

Indice

1	Interazione radiazione e materia	3
1.1	Tipi di interazione per ciascuna classe di particelle	3
1.2	Perdita di energia per ionizzazione	3
1.2.1	Perdita di energia nei materiali sottili	13
1.3	Lunghezza di radiazione	19
1.4	Scattering Coulombiano multiplo	20
1.5	Formazione del segnale	23
2	Fisica dei materiali semiconduttori	25
2.1	Gli elettroni nei nuclei	25
2.2	Teoria delle bande energetiche	26
2.2.1	Le bande energetiche dei semiconduttori	30
2.2.2	Cristallo, legami covalenti, intervallo proibito	32
2.2.3	Le lacune	35
2.3	I semiconduttori intrinseci	37
2.3.1	Densita' degli stati	37
2.3.2	Probabilita' di occupazione degli stati	37
2.4	I semiconduttori drogati (o <i>estrinsenci</i>)	41
2.4.1	Legge dell'azione di massa	44
2.5	Correnti: il trasporto dei portatori	50
2.5.1	Deriva	51
2.5.2	La resistivita' del semiconduttore	53
2.5.3	Diffusione	57
3	Giunzione <i>pn</i>	60
3.1	Distribuzione di carica, campo elettrico, potenziale elettrico in una giunzione <i>pn</i> generica	60
3.1.1	Correnti nella giunzione <i>pn</i> non polarizzata	67
3.1.2	Correnti nella giunzione <i>pn</i> direttamente polarizzata	68
3.1.3	Correnti nella giunzione <i>pn</i> inversamente polarizzata	70
3.2	Capacita' della giunzione	72
3.3	Svuotamento, sovratensione, rottura	82
4	Raccolta della carica e induzione del segnale	97
4.1	Raccolta della carica	97
4.2	Segnale indotto	97

5	Elaborazione del segnale	107
5.1	Fluttuazioni del segnale e rumore elettronico	109
5.2	Fattore di Fano	109
5.3	Introduzione al rumore elettronico	112
5.4	Quantità fondamentali nella trattazione del rumore	114
5.4.1	RMS	114
5.4.2	Rapporto segnale/rumore	114
5.4.3	Rumore espresso in elettroni equivalenti (ENC)	115
5.5	Analisi spettrale del rumore	115
5.6	Contributi alla potenza di rumore	119
5.7	Amplificatori e filtri	123
5.7.1	Effetto dell'amplificatore sul rapporto segnale/rumore	128
5.7.2	Il formatore	133
5.7.3	Effetto del formatore sul rumore in ingresso	136
5.7.4	Discussione delle caratteristiche dei formatori	137
6	Risoluzione spaziale dei rivelatori	142

1 Interazione radiazione e materia

L'azione del rivelatore dipende da come l'oggetto della rivelazione interagisce con il materiale del rivelatore. Una volta stabilito il processo che vogliamo osservare dobbiamo avere piena conoscenza dei meccanismi fondamentali dell'interazione delle particelle e radiazione in gioco, poiché è solo da questa interazione che possiamo rivelare le particelle.

1.1 Tipi di interazione per ciascuna classe di particelle

Le particelle possono essere rivelate solo tramite la loro interazione con la materia:

- *Particelle cariche*: Interazione Coulombiana tra campo della particella carica in movimento e campo generato dalla struttura nucleare o elettronica della materia attraversata
- *Adroni (carichi e neutri)*: Interazione forte con nuclei del materiale attraversato, con potenziale produzione di particelle secondarie.
- *Elettroni, positroni, fotoni*: Creazione di cascate elettromagnetiche e perdita di energia, con potenziale produzione di particelle secondarie.

Le particelle secondarie sono soggette a loro volta agli stessi tipi di interazione a seconda della loro classe di appartenenza.

L'interazione coulombiana riguarda sia l'interazione con i campi nucleari sia con i campi degli elettroni degli orbitali atomici. Tuttavia, il primo tipo di interazione è molto più raro, quindi il rivelatore si deve affidare unicamente al segnale generato dall'interazione con gli e^- .

1.2 Perdita di energia per ionizzazione

L'interazione elettromagnetica tra una particella carica e l'elettrone atomico nel materiale consiste in un trasferimento di energia che può risultare in:

- eccitazione dell'atomo: l' e^- acquista energia che gli permette di saltare su un orbitale atomico più esterno, portando l'atomo in uno stato eccitato

- ionizzazione dell'atomo: l' e^- acquista talmente tanta energia da riuscire a liberarsi dall'attrazione elettromagnetica che lo lega al nucleo, ionizzando l'atomo

La presenza di secondarie è un vantaggio perché permette di rivelare particelle che non generano direttamente un segnale rilevabile dal detector, ma presenta anche degli effetti di cui tener conto. La ionizzazione prevede la formazione di coppie di ioni, ovvero un atomo ionizzato e un e^- . Nel caso dei rivelatori a semiconduttore, l'atomo ionizzato non si ricombina facilmente e può essere usato per creare il volume di rivelazione. Gli e^- diventano invece portatori di carica liberi che possono muoversi e trasportare l'informazione dell'interazione originale.

Dal punto di vista di un rivelatore sensibile all'energia rilasciata, le particelle che è possibile rivelare sono:

- particelle cariche pesanti, con massa $m \gg m_{e^-}$ - ionizzazione
- elettroni abbastanza veloci - ionizzazione
- neutroni che liberano particelle cariche pesanti - ionizzazione secondarie
- γ e X che generano elettroni - ionizzazione secondarie

La maggior parte dell'energia persa per ionizzazione è ceduta tramite la liberazione di elettroni appartenenti agli orbitali atomici del materiale, che possono essere raccolti sotto forma di segnale. In caso di incontri molto ravvicinati tra la particella passante e gli e^- atomici, l'energia trasferita all' e^- strappato dall'atomo è così elevata che a loro volta essi perdono la propria energia per ionizzazione. In quest'ultimo caso sono chiamati δ -rays, hanno un range molto piccolo, creano una nuvola di cariche secondarie nella zona di ionizzazione primaria, determinando la cosiddetta "clusterizzazione" (raggruppamento) della ionizzazione.

A titolo esemplificativo, ricordiamo che la produzione dei δ -rays ha effetti non trascurabili sulla ricostruzione della posizione e dell'ampiezza del segnale primario. Questo ci fa intuire quanto complesso sia il processo di ricostruzione del segnale generato da una particella primaria.

La perdita media di energia lineare per unità di percorso attraversato nel materiale è detta anche potere di frenamento S :

$$S = -\frac{dE}{dx}$$

Ricavare la generalizzazione di questo risultato, la formula di Bethe-Bloch, risulta un procedimento piuttosto elaborato, tuttavia è di fondamentale importanza nello studio delle interazioni di particelle cariche con la materia.

Per dare una derivazione abbastanza intuitiva nel caso classico di Bloch possiamo immaginare che la particella carica di passaggio stia muovendosi con traiettoria rettilinea passante vicino all'atomo di materiale che vogliamo usare per rivelarla, passando a una distanza minima b da essa che è detta *parametro d'impatto* (vedi Fig. 1).

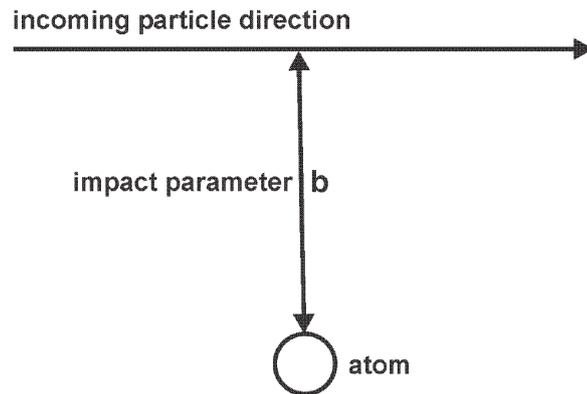


Figura 1: Passaggio della particella carica in prossimità dell'atomo del materiale attraversato alla distanza minima b , detta *parametro d'impatto* [Leroy-Rincoita]

Essa subirà l'interazione coulombiana prima di tutto con gli e- dell'atomo (vedi Fig. 2), fatte le doverose assunzioni secondo cui:

- la traiettoria della particella incidente è (almeno localmente) rettilinea
- la velocità della particella è molto maggiore della velocità degli e- atomici orbitanti (in sostanza e- è fermo rispetto alla particella).

Poiché l'interazione che avviene ad ogni istante si può scomporre in contributo perpendicolare e parallelo alla direzione della particella, si può isolare la componente che non si cancella, ovvero quella perpendicolare. L'interazione dunque dipende solo dalla distanza minima b (detta anche *parametro di impatto*) come $1/b^2$.

Sia le interazioni a grande distanza (b grande) sia ravvicinate (piccolo b) contribuiscono sia alla perdita di energia media, sia alla sua distribuzione, sia al valore più probabile.

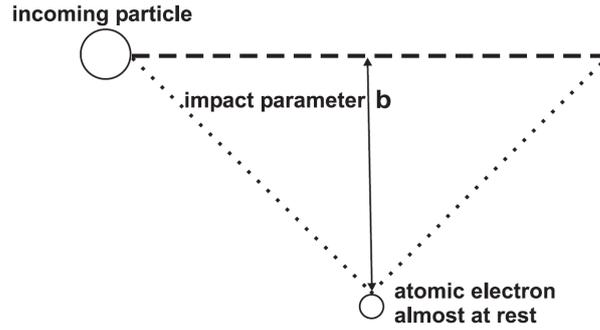


Figura 2: La particella carica passa con direzione rettilinea a distanza b dall'elettrone atomico che si considera a riposo rispetto alla velocità della particella.

Dalla definizione di Forza coulombiana, abbiamo che a meno di una costante k essa ha proprio questo andamento:

$$F = k \frac{ze^2}{b^2}$$

Il valore massimo del campo elettrico decresce all'aumentare della distanza b .

Il tempo di interazione, inoltre, è proporzionale alla distanza tra particella incidente e e^- dell'orbitale atomico e inversamente proporzionale alla velocità della particella incidente. Insomma $t \sim b/v$.

Se integriamo nel tempo questa forza otteniamo il momento trasferito

$$I \propto z \frac{e^2}{b^2} t = \frac{z}{b^2} e^2 \frac{b}{v} = \frac{ze^2}{bv}$$

. Un calcolo più raffinato risulta nel risultato

$$I = \frac{2ze^2}{bv}$$

Questo è anche il momento di rinculo dell' e^- , da cui si può calcolare l'energia cinetica acquisita dall' e^- :

$$W = \frac{I^2}{2m} = \frac{2z^2e^4}{mb^2v^2}$$

Ora dobbiamo passare dal contributo di un e^- a quello degli e^- compresi tra b e $b+db$ in una corona cilindrica circolare di raggio b , e poi integrare su tutto b per tener conto di tutte le interazioni (vedi Fig. 3).

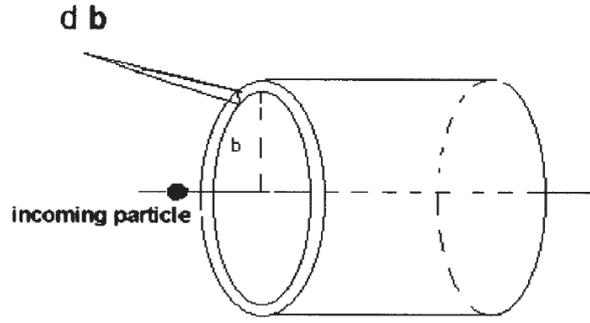


Figura 3: La particella carica passa con direzione rettilinea a distanza b dall'elettrone atomico che si considera a riposo rispetto alla velocità della particella.

Se n è la densità di e-, allora l'energia ceduta a tutti gli elettroni contenuti nel volume $dV = (2\pi b) db dx$ della corona cilindrica lungo un percorso dx è:

$$-dE(b) = \Delta E(b) n dV = \frac{2z^2 e^4}{mb^2 v^2} n (2\pi b) db dx = \frac{4\pi n z^2 e^4}{mv^2} \frac{db}{b} dx$$

Dal punto di vista della particella, l'energia persa per unità di percorso dx nell'intervallo di parametro di impatto considerato è dunque:

$$-\frac{dE}{dx} = \left(\frac{4\pi n z^2 e^4}{mv^2} \right) \frac{db}{b}$$

Se ora voglio integrare su tutto b , dovrò porre dei limiti di integrazione (altrimenti l'integrale diverge!): b è limitato dal fatto che il tempo di interazione non può essere molto più lungo del periodo τ tipico degli elettroni legati all'atomo, altrimenti non potrebbe influenzare l'energia dell'elettrone che avrebbe tempo di spostarsi.

$$b_{max} = \tau \gamma v = \frac{\gamma v}{\bar{\nu}} = \gamma \beta ch / I$$

dove abbiamo introdotto l'energia di eccitazione media $I = h\bar{\nu}$.

Per bassi valori il b è limitato dal fatto che l'approccio classico vale solo per distanze maggiori della distanza di de Broglie (non stiamo infatti considerando

la natura ondulatoria delle particelle):

$$b_{min} = \lambda/2 = h/2p_{cdm} = h/2m\gamma v = h/2\beta\gamma mc$$

Se integriamo tra b_{min} e b_{max} otteniamo:

$$-\frac{dE}{dx} = \frac{4\pi n z^2 e^4}{mv^2} \ln \left(\frac{b_{max}}{b_{min}} \right) = \frac{4\pi n z^2 e^4}{mv^2} \ln \left(\frac{v\gamma h 2m\gamma\beta c}{I h} \right) = \frac{4\pi n z^2 e^4}{mv^2} \ln \left(\frac{2m\gamma^2 v^2}{I} \right)^2$$

Se sostituiamo il valore della energia massima trasferita W_m , otteniamo l'espressione classica di Bohr che descrive ragionevolmente la perdita di energia media di una particella carica nel materiale per particelle pesanti come una particella α o nuclei piu' pesanti, ma non e' accurata per esempio per il protone e le particelle piu' leggere, visto che gli effetti quantistici diventano rilevanti.

Infine, usando il valore dell'energia massima trasferita $W_m =$ otteniamo:

$$-\frac{dE}{dx} = \frac{4\pi n z^2 e^4}{mv^2} \ln \left(\frac{2mv^2 W_m}{I^2(1 - \beta^2)} \right)$$

Da un trattamento quantistico dell'energia persa per collisione da una particella incidente pesante e a spin), si ottiene un termine aggiuntivo $-\beta^2$ che compare nella formula relativistica di Bethe:

$$-\frac{dE}{dx} = \frac{4\pi n z^2 e^4}{mv^2} \left(\ln \left[\frac{2mv^2 W_m}{I^2(1 - \beta^2)} \right] - \beta^2 \right) s$$

Posso esplicitare il raggio classico dell'elettrone $r_e = e^2/mc^2$ e esprimere la densita' come $n = Z\rho N_A/A$, con N_A numero di Avogadro, Z e A rispettivamente numero e peso atomico.

Con queste opportune sostituzioni, e aggiungendo due correzioni per l'effetto densita' ($-\delta/2$) e di Shell ($-2 C/Z$) si puo' riscrivere la formula come:

$$-\frac{dE}{dx} = 2\pi N_A m r_e^2 c^2 \rho \frac{Z}{A} \frac{z^2}{\beta^2} \left[\ln \left[\frac{2mv^2 W_m}{I^2(1 - \beta^2)} \right] - 2\beta^2 - \delta - 2\frac{C}{Z} \right] = \quad (1)$$

Sostituendo i valori delle costanti, la formula si puo' riscrivere come:

$$-\frac{dE}{dx} = 0.1535 \rho \frac{Z}{A} \frac{z^2}{\beta^2} \left[\ln \left[\frac{2mv^2 W_m}{I^2(1 - \beta^2)} \right] - 2\beta^2 - \delta - 2\frac{C}{Z} \right] [MeV/cm] \quad (2)$$

Grandezza	Significato
r_e	raggio classico dell'elettrone
N_A	numero di Avogadro
m	massa dell'elettrone
I	Potenziale di ionizzazione medio
Z, A	numero e peso atomici
ρ	densità del materiale
ze	carica della particella incidente
$\beta = v/c$	velocità della particella incidente
$\gamma = 1/\sqrt{1 - \beta^2}$	il fattore Lorentziano
δ	termine di correzione (effetto densità)
C	termine di correzione (Shell)
W_{max}	Massima energia cinetica cedibile per singolo urto

Se dividiamo per la densità ρ ed esprimiamo il potere di frenamento in unità di $MeV cm^2/g$, otteniamo il grafico per muoni positivi nel rame mostrato in fig. 4.

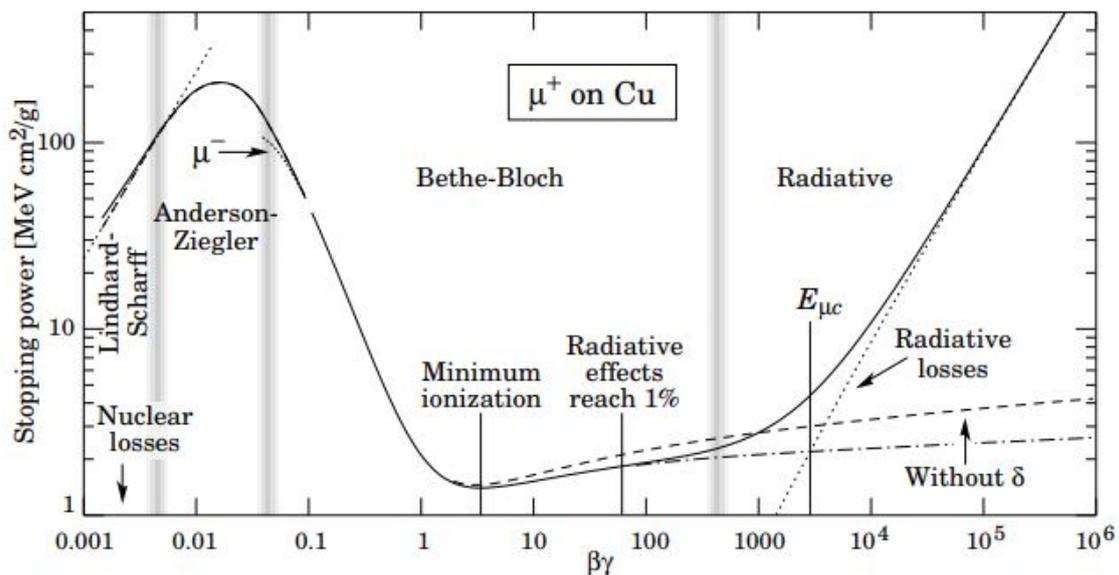


Figura 4: Perdita di energia per unità di percorso di muoni attraverso rame

La parte in parentesi quadre varia lentamente con l'energia della particella. Il fattore dominante è quindi quello moltiplicativo che è proporzionale a $1/\beta^2$. Generalmente, in regime non relativistico, la funzione decresce come $1/\beta^2$, ovvero è inversamente proporzionale all'energia. Qualitativamente, questo si deduce dal fatto che la particella lenta passa molto tempo nel cam-

po degli elettroni orbitanti. L'impulso e quindi l'energia trasferita percepiti sono quindi piu' elevati.

La funzione raggiunge **un minimo per** $\beta\gamma \sim 3 - 4$, ovvero $\beta \sim 0.96$. In questo punto la particella viene detta "minimo ionizzante" o "MIP". Per velocita' piu' alte, in teoria prevale il termine logaritmico (si parla di risalita relativistica) in quanto a regimi relativistici il campo elettrico si appiattisce e si estende nella direzione perpendicolare alla traiettoria, coinvolgendo piu' elettroni nell'interazione. Tuttavia gli atomi del materiale posizionati tra la particella e gli elettroni lontani sono a questo punto polarizzati, schermando in tal modo il contributo degli e^- piu' lontani. Per una particella relativistica dunque la risalita è molto lenta, proporzionale a $\ln(\beta\gamma)$. Tale correzione portata dal termine δ nella formula è anche detta correzione di densita', ed è ovviamente piu' accentuata nel caso di materiali piu' densi (solidi e liquidi) rispetto ai meno densi (gassosi), come visibile in fig. 5. Nella stessa figura è visibile come il rapporto Z/A influisce sulla traslazione delle curve su y . L'andamento vicino a quello di una costante, effetto della correzione di identita', compare per $\beta\gamma \sim 10$ (ricordiamo che in regime relativistico $\beta\gamma \sim \gamma$, in regime non relativistico è semplicemente $\beta\gamma \sim \beta$) 2 prende il nome di Plateau di Fermi, che studio' i contributi dagli elettroni lontani. Sul plateau di Fermi, $(dE/dx)_{plateau} \sim 1.05 - 1.1 (dE/dx)_{MIP}$ per i solidi, e $(dE/dx)_{plateau} \sim 1.5 (dE/dx)_{MIP}$ per i gas.

Notiamo che **non c'e' dipendenza diretta dalla massa**, quindi se dovessimo confrontare particelle diverse l'unica variabile rilevante è la carica z che compare a numeratore, elevata alla seconda potenza: una particella α per esempio avra' una perdita ~ 4 volte maggiore di una particella con carica unitaria.

La dimensionalità (espressa in unità di misura) della formula di Bethe-Bloch è

$$\left[-\frac{dE}{dx} \right] = \left[\frac{MeV}{cm} \right] \quad (3)$$

tuttavia è di gran lunga più utile riferirsi ad una lunghezza che non dipenda dalla densità del materiale, ossia lo *spessore massico* così definito:

$$\chi = \rho x \quad (4)$$

la cui dimensionalità (espressa in unità di misura) è:

$$\left[\chi \right] = \left[\frac{g}{cm^2} \right] \quad (5)$$

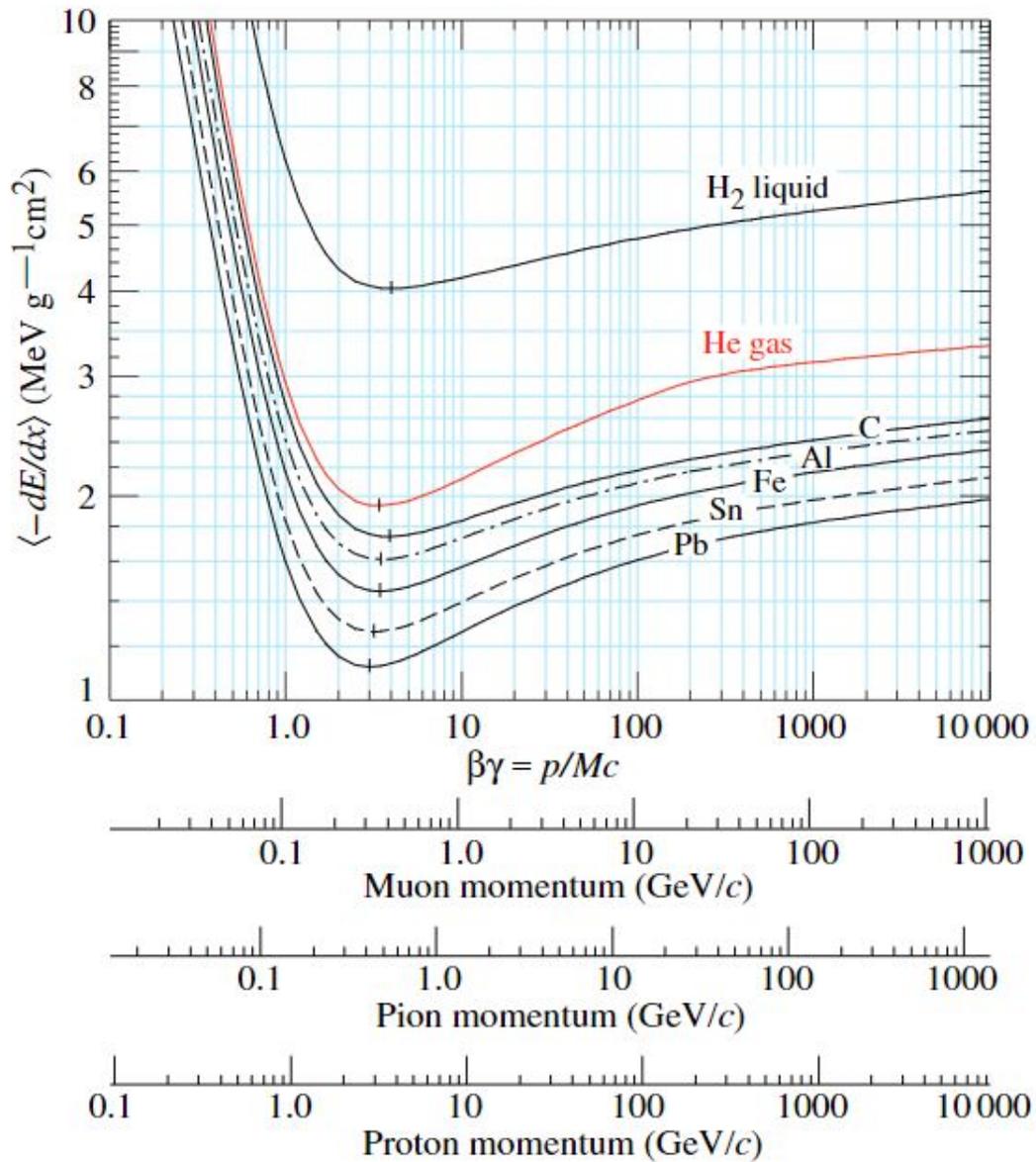


Figura 5: Perdita di energia per unita' di percorso attraverso diversi materiali

Esempio: Dato uno spessore $x = 1 \text{ mm}$ di silicio (Si) la cui densità vale $\rho = 2.33 \frac{\text{g}}{\text{cm}^3}$ lo spessore massico è dato dall'equazione 4, nel presente esempio si ha $\chi = 0.233 \frac{\text{g}}{\text{cm}^2}$. In pratica, invece di valutare la perdita di energia in un mm di silicio, si esprime la perdita di energia nel passaggio attraverso una certa quantità di materia, qui espressa come densità superficiale. In questo modo ci si rende indipendenti dalla densità dello specifico materiale. Quindi se per esempio ho $-dE/dx = 1 \text{ MeV cm}^2/\text{g}$, nel mio sensore avrò 0.233 MeV .

Avendo espresso dE/dx in $MeV\ cm^2/g$, la dipendenza dal materiale rientra solo grazie a I , l'energia di eccitazione media degli atomi del materiale, e ovviamente grazie a il numero e peso atomici del materiale. è tuttavia una dipendenza debole. Notiamo che in fig. 6 sull'asse x sono riportate particelle diverse e la scala si riferisce al momento delle particelle, una quantità dipendente dalla massa oltre che dalla velocità'. Si nota subito come la funzione risulti solo traslata sull'asse x per diverse particelle.

In fig. 6 si vede la curva di B-B per diverse particelle che attraversano uno spessore di aria. Un'ulteriore **correzione** e' quella delle basse energie,

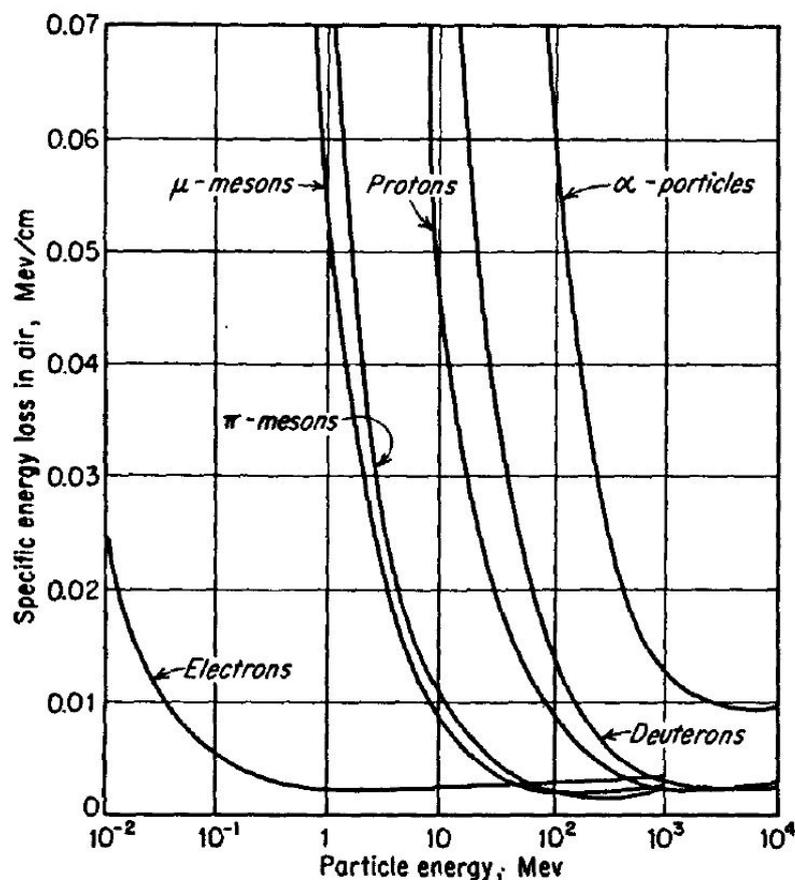


Figura 6: Perdita di energia di diverse specie di particelle cariche per unita' di percorso in aria

detta *di shell*: per protoni da 1 a 100 MeV, la velocità della particella interagente non è molto più veloce di quella degli elettroni atomici, quindi tale assunzione, che l'interazione coulombiana avvenga come se l'elettrone fosse fermo, cade, e il contributo proveniente da questi e- sarà minore. Sotto il

picco osservato a $\beta\gamma \sim 0.01$ l'interazione viene a mancare con un andamento in funzione di $1/Z$: piu' cariche ci sono, piu' piccole sono le correzioni.

1.2.1 Perdita di energia nei materiali sottili

La perdita di energia non è un meccanismo deterministico per la singola interazione, tuttavia puo' essere descritto in maniera statistica per il gran numero di interazioni a cui va incontro la particella che attraversa il materiale, come visto nella sezione precedente. L'energia media persa per unita' di percorso al passaggio nel materiale è descritta dalla Bethe-Bloch. Tale statistica è affetta da **grandi fluttuazioni**, sia nel numero di interazioni sia nell'energia persa per ogni interazione. Il numero di interazioni al variare del materiale attraversato segue una distribuzione poissoniana, quindi asimmetrica. Inoltre la particella puo' perdere piu' o meno energia a seconda del parametro di impatto rispetto all' e^- interagente. La maggior parte delle interazioni comporta una piccola perdita di energia per interazione, vicina al limite minimo. Eventi con interazione *head-on* (basso parametro di impatto) hanno statistica rara ma contribuiscono con una grande perdita di energia, vicina all'energia massima trasferibile. In condizioni di bassa statistica, ovvero quando ci sono poche interazioni, per esempio perche' le particelle sono molto energetiche o se gli spessori sono molto sottili, tali eventi hanno un peso maggiore, non vengono compensati dalle interazioni a bassa energia trasferita, e contribuiscono quindi a rendere asimmetrica la distribuzione statistica con coda alta verso l'alta energia.

La convoluzione della distribuzione probabilistica del numero di interazioni in funzione dello spessore e della distribuzione dell'energia persa per interazione porta alla distribuzione di probabilita' di energia totale persa da una particella carica in un dato spessore di materiale porta a una distribuzione molto diversa da quella Gaussiana. In effetti è Gaussiana solo vicino intorno al picco, ma poi la larghezza è ampia e asimmetrica e non cade come $1/\sqrt{X}$. Cio' rende l'energia persa media inutile per misurare l'energia depositata nel mezzo, fluttuera' in maniera non gaussiana. Il valore a massima probabilita' invece non fluttua e viene usato per misurare l'energia rilasciata. Bisogna tenere a mente che media e MPV non coincidono.

Un'ulteriore fattore che porta alla saturazione della **misura** dell'energia persa è la dimensione finita dei rivelatori usati: tale fattore limita la massima energia trasferibile, ragion per cui si parla di *restricted energy loss* ove

nella Bethe-Bloch compare una costante W_{cut} al posto di W_{max} dipendente da β , e che scala il termine in β^2 di un fattore $(1+W_{cut}/W_{max})$. Tale restrizione porta la curva di Bethe-Bloch ad avvicinarsi a un valore costante abbassando ulteriormente il Plateau di Fermi che avevamo affrontato parlando della correzione di densita'. Questo effetto è visibile nel confronto mostrato in fig. 7.

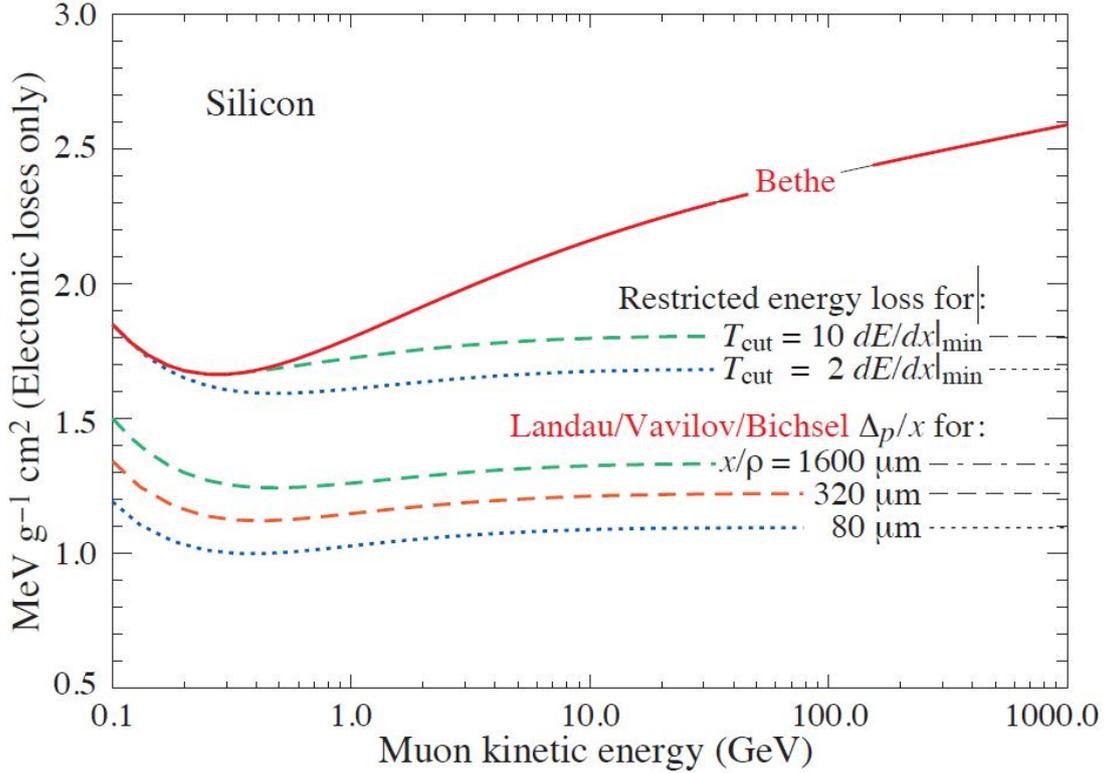


Figura 7: Bethe-Bloch completa confrontata con la versione ristretta. Altre rappresentazioni specifiche (Landau/Vavilov/Bichsel) riguardanti il valore più probabile di energia persa in materiali sottili visualizzati nel grafico sono trattati in sez. 1.2.1.

La distribuzione dell'energia persa negli spessori sottili è descritta dalla funzione di Landau (6), dal nome di chi la descrisse per primo:

$$L(\lambda) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(\lambda + e^{-\lambda})\right] \quad (6)$$

ed è riportata in grafico in figura 8.

Il parametro λ rappresenta la deviazione rispetto al valore più probabile di energia trasferita, scalato per un fattore parametrico che dipende da densita',

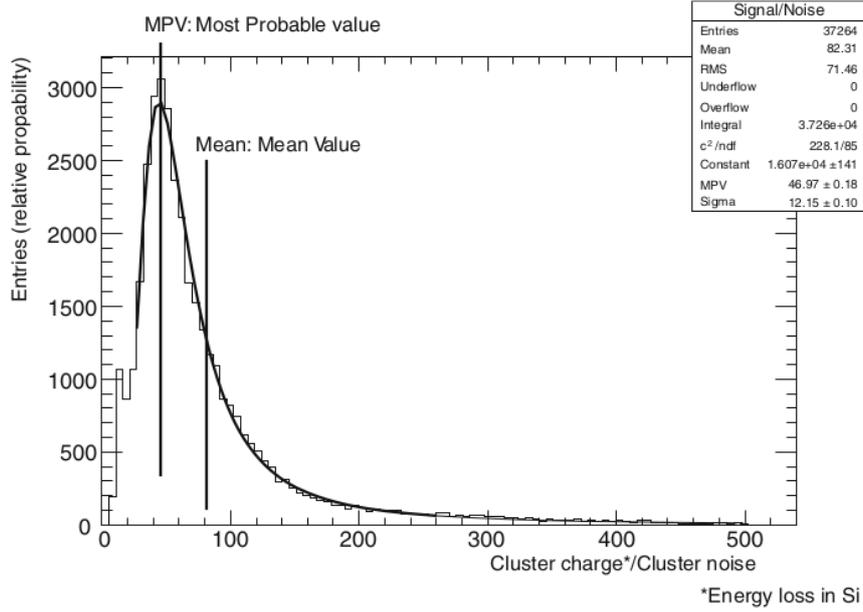


Figura 8: Grafico della distribuzione di Landau per raggi cosmici su silicio

lunghezza e da una costante caratteristica del materiale:

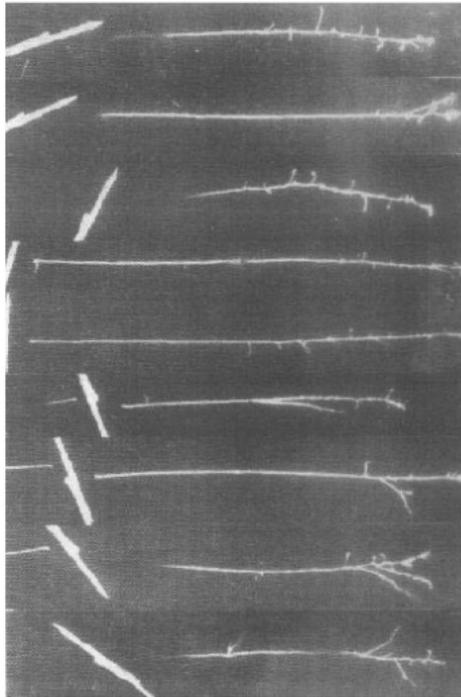
$$\lambda = \frac{\Delta E - \Delta E^w}{\xi} \quad (7)$$

dove ΔE è l'energia realmente persa nel materiale di spessore x , ΔE^w è l'energia piu' probabile persa nel materiale sottile e ξ è il parametro dipendente dalle caratteristiche del mezzo

$$\xi = \rho x k \quad (8)$$

dove ρ è la densità, x la distanza percorsa nel mezzo e k un parametro caratteristico.

Una particella che passa attraverso un materiale è interessata anche da eventi con un grande trasferimento di impulso e conseguente grande perdita di energia che, seppur rari, contribuiscono a modificare la distribuzione alzando la coda della distribuzione ad alte energie. Si suppone comunque che l'energia sia piccola rispetto all'energia della particella, cosicche' quest'ultima si puo' considerare costante: l'evento successivo avra' la particella nelle stesse condizioni cinematiche, e il processo sara' nuovamente soggetto alle stesse fluttuazioni nella perdita di energia.



From: <http://www.lateralscience.co.uk/cloud/diff.html>

Figura 9: I δ -rays si dipartono come rami laterali dalla ionizzazione principale.

In questi eventi ad alto trasferimento di energia vengono creati $\delta - rays$ che possono essere rilevati (Fig. 9).

La presenza di questi eventi infatti allunga la coda della distribuzione ad alte energie, rendendola asimmetrica, e spostando la media a valori piu' alti dal valore di picco, ovvero il valore piu' probabile (o *Maximum Probable Value* - *MPV*). In tali condizioni dunque, la descrizione di Bethe Bloch (perdita di energia media) sovrastima la perdita di energia piu' probabile.

Per spessori sottili la perdita media di energia è molto minore della massima energia trasferibile, vuoi perche' il materiale è sottile o perche' la particella è minimo ionizzante: $\langle \Delta E \rangle \ll E_{max}$. Si puo' trattare la distribuzione di probabilita' di perdita di energia con tre modelli diversi, a seconda del regime individuato dal rapporto tra energia persa in media attraverso uno spessore di materiale ΔE_{BB} e l'energia massima trasferibile per una singola interazione con un elettrone atomico ΔE_{Max} , che si indica con:

$$k = \frac{\Delta E_{BB}}{\Delta E_{Max}} \quad (9)$$

- $k \leq 0.01$ si usa la distribuzione di Landau. In questa situazione la perdita di energia è piccola rispetto all'energia massima trasferibile in una collisione (che si assume tendente all'infinito); la perdita di energia è grande se paragonata all'energia di legame degli elettroni, si trascurano quindi le perdite piccole di energia dovute alle collisioni lontane e c'è dunque un limite minimo sull'energia trasferibile nel singolo urto
- $k > 0.01$ Nel caso di perdite medie di energia paragonabili all'energia di legame (per esempio con silicio o TPC), non vale piu' l'assunzione di ΔE_{Max} che tende all'infinito e si usa la distribuzione di Vavilov. La distribuzione ha in questo caso code un po' meno asimmetriche il parametro λ della Landau viene modificato.
- $k \geq 10$ è valida la distribuzione di Gauss: l'energia persa in una collisione è una variabile casuale, le collisioni statisticamente indipendenti, l'energia totale persa è la somma di tutte le singole energie perse con la stessa distribuzione, ed è quindi gaussiana per il limite centrale.

Le distribuzioni misurate su spessori via via piu' sottili sono visibili in fig. 10. Nelle misure si può vedere come il MPV di energia persa per unita'

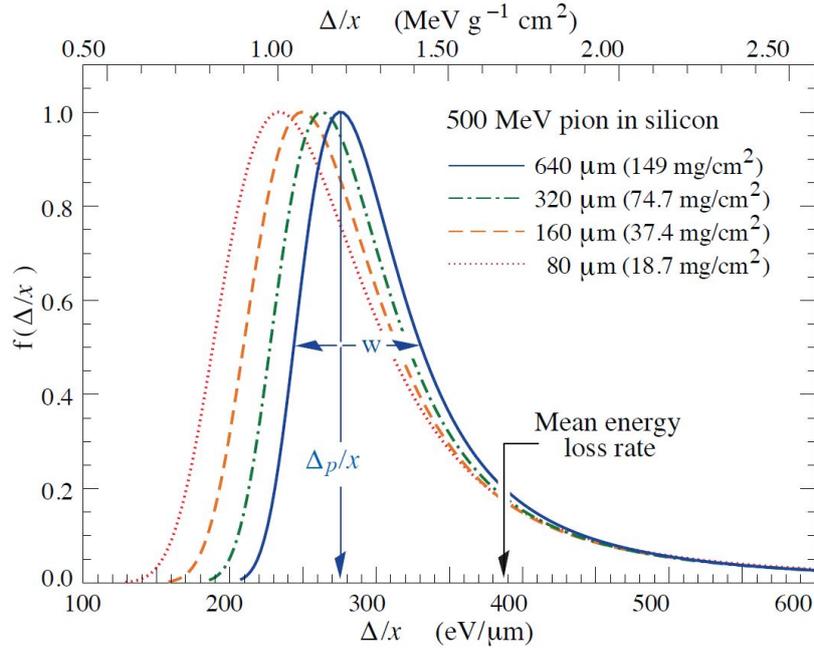


Figura 10: Perdita di energia di pioni da 500 MeV in 80-160-320-640 μm di silicio: distribuzione della perdita di energia per micron, normalizzate sul picco del MPV. Il tasso di perdita media di energia è costante.

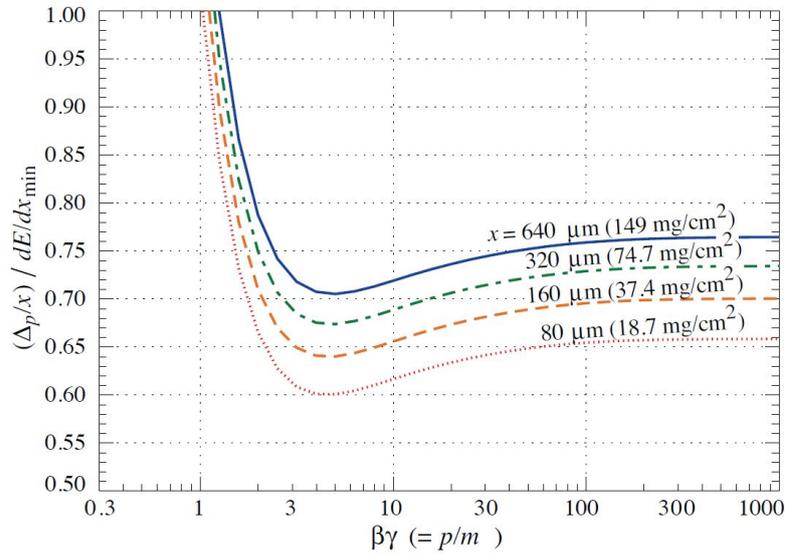


Figura 11: MPV per l'energia persa da pioni da 500 MeV in 80-160-320-640 μm di silicio, riscalata sulla perdita di energia media di una particella minimo ionizzante ($388 \text{ eV}/\mu\text{m}$)

di percorso per diversi spessori di materiali cresce al crescere dello spessore, con una media che rimane più alta.

Come si vede, nonostante la media sia fissa, il valore più probabile di tasso di perdita di energia diminuisce al diminuire dello spessore attraversato.

Visto che il picco MPV è fisso mentre la media è soggetta a fluttuazioni, si tende a usare metodi del tipo l'eliminazione degli eventi con il 30% più alto della distribuzione, per poi fittare gli altri con una gaussiana, o una media troncata con eliminazione del valore più alto, nel caso di pochissime misure.

La perdita di energia in **materiali compositi** o leghe si tratta come una sommatoria di perdite di energia in strati sottili di elementi puri diversi nella giusta proporzione:

$$\left\langle \frac{dE}{dx} \right\rangle = \sum w_j \left\langle \frac{dE}{dx} \right\rangle_j$$

1.3 Lunghezza di radiazione

La lunghezza di radiazione X_0 è un modo di rappresentare lo spessore di un materiale a seconda di come viene visto dalla particella che lo attraversa in termini di densità di interazioni. Equivale alla distanza su cui un elettrone riduce la sua energia iniziale E_0 di un fattore $1/e$ solo per emissione di radiazione di Bremsstrahlung, che per gli elettroni è quella che domina a energie medio-alte: $E(X_0) = \frac{E_0}{e}$. È il termine che compare nel numeratore dell'esponente della funzione di attenuazione dell'energia per perdita di radiazione

$$E(x) = E_0 e^{-\frac{x}{X_0}}$$

Da notare che questa non è un'attenuazione in intensità (ovvero in numero di fotoni contati dall'altro lato di uno strato di materiale su cui si è puntato un fascio di fotoni) ma in energia delle particelle. In letteratura, X_0 è riportata sempre per gli elettroni, ed è una caratteristica del materiale attraversato. In prima approssimazione, la lunghezza di radiazione si può calcolare come:

$$X_0 = \frac{716.4 \text{ g/cm}^2 A}{Z(Z+1) \ln(287/\sqrt{Z})}$$

con Z e A rispettivamente numero atomico e peso atomico del materiale attraversato.

Lunghezze di radiazione tipiche per alcuni materiali sono riportate in Tab. 1:

materiale	X_0 [cm]
acqua	36.0
aria	30 000
H	731 000
Si	9.4
Cu	1.43
Al	8.9

Tabella 1: Lunghezza di radiazione espressa in *cm*.

Cio' vuol dire che ci vuole uno spessore di acqua 4 volte piu' grande di quello di silicio perche' la particella perda la stessa energia, ovvero che introdurre dell'alluminio sulla traiettoria di una particella carica ha un effetto di attenuazione tre volte minore del rame.

Se andiamo a esprimere la X_0 in termini di densita' superficiale [g/cm^2], possiamo confrontare i contributi di vari elementi in maniera uniforme. Per esempio se io voglio introdurre un massimo di una lunghezza di radiazione sulla traiettoria di una particella, posso semplicemente sommare in contributi in tale unita', che rappresenteranno poi spessori fisici diversi.

1.4 Scattering Coulombiano multiplo

Una particella che attraversa la materia sara' dispersa da potenziali Coulombiani di nuclei ed elettroni incontrati, attraverso le reazioni di:

- Ionizzazione: perdita di energia per collisione con gli elettroni atomici
- Scattering Coulombiano: deviazione per interazione coi campi coulombiani dei nuclei

La deviazione totale in uscita dal materiale rispetto alla direzione di entrata della particella è il risultato di un grande numero di interazioni casuali con i campi Coulombiani dei nuclei incontrati, caratterizzati ognuno da una piccola deviazione di traiettoria. Per il Teorema del Limite Centrale, l'angolo di deviazione segue una statistica Gaussiana a piccoli angoli, ma per angoli piu' grandi segue invece la Teoria di Moliere, caratterizzata da un eccesso di eventi ad angolo piu' grande rispetto a una distribuzione normale, che si

Material	Z	A	X_0 [g/cm ²]	X_0 [cm]
Hydrogen	1	1.01	61.3	731 000
Helium	2	4.00	94	530 000
Lithium	3	6.94	83	156
Carbon	6	12.01	43	18.8
Nitrogen	7	14.01	38	30 500
Oxygen	8	16.00	34	24 000
Aluminium	13	26.98	24	8.9
Silicon	14	28.09	22	9.4
Iron	26	55.85	13.9	1.76
Copper	29	63.55	12.9	1.43
Silver	47	109.9	9.3	0.89
Tungsten	74	183.9	6.8	0.35
Lead	82	207.2	6.4	0.56
Air	7.3	14.4	37	30 000
SiO ₂	11.2	21.7	27	12
Water	7.5	14.2	36	36

Figura 12: Numero di carica, numero di massa atomica e lunghezza di radiazione espressa in g/cm^2 e cm per diversi materiali.

riflette in code piu' alte. Se Θ è l'angolo di uscita, la distribuzione è centrata in $\Theta = 0$ e ha una larghezza che si puo' esprimere come scarto quadratico medio dell'angolo di diffusione, in funzione del momento della particella e del materiale attraversato.

