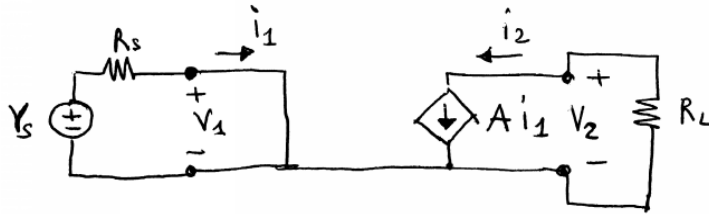


4. Il transistoro BJT

4.1 Generatore di corrente controllato in corrente

Una funzione essenziale nella maggior parte dei circuiti elettronici è rappresentata dal generatore controllato. Un generatore controllato è un dispositivo a due porte, quindi con quattro terminali, che si riducono tuttavia nella maggior parte dei casi a tre, perché due sono in comune tra ingresso e uscita. Prendiamo inizialmente in considerazione il generatore di corrente controllato in corrente, perché questo può essere approssimato tramite il transistoro bipolare a giunzione (Bipolar Junction Transistor, BJT). In forma ideale tale generatore può essere rappresentato con lo schema che segue.



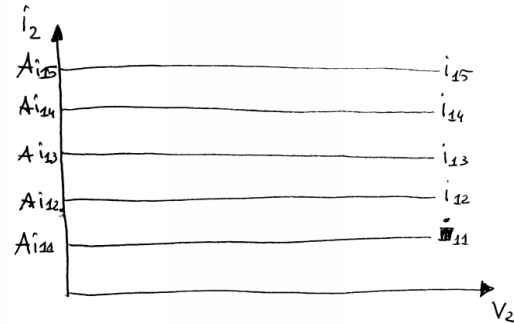
La quantità A che rappresenta il rapporto tra la corrente di uscita e quella di ingresso viene definito “guadagno di corrente”. Notiamo che l’impedenza di ingresso è stata scelta volutamente nulla in modo che la corrente di ingresso sia massima. Notiamo anche che il dispositivo è unidirezionale, cioè che le grandezze elettriche eventualmente applicate all’uscita non influenzano in alcun modo quelle in ingresso. Possiamo facilmente calcolare il rapporto tra la tensione v_s e quella v_2 di uscita:

$$i_1 = \frac{v_s}{R_S} \quad v_2 = -A i_1 R_L = -\frac{A R_L}{R_S} v_s.$$

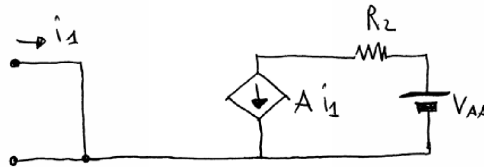
Se $A R_L / R_S$ è maggiore dell’unità $|v_2| > |v_s|$ e si ha quindi un guadagno di tensione. Se inoltre $A > 1$, come di solito accade in pratica, si ha anche un’amplificazione di corrente. Se il prodotto del guadagno di tensione per quello di corrente, $A^2 R_L / R_S$ è maggiore dell’unità si ha un guadagno di potenza: significa quindi che è possibile controllare con un piccolo segnale un segnale di potenza maggiore. Proprio questo effetto rappresenta la caratteristica peculiare dei cosiddetti “componenti attivi”, dei quali il BJT sarà il primo esempio che esamineremo. È importante sottolineare fino da ora che la potenza in più che si presenta all’uscita rispetto a quella in ingresso proviene dalla sorgente di alimentazione, che è sempre necessaria per il funzionamento dei dispositivi attivi. L’amplificazione di potenza è di fondamentale importanza non solo nel trattamento dei segnali analogici, ma anche in campo digitale, perché, insieme con la proprietà di unidirezionalità, consente la realizzazione di interruttori comandati che controllano correnti più grandi di quella di comando e quindi di reti che non danno luogo a degrado dei livelli logici.

Una descrizione molto utile del comportamento di un generatore comandato si ottiene tramite la rappresentazione grafica delle caratteristiche di ingresso e di uscita, sotto la forma di famiglie di curve. Le caratteristiche di ingresso vengono tracciate sul piano v_1 - i_1 e rappresentano il legame tra tali grandezze in funzione di una delle grandezze di uscita. Nel caso particolarmente semplice considerato non ha significato tracciare le caratteristiche di ingresso, dato che v_1 è sempre nulla. Per quanto

riguarda invece le caratteristiche di uscita, queste vengono rappresentate sul piano v_2-i_2 , in funzione della i_1 (per un numero discreto di valori della i_1). Per il generatore di corrente ideale controllato in corrente, le caratteristiche di uscita sono molto semplici: in corrispondenza di ogni valore della i_1 si ha una retta parallela all'asse delle ascisse e da esso distante Ai_1 , visto che la corrente di uscita è proporzionale a i_1 e indipendente dal valore di v_2 .



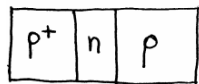
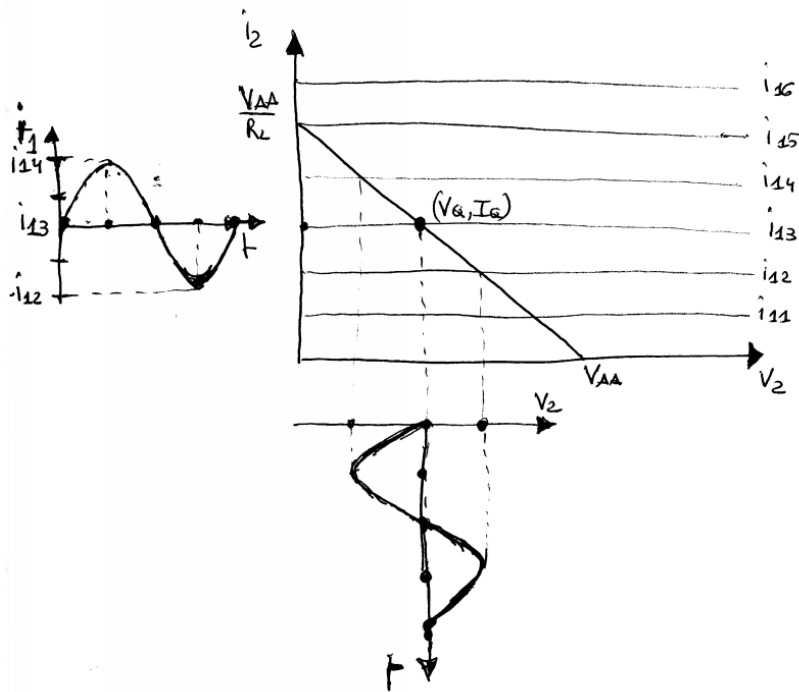
Se prendiamo in considerazione un circuito come quello di seguito rappresentato, in cui sull'uscita è collegato un generatore di tensione V_{AA} tramite un resistore R_2 , possiamo indicare la relazione tra v_2 e i_2 imposta dalla rete esterna per mezzo di una retta di carico, che può essere riportata sul piano delle caratteristiche. La presenza del generatore V_{AA} può sembrare a questo stadio artificiosa, dato che comporta semplicemente l'aggiunta di un valor medio alla tensione di uscita. Vedremo più avanti come, nel caso di un dispositivo reale, questa sia invece indispensabile per fornire l'energia necessaria al suo funzionamento.



Cerchiamo ora di determinare l'andamento della tensione di uscita in funzione della corrente i_1 . Al variare di i_1 il punto di lavoro si sposta lungo la retta di carico: rappresentiamo la corrente i_1 in funzione del tempo a lato della caratteristica di uscita e, per ogni valore di i_1 , andiamo a individuare la corrispondente caratteristica I_2-V_2 e la relativa intersezione con la retta di carico. In questo modo otteniamo il valore all'istante corrispondente della v_2 e lo possiamo riportare in un grafico posto sotto alle caratteristiche di uscita. Notiamo che la forma d'onda d'uscita è una replica amplificata, invertita di fase e traslata di quella di ingresso.

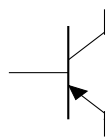
4.2 Il principio di funzionamento del BJT

Il transistor bipolare a giunzione consiste di due giunzioni pn poste una di seguito all'altra e orientate in senso inverso: si tratta quindi di tre regioni consecutive, una p , una n e una p nel caso di un dispositivo pnp e una n , una p e una n nel caso dei dispositivi npn . La caratteristica fondamentale che distingue il transistor da due giunzioni pn in serie "back-to-back" è rappresentata dallo spessore estremamente ridotto della zona centrale (non rappresentato correttamente in scala nel disegno

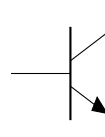


sottostante), che, come vedremo, dà luogo all'interazione tra le due giunzioni, alla base dell'“effetto transistor”.

L'elettrodo centrale viene definito base, mentre gli altri due sono denominati emettitore e collettore. Il drogaggio delle due regioni esterne non è uguale: in un transistor la zona di emettitore è significamente più drogata di quelle di base e di collettore; si indica infatti con p^+ nei transistori pnp e con n^+ nei transistori npn . Nella trattazione del modello fisico del transistor tutte le correnti vengono di solito scelte con verso entrante. I transistori vengono indicati negli schemi circuitali con i simboli sottoindicati, che differiscono, tra pnp e npn , solo per il verso della freccia dell'emettitore.

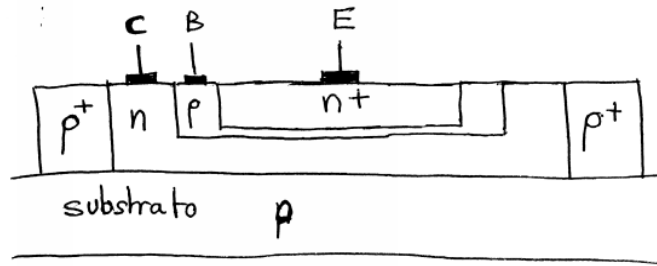


pnp



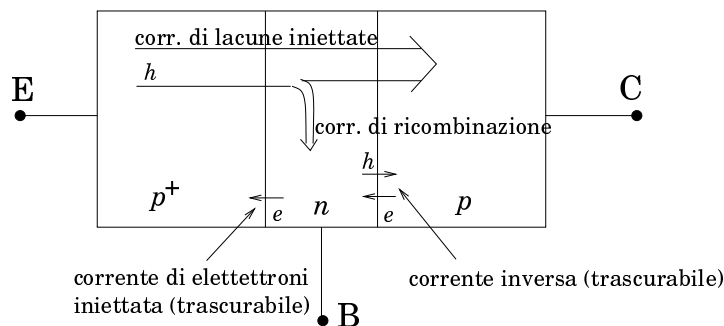
npn

Dal punto di vista costruttivo i BJT non vengono realizzati connettendo insieme pezzi separati di silicio con diverso drogaggio, anche perché in tal modo sarebbe ben difficile riuscire a ottenere gli spessori di base estremamente ridotti che sono richiesti per il corretto funzionamento. Un processo tecnologico standard per la fabbricazione dei transistori è quello planare, illustrato di seguito e basato sulla successiva realizzazione di strati con diverso drogaggio. L'esempio preso in considerazione è quello di un transistor npn .



Si parte da un substrato di silicio monocristallino di tipo p sul quale viene ottenuto uno strato n che rappresenta il collettore. Si realizza poi uno strato estremamente sottile di tipo p , che costituisce la base e infine viene ottenuta una regione n^+ che rappresenta l'emettitore. Lateralmente vengono realizzate delle regioni p^+ che hanno funzione di isolamento dai transistori adiacenti. La definizione laterale delle varie regioni viene ottenuta tramite tecniche litografiche. Con opportune procedure si ottengono anche dei contatti elettrici che raggiungono le tre regioni del transistor e rappresentano la connessione con il resto del circuito.

Passiamo ora a intraprendere uno studio più dettagliato del funzionamento del BJT. Nella nostra trattazione supporremo che, eccetto che nelle zone di svuotamento, la concentrazione di portatori nelle tre regioni sia sufficiente per poter considerare nulla la caduta di tensione attraverso ciascuna delle regioni stesse: le correnti che attraversano il transistor sono dunque tutte correnti di diffusione. Consideriamo la modalità di funzionamento più tipica, cioè quella nella quale la giunzione base-emettitore viene polarizzata direttamente, mentre quella base-collettore risulta polarizzata inversamente. In conseguenza della polarizzazione diretta tra base ed emettitore, una corrente di lacune viene iniettata dall'emettitore nella base, mentre una corrente di elettroni passa dalla base all'emettitore. Dato che, come già sottolineato, l'emettitore è molto più drogato della base, la corrente di lacune iniettata in base sarà molto più grande di quella di elettroni iniettata dalla base nell'emettitore, che può essere trascurata. Se la regione di base fosse lunga, la corrente di lacune iniettata dall'emettitore, darebbe luogo a una concentrazione in eccesso di lacune, che decaderebbe esponenzialmente con la distanza dalla zona di svuotamento. Essendo però la base corta rispetto alla lunghezza di diffusione per le lacune, solo poche lacune riescono a ricombinarsi, mentre la maggior parte raggiunge la zona di svuotamento tra base e collettore, dove le lacune vengono trascinate verso il collettore dal campo elettrico favorevole. Una grossa parte della corrente di lacune iniettata nella base dall'emettitore raggiunge quindi il collettore, mentre solo una piccola frazione dà luogo a ricombinazione in base.



