

1 Lezione del 02-03-26

1.1 Introduzione

Lo scopo del corso è quello di fornire gli elementi base per la comprensione e l'analisi dei circuiti elettronici sia analogici sia digitali.

In particolare, gli argomenti saranno:

- I **dispositivi a semiconduttore**, fra cui in particolare i *diodi*, i transistori bipolari (*BJT*), e i transistori ad effetto di campo (*MOSFET*). Vedremo come questi, ed in particolare l'ultimo (in particolare nella tecnologia *CMOS*), sono alla base dell'elettronica digitale (quindi lo sviluppo di elementi di elaborazione, memoria, ecc...);
- Lo studio dei **circuiti e sistemi analogici**, e quindi in particolare gli amplificatori a transistori, gli amplificatori operazionali, i regolatori di tensione, e i convertitori *A/D* e *D/A*;
- Lo studio dei **circuiti e sistemi digitali**, e quindi le porte logiche *CMOS* e bipolari, i flip-flop, e le memorie a semiconduttore. Qui riprenderemo gli argomenti visti ad esempio in reti logiche (le porte logiche), soffermandoci sull'implementazione effettiva di queste attraverso componenti elettronici.

Parte del corso consisterà nell'uso dello strumento **LTSPICE** per la simulazione dei circuiti studiati.

1.1.1 Note storiche

"L'elettronica è la scienza e la tecnologia del movimento delle cariche in un gas, nel vuoto o in un semiconduttore".

Riportiamo questa frase in quanto tutti i dispositivi che studieremo non si occupano di altro che influire sullo spostamento di cariche libere. Implementiamo diverse funzioni decidendo se determinate cariche devono essere bloccate, lasciate passare, ecc...

Quindi, l'elettronica è la disciplina che si occupa dello sviluppo di sistemi che sono capaci di gestire il movimento di cariche mobili.

1. L'elettronica inizia nel '900 attraverso componenti *a vuoto*, al tempo detti *valvole*. In particolare, nel 1904 Fleming inventa il **diodo a vuoto**, un componente capace di far passare o meno corrente fra 2 contatti. Nel 1906 De Forest inventa il **triolo a vuoto**, dove il terzo contatto serve a modulare la corrente che passa fra i primi 2.
2. Nel 1947, nei laboratori Bell viene inventato il **transistore bipolare** (*BJT*), mentre nel 1958 Kilby concepisce il **circuito integrato** (entrambe le scoperte vinceranno il Nobel per la fisica).
Agli anni immediatamente successivi risale la *legge di Moore*, che afferma che il numero di componenti sui chip raddoppia ogni anno (corretto a 2 anni nel 1975).
3. Lo sviluppo dei chip negli ultimi anni si è concentrato nelle mani di un numero ristretto di multinazionali (Intel, TSMC, ecc...). Questo deriva dal fatto che la

fabbricazione dei chip richiede risorse che sono diventate via via più costose nel tempo.

Diversi *trend* sono stati individuati negli anni:

- Le dimensioni dei circuiti stampati sui chip si sono ridotte: storicamente il metro di riferimento era la larghezza dei canali dei transistor MOSFET, che generalmente corrispondeva alla minore traccia che si poteva avere. Ad esempio, nel 2025 i moderni processi stanno intorno ai 2 nm;
- Il costo per transistor dei circuiti stampati si è ridotto di molto negli anni, portando ad avere circuiti con molti più transistor allo stesso costo (o a costi minori). Da qui si rimanda alla sopra riportata *legge di Moore*. Ad esempio, negli ultimi anni i chip che vengono fabbricati hanno un numero che sta intorno alle centinaia di miliardi di transistor.

I chip vengono realizzati attraverso il classico processo, a partire da **wafer** di silicio, che vanno a formare il *substrato* di semiconduttore sui quali i componenti elettronici vengono realizzati. In particolare:

- I wafer vengono formati attraverso un dato processo chimico, e sono formati da silicio purissimo;
- Sui wafer viene formato uno strato di diossido di silicio, un isolante;
- Attraverso la *fotolitografia*, lo strato di diossido di silicio viene tagliato in modo da esporre zone di silicio;
- Le zone esposte subiscono un processo di *doping* di Boro o altri elementi come il fosforo, che crea zone con diverse caratteristiche di conduttività (questo, come vedremo, è reso possibile dal fatto che il silicio è un *semiconduttore*);
- Il processo può quindi essere ripetuto per generare più strati, il wafer può essere *passivato* (ricoperto da uno strato protettivo), tagliato e messo in un *package*.

Oggi lo sviluppo della microelettronica si basa su due approcci:

- Approccio **System-on-Chip** (*SoC*): adottato per lo sviluppo dei componenti a contenuto digitale. Rispecchia la cosiddetta legge del "*more Moore*", quindi si basa sull'evoluzione quantitativa di tecnologie già provate (più transistor, ecc...);
- Approccio **System-in-Package** (*SiP*): adottato per lo sviluppo dei componenti a contenuto analogico, o in generale che si interfacciano col mondo esterno. Rispecchia la cosiddetta legge del "*more than Moore*", quindi mira alla diversificazione delle tecnologie e allo sviluppo di soluzioni alternative rispetto a quelle già provate.

1.2 Richiami di chimica

Alle basi dell'elettronica sta la chimica. Richiamiamo quindi alcuni concetti fondamentali.

1.2.1 Atomo

La visione dell'atomo passa attraverso più modelli:

- Modello di **Thomson**: si basa sull'osservazione delle particelle fondamentali e della loro carica ($+1 e$ per il protone e $-1 e$ per l'elettrone), e sul fatto che la carica complessiva dell'atomo risultava neutra.

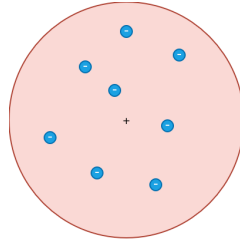


Figura 1: Atomo di Thomson

1.1: Modello di Thomson

Vede l'atomo come un insieme di elettroni carichi negativamente immersi in una "nube" carica positivamente (*plum pudding model*)

- Modello di **Rutherford**: si basa sull'osservazione che la carica positiva dell'atomo si trova concentrata al centro di esso.

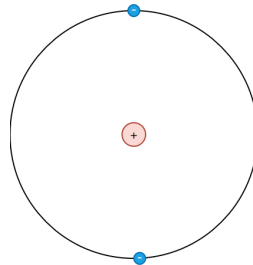


Figura 2: Atomo di Rutherford

1.2: Modello di Rutherford

Vede l'atomo come un nucleo carico positivamente molto piccolo, il cui volume circostante è occupato da elettroni;

- Modello di **Bohr** dell'atomo di idrogeno. Prendendo questo semplice elemento come riferimento, si può effettuare la semplificazione di avere un singolo elettrone (con carica $-1 e$) in un nucleo formato da un singolo protone ($+1 e$).

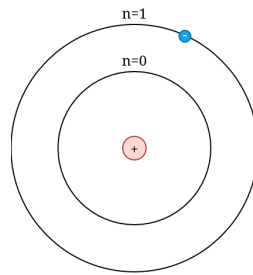


Figura 3: Atomo di idrogeno di Bohr

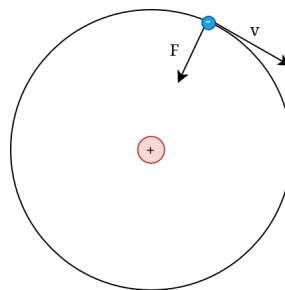
1.3: Modello di Bohr

All'elettrone vengono permessi solo certi stati di moto stazionario dove questo si trova ad un dato livello di energia. Quando un elettrone si trova in un dato stato, l'atomo non irradia energia.

Secondo Bohr le orbite degli elettroni erano circolari (sappiamo che questo è falso, in quanto conosciamo che la posizione degli atomi non è immediatamente conoscibile e si può descrivere con particolari funzioni d'onda dette *orbitali*).

Orbite di Bohr

Possiamo comunque sfruttare la meccanica classica per spingere avanti il modello di Bohr ed ottenere informazioni concrete sull'atomo, nel caso semplice dell'idrogeno.



1. Partiamo dalla meccanica classica. Assumiamo orbite circolari, e quindi la presenza di una certa forza F che mantenga l'elettrone di massa m su un'orbita a distanza r e velocità v del nucleo:

$$F = ma = -\frac{mv^2}{r}$$

Questa forza F non potrà essere che la forza di attrazione elettrostatica definita dalla *legge di Coulomb*, cioè:

$$F = k \frac{q_1 q_2}{r^2} \implies -\frac{mv^2}{r} = k \frac{q_1 q_2}{r^2}$$

Ora, q_1 e q_2 sono le cariche del nucleo ($+1 e$) e dell'elettrone in orbita ($-1 e$), per cui si ha l'espressione finale:

$$\frac{mv^2}{r} = \frac{ke^2}{r^2}$$

2. Abbiamo detto che secondo il modello di Bohr gli stati di energia sono *quantizzati*. Scegliamo quindi una costante \hbar (che chiameremo *costante di Planck*), e un intero n , e quantizziamo il momento angolare L :

$$L = mvr = n\hbar$$

Abbiamo quindi ottenuto il sistema di equazioni non lineari:

$$\begin{cases} \frac{mv^2}{r} = \frac{ke^2}{r^2} \\ mvr = n\hbar \end{cases}$$

Possiamo risolvere tale sistema per ricavare v ed r , cioè velocità a raggio dell'orbita:

$$v = \frac{ke^2}{n\hbar}, \quad r = \frac{n^2\hbar^2}{kme^2}$$

Note queste grandezze, si può ad esempio calcolare l'energia ad ogni livello semplicemente come l'energia cinetica:

$$E = \frac{1}{2}mv^2 = \frac{1}{2}m \frac{k^2e^4}{n^2\hbar^2}$$

Questo tipo di calcoli smette di essere accurato appena si introduce un nuovo elettrone (e quindi un nuovo protone per rendere stabile l'atomo), in quanto si vanno a introdurre, oltre alla forza di attrazione verso il nucleo, forse di interazione fra gli elettroni stessi.

Irradiazione

Abbiamo nominato il concetto di *irradiazione*. Quello che i primi studiosi dell'atomo notavano (e su cui basavano le loro ipotesi) è che quando si espongono gli atomi ad energia, questi irradiano, ovvero emettono *onde elettromagnetiche*. Il modello di Bohr spiega il fenomeno dell'irradiazione facendolo corrispondere con un cambio di livello di quantizzazione di un elettrone.

Facciamo alcune considerazioni preliminari sulle onde elettromagnetiche. Queste sono formate da perturbazioni ortogonali del campo elettrico (\vec{E}) e magnetico (\vec{B}). Si propagano nello spazio con velocità della luce c , per cui date frequenza f , periodo T e lunghezza d'onda λ si ha:

$$c = \lambda f = \frac{\lambda}{T}$$

Attraverso risultati sperimentali si ha che l'energia associata ad un onda elettromagnetica (e quindi quella liberata con un salto quantico) è direttamente proporzionale alla sua frequenza, cioè:

$$\Delta E = \hbar f = \hbar \frac{c}{\lambda} = E_1 - E_2$$

dove E_1 e E_2 sono i livelli di energia nell'atomo rispettivamente prima e dopo del salto.

Possiamo quindi usare la formula per E trovata prima, prendendo i numeri di quantizzazione n_1 e n_2 corrispondenti ai 2 stati nominati prima:

$$E_1 = \frac{1}{2}m\frac{k^2e^4}{n_1^2\hbar^2}, \quad E_2 = \frac{1}{2}m\frac{k^2e^4}{n_2^2\hbar^2}$$

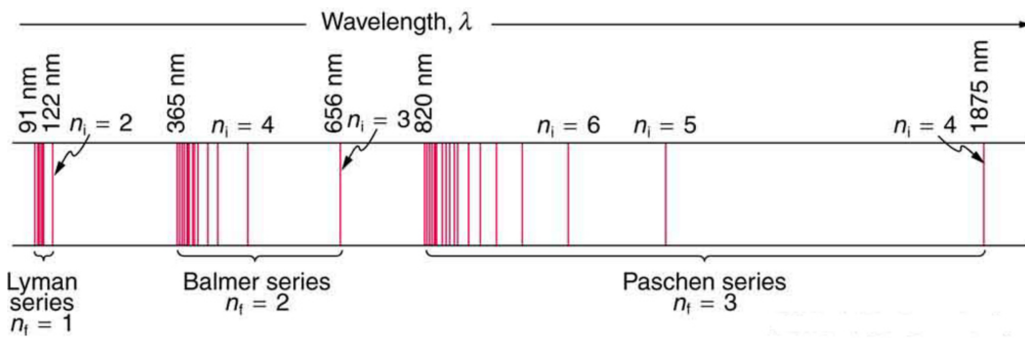
Sostituendo, si può quindi ricavare la relazione (detta equazione di *Rydberg*):

$$\frac{\hbar c}{\lambda} = \frac{1}{2}m\frac{k^2e^4}{n_1^2\hbar^2} - \frac{1}{2}m\frac{k^2e^4}{n_2^2\hbar^2} \implies \frac{1}{\lambda} = \frac{1}{2} \frac{mk^2e^4}{c\hbar^3} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) = R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right)$$

dove la costante $R_H = \frac{1}{2} \frac{mk^2e^4}{c\hbar^3}$ viene detta *costante di Rydberg*.

A partire da questa formula (accurata) si possono ricavare le lunghezze d'onda delle bande di emissione caratteristiche di elementi come l'idrogeno. In particolare, mantenendo n_1 costante e facendo scorrere n_2 da $n_1 + 1$ fino ad infinito, si ottiene uno spettro di emissione associato alla transizione di stato energetico da n_2 a n_1 dell'elettrone dell'idrogeno.

Vediamo queste serie tracciate su un grafico:



Notiamo come le prime serie ricavate dall'equazione di Rydberg hanno dei nomi sulla base del valore di n_1 :

n_1	n_2	Serie	Convergenza
1	$[2, \infty)$	Serie di <i>Lyman</i>	91.13 nm
2	$[3, \infty)$	Serie di <i>Balmer</i>	364.51 nm
3	$[4, \infty)$	Serie di <i>Paschen</i>	820.14 nm

Abbiamo quindi che il modello di *Bohr*, tralasciando la semplificazione delle orbite circolari, è il più accurato che abbiamo visto finora. In questo, ogni atomo individua 4 numeri:

- Il **numero quantico principale** n , che va da 1 al livello energetico che contiene l'elettrone di valenza più esterno dell'atomo (visto che raggio ed energia sono direttamente proporzionali, anche quello ad energia più alta);
- Il **numero quantico secondario** l , che va da 0 a $n - 1$, legato al momento angolare dell'atomo;

- Il **numero quantico magnetico** m_l , che va da $-l$ a l , specifica l'orbitale di appartenenza dell'elettrone;
- Il **numero quantico di spin magnetico** m_s , che può essere $-\frac{1}{2}$ e $\frac{1}{2}$.

Questi 4 numeri da soli specificano un qualsiasi elettrone appartenente all'atomo. Vale il **principio di esclusione** di Pauli, cioè:

1.4: Principio di esclusione di Pauli

Ogni stato elettronico deve avere numeri quantici diversi, o in altre parole più elettroni non possono occupare lo stesso stato.

Orbitale

Vale alla pena notare a questo punto che, come avevamo accennato, il modello basato sulle orbite di Bohr è inaccurato. In verità, vale il **principio di indeterminazione** di Heisenberg:

1.5: Principio di indeterminazione di Heisenberg

Non si può conoscere contemporaneamente sia la posizione che la quantità di moto di un elettrone.

Questo significa che non si può mai effettivamente parlare di *posizioni* di elettroni, ma di *distribuzioni di probabilità* nello spazio.

In particolare, determiniamo *funzioni d'onda* a valore $\in \mathbb{C}$ definite come:

$$\psi(\vec{r}) \in \mathbb{C}, \quad \vec{r} = (x, y, z, t) \in \mathbb{R}^4$$

dove r indica una posizione nello spazio-tempo a cui si valuta la funzione d'onda ψ . La probabilità di trovare un elettrone ad un dato punto \vec{r} sarà quindi:

$$|\psi(\vec{r})|^2$$

Rifacendoci ai numeri quantici nominati nello scorso paragrafo, abbiamo che:

- n , l e m_l , cioè rispettivamente i numeri quantici *principale*, *secondario* e *magnetico* individuano un orbitale.

In particolare, i primi 4 stati del numero l individuano quella che è la *forma* della distribuzione che l'orbitale individua nello spazio, ovvero:

l	Lettera	Forma dell'orbitale
0	s	<i>Sharp</i>
1	p	<i>Principal</i>
2	d	<i>Diffuse</i>
3	f	<i>Fundamental</i>

- m_s , cioè il numero quantico di *spin magnetico*, individua uno stato elettronico all'interno di un dato orbitale. In particolare, riguardo a quanto abbiamo detto prima (m_s può essere $\pm\frac{1}{2}$), si ha che ogni orbitale può essere occupato da al massimo 2 elettroni.

Dualità onda-particella

Abbiamo notato che:

- C'è un quadrato nella formula della probabilità di trovare un dato elettrone alla posizione \vec{r} , cioè $|\psi(\vec{r})|^2$;
- Questo è dato dal fatto che $\psi(\vec{r}) \in \mathbb{C}$, cioè la funzione d'onda, è una funzione complessa. Il quadrato:

$$|\psi(\vec{r})|^2 = \psi(\vec{r})\psi(\vec{r})^*$$

è quindi usato come una sorta di "norma" della funzione d'onda.

Queste considerazioni derivano dalla *dualità onda-particella* degli elettroni: in particolare sono legate al fenomeno dell'*interferenza*. Abbiamo finora trattato gli elettroni come particelle dotate di massa che occupavano un certo spazio. In particolare, dal principio di indeterminazione di Heisenberg (definizione 1.5), avevamo detto che era impossibile conoscere simultaneamente posizione e velocità di un elettrone, e dovevamo affidarci a distribuzioni di probabilità.

In realtà la situazione è più complicata. Si ha infatti che gli elettroni esibiscono un comportamento intermedio fra quello di un'onda e quello di una particella:

- Proprietà simili a un **onda**:
 - Gli elettroni non orbitano un nucleo come pianeti attorno ad una stella, ma invece esistono come *onde stazionarie* sul nucleo. Gli stati energetici minimi sono quindi equivalenti alla frequenza *fondamentale* di una corda, e gli stati energetici superiori alle varie frequenze *armoniche* date dai multipli della fondamentale;
 - Gli elettroni non sono mai in un singolo punto (x, y, z) dello spazio ad un dato istante temporale t , ma la probabilità di *interagire* con un elettrone ad un dato punto può essere ricavata dalla funzione d'onda (orbitale) $\psi(\vec{r})$ con $\vec{r} = (x, y, z, t)$. Come abbiamo visto, la probabilità si ricava prendendo il quadrato della funzione d'onda $|\psi(\vec{r})|^2$.
- Proprietà simili a una **particella**:
 - Il numero di elettroni che orbitano un atomo è sempre un intero;
 - Gli elettroni saltano fra gli orbitali come particelle: ad esempio, se un fotone colpisce un atomo, solo un corrispondente elettrone effettua un salto di energia.

La definizione della funzione d'onda come $\psi(\vec{r}) \in \mathbb{C}$ permette quindi di descrivere il comportamento *ondulatorio* degli elettroni, codificando in un unico numero informazione sia di ampiezza che di fase.

1.2.2 Conduttori, semiconduttori e isolanti

La conoscenza della distribuzione degli elettroni sui possibili orbitali nello stato di minima energia è importante per conoscere le proprietà degli elementi.

In particolare, si possono distinguere 3 categorie principali di elementi riguardo alla conduzione di corrente elettrica:

- **Conduttori**: che offrono elettroni liberi e quindi conducono corrente;

- **Isolanti:** che non offrono elettroni liberi e quindi non conducono corrente;
- **Semiconduttori:** che presentano caratteristiche intermedie fra conduttori e isolanti.

Gli elementi vengono categorizzati sulla base della configurazione dei loro orbitali nella **tavola periodica**. In questo, gli elementi vengono disposti man mano riempiendo i loro orbitali esterni. Essendo questi quelli che determinano le loro proprietà riguardo all'interazione con gli altri elementi, si ha che l'andamento degli elementi è periodico lungo la tavola (da cui il nome).

Riempimento degli orbitali

Prima di riportare il dettaglio della tavola periodica, dettagliamo il principio alla base del riempimento degli orbitali. Di base, vogliamo capire dove si va a disporre ogni nuovo elettrone (con esattezza, in quale orbitale) che introduciamo in un atomo (cioè che incrementiamo il numero atomico Z , che in realtà conterebbe i protoni nel nucleo).

Abbiamo visto nella sezione 1.2.1 cos'è un orbitale e quali sono i numeri che lo individuano (n , l e m_l), e il fatto che un orbitale può essere occupato al massimo da due elettroni con numero magnetico di spin opposto ($m_s = \pm \frac{1}{2}$).

Inoltre, avevamo visto le regole che definivano i range accettabili di n , l , e m_l , che riportiamo in breve:

- n determina i limiti degli altri valori, appartiene a $[1, +\infty)$;
- l appartiene a $[0, n + 1]$;
- m_l appartiene a $[-l, l]$.

Questo significa che per un dato n esiste un numero massimo di stati elettronici dato da:

$$2 \sum_{l=0}^{n-1} (2l + 1) = 2 \sum_{l=0}^{n-1} 2l + n = 4 \sum_{l=0}^{n-1} l + n = 4 \frac{n(n-1)}{2} + n = 2n^2 - n + n = 2n^2$$

dove il 2 deriva dal numero di spin m_s (2 elettroni per orbitale al massimo), la sommatoria da 0 a $n - 1$ scandaglia tutti i valori possibili di l fissato n , e $2l + 1$ scandaglia tutti possibili valori di m fissato n ed l .

Resta da capire come questi orbitali vengono riempiti. In questo ci aiuta il principio di *Aufbau*:

1.6: Principio di Aufbau

Per riempire gli orbitali atomici si procede come segue:

- Si calcola $n + l$, e si comincia da $n + l$ minimo. In caso di $n + l$ uguali, si sceglie prima l' n minore;
- Si riempiono gli m in ordine e in 2 passate, prima riempiendo con un elettrone di spin negativo $-\frac{1}{2}$ e poi con quello di spin positivo $+\frac{1}{2}$.

Il principio di Aufbau dà vita alla comune notazione della configurazione atomica, dove insieme di orbitali con n e l comune viene indicato come:

$$\langle n \rangle \langle l \text{ in forma simbolica} \rangle \langle \text{indice dell'elettrone} \rangle$$

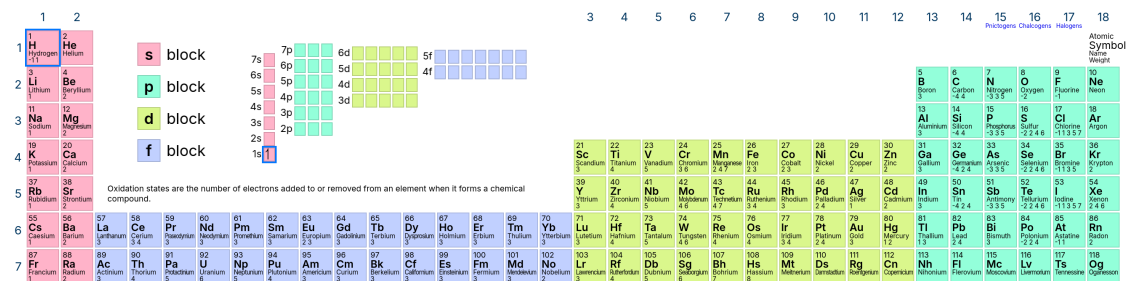
dove l'indice rispetta la formula vista prima riguardo al numero massimo di stati elettronici, e parte da 1. Per rappresentare un atomo si indica quindi la lista dei suoi insiemi di orbitali con n e l comune che hanno almeno un elettrone, rappresentato ad apice. Per sintesi, a volte si omettono le configurazioni di orbitali piene antepoendo il simbolo del primo gas nobile che si trova prima dell'elemento (che sappiamo avere gli orbitali pieni).

Tavola periodica

Abbiamo quindi introdotto come la tavola periodica presenta caratteristiche che cambiano in maniera, appunto, periodica, sulla base del riempimento dello strato di valenza (per quanto ci riguarda l'orbitale più esterno).

Adesso abbiamo visto come questi strati (orbitali) vengono effettivamente riempiti, e possiamo quindi vedere come spostandoci da sinistra verso destra, e dall'alto verso il basso, il numero atomico Z sale.

Riportiamo quindi una versione della tavola periodica che mette in evidenza la configurazione elettronica degli elementi:



Con i colori sono stati indicati gli orbitali che particolari elementi stanno riempiendo. Notiamo che i "salti" che incontriamo spostandoci verso destra sono dati dal principio di Aufbau, e in particolare dal fatto che scegliamo sempre l' n minore (a volte ci troviamo a fare dei "salti" all'indietro per riempire un orbitale più interno lasciato vuoto).

Possiamo implementare uno script in Python per applicare il principio di Aufbau (1.1) e ricavare le configurazioni elettroniche di ogni elemento dato il numero atomico Z :

```

1 # trova la configurazione elettronica dell'elemento di numero atomico z
2 def aufbau(z):
3     # genera tutte le combinazioni (n, l)
4     nls = [
5         (n, l)
6         for n in range(1, MAX_N + 1)
7         for l in range(0, n)
8     ]
9
10    # ordina le combinazioni secondo il principio di Aufbau
11    nls.sort(key=lambda x: (x[0] + x[1], x[0]))
12
13    # calcola gli stati elettronici per ogni combinazione
14    states = [
15        (n, l, m_l, m_s)
16        for n, l in nls
17        for m_l in range(-l, l + 1)
18        for m_s in [-0.5, 0.5]
19    ]
20

```

```

21 # restituisci i primi z stati elettronici
22 for i, elec in enumerate(states):
23     if i >= z:
24         break
25
26     yield elec

```

Uno script che usa questa funzione si può trovare negli appunti a `code/python/orbital.py`. Un esempio di esecuzione con $z = 14$ (numero atomico del silicio) dà il seguente risultato:

n	l	m_l	m_e
1	0	0	-0.5
1	0	0	0.5
2	0	0	-0.5
2	0	0	0.5
2	1	-1	-0.5
2	1	-1	0.5
2	1	0	-0.5
2	1	0	0.5
2	1	1	-0.5
2	1	1	0.5
3	0	0	-0.5
3	0	0	0.5
3	1	-1	-0.5
3	1	-1	0.5

Regola dell'ottetto

Una volta chiara la struttura della tavola periodica, possiamo interrogarci riguardo a regole empiriche come la regola dell'ottetto riguardo agli *elettroni di valenza*.

Di base, abbiamo che i primi 2 elementi (idrogeno ed elio, $1s^1$ e $1s^2$) riempiono il loro stato di valenza riempiendo l'orbitale s (con $n = 1$, l è fisso a 0). Passando ai prossimi 2 periodi, abbiamo che si passa a riempire anche l'orbitale p , per cui si hanno:

$$2 \cdot 2^2 - 2 = 8, \quad 2 \cdot 3^2 - 2 - 8 = 8$$

cioè 8 elettroni di valenza, da cui la regola dell'ottetto per il riempimento dello stato di valenza.

Proseguendo ai prossimi periodi si entra nel cosiddetto blocco d , contenente i *metalli di transizione*. Questi sono elementi che stanno riempiendo il loro orbitali d , e di cui quindi anche gli elettroni degli orbitali d partecipano allo stato di valenza. Per questi notiamo che la lunghezza dei periodi è 18 (non facciamo calcoli ma lo ricaviamo dalla tavola periodica e generalmente dal principio di Aufbau, 1.1). Questo si allinea col fatto che alcuni elementi (come ad esempio il ferro, $[\text{Ar}]4s^23d^6$) non obbediscono più alla regola dell'ottetto ma alla regola dei *18 elettroni*.

Drogaggio del silicio

Questo può essere un buon momento per accennare al drogaggio del silicio. Abbiamo che questo è di norma un cattivo semiconduttore (quindi ha poca capacità di condurre corrente, o più precisamente di fornire elettroni liberi).

Si procede quindi a introdurre impurità controllate per modificare le proprietà elettriche del materiale. Il boro viene utilizzato per il drogaggio di tipo p , creando lacune po-

sitive, mentre il fosforo o l'arsenico vengono usati per il drogaggio di tipo n , aumentando il numero di elettroni liberi.

1.2.3 Legami chimici

In natura, in condizioni ottimali, solo i gas nobili si trovano allo stato atomico mentre tutti gli altri elementi si presentano legati ad altri in svariate modalità.

Dal punto di vista chimico, i legami interatomici consistono nella formazione di una configurazione ad energia minima che di solito richiede la presenza nel guscio elettronico di 8 elettroni (regola dell'*ottetto*).

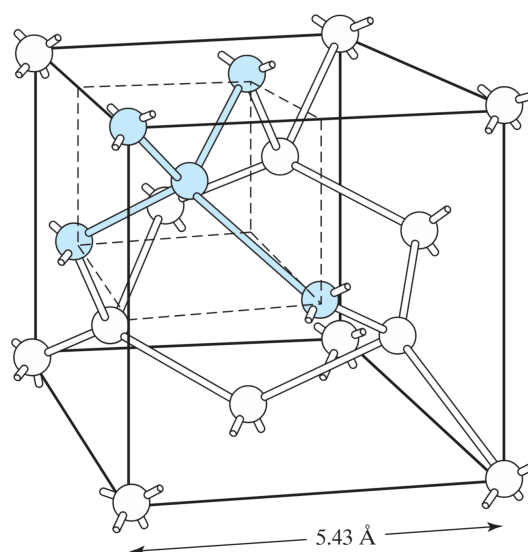
Esistono 3 tipi di legami chimici:

- Legame **ionico**: dove c'è un effettivo trasferimento di elettroni, solitamente finalizzato al completamento dell'ottetto. Per questo motivo accade quando un atomo ha il guscio quasi pieno e l'altro ha il guscio quasi vuoto. A questo punto l'attrazione fra gli atomi accade per la differenza di carica fra i 2 atomi;
- Legame **covalente**: ha luogo quando vengono messi a comune degli elettroni. Per questo accade quando 2 atomi hanno il guscio quasi pieno. Ci è di particolare interesse in quanto è il legame che forma il silicio;
- Legame **metallico**: è il legame tipico dei metalli, dove tutti gli elettroni vengono messi a comune e possono circolare liberamente. Questo rende i metalli dei conduttori particolarmente efficaci.

Legame del silicio

Possiamo usare quanto studiato finora per definire la struttura dei legami formati dal silicio. Abbiamo che nello stato base ci sono 4 elettroni di valenza (configurazione elettronica $[\text{Ne}]3s^23p^2$), per cui si possono formare legami covalenti fra un atomo di silicio e i 4 atomi adiacenti.

Questa configurazione porta alla formazione di una struttura cristallina tetraedrica, dove ogni atomo lega con 4 atomi adiacenti:



2 Lezione del 03-03-26

2.1 Conduzione

Iniziamo a dettagliare il fenomeno della conduzione attraverso un materiale. Supponiamo di avere un parallelepipedo di sezione S e lunghezza L .

2.1: Seconda legge di Ohm

Dalla seconda legge di *Ohm*, la sua abilità di condurre cariche di un corpo è inversamente proporzionale alla **resistenza**:

$$R = \rho \frac{L}{S}$$

dove ρ è una costante specifica al materiale detta *resistività*. Ricordiamo che questo deriva dalla formulazione propria di quella che chiameremo legge di *Ohm microscopica*:

$$J = \sigma E$$

dove J è la *densità di corrente* e σ è la *conducibilità* di un materiale, per cui:

$$I = JS = \sigma ES = \frac{1}{\rho} \frac{V}{L} S = \frac{V}{R}$$

dove si è usato la definizione di *potenziale elettrico* $V = EL$ sulla distanza L .

Abbiamo quindi ricavato la legge:

2.2: Prima legge di Ohm

Differenza di potenziale, corrente e resistenza sono legate dalla relazione:

$$V = IR$$

Riguardo a resistività e conducibilità, le cui unità di misura sono rispettivamente $\Omega \cdot \text{cm}$ e $\Omega \cdot \text{cm}^{-1}$, vale la relazione:

$$\rho = \frac{1}{\sigma}$$

immediata dalla definizione a partire dalla legge di Ohm.

2.1.1 Conduttori, semiconduttori e isolanti con Ohm

Sulla base della resistività si possono classificare i materiali in *conduttori*, *semiconduttori* e *isolanti* come accennato in 1.2.2.

In particolare, si riporta la tabella che individua i valori di resistività per diversi tipi di materiale:

Tipo	Resistività
Conduttore	$\rho < 10^{-2} \Omega \cdot \text{cm}$
Isolante	$\rho > 10^5 \Omega \cdot \text{cm}$
Semiconduttore	$10^{-2} \Omega \cdot \text{cm} < \rho < 10^5 \Omega \cdot \text{cm}$

Il vantaggio dei materiali semiconduttori, su cui noi ci concentriamo (in particolare noi assumiamo di usare il *silicio*, simbolo atomico Si e numero atomico 14) è che la loro

resistività può essere controllata in maniera tale da decidere se in una data occasione un semiconduttore deve comportarsi come un isolante o come un conduttore (come abbiamo accennato, il processo di *drogaggio* di 1.2.2).

2.1.2 Corrente di deriva in un conduttore

La **corrente di deriva** è quella che otteniamo applicando un campo elettrico \vec{E} ad un materiale.

Riprendiamo il parallelepipedo di 2.1, e assumiamo che sia composto da un materiale conduttore ($\rho < 10^{-2} \Omega \cdot \text{cm}$).

Prendiamo la definizione di corrente come quantità di carica che si sposta in una superficie nell'istante di tempo, cioè:

$$I = \frac{\Delta Q}{\Delta T}$$

Questo per noi significherà prendere una sezione del parallelepipedo e "contare" il numero di cariche che vi passano attraverso.

Modello di Drude-Lorentz

Il modello di *Drude-Lorentz* si pone di studiare il fenomeno della conduzione considerando quindi queste cariche libere all'interno di un metallo (prendiamo ad esempio il rame, che con configurazione elettronica $[\text{Ar}]3d^{10}4s^1$ offre un elettrone libero per atomo) come particelle di un gas perfetto.

- Se non si applica nessun campo \vec{E} , la corrente di deriva che passa attraverso il parallelepipedo sarà chiaramente 0. Notiamo che questo non deriva dal fatto che le cariche sono ferme, ma che si stanno muovendo in direzione casuale (e quindi statisticamente attraverso la nostra superficie entrano tante cariche quante ne escono).
- Procediamo quindi ad applicare un certo campo \vec{E} . Dalla definizione di campo elettrico, si ha che le cariche vengono sottoposte ad una forza (e rispettiva accelerazione) data da:

$$\vec{F} = -e\vec{E}, \quad \vec{a} = -\frac{e\vec{E}}{m}$$

Le traiettorie descritte dagli elettroni sono casuali, localmente in moto rettilineo uniforme finché questi non urtano altri elettroni. All'istante dell'urto modellizziamo la variazione di velocità (trascurando la forza relativa all'urto stesso) come:

$$\vec{v}_{i+1} = \vec{v}_i - \frac{e\vec{E}}{m}\tau$$

dove τ è il tempo necessario ad effettuare un *cammino libero medio* all'interno del conduttore.

Si ha quindi che la **velocità di deriva** media sarà data da:

$$\vec{v}_d = \langle \vec{v}_{i+1} \rangle = -\frac{e\vec{E}}{m}\tau$$

Dal punto di vista macroscopico, quello che avremo è che la velocità media è legata ad un parametro μ_m detto *mobilità*, e direttamente proporzionale al campo:

$$\vec{v}_d = -\mu_m \vec{E}$$

dove ricordiamo il segno negativo, in quanto gli elettroni si muovono nella direzione opposta al campo (sono cariche negative $-1 e$).

Da Drude-Lorentz a Ohm

Vediamo come riagganciarci alle leggi di Ohm a partire dal modello di Drude-Lorentz appena visto.

Osserviamo il percorso di una singola carica nel nostro materiale conduttivo, e prendiamo il tempo che impiega a passare da un capo all'altro del materiale come:

$$\Delta T = \frac{L}{\vec{v}_d}$$

detto *tempo di osservazione*. Riprendendo la definizione di corrente, e supponendo che nel materiale siano presenti N cariche elettroniche, potremo quindi dire:

$$I = \frac{\Delta Q}{\Delta T} = \frac{eN}{\Delta T} = \frac{eN}{L} v_d$$

Ci possiamo riportare anche alla densità di corrente J ponendo:

$$J = \frac{I}{S} = \frac{eNv_d}{LS} = \frac{eNv_d}{V_{ol}}$$

dove $V_{ol} = LS$ è il *volume* del parallelepipedo. Chiamiamo quindi n la *concentrazione* di carica all'interno del volume:

$$n = \frac{N}{V_{ol}}$$

in modo da poter dare quello che sarà il nostro nodo chiave tra il modello di Ohm e quello di Drude-Lorentz:

2.1: Legge microscopica di Ohm

La densità di carica in un materiale sottoposto ad un campo elettrico di modulo E è:

$$J = \frac{N}{V_{ol}} e v_d = n e v_d = n e \mu_n E$$

trascurando i segni (cioè assumendo che il campo sia perpendicolare al parallelepipedo conduttore considerato da principio).

Dal punto di vista vettoriale, riprendiamo la definizione della legge di Ohm, stavolta contando i segni:

$$\vec{J} = \sigma \vec{E}, \quad \sigma = n(-e) \cdot (-\mu_n) = n e \mu_n$$

dove si applica sempre la relazione con la resistività:

$$\rho = \frac{1}{\sigma} = \frac{1}{n e \mu_n}$$

Notiamo che appunti più accurati e tratti da un corso di fisica su questo argomento si possono trovare a <https://raw.githubusercontent.com/seggiani-luca/appunti-fis/68a83a26dacd1f545dcf697da0daabf257413325/master/master.pdf>. In particolare, negli appunti riportati si derivano completamente resistività e conduttività nel contesto del modello di Drude-Lorentz, oltre alla legge microscopica di Ohm (pagina 149).

In particolare, si mantiene la mobilità degli elettroni μ_e in funzione delle caratteristiche del materiale e del tempo di cammino libero medio τ :

$$\vec{v}_d = -\frac{e\vec{E}}{m}\tau = -\mu_m\vec{E} \implies \mu_m = \frac{e\tau}{m}$$

che sostituito nei nostri risultati dà la stessa cosa:

$$\rho = \frac{m}{ne^2\tau}, \quad \sigma = \frac{1}{\rho} = \frac{ne^2\tau}{m}$$

Abbiamo quindi derivato le leggi di Ohm a partire dalla densità di carica, e abbiamo visto come questa necessita solamente della conoscenza delle cariche libere in un materiale e la loro velocità di deriva.

Il modello di Drude-Lorentz ci permette quindi di stimare tale velocità di deriva, a partire da una visione delle particelle basata sulla meccanica classica e sulla forza di Lorentz applicata alle cariche.

Esempio numerico col rame

Fissiamo i concetti visti nell'ultima sezione con un esempio numerico. Poniamo di voler calcolare, a partire dal modello di Drude-Lorentz, la resistività del rame. Avevamo la formula:

$$\rho = \frac{m}{ne^2\tau}$$

e del rame conosciamo i valori:

Quantità	Simbolo	Valore
Massa dell'elettrone	m	$9.11 \times 10^{-31} \text{ kg}$
Densità atomica del rame	n	$8.47 \times 10^{28} \frac{\text{atomi}}{\text{m}^3}$
Carica dell'elettrone	e	$1.6 \times 10^{-19} \text{ C}$
Tempo di cammino libero medio	τ	$2.5 \times 10^{-14} \text{ s}$

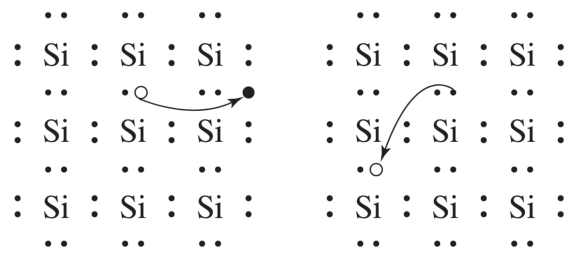
per cui:

$$\rho = \frac{9.11 \times 10^{-31}}{8.47 \times 10^{28} \cdot (1.6 \times 10^{-19})^2 \cdot 2.5 \times 10^{-14}} = 1.68 \times 10^{-8} \Omega \cdot \text{m}$$

che si allinea coi risultati sperimentali e classifica il rame come un ottimo conduttore, secondo quanto detto in 2.1.1.

2.1.3 Corrente di deriva in un semiconduttore

Negli scorsi esempi riguardo alla corrente di deriva abbiamo fatto uso di un materiale conduttore (in particolare del rame). In un semiconduttore, la situazione cambia in quanto i legami che si formano fra gli atomi non sono più metallici.



Prendiamo ad esempio il silicio, che forma legami di tipo covalente. Abbiamo detto in 1.2.3 che il silicio presenta 4 elettroni nello strato di valenza, per cui è capace di formare legami covalenti con altri 4 atomi di silicio in una struttura cristallina tridimensionale. Gli elettroni nel reticolo cristallino del silicio saranno quindi costretti a trovarsi nella zona del legame fra gli atomi, e non potremo usare l'approssimazione a gas perfetto valida per i metalli. La chiave sarà quindi data dalla **temperatura**.

- Di base, in una situazione dove la temperatura è $T = 0$ K gradi Kelvin, gli elettroni non sono liberi di muoversi nel reticolo e non si ha conduzione;
- Se invece si pone $T \neq 0$ K, si può avere conduzione attraverso gli elettroni liberi che vengono "portati" via dal legame. Questo è reso possibile dal fatto che l'energia fornita come aumento di temperatura all'atomo è tale da superare l'energia di legame, e quindi liberare un elettrone coinvolto nel legame.

Elettroni liberi in un semiconduttore

Avremo quindi bisogno di un modello per il numero di elettroni liberi in funzione della temperatura.