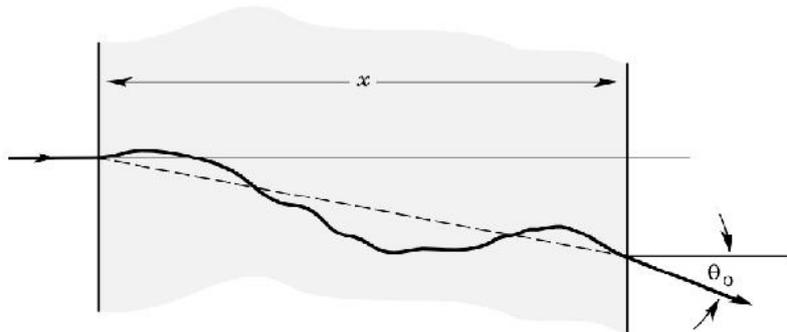


Figura 13: Schema dello scattering MCS.

In particolare, invece di considerare l'angolo di deviazione totale Θ (che sarebbe l'angolo di deviazione misurato sul piano immaginario contenente la direzione di entrata e quella di uscita), spesso conviene considerare la sua proiezione sui piani formati rispettivamente da x e y con la direzione iniziale della traiettoria della particella, dove x e y sono perpendicolari alla traiettoria

iniziale, visto che i rivelatori offrono normalmente una visione bi-dimensionale sullo schermo (x,y). Le due distribuzioni angolari saranno identiche. Lo scarto quadratico medio della proiezione dell'angolo di diffusione è:

$$\Theta_{rms}^{proj} = \sqrt{\langle \Theta^2 \rangle} = \frac{13.6 MeV}{\beta c p} \cdot z \sqrt{\frac{x}{X_0}} [1 + 0.038 \ln(x/X_0)]$$

ove p è il momento espresso in MeV/c, βc la velocità e z la carica della particella diffusa. x/X_0 è lo spessore del materiale attraversato in unità di lunghezze di radiazione. z è normalmente 1, e per piccoli spessori il termine logaritmico si può approssimare a 0.

Considerando invece l'angolo di deviazione totale invece, si può dimostrare che per piccole deflessioni, allora

$$\Theta_{rms}^{proj} = \frac{1}{\sqrt{2}} \Theta_{rms}^{space}$$

Quindi nello spazio tridimensionale, la formula approssimata per la deviazione angolare viene moltiplicata per fattore $\sqrt{2}$ e diventa:

$$\Theta_{rms}^{space} = \frac{19.2 MeV}{\beta c p} \sqrt{\frac{x}{X_0}}$$

Notiamo che la formula funziona se si esprime il momento in MeV/c e lo spessore in unità di lunghezze di radiazione, quindi con x e X_0 o nell'unità appropriata.

Excursus misura: Una facile maniera per determinare la curva della distribuzione è far passare dei protoni con una certa energia fissata attraverso uno strato di materiale e misurarne l'angolo proiettato sul piano di misura, contando i protoni che giungono sullo schermo di misura, assumendo che abbia una risoluzione spaziale adeguata. Se per angoli piccoli la distribuzione scende come $1/\Theta^2$, a grandi angoli scende come $1/\Theta^4$ (fig. 14).

Nel caso si abbiano strati di materiali diversi, un approccio è quello di sommare in quadratura i contributi di ogni strato in termini di scarto angolare. Tuttavia questo approccio sottostima costantemente la deviazione. Risulta dunque più accurato calcolare la lunghezza di radiazione e lo spessore attraversato per il composto attraversato, e applicare una sola volta la formula per lo scarto angolare.

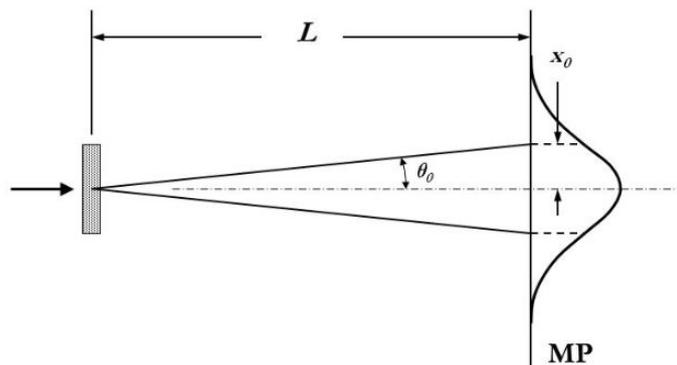


Figura 14: Schema della misura di MCS.

1.5 Formazione del segnale

I meccanismi principali di interazione tra particella carica e materia ci indicano come la dipendenza dalla densita' di materiale sia importante sia per il segnale generato sia per la deviazione della traiettoria al passaggio della particella nel materiale. La perdita di energia dipende infatti dalla concentrazione di e^- nel materiale, ovvero da

$$n = \frac{N_{e^-}}{Vol} \propto \frac{Z\rho}{A}$$

- L'obiettivo principale della progettazione e costruzione di un buon rivelatore consiste nel massimizzare il **rapporto segnale-rumore**, quindi c'e' la necessita' aumentare l'ampiezza del segnale, che dipende dalla minima energia necessaria per generare una unita' di segnale, e dal numero di interazioni in cui la particella produce il segnale nel **volume sensibile**.
- Si vuole pero' perturbare la traiettoria della particella il meno possibile, limitando lo scattering coulombiano multiplo attraverso la riduzione degli spessori attraversati. Prima di tutto i **volumi dei componenti accessori** (come i servizi e le infrastrutture), ma anche i volumi sensibili se si riesce a non sopprimere il segnale.
- Aggiungiamo anche il fattore dimensionale: la bassa densita' puo' implicare la necessita' di grandi volumi di materiale da attraversare per creare un segnale abbastanza ampio, e quindi rivelatori di **grandi dimensioni**

Si tratta quindi di trovare un compromesso tra efficienza di produzione del segnale e spessore effettivo del materiale sensibile. Possiamo qualificare i diversi tipi di materiale che costituiscono i volumi sensibili dei rivelatori in base a tre caratteristiche: la quantità di energia necessaria per produrre il segnale minimo (*quanto* di segnale), il tipo di segnale prodotto, la densità del materiale. Riassumiamo i tipi di materiale che compongono i rivelatori in Tab. 2: dalle loro caratteristiche possiamo intuire i pro e i contro di ognuna.

tipo	quanto di segnale	tipo di segnale	densità
gas	~ 30 eV	coppia $e^- - ione+$	bassa
semiconduttori (Si)	~ 3.6 eV	coppia $e^- - h^+$	alta
scintillatori	~ 100 eV	fotone	media

Tabella 2: Caratteristiche dei materiali rivelatori.

Si nota che i semiconduttori hanno il pregio di aver bisogno di solo 3.6 eV per poter produrre un quanto di segnale, e hanno una densità così alta da poter massimizzare la quantità di interazioni per unità di volume, e quindi limitare le dimensioni del rivelatore. Inoltre il segnale prodotto consiste in particelle cariche, alcune di esse libere di muoversi e trasportare l'informazione.

Cerchiamo di capire il processo che porta alla formazione del quanto di segnale nei semiconduttori.

2 Fisica dei materiali semiconduttori

I materiali si possono caratterizzare in base alle diverse proprietà elettriche. In particolare si può misurare l'intensità della corrente I che fluisce attraverso una barra di materiale omogeneo, di sezione uniforme e lunghezza nota, i cui estremi sono sottoposti alla differenza di potenziale elettrico V e dedurre la resistenza $R = V/I$, come descritto dalla Legge di Ohm. È facile verificare sperimentalmente che R varia in funzione dell'area A della sezione e della lunghezza L della barra secondo un parametro ρ chiamato *resistività*:

$$\rho = R \frac{A}{L} \quad (10)$$

La resistività nei materiali usati in campo elettronico e dei rivelatori varia di molti ordini di grandezza: si va da $\rho = 10^{-6} \Omega cm$ dell'alluminio usato nelle connessioni dei circuiti elettrici, al biossido di silicio SiO_2 usato per isolare e proteggere le parti superficiali e più esposte dei sensori, in cui $\rho = 10^{16} \Omega cm$, e alle plastiche $\rho = 10^{18} \Omega cm$.

In base alla resistività si possono quindi definire delle classi di materiali:

- $\rho \lesssim 10^{-2} \Omega cm$: conduttori
- $10^{-2} \Omega cm < \rho < 10^5 \Omega cm$: semiconduttori
- $\rho > 10^5 \Omega cm$: isolanti

La classe intermedia riguarda i semiconduttori, in cui la resistività può essere modificata in maniera controllata e con grande precisione la conduzione elettrica può essere a due tipi di portatori.

Si possono dedurre le caratteristiche elettriche del semiconduttore a partire dalla descrizione delle proprietà del sistema di elettroni in un atomo isolato e perturbando il sistema elettronico del solido per ottenere informazioni sulle caratteristiche del materiale.

2.1 Gli elettroni nei nuclei

In un atomo con numero atomico Z , per un e- convivono due tipi di campo di forze: quella attrattiva con i protoni del nucleo e quella repulsiva con l'insieme degli elettroni dell'atomo. La prima è radiale, diretta verso il nucleo, e ha potenziale

$$U(r) = \frac{Zq^2}{4\pi\epsilon_0 r}$$

dove r e' la distanza e- - nucleo, e dove l'en potenziale di riferimento e' presa all'infinito, come da convenzione. La seconda invece e' nulla per $Z=1$, ovviamente, e trascurabile per atomi con basso peso atomico. Nell'interpretazione classica, gli e- dell'atomo si muovono intorno al nucleo con orbite ellittiche con un fuoco nel nucleo. Questa interpretazione non e' accurata in quanto prevede:

- gli e- si muovono di moto accelerato, quindi dovrebbero irradiare energia, perdendola, e quindi diminuire la loro distanza dal nucleo fino a fermarsi nel nucleo stesso. Non sarebbero dunque stabili.
- l'energia degli e- dovrebbe essere continua, mentre sappiamo dagli spettri misurati sappiamo che e' quantizzata, ovvero si manifesta solo in determinate energie. Gli e- eccitati termicamente infatti si diseccitano emettendo radiazione elettromagnetica che puo' essere rivelata: i fotoni hanno frequenza tale che $hf = \Delta E$, e l'energia ΔE misurata non rappresenta uno spettro continuo ma dei valori determinati, ovvero delle *righe spettrali*, che corrispondono all'energia che distanza i livelli energetici su cui si situano gli elettroni.

La teoria quantistica dell'atomo idrogenoide parte dalla soluzione dell'eq di Shoedinger in presenza di un campo di forze coulombiano. Si ottengono, come per il caso monodimensionale, soluzioni di tipo legato, con proprieta' di quantizzazione. Le 3 dimensioni aggiungono piu' numeri quantici (m, n, l) piu' in numero di spin, che permette la compresenza di solo 2 e- sullo stesso livello energetico per il principio di Pauli.

2.2 Teoria delle bande energetiche

A partire dal modello dell'atomo di Bohr, conosciamo come interagisce un elettrone con il potenziale coulombiano di un nucleo atomico: esiste uno spettro di valori consentiti di energia che l'e- puo' assumere, corrispondente a

$$E_n = \frac{-Z^2 m_0 q^4}{8\epsilon_0^2 h^2 n^2} \quad (11)$$

al di sotto un'energia di riferimento assunta pari a 0. Nella formula: Z e' la carica netta del nucleo, q e' la carica elementare, m_0 e' la massa in quiete dell'elettrone, ϵ_0 e' la costante dielettrica nel vuoto, h la costante di Plank, e n un numero intero positivo. Per l'idrogeno, con $Z = 1$ i valori consentiti

sono $-13.6/n^2 eV$ al di sotto del livello di riferimento, con n intero, quindi: $-13.6 eV$, $-3.4 eV$, $-1.5 eV$, $-0.85 eV$, ecc ecc. Allo zero assoluto, il livelli a energia piu' bassi sono i primi ad essere occupati, con al massimo 2 e- per livello, seguendo il principio di esclusione di Pauli.

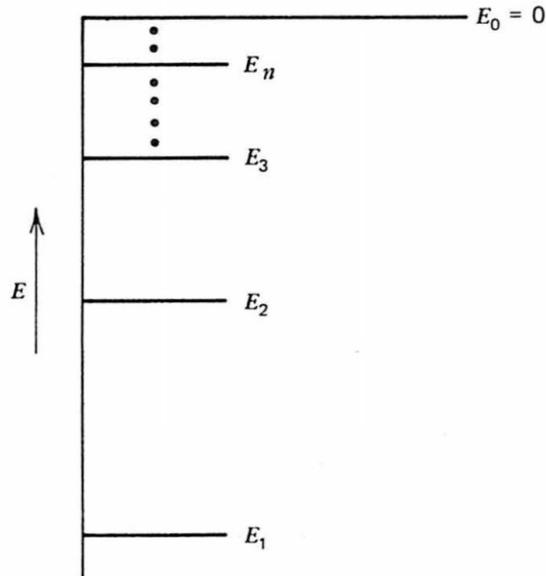


Figura 15: Livelli energetici dell'atomo idrogenoide.

Il cristallo regolare composto da atomi si puo' modellizzare come un profilo periodico di energia potenziale, costituito da una successione di buche di potenziale di altezza finita. Il moto di una particella in tale potenziale si puo' studiare risolvendo l'eq di Schrodinger.

Qui ci limitiamo a dedurre alcune conclusioni da discussioni qualitative. La modellizzazione come singola buca di potenziale offre quindi soluzioni

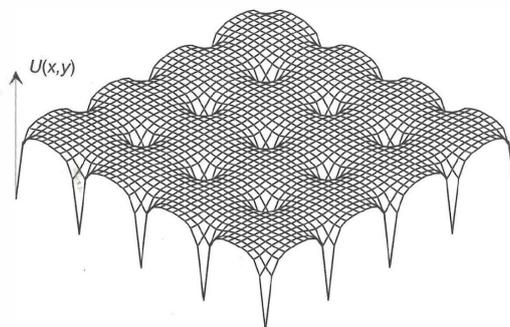


Figura 16: Enter Caption

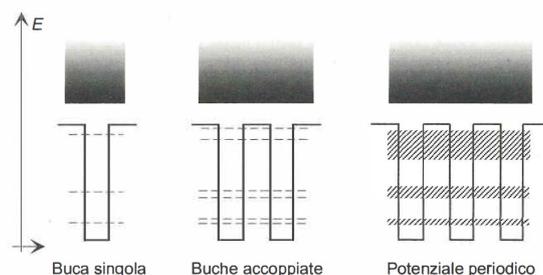


Figura 17: Enter Caption

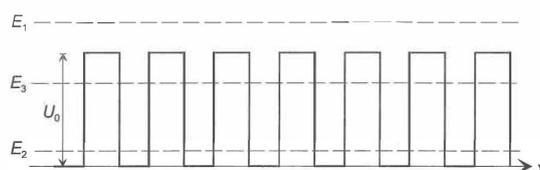


Figura 18: Enter Caption

quantizzate (particelle legate) e una banda di soluzioni continue (particelle libere). A causa dell'effetto tunnel, le particelle legate hanno una probabilita' non nulla di esistere al di fuori della buca.

Consideriamo ora solamente l'e- che occupa il livello con energia piu' alta, trascurando quelli occupati di energia inferiore. Quando due atomi sono isolati, l'elettrone associato a ogni atomo possiede un'energia di valore E_n . Se ora consideriamo due buche di potenziale, ovvero due atomi, a una certa distanza tra loro, l'elettrone del primo atomo sara' soggetto a una forza dovuta alla presenza del secondo nucleo, e il potenziale che determina i livelli di energie sara' quindi moidificato e cosi' i possibili valori di energia. Inoltre bisogna considerare il principio di esclusione di Pauli: per ogni atomo a un livello di energia corrispondono al piu' due e- di spin opposto: il sistema di due atomi quindi puo' contenere al piu' quattro elettroni in tutto, ma se vengono avvicinati, il livello puo' competere sempre a soli 2 e-, e dunque devono esistere due livelli di energia leggermente differente, ciascuno con due e-, in modo da garantire la coesistenza di 4 e-.

Quindi c'e' il doppio effetto di perturbare l'energia e di sdoppiare i livelli.

Il sistema presentera' quindi un insieme di stati legati, con energia pero' non coincidente con gli stati legati della singola buca per il principio di esclusione, ma che subiranno uno sdoppiamento, ovvero si differenzieranno in due livelli con energia leggermente piu' alta o piu' bassa di quella originale, in

modo che ogni livello sia occupabile solo da al piu' due e-. Inoltre si avra' sempre un continuo di energie corrispondenti a particelle libere. Si noti che i livelli piu' bassi dello spettro saranno poco sdoppiati (molto vicini in energia) perche' la probabilita' di effetto tunnel tra buche cosi' lontane e' piccolo, quindi la compresenza di e- a uguale energia e' bassa. Per i livelli piu' alti, occupati da particelle poco legate, tale probabilita' e' piu' alta, quindi piu' forte sara' lo sdoppiamento.

Introducendo un sistema con 3, 4 o molti atomi, ovvero buche di potenziale, i livelli energetici permessi per ogni buca sono triplicati, quadruplicati o moltiplicati. Al limite di un numero infinito di buche, come un cristallo di atomi, presenta un insieme di strutture a bande di energie permesse e bande proibite, che sono sempre piu' estese man mano che l'energia cresce. Il limite superiore delle bande e' l'energia per cui il le particelle sono libere, oltre cui lo spettro e' continuo.

Se il sistema e' composto da N atomi, l'originario livello di energia E_n si scinde in N livelli permessi, che costituiscono una banda di energia contenente al massimo $2N$ elettroni. Visto che $N \simeq 10^{22} \text{cm}^{-3}$ e l'estensione della banda e' di pochi eV, le differenze di energie tra stati sono cosi' piccole che qualunque eccitazione termica a temperatura ambiente permette agli e' di effettuare transizioni tra i livelli. E' sostanzialmente una banda continua di valori di energia permessi, occupabile al piu' da $2N$ e-.

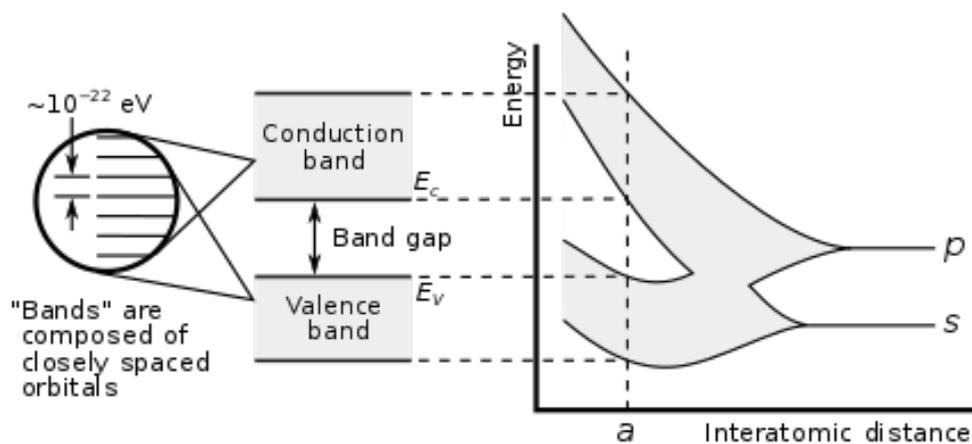


Figura 19: Configurazione delle bande energetiche in funzione della distanza interatomica (a destra) e struttura corrispondente nel cristallo (*in centro*). La distanza tra stati nella banda e' di 10^{-22} eV (*a sinistra*). Il parametro a equivale a 5.43 \AA

Le bande quindi individuano:

- un'energia minima di banda
- un'energia massima di banda
- eventuali intervalli di energia proibiti, o di *gap*, tra bande consecutive

Le bande e i limiti interessanti ai fini del comportamento elettrico sono quelle superiori:

- la banda contenente gli elettroni *di valenza* cioè quelli piu' esterni dell'atomo che contribuiscono ai legami chimici, e' detta appunto *banda di valenza*
- la banda immediatamente superiore, detta *banda di conduzione*
- l'estremo superiore della banda di conduzione e' detto *livello del vuoto*, ovvero l'energia minima che un e- deve possedere per sfuggire all'esterno del solido. Al di sopra di questo livello l'e- e' libero
- all'interno di ciascuna banda (di v. o c.) la relazione di dispersione dell'e- si puo' approssimare con una parabolica ma con massa efficace m^*

Accanto alle particelle libere e quelle legate, ci sono quindi le particelle quasi-libere, che si possono muovere per effetto tunnel da una buca all'altra, quindi sono particelle libere ma con relazione tra energia e momento diversa da quella quadratica:

$$E - E_0 \simeq \frac{1}{2m^*} |p - p_0|^2$$

dove m^* e' la massa efficace della particella, ovvero un parametro globale che considera il moto della particella semilegata come se fosse una particella libera che si muove nello spazio libero. Sono concetti fondamentali per lo studio del comportamento degli elettroni in una struttura cristallina come quella dei semiconduttori.

2.2.1 Le bande energetiche dei semiconduttori

Le proprieta' elettroniche di un materiale sono determinate dalla struttura a bande, e in particolare dalla presenza di intervalli proibiti e dalla loro estensione in energia (Fig. 20).

Come abbiamo visto, man mano che ci si allontana dall'energia di vuoto le bande diventano piu' estese. Per energie piu' elevate, l'elettrone e' meno legato, quindi ha un raggio di Bohr associato piu' grande, quindi diventa piu' influenzato da altri atomi, e quindi la sua energia ne e' piu' perturbata: di conseguenza c'e' una piu' ampia separazione tra i livelli originali che costituiscono la banda.

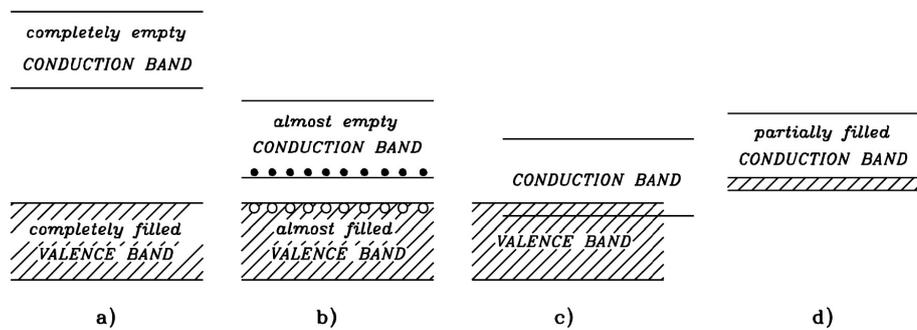


Figura 20: Le bande energetiche nei diversi tipi di materiale: a) isolante, b) semiconduttore, c) e d) metallo

Per esempio:

- un metallo alcalino con 1 solo elettrone di valenza nell'orbita piu' esterna, in un sistema composto da N atomi ravvicinati, avra' un livello energetico che degenera in una banda di energia. Nel caso piu' semplice, gli N elettroni di valenza occupano la meta' inferiore dei 2N stati della banda, lasciando stati vuoti vicino a stati occupati. Gli elettroni vicini all'energia massima possono essere facilmente eccitati o strappati fino a farli uscire dagli stati della banda e poi trasportati nel cristallo. Sono quindi conduttivi.
- i materiali che invece gli e- di valenza (cioe' dell'orbita piu' esterna) occupano completamente una banda permessa, che e' separata da un intervallo proibito dalla banda immediatamente superiore. In questo caso, la banda superiore a basse temperature e' completamente vuota. Se $E_g > 5eV$ allora tali materiali sono detti isolanti. Tale gap e' infatti pari infatti a 200 volte l'energia termica a temperatura ambiente o quella somministrabile con campi elettrici (decimi di eV).
- se l'intervallo proibito e' abbastanza piccolo ($\sim 1 eV$), esso e' abbastanza facile da superare per un elettrone con un'energia termica moderata

(40 volte la energia termica a temperatura ambiente), o con un campo elettrico applicato anche debole. Inoltre, l'aggiunta di impurezze opportune nel cristallo puo' aumentare in maniera controllata la conduttivita' del materiale. Questa pratica di aggiungere "impurezze" e' chiamata drogaggio, e ne parleremo profusamente in seguito. Materiali con questa configurazione di bande sono detti, appunto, semiconduttori.

A temperatura piu' alta dello zero assoluto, la banda di valenza non risulta completamente occupata, perche un piccolo numero die' possiete energia termica sufficiente per essere eccitato oltre il gap proibito. Nella banda superiore, quasi vuota, gli e- possono facilmente acquistare la poca energia sufficiente a originare una corrente quando sottoposti a campo elettrico. Essa e' la banda di conduzione.

2.2.2 Cristallo, legami covalenti, intervallo proibito

La struttura a reticolo cristallino di diamante, silicio e germanio prevede che ogni atomo del cristallo sia interessato da quattro legami covalenti con i primi vicini. Ogni legame corrisponde a due elettroni strettamente vincolati, uno per atomo. Allo zero assoluto tutti gli e' sono vincolati e nessuno e- disponibile per la conduzione in caso di campo esterno applicato. A temperature maggiori dello zero assoluto, alcuni legami si rompono e creano e- quasi liberi che possono contribuire alla corrente. Essi si posizionano sul limite basso della banda di conduzione. Lasciando uno stato libero da e', un altro e- puo' saltare al legame non occupato. Il moto risultante degli stati vacanti e' opposto a quello degli e-, e viene assimilato a moto di cariche positive.

Nei semiconduttori, quando un elettrone nella BV riceve energia e viene promosso in BC, lascia libera una buca in BV. Ciò può verificarsi, ad esempio, a causa di un fotone incidente, come a sinistra in fig. 21. Le buche riescono a muoversi nel materiale per sostituzione con elettroni di altri atomi, presenti in BV. In assenza di campo elettrico, a temperatura ambiente, elettroni e buche hanno **moti casuali** all'interno del materiale, e non si osserva una corrente netta. Se invece siamo in presenza di un **campo elettrico**, buche ed elettroni subiscono un moto di **deriva**, come a destra fig. 21. In questo modo, è possibile **raccogliere la carica** all'estremita del materiale, ottenendo un segnale elettrico in uscita.

Per il silicio, il **quanto di segnale** è di circa 3.6 eV, anche se la sua energia di gap ($E_g \simeq 1$ eV) è più piccola: idealmente, ci aspetteremmo lo stesso valore. Vediamo questo comportamento usando una radiazione di fotoni incidente sul

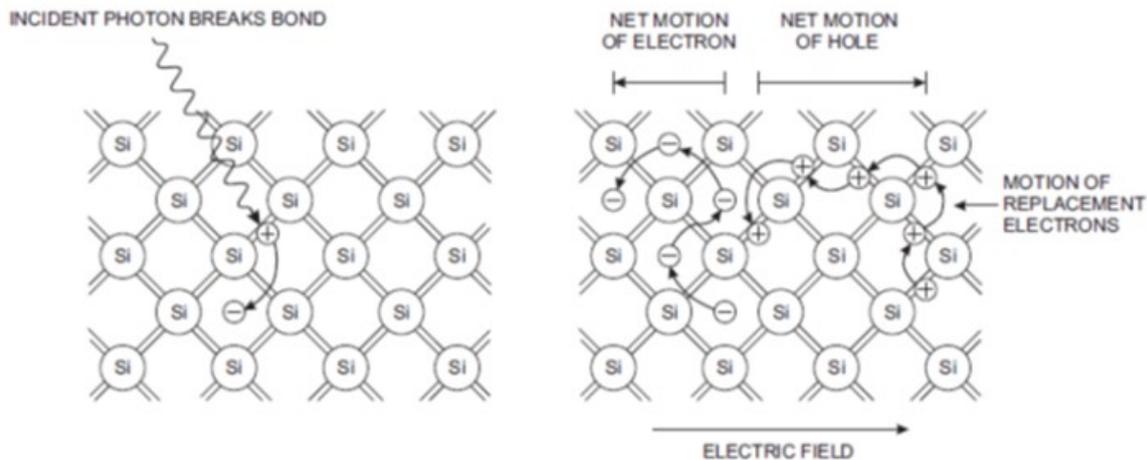


Figura 21: Struttura cristallina del Silicio, visualizzata in 2D: ogni legame e' formato da due e-, che appartengono ognuno a un atomo. Una particella incidente (o l'energia termica) puo' liberare un e- rompendo un legame, e creando un e- quasi libero disponibile a formare una corrente sotto l'azione di un campo elettrico. Il sito rimasto vacante, ovvero la lacuna, puo' essere occupata da un altro e- che a sua volta lascerà un altro sito vacante. Il processo si ripete, creando una corrente di lacune, in senso opposto a quello degli e-.

materiale: osserviamo un assorbimento (i.e. numero di elettroni estratti) che ha un andamento come quello nel grafico in fig. 22. Si tratta di un **fenomeno a soglia**, che avviene solo per energie del fotone $E_\gamma \geq E_g$. Ad un'iniziale fase di rapida crescita, ne segue una più lenta: parte dell'energia fornita viene **dissipata nel reticolo**, non contribuendo ad ulteriori ionizzazioni. Aumentando ancora l'energia dei fotoni, si osserva un flesso, oltre il quale la crescita diventa nuovamente rapida. L'assorbimento infine raggiunge una zona di **saturazione**, con delle fluttuazioni, ad un valore di energia che è quello identificato dal quanto di segnale.

L'**energia media di ionizzazione** I dipenderà dall'energia E_γ dei fotoni incidenti. In particolare, osserviamo che:

- per $E_\gamma \leq 4$ eV ho la produzione di una sola coppia elettrone - buca;
- per $E_\gamma \simeq 6$ eV l'energia di ionizzazione raggiunge il valore massimo $I \simeq 4.4$ eV;

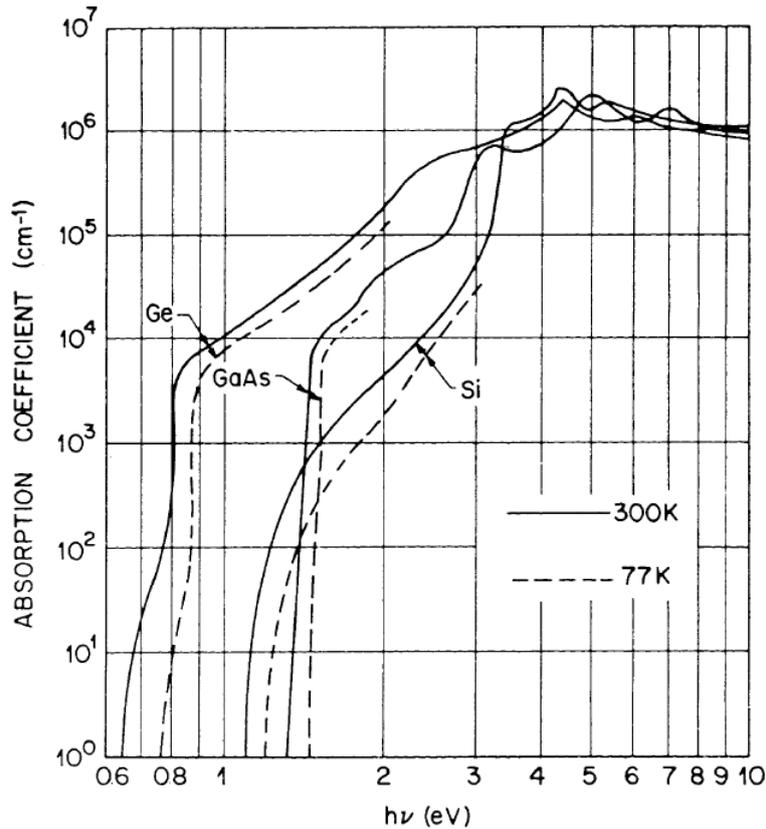


Figura 22: Coefficiente di assorbimento dei fotoni per unità di cammino nel materiale (a cui corrisponde il numero degli elettroni emessi) in funzione della loro energia $E_\gamma = h\nu$, per diversi materiali. Il fenomeno avviene per $E_\gamma \geq E_g$, con una rapida crescita iniziale. Segue una fase di crescita più lenta, a causa dell'energia persa per eccitazioni reticolari. Infine si ha ancora un'ultima fase a rapida crescita, prima della saturazione.

- per $E_\gamma \geq 1.5 \text{ keV}$ l'energia di ionizzazione assume il valore costante $I \simeq 3.6 \text{ eV}$, che è quello tipicamente usato per scopi pratici.

Per altri materiali semiconduttori, L'arseniuro di gallio (GaAs, detto anche *gallio-arsenio*), ha un gap di $E_g = 1.42 \text{ eV}$ a 300 K, e un'energia di ionizzazione media di 4.2 eV ; il germanio (Ge) ha un gap di $E_g = 0.67 \text{ eV}$ a 300 K, e un'energia di ionizzazione media di 2.5 eV .

Come rappresentato nel grafico in fig. 23, esiste una **relazione lineare** fra l'energia media di ionizzazione I e quella di gap E_g . Da un *fit* lineare, si ottiene l'equazione:

$$I = 2.8 E_g + 0.6 \text{ eV} \quad (12)$$

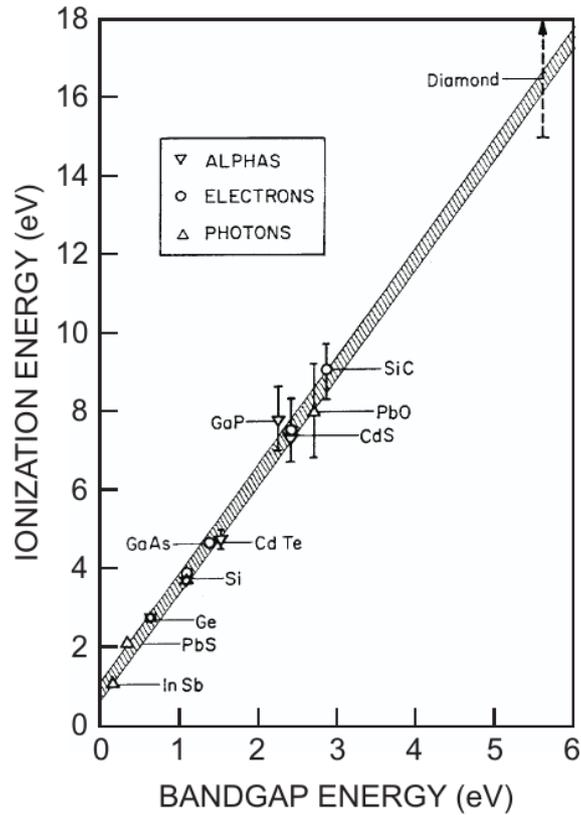


Figura 23: A sinistra: coefficiente di assorbimento dei fotoni (ovvero il numero di elettroni emessi), in funzione della loro energia $E_\gamma = h\nu$, a valori prossimi a quello di gap, per diversi semiconduttori. A destra: relazione fra energia di ionizzazione media I ed energia di gap E_g , per diversi semiconduttori.

Per molti materiali, in particolare per il silicio, si ha $I/E_g \simeq 30\%$, che è la frazione di energia effettivamente spesa per produrre segnale. Il restante $\simeq 70\%$ dell'energia viene invece dissipato in eccitazioni reticolari.

2.2.3 Le lacune

La **densità di corrente** \mathbf{J} (corrente per unità di area) che fluisce nella banda di conduzione è dato dalla somma dei contributi per unità di volume dei prodotti di carica q per velocità risultante dell' i -esimo elettrone v_i :

$$J_{bc} = \frac{I}{A} = \sum_{bc} (-q)v_i$$

La corrente e' comunque molto piu' piccola di quella in un metallo, poiche' il numero di e- in questa banda e' modesto.

Nel momento in cui un elettrone lascia il proprio orbitale, ovvero il proprio livello energetico, lascia un posto vacante chiamato *lacuna*. Essa non e' propriamente una carica, ma essendo determinata da una assenza di carica negativa, nell'economia globale di un sistema definito come il cristallo, e' assimilata a una carica positiva. Se le condizioni energetiche sono favorevoli, tale lacuna sar  presto occupata da un altro e-, che a sua volta lascerà libera un'altra lacuna. Applicando in maniera ricorsiva il processo, si vede come si pu  ottenere un movimento di lacune, ovvero di cariche positive, nel cristallo del semiconduttore. Il processo di movimento delle lacune, come vedremo, non e' un movimento diretto, ma e' conseguente del movimento degli elettroni: risulta quindi come un prodotto delle probabilit  di movimento degli elettroni, ed e' quindi un fenomeno pi  lento.

Quando in un semiconduttore un elettrone passa dalla banda di valenza alla banda di conduzione, lascia un livello della valenza libero, quindi una lacuna, che si pu  muovere nella banda di valenza secondo questo processo. La banda di valenza e' quindi responsabile per la conduzione di lacune, come la banda di conduzione lo e' per gli elettroni.

Se dovessimo esprimere matematicamente la densit  di corrente degli elettroni nella banda di valenza, dovrei fare lo stesso conto fatto per la banda di conduzione, con la difficult  di avere un gran numero di elettroni in banda di valenza da considerare. Sarebbe

$$J_{bv} = \sum_{bv} (-q)v_i$$

Per semplificare, posso considerare il moto di tutti gli elettroni in banda saturata e sottrarre il contributo degli elettroni che occuperebbero i soli stati vuoti, ovvero le lacune. La densit  di corrente diventa quindi:

$$J_{bv} = \sum_{\text{banda piena}} (-q)v_i - \sum_{\text{stati vuoti}} (-q)v_i = 0 + \sum_{\text{stati vuoti}} qv_i$$

Notiamo che gli elettroni nella banda di valenza sono in grande numero, e gli stati disponibili sono pochi. Ne risulta che gli e- hanno un moto diverso nella banda di valenza, e le loro singole velocit  saranno mediamente diverse da quelle in banda di conduzione.

Si puo' dunque esprimere il movimento delle cariche nella banda di valenza come moto di stati non occupati, ovvero lacune. Esse esistono solo in un

solido con struttura a bande, non esisterebbero in atomi liberi. L'energia delle lacune aumenta man mano che ci si muove verso il basso del diagramma delle bande (serve piu' energia alla lacuna per essere creata piu' in basso).

2.3 I semiconduttori intrinseci

2.3.1 Densita' degli stati

E' importante conoscere la densita' di stati possibili in un solido per poter calcolare in seguito le correnti di portatori di carica che useremo per trattare i semiconduttori nei vari regimi di funzionamento. La densita' di stati e' definita come il numero di stati disponibili (in banda di conduzione per gli e- e in banda di valenza per le lacune h+) per unita' di volume e di energia. Anche se il modello a bande ammette una distribuzione continua nelle due bande, il numero di livelli disponibili per unita' di energia in un volume finito e' limitato dal principio di esclusione.

Dalla fisica quantistica otteniamo che la densita' di stati $g_c(E)$ e $g_v(E)$ per unita' di energia per le bande di conduzione e valenza rispettivamente sono:

$$g_c(E) = \frac{dN_c}{dE} = \frac{4\pi}{h^3} (2m_e^*)^{3/2} \sqrt{E - E_c} \quad (13)$$

e considerando che l'asse dell'energia e' invertita per le lacune:

$$g_v(E) = \frac{dN_v}{dE} = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \sqrt{E_v - E} \quad (14)$$

In queste formule la massa efficace dell'e- e della lacuna nel semiconduttore e' definita come la massa che avrebbe un e- o una lacuna con lo stesso comportamento nel vuoto. La forma di tali distribuzioni compare nel secondo grafico da sinistra in Fig. 25.

2.3.2 Probabilita' di occupazione degli stati

Nei materiali semiconduttori puri, per ogni e- eccitato rimane una lacuna libera. Ovvero il numero n di e- in banda di conduzione e' uguale al numero p di lacune nella banda di valenza. Essendo gli e- dei fermioni, in base a considerazioni di base sul modo in cui gli stati di energia permessa sono occupati da particelle soggette al principio di esclusione di Pauli si ottiene che

la probabilita' di occupare uno stato con energia E e' dato dalla distribuzione di Fermi-Dirac:

$$f_e(E) = \frac{1}{e^{(E-E_F)/kT} + 1}$$

ove E_F e' il livello di riferimento detto di Fermi, a cui e' associata un'energia detta di Fermi, per il quale la probabilita' di occupazione vale:

$$f_e(E_F) = \frac{1}{e^{(E_F-E_F)/kT} + 1} = 1/2$$

L'energia del livello di Fermi E_F rappresenta l'ipotetico livello di energia che all'equilibrio termodinamico avrebbe il 50 % di probabilita' di essere occupato da un elettrone in qualunque istante considerato. Se consideriamo un grafico della funzione (Fig. 24) vedremo che la f ha valori prossimi all'unita' per energie inferiori a E_f , mostrando che gli stati piu' bassi sono occupati per lo piu'. A energie piu' elevate ha valori piccolissimi, e pochi e- occupano gli stati ad alta energia. Proviamo a cambiare T : per valori nulli, la funzione si riduce a una funzione gradino uguale a 0 per $E < E_f$ e 1 per $E > E_f$. Per T che cresce, la discesa da 1 a 0 si fara' sempre piu' graduale, cosa che si riflette nella possibilita' sempre piu' grande di avere stati non occupati nel limite superiore di banda di valenza e sempre piu' occupati nella banda di conduzione.

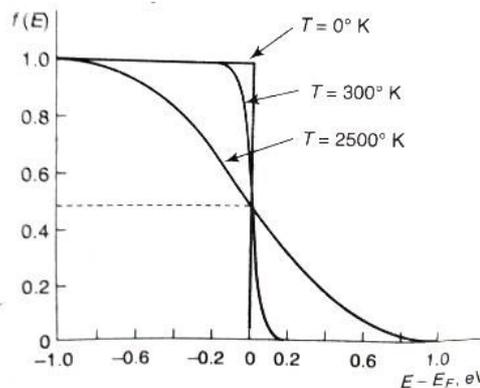


Figura 24: Distribuzione di Fermi-Dirac al variare della temperatura. A $T=0$ K e' una funzione gradino, con transizione sempre piu' graduale man mano che la T cresce.

La distribuzione di probabilita' della non occupazione degli stati, ovvero della presenza di lacune negli stati, e'

$$f_h(E) = 1 - \frac{1}{e^{(E-E_F)/kT} + 1} = \frac{1}{e^{(E_F-E)/kT} + 1}$$

Se c'è uguale concentrazione di e- e h+, come nel materiale semiconduttore intrinseco, allora $f_e(\frac{E_c+E_v}{2}) = f_h(\frac{E_c+E_v}{2})$ e quindi $E_F - \frac{E_c+E_v}{2} = \frac{E_c+E_v}{2} - E_F$ e $E_F = \frac{E_c+E_v}{2}$ e il livello di Fermi è proprio posto a metà dell'intervallo proibito.

Con un esercizio semplice possiamo calcolare che la probabilità che uno stato della banda di conduzione sia occupato da un e- nel silicio a 300 K per sola energia termica è circa $4.4 \cdot 10^{-10}$. Piccola, ma non trascurabile se consideriamo che la densità degli stati disponibili nella banda di conduzione è dell'ordine di 10^{19} . Come vedremo, quello che otterremo è una concentrazione di portatori nel silicio intrinseco di circa $1.45 \cdot 10^{10}/cm^3$.

La distribuzione di F-D non rappresenta da sola la popolazione di e- ad una certa energia. Dalla fisica quantistica abbiamo però la densità di stati disponibili in funzione dell'energia $g(E)$, che abbiamo trattato nel paragrafo 2.3.1. La distribuzione degli e- in funzione dell'energia si ottiene dal prodotto della densità di stati permessi in un piccolo intervallo di energia per la probabilità che questi stati siano occupati. La densità totale di elettroni n nella banda si ottiene poi integrando sull'energia, su tutta la banda di conduzione:

$$n = \int_{bc} f_e(E)g_c(E)dE$$

e analogamente per la densità di lacune p , integrando sulla banda di valenza:

$$p = \int_{bv} f_h(E)g_v(E)dE$$

come mostrato in Fig. 25.

Se la distanza tra E_F e il minimo della BC è abbastanza grande ($(E_c - E_F) \gg kT$) si può usare una formulazione statistica semplificata dal punto di vista matematico, che vale anche per particelle che non osservano il principio di Pauli. Questo è compatibile con l'assunzione che gli e- in banda di conduzione siano in numero molto ristretto in confronto al numero di stati disponibili, per cui il principio di esclusione non ha alcun effetto. Si tratta della distribuzione di Maxwell-Boltzmann, secondo la quale:

$$f_e(E) = \frac{1}{e^{(E-E_F)/kT} + 1} \approx e^{-\frac{(E-E_F)}{kT}}$$

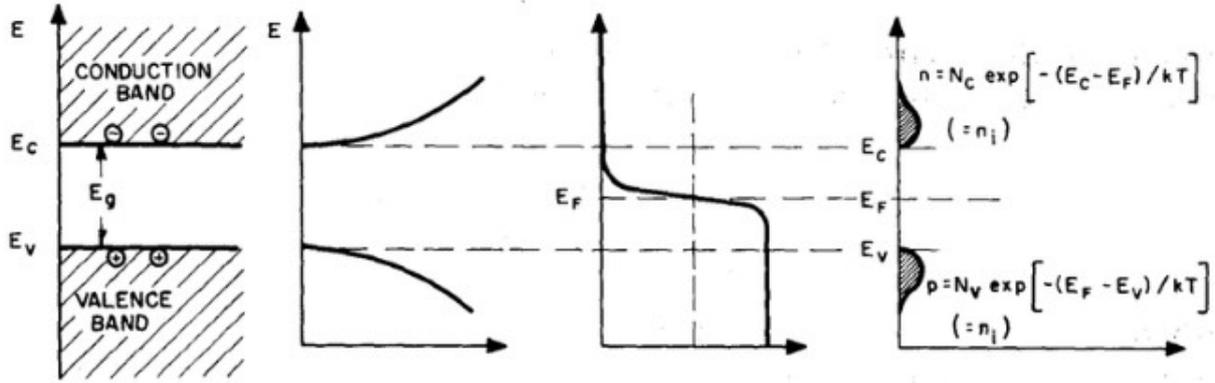


Figura 25: Semiconduttore intrinseco: struttura a bande, densita' degli stati disponibili, distribuzione di probabilita' dell'occupazione degli stati, distribuzione risultante degli stati occupati.

Inserendo tale definizione nell'integrale precedente otteniamo dopo numerose approssimazioni:

$$n = N_c e^{-\frac{(E_c - E_F)}{kT}} \quad (15)$$

con N_c densita' efficace di stati al bordo inferiore della banda di conduzione. Per definizione esso rappresenta tutti gli stati della banda di conduzione ma concentrati all'energia E_c .

Allo stesso modo si puo' procedere per la concentrazione di lacune:

$$p = N_v e^{-\frac{(E_F - E_v)}{kT}} \quad (16)$$

con tutti gli stati della banda di valenza concentrati all'energia E_v nella densita' efficace N_v .

Le quantita' N_c e N_v sono date da:

$$N_c = 2 \left(\frac{2\pi m_n^* kT}{h^2} \right)^{3/2} \propto T^{3/2} \quad (17)$$

e

$$N_v = 2 \left(\frac{2\pi m_p^* kT}{h^2} \right)^{3/2} \propto T^{3/2} \quad (18)$$

Nel silicio cristallino, a 300 K abbiamo $N_c = 2.8 \times 10^{19} \text{ cm}^{-3}$ e $N_v = 1.04 \times 10^{19} \text{ cm}^{-3}$.

A parte la differenza tra m_n^* e m_p^* , la cui origine non tratteremo in dettaglio, $N_c \simeq N_v$. Siccome $n = p$, avro che $(E_c - E_F) \simeq (E_F - E_v)$ e quindi

$E_F \simeq (E_c + E_v)/2$ come avevamo già visto: il livello di Fermi intrinseco sta a metà della banda proibita. Possiamo rinominarlo con il pedice i a indicare *intrinseco*: E_i .

Calcolo delle probabilità di occupazione Nel caso di silicio intrinseco, la probabilità di occupazione dello stato di energia minima della banda di conduzione a $T = 300\text{ K}$, $E_c - E_F = E_g/2 = 0.56\text{ eV}$, è:

$$P = \frac{1}{\exp\left(\frac{\Delta E}{kT}\right) + 1} \simeq \frac{1}{\exp\left(\frac{0.56}{0.026}\right) + 1} \simeq \frac{1}{2.26 \cdot 10^9 + 1} \simeq 4.4 \cdot 10^{-10} \quad (19)$$

da confrontarsi con $\mathcal{O}(10^{19})$ di stati disponibili.

2.4 I semiconduttori drogati (o *estrinseci*)

Un materiale semiconduttore puro è per definizione globalmente neutro, e per ogni e^- che passa dalla BV alla BC, c'è una lacuna corrispondente che si libera in BV. La concentrazione n di e^- nella banda di conduzione, che sono poi i portatori di carica negativi, è uguale alla concentrazione p di lacune nella banda di valenza. Questo materiale è chiamato *intrinseco*, e le sue concentrazioni vengono chiamate n_i e p_i rispettivamente.