La corrente di base è quindi sostanzialmente costituita dal flusso di elettroni che

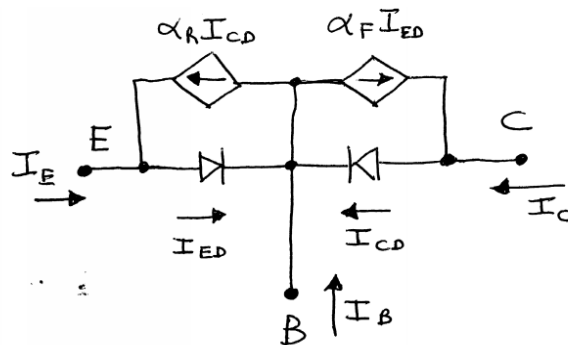
danno luogo alla ricombinazione delle lacune. Essendo la frazione di lacune iniettate dall'emettitore che si ricombina in base molto piccola, la corrente di base è molto più piccola di quelle di emettitore e di collettore. Una variazione percentuale della corrente di base si ripercuote dunque in modo proporzionale sulla corrente di collettore, dando luogo a un comportamento corrispondente a quello di un generatore di corrente controllato in corrente.

Le correnti all'interno di un transistor *pn*p possono dunque essere elencate nel modo seguente:

- corrente di lacune iniettata dall'emettitore nella base
- corrente di elettroni iniettata dalla base nell'emettitore (trascurabile)
- corrente di elettroni in base, corrispondente alla ricombinazione di una frazione delle lacune.
- corrente inversa attraverso la giunzione base-collettore (trascurabile)
- corrente di lacune inviata nel collettore, corrispondente alla frazione della corrente proveniente dall'emettitore che non si ricombina in base.

4.3 Le equazioni di Ebers-Moll

È possibile definire un modello matematico che descrive il comportamento per grandi segnali del transistor BJT, includendo tutte le componenti di corrente sopra menzionate. Tale modello, definito modello di Ebers-Moll dai nomi di coloro che lo proposero, è necessariamente non lineare e rappresenta il transistor come costituito da due giunzioni *pn* contrapposte con l'aggiunta di generatori di corrente comandati, che descrivono la porzione della corrente di emettitore trasferita al collettore e viceversa.



Dallo schema sopra rappresentato, valido per un transistor *pn*p, otteniamo le cosiddette equazioni di Ebers-Moll:

$$\begin{cases} I_E = I_{ED} - \alpha_R I_{CD} = I_{ES} \left(e^{\frac{V_{EB}}{V_T}} - 1 \right) - \alpha_R I_{CS} \left(e^{\frac{V_{CB}}{V_T}} - 1 \right) \\ I_C = I_{CD} - \alpha_F I_{ED} = -\alpha_F I_{ES} \left(e^{\frac{V_{EB}}{V_T}} - 1 \right) + I_{CS} \left(e^{\frac{V_{CB}}{V_T}} - 1 \right) \end{cases},$$

dove α_F rappresenta la frazione della corrente di emettitore che raggiunge il collettore, mentre α_R rappresenta la frazione della corrente di collettore (in caso di polarizzazione diretta della giunzione base-collettore) che viene trasferita all'emettitore. Sia α_R sia α_F risultano minori dell'unità (rappresentando una frazione di una quantità

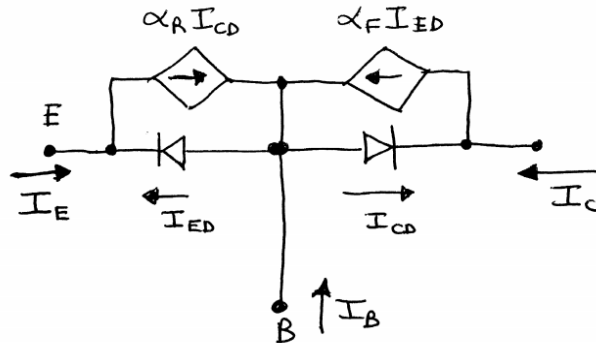
totale) e sono legati tra loro dalle cosiddette condizioni di reciprocità per il transistoro:

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

Per i transistori normalmente realizzati, nei quali la regione di emettitore è molto più drogata delle altre due regioni, si ha $0.98 \leq \alpha_F \leq 0.998$ e $0.4 \leq \alpha_R \leq 0.8$. Le correnti I_{ES} e I_{CS} sono dell'ordine di 10^{-15} A nei transistori al silicio. La corrente I_B può essere immediatamente ricavata dalle equazioni di Ebers-Moll applicando la legge di Kirchhoff ai nodi:

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

Nel caso di un transistoro *npn* il circuito equivalente di Ebers-Moll deve essere modificato, in modo da tenere in considerazione il diverso segno delle tensioni e delle correnti;



Le equazioni di Ebers-Moll per un transistoro *npn* diventano:

$$\begin{cases} I_E = -I_{ES} \left(e^{\frac{-V_{EB}}{V_T}} - 1 \right) + \alpha_R I_{CS} \left(e^{\frac{-V_{CB}}{V_T}} - 1 \right) \\ I_C = \alpha_F I_{ES} \left(e^{\frac{-V_{EB}}{V_T}} - 1 \right) - I_{CS} \left(e^{\frac{-V_{CB}}{V_T}} - 1 \right) \end{cases}.$$

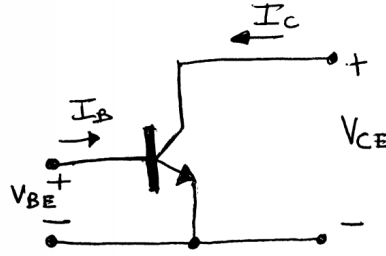
Sulla base del modello di Ebers-Moll possiamo definire le quattro zone di funzionamento possibili per il transistoro:

- Zona attiva diretta, con la giunzione base-emettitore polarizzata direttamente e quella collettore-base polarizzata inversamente. Questa è la condizione di funzionamento più utilizzata e quella in cui si hanno buone prestazioni dal punto di vista dell'amplificazione.
- Zona attiva inversa, con la giunzione base-emettitore polarizzata inversamente e quella collettore-base polarizzata direttamente. Dato che il transistoro non è simmetrico, le prestazioni sono in questo caso molto degradate rispetto a quelle in zona attiva diretta e pertanto i transistori non vengono quasi mai utilizzati in zona attiva inversa.
- Zona di interdizione, con ambedue le giunzioni polarizzate inversamente. In questo caso le correnti che attraversano il transistoro sono estremamente piccole e quindi lo stesso può considerarsi come un circuito aperto. Questo modo di funzionamento può essere utilizzato in applicazioni di tipo digitale per emulare un interruttore aperto.

- d) Zona di saturazione, con ambedue le giunzioni polarizzate direttamente. In questo caso la caduta di tensione tra collettore ed emettitore è molto piccola (pochi decimi di volt) e il comportamento del transistor assomiglia a quello di un corto circuito. Questo modo di funzionamento può essere impiegato in circuiti digitali per emulare un interruttore chiuso.

4.4 Caratteristiche a emettitore comune

Ci occuperemo ora della determinazione delle caratteristiche $I - V$ del transistor $n-p-n$ quando questo sia connesso nella configurazione a emettitore comune, vale a dire con l'emettitore a comune tra ingresso e uscita, la base sulla porta di ingresso e il collettore su quella di uscita. In tale configurazione le grandezze di ingresso sono la V_{BE} e la I_B , mentre quelle di uscita sono la V_{CE} e la I_C .



Le caratteristiche di uscita a emettitore comune descriveranno quindi la relazione tra I_C e V_{CE} , in funzione di I_B e saranno costituite da una famiglia di curve sul piano $V_{CE} - I_C$, ciascuna per un diverso valore della corrente I_B .

In zona di funzionamento attiva, se trascuriamo la corrente inversa che attraversa la giunzione base-collettore, possiamo affermare che la corrente di collettore è sostanzialmente pari alla porzione di quella di emettitore che ha superato la base: $I_C = -\alpha_F I_E$. L'espressione della corrente di base risulta quindi, dall'applicazione del principio di Kirchhoff ai nodi,

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

Sostituendo nella precedente equazione a I_E la sua espressione in funzione di I_C otteniamo

$$I_B = (1 - \alpha_F) \frac{I_C}{\alpha_F},$$

quindi

$$I_C = \frac{\alpha_F}{1 - \alpha_F} I_B = \beta_F I_B,$$

dove β_F è definito come

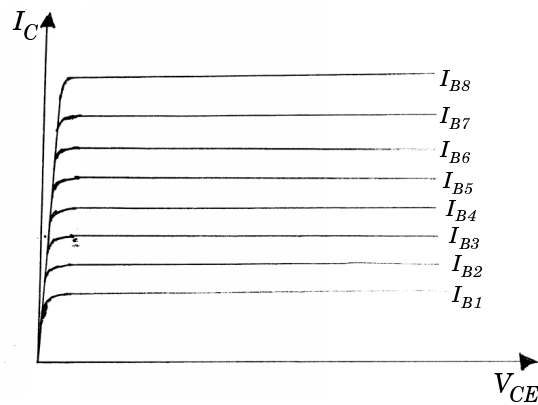
$$\beta_F = \frac{\alpha_F}{1 - \alpha_F}$$

ed è di solito indicato dai costruttori come h_{FE} (si noti che il pedice FE è maiuscolo e quindi questo parametro non deve essere confuso con h_{fe} che sarà introdotto più avanti e avrà tutt'altro significato).

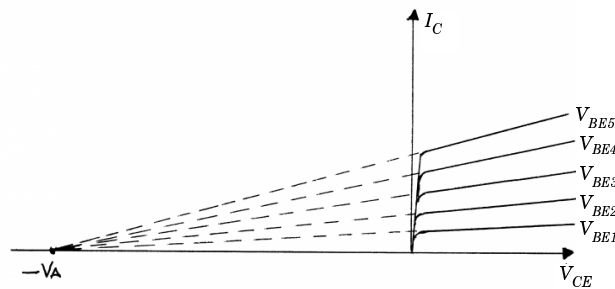
Se β_F fosse indipendente dal valore di I_C , sul piano delle caratteristiche di uscita si otterrebbero delle rette parallele ed equispaziate. In realtà β_F ha una certa dipendenza da I_C , per cui le caratteristiche non sono equispaziate. Inoltre esiste il

cosiddetto effetto Early, consistente nella riduzione della larghezza efficace della regione di base all'aumentare della polarizzazione inversa della giunzione base-collettore e il conseguente estendersi della zona di svuotamento relativa a tale giunzione. Una riduzione della lunghezza di base porta a una ridotta ricombinazione e quindi a un incremento della corrente di collettore. Pertanto, per valori di V_{CB} (e quindi di V_{CE}) più elevati la corrente di collettore aumenta a parità di corrente di base e quindi ciascuna delle curve corrispondenti alle caratteristiche di uscita risulta inclinata verso l'alto.

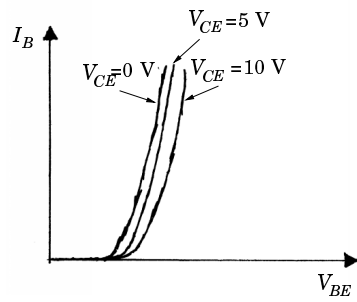
D'altro canto, per valori di tensione V_{CE} inferiori a V_γ la giunzione collettore-base non è più polarizzata inversamente e, per V_{CE} sufficientemente bassa (qualche decimo di volt) si entra in piena saturazione. In condizione di saturazione si perde il controllo della I_C da parte della I_B e tutte le caratteristiche tendono a sovrapporsi; in questa condizione il comportamento del transistor diventa fortemente non lineare. In figura rappresentiamo l'andamento delle caratteristiche di uscita a emettitore comune per un transistor *npn*



È possibile vedere che, prolungando verso sinistra con delle rette le caratteristiche di uscita costruite utilizzando V_{BE} come parametro invece di I_B , tutte tali rette intersecano l'asse delle ascisse in uno stesso punto che corrisponde alla cosiddetta "tensione di Early", in genere compresa tra -50 e -100 V.

