

ale-cci

Elettronica 1

July 17, 2020

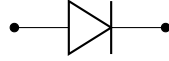
Contents

Diodo a giunzione PN	1
Diodo	1
Modello a soglia	2
Circuito limitatore di tensione	4
Circuito limitatore inferiore	5
Circuito rivelatore di massimo	5
Circuito rivelatore di minimo	7
Circuito limitatore di tensione superiore ed inferiore	8
Raddrizzatore a doppia semionda	10
Rivelatore di cresta	11
Diodi in regime dinamico	13
 Transistore bipolare a giunzione BJT	 17
Modello di Ebers e Moll	17
Amplificatore invertente di tensione per piccoli segnali	22
Approssimazione Transistor BJT	25
 Logica RTL: Transistor Resistor Logic	 29
Margine di immunità al rumore	29
Porta logica NOR	33
 Transistor in regione dinamica	 37
 Dispositivi e Circuiti Elettronici	 39
Lezione 4	46
Ragionamento su struttura diversamente drogata	47
Lezione 2019514	53

Applicazioni del transistor bipolare	54
Transistor JFET	56
Transistor MOSFET	57
Esempio circuito	65
NMOS a carico saturo	66
Modifica rete di pull-up con transistor a depletion	68
Tecnologia cmos	72
Prestazioni e qualità dell'invertitore	76
Amplificatore differenziale	81
Considerazioni su opamp	82
Amplificatore lineare non invertente	85
Circuito sommatore analogico	86
Circuito derivatore	86
.	89
RTL in regione dinamica	91

Diodo a giunzione PN

Diodo



Il diodo è un componente la cui corrente ha una formula esponenziale, dipendente dalla tensione applicatagli ai suoi capi:

$$I = I_s(e^{V/V_T} - 1)$$

Dall'espressione è possibile notare come per $V \rightarrow -\infty$, la corrente tende asintoticamente al valore $-I_s$, detto *corrente di saturazione*. Dato che I_s assume valori molto bassi, sull'ordine di 10^{-15} , si può considerare questa corrente trascurabile, e può essere considerata nulla.

Nel caso in cui $V > V_T$, il valore esponenziale della corrente inizia a farsi sentire, e scorre una forte corrente attraverso il componente.

È quindi possibile distinguere due regioni di funzionamento: una caratterizzata da una bassa corrente, ed una caratterizzata da una corrente elevata. Le due regioni prendono il nome di *polarizzazione diretta* e *polarizzazione inversa*.

Diremo che il diodo è spento (non fa passare corrente) in regione di polarizzazione inversa, mentre diremo che è acceso in regione di polarizzazione diretta.

Analisi del circuito raddrizzatore a singola semionda

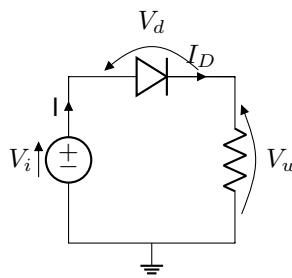


Figure 1: Raddrizzatore a singola semionda

Il circuito in figura è descritto dalle relazioni:

$$\begin{cases} V_i = V_d + V_u \\ I_D = I_s(e^{V_d/V_T} - 1) \\ I = I_D \end{cases}$$

V_T indica la tensione termica del diodo: $V_T = K \frac{T}{q}$, con K costante di Boltzmann, q la carica di un elettrone e T la temperatura del circuito (misurata a temperatura ambiente: $300K$).

Siccome è prodotto di costanti, V_T è considerabile come una costante che vale approssimativamente $26mV$.

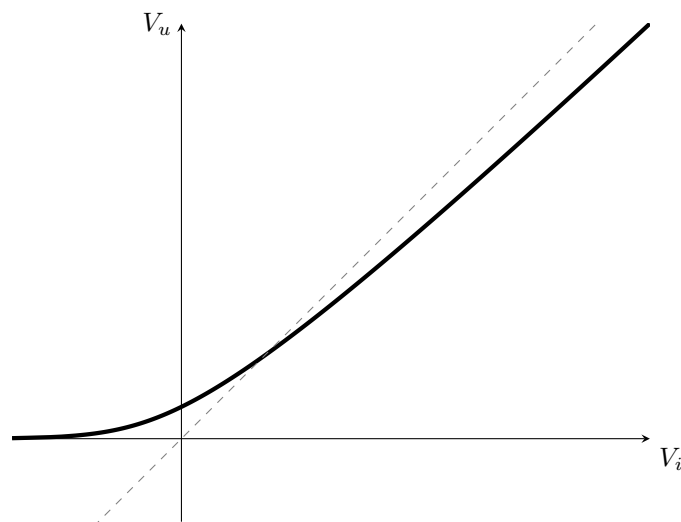
Svolgendo il sistema otteniamo l'equazione

$$V_u = RI_S(e^{\frac{V_i - V_u}{V_T}} - 1)$$

Risolviendo per V_i

$$V_i = V_u + V_T \ln(V_u + RI_S)$$

Otteniamo che V_i è composto dalla somma tra un componente lineare ed uno logaritmico, possiamo quindi tracciare un grafico approssimato di $V_i(V_u)$ e per ottenere $V_u(V_i)$ basta effettuare una simmetria sulla bisettrice:



Modello a soglia

Per semplificare la risoluzione dei circuiti con diodo, possiamo studiare le due regioni di funzionamento separatamente, approssimando i due tratti della caratteristica con due lineari: Il primo descritto dalla caratteristica $I = 0$ (diodo spento) valido per $V_d < V_\gamma$, ed il secondo una semiretta perpendicolare alle ascisse $V = V_\gamma$ a rappresentare la corrente costante a diodo acceso.

P. INVERSA (off)

$$\begin{cases} I = 0 \\ V < V_\gamma \end{cases}$$

P. DIRETTA (on)

$$\begin{cases} V = V_\gamma \\ I > 0 \end{cases}$$

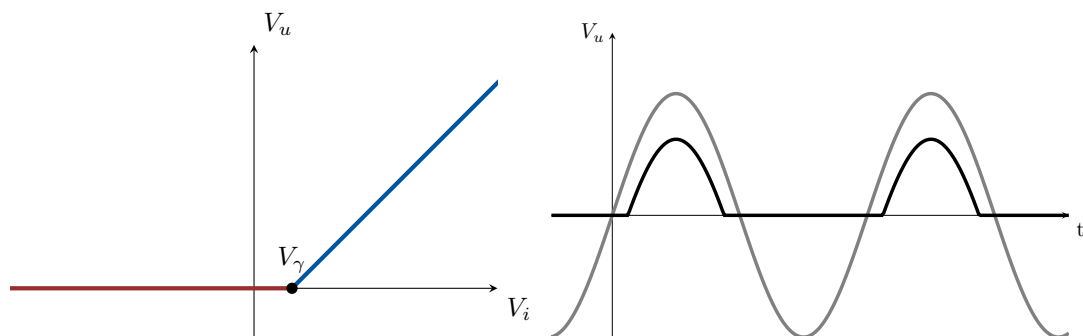
Questa semplificazione porta a dover risolvere il circuito due volte, una per ogni stato di funzionamento per cui il diodo si può trovare. Inoltre è necessario anche controllare le relative ipotesi di funzionamento.

Partendo dall'ipotesi che il diodo sia spento, $V_u = R \cdot I = 0$. Inoltre abbiamo che $V_d = V_i - V_u = V_i$, e per ipotesi di funzionamento $V_d < V_\gamma$, quindi il diodo è spento per $V_i < V_\gamma$.

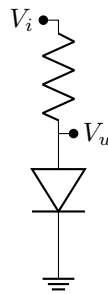
Studiando ora la seconda ipotesi, diodo acceso, abbiamo che: $V_u = V_i - V_d = V_i - V_\gamma$. E come condizioni di validità:

$$\begin{cases} V_u = R \cdot I \\ I \cdot R > 0 \\ V_u = V_i - V_d \end{cases} \Rightarrow V_i - V_d > 0 \Rightarrow V_i > V_d$$

Grafico dell'uscita ed andamento del circuito con ingresso sinusoidale



Circuito limitatore di tensione



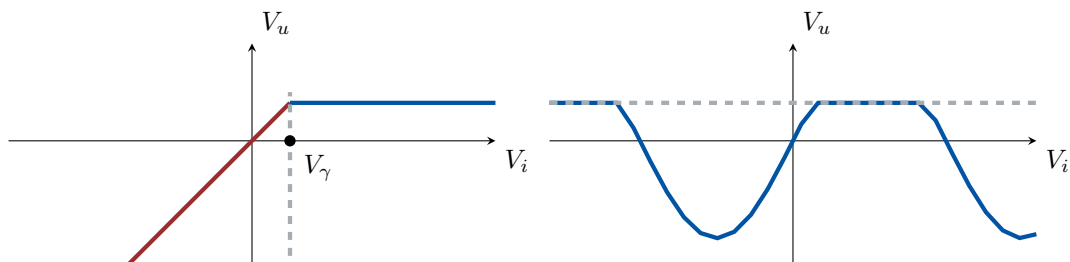
Diodo OFF

$$I = 0 \Rightarrow \begin{cases} V_i = V_u \\ V_u < V_\gamma \end{cases}$$

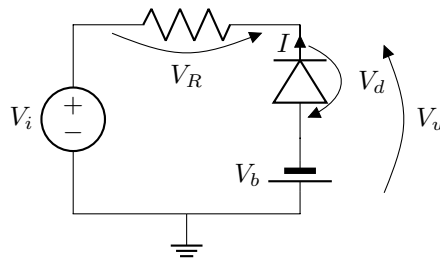
Diodo ON

$$\begin{cases} V_u = V_\gamma \\ I_D = \frac{V_i - V_\gamma}{R} > 0 \Rightarrow V_i > V_\gamma \end{cases}$$

Possiamo osservare che se il segnale in ingresso eccede V_γ , il diodo si accende e limita l'uscita a V_γ . Da questo il nome *circuito limitatore di tensione*.



Circuito limitatore inferiore



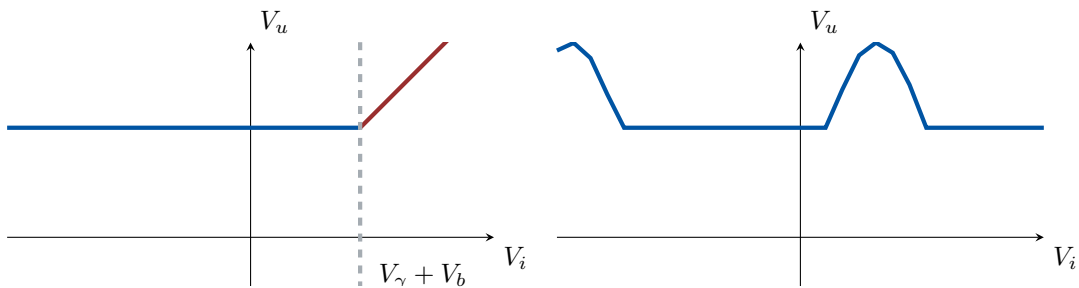
Diodo OFF

$$\begin{cases} I = 0 \\ V_d < V_\gamma \end{cases} \cup \begin{cases} V_i = V_u - V_R \\ V_u = V_b - V_d \end{cases} \Rightarrow \begin{cases} V_i = V_u \\ V_i > V_b - V_\gamma \end{cases}$$

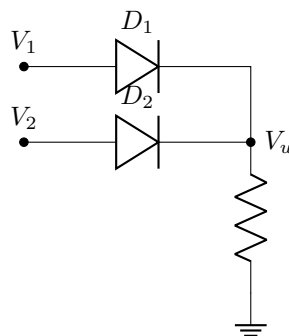
Diodo ON

$$\begin{cases} V_d = V_\gamma \\ I > 0 \end{cases} \cup \begin{cases} V_i = V_u - V_R \\ V_u = V_b - V_d \end{cases} \Rightarrow \begin{cases} V_u = V_b - V_\gamma \\ V_i < V_b - V_\gamma \end{cases}$$

Il circuito effettua una limitazione sui valori bassi, e la soglia di intervento è regolabile dal parametro V_b .



Circuito rivelatore di massimo



In questo caso, avendo due diodi, ciascuno descritto da un modello lineare a tratti, caratterizzato da due regioni distinte, abbiamo quattro regimi di funzionamento differenti.

Relazioni fondamentali

$$V_{d1} = V_1 - V_u$$

$$V_{d2} = V_2 - V_u$$

$$V_u = R \cdot I$$

$$I_1 + I_2 = I$$

D1 e D2 OFF

$$\begin{cases} V_u = 0 \\ V_1 < V_\gamma \\ V_2 < V_\gamma \end{cases}$$

D1 ON e D2 OFF

$$\begin{cases} V_u = V_1 - V_\gamma \\ V_1 > V_\gamma \\ V_1 > V_2 \end{cases}$$

D1 OFF e D2 ON

$$\begin{cases} V_u = V_2 - V_\gamma \\ V_2 > V_\gamma \\ V_2 > V_1 \end{cases}$$

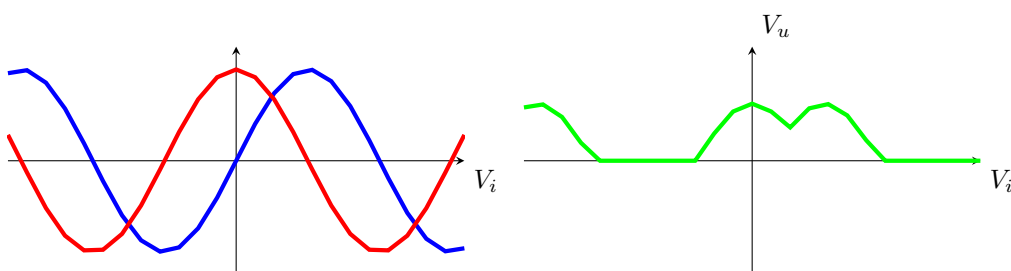
D1 ON e D2 ON

$$\begin{cases} V_1 = V_2 \\ V_1 > V_\gamma \\ V_2 > V_\gamma \end{cases}$$

Quando entrambi i diodi sono spenti, l'uscita è 0.

Quando la tensione V_1 è maggiore sia di V_γ , che di V_2 , la tensione di uscita segue il valore di V_1 a meno di una costante, V_γ . In maniera del tutto analoga, quando V_2 è maggiore di V_γ e V_1 , la tensione di uscita segue il valore di V_2 a meno di V_γ .

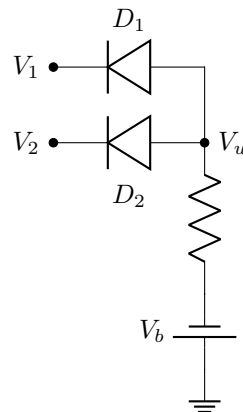
Quando V_1 e V_2 sono uguali, e sono entrambi maggiori di V_γ , allora l'uscita segue l'uno o l'altro a meno di una costante. In altre parole $V_u = \max \{0, V_1 - V_\gamma, V_2 - V_\gamma\}$.



Se si volesse estendere questo circuito per trovare il massimo tra tre ingressi, si potrebbe tranquillamente fare aggiungendo un' altro ramo in ingresso.

Nel caso in cui V_1 e V_2 siano segnali digitali, ovvero che possono solo assumere due valori V_h e V_l , il circuito si comporta come una porta logica **OR**.

Circuito rivelatore di minimo



Relazioni fondamentali

$$\begin{aligned}V_{d1} &= V_u - V_1 \\V_{d2} &= V_u - V_2 \\I &= I_1 + I_2 \\V_R &= R \cdot I = V_b - V_u\end{aligned}$$

D1 e D2 OFF

$$\begin{cases} V_u = V_b \\ V_1 > V_b - V_\gamma \\ V_2 > V_b - V_\gamma \end{cases}$$

D1 ON e D2 OFF

$$\begin{cases} V_u = V_1 + V_\gamma \\ V_1 < V_b - V_\gamma \\ V_1 < V_2 \end{cases}$$

D1 OFF e D2 ON

$$\begin{cases} V_u = V_2 + V_\gamma \\ V_2 < V_b - V_\gamma \\ V_2 < V_1 \end{cases}$$

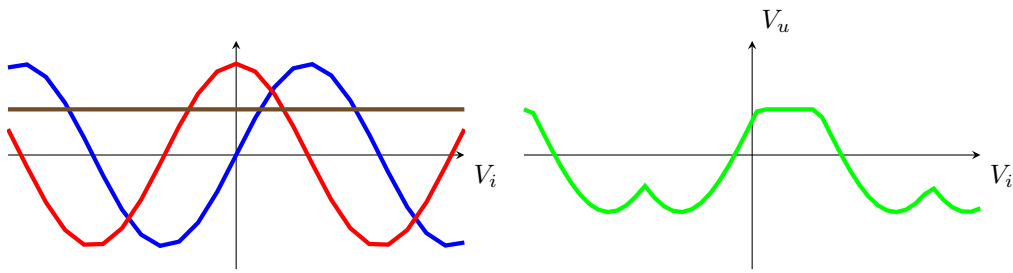
D1 e D2 ON

$$\begin{cases} V_u = V_1 + V_\gamma \\ V_1 = V_2 \\ V_1 < V_b - V_\gamma \end{cases}$$

Quando entrambi i segnali di ingresso sono superiori a $V_b - V_\gamma$, entrambi i diodi sono spenti e la tensione coincide con V_b . Se la tensione V_1 scende al di sotto di $V_b - V_\gamma$ ed è minore di V_2 , allora l'uscita segue V_1 a meno di una costante V_γ . Stesso succede quando V_2 scende al di sotto di $V_b - V_\gamma$. Se entrambe le tensioni di ingresso hanno lo stesso valore e sono al di sotto di $V_b - V_\gamma$, entrambi i diodi sono accesi e l'uscita segue l'uno o l'altro a meno di una costante.

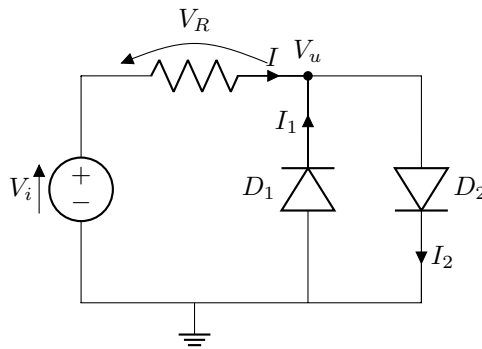
L'uscita V_u può essere vista come $V_u = \min \{V_1 + V_\gamma, V_2 + V_\gamma, V_b\}$.

Se si fa riferimento a segnali di tipo digitale, il circuito si comporta come una porta logica **AND**.



Circuito limitatore di tensione superiore ed inferiore

Negli ultimi due esempi abbiamo dovuto analizzare quattro casi, uno per ogni zona possibile in cui i due diodi del circuito potevano trovarsi. È evidente quindi, che nel caso più generale, con n diodi presenti nel circuito, il numero di casi da analizzare crescerebbe come 2^n . Quello che è importante da osservare è che non tutte le combinazioni sono significative dal punto di vista fisico. Alcune di esse possono essere escluse facendo ragionamenti a priori.



Equazioni generali

$$V_u + I \cdot R = V_i$$

$$V_1 = -V_u$$

$$V_2 = V_u$$

$$I + I_1 = I_2$$

D1 e D2 OFF

$$\begin{cases} V_i = V_u \\ V_i < V_\gamma \\ V_i > -V_\gamma \end{cases}$$

D1 ON e D2 OFF

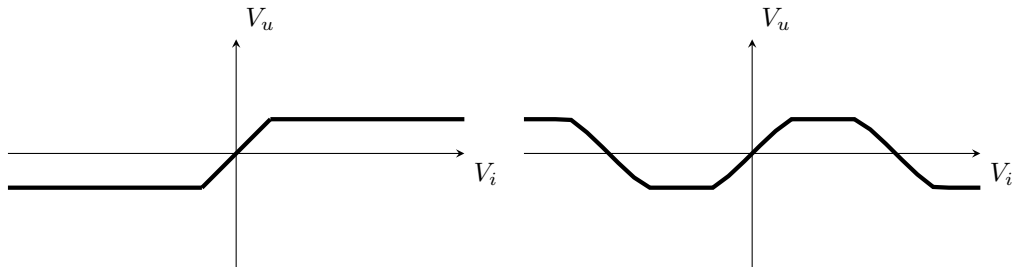
$$\begin{cases} V_u = -V_\gamma \\ V_i < -V_\gamma \end{cases}$$

D1 OFF e D2 ON

$$\begin{cases} V_u = V_\gamma \\ V_i > V_\gamma \end{cases}$$

D1 e D2 ON

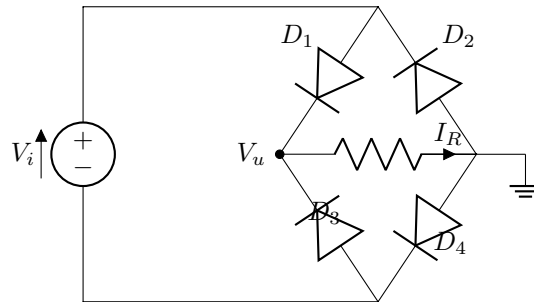
$$\begin{cases} V_u = -V_\gamma \\ V_u = V_\gamma \\ \text{Non verificabile} \end{cases}$$



Al momento dell'impostazione delle equazioni generali, si poteva direttamente notare che $V_1 = -V_2$, quindi entrambi i diodi non potevano essere accesi allo stesso tempo.

Per modificare le due soglie del raddrizzatore basta mettere in serie nel circuito due generatori di tensione.

Raddrizzatore a doppia semionda



Dal numero di diodi presenti nel circuito mi attendo $2^4 = 16$ combinazioni delle regioni di funzionamento del circuito.

Equazioni Generali

$$I = I_1 - I_2 = I_3 - I_4$$

$$V_i = V_1 + V_R + V_4$$

$$-V_i = V_3 + V_R + V_2$$

$$I_R = I_2 + I_4$$

$$I_R = I_3 + I_1$$

$$V_u + V_1 + V_2 = 0$$

$$V_u + V_3 + V_4 = 0$$

Risolviendo il circuito, otteniamo che le uniche soluzioni che hanno senso fisico sono le seguenti:

**D1,D4 OFF
D2,D3 ON**

$$\begin{cases} V_u = -2V_\gamma - V_i \\ V_i < -2V_\gamma \end{cases}$$

**D1,D2
D3,D4 OFF**

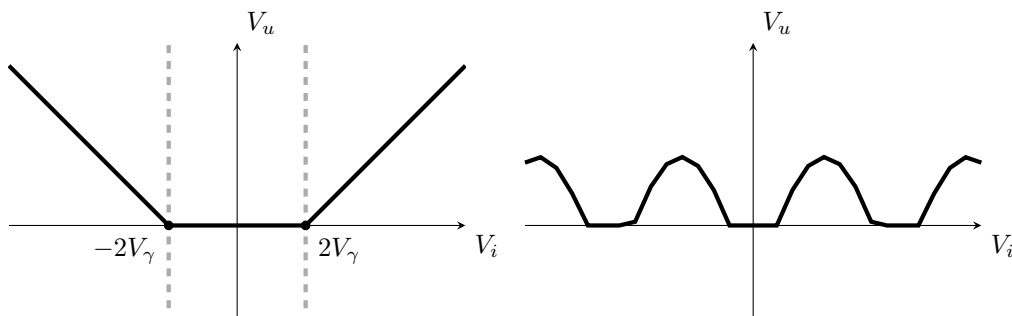
$$\begin{cases} V_u = 0 \\ V_i < 2V_\gamma \\ V_i > -2V_\gamma \end{cases}$$

**D1,D4 ON
D2,D3 OFF**

$$\begin{cases} V_u = V_i - 2V_\gamma \\ V_i > 2V_\gamma \end{cases}$$

Il circuito ha un comportamento analogo al raddrizzatore a singola semionda, ha di diverso un tratto a pendenza negativa. Ciò significa che per un valore negativo di $V_i < -V_\gamma$, l'uscita assume il valore positivo opposto. Quindi a differenza del circuito a singola semionda, che taglia la semionda negativa, questo circuito la trasforma in semionda positiva.

Il circuito raddrizzatore a doppia semionda è utilizzato per la trasformazione da corrente alternata a corrente continua.



Rivelatore di cresta

Quando ho in ingresso un segnale sinusoidale (a valor medio nullo), tutti i circuiti raddrizzatori visti fino ad ora, hanno la caratteristica di aver il valor medio della tensione in uscita, maggiore di zero. In particolare nel caso del raddrizzatore a doppia semionda, la trasformazione delle semionde negative, contribuisce ulteriormente al valor medio risultando in un valore maggiore rispetto al raddrizzatore a singola semionda. È stato anche accennato che questi circuiti sono utilizzati per la trasformazione di corrente alternata in corrente continua, resta comunque visibile dai grafici, che il segnale ottenuto in uscita dai circuiti è periodico e non assimilabile ad un segnale di tensione continua.

Quello che vogliamo ottenere ora è estrarre il valor medio della tensione dal segnale periodico in uscita. Attraverso le serie di Fourier possiamo ricostruire una qualunque funzione periodica attraverso una combinazione lineare di toni sinusoidali a frequenza decrescente.

In particolare ricordiamo che tra le armoniche ottenute dalla serie di Fourier, l'armonica con frequenza di $0Hz$ rappresenta il valor medio del segnale. Il nostro obiettivo diventa quindi quello di isolare la componente continua. Possiamo fare ciò attraverso un filtro passa-basso capace di fare passare le componenti a frequenza più bassa, filtrando quelle a frequenza più alta.

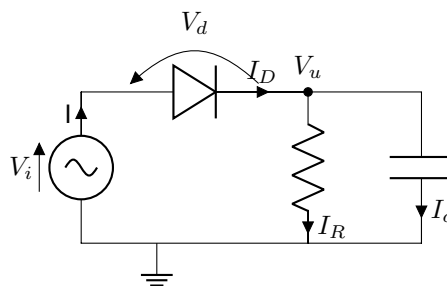


Figure 2: Raddrizzatore a singola semionda con filtro passa-basso

Equazioni Generali

$$I_D = I_R + I_C$$

$$I_R = \frac{V_u}{R}$$

$$I_C = C \frac{dV_u}{dt}$$

Nel caso del diodo acceso, il termine $\frac{V_u}{R}$, è positivo se V_u è positivo, quindi fino a quando siamo nel 1° o 2° quadrante. Mentre il secondo termine, essendo derivata di V_u è negativa in caso di segnale decrescente. Quindi il diodo è sicuramente acceso nell'intervallo $[0; \pi/2]$, mentre è sicuramente spento nell'intervallo $[\pi; \frac{3}{4}\pi]$.

Indichiamo con ωt_{off} il punto appartenente a $]\pi/2; \pi[$ in cui il diodo passa dallo stato ON allo stato OFF. Il punto ωt_{off} è facilmente calcolabile osservando che corrisponde al punto di spegnimento, quindi dal passaggio di $I_D > 0$ a $I_D = 0$.

$$\left. \begin{array}{l} V_d = V_\gamma \\ V_i - V_d - V_u = 0 \\ V_i = V_M \sin(\omega t) \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} V_u = V_M \sin(\omega t) - V_\gamma \text{ trascurabile} \\ \frac{dV_u}{dt} = V_M \cos(\omega t) \omega \end{array} \right.$$

$$I_D = \frac{V_M \sin(\omega t)}{R} + C V_M \omega \cos(\omega t)$$

$$\sin(\omega t_{\text{off}}) + C R \omega \cos(\omega t_{\text{off}}) = 0$$

Siccome il punto che cerchiamo appartiene all'intervallo $]\pi/2; \pi[$, possiamo tranquillamente dire che $\omega t_{\text{off}} = \arctan(-C R \omega)$, tenendo presente di prendere la soluzione in tale intervallo. Osserviamo inoltre il fatto che per $\omega R C$ crescente, il punto ωt_{off} tende a $\pi/2$.

Per il caso di diodo basta risolvere l'equazione differenziale $\frac{V_u}{R} + C \frac{dV_u}{dt} = 0$, ottenendo:

Diodo ON

$$\left\{ \begin{array}{l} \tan(\omega t) > -C R \omega \\ V_u = V_M \sin(\omega t) - V_\gamma \text{ trascurabile} \end{array} \right.$$

Diodo OFF

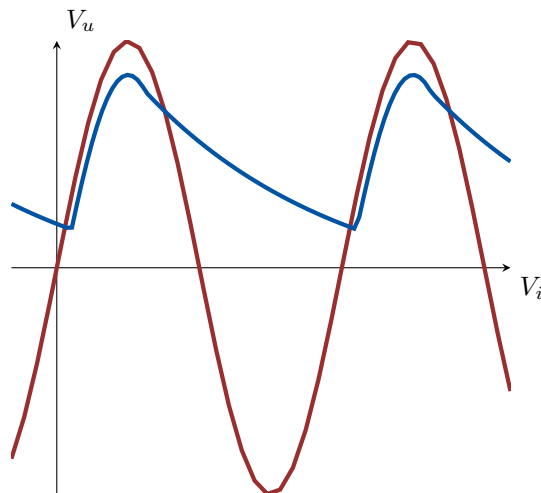
$$\left\{ \begin{array}{l} V_u(t) = V_u(t_{\text{off}}) e^{-\frac{1}{RC}(t-t_{\text{off}})} \\ V_u > V_M \sin(\omega t) - V_\gamma \text{ trascurabile} \end{array} \right.$$

Quindi al momento di spegnimento del diodo, l'andamento della tensione decade seguendo l'andamento di un esponenziale negativo.

