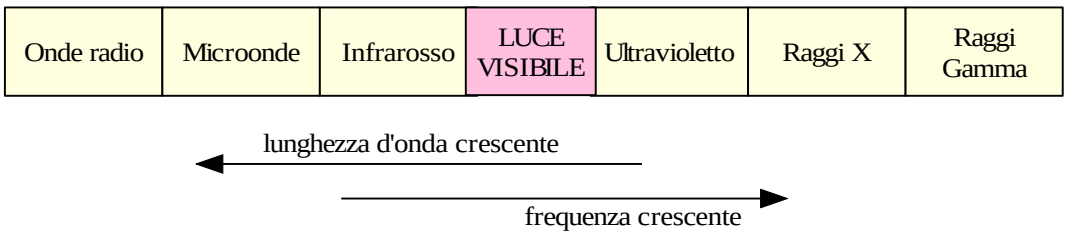
Appunti di Comunicazioni elettriche

Capitolo 8 - Trasmissione su fibre ottiche

INTRODUZIONE

In questo capitolo ci concentriamo sui sistemi di trasmissione che operano nel campo di frequenze del visibile e dell'infrarosso.

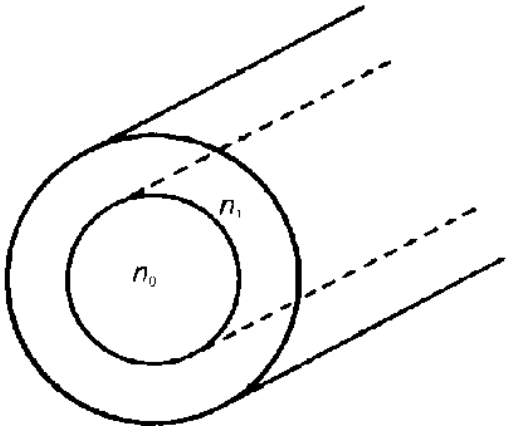
Ricordiamo, a questo proposito, che si definisce **spettro delle radiazione elettromagnetiche** l'insieme delle radiazioni caratterizzate da tutte le possibili lunghezze d'onda. Tale spettro copre un intervallo di lunghezze d'onda molto esteso, approssimativamente da 10^{-11} cm (valore minimo) e 10^{4} cm (valore massimo). A seconda dei valori delle lunghezze d'onda, si distinguono alcuni tipi di radiazioni, indicate nella figura seguente:



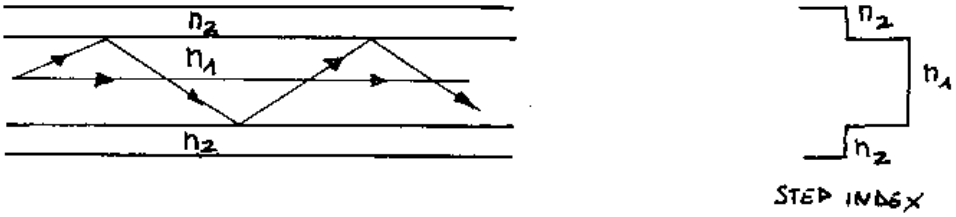
L'**infrarosso** è dunque la regione dello spettro elettromagnetico situata tra la regione della *luce visibile* e quella delle *microonde*, con lunghezza d'onda compresa tra 0.75μm e 1mm (le corrispondenti frequenze vanno da 300 GHz a 400 THz).

Nonostante sia possibile usare, su brevi tratte, la trasmissione irradiata, il mezzo di gran lunga più attraente, per l'uso di queste frequenze, è la **fibra ottica**, che lavora a frequenze dell'ordine di **10^{14} Hz** (100 THz).

La fibra ottica è una guida d'onda dielettrica. Nel caso più semplice, essa è costituita da un cilindro centrale, detto **nucleo** (a sezione circolare), di materiale dielettrico con indice di rifrazione n_0, rivestito da un involucro dielettrico, detto **mantello**, coassiale con il nucleo, con indice di rifrazione n_1 minore del precedente:

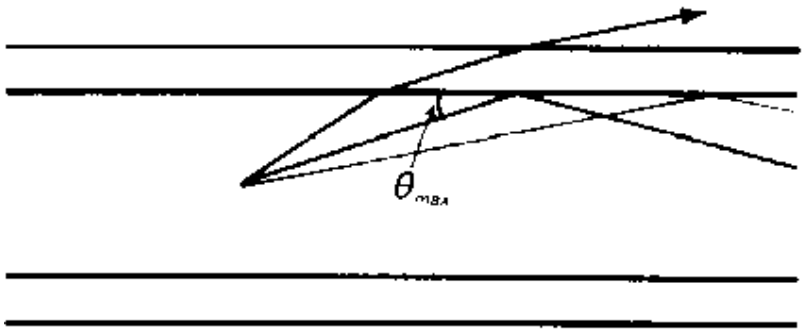


In particolare, quella indicata in figura è una **fibra step-index**, nella quale cioè le variazioni di indice di rifrazione dal nucleo al mantello sono brusche:

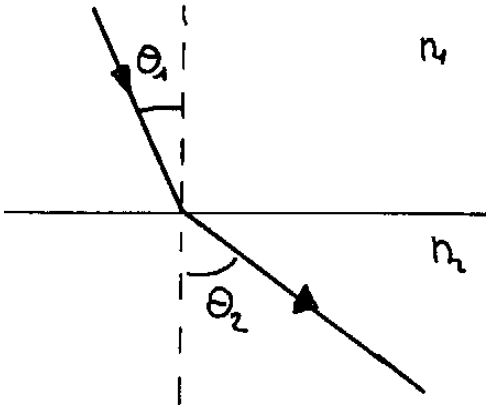


E' possibile dimostrare, usando le equazioni del campo elettromagnetico, che la differenza dei valori di indice di rifrazione, che di solito si esprime mediante la cosiddetta **apertura numerica**, data da $\Delta = \sqrt{n_0^2 - n_1^2}$, è essenziale se si vuole assicurare la propagazione nel nucleo dei **modi guidati**. *Per ciascuno di questi modi, esiste una frequenza critica (che dipende sia dalla geometria sia dall'apertura numerica D) al di sotto della quale il modo non può propagarsi.* Fa eccezione, a questa regola, solo il cosiddetto *modo fondamentale*, che è sempre presente.

Una descrizione semplice del meccanismo di propagazione all'interno di una fibra si può ottenere usando i concetti dell'*ottica geometrica*, come fatto nella figura seguente:



Ogni radiazione luminosa è rappresentata da un *raggio*, che ne individua la direzione di propagazione. Ogni raggio incide, sulla superficie di separazione tra nucleo e mantello, con un angolo di incidenza θ diverso. Sappiamo che ogni raggio subisce, in generale, sia il fenomeno della riflessione nello stesso mezzo da cui proviene sia il fenomeno della rifrazione nell'altro mezzo:



Se θ_1 è l'*angolo di incidenza* (relativo ad un mezzo con indice di rifrazione n_1) e θ_2 l'*angolo di rifrazione* (relativo ad un mezzo con indice di rifrazione n_2), i due angoli sono legati dalla nota **legge di Snell**, secondo la quale

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

Questa legge indica, in pratica, che *la radiazione tende ad incurvarsi, allontanandosi dalle zone meno dense (cioè con indice di rifrazione maggiore).*

Esiste un particolare **angolo critico** $\theta_{max} = \sqrt{n_0^2 - n_1^2}/n_0$ per l'incidenza: per incidenza con angolo $\theta=\theta_{max}$, la radiazione che emerge si propaga lungo la superficie di discontinuità dei due materiali; tutti i raggi che incidono invece con θ maggiore di θ_{max}, subiscono il fenomeno della **riflessione totale**, per cui rimangono all'interno del nucleo; i rimanenti raggi, infine, attraversano lo strato di separazione, dopo di che o vengono riflessi nuovamente nel mantello, nel quale subiscono una brusca attenuazione, oppure escono dalla guida.

Quindi, *all'interno del nucleo si realizza l'intrappolamento delle radiazioni che incidono con angolo maggiore di quello critico. Tutte le radiazioni che incidono con angolo minore di quello critico, vengono trasmesse nel mantello e si perdono.*

Se il materiale del nucleo è tale da dare attenuazione molto bassa (ossia non ci sono perdite per diffusione della radiazione o per effetto Joule), si ha una attenuazione molto più bassa di quella che si avrebbe in un cavo coassiale.

Osserviamo una cosa a questo proposito: in generale, noi sappiamo che le proprietà di un mezzo trasmissivo cambiano quando c'è una variazione percentuale sensibile della banda; nel caso delle fibre ottiche, però, *avendo una portante alla frequenza di circa 10^{14} Hz, si può tranquillamente assumere che le variazioni di banda siano piccole, per cui possiamo anche assumere che le proprietà del mezzo siano pressoché costanti.*

Nel momento in cui una **sorgente di radiazioni luminose** viene accoppiata ad una guida, parte della potenza luminosa viene iniettata nel nucleo e si propaga in esso.

Le fibre sono realizzare essenzialmente con biossido di silicio estremamente puro, il cui indice di rifrazione viene aumentato ricorrendo a tecniche di drogaggio (mediante germanio, fosforo o boro).

DISPERSIONE MODALE: FIBRE MULTIMODALI E MONOMODALI

Abbiamo detto che rimangono intrappolate nel nucleo tutte le radiazioni luminose che incidono con angolo superiore a quello critico θ_{max}. E' evidente che la distanza percorsa da ogni radiazione cambia seconda dell'angolo di incidenza. Dato, però, che la velocità di propagazione nel mezzo è costante (dato che è costante l'indice di rifrazione), ogni raggio uscirà dal nucleo con un ritardo di propagazione diverso dagli altri, a seconda del numero di riflessioni che esso ha subito.

Applicando le equazioni di Maxwell sull'elettromagnetismo, si ottiene che il numero dei raggi non è infinito, ma esiste un numero discreto di possibili angoli di incidenza. Ogni angolo di incidenza corrisponde ad un **modo** che si propaga nella fibra.

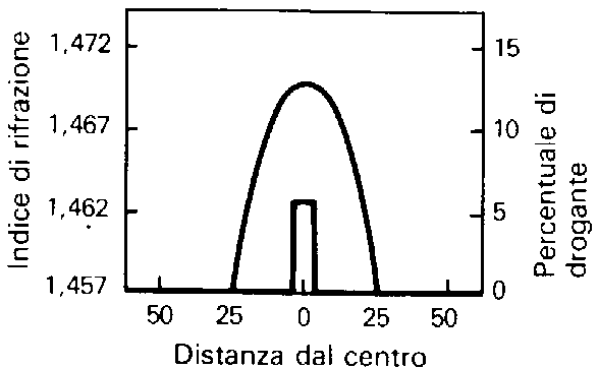
Una importante classificazione delle fibre è allora la seguente:

- si dicono **fibre multimodali** quelle in cui più modi possono propagarsi insieme;
- si dicono invece **fibre monomodali** quelle in cui un solo modo per volta si può propagare.