Il metodo di efficiente per controllare il numero di portatori di carica in un semiconduttore è l'introduzione di impurità per sostituzione degli atomi del cristallo con atomi di gruppi diversi (terzo e quinto), come schematizzato in Fig. 26.

n-type - Donori Il fosforo (P) ha 5 elettroni di valenza, di cui 4 saranno impiegati nei legami con gli atomi di silicio (Si) nel cristallo, e uno è solo debolmente legato al suo atomo di appartenenza a causa della carica positiva in eccesso nel nucleo, ma non forma legame covalente, ed è considerato libero di muoversi nel momento in cui gli viene fornita una piccola quantità di energia.

L'energia richiesta per rompere il legame del quinto elettrone è data dal potenziale coulombiano effettivo originato dal nucleo dell'atomo donatore: l' e^- è attratto dalla singola carica netta del nucleo del donatore schermata dagli effetti di polarizzazione del mezzo circostante composto da atomi di silicio.

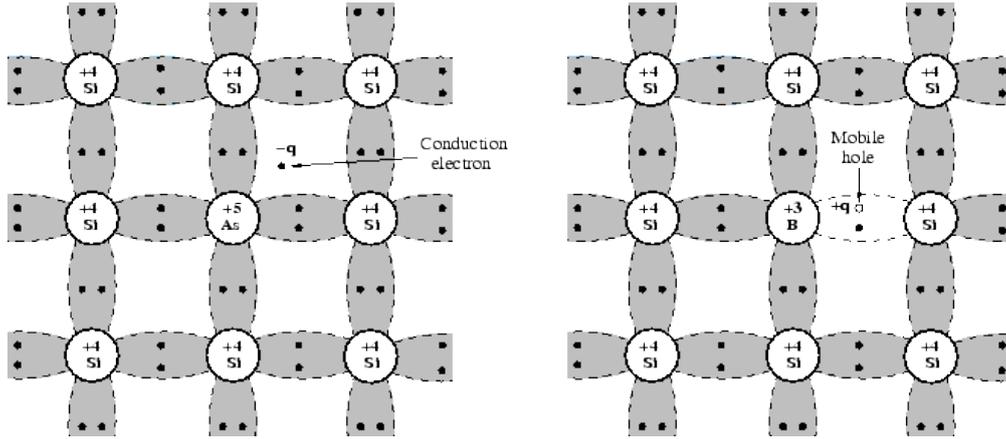


Figura 26: Schema bidimensionale della struttura cristallina del silicio in cui e' stato sostituito un atomo con uno di Arsenio (gruppo 5) a *sinistra*, e uno di Boro (gruppo 3) a *destra*. Compiono rispettivamente un e- e una lacuna relativamente liberi di muoversi, lasciando cosi' un atomo ionizzato nel cristallo.

Come detto, a partire dal modello dell'atomo di Bohr, conosciamo come interagisce un elettrone con il potenziale coulombiano di un nucleo atomico: esiste uno spettro di valori consentiti di energia che l'e- puo' assumere, corrispondente a

$$E_n = \frac{-Z^2 m_0 q^4}{8 \epsilon_0^2 h^2 n^2} \quad (20)$$

al di sotto un'energia di riferimento assunta pari a 0. Nella formula: Z e' la carica netta del nucleo, q e' la carica elementare, m_0 e' la massa in quiete dell'elettrone, ϵ_0 e' la costante dielettrica nel vuoto, h la costante di Plank, e n un numero intero positivo. Per l'idrogeno, con $Z = 1$ i valori consentiti sono $-13.6/n^2 eV$ al di sotto del livello di riferimento, con n intero, quindi: $-13.6 eV$, $-3.4 eV$, $-1.5 eV$, $-0.85 eV$, ecc ecc. Allo zero assoluto, il livelli a energia piu' bassi sono i primi ad essere occupati, con al massimo 2 e- per livello, seguendo il principio di esclusione di Pauli.

Teoricamente, l'energia di legame nel reticolo di semiconduttore per l'elettrone di un atomo drogante donatore, come per esempio il Fosforo, si puo' interpretare come l'energia dell'e- dell'atomo idrogenoide pero' scalata del rapporto tra la massa efficace e la massa a riposo dell'e-, e per la costante dielettrica relativa del silicio (al quadrato). Tale energia risulta dunque:

$$E = \frac{m_n^* q^4}{8 \epsilon_0^2 \epsilon_{rs}^2 h^2} = \frac{13.6 m_n^*}{\epsilon_{rs}^2 m_0} eV \quad (21)$$

con ϵ_{rs} costante dielettrica relativa del semiconduttore e m_n^* e' la massa efficace dell'elettrone nella banda di conduzione, che tiene conto delle interazioni quantistiche tra il reticolo cristallino e l'elettrone in movimento. Nel silicio, $\epsilon_{rs} = 11.7$ e $m_n^* = 1.1 m_0$. Il risultato e' $E = 0.1 \text{ eV}$. Ricordiamo che il valore della massa efficace dell'e- m^* considerato per il calcolo dell'energia di legame non e' lo stesso di quello considerato per il calcolo della mobilita', che vedremo piu' avanti.

Tale energia e' stata calcolata in maniera piu' accurata e poi misurata, e per un fosforo nel silicio il livello di tale elettrone in eccesso e' $\sim 0.044 \text{ eV}$ sotto il limite inferiore delle BC e rappresenta circa il 5% di E_g . E' molto facile per un e- in questo stato energetico raggiungere la BC ed essere trasportato.

La densita' volumica di donatori, ovvero il numero di atomi droganti per unita' di volume (cm^3) e' indicata con N_d . Un e- di un donatore passa anche a bassa temperatura in BC, lasciando non una lacuna nei legami covalenti del reticolo, ma localmente uno ione positivo fisso nel cristallo. Se il drogaggio e' di soli donatori, allora **i portatori di carica maggioritari saranno gli e-**, mentre ci saranno **solo poche lacune libere nella BC, quelle liberate per l'eccitazione degli e- del silicio** che ci sarebbe anche nel silicio intrinseco. **Si e' persa dunque la simmetria tra portatori che si aveva nel silicio intrinseco.**

Le cariche maggioritarie sono gli e- debolmente legati che sono passati in banda di conduzione senza lasciare una buca, lasciando solo un atomo ionizzato con carica fissa positiva nel sito nucleare corrispondente. I portatori di carica minoritaria sono solo le buche che possono crearsi da uno degli e- che rompendo un legame covalente si liberano dalla struttura di cui facevano parte. L'energia di legame per donori tipici e' dagli 0.044 eV per fosforo a 0.049 eV per l'arsenio e 0.039 eV per l'antimonio. L'energia di legame cosi' bassa fa si' che sia piu' probabile che il legame a bassa energia che connette il quinto e- al donore sia rotto piu' facilmente dei legami del silicio.

p-type - Accettori Allo stesso modo, aggiungendo per sostituzione un atomo del gruppo terzo, quale il boro, con 3 elettroni di valenza, uno dei legami covalenti col silicio sara' lasciato non saturato. Un altro e- da un legame vicino si puo' spostare per occuparlo, liberando un altro legame. L'assenza di legame si sposta.

Le cariche maggioritarie sono qui l'assenza di legame covalente portata dall'accettore, che generano uno stato non occupato con energia poco piu' alta del limite superiore della banda di valenza, a circa $\sim 0.044 \text{ eV}$, facilmente

occupabili da e- della banda di valenza, lasciando un atomo ionizzato negativamente. Le cariche minoritarie sono gli e- che possono sempre staccarsi dagli atomi di silicio.

Si puo' ottenere un effetto molto utile accostando diverse densita' dei due tipi di portatori di carica: meccanicamente con due materiali diversi o drogati diversamente, oppure compensando il drogaggio di un tipo con il drogaggio di tipo opposto. Vista la limitata precisione nelle concentrazioni di drogaggio ottenute con le tecniche in uso, si preferisce differenziare di almeno 1 ordine di grandezza due drogaggi diversi perche' siano effettivi.

2.4.1 Legge dell'azione di massa

In una situazione di equilibrio termodinamico ci sono fenomeni concorrenti che si equilibrano: per esempio e- che vengono liberati e e- che si ricombinano, o vibrazioni reticolari del cristallo che si trasferiscono all'energia degli e- e viceversa.

Ad una data temperatura si liberano e- lasciando sguarnite le buche. Il tasso di generazione delle coppie dipende dalla temperatura T ma non dal numero di portatori gia' presente, almeno in prima approssimazione: $G = f_1(T)$ La velocita' di ricombinazione, ovvero per un e- libero di andare a occupare una lacuna e neutralizzarla, dipende dalla concentrazione di e- nella banda di conduzione e di lacune nella banda di valenza, e della temperatura. $R = npf_2(T)$ All'equilibrio termodinamico, il tasso di generazione e quello di ricombinazione si equivalgono ($f_1(T) = npf_2(T)$) e quindi il prodotto delle concentrazioni e' funzione della sola temperatura $np = f_3(T)$.

Per un semiconduttore intrinseco, $n = p = n_i$ e dunque $n_i^2 = f_3(T)$. La concentrazione di portatori intrinseci dipende dalla T , sorgente di energia termica che eccita' i portatori attraverso l'intervallo proibito. Inoltre dipende da tale intervallo E_g , che appunto limita il passaggio di e-.

Si puo' scrivere

$$n_i^2 = N_v N_c \exp\left(\frac{-E_g}{kT}\right)$$

Per quanto N_v e N_c varino con la temperatura, la dipendenza e' molto piu' forte nell'esponenziale che domina. Nel caso del silicio, per ogni 8 K di aumento di temperatura sopra la temperatura ambiente, la concentrazione di portatori raddoppia (e come vedremo anche la corrente di buio, quella che esiste in assenza di eventi esterni).

Da quanto visto prima,

$$np = N_c N_v e^{\frac{-(E_c - E_F) - (E_F - E_v)}{kT}} = N_c N_v e^{\frac{-(E_c - E_v)}{kT}} = N_c N_v e^{\frac{-E_g}{kT}}$$

Poiche' la concentrazione di portatori intrinseci di un determinato semiconduttore e a una determinata T dipende solo da E_g caratteristica del semiconduttore e dalla T , il prodotto $np = n_i^2$ non dipende dal drogaggio estrinseco, e quindi vale sia per gli intrinseci sia per i drogati. Intuitivamente, l'abbondanza dei portatori maggioritari fa si' che parte di essi vada a ricombinarsi con i portatori minoritari in modo da mantenere costante il loro prodotto.

Noti N_c e N_v a 300 K, dalla legge di azione di massa e' possibile ricavare la concentrazione intrinseca:

$$n_i = \frac{\sqrt{N_{c300K} N_{v300K}}}{(300 \text{ K})^{3/2}} \cdot T^{3/2} e^{(-E_g/2kT)} \quad (22)$$

Semiconduttore drogato Avendo sostituito una frazione degli atomi dell'elemento che forma il semiconduttore intrinseco con elementi di altri gruppi della tavola periodica, ho ottenuto un semiconduttore drogato. Dal punto di vista costruttivo, il silicio viene fabbricato come tipo n (o p), quindi con la presenza di donori (o accettori). Successivamente per creare una porzione drogata del tipo opposto, quindi p (o n), si aggiungono atomi di elementi diversi che saranno rispettivamente accettori (o donori). Quindi nella pratica un semiconduttore drogato pu' includere sia atomi *accettori* sia *donatori* (o *donori*), in concentrazioni indicate rispettivamente con N_A e N_D . Sar' il bilanciamento tra le due concentrazioni a determinare poi se si tratta di un s.c. di tipo P o N.

La concentrazione effettiva di drogaggio N_{eff} e' quindi la differenza tra concentrazioni di donori e accettori

$$\Delta N = N_{eff} = |N_D - N_A|$$

Dunque:

se $N_D < N_A$ allora otteniamo un s.c. di tipo P,

se $N_D > N_A$ allora otteniamo un s.c. di tipo N.

Potremmo per chiarezza quindi indicare con N_{A_P} e N_{D_P} le concentrazioni di accettori e donori nel materiale di tipo P, e con N_{A_N} e N_{D_N} le concentrazioni di accettori e donori nel materiale di tipo N.

Come vedremo in seguito, a livello pratico per un dato volume di s.c. drogato, avremo che una delle due concentrazioni di drogaggio è preponderante sull'altro, e la concentrazione minoritaria diventa trascurabile:

nel tipo P: $N_{A_P} \gg N_{D_P}$ e quindi $N_{eff} \approx N_{A_P}$

nel tipo N: $N_{D_N} \gg N_{A_N}$ e quindi $N_{eff} \approx N_{D_N}$

Nei s.c. drogati, per conservare la relazione tra le concentrazioni di carica (legge di azione di massa) e per rispettare la neutralità elettrica del sistema, il livello di Fermi si allontana dal centro del gap, avvicinandosi al limite superiore della banda di valenza E_V nel tipo P, e al limite inferiore della banda di conduzione E_C nel tipo N. In altre parole nel tipo P, se confrontato con il materiale intrinseco, ci sono più lacune libere di trasferirsi, quindi è più probabile che una lacuna venga portata in banda di valenza. Nel tipo N, sempre rispetto al materiale intrinseco, ci sono più e^- a disposizione, quindi è più probabile che un e^- salti nella banda di conduzione.

Nella zona in equilibrio ci sono cariche positive localizzate negli atomi donori e nelle buche portatrici, e cariche negative negli atomi accettori e negli e^- portatori. Se c'è neutralità di carica in una regione dove tutti gli atomi donori e accettori sono stati ionizzati, allora

$$N_d + p = N_a + n$$

Applicando la legge di azione di massa avremo che

$$n - \frac{n_i^2}{n} = N_d - N_a$$

che si può risolvere per la concentrazione di elettroni n:

$$n = \frac{N_d - N_a}{2} + \left[\left(\frac{N_d - N_a}{2} \right)^2 + n_i^2 \right]^{1/2}$$

In un s.c. di tipo n $N_d > N_a$. La concentrazione di e^- dipende dall'eccesso netto di donori e accettori ionizzati, Quindi un tipo-p che contiene N_a accettori può essere convertito in tipo-n impiantando un eccesso di donori tale che $N_d > N_a$.

Per il silicio a temperatura ambiente, $n_i = 1.45 \times 10^{10} \text{cm}^{-3}$ mentre la densità efficace tipica di donori nel tipo n è tipicamente almeno $n_i = 10^{15} \text{cm}^{-3}$. Quindi $N_d - N_a \gg n_i$ e $n \sim (N_d - N_a)$. Di conseguenza

$$p = \frac{n_i^2}{n} \sim \frac{n_i^2}{N_d - N_a}$$

Con tali valori per N_{eff} otteniamo quindi $p = 2 \times 10^5 \text{cm}^{-3}$, ovvero 10 ordini di grandezza più piccola della popolazione di portatori maggioritari.

Dall'integrazione del prodotto di funzione di distribuzione e densità di stati già vista per l'intrinseco, possiamo allo stesso modo derivare la formula per la concentrazione di portatori. Per un materiale di tipo n non molto drogato, non essendoci molti stati occupati tra quelli consentiti in banda di conduzione, il livello di Fermi E_f è abbastanza lontano dal limite inferiore della banda di conduzione E_C , ovvero $(E_C - E_f) \gg kT$, e la funzione di distribuzione di Fermi si riduce alla più semplice distribuzione di Maxwell-Boltzmann.

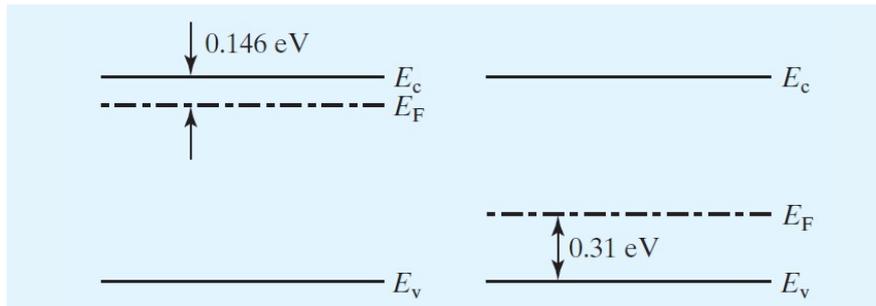


Figura 27: Posizione del livello di Fermi nel silicio estrinseco, drogato con $n = 10^{17} \text{cm}^{-3}$ a sinistra, e con $p = 10^{14} \text{cm}^{-3}$ a destra.

La visualizzazione usata per l'intrinseco, si può ora riproporre per il silicio drogato n e p (Fig. 28).

La concentrazione di portatori nella banda di conduzione per un silicio n moderatamente drogato è quindi:

$$n = N_c \exp \left[\frac{-(E_c - E_f)}{kT} \right]$$

Per un materiale di tipo p moderatamente drogato, la concentrazione di portatori lacune nella banda di valenza è

$$p = N_v \exp \left[\frac{-(E_f - E_v)}{kT} \right]$$

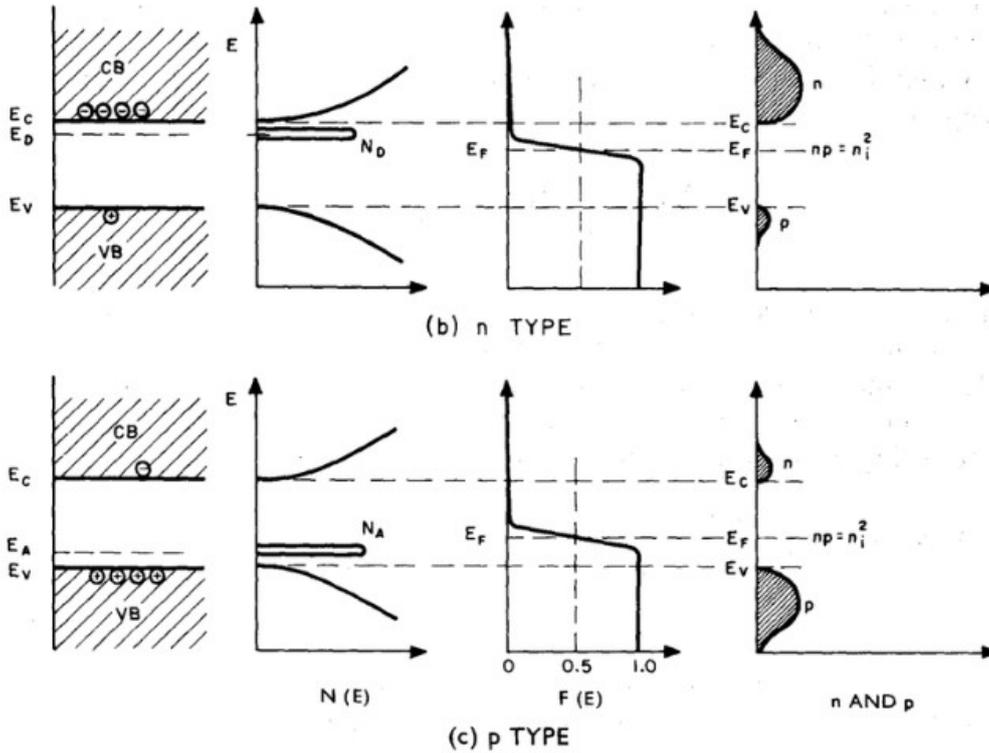


Figura 28: Semiconduttore estrinseco o drogato (*n* in *alto*, *p* in *basso*): struttura a bande, densita' degli stati disponibili, distribuzione di probabilita' dell'occupazione degli stati, distribuzione risultante degli stati occupati.

ove E_f e' il livello di Fermi nel materiale estrinseco.

Se nel s.c. intrinseco

$$n_i = N_c \exp \left[\frac{-(E_c - E_i)}{kT} \right] = N_v \exp \left[\frac{-(E_i - E_v)}{kT} \right]$$

in un s.c. estrinseco, ovvero drogato, possiamo scrivere allo stesso modo per il materiale moderatamente drogato *n*, la concentrazione dei portatori di carica maggioritari, ovvero gli e-, e':

$$n = n_i \exp \left[\frac{-(E_i - E_{f_n})}{kT} \right] = n_i \exp \frac{q\phi_n}{kT}$$

e per quello debolmente drogato *p*, la concentrazione di portatori maggioritari (lacune) e':

$$p = n_i \exp \left[\frac{-(E_{f_p} - E_i)}{kT} \right] = n_i \exp \frac{q\phi_p}{kT}$$

ove ho ridefinito la differenza tra livelli energetici come il potenziale associato a tale energia per la carica q nel caso di concentrazioni n e p : $q\phi_n = E_{f_n} - E_i$ e $q\phi_p = E_i - E_{f_p}$.

Quindi l'energia di separazione tra il livello di Fermi e il livello di Fermi intrinseco e' una misura della deviazione del materiale rispetto all'intrinseco. Visto che $E_f > E_i$ in un s.c. di tipo N , allora $n > n_i > p$.

Dal rapporto tra concentrazione intrinseca e estrinseca otteniamo:

$$\frac{q\phi_n}{kT} = \ln \frac{n}{n_i}$$

$$\frac{q\phi_p}{kT} = \ln \frac{p}{n_i}$$

Questi torneranno utili quando faremo interagire i due tipi di materiale: nella giunzione pn infatti la somma di ϕ_n e ϕ_p infatti rappresentera' la differenza di potenziale ai capi della giunzione. Inoltre, quando il livello di Fermi e' disturbato da radiazione incidente o da una polarizzazione applicata a una giunzione di materiali p ed n , una delle concentrazioni di portatori, o entrambe, vengono temporaneamente turbate e dunque i livelli di Fermi relativi ai due tipi di portatore variano rispetto a quello all'equilibrio. Per descrivere il **sistema fuori dall'equilibrio** si considerano i quasi-livelli di Fermi, che rappresenteranno la deviazione in potenziale dai livelli di Fermi all'equilibrio. I quasi-livelli di Fermi si definiscono per conservare le relazioni tra concentrazione di intrinseci e le concentrazioni di e' e lacune in condizioni di non equilibrio.

Quando il semiconduttore contiene un drogaggio molto alto, con concentrazioni N_a vicina a N_V o N_d vicina a N_C (per esempio 10^{19}cm^{-3} per il Si) allora non possiamo ignorare le limitazioni imposte dal principio di esclusione di Pauli. La Fermi-Dirac allora non si puo' approssimare da M-B e le equazioni per n e p non sono piu' valide. Per concentrazioni che superano gli stati sul fondo della banda, si parla di **semiconduttori degeneri** perche' il livello di Fermi ricade nelle bande e non nell'intervallo proibito.

Esercizio Ci si puo' domandare quale sia la probabilita' di transizione in banda di conduzione da tale stato posto nella banda proibita del Silicio, poco sotto E_c : quello stato e' lo stato di un e' debolmente legato all'atomo di P. Il problema si puo' approssimare prima calcolando la probabilita' di non ionizzazione dell'atomo di P, ovvero che l' e' stia nel suo sito, e poi calcolarne la probabilita' complementare.

Grandezza	unita'	Si	Ge	GaAs
Densita' atomica	cm^{-3}	5.0×10^{22}	4.22×10^{22}	2.21×10^{22}
E_g (300K)	eV	1.21	0.66	1.4-1.9
E_g (4.4K)	eV	1.21	0.74	1.52
m_n^* (en. legame)	m_e	1.08	0.55	0.068
m_n^* (condutt.)	m_e	0.26	0.12	0.068
m_p^* (en. legame)	m_e	0.81	0.3	0.5
m_p^* (condutt.)	m_e	0.386	0.3	0.5
N_c	cm^{-3}	2.8×10^{19}	1.04×10^{19}	4.7×10^{17}
N_v	cm^{-3}	1.04×10^{19}	6×10^{18}	7×10^{18}
n_i	cm^{-3}	1.45×10^{10}	2.4×10^{13}	9×10^6

Tabella 3: Quantita' tipiche per alcuni parametri caratteristici a $T = 300$ K per tre materiali semiconduttori comuni

2.5 Correnti: il trasporto dei portatori

I semiconduttori all'equilibrio sono corpi di carica complessivamente neutra, senza applicazione di potenziali esterni e dopo aver atteso un periodo di stabilizzazione abbastanza lungo da raggiungere l'equilibrio. I portatori sono particelle quasi libere, cioe' non associate ad alcuna particolare posizione reticolare ma comunque soggetti alle forze esercitate dal cristallo. Per questo si possono modellizzare come particelle dotate di una massa efficace minore di quella dell'e- che si muova nel vuoto, per tener conto di tali forze agenti invece nel semiconduttore. Nel silicio a $T = 300$ K tale massa diventa $m_{mobility}^* = 0.26 m_0$ ove m_0 e' la massa a riposo dell'e-. Si noti che questa massa efficace e' diversa da quella calcolata per la concentrazione di stati nella banda di conduzione, per esempio (che ricordiamo era circa $m_{density}^* = 1.08$ per l'elettrone).

Per la meccanica statistica, e- e lacune sono dotati dell'energia termica tipica delle particelle classiche libere, cioe' $\frac{1}{2}kT$ unita' di energia per ogni grado di liberta', con $k = 8.6 \times 10^{-5} eV/K$ costante di Boltzmann, e T temperatura assoluta. Per particelle puntiformi che si muovano in 3 dimensioni, l'energia termica sara' dunque

$$E_c = \frac{1}{2} m_n^* v_{th}^2 = \frac{3}{2} kT$$

(con velocita termica $v_{th} \simeq 10^7 cm/s$). La velocita' quadratica media rimarra' quindi legata alla temperatura approssimativamente da tale relazione.

Gli elettroni qui considerati sono in movimento attraverso il reticolo con velocità che ha direzione distribuita aleatoriamente, e soggetti a urti con il reticolo e con gli altri e-. Il moto è completamente casuale all'equilibrio termodinamico, la corrente risultante in ogni direzione è nulla. Gli urti con il reticolo provocano un trasferimento di energia tra e- e nuclei atomici. Indichiamo il tempo medio tra due collisioni calcolato sulla popolazione di e- come τ_{cn} .

Per le cariche in movimento nel semiconduttore, convivono due fenomeni legati al movimento delle cariche: *deriva* (nel caso ci sia un campo elettrico nel materiale) e *diffusione*, che andremo ora a trattare in dettaglio.

2.5.1 Deriva

In caso di applicazione dall'esterno di campi elettrici, i portatori di carica si muovono per **deriva** (*drift*). In questo caso il loro movimento netto è dovuto all'accelerazione che essi subiscono dal campo elettrico tra una collisione casuale e la prossima, e segue la direzione imposta dal campo elettrico E . Nel modello che stiamo seguendo, si assume che tutto l'impulso acquisito dalla particella per azione della forza del campo elettrico tra una collisione e l'altra (impulso è forza x tempo) sia trasferito al reticolo al momento dell'urto. La particella è quindi soggetta a continue accelerazioni dopo ogni urto. Questa trattazione intuitiva fa anche capire come la velocità dipenda linearmente dal campo elettrico fino a un certo limite, ma poi saturi a un massimo per campi elettrici molto intensi.

Quindi dopo un urto, con forza agente su e- pari a $-qE$ il momento acquisito è $m_n^*v_d$, quindi

$$-qE\tau_{cn} = m_n^*v_d$$

da cui deriva la proporzionalità tra la velocità di deriva e il campo elettrico applicato:

$$v_d = -\frac{qE\tau_{cn}}{m_n^*}$$

Tale fattore di proporzionalità è la **mobilità** per gli e-:

$$\mu_n = \frac{q\tau_{cn}}{m_n^*}$$

e per le lacune:

$$\mu_p = \frac{q\tau_{cp}}{m_p^*}$$

La mobilita' si esprime normalmente in $\text{cm}^2(\text{Vs})^{-1}$. Come si vede le mobilita' dipendono dal tempo medio tra collisioni e dalla massa efficace.

La velocita' media di deriva del portatore si ottiene da

$$\mathbf{v}_n = -\mu_n \mathbf{E}$$

$$\mathbf{v}_p = \mu_p \mathbf{E}$$

e il prodotto della carica per la concentrazione di portatori per tale velocita' ci dara' la (densita' di) corrente $J_{ndrift} = -qn v_n$ e $J_{pdrift} = qp v_p$.

Queste relazioni valgono per **campi abbastanza piccoli** da far si' che la componente della velocita' dovuta all'accelerazione del campo elettrico sia piccola rispetto alla velocita' termica e il tempo di collisione medio sia indipendente dal campo elettrico. Per **campi molto alti**, ove le energie dei portatori diventano molto alte rispetto all'energia termica, si osserva una deviazione importante rispetto alla linearita', finche' le velocita' di deriva diventano indipendenti dal campo elettrico a un valore di saturazione (Fig. 29).

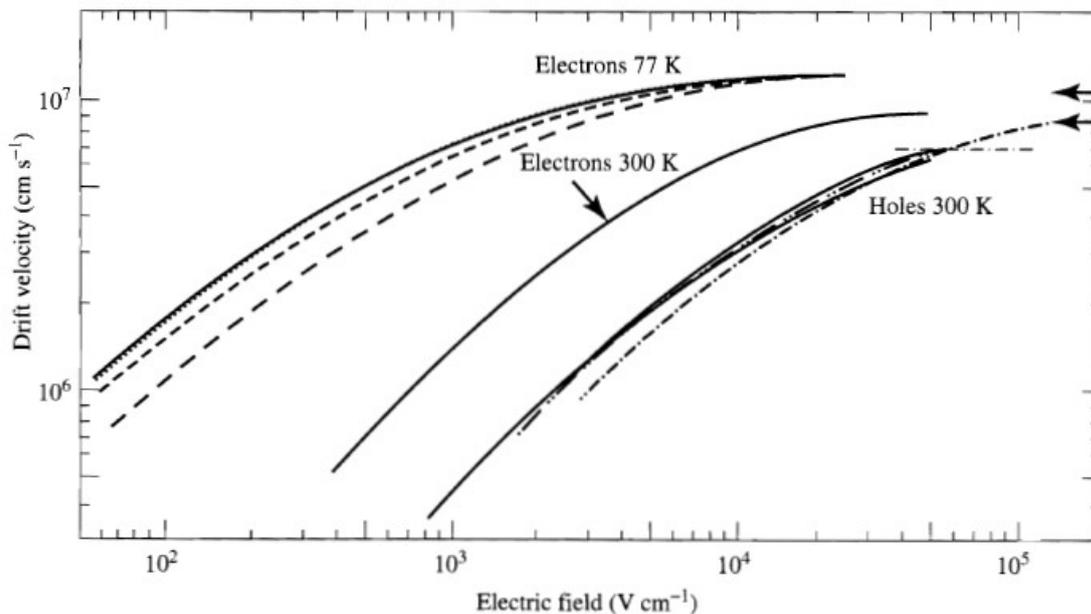


Figura 29: Andamento della velocita' di deriva in funzione del campo elettrico applicato. Si noti la saturazione della velocita' per campi piu' intensi di 10^3 V/cm per gli elettroni, e di 10^4 V/cm per le buche.

Le mobilita' di elettroni e buche μ_n e μ_p sono **dipendenti dalla temperatura e dalla concentrazione di drogaggio**, visto che risentono delle

collisioni con le imperfezioni del reticolo, a causa di vibrazioni termiche, difetti del cristallo e presenza di atomi droganti. A una data temperatura, nella nostra trattazione considereremo la mobilità di elettroni e buche come costanti.

Precisiamo che le distorsioni del reticolo, il quale, se perfettamente periodico e in stato stazionario, non dovrebbe diffondere i portatori, sono dovuti a più meccanismi concorrenti. Il tempo di cammino libero medio, e la mobilità, che a loro volta dipendono direttamente da esso, sono dominate dal fenomeno con tempo di cammino libero più piccolo.

Possiamo riassumere i meccanismi di diffusione così:

- **Vibrazioni del reticolo:** a temperature maggiori dello zero assoluto il reticolo vibra, producendo delle distorsioni nel campo visto dagli elettroni, che si può rappresentare con interazioni con pseudo-particelle a energia quantizzata dette *fononi* (0.063 eV di energia minima scambiata). Se questo meccanismo prevale, la mobilità ha una dipendenza del tipo $\approx T^{-n}$ dalla temperatura, con $1.5 < n < 2.5$
- **Atomi droganti, impurità e imperfezioni del cristallo:** queste entità diventano centri di diffusione che modificano il campo percepito dagli elettroni. Se questo meccanismo prevale, la mobilità ha una dipendenza lineare dalla temperatura $\approx T$. Per alte temperature questo meccanismo diventa meno importante in quanto gli e- riescono ad essere abbastanza veloci da non percepire la loro presenza.

La **dipendenza della mobilità dalla temperatura** ha dunque un massimo, che coincide con la temperatura alla quale il meccanismo di vibrazione del reticolo prende il sopravvento sul meccanismo di diffusione dai centri droganti (Fig. 30).

La **dipendenza della mobilità dal drogaggio** mette in evidenza come essa sia sensibile alla concentrazione di drogaggio solo per $N > 10^{15}$, ove il meccanismo dei centri di diffusione comincia a prevalere, ed è costante al di sotto di tale valore (Fig. 31). Per queste ragioni la mobilità ha un massimo in funzione della temperatura, e invece tende a saturare per campi molto intensi.

2.5.2 La resistività del semiconduttore

Voglio ora trovare una relazione fra la **mobilità** μ dei portatori di carica, e la **resistività** ρ del materiale.

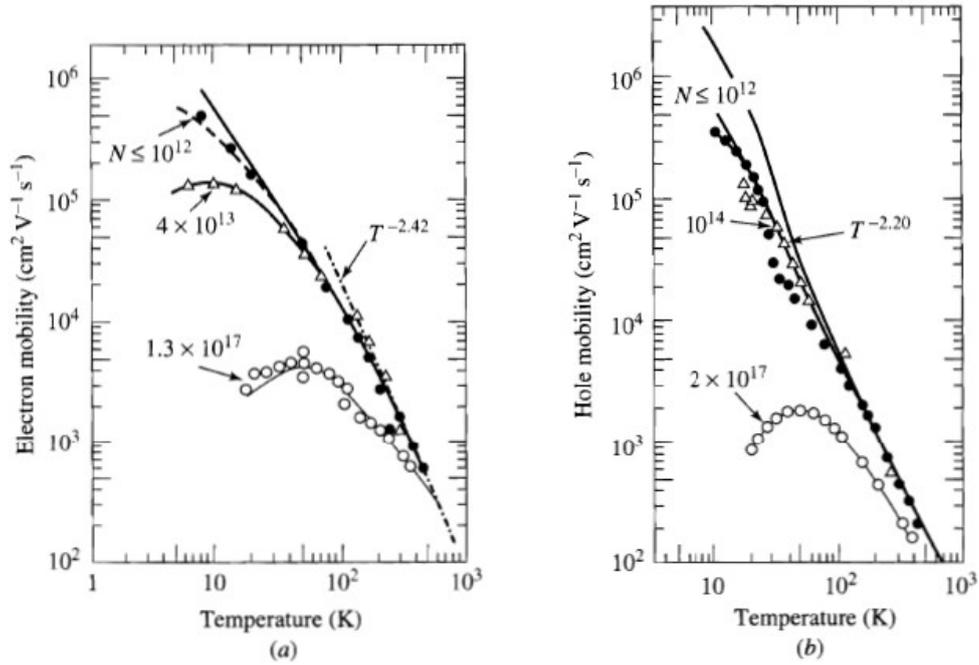


Figura 30: Andamento della mobilita' in funzione della temperatura per elettroni e buche con varie concentrazioni di drogaggio. Si noti come il prevalere di un determinato meccanismo a temperature diverse faccia prevalere un andamento crescente o decrescente, che risulta in un punto di massimo. Si noti anche come per concentrazioni piu' elevate tale punto di massimo si sposti a temperature piu' elevate.

Se considero una densità volumetrica n di portatori, ciascuno con carica q , che si muovono a velocità v nel materiale per deriva in un campo elettrico \mathbf{E} , posso scrivere la **densità di corrente** come $J = qnv$, e per la definizione di mobilita', possiamo esprimere la velocità dei portatori come $v = \mu E$.

Per **elettroni** e **buche**, avremo, rispettivamente:

$$J_n = qn_n v_n = qn_n \mu_n E \quad \text{e} \quad J_p = qn_p v_p = qn_p \mu_p E \quad (23)$$

La **densità di corrente totale** sarà data dalla somma dei contributi dovuti ad elettroni e buche. Nel caso di un **semiconduttore intrinseco**, abbiamo visto che le concentrazioni dei due tipi di portatori si equivalgono ($n_i = n_n = n_p$), possiamo quindi scrivere:

$$J = J_n + J_p = q(n_n \mu_n + n_p \mu_p) E = qn_i (\mu_n + \mu_p) E \quad (24)$$

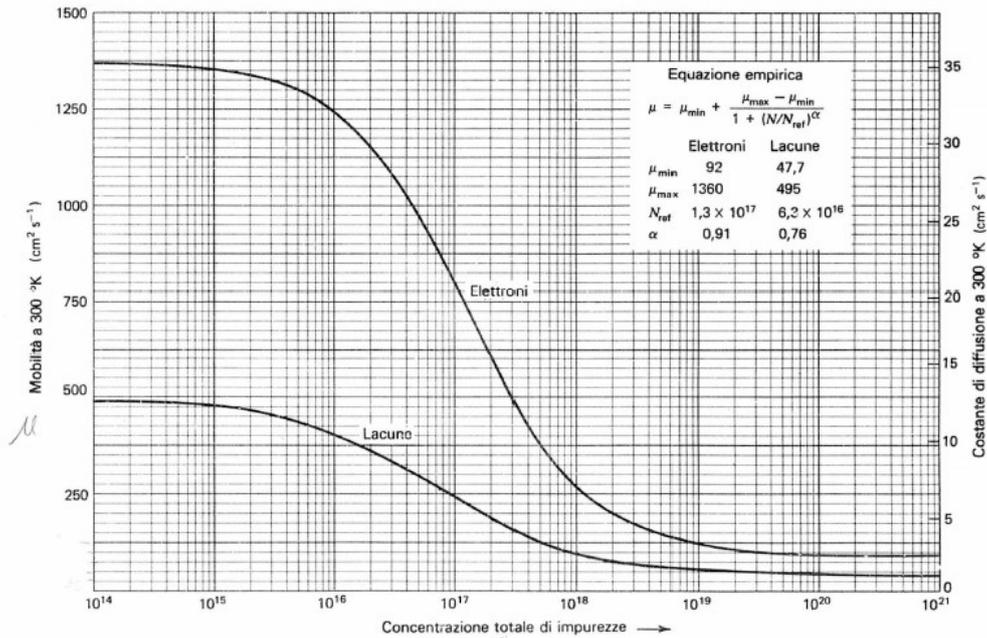


Figura 31: Mobilita' nel silicio a 300 K per elettroni e lacune. Si noti come a concentrazioni di drogaggio $N < 10^{15}$ la mobilita' non varia col drogaggio, in quanto prevalgono meccanismi di perdita di energia per interazione con le vibrazioni del reticolo. A concentrazioni piu' alte, prevale l'interazione con i centri di diffusione di drogaggio, e quindi la mobilita' scende al crescere della concentrazione.

Sappiamo anche che densita' di corrente in un conduttore e campo elettrico sono legati dalla relazione:

$$E = \frac{J}{\sigma} \quad (25)$$

ove σ e' la conduttivita', reciproco della resistivita' ρ .

Dunque:

$$J = \frac{1}{\rho} E = qn_i(\mu_n + \mu_p)E \quad (26)$$

da cui deriva

$$\rho = \frac{1}{qn_i(\mu_n + \mu_p)} \quad (27)$$

Osserviamo che la resistivita' aumenta al diminuire sia della concentrazione dei portatori, che della loro mobilita', come potevamo aspettarci.

La resistività è un parametro molto importante nei rivelatori: determinate tecnologie sono state storicamente sviluppate per materiali a bassa resistività; tali tecnologie non sono compatibili con nuove applicazioni per cui diventano indispensabili semiconduttori ad alte resistività, e potrebbe essere necessario uno sviluppo dedicato per adattarle alle nuove esigenze.

Usando la formula 27, possiamo ricavare la resistività del **silicio** e del **germanio**, entrambi intrinseci, a temperatura ambiente. Riassumiamo in tabella le variabili d'interesse: Notiamo che si tratta di valori di resistività

Semiconduttore	n_i [cm^{-3}]	μ_e [$\text{cm}^2(\text{Vs})^{-1}$]	μ_h [$\text{cm}^2(\text{Vs})^{-1}$]	ρ [$\Omega \text{ cm}$]
silicio (Si)	$1.4 \cdot 10^{10}$	1350	450	$2.5 \cdot 10^5$
germanio (Ge)	$2.4 \cdot 10^{13}$	3900	1900	45

ideale, in particolare abbiamo **trascurato** la presenza di **impurezze** nel cristallo ed **effetti di bordo**.

Esempio. Consideriamo un sensore al silicio con spessore $d \simeq 300 \mu\text{m}$ ed alimentato ad una tensione $V \simeq 30 \text{ V}$. Calcoliamo la velocità di deriva degli elettroni:

$$\begin{aligned} v_e = \mu_e E &= \mu_e \frac{V}{d} \simeq 1350 \frac{\text{cm}^2}{\text{Vs}} \frac{30}{3.00 \cdot 10^{-2}} \frac{\text{V}}{\text{cm}} = \\ &= 1350 \frac{\text{cm}^2}{\text{Vs}} 1.0 \cdot 10^3 \frac{\text{V}}{\text{cm}} = 1.4 \cdot 10^6 \frac{\text{cm}}{\text{s}} \end{aligned}$$

Il tempo che essi impiegheranno per percorrere tutta la giunzione sarà:

$$t_e = \frac{d}{v_e} \simeq \frac{3.00 \cdot 10^{-2} \text{ cm}}{1.4 \cdot 10^6 \frac{\text{cm}}{\text{s}}} = 2.1 \cdot 10^{-8} \text{ s} = 21 \text{ ns}$$

Sostanzialmente, i **tempi di deriva** dei portatori di carica sono dell'ordine dei **ns**, in particolare avremo:

- $t_e \simeq 20 \text{ ns}$ per gli elettroni;
- $t_h \simeq 60 \text{ ns}$ per le buche.

Questi saranno **dominanti** rispetto al **tempo necessario a produrre le ionizzazioni**, che è dell'ordine dei **ps**.

2.5.3 Diffusione

Ove esiste una distribuzione inhomogenea di portatori liberi di carica in un semiconduttore, trascurando ora tutti gli effetti del campo elettrico, le cariche mostreranno uno spostamento netto dalla regione con distribuzione piu' alta a quella con distribuzione piu' bassa. Nonostante ogni singola carica abbia spostamento medio nullo in assenza di campo elettrico, la probabilita' del movimento in questo senso e' piu' alta di quella in senso opposto, visto che ci sono piu' cariche che ne hanno possibilita'. Questo effetto e' chiamato **diffusione** (*diffusion*) e porta a uno smussamento della distribuzione di carica. Matematicamente la diffusione e' descritta dal prodotto di una costante di diffusione D_n, D_p per il gradiente di concentrazione $\nabla n, \nabla p$, che danno luogo a un flusso di cariche:

$$\mathbf{F}_n = -D_n \nabla n$$

$$\mathbf{F}_p = -D_p \nabla p.$$

Il flusso di portatori moltiplicato per la carica del portatore e' una corrente: $J_{n_{diff}} = -qF_n$ e $J_{p_{diff}} = qF_p$.

Questo contributo alla corrente non e' normalmente importante nei metalli, a causa dell'alta conducibilita'. Per i semiconduttori, invece, vista la debole conducibilita' e la disuniformita' di densita' di portatori e delle loro energie, essa e' determinante ai fini della conduzione.

Se ci mettiamo in un sistema semplificato in cui la concentrazione di portatori varia solo lungo una coordinata, per esempio x , e assumiamo che non ci sia ne' campo elettrico applicato ne' disuniformita' nella temperatura e quindi nell'energia media degli e-, l'unica variabile e' $n(x)$. Il movimento termico e' casuale. Se prendiamo un piano di riferimento a $x=0$, per intervalli di tempo pari al tempo medio di dispersione τ , gli e- che lo attraversano sono divisi da una coordinata $x = -l$ con l cammino libero medio, dato da $l = v_{th}\tau_{cn}$, sono: $\frac{1}{2}n(-l)v_{th}$ perche' se solo x e' la direzione di movimento casuale, meta' si saranno mossi avanti e meta' indietro. Allo stesso modo, da destra verso sinistra lo avranno attraversato in $\frac{1}{2}n(l)v_{th}$ creando un flusso risultante $F = \frac{1}{2}v_{th}[n(-l) - n(l)]$ ovvero, prendendo lo sviluppo in serie di Taylor, si riduce a $F = -v_{th}l\frac{dn}{dx}$. Se moltiplichiamo per la carica negativa, avro' che

$$J_n = -qF = qlv_{th}\frac{dn}{dx}$$

La corrente di diffusione e' proporzionale alla derivata spaziale della densita' di elettroni e sorge a causa di un gradiente di concentrazione. Per una

densita' di e- che aumenta con x , il gradiente e' positivo, come la densita' di corrente. Gli e- fluiranno in realta' dalla zona a piu' alta concentrazione verso quella a minore concentrazione, e visto che per convenzione la corrente e' nel verso opposto a quello degli e-, il segno piu' nella densita' di corrente risulta fisicamente corretto.

Applicando il teorema di equipartizione dell'energia in una dimensione, $\frac{1}{2}m_n^*v_t^2 = \frac{1}{2}kT$, utilizzando la definizione di mobilita'

$$\mu_n = \frac{q\tau_{cn}}{m_n^*}$$

e di cammino libero medio

$$l = v_t\tau_{cn}$$

otteniamo dunque:

$$J_n = q \left(\frac{kT}{q} \mu_n \right) \frac{dn}{dx} = q D_n \frac{dn}{dx}$$

ove ho definito questa costante di Diffusione:

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

che si vede proporzionale alla mobilita'. Tale relazione, derivata in questo contesto in maniera intuitiva, e' detta **relazione di Einstein**, mette in relazione le due quantita' importanti dei moti di deriva e di diffusione, e puo' essere derivata rigorosamente dalla meccanica statistica dei solidi.

Se aggiungessimo un campo elettrico E_x in direzione x , la densita' di corrente totale diventa:

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

ove sia la corrente di deriva e' diretta in senso positivo, in quanto movimento negativo di cariche negative, sia quella di diffusione, visto che rappresenta il movimento da zona ad alta concentrazione verso bassa concentrazione e cariche negative.

Per le lacune il discorso e' il medesimo, con la relazione di Einstein ancora valida:

$$D_p = \frac{kT}{q} \mu_p$$

e la densità di corrente:

$$J_p = q\mu_p p E_x - qD_p \frac{dp}{dx}$$

Qui la corrente di deriva di cariche positive nel senso del campo elettrico dà segno positivo, e la corrente di diffusione è in segno contrario contando la carica positiva opposta al verso del moto di diffusione per gradiente che sempre aumenta con x . Invece del gradiente di concentrazione, si può avere lo stesso effetto con concentrazione costante ma con portatori più energetici da una parte rispetto a un'altra, ovvero in presenza di un gradiente di temperatura.

3 Giunzione pn

Quando abbiamo i due tipi di semiconduttore conviventi nella stessa struttura (la *giunzione pn*), ci riferiamo generalmente alla concentrazione di accettori nel tipo P con N_A , e alla concentrazione effettiva di donori nel tipo N con N_D , visto che esse ne costituiscono le rispettive concentrazioni effettive. Per raggiungere l'equilibrio del sistema, le bande di energia si modificano in modo che il livello di Fermi del materiale drogato rimanga costante, mentre i limiti delle bande di valenza e conduzione si distorcono per mantenere le rispettive strutture invariate. Questo "dislivello" che si forma tra E_{C_P} e E_{C_N} equivale a una barriera di potenziale che gli e^- della BC del tipo N devono superare per potersi spostare nella BC del tipo P. Come sappiamo, la concentrazione dei portatori dipende dall'energia di Fermi dei materiali secondo le equazioni: È conveniente esprimere tali equazioni in termini dell'energia di Fermi del materiale intrinseco E_i , in modo da ritrovare un'equazione a noi nota, ovvero quella che esprime la concentrazione dei portatori di carica del materiale intrinseco in funzione dell'energia di Fermi, e tenere conto esplicitamente solo dello spostamento dell'energia di Fermi effettiva del materiale drogato rispetto a tale riferimento teorico. Tale dislivello in energia (diviso per la carica elementare) è anche chiamato potenziale di contatto o *built-in*, e viene indicato con V_{bi} . E' possibile ricavare la formula che descrive il potenziale elettrico lungo il materiale grazie all'equazione di Poisson, secondo cui la derivata seconda del potenziale equivale alla densità di carica fratto la permittività elettrica del materiale.

3.1 Distribuzione di carica, campo elettrico, potenziale elettrico in una giunzione pn generica

Consideriamo dunque una densità di carica caratteristica della *giunzione brusca (abrupt junction)*, ovvero in cui la densità di carica spaziale cambia secondo un profilo a scalino da nulla al suo massimo, sia per la zona svuotata in N sia in P.

Nel primo grafico dall'alto, in fig. 32, è riportata la **densità di carica** nel materiale, definita come:

$$\rho(x) = \begin{cases} -qN_p, & \text{se } -x_p \leq x \leq 0 \\ qN_n, & \text{se } 0 \leq x \leq x_n \\ 0, & \text{se } x < -x_p \text{ o } x > x_n \end{cases} \quad (28)$$

dove: q è la carica elementare, N_n ed N_p sono, rispettivamente, le **concentrazioni dei donori e degli accettori**, mentre x_n ed x_p sono le **profondità della zona di svuotamento** in direzione x , a partire dalla coordinata nulla (i.e. punto di giunzione ideale), verso la parte drogata n e verso quella drogata p , rispettivamente.