L'effetto della capacità diventa quindi evidente, La presenza della capacità fa sì che il diodo si spenga prima, tanto prima quanto più elevata la capacità, ed una volta che si è spento la tensione non segue più la sinusoide ma un esponenziale decrescente con costante di tempo dipendente da RC. Maggiore è il prodotto RC, minore è il decadimento.

L'idea di partenza era quella di trasformare la tensione alternata in tensione continua, e per RC sufficientemente grande, siamo in grado di approssimare un generatore di tensione continua.

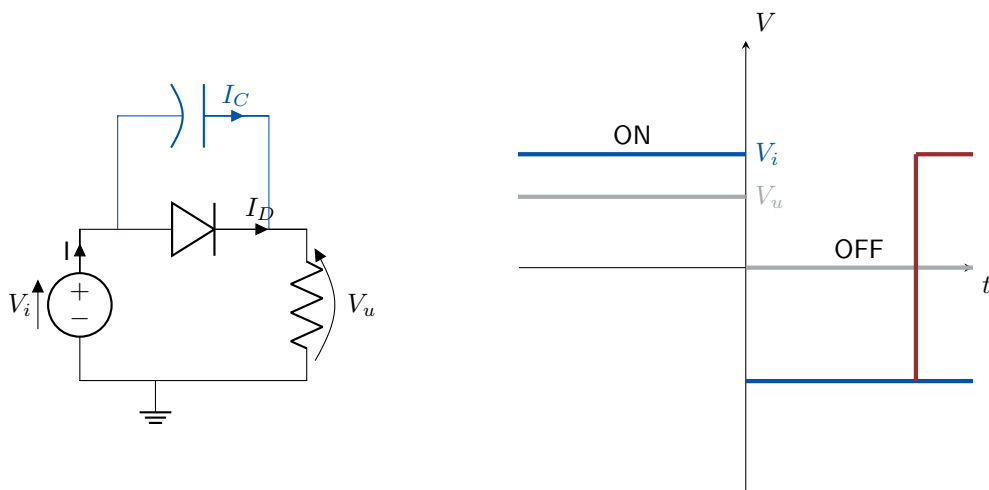
Questo circuito può essere utilizzato anche come un demodulatore di ampiezza, portando l'uscita a seguire l'andamento dell'ampiezza della sinusoide in ingresso.



Diodi in regime dinamico

Fin'ora abbiamo tracciato delle caratteristiche di trasferimento ingresso-uscita, presupponendo che in ogni istante fosse possibile determinare la tensione d'uscita in funzione di quella in ingresso indipendentemente dal tempo, come se fosse una relazione statica. Ma nel diodo, per passare dalla condizione di polarizzazione diretta, alla condizione di polarizzazione inversa, è necessario che avvenga uno spostamento di carica all'interno del dispositivo, ed è impossibile muovere elettrone in tempi nulli. Quindi a questo spostamento, che corrisponde alla commutazione di polarizzazione diretta ed inversa, è associato un ritardo.

Vogliamo stimare questo ritardo per capire se i risultati sino ad ora calcolati sono effettivamente realistici o devono essere a loro volta corretti alla luce di questa considerazione.



In grigio riportata la risposta statica del circuito in risposta al segnale in ingresso V_i .

Per calcolare il ritardo dobbiamo introdurre un modello nel diodo, che, in aggiunta a quanto descritto fino ad ora, tenga conto anche degli effetti di reazione appena introdotti. Introduciamo in parallelo al diodo una capacità, che ci permette di simulare il comportamento del diodo in regime dinamico. La relazione associata alla carica di questo condensatore parassita non è descritta dalla relazione

$Q = CV$, ma da una relazione non lineare, dove nel primo quadrante segue la relazione esponenziale $Q = Q_s(e^{V_d/V_t} - 1)$. È facile osservare che il rapporto $Q/I_D = Q_s/I_s$ è una costante delle dimensioni fisiche di un tempo, che indicheremo genericamente con la costante τ . Sintetizziamo quindi la regione di funzionamento della capacità in funzione diretta, come $Q = \tau I$.

La stessa relazione non vale in polarizzazione inversa, dove la carica non segue un asintoto, ma ha un'andamento simile a quello di una radice quadrata.

Dato che abbiamo approssimato l'andamento della corrente con un andamento lineare a tratti, faremo lo stesso con la carica. È del tutto evidente che la qualità di questa approssimazione è meno buona ma ci accontenteremo, mettendo in evidenza i momenti in cui questa approssimazione risulterà inadeguata.

Diodo OFF

$$\begin{cases} I_D = 0 \\ Q = 0 \\ V_d < V_\gamma \end{cases}$$

Diodo ON

$$\begin{cases} V_D = V_\gamma \\ I_D > 0 \\ Q > 0 \end{cases}$$

Ed in entrambe è valida $Q = \tau I$.

Per $t < 0$, siccome arriviamo da tempo $t = -\infty$, il circuito è ancora in regione statica, per questo la derivata della corrente è nulla e valgono ancora le equazioni del diodo in regione statica.

Stesso ragionamento è valido se attendo un tempo sufficientemente lungo, dove ogni fenomeno transitorio tenderà ad esaurirsi.

Equazioni generali

$$V_i - V_d - V_u = 0$$

$$I_d + I_c = I$$

$$I_c = \frac{dQ_u}{dt}$$

$$I_d = \frac{Q}{\tau}$$

Circuito a diodo acceso

$$\begin{cases} V_i = V_f > V_\gamma \\ V_u = V_f - V_\gamma \\ I = \frac{V_f - V_\gamma}{R} \\ Q = \frac{\tau}{R}(V_f - V_\gamma) \end{cases}$$

Circuito da $t > 0$ a $t \rightarrow \infty$

$$\begin{cases} V_u = 0 \\ I_c = 0 \\ I = 0 \\ Q = 0 \end{cases}$$

Condizioni iniziali $t = 0^+$

$$Q(0^+) = Q(0^-) = \frac{\tau}{R}(V_f - V_\gamma)$$

$$V_d(0^+) = V_d(0^-) = V_\gamma$$

$$V_u(0^+) = -V_R - V_\gamma$$

$$I(0^+) = \frac{-V_R - V_\gamma}{R}$$

Circuito a $t > 0$

$$\frac{dQ}{dt} + \frac{Q}{\tau} = -\frac{V_R + V_\gamma}{R}$$

da cui

$$Q(t) = \frac{\tau}{R}(V_f + V_R)e^{-\frac{t}{\tau}} - \frac{\tau}{R}(V_R + V_\gamma)$$

Analizzando l'equazione del transitorio di $Q(t)$ ottenuta, possiamo osservare che per $t = 0$, allora $Q(t) = \frac{\tau}{R}(V_f - V_\gamma)$, che corrisponde a $Q(0^+)$. Mentre per $t \rightarrow \infty$, otteniamo $-\frac{\tau}{R}(V_R + V_\gamma)$, che, essendo $V_R + V_\gamma$ positivo $Q(t) < 0$ per $t \rightarrow \infty$.

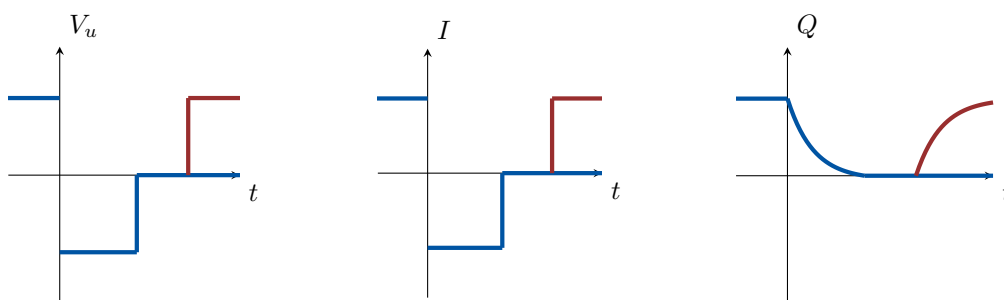
L'equazione $V_d = V_\gamma$, ipotesi utilizzata per il transitorio, è valida fino a quando il valore di Q è positivo, maggiore o al limite 0.

L'equazione del transitorio è valida fino a quando l'equazione del diodo una costante e pari a V_γ , $Q(t) > 0$ Il valore finale di $Q(t)$ può essere al minimo 0,

Il diodo, dovendo smaltire la carica positiva sulla giunzione, per un certo periodo di tempo tiene costante la tensione ai suoi capi, quindi ritarda a spegnersi. Chiamiamo questo tempo: tempo di storage. Per calcolare questo tempo dobbiamo trovare il punto in cui l'equazione della carica raggiunge il valore 0.

$$\begin{aligned} \frac{\tau}{R}(V_f + V_R)e^{-t_s/\tau} - \frac{\tau}{R}(V_R + V_\gamma) &= 0 \\ t_s &= \tau \ln \left(\frac{V_f + V_R}{V_R + V_\gamma} \right) \end{aligned}$$

Si vede subito che il rapporto è maggiore di 1, e che il tempo di storage, non dipende dal valore della resistenza, ma dipende unicamente dalla escursione di tensione e dalle caratteristiche fisiche τ del diodo.



Nel modello approssimato che stiamo utilizzando si impiega un tempo t_s per passare da $Q = Q(0^-)$, a $Q = 0$ perché abbiamo dovuto spostare le cariche interne al diodo fino ad arrivare al valore 0. Dato che nel modello approssimato che stiamo utilizzando non c'è ulteriore carica da spostare, il tempo del transitorio è necessariamente nullo, quindi la tensione è libera di variare istantaneamente.

Abbiamo già commentato che la qualità dell'approssimazione della carica è inferiore a quella della corrente, dato che in regione di polarizzazione inversa, una si comporta come radice e l'altra come esponenziale negativo che tende ad un valore di I_s pressoché 0.

Questo significa che l'approssimazione considera nulla la carica in polarizzazione inversa non è accuratissima, perché la carica essendo anche debolmente negativa, comporta un transitorio di scarica del condensatore, dove ci dobbiamo aspettare che il valore della resistenza R influirà sul tempo di scarica.

Abbiamo descritto il transitorio di spegnimento del diodo, descrivendolo in due tratti: il primo caratterizzato da una tensione costante $V_d = V_\gamma$ ed una carica che varia da un valore positivo ad un valore nullo in un tempo t_s , ed un secondo tratto dove la carica è costante e vale 0, e la tensione $V_d < V_\gamma$ che si compie in un tempo nullo.

Per determinare come funziona il passaggio di accensione del diodo, dovremo percorrere questi due tratti all'inverso, entrambi descritti dalle stesse equazioni. Dato che percorrendo il secondo tratto impieghiamo un tempo nullo, visto che non comporta spostamento di carica, nel primo tratto avremo ai capi del diodo una tensione costante pari a $V_f - V_\gamma$. Ma siccome non appena il condensatore comincia a caricarsi la tensione ai capi del diodo ha assunto già il suo valore definitivo, il transitorio inverso ha un comportamento totalmente differente. Mentre per spegnere il diodo è necessario un tempo di spegnimento t_s , per accendere il diodo (in termini di tensione in uscita) il tempo di accensione non è visibile.

Il transitorio del diodo è asimmetrico, è lento a spegnersi e molto rapido ad accendersi.

Abbiamo stimato che il tempo di spegnimento del diodo è di circa $20ns$, e questo tempo è pressoché indipendente dalla resistenza.

Le approssimazioni fatte sino ad ora sono del tutto accettabili quando il periodo del segnale è maggiore come ordine di grandezza al ritardo intrinseco del diodo. Diventa non trascurabile quando le frequenze del segnale sono notevolmente maggiori rispetto al ritardo. Al crescere della frequenza il capacitore tende a cortocircuitare il diodo.

Esiste quindi un limite alla frequenza massima di commutazione che posso richiedere ad un diodo. Questo è uno dei motivi per cui la frequenza di clock è limitata ad un valore massimo e che il periodo associato a quella frequenza deve essere sufficientemente lungo per permettere ai transienti del circuito di completarsi.

Transistore bipolare a giunzione BJT

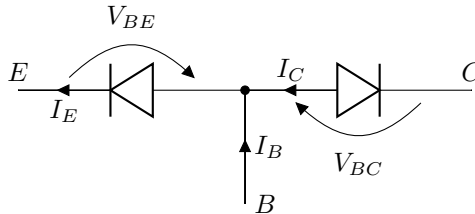


Figure 3: transistor bipolare

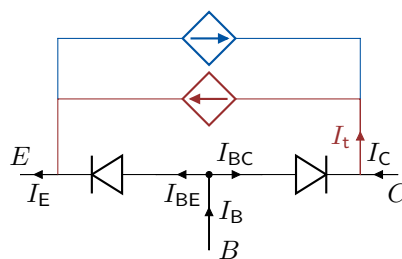
Modello di Ebers e Moll

Consideriamo $V_{BE} > 0$ e $V_{BC} < 0$. Di conseguenza il diodo tra base ed emettitore è polarizzato in regione diretta ed il diodo tra base e collettore è polarizzato in inversa, quindi spento. Analizzando il comportamento dei diodi ci aspettiamo che:

$$\begin{cases} I_E > 0 \\ I_C \approx 0 \\ I_B + I_C = I_E \end{cases} \Rightarrow I_B \approx I_E$$

Ma quando la distanza tra i due diodi è molto piccola entra in gioco l'effetto transistor, cambiando il funzionamento del circuito.

Effetto Transistore



Quando la distanza tra i due diodi (indicata con w) è ridotta, l'interazione tra le cariche cambia il comportamento del circuito; facendo valere le relazioni:

$$\begin{aligned} I_E &\approx -I_C \\ I_B &\approx 0 \end{aligned}$$

Nel momento in cui il diodo tra base ed emettitore è polarizzato in diretta, ed il diodo tra base e collettore è polarizzato in inversa, la corrente fluisce prevalentemente fra collettore ed emettitore a fronte di una corrente di base molto più piccola rispetto alle altre due regioni.

Possiamo simulare questa corrente tra collettore ed emettitore attraverso un generatore di corrente fittizio (rosso) che aggiunga alla corrente trascurabile descritta dal modello del diodo, una corrente dipendente dalla tensione V_{BE}

Questa nuova corrente prende I_T è detta *corrente dell'effetto transistor*.

Analogamente quando la polarizzazione dei diodi è invertita, per lo stesso principio si manifesta una corrente dal emettitore a collettore. Indichiamo questa seconda corrente dell'effetto transistor con un secondo generatore di corrente (riportato in blu).

Equazioni caratteristiche transistor

$$\begin{aligned} I_{BE} &= I_{BES} (e^{V_{BE}/V_T} - 1) \\ I_{BC} &= I_{BCS} (e^{V_{BC}/V_T} - 1) \\ I_t &= I_S (e^{V_{BE}/V_T} - 1) - I_S (e^{V_{BC}/V_T} - 1) \end{aligned}$$

Bisogna considerare che abbiamo utilizzato solo una particolare condizione di polarizzazione. Se studiasse il caso opposto, ovviamente otterremmo risultati simmetrici: vedremo che con $V_{BE} < 0$ e $V_{BC} > 0$ avremo una componente aggiuntiva di corrente ad effetto transistor, diretta in direzione opposta, rappresentabile anch'essa con un generatore di corrente pilotato.

Equazioni caratteristiche transistor (lineari. dipendenti)

$$\begin{aligned} I_B &= I_{BES} (e^{V_{BE}/V_T} - 1) - I_{BCS} (e^{V_{BC}/V_T} - 1) \\ I_E &= (I_S + I_{BES}) (e^{V_{BE}/V_T} - 1) - I_S (e^{V_{BC}/V_T} - 1) \\ I_C &= I_S (e^{V_{BE}/V_T} - 1) - (I_S + I_{BCS}) (e^{V_{BC}/V_T} - 1) \end{aligned}$$

Il transistor in figura 4 prevede due giunzioni ed ha il nome di **transistor npn**. Esiste anche il suo duale, **transistor pnp** al quale faremo solo un rapido cenno più avanti ma suo comportamento è del tutto identico a quello che stiamo discutendo.

Per la legge di Kirchhoff se conosciamo due differenze di potenziale ai lati del transistor, la terza è univocamente determinata. Lo stesso vale per le correnti.

Per determinare completamente il regime di funzionamento del transistor, occorre determinare le 6 grandezze: 3 correnti e 3 tensioni, attraverso 6 equazioni, due delle quali sono quelle di Kirchhoff appena indicate. Per trovare le altre quattro equazioni, utilizziamo lo stesso metodo che abbiamo applicato per determinare le equazioni del diodo: mettiamo un morsetto a terra e, fornendo un potenziale su uno dei

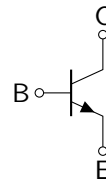


Figure 4: transistor npn

due morsetti rimanenti, misuriamo il potenziale sull'ultimo morsetto.

Ovviamente è possibile connettere il transistor in 3 modi differenti: emettitore, base e collettore comune. Ma di queste ultime due ce ne occuperemo più avanti.

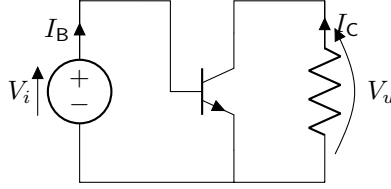


Figure 5: Connessione a Emettitore comune

Già osservando il circuito, possiamo notare che due incognite sono eliminate dalla equazione della tensione in ingresso $V_i = V_{BE}$, che possiamo considerare data, e dall'equazione $V_u = R \cdot I_C$. Ricordando che la corrente in ingresso corrisponde alla corrente di base, e che la corrente in uscita corrisponde alla corrente di collettore, per le ultime due equazioni residue possiamo utilizzare quelle fornite dalle caratteristiche del transistor elencate precedentemente.

$$I_B(V_{BE}, V_{CE}) = I_{BES}(e^{V_{BE}/V_T} - 1) - I_{BC}(e^{V_{BC}/V_T} - 1)$$

$$I_C(V_{BE}, V_{CE}) = I_S(e^{V_{BE}/V_T} - 1) - (I_S + I_{BCS})(e^{V_{BC}/V_T} - 1)$$

Siccome queste relazioni di correnti sono due funzioni di due variabili (V_{BE} e V_{CE}), è necessario un grafico in tre dimensioni per poterle rappresentare graficamente, e ciò non sarebbe pratico.

Per questo motivo riconduciamo queste espressioni ad una rappresentazione più semplice, riconducendoci ad una famiglia di curve parametriche ponendo V_{CE} come variabile indipendente e tracciando le funzioni al variare di V_{CE} (Figura 6).

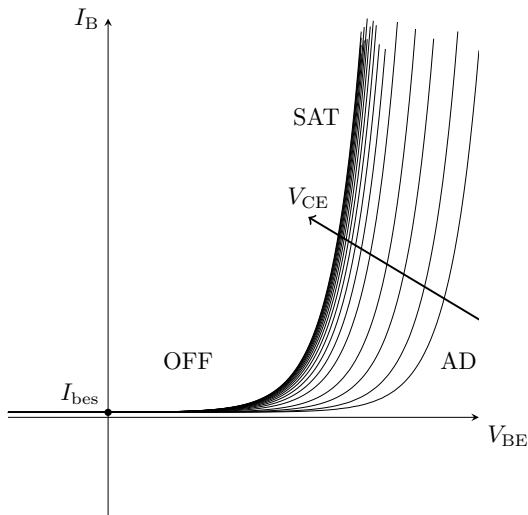


Figure 6: $I_B(V_{BE})$ al variare di V_{CE}

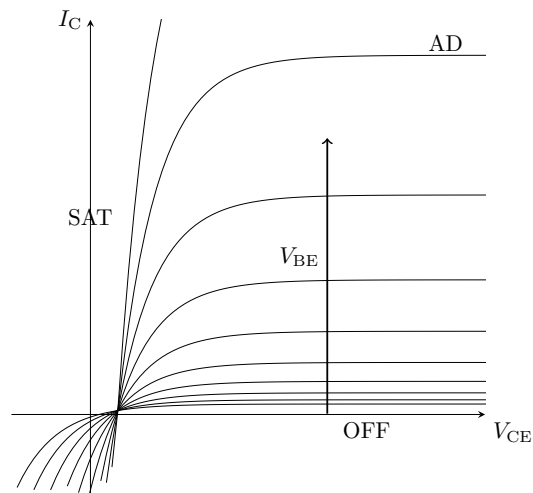


Figure 7: $I_C(V_{CE})$ al variare di V_{BE}

Il transistor è composto da 2 diodi, ognuno polarizzabile in regione diretta o regione inversa. Possiamo riconoscere quindi quattro regioni di funzionamento del transistor, dipendenti da V_{BE} e V_{BC} .

- BE on, BC off: Regione normale di funzionamento o regione di polarizzazione attiva diretta
- BE off, BC on: Regione di polarizzazione attiva inversa.
- BE off, BC off: Regione di interdizione, per brevità diremo che il transistor è spento
- BE on, BC on: Regione di saturazione.

Regione Attiva Diretta

Prendendo in considerazione la regione di funzionamento attiva diretta, considerando che $e^{V_{BE}/V_T} > e^{V_{BC}/V_T}$ siccome il primo termine è maggiore di 1 per $V_{BE} > 0$ ed il secondo minore di 1 per $V_{BC} < 0$, possiamo semplificare le relative equazioni caratteristiche del transistor in

$$\begin{aligned} I_C &= I_s (e^{V_{BE}/V_T} - 1) \\ I_B &= I_{BEs} (e^{V_{BE}/V_T} - 1) \\ I_E &= (I_s + I_{BEs}) (e^{V_{BE}/V_T} - 1) \end{aligned}$$

Osserviamo che in questa regione, tutte le correnti assumono la forma di esponenziale traslata in funzione della sola tensione V_{BE} . In questo caso particolare, il circuito in figura 5, la corrente in uscita I_C non è una funzione della tensione di uscita $V_u = V_{CE}$, quindi si comporta come un circuito generatore di corrente costante/controllata in funzione di V_i .

Inoltre, siccome tutte le correnti dipendono dallo stesso esponenziale, allora sono proporzionali tra loro

$$\frac{I_C}{I_B} = \frac{I_s}{I_{BEs}} = \beta_F$$

Per analogo ragionamento $I_E = (\beta_F + 1)I_B$, e $I_C = \frac{\beta_F}{\beta_F + 1}I_E = \alpha_F I_E$. Con α_F viene definita l'*efficienza di emettitore*: maggiore è la costante più la corrente di collettore assomiglia alla corrente di emettitore. Quindi siccome $\alpha_F = \frac{\beta_F}{\beta_F + 1}$, maggiore è β_F , più α_F è vicina ad 1, e migliore è il transistor.

Regione attiva inversa

Per ragionamento analogo alla regione precedente, $V_{BE} < 0$ e $V_{BC} > 0$ ci porta a trascurare i termini in funzione di V_{BE}

$$\begin{aligned} I_C &= -(I_S + I_{BCs})(e^{V_{BC}/V_T} - 1) \\ I_B &= I_{BCs}(e^{V_{BC}/V_T} - 1) \\ I_E &= -I_S(e^{V_{BC}/V_T} - 1) \end{aligned}$$

Indicando con $\beta_R = I_S/I_{BCs}$, abbiamo che

$$\begin{aligned} I_E &= -\beta_R I_B \\ I_C &= -(\beta_R + 1)I_B \\ I_E &= \frac{\beta_R}{1 + \beta_R} I_C = \alpha_R I_C \end{aligned}$$

Il crescere di β_R aumenta le prestazioni del transistor: $\alpha_R \approx 1$.

Nota: F e B a pedice, indicano *Forward* e *Reverse*

Regione di interdizione

$V_{BE} < 0$ e $V_{BC} < 0$ ci portano ad osservare che ciascuna delle espressioni esponenziali è trascurabile, quindi:

$$\begin{aligned} I_C &= I_{BCs} \\ I_B &= -I_{BEs} - I_{BCs} \\ I_E &= I_{BEs} \end{aligned}$$

Quindi in questa regione il modello non necessita di generatori controllati per essere descritto.

Regione di saturazione

$V_{BE} > 0$ e $V_{BC} > 0$ indica che nessun esponenziale è trascurabile. Tutte le correnti dipendono da entrambe le tensioni di funzionamento V_{BE} e V_{BC} .

Amplificatore invertente di tensione per piccoli segnali

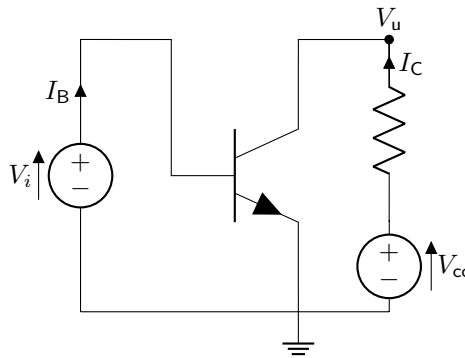
Utilizzando le approssimazioni appena ricavate nel circuito in figura 5

Regione attiva diretta

$$V_u = V_{BE} - V_{BC} > 0$$

$$V_u = -RI_C = -RI_s(e^{V_{BE}/V_T} - 1) < 0$$

Quindi il circuito non può funzionare in regione normale. Modifichiamo quindi il circuito aggiungendo un generatore di tensione V_{cc} tale che $V_u = V_{cc} - RI_C > 0$. Mantenendo compatibilità con la prima ipotesi.



Siccome ci siamo assicurati che il circuito funzioni in regione attiva diretta, verifichiamo ora se è possibile che tale circuito funzioni in regione attiva inversa:

Regione attiva inversa

$$V_u = V_{BE} - V_{BC} < 0$$

$$V_u = V_{cc} - RI_C > 0$$

Abbiamo dimostrato quindi che questo circuito, corretto appositamente per farlo lavorare in regione attiva diretta, non può lavorare in regione attiva inversa. Per funzionare in quest'ultima regione il generatore di tensione V_{cc} dovrebbe avere una tensione negativa, per questo motivo per analizzare di questo circuito, consideriamo solo tre regioni di funzionamento.

Regione di interdizione

$$I_C = I_{BEs} \approx 0$$

$$V_u = V_{cc}$$

$$V_{BE} = V_i < 0$$

$$V_{BC} = V_i - V_u < 0$$

Attiva diretta

$$V_{BE} = V_i > 0$$

$$V_u = V_{cc} - RI_s(e^{V_i/V_T} - 1)$$

In regione di saturazione abbiamo che

$$I_C = \frac{V_{CC} - V_{CE}}{R}$$

Ricordandoci che in figura 2.2 abbiamo già calcolato una relazione che lega I_C e V_{CE} . Tracciamo l'equazione di I_C appena trovata. Ciascuno dei punti di intersezione in tale figura, rappresenta quindi il luogo dei punti soluzione di questa equazione. Dal grafico quindi possiamo osservare che in regione di saturazione, la tensione di uscita continua a calare, ma tende asintoticamente ad un valore appena maggiore di 0.

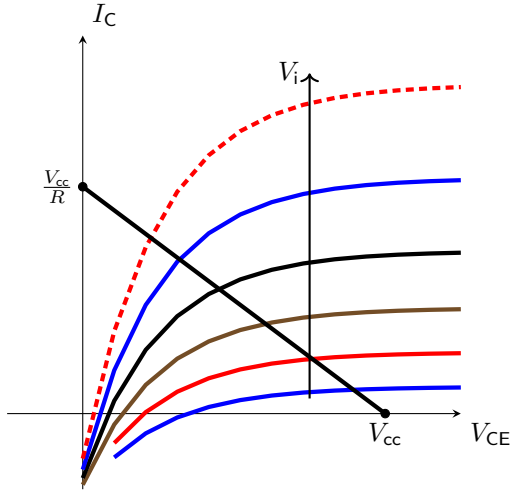


Figure 8: Grafico con ramo di carico

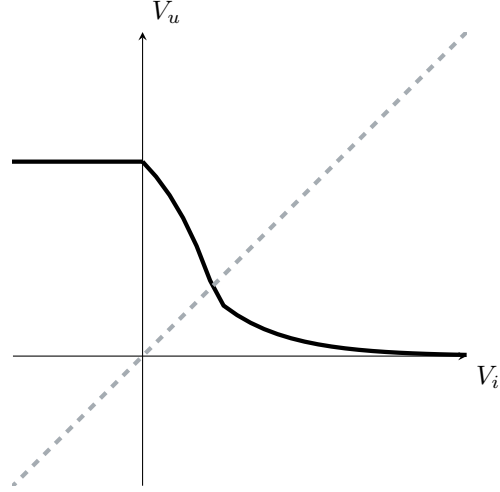


Figure 9: Andamento qualitativo di V_u in funzione di V_i

In particolare, il valore a cui asintoticamente tende la tensione in uscita è dato da

$$I_C = I_S(e^{V_{BE}/V_T} - 1) - (I_S + I_{BCS})(e^{V_{BC}/V_T} - 1) = 0$$

Dove entrambi i termini 1 sono trascurabili, siccome entrambi gli esponenziali sono maggiori di 1.

$$V_{CE} = V_T \ln \frac{1}{\alpha_R}$$

Quindi V_{CE} è strettamente positivo, dato che α_R è compreso tra 0 ed 1, quindi il suo reciproco è maggiore di 1, e rispettivo logaritmo è positivo. Inoltre dato che questa espressione non dipende da V_{BE} , significa che tutte le caratteristiche intersecano l'asse delle ascisse in corrispondenza di tale valore.

Considerando ora un punto (V_{i0}, V_{u0}) appartenente al tratto di polarizzazione attiva diretta, la retta tangente al grafico in tale punto ha equazione $V_u - V_{u0} = m(V_i - V_{i0})$, il termine m , uguale alla derivata di V_u rispetto a V_i , calcolata in (V_{i0}) , è chiamato "guadagno di tensione" ed è indicato dal simbolo A_V .

