

1

I TRANSISTORI

1.1 *Introduzione*

1.2 *Il transistor MOSFET*

- 1.2.1 Condizione di inversione e tensione di soglia
- 1.2.2 Resistore variabile
- 1.2.3 Pinch-off del canale ed effetto transistor
- 1.2.4 Modulazione della corrente con la tensione di Gate
- 1.2.5 Le curve caratteristiche dei MOSFET ad arricchimento
- 1.2.6 I simboli circuitali
- 1.2.7 Tensione di Early del MOSFET
- 1.2.8 Il MOSFET : un dispositivo in continua evoluzione

1.3 *Il transistor bipolare a giunzione - BJT*

- 1.3.1 Il principio di funzionamento
- 1.3.2 La corrente di Collettore
- 1.3.3 La corrente di Base
- 1.3.4 I simboli circuitali ed il funzionamento in zona attiva
- 1.3.5 Le curve caratteristiche
- 1.3.6 La transconduttanza
- 1.3.7 Effetto Early e resistenza di Collettore

1.1 INTRODUZIONE

In poco meno di un decennio, verso la fine degli anni Quaranta, la ricerca sui dispositivi a semiconduttore fece un grande balzo in avanti, passando dai semplici diodi rettificatori con ossidi di metalli, utilizzati ad esempio nei sistemi radar messi a punto negli anni precedenti alla seconda guerra mondiale, all'invenzione dei transistori a semiconduttore. Benché dietro allo sforzo scientifico che ha portato alla invenzione dei nuovi dispositivi ci fosse una motivazione applicativa (l'idea di voler sostituire gli interruttori meccanici delle centraline telefoniche con interruttori comandati più piccoli e più affidabili basati su principi *elettronici*), gran parte delle idee innovative furono il frutto dell'entusiasmo e delle fini capacità speculative di un gruppo di ricercatori della società dei telefoni americana: la *Bell Telephone*.

La prima conferma sperimentale della possibilità di ottenere un'amplificazione del segnale in un dispositivo a semiconduttore fu ottenuta il 16 dicembre 1947, quando John Bardeen e Walter H. Brattain fecero funzionare il primo transistor: il transistor con contatti a punte metalliche.

Il dispositivo era formato da un cristallo di germanio drogato n (Base) e da due contatti d'oro molto vicini tra di loro (Emettitore e Collettore). Collegando al Collettore un carico, i due ricercatori dimostrarono che variando la corrente al terminale di Base si riusciva ad ottenere una proporzionale variazione del potenziale del Collettore. Poiché il rapporto tra la variazione di potenziale del Collettore e la variazione della corrente di Base ha le dimensioni di una resistenza, gli autori descrissero il dispositivo come "transfer resistor" e coniarono il termine *transistor*. Il 23 dicembre 1947 i risultati degli esperimenti svolti furono comunicati alla direzione dei laboratori. Questa è la data universalmente considerata come data di nascita del transistor.



Ma fu un altro ricercatore dei laboratori Bell, il fisico William Shockley, a legare il proprio nome indissolubilmente ai dispositivi a semiconduttore. Si devono a lui infatti i maggiori meriti nello studio e comprensione dei singoli aspetti fisici del funzionamento del transistor:

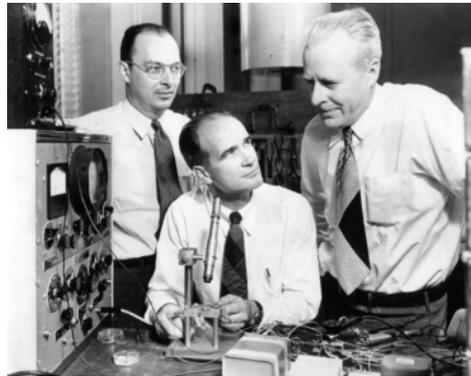
- (i) - l'iniezione dei portatori *minoritari* nelle giunzioni pn ;
- (ii) - il trasporto per diffusione che essi subiscono nella zona neutra, idea influenzata dai suoi studi con E. Fermi nel 1940 sulla diffusione dei neutroni nella grafite;
- (iii) - la rappresentazione a bande dei semiconduttori e l'introduzione del quasi-livello di Fermi (*imref*);
- (iv) - l'estendersi delle regioni svuotate nelle giunzioni polarizzate inversamente.

I risultati del suo studio ed una messe di idee innovative sono riassunti in cinque pagine del suo quaderno di laboratorio del 23 gennaio 1948. Nello scritto sono puntualizzati in forma chiara tutti gli aspetti del funzionamento del **transistore bipolare a giunzione**. Il testo costituì la base per la stesura del brevetto, presentato il 26 giugno 1948 e accettato il 25 settembre 1951 (U.S. Patent 2.569.347). A causa delle difficoltà tecnologiche che allora esistevano nelle tecniche di crescita e di drogaggio dei cristalli semiconduttori, si dovette attendere più di un anno prima che un dispositivo con tutte e sole giunzioni *pn* dimostrasse sperimentalmente la correttezza dei meccanismi di funzionamento ipotizzati.

Anche l'invenzione del **transistore unipolare ad effetto di campo** si deve a W. Shockley. L'idea di usare un campo elettrico esterno perpendicolare alla superficie del semiconduttore per controllare la densità dei portatori nel semiconduttore in prossimità della superficie fu infatti oggetto dei suoi studi fin dal 1943. Tuttavia la difficoltà di realizzare dispositivi in cui l'idea di principio non fosse mascherata da fenomeni fisici parassiti, non permise reali approfondimenti fino al 1952, quando pubblicò la proposta di un *transistore ad effetto di campo a giunzione*, con struttura concettualmente simile a quella di un moderno **JFET**, corredandola di uno studio approfondito del suo funzionamento. A questo scritto fece seguito, pochi mesi dopo, la conferma sperimentale delle proprietà descritte.

I **transistori MOSFET** dovettero invece attendere l'inizio degli anni Sessanta prima di essere realizzati, a causa della notevole dipendenza del loro funzionamento dalla qualità del cristallo usato e dalle caratteristiche della interfaccia ossido-semiconduttore. Alla loro realizzazione hanno contribuito gli sforzi di diversi laboratori di ricerca e di numerosi scienziati di differenti Università.

W. Shockley, W. H. Brattain e J. Bardeen (nella foto accanto) vennero insigniti, per le loro invenzioni, del Premio Nobel per la Fisica nel 1956.



La struttura dei MOSFET si mantenne simile a quella proposta dagli inventori e descritta nel paragrafo §1.2, per parecchi decenni fino a quando nuove esigenze emersero che stimolarono nuove soluzioni, come accenneremo nel paragrafo §1.2.8.

A suggello dell'importanza del transistor nella vita contemporanea si rifletta sul fatto che entro il 2025 ci saranno più transistor al silicio sul pianeta Terra che cellule umane.

1.2 IL TRANSISTORE MOSFET

Il funzionamento di un transistor MOSFET (*Metal Oxide Semiconductor Field Effect Transistor*) può essere illustrato facendo riferimento alla schematica struttura di un dispositivo a canale n , riportata nella Fig. 1.1. Il dispositivo è realizzato sulla superficie di un semiconduttore drogato p (substrato) ed è caratterizzato da due contatti n (Source -S- e Drain -D-) tra i quali è posto sopra al cristallo uno strato di dielettrico isolante (normalmente ossido di silicio, SiO_2) ricoperto da un contatto metallico detto di Gate -G-. Per semplicità supponiamo che nell'ossido non vi siano cariche elettriche e che il substrato sia tenuto allo stesso potenziale del contatto di Source, entrambi per esempio a massa.

I due contatti n di Source e Drain risultano isolati elettricamente tra loro. Infatti anche se venisse applicato un potenziale positivo all'elettrodo D ($V_S = V_{\text{SUB}} = V_G = 0$ e $V_D > 0$) la giunzione np tra Drain e substrato si polarizzerebbe inversamente e tra i due contatti n fluirebbe solo la debolissima corrente inversa dovuta ai pochi portatori minoritari presenti nelle due zone.

1.2.1 Condizione di inversione e tensione di soglia

Il funzionamento del dispositivo si fonda sulla possibilità di **variare**, agendo sul potenziale del Gate, **la concentrazione sotto l'ossido degli elettroni liberi** in modo da poter attivare un percorso continuo di elettroni tra i due contatti n . Se infatti si applica al Gate un potenziale positivo, si crea un campo elettrico perpendicolare tra Gate e substrato (da cui la classificazione del dispositivo come FET) che allontana le lacune verso il substrato e richiama gli elettroni minoritari verso l'interfaccia tra ossido e substrato (Fig. 1.2). L'arretramento delle lacune determina una zona svuotata di portatori maggioritari nella regione sotto l'ossido che si estende tanto più dalla superficie verso l'interno quanto più è alto il potenziale applicato al Gate. Se fosse solo questo l'effetto, l'isolamento tra S e D sarebbe addirittura accentuato. Ma contemporaneamente si ha l'aumento della concentrazione degli elettroni in un fine

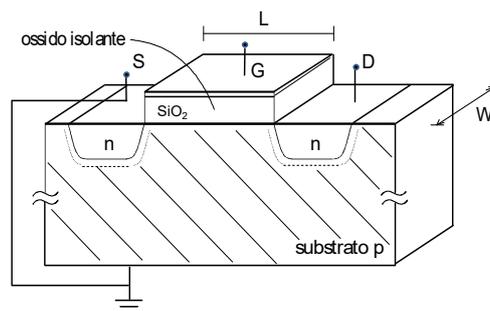


Fig. 1.1 Struttura schematica del MOSFET prototipo a canale n.

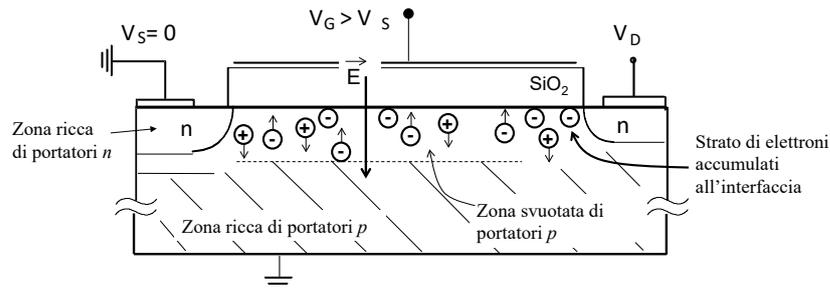


Fig. 1.2 Schematizzazione della formazione di un canale conduttivo di elettroni all'interfaccia ossido-semiconduttore in un nMOSFET quando al Gate viene applicata una tensione positiva rispetto al Source, che allontana le lacune e richiama elettroni.

strato sotto l'ossido che crea un collegamento continuo di elettroni (canale) tra Source e Drain. Questi due terminali, ricchi di elettroni perché drogati, cessano di essere isolati elettricamente e si collegano al canale. Se si applicasse un piccolo potenziale positivo V_D al Drain, tra i due contatti passerebbe corrente. Aumentando ulteriormente la polarizzazione positiva del Gate, la concentrazione di elettroni aumenta nel canale. Si intuisce quindi come il potenziale di Gate possa agire come variabile di controllo della conduttanza del dispositivo tra i due terminali di Source e Drain.

Il canale conduttivo tra Source e Drain è per convenzione ritenuto formato quando la concentrazione di elettroni sotto l'ossido giunge ad eguagliare la concentrazione che le lacune avevano lì prima dell'applicazione della tensione (cioè la concentrazione delle lacune maggioritarie nel substrato neutro). Il valore della tensione V_{GS} per cui si raggiunge questa condizione è detto **tensione di soglia**, V_T (**Threshold Voltage**) e quando si raggiunge tale valore, $V_{GS} = V_T$, si dice che si è raggiunta la **condizione di inversione**: i portatori che garantiscono la conducibilità dello strato sotto l'ossido sono di polarità opposta (nel nostro caso elettroni) ai portatori maggioritari lì presenti all'atto della fabbricazione (lacune).

La **TENSIONE di SOGLIA**, V_T , di un MOSFET è normalmente dell'ordine del Volt (0.3÷1V). Essa dipende dal drogaggio del substrato, dallo spessore dell'ossido e dalle cariche fisse eventualmente in esso presenti. Può variare fino ad un 20% da un lotto di fabbricazione ad un altro, mentre tra i MOSFET di uno stesso chip la variazione è in genere molto più piccola. Il valore della V_T può essere modificato in fase di fabbricazione mediante processi di impiantazione di droganti nel semiconduttore proprio e solo sotto l'ossido per cambiare localmente il drogaggio.

Si noti che tra il canale conduttivo di elettroni sotto l'ossido ed il substrato neutro si ha una zona svuotata, assicurando completo isolamento elettrico tra i portatori (elettroni) che scorrono nel canale e quelli maggioritari (lacune) del substrato.