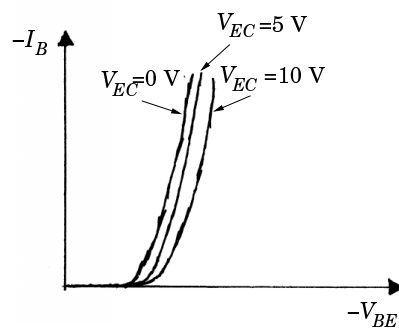
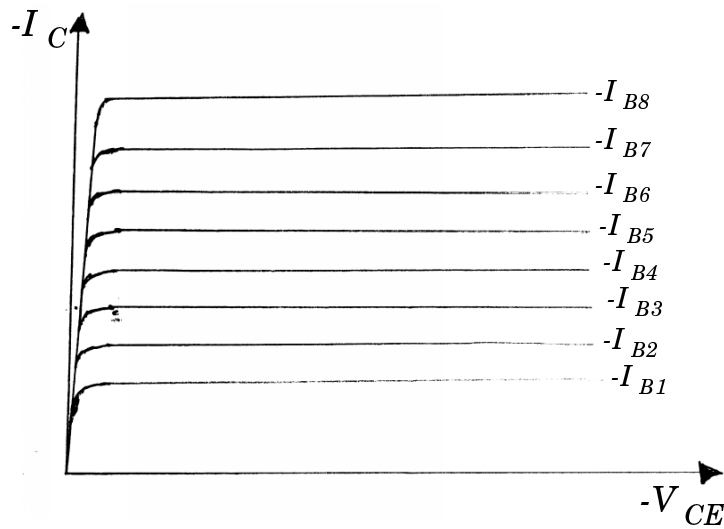


Le caratteristiche di ingresso a emettitore comune rappresentano la relazione tra I_B e V_{BE} in funzione della V_{CE} . Esse rappresentano il comportamento della giunzione base-emettitore e hanno quindi un andamento esponenziale simile a quello di un diodo; risentono inoltre dell'effetto Early. In particolare, se si incrementa V_{CE} , la lunghezza di base diminuisce, causando una minore ricombinazione e, di conseguenza, una diminuzione della corrente di base. Nella figura sottostante sono rappresentate le



caratteristiche di ingresso a emettitore comune per un transistor *npn* ed è possibile vedere chiaramente il risultato dell'effetto Early.

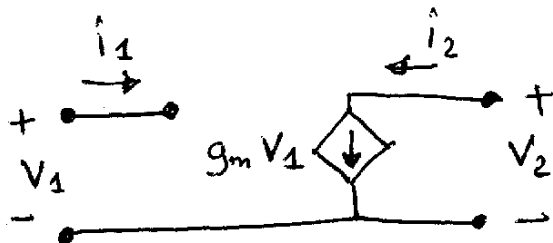
Le caratteristiche di ingresso e di uscita dei transistori *pnp* sono del tutto analoghe (con gli opportuni cambiamenti di segno) a quelle degli *npn*. Tali caratteristiche, riportate di seguito, si ottengono da quelle presentate per i transistori *npn* sostituendo I_B con $-I_B$, V_{BE} con $-V_{BE}$, V_{CE} con $-V_{CE}$, I_E con $-I_E$ e I_C con $-I_C$.



5. I transistori a effetto di campo

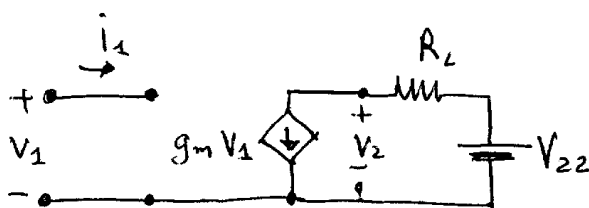
5.1 Generatore di corrente controllato in tensione

Nei transistori a effetto di campo, come suggerisce il nome stesso, il flusso di corrente viene controllato tramite un campo elettrico e, quindi, attraverso il valore della tensione applicata a un opportuno elettrodo di comando. Pertanto, così come nel caso dei BJT avevamo introdotto una rappresentazione idealizzata del funzionamento per mezzo di un generatore di corrente controllato in corrente, per lo studio dei transistori a effetto di campo introduciamo il concetto di generatore ideale di corrente controllato in tensione, la cui rappresentazione circuitale è indicata in figura.



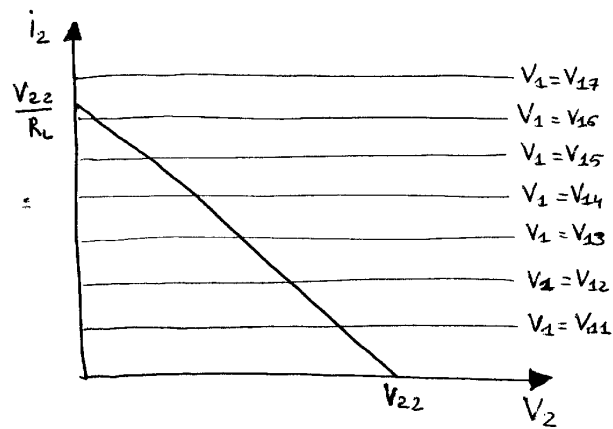
La quantità g_m rappresenta il rapporto tra la corrente di uscita e la tensione in ingresso. Dimensionalmente è quindi l'inverso di una resistenza ed è espressa in A/V o, più comunemente, in mA/V.

Possiamo anche in questo caso rappresentare le caratteristiche di uscita, sul piano v_2-i_2 , ottenendo una famiglia di curve in funzione del parametro costituito dalla tensione di ingresso. Tali curve altro non sono che rette parallele all'asse delle ascisse, poiché la corrente in uscita non dipende dalla tensione v_2 , ma soltanto dalla v_1 . Per questo tipo di generatore comandato possiamo pensare a una configurazione in cui sono presenti un generatore di alimentazione esterno V_{22} e una resistenza di carico R_L , come quella rappresentata nella figura sottostante.



È possibile tracciare sul piano delle caratteristiche di uscita una retta di carico, la cui pendenza corrisponde a $-1/R_L$ e che interseca l'asse delle ascisse in corrispondenza di V_{22} , come indicato sulle caratteristiche di seguito rappresentate.

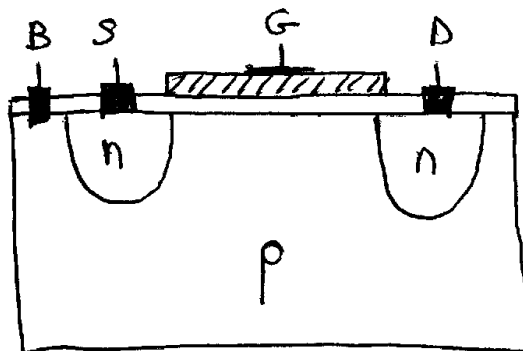
Al variare della tensione v_1 applicata in ingresso, il punto di lavoro si sposterà quindi lungo la retta di carico, dando luogo a una variazione della tensione di uscita v_2 che rappresenta una replica amplificata e invertita di fase della tensione di ingresso. Se le caratteristiche fossero equispaziate come quelle rappresentate in figura, la tensione di uscita presenterebbe una relazione perfettamente lineare con quella di ingresso; vedremo che in realtà i dispositivi a effetto di campo sono caratterizzati da una



dipendenza quadratica della corrente di uscita dalla tensione di ingresso, per cui il loro comportamento può considerarsi lineare solo per piccoli segnali.

5.2 Il transistor MOS

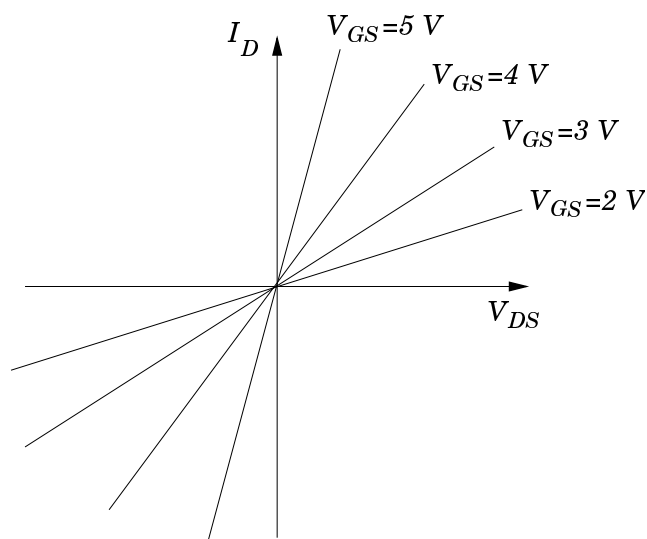
L'idea alla base del funzionamento del transistor MOS (Metal-Oxide-Semiconductor) è abbastanza semplice: si crea uno strato di cariche mobili in prossimità della superficie di un semiconduttore, tramite l'applicazione di un campo elettrico, per mezzo di un elettrodo di gate metallico, che "attira" le cariche verso la superficie stessa. Una descrizione appropriata del funzionamento del transistor MOS richiederebbe concetti avanzati di meccanica quantistica che esulano da questa trattazione, quindi sfrutteremo una descrizione intuitiva, ancorché non rigorosa, per svolgere lo studio del MOS. La struttura di un tipico transistor MOS a canale n è rappresentata nella figura seguente: in un blocco di silicio p vengono realizzate, a una certa distanza tra loro, due diffusioni di tipo n che rappresentano gli elettrodi di source e di drain; sopra la superficie del silicio viene ottenuto un sottile strato di ossido di silicio (dell'ordine della decina di nanometri o, più recentemente, dei nanometri) che funge da isolante; al di sopra dello strato isolante e tra source e drain viene realizzato uno strato metallico (o più recentemente di silicio policristallino fortemente drogato, in modo da renderlo conduttore) che funge da gate.



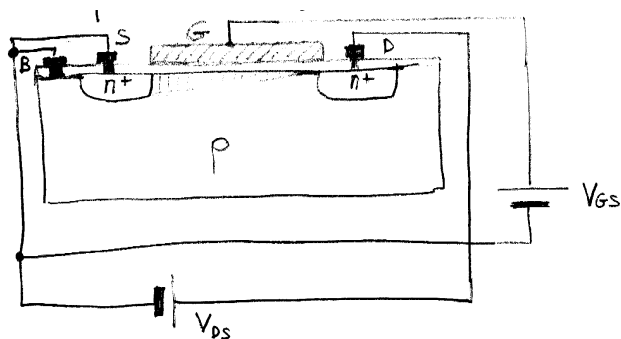
Se si applica una tensione positiva al gate rispetto al silicio p sottostante, vengono richiamati elettroni in prossimità della superficie, cosicché si ha, in un sottilissimo strato subito sotto la superficie stessa, un fenomeno di inversione: il silicio diventa localmente di tipo n , dando quindi luogo alla formazione di un canale conduttore che

mette in connessione source e drain. Il transistore MOS ha perciò 4 elettrodi: source, drain, gate e il quarto, denominato bulk, che corrisponde al silicio di substrato, quello nel quale si forma il canale. Nei dispositivi MOS non integrati l'elettrodo di bulk è di solito internamente collegato al source.

Il canale di portatori minoritari comincia a formarsi alla superficie quando la tensione tra gate e bulk supera un valore di soglia indicato con V_T . Poiché nei dispositivi discreti il bulk è di solito collegato con il source (come sopra accennato), la tensione tra gate e bulk corrisponde con la V_{GS} e come tale la indicheremo nel seguito. È importante sottolineare che la tensione V_T per i MOS non ha nulla a che fare con la V_T che abbiamo considerato nello studio dei BJT (pari a kT/q): la coincidenza dei nomi è del tutto fortuita. Per valori piccoli (inferiori a qualche decimo di volt) della V_{DS} , il canale del transistore MOS ha un andamento uniforme tra source e drain, con uno spessore, e conseguentemente una conducibilità, che cresceranno al crescere della V_{GS} . In questa condizione il MOS può quindi essere utilizzato come una resistenza variabile, controllata tramite la tensione V_{GS} , dando luogo a caratteristiche lineari sul piano $V_{CE} - I_C$.