Dato che la corrente $I_C(V_{i0}) = I_{c0} = I_S(e^{V_{i0}/V_T} - 1)$, e siamo in regione attiva diretta (quindi il termine 1 è trascurabile rispetto all'esponenziale), possiamo dire che $I_{c0} = I_S e^{V_{i0}/V_T}$. Quindi se calcolando il guadagno di tensione in V_{i0} otteniamo:

$$V_u = V_{CC} - R I_S (e^{V_i/V_T} - 1)$$

$$\left. \frac{dV_u}{dV_i} \right|_{V_{i0}} = -\frac{R I_S}{V_T} e^{V_{i0}/V_T} = -\frac{R}{V_T} I_{c0}$$

Per dare una stima a questo rapporto, prendiamo ad esempio un punto intermedio, in cui $V_{u0} \approx 2.5V$. Mentre la tensione termica V_T che compare a denominatore è dell'ordine di grandezza di $25mV$, quindi $A_V \approx -100$. Questo valore significa che in regione attiva diretta, le piccole variazioni di un segnale in ingresso, vengono amplificate di un fattore $A_V = 100$. Comportamento tipico di un amplificatore di tensione per piccoli segnali. Ovviamente, se il segnale di ingresso raggiunge la regione di saturazione, il segnale d'uscita non sarà più sinusoidale.

Facendo riferimento ad un segnale binario, il quale può assumere solamente i valori V_H molto grande e V_L prossimo a zero, allora il circuito si comporta come un invertitore.

Approssimazione Transistor BJT

Esattamente come abbiamo fatto nel caso del diodo, approssimiamo nello stesso modo le caratteristiche del transistor bipolare. Osservando i grafici in figura 6 e 7, possiamo formulare un modello lineare, valido nelle tre regioni di funzionamento, trascurando la regione inversa. Siccome dai calcoli svolti in precedenza, sappiamo che il grafico in figura 7 non passa per lo zero, chiamiamo tale punto $V_{CE_{SAT}} = 0.2V$.

Inoltre, siccome in saturazione entrambi i diodi sono in polarizzazione diretta

$$V_{CE} = V_{BE_{on}} - V_{BC_{on}} = V_{CE_{SAT}}$$

Sapendo già che in polarizzazione diretta $V_{BE_{on}} = V_\gamma$, possiamo dire che $V_{BC_{on}} = 0.55V = V'_\gamma$, quindi le tensioni base-emettitore e base-collettore in polarizzazione diretta, sono diverse tra di loro.

OFF $\begin{cases} I_B = 0, & V_{BE} < V_\gamma \\ I_C = 0, & V_{BC} < V'_\gamma \end{cases}$	AD $\begin{cases} I_B > 0, & V_{BE} = V_\gamma \\ I_C = \beta_F I_B > 0, & V_{CE} > V_{CE_{SAT}} \end{cases}$
SAT $\begin{cases} I_B > 0, & V_{BE} = V_\gamma \\ I_C < \beta_F I_B, & V_{CE} = V_{CE_{SAT}} \end{cases}$	

In questo caso, siccome abbiamo a che fare con due giunzioni, abbiamo bisogno di definire due correnti, e per ciascuna delle regioni abbiamo due disequazioni che descrivono la validità delle due equazioni.

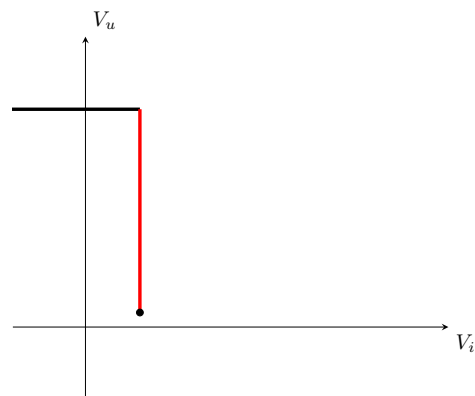
Studio del circuito utilizzando il modello approssimato

OFF $\begin{cases} V_u = V_{CC} \\ V_i < V_\gamma \end{cases}$	AD $\begin{cases} V_{CE_{SAT}} < V_u < V_{CC} \\ V_i = V_\gamma \end{cases}$	SAT $\begin{cases} V_i = V_\gamma \\ V_u = V_{CE_{SAT}} \end{cases}$
---	---	---

Questo modello rappresenta il tratto a pendenza elevata, con un tratto a pendenza infinita (tratto verticale). Se stiamo progettando un' amplificatore analogico, il quale si basa sul determinare il guadagno del circuito (A_V), non è una buona approssimazione.

Inoltre il modello rappresenta tutta la regione di saturazione con un unico punto, di coordinate $(V_\gamma, V_{CE_{SAT}})$ quindi è un'approssimazione non accettabile.

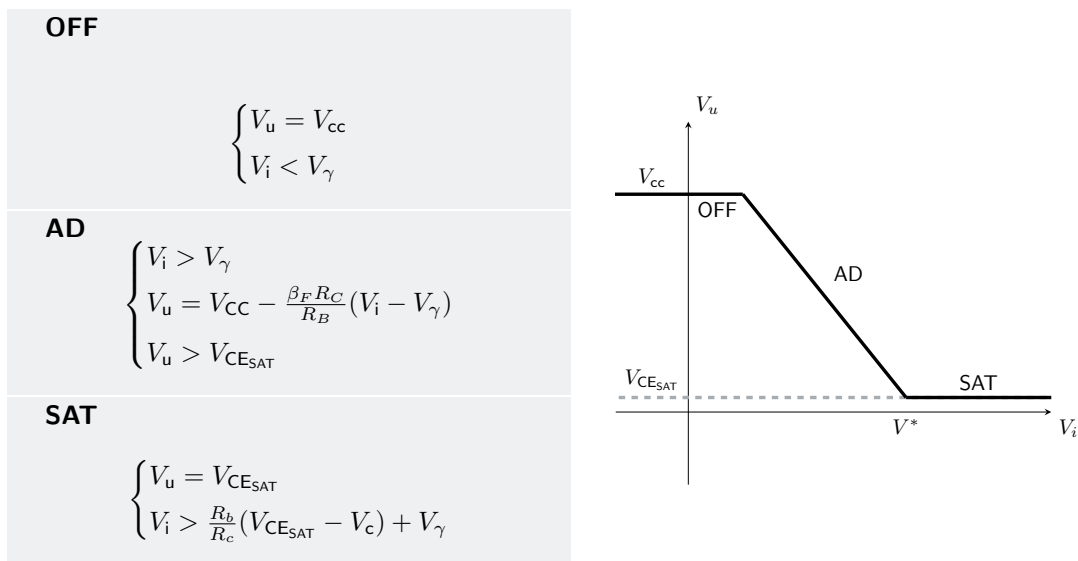
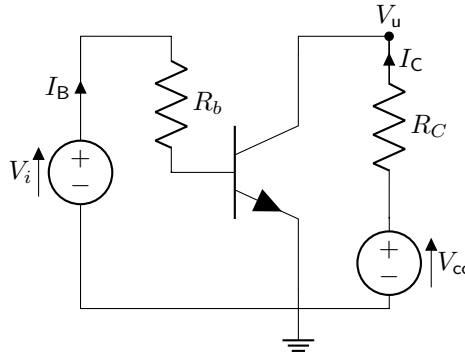
La spiegazione di questo fenomeno è che, siccome



$\lim_{V_i \rightarrow \infty} I_C = \infty$ non esiste un asintoto verticale, mentre noi stiamo approssimando tutte le possibili relazioni corrente-tensione con un'unica retta verticale.

Se osserviamo per quali ordini di grandezza di corrente, questo modello non rappresenta accuratamente il valore della tensione in uscita, sono dell'ordine dei giga ampere. Il circuito visto ora, non è realistico perché non ci sono limitazioni per i valori di corrente I_B , aggiungiamo quindi una resistenza alla base, per limitare la corrente, e rianalizziamo il circuito con il modello lineare.

Analisi dello stesso circuito con corrente limitata



Il modello a soglia descrive con un'ottima approssimazione i casi in cui la corrente si mantiene limitata. Il guadagno in questo modello lineare è esattamente il coefficiente angolare del tratto in regione attiva diretta: $-\beta_F R_C / R_B$. Dato che dipende solamente dai valori delle due resistenze del circuito, è possibile aumentare o diminuire arbitrariamente il guadagno introducendo una resistenza variabile nel circuito.

Valutazione metodo di approssimazione

Mettendo insieme l'equazione di tensione e corrente della maglia in ingresso:

$$\begin{aligned} V_i - R_B I_B - V_{BE} &= 0 \\ I_B &= I_{BES} (e^{V_{BE}/V_T} - 1) \end{aligned}$$

ottengo $V_i = R_B I_B + V_T \ln(I_B / I_{BES} + 1)$.

Si può vedere dall'espressione come la corrente di base influenza la tensione in ingresso con un termine logaritmico ed uno lineare, seguendo l'andamento del grafico nero in figura 10. Ad $I_B > 0$ la tensione in ingresso ha un andamento pressoché lineare, quindi una variazione lineare

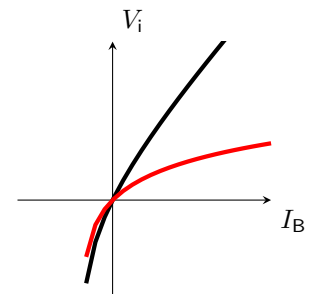


Figure 10: test

di tensione corrisponde ad una variazione lineare di corrente. Ciò significa che la corrente in ingresso è dissipata più facilmente sulla resistenza che sulla giunzione base-emettitore. Diverso è il caso precedente, rappresentato dal grafico rosso, dove per mancanza di resistenza a piccole variazioni della tensione in ingresso corrispondono grandi variazioni di corrente, e siccome manca la resistenza in ingresso, per raggiungere la stessa quantità di corrente è necessario raggiungere valori di tensione molto più elevati.

Riassumendo: Quando la giunzione è in serie ad una resistenza qualunque variazione in ingresso si scarica prevalentemente sulla resistenza consentendo un'approssimazione di tensione costante. Nel caso in cui manchi la resistenza, non si può pensare di applicare un modello che consideri la tensione in ingresso costante se quest'ultima è variabile per definizione.

Logica RTL: Transistor Resistor Logic

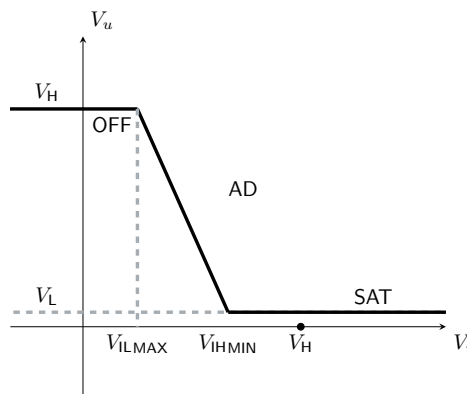
Se associamo a valori di tensione alti e bassi (V_H , V_L) una codifica logica 1 e 0, il circuito precedente si comporta come un invertitore logico.

Bisogna ricordare che anche il trasporto di carica (corrente) non è uniforme, ed è soggetto anch'esso a rumore. Il problema che ha questo circuito visto come amplificatore è che amplifica il rumore in ingresso ed il segnale in ingresso dello stesso fattore A_V . Inoltre dato che il segnale in ingresso varia di poco, la qualità del segnale d'uscita risente molto delle variazioni di rumore.

Diversamente lo stesso circuito visto come invertitore logico, anche a rumore elevato fa il suo lavoro.

L'immunità al rumore è caratterizzato dalle due caratteristiche a guadagno 0 (off e sat). Solo a fronte di valori estremamente alte di rumore, in grado di entrare nella fascia intermedia, si manifestano sul segnale in uscita.

Margine di immunità al rumore



Dal grafico sono distinguibili un valore basso $V_L \equiv V_{CE_{SAT}}$, ed un valore alto $V_H \equiv V_{CC}$. La caratteristica da invertitore è esprimibile come $V_u(V_H) = V_L$ e $V_u(V_L) = V_H$. Inoltre siccome questo circuito tollera il rumore, per rumore minore di δ_1 l'uscita V_u sarà ugualmente $V_u(V_L + \delta_1) = V_L$. Chiamiamo quindi $V_{ILMAX} = V_L + \delta_1$ il massimo valore della tensione d'ingresso che produce in uscita un V_H . Analogamente chiamo $V_{IHMIN} = V_H - \delta_2$ il massimo valore della tensione d'ingresso che produce ancora un'uscita V_L . Il margine di immunità al rumore N_M è il minimo tra δ_1 e δ_2 .

Dato che sappiamo già i valori di V_H e V_L , calcolando V_{IHMIN} come punto di intersezione tra la zona attiva diretta e di saturazione, otteniamo:

$$\delta_1 = V_{ILMAX} - V_L = 0.75 - 0.2 = 0.55V$$

$$\delta_2 = V_H - V_{IHMIN} = 5 - 1.23 = 3.77V$$

Da cui $N_M = 0.55V$. Per avere immunità al rumore è necessario che $\delta_1 = V_{ILMAX} - V_L$ e $\delta_2 =$

$V_H - V_{IHMIN}$ siano positivi sarà positiva anche la loro somma, per questo possiamo dire che

$$\frac{V_H - V_L}{V_{IHMIN} - V_{ILMAX}} > 1$$

Facendo sempre riferimento al grafico è facile capire che rapporto tra $V_H - V_L$ e $V_{IHMIN} - V_{ILMAX}$ è $|A_V|$. Per cui $|A_V| > 1$. Maggiore sarà il valore di $|A_V|$, migliore è la proprietà di filtro per il rumore di questo circuito.

Il risultato ottenuto dipende fortemente dall'approssimazione lineare utilizzata, dato il valore alto V_H ed il valore basso V_L sono stati definiti graficamente, in corrispondenza di V_{CC} e V_{CESAT} . Generalizziamo quindi questa definizione facendo riferimento ad altri circuiti.

Preso un andamento di circuito generico (es. figura 9) rimane ancora vero che esistono due regioni con guadagno basso ed una con guadagno maggiore di 1, ma la distinzione di V_H e V_L non è ovvia come nel caso precedente.

Per identificare il valore di V_L e V_H , facciamo il seguente ragionamento: considerando una serie di n invertitori, l'andamento del valore di uscita sarà identico per tutti gli invertitori di posto pari, mentre sarà invertito per tutti gli invertitori di posto dispari. Dato che gli invertitori di posto dispari dovranno corrispondere alla funzione inversa degli invertitori di posto pari, il loro grafico è uguale al grafico degli invertitori di posto pari ma specchiato rispetto alla bisettrice.

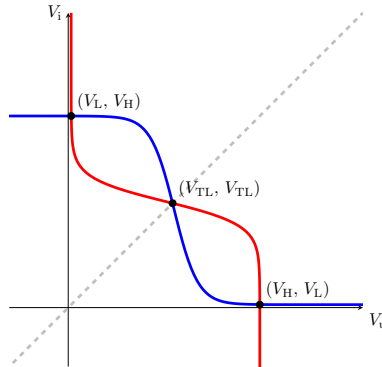


Figure 11: Grafico

Siccome è ovvio che quando tutti gli invertitori sono connessi in cascata, le loro condizioni devono essere soddisfatte, esistono solo 3 punti in queste caratteristiche che soddisfano le equazioni (vedi figura 11). Le coppie di punti (V_L, V_H) e (V_H, V_L) sono simmetrici per costruzione, per questo sono i valori V_H e V_L che stavamo cercando.

Definiamo questo punto l'escursione logica L_S (*Logic Swing*) come $V_H - V_L$.

Il punto (V_{TL}, V_{TL}) è tale per cui se posto in ingresso al primo invertitore, si propaga invariato fino al termine della catena. In realtà questa condizione è quasi impossibile, siccome ci troviamo in una zona con $|A_V| > 1$ ed al primo segnale di rumore veniamo spostati verso V_H o V_L , allontanandoci dal punto di precario equilibrio, detto anche metastabile.

Possiamo osservare che la qualità (intesa come distanza da V_H o V_L), aumenta, lungo la catena di

invertitori. Questa proprietà prende il nome di **proprietà rigenerativa del segnale**.

Per questo motivo, osservando anche in ingresso un valore compreso in un intorno di V_L , per la proprietà appena citata, esso viene trattato ugualmente come valore nominale basso. Il valore V_{TL} si comporta quindi come soglia logica, discriminando i valori alti dai valori bassi.

Questa proprietà è strettamente legata alla disuguaglianza $|A_V| > 1$. Nel caso quest'ultima non fosse verificata, osserviamo che il punto di intersezione delle due caratteristiche è uno solo (vedi figura da fare)

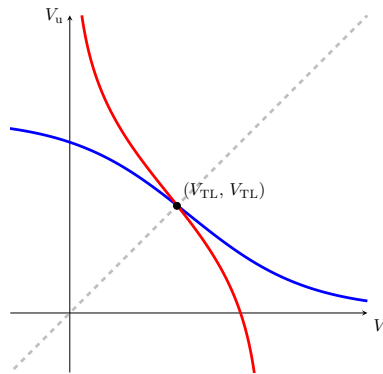


Figure 12: grafico per $|A_V| < 1$

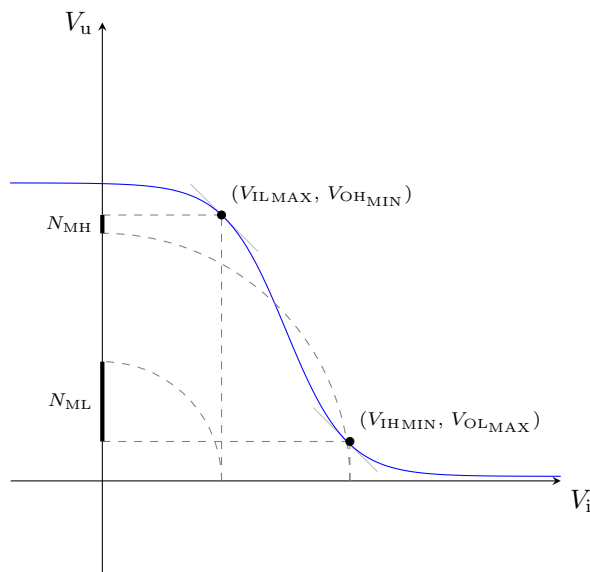
Quindi in caso $|A_V| < 1$ ogni punto in ingresso alla catena di invertitori, convergerebbe in (V_{TL}, V_{TL}) .

Margine di immunità ai disturbi

Rimane comunque vero che se il guadagno è minore di 1, il rumore in uscita è minore del rumore in ingresso, mentre se il guadagno è maggiore di 1, il rumore in ingresso è amplificato. Possiamo quindi identificare due punti, con guadagno $|A_V| = 1$, per suddividere i tratti di caratteristica, identificati da $|A_V| < 1$ e $|A_V| > 1$.

Utilizzando la stessa nomenclatura della logica RTL, chiamiamo V_{ILMAX} , il punto per cui il tratto $0 < x < V_{ILMAX}$ ha $|A_V| < 1$, e V_{IHMIN} il tratto in cui $V_{IHMIN} < x$ per cui $|A_V| < 1$. Di conseguenza $V_{ILMAX} < x < V_{IHMIN}$ è caratterizzato da $|A_V| > 1$.

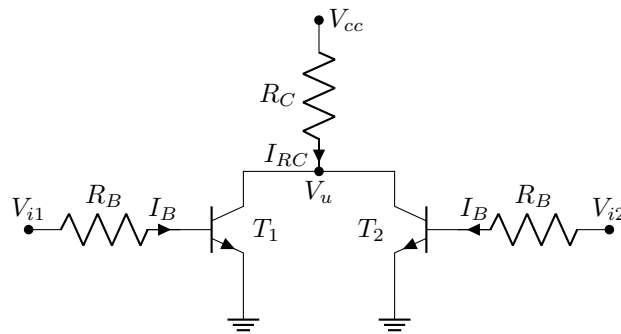
Indichiamo inoltre con $V_{OHMIN} = V_u(V_{ILMAX})$ il più piccolo valore dell'uscita associato all'attenuazione del rumore. Analogamente $V_{OLMAX} = V_u(V_{IHMIN})$.



La massima quantità di rumore accettabile per non uscire dalla regione di attenuazione sono $N_H = V_{OHMIN} - V_{IHMIN}$ e $N_L = V_{ILMAX} - V_{OLMAX}$. Quindi $N_M = \min(V_{OHMIN} - V_{IHMIN}, V_{ILMAX} - V_{OLMAX})$. Da questa costruzione risulta evidente che $L_S > N_H + N_L$, quindi nel caso in cui una delle due soglie sia controllata (ad esempio da una resistenza variabile) e dovesse aumentare di valore, l'altra sarebbe costretta a diminuire, per evitare che la somma superi L_S .

La condizione ottimale per avere massimo margine ai disturbi è quindi in caso in cui N_L e N_H siano simmetrici. (nel caso di approssimazione i punti coincidono con v_h e v_l , dato che per approssimazione non ci sono punti con derivata $= -1$, passa dal valore in modulo < 1 a modulo > 1)

Porta logica NOR



Supponendo i valori delle resistenze $R_C = 1k\Omega$ e $R_B = 10k\Omega$, e $\beta_F = 100$ vogliamo valutare il comportamento di questa rete al variare dei due ingressi V_{i1} e V_{i2} , prendendo il caso di segnali in ingresso digitali: $V_i = \{V_H, V_L\}$. Le combinazioni possibili in ingresso sono quindi enumerabili e studiabili individualmente.

Equazioni Generali

$$I_{RC} = I_{C1} + I_{C2}$$

$$V_i - R_B I_B - V_{BE} = 0$$

Nel caso di un qualsiasi transistor spento, considerando che $V_{BE} < V_\gamma$ per ipotesi, otteniamo $V_i < V_\gamma$.

Per $V_{i1} = V_{i2} = V_L$

$$V_i < V_\gamma$$

$$V_u = V_{cc}$$

Per $V_{i1} V_L, V_{i2} = V_H$

Osservando l'alta tensione ricevuta all'ingresso, è ragionevole ipotizzare che T_2 si trovi in regime di funzionamento saturo. Verificando quindi le ipotesi:

$$I_B = \frac{V_{cc} - V_\gamma}{R_B} = 0.425mA > 0$$

$$I_C = \frac{V_{cc} - V_{CE_{SAT}}}{R_C} = 4.8mA < \beta_F I_B = 42.5mA$$

Per cui in questo caso $V_u = V_{CE_{SAT}} = V_L$

Per $V_{i1} = V_{i2} = V_H$

Per ragionamento analogo al caso precedente, ci aspettiamo che entrambi i transistori siano saturi. Quindi $V_u = V_{CE_{SAT}} = V_L$

$$I_B = \frac{V_{cc} - V_\gamma}{R_B} = 0.425mA > 0$$

$$I_{RC} = \frac{V_{cc} - V_{CE_{SAT}}}{R_C} = 4.8mA < \beta_F I_B = 42.5mA$$

Essendo il circuito perfettamente simmetrico non c'è motivo di ipotizzare che $I_{C1} \neq I_{C2}$. Per cui $I_{C1} = I_{C2} = 4.8mA$, quindi sono verificate entrambe le condizioni $I_C < \beta_F I_B$.

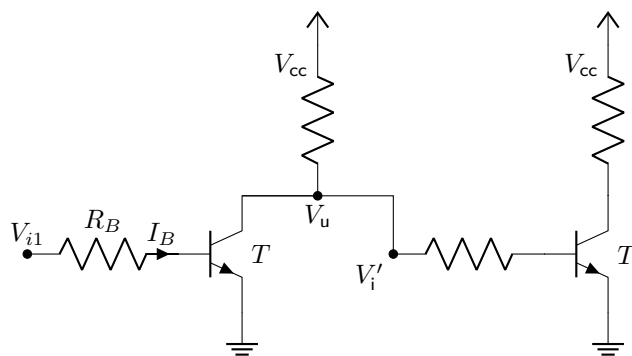
Questa rete si comporta come una porta logica NOR. La rete composta dai transistori T_1 e T_2 è definita come **rete di pulldown**, ovvero quando è accesa, trascina il valore di uscita verso il valore basso. Analogamente è possibile interpretare la rete composta dalla resistenza R_c come una **rete di pullup**: quando non è contrastata da reti di pulldown attive, l'uscita viene portata a V_H da R_c dove non circola corrente.

Osserviamo che questo circuito si comporta allo stesso modo se alla rete di pulldown è composta da più di due transistori, dato che la corrente I_{RC} sarebbe suddivisa equamente tra tutti i transistori accesi, mantenendo vera l'ipotesi $I_C < \beta_F I_B$. Quindi questo circuito è generalizzabile ad una porta NOR a numero arbitrario di ingressi.

L'insieme degli ingressi prende il nome di FAN-IN. Analogamente il numero delle uscite prende il nome di FAN-OUT.

L'operatore NOR rappresenta di per sé una famiglia funzionalmente completa. Per questo con logica RTL è possibile realizzare qualsiasi funzione combinatoria.

Manca da analizzare se la connessione in serie delle seguenti porte NOR, mantenga il valore del segnale.



Partendo ad analizzare la condizione per cui T OFF otteniamo che $V_u = V_{cc} - R_C I_B$ Nel caso in cui anche T' sia OFF, allora I_B' è 0 e di conseguenza:

$$V_u = V_{cc}$$

$$V_u = V_i' - \cancel{RI_B} = V_{BE}$$

Ma ciò non è possibile dato che se T' è OFF, allora $V'_{BE} < V_\gamma$. Quindi T' non può essere OFF se T è OFF. Quindi T' è necessariamente acceso, con $V'_{BE} = V_\gamma$. Dall'equazione di kirkoff al nodo V_u :

$$\frac{V_{cc} - V_u}{R_C} = \frac{V_u - V_\gamma}{R_B}$$

Ottenendo $V_u = \frac{R_B V_{CC} + R_C V_\gamma}{R_B + R_C} = < V_\gamma$ svolgendo la disequazione si ottiene: $V_\gamma < V_{CC}$. Sostituendo inoltre i dati utilizzati nel circuito precedente, otteniamo $V_u = 4.61V$

T in AD

$$V_u = V_{cc} - R_C \left\{ \beta_f \frac{V_i - V_\gamma}{R_B} + \frac{V_u - V_\gamma}{R_B} \right\}$$

Da cui

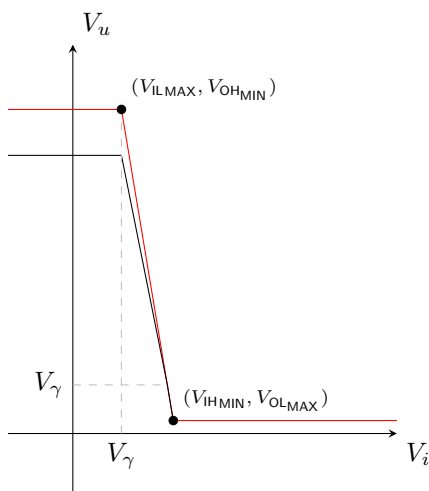
$$V_u = \frac{R_B V_{cc} + R_C V_\gamma}{R_B + R_C} - \frac{\beta_F R_C}{R_B + R_C} (V_i - V_\gamma)$$

Osserviamo come il termine blu è identico a quello calcolato in precedenza, nel tratto di regione attiva diretta, mentre il tratto rosso è un guadagno di modulo minore rispetto a quello di prima ($\frac{\beta_F R_C}{R_B}$).

Questa condizione è valida fino a quando o T passa alla regione di saturazione ($V_u = V_{CE\text{SAT}}$) o T' passa in regione attiva diretta ($V_u = V_\gamma$). Dato che $V_u < V_\gamma$ avverrà prima, sostituendo tale valore nell'equazione di V_u segue $V_i = 1.175V$

T in AD e T' OFF

Questo tratto di caratteristica corrisponde esattamente al precedente.



Delle quattro coordinate utilizzate per calcolare il margine di rumore, solo una è modificata dal fatto

che il FAN-OUT è modificato da 0 a 1. Il margine N_{ML} definito quindi per il livello basso rimane invariato: $N_{ML} = V_{ILMAX} - V_{OLMAX} = V_{\gamma} - V_{CEsat} = 0.55V$. Diversamente il margine N_{MH} varia: $N_{MH} = V_{OHMIN} - V_{IHMIN} = 4.61 - 1.23 = 3.37V$.

Il margine complessivo, rimanendo definito come il minimo tra N_{ML} e N_{MH} rimane invariato a $0.55V$.

Nel caso di un FAN-OUT generico ad n porte il valore N_{MH} tenderà a calare, dato che, la corrente del transistor di pullup è richiamata dagli altri componenti connessi.

Generalizzando con n componenti connessi che condividono V_u come tensione in ingresso, dato che tutti hanno la stessa V_i , tutti i transistori si troveranno nella stessa regione di funzionamento. Quindi le correnti in ingresso I_B necessariamente coincidono quindi $I_{RC} = I_C + nI'_B$.