Consideriamo allora una fibra multimodale: in base a quanto detto prima, ogni modo emerge dalla fibra con un proprio ritardo[1]: il ritardo differenziale con cui i modi emergono dalla fibra è la cosiddetta **dispersione modale**, misurata in *nsec/km*. E' ovvio che la dispersione modale è tanto maggiore quanto più lunga è la fibra.

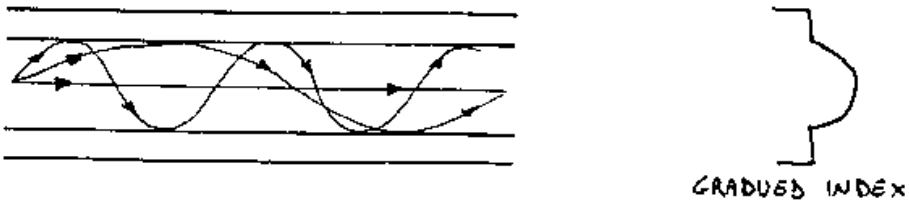
La dispersione modale costituisce chiaramente una limitazione all'impiego delle fibre: infatti, *pur avendo un mezzo con bassissima attenuazione, non lo si può usare per trasmettere a velocità molto alta, in quanto ...* Quindi, l'elemento limitante della capacità trasmissiva di una fibra ottica non è l'attenuazione, ma appunto la dispersione modale. Tipici valori dell'allargamento dell'impulso lungo la fibra, a causa della dispersione modale, sono dell'ordine di 0.3 nsec/km.

Si è cercato allora di minimizzare gli effetti della dispersione modale. Un modo abbastanza efficace di procedere è quello di realizzare un nucleo che, al posto di un indice di rifrazione costante, ne abbia uno decrescente dal centro (dove è massimo) alla periferia, come indicato nella figura seguente:



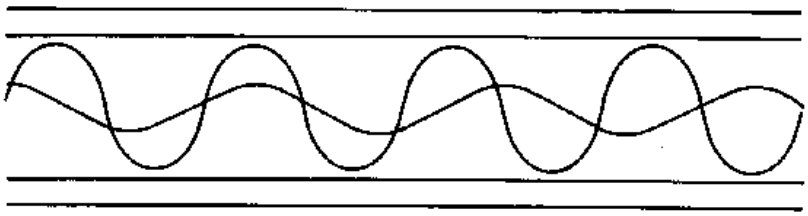
Queste sono le cosiddette **fibre gradued index**, nelle quali viene in pratica effettuato un drogaggio a più strati:

[1] In pratica, la dispersione modale corrisponde ad una dispersione dei tempi d'arrivo tra gli impulsi relativi a ciascun modo.



Con questo artificio, si ottengono diversi risultati:

- in primo luogo, le traiettorie dei raggi non sono più delle spezzate, ma si incurvano;
- in secondo luogo, il percorso dei raggi con angolo di incidenza più piccolo si svolge più vicino al centro: ne consegue che questi raggi, pur percorrendo un percorso più breve (come si nota bene nella figura seguente), lo compiono a velocità più bassa[1] dei raggi che seguono i percorsi più lunghi, cioè quelli che portano i raggi stessi ad urtare tra le regioni estreme del nucleo:



Questo fa si, chiaramente, che la dispersione modale risulti minore.

Quindi, *scegliendo il drogaggio della fibra in modo opportuno, si possono minimizzare gli effetti della dispersione modale.*

Il problema della dispersione modale non si pone, invece, nelle **fibre monomodali**[2], dove tutti i modi sono evanescenti tranne uno. *Affinché un solo modo possa propagarsi, è necessario però ridurre il diametro del nucleo a meno di 10mm,* il che ha un grosso svantaggio: quanto minore è il diametro del nucleo, tanto più difficile è l'iniezione, nel nucleo stesso, della potenza ottica. Si pone cioè il problema dell'**efficienza di iniezione**, intesa come rapporto tra la potenza effettivamente convogliata nel nucleo e quella totale prodotta: dato che non tutta la potenza prodotta viene inviata nel nucleo, ma una sua quota parte si perde nel mantello, questa efficienza non può che essere minore di 1 (ovviamente >0).

Un altro problema molto sentito nelle fibre monomodali è quello della cosiddetta **giuntatura**: dovendo porre in cascata due fibre ottiche, è evidentemente necessario allineare nel modo migliore possibile i due nuclei, in modo da perdere la minore potenza ottica possibile nel passare da una fibra all'altra ed è evidente che l'allineamento è tanto più importante quanto più piccolo è il diametro dei nuclei.

[1] Ricordiamo che la velocità di propagazione dei raggi è inversamente proporzionale alla radice quadrata dell'indice di rifrazione, per cui tale velocità sarà più bassa dove l'indice di rifrazione è maggiore, cioè appunto al centro del nucleo, mentre sarà via via minore man mano che ci si allontana dal nucleo stesso.

[2] Le fibre monomodali sono, attualmente, le uniche considerate nelle applicazioni

DISPERSIONE CROMATICA

Oltre alla dispersione modale, esaminata nel paragrafo precedente, bisogna tener conto della cosiddetta **dispersione cromatica**, dovuta al fatto che la velocità di propagazione della radiazione varia al variare della frequenza: infatti, la velocità di propagazione dipende dall'indice di rifrazione, il quale varia con la frequenza, provocando una non-linearità della *caratteristica di fase*[1] della fibra e quindi una variabilità con la frequenza della velocità di gruppo. Questo significa che le componenti spettrali di un segnale si propagano, nella fibra, con velocità diverse, il che comporta, evidentemente, che la loro ricomposizione in ricezione fornisca un segnale diverso da quello trasmesso.

DISPERSIONE SPAZIALE

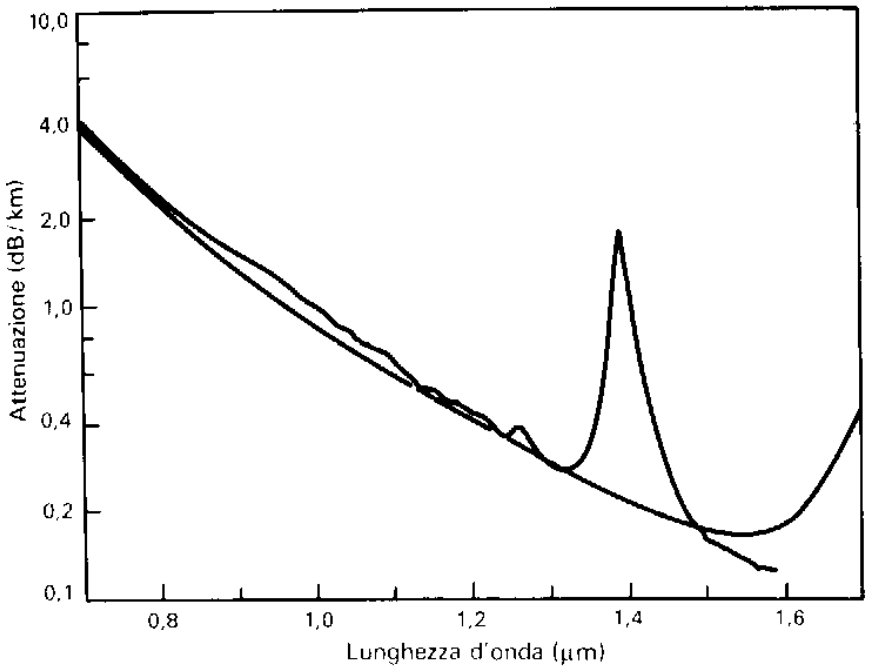
L'ultimo tipo di dispersione da considerare è dovuto al fatto che la superficie di discontinuità tra nucleo e mantello non potrà mai essere perfettamente cilindrica, ma presenterà delle discontinuità.

Ad ogni modo, esistono procedimenti tali da annullare questi termini di dispersioni. *Eliminando, allora, queste limitazioni, si giunge ad una fibra ottica con banda sostanzialmente illimitata.*

ATTENUAZIONE

Preoccupiamoci adesso dell'attenuazione nelle fibre ottiche.

L'onda ottica, propagandosi nella fibra, subisce una attenuazione dovuta a molteplici cause. La figura seguente ne mostra l'andamento in funzione della lunghezza d'onda[2], per una fibra di silice di alta qualità:



[1] Per "caratteristica di fase" intendiamo l'andamento della fase con la frequenza

[2] Ricordiamo che la lunghezza d'onda è inversamente proporzionale alla frequenza secondo la relazione $\lambda = c/f$

Nella regione in cui la fibra viene usata, se si eliminano assorbimenti particolari dovuti ad impurità (come per esempio il picco[1] che si nota per $\lambda \cong 1.4\mu m$), si trova che *la causa dominante di attenuazione è la diffusione di energia in conseguenza delle microvariazioni di indice di rifrazione dovute alle irregolarità nella struttura del materiale* (irregolarità che vengono tra l'altro accresciute dalla presenza di drogante).

E' importante osservare che l'attenuazione dovuta alla diffusione è proporzionale a $1/\lambda^4$. Accrescendo ulteriormente la lunghezza d'onda, l'attenuazione incomincia a salire a causa di assorbimenti dovuti a risonanze molecolari.

L'intervallo di utilizzazione della fibra è generalmente suddiviso in tre regioni, dette **finestre**, centrate approssimativamente a **0.85μm**, **1.4μm** e **1.55μm**:

- la **prima finestra** ha avuto le prime applicazioni, data la possibilità di reperire più facilmente sorgenti e rivelatori: infatti, l'energia del fotone corrispondente ad una lunghezza d'onda di 0.8μm è sufficiente a ionizzare atomi di silicio, per cui potevano funzionare fotodiodi e LED al silicio;
- la **seconda finestra** si trova nell'interno della lunghezza d'onda per cui la dispersione di annulla;
- la **terza finestra** è quella con attenuazione minima, che, nel caso monomodale illustrato in figura, ha un valore dell'ordine di **0.2 dB/km**.

La seconda e soprattutto la terza finestra sono le più promettenti per quanto riguarda l'attenuazione. E' però importante osservare che, *aumentando la lunghezza d'onda ai valori corrispondenti a tali due finestre, non si possono più usare dispositivi al silicio:* in particolare, per poter emettere in terza finestra è necessario utilizzare semiconduttori di tipo ternario.