2.3: Statistica degli elettroni liberi in un semiconduttore intrinseco

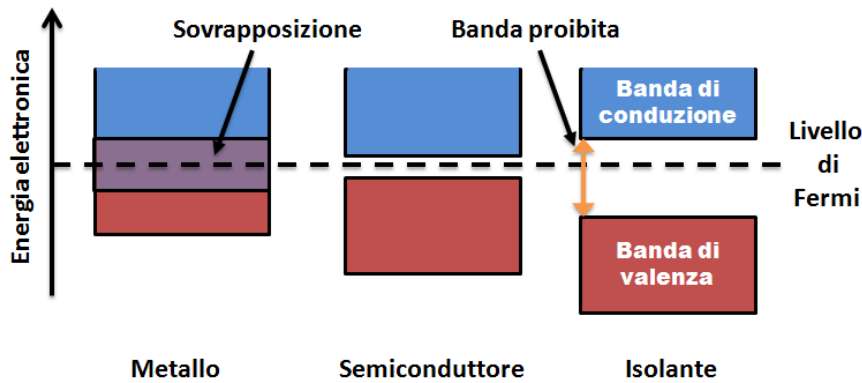
La densità di elettroni liberi in un semiconduttore intrinseco in funzione della temperatura T è data da:

$$n_i = BT^{\frac{3}{2}} e^{-\frac{E_G}{2k_B T}}$$

Ad esempio, abbiamo che a 300 K gradi Kelvin, il numero di elettroni liberi nel silicio è circa 10^{10} cm^{-3} .

Nella legge riportata n_i è la densità di elettroni liberi (analoga alla n dei conduttori), B è una costante dipendente dai materiali, E_G è l'intervallo di banda fra le cosiddette bande di *valenza* e *conduzione*, e k_B è una costante detta costante di Boltzmann.

Il modello che adottiamo per la statistica deriva dalla *teoria delle bande*, dove si ipotizza che in un sistema di particelle come quello del silicio intrinseco ad ognuna può essere associato un livello energetico (non necessariamente corrispondente ai livelli energetici degli elettroni negli orbitali). Questo livello energetico non può essere conosciuto per ogni particella, ma si possono fare delle *statistiche*. Riguardo all'intervallo di banda, abbiamo che tutti gli elementi presentano un livello energetico detto *livello di Fermi* necessario a fornire o sottrarre elettroni. Due bande, dette appunto banda di valenza e banda di conduzione, si trovano rispettivamente sotto e sopra tale livello.



- Nei **conduttori**, la banda di valenza e la banda di conduzione si sovrappongono, o comunque è presente un *intervallo di banda* E_G (cioè la distanza fra le due bande) molto piccolo;
- Negli **isolanti**, le bande di valenze e la banda di conduzione sono molto distanti (l'intervallo di banda E_G è consistente);
- Nei **semiconduttori** l'intervallo di banda E_G è intermedio fra conduttore e isolante: come abbiamo visto, compare nella formula per la densità di elettroni liberi n_i , dove crescendo diminuisce n_i in maniera esponenziale.

Capire il numero di portatori liberi equivale quindi a determinare una statistica per i livelli energetici e capire la probabilità che questi vadano a trovarsi nella zona di conduzione.

- La statistica applicabile in questo caso sarebbe quella di *Fermi-Dirac*, che in breve è:

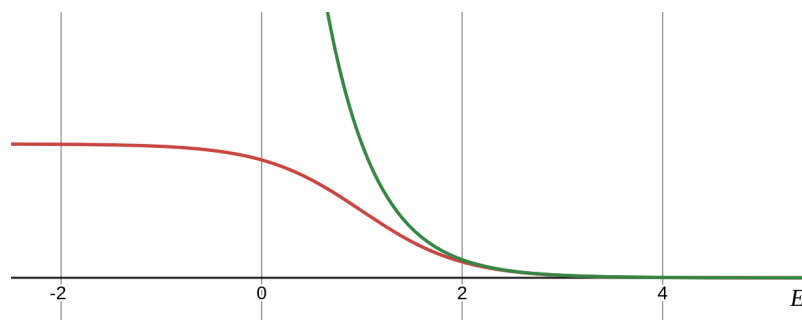
$$f_{\text{Fermi-Dirac}}(E) = \frac{1}{e^{\frac{E-E_F}{k_B T}} + 1}$$

Questa restituisce la probabilità che una particella si trovi allo stato energetico E (abbia energia E) rispetto al livello di Fermi E_F ;

- Nel caso $E - E_F \gg k_B T$ si può semplificare adottando la statistica di *Maxwell-Boltzmann*. Questa è una statistica adatta a particelle a temperatura più alta, basata sulla meccanica classica:

$$f_{\text{Maxwell-Boltzmann}}(E) = e^{-\frac{E-E_F}{k_B T}}$$

Vediamo su un grafico la differenza fra le statistiche $f_{\text{Fermi-Dirac}}(E)$ e $f_{\text{Maxwell-Boltzmann}}(E)$, mettendo in evidenza la similarità per grandi valori di E :



Per ricavare la formula in 2.2, quindi, prendiamo come livello $E = E_C$ alla base della banda di conduttività (questo è valido in quanto la maggior parte dei conduttori si concentra a tale livello). Se si ha che l'intervallo E_G è più o meno centrato sul livello di Fermi, seguirà che:

$$E_C - E_F = \frac{E_G}{2}$$

che è ciò che troviamo nella formula. Infine, per il termine $BT^{\frac{3}{2}}$, abbiamo che questo deriva sempre dalla statistica di Maxwell-Boltzmann. In particolare, abbiamo che il termine $f_{\text{Maxwell-Boltzmann}}(E)$ indica la *probabilità* di stati, mentre si dimostra che la *densità* effettiva di stati va esponenzialmente (di fattore $\frac{3}{2}$) con la temperatura. Il prodotto fra questi darà la densità di particelle nel livello energetico di conduzione.

Diciamo quindi che il silicio è un semiconduttore **intrinseco** in quanto ha naturalmente bassa conduttività. Di un semiconduttore drogato per ottenere maggiore conduttività si dice che è invece **estrinseco**.

Vedremo che in diversi casi, trattando di conduttori estrinseci (per noi principalmente il silicio drogato) stabiliremo relazioni matematiche che faranno riferimento al comportamento di tali semiconduttori in condizioni intrinseche. In questo, n_i si riferisce alla densità di elettroni liberi (come vedremo fra poco, di *coppie lacuna-elettrone*) nel silicio intrinseco (che obbedisce a $n = p = n_i$ e vale circa 10^{10} cm^{-3} a 300 Kelvin).

Generazione e ricombinazione termica

Vediamo le condizioni di equilibrio del numero di elettroni liberi in un semiconduttore.

- Il fenomeno visto finora prende il nome di **generazione termica** di elettroni liberi, in quanto chiaramente richiede di fornire calore al materiale semiconduttore le cui proprietà di conduzione vogliamo variare, in modo che questi si spostino nella fascia di conduzione;
- Si può osservare anche il fenomeno opposto, detto di **ricombinazione termica**, quando un elettrone perde energia e decade nuovamente, dalla fascia di conduzione, in un legame covalente.

Si ha che la quantità effettiva di elettroni liberi deriva da entrambi i fenomeni (che ricordiamo essere entrambi dipendenti esclusivamente dalla temperatura), per cui a grandi linee:

$$n_i = \text{generazione termica} + \text{ricombinazione termica}$$

cioè due processi attivi, uno di generazione, e uno di ricombinazione, sono in corso in un semiconduttore. Chiaramente, il primo porta alla formazione di più elettroni liberi, mentre il secondo annichilisce quelli già formati. Se il nostro obiettivo è quello di migliorare le proprietà di conduzione di un semiconduttore, preferiremo il primo.

Quanto abbiamo detto prima riguardo alle statistiche sui livelli energetiche vale quando ci si trova in condizioni di equilibrio dei processi di generazione e ricombinazione termiche. Questi sono infatti sempre in corso, e quando si stabilizzano le statistiche (e il numero di portatori liberi) diventa costante. In particolare, se chiamiamo G la frequenza di generazione termica, e R la frequenza di ricombinazione termica, si trova una condizione di equilibrio a:

$$G = R$$

che per il silicio ci dà il valore di 10^{10} cm^{-3} a 300 Kelvin.

Lacune elettroniche

Un fenomeno che non abbiamo visto è quello comportato dal fatto che un elettrone che viene liberato da un legame forma una **lacuna** o *vacuità* nel reticolo cristallino, cioè una regione di "vuoto" della carica. Le lacune possono spostarsi come gli elettroni, in quanto semplicemente può cambiare il legame su cui la lacuna si trova. In questo modo può essere più semplice considerare le lacune come cariche positive ($+1e$) che si spostano all'interno del reticolo cristallino al pari degli elettroni.

In verità, le lacune elettroniche nei semiconduttori sono importanti quanto gli elettroni, e abbiamo quindi 2 possibili trasportatori di carica:

- Gli **elettroni**, che hanno carica $-1e$. Ne indichiamo la densità con n ;
- Le **lacune**, che hanno carica $+1e$. Ne indichiamo la densità con p .

Abbiamo quindi l'immagine completa del fenomeno:

- La *generazione termica* dirà che, fornito calore ad un atomo di silicio, elettroni possono liberarsi dai loro legami covalenti spostandosi nella fascia di conduzione, e formando allo stesso tempo una lacuna elettronica. In un conduttore intrinseco, quindi, le coppie ($-1e, +1e$) di elettroni e lacune vanno di paripasso;
- Allo stesso modo, la *ricombinazione termica* sarà il fenomeno di ricadimento degli elettroni dalla fascia di conduzione ad uno dei legami covalenti della struttura cristallina del silicio. Si ha quindi che la coppia ($-1e, +1e$) di elettrone e lacuna viene annichilita.

Notiamo quindi che, secondo quanto detto sull'equilibrio di formazione di coppie elettrone/lacuna, e il fatto che queste in tale equilibrio sono sempre in coppia, riguardo alle densità di elettroni liberi e lacune elettroniche in un semiconduttore intrinseco vale:

$$n = p = n_i$$

Possiamo quindi riportarci alla legge di Ohm, dando una definizione della conduttività di un elemento semiconduttore basata su elettroni e lacune:

$$\sigma = n\mu_n e + p\mu_p e$$

dove μ_p è la *mobilità delle lacune*, che notiamo differisce da quella degli elettroni (già introdotta come μ_n). In generale, la differenza fra la mobilità dell'elettrone e quella delle lacune sarà quella che ci permetterà di variare la conduttività agendo sull'equilibrio stesso di elettroni e lacune, secondo il processo già accennato di cosiddetto *drogaggio*.

3 Lezione del 05-03-26

3.1 Richiami di Elettrotecnica

Dedichiamo questa lezione ad alcuni richiami di conoscenze passate. Notiamo che appunti più dettagliati di elettrotecnica possono trovarsi a <https://raw.githubusercontent.com/seggiani-luca/appunti-et/e63e7e9ac9d57bd7b3f690140c8b460d91e0b08e/master/master.pdf>.

Qui faremo solo un breve ripasso degli argomenti più importanti, cioè:

- Leggi di **Ohm** (già viste in 2.1);
- Leggi di **Kirchoff** sull'analisi circuitale;
- Teoremi di **Thevenin** e **Norton**;
- Componenti circuitali passivi, cioè *resistenze*, *capacità* ed *induttanze*.

Leggi di Ohm

In particolare, in 2.1 abbiamo ritrovato le leggi di Ohm, che riportiamo qui per completezza:

$$v = Ri, \quad i = Gv \quad (1^a \text{ legge di Ohm})$$

$$R = \frac{1}{G} = \rho \frac{l}{A}, \quad (2^a \text{ legge di Ohm})$$

Queste stabiliscono che la tensione v ai capi di un conduttore è proporzionale alla corrente i che percorre tale conduttore, di una costante R , detta *resistenza* (o il suo inverso G , detto *conduttanza*).

Leggi di Kirchoff

Vediamo quindi le leggi di analisi circuitale, a partire da:

3.1: Prima legge di Kirchoff (ai nodi)

La somma algebrica (con segno) delle correnti entranti in un qualsiasi nodo di un circuito è uguale a zero, ergo:

$$\sum_{n=1}^N i_n = 0$$

Per questa come sempre ricordiamo che le correnti hanno segno, per cui è importante prima stabilire i versi di tutte le correnti e quindi applicare la legge per evitare di incorrere in errori di segno.

Vediamo quindi la seconda legge:

3.2: Seconda legge di Kirchoff (alle maglie)

La somma algebrica delle cadute di tensione lungo un qualunque percorso chiuso di un circuito (detto *maglia*) è uguale a zero, ergo:

$$\sum_{n=1}^N v_n = 0$$

Spesso, in circuiti su cui sono presenti generatori, questo significa che la somma dei generatori di tensione è uguale alla caduta di potenziale sui terminali degli altri componenti. In altre parole: "La fem dei generatori è uguale alla somma algebrica delle cadute di tensione negli elementi circuitali passivi", ergo:

$$\sum_{i=0}^N E_i = \sum_{j=0}^M i R_j$$

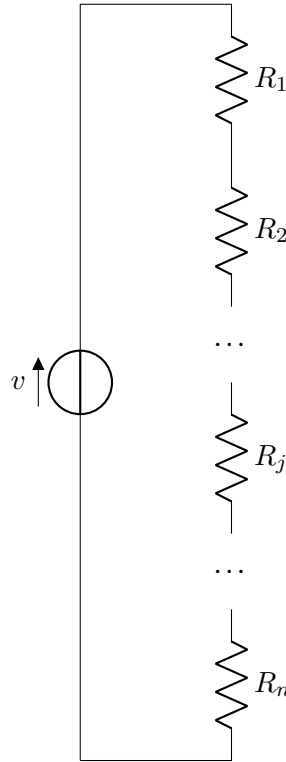
per generatori E_i e resistenze R_j su una maglia percorsa da corrente i .

Partitore di tensione

Abbiamo che la resistenza equivalente di resistenze serie è banalmente:

$$R_{eq} = R_1 + R_2 + \dots + R_j + \dots + R_n$$

Possiamo usare questa legge per studiare il *partitore di tensione*:



per cui la tensione ai capi di una resistenza in serie sarà data da, applicando la legge di Ohm (2.1):

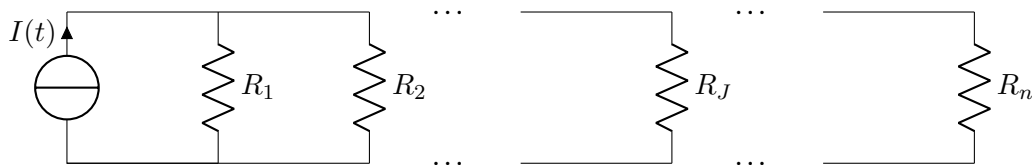
$$\begin{cases} i = \frac{v}{R_1+R_2+\dots}, & v_1 = iR_1 \implies v_1 = v \frac{R_1}{R_1+R_2+\dots} \\ i = \frac{v}{R_1+R_2+\dots}, & v_2 = iR_2 \implies v_2 = v \frac{R_2}{R_1+R_2+\dots} \\ \dots \end{cases}$$

Partitore di corrente

Per le resistenze in parallelo anziché le resistenze, sommiamo le conduttanze:

$$G_{eq} = G_1 + G_2 + \dots + G_j + \dots + G_n$$

Possiamo usare questa legge per studiare il *partitore di corrente*:



per cui la corrente che scorre su ogni resistenza sarà data da, sempre applicando la legge di Ohm:

$$\begin{cases} i = v(G_1 + G_2 + \dots), & i_1 = vG_1 \implies i_1 = i \frac{G_1}{G_1+G_2+\dots} \\ i = v(G_1 + G_2 + \dots), & i_2 = vG_2 \implies i_2 = i \frac{G_2}{G_1+G_2+\dots} \\ \dots \end{cases}$$

Ricordiamo quindi cosa significa effettivamente prendere la somma delle conduttanze:

$$G_{eq} = G_1 + G_2 = \frac{1}{R_1} + \frac{1}{R_2}, \quad R_{eq} = \left(\frac{1}{R_1} + \frac{1}{R_2} \right) = \frac{R_1 R_2}{R_1 + R_2}$$

ciò significa che il partitore di corrente si può prendere anche come:

$$i_j = i \frac{G_j}{\sum_i^N G_i} = i \frac{\prod_{i \neq j}^N R_i}{\sum_i^N R_i}$$

cioè basta prendere i prodotti di tutte le resistenze non coinvolte nel ramo su cui si calcola la corrente, al numeratore.

Principio di sovrapposizione degli effetti

Vediamo questa legge, che si applica a reti *lineari*:

3.3: Principio di sovrapposizione degli effetti

La risposta di una rete lineare che contiene più generatori indipendenti può essere ricavata considerando un singolo generatore per volta e sommando tutte le risposte così ottenute.

Facendo attenzione al fatto che parliamo di generatori *indipendenti* (i generatori dipendenti fanno eccezione).

Per considerare un singolo generatore per volta dobbiamo *disattivare* gli altri generatori. Ricordiamo che questo significa:

- Trasformare i generatori di *corrente* in **circuiti aperti**;
- Trasformare i generatori di *tensione* in **cortocircuiti**.

Notiamo che le reti che vedremo non sempre saranno lineari: in tal caso questa legge non si potrà applicare, e dovrà al limite essere applicata a versioni *linearizzate* di tali reti.

Teorema di Thevenin e Norton

Ricordiamo questi due teoremi:

3.4: Teorema di Thevenin

Un circuito lineare, rispetto ad una coppia dei suoi nodi, può essere trasformato in una combinazione di un generatore di tensione v_T e una resistenza R_{eq} .

In particolare, la tensione v è quella vista dai due terminali se si separa la rete dal resto del circuito e si mette fra loro un circuito aperto. La resistenza R è semplicemente la resistenza vista fra i 2 morsetti, disattivati i generatori indipendenti. Notiamo che nel caso ci siano generatori dipendenti, questi vanno mantenuti e bisogna introdurre un generatore di prova (di corrente o di tensione, si calcolerà la resistenza *vista* misurando l'altra grandezza e mettendola a rapporto).

3.5: Teorema di Norton

Un circuito lineare, rispetto ad una coppia dei suoi nodi, può essere trasformato in una combinazione di un generatore di corrente i_N e una resistenza R_{eq} .

In particolare, la tensione i è quella vista dai due terminali se si separa la rete dal resto del circuito e si mette fra loro un cortocircuito. La resistenza R_{eq} è analoga fra Thevenin e Norton.

Un corollario dei teoremi di Thevenin e Norton è che un generatore di corrente in parallelo a una resistenza equivale ad un generatore di tensione in serie alla stessa resistenza (salvo casi limite).

Capacità ed induttanze

Introduciamo altri due componenti circuitali, cioè le **capacità** e le **induttanze**.

- Definiamo la *capacità* come l'elemento passivo descritto dalla legge:

$$q = Cv$$

dove C viene detta **capacità** e si misura in F (*Farad*). Dal punto di vista temporale abbiamo quindi:

$$i = \frac{dq}{dt} = C \frac{dv}{dt}, \quad v = \frac{1}{C} \int_{t_0}^t i d\tau + v(t_0)$$

In DC, quindi, la capacità si comporta come un *circuito aperto*, e la *tensione* non può variare istantaneamente (porta energia).

Riguardo alle capacità in *serie*, abbiamo che queste si comportano nella maniera opposta alle resistenze, cioè:

$$C_{eq} = \left(\frac{1}{C_1} + \frac{1}{C_2} + \dots \right)^{-1}$$

mentre per le capacità in *parallelo* abbiamo:

$$C_{eq} = C_1 + C_2 + \dots$$

- Definiamo invece l'*induttanza* come l'elemento passivo descritto dalla legge:

$$\phi = Li$$

dove L viene detta **induttanza** e si misura in H (*Henry*). ϕ è invece il flusso magnetico. Dal punto di vista temporale abbiamo quindi:

$$v = \frac{d\phi}{dt} = L \frac{di}{dt}, \quad i = \frac{1}{L} \int_{t_0}^t v d\tau + i(t_0)$$

In DC, quindi, l'induttanza si comporta come un *cortocircuito*, e la *corrente* non può variare istantaneamente (porta energia).

Riguardo alle induttanze in *serie*, abbiamo che queste si comportano come le resistenze, cioè:

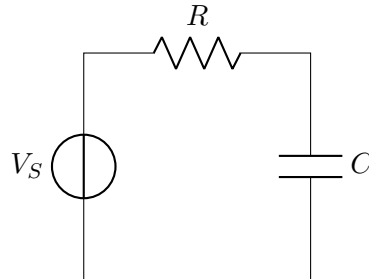
$$L_{eq} = L_1 + L_2 + \dots$$

mentre per le induttanze in *parallelo* abbiamo:

$$L_{eq} = \left(\frac{1}{L_1} + \frac{1}{L_2} + \dots \right)^{-1}$$

3.1.1 Transitorio del condensatore

Dopo aver dato le basi di elettrotecnica, vediamo la situazione in cui ci troviamo quando chiudiamo un circuito con una resistenza e una capacità (il celebre circuito **RC**).



Possiamo studiare questo circuito in una varietà di modi. Scegliamo il più semplice, partendo dall'impostare l'equazione di maglia:

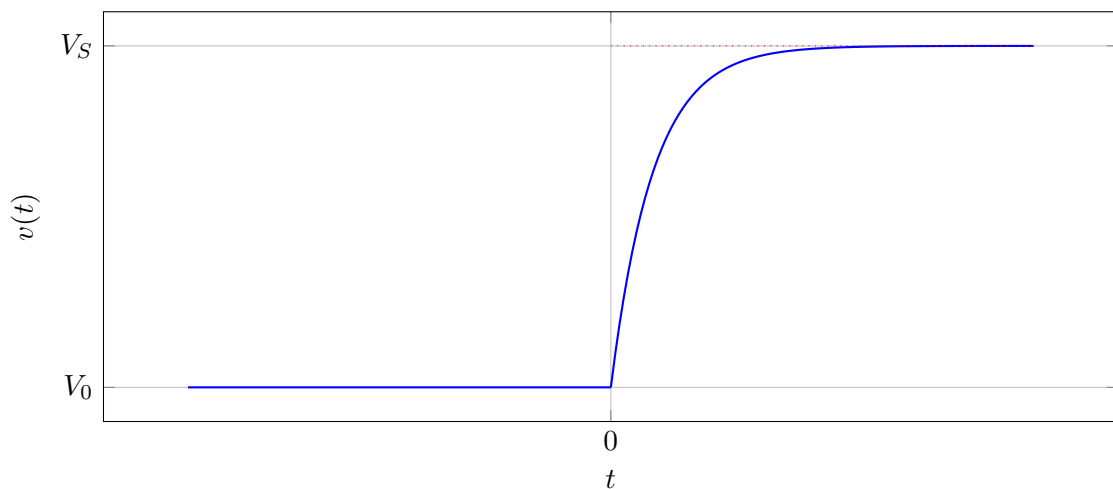
$$\frac{V_s}{R} = -\frac{v}{R} + C \frac{dv}{dt} \implies \frac{V_s - v}{R} = C \frac{dv}{dt} \implies \frac{dv}{v - V_s} = -\frac{dt}{RC}$$

separando le variabili. Risolviamo quindi integrando da entrambi i lati:

$$\ln(v - V_s) \Big|_{V_0}^{V(t)} = -\frac{t}{RC} \Big|_0^t \implies v(t) = V_s + (V_0 - V_s)e^{-\frac{t}{\tau}}$$

con $\tau = RC$ detto *tempo caratteristico* del circuito.

Questo dà la solita figura di crescita esponenziale della grandezza v , che riportiamo sul grafico:



dove le varie quantità sono:

- V_0 : tensione *iniziale* presente sul condensatore;
- V_S : tensione *finale* presente sul condensatore, cioè quella espressa dal generatore;
- τ : *tempo caratteristico* del circuito, determinato da RC .

4 Lezione del 06-03-26

Alla fine della lezione 2 abbiamo trattato il *silicio intrinseco*. In particolare, avevamo definito i fenomeni di:

- **Generazione termica:** sottoposto a calore, il silicio assorbe energia che rompe i legami portando alla formazione di coppie *elettrone/lacuna* (e/h , da h di *hole*, o come abbiamo detto in altri contesti coppie $(-1e, +1e)$);
- **Ricombinazione termica:** in tandem alla generazione termica, si ha un fenomeno di ricombinazione degli elettroni nei legami del silicio, che annichilisce le coppie elettrone/lacuna.

A conti fatti si aveva che il numero di portatori di carica (elettroni o lacune) disponibili alla conduzione era:

$$n = p = n_i = \text{generazione termica} + \text{ricombinazione termica}$$

in condizioni di equilibrio con le frequenze di generazione termica G e di ricombinazione termica R :

$$G = R$$

e la conduttività era quindi:

$$\sigma = n\mu_n e + p\mu_p e$$

che si otteneva a partire dall'espressione della velocità di deriva in funzione della mobilità elettronica μ_n o di lacuna μ_p , e dalla legge microscopica di Ohm (legge 2.1).

Stima della conduttanza intrinseca

Ricordiamo che a 300 K gradi Kelvin, il numero di elettroni liberi nel silicio è circa 10^{10} cm^{-3} . Prendendo i valori di mobilità elettronica e di lacuna del silicio:

Quantità	Simbolo	Valore
Mobilità elettronica	μ_n	$1350 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$
Mobilità delle lacune	μ_p	$480 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$

e svolgendo i calcoli si ha:

$$\sigma = 10^{10} \cdot 1350 \cdot (1.6 \times 10^{-19}) + 10^{10} \cdot 480 \cdot (1.6 \times 10^{-19}) = 3 \times 10^{-6} \Omega \cdot \text{cm}^{-1}$$

che di base rende il silicio intrinseco un cattivo conduttore.

4.1 Drogaggio del silicio

Possiamo quindi approfondire il concetto del **drogaggio** del silicio introdotto in 1.2.2. Di base, la legge che ci risulterà utile per valutare i portatori di carica di un conduttore non intrinseco sarà quella che lega il numero di portatori di carica negativi e positivi (elettroni e lacune) nel semiconduttore estrinseco alla concentrazione di elettroni (o ugualmente, lacune) nel corrispondente intrinseco:

4.1: Legge azione di massa

Il numero di portatori di carica negativi n e positivi p in un semiconduttore rispetta la relazione:

$$n \cdot p = n_i^2$$

dove n_i è la concentrazione di elettroni liberi nel silicio intrinseco.

Questo è in opposizione alla legge che valeva nei semiconduttori intrinseci, dove parlare di p o n era indifferente in quanto:

$$n = p = n_i$$

Quello che succede è che si può portare il silicio a formare legami con altri elementi, in particolare:

- Con quelli del quinto gruppo della tavola periodica (fra cui il fosforo, simbolo [P]), e l'arsenico, simbolo [As]), che chiamiamo drogaggio di tipo **n** (perché come vedremo aumenta il numero di portatori negativi, cioè elettroni);
- Con quelli del terzo gruppo della tavola periodica (fra cui il boro, simbolo [B]), che chiamiamo drogaggio di tipo **p** (perché come vedremo, in maniera duale al precedente, aumenta il numero di portatori positivi, cioè lacune).

Togliamoci anche la curiosità di come si svolge il processo di drogaggio del silicio dal punto di vista tecnico:

- Abbiamo che un primo metodo sfrutta il processo della *diffusione* (che vedremo nel dettaglio in 4.2 riguardo alla corrente di diffusione, ma qui agisce in un contesto chimico diverso). Questo è il processo più diffuso in laboratorio. In particolare, si stende un *film* di materiale drogante sul substrato di silicio, che viene poi riscaldato e quindi (appunto per *diffusione*) va a legare col substrato.
- Su scala industriale, invece, si mira ad usare un processo di *impiantazione ionica*. Questo metodo consiste nel bombardare il substrato con ioni di materiale drogante, forzando quindi in qualche modo i legami di atomi donatori o accettori. Notiamo che questo processo rovina il materiale, per cui in seguito si necessita di un processo di *annealing* termico che permetta la ricostruzione della struttura cristallina.

4.1.1 Drogaggio di tipo N

Vediamo cosa accade nel drogaggio con atomi del quinto gruppo. Ad esempio, prendiamo in considerazione il fosforo, che ha configurazione elettronica $[\text{Ne}]3s^23p^3$, e quindi 5 elettroni liberi nello strato di valenza.

Ciò che andiamo a fare è quindi sostituire una certa percentuale degli atomi di silicio nel reticolo cristallino con atomi di fosforo. Questi mettono i loro 4 elettroni al servizio dei legami covalenti con i 4 atomi di silicio adiacenti. Il quinto elettrone resta molto debolmente legato al nucleo di fosforo, per cui la maggior parte degli atomi ionizzano a temperatura ambiente (finora abbiamo parlato di $T = 300 \text{ K}$).

Chiamiamo N_D la concentrazione di tali atomi, detti *donatori*. Solitamente possiamo dire:

$$N_D = [10^{14}, 10^{21}] \text{ cm}^{-3}$$

dove notiamo che se $N_D = 10^{21}$ la concentrazione al centimetro cubo è quasi uguale alla mole, cioè abbiamo quasi interamente sostituito gli atomi di silicio.

Abbiamo quindi che in condizione di regime del silicio drogato con un elemento del quinto gruppo, i processi di *generazione* e *ricombinazione* termica avvengono comunque, ma parte integrante degli elettroni liberi vengono forniti dagli atomi donatori. In particolare, vale la stima:

$$n \approx N_D \implies p \approx \frac{n_i^2}{n} = \frac{n_i^2}{N_D}$$

visto che come abbiamo detto la maggior parte degli atomi donatori N_D ionizza.

Diciamo che gli n sono conduttori **maggioritari** (gli elettroni, crescono all'aumentare del drogaggio), mentre gli p sono conduttori **minoritari** (le lacune, diminuiscono all'aumentare del drogaggio).

Stima della conduttanza con drogaggio di tipo N

Prendiamo un valore $N_D = 10^{15} \text{ cm}^{-3}$ e calcoliamo nuovamente il valore della conduttanza. Notiamo che a seguito del processo di drogaggio i valori di mobilità elettronica e di lacuna cambiano, come vedremo in 4.1.3:

Quantità	Simbolo	Valore
Mobilità elettronica	μ_n	$1320 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$
Mobilità delle lacune	μ_p	$460 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$

Avremo quindi bisogno di ricavare n e p con le relazioni appena trovate:

$$n = N_D = 10^{15} \text{ cm}^{-3}, \quad p = \frac{(10^{10})^2}{10^{15}} = 10^5 \text{ cm}^{-3}$$

Si ha quindi che la conduttanza, riprendendo la formula dello scorso esempio, sarà:

$$\sigma = 10^{15} \cdot 1320 \cdot (1.6 \times 10^{-19}) + 10^5 \cdot 460 \cdot (1.6 \times 10^{-19}) = 0.211 \Omega \cdot \text{cm}^{-1}$$

cioè introducendo un doping del:

$$\frac{10^{15}}{10^{22}} \sim 10^{-7}$$

ovvero variando un atomo di silicio su 10^7 con un atomo del quinto gruppo, siamo riusciti ad ottenere un aumento della conduttanza di ben 5 ordini di grandezza.

4.1.2 Drogaggio di tipo P

Vediamo quindi cosa accade con drogaggio con un elemento del terzo gruppo. In questo caso prendiamo l'esempio del boro, che ha configurazione elettronica $[\text{He}]2s^22p^1$, e quindi 3 elettroni liberi nello strato di valenza.

Se procediamo col drogaggio al boro, quindi sostituendo alcuni atomi di silicio con atomi di boro, ciò che accade è che i 3 elettroni di valenza vanno a formare legami covalenti nel reticolo cristallino del silicio. Il quarto legame non può essere formato (in quanto quell'elettrone non c'è), per cui si forma naturalmente una lacuna elettronica. Anzi, abbiamo che gli atomi di silicio stessi ionizzano, cedendo elettroni all'atomo di boro (che diventa quindi ione negativo).

In questo caso abbiamo quindi la densità N_A di atomi, detti *accettori*. Solitamente possiamo dire:

$$N_A = [10^{14}, 10^{21}] \text{ cm}^{-3}$$

Abbiamo quindi che in condizione di regime del silicio drogato con un elemento del terzo gruppo, i conduttori maggioritari sono le lacune (in maniera totalmente duale al drogaggio del tipo N). In particolare, vale la stima:

$$p \approx N_A \implies n \approx \frac{n_i^2}{p} = \frac{n_i^2}{N_A}$$

data dal fatto che quasi ogni atomo di boro introdotto ionizza.

Notiamo che stavolta gli p sono conduttori **maggioritari** (le lacune, crescono all'aumentare del drogaggio), mentre gli n sono conduttori **minoritari** (gli elettroni, diminuiscono all'aumentare del drogaggio).

Stima della conduttanza con drogaggio di tipo P

Prendiamo un valore $N_A = 10^{15} \text{ cm}^{-3}$ e calcoliamo nuovamente il valore della conduttanza. Anche qui, a seguito del processo di drogaggio i valori di mobilità elettronica e di lacuna cambiano:

Quantità	Simbolo	Valore
Mobilità elettronica	μ_n	$1350 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$
Mobilità delle lacune	μ_p	$450 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$

Avremo quindi bisogno di ricavare n e p con le relazioni appena trovate:

$$n = \frac{(10^{10})^2}{10^{15}} = 10^5 \text{ cm}^{-3}, \quad p = N_A = 10^{15} \text{ cm}^{-3}$$

Si ha quindi che la conduttanza, riprendendo la formula dello scorso esempio, sarà:

$$\sigma = 10^5 \cdot 1350 \cdot (1.6 \times 10^{-19}) + 10^{15} \cdot 450 \cdot (1.6 \times 10^{-19}) = 0.072 \Omega \cdot \text{cm}^{-1}$$

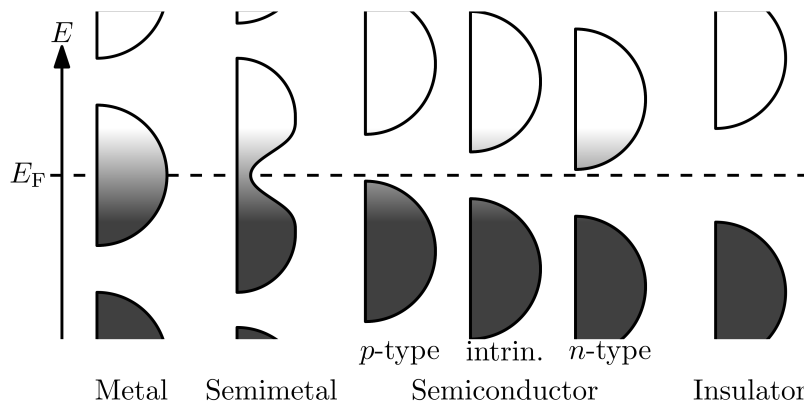
cioè introducendo un doping del:

$$\frac{10^{15}}{10^{22}} \sim 10^{-7}$$

ovvero variando un atomo di silicio su 10^7 con un atomo del terzo gruppo, siamo riusciti ad ottenere un aumento della conduttanza di 4 ordini di grandezza. Questo non è buono quanto quello che eravamo riusciti ad ottenere col drogaggio al fosforo (visto che la mobilità di lacuna è molto minore della mobilità elettronica), ma è comunque valido.

Legami alla teoria delle bande

In 2.1.3, parlando della corrente di deriva nei semiconduttori (intrinseci, prima del drogaggio, che abbiamo introdotto adesso) avevamo accennato alla *teoria delle bande*. L'idea fondamentale di questa era di considerare il reticolo cristallino di un semiconduttore, con elettroni e lacune libere, come un sistema di particelle dotate di stati energetici N . In particolare avevamo detto che il numero di portatori liberi in un semiconduttore era legato alla statistica di Maxwell-Boltzmann, usata per stimare la densità di conduttori sulla banda di conduzione (legge 2.2). Abbiamo che variando la composizione del reticolo (attraverso il drogaggio), si vanno a variare le caratteristiche di tali particelle e quindi le statistiche (seppur leggermente).



Vediamo adesso che l'effetto del drogaggio a livello di teoria delle bande è che il livello di Fermi (che avevamo visto nei semiconduttori si andava a trovare all'interno dell'intervallo energetico E_G fra la banda di valenza e la banda di conduzione) si sposta verso l'alto o verso il basso:

- Nel caso di drogaggio di tipo **p**, il livello di Fermi si sposta in basso: questo significa che gli elettroni diventano più "difficili" da eccitare, e si ha maggioranza di lacune;
- Nel caso di drogaggio di tipo **n**, il livello di Fermi si sposta in alto. Questo significa che si gli elettroni diventano più "facili" da eccitare, e si ha maggioranza elettronica.

Drogaggi multipli

Notiamo che si può effettuare un processo di drogaggio di tipo sia N che P. In particolare, se su un substrato di tipo N si fa un forte drogaggio di tipo P, effettivamente il tipo di drogaggio viene invertito. Viceversa, se su un substrato di tipo P si fa un forte drogaggio di tipo N, nuovamente il drogaggio viene invertito.

Cioè, in sostanza, si può ottenere silicio estrinseco di tipo P, N, o silicio intrinseco, sulla base del tipo di drogaggio che viene fatto anche in più passate sul substrato.

4.1.3 Stima delle mobilità

Riprendiamo la formula per la conducibilità:

$$\sigma = n\mu_n e + p\mu_p e = f(T)$$

Qui ricordiamo che compaiono:

- La *mobilità elettronica* dei portatori di carica negativi, cioè gli elettroni;
- La *mobilità di lacuna* dei portatori di carica positivi, cioè le lacune. Queste sarebbero legami che "saltano" da un atomo all'altro, che però possiamo semplicemente considerare come elettroni "positivi" che si spostano nel reticolo del semiconduttore.

Abbiamo che σ stessa è dipendente dalla temperatura T sulla base della variazione delle concentrazioni n e p . Si scopre che anche μ_n e μ_p sono dipendenti da T , e in particolare:

- μ_n, μ_p : diminuiscono all'aumentare di T (in particolare sono proporzionali a $T^{-\frac{3}{2}}$);
- n, p : qui dobbiamo considerare i vari casi di drogaggio o meno:
 - Per quanto riguarda il silicio **intrinseco**, vale:

$$n = p = n_i$$

e quindi n, p aumentano esponenzialmente come n_i (rispettando la legge 2.4). Questo significa che in media σ aumenta con T (in quanto le concentrazioni n, p aumentano più di quanto diminuiscono le mobilità μ_n, μ_p);

- Nel caso del silicio drogato, la situazione cambia. Consideriamo prima il caso del drogaggio di tipo **N**. Valgono le stime riguardo alle concentrazioni:

$$n \approx N_D, \quad p \approx \frac{n_i^2}{N_D}$$

cioè n resta costante sulla concentrazione di atomi donatori, mentre p aumenta con T . Notiamo che aumenta *poco*, in quanto di base p rappresenta la concentrazione di conduttori *minoritari* (le lacune). In questo caso, considerando anche la diminuzione di mobilità sulla base di T , si ha che complessivamente la conduttività *diminuisce* con T .

- Vediamo quindi il drogaggio di tipo **B**. Valgono le stime riguardo alle concentrazioni:

$$p \approx N_A, \quad n \approx \frac{n_i^2}{N_A}$$

cioè p resta costante sulla concentrazione di atomi accettori, mentre n aumenta con T . Notiamo che aumenta *poco*, in quanto di base n rappresenta la concentrazione di conduttori *minoritari* (gli elettroni). Come prima, considerando anche la diminuzione di mobilità sulla base di T , si ha che anche qui la conduttività *diminuisce* con T .

Notiamo che quanto detto sul silicio drogato vale per temperature alte *moderate*: a temperature T abbastanza alte si ha che il processo di generazione e ricombinazione termica (che ricordiamo era governato dalla statistica Maxwelliana a 2.3, esponenziale alla temperatura) supera la riduzione di mobilità. In questo nuovo regime, che diciamo regime *intrinseco* del silicio drogato, si ha che la conduttività ricomincia a salire.

4.2 Corrente di diffusione

Abbiamo quindi concluso la trattazione delle correnti di *deriva* nei conduttori e nei semiconduttori. Esiste un altro fenomeno di conduzione, particolarmente osservabile nei semiconduttori, detto **diffusione**. Il fenomeno della diffusione porta alla formazione della *corrente di diffusione*.

In generale, la diffusione è un meccanismo che avviene ogni volta che si ha un *gradiente di concentrazione* di un dato portatore (di carica, un gas, ecc...). Per via dell'*agitazione termica*, il sistema tende naturalmente ad uno stato di equilibrio dove la concentrazione è uguale in ogni punto dello spazio. Dinamicamente, questo significa che le zone di *alta* concentrazione diffondono verso le zone di *bassa* concentrazione. Le leggi che descrivono questi fenomeni sono le leggi di *Fick* sulle diffusioni di concentrazioni. Semplificando un po' si ha in sostanza:

4.2: Legge di Fick

Il flusso di una concentrazione ϕ è legato al suo gradiente come:

$$\vec{J} = -D\nabla\phi$$

dove D è la *costante di diffusione* di ϕ .

Un fenomeno di questo tipo avviene anche per gli elettroni e le lacune in un semiconduttore. Supponiamo di avere un corpo esteso sull'asse x , ed un certo gradiente di concentrazione degli elettroni n che varia sulla base di x (cioè va come una certa funzione

$n(x)$). Diciamo che $n(x)$ ha un picco per $x = 0$, cioè ha una forma del tipo:

$$n(x) = e^{-x}, \quad x \geq 0$$

In questo caso vicino a $x = 0$ si avrà una zona di alta generazione e ricombinazione termica. Questo porterà ad un fenomeno di diffusione che, appunto, diffonderà portatori di carica (elettroni) verso $x \gg 0$. Notiamo che questo porta alla formazione di una corrente (in quanto ci sono cariche che si spostano), nella direzione opposta a quella di diffusione. Ad esempio, se si ha una diffusione di elettroni e verso *destra*, si forma una densità di corrente \vec{J} rivolta verso sinistra.