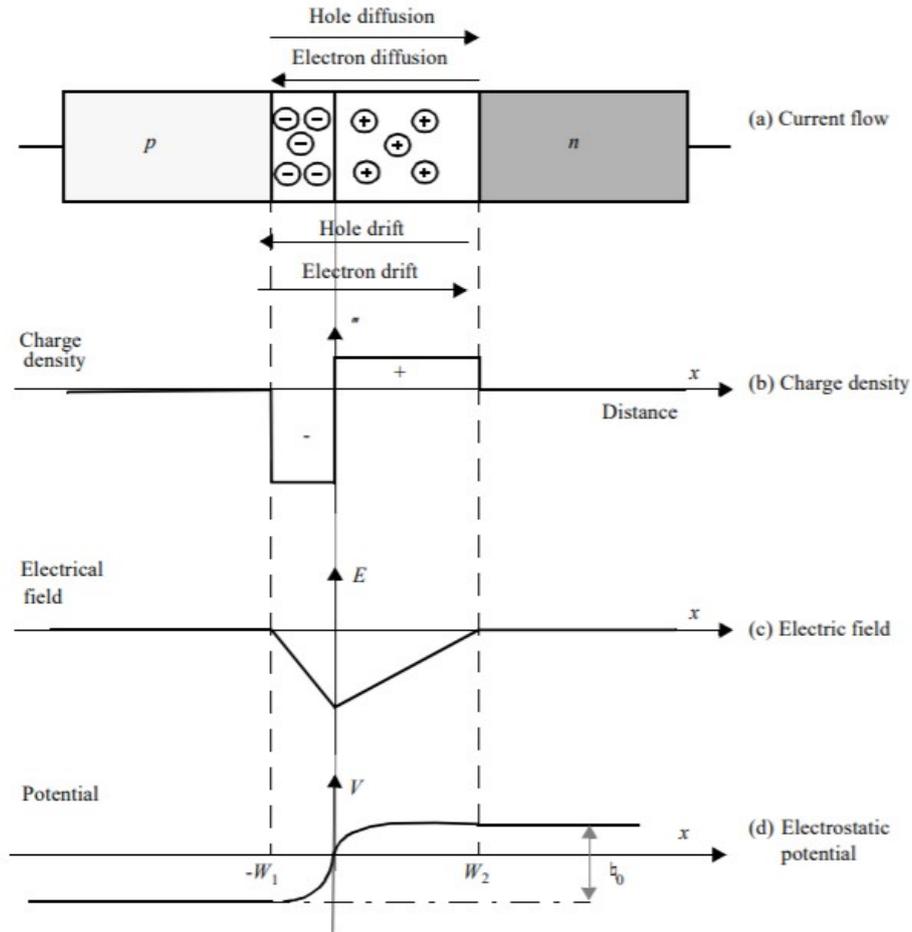


Figura 32: Forma della distribuzione di carica, del campo elettrico e del potenziale in una giunzione pn brusca.

Il **campo elettrico built-in** $E(x)$ è riportato nel grafico centrale di fig. 32. Possiamo ricavarlo integrando la prima equazione di Maxwell:

$$\frac{dE}{dx} = \frac{\rho(x)}{\epsilon} \quad \implies \quad E(x) = \int_{x_0}^x \frac{\rho(x')}{\epsilon} dx' + E(x_0) \quad (29)$$

Nelle formule seguenti, X_0 come termine di integrazione, rappresenta il punto in cui la funzione integrale è nota, quindi la condizione al contorno, e può assumere qualunque valore rispetto all'intervallo di integrazione.

Utilizzando la densità come definita in formula 28, otteniamo:

$$\begin{aligned} E_p(x) &= \frac{1}{\varepsilon} \int_{-x_p}^x (-q)N_p dx' = -\frac{qN_p}{\varepsilon}(x + x_p) \quad \text{per} \quad -x_p \leq x \leq 0 \\ E_n(x) &= \frac{1}{\varepsilon} \int_{x_n}^x qN_n dx' = \frac{qN_n}{\varepsilon}(x - x_n) \quad \text{per} \quad 0 \leq x \leq x_n \end{aligned} \quad (30)$$

in cui abbiamo tenuto conto della continuità del campo E ai confini della regione svuotata, oltre i quali il campo è nullo: $E(-x_p) = E(x_n) = 0$.

All'interfaccia ($x = 0$) vale la condizione di raccordo $E_p(x = 0) = E_n(x = 0)$. Eguagliando nell'origine le relazioni nella formula 30, ottengo:

$$-\frac{qN_p}{\varepsilon}x_p = \frac{qN_n}{\varepsilon}(-x_n) \quad \implies \quad N_px_p = N_nx_n \quad (31)$$

Questo ci dice che l'area sottesa al grafico della densità di carica è la stessa sia nella parte p che in quella n . Ciò è sensato, essendo uguale la quantità di elettroni liberi e di buche che si ricombinano nella zona di svuotamento. Una delle più usate configurazioni di giunzione pn è quella unipolare, ovvero con un drogaggio fortemente sbilanciato tra p ed n tale da ridurre il volume svuotato nella parte più drogata a uno strato superficiale, e invece estendere il volume svuotato nella parte meno drogata a tutto il volume disponibile.

Definendo lo **spessore di svuotamento** $W = x_n + x_p$, posso riscrivere la relazione 31 come:

$$\begin{aligned} N_px_p &= N_n(W - x_p) \quad \implies \quad (N_p + N_n)x_p = N_nW \\ N_nx_n &= N_p(W - x_n) \quad \implies \quad (N_p + N_n)x_n = N_pW \end{aligned} \quad (32)$$

Ottenendo dunque gli spessori delle due singole zone svuotate in funzione di W :

$$x_p = \frac{N_n}{N_p + N_n}W \quad \text{e} \quad x_n = \frac{N_p}{N_p + N_n}W \quad (33)$$

Il **potenziale elettrico** $\phi(x)$ è riportato nel grafico in basso di fig. 32. Possiamo ricavarlo integrando dalla definizione:

$$E(x) = -\frac{d\phi(x)}{dx} \quad \implies \quad \phi(x) = -\int_{x_0}^x E(x')dx' + \phi(x_0) \quad (34)$$

Usando le espressioni del campo elettrico ricavate nella formula 30, otteniamo:

$$\begin{aligned}
\phi_p(x) &= \frac{qN_p}{\varepsilon} \int_{-x_p}^x (x' + x_p) dx' + \phi_p(-x_p) = \\
&= \frac{qN_p}{\varepsilon} \left[\left(\int_{-x_p}^x x' dx' \right) + x_p \left(\int_{-x_p}^x dx' \right) \right] + \phi_p(-x_p) = \\
&= \frac{qN_p}{\varepsilon} \left[\frac{1}{2} (x^2 - (-x_p)^2) + x_p (x - (-x_p)) \right] + \phi_p(-x_p) = \quad (35) \\
&= \frac{qN_p}{2\varepsilon} [x^2 - x_p^2 + 2xx_p + 2x_p^2] + \phi_p(-x_p) = \\
&= \frac{qN_p}{2\varepsilon} (x^2 + 2xx_p + x_p^2) + \phi_p(-x_p) = \\
&= \frac{qN_p}{2\varepsilon} (x + x_p)^2 + \phi_p(-x_p)
\end{aligned}$$

$$\begin{aligned}
\phi_n(x) &= -\frac{qN_n}{\varepsilon} \int_{x_n}^x (x' - x_n) dx' + \phi_n(x_n) = \\
&= -\frac{qN_n}{\varepsilon} \left[\left(\int_{x_n}^x x' dx' \right) - x_n \left(\int_{x_n}^x dx' \right) \right] + \phi_n(x_n) = \\
&= -\frac{qN_n}{\varepsilon} \left[\frac{1}{2} (x^2 - x_n^2) - x_n (x - x_n) \right] + \phi_n(x_n) = \quad (36) \\
&= -\frac{qN_n}{2\varepsilon} [x^2 - x_n^2 - 2xx_n + 2x_n^2] + \phi_n(x_n) = \\
&= -\frac{qN_n}{2\varepsilon} (x^2 - 2xx_n + x_n^2) + \phi_n(x_n) = \\
&= -\frac{qN_n}{2\varepsilon} (x - x_n)^2 + \phi_n(x_n)
\end{aligned}$$

Scegliendo come riferimento il valore $\phi_p(-x_p) = 0$ e fissando il valore massimo del potenziale $\phi_n(x_n) = V_0$, otteniamo:

$$\begin{aligned}
\phi_p(x) &= \frac{qN_p}{2\varepsilon} (x + x_p)^2 & \text{per} & \quad -x_p \leq x \leq 0 \\
\phi_n(x) &= V_0 - \frac{qN_n}{2\varepsilon} (x - x_n)^2 & \text{per} & \quad 0 \leq x \leq x_n
\end{aligned} \quad (37)$$

Considerando anche la condizione di raccordo $\phi_p(x = 0) = \phi_n(x = 0)$, e confrontandolo con l'andamento visto per l'energia delle bande della giunzione pn, osserviamo che il gradiente di potenziale all'interno della zona di svuotamento favorisce il moto di deriva delle cariche libere, cosa che non accade nelle zone esterne, ove il gradiente di potenziale e' nullo.

Vediamo la **variazione massima di potenziale** V_0 , ossia la **tensione built-in**. Possiamo ricavarla imponendo la condizione di raccordo $\phi_p(x = 0) = \phi_n(x = 0)$, da cui segue che:

$$\frac{qN_p}{2\varepsilon}x_p^2 = V_0 - \frac{qN_n}{2\varepsilon}x_n^2 \quad \Longrightarrow \quad V_0 = \frac{q}{2\varepsilon}(N_px_p^2 + N_nx_n^2) \quad (38)$$

Utilizzando la relazione 31, è possibile eliminare, in alternativa x_n od x_p , ottenendo le due espressioni equivalenti:

$$\begin{aligned} V_0 &= \frac{q}{2\varepsilon} \left(N_px_p^2 + N_n \frac{N_p^2}{N_n^2} x_p^2 \right) = \frac{q}{2\varepsilon} N_px_p^2 \left(1 + \frac{N_p}{N_n} \right) = \frac{q}{2\varepsilon} N_px_p^2 \frac{N_p + N_n}{N_n} \\ V_0 &= \frac{q}{2\varepsilon} \left(N_p \frac{N_n^2}{N_p^2} x_n^2 + N_n x_n^2 \right) = \frac{q}{2\varepsilon} N_n x_n^2 \left(\frac{N_n}{N_p} + 1 \right) = \frac{q}{2\varepsilon} N_n x_n^2 \frac{N_p + N_n}{N_p} \end{aligned} \quad (39)$$

Avere una stima di tale grandezza può essere utile in quanto una sua **misura diretta**, ad esempio inserendo degli elettrodi all'interno della zona svuotata, andrebbe a **perturbare il potenziale** creato.

Invertendo le relazioni precedenti (formula 39) possiamo esprimere gli spessori di svuotamento delle due zone come:

$$x_p = \sqrt{\frac{2\varepsilon V_0}{q} \frac{N_n}{N_p(N_p + N_n)}} \quad \text{e} \quad x_n = \sqrt{\frac{2\varepsilon V_0}{q} \frac{N_p}{N_n(N_p + N_n)}} \quad (40)$$

Per cui lo spessore di svuotamento complessivo sarà:

$$\begin{aligned}
W = x_p + x_n &= \sqrt{\frac{2\varepsilon V_0}{q}} \frac{1}{\sqrt{N_p + N_n}} \left(\sqrt{\frac{N_n^2}{N_p N_n}} + \sqrt{\frac{N_p^2}{N_p N_n}} \right) = \\
&= \sqrt{\frac{2\varepsilon V_0}{q}} \frac{1}{\sqrt{N_p + N_n}} \left(\frac{N_p + N_n}{\sqrt{N_p N_n}} \right) = \sqrt{\frac{2\varepsilon V_0}{q}} \sqrt{\frac{N_p + N_n}{N_p N_n}} = \\
&= \sqrt{\frac{2\varepsilon V_0}{q} \left(\frac{1}{N_p} + \frac{1}{N_n} \right)} \quad (41)
\end{aligned}$$

Vediamo inoltre qual è il **modulo massimo del campo elettrico**, che si ha all'interfaccia della giunzione pn , ovvero $E_{MAX} = E(x = 0)$. Scegliendo di utilizzare la definizione del campo elettrico $E_n(x)$ (seconda relazione nella formula 30), unitamente all'espressione dello spessore di svuotamento x_n (seconda relazione nella formula 40), otteniamo:

$$\begin{aligned}
E_{MAX} = |E(0)| &= \frac{qN_n}{\varepsilon} x_n = \frac{qN_n}{\varepsilon} \sqrt{\frac{2\varepsilon V_0}{q} \frac{N_p}{N_n(N_p + N_n)}} = \sqrt{\frac{q^2 N_n^2 2\varepsilon V_0}{\varepsilon^2} \frac{N_p}{N_n(N_p + N_n)}} = \\
&= \sqrt{\frac{2qV_0}{\varepsilon} \frac{N_p N_n}{(N_p + N_n)}} = \sqrt{\frac{2qV_0}{\varepsilon} \left(\frac{1}{N_p} + \frac{1}{N_n} \right)^{-1} \frac{2V_0}{2V_0}} \\
&= \sqrt{\left[\frac{q}{2\varepsilon V_0} \left(\frac{1}{N_p} + \frac{1}{N_n} \right)^{-1} \right] 4V_0^2} = \frac{2V_0}{W} \quad (42)
\end{aligned}$$

In una giunzione pn non polarizzata c'è equilibrio tra la corrente di deriva dei portatori minoritari generati o iniettati nella zona di svuotamento (con e^- che si muovono guidati dal campo elettrico da n a p , le h^+ da p a n), e la corrente di diffusione, quindi per ricombinazione, dei portatori maggioritari (con h^+ che si muovono da n a p , e gli e^- da p a n).

Dopo aver ricavato Potenziale e Campo Elettrico a partire dalla densità di carica, passiamo a considerare i potenziali che agiscono su una carica q nel passaggio tra E_{fn} e E_i e' $\phi_n = q(E_{fn} - E_i)$ e quello tra E_{fp} e E_i e' $\phi_p = q(E_i - E_{fp})$. Come abbiamo visto graficamente e ricavato nel paragrafo precedente, c'è relazione tra i potenziali creati in ogni materiale drogato e

la densita' di drogaggio,

$$V_{bi} = \phi_p + \phi_n = \frac{kT}{q} \ln(n_n p_p / n_i^2) = \frac{kT}{q} \ln(N_D N_A / n_i^2)$$

3.1.1 Correnti nella giunzione pn non polarizzata

Come detto, mettendo insieme gli effetti di deriva e diffusione sulle cariche otteniamo le densita' di corrente per i due portatori si ottiene:

$$\mathbf{J}_n = q\mu_n n E + qD_n \nabla n$$

$$\mathbf{J}_p = q\mu_p p E - qD_p \nabla p$$

Per comprendere che cosa sia la costante di diffusione possiamo considerare le formule precedenti in una dimensione e porre la condizione al contorno che il sistema sia all'equilibrio, ovvero che la densita' di corrente totale di cariche negative sia zero $J_n = 0$:

$$\mu_n \frac{dV}{dx} = -D_n \frac{dn}{dx}$$

e quindi

$$\frac{1}{n} \frac{dn}{dx} = \frac{\mu_n}{D_n} \frac{dV}{dx}$$

con V potenziale associato al campo elettrico di deriva. Prese due zone del semiconduttore corrispondenti alle posizioni x_1 e x_2 con due concentrazioni diverse n_1 e n_2 possiamo integrare sulla posizione:

$$\ln \left(\frac{n_2}{n_1} \right) = -\frac{\mu_n}{D_n} (V_2 - V_1)$$

e quindi

$$\frac{n_2}{n_1} = -e^{\frac{\mu_n}{D_n} (V_2 - V_1)} \quad (43)$$

D'altra parte, avevamo usato la distribuzione statistica di Boltzmann per definire la concentrazione di portatori in funzione del livello energetico di Fermi del materiale drogato:

$$n = n_i e^{\frac{E_F - E_i}{kT}}.$$

Possiamo dunque scrivere per due concentrazioni diverse:

$$\frac{n_2}{n_1} = \frac{n_i e^{\frac{E_{F_2} - E_i}{kT}}}{n_i e^{\frac{E_{F_1} - E_i}{kT}}} = e^{\frac{E_{F_2} - E_{F_1}}{kT}} = e^{q \frac{V_2 - V_1}{kT}} \quad (44)$$

ove V e' il potenziale corrispondente alla differenza tra energia di Fermi alla concentrazione considerata e energia di Fermi intrinseca. Uguagliando dunque le equazioni 43 e 44 otteniamo l'equazione di Einstein:

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

da cui si ricavano le definizioni dei coefficienti di diffusione:

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

$$D_p = \frac{\mu_p}{q} kT$$

3.1.2 Correnti nella giunzione pn direttamente polarizzata

In una giunzione pn direttamente polarizzata, la zona di svuotamento si assottiglia e il campo elettrico di deriva decresce, portando a una diminuzione di corrente di deriva. Il gradiente di concentrazione invece rimane invariato, quindi la corrente di diffusione supera quella di deriva portando il bilancio a una corrente netta. Tale corrente porta a un eccesso di elettroni nella regione p e a un eccesso di lacune nella regione n , in quanto non c'è più una barriera di potenziale a controbilanciare il fenomeno. Localmente dunque si ha una deviazione dall'equilibrio, per cui $pn > n_i^2$. L'equilibrio viene presto ripristinato dalla ricombinazione.

Per le buche che fluiscono nella regione n , esse passeranno il sottile volume svuotato con poche perdite dovute alla ricombinazione, dato che la concentrazione di elettroni e' piccola rispetto al bulk (e' svuotato dai portatori). Arrivato all'interfaccia tra regione svuotata e non svuotata in n , la concentrazione di e^- a disposizione per ricombinarsi aumenta e quella di lacune decresce con la distanza, in funzione della probabilita' di ricombinazione, espressa come lunghezza di diffusione. Alla fine, tutte le lacune si ricombinano con gli elettroni, che sono forniti in quantita' illimitata dall'elettrodo connesso all'alimentatore. Sul lato p le buche verranno fornite dall'elettrodo per sostenere un meccanismo analogo.

La distribuzione delle cariche negative nella regione p in stato stazionario si ricava risolvendo l'equazione di diffusione:

$$D_n \frac{d^2 n_p}{dx^2} - \frac{n_p - n_{p0}}{\tau_{cn}} = 0 \quad (45)$$

dove:

- D_n è la costante di diffusione degli elettroni,
- n_p è la concentrazione di elettroni nella regione *p-type*,
- n_{p0} è la stessa concentrazione ma all'equilibrio,
- τ_{cn} è il tempo di ricombinazione degli elettroni corrispondente ad una lunghezza di diffusione,
- lunghezza di diffusione $L_n = \sqrt{D_n \tau_{cn}}$.

Ponendo condizioni al contorno per cui la concentrazione di e^- in eccesso svanisce a grandi distanze $n_p(\infty) = n_{p0}$, e per cui i portatori iniettati all'origine della regione svuotata $x = 0$ con concentrazione $n_p(0)$, si ottiene la soluzione:

$$n_p(x) = n_{p0} + (n_p(0) - n_{p0}) e^{-x/L_n}.$$

L'equazione di continuita' afferma che l'incremento del numero di portatori di carica di un determinato tipo in un qualunque piccola regione del semiconduttore e' dato dalla differenza tra generazione e ricombinazione nel volume e flusso entrante attraverso la superficie del volume. Con opportune assunzioni, e usando i livelli di Fermi per i semiconduttori drogati e la relazione di Einstein, si possono connettere le zone neutra e svuotata di diffusione e di svuotamento e uguagliare le correnti entranti e uscenti, ottenendo le equazioni per le densita' di corrente:

$$\begin{cases} J_p = \frac{qD_p n_i^2}{N_D L_p} (e^{\frac{qV}{kT}} - 1) \\ J_n = \frac{qD_n n_i^2}{N_A L_n} (e^{\frac{qV}{kT}} - 1) \end{cases} \quad (46)$$

La densità di corrente totale è data dalla somma delle due, ed e' chiamata Equazione di Shockley o *di diodo*:

$$J_{tot} = J_p + J_n = J_0 (e^{\frac{qV}{kT}} - 1) \quad (47)$$

dove

$$J_0 = qn_i^2 \left(\frac{D_n}{N_A L_n} + \frac{D_p}{N_D L_p} \right) \quad (48)$$

e' detta anche corrente di saturazione, o J_S .

Tale equazione ci dice inoltre che:

- la corrente dipende da E_g solo attraverso la concentrazione intrinseca n_i ,
- i contributi degli e^- e delle h^+ non sono uguali,
- J fluisce sempre, per ogni V , ma ha un *ginocchio* determinato da J_0
- materiali anche con E_g diversi hanno correnti uguali se J_0 e' uguale

3.1.3 Correnti nella giunzione pn inversamente polarizzata

Aggiungendo un potenziale esterno dello stesso segno di quella *built-in*, cioè portando il lato N a una tensione ancora maggiore rispetto a P, posso operare la giunzione in polarizzazione inversa: le h^+ in p saranno attratte dall'elettrodo al capo del lato p esponendo ioni negativi, viceversa gli e^- di n saranno attratti dall'elettrodo al capo del lato n . La precedente trattazione delle correnti non aveva imposto limiti su V . Proviamo quindi a considerare la polarizzazione negativa della giunzione, in cui $V < 0$. Per $|V| \gg kT$ in polarizzazione inversa, l'esponenziale tende a zero, e la densità di corrente totale tende a:

$$J_{tot} = -J_0 = -qn_i^2 \left(\frac{D_n}{N_A L_n} + \frac{D_p}{N_D L_p} \right) \quad (49)$$

per esempio, nel caso di una giunzione pn di silicio con concentrazioni $N_A \sim N_D \sim 10^{16} \text{ cm}^{-3}$ a 300 K ottengo J_{tot}

$$= -1.6 \cdot 10^{-19} \text{ C} \cdot (10^{10} \text{ cm}^{-3})^2 \left(\frac{36 \text{ cm}^2/\text{s}}{10^{16} \text{ cm}^{-3} \cdot 0.05 \text{ cm}} + \frac{12 \text{ cm}^2/\text{s}}{10^{16} \text{ cm}^{-3} \cdot 0.03 \text{ cm}} \right) \quad (50)$$

In generale, per concentrazioni di drogaggio qualunque su una giunzione di area $10 \times 10 \text{ cm}^2$ ottengo una corrente:

$$I \simeq 10^{-14} \sim 10^{-18} \text{ A} \quad (51)$$

Tale corrente e' determinata quindi solo dalle concentrazioni di drogaggio $N_{A,D}$, dalle costanti di diffusione $D_{n,p}$ e dalle lunghezze di ricombinazione $L_{n,p}$, e non dipende da V .

Nelle giunzioni reali, pero', misuriamo una corrente di ordini di grandezza piu' grande, e che cresce con lo spessore svuotato, quindi non costante. Ci sono infatti alcuni fenomeni che contribuiscono in maniera non trascurabile

alla corrente, e che dipendono dalla presenza di imperfezioni nel cristallo che aumentano la corrente di *leakage*. Andiamo a considerarli brevemente.

Generazione nella zona di svuotamento Nella zona di svuotamento ci sono centri di cattura e ricombinazione ma non cariche a disposizione per questi processi. Quello che invece e' presente e' la **corrente di generazione** dipendente dai centri emettitori. In assenza di cattura, essi possono solo emettere in maniera alternata h^+ e e^- . Prendiamo dunque la variazione della concentrazione di centri di emissione nel tempo. Essa si puo' ricavare dalla differenza tra tassi di cattura ed emissione, a partire dalle distribuzioni probabilistiche di occupazione degli stati energetici. Risulta dunque un tasso di generazione:

$$\frac{dN_e}{dt} = \frac{n_i}{2\tau_g}$$

ove τ_g e' il tempo caratteristico di generazione. Si puo' dimostrare che solo stati vicini al livello di Fermi intrinseco E_F contribuiscono significativamente al tasso di generazione. Intuitivamente, visto che entrambe le probabilita' di emissione per e^- e h^+ aumentano esponenzialmente rispetto al limite della banda relativa, la probabilita' di un'emissione sequenziale di lacuna ed elettrone e' massima a meta' *gap*.

Per centri di emissione distribuiti uniformemente sullo spessore w svuotato, la corrente di generazione e' dunque:

$$J_g = q \frac{dN_e}{dt} w = \frac{qn_i w}{2\tau_g}$$

$$J_{tot} = J_R (e^{\frac{qV}{kT}} - 1) \quad (52)$$

dove

$$J_R = qn_i^2 \left(\frac{1}{N_A} \sqrt{\frac{D_n}{\tau_{cn}}} + \frac{1}{N_D} \sqrt{\frac{D_p}{\tau_{cp}}} \right) + q \frac{n_i w}{2\tau_g} \quad (53)$$

ovvero:

$$-J = J_R = J_0 + J_g = qn_i^2 \left(\frac{D_n}{N_A L_n} + \frac{D_p}{N_D L_p} \right) + q \frac{n_i W}{2\tau_g} \quad (54)$$

All'aumentare della temperatura, la corrente di diffusione prevale velocemente.

3.2 Capacita' della giunzione

Abbiamo visto che i valori tipici delle concentrazioni di drogaggio sono $10^{12\div 20} \text{ cm}^{-3}$, con un profilo che, rispetto alla coordinata della giunzione, è **idealmente** a gradino (come riportato nel grafico in alto di fig. 32). Nella realtà tale profilo è piu' graduale, variando in modo **continuo**. In prima approssimazione, i profili a gradino rappresentano bene il comportamento delle giunzioni reali comuni.

Ricordiamo i passaggi principali con cui si giunge alla formazione di una giunzione pn:

1. Per **diffusione** le cariche libere di tipo opposto, presenti nei rispettivi lati, tendono a ricombinarsi;
2. Avendo segno opposto, le cariche libere si **neutralizzano**, lasciando **ionizzati** gli atomi, fissi nelle loro posizioni reticolari, a cui erano precedentemente legate. Tale processo inizia dapprima a ridosso della giunzione, per poi coinvolgere atomi man mano più periferici;
3. Si viene così a creare una **zona di svuotamento** in cui troviamo un egual numero di atomi ionizzati negativamente, nel lato p , e positivamente, nel lato n ;
4. Il processo diffusivo prosegue fintantoché il **campo elettrico built-in** che si viene a creare grazie agli atomi ionizzati nella zona di svuotamento, forma una barriera di potenziale che impedisce il passaggio di ulteriori cariche libere.

Ci troviamo in una condizione di **equilibrio termico**, in cui la diffusione è contrastata dal potenziale *built-in* che si è venuto a creare ($V_{bi} \simeq 0.7 \text{ V}$, nel caso del silicio).

La corrente totale che fluisce nella giunzione (\vec{J}_{tot}), è data dal contributo delle componenti di **deriva** e di **diffusione**, ovvero $\vec{J}_{tot} = \vec{J}_{der} + \vec{J}_{diff}$. Nel dettaglio:

- la **corrente di diffusione** (\vec{J}_{diff}) è unicamente dovuta alla presenza di un **gradiente di concentrazione** delle cariche libere, che tendono a muoversi verso la regione a concentrazione piu' bassa, che ospita una concentrazione di segno opposto piu' alta, facilitando la ricombinazione. Anche iniettando cariche da un circuito esterno, tale ricombinazione

mantiene un gradiente di concentrazione che alimenta la diffusione. Per tale motivo tale corrente è detta anche **di ricombinazione** (\vec{J}_{ric});

- la **corrente di deriva** (\vec{J}_{der}) dipende invece solo dalle **cariche che si muovono nel campo elettrico** all'interno della zona svuotata, in gran parte generate in tale zona, e per tale motivo è detta anche **corrente di generazione** (\vec{J}_{gen}).

All'equilibrio termico, osserviamo che la corrente totale è nulla, ovvero $\vec{J}_{tot} = \vec{0}$. Ciò comporta che le due componenti, in generale non-nulle, debbano compensarsi: $\vec{J}_{der} = -\vec{J}_{diff}$.

Se inserisco un generatore di tensione in polarizzazione **diretta** (come nello schema al centro in fig. 34), sto sostanzialmente **introducendo elettroni nel lato n e buche nel lato p**. Per cui fornisco portatori liberi di carica, che, diffondendo nella zona svuotata, ripristinano la neutralità degli ioni da entrambi i lati della giunzione. Di conseguenza il campo elettrico cala in intensità, e **la barriera di potenziale diminuisce**.

La corrente totale non è nulla ($\vec{J}_{tot} \neq \vec{0}$), ed il contributo dominante è dovuto alla componente di **diffusione**. La componente di deriva non influisce apprezzabilmente per due ragioni:

- avendo ridotto il volume della zona svuotata, si è ridotta anche la probabilità di generare coppie di cariche libere;
- si è ridotto anche il potenziale, e dunque il modulo del campo elettrico, ossia ciò che causa del moto delle cariche.

In questo regime, il materiale si comporta da **conduttore**. La corrente che fluisce lungo la giunzione, in funzione del potenziale introdotto, ha un andamento inizialmente lineare (per i primi decimi di Volt), a cui segue una crescita esponenziale, come visibile nel grafico in fig. 33. La tensione applicabile senza effetti dirompenti sulla corrente, e senza quindi che la caduta di potenziale stessa esaurisca l'intera tensione applicata, è limitata dal valore built-in.

Consideriamo le correnti in *forward* e *direct bias* da un punto di vista osservativo. Rispetto alla situazione iniziale di equilibrio termico, il potenziale esterno aumenta la concentrazione di cariche libere in entrambi gli estremi della giunzione. Riprende dunque il moto di diffusione, che spinge le cariche a muoversi verso i rispettivi lati opposti lungo la giunzione, portandole

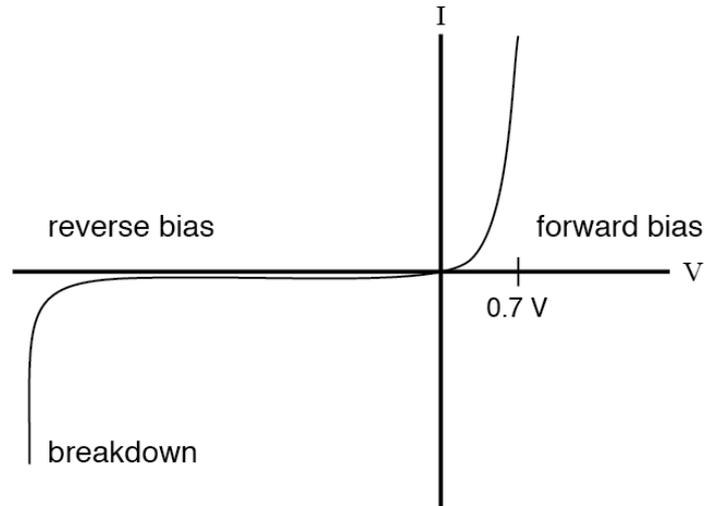


Figura 33: Curva di tensione-corrente per una giunzione pn in silicio. A potenziali positivi (*forward bias*) la giunzione si comporta da conduttore: la corrente cresce molto velocemente in funzione della tensione. A potenziali negativi (*reverse bias*) la giunzione si comporta da isolante: al crescere del potenziale, la corrente rimane ad un piccolo valore costante (corrente di buio, o *dark current*). Se il potenziale aumenta sino al valore di rottura (*breakdown*), si ha nuovamente conducibilità, e la corrente aumenta rapidamente.

a ricombinarsi. A differenza dell'equilibrio termico, in cui le ricombinazioni portavano ad un incremento del volume della zona svuotata, in questo caso il generatore provvede a portare nuove cariche libere ad entrambi gli estremi della giunzione, ripristinandone la concentrazione iniziale, ed il processo si ripete. In tal modo si osserva il passaggio di una corrente nel circuito.

Vediamo cosa succede quando inserisco il generatore di tensione in polarizzazione **inversa**. Per comodità, esprimiamo il potenziale complessivo come la somma quello fornito dal generatore (V_R) e quello intrinseco della giunzione (V_{bi}):

$$V = V_R + V_{bi} \quad (55)$$

Come si può osservare in fig. 34, **la barriera di potenziale cresce**, e quindi aumenta anche il modulo del campo elettrico presente nella zona svuotata. Questo è proporzionale all'entità delle cariche che lo generano (Q) all'interno della zona svuotata. Tale quantità dipende dalla concentrazione del drogaggio (N_q , i.e. quanti ioni positivi e negativi ci sono nella zona svuotata, per unità di volume) moltiplicata per lo spessore di svuotamento (W ,

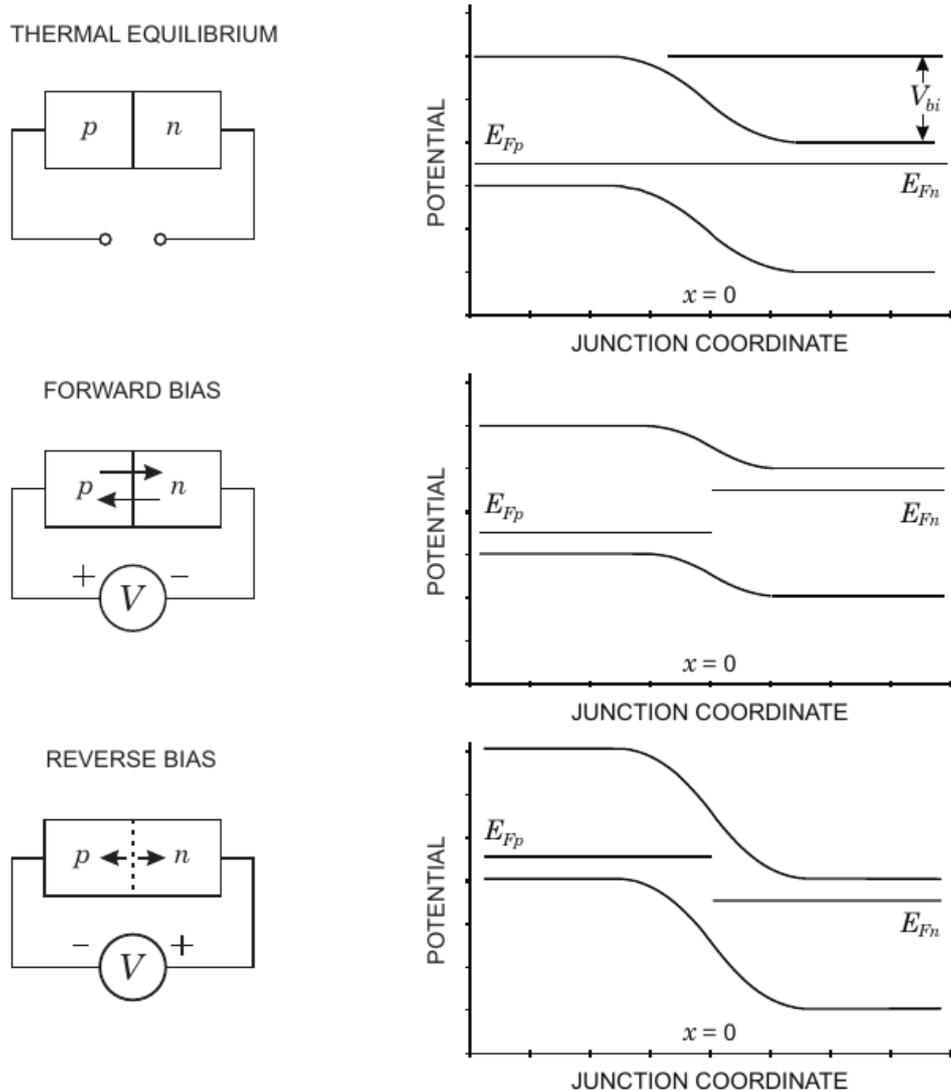


Figura 34: Confronto tra: giunzione pn all'equilibrio termico, in polarizzazione diretta (*forward bias*), ed inversa (*reverse bias*). All'equilibrio, le bande si piegano per eguagliare l'energia di Fermi su entrambi i lati. In polarizzazione diretta, la differenza di potenziale si riduce, favorendo il passaggio di elettroni. In polarizzazione inversa invece, la differenza di potenziale aumenta, bloccando il passaggio di elettroni.

assumendo che la sezione trasversa della giunzione sia costante). Possiamo dunque scrivere:

$$E \propto Q \propto N_q W \quad (56)$$

Siccome la concentrazione di cariche non cambia, l'incremento al campo elettrico può essere unicamente dovuto ad un **aumento dello spessore della**

zona svuotata. Quindi, osservandolo in funzione della tensione, avremo:

$$W(V_R + V_{bi}) > W(V_{bi}) \quad (57)$$

In sostanza, attraverso il generatore di tensione, stiamo inserendo cariche libere di segno opposto rispetto a quelle già presenti in ciascuno dei due lati a causa del drogaggio. Abbiamo quindi un maggior numero di ricombinazioni alle due periferie della zona svuotata, con un aumento dello spessore W di svuotamento.

In questo regime, la corrente di diffusione è ancora più piccola di quella che avevamo all'equilibrio termico, per cui è trascurabile ($\vec{J}_{diff} = \vec{0}$). Avendo un campo elettrico più intenso, ci potremmo aspettare un contributo significativo da parte della corrente di deriva: ciò in realtà non avviene poiché la zona svuotata è sostanzialmente **priva di cariche libere**. Ciò che si osserva, all'aumentare della tensione di *reverse bias*, è una corrente di deriva molto piccola ($|\vec{J}_{der}| \gtrsim 0$), che inizialmente cresce in modo lineare, e presto si stabilizza ad un valore pressoché costante.

Tale corrente viene anche chiamata **corrente di buio** (*dark current*), o di perdita (*leakage current*). Essa rappresenta il rumore di fondo da cui si vuole differenziare il segnale. Diversi fenomeni, quali la temperatura e il danno da radiazione, possono contribuire ad aumentare questa corrente, e dunque è necessaria una grande attenzione al suo studio e all'implementazione di metodi per la sua riduzione.

Aumentando in maniera considerevole la tensione di polarizzazione inversa, si raggiunge il valore di **rottura** (o *breakdown*), e si ha un aumento esponenziale della corrente.

La giunzione *pn* è in definitiva un oggetto che può comportarsi da conduttore o da isolante, a seconda della tensione con cui viene alimentata. Nel nostro caso, la useremo come rivelatore di particelle, quindi in *reverse bias*. La giunzione *pn* è poi usata in una moltitudine di applicazioni molto diverse, quali **diodi, transistor, amplificatori**.

Le grandezze già introdotte per le giunzioni in regime di equilibrio termico, che ridefiniremo aggiungendovi il **pedice 0**, possono facilmente essere riadattate al regime di *reverse bias*. In particolare, gli **spessori di svuotamento** (formule 40 e 41) possono essere scritti come:

$$\begin{aligned}
x_p &= \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_D}{N_A(N_A + N_D)}} = x_{p,0} \sqrt{\frac{V_R + V_{bi}}{V_{bi}}} = x_{p,0} \sqrt{1 + \frac{V_R}{V_{bi}}} \\
x_n &= \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_A}{N_D(N_A + N_D)}} = x_{n,0} \sqrt{\frac{V_R + V_{bi}}{V_{bi}}} = x_{n,0} \sqrt{1 + \frac{V_R}{V_{bi}}} \\
W &= \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_A + N_D}{N_A N_D}} = W_0 \sqrt{\frac{V_R + V_{bi}}{V_{bi}}} = W_0 \sqrt{1 + \frac{V_R}{V_{bi}}}
\end{aligned} \tag{58}$$

Mentre il **campo elettrico massimo** (formula 42) viene ad essere:

$$\begin{aligned}
E_{MAX} &= \sqrt{\frac{2q(V_R + V_{bi})}{\varepsilon} \frac{N_A N_D}{(N_A + N_D)} \frac{2(V_R + V_{bi})}{2(V_R + V_{bi})}} = \\
&= \sqrt{\left[\frac{q}{2\varepsilon(V_R + V_{bi})} \frac{N_A N_D}{(N_A + N_D)} \right] 4(V_R + V_{bi})^2} = \\
&= \frac{2(V_R + V_{bi})}{W}
\end{aligned} \tag{59}$$

Esempio. Considero la giunzione *pn* in silicio vista in precedenza (a temperatura ambiente $T_0 = 300$ K, con $N_A = 5.0 \cdot 10^{15} \text{ cm}^{-3}$ e $N_D = 5.0 \cdot 10^{16} \text{ cm}^{-3}$), per la quale avevamo già trovato che $x_{p,0} = 0.42 \text{ }\mu\text{m}$ e $x_{n,0} = 0.042 \text{ }\mu\text{m}$.

Applicando una differenza di potenziale in regime di *reverse bias* $V_R = 4.0$ V ai capi della giunzione, gli spessori di svuotamento diventano:

$$\begin{aligned}
x_p &= x_{p,0} \sqrt{\frac{V_R + V_{bi}}{V_{bi}}} = 0.42 \cdot \sqrt{\frac{4.0 + 0.7}{0.7}} \text{ }\mu\text{m} = 0.42 \cdot 2.6 \text{ }\mu\text{m} = 1.1 \text{ }\mu\text{m} \\
x_n &= x_{n,0} \sqrt{\frac{V_R + V_{bi}}{V_{bi}}} = 0.042 \cdot \sqrt{\frac{4.0 + 0.7}{0.7}} \text{ }\mu\text{m} = 0.042 \cdot 2.6 \text{ }\mu\text{m} = 0.11 \text{ }\mu\text{m}
\end{aligned}$$

Ancora, se aumentiamo il potenziale a $V_R = 8.0$ V:

$$x_p = x_{p,0} \sqrt{\frac{V_R + V_{bi}}{V_{bi}}} = 0.42 \cdot \sqrt{\frac{8.0 + 0.7}{0.7}} \mu\text{m} = 0.42 \cdot 3.5 \mu\text{m} = 1.5 \mu\text{m}$$

$$x_n = x_{n,0} \sqrt{\frac{V_R + V_{bi}}{V_{bi}}} = 0.042 \cdot \sqrt{\frac{8.0 + 0.7}{0.7}} \mu\text{m} = 0.042 \cdot 3.5 \mu\text{m} = 0.15 \mu\text{m}$$

Esempio. Voglio calcolare la tensione esterna V_R che devo applicare ai capi di una giunzione pn per ottenere un valore di campo massimo pari ad $E_{MAX} = 7.2 \cdot 10^4$ V/cm. Stiamo operando a temperatura ambiente, con concentrazioni di accettori e donori rispettivamente pari a $N_A = 3.0 \cdot 10^{16}$ cm⁻³ e $N_D = 5.0 \cdot 10^{15}$ cm⁻³.

In tal caso basta invertire opportunamente la relazione 59, che, elevata al quadrato, ci da:

$$E_{MAX}^2 = \frac{2q(V_R + V_{bi})}{\epsilon} \frac{N_A N_D}{(N_A + N_D)} \implies V_R + V_{bi} = \frac{\epsilon E_{MAX}^2}{2q} \frac{(N_A + N_D)}{N_A N_D}$$

Sapendo inoltre che $V_{bi} = 0.7$, abbiamo dunque:

$$\begin{aligned} V_R &= \frac{\epsilon E_{MAX}^2}{2q} \frac{(N_A + N_D)}{N_A N_D} - V_{bi} = \\ &= \frac{1.035 \cdot 10^{-12} \text{ C V}^{-1} \text{ cm}^{-1} \cdot (7.2 \cdot 10^4 \text{ V cm}^{-1})^2}{2 \cdot 1.6 \cdot 10^{-19} \text{ C}} \frac{(3.0 \cdot 10^{16} + 5.0 \cdot 10^{15}) \text{ cm}^{-3}}{(3.0 \cdot 10^{16} \cdot 5.0 \cdot 10^{15}) \text{ cm}^{-6}} - 0.7 \text{ V} = \\ &= \frac{1.035 \cdot 7.2^2 \cdot 3.5}{2 \cdot 1.6 \cdot 3.0 \cdot 5.0} 10^{(-12+8+16+19-31)} \text{ C}^{(+1-1)} \text{ V}^{(-1+2)} \text{ cm}^{(-1-2-3+6)} - 0.7 \text{ V} = \\ &= 4.0 \cdot 10^0 \text{ V} - 0.7 \text{ V} = 3.3 \text{ V} \end{aligned}$$

Vediamo ora come si modifica la **carica** inclusa nella zona svuotata in funzione del potenziale esterno V_R applicato, in *reverse bias*. Tenendo anche conto della geometria del problema, è utile definire la **carica per unità di superficie** (Q') e la **capacità per unità di superficie** (C'), come:

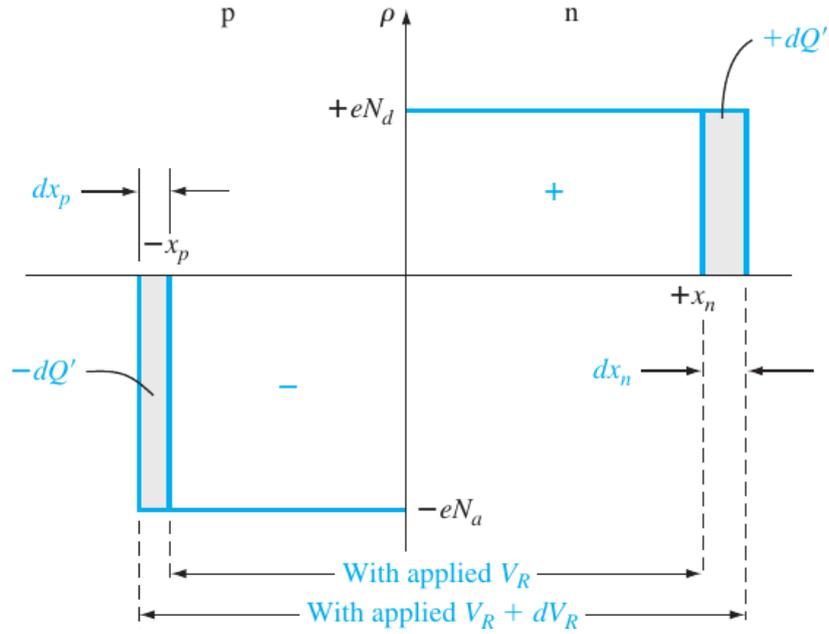


Figura 35: Incrementando il potenziale esterno V_R in regime di *reverse bias*, la zona di svuotamento si allarga, ed aumenta la carica in essa contenuta.

$$Q' = \frac{Q}{A} \quad \text{e} \quad C' = \frac{C}{A} = \frac{dQ'}{dV_R} \quad (60)$$

dove A rappresenta la sezione trasversale della giunzione, rispetto alla coordinata x . Dimensionalmente avremo che $[Q'] = \text{C cm}^{-2}$, mentre $[C'] = \text{C V}^{-1} \text{ cm}^{-2} = \text{F cm}^{-2}$.

La carica presente nel lato p della zona svuotata è identica, in modulo, a quella presente nel lato n , ovvero $Q' = qN_D x_n = qN_A x_p$. Tenendo in considerazione l'espressione trovata per x_n (in formula 58), posso calcolare la capacità per unità di superficie come:

$$\begin{aligned}
C' &= \frac{dQ'}{dV_R} = \frac{d}{dV_R}(qN_D x_n) = qN_D \frac{dx_n}{dV_R} = \\
&= qN_D \frac{1}{2} \left[\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_A}{N_D(N_A + N_D)} \right]^{-1/2} \left[\frac{2\varepsilon}{q} \frac{N_A}{N_D(N_A + N_D)} \right] = \\
&= \sqrt{q^2 N_D^2 \frac{1}{4} \frac{q}{2\varepsilon(V_R + V_{bi})} \frac{N_D(N_A + N_D)}{N_A} \frac{4\varepsilon^2}{q^2} \frac{N_A^2}{N_D^2(N_A + N_D)^2}} = \\
&= \sqrt{\frac{q\varepsilon N_A N_D}{2(N_A + N_D)(V_R + V_{bi})}} \tag{61}
\end{aligned}$$

Tenendo conto dell'espressione trovata per W (nella formula 58), posso anche scriverla come:

$$C' = \sqrt{\frac{q\varepsilon N_A N_D}{2(N_A + N_D)(V_R + V_{bi})} \frac{\varepsilon}{\varepsilon}} = \sqrt{\varepsilon^2 \underbrace{\frac{q\varepsilon N_A N_D}{2(N_A + N_D)(V_R + V_{bi})}}_{=W^{-1}}} = \frac{\varepsilon}{W} \tag{62}$$

Esempio. Voglio calcolare la capacità per unità di superficie (C') nel caso della giunzione dell'esempio precedente, con l'unica differenza che le concentrazioni di accettori e donori sono invertite, ovvero $N_A = 5.0 \cdot 10^{15} \text{ cm}^{-3}$ e $N_D = 3.0 \cdot 10^{16} \text{ cm}^{-3}$. Notiamo che la formula 59 non cambia invertendo le due concentrazioni, quindi la tensione di alimentazione in *reverse bias* rimane $V_R = 3.3 \text{ V}$, con $V_{bi} = 0.7 \text{ V}$.