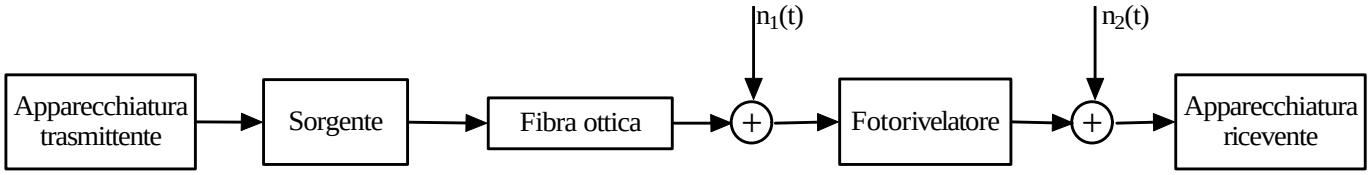
Valori dell'ordine di 0.2 dB/km non sono in alcun modo ottenibili su mezzi trasmissivi metallici come i cavi coassiali, a parità di dimensioni trasversali: si può affermare che il rapporto attenuazione/dimensione trasversale è, nella fibra, 450 volte più piccolo che in un cavo coassiale adoperato per una capacità trasmissiva di 140 Mbit/sec. Naturalmente, la situazione diventa ancora più favorevole alla fibra aumentando la velocità. E' questa una delle ragioni che rendono le fibre ottiche tanto interessanti per il futuro, unitamente al fatto che *l'attenuazione si può considerare pressoché costante in bande di frequenza estremamente ampie:* tanto per avere una idea, sempre con riferimento alla figura di prima, consideriamo il valore dell'attenuazione in corrispondenza di $\lambda = 1.55\mu m$; spostandoci di 1nm (cioè 3 ordini di grandezza in meno rispetto al μm) di lunghezza d'onda, l'attenuazione è praticamente ancora la stessa, ma una variazione di lunghezza d'onda di 1nm rispetto a 1.55μm equivale ad una variazione di frequenza di 125 GHz attorno ad una frequenza di $2 \cdot 10^{14}$ Hz, per cui possiamo ritenere che l'attenuazione sia la stessa in un intervallo di frequenza, attorno a $2 \cdot 10^{14}$ Hz, ampio circa 250 GHz.

Diciamo inoltre che, *proprio grazie ai valori estremamente bassi dell'attenuazione, è possibile realizzare, con le fibre ottiche, collegamenti su lunga distanza senza apparecchiature rigenerative intermedie.* D'altra parte, a fronte di attenuazioni così piccole, diventano rilevanti le attenuazioni dovute alle giunzioni, cui abbiamo accennato prima. Tra l'altro, tali giunzioni costituiscono anche delle discontinuità che possono innescare ulteriori modi anche nelle fibre monomodo, causando così perdite di potenza.

[1] Questo picco, a lunghezza d'onda di circa 1.4μm, risulta dovuto all'ossidrile OH, cioè si tratta del picco di assorbimento dovuto all'acqua: questo significa che è necessario proteggere la fibra dall'umidità.

SORGENTI

La bassissima attenuazione della fibra consente, negli attuali sistemi commerciali di tipo numerico binario, l'adozione di una struttura molto semplice, detta a **rivelazione diretta incoerente** per motivi che chiariremo più avanti, del tipo indicato in figura:



In questa struttura, i componenti ottici fondamentali sono la **sorgente ottica** ed il **fotorivelatore**. In questo paragrafo cominciamo ad occuparci delle sorgenti ottiche.

I dispositivi trasmettitori usati sono di due tipi: diodi LED (dove LED sta per Light Emitting Diode) e diodi laser. Entrambi questi dispositivi funzionano generalmente secondo una modulazione del tipo **OOK** (*On Off Keying*, ossia "o tutto o niente"): questo significa che vengono pilotati in modo da assegnare la massima intensità di radiazione ad uno dei simboli e intensità zero (o quasi zero) all'altro.

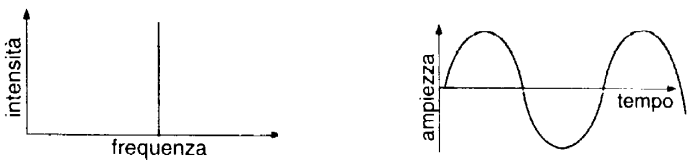
I **diodi LED**, eventualmente ad alta intensità, sono di più semplice impiego e di costo ridotto. Tuttavia, essi hanno diverse limitazioni, dovute all'incoerenza della luce emessa, alla notevole larghezza di riga (50 ÷ 100 nm) ed alla limitazione di banda dovuta all'eccessivo tempo di spegnimento.

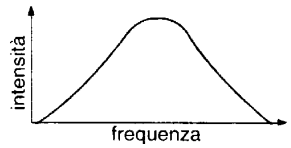
I **diodi laser** (**LD**), invece, hanno, oltre alla *coerenza spaziale* e alla radiazione direzionale (il che permette di iniettare più potenza nella fibra, aumentando così l'efficienza di iniezione), una purezza spettrale migliore (circa uguale a 2nm).

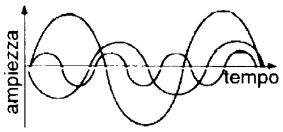
Richiami sulla luce coerente

Il *diodo laser* è un dispositivo che permette di ottenere fasci molto intensi di luce che, a differenza della luce ordinaria (per esempio quella del Sole o quella di una candela) presentano 2 caratteristiche fondamentali: si tratta di luce **monocromatica**, cioè tutta di una stessa lunghezza d'onda, e **coerente**, ossia i fotoni (o le onde) risultano tutti in fase nello stesso istante (la luce ordinaria è invece un insieme di lunghezze d'onda diverse, ossia non è monocromatica, e, inoltre, i vari fotoni interferiscono tra di loro e viaggiano in direzioni diverse; in altre parole, si disperdono nello spazio e non si mantengono in fase tra di loro, rendendo quindi la luce *incoerente*). Vediamo di capire bene cosa si intende per *luce coerente*.

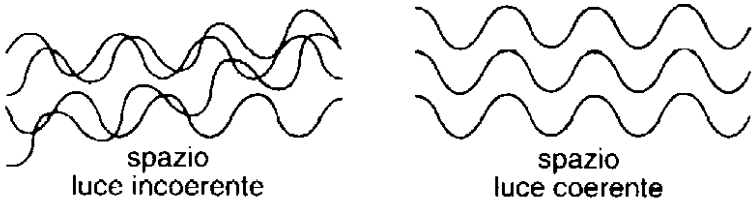
I pacchetti d'onda che compongono un fascio luminoso (stiamo adottando la descrizione ondulatoria della luce) hanno in genere una fase diversa, il che significa che i picchi e i ventri non coincidono:







Un fascio di luce è coerente (da un punto di vista spaziale) quando i vari pacchetti d'onda mantengono costanti nel tempo le relative differenze di fase. Si parla poi di coerenza temporale quando le fasi sono uguali (cioè quindi quando le differenze di fasi sono costantemente pari a 0).
In una normale sorgente luminosa, gli atomi emettono i vari pacchetti d'onda in modo non correlato e con lunghezze d'onda diverse: ne risulta che, tra una parte e l'altra della sorgente, le differenze di fase cambiano continuamente e casualmente.
In un laser, invece, il meccanismo della emissione stimolata fa sì che gli atomi emettano tutti in sincronia ed alla stessa lunghezza d'onda, per cui le differenze di fase sono mantenute costanti:



Il **laser** è costituito da un meccanismo di generazione di luce all'interno di una cavità risonante, chiusa agli estremi da due specchi semiriflettenti. La radiazione rimbalza su tali specchi, acquistando potenza ad ogni passaggio. Quando la potenza è sufficientemente elevata, la radiazione emerge, dando appunto luogo al **fascio laser**. Il dispositivo funziona fino a quando gli specchi sono in grado di riflettere la radiazione in modo costruttivo: questo accade in tutti i modi possibili di risonanza della cavità.

Dato che, per avere una certa efficienza di generazione della luce, il laser deve essere lungo parecchie lunghezze d'onda, ne consegue che le frequenze di risonanza sono abbastanza vicine tra loro. Effettuando infatti una analisi spettrale dell'emissione di un LD, si ottiene una concentrazione di potenza su bande di frequenza molto strette ma molto ravvicinate tra loro, come indicato nella figura seguente:

Se si volesse ripulire la radiazione, si possono realizzare, sulla superficie degli specchi, delle rugosità: così facendo, infatti, le risonanze che non servono vengono fatte interferire in modo distruttivo. Questi particolari laser vengono detti a **DFB** (che sta per *Laser a Feedback Distruttivo*) e con essi si riescono ad ottenere bande di 10 MHz e quindi bande relative molto piccole.

La velocità di trasmissione nelle fibre ottiche è limitata proprio dalla capacità delle sorgenti, soprattutto i LED, di pilotare con transizioni molto rapide e brusche.

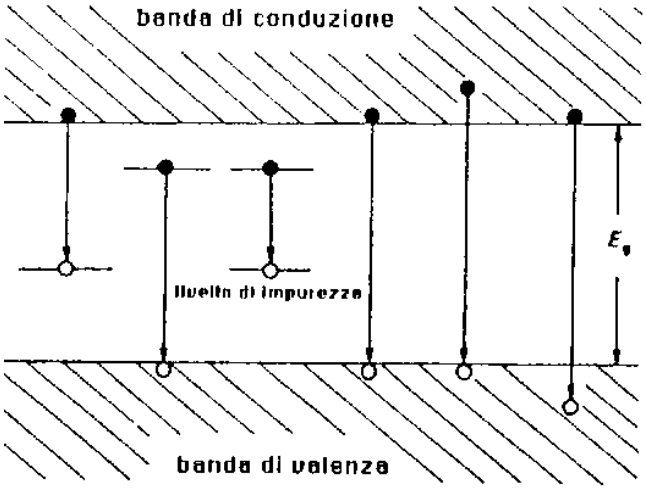
Cenni ai dispositivi optoelettronici

*I **dispositivi optoelettronici** sono quelli che convertono la luce in energia o viceversa* e, in alcuni casi, i due fenomeni possono aver luogo anche nello stesso dispositivo.

L'uso di ottica ed elettronica, piuttosto che della sola elettronica, è dettato dalle esigenze in termini di larghezza di banda, particolarmente nel campo delle telecomunicazioni e della elettronica di commutazione veloce usata per i computer. In particolare, le capacità parassite associate alle interconnessioni elettriche costituiscono un serio limite per la massima frequenza di funzionamento dei dispositivi e circuiti elettronici a larga banda, per cui si rende necessario l'uso di interconnessioni ottiche e di dispositivi optoelettronici, in grado cioè di trasformare segnali elettrici in segnali ottici e viceversa.

