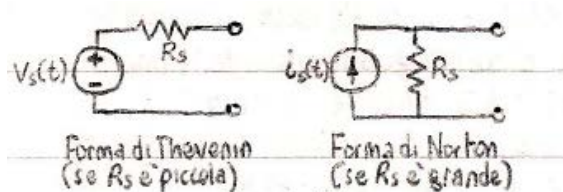


# Elettronica

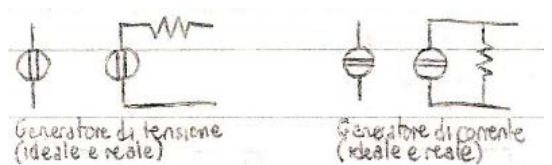
I segnali contengono informazioni: per estrarle, l'osservatore deve effettuare una *elaborazione del segnale*, attraverso un sistema elettronico. Perché questo sia possibile occorre quindi necessariamente convertire il segnale in un *segnale elettrico* (una tensione o una corrente), attraverso dei *trasduttori*: un segnale quindi è



una grandezza che varia nel tempo, che inoltre può essere rappresentata anche mediante il suo spettro in frequenza (come somma di segnali sinusoidali di differenti frequenze ed ampiezze). I segnali possono essere *analogici* o *digitali*, a seconda se essi assumono molteplici valori in ampiezza in diversi istanti di tempo o valori ben precisi. Un segnale

analogico può essere trasformato in digitale mediante un *campionamento*.

## Richiami sui circuiti lineari



Legge Ohm:  $V=RI$

Principio di equivalenza tra generatori: i due generatori coincidono con  $V_g=RI_g$

Leggi di Kirchhoff: 1) la somma algebrica delle correnti

entranti in un nodo del circuito è uguale a zero:

$$\sum_{i=1}^n I_i = 0$$

2) la somma algebrica delle differenze di potenziale che s'incontrano spostandosi su un circuito lungo una linea chiusa e finita è nulla. Le tensioni vanno prese positive se concordi con il verso di spostamento:

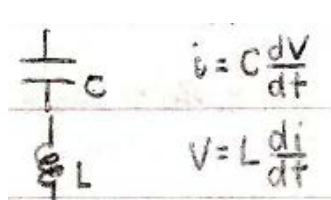
$$\sum_{i=1}^n V_i = 0$$

Teorema di Thevenin: un sistema lineare a una porta è equivalente a una rete con un generatore di tensione e una resistenza.

Teorema di Norton: un sistema lineare a una porta è equivalente a una rete con un generatore di corrente e una resistenza.

Generatori controllati: i valori delle grandezze elettriche sono legate ad un'altra grandezza elettrica presente nel circuito.

## Elementi reattivi



impedenza:  $z = \frac{1}{j\omega C}$  (risp. in frequenza)

$V_c(0^+) = V_c(0^-)$  e  $V_c(t) = V(\infty) - [V(\infty) - V(0)]e^{-\frac{t}{\tau}}$  con  $\tau = RC$  (al gradino)

Impedenza:  $z = j\omega L$  (risp. in frequenza)

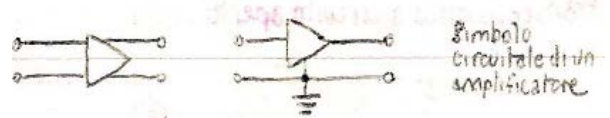
$I_L(0^+) = I_L(0^-)$  e  $I_L(t) = I_L(\infty) - [I_L(\infty) - I_L(0)]e^{-\frac{t}{\tau}}$  con  $\tau = LC$  (al gradino)

Resistenze in serie:  $R_{eq} = R_1 + R_2 + \dots + R_n$

Resistenze in parallelo:  $\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \dots + \frac{1}{R_n} \rightarrow R_{eq} = \frac{R_1 * R_2 * \dots}{R_1 + R_2 + \dots}$

## Amplificatori

La prima necessità nell'elaborazione dei segnali è quella di amplificarli (poiché i trasduttori forniscono segnali piccoli, di energie basse): ciò avviene tramite un *amplificatore*, che deve essere dotato di *linearità* per evitare distorsioni del segnale.



Un tale amplificatore è formato dalla relazione  $V_o(t) = A$

$V_i(t)$  dove  $A$ , una costante, rappresenta il *guadagno dell'amplificatore*. L'amplificatore di segnali è una rete due porte, e spesso viene rappresentato con un terminale in comune tra la porta in ingresso e quella di uscita, detto *massa del circuito*. Un amplificatore riceve in ingresso un segnale  $V_i(t)$  e fornisce in uscita, attraverso una resistenza di carica  $R_L$  un segnale  $V_o(t)$ : si definisce *guadagno di tensione* la relazione  $A_v = \frac{V_o}{V_i}$ . Si

definisce *guadagno di potenza* (che differenzia l'amplificatore da un trasformatore) la relazione:

$$A_p = \frac{V_o I_o}{V_i I_i} = \frac{\text{potenza trasferita al carico } (P_L)}{\text{potenza in ingresso } (P_i)} \quad \text{con } I_o = \frac{V_o}{R_L}$$

(corrente fornita al carico da  $A$ ).

Definendo inoltre il *guadagno di corrente*  $A_i = \frac{I_o}{I_i}$  si ricava che  $A_p = A_v A_i$ .

Poiché l'energia trasferita al carico è maggiore di quella assorbita dal generatore di segnale, è evidente che debba essere fornita energia

addizionale all'amplificatore: ciò avviene tramite il collegamento di generatori in continua, detti *alimentatori*.

Un *amplificatore* richiede due alimentatori: uno positivo di valore  $V_1$  (il cui polo negativo è messo a massa) e uno negativo di valore  $V_2$  (il cui polo positivo è messo a massa). Essi erogano corrente, rispettivamente  $I_1$  e  $I_2$ , per cui la potenza fornita all'amplificatore è:  $P_{dc} = V_1 I_1 + V_2 I_2$ .

Indicando con  $P_{diss}$  la potenza dissipata all'interno dell'amplificatore, si ha il bilancio energetico:

$$P_{dc} + P_i = P_L + P_{diss}$$

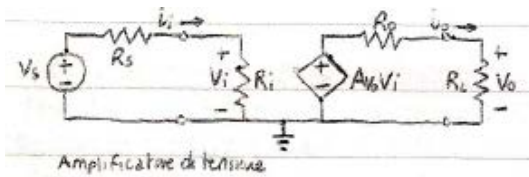
Dove  $P_L$  (potenza fornita al carico). Siccome  $P_i$  (la potenza assorbita) di solito è molto bassa, l'*efficienza* dell'amplificatore è definita come  $\eta = \frac{P_L}{P_{dc}} * 100$ .

Un amplificatore lavora in regime lineare solo per un limitato intervallo di tensioni di ingresso e di uscita: se alimentato da due generatori la tensione di uscita non può superare un limite positivo  $L^+$  e uno negativo  $L^-$ , oltre i quali l'amplificatore è in *saturazione*. In particolare il segnale di ingresso deve essere mantenuto nell'intervallo  $\frac{L^-}{A_v} < V_i < \frac{L^+}{A_v}$ . Nella realtà gli amplificatori presentano caratteristiche non sempre lineari: ad

esempio quelli alimentati da un solo generatore hanno il punto di lavoro decentrato rispetto all'origine. Per ovviare a ciò pur mantenendo l'ampiezza del segnale bassa, occorre *polarizzare* il circuito, applicando una tensione costante  $v_1$  che si somma alla tensione del segnale:  $v_i(t) = V_1 + v_i(t)$ ; in uscita si avrà  $v_o(t) = V_0 + v_o(t)$ , cioè  $V_o(t) = A_v v_i(t)$ , dove  $A_v$  è la pendenza lineare della caratteristica, cioè  $A_v = \left. \frac{dv_o}{dv_1} \right|_{in Q}$ .

Esistono diversi tipi di amplificatori, ma in particolare ne vengono analizzati quattro: L'amplificatore di tensione, di corrente, di transresistenza e di transconduttanza.

L'**amplificatore di tensione** è costituito da un generatore di tensione controllato in tensione con un fattore di amplificazione  $A_{v0}$ , una resistenza di ingresso  $R_i$  e una d'uscita  $R_o$ . Usando la regola del partitore di tensione si ha



$$V_0 = A_{v0} v_i \frac{R_L}{R_L + R_o}$$

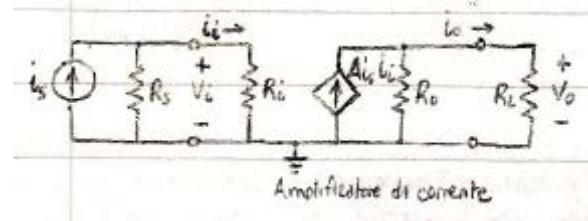
quindi il guadagno di tensione è  $A_v = \frac{V_0}{V_i} = A_{v0} \frac{R_L}{R_L + R_o}$ .

Affinché il guadagno non venga ridotto, la resistenza di uscita  $R_o$  deve essere molto più piccola della resistenza di carico  $R_L$ ; inoltre poiché  $R_L \rightarrow \infty$  si ha  $A_v = A_{v0}$ .  $A_v$  è detto **guadagno di tensione a circuito aperto**. Quando la resistenza d'ingresso  $R_i$  ha valore finito, all'ingresso dell'amplificatore è presente un altro partitore di tensione, quindi solo una frazione del segnale d'ingresso  $V_s$  arriva in ingresso all'amplificatore, cioè  $v_i = v_s \frac{R_i}{R_i + R_s}$ , per cui per non perdere una parte del segnale d'ingresso, la resistenza di ingresso deve essere molto maggiore di quella del generatore, cioè  $R_i \gg R_s$  (e  $R_o \ll R_L$ ). Un caso particolare è il **buffer di tensione**: uno stadio separatore da utilizzare quando  $R_s \gg R_L$  e quindi quando si avrebbe una forte attenuazione del segnale; in tal caso si utilizza questo tipo di amplificatore, con elevata resistenza d'ingresso (tale che  $R_i \gg R_s$ ) e una bassa resistenza di uscita (tale che  $R_o \ll R_L$ ), pur avendo un guadagno unitario  $A_{vs}=1$ , ma tale che  $V_0 \cong V_s$ .

Il guadagno di tensione complessivo:  $A_{vs} = \frac{V_0}{V_s} = A_{v0} \frac{R_L}{R_L + R_o} * \frac{R_i}{R_i + R_s}$ ;  $V_0 = V_s \frac{R_i}{R_i + R_s} * \frac{R_L}{R_L + R_o} \rightarrow V_0 \cong V_s$ ,

se  $R_i \rightarrow \infty$ ,  $R_o \rightarrow 0$ ,  $A_{vs} \rightarrow 1$ .

Sebbene sia molto più usato l'amplificatore di tensione, è importante anche l'**amplificatore di corrente**: esso è costituito da un generatore di corrente controllato in corrente con amplificatore  $A_{is}$ , Resistenza in ingresso  $R_i$  e di uscita  $R_o$ . Usando la regola del partitore di corrente si ottiene  $I_0 = A_{is} i_i \frac{R_o}{R_L + R_o}$



quindi il guadagno di corrente è  $A_i = \frac{i_0}{i_i} = A_{is} \frac{R_o}{R_L + R_o}$ .

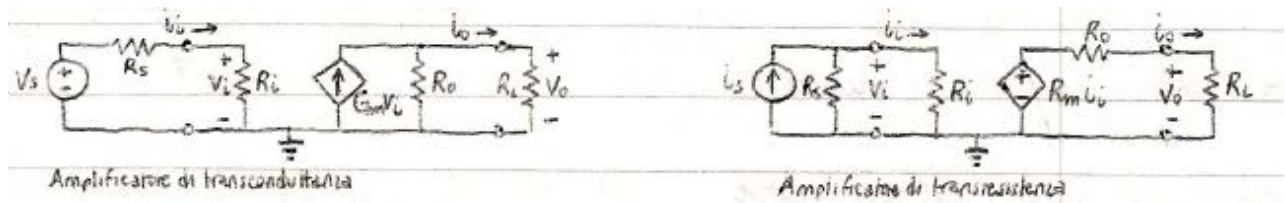
Per evitare perdita di guadagno, la resistenza di uscita  $R_o$  deve essere molto maggiore della resistenza di carico  $R_L$ ; inoltre se  $R_L = 0$  si ha  $A_i = A_{is}$ , quindi  $A_{is}$  è chiamata **guadagno di corrente in corto circuito**.

Anche in ingresso vi è un partitore di corrente, e solo una sua frazione giunge in ingresso all'amplificatore:  $i_i = I_s \frac{R_s}{R_i + R_s}$ , quindi per non perdere parte del segnale di ingresso, la resistenza di ingresso deve essere molto minore di quella del generatore, cioè  $R_i \ll R_s$  (e  $R_o \gg R_L$ ). Un caso particolare è il **buffer di corrente**: si utilizza uno stadio separatore quando si ha alta impedenza d'uscita ( $R_s \ll R_L$ ) con guadagno unitario, ma tale che  $I_0 \cong I_s$ .

$A_i = \frac{i_0}{i_s} = A_{is} \frac{R_o}{R_L + R_o} * \frac{R_s}{R_i + R_s}$ ;  $I_0 = I_s \frac{R_s}{R_i + R_s} * \frac{R_o}{R_L + R_o} \rightarrow I_0 \cong I_s$ , se  $R_i \rightarrow 0$ ,  $R_o \rightarrow \infty$ ,  $A_i \rightarrow 1$ .

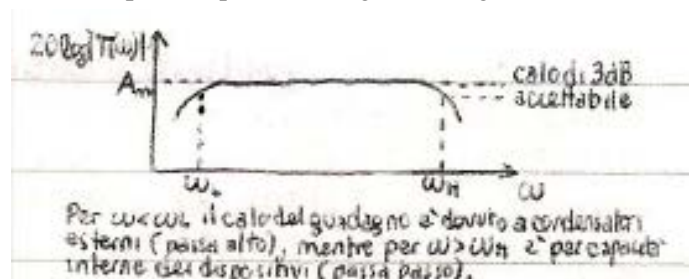
Vi sono infine l'**amplificatore di Transconduttanza**, che ha un segnale di tensione in ingresso e fornisce un segnale di corrente in uscita e il cui parametro di guadagno  $G_m$  è detto **transconduttanza di corto circuito** (amplificatore ideale per  $R_0 = \infty$   $R_i = \infty$ ); e l'**amplificatore di transresistenza**, che ha un segnale di corrente in ingresso e fornisce un segnale di tensione in uscita e il cui parametro di guadagno  $R_m$  è detto **transresistenza a circuito aperto** (amplificatore ideale per  $R_0 = 0$   $R_i = 0$ ).

I quattro modelli di amplificatori sono *unidirezionali*, cioè, il segnale fluisce in una sola direzione, dall'ingresso all'uscita.



Individuare la **risposta in frequenza** di un amplificatore significa esaminare la sua risposta a segnali d'ingresso sinusoidali di varia frequenza. Ad un ingresso di questo tipo corrisponde in uscita un segnale con stessa frequenza  $\omega$ , ma in generale un'ampiezza diversa  $V_o$  rispetto a quella del segnale d'ingresso  $V_i$  e uno sfasamento  $\phi$ : indicando con  $T(\omega)$  la funzione di trasferimento dell'amplificatore, si ha che la risposta dell'amplificatore è completamente descritta da  $|T(\omega)| = \frac{V_o}{V_i}$  e la fase  $\angle T(\omega) = \phi$ .

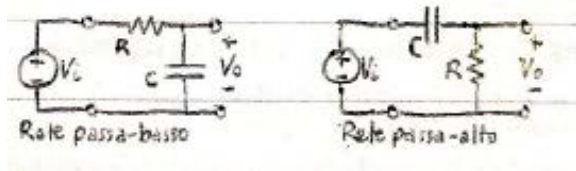
Quindi individuare la risposta in frequenza significa ricavare questi due valori al variare della frequenza  $\omega$ . L'amplificazione rimane costante in



un intervallo di frequenza tra  $\omega_L$  e  $\omega_H$ , mentre al di sotto di  $\omega_L$  e al di sopra di  $\omega_H$  è minore, poiché il guadagno decresce: tale intervallo è detto **larghezza di banda** o **banda passante**. Un fattore di merito per l'amplificatore è il prodotto **Banda \* Guadagno**:

$$(\omega_H - \omega_L) * A_m = \text{costante}.$$

Per determinare la **risposta in frequenza** di un amplificatore, è utile conoscere quella delle reti STC, con una singola costante di tempo: esse si possono dividere per lo più in due categorie, caratterizzate da risposte differenti; **passa-basso** e **passa-alto**. Infatti, sapendo che entrambe le funzioni di trasferimento si possono esprimere mediante il partitore di tensione (formato da una resistenza e un condensatore con impedenza al variare della frequenza:



$Z = \frac{1}{j\omega}$ ), nel caso di una rete passa-basso diminuisce all'aumentare della frequenza e tende a zero per  $\omega$  che tende all'infinito (quindi fa passare le basse frequenze attenuandole poco o non attenuandole, mentre attenua quelle ad alta frequenza), invece nel caso di una rete passa-alto diminuisce con  $\omega$ , annullandosi per  $\omega=0$ . Nei due casi, governati dalla costante di tempo  $\tau=RC$ , si ha:

$$V_o = \frac{1/sC}{R + 1/sC} * V_i = \frac{1}{1 + sRC} * V_i \rightarrow T(s) = \frac{1}{1 + sRC} \rightarrow \text{con } \omega_0 = \frac{1}{\tau} = \frac{1}{RC} \rightarrow T(j\omega) = \frac{1}{1 + j(\omega/\omega_0)}$$

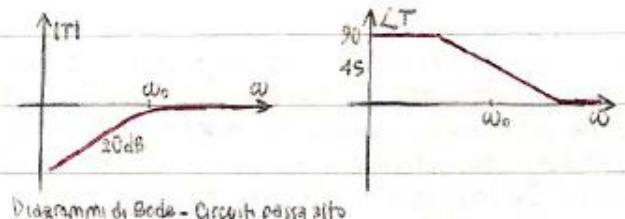
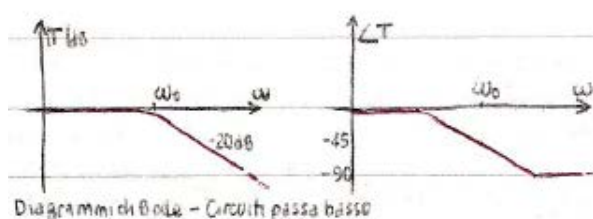
**Passa-Basso** :

$$\begin{cases} \text{se } \omega \ll \omega_0 \rightarrow |T| = 1, |T|_{db} = 0 \\ \text{se } \omega = \omega_0 \rightarrow |T| = \frac{1}{\sqrt{2}}, |T|_{db} = -3 \text{ db} \\ \text{se } \omega \gg \omega_0 \rightarrow |T| = \frac{\omega_0}{\omega} \end{cases}$$

$$V_o = \frac{R}{R + 1/sC} * V_i = \frac{1}{1 + 1/sRC} * V_i \rightarrow T(s) = \frac{1}{1 + 1/sRC} \rightarrow \text{con } \omega_0 = \frac{1}{\tau} = \frac{1}{RC} \rightarrow T(j\omega) = \frac{1}{1 - j(\omega_0/\omega)}$$

**Passa-Alto** :

$$\begin{cases} \text{se } \omega \ll \omega_0 \rightarrow |T| = \frac{\omega}{\omega_0} \\ \text{se } \omega = \omega_0 \rightarrow |T| = \frac{1}{\sqrt{2}}, |T|_{db} = -3 \text{ db} \\ \text{se } \omega \gg \omega_0 \rightarrow |T| = 1, |T|_{db} = 0 \end{cases}$$



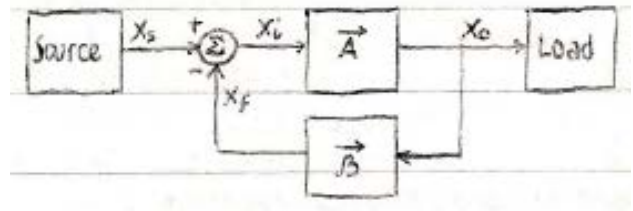
Gli amplificatori possono essere classificati in base alla forma della loro risposta in ampiezza; essi possono essere **accoppiati direttamente** (per cui vengono amplificate anche le basse frequenze), **accoppiati capacitivamente** (tramite condensatori, per cui è amplificato solo un intervallo in frequenza) o possono essere **amplificati accordati** (che hanno come risposta un picco attorno ad una frequenza, le altre sono filtrate).

**La controreazione** (feedback) può essere sia *positiva* che *negativa*. Nel progetto degli amplificatori la controreazione negativa è usata per avere delle proprietà:

- 1) *Stabilizzare il guadagno*, cioè rendere il guadagno meno sensibile alle variazioni del circuito (ad esempio della temperatura).
- 2) *Ridurre la distorsione*: fa sì che il segnale d'uscita sia direttamente proporzionale all'ingresso (cioè il guadagno è indipendente dall'ampiezza).
- 3) *Ridurre l'effetto del rumore*: minimizzare il contributo da parte dei segnali elettrici indesiderati.
- 4) *Controllare le impedenze di ingresso e di uscita*.
- 5) *Estendere la banda passante dell'amplificatore*.