Diffusione elettronica

Cerchiamo di calcolare questa densità di corrente per i soli elettroni:

$$\vec{J} = -(-e)D_n \left(\frac{\partial n}{\partial x} \right) \hat{i}_x = eD_n \frac{\partial n}{\partial x} \hat{i}_x$$

dove D_n è la costante di diffusione vista nella legge di Fick, una quantità analoga alla *mobilità* considerata per la corrente di deriva e dipendente dalla deriva data dal fenomeno di diffusione. Il segno meno alla derivata parziale, come abbiamo detto, è dato dal fatto che gli elettroni si spostano in direzione opposta alla corrente. Questa legge deriva direttamente dalla legge di Fick (4.2), dove $\frac{\partial n}{\partial x}$ non è altro che la versione bidimensionale del gradiente $\nabla\phi$ (in questo caso ∇n).

Diffusione di lacune

Possiamo calcolare la corrente di diffusione anche considerando solo i portatori positivi (le lacune):

$$\vec{J} = -eD_p \left(\frac{\partial n}{\partial x} \right) \hat{i}_x = -eD_p \frac{\partial n}{\partial x} \hat{i}_x$$

dove D_p è la costante di diffusione di lacuna (per distinguerla dalla costante di diffusione elettronica D_n). Chiaramente, stavolta le cariche si prendono come positive.

4.2.1 Stima delle costanti di diffusione

Potremmo voler stimolare le costanti di diffusione, o almeno legarle a costanti che conosciamo (magari le mobilità elettroniche e di lacuna μ_n, μ_p). Per questo ci torna utile la seguente legge:

4.3: Legge di Einstein delle costanti di diffusione

Riguardo alle costanti di *diffusione* e le *mobilità* di semiconduttori, possiamo applicare la stima:

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{k_B T}{e} = V_T$$

dove K_B è la costante di *Boltzmann* già vista nella legge 2.1, e V_T è la cosiddetta *tensione termica*, che a $300K$ vale circa $V_T = 26 mV$. Non diamo l'origine di questa formula (deriva dagli studi di Einstein sul *moto browniano*), ma vediamo che la proporzione:

$$D_n : \mu_n = D_p : \mu_p$$

sembra ragionevole, in quanto particelle più mobili saranno più disposte a diffondere.

4.2.2 Combinazione di correnti di drift e di diffusione

Completiamo quindi il nostro modello combinando le densità di corrente di *drift* e di *diffusione*, che chiamiamo rispettivamente \vec{J}_{Drift} e $\vec{J}_{\text{Diffusione}}$.

- Prendendo le sole cariche elettroniche, cioè i portatori negativi, si ha:

$$\vec{J}_N = \vec{J}_{N,\text{Drift}} + \vec{J}_{N,\text{Diffusione}} = ne\mu_n\vec{E} + eD_n\frac{\partial n}{\partial x}\hat{i}_x$$

ricavato dalla legge microscopica di Ohm al primo termine (legge 2.3), e la legge di Fick (4.2) al secondo.

- Prendendo le sole cariche positive (le lacune elettroniche), cioè i portatori positivi, si ha:

$$\vec{J}_P = \vec{J}_{P,\text{Drift}} + \vec{J}_{P,\text{Diffusione}} = pe\mu_p\vec{E} - eD_P\frac{\partial n}{\partial x}\hat{i}_x$$

in maniera totalmente analoga alla scorsa derivazione.

La densità di corrente complessiva sarà quindi quella ottenuta combinando i nostri modelli:

$$\vec{J}_N + \vec{J}_P = \vec{J}_{N,\text{Drift}} + \vec{J}_{N,\text{Diffusione}} + \vec{J}_{P,\text{Drift}} + \vec{J}_{P,\text{Diffusione}} = ne\mu_n\vec{E} + eD_n\frac{\partial n}{\partial x}\hat{i}_x + pe\mu_p\vec{E} - eD_P\frac{\partial n}{\partial x}\hat{i}_x$$

Tutte queste caratteristiche, quindi (drift e diffusione) vanno tenute in considerazione per stimare in maniera corretta la densità di corrente, in particolare in un semiconduttore.

5 Lezione del 10-03-26

5.1 Giunzione P/N

Fenomeni elettrici interessanti accadono quando si collegano 2 materiali semiconduttori (silicio) drogati con un processo rispettivamente di tipo P e di tipo N.

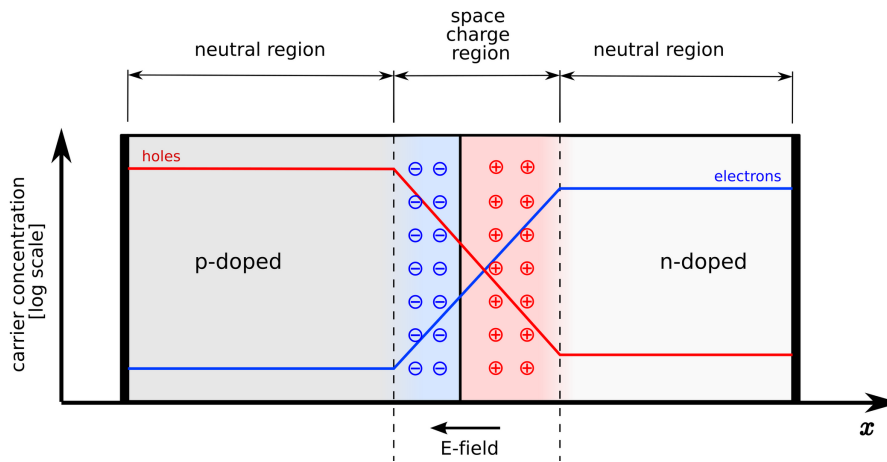
Supponiamo quindi di avere un corpo rettangolare diviso in due regioni (una di tipo P e una di tipo N), collegato ad un anodo A dal lato P e ad un catodo K dal lato N.

- Ricordiamo che il silicio di tipo P è un silicio drogato con atomi accettori di elettroni (quindi del terzo gruppo, ad esempio boro). Immaginiamo in questo caso di avere una concentrazione di $N_A = 10^{17} \text{ cm}^{-3}$ atomi accettori:

$$p_P = N_A = 10^{17} \text{ cm}^{-3}, \quad n_P = \frac{n_i^2}{N_A} = \frac{10^{20}}{10^{17}} = 10^3 \text{ cm}^{-3}$$

- Il silicio di tipo N è invece un silicio drogato con atomi donatori di elettroni (quindi del quinto gruppo, ad esempio fosforo o arsenico). Immaginiamo in questo caso di avere una concentrazione di $N_D = 10^{16} \text{ cm}^{-3}$ atomi donatori:

$$n_N = N_D = 10^{16} \text{ cm}^{-3}, \quad p_N = \frac{n_i^2}{N_D} = \frac{10^{20}}{10^{16}} = 10^4 \text{ cm}^{-3}$$



In questo caso abbiamo che le concentrazioni p e n di lacune ed elettroni liberi rispettivamente variano lungo l'asse x del corpo, ed hanno un punto di incrocio nel punto centrale (chiamiamolo ϕ). La corrente totale \vec{I} che passa nel punto ϕ sarà quindi nulla, come lo sarà la densità di corrente \vec{J} .

Corrente di diffusione

Ciò che accadrà in una configurazione di questo tipo è però che le lacune si sposteranno da sinistra verso destra (in quanto c'è una maggiore concentrazione di lacune piuttosto che di elettroni liberi, $10^{17} > 10^{16}$).

- Avevamo da 4.2 la legge di Fick della diffusione, per cui la densità di corrente di diffusione delle lacune sarà:

$$\vec{J}_{P,Diffusione} = -eD_P \frac{\partial p_{\hat{x}}}{\partial x}$$

- Allo stesso modo, la corrente di diffusione degli elettroni liberi sarà:

$$\vec{J}_{N,Diffusione} = +eD_N \frac{\partial n_{\hat{x}}}{\partial x}$$

Da quanto detto rispetto alla maggiore concentrazione di lacune piuttosto che di elettroni liberi, si avrà per forza che:

$$\vec{J}_{P,Diffusione} + \vec{J}_{N,Diffusione} \neq 0$$

Questo è apparentemente assurdo, in quanto avevamo imposto $\vec{J} = 0$. Dovrà quindi esistere una corrente di deriva che va a bilanciare le correnti di diffusione.

Corrente di deriva

Perchè esista una corrente di deriva deve esistere un campo elettrico. Ciò che accade è quindi che si va a formare un campo elettrico \vec{E} interno alla giunzione, che forma le correnti di deriva che bilanciano le correnti di diffusione (in modo che nel complessivo si obbedisca a $\vec{J} = 0$).

Ingrandiamo quindi la zona della giunzione fra P ed N.

- Nella regione P avremo una serie di ioni *accettori* fissi. In condizioni di equilibrio (quindi prima di collegare la regione N), ad ognuno di questi sarà associata una lacuna libera di drogaggio, per cui la carica complessiva sarà neutra;

- Allo stesso modo nella regione N avremo una serie di ioni *donatori* fissi. In condizioni di equilibrio (quindi prima di collegare la regione P), ad ognuno di questi sarà associato un elettrone libero di drogaggio, per cui la carica complessiva sarà anche qui neutra.

Nel momento in cui si va a formare la giunzione, avremo che le lacune nella regione P vengono a spostarsi nella regione N ed a ricombinarsi con gli elettroni liberi (in verità sono gli ioni di silicio negativi che forniscono elettroni agli ioni di silicio positivi della regione P, ma continuiamo a considerare le lacune come particelle). Allo stesso modo, gli elettroni liberi nella regione N si spostano nella regione P ed a ricombinarsi con le lacune libere.

Questo processo si svolge a livello di atomi di silicio: gli ioni fissi (accettori e donatori rispettivamente nelle regioni P ed N) resteranno tali, ma la loro carica non sarà più bilanciata, per cui si genereranno 2 regioni con carica non neutra. In particolare:

- Nella regione P, a maggioranza di ioni *accettori* fissi, resteranno questi (carichi negativamente) e quindi si avrà carica *negativa*;
- Nella regione N, a maggioranza di ioni *donatori* fissi, resteranno questi (carichi positivamente) e quindi si avrà carica *positiva*.

Zona di svuotamento

Indichiamo con w la lunghezza della combinazione delle 2 regioni così formate. Chiamiamo tale regione **zona di svuotamento**. All'interno della zona di svuotamento si andrà a formare un campo \vec{E} . Da quanto abbiamo detto riguardo alla corrente complessiva e alle correnti di diffusione, dovrà essere che la situazione di *equilibrio* del campo \vec{E} è quella che porta alla formazione di correnti di deriva bilanciano tali correnti di diffusione. In simboli si parte dal dire:

$$\vec{J} = \vec{J}_P + \vec{J}_N = 0 \implies \begin{cases} \vec{J}_P = 0 \\ \vec{J}_N = 0 \end{cases}$$

Questo in quanto non si può avere accumulazione di lacune nella regione N (che è quello che accadrebbe con $\vec{J}_P > 0$), e allo stesso modo non si può avere accumulazione di elettroni nella regione P (che è quello che accadrebbe con $\vec{J}_N > 0$). Resta che l'unica possibilità è che siano entrambe 0.

Possiamo quindi porre separatamente:

$$\begin{cases} \vec{J}_N = +eD_N \frac{\partial n}{\partial x} \hat{i}_x + ne\mu_N \vec{E} = 0 \\ \vec{J}_P = -eD_P \frac{\partial p}{\partial x} \hat{i}_x + pe\mu_P \vec{E} = 0 \end{cases}$$

e quindi risolvere le 2 equazioni ottenute per ottenere 2 equazioni (entrambe valide) per il campo che si va a formare nella zona di svuotamento:

$$\begin{cases} E = -\frac{D_n}{n\mu_n} \frac{dn}{dx} \\ E = +\frac{D_p}{p\mu_p} \frac{dp}{dx} \end{cases}$$

dove ci siamo riportati a scalari prendendo tutti i vettori paralleli. Possiamo quindi prendere la legge di Einstein delle costanti di diffusione vista a 4.3, che in particolare diceva:

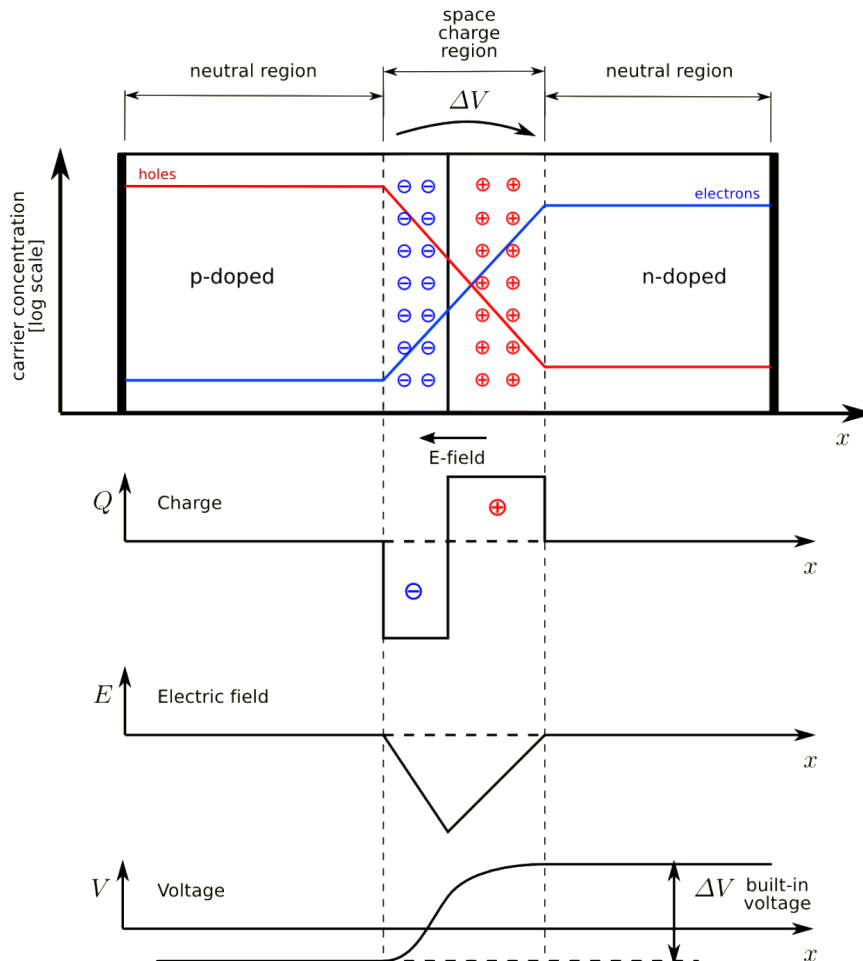
$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{k_B T}{e}$$

Usiamo tale stima per esprimere diversamente i rapporti costante di diffusione / mobilità:

$$\begin{cases} E = -\frac{k_B T}{e} \frac{1}{n} \frac{dn}{dx} \\ E = +\frac{k_B T}{e} \frac{1}{p} \frac{dp}{dx} \end{cases}$$

che è la formula del campo all'interno della zona di svuotamento. Vedremo come queste definizioni ci torneranno utili a breve, nel calcolo della *barriera di tensione*.

Descrizione qualitativa del campo



Vediamo quindi la descrizione qualitativa del campo che si va a formare. Prima di tutto, abbiamo che nella zona di svuotamento si avranno una prima regione negativa e quindi una regione positiva (chiamiamo $Q(x)$ la funzione di carica). Se l'equilibrio è rispettato, dovrà essere che l'integrale della carica nella zona negativa dovrà essere uguale a quello nella zona positiva. Questo va specificato in quanto, visto che le lacune sono più degli elettroni, si ha che la barriera ϕ fra le due regione è leggermente spostata verso la regione P.

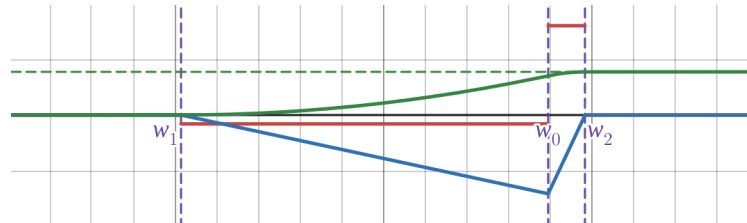
Una volta nota la carica $Q(x)$, possiamo calcolare il campo $E(x)$ applicando la definizione come:

$$\frac{dE(x)}{dx} = -\frac{Q(x)}{S\epsilon_0}$$

dove S è la sezione della giunzione e ϵ_0 è la permittività dello spazio. Come vediamo, questo significherà che abbiamo un picco di campo elettrico \vec{E} (negativo, dalla zona N

alla zona P) nel punto ϕ (abbiamo che in verità). Ai limiti della zona di svuotamento andranno considerati gli opportuni effetti di bordo.

Per curiosità, un modello matematico di questa situazione (basato anche sui calcoli della sezione che segue) è stato sviluppato a <https://www.desmos.com/calculator/hhrxntwkb1>.



Nella figura abbiamo il risultato del modello per la regione di svuotamento w , dove w_1 è l'inizio della regione, w_2 la fine, e w_0 è la barriera all'interno della zona w fra regione P e regione N (sostanzialmente la nostra ϕ). Abbiamo:

- In *rosso* la carica $Q(x)$ all'interno della giunzione, da cui si nota che si hanno quantità uguali di carica a sinistra e a destra di w_0 , in quanto:

$$\begin{cases} q_P = -q w_P N_D A = -q(w_0 - w_1) N_D A \\ q_N = q w_N N_A A = q(w_2 - w_0) N_A A \end{cases}$$

e quindi valgono le relazioni:

$$\frac{w_P}{N_A} = \frac{w_D}{N_D}, \quad w_P + w_D = (w_0 - w_1) + (w_2 - w_0) = w_2 - w_1 = w$$

cioè le lunghezze delle regioni sono tali per cui viene occupata l'intera larghezza w , con bias verso la regione a densità maggiore, e le quantità di carica sull'unità di lunghezza rispondono in maniera da mantenere la carica da entrambi i lati uguale;

- In *verde* il campo $E(x)$ all'interno della giunzione, calcolato direttamente dall'integrale:

$$E(x) = \int_0^x \frac{Q(r)}{A \epsilon_0} dr$$

- Infine, in *blu* il potenziale "built-in" $V_b(x)$, calcolato direttamente dall'integrale:

$$V(x) = - \int_0^x E(r) dr$$

da dove notiamo che questo corrisponde (senza imporre nessun valore) al valore della barriera di potenziale che calcoleremo nella prossima sezione.

Barriera di potenziale

Infine, possiamo interrogarci riguardo al potenziale $V(x)$. In generale, vale:

$$E(x) = - \frac{dV(x)}{dx}$$

Questo, nella giunzione appena vista, implica la seguente legge:

5.1: Barriera di potenziale

In una giunzione P/N, per via del campo E che si forma nella zona di svuotamento, si avrà la formazione di un campo V_0 .

$$\Delta V = V_0 = \frac{k_B T}{e} \ln \left(\frac{N_A N_D}{n_i^2} \right)$$

Chiamiamo questa differenza di potenziale ΔV che si misura fra la regione prossima all'anodo e la regione prossima al catodo della giunzione **potenziale "built-in"** interno alla giunzione.

Il calcolo di ΔV è banale note le espressioni del campo E viste 2 sezioni fa. Possiamo procedere usando l'espressione sia in funzione dei portatori negativi n che in funzione dei portatori positivi p .

- Se si sceglie di proseguire con i portatori n , si ha l'espressione del campo:

$$E = -\frac{k_B T}{e} \frac{1}{n} \frac{dn}{dx}$$

dalla definizione del potenziale si ottiene di dover integrare:

$$E = -\frac{dV}{dx} \implies V = -\int_0^l E dx = \int_{\frac{n_i^2}{N_A}}^{N_D} \frac{k_B T}{e} \frac{1}{n} \frac{dn}{dx} dx$$

Soffermiamoci un istante sugli estremi di integrazione. Abbiamo che finiamo per integrare su dn , per cui dobbiamo prendere la variazione di n sul percorso dove calcoliamo il potenziale, cioè dall'estremo N all'estremo P della giunzione.

- All'estremo N della giunzione si ha che, per il doping di tipo N, la densità elettronica è semplicemente:

$$n = N_D$$

- All'estremo P della giunzione si ha che, per il doping di tipo P e la legge massa-azione (4.1), la densità elettronica è:

$$n = \frac{n_i^2}{N_A}$$

da cui gli estremi che vediamo nell'integrale. Risolvendo si ottiene quindi:

$$V = \int_{\frac{n_i^2}{N_A}}^{N_D} \frac{k_B T}{e} \frac{1}{n} dn = \frac{k_B T}{e} \ln(n) \Big|_{\frac{n_i^2}{N_A}}^{N_D} = \frac{k_B T}{e} \ln \left(\frac{N_A N_D}{n_i^2} \right)$$

che è quanto stavamo cercando.

- Il processo con i portatori p è analogo, si ha l'espressione del campo:

$$E = +\frac{k_B T}{e} \frac{1}{p} \frac{dp}{dx}$$

e l'integrale è:

$$V = - \int_0^l E dx = - \int_{N_A}^{\frac{n_i^2}{N_D}} \frac{k_B T}{e} \frac{1}{p} \frac{dp}{dx} dx$$

Gli estremi di integrazione sono scelti come prima. Abbiamo che finiamo per integrare su dp , per cui dobbiamo prendere la variazione di p sul percorso dove calcoliamo il potenziale, cioè dall'estremo N all'estremo P della giunzione.

- All'estremo N della giunzione si ha che, per il doping di tipo N e la legge massa-azione, la densità di lacuna è:

$$n = \frac{n_i^2}{N_D}$$

- All'estremo P della giunzione si ha che, per il doping di tipo P, la densità di lacuna è semplicemente:

$$p = N_A$$

da cui gli estremi che vediamo nell'integrale. Risolvendo si ottiene quindi:

$$V = - \int_{N_A}^{\frac{n_i^2}{N_D}} \frac{k_B T}{e} \frac{1}{p} dp = \frac{k_B T}{e} \ln(p) \Big|_{N_A}^{\frac{n_i^2}{N_D}} = \frac{k_B T}{e} \ln \left(\frac{N_A N_D}{n_i^2} \right)$$

che è nuovamente quanto stavamo cercando.

In verità, fra l'anodo e la giunzione con la regione P si viene a formare un **potenziale di contatto** V_1 , e allo stesso modo fra il catodo e la giunzione con la regione N si viene a formare un'altro potenziale di contatto V_2 . Si dimostra che:

$$V_1 + V_2 + V_0 = 0$$

per cui è impossibile misurare direttamente il potenziale built-in ΔV a partire da anodo e catodo. Notiamo però che tale potenziale esiste comunque all'interno della giunzione.

Notiamo quindi che nel complesso:

- Una lacuna che vuole spostarsi dalla regione P alla regione N deve vincere il potenziale "built-in" ΔV . Questo perché l'energia potenziale di lacuna è:

$$U = eV$$

e spostarsi da sinistra verso destra significa aumentare V , quindi aumentare propria energia potenziale;

- Un elettrone libero che vuole spostarsi dalla regione N alla regione P deve vincere sempre lo stesso potenziale "built-in" ΔV . Questo perché l'energia potenziale elettronica è:

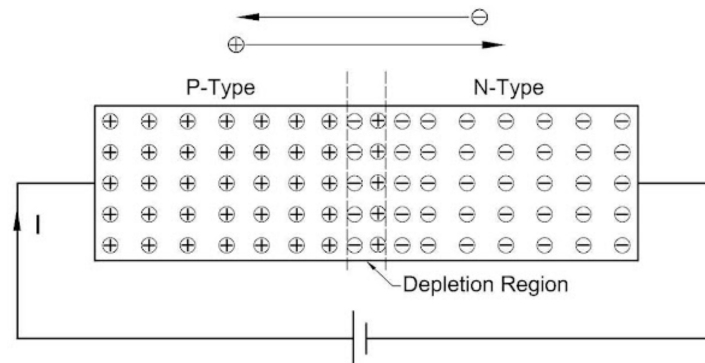
$$U = -eV$$

e spostarsi da destra verso sinistra significa diminuire V , quindi aumentare la propria energia potenziale;

Si ha quindi che il potenziale "built-in" funge in qualche modo da *barriera di potenziale* per la corrente che viaggia da sinistra verso destra lungo la giunzione.

Bias in avanti

Vediamo cosa accade quando si collega la giunzione P/N ad un generatore di potenziale V che ponga l'anodo a tensione positiva e il catodo a tensione negativa. Si dice in questo caso che si dà un **bias in avanti** alla giunzione.



Il primo effetto che osserviamo è che la barriera di potenziale diminuisce, in quanto introduciamo un nuovo potenziale V che va a diminuire ΔV :

$$\Delta V = V_0 - V$$

Questo avrà un effetto sulle correnti di diffusione e di deriva che hanno luogo all'interno della zona di svuotamento della giunzione. In particolare:

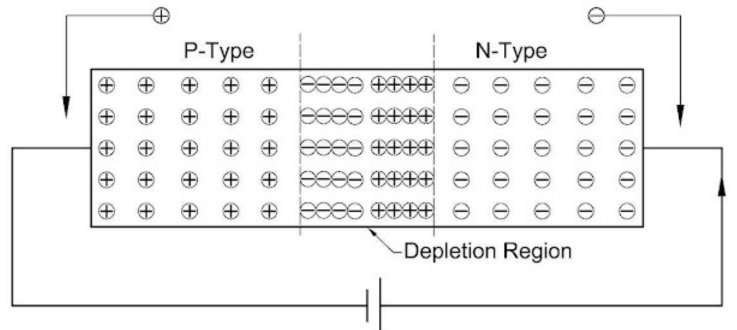
- Le densità di corrente di *diffusione* $\vec{J}_{N,Diffusione}$ e $\vec{J}_{P,Diffusione}$ aumenteranno in quanto sarà ridotta la barriera di potenziale che gli elettroni e le lacune rispettivamente dovranno vincere;
- Per quanto riguarda le densità di corrente di *deriva* $\vec{J}_{N,Drift}$ e $\vec{J}_{P,Drift}$, invece si avrà che queste sono piccole e restano costanti. Questè perchè le cariche che si spostano nella zona di svuotamento per effetto del campo che vi si forma sono le cariche minoritarie di ciascuna regione:
 - Per effetto del campo elettrico E formato, le lacune della regione N (portatrici minoritarie) derivano verso la regione P di minore potenziale;
 - Allo stesso modo, gli elettroni della regione P (portatori minoritari) derivano verso la regione N di maggiore potenziale.

Visto che la corrente di diffusione aumenta esponenzialmente, si ha che la conduzione può effettivamente accadere nella regione di svuotamento, e avviene con maggioranza di corrente di diffusione (piuttosto che di drift). Il risultato sarà che la densità di corrente complessiva \vec{J} non sarà più nulla, ma invece:

$$\vec{J} \neq 0$$

Bias all'indietro

Vediamo la situazione opposta, cioè quella dove si collega la giunzione P/N ad un generatore di potenziale V che ponga l'anodo a tensione negativa e il catodo a tensione positiva. Si dice in questo caso che si dà un **bias all'indietro** alla giunzione.



Il primo effetto che osserviamo è che in questo la barriera di potenziale aumenta, in quanto introduciamo un nuovo potenziale V che va ad aumentare ΔV :

$$\Delta V = V_0 + V$$

Come nella situazione di bias in avanti, si avrà un effetto sulle correnti di diffusione e di deriva che hanno luogo all'interno della zona di svuotamento della giunzione. In particolare:

- Le densità di corrente di *diffusione* $\vec{J}_{N,Diffusione}$ e $\vec{J}_{P,Diffusione}$ diminuiranno in quanto sarà aumentata la barriera di potenziale che gli elettroni e le lacune rispettivamente dovranno vincere;
- Per quanto riguarda le densità di corrente di *deriva* $\vec{J}_{N,Drift}$ e $\vec{J}_{P,Drift}$, invece si avrà che queste sono piccole e restano costanti secondo le stesse motivazioni della scorsa sezione.

Visto che la corrente di diffusione diminuisce esponenzialmente, si ha che la conduzione può effettivamente accadere nella regione di svuotamento, e avviene con maggioranza di corrente di drift, cioè questa riesce a vincere la corrente di diffusione. Nuovamente si ha che la densità di corrente complessiva \vec{J} non sarà più nulla, ma invece:

$$\vec{J} \neq 0$$

Questa densità di corrente sarà quindi dominata dal drift, e sarà *piccola*, rivolta da catodo verso anodo (quindi all'opposto della corrente vista nel bias in avanti).

Larghezza della zona di svuotamento

Notiamo che la *larghezza* della zona di svuotamento dipende dalla V che imponiamo nei bias in *avanti* e all'*indietro*. In particolare, dipende da:

$$w = \sqrt{\frac{2\epsilon_0}{e} \left(\frac{1}{N_A} + \frac{1}{N_d} \right) (V_0 - V)}$$

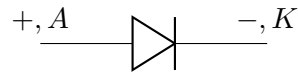
Per cui:

- Nel caso del bias in avanti, si ha una *diminuzione* della larghezza della zona di svuotamento w ;
- All'inverso, nel caso del bias all'indietro, si ha un *aumento* della larghezza della zona di svuotamento w ;

6 Lezione del 13-03-26

6.1 Diodi

Le giunzioni P/N come le abbiamo studiate nell'ultima lezione formano quei componenti elettronici noti come **diodi**, col simbolo del componente circuitale:



vediamo come si prende la corrente I_D come la corrente lungo il diodo, da + (anodo A) a - (catodo K), e il voltaggio V_d come quello che si trova fra + e -.

I diodi sono governati dalla cosiddetta legge di *Shockley*:

6.1: Legge di Shockley

La corrente I_D su un diodo sottoposto a tensione V_D dipende da:

$$I_D = I_S \left(e^{\frac{V_D}{\eta V_T}} - 1 \right)$$

dove I_S è la *corrente inversa di saturazione*, η è il *fattore di idealità*, e V_T è la *tensione termica*:

$$V_T = \frac{k_B T}{e}$$

già vista nella legge di Einstein a 4.3, e già usata per il calcolo del potenziale "built-in" della giunzione P/N in 5.1.

Tracciamo quindi il grafico tensione-corrente del diodo, prendendo i seguenti intervalli:

- Per $V_D = 0$, avremo $I_D = 0$;
- Per V_D positivo, il termine esponenziale domina, e si ha:

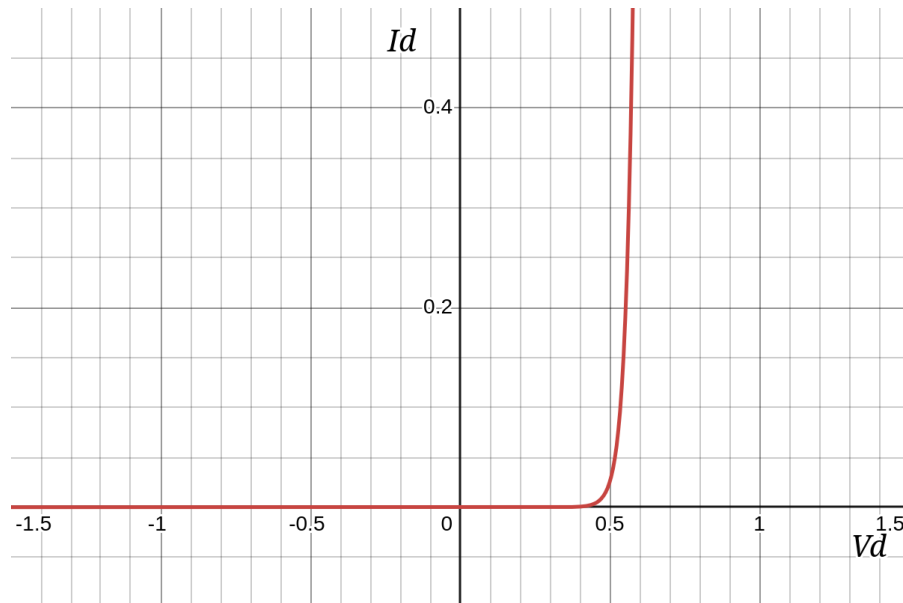
$$I_D \approx I_S e^{\frac{V_D}{\eta V_T}}$$

cioè la crescita è sostanzialmente esponenziale;

- Per V_D negativo, il termine esponenziale è molto piccolo, e si ha:

$$I_D \approx -I_S$$

Col fatto che I_S è solitamente molto piccola (intorno a 10^{-10} A), si ha quindi che il valore di I_D per $V_D < 0$ è piccolo e costante.



Abbiamo quindi visto il comportamento tipico del diodo, che è chiaramente un componente *non lineare*: per $V_D > 0$ a cui si sottopone il diodo, la corrente cresce in maniera esponenziale, mentre per $V_D < 0$ quasi non passa corrente. Possiamo dire che, in qualche modo, il diodo "rettifica" la corrente che ci passa.

6.1.1 Saturazione e breakdown

Vediamo alcuni fenomeni particolari che il diodo presenta agli estremi di tensione:

- Notiamo subito che la legge esponenziale per $V_D > 0$ non è soddisfatta in assoluto: esiste una certa soglia massima di **saturazione**, oltre la quale I_D resta costante o addirittura decresce.
- Per il caso opposto, cioè $V_D < 0$, il diodo entra in una fase di cosiddetto **breakdown** (reversibile). Abbiamo infatti che a V_D molto negativi, oltre una certa soglia la corrente passa da I_S costante ad un valore esponenzialmente grande, cioè una grande corrente si trova a scorrere sul diodo. Per questo motivo, spesso si mettono resistenze in serie ai diodi, in modo da evitare di superare i limiti di potenza nel caso di fenomeni di breakdown.

Per spiegarci tale fenomeno, ricordiamo che la corrente in polarizzazione inversa è principalmente una corrente di drift, cioè causa dei portatori minoritari di ciascuna regione (P ed N) che viaggiano lungo il potenziale "built-in" (formato sulla giunzione dagli ioni fissi), aumentato dal potenziale di polarizzazione. Dovrà quindi esistere un qualche fenomeno che aumenta il numero di portatori minoritari a grandi valori di V_D . In verità questi effetti sono due:

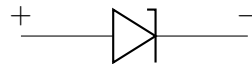
- Effetto **Zener**: in opportune situazioni di drogaggio, il campo elettrico nella zona di svuotamento è abbastanza forte da rompere alcuni legami covalenti, portando alla formazione di nuovi portatori;
- Effetto **valanga**: gli elettroni in moto nella giunzione P/N possono fornire energia tale agli atomi contro cui collidono da liberare altri atomi (*ionizzazione per urto*). In particolare, vediamo che se partiamo con un elettrone e_0 , dopo

un urto otteniamo 2 nuovi portatori (e_1 e la relativa lacuna h_1). Entrambi di questi portatori (anche la lacuna, ma chiaramente spostandosi in direzione opposta: questo potrebbe sembrare strano, ma dobbiamo pensare alle lacune come particelle positive, dove l'energia immagazzinata è quella che viene tolta agli elettroni che vengono spostati) possono quindi continuare a collidere e a generare nuovi portatori. Questo fenomeno ha quindi crescita esponenziale, cioè all i -esimo urto partendo da un solo elettrone e_0 si hanno:

$$[e, p] = 3^i$$

portatori di carica.

Esistono particolari tipi di diodi, detti **diodi Zener**, che sfruttano l'effetto di breakdown. Il loro simbolo del componente circuitale è:



e la loro caratteristica è che l'effetto di breakdown avviene a tensioni molto più piccole del normale (ordine di 100 V nei diodi normali, ordine dei 10 V o meno nei diodi Zener). Spesso i diodi Zener vengono fatti lavorare nella loro zona di breakdown, in quanto in questo modo possono fare da riferimento per la tensione (se la curva di corrente in fase di breakdown fosse perfettamente verticale, sarebbero effettivamente generatori di tensione).

6.1.2 Temperatura nei diodi

Vediamo qual'è l'effetto della temperatura sul comportamento dei diodi.

- Nel caso di diodi ad effetto *Zener*, all'aumentare della temperatura T la tensione di breakdown V_{BR} aumenta (cioè l'effetto di breakdown avviene ad una tensione maggiore);
- Nel caso di diodi ad effetto *Valanga*, prendiamo in considerazione il fatto che se si aumenta T la mobilità elettronica μ_n e di lacuna μ_p diminuiscono (aumenta la frequenza degli urti).

Potremmo pensare che questo porti ad un maggiore effetto valanga, ma il risultato effettivo è che il tempo di cammino medio τ , e quindi la lunghezza del cammino medio stesso (abbiamo parlato di cammini liberi discutendo il modello di Drude-Lorentz in 2.1.2), diminuiscono, per cui diminuisce l'energia che i portatori riescono ad accumulare e quindi trasferire agli atomi in fase di urto. Il risultato è quindi che la tensione di breakdown V_{BR} diminuisce (cioè l'effetto di breakdown avviene ad una tensione minore).

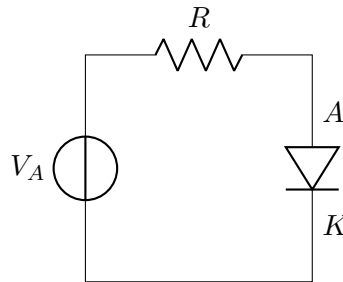
Esistono diodi particolari, all'interno della famiglia dei diodi Zener, costruiti per presentare entrambi gli effetti. Solitamente infatti si ha che:

- I diodi che stanno fra 0 e 5 volts di tensione di breakdown V_{BR} presentano principalmente l'effetto Zener;
- I diodi che stanno fra 7 e $+\infty$ volts di tensione di breakdown V_{BR} presentano principalmente l'effetto Zener;

- I diodi che stanno nella regione di mezzo (fra 5 e 7 volts) della tensione di breakdown V_{BR} presentano principalmente entrambi gli effetti. Il risultato è che i 2 effetti in qualche modo si bilanciano, e la tensione V_{BR} resta perlopiù costante. Questo è utile nel caso di diodi Zener che, come abbiamo detto nella scorsa sezione, servono a fare da riferimento di tensione.

6.1.3 Circuito RD

Vediamo quindi il diodo da un punto di vista circuitale, risolvendo un semplice circuito, formato da un resistore in serie ad un diodo:



In casi come questo, visto che si parla di componenti non lineari, molti dei teoremi e delle leggi che conosciamo dall'elettrotecnica non valgono più. Dobbiamo quindi trovare dei metodi di soluzione alternativi.

- **Metodo numerico**

Anche se non valgono più molti teoremi, valgono sempre, invece, le leggi di analisi circuitale di Kirchoff (viste nella lezione 3). Imponendo ad esempio la seconda legge sulla maglia del circuito si ha:

$$\begin{cases} V_A = RI_D + V_D \\ I_D = I_s \left(e^{\frac{V_D}{\eta V_T}} - 1 \right) \end{cases}$$

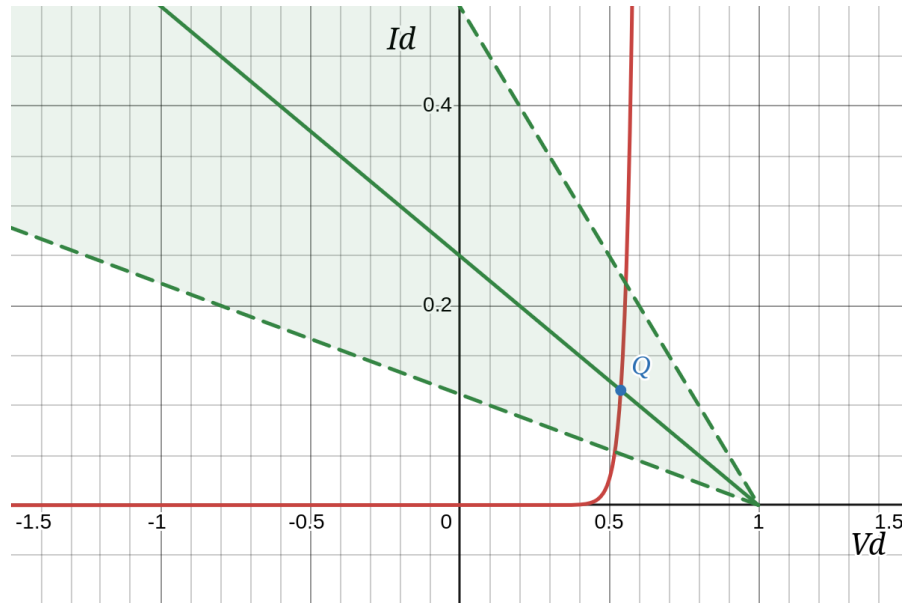
Questo sistema è però fortemente non lineare (presenta una funzione esponenziale), ed è quindi adatto particolarmente a metodi *numerici* (anche se questo piccolo esempio è risolvibile, circuiti più complessi con più diodi probabilmente non lo saranno). Per questo motivo, lo chiamiamo metodo **numerico**.

- **Metodo grafico**

Una possibile soluzione, in casi semplici, è data dal metodo grafico. Potremo infatti ricavare I_D da entrambi le equazioni dello scorso sistema:

$$\begin{cases} I_D = \frac{V_A - V_D}{R} \\ I_D = I_s \left(e^{\frac{V_D}{\eta V_T}} - 1 \right) \end{cases}$$

cioè rispettivamente ottenere una retta ed un esponenziale, che potremo quindi intersecare su un grafico per trovare valori approssimati di I_D al particolare V_D (formando la coppia $((V_{DQ}, I_{DQ}))$ che risolve il sistema).



Questo punto di intersezione viene detto punto di **lavoro** o di *quiescenza*, ed indicato con Q .

Notiamo che la retta:

$$I_D = \frac{V_A - V_D}{R}$$

dove V_A è il potenziale esterno e V_D è quello sul diodo, viene chiamata **retta di carico** (e l'intero metodo prende anche il nome di *metodo della retta di carico*).

Sempre dal grafico, notiamo che variando R si ottengono diversi grafici della retta di carico, che mostrano direttamente la variazione della corrente al punto di quiescenza Q . In particolare:

- Al diminuire di R , si ottiene una retta del fascio con coefficiente angolare più grande (in modulo), e quindi si interseca un punto più alto della curva di Shockley (maggiore corrente I_D sul diodo);
- All'aumentare di R , si ottiene una retta del fascio con coefficiente angolare più piccolo (in modulo), e quindi si interseca un punto più basso della curva di Shockley (minore corrente I_D sul diodo);

• Metodo approssimato

Notiamo che gli ultimi due metodi che abbiamo visto (metodo *numerico* e metodo *grafico*) sono metodi *non approssimati* e quindi esatti (che per noi significa ad altissima precisione). Vediamo quindi che esistono metodi ad alta precisione, ma approssimati, che rinunciano alla legge di Shockley (che non sappiamo risolvere) e quindi danno vita a sistemi risolvibili. Questi vengono detti modelli del diodo per **grandi segnali**.