*Si definisce **elettroluminescenza** l'emissione di radiazione ottica (luce) causata dalla applicazione di un campo elettrico o di una corrente su un materiale sensibile.* Le applicazioni dei **dispositivi elettroluminescenti** riguardano essenzialmente campi in cui le caratteristiche spettrali della radiazione emessa, e quindi l'andamento dell'intensità ottica di uscita in funzione della lunghezza d'onda, non sono particolarmente importanti, il che significa che è tollerata una caratteristica di emissione larga, a differenza di quanto avviene nelle sorgenti di luce coerente, come i *diodi laser*, in cui la caratteristica spettrale è praticamente monocromatica.

I tipi di transizioni energetiche che avvengono, in un semiconduttore drogato, tra la banda di valenza e quella di conduzione e che possono dar luogo a emissione di luce sono indicate nella figura seguente:



Le prime tre da sinistra rappresentano transizioni in cui l'energia di passaggio è minore del gap di banda proibita (E_g) caratteristico del materiale: questo perché in tutti e tre i casi sono coinvolti livelli energetici situati all'interno del gap di banda proibita. Le altre tre transizioni, invece, consistono in un passaggio diretto, da banda a banda, con salti energetici superiori al gap di banda proibita.

Non tutte le transizioni mostrate rilasciano normalmente fotoni: *si parla di **materiale luminescente efficiente** quando le **transizioni radiative**, cioè appunto quelle accompagnate da emissione di fotoni, predominano sulle altre, dette non radiative.*

Esiste una precisa legge che lega l'intervallo Eg di energia proibita alla lunghezza dell'emissione fotonica corrispondente:

$$E_g = hf = h\frac{c}{\lambda}$$

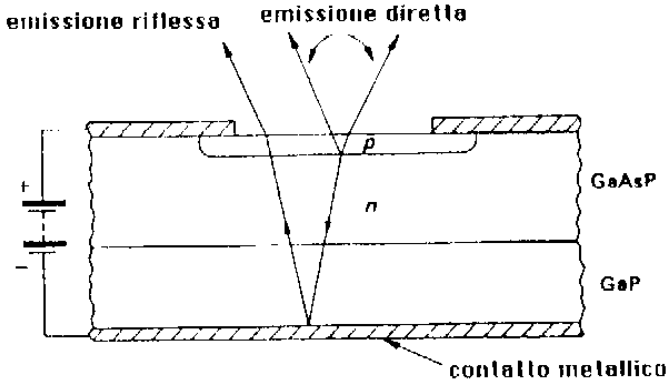
dove h è la costante di Planck, *f* la frequenza e λ la lunghezza d'onda del fotone emesso. Questa relazione consente dunque il calcolo della lunghezza d'onda di emissione di un cristallo semiconduttore avente gap proibito E_g:

$$\lambda = h\frac{c}{E_g}$$

Tipicamente, per l'emissione nel rosso, cioè con λ=635 nm, si richiede E_g circa pari a 1.95 eV.

Il LED

Il *diodo emettitore di luce* o **LED** (Light Emitting Diode) è un tipico dispositivo elettroluminescente, costituito da una giunzione pn in grado di emettere radiazioni spontanee nella regione dell'ultravioletto, del visibile o dell'infrarosso. Una tipica struttura di un LED è quella indicata nella figura seguente:



C'è uno strato di materiale semiconduttore composto come il GaAsP, cresciuto su un substrato di GaP, che è un materiale semiconduttore III-V. L'adozione di un substrato trasparente, come quello in GaP, permette l'emissione di luce più efficiente, visto che si ottiene la somma della luce emessa direttamente e di quella emessa verso il basso, nel substrato appunto, e che viene riflessa verso l'alto dal contatto metallico inferiore. Viene anche diffuso uno strato di tipo p al di sopra dello strato in GaAsP di tipo n, in modo da realizzare una giunzione pn.

L'alta efficienza di emissione di luce si può ottenere quando il LED è polarizzato direttamente, in quanto la corrente diretta risultante produce l'iniezione di elettroni di conduzione nella regione di giunzione. Quando questi elettroni di conduzione si ricombinano, l'energia viene rilasciata sotto forma di fotoni, con lunghezza d'onda data da $\lambda = h\frac{c}{E_g}$, come visto prima.

Il gran numero di fotoni emessi in questo dispositivo dipende dal fatto che si usa un materiale (il GaAsP) in cui la struttura a bande favorisce, con alta probabilità, il verificarsi di una transizione radiativa rispetto ad una non radiative. Questa caratteristica si ha comunque anche nel GaAs e, più in generale, nei cosiddetti *semiconduttori diretti*, nei quali cioè la transizione di un elettrone da una banda all'altra non richiede alcuna variazione del numero d'onda. Ci sono invece altri semiconduttori, come il germanio o lo stesso silicio (*semiconduttori indiretti*), nei quali la transizione radiativa è molto meno probabile di quella non radiativa ed è per questo che essi vengono usati essenzialmente come materiali rivelatori di radiazione incidente (quindi usati come **fotorivelatori**) in un limitato intervallo di lunghezze d'onda.

Il numero di fotoni emessi dal LED per unità di tempo è proporzionale al numero di elettroni iniettati, per unità di tempo, nella banda di conduzione del diodo; di conseguenza, l'intensità della luce in uscita risulta proporzionale alla corrente in ingresso, fino ad un livello limite di saturazione, che si verifica quando il numero di elettroni iniettati nella banda di conduzione nell'unità di tempo supera la velocità di transizione degli elettroni dalla banda di conduzione a quella di valenza.

La larghezza dell'emissione spettrale di un tipico LED è piuttosto elevata, intorno ai 40 nm: il motivo è che il rilascio di fotoni in uscita avviene in modo disordinato (si parla di **emissione incoerente**).

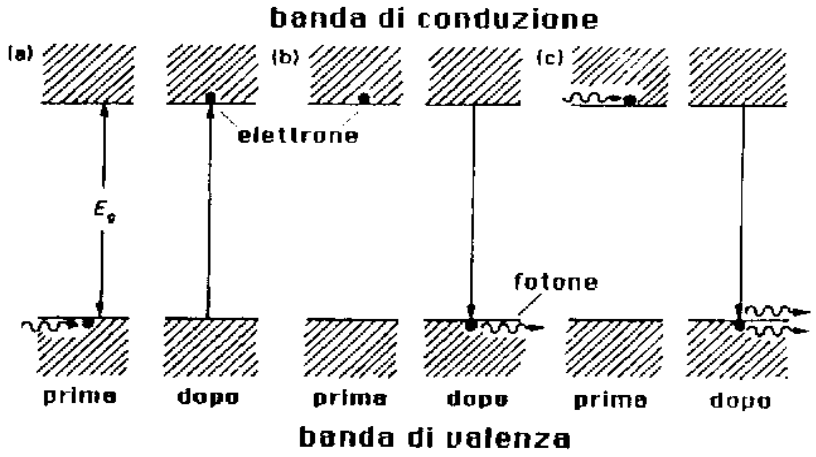
Il diodo LASER

Il **diodo LASER** fu ideato allo scopo di migliorare le prestazioni dei LED e rappresenta un emettitore di luce coerente realizzato in materiale semiconduttore. Il termine *LASER* significa Light Amplification by Stimulated Emissione of Radiation, ossia amplificazione della luce per mezzo di emissione stimolata della radiazione.

Tutti i semiconduttori impiegati nei LASER sono *semiconduttori diretti*: questo è necessario in quanto in questi materiali, proprio per il fatto che la transizione da una banda all'altra non richiede una variazione del numero d'onda $\vec{K} = (2\pi/\lambda)\vec{i}$, le transizioni radiative sono di gran lunga più probabili di quelle non radiative.

Le lunghezze d'onda tipiche dell'emissione LASER coprono la banda compresa tra 0.3 μm e oltre 30 μm.

Il processo di emissione stimolata fu studiato per la prima volta da Einstein. Nella figura seguente sono indicate le 3 possibili situazioni che si verificano nella transizione di un elettrone tra la banda di valenza e la banda di conduzione di un semiconduttore:



Nella prima situazione, un fotone collide con un atomo del materiale: il fotone può essere assorbito, fornendo la sua energia ad un elettrone, che può quindi passare nella banda di conduzione, se naturalmente l'energia assorbita è uguale o superiore all'intervallo di energia proibita del materiale.

Se l'elettrone si ricombina con un altro atomo (seconda situazione), ritorna nella banda di valenza e si ha perciò un rilascio spontaneo di energia, cioè l'*emissione spontanea* di un fotone avente energia pari a quella persa dall'elettrone.

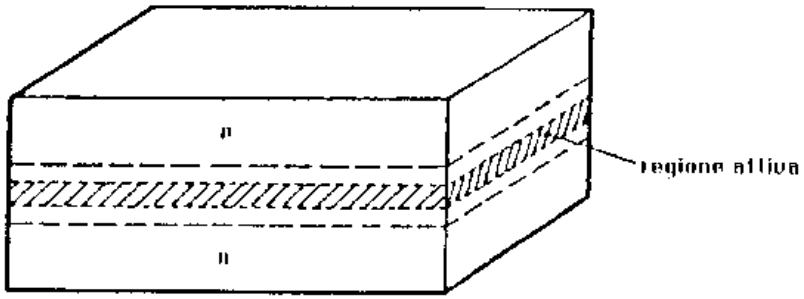
Infine, se un fotone collide con un elettrone situato già in banda di conduzione, si ha proprio il fenomeno della **emissione stimolata**, in cui due distinti fotoni vengono rilasciati in fase.

E' chiaro che, *se l'emissione stimolata di elettroni deve costituire il fenomeno dominante, è necessario che la densità di elettroni nel livello energetico superiore sia maggiore di quella relativa al livello energetico inferiore.* Questa condizione si chiama **inversione di popolazione**, in quanto sappiamo bene che, in condizioni di equilibrio, si verifica la situazione opposta (gli elettroni tendono sempre a portarsi su livelli energetici inferiori). Quindi, per accrescere l'emissione stimolata necessaria per il funzionamento di un laser a semiconduttore, bisogna realizzare l'inversione di popolazione.

Per capire come ottenere questo effetto, consideriamo il funzionamento di una giunzione pn formata da semiconduttori degeneri (cioè notevolmente drogati, con concentrazione di drogante dell'ordine di 10^{19} e anche più). Applicando una polarizzazione diretta alla giunzione, gli elettroni e le lacune vengono iniettati, attraverso la giunzione, in numero tale che la velocità dell'emissione stimolata sia sempre maggiore della velocità con cui i fotoni emessi vengono riassorbiti da tutte le perdite presenti nel dispositivo; in tal modo, si ottiene una azione netta di emissione di radiazione LASER all'uscita: ciò può essere ottenuto applicando una polarizzazione diretta di intensità sufficiente da provocare una intensa iniezione di portatori, cioè forti concentrazioni di lacune ed elettroni nella RCS. In altre parole, si viene a creare una regione contenente una elevata concentrazione di elettroni nella banda di conduzione ed una elevata concentrazione di lacune in banda di valenza, provocando appunto l'inversione di popolazione.