Utilizzando la formula 61, ricaviamo immediatamente che:

$$\begin{aligned}
C' &= \sqrt{\frac{q\varepsilon N_A N_D}{2(N_A + N_D)(V_R + V_{bi})}} = \\
&= \sqrt{\frac{1.6 \cdot 10^{-19} \text{ C} \cdot 1.035 \cdot 10^{-12} \text{ C V}^{-1} \text{ cm}^{-1} \cdot 5.0 \cdot 10^{15} \text{ cm}^{-3} \cdot 3.0 \cdot 10^{16} \text{ cm}^{-3}}{2 \cdot (5.0 \cdot 10^{15} \text{ cm}^{-3} + 3.0 \cdot 10^{16} \text{ cm}^{-3}) \cdot (3.3 \text{ V} + 0.7 \text{ V})}} = \\
&= \sqrt{\frac{1.6 \cdot 1.035 \cdot 5.0 \cdot 3.0}{2 \cdot 3.5 \cdot 4.0} 10^{(-19-12+15+16-16)} \text{ C}^{(1+1)} \text{ V}^{(-1-1)} \text{ cm}^{(-1-3-3+3)}} = \\
&= \sqrt{0.90 \cdot 10^{-16} \text{ C}^2 \text{ V}^{-2} \text{ cm}^{-4}} = 0.95 \cdot 10^{-8} \frac{\text{C}}{\text{V cm}^2} = 9.5 \cdot 10^{-9} \frac{\text{F}}{\text{cm}^2} = 9.5 \frac{\text{nF}}{\text{cm}^2}
\end{aligned}$$

Applicando al caso realistico di un sensore avente area:

$$A = 100 \mu\text{m} \cdot 100 \mu\text{m} = (1.00 \cdot 10^{-2} \text{ cm})^2 = 1.00 \cdot 10^{-4} \text{ cm}^2$$

la sua capacità (C) sarà:

$$C = C' A = 9.5 \cdot 10^{-9} \frac{\text{F}}{\text{cm}^2} 1.00 \cdot 10^{-4} \text{ cm}^2 = 9.5 \cdot 10^{-13} \text{ F} = 0.95 \cdot 10^{-12} \text{ F} = 0.95 \text{ pF}$$

Vediamo ora il caso delle **giunzioni unilaterali**, ovvero nelle quali **una delle due concentrazioni è molto maggiore dell'altra**.

Se considero $N_A \gg N_D$ ho che:

$$\frac{N_D}{N_A(N_A + N_D)} \sim 0 \quad , \quad \frac{N_A}{N_D(N_A + N_D)} \sim \frac{1}{N_D} \quad \text{e} \quad \frac{N_A + N_D}{N_A N_D} \sim \frac{1}{N_D} \quad (63)$$

Quindi le relazioni nelle formule 58, 61 e 62 si approssimano come:

$$\begin{aligned} x_p &= \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_D}{N_A(N_A + N_D)}} \simeq 0 \\ x_n &= \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_A}{N_D(N_A + N_D)}} \simeq \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q N_D}} \\ W &= \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_A + N_D}{N_A N_D}} \simeq \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q N_D}} \simeq x_n \\ C' &= \sqrt{\frac{q\varepsilon N_A N_D}{2(N_A + N_D)(V_R + V_{bi})}} = \sqrt{\frac{q\varepsilon N_D}{2(V_R + V_{bi})}} \simeq \frac{\varepsilon}{x_n} \end{aligned} \quad (64)$$

In modo del tutto analogo, se invece considero $N_D \gg N_A$ ho che:

$$\frac{N_D}{N_A(N_A + N_D)} \sim \frac{1}{N_A} \quad , \quad \frac{N_A}{N_D(N_A + N_D)} \sim 0 \quad \text{e} \quad \frac{N_A + N_D}{N_A N_D} \sim \frac{1}{N_A} \quad (65)$$

In questo caso, le relazioni nelle formule 58, 61 e 62 si approssimano come:

$$\begin{aligned}
x_p &= \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_D}{N_A(N_A + N_D)}} \simeq \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{qN_A}} \\
x_n &= \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_A}{N_D(N_A + N_D)}} \simeq 0 \\
W &= \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{q} \frac{N_A + N_D}{N_A N_D}} \simeq \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{qN_A}} \simeq x_p \\
C' &= \sqrt{\frac{q\varepsilon N_A N_D}{2(N_A + N_D)(V_R + V_{bi})}} = \sqrt{\frac{q\varepsilon N_A}{2(V_R + V_{bi})}} \simeq \frac{\varepsilon}{x_p}
\end{aligned} \tag{66}$$

In entrambi i casi, si osserva che:

- $W \propto \sqrt{V_R + V_{bi}}$, ovvero il volume sensibile cresce con la radice del potenziale complessivo;
- $C' \propto 1/\sqrt{V_R + V_{bi}}$, ovvero la capacità decresce con la radice del potenziale complessivo.

In pratica, possiamo considerare la quantità $1/C'^2$, che risulta lineare nella tensione V_R di *reverse bias*, del tipo $a + bV_R$. Usando le relazioni per C' nelle rispettive formule 64 e 66, otteniamo:

$$\begin{aligned}
\frac{1}{C'^2}(V_R) &= \frac{2(V_R + V_{bi})}{q\varepsilon N_D} = \underbrace{\frac{2V_{bi}}{q\varepsilon N_D}}_{=a} + \underbrace{\frac{2}{q\varepsilon N_D}}_{=b} V_R & \text{se } N_A \gg N_D \\
\frac{1}{C'^2}(V_R) &= \frac{2(V_R + V_{bi})}{q\varepsilon N_A} = \underbrace{\frac{2V_{bi}}{q\varepsilon N_A}}_{=a} + \underbrace{\frac{2}{q\varepsilon N_A}}_{=b} V_R & \text{se } N_D \gg N_A
\end{aligned} \tag{67}$$

3.3 Svuotamento, sovratensione, rottura

D'ora in poi non distingueremo fra la capacità C e la capacità per unità di sezione $C' = C/A$, visto che esse differiscono solo per un fattore costante (la sezione trasversa A della giunzione).

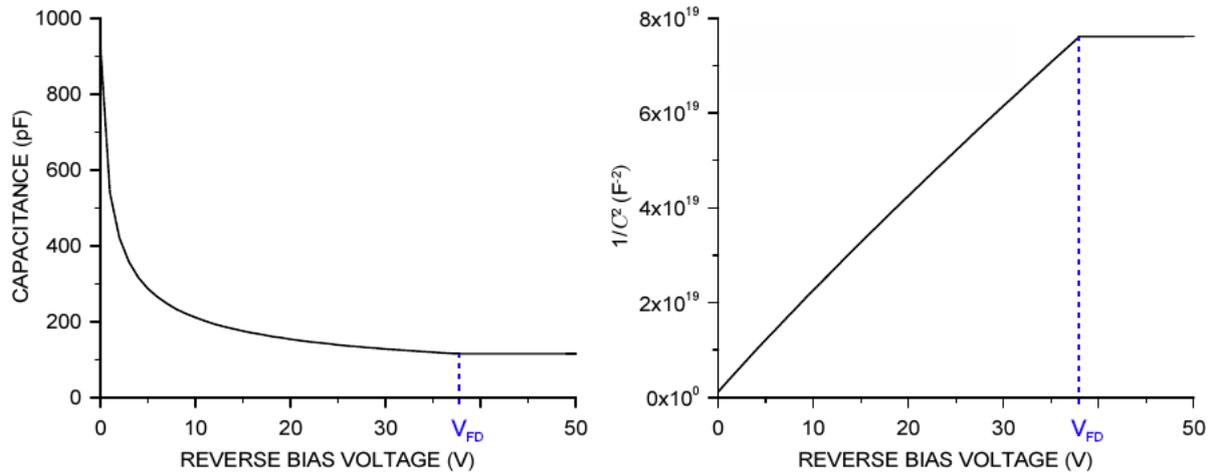


Figura 36: Grafici della capacità C (a sinistra) e del suo reciproco al quadrato $1/C^2$ (a destra) in funzione della tensione di *reverse bias* V_R , per una giunzione *pn* unilaterale, avente concentrazione $N = 5 \cdot 10^{12} \text{ cm}^{-3}$ del drogante maggioritario. Aumentando V_R oltre al valore della tensione di massimo svuotamento V_{FD} , la capacità raggiunge il suo valore minimo e rimane costante.

Abbiamo già visto che la **capacità** di una giunzione *pn* è proporzionale allo **spessore di svuotamento**, il quale è proporzionale a sua volta dalla **tensione** di alimentazione in regime di *reverse bias*, ossia: $C \propto (1/W) \propto (1/\sqrt{V_R})$.

Aumentando la tensione V_R , la zona di svuotamento cresce, fino ad eguagliare il volume del semiconduttore stesso. Tale condizione viene raggiunta quando $V_R = V_{FD}$, che viene definita **tensione svuotamento totale** (*full depletion voltage*).

Aumentando ulteriormente la tensione di alimentazione a valori $V_R \geq V_{FD}$, la capacità resta **costante**, come si può osservare nel grafico riportato a sinistra in fig. 36. Tale andamento è ancora più evidente se consideriamo la variabile $1/C^2$, il cui grafico in funzione di V_R è riportato a destra in fig. 36.

Consideriamo il caso di una giunzione unilaterale p^+n , dove l'**apice** + davanti al tipo di drogaggio (*p*, nel nostro caso) indica che esso è molto maggiore rispetto a quello del tipo opposto (*n*, nel nostro caso). Per quanto visto in formula 64, $W \simeq x_n$, mentre $x_p \simeq 0$. La zona svuotata dunque si estende prettamente nel lato *n*, mentre dall'altro lato avremo una piccolissima zona che rimane pesantemente drogata di tipo *p*.

Derivando rispetto a V_R l'espressione 67 ottenuta per $1/C^2$, otteniamo:

$$\frac{d}{dV_R} \left(\frac{1}{C^2} \right) = \frac{2}{q\varepsilon N_D} \quad (68)$$

Tale derivata corrisponde alla pendenza della retta nel grafico a destra in fig. 36. Possiamo ricavare il suo valore mediante un *fit* a retta su delle misure di capacità in funzione della tensione applicata. Ciò permette di ottenere la **concentrazione dei portatori minoritari** N_D in funzione di quantità note, o comunque misurabili:

$$N_D = \frac{2}{q\varepsilon} \left[\frac{d}{dV_R} \left(\frac{1}{C^2} \right) \right]^{-1} \quad (69)$$

Di solito, tale concentrazione viene espressa in termini della **resistività** ρ . Invertendo la formula ??, abbiamo che $N_D = (\rho q \mu_e)^{-1}$, quindi:

$$\rho = \frac{\varepsilon}{2\mu_e} \frac{d}{dV_R} \left(\frac{1}{C^2} \right) \quad (70)$$

Volendo misurare lo spessore di svuotamento, possiamo usare la relazione 62, da cui:

$$C = \frac{\varepsilon}{W} \quad \text{e} \quad \frac{1}{C} = \frac{W}{\varepsilon} \quad (71)$$

Effettuandone le misure, si ottengono due grafici simili a quelli visti in fig. 36. Ragionando in modo analogo a quanto fatto prima, possiamo ricavare lo **spessore di svuotamento massimo**, che chiameremo W_{FD} .

Abbiamo quindi dei metodi per **misurare indirettamente** le tre grandezze viste. Ciò è utile a **fini pratici** poiché il produttore della giunzione spesso non è in grado di fornire misure dirette di tali grandezze: in tal modo esse possono comunque essere **misurate a posteriori**.

Notiamo che gli andamenti visti nei grafici in fig. 36 presentano solo idealmente un punto angoloso per $V_R = V_{FD}$. Nella realtà si osserva che essi sono raccordati in modo **continuo**: ciò è dovuto al fatto che anche il profilo di drogaggio varia in modo continuo fra i due lati della giunzione. Ad ogni modo, dal secondo grafico è comunque possibile ricavare la tensione di svuotamento massimo V_{FD} , come l'intersezione tra i due *fit* a retta, fatti sui rispettivi andamenti.

Nel caso in cui la giunzione venga utilizzata come rivelatore, il **rilascio di carica** che avviene entro il volume svuotato (sensibile) dipende dalla sua **capacità**. Come visibile nei grafici in fig. 37, aumentando la tensione V_R di

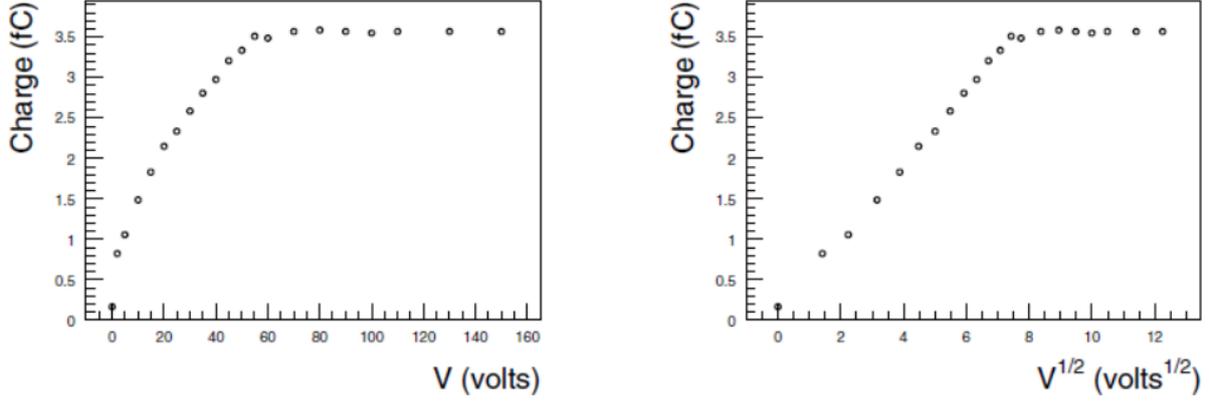


Figura 37: Grafici della carica raccolta in un rivelatore al silicio (dello spessore di $300 \mu\text{m}$ ed area di 1 cm^2) esposto ad un fascio di elettroni (ad energia $\geq 2 \text{ MeV}$), in funzione della tensione di alimentazione in *reverse bias*, a sinistra, e della sua radice, a destra.

alimentazione in *reverse bias*, aumenta anche la capacit , e quindi la carica raccolta, sino al raggiungimento della tensione di massimo svuotamento, oltre la quale rimane invece costante.

Vediamo ora come cambia il **campo elettrico** entro la zona svuotata di una giunzione p^+n . Innanzitutto, abbiamo gi  visto (formula 64) che $x_p \simeq 0$, mentre:

$$x_n \simeq W \simeq \sqrt{\frac{2\varepsilon(V_R + V_{bi})}{qN_D}} \Rightarrow W^2 \simeq \frac{2\varepsilon(V_R + V_{bi})}{qN_D} \Rightarrow \frac{qN_D}{\varepsilon} \simeq \frac{2(V_R + V_{bi})}{W^2} \quad (72)$$

La relazione 30, vista nel caso di una giunzione all'equilibrio termico, pu  dunque essere riadattata per ricavare il **modulo del campo elettrico** in funzione della tensione e dello spessore di svuotamento, come:

$$E(x) = \frac{qN_D}{\varepsilon}(x - x_n) \simeq \frac{2(V_R + V_{bi})}{W^2}(W - x) = \frac{2(V_R + V_{bi})}{W} \left(1 - \frac{x}{W}\right) \quad (73)$$

Se la giunzione ha spessore d , lo svuotamento totale si verifica per $W = W_{FD} = d$, quasi interamente dal lato n . Ad esempio, nel caso dei tracciatori di nostro interesse, aventi spessore $d \sim 10^2 \mu\text{m}$, avremo che $x_p \sim 10^{-1} \mu\text{m}$, trascurabile dunque rispetto ad x_n .

Sapendo che lo svuotamento totale si ottiene applicando una tensione $V_R = V_{FD}$, l'espressione per W , ricavata in formula 64, possiamo scrivere:

$$d = \sqrt{\frac{2\varepsilon(V_{FD} + V_{bi})}{qN_D}} \Rightarrow d^2 = \frac{2\varepsilon(V_{FD} + V_{bi})}{qN_D} \Rightarrow d^2 = \frac{2\varepsilon}{qN_D}V_{FD} + \frac{2\varepsilon}{qN_D}V_{bi} \quad (74)$$

In tal modo otteniamo un'espressione per la tensione V_{FD} di svuotamento totale in funzione di grandezze note, o comunque misurabili:

$$V_{FD} = \frac{qN_D}{2\varepsilon}d^2 - V_{bi} \quad (75)$$

Viene anche definita la **tensione di svuotamento interno** (*internal depletion voltage*), come:

$$V_{di} = V_{FD} + V_{bi} \quad (76)$$

Negli scopi pratici, la tensione *built-in* risulta essere **trascurabile** rispetto a quella fornita dall'esterno, ossia $V_{FD} \gg V_{bi}$, per cui potremo approssimare $V_{di} \simeq V_{FD}$. Ad esempio, i tracciatori al silicio di nostro interesse sono alimentati con tensioni di svuotamento completo $V_{FD} = (60 \div 100)$ V, quindi $V_{bi} \simeq 0.7$ V risulta almeno due ordini di grandezza più piccola, dunque trascurabile in somma.

Quando la tensione di alimentazione supera quella di svuotamento completo, ovvero $V_R > V_{FD}$, si parla di **sovrasvuotamento** (o *overbias*). In tale condizione lo spessore W della zona svuotata **non può aumentare ulteriormente**, siccome esso ha già raggiunto il suo valore massimo, pari allo spessore d della giunzione stessa.

Come visibile nei grafici di fig. 38, il modulo del campo elettrico E è massimo all'interfaccia ($x = 0$) e **decrece linearmente** con x . La **quota** della retta varia con la radice della tensione applicata ($E \propto \sqrt{V_R + V_{bi}}$, come in formula 59). Essendo $x_p \simeq 0$, l'iniziale fase di crescita lineare per $-x_p \leq x \leq 0$ viene trascurata.

Nel grafico a sinistra in fig. 38 vediamo che, per tensioni applicate $V_R < V_{FD}$, il campo elettrico si annulla in $x = W < d$. In quello di destra vediamo invece il caso per $V_R > V_{FD}$. Qui il campo elettrico **non si annulla** mai, ma raggiunge un **valore minimo** sull'elettrodo n , ovvero $E(x = d) = E_{min} > 0$, dovuto al contributo della sovratensione, che possiamo scrivere come $V_R - (V_{FD} + V_{bi})$. In questo caso, il modulo del campo elettrico sarà quindi:

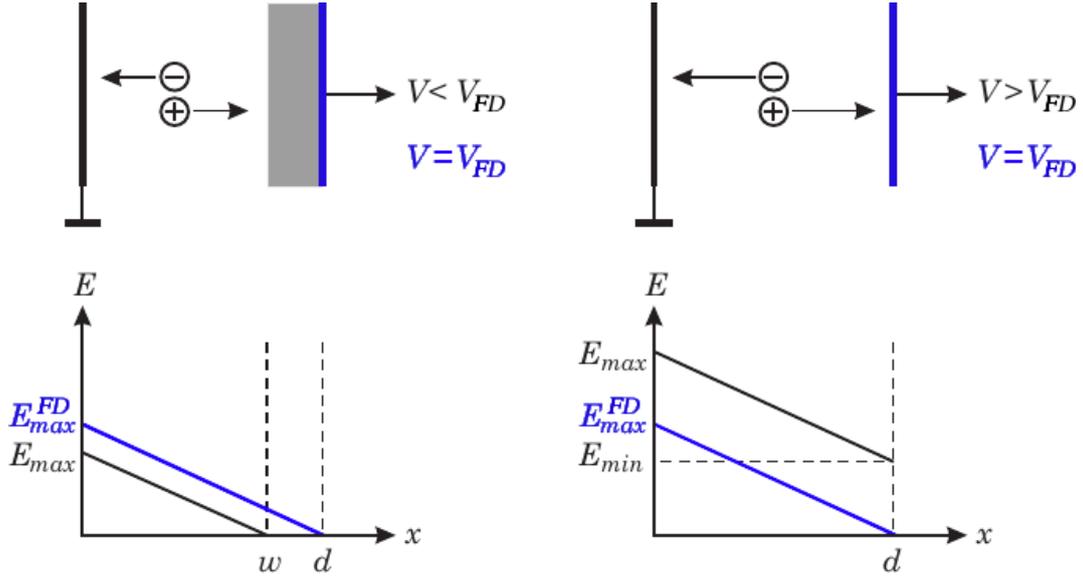


Figura 38: Campo elettrico nella zona svuotata di una giunzione p^+n di spessore d . A sinistra nel caso di svuotamento parziale (ad una tensione applicata $V < V_{FD}$), a destra nel caso di *overbias* (ad una tensione applicata $V > V_{FD}$). Per confronto, viene riportato anche il caso di svuotamento completo (ad una tensione applicata $V = V_{FD}$).

$$E(x) = \frac{2(V_{FD} + V_{bi})}{d} \left(1 - \frac{x}{d}\right) + \frac{V_R - (V_{FD} + V_{bi})}{d} \quad (77)$$

dove il primo termine discende dalla formula 73, e descrive la decrescita lineare fra $x = 0$ ed $x = d$, mentre il secondo termine è l'*offset* costante dovuto al contributo della sovratensione.

Ci preoccupiamo adesso di ricavare il **tempo di raccolta** delle cariche liberate all'interno della zona svuotata di una giunzione p^+n . Nel caso di **svuotamento parziale**, per $V_R < V_{FD}$, possiamo scrivere il campo elettrico (sempre riadattando la relazione 30, vista nel caso di una giunzione all'equilibrio termico) come:

$$E(x) = E_0 \cdot (W - x) \quad \text{con} \quad E_0 \equiv -\frac{qN_D}{\varepsilon} \quad (78)$$

Usando la definizione di **mobilità**, possiamo scrivere la **velocità di deriva** delle cariche come:

$$v(x) = \mu E(x) = \mu E_0 \cdot (W - x) \quad (79)$$

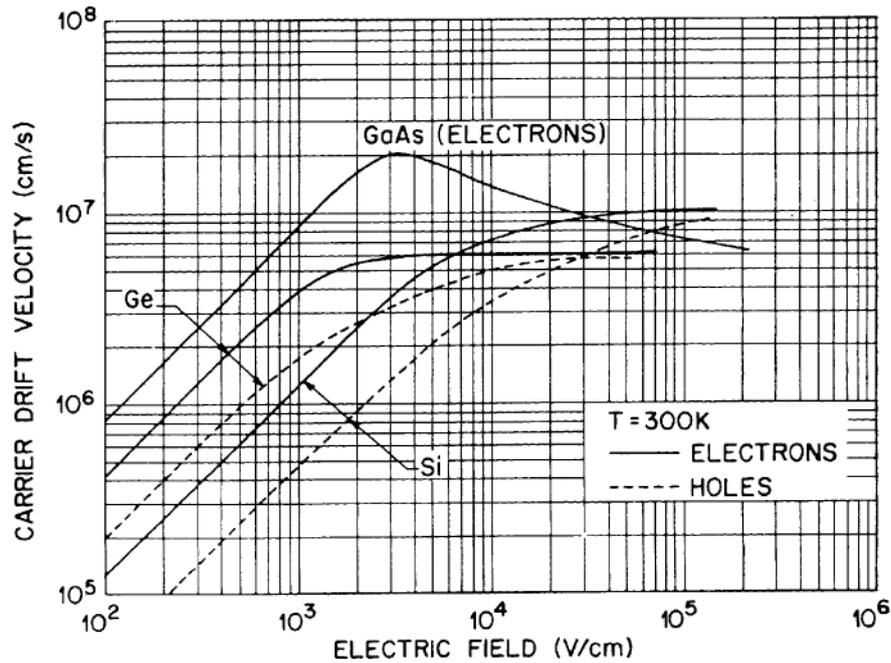


Figura 39: Velocità dei portatori di carica (elettroni e buche) per diversi semiconduttori, a temperatura ambiente ($T_0 = 300 \text{ K}$), in funzione del campo elettrico.

Come si vede nel grafico di fig. 39, la mobilità dei portatori di carica dipende dal campo elettrico che si crea nel materiale. Nel caso del silicio a temperatura ambiente, osserviamo che:

- per $E \lesssim 10^3 \text{ V/cm}$, la mobilità è circa **costante**;
- per $E \gtrsim 10^4 \text{ V/cm}$, essa **decresce**. Gli elettroni a questi regimi infatti gli sentono l'**interazione con il reticolo** cristallino del materiale in cui si muovono, e ivi perdono energia con emissione di **fononi**;
- per $E \gtrsim 10^5 \text{ V/cm}$, persiste l'andamento decrescente, con $\mu \propto 1/E$. A questo regime la **velocità** dei portatori diventa **costante**, con $v \simeq 10^7 \text{ cm/s}$.

Per ricavare il tempo di deriva, sarà sufficiente integrare il reciproco della velocità (formula 79) nella coordinata spaziale x . Per i nostri scopi possiamo trascurare la dipendenza della mobilità dal campo elettrico, e dunque dalla posizione x . Come riferimenti posizionali, avremo:

- in $x = 0$ l'elettrodo p , verso il quale migrano le buche;

- in $x = W$ l'elettrodo n , verso il quale migrano gli elettroni;
- in $x = x_0$, con $0 \leq x_0 \leq W$, il punto di origine del rilascio di carica.

Il tempo impiegato dal portatore di carica per portarsi dalla posizione di origine x_0 alla generica posizione x (interna alla giunzione) è calcolabile come:

$$t(x) = \int_{x_0}^x \frac{dx'}{v(x')} = \frac{1}{\mu E_0} \int_{x_0}^x \frac{dx'}{W - x'} = -\frac{1}{\mu E_0} \ln(w - x)|_{x_0}^x = \frac{\varepsilon}{\mu q N_D} \ln \left(\frac{W - x}{W - x_0} \right) \quad (80)$$

Nel caso di una **lacuna** (h) che migra dal punto di origine (in $x = x_0$) all'elettrodo p (in $x = 0$), il tempo di deriva si ottiene applicando la formula 80 appena vista:

$$t_h(0) = \frac{\varepsilon}{\mu_h q N_D} \ln \left(\frac{W}{W - x_0} \right) = \tau_h \ln \left(\frac{W}{W - x_0} \right) \quad \text{con} \quad \tau_h \equiv \frac{\varepsilon}{\mu_h q N_D} \quad (81)$$

dove τ_h viene definito **tempo caratteristico delle lacune**. Confrontando alcuni valori numerici, si osserva come variano i tempi di deriva, ad esempio:

- se il rilascio di carica avviene **al centro** dello spessore svuotato, ovvero $x_0 = 50\%W$, allora $t_h(0) = 0.7\tau_h$;
- se invece avviene nei pressi dell'**elettrodo opposto**, come potrebbe essere $x_0 = 95\%W$, allora $t_h(0) = 3.0\tau_h$.

Nel caso di un **elettrone** (e) che si muove dal punto di origine (in $x = x_0$) all'elettrodo n (in $x = W$), la formula 80 **non ha una soluzione** diretta, poiché porterebbe a tempi divergenti: $t_e(W) \propto \ln(W - W) = \ln(0) = -\infty$. Il problema può essere comunque trattato, invertendo la formula stessa per esprimere la posizione in funzione del tempo, come:

$$\ln \left(\frac{W - x(t)}{W - x_0} \right) = -\frac{t}{\tau_e} \Rightarrow \exp \left(-\frac{t}{\tau_e} \right) = \frac{W - x(t)}{W - x_0} \Rightarrow W - x(t) = (W - x_0) e^{-t/\tau_e}$$

$$x(t) = W - (W - x_0) e^{-t/\tau_e} \quad \text{con} \quad \tau_e \equiv \frac{\varepsilon}{\mu |q| N_D} \quad (82)$$

Notare che il **tempo caratteristico degli elettroni** (τ_e) viene **definito positivo**, per cui il **segno negativo** dovuto alla carica elementare è riportato in modo **esplicito**. Qualora l'elettrone venga generato sull'elettrodo p , ovvero $x_0 = 0$, avremo:

$$x(t) = W(1 - e^{-t/\tau_e}) \quad (83)$$

Osserviamo che, secondo questa descrizione, un elettrone che viene rilasciato in $0 \leq x_0 < W$, non raggiunge mai l'elettrodo n per sola deriva.

Un elettrone però, con il solo moto di deriva, può comunque arrivare in prossimità dell'elettrodo n in tempi relativamente brevi. Usando la relazione 83, possiamo stimare il tempo t' necessario a percorrere il 95% dello spessore svuotato:

$$1 - e^{-t'/\tau_e} = \frac{95}{100} \Rightarrow e^{-t'/\tau_e} = \frac{5}{100} \Rightarrow -\frac{t'}{\tau_e} = \ln\left(\frac{5}{100}\right) = \ln\left(\frac{1}{20}\right) = -\ln(20)$$

$$t' = \tau_e \ln(20) \simeq 3.0\tau_e \quad (84)$$

A questo punto il moto di **diffusione**, che comunque accompagna sempre la deriva, permetterà all'elettrone di giungere comunque all'elettrodo n .

Osserviamo che il **tempo caratteristico** dipende sia dalla **mobilità** μ del portatore considerato, che dalla **concentrazione** N_D dei portatori minoritari. Mentre **non** dipende né dalla **tensione** di alimentazione V_R , né dallo **spessore svuotato** W .

Esprimiamo ora il tempo caratteristico in funzione della **resistività** del materiale. Secondo la formula ??, essa è definita come $\rho = (\mu q N_D)^{-1}$, quindi abbiamo che:

$$\tau = \varepsilon \rho \quad (85)$$

Per il silicio abbiamo $\varepsilon = 1.035 \cdot 10^{-12} \text{ C V}^{-1}\text{cm}^{-1}$. Nella pratica, la resistività ρ si trova spesso espressa in $\text{k}\Omega \cdot \text{cm}$. Riscriviamo dunque ε in modo opportuno, utilizzando la legge di Ohm. Dimensionalmente, abbiamo $[V] = [R] \cdot [I]$, ossia $\text{V} = \Omega \cdot \text{A} = \Omega \cdot \text{C} \cdot \text{s}^{-1}$, da cui:

$$\varepsilon = 1.035 \cdot 10^{-12} \frac{\text{C}}{\text{V} \cdot \text{cm}} = 1.035 \frac{10^{-9}}{10^3} \frac{\text{C} \cdot \text{s}}{\Omega \cdot \text{C} \cdot \text{cm}} = 1.035 \frac{(10^{-9} \text{ s})}{(10^3 \Omega) \cdot \text{cm}} = 1.035 \frac{\text{ns}}{\text{k}\Omega \cdot \text{cm}} \quad (86)$$

Abbiamo già visto che $\mu_h \simeq \mu_e/3$, dunque essendo $\rho \propto (1/\mu)$, la resistività per le **buche** sarà $\rho_h \simeq 3\rho_e$. Ed anche il suo tempo caratteristico sarà $\tau_h = \varepsilon\rho_h \simeq 3\varepsilon\rho_e = 3\tau_e$.

Esempio. Consideriamo una giunzione p^+n con resistività $\rho_e = 10.0 \text{ k}\Omega \text{ cm}$. Possiamo calcolare i tempi caratteristici per elettroni e buche dalla resistività, come:

$$\tau_e = \varepsilon\rho_e = 1.035 \frac{\text{ns}}{\text{k}\Omega \text{ cm}} 10.0 \text{ k}\Omega \text{ cm} = 10.5 \text{ ns} \quad \text{e} \quad \tau_h = 3\tau_e = 3 \cdot 10.5 \text{ ns} = 31.5 \text{ ns}$$

Per le considerazioni fatte in precedenza, abbiamo che entrambi i tipi di portatori di carica percorrono il 95% dello spessore svuotato in un tempo $t \simeq 3\tau$, che approssimiamo al tempo necessario a percorrerlo per intero. Per cui avremo:

$$t_e = 3\tau_e = 31.5 \text{ ns} \quad \text{e} \quad t_h = 3\tau_h = 94.5 \text{ ns}$$

Ci preoccupiamo ora di calcolare i **tempi di raccolta** delle cariche nel caso in cui la giunzione p^+n sia **sovrasvuotata**. Vedremo che saranno **più bassi** rispetto a quelli che si hanno nel caso di svuotamento parziale. Per prima cosa, scriviamo il campo elettrico in forma analoga a quella vista per il suo modulo (formula 77):

$$E(x) = - \left[E_0 \cdot \left(1 - \frac{x}{d} \right) + E_1 \right] \quad (87)$$

In riferimento al grafico a destra in fig. 38, qui $E_0 = E_{max}^{FD} = E_{max} - E_{min}$ rappresenta il valore massimo del campo, che otteniamo in condizione di svuotamento completo, mentre $E_1 = E_{min}$ è l'*offset* costante, dovuto alla tensione di *overbias*.

Anche in questo caso, usiamo la definizione di **mobilità** (formula ??) per scrivere la **velocità di deriva** delle cariche come:

$$v(x) = \mu E(x) = -\mu \left[E_0 \cdot \left(1 - \frac{x}{d} \right) + E_1 \right] \quad (88)$$

Quindi il tempo di deriva dei portatori di carica si può ricavare come:

$$\begin{aligned}
t(x) &= \int_{x_0}^x \frac{dx'}{v(x')} = -\frac{1}{\mu} \int_{x_0}^x \left[E_0 \cdot \left(1 - \frac{x'}{d} \right) + E_1 \right]^{-1} dx' = \\
&= -\frac{1}{\mu} \left(-\frac{d}{E_0} \right) \ln \left(E_0 + E_1 - E_0 \frac{x'}{d} \right) \Big|_{x_0}^x = \\
&= \frac{d}{\mu E_0} \ln \left(\frac{E_0 + E_1 - E_0 \frac{x}{d}}{E_0 + E_1 - E_0 \frac{x_0}{d}} \right)
\end{aligned} \tag{89}$$

Nel caso di una **buca** che viene generata in $x_0 = d$ (ovvero al limite della zona svuotata, sull'elettrodo n), essa compirà un moto di deriva lungo tutta la giunzione, per venir raccolta in $x = 0$, in un tempo:

$$t_h(0) = \frac{d}{\mu_h E_0} \ln \left(\frac{E_0 + E_1 - E_0 \frac{0}{d}}{E_0 + E_1 - E_0 \frac{d}{d}} \right) = \frac{d}{\mu_h E_0} \ln \left(\frac{E_0 + E_1}{E_1} \right) = \frac{d}{\mu_h E_0} \ln \left(1 + \frac{E_0}{E_1} \right) \tag{90}$$

Si osserva che aumentando la tensione di *overbias*, ovvero il campo E_1 , il termine logaritmico si riduce, assieme al tempo di deriva. Ricordando che $\ln(1+x) \sim x$ quando x è piccolo, otteniamo:

$$t_h(0) = \frac{d}{\mu_h E_0} \frac{E_0}{E_1} = \frac{d}{\mu_h E_1} \quad \text{quando} \quad E_1 \gg E_0 \tag{91}$$

La carica si muoverà dunque come fosse sotto l'azione soltanto della componente costante E_1 del campo elettrico.

Consideriamo invece un **eletttrone** generato in $x_0 = 0$, ossia al limite della giunzione sull'elettrodo p , che compie un moto di deriva verso l'elettrodo opposto, in $x = d$. Osservando che **mobilità** sarà **negativa** a causa del segno della carica elementare, avremo un tempo di deriva pari a:

$$\begin{aligned}
t_e(d) &= -\frac{d}{\mu_e E_0} \ln \left(\frac{E_0 + E_1 - E_0 \frac{d}{d}}{E_0 + E_1 - E_0 \frac{0}{d}} \right) = -\frac{d}{\mu_e E_0} \ln \left(\frac{E_1}{E_0 + E_1} \right) = \\
&= -\frac{d}{\mu_e E_0} \left[-\ln \left(\frac{E_0 + E_1}{E_1} \right) \right] = \frac{d}{\mu_e E_0} \ln \left(1 + \frac{E_0}{E_1} \right)
\end{aligned} \tag{92}$$

Valgono dunque le stesse considerazioni viste per le buche. In particolare, avremo anche qui:

$$t_e(d) = \frac{d}{\mu_e E_0} \frac{E_0}{E_1} = \frac{d}{\mu_e E_1} \quad \text{quando} \quad E_1 \gg E_0 \quad (93)$$

Vogliamo ora calcolare i **tempi tipici** di deriva, come fatto nel caso precedente. Ricordiamo innanzitutto che i parametri definiti in formula 77, a partire dalla 87, sono:

$$E_0 = \frac{2(V_{FD} + V_{bi})}{d} \quad \text{ed} \quad E_1 = \frac{V_R - (V_{FD} + V_{bi})}{d} \quad (94)$$

A partire da questi, posso dunque riscrivere il tempo di raccolta totale t_c per buche ed elettroni (formule 90 e 92) in funzione della **tensione**, come:

$$\begin{aligned} t_c &= \frac{d}{\mu E_0} \ln \left(\frac{E_0 + E_1}{E_1} \right) = \\ &= \frac{d^2}{2\mu(V_{FD} + V_{bi})} \ln \left[\frac{2(V_{FD} + V_{bi}) + V_R - (V_{FD} + V_{bi})}{V_R - (V_{FD} + V_{bi})} \right] = \\ &= \frac{d^2}{2\mu(V_{FD} + V_{bi})} \ln \left[\frac{V_R + (V_{FD} + V_{bi})}{V_R - (V_{FD} + V_{bi})} \right] \end{aligned} \quad (95)$$

in cui la generica mobilità μ dev'essere sostituita con quella degli elettroni o delle buche, a seconda dei casi.

Esempio. Consideriamo il caso di un tipico sensore, costituito da una giunzione p^+n in silicio avente uno spessore $d = 300 \mu\text{m}$ ed una tensione di svuotamento interno $V_{di} = 30 \text{ V}$. Supponiamo di applicare una tensione esterna in *reverse bias* $V_R = 2V_{di} = 60 \text{ V}$. Possiamo calcolare il tempo di deriva totale per le buche applicando la formula 95:

$$\begin{aligned} t_{c,h} &= \frac{d^2}{2\mu_h V_{di}} \ln \left[\frac{V_R + V_{di}}{V_R - V_{di}} \right] = \frac{(3.00 \cdot 10^{-2} \text{ cm})^2}{2 \cdot 450 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1} \cdot 30 \text{ V}} \ln \left[\frac{60 \text{ V} + 30 \text{ V}}{60 \text{ V} - 30 \text{ V}} \right] = \\ &= \frac{(3.00)^2}{2 \cdot 4.50 \cdot 3.0} \ln \left(\frac{90}{30} \right) 10^{(-4-2-1)} \text{ cm}^{(2-2)} \text{ V}^{(+1-1)} \text{ s} = 0.33 \ln(3) 10^{-7} \text{ s} = \\ &= 0.36 \cdot 10^{-7} \text{ s} = 36 \cdot 10^{-9} \text{ s} = 36 \text{ ns} \end{aligned}$$

Per gli elettroni inoltre, ricaviamo immediatamente che:

$$t_{c,e} = \frac{\mu_h}{\mu_e} t_{c,h} = \frac{t_{c,h}}{3} = 12 \text{ ns}$$

Osserviamo che si tratta di tempi decisamente più bassi rispetto a quelli calcolati, nell'esempio precedente, nel caso di svuotamento parziale, dove avevamo $t_e = 31.5$ ns e $t_h = 94.5$ ns.

Nel caso dei rivelatori, si vuole **minimizzare** il tempo di raccolta del segnale, e quindi si opera in regime di **sovrasvuotamento**, aumentando la tensione a valori $V_R > V_{FD}$. Come rappresentato nel grafico in fig. 33, esiste un limite superiore, rappresentato dalla **tensione di rottura**, o di *breakdown*, V_{BD} , oltre il quale la corrente cresce esponenzialmente. Dunque la regione operativa del sensore è limitata all'intervallo di tensioni $V_{FD} < V_R < V_{BD}$.

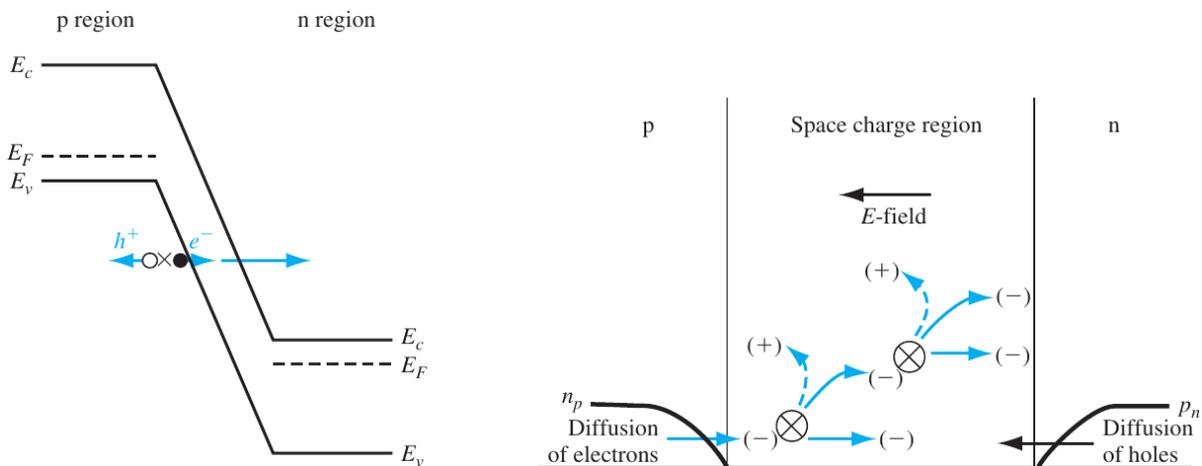


Figura 40: Grafici delle bande per l'effetto Zener (a sinistra) e rappresentazione schematica dell'effetto valanga (a destra).

Il *breakdown* può avvenire secondo due modalità:

- per **effetto Zener**, rappresentato nel grafico a sinistra di fig. 40. Esso avviene poiché l'eccessiva tensione di sovrasvuotamento porta la BV del lato *p* molto vicina alla BC del lato *n*: la natura quantistica dell'elettrone prevede che esso possa passare, con una certa probabilità, dalla prima alla seconda per **effetto tunnel**. Tale meccanismo non è controllabile, e si crea dunque un passaggio di corrente attraverso la giunzione;
- per **effetto valanga**, rappresentato nel grafico a destra di fig. 40. In questo caso, l'intensità del campo elettrico entro la zona svuotata è molto alta. Può capitare che un elettrone di valenza nel lato *p* entri,

per diffusione, nella zona svuotata, oppure che in essa si generi una coppia elettrone-buca per effetto termico. Quando ciò accade, queste cariche subiscono una forte accelerazione a causa del campo molto elevato. L'elettrone, migrando verso il lato n , acquisisce tanta energia da poter ionizzare gli atomi del materiale che incontra sulla sua traiettoria, creando così altre coppie elettrone-buca. Queste cariche subiscono lo stesso processo, instaurando un meccanismo "a valanga" non controllabile, con conseguente passaggio di corrente lungo la giunzione.

Il *breakdown* è un fenomeno a soglia che si verifica per campi elettrici di intensità $E \geq E_C$, dove E_C viene definito **campo critico**. Possiamo calcolare la tensione di rottura V_{BD} in funzione di E_C attraverso la relazione 73 (al quadrato), nella quale approssimiamo lo spessore x_n come in formula 64, considerando $V_R + V_{bi} \simeq V_{BD}$ come tensione totale. Otteniamo:

$$E_C^2 = \frac{q^2 N_D^2}{\varepsilon^2} x_n^2 = \frac{q^2 N_D^2}{\varepsilon^2} \frac{2\varepsilon V_{BD}}{q N_D} = \frac{2q N_D}{\varepsilon} V_{BD} \quad \Rightarrow \quad V_{BD} = \frac{\varepsilon E_C^2}{2q N_D} \quad (96)$$

Come visibile nel grafico a sinistra in fig. 41, il campo critico cresce all'aumentare della concentrazione dei portatori minoritari N_D . La rottura avviene per effetto valanga fino a concentrazioni di drogaggio $N_D \lesssim 10^{17} \text{ cm}^{-3}$. L'effetto Zener (*tunneling* quantistico) prevale invece quando esse raggiungono valori più elevati, ovvero per $N_D \gtrsim 10^{18} \text{ cm}^{-3}$.

Descriviamo meglio l'**effetto valanga**. Consideriamo una corrente iniziale I_{n0} causata dagli elettroni che riescono ad entrare nella zona svuotata dal lato p per moto di diffusione. A causa delle ionizzazioni a valanga, questa corrente $I_n(x)$ aumenta lungo la coordinata x della giunzione, raggiungendo un massimo in $x = W$, come visibile nel grafico a destra in fig. 41. Parallelamente, si generano altrettante buche che migrano verso il lato p della giunzione, dando origine ad una corrente $I_p(x)$, speculare rispetto alla precedente. Avremo una corrente totale $I = I_n(x) + I_p(x)$ che fluisce lungo la giunzione, **costante** in ogni punto x . Definiamo quindi il **fattore di moltiplicazione** M tale che:

$$I_n(W) = M I_{n0} \quad (97)$$

Si tratta di una relazione particolarmente utile quando siamo in grado di operare separatamente con le correnti delle due specie. In tal caso, ad esempio, l'ultima relazione può essere utilizzata nel caso si voglia **amplificare** la

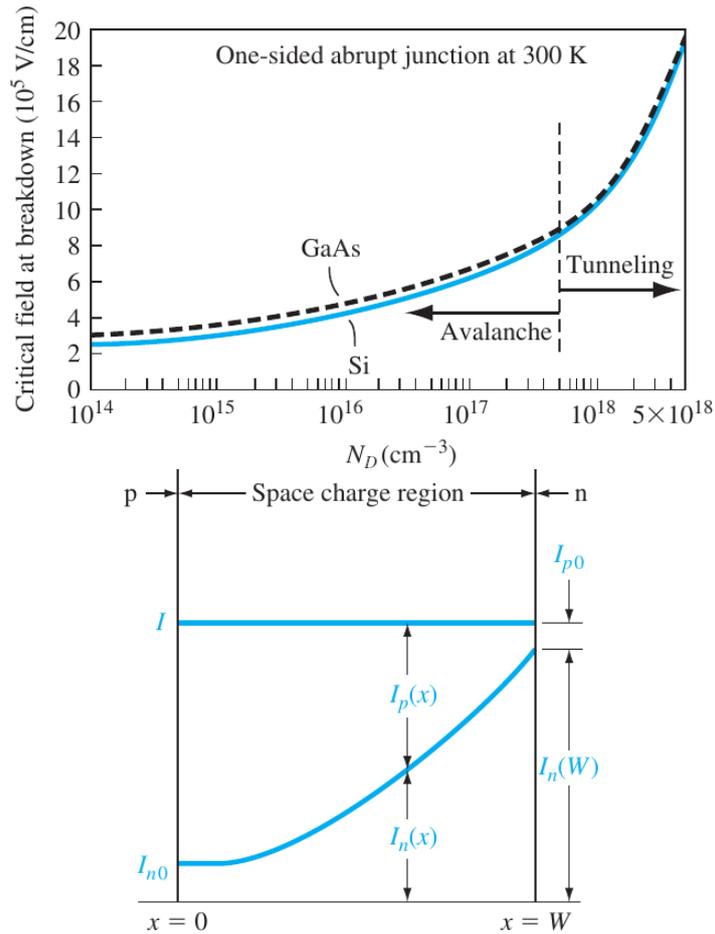


Figura 41: A sinistra: campo critico in funzione della concentrazione dei portatori minoritari, per una giunzione p^+n a temperatura ambiente. A destra: profilo della corrente per effetto valanga nella regione di svuotamento.

corrente degli elettroni che migrano dal lato p verso n .

4 Raccolta della carica e induzione del segnale

4.1 Raccolta della carica

Molto spesso i rivelatori di nostro interesse vengono operati in regime di **sovrasvuotamento**. Il potenziale di polarizzazione inverso applicato, detto anche *reverse bias*, viene indicato con V_R o V_b . Indicando con $V_{di} = V_{FD} + V_{bi}$ il potenziale di svuotamento interno (formula 76), il **tempo di raccolta** delle cariche, ricavato in formula 95, può dunque essere riscritto, separatamente per elettroni e buche, come:

$$t_{c,e} = \frac{d^2}{2\mu_e V_{di}} \ln \left(\frac{V_b + V_{di}}{V_R - V_{di}} \right) \quad \text{e} \quad t_{c,h} = \frac{d^2}{2\mu_h V_{di}} \ln \left(\frac{V_b + V_{di}}{V_R - V_{di}} \right) \quad (98)$$

4.2 Segnale indotto

Ci chiediamo ora quale sia l'**istante iniziale** in cui si osserva la corrente di segnale in lettura, ovvero l'informazione reale che posso trasportare fuori dal rivelatore e usare come dato. Come vedremo, non e' necessari aspettare il momento in cui le cariche raggiungono l'elettrodo di raccolta per ottenere il primo segnale in lettura: si osserva una corrente di segnale già nell'istante in cui le cariche iniziano a muoversi nella giunzione.

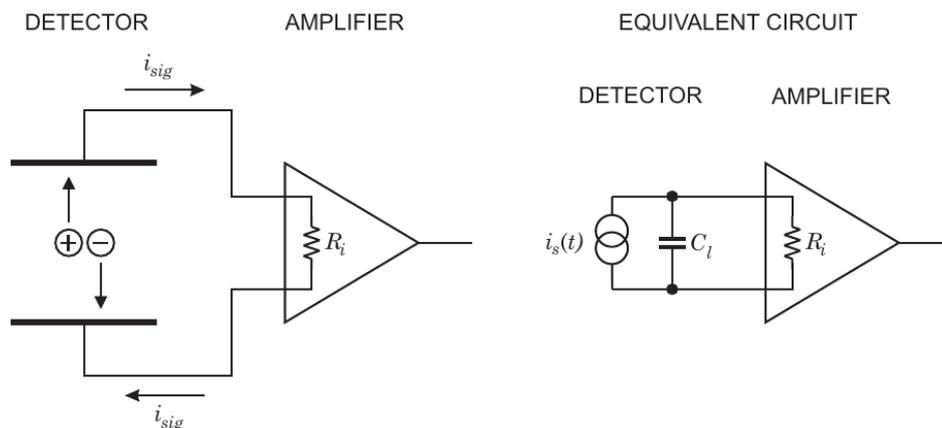


Figura 42: Il movimento delle cariche nel rivelatore induce una corrente di segnale nel circuito di acquisizione (a sinistra). Viene riportato, a destra, anche il circuito equivalente.

Per analizzare il problema, consideriamo il circuito schematizzato in fig. 42. Nella giunzione si originano le cariche che generano la **corrente di segnale** $i_s(t)$, la quale poi arriva all'**amplificatore** connesso alla cella di rivelazione per poi essere trasmessa al resto della catena del sistema di acquisizione. Nella schematizzazione del circuito, sono visibili esplicitamente anche la capacità di carico, o *load capacitance*, C_l , ovvero la capacità equivalente introdotta dalla giunzione stessa, e la **resistenza interna** R_i dell'amplificatore, che come vedremo giocherà un ruolo importante nel processo di amplificazione e trasmissione del segnale.

