

CAPITOLO XII

DISPOSITIVI DI COMMUTAZIONE

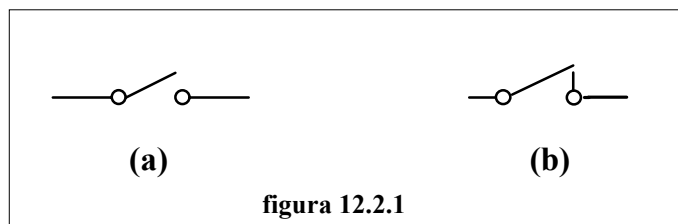
12.1) Introduzione.

La validita' di quanto e' stato esposto nei precedenti capitoli non dipende in alcun modo dal tipo dei dispositivi che vengono impiegati nella realizzazione circuitale. Tuttavia le caratteristiche reali dei componenti che nella pratica vengono usati, impongono vincoli non trascurabili al progettista.

E' quindi opportuno prendere in esame i piu' comuni dispositivi, mettendone in evidenza le principali caratteristiche.

12.2) Interruttori e rele'.

Il piu' semplice dispositivo di commutazione e' l'interruttore, cioe' un dispositivo capace di interrompere o chiudere un circuito. Un rele' e' invece un interruttore azionato da un elettromagnete. La posizione assunta dai suoi contatti, quando nella bobina dell'elettromagnete non circola corrente, viene detta posizione normale. Esistono pertanto rele' con **contatti normalmente aperti**, detti anche **contatti di lavoro**, e rele' con **contatti normalmente chiusi**, detti anche **contatti di riposo**. Il simbolo con cui vengono rappresentati questi contatti e' riportato rispettivamente in fig. 12.2.1 (a) e 12.2.1 (b).



La velocita' operativa dei rele' e' eccessivamente bassa perche' essi possano essere utilizzati nelle moderne apparecchiature digitali. Tuttavia essi trovano ancora largo impiego in numerose apparecchiature periferiche.

Un'interessante proprieta' dei rele' e' la loro natura bilaterale, che da' luogo ad alcuni particolari problemi, che verranno presi in considerazione nei paragrafi che seguono.

12.3) Caratteristiche basilari dei circuiti a rele'.

Come si e' gia' accennato in precedenza, la teoria della commutazione e' stata in origine sviluppata per la progettazione dei circuiti di commutazione telefonica che utilizzavano interruttori e rele'. Al giorno d'oggi i rele' sono stati largamente rimpiazzati dai circuiti elettronici nella maggior parte dei circuiti digitali; rimangono tuttavia alcune situazioni in cui i rele' sono i dispositivi piu' pratici per realizzare circuiti logici, come ad esempio nei casi in cui si debbano controllare potenze non indifferenti.

Un rele' e' realizzato con un elettromagnete e un pacco di contatti. I contatti operano quando nella bobina del rele' viene fatta circolare corrente. Si e' gia' parlato al paragrafo precedente dei contatti normalmente aperti (NO) e dei contatti normalmente chiusi (NC); e' opportuno introdurre un terzo tipo di contatto, detto contatto di scambio, realizzato combinando un contatto NO e uno NC, come illustrato in fig. 12.3.1

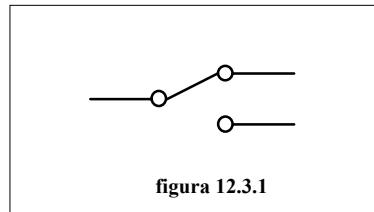


figura 12.3.1

Ciascun rele' puo' contenere nel suo pacco di contatti contatti di tutti i tre tipi illustrati.

La forma standard di un circuito a rele' e' riportata in fig. 12.3.2. Gli ingressi sono i segnali applicati alle bobine e il valore logico 1 viene associato alla circolazione di corrente nella bobina. Nella particolare realizzazione illustrata, cio' viene ottenuto connettendo un ingresso a terra. La funzione di commutazione viene realizzata interconnettendo opportunamente i contatti del rele'. Le uscite vengono chiamate trasmissioni T_i , e assumono valore 1 se sono connesse a terra attraverso la rete, 0 altrimenti.

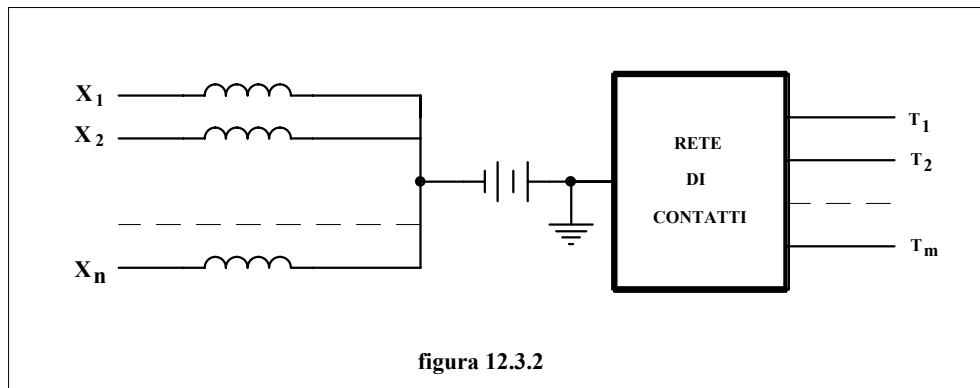
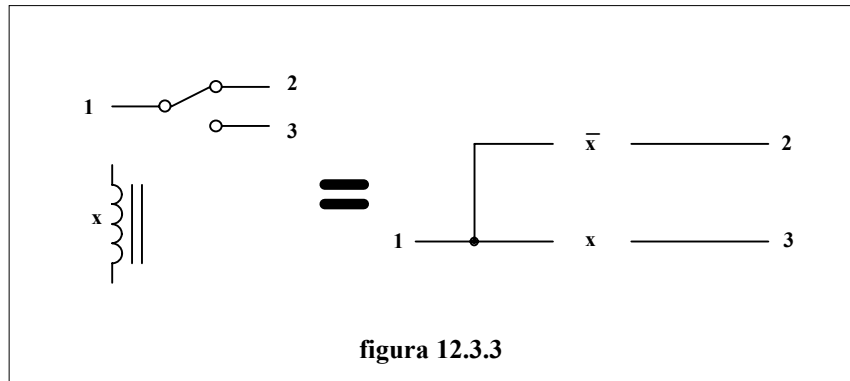


figura 12.3.2

Si e' gia' visto in precedenza che nell'analogia di Shannon qualsiasi circuito serie-parallelo a contatti puo' essere descritto da una funzione di commutazione e, inversamente, che ad ogni funzione di commutazione puo' esser fatto corrispondere un circuito di commutazione a contatti.

Un'interessante proprieta' dei circuiti a rele' rispetto agli altri tipi di circuito di commutazione e' che la loro velocita' operativa e' in pratica indipendente dal numero di livelli della funzione di commutazione; infatti, essendo la connessione tra ingresso e uscita della rete di contatti realizzata in pratica solo con conduttori, il ritardo di propagazione e' trascurabile e la velocita' operativa e' in sostanza determinata solo dai tempi di commutazione dei rele'.

La semplificazione dei circuiti a rele' tende a ridurre il numero di contatti; sarebbe evidentemente desiderabile ridurre anche il numero di rele', ma questo numero e' determinato dal numero di variabili richiesto e il piu' delle volte tale parametro non e' controllabile. Una via per ridurre il numero di contatti e' quella di combinare quelli normalmente aperti e quelli normalmente chiusi in contatti di scambio, ove possibile. Una rappresentazione schematica del contatto di scambio e' riportata in fig. 12.3.3.



Un altro metodo di semplificazione, peculiare dei circuiti a relè, fa uso del carattere bilaterale dei loro contatti. Essi infatti conducono in ambedue le direzioni, mentre i circuiti elettronici normalmente non lo possono fare; talvolta la bilateralità può essere vantaggiosamente sfruttata.

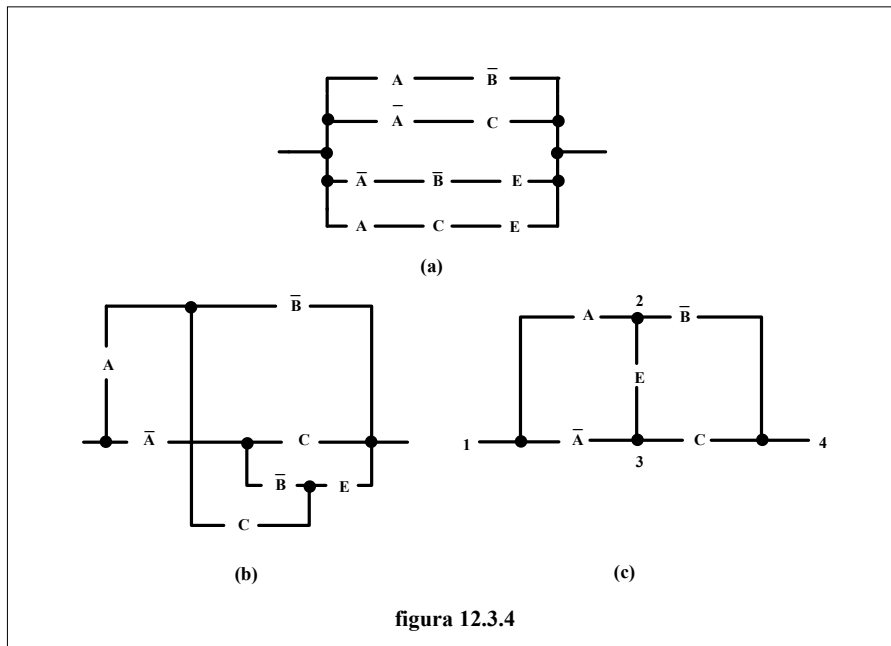
Si assuma, ad esempio, di voler realizzare la funzione:

$$T(A, B, C, D, E) = A \cdot \bar{B} + \bar{A} \cdot C + \bar{A} \cdot \bar{B} \cdot E + A \cdot C \cdot E$$

La realizzazione diretta da' origine al circuito di fig. 12.3.4 (a); per semplice ispezione poi il circuito può essere portato nella forma di fig. 12.3.4 (b).

Sfruttando poi la bilateralità del contatto E, ci si può ridurre al circuito di fig. 12.3.4 (c), che prende il nome di circuito a ponte o a traliccio.

Il percorso 1-3-4 realizza il termine $\bar{A} \cdot C$, quello 1-2-4 il termine $A \cdot \bar{B}$, quello 1-2-3-4 il termine $A \cdot C \cdot E$ ed infine quello 1-3-2-4 il termine $\bar{A} \cdot \bar{B} \cdot E$.



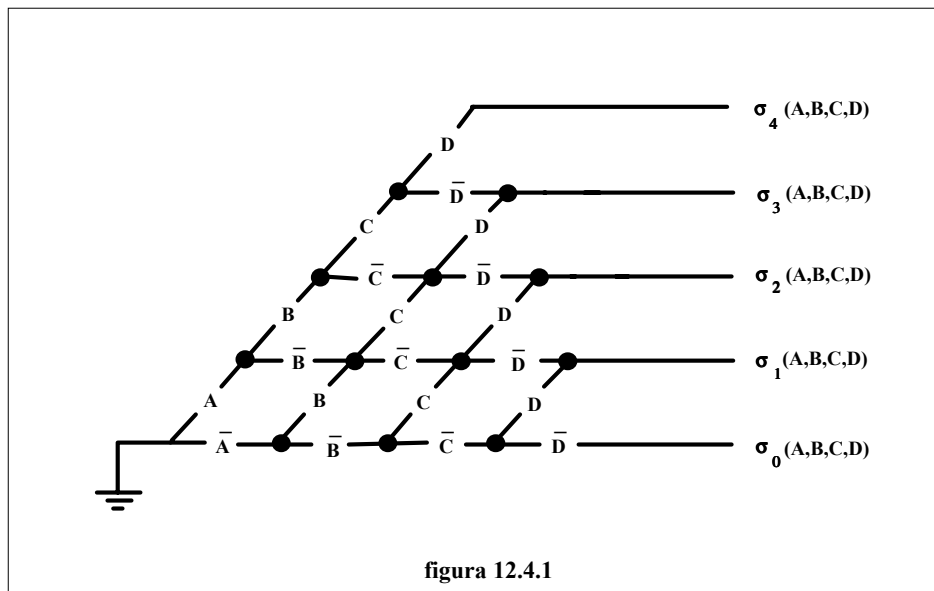
La domanda che ci si può porre a questo punto è allora come proceda il progetto di un circuito a relè. Il modo più semplice è quello di utilizzare le normali tecniche di

semplificazione per ottenere la forma minima a due livelli. Solo successivamente, con una combinazione intuitiva di ispezione visuale e manipolazione algebrica, i contatti vengono eliminati o trasformati in contatti di scambio, non appena cio' sia possibile. Infine va esaminato se non sia possibile passare ad una struttura a traliccio.