1.2.2 Resistore variabile

Lo strato di cariche mobili nel canale e l'elettrodo di Gate formano le armature di un condensatore, il cui dielettrico è l'ossido (Fig.1.3a). Un'analisi dettagliata del sistema metallo-ossido-semiconduttore (MOS) mostrerebbe come, all'aumentare della tensione di Gate oltre il valore di soglia V_T , la tensione in eccesso ($V_{GS}-V_T$), spesso chiamata **tensione di overdrive** $V_{ov}=V_{GS}-V_T$, vada a cadere quasi interamente ai capi dell'ossido. Pertanto la concentrazione di carica nel canale per unità di area cresce linearmente con la tensione secondo la nota relazione di un condensatore:

$$Q'_n = C'_{ox} (V_{GS} - V_T) \quad (1.1)$$

dove $C'_{ox} = \epsilon_{ox} / t_{ox}$ è la capacità per unità di superficie del gate e t_{ox} lo spessore dell'ossido. Corrispondentemente la resistenza del canale tra Source e Drain, costituita dal film di elettroni spesso S_n sarebbe:

$$R = \rho \cdot \frac{L}{S_n \cdot W} = \frac{1}{\mu_n \cdot q \cdot n} \cdot \frac{L}{S_n \cdot W}$$

dove L e W sono rispettivamente la lunghezza (lungo la direzione di moto dei portatori, tra S e D) e la larghezza del canale (Fig.1.1), μ_n è la mobilità degli elettroni

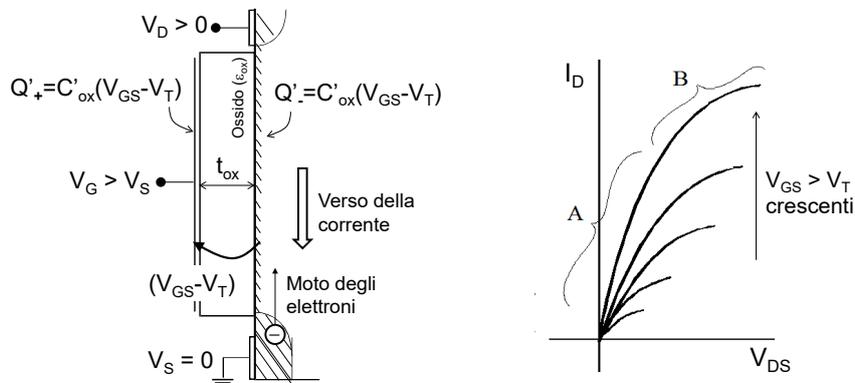


Fig. 1.3 a) Comportamento del MOS, analogo ad un condensatore a facce piane parallele e b) curve caratteristiche di un MOSFET a canale n nella zona ohmica. La corrente scorre dal potenziale maggiore a massa, dall'alto verso il basso per comodità di rappresentazione.

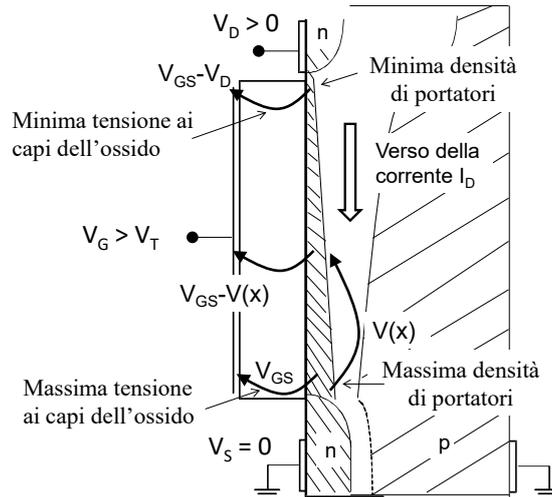


Fig. 1.4 *Cadute di tensione lungo il canale di un nMOSFET nella zona ohmica e conseguente distribuzione della concentrazione dei portatori visualizzata con uno spessore variabile del tratteggio.*

ed n la concentrazione di elettroni in un volume unitario. Poiché $Q_n' = q \cdot n \cdot S_n$ si ottiene

$$R = \frac{1}{\mu_n \cdot Q_n'} \cdot \frac{L}{W} = \frac{1}{\mu_n} \cdot \frac{L}{W} \cdot \frac{1}{C_{ox}} \cdot \frac{1}{V_{GS} - V_T} \quad (1.2)$$

Al variare della tensione di Gate, in accordo con la (1.1), la resistenza di canale può essere aumentata o diminuita. Applicando una tensione positiva al Drain, nel canale fluisce una corrente tanto maggiore quanto più è alta la concentrazione superficiale di elettroni.

La Fig. 1.3b mostra ciò: se $V_{GS} > V_T$ e $V_{DS} > 0$, la corrente nel dispositivo varia secondo la famiglia di curve caratteristiche mostrate. Per piccole tensioni V_{DS} (tratto A delle caratteristiche) la crescita della corrente è praticamente lineare perché la conducibilità del canale è praticamente costante lungo tutta la sua estensione tra S e D. Questa regione di funzionamento è detta **zona ohmica del MOSFET** proprio perché il dispositivo si comporta come un semplice resistore.

Quando però si aumenta ulteriormente il potenziale del Drain, l'ulteriore aumento della corrente circolante segue una legge meno che lineare (tratto B della caratteristica). Per capirlo basta concentrarsi sulla tensione ai capi dell'ossido nei vari punti del canale (Fig.1.4). Infatti, la caduta di potenziale lungo il canale determinata dal passaggio della corrente fa diminuire la tensione ai capi dell'ossido man mano che ci si sposta verso il Drain: vicino al Source la tensione ai capi

dell'ossido sarà massima e pari a V_{GS} ; vicino al Drain la tensione sarà minima e pari a $(V_{GS}-V_T)$. Quindi, in accordo con la (1.1), la densità di carica nel canale diminuisce procedendo dal Source verso il Drain e raggiunge il suo minimo all'estremità del canale adiacente al Drain dove la tensione ai capi dell'ossido è minima. Questa diminuzione fa sì che, aumentando la tensione V_{DS} , la resistenza media del canale aumenti e quindi che la corrente tenda ad aumentare ma meno che proporzionalmente con V_{DS} , come appunto si vede nel tratto B delle curve della Fig.1.3b

1.2.3 Pinch-off del canale ed effetto transistor

Aumentando ulteriormente V_{DS} si arriva alla situazione in cui la caduta di potenziale ai capi dell'ossido tra Gate e Drain diventa pari o inferiore a V_T e quindi non sia più sufficiente per avere canale conduttivo all'estremo verso il Drain. Questa situazione, visualizzata nella Fig.1.5, è detta condizione di *strozzamento (pinch-off) del canale* ed indichiamo con P il punto lungo il canale dove questo si realizza. La tensione tra Drain e Source a cui si instaura il *pinch-off* è detta tensione di saturazione, $V_{Dsat}=V_{GS}-V_T$.

Si noti che il *pinch-off* del canale conduttivo non impedisce il passaggio della corrente. Infatti :

- la tensione del punto P vale $V_P=(V_{GS}-V_T)$ e quindi la tensione ai capi del canale conduttivo vale esattamente $(V_{GS}-V_T)$;

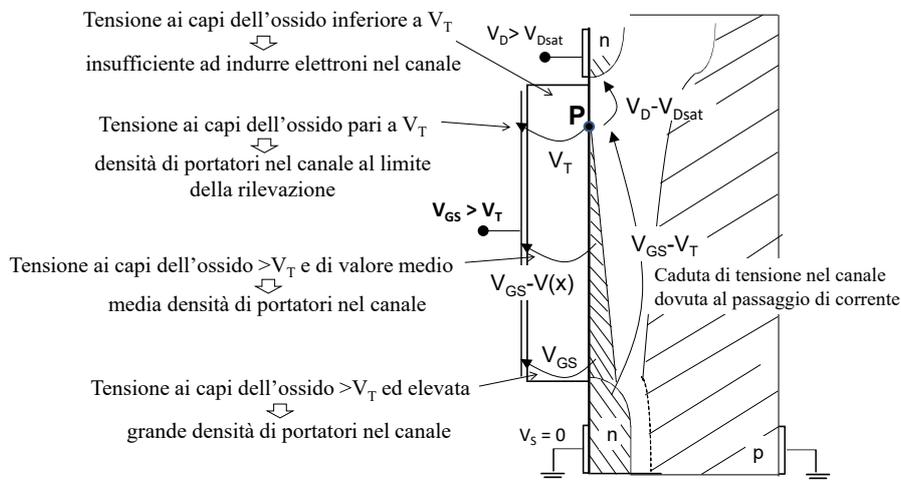
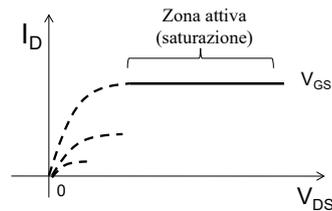


Fig. 1.5 Distribuzione delle cadute di tensione ai capi dell'ossido e corrispondente distribuzione di concentrazione di elettroni nel canale.

- il rapporto tra la tensione del punto P, ($V_{GS}-V_T$), e la resistenza del canale determina una corrente I_D che fluisce tra Drain e Source;
- gli elettroni che dal Source giungono nel punto P continuano ad essere accelerati verso il Drain da una differenza di potenziale ancora favorevole alla continuazione del loro moto, in una zona di cristallo dove non ci sono altri portatori.

Per $V_{DS} \geq V_{DSat}$ la situazione non cambia e si avrà sempre un punto P nel canale molto vicino al Drain a tensione $V_P=(V_{GS}-V_T)$ (è la tensione minima ai capi dell'ossido per invertire il canale), mentre la tensione di Drain in eccesso ($V_{DS}-V_{DSat}$) va a cadere tra il punto P ed il morsetto di Drain, determinando eventualmente un allargamento locale della regione svuotata ed una accelerazione maggiore dei portatori.

La corrente che fluisce nel transistor **rimane** quindi pressoché **costante** nonostante l'aumento di V_{DS} (si dice che il transistor opera in **saturazione**) ed è sempre data dal rapporto tra $V(P)=(V_{GS}-V_T)$ e la resistenza di canale, entrambi pressoché costanti, come visualizzato nella curva caratteristica accanto (e in Fig.1.6).



1.2.4 Modulazione della corrente con la tensione di Gate

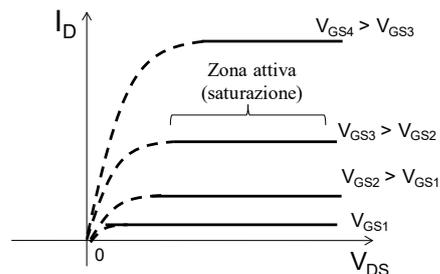
Consideriamo ora cosa accade **umentando la tensione V_{GS}** ma lasciando V_{DS} ad un valore alto, superiore a V_{DSat} :

1) **aumenta la carica nel canale facendone diminuire la resistenza.**

Infatti, l'aumento di tensione al Gate aumenta il campo elettrico perpendicolare ed, in base alla (1.1), viene indotta più carica nel canale.

2) **aumenta la tensione ai capi del canale tra il punto P ed il Source;**

Infatti, poiché con il dispositivo in saturazione la differenza di potenziale tra l'elettrodo di Gate ed il punto P rimane sempre pari a V_T , ogni variazione del potenziale sul Gate si trasmette rigidamente come una identica variazione del potenziale del punto P. In definitiva, attraverso un morsetto esterno (il Gate), si riesce a controllare il potenziale di un punto interno del dispositivo (il punto P) e conseguentemente il valore della tensione ai capi della resistenza del canale.



La combinazione dei due effetti determina un aumento di I_D proporzionale al quadrato della tensione ($V_{GS}-V_T$), secondo la relazione:

$$I_D = \frac{1}{2} \mu_n \cdot C'_{ox} \frac{W}{L} (V_{GS} - V_T)^2 = k(V_{GS} - V_T)^2 \quad (1.3)$$

dove il parametro di proporzionalità k ha le dimensioni $[A/V^2]$. La Fig. 1.6(b) mostra in forma grafica l'andamento della (1.3). Essa è la relazione da usare per valutare la corrente in un MOSFET in saturazione al variare del comando V_G .

UN ESERCIZIO DI TECNOLOGIA

Si pensi di usare per la realizzazione di un circuito integrato una tecnologia CMOS che permetta la definizione di geometrie con dimensioni minime di 90nm e spessori minimi per l'ossido di gate di 10nm.

Calcolare il valore del coefficiente k dei transistori nMOSFET e pMOSFET ottenibili con questa tecnologia ed aventi le dimensioni minime possibili.