Consideriamo una carica puntiforme q che compie uno spostamento Δx lungo la coordinata x della giunzione, il cui spessore di svuotamento è W . Viene allora **indotta istantaneamente** una carica ΔQ sull'elettrodo proporzionale alla distanza percorsa dalla carica nella giunzione, ovvero tale che:

$$\Delta Q : q = \Delta x : W \quad \Longrightarrow \quad \Delta Q = q \frac{\Delta x}{W} \quad (99)$$

L'induzione della carica sull'elettrodo è spiegata come effetto del flusso del campo elettrico generato dalla carica originale in movimento. Considero inizialmente una carica puntiforme q nelle vicinanze di un **elettrodo planare infinitamente esteso**. Le linee di campo che escono dalla carica puntiforme, entrano tutte nell'elettrodo. Il **teorema di Gauss** ci dice che dal calcolo del flusso del campo elettrico \vec{E} attraverso una superficie chiusa S che include unicamente la carica puntiforme q , otteniamo:

$$\oint_S \vec{E} \cdot d\vec{s} = \frac{q}{\epsilon} \quad (100)$$

Se invece integriamo su una superficie chiusa che contiene solo l'elettrodo, il flusso del campo è lo stesso del caso precedente, ma il verso è entrante, per cui otteniamo $-q$, ovvero la carica indotta sull'elettrodo.

Consideriamo invece il caso in cui gli elettrodi siano due. Se la carica puntiforme q è posta equidistante dai due elettrodi (come a sinistra in fig. 43), il flusso del campo attraverso le superfici S_1 ed S_2 è uguale (i.e. lo stesso numero di linee di campo terminano su ciascuno dei due elettrodi), e la carica indotta su ciascuno di essi è $-q/2$. Se invece la carica puntiforme q è posta più vicina all'elettrodo inferiore (come a destra in fig. 43), un maggior numero di linee di campo termina su quest'ultimo, dunque il flusso del campo attraverso la superficie gaussiana S_2 è più grande di quello attraverso S_1 , ed anche la

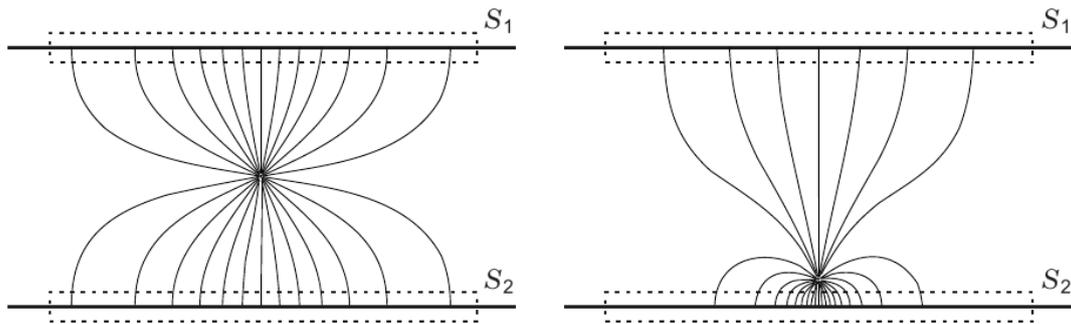


Figura 43: Flusso del campo elettrico generato da una carica puntiforme q posta fra due elettrodi piani paralleli ed infinitamente estesi. Vediamo che esso varia in funzione della posizione di q .

carica indotta su questo elettrodo è maggiore di quella indotta sull'elettrodo superiore.

Se connettiamo i due elettrodi in un circuito, siamo in grado di osservare e misurare una variazione di carica indotta Δq sugli elettrodi, e dunque il passaggio di una **corrente elettrica** nel circuito.

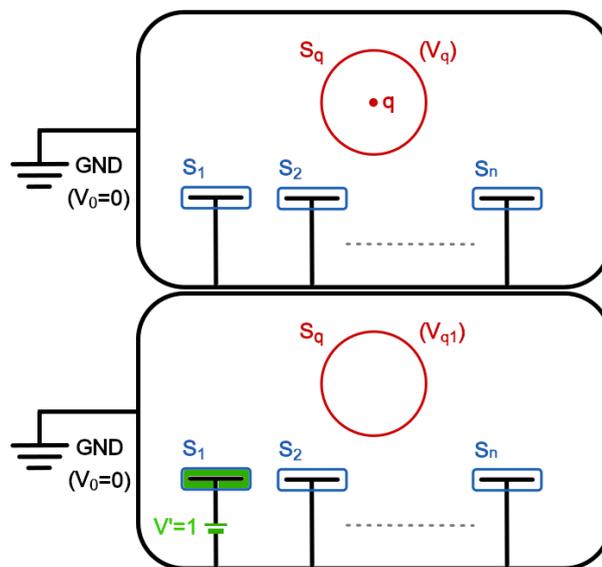


Figura 44: Schemi circuitali per il teorema di Ramo. A sinistra: n elettrodi messi a terra, ed una carica puntiforme q . A destra: viene rimossa la carica puntiforme ed il primo elettrodo è elevato al potenziale $V' = 1$, introducendo un generatore di tensione costante.

Quanto visto è formalizzato nel **teorema di Ramo**. Dimostrazione dettagliata e' reperibile nei libri di testo.

In breve, si calcola la carica Q indotta su un elettrodo dalla presenza di una carica puntiforme q come:

$$Q = -q\Phi \quad (101)$$

con il potenziale peso Φ definito nello spazio fissando ad 1 il potenziale sull'elettrodo considerato, ed a 0 quello sugli elettrodi rimanenti. Definendo inoltre il **campo peso** (*weighting field*) $\vec{E}_w = -\vec{\nabla}\Phi$, possiamo calcolare la **corrente** i che viene indotta sull'elettrodo come:

$$i = \frac{dQ}{dt} = \frac{d}{dt}(-q\Phi) = -q\frac{d\Phi}{dt} = -q\frac{d\vec{x}}{dt} \cdot \vec{\nabla}\Phi = q\vec{v} \cdot \vec{E}_w \quad (102)$$

Nel caso in cui la carica q compia un moto da \vec{x}_1 ad \vec{x}_2 , la carica indotta sull'elettrodo sarà:

$$\Delta Q = -q[\Phi(\vec{x}_2) - \Phi(\vec{x}_1)] \quad (103)$$

Osserviamo complessivamente che:

- il **campo elettrico** \vec{E} determina la **traiettoria** e la **velocità** del moto di **deriva** della carica;
- il **campo peso** \vec{E}_w dipende solo dalla **geometria** dell'apparato, e determina come il moto della carica si accoppia ad uno degli elettrodi;
- in generale $\vec{E}_w \neq \vec{E}$. Solo nelle configurazioni a due elettrodi essi hanno la stessa forma.

Consideriamo il caso delle giunzioni semplici fin ora viste. Si tratta di una configurazione a due elettrodi, schematizzata a sinistra in fig. 45. Vediamo che il campo elettrico è uniforme, ed il campo peso ha la stessa direzione, con rispettivi moduli:

$$E = \frac{V_b}{d} \quad \text{ed} \quad E_w = \frac{1}{d} \quad (104)$$

Una carica puntiforme q compie fra gli elettrodi un moto di deriva lungo la direzione del campo elettrico, con velocità calcolabile come:

$$v = \mu E = \mu \frac{V_b}{d} \quad (105)$$

Attraverso la formula 102 ricaviamo la corrente circolante fra gli elettrodi:

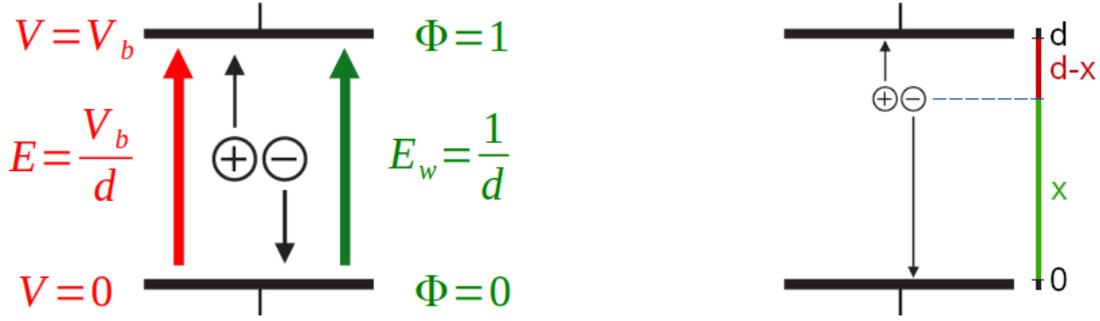


Figura 45: A sinistra: schema della configurazione a due elettrodi piani paralleli, posti a distanza d . Vengono riportati i valori di potenziale e potenziale peso, assieme al campo elettrico e al campo peso. A destra: schema del sistema di riferimento per il moto di deriva delle cariche all'interno degli elettrodi.

$$i = qvE_w = q\mu \frac{V_b}{d} \frac{1}{d} = \frac{q\mu V_b}{d^2} \quad (106)$$

Osserviamo che si tratta di una corrente costante, che circola per la durata del tempo di raccolta della carica q , ovvero tra l'istante in cui essa viene rilasciata e quello in cui essa giunge all'elettrodo, ricombinandosi. Il tempo di raccolta massimo è quello in cui la carica si origina su un elettrodo, e deriva lungo tutto lo spessore d della giunzione:

$$t_c = \frac{d}{v} = \frac{d}{\frac{\mu V_b}{d}} = \frac{d^2}{\mu V_b} \quad (107)$$

La carica indotta sull'elettrodo si ottiene integrando la corrente sul tempo di raccolta:

$$Q = it_c = \frac{q\mu V_b}{d^2} \frac{d^2}{\mu V_b} = q \quad (108)$$

Notiamo che l'azione **a distanza** del meccanismo di induzione della carica, permette di ottenere un segnale anche nel caso in cui q **non raggiunga l'elettrodo**. In tal caso **avremo comunque un segnale**, seppur incompleto (i.e. viene indotta una carica Q più piccola all'elettrodo). Ciò può accadere, come vedremo meglio in seguito, a causa delle **trappole**, ossia di impurità nella struttura cristallina, atomi dislocati, siti vacanti, che possono inibire il moto delle cariche libere.

Vediamo il caso generico, in cui il rilascio di carica avviene in un generico punto x entro la zona svuotata, come riportato nello schema a destra in fig. 45. Elettroni e buche migrano in direzioni opposte, con **tempi di raccolta**:

$$t_{c,e} = \frac{x}{v_e} = \frac{x}{\mu_e \frac{V_b}{d}} = \frac{xd}{\mu_e V_b} \quad \text{e} \quad t_{c,h} = \frac{d-x}{v_h} = \frac{d-x}{\mu_h \frac{V_b}{d}} = \frac{(d-x)d}{\mu_h V_b} \quad (109)$$

Osserviamo che i due tipi di portatori hanno carica di segno opposto, e muovendosi anche in direzioni opposte, contribuiscono ad una corrente netta che fluisce tutta nella stessa direzione. Indichiamo con i_e la corrente dovuta agli elettroni che migrano verso l'elettrodo n , ivi inducendo una carica Q_e , mentre con i_h indichiamo quella dovuta alle buche, che migrano verso l'elettrodo p , ivi inducendo una carica Q_h . Avremo dunque:

$$\begin{aligned} Q_e &= i_e t_e = \frac{q\mu_e V_b}{d^2} \frac{xd}{\mu_e V_b} = q \frac{x}{d} \\ Q_h &= i_h t_h = \frac{q\mu_h V_b}{d^2} \frac{(d-x)d}{\mu_h V_b} = q \frac{d-x}{d} = q \left(1 - \frac{x}{d}\right) \end{aligned} \quad (110)$$

Osserviamo che **in ogni istante la carica indotta su ciascun elettrodo è diversa**, poiché diversa è la velocità di deriva per i due tipi di cariche in moto, sebbene quella totale rimanga sempre $q = Q_e + Q_h$.

Per apprezzare questo comportamento, consideriamo il caso semplificato di una **carica rilasciata a metà dello spessore svuotato**, ossia in $x = d/2$. Calcoliamo il tempo di deriva degli elettroni:

$$t_{c,e} = v_e^{-1} \frac{d}{2} = \frac{d}{\mu_e V_b} \frac{d}{2} = \frac{d^2}{2\mu_e V_b} \quad (111)$$

Utilizzando la formula 106 per la corrente indotta, possiamo vedere che, in questo tempo, la carica indotta sull'elettrodo n sarà:

$$Q_e(t_{c,e}) = i_e t_{c,e} = \frac{q\mu_e V_b}{d^2} \frac{d^2}{2\mu_e V_b} = \frac{q}{2} \quad (112)$$

come ci aspettavamo. Nello stesso tempo, la carica indotta all'elettrodo p sarà invece:

$$Q_h(t_{c,e}) = i_h t_{c,e} = \frac{q\mu_h V_b}{d^2} \frac{d^2}{2\mu_e V_b} = \frac{\mu_h q}{\mu_e 2} = \frac{1}{3} \frac{q}{2} = \frac{q}{6} \quad (113)$$

Il tempo di raccolta delle buche è:

$$t_{c,h} = v_h^{-1} \frac{d}{2} = \frac{d}{\mu_h V_b} \frac{d}{2} = \frac{d^2}{2\mu_h V_b} \quad (114)$$

per cui posso calcolare il tempo $\Delta t = t_{c,h} - t_{c,e}$ necessario per raccogliere le buche rimanenti dopo che gli elettroni sono già stati raccolti, come:

$$\begin{aligned} \Delta t = t_{c,h} - t_{c,e} &= \frac{d^2}{2\mu_h V_b} - \frac{d^2}{2\mu_e V_b} \frac{\mu_h}{\mu_h} = \frac{d^2}{2\mu_h V_b} \left(1 - \frac{\mu_h}{\mu_e}\right) \simeq \\ &\simeq \frac{d^2}{2\mu_h V_b} \left(1 - \frac{1}{3}\right) = \frac{d^2}{2\mu_h V_b} \frac{2}{3} = \frac{d^2}{3\mu_h V_b} \end{aligned} \quad (115)$$

Posso quindi calcolare la carica positiva ancora da raccogliere come:

$$\Delta Q_h = i_h \Delta t = \frac{q\mu_h V_b}{d^2} \frac{d^2}{3\mu_h V_b} = \frac{q}{3} \quad (116)$$

Possiamo verificare che la carica totale positiva raccolta, sarà:

$$Q_h(t_{c,h}) = Q_h(t_{c,e}) + \Delta Q = \frac{q}{3} + \frac{q}{6} = \frac{q}{2} \quad (117)$$

Se devo quindi fare una misura accurata del **rilascio energetico totale**, devo utilizzare un tempo d'integrazione almeno uguale a quello di raccolta delle buche. Viceversa, se sono interessato a misurare solo il **tempo di arrivo** del segnale, sarà sufficiente integrare sul tempo di raccolta dei soli elettroni, che danno un segnale molto piccato rispetto a quello dovuto alle buche: la risoluzione temporale sarà migliore, a scapito dell'informazione sulla carica totale rilasciata.

Notiamo che le correnti indotte sui due elettrodi hanno stesso modulo ma polarità opposta, ovvero $i_1 = -i_2$. Più in generale, vale l'equazione di **continuità** (legge di Kirchhoff). Essa afferma che, in ogni punto del circuito, la somma delle correnti indotte su ciascun k -esimo elettrodo è nulla, ovvero:

$$\sum_k i_k = 0 \quad (118)$$

Nel caso in cui ci siano più elettrodi, come nei rivelatori a **microstrisce**, la corrente indotta su un elettrodo deve compensare la somma di quelle indotte su tutti gli altri.

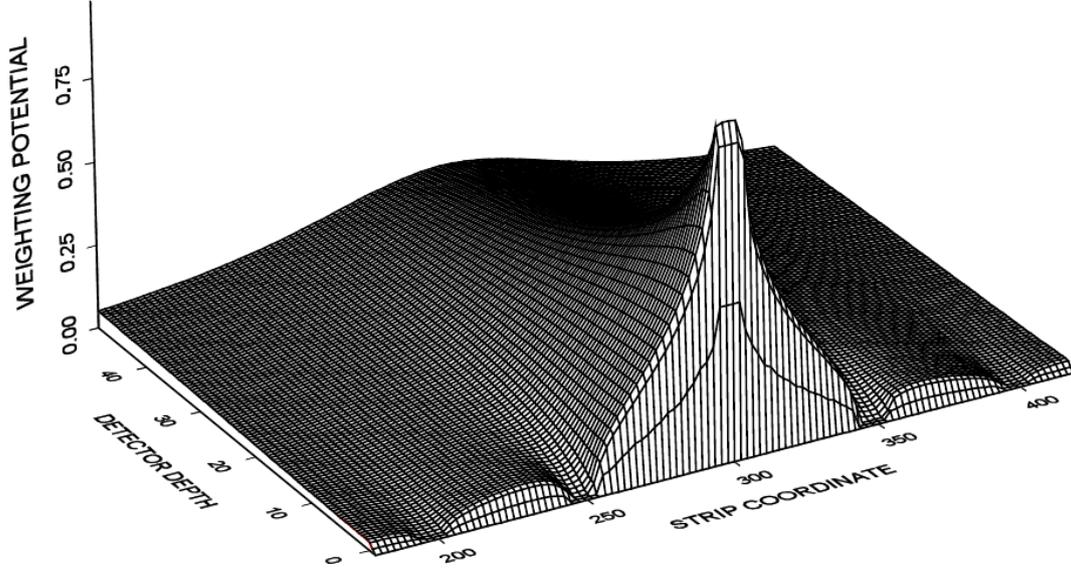


Figura 46: Potenziale peso di un rivelatore a *microstrip* di silicio avente spessore $d = 300 \mu\text{m}$ e *pitch* $w = 50 \mu\text{m}$. La striscia centrale è a potenziale unitario, mentre le altre sono a zero. Viene mostrata la profondità solo fino a $50 \mu\text{m}$ nel materiale.

Questo tipo di sensori è caratterizzato da una distanza w fra i centri delle strisce, detta *pitch*, oltre che dallo spessore d della giunzione. Per avere un buon comportamento del sensore, richiediamo che il *pitch* sia minore dello spessore, tipicamente $w \leq d/3$. Immaginiamo che il sensore abbia entrambi gli elettrodi p ed n segmentati in strisce sovrapposte e parallele fra loro. All'interno di esse, tranne che in prossimità degli elettrodi, possiamo approssimare il campo elettrico come uniforme, con un potenziale che varia linearmente lungo la coordinata perpendicolare alla giunzione. Ben diverso è invece il potenziale peso, che è illustrato nel grafico di fig. 46.

Consideriamo un rilascio di carica nel punto di coordinata x_0 lungo la perpendicolare alla k -esima striscia. In questo punto si origina una coppia elettrone-buca, con cariche $-q$ e $+q$, rispettivamente. Se x_p ed x_n sono le coordinate dei due elettrodi, possiamo calcolare la carica Q_k indotta, usando l'equazione 103:

$$\begin{aligned}
 Q_k &= q[\phi_k(x_p) - \phi_k(x_0)] - q[\phi_k(x_n) - \phi_k(x_0)] = \\
 &= q[\phi_k(x_p) - \phi_k(x_0) - \phi_k(x_n) + \phi_k(x_0)] = \\
 &= q[\phi_k(x_p) - \phi_k(x_n)]
 \end{aligned}
 \tag{119}$$

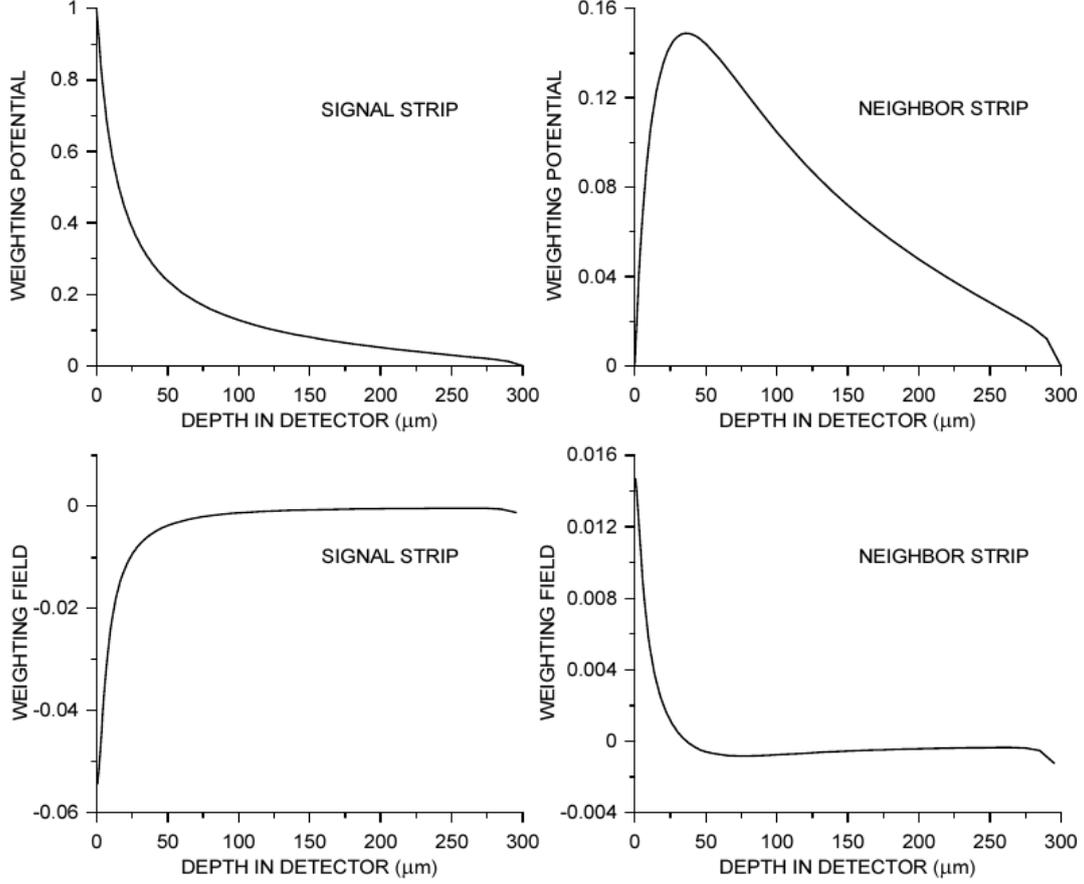


Figura 47: Profili del potenziale peso (in alto) e del campo peso (in basso) lungo il piano verticale a metà larghezza della striscia. A sinistra, per la striscia k -esima, su cui viene letto il segnale, mentre a destra per la sua vicina $(k + 1)$ -esima. Gli elettrodi p hanno coordinata $x_p = 0 \mu\text{m}$, mentre quelli n sono in $x_n = 300 \mu\text{m}$. Poniamo ad 1 il potenziale peso dell'elettrodo p della k -esima striscia, 0 per tutti gli altri.

Supponiamo di leggere il segnale proveniente dall'elettrodo p della k -esima striscia, la quale ha un profilo di potenziale peso come nel grafico in alto a sinistra di fig. 47. Osservando che $\phi_k(x_p) = 1$ e $\phi_k(x_n) = 0$, dalla precedente relazione, otteniamo:

$$Q_k = q[1 - 0] = q \quad (120)$$

Se invece misuriamo il segnale proveniente dall'elettrodo vicino $(k + 1)$ -esimo, il potenziale peso è quello riportato nel grafico in alto a destra di fig. 47, per cui, siccome $\phi_{k+1}(x_p) = \phi_{k+1}(x_n) = 0$, otteniamo:

$$Q_{k+1} = q[0 - 0] = 0 \quad (121)$$

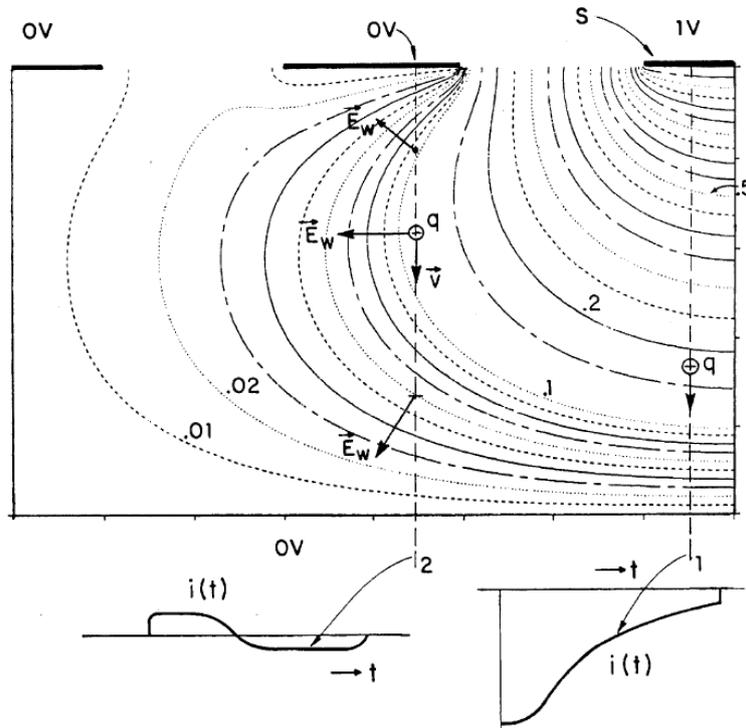


Figura 48: Rappresentazione del campo peso (linee di campo) per un sensore a microstrisce, visto di profilo. La striscia in lettura è quella più a destra. In basso, viene riportata la corrente indotta per l'elettrodo in lettura e per il suo vicino immediatamente a sinistra.

In generale, come rappresentato nel grafico in fig. 48, vediamo che quando una carica non viene raccolta all'elettrodo su cui effettuiamo la misura, su di esso viene comunque indotta una corrente, che però cambia segno durante il moto della carica nella giunzione. Il suo integrale nel tempo, che rappresenta la carica raccolta su tale elettrodo, è nullo.

Essendo $i \propto E_w$ (dalla formula 102), e in particolare al prodotto scalare del vettore velocità della carica con il vettore campo peso, possiamo analizzare il problema anche in termini di campo peso. Come visibile nei due grafici in basso in fig. 47, il campo peso della striscia k -esima, che è quella in lettura, presenta un andamento crescente sul percorso di deriva delle cariche, dunque l'integrale su tale cammino è non nullo. Ciò non accade per la striscia vicina, $(k + 1)$ -esima, che non viene letta: qui il campo peso cambia segno lungo il percorso, ed il suo integrale su tale cammino è nullo.

Quanto visto è vero se entrambi i portatori di carica vengono raccolti. Se, ad esempio, non siamo in grado di raccogliere le buche, l'equazione 119 si riduce a:

$$Q_k = -q[\phi_k(x_n) - \phi_k(x_0)] = q[\phi_k(x_0) - \phi_k(x_n)] \quad (122)$$

e per la strip non in lettura:

$$Q_{k+1} = -q[\phi_{k+1}(x_n) - \phi_{k+1}(x_0)] = q[\phi_{k+1}(x_0)] \quad (123)$$

Si tratta di un contributo, in generale, non nullo, essendo $\phi_k(x_0) \neq 0$, motivo per cui osserveremo comunque un segnale, seppur debole.

5 Elaborazione del segnale

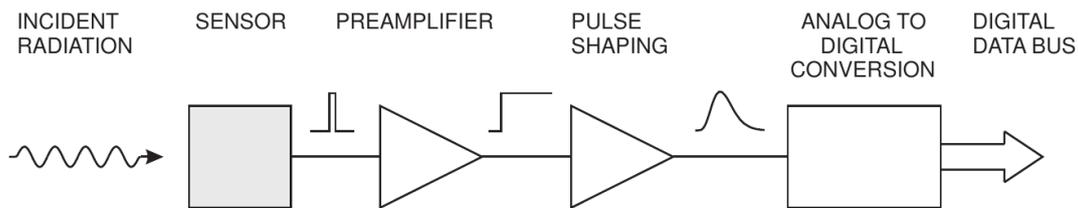


Figura 49: Schema delle funzionamento di base di un rivelatore: la radiazione incidente produce un segnale elettrico, che viene processato attraverso un preamplificatore ed un formatore, per venire infine digitalizzato, attraverso un ADC, e salvato poi in memoria. Si noti la tipica forma assunta dal segnale dopo ogni stadio di elaborazione.

Vediamo ora cosa succede al segnale dopo la raccolta della carica. Di solito, i segnali in uscita dal sensore hanno generalmente una durata dell'ordine dei ns. Nei casi in cui si punti ad un segnale con alta risoluzione temporale, essi possono scendere fino a durate dell'ordine di 10 ps. In generale, essi hanno comunque un'ampiezza molto piccola, e necessitano dunque di **amplificazione** e di una forma che ne permetta il trattamento successivo. La catena di trasporto ed elaborazione del segnale induce una fluttuazione nel numero di cariche che compongono il segnale, che corrisponde al rumore elettronico.

Come vediamo nello schema di fig. 49, il primo stadio di trattamento del segnale è costituito dal **preamplificatore**, che viene posto quanto più vicino

possibile al sensore, per limitare il piu' possibile dispersioni e deformazioni prima di essere trattato. Qui il segnale viene amplificato, potenzialmente trasformandone la natura da impulso di cariche a differenza di potenziale, possibilmente aumentandone l'ampiezza per un fattore di guadagno.

Successivamente, il segnale viene trasferito al **formatore**, il cui scopo è di conferire una forma che sia vantaggiosa per confrontare segnali fisici con segnali di puro rumore, e per filtrare questi ultimi, sia per rendere i segnali facilmente misurabili e campionabili dall'elettronica di acquisizione a valle. Il segnale acquisisce una forma, con un fronte di salita ed uno di discesa ben definiti nel dominio del tempo, corrispondenti a caratteristiche precise nel dominio delle frequenze. La riduzione della banda di frequenze per limitare il contributo totale del rumore, infatti, porta necessariamente a sopprimere le alte frequenze, e quindi ad estendere la durata dei segnali fisici in modo da renderli distinguibili da quelli di rumore. D'altro canto, la necessita' di rivelare segnali con alta frequenza porta a limitare la durata del segnale per evitare che segnali successivi sullo stesso canale si sovrappongano e sommino le proprie cariche perdendo parte dell'informazione (*pile-up*). Il formatore garantisce anche che il segnale cosi' formato mantenga il proprio valore massimo per una certa durata senza singolarita', in modo da poter misurato negli stadi successivi.

Il segnale viene infine misurato (nel valore massimo o con un suo campionamento) dall'**ADC** (*Analog to Digital Converter*), costituito nella forma piu' semplice da una serie di comparatori, di discriminatori a soglia via via più elevata. In tal modo vengono determinati degli intervalli di ampiezza, fra una soglia e quella immediatamente successiva, che determinano anche la risoluzione su questa misura.

5.1 Fluttuazioni del segnale e rumore elettronico

L'osservabile misurata contiene delle incertezze intrinseche che compaiono prima ancora di trattare il segnale:

- abbiamo parlato per esempio dell'energia persa per unità di percorso nei semiconduttori, che non è un valore deterministico ma statistico, con distribuzione anche non gaussiana a cui è associata una certa larghezza, che si può utilizzare come stima dell'incertezza;
- abbiamo osservato come tale distribuzione varia per materiali sottili, modificandone anche l'incertezza sulla misura;
- inoltre la risoluzione energetica di un rivelatore, teoricamente definita come il rapporto R tra fluttuazione ΔE dell'energia persa dalla particella ed il valore medio E di tale energia ($R = \Delta E/E$), beneficia del fatto che la fluttuazione nel numero di cariche generate è minore della semplice varianza statistica $\sigma_Q = \sqrt{FN_Q}$, con $F < 1$ detto Fattore di Fano. F è specifico del materiale e della tecnologia usati nel rivelatore, e vale $F \sim 0.1$ nel caso di silicio, germanio, gallio-arsenio e diamante.
- infine, la carica così generata, comprensiva della fluttuazione di cui l'evento in questione è affetto, subisce un'ulteriore modifica che è effetto del cosiddetto rumore elettronico, ovvero della variazione del numero e della velocità degli elettroni durante la raccolta;
- il rumore elettronico influenzerà poi il segnale anche nella catena di componenti elettronici a valle del sensore.

5.2 Fattore di Fano

Senza darne dimostrazione completa, riassumiamo qui l'origine delle fluttuazioni corrette tramite il fattore di Fano.

Come rappresentato nella parte sinistra di fig. 50, se una radiazione monocromatica viene completamente assorbita nel rivelatore, produce un segnale di risposta con spettro non monocromatico: il numero di cariche generate ha una distribuzione estesa. Con un **grande numero** di collisioni **indipendenti** tra la radiazione incidente e il materiale, la distribuzione del segnale generato è **gaussiana** (per il **teorema del limite centrale**) e centrata sul valore V_0 corrispondente all'energia iniziale E_0 , ma con una larghezza finita,

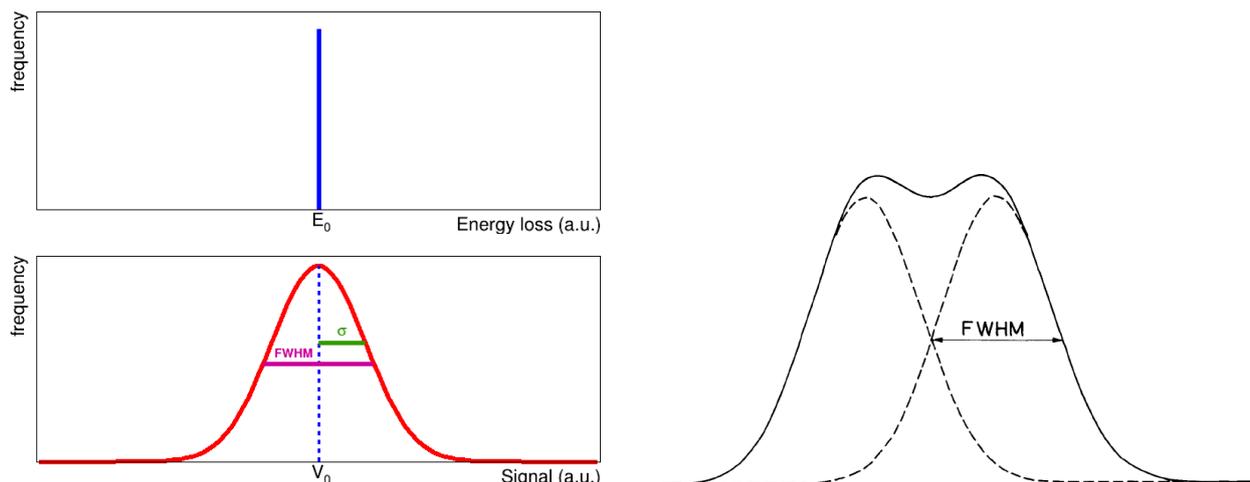


Figura 50: A sinistra: in basso, la risposta del rivelatore in seguito al completo assorbimento di una radiazione monocromatica, il cui spettro è riportato in alto. A destra: due segnali sono risolvibili se separati da almeno una FWHM, come nel caso qui riportato, in cui abbiamo due picchi gaussiani (linee tratteggiate) e la loro somma (linea continua).

quantificabile mediante la deviazione standard σ . Nella pratica si utilizza la **larghezza a metà altezza** (FWHM, *Full Width at Half Maximum*), che è utilizzabile anche nel caso in cui la distribuzione non sia gaussiana. Con tale quantità, possiamo definire la **risoluzione** come:

$$R = \frac{\text{FWHM}}{E} \quad (124)$$

Per cui potremo considerare due picchi come **risolvibili** soltanto qualora gli spettri delle loro rispettive distribuzioni siano separati per almeno una FWHM, come rappresentato a destra in fig. 50.

Nel caso della distribuzione **gaussiana**, possiamo calcolare la FWHM in funzione della deviazione standard σ , come:

$$\text{FWHM} = 2\sigma\sqrt{2\ln(2)} \simeq 2.35\sigma \quad (125)$$

Da cui possiamo ricavare la risoluzione energetica:

$$R = \frac{\text{FWHM}}{E} \simeq \frac{2.35\sigma}{E} \quad (126)$$

Osserviamo però che non è corretto trattare le singole collisioni tra la radiazione incidente e le cariche del materiale come indipendenti. La particella incidente viene infatti completamente assorbita nel materiale, e la sua

energia iniziale deve dunque eguagliare la somma di quelle perse in ciascuna collisione. Ciò rappresenta un **vincolo**, per il quale le collisioni risultano essere **correlate** fra loro. Per questo motivo l'adozione della **distribuzione gaussiana** è giustificata solo parzialmente, come **limite teorico**. La **risoluzione teorica** (R_{teor}), è dunque quella già ricavata nella formula 126, ed è diversa da quella **sperimentale** (R_{exp}). Avremo che:

$$R_{exp} = \frac{2.35\sigma}{E} \sqrt{F} < R_{teor} \quad (127)$$

dove F è il **fattore di Fano**, che permette di correggere la risoluzione teorica, ottenendo quella sperimentale. Sperimentalmente, avremo dunque una fluttuazione minore, ovvero una migliore risoluzione, rispetto alla previsione teorica.

Infatti, se il numero di collisioni N è una variabile **poissoniana**, valore medio e deviazione standard saranno, rispettivamente, $\mu = N$ e $\sigma = \sqrt{N}$. Correggendo quest'ultima per il fattore di Fano, si ottiene la deviazione standard misurata $\sigma = \sqrt{NF}$, per cui:

$$\frac{\sigma}{\mu} = \frac{\sqrt{NF}}{N} = \sqrt{\frac{F}{N}} \quad (128)$$

Il **fattore di Fano** può quindi essere definito come:

$$F \equiv \frac{\text{varianza osservata su } N}{\text{varianza predetta su } N} = \frac{\sigma_{exp}^2}{\sigma_{teor}^2} \quad (129)$$

Diamo una breve spiegazione di questo effetto. Quando una particella perde un'energia E nota, produce un numero N di ionizzazioni. Trattando tali eventi come indipendenti, il loro numero dovrebbe fluttuare come \sqrt{N} , seguendo la statistica poissoniana. Nella realtà però questi eventi sono correlati, visto che la somma totale dell'energia persa E è nota, e che c'è una porzione di energia che viene usata per le ionizzazioni ed una per le interazioni con il reticolo, ognuna di questi fenomeni accompagnati dalla propria fluttuazione caratteristica. La correzione introdotta da questa correlazione è appunto il fattore di Fano.

Siccome il numero di cariche generate è $N = E/E_{ion}$, dove E_{ion} è l'energia di ionizzazione media e E la perdita di energia, la **fluttuazione del numero di cariche** generate può essere scritta come:

$$\Delta N = \sqrt{NF} = \sqrt{\frac{EF}{E_{ion}}} \quad (130)$$

La **fluttuazione relativa dell'energia** è dunque legata a quella del numero di cariche, per cui scriveremo:

$$\frac{\Delta E}{E} = \frac{\Delta N}{N} = \frac{\sqrt{EF/E_{ion}}}{E/E_{ion}} = \frac{\sqrt{F}}{\sqrt{E/E_{ion}}} = \sqrt{\frac{FE_{ion}}{E}} \quad (131)$$

e prendendo la larghezza a metà altezza:

$$R = \frac{\Delta E_{FWMH}}{E} = 2.35 \sqrt{\frac{FE_{ion}}{E}} \quad (132)$$

5.3 Introduzione al rumore elettronico

Come sappiamo, il segnale raccolto sugli elettrodi collegati al nostro sistema di trattamento e misura è indotto dal movimento di elettroni per deriva, come previsto dal teorema di Ramo (4). Se consideriamo quindi n elettroni che si muovono tra due elettrodi per un tratto s di circuito, in un tempo t , con velocità media v , otteniamo che sugli elettrodi viene indotta una corrente i , che possiamo esprimere come:

$$i = \frac{nq}{t} = \frac{nvq}{s} \quad (133)$$

La sua fluttuazione media $\langle \delta i \rangle$ dipende sia da quella della velocità v degli elettroni, sia da quella del loro numero n . Queste due componenti non sono correlate statisticamente, quindi le loro fluttuazioni vengono sommate in quadratura:

$$\langle \delta i \rangle^2 = \left(\frac{nq}{s} \langle \delta v \rangle \right)^2 + \left(\frac{qv}{s} \langle \delta n \rangle \right)^2 \quad (134)$$

La fluttuazione δv della velocità dei portatori avviene a causa dell'eccitazione termica di questi ultimi. Si tratta dunque di **rumore termico**, il cui spettro in frequenza è uniforme, per cui è un rumore **bianco**.

La fluttuazione δn del numero di portatori può avvenire per diversi motivi. Uno fra questi, è che la corrente si forma a causa di singoli eventi discreti di generazione e ricombinazione, che non sono correlati gli uni agli altri, ed il cui numero può dunque fluttuare casualmente. Tale componente viene definita

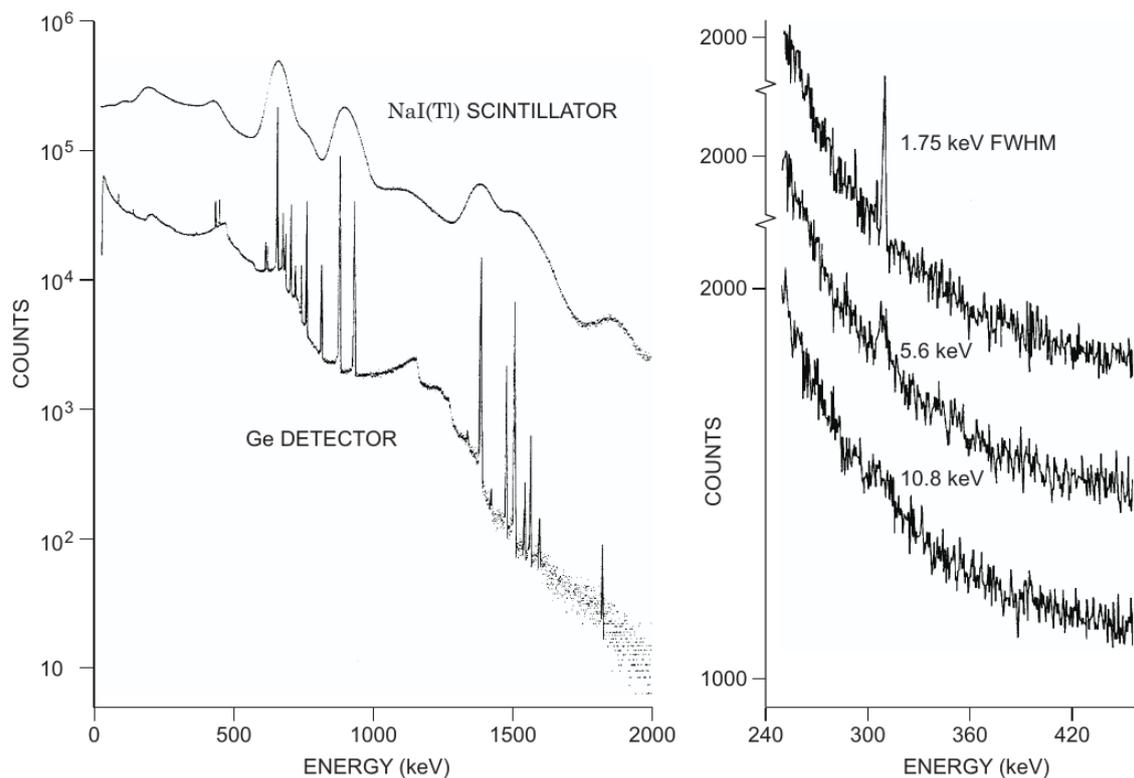


Figura 51: Effetti della risoluzione sulla misura di spettri energetici. A sinistra: confronto fra gli spettri energetici di raggi gamma misurati con un rivelatore a scintillazione ed uno a semiconduttore. La migliore risoluzione di quest'ultimo permette di visualizzare la struttura nel dettaglio, distinguendo i diversi picchi dei rispettivi segnali. A destra: una risoluzione migliore, qui espressa in termini di larghezza a metà altezza (FWHM), incrementa anche il rapporto segnale-rumore. Con risoluzioni migliori è possibile distinguere il picco segnale dal rumore.

rumore shot. Anche in questo caso lo spettro in frequenza è uniforme, e quindi si tratta di rumore **bianco**.

Il numero di cariche può fluttuare anche a causa delle impurità e delle imperfezioni del cristallo. La sovrapposizione di fenomeni di **intrappolamento** con tempi caratteristici diversi (i.e. vite medie), dà luogo ad un **rumore**, in particolare alle **basse frequenze**. Si tratta di rumore **rosa**, poiché il suo spettro è $\propto 1/f^\alpha$, dove f indica la frequenza, mentre l'esponente α è tipicamente nell'intervallo $0.5 \div 2$.

5.4 Quantità fondamentali nella trattazione del rumore

Alle fluttuazioni intrinseche al segnale di cui abbiamo appena parlato, si sommano le fluttuazioni che subisce il segnale nel suo trasferimento e trattamento, che normalmente vengono riassunti con *rumore elettronico*. Introduciamo alcune quantità fondamentali usate per la trattazione di tale rumore.

5.4.1 RMS

Introduciamo brevemente il **valore quadratico medio** (RMS, *root mean square*), o *valore efficace*. Nel caso di un insieme di N valori discreti $\{x_1, \dots, x_N\}$, esso viene definito come la radice della media quadratica, ovvero:

$$x_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (135)$$

Possiamo definirlo anche per una funzione continua del tempo $f(t)$. Nel caso di un intervallo temporale finito $T_1 \leq t \leq T_2$ avremo:

$$f_{\text{RMS}} = \sqrt{\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} [f(t)]^2 dt} \quad (136)$$

Nel caso lo si consideri invece su tutto l'asse temporale, avremo:

$$f_{\text{RMS}} = \lim_{T \rightarrow \infty} \sqrt{\frac{1}{2T} \int_{-T}^T [f(t)]^2 dt} \quad (137)$$

Qualora la funzione venga campionata ad istanti di tempo successivi, come accade nelle nostre misure, si può approssimare l'RMS riconducendosi al caso discreto (formula 135).

5.4.2 Rapporto segnale/rumore

Possiamo ora definire il **rapporto segnale-rumore** (SNR, o S/N, *signal-to-noise ratio*) come:

$$\text{S/N} = \frac{\langle s^2(t) \rangle}{\langle n^2(t) \rangle} \quad (138)$$

ovvero come il rapporto fra la media quadratica del segnale $s(t)$ e quella del rumore $n(t)$.

5.4.3 Rumore espresso in elettroni equivalenti (ENC)

Il rumore si può misurare con la **carica equivalente di rumore** (ENC, *equivalent noise charge*), cioè con il numero di elettroni che producono un segnale equivalente al valore quadratico medio del rumore, ovvero quando il rapporto segnale-rumore è $S/N = 1$. In pratica, consideriamo che nel sensore venga rilasciata una carica Q_S di segnale, e che S/N sia il rapporto segnale-rumore, allora la carica equivalente di rumore sarà:

$$\text{ENC} \equiv Q_n = \frac{Q_S}{S/N} \quad (139)$$

Conoscendo l'energia media di ionizzazione ε_i del materiale di cui è composto il sensore, possiamo ricavare facilmente il rumore in termini energetici:

$$\Delta E_n = \varepsilon_i \text{ ENC} \quad (140)$$

Come vedremo fra poco, risulta conveniente analizzare i singoli contributi al rumore in termini di tensione e corrente: l'ENC viene calcolata sul rumore complessivo, ed in tal senso non è una quantità primitiva, ma derivata.

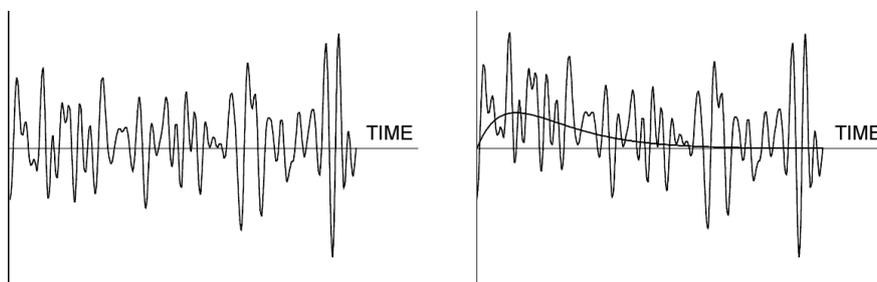


Figura 52: A sinistra: forma d'onda del rumore casuale. A destra: forma d'onda del segnale sovrapposto al rumore. Viene riportata anche la forma del segnale separata dal rumore. Il rapporto segnale-rumore è $S/N = 1$, ovvero l'ampiezza del picco del segnale coincide con l'RMS del rumore.