Come non esiste alcuna procedura sistematica che permetta di individuare la forma migliore per una funzione di commutazione, cosi' non esiste alcun modo per determinare il circuito minimo in una realizzazione a rele' o per giudicare se un dato circuito sia minimo. Un buon progetto a rele' deriva essenzialmente dal grado di esperienza del progettista.

12.4) Realizzazione a rele' delle funzioni simmetriche.

Le funzioni simmetriche possono essere realizzate in maniera notevolmente economica usando reti di contatti. Si noti infatti che le cinque uscite della rete di fig. 12.4.1 sono le cinque funzioni simmetriche elementari di quattro variabili.



Qualsiasi funzione simmetrica puo' quindi essere realizzata combinando assieme alcune di queste uscite e si deve far osservare che quando alcune uscite vengono connesse tra di loro alcuni contatti diventano ridondanti e possono essere eliminati.

E' evidente che tale discorso non risulta limitato alle funzioni simmetriche di quattro variabili, ma puo' essere esteso ad un numero di variabili qualsiasi. L'argomento e' stato affrontato ed esposto in dettaglio (*) da parecchi autori.

Un altro speciale tipo di rete che puo' essere facilmente realizzato con rele' e' la struttura ad albero binario di fig. 12.4.2.

Si noti che la struttura illustrata realizza un selettore di tre variabili; l'estensione ad un qualsiasi numero di variabili e' immediata.

(*) **Marcus M.P.** "Switching Circuits for Engineers" Prentice - Hall, Englewood Cliffs, N.J. 1962

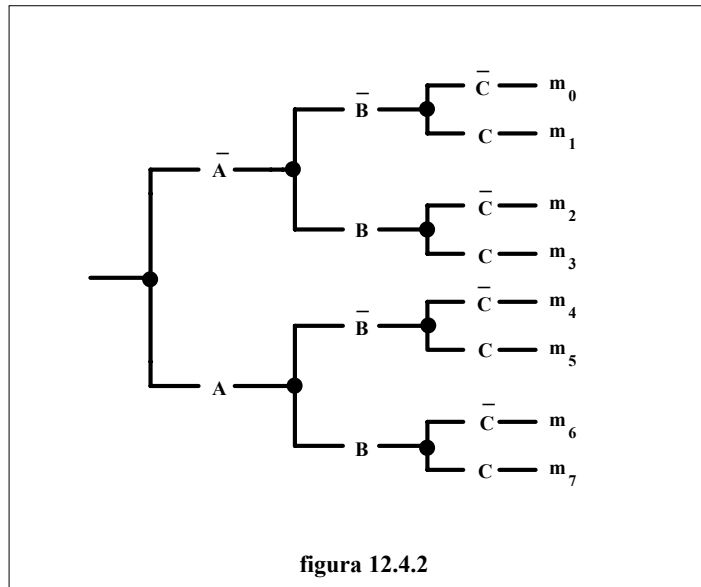


figura 12.4.2

12.5) I rele' nei circuiti sequenziali.

Tra i vari tipi di rele' disponibili esistono i rele' a tenuta (latching relays), cioe' quelli nei quali l'una o l'altra delle due posizioni puo' essere mantenuta anche quando l'eccitazione viene a cessare. Questi dispositivi operano come un flip-flop, in quanto quando uno di essi viene posizionato in uno dei suoi due stati, dando corrente a una delle sue due bobine, vi rimane finche' l'altra sua bobina non viene eccitata. Parimenti due rele' ordinari possono essere collegati in modo incrociato per formare elementi di memoria allo stesso modo in cui si collegano porte NAND e NOR. I sistemi che usano rele' a tenuta possono essere considerati circuiti sequenziali sincroni o a impulsi, tuttavia l'uso di questi componenti non e' molto diffuso presso i progettisti di circuiti a rele'. Piu' frequentemente alla rete combinatoria del rele' viene aggiunta una reazione in modo da formare un circuito sequenziale funzionante in modo fondamentale. In fig. 12.5.1 e' illustrata una configurazione circuitale che si comporta come un flip-flop SR, con l'ingresso A corrispondente al set e il B al reset.

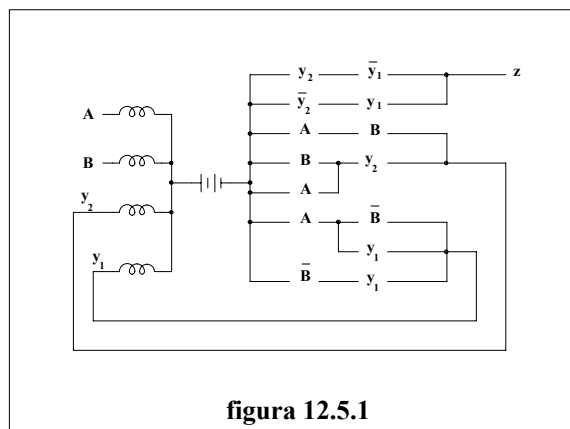


figura 12.5.1

Si noti che il ritardo introdotto dalla commutazione di un rele' e' di parecchi ordini di grandezza maggiore del ritardo dei circuiti elettronici ed e' in pratica uguale per tutte le uscite.

Di conseguenza la commutazione dei contatti dovuta alla transizione degli ingressi sarà sempre completata prima che un'ulteriore transizione possa avvenire a causa della variazione delle variabili di stato. Pertanto, a causa di questo ritardo, che si può immaginare concentrato tutto nei loop di reazione, le alee essenziali non sono un problema nei circuiti a rele'.

La tavola di flusso che corrisponde al circuito di fig. 12.5.1 è riportata in fig. 12.5.2. Le equazioni di eccitazione e di uscita sono:

$$y_2 = A \cdot B + B \cdot y_2 + A \cdot y_2$$

$$y_1 = A \cdot \overline{B} + A \cdot y_1 + \overline{B} \cdot y_1$$

$$z = y_2 \cdot y_1 + y_2 \cdot y_1$$

Poiché le uscite dipendono solo dalla storia delle transizioni di ingresso, la tavola di stato minima rimane una tavola di Moore e si ha un circuito privo di alee senza alcuna particolare precauzione.

		AB				z
		00	01	11	10	
y ₁ y ₂	00	(00)	(00)	10	01	0
	01	(01)	00	11	(01)	1
	11	01	10	(11)	(11)	0
	10	00	(10)	(10)	11	1

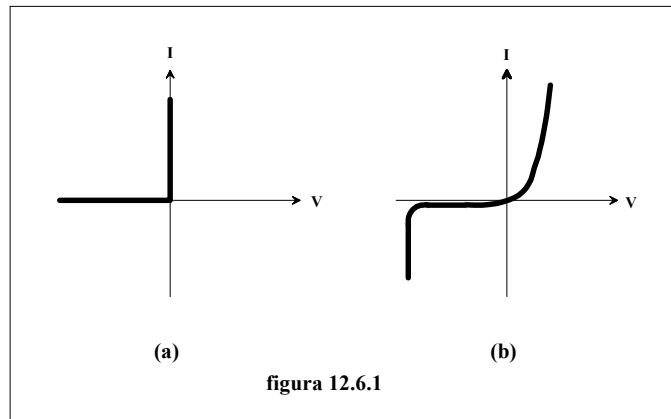
figura 12.5.2

12.6) I circuiti logici.

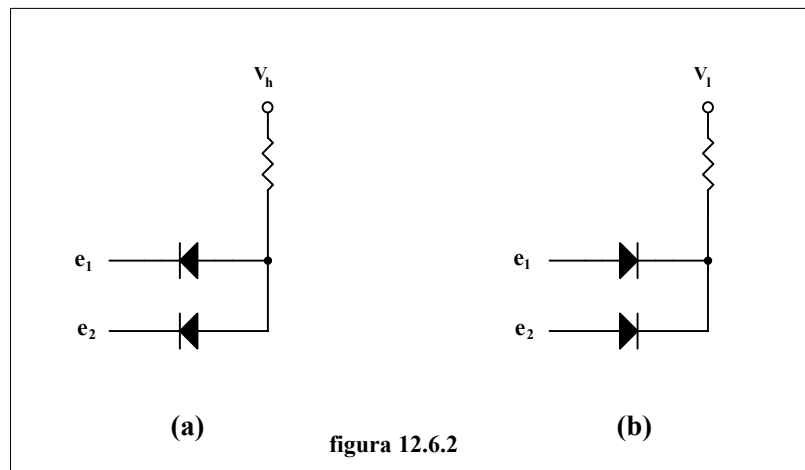
Gli interruttori e i rele' hanno avuto una grande importanza storica, ma al giorno d'oggi la maggior parte dei sistemi digitali è realizzata con dispositivi elettronici, essenzialmente diodi semiconduttori e transistori.

In termini ideali un diodo ha una resistenza nulla in senso diretto ed infinita in senso inverso e la sua caratteristica è quella riportata in fig. 12.6.1 (a). Nella realtà le due ipotesi enunciate non sono affatto verificate e l'effettiva caratteristica di un diodo reale è quella di fig. 12.6.1 (b).

Senza entrare in eccessivo dettaglio si può ricordare che per un diodo reale la resistenza diretta non è nulla, mentre quella inversa non è infinita. Inoltre per tensioni inverse sufficientemente grandi si ha il fenomeno della scarica zener per effetto del quale la corrente sale rapidamente.



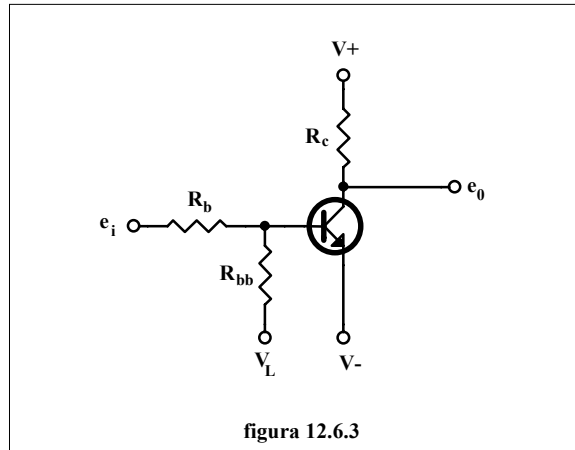
I circuiti per le due forme base di gate a diodi sono riportati in fig. 12.6.2. Il loro funzionamento e' intuitivo e la funzione realizzata da ciascuno d'essi dipende dalla logica, positiva o negativa, con cui operano. E' facile verificare che in logica positiva il primo circuito realizza la funzione logica AND mentre il secondo realizza quella OR.



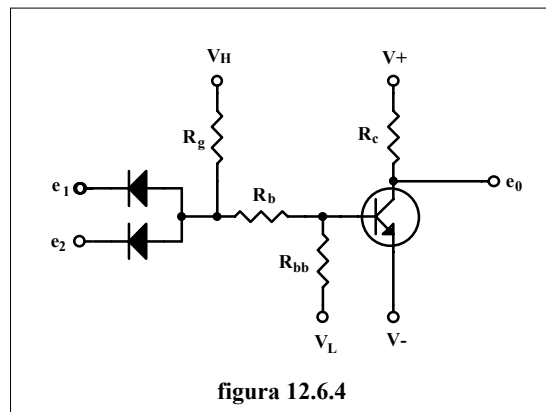
E' interessante far notare che se i diodi fossero ideali non ci sarebbe alcun limite al numero di ingressi di ciascun gate ne' alcun limite al numero di altri circuiti analoghi che potrebbero essere connessi all'uscita di un singolo gate. Invece con diodi reali le cadute ai capi dei diodi conduttori e le correnti inverse portano ad una continua degradazione dei livelli dei segnali via via che la complessita' e il numero di stadi in cascata aumentano.