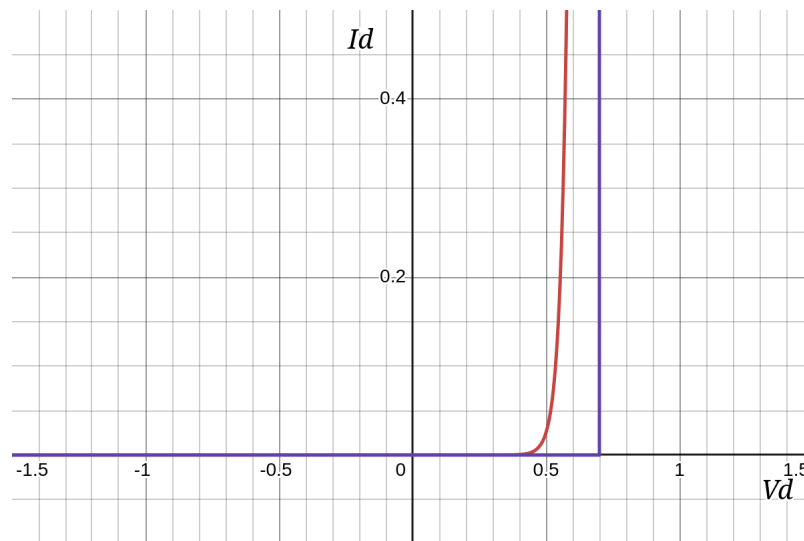
6.1.4 Modelli del diodo per grandi segnali

Vediamo quindi quali sono questi 3 modelli approssimati del diodo per *grandi segnali*.

• Modello a caduta costante

Nel modello a **caduta costante** sfruttiamo il fatto che I_D è quasi nulla per V_D minore di una certa soglia (solitamente $V_\gamma = 0.7\text{ V}$). Prendiamo quindi il diodo come

un circuito *aperto* al di sotto di tale soglia ($V_D < V_\gamma$), e di lì in poi ($V_D > V_\gamma$) lo approssimiamo come un *generatore di tensione* (qualsiasi corrente alla stessa tensione, cioè sul grafico una linea verticale).



Dal punto di vista dei calcoli, un modello di questo tipo (sostanzialmente binario) si risolve facendo una prima ipotesi riguardo al regime di comportamento del diodo (aperto o generatore di tensione), quindi risolvendo il circuito, e controllando se il risultato è fisicamente valido. Se non lo è, chiaramente l'ipotesi da fare sarà stata l'altra.

Questo si può fare grazie al *teorema dei circuiti*:

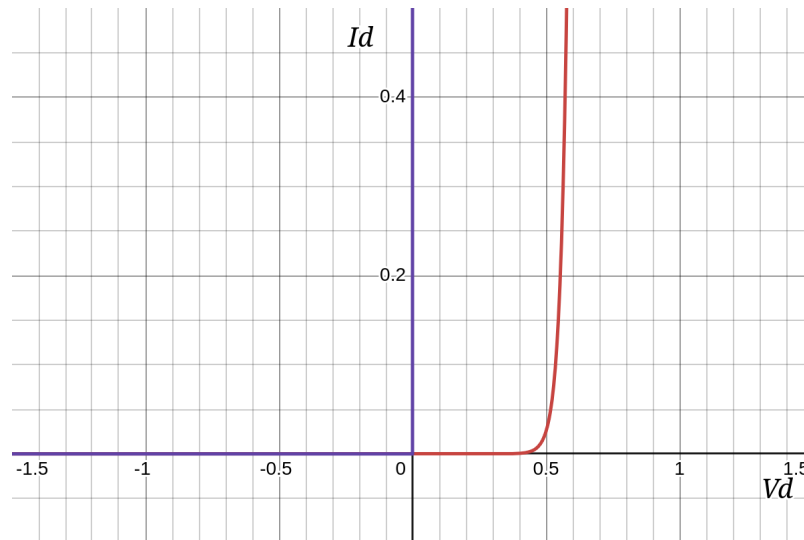
6.1: Teorema dei circuiti

Che in un circuito ci siano elementi lineari e non, se questi hanno tutti andamento monotono crescente, esiste una e una sola soluzione del circuito.

Notiamo però che questo metodo diventa difficile da applicare per molti diodi, in quanto si hanno 2 opzioni per diodo, che risultano in 2^n possibilità di ipotesi iniziali per n diodi.

- **Modello del diodo ideale**

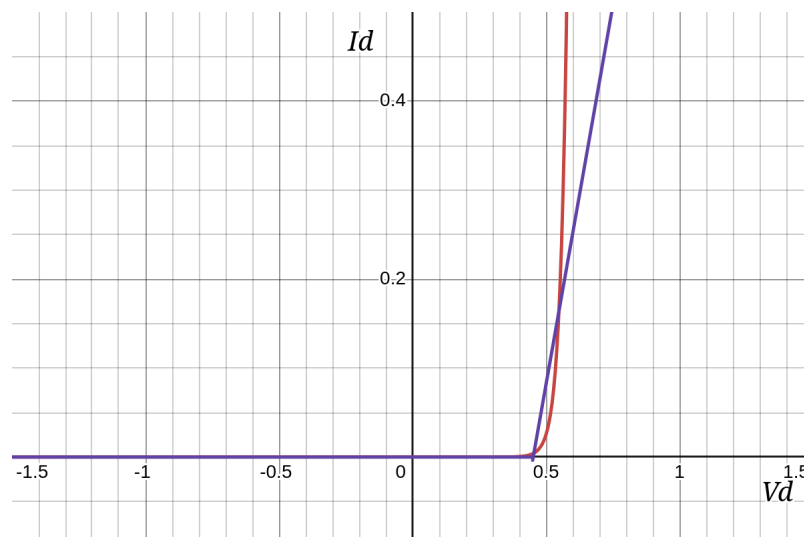
Questo modello equivale a quello della caduta costante, ma con V_γ preso nullo. Questo significa che per $V_D < 0$ il diodo modellizzato si comporterà come un circuito *aperto*, mentre per $V_D > 0$ si comporterà come un circuito aperto. Notiamo che per fare questo, trascuriamo completamente la caduta di potenziale sul diodo: in questo il modello del diodo ideale è più semplice di quello a caduta costante (risulta valido in circuiti dove le cadute di potenziale sono più grandi di quelle del diodo, cioè di $V \approx 0.7$ V).



Dal punto di vista dei calcoli, un modello di questo tipo è sempre binario, e si risolve facendo una prima ipotesi riguardo al regime di comportamento del diodo (aperto o cortocircuito), quindi risolvendo il circuito, e controllando se il risultato è fisicamente valido. Se non lo è, chiaramente l'ipotesi da fare sarà stata l'altra.

- **Modello lineare a tratti**

In questo caso cerchiamo un'approssimazione migliore della caduta costante, lineare, che cerchi di "inseguire" in qualche modo l'andamento dell'esponenziale, almeno vicino a $V_\gamma \approx 0.7$ V. In questo caso, per $V_D < V_\gamma$, il diodo modellizzato si comporta sempre come un circuito *aperto*, mentre per $V_D > V_\gamma$ si comporta come una serie *generatore di tensione - resistenza*.



Dal punto di vista dei calcoli questo modello è analogo ai precedenti, ma presenta delle complessità maggiori: andiamo a sostituire il diodo (che è un solo componente circuitale) con 2 componenti (generatore di tensione e resistenza), cosa che comporta una grande complicazione di circuiti con molti diodi.

7 Lezione del 16-03-26

Abbiamo visto nella scorsa lezione come trattare i *diodi* all'interno dei nostri circuiti. In particolare, avevamo dato una soglia $V_\gamma = 0.7\text{ V}$ di "attivazione" del diodo.

- Diciamo che quando questo si trova oltre la soglia, e quindi si comporta da conduttore, siamo nell'ipotesi di *conduzione* del diodo;
- Diciamo che quando questo si trova al disotto la soglia, e quindi si comporta da aperto, siamo nell'ipotesi di *interdizione* del diodo;

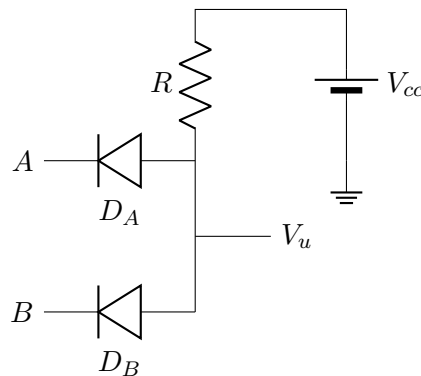
Risolvere circuiti contenenti diodi significa quindi fare ipotesi sulla conduzione o interdizione del diodo, e quindi verificare se tali ipotesi sono vere.

7.1 Logica a diodi

Iniziamo a vedere alcuni semplici circuiti in logica diodi, a partire dalla realizzazione di alcune **porte logiche**. Questo ci permetterà di mettere alla prova le tecniche per la risoluzione dei circuiti con diodi che abbiamo visto nelle ultime sezioni.

7.1.1 Studio di una porta AND

Vediamo come primo esempio di studio dei circuiti basati su diodi, l'analisi di un circuito che implementa una porta **AND**. Il circuito che prendiamo è il seguente:



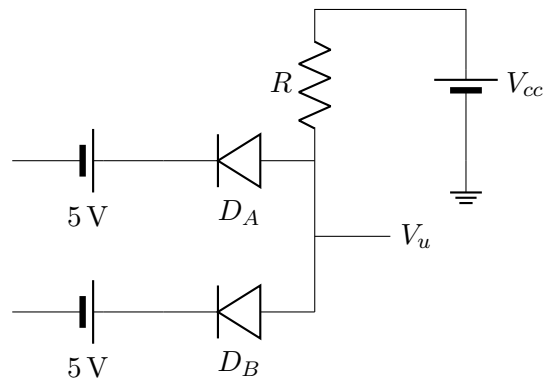
Abbiamo quindi due diodi collegati ai capi A e B , di cui assumiamo di poter variare il voltaggio V_A e V_B rispettivamente. V_{cc} sarà la nostra sorgente di alimentazione, fissa a 5 V . Il valore di R non è particolarmente significativo, ma come tutte le resistenze assumiamo che sia $R > 0$. V_u sarà il valore di uscita che vogliamo modulare con V_A e V_B .

Vediamo quindi tutte le possibili combinazioni di A e B , assumendo di fissarli o a 0 V o a 5 V . Il risultato che ci aspettiamo di ottenere è quello dato dalla seguente tabella:

V_A	V_B	V_u
5 V	5 V	5 V
0 V	5 V	0 V
5 V	0 V	0 V
0 V	0 V	0 V

- $V_A = 5\text{ V}$, $V_B = 5\text{ V}$

La situazione si modella come segue:



In questo caso vediamo che i diodi D_A e D_B hanno $+5\text{ V}$ dal lato sinistro, e qualcosa di simile a $+5\text{ V}$ dal lato destro (dalla sorgente V_{cc}). In verità potremmo prevedere di avere una caduta di potenziale su R , ma in ogni caso questo porterà il voltaggio V' a destra dei diodi a:

$$V' = V_{cc} - IR < V_{cc}$$

per cui in ogni caso i diodi si trovano a voltaggio 0, o addirittura in polarizzazione negativa. Questo significherà che possiamo considerarli come entrambi interdetti. Se i diodi sono interdetti, allora il circuito non è neanche chiuso, e il voltaggio sul nodo V_u sarà uguale al V_{cc} , cioè:

$$V_u = 5\text{ V}$$

Verifica

La verifica in caso di diodi presi *interdetti* si fa ponendo:

$$V_{AK} < V_\gamma$$

dove V_{AK} è la differenza di potenziale fra anodo e catodo del diodo preso interdetto. In questo caso avremo che sia D_A che D_B sono stati presi interdetti, per cui:

$$V_{D_A}, V_{D_B} < V_\gamma \approx 0.7\text{ V}$$

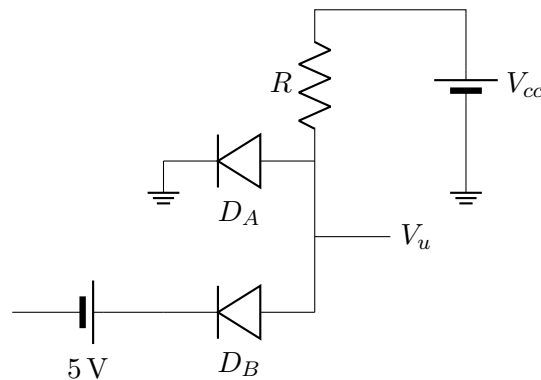
In questo caso, come abbiamo detto, i diodi sono a voltaggio 0 o in polarizzazione negativa, per cui:

$$V_{D_A}, V_{D_B} < 0\text{ V} \implies V_{D_A}, V_{D_B} < V_\gamma \approx 0.7\text{ V}$$

e l'ipotesi è verificata.

- $V_A = 0\text{ V}$, $V_B = 5\text{ V}$

La situazione si modella come segue:



In questo caso possiamo fare l'ipotesi che il diodo D_B , come nel caso precedente, sia in interdizione. Il diodo D_A , invece, si troverà a chiudere un circuito diretto fra V_{cc} e ground, per cui potremmo prenderlo come in conduzione.

In tal caso, sarà immediato notare (preso il modello a diodo ideale) che a sinistra e a destra di D_A siamo collegati direttamente al ground, per cui il voltaggio al nodo V_u è nullo, cioè:

$$V_u = 0\text{ V}$$

Verifica

La verifica in caso di diodi presi *conduttori* si fa ponendo:

$$I_d > 0$$

dove I_d è la corrente sul diodo. Nel nostro circuito, abbiamo che la corrente sul diodo in conduzione D_A è:

$$I_{D_A} = \frac{V_{cc}}{R} > 0$$

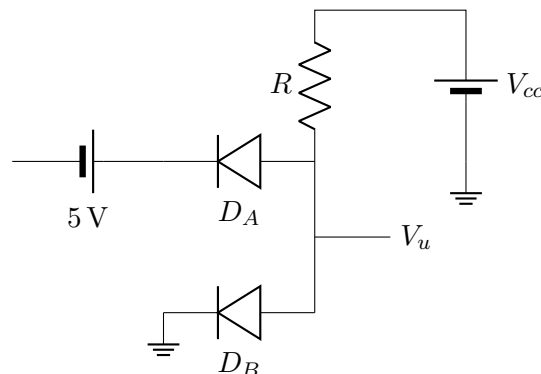
in quanto $V_{cc} > 0$ e R , come tutte le resistenze, è $R > 0$.

Per quanto riguarda D_B , invece, lo abbiamo preso in interdizione e quindi applichiamo la formula di verifica del caso precedente:

$$V_{AK} = V_{D_B} = -5\text{ V} < V_\gamma$$

- $V_A = 5\text{ V}$, $V_B = 0\text{ V}$

La situazione si modella come segue:



Qui vale l'ipotesi opposta alla precedente: D_A sarà in interdizione mentre D_B sarà in conduzione (in quanto chiude un circuito diretto fra V_{cc} e ground).

Come prima, vale che a sinistra e a destra di D_B siamo collegati direttamente al ground, per cui il voltaggio al nodo V_u è nullo, cioè:

$$V_u = 0 \text{ V}$$

Verifica

La verifica si fa in maniera analoga al caso precedente. Nel nostro circuito, abbiamo che la corrente sul diodo in conduzione D_B è:

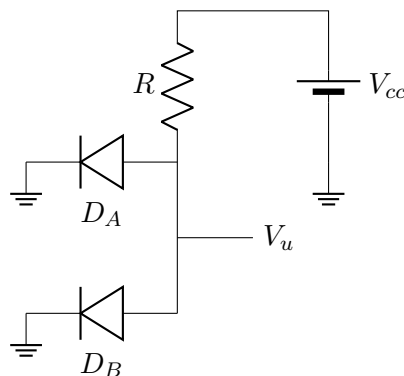
$$I_{D_B} = \frac{V_{cc}}{R} > 0$$

che è subito dimostrato, mentre per quanto riguarda D_A lo abbiamo preso in interdizione e quindi applichiamo la formula di verifica del caso precedente:

$$V_{AK} = V_{D_B} = -5 \text{ V} < V_\gamma$$

- $V_A = 0 \text{ V}$, $V_B = 0 \text{ V}$

La situazione si modella come segue:



In questo caso ci si aspetta che entrambi i diodi D_A e D_B siano in conduzione, in quanto fanno parte di un circuito diretto da V_{cc} al ground.

Abbiamo quindi lo stesso risultato di prima:

$$V_u = 0 \text{ V}$$

dato dal fatto che V_u si trova sullo stesso polo di D_A e D_B (presi come ideali), collegato al ground.

Verifica

La verifica si fa verificando l'ipotesi di conduzione, e quindi prendendo:

$$I_d > 0, \quad I_{D_A}, I_{D_B} > 0$$

che è verificata considerando che sui due rami passa la corrente:

$$I_R = \frac{V_{cc}}{R} > 0$$

Adesso, verificati tutti i casi, assegnamo dei valori logici ai voltaggi V_A , V_B e V_u . Diciamo ad esempio che:

- 0 V corrisponde al valore logico *falso*;
- 5 V corrisponde al valore logico *vero*.

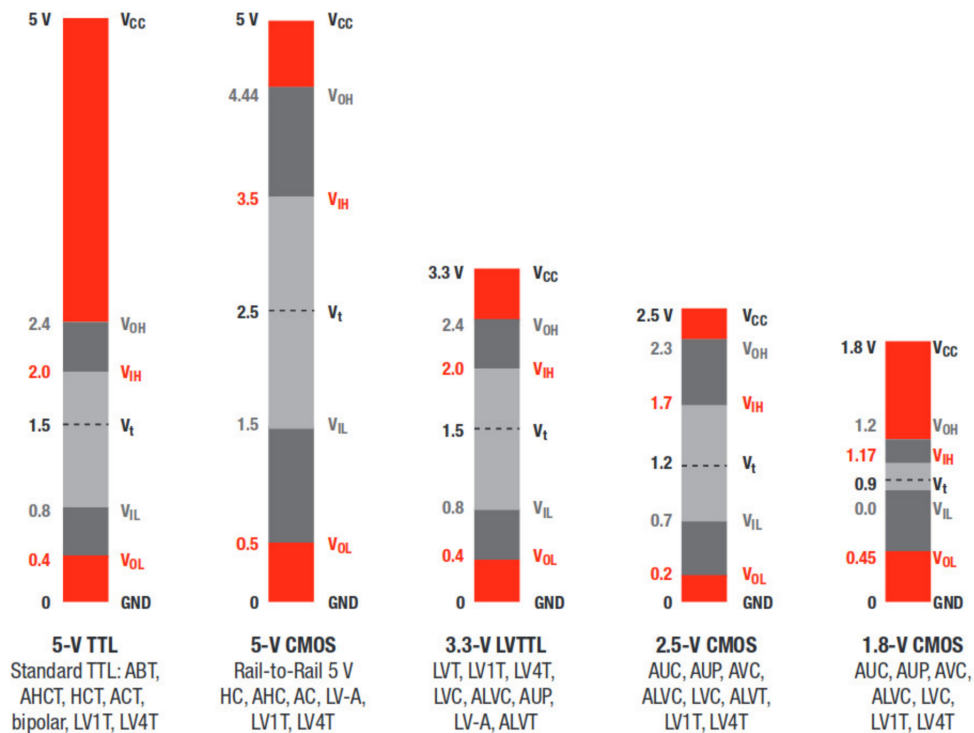
Applicando tale associazione alla tabella dei voltaggi riportata sopra otteniamo:

A	B	u
1	1	1
0	1	0
1	0	0
0	0	0

Vediamo quindi che il circuito implementa la tavola di verità dell'operazione di congiunzione logica, e quindi l'AND logico. Notiamo che questo è il modo in cui implementiamo tutte le *porte logiche*:

- Nel dominio *analogico*, creiamo circuiti che manipolano voltaggi e correnti in modo da ottenere i risultati desiderati;
- Solo dopo, assegnamo a determinati range di corrente o voltaggio dei valori logici, in modo da implementare operazioni nel cosiddetto dominio *logico*.

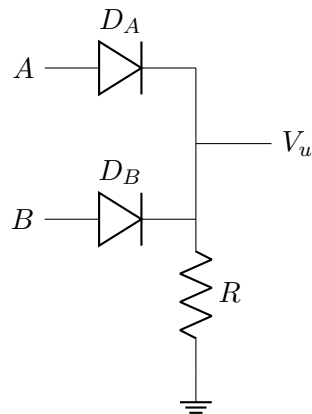
In tal caso, visto che i valori analogici sono *reali* $\in \mathbb{R}$, e non valori discreti, dobbiamo definire dei *range* di validità dei valori falsi o veri rappresentati dal voltaggio. Fra i due ci sarà una zona di indeterminazione, che vorremo evitare, e tollereremo solo in particolari situazioni transienti.



In particolare, noi abbiamo scelto 5 V come riferimento per il *vero* logico. Questo è il valore tipico della logica a diodi o a transistor (**TTL**, *Transistor-Transistor Logic*). Le moderne applicazioni **CMOS** (*Complementary Metal-Oxide-Semiconductor*) usano invece valori come 3.3 V, 2.5 V o addirittura 1.8 V.

7.1.2 Studio di una porta OR

Come abbiamo visto una porta *AND*, vediamo a scopo di esempio anche una porta **OR**. Il circuito che prendiamo è il seguente:



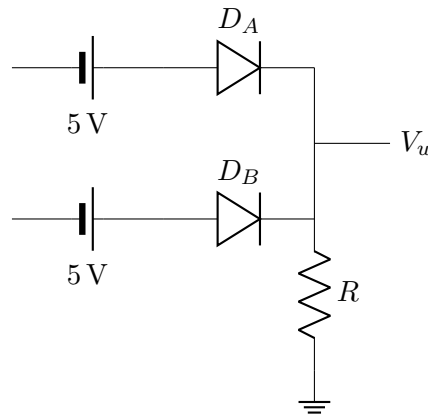
In questo caso abbiamo sempre i morsetti *A* e *B*, e i relativi diodi D_A e D_B , che faranno da ingressi della porta, e il morsetto V_u , che farà da uscita. Notiamo invece la particolarità di non avere un generatore di tensione V_{cc} : la porta è quindi completamente *passiva*.

Come prima, quindi, vediamo tutte le possibili combinazioni di *A* e *B*, assumendo di fissarli o a 0 V o a 5 V. Il risultato che ci aspettiamo di ottenere è quello dato dalla seguente tabella:

V_A	V_B	V_u
5 V	5 V	5 V
0 V	5 V	5 V
5 V	0 V	5 V
0 V	0 V	0 V

- $V_A = 5\text{ V}$, $V_B = 5\text{ V}$

La situazione si modella come segue:



In questo caso entrambi i diodi D_A e D_B saranno in conduzione, in quanto chiuderanno un circuito con il ground attraverso R .

Avremo quindi che V_u è parte del polo formato da D_A e D_B , e la sua tensione sarà di conseguenza uguale a V_{D_A} e V_{D_B} :

$$V_u = 5\text{ V}$$

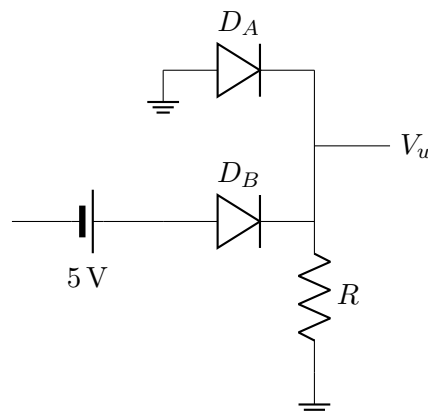
Verifica

Per ipotesi di conduzione verifichiamo che la corrente sia non negativa, per cui:

$$I_d > 0, \quad I_{D_A} = I_{D_B} = \frac{5\text{ V}}{R} > 0$$

- $V_A = 0\text{ V}$, $V_B = 5\text{ V}$

La situazione si modella come segue:



In questo caso D_A sarà in interdizione, portato a 5 V sul catodo dal diodo D_B (che di per sè sarà in conduzione). Vediamo qui l'utilità del diodo ad impedire che la corrente passi all'opposto della direzione desiderata (cioè imporre che non passi dal diodo D_B a 5 V al ground a sinistra di D_A).

La situazione in uscita non cambia in quanto il polo su V_u resta pilotato da voltaggio applicato a D_B , e quindi:

$$V_u = 5 \text{ V}$$

Verifica

Abbiamo fatto l'ipotesi di interdizione su D_A , per cui applichiamo la verifica del voltaggio:

$$V_{D_A} < 0 \implies V_{D_A} < V_\gamma \approx 0.7 \text{ V}$$

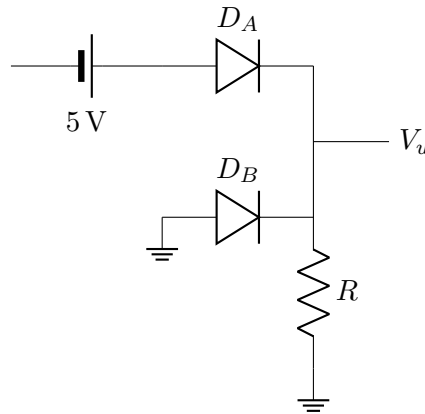
in quanto abbiamo detto D_A è in polarizzazione negativa.

Per quanto riguarda D_B , invece, vogliamo verificare $I_d > 0$:

$$I_d = I_{D_B} > 0 = \frac{5 \text{ V}}{R} > 0$$

- $V_A = 5 \text{ V}$, $V_B = 0 \text{ V}$

La situazione si modella come segue:



In questo caso D_B sarà in interdizione, portato a 5 V sul catodo dal diodo D_A (che di per sè sarà in conduzione), in una situazione completamente opposta alla precedente.

La situazione in uscita ancora una volta non cambia in quanto il polo su V_u resta pilotato da voltaggio applicato a D_A , e quindi:

$$V_u = 5 \text{ V}$$

Verifica

Abbiamo fatto l'ipotesi di interdizione su D_B , per cui applichiamo la verifica del voltaggio:

$$V_{D_B} < 0 \implies V_{D_B} < V_\gamma \approx 0.7 \text{ V}$$

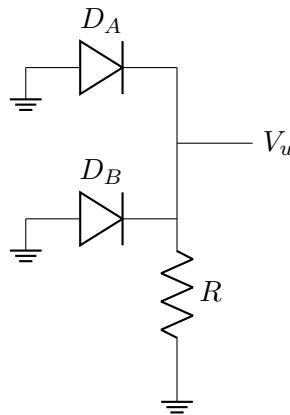
in quanto abbiamo detto D_B è in polarizzazione negativa.

Per quanto riguarda D_A , invece, vogliamo verificare $I_d > 0$:

$$I_d = I_{D_A} > 0 = \frac{5 \text{ V}}{R} > 0$$

- $V_A = 0\text{ V}$, $V_B = 0\text{ V}$

La situazione si modella come segue:



In questo caso la rete è triviale in quanto non presenta nemmeno un generatore di tensione. Tutto è connesso al ground, e quindi tutto è a 0 V . Possiamo assumere che diodi siano in interdizione, in quanto il voltaggio che hanno ad entrambi i lati è nullo.

Verifica

Per i diodi D_A e D_B in interdizione, e tutta la rete a 0 V , la verifica è banale:

$$V_{D_A} = V_{D_B} = 0 < V_\gamma \approx 0.7\text{ V}$$

Risulta immediato, applicando lo stesso schema di valori logici di prima, che abbiamo ricavato la tabella:

A	B	u
1	1	1
0	1	1
1	0	1
0	0	0

cioè la tabella di verità di una porta OR, che è ciò che ci eravamo prefissati di fare.

7.1.3 Pregi e difetti della logica a diodi

Abbiamo quindi visto alcuni semplici circuiti digitali in logica a diodi. Abbiamo che questi sono molto semplici (sia da capire che da realizzare), in quanto coinvolgono solo componenti semplici (diodi, resistenze, ecc...). Nonostante ciò, abbiamo che la logica a diodi ha delle problematiche:

- La corrente in entrata alle porte $I \neq 0$ può essere diversa da 0: questo è problematico in quanto ci piacerebbe avere una "interfaccia" comune ai nostri componenti (magari codificando i valori logici solo attraverso la tensione, e non attraverso la corrente);
- Visto che i circuiti reali non sono ideali, si ha un certo grado di *degradazione* del segnale. Ciò significa che la corrente e le tensioni in uscita dalle porte non sono

pari al V_{cc} di 5 V, ma leggermente minori. Il risultato sarà che collegando diverse porte in *cascata* si avrà una degradazione del segnale che eventualmente renderà i nostri circuiti inutilizzabili.

Questo problema è eliminato ad esempio dalla logica CMOS), che invece compie un'operazione di *rigenerazione* del segnale sulle porte;

- Infine, con la sola logica a diodi, non si ha a disposizione nessun tipo di *inverter* (porta logica NOT). Vediamo che questo non ci permette di realizzare tutta la logica booleana, e quindi rende la logica a diodi effettivamente per i calcolatori.

Una soluzione sarà data dalla logica a transistor, attraverso la quale è semplice realizzare ad esempio la porta **NAND**, che dagli anni '50 è stata alla base della logica digitale (ricordiamo che una NAND può formare una NOT connettendo gli ingressi fra di loro);

Inoltre, ci sono diverse caratteristiche da tenere a mente quando si progettano circuiti (in qualsiasi tipo di logica). Ad esempio:

- Il livello di *porte in cascata* è da considerare sia che si abbiano i problemi di degradazione della logica a diodi che no, in quanto si hanno tradeoff in termini di tempo di transizione, o spazio occupato dalle porte. Dai corsi di reti logiche, ricordiamo che esistono svariate tecniche di decomposizione dei circuiti logici, ricavate dalla logica booleana, che possono permetterci di realizzare lo stesso circuito logico usando più o meno porte;
- Anche i tradeoff legati alla *potenza* consumata dai circuiti in proporzione alla loro velocità sono da considerare: in generale, vorremo velocità più grandi possibili per diminuire i tempi di risposta, ma questo non viene se non compromettendo sulla potenza usata. Ad esempio, nella realizzazione di memorie digitali, sfruttiamo capacità. Per caricare più velocemente una capacità, è intuitivo che bisogna aumentare la corrente in entrata I . Ora, se la potenza è data da:

$$P = I^2 R$$

allora sarà semplice notare che maggiori velocità di carica corrispondono in linea di massima a maggiore potenza usata;

- Infine, consideriamo l'impatto che la struttura delle porte può avere sulla *superficie* del circuito. In generale i circuiti si pagano al mm^2 , per cui realizzare lo stesso circuito su una *footprint* minore può portare a risparmi energetici. Inoltre, la distanza fra i componenti, il loro numero, ecc... possono ancora influire sui tempi di risposta e di transizione, particolarmente importanti nei circuiti digitali. Anche in questo caso ci aiuta passare dalla logica a diodi o transistor ad altri tipi di logica come CMOS.

7.2 Circuiti rettificatori

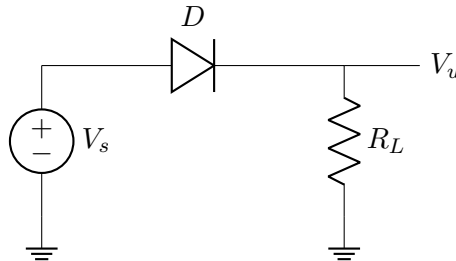
Un'applicazione tipica dei diodi è quella della **rettificazione** della tensione. La rettificazione è un'operazione che porta una corrente alternata in una corrente continua, o comunque porta una tensione che oscilla attorno ad uno 0 di potenziale a stare tutta sopra tale 0.

Circuiti di questo tipo sono utili ad esempio per ottenere generatori di tensione stabili dalla comune alimentazione casalinga (311 V di picco, in Italia, dal valore efficace

220 V moltiplicato per $\sqrt{2}$), che non solo è in corrente alternata sinusoidale, ma è anche piena di disturbi ed interferenze date dalla rete elettrica.

7.2.1 Circuito rettificatore ideale

Il circuito rettificatore più semplice che possiamo disegnare è il seguente:



Di questo circuito notiamo 2 componenti fondamentali:

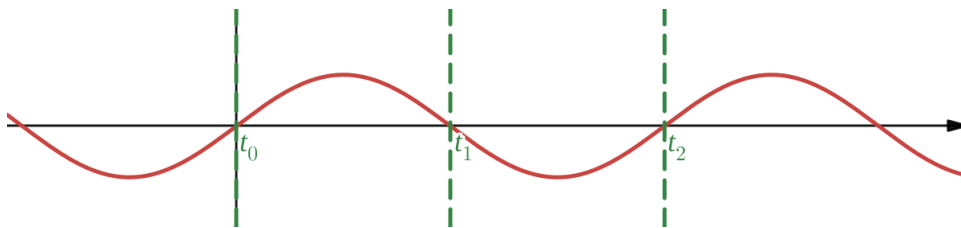
- Il generatore di tensione V_s . Questo equivarrà alla nostra sorgente da rettificare (ad esempio la comune presa di casa). Possiamo dire che il voltaggio di V_s è descritto da:

$$V_s = V_m \sin(\omega t), \quad \omega = 2\pi f$$

ricordiamo che per la presa di casa V_m vale circa 311 V, mentre f vale circa 50 Hz;

- La resistenza R_L , che prendiamo come resistenza di *carico*, da *L*, "load". Questa rappresenterà l'applicazione effettiva che sfrutterà il voltaggio rettificato dal diodo *D*.

Per studiare il circuito prendiamo due range temporali diversi, calcolati a partire dall'espressione V_s : uno dove questa è positiva, e l'altro dov'è negativa.



- Il primo range sarà dato da $t_0 = 0$ a t_1 , inteso come il punto dove V_s interseca per la prima volta l'asse delle ascisse. Questo sarà quindi il range dove V_s è *positiva*;
- Il primo range sarà dato da t_1 (calcolato prima) a t_2 , inteso come il punto dove V_s interseca per la seconda volta l'asse delle ascisse (cioè V_s completa un ciclo). Questo sarà quindi il range dove V_s è *negativa*;

Lo studio del comportamento del circuito per tutti gli altri range temporali sarà quindi derivabile dalla periodicità di V_s (le due situazioni che studieremo si ripeteranno ciclicamente nel tempo).

Vediamo quindi il comportamento del circuito in entrambe le situazioni, preso *D* come diodo *ideale*:

- $t_0 < t < t_1$

In questo caso risulterà che D è posto fra un potenziale positivo all'anodo e il ground al catodo. Segue che molto probabilmente sarà in conduzione, e in tal caso V_u vedrà il voltaggio V_s :

$$V_u = V_s = V_m \sin(\omega t)$$

Verifica

Abbiamo posto D in conduzione, quindi imponiamo corrente positiva:

$$I_D = \frac{V_s}{R_L} > 0$$

per cui l'ipotesi è verificata.

- $t_1 < t < t_2$

In questo caso notiamo che D si trova in polarizzazione negativa, per cui probabilmente sarà in interdizione. In tal caso sul lato destro del circuito non passerà nessuna corrente, e V_u sarà nulla:

$$V_u = 0 \text{ V}$$

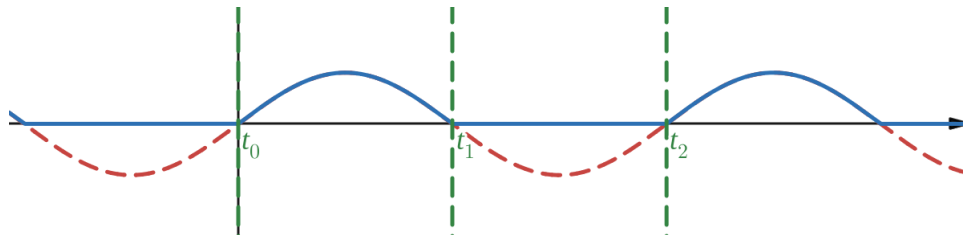
Verifica

Abbiamo posto D in interdizione, quindi imponiamo tensione sotto la soglia di interdizione:

$$V_D = -5 \text{ V} < V_\gamma \approx 0.7 \text{ V}$$

per cui l'ipotesi è verificata.

Abbiamo quindi ottenuto che il circuito rettificatore, almeno nel caso ideale, equivale ad un *cortocircuito* per tensioni positive, e ad un *circuito aperto* per tensioni negative. Il risultato è che la forma d'onda del generatore V_s viene rettificata:

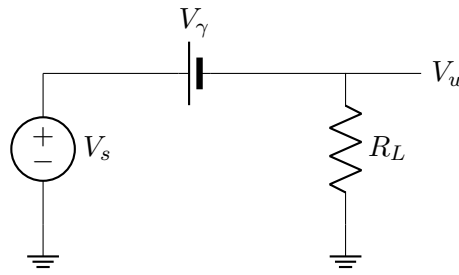


Questo tipo di rettifica viene detta *single phase rectification*, in quanto l'onda sotto lo 0 di potenziale viene "troncata" completamente.

7.2.2 Circuito rettificatore a caduta costante

Vediamo adesso cosa accade se studiamo il circuito rettificatore sfruttando un modello più accurato del diodo di quello ideale, cioè il modello a *caduta costante*. In tal caso notiamo 2 differenze principali:

1. Innanzitutto, quando entra in conduzione, il diodo presenta una caduta di potenziale V_γ (che modellizziamo con un generatore di tensione):



Notiamo che prendiamo V_γ come:

$$V_\gamma \approx 0.7 \text{ V}$$

secondo quanto detto finora. Questo significherà che il voltaggio in uscita V_u , preso D in conduzione, non sarà più equivalente a V_s , ma avremo da considerare la caduta:

$$V_u = V_s - V_\gamma, \quad I_D = \frac{V_s - V_\gamma}{R}$$

e la corrente cambia di conseguenza.

- La seconda differenza diventa chiara nel momento in cui proviamo a svolgere la verifica del caso in conduzione. Dovrà infatti essere che:

$$I_D = \frac{V_s - V_\gamma}{R} > 0 \implies V_s > V_\gamma$$

cioè l'ipotesi di conduzione vale soltanto nel caso in cui V_s riesce a superare la soglia V_γ . Questo significa che avremo due fenomeni sulla forma d'onda di V_u :

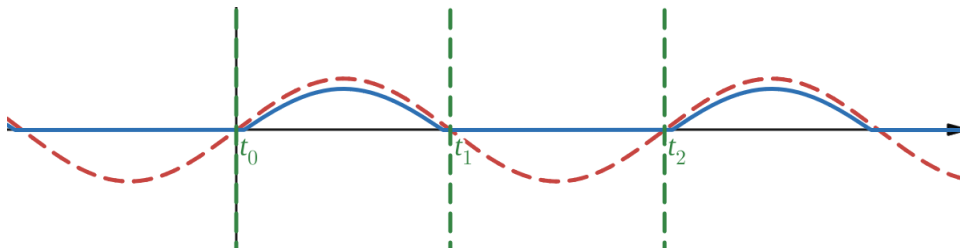
- Un *ritardo di conduzione* sul fronte dell'onda che parte da 0 V;
- Un *anticipo di spegnimento* sul fronte dell'onda che va a ricadere su 0 V.

Ritardo di conduzione ed anticipo di spegnimento sono identici ed uguali a Δt . Calcoliamo tale intervallo notando che vogliamo risolvere:

$$V_s = V_\gamma = V_m \sin(\omega t), \quad \sin(\omega t) = \frac{V_\gamma}{V_m}, \quad \omega t = \sin^{-1} \left(\frac{V_\gamma}{V_m} \right)$$

$$\implies \Delta t = \frac{1}{\omega} \sin^{-1} \left(\frac{V_\gamma}{V_m} \right)$$

Visualizziamo quindi su un grafico il risultato dell'adozione del diodo a caduta costante come modello nel rettificatore.



Abbiamo che i picchi del segnale dopo il rettificatore sono stati ridotti (appunto dalla caduta costante), e abbiamo dei ritardi e degli anticipi della regione in cui il segnale viene troncato al di sopra di 0 V (in sostanza passiamo più tempo vicino agli 0 V).

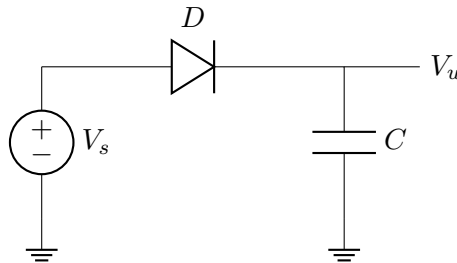
8 Lezione del 17-03-26

Nella scorsa lezione abbiamo visto come usare il diodo in alcuni circuiti elementari, fra cui i *rettificatori*. In particolare, finora abbiamo considerato il comportamento dei rettificatori sottoposti ad un carico resistivo R_L , secondo due modelli (ideale e caduta costante).

Abbiamo però che siamo riusciti soltanto a realizzare un'operazione di rettifica *single phase*, mentre l'obiettivo che abbiamo quando rettifichiamo un segnale è probabilmente quello di portarlo a essere il più vicino possibile al DC (come avevamo già detto in 7.2, rettificare l'alimentazione casalinga). Per fare ciò, vediamo cosa succede se si sottopone il rettificatore ad un elemento di filtraggio *reattivo*.