Nell'ipotesi che il transistor T sia spento:

$$I_{RC} = \frac{V_{cc} - V_u}{R_C}$$

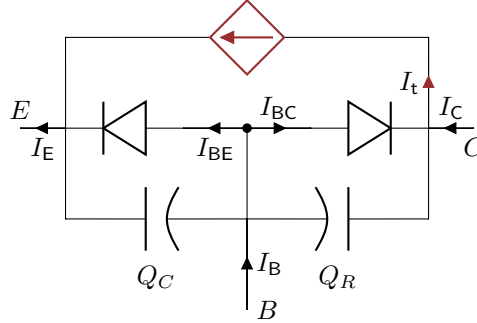
$$I'_B = \frac{V_u - V_{\gamma}}{R_B}$$

Ottenendo la generica relazione:

$$V_u(n) = V_{OHMIN}(n) = \frac{R_B V_{cc} + n R_C V_{\gamma}}{R_B + n R_C}$$

Da cui $N_{MH} = V_{OHMIN}(n) - 1.23$. Definiamo con FAN-OUTmax il massimo punto in per cui il valore di N_{MH} rimane superiore a N_{ML} . Risolvendo l'equazione $0.55 = V_{OHMIN}(n) - 1.23$ otteniamo $n = 31.26$, per questo FAN-OUTmax è 31.

Transistor in regione dinamica



A pagina 13 è stato analizzato il diodo in regime dinamico.

Effettuiamo quindi un'analisi analoga al transistore bipolare, utilizzando come formule, quelle del modello di Ebers e Moll.

Dato che i termini esponenziali compaiono spesso, introduciamo per semplicità le seguenti notazioni:

$$X_{BE} = (e^{V_{BE}/V_T} - 1) = \frac{I_F}{I_S}$$

$$X_{BC} = (e^{V_{BC}/V_T} - 1) = \frac{I_R}{I_S}$$

Inoltre indichiamo con $I_F = I_S X_{BE}$ e $I_R = I_S X_{BC}$ le correnti forward e backward che si manifestano. Ricordandoci che $\alpha_R = \frac{I_S}{I_S + I_{BCS}}$ e $\alpha_F = \frac{I_S}{I_S + I_{BES}}$, le formule del modello sono esprimibili come:

$$I_C = I_F - \frac{I_R}{\alpha_R}$$

$$I_E = \frac{I_F}{\alpha_F} - I_R$$

$$I_B = \frac{I_F}{\beta_F} + \frac{I_R}{\beta_R}$$

Queste equazioni descrivono le tre correnti statiche del transistore, alle quali vogliamo aggiungere le correnti dinamiche, come nel caso del diodo.

Aggiungiamo quindi due capacità parassita alla giunzione, chiamate rispettivamente Q_F e Q_R .

Introduciamo quindi relazioni analoghe a descrivere le due cariche:

$$Q_F = Q_{Fs} X_{BE} = Q_{Fs} \frac{I_F}{I_S}$$

$$Q_R = Q_{Rs} X_{BC} = Q_{Rs} \frac{I_R}{I_S}$$

Da cui seguono le relazioni:

$$\frac{Q_F}{I_F} = \frac{Q_{Fs}}{I_S} = \tau_F$$

$$\frac{Q_R}{I_R} = \frac{Q_{Rs}}{I_S} = \tau_R$$

Sostituendo queste relazioni alle formule in regime statico otteniamo il modello:

$$\begin{aligned} I_C &= \frac{Q_F}{\tau_F} - \frac{Q_R}{\alpha_R \tau_R} \\ I_E &= \frac{Q_F}{\alpha_F \tau_F} - \frac{Q_R}{\tau_R} \\ I_B &= \frac{Q_F}{\beta_F \tau_F} - \frac{Q_R}{\beta_R \tau_R} \end{aligned}$$

In altre parole, siccome tutte le correnti e cariche dipendono linearmente dagli stessi esponenziali, è possibile ricavare una relazione che lega le correnti alle cariche.

Queste equazioni descrivono ancora il modello statico del transistor, ma lega però il valore statico della carica immagazzinata alle due giunzioni. Se ci muoviamo in regime dinamico, non è più vero che la corrente sui condensatori è nulla, ma è esprimibile come $\frac{dQ_F}{dt}$ e $\frac{dQ_R}{dt}$, ottenendo le nuove relazioni:

$$\begin{aligned} I_C &= \frac{Q_F}{\tau_F} - \frac{Q_R}{\alpha_R \tau_R} - \frac{dQ_R}{dt} \\ I_E &= \frac{Q_F}{\alpha_F \tau_F} - \frac{Q_R}{\tau_R} + \frac{dQ_F}{dt} \\ I_B &= \frac{Q_F}{\beta_F \tau_F} - \frac{Q_R}{\beta_R \tau_R} + \frac{dQ_F}{dt} + \frac{dQ_R}{dt} \end{aligned}$$

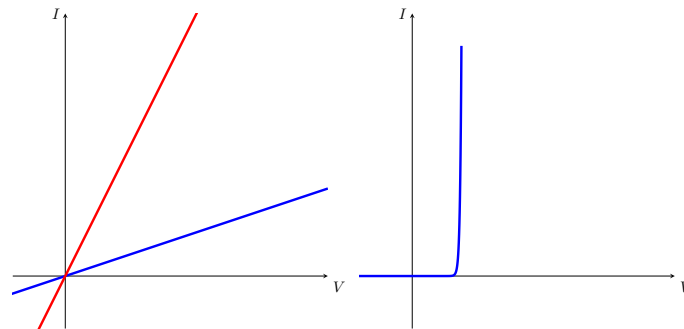
Chiameremo questo modello: **modello a controllo di carica**, nel senso che le espressioni sono dipendenti dalla carica e non più dalla tensione.

Dispositivi e Circuiti Elettronici

La corrente elettrica $I = \frac{dQ}{dt}$ è generata da movimenti di carica. Uno tra i diversi in moto per spostare delle cariche è attraverso un campo elettrico.

La densità di corrente elettrica J è legata al campo elettrico E dalla conducibilità elettrica σ attraverso la legge di ohm $\vec{J} = \sigma \vec{E}$. I materiali sono classificabili dal loro valore di σ , i materiali isolanti sono caratterizzati da bassi valori di σ , corrispondenti a bassi valori di corrente. Valori di σ alti comportano un'alta corrente, caratteristica dei materiali conduttori.

I semiconduttori sono materiali con conducibilità elettrica σ variabile (ad esempio in funzione della temperatura). Guardando il caso del componente resistenza, la corrente ai capi di essa dipende dall'inverso della costante R , quindi, a tensione costante, maggiore è la resistenza, minore è la corrente misurata ai due capi del componente.



In figura sono riportate le relazioni corrente-tensione di una resistenza ed un diodo. È possibile osservare come il comportamento della corrente del diodo è completamente differente.

Modello fisico

Gli elettroni di ciascun materiale orbitano attorno ai rispettivi nuclei, in stato di equilibrio. In stato di equilibrio, gli elettroni di un materiale orbitano attorno al nucleo, la complessiva somma delle cariche risulta nulla.

Gli elettroni attratti dalla forza coulombiana, orbitano attorno al nucleo. Ad ogni orbita corrisponde una velocità di percorrenza, quindi un'energia cinetica. L'ampiezza di ogni orbita dipende inversamente dal quadrato della distanza dal nucleo (formula forza di coulomb).

In definitiva, possiamo

Maggiore è la grandezza dell'orbita degli elettroni, maggiore è la velocità. Inoltre maggiore è la grandezza dell'orbita degli elettroni, maggiore è la velocità degli elettroni, quindi la loro energia cinetica.

All'ampiezza dell'orbita è quindi associato inversamente la forza attrattiva del nucleo, ed un'energia cinetica legata alla velocità di percorrenza.

Possiamo quindi, gli elettroni interni ad un atomo generico in funzione dell'energia da essi posseduta, ottenendo un diagramma simile a quello in figura ..., distinta da diversi livelli



Ad ogni elettrone è dunque associato un livello sull'asse delle energie. Per il principio di quantizzazione, dell'energia, non tutti i possibili livelli di energia sono possibili. L'asse quindi non è continua, ma quantizzata.

È importante notare che le orbite degli elettroni attorno agli atomi non sono descritte dalla meccanica classica, ma dalla meccanica quantistica, dove gli elettroni sono descritti attraverso forme sinusoidali.

Per questo motivo è possibile associare agli elettroni una lunghezza d'onda λ , dove nel caso di ipotesi stazionaria, è necessario che la circonferenza dell'orbita sia multiplo di λ .

Questo spiega a grandi linee la presenza di valori permessi e proibiti sull'asse delle energie.

Principio di esclusione di Pauli

Ciascun livello energetico è occupabile al più da due elettroni (con spin opposto). In altre parole, il numero di elettroni che possono avere una certa distanza dal nucleo è finito.

Da questo è possibile dedurre che è possibile occupare interamente i livelli di energia.

Spostare corrente vuol dire spostare elettroni, cambiandone la velocità, quindi aumentandone l'energia cinetica, vincolata dai livelli di energia e dal principio di esclusione.

Per muovere un'elettrone quindi serve almeno l'energia per raggiungere il livello energetico libero più vicino.

Nel momento in cui due atomi diventino abbastanza vicini da interagire, quello che mi posso aspettare è che gli elettroni dei rispettivi atomi esercitino una forza repulsiva, con l'effetto di modificare l'orbita dei due elettroni.

Considerando quindi un unico diagramma energetico per il sistema dei due atomi, quello che ottengo è che i livelli non si sovrappongono ma si scostano leggermente, mantenendo comunque i principi di quantizzazione ed esclusione.

Generalizzando il discorso con un'interazione di n atomi, otteniamo che ai precedenti livelli energetici, corrisponderà una moltitudine di livelli permessi, tra loro poco differenti, chiamata *banda permessa*. I valori di energia tra due bande permesse, prendono il nome di *banda proibita*.

È importante sottolineare che i valori interni alla banda permessa rispettano ancora il principio di esclusione e di quantizzazione, quindi è possibile che un'intera banda sia occupata da elettroni.

In condizione di quiete, gli elettroni tendono spontaneamente ad occupare i livelli con minore energia, quindi quelli più bassi.

La statistica di fermi indica un valore limite (*Livello di Fermi*) che indica, in assenza di perturbazione, in termini probabilistici, tutti i livelli che sotto tale valore risultano occupati.

La posizione di questo livello diventa quindi fondamentale per indicare le condizioni di trasporto di carica del materiale. Nel caso di materiali conduttori, il livello di fermi ricade internamente ad una banda permessa, l'energia richiesta per spostare elettroni da livelli energetici occupati a livelli energetici liberi, è dipendente dalla loro distanza, quindi molto bassa.

Nel caso di materiali isolanti, il livello di fermi ricade all'interno di una banda proibita, l'energia sufficiente richiesta (energy gap) è quella per scavalcare l'intera banda proibita, quindi molto superiore al caso precedente.

I materiali isolanti sono dunque anch'essi soggetti a fenomeni di scarica elettrica.

I semiconduttori

Il semiconduttore ha una struttura simile a quella dell'isolante, quello che differisce è l'ampiezza del gap, richiedendo una quantità di energia bassa per effettuare il "salto" della banda proibita. Se la quantità di energia ricevibile dalla temperatura dall'ambiente è pari o superiore al gap, diventa facile che elettroni passino da una banda energetica superiore. Elettroni quindi nella banda permessa superiore, trovandola completamente vuota, richiedono a loro volta poca energia per muoversi da un livello energetico all'altro, fornendo al materiale caratteristiche di un conduttore. Chiameremo quindi questa nuova banda, banda di conduzione, mentre chiameremo la vecchia banda, banda di valenza.

La banda di valenza, avendo anch'essa livelli svuotati da elettroni spostati in banda di conduzione, richiederà anch'essa poca energia per effettuare salti di gap interni alla banda. La conducibilità elettrica aumenta quindi con la temperatura del circuito.

Riassumendo quindi, un semiconduttore il livello di fermi interseca una banda proibita, la cui ampiezza è sufficientemente piccola da essere probabile l'effetto di scavalco con la sola energia termica. Il semiconduttore ha quindi "due" bande di conduzione.

Seconda lezione

La distanza dal nucleo cresce col crescere dell'energia, andando a vedere gli elettroni sullo strato di valenza, sono quelli più distanti dal nucleo. La forma del reticolo di atomi è determinata dal numero di elettroni disponibili a collegarsi agli atomi vicini. Tutti i materiali semiconduttori sono materiali della quarta colonna della tavola periodica, i quali possiedono quattro elementi nell'orbita di legame, tra essi prevale il silicio.

In alternativa agli elementi della tavola periodica è possibile formare delle leghe tra elementi della terza e quarta colonna o quarta e quinta, come arseniuro di gallio.

Schematizzando il reticolo del silicio su una mappa in due dimensioni (figura ?) La promozione da banda di valenza a banda di conduzione, significa che esso è meno legato al nucleo originario. È talmente poco legato al nucleo che attraverso una forza di un campo magnetico può spostarsi internamente al reticolo. Tali elettroni vengono definiti come elettroni liberi.

Il materiale in condizione di quiete è neutro, ovvero ha tanta carica positiva quanta negativa. Al momento in cui un'elettrone libero esce da una regione, quella regione non ha più carica nulla ma leggermente positiva. Possiamo quindi immaginare il moto dell'elettrone come uno spostamento di carica nello spazio, quindi una corrente.

Lo spazio lasciato vuoto dall'elettrone, è successivamente occupato da altri elettroni liberi. Creando uno spostamento a catena degli elettroni.

Alla banda di conduzione è associato il movimento di una carica negativa, alla banda di valenza è associato il movimento della lacuna, carica positiva generata dallo spostamento dell'elettrone.

In un conduttore esiste solo un tipo di portatore di carica (elettroni), mentre nei semiconduttori esistono portatori di carica positiva e portatori di carica negativa.

È possibile modellare il movimento della lacuna come il movimento di una fittizia particella fisica, dotata di una massa (maggiore di quella dell'elettrone perché si muove più lentamente)

Per misurare la corrente è necessario conoscere la quantità di elettroni e lacune in un determinato volume. Indicheremo quindi con n il numero di elettroni per unità di volume¹. Analogamente definiamo la concentrazione di lacune con p come il numero di lacune per unità di volume.

Per densità di carica degli elettroni si calcola con $-qn$, mentre per lacune qn . Ad ogni elettrone libero in banda di valenza, corrisponde una lacuna in banda di conduzione, quindi necessariamente all'equilibrio $p = n$.

Possiamo quindi definire come evento di generazione il momento in cui un'elettrone abbandona la banda di valenza generando una lacuna, mentre il fenomeno duale, il passaggio da banda di conduzione a banda di valenza viene chiamato evento di ricombinazione.

Indichiamo quindi con G il numero di coppie elettrone-lacune generate nell'unità di volume e nell'unità di tempo. Il tasso di ricombinazione è indicato con R .

All' equilibrio $n = p = n_i(T)$ e $R = G \geq 0$ il numero di elettroni e lacune costante prende il nome di n_i (concentrazione intrinseca del materiale).

Sostituendo ad alcuni atomi di silicio, con atomi della 5a colonna, ad es fosforo. In questo modo avendo un'elettrone di legame in più (ed un protone in più), siccome i legami sono tutti occupati con gli atomi di silicio adiacenti, otteniamo un'elettrone che non contribuisce ad alcun legame nel reticolo.

La sostituzione di alcuni atomi di silicio con atomi della 3a o 5a colonna prende il nome di drogaggio.

La rara sostituzione di atomi di silicio, non varia la struttura del reticolo cristallino originario.

L'elettrone nella fascia più esterna non essendo legato agli altri se riceve energia sufficiente può liberarsi e comportarsi come una carica negativa mobile. È ancora vero che lascia alle sue spalle una carenza di carica negativa, ma la carica positiva è associata alla presenza del protone nel nucleo e non è in grado di spostarsi. In questo caso si genera un'elettrone mobile, senza generare lacune mobili. In questo caso $p \neq n$. Gli atomi della quinta colonna prendono quindi il nome di atomi droganti di tipo donatore.

Tutto questo è rappresentabile nel diagramma delle energie inserendo un livello "donatore" interno alla banda proibita permettendo che l'evento di liberazione dell'elettrone richieda meno energia, rendendolo

¹Misurata per numero di elettroni per centimetro cubo

ancora più probabile a temperatura ambiente.

Chiameremo N_D la concentrazione di atomi donatori, e per mantenere la struttura del cristallino $N_D \ll 10^{22}$.

Analogamente prendendo un' elemento della terza colonna (es. boro) nel reticolo viene rimosso un legame, fornendo energia al reticolo e spostando elettroni, essi andranno ad occupare del legame mancante, generando lo spostamento di una carica positiva. La carica negativa è fissa perché legata alla struttura del boro.

Chiameremo il boro atomo accettore, perché capace di ionizzare il reticolo negativamente. Con N_A indicheremo la concentrazione degli atomi accettori per unità di volume. Interpretando l'evento sul diagramma energetico, sarebbe come un livello energetico che ricade nella banda proibita molto vicino alla banda di valenza. Generando una lacuna mobile in tale banda.

In questo caso $n < p$.

Terza Lezione

Abbiamo interpretato il diverso comportamento delle due bande, elettroni liberi che si muovono tra livelli energetici nella banda di conduzione, e lacune alle quali è associato un significato di carica positiva mobile.

In un materiale intrinseco, elettroni e lacune sono creati sempre in coppia, quindi necessariamente la concentrazione di elettroni e lacune sono uguali alla concentrazione intrinseca, dipendente dalla temperatura.

Inoltre in condizioni di equilibrio il tasso di generazione G e di ricombinazione R devono essere uguali.

L'equivalenza tra elettroni e lacune interna al materiale può essere modificata attraverso drogaggi con atomi donatori (N_D) o accettori (N_A). La concentrazione elettroni-lacune in materiali drogati è chiamata concentrazione estrinseca.

Supponendo di conoscere la concentrazione tra atomi donatori ed accettori, calcoliamo ora i nuovi valori di p ed n interni al materiale.

Osservando il caso della concentrazione intrinseca, osserviamo che necessariamente vale il prodotto $pn = n_i^2$. Questa relazione è valida non solo per materiali intrinseci, ma anche per materiali estrinseci.

Possiamo pensare che il tasso di generazione G sia dipendente unicamente dalla temperatura: $g(T)$. Mentre il tasso di ricombinazione R , l'occupazione di una lacuna da parte di un'elettrone, è logico pensare che dipenda dal numero di elettroni presenti nel materiale e dalla temperatura: $R = pnr(T)$. Il tasso di generazione non dipende dalla concentrazione effettiva di p ed n , perché per evitare che si effettui una generazione, tutti gli elettroni in banda di valenza dovrebbero essersi spostati in banda di conduzione. Ipotesi irragionevole a temperatura ambiente.

Ricordandoci che in condizioni di equilibrio $G = R$, otteniamo che $g(T) = pnr(T)$. Quindi $pn = g(T)/r(T) = f(T)$ è dipendente unicamente dalla temperatura. Quindi a temperatura costante il valore pn è costante e vale n_i^2 .

Inoltre se è vero che a livello globale la quantità di carica positiva equivale alla carica negativa, allora

la densità di carica ρ in tutto il volume è costante e pari a 0.

$$\rho = -qn + qp + qN_D - qN_A = q(p - n + N_D - N_A) = 0$$

La densità di carica dipende dal numero di elettroni n e lacune p con rispettiva carica, ed il numero di atomi accettori e donatori interni ad un materiale, ad ognuno dei quali è associata rispettivamente una carica fissa in modulo q .

Unendo la relazione con $pn = n_i$, ricaviamo:

$$n = \frac{(N_D - N_A) \pm \sqrt{(N_D - N_A)^2 + 4n_i^2}}{2}$$

Siccome è evidente che non può valere il segno meno in quanto condurrebbe ad un valore di n negativo, ottengo un'espressione del valore di n dipendente solo da valori noti.

Con questo ragionamento otteniamo anche il valore di p :

$$p = \frac{(N_A - N_D) + \sqrt{(N_A - N_D)^2 + 4n_i^2}}{2}$$

Dalle espressioni si può osservare come compare sempre la differenza dei valori N_D ed N_A , il fenomeno prende il nome di principio di compensazione, il quale dice che non è importante il singolo valore di N_D ed N_A , quello che conta è sempre la loro differenza.

Nel caso di una concentrazione con forte sbilanciamento, es $N_D \gg N_A \gg n_i$, nelle equazioni in precedenza, posso trascurare i termini N_A ed n_i , ottenendo $n = N_D$. Quindi se la concentrazione del drogante è molto maggiore rispetto alla concentrazione intrinseca, allora la concentrazione di elettroni è dovuta praticamente solo ad esso. Siccome vale $pn = n_i^2$, abbiamo che $p = n_i^2/N_D$. Segue che il trasporto della carica, avviene quasi unicamente attraverso elettroni.

Chiameremo questo materiale "estrinseco di tipo n", ad indicare la larga prevalenza degli elettroni.

Posso effettuare gli stessi ragionamenti nel caso duale ($N_A \gg N_D \gg n_i$), ottenendo $p = N_A$ e $n = n_i^2/N_A$. Chiameremo questo materiale "estrinseco di tipo p".

Studio conduzione materiale uniforme, estrinseco di tipo n

Conoscendo la conduzione del materiale $n \approx N_D$, ed indicando con S la sezione trasversale del materiale, indicheremo $J = I/S$. Ricordando che $I = dQ/dt$, otteniamo:

$$J = \frac{1}{S} \frac{dQ}{dt}$$

Sapendo che il calcolo elettrico E esercita una forza di coulomb pari a $F = -qE$, ed $F = ma$, otteniamo $a = -qE/m$, quindi accelerazione costante a campo elettrico costante. Quindi un'elettrone dovrebbe muoversi di moto uniformemente accelerato.

Ma ciò non è vero perché l'elettrone è soggetto anche alle forze interne al reticolo cristallino. Quindi è soggetto ad una forza repulsiva. Quello che si può vedere sperimentalmente è che la velocità V_n è costante ed esprimibile come $V_n = -\mu_n E$, con μ_n costante di mobilità elettronica.

Gli elettroni si muovono quindi in direzione opposta al campo elettrico, con velocità $|V_n|$. Il percorso che percorre una singola particella in un piccolo intervallo di tempo è esprimibile quindi come $dx = -\mu_n E dt$.

La carica dQ è quella che nello stesso intervallo di tempo dt attraversa la sezione del materiale, è possibile pensarla come $nqSdx$ (carica del numero di portatori nel volume Sdx).

Quindi la densità J è esprimibile come:

$$\begin{cases} \frac{dx}{dt} = \mu_n E \\ J = \frac{1}{S} \frac{dQ}{dt} \\ Q = nqSdx \end{cases} \Rightarrow J_n = \frac{1}{S} \frac{nqS\mu_n E dt}{dt} = nq\mu_n E$$

Vale ancora la relazione di ohm in forma locale $J = \sigma E$, con $\sigma = nq\mu_n = q\mu_n N_D$. È possibile concludere che $N_D \propto \sigma$. Quindi possiamo variando dinamicamente il valore di N_D si può ottenere un conduttore o un isolante.

Quindi solo variando il valore di N_D in una parte del materiale posso formare due regioni conduttive, separate da una isolante.

Conduzione materiali estrinseci di tipo p

Con ragionamento del tutto analogo al precedente, si ottiene che la velocità di mobilità delle lacune è pari a $V_p = \mu_p E$. Nel dettaglio $\mu_p \approx \frac{2}{3}\mu_n$, sottolineando che le lacune si comportano come particelle fittizie che si muovono più lentamente degli elettroni.

Ottenendo

$$J_p = \frac{1}{S} \frac{qp\mu_p E S dx}{dt} = qp\mu_p E$$

In questo caso la conducibilità elettrica è determinata da N_A .

La totale conducibilità elettrica è possibile esprimerla come $J = J_n + J_p$. J conta la quantità di carica che attraversa la sezione, quindi gli elettroni portano contributo positivo muovendo carica negativa in senso opposto, e le lacune portano contributo positivo.

$$J = J_n + J_p = nq\mu_n E + qp\mu_p E$$

Materiale drogato non uniformemente

Nel caso in cui il campo elettrico esterno sia nullo, gli elettroni si spostano ugualmente per il moto browniano. Possiamo pensare alla carica dQ come $dQ_1 - dQ_2$ ovvero la carica dovuta agli elettroni che escono dalla sezione di volume, e la carica degli elettroni entranti nella sezione di volume. Chiamata quindi n_1 la quantità di elettroni presente nel primo volume, e n_2 la quantità di elettroni nel secondo; siccome le cariche hanno un 50% di probabilità di entrare nel volume, si possono esprimere le cariche come:

$$\begin{aligned} dQ_1 &= -\frac{qn_1 dV}{2} \\ dQ_2 &= -\frac{qn_2 dV}{2} \end{aligned}$$

Ricordando che il volume $dV = SV_{th}dt$ si ottiene che:

$$J_n = \frac{1}{S} \frac{-qn_1 SV_{th}dt + qn_2 SV_{th}dt}{2dt} = q(n_2 - n_1)V_{th} \neq 0$$

Ma siamo giunti ad un assurdo, in quanto avremmo trovato un modo per generare corrente senza spendere energia.

Effettuando i calcoli più accuratamente è possibile ottenere $J_n = qD_n \frac{dn}{dx}$. Quello che accade è il fenomeno della diffusione, infatti tende a muovere la carica, dalla regione di concentrazione più alta a quella più bassa. Il sistema evolve verso una distribuzione uniforme.

Lo spostamento degli elettroni lascia cariche positive nella regione con maggior carica negativa, portando una carica negativa nell'altra regione.

Otteniamo quindi un campo elettrico nella direzione della regione con meno cariche negative, pari a $F = -qE$.

Lezione 4

Analogamente al caso precedente, studiamo un materiale estrinseco di tipo p, la cui concentrazione di lacune non è costante, ottenendo:

$$J_p = \frac{1}{S} \frac{d(Q_1 - Q_2)}{dt} = \frac{q(p_1 - p_2)V_{th}}{2}$$

contrariamente al caso degli elettroni, se $p_2 > p_1$ otteniamo $J_p < 0$. In questo, effettuando i calcoli in maniera più accurata, otterremo che $J_p = -qD_p \frac{dp}{dx}$.

Abbiamo introdotto quindi i funzionamenti di tipo ohmico: $J_n = q\mu_n nE$ e $J_p = q\mu_p pE$, la cui direzione è concorde; ed un meccanismo di tipo diffusivo: $J_n = qD_n \frac{dn}{dx}$ e $J_p = -qD_p \frac{dp}{dx}$, la discordanza di segno è dovuta al bilanciamento della carica.

Non possiamo considerare i due funzionamenti come uno indipendente dall'altro, in quanto porterebbe a conclusioni assurde, esprimiamo quindi la combinazione degli effetti:

$$\begin{aligned} J_n &= q\mu_n nE + qD_n \frac{dn}{dx} \\ J_p &= q\mu_p pE + qD_p \frac{dp}{dx} \end{aligned}$$

Quindi un conduttore prevede un solo meccanismo di trasporto di carica, per un meccanismo unicamente di tipo ohmico. In un semiconduttore teniamo conto di due specie di carica e di due specie di carica e due diversi meccanismi di trasporto di carica: trasporto ohmico e diffusivo.

Il modello prende infatti il nome di ohmico-diffusivo (*drift-diffusion*).

D_n indica la reattività dell'elettrone alla temperatura.

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

Analogamente $D_p = \frac{kT}{q} \mu_p$. Queste prendono il nome di relazioni di Einstein, dove k è la costante di Boltzmann, q è la carica dell'elettrone e T è la temperatura.

Quello che dicono queste relazioni è che il coefficiente di diffusione e la mobilità dell'elettrone, sono legate. Ricordando che la mobilità degli elettroni è superiore a quella delle lacune, analogamente il fattore D_n è maggiore di D_p .

Ricordandoci che $E = -\frac{d\phi}{dx}$ prendiamo in considerazione, un materiale, tale per cui la concentrazione varia in base alla coordinata x : $n(x)$. Siccome il campo elettrico E , costante ed indipendente dalla posizione x , prendiamo un punto x_1 interno al materiale e lo usiamo come origine per misurare il campo elettrico. In tale punto avremo che $\varphi = 0$, n_0 e p_0 .

Allo stesso identico modo posso scegliere un punto x_2 , caratterizzato da un $\varphi(x)$, $n(x)$ e $p(x)$.

All'equilibrio $pn = n_i^2$, otteniamo che, dalla funzione di $J_n = q\mu_n nE + qD_n \frac{dn}{dx} \approx 0$, otteniamo:

$$\frac{kT}{nq} \frac{dn}{dx} = \frac{d\varphi}{dx}$$

Integrando la relazione sui margini $\varphi(0)$ e $\varphi(x)$, ottengo:

$$\varphi(x) = \frac{kT}{q} \ln \frac{n(x)}{n_0}$$

Da cui

$$n(x) = n_0 e^{\frac{q\varphi(x)}{kT}}$$

Inoltre dalla relazione $p(x)n(x) = n_i^2$, otteniamo la relazione duale $p(x) = p_0 e^{-q\varphi(x)/kT}$ la inversa $\varphi(x) = -\frac{kT}{q} \ln \frac{p(x)}{p_0}$.

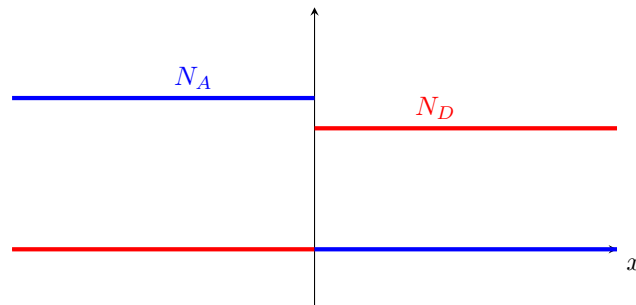
Dalle equazioni di Maxwell, per i materiali uniformi: $\frac{d^2\varphi}{dx^2} = -\frac{\rho}{\epsilon}$

Ragionamento su struttura diversamente drogata

Supponiamo un unico cristallo di semiconduttore, drogato per metà in con atomi accettori, e per metà con atomi donatori. Il materiale prende il nome di giunzione pn.