Questo inconveniente della logica a diodi puo' essere superato introducendo degli opportuni amplificatori, di solito realizzati a transistori. Tali amplificatori lavorano quasi sempre tra saturazione e interdizione, in modo da assicurare una notevole stabilita' dei livelli di tensione associati a ciascun stato logico, rendendo nel contempo disponibile una notevole escursione di tensione in corrispondenza alla commutazione. Alle velocita' di commutazione piu' elevate tuttavia puo' essere necessario ricorrere ad amplificatori non saturati, per motivi che verranno illustrati piu' avanti.

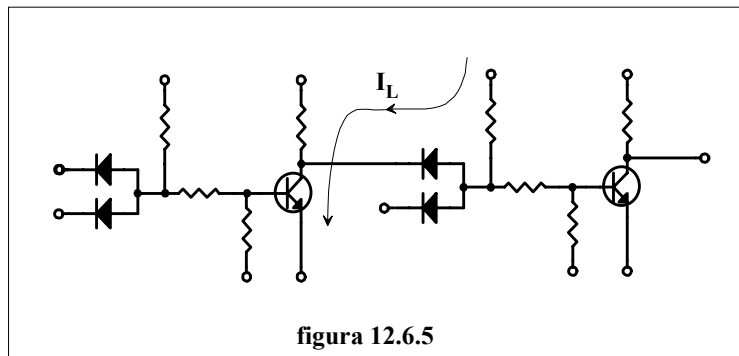
Il circuito base dell'invertitore, nel caso di transistori NPN, e' quello di fig. 12.6.3. Si tenga presente che utilizzando transistori al silicio spesso R_{bb} viene omessa e V_{-} vale zero.



Per correggere le deficienze di carico di un gate a diodi puo' essere sufficiente collegare alla sua uscita un invertitore, come mostrato in fig. 12.6.4.



Il funzionamento e' evidente e realizza in logica positiva la funzione NAND, in logica negativa la funzione NOR. E' bene far osservare che l'aggiunta di un invertitore diminuisce sostanzialmente i problemi di carico di un gate a diodi, ma non li elimina completamente. Si consideri infatti la situazione di fig. 12.6.5, in cui un gate NAND pilota un secondo gate dello stesso tipo.



Quando l'uscita del primo gate e' a livello basso, cioe' quando il relativo transistore e' saturato, in esso fluisce, come illustrato in fig. 12.6.5, una corrente di carico I_L . Se altri gate

vengono connessi in parallelo sull'uscita, essi contribuiscono ad aumentare I_L . Se tale corrente supera il limite determinato dalla corrente di base e dal guadagno di corrente del transistor, quest'ultimo esce di saturazione e la tensione di uscita sale, superando facilmente il livello logico basso. Anche il numero massimo di ingressi per ciascun gate risulta aumentato con l'introduzione dell'invertitore; infatti proprio per la presenza del transistor non e' piu' necessario assicurare che il livello alto di tensione in uscita dal gate a diodi non scenda al di sotto di un determinato livello, ma e' sufficiente che alla base del transistor venga fornita una corrente sufficiente a mantenerlo in saturazione. Pertanto la resistenza non infinita dei diodi polarizzati inversamente pesa molto meno che non nel caso in cui non ci sia in uscita l'invertitore.

C'e' infine da osservare che l'introduzione di un amplificatore a transistori elimina completamente la degradazione del segnale che si ha, via via che aumenta il numero degli stadi in cascata.

12.7) Velocita' e ritardo nei circuiti logici.

La velocita' operativa dei circuiti logici e' generalmente un parametro importante, soprattutto nelle applicazioni di calcolo automatico. Per tale ragione gran parte degli sforzi dedicati al campo della progettazione circuitale tendono a realizzare dispositivi sempre piu' veloci.

Malgrado che questo aspetto riguardi piu' i progettisti di circuiti che non i progettisti logici, la conoscenza, sia pure qualitativa, della natura delle transizioni e dei ritardi introdotti dalle strutture circuitali con cui si ha a che fare gioca un ruolo importante nella teoria logica dei circuiti sequenziali.

La fig. 12.7.1 illustra un invertitore nella configurazione a emettitore comune e la sua risposta ad un impulso positivo applicato alla base, tale da portarlo da interdizione a saturazione e viceversa.

Per determinare questa risposta e' necessario considerare un certo numero di fenomeni fisici complessi. Vale la pena di richiamare brevemente alcuni di essi.

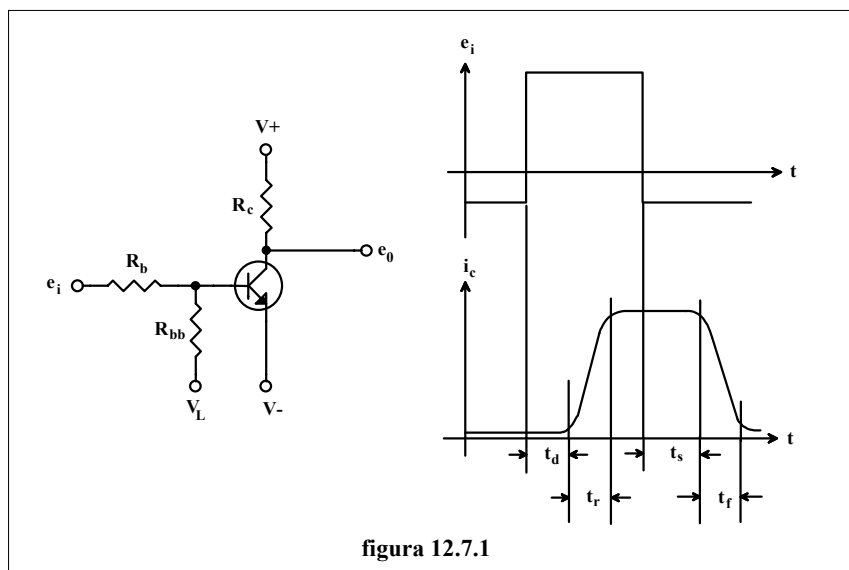


figura 12.7.1

A partire dall'istante in cui inizia l'impulso di base, vi e' un tempo di ritardo t_d prima che la corrente di collettore inizi a salire. Questo ritardo e' determinato essenzialmente dalla

capacita' totale base-emettitore del transistor, che deve venir caricata ad una tensione lievemente positiva prima che il transistor possa iniziare a condurre. Dopo l'inizio della conduzione si ha un tempo di salita finito t_r determinato principalmente dalla capacita' di collettore e dalle caratteristiche intrinseche del transistor.

Sul fronte di discesa dell'impulso di ingresso si ha un primo tempo t_s , detto tempo di restituzione, dovuto sia alla capacita' base-emettitore sia all'accumulo di cariche minoritarie nello spazio di base. Questo tempo e' forse il piu' importante tra i tempi di commutazione ed e' comunque notevolmente piu' grande che non t_d .

Infine vi e' un tempo di discesa finito t_f dovuto anch'esso, come il tempo di salita, alla capacita' di collettore e alle caratteristiche intrinseche del transistor.

Allo scopo di specificare il ritardo cui il segnale e' sottoposto nell'attraversare i gate i tempi t_r e t_f sono significativi solo se i corrispondenti tempi del segnale presente all'ingresso sono trascurabili. E' questo tuttavia un caso abbastanza raro, poiche' i circuiti logici sono normalmente pilotati da circuiti analoghi e di conseguenza i tempi di salita e di discesa sono dello stesso ordine di grandezza sia all'ingresso che all'uscita. Essi tuttavia, pur non essendo completamente significativi, influenzano il valore del tempo di ritardo a seconda di come questo tempo di ritardo viene misurato. E' necessario quindi specificare i punti delle forme d'onda di ingresso e di uscita usati per definire i tempi di ritardo stessi.

A questo riguardo e' necessario notare che i transistori usati nei circuiti logici lavorano normalmente in regime di saturazione. Anche se il segnale di ingresso varia sull'intero campo di valori ammissibili, solo un piccolo intervallo di questo campo e' quello che determina la commutazione tra un livello logico e l'altro. Per i principali dispositivi logici tale intervallo si aggira su 0.1 volt e quindi i tempi di salita e discesa del segnale di uscita saranno totalmente determinati dai parametri interni del circuito logico a meno che il segnale di ingresso non vari talmente lentamente da permanere in questo intervallo per tempi notevolmente maggiori dei tempi di commutazione nominali.

In definitiva i tempi di ritardo di un gate vengono misurati, come e' illustrato in fig. 12.7.2, pilotando l'ingresso con un impulso che abbia tempo di salita e discesa dello stesso ordine di grandezza di quelli del segnale di uscita ed eseguendo la misura con riferimento a punti ben specificati, quali ad esempio possono essere i tempi all'emivalore.

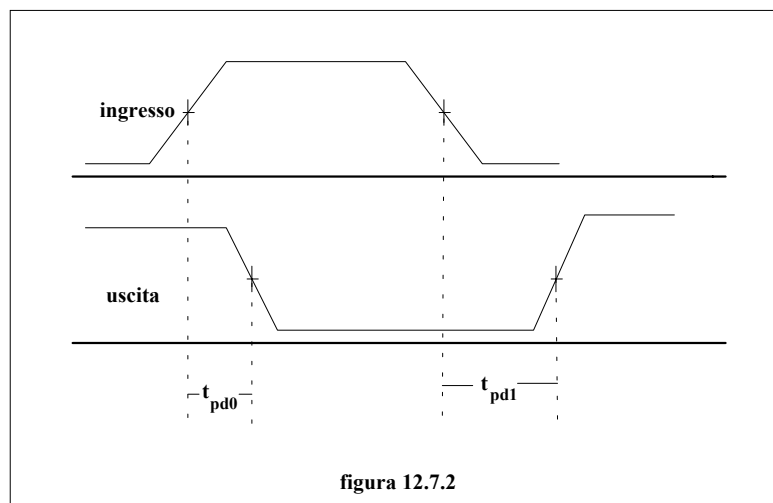


figura 12.7.2

Le due transizioni di solito non danno luogo a due tempi di ritardo uguali, com'e' logico aspettarsi dal momento che i fenomeni fisici interessati dalle due commutazioni sono diversi.

Per quanto riguarda la notazione per i tempi di ritardo essa purtroppo non e' standardizzata. La piu' comune indica con t_{pd0} o con t_{pd-} il ritardo associato al fronte di discesa dell'uscita, con t_{pd1} o t_{pd+} quello associato al fronte positivo. Il tempo di ritardo medio t_{pd} e' definito come media aritmetica di questi due tempi ed e' quello che piu' comunemente viene fornito dai costruttori tra i dati caratteristici.

12.8) I circuiti logici integrati.

Tutti i circuiti integrati digitali, con l'eccezione di alcune realizzazioni a larga scala, quali ad esempio le memorie o i microprocessori, sono riuniti in un certo numero di famiglie logiche. Ciascuna famiglia e' basata su un particolare tipo di circuito elementare usato in tutti i dispositivi della famiglia per realizzare invertitori, gate di diverso tipo ed elementi di memoria quali i flip-flop. Ciascuna famiglia inoltre include elementi a media e talvolta a larga scala di integrazione.

Tutti i dispositivi di una famiglia logica sono tra di loro compatibili, nel senso che i livelli logici sono gli stessi, possono essere alimentati con la stessa tensione e l'uscita di un elemento e' in grado di pilotare direttamente almeno un ingresso di qualsiasi altro elemento della stessa famiglia.

Al contrario, poiche' ogni famiglia logica usa un diverso circuito elementare, esse non sono di solito compatibili tra di loro, anche nel caso in cui alcune delle loro caratteristiche coincidano.

Ciascuna famiglia comprende, in aggiunta agli elementi base SSI, MSI e LSI, un certo numero di circuiti ausiliari, quali ad esempio traslatori di livello (per scopi di interfacciamento con altre famiglie), circuiti di condizionamento del segnale (trigger di Schmitt e multivibratori), circuiti di pilotaggio di visualizzatori e vari altri circuiti per applicazioni speciali. Il numero di questi circuiti ausiliari dipende dalla diffusione della famiglia stessa; quella a tutt'oggi piu' numerosa comprende oltre 300 diversi dispositivi, quella piu' povera meno di 50.