Con riferimento alla Fig. 1.1, per il calcolo del coefficiente k dei MOSFET basta ricordare che:

- la mobilità dei portatori all'interfaccia tra silicio ed ossido di silicio è pari a $\mu_n \cong 1000 \text{cm}^2/\text{Vs}$ e $\mu_p \cong 350 \text{cm}^2/\text{Vs}$ rispettivamente per elettroni e lacune;
- la costante dielettrica dell'ossido di silicio (SiO_2) è pari a $\epsilon_{ox} = \epsilon_o \epsilon_r$ dove $\epsilon_o = 8,8 \times 10^{-14} \text{F/cm}$ è la costante dielettrica del vuoto ed $\epsilon_r = 3.9$ è la costante dielettrica relativa dell'ossido di silicio. Ne risulta $\epsilon_{ox} = 34 \times 10^{-14} \text{F/cm}$.
- la capacità per unità di area, $C'_{ox} = \epsilon_{ox} / t_{ox}$, dove t_{ox} è lo spessore dell'ossido di silicio utilizzato come dielettrico di gate, è quindi pari a 340nF/cm^2 . Nel fare questo conto, ed anche i successivi sulle dimensioni del dispositivo, si faccia attenzione a riconduurre le dimensioni geometriche ai cm.
- le dimensioni minime possibili per realizzare gli elementi del dispositivo (dimensione di un elettrodo, distanza tra due elettrodi, ecc.) con la tecnologia indicate sono 90nm, e quindi $W=90\text{nm}$ e $L=90\text{nm}$.

Utilizzando tali dati si ottiene $k_n = 170 \mu\text{A}/\text{V}^2$ e $k_p = 60 \mu\text{A}/\text{V}^2$.

1.2.5 Le curve caratteristiche dei MOSFET ad arricchimento

Nella Fig.1.6(a) sono riportate le **curve caratteristiche** di un ideale MOSFET a canale n . Si noti la zona di saturazione (chiamata anche "regione attiva" o "regione di pinch-off") in cui la corrente portata dal transistor rimane costante pur variando la tensione tra Drain e Source. Essa è la regione di funzionamento più interessante ed è quella in cui il MOSFET viene polarizzato quando impiegato come amplificatore.

Si noti anche che per $V_{GS} \leq 1V$ in questo esempio il transistor è interdetto (non scorre corrente), valore che quindi costituisce la sua tensione di soglia, $V_T \cong 1V$. Poiché per $V_{GS} = 0$ il transistor non conduce, si dice che esso è *normalmente spento* (*normally off*). Solo applicando una tensione di Gate maggiore della tensione di soglia si arricchisce il canale di portatori mobili, la conducibilità del canale aumenta e nel dispositivo fluiscono correnti sempre più intense. Questo comportamento classifica il dispositivo come un **MOSFET a canale n ad arricchimento** (*enhancement*).

Nella Fig.1.6b è riportato il grafico della relazione (1.3), corrispondente ad un ramo di parabola con vertice in $V_{GS} = V_T$, per un dato valore di V_{DS} . Essa è detta **curva transcaratteristica** perché lega la tensione "all'ingresso" del MOSFET, V_{GS} , alla corrente I_D che scorre "in uscita" dal dispositivo (da cui "trans-")

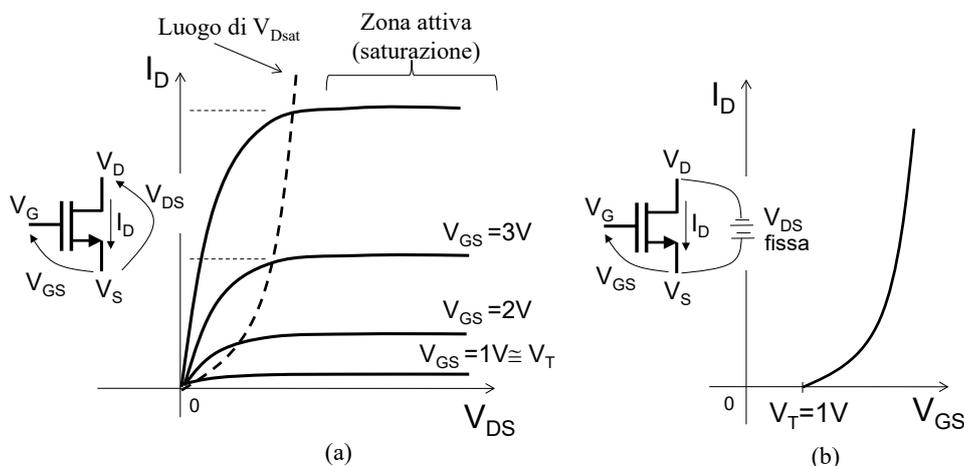


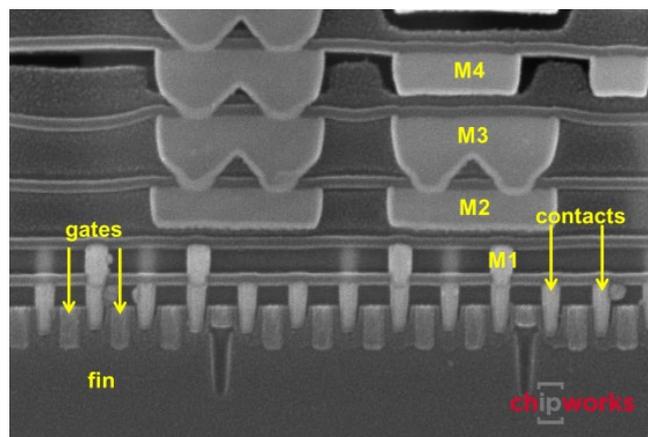
Fig. 1.6 MOSFET ideale ad arricchimento a canale n avente $V_T = 1V$: a) curve caratteristiche, il cui parametro è V_{GS} ; b) curva transcaratteristica (Eq.(1.3)) in zona di saturazione per una fissata $V_{DS} > V_{DSsat}$

I MOSFET a canale p , in cui i portatori mobili nel canale sono lacune, sono realizzati a partire da un substrato n su cui i contatti di Source e di Drain sono di tipo p . La tensione di soglia è negativa poichè il Gate deve essere polarizzato negativamente rispetto al Source per attrarre lacune sotto l'ossido a formare il canale conduttivo. Le curve caratteristiche di un **MOSFET a canale p ad arricchimento** sono riportate nell' esercizio E 1.1.

Si noti che il funzionamento del MOSFET, come di tutti gli altri FET esistenti realizzati anche con altri materiali, è legato al trasporto dei portatori *maggioritari* in un canale conduttivo. Inoltre si noti che **la corrente stazionaria assorbita dal Gate di un MOS è nulla.**

Dove è il transistor ?

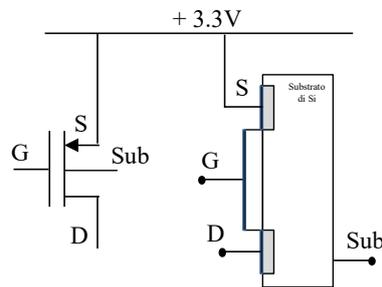
Qualche anno fa, ad un congresso internazionale, venne proiettata un'immagine simile alla seguente e venne chiesto al pubblico "Where is the transistor?"



La foto al microscopio elettronico rappresenta la sezione di un chip di un microcontrollore. Essa sottolinea l'importanza sempre maggiore che oggi hanno le connessioni metalliche (rappresentate in figura dai livelli di metallo M1, M4..) necessarie per collegare tra loro i moltissimi singoli transistori presenti nel chip. Potete vedere alcuni transistori in basso sotto i gates le cui piccole dimensioni sono trascurabili rispetto allo spazio occupato dalle interconnessioni !

P – MOSFET

Riflettiamo su come funziona un transistor pMOSFET rispondendo alle seguenti domande:



Che tipo di substrato scegliereste per realizzarlo, drogato n (n-type) o drogato p (p-type) ?

Con quale drogaggio devono essere fatti i contatti di Source e Drain, n-type o p-type ?

A che potenziale va collegato il substrato ? Perché ?

Cosa si intende quando si dice che il transistor ha una $V_T=0.5V$?

Quale valore di tensione tra questi (3.8V; 3.3V; 2.8V) devo applicare al Gate dell'esempio sopra per iniziare ad indurre carica nel canale ?

In che direzione devo muovere la tensione al Gate per aumentare la corrente nel dispositivo, fissata una piccola differenza di potenziale tra S e D, ad esempio di 50mV ?

A quali valori di tensione posso porre il Drain per assicurare il funzionamento del MOSFET in saturazione ?

A quale tensione sta il punto di pinch-off del canale quando $V_G=2.5V$ (sempre con MOSFET con $V_T=0.5V$) ?

A quale tensione sta il punto di pinch-off del canale quando $V_G=2V$? E' giusto che rispetto a prima la corrente I_D aumenti ? Perché ?

Perché la corrente nel MOSFET aumenta quadraticamente e non linearmente ?

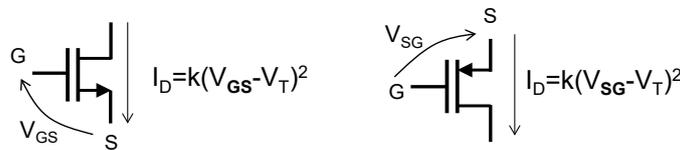
1.2.6 I simboli circuitali

La figura seguente riporta alcuni dei simboli circuitali più usati dei MOSFET ad arricchimento a canale n ed a canale p .



Il terminale di substrato è spesso cortocircuitato con il Source (in modo da assicurare che le giunzioni di Source e di Drain non vadano mai in diretta con il substrato) e può non essere disponibile come contatto esterno, nel qual caso esso può non apparire nel simbolo del dispositivo.

Per convenzione mutuata dalla teoria delle reti elettriche, molti testi assumono come positiva la corrente entrante nel Drain. Attenendosi ad essa, nella (1.3) k ha segno positivo per gli n MOSFET mentre ha segno negativo per i p MOSFET. In verità è molto più comodo pensare al funzionamento reale del dispositivo, **indicando le tensioni e le correnti con il loro verso fisico** (V_{SG} in p MOSFET e V_{GS} in n MOSFET, I_D dal Source al Drain in p MOSFET e I_D dal Drain al Source in n MOSFET) **in modo da trattare sempre grandezze positive**.



Anche k e V_T dei p MOSFET o dei n MOSFET è più comodo considerarli sempre valori positivi legati alla fisica dei meccanismi di funzionamento dei MOSFET. Così facendo, il valore di k determinerà il valore effettivo della corrente I_D nel verso fisico in cui effettivamente scorre nel transistor. Per quanto riguarda V_T basterà prendere il verso fisico giusto delle tensioni di comando per avere sempre il corretto valore di tensione di overdrive, $V_{od}=(V_{GS}-V_T)$ in n MOSFET e $V_{od}=(V_{SG}-V_T)$ in p MOSFET, utilizzando sempre V_T positivo.

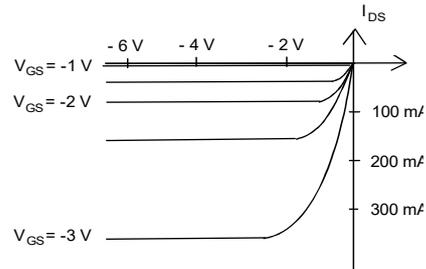
La (1.3) diventerebbe per il p MOSFET:

$$I_D = \frac{1}{2} \mu_p \cdot C'_{ox} \frac{W}{L} (V_{SG} - V_T)^2 = k(V_{SG} - V_T)^2 \quad (1.4)$$

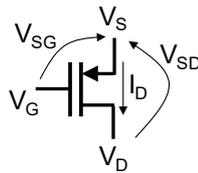
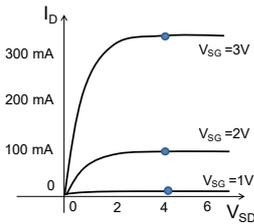
rispettando l'effettivo funzionamento fisico come indicato nella figura sopra.

Affinché i dispositivi operino in saturazione, e quindi valgano le (1.3) e la (1.4), bisogna accertarsi sempre che la tensione tra G e D determini il pinch-off del canale verso il Drain.

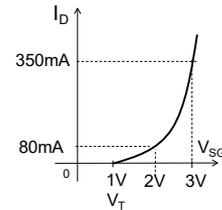
E 1.1 Partendo dalle curve caratteristiche del p-MOSFET ad arricchimento della figura accanto disponibili sul data-sheet del componente fornito dal fabbricante, tracciare la curva transcaratteristica e ricavare il valore di V_T .



