DIODI
e
TRANSISTOR

Capitolo 1

IL DIODO
E LE SUE APPLICAZIONI

1.1 *Semiconduttori*

I semiconduttori sono un vasto gruppo di sostanze con caratteristiche intermedie tra i materiali conduttori e quelli isolanti; la caratteristica che li rende interessanti per l'elettronica è il fatto che esiste modo di far sì che in alcune circostanze si comportino da isolanti ed in altre da conduttori. La cosa più importante è la possibilità di controllare (pilotare) tale cambiamento di caratteristiche.
I ***semiconduttori*** presentano le seguenti ***caratteristiche*:**

1. *la resistività è intermedia tra quella di un conduttore e quella di un isolante (ovvero* $10^5/10^4 \Omega \times m$ *10^5 e 10^4 Ω*m [tra 100.000 e 10.000 Ω per metro lineare]);*
2. *in tali materiali la resistività* ***diminuisce*** *con l'aumento della temperatura T (mentre nei conduttori* ***aumenta****!);*
3. *tali materiali presentano una grande sensibilità al "drogaggio"**:** aggiungendo al semiconduttore puro poche parti per milione (ppm) di sostanze estranee si genera una forte variazione della conduttività;*
4. *i semiconduttori presentano una grande efficienza nell'effetto fotoelettrico (generazione diretta di corrente in presenza di un segnale luminoso e generazione diretta di luce in presenza di una adeguata corrente);*
5. *presentano grande potere termoelettrico;*
6. *è infine possibile pilotare il cambio di comportamento da conduttore a isolante a mezzo di un adeguato campo elettrico.*

I semiconduttori "naturali" comprendono elementi inorganici ed organici, cristallini (solidi) e non cristallini (liquidi e gas). I semiconduttori più noti e più usati attualmente nelle applicazioni pratiche in elettronica sono il ***silicio*** (Si) e i1 ***germanio*** (Ge) accanto ad altri composti più rari, tutti ottenuti da elementi del ***III*** e del ***IV gruppo*** del sistema periodico (vedi ***CHIMICA***, ***Tavola di Mendelejev***, ovvero ***Tavola periodica degli elementi***).

1.2 *La teoria delle bande di energia*

La teoria delle bande di energia spiega i legami tra le diverse particelle che compongono gli atomi (protoni e neutroni [nucleo] ed elettroni); il nucleo è formato da protoni e neutroni ed è molto compatto. L'atomo di ogni elemento presenta un caratteristico numero di protoni (carica positiva) e di neutroni (carica neutra, ma ognuno con massa uguale al protone). Gli elettroni hanno una massa circa 1.800 volte inferiore a quella dei protoni e dei neutroni; ogni elettrone, tuttavia, possiede una carica elettrica negativa che equilibra esattamente quella (positiva) di un protone; gli elettroni ruotano attorno al loro nucleo come i satelliti ruotano attorno al loro sole; gli elettroni sono in numero uguale ai protoni (l'atomo è normalmente elettricamente neutro), ma sono "divisi" in orbite; quelli che occupano l'orbita più esterna sono quelli che possono (eventualmente) generare i *legami di valenza* (chimici) o possono venire "spostati", generando "ioni" e quindi corrente elettrica.

La forza che lega al nucleo gli elettroni, in particolare quelli che occupano l'orbita più esterna, non è uguale in tutti gli atomi. È questo che crea la differenza tra i materiali conduttori e quelli isolanti. In alcune sostanze (ed in particolare negli isolanti), gli elettroni formano tante piccole nubi di cariche negative, relativamente stabili, che avvolgono ciascun nucleo (caricato positivamente); in questi materiali non è praticamente possibile spostare elettroni; non è quindi possibile creare quel flusso di ioni negativi [elettroni] che è detta "corrente elettrica". In sostanze

come il legno, il vetro, la ceramica, il carbonio, l'aria, gli elettroni sono fortemente legati al nucleo e non possono pertanto circolare liberamente all'interno della materia. Tali sostanze sono dette *isolanti* perché, in condizioni normali, in essi non è possibile innescare una corrente elettrica.

In altre sostanze, definite conduttori (ed in particolare nei metalli), gli elettroni sono debolmente legati al nucleo; in particolare, gli elettroni dell'orbita più esterna, essendo dotati di notevole energia, durante il loro moto (dovuto all'agitazione termica, che è funzione della temperatura) possono "sganciarsi" e muoversi liberamente, in modo disordinato, passando da un atomo all'altro. Invece di formare tante piccole nubi negative che avvolgono ciascun nucleo, nei metalli, questi elettroni "liberi" formano un'unica nube negativa che avvolge l'intelaiatura cristallina del metallo, formata dalla parte interna degli atomi, cioè da ioni positivi; è pertanto impossibile capire a quale atomo appartenga ogni singolo elettrone. Questa situazione prende il nome di legame metallico. È questa la ragione per cui i materiali così costituiti sono buoni conduttori di elettricità.

Ricordiamo di nuovo che nelle sostanze dette isolanti come ad esempio il legno, il vetro, la ceramica, il carbonio, l'aria, gli elettroni risultano molto più fortemente legati al nucleo e pertanto non possono circolare liberamente all'interno della materia.

La teoria delle bande di energia (messa a punto da A.H.Wilson nel 1931 e che qui ovviamente non si intende approfondire) dimostra che semiconduttori ed isolanti hanno una medesima struttura, Tuttavia presentano un'importante differenza**:** quando gli atomi di una sostanza si dispongono in una struttura ordinata che si ripete indefinitamente, dando luogo ad un cristallo, i livelli energetici degli elettroni più esterni interagiscono e se ne formano di nuovi, molto vicini tra loro, raggruppati in fasce, o bande, di energia, separate da intervalli proibiti chiamati bande proibite. La loro distanza è corrispondente all'energia necessaria ad un elettrone per saltare da una banda all'altra.

Allo zero assoluto (0°K), in un isolante e in un semiconduttore gli elettroni più esterni sono legati al nucleo e occupano la banda di valenza, cosiddetta perché con queste energie gli elettroni esterni assicurano i legami chimici con gli altri atomi contigui (come meglio si dirà tra poco). Al di sopra di questa temperatura, la banda di conduzione è completamente vuota. Le due bande sono separate da una banda di livelli proibiti.

La grande differenza tra isolanti e semiconduttori sta in questo: *la banda proibita nei semiconduttori è molto più piccola che negli isolanti*. Ciò significa che a temperature superiori allo zero assoluto, ***nei semiconduttori*** un certo numero di elettroni ***passa effettivamente dalla banda di valenza a quella di conduzione***.

Ciò spiega la particolare dipendenza della resistività dei semiconduttori dalla temperatura; un aumento di temperatura implica un aumento di elettroni liberi e pertanto una diminuzione della resistività (al contrario di quanto accade nei metalli). Nei metalli, ove non esiste la banda proibita, gli elettroni esterni sono anche elettroni di conduzione; nei metalli, quindi, l'esistenza dei portatori di carica elettrica non dipende dalla temperatura.

1.3 *Silicio e germanio*

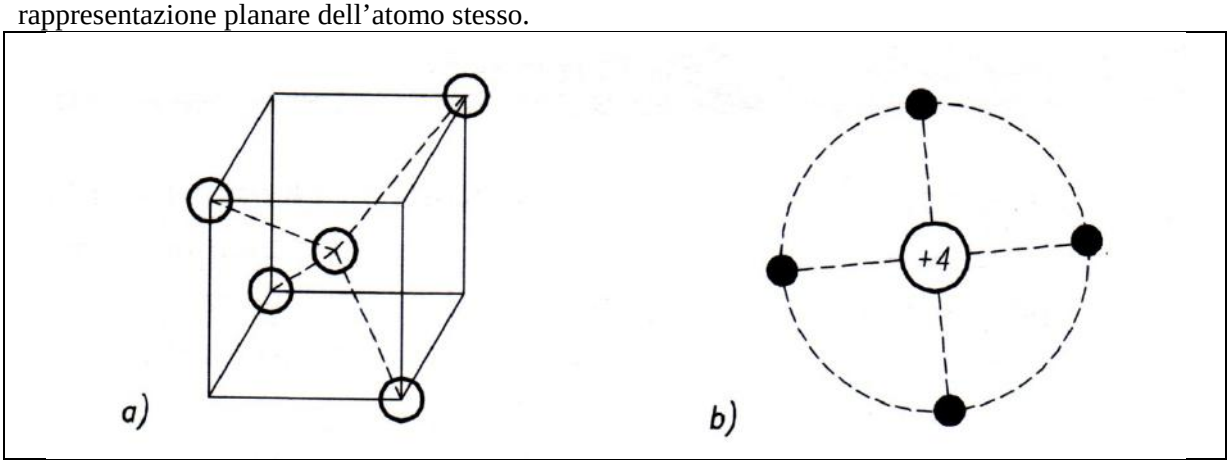
I semiconduttori sono tutti costituiti da *elementi tetravalenti* (***TETRA*** vuol dire quattro in greco), così chiamati perché ogni atomo di questi elementi è stabile quando condivide con gli atomi vicini *quattro elettroni*, detti appunto elettroni *di valenza*. Importanti nelle applicazioni pratiche sono soprattutto il silicio (scoperto nei 1824) ed il germanio (scoperto nel 1886).

L'energia media di un elettrone a temperatura ambiente è di ***0,025 eV*** (elettronVolt). La ampiezza della banda proibita del silicio è pari a 1,1 V**.** Per diventare elettrone di conduzione, un elettrone di silicio deve acquistare un'energia 45 volte superiore a quella media. Nel caso del germanio, la banda proibita ha un'ampiezza inferiore, pari a ***0,72 V*** e pertanto presenta una conducibilità sensibilmente maggiore del silicio**:** a temperatura ambiente, la **resistività** del silicio è **100 milioni** di volte quella del rame, mentre quella del germanio è "solo" **30 milioni** di volte superiore. Il germanio presenta dunque caratteristiche migliori rispetto al silicio, ma è molto meno diffuso, trattandosi di un materiale molto più raro e quindi molto più costoso.

- **NOTA:** Gli isolanti presentano una banda proibita pari ad alcuni ***eV*** (il diamante: 5,33 ***eV***). Si tratta di un'energia enorme ed è pertanto estremamente improbabile la formazione di coppie elettroni-lacune (ovvero corrente elettrica) per via termica.

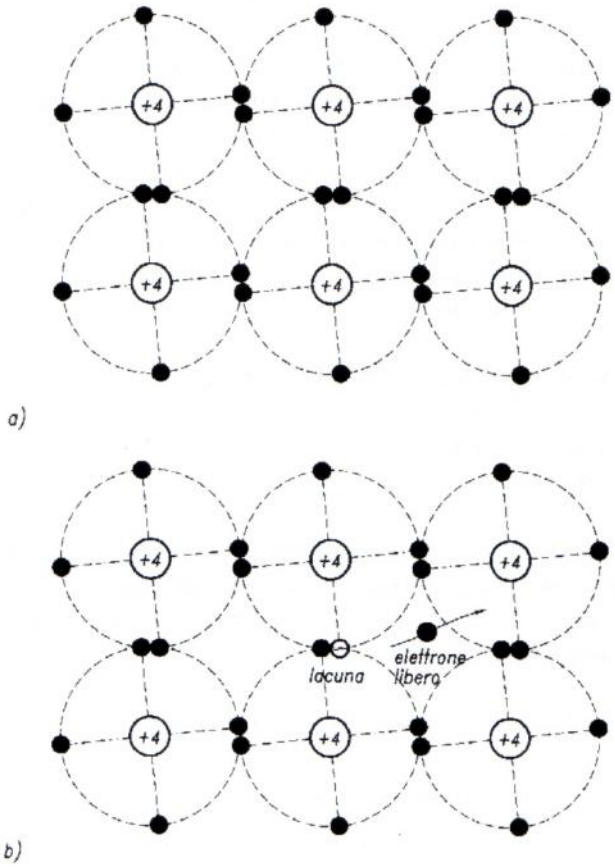
I semiconduttori, durante la solidificazione, cristallizzano in una struttura simile a quella del diamante, formando celle tetraedriche nelle quali ogni atomo si lega nel reticolo cristallino con altri quattro atomi dello stesso elemento.

La figura ***1a*** è una rappresentazione prospettica del cristallo, mentre la figura ***1b*** è una rappresentazione planare dell'atomo stesso.



- **NOTA:** La maggior parte dei materiali solidi con cui abbiamo a che fare sono in realtà aggregati di microcrístalli di dimensioni variabili (detti grani cristallini), la cui dimensione dipende dal processo di fabbricazione. In questo caso si parla di solidi policristallini. La rappresentazione di un materiale come cristallo ordinato, è comunque inadatta alle "alte" temperature; gli atomi di un qualunque cristallo, indipendentemente dal materiale, vibrano continuamente e molto velocemente attorno alla loro posizione di equilibrio; l'ampiezza di tale vibrazione è funzione della temperatura.

1.4 *Elettroni e lacune*



Il legame esistente in ogni cella elementare del cristallo riguarda un singolo atomo ed i quattro atomi vicini; è questa la cella elementare del cristallo; il legame tra gli atomi consiste nel mettere in comune una coppia di elettroni (doppietto elettronico).

É un po' come se ciascun atomo giocasse quattro partite a tennis contemporaneamente con i quattro atomi vicini, scambiando continuamente con ciascuno di essi due palline (cioè due elettroni). Questo tipo di legame tra atomi è detto legame covalente (Fig. ***2a***).

Quando, in seguito all'agitazione termica, un elettrone effettua una transizione dalla banda di valenza e quella di conduzione, lascia una lacuna nel legame covalente (ovvero una buca nella banda di valenza) (Fig. ***2b***).

Perdendo un elettrone, l' atomo (ma anche il cristallo) si squilibra elettricamente ed acquista una carica

positiva (ione +); tende quindi ad attirare le cariche negative più vicine.

Accade allora che un elettrone di un atomo vicino, pur non avendo energia sufficiente per diventare elettrone di conduzione, possa avere però l'energia sufficiente per spostarsi fino ad occupare 1a lacuna. Ma spostandosi, l' elettrone crea una nuova lacuna nell'atomo di provenienza, che a sua volta diventa positivo ed attira un altro elettrone vicino, e così via. Pertanto, ponendo il semiconduttore a temperatura superiore a quella ambiente ed entro un campo elettrico (ovvero collegandolo ai morsetti di un generatore di forza elettro motrice [f.e.m.] continua), mentre gli elettroni negativi si spostano dal polo negativo al polo positivo, le lacune positive si spostano nella direzione opposta, facilitando lo spostamento di altri elettroni non di conduzione. Il risultato complessivo è un'unica corrente elettronica dovuta ai due contributi (movimento di lacune e di elettroni).

- **NOTA IMPORTANTE:** Si vede dunque una seconda grande ***differenza*** tra i ***semiconcuttori*** ed i ***metalli***. Mentre in questi ultimi la conduzîone elettrica è solo elettronica, nei semiconduttori è dovuta sia agli elettroni liberi sia alle lacune, che nel loro moto creano un ulteriore flusso di elettroni che si va a sommare all'altro.

In un cristallo di silicio, la corrente elettrica è dovuta per il 75% al moto degli elettroni di conduzione e per il 25 % allo spostamento delle lacune.

1.5 *Il drogaggio*

A 100°C la conducibilità del silicio è circa 8 volte maggiore di quella a temperatura ambiente, che rimane comunque 100 milioni di volte inferiore a quella del rame nelle stesse condizioni. Oltre che con la temperatura, nei semiconduttori è possibile aumentare i portatori di carica liberi (e pertanto la conducibilità), ***introducendo delle impurezze all'interno della struttura cristallina*** (**che deve comunque essere prima ottenuta ad altissimo grado di purezza**). La presenza di queste impurezze riesce a produrre un “notevole” effetto "eccitante" sui meccanismi della conduzione elettrica. ***Un semiconduttore che ha subito questo trattamento si dice drogato***.

Supponiamo di inserire uniformemente nel semiconduttore piccole quantità di una sostanza pentavalente (ad esempio 1 atomo su ogni milione di atomi di silicio ***= 1 ppm***); si potrebbe utilizzare, ad esempio l'arsenico, oppure 1'antimonio o il fosforo. Questi atomi pentavalenti vanno a sostituire nella struttura cristallina altrettanti atomi tetravalenti del semiconduttore. Poiché l'arsenico tende a stabilire cinque legami con gli atomi vicini, mentre soltanto quattro dei suoi elettroni sono messi in comune con atomi contigui. ***I1 quinto elettrone resta libero di muoversi nel cristallo e costituisce pertanto un portatore di carica negativa che va pertanto ad aumentare la popolazione della banda di conduzione***. Gli elettroni liberi dell'arsenico alterano dunque il naturale equilibrio tra elettroni liberi e lacune libere del semiconduttore**:** gli elettroni diventano prevalenti sulle lacune. Gli elettroni liberi vagano disordinatamente nel cristallo drogato, assieme a quelli "naturali", prodottisi a causa dell’agitazione termica.

Un *semiconduttore* con queste caratteristiche è detto *di tipo n*, perché prevalgono i portatori di carica negativa.

Se si introduce invece una sostanza ***trivalente*** (ad esempio il boro, l'alluminio, il gallio o l'indio) si crea una lacuna di legame covalente. ***Le lacune prevalgono sugli elettroni*** e si parla di *semiconduttore di tipo p*, perché prevalgono i portatori di carica positiva.

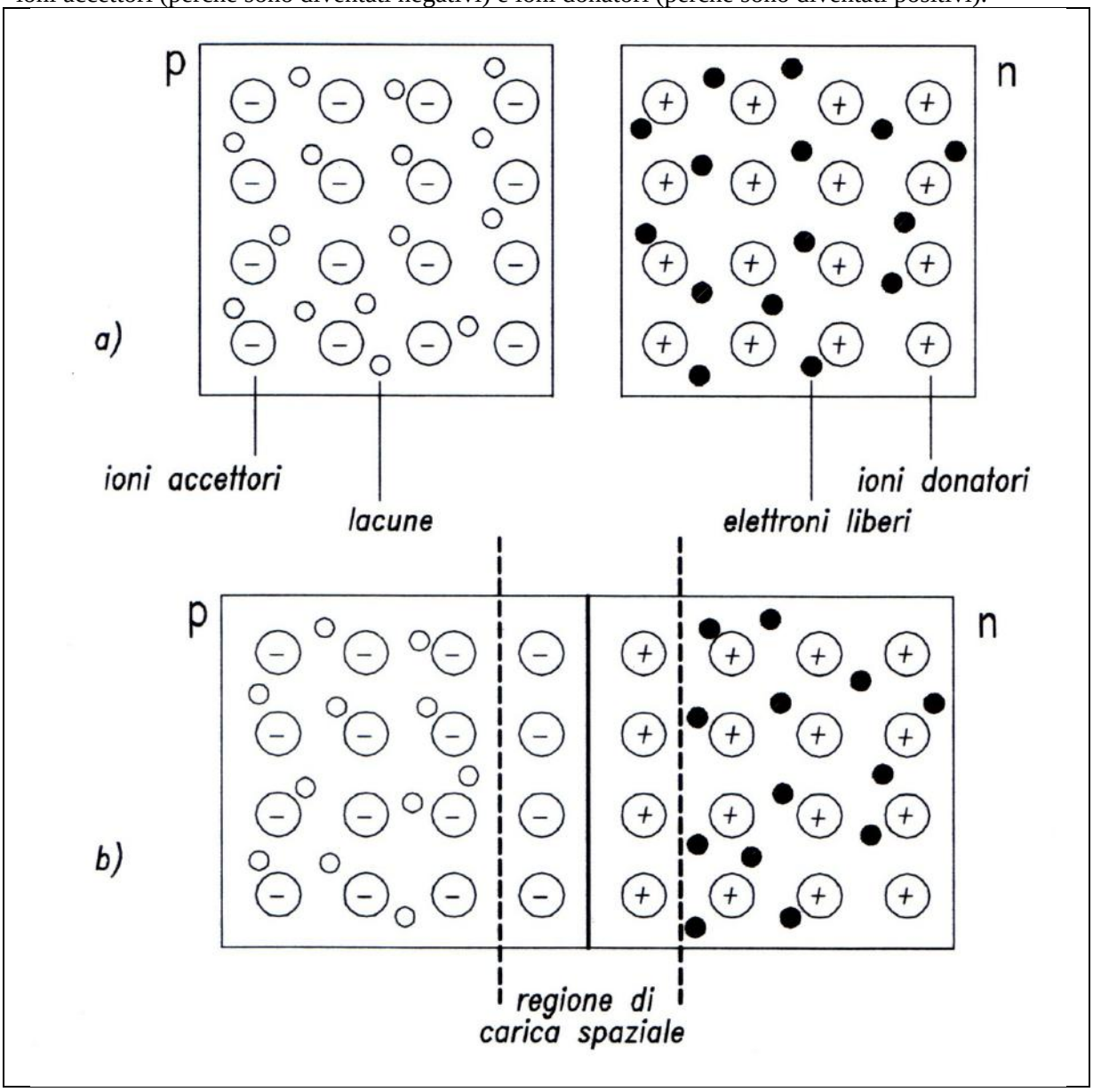
- **Nota:** In un semiconduttore drogato, il rapporto tra i portatori di carica dovuti al drogaggio (elettroni liberi) e quelli naturalmente presenti è nell’ordine di 10^5. Si dice pertanto che ***elettroni liberi*** (nel caso *n*) o ***lacune libere*** (nel caso *p*), in entrambi i casi dovuti al drogaggio, vengono definiti ***portatori di carica maggioritari***.

Un campo elettrico esterno introduce in questi movimenti caotici una piccola componente unidirezionale**:** gli elettroni si dirigono verso il polo positivo (***anodo***) e le lacune verso il polo negativo (***catodo***). Il movimento caotico dipende dalla temperatura e conferisce alle cariche velocità che arrivano a 106 m/s. La componente unidirezionale conferita da un campo elettrico esterno nei solidi è di qualche mm/s. Quindi i1 tentativo di moto ordinato e unidirezionale indotto da un campo elettrico deve fare i conti con un moto disordinato enormemente più intenso.

Negli studi avanzati sulla struttura della materia si parla di "***gas di elettroni***" perché essi manifestano proprietà tipiche dei gas**:** velocissimo moto casuale e tendenza ad occupare tutto lo spazio a disposizione (diffusione). In modo analogo, si potrà parlare di ***gas di lacune***.

1.6 La giunzione PN *non polarizzata*

Consideriamo due semiconduttori drogati in modo diverso (uno di *tipo p* e l'altro di *tipo n*). In Fig. ***3a*** sono riportati i portatori maggioritari nelle due zone e gli ioni cui sono dovuti, chiamati ioni accettori (perché sono diventati negativi) e ioni donatori (perché sono diventati positivi).



Se i due materiali sono posti a contatto in modo da formare una giunzione, la diversa concentrazione di portatori maggioritari di carica nei due materiali promuove una diffusione delle cariche dalla zona in cui sono maggioritarie nella zona in cui sono minoritarie, dove scompaiono ricombinandosi con cariche opposte.

Il risultato di questo processo è di creare una zona di ioni negativi senza la carica positiva vicino (o *ioni n*) nella *zona p*, e una zona di ioni positivi senza la carica negativa vicino (cioè *ioni p*) nella *zona n*. Questa regione è detta ***zona di carica spaziale***, o ***zona di svuotamento*** (*depletion layer*); dimensionalmente, il suo spessore è dell'ordine del micron.

Il campo elettrico (barriera di potenziale) che si viene a formare ha valori tipici**:**

0,2 V per il germanio, 0,6 V per il silicio.

È proprio questa differenza di potenziale ad impedire che, prima o poi tutti gli elettroni della zona *n* si diffondano nella zona *p* (e viceversa per le lacune). La barriera di potenziale ostacola ogni ulteriore arrivo di cariche maggioritarie ma favorisce il flusso di cariche minoritarie (lacune da *n* verso *p*, elettroni da *p* verso *n*) generate per agitazione termica (flusso chiamato **corrente di *drift***). A temperatura ambiente essa è normalmente trascurabile (dell'ordine di **qualche *mA*** [milli ampere] per il germanio e di **qualche *nA*** [nano ampere] per il silicio).

Il flusso dei portatori minoritari abbassa la barriera di potenziale, che innesca un nuovo flusso di portatori maggioritari, i quali però innalzano la barriera di potenziale. Si crea un sistema autoregolato in cui la barriera di potenziale si regola automaticamente in modo che la corrente dei portatori maggioritari sia esattamente uguale e contraria a quella dei portatori minoritari, per dar luogo ad una situazione di polarizzazione complessiva nulla.

La giunzione *PN* è anche chiamata *diodo a semiconduttore*.

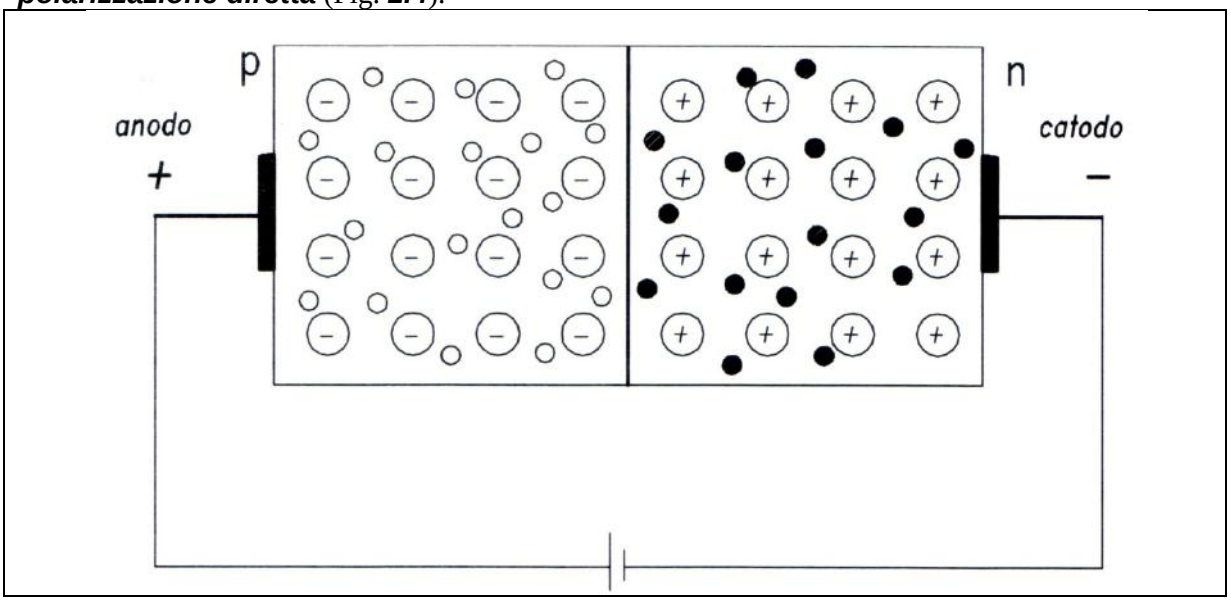
1.7 La giunzione PN *polarizzata*

Colleghiamo ora il polo positivo e quello negativo di un generatore di f.e.m. continue, rispettivamente al terminale chiamato ***catodo (-)*** del materiale di *tipo n* e al terminale chiamato ***anodo (+)*** del materiale di *tipo p*. Per valori non troppo elevati della f.e.m. della batteria, l'energia che questa conferisce alle cariche libere è sempre molto più piccola dell'energia termica che le fa muovere in senso caotico e che determina il meccanismo della diffusione. Pertanto, si crea un fenomeno di diffusione in senso contrario a quello prima descritto**:** i portatori maggioritari diffondono verso i poli del generatore, allargando la regione di carica spaziale e pertanto aumentando la barriera di potenziale.

In queste condizioni (**dette di *polarizzazione inversa***) attraverso la giunzione vi è solo il flusso dei portatori minoritari, di origine termica e non elettrica. La loro modestissima corrente tende velocemente alla saturazione e non aumenta aumentando la tensione inversa.

Detto in altri termini, la zona di giunzione si comporta come un condensatore, caricato il quale cessa ogni flusso di portatori maggioritari, e pertanto la corrente nell'intero circuito. Vi è solo la trascurabile corrente inversa.

Invertendo il collegamento della giunzione ai poli del generatore si ha la cosiddetta ***polarizzazione diretta*** (Fig. ***1.4***).



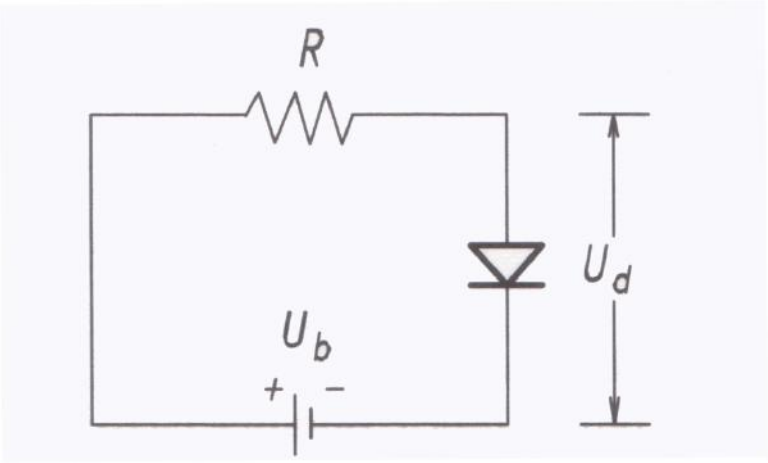
Se l'energia conferita dalla batteria è sufficiente, gli elettroni della *zona n* vicini alla zona di svuotamento riescono ad attraversare la barriera di potenziale entrando nella *zona p*; in contemporanea, le lacune della *zona p* riescono a penetrare nella *zona n*.

Mentre gli elettroni risalgono la differenza di potenziale si incontrano con lacune che la stanno discendendo. Quando un elettrone e una lacuna si trovano molto vicini il primo è catturato dalla seconda e i due portatori di carica si annullano. Ma questa perdita è compensata dall'afflusso

di nuovi elettroni e nuove lacune spinti all'interno della giunzione dal generatore, in modo tale che la corrente continua a fluire nel circuito.

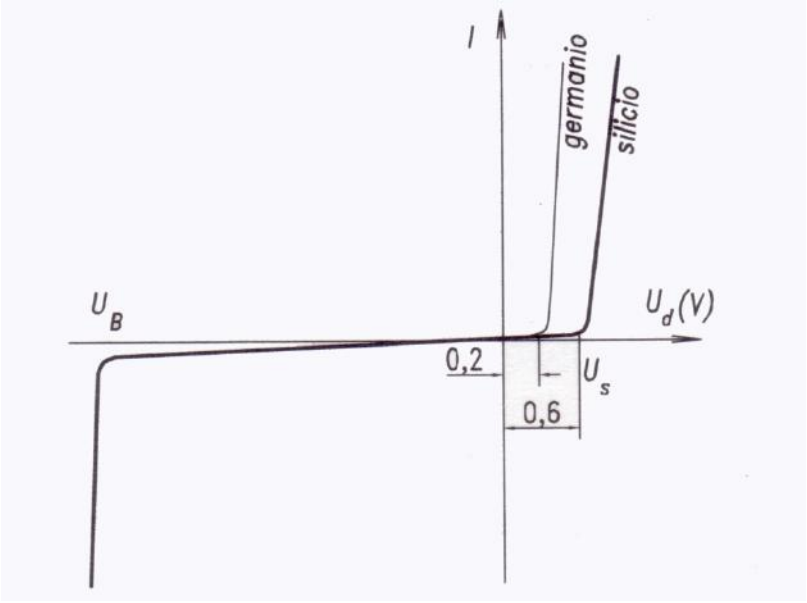
- **NOTA:** La *giunzione pn* (***diodo***) é dunque un ***dispositivo a senso unico*** che *consente il passaggio della corrente* (convenzionale) solo in un senso: **dal cristallo *p* al cristallo *n***.

Il diodo a semiconduttore svolge pertanto la funzione di un interruttore che è chiuso o aperto a seconda del segno della tensione applicata. Per uno abituato a ragionare in termini di pneumatica, è il corrispettivo elettrico della valvola unidirezionale**:** **se la corrente è continua**, il diodo le permette di circolare in un senso, nell'altro no.



Occorre, tuttavia, fare molta attenzione perché se la corrente è eccessiva (ovviamente in rapporto al particolare diodo impiegato), il componente si scalda rapidamente, raggiungendo temperature che ne provocano la fusione. Pertanto, per preservare la vita stessa del componente, occorre limitare le correnti in gioco; il sistema più semplice è di aggiungere al circuito diodo-generatore una resistenza in funzione di limitatrice di corrente.

Ne risulta il circuito di polarizzazione illustrato qui sopra (Fig.***1.5***), che riporta il simbolo normalmente impiegato negli schemi elettronici per il diodo.



La curva voltamperometrica (Fig. l6.6) mette in evidenza che nel diodo la relazione tra la corrente e la tensione diretta è non lineare. La curva è ben rappresentata dalla funzione di tipo esponenziale ***$U_d(I)$*** , detta ***equazione di Shockley***. Si nota immediatamente che la corrente sale in modo deciso solo dopo che la tensione diretta ha superato un caratteristico *valore di soglia*, che vale circa *0,2 V per il germanio* e *0,6 V per il silicio*.

⇦ Fig ***1.6***

Una volta innescata la conduzione, la corrente circolante non dipende tanto dalla tensione applicata al diodo, che rimane approssimativamente costante (ovvero pari a 0,2 o 0,6 V), quanto dalle condizioni del circuito esterno (batteria-resistenza).

Il diagramma mette inoltre in evidenza che in condizioni di polarizzazione inversa vi è una corrente inversa trascurabile, dovuta solo ai portatori minoritari. La cosa si presta a due considerazioni**:**

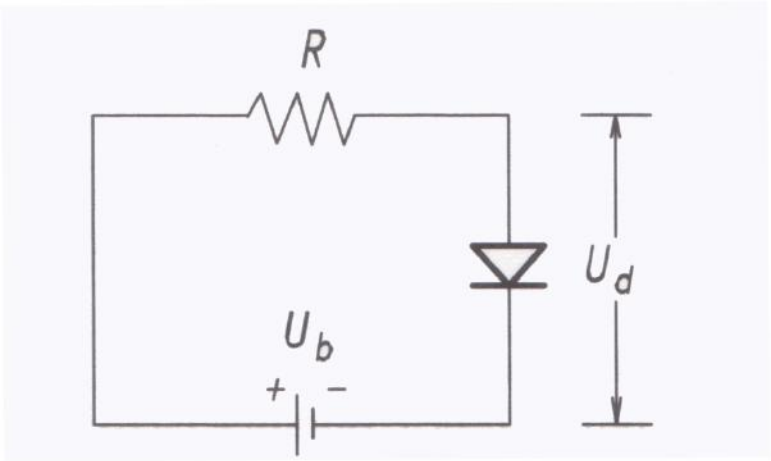
1) la corrente che attraversa un diodo sotto tensione è trascurabile, ma non è mai nulla;

2) se invece la tensione inversa raggiunge un valore alto e caratteristico (detto *tensione di breakdown*, che è dell'ordine delle centinaia di volt), la corrente inversa assume immediatamente valori elevatissimi , se non viene limitata da resistenze esterne. Si instaura un fenomeno che sarà meglio descritto più avanti.

1.8 Punto di lavoro del diodo

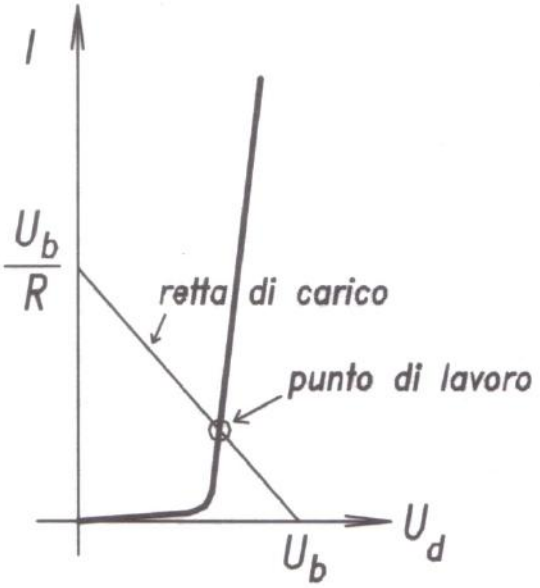
Le leggi di Kirchhoff, che regolano i circuiti elettrici, hanno validità del tutto generale e si applicano pertanto anche a maglie che contengono elementi non lineari come i diodi. Con riferimento al circuito di polarizzazione della Fig.***1.5*** (di nuovo riprodotta qui sotto), la legge delle maglie conduce alla equazione**:**

$$U_b - U_d = R \cdot I$$



Risolvendo il sistema con l'equazione ***Ud(1)*** si ottiene la coppia di valori (tensione sul diodo e corrente) che costituisce il cosiddetto punto di lavoro. Rispetto al caso di componenti tutti lineari, come le resistenze, la differenza è qui costituita dal fatto che non è disponibile una relazione lineare tra tensione del diodo e corrente che lo attraversa, come nel caso delle resistenze.

🡅 Fig. ***1.5*** / 🡇 Fig. ***1.7***



Il punto di lavoro del diodo potrà essere dunque determinato in modo agevole solo graficamente, come intersezione (Fig. 1.7) tra la curva voltamperometrica e la cosiddetta retta di carico data dall'equazione:

$$I = \frac{U_b}{R} - \frac{U_d}{R}$$

che è del tipo

$$y = a - bx$$

e che pertanto è rappresentata da una retta con angolo di inclinazione maggiore di 90°.

1.9.1 Effetto valanga

Riprendiamo il discorso introdotto nel paragrafo ***1.7*** a proposito di due distinti problemi**:** ***1)*** gli effetti distruttivi causati da una corrente eccessiva, ***2)*** collegando questo effetto a quello dell'impennata nell'intensità del passaggio di corrente generato da uno sconsiderato aumento di tensione ai capi del diodo. Come già detto, in condizioni di polarizzazione inversa, a temperatura ambiente, si genera solo un piccolissimo numero di coppie elettroni-lacune; attraverso la ***giunzione pn*** polarizzata inversamente si determinano piccolissimi valori di corrente inversa, tanto piccoli che di solito questa corrente viene, infatti, trascurata (per chiarezza rivedi la Fig. ***1.6***).

Tuttavia, se la tensione inversa raggiunge valori molto elevati, questi elettroni possono essere accelerati molto fortemente verso il polo positivo. Inevitabilmente essi urteranno contro altri elettroni che stanno nella banda di valenza, facendoli "saltare" nella banda di conduzione; ciascuno di questi, a sua volta, può urtare ed accelerare violentemente altri elettroni, strappandoli dalla banda di valenza. Si viene così a generare quasi istantaneamente un effetto valanga, che fa raggiungere alla corrente inversa di elettroni liberi valori enormi (guarda ancora la Fig. ***1.6***).

Il componente, che è già sottoposto a elevati valori di tensione, viene attraversato di colpo da una corrente di forte intensità; improvvisamente è costretto a dissipare una potenza altissima; è pertanto del tutto probabile che si scaldi fino ad andare fuori uso.