Le famiglie piu' comuni e piu' diffuse sono:

- TTL** - logica transistori - transistori
- DTL** - logica diodi - transistori
- HTL** - logica a elevata soglia
- RTL** - logica resistori - transistori
- ECL** - logica ad accoppiamento di emettitore chiamata anche logica a commutazione di corrente.
- CMOS** - logica a transistori MOS complementari.

Le prime cinque famiglie usano transistori bipolari, l'ultima transistori unipolari.

In aggiunta alle famiglie logiche esiste un'ampia varieta' di dispositivi realizzati in tecnologia MOS, normalmente a canale P e a rinforzo. Tuttavia tale tecnologia e' usata quasi esclusivamente per la produzione di circuiti LSI specializzati, quali memorie, microprocessori e circuiti "custom". Per tale motivo i circuiti MOS non sono inquadrati in una famiglia logica, ma sono di solito considerati prodotti individuali a larga scala di integrazione.

Nel discutere le famiglie logiche e' infine importante porre l'accento sul fatto che il loro sviluppo avviene in un ambiente dallo sviluppo tecnologico estremamente veloce e di intensa competizione. Tale fatto ha portato ad un rapido sviluppo dei componenti e, a causa della vitalita' del mercato, al nascere e all'affermarsi di un notevole gruppo di produttori. Al giorno

d'oggi, quando uno dei produttori ha successo nell'introdurre e commercializzare una determinata famiglia logica, essa viene immediatamente prodotta su licenza dalle altre compagnie, eventualmente completata con nuovi dispositivi o migliorata nelle sue caratteristiche.

12.9) Caratteristiche dei circuiti integrati digitali.

Ogni famiglia logica viene progettata per una differente applicazione ed ha in alcune sue caratteristiche i suoi punti di forza, in altre i suoi inconvenienti.

Poiche' il circuito base di ciascuna famiglia e' differente, i relativi dispositivi hanno caratteristiche diverse. Ad esempio i circuiti dei calcolatori di grosse dimensioni devono operare molto velocemente e di conseguenza per la loro realizzazione si useranno componenti di famiglie logiche veloci; se al contrario si prevedesse per il circuiti lunghi periodi di funzionamento con alimentazione fornita da batterie, allora diventa molto piu' importante la caratteristica di un basso consumo.

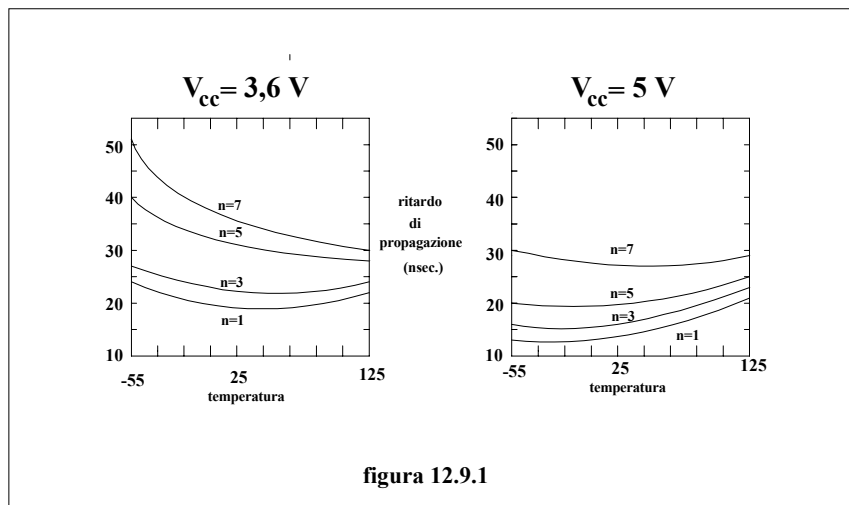
Sebbene una valutazione corretta dei circuiti digitali integrati richieda un attento studio dei dati forniti dal costruttore, si riportano qui di seguito le definizioni di quelle caratteristiche che si ritengono maggiormente interessanti e significative.

FAN-IN. E' il numero di ingressi che controllano un operatore logico. Molto spesso sono disponibili degli ingressi di espansione che permettono di aumentare il fan-in. L'eccessivo uso di ingressi di espansione fa tuttavia aumentare la capacita' di ingresso e di conseguenza il tempo di propagazione.

FAN-OUT. E' la misura del carico che puo' essere connesso all'uscita di un gate o di qualsiasi elemento logico della famiglia. Poiche' il carico e' normalmente dato da uno o piu' ingressi di altri elementi logici della stessa famiglia, il fan-out e' usualmente identificato dal numero di ingressi che possono venir connessi all'uscita dell'elemento logico considerato. Piu' esattamente ciascuna famiglia logica ha "un'unita' di carico" espressa in milliampere che rappresenta la corrente necessaria ad attivare un tipico circuito di quella famiglia logica. Il fan-out viene ottenuto normalizzando tramite l'unita' di carico la massima corrente di uscita. In sostanza quindi, quando ad esempio si parla di fan-out pari a 10, cio' sta a significare che puo' venir pilotato un numero di ingressi tale da costituire un fattore di carico normalizzato pari a 10.

LIVELLI LOGICI. L'uscita di un circuito digitale puo' assumere normalmente due soli stati, detti rispettivamente alto e basso. I corrispondenti livelli di tensione vengono indicati di solito con V_H e V_L e variano da famiglia a famiglia logica. E' opportuno tuttavia far notare che alcuni circuiti digitali vengono progettati per lavorare a tre stati e il terzo stato corrisponde a un circuito aperto.

RITARDO DI PROPAGAZIONE E VELOCITA'. Il ritardo di propagazione misura quanto rapidamente si propaghi la commutazione di un livello logico di ingresso verso l'uscita. Più esattamente si definiscono di solito due ritardi di propagazione che tengono conto della dissimmetria di funzionamento nelle due commutazioni: il ritardo di propagazione per l'uscita che commuta da livello basso a livello alto è indicata con t_{pLH} o t_{pd+} , quello relativo alla commutazione opposta con t_{pHL} o t_{pd-} . Ambedue i tempi vengono misurati tra i punti al 50% dei segnali di ingresso e di uscita. La semisomma di tali tempi prende il nome di ritardo medio di propagazione ed è indicata normalmente con t_{av} o t_{pd} . Il tempo di salita t_r e di discesa t_f sono invece misurati tra il 10% e il 90% del fronte di salita e rispettivamente di discesa dell'uscita. I tempi di propagazione variano con la temperatura, la tensione di alimentazione e il carico. Alcuni esempi sono riportati nei grafici di fig. 12.9.1. La velocità è direttamente correlata con il ritardo di propagazione e misura la massima frequenza cui può commutare un flip-flop senza commettere commutazioni errate.



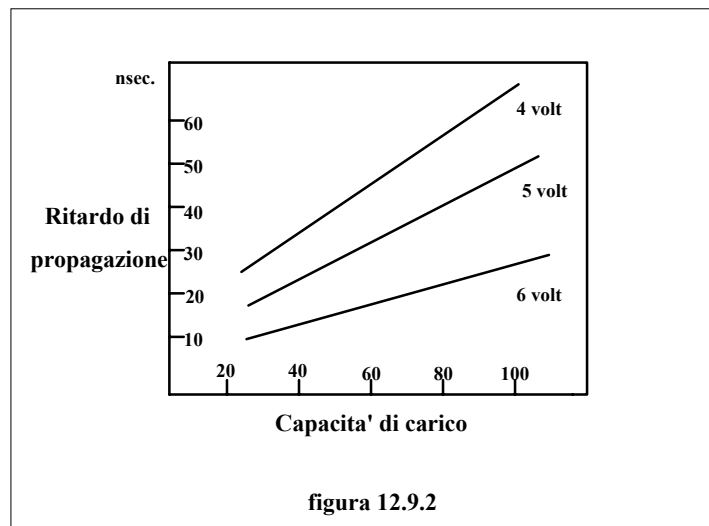
DISSIPAZIONE DI POTENZA. Come in qualsiasi altro circuito elettronico la potenza dissipata viene misurata in watt o in milliwatt. Essa rappresenta la potenza che effettivamente si dissipa nel dispositivo e non comprende quindi la potenza che viene eventualmente fornita ad altri dispositivi. Di regola essa indica nei circuiti integrati la potenza che si dissipa in un gate che lavora per metà del tempo allo stato alto e per l'altra metà allo stato basso e permette di valutare la potenza totale dissipata dal circuito ai fini del dimensionamento dell'alimentatore.

IMMUNITA' AL RUMORE. Misura la quantita' di segnale disturbante che puo' essere sovrapposto a un segnale logico senza causare un'errata commutazione dell'uscita. E' necessario distinguere tra immunita' al rumore **statica**, relativa a segnali che varino lentamente o ad impulsi sufficientemente lunghi e immunita' al rumore **dinamica**. E' abbastanza evidente infatti che un impulso di ingresso di durata comparabile con il ritardo di propagazione potra' avere un'ampiezza anche notevole senza dar luogo ad un'errata commutazione dell'uscita. La medesima ampiezza sarebbe invece intollerabile se il segnale di disturbo fosse a bassa frequenza. L'immunita' al rumore e' anch'essa funzione della temperatura, della tensione di alimentazione e dal carico.

CAMPO DI TEMPERATURE DI UTILIZZO. E' il campo di temperatura entro il quale un circuito integrato puo' operare senza che alcuna delle sue caratteristiche nominali sia compromessa.

ATTITUDINE AL PILOTAGGIO DI CARICHI CAPACITIVI.

L'attitudine a pilotare carichi capacitivi e' un'importante caratteristica dei circuiti digitali integrati, soprattutto quando un'uscita deve pilotare delle interconnessioni realizzate con cavi di una certa lunghezza.



Questa attitudine dipende direttamente dall'impedenza d'uscita del circuito e quanto minore e' questa impedenza, tanto meglio un carico capacitivo viene pilotato. Quando infatti ad un circuito viene applicato un carico capacitivo i tempi di salita e di discesa vengono largamente influenzati dal carico stesso. Il tempo di propagazione viene aumentato

in maniera proporzionale all'aumento della costante di tempo del circuito causata dal carico capacitivo. In fig. 12.9.2 e' riportato il tempo di propagazione in funzione della capacita' di carico e della tensione di alimentazione per un tipico circuito appartenente alla famiglia DTL.

PROTEZIONE DAL CORTO CIRCUITO. Impedisce che un circuito, la cui uscita sia stata connessa direttamente a terra, venga danneggiato. Tale proprieta' puo' essere riconosciuta semplicemente esaminando il circuito e riconoscendo se nel circuito d'uscita vi sia o meno una resistenza connessa tra alimentazione e terminale di uscita, destinata a limitare la corrente nel caso di corto circuito.

TOLLERANZA SULLE TENSIONI DI ALIMENTAZIONE. Definisce il campo di tensioni di alimentazione entro il quale le caratteristiche del dispositivo sono garantite.

TIPO DI CIRCUITO LOGICO. Dipende dal numero di circuiti elementari integrati su un unico substrato di silicio. Si possono distinguere le tre principali categorie:

Circuiti SSI (small scale integration) con un numero di porte logiche elementari che puo' raggiungere le 12 porte per involucro.

Circuiti MSI (Medium scale integration) con capacita' comprese tra le 12 e le 100 porte logiche.

Circuiti LSI (Large scale integration) con capacita' superiore alle 100 porte logiche.

La realizzazione SSI copre il campo dei singoli elementi logici e delle reti logiche piu' semplici, quelle MSI le reti logiche di dimensioni medio - grandi, quali convertitori A/D, multiplexer, ecc. Infine le realizzazioni LSI sono destinate a reti logiche di dimensioni molto ampie, quali memorie, unita' aritmetico-logiche, ecc.