Conviene analizzare il comportamento del transistor pensando alle tensioni ed alle correnti come positive nel verso in cui sono effettivamente presenti nel transistor, e quindi ridisegnare le curve caratteristiche nel seguente modo.



Mettendosi ad una definita V_{SD} , ad esempio $V_{SD}=4V$, si leggono i valori di corrente I_D alle varie V_{SG} . Questi costituiscono i valori con cui disegnare il grafico $I_D - V_{SG}$ (transcaratteristica). Ne risulta una V_T prossima a 1V.



Simmetria della struttura del MOSFET

La struttura del MOSFET ideale (Fig.1.1) è simmetrica, nel senso che il Source ed il Drain possono essere interscambiati. Questo aspetto viene a volte sottolineato dall'uso di un simbolo circuitale senza la freccia sul Source.



Sarà poi il collegamento nel circuito a definire quale contatto si comporti effettivamente da Source e quale da Drain. Non sempre facile da capire in un circuito complesso !

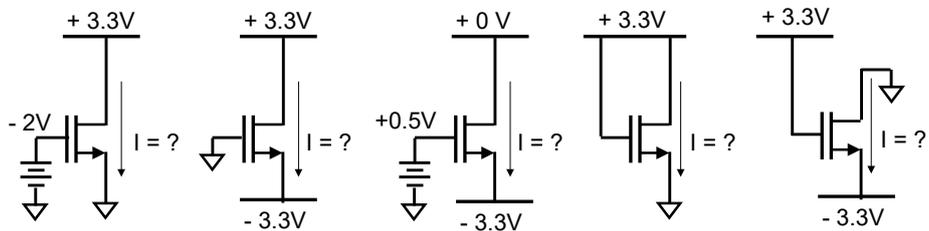
Un MOSFET può non essere simmetrico qualora siano differenti le distanze tra il gate ed i due contatti di S e di D o addirittura nei transistori a simmetria circolare, con ad esempio il S al centro attorniato da un anello di Drain.

FlashTest sul MOSFET : Ohmico, saturato, interdetto

Quali tensioni possono esserci ai tre morsetti di un MOSFET affinché funzioni correttamente ?

I seguenti quiz dovrebbero aiutarvi a prendere confidenza con questi aspetti fondamentali di uso di un MOSFET in un circuito elettronico.

ES.1 Si supponga che i MOSFET abbiano $V_T=0.8V$ e $k=300\mu A/V^2$:



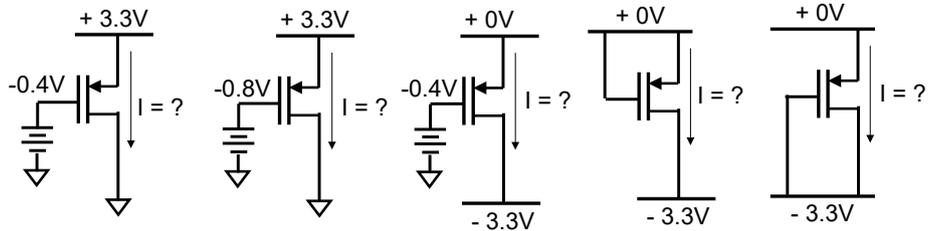
Il MOSFET è in saturazione (cioè nella corretta zona di funzionamento) ?

- si no si no si no si no si no

Se si, quanta corrente I scorre ?

$I = \dots\dots\dots$ $I = \dots\dots\dots$ $I = \dots\dots\dots$ $I = \dots\dots\dots$ $I = \dots\dots\dots$

ES.2 Si supponga che i MOSFET abbiano $V_T=0.7V$ e $k=100\mu A/V^2$:



Il MOSFET è in saturazione (cioè nella corretta zona di funzionamento) ?

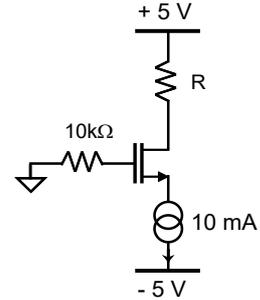
- si no si no si no si no si no

Se si, quanta corrente I scorre ?

$I = \dots\dots\dots$ $I = \dots\dots\dots$ $I = \dots\dots\dots$ $I = \dots\dots\dots$ $I = \dots\dots\dots$

E 1.2

- (a) Calcolare nel circuito accanto a quale tensione si pone il Source del MOSFET, avente $V_T=1.5V$ e $k=10mA/V^2$.
- (b) Indicare quale (o quali) tra i seguenti valori di resistenze collegate al Drain farebbe uscire il MOSFET dalla zona di saturazione: $R=80\Omega$, 500Ω , 640Ω , 720Ω
- (c) Calcolare di quanto varierebbe la corrente di Drain del MOSFET se k fosse $20mA/V^2$.



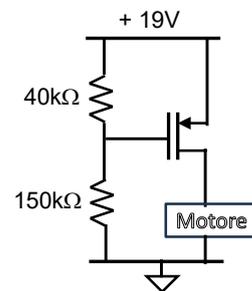
a- Poiché non scorre corrente nel Gate, non c'è caduta di tensione sulla resistenza da $10k\Omega$ e quindi $V_G=0V$. Il generatore di corrente forza una corrente di $10mA$ nel MOSFET e quindi il Source deve portarsi alla tensione di $-2.5V$ per lasciarla scorrere.

b- Affinché il MOSFET funzioni nella zona di saturazione, il canale deve essere in *pinch-off* dal lato del Drain. Pertanto la tensione al Drain deve essere $V_D > -1.5V$. L'unico valore di R che deve essere scartato è pertanto 720Ω che determinerebbe $V_D = -2.2V$ e creerebbe un canale conduttivo in tutto il transistor portandolo ad operare in zona ohmica. Gli altri valori di R invece assicurano la saturazione.

c- La corrente nel Drain non varia perché imposta dal generatore di corrente. Al variare di k varierà solamente la tensione a cui si porta il Source per lasciare passare i $10mA$ imposti dal generatore di corrente.

E 1.3

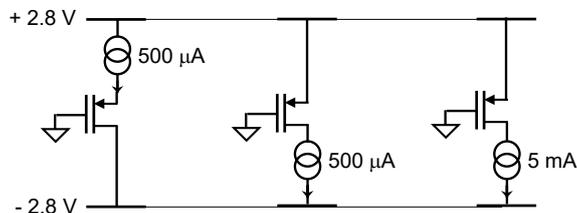
- a) Quale valore di $k = \frac{1}{2}\mu_p \cdot C'_{ox} \frac{W}{L}$ deve avere il MOSFET nel circuito accanto affinché la corrente inviata al motore elettrico sia di $1A$? Si supponga di usare una tecnologia che realizzi MOSFET con $V_T=2V$.
- b) Quale è la massima resistenza che può mostrare il motore per non fare uscire il MOSFET dalla saturazione?



- a) $K=250mA/V^2$.
- b) $R_{motore}=17\Omega$.

Si può FORZARE CORRENTE in un MOSFET ?

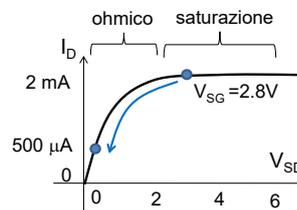
In cosa differiscono i seguenti circuiti, i cui MOSFET abbiano $V_T=0.8V$ e $k=500 \mu A$? Riflettete prima di leggere la risposta.



Ricordiamoci sempre che *in saturazione* la corrente I_D nel MOSFET è definita dalla sua tensione di comando V_{SG} secondo la relazione $I_D = k(V_{SG} - V_T)^2$

Nel circuito a sinistra, V_{SG} non è imposta. La corrente I , se non riuscisse a passare nel transistor, accumulerebbe carica positiva al Source facendone alzare il valore in tensione. Questo porta V_{SG} ad aumentare, attivando corrente I_D nel MOSFET. Questo fenomeno raggiunge il perfetto equilibrio quando al nodo di Source $I_D=I$, cioè quando V_S si sarà portato esattamente a $V_S=1.8V$. A posteriori si controlla che il MOSFET sia correttamente in saturazione, situazione assicurata qui dalla alimentazione a $-2.8V$. Concludiamo dicendo che è **possibile forzare corrente nel Source di un MOSFET**.

Nel circuito al centro, V_{SG} è imposta e vale $V_{SG}=2.8V$. Quindi, se in saturazione, la corrente sarà fissata a $I_D=2mA$. Ma così il bilancio di correnti al Drain non è soddisfatto! La corrente in più da MOSFET fa alzare il potenziale del Drain. Questo però fa uscire il MOSFET dalla saturazione e lo porta in ohmico. Il Drain sale in tensione fino a quando il MOSFET in ohmico produce esattamente la stessa corrente richiamata dal generatore, $500\mu A$.



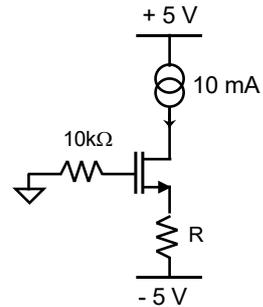
Nel circuito a destra la situazione è simile alla precedente ma ora il bilancio delle correnti al nodo di Drain fa scendere la sua tensione, andando a diminuire la tensione ai capi del generatore. In questo caso bisogna pensare che il generatore di corrente sia reale e che quando la tensione ai suoi capi diventa molto piccola, la corrente generata non riesca più ad essere quella nominale di $5mA$ ma diminuisca, in questo esempio fino proprio a $I=2mA$. Vedremo nei prossimi capitoli che i generatori di corrente fatti con transistori consentono proprio questi aggiustamenti.

Concludiamo dicendo che **NON è indolore forzare corrente nel Drain di un MOSFET**.

E 1.4 In questo circuito, il cui MOSFET abbia come in E1.2 $V_T=1.5V$ e $k=10mA/V^2$, discutere le seguenti situazioni:

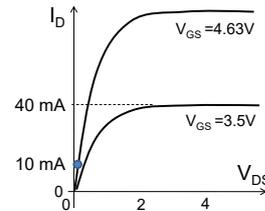
- Se $R=37\Omega$, sapresti dire se il MOSFET è in saturazione o no? Quali tensioni avrebbero il Gate, Source e Drain?
- Se $R=1200\Omega$, sapresti dire se il MOSFET è in saturazione o no? Quali tensioni avrebbero il Gate, Source e Drain?

Riflettere sulla situazione fisica nel dispositivo e commentare le criticità trovate.



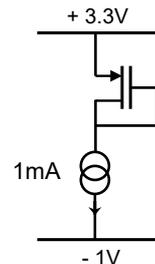
La situazione di funzionamento di questo circuito è completamente diversa da quella del circuito in E1.2.

- Con $R=37\Omega$ la corrente nel MOSFET, se fosse in saturazione, sarebbe di circa 40mA. Questo valore sarebbe ben maggiore dei 10mA provenienti dall'alto. Pertanto, il morsetto di Drain si abbasserà in potenziale fino a quando è soddisfatto il bilancio di corrente al nodo di Drain. Questo avverrà quando il MOSFET sarà entrato in zona ohmica fino a portare esattamente 10mA. Il fenomeno è mostrato nel grafico delle curve caratteristiche, in cui il MOSFET si assesta sulla curva con $V_{GS}=4.63V$ ed appunto $I_D=10mA$.



- Con $R=1200\Omega$, la situazione è diversa: la corrente nel MOSFET se fosse in saturazione sarebbe di 2.5mA, minore di 10mA, e quindi il bilancio al Drain tende ad innalzare il suo potenziale fino a +5V nel tentativo di fare funzionare male il generatore di corrente e fargli portare esattamente la corrente di 2.5mA definita dal MOSFET. Notate che il MOSFET è in saturazione.

E 1.5 Verificare che in questo circuito sia possibile forzare corrente nel Drain del MOSFET e calcolare a quale tensione si trova il Drain ($V_T=0.6V$, $k=4mA/V^2$).



$$V_D = V_G = +2.2V$$

Come scegliere la dimensione di un MOSFET

La corrente portata dal transistor in saturazione segue, come visto nella (1.3) o (1.4), la relazione :

$$I_D = \frac{1}{2} \mu_n \cdot C'_{ox} \frac{W}{L} (V_{GS} - V_T)^2$$

A seconda delle applicazioni, la corrente richiesta in un circuito può variare da pochi nA (nei circuiti elettronici del cellulare ad esempio e nella maggior parte dei circuiti di supporto al digitale dove si vuole consumare la minima potenza elettrica possibile) fino a parecchie centinaia di Ampère (nel comando del motore nelle automobili elettriche, ad esempio, dove viceversa è essenziale gestire enormi quantità di energia).

Come fare a soddisfare esigenze così diverse ?