Tutti questi effetti positivi vengono ottenuti a prezzo di una riduzione del guadagno: tale fattore è detto *tasso di controreazione* ( $\beta$ ). L'amplificatore ad anello aperto

ha un guadagno  $A$ , quindi l'uscita  $X_0$  è legata all'ingresso  $X_i$  tramite la relazione:  $X_0 = A X_i$ . L'uscita è inviata ad un carico e ad una rete di controreazione, che produce una replica  $X_f$  dell'uscita:  $X_f = \beta X_0$ . Il segnale di retroazione  $X_f$  viene sottratto a quello



fornito dal generatore  $X_s$ , per alterare il segnale di ingresso dell'amplificatore:  $X_i = X_s - X_f$  (questa sottrazione rende negativa la controreazione). Il guadagno dell'amplificatore controreazionato si ottiene combinando le equazioni:  $A_f = \frac{X_0}{X_f} = \frac{A}{1 + \beta A}$ , dove la quantità  $A\beta$  è detta *guadagno ad anello* (che deve essere positiva se è controreazione negativa). Esso è minore di un fattore  $1 + A\beta$ , se  $A\beta \gg 1$  allora  $A_f \cong \frac{1}{\beta}$  (cioè dipende solo dalla controreazione) e inoltre  $X_f \cong X_s$ , essendo  $X_f = \frac{A\beta}{1 + \beta A} X_s$ .

Per determinare analiticamente la stabilizzazione del guadagno si differenziano entrambi i membri

$$A_f = \frac{A}{1 + \beta A}, \text{ ottenendo: } \frac{dA_f}{dA} = \frac{1 + \beta A - \beta A}{(1 + \beta A)^2} \rightarrow dA_f = \frac{dA}{(1 + \beta A)^2}$$

che, dividendola per l'equazione di partenza porta a:  $\frac{dA_f}{dA} = \frac{1}{(1 + \beta A)} * \frac{dA}{A}$ .

Si evince quindi che la variazione di  $A_f$  (dovuta a variazioni del circuito) è minore della variazione di  $A$  del fattore di controreazione  $1 + \beta A$ , detto anche *fattore di stabilizzazione*.

Per quanto riguarda l'allargamento della banda (ad esempio per un amplificatore di alte frequenze) si ha:

$A(s) = \frac{A_m}{1 + s/\omega_H}$  con  $A_m$  che rappresenta il guadagno di banda intermedio e  $\omega_H$  la frequenza di taglio allora applicando a questo amplificatore la controreazione otteniamo un guadagno ad anello chiuso:

$$A_f(s) = \frac{A(s)}{1 + \beta A(s)} \rightarrow \text{sostituendo } A(s): A_f(s) = \frac{A_m / (1 + A_m \beta)}{1 + s/\omega_H (1 + A_m \beta)}.$$

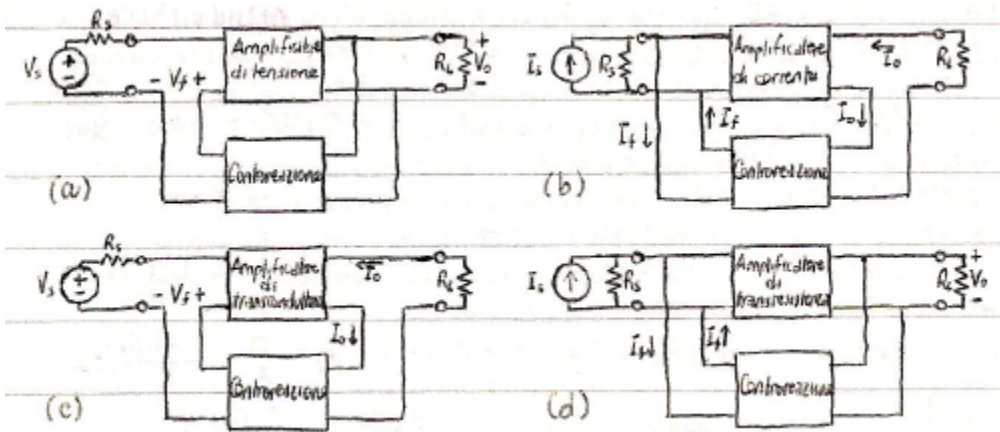
Da cui si vede una diminuzione del guadagno a centro-banda ( $A_m / (1 + A_m \beta)$ ).

E un aumento della frequenza di taglio ( $\omega_{Hf} = \omega_H (1 + A_m \beta)$ ).

Analogamente si calcola per le basse frequenze: in ogni caso si mantiene costante il prodotto Banda\*Guadagno.



A seconda del tipo di amplificatore vi è il tipo di controreazione appropriato (in totale quindi quattro tipi).



Nel caso di un amplificatore di tensione, la grandezza da prelevare in uscita è una tensione, quindi il segnale  $X_f$  deve essere una tensione che deve essere sommata a quella del generatore di segnale: è una controreazione **serie-parallelo** (fig (a) ). In un

amplificatore di corrente, la grandezza da prelevare in uscita è una corrente, quindi anche il segnale di controreazione dovrebbe essere una corrente, da essere inserita in parallelo alla corrente del generatore: è una controreazione **parallelo-serie** (fig(b)). Infine per un amplificatore di transconduttanza, che ha in ingresso una tensione e in uscita una corrente, si parla di controreazione **serie-serie** (fig(c)) mentre per un amplificatore di transresistenza, che ha in ingresso una corrente e in uscita una tensione, si ha una controreazione **parallelo-parallelo** (fig(d)).

La controreazione serie-parallelo inoltre porta ad un aumento della resistenza d'ingresso e ad una diminuzione di quella d'uscita. Sia il guadagno ad anello chiuso  $A_f = \frac{V_o}{V_f} = \frac{A}{1+\beta A}$

sapendo che  $V_s = V_i + \beta V_o = V_i + A\beta V_i = V_i (1 + A\beta)$  per la controreazione, si ha:

$$R_{if} = \frac{V_s}{I_i} = \frac{V_s}{V_i/R_i} = R_i \frac{V_s}{V_i}$$

Allora sostituendo  $V_s$ :  $R_{if} = R_i \frac{V_i}{V_i} (1 + A\beta) = R_i (1 + A\beta)$  (cioè è maggiore del fattore  $1 + A\beta$ ).

Mentre per individuare la resistenza di uscita, si pone  $V_s$  a zero e si applica una tensione di prova in uscita all'amplificatore ( $V_x$ ):

$$R_{of} = \frac{V_x}{I_x} \rightarrow \text{sapendo che } V_x = I_x R_o + A V_i = I_x R_o - A\beta V_x, \text{ cioè } V_x (1 + A\beta) = I_x R_o$$

Allora ricavando e sostituendo  $V_x$ :  $R_{of} = \frac{R_o}{1 + A\beta}$  (minore per  $1 + A\beta$ ).

In generale però si ha che la relazione tra  $R_{if}$  e  $R_i$  (e  $R_{of}$  e  $R_o$ ) dipende solo dal metodo del confronto: il confronto in serie quindi aumenta sempre la resistenza d'ingresso (o di uscita) e quello in parallelo diminuisce (anche quella di uscita). Infatti, nel caso di controreazione serie-serie, sapendo che  $A = \frac{I_o}{V_i}$ ;

$$\beta = \frac{V_f}{I_o} \text{ e } A_f = \frac{I_o}{V_f} = \frac{A}{1 + A\beta} \text{ con calcoli analoghi ai precedenti si ricava che } R_{if} = R_i (1 + A\beta);$$

per calcolare la resistenza di uscita invece si pone  $V_s = 0$  e si interrompe il circuito di uscita e si applica una corrente di prova  $I_t$  (al posto di  $R_i$ ). in questo caso  $V_i = -V_f = -\beta I_t$ ,

quindi ai capi dell'amplificatore:  $V = (I_t - A V_i) R_o = (I_t + A\beta I_t) R_o$  e, sapendo che  $R_{of} = \frac{V}{I_t}$ , sostituendo  $V$  si ricava  $R_{of} = (1 + A\beta) R_o$  (un aumento della resistenza d'uscita, essendo serie-serie).

Infine, per la configurazione parallelo-parallelo (per cui si ha  $A_f = \frac{V_o}{I_s} = \frac{A}{1 + A\beta}$ )

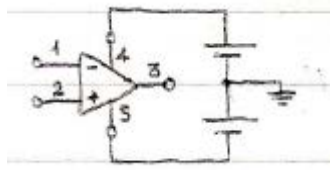
le resistenze sono pari a:  $R_{if} = \frac{R_i}{1 + A\beta}$  e  $R_{of} = \frac{R_o}{1 + A\beta}$

mentre per quella parallelo-serie ( per cui si ha  $A_f = \frac{I_o}{I_s} = \frac{A}{1 + A\beta}$  ) le resistenze sono pari a:  $R_{if} = \frac{R_i}{1 + A\beta}$  e

$R_{of} = (1 + A\beta) R_o$ .

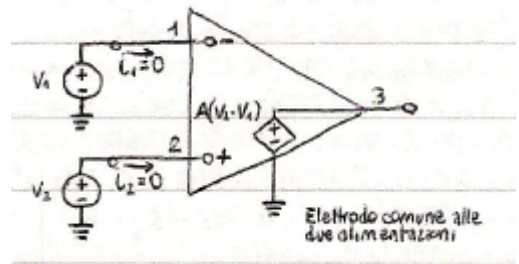
## Amplificatori operazionali

L'amplificatore operazionale ha tre terminali principali: due d'ingresso e uno d'uscita: a maggior parte di essi inoltre richiede due alimentazioni in continua, quindi sono presenti altri due terminali, rispettivamente

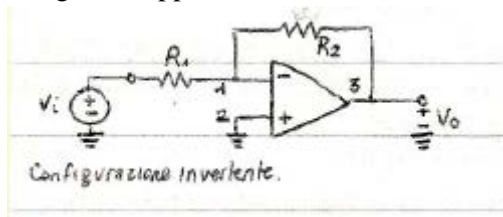


connessi a una tensione positiva  $V^+$  e una negativa  $V^-$ , che hanno massa in comune: è importante notare che nessun terminale dell'amplificatore operazionale è connesso fisicamente a massa. **L'amplificatore operazionale ideale** rileva la differenza tra i segnali di tensione applicati ai suoi due terminali di ingresso ( $V_2 - V_1$ ), moltiplica tale differenza per un numero  $A$  e presenta in uscita la tensione  $A(V_2 - V_1)$  ( $V_1$  sta per tensione tra terminale 1 e massa,

analogamente a  $V_2$ ). L'amplificatore operazionale non assorbe corrente sui terminali d'ingresso ( $i_1=0$ ,  $i_2=0$ ), cioè la resistenza di ingresso di un amplificatore operazionale è infinita ( $R_i = \infty$ ); in uscita la tensione tra il terminale e la massa è sempre pari a  $A(V_2 - V_1)$ , indipendentemente dalla corrente fornita all'impedenza di carica, cioè la resistenza d'uscita è nulla ( $R_0=0$ ). L'uscita inoltre è in fase con  $V_2$  (hanno lo stesso segno) ed è in opposizione di fase con  $V_1$  (hanno segno opposto): perciò il terminale 1 è detto **terminale di ingresso invertente**, il terminale 2 è detto **terminale di ingresso non invertente**. L'amplificatore operazionale ideale è a **ingresso differenziale** ( $V_2 - V_1$ ) e **uscita sbilanciata**: se  $V_2 = V_1$ , allora l'uscita  $V_0$  è idealmente nulla (questa proprietà è detta **reiezione di modo comune**). Inoltre, è caratterizzato da un guadagno che rimane costante da frequenza zero fino ad una frequenza infinita: amplificano cioè i segnali di ogni frequenza con lo stesso guadagno (cioè sono sempre saturi). Poiché il guadagno  $A$  è idealmente finito, l'amplificatore operazionale non viene usato nella configurazione ad anello aperto ma con una controeazione dove una controeazione collega l'uscita con l'ingresso dell'amplificatore opzionale.



In figura è rappresentata una **controeazione negativa**: il resistore  $R_2$  è connesso tra il terminale di uscita (3)



dell'amplificatore ed il terminale di ingresso invertente (o negativo). Se fosse stato connesso invece tra il terminale di uscita e il terminale di ingresso non invertente sarebbe stata una controeazione positiva. Il terminale 2 inoltre è posto a massa ed è presente un resistore  $R_1$  tra il terminale 1 e il generatore di segnale di tensione  $V_i$ ; l'uscita è presa tra il terminale 3 e massa.

Questo tipo di configurazione è detto **configurazione invertente** (poiché il guadagno ad anello chiuso è associato all'ingresso 1 negativo). Assumendo che l'amplificatore sia ideale, si vuole analizzare il **guadagno ad anello chiuso**  $G = \frac{V_0}{V_i}$ . Poiché  $A \rightarrow \infty$ , per definizione si ha  $V_2 - V_1 = \frac{V_0}{A} = 0$ , e da ciò segue che

$V_2 \cong V_1$ , cioè la tensione  $V_1$  è molto prossima a  $V_2$  essendo il guadagno  $A$  prossimo a infinito (si tratta di un **cortocircuito virtuale**, cioè qualsiasi tensione sia presente sul terminale 1 è presente automaticamente anche sul terminale 2); siccome il terminale 2 è connesso a massa si ha  $V_2=0$ , quindi  $V_1 \cong 0$  (il terminale è una **massa virtuale**, cioè è a tensione nulla anche se non è fisicamente connesso a massa). La corrente  $i_1$  che scorre su  $R_1$ :  $i_1 = V_i \cdot \frac{V_1}{R_1} = \frac{V_i}{R_1}$  non può entrare nell'amplificatore operazionale (essendo ideale ha impedenza d'ingresso infinita), che quindi scorre attraverso  $R_2$  nel terminale 3 ( $i_2$ ). Si ha quindi  $i_1 = i_2$ , da ciò:

$V_0 = V_1 - i_2 R_2 = 0 - \frac{V_i}{R_1} R_2 = -\left(\frac{V_i}{R_1}\right) R_2$ , quindi il guadagno ad anello chiuso  $G = \frac{V_0}{V_i} = -\frac{R_2}{R_1}$  dove il segno negativo (config. invertente) indica in uscita un'inversione del segnale (sfasamento  $180^\circ$ ): esso risulta dipendere solo dai componenti passivi esterni  $R_1$  e  $R_2$  ed è idealmente indipendente dal guadagno dell'amplificatore operazionale, poiché  $A \rightarrow \infty$ .



Nell'ipotesi che il guadagno ad anello aperto  $A$  sia finito, indicato con  $V_0$  la tensione d'uscita, la tensione tra i due terminali d'ingresso dell'amplificatore operazionale è pari a  $\frac{V_0}{A}$ , ed in particolare poiché l'ingresso positivo è connesso a massa, la tensione al terminale d'ingresso negativo è pari a  $-\frac{V_0}{A}$ . La corrente  $i_1$  che

scorre su  $R_1$  è pari a:  $i_1 = \frac{V_i - (-\frac{V_0}{A})}{R_1} = \frac{V_i + \frac{V_0}{A}}{R_1}$  da cui sapendo che  $i_1 = i_2$  (cioè  $i_1$  scorre interamente su  $R_2$ ):

$$V_0 = -i_2 R_2 + V_1 = -\frac{V_0}{A} - i_2 R_2 = -\frac{V_0}{A} - \left(\frac{V_i + \frac{V_0}{A}}{R_1}\right) R_2 \text{ quindi } V_0 = -V_i \frac{R_2}{R_1} - \frac{V_0}{A} \frac{R_2}{R_1} - \frac{V_0}{A} \text{ quindi}$$

$$V_0 = \left(1 + \frac{1}{A} \frac{R_2}{R_1} + \frac{1}{A}\right) = -V_i \frac{R_2}{R_1} \text{ da cui il guadagno ad anello chiuso: } G = \frac{V_0}{V_i} = \frac{-\frac{R_2}{R_1}}{1 + \frac{1}{A} \left(1 + \frac{R_2}{R_1}\right)} \text{ e la dipendenza di } A \text{ è}$$

minima se  $1 + \frac{R_2}{R_1} \ll A$ . Infine si ha  $R_{in} = \frac{V_i}{I_1} = R_1$  (poiché  $i_1 = \frac{V_i}{R_1}$ ) e  $R_0 = 0$ .

Invece di usare due resistori  $R_1$  e  $R_2$ , si possono utilizzare due impedenze  $Z_1$  e  $Z_2$ : in generale quindi la funzione di trasferimento ad anello chiuso è pari a  $\frac{V_0}{V_i} = -\frac{Z_2}{Z_1}$ .

Come caso particolare si può considerare  $Z_1 = R$  e  $Z_2 = \frac{1}{sC}$  (cioè

$Z_1$  è un resistore e  $Z_2$  un condensatore), da cui  $\frac{V_0}{V_i} = -\frac{1}{sRC}$  che

per frequenze reali  $s = j\omega$  diventa  $\frac{V_0}{V_i} = -\frac{1}{j\omega RC}$ . Questa

corrisponde ad una integrazione, cioè  $V_0(t)$  risulterà integrale di

$V_i(t)$ : infatti, nel dominio del tempo si ha  $I_1 = \frac{V_i(t)}{R}$ .

Se all'istante  $t=0$  la tensione ai capi del condensatore è  $V_C$ , allora si ha:  $V_0(t) = V_C - \frac{1}{RC} \int_0^t V_i(t) dt$  con  $RC$

*costante di tempo di integrazione*. Quindi  $V_0(t)$  è l'integrale nel tempo di

$V_i(t)$ , e la tensione  $V_0$  è la condizione iniziale del processo: questo

circuito integrale è invertente a causa del segno negativo associato alla

funzione ed è detto *integratore di Miller* (o *integratore invertente*). Esso

ha una risposta in ampiezza che è identica a una rete passa-basso però

con frequenza di taglio  $\omega_0$  nulla: ciò significa che vi corrisponde un

guadagno ad anello chiuso infinito (cioè per  $\omega = 0$  si ha  $A \rightarrow \infty$ ). Per

rendere finito il suo valore occorre connettere una resistenza  $R_2$  di valore elevato in parallelo al condensatore dell'integrale (che diventa però non

ideale), per cui si ha  $Z_1 = R_1$ ,  $Z_2 = R_2 \parallel \frac{1}{sC} = \frac{R_2 \cdot \frac{1}{sC}}{R_2 + \frac{1}{sC}} = \frac{R_2}{1 + sR_2 C}$  dove  $-\frac{R_2}{R_1}$  è il

guadagno ad anello chiuso dell'amplificatore.

Un altro caso particolare è il circuito che compie operazione di derivazione; si ha per  $Z_1 = \frac{1}{sC}$  e  $Z_2 = R$  (cioè  $Z_1$  è un condensatore e  $Z_2$  un Resistore) da cui  $\frac{V_0}{V_i} = -\frac{Z_2}{Z_1} = -sRC$ , che per

frequenze reali  $s = j\omega$  diventa  $\frac{V_0}{V_i} = -j\omega RC$ : ciò corrisponde ad

un'operazione di differenziazione.

Infatti si ha  $i_1(t) = C \frac{dV_i(t)}{dt}$ , da cui  $V_0(t) = -i_1(t)R = -RC \frac{dV_i(t)}{dt}$ . Questo

tipo di circuito è detto quindi *derivatore*: ha una risposta in ampiezza

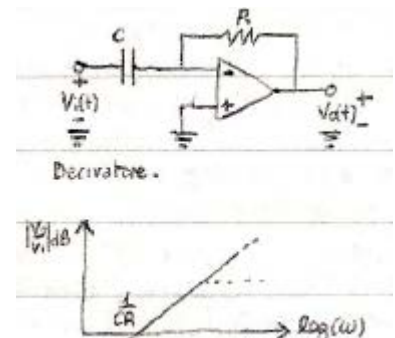
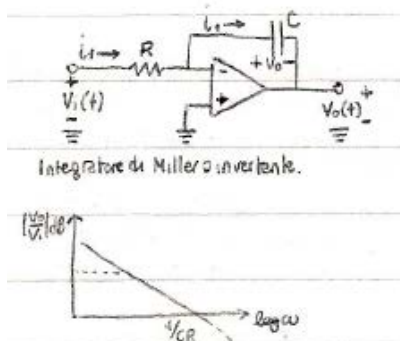
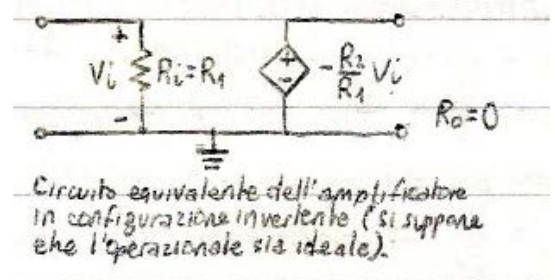
identica a quella della rete STC passa-alto, però con frequenza di taglio

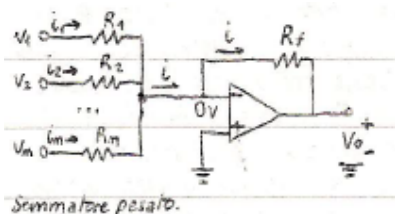
infinita; maggiore è la frequenza del segnale più esso viene amplificato.