TIPO DI INVOLUCRO. Esistono numerosi tipi di involucro per i circuiti integrati: I piu' comuni sono il TO-5, il flat pack, il dual in line e il SOT. Il TO-5 e' un involucro metallico, in pratica uguale a quello utilizzato per i transistori, con 8,10 o 12 piedini. Il flat pack e' un involucro misto metallo ceramica di forma rettangolare appiattita. Il flat pack viene prodotto in due dimensioni con 10, 12 o 14 piedini; questi a loro volta possono essere terminali piatti flessibili e dorati o collegamenti a sezione circolare in rame. L'involucro dual in

line e' forse il piu' diffuso, pur essendo di maggiori dimensioni che non gli altri due. I piedini, di tipo rigido, sono in numero di 14, 16, 24 o piu', ma proprio per la loro rigidita' sono particolarmente adatti ad un montaggio automatico su circuito stampato. Infine il SOT e' l'involucro utilizzato nelle tecnologie di montaggio superficiale. Esistono ovviamente numerosi altri involucri e al giorno d'oggi non e' raro trovare contenitori con un numero di piedini di collegamento che supera abbondantemente il centinaio

I parametri descritti presentano il difetto di un'eccessiva schematizzazione, con l'aggravante che ciascun costruttore tende ad enfatizzare gli aspetti positivi dei suoi prodotti e a minimizzarne quelli negativi.

Vi sono altri parametri oltre quelli descritti che potrebbero determinare la scelta di una famiglia logica, ad esempio la varieta' dei componenti della famiglia o la loro reperibilita' presso diverse fonti, la compatibilita' con i livelli di segnale o con le tensioni di alimentazione di altre famiglie o ancora la loro resistenza ad ambienti particolarmente sfavorevoli e/o aggressivi.

Infine come parametro di notevole interesse si puo' citare il prodotto della potenza dissipata per il tempo medio di propagazione. Anche se raramente riportato sulle caratteristiche, tale parametro misura in un certo qual modo la "bonta'" di una famiglia logica, che si puo' ritenere inversamente proporzionale al valore del prodotto stesso.

12.10) Immunita' al rumore.

La conoscenza del margine di rumore dei dispositivi digitali e' di estrema importanza per il progettista, sia allo scopo di determinare i segnali di ingresso e di uscita del sistema, sia per decidere se l'apparecchiatura realizzata debba o no essere schermata.

Cosa significhi effettivamente il margine di rumore in un sistema digitale e' facilmente spiegabile sulla base della caratteristica di trasferimento ingresso-uscita dei dispositivi usati.

Si consideri un semplice sistema reazionato costituito da due invertitori identici. Sul grafico di fig. 12.10.1 (b) sono riportate sia la caratteristica del dispositivo 1 che quella del dispositivo 2, con l'avvertenza tuttavia che il segnale di ingresso X_2 e' riportato sull'asse verticale, e il corrispondente segnale di uscita su quello orizzontale, al contrario pertanto di quanto fatto per il dispositivo 1.

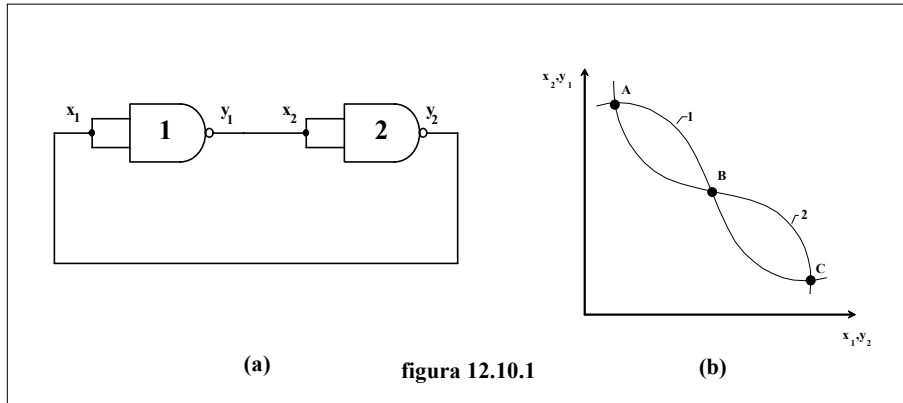
Tale modo di operare appare tuttavia naturale quando si consideri che il segnale di uscita di un gate coincide con il segnale di ingresso dell'altro gate. Le due caratteristiche si intersecano in tre punti, A, B e C che rappresentano le sole tre condizioni in cui il circuito puo' trovarsi; infatti solo in questi punti

$$X_2 = Y_1 \quad \text{e} \quad X_1 = Y_2$$

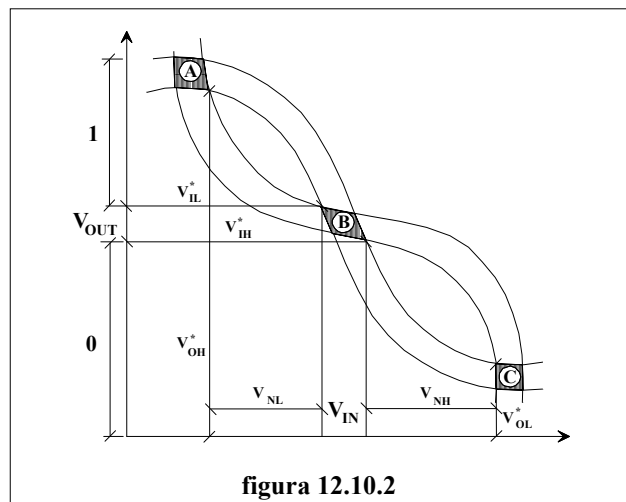
Se il dispositivo 1 si trova nel punto A allora l'ingresso e l'uscita del dispositivo 2 sono ancora specificati dal punto A.

Il sistema, a rigore, potrebbe trovarsi in uno qualsiasi dei tre punti; tuttavia la condizione di funzionamento nel punto B e' instabile e qualsiasi perturbazione, per quanto piccola, portera' il sistema in A o in C.

Nella realtà i due dispositivi non saranno mai identici. La caratteristica di trasferimento non potrà dunque essere rappresentata da un singola curva, ma sarà un involuppo contenente tutte le possibili caratteristiche. Le dimensioni dell'involuppo sono determinate dai limiti che vengono fissati per l'uscita, quando viene compiuto l'esame di accettazione in fase di produzione.



Generalmente viene scelta in modo arbitrario la massima tensione di uscita V_{OL} accettabile come valore logico 0. A tale tensione viene aggiunta una banda di guardia e la tensione totale ($V_{OL} +$ banda di guardia) viene applicata all'ingresso. La tensione di uscita che se ne ricava è chiamata V_{OH} , cioè la più bassa tensione possibile accettabile come 1 logico. Questi due parametri V_{OL} e V_{OH} vengono utilizzati per scegliere i dispositivi in sede di produzione e determinano l'involuppo delle caratteristiche di trasferimento. In fig. 12.10. 2 sono riportati due di questi involuppi per il semplice sistema reazionato che è stato preso in considerazione. Le tensioni rappresentative del punto di lavoro di ciascun dispositivo possono in questo caso cadere nelle tre aree tratteggiate A, B e C, intersezioni tra i due involuppi. È evidente che anche in questo caso tutti i punti dell'area B rappresentano situazioni di instabilità.



L'ordinata del punto al vertice sinistro alto dell'area C rappresenta la massima tensione allo stato basso che un dispositivo può avere quando viene pilotato da un altro dispositivo della stessa famiglia. Tale tensione viene indicata con V_{OL}^* .

Un dispositivo pilotato con V_{OL}^* avra' come tensione di uscita V_{OH}^* quella rappresentata dall'ordinata del vertice destro basso dell'area A.

I due livelli cosi' definiti sono le uscite dei due dispositivi interconnessi, nell'ipotesi che durante la loro selezione si siano incontrati due dispositivi al limite estremo del campo di tollerabilità, in modo tale che il sistema reazionato lavori nelle peggiori condizioni possibili.

Per un qualsiasi motivo tuttavia, ad esempio per la presenza di un disturbo, il segnale di ingresso puo' spostarsi dalla regione A verso la regione B. Malgrado cio', purché non si entri nella regione B stessa, il segnale sara' ancora interpretato come segnale logico basso. In corrispondenza la tensione d'uscita tendera' al vertice opposto della regione B, ma non dara' luogo ad un cambio di stato del sistema. Al cessare della causa perturbante il sistema ritornera' nelle condizioni originarie.

I vertici opposti della regione B rappresentano pertanto i valori estremi della tensione di ingresso che non danno luogo ad una commutazione del sistema; tali valori vengono indicati con V_{IL}^* e V_{IH}^* e rappresentano rispettivamente il massimo valore di ingresso sicuramente riconosciuto come uno zero logico e il minimo valore riconosciuto come un uno logico. Il margine di rumore del sistema e' la distanza che intercorre tra le uscite dello stesso sistema supposto privo di rumore e le due tensioni limite di ingresso V_{IL}^* e V_{IH}^* .

Queste differenze, indicate in fig. 12.10.2 con V_{NH} e V_{NL} sono dette rispettivamente **margine di rumore allo stato alto** e **margine di rumore allo stato basso** e possono ovviamente essere differenti.

Si noti che gli involucri si ampliano assegnando specifiche di tolleranza meno pesanti; inoltre se la pendenza della zona di transizione diminuisce, la regione B si estende verso A e C, diminuendo i margini di rumore.

Sui dati caratteristici forniti dai costruttori sono assegnate le quantita' V_{OH} , V_{OL} , V_{IH} e V_{IL} che non coincidono perfettamente con le quantita' asteriscate fin qui introdotte. Si ricordi che per le quantita' asteriscate valgono le seguenti definizioni:

V_{OL}^*	Arbitrario
V_{IL}^*	$V_{OL}^* +$ la banda di guardia V_{NL}
V_{OH}^*	uscita minima allo stato alto quando l'ingresso e' V_{IL}^*
V_{IH}^*	$V_{OH}^* -$ la banda di guardia V_{NH}

Le grandezze fornite dai costruttori sono cautelative rispetto a quelle asteriscate; piu' precisamente si puo' dire che:

$$V_{IL} = V_{IL}^* \qquad V_{IH} = V_{IH}^*$$

mentre V_{OH} si trova sulla curva di trasferimento appena a destra e al di sotto della zona A e V_{OL} appena a sinistra e al di sopra della zona C.

I margini di rumore apparenti, calcolati come:

$$V_{NL} = V_{IL} - V_{OL} \qquad V_{NH} = V_{OH} - V_{IH}$$

sono pertanto inferiori a quelli reali, sia per i valori piu' pessimistici di V_{OL} e V_{OH} , sia perche' normalmente i dispositivi usati ben difficilmente lavoreranno nelle peggiori condizioni, identificate dai vertici della regione B e da quello inferiore e superiore rispettivamente della regione A e C.

12.11) Margine di rumore dinamico.

Nella discussione sul margine di rumore del paragrafo precedente si e' implicitamente assunto che la caratteristica di trasferimento di un gate dipenda unicamente dall'ampiezza del segnale di ingresso e non dalla frequenza o dalla durata dell'impulso di ingresso.

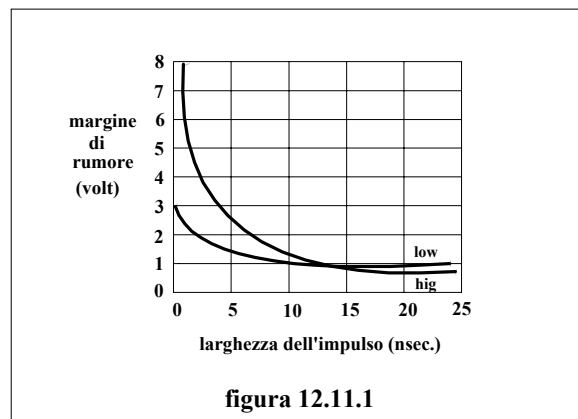
Si ricordi tuttavia che, come e' stato illustrato al paragrafo 12.7, un circuito digitale reale non puo' rispondere immediatamente ad una variazione del segnale di ingresso; inoltre diversi dispositivi hanno una caratteristica di trasferimento dotata di isteresi.

A causa di queste caratteristiche il margine di rumore dinamico, detto anche **margine di rumore in corrente alternata**, e' di solito ben diverso da quello in corrente continua.