8.0.1 Circuito rettificatore con condensatore

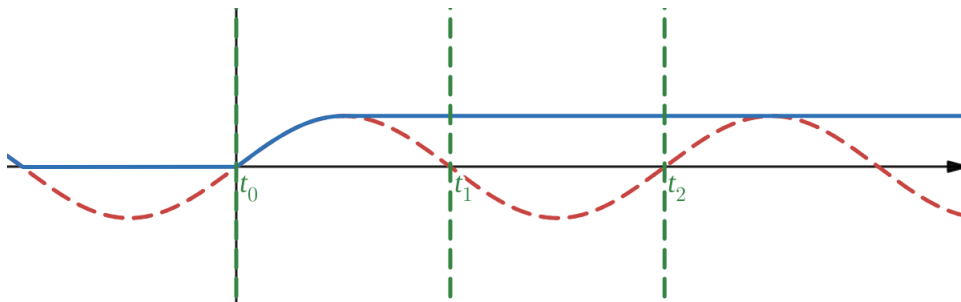
Vediamo quindi cosa succede se mettiamo a valle del rettificatore un *condensatore*:



Quello che accadrà in questo caso è che il condensatore, di equazione:

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

ci porterà in una situazione dove:



1. Da t_0 a $\frac{t_1}{2}$ (considerati gli estremi della scorsa sezione, $\frac{t_1}{2}$ è il picco di tensione) la tensione V_s sarà positiva, e lo sarà anche la V_u , che la inseguirà con i vari dettagli che avevamo detto (una piccola caduta di potenziale e il ritardo di conduzione). Notiamo che questo processo non andrà solo ad aumentare V_u , ma anche a caricare il condensatore C ;
2. Da $\frac{t_1}{2}$ in poi, per via della capacità C (ormai carica) che vorrà mantenere il voltaggio costante, avremo che il voltaggio V_u sarà pressappoco costante su t .

In questo caso ideale, il valore della capacità C non sarà particolarmente significativo finché questa verrà caricata da V_s abbastanza in fretta da raggiungere il picco in $\frac{t_1}{2}$.

Questo perché non ci sarà nessun elemento resistivo (nel caso ideale) che potrà fungere da "sfogo" per la capacità, che andrebbe alternativamente a scaricarsi (cosa che vedremo fra poco).

Per fissare il meccanismo rispetto all'analisi del comportamento dei diodi, ipotizziamo che il diodo si trovi in conduzione finché:

$$I_D > 0 \implies \frac{dV_u}{dt} > 0$$

e quindi la V_u caricherà come V_s fino a $\frac{t_1}{2}$. Dopo aver attraversato il primo picco, il condensatore non caricherà più, il diodo si trasformerà in un aperto, e quindi avremo:

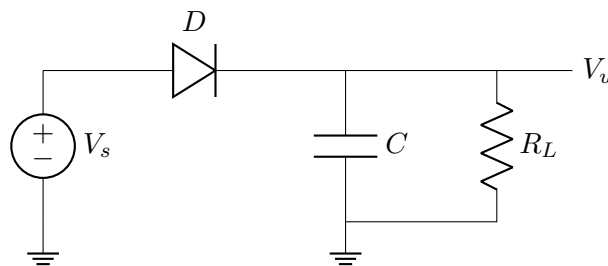
$$I_C = I_D = 0$$

Cioè da $\frac{t_1}{2}$ in poi il diodo sarà sempre in inversa, meno i punti di cresta dove starà alla soglia di conduzione ma a corrente nulla.

Questo ci sarà utile se il nostro obiettivo era quello di trasformare una tensione alternata in una tensione continua, cioè effettivamente rettificarla ad una sorgente di tensione costante nel tempo. Notiamo però che non abbiamo totalmente risolto il problema, in quanto potrebbero sempre esserci irregolarità nella retta da $\frac{t_1}{2}$ in poi date dalla scarica del condensatore o dal rumore del segnale V_s .

8.0.2 Circuito rettificatore con filtro RC

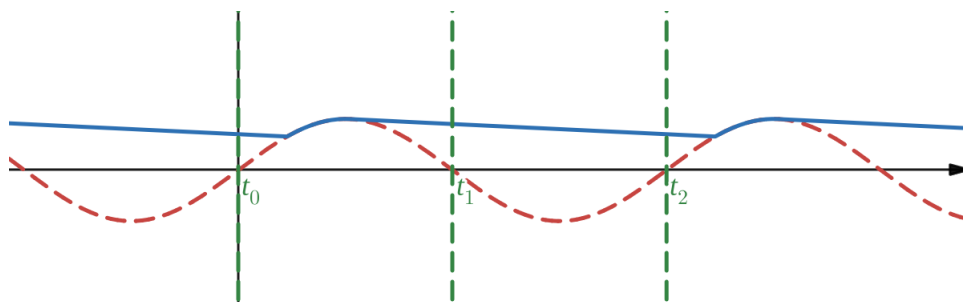
Vediamo nel dettaglio cosa intendiamo. Se reintroduciamo una resistenza di carico, in parallelo al condensatore, otteniamo la seguente configurazione:



In questo caso il condensatore, da $\frac{t_1}{2}$ in poi, avrà modo di scaricarsi su R_L . In altre parole, quando il diodo entrerà in interdizione, la tensione sul condensatore tenderà a scaricarsi sulla resistenza seguendo l'equazione tipica del circuito RC:

$$V_u \propto e^{-\frac{t}{\tau}}, \quad \tau = RC$$

Ciò significherà che esisterà un istante, precedente al t_0 del ciclo successivo, dove V_s salirà sopra il livello raggiunto da V_u , ricaricherà il condensatore (alzando il voltaggio nuovamente fino al valore di picco), permettendo che il ciclo continui.

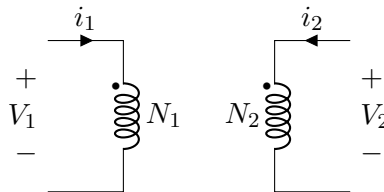


Chiamiamo questo voltaggio di caduta **ripple voltage** del rettificatore.

8.0.3 Circuito rettificatore a doppia semionda

L'idea che abbiamo guardando al grafico della scorsa sezione è che, se si potessero combinare i contributi di due diodi (uno in una direzione, e uno nell'altra, rispetto a V_s), allora si potrebbe "riempire" il calo di tensione successivo a $\frac{t_1}{2}$ con un contributo in tensione dato dal diodo invertito.

Per formalizzare quello che si è detto, introduciamo un componente detto **trasformatore**, dal simbolo:



Questo è un componente magnetico che varia il valore efficace della tensione alternata tramite il *rapporto spire*:

$$n = \frac{N_1}{N_2}$$

preservando la potenza ideale tra avvolgimento *primario* e *secondario*.

Prenderemo in considerazione soltanto i trasformatori ideali, per i quali il flusso magnetico totale può essere considerato trascurabile, per cui:

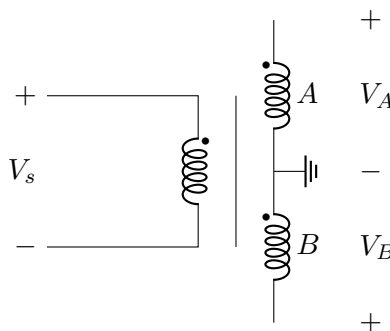
- La forza **magnetomotrice** totale dovrà essere nulla per mantenere un flusso finito (riluttanza ideale nulla):

$$N_1 i_1 + N_2 i_2 = 0 \implies \frac{i_2}{i_1} = -\frac{N_1}{N_2} = -n$$

- Vale la legge di **Faraday** alle tensioni:

$$V_1 = N_1 \frac{d\phi}{dt}, V_2 = N_2 \frac{d\phi}{dt} \implies \frac{V_1}{V_2} = \frac{N_1}{N_2} = n$$

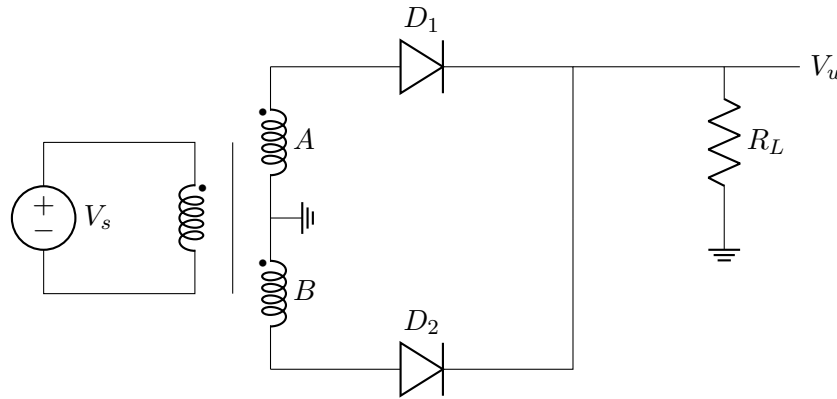
Notiamo che ciò di cui abbiamo effettivamente bisogno è di un trasformatore a **presa centrale** (o a *massa centrale*, o ancora in inglese *center tap*), cioè un trasformatore che ha il punto centrale di uno degli avvolgimenti ancorato a massa.



Questo significa che le due estremità che resteranno esposte (V_A e V_B) risulteranno, sottoposte al voltaggio V_s sull'altro avvolgimento, alternativamente in tensione.

Una caratteristica fondamentale da notare del trasformatore a presa centrale è che i voltaggi V_A e V_B visti dai due estremi saranno la metà del voltaggio che avremo sull'intero trasformatore (preso $n = 1$, di V_s). Per avere un comportamento simile ad un trasformatore normale, quindi, converrà prendere qualcosa come $n \approx 0.5$.

Vediamo quindi com'è disposto il circuito rettificatore a *doppia semionda* che sfrutta il trasformatore a massa centrale:



Quello che otterremo in questa situazione è che durante il primo semiperiodo, cioè da t_0 a t_1 , si avrà $V_s > 0$, e quindi D_1 sarà in conduzione e D_2 in interdizione. Durante il secondo semiperiodo, invece, cioè da t_1 a t_2 , si avrà $V_s < 0$, e quindi D_2 sarà in conduzione e D_1 in interdizione.

La verifica si ha analizzando i due semiperiodi:

- $t_0 < t < t_1$

Per il primo semiperiodo, poniamo la condizione di conduzione a D_1 e di interdizione a D_2 , per cui la corrente scorre in senso orario sulla maglia formata dai diodi.

Verifica La verifica si fa imponendo la condizione di conduzione a D_1 e di interdizione a D_2 :

$$I_{D_1} = \frac{V_A}{R_L} > 0, \quad V_{AK2} = -V_B - V_A < 0$$

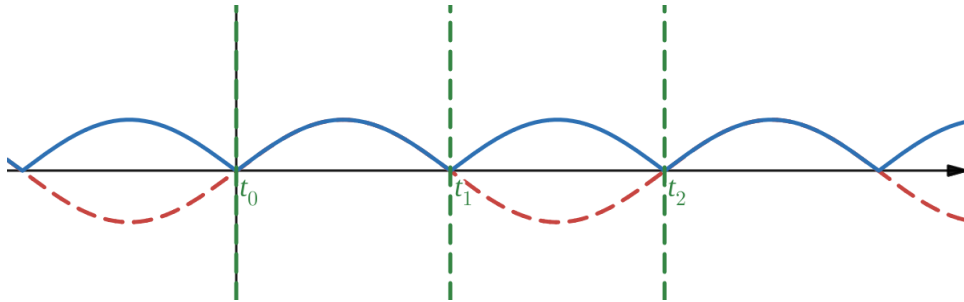
- $t_1 < t < t_2$

Per il secondo semiperiodo, poniamo la condizione di conduzione a D_2 e di interdizione a D_1 , per cui la corrente scorre in senso antiorario sulla maglia formata dai diodi.

Verifica La verifica si fa imponendo la condizione di conduzione a D_2 e di interdizione a D_1 :

$$I_{D_2} = -\frac{V_B}{R_L} > 0, \quad V_{AK1} = V_A - (-V_B) < 0$$

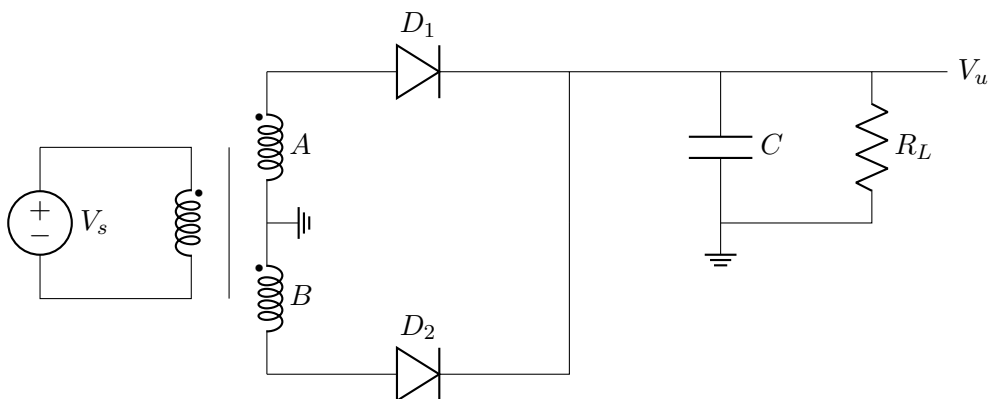
Il risultato sulla forma d'onda è che otteniamo un valore non rettificato in *single phase*, ma in qualche modo il valore assoluto del segnale da V_s , cioè:



Chiamiamo questo tipo di rettifica *full wave rectification*, per distinguerla appunto dalla *single phase rectification* vista in 7.2.1. Abbiamo quindi riempito il secondo semiperiodo come ci eravamo prefissi di fare.

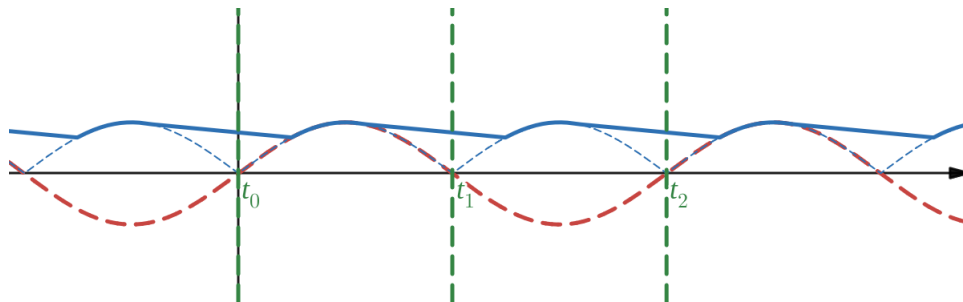
8.0.4 Circuito rettificatore a doppia semionda con filtro RC

Concludiamo reintroducendo il condensatore C per "stabilizzare" in qualche modo il voltaggio. Non facciamo il caso con solo C , in quanto sarebbe banale ed uguale al caso senza doppia semionda, e passiamo direttamente al filtro RC:



Quello che otteniamo è quindi l'effetto combinato della doppia semionda e del condensatore. Anche in questo caso, quando il diodo entrerà in interdizione, la tensione sul condensatore tenderà a scaricare sulla resistenza seguendo l'equazione tipica del circuito RC. Ciò significherà che gli istanti dove V_s salirà sopra il livello raggiunto da V_u e quindi ricaricherà il condensatore (alzando il voltaggio nuovamente fino al valore di picco) accadranno con frequenza doppia (in corrispondenza sia di t_0 che di t_1 , più opportune fasi).

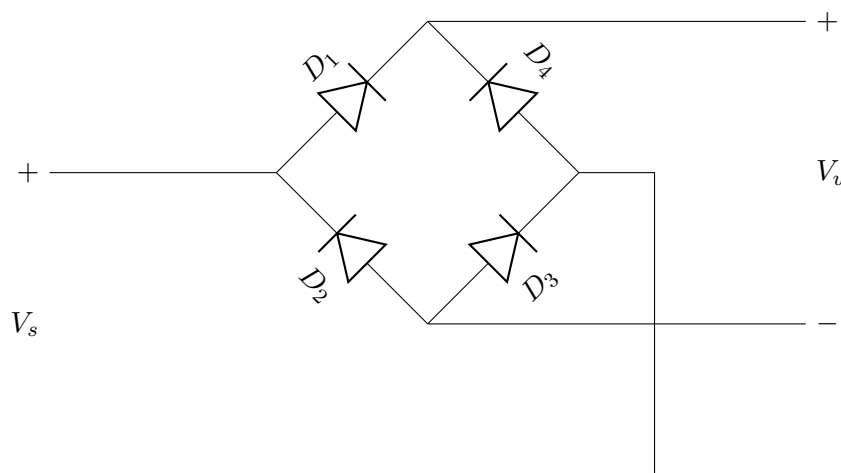
Questo ci permetterà di ridurre al minimo l'effetto del ripple voltage, portando ad una forma d'onda che è molto vicina all'essere costante. Chiaramente, all'aumentare di C "appiattiamo" sempre più la forma d'onda.



8.0.5 Rettificatore a ponte di Graetz

Una soluzione alternativa alla doppia semionda col trasformatore (che chiamiamo anche *presa centrale*) è data dall'utilizzo di una particolare configurazione di 4 diodi, detta ponte di *Graetz*. Questo ci permetterà di risparmiare sul trasformatore a presa centrale (ingombrante e costoso).

Il circuito del ponte di Graetz è il seguente:



Abbiamo che a sinistra introduciamo una corrente *alternata* V_s e a destra estraiamo una corrente *continua* V_u (o almeno rettificata in *full wave*, vedremo in seguito metodi analoghi a prima per stabilizzare fino a DC).

Studiamo quindi il circuito nei due regimi, cioè nei due semiperiodi:

- $t_0 < t < t_1$

Nel primo semiperiodo, i diodi D_1 e D_3 sono in conduzione, mentre i diodi D_2 e D_4 sono in interdizione. Questo significa che avremo percorsi chiusi dal + di V_s al + di V_u , e dal - di V_s al - di V_u .

Prenderemo quindi la parte positiva del segnale V_s , con:

$$V_u = V_s$$

Verifica La verifica per i diodi D_1 e D_3 si fa, ponendo una certa resistenza di carico R_L su V_u , notando che si va a formare un circuito chiuso tale che:

$$I = \frac{V_s}{R_L}$$

che rispetta la condizione di conduzione.

Per gli altri diodi dovremo imporre le condizioni di interdizione sul voltaggio:

$$V_{AK2} = -V_s, \quad V_{AK4} = -V_s$$

dalle considerazioni fatte prima.

- $t_1 < t < t_2$

Nel secondo semiperiodo, saranno invece i diodi D_2 e D_4 sono in conduzione, mentre i diodi D_1 e D_3 sono in interdizione. Questo significa che avremo percorsi chiusi dal $-$ (ora positivo) di V_s al $+$ di V_u , e dal $+$ di V_s (ora negativo) al $-$ di V_u .

Prenderemo quindi la parte negativa rettificata del segnale V_s , con:

$$V_u = -V_s$$

Verifica La verifica per i diodi D_2 e D_4 si fa come prima ponendo una certa resistenza di carico R_L su V_u , notando che si va a formare un circuito chiuso tale che:

$$I = \frac{V_s}{R_L}$$

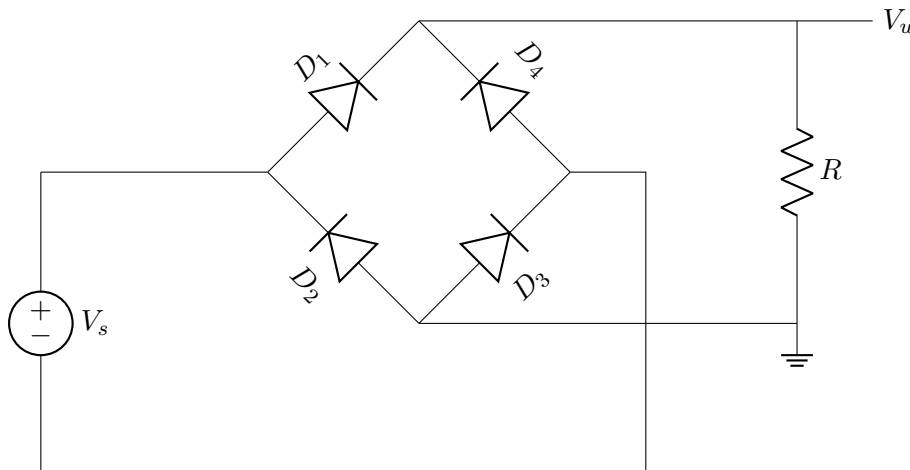
che rispetta la condizione di conduzione.

Per gli altri diodi dovremo imporre le condizioni di interdizione sul voltaggio:

$$V_{AK1} = -V_s, \quad V_{AK3} = -V_s$$

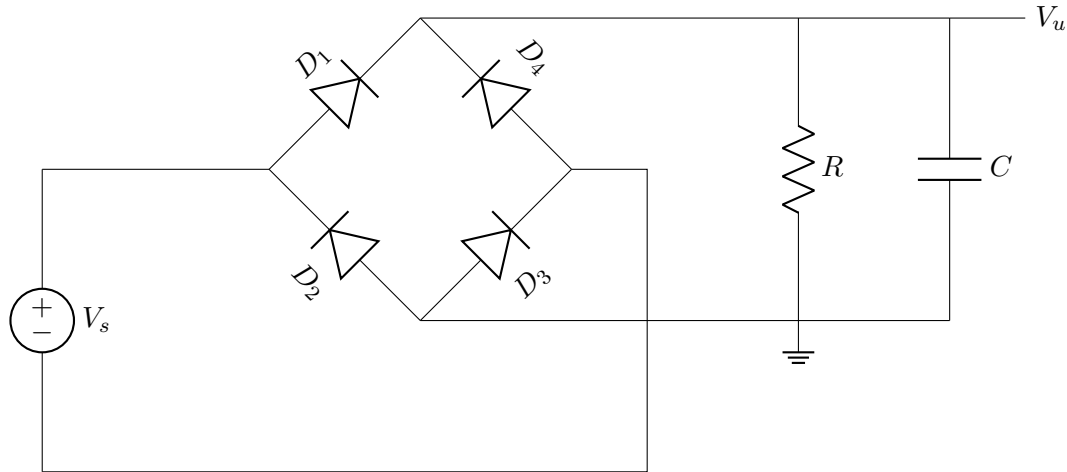
dalle considerazioni fatte prima.

Il risultato è quindi che abbiamo ottenuto un circuito esattamente uguale al rettificatore a doppia semionda visto in 8.0.3, ma risparmiandoci il trasformatore a presa centrale. Anche il grafico della forma d'onda sarà lo stesso di 8.0.3. Nello specifico, tracciamo il circuito come:



8.0.6 Rettificatore a ponte di Graetz con filtro RC

Come per il circuito a rettificatore a trasformatore a presa centrale, il ponte di Graetz può essere migliorato introducendo uno stage di filtraggio RC che stabilizzi la tensione V_u in uscita:



Anche in questo caso, questa soluzione ci permetterà di ridurre al minimo l'effetto del ripple voltage, portando ad una forma d'onda che è molto vicina all'essere costante. Chiaramente, all'aumentare di C "appiattiamo" sempre più la forma d'onda. Il grafico della forma d'onda sarà lo stesso di 8.0.4.

8.0.7 Confronto fra presa centrale e ponte di Graetz

Confrontiamo le due soluzioni che abbiamo dato per la rettifica del segnale.

- La soluzione a **presa centrale** per la rettifica della tensione richiede un trasformatore ingombrante e costoso, di cui solo metà avvolgimento per volta è effettivamente in operazione;
- La soluzione a **ponte di Graetz** invece utilizza 4 diodi che lavorano in coppia per risolvere lo stesso problema. Solitamente i diodi sono compatti e leggeri rispetto al trasformatore della presa centrale, per cui si ottiene un circuito più compatto e più leggero.

Notiamo quindi in termini numerici quali sono le differenze principali:

Caratteristica	Presa centrale	Ponte di Graetz
Diodi necessari	2	4
PIV (Tensione Inversa)	$2V_s$	V_s
Ingombro/Peso	Elevato (Ferro/Rame)	Ridotto (Efficiente)
Caduta di tensione	$1 \times V_\gamma$	$2 \times V_\gamma$

- Il numero di *diodi necessari* rappresenta una metrica ovvia: il ponte di Graetz è più costoso in termini di diodi. Questo è però giustificato dal fatto che il trasformatore a presa centrale è solitamente ingombrante e costoso;

- La **PIV**, o *tensione inversa*, è la tensione inversa a V_s necessaria a destabilizzare il rettificatore. Abbiamo che questa è doppia nel rettificatore a presa centrale, dove dobbiamo vincere entrambi gli estremi di avvolgimenti, mentre è unitaria per il rettificatore a ponte;
- Anche l'*ingombro/peso* del componente rettificatore finale rappresenta una metrica triviale, che abbiamo già discusso notando i vantaggi del ponte a diodi rispetto al trasformatore con tap;
- Infine, notiamo che il trasformatore è leggermente più efficiente per quanto riguarda le cadute di potenziale $V_\gamma \approx 0.7\text{ V}$ sui diodi. Questo è chiaro dal fatto che nel primo caso dobbiamo passare attraverso al più un diodo in conduzione, mentre nel ponte si passa sempre per 2 diodi in conduzione (raddoppiando la caduta complessiva).

In conclusione, il ponte di Graetz è lo standard moderno, in quanto il costo di due diodi extra è trascurabile rispetto al risparmio su peso e dimensioni del trasformatore.

Detto questo, il ponte di Graetz non è perfetto: a parte il problema del costo aggiunto dei 4 (2 in più rispetto alla presa centrale) diodi, abbiamo che l'uscita di un ponte di Graetz è unidirezionale, ma fortemente pulsante. Questo significa che si ha comunque un effetto consistente del ripple voltage. Abbiamo che questa fluttuazione di tensione (ondulazione) è inaccettabile per la maggior parte dei circuiti elettronici. Ad esempio:

- Nei circuiti *audio*, il ripple voltage comporta un ronzio a 100 Hz (in Europa, 120 Hz negli Stati Uniti) negli altoparlanti;
- Nei circuiti *digitali* si potrebbero avere errori logici: microprocessori che lavorano ad alta frequenza potrebbero riavviarsi se sottoposti a variazione della tensione di alimentazione;
- Si potrebbero avere errori di precisione nei circuiti di *misura*, ovvero imprecisioni dei sensori utilizzati.

La strategia per risolvere questo problema sarà data in due passaggi:

1. **Filtraggio:** Adottiamo un condensatore per "livellare" le valli di tensione (e questo lo abbiamo già visto, introducendo il filtro RC nei rettificatori studiati);
2. **Regolazione:** Introduciamo un diodo *Zener* (visti in 6.1.1) per rendere l'uscita immune alle variazioni del carico. Avevamo già introdotto, infatti, come un diodo di Zener in regime di breakdown può comportarsi come un generatore di tensione costante.

8.0.8 Rettificatore con diodo Zener

Vediamo nello specifico come i diodi **Zener** vengono usati nei rettificatori. Abbiamo che questi sono, di base, componenti con le seguenti caratteristiche:

- A differenza dei diodi comuni, il diodo Zener è progettato per lavorare stabilmente nella regione di rottura (*breakdown*) senza danneggiarsi;
- In questa zona, la curva *i-v* diventa quasi verticale: grandi variazioni di corrente corrispondono a variazioni minime di tensione.

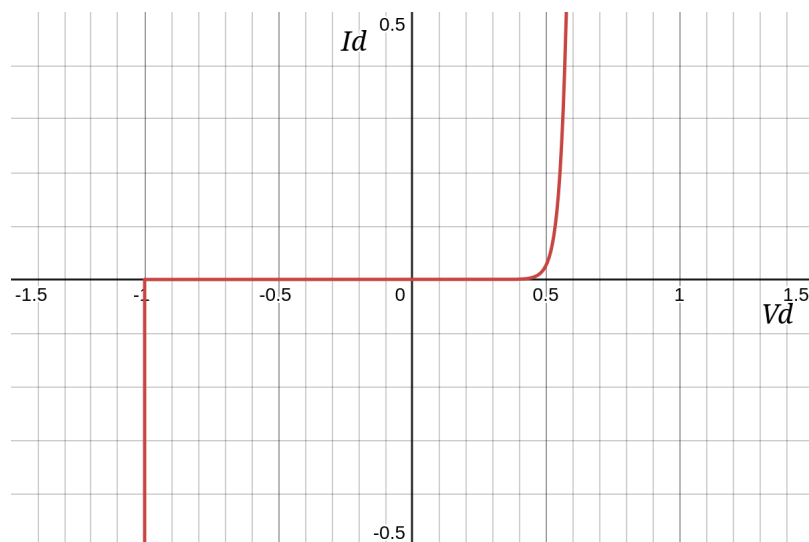
Ciò significa che il diodo Zener si comporta come segue:

- Si comporta come un normale diodo in polarizzazione diretta.
- In polarizzazione inversa, blocca la corrente fino a V_Z .
- Superata V_Z , la tensione ai capi del diodo resta circa costante a V_Z .

Questo significa che possiamo utilizzare il diodo Zener per effettuare 2 operazioni principali:

- **Riferimento:** il diodo in regime di breakdown fornisce una tensione costante e nota;
- **Protezione:** introducendo un diodo Zener possiamo imitare i picchi di tensione.

Avremo bisogno di modificare il modello a caduta tensione che avevamo introdotto per il diodo, introducendo la zona di breakdown per voltaggi negativi.



Il modello a caduta costante verrà quindi variato introducendo le 3 regioni operative:

1. Conduzione in polarizzazione diretta:

$$V_{AK} = V_{\gamma}, \quad \text{verificare } I_D \geq 0$$

In questa zona il diodo si comporta da generatore di tensione in polarizzazione V_{γ} positiva;

2. Interdizione:

$$I_D = 0, \quad \text{verificare } -|V_Z| < V_{AK} < V_{\gamma}$$

In questa zona il diodo si comporta come un aperto;

3. Conduzione in polarizzazione inversa (breakdown):

$$V_{AK} = -|V_Z|, \quad \text{verificare } I_D \leq 0$$

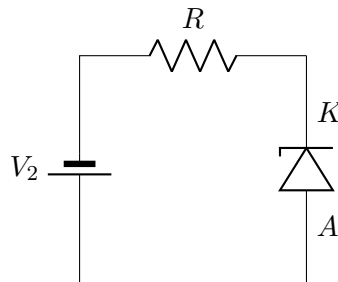
Questa è la zona di funzionamento utilizzata per il regolatore. In questa zona il diodo si comporta da generatore di tensione in polarizzazione V_z negativa; caso

9 Lezione del 20-03-26

Veniamo a come regolare la tensione usando i diodi Zener.

9.0.1 Lavorare in breakdown

Poniamo il seguente circuito, formato da un diodo di zener sottoposto ad un certo voltaggio V_2 in polarizzazione inversa. Viene da sé che questo verrà messo a valle del circuito rettificatore filtrato RC che abbiamo visto alla fine della scorsa lezione.

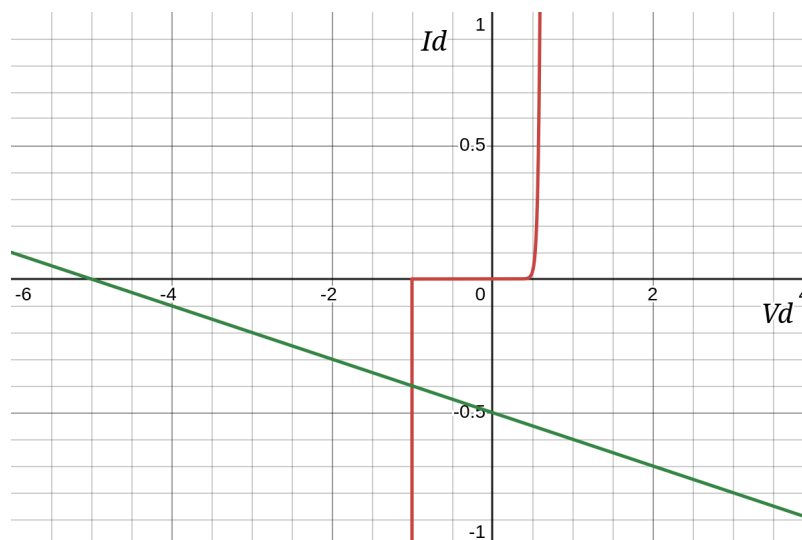


In questo caso V_2 sarà una tensione con una componente DC, che vogliamo isolare dal resto del segnale (rumore, e nel contesto dei rettificatori di cui parlavamo alla scorsa lezione, *ripple voltage*).

Prendiamo quindi la retta di carico (considerata per la prima volta in 6.1.3 nello studio del diodo), che sarà:

$$I_D = -\frac{V_2 + V_D}{R}$$

in quanto il diodo zener è in polarizzazione negativa rispetto a V_2 . Abbiamo quindi una retta la cui intercetta alle ascisse è variata da V_2 , mentre il coefficiente angolare è variato da R . Su un grafico, essa appare come:



cioè per qualsiasi valore di V_D negativa oltre la V_Z di breakdown imposta sullo zener, questo esprimerà una corrente costante (pari a V_Z).

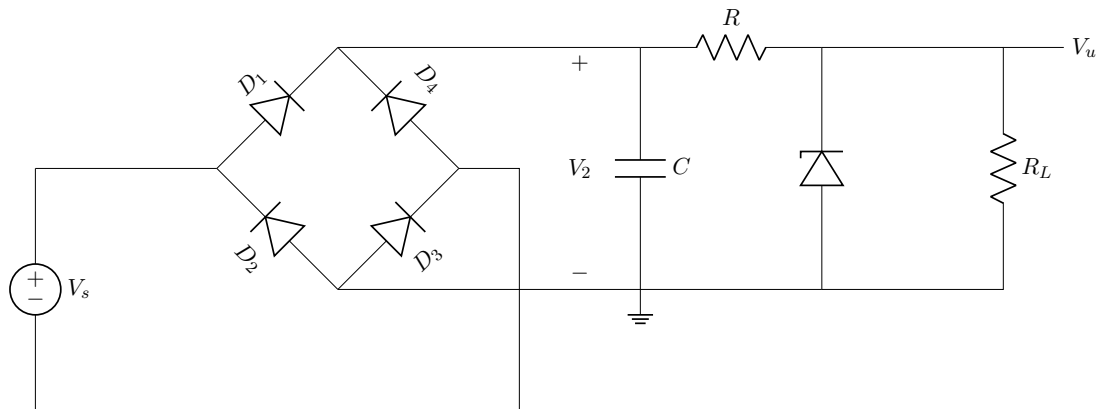
Per facilitare l'analisi, indichiamo $I_Z = -I_D$. Notiamo che, come sempre, vale:

$$P_Z = V_Z I_Z$$

Sarà quindi che la potenza dissipata sul diodo è il prodotto fra la corrente che ci passa e la tensione dissipata. Usando diodi zener, che lavorano in regioni dove la corrente può crescere anche molto, dobbiamo stare attenti a mantenere P_Z entro le soglie di operatività dello zener. Altrimenti si avrebbero aumenti della temperatura, che da quanto discusso parlando delle giunzioni PN sappiamo ci porteranno a fallimento del diodo.

9.0.2 Rettificatore con regolatore di tensione

Possiamo quindi completare il design del nostro rettificatore introducendo il circuito di regolazione della tensione che abbiamo appena studiato. Questo porterò al seguente schema:



Lo scopo del componente che abbiamo aggiunto, cioè la coppia R (che si va a differenziare dalla R_L di carico) e diodo di zener, è quello di regolare la tensione V_2 (quella in uscita dal ponte di Graetz filtrato RC, cioè la tensione sulla capacità C).

Per studiare tale circuito facciamo l'ipotesi che lo zener lavori in breakdown, cioè:

$$I_Z > 0 \implies V_u = V_Z$$

e quindi la tensione in uscita dal circuito V_u è pari alla tensione di breakdown V_Z del diodo.

Verifica

Questo si verifica calcolando la corrente sullo zener, in modo da verificare che questo si trovi in regime di breakdown (corrente negativa):

$$I_R = \frac{V_2 - V_Z}{R}, \quad I_L = \frac{V_Z}{R_L} \implies I_Z = I_R - I_L = \frac{V_2 - V_Z}{R} - \frac{V_Z}{R_L} > 0$$

dove I_R è la corrente sulla resistenza R diretta verso destra, I_L è la corrente sulla resistenza di carico R_L diretta verso il basso. Da quanto detto prima rispetto allo zener, $I_Z = I_R - I_L$ è un'affermazione della prima legge di Kirchoff.

Abbiamo quindi imposto la $I_Z > 0$, che era la nostra ipotesi di funzionamento del diodo in regime di breakdown. Nel caso in cui la disequazione non sia soddisfatta, significherà che non siamo riusciti a portare lo zener nel regime di breakdown, e quindi la tensione V_u che misuriamo è nulla.

Scelta della resistenza

Interrogiamoci su come si andrà a scegliere la R nel circuito di regolazione a diodo di zener. Questa resistenza è proprio quella che sta a limitare la $P_Z = V_Z I_Z$ dello scorso

paragrafo, e quindi la corrente che scorre sul diodo. Vogliamo mantenere questa limitata in modo che il diodo operi nei suoi regimi validi.

Abbiamo che la tensione a valle del ponte di Graetz filtrato RC è soggetta a ripple, per cui possiamo dire:

$$V_{2,\min} \leq V_2 \leq V_{2,\max}$$

Nel caso peggiore, il carico non richiede corrente, per cui $R_L \rightarrow +\infty$, o anzi la sostituiamo direttamente con un aperto. In tal caso tutta la corrente si trova sul zener:

$$I_{Z,\max} = \frac{V_{2,\max} - V_Z}{R}$$

Dal datasheet dello zener avremo una certa potenza $P_{Z,\max}$, che è la *massima* tollerabile dal diodo. Imporremo quindi che tale potenza non venga superata:

$$V_Z I_{Z,\max} < P_{Z,\max} \implies R^* > \frac{(V_{2,\max} - V_Z) V_Z}{P_{Z,\max}}$$

cioè R^* sarà la resistenza *minima* che potremmo mettere nel circuito regolatore in modo da evitare che questo bruci.

Al carico R_L , invece, quando questo richiede corrente, avremo al limite $I_Z = 0$, per cui potremo offrire almeno la corrente I_L :

$$I_{L,\max} = \frac{V_{2,\min} - V_Z}{R}$$

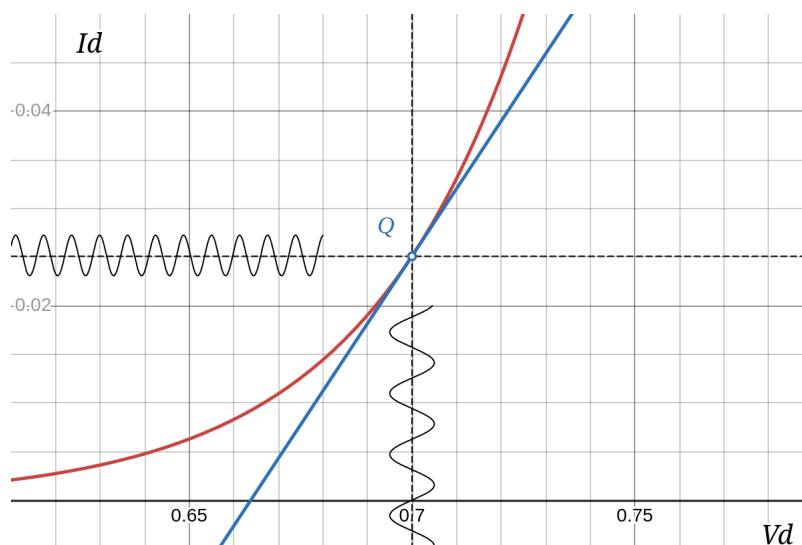
9.1 Modello del diodo ai piccoli segnali

Finora abbiamo trattato il diodo ai grandi segnali, cioè per cadute di tensione $\Delta V \gg V_\gamma \approx 0.7\text{ V}$ o comunque confrontabili. Vediamo invece come trattare il diodo per i piccoli segnali, cioè quelli $\Delta V \ll V_\gamma$.

In tal caso riprenderemo la legge di Shockley, vista in 6.1, e quindi l'andamento esponenziale della corrente sul diodo:

$$I_D = I_S \left(e^{\frac{V_D}{\eta V_T}} - 1 \right)$$

dove I_S era la corrente di saturazione, η il fattore di idealità e V_T la tensione termica. Per un piccolo segnale è inutile prendere l'intero esponenziale, e si può invece prendere un punto centrale Q e linearizzare la risposta attorno ad esso.



Più nel dettaglio, se le variazioni del segnale ($v_d(t)$) sono sufficientemente piccole rispetto alla tensione termica (V_T), la curva esponenziale del diodo può essere approssimata con un segmento rettilineo (la tangente nel punto Q , dove Q sta per *Quiescenza*). In questo modo potremo dire che esiste una legge del tipo:

$$\begin{cases} V_D(t) = V_{DQ} + v_d(t) \\ I_D(t) = I_{DQ} + i_d(t) \end{cases}$$

per tensione e corrente sul diodo in relazione a Q .

Questo ci dà due vantaggi:

1. **Semplificazione** del circuito: anziché risolvere equazioni trascendentali, cioè la funzione esponenziale, trattiamo il diodo come una semplice resistenza r_d , con:

$$i_d = g_d v_d, \quad v_d = r_d i_d \quad \text{con} \quad r_d = \frac{1}{g_d}$$

dove la g_d (la *conduttività* dell'approssimazione a resistenza) non è altro che la pendenza della retta linearizzante;

2. Inoltre, sfruttare un modello linearizzato ci permette di applicare il principio di **sovrapposizione** degli effetti. In altre parole, in un circuito linearizzato, possiamo separare nettamente l'analisi in continua (che fissa il punto Q di quiescenza) dall'analisi in alternata (che elabora il segnale).

Linearizzazione su Q

Vediamo quindi a come ricavare i valori linearizzati, cioè ricaviamoci espressioni esplicite per V_{DQ} , I_{DQ} e la conduttanza g_d (r_d è facilmente ricavabile da $\frac{1}{g_d}$).