Studio della struttura all'equilibrio

Supponiamo una differente distribuzione di atomi accettori e donatori. Prendendo quindi come origine il punto intermedio tra le due giunzioni, otteniamo:



Supponiamo che il dispositivo sia abbastanza lungo, da supportare l'ipotesi che dopo una certa distanza, l'effetto della giunzione risulti trascurabile, in modo che sia possibile esprimere le relazioni dei due estremi indipendentemente. Quindi, che esistano due coordinate w_p e w_n , che oltre a quelle coordinate, le due regioni si comportano come regioni uniformemente drogate di tipo p ed n .

Zona neutra p

Per $x < -w_p$, la popolazione di portatori maggioritaria sono le lacune, quindi $p = N_A$ e $n = \frac{n_i^2}{N_A}$

$$\rho = q(N_D - N_A + p - n) \approx 0$$

$$J_p = q\mu_p p E - qD_n \frac{dp}{dx} = 0$$

Essendo la densità di carica pari a 0, e la variazione di atomi portatori nulla, allora necessariamente $E = 0$

Siccome E dipende dalla derivata di φ rispetto ad x allora φ è costante. Prendiamo questo valore, come valore di riferimento, per i successivi φ , quindi $\varphi = 0$.

Zona neutra n

Per $x > w_n$, $n = N_D$ e $p = \frac{n_i^2}{N_D}$. Inoltre $\rho = q(N_D - n) \approx 0$.

$$J_n = q\mu_n n E + qD_n \frac{dn}{dx} = 0$$

Per analogo ragionamento $E = 0$ e φ è costante, calcolabile attraverso la formula

$$\frac{kT}{q} \ln \frac{n(x)}{n_0} = \frac{kT}{q} \ln \frac{N_D N_A}{n_i^2}$$

Indicheremo questo valore con Ψ_B . Siccome i valori di N_D ed N_A sono entrambi molto maggiori di n_i , allora Ψ_B è positiva e molto maggiore di 1. Inoltre siccome dipende dal logaritmo di N_D ed N_A , per variazioni di quest'ultimi è assimilabile ad una costante.

Regione svuotata

Per $-w_p < x < w_n$, ipotizzando che il potenziale sia monotono in questo intervallo, possiamo presupporre che $0 < \varphi(x) < \Psi_B$.

Quindi in tutto l'intervallo $\varphi(x) > 0$, da cui moltiplicando per la quantità positiva $\frac{q}{kT}$, otteniamo

$$p_0 e^{-\frac{q\varphi}{kT}} \ll p_0 = N_A$$

Il cui primo termine è $p(x)$. Analogamente, da ipotesi di funzione monotona, posso dire che $\varphi(x) < \Psi_B$, e moltiplicandola per la stessa quantità positiva $\frac{q}{kT}$ ottengo:

$$n(x) = n_0 e^{\frac{q\varphi(x)}{kT}} \ll n_0 e^{\frac{q\Psi_B}{kT}} = N_D$$

Da cui i valori di n e p sono rappresentati graficamente da valori prossimi a 0.

Regione svuotata e regione di carica

Per $-w_p < x < 0$ La densità di carica non sarà più nulla, ma siccome $p(x) \ll N_A$, $n(x) \ll N_D$, in base all'intervallo in cui mi trovo, per la relazione $\rho = q(N_D - N_A + p - n)$, la densità di carica potrà valere $-qN_A$ o qN_D . Per la relazione di poisson: $\frac{dE}{dx} = -\frac{\rho}{\epsilon}$. Sostituendo rispettivo valore di ρ ed integrando nei relativi estremi, si ottiene il campo elettrico:

$$E = -\frac{qN_A}{\epsilon}(x + w_p)$$

Nel caso particolare $E(0^-) = \frac{-qN_A w_p}{\epsilon}$

Dalla relazione $E = -\frac{d\phi}{dx}$, integrando sugli estremi si ottiene:

$$\phi(x) = \frac{qN_A}{\epsilon} \frac{(x + w_p)^2}{2}$$

con $\phi(0^-) = \frac{qN_A}{2\epsilon} w_p^2$

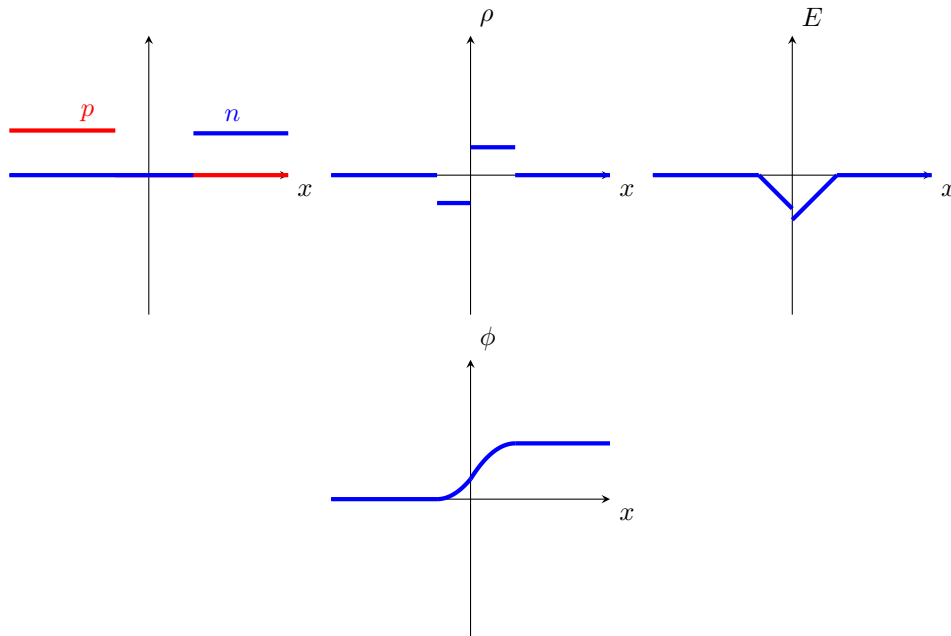
Per $0 < x < w_n$, con analoghi ragionamenti, $\rho = qN_D$ ed il campo elettrico E e potenziale ϕ saranno rispettivamente:

$$E(x) = \frac{qN_D}{\epsilon}(x - w_n)$$

$$E(0^+) = -\frac{qN_D}{\epsilon} w_n$$

$$\phi(x) = \Psi_B - \frac{qN_D}{2\epsilon}(x - w_n)^2$$

$$\psi(0^+) = \Psi_B - \frac{qN_D}{2\epsilon} w_n^2$$



Se è presente una corrente diffusiva, è presente anche una corrente opposta di trascinamento ohmico. La regione $-w_p < x < 0$ è sede di carica negativa fissa, e la controparte $0 < x < w_n$ è ricca di atomi droganti, che avendo donato elettroni per effetto di diffusione, portano una regione di carica positiva fissa.

Le cariche fisse positive e negative, rappresentano un dipolo di carica, e generano un campo elettrico

rivolto verso la regione p. Per sostenere tale campo elettrico è necessaria una differenza di potenziale.

La condizione di equilibrio è raggiungibile in due modi: o sia la componente ohmica e diffusiva sono nulle, o opposte. In particolare lo spostamento di elettroni per diffusione forma un campo elettrico ohmico che all'equilibrio è uguale ed opposto.

Φ_B prende infatti il nome di potenziale di barriera, perché è necessario a sostenere il campo elettrico necessario a bilanciare la diffusione.

Lezione 5 - 12

Il modello appena analizzato si basa sull'esistenza di due coordinate w_p e w_n , che discriminano regioni tra vicine e lontane alla giunzione di carica.

Osservando che $E(0-) = E(0+)$, dalle relazioni, si ricava che $-qN_A w_p = -qN_D w_n$, considerato che dal grafico della densità di carica ρ , qN_A e w_p sono base ed altezza del rettangolo

L'equazione esprime l'uguaglianza tra le due aree dei rettangoli del primo e terzo quadrante. Integrando quindi densità di carica sul volume, moltiplico entrambi i rettangoli per la sezione trasversale S , ottengo che il volume dei due cilindri $w_p S$, moltiplicato alla densità di carica qN_A , ottenendo le cariche Q_p e Q_n associate alla regione svuotata.

In questa regione quindi, il dispositivo si comporta come un condensatore. La differenza è che la carica del condensatore è concentrata su due armature, questa carica è distribuita nel volume.

Questa relazione esprime che w_p e w_n sono legati tra loro da una relazione.

Ponendo l'uguaglianza tra le due formule del potenziale Ψ nell'origine, otteniamo:

$$\Psi_B = \frac{q}{w\varepsilon} \left(N_A \cdot \frac{N_D^2}{N_A^2} + N_D \right) w_n^2 = \frac{qN_D^2}{2\varepsilon} \left(\frac{1}{N_A} + \frac{1}{N_D} \right) w_n^2$$

Da cui si ricava il valore di w_n come:

$$w_n = \frac{1}{N_D} \sqrt{\frac{2\varepsilon\Psi_B}{q\left(\frac{1}{N_D} + \frac{1}{N_A}\right)}}$$

e w_p segue dalla relazione $w_p N_A = w_n N_D$.

Dipendentemente dalla concentrazione dei droganti, i valori di w_p e w_n assumono valori dell'ordine di $10^{-6}m$ o inferiori, quindi l'ipotesi di esistenza delle zone neutre, sufficientemente lontane dalla giunzione, è soddisfatta appena le dimensioni fisiche dei dispositivi eccedono i decimi di micron.

L'alta densità di lacune nella regione p , tende per effetto di diffusione, a muoversi verso la regione n . Allo stesso tempo, l'alta densità di elettroni tendono a spostarsi nella regione p . Per questa condizione la corrente non può essere nulla. Difatti ad un dipolo di carica è associato un campo elettrico, col verso rivolto dalla carica positiva, alla negativa.

Il campo elettrico è intrinseco. Le lacune si muovono per effetto di diffusione da p ad n , gli elettroni si muovono per il campo elettrico da n a p per effetto di trascinamento.

L'effetto del campo elettrico contrasta il fenomeno della diffusione, ed all'equilibrio le due componenti sono bilanciate. Nell'ipotesi in cui non ci sia trascinamento, ci sarebbe solo diffusione, non contrastata, portando una corrente non nulla, contraddicendo l'ipotesi di equilibrio.

Analogamente se la regione svuotata si estendesse infinitamente, non si avrebbe diffusione in quanto non variabile, ma solo il trascinamento ohmico. I due fenomeni quindi esistono insieme.

La diffusione agisce sulle lacune, portandole dalla regione p , alla n , per questo prende il nome di diffusione di portatori maggioritari. Il trascinamento, invece agisce sugli elettroni: i portatori minoritari.

$$\Psi_B \approx 0.5V$$

Questa differenza di potenziale non è in grado di fornire energia internamente ad una maglia, è la reazione intrinseca del materiale che si oppone generare corrente senza sorgente di energia.

Ipotizzando quindi di cortocircuitare la giunzione, per lo stesso motivo per cui esiste una differenza di potenziale tra le due regioni polarizzate diversamente, esistono due differenze di potenziale "di contatto" ai capi della giunzione, collegata al filo conduttore in metallo.

Calcoli in regione dinamica

Spostandoci dalla condizione di equilibrio, ovvero applicando un generatore di tensione alla maglia, ne consegue che la corrente interna non è nulla.

I due potenziali di barriera Ψ_{cn} e Ψ_{cb} , possono essere considerati costanti, e uguali alla differenza di potenziale, calcolata precedentemente in condizione di equilibrio. Questa condizione prende il nome di "condizione di contatto ohmico", ovvero che la tensione è indipendente dalla corrente.

Quando un flusso di carica attraversa una regione p o n , la conducibilità elettrica $\sigma = qN_A\mu_p$. Dalla conducibilità si ricava la resistività ρ come $\rho = \frac{1}{\sigma}$, e quindi la resistenza vale: $R = \rho \frac{l}{S}$, considerando in questo modo la regione p come una resistenza. Analogo ragionamento vale per la regione n .

Alle regioni p ed n corrisponde quindi una differenza di potenziale: Ψ_p e Ψ_n .

L'equazione di kirchoff alla maglia diventa quindi:

$$V + \Psi_{cp} - \Psi_p - \Psi_B - \Psi_n - \Psi_{cn} = 0$$

Con V pari alla tensione applicata alla maglia. Prendendo come ipotesi che le differenze di potenziale Ψ_p e Ψ_n siano molto più piccole del potenziale di barriera Ψ_b , possiamo semplificare l'equazione a $V + \Psi_{cp} + \Psi_B - \Psi_{cn} = 0$.

Questa ipotesi di semplificazione è quindi anche una condizione di validità dell'equazione ottenuta.

Ricordando che i due potenziali sono costanti, all'equilibrio: $0 + \Psi_{cp0} + \Psi_{B0} - \Psi_{cn0} = 0$, da cui $\Psi_{B0} = \Psi_{cn0} - \Psi_{cp0} = \Psi_{cn} - \Psi_{cp}$.

Riportato il risultato ottenuto all'equazione precedente:

$$V - \Psi_{B0} + \Psi_B = 0$$

Sotto ipotesi di validità, la barriera di potenziale $\Psi_B = \Psi_{B0} - V$. Applicare una tensione positiva, vuol dire esattamente ridurre la barriera di potenziale dello stesso valore.

Ricordando che la barriera di potenziale è legata al campo elettrico attraverso la formula $\Psi_B = -\int_{w_p}^{w_n} E dx$, al diminuire del potenziale, diminuisce anche l'area del campo elettrico.

Inoltre, siccome la formula del campo elettrico è $E(x) = -\frac{qN_A}{\epsilon}(x + w_p)$, ed il suo coefficiente angolare è costante, allora per fare in modo che l'area del rettangolo diminuisca, i valori w_n e w_p devono per forza avvicinarsi all'origine.

Applicare una tensione positiva significa che il campo elettrico, non crea un effetto di trascinamento sufficiente a bilanciare l'effetto di diffusione. Per questa ragione si crea una corrente I positiva che circola nella maglia. Analogamente se una tensione negativa viene applicata, l'effetto di trascinamento prevale sulla diffusione, favorendo uno spostamento delle lacune nella direzione del campo (negativa), e gli elettroni nella regione concorde, risultando quindi in una corrente negativa.

Per $V > 0$ la diffusione trasporta portatori maggioritari, risultando quindi in una corrente di maggioritari. Per $V < 0$ vengono spostate lacune dalla regione n ed elettroni dalla regione p , risultando in una corrente di portatori minoritari. Quindi per tensione positiva, la corrente risulta molto forte in quanto sono presenti molti portatori di carica. Con la stessa differenza di potenziale ma negativa, si ottiene lo stesso effetto, ma il numero di portatori è estremamente inferiore, risultando in una corrente debole.

Da questo segue l'equazione del diodo $I_s(e^{V/V_T} - 1)$. La zona di polarizzazione diretta corrisponde alla tensione positiva, ed una corrente di portatori maggioritari, e la polarizzazione inversa alla tensione negativa quindi ad una corrente di portatori minoritari.

Verifica delle ipotesi

Tutto questo è vero se sono verificate le ipotesi imposte nella dimostrazione:

- L'esistenza di due zone sufficientemente distanti dalla giunzione, già verificata
- Le cadute sulle zone neutre siano trascurabili

Effettuando una simulazione, sul comportamento del diodo, otteniamo il risultato che l'espressione esponenziale della corrente è valida solamente per valori di tensione compresi in un certo intervallo. Infatti fintanto che la corrente è piccola, la caduta di potenziale sulle due regioni (ipotizzate come resistenze), è piccola e trascurabile, al crescere della corrente l'ipotesi non è più vera.

Inoltre, per tensioni eccessivamente negative, la corrente cresce esponenzialmente, distaccandosi dal comportamento pressoché nullo atteso dal modello. Questo è dovuto all'alto valore del campo elettrico, che innesca fenomeni di generazione elettrone-lacuna (generazione per impatto), aumentando il numero di portatori di carica e di conseguenza la corrente. Prende il nome di "fenomeno di breakdown".

Il fenomeno di breakdown, fornendo una corrente esponenziale sotto certo livello di tensione negativa, è utilizzato da un tipo di dispositivi che prende il nome di diodi zener, caratterizzati dal controllo preciso del livello negativo di tensione.

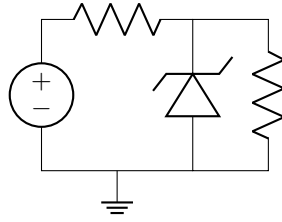


Figure 13: Limitatore di tensione con diodo zener

Il passaggio tra regione di polarizzazione diretta ed inversa del diodo non è immediata, in quanto richiede dello spostamento di carica. Per questo è stato introdotto un tempo di reazione

Lezione 2019514

Siccome in polarizzazione diretta si restringe la giunzione svuotata

Carica fissa associata allo svuotamento: $Q = -qN_A S w_p$, con all'equilibrio w_p noto. Supponendo di spostarci dalla condizione di equilibrio sufficientemente poco da poter considerare w_p esprimibile ancora dalla formula:

$$w_p = \frac{1}{N_A} \sqrt{\frac{2\varepsilon\Psi_{B0}}{q\left(\frac{1}{N_D} + \frac{1}{N_A}\right)}}$$

Nella nuova condizione, siccome varia il potenziale di barriera, la formula diventa

$$w_p(V) = \frac{1}{N_A} \sqrt{\frac{2\varepsilon(\Psi_{B0} - V)}{q\left(\frac{1}{N_D} + \frac{1}{N_A}\right)}}$$

Sostituendo w_p all'equazione della carica Q , si può osservare che la funzione non passa per l'origine

$$Q(V) = -qS \sqrt{\frac{2\varepsilon(\Psi_{B0} - V)}{q\left(\frac{1}{N_D} + \frac{1}{N_A}\right)}}$$

Calcolando la carica specifica (Sezione trasversale unitaria, $S = 1$) per $V = 0$, si ottiene la formula

$$Q(V) = -\sqrt{\frac{2q\varepsilon}{\frac{1}{N_D} + \frac{1}{N_A}}} \sqrt{\Psi_{B0} - V} = -M \sqrt{\Psi_{B0} - V}$$

Osservando quindi che ha un andamento di radice quadrata, e che $Q = 0$ per $V = \Psi_{B0}$. Ipotizzando una tensione maggiore di Ψ_{B0} , il modello prevede una condizione assurda, quindi ci si può aspettare che anche le condizioni non siano più soddisfatte.

Infatti spostandoci verso la regione di polarizzazione diretta, la regione svuotata si restringe, e la componente diffusiva porta le lacune a muoversi dalla regione p alla regione n . Il modello che si basa sull'ipotesi di completo svuotamento non funziona più.

Conoscendo la carica, attraverso la sua derivata, conosco anche la formula della corrente. Per questo è possibile aggiungere una costante al valore della carica, senza far variare l'espressione della corrente. Per questa proprietà trasliamo il grafico della carica per farlo passare per l'origine.

La quantità di carica che si può trovare in una certa zona, dipende dalla probabilità di ricombinazione delle lacune. La carica per $V < 0$ è fissa e di svuotamento, per $V > 0$ è carica mobile, dovuta all'effetto di diffusione.

Ottenere la carica complessiva nel volume è necessario integrare la quantità di carica qn su tutto il volume, ovvero tutta la regione n , dato che nella regione

$$Q = \int_0^\infty qn S dx$$

Siccome siamo in regione di polarizzazione diretta, la corrente diffusiva prevale sulla corrente ohmica, e possiamo. Siccome la componente di corrente dipenda da una forma esponenziale, essendo p dipendente dalla derivata della corrente, anche p sarà in forma esponenziale, quindi la concentrazione di elettroni dipende esponenzialmente da V , ed integrando quella concentrazione si arriva alla conclusione che anche Q è dipendente esponenzialmente da V .

Sunto

In polarizzazione inversa, abbiamo ricavato l'espressione della carica efficace, dipendente da un termine di radice quadrata, ed associata alla carica fissa di svuotamento. Dall'altra parte in polarizzazione diretta, l'andamento è di tipo esponenziale, e dipende dalla carica mobile ed associata al movimento diffusivo.

Mentre la qualità dell'approssimazione della regione in polarizzazione diretta è tanto buona quanto l'approssimazione della corrente, in polarizzazione inversa la componente radice quadrata è approssimata da una retta parallela alle ascisse, ovviamente riportando un errore elevato per $V \rightarrow -\infty$

Per questo motivo viene introdotto in parallelo un condensatore di capacità variabile $C(V) = \frac{dQ}{dV}$, facendo comportare il circuito diodo-condensatore, per piccoli segnali, come un condensatore variabile.

Applicazioni del transistor bipolare

Riprendiamo il transistor bipolare e l'effetto transistor descritto a pagina 17 con relazione:

$$I_E \approx I_C \gg I_B$$

Un transistor npn, è formato da due regioni n agli estremi, ed una regione p intermedia. Non è detto che il drogaggio tra collettore ed emettitore sia uguale.

La distanza w dai due diodi è quindi la distanza tra le due giunzioni p . In condizione di equilibrio si forma in p ed ai suoi capi, una regione svuotata di elettroni.

Andando in regione di polarizzazione diretta, la giunzione base emettitore, vede la regione svuotata ai suoi capi restringersi rispetto alla giunzione precedente, mentre avendo polarizzata in inversa la giunzione base collettore, la rispettiva giunzione si allarga.

La regione non è più svuotata ma è sede di un fenomeno di diffusione, dove la carica si diffonde e si riduce gradualmente per effetto di ricombinazione.

Con distanza w ridotta, all'equilibrio, la situazione non cambia, ma in regione dinamica, essendo le giunzioni più vicine si ha un alto numero di elettroni che non si ricombinano, portando una corrente non trascurabile di portatori minoritari all'altra regione. Quindi fintanto che le giunzioni sono distanti lavorando indipendentemente, quando si avvicinano si ha un effetto di interazione, dove l'emettitore emette elettroni nella giunzione di base, dove non percorrono abbastanza spazio per potersi tutti ricombinare, e buona parte passa quindi alla regione di collettore per effetto di campo.

Più piccola è w maggiore è l'effetto transistor.

Se il transistor fosse simmetrico, allora il funzionamento in regione diretta ed inversa sarebbe identico, ma per via di un differente drogaggio ne risulta che la regione diretta significativamente più efficiente della regione inversa.

L'effetto di ricombinazione è una corrente che non aiuta opposta, chiamata corrente di base. La corrente di emettitore è dovuta agli elettroni che viaggiano dalla zona di emettitore al collettore, per effetto di campo, alle lacune che si spostano dalla base all'emettitore, che si oppone all'effetto transistor. L'obiettivo è rendere la corrente formata dagli elettroni, molto maggiore rispetto alla corrente di base portata dallo spostamento di lacune.

Per fare questo, indichiamo con N_E la concentrazione di atomi donatori nella regione di emettitore, la concentrazione di lacune maggioritarie N_B nella regione di base. Quindi per far sì che la corrente di elettroni sia molto maggiore rispetto alla concentrazione di lacune, bisogna far sì che $N_E \gg N_B$.

Il parametro $\beta_F \gg 1$ e $\alpha_F \approx 1$.

Per far sì che il transistor funzioni ugualmente bene in regione attiva inversa, bisognerebbe arrivare alla condizione che $N_C \gg N_B$, ma dal grafico corrente tensione (I_C , V_{CE}), sono presenti dei tratti dove la corrente è costante ed indipendente dalla tensione. Al variare della V_{CE} l'ampiezza della regione svuotata $w_p(V_{BC})$ è dipendente da V_{BC} secondo ragionamenti fatti in precedenza per il diodo.

Aumentare la tensione V_{BC} porta ad aumentare V_{CE} e quindi aumentare la grandezza della regione svuotata. Portando quindi ad aumentare il cammino degli elettroni in regione svuotata e diminuendo la corrente (Effetto Early).

Quando la regione di base scompare completamente perché sovrapposta dalla regione svuotata, l'effetto prende il nome di "punch-through". L'estensione della regione svuotata è inversamente proporzionale al drogaggio. ($qN_A w_p = qN_D w_n$).

Per minimizzare l'effetto early ed il "punch-through" occorre che il drogaggio della regione di base N_B sia molto superiore al drogaggio del collettore N_C , portando l'asimmetria di drogaggio nel transistor.

Sommariamente $N_E \gg N_B \gg N_C$. Esiste quindi un verso privilegiato per lo scorrimento di corrente.

Il transistor non si comporta come "interruttore ideale" in quanto la corrente di base dovrebbe essere nulla.

Transistor JFET

Supponiamo di avere un blocco di materiale univorme di dimensioni $L \times w \times h$ e drogato uniformemente con concentrazione N_D . Se applicato un campo elettrico ai suoi estremi,

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

Permettendo al dispositivo di comportarsi come buon conduttore o isolante in base al valore di N_D . La resistenza associata ricordiamo che è esprimibile come $R = \rho \frac{L}{hw}$.

Presupponendo di applicare una differenza di potenziale V_L ai suoi estremi, possiamo calcolare la corrente attraverso la legge di ohm:

$$I = \frac{V_L}{R}$$

Prendendo una giunzione pn ed applicando agli estremi delle due regioni una differenza di potenziale V_T , sappiamo che nella giunzione si forma una regione svuotata, la cui ampiezza tende a restringersi in polarizzazione diretta ed allargarsi in polarizzazione inversa.

L'ampiezza w_n ricordiamo che è esprimibile come:

$$w_n = \frac{1}{N_A} \sqrt{\frac{2\varepsilon}{q(\frac{1}{N_D} + \frac{1}{N_A})}} \sqrt{\Psi_{B0} - V}$$

Immaginando ora di mettere sopra al blocco, drogato uniformemente con concentrazione N_D e con una tensione agli estremi V_L , una lastra di materiale p drogato con concentrazione N_A , ed applicare una tensione V_T dall'alto verso il basso. Si formano quindi due correnti I_T e I_L .

Supponendo $V_T < 0$, quindi $V_T \approx 0$, sotto la giunzione si forma una regione svuotata di portatori mobili. La conducibilità della regione svuotata è molto bassa, comportandosi come un isolante. Per calcolare la nuova resistenza del materiale non bisogna considerare la regione svuotata:

$$R = \rho \frac{L}{(h - w_n)w}$$

Agendo quindi sulla tensione V_T si è in grado di allargare la regione svuotata ed aumentare quindi il valore della resistenza R e quindi la corrente I_L .

Questo dispositivo approssima molto di più l'effetto di valvola idraulica, e prende il nome di *Field Effect Transistor* o *JFET*. La regione non svuotata prende il nome di "canale".

Il dispositivo non è più bipolare, ma unipolare, in quanto il trasporto di carica è dovuto solo agli elettroni.

Non è molto utilizzato nei circuiti digitali, in quanto per svuotare completamente la giunzione occorrono valori di V_T molto bassi.

Transistor MOSFET

Ricordando la struttura fisica di un condensatore: due lamine di metallo con in mezzo uno strato di materiale isolante, osserviamo che in condizioni statiche la corrente è necessariamente nulla. Utilizziamo quindi questo principio per creare un nuovo tipo di transistor, dove la corrente di base potrà essere nulla.

Sostituendo la lamina di metallo inferiore con del materiale drogato uniformemente, ad esempio silicio drogato con concentrazione N_A , ed utilizzando come materiale isolante ad esempio l'ossido di silicio, otteniamo un condensatore MOS (*Metallo ossido semiconduttore*).

Chiameremo la regione di semiconduttore, "regione di substrato" o bulk. La regione in metallo "gate", simboleggiando la regione che controlla il flusso di carica, ed immaginiamo di applicare una tensione V_{GB} tra gate e bulk.

Immaginando che la differenza di potenziale sia $V_{GB} > 0$, le lacune interne alla regione di bulk vengono trascinate nella direzione del campo elettrico, quindi verso il basso. Creando quindi un eccesso di carica positiva sulla superficie gate, una carica negativa distribuita spazialmente nella regione di bulk, dovuta agli ioni fissi ed una carica mobile negativa superficiale, dovuta al campo elettrico.

Più è intenso il campo elettrico, maggiore è la concentrazione degli elettroni sulla superficie. Aumentare la concentrazione di elettroni porta ad aumentare la conducibilità elettrica e quindi un diminuire della resistenza.

Sotto la regione con alta concentrazione di elettroni, si ha anche una regione svuotata di portatori di carica, con alta resistività e quindi isolante.

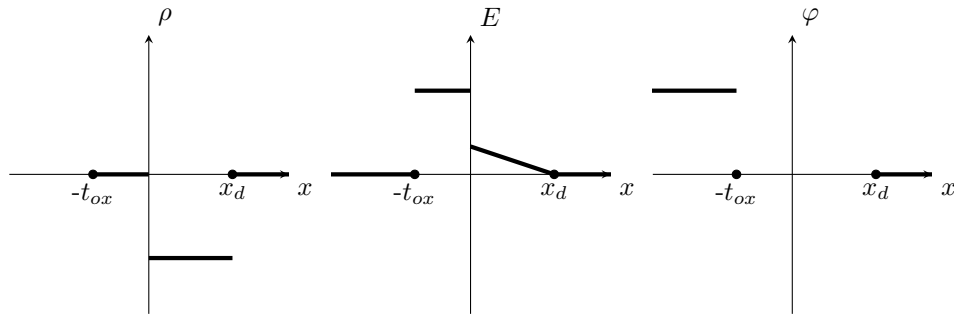
In questo caso viene ottenuto un canale n, a partire da un substrato p.