Il progettista che decide di impiegare un diodo deve tenere conto non solo della massima corrente diretta sopportabile dal diodo, ma anche della massima tensione inversa, per evitare che si verifichi la distruttiva valanga di elettroni. Il valore a cui si verifica questo fenomeno è detto tensione inversa di rottura (*breakdown reverse*), indicata dalla sigla U_{BR} .

Normalmente questa tensione vale alcune centinaia di volt. Ad esempio, un diodo molto usato ed economico, il *tipo **1N4007***; quando funziona in polarizzazione diretta sopporta una corrente massima di ***1 A*** ed ha un valore di U_{BR} pari a ***1000 V***.

1.9.2 Diodo Zener

Tuttavia anche un fenomeno nocivo come l'effetto valanga, se essere opportunamente controllato, può essere utilmente sfruttato.

Se si realizza una *giunzione **pn*** con materiali ad alto drogaggio e si polarizza inversamente questa giunzione, attorno ad essa si crea una stretta regione di carica spaziale con un campo molto elevato. In tal modo, anche un modesto valore della tensione inversa può far sì che una valanga di elettroni siano strappati dalla banda di valenza della *zona p*.

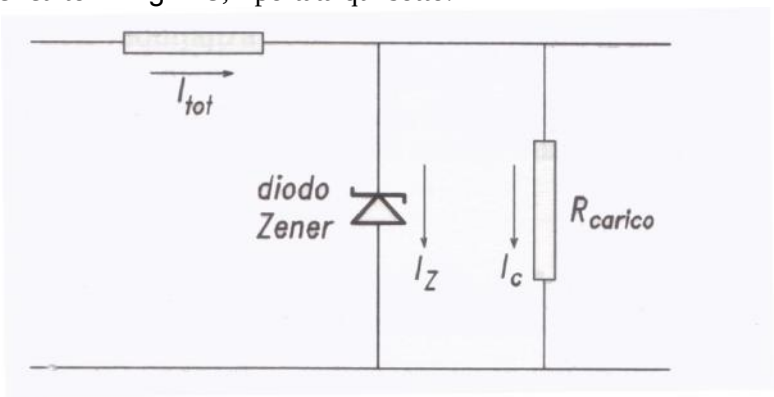
Questo fenomeno prende il nome di ***effetto Zener***. I diodi fabbricati per farlo avvenire in modo controllato si chiamano *diodi Zener*.

NOTA BENE: L' aspetto esteriore di un *diodo Zener* è identico a quello di un diodo normale; non solo**: il normale collaudo tramite tester non permette di distinguere un *diodo Zener* da uno normale; è necessario realizzare un circuito di prova per verificare se esista o meno conduzione in senso inverso.**

Per questi diodi la valanga di elettroni non è distruttiva perché la tensione è bassa e perché in serie è disposta un'opportuna resistenza limitatrice (come già accennato). **Quando si acquista un diodo Zener bisogna specificare la tensione U_{BR} e la potenza massima dissipabile**. In commercio esiste una gamma vastissima di *diodi Zener*, con tensioni U_{BR} da qualche volt a qualche decina di volt.

1.9.3 Lo Zener come stabilizzatore di tensione

Il diodo Zener trova larga applicazione come stabilizzatore di tensione. Consideriamo il circuito in Fig. ***1.8***, riportata qui sotto.



⇦ Fig. ***1.8***

Eventuali variazioni della resistenza di carico o della tensione di alimentazione faranno variare la corrente sul carico I_C e la corrente totale I_{TOT}, e conseguentemente la corrente I_Z derivata dallo Zener. Tuttavia, alla *tensione di breakdown* la tensione ai capi del diodo è pressoché indipendente dalla corrente che lo attraversa.

In tal modo si ottiene un effetto di stabilizzazione della tensione sul carico sia alla presenza delle variazioni del carico che della tensione di alimentazione. Poiché in ogni caso**:**

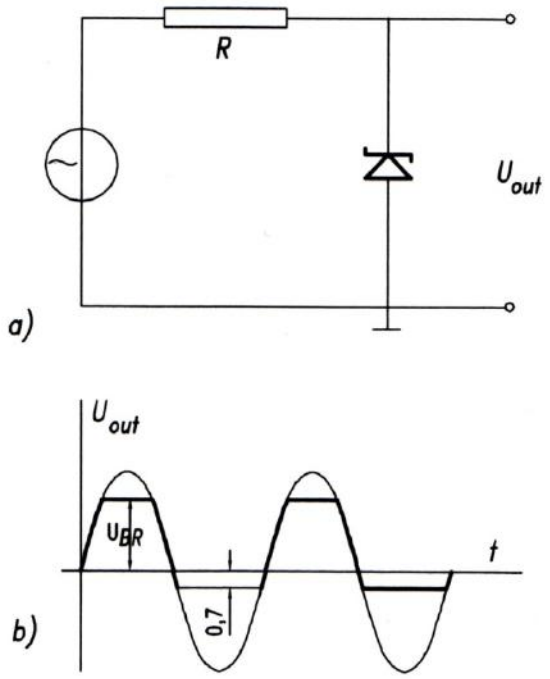
$$I_{TOT} = I_Z + I_C$$

Nel diagramma si vede come al variare delle correnti I_{TOT} e I_C, varia anche I_Z; dunque un *diodo Zener* si comporta come una resistenza automaticamente variabile, col pregio di mantenere in ogni caso costante la tensione ai capi.

1.9.4 *Lo Zener come limitatore di tensione*

Un'altra notevole applicazione del diodo Zener è come limitatore di tensione. In Fig. ***1.9a*** è illustrato un circuito con un generatore di tensione alternata avente una determinata ampiezza. Ad esso è collegato un circuito formato da una resistenza con in serie un *diodo Zener* avente l'*anodo* (polo **+**) collegato a massa. Ai capi di questo è prelevata la tensione di uscita. Se polarizzato in senso diretto un *diodo Zener* non differisce da uno normale; le semionde negative, invece, lo polarizzano inversamente, determinando ai suoi capi una tensione compresa tra *0,6* e *0,7 V*.

Quando polarizzato in senso inverso, per piccoli valori di tensione lo Zener non conduce, quindi nella resistenza non circola corrente, e perciò la tensione di uscita ricopia fedelmente la tensione del generatore. Raggiunto il valore di ***UBR***, il diodo conduce e stabilizza la tensione a tale valore. In uscita si trova dunque il segnale "tagliato" al valore *UBR* (ovvero della tensione ***inversa*** di rottura o *breakdown reverse*).



⇦ Fig. ***1.9a*** ed ***1.9b***

1.10 Diodo di libera circolazione

Consideriamo l'apertura di un circuito induttivo come quello illustrato in Fig. ***1.10a***.

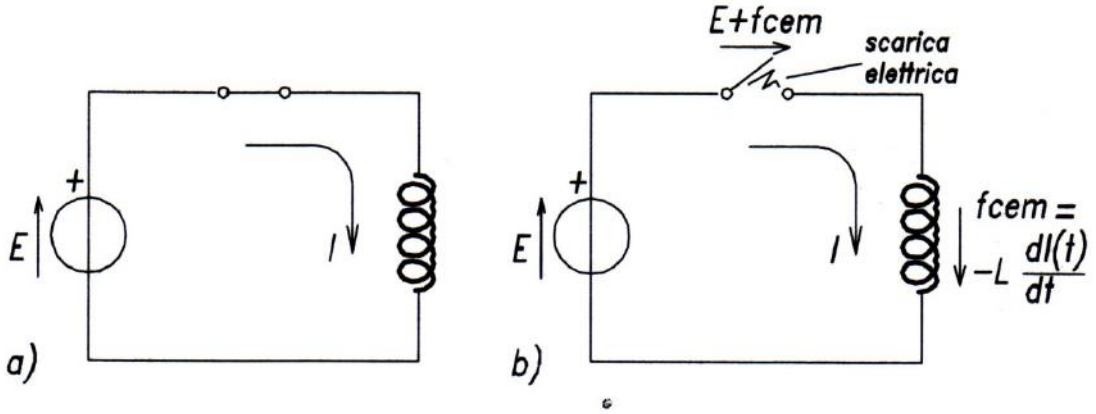


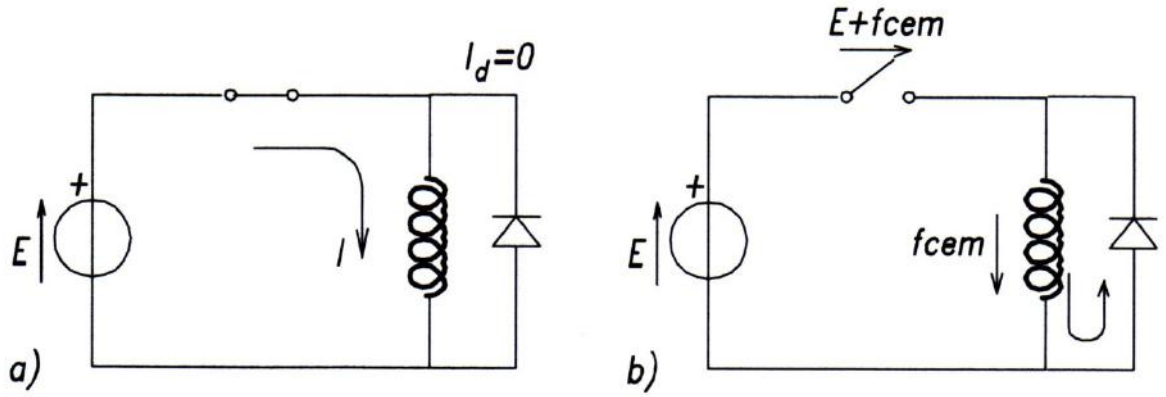
Fig. ***1.10a*** Fig. ***1.10b***

Quando si apre l'interruttore (spegnimento) la corrente che percorre l'induttanza tende ad estinguersi; si allora crea una ***forza controelettromotrice*** (***f. c. e. m.***) di valore elevato che tende a contrastare l'estinguersi della corrente.

Ai capi dell'interruttore si stabilisce pertanto una tensione ***E + f. c. e. m.*** così elevata da far scoccare un arco elettrico (Fig. ***1.10b***). In tal modo, la corrente può, anche se solo per un attimo,

continuare a fluire anche attraverso l'aria ionizzata; tuttavia, lo scoccare dell'arco dissipa l'energia accumulata nel campo magnetico dell'induttore. Ogni volta che si apre il circuito si ha una scarica che genera l'arco elettrico di cui sopra; nel tempo i contatti , dell'interruttore si danneggiano (tutti possono constatare gli evidenti segni di bruciature presenti sui contatti degli interruttori usati).

Un modo per eliminare il problema consiste nel collegare un diodo in parallelo (Fig. ***1.11a***). Nel normale funzionamento, il diodo è interdetto (è polarizzato inversamente, la corrente non lo attraversa), pertanto la sua presenza non si nota. All'apertura del circuito, la ***f. c. e. m***. che si forma ai capi dell'induttanza polarizza direttamente il diodo**:** in questo modo, la corrente del carico trova un

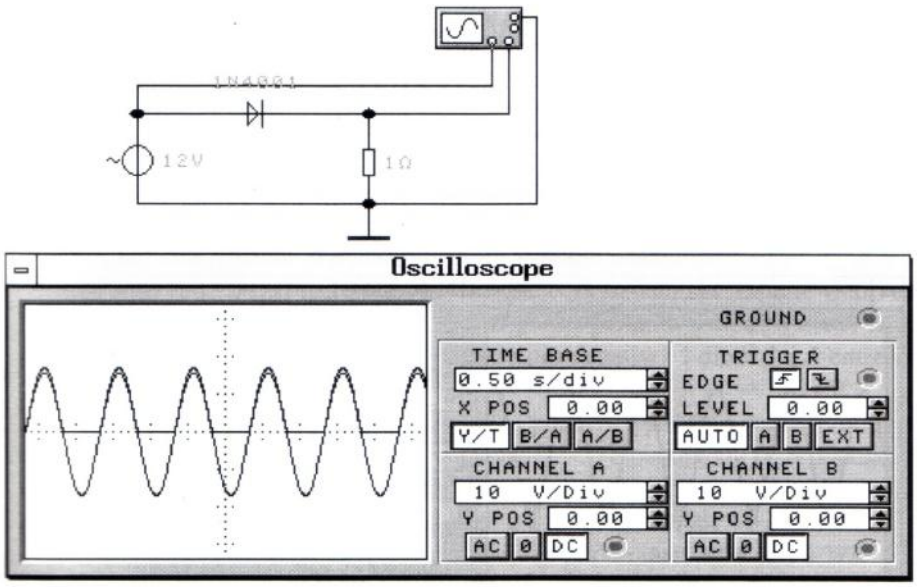


percorso alternativo, lungo il quale continua a circolare finché l'energia del campo magnetico non si estingue. L'arco elettrico non si forma. Ciò, tuttavia, implica che debba trascorrere un certo tempo, ben preciso, tra uno spegnimento e la successiva accensione; ***ciò significa che l'operazione di accensione/ spegnimento non può essere ripetuta con una rapidità inferiore al tempo di estinzione della corrente nel circuito alternativo in cui è montato il diodo***. Per diminuire tale ***costante di tempo*** (vedi), è possibile inserire nel ramo del diodo una resistenza di valore opportuno.

16.11 Raddrizzatori

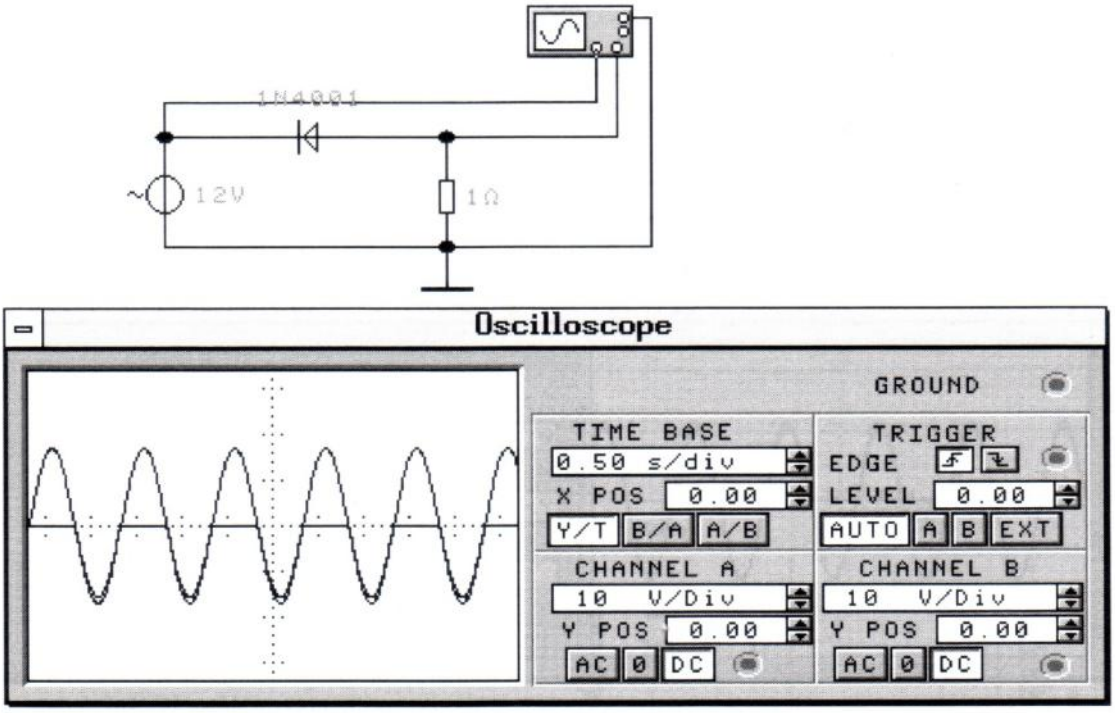
La funzione fondamentale di un diodo è di raddrizzare la corrente, cioè di farla scorrere sempre nello stesso verso.

In Fig.16.12 è riportato un semplice circuito, con un generatore di f.e.m. alternate collegato alla serie di una resistenza con un diodo.



Il generatore spinge la corrente in un verso e nell' altro nell' alternarsi delle semionde; il diodo lascia passare la corrente solo quando fluisce dall' anodo verso il catodo; in questo caso esso determina una caduta di tensione Ud dovuta alla barriera di potenziale formatasi nelle giunzione. Ai capi della resistenza si troveranno pertanto solo le semionde positive, lievemente attenuate, come si può costatare con un oscilloscopio a doppia traccia.

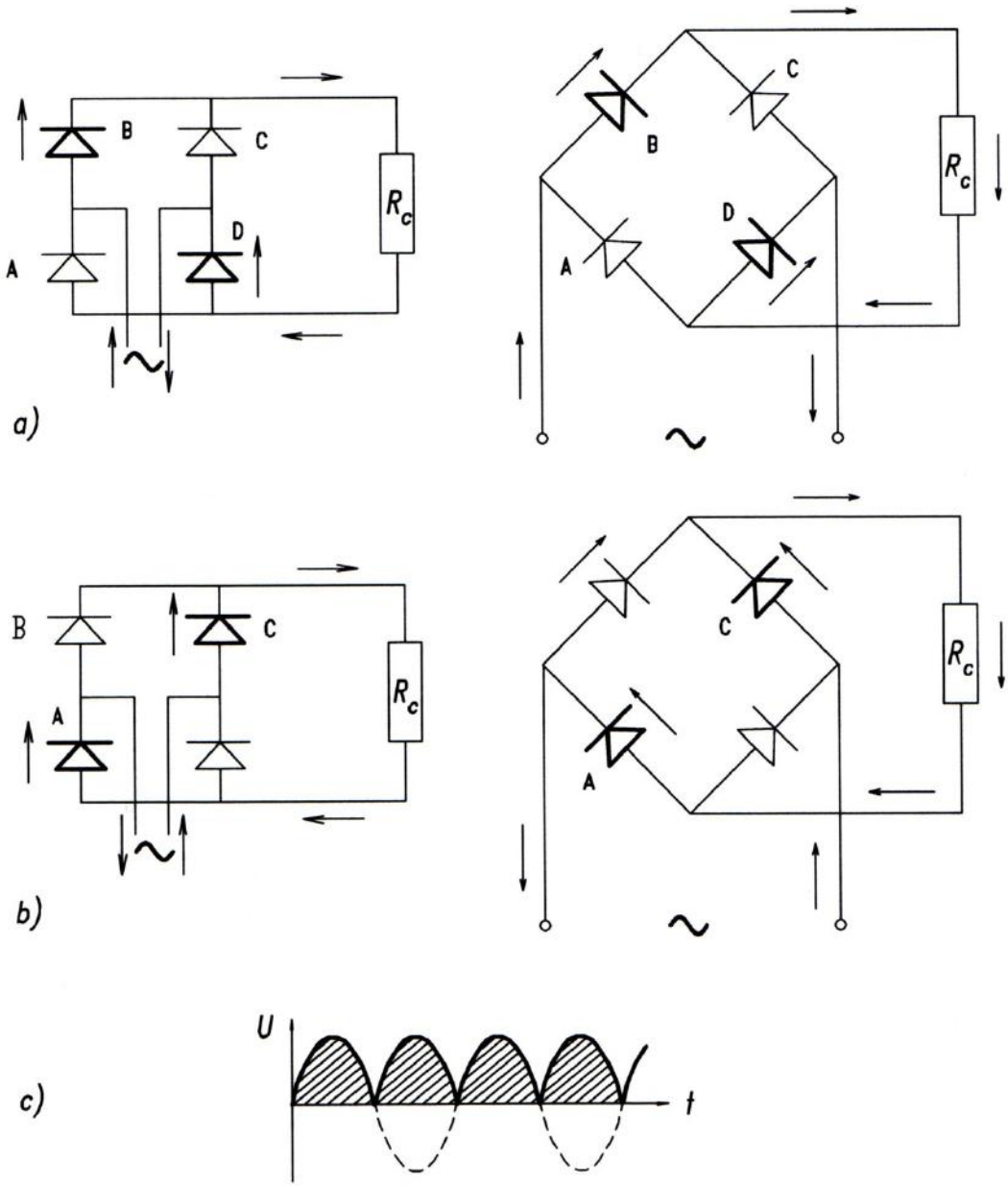
Se l'ampiezza della tensione alternata è di parecchi volt, l'effetto di attenuazione diventa trascurabile.
Invertendo le polarità del diodo, ai capi della resistenza si possono prelevare solo le semionde negative, come illustrato in Fig. 16.13.



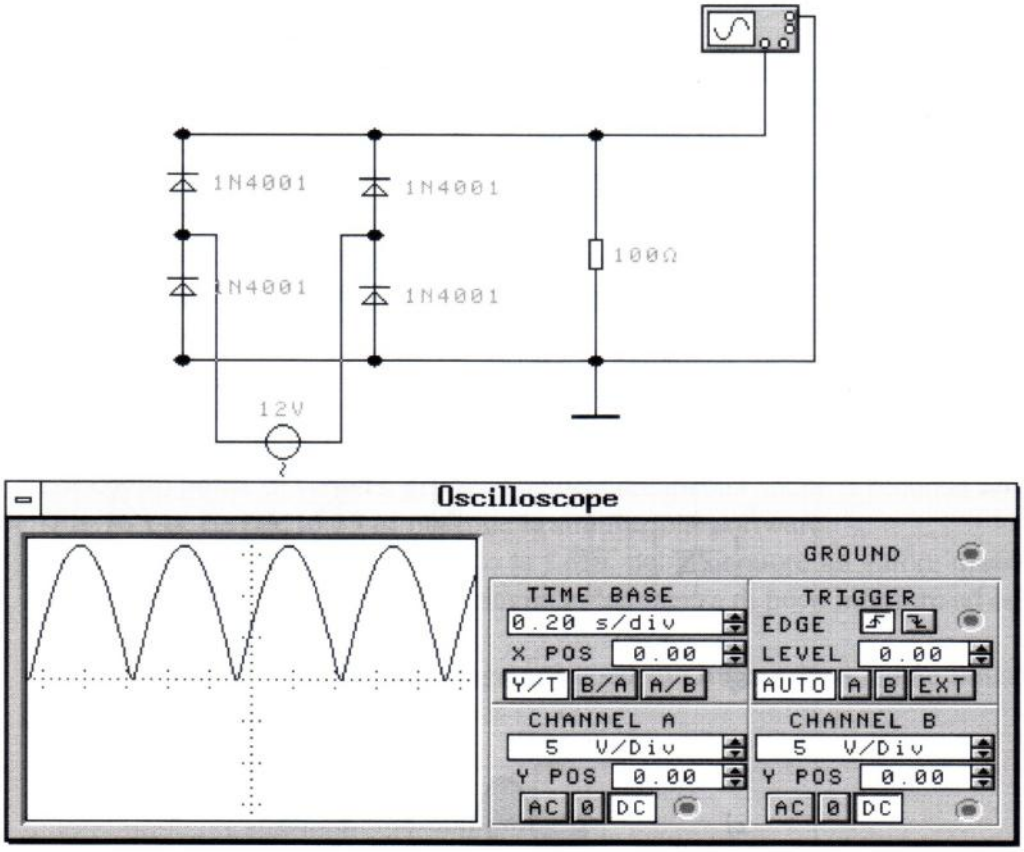
In entrambi i casi si ottiene sulla resistenza di carico una corrente niente affatto costante, ma unidirezionale, e ai capi della resistenza vi è sempre la stessa polarità. Il raddrizzamento può avvenire anche nel caso in cui la tensione del generatore non sia sinusoidale ma triangolare, a dente di sega ecc.

Il circuito descritto è molto semplice ma presenta l'inconveniente di lasciare inattivo il generatore durante la semionda eliminata.
Un circuito di raddrizzamento più conveniente è la disposizione di quattro diodi nel cosiddetto ponte di Graetz, grazie al quale è recuperata anche la seconda semionda (Fig.16.14).
La Fig. 16.15 ne propone la simulazione software.
Si noti, nella Fig. 16.15, che avendo la f.e.m. del generatore un valore massimo di 16,97 V ($12 \cdot \sqrt{2}$) la tensione di uscita ha un'ampiezza di poco inferiore ad esso.

Oltre i 10 kW di potenza si impiegano i ponti di Graetz trifasi, come illustrato in Fig. 16.16.



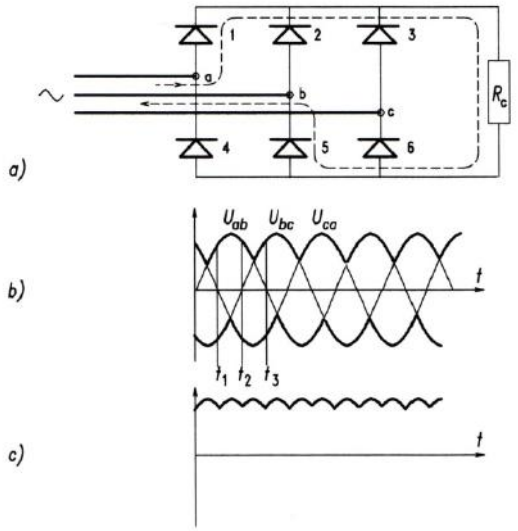
In ogni istante vi sono sempre due diodi che conducono in serie con il carico: sono quelli che sono interessati dalla stessa tensione concatenata che raggiunge, rispetto alle altre due tensioni, il massimo valore assoluto istantaneo.
Nell'intervallo 0 – t1 (Fig. 16.16b) la tensione concatenata che assume valori assoluti maggiori è la U_{bc}; essendo $U_{bc} < 0$, sarà $U_{cb} > 0$ e pertanto il diodo 3 avrà l'anodo più positivo, mentre il diodo 5 avrà il catodo più negativo: i diodi 3 e 5 entreranno in conduzione.
Nell'intervallo tl-t2 prevale la tensione U_{ab} positiva e pertanto condurranno i diodi 1 e 5. Nell'intervallo t2-t3 prevale la tensione $U_{ca} < 0$ e pertanto la $U_{ac} > 0$: condurranno i diodi 1 e 6. Nell'intervallo successivo (pari, come gli altri a 1/6 del periodo T) condurranno i diodi 2 e 6; quindi i diodi 2 e 4; infine i diodi 3 e 4.
Come si vede, ciascun diodo è in conduzione per un intervallo pari a T/3.
Nella Fig. 16.16b con tratto grosso è stata disegnata la tensione di uscita per ogni gruppo di raddrizzatori, mentre nella Fig. l6.l6c è indicata la tensione complessiva fornita dal ponte di raddrizzatori (questo diagramma si ottiene ribaltando il diagramma nel semipiano negativo ed eliminando le sovrapposizioni).

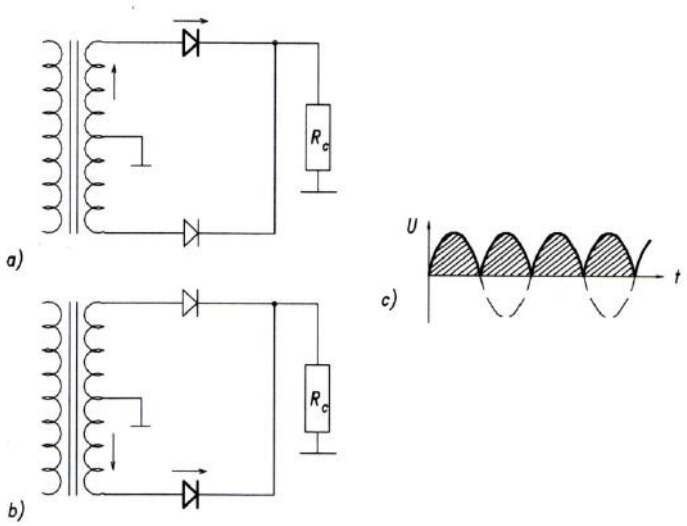


Il circuito descritto in Fig. l6.14 costituisce il più diffuso raddrizzatore a doppia semionda per i notevoli vantaggi che comporta: piccola percentuale di armoniche, elevato rapporto tra potenza raddrizzata e potenza apparente al secondario del trasformatore di ingresso (quindi minimo dimensionamento del trasformatore).

Il raddrizzatore a doppia semionda può essere realizzato anche senza ricorrere ai diodi, sfruttando le proprietà del trasformatore.
Si consideri (Fig. 16.17) un trasformatore con numero di spire doppio rispetto a quello necessario, e recante una presa a metà dell'avvolgimento, detto perciò a presa centrale. Ad esempio, volendo prelevare una tensione a 12 V si utilizzerà un trasformatore con secondario che fornisce 24 V

I terminali del secondario presenteranno, rispetto alla presa centrale, due tensioni alternate uguali ma sfasate di 180°, di modo che in ogni istante la loro differenza sia proprio, nell'esempio fatto, di 24 V





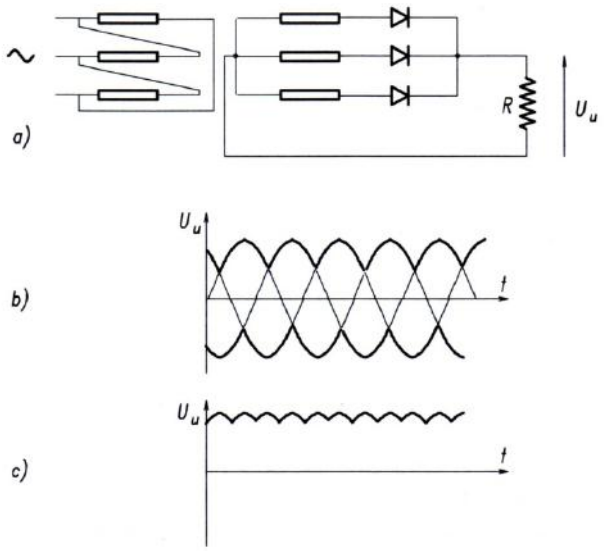
In Fig. I6.I7a il terminale superiore è positivo e perciò il diodo ad esso collegato conduce; contemporaneamente, il terminale inferiore è negativo ed il diodo collegato non conduce. La situazione si inverte quando si inverte la semionda (Fig. 1 b.17b).

Nel progetto del trasformatore bisogna tenere conto del fatto che il secondario lavora metà per volta e quindi dimensionare il nucleo per una potenza al secondario corrispondente a metà avvolgimento.

I due raddrizzatori descritti sono largamente utilizzati e sono ragioni economiche a far preferire l'uno o altro. Il raddrizzatore con trasformatore a presa centrale consente il risparmio di due diodi ma presenta il costo del trasformatore.

La versione trifase del raddrizzatore a presa centrale è illustrata in Fig. I6.18.

Per chiudere questo paragrafo, si deve accennare al problema della massima tensione inversa dei diodi, affinché non si verifichi un eventuale fenomeno di valanga distruttiva.
Nel raddrizzatore a semionda, durante le semionde in cui non conduce, all'anodo vi è una tensione massima pari al massimo valore negativo della tensione sinusoidale. Nello stesso istante, al catodo vi è la tensione di uscita che è mediamente di poco inferiore al valore massimo positivo della tensione applicata. Perciò, la massima tensione inversa che il diodo sopporta ai suoi capi è 2 Umax
Ad esempio, se un diodo lavora con una tensione di rete a 200 V efficaci, corrispondenti ad un picco di circa 311 V deve essere in grado di sopportare una tensione inversa massima di 622 V
Per un raddrizzatore a presa centrale monofase il problema è identico, essendo esso formato da due raddrizzatori a semionda distinti.



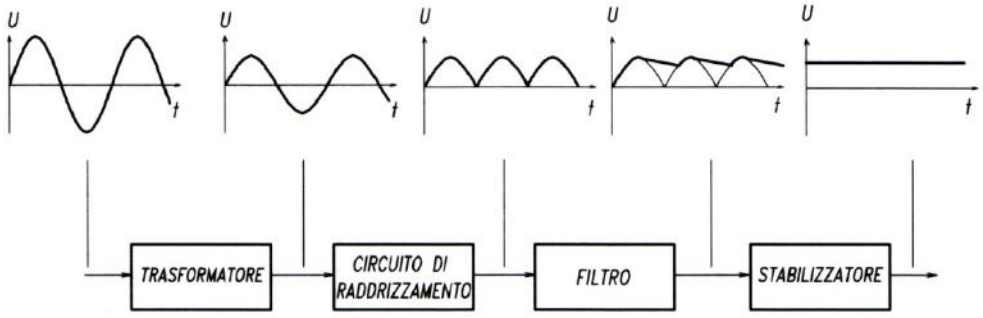
Nel caso del ponte di Graetz, invece, siccome i diodi lavorano sempre a due a due, ciascuno di essi si suddivide la tensione massima inversa: pertanto, ciascuno dei due diodi deve sopportare una tensione massima inversa pari al valor massimo della tensione sinusoidale in ingresso.

16.12 Alimentatori

Quasi tutte le apparecchiature elettroniche per funzionare hanno bisogno dì una o più . tensioni di alimentazione continue:.

Le batterie, che forniscono questo tipo di f.e.m., sono adatte per apparecchiature portatili, mentre per gli usi fissi conviene ricavare le f.e.m. continue dalle f.e.m. alternate rese disponibili dalla capillare rete di distribuzione dell'energia elettrica. Si chiamano alimentatori quegli apparecchi che assorbendo energia dalla corrente alternata di rete, riescono ad erogare in uscita una o più tensioni continue al voltaggio occorrente, positivo o negativo.

La struttura di un semplice alimentatore prevede (Fig. 16.19) un trasformatore che abbassa (o qualche eleva) la tensione di rete monofase a 220 V per dare in uscita il valore desiderato.
Subito a valle troviamo un raddrizzatore che rende la corrente unidirezionale: esso può essere a semionda o a doppia onda.
Segue infine un filtro livellatore che ha il compito di rendere quanto più possibile costante la tensione fornita.
Gli alimentatori migliori prevedono in cascata al filtro anche un circuito stabilizzatore, che ha il compito di fare in modo che la tensione fornita non oscilli (in conseguenza di variazioni del carico) oltre una soglia prestabilita.



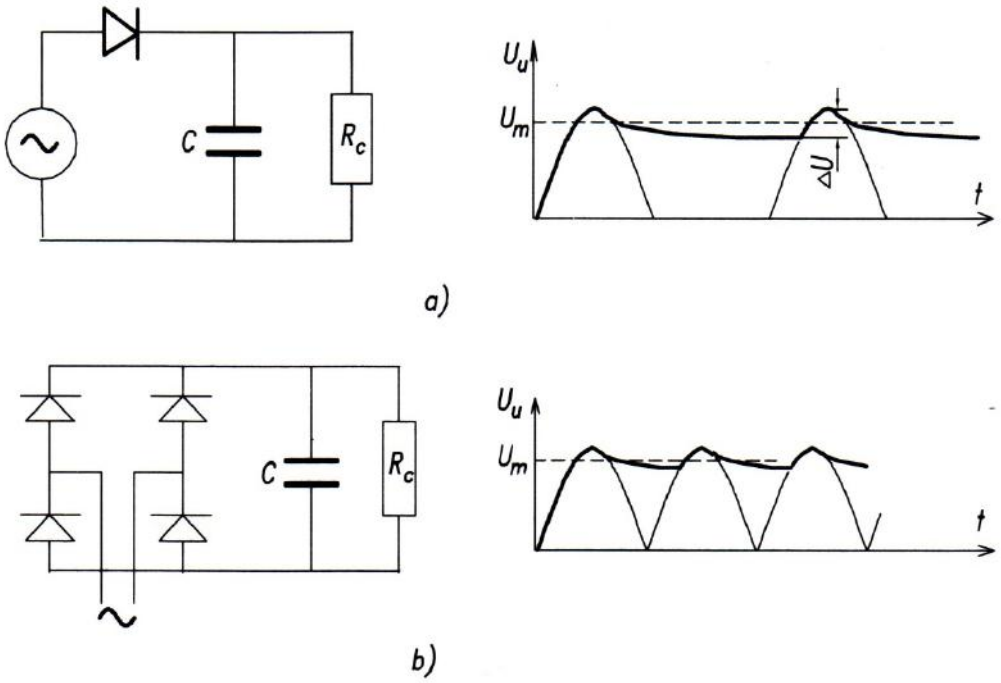
Il più semplice filtro livellatore è costituito da un condensatore di grande capacità disposto in parallelo al carico (Fîg.16.20).

Con riferimento alla Fig. 16,20a, se non ci fosse il carico U di resistenza Rc, il condensatore si caricherebbe al valore di picco della tensione e manterrebbe questa carica, risultando il diodo interdetto. Al contrario, in presenza della resistenza di carico, il condensatore può scaricarsi attraverso questa resistenza.
Per una scarica molto lenta, e quindi per una tensione ben livellata, la costante di tempo Rc · C deve essere elevata rispetto al periodo della tensione alternata da raddrizzare. Nel raddrizzatore a semionda il condensatore è caricato ogni 20 ms, cioè ogni periodo della tensione di rete. Nel raddrizzatore a doppia onda (o ad onda intera), il condensatore è invece caricato ogni 10 ms, cioè ogni mezzo periodo della tensione di rete.

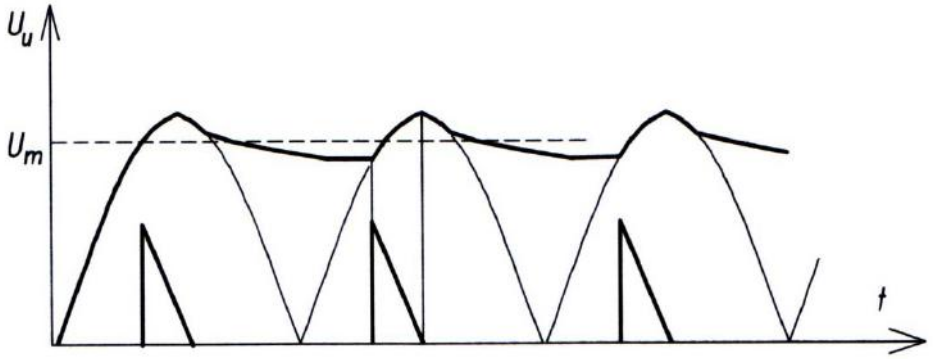
L' alimentatore è tanto migliore quanto più è piccola è la variazione massima ΔV di tensione, detta ondulazione (in inglese: ripple).
In pratica, per valutare la bontà di un alimentatore si fa riferimento al fattore di ripple, definito come il rapporto tra il valore efficace dell'ondulazione e il valor medio della tensione di uscita:

$$K_r = \frac{\Delta U_{eff}}{U_m}$$

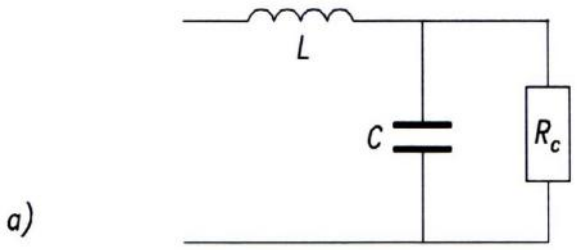
In effetti, una variazione di 1 V su un valor medio di 200 V è piccola, mentre non lo è la stessa variazione su un valor medio di 20 V

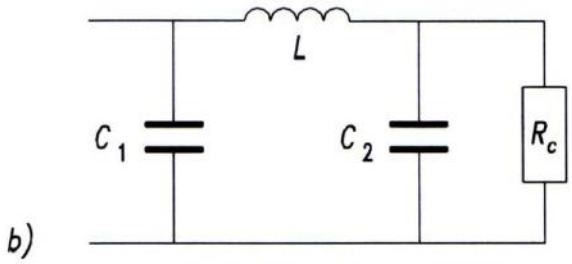


La Fig. I6.21 illustra l'andamento della corrente fornita dal generatore per ricaricare ogni volta il condensatore.