Ciò rende il derivatore un "amplificatore di rumore", poiché le interferenze

hanno quasi sempre frequenza maggiore: affinché possa lavorare effettivamente come derivatore (e in

particolare solo per le basse frequenze, mentre per quelle alte si tratta di un amplificatore) occorre connettere una piccola resistenza in serie al condensatore (che rende però il derivatore non ideale).





Sommatore pesato.

Come ultima applicazione della configurazione invertente si ha il **sommatore pesato**. Si ha in questo circuito una resistenza  $R_f$  nell'anello di controreazione negativa, ma più segnali di ingresso  $V_1, V_2, \dots, V_n$  applicati su resistori  $R_1, R_2, \dots, R_n$  connessi al terminale invertente dell'amplificatore operazionale: poiché esso è ideale, presenta in ingresso massa virtuale, quindi corrente  $i = i_1 + i_2 + \dots + i_n$  (con  $i_1 = \frac{V_1}{R_1}$ ,  $i_2 = \frac{V_2}{R_2}$ ,  $\dots$ ,  $i_n = \frac{V_n}{R_n}$ ) è costretta a

scorrere in  $R_f$ . La tensione di uscita quindi è pari a:  $V_o = 0 - iR_f = -iR_f = -(\frac{R_f}{R_1}V_1 + \frac{R_f}{R_2}V_2 + \dots + \frac{R_f}{R_n}V_n)$  cioè è somma pesata degli ingressi.

La seconda configurazione possibile con una controreazione negativa è la **configurazione non invertente**: in questo caso il segnale di ingresso  $V_i$  è applicato al terminale d'ingresso positivo dell'amplificatore operazionale mentre un terminale di  $R_1$  è connesso a massa. Per determinare il **guadagno ad anello chiuso**

$G = \frac{V_o}{V_i}$  si assume che l'amplificatore operazionale ideale con guadagno finito e corto circuito virtuale tra i terminali di ingressi, quindi

$V_2 - V_1 = \frac{V_o}{A} = 0$  per  $A \rightarrow \infty$ , da cui  $V_1 = V_2$ , dove  $V_2 = V_i$ . La corrente

che scorre in  $R_1$  si può determinare come:  $i_1 = \frac{V_i}{R_1}$  che, a causa dell'impedenza infinita d'ingresso

dell'amplificatore operazionale, scorre interamente su  $R_2$ . La tensione di uscita quindi è

$V_o = V_i + i_2 R_2 = V_i + \frac{V_i}{R_1} R_2$  da cui si ricava il guadagno ad anello chiuso  $G = \frac{V_o}{V_i} = 1 + \frac{R_2}{R_1}$ : in particolare è un

valore positivo, da cui il nome non invertente. Supponendo che il guadagno ad anello aperto sia finito, si ha

$V_2 - V_1 = \frac{V_o}{A}$ , da cui essendo  $V_2 = V_i$  allora  $V_1 = V_i - \frac{V_o}{A}$ .

Sapendo che  $i_1 = \frac{V_1}{R_1} = \frac{V_i - \frac{V_o}{A}}{R_1}$  e che  $V_o = V_1 + i_2 R_2 = i_1 R_1 + i_2 R_2$  (con  $i_1 = i_2$ ) si

ricava:  $V_o = \frac{V_i - \frac{V_o}{A}}{R_1} (R_1 + R_2) = V_i (1 + \frac{R_2}{R_1}) - \frac{V_o}{A} (1 + \frac{R_2}{R_1}) \rightarrow$

$\rightarrow V_o = [1 + \frac{1}{A} (1 + \frac{R_2}{R_1})] V_i = V_i (1 + \frac{R_2}{R_1})$  quindi il guadagno ad anello chiuso è

$G = \frac{V_o}{V_i} = \frac{1 + \frac{R_2}{R_1}}{1 + \frac{1}{A} (1 + \frac{R_2}{R_1})}$  e la dipendenza da  $A$  è minima se  $1 + \frac{R_2}{R_1} \ll A$ . Infine si ha  $R_{in} = \infty$  e  $R_o = 0$ .

Nel caso in cui si voglia avere un trasformatore di impedenza o un amplificatore di potenza anziché un amplificatore di tensione è possibile trasformare questa configurazione in un buffer: imponendo  $R_1 = \infty$  e  $R_2 = 0$  si ottiene un amplificatore a guadagno unitario, e questo tipo di circuito è detto **inseguitore di tensione**, per il quale nel caso ideale si ha  $V_o = V_i$ ,  $R_{in} = \infty$  e  $R_o = 0$ .

Nella realtà il guadagno differenziale ad anello aperto di un amplificatore operazionale non è infinito, ma è elevato per le basse frequenze e diminuisce con l'aumentare della frequenza (comportamento passa-basso, amplificatori operazionali **compensati internamente**). Si ha  $A(s) = \frac{A_0}{1 + \frac{s}{\omega_b}}$  cioè per frequenze reali  $S = j\omega$ :

$A(j\omega) = \frac{A_0}{1 + \frac{j\omega}{\omega_b}}$ . Per frequenze  $\omega \gg \omega_b$  il guadagno si approssima con  $A(j\omega) = \frac{A_0 \omega_b}{j\omega}$  da cui si evince che il

guadagno diventa unitario per una frequenza  $\omega_t = A_0 \omega_b$  detta **banda di guadagno unitario** (che si indica anche con  $f_t = \frac{\omega_t}{2\pi}$ ). Analizzando quindi la risposta in frequenza degli amplificatori ad anello chiuso si ricava:

•  $\frac{V_o}{V_i} = \frac{\frac{R_2}{R_1}}{1 + \frac{1}{A} (1 + \frac{R_2}{R_1})}$  da cui sostituendo  $A(j\omega)$  e con  $A_0 \gg 1 + \frac{R_2}{R_1}$  diventa:  $\frac{V_o}{V_i} = \frac{\frac{R_2}{R_1}}{1 + \frac{j\omega}{\omega_t / [1 + \frac{R_2}{R_1}]}}$

**Configurazione invertente**

•  $\frac{V_o}{V_i} = \frac{1 + \frac{R_2}{R_1}}{1 + \frac{1}{A} (1 + \frac{R_2}{R_1})}$  da cui sostituendo  $A(j\omega)$  e con  $A_0 \gg 1 + \frac{R_2}{R_1}$  diventa:  $\frac{V_o}{V_i} = \frac{1 + \frac{R_2}{R_1}}{1 + \frac{j\omega}{\omega_t / [1 + \frac{R_2}{R_1}]}}$

**Configurazione non invertente**

## Generatori di segnale e generatori di forme d'onda

spesso si presenta la necessità di avere segnali con forme d'onda standard: per ottenerli si possono utilizzare due diversi approcci, cioè utilizzare un anello di controreazione positiva oppure modellando onde triangolari (fornite da appositi generatori). Nel primo caso si usa un amplificatore ed una rete selettiva in frequenza connessa ad un anello di controreazione positiva. Il guadagno ad anello chiuso in questo caso è

$$A_f(s) = \frac{A(s)}{1-A(s)\beta(s)} \text{ da cui il guadagno ad anello (ignorando il segnale negativo): } L(s) = A(s)\beta(s) \text{ da cui}$$

l'equazione caratteristica  $1-L(s)=0$ . Se questo quindi risulta essere pari a 1 ( $A\beta = 1$ ), si ha  $A_f = \infty$ , per cui il circuito presenta uscita finita in presenza di ingresso nullo (definizione di oscillatore), per cui per fornire oscillazioni sinusoidali deve essere  $L(j\omega_0) = A(j\omega_0)\beta(j\omega_0) = 1$  (oscillazioni di frequenza  $\omega_0$ ) si ha infatti (con  $X_s=0$ ):  $X_f = \beta x_0$ , con  $A X_f = x_0$  si ha  $A\beta X_f = x_0$  da cui  $A\beta = 1$ . Gli oscillatori non lineari fanno uso di una speciale classe di circuiti, detti multivibratori.

Il **multivibratore bistabile** ha due stati stabili e si muove da uno stato all'altro mediante opportuno impulso di eccitazione definito Trigger, motivo per cui il multivibratore bistabile è anche noto come trigger di Schmitt.

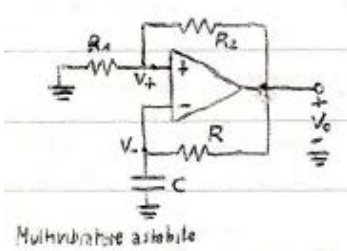
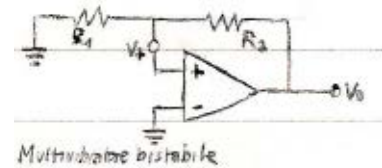
Si tratta di un amplificatore operazionale controreazionale positivamente. Supposto  $V_+$  inizialmente nullo, ad un suo incremento positivo corrisponde un aumento del segnale di uscita  $V_0$ : il partitore di tensione riporta sul terminale di ingresso positivo  $V_+$  una frazione  $\beta = \frac{R_1}{(R_1+R_2)}$  del segnale in uscita  $V_0$  che, se  $A\beta > 1$ , sarà maggiore dell'incremento originario di  $V_+$ ; questo processo rigenerativo continua fino a portare in saturazione positiva l'amplificatore operazionale al livello  $L^+$ , quindi  $V_+$  diventa  $\frac{L_+ R_1}{(R_1+R_2)}$ . Se invece di un incremento positivo  $V_+$  avesse ricevuto un incremento negativo, l'amplificatore operazionale sarebbe finito in saturazione negativa con  $V_0 = L_-$  e  $V_+ = \frac{L_- R_1}{(R_1+R_2)}$ .

Per far cambiare stato al circuito bistabile è possibile applicare una tensione di ingresso o sul terminale invertente o su quello non invertente. Nel primo caso, supponendo che  $V_0$  si trovi al livello  $L_+$  e  $V_+$  a  $\beta L_+$ , la tensione  $V_i$  non ha effetto fino a che non si supera il valore di  $V_+$ , cioè  $\beta L_+$  oltre il quale si forma una tensione negativa tra i terminali di ingresso, amplificata in uscita o riportata in ingresso tramite processo rigenerativo fino a che il circuito cambia completamente stato; analogamente se  $V_0 = L_-$  e  $V_+ = \beta L_-$  l'ingresso  $V_i$  non ha effetto fino a che non scende sotto il valore  $\beta L_-$  (tale segnale di ingresso può anche essere solo un impulso, per dare via al processo). Sia che per tensioni di ingresso  $\beta L_- < V_i < \beta L_+$  l'uscita può essere  $L_+$  o  $L_-$ , a seconda dello stato in cui si trova il circuito, quindi l'uscita si determina dal precedente segnale di ingresso (elemento di memoria). Nel secondo caso invece si ha  $V_+ = V_i \frac{R_2}{(R_1+R_2)} + V_0 \frac{R_1}{(R_1+R_2)}$  per cui se  $V_0 = L_+$  per far commutare il circuito  $V_i$  deve assumere valori negativi per far diventare  $V_+$  negativa: tale soglia si ottiene sostituendo  $V_0 = L_+$  e  $V_+ = 0$ , ricavando  $V_i = V_{TL} = -L_+ \frac{R_1}{R_2}$ . Analogamente se  $V_0 = L_-$  si ottiene  $V_{TH} = -L_- \frac{R_1}{R_2}$ . Questo tipo di circuito può essere utilizzato come un comparatore, che funziona in modo ottimale anche in

presenza di interferenze (grazie all'intervallo)  $V_{TL} < V_i < V_{TH}$ ).

Un **multivibratore astabile** si ottiene collegando un circuito RC in controreazione al multivibratore bistabile invertente. Supponendo che si trovi nello stato  $L_+$ , il condensatore si carica verso questo livello attraverso il resistore R, quindi la tensione ai capi di C cresce esponenzialmente verso  $L^+$  con costante di tempo  $\tau = RC$ , fino a raggiungere la soglia positiva  $V_{TH}$ : successivamente il multivibratore bistabile commuta stato, e ciò porta il condensatore a scaricarsi e la sua tensione  $V_-$  decrescerà esponenzialmente

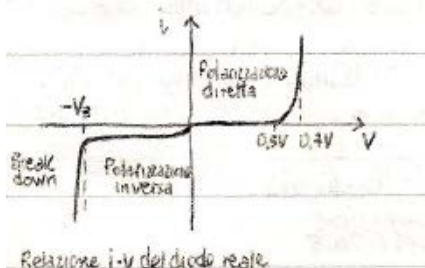
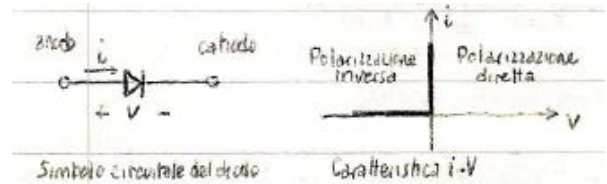
verso  $L_-$ , fino a raggiungere la soglia negativa  $V_{TL}$  (che porta ad una nuova commutazione di stato, e quindi ad un nuovo ciclo). In questo modo è possibile formare delle onde quadre (che variano al variare di C e R), ma è possibile avere anche delle onde triangolari, sostituendo il circuito passo-basso RC con un integratore che carica e scarica linearmente il condensatore.



Infine si ha il **multivibratore monostabile**, che ha uno stato stabile, nel quale può rimanere indefinitamente e uno stato quasi-stabile in cui si porta in seguito ad un impulso di Trigger, restandovi per un intervallo di tempo predeterminato uguale alla durata desiderata dell'impulso di uscita: alla fine il multivibratore monostabile ritorna nel suo stato stabile, in attesa di un altro segnale di Trigger.

## Diodi

molte operazioni di elaborazione dei segnali possono essere implementate solo attraverso circuiti non lineari: il più semplice e fondamentale elemento di un circuito non lineare è il **diodo** componente a due terminali come un resistore, che però presenta una relazione corrente – tensione (i-V) non lineare. Dal punto di vista ideale il comportamento del diodo è il seguente: se viene applicata una tensione negativa allora non scorre alcuna corrente ed il diodo si comporta come un circuito aperto (diodo **polarizzante inversamente**, viene detto **interdetto**), mentre se viene fatta scorrere corrente positiva ai suoi capi vi è una caduta di potenziale nulla, cioè il diodo si comporta come un cortocircuito (diodo **polarizzante direttamente**, viene detto **in conduzione**). Il circuito esterno deve essere progettato in modo da limitare a valori predeterminati la corrente che scorre e la tensione inversa. Il terminale positivo del diodo è detto **anodo**, quello negativo **catodo**, e presenta una funzione caratteristica **lineare a tratti**.



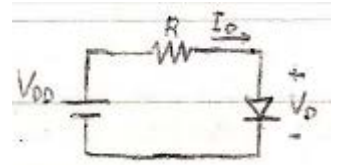
Dal punto di vista reale, la curva caratteristica del diodo è formata da tre diverse regioni; quella di polarizzazione diretta (per  $V > 0$ ), quella di polarizzazione inversa (per  $V < 0$ ) e la regione di **breakdown**, determinata per  $V < -V_z$ . Nella regione di polarizzazione diretta, quando la tensione ai capi del diodo è positiva, la caratteristica i-V si può approssimare con  $i = I_S(e^{\frac{v}{nV_T}} - 1)$ :  $I_S$  è una costante dipendente dai parametri costruttivi del diodo (e dell'area della giunzione) e dalla

temperatura ed è chiamata **corrente di saturazione inversa**; la tensione  $V_T$  è una costante detta **tensione termica** definita come  $V_T = \frac{kT}{q}$  (con  $K$  costante di Boltzmann,  $T$  temperatura assoluta,  $q$  valore assoluto della carica dell'elettrone); la costante  $n$  assume il valore 1 o il valore 2 in base al materiale e alla struttura fisica del diodo. La corrente per valori di tensione minori di circa 0.5V è trascurabile (detta **tensione di soglia**); successivamente subisce un rapido incremento, tanto che un diodo in piena conduzione ha la tensione ai suoi capi in campo ristretto, circa tra 0.6V e 0.8V, per cui viene assunto che qualunque sia la corrente in conduzione, ai capi del diodo vi è una tensione pari a 0.7V.

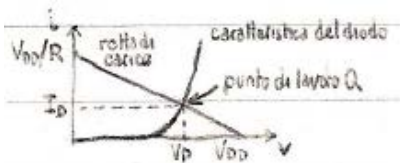
Si entra nella regione di polarizzazione inversa quando la tensione ai capi del diodo diventa negativa: quindi se  $V < 0$  ed è ampiezza maggiore di  $V_T$  il termine esponenziale diventa trascurabile per cui la corrente dei diodi in questo caso è  $i = -I_S$ , cioè la corrente in direzione inversa è una costante pari ad  $I_S$  (fortemente variabile con la temperatura). Infine si entra nella regione di breakdown se l'ampiezza della tensione inversa supera un valore di soglia tipico del particolare diodo, detto **tensione di breakdown**, indicata con  $V_z$  (dove  $Z$  sta per Zener). La corrente inversa cresce rapidamente mentre il corrispondente aumento della tensione ai capi del diodo rimane piuttosto contenuto (la potenza dissipata deve essere regolata dal circuito esterno).



Per analizzare il funzionamento del diodo nei circuiti si prenda in considerazione quello figura a lato formato da un generatore di tensione continua  $V_{DD}$  una resistenza  $R$  e un diodo: si vuole determinare la corrente  $I_D$  che scorre nel diodo e la tensione  $V_D$  ai suoi capi. Il diodo è polarizzato direttamente e assume



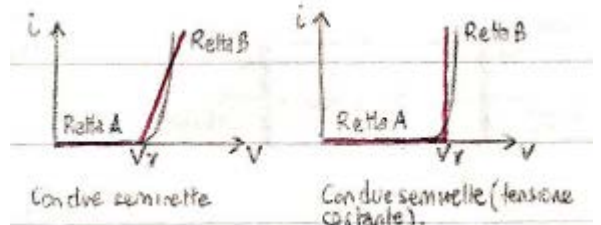
$V_{DD} \geq 0,5V$ . La corrente nel diodo è più grande di  $I_S$  e si ha:  $I_D \cong I_S(e^{\frac{V_D}{nV_T}})$ . L'equazione alla maglia che governa il circuito è:  $V_{DD} = RI_D + V_D$ , cioè  $I_D = \frac{V_{DD} - V_D}{R}$  per cui si hanno due equazioni nelle due incognite  $I_D$  e  $V_D$ ; per ottenere la soluzione si può usare l'analisi iterativa (usando iterativamente le equazioni  $I_D = \frac{V_{DD} - V_D}{R}$  e  $V_2 - V_1 = 2,3nV_T \log \frac{I_2}{I_1}$  fino ad ottenere risultati simili ai precedenti. La seconda equazione si ottiene a



partire da  $I_D \cong I_S(e^{\frac{V_D}{nV_T}})$  calcolata in due tensioni di  $V_D$ ,  $V_1$  e  $V_2$  e poi facendone il rapporto) oppure l'analisi grafica: quest'ultima si ricava riportando nel piano  $i-V$  le curve che descrivono le due equazioni che governano il circuito e individuandone l'intersezione nel punto  $q$ , detto **punto di lavoro**, cioè tra la caratteristica del diodo e la **retta di carico**.

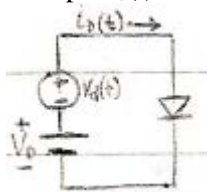
La natura non lineare del diodo comunque complica l'analisi dei circuiti che li contengono, perciò occorre determinare una relazione lineare che ne determini il comportamento. Un primo tentativo è quello di approssimare la curva esponenziale con due semirette, la linea A con pendenza nulla e la linea B con pendenza  $\frac{1}{r_D}$ ,

cioè: A:  $\begin{cases} i_D = 0 \\ V_D \leq V_Y \end{cases}$  B:  $\begin{cases} i_D = \frac{V_D - V_Y}{r_D} \\ V_D \geq V_Y \end{cases}$  oppure B:  $\begin{cases} i_D > 0 \\ V_D = V_Y + i_D r_D \end{cases}$



Il circuito corrispondente è un interruttore aperto per lo stato A, un generatore di tensione  $V_Y$  con resistenza  $r_D$  per lo stato B. Un modello ancora più semplice può essere

ottenuto mediante una semiretta verticale che approssimi la parte più rapidamente crescente della curva esponenziale, cioè il diodo mostra una caduta di potenziale costante  $V_Y$  (per lo stato B:  $i_D > 0$  e  $V_D = V_Y$ ): il circuito equivalente corrispondente è un interruttore per lo stato A e un generatore di tensione per lo stato B. In alcune applicazioni un diodo viene polarizzato mentre un segnale alternato viene sovrapposto alle quantità in continua: in questo caso la migliore rappresentazione del diodo è ottenuto mediante una resistenza uguale all'inverso della pendenza della tangente alla caratteristica nel punto di lavoro. Si prenda in considerazione un circuito in cui a una tensione continua  $V_D$  applicata al diodo viene sovrapposto un segnale variabile nel tempo  $v_d(t)$ . In assenza del segnale, la tensione ai capi del diodo è pari proprio a  $V_D$  e condurrà una corrente



continua  $I_D = I_S(e^{\frac{V_D}{nV_T}})$ ; quando viene applicato il segnale, ai capi del diodo si avrà  $v_D(t) = V_D + v_d(t)$  e la corrente corrispondente pari a:

$$i_D(t) = I_S(e^{\frac{V_D(t)}{nV_T}}) = I_S(e^{\frac{V_D + v_d(t)}{nV_T}}) = I_D + \frac{I_D}{nV_T} v_d(t) = I_D + i_d(t) \text{ con } i_d(t) = \frac{I_D}{nV_T} v_d(t), \text{ componente di}$$

segnale della corrente direttamente proporzionale alla tensione di segnale  $V_d(t)$ . Il rapporto tra  $i_d(t)$  e  $V_d(t)$  viene detta **conduttanza per piccoli segnali**, e il suo inverso è detto **resistenza per piccoli segnali**:  $r_d = \frac{nV_T}{I_D}$ . In particolare risulta che la pendenza della curva del diodo per  $i = I_D$  è pari a  $\frac{1}{r_d}$  cioè risulta

$$r_d = \frac{1}{\left[ \frac{\partial i_D}{\partial V_D} \right]_{i_D = I_D}}. \text{ La retta tangente si può esprimere come } i_D = \frac{1}{r_d} (V_D - V_{D0}), \text{ dove } V_{D0} \text{ è la sua intersezione}$$

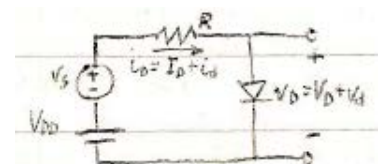
con l'asse  $x$  di  $V$ . A questa equazione corrisponde un circuito equivalente contenente un diodo, da cui si ricava:  $r_d i_D = V_D - V_{D0} \rightarrow V_D = r_d i_D + V_{D0} = (I_D + i_D) r_d + V_{D0} = V_D + i_d r_d$ . per illustrare l'applicazione del modello del diodo per piccoli segnali si consideri un circuito in cui è presente una tensione del segnale  $V_S$  in serie ad un generatore di tensione continua  $V_{DD}$  (in cui il diodo polarizzato direttamente, presenta ai suoi capi una tensione  $V_{D0}$  e una resistenza  $r_d$ ).