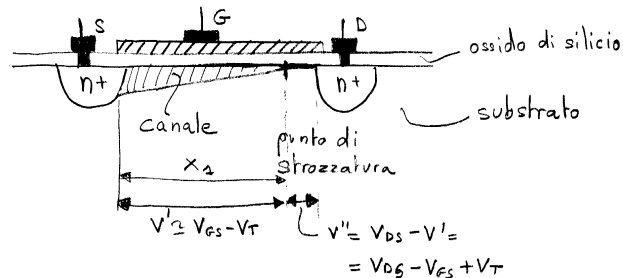


Se V_{DS} è maggiore di qualche decimo di volt, il comportamento del canale non è più quello di una resistenza lineare, poiché l'ampiezza verticale dello stesso verrà a dipendere dalla posizione considerata tra source e drain, come rappresentato nella figura seguente.



L'effettiva tensione tra gate e substrato dipende dalla posizione lungo il canale: sarà pari a V_{GS} in corrispondenza del source e diminuirà avvicinandosi al drain. Se $V_{GS} -$

$V_{DS} < V_T$ esisterà un punto in cui la tensione tra gate e substrato risulta pari a V_T (la tensione di soglia per la formazione del canale) e oltre tale punto risulta inferiore a V_T , quindi il canale viene strozzato. Per $V_{GS} - V_{DS} = V_T$ la strozzatura si trova in corrispondenza dell'estremità di drain; all'aumentare di V_{DS} si sposta verso sinistra, ma di una quantità piccola relativamente alla lunghezza del canale, almeno per i valori di V_{DS} normalmente utilizzati. Quindi si ha un tratto di canale di lunghezza pari quasi alla distanza tra drain e source ai capi del quale, una volta superata la V_{DS} per cui $V_{GS} - V_{DS} < V_T$, esiste una tensione pari a $V_{GS} - V_T$, sostanzialmente indipendente da ulteriori variazioni di V_{DS} . La corrente che lo attraversa è quindi anch'essa pressoché indipendente da V_{DS} e raggiunge la diffusione di drain attraverso la zona strozzata. Sul piano delle caratteristiche di uscita questo corrisponde ad avere caratteristiche pressoché orizzontali al di sopra del valore di V_{DS} per cui avviene lo strozzamento. Questa zona delle caratteristiche viene definita "zona di saturazione"; si noti che l'uso del termine "saturazione" in questo caso non ha nulla a che fare con quello che si ha nel caso del transistor BJT. La saturazione nel caso del BJT indica una condizione di funzionamento in cui tutte e due le giunzioni sono polarizzate direttamente e la corrente di base non ha quasi più alcuna influenza su quella di collettore; nel caso del transistor MOS la saturazione corrisponde alla condizione di funzionamento attivo che consente di ottenere un'amplificazione. Vediamo ora una rappresentazione grafica di quanto è stato fin qui descritto.



All'aumentare di V_{DS} , la distanza X_1 si riduce, anche se di poco, e questo fatto dà luogo all'inclinazione delle caratteristiche di uscita. Dato che tale riduzione è piuttosto piccola e che la tensione V' è praticamente costante, la corrente I_{DS} si può considerare, almeno in prima approssimazione, indipendente da V_{DS} nella zona di saturazione.

In condizioni di saturazione la corrente è data da

$$I_D = k \left(\frac{W}{L} \right) (V_{GS} - V_T)^2,$$

dove k rappresenta un parametro di processo, W è la larghezza in direzione perpendicolare al disegno e L è la lunghezza del canale, cioè la distanza tra drain e source. Il parametro di processo k ha le dimensioni di una corrente divisa per una tensione al quadrato e può essere espresso come

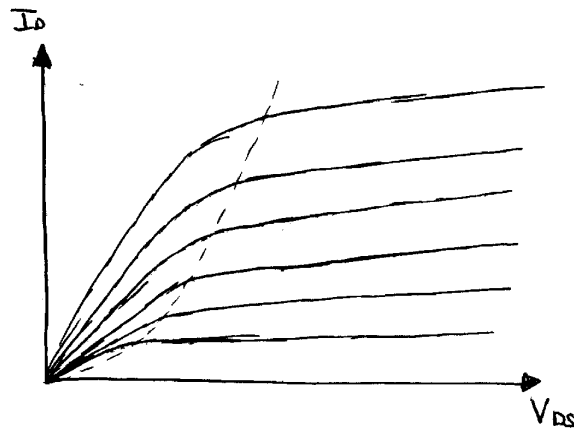
$$k = \mu_n \frac{C_{ox}}{2},$$

dove C_{ox} è la capacità dell'ossido per unità di superficie ($C_{ox} = \epsilon/T_{ox}$), con ϵ pari alla costante dielettrica dell'ossido e T_{ox} pari allo spessore dell'ossido) e μ_n è la mobilità dei portatori interessati alla conduzione tra source e drain, in questo caso elettroni.

Tale espressione è valida non solo per $V_{DS} = V_{GS} - V_T$, ma anche per valori superiori di V_{DS} , se facciamo l'ipotesi di corrente in zona di saturazione indipendente da V_{DS} . Poiché all'entrata in saturazione $V_{DS} = V_{GS} - V_T$, possiamo ricavare facilmente l'espressione della curva che separa sul piano delle caratteristiche di uscita, la regione cosiddetta "triode" (per V_{DS} minore di quella di saturazione) e quella di saturazione: si tratta di una parabola descritta dall'equazione

$$I_D = k \left(\frac{W}{L} \right) V_{DS}^2.$$

Le caratteristiche di uscita così ottenute sono rappresentate nella figura seguente.



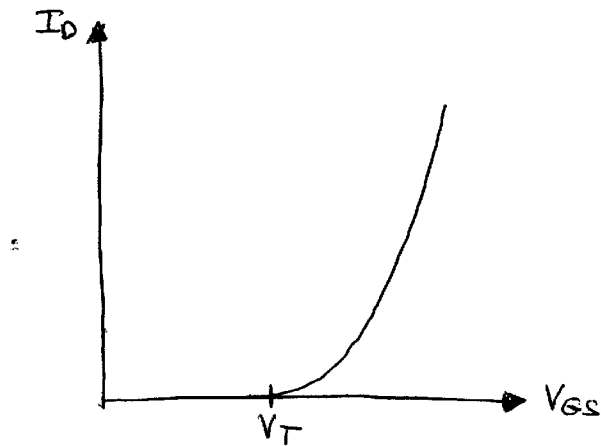
Se, nell'espressione della corrente in saturazione, vogliamo includere anche l'effetto di accorciamento del canale all'aumentare della V_{DS} , che dà luogo a una leggera inclinazione verso l'alto delle caratteristiche di uscita, possiamo usare l'espressione di tipo fenomenologico

$$I_D = k \left(\frac{W}{L} \right) (V_{GS} - V_T)^2 (1 + \lambda V_{DS}),$$

dove il parametro λ rappresenta l'inverso di una tensione, che viene di solito chiamata tensione di Early, in analogia a quanto accade nei BJT, anche se in questo caso il fenomeno fisico che porta all'inclinazione delle caratteristiche di uscita è completamente diverso da quello della riduzione della lunghezza efficace di base dei BJT.

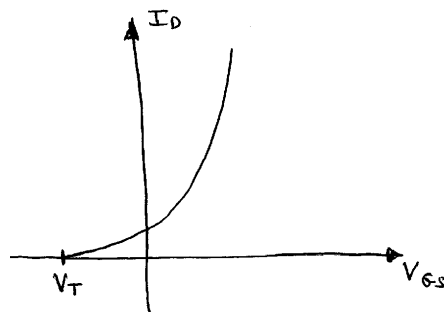
Sulla base delle espressioni finora derivate è anche possibile tracciare la transcaratteristica ingresso-uscita per un transistor MOS: si tratta di una parabola che inizia per $V_{GS} = V_T$. Di solito si trascura la dipendenza della transcaratteristica dalla V_{DS} e si traccia un'unica curva, indipendentemente dal valore di V_{DS} .

Per i transistori MOS esistono versioni complementari a canale p , nelle quali il substrato è di tipo n e le diffusioni di source e di drain sono di tipo p . In tali transistori il trasporto tra source e drain è ottenuto tramite una corrente di lacune, che vengono indotte alla superficie tramite l'applicazione di una tensione negativa tra gate e bulk. Pertanto le caratteristiche di un transistor a canale p sono equivalenti a quelle di un transistor a canale n , purché si scambino i segni di correnti e tensioni. I MOS a canale p hanno prestazioni inferiori, a parità di altre caratteristiche, rispetto ai corrispondenti transistori a canale n , in conseguenza della ridotta mobilità delle lacune rispetto agli elettroni. Si può compensare peraltro tale ridotta mobilità



con un aumento della larghezza di canale W , cosa che si fa molto frequentemente, dato che la disponibilità di transistori MOS complementari è alla base della tecnologia CMOS, che rappresenta, come vedremo più avanti, uno dei motori trainanti dell'attuale industria microelettronica.

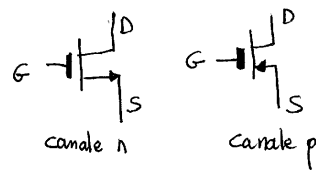
I transistori MOS finora visti sono di tipo ad arricchimento (o enhancement), poiché per tensione di gate nulla il canale non esiste e la sua formazione è una conseguenza dell'arricchimento di portatori del tipo opportuno operato tramite l'applicazione di una tensione di gate. Esistono anche transistori di tipo diverso, nei quali il canale è preesistente, ottenuto tramite impiantazione di droganti nello strato superficiale e lo si può far scomparire con l'applicazione di un'opportuna tensione di gate. Tali transistori si definiscono a svuotamento (o depletion). Nel caso del transistorore depletion a canale n la tensione di soglia V_T risulta quindi negativa e la caratteristica ingresso-uscita risulta traslata all'indietro rispetto a quella dell'equivalente enhancement:



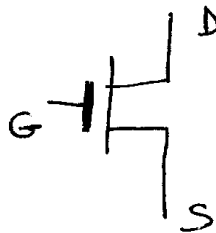
Le caratteristiche di uscita sono praticamente equivalenti a quelle del transistorore enhancement, con l'unica variante dei valori di tensione V_{GS} corrispondenti a ciascuna di tali caratteristiche, che risultano traslati verso il basso.

Rovesciando i segni di correnti e tensioni possono facilmente ottenersi le caratteristiche dei transistori a canale p , sia enhancement sia depletion.

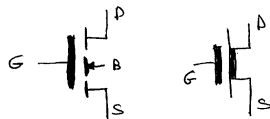
Esistono diversi simboli circuitali per i transistori MOS. Quelli più comunemente usati non danno informazioni sul tipo (ad arricchimento o a svuotamento), ma indicano soltanto se il canale è p o n :



Si utilizza talvolta anche un simbolo ancor più generale, valido sia per MOS a canale n sia per MOS a canale p :

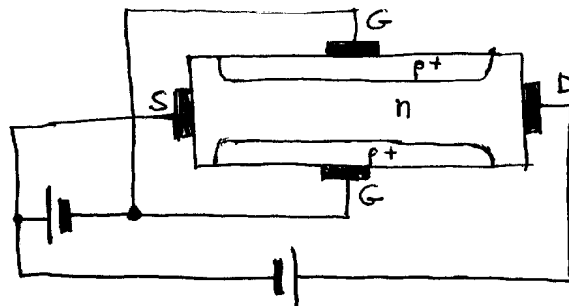


Esistono infine dei simboli, molto poco usati, che distinguono tra MOS ad arricchimento (a sinistra) e MOS a svuotamento (a destra):

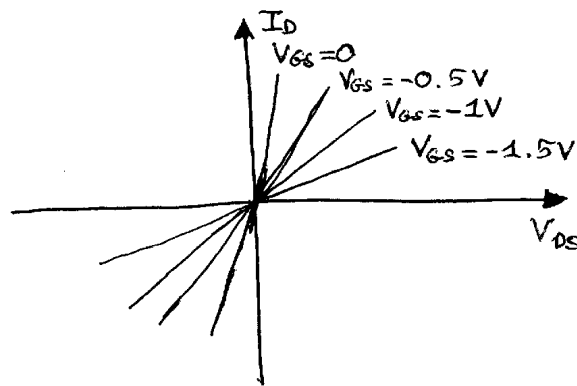


5.3 Il transistorore JFET

Il transistorore JFET (Junction Field Effect Transistor) è il primo transistorore a effetto di campo che fu realizzato e contiene, come è chiaro dal nome, una giunzione p - n . Prendiamo dapprima in considerazione il transistorore JFET a canale n , il cui funzionamento può essere compreso in base allo schema di principio riportato nella figura seguente.