Due altri tipi di filtro consentono di ottenere tensioni ancor più livellate.
Il filtro LC serie è un quadripolo che si pone in parallelo al carico come illustrato in Fig. I6.22a. Si può dimostrare che questo filtro è efficiente nel caso di tensioni raddrizzate con frequenze più elevate di quella di rete.
Il filtro a π deriva il nome dalla figura formata dai due condensatori e dall'induttore centrale (Fig. 1 b.22b). Esso può essere considerato come un filtro LC serie in parallelo ad un filtro capacitivo.
Quando i più grossi condensatori raggiungevano a stento poche decine di microfarad, il filtro a π era il migliore. Oggi, senza troppo ingombro è facile reperire condensatori da decine di migliaia di microfarad e perciò per livellare la tensione raddrizzata molte volte basta la semplice soluzione dell'unico condensatore.

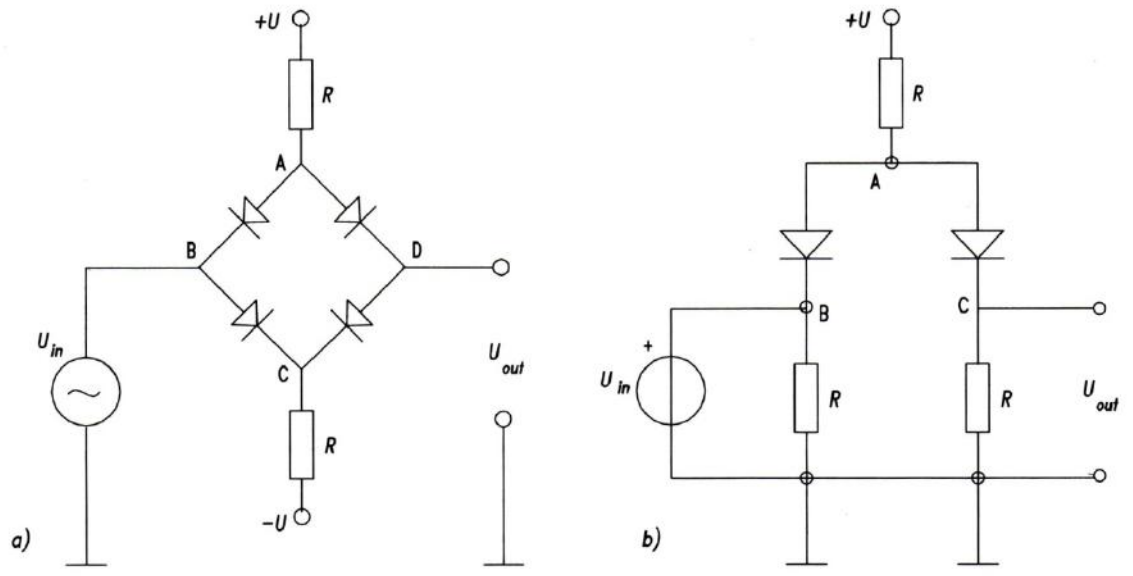




16.13 Circuito campionatore

Il campionamento di un segnale è una tecnica molto usata in telefonia ed in altri campi dell'elettronica. Essa consiste nel leggere un segnale non con continuità ma con cadenze ben precise. In fig. 16.23 sono riportati due semplici circuiti in grado di svolgere questa operazione.
.
Con riferimento al circuito della fig. 16.23a, se le tensioni +U e -U sono nulle, data la presenza dei due diodi che lasciano passare la corrente solo in senso diretto, in uscita non si ha alcun segnale. Se invece le due tensioni assumono un dato valore (costante), i diodi conducono. In tal caso, la caduta tra A e B è identica alla caduta tra A e D, perciò B e D sono allo stesso potenziale. Ma poiché B è al potenziale del generatore U_{in}, su D la tensione di uscita U_{out} ripete quella del generatore.
In tal modo, il circuito si comporta come una porta per segnale analogici: le tensioni + U e -U danno il consenso al passaggio di un segnale analogico. I diodi servono a bloccare il segnale nel caso manchi il consenso.
Ovviamente, l'ampiezza del segnale di ingresso deve essere inferiore all' ampiezza delle tensioni + U e - U decurtata della caduta su ciascun diodo.

Ora, se le tensioni di campionamento +U e -U sono impulsive, il segnale di ingresso sarà trasferito a pezzetti. Lo scopo di questa strategia può essere quello di trasmettere più segnali sullo stesso cavo: se ne può trasmettere uno durante i tempi morti degli altri.



Questa tecnica, detta di chopper, è utilizzata ad esempio negli oscilloscopi a doppia traccia: qui, un unico fascetto di elettroni deve descrivere due diverse forme d'onda e ciò è fatto saltando continuamente da un segnale all'altro e suddividendoli in pezzetti talmente piccoli da fare sembrare continui entrambi i diagrammi.

Il circuito della fig. 16.23b è più semplice e serve a campionare segnali di un solo segno. La tensione di campionamento + U è una sola; negli istanti in cui essa è nulla il punto C è a tensione nulla e i due diodi bloccano qualsiasi segnale che da B voglia passare in C. Quando +U ha valore costante positivo, i due diodi conducono e determinano ai loro capi la stessa caduta di potenziale, per cui $U_B = U_c$. In questa condizione la tensione di uscita U_{out} è uguale alla tensione del generatore U_{in}.

16.14 Diodo LED

Un diodo LED (Light Emitting Diode) è un diodo che quando è attraversato da corrente emette una radiazione luminosa monocromatica, che può essere rossa, verde o gialla.
L'emissione di energia raggiante è diretta conseguenza delle ricombinazioni elettroni -lacune che si hanno nella zona di giunzione: questa ricombinazione è infatti la ricaduta dell'elettrone verso una banda di energia inferiore, con conseguente emissione dell'energia in più.

In un LED i materiali sono scelti in modo da presentare una netta distinzione dei livelli energetici, in modo da fornire una elevata energia raggiante (fotoni). Questa energia si trasmette ad una frequenza che cade nel campo del visibile.
Per la legge di Einstein, infatti, l'energia E associata ad un'onda elettromagnetica di frequenza f è data dalla:

$E = h * f$

essendo h la costante universale di Plank. Facendo riferimento alla lunghezza d'onda l ed esprimendo l'energia in elettronVolt (eV):

$\lambda = 1,24 / E * [\mu m]$

In questi materiali, pertanto, l è compreso tra i valori 0,3 mm, corrispondente alla luce violetta, e 0,75 mm, corrispondente alla luce rossa. Ad esempio, la luce gialloverde corrisponde a valori 0,5 - 0,6 mm.

In commercio esistono LED che emettono luce gialla, verde e rossa, ed anche infrarossa.
Ad esempio, l' arseniuro di gallio e alluminio (Ga Al As) presenta un salto tra banda di valenza e banda di conduzione di 1,86 eV e ad ogni ricombinazione è emessa radiazione alla frequenza di 0,667 mm, che corrisponde a luce visibile di colore rosso.

Altre miscele presentano un salto lievemente più elevato e la radiazione emessa ha una lunghezza d'onda lievemente più corta, corrispondente al giallo (0,575 mm) o al verde (0,550 mm).
La luce emessa da un LED è abbastanza monocromatica (cioè di un solo colore) perché in tutte le ricombinazioni gli elettroni effettuano lo stesso salto energetico.

I diodi LED, proprio perché presentano una più marcata differenza di energia tra banda di valenza e banda di conduzione, rispetto ai diodi al silicio o al germanio, creano nella giunzione una barriera di potenziale più elevata. Per i LED rossi la d.d.p. di barriera (quindi la U_d ai capi del LED durante la conduzione) è circa 1,6 V (contro 0,6 V per il silicio e 0,2 V per il germanio).
La corrente diretta sufficiente a fare illuminare normalmente un diodo LED si aggira sui 20 mA, perciò la potenza dissipata è di 32 mW.

Rispetto ad una normale lampadina, il LED non può certo competere per potenza luminosa, però esso è certamente da preferire perché trasforma in modo più efficiente l'energia che gli si fornisce; come noto, una lampadina trasforma invece la maggior parte dell'energia che assorbe in calore.
Ma la caratteristica che sta alla base della capillare diffusione dei LED è la rapidità con la quale può essere acceso o spento, laddove il filamento di una lampadina per riscaldarsi e raffreddarsi ha bisogno di decimi o centesimi di secondo, un intervallo enormemente lungo per l'elettronica.

16.15 Fotodiodo

Un fotodiodo sfrutta il fenomeno duale del LED: un fotone proveniente dall'esterno può colpire in una giunzione pn un elettrone di valenza conferendogli l'energia per passare nella banda di conduzione: in tal modo, un fotone crea una coppia elettrone - lacuna portatori di carica. Quello descritto è il cosiddetto effetto fotoelettrico.

In una giunzione pn polarizzata direttamente vi è un notevole passaggio di cariche maggioritarie e perciò eventuali coppie di portatori di carica prodotte da fotoni esterni si confonderebbero nel mare dei portatori ordinari.
Al contrario, in una giunzione polarizzata inversamente non vi è movimento dei portatori maggioritari e vi è solo la modestissima corrente inversa dei portatori minoritari: in questo caso i nuovi portatori creati dai fotoni fanno sentire il loro effetto, incrementando la corrente inversa in modo direttamente proporzionale all'intensità della radiazione incidente.
Se la tensione inversa che polarizza il diodo è adeguatamente grande, prossima perciò alla tensione di breakdown, le coppie di portatori creati dai fotoni possono innescare nel circuito una corrente inversa notevole (non più proporzionale alla radiazione incidente), che può essere innescata anche da radiazioni poco intense.

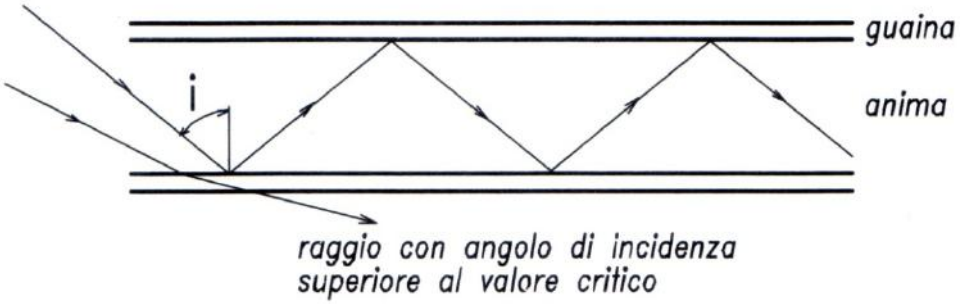
16.16 Trasmissioni a fibra ottica

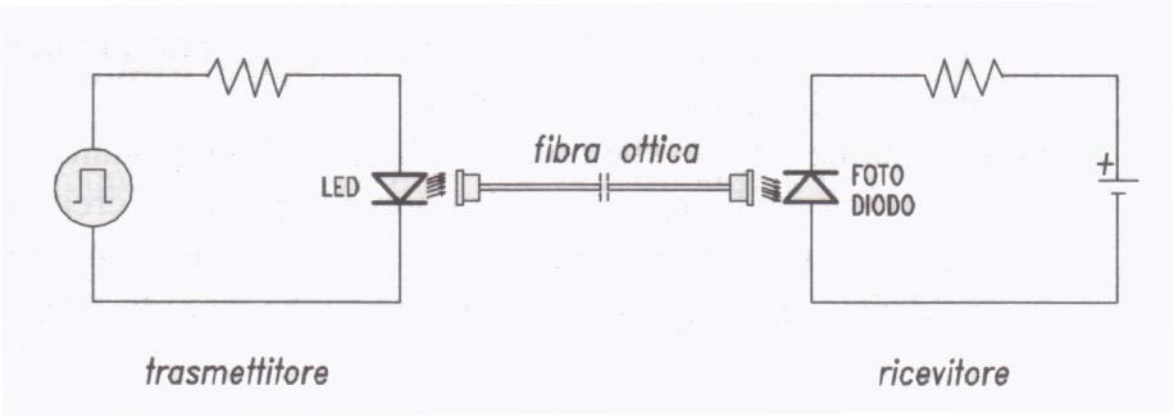
Il sistema fotoemettitore - fotoricevitore costituito da un LED e da un fotodiodo ha aperto tutto un mondo di applicazioni, ed in particolare le telecomunicazioni con fibre ottiche.

Alla base della costruzione delle fibre ottiche vi sono vetri e resine straordinariamente trasparenti (conserverebbero la trasparenza anche con spessori di chilometri). Essi sono formati a fili molto sottili (decine o centinaia di micron) nei quali la parte più interna (anima, in lingua inglese core: fig. 16.24) ha un indice di rifrazione più alto di quella esterna. In questo modo, un raggio luminoso iniettato, con un opportuno angolo di entrata, si riflette totalmente all'interno del lungo filo, qualunque configurazione questo assuma, fuoriuscendo solamente all'altra estremità.
Questo filo di vetro a due strati, rivestito da una guaina protettiva, è chiamato fibra ottica.

Il sistema di trasmissione (fig. 16.25) prevede un LED che emette luce (a raggi infrarossi) ad una estremità e un fotodiodo ricevitore dall'altra. Si noti, nella fig. 16.25, come i simboli del LED e del fotodiodo siano gli stessi del diodo normale, con accanto piccole frecce a indicare la radiazione elettromagnetica.

Grazie alla eccellente trasparenza delle fibre ottiche può essere coperta una distanza di alcuni chilometri.





La velocità con la quale è possibile accendere e spegnere il LED e la risposta altrettanto veloce del fotodiodo consentono di inviare impulsi in rapidissima successione, fino a 50 MHz.
In tal modo, possono essere trasmessi lungo la fibra ottica rapidissimi impulsi di corrente (rappresentativi, secondo un opportuno codice binario, di informazioni) trasformati in impulsi luminosi.

A differenza dei segnali elettrici lungo conduttori metallici, i segnali luminosi sono del tutto insensibili ai disturbi elettromagnetici. Inoltre, perdono meno potenza. Infine, nel punto di rilevamento possono essere raggiunte temperature anche molto elevate.
Da un punto di vista della riservatezza delle informazioni, mentre un segnale elettrico in un filo può essere letto in qualsiasi punto (con un voltmetro o altro), il segnale elettrico può essere letto solo all'arrivo.

16.17 Diodo laser

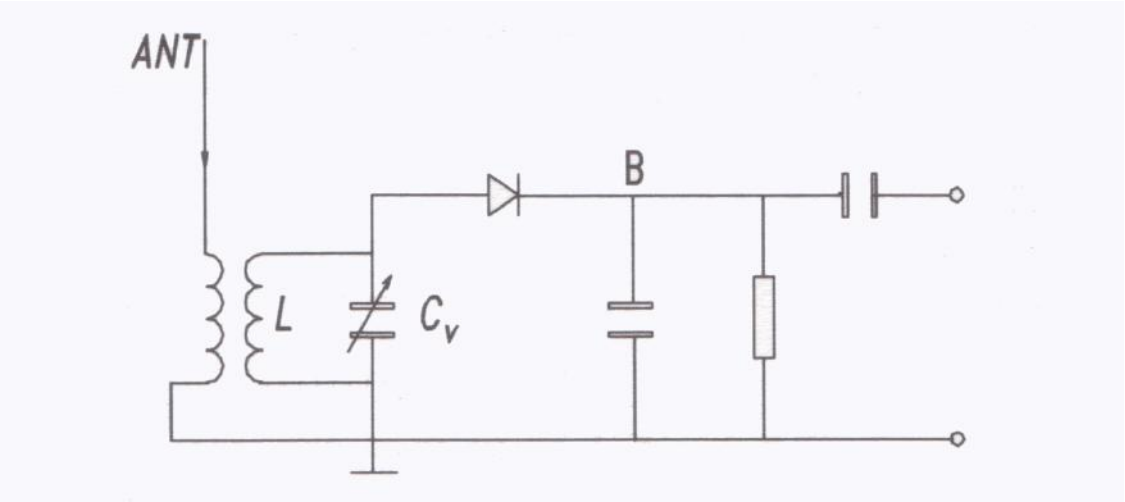
Sono state realizzate giunzioni pn in grado di emettere dei veri e propri raggi laser di piccolissima potenza.
Essi sono utilizzati nelle trasmissioni con fibra ottica e sono alla base anche dei compact disk.
In questi, l'informazione (parola, immagine, suono) è stata codificata, con un raggio laser più potente, sotto forma di successione di buche (pitch) oppure in mancanza di esse, lungo una lunghissima spirale che inizia al centro del disco.
In lettura, la luce laser emessa dalla giunzione pn colpisce i punti della spirale e dove trova una buca non è riflessa; un circuito di rilevamento opportuno converte questo fatto nel bit 1. In mancanza di buca, la luce laser è riflessa ed il circuito rilevatore converte l'informazione in bit 0.

16.18 Altre applicazioni dei diodi

I diodi sono protagonisti di molte altre applicazioni. Ci limitiamo qui a citare i circuiti moltiplicatori di tensione alimentati con tensione alternata dalla quale si ottiene una tensione continua secondo un certo fattore intero di moltiplicazione.

Inoltre non possiamo ignorare l'applicazione fondamentale della radiotecnica: la rivelazione di un segnale.

Come sappiamo, il suono o 1'immagine televisiva sono trasmesse con la tecnica della modulazione di ampiezza: l'onda da trasmettere (suono o immagine) è utilizzata per modulare l'ampiezza di un'onda ad alta frequenza, chiamata portante.
Ad una distanza anche notevole, un'antenna (fig. 16.26) investita dal campo elettromagnetico dell'onda portante genera ai suoi capi un segnale avente la stessa frequenza della portante. Il segnale dell'antenna va ad alimentare l'avvolgimento primario di un trasformatore.
Il secondario è collegato in parallelo ad un condensatore e forma con esso un circuito LC parallelo che può mettersi a risonare col segnale catturato dall'antenna. Agendo sui parametri del circuito risonante è possibile selezionare la frequenza di risonanza, in modo che sia catturato il segnale che interessa tra tutti quelli che investono l'antenna.
Il segnale in uscita dal circuito LC passa quindi ad uno stadio raddrizzatore a semionda con filtro capacitivo. Nel punto B del circuito è dunque presente il segnale modulante "estratto" da quello "portante" che era servito per il trasporto.
Esso presenta ancora una componente continua pari al valor medio dell'onda portante: per eliminare questa componente si utilizza un secondo condensatore. Collegando al circuito della fig. 16.26 un amplificatore audio con un altoparlante si ottiene un elementare radioricevitore.



LABORATORIO DI ELETTRONICA

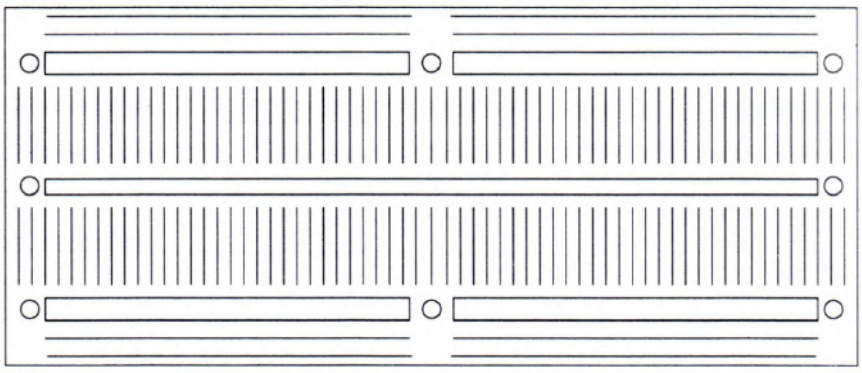
Per la prova rapida di circuiti elettronici si impiega comunemente una base di montaggio chiamata breadboard.
Essa si presenta come una tavoletta di plastica bucherellata. Nei fori vi sono due linguette metalliche la cui funzione è di realizzare il contatto elettrico con un filo elettrico opportunamente "spelato" che vi sia inserito.
Alcuni fori sono tra loro collegati attraverso minuscole barrette metalliche poste all'interno della breadboard. I collegamenti possono essere verificati con un multimetro predisposto per misure di resistenza (con una portata qualunque), due coccodrilli e due spezzoni di filo: se il multimetro dà indicazione di resistenza fuori portata significa che tra i due fori non vi è contatto; al contrario, se è registrata una resistenza molto bassa, tra i due fori vi è connessione elettrica.
Si può facilmente verificare lo schema dei collegamenti interni indicato in fig. 16.27.

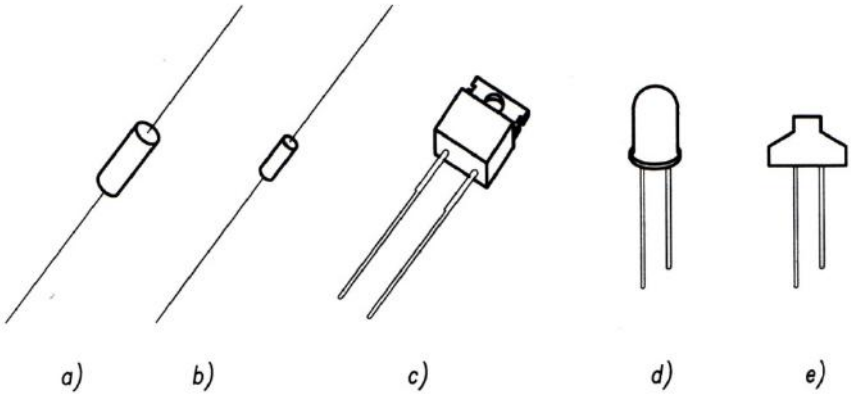
L'assemblaggio dei componenti su una breadboard permette rapide connessioni ed altrettanto rapide modifiche sul circuito. Inoltre, questa tecnica consente l'utilizzo degli stessi componenti per la realizzazione di diversi circuiti elettronici, per studi e sperimentazioni.
Le alternative, più complesse e pertanto da noi ignorate, sono costituite dalla realizzazione dei circuiti con collegamenti effettuati a mano, oppure con la tecnica dei circuiti stampati.

I diodi sono disponibili in contenitori la cui forma e le cui dimensioni sono funzione della massima potenza che possono dissipare: I costruttori classificano i diodi in due grandi categorie:

- diodi disegnale;
- diodi di potenza.

La fig. 16.28 illustra alcune forme tipiche, sia per diodi normali (fig. 16.28a,b,c) che per LED (Fig. 16.28d,e).





In generale, le norme prescrivono due metodi per siglare i componenti a semiconduttore.
Secondo il primo metodo, si usa un prefisso seguito da un numero ed eventualmente una lettera. Vediamo alcuni significati:

1N indica un semiconduttore con una giunzione, come i diodi;
2N indica un semiconduttore con due giunzioni, come un transistor;
3N indica un semiconduttore con tre giunzioni;
4N-6N indica un dispositivo optoelettronico.

Il numero di serie identifica il singolo componente, mentre la lettera aggiuntiva identifica una particolare caratteristica del componente. Ad esempio:

1N4148 diodo a giunzione al silicio;
2N2222 transistor di segnale;
3N128 transistor JFET;
4N25 fotoaccoppiatore.

Il secondo metodo utilizza due o più lettere seguite da un numero di serie. La prima lettera identifica il materiale semiconduttore:
A germanio;
B silicio;
C arseniuro di gallio;
D antimonio di indio;
R elementi fotosensibili.

La seconda lettera identifica il tipo di dispositivo elettronico:

A diodo di rivelazione;
B diodo varicap;
C transistor per basse frequenze;
D transistor di potenza per basse tensioni;
E diodo tunnel;
F transistor per alte frequenze;
L transistor di potenza per alte frequenze;
P fotoelemento;
R tyristor;
S transistor a commutazione;
T TRIAC, SCR di potenza;
U transistor di potenza per commutazione;
X varactor;
Y diodo rettificatore di potenza;
Z diodo Zener.

16.19 Collaudo di un diodo col multimetro

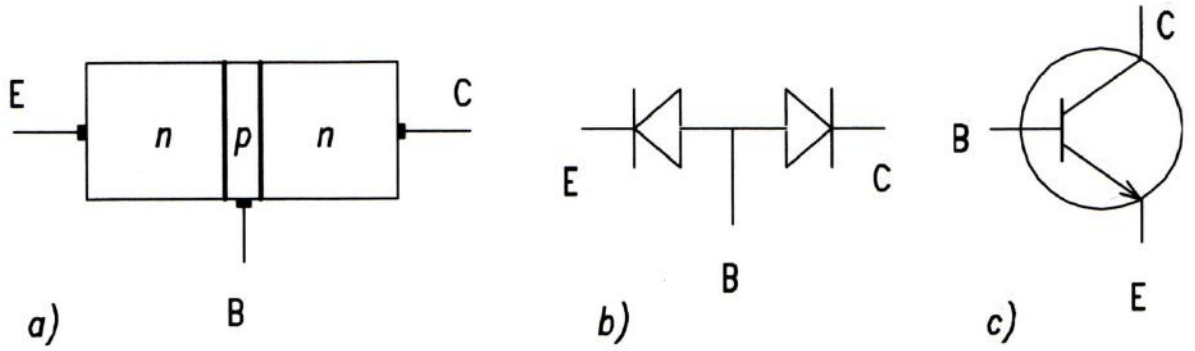
Nei multimetri di vecchia generazione, doveva essere utilizzata la funzione come ohmmetro: misurando la resistenza del diodo, in un senso essa doveva essere molto bassa e nell'altro molto alta. Nei multimetri moderni, specie se digitali, vi è una specifica funzione di controllo diodi. Inserendo i terminali del diodo entro apposite prese è attivato un piccolo generatore di corrente (generalmente da 1 mA) e lo strumento si predispone a leggere la tensione in mV ai capi del componente. In senso diretto si dovrà trovare circa 600 mV per i diodi al silicio e 200 mV per quelli al germanio; in senso inverso i diodi si devono comportare tutti come circuiti aperti.

Il transistor e le sue applicazioni

17.1 Il transistor a giunzione

Il transistor a giunzione, o transistor bipolare, indicato con la sigla BJT, da Bipolar Junction Transistor, è costituito da due giunzioni pn, nella configurazione pnp o npn; quest'ultima è quella più frequentemente utilizzata, e ad essa perciò ci riferiremo. I terminali (fig. 17.1) che permettono di accedere alle tre zone di semiconduttore si chiamano, rispettivamente, emettitore (E), base (B) e collettore (C).
In Fig.I7.1b è illustrata la corrispondente schematizzazione a diodi, mentre la fig. I7.lc riporta il simbolo normalmente impiegato negli schemi elettronici.

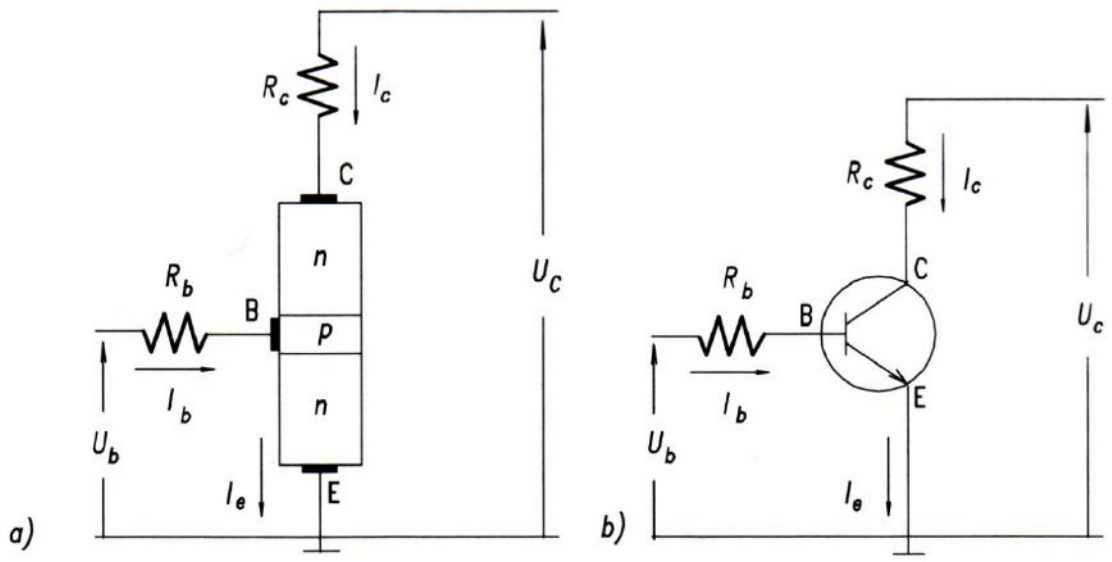
Il dispositivo funziona nel modo desiderato se l'emettitore è molto drogato e se:
- la regione di base è molto sottile;
- la regione di base è pochissimo drogata;
- il collettore è ancora meno drogato della base.



Per quanto riguarda la polarizzazione, si possono creare tre configurazioni, ma la più frequente è quella ad emettitore comune cosiddetta perché l'emettitore è in comune al circuito d'ingresso, a due poli, ed al circuito d'uscita, a due poli. Il transistor in questa configurazione è pertanto un quadripolo anche se ha tre terminali.
Per ottenere il funzionamento desiderato, la giunzione EB deve essere polarizzata direttamente e la giunzione CB inversamente. Si ottiene il circuito illustrato in fig. 17.2.
Infine, la tensione U_{CE} deve essere di un ordine di grandezza superiore alla tensione U_{BE}.

Ciò premesso, poiché la prima giunzione EB è polarizzata direttamente, un flusso di elettroni (portatori maggioritari nella zona n) la attraversa, entrando nella zona p di base, dove diventano portatori minoritari. Se la tensione U_{BE} è abbastanza elevata (in ogni caso superiore al valore di soglia della giunzione), il flusso di elettroni iniettati dall'emettitore nella base è intenso.
In contemporanea, vi è un flusso di lacune dalla zona di base all'emettitore, ma poiché la base è pochissimo drogata, gli elettroni in senso opposto saranno molto superiori di numero.

Ora, poiché la base è molto sottile, la quasi totalità degli elettroni iniettati dall'emettitore (corrente di emettitore I_e) giungerà al collettore, dove troveranno la polarizzazione inversa che ne favorisce il moto verso il collettore, di modo che nel circuito del carico, dove è la resistenza di carico R_c, fluirà la corrente di collettore I_c.
Si noti che nella fig. I7.2 sono indicati i versi convenzionali della corrente elettrica (cioè i versi delle cariche positive).



La cosiddetta corrente di base I_b è dovuta sia al contemporaneo transito di lacune da B verso E, sia alle lacune richiamate dal circuito di ingresso a causa della ricombinazione elettroni-lacune che avviene nella base. Si tratta di due contributi estremamente ridotti e pertanto la corrente di base è molto piccola.

Per la legge dei nodi:

$$I_b = I_e - I_c$$

Si potrà ritenere:

$$I_c = \alpha\ I_e$$

con a un coefficiente dipendente dal tipo di semiconduttore, ma in ogni caso poco inferiore a 1: di solito tra 0,95 e 0,99. Infatti, in molte applicazioni si suppone direttamente che $I_c = I_e$.

Sostituendo la precedente espressione:

$$I_b = (1 - \alpha) * I_e = (1 - \alpha)\ I_c/\alpha \quad => \quad I_c = \alpha / (1 - \alpha) * I_b$$

In sintesi:

$$I_c = \beta\ I_b$$

Con $\alpha = 0{,}99$ si avrebbe $I_c = 99 I_b$, dal che si vede immediatamente l'effetto di amplificazione della corrente di base operata dal transistor.
Nei data sheet dei transistor commerciali, al posto di β è indicato il parametro h_{FE}, chiamata guadagno del transistor.

Concludendo:
La U_{BE}, e pertanto la I_b, controlla la barriera di potenziale della giunzione BE e determina quindi il numero di elettroni che sono iniettati dall'emettitore alla base e che saranno messi in moto, a meno di piccole perdite, nel circuito del carico.

Il generatore di questo circuito avrà la funzione di rigenerare l'energia di questi elettroni, perduta nell'attraversamento della resistenza di carico, riportandoli all'emettitore.

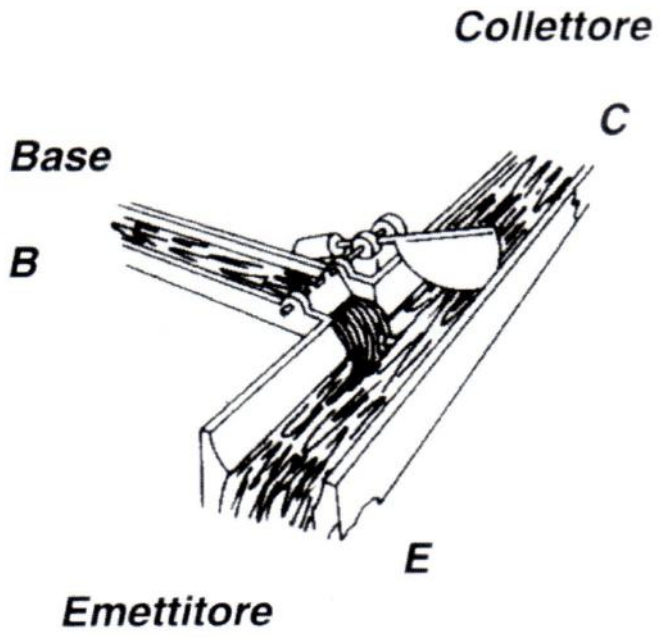
Piccole variazioni della corrente di base (sempre dell'ordine dei microampere) si traducono in grandi variazioni della corrente di collettore (sempre dell'ordine dei milliampere o più).
L'analogia idraulica illustrata in fig. I7.3 mostra come una piccola corrente di base sia in grado, grazie al contrappeso, di pilotare o modulare (non di generare!) la ben più grande corrente di collettore.

Un transistor si comporta pertanto come una resistenza variabile, ed il suo nome può essere considerato acronimo di transformer resistor.

Da un altro punto di vista, si può affermare che il transistor porta corrente da una zona a bassa resistenza (la giunzione BE polarizzata direttamente) ad una zona ad alta resistenza (la giunzione collettore - base polarizzata inversamente). Ed il nome può essere fatto discendere dall'espressione transfer resistor.

Il transistor pnp funziona in modo analogo, ma con tensioni e correnti di segno opposto.

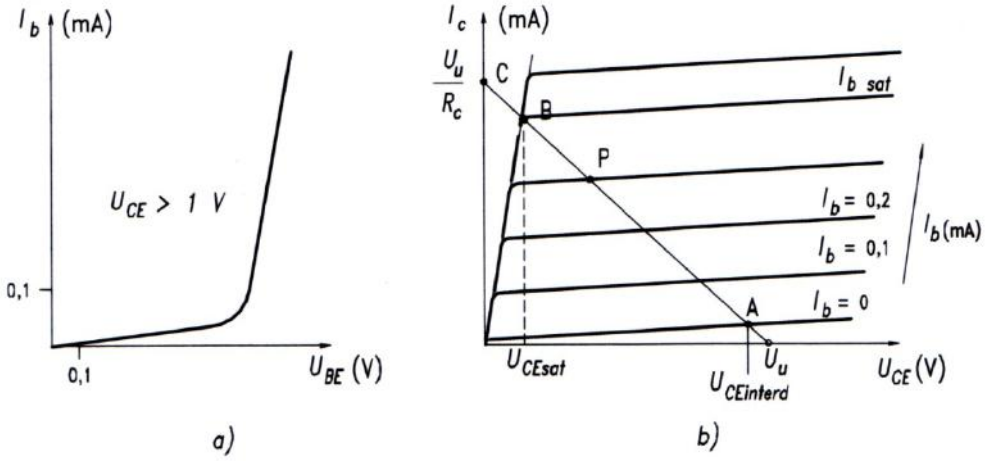
17.2 Curve caratteristiche del transistor a giunzione



La curva caratteristica della giunzione BE, detta anche curva caratteristica di ingresso, non differisce da quella di un normale diodo (Fig.I7.4a). Si nota che per valori superiori al valore di soglia caratteristico della giunzione np, la tensione U_{BE} si

mantiene approssimativamente costante al variare della corrente di base. In pratica, per applicazioni a bassa potenza, si può ritenere per transistor al silicio:

$$U_{BE} \equiv \text{costante} = 0{,}65\text{V}$$



Più interessante è invece la curva caratteristica di uscita, o meglio la famiglia di curve caratteristiche che esprimono la relazione tra tensione di collettore U_{CE} e corrente di collettore per ogni valore della corrente di base (fig. I7.4b).
Il punto di lavoro del transistor è dato dalla intersezione della curva caratteristica con la retta di carico data dall'equazione della maglia di uscita:

$I_c = U_c/R_c - 1/R_c * U_{CE}$

Per $I_b = 0$ la corrente di collettore è praticamente nulla, salvo la piccolissima corrente inversa di saturazione (I_{CE0}) della giunzione BC. La tensione U_{CE} è pertanto pari alla f.e.m. esterna U_C, a meno della piccola caduta su R_c dovuta alla I_{CE0}.
Il punto di lavoro è il punto A: il transistor si comporta come un interruttore aperto, ed il suo stato è detto di interdizione.
In assenza di corrente di base, i potenziali di B ed E sono uguali, e pertanto le giunzioni BE e BC sono ora polarizzate entrambe inversamente.

Se viene aumentata la corrente di base, polarizzando direttamente la giunzione BE, la corrente del collettore aumenta proporzionalmente ($b * I_b$). La corrente di collettore provocherà una maggiore caduta sulla resistenza R_c, così che la tensione U_{CE} diminuirà, ma non in modo tale da mutare lo stato di polarizzazione inversa della giunzione BC.
Aumentando la corrente di base il punto di funzionamento si sposta dunque lungo la AB, mantenendo in ogni punto la proporzione $I_c = b * I_b$. Questa è la cosiddetta zona di comportamento lineare del transistor, o zona nella quale il transistor si comporta come una resistenza variabile.
Dato l'effetto di amplificazione della corrente di base, si parla anche di zona attiva.