Per fare ciò prendiamo lo sviluppo di Taylor attorno a Q dell'equazione di Shockley, che chiamiamo $S(V_D)$:

$$I_D(t) = S(V_D(t)) = S(Q + v_d(t)) = S(Q) + \frac{\partial}{\partial V_D} S(Q) v_d(t) + \frac{1}{2} \frac{\partial^2}{\partial V_D^2} S(Q) v_d^2(t) + \dots$$

e tronchiamo al termine lineare, cioè:

$$\approx S(Q + v_d(t)) = S(Q) + \frac{\partial}{\partial V_D} S(Q) v_d(t) = I_{DQ} + g_d v_d(t) = I_{DQ} + i_d(t)$$

Abbiamo quindi la forma linearizzata di $I_D(t)$, dove notiamo:

$$I_{DQ} = S(Q), \quad i_d(t) = g_d v_d(t)$$

Allo stesso modo, per come abbiamo scelto Q , potremo dire che semplicemente:

$$V_{DQ} = Q, \quad v_d(t) = r_d i_d(t)$$

In genere vale poi:

$$g_d = \frac{\partial}{\partial V_D} S(Q), \quad r_d = \frac{1}{g_d} = \left(\frac{\partial}{\partial V_D} S(Q) \right)^{-1}$$

chiamate rispettivamente *conduttanza differenziale* e *resistenza differenziale*.

Questa selezione di variabili ci assicura che:

1. Prendiamo il comportamento su $Q \rightarrow S(Q)$, cioè $V_{DQ} \rightarrow I_{DQ}$, e lo fissiamo come contributo costante;
2. Troviamo espressioni dei piccoli segnali $v_d(t)$ e $i_d(t)$ che permettono di trattare il diodo come un resistore a conduttanza g_d ;
3. Sommare il contributo DC I_{DQ} al piccolo segnale $i_d(t)$ dà un'espressione valida per la corrente sul diodo $I_D(t)$, e allo stesso modo V_{DQ} più il piccolo segnale $v_d(t)$ dà un'espressione valida per la tensione sul diodo $V_D(t)$. Queste espressioni valide sono poi le stesse di prima:

$$\begin{cases} I_D(t) = I_{DQ} + i_d(t) \\ V_D(t) = V_{DQ} + v_d(t) \end{cases}$$

e siamo felici.

Collegamento alla legge di Shockley

Smettiamo quindi di parlare della $S(V_D)$, e riprendiamo l'espressione propria della legge di Shockley:

$$I_D = I_S \left(e^{\frac{V_D}{\eta V_T}} - 1 \right)$$

Ciò che vogliamo fare è trovare espressioni per I_{DQ} , V_{DQ} e g_d (come prima, r_d è diretta da g_d) che derivino dai parametri della legge di Shockley (cioè la corrente di saturazione I_S , il fattore di idealità η , e la tensione termica V_T).

- V_{DQ} è banale in quanto è semplicemente la tensione di linearizzazione Q (dovremmo ormai aver capito che Q è una tensione e $V_{DQ} = Q$);
- I_{DQ} si ricava abbastanza facilmente dalla legge di Shockley stessa, per cui:

$$I_{DQ} = I_S \left(e^{\frac{Q}{\eta V_T}} - 1 \right)$$

- Infine, g_d era stata espressa come:

$$g_d = \frac{\partial}{\partial V_D} S(Q) = \frac{\partial}{\partial V_D} I_S \left(e^{\frac{Q}{\eta V_T}} - 1 \right)$$

con S la legge di Shockley. Dobbiamo quindi calcolare questa derivata:

$$\frac{\partial}{\partial V_D} I_S \left(e^{\frac{V_D}{\eta V_T}} - 1 \right) = \frac{I_S}{\eta V_T} e^{\frac{V_D}{\eta V_T}}$$

per cui sostituendo Q e ricordando $Q = V_{DQ}$ si ha:

$$\begin{aligned} g_d &= \frac{\partial}{\partial V_D} S(Q) = \frac{I_S}{\eta V_T} e^{\frac{V_{DQ}}{\eta V_T}} = \frac{I_S \left(e^{\frac{V_{DQ}}{\eta V_T}} - 1 \right) + I_S}{\eta V_T} \\ &= \frac{I_{DQ} + I_S}{\eta V_T} \approx \frac{I_{DQ}}{\eta V_T}, \quad I_{DQ} \gg I_S \end{aligned}$$

Quest'ultima espressione richiede alcune spiegazioni. Innanzitutto, il passaggio dove espandiamo la forma della derivata prima è matematicamente valido (I_S si annulla con il -1). La semplificazione che facciamo in seguito, invece, vale quando $I_{DQ} \gg I_S$, e ricordando che il termine a sinistra della somma è I_{DQ} stesso.

Abbiamo quindi un'espressione per g_d in funzione del punto di linearizzazione. L'espressione di r_d è immediata:

$$g_d = \frac{I_{DQ}}{\eta V_T}, \quad r_d = \frac{\eta V_T}{I_{DQ}}$$

Verifica della validità

Abbiamo trascurato i termini superiori nello sviluppo della serie di Taylor per poter linearizzare attorno a Q . Interrogiamoci riguardo alla validità di questa approssimazione. Di base, questa sarà valida se potremo dire:

$$\left| \frac{\partial}{\partial V_D} S(Q) v_d(t) \right| \gg \left| \frac{1}{2} \frac{\partial^2}{\partial V_D^2} S(Q) v_d^2(t) \right|$$

cioè il termine di primo grado è di molto maggiore rispetto al termine di secondo grado. Riportandoci al dominio delle derivate vere e proprie della legge di Shockley (non riportiamo il calcolo della derivata seconda in quanto tedioso):

$$\frac{I_{DQ} + I_S}{\eta V_T} \gg \frac{I_{DQ} + I_S}{2(\eta V_T)^2} \cdot |v_d(t)| \implies |v_d(t)| \ll 2\eta V_T$$

Abbiamo quindi ricavato una condizione perché la nostra approssimazione sia valida. A temperatura ambiente (27°C) si ha che:

$$|v_d(t)| \ll 57.2 \text{ mV}$$

cioè per segnali con ampiezze di pochi milli-volt la nostra approssimazione (linearizzazione) è valida.

10 Lezione del 24-03-26

Alla fine della scorsa lezione abbiamo visto un modello per il diodo ai *piccoli segnali*, che consisteva nella linearizzazione attorno ad un punto di riposo Q . Questo corrispondeva effettivamente alla separazione del segnale studiato in una parte DC (data dal punto di linearizzazione Q), e una parte AC che studiavamo come piccolo segnale.

Vediamo quindi come sfruttare questo meccanismo per risolvere circuiti.

10.0.1 Risoluzione di circuiti con diodi ai piccoli segnali

Abbiamo quindi che un circuito con un diodo sottoposto ad un piccolo segnale può essere risolto prima considerando la parte DC del segnale secondo il classico modello ai grandi segnali (e.g. caduta costante), e quindi trattando lo specifico piccolo segnale. I due risultati si possono quindi combinare sfruttando il principio di sovrapposizione degli effetti.

Il procedimento complessivo sarà:

1. Risolviamo prima il circuito nel punto di riposo con il modello del diodo per grandi segnali (e.g. grandi segnali) per trovare il comportamento al punto di riposo Q , o nello specifico la coppia (V_Q, I_Q) , di cui ricordiamo:

$$V_{DQ} = Q, \quad I_{DQ} = I_S \left(e^{\frac{Q}{\eta V_T}} - 1 \right)$$

2. Quindi si potrà risolvere il circuito per il piccolo segnale usando il modello del diodo linearizzato, di cui ricordiamo:

$$i_d = g_d v_d, \quad v_d = r_d i_d \quad \text{con} \quad g_d = \frac{I_{DQ}}{\eta V_T}, \quad r_d = \frac{\eta V_T}{I_{DQ}}$$

cioè il diodo viene trattato come una semplice resistenza r_d . Notiamo che g_d (e ugualmente r_d) dipende dai parametri trovati sul punto di riposo, per cui bisogna prima fare l'analisi DC e solo dopo l'analisi AC.

Chiaramente dovremo tenere a mente 2 cose:

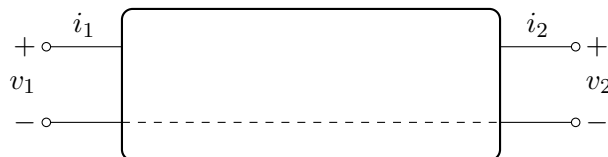
1. Prima di applicare questo processo si deve dividere il segnale agente sul circuito in una coppia DC, AC, cioè una componente costante (diciamo V_{DC}) e una componente che rappresenterà il piccolo segnale (diciamo V_{AC}).

Questo si farà *spengendo* i segnali AC nell'analisi DC, e quindi usando un modello come la caduta costante, e ugualmente spengere i generatori DC nell'analisi AC, usando i parametri trovati dall'analisi AC per determinare i parametri di risposta del circuito ai piccoli segnali.

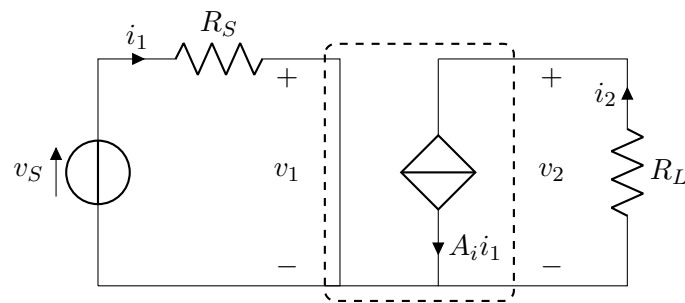
2. Applicando il principio di sovrapposizione degli effetti ad un circuito sostanzialmente non lineare, dovremmo commettere un errore. Questo è però giustificato dal fatto che se il segnale è piccolo, sostituiamo la curva esponenziale con la sua tangente nel punto Q . Una *retta* in questo caso è un modello lineare abbastanza vicino al comportamento reale del componente.

10.1 Transistore bipolare

Iniziamo a vedere il **transistore bipolare** (o **BJT**, *Bipolar Junction Transistor*). Questo è un dispositivo a 2 porte, di cui però colleghiamo i negativi delle porte assieme, formando effettivamente un *tripolo*.



Idealmente viene rappresentato da un *generatore di corrente controllato in corrente* (**GCCC**):



Nel circuito abbiamo rappresentato:

- A sinistra una *sorgente esterna*, cioè v_S , con la sua resistenza associata R_S ;
- A destra una *resistenza di carico*, che indichiamo come R_L .

Di base, un transistor bipolare sarà quindi un circuito che accetta una corrente a sinistra, e impone una corrente pari a questa ma modulata di un certo guadagno A_i (che vedremo subito) a destra (nell'esempio specifico, accetta una corrente imposta dal generatore v_S e fa passare una corrente sul carico R_L).

Considerazioni sui guadagni

Vediamo nel dettaglio cosa rappresenta il parametro A_i . Di un transistor possiamo calcolare il *guadagno di corrente complessivo* A_i come:

$$i_2 = A_i i_1 \implies A_i = \frac{i_2}{i_1} = \frac{i_2}{i_S}$$

visto che la corrente sul generatore v_S è i_1 .

Nel suo uso generico, come abbiamo visto, si accoppia un transistor con una rete esterna che comprende un generatore di segnale di bassa potenza e un carico. L'accoppiamento con la rete esterna (in particolare l'esistenza delle resistenze R_S alla sorgente e R_L al carico) permette di definire un *guadagno di tensione complessivo* A_v :

$$v_2 = -R_L i_2 = -R_L A_i i_1 = -R_L A_i \frac{v_S}{R_S}$$

per cui possiamo dire, analogamente al guadagno di corrente complessivo A_i :

$$A_v = \frac{v_2}{v_S} = -\frac{R_L A_i \frac{v_S}{R_S}}{v_S} = -\frac{R_L}{R_S} A_i$$

Chiamiamo questi guadagni anche *attenuazioni* o *amplificazioni* sulla base dei loro valori:

$$A_i = \frac{i_{out}}{i_{in}} \quad (\text{guadagno di corrente}) \quad \begin{cases} A_i > 1 & (\text{amplificazione di corrente}) \\ A_i < 1 & (\text{attenuazione di corrente}) \end{cases}$$

$$A_v = \frac{v_{out}}{v_{in}} \quad (\text{guadagno di tensione}) \quad \begin{cases} A_v > 1 & (\text{amplificazione di tensione}) \\ A_v < 1 & (\text{attenuazione di tensione}) \end{cases}$$

Considerazioni di potenza

Parliamo quindi della potenza erogata da un lato e dall'altro del transistor. Abbiamo che:

- Sul *carico* R_L , eroghiamo la potenza:

$$P_L = -v_2 i_2$$

- La *sorgente*, di per sé, eroga la potenza:

$$P_S = v_s i_1$$

Possiamo quindi definire un guadagno complessivo anche sulle *potenze* A_P , che sarà uguale a:

$$A_P = \frac{P_L}{P_S} = -\frac{v_2 i_2}{v_s i_1} = -A_i A_v = A_i^2 \frac{R_L}{R_S} = -A_v^2 \frac{R_S}{R_L}$$

Notiamo quindi una prima particolarità dei transistor: i componenti *attivi* (come i **BJT** o i **MOSFET**, *Metal–Oxide–Semiconductor Field-Effect Transistor*) permettono di ottenere valori di $A_P > 1$, cioè ricavare più potenza al carico di quanta ne viene erogata dalla sorgente (chiaramente sotto alimentazione esterna).

Questo è il primo caso di utilizzo del transistor, cioè come *amplificatore* di potenza.

10.1.1 Modello fisico del transistor

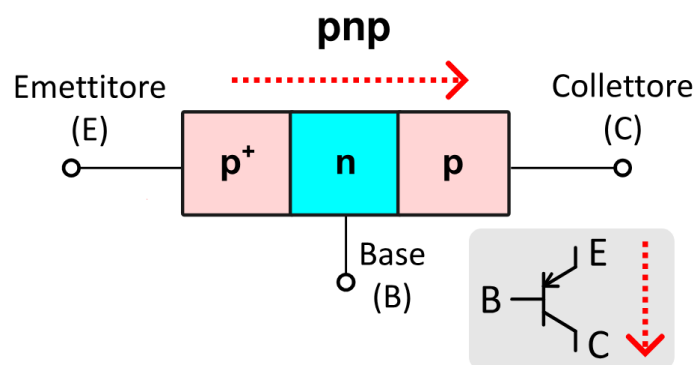
Finora abbiamo trattato il transistor come un componente idealizzato, che si comportava come un GCCC, e abbiamo quindi calcolato i valori di guadagno in *corrente*, *tensione* e *potenza*.

Adesso vediamo come, dal punto di vista fisico, si realizza un transistor a giunzione bipolare. Innanzitutto, *bipolare* significa due giunzioni PN poste una di seguito all'altra e orientate in senso inverso.

Per questo motivo esistono 2 tipi principali di transistor bipolare: il transistor **PNP** (che lega una giunzione PN ad una giunzione NP) e il transistor **NPN** (che lega una giunzione NP ad una giunzione PN):

- **Transistore PNP**

Viene formato da una giunzione PN e una giunzione NP:

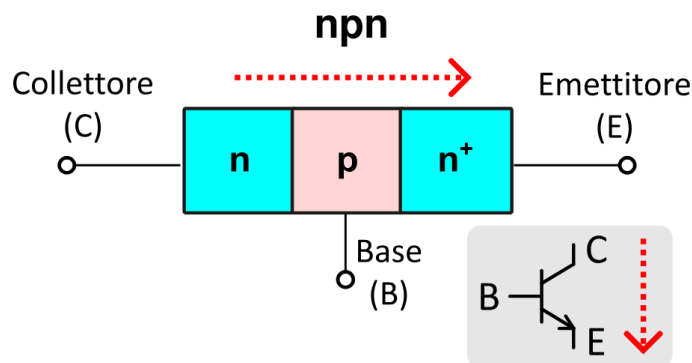


Nel transistor PNP abbiamo quindi una regione P^+ fortemente drogata (da cui il + ad apice), una regione N leggermente drogata, e una regione P meno drogata della prima regione P^+ .

I poli connessi a sinistra e destra sono detti *emettitore* (E) e *collettore* (C). La regione N è ancorata alla *base* B. Il grafico in basso a destra mostra la direzione della corrente tipica (da emettitore a collettore) del transistor PNP;

• **Transistore NPN**

Viene formato da una giunzione NP e una giunzione PN:



Nel transistore NPN abbiamo quindi una regione N, una regione P leggermente drogata, e una regione N⁺ più drogata della prima regione N.

I poli connessi a sinistra e destra sono stavolta *collettore* (C) ed *emettitore* (E). Stavolta la regione P è ancorata alla *base* B. Il grafico in basso a destra mostra la direzione della corrente tipica (da collettore ad emettitore) del transistore NPN;

Riguardo ai poli che abbiamo nominato, vale che in genere:

- L'*emettitore* (E) si occupa di *emettere* i portatori di carica;
- Il *collettore* (C) si occupa di *raccogliere* i portatori di carica;
- La *base* (B) si occupa di *modulare* il passaggio dei portatori di carica fra le zone emettitore e collettore.

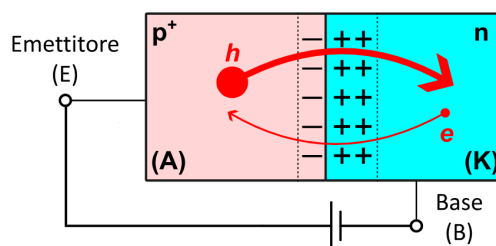
Il risultato è che in un transistore PNP una corrente passa fra collettore ed emettitore quando si applica una piccola corrente *negativa* alla base, mentre in un transistore NPN una corrente passa fra emettitore e collettore quando si applica una piccola corrente *positiva* alla base.

Trattazione del caso PNP

Continuiamo la trattazione fisica nel dettaglio, prendendo a riferimento il transistore PNP (la descrizione del NPN sarà pressapoco analoga). Consideriamo quindi un modello semplificato che prende le due giunzioni (PN ed NP) separatamente. Inoltre, assumiamo che fra base ed emettitore sia imposto un voltaggio positivo (polarizzazione positiva della regione PN), e che fra base e collettore sia imposto un voltaggio negativo (polarizzazione negativa della regione NP).

• **Giunzione PN**

A sinistra il transistore PNP si presenta come segue:



Abbiamo quindi la giunzione PN che viene messa in polarizzazione *diretta* dal voltaggio applicato fra emettitore e base.

Questa è esattamente la situazione che abbiamo studiato in 5.1, e porta questa parte del transistor a comportarsi come un diodo in conduzione: applicando voltaggio in polarizzazione diretta i conduttori maggioritari della zona P (le lacune) si sposteranno nella regione N, la barriera di potenziale si abbasserà, la zona di svuotamento si restringerà, e si avrà quindi conduzione di corrente da E a B (con crescita esponenziale all'aumentare del voltaggio applicato).

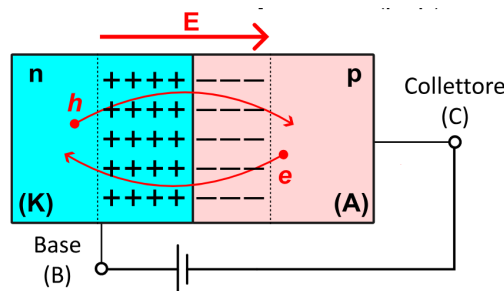
Inoltre, nel PNP si ha che la zona P⁺ è molto drogata, cioè:

$$N_A \gg N_D$$

e quindi il flusso di *h* è molto maggiore del flusso di *e*. L'effetto complessivo di questa situazione sarà quindi che un grande numero di lacune *h* verranno spinte (a volte si dice *iniettate*) nella regione N centrale.

• **Giunzione NP**

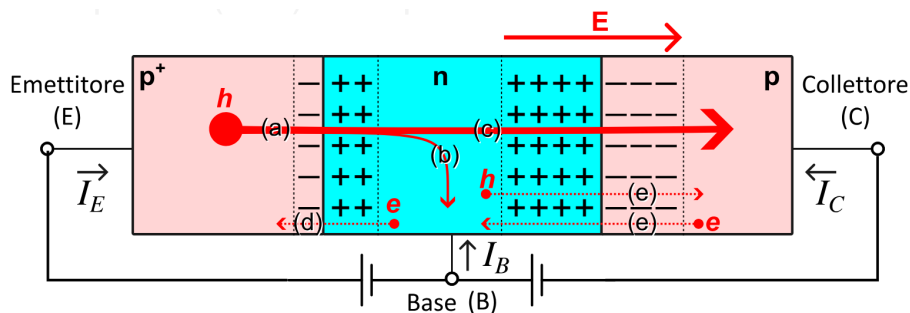
A destra il transistor PNP si presenta come segue:



In questo caso abbiamo sempre una giunzione PN, stavolta in polarizzazione *inversa* fra collettore e base.

Prendiamo per adesso l'effetto ottenuto trascurando la giunzione PN a sinistra (che sta spingendo portatori di carica nella zona N). Visto che siamo in conduzione inversa (sempre da 5.1), la barriera di potenziale si alzerà, la zona di svuotamento si allargherà, e si avrà una corrente di deriva data dai portatori minoritari, ovvero le lacune, limitata dalla generazione termica ed indipendente dal campo *E*. Questa corrente sarà di per sé piccola (saremo praticamente nel caso del diodo in interdizione).

Vediamo quindi cosa accade se si combinano le 2 giunzioni appena viste nel componente che è il transistor PNP. Nel complesso, questo si mostra come segue:



Il regime di operazione visto finora viene detto *Zona Attiva Diretta (ZAD)*, cioè si ha la giunzione BE (cioè la giunzione PN collegata ad emettitore e base) in polarizzazione diretta, e la giunzione BC (cioè la giunzione NP collegata a base e collettore) in polarizzazione inversa. In simboli:

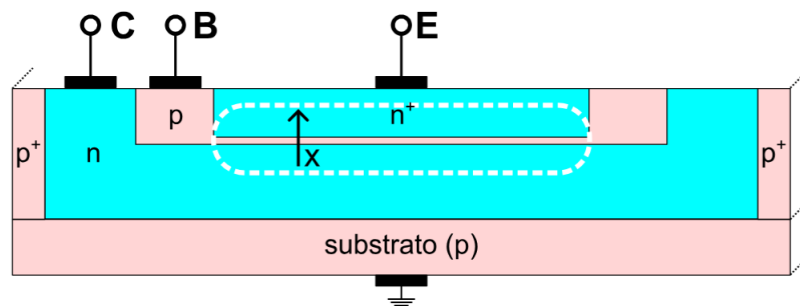
$$V_{EB} > 0, \quad V_{CB} < 0$$

Da quanto detto prima, questo significherà che:

1. Ci sarà un grande flusso di lacune h spinte (iniettate) da E in B. Questo sarà dato dalla corrente di diffusione (cioè dalla giunzione PN in polarizzazione diretta) dei portatori maggioritari della regione P fortemente drogata;
2. Le h si troveranno quindi all'interno della regione N (cioè la base), dove rappresenteranno le portatrici minoritarie. Il loro destino qui sarà duplice:
3. Una piccola porzione degli h nella base si ricombineranno nella base, dove si formerà una zona neutra.

In altre parole, un'alta concentrazione di e (data dal drogaggio N e dal fatto che il campo E applicato riporta gli n all'interno della zona N) nella base porterà all'annularsi di alcune delle lacune h entranti, e quindi renderà la zona complessivamente neutra. Come approfondiremo, questo meccanismo deve essere supplementato da altri elettroni che entrano nella base, attraverso una corrente uscente sulla stessa;

4. La parte rimanente degli h arriverà alla zona di svuotamento BC. Qui sarà presente un forte campo elettrico: in particolare, gli atomi accettori rimasti nella zona P di destra formeranno una forte carica negativa oltre la zona di svuotamento della giunzione NP.



Questo, unito al fatto che la zona N è molto piccola (come si nota dalla figura, che mostra la realizzazione effettiva del PNP su un substrato di silicio drogato P), porterà la maggior parte delle lacune h a superare completamente la zona N ed arrivare direttamente al collettore. Queste sono le cariche che riforniscono la giunzione BC in inversa.

Notiamo un altro dettaglio riguardo al fatto che la zona N (nel PNP, altrimenti nell'NPN sarebbe la zona P) è molto ristretta: questo, oltre al fatto che permette ai portatori di attraversare la base per raggiungere il collettore, impedisce al transistor di comportarsi semplicemente come 2 diodi accoppiati fra di loro (che è quello che ci aspetteremmo da due giunzioni PN unite fra di loro).

Abbiamo quindi chiarito come mai i nomi e i funzionamenti delle regioni del transistor:

- L'*emettitore* (**E**) si occupa appunto di *emettere* i portatori di carica: fa questo (nel PNP) agendo da giunzione PN polarizzata positivamente, dove naturalmente si forma un flusso di portatori maggioritari verso la regione N.
- Il *collettore* (**C**) si occupa invece di *raccogliere* i portatori di carica. Questo atto di "raccoglimento" non è altro che l'attrazione dei portatori (nel PNP le lacune) finiti nella regione neutra attraverso la carica disposta sulla zona di svuotamento;

Infine, sulla *base* (**B**) si potrà *modulare* il passaggio dei portatori di carica fra le zone emettitore e collettore. Questo verrà fatto agendo sulla neutralità della base stessa: la corrente I_B rifornirà i portatori (nel PNP gli elettroni) necessari per mantenere la neutralità di carica nella zona neutra. Se non li rifornisse (cioè nel caso $I_B = 0$), i portatori si accumulerebbero in base generando un potenziale che ostacolerebbe l'ulteriore iniezione di portatori ($I_E = 0$): la zona neutra della base sparirebbe. Sia quindi chiaro che la corrente I_B è di fondamentale importanza per il funzionamento del transistor stesso.

Abbiamo quindi ottenuto una particolare situazione dove la corrente piccola sulla base controlla una corrente elevata fra emettitore e collettore. Quest'ultima è una versione amplificata della corrente di base. In altre parole, abbiamo realizzato il GCCC che ci eravamo prefissati di ottenere.

Trattazione del caso NPN

Senza raggiungere lo stesso livello di dettaglio (la trattazione completa fatta sul PNP è più che sufficiente) vediamo in breve il processo, analogo, che si svolge nel transistor NPN. Innanzitutto, abbiamo che in un transistor NPN l'emettitore si trova a valle della corrente (si producono elettroni, che si spostano nella direzione opposta alla corrente). Questo significherà che si dovrà assumere la prima regione NP in polarizzazione negativa, e la successiva regione PN in polarizzazione positiva.

- **Giunzione NP**

Vediamo quindi l'NPN al contrario (dal $-$ della corrente al $+$, partendo dalla giunzione NP. Questa giunzione (emettitore-base) è in polarizzazione *diretta*.

Questa è la stessa situazione studiata nei diodi in conduzione: applicando un voltaggio diretto, i portatori maggioritari della regione N (gli elettroni) si spostano verso la regione P, la barriera di potenziale si abbassa, la zona di svuotamento si restringe, e si ha conduzione di corrente da E a B.

Inoltre, nell'NPN si ha che la regione N (emettitore) è fortemente drogata e quindi il flusso di elettroni è molto maggiore del flusso di lacune. L'effetto complessivo è che un grande numero di elettroni vengono *iniettati* nella regione P centrale (la base).

- **Giunzione PN**

Questa giunzione (base-collettore) è in polarizzazione *inversa*.

Se consideriamo solo questa giunzione, la barriera di potenziale aumenta, la zona di svuotamento si allarga, e si ha una corrente di deriva dovuta ai portatori minoritari (gli elettroni nella regione P), limitata dalla generazione termica.

Questa corrente è molto piccola (caso di diodo in interdizione).

Combinando le due giunzioni si ottiene il funzionamento del transistor NPN in *Zona Attiva Diretta* (ZAD) (come prima per il PNP).

$$V_{BE} > 0, \quad V_{BC} < 0$$

1. Ci sarà un grande flusso di elettroni n spinti (iniettate) da E in B. Questo sarà dato dalla corrente di diffusione (cioè dalla giunzione NP in polarizzazione diretta) dei portatori maggioritari della regione N fortemente drogata;
2. Gli n si troveranno quindi all'interno della regione P (cioè la base), dove rappresenteranno i portatori minoritarie. Come per il PNP, il loro destino qui sarà duplice:
3. Una piccola porzione degli e nella base si ricombineranno nella base, dove si formerà una zona neutra.
4. La parte rimanente degli e arriverà alla zona di svuotamento BC. Qui sarà presente un forte campo elettrico: in particolare, gli atomi donatori rimasti nella zona N di sinistra formeranno una forte carica negativa oltre la zona di svuotamento della giunzione PN. Questo, unito al fatto che la zona P è molto piccola, porterà la maggior parte degli elettroni e a superare completamente la zona P ed arrivare direttamente al collettore.

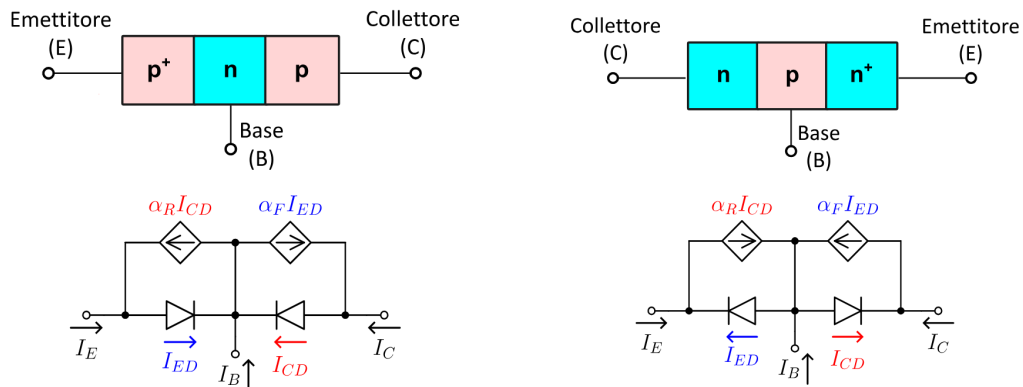
Le regioni (e i loro funzionamenti) saranno quindi analoghi a quelli del transistor PNP (se non invertiti spazialmente).

In questo caso, la corrente di base I_B fornirà le lacune necessarie per compensare la ricombinazione degli elettroni nella base. Se $I_B = 0$ si accumulerebbero cariche nella base, generando un campo elettrico che ostacolerebbe ulteriori iniezioni. Tuttavia, il transistor può comunque condurre in altre condizioni di polarizzazione: la base corrente non è un "abilitatore", ma una conseguenza della ricombinazione.

Si ottiene quindi una configurazione in cui una piccola corrente di base controlla una corrente molto maggiore tra emettitore e collettore, realizzando il principio di amplificazione del transistor. Di nuovo, abbiamo realizzato il GCCC che ci eravamo prefissati di ottenere.

10.1.2 Modello di Ebers-Moll

Il modello di *Ebers-Moll* per il transistor bipolare permette di descrivere questo componente ai grandi segnali. Il modello si deriva direttamente dalla rappresentazione fisica del transistor, riportata ad una serie di componenti governati da legge elettriche (e quindi per cui sappiamo scrivere e risolvere equazioni).



Modellizziamo quindi il transistore come una coppia di diodi che pilotano generatori di corrente (per realizzare appunto il GCCC). Questi sono:

- Rivolti da E a B e da E a C nel caso del transistore PNP;
- Rivolti da B ad E e da B a C nel caso del transistore NPN.

I generatori di corrente chiudono le maglie seguendo le direzioni dei diodi. Le costanti di proporzionalità sono 2, cioè α_F e α_R :

- α_F è la frazione di corrente *diretta*, con:

$$0.98 < \alpha_F < 0.998$$

- α_R è la frazione di corrente *inversa*, con:

$$0.4 < \alpha_R < 0.8$$

Fra queste costanti, a causa del diverso drogaggio delle regioni P (nel PNP) o N (nel NPN), saranno diverse fra di loro, e in particolare varrà:

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

Dove I_{ES} e I_{CS} sono le correnti di saturazione dei diodi sull'emettitore e sul collettore. Solitamente avremo:

- In un PNP che α_F (da *Forward*) $\gg \alpha_R$ (da *Reverse*), cioè facendo scorrere corrente fra emettitore e base (sul diodo I_{ED}) si impone una grande corrente fra emettitore e collettore (generatore di corrente $\alpha_F I_{ED}$);
- In un NPN si ha la stessa cosa, cioè facendo scorrere corrente fra base ed emettitore (sul diodo I_{ED}) si impone una grande corrente fra collettore ed emettitore (generatore di corrente $\alpha_F I_{ED}$);

Trattazione del caso PNP

Come per la trattazione fisica, prendiamo ad esempio specifico il transistore PNP. Possiamo discutere il modello di Ebers-Moll introducendo l'equazione di Shockley per i diodi:

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

e quindi scrivendo le equazioni nodali:

$$\begin{cases} I_E = I_{ED} - \alpha_R I_{CD} \\ I_C = I_{CD} - \alpha_F I_{ED} \\ I_B = -I_E - I_C \end{cases}$$

che sostituendo ci dà:

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

Sempre nel PNP, in *zona attiva diretta*, abbiamo:

$$V_{EB} \gg V_T, \quad V_{CB} \ll 0$$

cioè regione PN in polarizzazione diretta e regione NP in polarizzazione inversa, per cui possiamo approssimare gli esponenziali:

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

Semplificando ancora possiamo rimuovere dei termini (in quanto gli esponenziali vincono sui termini lineari):

$$\begin{cases} I_E = I_{ES} e^{\frac{V_{EB}}{V_T}} \\ I_C = -\alpha_F I_{ES} e^{\frac{V_{EB}}{V_T}} \end{cases}$$

A questo punto, mettendo a comune la I_E , si ottiene:

$$\begin{cases} I_C = -\alpha_F I_E \\ I_B = -(I_E + I_C) = -(1 - \alpha_F) I_E \end{cases}, \quad I_E = I_{ES} e^{\frac{V_{EB}}{V_T}}$$

Ci siamo quindi riportati al modello a *generatore di corrente pilotato in corrente*. Possiamo ricavare un'espressione per il *fattore di proporzionalità* fra la corrente applicata alla base e quella misurata al collettore, cioè se vogliamo l' A_i della prima trattazione fatta in 1.1:

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

Abbiamo quindi che, almeno nella ZAD, il BJT si comporta effettivamente da GCCC, con costante di proporzionalità β_f per la corrente al collettore rispetto alla corrente imposta alla base (quella che prendiamo come modulante).

Trattazione del caso NPN

Trattiamo brevemente anche il transistor PNP. Introduciamo quindi l'equazione di Shockley per i diodi:

$$\begin{cases} I_{ED} = I_{ES} \left(e^{\frac{V_{BE}}{V_T}} - 1 \right) \\ I_{CD} = I_{CS} \left(e^{\frac{V_{BC}}{V_T}} - 1 \right) \end{cases}$$

e scrivendo le equazioni nodali:

$$\begin{cases} I_E = -I_{ED} + \alpha_R I_{CD} \\ I_C = -I_{CD} + \alpha_F I_{ED} \\ I_B = -I_E - I_C \end{cases}$$

che sostituendo ci dà:

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

Sempre nel NPN, in *zona attiva diretta*, abbiamo:

$$V_{BE} \gg V_T, \quad V_{BC} \ll 0$$

cioè regione NP in polarizzazione diretta e regione PN in polarizzazione inversa, per cui possiamo approssimare gli esponenziali:

$$\begin{cases} I_E = -I_{ES}e^{\frac{V_{BE}}{V_T}} + \alpha_R I_{CS}(-1) \\ I_C = -I_{CS}(-1) + \alpha_F I_{ES}e^{\frac{V_{BE}}{V_T}} \end{cases} \implies \begin{cases} I_E = -I_{ES}e^{\frac{V_{BE}}{V_T}} - \alpha_R I_{CS} \\ I_C = +I_{CS} + \alpha_F I_{ES}e^{\frac{V_{BE}}{V_T}} \end{cases}$$

Semplificando ancora possiamo rimuovere dei termini (in quanto gli esponenziali vincono sui termini lineari):

$$\begin{cases} I_E = -I_{ES}e^{\frac{V_{BE}}{V_T}} \\ I_C = +\alpha_F I_{ES}e^{\frac{V_{BE}}{V_T}} \end{cases}$$

A questo punto, mettendo a comune la I_E , si ottiene:

$$\begin{cases} I_C = -\alpha_F I_E \\ I_B = -(I_E + I_C) = -(1 - \alpha_F)I_E \end{cases}, \quad I_E = -I_{ES}e^{\frac{V_{BE}}{V_T}}$$

Che dà lo stesso identico modello a *generatore di corrente pilotato in corrente* del PNP, con *fattore di proporzionalità*:

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

Potrebbe sembrarci strano prendere I_C anziché I_E per il calcolo del fattore di moltiplicazione della corrente, in quanto sappiamo che in un transistor NPN la corrente va da collettore ad emettitore, e non viceversa (l'emettitore emette elettroni e opposti al flusso di corrente). Come vedremo in seguito, però, in fase di utilizzo i transistori vengono solitamente messi ad *emettitore comune*, cioè emettitore E e base B vengono messi a comune e si prende la corrente di carico sul collettore C. Il funzionamento di transistor PNP e NPN sarà quindi identico, se non per la direzione della corrente che viene invertita (il ruolo delle regioni E, B e C resta tale).

10.1.3 Regioni di funzionamento del BJT

Notiamo innanzitutto qual'è il valore indicativo di questa costante di proporzionalità β_F (in diretta) ricavata nella scorsa sezione.

Avevamo detto che le costanti di proporzionalità del modello di Ebers-Moll erano le seguenti:

- α_F è la frazione di corrente *diretta*, con:

$$0.98 < \alpha_F < 0.998$$

- α_R è la frazione di corrente *inversa*, con:

$$0.4 < \alpha_R < 0.8$$

Si ricava allora immediatamente:

$$\beta_F = \frac{\alpha_F}{1 - \alpha_F} \implies 50 < \beta_F < 500$$

Iniziamo quindi a vedere le *regioni* di operazione dei BJT. Finora abbiamo visto solo la regione *ZAD* (*Zona Attiva Diretta*), che corrispondeva a polarizzazione EB diretta e polarizzazione CB inversa nel PNP, e polarizzazione BE diretta e polarizzazione BC inversa nel NPN. Abbiamo, in verità, che un transistor può trovarsi in diverse regioni a seconda delle polarizzazioni delle due giunzioni PN. Vediamole nel dettaglio:

Zona	Impiego	Polarizzazione PNP	Polarizzazione NPN
<i>Zona Attiva Diretta</i> (ZAD)	Amplificatore efficiente	EB diretta, CB inversa	BE diretta, BC inversa
<i>Zona Attiva Inversa</i> (ZAI)	Amplificatore inefficiente	EB inversa, CB diretta	BE inversa, BC diretta
<i>Interdizione</i> (OFF)	Interruttore aperto	EB inversa, CB inversa	BE inversa, BC inversa
<i>Saturazione</i> (ON)	Interruttore chiuso	EB diretta, CB diretta	BE diretta, BC diretta

Commentiamo brevemente queste zone. Abbiamo che:

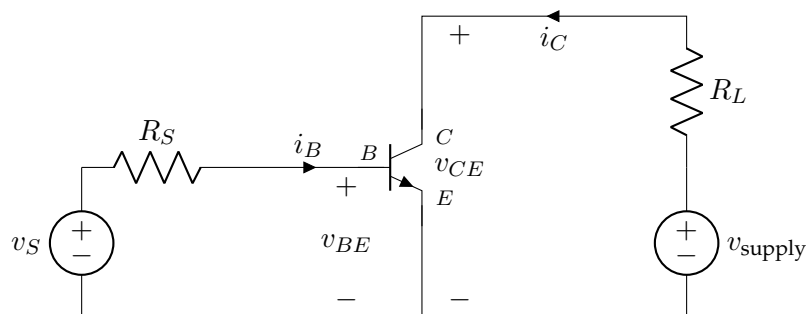
- Nell'elettronica **analogica**, dove il transistor deve fare da componente *amplificatore* di corrente, resteremo nella ZAD (per ottenere un amplificatore efficiente, dato che $\alpha_F > \alpha_R$ e l'alternativa zona ZAI sfrutterebbe la costante di proporzionalità minore; notiamo inoltre che la differenza fra α_F ed α_R è voluta e data dei diversi livelli di drogaggio delle zone esterne del transistor);
- Nell'elettronica **digitale**, dove il transistor fa da *interruttore* controllato in corrente, vorremo oscillare fra le zone di interdizione e saturazione.

10.1.4 Caratteristiche ad emettitore comune

Vediamo quindi come si studia un circuito che mostra il transistor (PNP o NPN) nel suo caso di applicazione tipico, cioè come amplificatore di potenza per una certa sorgente di segnale.

Innanzitutto, vediamo il circuito che vogliamo studiare sia con transistor NPN che PNP.

- Il circuito con transistor NPN si presenta come segue:



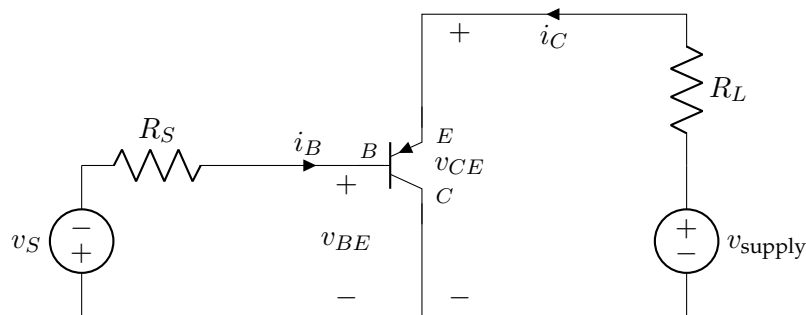
Il funzionamento del circuito è il seguente:

1. Il generatore v_S a sinistra produce un (piccolo) segnale, che passa attraverso la resistenza (che associamo al circuito generatore del segnale) R_S e quindi entra positivo nella base dell'NPN;
2. Il generatore v_{supply} a destra produce una (grande) differenza di potenziale. Il transistor NPN, quindi, traduce tale differenza di potenziale in una corrente I_C che viene fatta passare sulla resistenza di carico R_L . Notiamo che la differenza di potenziale v_{CE} sul transistor dipende anche dalla differenza di potenziale imposta alla base, cioè vale:

$$v_{CE} = v_{BE} + V_{CB} = v_{BE} - V_{BC}$$

Vediamo quindi il componente circuitale del transistor NPN, e la scelta canonica delle regioni C (collettore) ed E (emettitore). In questo caso abbiamo che la corrente i_C entra nel collettore C ed esce dall'emettitore E, messo a comune con la maglia della base B. Era questa la convenzione di corrente che avevamo già mostrato per l'NPN: l'emettitore emette elettroni e , che viaggiano verso il collettore, per cui la direzione della corrente convenzionale è opposta.