Non certo usando sempre e solo uno stesso transistor e cambiando la tensione di comando tra Gate e Source. Essa è normalmente limitata da ragioni pratiche e tecnologiche a pochi Volt oltre cui il fine ossido di Gate verrebbe danneggiato da un eccessivo campo elettrico.

Quello che si fa è progettare MOSFET con l'opportuna **dimensione W**. Essa è l'effettiva grandezza che il progettista sceglie per definire l'ordine di grandezza della corrente portata dal transistor: circuiti a basso consumo e con piccole correnti di polarizzazione utilizzano MOSFET con una W molto piccola, di poche decine di nm; circuiti di grande potenza e grandissime correnti di polarizzazione utilizzano transistori di grandi dimensioni, con W che raggiungono i centimetri. Tutti operanti a tensioni V_{GS} di pochi Volts !

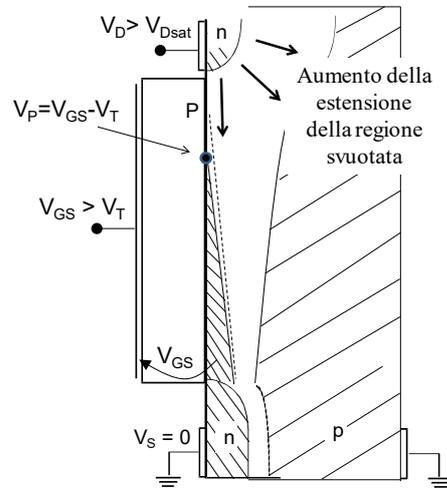
La **dimensione L**, cioè la lunghezza del canale tra S e D, viene tendenzialmente scelta piccola, spesso della minima dimensione consentita dalla tecnologia, così da avere il minore tempo di transito possibile, cioè la velocità di risposta la più elevata possibile, e le minori capacità possibili, sempre per avere una banda passante la più elevata possibile. Vedremo questi aspetti nei prossimi capitoli. La dimensione L viene scelta grande solo in casi particolari, ad esempio quando si vuole avere una elevata tensione di Early (vedi il prossimo paragrafo) cioè avere una grande resistenza di Drain.

1.2.7 Tensione di Early del MOSFET

La situazione ideale sintetizzata dalla (1.3) e rappresentata dalle curve caratteristiche della Fig.1.6 non trova preciso riscontro nelle caratteristiche ottenute sperimentalmente su un dispositivo reale. Normalmente infatti **la reale corrente erogata nella zona di saturazione non è costante ma tende ad aumentare leggermente all'aumentare della tensione V_{DS}** (Fig.1.7).

Questo effetto è dovuto al fatto che aumentando la tensione V_{DS} aumenta la tensione inversa della giunzione Drain-Substrato e conseguentemente aumenta l'estensione della zona svuotata al Drain. Essa "spinge" il punto P di *pinch-off* leggermente verso il Source (fenomeno chiamato di "modulazione della lunghezza di canale").

In conseguenza di ciò, pur tenendo fissi i potenziali V_S e V_G , un leggero accorciamento del canale conduttivo ne diminuisce la resistenza e, rimanendo invariata la caduta di tensione tra il Source ed il punto P di *pinch-off*, fa aumentare la corrente nel canale.



Questo effetto è indesiderato, proprio perché rende la corrente fornita al carico un po' dipendente anche dal carico stesso (attraverso la tensione ai suoi capi) e non solo dalla tensione di comando V_{GS} . Supponendo per semplicità che le nuove

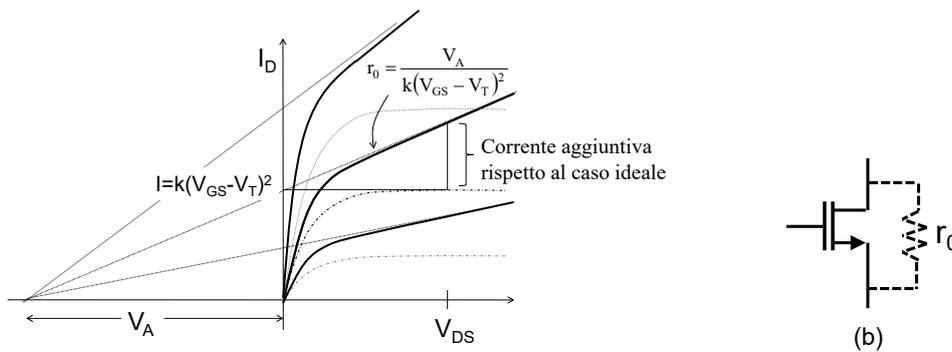


Fig. 1.7 Curve caratteristiche reali di un MOSFET e calcolo della sua resistenza equivalente tra Drain e Source; (b) corrispondente visualizzazione nel simbolo circuitale.

curve caratteristiche (Fig.1.7) siano comunque delle rette nella zona di saturazione, c'è un'evidenza pratica che le curve convergano tutte con buona approssimazione in un unico punto lungo l'asse V_{DS} negativo. Chiamiamo questa tensione "tensione di Early", V_A . L'inverso della pendenza delle curve caratteristiche nella zona di saturazione ha le dimensioni di una resistenza. Essa può essere considerata come una resistenza "nascosta" all'interno del transistor che fornisce un contributo aggiuntivo di corrente proporzionale alla tensione ai suoi capi, V_{DS} , ed è per questo stata indicata con r_0 nel simbolo circuitale della Fig.1.7b. Essa è facilmente stimabile in valore considerando il triangolo nel secondo quadrante:

$$r_0 = \frac{V_A}{k(V_{GS} - V_T)^2} \quad (1.5)$$

Notare che mentre V_A è un dato che caratterizza il transistor, e quindi è fornito dal costruttore, il valore di r_0 varia al variare della tensione di comando V_{GS} perché variano le pendenze delle curve: più corrente è fatta circolare nel MOSFET (V_{GS} più grandi) e più piccola risulta la resistenza r_0 . Le curve caratteristiche infatti si aprono a ventaglio, come potete osservare in Fig.1.7.

La frazione di corrente aggiuntiva circolante nel transistor rispetto al caso ideale (curve orizzontali) è facilmente calcolabile quantificando il triangolo evidenziato nella Fig.1.7, così da ottenere la reale corrente circolante in un MOSFET come:

$$I_D \cong k(V_{GS} - V_T)^2 + \frac{k(V_{GS} - V_T)^2}{V_A} \cdot V_{DS} = k(V_{GS} - V_T)^2 \cdot \left(1 + \frac{V_{DS}}{V_A}\right) \quad (1.6)$$

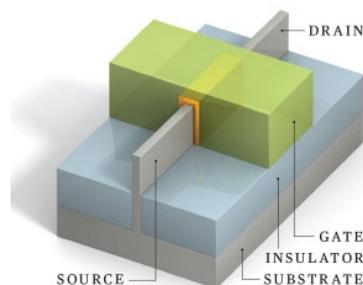
Questa espressione della corrente di Drain è quella usata effettivamente nel **simulatore circuitale SPICE** per il calcolo di primo livello della corrente circolante in un MOSFET, in cui viene richiesto il parametro $LAMBDA=1/V_A$.

1.2.8 Il MOSFET : un dispositivo in continua evoluzione

I transistori MOS si sono evoluti tantissimo nei decenni spinti dall'esigenza di ridurre le dimensioni in modo da velocizzarne la risposta in frequenza grazie ad una distanza minore tra S e D e da poterne disporre in numero maggiore sulla stessa area riducendo quindi i costi a pari funzionalità. Dagli anni 60 ad oggi i progressi nella tecnologia dei chip sono stati straordinari ed in linea con la Legge di Moore, l'osservazione che il numero di transistor su un chip raddoppia all'incirca ogni due anni.

Ma già dagli anni '90 si è intuito che si sarebbe presto giunti alla fine di questo percorso di riduzione delle dimensioni, perché si sarebbero raggiunte le dimensioni degli stessi pochi atomi che avrebbero composto il dispositivo. Da allora tantissime idee sono state esplorate e continuano ad essere proposte per poter continuare a migliorare le prestazioni dei transistori senza necessariamente ridurre le dimensioni lineari.

Un contributo importante l'ha dato Hu Chenming, a lungo professore alla University of California, Berkeley (USA). Hu si stava scontrando con la crescente difficoltà a spegnere i transistori via via che questi venivano scalati nelle dimensioni e alimentati con tensioni sempre più basse, in quegli anni ormai prossime ad 1V. Queste crescenti perdite dovute a correnti tuttavia circolanti anche a transistori spento rendevano i circuiti digitali sempre meno efficienti dal punto di vista energetico. Per migliorare quindi l'efficacia del Gate nell'aprire e soprattutto chiudere il canale, Hu propose di realizzare un Gate che avvolgesse su almeno tre lati un canale tridimensionale così da controllarlo con maggiore efficacia. Per fare questo dovette concepire una tecnologia costruttiva che andasse al di là della classica realizzazione planare per invece diventare tridimensionale. Il transistoro da lui concepito e realizzato venne chiamato FINFET, per richiamare nel nome la pinna dello squalo (fin) che ben veniva evocata dalla struttura del dispositivo, come evidente nella figura accanto. Il primo prodotto industriale che sfruttasse questa tecnologia uscì nel 2011 grazie a Intel. L'odierno processore i7 contiene miliardi di FINFET delle dimensioni minime di 14nm.

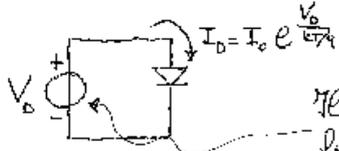


1.3 IL TRANSISTORE BIPOLARE A GIUNZIONE - BJT

TRANSISTORE BIPOLARE

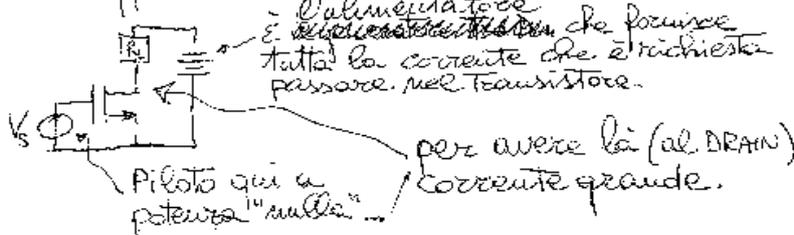
1

① Limite di un bipolo:



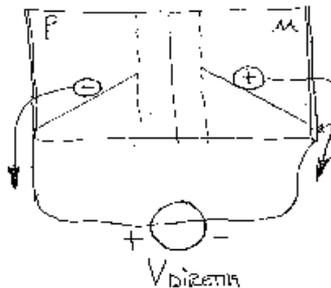
Il generatore di "segnale" deve fornire tutta la potenza utilizzata dal diodo.

Vantaggio di un transistor:



l'alimentazione è ~~separata~~ che fornisce tutta la corrente che è richiesta passare nel transistor.

Partendo dal diodo:

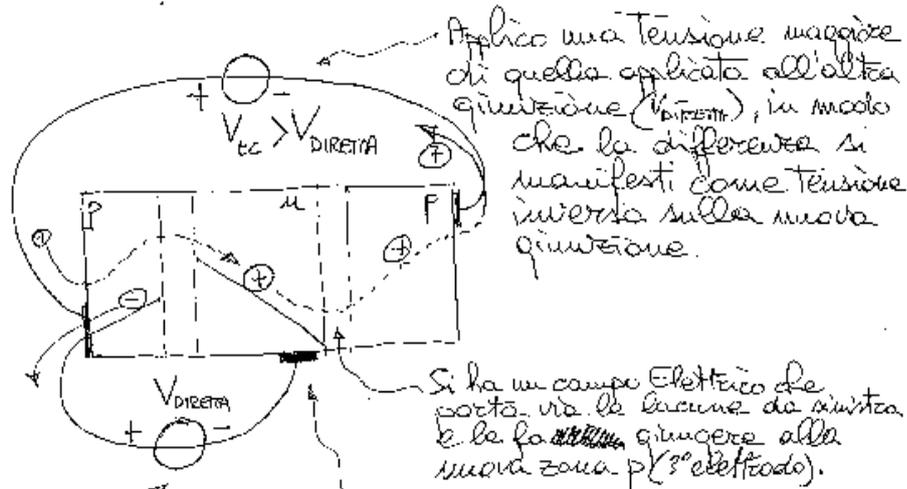


Elettrodo che:

- 1 - definisce il potenziale della zona neutra (che è conduttiva)
- 2 - raccoglie le cariche che stanno diffondendo perché si trova lungo il loro percorso.

Vorrei che i morsetti che modulano la tensione tra p-n NON siano anche quelli che forniscono la corrente. \Rightarrow DEVO TROVARE I POTENZIALI ALTROVE

② Uso un'altra giunzione, questa volta polarizzata in ~~la~~ inversa, per raccogliere la carica:



Con questo generatore controllor, come per il diodo, la tensione ai capi della giunzione. Quindi continuo a controllare la quantità di elettroni e di lacune iniettate nelle corrispondenti zone neutre..