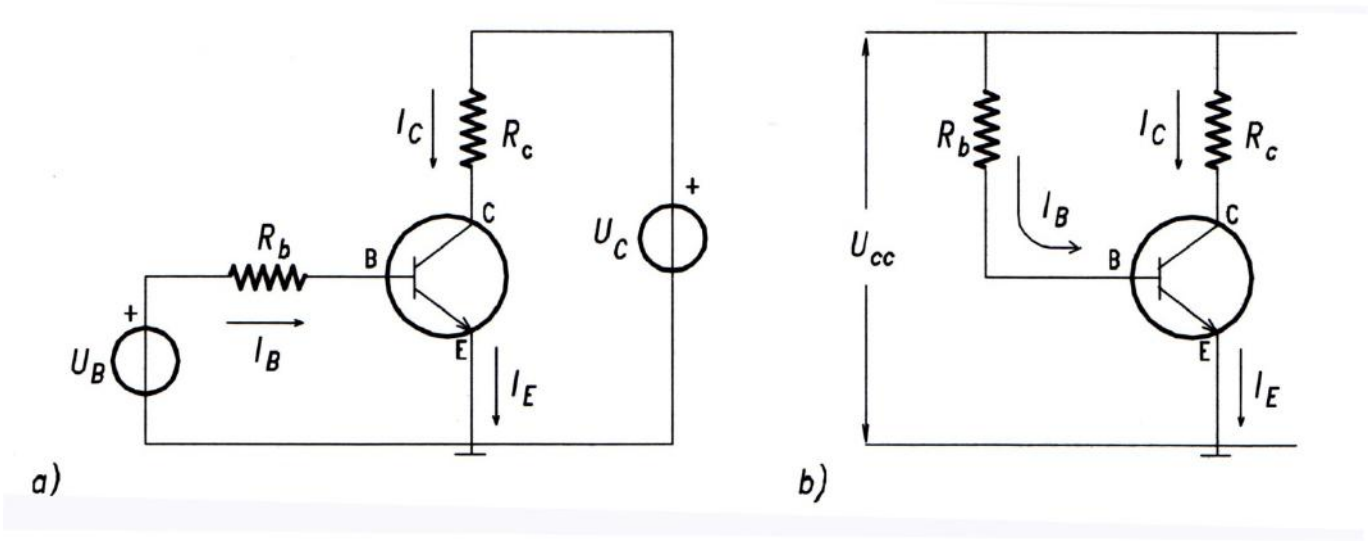
Aumentando la corrente di base oltre il valore I_{bsat}, la U_{CE} diventa inferiore a U_{BE}, la giunzione BC risulta anch'essa polarizzata direttamente e la corrente di base perde il controllo sulla corrente di collettore, che cessa pertanto di aumentare anche aumentando la I_b.
L'insieme di questi punti così raggiunti costituisce la cosiddetta zona di saturazione, nella quale il transistor si comporta come una semplice interruttore chiuso. Se il transistor presentasse resistenza nulla (transistor ideale), cioè si comportasse come un perfetto cortocircuito, la tensione U_{CEsat} sarebbe pari a zero e la corrente di collettore assumerebbe il valore (punto C del diagramma in fig. I7.4b):

$$I_{c,sat} = U_c/R_c$$

In un transistor reale raramente la U_{CEsat} supera il valore di 0,5V.

17.3 Autopolarizzazione

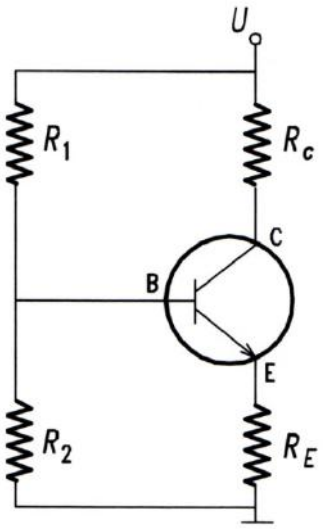
Fin qui è stato considerato il circuito di polarizzazione a due tensioni, richiamato in fig. 17.5a. Dato che l'impiego di due diverse alimentazioni risulta spesso scomodo, si può ricorrere al circuito ad una sola alimentazione illustrato in fig. 17.5b.



Un modo equivalente per realizzare 1' autopolarizzazione è illustrato in fig. 1 7.6.
Se si trascura la corrente assorbita dalla base, la tensione applicata alla base è facilmente calcolabile:

$$U_B = R_2 / (R_1 + R_2) * U$$

Nel circuito della fig. 17.6 la resistenza R_E svolge una funzione di stabilizzazione del punto di lavoro contrastando le variazioni di temperatura. Si può dimostrare che se I_c dovesse ad esempio aumentare, aumenterebbe anche la caduta su R_E, il che comporterebbe una diminuzione di I_B e quindi una conseguente diminuzione di I_C. Una resistenza R_E andrebbe posta anche nel circuito della fig. 17.5.



17.4 Transistor BJT come amplificatore per piccoli segnali

In generale, si chiama segnale elettrico una informazione, un messaggio, codificato tramite una tensione o una corrente che variano nel tempo.

Essi sono generati dai trasduttori, dispositivi che producono una grandezza elettrica proporzionale alla grandezza di ingresso. Ad esempio, un potenziometro trasforma uno spostamento in una tensione variabile, una termocoppia trasforma una temperatura in una tensione variabile, un microfono trasforma il suono in una tensione o corrente variabile ecc.
Spesso, tuttavia, i segnali prodotti dai trasduttori presentano una ampiezza molto modesta e si rende necessaria la loro amplificazione per assicurare una trasmissione ed un successivo trattamento più affidabili.
Il principio di conservazione dell'energia non può essere comunque violato: il dispositivo che amplifica la potenza non può creare energia, ma si limita a modulare l'energia fornita da una sorgente esterna, in questo caso costituita da una batteria esterna.

Un amplificatore è sempre un quadripolo ai cui due morsetti di entrata è applicato il segnale da amplificare e che fungerà da modulatore dell'energia esterna; l'energia tratta da questa fonte esterna si ritroverà ai due morsetti di uscita.

Nota: *per questa sua funzione di amplificazione il quadripolo è detto attivo. Quadripoli passivi si limitano invece a trasferire potenza dai morsetti di entrata a quelli di uscita senza alterarne il valore. Un esempio importante è costituito dal trasformatore, il quale ha lo scopo di trasferire potenza modificandone il valore dei parametri (tensione e corrente): infatti, se eleva la tensione, nella stessa proporzione riduce la corrente, e viceversa.*

Ciò premesso vediamo come un transistor BJT fatto lavorare nella zona lineare costituisca un efficiente amplificatore di piccoli segnali. Questa costituisce pertanto una delle più importanti applicazioni dei transistor.

Si consideri il circuito in fig. 17.7a.
Il generatore di tensione U_B serve per la polarizzazione del transistor, mentre U_S è il generatore che produce il segnale di ingresso simulando un traduttore.
A patto di utilizzare il transistor nella zona di funzionamento lineare è possibile applicare il principio della sovrapposizione degli effetti.
Supponendo attivo solo il generatore U_B si determina il punto di lavoro P del transistor. Quando si attiva anche il generatore di segnale, la tensione U_S produce una variazione della corrente di base rispetto al valore I_{B0} corrispondente al punto P:

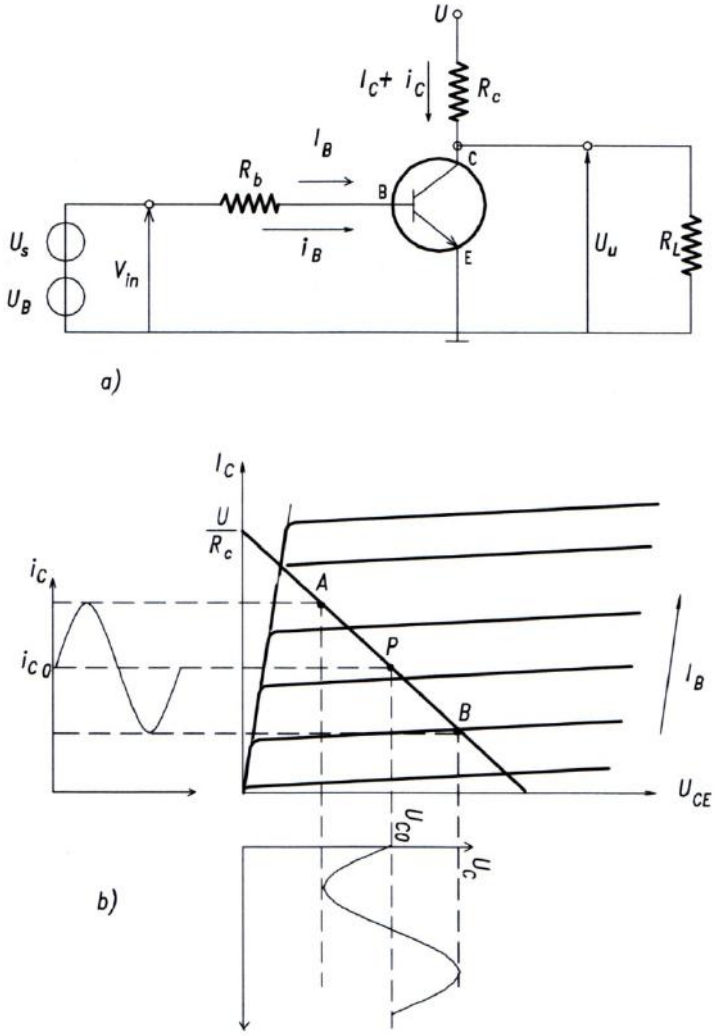
$i_B(t) = I_{B0} + i_b(t)$

In corrispondenza ai diversi valori della corrente di base, il punto di lavoro assume tutti i valori compresi tra i punti A e B corrispondenti alla massima e minima corrente di base.
Se il segnale di ingresso è sinusoidale, il punto di funzionamento si sposta con moto armonico tra i punti A e B.
Poiché ad ogni punto di funzionamento corrisponde una coppia di valori U_{CE} e i_C, anche queste due grandezze oscilleranno sinusoidalmente nel tempo, come si vede in fig. l7.7b:

$i_C(t) = I_{C0} + i_C(t)$; $U_{CE}(t) = U_{CE0} + u_{CE}(t)$

Dove il primo termine costante rappresenta la componente continua di polarizzazione, mentre il secondo è la componente di uscita dovuta al solo segnale di ingresso.

Si noti, dal diagramma della fig. 17.7b, che la tensione di uscita $U_{CE}(t)$ è sfasata di 90° sulla corrente di uscita $i_c(t)$: infatti, quando in un BJT la i_c diminuisce la U_{CE} aumenta e viceversa. L amplificatore è pertanto detto invertente.

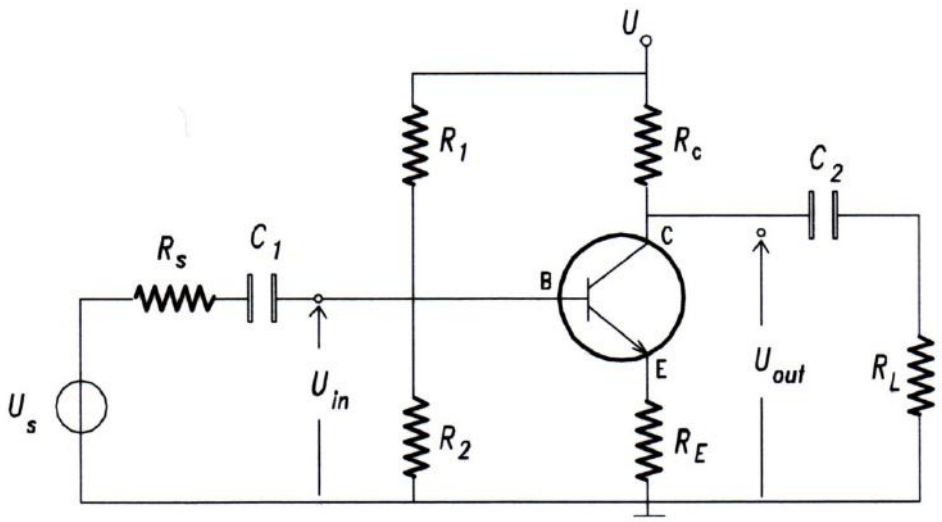


in tal modo, una corrente di ingresso di pochi microampere può produrre una variazione della corrente di collettore dell'ordine dei milliampere ed una variazione della tensione di uscita di diversi volt.

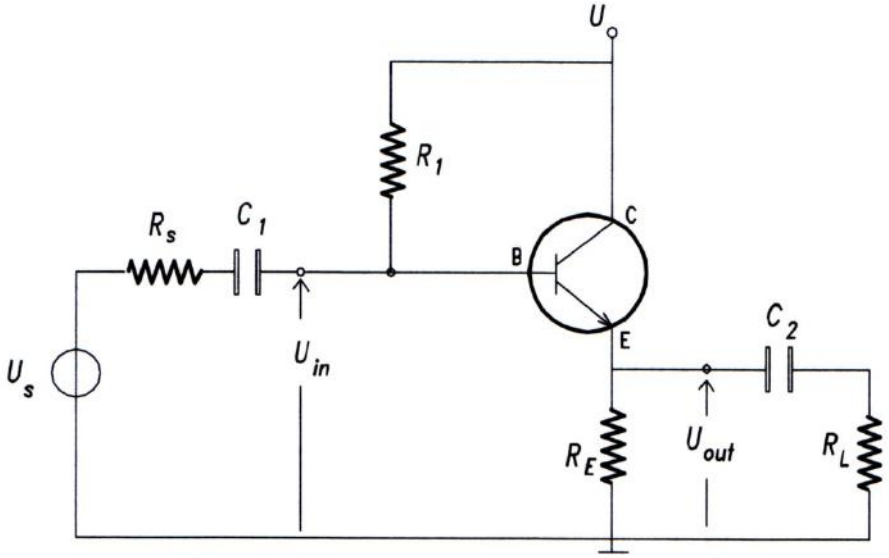
Ai segnali di ingresso e di uscita sono associate delle potenze. Amplificando le correnti e le tensioni, il dispositivo amplifica dunque la potenza in entrata. Il principio di conservazione dell'energia non è tuttavia violato: infatti, l'amplificazione di potenza avviene a scapito dell'energia fornita dall'alimentazione.

Si comprende l'importanza della scelta del punto di polarizzazione: una scelta sbagliata (il punto P troppo vicino al punto di interdizione o di saturazione) implica il taglio di valori di corrente e tensione di collettore, perché non risulta possibile una loro escursione completa. Solitamente si fa in modo che il punto di lavoro sia a metà della retta di carico.

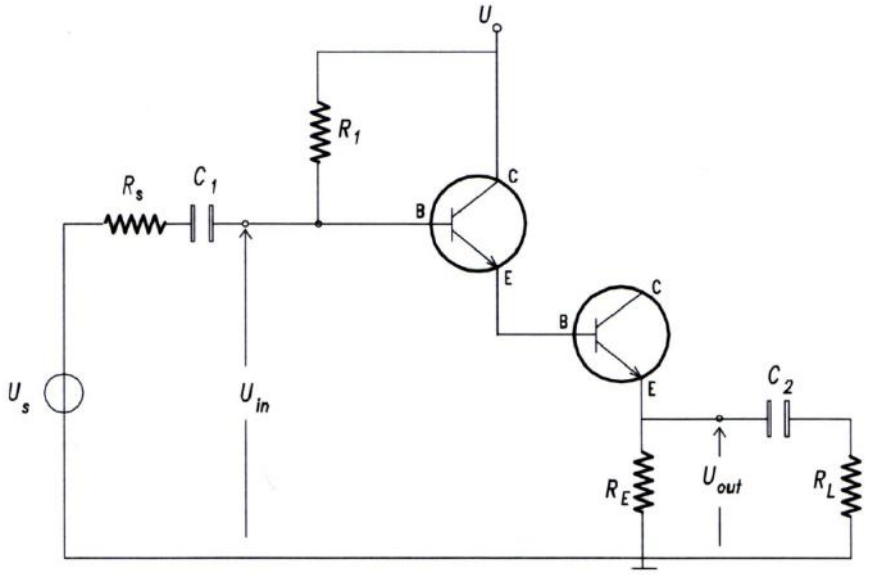
In fig. 17.8 è riportato un circuito più completo.
Il condensatore C_1 serve per eliminare la componente continua dovuta alla polarizzazione: il segnale, essendo variabile nel tempo con frequenza f "vedrà" il condensatore come una modesta reattanza capacitiva e si trasmetterà quindi sulla base del BJT; le correnti continue dovute alla polarizzazione "vedranno" invece il condensatore come un circuito aperto e non potranno interessare la base.
In modo analogo, il condensatore C_2 serve per eliminare dal segnale U_{out} di uscita la componente continua dovuta alla polarizzazione.
Nel circuito, la resistenza R_s rappresenta la resistenza del generatore del segnale da amplificare.



Il circuito della fig. 17.9 propone la struttura dell'amplificatore a collettore comune. Si può dimostrare che esso non amplifica la tensione, ma si limita a riportarla in uscita. Si chiama pertanto inseguitore di tensione (emitter follower). La grande utilità di questo circuito sta nel fatto che la corrente erogata dal generatore di segnale è molto più piccola di quella disponibile su R_E; in altre parole, questo circuito è un grande amplificatore di corrente. Se R_l è molto più grande delle altre resistenze, l'amplificazione di corrente è pari al fattore h_{FE}+l.



Una applicazione dell'inseguitore di emettitore è la cosiddetta connessione Darlington, illustrata in fig. 17.10. Si tratta di due emitter follower collegati in cascata. In tal modo, la corrente che entra nella base del primo transistor è amplificata nel rapporto (h_{FE}+l) ed altrettanto avviene nel secondo transistor. Essendo h_{FE} dell'ordine del centinaio ne derivano rapporti di amplificazione della corrente dell'ordine delle decine di migliaia.



Perciò, con pochi milliampere in ingresso si possono avere alcuni ampere in uscita. I due transistor non sono uguali: il secondo deve essere adatto a elevate correnti. In commercio è reperibile una vasta gamma di componenti connessi in configurazione Darlington, con rapporti di amplificazione da 1000 a 5000. L' applicazione tipica è lo stadio finale dell'amplificazione di potenza per determinati attuatori (ad esempio altoparlanti).

Per concludere accenniamo brevemente all'amplificatore a base comune. Esso si dimostra un buon amplificatore di tensione (paragonabile al transistor ad emettitore comune), mentre non amplifica per nulla la corrente.
Esso presenta il difetto di avere una bassa impedenza di ingresso, e pertanto carica eccessivamente il circuito a monte facendogli erogare una eccessiva corrente; ed inoltre un'alta impedenza di uscita, e pertanto fornisce una modesta corrente al circuito a valle. Per queste caratteristiche, questa configurazione è utilizzata solo in casi particolari.

Banda passante

Gli amplificatori visti in precedenza erano basati sull'impiego del transistor come regolatore di corrente, che è in ogni caso unidirezionale (dal collettore all'emettitore negli npn). L'elaborazione di segnali alternati è stata possibile grazie al "trucco" della polarizzazione, cioè sovrapponendo al segnale alternato una componente continua in modo che durante il funzionamento le tensioni e le correnti si abbassano e si innalzano ma restano sempre dello stesso segno.
Nei circuiti considerati i due condensatori li abbiamo considerati perfettamente trasparenti ai segnali alternati. Ma, in realtà, essi oppongono una reattanza che è funzione della frequenza del segnale alternato:

$$X_c = 1 / (2\pi * f * C)$$

Alle basse frequenze (caso ad esempio dei segnali acustici, con frequenze da 20 a 16000 Hz) la reattanza può assumere valori tali da provocare una attenuazione del guadagno G dell'amplificatore, definito come rapporto tra le tensioni di uscita e di ingresso:

$$G = U_{out} / U_{in}$$

La funzione che esprime l'andamento del guadagno in funzione della frequenza è spesso diagrammata in scala logaritmica. Ancora meglio, l'unità di misura del guadagno è il decibel, definito come:

$$1 \text{ dB} = 20 \log G$$

In tal modo, una attenuazione dal valore normale ad un valore pari al 70% di quello implica un abbassamento di: .

$$20\log (0{,}707G/G) = 20\log 0{,}707 \equiv -3 \text{ dB}$$

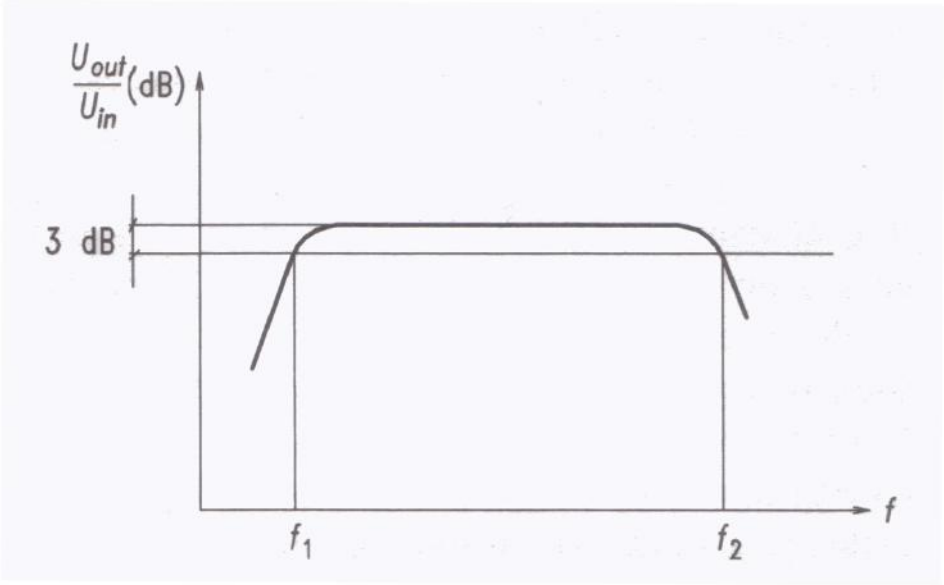
Si definisce banda passante di un amplificatore l'intervallo di frequenze entro le quali il guadagno non subisce una attenuazione superiore a 3 dB. La banda passante è definita da due valori di frequenza, chiamati frequenza di taglio inferiore f_1 e superiore, f_2 (Fig17.11).
Per quanto riguarda l'attenuazione alle alte frequenze, la loro origine va individuata nella capacità parassita che si determina nella giunzione BC polarizzata inversamente, che sappiamo si comporta come un condensatore.

Un comportamento come quello descritto è indicato come filtro passa alto, perché presenta guadagno costante per frequenze elevate (comunque non superiori alla f2).

17.5 Transistor BJT come amplificatore di potenza

Nel campo dei piccoli segnali, l'importanza dell'amplificatore a BJT è andata progressivamente diminuendo, a vantaggio di nuove soluzioni integrate, più efficienti e più stabili.

I transistor BJT mantengono invece un ruolo di primaria importanza nel campo degli amplificatori per alte potenze..

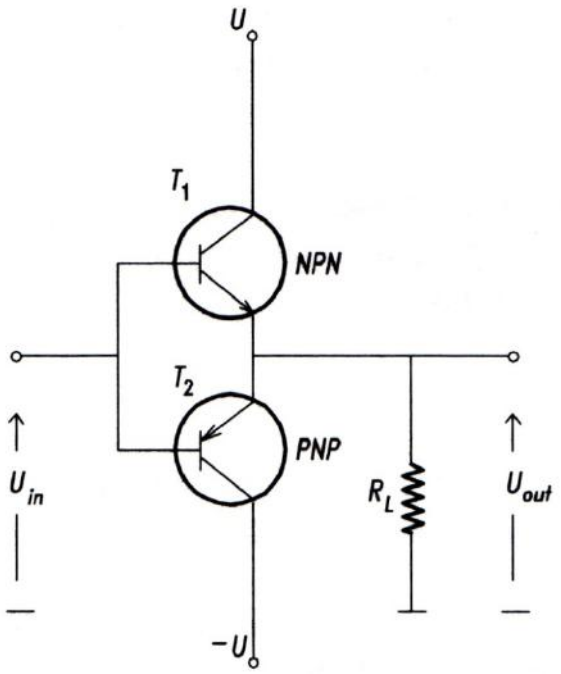


I circuiti integrati forniscono infatti forti guadagni di tensione ma associate a correnti sempre molto deboli.
Un amplificatore integrato, ad esempio, non potrebbe mai fornire la potenza necessaria al comando di un motore in corrente continua, se non nel caso di piccola potenza.
Esistono in commercio BJT di potenza in grado di erogare potenze anche dell'ordine delle centinaia di watt, in grado di pilotare una vasta serie di attuatori (piccoli motori elettrici, solenoidi, elettrovalvole, altoparlanti ecc.).
In generale, gli amplificatori di potenza (non solo quelli a BJT) si classificano in due grandi categorie:

- in classe A: quando il BJT è permanentemente in conduzione;
- in classe B: il transistor conduce solo per ogni metà del periodo del segnale;
- in classe C: il transistor conduce per meno di un semiperiodo del segnale applicato.

Appartengono alla classe A gli amplificatori a BJT del tipo di quelli visti per piccoli segnali, con la differenza che qui essi assumono valori più elevati (correnti dell'ordine di alcuni ampere).
Si può dimostrare che la massima potenza utile è pari al 25% di quella fornita dal generatore. Queste configurazioni lavorano perciò con rendimento massimo del 25%.
Ora, nel caso di piccoli segnali, questa perdita di energia non è significativa, ma non altrettanto può dirsi nel caso di grandi potenze.
Questo modo di amplificare il segnale, interamente e con piccola distorsione, ma con modestissimo rendimento è detto classe di amplificazione di tipo A.

I dispositivi destinati a funzionare in classe B sono polarizzati in modo che il punto di riposo coincida con il punto di interdizione: in tal modo, la potenza dissipata a riposo è nulla (in classe A sarebbe pari alla metà della potenza fornita dalla esterna) e la U_{CE} può compiere una escursione massima, ma metà semionda sarebbe perduta. Per recuperarla si dispone un secondo BJT pnp (se il primo era npn) in serie, come mostrato nella fig. 17.12. Il segnale da amplificare è applicato ad entrambi i transistor, i quali però conducono in fasi successive: la semionda positiva mette in conduzione T_1 mentre interdice T_2, la semionda negativa fa il contrario.



Il funzionamento dell'amplificatore in classe B presenta la cosiddetta distorsione di crossover (di attraversamento). I due BJT passano continuamente dallo stato di conduzione a quello di interdizione. Tuttavia, il passaggio in conduzione avviene solo dopo che la tensione di comando supera il valore di soglia tipico delle giunzioni pn (= 0,6 V per il silicio).
Questo significa che una parte del segnale di ingresso, e precisamente tutta quella parte il cui valore è al di sotto di tale soglia, non è trasmesso all'uscita a causa del prolungarsi dell'interdizione dei due BJT.

Il segnale di uscita risulta così sensibilmente distorto rispetto alla forma d'onda d'ingresso.
Per eliminare la distorsione di crossover di solito si polarizzano le basi dei due BJT con un leggero potenziale, in modo da tenere i due dispositivi sempre sulla soglia della conduzione.

17.6 Transistor BJT come interruttore (classe C)

Nelle applicazioni descritte in precedenza (classi A e B) il punto di lavoro si muoveva lungo la retta di carico, attraversandone tutti i punti in modo che le tensioni e le correnti potessero assumere una infinita gamma di valori compresi tra un massimo e un minimo (in ogni caso tali che il loro prodotto non fosse superiore alla massima potenza dissipabile).

Nel funzionamento come interruttore, detto in classe C, il componente salta dallo stato di interdizione allo stato di saturazione in un tempo che, seppur finito, è tanto breve da essere trascurabile.

Il tipo npn è solitamente preferito proprio perché commuta più rapidamente del pnp, e inoltre perché è di costruzione più semplice.
Come sappiamo, in interdizione la corrente I è nulla e perciò anche se la tensione U_{CE} può assumere valori anche abbastanza grandi (sempre entro il limite della tensione di breakdown U_{BR}), la potenza dissipata è nulla.

In saturazione la tensione U_{CE} è prossima allo zero e pertanto di nuovo la potenza dissipata è nulla anche se la corrente di saturazione raggiunge valori elevati. Pertanto, in questo tipo di funzionamento, il transistor ha un rendimento teoricamente pari al 100%. In pratica, è prossimo a questo valore.

Nel funzionamento in classe C a rigore non si può parlare di amplificazione, se non nel senso che con una piccolissima potenza spesa si può controllare, in modo tutto/niente una potenza più grande.

Il controllo a due stati (on-off, tutto - niente) di attuatori (motori elettrici, cilindri e motori pneumatici ed oleodinamici, solenoidi, altoparlanti, LED di segnalazione ecc.), è alla base del vasto settore dell'automazione industriale binaria: elettrica, elettropneumatica, elettrooleodinamica.

Un secondo, vastissimo, campo di applicazione è costituito dai circuiti e dispositivi dell'elettronica digitale.

Come già sappiamo, con i transistor fatti funzionare in modo on-off {modo switch) si costruiscono le porte logiche, e pertanto tutte le circuiterie di elaborazione e memorizzazione, realizzate sotto forma di circuito integrato.

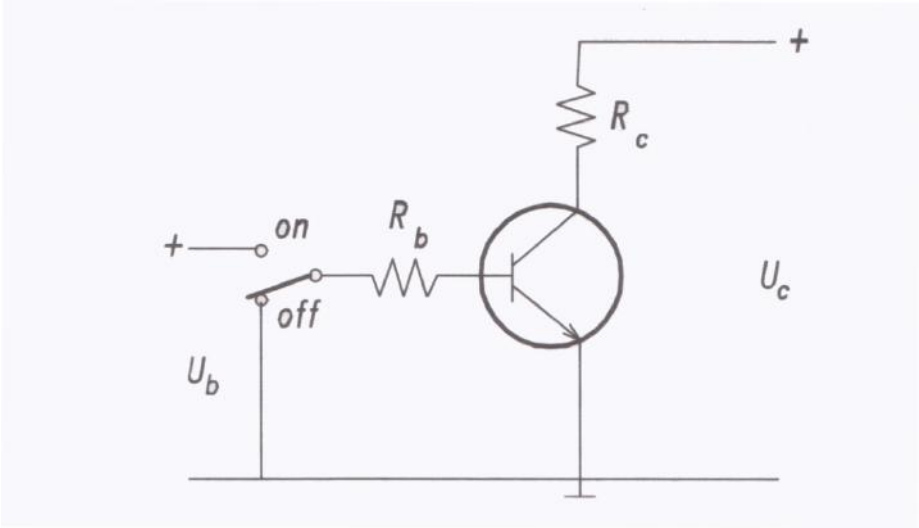
In un microprocessore, ad esempio, lavorano tantissimi componenti (fino a 6 milioni) integrati in una superficie di pochi cm^2. Essi possono stare così ravvicinati proprio perché lavorano a due stati e quindi la potenza dissipata da ciascuno è bassissima.

Altre applicazioni, meno comuni, degli amplificatori in classe C si hanno ad esempio nel campo delle telecomunicazioni e più in generale nel campo delle alte frequenze.

In pratica, per far funzione un BJT in classe C è sufficiente disporre o meno il collegamento della base ad un generatore U_b, come in fig. 17.13.
Nel circuito di polarizzazione sono in genere noti i valori di U_b, U_c e R_c.
Il calcolo del circuito di polarizzazione si riduce pertanto al calcolo della resistenza di base. Il suo valore deve essere tale che si possa generare la corrente di base I_b necessaria a portare in saturazione il componente.
Tale cioè che la I_b possa creare una I_c che, provocando la caduta di quasi tutta la U_c sulla R_c, ponga U_{CE} prossima al valore 0,5 V.

In saturazione, assumendo un certo valore per U_{CEsat}, si calcola la corrispondente corrente di collettore dall'equazione della maglia di uscita:

$$I_{c,sat} = (U_c - U_{CE,sat}) / R_c$$



È dunque necessaria una corrente di base:

$$I_b = I_c / h_{FEmin}$$

essendo h_{FEmin} il minimo valore del guadagno di corrente garantito dal costruttore, e reperibile sul data sheet del componente. Si fa riferimento a questo valore per garantire che il circuito funzioni anche nella condizione più sfavorevole. Dall'equazione della maglia di base è infine possibile calcolare il valore della necessaria resistenza di base, assumendo per U_{BEsat} un valore attorno a 0,6 V:

$$R_b = (U_b - U_{BEsat}) / I_b$$

Nel pilotaggio di carichi in modo on-off, l'attuatore non sempre può essere disposto nel circuito di collettore di un BJT seppur di potenza, perché la potenza messa a disposizione è normalmente dell'ordine dei watt.
Nei casi di maggior potenza, il BJT costituisce l'interruttore di un circuito nel quale è presente il carico ed una adeguato generatore di f.e.m. di alimentazione.
I vantaggi dell'impiego di un transistor come interruttore sono molteplici. In primo luogo, si tratta di un componente completamente statico, e quindi libero da tutti i difetti degli interruttori elettromeccanici costituiti dai relè: ampi tempi di commutazione, rimbalzi dei contatti, usura.
In secondo luogo, esso può essere agevolmente comandato tramite le deboli correnti emesse da circuiti elettronici di controllo quali schede a microprocessore, personal computer, controllori programmabili ecc.
Ritorneremo su questi argomenti più avanti.

17.7 Transistor a effetto di campo

In pratica nella realizzazione dei circuiti integrati interessa risparmiare lo spazio occupato dal componente e minimizzare la potenza dissipata. Per questi motivi, al posto del transistor bipolare, viene di norma utilizzato il transistor a effetto di campo.

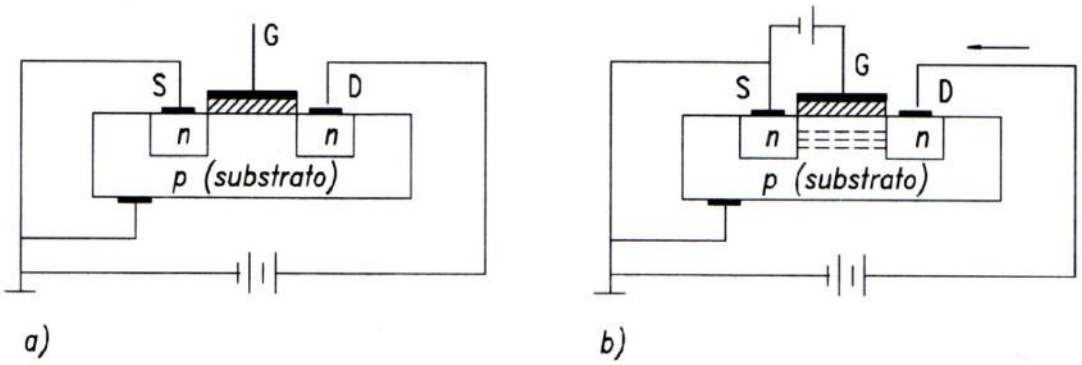
Il transistor a giunzione ad effetto di campo, o JFET (Junction Field Effect Transistor), presenta tre terminali chiamati gate (G, analogo della base del BJT), source (S, analogo dell'emettitore del BJT) e drain (D, analogo del collettore del BJT).

Il JFET modula la corrente che lo attraversa attraverso l'allargamento od il restringimento di un canale di passaggio: a seconda del suo drogaggio si hanno i tipi detti a canale n (analogo del tipo npn BJT) e a canale p (analogo del tipo pnp BJT). La tecnologia di realizzazione più importante è indicata con la sigla MOSFET, da Metal-Oxide-Semiconductor FET. Di essa sono disponibili due versioni: a svuotamento (depletion) e ad arricchimento (enhancement).
La versione enhancement a canale n (NMOS) si costruisce da una lastrina di silicio di tipo p, sulla quale si realizzano due piccole isole tipo n, collegate ai terminali S e D; il terminale G è isolato dal substrato p mediante ossido di silicio (di qui il nome: metal-oxide-semiconductor).
Lo stato di ossido di silicio si può considerare come il dielettrico di un condensatore: infatti il metallo ed il substrato p sono due conduttori; ed il gate è collegato ad una delle armature di questo condensatore.

Se gli elettrodi di source e di substrato (attraverso una porta di massa = bulck gate) sono collegati a massa mentre il drain è collegato ad un potenziale positivo (fig. 17.14a), non si genera alcun percorso conduttivo, perché source e drain formano col substrato due giunzioni in opposizione.
Se al gate è applicato un potenziale positivo rispetto al source ($U_{GS} > 0$), la tensione "carica" il condensatore e gli elettroni liberi (cariche di minoranza) presenti nel substrato sono richiamati in direzione del gate (mentre le cariche positive sono allontanate). L'accumulo di cariche negative crea una zona di tipo n, cioè un canale di collegamento tra le isole di tipo n, la cui ampiezza dipende dalla tensione U_{GS}.
A questo punto, l'applicazione di una tensione U_{DS} tra drain e source produce una corrente I_D analoga alla corrente di collettore di un BJT. A parità di tensione U_{DS}, l'intensità della I_D è proporzionale alla tensione U_{GS}.

Il parametro che indica il guadagno dei MOSFET è la transconduttanza, definita come rapporto tra corrente di drain e tensione gate-source:

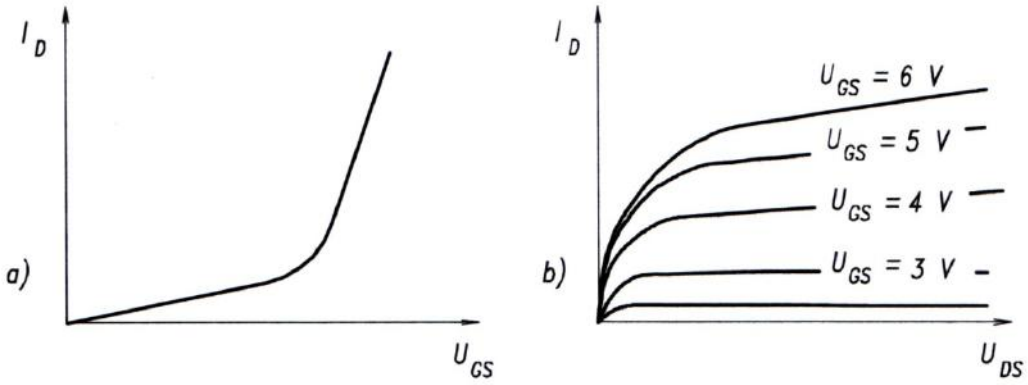
$$g_m = I_D / U_{GS}$$



Come per tutti i dispositivi a semiconduttore, essa è variabile con la temperatura e con le condizioni di funzionamento.

Dalla curva caratteristica in fig. 17.15a si nota come la corrente di drain sia trascurabile finché la tensione di gate si mantiene al di sotto di un determinato valore di soglia, tipicamente pari a 1,5 V. Le curve caratteristiche di uscita, illustrate in fig. I7.I4b, mostrano anch'esse un comportamento del tutto analogo a quello di un BJT, a patto di scambiare la tensione U_{GS} con la corrente di base e la tensione U_{DS} con la tensione collettore - emettitore.

A differenza di quanto accade in un BJT, dove il terminale di controllo, cioè la base, assorbe corrente, in un MOS il terminale di controllo costituito dal gate ha una natura capacitiva (cioè presenta una resistenza di ingresso infinita) e pertanto non assorbe alcuna corrente continua. Nello stato di conduzione il transistor si comporta come una resistenza dell'ordine dei 1000Ω.



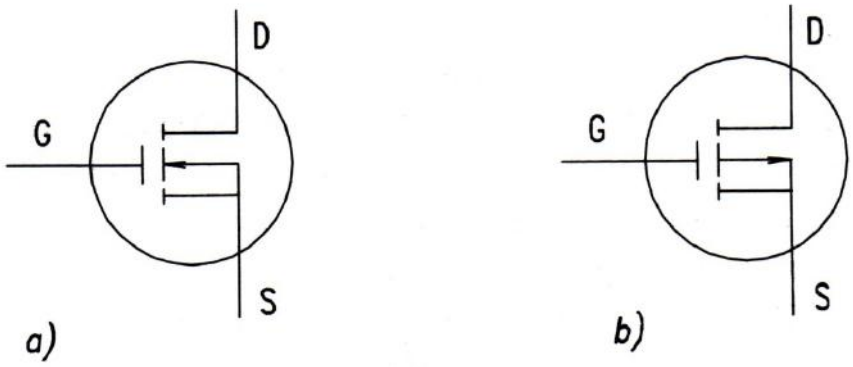
Questo comportamento rende i transistor MOS particolarmente indicati per la realizzazione di circuiti integrati a grande e grandissima scala di integrazione, dove invece i BJT sono da escludere perché darebbero luogo a gravissimi problemi di surriscaldamento.

