputer development systems) offerti in commercio dagli stessi fabbricanti del microprocessore o da societa' del settore, che permettono sia l'addestramento che la realizzazione di sistemi funzionali completi. Tuttavia a Roma abbiamo preferito sviluppare un sistema ancora piu'

semplice, che chiameremo scheda didattica (Fig. 8.10), particolarmente adatta per esperienze sullo Z80 realizzate dagli studenti. Essa verra' descritta nel seguito, e ci servira' come base per la realizzazione di alcune semplici applicazioni. La scheda, dotata di una memoria CMOS statica da 2 kbytes, consente di scrivere ed eseguire piccoli programmi, collegare e gestire periferiche, visualizzare lo stato dei bus. E' corredata da un generatore di clock da 1 Mhz, ma puo' accettare un clock esterno di qualunque frequenza per consentire la visualizzazione dello stato delle varie linee.

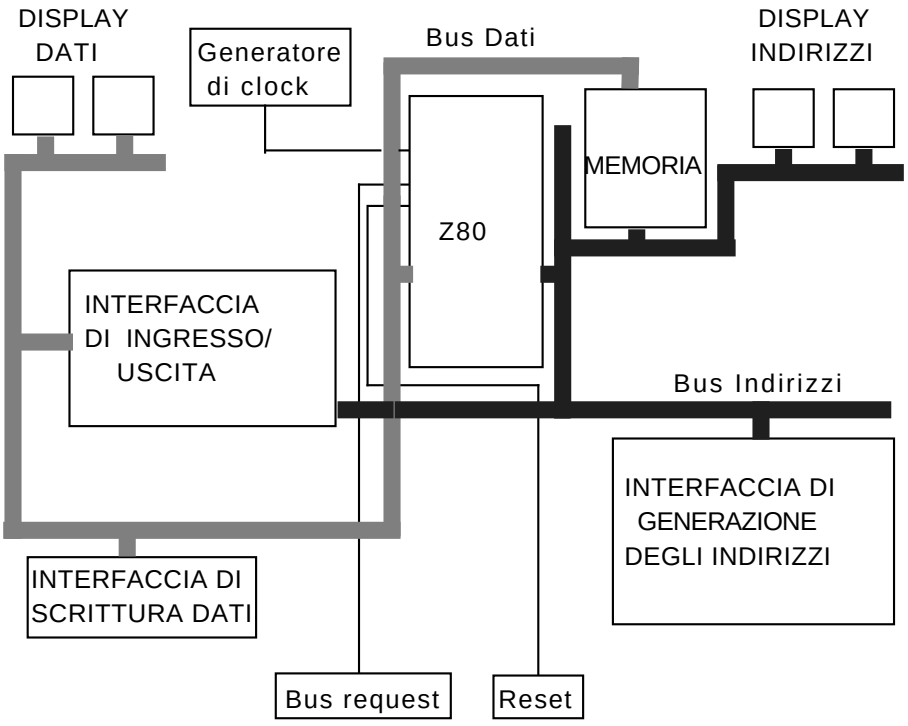

Figura 8.10: *La scheda didattica Z80: schema a blocchi.*

8.4.1 Descrizione circuitale

La CPU Z80 ha il bus dei dati connesso ad una RAM CMOS statica da 2 Kbytes ed all'interfaccia per la scrittura di dati; e' inoltre previsto un connettore pe il bus alla basetta per esperimenti. Il bus degli indirizzi (si utilizzano solo le linee $A_0 \ldots A_7$) e' collegato alla RAM, all'interfaccia di generazione degli indirizzi ed a un decodificatore (74LS138) per la selezione dei dispositivi di I/O che possono essere allocati sulla basetta per esperimenti. I segnali di controllo utilizzati sono $\overline{MREQ}$, $\overline{RD}$, $\overline{WR}$ e $\overline{BUSAK}$ per la temporizzazione delle operazioni in memoria, $\overline{BUSRQ}$ e $\overline{BUSAK}$ per la gestione dei bus, $\overline{IORQ}$ per l'interfaccia di ingresso/uscita, $\overline{RESET}$ per l'inizializzazione e $\overline{HALT}$ per visualizzare lo stato di (eventuale) attesa della CPU. I restanti segnali non sono utilizzati e i corrispondenti piedini sono lasciati aperti o collegati, se necessario, a $+5\ V$ tramite un resistore di pull-up.