L'azione della corrente di pilotaggio applicata dall'esterno al laser è proprio quella di determinazione l'inversione di popolazione. Un valore di corrente diretta sufficiente a questo scopo è di circa 25000 A/cm^2.

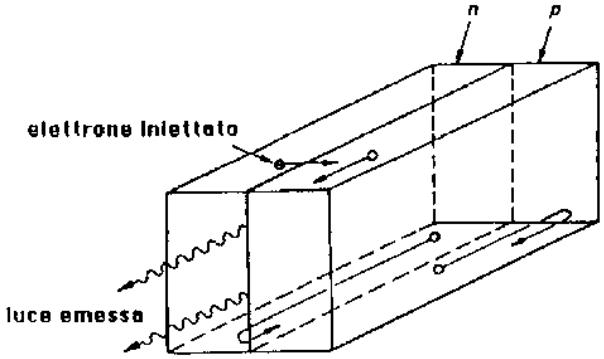
La regione in cui si verifica l'inversione di popolazione è una stretta striscia intorno all'interfaccia geometrica della giunzione e prende il nome di **regione attiva**:



L'addensamento di elettroni nella regione attiva, dovuto appunto all'inversione di popolazione, è accompagnato da un ulteriore pregio e cioè da un aumento dell'indice di rifrazione rispetto alle zone circostanti: questo determina un comodo effetto guidante, comunque debole, sulla radiazione luminosa emessa.

In definitiva, *per avere emissione di luce coerente occorre che si verifichino due distinte condizioni: l'inversione della popolazione e la continua presenza di un **fotoni** (cioè un campo ottico) che collidano con gli elettroni nella regione attiva.*

In base a queste esigenze, il più semplice diodo LASER è offerto dalla giunzione pn realizzare su GaAs, illustrata nella figura seguente:



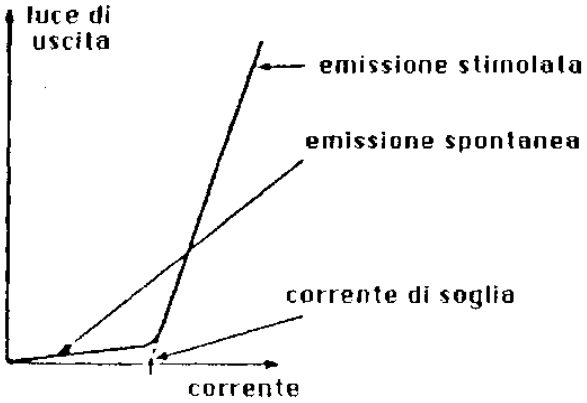
Dato che in entrambi i lati della giunzione è presente lo stesso materiale, si parla di **LASER ad omogiunzione**.

Si fabbrica la giunzione in modo da avere 2 piani paralleli, lisci e perpendicolari alla giunzione stessa; uno dei due piani si comporta da superficie riflettente, mentre l'altro da superficie semi-riflettente, in modo da consentire l'emissione della radiazione LASER: parte della luce emessa nella regione della giunzione viene riflessa avanti e dietro, formando così un'onda stazionaria nella regione attiva. Tale struttura è chiamata **cavità di Fabry-Perot**. Vediamo come funziona.

Quando la giunzione è polarizzata direttamente, gli elettroni di conduzione attraversano lo strato n e vanno nello strato p (ovviamente sempre nella banda di conduzione); entro breve tempo, si ricombinano con le lacune situate nella banda di valenza della zona di tipo p, rilasciando così fotoni con energia pari ad E_g. Dato che i fotoni si riflettono avanti e indietro tra le due facce estreme del cristallo, la probabilità che un fotone collida con un elettrone in banda di conduzione risulta aumentata: si ottiene così il già descritto fenomeno della emissione stimolata, che comporta il rilascio di due fotoni in fase tra loro.

Questo spiega la grande coerenza della radiazione emessa da un diodo LASER, con una larghezza spettrale di circa 1 nm, che rappresenta un miglioramento decisivo, di circa due ordini di grandezza, rispetto all'emissione di un LED.

Osserviamo, infine, che, studiando la variazione dell'intensità di luce in uscita dal LASER ad omogiunzione in funzione della corrente di pilotaggio, si ottiene un diagramma del tipo seguente:



La cosa più interessante da osservare, in questo diagramma, è la presenza di un valore di soglia della corrente, al di sotto del quale le perdite nel mezzo attivo prevalgono sul guadagno e l'emissione è solo spontanea. Al di sopra del valore di soglia, invece, prevale l'emissione stimolata.

Richiami storici: laser e maser

Lo sviluppo del **laser** (1960) fu reso possibile dalla precedente realizzazione del cosiddetto **maser** (1954), il cui nome è l'acronimo di *Microwave Amplification by Stimulated Emission of Radiation*, ossia amplificazione di microonde mediante emissione stimolata della radiazione. Il maser lavora dunque nella regione delle microonde anziché in quella delle radiazioni luminose (per questo i maser si utilizzano in genere nei radiotelescopi per amplificare segnali molto deboli come quelli stellari oppure nelle antenna per ricevere i segnali dai satelliti).

SCHEMA GENERALE DI UN SISTEMA DI TRASMISSIONE SU FIBRA OTTICA

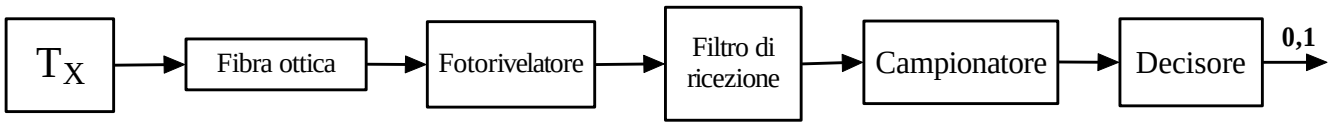
Abbiamo già mostrato lo schema generale di un sistema di trasmissione (numerica) in fibra ottica. Vediamo adesso di arrivarci per gradi.

In primo luogo, abbiamo detto che le apparecchiature trasmittenti sono dei generatori di luce (diodi LED o diodi LASER) che funzionano generalmente con una modulazione del tipo OOK, ossia trasmettono una certa potenza ottica P_T quando deve essere trasmesso un 1 e niente (o quasi niente, per i motivi che vedremo) quando deve essere trasmesso uno 0. Sia nei LED sia nei LASER, l'intensità (cioè la potenza) della luce emessa viene controllata semplicemente mediante la corrente di eccitazione.

L'eventuale potenza ottica trasmessa viene iniettata nella fibra e si propaga lungo essa, giungendo, inevitabilmente attenuata, al terminale ricevente. Qui è necessario disporre di un dispositivo che sia in grado di rivelare la potenza ottica in arrivo e di trasformarla in un segnale elettrico: questo dispositivo sarà dunque un **fotorivelatore** e ne vedremo le caratteristiche.

All'uscita del fotorivelatore abbiamo dunque un segnale elettrico che può essere trattato con il normale sistema filtro-campionatore-decisore ampiamente visto in precedenza per i sistemi numerici.

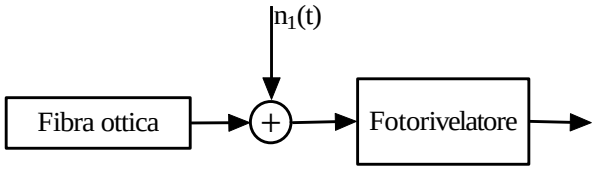
In definitiva, quindi, lo schema di principio è del tipo seguente:



Possiamo subito osservare che il fotorivelatore svolge una funzione analoga a quella del demodulatore nei sistemi a microonde, con due differenze fondamentali: la prima è, ovviamente, nella frequenza di lavoro (frequenza ottica nei sistemi a fibra e frequenza minore nei sistemi a microonde); la seconda, che più ci interessa, è che, ***mentre il demodulatore fornisce una risposta proporzionale all'ampiezza del campo elettrico in ingresso, il fotorivelatore fornisce in uscita una risposta proporzionale alla potenza del campo incidente.***

Tornando adesso allo schema, è evidente che esso manca di una cosa fondamentale, ossia il rumore. Dobbiamo allora capire che tipo di rumore è presente in un apparato del genere e dove esso è presente.

Il primo contributo di rumore è senz'altro il **rumore termico** $n_1(t)$ generato dal mezzo trasmissivo. Dobbiamo dunque aggiungere il solito rumore additivo a valle della fibra:



Tuttavia, questo è un rumore termico sovrapposto ad una *portante ottica* (cioè ad una portante sinusoidale a frequenza ottica, dell'ordine di 10^{14} Hz), per cui richiede delle precisazioni: la fibra, ricoperta dal suo strato protettivo, è costituita da materiale a temperatura ambiente T_0, per cui la temperatura equivalente di rumore alla sua uscita si può assumere pari a T_0. Tuttavia, per un sistema che lavora a frequenze dell'ordine di 10^{14} Hz, non possiamo più utilizzare, per la densità spettrale di potenza del rumore termico, l'espressione kT_0 sempre usata in precedenza: il motivo è che, ***alle frequenze ottiche, nello stimare la suddetta densità spettrale bisogna necessariamente tenere conto dei fenomeni quantistici.*** Allora, l'espressione corretta della densità spettrale di potenza del rumore termico è la seguente:

$$h_n(f) = \frac{hf}{e^{\frac{hf}{kT}} - 1}$$

dove h è la *costante di Planck*. Da questa espressione si giunge all'espressione classica $h_n=kT$ solo se il termine esponenziale a denominatore è molto piccolo, ossia se $hf<<kT$, ossia quindi se $f<10^{12}$ Hz. Per frequenze dell'ordine di 10^{14} Hz, invece, quella espressione non può essere semplificata. Tuttavia, facendo qualche semplice calcolo, si trova quanto segue: a temperatura ambiente $T=T_0$, mentre la densità spettrale di potenza di rumore termico di bassa frequenza kT fornisce una potenza di rumore termico per Hz di -174 dBm, alle frequenze ottiche tale valore scende a **-233dBm**. Questo è un valore sicuramente trascurabile ai fini del dimensionamento di un sistema di trasmissione, per cui possiamo affermare che, ***in un sistema su fibra ottica, è lecito trascurare il rumore termico in uscita dalla fibra stessa.***

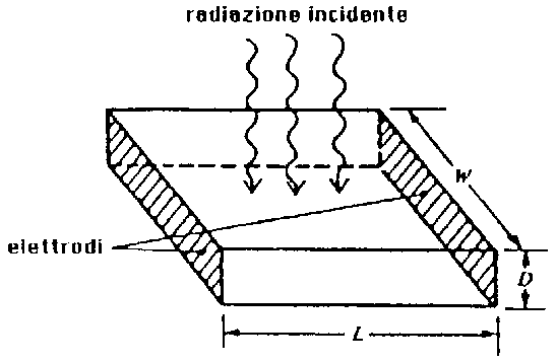
Fino, quindi, all'uscita del fotorivelatore, non abbiamo alcuna sorgente di rumore. All'uscita del fotorivelatore, invece, il segnale viene riportato, come vedremo tra un attimo, a frequenze molto più basse, per cui eventuali sorgenti di rumore termico andranno incluse e trattate nel modo a noi familiare, ossia considerando $h_n=kT$.