L'equazione alla maglia porta a  $V_{DD} + V_S = i_D R + V_{D0} + i_D r_d$  quindi

$$V_{DD} + V_S = (I_D + i_D) R + V_{D0} + (I_D + i_D) r_d = I_D R + (V_{D0} + I_D r_d) + i_d (R + r_d) \text{ con}$$

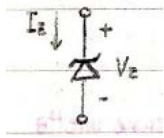
$$V_D = V_{D0} + I_D r_d. \text{ Da ciò è possibile separare le componenti in continua da}$$

$$\text{quelle di segnale, da cui: } V_{DD} = I_D R + V_D \text{ e } V_S = i_d (R + r_d).$$

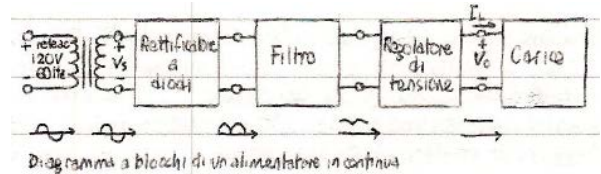




La curva i-v molto rapida nella regione di breakdown e la caduta di potenziale pressoché costante associata ad essa suggerisce che i diodi funzionanti nella regione di breakdown possano essere usati nel progetto di regolatori di tensione (circuiti che devono fornire una tensione costante e continua tra i suoi terminali di uscita): questo tipo di diodo viene chiamato **diodo zener**, in cui la corrente scorre entrando nel catodo, ed il catodo è positivo rispetto all'anodo, quindi  $I_Z$  e  $V_Z$  risultano avere valori positivi. In risposta ad una variazione  $\Delta I$  della corrente, la tensione dello zener cambia di valore  $\Delta V$  che dipende da  $\Delta I$  tramite la relazione  $\Delta V = r_z \Delta I$  dove  $r_z$  è l'inverso della pendenza della curva ed è detta **resistenza dinamica** dello zener (più è piccola minore è la pendenza della curva e il diodo è ideale). Il circuito equivalente può essere rappresentato da un generatore di tensione  $V_Z$  e resistenza in serie  $r_z$ .

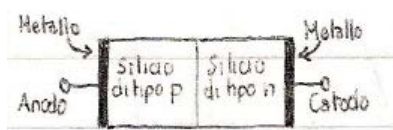


Uno dei più importanti campi di applicazione dei diodi è quello del progetto di circuiti raddrizzatori: il diodo rettificatore è un blocco costitutivo essenziale per gli alimentatori in continua che forniscono potenza agli apparati elettronici. Il primo blocco è il **trasformatore di potenza**, che con un opportuno cambio di spire  $\frac{N_2}{N_1}$  si



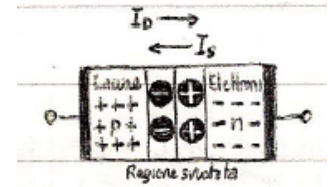
riduce la tensione di ingresso per ottenere il valore desiderato all'uscita dell'alimentatore: esso è collegato al rettificatore a diodi, che converte la sinusoide di ingresso  $v_s$  in un'uscita monopolare; quest'ultima avendo ancora natura pulsante e non essendo quindi adatta a fungere da sorgente, viene filtrata con un **filtro** che ne diminuisce le variazioni in ampiezza; infine per stabilizzare la tensione viene utilizzato un **regolatore di tensione**. Il rettificatore può essere a **mezz'onda**, che utilizza mezzo periodo della sinusoide d'ingresso (essendo costituite da un solenoide, viene preso in considerazione solo il semi-periodo che la polarizza direttamente, infatti quando è polarizzato inversamente non passa corrente), con l'accortezza che il segnale d'ingresso non abbia una tensione di picco che mandi in breakdown il diodo, oppure a **doppia semionda**, che utilizza l'intero periodo della sinusoide di ingresso invertendo l'onda nel semi-periodo in cui è negativa (utilizzando due diodi connessi a due ingressi uguali in modo che risultino alternativamente uno polarizzato direttamente e uno inversamente oppure utilizzando quattro diodi in una configurazione a ponte, però con un solo ingresso). Il filtro invece si ottiene collegando un condensatore in parallelo alla resistenza di carico  $R_L$ : esso si carica fino alla massima tensione del segnale d'ingresso e successivamente comincia a scaricarsi attraverso  $R_L$  fino a che la tensione del segnale di ingresso supera quella attuale del condensatore, quindi si ricarica nuovamente fino al picco massimo. Per evitare che la tensione d'uscita decresca troppo rapidamente si ricava una costante di tempo  $CR$  molto grande e tale che  $CR \gg T$ , dove  $T$  è il periodo della sinusoide di ingresso. Il regolatore di tensione infine si ottiene con un diodo zener in parallelo al carico.

Il diodo è fondamentalmente una **giunzione pn**: è formata da un materiale semiconduttore p a stretto contatto con un semiconduttore di tipo n (si tratta di due zone di differente drogaggio di uno stesso materiale



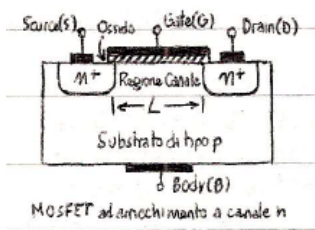
semiconduttore, ad esempio il silicio), alle cui estremità, cioè ai terminali del diodo, sono presenti due contatti metallici. Un cristallo di silicio puro ha una struttura regolare in cui gli atomi sono trattenuti in posizione da **legami covalenti** formati dai quattro elettroni di valenza propri degli atomi di silicio: per poter avere **elettroni liberi**, occorre fornirgli energia sufficiente per abbandonare l'atomo di origine, spezzando il legame covalente e in questo modo si riscontra nell'atomo stesso la presenza di una carica positiva uguale in modulo a quella dell'elettrone, lasciata scoperta detta **lacuna**. Sia gli elettroni liberi che le lacune si muovono casualmente lungo tutta la struttura del cristallo (le lacune vengono colmate da elettroni di altri atomi vicini, che a loro volta creano altre lacune, determinando lo spostamento): all'equilibrio termico la concentrazione di elettroni liberi  $n$  è uguale alla concentrazione di lacune  $p$ . Il movimento avviene tramite il meccanismo della diffusione e della deriva: la diffusione è associata al moto causale dovuto all'agitazione termica (della regione ad alta concentrazione a quella con minore concentrazione) che dà vita alla **corrente di diffusione**, mentre la deriva si verifica quando viene applicato un campo elettrico ai capi di un campione di silicio, dando vita alla **corrente di deriva**.

Il drogaggio di un cristallo di silicio per trasformarlo in tipo p (dove le lacune sono maggioritarie) oppure in tipo n (dove gli elettroni sono maggioritari) si ottiene introducendo in esso un piccolo numero di atomi di impurezze: rispettivamente riducendo atomi trivalenti  $N_A$  (che all'equilibrio termico portano ad avere un numero  $P_p \cong N_A$  di lacune e, poiché  $n_p P_p = n_{tot}^2$  un numero di elettroni  $n_p = \frac{n_{tot}^2}{P_p} = \frac{n_{tot}^2}{N_A}$ ) e atomi pentavalenti  $N_D$  (che all'equilibrio termico portano ad avere un numero  $n_n \cong N_D$  di elettroni e, poiché  $n_n P_n = n_{tot}^2$ , un numero di lacune  $p_n = \frac{n_{tot}^2}{N_D}$ ). In un diodo a giunzione con i terminali esterni non connessi, le lacune diffondono attraverso la giunzione del lato p verso il lato n e analogamente gli elettroni dal lato n verso il lato p: queste due componenti si sommano per formare la corrente di diffusione  $I_D$ , la cui direzione va da p ad n. Le lacune che diffondono verso la zona n si ricombinano con alcuni degli elettroni maggioritari presenti in quella zona ed escono di scena; ciò comporta la scomparsa di alcuni elettroni liberi dal materiale di tipo n, che lascia scoperta una corrispondente carica positiva disposta in una regione lungo la giunzione (dove avviene la ricombinazione), e analogamente succede per gli elettroni che diffondono verso la zona p. Si forma cioè una **regione svuotata dai portatori** da ambo le parti della giunzione, con il materiale n carico positivamente e quello p negativamente; le cariche ai lati della regione di svuotamento fanno in modo che si stabilisca un campo elettrico nella regione stessa che agisce come **barriera** per la corrente di diffusione  $I_D$ . È presente inoltre una corrente di deriva  $I_S$ : alcune delle lacune generate termicamente si muovono per diffusione fino al bordo della regione di svuotamento, dopodiché il campo elettrico formatosi le spazza via verso l'altra estremità nel lato p, e analogamente succede per gli elettroni che si formano nel materiale p; la somma di queste due componenti crea la corrente di deriva  $I_S$ , la cui direzione va dalla regione n alla regione p. Se i terminali esterni non sono connessi, non può esistere corrente esterna: quindi per mezzo della tensione di barriera, si ha  $I_D = I_S$ , e la tensione ai morsetti sarà nulla. L'effetto della **polarizzazione inversa** (attraverso ad esempio un generatore di corrente I, che va dal materiale p verso quello n, cui è generata quindi una tensione inversa) è quello di portare gli elettroni ad abbandonare il materiale p attraverso il circuito esterno: ciò provoca un aumento del numero delle cariche localizzate positive che rimangono scoperte nel materiale n e un aumento delle cariche localizzate negative nel materiale p, quindi un aumento della regione di svuotamento e quindi un aumento della barriera di potenziale, con la conseguente diminuzione della corrente di diffusione  $I_D$ , all'equilibrio si avrà una corrente inversa pari a  $I = I_S - I_D$ . Se si applica una corrente  $I > I_S$  (o una tensione V maggiore di  $V_Z$ ), la giunzione sfocia nella zona di breakdown: i meccanismi relativi possibili sono due, l'**effetto zener** (il campo elettrico aumenta al punto di poter rompere legami covalenti e generare coppie elettrone-lacune che vengono spazzati nelle zone opposte dal campo elettrico stesso, contribuendo alla corrente I) e l'**effetto valanga** (il campo elettrico agisce sui portatori minoritari della regione svuotata, fornendo loro energia sufficiente per spezzare i legami degli atomi che vanno ad urtare). La **polarizzazione diretta** invece produce con un processo analogo un assottigliamento della regione di svuotamento, quindi una riduzione della barriera di potenziale. Ciò permette ad un numero maggiore di lacune e di elettroni di attraversarla, portando ad un aumento della corrente di diffusione  $I_D$ , e all'equilibrio si avrà una corrente pari a  $I = I_D - I_S$ .



## Transistor

Il transistor ad effetto di campo (FET) si basa sul principio del campo elettrico per controllare un flusso di corrente: inoltre, poiché tale corrente è dovuta a un solo tipo di portatori (elettroni o lacune) è noto anche

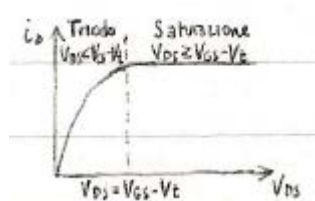
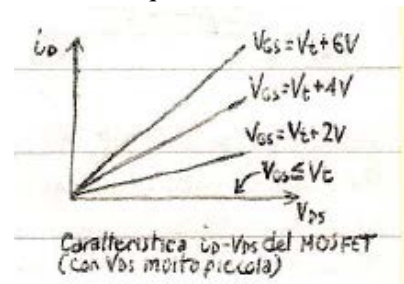


come **transistor unipolare**, in particolare viene utilizzato un transistor di tipo metallo-ossido-semiconduttore (MOS), e la sua configurazione più usata è quella ad arricchimento. In questo caso il transistor è realizzato su un substrato di tipo  $p$ , e in esso ricavate due regioni di tipo  $n$  fortemente drogate ( $n^+$ ), indicate come **source** e **drain**; sulla superficie viene fatto crescere uno strato di biossido di silicio, ottimo isolante elettrico, e su di esso viene posto del metallo per ottenere l'elettrodo di **gate** (analogamente per source, drain e substrato, detto body), che è

eletticamente isolato dal corpo del dispositivo. Il substrato forma giunzioni  $pn$  con le regioni di source e di drain, che normalmente sono polarizzate inversamente: la corrente che scorre tra queste due regioni può essere controllata dalla tensione applicata al gate (per la regione di canale sono importanti la lunghezza  $L$  e la larghezza  $W$ ) quando sul gate non viene applicata una tensione di polarizzazione, tra drain e source sono presenti due diodi in serie connessione dorso a dorso, che impediscono il passaggio di corrente anche quando viene applicata una tensione  $v_{DS}$ : se si applica una tensione positiva sul gate  $v_{GS}$ , applicata tra gate e source, con drain e source connessi a massa, le lacune libere nella regione sotto il gate vengono spinte verso il basso, lasciando quindi scoperte delle cariche negative; contemporaneamente vengono attirati nella medesima regione elettroni dalle zone  $n^+$ . In prossimità della superficie del substrato si viene a creare quindi una regione  $n$  che collega source e drain, quindi applicando una tensione  $v_{DS}$ , attraverso questa regione  $n$  potrà scorrere una corrente trasportata dagli elettroni: questa configurazione è detta **Mosfet a canale n** o **transistor Nmos**. Il valore di  $v_{GS}$  per cui nella regione di canale si accumula un numero di elettroni liberi sufficiente per formare il canale di conduzione è chiamato **tensione di soglia** (threshold)  $V_t$ : essa permette quindi la creazione di un campo elettrico verticale tra gate e substrato (che agiscono come maglie di un condensatore). Ora avendo indotto il canale, una tensione positiva  $v_{DS}$  molto piccola fa scorrere una corrente  $i_D$  in verso opposto agli elettroni che viaggiano dal source al drain, la sua intensità dipende dalla densità degli elettroni nel canale, che a sua volta dipende dalla tensione  $v_{GS}$  (in particolare se  $v_{GS}=V_t$ , il canale è appena creato e la corrente trascurabile, mentre per valori  $v_{GS}>V_t$  il canale diventa più spesso) e nello specifico la corrente  $i_D$  è proporzionale all'**eccesso** della tensione di gate:  $v_{GS}-V_t$  e anche alla tensione  $v_{DS}$ . In questo caso il mosfet sta funzionando come una resistenza lineare il cui valore è controllato

da  $v_{GS}$ : per  $v_{GS}\leq V_t$  la resistenza è infinita (la corrente è nulla) e

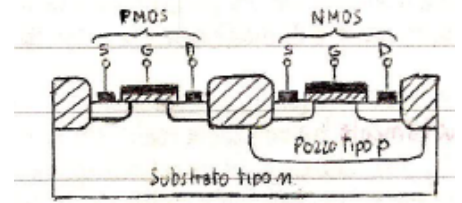
decresce non appena  $v_{GS}$  supera  $V_t$ , i valori per i quali il canale viene arricchito (per questo motivo è detto **Mosfet di arricchimento**); inoltre si ha  $I_S=I_D$ . Se si mantiene costante  $v_{GS}$  ad un valore superiore di  $V_t$  e se si aumenta  $v_{DS}$ , si osserva che  $v_{DS}$  si presenta come caduta di potenziale, cioè lungo il canale a partire dal source la tensione cresce da zero a  $v_{DS}$ , perciò la tensione tra gate e i vari punti del canale diminuisce dal valore  $v_{GS}$  (che assume all'estremità



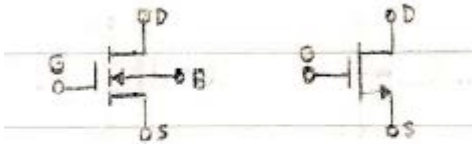
source) fino a  $v_{GS}-v_{DS}$  (che assume all'estremità drain); di conseguenza, poiché lo spessore del canale dipende da questa tensione, il canale non sarà più uniforme, ma avrà spessore maggiore in corrispondenza del source minore vicino al drain, e all'aumentare di  $v_{DS}$  il canale diventa sempre più appuntito (e la sua resistenza aumenta corrispondentemente, la relazione non è più lineare).

Se  $v_{DS}$  è aumentata fino ad avere  $v_{GS}-v_{DS}=V_t$  e quindi  $v_{DS}=v_{GS}-V_t$ , lo spessore del canale all'estremità di drain diminuisce fino quasi ad annullarsi e si dice che il canale è in **pinch-off** (strozzato): aumentandola ulteriormente,  $v_{DS}$  produce un effetto nullo sul canale, e la corrente si mantiene costante al valore raggiunto, e si dice che il mosfet entra nella **regione di saturazione** (da osservare che ad ogni valore di  $v_{GS}>V_t$  corrisponde un diverso valore di  $v_{DS}$  che porta il transistor in saturazione), mentre per valori  $v_{DS}<v_{GS}-V_t$  si è in **regione di triodo**.

Un **mosfet ad arricchimento a canale p** è realizzato su un substrato di tipo n con regioni  $p^+$  per il drain e il source: funziona nella stessa maniera di un componente a canale n tranne per il fatto che  $v_{GS}$  e  $v_{DS}$  sono negative e la tensione di soglia  $V_t$  è negativa. La tecnologia PMOS (soppiantata da quella NMOS più veloce ed economica) viene utilizzata nei circuiti a **MOS complementari** o **CMOS**, che impiega transistor MOS di entrambi i tipi: mentre il PMOS è implementato direttamente nel substrato di tipo n, il transistor NMOS è fabbricato in una regione p appositamente creata, chiamata pozzo P. È possibile anche un'altra configurazione con un substrato di tipo p e con il transistor di tipo p realizzato in un pozzo di tipo n.