Elettodo che:

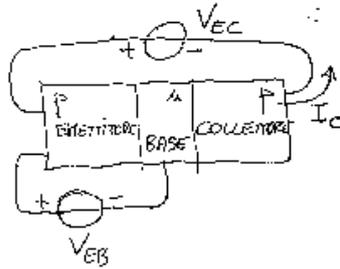
1- definisce il potenziale di tutta la zona n, grazie all'alto drogaggio n che la rende simile ad un conduttore.

2- non raccoglie le lacune perché non è posto lungo il loro percorso di moto per diffusione.

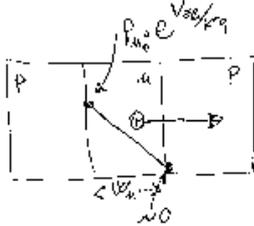
- LE LACUNE FANNO TUTTO IL PERCORSO LUNGO IL NUOVO DISPOSITIVO, però...
- LE LACUNE NON LE FORNISCO CON IL GENERATORE $V_{DIRETTA}$, però...
- IL GENERATORE $V_{DIRETTA}$ TU CONTROLLA IL FLUSSO DI LACUNE.

13

③ Un po' di nomenclatura:



④ Calcolo la corrente di lacune (I_C):



Come per il diodo:

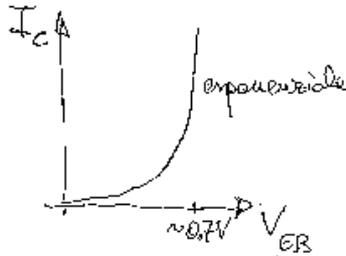
$$I_{\oplus} = q \cdot D_p \cdot \frac{p_{n0} e^{V_{EB}/kT} - 0}{W_B} \cdot Area$$

La corrente è unicamente data da questa espressione. Le zone di Elettrodo e di COLLETTORE non impongono alcun flusso di lacune, ma semplicemente fanno passare quella dettata dalla DIFFUSIONE NELLA BASE.

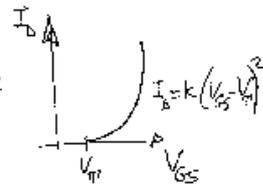
finito dal costruttore

$$I_C = q D_p \frac{p_{n0}}{W_B} \cdot Area \cdot e^{\frac{V_{EB}}{kT}}$$

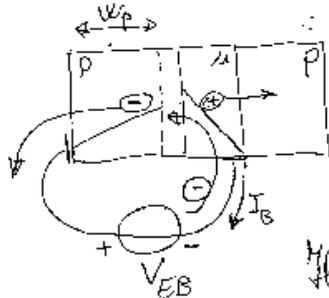
Praticamente indipendente da V_{EC}



Ricordarsi in MOSFET:



⑤ Come la mettiamo con gli elettroni? 4



Gli elettroni iniettati dalla base nell'emettitore circolano nella maglia del generatore V_{EB} , ohmicamente. Peccato!

Il comando non è a potenza nulla, come nel MOSFET ($V_{GS} \cdot I_{GATE} = 0$),

N.B. il moto degli elettroni non interferisce in alcun modo con quello della lacuna. primo. descritto.

⑥ Calcolo la corrente di base (I_B)

$$I_B = q D_n \frac{\mu_{p0} e^{-V_{EB}/kT} - \mu_{p0}}{W_p} \cdot Area = \boxed{q D_n \frac{\mu_{p0} \cdot Area \cdot e^{-V_{EB}/kT}}{W_p}}$$

Questa corrente è inevitabile, come visto, e deve essere ridotta al minimo con particolari attenzioni costruttive (W_p lunga, μ_{p0} piccola e altro).

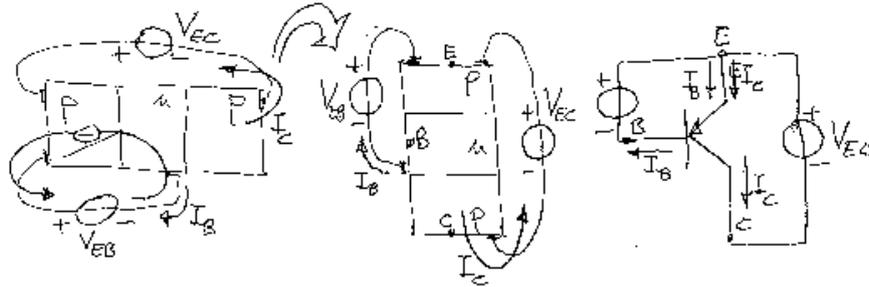
⑦ Importante il rapporto:

$$\beta = \frac{I_C}{I_B} = \frac{D_p \mu_{n0} W_p}{D_n \mu_{p0} W_n} \quad \text{che dipende solo da termini legati alla costruzione.}$$

Benché le singole correnti (I_C, I_B) siano definite da termini (μ_n, μ_p) molto variabili da esemplare a esemplare (anche di parecchi ordini di grandezza), il loro rapporto è molto meglio controllato (vari di poche decine di % tra esemplari dello stesso lotto).

8) Simbolo circuitale e modo d'uso

5



9) Verificare sempre che:

- giunzione EMETTITORE-BASE sia in DIRETTA
- " " BASE-COLLEITORE sia in INVERSA.

1.3.1 Il principio di funzionamento

Il transistor bipolare a giunzione (BJT - *Bipolar Junction Transistor*) è un dispositivo costituito da tre zone di semiconduttore contigue, chiamate **Elettore**, **Base** e **Collettore**, in cui quella centrale è drogata di segno opposto alle altre due in modo da formare due giunzioni *pn*, una tra Elettore e Base e l'altra tra Base e Collettore. Ognuna delle 3 zone di semiconduttore è collegata ad un proprio elettrodo metallico tramite un contatto Ohmico in modo da poter applicare localmente la tensione e raccogliere la corrente. Nel funzionamento normale del BJT, la giunzione Elettore-Base è polarizzata direttamente e quella Base-Collettore è polarizzata inversamente.

Il principio di funzionamento di un transistor *npn* è illustrato nella Fig.1.8 in cui è riportata la sezione di un dispositivo realizzato con tecnologia planare. Gli elettroni, iniettati nella Base dalla giunzione Elettore-Base polarizzata direttamente, diffondono verso la giunzione di Collettore (dall'alto verso il basso

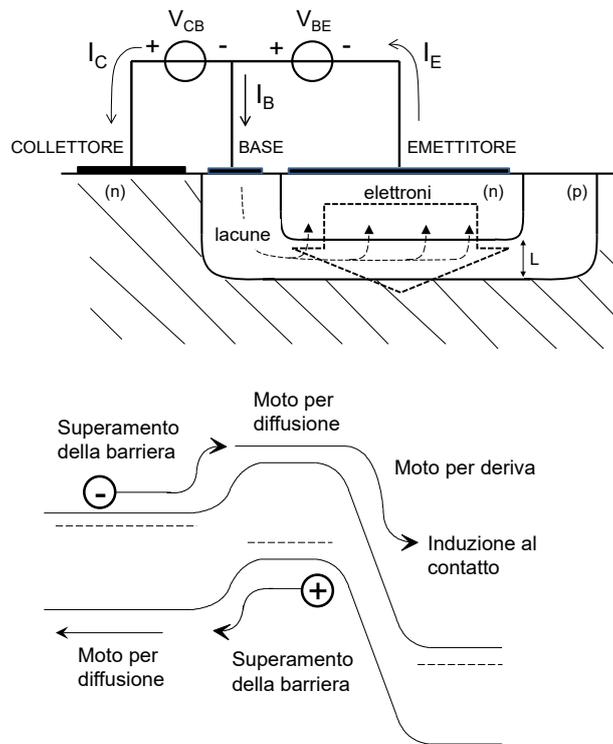


Fig. 1.8 Percorsi dei singoli portatori (elettroni e lacune) sia nella sezione che nel diagramma a bande di energia potenziale di un transistor *npn* prototipo.

nella sezione). Se lo spessore della Base neutra, L , è molto più piccolo della lunghezza di diffusione degli elettroni, la quasi totalità degli elettroni iniettati raggiunge la giunzione Base-Collettore. Essendo questa giunzione polarizzata inversamente, il campo elettrico nella sua regione svuotata è tale da continuare a favorire il moto degli elettroni e trascinarli effettivamente nel Collettore. Nella sezione della Fig.1.8 il Collettore è costituito da tutta la regione tratteggiata: essendo essa drogata, può essere considerata tutta conduttiva e quindi gli elettroni raccolti al Collettore sono immediatamente disponibili al contatto di Collettore in superficie. In un buon transistor bipolare, quindi, il flusso di elettroni raccolti al Collettore (che definisce la corrente I_C) è praticamente uguale al flusso di elettroni partiti dall'Emettitore.

Questo movimento degli elettroni dall'Emettitore al Collettore può anche essere illustrato facendo riferimento ai diagrammi energetici lungo le giunzioni (Fig.1.8 in basso). La polarizzazione diretta, V_{BE} , della giunzione Base-Emettitore causa un corrispondente abbassamento della barriera energetica vista dagli elettroni che si trovano nell'Emettitore. Essi, quindi, sono iniettati nella regione di Base. Essendo questa una zona drogata, essa è conduttiva e pertanto non vi è lungo di essa una caduta di potenziale significativa, come messo in luce dalla costanza del livello di energia nel grafico. Essendo il campo elettrico quasi inesistente, nella Base i portatori iniettati dall'Emettitore si muovono per diffusione, stimolati cioè dal gradiente di concentrazione lungo l'asse Emettitore-Collettore (tanti elettroni nella parte di Base adiacente all'Emettitore e pochi elettroni nella parte di Base adiacente al Collettore). Alla giunzione Base-Collettore infatti, come già accennato, gli elettroni sentono la ulteriore diminuzione dell'energia potenziale, dovuta alla polarizzazione inversa, V_{CB} , che li trascina via dalla Base dentro nel Collettore, dove diventano maggioritari così da indurre immediatamente il segnale nel filo esterno di contatto del Collettore.

La diminuzione della barriera di energia tra Emettitore e Base consente contemporaneamente anche alle lacune maggioritarie della Base di essere iniettate dalla Base nell'Emettitore. Quando in Emettitore, le lacune si trovano ad essere minoritarie in una zona altamente drogata (e quindi senza presenza di campo elettrico come è effettivamente l'Emettitore) e si muoveranno per diffusione (al lato verso la Base le lacune sono tante, al lato verso il contatto ohmico esterno la densità è prossima a zero) attraverso l'Emettitore fino a concludere il loro moto al contatto di Emettitore, inducendo un ugual segnale sul filo del contatto. Ad ogni lacuna iniettata dalla Base nell'Emettitore corrisponde una lacuna richiamata dal filo di contatto nella Base per conservare la neutralità elettrica della regione di Base, dando luogo a quella che chiameremo **corrente di Base**, I_B . Poiché normalmente l'Emettitore è molto più drogato della Base, il flusso di portatori (I_B) iniettati dalla Base verso l'Emettitore è molto minore del flusso (I_C) di elettroni iniettati dall'Emettitore verso la Base che poi raggiungono il Collettore.

Tra le correnti ai morsetti del dispositivo vale quindi la seguente relazione di bilancio:

$$I_E = I_C + I_B \quad (1.7)$$

Per quanto visto, il rapporto I_C/I_E ha un valore molto vicino all'unità ed è indicato con il simbolo α

$$\alpha = \frac{I_C}{I_E} \quad (1.8)$$

Esso rende conto del trasporto principale di portatori lungo la direttrice Emittitore-Collettore del transistor.

Il rapporto tra la corrente di Collettore e la corrente di Base rende conto invece del guadagno di corrente del transistor bipolare ed è indicato con il simbolo h_{FE} oppure β :

$$\beta = \frac{I_C}{I_B} \quad (1.9)$$

In commercio si trovano transistori con valori di β compresi tra 10 (transistori di potenza) e 5000 (transistori di segnale a superbeta). Normalmente sono impiegati transistori con $\beta=100 \div 400$. Da ultimo, in base alla (1.8) e ricordando la definizione di α è facile ricavare che $\beta = \alpha/(1-\alpha)$ e $\alpha = \beta/(\beta+1)$. Se $\beta=100$ si ha $\alpha=0.99$.