In un circuito digitale correttamente progettato, che operi con impulsi di durata notevolmente superiore al tempo medio di propagazione dei vari gate e con una bassa frequenza di ripetizione, i margini di rumore in corrente continua e in corrente alternata non differiscono in maniera apprezzabile. Non appena pero' la larghezza dell'impulso diminuisce e si avvicina al tempo di propagazione, il margine di rumore cambia in funzione di come sono realizzati circuitualmente i vari gate. Se l'accoppiamento tra i vari stadi e' di tipo capacitivo, il margine di rumore puo' anche diminuire al diminuire della durata dell'impulso; tuttavia nella maggior parte dei casi il margine di rumore va via via aumentando, poiche' il circuito non e' in grado di rispondere istantaneamente a causa della carica accumulata nello spazio di base e a causa dei carichi esterni. A frequenze molto alte e con larghezze dell'impulso di ingresso molto piccole si puo' poi avere un effetto di polarizzazione del dispositivo, che puo' diminuire nuovamente il margine di rumore.

Caratteristica di notevole interesse in relazione al margine di rumore in corrente alternata e' infine la larghezza degli impulsi che il circuito stesso genera, agendo come generatore di rumore. Essi sono causati dai ritardi dei segnali che si propagano lungo diversi itinerari e dalle cadute resistive o reattive dovute alle impedenze distribuite del circuito. Questi disturbi autogenerati sono di solito molto stretti e con elevata frequenza di ripetizione.

In fig. 12.11.1 e' riportato l'andamento del margine di rumore in funzione della larghezza dell'impulso per dispositivi della famiglia TTL standard.



12.12) Caratteristiche dei flip-flop e dei registri.

Come si e' gia' accennato al capitolo V, il piu' semplice elemento di memoria puo' essere realizzato con due gate di tipo NAND o NOR connessi con accoppiamento incrociato. Molto spesso esiste poi anche un ingresso di abilitazione (enable) che molto impropriamente viene spesso chiamato clock.

Il collegamento incrociato di due gate da' luogo a un flip-flop che viene detto di tipo **trasparente**, in quanto l'uscita risponde immediatamente alle variazioni degli ingressi. In tal caso, riportando le uscite del dispositivo agli ingressi, si potrebbero addirittura generare condizioni di oscillazione.

Elementi **non trasparenti** sono viceversa i flip-flop master-slave, in cui la memorizzazione dell'informazione si ha durante il periodo di clock, mentre la variazione dell'uscita si ha nell'intervallo tra due successivi impulsi di clock, oppure i flip-flop azionati sui fronti.

A causa della non trasparenza l'uscita di questi dispositivi puo' essere riportata all'ingresso senza particolari precauzioni e senza dar luogo a condizioni pericolose. Non sono pertanto richiesti schemi di temporizzazione multifase e il progetto del sistema ne risulta notevolmente semplificato.

Per il corretto uso di questi dispositivi il progettista deve conoscere tuttavia l'esatto significato dei parametri di temporizzazione (tempo di ritardo e specifiche di temporizzazione dell'ingresso) e metterli in relazione al fenomeno del "clock skew".

12.12.1) Tempo di propagazione.

L'uscita di qualsiasi flip-flop o registro cambia unicamente in corrispondenza ad una delle transizioni del segnale di clock; nei circuiti piu' moderni tale transizione e' quella dallo stato basso allo stato alto.

Il ritardo tra questo fronte e la variazione corrispondente dell'uscita e' il tempo di ritardo del dispositivo e viene indicato con t_{pd} . Molto spesso tuttavia i tempi sono due, quello tra il fronte attivo del clock e la commutazione positiva dell'uscita, indicato con t_{pd+} o t_{pLH} , e quello tra il fronte attivo del clock e la commutazione negativa dell'uscita, indicato con t_{pd-} o t_{pHL} . Sui dati caratteristici sono di solito indicati sia il valore tipico che il valore massimo di tali parametri.

12.12.2) Specifiche di temporizzazione degli ingressi.

I flip-flop sensibili ai fronti sono normalmente condizionati dai valori che i segnali di ingresso possiedono immediatamente prima del fronte attivo del clock. Per un corretto funzionamento i dati di ingresso devono quindi essere presentati con un certo anticipo rispetto a questo fronte ed esiste pertanto un intervallo di tempo critico immediatamente precedente il clock, durante il quale essi non devono cambiare. L'istante nel quale il valore degli ingressi assume importanza viene detto tempo di "**set-up**", e' critico e viene di solito indicato con t_s . Poiche' tutti i parametri circuitali variano con la temperatura, le tensioni di alimentazione ed in funzione delle tolleranze di fabbricazione e del carico, anche il tempo di set-up ha una certa dispersione. E' importante assicurarsi che i livelli di ingresso rimangano stabili all'interno della finestra compresa tra il piu' lungo tempo di set-up, che rappresenta il piu' lento dei dispositivi e il piu' breve, che ne rappresenta il piu' veloce.

Purtroppo i costruttori hanno usato via via diverse nomenclature, talvolta in contrasto, per indicare i due bordi della finestra di temporizzazione appena definita. Il piu' lungo tempo di set-up viene spesso chiamato $t_{s(max)}$, ma spesso anche $t_{s(min)}$ in quanto rappresenta la durata minima dell'intervallo di tempo durante il quale gli ingressi devono rimanere costanti prima dell'impulso di clock. Il piu' breve tempo di set-up viene detto $t_{s(min)}$, ma spesso senza alcuna logica $t_{s(max)}$.

La Fairchild usa il termine "**tempo di disimpegno**" (release time t_r), altri fabbricanti lo chiamano - t_{hold} (tempo negativo di tenuta). Quest'ultimo nome prende origine dal fatto che nei circuiti piu' anziani il ritardo di propagazione del clock era superiore a quello degli

ingressi e di conseguenza questi ultimi dovevano rimanere stabili anche dopo il fronte attivo del clock (tempo di tenuta positivo). Tutti i circuiti moderni hanno tuttavia tempi di set-up positivi e di conseguenza tempi di tenuta negativi.

A differenza dei flip-flop comandati sui fronti, nei flip-flop master-slave a livelli il master accetta i dati di ingresso durante tutto il semiperiodo attivo del clock, mentre lo slave commuta in funzione del contenuto del master nell'intervallo tra un impulso di clock e l'altro. Pertanto in questi flip-flop per un corretto funzionamento e' indispensabile che gli ingressi rimangano stabili durante l'intero semiperiodo attivo del clock. E' opportuno quindi che la durata dell'impulso di clock sia la minima possibile, compatibilmente con i dati forniti dal costruttore.

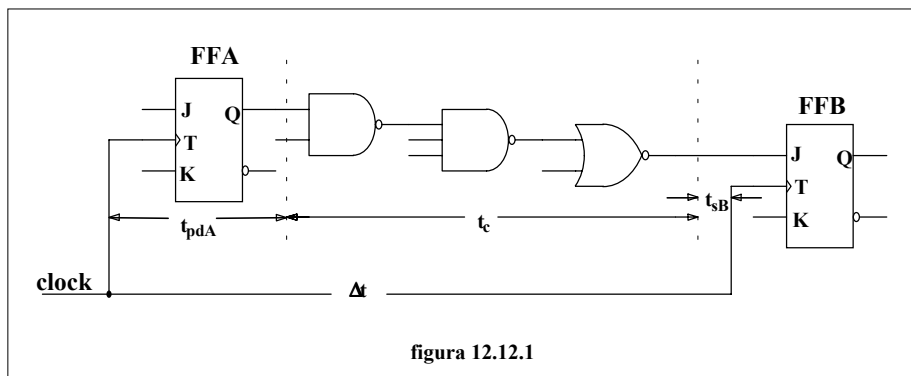
Qualsiasi sia poi il dispositivo, comandato a fronti o a livelli, e' necessario assicurarsi che i tempi di salita e di discesa non superino determinati valori. Infatti anche se i dispositivi non sono direttamente sensibili all'inclinazione dei fronti, se tali tempi fossero molto lunghi, il segnale di clock si troverebbe in prossimita' dei livelli di soglia per tempi superiori al ritardo di propagazione del flip-flop, e si potrebbero avere, in presenza di disturbi anche piccoli, determinati ad esempio da diafonia, delle commutazioni multiple.

E' opportuno pertanto che il segnale di clock abbia dei tempi di commutazione non superiori a 5 - 6 volte il ritardo di propagazione del dispositivo.

12.12.3) Il clock skew.

Il fenomeno del "clock skew", cui si e' gia' accennato in precedenza, e' proprio di sistemi digitali sincroni che utilizzino componenti ad alta velocita'. Poiche' tale fenomeno e' indipendente dalla frequenza di ripetizione del clock, ma e' originato solamente dalla velocita' dei componenti, esso puo' essere causa di spiacevoli sorprese anche in sistemi a bassa velocita' operativa. In teoria in un sistema sincrono tutti gli ingressi di clock dovrebbero essere attivati simultaneamente; in realta', ed in particolare nei sistemi di grande dimensione, cio' e' praticamente impossibile. Esistono sempre delle differenze temporali nella distribuzione del segnale, che determinano malposizionamenti del clock (clock skew); esse possono essere determinate da differenti tempi di propagazione sulle linee e/o da differenze nei tempi di ritardo dei vari gate che pilotano linee di clock diverse.

Si consideri, a titolo di esempio, il semplice caso illustrato in fig. 12.12.1, in cui un primo flip-flop FFA pilota un secondo flip-flop FFB attraverso una rete combinatoria. Il tempo di ritardo totale della rete combinatoria sia t_c .



Se l'impulso di clock giunge a FFB ritardato rispetto a FFA, per un funzionamento corretto e' necessario che il ritardo sia inferiore a:

$$\Delta t_{\max} = t_{pdA} + t_c + t_{sB}$$

Se esso e' maggiore il flip-flop FFB vede come ingresso la nuova uscita della rete combinatoria, che nel frattempo puo' essere variata, dando luogo ad un malfunzionamento. Si noti che in sede di progetto e' necessario usare per t_{pd} , t_c e t_s i valori minimi; purtroppo molto spesso questi non vengono specificati tra i dati caratteristici.

Vi sono due provvedimenti che si possono prendere per evitare i problemi relativi al clock skew. Il primo e' quello di usare l'accorgimento di pilotare tutti gli ingressi di clock da un'unica sorgente, eventualmente mettendo in parallelo piu' gate per aumentarne il fan-out.

Il secondo consiste nel distribuire il clock in maniera opportuna, in modo che i ritardi di propagazione del clock giochino in senso opposto al ritardo di propagazione dei dati. Si fa in modo cioe' che l'ultimo dispositivo della catena riceva il clock per primo.

Qualora cio' non fosse possibile, come ad esempio nei registri a circolazione (ring counter), si puo' inserire nel flusso dei dati un flip-flop con polarita' opposta.

12.13) Criteri di massima per la scelta della famiglia logica.

Nessuna famiglia logica e' ottima in rapporto a tutti i parametri che sono stati presi in considerazione nei paragrafi precedenti. Di solito anzi una famiglia e' ottima solo dal punto di vista di alcune sue caratteristiche, mentre e' carente in altre. Ad esempio l'aumento della velocita' operativa gioca un ruolo negativo nei confronti della dissipazione di potenza ed e' normale per qualsiasi circuito che lo sviluppo e il miglioramento di certe sue caratteristiche vadano a detrimento di altre.

In ciascuna famiglia logica pertanto vi e' una caratteristica predominante che tende ad essere utilizzata in quelle applicazioni in cui e' essenziale.

La mutua interazione tra le caratteristiche e' tipica di tutti i circuiti; storicamente i progettisti hanno dapprima tentato di ottimizzare le caratteristiche nel loro complesso (famiglie ad uso generale) e solo successivamente hanno sviluppato alcune caratteristiche a spese delle altre (famiglie per applicazioni particolari) oppure hanno sviluppato nuove famiglie basate su un circuito diverso o su una diversa tecnologia costruttiva.

La TTL standard, ad esempio, e' una famiglia ad uso generale, mentre la TTL low power o la TTL ad alta velocita' sono sottofamiglie in cui una delle caratteristiche della serie di uso generale e' stata ottimizzata.

Nel capitolo che segue verranno discusse le sei famiglie principali, illustrando in ciascun caso il circuito fondamentale della famiglia e descrivendone le principali caratteristiche.