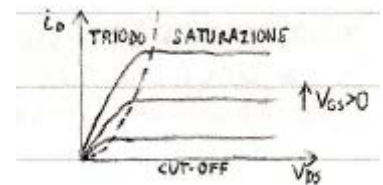
Il simbolo circuitale è piuttosto descrittivo: viene principalmente usato quello semplificato, in cui il substrato B è collegato al source S, l'elettrodo di gate G è staccato dalla linea



MOSFET ad arricchimento a canale n (B collegato a S)

del canale per rappresentare il fatto che è isolato dal corpo del componente, mentre la freccia indica la direzione normale del flusso di corrente, distingue il source dal drain e indica la polarità del dispositivo (in questo caso è a canale n). Le caratteristiche  $i_D$ - $v_{DS}$ , che sono una famiglia di curve per diversi valori di  $V_{GS}$ , indicano che esistono tre regioni di funzionamento: la regione di

interdizione (cut-off), la regione di triodo e la regione di saturazione. La regione di saturazione viene usata se il FET deve lavorare come un amplificatore, mentre le regioni di triodo e di interdizione sono utilizzate per farlo lavorare come interruttore. Il transistor è in cut-off quando  $v_{GS} < V_t$ ; per farlo lavorare nella regione di triodo occorre prima indurre un canale ( $v_{GS} \geq V_t$ ) e poi mantenere  $V_{DS}$  abbastanza piccola da non provocare la



strozzatura del canale, assicurando che la tensione tra gate e drain sia  $v_{GD} > V_t$  cioè sapendo che

$v_{GD} = v_{GS} = v_{SD} = v_{GS} - v_{DS}$ , deve risultare:  $v_{GS} - v_{DS} > V_t$  da cui  $v_{DS} < v_{GS} - V_t$ . Nella regione di triodo, le

caratteristiche  $i_D$ - $v_{DS}$  possono essere descritte approssimativamente dalla relazione  $i_D = K[2(v_{GS} - V_t) v_{DS} - v_{DS}^2]$

in cui  $K = \frac{1}{2} \mu_n C_{ox} \left(\frac{W}{L}\right)$  è un parametro proprio del componente, con  $\mu_n$  costante fisica detta **mobilità dell'elettrode**,  $C_{ox}$  capacità del condensatore tra il gate e il substrato  $L$  è la lunghezza del canale e  $W$  la

larghezza; siccome la quantità  $\frac{1}{2} \mu_n C_{ox}$  è una costante, il parametro  $K$  è determinato dal fattore  $\frac{W}{L}$ . Se  $v_{DS}$  è piccola tale da poter trascurare il termine  $v_{DS}^2$  le caratteristiche in prossimità dell'origine sono descritte da

$i_D \cong 2K(v_{GS} - V_t) v_{DS}$ , cioè da una relazione lineare attraverso una resistenza  $r_{DS} = \frac{v_{DS}}{i_D} = [2K(v_{GS} - V_t)]^{-1}$

controllata da  $v_{GS}$ . Per far lavorare il Mosfet nella regione di saturazione occorre innanzitutto indurre un

canale ( $v_{GS} \geq V_t$ ) che deve essere poi strozzato all'estremità di drain portando  $v_{DS}$  ad un valore tale che la

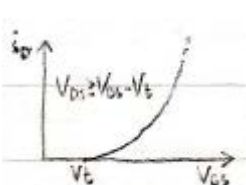
tensione tra il gate e il drain sia sotto  $V_t$ :  $v_{GD} \leq V_t$  cioè, scritto in termini di  $v_{DS}$ ,  $v_{DS} \geq v_{GS} - V_t$ . Il confine tra

regione di triodo e regione di saturazione è descritto da  $v_{DS} = v_{GS} - V_t$ , quindi sostituendo questo valore

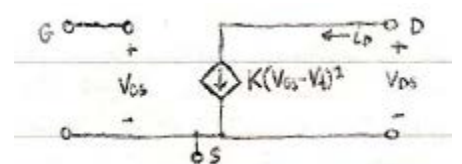
nell'equazione della curva caratteristica si ottiene il valore di saturazione della corrente  $i_D$ :  $i_D = K(v_{GS} - V_t)^2$

che in particolare è indipendente dalla tensione di drain  $v_{DS}$  ed è determinato dalla tensione di gate  $v_{GS}$

secondo una relazione quadratica. Un Mosfet in saturazione quindi si comporta come un generatore ideale di



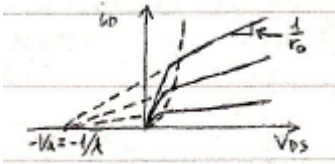
corrente il cui valore è controllato da  $v_{GS}$ . Va però considerato il fatto che la completa indipendenza di  $i_D$  da  $v_{DS}$  in saturazione è un'idealizzazione basata sulla premessa che, una volta strozzato il canale, ulteriori incrementi di  $v_{DS}$  non abbiano



effetti sul canale: in realtà aumentando  $v_{DS}$  il punto in cui si verifica la strozzatura si sposta verso il source, perciò la lunghezza del canale si riduce dando luogo ad una **modulazione della lunghezza del canale**.

Poiché  $K$  è inversamente proporzionale alla lunghezza  $L$ , al crescere di  $v_{DS}$  si verifica un aumento di  $K$  e quindi di  $i_D$ : questa piccola dipendenza lineare di  $i_D$  da  $v_{DS}$  si può esprimere mediante un fattore  $1 + \lambda v_{DS}$  tramite  $i_D = K(v_{GS} - V_t)^2(1 + \lambda v_{DS})$  dove la costante positiva  $\lambda$  è un parametro proprio del Mosfet.





Le caratteristiche  $i_D - v_{DS}$  in saturazione, che hanno andamento rettilineo.

Intrinsecano l'asse  $v_{DS}$  in un unico punto  $v_{DS} = \frac{-1}{\lambda} = -V_A$ , dove  $V_A$  è una tensione positiva caratteristica del transistor. La modulazione di lunghezza del canale fa sì che la resistenza d'uscita in saturazione assuma un valore finito,

dato da  $r_0 = \left[ \frac{\delta i_D}{\delta v_{DS}} \right]^{-1} = [\lambda K (v_{GS} - V_t)^2]^{-1}$  che può essere approssimato con  $r_0 \cong [\lambda I_D]^{-1}$  dove  $I_D$  è la

corrente corrispondente al particolare valore di  $v_{GS}$  per cui  $r_0$  è stato calcolato. Si può anche scrivere come  $r_0 \cong \frac{V_A}{I_D}$ ; nel circuito equivalente per grandi segnali (vedi sopra) occorre quindi porre una resistenza  $r_0$  in

parallelo al generatore di corrente. Per quanto riguarda il Mosfet a canale P, le considerazioni sono analoghe: per indurre un canale è necessario una tensione di gate più negativa di  $V_t$ , cioè  $v_{GS} \leq V_t$ ; ed una tensione di drain più negativa della tensione di source. Per ottenere il funzionamento nella regione di triodo si deve avere  $v_{DS} \geq v_{GS} - V_t$ ; la corrente  $i_D$  è data dalla stessa equazione valida per gli NMOS:

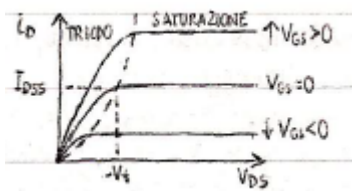
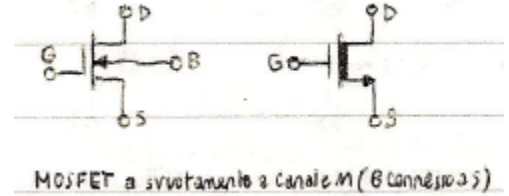
$i_D = K[(v_{GS} - V_t)^2 - v_{DS}^2]$  tenendo conto che  $v_{GS}$ ,  $v_t$  e  $v_{DS}$  sono negative e che  $K = \frac{1}{2} \mu_p C_{ox} \left( \frac{W}{L} \right)$ , dove  $\mu_p$ , la mobilità delle lacune nel canale indotto è  $\mu_p \cong \frac{1}{2} \mu_n$ . Per lavorare in saturazione invece deve essere

$v_{DS} \leq v_{GS} - V_t$  e la corrente  $i_D$  è data dalla stessa equazione valida per gli NMOS:

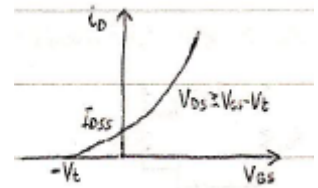
$i_D = K(v_{GS} - V_t)^2 (1 + \lambda v_{DS})$ . In molte applicazioni il substrato B è connesso all'elettrodo di source in modo da mantenere sempre in polarizzazione inversa la giunzione pn tra il substrato e il canale indotto; nei circuiti integrati tuttavia il substrato è comune a molti transistor MOS, e per mantenere questa giunzione il substrato viene connesso all'alimentazione più negativa in un circuito NMOS e a quella più positiva in un circuito PMOS, introducendo così una tensione di polarizzazione inversa tra il source e il substrato  $V_{SB}$  che influenza il funzionamento del dispositivo. Per esempio, ponendo il substrato NMOS ad una tensione negativa rispetto al source, la polarizzazione inversa fa allargare la zona di svuotamento riducendo quindi lo spessore del canale, e per far tornare la situazione come in precedenza occorre aumentare  $v_{GS}$ : in pratica  $V_{SB}$  porta ad un cambiamento della tensione di soglia, secondo la legge:  $V_t = V_{t0} + \gamma [\sqrt{2\theta_f + V_{SB}} - \sqrt{2\theta_r}]$ , (con  $V_{t0}$  tensione di soglia per  $V_{SB} = 0$ ,  $\gamma$  parametro del processo e  $\theta_f$  parametro fisico): da ciò si osserva che un aumento di  $V_{SB}$  porta ad un incremento di  $V_t$ , che a sua volta porta ad un incremento di  $i_D$ , con  $v_{GS}$  invariata. Il substrato si comporta quindi come un altro gate per il Mosfet, dando luogo al fenomeno di **body effect**. Va infine considerato che sia  $V_t$  che  $K$  sono sensibili alla temperatura (all'aumentare della temperatura l'effetto complessivo è di ridurre la corrente  $i_D$ ) e che a tensioni elevate possono verificarsi fenomeni di breakdown a valanga tra drain e substrato, di punch-through (perforazione) del canale nella zona di svuotamento e di breakdown dell'ossido (in questo caso si rompe il componente).



Il Mosfet a svuotamento ha una struttura simile a quella del Mosfet ad arricchimento, ma con un'importante differenza: Il **Mosfet a svuotamento** ha un canale realizzato fisicamente. In un transistor a canale n quindi è presente alla sommità del substrato p una regione di tipo n che connette le regioni n<sup>+</sup> di source e drain, perciò in presenza di una tensione  $v_{DS}$  la corrente  $i_D$  può scorrere anche se  $v_{GS}=0$ . Comunque la profondità e conducibilità del canale possono essere controllati da  $v_{GS}$ : applicando una tensione  $v_{GS}$  positiva vengono attratti altri elettroni nel canale, che risulta arricchito, mentre applicando una tensione  $v_{GS}$  negativa gli elettroni vengono allontanati dal canale, rendendolo più sottile e facendone diminuire la conducibilità (fino ad una tensione di soglia per la quale il canale è completamente svuotato e  $i_D$  è ridotta a zero anche in presenza di  $v_{DS}$ ). Le caratteristiche  $I_D-v_{DS}$  sono simili a quelle del componente ad arricchimento. Nel simbolo circuitale è presente una zona annerita per indicare la presenza di un canale permanente. Come anche nel Mosfet ad arricchimento è presente un'identica dipendenza di  $i_D$  da  $V_{DS}$ ; anche in questo caso inoltre, per funzionare nella regione di triodo deve valere:

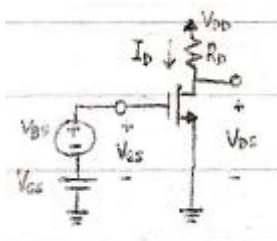


$v_{DS} \leq v_{GS} - V_t$ , e per funzionare nella regione di saturazione deve valere:  $v_{DS} \geq v_{GS} - V_t$ . Le caratteristiche corrente-tensione del Mosfet a svuotamento sono descritte dalle stesse equazioni per il componente ad arricchimento con la sola differenza che  $V_t$  è negativo (per la configurazione a canale n): un altro parametro che



descrive il mosfet a svuotamento è il calore della corrente di drain in saturazione per  $v_{GS} = 0$ , indicato con  $I_{DSS}$ . Il funzionamento dei transistor PMOS a svuotamento è del tutto analogo con la differenza che la polarità di tutte le tensioni (compresa  $V_t$ ) sono invertite.

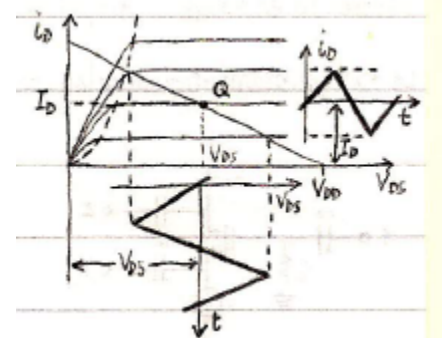
I transistor ad effetto di campo possono essere utilizzati come amplificatori. Si prenda un Mosfet a canale n ad arricchimento polarizzato idealmente da una batteria  $V_{GS}$ : a tale tensione continua tra gate e source è sovrapposto un segnale variabile nel tempo  $v_{gs}$  che è quello che si desidera amplificare. La tensione istantanea complessiva tra gate e source è quindi  $v_{GS} = V_{GS} + v_{gs}$ : in ogni istante il punto di lavoro si troverà sulla curva  $i_D - V_{DS}$  corrispondente al particolare valore di  $V_{GS}$  ed è determinato dal  $V_{DD}$  e  $R_D$  in base alla relazione  $V_{DS} = V_{DD} - R_D i_D$  che si può scrivere come  $i_D = \frac{V_{DD}}{R_D} - \frac{1}{R_D} V_{DS}$ . Questa è un'equazione



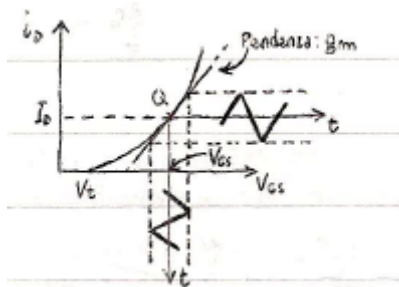
rappresentabile attraverso una retta sul piano  $i_D - V_{DS}$  che interseca l'asse delle ascisse in corrispondenza del valore  $V_{DD}$  ed ha una pendenza uguale a  $-\frac{1}{R_D}$  (poiché

$R_D$  è la resistenza di carico dell'amplificatore questa retta è detta **retta di carico**): il punto di lavoro istantaneo del Mosfet si troverà quindi all'intersezione tra questa retta e la curva  $i_D - V_{DS}$  relativa al valore istantaneo di  $V_{GS}$  con

coordinate  $i_D$  e  $V_{DS}$ . In particolare in assenza del segnale di ingresso  $V_{gs}$  il Mosfet si troverà nel punto contrassegnato con Q, che viene chiamato **punto di polarizzazione in continua o punto di riposo**: quando si applica l'onda triangolare  $V_{gs}$  il punto di lavoro istantaneo si muove lungo la retta di carico in corrispondenza della tensione istantanea complessiva  $V_{GS}$ . Tale andamento è rappresentato sia dalla forma d'onda risultante della corrente (alla corrente continua  $I_D$  è sovrapposta una componente  $i_D$  variabile nel tempo) che da quella della tensione (sovrapposta alla corrente in continua  $V_{DS}$  è presente una componente di segnale, a forma triangolare), che in particolare costituisce il segnale d'uscita e risulta una replica amplificata del segnale d'ingresso.



Scegliendo opportunamente il punto di lavoro Q e mantenendo piccola l'ampiezza del segnale d'ingresso è possibile quindi ottenere un amplificatore quasi lineare con un Mosfet: per far ciò il punto di lavoro è stato confinato nella regione di saturazione, dove il Mosfet lavora come generatore di corrente controllato da  $V_{GS}$ : se si permette al punto di lavoro di lasciare questa regione (ad esempio usando una resistenza  $R_D$  molto grande tale punto entra nella regione di triodo) si possono verificare notevoli distorsioni non lineari in uscita. Dal punto di vista algebrico, quando il segnale d'ingresso  $V_{gs}$  viene posto a zero, la corrente continua di drain  $I_D$  e la tensione continua di drain  $V_{DS}$  (o semplicemente  $V_D$  se il source è collegato a massa) sono date dalle relazioni:  $I_D = K(V_{GS} - V_t)^2$  e  $V_D = V_{DD} - R_D I_D$  (trascurando la modulazione di canale); quando il segnale viene sovrapposto a  $V_{GS}$  la tensione istantanea complessiva tra gate e source è pari a  $v_{gs} = V_{GS} + v_{gs}$  e



corrispondentemente la corrente istantanea complessiva è:

$$i_D = K(V_{GS} - V_t)^2 = K(V_{GS} + v_{gs} - V_t)^2 =$$

$= K(V_{GS} - V_t)^2 + 2K(V_{GS} - V_t)v_{gs} + Kv_{gs}^2$  dove il primo termine è la componente in continua della corrente, il secondo termine rappresenta una componente direttamente proporzionale al segnale di ingresso  $v_{gs}$ ,

mentre l'ultima componente, che è proporzionale al quadrato del segnale di ingresso, rappresenta un termine di distorsione non lineare, quindi non

desiderata; per ridurre la distorsione il segnale di ingresso deve soddisfare la condizione di piccolo segnale, cioè  $v_{gs} \ll 2(V_{GS} - V_t)$ , in questo modo è possibile trascurare l'ultimo termine ed esprimere  $i_D$  come

$i_D = I_D + i_d$  dove la corrente di segnale  $i_d$  è data da  $i_d = 2K(V_{GS} - V_t)v_{gs}$ . La costante che mette in relazione

$i_d$  con  $v_{gs}$  è la **transconduttanza  $g_m$** :  $g_m = \frac{i_d}{v_{gs}} = 2K(V_{GS} - V_t)$ . Si osserva che  $i_D = K(V_{GS} - V_t)^2$  quindi si

ha che  $g_m = \frac{\partial i_D}{\partial V_{GS}}|_{V_{GS}=V_{GS}}$  cioè è la pendenza della caratteristica  $i_D - V_{GS}$  nel punto di lavoro. Sostituendo il valore di K si ricava:

$g_m = (\mu_n C_{ox}) \left(\frac{W}{L}\right) (V_{GS} - V_t)$ : ciò indica che  $g_m$  dipende dal rapporto  $\frac{W}{L}$  del transistor, quindi per ottenere una transconduttanza grande il dispositivo deve essere corto e largo;  $g_m$  inoltre è proporzionale alla differenza di tensione  $\Delta V = V_{GS} - V_t$  ma se si aumenta  $V_{GS}$  per aumentare  $g_m$  diminuisce la dinamica del segnale

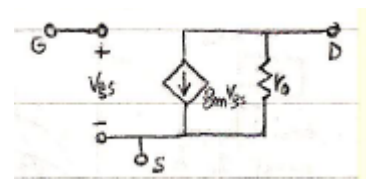
d'uscita. Un'altra espressione utile della transconduttanza può essere ottenuta sostituendo  $\sqrt{\frac{I_D}{K}}$  al posto di

$(V_{GS} - V_t)$  e poi sostituendo nuovamente K con  $\frac{1}{2}\mu_n C_{ox} \left(\frac{W}{L}\right)$  da cui risulta

$$g_m = \sqrt{2\mu_n C_{ox}} \sqrt{\frac{W}{L}} \sqrt{I_D}$$

. È possibile esprimere la tensione complessiva del circuito come  $V_D = V_{DD} - R_D I_D$ ; la componente di segnale della tensione di drain quindi è:  $V_d = -i_d R_D = -g_m R_D V_{gs}$  da cui si deduce che il

guadagno di tensione è dato da  $A_V = \frac{V_d}{V_{gs}} = -g_m R_D$ . Il segno negativo



indica che il segnale d'uscita  $V_d$  è sfasato di  $180^\circ$  rispetto al segnale d'ingresso  $V_{gs}$ . È importante notare

inoltre che, nell'ipotesi di piccoli segnali, le quantità di segnale ( $i_d$  e  $V_d$ ) si trovano sovrapposte a quelle continue ( $I_D$  e  $V_D$ ), per cui è possibile separare i calcoli in continua relativi alla polarizzazione da quelli sui segnali. il FET in pratica si comporta come un generatore di corrente controllato in tensione: riceve un

segnale  $V_{gs}$  tra gate e source e fornisce una corrente  $g_m V_{gs}$  sul terminale di drain (sia la resistenza di ingresso che quella di uscita sono molto alte e si possono assumere come infinite). Nell'analisi di un amplificatore

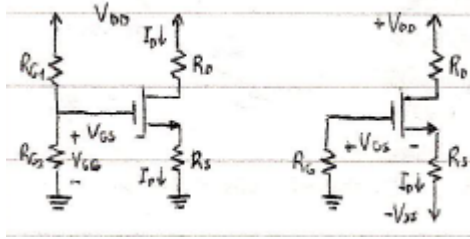
FET, il Fet può essere sostituito dal circuito equivalente, lasciando invariato il resto del circuito tranne per il fatto che i generatori di tensione continua vengono sostituiti con dei cortocircuiti: per includere l'effetto di

modulazione della lunghezza del canale si introduce una resistenza  $r_o = \frac{|V_A|}{I_D}$  in parallelo al generatore

controllato. È importante notare che i parametri  $g_m$  ed  $r_o$  dipendono dal punto di lavoro in continua del Fet

ed il guadagno diventa:  $\frac{V_d}{V_{gs}} = -g_m (R_D // r_o)$ .

Nel progetto di un amplificatore a FET occorre innanzitutto ricercare una polarizzazione prevedibile e stabile: gli schemi di polarizzazione che vengono usati sono due. Si può ottenere una polarizzazione



utilizzando la *controreazione ottenuta mediante resistenza di source*, sia per un circuito ad alimentazione singola, sia ad alimentazione doppia. Nel primo caso, con il partitore di tensione si ha il gate alimentato dalla tensione costante:  $V_{GG} = V_{DD} \left( \frac{R_{G2}}{R_{G1} + R_{G2}} \right)$  che, in assenza di  $R_S$  viene applicata direttamente tra gate e source, e la corrispondente corrente  $I_D$  dipenderà fortemente da  $V_{GG}$ ,  $K$  e

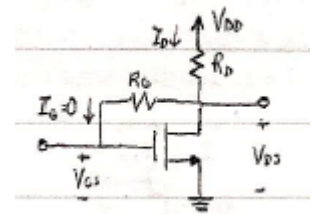
$V_t$ ; inserendo  $R_S$  si ottiene:  $V_{GG} = V_{GS} + I_D R_S$ , cioè  $I_D = \frac{V_{GG}}{R_S} - \frac{1}{R_S} V_{GS}$ . Nel secondo caso il funzionamento

è lo stesso, a patto di sostituire  $V_{GG}$  con  $V_{SS}$ , per cui vale l'equazione  $I_D = \frac{V_{SS}}{R_S} - \frac{1}{R_S} V_{GS}$ . In entrambi i casi è assunto che il valore di  $R_D$  fosse stato scelto in modo tale che il dispositivo funzionasse in saturazione:

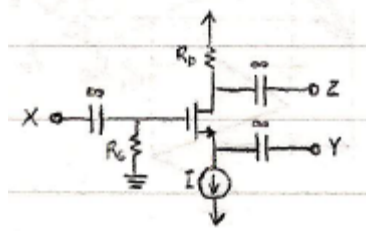
$I_D = K(V_{GS} - V_t)^2$ . Si può ottenere una polarizzazione utilizzando la *controreazione tra gate e drain*: cioè viene connesso tra gate e drain una resistenza  $R_G$ , usualmente di valore elevato, per cui essendo  $I_G = 0$ , la tensione continua di gate sarà uguale alla tensione continua di drain  $V_D = V_G$ . Dal

circuito a lato si ricava che  $v_{DS} = V_{DD} - i_D R_D$ , cioè  $i_D = \frac{V_{DD}}{R_D} - \frac{1}{R_D} v_{DS}$  e,

risolvendo questa equazione insieme a quella che descrive il luogo dei punti  $v_{DS} = v_{GS}$ , cioè  $i_D = K(V_{DS} - V_t)^2$ , si ricavano i valori di  $i_D$  e  $V_{DS}$  nel punto di lavoro. In entrambi i casi gli aumenti indesiderati di  $I_D$  sono quindi bilanciati dalle cadute di tensione su  $R_S$  e  $R_G$ .