Per utilizzare questo condensatore come transistor basta applicare una differenza di potenziale longitudinale, per spostare gli elettroni presenti nel canale. La tensione V_{GB} si occupa quindi di gestire la concentrazione dei portatori, e la tensione V_T per spostare gli elettroni nel canale.

Per collegare tutti i contatti da un'unica parte del dispositivo, vengono create due regioni fortemente drogate, in modo da funzionare da conduttori per collegare i due terminali (source e drain) da dove verrà applicata la tensione V_T .

La creazione di questi due regioni di fortemente drogati si comporta come un diodo, quindi all'equilibrio sarà circondata da una regione svuotata. L'intero dispositivo quindi è completamente circondato da una regione isolata, e non necessita di isolamento esterno, si parla di dispositivo "autoisolante", caratteristica fondamentale per avere molti dispositivi in poco spazio.

Diversamente dal transistor bipolare, non c'è differenza tra source e drain.



Studio dei fenomeni di campo del mosfet

Per studiare il comportamento del transistor nmos, supponiamo l'esistenza di una regione sufficientemente lontana dall'interfaccia ossido-semiconduttore da poter descrivere la regione di semiconduttore indipendentemente dall'interfaccia.

Zona neutra

Per $x > x_d$, anche avendo applicato una differenza di potenziale, la corrente nella maglia è necessariamente nulla, in quanto non passa corrente attraverso lo strato isolante. Possiamo considerare quindi questa come una condizione di equilibrio.

$$\begin{cases} n = n_0 e^{\frac{q\varphi}{kT}} \\ p = p_0 e^{-\frac{q\varphi}{kT}} \\ pn = n_i^2 \end{cases} \Rightarrow \begin{cases} p \approx N_A \\ n = \frac{n_i^2}{N_A} \end{cases}$$

Dalla relazione della resistività: $\rho = q(N_D - N_A + p - n) = 0$. Inoltre sappiamo che la densità di corrente $J_p = q\mu_p p E - qD_p \frac{dp}{dx}$, siccome $p = N_A$ ed è costante, da cui derivata nulla, allora siccome la densità di corrente deve essere necessariamente nulla (non passa corrente), allora ne segue che il campo elettrico E è anch'esso nullo.

$$E = -\frac{d\varphi}{dx} = 0 \Rightarrow \varphi = \text{costante} = 0$$

Risulta quindi una regione neutra.

Regione metallica

Per $x < t_{ox}$, siccome è un conduttore ideale, il potenziale è costante $\varphi = \varphi_M$ ed il campo elettrico è nullo.

Indichiamo con Ψ_{MS} il potenziale di contatto, possiamo calcolare il valore del potenziale del metallo come $\varphi_M = 0 - \Psi_{MS} + V_G$. Possiamo indicarlo anche con V'_G in quanto è il valore di V_G diminuito di una costante.

In questa regione la resistività non ha significato, quindi si può evitare di calcolare.

Regione di ossido

Per $-t_{ox} < x < 0$, dall'equazione di poisson:

$$\frac{d\varepsilon E}{dx} = \rho$$

Ipotizzando che non ci sia carica interna all'ossido, possiamo considerare $\rho = 0$. (Ipotesi non utilizzabile se si tiene conto di eventuali impurità dell'ossido).

Da cui ricaviamo che il campo elettrico interno alla regione è costante, e lo indichiamo con E_{ox} .

$$\begin{aligned} E_{ox} &= -\frac{d\varphi}{dx} \\ \int_{-t_{ox}}^x E_{ox} dx &= -\int_{\varphi(-t_{ox})}^{\varphi(x)} \frac{d\varphi}{dx} dx \\ E_{ox}(x + t_{ox}) &= -(\varphi(x) - \varphi(-t_{ox})) \\ \varphi(x) &= V_G' - E_{ox}(x + t_{ox}) \end{aligned}$$

In particolare posso osservare che nell'origine è pari a $\varphi_S = V_G' - E_{ox}t_{ox}$, rendendo quindi esprimibile il campo elettrico come $E_{ox} = \frac{V_G' - \varphi_S}{t_{ox}}$.

Regione di semiconduttore

In $0 < x < x_d$, ipotizzando che, come nella giunzione pn, il potenziale abbia un andamento monotono, posso dire che $0 < \varphi(x) < \varphi_S$.

Ricordando l'espressione della concentrazione di lacune riportata in precedenza, se il potenziale è positivo, ne segue che l'esponenziale $e^{-q\varphi/kT} < 1$, in particolare moltiplicando entrambi i membri della disequazione per p_0 , ottengo che: $p(x) \ll p_0 \approx N_A$.

Analogo ragionamento si può fare per la concentrazione di elettroni, partendo dalla formula $\varphi(x) < \varphi_S$, ottenendo

$$n(x) = n_0 e^{\frac{q\varphi(x)}{kT}} \ll n_0 e^{\frac{q\varphi_S}{kT}} = n_s$$

Con n_s la concentrazione di elettroni all'origine, i.e. l'interfaccia. Ipotizzando che rimanga vera l'equazione $n(x) \ll N_A$, possiamo dire che questa regione è completamente svuotata sia dai portatori maggioritari, che dai portatori minoritari. Ottenendo $\rho = -qN_A$.

Dall'equazione di poisson otteiamo:

$$\begin{aligned} \frac{dE}{dx} &= -\frac{qN_A}{\varepsilon_S} \\ \int_x^{x_d} \frac{dE}{dx} dx &= \int_x^{x_d} -\frac{qN_A}{\varepsilon_S} dx \\ E(x) &= \frac{qN_A}{\varepsilon_S}(x_d - x) \end{aligned}$$

Si può osservare come la retta non è continua con il punto precedente, a cause dei differenti valori che assume la costante dielettrica: $\varepsilon_{ox} \approx 3.9\varepsilon_0$ e $\varepsilon_S = 11.7\varepsilon_0$. Integrando l'equazione di poisson

nell'intorno di 0, ottenendo

$$\int_{0-}^{0+} \frac{d\varepsilon E}{dx} dx = \varepsilon_S E(0^+) - \varepsilon_{ox} E_{ox} \Rightarrow E(0^+) = \frac{\varepsilon_{ox}}{\varepsilon_S} E_{ox}$$

$$\begin{aligned} E(x) &= \frac{qN_A}{\varepsilon_S} (x_d - x) = -\frac{d\varphi}{dx} \\ \int_x^{x_d} \frac{d\varphi}{dx} dx &= \int_x^{x_d} \frac{qN_A}{\varepsilon_S} (x - x_d) dx \\ \varphi(x) &= \frac{qN_A}{2\varepsilon_S} (x - x_d)^2 \end{aligned}$$

Siccome il campo elettrico è derivata del potenziale, discontinuità di campo implica una discontinuità di pendenza del potenziale.

Attraverso l'espressione ricavata è possibile calcolare il valore del potenziale all'origine φ_S , analizzando l'espressione possiamo verificare le ipotesi di validità presupposte in precedenza:

$$\begin{aligned} \varphi(x) &= \frac{qN_A}{2\varepsilon_S} x_d^2 = \varphi_S \\ x_d &= \sqrt{\frac{2\varepsilon_S \varphi_S}{qN_A}} \end{aligned}$$

La distanza x_d ha ancora dimensioni dell'ordine di micron, rendendo valida l'ipotesi di esistenza di una regione neutra non perturbata.

Studio della concentrazione di lacune ed elettroni sulla superficie del semiconduttore

Guardiamo al variare del potenziale incognito φ_S come varia la concentrazione superficiale degli elettroni e delle lacune.

Partendo dalla condizione particolare, $\varphi_S = 0$, la concentrazione $n_s = n_0 = n_i^2/N_A$, mentre $p_s = p_0 = N_A$. Questo è chiamata condizione di banda piatta, perché il potenziale si riduce ad una retta piatta sull'asse delle ascisse, in particolare $V'_G = V_G - \Psi_{MS} = 0 \Rightarrow V_G = \Psi_{MS}$.

Aumentando il potenziale, si può presupporre che la concentrazione di elettroni cresca, mentre la concentrazione di lacune diminuisca. Accadrà quindi che per un particolare valore superficiale, la concentrazione di lacune, corrisponda alla concentrazione di elettroni. Chiamiamo questo particolare valore φ_S^* .

$$\begin{aligned} n_0 e^{\frac{q\varphi_S^*}{kT}} &= p_0 e^{-\frac{q\varphi_S^*}{kT}} \\ e^{2\frac{q\varphi_S^*}{kT}} &= \left(\frac{N_A}{n_i}\right)^2 \\ \varphi_S^* &= \frac{kT}{q} \ln\left(\frac{N_A}{n_i}\right) = \varphi_F \end{aligned}$$

φ_F prende il nome di potenziale di fermi. Dove la concentrazione di elettroni e lacune si equivalgono.

Prende il nome di fenomeno di inversione, il momento in cui gli elettroni diventano i portatori maggioritari sulla superficie, rispetto alle lacune. Diversamente la condizione per cui la concentrazione di lacune rimane superiore alla concentrazione di elettroni prende il nome di regione di svuotamento.

Diventa importante, calcolare il punto in cui gli elettroni diventano i portatori maggioritari, eguagliando il numero originale di lacune, entrando quindi in una regione di forte inversione, in contrasto con la precedente regione di debole inversione. Volendo calcolare il valore di φ^* corrispondente:

$$\frac{p_0}{n_0} = e^{\frac{q\varphi_S^*}{kT}}$$

$$\varphi_S^* = 2\varphi_F$$

Ottenendo una relazione $p_S < n_0 \ll p_0 < n_S$.

In caso di potenziale negativo, la popolazione di lacune aumenta maggiormente e quella di elettroni diminuisce, entrando nella regione chiamata di "accumulazione dei portatori maggioritari".

Per calcolare il valore di $\varphi_S(V'_G)$, ovvero la concentrazione di elettroni in superficie, non si può riutilizzare l'equazione di Poisson, presupponendo la regione svuotata di portatori di carica. È necessario utilizzare un'equazione di Poisson più precisa:

$$\frac{dE}{dx} = q \frac{(N_D - N_A + p_0 e^{-\frac{q\varphi}{kT}} - n_0 e^{\frac{q\varphi}{kT}})}{\varepsilon_S} = -\frac{d^2\varphi}{dx^2}$$

Per risparmiare tempo non risolviamo l'equazione ma andiamo direttamente al risultato, ottenendo un'andamento caratterizzato da tre regioni: Una lineare, una esponenziale, ed una satura. I valori negativi, lineari corrispondono al regime di accumulazione. Il punto di transizione tra regione esponenziale e satura corrisponde a $2\varphi_F$. I valori positivi, lineari corrispondono al regime di accumulazione. Il punto di transizione tra regione esponenziale e satura corrisponde a $2\varphi_F$.

Al corrispondente potenziale di saturazione, chiamiamo la rispettiva tensione V'_T , come tensione di soglia. (Il primo è presente perchè fa riferimento a V'_G)

La regione di saturazione giustifica l'ipotesi presa in precedenza, che il valore n_S non superi eccessivamente il valore di N_A .

Possiamo dire che per valori di $V'_G < V'_T$, il canale non si è formato (off), mentre per valori superiori, indichiamo che il canale si è formato (transistor on).

Calcolo del valore della tensione di soglia

$$\begin{cases} E(0^+) = \frac{\varepsilon_{ox}}{\varepsilon_S t_{ox}} (V'_G - \varphi_S) \\ E(0^+) = \frac{qN_A}{\varepsilon_S} x_d \end{cases} \Rightarrow x_d = \sqrt{\frac{2\varepsilon_S \varphi_S}{qN_A}}$$

$$\frac{\sqrt{2qN_A\varepsilon_S\varphi_S}}{\frac{\varepsilon_{ox}}{t_{ox}}} = V'_G - \varphi_S = \gamma$$

In questa condizione $\varphi_S = 2\varphi_F$, ottenendo :

$$V'_T = \gamma\sqrt{2\varphi_F} + 2\varphi_F$$

Ricordando che $V'_T = V_T - \Psi_{MS}$:

$$V_T = \Psi_{MS} + \gamma\sqrt{2\varphi_F} + 2\varphi_F$$

Concludendo che la tensione di soglia, discrimina le tensioni di gate al di sotto della quale il canale non è formato. Possiamo vedere la formula come la somma di tre componenti: La condizione di banda piatta, la condizione di svuotamento del canale ed il raggiungimento della forte inversione.

Inserendo una carica negativa interna all'ossido, allora una delle cariche positive del gate è impiegata a bilanciare tale carica, diverse cariche negative richiamano meno cariche negative nel canale, e ritardano il processo di raggiungimento di forte inversione.

In questo caso, la tensione di soglia di un condensatore è modificabile, in alle cariche presenti nell'ossido.

Le cariche in gioco sono $Q_M = Q_i + Q_B$, carica del gate, carica di interfaccia e carica di volume.

Andamento di Q_i e Q_B in funzione di V_G

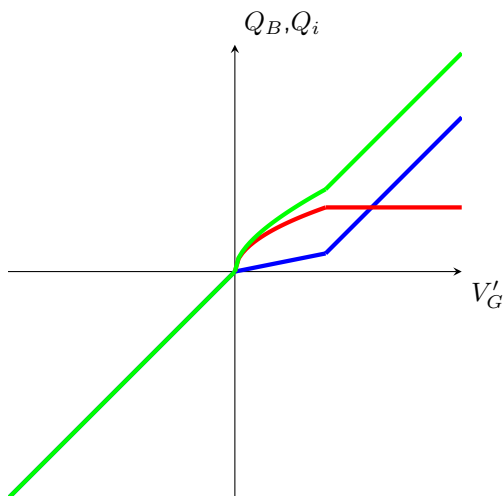
Per $V'_G < V'_T$ il potenziale $\varphi_S < 2\varphi_F$, ed è una regione svuotata di portatori mobili, quindi $n \ll N_A$.

La carica Q_i esce nell'intervallo, ma di una quantità non percepibile, mentre Q_B è data da $Q_N = -qN_Ax_dS$, considerando la superficie S unitaria, siccome x_d dipende da un andamento di radice quadrata rispetto a φ_S , e φ_S nel tratto da 0 a V'_T ha un'andamento pressochè rettilineo, quindi Q_B ha un andamento tipo radice quadrata.

Al momento di formazione del canale, per $V'_G > V'_T$, la carica Q_i cresce linearmente come un normale condensatore, in quanto ad ogni nuova carica positiva formata nella regione metallica, deve corrispondere una carica negativa associata nella regione di canale, in quanto la regione è già svuotata. L'andamento di Q_i è quindi simile a quello di un condensatore, ma non passa per l'origine, rimanendo esprimibile dalla formula:

$$Q_i = C_{ox}(V'_G - V'_T) = C_{ox}(V_G - V_T)$$

Il condensatore MOS ha quindi due regimi di funzionamento diversi, un primo quando il canale non è formato, ed un secondo, dove la carica associata allo svuotamento non cambia, e la variazione di carica avviene solo nel canale.



La capacità C della formula del condensatore mos, dipende da $C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} S$, ricordando che consideriamo la superficie S unitaria. Posso tracciare la capacità equivalente come la variazione di capacità del condensatore, la variare del potenziale, ottenendo due valori costanti nei tratti rettilinei, ed un "infossamento" nel tratto intermedio, dove cala per poi tornare costante.

Il massimo valore di capacità lo otteniamo quindi come il valore costante C_{ox} . Siccome maggiore è la capacità, maggiore è il ritardo interno al circuito. Inoltre questo transistor fornisce un modo di ottenere una capacità variabile, al variare della tensione di polarizzazione.

Oss

Sono presenti due campi interni al transistor mos, uno E_x , che si occupa della formazione del canale, ed uno E_y , che si occupa dello spostamento della corrente.

Se si scompone il condensatore infinitesimalmente, e si guarda la quantità di carica presente in ogni frammento, otteniamo che

$$Q_i = C_{ox}(V_G - \varphi(y) - V_T)$$

Con $\varphi(y)$ crescente più ci si sposta dal source al drain. Quindi la carica non è distribuita uniformemente sul condensatore, ed è più concentrata verso il source.

Per utilizzare sempre un modello monodimensionale, utilizziamo un'ipotesi di profilo graduale, ipotizzando che le variazioni di carica siano sufficientemente piccole, si può modellare il condensatore come una serie di condensatori con carica distribuita costantemente. In questo modo si può dire che il trasporto della carica, si verifica sempre con la formula utilizzata in precedenza.

Quello che si vuole calcolare è quindi l'andamento della corrente I_D in funzione di V_{GS} e V_{DS} .

Saltando direttamente alla conclusione, si dimostra che per $V_{GS} < V_T$, allora $I_D = 0$ e per $V_{GS} > V_T$, allora

$$I_D = \beta \left\{ (V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right\}$$

Il coefficiente $\beta = C_{ox}\mu_n \frac{w}{L}$, con μ_n mobilità degli elettroni, w la profondità del condensatore ed L la lunghezza del condensatore.

Per tracciare l'andamento in figura, possiamo osservare che la corrente I è esprimibile in due termini, dove il primo è una retta, dove V_{GS} funziona da coefficiente angolare, mentre il secondo è un' arco di parabola negativo.

Osservando il grafico esistono tratti con tensione positiva e corrente negativa, quindi calcolando la potenza dissipata $P = VI$ si ottiene un valore negativo, che risulta impossibile in quanto andrebbe in contraddizione col principio di conservazione dell'energia.

Osservando un modello sperimentale, in corrispondenza del punto di massimo l'approssimazione utilizzata non funziona più in quanto il valore di massimo una volta raggiunto rimane costante, e non diminuisce.

Calcolando il valore del punto di massimo:

$$\frac{dI_D}{dV_{DS}} = V_{GS} - V_T - V_{DS} = 0$$

$$V_{DS} = V_{GS} - V_T$$

Ricordando il modello di approssimazione, esprimendo la carica nel punto di massimo:

$$\bar{Q}_i = C_{ox}(V_{GS} - V_{DS} - V_T) = 0$$

Otteniamo quindi che nella sezione del punto di massimo, la carica si annulla. Questa condizione prende il nome di "pinch-off". Il raggiungimento di questa regione di pinch-off fa fallire il presupposto di svuotamento completo in quanto continua ad esistere corrente anche in assenza apparente di carica.

Studio della condizione di pinch-off

Il fatto che la carica $Q_i(L) \rightarrow 0$, siccome la corrente deve rimanere costante, è necessario che la velocità degli elettroni $v_n \rightarrow \infty$, che risulterebbe assurdo. Ricordando che la velocità degli elettroni è $V_n = -\mu_n E$ allora il campo elettrico dovrebbe tendere all'infinito. Per poi tornare a valori bassi una volta raggiunta la zona drogata n .

$$E_y = -\frac{d\varphi}{dy} \Rightarrow \int_0^L E_y dy = -\int_{\varphi(0)}^{\varphi(L)} \frac{d\varphi}{dy} dy$$

Il primo integrale è interpretabile come l'area sottesa del campo elettrico da 0 ad L . Il secondo come la differenza di potenziale ai capi del canale, pari a V_{DS} . Siccome V_{DS} è un valore finito, allora l'area sottesa dalla curva è anch'essa finita. Quindi la lunghezza del tratto di campo elettrico che tende all'infinito deve necessariamente tendere a 0. Indipendentemente dal valore di V_{DS} .

Ci si può immaginare quindi il canale come un partitore. Dove la regione con il canale formato è esprimibile da una resistenza. Mentre la regione strozzata, di "pinch-off" con un'altra resistenza, e siccome è presente poca carica, la conducibilità è piccola e la resistenza equivalente quindi è superiore alla precedente. Di conseguenza, dato che la resistenza di pinch-off è molto superiore alla resistenza di canale, ogni ulteriore variazione verrà assorbita principalmente dalla regione di pinch-off, non ripercuotendosi sulla resistenza di canale.

OFF

$$\begin{cases} V_{GS} < V_T \\ I_D = 0 \end{cases}$$

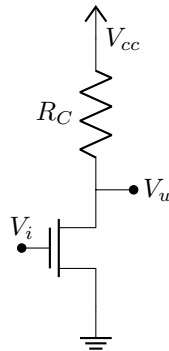
Lineare

$$\begin{cases} V_{GS} > V_{DS} + V_T \\ I_D = \beta_n \left\{ (V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right\} \end{cases}$$

Saturazione

$$\begin{cases} V_T < V_{GS} < V_{DS} + V_T \\ I_D = \frac{\beta_n}{2} (V_{GS} - V_T)^2 \end{cases}$$

Esempio circuito



Per $V_i < V_{GS}$ il transistor è spento, quindi $I_D = 0$, da cui:

$$\begin{cases} V_u = V_{cc} \\ V_i < V_{GS} \end{cases}$$

Supponendo NMOS saturo:

$$\begin{cases} V_i > V_T \\ V_u > V_i - V_T \end{cases} \Rightarrow V_T < V_i < V_u + V_T$$

La regione di validità è sopra la retta $V_i = V_T - V_u$. La formula che descrive V_u è un arco di parabola decrescente:

$$V_u = V_{cc} - \frac{R\beta}{2} (V_i - V_T)^2$$

In regione di funzionamento lineare:

$$\begin{cases} V_i > V_T \\ V_u < V_i - V_T \end{cases}$$

Ed il valore della tensione in uscita è descritto dalla formula:

$$V_u = V_{cc} - R\beta \left\{ (V_i - V_T) V_u - \frac{V_u^2}{2} \right\}$$

Dalle condizioni di esistenza e dall'espressione della corrente, possiamo notare come in questo tratto, $I_D > 0$.

Natura ratioed

In regione di saturazione, il guadagno A_V è descritto dalla formula:

$$A_V = -\beta R(V_i - V_T)$$

e dipende dal punto V_i in cui viene calcolato, e dal valore di β , ovvero dalle caratteristiche specifiche (ϵ_{ox} , t_{ox} , μ_n , w ed L) del transistor.

In termini progettuali è possibile variare le dimensioni fisiche del canale w ed L , ed il valore della resistenza R , dipendente da rispettiva lunghezza L_R e ampiezza w_R .

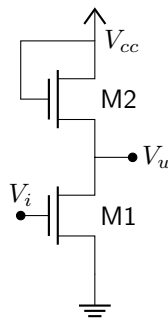
$$\beta R = c_{ox} \mu_n \frac{\rho}{t} \frac{w_M}{L_M} \frac{L_R}{w_R} = k \frac{w_M/L_M}{w_R/L_R}$$

Siccome il guadagno dipende dal rapporto tra i fattori di forma (w/L), l'invertitore è detto di tipo "ratioed" o dimensionato, in quanto per variare il valore del guadagno per aumentare l'immunità ai disturbi, è richiesto di cambiare le dimensioni dell'nmos.

Nel circuito il transistor prende il nome di rete di pull-down e la resistenza prende il nome di pull-up.

NMOS a carico saturo

Sostituendo alla rete di pull-up un secondo transistor, in modo da simulare una resistenza con più alti valori resistivi otteniamo il seguente circuito:



Il transistor nella rete di pull-up è detto connesso a diodo: facendo riferimento al transistor bipolare, il cortocircuito tra le giunzioni p ed n lo fa comportare come un diodo.

Questa connessione è caratterizzata da $V_{GS} = V_{DS}$, facendo sì che quando il transistor è acceso, funziona solo in regione di saturazione.

Per questa ragione l'invertitore prende il nome di *nmos a carico saturo*.

Analisi del modello

Per M_1 OFF

$$\begin{cases} V_{GS1} = V_i < V_T \\ I_{D1} = I_{D2} = 0 \end{cases}$$

Siccome la corrente è nulla ed M_2 se acceso può funzionare solo in regione lineare, di conseguenza M_2 è spento.

$$\begin{cases} V_i < V_T \\ V_u = V_{cc} - V_{DS2} = V_{CC} - V_T \end{cases}$$

Per M_1 saturo

$$\begin{cases} V_i > V_T \\ V_u > V_i - V_T \\ I_{D1} = \frac{\beta_1}{2}(V_i - V_T)^2 \end{cases}$$

Dato che la corrente I_{D1} è positiva, M_2 necessariamente è in regione di saturazione:

$$\frac{\beta_1}{2}(V_i - V_T)^2 = \frac{\beta_2}{2}(V_{cc} - V_u - V_T)^2$$

Ottenendo la relazione:

$$V_u = V_{cc} - V_T - \sqrt{\frac{\beta_1}{\beta_2}}(V_i - V_T)$$

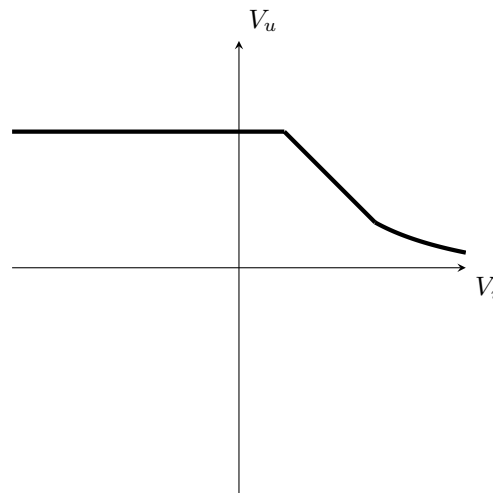
Per M_1 lineare

$$\begin{cases} I_{D1} = \beta_1 \left((V_i - V_T)V_u - \frac{V_u^2}{2} \right) \\ V_i > V_T \\ V_u < V_i - V_T \end{cases}$$

Ancora una volta la corrente I_{D1} è positiva, quindi M_2 si trova in regione di saturazione:

$$I_{D1} = \beta_1 \left\{ (V_i - V_T)V_u - \frac{V_u^2}{2} \right\} = \frac{\beta_2}{2}(V_{cc} - V_u - V_T)^2 = I_{D2}$$

Che ci risparmiamo i calcoli dicendo che risulta in un tratto ad andamento decrescente.



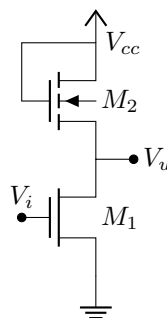
Osservazioni sul circuito

Come visibile nella regione di saturazione di M_1 , il rapporto tra i coefficienti β_1 e β_2 influenza direttamente il guadagno. Di conseguenza anche questo circuito è di tipo ratioed.

Modifica rete di pull-up con transistor a depletion

È possibile modificare il transistor mos, cambiando il drogaggio del semiconduttore, aggiungendo uno strato drogato con atomi donatori, in modo da rendere il canale già formato in partenza ed utilizzare la tensione V_{GB} per svuotarlo. Questo tipo di transistor prende il nome di transistor a depletion, mentre quello utilizzato fino ad ora prende il nome di transistor ad enhancement.

Connettendo a diodo questo nuovo tipo di transistor, è vera la condizione di funzionamento lineare: $V_{GS} > V_{DS} - |V_T|$, mentre la condizione di saturazione non è permessa.



M_2 è spento se $V_{GS2} = V_{cc} - V_u < V_{T2}$, ovvero $V_u > V_{cc} + |V_{T2}|$. Quindi se la tensione di uscita supera il valore di alimentazione. Non potendosi verificare questa condizione, M_2 funziona unicamente in regione lineare.

Per M_1 spento, allora $I_{D2} = 0$, quindi $V_{DS2} = 0$ quindi $V_u = V_{cc}$

Per M_1 in regione di saturazione si ottiene ancora una relazione analoga alla precedente

$$I_{D2} = \beta_2 \left\{ (V_{cc} - V_u - V_T)(V_{cc} - V_u) - \frac{(V_{cc} - V_u)^2}{2} \right\}$$

$$= \beta_2 \left(\frac{x^2}{2} + x|V_{T2}| \right)$$

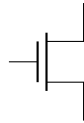
Questo circuito a differenza del precedente, ha il vantaggio di avere al valore alto lo stesso valore di V_{cc} .

Transistore pmos

L'unica differenza è che il semiconduttore è drogato con atomi donatori N_D , il canale è formato dalle lacune, essendo portatori di carica positiva, il potenziale più basso è al drain e non al source. Il funzionamento rimane lo stesso.

La mobilità delle lacune è inferiore di quella degli elettroni, risultando in un coefficiente β_p peggiore rispetto a quello di β_n

NMOS



OFF: $V_{GS} < V_{Tn}$ e $I_D = 0$

SAT:

$$V_{Tn} < V_{GS} < V_{DS} + V_{Tn}$$

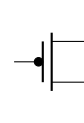
$$I_D = \frac{\beta_n}{2} (V_{GS} - V_{Tn})^2$$

LIN:

$$V_{GS} > V_{DS} + V_{Tn}$$

$$I_D = \beta_n \left\{ (V_{GS} - V_{Tn})V_{DS} - \frac{V_{DS}^2}{2} \right\}$$

PMOS



OFF: $V_{GS} > V_{Tp} < 0$ e $I_D = 0$

SAT:

$$V_{DS} + v_{Tp} < V_{GS} < V_{Tp}$$

$$I_D = \frac{\beta_p}{2} (V_{GS} - V_{Tp})^2$$

LIN:

$$V_{GS} < V_{DS} + V_{Tp}$$

$$I_D = \beta_p \left\{ (V_{GS} - V_{Tp})V_{DS} - \frac{V_{DS}^2}{2} \right\}$$

Limitandoci ad analizzare i transistori pmos ad arricchimento, ovvero $V_{Tp} < 0$, allora è possibile esprimere le equazioni come:

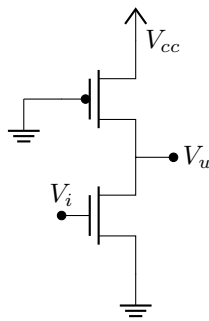
OFF: $V_{SG} < |V_{Tp}|$ e $I_D = 0$

SAT: $V_{DS} + |V_{Tp}| > V_{SG} > |V_{Tp}|$ e $I_D = \frac{\beta_p}{2} (V_{SG} - |V_{Tp}|)^2$

LIN: $V_{SG} > V_{SD} + |V_{Tp}|$ e $I_D = \beta_p \left\{ (V_{SD} - |V_{Tp}|)V_{SD} - \frac{V_{SD}^2}{2} \right\}$

La struttura delle equazioni rimane quindi essenzialmente la stessa, variano solamente i componenti.