FUNZIONAMENTO DEL FOTORIVELATORE

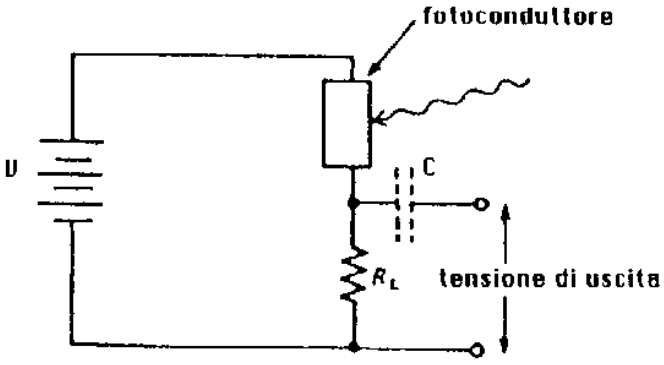
Andiamo adesso a vedere come funziona il fotorivelatore.

Cenni ai dispositivi fotoconduttori

Intanto, un **dispositivo fotoconduttore** consiste semplicemente in uno strato di semiconduttore dotato di contatti ohmici ad entrambe le estremità:



Quando la luce incidente colpisce la superficie del fotoconduttore, coppie elettrone-lacuna sono generate sia attraverso transizioni banda-banda, sia per transizioni che coinvolgono livelli energetici situati nella banda proibita. Queste nuove coppie elettrone-lacune producono evidentemente un aumento della conduttività.
Applicando allora una tensione tra i contatti ohmici, fluisce una corrente tanto maggiore quanto maggiore è la luce incidente: ad ogni modo, l'aumento di corrente ottenuto in presenza di luce incidente è molto piccolo (prende il nome di **fotocorrente**) e, per rivelarlo, è necessario porre il fotoconduttore in serie ad un circuito comprendente una tensione di alimentazione ed una resistenza di carico R_L:



Così facendo, variazioni della resistenza del fotoconduttore dovute alla radiazione luminosa incidente, si manifestano come variazioni della tensione ai capi del resistore, che quindi può essere misurata. Se, poi, interessa solo la componente alternata di questa tensione, cioè appunto quella derivante dalla luce incidente, allora si può pensare di porre un condensatore di blocco, come indicato nella figura precedente.
La fotocorrente generata dipende ovviamente dall'efficienza con cui il materiale converte l'energia luminosa incidente in energia elettrica: si definisce allora una **efficienza quantica** , pari al numero di coppie elettrone-lacuna generate da ciascun fotone incidente. E' anche importante l'assorbimento che la radiazione subisce all'interno del materiale.

I rivelatori di luce utilizzati nei sistemi di comunicazione su fibra ottica sono generalmente dei **fotodiodi**, che sfruttano la capacità ionizzante della radiazione che si propaga lungo la fibra: tale radiazione, infatti, contribuisce a formare coppie elettrone-lacuna nella zona di svuotamento di un diodo a giunzione polarizzato. Nel momento in cui si generano l'elettrone e la lacuna, il campo elettrico interno li separa immediatamente e, se generati nei pressi del limite della zona di svuotamento, possono uscire, alimentando così una tensione esterna a circuito aperto oppure una corrente in un circuito chiuso.

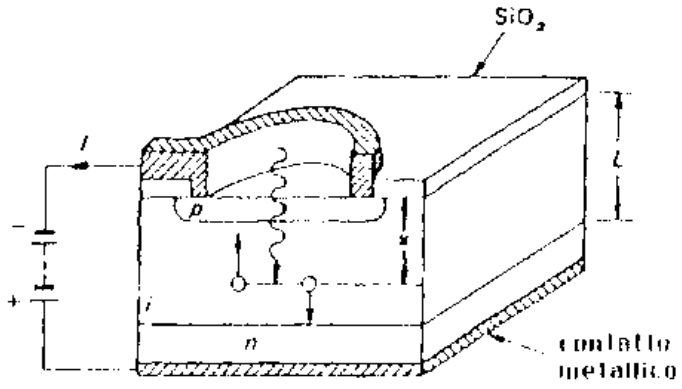
Questo fenomeno può essere maggiormente stimolato nel caso in cui la giunzione sia polarizzata inversamente: in tal caso, le coppie elettrone-lacuna, prodotte dall'assorbimento dei fotoni, vengono rapidamente allontanati dal forte campo elettrico, dando luogo ad una fotocorrente stazionaria.

E' chiaro che, *affinché un fotone possa generare una coppia elettrone-lacuna, è necessario che possieda una energia maggiore o al più uguale a quella necessaria a ionizzare il singolo atomo.* E' evidente che non tutti i fotoni hanno questa energia, per cui non tutti i fotoni ionizzano gli atomi: si è trovato che la probabilità che il fotone in arrivo crei una coppia elettrone-lacuna dipende dal volume a disposizione della radiazione per trovare atomi disponibili alla ionizzazione. Sulla scorta di ciò, i fotorivelatori sono generalmente dei **diodi PIN**, in cui è presente una regione intrinseca tra la zona p e la zona n: così facendo, infatti, *aumenta la zona di svuotamento ed aumenta quindi l'efficienza, ossia il numero di fotoni, tra quelli che arrivano, che creano portatori di carica (cioè appunto coppie elettrone-lacuna).*

Il problema è che, a fronte di questa aumentata efficienza, corrisponde una maggiore lentezza nella risposta del dispositivo al singolo fotone: infatti, la coppia elettrone-lacuna impiegherà più tempo ad attraversare la zona di svuotamento, per cui la banda del dispositivo risulterà ridotta.

Quindi, il requisito di aumentare l'efficienza tramite un allargamento della regione di svuotamento è in contrasto con il requisito di far lavorare il dispositivo ad alta frequenza, che si otterrebbe restringendo la zona di svuotamento: i diodi PIN sono appunto una soluzione di compromesso tra queste esigenze contrastanti.

Nella figura seguente è mostrata la tipica struttura di un diodo PIN:



Lo strato di tipo p è molto sottile, in modo che solo un numero trascurabile di fotoni sia assorbito in questa regione. La regione di svuotamento è invece spessa abbastanza da catturare la maggior parte della luce incidente, ma non tanto da causare un aumento eccessivo del tempo di attraversamento da parte dei portatori e quindi una riduzione sensibile della larghezza di banda.

Ad ogni modo, l'uscita del fotorivelatore è un impulso di corrente inviato nel circuito esterno. L'area dell'impulso è pari alla carica del portatore[1].

Cominciamo ad analizzare la cosa da un punto di vista matematico. Indichiamo con P(t) la potenza (generalmente funzione del tempo) della radiazione incidente; ricordiamo inoltre che l'energia che ciascun fotone "porta con sé" è data da

$$E_{\text{fotone}} = hf = h\frac{c}{\lambda}$$

Allora, il numero medio di fotoni ricevuto nell'unità di tempo è

$$N(t) = \frac{P(t)}{E_{\text{fotone}}} = \frac{P(t)}{hf}$$

Se ciascun fotone in arrivo generasse un portatore di carica (cioè una coppia elettrone-lacuna), allora N(t) sarebbe anche il numero medio di portatori generati nell'unità di tempo dal fotorivelatore; al contrario, solo una parte della potenza incidente penetra nella regione arriva del fotodiodo e, inoltre, come detto prima, non tutti i fotoni generano un portatore: di conseguenza, dobbiamo tener conto di una **efficienza quantica** (simbolo η), in base alla quale scrivere che il numero medio di portatori generati nell'unità di tempo dal fotorivelatore è

$$\chi(t) = \eta N(t) = \eta\frac{P(t)}{hf}$$

Nel caso ideale, sarebbe $\eta=1$; nella realtà, invece, un valore tipico è $\eta=0.8$.

Noto $\chi(t)$, possiamo valutare il valore medio statistico della corrente prodotta dal fotorivelatore:

$$I(t) = q\chi(t) = \frac{\eta q}{hf}P(t)$$

Generalmente, il fattore $\frac{\eta q}{hf}$ viene indicato con ρ e prende il nome di **responsività** del fotodiodo: possiamo dunque concludere che

$$\boxed{I(t) = \rho P(t)}$$

Come già preannunciato in precedenza, *l'uscita del fotorivelatore è dunque proporzionale alla potenza ottica in ingresso: il coefficienza di proporzionalità è r.*

A questo punto, facciamo il discorso seguente: intanto, supponiamo che la trasmissione avvenga inviando luce (durante un intervallo T e con una certa potenza P_T) quando c'è da trasmettere un 1 e niente quando c'è da trasmettere uno 0; con questa scelta, per capire, in ricezione, quale sia il simbolo trasmesso, basta capire se è arrivato almeno un fotone: se si, allora il simbolo trasmesso è 1, altrimenti è 0. Di conseguenza, se ci poniamo nel caso ideale in cui il fotorivelatore non emette alcuna corrente in assenza di luce incidente, è evidente che non ci sarebbe possibilità di sbagliare. Al contrario, invece, si può sbagliare, fondamentalmente per due motivi:

[1] Ricordiamo infatti che I=q/t.

- il primo è che il fotorivelatore non sarà mai ideale, il che significa che esso genera una certa corrente (detta **corrente di buio**) anche quando non c'è alcuna luce incidente;
- il secondo è invece relativo all'attenuazione della fibra: *a causa di tale attenuazione, non tutti i fotoni trasmessi giungono in ricezione, per cui il fotodiodo sarà comunque eccitato da un numero di fotoni minore del numero di fotoni trasmessi;* allora, se i fotoni emessi in trasmissione sono particolarmente pochi, è possibile che nessuno di essi giunga in ricezione, per cui il fotodiodo (ammesso che sia ideale) non fornirebbe alcuna corrente, facendo optare il decisore per uno 0 quando invece era stato trasmesso un 1.

Cerchiamo allora di capire, da un punto di vista statistico, quanti fotoni giungono al fotorivelatore.