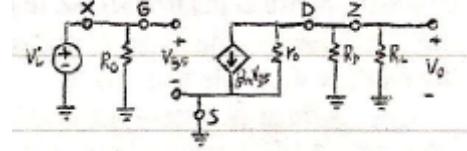
Le configurazioni fondamentali con cui può essere realizzato un amplificatore usando FET sono tre, e



ognuna offre vantaggi peculiari: in tutti e tre i casi comunque viene utilizzato lo stesso schema di polarizzazione, cioè un generatore di corrente continua connesso all'alimentazione continua, un resistore  $R_G$  che connette il gate a massa, fissando così  $V_0$  a zero, una resistenza  $R_D$  che collega il drain all'alimentazione positiva, che assicura che il transistor funzioni sempre nella regione di saturazione. La configurazione **source comune** si ottiene connettendo l'elettrodo

Y a massa, il segnale d'ingresso viene connesso al gate e la resistenza di carico al drain (nel circuito equivalente il generatore di corrente di polarizzazione è sostituito da un circuito aperto): in questo caso si ha  $R_{IN} = R_G$ ,  $R_{OUT} = R_D // r_0$  e il guadagno pari a

$$A_V = \frac{V_0}{V_i} = -g_m (R_L // R_D // r_0) \text{ o a circuito aperto senza tener conto di } R_L \text{ } A_{V0} = -g_m (R_D // r_0); \text{ il}$$



guadagno quindi si può ottenere mediante il partitore di tensione:

$$A_V = A_{V0} \left( \frac{R_L}{R_L + R_{out}} \right). \text{ La configurazione a } \textbf{gate comune} \text{ si ottiene}$$

connettendo l'elettrodo X a massa, il segnale di ingresso viene connesso al source e la resistenza di carico al drain; trascurando momentaneamente  $r_0$  si osserva che  $V_{gs} = -V_i$ , dunque dal

generatore di segnale viene assorbita una corrente  $gmV_i$  quindi la

resistenza di ingresso è semplicemente  $\frac{1}{gm}$ ; osservando che la corrente di drain è  $gmV_{gs} = -gmV_i$  e quindi

$V_0 = gmV_i(R_L // R_D)$ , il guadagno di tensione è  $A_V = \frac{V_0}{V_i} = gm(R_L // R_D)$ , invece tenendo conto della resistenza

$r_0$  si ha  $R_{IN} = \frac{1}{gm}$  e  $R_{OUT} = R_D // r_0$  e il guadagno pari a

$A_V = gm(R_L // R_D // r_0)$  che si differenzia dalla configurazione precedente dal fatto che non c'è inversione di segno. Infine la

configurazione a **drain comune** si ottiene connettendo l'elettrodo Z a massa, il segnale di ingresso viene connesso al gate e la resistenza di carico al source: la resistenza di ingresso è  $R_{IN} = R_G$ , mentre la

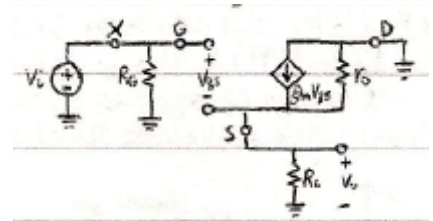
resistenza d'uscita può essere trovata cortocircuitando la sorgente di segnale  $V_i$  e applicando la tensione di

prova  $V_Y$  al source; in questo caso  $V_{gs} = -V_Y$ ,  $i_Y = -gmV_{gs} + \frac{V_Y}{r_0} = gmV_Y + \frac{V_Y}{r_0}$  quindi si ha  $R_{OUT} = \frac{1}{gm} // r_0$

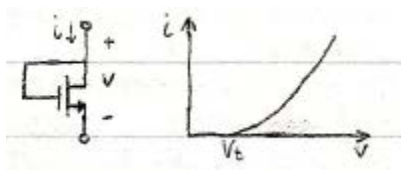
(se  $r_0 \gg \frac{1}{gm}$  allora  $R_{OUT} \cong \frac{1}{gm}$ ). Per quanto riguarda il guadagno di tensione senza contare  $R_L$  si ha

$V_0 = V_S = gmV_{gs}r_0$  e sapendo che  $V_i = V_{gs} + V_0$  si ricava  $V_i = \frac{V_0}{gmr_0} + V_0$  quindi il guadagno ad anello aperto è

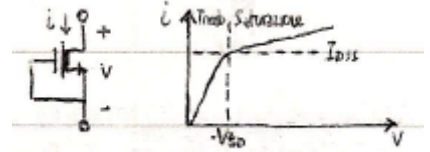
$$A_{V0} = \frac{1}{1 + \left( \frac{1}{gmr_0} \right)} \text{ mentre il guadagno di tensione è } A_V = A_{V0} \frac{R_L}{R_L + R_{out}}.$$



Gli amplificatori integrati in tecnologia Mos utilizzano altri transistor Mos come elementi di carico, al posto dei resistori: vengono impiegate due diverse tecnologie, cioè gli NMos (ad arricchimento e svuotamento) e i CMos (ad arricchimento). Nel primo caso vengono utilizzati due elementi di carico, cioè MosFet ad

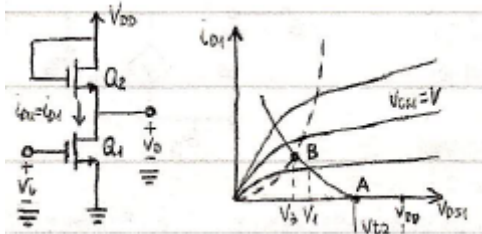


arricchimento con il drain collegato al gate ed il MosFet a svuotamento con il gate collegato al source; il primo transistor è caratterizzato da una caratteristica  $i-v$ :  $i = K(v - V_t)^2$  ed è

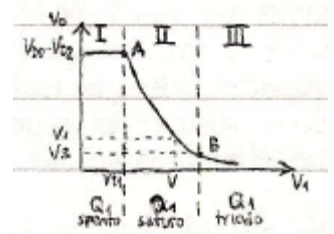


sempre in saturazione (con resistenza  $\frac{1}{gm}$ ), mentre il secondo per essere in saturazione, la tensione tra i due terminali deve superare  $-V_{tD}$  tensione di soglia; poiché nella regione di triodo è  $i = k(-2V_{tD}v - v^2)$ , per  $v = V_{tD}$  si ha  $i = KV_{tD}^2 = I_{DSS}$ . Tenendo conto della modulazione di canale invece si ha  $i \cong KV_{tD}^2(1 + \frac{v}{V_A})$ . Per quanto riguarda l'amplificatore **NMOS con carico ad arricchimento** occorre notare che

la corrente  $i_{D1}$  è la stessa che scorre nel componente di carico  $Q_2$ , che  $V_{DS1} = V_0$  e che  $V_{GS1} = V_i$ . Alle caratteristiche statiche di  $Q_1$  viene sovrapposta la curva di carico così la caratteristica di trasferimento risulta data dai punti di intersezione della curva di carico con le curve caratteristiche  $i_{D1}-V_{DS1}$ ; la curva risultante mostra tre regioni ben definite. Nella regione I il transistor pilota  $Q_1$  è in interdizione, poiché  $V_i < V_{t1}$  mentre  $Q_2$  è in saturazione, quindi la tensione ai suoi capi è  $V_{t2}$  e la tensione d'uscita è  $V_{DD} - V_{t2}$ , nella regione II  $Q_1$



è in saturazione e la curva risulta lineare; infine nella regione III  $Q_1$  entra nella regione di triodo. Nell'ipotesi che in saturazione entrambi i dispositivi abbiano resistenza infinita (cioè curve caratteristiche orizzontali) e che

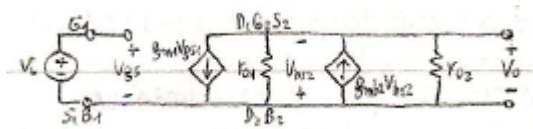


abbiano stessa tensione di soglia  $V_t$ , si ha che, quando  $Q_1$  è in saturazione:  $i_{D1} = K_1(V_{GS1} - V_t)^2$  e poiché  $i_{D1} = i_{D2} = i_D$  e  $V_{GS1} = V_i$ , l'equazione si può scrivere come:  $i_D = K_1(V_i - V_t)^2$ . Il funzionamento di  $Q_2$  è descritto da  $i_D = K_2(V_{GS2} - V_t)^2$  che, poiché  $V_{GS2} = V_{DD} - V_0$ , si può riscrivere come  $i_D = K_2(V_{DD} - V_0 - V_t)^2$ . Combinando queste due equazioni si ottiene:

$V_0 = \left( V_{DD} - V_t + \sqrt{\frac{K_1}{K_2}} V_t \right) - \sqrt{\frac{K_1}{K_2}} V_i$  che è la relazione tra  $V_i$  e  $V_0$  relativa alla regione II; il guadagno

dell'amplificatore è:  $A_V = -\sqrt{\frac{K_1}{K_2}} = -\sqrt{\frac{(W/L)_1}{(W/L)_2}}$  quindi è fissato una volta scelte le geometrie dei due

componenti, e per avere un guadagno elevato  $Q_1$  deve essere corto e largo e  $Q_2$  lungo e stretto. Dal punto di vista circuitale, poiché la caduta di tensione ai capi del generatore di corrente  $gm_2 V_{GS2}$  è  $V_{GS2}$ , questo può



essere sostituito da una resistenza  $\frac{1}{gm_2}$  quindi si ha la tensione

d'uscita:  $V_0 = -gm_1 V_{GS1} \left[ \left( \frac{1}{gm_2} \right) // r_{o1} // r_{o2} \right]$ ; poiché

$V_{GS1} = V_i$  il guadagno è  $A_V = \frac{V_0}{V_i} = \frac{-gm_1}{gm_2 + \frac{1}{r_{o1}} + \frac{1}{r_{o2}}}$  e se

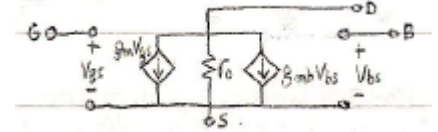
$r_{o1}$  e  $r_{o2}$  sono molto maggiori di  $\frac{1}{gm_2}$  allora si ha  $A_V = -\frac{gm_1}{gm_2}$  che riconduce a  $A_V = -\sqrt{\frac{K_1}{K_2}}$ .



Occorre tener presente però che  $Q_1$  e  $Q_2$  condividono lo stesso substrato, normalmente collegato

all'alimentazione più negativa, quindi per  $Q_2$  il substrato è al potenziale di massa mentre il source non lo è, e risente dunque dell'effetto body poiché il substrato non è collegato al source e si presenta una tensione tra body e source  $V_{BS}$  che dà luogo ad una componente della corrente di drain indicata come  $g_{mb} V_{BS}$ . Poiché  $i_D$  dipende da  $V_{BS}$  attraverso la dipendenza

di  $V_t$  da  $V_{BS}$ , si ricava che  $g_{mb} = \chi g_m$ . A lato è presente il circuito equivalente di un Mosfet in cui il substrato non è collegato al source.



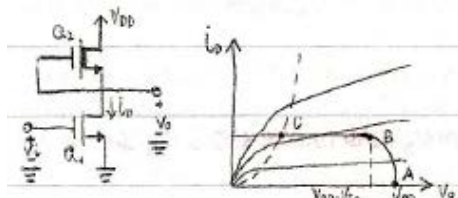
Prendendo quindi in considerazione l'effetto body per l'amplificatore con carico ad arricchimento occorre tener conto di un generatore controllato  $g_{mb2} V_{BS2}$  che può essere sostituito da una resistenza  $\frac{1}{g_{mb2}}$  poiché ai suoi

capi vi è una tensione  $V_{BS2}$  quindi si ha:  $V_0 = -g_{m1} V_{GS1} \left[ \left( \frac{1}{g_{m2}} \right) // \left( \frac{1}{g_{mb2}} \right) // r_{o1} // r_{o2} \right]$  da cui essendo

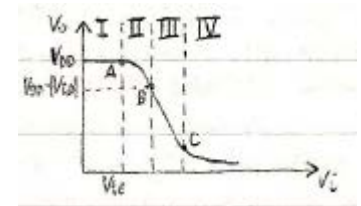
$$V_{GS1} = V_i: A_V = \frac{-g_{m1}}{(g_{m2} + g_{mb2} + \frac{1}{r_{o1}} + \frac{1}{r_{o2}})}$$

Assumendo che  $r_{o1}$  e  $r_{o2}$  siano più grandi rispetto a  $\frac{1}{g_{m2}}$  si può approssimare con  $A_V = \frac{-g_{m1}}{(g_{m2} + g_{mb2})} = \frac{-g_{m1}}{g_{m2}} \frac{1}{1 + \chi}$

cioè il guadagno è ridotto. Un amplificatore **Nmos con carico a svuotamento** ha caratteristiche superiori



rispetto a quello con carico ad arricchimento. Dalla caratteristica risultante (che si può individuare come nel caso del carico ad arricchimento) si possono individuare 4 regioni diverse: nella regione I, per  $V_i < V_{tE}$  dove  $V_{tE}$  è la soglia del transistor  $Q_1$  il transistor  $Q_1$  è interdetto mentre  $Q_2$  è in triodo,

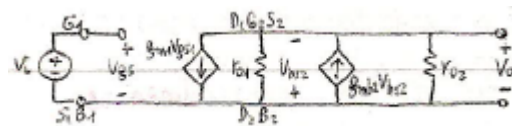


finché  $V_0 < V_{DD} - |V_{tD}|$  dopo tale valore si entra nella regione III dove

entrambi i transistor sono in saturazione e hanno resistenza di uscita elevata, che dà luogo a guadagno elevato; infine nella regione IV  $Q_1$  è in triodo mentre

$Q_2$  è in saturazione, quando  $V_0$  diventa minore di  $V_i$  di una quantità  $V_{tE}$ . Se l'amplificatore viene polarizzato in modo da funzionare nella regione III, allora il guadagno è pari a:  $A_V = \frac{V_0}{V_i} = -g_{m1} [r_{o1} // r_{o2}]$ . Nella

pratica però, tenendo conto dell'effetto body nel transistor  $Q_2$ , si ottiene un guadagno più basso: tra il



substrato ed il source di  $Q_2$  compare un segnale di tensione

pari a  $-V_0$ . Nel circuito equivalente, il generatore di corrente

$g_{mb2} V_{BS2}$  può essere sostituito da una resistenza  $\frac{1}{g_{mb2}}$  quindi

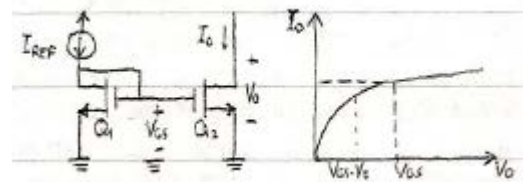
si ha la tensione d'uscita  $V_0 = -g_{m1} V_{GS1} \left[ \frac{1}{g_{mb2}} // r_{o1} // r_{o2} \right]$  e poiché  $V_{GS1} = V_i$  il guadagno:

$A_V = \frac{V_0}{V_i} = -g_{m1} \left[ \frac{1}{g_{mb2}} // r_{o1} // r_{o2} \right]$ . Poiché in generale  $\frac{1}{g_{mb2}}$  è più piccola di  $r_{o1}$  e  $r_{o2}$ , si ha

$A_V = \frac{-g_{m1}}{g_{mb2}} = \frac{-g_{m1}}{g_{mb2} \chi}$  cioè  $A_V = - \sqrt{\left( \frac{W}{L} \right)_1} * \frac{1}{\chi}$  quindi il guadagno è maggiore rispetto a quello con carico

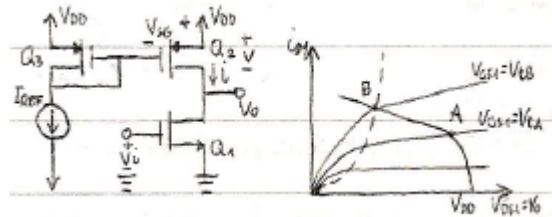
ad arricchimento di un fattore  $\frac{1 + \chi}{\chi}$ .

Sia nei circuiti Nmos che Cmos viene prodotta una corrente continua di riferimento stabile e prevedibile, utilizzata per ottenere correnti continue ad essa proporzionali con cui polarizzare i vari transistor: l'elemento che fornisce tali correnti continue è lo **specchio di corrente**, Esso è formato da due Mosfet ad arricchimento  $Q_1$  e  $Q_2$  con medesima tensione di soglia  $V_t$ :



$Q_1$  è pilotato dalla corrente di riferimento  $I_{REF}$ , mentre la corrente d'uscita  $I_0$  è presa sul drain di  $Q_2$ , che deve essere in saturazione. Per  $Q_1$  si ha  $I_{REF} = K_1(V_{GS} - V_t)^2$  e poiché  $Q_2$  è in parallelo a  $Q_1$  essi hanno la stessa tensione  $V_{GS}$ , quindi:  $I_0 = K_2(V_{GS} - V_t)^2$ , trascurando la resistenza d'uscita; combinando le equazioni si ha  $I_0 = I_{REF} \frac{K_2}{K_1} = \left(\frac{W}{L}\right)_2 \frac{K_2}{K_1}$ , quindi idealmente  $I_0$  è multiplo di  $I_{REF}$  con un valore determinato dalla geometria

dei componenti (in pratica però solo se la tensione sul drain di  $Q_2$  è uguale a  $V_{GS}$ ). Questa configurazione



viene usata per gli amplificatori **Cmos** (disponibili sia a canale  $n$  sia a canale  $p$ , realizzati in modo da eliminare l'effetto body): in questo circuito  $Q_1$  e  $Q_2$  sono componenti a canale  $p$  collegati a specchio di corrente e pilotati dalla corrente  $I_{REF}$ , quindi  $Q_2$  si comporta come generatore di corrente; funziona nella regione di saturazione se la tensione

sul suo drain risulta più bassa di quella del suo source ( $V_{DD}$ ) di una quantità

almeno pari a  $V_{SG} - |V_{tp}|$  e, una volta in saturazione  $Q_2$  ha un'elevata

resistenza d'uscita  $r_{02} = \frac{|V_1|}{I_{REF}}$  ( $Q_2$  è il carico attivo del transistor  $Q_1$ ). Dalla

caratteristica di trasferimento risultante si osservano quattro distinte regioni di

funzionamento: nella regione I  $Q_1$  è interdetto, nella regione II  $Q_1$  è in

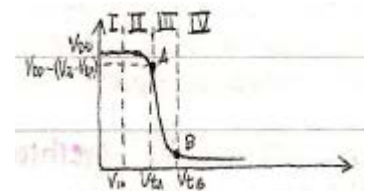
saturazione e  $Q_2$  in triodo, nella regione III sia  $Q_1$  che  $Q_2$  sono in saturazione mentre nella regione IV  $Q_1$  è in

triode e  $Q_2$  in saturazione. Nella regione III, che è quella di interesse, il guadagno è

$A_V = -\frac{V_0}{V_i} = -g_{m1}[r_{01}/r_{02}]$  e, poiché  $Q_1$  funziona con una corrente continua in polarizzazione uguale a

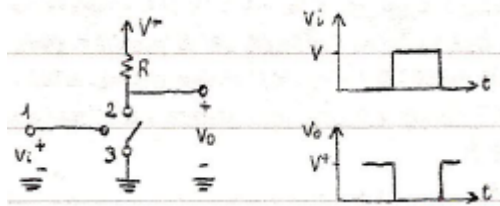
$I_{REF}$ , si ha:  $g_{m1} = \sqrt{(2\mu_n C_{ox}) \left(\frac{W}{L}\right)_1 I_{REF}}$ .

Sapendo inoltre che  $r_{01} = r_{02} = \frac{|V_1|}{I_{REF}}$  si ottiene  $A_V = -\frac{\sqrt{K_n} |V_1|}{\sqrt{I_{REF}}}$ .