Interfaccia di generazione degli indirizzi

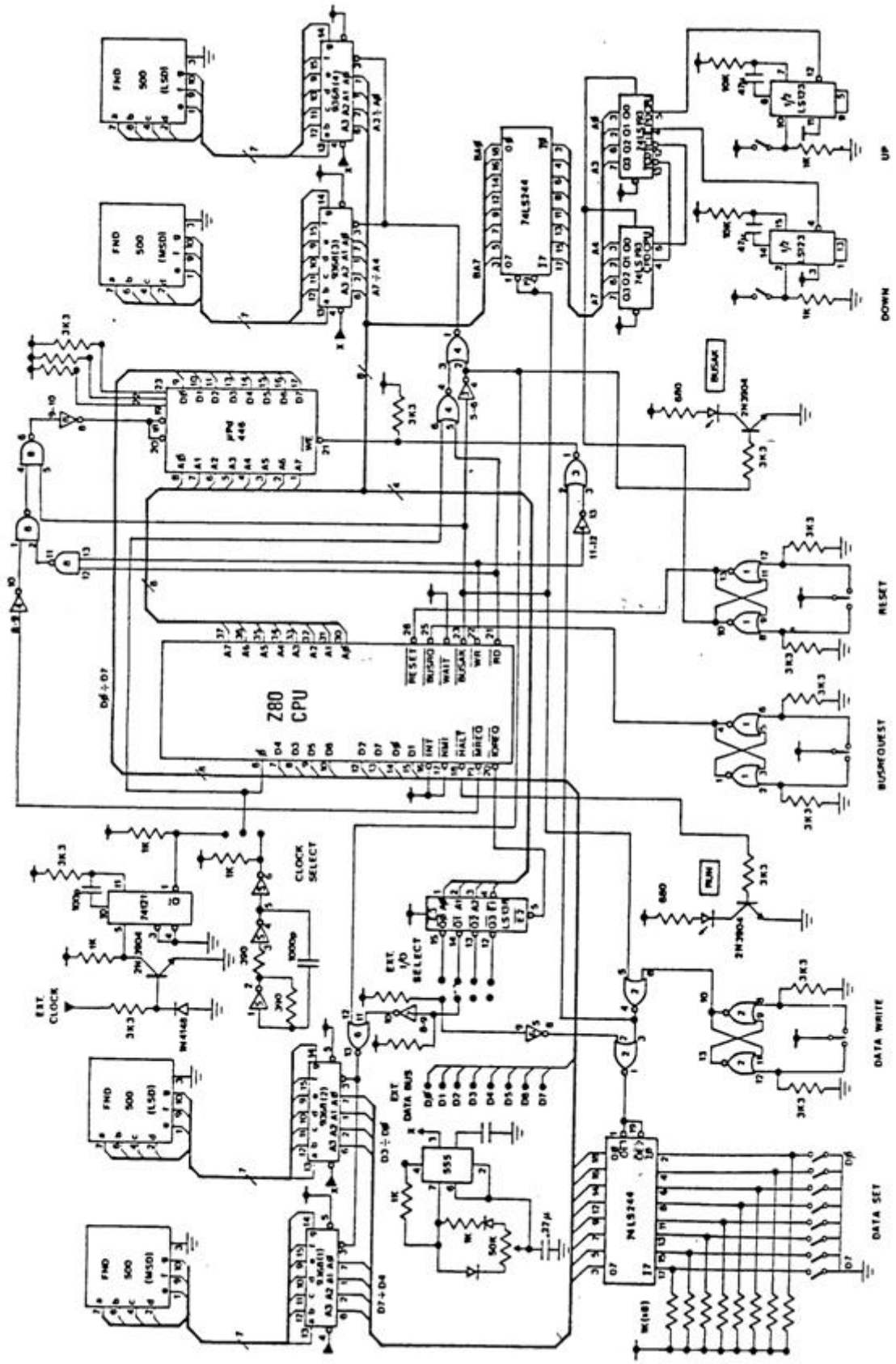

Figura 8.11: *La scheda didattica Z80: diagramma circuitale*

Il compito di questo blocco e' quello di generare, in sequenza, gli indirizzi da porre sul bus relativo. Tali indirizzi sono visualizzati su un display esadecimale. L'indirizzo viene fornito dalle uscite di 2 contatori up/down (74LS193) collegati in cascata e comandati da un monostabile doppio (74LS123). Due tasti (UP e DOWN) consentono di incrementare o decrementare il conteggio. Il comando di reset generale azzera il contenutoi dei contatori, mentre, mantenendo premuto uno dei tasti, si genera un treno di impulsi che consente il rapido incremento o decremento degli indirizzi generati. Gli indirizzi ottenuti vengono inseriti sul bus tramite un buffer (74LS244) abilitato dal segnale $\overline{BUSAK}$ della CPU; cio' evita interferenze con la normale attivita' della CPU.

Interfaccia di scrittura dati in memoria
I dati da scrivere in memoria vengono predisposti mediante 8 interruttori a levetta, collegati al bus mediante un buffer (74LS244) abilitato dallAND logico tra $\overline{BUSAK}$ e il comando di scrittura $\overline{DATAWRITE}$; tale comando e' fornito da un flip-flop SR attivato da un pulsante.

Visualizzazione dei dati e degli indirizzi
La visualizzazione dei dati e' ottenuta da due decodifiche esadecimali (9368) che pilotano due display a sette segmenti; esse sono abilitate dal segnale di $\overline{BUSAK}$.
La visualizzazione degli indirizzi e' effettuata in modo analogo, ma la sua abilitazione e' piu' complessa. Infatti essa e' data da

$$\overline{BUSAK} + (\overline{RD} \cdot CLOCK)$$

In questo modo il contenuto del bus e' visualizzato sia quando si ha il controllo manuale del bus, sia durante la fase di lavoro della CPU. In quest'ultimo caso l'AND tra il clock e il segnale di $\overline{RD}$ e' necessario per eliminare gli indirizzi di refresh, che lo Z80 presenta sul bus alternati agli indirizzi veri.
Un oscillatore realizzato con un 555 abilita ad intervalli regolari tutte le decodifiche, con una frequenza abbastanza adeguata da sfruttare l'effetto di persistenza dell'immagine, e nello stesso tempo riduce la potenza dissipata dai display.