5.5 Analisi spettrale del rumore

Vediamo ora i metodi per effettuare l'**analisi spettrale del rumore**. Come abbiamo già detto, il rumore è una variabile **casuale**, di cui non possiamo

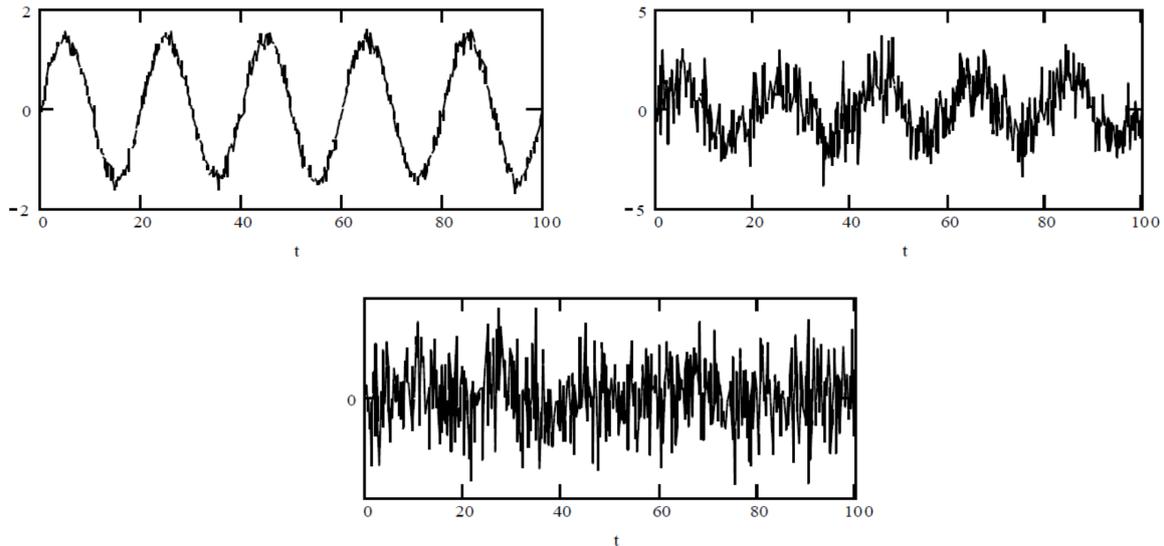


Figura 53: Segnale sinusoidale sommato a rumore gaussiano, per tre diversi valori di rapporto segnale-rumore. Nel grafico in alto a sinistra $S/N = 10$, in quello in alto a destra $S/N = 1$, mentre in quello in basso al centro $S/N = 0.1$. Osserviamo che risulta sempre più difficile distinguere il segnale dal rumore al diminuire di tale rapporto.

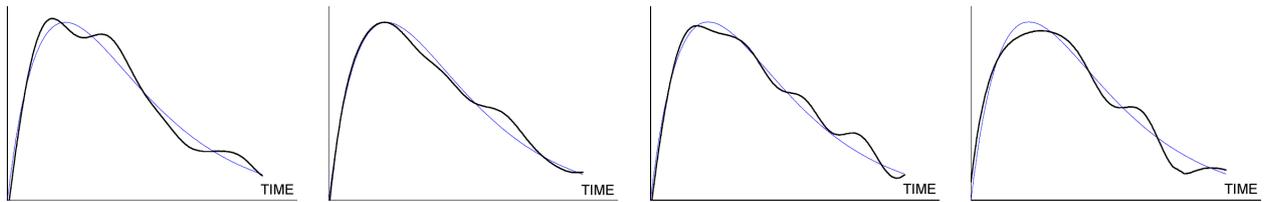


Figura 54: Effetti del rumore sulla risoluzione temporale. Vediamo quattro esempi di segnali affetti rumore, con $S/N = 20$. Per confronto, viene riportato anche il segnale ideale, senza rumore, che è sempre lo stesso. Se il tempo del segnale è dato dal superamento di una soglia fissata in ampiezza, ripetendo la misura otterremo, in generale, tempi diversi.

prevedere l'evoluzione temporale. Possiamo però individuare delle proprietà in senso statistico, si tratta infatti di un processo:

- **stazionario**, ovvero le cui proprietà statistiche rimangono costanti nel tempo;
- **ergodico**, ovvero le cui proprietà statistiche rimangono costanti tra diverse misure (i.e. la media su tutte le possibili realizzazioni coincide con la media temporale).

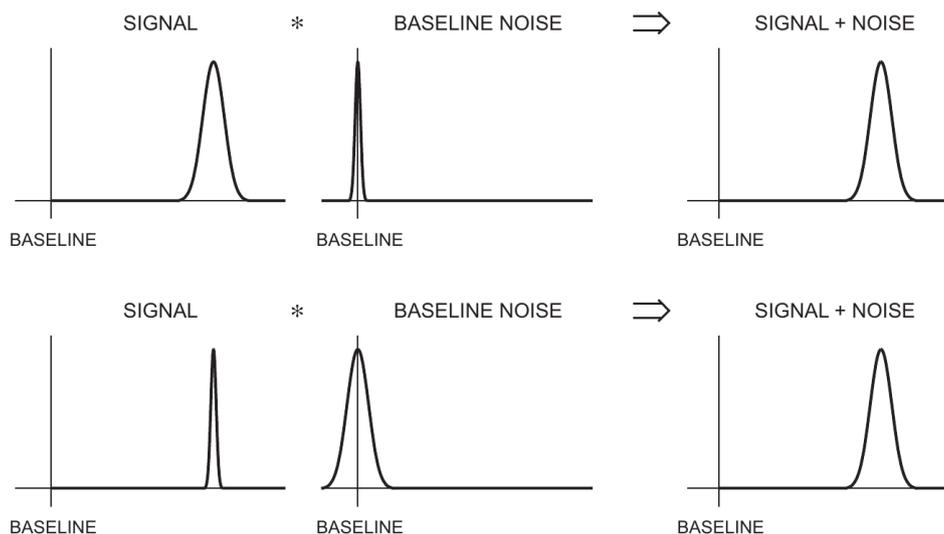


Figura 55: Le fluttuazioni del segnale e del rumore si sommano in quadratura. Se il segnale ha una varianza grande (schema in alto), le fluttuazioni del rumore sono solitamente trascurabili. In tal caso è ammissibile un'elettronica più rumorosa, come avviene, ad esempio, negli scintillatori o nelle camere proporzionali. Se invece il segnale ha una varianza piccola (schema in basso), è cruciale avere un'elettronica silenziosa, per mantenere una buona risoluzione del segnale.

Preliminarmente, definiamo la **media temporale** di una funzione del tempo $x(t)$ come:

$$\bar{x} = \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt \quad (141)$$

Solitamente, tranne nei casi in cui la funzione $x(t)$ sia una sinusoidale, possiamo estendere la media temporale su tutto l'asse dei tempi, ridefinendo:

$$\bar{x} = \lim_{T \rightarrow \infty} \left[\frac{1}{T} \int_{-T/2}^{T/2} x(t) dt \right] \quad (142)$$

Ricordiamo che la potenza in un circuito resistivo può essere scritta come $P = VI = RI^2 = V^2/R$. Possiamo quindi definire la **potenza fisica** del segnale come:

$$P_{phy} = \lim_{T \rightarrow \infty} \left[\frac{1}{T} \int_{-T/2}^{T/2} RI^2(t) dt \right] = \lim_{T \rightarrow \infty} \left[\frac{1}{T} \int_{-T/2}^{T/2} \frac{V^2(t)}{R} dt \right] \quad (143)$$

ovvero $P_{phy} = R\overline{I^2} = \overline{V^2}/R$. Esprimendo il segnale in uscita in termini di tensione, chiamiamo $x(t) = V(t)$. Generalizzando, possiamo ora definire la **potenza del segnale** $x(t)$ come:

$$P = \overline{x^2} = \lim_{T \rightarrow \infty} \left[\frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt \right] \quad (144)$$

Notiamo che non si tratta della potenza fisica, per ottenere la quale dovrei dividere per la resistenza: $P_{phy} = P/R$.

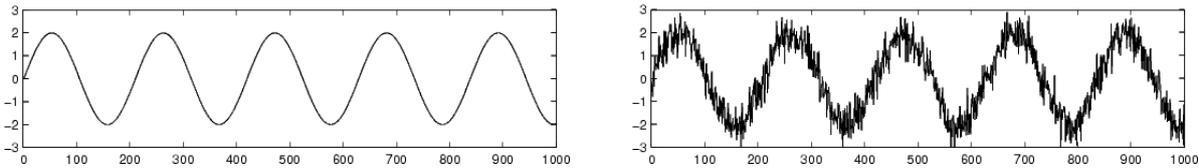


Figura 56: Confronto fra un segnale sinusoidale ideale (a sinistra) ed il medesimo affetto da rumore casuale (a destra). Nei grafici abbiamo l'ampiezza in ordinata ed il tempo in ascissa. Notiamo che il rapporto segnale-rumore nel secondo grafico è abbastanza grande, per cui i due segnali hanno circa la stessa potenza.

Come visibile dai grafici in fig. 56, un segnale con rapporto segnale-rumore sufficientemente elevato avrà una potenza simile a quella del segnale ideale, privo di rumore.

Il rumore elettronico analizzato nel dominio delle frequenze ha comportamenti che variano al variare dell'origine della componente considerata. A loro volta, i componenti attivi dell'elettronica di elaborazione del segnale a valle del sensore hanno una risposta dipendente dalla frequenza. Operano cioè come dei filtri, in grado di esaltare un certo intervallo di frequenze e sopprimerne altre.

Per poter trattare con facilità l'effetto del circuito sul segnale (e sul rumore) devo dunque passare al **dominio delle frequenze**.

A tale scopo, definiamo la **densità spettrale di potenza** (PSD, *power spectral density*) $S_x(f)$ di un segnale x in funzione della frequenza f . La **potenza del segnale** x in una banda di frequenze Δf viene indicata con $P_{x\Delta f}$.

Cerchiamo di trattare il nostro segnale complessivo nello spazio delle frequenze, e per un dato intervallo di frequenze vogliamo sapere quale sia il contributo di potenza, ovvero qual è l'energia del sistema apportato da queste frequenze, e fare poi una scansione di tale quantità su tutte le frequenze.

Posso immaginare di prendere per esempio 1 Hz di intervallo e calcolare la potenza media su quell'intervallo, poi spostare l'intervallo di 1 Hz e ricalcolare. I punti calcolati riportati in grafico in funzione della frequenza centrale del mio intervallo costruiranno l'andamento di tale potenza spettrale in funzione della frequenza. Se cambiamo l'estensione dell'intervallo, il contributo di ogni intervallo cambierà'. Possiamo quindi normalizzare il contributo dividendolo per l'estensione dell'intervallo e far tendere a zero l'estensione dell'intervallo, in modo da rendere la quantità calcolata una densità'. Abbiamo quindi ottenuto la **densità spettrale di potenza**.

Quindi avremo:

$$S_x(f) \equiv \lim_{\Delta f \rightarrow 0} \frac{P_x \Delta f}{\Delta f} \quad (145)$$

La potenza totale del segnale si ottiene integrando la densità spettrale di potenza su tutte le frequenze, cioè:

$$P = \int_{-\infty}^{+\infty} S_x(f) df \quad (146)$$

5.6 Contributi alla potenza di rumore

La trattazione che abbiamo visto in precedenza riguardava la potenza del segnale, che però non coincide con la sua **potenza fisica** (come da confronto fra le relazioni 144 e 143). A seconda del tipo di segnale con cui stiamo operando, in tensione oppure in corrente, possiamo sostituire la densità spettrale di potenza (non fisica) del rumore (da cui il pedice n , per *noise*), con quella **in tensione** oppure **in corrente**, ovvero:

$$\frac{dP}{df} \quad \longmapsto \quad \frac{1}{R} \frac{dv_n^2}{df} \quad \text{oppure} \quad R \frac{di_n^2}{df} \quad (147)$$

Per convenzione, il rumore quindi si descrive in termini di densità spettrali di tensione e di corrente:

$$e_n^2 \equiv \frac{dv_n^2}{df} \quad \text{e} \quad i_n^2 \equiv \frac{di_n^2}{df} \quad (148)$$

La notazione tende a confondere inizialmente, ma i_n sarà usato ora solo nell'accezione di densità spettrale di corrente di rumore. Sarà dunque

conveniente esplicitare la densità spettrale di tensione e_n e corrente i_n di rumore, a seconda che il rumore produca un segnale in tensione o in corrente. Attenzione che esse sono quantità espresse in unità di V/\sqrt{Hz} e A/\sqrt{Hz} , rispettivamente.

Il rumore totale in uscita è ottenuto integrando la potenza di rumore sugli intervalli di frequenza rilevanti, ovvero la larghezza di banda (*bandwidth*) del sistema. Dato che la potenza è proporzionale alla tensione o alla corrente al quadrato, il rumore in uscita di un amplificatore che ha guadagno $A(f)$ dipendente dalla frequenza è:

$$v_{no}^2 = \int_0^{+\infty} e_n^2 A^2(f) df \quad (149)$$

$$i_{no}^2 = \int_0^{+\infty} i_n^2 A^2(f) df \quad (150)$$

Abbiamo visto che esistono tre principali categorie di rumore elettronico: quello **termico** (bianco), quello **shot**, e quello **dipendente dalla frequenza** (rosa). Esistono anche altre sottocategorie di rumore che possono comunque essere ricondotte alle tre succitate. Ci preoccupiamo ora di ricavarne le rispettive densità spettrali in frequenza.

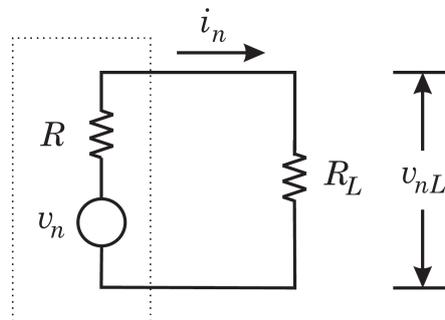


Figura 57: Circuito equivalente per valutare il rumore termico generato da una resistenza reale, rappresentata nel blocco tratteggiato. Essa è idealmente schematizzata in una resistenza R ideale, in serie con il generatore della tensione v_n di rumore. La corrente i_n circolante, viene dissipata sulla resistenza di carico R_L , ai capi della quale la caduta di potenziale è v_{nL} .

Vediamo il **rumore termico** generato da una resistenza in serie nel circuito. Possiamo schematizzarla con il circuito equivalente rappresentato nell'area tratteggiata in fig. 57.

Gli spettri in frequenza, in termini di potenza, tensione e corrente, del rumore termico, sono rispettivamente:

$$\frac{dP}{df} = 4kT \quad , \quad \frac{dv_n^2}{df} = 4kTR \quad , \quad \frac{di_n^2}{df} = \frac{4kT}{R} \quad (151)$$

Notiamo che, secondo questa trattazione, la potenza cresce linearmente con l'ampiezza della banda di frequenza, per cui diverge se consideriamo tutte le frequenze. In realtà ciò non accade, ed il modello di corpo nero utilizzato per descrivere la resistenza dev'essere modificato, introducendo termini correttivi che mandano a zero la densità spettrale di potenza alle alte frequenze, garantendo la convergenza della potenza integrata. Nei circuiti reali, inoltre, lo spettro in frequenza del rumore è comunque limitato, per cui il risultato ricavato rimane ugualmente valido in prima approssimazione.

Vediamo ora il **rumore shot**. Esso è dovuto alla fluttuazione del numero di portatori di carica che compongono il segnale. Consideriamo un condensatore vuoto, alimentato con una tensione costante, tale da riuscire ad estrarre elettroni dal catodo. Questi poi derivano lungo il campo, verso l'anodo, dando luogo ad una corrente elettrica. Ogni ionizzazione avviene in modo casuale e scorrelato dalle altre, seguendo la statistica di Poisson, e produce un segnale impulsivo di una durata caratteristica t . Questo impulso può essere rappresentato come una delta di Dirac, se osservato su tempi $T \gg t$, ovvero, in termini di frequenza ($f = 1/T$), per $f \ll 1/t$. Passando dal dominio dei tempi a quello delle frequenze, osserviamo che la trasformata di Fourier della delta è costante. Ciò significa che la delta possiede tutte le frequenze, distribuite in modo uniforme, ha cioè uno spettro "bianco".

Ricordando che $j_n^2 = di_n^2/df$, ed identificando la corrente media $I = Nq$, abbiamo:

$$\frac{di_n^2}{df} = 2Iq \quad (152)$$

Osserviamo che il rumore *shot* diventa rilevante alle alte frequenze. In questo regime, infatti, il tempo di osservazione T diventa tanto piccolo da essere confrontabile con il tempo caratteristico t della singola ionizzazione, ovvero $T \sim t$. Equivalentemente, stiamo dicendo che la frequenza aumenta, ovvero $f = T^{-1} \sim 1/t$. In questo modo aumenta anche la risoluzione temporale sui singoli impulsi dovuti al passaggio di ciascun elettrone, e l'effetto di "granularità" aumenta.

Vediamo ora il **rumore alle basse frequenze**. Esso è causato dal fenomeno del *trapping* (descritto in ??), in presenza del quale, ricordiamo, la

carica che si muove nel mezzo per deriva (i.e. a velocità costante), si riduce, dopo un tempo t , ad un fattore $e^{-t/\tau}$, ove τ è il parametro di vita media del portatore.

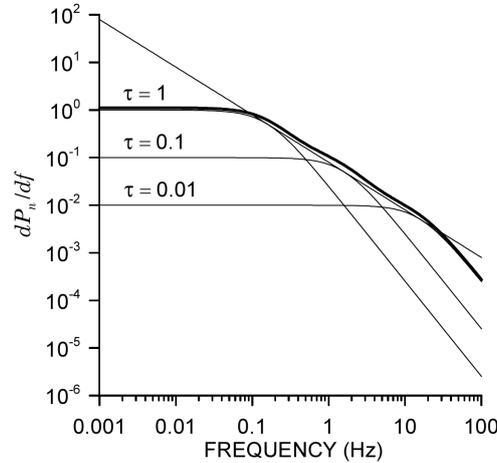


Figura 58: Intrappolamento con tre vite medie distinte, per ciascuna delle quali è riportata (linea sottile) la densità spettrale di rumore, che decade come $1/f^2$. Quella del rumore complessivo (linea grossa), è data dalla somma delle prime tre, e decade, in prima approssimazione, come $1/f$. Viene riportato, per confronto anche quest'ultimo andamento.

Quando la vita media assume un unico valore, la densità spettrale del rumore è costante alle basse frequenze, per poi decadere come $1/f^2$. Nelle situazioni reali, però, la vita media assume molteplici valori. La densità spettrale complessiva si ottiene sommando le singole densità spettrali su tutti i diversi valori di vita media. Come visibile nel grafico in fig. 58, già con soli tre valori diversi di vita media, lo spettro del rumore complessivo assume un andamento che decade verosimilmente come $1/f$. Possiamo scrivere dunque:

$$S(f) = \frac{k}{f} \quad (153)$$

Chiaramente, anche la densità spettrale della tensione del rumore avrà lo stesso andamento:

$$\frac{dv_n^2}{df} = \frac{A}{f} \quad (154)$$

che, integrata sullo spettro delle frequenze ci da:

$$v_n^2 = \int_{f_1}^{f_2} \frac{A}{f} df = A \ln \left(\frac{f_2}{f_1} \right) \quad (155)$$

Notiamo che tale andamento è ideale: sperimentalmente osserviamo uno spettro che decade come $1/f^\alpha$, con α compreso fra 0.5 e 2.

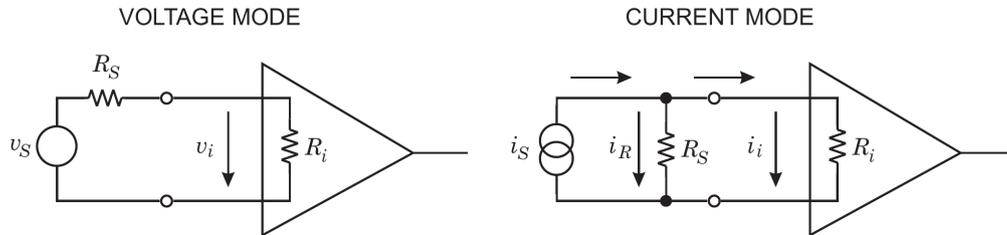


Figura 59: A seconda del circuito in cui è inserito, l'amplificatore può operare in tensione (a sinistra), oppure in corrente (a destra).

5.7 Amplificatori e filtri

Discutiamo ora nel dettaglio gli **amplificatori**. A seconda di come stiamo leggendo il segnale, in tensione oppure in corrente, possiamo scegliere il circuito più adatto per amplificare l'una o l'altra grandezza, come rappresentato negli schemi di fig. 59.

Vediamo come amplificare la **tensione**. Nel circuito a sinistra in fig. 59, il generatore di tensione reale è composto da uno ideale, che produce una tensione v_S , messo in serie con la resistenza interna R_S . Anche l'amplificatore presenta una sua resistenza interna R_i . Per prima cosa vogliamo ricavare la tensione v_i che arriva all'amplificatore, in funzione di v_S . Utilizziamo le leggi di Kirchhoff, considerando che la corrente i è uguale su tutto il circuito a monte dell'amplificatore, che non presenta nodi. Dalla legge di Ohm, la tensione in ingresso all'amplificatore può essere espressa come:

$$v_i = iR_i \quad (156)$$

La tensione in uscita dal generatore (ideale) è pari alla somma delle cadute di potenziale sulle altre resistenze, ovvero:

$$v_S = iR_i + iR_S = i(R_i + R_S) \quad (157)$$

Dividendo la prima equazione per la seconda, si elimina la corrente i , ottenendo la frazione di tensione del generatore che arriva in ingresso all'amplificatore:

$$\frac{v_i}{v_S} = \frac{R_i}{R_i + R_S} \quad (158)$$

Osserviamo che, per avere $v_i = v_S$, bisogna che sia $R_i \gg R_S$, ovvero un amplificatore di tensione ideale ha resistenza interna $R_i \rightarrow \infty$. Nei casi reali abbiamo però più amplificatori a cascata, in cui la resistenza interna di un amplificatore gioca il ruolo di resistenza interna della sorgente nello stadio successivo. Ciò richiede dunque di calibrare le resistenze interne degli amplificatori.

Vediamo ora come amplificare la **corrente**. Un generatore reale di corrente può essere schematizzato con un generatore ideale che produce una corrente i_S messo in parallelo con una resistenza R_S . Ai suoi capi viene collegato l'amplificatore, a formare un circuito come quello a destra in fig. 59. Anche in questo caso, vogliamo ricavare la corrente i_i in ingresso all'amplificatore, in funzione di i_S . Chiamando i_R la corrente che passa nel ramo della resistenza interna del generatore (R_S), la legge dei nodi di Kirchhoff, ci permette di scrivere:

$$i_S = i_R + i_i \quad (159)$$

Mentre per le cadute di tensione sulla maglia più a destra, abbiamo:

$$i_R R_S = i_i R_i \quad \implies \quad i_R = i_i \frac{R_i}{R_S} \quad (160)$$

Sostituendo questo risultato nell'equazione precedente, otteniamo:

$$i_S = i_i \frac{R_i}{R_S} + i_i = i_i \left(\frac{R_i}{R_S} + 1 \right) = i_i \left(\frac{R_i + R_S}{R_S} \right) \quad (161)$$

Quindi la frazione di corrente del generatore che arriva in ingresso all'amplificatore è:

$$\frac{i_i}{i_S} = \frac{R_S}{R_i + R_S} \quad (162)$$

Contrariamente al caso precedente, per ottenere $i_i = i_S$, dovremo avere $R_S \gg R_i$, ovvero un amplificatore di corrente ideale richiede un generatore con resistenza interna $R_S \rightarrow \infty$.

Quindi un amplificatore, per poter essere usato in modalità di tensione richiederà una resistenza interna grande, viceversa per essere usato nella modalità di corrente, richiederà una resistenza piccola (dove grande e piccolo è sempre inteso in senso relativo, rispetto alla resistenza interna della sorgente).

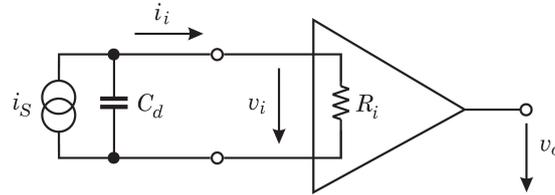


Figura 60: Circuito di amplificazione con una sorgente capacitiva, come un sensore.

I sensori di nostro interesse si comportano come **sorgenti capacitive**. Il circuito in cui vengono inseriti è rappresentato in fig. 60. Un sensore viene idealmente schematizzato come un generatore di corrente i_s di segnale, in parallelo con una capacità C_d . Il rilascio di una carica Q_s nel sensore produce una corrente impulsiva i_s che dura per tutto il tempo di raccolta della carica t_c :

$$Q_s = \int_0^{t_c} i_s(t) dt = i_s t_c \quad (163)$$

Se v_i è la tensione in ingresso all'amplificatore, possiamo scrivere la tensione in uscita come $v_o = A_v v_i$, dove A_v è il **fattore di guadagno in tensione**. Confrontando il tempo di raccolta t_c con il tempo caratteristico $R_i C_d$ del circuito a monte dell'amplificatore, questo opererà in modalità di corrente oppure di tensione. Nello specifico:

- per $R_i C_d \ll t_c$, la capacità del sensore si scarica rapidamente (rispetto al tempo di raccolta). In questo caso, la corrente in ingresso all'amplificatore sarà quasi uguale a quella che esce dal sensore, ossia $i_i \simeq i_s$. In uscita dall'amplificatore otterremo un segnale in tensione che segue, istante per istante, l'andamento della corrente in ingresso: $v_o(t) \propto i_s(t)$. In tale regime, il sistema complessivo sensore-amplificatore, opera in **modalità di corrente**;
- per $R_i C_d \gg t_c$, la capacità del sensore si scarica molto lentamente (rispetto al tempo di raccolta). La carica viene prima immagazzinata nella capacità, che raggiunge una tensione massima $V_i = Q_s / C_d \propto \int_0^{t_c} i_s(t) dt$,

dopo il tempo di raccolta t_c . Successivamente, il condensatore si scarica sulla resistenza: la tensione ai capi di quest'ultima (i.e. in ingresso all'amplificatore) è $v_i(t) = V_i \exp(-t/R_i C_d)$. In uscita dall'amplificatore otteniamo una tensione $v_o(t) \propto v_i(t)$. In questo regime, si perde informazione sulla corrente in uscita dal sensore, ed il sistema complessivo sensore-amplificatore opera in **modalità di tensione**.

In entrambi i casi, l'amplificatore apporta un fattore di **guadagno sempre in tensione**. La differenza fra i due regimi sta nel fatto che in ingresso all'amplificatore può arrivare un impulso in tensione che, o segue istantaneamente l'andamento della corrente (modalità di corrente), o segue il decadimento della carica totale, ottenuta per integrazione della corrente nel tempo di raccolta (modalità di tensione).

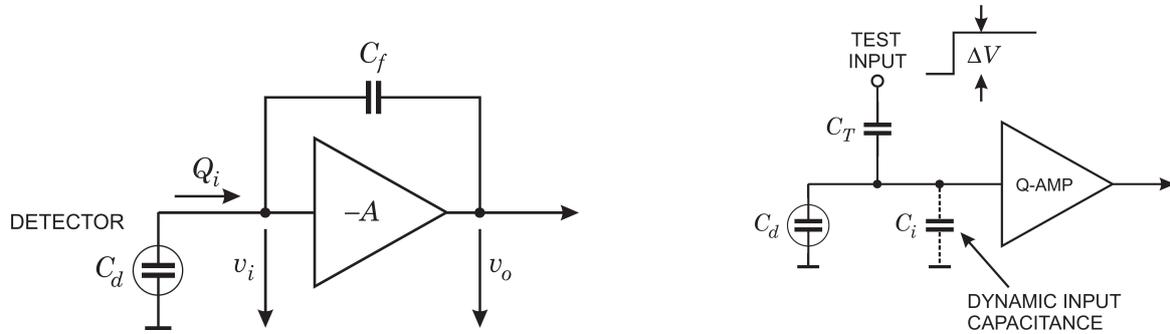


Figura 61: A sinistra: circuito di amplificazione con sorgente capacitiva (sensore) e capacità di *feedback*. A destra: circuito per effettuare la calibrazione dell'amplificatore. La capacità di test viene utilizzata per introdurre una quantità di carica nota nel circuito.

Inserendo una capacità C_f di *feedback* in parallelo con l'amplificatore, come nel circuito a sinistra in fig. 61, possiamo ottenere un **amplificatore di carica** (o *charge-sensitive amplifier*). L'amplificatore in questione ha resistenza interna R_i molto alta, idealmente infinita, ed è utilizzato come invertitore, ovvero con guadagno negativo:

$$\frac{dv_o}{dv_i} = -A \quad \Longrightarrow \quad v_o = -Av_i \quad (164)$$

La tensione ai capi della capacità di *feedback* sarà $v_f = v_i(A + 1)$, con una carica immagazzinata di $Q_f = C_f v_f = C_f v_i(A + 1)$. Dal momento che la resistenza R_i è grande, non vi è alcuna corrente che passa attraverso l'amplificatore, e tutta la carica iniziale Q_i viene immagazzinata nel condensatore, ovvero $Q_f = Q_i$. Da tali considerazioni, possiamo quindi scrivere:

$$\frac{Q_i}{v_i} = C_f(A + 1) \equiv C_i \quad (165)$$

dove C_i viene chiamata **capacità dinamica in ingresso**. Ci interessa calcolare il **fattore di amplificazione di carica**, ovvero quel valore della tensione in uscita che otteniamo per una carica unitaria in ingresso:

$$A_Q \equiv \frac{v_o}{Q_i} = \frac{Av_i}{C_i v_i} = \frac{A}{C_i} = \frac{A}{A + 1} \frac{1}{C_f} \quad (166)$$

Questo parametro ci da una misura di quanto ampio è il segnale in tensione che esce dall'amplificatore, a fronte di una determinata carica in ingresso. Osserviamo che esso dipende unicamente dal guadagno dell'amplificatore e dalla capacità di *feedback*. Inoltre, nel limite per $A \gg 1$, abbiamo che $A_Q \sim 1/C_f$, ovvero la dipendenza si limita alla sola capacità di *feedback*.

Valutiamo ora il rapporto fra la carica che viene effettivamente misurata (Q_i , che arriva in ingresso all'amplificatore) e quella di segnale (Q_s , rilasciata nel sensore):

$$\frac{Q_i}{Q_s} = \frac{Q_i}{Q_i + Q_d} = \frac{C_i}{C_i + C_d} = \frac{1}{1 + \frac{C_d}{C_i}} \quad (167)$$

Per avere $Q_i \simeq Q_s$, bisogna scegliere $C_i \gg C_d$, ovvero una capacità dinamica in ingresso all'amplificatore che sovrasti quella del sensore: ciò fa sì che la carica non rimanga nel sensore, ma migri verso l'amplificatore.

Per calibrare l'amplificatore (i.e. per conoscere il suo fattore di guadagno) si ricorre ad un circuito come quello riportato a destra in fig. 61, che implementa una **capacità di test** C_T . Si tratta di una componente normalmente già integrata nei sensori più recenti. Applicando una tensione costante ΔV ai capi di C_T , stiamo introducendo nel circuito una quantità nota di carica, pari a $C_T \Delta V$. Ad essere più precisi, questa carica non arriva completamente all'amplificatore: parte fluisce anche in quella del sensore (C_d). Calcoliamo dunque la capacità equivalente C_{eq} del circuito. Dallo schema, vediamo che C_T è in serie sia con C_d , che con C_i , le quali sono collegate in parallelo fra loro, quindi:

$$C_{eq} = \left[\frac{1}{C_T} + \frac{1}{C_i + C_d} \right]^{-1} = \left[\frac{1}{C_T} \left(1 + \frac{C_T}{C_i + C_d} \right) \right]^{-1} = \frac{C_T}{1 + \frac{C_T}{C_i + C_d}} \quad (168)$$

Per massimizzare la frazione di carica introdotta che arriva all'amplificatore, bisogna scegliere una capacità di test con $C_i \gg C_T$ (analogamente a quanto visto per la capacità del sensore, per la quale $C_i \gg C_d$). In tal modo, il termine $C_T/(C_i + C_d)$ è piccolo, e ricordando che $1/(1+x) \sim (1-x)$ per $x \rightarrow 0$, possiamo approssimare il risultato precedente come:

$$C_{eq} \simeq C_T \left(1 - \frac{C_T}{C_i + C_d} \right) \quad (169)$$

Quindi la carica di test Q_T che arriva all'amplificatore sarà:

$$Q_T = C_{eq} \Delta V = \frac{C_T}{1 + \frac{C_T}{C_i + C_d}} \Delta V \simeq C_T \left(1 - \frac{C_T}{C_i + C_d} \right) \Delta V \quad (170)$$

La formula suggerisce che è necessario collegare il sensore (C_d) all'elettronica per tarare tale carica di test.

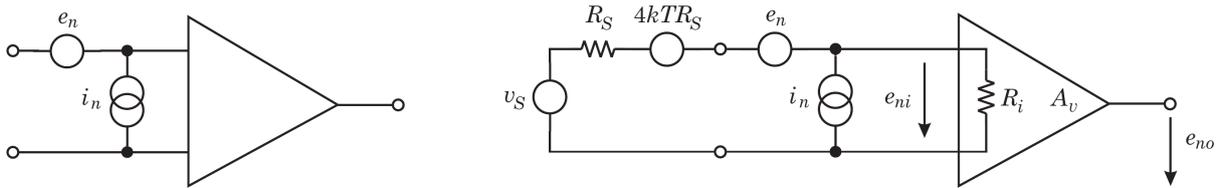


Figura 62: A sinistra: il rumore in ingresso all'amplificatore è determinato dai contributi in tensione ed in corrente, schematizzati dai due generatori, rispettivamente e_n ed i_n . A destra: allo stesso circuito viene aggiunta una sorgente di segnale v_S , con carico resistivo R_S , che, a sua volta, è fonte di rumore termico.

5.7.1 Effetto dell'amplificatore sul rapporto segnale/rumore

Vogliamo ora analizzare le proprietà del **rumore in ingresso all'amplificatore**. Esso può essere completamente descritto in termini di sorgente di tensione di rumore e di sorgente di corrente di rumore all'input dell'amplificatore. Invece che specificare il noise totale su tutta la banda, l'ampiezza di ogni sorgente di rumore è caratterizzata dalla sua densità spettrale. Questo è conveniente perché gli effetti delle impedenze dipendenti dalla frequenza nel circuito di ingresso e della larghezza di banda dell'amplificatore possono essere valutati separatamente.

Per brevità, parleremo semplicemente di *rumore*, riferendoci alla sua densità spettrale. Useremo inoltre le seguenti notazioni per la densità spettrale in tensione e quella in corrente del rumore:

$$e_n^2 \equiv \frac{dv_n^2}{df} \quad \text{e} \quad i_n^2 \equiv \frac{di_n^2}{df} \quad (171)$$

Come visibile nel circuito a sinistra in fig. 62, abbiamo una sorgente di rumore in tensione e_n ed una in corrente i_n (quindi indicate attraverso le radici delle rispettive densità di potenza). Come unità di misura, nelle applicazioni di nostro interesse, avremo tipicamente:

$$[e_n] = \frac{\text{nV}}{\sqrt{\text{Hz}}} \quad \text{e} \quad [i_n] = \frac{\text{fA}}{\sqrt{\text{Hz}}} \quad \text{oppure} \quad \frac{\text{pA}}{\sqrt{\text{Hz}}} \quad (172)$$

Nel circuito a destra in fig. 62 abbiamo anche inserito una sorgente reale di segnale in tensione, scomposta in un generatore di tensione ideale, con la resistenza R_S inserita in serie, che è fonte di rumore termico, con spettro di potenza in tensione pari a $4kTR_S$ (come calcolato nelle relazioni 151). Teniamo conto anche della resistenza interna dell'amplificatore R_i , il cui valore sarà molto grande, sebbene non infinito, come nel caso ideale. L'amplificazione viene effettuata in tensione, con un fattore di guadagno A_v . Con e_{no} chiamiamo il rumore in tensione all'uscita dell'amplificatore, che è proporzionale a tutto il rumore che esso riceve in ingresso. Partiamo dal circuito ancora privo della sorgente di segnale, ovvero come quello a sinistra in fig. 62. Se ipotizziamo di avere **resistenza interna infinita**, $R_i = \infty$, non avremo alcun rumore dovuto alla corrente, poiché stiamo inibendo il suo passaggio nell'amplificatore. Possiamo scrivere il rumore in uscita come:

$$e_{no} = e_n A_v \quad (173)$$

Inseriamo ora la sorgente, per ottenere la configurazione completa, come a destra in fig. 62. Vediamo che la sorgente di rumore in corrente, i_n , è collegata alla resistenza R_S interna alla sorgente, grazie alla quale produce un rumore in tensione, quantificabile come $i_n R_S$. Le diverse sorgenti di rumore non sono correlate fra loro, dunque il rumore totale è dato dalla loro somma in quadratura, ovvero:

$$e_{ni}^2 = 4kTR_S + e_n^2 + (i_n R_S)^2 \quad (174)$$

Applicando la relazione 173, il rumore in uscita dall'amplificatore sarà:

$$e_{no}^2 = A_v^2 e_{ni}^2 = A_v^2 [4kTR_S + e_n^2 + (i_n R_S)^2] \quad (175)$$

Introducendo ora il segnale v_S prodotto dalla sorgente, avremo in uscita dall'amplificatore un segnale in tensione $S = A_v v_S$, affetto da un rumore $N = e_{no}$. Quindi il rapporto segnale-rumore sarà:

$$\left(\frac{S}{N}\right)^2 = \frac{A_v^2 v_S^2}{A_v^2 [4kTR_S + e_n^2 + (i_n R_S)^2]} = \frac{v_S^2}{4kTR_S + e_n^2 + i_n^2 R_S^2} \quad (176)$$

In questo caso, sia il segnale che il rumore vengono amplificati per lo stesso fattore di guadagno, ciò fa sì che il rapporto segnale-rumore sia indipendente da tale fattore.

Nel caso reale, avremo una **resistenza interna finita**, $R_i < \infty$: vogliamo vedere se e come cambia il rapporto segnale-rumore. Per prima cosa calcoliamo il segnale in tensione v_{Si} in ingresso all'amplificatore (i.e. ai capi della sua resistenza interna), distinguendolo da v_S , ovvero il potenziale prodotto dal generatore ideale. Chiamando i_S la corrente che fluisce sulla maglia più esterna (i.e. togliendo il ramo con la sorgente i_n di rumore in corrente), avremo che $v_{Si} = R_i i_S$, ovvero $i_S = v_{Si}/R_i$. Bilanciamo dunque le cadute di potenziale sulla maglia, per ottenere v_{Si} :

$$0 = v_S - R_S i_S - v_{Si} = v_S - R_S \frac{v_{Si}}{R_i} - v_{Si} = v_S - v_{Si} \left(1 + \frac{R_S}{R_i}\right) = v_S - v_{Si} \frac{R_i + R_S}{R_i}$$

$$v_{Si} = v_S \frac{R_i}{R_i + R_S} \quad (177)$$

Rispetto al caso precedente, stiamo introducendo il fattore correttivo $R_i/(R_i + R_S)$ che rimodula la tensione in ingresso all'amplificatore. Naturalmente, nel limite per $R_i \rightarrow \infty$, riotteniamo il caso di prima, con $v_{Si} = v_S$. Anche le due fonti di rumore in tensione subiscono tale trasformazione.

La sorgente di rumore in corrente, i_n , produce ora un passaggio di corrente sia nella resistenza interna della sorgente, R_S , ma anche in quella interna dell'amplificatore R_i . In tale configurazione, la resistenza equivalente, come vista dalla sorgente di rumore in corrente, è data dal parallelo delle due. Il rumore in tensione dovuto a tale sorgente si può scrivere come:

$$i_n \left(\frac{1}{R_i} + \frac{1}{R_s} \right)^{-1} = i_n \frac{R_i R_s}{R_i + R_s} \quad (178)$$

Per quanto detto, il rumore totale in ingresso all'amplificatore viene ad essere:

$$e_{ni}^2 = (4kTR_s + e_n^2) \left(\frac{R_i}{R_i + R_s} \right)^2 + i_n^2 \left(\frac{R_i R_s}{R_i + R_s} \right)^2 \quad (179)$$

Possiamo quindi ricalcolare il rapporto segnale-rumore in uscita dall'amplificatore:

$$\left(\frac{S}{N} \right)^2 = \frac{A_v^2 v_{Si}^2}{A_v^2 e_{ni}^2} = \frac{v_S^2 \left(\frac{R_i}{R_i + R_s} \right)^2}{(4kTR_s + e_n^2) \left(\frac{R_i}{R_i + R_s} \right)^2 + i_n^2 \left(\frac{R_i R_s}{R_i + R_s} \right)^2} = \frac{v_S^2}{(4kTR_s + e_n^2) + i_n^2 R_s^2} \quad (180)$$

Osserviamo subito che questo valore è identico a quello che avevamo calcolato per $R_i = \infty$ (in formula 176). Quindi il rapporto segnale-rumore in uscita dall'amplificatore non dipende dalla sua resistenza interna, e la trattazione del rumore fatta nel caso ideale, resta valida anche in quello reale. Questo ragionamento si può estendere al caso più generico, che tiene conto dell'impedenza interna dell'amplificatore.

Nel caso in cui le sorgenti di rumore e_n ed i_n siano fra loro **correlate**, il rumore totale in tensione che entra nell'amplificatore (formula 174) viene riscritto come:

$$e_{ni}^2 = 4kTR_s + e_n^2 + (i_n R_s)^2 + 2C e_n i_n R_s \quad (181)$$

dove C è il coefficiente di correlazione tra le due fonti di rumore.

Ricordiamo che $e_n^2 = e_n^2(f)$ è una densità spettrale in frequenza della potenza (non fisica) del rumore in tensione, come definita in formula 171. Pertanto se vogliamo calcolare la potenza totale del rumore in uscita dall'amplificatore, dobbiamo integrare $e_{no}^2(f)$ sulle frequenze:

$$v_{no}^2 = \int_0^{+\infty} e_{no}^2(f) df = \int_0^{+\infty} e_{ni}^2(f) |A_v|^2 df \quad (182)$$

Il fattore di guadagno A_v dell'amplificatore ha un modulo ed una fase, in generale, dipendenti dalla frequenza, si tratta dunque di una quan-

tità complessa, ed è dunque corretto considerarla con il suo modulo quadro $|A_v|^2 = A_v^* A_v$.

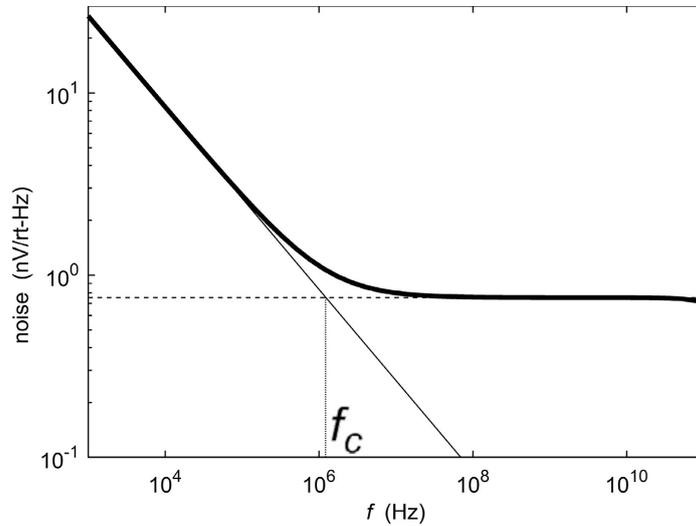


Figura 63: Spettro in frequenza del rumore in uscita dall'amplificatore, in scala doppio-logaritmica. Sono riportate le due componenti, rosa e bianca, ed il rumore complessivo, nonché la frequenza f_c di *corner*.

Gli amplificatori comunemente hanno uno spettro di rumore che decade come $1/f$ (i.e. rosa) alle frequenze basse, mentre è costante (i.e. bianco) alle alte frequenze. Ricordiamo che:

- i rumori bianchi hanno densità spettrale di tensione che è costante in frequenza (come nel caso del rumore termico, in formula 151). Quindi la potenza totale, in tensione, è $v_n^2 \propto f$, ovvero $v_n \propto \sqrt{f}$;
- i rumori rosa hanno densità spettrale di tensione che decade come $1/f$. Hanno una potenza totale, in tensione, che è quella calcolata in formula 155, ovvero $v_n^2 = A_f \ln(f_{max}/f_{min})$, che dunque dipende dal rapporto tra la frequenza massima f_{max} e quella minima f_{min} del loro spettro.

Lo spettro complessivo del rumore dovuto all'amplificatore è riportato in fig. 63. Il valore di frequenza f_c in cui le due componenti spettrali si equivalgono, è detto **frequenza di corner**. Per frequenze $f > f_c$ domina il rumore bianco, mentre per $f < f_c$ domina quello rosa.

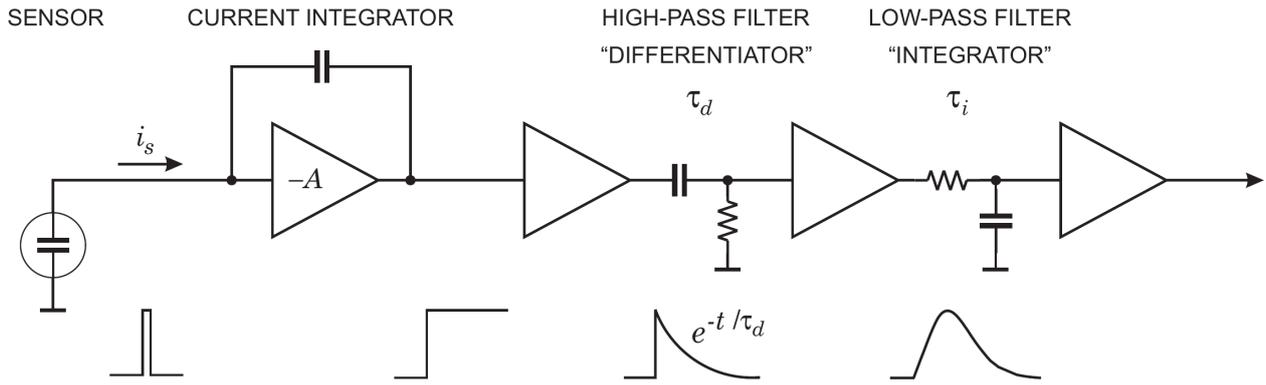


Figura 64: Circuito completo con sensore, preamplificatore e formatore. Vengono riportate le forme dei segnali in uscita da ciascuna componente.

5.7.2 Il formatore

Vediamo ora il **formatore** (o *shaper*). Nello schema di fig. 64, vediamo che dal sensore esce un segnale di durata molto breve. Esso arriva al preamplificatore, che produce un segnale in uscita a gradino, che decade in un tempi lunghi, a seconda della capacità di *feedback* utilizzata. Per tale ragione, il preamplificatore agisce da **integratore** della corrente in ingresso (idealmente, la funzione a gradino è l'integrale della delta di Dirac).

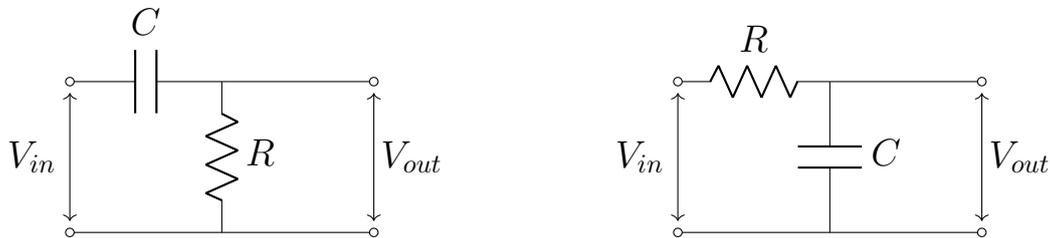


Figura 65: Filtri passivi: passa-alto (sinistra) e passa-basso (destra).

Vediamo preliminarmente i **filtri passa-alto** e **passa-basso**. Essi sono rappresentati nei circuiti in fig. 65. Dato un segnale in ingresso V_{in} , vogliamo vedere come cambia la quantità $A_v(\omega) = V_{out}/V_{in}$, detta funzione di trasferimento, al variare della frequenza ω del segnale in ingresso.

In entrambi i circuiti, l'impedenza totale Z ai capi degli elettrodi del segnale in ingresso V_{in} , è data dal collegamento in serie della resistenza e del condensatore:

$$Z = R + \frac{1}{i\omega C} = R - \frac{i}{\omega C} = \frac{\omega RC - i}{\omega C} \quad (183)$$

Se preleviamo la tensione d'uscita V_{out} ai capi dell'elemento resistivo, come a sinistra fig. 65, otteniamo un filtro passa-alto. Avendo una tensione d'uscita pari a $V_{out} = RI$, la funzione di trasferimento viene ad essere:

$$A_v(\omega) = \frac{V_{out}}{V_{in}} = \frac{R}{Z} = \frac{\omega RC}{\omega RC - i} = \frac{\omega\tau}{\omega\tau - i} \frac{\omega\tau + i}{\omega\tau + i} = \frac{(\omega\tau)^2}{(\omega\tau)^2 + 1} + i \frac{\omega\tau}{(\omega\tau)^2 + 1} \quad (184)$$

dove $\tau = RC$ è il tempo caratteristico del circuito. Il suo modulo è:

$$|A_v(\omega)| = \sqrt{A_v^*(\omega)A_v(\omega)} = \sqrt{\frac{(\omega\tau)^4 + (\omega\tau)^2}{[(\omega\tau)^2 + 1]^2}} = \sqrt{\frac{(\omega\tau)^2[(\omega\tau)^2 + 1]}{[(\omega\tau)^2 + 1]^2}} = \frac{\omega\tau}{\sqrt{(\omega\tau)^2 + 1}} \quad (185)$$

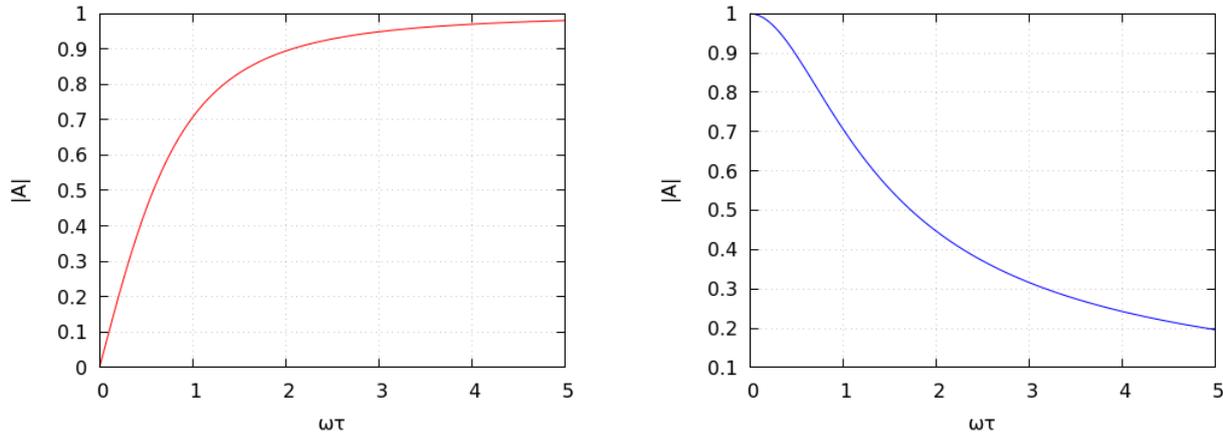


Figura 66: Modulo della funzione di trasferimento per un filtro passa-alto (sinistra) e passa-basso (destra).