## Circuiti Digitali

In un circuito digitale i segnali assumono un numero limitato di valori, e quelli più comuni utilizzano due soli valori: il livello “alto” e il livello “basso”, indicati con i simboli 1 e 0: in generale però, per tener conto di tutti gli effetti che possono cambiare i livelli della tensione di segnale, alle variabili binarie vengono assegnati due intervalli di tensione distinti (compreso tra  $V_{L1}$  e  $V_{L2}$  per lo 0 logico e  $V_{H1}$  e  $V_{H2}$  per l'1 logico), separati da un intervallo di indeterminazione in cui si suppone che la tensione non possa mai cadere. I circuiti integrati sono classificati in diverse famiglie (principalmente la Nmos e la Cmos) e la scelta di una famiglia logica si basa su diverse considerazioni: flessibilità logica, velocità di funzionamento, disponibilità di funzioni complesse, immunità al rumore, intervallo di temperature di funzionamento, potenza dissipata e costo. Secondo la loro complessità si ricavano quattro tipi di circuiti integrati: SSI, MSI, LSI, VLSI. Il componente base che tratta i segnali H e L è l'**invertitore logico**, che è un interruttore controllato in tensione:



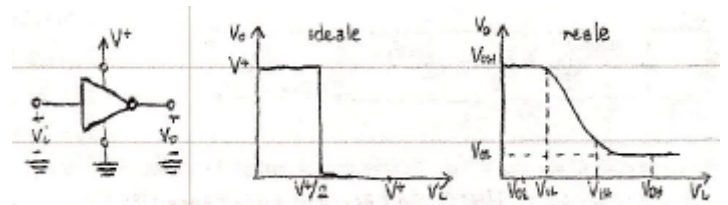
l'interruttore collegato tra i punti 2 e 3 è controllato dal segnale di ingresso applicato tra i punti 1 e 3 (il terminale 3 è connesso a massa), mentre la tensione d'uscita è prelevata tra il terminale 2 e massa. Quando  $V_i$  è bassa l'interruttore è aperto e  $V_0$  è alta (pari a  $V^+$ ) mentre se  $V_i$  è alta l'interruttore è chiuso e l'uscita è bassa (implementa l'operazione NOT di inversione). Dal punto

di vista reale però occorre tener presente vari aspetti: l'elettrodo di ingresso assorbe un po' di corrente dal generatore che lo pilota; l'interruttore non è ideale, ma quando è chiuso non si comporta come un cortocircuito ma presenta una resistenza di chiusura finita e, a volte, una caduta di tensione aggiuntiva, quindi  $V_0$  non sarà esattamente 0; l'interruttore non commuta istintivamente, nel senso che gli invertitori reali non mostrano una soglia di commutazione ben definita. In una caratteristica di trasferimento ideale, la

tensione di soglia è  $V_{th} = \frac{1}{2} V^+$ ; al di sotto di

essa i segnali sono interpretati come bassi e l'uscita è  $V^+$ , al di sopra sono considerati alti e l'uscita è nulla. Nella caratteristica reale la

tensione di soglia non è più ben definita ed esiste



una regione di transizione tra gli stati alto e basso, mentre i livelli di uscita alto ( $V_{0H}$ ) e basso ( $V_{0L}$ ) non sono pari a  $V^+$  e 0V. Le tre regioni distinte sono quindi: ingresso basso ( $V_i < V_{1L}$ ), nella regione di transizione ( $V_{1L} < V_i < V_{1H}$ ), in ingresso alto ( $V_i > V_{1H}$ ): le tensioni minori di  $V_{1L}$  sono interpretate come 0 logico (quindi  $V_{1L}$  è il massimo valore ammesso per lo 0 logico) e analogamente tensioni d'ingresso maggiori di  $V_{1H}$  sono interpretate come 1 logico (quindi  $V_{1H}$  è il minimo valore ammesso per l'1 logico).

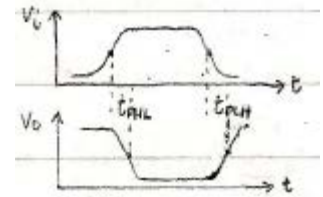
Un grosso vantaggio dei circuiti digitali binari è la tolleranza alle variazioni del segnale d'ingresso, cioè un'immunità al rumore (segnali indesiderati): la differenza  $NM_H = V_{0H} - V_{1H}$  è definita **margine di rumore all'1 logico** (cioè l'uscita di una porta non viene modificata dal rumore sovrapposto fin quando l'ampiezza della tensione di rumore non supera la soglia  $V_{0H} - V_{1H}$ ) e la differenza  $NM_L = V_{0L} - V_{1L}$  è definita **margine di rumore allo 0 logico** (con  $V_{0H}$  minima tensione d'uscita per il valore alto,  $V_{1H}$  minima tensione d'ingresso riconosciuta come 1 logico,  $V_{0L}$  massima tensione d'uscita a livello basso e  $V_{1L}$  massima tensione d'ingresso riconosciuta come 0 logico); per massimizzare i margini di rumore dovrebbe essere  $V_{1L} = V_{1H}$  (un valore al centro dell'intervallo  $V_{0L} - V_{0H}$ ) quindi che la regione di transizione sia caratterizzata da un guadagno elevato, e la commutazione dovrebbe avvenire al centro dell'esecuzione logica, quindi con  $V_{0L}$  più piccola possibile e  $V_{0H}$  più elevata possibile (fino a  $V^+$ ).

Un parametro di interesse nella progettazione dei circuiti logici è la **potenza dissipata**, che permette di stabilire la corrente che l'alimentatore deve essere in grado di fornire al circuito; essa è formata da due componenti. La potenza **statica** è quella dissipata quando il circuito non cambia: se  $V_0 = 1$  allora  $I_H = 0$  e quindi la potenza dissipata è nulla, mentre se  $V_0 = 0$  allora  $I_L = \frac{V^+}{R}$  e quindi la potenza dissipata è  $\left(\frac{V^+}{R}\right) V^+$ ; se si assume che mediamente la porta si trovi per metà del tempo in ciascuno dei due stati, la potenza statica media dissipata è  $\frac{1}{2} \left(\frac{V^{+2}}{R}\right)$ . Per la potenza **dinamica** si considera un carico capacitivo  $C_L$ : assumendo che inizialmente l'ingresso sia alto e l'interruttore sia chiuso, la capacità è inizialmente scarica; rendendo  $v_i$  basso, l'interruttore si apre e  $v_0$  cresce verso  $V^+$  con legge esponenziale.

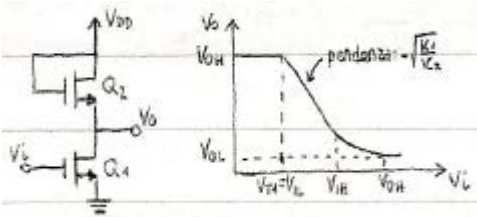
La corrente di carica scorrerà attraverso R producendo dissipazione: sapendo che l'energia assorbita dall'alimentatore :  $\int V + i dt = V + \int i dt = V + a$  (con  $a = C_L V^+$  carica fornita al condensatore), da cui risulta pari a  $C_L (V^+)^2$  e che al termine della transizione l'energia immagazzinata nel condensatore è  $\frac{1}{2} C_L (V^+)^2$ , in totale l'energia dissipata è  $\frac{1}{2} C_L (V^+)^2$ . Successivamente, riportando  $v_i$  ad un valore alto, l'interruttore si chiude e la capacità si scarica sulla resistenza dell'interruttore, che quindi dissipa potenza pari a quella accumulata dal condensatore, cioè  $\frac{1}{2} C_L (V^+)^2$ : in un ciclo completo quindi la potenza dissipata è  $C_L (V^+)^2$  ( e pari a  $f C_L (V^+)^2$  se i cicli sono  $f$ ). A causa dei fenomeni dinamici associati alla commutazione di un transistor, l'interruttore non può rispondere istantaneamente a  $v_i$ ; inoltre la capacità di carico all'uscita rende la forma d'onda  $v_o$  diversa da quella di un impulso ideale ed è presente anche un ritardo tra gli impulsi d'ingresso e di uscita, definito come *ritardo di propagazione*, vi sono diversi modi per esprimere tale ritardo: di solito viene fornito il tempo che intercorre tra l'istante in cui la forma d'onda in ingresso assume il 50% del valore finale e l'istante in cui la forma d'onda d'uscita raggiunge il 50% del corrispondente valore finale; questi tempi sono indicati come  $t_{PHL}$  e  $t_{PLH}$ , per cui il ritardo di propagazione si definisce come loro valor medio:

$t_P = \frac{1}{2} (t_{PHL} + t_{PLH})$ , mentre i tempi di salita e di discesa vengono misurati tra

il 10% e il 90% della transizione. Ciò che si desidera è un basso consumo di potenza e un basso tempo di ritardo, ma queste due caratteristiche sono in contrasto tra loro, poiché si riduce la dissipazione diminuendo la corrente di alimentazione, ma aumentando così il ritardo della portata: un fattore di merito quindi è il *prodotto potenza-ritardo*, definito come  $DP = t_P P_D$  (più è basso più è efficiente).



Per quanto riguarda l'**invertitore Nmos con carico ad arricchimento**, quando il segnale d'ingresso si trova allo 0 logico, cioè ad una tensione minore della tensione di soglia  $V_{T1}$  di  $Q_1$ , quest'ultimo si trova in



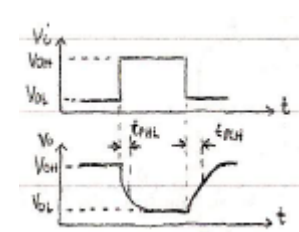
interdizione e la tensione d'uscita risulta alta, con il valore:

$V_{OH} = V_{DD} - V_{T2}$ , cioè l'uscita è minore di  $V_{DD}$  del fattore  $V_{T2}$  (è un difetto: riduce l'escursione di tensione tra gli stati e il margine di rumore), mentre se  $v_i$  è al valore 1 logico, cioè pari a

$V_{DD} - V_{T2}$ . L'uscita risulta  $V_{OL}$  ( $Q_1$  è in triodo e  $Q_2$  in saturazione). La pendenza della caratteristica nella regione di transizione è pari a

$-\sqrt{\frac{K_1}{K_2}}$  dove  $K_R = \frac{K_1}{K_2} = \left(\frac{W}{L}\right)_1$  è detto rapporto geometrico, quindi deve avere una regione di transizione stretta

(e margine di rumore accettabile) deve essere almeno  $K_R > 8$ . L'effetto body è significativo solo quando  $V_0$  è alto e provoca semplicemente una riduzione del valore di  $V_{OH}$ . Per quanto riguarda il funzionamento dinamico, si pone una capacità di carico  $C$  (in cui sono fatti conglobare tutti i vari effetti capacitivi per semplicità) e si assume che l'impulso d'ingresso sia ideale, con tempi di salita e di discesa nulla: prima che venga applicato l'impulso di ingresso, si ha  $v_i = V_{OL}$  e  $v_0 = V_{OH}$ , quindi l'invertitore lavora nel punto B e il condensatore è carico al valore  $V_{OH}$ . Quando  $v_i$  va al valore alto  $V_{OH}$ , il transistor  $Q_1$  entra in regione attiva e, poiché il condensatore non si scarica istantaneamente, il punto di lavoro salta da B a D;  $Q_1$  assorbe una corrente alta, scaricando il condensatore, e il punto di lavoro si sposta lungo la curva verso il punto A, dove  $v_0 = V_{OL}$ . Il ritardo di propagazione  $t_{PHL}$  è il tempo impiegato dal punto di lavoro per passare da D a N:

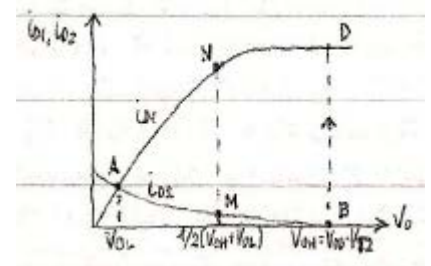


poiché la corrente di scarica del condensatore è in ogni istante la differenza tra  $i_{D1}$  e  $i_{D2}$ , il suo

valore medio è pari a  $I_{HL} = \frac{i_{D1}(D) + i_{D1}(N) - i_{D2}(M)}{2}$ ,

quindi si può calcolare il ritardo di propagazione:

$t_{PHL} = \frac{C[V_{OH} - \frac{1}{2}(V_{OH} + V_{OL})]}{I_{HL}}$ . Successivamente, quando



$V_i$  scende al valore  $V_{OL}$ ,  $Q_1$  si spegne immediatamente e la capacità  $C$  viene caricata dalla corrente fornita da  $Q_2$ , ed il punto di lavoro si sposta da A verso B: sapendo che la corrente

media di carica è:  $I_{LH} = \frac{i_{D2}(A) + i_{D2}(M)}{2}$  si ricava  $t_{PLH} = \frac{C[\frac{1}{2}(V_{OH} + V_{OL}) - V_{OL}]}{I_{LH}}$ . Il ritardo di propagazione può

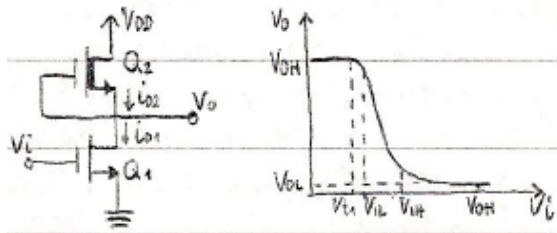
quindi essere individuato facendo la media dei valore trovati per  $t_{PHL}$  e  $t_{PLH}$ . Il prodotto ritardo-potenza della famiglia logica Nmos con carico ad arricchimento può essere calcolato moltiplicando il ritardo di propagazione per la potenza dissipata staticamente (pari a  $\frac{1}{2} V_{DD} I_D^{(V_0=0)}$ ): assumendo  $V_{OL} \cong 0$ , si ha che

$I_{LH} \cong \frac{5}{8} K_2 (V_{DD} - V_t)^2$ , quindi  $t_{PLH} = \frac{0.8 C}{K_2 (V_{DD} - V_t)}$  da cui, poiché  $t_{PHL} \ll t_{PLH}$ , si ricava che  $t_P \cong \frac{1}{2} t_{PLH}$ ;

siccome per  $V_{OL} \cong 0$  la dissipazione media di potenza è:  $P_D \cong \frac{1}{2} K_2 (V_{DD} - V_t)^2 V_{DD}$  si ottiene il prodotto ritardo-potenza:  $D_P \cong 0.2 C V_{DD} (V_{DD} - V_t)$ , e può quindi essere ridotto diminuendo  $C$  oppure  $V_{DD}$  (ma diminuendo  $V_{DD}$  si riduce l'escursione del segnale ed il margine di rumore).



Un **invertitore Nmos con carico a svuotamento** ha un guadagno più elevato, con caratteristica di trasferimento più brusca e margini di rumore migliorati, ottenuti inoltre usando un fattore di forma  $K_R$



minore: il prezzo da pagare è un passo aggiuntivo del processo. In questo caso però ha un ruolo importante l'effetto-body, che non permette al carico a svuotamento di funzionare come generatore di corrente costante per un considerevole intervallo di  $v_0$  (e quindi una caratteristica di trasferimento caratterizzata da una repentina transizione tra

gli stati. Analiticamente la corrente nel carico a svuotamento è data da  $i_{D2} = K_2|V_t|^2$  per  $V_0 \leq V_{DD} - |V_t|$  e  $i_{D2} = K_2[2|V_t|(V_{DD} - V_0) - (V_{DD} - V_0)^2]$  per  $V_0 \geq V_{DD} - |V_t|$ . Per quanto riguarda i margini di rumore, si ha che  $V_{0H} = V_{DD}$ ; per individuare  $V_{0L}$ , si pone  $V_i = V_{0H}$  (quindi  $V_0 = V_{0L}$ ) e lo si ricava da  $i_{D1} = i_{D2}$ , cioè  $K_1[2(V_{GS1} - V_{t1})V_{DS1} - V_{0H}^2] = K_2(V_{GS2} - V_{t2})^2$  con  $V_{GS2} = 0$ ; per ottenere  $V_{1H}$  si deriva  $V_0$  rispetto a  $V_i$  a partire sempre da  $i_{D1} = i_{D2}$  imponendo  $V_i = V_{1H}$  e  $\frac{\delta V_0}{\delta V_i} = -1$ ; analogamente per ottenere  $V_{1L}$  si deriva

$V_0$  rispetto a  $V_i$  a partire sempre da  $i_{D1} = i_{D2}$  e sostituendo  $V_i = V_{1L}$  e  $\frac{\delta V_0}{\delta V_i} = -1$  (avendo così due equazioni in due incognite come prima). Dal punto di vista del funzionamento dinamico, in presenza di una capacità C, i ritardi di propagazione  $t_{PLH}$  e  $t_{PHL}$  possono essere calcolati determinando le correnti medie per caricare e scaricare il condensatore, come nel caso precedente, da cui  $t_{PHL} \cong \frac{C}{2} \frac{V_{DD} - V_{0L}}{I_H} = \frac{C(V_{DD} - V_{0L})}{2K_1(V_{DD} - V_{t1})^2}$  e

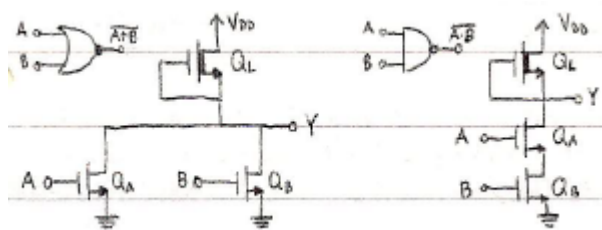
$t_{PLH} \cong \frac{C}{2} \frac{V_{DD} - V_{0L}}{I_L} = \frac{C(V_{DD} - V_{0L})}{2K_2(V_{t2})^2}$ ; il carico a svuotamento fornisce

correnti più alte su un intervallo  $v_0$  più ampio, consentendo di caricare la capacità più velocemente e di avere un ritardo  $t_{PLH}$  leggermente minore. Sapendo inoltre la dissipazione statica media (paria a  $\frac{1}{2}K_2|V_{t2}|^2V_{DD}$ ), si ricava il prodotto ritardo-potenza che è

circa pari a:  $DP \cong \frac{1}{8\alpha C V_{DD}^2}$  dove  $\alpha$  è una frazione minore ma prossima all'unità, che tiene conto delle variazioni di  $V_{t2}$  con  $V_0$ .

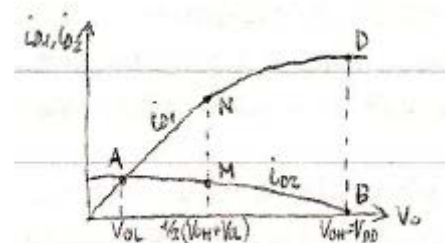
Rispetto all'invertitore con carico ad arricchimento, l'invertitore Nmos con carico a svuotamento è

caratterizzato da un margine di rumore più alto e una maggiore velocità di funzionamento, pur occupando

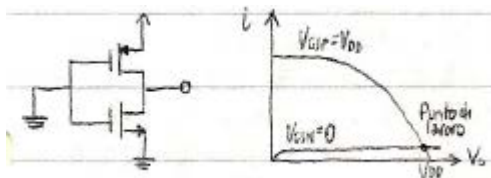
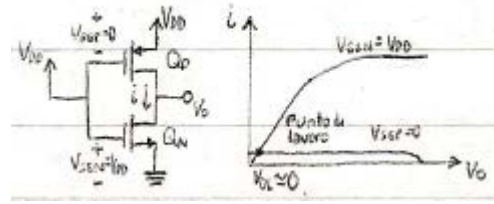


area minore: per questo i circuiti logici e di memoria Nmos utilizzano questa tecnologia. Il funzionamento di una porta **NOR** in tecnologia Nmos con carico a svuotamento è il seguente: se la tensione ad uno dei qualsiasi terminali d'ingresso è alta (a  $V_{DD}$ ) allora il transistor corrispondente è in conduzione e la tensione d'uscita è bassa ( $V_{0L}$ ); l'uscita risulta alta ( $V_{DD}$ ) solo se i

due ingressi sono contemporaneamente bassi (e i transistor interdetti). Nella porta **NAND** invece i transistor di ingresso sono in serie e non in parallelo: l'uscita è passa solo quando entrambi i transistor sono in conduzione, cioè quando tutti e due gli ingressi sono contemporaneamente alti. Nelle porte NAND, per mantenere la tensione d'uscita al valore  $V_{0L}$  quando i transistor sono in conduzione, questi devono avere larghezza doppia di quella del transistor invertitore, in modo che presentino lo stesso rapporto  $\frac{W}{L}$  quindi a parità di ingressi una porta NAND occupa più spazio di una porta NOR.