- Il circuito con transistor PNP si presenta invece come segue:



Il funzionamento del circuito è analogo al caso con il PNP, con la sola differenza sulla direzione del generatore v_S .

Vediamo quindi il componente circuitale del transistor PNP, e la scelta canonica delle regioni C (collettore) ed E (emettitore). In questo caso abbiamo che la corrente i_C entra nell'emettitore E ed esce dal collettore C, messo a comune con la maglia della base B. Era questa la convenzione di corrente che avevamo già mostrato per il PNP l'emettitore emette lacune h , che viaggiano verso il collettore, per cui la direzione della corrente convenzionale è diretta.

Da qui in poi continueremo con l'esempio dell'NPN, in quanto il segno della corrente i_B è più immediato (è positivo, cioè corrente entra nel transistor).

Il tipo di modellizzazione che vorremo ottenere è quindi quello della *caratteristica di ingresso*, cioè la relazione della corrente i_B ai voltaggi imposti V_{BE} e V_{CE} , e della *caratteristica di uscita*, cioè la relazione della corrente i_C al voltaggio imposto V_{CE} e alla corrente i_B che riusciamo ad ottenere alla base.

Notiamo che questo è in accordo con quanto detto in generale riguardo al transistor come componente GCCC (generatore di corrente controllato in corrente): agendo sulla corrente i_B si varia la neutralità della zona centrale del transistor (per l'NPN, la zona

P), e quindi si modula la corrente i_C che passa fra emettitore e collettore (nell'NPN in senso negativo).

Notiamo che questo è in accordo anche con il cosiddetto *principio di accoppiamento*:

10.1: Principio di accoppiamento

In un dispositivo a tre terminali (quadripolo), lo stato elettrico di una porta è intrinsecamente dipendente da una variabile della porta opposta, definendo così l'effetto di controllo o amplificazione.

Caratteristica di ingresso

Vediamo la caratteristica di ingresso. Questa sarà, come abbiamo detto, la relazione della corrente i_B ai voltaggi imposti V_{BE} e V_{CE} . In simboli:

$$I_B = f(V_{BE}, V_{CE})$$

Cioè la corrente sarà funzione dei voltaggi imposti a base e emettitore-collettore. Possiamo ricavare la relazione esplicita riprendendo la discussione fatta per l'NPN in 10.1.2, cioè sfruttando il modello di *Ebers-Moll*.

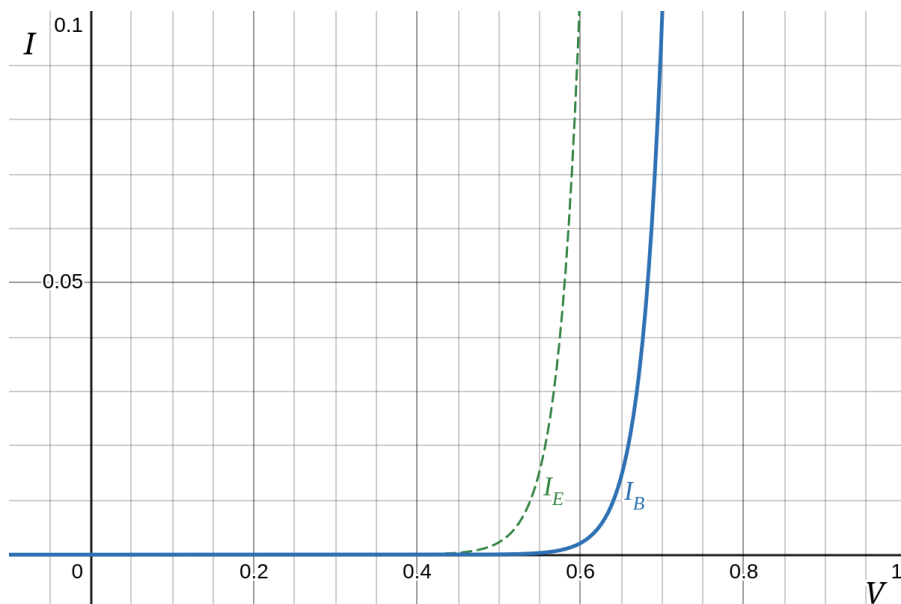
In tal caso abbiamo:

$$I_B = -(1 - \alpha_F)I_E, \quad I_E = -I_{ES}e^{\frac{V_{BE}}{V_T}}$$

visto che ci interessiamo alle risposte, e sappiamo che la corrente è uscente dal transistor in I_E (opposta alla direzione in cui la prendiamo, rivolta verso il transistor), passiamo ai valori assoluti:

$$I_B = (1 - \alpha_F)|I_E| = (1 - \alpha_F)I_{ES}e^{\frac{V_{BE}}{V_T}}$$

Avremo tra l'altro che questa è anche la curva di risposta del PNP.



Vediamo quindi che questa è analoga alla caratteristica del diodo, cioè presenta una crescita esponenziale legata al voltaggio applicato. Per semplicità, poi, consideriamo di

trovarci in ZAD per V_{BE} imposta $> 0.7\text{ V}$, cioè circa V_γ . In ZAD, quindi semplifichiamo dicendo:

$$I_B = f(V_{BE})$$

cioè trascurando la dipendenza da V_{CE} . Vedremo più avanti che in realtà questa è una semplificazione, valida per la ZAD, ma non per altre zone (soprattutto vedremo come si passa dalla regione ZAD alla regione di saturazione).

Caratteristica di uscita

Veniamo quindi alla caratteristica di uscita. Questa sarà, come abbiamo detto, la relazione della corrente i_C al voltaggio imposto V_{CE} e alla corrente i_B che riusciamo ad ottenere alla base. In simboli:

$$I_C = f(I_B, V_{CE})$$

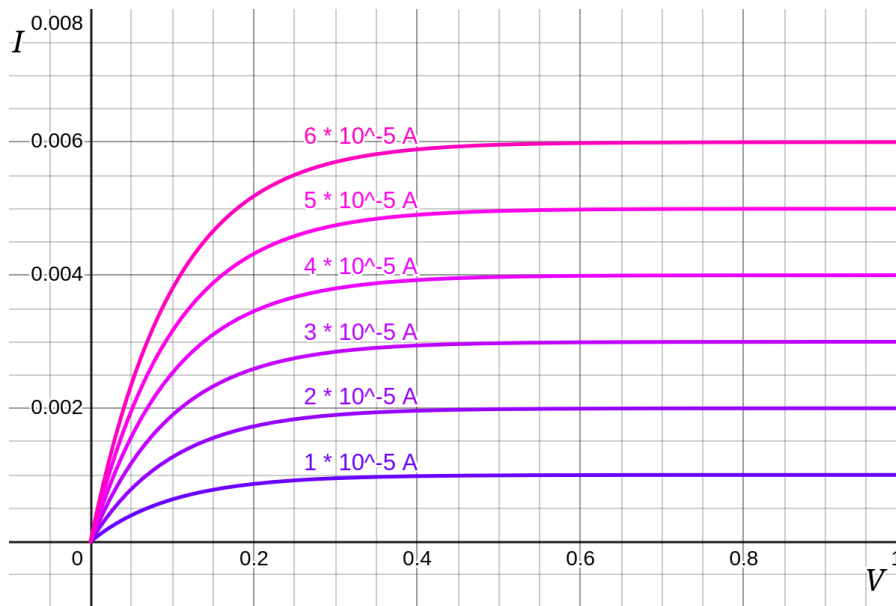
Cioè la corrente sarà funzione della corrente i_B che otteniamo, e del voltaggio emettitore-collettore. Possiamo ricavare la relazione esplicita riprendendo sempre il modello di Ebers-Moll.

In tal caso abbiamo:

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

totalmente lineare, che era ciò che volevamo ottenere per avere un transistor ideale più vicino possibile ad un GCCC. Anche qui, visto che ci interessiamo alle risposte, e sappiamo che la corrente è entrante nel transistor in I_B (concorde alla direzione in cui la prendiamo, rivolta verso il transistor), passiamo ai valori assoluti:

$$I_C = \beta_F |I_B| = \beta_F I_B$$



Vediamo come questa curva è accurata solo per valori sufficientemente grandi di V_{CE} . Per valori sufficientemente bassi notiamo che I_C collassa a zero, secondo un fenomeno detto *piegamento delle curve*. Chiamiamo il valore di V_{CE} dove il piegamento inizia ad essere consistente $V_{CE, sat}$, cioè il punto di saturazione dove il transistor inizia a passare dalla regione ZAD alla regione di saturazione.

Riassumendo, vediamo che in ZAD la dipendenza da V_{CE} è scomparsa, ma come abbiamo anticipato questa è una semplificazione. Al di sotto della $V_{CE, sat}$, infatti, entriamo in zona di saturazione, dove la I_C crolla a zero.

Zona di saturazione

Vediamo nel dettaglio come si entra nella zona di saturazione. Abbiamo detto che i nostri valori in uscita dipendono dal voltaggio V_{CE} , e che nella regione ZAD questa dipendenza poteva essere trascurata prendendo la dipendenza lineare da I_B (abbiamo detto che in verità I_B dipende leggermente da V_{CE} , soprattutto fuori dalla ZAD, per cui la dipendenza non è totalmente lineare). In altre parole, in regime ZAD (sopra gli 0.7 V circa) il collettore "raccolge" tutti i portatori che arrivano dalla base, e il BJT si comporta come un generatore di corrente ideale.

Specifichiamo quindi V_{CE} nell'esattezza:

$$V_{CE} = V_{BE} - V_{BC}$$

per cui la stessa tensione che applichiamo al collettore-emettitore del transistor dipende dalla tensione in BC e quindi dalla caduta data dalla corrente I_B . Ciò che succede a livello generale è riducendo la V_{CE} la giunzione BC (la giunzione PN del NPN) passa dall'essere in polarizzazione *inversa* alla polarizzazione *diretta*. Possiamo darci una spiegazione fisica di questo fenomeno dicendo che la giunzione BC diretta inietta portatori opposti a quelli dell'emettitore. Come risultato, la corrente risultante I_C diminuisce drasticamente.

Vediamo i range di tensione nello specifico.

- Vediamo che il fenomeno di piegamento delle curve inizia verso la tensione di conduzione del diodo, cioè:

$$V_\gamma \approx 0.7 \text{ V}$$

- Il punto di saturazione vero e proprio è invece raggiunto nella $V_{CE, \text{sat}}$, che è leggermente minore:

$$V_{CE, \text{sat}} \approx 0.1 - 0.3 \text{ V}$$

Potremmo chiederci come mai tale valore non è pari a zero. In saturazione, si ha che:

$$V_{CE, \text{sat}} = V_{BE} - V_{BC}$$

Essendo la giunzione BC meno drogata, la sua caduta (di circa 0.4 V) è inferiore alla BE (0.7 V, situazione che lascia un residuo positivo).

10.1.5 Effetto Early

Il modello di Ebers-Moll, e quindi tutti gli ulteriori che abbiamo fatto su transistor PNP ed NPN ad emettitore comune, trascurano un effetto che notiamo sulle caratteristiche sia di ingresso che di uscita dipendente dal voltaggio V_{CE} (come anticipato più volte nelle scorse sezioni). Questa è una generale dipendenza direttamente proporzionale delle correnti a V_{CE} , che chiamiamo effetto di *Early*.

Caratteristica di ingresso

Riprendiamo la caratteristica di ingresso. Questa era la relazione della corrente i_B ai voltaggi imposti V_{BE} e V_{CE} . In simboli:

$$I_B = f(V_{BE}, V_{CE})$$

Avevamo detto che la dipendenza da V_{CE} era trascurabile in ZAD, e quindi la risposta era paragonabile a quella del diodo per $V_{CE} > V_\gamma \approx 0.7 \text{ V}$.

Vediamo in verità che si nota una riduzione di I_B all'aumentare di V_{CE} (cosa che porta anche alla non linearità delle I_B ottenute a diverse tensioni V_{BE}). Comunque, ricordiamo che questo è un effetto poco evidente, e secondario rispetto al prossimo che vedremo.

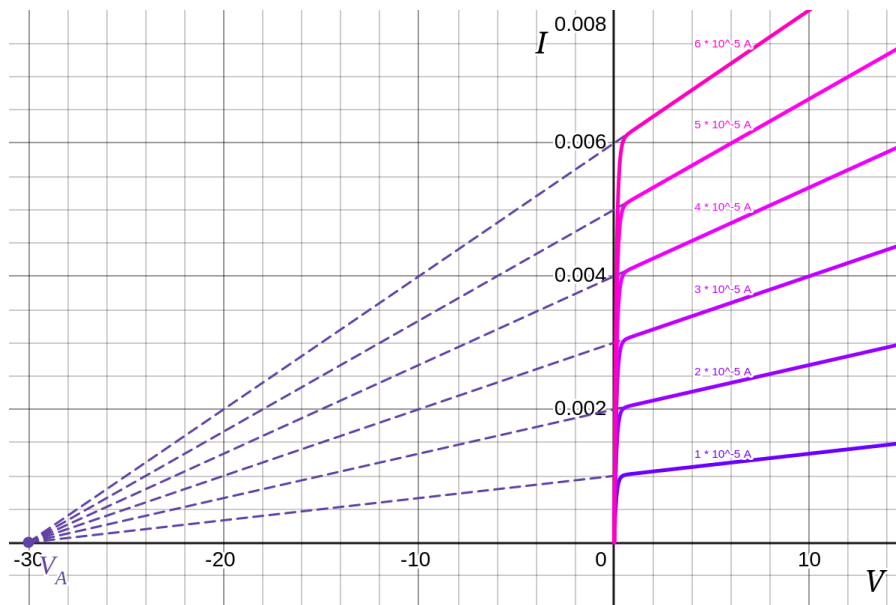
Caratteristica di uscita

L'effetto di Early più consistente si ha sulla caratteristica di uscita. Questa era la relazione della corrente i_C al voltaggio imposto V_{CE} e alla corrente i_B che riusciamo ad ottenere alla base. In simboli:

$$I_C = f(I_B, V_{CE})$$

Di questa avevamo detto che la dipendenza da V_{CE} era visibile solamente per $V_{CE} < V_{\gamma}$, e in particolare nel punto di saturazione $V_{CE, sat}$ dove la giunzione BC passava in polarizzazione diretta e si entrava quindi nel regime di saturazione del transistor.

Vediamo che esiste un'ulteriore dipendenza in V_{CE} , non mostrata dal modello di Ebers-Moll. Prolungando le caratteristiche di uscita in I_C si ottiene un punto di convergenza sull'asse della V_{CE} :



Questo punto viene detto *tensione di Early* V_A (o $-V_A$, a seconda di come scegliamo la convenzione di segno). Vediamo come le correnti I_C , che seguono la convergenza a alla tensione di Early, mostrano una dipendenza lineare alla tensione V_{CE} .

Possiamo darci una spiegazione fisica dell'effetto di Early considerando la larghezza della zona di svuotamento all'interno di una giunzione PN, vista per la prima volta in 5.1:

$$w = \sqrt{\frac{2\epsilon_0}{e} \left(\frac{1}{N_A} + \frac{1}{N_d} \right) (V_0 - V_{BC})}$$

dove notiamo che abbiamo scelto come la nostra tensione di bias la V_{BC} imposta alla base del transistor.

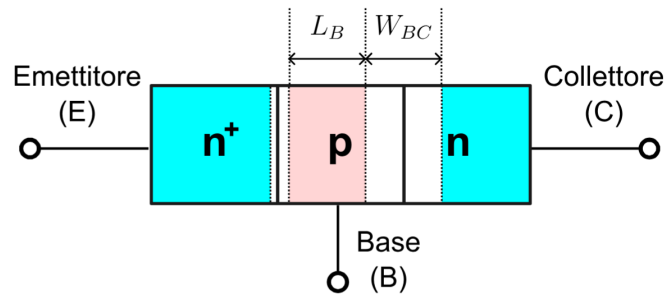
Ricordiamo quindi che:

$$V_{BC} = -V_{CB}, \quad V_{CB} = V_{CE} - V_{BE}$$

da quanto detto in 10.1.4 riguardo ai circuiti ad emettitore comune. Quindi, ricordiamo che in ZAD si ha che V_{BE} è circa costante, in particolare $\sim V_{\gamma}$ come dal diodo.

L'effetto sarà che all'aumentare di V_{CE} , con V_{BE} fissata, non può che aumentare V_{CB} . Se V_{CB} aumenta, aumenta la larghezza delle zone di svuotamento ai margini della regione centrale del transistor w . Diminuirà quindi anche la larghezza della regione neutra del transistor (che viene "mangiata" dalle zone di svuotamento che vanno ad allargarsi). In questa situazione i portatori maggioritari hanno minore probabilità di ricombinarsi, e quindi una parte maggiore di loro raggiungerà il collettore, contribuendo alla corrente I_C .

Possiamo vedere un grafico che mostra questo esempio su un NPN:



Riassumendo, abbiamo che al ridursi della lunghezza della porzione di base neutra (L_B), diminuisce la probabilità di ricombinazione in base (I_B , effetto Early sulla caratteristica in entrata), e quindi più lacune (almeno nel NPN) emesse raggiungono il collettore (I_C , effetto Early sulla caratteristica in uscita).

Dal punto di vista pratico, l'effetto di Early sulla caratteristica di uscita significa che esiste una resistenza in uscita finita in un transistor BJT. Il risultato è che la regione attiva (la ZAD) non è perfettamente costante in V_{CE} , ma presenta una piccola resistenza:

$$r = \frac{V_A}{I_C}$$

Vorremo quindi tenere conto di tale resistenza nella progettazione di circuiti che sfruttano i transistor, ad esempio gli amplificatori. Questo perché la resistenza può introdurre *distorsione* nel segnale in uscita (dato dal fatto che la caratteristica in uscita non è costantemente proporzionale alla I_B , e quindi non lo è neanche il guadagno). Per ovviare a questo problema abbiamo 2 opzioni:

- Scegliere transistor con grande tensione di *Early* V_A : questo ha un effetto negativo sul guadagno (aumenta la resistenza) ma rende anche la risposta in uscita più lineare con l'ingresso I_B , riducendo la distorsione;
- Mantenere la I_C bassa, tenendo a sua volta bassa la I_B di modulazione, dato che ricordiamo:

$$I_C = \beta_F I_B$$

Questo, come per la soluzione precedente, ha un effetto negativo sul guadagno (aumentiamo la resistenza e inoltre inviamo una corrente minore alla base del transistor). Sempre come prima, però, riduciamo la distorsione data dall'effetto di Early e quindi otteniamo dal transistor un'amplificazione di qualità migliore.

Abbiamo quindi visto come entrambe le soluzioni proposte presentano un pro (diminuire la distorsione da effetto Early) ed un contro (ridurre il guadagno in corrente A_i dell'amplificatore). Questo è un classico esempio di *tradeoff* ingegneristico nella progettazione di un circuito elettronico: avere una buona qualità significa avere minore guadagno, e viceversa.

11 Lezione del 27-03-26

Nella scorsa lezione abbiamo introdotto i transistori **BJT** (*Bipolar Junction Transistor*). Iniziamo adesso a vedere come risolvere circuiti che contengono tali componenti.

11.0.1 Risoluzione di circuiti con BJT

Avremo a disposizione diversi metodi di risoluzione per determinare la risposta del transistore ai *grandi segnali* (come avevamo per il diodo a 6.1.3). Innanzitutto, ai grandi segnali significa determinare il punto di riposo, cioè le grandezze:

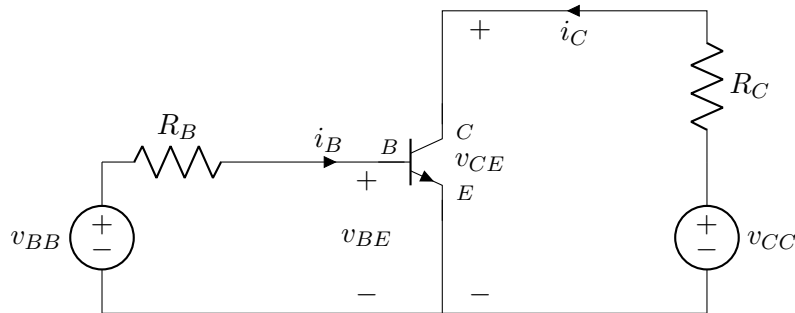
$$V_{BEQ}, I_{BQ}, V_{CEQ}, I_{CQ}$$

cioè tensioni e correnti di base e di collettore in condizioni di equilibrio.

I metodi che avremo saranno quindi:

- Metodi *analitici*: basati sul modello di Ebers-Moll o eventuali modelli semplificati derivanti da questo;
- Metodi *grafici*: consistono nella tracciatura della retta di carico (vista sempre in 6.1.3) sulle caratteristiche di ingresso e di uscita del BJT (viste nella scorsa lezione, legano le correnti di base e di collettore alle relative tensioni).

Vediamo ad esempio il seguente circuito:



già introdotto alla scorsa lezione per la discussione delle caratteristiche di ingresso e di uscita ad emettitore comune.

Possiamo quindi impostare le equazioni alle maglie:

$$\begin{cases} V_{BB} = R_B I_B + V_{BE} \\ V_{CC} = R_C I_C + V_{CE} \end{cases}$$

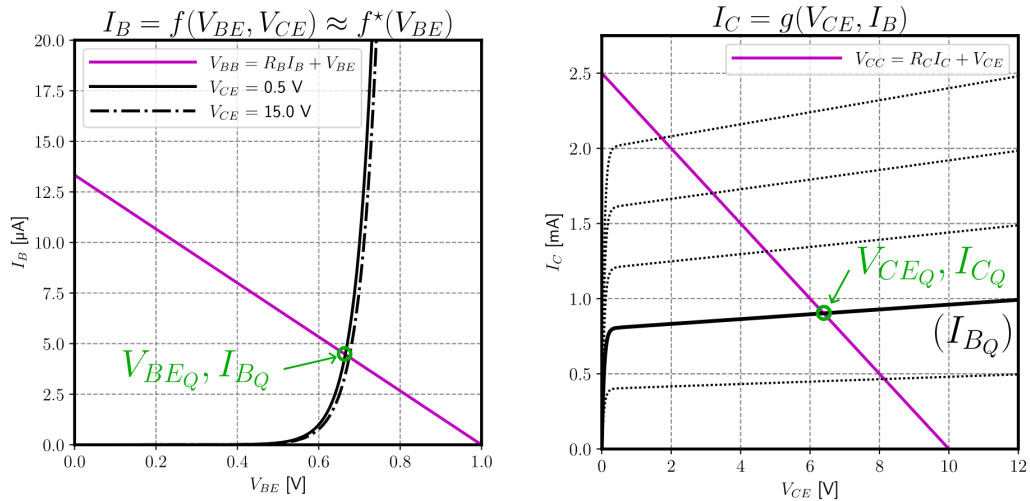
e le caratteristiche di ingresso e di uscita del BJT:

$$\begin{cases} I_B = f(V_{BE}, V_{CE}) = I_{BES} e^{\frac{V_{BE}}{V_T}} \\ I_C = f(I_B, V_{CE}) = \beta_F I_B \end{cases}$$

- **Metodo grafico**

Proviamo a risolvere il circuito col metodo grafico. Per fare ciò prenderemo la retta di carico ricavata dall'equazione della maglia su V_{BB} , e la tratteremo sul grafico sovrapposta alla curva di risposta in ingresso del BJT.

Quindi, ricavato un valore valido di I_B , possiamo selezionare tale valore sul grafico della I_C in funzione di V_{CE} al variare di I_B . Sempre per intersezione di grafici, troveremo la corrente I_C di collettore.



Questo metodo ci dà un errore dato dal fatto che, presa la caratteristica di ingresso approssimata:

$$I_B = f(V_{BE}, V_{CE}) = (1 - \alpha_F) I_{ES} e^{V_{BE}/T}$$

trascuriamo la V_{CE} per il calcolo di I_B (effetto di Early trascurabile). Questo ci è spesso possibile in quanto la differenza effettiva fra i valori che otteniamo è minima (dal grafico si vede la minima differenza variando V_{CE} da 0.5 V a 15 V);

• **Metodo analitico**

Per il metodo analitico, assumiamo di trovarci in ZAD e prendiamo il modello di Ebers-Moll semplificato:

$$I_{BQ} = (1 - \alpha_F) I_{ES} e^{\frac{V_{BEQ}}{V_T}}$$

in tal caso l'effetto di Early all'ingresso è automaticamente trascurato (il modello di Ebers-Moll non lo considera a priori). Possiamo quindi sostituire questa nell'equazione della retta di carico, per ottenere:

$$V_{BB} = R_B (1 - \alpha_F) I_{ES} e^{\frac{V_{BEQ}}{V_T}} + V_{BEQ}$$

Notiamo però che questa è un'equazione trascendente con incognita V_{BEQ} , di difficile risoluzione analitica. Il meglio che possiamo fare è trovare una soluzione approssimata per via numerica (a mano o sfruttando un calcolatore);

• **Metodo analitico semplificato**

Prendiamo quindi un modello a caduta costante della giunzione BE (assunto il BJT in ZAD). Questo sarà esattamente uguale al modello a caduta costante del diodo preso in 6.1.3, con:

$$V_{BEQ} = V_\gamma = 0.7 \text{ V}$$

Sostituendo ciò nell'equazione della retta di carico (cioè semplicemente risolvendo la maglia), si ottiene:

$$I_{BQ} = \frac{V_{BB} - V_\gamma}{R_B}$$

cioè siamo riusciti a semplificare di molto (approssimando) un'equazione trascendente con incognita V_{BEQ} , rendendola qualcosa che siamo effettivamente capaci di risolvere.

A questo punto l'effetto Early viene trascurato anche sulla caratteristica di uscita, quindi si usa quindi la semplice relazione:

$$I_{CQ} = \beta_F I_{BQ}$$

La V_{CEQ} si avrà sempre dalle equazioni di maglia:

$$V_{CEQ} = V_{CC} - R_C \beta_F I_{BQ}$$

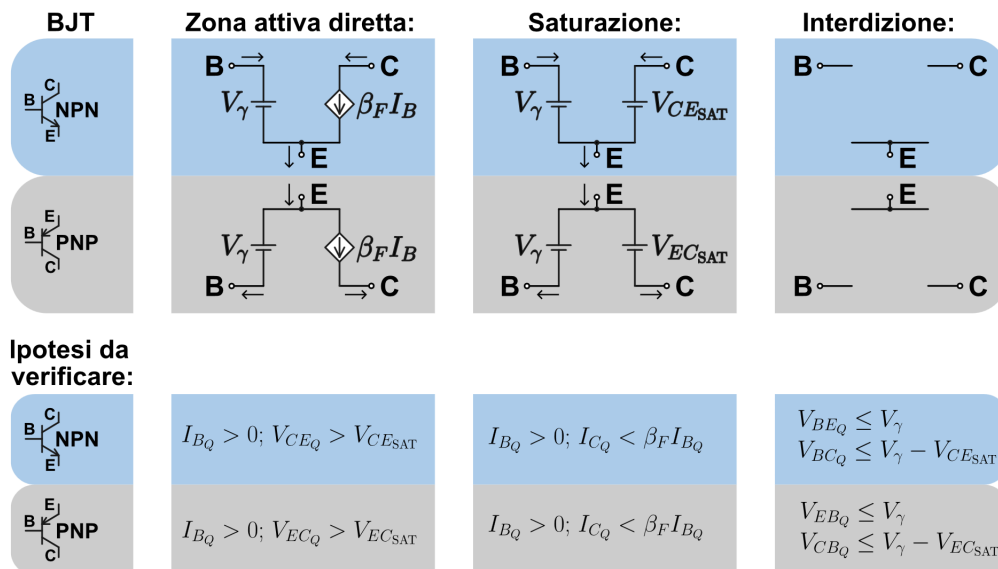
verificando che $V_{CEQ} > V_{CE,sat} \approx 0.2V$.

Avremo quindi trovato, in maniera approssimata, tutte le condizioni a riposo, cioè le grandezze:

$$V_{BEQ}, I_{BQ}, V_{CEQ}, I_{CQ}$$

• **Metodo analitico semplificato generale**

Estendiamo quindi il metodo di risoluzione alle regioni di saturazione e interdizione (oltre la ZAD), mostrando anche le condizioni da verificare, nel seguente schema:



Discutiamo lo schema:

– **Zona Attiva Diretta**

Questo è il modello che abbiamo discusso finora, dove assumiamo la caduta fra base ed emettitore come costante (pari a $V_\gamma \approx 0.7V$), e usiamo la semplice approssimazione lineare $I_C = \beta_F I_B$ per la corrente al collettore. La situazione al PNP è la stessa ma invertendo collettore ed emettitore.

Notiamo che in questo caso dobbiamo verificare che la corrente alla base sia maggiore di 0 (cioè siamo in transistor aperto), ovvero $I_{BQ} > 0$, e che la tensione collettore-emettitore sia maggiore di quella di saturazione, ovvero $V_{CEQ} > V_{CEsat}$.

– Saturazione

In saturazione sostituiamo il generatore di corrente al collettore con una tensione costante pari alla tensione di saturazione (che assumiamo già raggiunta)

$$V_{CE\text{ sat}}.$$

In questo caso le condizioni da soddisfare saranno che la corrente alla base sia comunque maggiore di 0, ovvero $I_{BQ} > 0$, e che anche la corrente al collettore sia minore del valore in ZAD, cioè $I_C < \beta_F I_B$;

– Interdizione

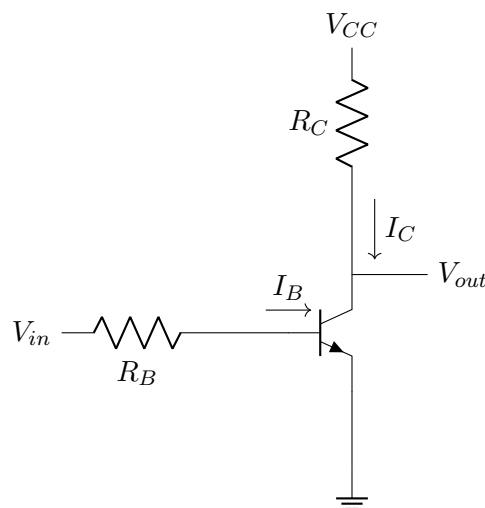
In interdizione banalmente prendiamo circuiti aperti.

Qui le condizioni da soddisfare saranno semplicemente che le tensioni a base-emettitore e base-collettore siano:

$$V_{BEQ} \leq V_\gamma, \quad V_{BCQ} \leq V_\gamma - V_{CE\text{ sat}}$$

11.1 Analisi di un inverter a carico resistivo

Vediamo quindi un primo circuito formato da transistor, formato da un singolo transistor NPN:



Assumiamo che la tensione di ingresso V_{in} varia da 0 fino al valore V_{CC} (ad esempio, 5 V) a passi discreti, e che il valore della tensione di uscita viene registrata una volta esauriti i transistori. In questo caso abbiamo che il BJT attraversa varie zone di funzionamento (interdizione, ZAD e saturazione) in diversi intervalli della V_{in} applicata.

- $0 \leq V_{in} < V_\gamma$

Qui facciamo l'ipotesi del BJT interdetto, cioè poniamo:

$$I_C = I_B = 0, \quad V_{out} = V_{CC}$$

Verifica

La verifica si fa ponendo che le tensioni a base-emettitore e base-collettore siano minori dei valori di riferimento, cioè dire:

$$V_{BE} = V_{in} \leq V_\gamma$$

che è immediato dalle ipotesi fatte, e:

$$V_{BC} = V_{BE} - V_{CE} = V_{in} - V_{CC} < V_{\gamma} - V_{CC} < V_{\gamma} - V_{CE\text{sat}}$$

Questo è verificato dal fatto che V_{CC} è sicuramente maggiore della tensione di saturazione.

In questa situazione stiamo fornendo una tensione *piccola* a V_{in} , e stiamo ottenendo indietro una tensione ON: abbiamo quindi invertito con successo il segnale;

- $V_{\gamma} \leq V_{in} < V_{in\text{sat}}$
In questo caso facciamo l'ipotesi del transistor in ZAD, per cui sfruttiamo il modello analitico semplificato:

$$I_B = \frac{V_{in} - V_{\gamma}}{R_B}, \quad I_C = \beta_F I_B$$

da questo calcoliamo la tensione di uscita come:

$$V_{out} = V_{CC} - R_C I_C = V_{CC} - R_C \beta_F \frac{V_{in} - V_{\gamma}}{R_B}$$

Verifica

La verifica si fa ponendo che la corrente I_B alla base sia maggiore di zero, cosa verificata per ipotesi dal fatto che:

$$I_B = \frac{V_{in} - V_{\gamma}}{R_B} > 0$$

Si deve quindi verificare che la V_{CE} sia maggiore della tensione di saturazione, cioè:

$$V_{CE} > V_{CE\text{sat}}$$

Questa ipotesi è verificata finché scegliamo una $V_{in\text{sat}}$ che la rispetti. Calcoliamone quindi il valore, ponendo:

$$V_{CE\text{sat}} = V_{CC} - \frac{R_C \beta_F}{R_B} (V_{in\text{sat}} - V_{\gamma}) \implies V_{in\text{sat}} = V_{\gamma} + \frac{R_B}{R_C \beta_F} (V_{CC} - V_{CE\text{sat}})$$

- $V_{in} \geq V_{in\text{sat}}$

Superando la $V_{in\text{sat}}$ appena calcolata, il transistor entrerà chiaramente in saturazione, per cui prenderemo il suo modello come tale e ricaveremo:

$$I_B = \frac{V_{in} - V_{\gamma}}{R_B}, \quad I_C = \frac{V_{CC} - V_{CE\text{sat}}}{R_C}, \quad V_{out} = V_{CE\text{sat}}$$

Verifica

La verifica qui si fa ponendo la corrente $I_B > 0$, e la corrente al di sotto di quella in ZAD, per cui:

$$I_B > 0$$

che è verificato come nel caso precedente, e:

$$I_C \leq \beta_F I_B$$

Questo è immediato guardando alla caratteristica di uscita (la corrente al di sopra di quella di saturazione, nella zona di "piegamento delle curve", raggiunge un picco più o meno costante, e se siamo saturazione ne siamo sicuramente al di sotto). Possiamo comunque verificarla matematicamente ponendo una differenza di tensione dalla tensione di saturazione all'ingresso $V_{in\,sat}$, per cui:

$$V_{in} = \Delta V_{in} + V_{in\,sat}, \quad \Delta V_{in} > 0$$

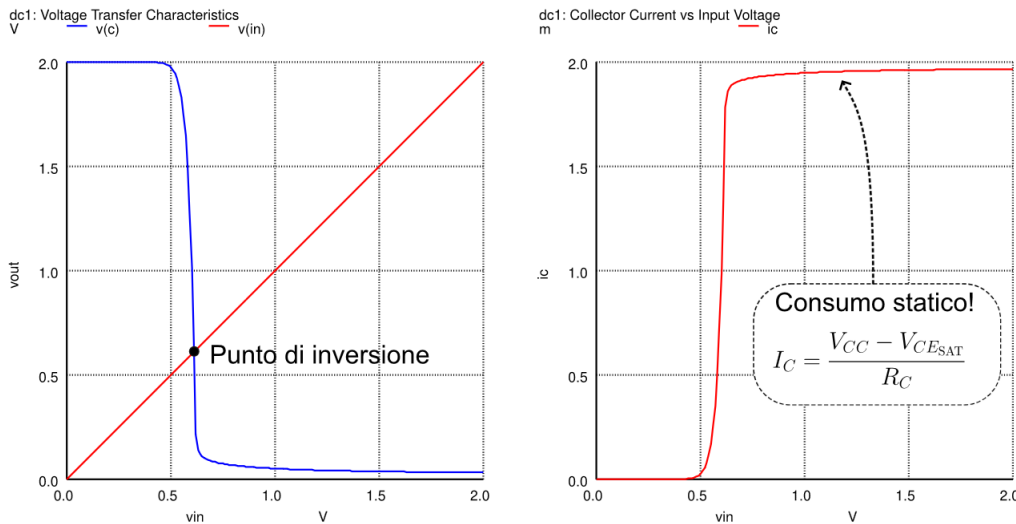
$$I_B = \frac{\Delta V_{in}}{R_B} + \frac{V_{in\,sat} - V_\gamma}{R_B} = \frac{\Delta V_{in}}{R_B} + \frac{V_{CC} - V_{CE\,sat}}{\beta_F R_C}$$

$$\beta_F I_B = \frac{\beta_F \Delta V_{in}}{R_B} + \frac{V_{CC} - V_{CE\,sat}}{R_C} = \frac{\beta_F \Delta V_{in}}{R_B} + I_C$$

che è quanto volevamo dimostrare (guardando agli estremi).

Notiamo che a questo punto il nostro inverter è a livello logico OFF, fornito un ingresso ON: abbiamo quindi invertito la tensione in entrata (che è ciò che dovrebbe fare un inverter).

Dopo aver discusso la modellizzazione matematica dell'inverter, vediamo il grafico dell'uscita V_{out} in funzione della tensione V_{in} di ingresso, per studiarne il comportamento al variare dei parametri di ingresso e quindi le relazioni fra i livelli logici.



Ciò che notiamo è che per tensioni in ingresso basse, come abbiamo detto, otteniamo valori logici in uscita alti, e per tensioni di ingresso alte, otteniamo valori logici in uscita bassi. Questo è in sintonia con il funzionamento che ci aspettiamo dall'inverter. Notiamo anche che nella zona centrale abbiamo una decrescita della tensione V_{out} (perché ci troviamo nella regione ZAD e ci stiamo spostando verso la saturazione), fino ad un punto (cosiddetto di *inversione*) incontrato quando la V_{in} incrocia la V_{out} .

Notiamo che questo design dell'inverter solleva uno dei punti critici che avevamo notato in 7.1.3: quando siamo nello stato V_{in} alto (e quindi V_{out} basso), stiamo consumando corrente a vuoto, cioè la corrente di collettore che avevamo considerato nella regione di saturazione:

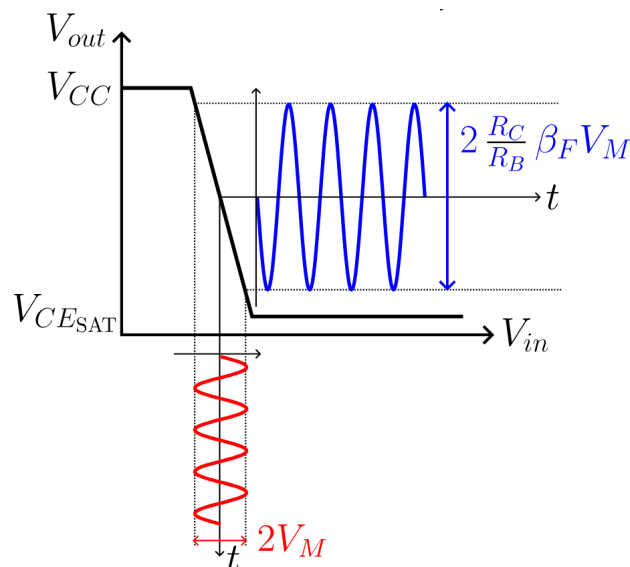
$$I_C = \frac{V_{CC} - V_{CE\text{ sat}}}{R_C}$$

Questa caratteristica, che purtroppo è necessaria per avere il corretto funzionamento dell'inverter, porta ad un consumo di energia costante ogni volta che l'inverter è in inversione con segnale V_{out} basso.

11.2 Amplificazione coi BJT

Vediamo, considerando sempre lo stesso circuito (presentato come un inverter) della sezione precedente, come un transistor BJT può comportarsi da *amplificatore* di potenza per un dato segnale di ingresso.

Si ha infatti che l'inverter a carico resistivo si comporta come un amplificatore di tensione invertente se l'ingresso è polarizzato nella zona di transizione. Questo si ricava direttamente dal grafico della V_{out} in funzione della V_{in} riportato prima:



cioè introducendo alla V_{in} un segnale ristretto alla regione che porta il transistor a lavorare in ZAD, che ricordiamo era:

$$V_\gamma \leq V_{in} < V_{in\text{ sat}}$$

si ottiene sulla V_{out} lo stesso segnale, invertito, ed amplificato da $V_{CE\text{ sat}}$ a V_{CC} . L'amplificazione di tensione sarà la seguente:

$$A_V = -\beta_F \frac{R_C}{R_B}$$

Notiamo che un difetto di questo circuito è che l'amplificazione può variare fortemente con i parametri del transistor e la temperatura. Per questo avremo bisogno di realizzare reti amplificatrici migliori.

11.2.1 Note sugli amplificatori

Facciamo quindi una trattazione dettagliata di cos'è un *amplificatore* in primo luogo. Di base, un amplificatore sarà un dispositivo elettronico formato da due porte:

- Una porta sarà rivolta verso un generatore di *segnale* V_s , con la sua resistenza R_s . Solitamente il generatore di segnale è piccolo (rappresenta il segnale che vogliamo amplificare). Chiamiamo le grandezze su questa porta V_i e i_i ;
- L'altra porta sarà rivolta verso un *carico*, rappresentato dall'impedenza di carico R_L . Questa vorrà vedere una versione amplificata del segnale erogato da V_s . Esistono diversi tipi di amplificazione, fra cui amplificazione di *tensione* e amplificazione di *corrente*. In generale, comunque, vogliamo avere un'amplificazione della *potenza* al carico, da:

$$P = VI = \frac{V^2}{R} = I^2 R$$

Chiamiamo le grandezze su questa porta V_o e i_o ;

Ci deve essere chiaro che un amplificatore è sempre un componente *attivo*, cioè oltre alle porte di ingresso e di uscita di cui abbiamo parlato, dovrà essere connesso all'alimentazione con V_{CC} e V_{EE} .