Analizzando il circuito con un transistore



La rete di pull-down rimane la stessa, ed è spenta per $V_i < V_{Tn}$ ed è saturo per $V_u > V_i - V_{Tn}$ e lineare per $V_u < V_i - V_{Tn}$.

Diversamente il transistor pmos non è mai spento in quanto $V_{cc} > |V_{Tp}|$. È saturo quando $V_{cc} < V_{cc} - V_u + |V_{Tp}|$ ovvero $V_u < |V_{Tp}|$.

Se l'nmos è spento, allora la corrente è nulla. la corrente sul ramo del transistor pmos è nulla se il transistor si trova in regione lineare o è spento. Quindi da precedente osservazione l'nmos funziona in regione lineare. Ponendo a zero la corrente in regione lineare, allora otteniamo che $V_{SD} = 0$, quindi $V_u = V_{cc}$

Nella regione in cui l'nmos è saturo ed il pmos è lineare, mettiamo a sistema le equazioni delle due correnti ottenendo:

$$(V_{cc} - V_u)^2 - 2(V_{cc} - |V_{Tp}|)(V_{cc} - V_u) + \frac{\beta_n}{\beta_p}(V_i - V_{Tn})^2 = 0$$

Da cui:

$$V_u = |V_{Tp}| \pm \sqrt{(V_{cc} - |V_{Tp}|)^2 - \frac{\beta_n}{\beta_p}(V_i - V_{Tn})^2}$$

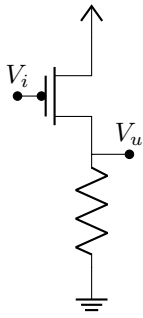
Per $V_i = V_{Tn}$, siccome sappiamo che l'uscita V_u deve essere uguale a V_{cc} per continuità, allora necessariamente la radice con soluzione deve essere positiva.

Risolvendo graficamente l'equazione si ottiene sempre il grafico di un'invertitore, dove la tensione d'uscita tende asintoticamente al valore 0. Nuovamente il guadagno di tensione è funzione dipendente da β_n e β_p .

Questo invertitore prende il nome di invertitore pseudo-nmos, per via che il transistor a canale p è utilizzato come rete di pull-up.

Per utilizzare transistori nmos e pmos nello stesso circuito, dato che richiedono substrati con drogaggi di tipo differente, è necessario tenere in mente che la cosa che differenzia una regione drogata di tipo p ed una di tipo n è la differenza di atomi drogati. È possibile drogare selettivamente parti di un materiale già drogato.

Nello stesso substrato è quindi possibile creare transistori di tipo n e di tipo p, questa tecnologia prende il nome di CMOS (*Complementary MOS*)



pmos è spento per $V_{SG} < |V_{Tp}|$, quindi $V_i > V_{cc} - |V_{Tp}|$ ed $I_D = 0$ quindi $V_u = 0$.

In saturazione $V_u < V_i + |V_{Tp}|$ e

$$V_u = \frac{\beta_p R}{2} (V_{cc} - V_i - |V_{Tp}|)^2$$

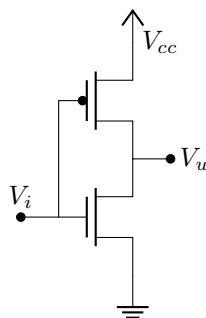
$I_D = \frac{V_{cc} - V_{SD}}{R}$, intersecata con il grafico della caratteristica, si osserva che all'aumentare di V_{SG} , la tensione V_{SD} tende asintoticamente a 0, quindi V_u tende a V_{cc}

Confrontando i risultati ottenuti con l'nmos, per ingresso $V_i = V_L$, l'uscita dell'nmos si porta a V_H indipendentemente dai fattori di forma e la corrente è nulla. Diversamente per uscita $V_u = V_H$ l'uscita dell'nmos dipende dai fattori di forma e la corrente è positiva. Per l'nmos quindi il comportamento è ideale per $V_i = V_L$, con potenza dissipata nulla, ha una potenza statica solamente per $V_u = V_H$. L'uscita alta è la migliore possibile.

Il pmos è l'esatto opposto portando all'uscita bassa $V_i = 0$, la potenza statica dissipata per tenere il valore basso in uscita è nulla. La tensione alta del pmos non raggiunge mai il valore massimo ed il valore dipende dal fattore di forma, dissipando potenza statica.

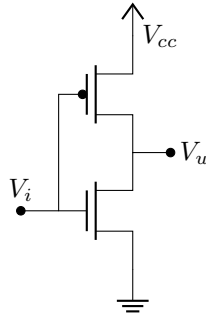
L'nmos è un pull-down attivo intelligente in quando non serve si spegne, non dissipando potenza. Il pmos è una rete di pull-up intelligente.

Si può intuire quindi che utilizzando i due componenti nella loro rete ideale otterremo migliori prestazioni, rendendo il risultato indipendente dal fattore di forma (ratioless). Questo circuito prende il nome di invertitore cmos.

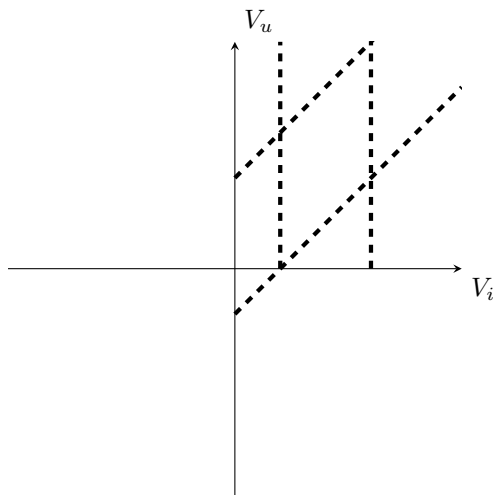


Tecnologia cmos

Facendo riferimento all'invertitore nmos citato al capitolo precedente:



Il transistor a canale n ke' spento per $V_i < V_{Tn}$ e saturo quando $V_i < V_u + V_{Tn}$



In regione di saturazione: $V_{SGp} < V_{SDp} + |V_{Tp}|$ da cui $V_n < V_i + |V_{Tp}|$

In caso di nmos spento, $I_{DP} = 0$, quindi il pmos funziona in regione lineare o è spento. Avendo ipotizzato che $V_{cc} - |V_{Tp}| > V_{Tn}$, allora non esiste regione del piano in cui entrambi i transistor sono simultaneamente spenti. Quindi il transistor p deve essere in lineare.

$$I_{DP} = \beta_p \left\{ V_{SGp} - |V_{Tp}| \right\} V_{SDp} - \frac{V_{SDp}^2}{2} \Bigg\} = \beta_p V_{SDp} \left(V_{SGp} - |V_{Tp}| - \frac{V_{SDp}}{2} \right)$$

Quindi $V_{SDp} = 0$ e $V_u = V_{cc}$.

in caso di pmos spento: $I_{Dp} = 0$, quindi necessariamente l'nmos è spento o lineare, ma siccome è già stato detto che entrambi i transistori non possono essere spenti allo stesso tempo, allora n è in lineare. Quindi

$$\beta_n V_u \left\{ V_i - V_{Tn} - \frac{V_u}{2} \right\} = 0$$

Da cui $V_u = 0$

Allora il valore massimo è il massimo possibile e non c'è corrente statica, ed il valore basso è il minimo possibile ed ancora non c'è corrente statica.

nella regione in cui entrambi i transistori sono in regione di saturazione: (prese solo le radici positive per ipotesi)

$$\sqrt{\frac{\beta_n}{\beta_p}} (V_i - V_{Tn}) = V_{cc} - V_i - |V_{Tp}|$$

In questa espressione non compare V_u , quindi l'unica soluzione possibile è un V_i costante

$$V_i = \frac{V_{cc} - |V_{Tp}| + \theta V_{Tn}}{1 + \theta}$$

Ottenendo un guadagno $|A_V| \rightarrow \infty$! Questo è il motivo per cui prende il nome di ratio-less.

Indipendentemente dalla dimensione dei dispositivi si comporta comunque come un invertitore.

Caso particolare

Supponiamo la condizione di complementarietà perfetta: $\beta_n = \beta_p$ e $V_{Tn} = |V_{Tp}|$, quindi $\theta = 1$

$$\bar{V}_i = \frac{V_{dd}}{2}$$

Ottenendo il massimo valore di margine di immunità ai disturbi.

Calcolo del margine di immunità ai disturbi

Per calcolare il guadagno è necessario calcolarlo nella curva in cui si passa da guadagno 0 a guadagno infinito. Utilizzando il caso particolare, calcoliamo le coordinate dei punti della caratteristica in cui la derivata è -1.

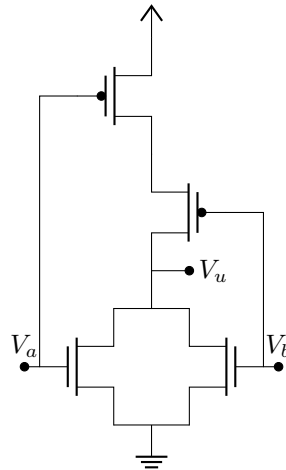
Siccome n è in regione lineare e p è in saturazione, allora:

$$\begin{aligned} \beta \left\{ (V_i - V_T) V_u - \frac{V_u^2}{2} \right\} &= \frac{\beta_2}{2} (V_{cc} - V_i - V_T)^2 \\ V_u + (V_i - V_T) \frac{dV_u}{dV_i} - V_u \frac{dV_u}{dV_i} &= (V_{cc} - V_i - V_T) \\ 2V_u - V_i + V_T &= -V_{cc} + V_i + V_T \\ V_u &= V_i - \frac{V_{cc}}{2} \end{aligned}$$

Intersecando la condizione della retta ottenuta con la curva iniziale, otteniamo il valore di V_{IHMIN} e V_{OLMAX} .

Porta NOR

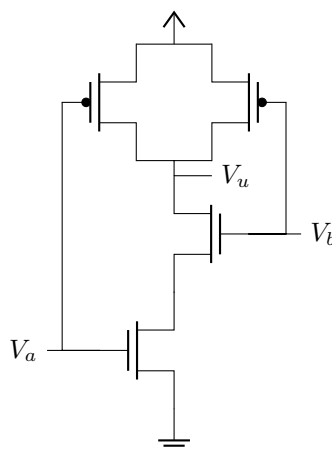
Dall'espressione logica $y = \overline{a + b}$ sappiamo che la rete di pull-up deve portare l'uscita ad 1, quando sia a che b sono a 0, mentre la rete di pull-down deve portare l'uscita a 0 quando almeno uno dei due ingressi è al valore alto.



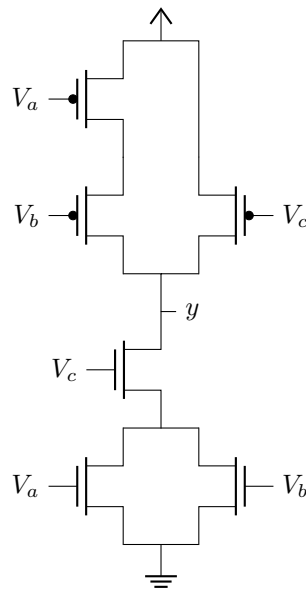
a	b	M1	M2	M3	M4	PD	PU	y
0	0	OFF	OFF	ON	ON	OFF	ON	1
0	1	OFF	ON	OFF	ON	ON	OFF	0
1	0	ON	OFF	ON	OFF	ON	OFF	0
1	1	ON	ON	OFF	OFF	ON	OFF	0

Porta nand

Analogamente, dall'espressione, logica $y = \overline{a \cdot b}$, il pull-up deve portare l'uscita ad 1 quando almeno 1 tra a e b è a 0, mentre la rete di pull down porta l'uscita a 0 quando entrambi gli ingressi sono ad 1

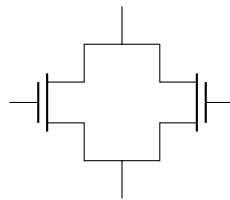


Inoltre bisogna notare che non sempre è ottimale esprimere le espressioni in funzione di nand o nor, ma è possibile fare la funzione direttamente con i transistori, risparmiando componenti, ad esempio $y = \overline{(a + b)c}$:



Risparmiando un numero di transistori inferiore alla metà. Con questo tipo di logica, per n ingressi, sono richiesti $2n$ transistori.

Transistori in parallelo



Ai due transistori connessi in parallelo, è possibile sostituire un'unico transistorore con stessa tensione di soglia V_T ed un fattore β_{eq} calcolabile come:

$$I_{Deq} = (\beta_1 + \beta_2) \left\{ (V_{GS} - V_T) V_{DS} \frac{V_{DS}^2}{2} \right\}$$

$$\beta_{eq} = \beta_1 + \beta_2$$

Analogamente se si dispongono due transistori, caratterizzati dalla stessa tensione di soglia V_T , se la tensione in ingresso è equivalente, allora è possibile sostituire ai due transistori un singolo transistorore con

$$\beta_{eq} = \frac{\beta_1 \beta_2}{\beta_1 + \beta_2}$$

Nel caso $\beta_1 = \beta_2$: $\beta_{eq} = \beta_2$

Una qualsiasi rete cmos è quindi sempre riconducibile ad un invertitore cmos equivalente.

In una rete cmos se un transistorore è spento, quello conta come un circuito staccato.

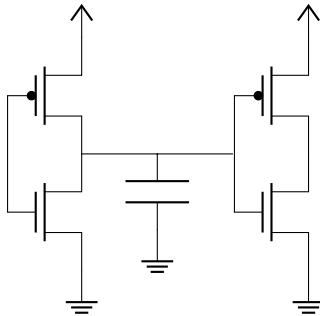
Prestazioni e qualità dell'invertitore

Il circuito non presenta potenza statica, ma una dinamica, in quanto la corrente è diversa da 0 al momento del passaggio di stato. Inoltre in quanto è presente uno spostamento di carica è presente anche un ritardo.

Per calcolare questi valori, prendiamo la capacità peggiore C_o del transistor. Il modello del transistor MOS deve tenere conto anche di altre complicazioni, come la presenza di una capacità parassita non lineare tra source e bulk e tra bulk e drain. Una tra source e gate ed una tra gate e drain. In altre parole ogni collegamento interno ad un circuito contiene una rispettiva capacità parassita tra i due nodi.

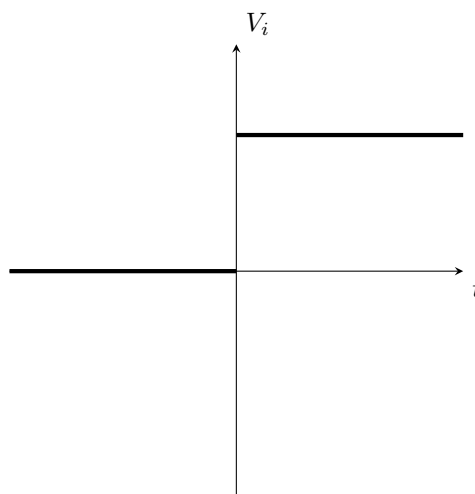
Per semplificare i calcoli, diciamo che queste capacità sono correlate alla capacità del condensatore C_o .

Calcolo del ritardo tra due invertitori in cascata



È possibile concentrare tutte le capacità parassite dei condensatori, in un'unica capacità, situata nella connessione tra i due invertitori. La capacità semplificata dipende da $C_L = KC_0 + C_{wire}$, ovvero parte delle capacità parassite dei transistori e la capacità parassita del filo. Il termine prevalente è kC_0

Presupponiamo che la tensione in ingresso abbia un andamento del tipo:



per $t < 0$ e $V_i = 0$ siamo in condizione statica, ed è già nota dalla caratteristica: $V_u = V_{cc}$.

Per $t > 0$ $V_i = V_{cc}$, presupponendo un tempo infinito, in cui si annullano gli effetti del transitorio, conosciamo già i valori della caratteristica: $V_u = 0$.

Dato che serve un tempo infinito per raggiungere il valore $V_u = 0$, consideriamo il tempo di propagazione, il tempo necessario per portare l'uscita $V_u = V_{cc}/2$.

Per $t = 0^+$, $V_{GS} = V_i = V_{cc} > V_T$, quindi l'nmos è necessariamente acceso, mentre $V_{SGp} = V_{cc} - V_i = 0 < |V_{Tp}|$ il pmos si spegne immediatamente.

Quindi la corrente dell'nmos I_D è uguale alla corrente che passa dal condensatore.

$$I_{DN} = -I_C = -C_L \frac{dV_u}{dt}$$

Siccome la formula della corrente varia in base al regime di funzionamento dell'nmos, il transitorio può essere scomposto in due tratti: da V_{cc} a $V_{cc} - V_T$ in regione satura, ed un secondo tratto il regione lineare fino a raggiungere il valore $V_{cc}/2$

Quindi nel tratto in regione satura:

$$I_D = \frac{\beta_n}{2} (V_{GS} - V_T)^2 = -C_L \frac{dV_u}{dt}$$

$$\int_0^t \frac{\beta_n}{2} (V_{GS} - V_T)^2 dt = \int_{V_u(0)=V_{cc}}^{V_u(t)} -C_L \frac{dV_u}{dt} dt$$

Ottenendo un'andamento lineare decrescente

$$V_u(t) = V_{cc} - \frac{\beta_n}{2C} (V_{cc} - V_T)^2 t$$

Ponendo la formula uguale a $V_{cc} - V_T$ ricavo la durata del primo transitorio:

$$t_1 = \frac{2C_L}{\beta_n} \frac{V_T}{(V_{cc} - V_T)^2}$$

Il secondo tratto del transitorio è caratterizzato dalla regione lineare, valido da $V_{cc} - V_T$ a $\frac{V_{cc}}{2}$. Risolvendo l'equazione utilizzando la diversa espressione della corrente ottengo ($m = 2(V_{cc} - V_T)$)

$$\frac{2C_L}{\beta} \frac{dV_u}{dt} = V_u(V_u - m)$$

$$\frac{2C_L}{\beta} \frac{1}{V_u(V_u - m)} \frac{dV_u}{dt} = 1$$

$$\int_{t_1}^{t_{phl}} \frac{2C_L}{\beta} \frac{1}{V_u(V_u - m)} \frac{dV_u}{dt} dt = \int_{t_1}^{t_{phl}} dt$$

$$\frac{2C_L}{\beta_n} \frac{1}{2(V_{cc} - V_T)} \ln \left(3 - 4 \frac{V_T}{V_{cc}} \right) = t_{phl} - t_1$$

Mettendo a sistema l'equazione appena ottenuta con t_1 calcolato in precedenza, otteniamo:

$$t_{phl} = \frac{2C_L}{\beta_n(V_{cc} - V_T)} \left\{ \frac{V_T}{(V_{cc} - V_T)} + \frac{1}{2} \ln \left(3 - 4 \frac{V_T}{V_{cc}} \right) \right\}$$

Ipotizzando che $V_{cc} \gg V_T$ l'espressione si semplifica notevolmente in

$$\frac{C_L}{\beta_n V_{cc}} \ln(3) \approx \frac{C_L}{\beta_n V_{cc}}$$

La capacità C_L che è una tra le componenti principali del ritardo è composta da C_{MOS} , capacità parassita legata al transistor e da C_{wire} , capacità parassita del filo. Prendendo come presupposto che $C_{MOS} \gg C_{wire}$, allora possiamo riscrivere l'espressione del tempo di propagazione come:

$$t_p = \frac{k C_{ox} w L}{c_{ox} \mu_n \frac{w}{L} V_{cc}} = \frac{k}{\mu_n} \frac{L^2}{V_{cc}}$$

Da cui possiamo osservare come sia la capacità, sia il fattore w non continuo nell'espressione del tempo di reazione. Si osserva quindi che la componente principale è la lunghezza del canale L . Una volta ridotta il più possibile la lunghezza L , l'unico parametro rimanente è la tensione V_{cc} , aumentando la tensione però ci si può aspettare un aumento di potenza.

Diversamente se la capacità parassita del filo prevale sulla capacità interna si sviluppano altri termini progettuali attraverso circuiti di buffer (spiegato dopo).

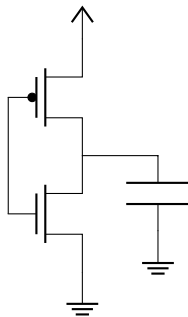
Consumo di potenza

Abbiamo già discusso il fatto che la potenza statica dell'invertitore è nulla. Ma la corrente di polarizzazione inversa, interna al transistor, seppur piccola, può essere significativa in circuiti "fermi" risultando, in circuiti complessi come componente fondamentale della potenza statica.

La dispersione principe dei dispositivi che analizziamo in questo corso è la potenza dinamica. Siccome non ha senso calcolare la potenza in un determinato istante, faremo riferimento ad una potenza media:

$$\bar{P} = \frac{1}{T} \int_0^T P_{dd} dt$$

P_{dd} è la potenza istantanea erogata dal generatore e corrisponde a $P_{dd} = V_{dd} I_D$



Dal circuito possiamo osservare come la corrente di pull-up $I_D = I_C + I_{Dn}$. Passando da un valore di uscita basso, ad un valore di uscita alto il valore della corrente assume un valore diverso da 0, quindi anche in assenza del condensatore si dissipa potenza solamente per il fatto di avere per un'istante entrambi i condensatori accesi. La potenza dissipata in questo modo prende il nome di potenza di cortocircuito P_{cc} , con rispettiva corrente I_{cc} . L'altra componente della potenza dissipata è quella

associata al carico, utilizzata per caricare il condensatore parassita C_L , prendendo il nome di potenza di carico P_L .

Si può dimostrare che P_{cc} dipende solo da quando il segnale d'ingresso è compreso tra $V_T < V_i < V_{dd} - V_T$, assumendo la forma di

$$\overline{P_{cc}} = \frac{\beta}{12} \frac{t_R}{T} (V_{dd} - 2V_T)^3$$

Di conseguenza se il tempo impiegato dal segnale d'ingresso per portarsi al valore alto tende a 0, allora anche $P_{cc} \rightarrow 0$.

Per calcolare solamente la componente di potenza di carico, basta annullare la potenza di cortocircuito, assumendo che il segnale d'ingresso abbia un andamento a gradino.

Prendendo come riferimento un segnale d'ingresso ad'onda quadrata con periodo T , nel primo periodo da 0 a $T/2$, il segnale d'ingresso è alto, quindi $V_i = V_{cc}$, di conseguenza l'nmos è acceso ed il pmos è spento. Nel secondo periodo il segnale d'ingresso è basso, quindi l'nmos è spento ed il pmos è acceso.

Quindi

$$\overline{P_L} = \frac{1}{T} \int_0^T V_{dd} (I_{dn} + I_C) dt = \overline{P_n} + \overline{P_p} + \overline{P_C}$$

osservando che il sistema è periodico, tutta la potenza assorbita dal circuito in un periodo, deve essere dissipata dal circuito nel periodo successivo, quindi la potenza media è composta dalla potenza dei singoli componenti del circuito: nmos, pmos e condensatore

$$\overline{P_C} = \frac{1}{T} \int_0^T V_u I_C dt = \frac{1}{T} \int_0^T V_u C \frac{dV_u}{dt} dt = \frac{C}{T} \left\{ \frac{V_{dd}^2}{2} - \frac{V_{dd}^2}{2} \right\} = 0$$

Nella fase di pull-up il condensatore si carica e nella fase di pulldown rilascia energia, risultando in un bilancio energetico nullo.

$$\overline{P_n} = \frac{1}{T} \int_0^T V_{DS} I_D dt = \frac{1}{T} \int_0^{T/2} V_u (-C) \frac{dV_u}{dt} dt = -\frac{C}{T} \left\{ \frac{0^2}{2} - \frac{V_{dd}^2}{2} \right\} = \frac{C}{T} \frac{V_{dd}^2}{2}$$

$$\overline{P_p} = \frac{1}{T} \int_0^T V_{SD} I_D dt = \frac{1}{T} \int_{T/2}^T (V_{dd} - V_u) I_D dt = \frac{C}{T} \frac{V_{dd}^2}{2}$$

Concludendo che l'energia nella fase di pull-up è per metà dissipata e per metà immagazzinata nel condensatore, nella fase di pull-down l'energia immagazzinata viene dissipata attraverso la potenza richiesta dalla rete di pull-down nell'altra fase del transitorio.

$$\overline{P_L} = \frac{C_L V_{dd}^2}{T} = C_L V_{dd}^2 f$$

aumentando V_{dd} il circuito è quindi più veloce ma spende molta più potenza. Inoltre maggiore è la frequenza, maggiore è l'energia dissipata. Per questo motivo a bassi livelli di energia le prestazioni dei calcolatori diminuiscono.

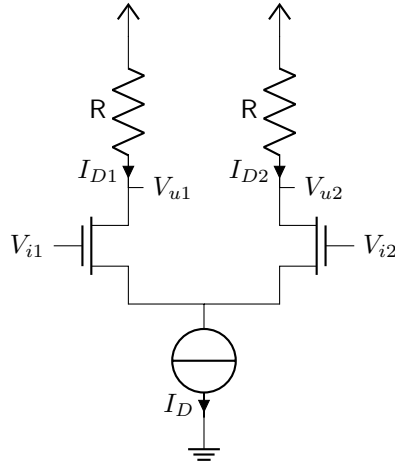
Nel confronto tra le due componenti di potenza, P_{cc} e P_L , quello che si vede facilmente è che aumentando la frequenza, tipicamente il termine P_L prevale su P_{cc} .

Il prodotto tra il tempo richiesto per effettuare un'operazione e la potenza media richiesta, può essere interpretato come l'energia richiesta da una singola operazione:

$$t_p \overline{P_L} = \frac{C_L}{\beta V_{dd}} \frac{C_L V_{dd}^2}{T} = \frac{C_L^2 V_{dd}}{\beta T}$$

Diminuire la capacità ha un effetto benefico su entrambi i parametri.

Amplificatore differenziale



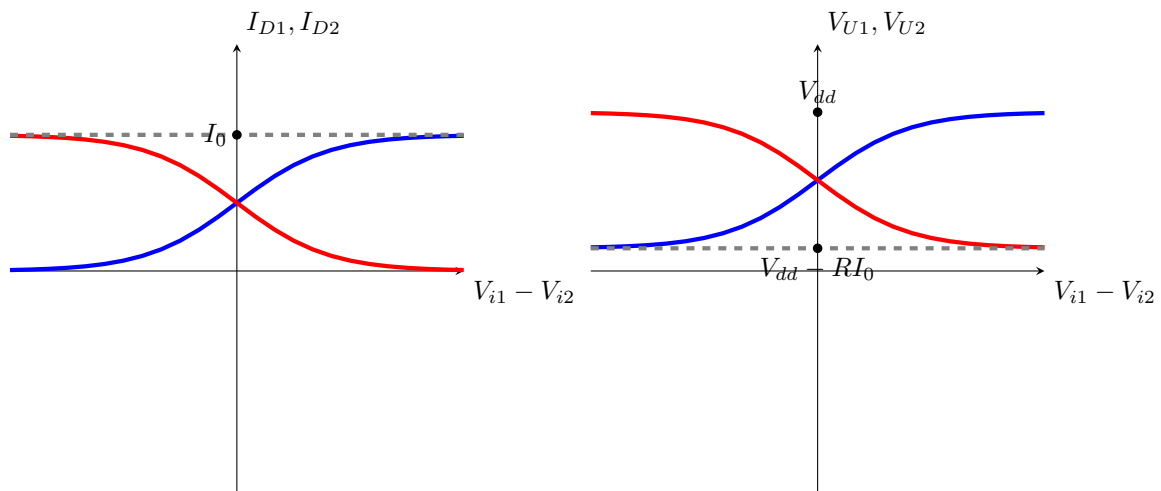
Da Kirkoff $I_{D1} + I_{D2} = I_D$

$$V_{i1} - V_{i2} = V_{GS1} - V_{GS2}$$

Questo ti porta a dire che se $V_{i1} > V_{i2}$, allora $V_{GS1} > V_{GS2}$ e di conseguenza $I_{D1} > I_{D2}$. E siccome la tensione di uscita è scritta nella forma $V_u = V_{dd} - RI_D$, ne segue che $V_{u1} < V_{u2}$.

Se le due tensioni di ingresso sono identiche, non c'è motivo di pensare che le correnti sui due rami non siano identiche, siccome il circuito è simmetrico. Diversamente se una delle due tensioni è maggiore ci si può aspettare che la corrente maggiore vada dal ramo con tensione maggiore.

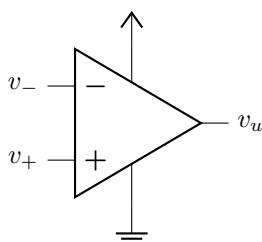
Questo succede fino al limite in cui tutta la corrente circola su un unico ramo del circuito.