Intanto, questi fotoni arrivano in modo del tutto indipendente e in modo del tutto casuale. *Si può allora stimare che il processo "fotoni in arrivo in seguito alla trasmissione di potenza P_T" sia un* ***processo di Poisson***. Questo significa che la probabilità che nel periodo T arrivino k fotoni è

$$P(k) = \frac{(\chi T)^k}{k!} e^{-\chi T}$$

dove χ è il numero medio di fotoni che arrivano nell'unità di tempo. In particolare, si è considerato χ indipendente dal tempo (quando invece prima avevamo posto $\chi(t)$) in quanto, supponendo di trasmettere impulsi rettangolari (con potenza costante durante ciascun impulso) per rappresentare il simbolo 1, χ è costante durante l'impulso. Si è inoltre considerata, per semplicità, una efficienza quantica η unitaria.

Osserviamo inoltre che la quantità χT rappresenta il numero medio di fotoni che giungono in un periodo di cifra T corrispondente, ovviamente, alla trasmissione di 1 (nel senso che, in un periodo di cifra corrispondente alla trasmissione di uno 0, non arriverebbe alcun fotone).

Possiamo adesso stimare la probabilità media di errore, che va calcolata con la solita formula generale

$$p(\varepsilon) = P(\varepsilon \mid 0T)P(0T) + P(\varepsilon \mid 1T)P(1T)$$

In primo luogo, sempre nell'ipotesi di trascurare la corrente di buio del fotorivelatore, è evidente che non c'è possibilità di errore quando è stato trasmesso uno 0, in quanto non arriva comunque alcun fotone in ricezione: quindi $P(\varepsilon|0T)=0$ e $p(\varepsilon) = P(\varepsilon \mid 1T)P(1T)$.

In secondo luogo, la probabilità di sbagliare quando è stato trasmesso un 1 corrisponde alla probabilità che il fotorivelatore non riceva in ingresso alcuna luce incidente, ossia alcun fotone: si tratta quindi di P(k) con k=0, per cui

$$P(\varepsilon \mid 1T) = P(k=0) = e^{-\chi T} \longrightarrow p(\varepsilon) = e^{-\chi T} P(1T)$$

Supponendo infine che i simboli 0 ed 1 siano equiprobabili, è ovvio che $P(1T)=P(0T)=0.5$, per cui concludiamo che la probabilità di errore, in un sistema su fibra ottica, è

$$\boxed{p(\varepsilon) = 0.5e^{-\chi T}}$$

Questa formula ci consente di valutare il numero medio di fotoni, per periodo di cifra, che è necessario ottenere in ricezione per ottenere una prefissata probabilità di errore: per esempio, considerando una probabilità di errore $p(\varepsilon)=10^{-9}$, si trova che

$$N_{min} = \chi T = -\log_e \frac{p(\varepsilon)}{0.5} = \log_e \frac{0.5}{p(\varepsilon)} \cong 20$$

In conclusione, se vogliamo sbagliare non più di 1 bit per ogni miliardo, abbiamo bisogno, in ricezione, di almeno 20 fotoni per periodo di cifra, relativamente alla trasmissione di 1.

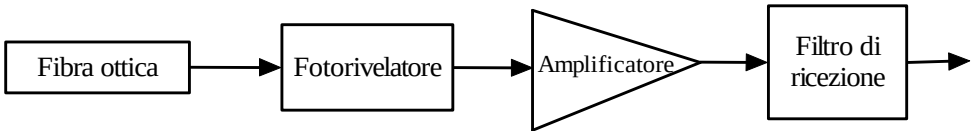
Se i simboli sono equiprobabili, come già supposto, il *numero medio di fotoni per simbolo* è

$$N = N_{min} \cdot P(1T) = N_{min} \cdot P(0T) = N_{min} \cdot 0.5 = 10$$

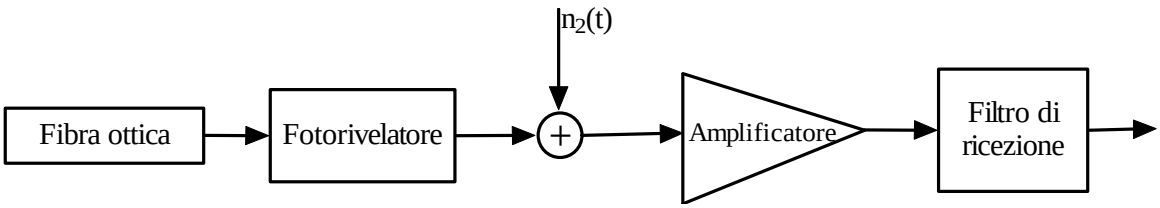
DIMENSIONAMENTO DI UN SISTEMA REALE

Le considerazioni del paragrafo precedente valgono, dunque, nel caso ideale in cui il fotorivelatore non abbia corrente di buio e tutti i fotoni sia efficaci (h=1) ai fini della generazione di coppie elettrone-lacuna. Al contrario, sappiamo che $\eta<1$ e sappiamo anche che la corrente di buio è inevitabile (in quanto, a causa della agitazione termica, casualmente qualche atomo si ionizza da solo, generando appunto una piccola corrente).

A queste considerazioni, dobbiamo aggiungerne anche un'altra molto importante: la corrente I(t) generata dal fotorivelatore a seguito della ionizzazione dei suoi atomi è comunque una corrente molto bassa, che deve essere necessariamente amplificata prima di arrivare in ingresso al filtro di ricezione:



Questa amplificazione introduce allora nuovamente un rumore termico, che in questo caso dobbiamo considerare per il semplice motivo che il segnale, in questo punto della catena, non è più a frequenze ottiche, per cui ha una propria rilevanza:



Sulla base di tutte queste considerazioni, andiamo allora a dimensionare un sistema reale.

Intanto, abbiamo prima visto che, considerando l'efficienza quantica η, il valore medio statistico della corrente prodotta dal rivelatore è

$$I(t) = q\chi(t) = \frac{\eta q}{hf} P(t) = \rho P(t)$$

L'impulsino di corrente I(t) è generato in modo casuale, il che significa che l'uscita del rivelatore non è un segnale deterministico, con del rumore sovrapposto, ma un processo stocastico con

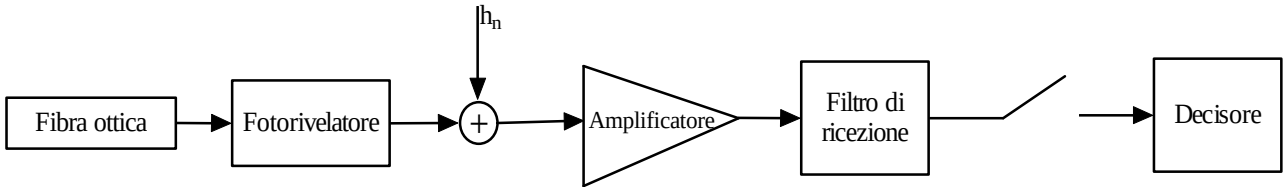
statistica di Poisson. Questo processo di Poisson va in ingresso al filtro di ricezione e non siamo in grado di dire cosa venga fuori dal filtro: infatti, solo se la statistica in ingresso fosse gaussiana, saremmo autorizzati a dire che anche la statistica in uscita sarebbe gaussiana. In tutti gli altri casi, invece, per calcolare la statistica del primo ordine all'uscita del filtro dovremmo necessariamente conoscere la statistica, di tutti gli ordini, del processo in ingresso. Questo sembrerebbe quindi un ostacolo insormontabile per il dimensionamento.

Al contrario, *possiamo procedere in modo leggermente forzato, dimostrando che, di fatto, il processo in uscita dal filtro è ancora una volta gaussiano*.

Abbiamo ormai capito che, in ricezione, è necessario comunque più di un fotone per optare per un 1 trasmesso. Ogni impulsino di corrente prodotto dal rivelatore viene amplificato e poi filtrato dal filtro. All'uscita da tale filtro, esso avrà quindi una banda che è circa pari alla frequenza di cifra $f_S=1/T$, a prescindere da quanto fosse all'ingresso. Se nel periodo di cifra T abbiamo più di un fotone (generalmente ne otteniamo da 100 a 1000), possiamo affermare che il processo di Poisson, che era rappresentato da impulsini che arrivavano casualmente nel tempo in numero χ per secondo, una volta filtrato è tale che al posto di ogni impulsino si abbiamo le risposte all'impulso del filtro stesso, che vanno sommate tra loro. All'uscita del filtro, quindi, si ha la combinazione lineare, con pesi più o meno paragonabili, di un gran numero di impulsini. Se la distanza temporale media tra due impulsi è molto piccola rispetto al periodo di cifra (ossia se $\mathbf{1/\chi<<T}$), allora praticamente si sovrappongono molte risposte all'impulso. L'uscita del filtro viene poi campionata: l'uscita del campionatore, cioè appunto la variabile casuale ottenuta campionando il processo di Poisson filtrato) è, dunque , istante per istante, la combinazione lineare di un gran numero di variabili, indipendenti tra loro e pesate in modo che nessuna sia dominante rispetto all'altra. Invocando, allora, il *teorema del limite centrale*, possiamo affermare che tale processo sia quasi gaussiano, per cui lo consideriamo come tale.

In base a queste considerazioni, assumiamo dunque che il processo in uscita dal filtro sia gaussiano: dovremo in seguito verificare che questa assunzione sia congruente con i risultati che ci accingiamo ad ottenere, ossia quindi dovremo verificare che tali risultati legittimino la relazione $1/\chi<<T$.

Riprendiamo dunque lo schema dell'apparato ricevitore:



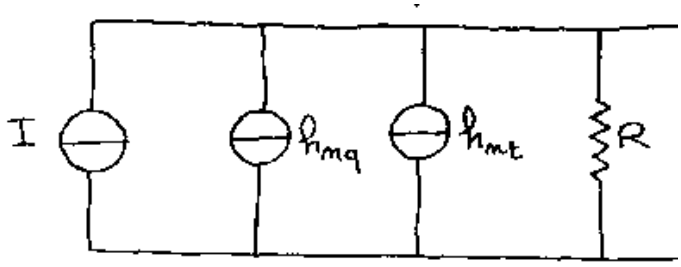
Il fotorivelatore può essere schematizzato come un generatore di corrente: in tal modo, il segnale (cioè la parte "desiderata" del processo all'uscita del rivelatore) è il valor medio della corrente in uscita dal rivelatore stesso: avendo trovato prima che il valore statistico medio della corrente è $I(t)=\rho P(t)$, il valor medio, se P_R è la potenza media in ingresso al rivelatore, sarà $\mathbf{I=\rho P_R}$.