1.3.2 La corrente di Collettore

Per ricavare le relazioni funzionali tra le tensioni applicate ai morsetti del transistor e le correnti circolanti si faccia riferimento alla Fig.1.9 in cui sono riportati i profili di concentrazione degli elettroni nella Base e delle lacune nell'Emittitore di un transistor *npn* quando la giunzione Emittitore-Base è polarizzata in diretta e la giunzione Base-Collettore è polarizzata in inversa, come deve essere nella condizione normale di funzionamento. Trascurando la eventuale (e comunque piccolissima) ricombinazione degli elettroni con le lacune maggioritarie della Base durante il loro moto da Emittitore a Collettore, il profilo della densità di elettroni è rettilineo lungo la Base con:

i) un valore di concentrazione al bordo della giunzione con l'Emittitore pari a:

$$n(0) = \frac{n_i^2}{N_A} e^{V_{BE}/V_{th}} \quad (1.10)$$

dove n_i è la concentrazione intrinseca e N_A è la concentrazione dei droganti accettori nella Base;

ii) un valore di concentrazione all'estremità della Base verso il Collettore quasi nulla, perché gli elettroni vengono trascinati via verso il Collettore dal campo elettrico della giunzione polarizzata inversamente.

La corrente di Collettore è pari alla corrente di diffusione degli elettroni nella Base e quindi è proporzionale alla pendenza del loro profilo $n(x)$ secondo la relazione:

$$I_C = qAD_n \frac{dn(x)}{dx}$$

dove A è la sezione del dispositivo e D_n è il coefficiente di diffusione degli elettroni. Sostituendo si ottiene:

$$I_C = qAD_n \frac{n_i^2}{W_B N_A} e^{V_{BE}/V_T} = I_S e^{V_{BE}/V_T} \quad (1.11)$$

dove W_B è la lunghezza della zona neutra della Base. La relazione (1.11) ci dice che al variare della tensione diretta V_{BE} del transistor, la corrente di Collettore varia seguendo la legge esponenziale tipica del diodo Emettore-Base. Per come abbiamo ricavato la (1.11), essa ci dice anche un'altra cosa fondamentale e cioè che la corrente I_C è sostanzialmente indipendente dal valore della tensione del Collettore: infatti, al variare della tensione V_{BC} inversa varia l'entità del campo elettrico nella regione svuotata tra Base e Collettore ma non modifica la capacità di raccolta degli elettroni da parte del Collettore stesso.

Si noti come la corrente I_S , detta **corrente di saturazione inversa**, sia proporzionale all'area del transistor. Quindi, transistori realizzati con gli stessi processi tecnologici e polarizzati a pari V_{BE} , conducono correnti proporzionali alle relative aree di Emettore.

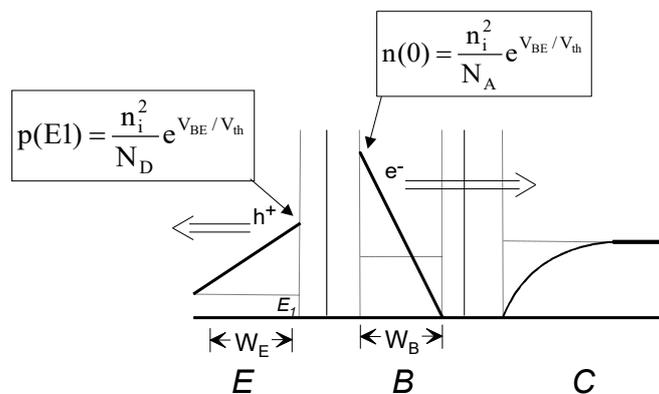


Fig. 1.9 Distribuzione della concentrazione dei portatori minoritari in un transistor npn polarizzato in zona attiva diretta.

1.3.3 La corrente di Base

Come già accennato, contemporaneamente alla iniezione di elettroni dall'Emettitore alla Base, quando si applica una tensione V_{BE} diretta si ha anche una iniezione di lacune dalla Base nell'Emettitore. La concentrazione di lacune appena entrati in Emettitore (E_1 nella figura) è pari, in analogia con la (1.10), a:

$$p(E_1) = \frac{n_i^2}{N_D} e^{V_{BE}/V_{th}} \quad (1.12)$$

Al contatto metallico di Emettitore invece non ci sarà alcun accumulo di lacune, perché tutte quelle che lì vi giungono vengono subito portate via dal filo del contatto. La distribuzione di lacune lungo l'Emettitore segue quindi il profilo lineare visibile nella Fig.1.9 e permette di calcolare il valore della corrente di lacune che si sta muovendo per diffusione:

$$I_B = qAD_p \cdot \text{grad}(p(x)) = qAD_p \frac{n_i^2}{W_E N_D} e^{V_{BE}/V_T} = I_K e^{V_{BE}/V_T} \quad (1.13)$$

dove W_E è lo spessore dell'Emettitore. Questo contributo costituisce la Corrente di Base, I_B , del transistor perché le lacune iniettate in Emettitore devono essere costantemente rifornite dal contatto metallico della Base. La (1.13) ci mostra che anche la corrente di Base ha un andamento esponenziale con la tensione di comando applicata tra Base ed Emettitore, V_{BE} . Il β del transistor:

$$\beta = \frac{I_C}{I_B} = \frac{D_n N_D W_E}{D_p N_A W_B}$$

dipende solo da termini legati alla costruzione, quali dimensioni e drogaggi. Benché ognuno di questi termini può variare molto da lotto a lotto di fabbricazione, il loro rapporto tende ad essere ragionevolmente controllato, per cui il valore di β di transistori dello stesso tipo differisce al più di qualche decina di %.

1.3.4 I simboli circuitali ed il funzionamento in zona attiva diretta

Nella Fig.1.10 sono riportati i simboli circuitali dei transistori BJT *pnp* ed *npn*. La freccia è sempre sul terminale di Emettitore ed indica il verso della corrente. In un transistor *pnp* la corrente è equiversa al flusso di lacune dall'Emettitore verso il Collettore, mentre in un transistor *npn* la corrente scorre dal Collettore all'Emettitore nel verso opposto all'effettivo flusso di elettroni.

Nella figura sono anche indicati i versi delle tensioni tra i terminali quando il dispositivo funziona in **zona attiva diretta** in cui la giunzione Base-Emettitore è polarizzata direttamente e la giunzione Base-Collettore è polarizzata inversamente.

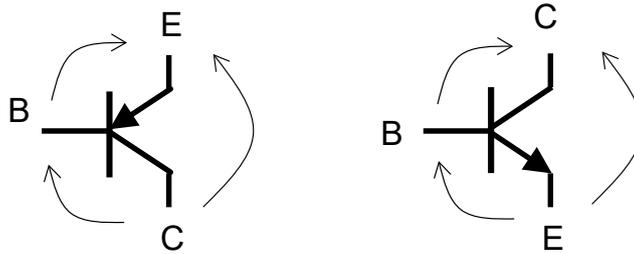
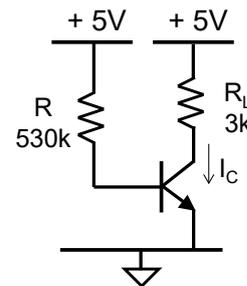
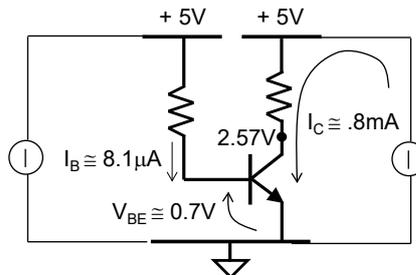


Fig. 1.10 Simboli circuitali dei transistori bipolari a giunzione: a) pnp; b) npn.

E 1.4 Ricavare le tensioni ai capi del transistor collegato come nella figura e le correnti in esso circolanti. Verificare che il transistor effettivamente si trovi a funzionare in zona attiva diretta. Si supponga che il BJT abbia $\beta=100$.



La Base è connessa all'alimentazione positiva attraverso un resistore R, per cui la giunzione Base-Elettore è certamente polarizzata direttamente e, come accade ai capi di un semplice diodo, la tensione V_{BE} si porta a circa 0.7V. La corrente di Base è quindi $I_B \cong 4.3V/R = 8.1\mu A$ e la corrente di Collettore $I_C = \beta I_B = 0.81mA$. Il potenziale del morsetto di Collettore è quindi pari a $V_C = 5V - 2.43V = 2.57V$.



È facile verificare che la giunzione Base-Collettore è polarizzata inversamente di $2.57V - 0.7V = 1.87V$ e che quindi il BJT opera effettivamente in zona attiva diretta. Si noti come una differente scelta del valore di V_{BE} (ad esempio $V_{BE} = 0.67V$ o $V_{BE} = 0.72V$, perché in effetti esso non ci è rigorosamente noto) non avrebbe condotto ad una valutazione significativamente diversa delle correnti circolanti.

Si supponga ora di far variare il potenziale dell'alimentatore connesso alla resistenza R di Base dell'esercizio precedente. Si pensi di diminuire il valore della tensione erogata da 5V a 4V. La corrente di Base diminuisce da $8.1\mu\text{A}$ a $6.2\mu\text{A}$. La variazione $\Delta I_B = 1.9\mu\text{A}$ determina una diminuzione $\beta \Delta I_B = 0.19\text{mA}$ della corrente di Collettore ed una diminuzione di $\beta \Delta I_B R_L = 0.57\text{V}$ della caduta di tensione ai capi del resistore di carico. In definitiva la variazione ΔI_B determina un aumento del potenziale del nodo di Collettore che passa da 2.57V a 3.14V.

Questo è stato l'effetto inizialmente osservato da J. Bardeen e W. H. Brattain: variando la corrente al morsetto di Base, varia la corrente che fluisce nel circuito principale, costituito dalla maglia che contiene R_L , il Collettore, l'Emettitore e l'alimentazione da 5V. Il dispositivo attivo si comporta quindi come una valvola in un condotto idrico: la posizione della valvola controlla la portata del condotto principale. In questo caso la variazione della piccola corrente di Base fa variare la corrente ben maggiore che fluisce tra Collettore ed Emettitore.

Come variabile di comando del transistor si può anche pensare alla tensione V_{BE} tra Base e Collettore: variando di poco la tensione V_{BE} infatti si riesce a variare di molto la corrente I_C e quindi la tensione V_C , con un effetto di "amplificazione" molto interessante.

I transistori sono quindi dispositivi con cui si può controllare e modulare la corrente che fluisce tra Emettitore e Collettore agendo su un opportuno elettrodo di comando, la Base del BJT. Ai loro morsetti si ha una amplificazione di potenza, nel senso che spendendo una piccola potenza per il controllo si modula l'erogazione di una potenza ben maggiore ai capi del carico.

Si noti come, fissata la corrente che deve fluire in un carico, all'aumentare del β del transistor bipolare diminuisce la corrente di Base e quindi la potenza dissipata nel ramo di controllo. Ecco quindi perchè il β è un parametro di merito del transistor.

1.3.5 Le curve caratteristiche

Zona attiva diretta - La relazione esponenziale (1.11) che lega la tensione di comando V_{BE} alla corrente al terzo morsetto (Collettore) è riportata nella Fig.1.11a. Poichè essa lega la corrente erogata ad una grandezza della maglia di controllo questa curva è detta *curva transcaratteristica* del BJT e ricalca l'andamento tipico della relazione corrente-tensione di una giunzione *pn*. Diversamente, fissata la corrente di Base (o equivalentemente la tensione V_{BE}) è possibile tracciare la dipendenza della corrente I_C con la tensione V_{CE} . Questa è chiamata *curva caratteristica* ed un suo esempio è mostrato nella Fig.1.11b.

Nelle applicazioni in cui è impiegato per amplificare dei segnali, il transistor è polarizzato nella zona attiva diretta, ovvero nella regione in cui le curve caratteristiche sono quasi piatte (I_C praticamente indipendente dalla tensione V_{CE}) e la corrente circolante è βI_B . In questa regione il transistor si comporta come un

generatore di corrente comandato dalla corrente di Base I_B o dalla tensione V_{BE} . La massima corrente erogabile dal transistor è solo limitata dalla massima potenza termica dissipabile dal dispositivo, mentre la massima tensione V_{CE} deve essere inferiore alla tensione di rottura (*breakdown*) della giunzione Collettore-Base indicata dal costruttore.