Comunque sia, una volta che si sia deciso di utilizzare una determinata famiglia logica, la tendenza e' quella di mantenersi nell'ambito di tale famiglia anche per tutte le realizzazioni successive. Si ottengono in tal modo diversi vantaggi, quali ad esempio maggiori sconti per quantita', un miglior utilizzo delle apparecchiature di sviluppo e il non dover riaddestrare il personale con troppa frequenza. A tutto cio' poi si puo' aggiungere la considerazione che se una linea di prodotti viene largamente accettata, si genera una spinta verso l'espansione e il miglioramento di tale linea. Il costruttore tende cioe' ad offrire una maggior quantita' di servizi, quali letteratura in argomento e bollettini applicativi ed a produrre tutta una serie di sottosistemi standardizzati quali alimentatori e dispositivi ausiliari (display, convertitori A/D e D/A, elementi di memoria, ecc.).

In definitiva una famiglia logica di successo tende ad introdursi in modo definitivo sul mercato, autoperpetuando il proprio successo.

12.14) Cenni sulle tecnologie costruttive dei circuiti integrati.

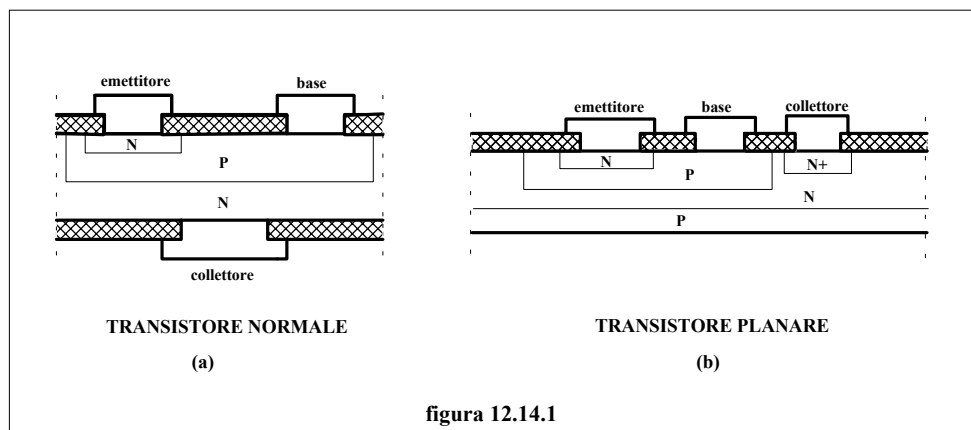
12.14.1) Circuiti integrati bipolari.

12.14.1.a) I transistori.

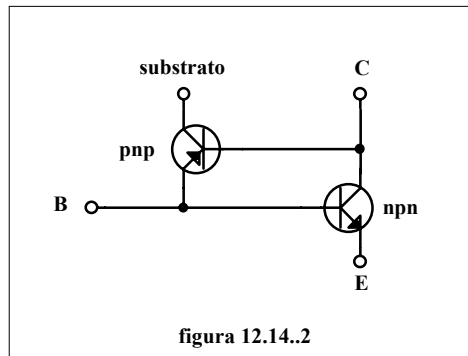
La tecnologia normalmente usata per la costruzione dei circuiti integrati monolitici bipolari e' quella planare epitassiale. Tale tecnologia si e' sviluppata negli anni 60 e si basa su due fatti fondamentali.

- 1) Sulla superficie di una fetta di silicio e' facilmente realizzabile uno strato di biossido di silicio, sia con reazioni di tipo "dry" in presenza di O_2 che di tipo "steam" in presenza di H_2O . L'ossido di silicio presenta ottime caratteristiche dielettriche ed assenza di difetti superficiali, specialmente quando venga realizzato in strati spessi 1 - 2 micron, anche se attualmente e' gia' notevolmente diffusa la tecnologia a 0,8 micron.
- 2) Il coefficiente di diffusione dei piu' comuni donatori (fosforo, arsenico, antimonio) e del piu' comune accettore (boro) e' di due o tre ordini di grandezza minore nel biossido che non nel silicio. Pertanto l'ossido puo' essere utilizzato come maschera per proteggere determinate zone dalla diffusione, rimuovendo con procedimenti fotolitografici l'ossido dove cio' e' necessario.

In fig. 12.14.1 (a) e' riportata la struttura tipica di un transistore planare epitassiale. Il transistore integrato differisce da quest'ultimo in quanto e' una struttura a quattro strati e tre giunzioni, come illustrato in fig. 12.14.1 (b). Lo strato in piu', rispetto alla normale struttura di un transistore, deriva dall'esigenza di costruire tutto il circuito su un substrato comune, che non e' presente nelle realizzazioni a componenti discreti.



Lo strato addizionale fa si' che il transistore integrato differisca da quello tradizionale per quanto riguarda i parametri parassiti. Si ha infatti, quale ulteriore elemento parassita, un transistore PNP, la cui base e il cui emettitore sono rispettivamente collettore e base del transistore integrato e il cui collettore e' il substrato. La struttura di un transistore integrato e' quindi, trascurando gli altri parametri parassiti, quella illustrata in fig. 12.14.2.

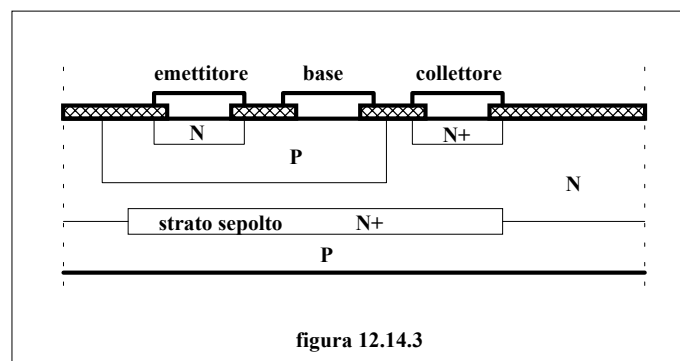


Il transistor parassita assume un ruolo diverso a seconda che si abbia a che fare con un circuito integrato lineare o con uno digitale. In un circuito digitale a logica saturata la tensione di collettore del transistor NPN è minore di quella di base. Pertanto la giunzione base-emettitore del transistor parassita PNP è polarizzata direttamente, mentre il suo collettore è connesso al punto più negativo del circuito. Il transistor parassita lavora quindi in zona attiva e di conseguenza un'elevata aliquota della corrente di pilotaggio del transistor NPN si perde nel substrato. Allo scopo di minimizzare tale inconveniente il guadagno di corrente del transistor parassita viene portato a valori approssimativamente pari a 0.01 con una diffusione addizionale di oro che ha lo scopo di ridurre drasticamente il tempo medio di vita delle cariche minoritarie.

Per i circuiti lineari invece il transistor parassita non dà un grande fastidio poiché il transistor NPN opera normalmente lontano dalla saturazione mantenendo quindi interdetti il transistor PNP.

In effetti la struttura illustrata per il transistor integrato non è la più adatta per gli impieghi nel campo della commutazione. Infatti, poiché il contatto di collettore deve essere ricavato sulla faccia superiore del wafer di silicio anziché su quella inferiore, la resistenza in serie al collettore è piuttosto alta e di conseguenza è alta la V_{CES} , tensione di saturazione collettore-emettitore.

Per ovviare a tale inconveniente si usa il cosiddetto procedimento dello "strato sepolto", che consiste nel formare nella zona N di collettore una zona a drogaggio più elevato e quindi di minor resistività, come illustrato in fig. 12.14.3.



Si può pertanto concludere che a causa della diffusione d'oro e della necessità di realizzare lo strato sepolto il costo di un transistor per impieghi digitali è più alto che non quello di un transistor per impieghi lineari.

12.14.1.b) I diodi.

Poiche' un diodo altro non e' se non una giunzione PN, da un transistor integrato si possono ricavare diodi in due modi diversi, utilizzando cioe' la giunzione base-emettitore oppure quella base-collettore.

Diodo base emettitore.

Con riferimento alla fig. 12.14.4 (a) appare evidente che il circuito equivalente del diodo e' quello di fig. 12.14.4 (b). C'e' la possibilita' di eliminare l'effetto del transistor parassita senza ricorrere alla diffusione d'oro semplicemente cortocircuitando tra di loro base e collettore, come illustrato in fig. 12.14.4 (c).

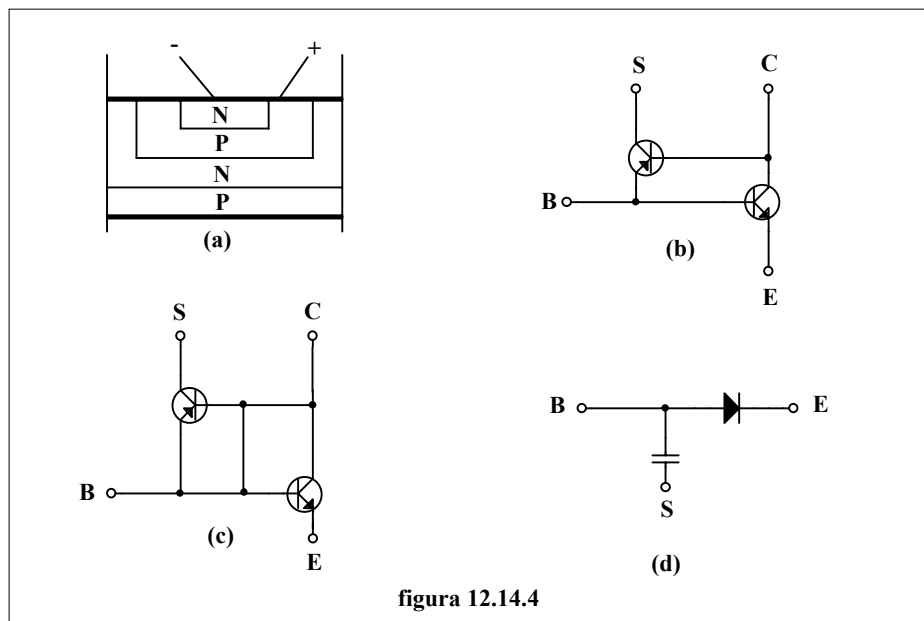


figura 12.14.4

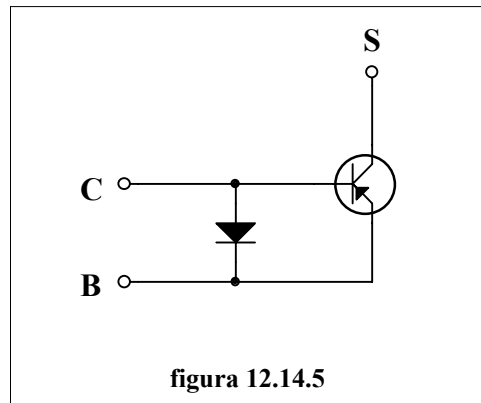
In tal caso la giunzione tra collettore e substrato risulta polarizzata inversamente e puo' essere considerata una semplice capacita' di basso valore. Il relativo circuito equivalente e' riportato in fig. 12.14.4 (d).

Diodo base collettore.

Qualora si utilizzi il diodo base collettore non vi e' alcun accorgimento circuitale che permetta di trascurare il transistor parassita ed e' giocoforza ricorrere alla diffusione addizionale di oro. Il circuito equivalente e' riportato in fig. 12.14.5.

Questo svantaggio tuttavia e' compensato dal fatto che la tensione di breakdown della giunzione di collettore e' piu' elevata di quella della giunzione di emettitore (30 - 40 volt contro 6 - 7 volt) a causa della diversa concentrazione delle impurita' nelle regioni di emettitore e di collettore.

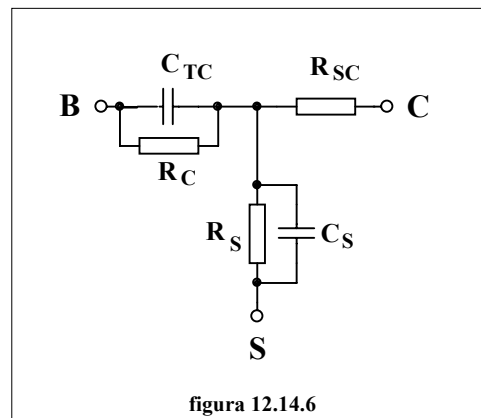
Il diodo base - collettore viene quindi usato, accompagnato da diffusione d'oro, dove sia necessaria una tensione di breakdown relativamente alta.