L'apparire dei transistor MOS ha, tra le altre cose, favorito la grande diffusione di calcolatrici tascabili e di orologi digitali.

Il transistor qui discusso è detto ad arricchimento perché il canale conduttivo è costruito arricchendo di elettroni liberi la zona di substrato compresa tra le due isole N.

Nel transistor PMOS ad arricchimento le due isole sono di tipo p ed il substrato di tipo n. Il canale conduttivo si ottiene arricchendo il substrato di cariche positive (lacune). La fig. 17.16 mostra i simboli normalmente impiegati per i due tipi di transistor.
La linea tratteggiata ricorda che non esiste canale conduttivo tra drain e source senza l'applicazione di una tensione opportuna di gate (positiva nel tipo NMOSFET e negativa nel PMOSFET). La separazione tra il terminale di gate e gli altri due terminali indica che tra gate e canale esiste nello stato OFF del dispositivo una resistenza elevatissima (dell'ordine di 10^{10}Ω), dovuta allo strato di ossido di silicio.

Poiché la corrente è dovuta solo agli elettroni (solo alle lacune nel tipo a canale p), il transistor MOSFET è detto unipolare.



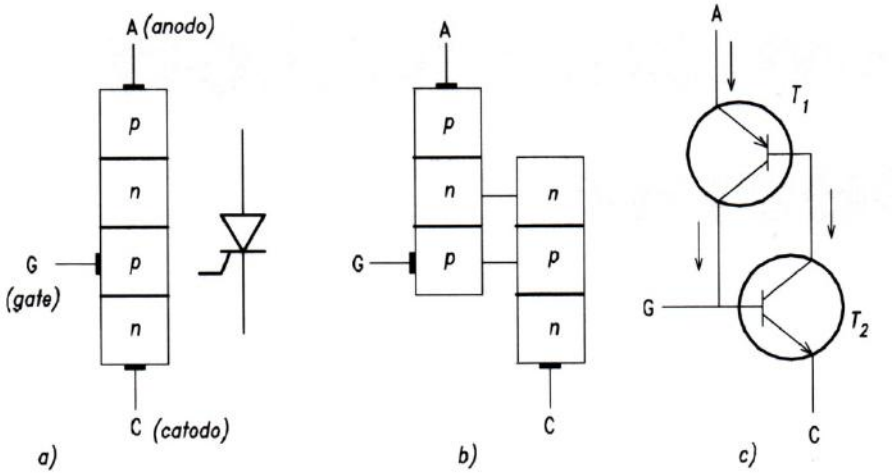
Semiconduttori speciali

18.1 Diodi multigiunzione

I cosiddetti diodi multigiunzione sono costituiti da quattro strati di materiale semiconduttore drogati alternativamente p e n.
Negli ultimi decenni sono stati realizzati e perfezionati dispositivi con una struttura simile a quella dei diodi multigiunzione, ma che presentano dei terminali di controllo collegati agli strati interni del dispositivo. Questi dispositivi formano la famiglia dei tiristori, nome che deriva da una parola greca che significa "porta". Appartengono a questa famiglia tre importanti dispositivi: SCR, GTO e TRIAC, impiegati nel campo delle alte potenze.

I rettificatori controllati al silicio, o SCR (Silicon Controlled Rectifier) presentano la struttura ed il simbolo convenzionale illustrati in fig. 18.1 a: dei tre terminali quello denotato con la sigla G (gate) è il terminale di controllo.
Per meglio comprenderne il funzionamento conviene fare riferimento ai modelli equivalenti illustrati in fig. 18.1b e fig. 18.lc.

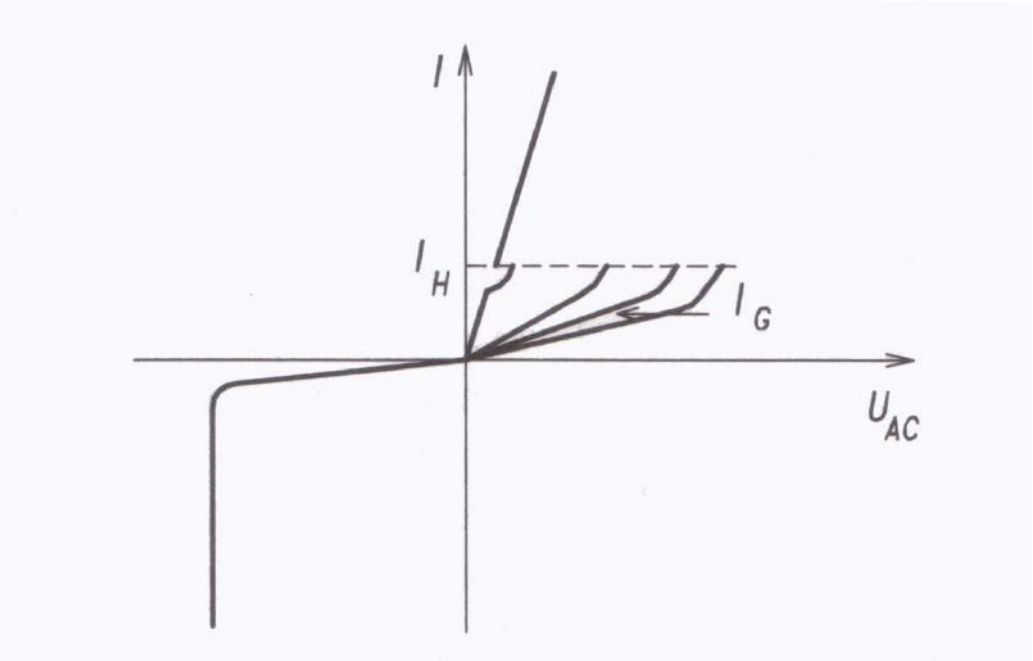
Applicando una tensione negativa tra anodo e catodo, il dispositivo rimane in stato d'interdizione, anche se è applicata una tensione negativa tra gate e catodo.
Se invece si applica una tensione positiva tra anodo e catodo ed inoltre si applica una tensione positiva tra gate e catodo, il dispositivo si porta nello stato di conduzione, presentando una debole resistenza al passaggio di corrente.
Infatti, la tensione U_{GC} è tale da portare in saturazione il transistor T_2 (npn): ciò provoca una corrente di base nel transistor T_1 (pnp) che è portato anch'esso in saturazione; la corrente di collettore di questo può essere pertanto applicata alla base di T_2 confermandone lo stato di saturazione.



Come si vede, una volta innescato, lo stato di saturazione dei due transistor si autosostiene, così che può anche essere soppressa la tensione di gate.

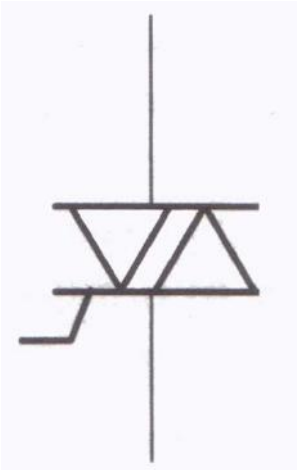
In altre parole, l'SCR si comporta come un normale diodo con la differenza che ora la polarizzazione diretta non è più sufficiente alla conduzione: é necessario che in concomitanza sia applicata una tensione positiva, anche di tipo impulsivo, tra gate e catodo.

L'unico sistema per riportare in interdizione l' SCR è quello di annullare, anche per un brevissimo istante, la corrente che vi circola. In pratica, abbassarla al di sotto del limite di automantenimento (I_H in fig. 18.2).
La tensione U_{GC} di innesco assume valori differenti secondo il valore della corrente di gate: più alto è il valore della corrente e più bassa è la tensione necessaria all'innesco dell' SCR. Abbiamo così una famiglia di curve dipendenti dal valore della corrente di gate, illustrate in fig. 18.2.



Il GTO (Gate Turn-Off) ha una struttura del tutto simile a quella di un SCR, con la differenza che essendo costruito con una particolare tecnologia avanzata, può essere "spento" tramite un impulso negativo applicato tra gate e catodo.
Il GTO può sopportare sovraccarichi notevoli di tensione e corrente (2500 V e 1000 A) e si comporta praticamente come un interruttore ideale.

La struttura del TRIAC (TRiode Alternate Current) è riconducibile a due SCR in antiparallelo come risulta dal simbolo convenzionale della fig. I8.3, e pertanto presenta la caratteristica di poter essere portato in conduzione in entrambe le direzioni.



Può essere polarizzato con tensioni di gate sia positive sia negative, perché in tal modo entra in conduzione l'uno o 1'altro dei due SCR. Anche il TRIAC presenta i problemi di disinnesco dell' SCR.

La fondamentale applicazione del TRIAC è come interruttore a stato solido: portandolo in stato ON può essere chiuso un circuito comprendente utilizzatori in corrente alternata, anche di notevole potenza (nei quali non possono essere impiegati i transistor BJT). In questo impiego, il TRIAC sostituisce vantaggiosamente i contattori ed i relè elettromagnetici, che presentano diversi svantaggi legati alla costituzione elettromeccanica: ritardo nella commutazione (perciò scarsa frequenza di lavoro), falsi contatti, usura ecc.

18.2 Raddrizzatori ca/cc

Al capitolo precedente abbiamo visto come una delle fondamentali applicazioni dei diodi di potenza sia di raddrizzare (rettificare) una tensione alternata, producendo una tensione impulsiva che può essere poi opportunamente livellata fino a farla diventare continua e di valore dipendente dalle caratteristiche della tensione alternata di ingresso (ed anche dei parametri del filtro di livellamento).

Vi sono tuttavia delle applicazioni in cui è necessario controllare il valore della tensione.
Il caso tipico è costituito dall'alimentazione dei motori in corrente continua. Come sappiamo, la velocità di rotazione di questi motori è direttamente proporzionale (a parità di flusso induttore) alla tensione continua applicata alle spazzole. Per ottenere una velocità variabile è dunque necessario poter inviare alle spazzole una tensione continua variabile. Inoltre, brusche frenature e inversioni di marcia sono possibili invertendo le polarità della tensione di alimentazione.

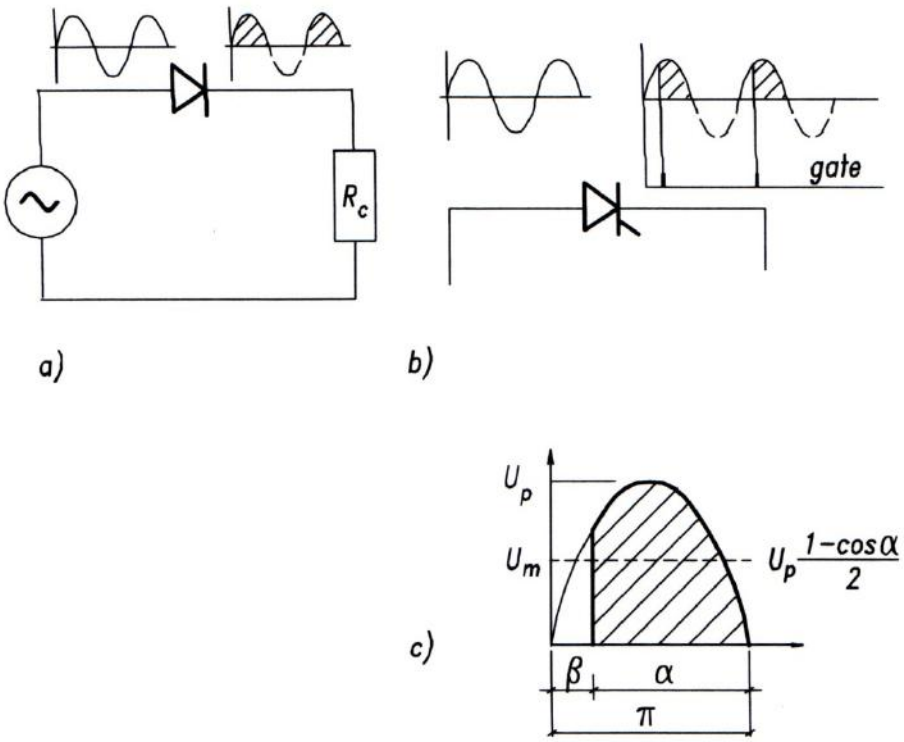
I moduli di alimentazione che impiegano diodi controllati al silicio al posto dei diodi normali offrono una efficace soluzione a questi problemi.

In fig. 18.4a è stata richiamata la struttura del raddrizzatore a semionda che impiega un diodo normale. La corrente arriva dunque all'utilizzatore (ad esempio il motore cc) in forma pulsante e con un valore medio costante.
Utilizzando invece un diodo SCR (fig. 18.4b), con impulsi di gate (prodotti da un apposito generatore di impulsi) con la stessa frequenza della corrente alternata e diversi valori di sfasamento, è possibile ottenere diversi valori di U (vedi fig. 18.4b).

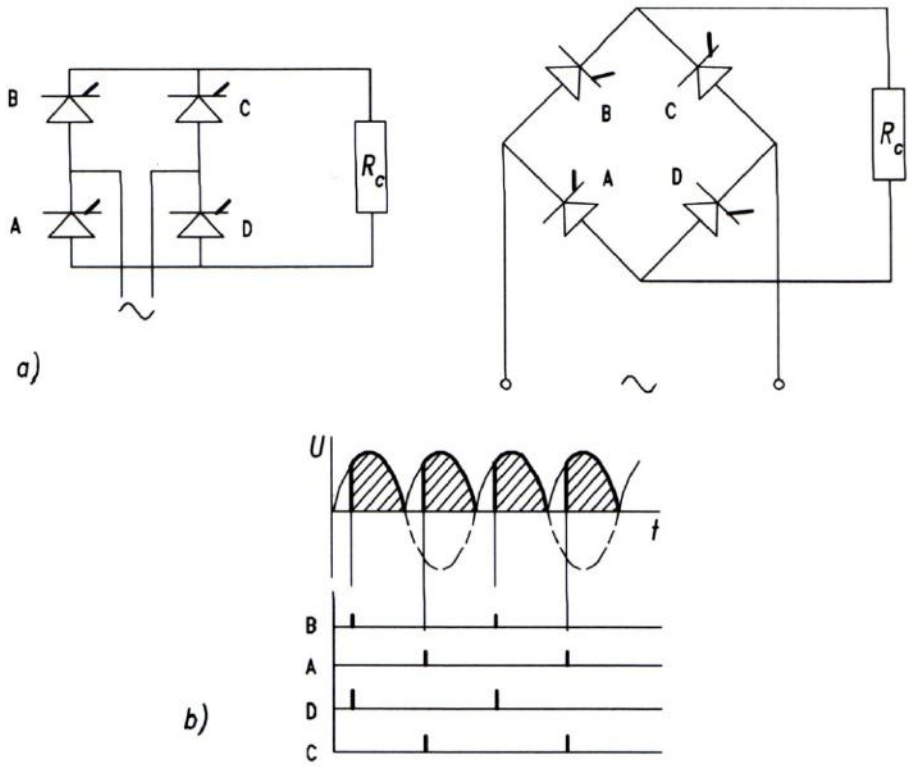
Considerando carichi puramente resistivi, con riferimento alla fig. 18.4c, si ha:

$U_m = U_p (1 – \cos\alpha) / 2\pi$
.
nella quale essendo β = angolo di innesco e α = angolo di conduzione =180° - β. Variando l'angolo di conduzione si ha dunque un valore della tensione di alimentazione variabile da zero al valore massimo U_p / π (per α =180°, corrispondente a quello che si ottiene col normale raddrizzatore a semionda).
Il campo di funzionamento soddisfacente dei diodi SCR arriva fino alla tensione di 500 V con una intensità di corrente di 1000 ampere.
Lo spegnimento dell' SCR non pone problemi: avviene automaticamente al sopraggiungere di ogni semionda negativa.

Il raddrizzatore a doppia semionda può essere semicontrollato o totalcontrollato: rispettivamente, nel ponte di Graetz vi saranno due SCR oppure quattro SCR.
In entrambi i casi si ottengono le stesse forme d'onda, ma col secondo è possibile anche il funzionamento a recupero di energia: quando è il carico a generare energia elettrica (ad esempio nelle fasi di frenatura), il ponte totalcontrollato permette il flusso di energia "all'indietro", verso il generatore.

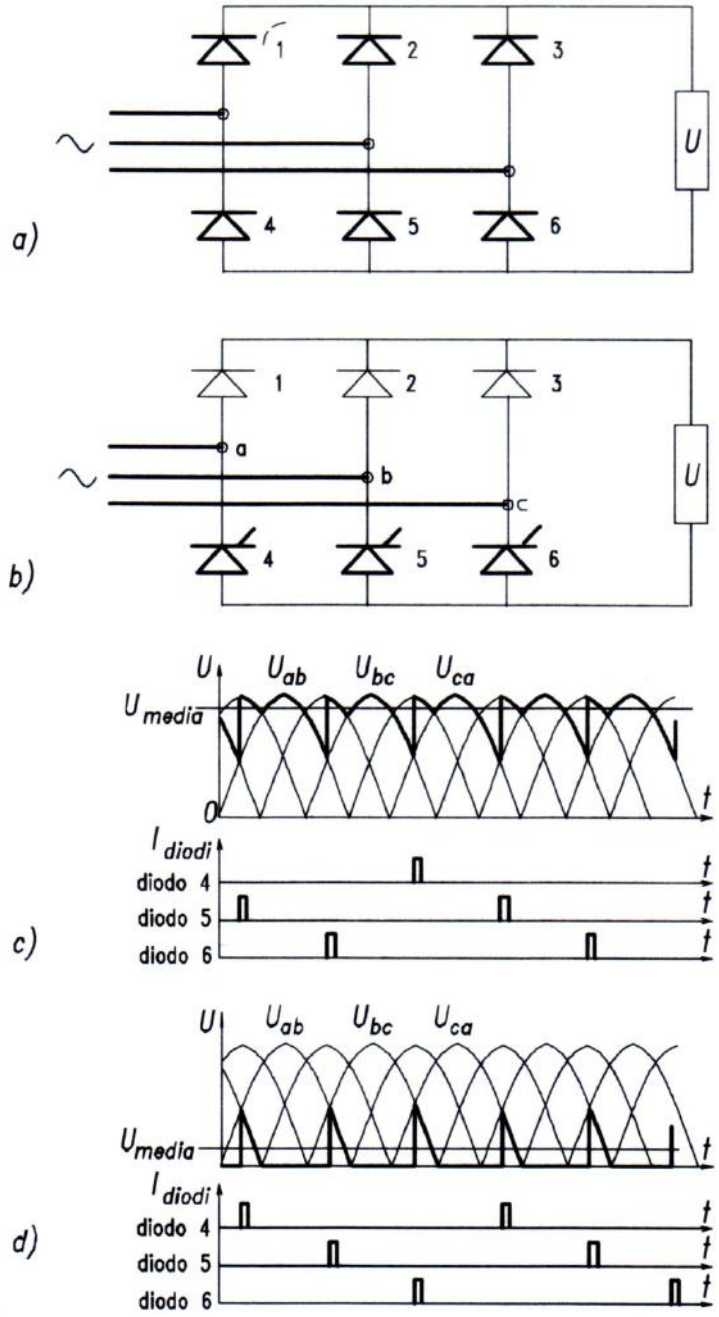


Come ben si comprende, nel caso del ponte totalcontrollato (fig. 18.5a,b), gli impulsi di controllo degli SCR devono essere inviati contemporaneamente ai due diodi nei rami opposti del ponte durante una alternanza della tensione di alimentazione ed agli altri due durante la seconda alternanza, come ben si vede in fig. 18.5c.



Come sappiamo, oltre i 10 kW di potenza risultano più economici i ponti di Graetz trifase. Il circuito della fig. I8.6a, che utilizza diodi normali, produce una tensione di valor medio fisso. È il circuito più utilizzato quando non sia richiesta la variazione della tensione media in uscita. Presenta infatti una piccola percentuale di armoniche, elevato rapporto potenza raddrizzata/potenza apparente al secondario del trasformatore di alimentazione, quindi minimo dimensionamento di questo.
La variazione della tensione media in uscita dal ponte si ottiene col ponte semicontrollato, che impiega tre diodi normali e tre diodi controllati, secondo lo schema della fig. 18.6b.

Variando l'angolo di innesco dei diodi controllati si può ottenere una tensione continua variabile tra zero ed un valor massimo. Nelle Figg.18.6b e 18.6c sono illustrati due casi particolari, rispettivamente per un angolo di innesco dei diodi di 30° e 150°.



I circuiti raddrizzatori analizzati in precedenza forniscono una tensione media sempre positiva ed una corrente sempre unidirezionale sul carico, stante il modo di funzionare dei diodi, e risulta pertanto

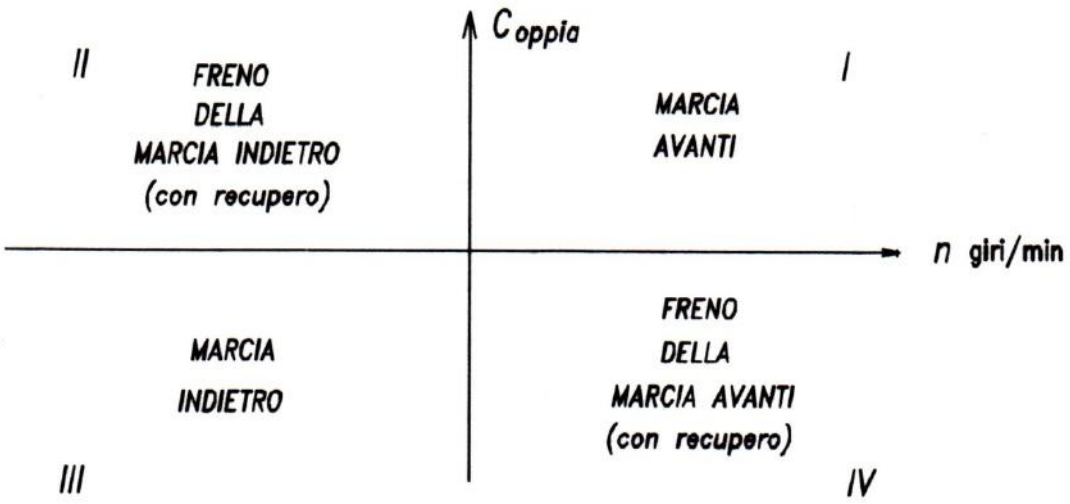
impossibile effettuare la frenatura a recupero, che richiede l'inversione della polarità della tensione di alimentazione.
In altre parole, è possibile soltanto il funzionamento in un solo quadrante (marcia avanti o marcia indietro: vedi fig. I8.7).

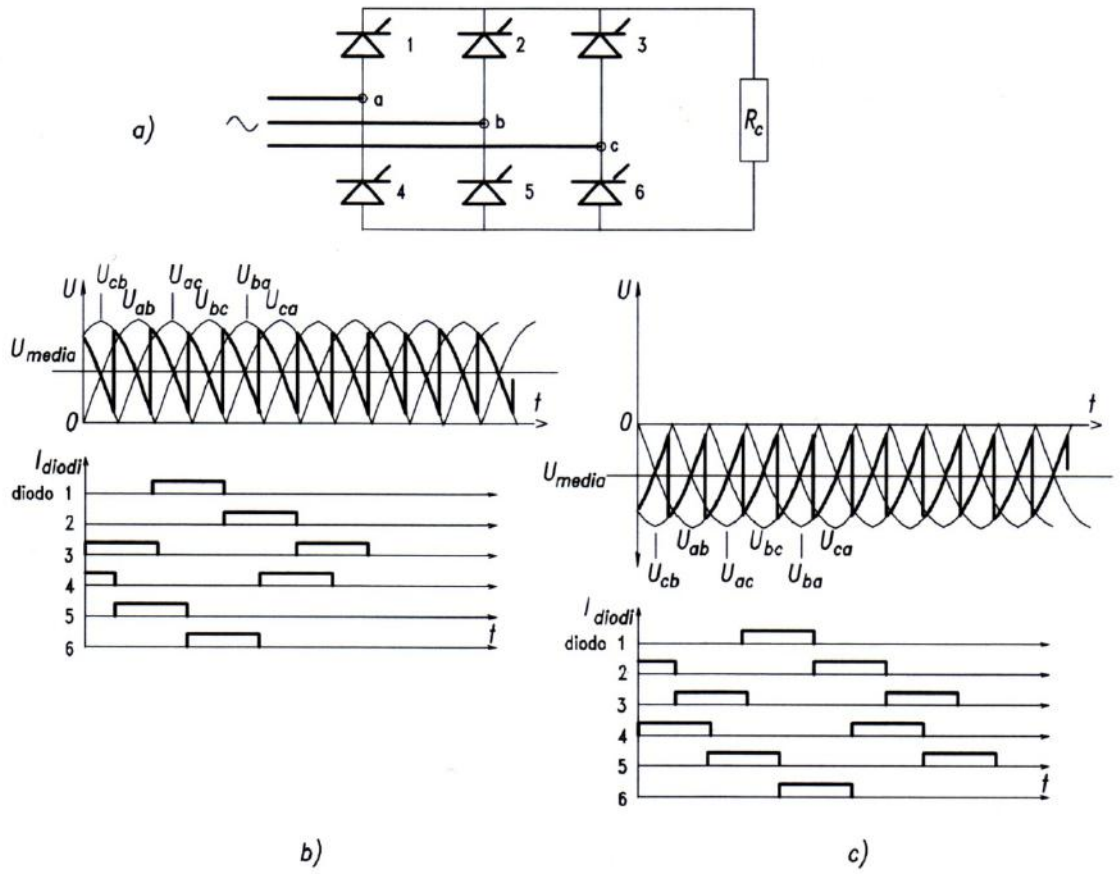
Per ottenere i1 funzionamento su due quadranti (marcia in un senso e frenatura a recupero del senso di marcia opposto: quadranti I - II o III - IV) è necessario ricorrere ad un ponte totalcontrollato, in cui ci tutti i diodi siano controllati, come illustrato in fig. 18.8a.

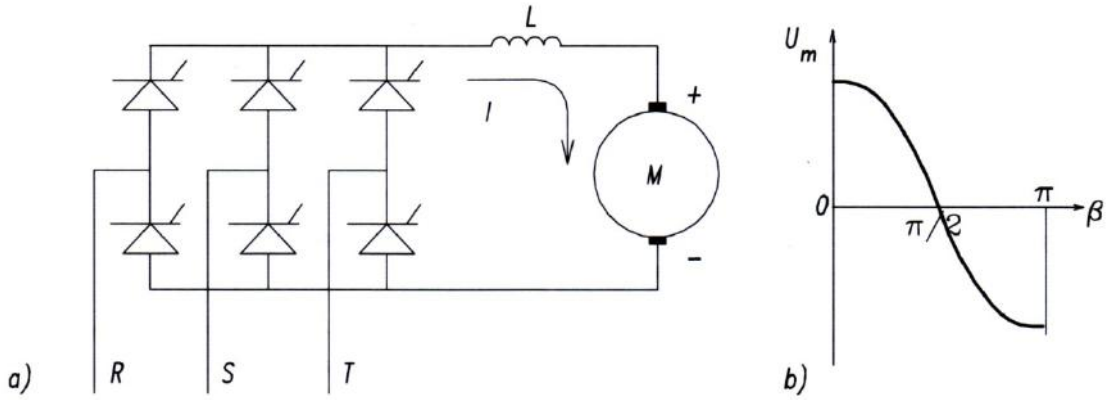
Per angoli di innesco variabili tra 0 e π/2 si ottengono tensioni medie positive, mentre per angoli di innesco tra π/2 e π si ottengono tensioni medie negative. In fig. 18.8b e 18.8c sono illustrati due esempi, rispettivamente per l'angolo di 50° e di 130°.
Il diagramma completo illustrato in fig. I8.9.

Stante la presenza dei diodi, la corrente circolante nel ponte è in ogni caso unidirezionale, e così la coppia che da essa dipende: si ottiene pertanto un funzionamento su due quadranti: I e II oppure III e IV.
Se il carico trascina il motore in senso inverso a quello della sua marcia normale (caso ad esempio del montacarichi), la coppia motrice svolge una azione frenante. Tale azione è resa rigenerativa (o a recupero di energia) facendo in modo che in uscita dal ponte la tensione sia negativa, ciò che si ottiene predisponendo un angolo di accensione maggiore di 90°.
La corrente unidirezionale sul ponte è ora spinta dal carico stesso e smistata ciclicamente su due dei tre conduttori di fase della rete alternata di alimentazione.

Per ottenere il funzionamento sui quattro i quadranti è necessario poter invertire il senso della corrente circolante nel motore, e pertanto la coppia motrice. Si dispongono allo scopo due ponti totalcontrollati in antiparallelo, secondo quanto illustrato







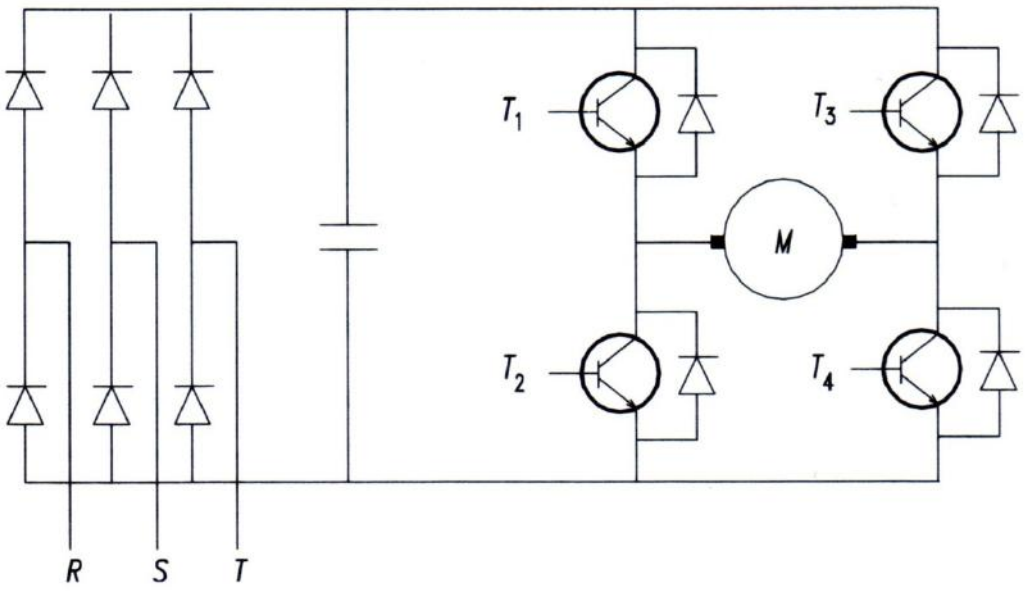
in Fig. 18.10: un ponte assicura la marcia avanti e la frenatura rigenerativa della marcia indietro, mentre l'altro ponte assicura la marcia indietro e la frenatura rigenerativa della marcia avanti.

Come è evidente, il circuito di controllo dei due ponti deve assicurare, oltre che la giusta sequenza di impulsi di gate agli SCR, anche che i due ponti non lavorino mai contemporaneamente, e che entrambi siano spenti nella commutazione dall'uno all'altro, onde evitare cortocircuiti.

Il funzionamento nei 4 quadranti permette di raggiungere la massima velocità in senso diretto e la massima in senso inverso. La frenatura rigenerativa permette il rapido passaggio dall'una all'altra, cioè la rapida inversione del senso di marcia.

Per ottenere le massime velocità di risposta, al posto degli SCR si impiegano i transistor, tuttavia a scapito della potenza massima controllabile.

In Fig. 18.11 è illustrato uno schema semplificato. Il ponte trifase a diodi produce una tensione continua a valor fisso. I quattro transistor (qui sono BJT ma possono anche essere MOS) del cosiddetto ponte ad H producono la circolazione di corrente in un verso (grazie alla contemporanea conduzione dei transistor T_1 e T_4) oppure in quello opposto grazie alla contemporanea conduzione dei transistor T_2 e T_3).
La variazione di tensione si ottiene dall'opportuno pilotaggio dei transistor del ponte ad H. Lo stadio pilota di tale ponte utilizza la tecnica della modulazione della larghezza degli impulsi, o tecnica PWM (Pulse Width Modulation): variando il tempo di accensione dei transistori si ottiene una tensione come una successione di impulsi, positivi o negativi, di larghezza variabile, in modo da ottenere una tensione continua di valor medio variabile.



Ogni transistor presenta un diodo di libera circolazione in parallelo: il diodo protegge il transistor dalle sovratensioni che insorgono ogni volta che passa nello stato di interdizione a causa dell'induttanza del motore.

Tipi di controllo dei motori cc

La discussa alimentazione di un motore cc a tensione variabile realizza un tipo di controllo del motore detto a coppia costante.
La coppia motrice del motore fornita dal motore è data dalla relazione:

$$C_m = K\Phi I$$

Essendo K una costante caratteristica del motore, Φ il flusso induttore e I la corrente circolante nel circuito di indotto o di armatura del motore.
Ora, il flusso è mantenuto costante attraverso una alimentazione costante al circuito di eccitazione (ponte raddrizzatore a diodi normali) oppure ricorrendo alla costruzione con magnete permanente.
L'equazione del circuito del motore si ottiene applicando la legge delle maglie al circuito di indotto:

$$U - E = R_a * I \quad \Rightarrow \quad U - k\Phi n = R_a * I$$

essendo E la forza controelettromotrice indotta. Aumentando la tensione di armatura U aumenta la velocità di rotazione n e pertanto anche la f.c.e.m. E: in tal modo, la corrente che circola nel motore rimane pressoché costante.
Si vede dunque che essendo costanti tutti i parametri da cui dipende la coppia motrice, questa rimane sostanzialmente costante al variare della velocità.

Si noti che nel suo funzionamento il motore adegua in ogni caso, e di volta in volta, la coppia motrice a quella resistente richiesta dal carico. Parlando di azionamento a coppia costante si intende pertanto dire più precisamente: "a coppia massima disponibile costante". Si tratta della coppia che il motore potrebbe erogare disponendo il flusso al suo valore massimo (coincidente di solito col valore nominale).

Vi sono tuttavia altri modi per controllare un motore in corrente continua. Poiché la velocità è data dalla relazione approssimata:

$$n = U / k\Phi$$

si vede che la variazione di velocità si può conseguire anche mantenendo costante la tensione di armatura e facendo variare (tramite ponte a SCR) il flusso induttore (ovviamente nel caso in cui il motore presenti l'avvolgimento induttore e non i magneti permanenti).
In questo caso, la coppia motrice segue la variazione del flusso e la grandezza che si mantiene costante è la potenza erogata dal motore.
Anche qui possono ripetersi considerazioni analoghe a quelle svolte per il controllo a coppia costante e definire il controllo di velocità a flusso variabile come controllo a potenza massima disponibile costante.

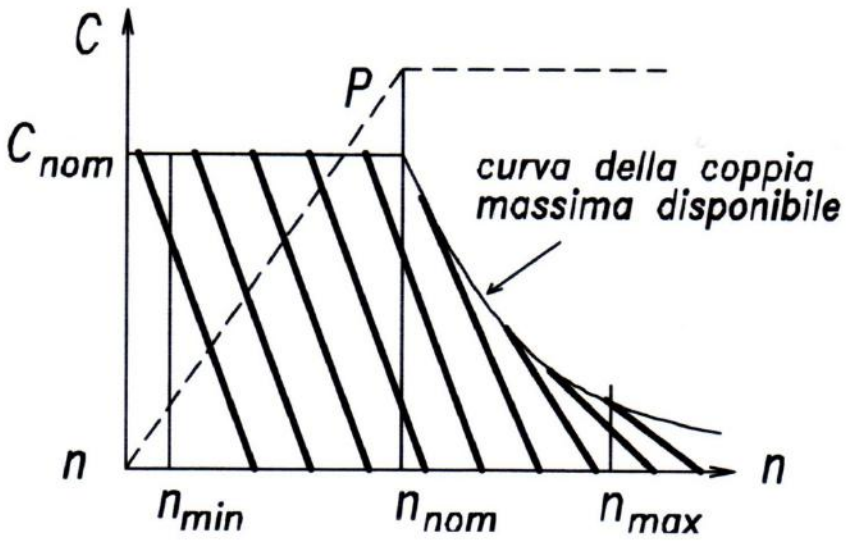
Se il campo di velocità è molto esteso è necessario ricorrere al controllo misto. Per bassi regimi di rotazione si imposta il flusso al valore nominale e si fa variare la velocità aumentando la tensione di armatura. Quando anche questa raggiunge il valore massimo applicabile (nominale), ogni ulteriore aumento di velocità può avvenire soltanto mediante diminuzione del flusso dal valore nominale.
Si vede dunque come un primo campo di regolazione della velocità sia caratterizzato dalla coppia massima disponibile costante e come il successivo sia invece a potenza massima disponibile costante.

Si veda la Fig. 18.12 relativa al primo quadrante. Si noti la variazione delle curve caratteristiche meccaniche.

18.3 Convertitori ca/ca (inverter)

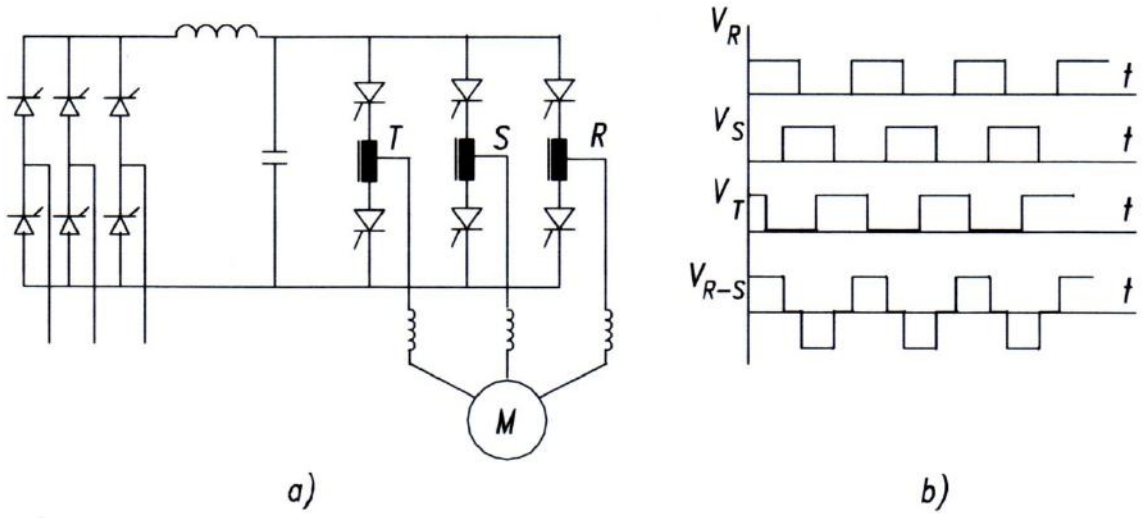
La velocità di un motore asincrono trifase è data dalla relazione:

$$n = 60 * f / p$$



essendo f la frequenza delle tensioni statoriche di alimentazione e p il numero di coppie polari.
La variazione di velocità si può conseguire più facilmente e con continuità variando la frequenza di alimentazione. L'inversione del senso di marcia si ottiene poi invertendo il campo rotante.
Affinché il flusso per polo rimanga costante è tuttavia necessario che la variazione di frequenza sia attuata contestualmente ad una proporzionale variazione della tensione (concatenata) di alimentazione. Deve cioè risultare U / f = costante.
Le descritte operazioni di comando possono essere agevolmente effettuate mediante un modulo di alimentazione chiamato inverter, che presenta una uscita variabile in tensione e frequenza.
L' inverter di tipo tradizionale (introdotto negli anni '60) è costituito da due stadi (Fig. 18.13a). Il primo stadio è un ponte totalcontrollato che genera una tensione continua di valore variabile. Il secondo stadio (1' inverter vero e proprio) è un ponte di SCR che converte tale tensione in una tensione di data frequenza.