Memoria
La memoria viene indirizzata attraverso gli 8 bit meno significativi del bus; gli ulteriori 3 piedini di cui essa dispone (A_8, A_9 A_{10}) sono posti a+5 V con dei resistori di pull-up ed hanno la possibilita' di essere collegati a massa con dei ponticelli. In questo modo possono essere utilizzati solo 256 bytes per volta, sufficienti tuttavia per gli scopi didattici. con tutti e 3 i ponticelli inseriti gli indirizzi disponibili vanno da 000 a $0FF$; senza nessun ponticello da 700 a $7FF$.
La memoria viene utilizzata sia dalla CPU, sia per l'inserimento manuale dei programmi e dei dati; la sua abilitazione e' quindi data da:

$$\overline{BUSAK} + \overline{MREQ} \cdot (\overline{RD} + \overline{WR})$$

in questo modo non vengono ricevuti gli indirizzi di refresh, irrilevanti per una memoria statica.

Interfaccia di ingresso/uscita
I circuiti di I/O vengono indirizzati attraverso un decodificatore (74LS138) abilitato dal segnale $\overline{IORQ}$ della CPU e dal bit A_3 del bus degli indirizzi. Si utilizzano solo le tre linee A_0, A_1, A_2 e quindi si possono abilitare fino ad 8 periferiche. Tuttavia vengono utilizzate solo le prime quattro uscite del decodificatore ($Q_0 \dots Q_3$), che possono essere connesse alla basetta per esperimenti. E' inoltre previsto, tramite opportuni ponticelli, la possibilita' di collegare la sezione di generazione degli indirizzi come periferica di ingresso, e quella di visualizzazione dei dati come periferica d'uscita. In questo modo si possono realizzare semplici operazioni di input/output.

Generatore di clock
Il segnale di clock puo' essere prelevato tramite un commutatore ($CLOCK\ SELECT$) da due diverse sezioni. La prima e' costituita da un oscillatore a 1 Mhz formato da due invertitori (7404) con componenti opportuni, seguito da un buffer. La seconda sezione accetta un clock esterno fornito da un generatore che, attraverso uno stadio separatore, pilota un multivibratore (74121) che fornisce il segnale TTL per la CPU. Poiche' l'ingresso del multivibratore e' a trigger di Schmitt e' possibile pilotare il sistema con impulsi di breve durata (circa 200 ns formati dalla rete RC) e qualunque frequenza, mentre la CPU e' protetta da eventuali sovratensioni.
E' anche possibile, mediante un opportuno ponticello, utilizzare come clock esterno a bassa frequenza (100 Hz) il segnale di blanking dei display a sette segmenti.

Per immettere un programma nel sistema e farlo eseguire si deve quindi operare nel seguente modo:

1. prendere il controllo del bus mediante l'interruttore $BUSREQUEST$; l'accensione di un led verde ci conferma che la CPU ha riconosciuto e accettato la nostra richiesta;
2. mediante i pulsanti UP e DOWN posizionare il contatore degli indirizzi nella locazione di memoria da cui si desidera far partire il programma; l'indirizzo selezionato appare sul display;
3. impostare, mediante gli 8 interruttori, il codice binario dell'istruzione che si vuole inserire in memoria;
4. scrivere in memoria il dato impostato mediante il pulsante DATA WRITE;
5. incrementare di uno il contatore degli indirizzi e ripetere i passi 3,4 e 5 fino al termine del programma da inserire;
6. l'esattezza dei dati memorizzati puo' essere verificata locazione per locazione, decrementando il contatore degli indirizzi;
7. restituire i bus alla CPU mediante l'interruttore BUSREQUEST (il led verde si spegne);
8. premere per un istante il pulsante di RESET; si accende il led rosso di RUN e la CPU cerca la prima istruzione da eseguire nella locazione 0000.

Bibliografia

[1] R.Cervellati - D. Malosti: ELETTRONICA - Esercitazioni per il Laboratorio di Fisica
La Goliardica Editrice - Roma

[2] J.Millman: Circuiti e sistemi microelettronici
Boringhieri, 1985

[3] P.Horowitz - W.Hill: The Art of Electronics
Cambridge University Press, 1989

[4] T.C.Hayes - P.Horowitz: Student Manual for the Art of Electronics
Cambridge University Press, 1989

[5] R.Cervellati - P.Monacelli - S.Petrarca: Lezioni per il corso di Microprocessori
Dispense edite dal Dipartimento di Fisica - Universita' La Sapienza- Roma

[6] I.Vannucci: "Introduzione alla Scheda Didattica Z80"
Nota interna 870 (9/10/1986) - Dipartimento di Fisica - Universita' La Sapienza-Roma