12.14.1.c) Capacita'.

In tecnica integrata i condensatori possono venir realizzati ricorrendo alle capacita' di transizione delle due giunzioni polarizzate inversamente. Tali capacita' sono ovviamente polari e dipendono dalla tensione applicata, che deve necessariamente essere minore della relativa tensione di breakdown. Il circuito equivalente e' analogo in ambedue i casi ed e' riportato in fig. 12.14.6. E' interessante notare l'esistenza di una capacita' C_s verso il substrato, in quanto anche il diodo collettore-substrato risulta polarizzato inversamente.

E' evidente che non e' possibile realizzare ne' capacita' troppo elevate, per l'eccessivo ingombro che ne deriverebbe, ne' troppo piccole, tali cioe' da essere trascurabili rispetto a C_s . Un valore abbastanza comune e' quello di 3000 pF/mm^2 con superfici utili di solito non superiori a $0.03 - 0.04 \text{ mm}^2$.



12.14.1.d) Resistori.

I resistori vengono ricavati sfruttando la resistivita' propria dei materiali semiconduttori e le loro dimensioni dipendono dalla purezza del materiale e dalla precisione di lavorazione. Quale materiale resistivo si usa di solito la regione di base, cioe' la regione P. In fig. 12.14.7 e' riportata la struttura di un resistore integrato; si vede immediatamente dal circuito equivalente che oltre al resistore si hanno anche due capacita' e un transistor parassita.

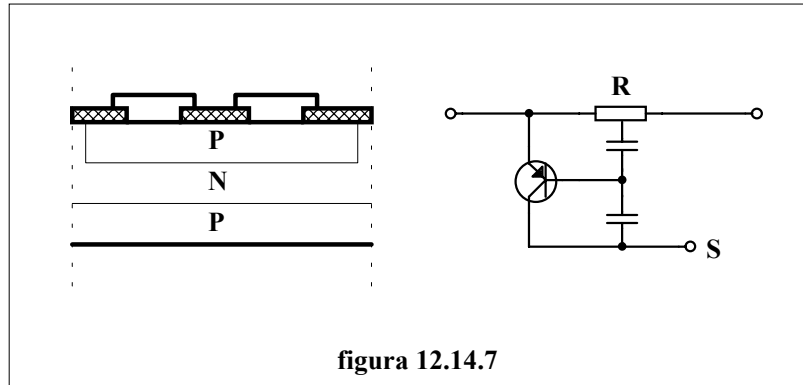


figura 12.14.7

In particolare a causa di quest'ultimo vi puo' essere una dispersione di corrente verso il substrato; per ridurre a valori trascurabili tale dispersione si connette normalmente la regione N al punto piu' positivo del circuito. Il circuito equivalente in tal caso si riduce unicamente a un resistore e ad una capacita' di valore molto modesto (circa 0.5 picofarad) verso lo strato N.

12.14.1.e) Suddivisione in isole.

Un circuito integrato completo e' realizzato in regioni, dette isole, di dimensioni relativamente grandi, che contengono uno o piu' componenti. Queste isole sono separate tra di loro da diffusioni profonde P ad alta concentrazione che vengono ad interrompere le regioni sottostanti N, come illustrato in fig. 12.14.8.

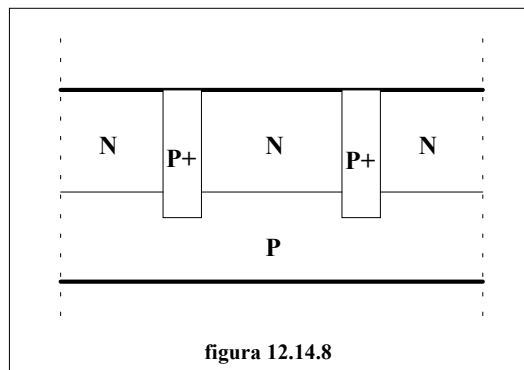


figura 12.14.8

In sostanza tali diffusioni P profonde sono necessarie ad isolare tra di loro quelle porzioni della regione N che non si trovano alla stessa tensione.

12.14.1.f) Processo di fabbricazione.

Il circuito integrato viene ricavato da una lastrina di silicio di tipo P, il cui spessore e' mediamente 200 micron, levigata con mezzi chimici. Da una lastrina si ricava un certo numero di circuiti completi con i seguenti passi di lavorazione:

- 1) Con processi termici viene formato sulla lastrina uno strato di biossido di silicio SiO_2 .
- 2) Sullo strato di biossido di silicio viene depositata una pellicola di materiale fotosensibile e dopo aver disposto sulla lastrina una prima

maschera si impressiona il materiale fotosensibile a luce ultravioletta. Con il successivo sviluppo le parti non impressionate vengono asportate.

- 3) Mediante azione chimica lo strato di SiO_2 rimasto scoperto viene asportato.
- 4) Rimossa la pellicola fotosensibile, dopo aver ripulito tutta la superficie, si diffonde arsenico nelle zone non protette dal biossido, formando in tal modo gli strati sepolti.
- 5) Dopo aver eliminato tutto lo strato di biossido, si fa crescere con procedimento epitassiale, uno strato superficiale di semiconduttore N, allo stesso grado di purezza del substrato.

I passi successivi del processo sono molto simili e consistono in:

- 6) Pulitura ed ossidazione superficiale.
- 7) Deposizione di materiale fotosensibile.
- 8) Suddivisione, mediante mascheratura, esposizione ad ultravioletti, sviluppo e attacco chimico, della superficie in isole.
- 9) Diffusione profonda di boro per ottenere la suddivisione in isole.
- 10) Pulitura ed ossidazione superficiale.
- 11) Incisione, come ai punti 7 e 8, delle aperture necessarie alla formazione delle basi e delle resistenze.
- 12) Diffusione di boro per le basi e le resistenze.
- 13) Pulitura ed ossidazione superficiale.
- 14) Incisione, come ai punti 7 e 8, delle aperture necessarie alla formazione degli emettitori e dei collettori.
- 15) Diffusione degli emettitori e delle zone N^+ dei collettori.
- 16) Deposizione di uno strato uniforme di alluminio sulla lastrina.
- 17) Fotoincisione dell'alluminio in modo da formare i collegamenti tra le isole ed i punti di connessione con i reofori.
- 18) Taglio della lastrina per ricavarne i singoli circuiti.

12.14.2) Circuiti integrati unipolari.

I transistori MOS vengono chiamati dispositivi unipolari poiché la corrente che durante il loro funzionamento attraversa la regione di canale è costituita da un unico tipo di portatori, a differenza di quanto accade nel funzionamento dei transistori a giunzione. Nel caso di transistori MOS a canale N i portatori sono elettroni, mentre nei MOS a canale P sono lacune.

I circuiti integrati unipolari possono contenere sia transistori di un solo tipo, sia di ambedue i tipi. In quest'ultimo caso vengono chiamati circuiti MOS complementari o CMOS.

È interessante notare che le tecnologie che usano un solo tipo di transistori unipolare (NMOS, PMOS) sono competitive solo se usate a larga scala di integrazione. Infatti i circuiti

MOS a singolo canale, quando utilizzati per la realizzazione di gate individuali, offrono scarsi vantaggi rispetto ai circuiti bipolari; sono infatti notevolmente più lenti di questi ultimi e soprattutto lavorano a tensioni e con valori di impedenza notevolmente diversi. È necessario pertanto interfacciare sia i segnali di ingresso che i segnali di uscita per renderli compatibili con i circuiti bipolari; il costo dei circuiti di interfaccia diviene accettabile solo per circuiti a larga scala.

Di conseguenza dei tre tipi di circuito, NMOS, PMOS e CMOS solo gli ultimi vengono in pratica utilizzati per realizzare gate e flip - flop.

Esistono diverse modalità realizzative dei MOS utilizzati. Per quanto riguarda i gate essi possono venir realizzati con strato metallico o con silicio policristallino. Poi le tecniche di diffusione possono essere diverse. Una delle più affermate, soprattutto per le precisioni raggiungibili nei parametri degli elementi prodotti, è quella di impianto ionico. Infine anche il substrato, normalmente silicio, può essere realizzato in zaffiro in modo da ridurre le scabrosità superficiali, aumentando la precisione di lavorazione.

12.14.2.a) Tecnologia di impianto ionico.

L'impianto ionico consiste nel drogare le fette di silicio bombardandole con ioni accelerati ad alta velocità in modo che essi penetrino in profondità. Tale metodo si distacca completamente dai sistemi di diffusione a caldo e permette un'estrema controllabilità del processo fino a livelli di drogaggio molto bassi, in modo da ottenere un'alta uniformità. In realtà la tecnologia di impianto ionico non è sostitutiva dei metodi classici di drogaggio, ma un sistema "**a freddo**" completamente diverso, che consente di ottenere prestazioni non raggiungibili con i sistemi a caldo.

L'apparecchiatura utilizzata consiste in cinque parti principali. Vi è la sorgente di ioni costituita da una regione in cui si forma il gas da cui devono essere tratte le particelle di drogante; gli ioni sono ottenuti per impatto di elettroni sul gas da ionizzare e la probabilità di collisione elettrone-gas viene aumentata ricorrendo a campi magnetici. In prossimità della sorgente vi è poi un acceleratore a potenziale variabile tra i 20 e i 200 o più kilovolt. All'uscita dell'acceleratore vi è uno spettrometro di massa che permette di selezionare solamente quegli atomi che si vogliono utilizzare come droganti.

Ad esempio, volendo utilizzare come drogante il boro, si partirà dal trifluoruro di boro e mediante lo spettrometro di massa si selezionerà il solo boro di peso atomico 11, scartando gli altri isotopi di boro e gli ioni di fluoro. Gli ioni che si ottengono, riuniti in un fascio ben focalizzato e dotati di velocità dell'ordine di 10^5 m/sec, sono in grado di penetrare abbastanza profondamente, qualche diecina di micron, nella fetta di silicio da drogare. Per ottenere un drogaggio uniforme, esiste infine un sistema di deflessione del fascio secondo due assi perpendicolari.

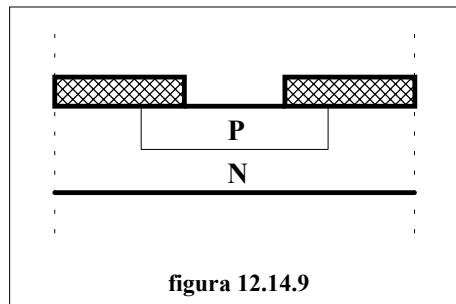
Il vantaggio fondamentale dell'impianto ionico risiede nel fatto che si possono regolare in modo indipendente lo spessore dello strato drogato, modulando tramite la tensione di accelerazione la velocità del fascio, e la concentrazione, sia controllando la durata del processo di impianto, sia modulando l'intensità del fascio. C'è poi da aggiungere che si può ottenere il drogaggio desiderato non solo in superficie, ma ad ogni profondità, e che il livello stesso del drogaggio può essere bassissimo.

Infine è importante far rilevare che la fetta di silicio può essere drogata anche attraverso uno strato di ossido, lasciando tuttavia la superficie nelle migliori condizioni.

12.14.2.b) Processo di fabbricazione.

Il materiale di partenza e' di solito una fetta di silicio orientata secondo gli assi cristallografici 100, di resistivita' compresa tra 1 e 2 ohm.cm. Su questo substrato possono essere realizzati direttamente transistori MOS di tipo P, mentre per quelli di tipo N e' necessario ottenere localmente un cambiamento della conduttivita' del materiale di base, ottenuta mediante la formazione di sacche P (P - well) nel substrato N.

Nelle tecnologie tradizionali, con drogaggio per diffusione, l'agente drogante, generalmente boro, raggiunge il silicio attraverso gli spazi aperti nello strato superficiale di ossido e viene diffuso ad una profondita' di qualche micron, in modo da portare la concentrazione superficiale al valore voluto, come illustrato in fig. 12.14.9.



Successivamente vengono aperte le fessure per la diffusione P+, con cui vengono formate le regioni di source e di drain dei transistori a canale P, le resistenze integrate e i cosiddetti "channel stoppers" dei transistori a canale N (fig.12.14.10).