Questo tipo di inverter utilizza il metodo dell'onda quadra per il pilotaggio degli SCR, ed è perciò detto inverter PAM.

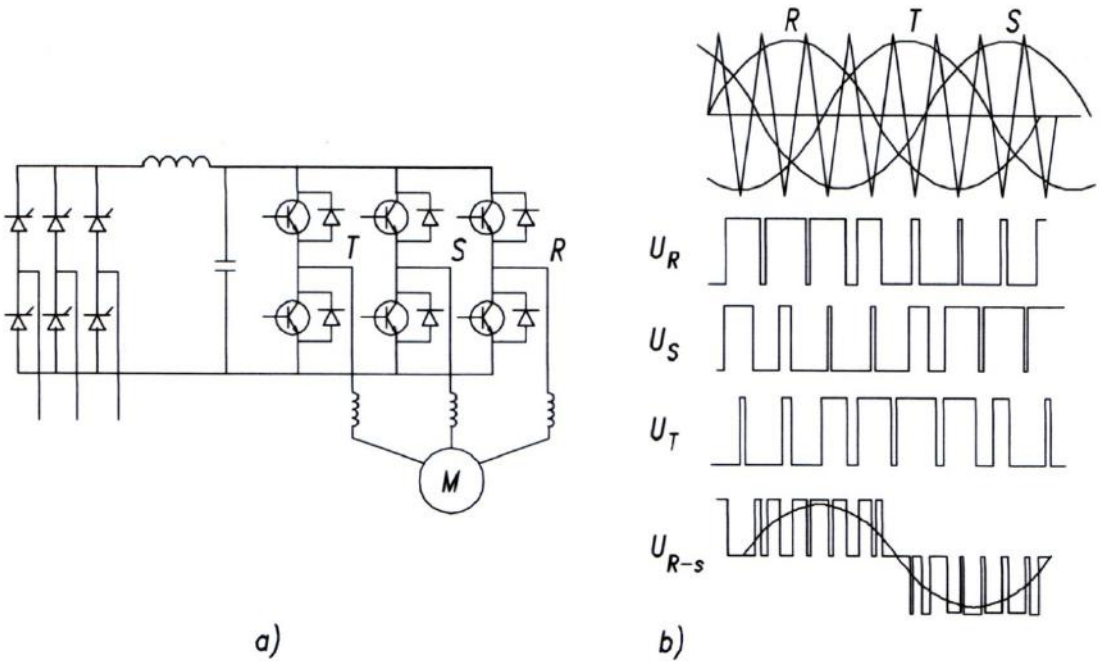


I dispositivi di ciascuna coppia sono fatti condurre alternativamente, rendendo possibile la presenza di tre tensioni di forma rettangolare alle uscite R, S, T.
Dalla differenza algebrica tra la tensione in R ed in S, ad esempio (vedi Fig. 18.13b), risulta la forma d'onda che rassomiglia, gradini a parte, ad una sinusoide.
Se l'innesco di ciascuna coppia avviene con intervalli regolari con uno sfasamento di 1/3 di periodo uno dall'altro, le tre uscite risultano sfasate reciprocamente di 120° (sistema simmetrico di tensioni).
La frequenza delle tensioni può essere variata variando la frequenza dei treni di impulsi di comando agli SCR.

L' inverter PAM è anche detto a tensione impressa perché un grosso filtro capacitivo fissa rigidamente la tensione di alimentazione al secondo stadio, cosicché la tensione di uscita verso il motore non è influenzata dalla natura del carico. Per un maggiore efficienza, il motore deve essere a basso scorrimento.

Questo tipo di convertitore c.a.-c.a. ha avuto il pregio di essere semplice e silenzioso, ma per contro presentava il difetto di scarsa efficienza a causa delle armoniche generate e dalla bassa coppia fornita al motore, coppia oltretutto soggetta ad oscillazioni.

Modernamente, 1' inverter è costruito con lo schema della Fig. 18.14a, con un ponte di transistor pilotati con la tecnica PWM sinusoidale (inverter PWM). In modo analogo a quando spiegato in precedenza (vedi la Fig. 18.14b), è possibile ottenere



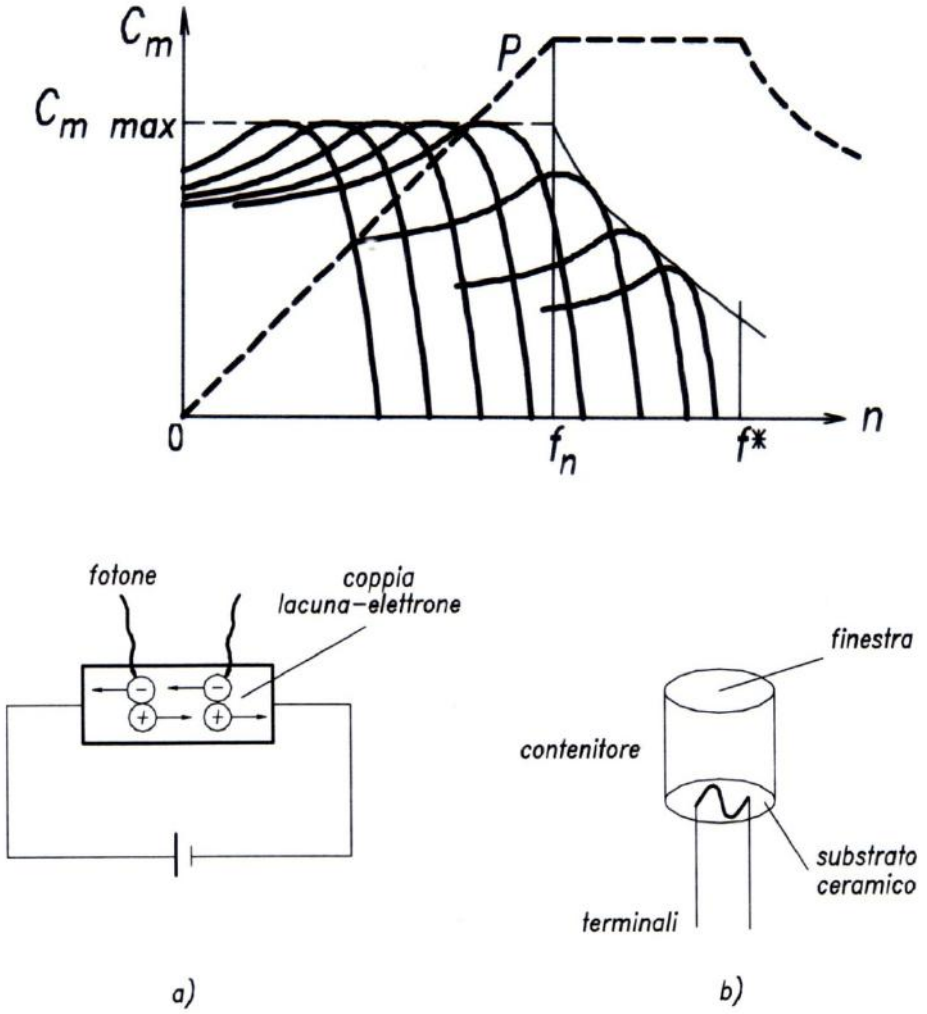
un sistema di tensioni più somiglianti ad un'onda sinusoidale di quanto non consentito dall' inverter PAM.
La doppia regolazione della tensione e della frequenza è demandata alla sezione di controllo dei transistor, la quale controlla elettronicamente entrambe modificando la frequenza e l'ampiezza della modulante.

Attualmente, gli inverter PWM sono i più diffusi, soprattutto per potenze fino a 100 kW, mentre quelli ad onda quadra sono utilizzati nelle applicazioni al altissima velocità (frequenza di inverter >200 Hz).
Modernamente, il controllo della sezione invertitrice è affidato ad algoritmi memorizzati su EPROM (controllo digitale).

Con gli inverter descritti il motore asincrono trifase può lavorare sui quattro quadranti di funzionamento; in altre parole, gli è consentito il funzionamento a recupero di energia nei quadranti II e IV.
Inoltre, con 1' azionamento a U / f con la frequenza che varia da 0 al valore nominale, la coppia massima disponibile rimane costante (le curve caratteristiche meccaniche traslano mantenendo la propria forma). Oltre la frequenza nominale, la diminuzione della corrente statorica e pertanto della corrente rotorica responsabile della coppia non può più essere compensata: la coppia diminuisce in un primo tratto quasi linearmente: in questo tratto si ha pertanto un funzionamento a potenza massima disponibile costante. Superato questo tratto, le reattanze diventano così elevate che la coppia diminuisce più rapidamente e la potenza varia inversamente alla frequenza.
La Fig. 18.15 illustra il controllo misto di velocità nel primo quadrante.

18.4 Fotoresistenze

In un rivelatore fotoresistivo, o fotoresistenza, i fotoni incidenti liberano elettroni dai rispettivi atomi (Fig. 18.16a), producendo coppie lacuna-elettrone libero. Gli elettroni così liberati diventano disponibili alla conduzione elettrica, diminuendo la resistenza del materiale in cui scorrono: in un circuito nel quale è presente un generatore scorre così una corrente che è funzione dell'energia luminosa che investe la fotoresistenza.



L'energia necessaria per rimuovere un elettrone dal proprio atomo dipende dal tipo di materiale. La relazione tra l'energia e la lunghezza d'onda della luce incidente è data dall'equazione:

$E = 12400 / \lambda$

nella quale l'energia E è misurata in elettronvolt (eV) e la lunghezza d'onda λ è misurata in ångstrom (Å).
Considerando lunghezze d'onda nel campo tra i 5000 e 60000Å, risulta una energia compresa tra 0,2 e 2,4 eV.
La Fig. 18.16b illustra una concreta realizzazione del dispositivo, che può essere utilizzato per rivelare una radiazione luminosa incidente, come i fotodiodi.

Dal punto di vista elettrico, la caratteristica più importante di una fotoresistenza è il rapporto tra la resistenza misurata in condizione di oscurità e quella misurata in piena luce. Questo valore è tipicamente di 1000: l.

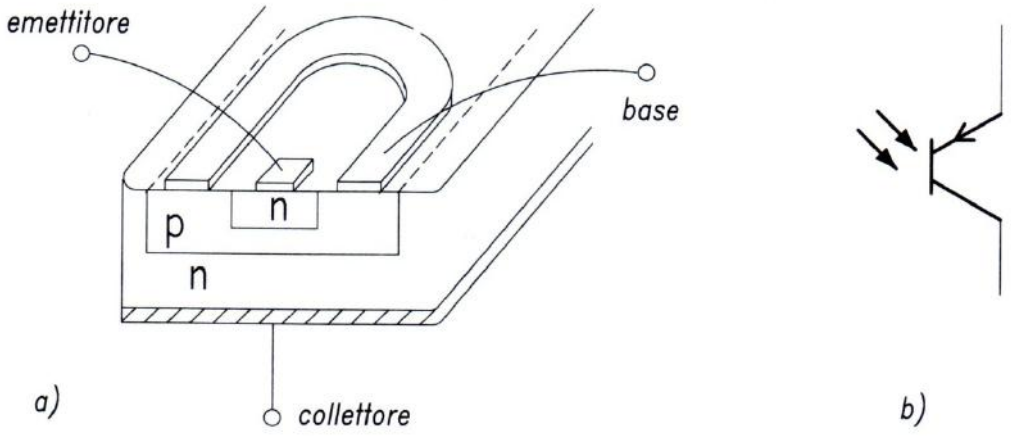
18.5 Celle solari

Se un fotodiodo non è polarizzato, l'esposizione alla luce determina una f.e.m. ai suoi capi, così che se l'anodo è connesso al catodo attraverso una resistenza di carico, nel circuito così formato scorrerà una corrente proporzionale all'intensità luminosa. Un fotodiodo utilizzato per convertire energia luminosa in energia elettrica prende il nome di cella solare.
Poiché il dispositivo non è polarizzato, per rimuovere le lacune e gli elettroni dalla zona di svuotamento è necessario che questa sia la più sottile possibile, altrimenti vi sarebbe ricombinazione al suo interno.
Deve essere inoltre dotato della più ampia superficie di ricezione dell'energia raggiante.
Una superficie elevata ed una sottile regione di svuotamento originano però una alta capacità, così che la risposta del dispositivo alle variazioni dell'intensità luminosa è molto lenta.
Una cella solare è pertanto del tutto inadatta ad essere utilizzata come fotorivelatore nella maggioranza delle applicazioni.
Le celle solari hanno trovato la loro principale applicazione sui satelliti come sorgenti di energia per gli apparati elettrici di bordo.

18.6 Fototransistor

I fotodiodi presentano l'inconveniente di generare deboli segnali elettrici, i quali richiedono pertanto uno stadio di amplificazione a valle.
L'impiego di un fototransistor (Fig. 18.17) risolve il problema, perché esso contiene una giunzione pn che può essere utilizzata come fotodiodo, ed inoltre presenta la possibilità di amplificazione.

La luce incidente sulla giunzione base-collettore provoca una fotocorrente nella base, e questa corrente è amplificata di h_{FE} volte per produrre la corrente di collettore.
Il fototransistor presenta tuttavia una minore prontezza di risposta alle variazioni dell'intensità luminosa.
In molti fototransistor il terminale di base non è accessibile all'esterno, mentre in altri lo è, per permettere un ulteriore controllo del suo funzionamento.



18.7 Fotoaccoppiatori

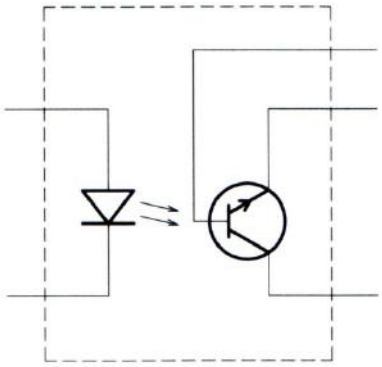
I fotoaccoppiatori sono realizzati unendo in uno stesso contenitore un diodo LED ed un elemento fotosensibile che nella maggioranza dei casi è un fototransistor, ma che può anche essere un fotoSCR, o un fotoTRIAC, elementi nei quali la corrente di controllo è generata fotoelettricamente.
Quando il LED è polarizzato direttamente emette radiazione elettromagnetica che è captata dall'elemento fotosensibile, il quale entra in conduzione.

Il fondamentale vantaggio di questo dispositivo quello di accoppiare due circuiti elettrici con un perfetto isolamento elettrico tra essi.

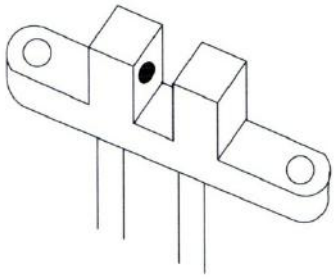
Mutando punto di vista, il dispositivo disaccoppia (o isola) un segnale di ingresso dal circuito che deve elaborarlo: viene perciò chiamato anche optoisolatore.
La tensione massima che può essere applicata senza pregiudicare l'isolamento è dell'ordine dei 7500 V alternati.

Oltre la massima tensione di isolamento, il parametro che caratterizza questo componente è il CMR (Current Transfer Ratio), rapporto percentuale tra corrente emessa dal fotorivelatore e corrente assorbita dal diodo LED.

Non esiste un vero e proprio simbolo grafico per 1' optoisolatore, perché non è un vero e proprio dispositivo quanto piuttosto un insieme di dispositivi. Il simbolo normalmente utilizzato pertanto l'insieme dei simboli dei dispositivi componenti (Fig.18.18). Per lo stesso motivo, la sigla identificativa è formata dall'unione delle sigle: in genere OPT. La Fig. 18.18b riporta un tipico contenitore utilizzato in applicazioni di rivelazione: i due bracci contengono il LED e l'elemento fotosensibile: quando un oggetto (ad esempio un disco forato) interrompe il raggio luminoso emesso dal LED il dispositivo commuta.



a)



b)

Amplificatore operazionale

19.1 Generalità

L'amplificatore operazionale (amp-op) è un amplificatore in continua a più stadi, disponibile sotto forma di circuito integrato, utilizzato per segnali in continua o alternata ed inoltre per effettuare su di essi vari tipi di operazioni (di qui il nome) con l'aggiunta di pochi elementi passivi (soprattutto resistenze).

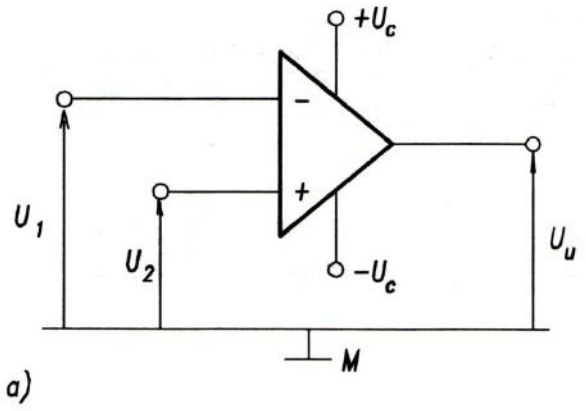
Il dispositivo presenta due ingressi:
- ingresso non invertente, denotato col simbolo "+",
- ingresso invertente, denotato col simbolo "-"

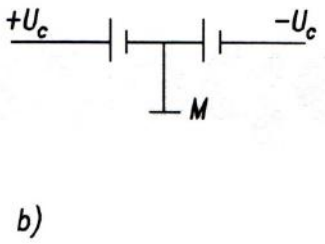
e una uscita (Fig.19.1 a). Di solito, è richiesta una alimentazione duale, con due tensioni uguali e contrarie, che possono essere ottenute da due sorgenti uguali con le connessioni indicate in Fig. l9.lb.

Si definisce guadagno dell'amplificatore operazionale (amp-op) il rapporto tra la tensione di uscita U_u e la tensione differenziale di ingresso ($U_d = U_1 - U_2$):

$$G = U_u / U_d$$

L'amplificatore operazionale ideale presenta le seguenti caratteristiche:
- resistenza di ingresso infinita ($R_{in} = \infty$)
- resistenza di uscita nulla ($R_{out} = 0$)
- guadagno infinito ($G = \infty$).





L'ultima caratteristica implica che l'applicazione di una tensione differenziale anche piccolissima comporta una tensione infinita in uscita. In pratica, poiché la tensione non potrà in ogni caso superare quella di alimentazione UC, l'uscita finisce per assumere il valore U_c oppure - U_c, Si afferma che l'amplificatore va in condizioni di saturazione. La curva caratteristica dell'amplificatore ideale è illustrata in Fig. 19.2a.

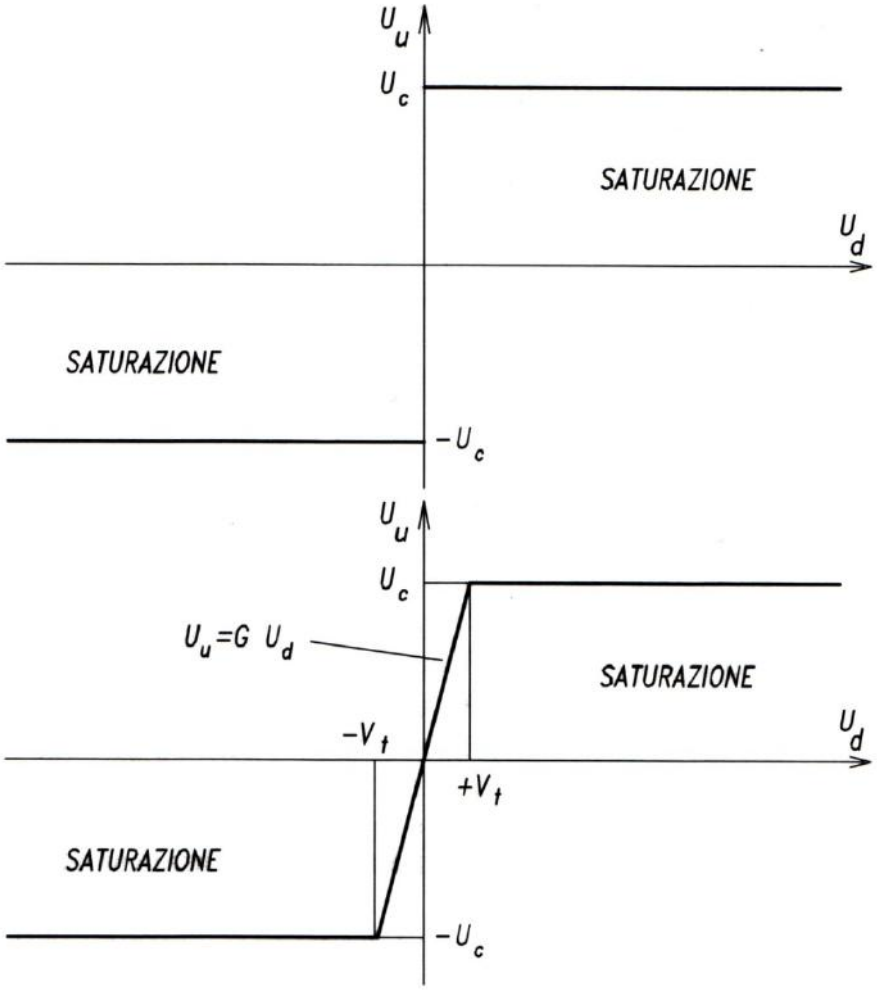
In pratica, il guadagno di un amp-op reale non è infinito.
Se un amp-op reale presenta un guadagno pari a 20 milioni ed è alimentato con una tensione duale di 15 V, raggiunge la saturazione con un ingresso di:

$U_d = U_u / G = 15 / 20000000 = 0{,}75\mu V$

Questo significa che in un intorno di 1,5 μV il comportamento del dispositivo è lineare, così come indicato in Fig. 19.2b.

Per ottenere un comportamento lineare entro un intorno abbastanza grande (dell'ordine di 0,3 V) è necessario ridurre il guadagno, ciò che si ottiene predisponendo una retroazione negativa dell'uscita sull'ingresso invertente.
L'azione dell'uscita sull'ingresso (retroazione) è detta negativa perché il segnale proveniente dall'uscita si sottrae a quello proveniente dall'ingresso: in altre parole, la retroazione viene di norma realizzata sull'ingresso invertente.
Si può dimostrare che il guadagno ad anello chiuso dipende solo dal guadagno del blocco di ritorno.
Poiché il guadagno si mantiene in ogni caso molto elevato, si potrà sempre assumere che le tensioni rispetto a massa dei due terminali di ingresso siano uguali.

L' amp-op è un elemento universale di reti nelle quali il tipo e la posizione degli elementi passivi determinano il tipo di funzionamento.
Vedremo di seguito le configurazioni più comuni dell' amp-op con retroazione negativa, assumendo per semplicità che possa essere assimilato ad un amp-op ideale. Per semplicità, sarà omessa l'indicazione dell'alimentazione dell' amp-op.



Si vedranno infine le applicazioni nelle quali interessa proprio il comportamento non lineare dell' amp-op non retroazionato.

19.2 Amp-op invertente

L' amp-op funziona come amplificatore invertente nella configurazione illustrata in Fig. I9.3a. Applicando la legge di Ohm ai due resistori:

I1 = (Ui – UP) / R1 ; I2 = (UP – Uu) / R2

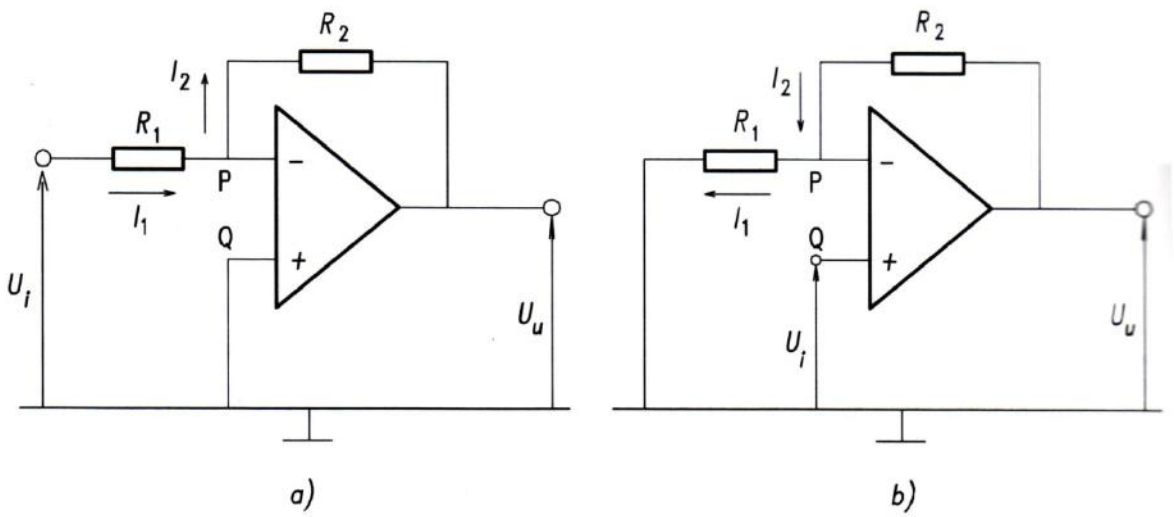
Per le ipotesi ammesse, UP = UQ = 0 (il punto P è detto massa virtuale poiché Q è collegato a massa). Dato poi che la corrente assorbita in ingresso dall' amp-op è nulla (causa il valore infinito della resistenza di ingresso), si deduce che la corrente entrante nel nodo P è uguale alla corrente uscente. Pertanto, dall'uguaglianza:

$(U_i – U_P) / R_1 = (U_P – U_u) / R_2$

si trae:

$U_u = - (R_2 / R_1) U_i$

L' amp-op retroazionato manifesta dunque in questo caso un guadagno finito che dipende solo dal valore delle resistenze esterne, e può essere fatto variare senza riprogettare l'amplificatore.
Il segno negativo dimostra come la funzione di amplificazione si accompagni alla funzione di inversione del segno della tensione di ingresso.
In particolare, se $R_1 = R_2$, 1'amp-op funziona come semplice invertitore.
Inoltre, se R_2/R_1 è inferiore all'unità 1'amp-op attenua il segnale di ingresso, anziché amplificarlo.



19.3 Amp-op non invertente

L' amp-op funziona come amplificatore non invertente nella configurazione della Fig. 19.3b. Ripetendo considerazioni analoghe a quelle svolte in precedenza, risulta ($U_P = U_Q = U_i$):

$(U_i – 0) / R_1 = (U_u – U_i) / R_2$

dalla quale si trae:

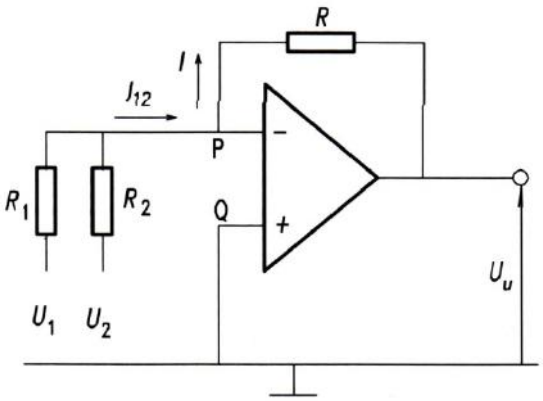
$U_u = U_i (1 + R_2 / R_1)$

I1 guadagno ad anello chiuso dell' amp-op è dunque sempre maggiore di uno qualunque valore abbiano le resistenze.

19.4 Amp-op sommatore invertente

Considerando la configurazione illustrata in Fig. 19.4 con le solite ipotesi si ha:

$$[(U_1 - U_P) / R_1] + [(U_2 - U_P) / R_2] = (U_P - U_u) / R$$



dalla quale, essendo $U_P = 0$, si trae:

$$U_u = - (R / R_1 * U_1 + R / R_2 * U_2)$$

In particolare, se $R_1 = R_2 = R$, risulta:

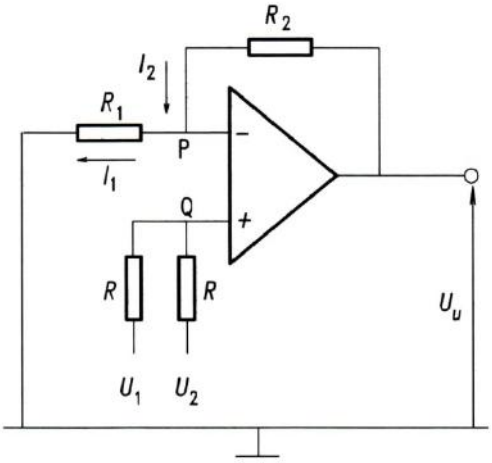
$$U_u = - (U_1 + U_2)$$

Si vede dunque che in questa configurazione 1' amp-op effettua la somma delle tensioni in ingresso e poi ne inverte il segno.

19.5 Amp-op sommatore non invertente

La configurazione della Fig. 19.5 1' amp-op funziona invece come sommatore non invertente. Si può infatti dimostrare che se $R_1 = R_2 = R$, si ha:

$$U_u = (U_1 + U_2)$$



19.6 Amp-op amplificatore differenziale (sottrattore)

La configurazione illustrata in Fig. 19.6 effettua l'amplificazione della differenza tra le tensioni poste in ingresso.

La tensione tra il punto Q e la massa si può determinare applicando il principio del partitore di tensione:

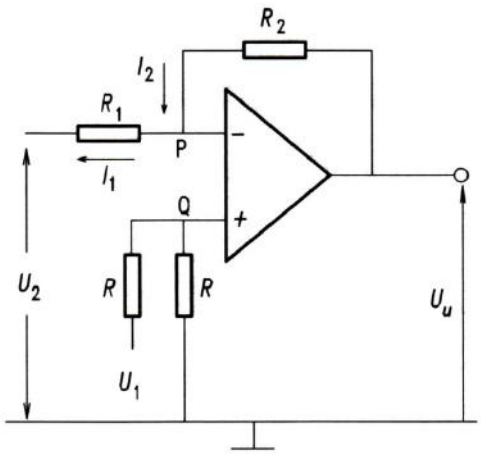
$$U_Q = R / (R + R) * U_1 = U_1 / 2 = U_P$$

Procedendo ora al solito modo:

$$(U_1/2 – U_2) / R_1 = (U_u – U_1/2) / R_2$$

dalla quale, con facile passaggio:

$$U_u = (1 + R_2 / R_2) * U_1/2 – R_2 / R_1 * U_2$$



Pertanto, se $R_1 = R_2 = R$:

$$U_u = U_1 – U_2$$

19.7 Amp-op amplificatore integratore

Dalla definizione di capacità C di un condensatore, la quantità di carica Q accumulata sulle armature è proporzionale alla tensione U applicata:

$$Q = C * U$$

La variazione elementare di carica conseguente alla variazione infinitesima di tensione è pertanto:

$$dQ = C * dU$$

Considerando le variazioni nell'unità di tempo:

$$dQ / dt = C * dU / dt$$

In questa relazione, il primo termine rappresenta la definizione dell'intensità della corrente elettrica, e pertanto:

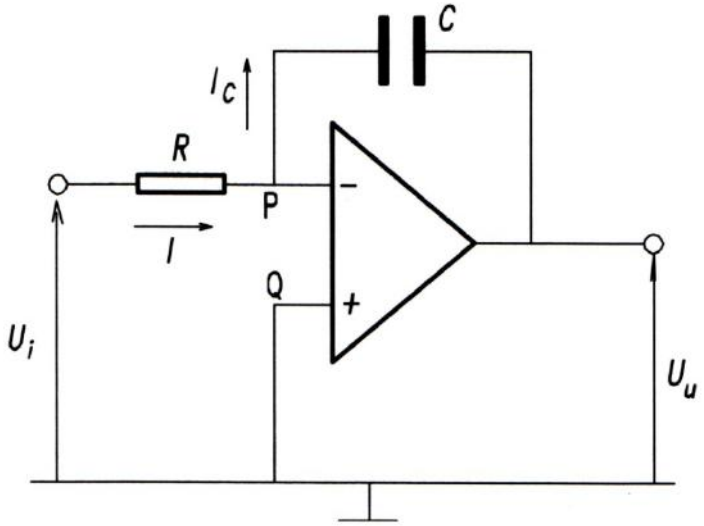
$$I(t) = C * dU / dt = I_C$$

Con riferimento al circuito della Fig. I9.7 con le solite ipotesi:

$$(U_i – U_P) / R = C * dU / dt = C * d(U_P – U_u) /dt$$

Ed essendo $U_P = 0$, risulta:

$$U_i / R = -C * dU_u /dt \quad \rightarrow \quad dU_u = - (1/ RC) * U_i * dt$$



Integrando entrambi i membri si ottiene infine:

$$U_u = - (1/RC) \int U_i (t) dt + K$$

Si vede dunque che la tensione di uscita è proporzionale all'integrale della tensione di ingresso. La costante K coincide con la tensione ai capi del condensatore all'istante t = 0. Se si suppone il condensatore inizialmente scarico, K = 0.

Nel caso si applichi in ingresso una tensione sinusoidale E * sen(ω * t), all'uscita si avrà:

$$Uu = (E / \omega * R * C) * \cos(\omega * t)$$

Considerando i valori efficaci, risulta pertanto un guadagno di tensione:

$$G = U_u / U_i = 1/ \omega * R * C$$

Grazie alla presenza del termine ω al denominatore, l'integratore risulta poco sensibile ai disturbi in alta frequenza dovuti al rumore.

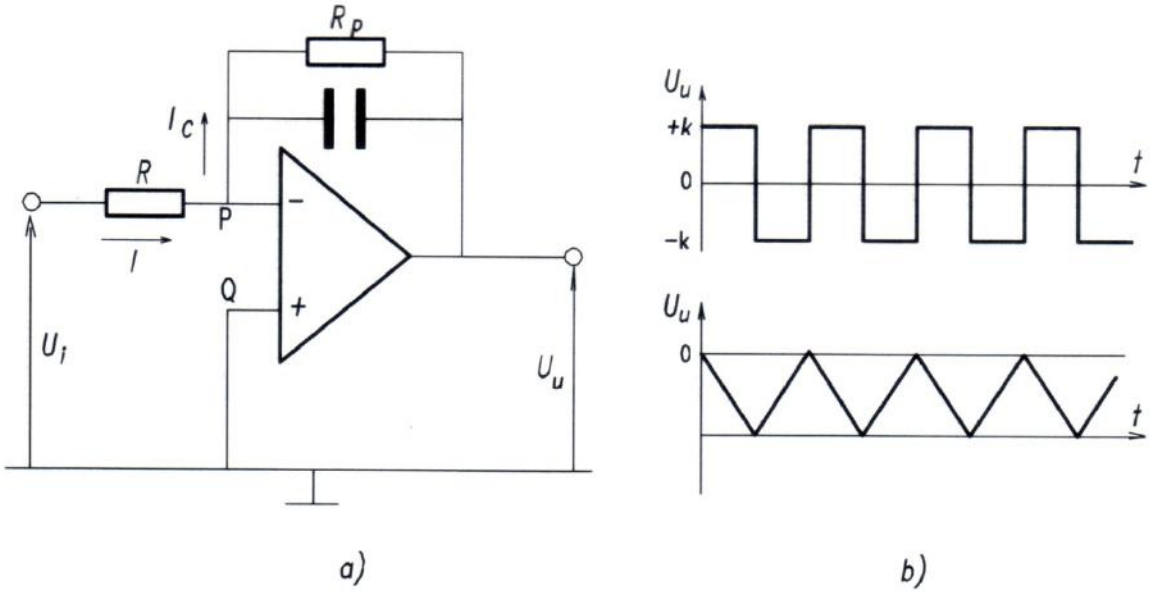
Il guadagno vale 1 alla frequenza 1/2πRC ed aumenta linearmente al diminuire della frequenza. Ciò comporta una amplificazione molto alta dei segnali a bassa frequenza ed una amplificazione teoricamente infinita per le componenti continue indesiderate del segnale di ingresso, ciò che porta 1' amp-op in saturazione (cioè fuori del comportamento lineare).
Per ridurre il guadagno alle basse frequenze si impiega il circuito della Fig. I9.8a, con una resistenza R_P di alto valore in parallelo al condensatore: al di sotto della frequenza 1/2π RC la reattanza capacitiva diventa trascurabile nei confronti della resistenza, ed il circuito si comporta come un amplificatore invertente con guadagno $-R_P$ / R. Affinchè il circuito mantenga buone caratteristiche di linearità è sufficiente che sia:

$$R_P > 10 * R$$

La Fig. 19.8b illustra la risposta del circuito integratore a segnali d'ingresso ad onda quadra. L'andamento dell'uscita è facilmente deducibile se si considera che l'ingresso assume valore + k in un semiperiodo e valore -k nel semiperiodo successivo. Nella prima semionda:

$$U_u (t) = - (1 / R * C) k * t$$

e si ha il primo tratto discendente. Nella semionda successiva, la costante di integrazione assume il valore di tensione raggiunto dalla precedente. E così via.



19.8 Amp-op amplificatore derivatore

Con riferimento alla Fig. 19.9, utilizzando un risultato conseguito nel paragrafo precedente:

$$I_C = C [d(U_i – U_P) / dt] = C (dU_i / dt) \quad ; \quad I = -(U_u / R)$$

E poiché le due correnti sono uguali, risulta:

$$U_u = -R * C (dU_i / dt)$$

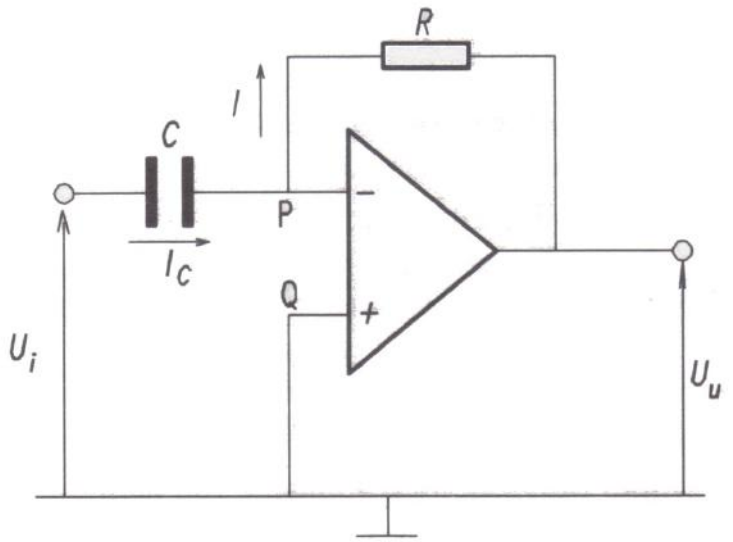
dalla quale si vede come la tensione di uscita sia proporzionale alla derivata della tensione di ingresso.