Definiamo diversi parametri riguardo agli amplificatori, fra cui:

- **Guadagno di tensione** A_V
Questa equivarrà all'aumento di tensione che vediamo sul carico rispetto al segnale, cioè:

$$A_V = \left. \frac{V_o}{V_i} \right|_{R_L \rightarrow +\infty}$$

Il limite della resistenza ad infinito significa che sostanzialmente prendiamo la tensione imposta alla porta di uscita quasi fosse un aperto;

- **Guadagno di corrente** A_i
Allo stesso modo, possiamo definire l'aumento di corrente che vediamo sul carico rispetto al segnale, cioè:

$$A_i = \frac{i_o}{i_i}$$

- **Impedenza di ingresso** R_i
Questa sarà la resistenza (generalizzata ad impedenza per segnali non DC) vista all'ingresso dell'amplificatore, e cioè:

$$R_i = \frac{V_i}{i_i}$$

- **Impedenza di uscita** R_o
Infine, questa sarà la resistenza (generalizzata ad impedenza per segnali non DC) vista all'uscita dell'amplificatore, e cioè:

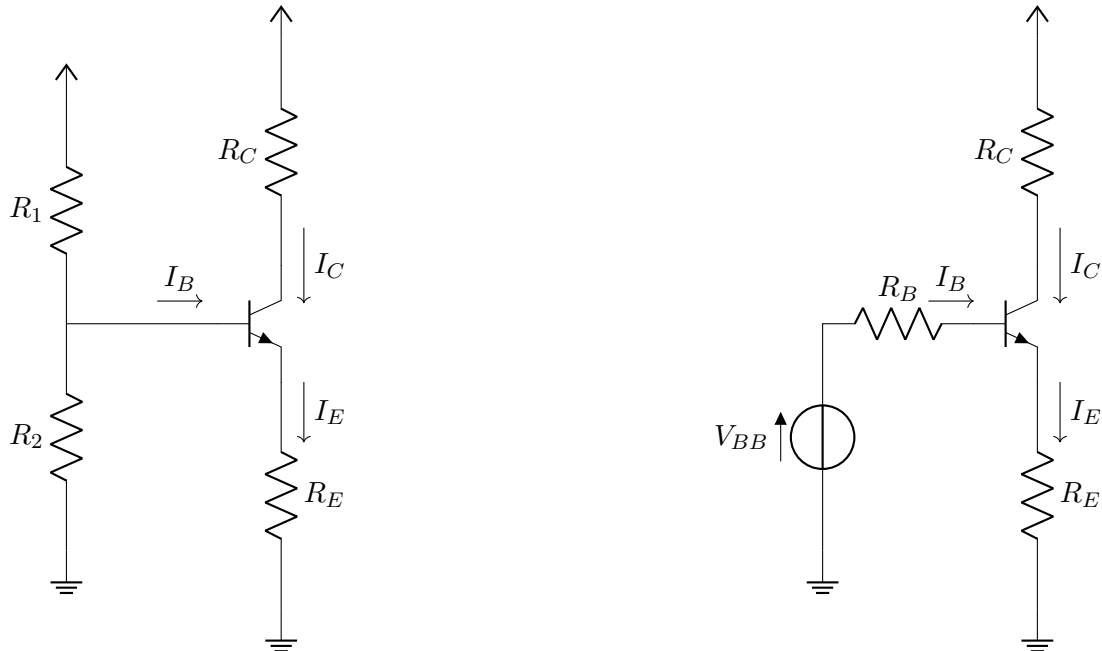
$$R_o = \left. \frac{V_o}{i_o} \right|_{V_s=0}$$

Notiamo che proviamo questa impedenza staccando il generatore di segnale alla porta di ingresso.

11.2.2 Rete a 4 resistori

Abbiamo quindi detto che è necessario fissare una corrente continua costante di collettore o di emettitore che sia poco sensibile alle variazioni della temperatura e dei parametri del transistor.

Consideriamo allora la seguente rete (a 4 resistori):



Questo circuito è il primo tipo di amplificatore che vedremo, basato sul transistor di tipo BJT ad *emettitore comune* (che avevamo iniziato a vedere in 10.1.4).

Il segnale si va ad immettere sul nodo centrale alle due resistenze R_1 ed R_2 , cioè alla base del BJT. Solitamente, si usa l'accoppiamento in AC (anziché in DC, in quanto ci permette di non modificare il punto di riposo): vedremo in 12.2 cosa si intende di preciso. L'idea del partitore formato da R_1 e R_2 è quello di stabilire un *bias*, cioè uno scostamento fisso della tensione alla base del transistor che determina il punto di riposo o *quiescenza* Q del circuito. A questo sommeremo un *piccolo* voltaggio di segnale V_s che verrà effettivamente amplificato.

Ciò che facciamo innanzitutto riportarci all'equivalente di Thevenin del partitore di tensione a sinistra, ricordando che:

$$V_{BB} = \frac{R_2}{R_1 + R_2} V_{CC} \implies R_B = \frac{R_1 R_2}{R_1 + R_2}$$

Questo ci darà la topologia a destra, stabilita la resistenza vista R_B e la tensione V_{BB} all'ammissione dell'amplificatore (il cosiddetto bias che stavamo cercando).

Quindi facciamo l'ipotesi di funzionamento in ZAD (che era la stessa ipotesi dell'amplificatore di prima), imponendo il modello ideale e quindi ricavando:

$$\begin{cases} V_{BB} = R_B I_B + V_\gamma + R_E I_E \\ V_{CC} = R_C I_C + V_{CE} + R_E I_E \\ I_C = \beta_F I_B \\ I_E = I_C + I_B = (\beta_f + 1) I_B \end{cases}$$

Risolvendo tale sistema si ottiene:

$$I_E = (\beta_f + 1)I_B = \frac{V_{BB} - V_\gamma}{\frac{R_B}{\beta_F + 1} + R_E}$$

che sotto determinate ipotesi ci dà il risultato:

$$\begin{cases} V_{BB} \gg V_\gamma \\ R_E \gg \frac{R_B}{\beta_F + 1} \end{cases} \implies I_E \approx I_C \approx \frac{V_{BB}}{R_E}$$

Cioè la corrente di emettitore e, quindi, quella di collettore, sono determinate dai componenti esterni collegati al transistor.

12 Lezione del 30-03-26

12.1 Modello del BJT ai piccoli segnali

Iniziamo a vedere come studiare un modello del BJT adatto all'analisi dei *piccoli segnali*.

Anche qui adotteremo un approccio simile a quello visto in 9.1, cioè dove linearizziamo attorno ad un punto di equilibrio. Iniziamo dalle caratteristiche di ingresso e di uscita ormai ampiamente viste:

$$\begin{cases} V_{BE} = f(I_B, V_{CE}) \\ I_C = g(I_B, V_{CE}) \end{cases}$$

e prevediamo un punto di linearizzazione Q (o di *quiescenza*) dei parametri I_B e V_{CE} :

$$\begin{cases} I_B(t) = I_{BQ} + i_b(t) \\ V_{CE}(t) = V_{CEQ} + v_{ce}(t) \end{cases}$$

per cui le caratteristiche di ingresso e uscita diventeranno:

$$\begin{cases} V_{BE} = f(I_{BQ} + i_b(t), V_{CEQ} + v_{ce}(t)) \\ I_C = g(I_{BQ} + i_b(t), V_{CEQ} + v_{ce}(t)) \end{cases}$$

Procediamo quindi ad esprimere, sempre come avevamo fatto in 9.1, la decomposizione di Taylor troncata al termine lineare di entrambe le caratteristiche:

$$\begin{cases} V_{BE} = f(I_{BQ}) + \left. \frac{\partial f}{\partial i_B} \right|_{I_{BQ}, V_{CEQ}} \cdot i_b(t) + \left. \frac{\partial f}{\partial v_{ce}} \right|_{I_{BQ}, V_{CEQ}} \cdot v_{ce}(t) = V_{BEQ} + v_{be}(t) \\ I_C = g(I_{BQ}) + \left. \frac{\partial g}{\partial i_B} \right|_{I_{BQ}, V_{CEQ}} \cdot i_b(t) + \left. \frac{\partial g}{\partial v_{ce}} \right|_{I_{BQ}, V_{CEQ}} \cdot v_{ce}(t) = I_{CQ} + i_c(t) \end{cases}$$

dove i termini linearizzati sono::

$$\begin{cases} v_{be}(t) \approx \left. \frac{\partial f}{\partial i_B} \right|_Q \cdot i_b(t) + \left. \frac{\partial f}{\partial v_{ce}} \right|_Q \cdot v_{ce}(t) \\ i_c(t) \approx \left. \frac{\partial g}{\partial i_B} \right|_Q \cdot i_b(t) + \left. \frac{\partial g}{\partial v_{ce}} \right|_Q \cdot v_{ce}(t) \end{cases}$$

Parametri ibridi

I parametri differenziali in forma:

$$\frac{\partial\{f, g\}}{\partial\{I_B V_{CE}\}} \Big|_Q$$

si prestano ad una parametrizzazione a **parametri ibridi**, già vista ad esempio nei testi di elettrotecnica. In particolare possiamo definire i parametri:

$$\begin{cases} h_{ie} = \bar{h}_{11} = \frac{\partial f}{\partial I_B} \Big|_Q \\ h_{re} = \bar{h}_{12} = \frac{\partial f}{\partial V_{CE}} \Big|_Q \\ h_{fe} = \bar{h}_{21} = \frac{\partial g}{\partial I_B} \Big|_Q \\ h_{oe} = \bar{h}_{22} = \frac{\partial g}{\partial V_{CE}} \Big|_Q \end{cases}$$

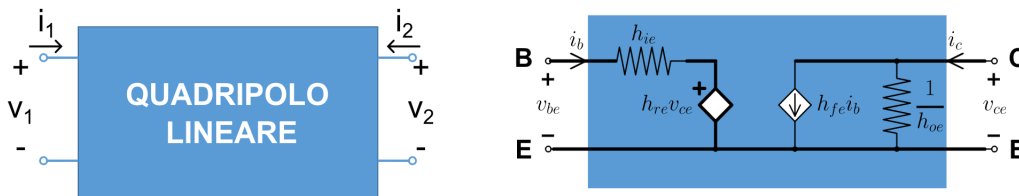
Ricordiamo dai testi che la forma generale di una sintesi a parametri h è:

$$\begin{cases} \dot{V}_1 = \bar{h}_{11}\dot{I}_1 + \bar{h}_{12}\dot{V}_2 \\ \dot{I}_2 = \bar{h}_{21}\dot{I}_1 + \bar{h}_{22}\dot{V}_2 \end{cases}$$

o come matrice:

$$\begin{pmatrix} \dot{V}_1 \\ \dot{I}_2 \end{pmatrix} = \bar{h} \begin{pmatrix} \dot{I}_1 \\ \dot{V}_2 \end{pmatrix}$$

Su un transistor prendiamo la corrente i_1 come la I_B entrante alla base,, la i_2 come la I_C entrante al collettore, e chiaramente l'emettitore comune. Questo è meglio esemplificato dal seguente schema:



Le componenti di \bar{h} si ricavano quindi come:

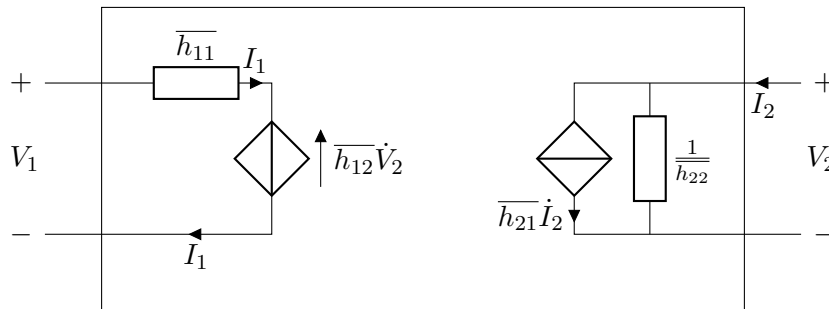
$$h : \begin{pmatrix} h_{ie} = \bar{h}_{11} = \frac{\dot{V}_1}{\dot{I}_1} \Big|_{\dot{V}_2=0} & h_{re} = \bar{h}_{12} = \frac{\dot{V}_1}{\dot{V}_2} \Big|_{\dot{I}_1=0} \\ h_{fe} = \bar{h}_{21} = \frac{\dot{I}_2}{\dot{I}_1} \Big|_{\dot{V}_2=0} & h_{oe} = \bar{h}_{22} = \frac{\dot{I}_2}{\dot{V}_2} \Big|_{\dot{I}_1=0} \end{pmatrix}$$

Questo tipo di sintesi ha un significato interessante sui vari parametri, e in particolare su dispositivi come il BJT, in quanto associa ogni parametro ad una particolare proprietà del transistor:

- \bar{h}_{11} : impedenza di ingresso;
- \bar{h}_{12} : guadagno di tensione inverso;
- \bar{h}_{21} : guadagno di corrente diretto;
- \bar{h}_{22} : ammettenza di uscita.

Circuito equivalente

Il circuito equivalente che possiamo formare da una rappresentazione a parametri ibridi è il seguente:



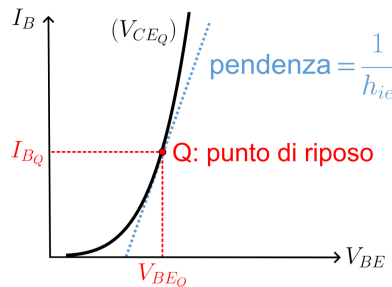
da cui l'impedenza $\overline{h_{11}}$ e l'ammettenza $\overline{h_{22}}$ vengono rese come impedenze rispettivamente in serie e in parallelo, l'inverso dell'amplificazione di tensione $\overline{h_{12}}$ come il coefficiente di un generatore di tensione pilotato, e l'amplificazione di corrente $\overline{h_{21}}$ come il coefficiente di un generatore di corrente pilotato.

Significato dei parametri

Approfondiamo qual'è il significato dal punto di vista fisico dei parametri ibridi appena calcolati sul BJT.

- **Impedenza di ingresso** $h_{ie} = \overline{h_{11}}$

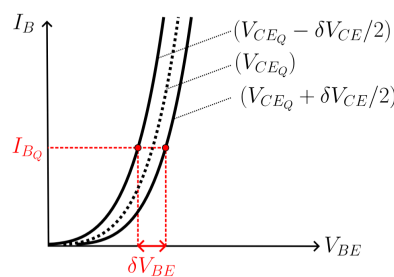
Questo parametro corrisponde al rapporto fra la tensione V_{BE} alla base e la corrente I_B che riusciamo a far scorrere. In questo, è una misura diretta della pendenza della curva di risposta della caratteristica in ingresso:



Risulta analoga alla resistenza differenziale del diodo (con la giunzione BE in ZAD).

- **Guadagno di tensione inverso** $h_{re} = \overline{h_{12}}$

Questo parametro rappresenta il guadagno di tensione che abbiamo sulla porta di ingresso (la base-emettitore). In questo modella il risultato dell'effetto Early sulla caratteristica di ingresso, cioè la dipendenza dal voltaggio V_{CE} alla porta di uscita.

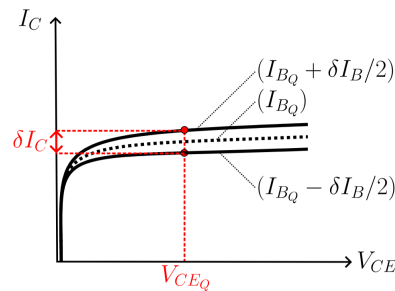


- **Guadagno di corrente diretto** $h_{fe} = \overline{h_{21}}$

Questa è la costante che lega la corrente in entrata alla porta di ingresso (la base) alla corrente in I_C . In questa è direttamente legata al fattore di guadagno in corrente β_F :

$$I_C = \beta_F I_B$$

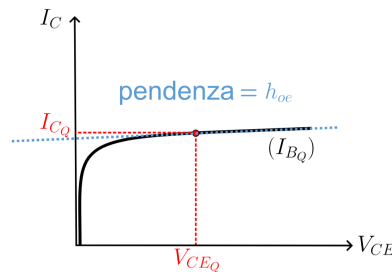
anche se notiamo che nel parametro h_{fe} reale, ottenuto linearizzando il transistor su un punto di riposo Q , h_{fe} è leggermente superiore a β_F .



- **Ammettenza di uscita** $h_{oe} = \overline{h_{22}}$

Infine, questo parametro (o meglio il suo opposto) rappresenta l'impedenza sulla porta di uscita, quindi la piccola resistenza che vi associavamo per causa dell'effetto di Early.

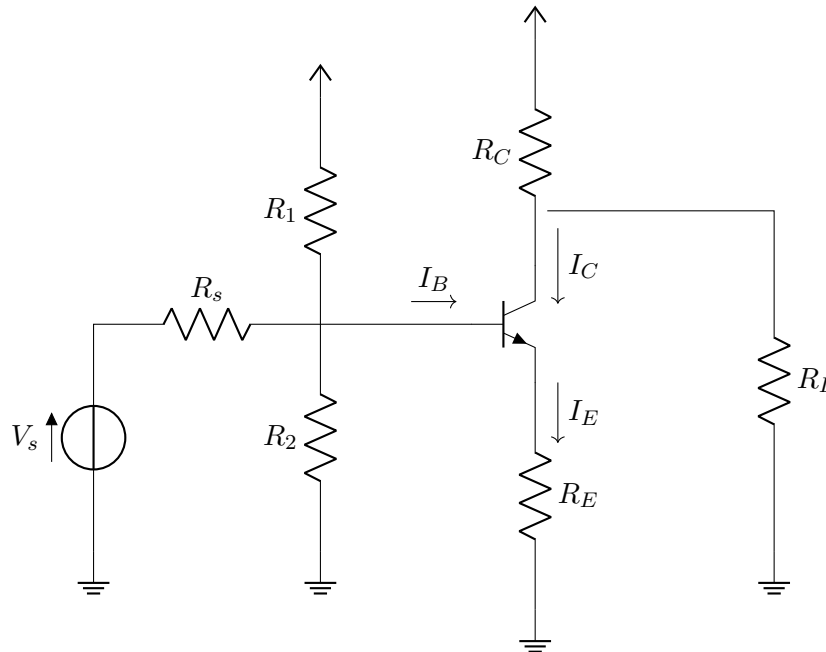
In termini grafici, è legato alla pendenza della curva della caratteristica di uscita nel punto di linearizzazione:



Abbiamo quindi trovato un altro modello per il transistor BJT, che possiamo affiancare ad i modelli già trovati per i BJT PNP e NPN. Notiamo che il modello è equivalente per entrambi i tipi di transistor, ovvero sia il PNP che l'NPN collassano allo stesso modello (e quindi pressappoco lo stesso comportamento) ai piccoli segnali.

12.2 Accoppiamento AC

Riprendiamo l'amplificatore a 4 resistori visto in 11.2.1, e vediamo cosa accade se si collega il segnale direttamente alla base del BJT, e il carico direttamente al collettore. Chiamiamo questo accoppiamento anche *accoppiamento DC*.



Notiamo subito che questa situazione è fallimentare. Possiamo applicare Thevenin alla rete di sinistra (quella collegata alla base del BJT) per trovare il voltaggio equivalente alla base. Inoltre, possiamo calcolare separatamente i casi con un solo generatore (V_{cc} e V_s) attaccato, e quindi applicare la sovrapposizione degli effetti:

- Per $V_s \neq 0$ e $V_{cc} = 0$ abbiamo che la coppia R_1, R_2 è collegata in parallelo a massa, e quindi la V_{THEV} equivalente si calcola come:

$$V'_{THEV} = V_s \frac{R_1 || R_2}{R_1 || R_2 + R_s}$$

cioè la tensione data dal partitore di tensione formato dalla coppia $R_1 || R_2$ e R_s ;

- Per $V_s = 0$ e $V_{cc} \neq 0$ abbiamo invece che la coppia R_s, R_2 è collegata in parallelo a massa, e quindi la V_{THEV} equivalente si calcola come:

$$V''_{THEV} = V_{cc} \frac{R_s || R_2}{R_s || R_2 + R_1}$$

cioè la tensione data dal partitore di tensione formato dalla coppia $R_s || R_2$ e R_1 .

Trovata la tensione alla base, calcoliamo con lo stesso metodo la tensione sull'alimentazione del transistor (al collettore):

$$V_{cc} \frac{R_L}{R_L + R_C}$$

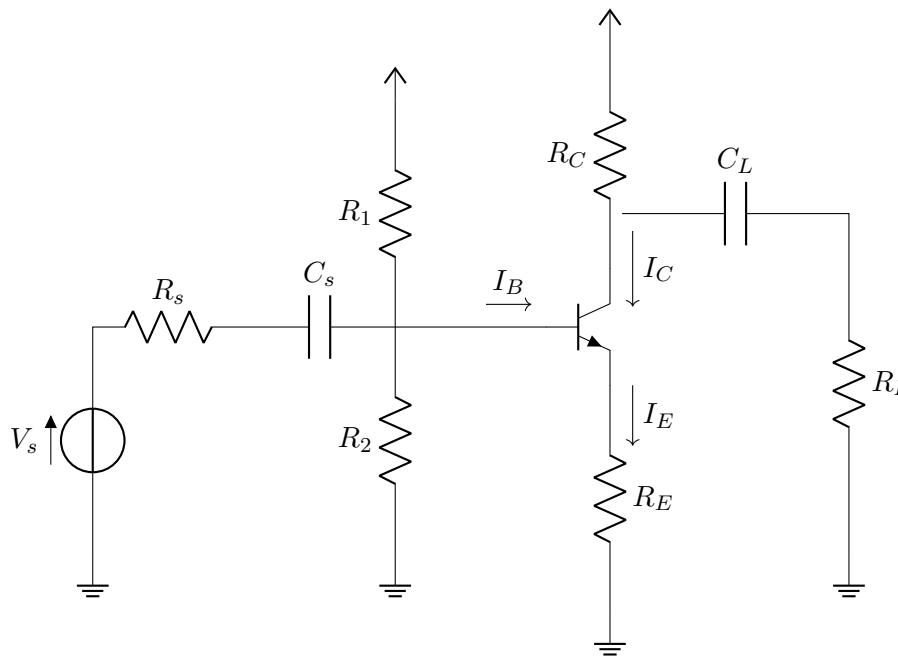
dato dal partitore di tensione formato da R_L al carico e R_C al collettore, che si trovano in serie.

Seguiranno quindi alcune considerazioni:

- Se $R_L \ll R_0$, l'alimentazione $V_{cc} \frac{R_L}{R_L + R_C}$ andrà a tendere a 0, e quindi tenderà a zero anche la V_{CEQ} . Il transistor BJT lavorerà quindi nel regime di saturazione, dove abbiamo detto non è efficiente come amplificatore;
- Se $R_S \ll R_1, R_2$, la tensione alla base tenderà a zero (V_s spento per sovrapposizione degli effetti). Se anche la base va a zero, il transistor torna a lavorare poco, e addirittura si spegne per $V_{BE} < V_\gamma$.

Abbiamo quindi che, in generale, si presenta il seguente problema: l'accoppiamento DC interferisce con la polarizzazione (il punto Q) del BJT.

La soluzione che daremo a questo problema sarà quello di usare il cosiddetto *accoppiamento AC*:



In questo accoppiamento mettiamo due condensatori (C_s al segnale e C_L al carico), detti *condensatori di disaccoppiamento*. Abbiamo quindi che come nell'accoppiamento DC sia il generatore di sorgente che il carico vengono connessi al BJT (con lo schema di polarizzazione a 4 resistori), ma stavolta tramite i condensatori di disaccoppiamento C_s, C_L .

Vediamo come questi risolvono il problema:

- In corrente continua (DC), i condensatori offrono una impedenza infinita, aprendo le maglie relative alla sorgente e al carico. Il punto Q del BJT viene quindi determinato unicamente dallo schema di polarizzazione a 4 resistori;
- In alternata (AC), i condensatori si considerano corto-circuiti, permettendo al segnale (tempo-variante) di attraversare lo stadio amplificatore e raggiungere il carico (questo è il regime di operazione che vogliamo dall'amplificatore).

Abbiamo quindi risolto il problema. Vediamo però che le frequenze per cui i condensatori possono considerarsi in cortocircuito dipendono dalla resistenza vista da ciascuno di essi:

$$\omega_{z,X} = \frac{1}{R_X C_X}$$

È possibile calcolare R_X dal circuito: quindi decidere $\omega_{z,X}$ sarà fatto dimensionando opportunamente la capacità (selezionando C_X).

Questa è in verità una semplificazione: la visione di insieme pone che l'effetto collettivo di tutte le capacità nel circuito (anche quelle intrinseche ai componenti) sia equivalente a quello di un cortocircuito per i segnali alla frequenza di operazione del circuito. In questo modo, quando si studia ad esempio l'amplificatore (o comunque qualsiasi circuito che elabora segnali) si possono trascurare le capacità ed assumerle cortocircuiti.

12.3 Riassunto sull'analisi dei circuiti non lineari

Facciamo quindi un riassunto su come gestire i circuiti non lineari, in DC e in AC:

- **DC (punto di riposo)**

In DC, l'analisi mira perlopiù a trovare il punto di riposo Q di componenti (come ad esempio un BJT). In questo, consiste nel:

1. Disattivare i generatori di segnale ($v(t)$ in cortocircuito, $i(t)$ in circuito aperto);
2. Sostituire i condensatori con un circuito aperto e le induttanze con un circuito chiuso (come dalla classica analisi *steady-state* dei circuiti lineari);
3. Sostituire i componenti non lineari con il loro modello per grandi segnali.

- **AC (media frequenza)**

In AC, l'analisi mira invece a calcolare l'effetto dei segnali per i componenti non lineari, già assunti a riposo (quiescenza) e quindi linearizzati. In questo, consiste nel:

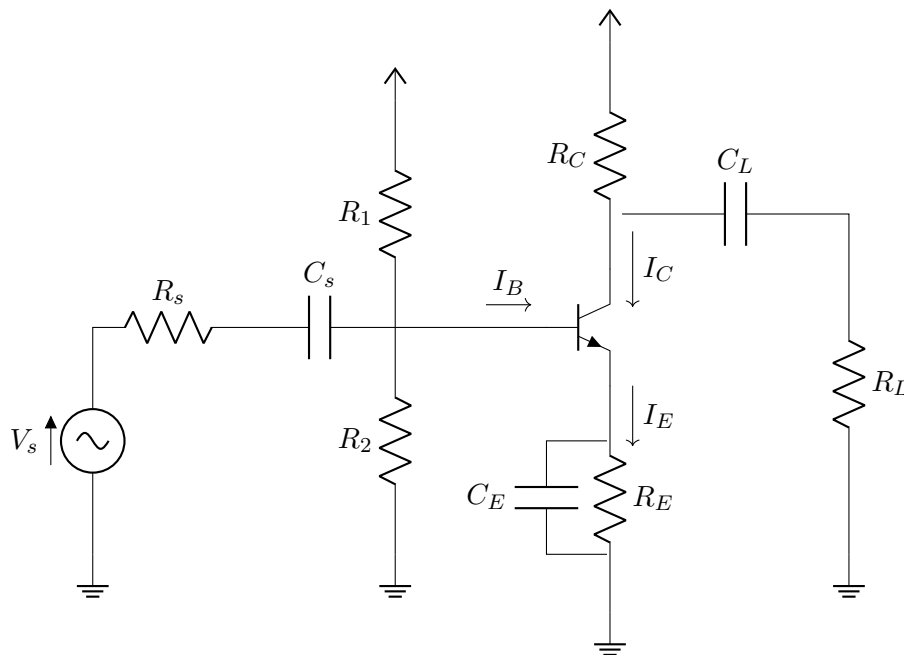
1. Disattivare i generatori costanti ($V(t)$ in cortocircuito, $I(t)$ in circuito aperto);
2. Sostituire i condensatori con un circuito chiuso e le induttanze con un circuito aperto, assumendo, come detto sopra, di aver tarato capacità per lasciar passare i segnali alla frequenza che ci interessa;
3. Sostituire i componenti non lineari con il loro modello per piccoli segnali. Notiamo che i condensatori intrinseci ai componenti non lineari vengono considerati in circuito aperto.

13 Lezione del 31-03-26

13.0.1 Analisi dell'amplificatore ad emettitore comune

Siamo quindi pronti ad effettuare un'analisi vera e propria dell'amplificatore ad emettitore comune. Lo avevamo introdotto in 11.2.2 parlando di amplificatori e dell'accoppiamento termico dei BJT; quindi lo avevamo ripreso nella scorsa lezione per parlare dell'accoppiamento AC. Adesso proviamo effettivamente a studiarlo come circuito non lineare nei regimi di operazione AC e DC.

Il circuito da studiare sarà quindi il seguente, dove notiamo dell'aggiunta della capacità ad emettitore C_E (la sua motivazione ci sarà chiara a breve):

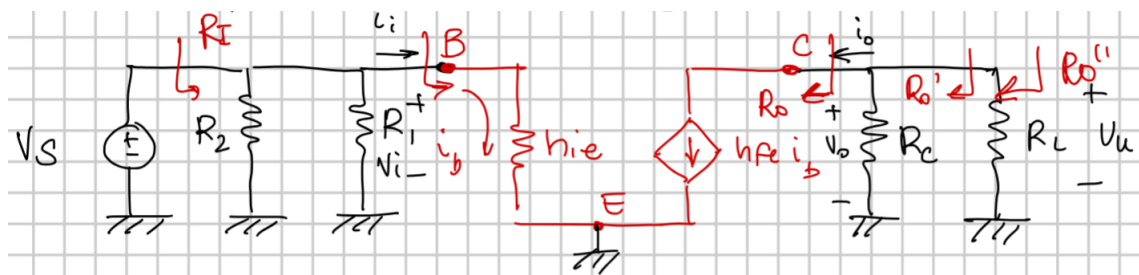


Notiamo che faremo due trattazioni:

- La prima prenderà considererà la capacità di accoppiamento C_E all'emettitore, e mostrerà come questa aumenta la capacità di amplificazione del circuito;
- La seconda invece non la considererà, e mostrerà come questo riduce l'amplificazione effettiva.

Con capacità di accoppiamento

Consideriamo la componente DC già calcolata col punto di quiescenza determinato dal bias di R_1 e R_2 . In questo caso potremo passare all'analisi AC considerare tutte le capacità di accoppiamento come cortocircuito, inclusa quella ad R_E , e sostituire il transistor con il modello ai piccoli segnali. Avremo quindi l'impedenza di ingresso h_{ie} e la sorgente di corrente pilotata in h_{fe} da i_b (per la legge $i_c = h_{fe}i_b$). Il circuito equivalente (cortesia della collega Imane Rajy) sarà:



Vogliamo calcolare i 4 parametri (guadagni di corrente e tensione, impedenze di ingresso e uscita) considerati per gli amplificatori in 11.2.1.

- **Guadagno di corrente** A_i

Questo sarà immediato dalla corrente imposta all'uscita dal transistor, cioè:

$$i_o = i_c = h_{fe}i_b$$

e la corrente di ingresso, che con R_E non considerata (dal corto su C_E per segnali AC) consiste solamente in:

$$i_i = i_b$$

Il guadagno di corrente sarà quindi il rapporto:

$$A_i = \frac{i_o}{i_i} = \frac{h_{fe}i_b}{i_b} = h_{fe}$$

cioè il parametro h_{fe} , senza sorprenderci particolarmente, rappresenta il guadagno in corrente dell'amplificatore (come rappresentava il guadagno di corrente diretto del BJT stesso in parametri ibridi).

- **Guadagno di tensione** A_V

Per il guadagno di tensione avremo bisogno di un'indicazione per la tensione in uscita all'amplificatore. Questa sarà data, ad esempio, dalla caduta su R_C (trascurando la resistenza di carico R_L):

$$v_o \Big|_{R_L \rightarrow +\infty} = -h_{fe}i_b R_C$$

mentre la tensione in ingresso sarà semplicemente il segnale fornito:

$$v_i = V_s$$

Cerchiamo quindi una forma chiusa in V_s della v_o in modo da esprimere tutto in funzione di V_s e ricavare un guadagno. Per fare ciò prendiamo la i_b in funzione della V_s :

$$i_b = \frac{V_s}{h_{ie}}$$

e sostituiamo in v_o , da cui il guadagno:

$$v_o = -h_{fe}R_C \frac{V_s}{h_{ie}} \implies A_V = -\frac{h_{fe}R_C}{h_{ie}}$$

Notiamo qui una particolarità di questo tipo di amplificatore: abbiamo che è *invertente* (il cosiddetto **inverting amplifier**), cioè la tensione in uscita è sfasata di 180° rispetto alla tensione in entrata. Inoltre, notiamo che il guadagno calcolato è quello visto all'uscita dell'amplificatore, posta la R_L ad infinito (cioè come circuito aperto). Chiaramente collegando una resistenza di carico diversa si avrebbe un calo di tensione.

- **Impedenza di uscita** R_o

Per calcolare l'impedenza di uscita si pone un generatore di prova V_p all'uscita e si spegne il generatore di segnale V_s . A questo punto si può considerare la resistenza vista al collettore del transistor BJT, a valle della resistenza R_C , o a valle della resistenza R_L .

Al collettore del BJT si avrà, spento il generatore V_s , corrente i_b nulla e di conseguenza corrente i_c nulla, per cui in:

$$R_o = \left. \frac{v_o}{i_o} \right|_{i_o \rightarrow 0} = +\infty$$

cioè impedenza infinita.

Avremo quindi la sola R_C nel caso in cui si consideri la resistenza alla R_C in sé per sé, o dal parallelo:

$$R_o = \frac{v_o}{i_o} = R_C || R_L$$

nel caso in cui si consideri anche la resistenza di carico.

• **Impedenza di ingresso R_i**

L'impedenza di ingresso verrà calcolata in maniera simile, ponendo un generatore di prova al posto di V_s . In questo caso, visto che l'emettitore è collegato direttamente a massa, si avrà che tutto il lato destro del circuito è irrilevante al calcolo dell'impedenza in ingresso. Vedremo che togliendo la capacità C_E questo non sarà più il caso.

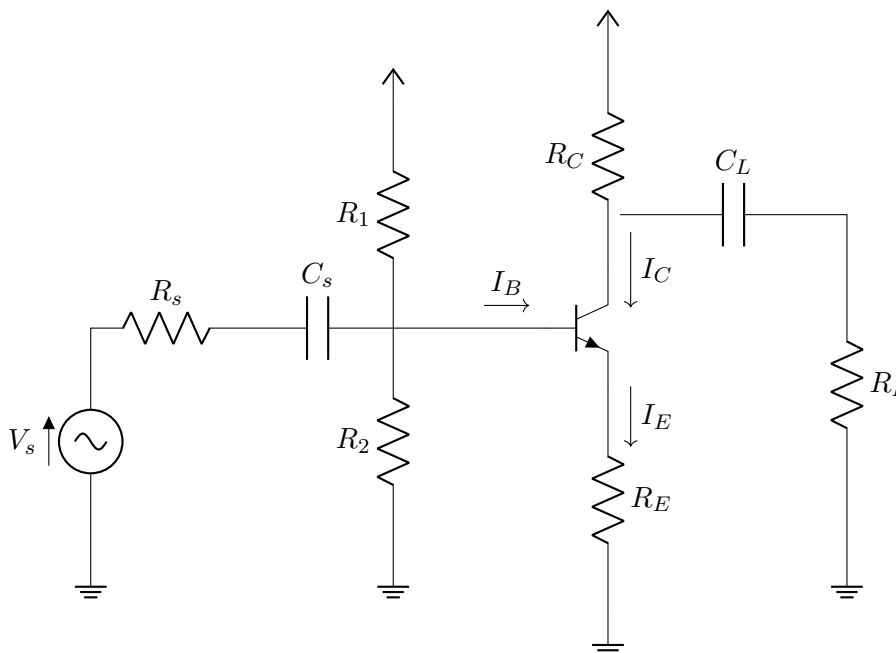
Prendiamo quindi il parallelo fra la resistenza associata al modello dei piccoli segnali del transistor, e le resistenze del partitore di bias R_1 e R_2 :

$$R_i = \frac{v_i}{i_i} = R_1 || R_2 || h_{ie}$$

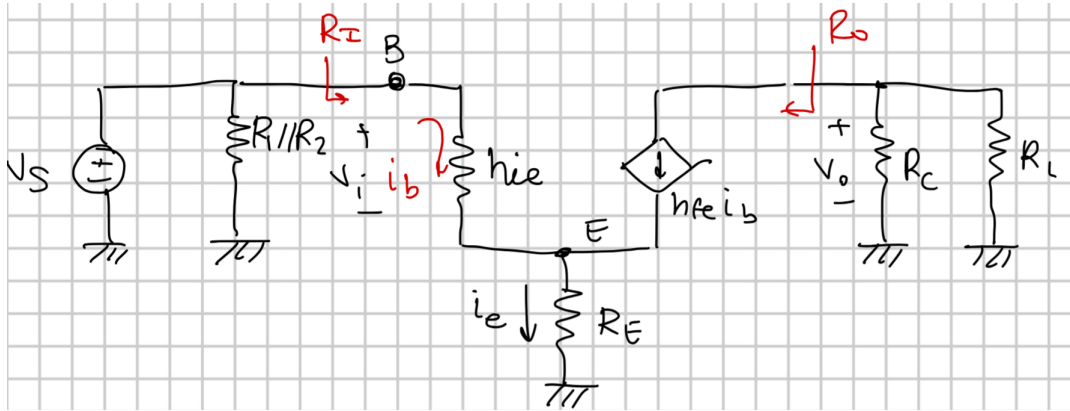
Abbiamo quindi visto come introdurre la capacità di accoppiamento C_E all'emettitore ci permette di avere disaccoppiamento in AC della resistenza R_E , e quindi semplificazioni nei calcoli e un aumento del guadagno di tensione.

Senza capacità di accoppiamento

Vediamo adesso il caso con questa capacità rimossa, cioè preso il circuito:



In tal caso il circuito equivalente che otterremo in AC prevederà la resistenza R_E all'emettitore del transistor, in quanto non potremmo semplificarla via col parallelo a C_E :



Ricalcoliamo quindi gli stessi parametri dell'amplificatore, vedendo come variano.

- **Guadagno di corrente A_i**

Questa non cambia, in quanto le correnti in entrata e in uscita restano le solite:

$$i_i = i_b, \quad i_o = h_{fe}i_b$$

da cui il guadagno corrente legato al parametro ibrido:

$$A_i = \frac{i_o}{i_i} = \frac{h_{fe}i_b}{i_b} = h_{fe}$$

- **Guadagno di tensione A_V**

Cose più interessanti accadono per il guadagno di tensione. Abbiamo che non potremo calcolare solo la caduta sulla resistenza h_{ie} all'ingresso, ma dovremmo tenere conto anche della caduta su R_E :

$$v_i = h_{ie}i_b + V_E$$

Le cose all'uscita restano invece uguali, cioè la tensione di uscita è:

$$v_o \Big|_{R_L \rightarrow +\infty} = -h_{fe}i_b R_C$$

preso come prima R_L come un circuito aperto (tensione imposta a vuoto senza carichi resistivi).

Vediamo quindi come calcolare questa V_E . Questa dipenderà dalla corrente che passa su R_E dalla prima legge di Kirchoff, cioè i_b e $i_c = h_{fe}i_b$:

$$V_E = (i_b + h_{fe}i_b)R_E$$

per cui la v_i totale (e di conseguenza il guadagno in tensione) risultano:

$$\begin{aligned} v_i = i_b h_{ie} + (i_b + h_{fe})R_E &\implies A_v = \frac{v_o}{v_i} \Big|_{R_L \rightarrow +\infty} \\ &= -\frac{h_{fe}i_b R_C}{i_b h_{ie} + (i_b + h_{fe})R_E} = -\frac{h_{fe}R_C}{h_{ie} + (1 + h_{fe})R_E} \end{aligned}$$

Discutiamo questo risultato. L'amplificatore è sempre invertente, in quanto rimane il segno negativo. Si ha però una diminuzione del guadagno ottenuto dato dal componente:

$$(1 + h_{fe})R_E$$

Ci potremmo chiedere come mai non includere sempre la capacità C_E visto che questa aumenta il guadagno. La risposta è che, rimossa la C_E , e assunto R_E sufficientemente grande, si ha approssimativamente:

$$A_v = -\frac{h_{fe}R_C}{h_{ie} + (1 + h_{fe})R_E} \sim -\frac{R_C}{R_E}$$

cioè il calcolo del guadagno A_v diventa triviale. In situazioni dove non è fondamentale ottenere un grande guadagno, ma piuttosto tararlo con precisione agendo su R_E e R_C , e quindi più semplice non includere la capacità di accoppiamento C_E e sfruttare l'approssimazione $\frac{R_C}{R_E}$ per il guadagno.

- **Impedenza di uscita R_o**

L'impedenza in uscita è esattamente analoga alla situazione vista con la capacità di accoppiamento C_E : si pone un generatore di prova V_p all'uscita e si spegne il generatore di segnale V_s . A questo punto si può considerare la resistenza vista al collettore del transistor BJT, a valle della resistenza R_C , o a valle della resistenza R_L .

Al collettore del BJT si avrà, spento il generatore V_s , corrente i_b nulla e di conseguenza corrente i_c nulla, per cui di nuovo impedenza infinita. Avremo quindi la sola R_C nel caso in cui si consideri la resistenza alla R_C in sé per sé, o dal parallelo: $R_C || R_L$ nel caso in cui si consideri anche la resistenza di carico.

- **Impedenza di ingresso R_i**

Veniamo quindi l'impedenza di ingresso, che avevamo anticipato risentiva dal lato uscita del circuito quando si rimuoveva la capacità di accoppiamento. Abbiamo infatti che ponendo il generatore di prova V_p al posto di V_s si ottiene la stessa identica corrente su R_E (e quindi caduta V_E) che avevamo visto per il calcolo del guadagno in tensione. Risulta quindi che:

$$V_p = h_{ie}i_p + (1 + h_{fe})R_E i_p$$

per cui l'impedenza vista è:

$$R_i = \frac{V_p}{i_p} = h_{ie} + (1 + h_{fe})R_E$$

Questo è esattamente lo stesso termine trovato per il guadagno in tensione A_v , e ci mostra come è possibile che l'uscita del circuito influenzi l'impedenza vista all'ingresso. Chiamiamo la legge trovata **regola di riflessione della resistenza**.

Infine notiamo che, chiaramente, questa impedenza viene calcolata prima delle resistenze R_1 e R_2 .