Zona di saturazione - La primissima parte delle curve caratteristiche, per tensioni V_{CE} inferiori a 0.7V circa, corrisponde al caso in cui la giunzione Base-Collettore è polarizzata direttamente. In questo caso l'iniezione dei portatori minoritari nella Base avviene da entrambe le giunzioni. I minoritari in Base hanno un profilo ancora rettilineo come quello della Fig. 1.9, ma la concentrazione di minoritari alla giunzione Base-Collettore non è più nulla. La pendenza del profilo è minore rispetto a quella che si avrebbe se il transistor funzionasse in zona attiva, e quindi anche la corrente di Collettore è minore. Questa condizione di funzionamento, in cui la Base è invasa (*saturata*) dai portatori minoritari provenienti anche dal Collettore, costituisce la cosiddetta zona di saturazione di un BJT, che è bene evitare quando si progettano circuiti analogici. Considerando una singola curva caratteristica di Fig. 1.11b ad I_B costante, la condizione di saturazione è quindi identificata dalla diseuguaglianza:

$$I_C < \beta I_B$$

Zona di interdizione - Quando la tensione V_{BE} è molto piccola, o addirittura si ha una polarizzazione inversa della giunzione Elettore-Base, la corrente I_C è praticamente nulla e si dice che il transistor bipolare è nella zona di interdizione.

Nelle applicazioni in cui il BJT è usato per amplificare segnali, le condizioni di saturazione e di interdizione sono da evitare. Il funzionamento in saturazione ed

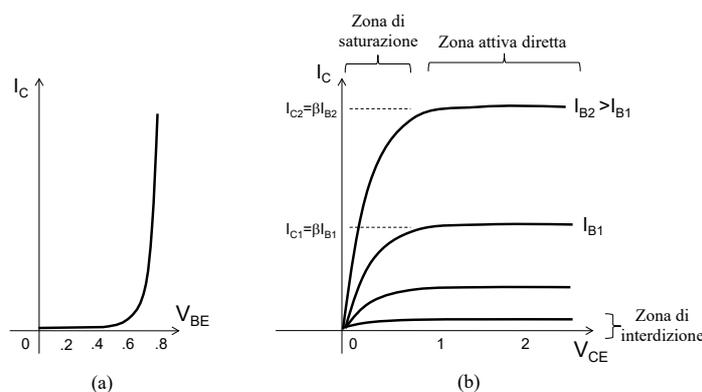


Fig. 1.11 Esempio di (a) curva transcaratteristica e (b) di famiglia di curve caratteristiche di un transistor bipolare npn.

in interdizione è invece utilizzato nelle applicazioni digitali ed in generale quando il transistor è utilizzato come interruttore. Si intuisce infatti come la condizione di interdizione corrisponda allo stato di interruttore aperto, mentre la condizione di saturazione, in cui i potenziali dell'Emettitore e del Collettore sono tra loro vicini (circa 0.2V), corrisponda allo stato di interruttore chiuso.

Attenzione !

Purtroppo si indica con lo stesso termine di *saturazione* due condizioni di funzionamento nel BJT e nel FET completamente diverse, corrispondenti a due tratti diversi delle relative curve caratteristiche.

Nei MOSFET la saturazione corrisponde al tratto di curva caratteristica in cui la corrente di Drain non varia più il suo valore (è saturata) nonostante la variazione della tensione di Drain. È questa la zona di funzionamento di maggior interesse per le applicazioni circuitali analogiche lineari.

Nei BJT invece, la saturazione corrisponde ad un accumulo eccessivo (saturazione) di portatori nella Base provenienti anche dal Collettore ed è in genere una zona di funzionamento da evitare.

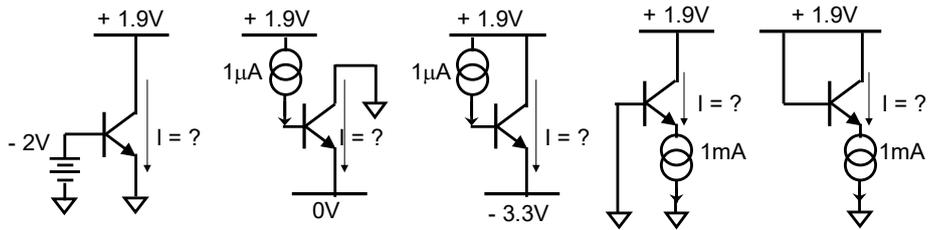
PRATICA di POLARIZZAZIONE

Quali tensioni possono esserci ai tre morsetti di un BJT affinché funzioni correttamente ?

Quali tensioni non possono essere applicate ai tre morsetti in quanto farebbero lavorare il BJT in una zona diversa dalla zona attiva diretta ?

I seguenti quiz dovrebbero aiutarvi a prendere confidenza con questi aspetti fondamentali di un BJT in un circuito elettronico.

ES.1 Si supponga che i BJT abbiano $\beta=200$:



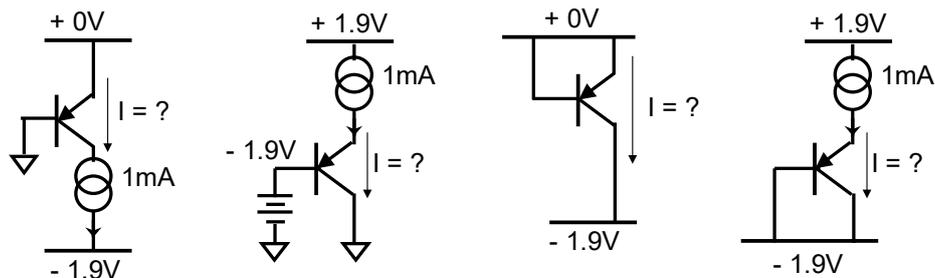
Il BJT è nella corretta zona di funzionamento ?

- si no si no si no si no si no

Se si, quanta corrente I scorre ?

- I = I = I = I = I =

ES.2 Si supponga che i BJT abbiano $\beta=200$:



Il BJT è nella corretta zona di funzionamento ?

- si no si no si no si no

Se si, quanta corrente I scorre ?

- I = I = I = I =

1.3.6 La transconduttanza

Differenziando la relazione (1.11) rispetto alla tensione di comando V_{BE} , si ottiene la transconduttanza del dispositivo in zona attiva diretta:

$$g_m = \frac{\partial I_C}{\partial V_{BE}} = \frac{I_S e^{\frac{V_{BE}}{V_T}}}{V_{th}} = \frac{I_C}{V_{th}} \quad (1.14)$$

che definisce quanto varia la corrente che fluisce nel Collettore a causa di una piccola variazione della tensione di comando V_{BE} . La transconduttanza dipende dal valore della corrente di polarizzazione I_C ed aumenta proporzionalmente con I_C . Questo è comprensibile perché la transconduttanza è la pendenza della curva esponenziale e quindi aumenta all'aumentare di V_{BE} , cioè di I_C come graficamente evidente dalla Fig.1.12. Ad esempio, con $I_C=1\text{mA}$ e ricordandosi che $V_{th}=25\text{mV}$ si ottiene $g_m=40\text{mA/V}$.

1.3.7 Effetto Early e resistenza di Collettore

Le curve caratteristiche nella zona attiva di un BJT reale non sono perfettamente parallele all'asse delle tensioni ma leggermente inclinate come nella Fig.1.13. Il Collettore quindi non si comporta come un generatore di corrente ideale, erogando una corrente dipendente solo dalla giunzione Emittitore-Base, ma fornisce una corrente che risente, anche se poco, del suo stesso valore di tensione. Per illustrare le ragioni di questa non idealità del transistor BJT si faccia riferimento al transistor *npn* prototipo della Fig.1.14. Si tenga fissa la tensione V_{BE} (equivalente a muoversi lungo una definita curva caratteristica nella Fig.1.13). All'aumentare della tensione V_{CE} , aumenta la polarizzazione inversa della giunzione Base-

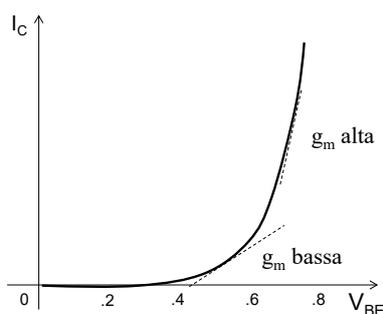


Fig. 1.12 Esempio di (a) curva transcaratteristica e (b) di famiglia di curve caratteristiche di un transistor bipolare npn.

Collettore e quindi aumenta lo spessore della zona di carica spaziale della giunzione stessa. Corrispondentemente diminuisce lo spessore neutro della Base, W . Questo produce una aumentata pendenza della distribuzione dei minoritari nella Base che comporta un aumento della corrente di collettore anche se V_{BE} resta costante, come evidenziato nella curva caratteristica. Questo effetto di modulazione della lunghezza della Base al variare della tensione V_{CE} è noto come **effetto Early**.

Quantitativamente esso si esprime nel fornire un valore di tensione V_A , detta **tensione di Early**, che identifica il punto nella Fig.1.13 dove con buona approssimazione convergono le curve caratteristiche. Poiché il grafico ha sugli assi una tensione ed una corrente, è intuitivo definire la "resistenza di collettore" equivalente del BJT come:

$$r_o \cong \frac{V_A}{\beta \cdot I_B} = \frac{V_A}{I_0 \cdot e^{\frac{qV_{BE}}{kT}}} \quad (1.15)$$

e rappresentarla nel simbolo circuitale della Fig.1.13 da una resistenza di uscita r_o tra Collettore ed Emettore. Si noti come il valore di r_o dipenda anche dalla corrente a cui il transistor opera: maggiore è la corrente circolante maggiore è la pendenza della curva e quindi minore è il valore di r_o corrispondente. Ciò proprio perché V_A è un valore fisso, in prima approssimazione definito dalla tecnologia costruttiva. La tensione di Early di un BJT ha un valore tipico compreso tra 50V÷200V che, per correnti di collettore intorno ad 1mA, corrisponde ad una resistenza di Collettore di 50 kΩ÷200kΩ.

Per ricavare la effettiva corrente circolante nel transistor date le tensioni ai morsetti, basta sommare al valore ideale (quello che si avrebbe per $V_A = \infty$) il contributo aggiuntivo dovuto alla pendenza (per comodità si considera il triangolo)

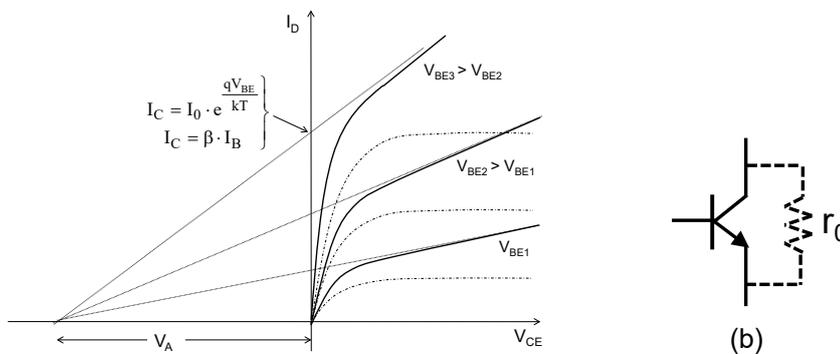


Fig. 1.13 (a) Curva caratteristica di un transistor bipolare reale con in evidenza la tensione di Early V_A e (b) visualizzazione della resistenza equivalente prodotta.

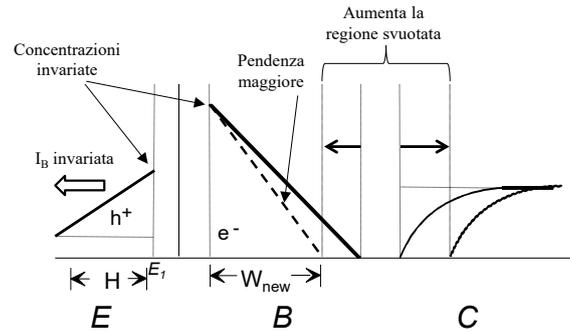


Fig. 1.14 Effetto di un aumento della tensione inversa tra Collettore e Base in un transistor npn.

$$I_C = I_0 \cdot e^{\frac{q(V_{BE} + v_{bc})}{kT}} + I_0 \cdot e^{\frac{q(V_{BE} + v_{bc})}{kT}} \frac{V_{CE}}{V_A} = I_0 \cdot e^{\frac{q(V_{BE} + v_{bc})}{kT}} \cdot \left(1 + \frac{V_{CE}}{V_A} \right)$$

Concludendo questo primo capitolo hai concluso la tua prima fatica. Sentiti come Ercole che per espiare il fatto di essersi reso colpevole della morte della sua famiglia, viene convinto dall'oracolo a mettersi al servizio di Euristeo ed a compiere una serie di imprese, ottenendone alla fine come compenso l'immortalità.

La prima fatica di Ercole fu catturare e uccidere il leone di Nemea, una bestia dalla pelle durissima che non poteva essere scalfita, che sbranava e divorava animali e pastori. Ercole scelse di battersi a mani nude contro la belva, cingendogli il collo con forza e riuscendo a stringerla fino a soffocarla. Dopo aver ucciso il leone Ercole conció la pelle invincibile del leone e ne fece un mantello per se stesso, mentre della testa ne fece un elmo. La prima fatica era compiuta.