La corrente scorre tra l'elettrodo di source e quello di drain attraverso un pezzo di semiconduttore drogato n , nel quale sono realizzate due diffusioni di tipo p^+ , che costituiscono l'elettrodo di gate. Essendo le diffusioni di gate più drogate, la zona di svuotamento si estende prevalentemente nella regione n . Se supponiamo di applicare una tensione molto piccola tra drain e source e una tensione negativa V_{GS} tra gate e source, in modo da polarizzare inversamente la giunzione, la corrente scorrerà in un canale con dimensione verticale variabile al variare di V_{GS} , compreso tra i bordi delle due zone di svuotamento dovute alle diffusioni p^+ . Per piccoli valori della tensione V_{DS} applicata tra drain e source, il JFET si comporta quindi in modo simile a quanto già visto per il MOS, come una resistenza il cui valore può essere variato per mezzo della V_{GS} : in questo caso tanto più grande è in modulo il valore della V_{GS} tanto maggiore risulta il valore di resistenza presente tra drain e source. Tutto ciò può essere rappresentato tramite le caratteristiche di uscita nella regione per V_{DS} piccola, che sono sostanzialmente delle rette che confluiscono nell'origine, come illustrato nella figura seguente. La pendenza della caratteristica corrisponde all'inverso della resistenza e diminuisce all'aumentare in modulo della V_{GS} .

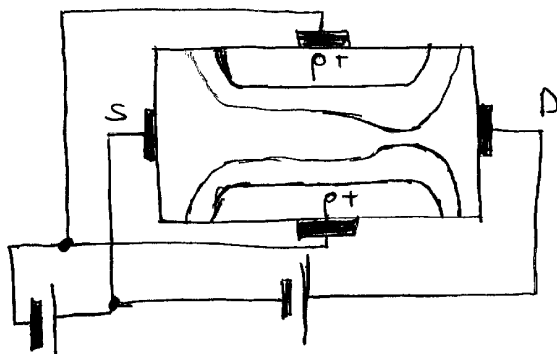


Per V_{GS} abbastanza negativa le due zone di svuotamento si congiungono e il canale scompare, per cui la corrente si annulla: il valore di V_{GS} per il quale questo avviene si definisce tensione di pinch-off e si indica con V_P . Per tensioni V_{GS} in modulo maggiori di V_P il transistor JFET si comporta come un interruttore aperto.

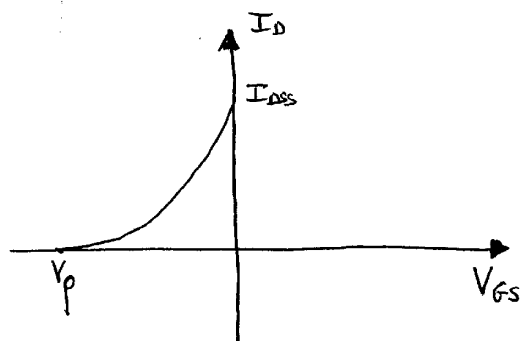
Quando si applica invece una V_{DS} non trascurabile (superiore a qualche decimo di volt), il comportamento del dispositivo non può più essere assimilato a quello di una resistenza lineare: per una V_{DS} significativa, la differenza di potenziale tra drain e canale risulta maggiore in modulo di quella tra source e canale, per cui le zone di svuotamento che definiscono il canale stesso assumono una forma asimmetrica, risultando più ampie in corrispondenza della regione di drain.

Al crescere di V_{DS} , il canale diviene sempre più stretto fino a risultare strozzato in corrispondenza del drain, come indicato nella figura seguente. In questa condizione il canale si estende dall'estremità di source fino al punto dove inizia la strozzatura. Si ha in tal caso un fenomeno di saturazione del tutto analogo a quello già visto per i transistori MOS. Si ha quindi un comportamento che approssima quello del generatore ideale di corrente controllato in tensione. Le caratteristiche corrispondenti a valori equispaziati della tensione V_{GS} non sono però equispaziate: si può dimostrare che in saturazione la corrente di drain I_D in un JFET ha una dipendenza quadratica dalla tensione V_{GS} , secondo la seguente equazione:

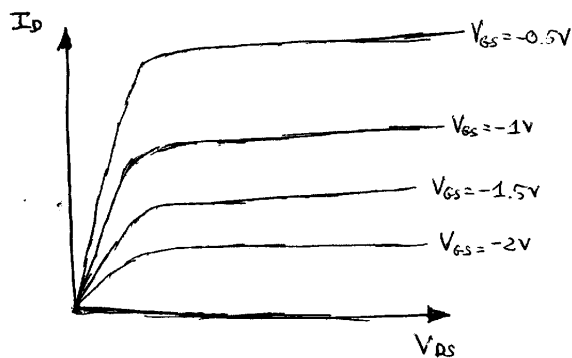
$$I_D = I_{DSS} \left(1 - \frac{V_{GS}}{V_P} \right)^2,$$



dove I_{DSS} è la corrente di drain che si ha per V_{GS} nulla. Tale relazione rappresenta la transcaratteristica ingresso-uscita e può essere riportata graficamente sul piano $V_{GS}-I_D$.

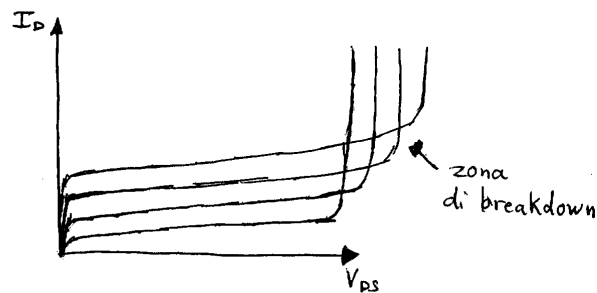


In prima approssimazione possiamo assumere che la transcaratteristica sia indipendente dal valore di V_{DS} (cosa che faremo in tutte le applicazioni numeriche), che corrisponde ad assumere, come abbiamo fatto in precedenza, orizzontali le caratteristiche di uscita in saturazione. In realtà queste non sono perfettamente orizzontali, a causa del fatto che il punto in cui inizia la strozzatura si sposta leggermente verso il source all'aumentare della tensione V_{DS} per cui la stessa tensione V_C viene applicata a un canale di lunghezza minore, dando così luogo a una corrente di drain maggiore. L'andamento delle caratteristiche di uscita in zona di saturazione è perciò inclinato, in modo simile a quanto accade nei BJT a causa dell'effetto Early. Le caratteristiche di uscita di un JFET a canale n hanno perciò l'aspetto sotto rappresentato.



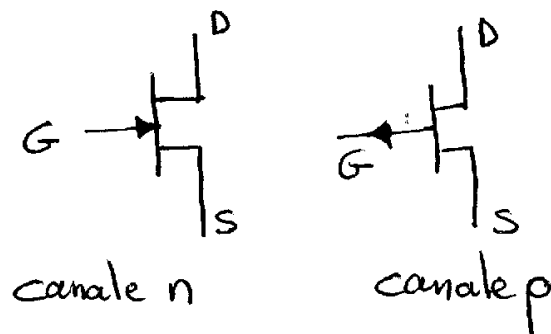
La giunzione tra gate e canale è polarizzata inversamente, quindi viene attraversata da una corrente molto piccola di valore compreso tra le decine di picoampere e qualche

nanoampere, a seconda dell'area di giunzione e della temperatura. Per V_{DS} elevata la tensione inversa applicata sulla giunzione tra gate e canale può superare il valore di breakdown e si assiste in tal caso a un rapido aumento della corrente di drain. È da notare che il breakdown avviene per valori di V_{DS} tanto minori quanto più grande è, in modulo, la V_{GS} . Questo perché la tensione effettivamente presente ai capi della giunzione nell'area di drain risulta pari a $V_{DS} - V_{GS}$.



Finora abbiamo parlato di transistori JFET con canale n e gate realizzato con una diffusione p^+ ; è possibile avere anche una realizzazione di tipo complementare, con canale p e diffusioni di gate di tipo n^+ . Il funzionamento è del tutto analogo, con la sola differenza che tutte le tensioni e le correnti avranno segno invertito: la tensione V_{GS} dovrà essere positiva, mentre la V_{DS} sarà negativa. Le prestazioni dei transistori JFET a canale p differiscono, a parità di drogaggi e di caratteristiche geometriche, da quelle dei JFET a canale n , a causa della diversa mobilità delle lacune rispetto agli elettroni. In genere le prestazioni dei JFET a canale n sono dunque migliori, ed è questo il motivo per cui essi risultano molto più utilizzati.

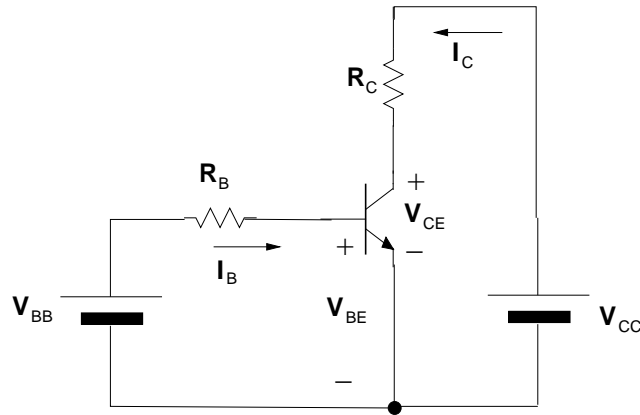
Il simbolo circuitale del JFET è sotto rappresentato, sia per il tipo a canale n sia per quello a canale p .



6. Reti di polarizzazione

6.1 Reti di polarizzazione per transistori BJT

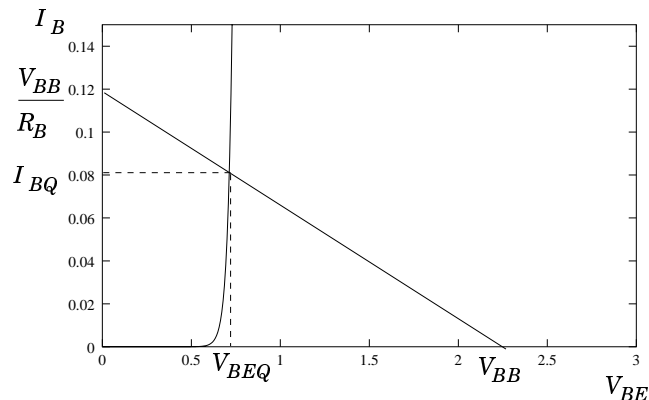
La rete di polarizzazione più semplice possibile da un punto di vista concettuale è quella di seguito riportata:



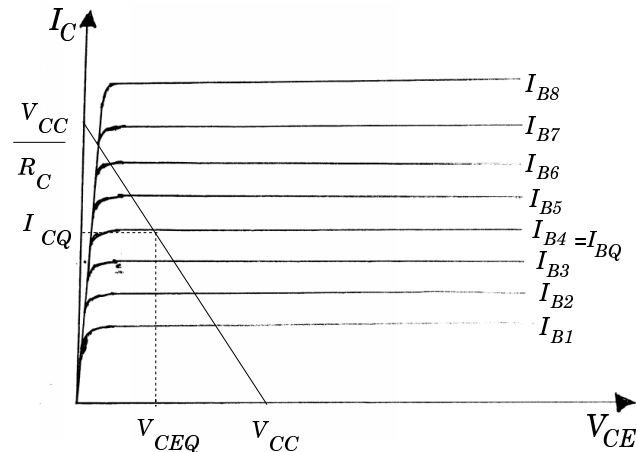
Le quantità incognite sono V_{CE} , I_C , V_{BE} , I_B . Abbiamo quindi la necessità di quattro equazioni che legano tra loro queste quantità. Tali equazioni possono essere ricavate dalle relazioni tra le varie grandezze imposte dal circuito esterno e dalle relazioni imposte dal transistoro stesso.

$$\begin{aligned}V_{BE} &= g(I_B, V_{CE}) \\ I_C &= f(I_B, V_{CE}) \\ V_{BB} &= R_B I_B + V_{BE} \\ V_{CC} &= R_C I_C + V_{CE},\end{aligned}$$

dove le prime due equazioni rappresentano le caratteristiche del transistoro (rispettivamente di ingresso e di uscita) e le altre sono semplicemente le equazioni alle due maglie dalle quali il circuito è costituito. Se le caratteristiche di ingresso corrispondenti a diversi valori di V_{CE} non differiscono significativamente tra loro, possiamo determinare il valore di riposo I_{BQ} della I_B tracciando sul grafico delle caratteristiche di ingresso stesse la retta di carico definita dalla terza equazione.