La tecnologia CMOS è quella più diffusa per i circuiti logici. L'**Invertitore Cmos** utilizza due Mosfet ad arricchimento accoppiati: uno a canale n,  $Q_n$  e l'altro a canale P,  $Q_p$ : il substrato di ogni FET è collegato al suo source e quindi l'effetto-body non è presente. Si considera il transistor  $Q_n$  come transistor pilota e il transistor  $Q_p$  come carico: nel caso in cui  $V_i = V_{DD}$ , il punto di lavoro si troverà all'intersezione fra la curva caratteristica  $i_D - V_{DS}$  per  $Q_n$  con  $V_{GSN} = V_{DD}$  (con  $V_{DSN} = V_0$ ), e la curva di carico  $i_D - V_{SD}$  di  $Q_p$  per  $V_{GSP} = 0$  e, poiché  $V_{SGP} < V_t$  e la curva di carico sarà a corrente quasi nulla, sarà caratterizzato da una corrente quasi nulla e tensione d'uscita prossima a 0; ciò significa che la dissipazione di potenza è molto piccola (si ha  $Q_n$  ON e  $Q_p$  OFF). Quando



invece  $V_i = 0V$ ,  $Q_n$  funziona con  $V_{GSN} = 0$ , quindi la sua curva caratteristica  $i_D - V_{DS}$  è una retta a corrente nulla, mentre per  $Q_p$  la curva di carico funziona per  $V_{SGP} = V_{DD}$ ; nel punto di lavoro quindi la tensione è circa  $V_{DD}$  e la corrente che scorre è quasi nulla, quindi anche in questo caso la dissipazione di potenza è

molto contenuta. L'invertitore CMOS si comporta come un elemento logico ideale: la tensione d'uscita è quasi uguale a 0 e  $V_{DD}$ , mentre la potenza dissipata è quasi nulla. Per individuare la caratteristica di trasferimento completa occorre individuare i punti critici della risultante curva; a tal proposito occorre tenere presente le relazioni  $i - V$  di  $Q_n$  (cioè:  $i_{DN} = K_N [2 (V_i - V_{tn}) V_0 - V_0^2]$  per  $V_0 \leq V_i - V_{tn}$  e  $i_{DN} = K_N (V_i - V_{tn})^2$  per  $V_0 \geq V_i - V_{tn}$ )

e di  $Q_p$  (cioè:  $i_{DP} = K_P [2 (V_{DD} - V_i - |V_{tp}|) (V_{DD} - V_0) - (V_{DD} - V_0)^2]$

per  $V_0 \geq V_i + |V_{tp}|$  e  $i_{DP} = K_P (V_{DD} - V_i - |V_{tp}|)^2$  per  $V_0 \leq V_i + |V_{tp}|$ ), sapendo che generalmente si ha  $V_{tn} = |V_{tp}| = V_t$  e  $K_n = K_p = K$  (così l'invertitore può erogare la stessa corrente in entrambe le direzioni). La caratteristica di trasferimento può essere divisa in

cinque segmenti, corrispondenti a cinque diversi modi di funzionamento di  $Q_n$  e  $Q_p$  (fino ad A,  $Q_n$  è OFF; da A a B,  $Q_n$  è in saturazione e  $Q_p$  in triodo; da B a C entrambi in saturazione; da C a D  $Q_n$  è in triodo e  $Q_p$  in saturazione, dopo D  $Q_p$  è OFF);

trascurando la resistenza d'uscita finita, il segmento BC è verticale e il guadagno in questa regione è infinito, con  $V_0(B) = \frac{V_{DD}}{2} + V_t$  e

$V_0(C) = \frac{V_{DD}}{2} - V_t$ . Per determinare  $V_{1H}$  basta osservare che  $Q_n$  è in triodo e  $Q_p$  in saturazione, quindi eguagliare le rispettive correnti

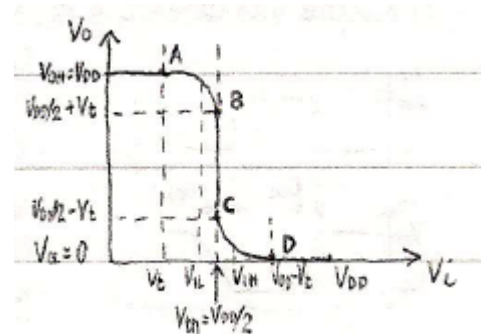
ottenendo:  $i_{DN} = i_{DP}$  cioè  $2 (V_i - V_t) V_0 - V_0^2 = (V_{DD} - V_i - V_t)^2$ ; derivare entrambi i membri rispetto a  $V_i$ :  $2(V_i - V_t) \frac{\delta V_0}{\delta V_i} + 2V_0 - 2V_0 \frac{\delta V_0}{\delta V_i} = -2 (V_{DD} - V_i - V_t)$  da cui, sostituendo  $V_i = V_{1H}$  e  $\frac{\delta V_0}{\delta V_i} = -1$  si ricava

che  $V_0 = V_{1H} - \frac{V_{DD}}{2}$ . Sostituendo questo valore nell'equazione di partenza e imponendo  $V_i = V_{1H}$  si ottiene quindi  $V_{1H} = \frac{1}{8} (5V_{DD} - 2V_t)$ . Per calcolare  $V_{1L}$  si procede in maniera analoga oppure si può usare la

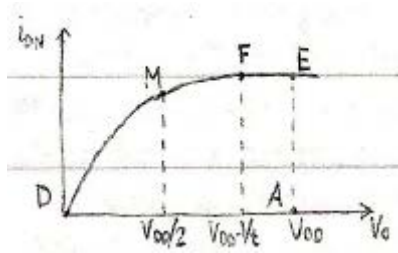
relazione di simmetria:  $V_{1H} - \frac{V_{DD}}{2} = \frac{V_{DD}}{2} - V_{1L}$  utilizzando  $V_{1H}$ ;  $V_{1L} = \frac{1}{8} (3V_{DD} + 2V_t)$ ; a questo punto è

possibile calcolare i margini di rumore  $NM_H = V_{OH} - V_{1H}$  e  $NM_L = V_{1L} - V_{OL}$  (che, per la simmetria, risultano uguali). Da notare che mentre l'invertitore cambia stato, attraverso la connessione in serie dei due Mos scorre corrente (il cui picco è  $V_i = \frac{V_{DD}}{2}$ ); questa corrente provoca dissipazioni di potenza nel

funzionamento dinamico, che è significativa quando l'invertitore è connesso a una capacità di carico C, per cui si ha:  $P_D = f C V_{DD}^2$ , con  $f$  frequenza con cui commuta l'invertitore.



In presenza di carichi capacitivi, l'invertitore Cmos presenta tempi di salita e discesa uguali, essendo il



circuito simmetrico, per cui basta considerare uno solo dei processi di commutazione, considerando come ingresso un impulso ideale: si suppone all'istante  $t = 0^-$  la tensione d'uscita uguale a  $V_{DD}$  e il condensatore carico a tale valore; all'istante  $t=0$ , l'impulso d'ingresso passa da  $V_i = V_{OL} = 0$  a  $V_i = V_{OH} = V_{DD}$  mandando  $Q_P$  in interdizione. Il punto di lavoro per  $t=0^+$  è il punto E, in cui  $Q_n$  si trova in saturazione e fa scorrere una corrente molto grande; scaricando C, la corrente rimane

costante finché  $V_O = V_{DD} - V_t$  (punto F), e questa prima porzione di scarica è

$t_{PHL1} = \frac{C(V_{DD} - (V_{DD} - V_t))}{K_N(V_{DD} - V_t)^2} = \frac{CV_t}{K_N(V_{DD} - V_t)^2}$ . Oltre il punto F il transistor  $Q_P$  è sempre OFF mentre  $Q_n$  entra nella regione di triodo; quest'altra porzione di intervallo di scarica è descritta da  $i_{DN} dt = -C dV_O$  e, poiché

$$i_{DN} = K_N[2(V_i - V_{tn})V_O - V_O^2] \text{ con } V_i = V_{DD}, \text{ si ricava: } \frac{-K_N}{C} dt = \left[ \frac{1}{2(V_{DD} - V_t)} \right] \left[ \frac{dV_O}{\left( \frac{V_O^2}{2(V_{DD} - V_t)^2} \right) - V_O} \right]. \text{ Per}$$

trovare il tempo necessario perché  $V_O$  scenda da  $V_{DD} - V_t$  al punto di mezzo  $\frac{V_{DD}}{2}$ , si integrano entrambi i

membri, ottenendo  $t_{PHL2} = \frac{C}{2K_N(V_{DD} - V_t)} \ln \frac{3V_{DD} - 4V_t}{V_{DD}}$ . Sommando i due termini ottenuti si ricava il ritardo

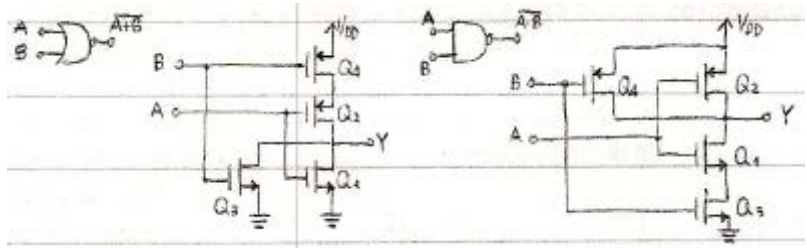
complessivo e, sapendo che di solito  $V_t \cong 0.2 V_{DD}$ , si semplifica con:  $t_{PHL} = \frac{0.8 C}{K_N V_{DD}}$  (analogamente si ha

$t_{PLH} = \frac{0.8 C}{K_P V_{DD}}$ ). Il prodotto ritardo-potenza è direttamente proporzionale alla velocità di commutazione e può essere ridotto lavorando a basse frequenze, oltre a poter diminuire la capacità di carico e la tensione di alimentazione come per gli Nmos.

Le porte logiche Cmos si ottengono modificando il circuito di base dell'invertitore: considerando la porta

**NOR**, il cuore della porta è

l'invertitore  $Q_1 - Q_2$  e la porta è ottenuta aggiungendo il Mosfet  $Q_3$  a canale n in parallelo e il Mosfet  $Q_4$  a canale P in serie (per ogni ingresso addizionale), mentre la porta **NAND** è ottenuta collegando un Mosfet a

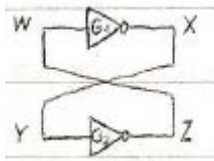


canale n in serie ed un Mosfet a canale p in Parallelo (per ogni ingresso addizionale). L'uscita del primo circuito è alta ( $V_{DD}$ ) se sia  $Q_2$  che  $Q_4$  sono in conduzione, cioè quando A e B sono contemporaneamente bassi, mentre nel secondo circuito l'uscita è bassa solo quando  $Q_1$  e  $Q_3$  sono entrambi in conduzione, cioè

quando A e B sono entrambi alti. In generale quindi le porte Cmos occupano area maggiore di quelle Nmos; inoltre si realizzano porte Cmos in modo che erogino stessa corrente in entrambe le direzioni, occorre tener conto che transistor in parallelo forniscono una corrente che è somma delle singole fornite mentre transistor in serie forniscono la stessa corrente che fornirebbe un solo transistor, quindi per una porta NOR deve essere:

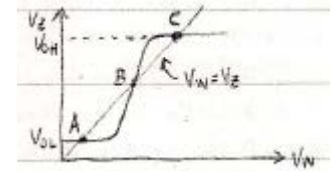
$$\left(\frac{W}{L}\right)_P \cong 2N\left(\frac{W}{L}\right)_N \text{ e per una NAND: } \left(\frac{W}{L}\right)_N \cong \frac{N}{2}\left(\frac{W}{L}\right)_P.$$

La memoria è una parte molto importante dei sistemi digitali; i circuiti logici con memoria sono chiamati circuiti sequenziali, cioè l'uscita dipende non solo dal valore attuale degli ingressi ma anche da quelli precedenti (è necessario un generatore di temporizzazione, il clock). L'elemento di memoria fondamentale è



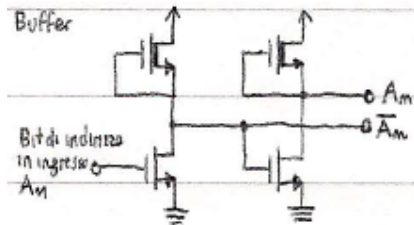
il **latch** (bistabile), formato da due invertitori interallacciati,  $G_1$  e  $G_2$ , che formano una controreazione positiva. Per studiarne il funzionamento si può interrompere l'anello e applicare un ingresso ad uno dei due invertitori, ad esempio il segnale  $V_W$  a  $G_1$ : la caratteristica di trasferimento di  $V_Z$  in funzione di  $V_W$  è suddivisa in tre segmenti, di cui

quello centrale rappresenta la transizione (se l'anello è chiuso si ha la retta  $V_W = V_Z$ ) e i punti di lavoro sono A e C, che sono stabili, cioè il latch vi rimane per un tempo indefinito, e B che è instabile; infatti se fosse polarizzato nel punto B, con un aumento di  $V_W$  si avrebbe un aumento di  $V_X$  (pari al prodotto  $V_W$  per il guadagno  $G_1$ ) che, applicato a  $G_2$ , porterebbe ad un ulteriore aumento

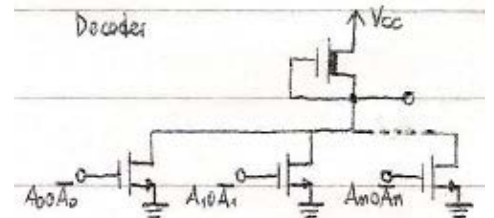


di segnale nel punto Z:  $V_Z$  è legato a  $V_W$  attraverso il guadagno di anello nel punto B (dato dalla pendenza della curva nel punto B) e aumenta finché il punto di lavoro si sposta su C (dove il guadagno d'anello è 0); con una diminuzione di  $V_W$  invece il punto di lavoro si sarebbe spostato su A. Nel punto C  $V_W$  è alto,  $V_X$  è basso,  $V_Y$  è basso e  $V_Z$  è alto, mentre nel punto A i valori sono invertiti; se si considerano come uscita i punti X e Z si osserva che il latch è un circuito bistabile con due diverse uscite complementari, e può mantenere in memoria un solo bit di informazione: attraverso un meccanismo di trigger il latch è forzato a cambiare stato e il latch unito al circuito di trigger forma un flip-flop. Il più semplice flip-flop è quello **set-reset (SR)**, formato da due porte NOR incrociate; il secondo ingresso di ciascuna porta serve come ingresso forzante, il set S porta lo stato ad 1 e il reset R porta lo stato a 0; l'uscita memorizza 1 quando  $a$  è alto e  $\bar{a}$  è basso, mentre memorizza 0 se  $a$  è basso e  $\bar{a}$  è alto. Se  $S=1$ ,  $R=0$  allora l'uscita è 1; se  $S=0$ ,  $R=1$  allora l'uscita è 0; se  $S=0$ ,  $R=0$ , il flip-flop mantiene lo stato precedente; la combinazione  $S=1$ ,  $R=1$  non è consentito poiché porta il flip-flop in uno stato indefinito. È possibile realizzare un flip-flop SR anche utilizzando due porte NAND. Un tipo di flip-flop particolare è il flip-flop **D in tecnologia Cmos**; sono presenti due ingressi, cioè l'ingresso del dato (D) e di clock (C) e le due uscite  $a$  e  $\bar{a}$ ; quando il clock è basso il flip-flop è a riposo e qualunque cambiamento su D non produce effetti, mentre se il clock è alto il flip-flop acquisisce il valore sulla linea D (alcuni contengono gli ingressi di set e reset). È formato da due flip-flop SR, chiamati master e slave, formati da due porte NOR e una di trasmissione inserita nell'anello di controreazione; sono presenti anche una porta di trasmissione tra l'ingresso D e il master e una porta di trasmissione tra il master e lo slave. Quando il clock è basso lo slave è isolato dal master ed il suo anello è chiuso, consentendogli di mantenere lo stato, mentre, l'anello del master è aperto e la sua uscita è il complemento di D, mentre quando il clock è alto il master viene scollegato dall'ingresso D e il suo anello chiuso, mentre l'anello dello slave viene aperto e riceve l'uscita del master, quindi l'uscita assume valore complementare a  $a$  (quindi il valore di D).

La memoria di un computer può essere divisa in due parti, memoria principale e memoria di massa; la memoria principale, ad accesso più rapido, è di solito una **memoria ad accesso casuale (RAM)**, con tempo di accesso in scrittura e lettura indipendenti dalla localizzazione dell'informazione (mentre la memoria di massa di solito è di tipo sequenziale). Il cuore di un chip di memoria è costituito dalle celle in cui sono immagazzinati i bit, cioè ogni cella di memoria è un circuito capace di immagazzinare un bit, e sono organizzate in matrice quadrata (ogni cella è selezionata rendendo alti la riga e la colonna corrispondente, queste sono selezionate attraverso i rispettivi decoder). Gli ingressi di indirizzo di un chip RAM sono di



solito interfacciati attraverso un invertitore: ad esempio in un RAM Nmos vengono utilizzati i circuiti con carico a svuotamento, che forniscono in uscita il bit d'ingresso e il suo complemento. Ogni decoder d'indirizzo è di solito un circuito combinatorio, come la porta NOR: la sua uscita è connessa a una riga di memoria (se



è un decoder per li indirizzi di memoria di riga) ed i suoi ingressi saranno connessi alla opportuna combinazione dei bit d'ingresso e dei loro complementi. Esistono due tipi di RAM Mos: statiche e dinamiche.

Le prime utilizzano come celle di memoria dei flip-flop, mentre le seconde memorizzano dati su delle capacità. La cella di una RAM statica è formata da un flip-flop che contiene due invertitori accoppiati in modo incrociato e da due transistor di accesso  $a_S$  e  $a_G$ , che si accendono quando viene selezionata la riga corrispondente e connettono il flip-flop sia alla linea di colonna D sia alla linea di colonna negata  $\bar{D}$ ; la tensione di segnale tra D e  $\bar{D}$  viene fornito all'amplificatore di lettura dalla colonna (uno per colonna), l'unico attivo, e la sua uscita verrà connessa alla linea data-output, di uscita dei dati, del chip (operazione di lettura). I dati che devono essere scritti sono trasferiti sulle linee D e  $\bar{D}$  (il bit e il suo complemento) e i transistor e i transistor in conduzione  $a_S$  e  $a_G$  li portano i ingresso al flip-flop (operazione di scrittura). Da osservare che i tempi di salita e di discesa dei segnali delle linee D e  $\bar{D}$  sono quindi determinati dalle correnti che scorrono nei transistor dei flip-flop, e dalle capacità delle linee D e  $\bar{D}$  stesse: questi contribuiscono al tempo di accesso in RAM. Le RAM statiche possono mantenere il loro contenuto indefinitamente, e non è essenziale per il funzionamento il clock. In una RAM dinamica uno zero logico è rappresentato da assenza di carica, quindi una tensione prossima allo zero, sul condensatore, mentre un 1 logico è rappresentato da una tensione prossima a quella di alimentazione. A causa di effetti di perdita, il condensatore tende a perdere la sua carica, quindi per il concetto funzionamento della RAM dinamica si usa l'operazione di refresh (le celle vengono lette e riscritte) ogni 2-4 ms, e ciò implica la presenza di un clock. Ogni cella è formata da un solo Mosfet ad arricchimento a canale n, transistor di accesso, e da un condensatore di memoria: in fase di lettura, il condensatore è connesso alla corrispettiva linea D, quindi in parallelo alla sua capacità; se viene letto 1 la tensione sulla capacità di linea subisce un incremento, che viene amplificato e portato in uscita, altrimenti rimane invariata. Il segnale amplificato viene anche re-impostato sul condensatore di memoria, ripristinando il livello originale, e in questo modo vengono rinfrescate tutte le celle della riga selezionata. La fase di scrittura è analoga, cioè nel memorizzare il dato su una cella, tutte le celle di riga vengono rinfrescate.



Le **memorie di sola lettura (ROM)** invece contengono dei dati fissi, mantengono le informazioni anche in assenza di alimentazione: può essere vista come un circuito combinatorio in cui l'ingresso è l'insieme dei bit di indirizzo e l'uscita è l'insieme dei bit letti nella localizzazione indirizzata. Una ROM in tecnologia Mosfet è formata da una matrice di transistor ad arricchimento i cui gate sono connessi alla rispettiva riga. I source sono a massa e i drain alle linee di bit: ogni linea di bit è connessa all'alimentazione attraverso un transistor di carico a svuotamento. All'interno delle celle che memorizzano uno 0 è presente un Mosfet, mentre nelle celle che memorizzano 1 non è presente alcun dispositivo: portando a livello alto la tensione di riga infatti i transistor andranno in conduzione, portando a 0 la loro uscita, mentre quelle senza transistor rimarranno alla tensione di alimentazione grazie al transistor di carico a svuotamento. Questo comunque è un tipo di ROM determinato al momento di fabbricazione (quindi personalizzato); esistono delle ROM programmabili a maschera, in cui la superficie è rivestita da uno strato di alluminio e quindi viene asportato selettivamente, lasciando il metallo solo dove si desiderano delle connessioni: si tratta di ROM programmabili nell'ultima fase di programmazione. Vi sono poi le PROM, programmabili dall'utilizzatore ma una sola volta, bruciando i fusibili desiderati attraverso una forte corrente. Le EPROM sono ROM cancellabili, quindi programmabili un numero di volte indefinito: ogni cella è costituita da un Mosfet ad arricchimento a canale n, con due gate, uno dei quali non è elettricamente connesso a nessun'altra parte del circuito e per questo è chiamato gate flottante (che mantiene intrappolata la carica al suo interno anche in assenza di alimentazione, una volta programmato- se c'è carica allora è uno 0 logico, altrimenti se il transistor è interdetto è un 1 logico): per cancellare la programmazione (quindi eliminare la carica immagazzinata) occorre usare luce ultravioletta per un tempo prefissato, che fornisce energia agli elettroni intrappolati sufficiente a superare la barriera di potenziale e tornare nel substrato.