Questo circuito è un amplificatore che varia con la differenza dei segnali d'ingresso. Chiamando $V_{i1} - V_{i2} = V_{id}$ e $\frac{V_{i1} + V_{i2}}{2} = V_{ic}$, posso sempre determinare i valori V_{i1} e V_{i2} attraverso differenza e valore medio. La media degli ingressi prende il nome di "componente di modo comune", ed è indicata con V_{ic} . È possibile definire un guadagno $A_d = \frac{dV_u}{dV_{id}}$, calcolando lo stesso guadagno per le

componenti di modo comune: Tenendo i due segnali identici in ingresso, si ottiene: $I_{D1} = I_{D2} = \frac{I_0}{2}$, indipendentemente dal valore di corrente $V_{u1} = V_{u2} = k$. Da cui $A_c = \frac{dV_u}{dV_{ic}} = 0$, quindi questo circuito è in grado di amplificare la componente differenziale, ignorando completamente le variazioni in modo comune.

Considerazioni su opamp

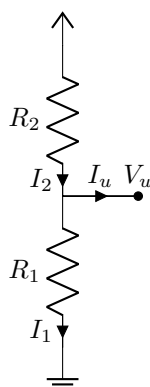


È possibile descrivere la relazione in uscita della funzione a gradino in tre tratti differenti, una regione centrale, in cui il guadagno tende all'infinito, descritto dall'equazione di validità $V_{id=0}$ per $-V_M < V_u < V_M$; un tratto di saturazione positiva, vero per $V_{id} > 0$, dove $V_u = V_M$, ed un tratto di saturazione negativa, $V_u = -V_M$, vero per $V_{id} < 0$.

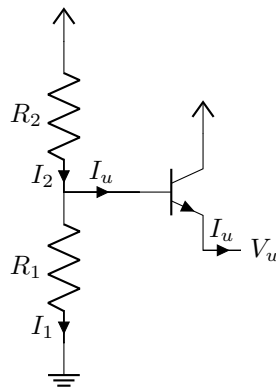
Per trasformare la funzione in uscita dell'operatore operazionale in una funzione a gradino, occorre amplificare l'uscita corrente e traslarla sul nuovo intervallo. Infatti mettendo l'uscita in ingresso a due invertitori il valore, appena positivo, viene tradotto in un valore positivo alto.

Chiamando le corrente di ingresso sui rami I^+ e I^- , è facile capire che sono nulle, siccome sono correnti di gate dei transistori mos. Inoltre la tensione in uscita non dovrebbe essere influenzata dai circuiti connessi a carico, in modo da far funzionare questo componente come un generatore ideale di tensione, controllato dalla differenza di tensione in ingresso.

Analizzando il seguente circuito:



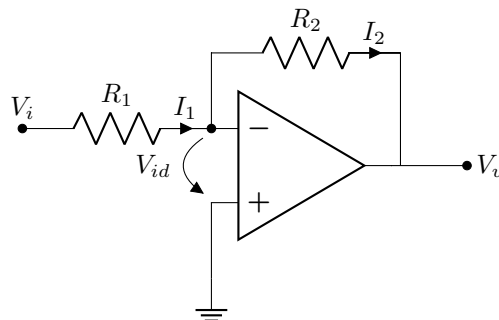
Otteniamo la formula della tensione in uscita $V_u = \frac{R_1}{R_1 + R_2}(V_{dd} - R_2 I_u)$, da cui possiamo osservare che se la corrente in uscita I_u è 0, il circuito si comporta esattamente come un partitore di corrente. Per ottenere questo risultato, colleghiamo un transistoro bipolare



Il quale deve necessariamente funzionare in regione attiva diretta, in quanto la corrente di emettitore $I_u > 0$ e non può essere saturo, in quanto la giunzione della base collettore è negativa. Quindi la corrente di collettore $I_c = \beta_F I_B$ e $I_E = (\beta_F + 1) I_B$, siccome β_F è grande, allora la corrente I_B è approssimativamente 100 volte minore della corrente I_E . Generando una tensione in uscita pressoché indipendente dalla corrente a carico.

Le tre caratteristiche ideali di questo circuito sono che dipende unicamente dalla differenza delle tensioni in ingresso, e la tensione e corrente a carico no influiscono sull'uscita.

Amplificatore invertente



Il ramo con la corrente R_2 prende il nome di ramo in retroazione, siccome collega l'uscita all'ingresso. Studiamo il circuito nelle tre regioni di funzionamento, ricordando le relazioni:

Alto guadagno:

Saturazione positiva

Saturazione negativa

$$\begin{array}{lll}
 \left\{ \begin{array}{l} V_{id} = 0 \\ -V_M < V_u < V_M \end{array} \right. & \left\{ \begin{array}{l} V_{id} > 0 \\ V_u = V_M \end{array} \right. & \left\{ \begin{array}{l} V_{id} < 0 \\ V_u = -V_M \end{array} \right. & \text{Con } I^+ = \\
 I^- = 0 \text{ e } V_{id} = V^+ - V^- \text{ sempre valide.} & & &
 \end{array}$$

In regione di alto guadagno, $V^+ = V^- = 0$, in quanto la tensione V^- è vincolata dal potenziale di terra, portando un riferimento di terra virtuale, in quanto il nodo non è connesso veramente a terra.

Applicando kirkoff al nodo, allora $I_1 = I_2$ implicando

$$V_u = -\frac{R_2}{R_1} V_i$$

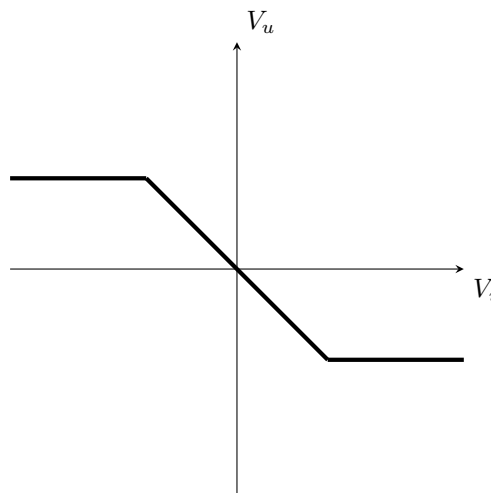
Valida per $V_i > -\frac{R_1}{R_2} V_M$ e $V_u < \frac{R_1}{R_2} V_M$

Come osservabile dalla relazione della tensione in uscita, in questa regione di alto guadagno, è possibile variare il valore del guadagno semplicemente cambiando le proporzioni tra le due resistenze.

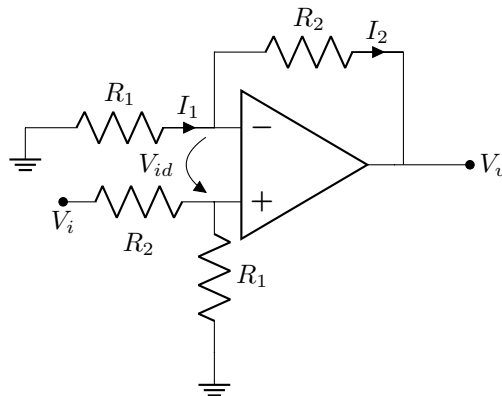
In regime di saturazione positiva, allora $V^+ > V_-$ e la corrente in ingresso è nulla, implicando $I_1 = I_2$ e dalle equazioni di ohm:

$$V^- = \frac{V_i R_2 + V_u R_1}{R_1 + R_2}$$

È da controllare $V^- < 0$, rendendo vera l'equazione per $V_i < -\frac{R_1 V_M}{R_2}$, lo stesso accade in regione di saturazione negativa.



Amplificatore lineare non invertente



In regione di alto guadagno,

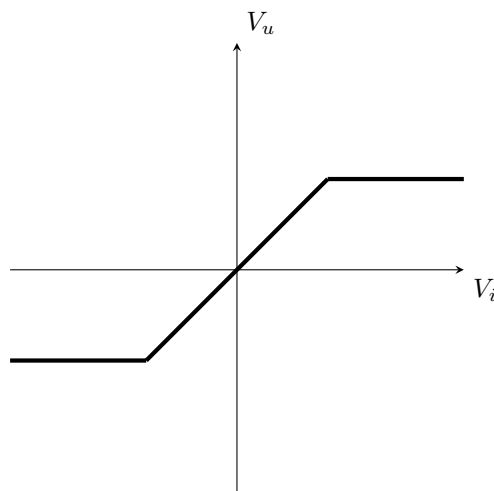
$$V^+ = V_i \frac{R_2}{R_1 + R_2}$$

$$V^- = \frac{V^- - V_u}{R_2}$$

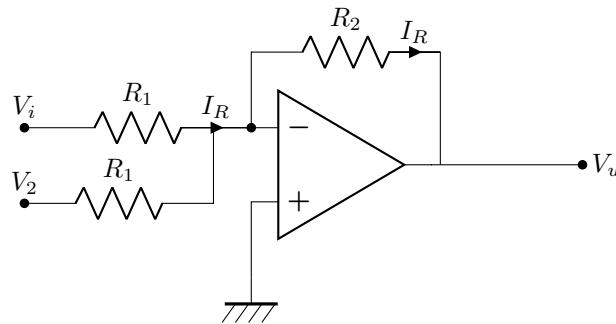
Da cui la relazione: $V_u = \frac{R_2}{R_1} V_i$, che è la stessa relazione di prima, ma con segno opposto. Calcolando le condizioni di validità:

$$\frac{R_2}{R_1} V_M < V_i < \frac{R_2}{R_1} V_M$$

I calcoli sono gli stessi del precedente, risultando in un amplificatore lineare non invertente.



Circuito sommatore analogico



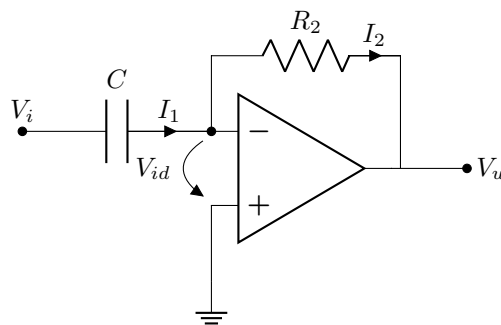
In regione di alto guadagno, unendo le equazioni delle singole correnti con l'equazione di kirchoff al nodo:

$$I_1 + I'_1 = I_2$$
$$\frac{V_u + V_{i2}}{R_1} = -\frac{V_u}{R_2}$$

Da cui $V_u = -\frac{R_2}{R_1}(V_{i1} + V_{i2})$, risultato estremamente importante perché indica la funzione di una somma tra le due tensioni in ingresso. Per questo motivo, il circuito prende il nome di sommatore analogico.

Questo risultato ovviamente è indipendente dal numero di ingressi, e variando il valore delle resistenze legate agli ingressi è possibile fare una somma pesata dei segnali in ingresso.

Circuito derivatore

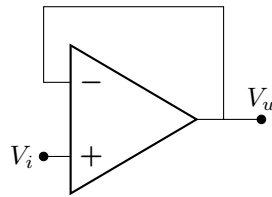


In regione attiva diretta $I_R = -\frac{V_u}{R}$ e $I_C = C \frac{dV_i}{dt}$, da cui:

$$V_u = -RC \frac{dV_i}{dt}$$

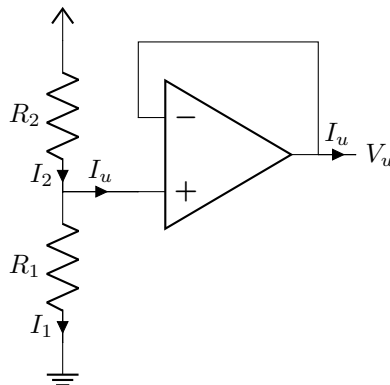
Il circuito è quindi in grado di calcolare la derivata del segnale in ingresso. Mettendo un condensatore sul ramo di uscita ed una resistenza nel ramo in ingresso si ottiene un circuito integratore.

Stadio separatore



In regione attiva diretta, è facile calcolare $V_u = V_i$ per $-V_M < V_i < V_M$.

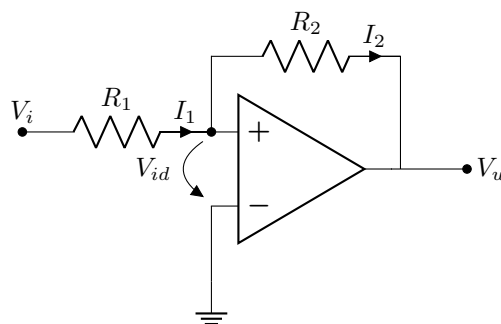
In questo modo è possibile leggere il valore di V_i a corrente d'ingresso nulla, erogando una corrente arbitraria. Questo principio è possibile utilizzarlo per creare un generatore di tensione



L'applicazione più importante è in strumenti di misura, essenziale per assorbire una corrente nulla dal circuito in oggetto.

Test

Ipotizzando di invertire la polarità dell'amplificatore operazionale, analizziamo il seguente circuito:



In regione di alto guadagno, $V^+ = V^- = 0$ e $I_1 = I_2$, $I_1 = V_i/R$ e $I_2 = -V_u/R_2$ che ci porta a dire:

$$V_u = -\frac{R_2}{R_1}V_i$$

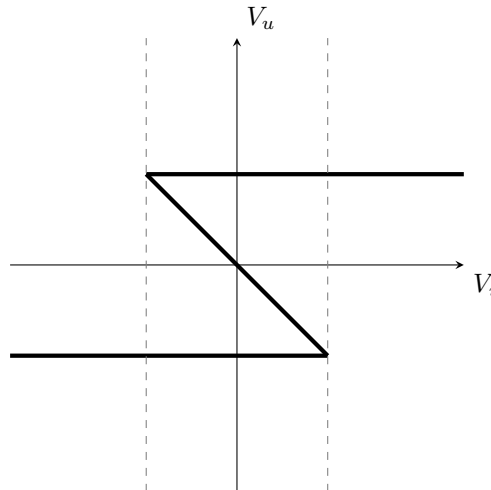
e le condizioni di esistenza di rimangono le stesse:

$$-\frac{R_2}{R_1}V_M < V_i < \frac{R_2}{R_1}V_M$$

In regione di saturazione positiva $V_u = V_M$ e $V_{id} > 0$, quindi $V^+ > 0$ Ma calcolando le condizioni di esistenza otteniamo:

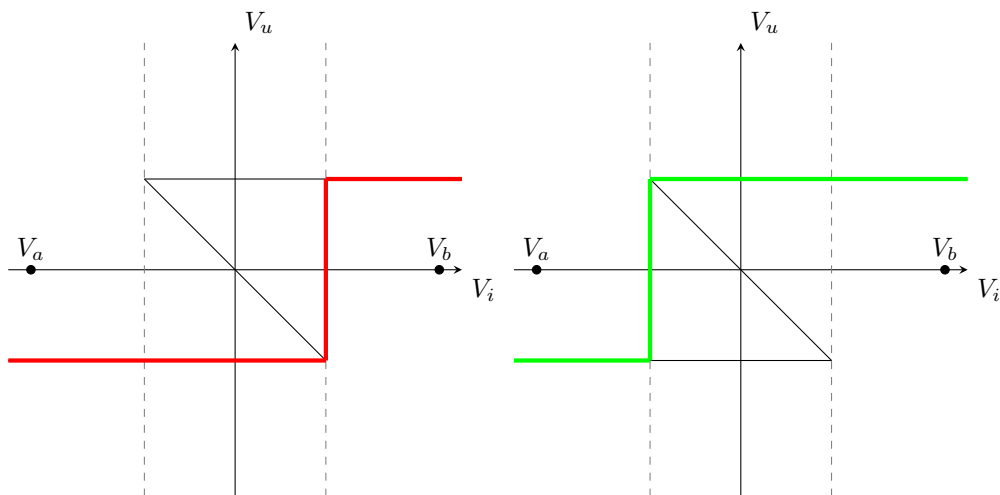
$$V_i > -\frac{R_1V_M}{R_2}$$

Mentre la regione di alto guadagno non cambia aspetto, le regioni di saturazione cambiano funzionamento.



La relazione tra ingresso ed uscita non è più funzionale, richiede un'ulteriore ragionamento per comprendere quale tra i tre valori vengono restituiti dal circuito nella regione compresa tra $-V^*$ e V^* .

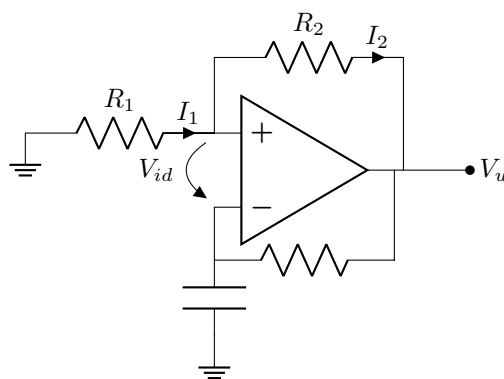
Ragioniamo quindi in regime dinamico, partendo da $V_a < -V^*$ spostandoci a $V_b > V^*$. Fintanto che $V_a < -V^*$, il valore è costante raggiunto il valore critico, per continuità l'uscita si tiene costante vino al punto V^* . Appena raggiunto il punto V^* è ritorna ad essere presente una sola soluzione, portando l'uscita al valore positivo.



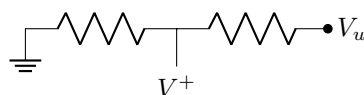
Seguendo analogo ragionamento in verso opposto, si nota che per andare da b ad a e da a a b , si effettuando due percorsi diversi, indicati in rosso ed in verde. Questo andamento prende il nome di ciclo di isteresi, ed il circuito prende il nome di circuito "trigger" di Shmitt.

Per forzare l'uscita alta serve forzare un valore sufficientemente positivo. La caratteristica di questo circuito è che quando si triggera ed il valore si porta positivo, è difficile che si spenga a causa di rumore. Questo è utile ad esempio in un circuito dove una lampada si spegne quando raggiunge un certo livello di luce.

Questo è il primo circuito che presenta un elemento di memoria.



Immaginiamo che inizialmente la capacità sia scarica, e che per un qualunque motivo, ci troviamo nella condizione di alto guadagno, con $V_u = V_M$. Ricordando che la corrente in ingresso è comunque nulla, possiamo separare il ramo e vederlo come uno risolvibile attraverso la formula del partitore:



Quindi fintanto che $V_u = V_M$ allora

$$V^+ = \frac{R_q}{R_1 + R_2} V_M > 0$$

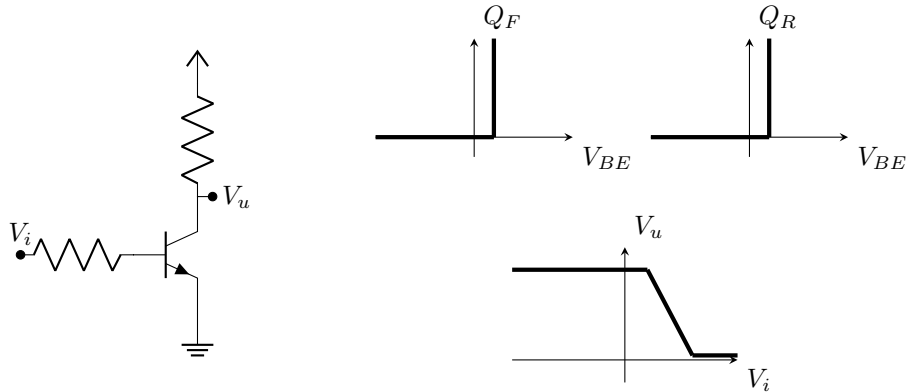
Quindi $V_{id} > 0$, in accordo con le ipotesi. Analizzando l'altro ramo con il condensatore, siccome la tensione V_u è positiva, e sul ramo circola una corrente, questa ha l'effetto di caricare il condensatore C , aumentando la tensione V^- . Fintanto che $V_{id} > 0$, ovvero $V^+ > V^-$ allora l'amplificatore lavora in alto guadagno. Nel momento in cui $V_{id} < 0$, allora $V^- > V^+$ e l'uscita si porta con un ritardo al valore $-V_M$.

Il condensatore vedendo variato il valore di V_u da V_M a $-V_M$ inizia a scaricarsi, fino a quando $V^- > V_u$. Raggiunto quel valore l'uscita si riporta a $+V_M$ e si ripete il ciclo.

Analizzando la stabilità di questo circuito infatti, il valore di $V_i = 0$, otteniamo che il circuito è instabile. In questo modo si ottiene un generatore di un segnale periodico, (onda quadra) dipendente dal valore delle resistenze e dalla capacità del condensatore.

Analogamente, mettendo a cascata un numero dispari di invertitori, e chiudendoli ad anello, si ottiene un oscillatore ad anello.

RTL in regione dinamica



Ricordandoci i modelli di approssimazione utilizzati, descritti dalle equazioni:

$$Q_F := \begin{cases} Q_+ = 0 & \text{per } V_{BE} < V_\gamma \\ V_{BE} = V_\gamma & \text{per } Q_F > 0 \end{cases}$$

$$Q_R := \begin{cases} V_{BC} = V'_\gamma & \text{per } Q_R > 0 \\ Q_R = 0 & \text{per } V_{BC} < V'_\gamma \end{cases}$$

Ricordiamo inoltre che $V_\gamma - V'_\gamma = V_{CE_{SAT}}$ quando entrambe le giunzioni sono polarizzate in diretta. E l'uscita si porta al valore alto quando il transistor è spento, ed il valore basso quando è in saturazione.

Tenendo presente questi dati calcoliamo il comportamento del circuito in regime dinamico, presupponendo un'ingresso a gradino, dove

$$V_i(t) = \begin{cases} 0 & \text{per } t < 0 \\ V_{cc} & \text{per } t > 0 \end{cases}$$

Per $t < 0$, siccome siamo in regioni statiche conosciamo già il valore dell'uscita: $V_u = V_{cc}$.

Quindi, sapendo che il transistor è spento, $I_B = 0$, e $V_i = V_{BE} = 0$. Guardando in corrispondenza del grafico il valore della carica: $Q_F = 0$

Siccome $V_{BC} = V_{BE} - V_{cc} = -V_{cc}$, allora $Q_R = 0$

Per $t \rightarrow \infty$ il transistorore si trova in regione di saturazione, quindi $V_u = V_{CE_{SAT}}$, ed essendo $V_{BE} = V_\gamma$, allora $Q_F > 0$ e $V_{BC} = V'_\gamma$ quindi $Q_R > 0$.

$$I_B = \frac{V_{cc} - V_\gamma}{R_B} = \frac{Q_F}{\beta_F \tau_F} + \frac{Q_R}{\beta_B \tau_R}$$

$$I_C = \frac{V_{cc} - V_{CE_{SAT}}}{R_C} = \frac{Q_F}{\tau_F} - \frac{Q_R}{\alpha_R \tau_R}$$

Da cui, indicando con $\frac{1}{M} \frac{1}{\tau_R} \left(\frac{1}{\alpha_R} + \frac{\beta_F}{\beta_R} \right)$ si può scrivere:

$$Q_R = M \left(\frac{\beta_F (V_{cc} - V_\gamma)}{R_B} - \frac{(V_{cc} - V_{CE_{SAT}})}{R_C} \right)$$

$$\frac{Q_F}{\tau_F} = \frac{V_{cc} - V_{CE_{SAT}}}{R_C} + \frac{Q_R}{\alpha_R \tau_R}$$

Siccome la carica Q_R deve esser positiva, allora controllando la condizione, si ottiene:

$$V_{CE_{SAT}} < V_{cc} - \frac{\beta_F R_C}{R_B} (V_{cc} - V_\gamma)$$

dove la condizione a destra della disuguaglianza è il tratto obliquo della retta del grafico di $V_u(V_i)$, tanto più la condizione è verificata, tanto più è maggiore la carica Q_R associata al transistorore e maggiore è la durata del transitorio. Diminuire la carica va quindi in conflitto con l'obiettivo di diminuire il margine ai ritardi del transitorio.

Ovviamente, nel transitorio della carica da $Q = 0$ a Q verticale, (vedi grafico carica) sarà necessario passare in un punto intermedio (C).

$$V_{BC} = V_{BE} - V_{CE}$$

$$V_{BC} = V_\gamma - V_{cc}$$

Descriviamo quindi il transitorio da A a B , separandolo in due intervalli: Un primo tratto, da A a C , dove sono vere le equazioni:

$$Q_F = 0$$

$$Q_R = 0$$

$$V_{BE} < V_\gamma$$

$$V_{BC} < V'_\gamma$$

Un secondo tratto da C a D :

$$Q_F > 0$$

$$Q_R = 0$$

$$V_{BE} = V_\gamma$$

$$V_{BC} < V'_\gamma$$

Ed un terzo tratto da D a B :

$$\left. \begin{aligned} Q_F &> 0 \\ Q_R &> 0 \\ V_{BE} &= V_\gamma \\ V_{BC} &> V'_\gamma \end{aligned} \right\} \Rightarrow V_{CE} = V_{CE_{SAT}}$$

Analisi dei tratti della caratteristica

Nel primo tratto, non essendoci spostamento di carica il transitorio è istantaneo. In un tempo potenzialmente nullo, la tensione V_{BE} , raggiunge il valore di V_γ .

Nel secondo tratto, necessariamente la corrente $I_B = \frac{Q_F}{\beta_F \tau_F} + \frac{Q_R}{\beta_R \tau_R} + \frac{dQ_F}{dt} + \frac{dQ_R}{dt}$, siccome la carica Q_R è nulla, allora viene espressa come:

$$I_B = \frac{Q_F}{\beta_F \tau_F} + \frac{dQ_F}{dt}$$

Inoltre, dall'equazione di kirchoff, sappiamo che $I_B = \frac{V_u - V_\gamma}{R_B}$. Dall'equazione della carica I_B precedente possiamo scrivere:

$$\frac{dQ_F}{dt} = I_B - \frac{Q_F}{\beta_F \tau_F} = -\frac{1}{\beta_F \tau_F} (Q_F - M)$$

Ed integrando si ottiene:

$$\ln \frac{Q_F(t) - M}{-M} = -\frac{t}{\beta_F \tau_F}$$

Da cui:

$$Q_F(t) = M(1 - e^{-\frac{t}{\beta_F \tau_F}}) = \frac{\beta_F \tau_F (V_{cc} - V_\gamma)}{R_B} (1 - e^{-\frac{t}{\beta_F \tau_F}})$$

Tutto questo è vero fino a quando non viene raggiunto il punto D , in cui $V_{BE} = V_\gamma$ e $V_{BC} = V'_\gamma$, quindi $V_{CE} = V_u = V_{CE_{SAT}}$. Scrivendo l'equazione di V_{cc} , otteniamo:

$$V_u = V_{cc} - R_C I_C = V_{cc} - \frac{\beta_F R_C}{R_B} (V_{cc} - V_\gamma) (1 - e^{-\frac{t}{\beta_F \tau_F}})$$

Analizzando il terzo tratto del transitorio, dove $V_{BE} = V_\gamma$, $V_{BC} = V'_\gamma$ e $Q_F, Q_R > 0$, facendo sempre riferimento all'equazione della corrente I_C ed I_B , dove questa volta tutte le componenti sono significative.

Inoltre siccome $V_{BE} = V_\gamma$ e $V_{BC} = V'_\gamma$, l'uscita $V_u = V_{CE} = V_\gamma - V'_\gamma = V_{CE_{SAT}}$. Unendo le equazioni delle due correnti quindi otteniamo il sistema:

$$\begin{cases} I_C = \frac{Q_F}{\tau_F} - \frac{Q_R}{\alpha_R \tau_R} - \frac{dQ_R}{dt} = \frac{V_{cc} - V_{CE_{SAT}}}{R_C} \\ I_B = \frac{Q_F}{\beta_F \tau_F} + \frac{Q_R}{\beta_R \tau_R} + \frac{dQ_F}{dt} + \frac{dQ_R}{dt} = \frac{V_{cc} - V_\gamma}{R_B} \end{cases}$$

Siccome sappiamo già che l'uscita V_u è costante ed indipendente dal tempo, la carica continua a crescere e sappiamo già a quali valori tende asintoticamente. Quindi dopo un tempo sufficientemente lungo abbiamo già calcolato il valore della caratteristica.

Studio transitorio "gradino in discesa"

Studiamo il transitorio percorrendo i punti da B ad A . Nel tratto iniziale, caratterizzato da $Q_F, Q_R > 0$ abbiamo $V_u = V_{CE_{SAT}}$. Quando il valore di Q_R raggiunge 0, allora il punto coincide con il punto D della caratteristica. Esattamente come il diodo, esso prima di spegnersi mantiene ai suoi capi una tensione costante. Calcolando nell'intervallo la corrente di base:

$$I_B = \frac{V_i = V_{BE}}{R_B} = -\frac{V_\gamma}{R_B}$$

Una corrente negativa, che ricorda il tratto di spegnimento del diodo, in quanto stiamo passando dalla regione di saturazione, dove entrambe le giunzioni sono polarizzate in diretta, alla regione normale, dove la regione è polarizzata in inversa. Nell'intervallo anche la carica Q_R , varia secondo una relazione non ancora definita nel dettaglio, senza raggiungere 0.

Analizzando il secondo tratto del transitorio, da D a C , caratterizzato da $V_{BE} = V_\gamma$, $V_{BC} < V'_\gamma$ di conseguenza $V_u > V_{CE_{SAT}}$. $Q_F > 0$ e $Q_R = 0$. Possiamo quindi scrivere un'equazione simile a quella espressa in precedenza ricordando la formula della corrente:

$$I_B = \frac{Q_F}{\beta_F I_F} + \frac{Q_R}{\beta_R \tau_R} + \frac{dQ_F}{dt} + \frac{dQ_R}{dt} = \frac{Q_F}{\beta_F I_F} + \frac{dQ_F}{dt} = -\frac{V_\gamma}{R_B}$$
$$\frac{dQ_F}{dt} =$$