A questo punto possiamo costruire la retta di carico corrispondente alla Eq. (4) sul piano delle caratteristiche di uscita e individuare i valori di riposo della I_{CQ} della I_C e V_{CEQ} della V_{CE} dall'intersezione di tale retta con la caratteristica relativa alla I_{BQ} prima determinata, completando così la determinazione del punto di lavoro del transistor:

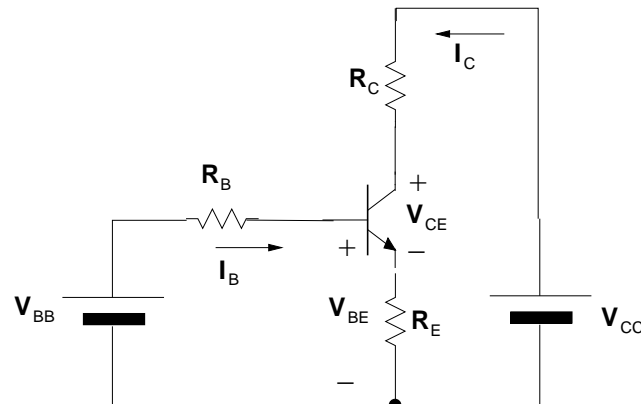


La procedura seguita può essere semplificata notando che in zona attiva diretta la V_{BE} non si discosta da V_γ più di 0.1-0.15 V, per cui possiamo assumere, senza commettere errori significativi, $V_{BE} \simeq V_\gamma$. In tal caso non è più necessario ricorrere alle caratteristiche di ingresso. Dalla equazione delle cadute di tensione sulla maglia di ingresso otteniamo direttamente

$$I_B = \frac{V_{BB} - V_\gamma}{R_B},$$

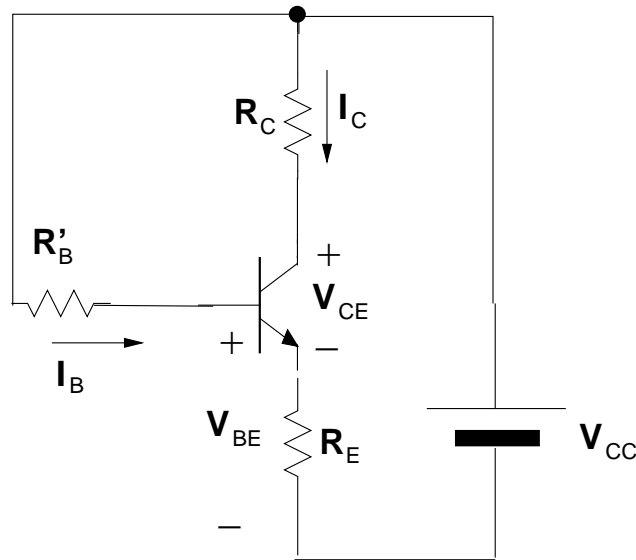
e per il resto si procede nel modo già visto.

Il circuito di polarizzazione esaminato presenta il vantaggio della facilità di analisi, ma ha anche gravi inconvenienti, tali da renderlo raramente utilizzato in pratica. Gli inconvenienti consistono nel fatto che il punto di lavoro è fortemente legato ai parametri del dispositivo e, soprattutto, varia al variare della temperatura. Infatti, se per esempio la temperatura aumenta, diminuisce la V_{BE} necessaria per avere una data I_B . Di conseguenza aumenta la corrente di base e quindi anche quella di emettitore, causando un ulteriore incremento della temperatura della giunzione. Si ha quindi un meccanismo che conduce a forti scostamenti dal punto di riposo e può portare anche alla distruzione del dispositivo. Inserendo una resistenza in serie all'emettitore si ottiene un notevole miglioramento, perché un eventuale incremento di corrente porta a un aumento della caduta sulla resistenza di emettitore e quindi a una diminuzione della V_{BE} , che contrasta l'originario aumento di corrente:



Il calcolo del punto di riposo in presenza di una resistenza di emettitore è più complesso, poiché le due maglie, quella di uscita e quella di ingresso, non sono più disaccoppiate, quindi non è più possibile calcolare la I_B indipendentemente dalla conoscenza della I_C . Per una valutazione esatta del punto di riposo bisogna ricorrere a procedure grafiche abbastanza complesse o a procedure numeriche iterative, che non prenderemo in esame. Considereremo invece un caso particolarmente semplice, che corrisponde alla maggioranza delle situazioni che si incontrano nella pratica.

Innanzitutto va detto che nei circuiti di polarizzazione visti finora è necessario disporre di due generatori distinti: V_{BB} e V_{CC} . Ciò rappresenterebbe in pratica un grosso problema e si cerca invece di avere un unico generatore di alimentazione per molti transistori. Vediamo come si può passare da due generatori di alimentazione a uno soltanto nel circuito appena considerato: basta utilizzare V_{CC} al posto di V_{BB} e variare R_B opportunamente, in modo da ottenere la stessa corrente di base che si sarebbe ottenuta con V_{BB} . Definiamo R'_B la nuova R_B .

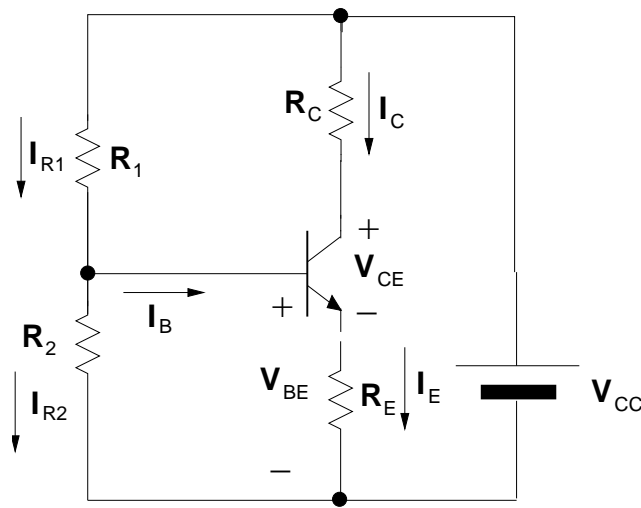


L'altro inconveniente prima citato consiste nel fatto che il punto di lavoro dipende fortemente dalle caratteristiche intrinseche del transistor, le quali possono variare notevolmente, a causa della dipendenza da parametri costruttivi come la larghezza di base che non possono essere controllati con grande precisione. Il circuito che più spesso si utilizza per la polarizzazione dei BJT risolve anche questo problema e si basa sull'impiego di un partitore "pesante" per fissare la tensione di base.

Per partitore pesante si intende un partitore realizzato con resistenze tali da far passare una corrente molto maggiore di quella che da tale partitore viene derivata, in questo caso particolare la corrente di base. In una tale situazione la tensione che si ottiene con il partitore è sostanzialmente determinata dal rapporto delle resistenze. Perché questa approssimazione sia valida è necessario che la corrente nelle resistenze sia almeno venti volte più grande di quella di base. Di solito, nella risoluzione dei circuiti, si fa l'ipotesi di partitore pesante e poi si verifica al termine se la I_B trovata è consistente con l'ipotesi stessa.

Nell'ipotesi partitore pesante possiamo dunque scrivere che la tensione V_B della base rispetto a massa è data da

$$V_B = V_{CC} \frac{R_2}{R_1 + R_2}.$$



Se supponiamo che la tensione V_{BE} sia pari a V_γ , possiamo subito determinare V_E : $V_E = V_B - V_\gamma$. Quindi

$$I_E = \frac{V_B - V_\gamma}{R_E}.$$

Si noti che in questo caso I_E è stata scelta come uscente dall'emettitore, in modo che abbia segno positivo, per una maggiore comodità nei calcoli. Facciamo ora l'ipotesi, anch'essa da verificare al termine, che $I_B \ll I_C$. In tal caso $I_E \simeq I_C$ e quindi possiamo calcolare la caduta su R_C utilizzando il valore di I_E che abbiamo ottenuto:

$$V_C = V_{CC} - R_C I_C \simeq V_{CC} - R_C I_E.$$

Dunque

$$V_{CE} = V_{CC} - R_C I_C - V_B + V_\gamma.$$

Essendo a questo punto note V_{CE} e I_C , possiamo determinare I_B dalle caratteristiche di uscita e emettitore comune, individuando il punto di riposo identificato da tali valori e ottenendo la I_B corrispondente per interpolazione. Dobbiamo ora verificare che siano effettivamente soddisfatte le disuguaglianze relative alle ipotesi precedentemente fatte:

$$I_B \ll I_{R1}$$

$$I_B \ll I_C.$$

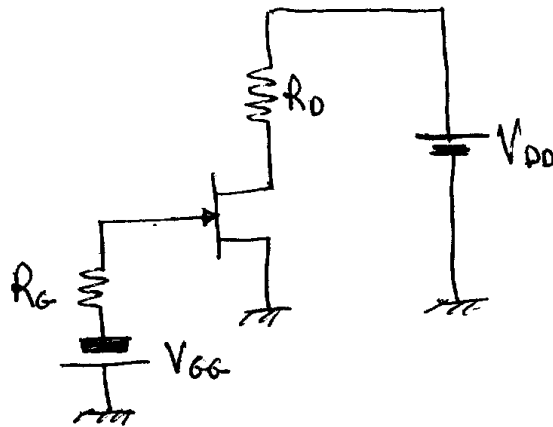
Si noti che per soddisfare la prima delle due disuguaglianze non si può incrementare arbitrariamente I_{R1} , diminuendo il valore delle resistenze del partitore, perché altrimenti avremmo un eccessivo consumo di corrente e un'eccessiva dissipazione di potenza sul partitore. Sottolineiamo anche il fatto che con l'ipotesi di partitore pesante sovrastimiamo la corrente di base (poiché ipotizziamo un valore di V_{BE} maggiore di quello reale), quindi non c'è il rischio che la verifica a posteriori ci tragga in inganno (l'effettiva corrente di base sarà comunque minore di quella stimata).

Dai risultati ottenuti nell'analisi del circuito di polarizzazione con partitore pesante risulta evidente come il punto di lavoro sia fissato, eccetto che per la I_B , dai valori delle quattro resistenze utilizzate e dipenda molto poco dalle caratteristiche intrinseche del transistor, risolvendo quindi il problema che era stato precedentemente messo in evidenza.

6.2 Reti di polarizzazione per transistori a effetto di campo

La polarizzazione dei transistori a effetto di campo si ottiene con reti molto semplici, grazie al fatto che l'assorbimento di corrente da parte del gate è trascurabile e che il problema della fuga termica non si pone. Prendiamo innanzitutto in esame quest'ultimo punto: avevamo visto che nel caso dei BJT la resistenza di emettitore era indispensabile per ottenere una stabilizzazione del punto di lavoro del transistor. Questa era la conseguenza del fatto che nel BJT la corrente è trasportata prevalentemente da portatori minoritari, la cui concentrazione dipende fortemente dalla temperatura. Nel caso dei transistori a effetto di campo la corrente viene trasportata da portatori maggioritari la cui concentrazione è praticamente indipendente dalla temperatura e la cui mobilità diminuisce all'aumentare della temperatura. Pertanto nei transistori a effetto di campo la corrente tende a diminuire all'aumentare della temperatura, con una conseguente stabilizzazione termica intrinseca.

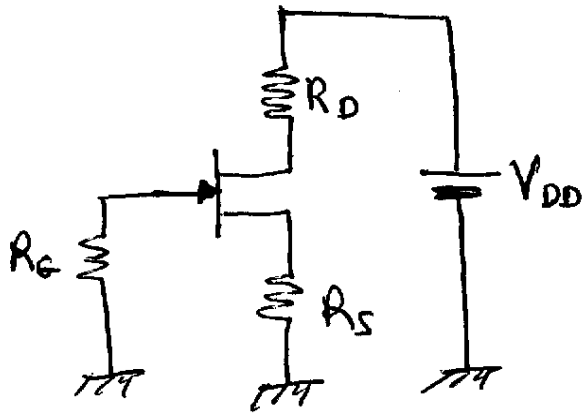
Per polarizzare un transistor JFET sarebbe quindi sufficiente uno schema del tipo riportato in figura, dove V_{DD} fornisce la tensione di polarizzazione necessaria per il drain e V_{GG} fornisce l'opportuna tensione negativa al gate (la presenza della resistenza R_G è necessaria per evitare di cortocircuitare l'eventuale generatore di segnale connesso al gate).