Come è facile aspettarsi, questo circuito manifesta un comportamento del tutto opposto al circuito integratore. Nel caso si applichi in ingresso una tensione sinusoidale E * sen(ω * t), all'uscita si avrà:

$$Uu = - \omega RC * \cos(\omega * t)$$

Considerando i valori efficaci, risulta pertanto un guadagno di tensione:

$$G = U_u / U_i = - \omega RC$$



Il guadagno vale uno alla frequenza 1/2π RC ed aumenta linearmente al crescere della frequenza. Ciò comporta una amplificazione molto alta dei segnali ad alta frequenza, spesso costituiti da rumori indesiderati.
Per limitare il guadagno alle alte frequenze si impiega il circuito della Fig. 19.10a: sopra la frequenza di taglio 1/2π RC nel ramo d'ingresso la capacità diventa trascurabile rispetto alla resistenza ed il circuito si comporta come un amplificatore invertente con guadagno $-R / R_S$.
Nel progetto si fa in modo che la costante di tempo del derivatore (R * C) sia uguale al periodo del segnale da derivare. La resistenza R_S si sceglie in modo da avere una frequenza di taglio di valore almeno 10 volte superiore a quella del segnale da derivare, e la resistenza R_O uguale al parallelo tra R_S e R.
La Fig. 19.1 Ob illustra la risposta del circuito derivatore all'onda triangolare applicata in ingresso: a riprova del comportamento duale rispetto al circuito integratore, la tensione di uscita assume l'aspetto di un'onda quadra.

19.9 Amp-op inseguitore di tensione

Consideriamo ora un primo esempio di circuito particolare realizzato con amp-op. Esaminando il guadagno del circuito amplificatore non invertente si nota che esso diventa unitario quando il rapporto R_2 / R_1 diventa nullo. Ciò è realizzabile col circuito della Fig. 19.11 b, nel quale $R_2 = 0$ e R 1= ∞.

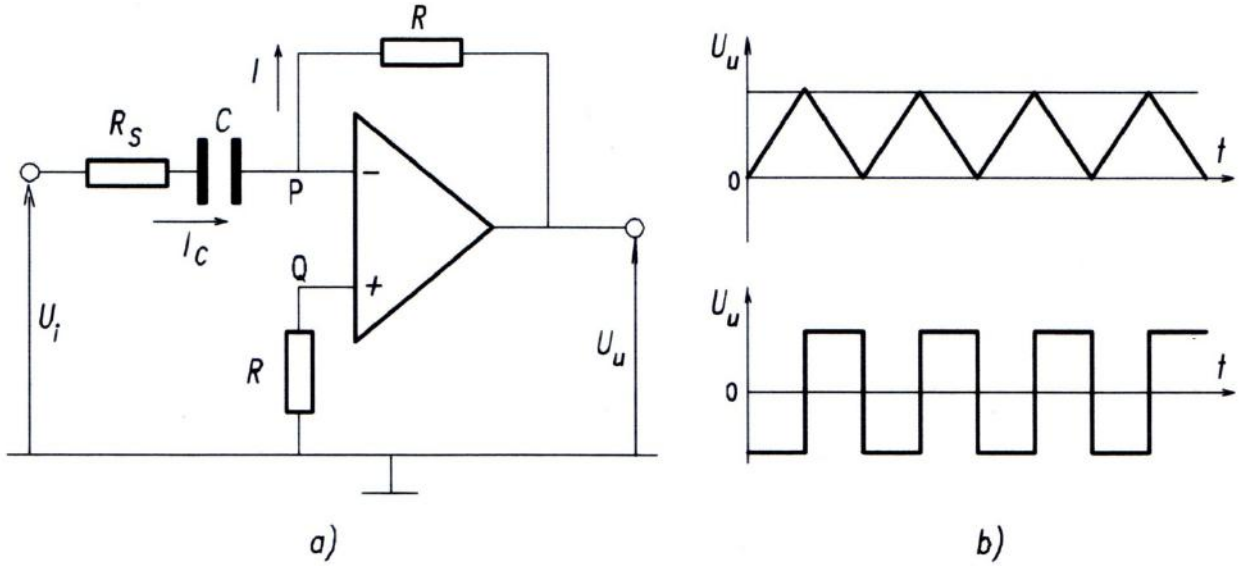
La tensione di ingresso è riportata in uscita ($U_u = U_i$) e, grazie alle caratteristiche dell' amp-op, ciò avviene senza prelevare corrente dal circuito a monte e senza caricare il circuito a valle.

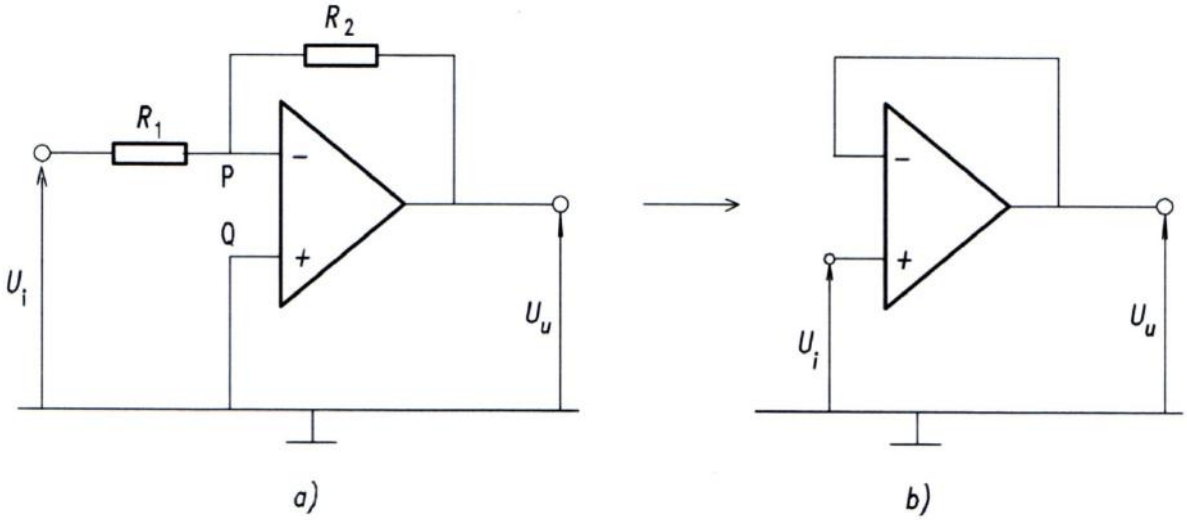
L' amp-op come inseguitore di tensione è pertanto utilizzato come stadio separatore, o buffer, tra due circuiti in cascata.

Un'interessante applicazione è costituita dal collegamento tra un potenziometro ed un carico.
Se nessun carico è collegato all'uscita del potenziometro sussiste una relazione lineare tra tensione di uscita e posizione x del cursore mobile (Fig. 19.12a).
Se invece si collega un carico, la relazione si modifica nel modo seguente (Fig. 19.12b):

da: $U_u = x * U$ a: $U_u = U [x / (xR / R_C + 1)]$

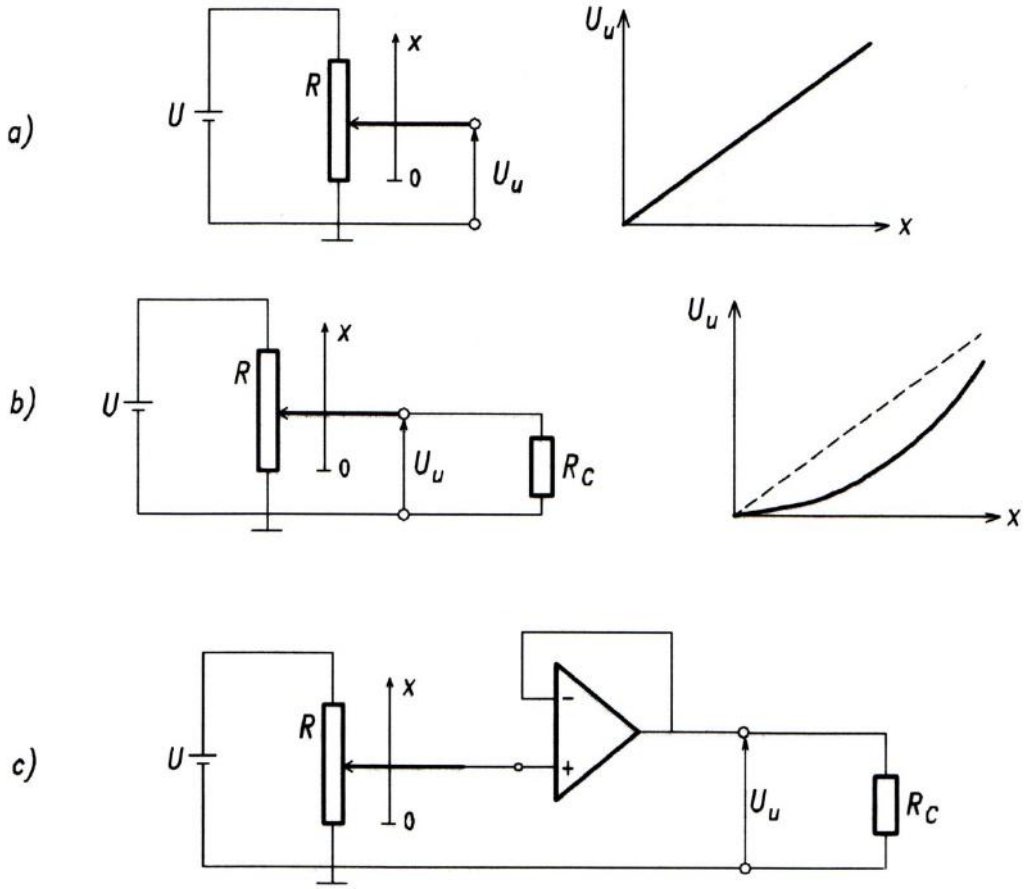
La seconda relazione è riducibile alla prima solo se $R_C >> R$: proprio a questo provvede l'inseguitore, mostrando al potenziometro la sua altissima resistenza di ingresso (Fig. 19.12c).


R
I
R_S
C
P
I_C
Q
U_i
R
U_u
a)
U_u
0
t
U_u
0
t
b)


R_2
R_1
P
Q
U_i
U_u
a)
U_i
U_u
b)

19.10 Amp-op per ponti di misura

Come secondo circuito particolare consideriamo l'amplificatore per ponti di misura. Come noto, piccole variazioni di una resistenza in un ponte di Wheatstone provocano una tensione di sbilanciamento del ponte. Molto spesso, tale tensione è molto piccola, per cui si rende necessaria la sua amplificazione.



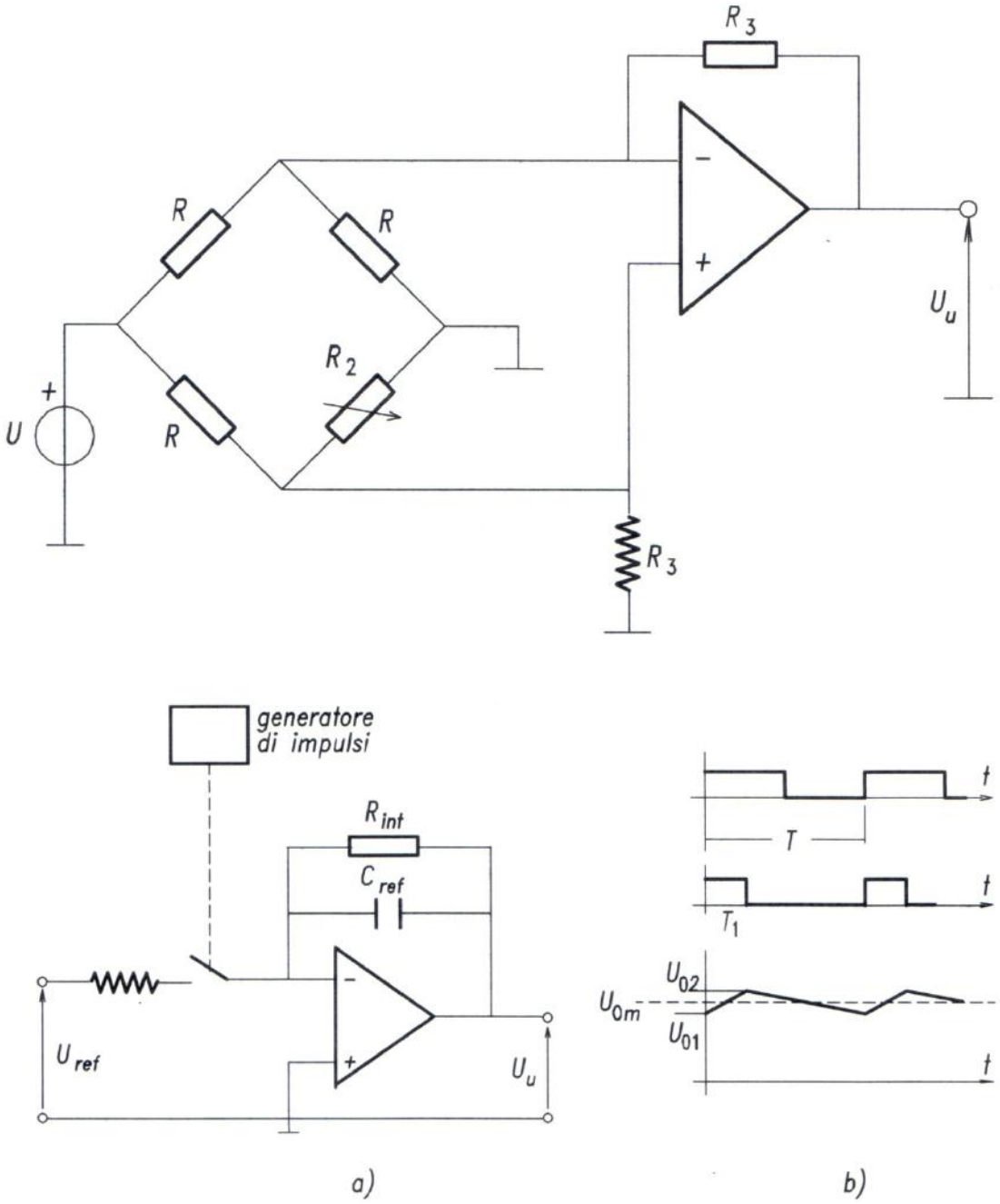
Nel caso del circuito illustrato in Fig. 19.13 si può dimostrare che:

$$U_u = U\{[R_3 / R]\ [(R_3 * R_2 – R * R_3) / (R_3 * R_2 + R * R_2 + R * R_2)]\}$$

Dalla quale si trae, correttamente, che per $R_2 = R$ si ha $U_u = 0$, cioè l'equilibrio del ponte.

19.11 Convertitore frequenza-tensione

Come terza applicazione particolare consideriamo il convertitore frequenza-tensione, il quale produce una tensione continua proporzionale alla frequenza degli impulsi di comando. La Fig. 19.14a ne illustra lo schema essenziale.
Con riferimento al funzionamento a regime, ad ogni passaggio per lo zero, la chiusura dell'interruttore (in pratica un transistor) fa caricare il condensatore per l'intervallo T_1 (Fig. 19.14b), di modo che la tensione ai suoi capi sale dal valore U_{01} al valore U_{02}.
Durante il successivo intervallo $T - T_1$ la capacità C_{ref} si scarica attraverso la resistenza R_{int}. Se la carica persa è pari a quella acquistata durante il periodo T_1 di carica del condensatore, si raggiunge un equilibrio e quindi una tensione in uscita di valore medio U_{0m} non nullo.



La carica immagazzinata vale:

$$q_i = C_{ref} * U_{ref}$$

La carica persa vale:

$$q_P = (U_{0m} / R_{int}) * T$$

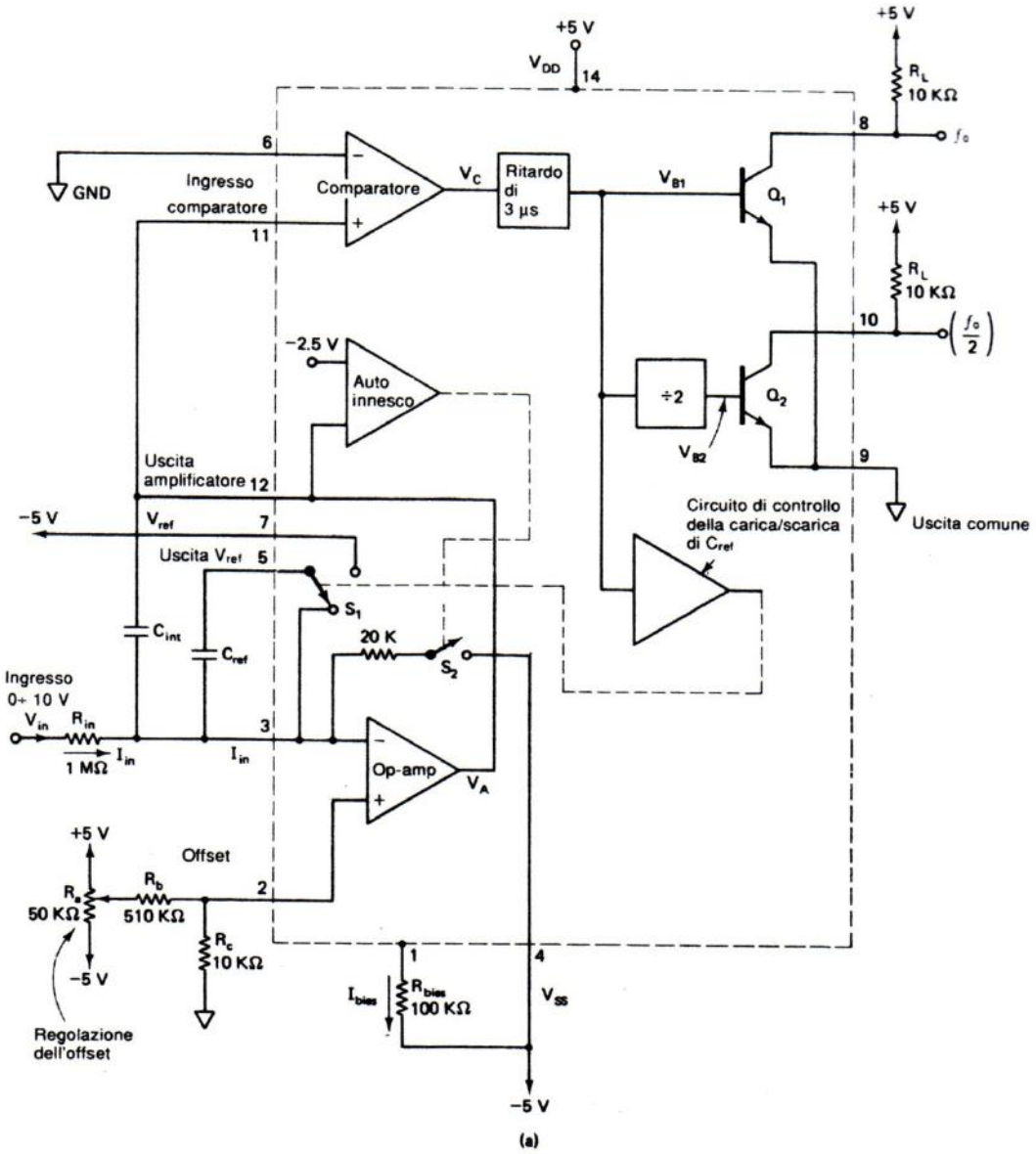
Uguagliando e considerando che 1/T = f (frequenza):

$$U_{0m} = U_{ref} * C_{ref} * R_{int} * f$$

Cioè:

$$U_{0m} = K * f$$

Per i motivi detti, questo convertitore è detto a bilanciamento di carica.
Un convertitore di frequenza sotto forma integrata è il CI 9400 della Teledyne Semiconductor, il quale può essere utilizzato (come d'altra parte tutti i convertitori f - V) anche per la conversione tensione-frequenza.
La Fig. 19.15 illustra lo schema interno del citato CI 9400.



19.12 Altre applicazioni lineari dell'amp-op

L'amplificatore operazionale è veramente un elemento universale, il quale può essere utilizzato in numerosissime circostanze applicative. Qui se ne citano solo alcune, rimandando ai testi specializzati ogni ulteriore approfondimento.

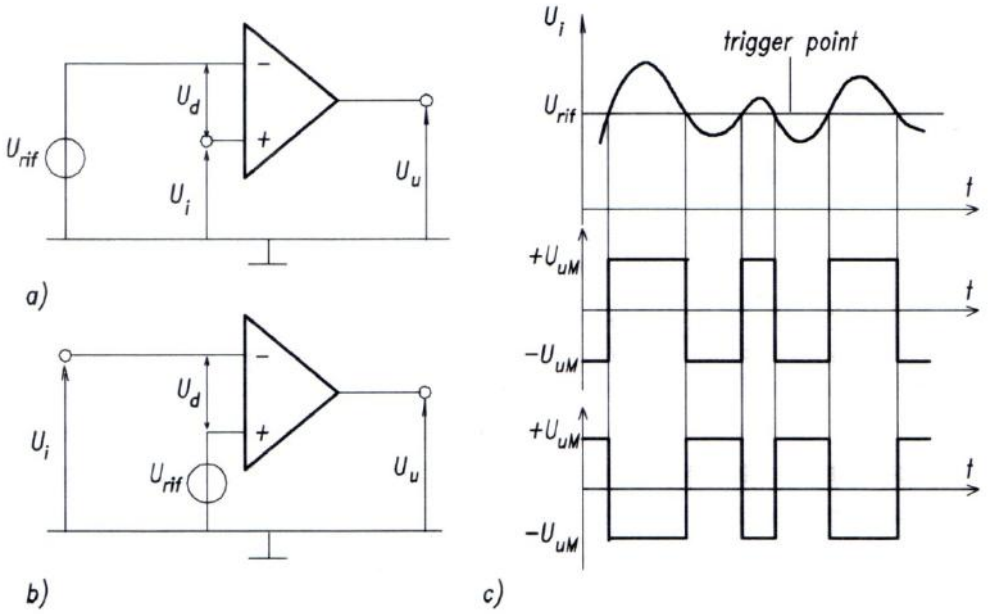
Con gli amp-op è innanzi tutto possibile realizzare filtri attivi aventi un comportamento del tutto simile a quello dei filtri passivi: filtri passa basso, passa alto, passa banda, elimina banda, risonanti ecc.
Un amplificatore logaritmico fornisce in uscita una tensione che è funzione logaritmica di quella in ingresso. Un amplificatore antilogaritmico dà in uscita la funzione inversa, cioè la funzione esponenziale della tensione di ingresso.
Un oscillatore è un circuito che produce in uscita un segnale periodico. Una categoria molto importante di tali dispositivi è quella degli oscillatori sinusoidali. L'oscillatore in quadratura ha due uscite, una variabile secondo la funzione seno e l'altra variabile secondo la funzione coseno.
Gli alimentatori stabilizzati forniscono una tensione continua indipendente (entro certi limiti) dal carico. Lo stadio stabilizzatore può essere realizzato anche con operazionali.

19.13 Applicazioni non lineari dell'amp-op

Vi sono moltissime applicazioni che sfruttano il comportamento non lineare dell' ampop non retroazionato, e cioè la commutazione a scatto quando si presenta all' ingresso una tensione differenziale diversa da zero.
Di seguito vedremo le applicazioni più interessanti nei riguardi dell'automazione industriale.

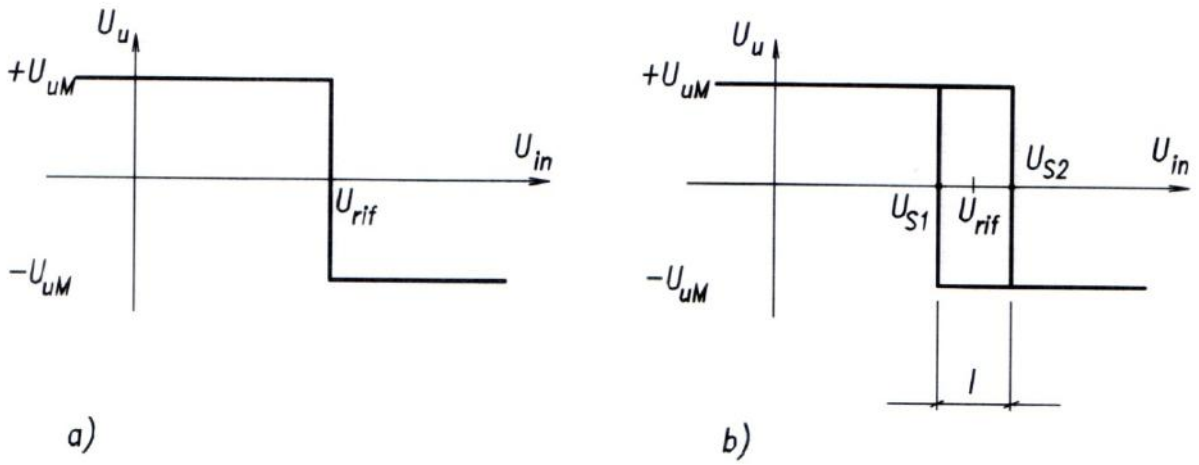
Comparatore (trigger)

I1 comparatore confronta la tensione di ingresso, applicata ad un ingresso dell' ampop, con una tensione di riferimento, applicata all'altro ingresso.
A seconda che la tensione di ingresso sia applicata al terminale non invertente o non invertente, si hanno le configurazioni illustrate nelle Figg.19.16a e 19.16b.
Data la ristrettissima zona lineare, l' amp-op commuta la propria uscita a scatto, ogni qual volta la tensione di ingresso si discosta di pochissimo dalla tensione di riferimento. La Fig. 19.16c illustra le risposte del comparatore non invertente ed invertente.



Comparatore con isteresi (trigger di Schmitt)

La configurazione precedente può dar luogo a commutazioni indesiderate, per lievi oscillazioni del segnale attorno al trigger point, dovute a disturbi.
Si utilizza allora un comparatore con isteresi, detto trigger di Schmitt, che presenta la curva caratteristica illustrata in Fig. 19.17b.



Si introducono due distinti livelli di riferimento U_{S1} e U_{S2}, così che, in configurazione invertente:

$$U_u = U_{uM} \quad \text{quando} \quad U_i < U_{S1} \quad \text{e} \quad U_u = -U_{uM} \quad \text{quando} \quad U_i > U_{S2}$$

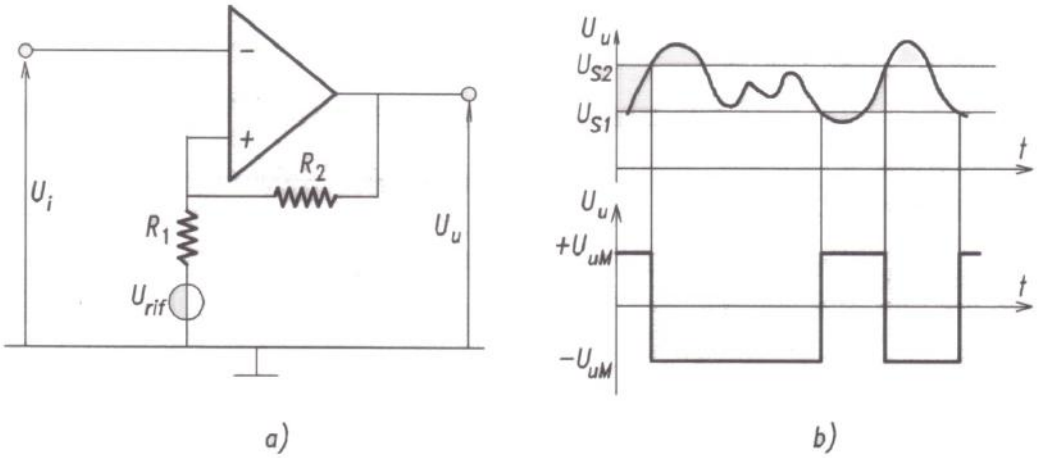
La differenza tra i due valori citati si chiama isteresi.
Un trigger di Schmitt puà essere realizzato con la configurazione illustrata in Fig. 19.18a, nella quale:

$$U_{S2} = U_{rif}\,[R_2 / (R_1 + R_2)] + U_{uM}\,[R_1 / (R_1 + R_2)] \quad ; \quad U_{S1} = U_{rif}\,[R_2 / (R_1 + R_2)] - U_{uM}\,[R_1 / (R_1 + R_2)]$$

cui corrisponde un'isteresi:

$$I = 2\,U_{uM}\,[R_1 / (R_1 + R_2)]$$

La Fig. 19.18b illustra la risposta del trigger di Schmitt ad un segnale con componenti spurie.

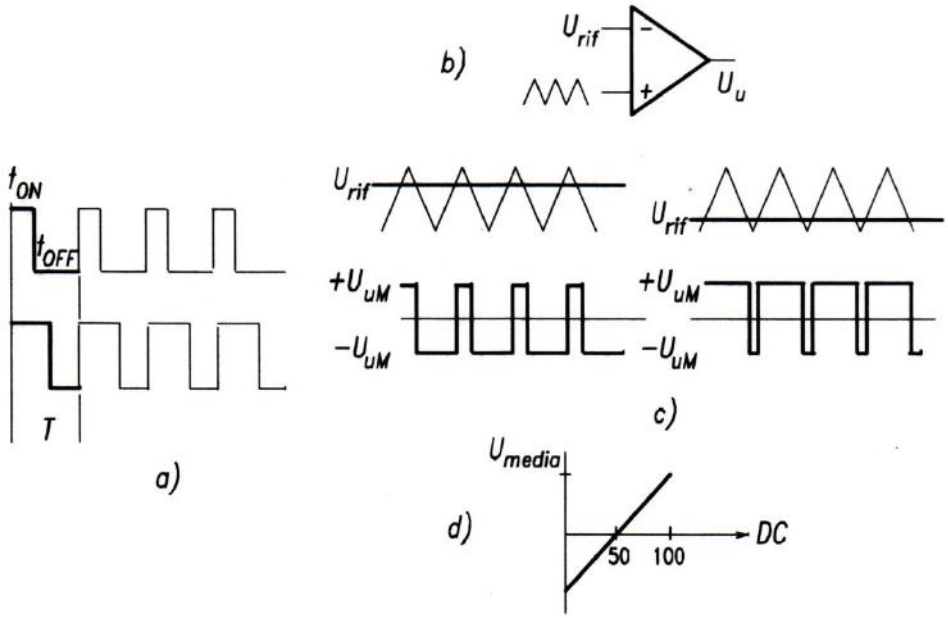


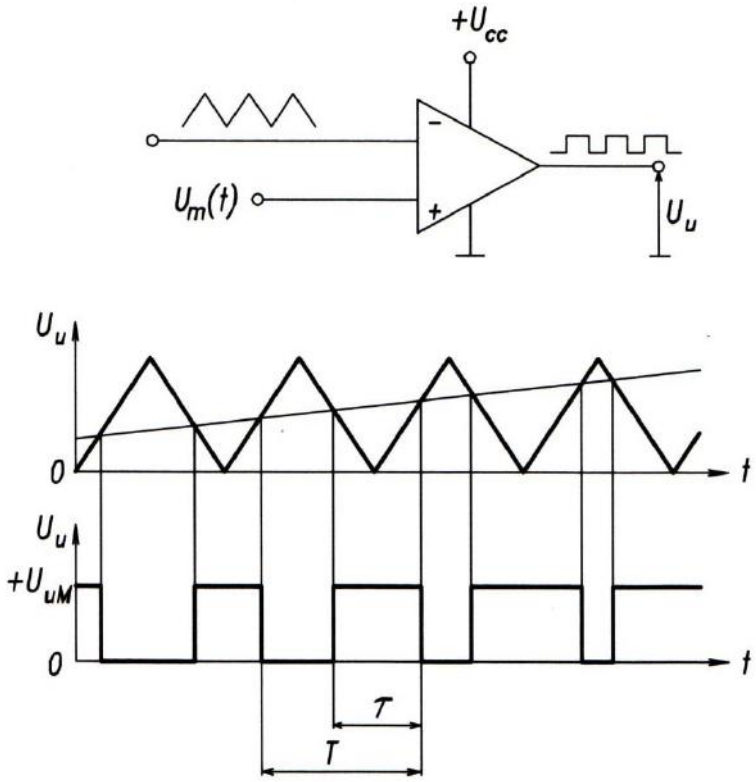
Modulatore PWM

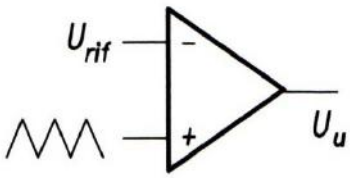
In molte circostanze risulta necessario disporre di un treno di impulsi a periodo costante ma del quale si deve far variare il cosiddetto duty cicle (o ciclo di lavoro, DC), cioè il rapporto tra il tempo di conduzione e quello di interdizione. In Fig. 19.19a sono illustrate due onde quadre con pari periodo ma diverso duty cicle.

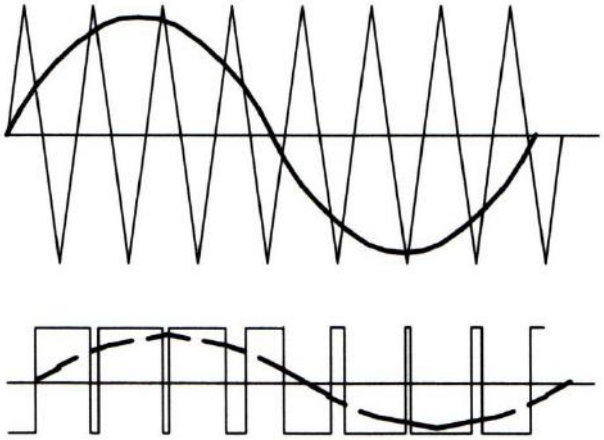
I1 dispositivo che genera tali onde è chiamato modulatore PWM, da Pulse Width Modulator ed è costituito essenzialmente da un comparatore che confronta un'onda a dente di sega con una tensione di riferimento costante ma di valore variabile (Fig. I9.I9b), detta modulante.
In Fig. 19.19c sono illustrate due onde quadre corrispondenti a due diversi valori della tensione di riferimento. Nei due casi si ottengono due tensioni di valor medio differente: in generale, il valor medio è legato al duty cicle (percentuale) dalla relazione illustrata in Fig. 19.19d.
In tal modo, una tensione continua a valore variabile è fatta dipendere dal valore di una modesta tensione di riferimento. Un'applicazione può essere il controllo della velocità di un motore c.c., la quale come sappiamo dipende dalla tensione continua applicata alle spazzole.

Se la tensione di riferimento è variabile anziché continua si ottiene un'onda a duty cicle variabile. In Fig. 19.20 è illustrato il caso di un riferimento crescente linearmente (a differenza che in tutti i casi trattati, qui 1' amp-op funziona con alimentazione singola).
Un caso molto importante nelle applicazioni è quello nel quale la modulante è un'onda sinusoidale: si ottiene un'onda quadra che approssima molto bene un'onda sinusoidale, come illustrato nella Fig. 19.21. L'applicazione più comune è il controllo della velocità dei motori asincroni trifase tramite variazione della frequenza. Come sappiamo, in questo caso, a rigore, si deve inviare un sistema trifase di tensioni sinusoidali.


b)
U_{rif}
U_u
t_{ON}
t_{OFF}
T
a)
U_{rif}
$+U_{uM}$
$-U_{uM}$
c)
U_{media}
d)
DC
50
100


$+U_{cc}$
$U_m(t)$
U_u
0
t
$+U_{uM}$
τ
T





Altre applicazioni non lineari dell'amp-op

Anche nel campo delle applicazioni non lineari, le applicazioni dell' amp-op sono molto numerose.
Oltre a quelle citate, se ne ricordano qui solo altre due. I circuiti limitatori limitano la tensione di uscita ad un valore determinato (comunque inferiore alla tensione di alimentazione). Il vantaggio dell'impiego dell' amp-op sta nell'indipendenza delle caratteristiche dal carico in uscita. Molto importanti nelle applicazioni sono poi i generatori di forme d'onda: onda quadra a duty cicle costante o variabile, onde triangolari.

Sistemi di comando digitali

20.1 Generalità

Fino a qualche decina d'anni fa, i sistemi di comando d'automatismi sequenziali erano realizzati con robusti pannelli a relè: gli ingressi erano connessi a sensori di vario tipo (finecorsa, switch, termostati, fotocellule ecc.) e le uscite azionavano attuatori come elettrovalvole, motori, resistenze di riscaldamento ecc.
Agli schemi di cablaggio si ricorreva in caso di diagnostica dei guasti oppure per effettuare modifiche al processo produttivo.

L'introduzione dei circuiti integrati a bassa e media scala di integrazione permise un vero e proprio salto qualitativo, sotto ogni aspetto saliente: velocità, compattezza, consumo, costo e manutenzione.
Tuttavia, anche le logiche cablate in tecnologia elettronica mantenevano la fondamentale caratteristica della rigidità: le funzioni svolte dal sistema risultavano fissate in modo definitivo all'atto del cablaggio dei componenti, e modifiche anche lievi richiedevano onerosi interventi di ricablaggio.

All'inizio degli anni '70 tale caratteristica di rigidità divenne sempre più insostenibile, man mano che i mercati proponevano una domanda mutevole e capricciosa, che decretava la fine di una produzione industriale improntata alla grande serie di prodotti tutti uguali o con lievissime differenze.
Così, il fondamentale requisito, richiesto ad un impianto di produzione, divenne sempre più la flessibilità, anche a costo di una perdita in produttività (che è massima nelle macchine automatiche monoscopo).

La soluzione fu infine trovata con i microprocessori ed i microcomputer.

Ad un microcomputer industriale fanno capo le linee di segnale provenienti dai sensori puntati sul sistema sede del processo controllato (eventualmente, anche quelle provenienti dall'operatore e dall'ambiente di lavoro). Da esso escono le linee di segnale dirette agli attuatori.
Le equazioni logiche che rappresentano il compito dell'automatismo sono articolate in un'opportuna successione di istruzioni logiche elementari: il programma.

Così, anziché costruire il circuito che realizza determinate funzioni (logiche) di controllo, si prende una logica elettronica programmabile (un microprocessore) e la si programma.
Il vantaggio è evidente: una circuiteria standard (il microprocessore), di basso costo e facilmente interfacciabile ai dispositivi sensoriali e di potenza, può essere impiegata per compiti diversissimi, semplicemente mutando il programma di elaborazione. .
E poi possibile operare piccole modifiche al compito assegnato, semplicemente modificando il programma.
Ma, soprattutto, con una logica programmabile si può utilizzare, per implementare le funzioni di controllo, il set di potenti istruzioni del microprocessore, il quale offre le funzioni corrispondenti a quelle svolte dai componenti tradizionali (ad esempio temporizzatori e contatori), e molte altre di trattamento numerico e logico.