Nel grafico a sinistra in fig. 66 vediamo l'andamento della funzione di trasferimento al variare di $\omega\tau$. Per $\omega\tau \ll 1$, ovvero per frequenze $\omega \ll 1/\tau$, la funzione di trasferimento si annulla, viceversa, per $\omega\tau \gg 1$, ovvero per frequenze $\omega \gg 1/\tau$, essa tende ad 1. Ciò significa che i segnali alle basse frequenze verranno *tagliati*: abbiamo realizzato un filtro passa-alto.

Nel caso di un filtro passa-basso, il circuito è lo stesso, ma preleviamo la tensione d'uscita V_{out} ai capi del condensatore. L'impedenza complessiva Z è la stessa di prima (formula 183), mentre la tensione d'uscita è $V_{out} = I/(i\omega C)$, dunque la funzione di trasferimento è:

$$A_v(\omega) = \frac{V_{\text{out}}}{V_{\text{in}}} = \frac{\frac{1}{i\omega C}}{Z} = \frac{1}{i\omega C} \frac{\omega C}{\omega RC - i} = \frac{1}{1 + i\omega\tau} \frac{1 - i\omega\tau}{1 - i\omega\tau} = \frac{1}{(\omega\tau)^2 + 1} - i \frac{\omega\tau}{(\omega\tau)^2 + 1} \quad (186)$$

avente modulo pari a:

$$|A_v(\omega)| = \sqrt{A_v^*(\omega)A_v(\omega)} = \sqrt{\frac{1 + (\omega\tau)^2}{[(\omega\tau)^2 + 1]^2}} = \frac{1}{\sqrt{(\omega\tau)^2 + 1}} \quad (187)$$

Nel grafico a destra in fig. 66 vediamo l'andamento della funzione di trasferimento al variare di $\omega\tau$. Per $\omega\tau \ll 1$, ovvero per frequenze $\omega \ll 1/\tau$, la funzione di trasferimento tende ad 1, viceversa, per $\omega\tau \gg 1$, ovvero per frequenze $\omega \gg 1/\tau$, essa tende a 0. Ciò significa che i segnali alle alte frequenze verranno *tagliati*: abbiamo realizzato un filtro passa-basso.

Ritornando al circuito in fig. 64, all'uscita del preamplificatore abbiamo un segnale a gradino che decade in tempi molto lunghi. Inseriamo dunque, nell'ordine:

- un **filtro passa-alto**, che fa decadere il segnale in tempi più rapidi, come e^{-t/τ_d} . Un taglio alle basse frequenze corrisponde infatti ad un taglio sui tempi più lunghi. Tale filtro è detto anche **differenziatore**, poiché, per un segnale in ingresso di breve durata (i.e. con spettro di frequenza molto esteso), ne restituisce la derivata;
- un **filtro passa-basso**, che produce un fronte di salita meno ripido, con andamento $(1 - e^{-t/\tau_i})$. Un taglio alle alte frequenze corrisponde infatti ad un taglio sui tempi più brevi. Tale filtro è detto anche **integratore**, poiché, per un segnale in ingresso di breve durata (i.e. con spettro di frequenza molto esteso), ne restituisce l'integrale.

Siccome il filtro passa-alto è costituito da un circuito CR , mentre quello passa-basso è costituito da un circuito RC , questo tipo di formatore è detto $CR - RC$. Ne esistono di più sofisticati, aventi però le medesime caratteristiche di base: un limite per le basse frequenze ed uno per le alte.

La risposta in frequenza del formatore è data dal prodotto $G(f) = G_{\text{int}}(f) \cdot G_{\text{dif}}(f)$, ovvero fra quella dell'integratore e del differenziatore. L'antitrasformata di Fourier di tale prodotto corrisponde al prodotto di convoluzione, per cui il segnale in uscita si scrive come:

$$V_o(t) = V_i(t) * g_{int}(t) * g_{dif}(t) \quad (188)$$

Nel caso particolare di un segnale in ingresso a gradino unitario:

$$V_i(t) = \begin{cases} 1 & \text{per } t > 0 \\ 0 & \text{per } t \leq 0 \end{cases} \quad (189)$$

il segnale in uscita sar :

$$V_o(t) = \frac{\tau_d}{\tau_d - \tau_i} \left[e^{-t/\tau_d} - e^{-t/\tau_i} \right] \quad (190)$$

Nel caso specifico in cui $\tau_d = \tau_i = \tau$, che usiamo per facilitare la trattazione, abbiamo:

$$V_o(t) = \frac{t}{\tau} e^{-t/\tau} \quad (191)$$

il cui massimo si ha per $t = \tau$, che prende anche il nome di tempo di picco. Il tempo di formazione, o *shaping time*, e' legato a tale tempo di picco.

Sperimentalmente, i formatori reali introducono un rumore che   solo del $\sim 36\%$ maggiore rispetto a quello del modello ideale, qui descritto: data la sua semplicit , pu  essere dunque utilizzato per produrre stime realistiche.

5.7.3 Effetto del formatore sul rumore in ingresso

Per valutare l'azione del formatore sul rumore in ingresso dovremo seguire i seguenti passaggi nell'analisi:

1. Determinare la tensione di rumore in funzione della frequenza che si presenta all'ingresso dell'amplificatore da tutte le singole sorgenti di rumore.
2. Integrare sulla risposta in frequenza di un formatore CR-RC per determinare il rumore totale in uscita.
3. Determinare la carica di rumore equivalente.

La carica di rumore equivalente tiene conto i contributi delle sorgenti di rumore con una dipendenza di questo tipo:

$$Q_n^2 = i_n^2 T F_i + C_i^2 v_n^2 \frac{1}{T} F_v + C_i^2 A_f F_{vf} \quad (192)$$

con i tre termini che descrivono i contributi rispettivamente della corrente di rumore bianco e delle basse frequenze.

I fattori di forma F_i , F_v e F_{vf} dipendono dal rapporto τ_d/τ_i tra costanti di tempo di integrazione e differenziazione, piu' che dai loro valori assoluti, e dal tempo di picco T , e sono caratteristici della tipologia di formatore usato.

I contributi al rumore dipendono quindi dal tempo caratteristico di picco T .

- Il rumore di corrente è indipendente dalla capacità del rivelatore, in linea con il concetto di "conteggio degli elettroni".
- Il rumore di tensione aumenta con la capacità del rivelatore (tensione del segnale ridotta).
- Il rumore $1/f$ è indipendente dal tempo di formazione. In generale, il rumore totale di un rumore $1/f$ dipende dal rapporto tra le frequenze di taglio superiore e inferiore, non dalla larghezza di banda assoluta del rumore. Se τ_d e τ_i sono scalati dallo stesso fattore, questo rapporto rimane costante.

La carica di rumore equivalente Q_n assume un minimo quando i contributi di rumore di corrente e di tensione sono uguali.

5.7.4 Discussione delle caratteristiche dei formatori

Vediamo alcuni grafici esplicativi delle caratteristiche del formatore. In quello a sinistra in fig. 68 stiamo aumentando la costante di tempo d'integrazione τ_{int} del filtro passa-basso, tenendo fissata quella di differenziazione τ_{diff} . Sapendo che il limite superiore della banda passante di frequenze è dato da $1/\tau_{int}$, esso si riduce all'aumentare di τ_{int} . Per cui stiamo tagliando una parte via via maggiore dello spettro alle alte frequenze, con conseguente riduzione del rumore. Vediamo che il segnale decade più lentamente e diminuisce anche l'altezza del picco. Ciò può avere effetti a seconda di come opera l'ADC a valle del formatore: esso infatti può integrare tutto il segnale in un intervallo temporale prefissato, oppure può misurarne solo l'ampiezza del picco.

Nel grafico a destra in fig. 68 viene viceversa tenuta costante τ_{int} , variando τ_{diff} . Trattandosi di un filtro passa-alto, esso determina un limite inferiore per la banda passante di frequenze, che è dato da $1/\tau_{diff}$. Partendo da $\tau_{diff} = \infty$, tale limite inferiore è posto a 0: le basse frequenze non vengono tagliate,

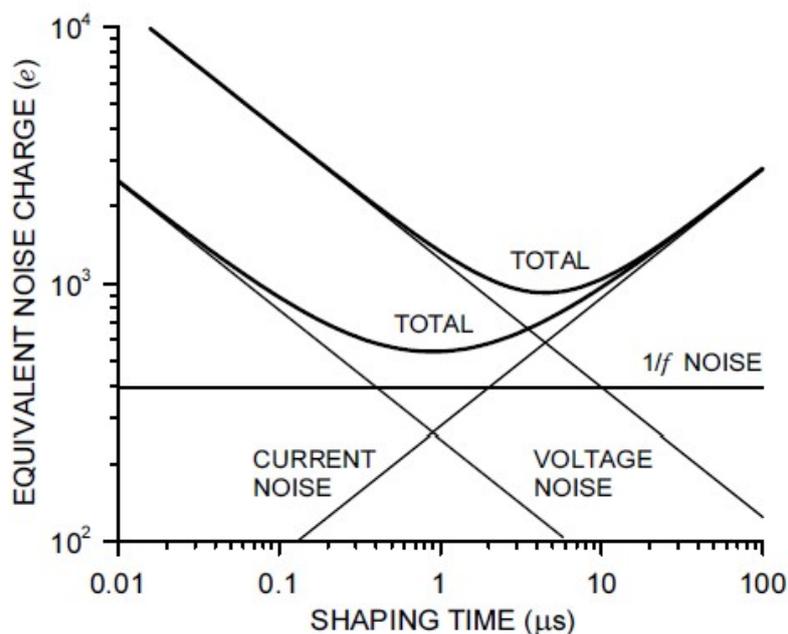


Figura 67: Carica di rumore equivalente rispetto al tempo di formazione. Per tempi di formazione brevi (larghezza di banda elevata), l'ENC è dominato dal rumore di tensione, mentre per tempi di formazione lunghi (tempi di integrazione lunghi) dominano i contributi di rumore di corrente. Il rumore totale assume un minimo quando i contributi di corrente e di tensione sono uguali. Il contributo di rumore "1/f" è indipendente dal tempo di formazione e appiattisce il minimo di rumore. L'aumento del contributo di rumore di tensione o di corrente sposta il minimo di rumore. L'aumento del rumore di tensione è mostrato come esempio.

e dunque il segnale in uscita preserva, sui tempi lunghi, le caratteristiche di quello in ingresso. In particolare osserviamo che viene raggiunto il picco massimo, e il tempo di decadimento risulta molto lungo. Diminuendo τ_{diff} , il limite inferiore della banda di frequenze aumenta: tagliando una parte via via maggiore dello spettro alle basse frequenze, diminuisce il tempo di decadimento del segnale, ma anche il valore del picco massimo.

Consideriamo due diversi valori del rapporto segnale-rumore $(SNR)_1$ ed $(SNR)_2$, a cui corrispondono due valori di carica equivalente di rumore $(ENC)_1$ ed $(ENC)_2$, rispettivamente. Stando alla definizione in formula 139 della carica equivalente di rumore, a parità di carica di segnale, si conserva il prodotto $(SNR) \cdot (ENC)$. Quindi, nel nostro caso, avremo:

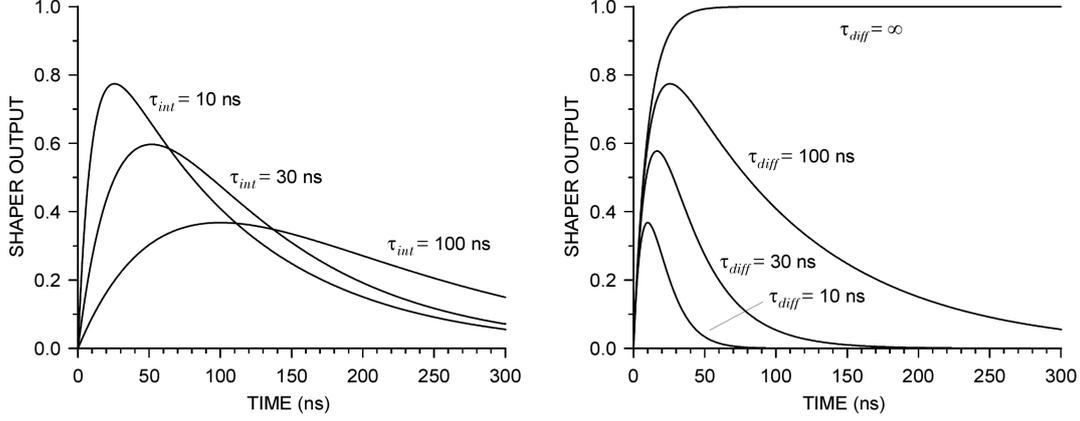


Figura 68: Ampiezza del segnale in uscita dal formatore in funzione del tempo: a sinistra, per diversi valori di τ_{int} (costante di tempo d'integrazione), con fissata $\tau_{diff} = 100$ ns (costante di tempo di differenziazione); a destra, viceversa, per diversi valori di τ_{diff} , con fissata $\tau_{int} = 10$ ns.

$$(\text{SNR})_1 \cdot (\text{ENC})_1 = (\text{SNR})_2 \cdot (\text{ENC})_2 \quad \Rightarrow \quad (\text{ENC})_2 = \frac{(\text{SNR})_1}{(\text{SNR})_2} \cdot (\text{ENC})_1 \quad (193)$$

A sinistra in fig. 69, a parità di costante di tempo di differenziazione, stiamo aumentando quella di integrazione (i.e. riduciamo il limite superiore dello spettro delle frequenze, tagliandone una parte via via maggiore, con conseguente riduzione del rumore). Passando da $\tau_{int} = 10$ ns a $\tau_{int} = 100$ ns, vediamo che il picco del segnale (grafico centrale) si dimezza, mentre il rumore (grafico in alto) diventa circa 1/4, per cui il rapporto segnale-rumore raddoppia. Quindi l'ENC (grafico in basso), secondo la relazione in formula 193, si dimezza.

A destra in fig. 69, a parità di costante di tempo di integrazione, stiamo aumentando quella di differenziazione (i.e. diminuiamo il limite inferiore della banda passante di frequenze, che, allargandosi, lascia entrare più rumore). Passando da $\tau_{diff} = 10$ ns a $\tau_{diff} = 100$ ns vediamo che il picco del segnale (grafico centrale) raddoppia, mentre il rumore (grafico in alto) aumenta del $\sim 0.3 = 30\%$. Il rapporto segnale-rumore scala di un fattore $2/(1 + 0.3) \simeq 1.54$, cui corrisponde un aumento del $(1.54 - 1) = 0.54 = 54\%$. L'ENC invece scala di un fattore $(1 + 0.3)/2 = 0.65$, cui corrisponde una diminuzione del $(1 - 0.65) = 0.35 = 35\%$.

Per ridurre il tempo di discesa del segnale dopo il picco massimo, si può ricorrere all'aggiunta di più integratori a valle del differenziatore. Tale confi-

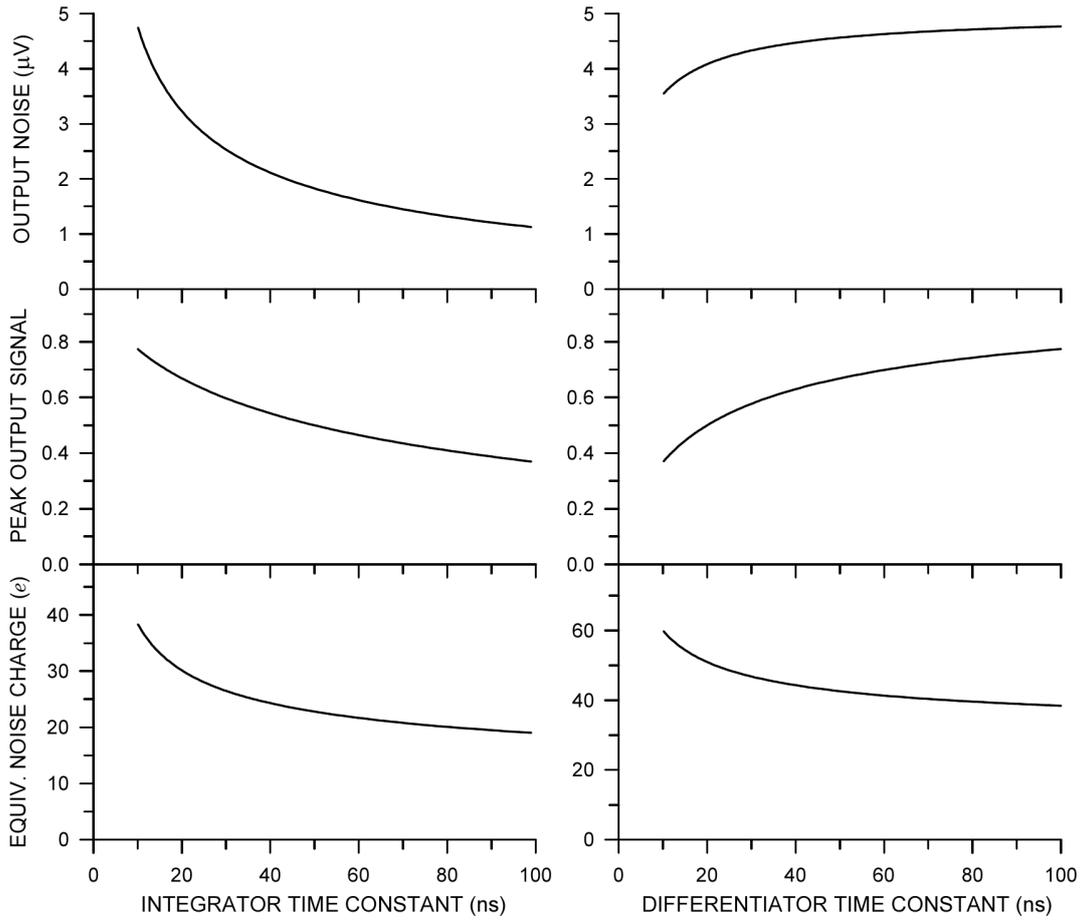


Figura 69: Rumore (in alto), picco del segnale (al centro) e rumore equivalente in elettroni (in basso): a sinistra, in funzione della costante di tempo d'integrazione τ_{int} , tenendo fissa quella di differenziazione ($\tau_{diff} = 100$ ns); a destra, viceversa, in funzione della costante di tempo di differenziazione τ_{diff} , tenendo fissa quella d'integrazione ($\tau_{int} = 10$ ns).

gurazione viene chiamata $CR - nRC$, dove n indica il numero di integratori a valle del differenziatore. Consideriamo che differenziatore ed integratori abbiano tutti la stessa costante di tempo $\tau_{diff} = \tau_{int} = \tau$. Mandando in ingresso un segnale a gradino, quello in uscita si ottiene generalizzando la relazione 191, come:

$$V_o(t) = \left(\frac{t}{\tau}\right)^n e^{-t/\tau} \quad (194)$$

Il picco si trova ad un tempo $t_{peak} = n\tau$, ed ha un valore $V_{peak} = n^n e^{-n}$, che diverge rapidamente all'aumentare di n . Scegliendo opportuni valori di capacità e resistenza, possiamo riscalarare la costante di tempo in funzione

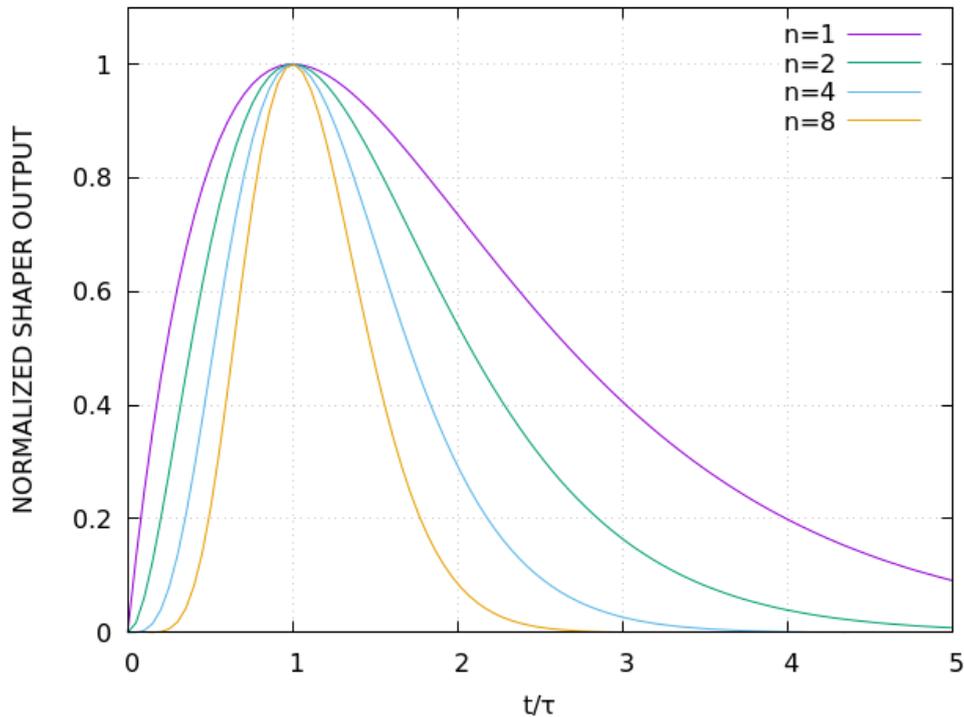


Figura 70: Forma del segnale in uscita dal formatore, quando esso è costituito da un differenziatore ed n integratori (circuito $CR - nRC$), tutti con stessa costante di tempo $\tau_{diff} = \tau_{int} = \tau$.

del numero n di integratori, come τ/n . In tal modo fissiamo la posizione del picco in $t_{peak} = \tau$, mentre la sua altezza rimane quella di prima. Se ora normalizziamo ad 1 anche quest'ultima, otteniamo proprio il grafico in fig. 70. Osserviamo che, all'aumentare di n , il segnale in uscita dal formatore decade in tempi via via più brevi, assumendo una forma più simmetrica e pronunciata in prossimità del picco. Ciò si riflette sulla risoluzione temporale, ed è utile per evitare fenomeni di *pile-up* nel caso in cui il rateo di eventi sia elevato.

Nelle applicazioni sperimentali, i sensori hanno dei circuiti integrati che permettono di variare la costante di tempo di *shaping* τ in un determinato intervallo. Si effettuano sia delle misure di segnale, irradiando il sensore con una sorgente, sia di rumore, ovvero in assenza della sorgente, determinando dunque il rapporto segnale-rumore. Effettuando tali misure per diversi valori di τ , possiamo stabilirne il valore più opportuno, ossia quello che minimizza il rapporto segnale-rumore.

6 Risoluzione spaziale dei rivelatori

Una volta affrontato il problema della rivelazione della particella possiamo cominciare a pensare di soddisfare ulteriori requisiti, quali per esempio il raggiungimento di una buona **risoluzione spaziale**, per esempio ai fini di operare un tracciamento della particella.

Per determinare la posizione di una particella che attraversa un sensore, possiamo utilizzare una **segmentazione naturale**, ovvero una configurazione con più rivelatori di piccola dimensione di cui conosciamo la posizione, ad esempio disponendoli su una griglia, e vedere quale di essi produce segnale. Questo modo di operare tuttavia è **poco efficiente** poiché si moltiplicano, oltre ai sensori, anche i volumi non sensibili (sostegni, circuiti di alimentazione e di raccolta dei dati, sistemi di raffreddamento, etc...). Risulta invece più efficiente **segmentare internamente** il volume sensibile di un singolo rivelatore, per esempio dividendo la giunzione in strisce parallele, come nel caso dei rivelatori a **microstrisce** in cui la superficie è segmentata lungo una coordinata, oppure in una matrice di **pixel** in cui la superficie è segmentata in entrambe le coordinate, come appunto per i rivelatori a pixel. Ognuna delle strisce, o dei pixel, costituisce in questo caso una cella sensibile, il cui funzionamento si basa sul meccanismo della giunzione *pn*.

Vediamo quindi la **risoluzione spaziale** dei rivelatori internamente segmentati in **celle**, su una sola dimensione. La prima caratteristica rilevante è la larghezza p della singola cella (in questo caso striscia) in cui è segmentata la coordinata. Tale larghezza è detta **passo** (o *pitch*). Una particella ha a priori probabilità uniforme di colpire qualunque punto della dimensione trasversa della striscia.

Altra caratteristica importante è la **modalità di lettura del segnale**, che può essere:

- **digitale**, quando in uscita da ogni cella abbiamo solo un segnale a due stati, 1 o 0, a seconda che il segnale raccolto abbia superato la soglia impostata o meno;
- **analogica**, se invece viene misurata l'ampiezza del segnale, in tutta la sua durata temporale.

La lettura digitale si ottiene utilizzando un **discriminatore**, che produce in uscita un segnale digitale qualora quello analogico in ingresso superi in ampiezza una soglia prefissata.

Gioca un ruolo importante anche la **condivisione di carica**. Per una particella incidente con direzione perpendicolare al sensore entro il perimetro di una determinata cella, se il rilascio di carica avviene vicino al confine con le celle adiacenti, la rete resistiva o capacitiva esistente tra celle adiacenti (o anche piu' lontane) puo' far si' che la carica venga in realta' raccolta non solo dall'elettrodo della cella in cui la carica e' stata rilasciata, ma anche dalle altre. Possiamo determinare la **dimensione del cluster**, ovvero del gruppo di celle adiacenti che vengono attivate. In lettura **analogica**, la valutazione del **baricentro del rilascio di carica** permette di determinare la posizione in cui è passata la particella con maggiore precisione.

Ad esempio, se vengono attivate due celle, che danno segnali di eguale ampiezza, ne deduciamo che la particella è passata a metà fra di esse. Altrimenti, se il segnale su una cella è più grande, mentre quello sull'altra è più piccolo, allora la particella sarà passata più vicina al centro della prima, con un peso proporzionale al segnale raccolto nella singola cella.

Infine, una particella che incide con angolo di verso da 90° , attivera' naturalmente piu' celle vicine, anche in assenza di condivisione di carica: conoscere quanta carica è stata rilasciata su ciascuna di esse è cruciale per ricostruire in modo accurato la traiettoria.

Consideriamo ora il caso di un sensore a strisce, segmentato su una sola dimensione, con le seguenti caratteristiche:

1. il *pitch* p delle strisce è costante;
2. non c'è condivisione di carica (i.e. le singole strisce sono perfettamente isolate fra loro);
3. la lettura viene effettuata in modalità digitale.

Supponendo che la traiettoria della particella sia sufficientemente perpendicolare da passare attraverso una sola striscia, e non avendo condivisione di carica tra strisce vicine, la carica di segnale generata appena sotto la superficie viene raccolta soltanto sulla striscia in corrispondenza della quale è transitata la particella. In assenza di ulteriori informazioni, possiamo solo supporre che la posizione misurata sia uniformemente distribuita lungo la larghezza p della striscia interessata. Se essa è centrata in $x = 0$, possiamo dunque quantificare l'incertezza mediante la varianza:

$$\sigma_x^2 = \langle \Delta x^2 \rangle = \int_{-p/2}^{p/2} \frac{x^2}{p} dx = \frac{1}{p} \left[\frac{x^3}{3} \right]_{-p/2}^{p/2} = \frac{1}{3p} \left[\frac{p^3}{8} - \left(-\frac{p^3}{8} \right) \right] = \frac{1}{3p} \frac{2p^3}{8} = \frac{p^2}{12} \quad (195)$$

Segue un errore, in termini di deviazione standard sulla misura, di $\sigma_x = p/\sqrt{12}$. Questo valore segna il limite massimo per la risoluzione spaziale di un rivelatore uniformemente segmentato. Tale stima viene normalmente usata come valore pessimistico, o limite superiore, della risoluzione spaziale, prodotta dalla sola segmentazione. Esso ovviamente puo' essere migliorato in presenza di lettura analogica e di condivisione di carica.

Per migliorare la risoluzione possiamo ridurre dunque la dimensione p del passo. Esistono però limitazioni di carattere pratico. Ad esempio, non possiamo realizzare una cella che sia più piccola degli stessi contatti necessari a portarne fuori il segnale.

Oltre a ciò, esistono anche limitazioni intrinseche, derivanti dalla fisica del rivelatore: consideriamo ora della carica generata lontano dalla superficie del sensore, che abbia dunque maniera di spostarsi per deriva verso gli elettrodi. Il moto di diffusione delle cariche che coesiste durante la raccolta per deriva contribuisce a sparpagliare la carica durante lungo il loro percorso. Tale sparpagliamento e' un limite intrinseco alla risoluzione, nel senso che per quanto io diminuisca il passo di cella, la nuvola di cariche giunte agli elettrodi sara' comunque piu' larga rispetto al volume in cui esse erano state generate inizialmente. In questo caso non è vantaggioso scegliere un valore di *pitch* inferiore alla larghezza della nube di carica che giunge all'elettrodo, cosa che, di fatto, produrrebbe segnale su due (o più) strisce senza un miglioramento della risoluzione. Questo effetto quindi puo' dare un'indicazione utile per determinare il passo minimo tra celle.

Vediamo di descrivere meglio l'effetto della **diffusione**. In riferimento alla fig. 71, il moto di deriva avviene lungo la coordinata y della giunzione (i.e. perpendicolare al piano degli elettrodi), mentre x è la coordinata trasversa (i.e. parallela al piano degli elettrodi). Lungo questa coordinata, per diffusione, si produce un allargamento gaussiano della nube di carica, la cui deviazione standard viene quantificata come:

$$\sigma_x = \sqrt{2Dt_c} = \sqrt{2 \frac{\mu kT}{q} t_c} \quad (196)$$

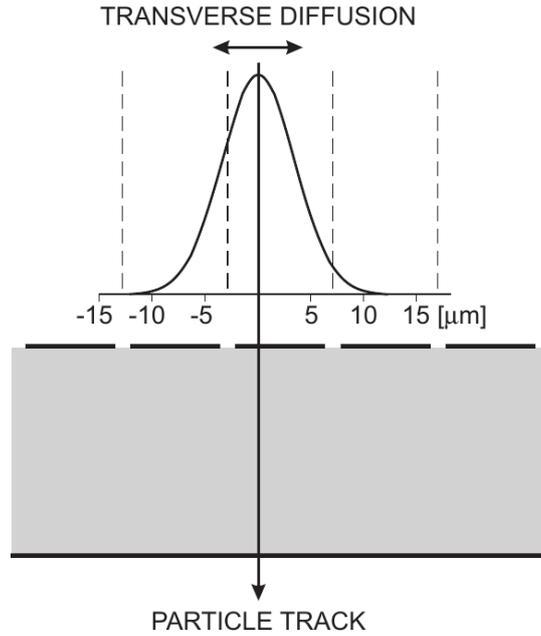


Figura 71: Diffusione trasversa (in direzione x) della carica sugli elettrodi di raccolta di un sensore a strisce con *pitch* di 20 μm . I limiti degli elettrodi su cui viene raccolta la carica sono evidenziati da linee tratteggiate.

dove t_c è il tempo di raccolta della carica, mentre D è il coefficiente di diffusione, che è stato riscritto nell'ultimo passaggio in funzione della mobilità μ , utilizzando la nota relazione di Einstein.

Se d è lo spessore del rivelatore, nel caso tipico di una giunzione unilaterale a gradino, la tensione di svuotamento completo in può essere scritta, utilizzando le relazioni 64,

$$d \simeq \sqrt{\frac{2\epsilon V}{qN_D}} \quad (197)$$

come:

$$V = \frac{qN_D d^2}{2\epsilon} \quad (198)$$

Il campo elettrico sul percorso, può essere approssimato come campo medio:

$$E \simeq \frac{V}{d} = \frac{qN_D d}{2\epsilon} \quad (199)$$

Ipotizzando che la carica venga rilasciata a **metà spessore del rivelatore**, essa viene raccolta in un tempo pari a:

$$t_c = \frac{d/2}{v_{drift}} = \frac{d}{2\mu E} = \frac{d}{2} \frac{2\epsilon}{\mu q N_D d} = \frac{\epsilon}{\mu q N_D} \quad (200)$$

Utilizzando questo risultato nell'equazione 196, otteniamo:

$$\sigma_x = \sqrt{2 \frac{\mu k T}{q} t_c} = \sqrt{\frac{2\mu k T}{q} \frac{\epsilon}{\mu q N_D}} = \sqrt{\frac{2k T \epsilon}{q^2 N_D}} \quad (201)$$

che non dipende dal particolare tipo di portatore di carica. La scelta del *pitch* p sarà condizionata dallo sparpagliamento delle cariche, per cui richiediamo $\sigma_{x,p} \sim \sigma_{x,d}$.

Esempio. Consideriamo un sensore costituito da una giunzione unilaterale p^+n operata a temperatura ambiente. Esso ha uno spessore $d = 300 \mu\text{m}$, ed una concentrazione di donori dell'ordine di $N_D \sim 10^{12} \text{ cm}^{-3}$. La tensione da applicare per ottenere lo svuotamento completo è:

$$\begin{aligned} V &= \frac{q N_D d^2}{2\epsilon} = \frac{1.6 \cdot 10^{-19} \text{ C} \cdot 10^{12} \text{ cm}^{-3} \cdot (3.00 \cdot 10^{-2} \text{ cm})^2}{2 \cdot 1.035 \cdot 10^{-12} \text{ F cm}^{-1}} = \\ &= \frac{1.6 \cdot 9.00}{2 \cdot 1.035} \cdot 10^{-19+12-4+12} \text{ V} = 6.83 \cdot 10^{-1} \text{ V} = 68.3 \text{ V} \end{aligned}$$

Il tempo di raccolta delle buche, prodotte a metà dello spessore della giunzione, è:

$$\begin{aligned} t_{c,h} &= \frac{\epsilon}{\mu_h q N_D} = \frac{1.035 \cdot 10^{-12} \text{ F cm}^{-1}}{450 \text{ cm}^2 (\text{V s})^{-1} \cdot 1.60 \cdot 10^{-19} \text{ C} \cdot 10^{12} \text{ cm}^{-3}} = \\ &= \frac{1.035}{4.50 \cdot 1.60} \cdot 10^{-2-12+19-12} \text{ s} = 0.146 \cdot 10^{-7} \text{ s} = 14.6 \cdot 10^{-9} \text{ s} = 14.6 \text{ ns} \end{aligned}$$

Tenendo conto che $kT = 0.026 \text{ eV}$, il coefficiente di diffusione delle buche è pari a:

$$D_{x,h} = \mu_h \frac{kT}{q} = 450 \frac{\text{cm}^2}{\text{V s}} \cdot 0.026 \text{ V} = 11.7 \frac{\text{cm}^2}{\text{s}}$$

Segue la deviazione standard per la diffusione:

$$\sigma_x = \sqrt{2 D_{x,h} t_{c,h}} = \sqrt{2 \cdot 11.7 \frac{\text{cm}^2}{\text{s}} \cdot 1.46 \cdot 10^{-8} \text{ s}} = \sqrt{2 \cdot 11.7 \cdot 1.46 \cdot 10^{-4}} \text{ cm} = 5.84 \mu\text{m}$$

Usando la relazione 195, vediamo che il *pitch* ottimale sarà dell'ordine di:

$$p \simeq 2\sqrt{3} \sigma_{x,d} = 2\sqrt{3} 5.84 \mu\text{m} \simeq 20 \mu\text{m}$$

Nel caso di un rivelatore a microstrisce avente lunghezza complessiva $l = 10 \text{ cm}$, il numero di canali richiesti sarà dell'ordine di:

$$N_{ch} \sim \frac{l}{p} = \frac{10 \text{ cm}}{20 \mu\text{m}} = \frac{10^5 \mu\text{m}}{20 \mu\text{m}} \sim 10^{4\div 5}$$

Ciascun canale viene letto ed alimentato separatamente, quindi un'elevata segmentazione comporta notevoli problematiche di gestione, sia per quanto riguarda l'elettronica esterna, sia per la dissipazione del calore. Vogliamo dunque trovare un metodo che ci permetta di ottenere la medesima informazione, ma senza dover leggere tutti i canali.

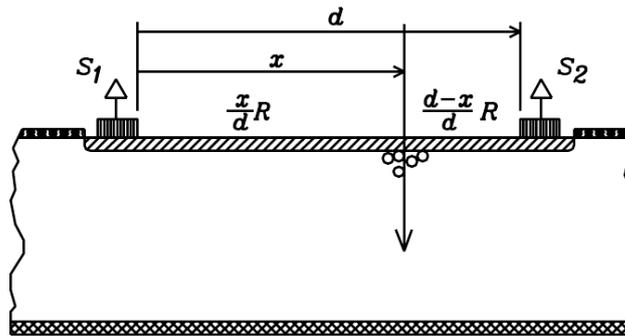


Figura 72: Misura della posizione, con l'utilizzo della divisione resistiva della carica.

A tale scopo, si può utilizzare la **divisione resistiva della carica**, descritta in fig. 72. Vediamo che i due elettrodi di raccolta sono posti ad una distanza d l'uno dall'altro. Lo strato superficiale che li separa costituisce un mezzo resistivo, di resistenza R . La carica viene rilasciata ad una distanza x dal primo, ovvero ad una distanza $d - x$ dal secondo. La quantità di carica raccolta da ciascuno dei due elettrodi sarà proporzionale al reciproco della resistenza che essa incontra sul suo cammino, per cui avremo:

$$S_1 : \frac{1}{R_1} = S_2 : \frac{1}{R_2} \quad \text{i.e.} \quad S_1 R_1 = S_2 R_2 \quad (202)$$

I due valori di resistenza possono essere scritti in funzione di R , come:

$$R_1 = \frac{x}{d}R \quad \text{ed} \quad R_2 = \frac{d-x}{d}R \quad (203)$$

che ci permette di riscrivere la relazione di prima come:

$$S_1x = S_2(d-x) \quad (204)$$

La posizione x del rilascio di carica può quindi essere ricavata da quest'ultima, in funzione dell'ampiezza dei segnali S_1 ed S_2 misurati, e della distanza nota d fra gli elettrodi, come:

$$x = \frac{d S_2}{S_1 + S_2} \quad (205)$$

L'errore sulla posizione può essere calcolato propagando quelli sulle ampiezze dei due segnali:

$$\begin{aligned} \langle \Delta x \rangle &= \frac{\partial x}{\partial S_1} \Delta S_1 + \frac{\partial x}{\partial S_2} \Delta S_2 = \\ &= -\frac{d S_2}{(S_1 + S_2)^2} \Delta S_1 + \frac{d (S_1 + S_2) - d S_2}{(S_1 + S_2)^2} \Delta S_2 = \\ &= \left(\frac{d}{S_1 + S_2} - \frac{d S_2}{(S_1 + S_2)^2} \right) \Delta S_2 - \frac{d S_2}{(S_1 + S_2)^2} \Delta S_1 = \\ &= \left(\frac{d}{S_1 + S_2} - \frac{x}{S_1 + S_2} \right) \Delta S_2 - \frac{x}{S_1 + S_2} \Delta S_1 = \\ &= (d-x) \frac{\Delta S_2}{S_1 + S_2} - x \frac{\Delta S_1}{S_1 + S_2} \end{aligned} \quad (206)$$

da cui segue lo scarto quadratico medio:

$$\langle \Delta x^2 \rangle = \frac{1}{(S_1 + S_2)^2} \left[(d-x)^2 \langle \Delta S_2^2 \rangle - 2x(d-x) \langle \Delta S_1 \Delta S_2 \rangle + x^2 \langle \Delta S_1^2 \rangle \right] \quad (207)$$

Gli errori ΔS_1 e ΔS_2 sono dovuti al rumore, le cui sorgenti principali sono:

1. le **resistenze**, R_1 ed R_2 , che sono totalmente correlate, poiché la carica che non passa per la prima, passa attraverso la seconda, e viceversa. Nota la carica rilasciata, i rumori dovuti alle due resistenze si compensano, per cui $\Delta S_{R1} = -\Delta S_{R2} \equiv \Delta S_R$, ed il termine di correlazione è $\langle \Delta S_{R1} \Delta S_{R2} \rangle = -\langle \Delta S_R^2 \rangle$;

2. gli **amplificatori**, a_1 ed a_2 , che sono sostanzialmente scorrelati, per cui $\langle \Delta S_{a1} \Delta S_{a2} \rangle = 0$. Assumendo che entrambi producano rumore con le medesime proprietà, possiamo scrivere $\Delta S_{a1} = \Delta S_{a2} \equiv \Delta S_a$.

Tenendo conto di queste sorgenti di rumore nella formula 207, abbiamo:

$$\begin{aligned}
\langle \Delta x^2 \rangle &= \frac{1}{(S_1 + S_2)^2} \left[(d-x)^2 \langle \Delta S_{R2}^2 \rangle - 2x(d-x) \langle \Delta S_{R1} \Delta S_{R2} \rangle + x^2 \langle \Delta S_{R1}^2 \rangle + \right. \\
&\quad \left. + (d-x)^2 \langle \Delta S_{a2}^2 \rangle - 2x(d-x) \langle \Delta S_{a1} \Delta S_{a2} \rangle + x^2 \langle \Delta S_{a1}^2 \rangle \right] = \\
&= \frac{1}{(S_1 + S_2)^2} \left[\langle \Delta S_R^2 \rangle [(d-x)^2 - 2x(d-x) + x^2] + \langle \Delta S_a^2 \rangle [(d-x)^2 + x^2] \right] = \\
&= \frac{1}{(S_1 + S_2)^2} \left[\langle \Delta S_R^2 \rangle d^2 + \langle \Delta S_a^2 \rangle [d^2 - 2xd + 2x^2] \right] = \\
&= \frac{d^2}{(S_1 + S_2)^2} \left[\langle \Delta S_a^2 \rangle \left(1 - 2\frac{x}{d} + 2\frac{x^2}{d^2} \right) + \langle \Delta S_R^2 \rangle \right]
\end{aligned} \tag{208}$$

Osserviamo che l'errore dipende dalla particolare coordinata x del punto di passaggio della particella. In particolare, risulta massimo in corrispondenza degli elettrodi (in $x = 0$ ed $x = d$) e minimo a metà fra essi (in $x = d/2$). Da questa analisi la divisione resistiva risulta dunque di difficile applicazione per ottenere una misura con un'adeguata risoluzione a prescindere dalla posizione stessa da misurare.

Per ovviare a questo problema, si può leggere il segnale in regime di **divisione capacitiva della carica**. Come visibile in fig 73, data la geometria del sensore e degli impianti che formano le singole giunzioni pn , al suo interno si instaurano automaticamente delle capacità. Ogni striscia presenta infatti una **capacità con il backplane** C_b fra l'elettrodo superiore e quello inferiore, inoltre fra strisce adiacenti si formano **capacità interstrip** C_{ss} . Ulteriori accoppiamenti capacitivi con le strisce non adiacenti provocano una divisione di carica anche a più grande raggio, di cui è necessario tener conto per una trattazione accurata.

Chiamiamo con C_1 e C_2 le capacità equivalenti che la carica di segnale vede a sinistra e a destra del suo punto di origine, sino agli elettrodi, sui quali vengono rispettivamente letti i segnali S_1 ed S_2 . Avremo che:

$$S_1 : C_1 = S_2 : C_2 \quad \text{i.e.} \quad S_1 C_2 = S_2 C_1 \tag{209}$$

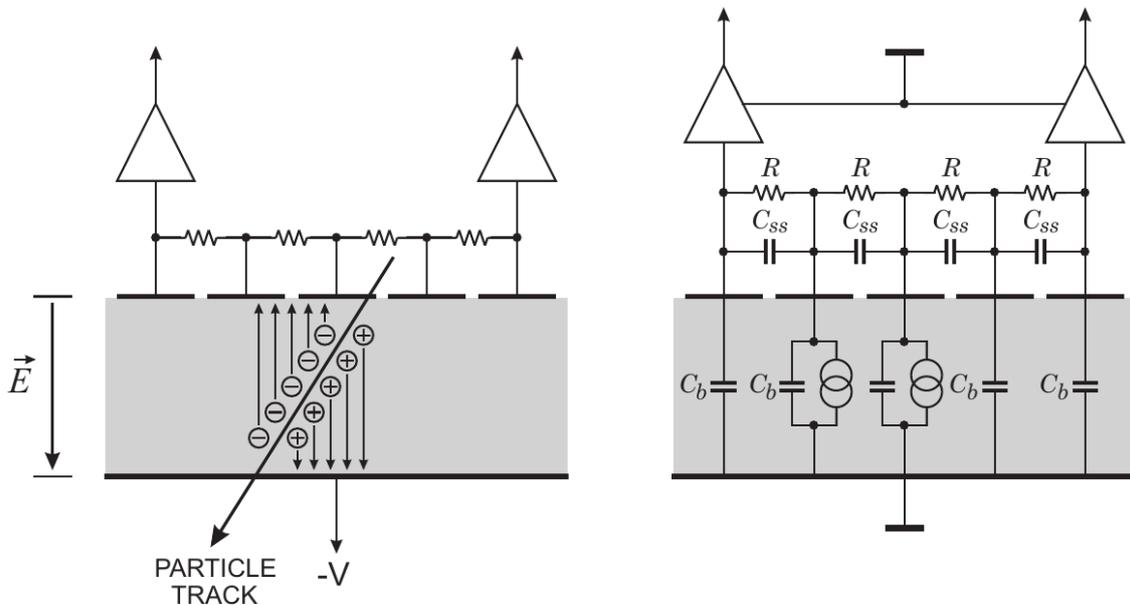


Figura 73: Lettura del segnale con divisione capacitiva della carica. A sinistra: schema reale del sensore, in cui è visibile la traccia della particella ionizzante, che rilascia la carica sul secondo e terzo elettrodo. A destra: schema circuitale del sensore: i generatori di corrente rappresentano il rilascio di carica, C_b è la capacità fra l'elettrodo (superiore) ed il *backplane* (inferiore), mentre C_{ss} è la capacità interstrip, ossia *strip-to-strip*, fra due strisce adiacenti.

Osserviamo che per trasportare la carica agli amplificatori, bisogna fare in modo che non venga raccolta sul *backplane*: per tale ragione si richiede che sia almeno $C_{ss} \sim 10C_b$.

La presenza di **elementi resistivi** è dovuta allo strato di ossido di silicio che usualmente ricopre il sensore. Esso è necessario sia per esigenze meccaniche, sia come isolante elettrico. L'alimentazione delle strisce non può dunque avvenire sfruttando unicamente i contatti degli elettrodi di raccolta, infatti, come dallo schema in figura, si otterrebbe un partitore di tensione: in tal caso le strisce verrebbero alimentate a tensioni diverse, e la carica sarebbe raccolta soltanto su quella a potenziale maggiore. Per tale motivo è necessario alimentare singolarmente ogni striscia, cosa che comporta la presenza di molti contatti di alimentazione.

Esempio. Vediamo il caso pratico dei rivelatori Alice-SSD (*Silicon-Strip Detector*). Si tratta di rivelatori a microstrisce con *pitch* $p \simeq 95 \mu\text{m}$ ed una lunghezza $l = 7 \text{ cm}$ del lato segmentato, per cui vi sono all'incirca

768 canali. Si tratta di un numero relativamente piccolo, per cui è materialmente possibile leggere ciascuno di questi canali, in accoppiamento resistivo della carica.

Se il primo elettrodo è posizionato in x_1 mentre il secondo è posizionato in x_2 , allora la posizione x di passaggio della particella può essere calcolata dal **baricentro del rilascio di carica**, utilizzando il risultato in formula 204, come:

$$x = \frac{S_1 x_1 + S_2 x_2}{S_1 + S_2}$$

Tale sistema, in gergo, costituisce un *cluster* di dimensione 2.

Il limite inferiore per la risoluzione spaziale sarà dato dal rapporto fra il *pitch* ed il rapporto segnale-rumore. Siccome per i sensori ALICE-SSD è $S/R = 30$, otteniamo:

$$\Delta x \geq \frac{p}{S/R} = \frac{95 \mu\text{m}}{30} \simeq 3 \mu\text{m}$$

Come ultima considerazione, notiamo che, rispetto all'accoppiamento resistivo, quello capacitivo richiede una **misura molto precisa delle capacità** in gioco, al fine di avere un basso errore sulla posizione. In tal caso, la componente più rilevante dell'errore è dovuta al **rumore elettronico** introdotto dall'amplificatore.

Nel tentativo di ottimizzare la risoluzione della posizione, è necessario considerare anche altri effetti. Le particelle minimo ionizzanti non depositano carica in modo uniforme lungo la traccia, quindi tracce che incidono nella stessa posizione mostreranno centri di carica indotta variabili. Un ulteriore limite alla risoluzione della posizione è imposto dagli elettroni delta-rays energetici formati lungo la traiettoria della traccia, che possono alterare significativamente il centroide della carica.