Un tale circuito non risulta però conveniente dal punto di vista pratico, perché contiene due generatori di tensione, uno dei quali deve fornire una tensione negativa rispetto a massa e non può quindi essere realizzato con una semplice partizione della tensione di alimentazione V_{DD} . Per tale motivo questo circuito non è quasi mai utilizzato in pratica e al suo posto si realizza quello riportato di seguito, consistente nella cosiddetta autopolarizzazione del JFET:

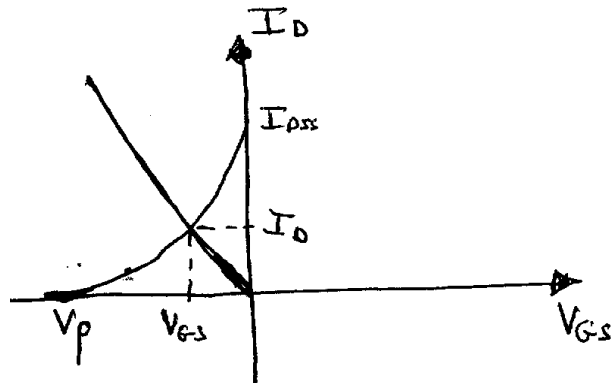
In questo circuito la resistenza R_S dà luogo a una caduta di tensione corrispondente alla V_{GS} necessaria per la polarizzazione (si parla proprio per questo motivo di autopolarizzazione), mentre la R_G ha la funzione di fissare il gate al potenziale di massa, senza peraltro cortocircuitare il gate stesso a massa. La R_G può anche avere dei valori piuttosto elevati, dell'ordine dei megaohm, senza che su di essa si manifesti una caduta di tensione misurabile, dato il piccolissimo valore della corrente di gate (che ricordiamo corrisponde alla corrente attraverso una giunzione polarizzata inversamente). Possiamo scrivere un'equazione molto semplice:

$$I_D = -\frac{V_{GS}}{R_S},$$



che corrisponde, sul piano della transcaratteristica ingresso-uscita, a una retta, passante per l'origine e con pendenza pari a $-1/R_S$. Il punto di lavoro corrisponderà all'intersezione tra tale retta e la transcaratteristica stessa. Per quanto riguarda la maglia di uscita è possibile scrivere un'altra semplice equazione, che ci permette di calcolare la V_{DS} :

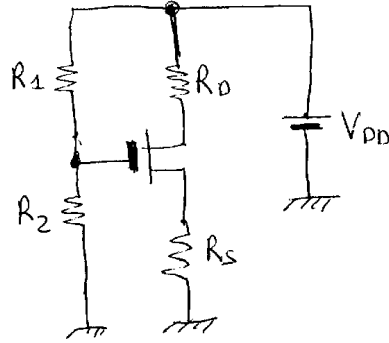
$$V_{DD} = I_D(R_D + R_S) + V_{DS}.$$



È quindi possibile realizzare un circuito che approssima un generatore ideale di corrente utilizzando semplicemente un transistor JFET e una resistenza, come nello schema di seguito riportato. Il valore della corrente erogata dipenderà dalla transcaratteristica del dispositivo e dal valore della resistenza di source.



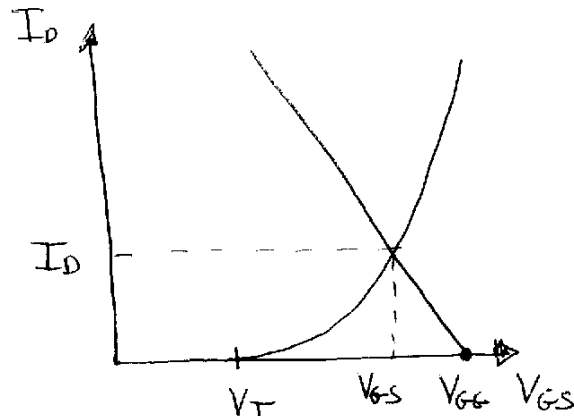
Per la polarizzazione dei transistori MOS si ricorre a schemi di analogia semplicità. Consideriamo per esempio il caso di un transistore MOS a canale n ad arricchimento.



Il partitore formato da R_1 e R_2 consente di applicare al gate del transistore MOS una tensione V_{GG} pari a $V_{DD}R_2/(R_1 + R_2)$. Poiché $V_{GS} = V_{GG} - I_D R_S$, avremo l'equazione

$$I_D = -\frac{1}{R_S}V_{GS} + \frac{V_{GG}}{R_S},$$

che rappresenta una retta sul piano della transcaratteristica di uscita. L'intersezione tra tale retta e la transcaratteristica individuerà il punto di riposo cercato.



È chiaro che, almeno dal punto di vista del funzionamento in continua, si può fare a meno senza problemi della resistenza di source R_S , dato che la tensione V_{GS} deve comunque essere positiva per questo tipo di dispositivo e non esistono problemi di stabilizzazione termica.

7. Funzionamento linearizzato per piccoli segnali

7.1 Modello linearizzato per transistori BJT

Come nel caso già visto per il diodo, è possibile sviluppare un modello linearizzato anche per i componenti attivi e, in particolare, anche per i transistori BJT. Questo modello è valido soltanto per piccoli spostamenti intorno al punto di lavoro, tali da poter approssimare il comportamento del transistoro considerando i soli termini del primo ordine dello sviluppo in serie (nel caso del transistoro si tratta, a differenza del diodo, dello sviluppo in serie di funzioni di più variabili).

Il transistoro è un quadripolo, quindi il suo modello linearizzato dovrà essere rappresentabile con un quadripolo lineare. Il comportamento elettrico di un quadripolo lineare può essere completamente definito tramite le relazioni tra quattro grandezze: i_1 , v_1 , i_2 , v_2 , le quali sono scelte con i versi illustrati nella figura seguente:



Eccetto che in casi particolari, è di solito possibile esprimere due a scelta delle quattro grandezze in funzione delle due rimanenti. Le relazioni tra tali grandezze risulteranno anch'esse lineari. Scegliamo di esprimere v_1 e i_2 in funzione di i_1 e v_2 (questa è la scelta di solito fatta per la rappresentazione linearizzata dei transistori). Otterremo:

$$\begin{cases} v_1 = h_{11}i_1 + h_{12}v_2 \\ i_2 = h_{21}i_1 + h_{22}v_2. \end{cases}$$

I parametri h_{ij} che definiscono questo modello sono detti parametri ibridi (perché le grandezze indipendenti sono di tipo tra loro diverso: una tensione e una corrente) e si indicano con la lettera h da “hybrid”.

Esaminiamo in dettaglio ciascuno dei parametri h , individuando le relative dimensioni:

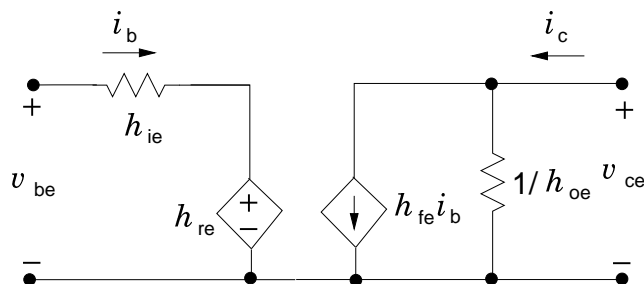
$$\begin{aligned} h_{11} = h_i(\text{input}) &= \left. \frac{v_1}{i_1} \right|_{v_2=0} && [\Omega] \\ h_{12} = h_r(\text{reverse}) &= \left. \frac{v_1}{v_2} \right|_{i_1=0} && \text{rapporto adimensionale} \\ h_{21} = h_f(\text{forward}) &= \left. \frac{i_2}{i_1} \right|_{v_2=0} && \text{rapporto adimensionale} \\ h_{22} = h_o(\text{output}) &= \left. \frac{i_2}{v_2} \right|_{i_1=0} && [\Omega^{-1}] \end{aligned}$$

La notazione con pedici letterali indica l'effetto connesso a ciascuno dei parametri h : h_i esprime il rapporto tra la tensione e la corrente in ingresso, quindi una quantità relativa alla sola maglia di ingresso; h_r indica l'azione della tensione di uscita su quella di ingresso, quindi un effetto di tipo inverso; h_f rappresenta il rapporto tra la corrente nella maglia di uscita e quella nella maglia di ingresso, dunque l'effetto

diretto e desiderato; h_o infine esprime il rapporto tra la corrente e la tensione di uscita ed è quindi relativo alla sola maglia di uscita. Di solito viene preso in considerazione il quadripolo corrispondente al montaggio del transistor a emettitore comune, vale a dire con l'emettitore a comune tra l'ingresso e l'uscita. In questa configurazione $v_1 = v_{be}$, $v_2 = V_{ce}$, $i_1 = i_b$ e $i_2 = i_c$. Per specificare che i parametri h si riferiscono al montaggio a emettitore comune si aggiunge un secondo pedice e :

$$\begin{cases} v_{be} = h_{ie} i_b + h_{re} v_{ce} \\ i_c = h_{fe} i_b + h_{oe} v_{ce} \end{cases}$$

Poiché le relazioni appena viste definiscono un quadripolo lineare, possiamo rappresentare tale quadripolo con un circuito equivalente, cercando di identificare il significato fisico dei parametri ibridi:



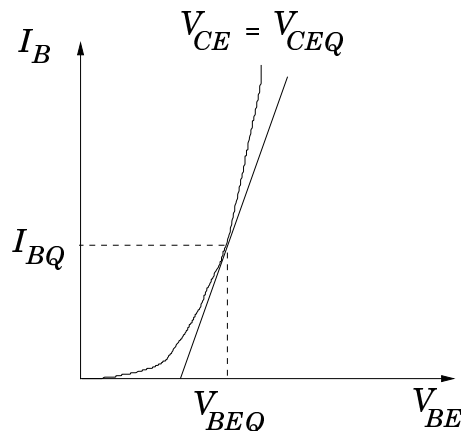
Fino a ora abbiamo rappresentato i parametri ibridi come rapporti tra le grandezze dinamiche, vale a dire come rapporti tra piccole variazioni delle grandezze totali. Possiamo quindi anche indicarli e calcolarli come limite dei rapporti incrementali delle grandezze totali e, di conseguenza, come derivate parziali di tali grandezze (nelle equazioni che seguono useremo, per conformità con l'uso comune, la notazione con simbolo e pedice maiuscolo per indicare le grandezze totali, anche se sarebbe più corretto indicarle con simbolo minuscolo e pedice maiuscolo, come si fa in generale nell'analisi dei circuiti). Per esempio, possiamo scrivere h_{fe} come:

$$h_{fe} = \lim_{\delta I_B \rightarrow 0} \frac{\delta I_C}{\delta I_B} \Big|_{V_{CE}=V_{CEQ}} = \frac{\partial I_C}{\partial I_B} \Big|_{V_{CE}=V_{CEQ}},$$

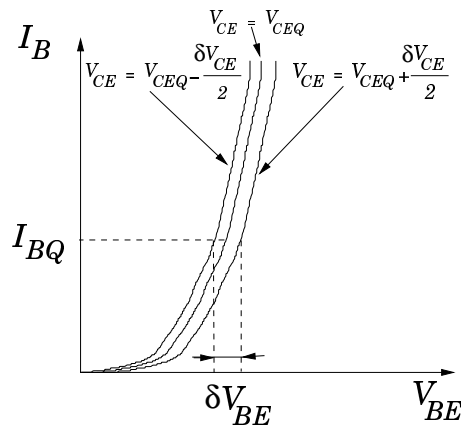
dove la derivata parziale viene valutata per V_{CE} costante e pari al valore nel punto di riposo, che corrisponde ad assumere una variazione nulla rispetto al punto di riposo, e quindi una v_{ce} nulla.

Vediamo ora come si determinano i parametri ibridi a emettitore comune a partire dalle caratteristiche del transistor, facendo riferimento, per semplicità, al caso di un transistor NPN. Consideriamo dapprima una procedura esclusivamente grafica, basata sulle caratteristiche di ingresso e di uscita.

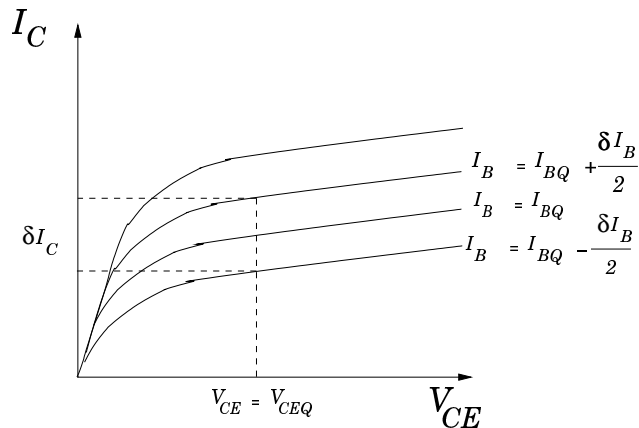
Dalla definizione, h_{ie} può essere approssimato con il rapporto tra la variazione della V_{BE} e quella della I_B , per V_{CE} costante e pari al valore di riposo. Si tratta quindi di prendere in considerazione la caratteristica di ingresso a emettitore comune per $V_{CE} = V_{CEQ}$ e misurare la pendenza della tangente nel punto corrispondente a V_{BEQ} . L'inverso di tale tangente corrisponde proprio a h_{ie} .