A fronte del citato vantaggio della flessibilità, i microcomputer presentano tutti il difetto della esecuzione rigidamente sequenziale delle istruzioni, conseguenza delle caratteristiche software-hardware dell'architettura Von Neumann. È cioè esclusa qualunque forma di "parallelismo", e l'elaborazione delle informazioni ne risulta notevolmente rallentata. Di modo che, se i segnali variano molto rapidamente, possono anche perdersi delle informazioni.

Con la logica elettronica cablata si possono invece adottare tutti i circuiti che occorrono, e farli lavorare in parallelo: le uscite che vedono soddisfatte le condizioni di attivazione sono attivate contemporaneamente. I circuiti che in tal modo si ottengono presentano pertanto la massima velocità nel trattamento dei segnali. Di qui il nome di logica "sparsa": capacità di elaborazione sparsa nel circuito, e non "concentrata" in un unico processore di elaborazione.
Vedremo tuttavia (nel terzo volume) che la discussa caratteristica delle logiche programmabili raramente costituisce un serio limite.

20.2 Comportamento elettrico dei circuiti integrati

L'architettura hardware e software di un computer utilizzato come sistema di comando non differisce sostanzialmente da quella di un comune personal computer.

Le differenze stanno nella qualità-degli isolamenti e nella robustezza meccanica dei contenitori e degli alloggiamenti, ciò che rende questi computer adatti a operare in ambienti con temperature fino a 60 - 70 C°, con umidità relative fino al 95% ed anche alla presenza di forti campi magnetici.
Ma è la sezione di I/O la parte funzionale dove maggiori sono le differenze.

La sezione di I/O deve assicurare il dialogo con sensori e con attuatori, i quali funzionano con tensioni e correnti molto diverse, prevenendo errate interpretazioni dei segnali ed evitando che sovraccarichi nei circuiti a valle possano compromettere l'integrità della parte elettronica.

È dunque di fondamentale importanza conoscere il comportamento elettrico dei circuiti integrati: le tensioni di lavoro e le correnti massime che possono assorbire ed emettere.

Famiglia TTL

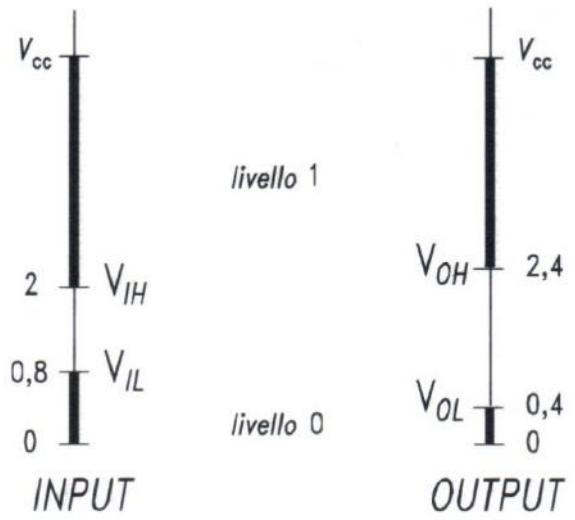
Come sappiamo, tutti i circuiti integrati digitali, salvo casi particolari, si classificano in famiglie logiche, secondo il tipo di transistor utilizzato, la tecnologia di realizzazione delle resistenze e del tipo di circuito impiegato per realizzare le porte logiche. Al giorno d'oggi, le famiglie logiche di gran lunga più diffuse sono la TTL e la CMOS, utilizzanti solo transistor e resistenze.

La TTL, da Transistor-Transistor-Logic, deriva il proprio nome dal fatto che gli stadi di ingresso e di uscita delle porte sono costituiti da transistor, i quali sono di tipo BJT.

La tensione continua di alimentazione di tutti gli integrati TTL è fissata nel valore di 5V ± 5%. La tensione massima che può essere applicata agli ingressi è di 5,5 V.
I livelli di tensione corrispondenti ai livelli logici 0 e I, per gli ingressi e le uscite, sono indicati in Fig.20.1.

La famiglia TTL è stata realizzata per la prima volta dalla Texas Instruments, con la sigla generale SN (Semiconductor Network) per tutti i chip, e due serie denominate rispettivamente 54 e 74; la prima per applicazioni militari (nel campo di temperatura da -55 C° a +125 C°), la seconda per le normali applicazioni commerciali (nel campo di temperatura da 0 C° a 70 C°).
Entrambe le serie sono realizzate in diverse versioni, di seguito indicate con riferimento alla serie 74:

- standard (74);
- a bassa potenza di dissipazione (74L);
- ad alta velocità di commutazione, o tipo Schottky (74S);
- a bassa potenza di dissipazione ed altissima velocità (74LS);
- advanced Schottky (AS);
- advanced low-power Schottky (ALS).



Nella versione L l'unica differenza rispetto alla versione standard consiste nell'impiego di resistenze più elevate, così che la potenza dissipata è circa 1/10 di quella degli integrati standard.
La serie S è la più veloce perché impiega opportunamente dei diodi Schottky, i quali impediscono ai transistor di entrare in saturazione, e pertanto ne aumentano la velocità di commutazione.
Le versioni standard ed LS sono le più facilmente reperibili sul mercato.
La Tabella 20.1 mette a confronto i valori massimi consigliati di correnti emesse e assorbite per le versioni standard ed LS della serie 74, con la convenzione di ritenere positiva una corrente entrante e negativa una corrente uscente dalla logica.

Tabella 20.1 Correnti emesse/assorbite da un CI in tecnologia TTL standard e LS

Correnti (mA)	Serie 74	Serie 74LS	Descrizione
I_{OH}	-0,4	-0,4	Massima corrente emessa da un'uscita a livello alto
I_{OL}	16	8	Massima corrente assorbita da un ingresso a livello basso
I_{IH}	40	20	Massima corrente assorbita da un ingresso a livello alto
I_{IL}	-1,6	-0,4	Massima corrente emessa da un ingresso a livello basso

Un ingresso non collegato equivale ad un ingresso a livello I. È buona norma non lasciare mai ingressi fluttuanti perché potrebbero raccogliere disturbi di varia natura, alterando il normale funzionamento della porta.

Oltre la Texas Instruments, diversi altri costruttori producono CI TTL nelle diverse serie ora viste. Anch'essi sono identificabili attraverso la sigla stampigliata sul package: National Semiconductor (DM = Digital Monolitic), Motorola (MC), Fairchild (FSC), RCA (CD), SGS (HCC, HCT).

Famiglia CMOS

Quando furono introdotti i CI TTL, svariati problemi impedivano l'utilizzo pratico dei transistor MOS. Quando infine tali problemi furono superati nacquero le famiglie logiche PMOS ed NMOS, per giungere alla realizzazione della prima serie CMOS (Complementary-MOS), la 4000A, realizzata dalla RCA nel 1968, e successivamente migliorata con la 4000B .
I CMOS sono più veloci e consumano meno potenza delle altre famiglie MOS, ma rispetto a queste presentano una maggiore complessità di fabbricazione e perciò minore densità di integrazione.
Il processo di fabbricazione è invece più semplice rispetto a quello dei TTL; inoltre, consumano solo una frazione della potenza richiesta da questi e sono quindi i più adatti nelle applicazioni in cui l'alimentazione è fornita da una batteria. I CMOS sono in genere più lenti dei TTL, sebbene le nuove serie CMOS ad alta velocità possano competere con le serie 74 e 74LS.
I CMOS si sono imposti soprattutto nell'area a media scala di integrazione (MSI); sono preferiti i MOS nelle applicazioni che richiedono le massime prestazioni in LSI.

Le serie 4000 operano con una qualunque tensione appartenente all'intervallo da 3 a 15 V.
Per gli ingressi , i valori limite sono fissati come percentuali della tensione di alimentazione, come risulta in tabella 20.2

Tabella 20.2 Tensioni corrispondenti ai livelli logici nella famiglia CMOS

Tensioni	Valori (V)	Descrizione
V_{OL}	0	Massima tensione per il livello basso in uscita
V_{OH}	V_D	Minima tensione per il livello alto in uscita
V_{IL}	30% V_D	Massima tensione per il livello basso in ingresso
V_{IH}	70% V_D	Minima tensione per il livello alto in ingresso

Gli ingressi CMOS presentano una resistenza elevatissima. Pertanto, in condizioni statiche, i CMOS non dissipano praticamente alcuna potenza. La potenza è dissipata nella fase di commutazione. La potenza media dissipata dipende pertanto dalla frequenza delle commutazioni

Molto spesso, i CMOS vengono impiegati assieme ai TTL in un medesimo circuito: in questo caso può essere utilizzata un'unica alimentazione a 5 V. Risultano i seguenti valori limite:

V_{IL} = 1,5V ; V_{IH} = 3,5V

Perciò, quando un CI in CMOS è alimentato con una tensione di 5 V riconoscerà come livello basso qualsiasi valore tensione di ingresso minore di 1,5 V, e come livello alto qualsiasi tensione di ingresso maggiore di 3,5 V.

Anche gli ingressi CMOS devono essere sempre collegati (non fluttuanti), a massa o all'alimentazione: ciò previene l'accumulo di cariche elettrostatiche e la conseguente interpretazione errata dei segnali logici e soprattutto previene la possibilità della conduzione accrescendo la potenza dissipata ed esponendo il CI al rischio di danneggiamento per sovratemperatura.

La serie 74C ha caratteristiche del tutto analoghe a quelle delle serie 4000, ed è del tutto compatibile, a livello di piedini e di funzioni, con i TTL aventi medesimo numero. La serie 74HC è la serie migliorata della precedente ed il miglioramento riguarda la velocità di commutazione, che è 10 volte maggiore, paragonabile a quella dei TTL 74LS. La serie 74HCT è un'altra serie ad alta velocità che ha la caratteristica di poter essere pilotata direttamente da un'uscita TTL.

20.3 Interfacciamento con sensori logici

I sensori e gli attuatori si classificano nelle due categorie dei dispositivi:

- logici
- analogici

Al funzionamento dei primi sono connessi segnali logici, mentre al funzionamento dei secondi sono connessi segnali continui o analogici.

Nei sensori logici si verifica la commutazione di un contatto in corrispondenza ad un determinato evento, che può essere il raggiungimento:

- di una determinata posizione (ad esempio di finecorsa);
- di un determinato livello liquido;
- di una determinata pressione o temperatura (pressostato, termostato);
- di un determinato livello di vuoto (vacuostato);
- e così via.

Il contatto è inserito in un circuito con una propria alimentazione ed un'opportuna resistenza, in modo che la circolazione di corrente in questo circuito possa essere trasformata in un segnale di tensione in ingresso a qualche circuito digitale (TTL o CMOS).

Le indicazioni date al paragrafo precedente permettono il progetto dei circuiti di interfacciamento diretto a circuiti integrati.

Tuttavia, di norma si opera affinché sussista una separazione galvanica tra circuito di ingresso e circuiti interni elettronici, onde risolvere il problema delle diverse tensioni e correnti ed inoltre per preservare la logica elettrica da sovratensioni che possono sempre generarsi nei circuiti sensoriali.

Nel caso di sensori funzionanti in corrente continua la soluzione è di utilizzare un optoisolatore, secondo quanto indicato nello schema di principio della Fig.20.2. La corrente circolante nel circuito del sensore eccita il LED il quale a sua volta satura il fototransistor: si chiude in tal modo un circuito interno.

Se il sensore funziona in corrente alternata, si renderà necessario uno stadio raddrizzatore della corrente.
In alternativa, si può utilizzare un minirelè: questa soluzione presenta tutti gli svantaggi dell'avere un dispositivo con parti mobili, ma in compenso è adatta per correnti sia continue sia alternate.

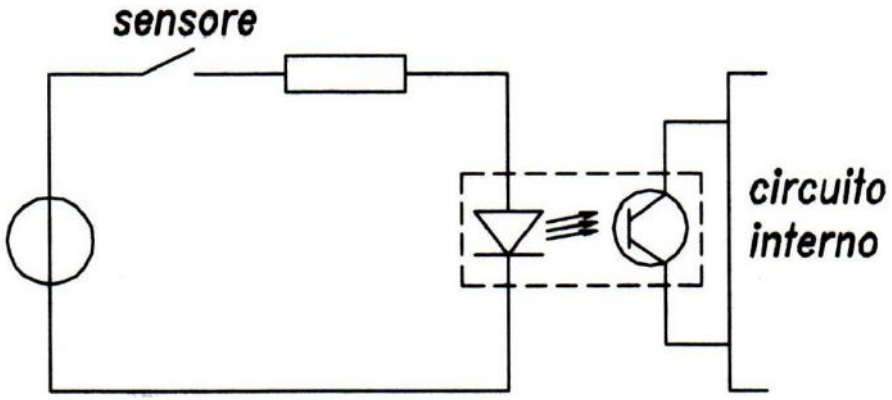
Sensori come fotoresistenze e fotodiodi danno luogo a segnali variabili nel tempo, e devono essere convertiti in formato digitale mediante un trigger di Schmitt.

20.4 Interfacciamento con attuatori logici

I dispositivi pilotati da una logica elettronica programmabile costituiscono "carichi" elettrici riconducibili alle seguenti categorie:

- carichi di tipo resistivo: lampade a incandescenza, elementi elettrici scaldanti, LED ecc.; vanno alimentati con adeguati valori di corrente e tensione, cioè con adeguata potenza;
- carichi di tipo induttivo: bobine di relè e di elettrovalvole, motori, altoparlanti ecc.; assorbono in genere una notevole corrente;
- carichi di tipo capacitivo: praticamente non assorbono corrente, ma spesso necessitano di particolari tensioni di alimentazione: tubi a raggi catodici, display a gas ecc.

Le indicazioni date al paragrafo precedente permettono il progetto dei circuiti di interfacciamento diretto a circuiti integrati, con riguardo sia alle tensioni sia alle correnti necessarie al carico in continua.
Anche qui, tuttavia, viene di norma perseguito l'obiettivo della separazione galvanica tra i circuiti interni alla logica ed i circuiti degli attuatori comandati.



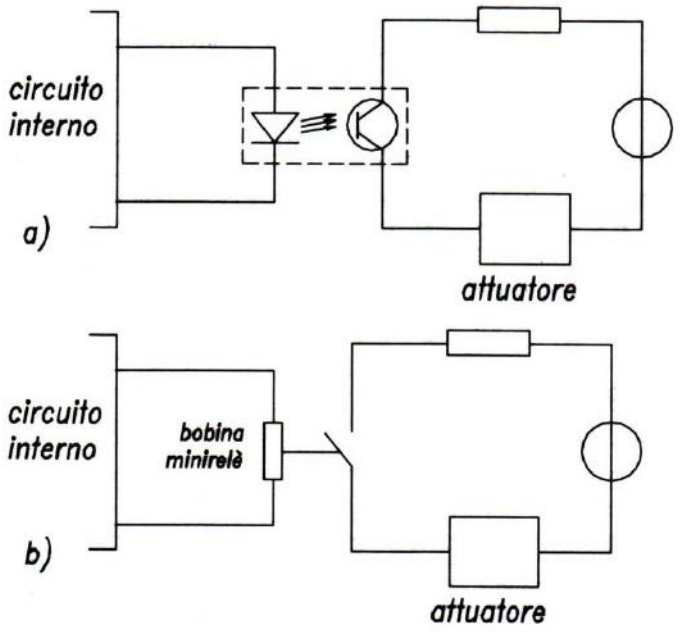
La soluzione è di chiudere un contatto entro un circuito di alimentazione del carico, fornito della opportuna sorgente d'energia elettrica. Il contatto può essere un fototransistor per modesti carichi in continua, oppure il contatto di un minirelè, per carichi in continua o alternata. Si vedano gli schemi di principio della Fig.20.3.

In alternativa al minirelè può essere utilizzato un TRIAC, il quale presenta il vantaggio di essere un dispositivo completamente statico.

20.5 Interfacciamento con sensori e attuatori analogici

Spesso, le informazioni che una logica digitale deve elaborare non sono di semplice presenza-assenza, così come gli ordini che deve emettere non sono di semplice attivazione-disattivazione.
Una branca molto importante dell'automazione industriale è quella del controllo di un processo continuo: il computer di governo ha il compito di mantenere determinate grandezze a valori prefissati, oppure imporre loro determinate variazioni nel tempo. I1 valore effettivo di tali grandezze viene quindi continuamente rilevato e confrontato con il valore desiderato. In caso di scostamento, il computer genera un'azione correttrice proporzionale allo scostamento (o errore), attraverso un certo attuatore.

Vi è dunque la necessità in primo luogo di trasformare l'informazione sensoriale (una temperatura, una pressione, uno spessore ecc.) dandole una forma elettrica.
I dispositivi che rilevano una certa grandezza e generano un segnale elettrico proporzionale sono chiamati trasduttori.



Se il trasduttore non genera di per sé un segnale sufficientemente grande per essere trattato dalla circuiteria di conversione, il livello di tensione è amplificato. Un filtro attivo elimina poi le componenti di segnale di bassa e alta frequenza; ulteriori circuiti di condizionamento "formatteranno" infine il segnale.
Questo segnale deve essere poi convertito in forma digitale, come gruppo di impulsi elettrici. A ciò provvedono i convertitori Analogico/Digitale (A/D Converter, o ADC = Analog/Digital Converter), utilizzando un apposito codice binario.
L'evoluzione della grandezza appare dunque al computer come una serie di numeri binari.

Ovviamente, la conversione non si rende necessaria se i trasduttori generano direttamente un'uscita digitale: per questo motivo la tecnologia elettronica ha studiato e realizzato trasduttori digitali.

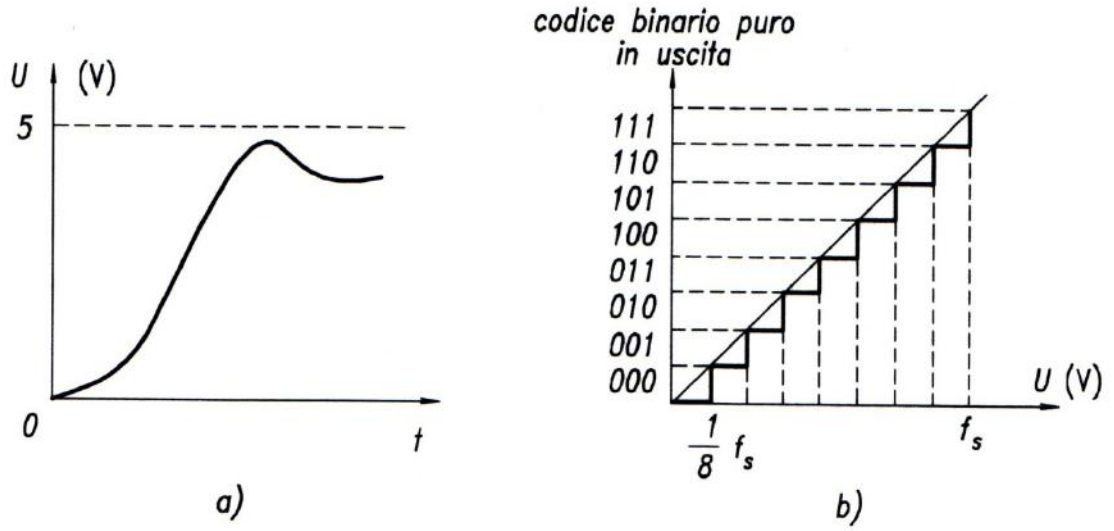
Reciprocamente, dal lato uscita, il computer emette gruppi di segnali elettrici che rappresentano (secondo un certo codice binario) il valore dell'azione di comando. E vi è pertanto la necessità di ricostruire una tensione elettrica di valore corrispondente, che possa pilotare 1' attuatore.
A questa funzione provvedono i convertitori Digitale/Analogico (D/A converter, o DAC = Digital-Analog Converter). Anche in questo caso, la presenza del D/A non necessaria se gli attuatori analogici accettano segnali di comando di tipo digitale: tipico il caso del motore passo - passo (step - motor).

Principi della conversione analogico-digitale

Come si è detto, un trasduttore genera un segnale di tensione che varia nell'identico modo della grandezza posta al suo ingresso (segnale perciò detto analogico). L'equazione di funzionamento (se nota) o la curva caratteristica del trasduttore, rilevata sperimentalmente, permettono di dedurre i valori della grandezza dai valori di tensione in uscita.
I diversi convertitori A/D accettano segnali di tensione compresi entro un determinato campo di variazione, il quale può essere simmetrico (ad esempio: 5V,10V) o asimmetrico (0÷5V, 0÷10V). Per convertire segnali esterni al campo specificato, si può utilizzare uno stadio di adattamento iniziale.
La Fig.20.4a illustra la generica variazione di un segnale di tensione entro il campo 0÷5V.

Ora, se il convertitore genera in uscita un'informazione a n bit, i diversi valori binari che può emettere sono 2^n. Indicando con f_s l'ampiezza del campo di variazione, questa è pertanto divisa in 2^n intervalli cui è associata una delle possibili parole binarie. Così, l'ADC genera in ogni istante la parola binaria corrispondente all'intervallo di tensione nel quale cade il segnale di ingresso.

Il codice utilizzato per la conversione è illustrato dalla curva caratteristica dell'ADC.



La Fig.20.4b illustra il caso di un convertitore a 3 bit che utilizza il codice binario puro, molto frequente nelle applicazioni.

Dunque, il campo di variazione è suddiviso in intervalli chiamati quanti di tensione, è cioè quantizzato. E l'informazione che giunge al computer non è continua ma, appunto, quantizzata, perché fintanto che la variazione all'interno di un certo quanto di tensione non è avvertita, non produce alcuna variazione di codice.

Riconsiderando l'esempio della Fig.20.4b, se in uscita si ha un gruppo di segnali corrispondente al codice binario puro 101, significa che il valore di tensione in ingresso è compreso nel quinto gradino. Il valore di ingresso dunque risulta:

$$U_i = 5/8\ f_s$$

In generale, se x è il numero decimale corrispondente al codice binario e f_s la tensione di fondo scala, si ha:

$$Ui = x/2^n * f_s$$

Il numero di bit del codice determina quindi la risoluzione del convertitore: la minima variazione della grandezza che dà luogo ad una variazione del codice. Con 8 bit il campo di funzionamento risulterebbe diviso in 256 parti e la risoluzione sarebbe 1/256 del fondo scala.
La metà della risoluzione indica l'errore dello strumento. Ad esempio, se la risoluzione è 0,1V,1' errore è di 0,05V; infatti, questa è la quantità che può far "scattare" i valori di 0,1 in 0,1.

Ora, in generale, una grandezza può assumere una qualunque variazione nel tempo ed è necessario "ricostruirla" fedelmente come serie di valori binari codificati.
Si deve in altre parole decidere con quale frequenza "campionare" la grandezza in ingresso al convertitore.
In base al Teorema di Fourier, un generico segnale v(t) può sempre essere ricondotto alla somma di un numero infinito di sinusoidi, dette armoniche, ciascuna con pulsazione pari ad un multiplo intero di una pulsazione detta fondamentale (relativa alla prima armonica). La rappresentazione del segnale tramite le sue armoniche prende il nome di spettro del segnale. Particolare importanza riveste la massima escursione in frequenza di detto spettro, che è definita banda del segnale.
In pratica, per caratterizzare il segnale sono sufficienti le prime tre o quattro armoniche, dette armoniche fondamentali, e la banda del segnale riferita ad esse.

Il Teorema di Shannon afferma che se la frequenza f con cui sono prelevati i campioni è maggiore od almeno uguale al doppio della banda B del segnale di partenza è possibile ricostruire (tramite appositi filtri) detto segnale dalla sua rappresentazione campionata.

In modo equivalente, l'intervallo tra due successivi campionamenti deve essere:

$$T_{camp} < 1/2B$$

Deve essere pertanto scelto un convertitore con un tempo di conversione T_{conv} inferiore all'intervallo di campionamento:

$$T_{conv} < T_{camp}$$

Il tempo di conversione dipende soprattutto dal metodo utilizzato per la conversione.

Poiché infine il segnale di ingresso al convertitore deve rimanere costante per tutto i tempo della conversione, in alcuni tipi di convertitore (vedi più avanti) è necessario un circuito di "tenuta" (hold), o meglio di memoria analogica (costituito da un condensatore) del campione (sample).
In realtà, anche l'operazione di campionamento (sampling) richiede un tempo non nullo, del quale bisogna tenere conto nella scelta del convertitore.

Principi della conversione digitale – analogico

Un convertitore digitale/analogico è un dispositivo che riceve al suo ingresso un insieme di segnali elettrici e fornisce in uscita il segnale analogico (tensione o corrente) che gli corrisponde secondo il codice binario adottato, e secondo un segnale interno di riferimento U_{rif} . Valgono così tutte le considerazioni già svolte, solo che ora grandezze di ingresso e di uscita invertono i ruoli, ed al posto della tensione f_s è necessario considerare la tensione U_{rif}.

Nel caso il codice sia il binario puro, il convertitore genera una tensione che è una frazione di U,
f. pari alla frazione tra il numero decimale in ingresso ed il numero
massimo:

$$U_u = (x / 2^n) U_{rif}$$

Il valore decimale di ingresso è calcolato per il codice a tre bit (b_2, b_l, b_0) come segue:

$$x=b_2 * 2^2 +b_l * 2^1 +b_0 * 2^0$$

Sostituendo, per n = 3:

$$U_u = (U_{rif} / 8) * (4*b_2 + 2 * b_l + 1 * b_0)$$

si vede che ciascun bit dà un contributo all'uscita proporzionale al suo "peso".
Il bit meno significativo, avente "peso" minore, è indicato con la sigla LSB (Less Significant Bit); mentre quello più significativo, avente "peso" maggiore, è indicato con la sigla MSB (Most Significant Bit). Nel caso a tre bit, rispettivamente boe b2. I1 circuito che genera il riferimento di tensione U,f. può essere sia interno che esterno al convertitore, in funzione dei requisiti di stabilità e risoluzione.
In generale, utilizzando la tecnologia CMOS, poiché è possibile sfruttare i transistori bipolari parassiti messi a disposizione dalla tecnologia, si possono realizzare riferimenti di tensione ragionevolmente accurati.

Tecniche della conversione A/D e D/A

Vi sono diverse tecniche di conversione A/D, le quali vanno valutate in relazione ai due parametri fondamentali che sono la risoluzione espressa in numero di bit del codice e la velocità di campionamento espressa in ksample/s o in Msample/s.
Tali parametri tendono spesso ad essere in conflitto: convertitori ad alta velocità presentano un risoluzione bassa, e viceversa.

Il convertitore flash (o convertitore parallelo) è costituito da un partitore di tensione alimentato con la tensione di fondo scala; ad ogni uscita intermedia del partitore vi è un valore di tensione pari ad un

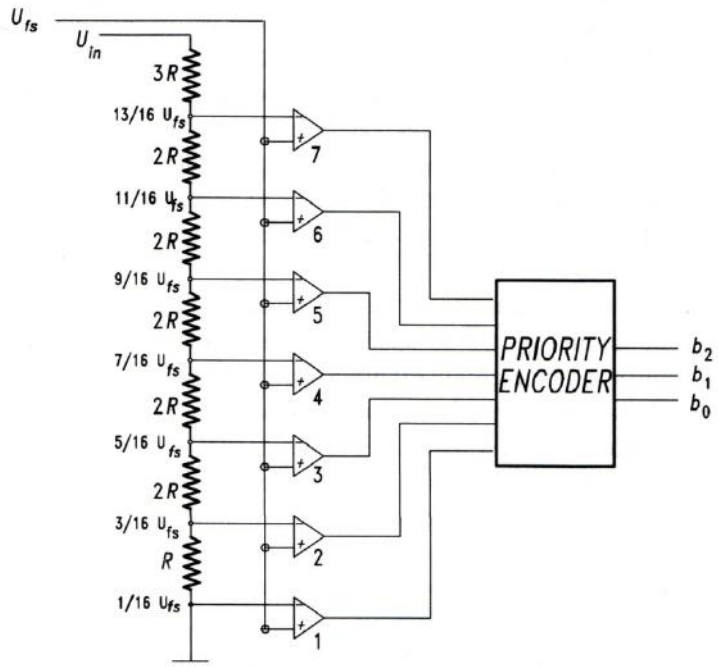
gradino di tensione codificabile col codice a n bit. I1 caso della Fig.20.5 si riferisce ad un convertitore a 3 bit.
In realtà, in ogni uscita intermedia del partitore non vi è proprio il valore corrispondente ad uno degli 8 gradini del fondo scala. Si è infatti utilizzata la tecnica dello spostamento della caratteristica del convertitore per dimezzare l'errore dello stesso, che passa da Ut,,/8 a Ufs/16 in valore assoluto, come ben si vede dalla Fig.20.6.

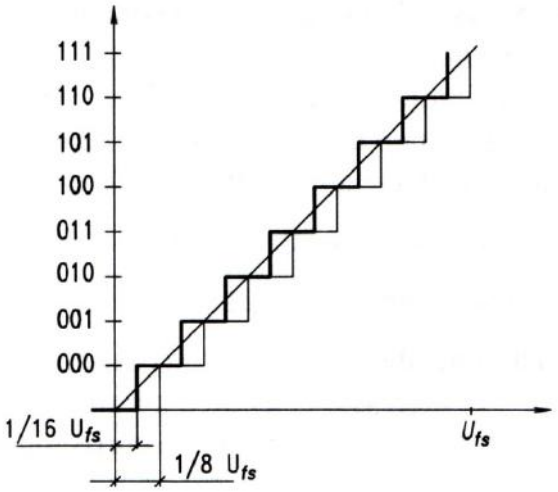
Ciascun terminale di uscita del partitore è posto in ingresso ad un comparatore analogico. All'altro ingresso di tutti i comparatori è posta la tensione che si vuole codificare in forma digitale.
Quando il circuito è operativo, tutti i comparatori sono contemporaneamente impegnati a confrontare la tensione ai loro capi con quella di ingresso U_{in}. Il risultato del raffronto si traduce nello stato logico di uscita corrispondente:

- stato 0 se la tensione di ingresso è minore di quella prelevata dal partitore;
- stato 1 se la tensione di ingresso è maggiore di quella prelevata dal partitore.

I1 priority encoder genera poi il codice che corrisponde al gradino in cui si situa il valore U_{in} in ingresso: considera infatti il comparatore a stato 1 avente "peso" più elevato e fornisce il codice binario corrispondente.

I convertitori flash sono dei dispositivi campionatori e non richiedono normalmente un circuito sample-hold, sebbene la sua presenza possa migliorare le prestazioni in alcune applicazioni.
Lo svantaggio principale di questo tipo di convertitore è costituito dal numero di comparatori richiesti: 2^n -1 per un codice a n bit. Il limite pratico è costituito dal convertitore a 10 bit, che richiede 1023 comparatori; oltre i 10 bit risulterebbero insostenibili le dimensioni ed il costo di produzione.

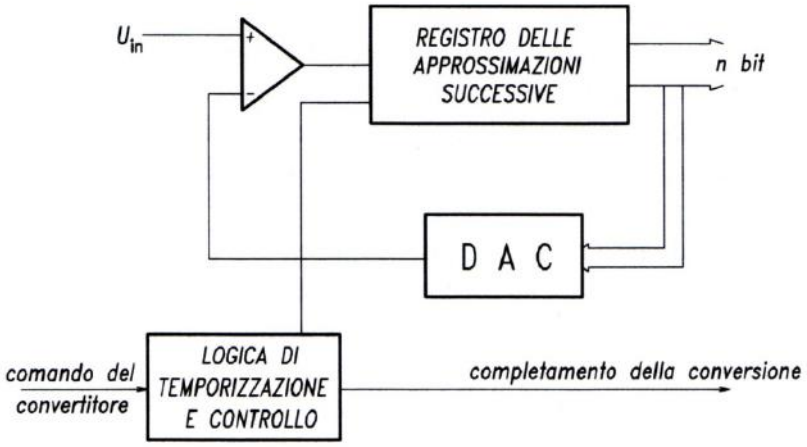


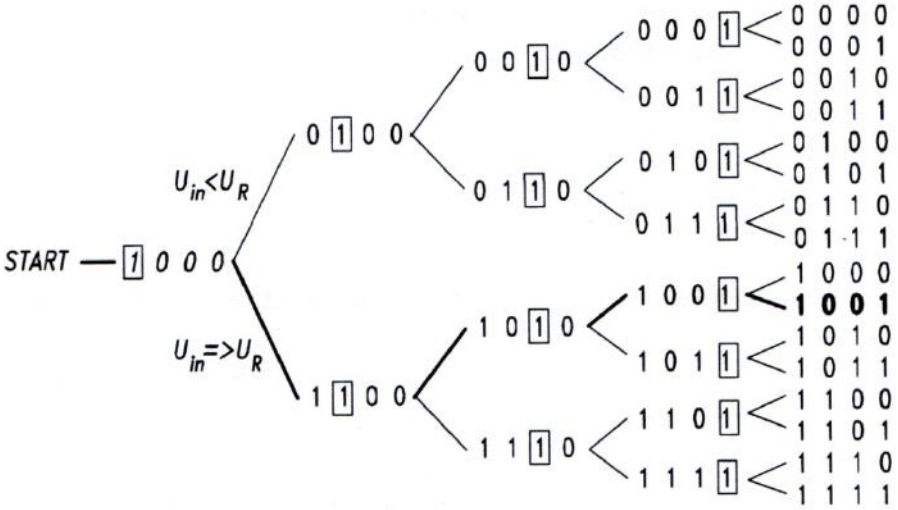


D'altra parte, si tratta del convertitore in assoluto più veloce: da 1 Msample/s fino a 500 Msample/s. Grazie a questa caratteristica, trova largo impiego in settori quali le comunicazioni, i radar, gli oscilloscopi digitali, gli analizzatori di forme d'onda, e gli elaboratori di segnali video.

Il convertitore ad approssimazioni successive è il tipo di convertitore di gran lunga più diffuso, perché è in grado di soddisfare un gran numero di esigenze. Usato in apparecchiature di qualunque tipo, dai modem all'elettronica per missili, deve la sua diffusione alla capacità di combinare valori relativamente alti di risoluzione (fino a 16 bit) con alte velocità massime di 1 Msample/s.
Questo convertitore appartiene alla categoria dei cosiddetti convertitori a retroazione: una rete logica sequenziale genera, secondo un'opportuna strategia, tutte le configurazioni binarie previste dal codice; tali configurazioni sono convertite da un DAC in una tensione che è confrontata con la tensione di ingresso; quando il comparatore cambia stato, la logica sequenziale interrompe la sequenza congelando il dato binario in uscita.
La struttura del convertitore è illustrata in Fig.20.7. La logica sequenziale è costituita da un registro ad approssimazioni successive il quale implementa quello che in informatica è chiamato algoritmo di ricerca dicotomica: si tratta di reperire- iell' insieme dei quanti di tensione il quanto in cui cade la tensione di ingresso, rvabilendo in quale metà si trova, poi in quale metà della metà, e così via. Come si vede, i bit sono aggiustati ad uno ad uno a partire dal MSB. In Fig.20.8 è illustrato un esempio, con tensione di ingresso nel nono gradino.
La presenza di un circuito sample-hold si rende talvolta necessario in relazione alla natura del segnale analogico da campionare.

Una realizzazione sotto forma di circuito integrato del convertitore ad approssimazione successive è costituito dall' ADC0801, un 8 bit in tecnologia CMOS.





I convertitori subranging combinano elementi della conversione flash con quelli della conversione sequenziale e permettono di ottenere prestazioni di compromesso: risoluzione fino a 16 bit e velocità fino a 40 Msample/s. Richiedono un circuito di sample-hold.

In netto contrasto con i convertitori fin qui descritti, i convertitori ad integrazione lavorano tipicamente a velocità dell'ordine dei 10 sample/s. Impiegati per le misure di precisione di segnali lentamente variabili, i convertitori integratori trovano il loro impiego in voltmetri e multimetri digitali e nella elaborazione dell'uscita di alcuni trasduttori.
Sono basati sulla generazione di una rampa di tensione lineare, ottenuta mediante un integratore che lavora su una tensione di riferimento, che è confrontata con la tensione in ingresso. Il tempo necessario perché la rampa raggiunga il valore dell'ingresso,misurato contando gli impulsi di un opportuno clock, è proporzionale alla tensione da misurare.
I convertitori a doppia rampa confrontano due rampe, una di carica di un condensatore verso la tensione di riferimento, l'altra di scarica verso la tensione incognita: in questo modo si ovvia alla imprecisione del primo metodo, basato sulla stabilità della pendenza della rampa, che dipende dalla stabilità nel tempo dei componenti integratori.

Per quanto riguarda la conversione D/A, la tecnica a resistenze pesate è la più semplice e quella che fornisce il fondamento per le altre.
Consideriamo il circuito della Fig.20.9, relativo ad un codice a 3 bit.

Come sappiamo dallo studio dell'amplificatore operazionale, il punto A è a massa virtuale e poiché l'ingresso dell' amp-op presenta resistenza infinita, lungo la resistenza di retroazione passa la somma delle correnti di ingresso:

$$I = I_0 + I_1 + I_2$$

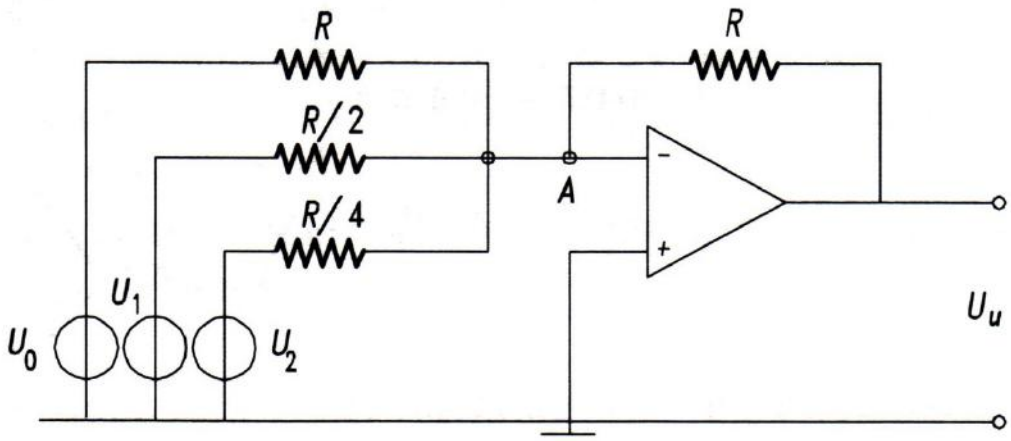
Ed essendo:

$$I = -(U_u / R) \quad ; \quad I_0 = U_0 / R \quad ; \quad I_1 = U_1 / (R/2) \quad ; \quad I_2 = U_2 / (R/4)$$

Risulta:

$$U_u = -(U_0 + 2U_1 + 4U_2)$$

In altre parole, la tensione di uscita dipende dalle tensioni di ingresso (che possono essere presenti o meno: bit 1 oppure bit 0 del codice) con "pesi" proporzionali alla posizione nel codice in ingresso.
Il convertitore a resistenze pesate offre il vantaggio di una struttura estremamente semplice, ma le resistenze devono essere di valore perfettamente calibrato, cosa non facile ad ottenersi.


R
R
R/2
R/4
A
U1
U0
U2
Uu