Dalla definizione, h_{re} può essere approssimato con il rapporto tra la variazione della V_{BE} e quella della V_{CE} , per I_B costante e pari al valore di riposo. Si tratta quindi di prendere in considerazione le caratteristiche di ingresso a emettitore comune e valutare la differenza tra le V_{BE} che si ottengono sommando o sottraendo $\delta V_{CE}/2$ a V_{CEQ} . Il rapporto tra tale variazione δV_{BE} e quella della tensione di uscita δV_{CE} ci fornisce il valore di h_{re} .

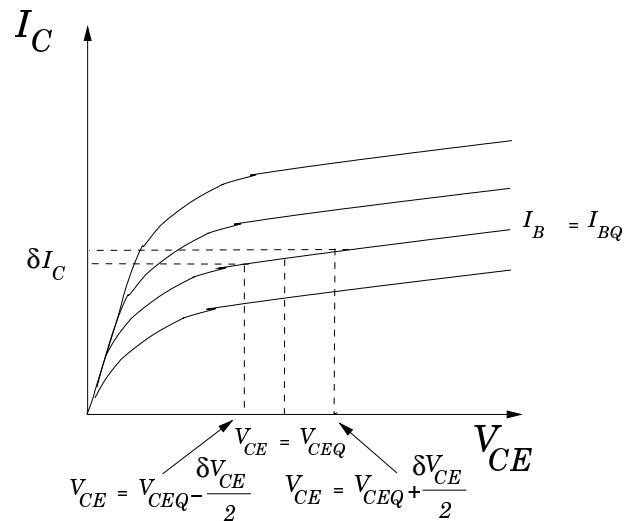


Il parametro h_{fe} può essere approssimato con il rapporto tra la variazione della I_C e quella della I_B , per V_{CE} costante e pari al valore di riposo. Dobbiamo quindi prendere in considerazione le caratteristiche di uscita a emettitore comune e valutare la differenza tra le I_C che si ottengono sommando o sottraendo δI_B a I_{BQ} . Il rapporto tra tale variazione δI_C e quella della corrente di ingresso δI_B ci fornisce il valore di h_{fe} .



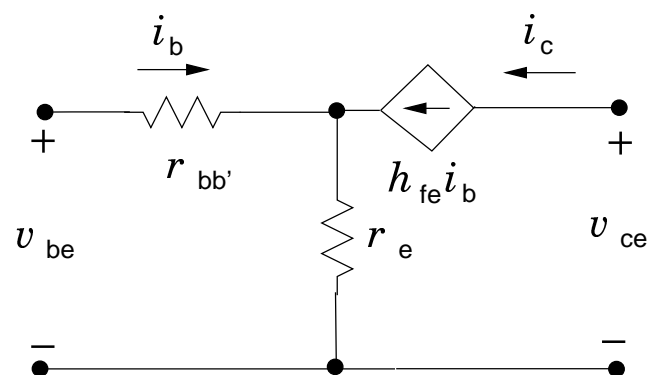
Infine il parametro h_{oe} può essere approssimato con il rapporto tra la variazione della

I_C e quella della V_{CE} , per I_B costante e pari al valore di riposo. Dobbiamo quindi prendere in considerazione le caratteristiche di uscita a emettitore comune e valutare la pendenza della tangente alla caratteristica per $I_B = I_{BQ}$ nel punto di riposo. Il valore di h_{oe} corrisponderà proprio a tale pendenza. Di solito la tangente è quasi orizzontale, dato il piccolo valore di h_{oe} , quindi conviene considerare la variazione di I_C su un intervallo molto ampio, per ottenere una precisione accettabile.



Nella pratica h_{oe} e h_{fe} si valutano effettivamente con la procedura vista, h_{re} si pone di solito pari a zero e h_{ie} si valuta invece in modo diverso, perché le caratteristiche di ingresso sono raramente disponibili.

Innanzitutto dobbiamo precisare che h_{ie} è costituita da due componenti in serie, $r_{bb'}$ e $r_{b'e}$. La componente $r_{bb'}$ corrisponde a una resistenza indesiderata, dovuta alla porzione di base che va dal contatto di ingresso alla zona attiva (base intrinseca). Tale resistenza parassita può avere valori compresi tra pochi ohm e un kilohm, a seconda delle caratteristiche costruttive del transistor (di solito è minore nei transistori di potenza maggiore). La $r_{b'e}$ è invece l'effettiva resistenza associata con il funzionamento del transistor. Per calcolare $r_{b'e}$ consideriamo un modello per le variazioni per il transistor in cui introduciamo una resistenza differenziale r_e a comune tra la maglia di ingresso e quella di uscita:



Tale resistenza differenziale è quella della giunzione base-emettitore, che può essere valutata a partire dalle equazioni di Ebers-Moll:

$$I_E = I_{ES} \left(e^{V_{BE}/V_T} - 1 \right) \simeq I_{ES} e^{V_{BE}/V_T}.$$

Il termine relativo al generatore comandato dalla corrente nel diodo di collettore non è stato indicato data la polarizzazione inversa della giunzione collettore-base. Si può inoltre trascurare l'unità rispetto a e^{V_{BE}/V_T} , dato che $V_{BE} \simeq V_\gamma$ e quindi $e^{V_{BE}/V_T} \gg 1$.

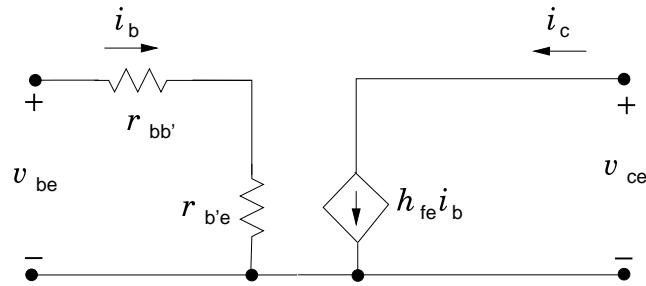
Possiamo quindi ottenere il reciproco di r_e derivando la corrente di emettitore rispetto alla tensione base intrinseca-emettitore:

$$\frac{1}{r_e} = \frac{\partial I_E}{\partial V_{BE}} = \frac{I_{EQ}}{V_T},$$

dove I_{EQ} è la corrente di emettitore nel punto di lavoro. Inoltre, poiché $I_{EQ} \simeq I_{CQ}$,

$$r_e = \frac{V_T}{I_{EQ}} \simeq \frac{V_T}{I_{CQ}}.$$

Uguagliamo la v_{be} nel circuito appena esaminato a quella nel circuito equivalente a parametri ibridi, semplificato trascurando h_{re} e h_{oe} :



Per il circuito con r_e abbiamo:

$$v_{be} = [r_{bb'} + r_e(h_{fe} + 1)]i_b,$$

mentre per il circuito a parametri ibridi otteniamo:

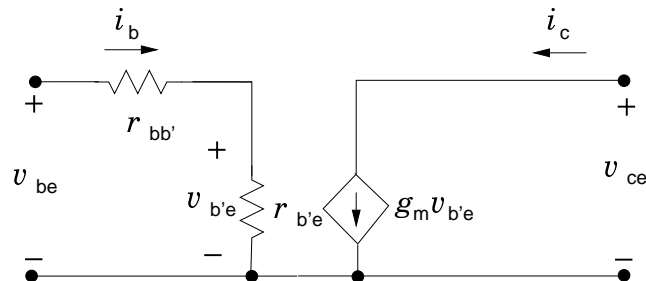
$$v_{be} = (r_{bb'} + r_{b'e})i_b.$$

Dal confronto tra queste due equazioni si ricava che

$$r_{b'e} = r_e(h_{fe} + 1) \simeq \frac{V_T}{I_{CQ}}(h_{fe} + 1) \simeq \frac{V_T}{I_{CQ}}h_{fe}.$$

L'ultima equazione ci permette di calcolare $r_{b'e}$ a partire dal valore di riposo della corrente di collettore e da quello di h_{fe} . Per trovare h_{ie} sarà sufficiente sommare $r_{b'e}$ e $r_{bb'}$, supponendo quest'ultima non dipendente dal punto di riposo, dato che è legata soltanto a parametri costruttivi del transistor.

Talvolta il generatore di corrente comandato in uscita viene rappresentato, invece che come un generatore di corrente comandato in corrente, come un generatore di corrente comandato in tensione, secondo lo schema sotto rappresentato.



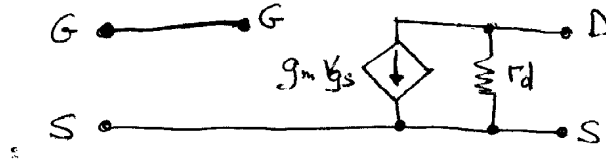
La tensione di controllo è $v_{b'e}$, corrispondente alla differenza di potenziale presente ai capi della $r_{b'e}$. È semplice determinare la relazione tra g_m e i parametri precedentemente considerati, imponendo l'uguaglianza delle correnti di collettore nei due circuiti:

$$h_{fe}i_b = g_m v_{b'e}.$$

Dato che $v_{b'e} = r_{b'e}i_b$, otteniamo che $g_m = h_{fe}/r_{b'e}$. Il parametro g_m ha le dimensioni dell'inverso di una resistenza, quindi Ω^{-1} o S (siemens). Spesso si utilizza come unità di misura per g_m l'ampere su volt (A/V) o il mA/V (dimensionalmente equivalenti all'inverso di una resistenza).

7.2 Modello linearizzato per transistori a effetto di campo

Si utilizza lo stesso modello linearizzato, molto semplice, per tutti i transistori a effetto di campo (sia MOSFET sia JFET). Esso è costituito da un generatore di corrente comandato in tensione, con in parallelo una resistenza r_d , che rappresenta l'effetto dovuto al fenomeno di accorciamento del canale al crescere di V_{DS} .



Il valore della transconduttanza g_m (dell'ordine di qualche mA/V per la maggior parte dei transistori a effetto di campo) si può ricavare dalla transcaratteristica fornita in forma grafica determinando la pendenza della tangente alla stessa nel punto di lavoro, che corrisponde proprio a g_m . se la caratteristica è invece fornita in maniera analitica, g_m si può ottenere derivando I_D rispetto a V_{GS} . Per un JFET otteniamo

$$g_m = \frac{-2I_{DSS}}{V_P} \left(1 - \frac{V_{GS}}{V_P}\right).$$

Questa formula è valida per JFET sia a canale p sia a canale n e dà luogo a un valore di g_m sempre positivo. È importante sottolineare che se la caratteristica è fornita in modo grafico è opportuno utilizzare il metodo della tangente per ricavare g_m , invece di determinare V_P e I_{DSS} per poi procedere con l'espressione analitica. Quest'ultimo metodo, assai sconsigliabile, può portare a risultati significativamente diversi da quello grafico perché la transcaratteristica non è esattamente una parabola (l'espressione analitica è soltanto un'approssimazione).

Il valore della resistenza differenziale r_d si ricava andando a valutare la pendenza della tangente alle caratteristiche di uscita in corrispondenza del punto di lavoro: r_d corrisponde all'inverso del coefficiente angolare di tale tangente.

Per quanto riguarda i transistori MOS, se la transcaratteristica è fornita in modo grafico si procede in modo analogo a quanto già visto per i JFET. Se, invece, la transcaratteristica è data in forma analitica, la si deriva rispetto a V_{GS} , ottenendo

$$g_m = 2k \left(\frac{W}{L}\right) (V_{GS} - V_T).$$

La r_d dei transistori MOS si ricava come già visto per i JFET se le caratteristiche di uscita sono disponibili in forma grafica. Se sono invece disponibili nella forma dell'espressione fenomenologica che include la tensione di Early, si ottiene come inverso della derivata di tale espressione rispetto a V_{DS} :

$$r_d = \left(\frac{dI_D}{dV_{DS}} \right)^{-1} = \frac{1 + \lambda V_{DS}}{\lambda I_D}.$$

Se $\lambda V_{DS} \ll 1$, abbiamo $r_d \simeq 1/(\lambda I_D)$.