Sovrapposto a questo segnale, c'è un disturbo, che rappresenta le fluttuazioni attorno al valor medio I. A cosa corrisponde questo disturbo? Ci sono varie cause di rumore da portare in conto:

- in primo luogo, un termine di **rumore quantico** dovuto all'efficienza quantica η, ossia al fatto che non tutti i fotoni in uscita dalla fibra riescono a ionizzare gli atomi del diodo: avendo deciso di schematizzare il fotorivelatore come generatore di corrente, ci conviene modellare anche il rumore quantico come un generatore di corrente i_q, con densità spettrale (da calcolare) che indichiamo con h_{nq};

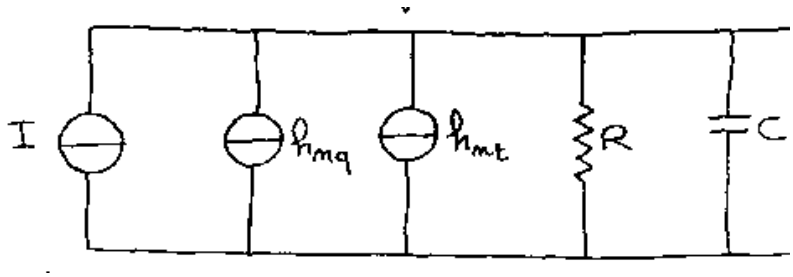
- in secondo luogo, dovremmo anche considerare la *corrente di buio*, il cui valor medio andrebbe sommato ad I e la cui variazione andrebbe sommata a i_q: possiamo però trascurare, con buona approssimazione, questo contributo;
- in terzo luogo, dobbiamo tener conto che, per polarizzare inversamente il diodo rivelatore, abbiamo bisogno di un circuito di polarizzazione, che inevitabilmente conterrà una resistenza[1], cioè quindi un generatore di **rumore termico**: se R è il valore di questa resistenza, potremo modellarne il contributo di rumore termico mediante un generatore di corrente in con densità spettrale $4kT_0 / R$;
- ancora, dobbiamo considerare l'amplificatore posto a monte del filtro: tale amplificatore presenta una propria resistenza interna e quindi una rumorosità (sempre rumore termico) che possiamo riportare in ingresso, adoperando ancora una volta il concetto di fattore di rumore; d'altra parte, possiamo inglobare questo rumore (che è quello generato dalla parte resistiva dell'impedenza di ingresso dell'amplificatore) nel rumore prodotto dalla resistenza di polarizzazione, considerando una densità spettrale pari a $h_{nt} = 4kFT_0 / R$.

Così facendo, siamo dunque pervenuti ad un circuito, che rappresenta appunto il rivelatore e la rumorosità dell'amplificatore, fatto nel modo seguente:



Questo schema può essere ulteriormente perfezionato tenendo conto che un diodo polarizzato inversamente presenta un comportamento capacitivo (legato alla dipendenza, dalla tensione applicata, della carica presente nella regione di svuotamento). Possiamo allora modellare questo comportamento mediante una capacità C in cui possiamo anche inglobare la parte reattiva dell'impedenza di ingresso dell'amplificatore.

Lo schema cui fare riferimento è dunque il seguente:



[1] Questa resistenza non va assolutamente realizzata in granuli di carbonio, ma con tecnologia tale da non far comparire il flicker noise (cioè il rumore a bassa frequenza): all'uscita del rivelatore, infatti, ciò che interessa sono le basse frequenza.

L'unico termine che ancora non abbiamo valutato quantitativamente è la densità spettrale del rumore quantico: ci ricordiamo allora che tale densità spettrale (monolatera) ha espressione $h_{nq} = 2qI$.

Fatte tutte queste premesse, possiamo procedere al dimensionamento, che per il blocco filtro-campionatore-decisore va condotto così come in tutti i casi considerati in precedenza.

Il punto di partenza è la fissata p(ε) a valle del decisore. Per passare a monte del decisore, ci basta considerare la relazione Q(γ)=p(ε), dove $\gamma = \frac{V_1 - V_0}{2\sigma_n}$, avendo indicato con V_1 e V_0 i livelli che si misurerebbero, in uscita dal campionatore, in assenza di rumore (caratterizzato da una deviazione standard σ_n, uguale sia all'uscita sia all'ingresso del campionatore).

Se fosse $p(\varepsilon)=10^{-6}$, sappiamo che si otterrebbe $(\gamma)_{dB} = 20\log_{10}\gamma = 13.5$ dB. Dato che, invece, vogliamo $p(\varepsilon)=10^{-9}$, dobbiamo considerare $(\gamma)_{dB} = 16.5$ dB.

Possiamo adesso passare al rapporto segnale/rumore $\left.\frac{S}{N}\right|_U$ a valle del filtro di ricezione, definito come rapporto tra la potenza di picco $P_{S,p}$ del segnale e la potenza media $P_{N,m}$ di rumore: dato che abbiamo scelto una *codifica di tipo ortogonale*, sappiamo che γ^2 è numericamente uguale ad un quarto del rapporto S/N a valle del filtro, per cui

$$\frac{1}{4}\left.\frac{S}{N}\right|_U = \frac{1}{4}\frac{P_{S,p}}{P_{N,m}} = \gamma^2 \longrightarrow \left.\frac{S}{N}\right|_U = \frac{P_{S,p}}{P_{N,m}} = 4\gamma^2$$

Per passare all'ingresso del filtro di ricezione, dobbiamo scegliere il tipo di filtro. Se scegliamo il *filtro adattato*, sappiamo che il rapporto S/N all'uscita del filtro è numericamente pari al rapporto S/N in ingresso, a patto di definire quest'ultimo come rapporto tra la potenza media di segnale e la potenza di rumore che cade nella *banda convenzionale* $f_S/2$:

$$\left.\frac{S}{N}\right|_U = \left.\frac{S}{N}\right|_{IN} = \frac{P_R}{h_n \frac{f_S}{2}}$$

dove h_n è la densità spettrale di potenza totale del rumore sovrapposto al segnale, mentre P_R è la potenza media (di segnale) ricevuta.

In base a quanto detto prima, questo rapporto è numericamente pari a $4\gamma^2$; se però consideriamo il disadattamento del filtro, adottando un fattore di forma di 0.5dB, concludiamo che

$$\left.\frac{S}{N}\right|_U = \left.\frac{S}{N}\right|_{IN} = \frac{P_R}{h_n \frac{f_S}{2}} = 4\gamma^2 10^{\frac{0.5}{10}} = 4\gamma^2 10^{0.05}$$

(dove abbiamo espresso il tutto in unità naturali).

Siamo dunque giunti a monte del filtro di ricezione, dove giunge il segnale (eventualmente amplificato) prodotto dal fotorivelatore, con sovrapposto sia il rumore quantico dovuto allo stesso fotorivelatore sia il rumore termico dovuto alla polarizzazione del fotorivelatore e all'impedenza di ingresso dell'amplificatore:

- per quanto riguarda il segnale, avendo detto che il valor medio della corrente in uscita dal fotorivelatore è I, la potenza sarà $P_R=I^2$;

- *per quanto riguarda, invece, il rumore, abbiamo due contributi che, essendo statisticamente indipendenti, si sommano in potenza:* dobbiamo perciò sommare la potenza media di rumore quantico e la potenza media di rumore termico, entrambe valutate nella banda convenzionale $f_S/2$.

Abbiamo dunque che

$$\frac{S}{N}=\frac{P_R}{h_n\frac{f_S}{2}}=\frac{I^2}{\left(2qI+\frac{4kT_0F}{R}\right)\frac{f_S}{2}}=4\gamma^2 10^{0.05}$$

In realtà, questa uguaglianza è un po' forzata: infatti, essa vale se si assume, implicitamente, che la densità spettrale di rumore abbia una statistica di Gauss. Questo, invece, non è vero, in quanto, sommando le due variabili indipendenti (rumore quantico e rumore termico), si sommano anche le statistiche del secondo ordine (varianze), ma ciò non significa che il risultato avrà la statistica di una delle due, cioè gaussiano.

Ad ogni modo, noi continuiamo a ritenere valida l'approssimazione relativa al teorema del limite centrale, per cui accettiamo quella formula. Il nostro scopo è quello di calcolare il numero medio di fotoni da trasmettere per ogni periodo di cifra e di confrontarlo con il valore $N_{min}=20$ trovato in precedenza per il caso ideale.

Intanto, sostituiamo l'uguaglianza di prima con una disuguaglianza, imponendo che S/N sia maggiore o al più uguale al valore $4\gamma^2 10^{0.05}$: quindi

$$\frac{S}{N}=\frac{I^2}{\left(2qI+\frac{4kT_0F}{R}\right)\frac{f_S}{2}}\geq 4\gamma^2 10^{0.05}$$

Ricordando adesso che il periodo di cifra è $T=1/f_S$, possiamo porre $f_S=1/T$ e portare T al numeratore:

$$\frac{S}{N}=\frac{I^2T}{\left(2qI+\frac{4kT_0F}{R}\right)\frac{1}{2}}\geq 4\gamma^2 10^{0.05}$$

In questo modo, I^2T è l'energia della forma d'onda rettangolare all'uscita del fotorivelatore (il quale, lo ricordiamo, produce impulsi rettangolari di corrente di valor medio I).

A questo punto, per applicare quella formula, dobbiamo necessariamente fissare il valore della resistenza di polarizzazione R. E' evidente che, se rendessimo R molto elevata, otterremmo una riduzione della corrente di rumore termico, la cui densità spettrale vale infatti $\frac{4kT_0F}{R}$. Non solo, ma un valore elevato di R è opportuno anche per ottimizzare le prestazioni del diodo fotorivelatore: infatti, se R è grande, risulta maggiore la quota parte di corrente (segnale) che circola nell'amplificatore. Il problema, però, è che tale resistenza R forma, con la capacità C in parallelo, un filtro passa-basso, la cui frequenza di taglio è notoriamente $f_T=\frac{1}{2\pi}\frac{1}{RC}$: allora, fissato il valore di C (che è dell'ordine dei pF, visto che rappresenta la capacità di un diodo polarizzato inversamente), quanto maggiore è R tanto minore è la frequenza di taglio del filtro, il che non permetterebbe di trasmettere ad una frequenza di cifra f_S sufficientemente elevata.

Allora, dobbiamo dimensionare R ponendoci, come scopo principale, quello di poter trasmettere con una velocità di trasmissione pari proprio a f_S. Per ottenere questo, basta fare in modo che la frequenza di taglio dell' RC sia molto prossima al valore f_S: imponiamo perciò che risulti

$$f_T = \frac{1}{2\pi}\frac{1}{RC} = \xi f_S$$

dove il coefficiente ξ deve essere reso circa unitario:

$$R = \frac{1}{2\pi C \xi f_S} = \frac{T}{2\pi C \xi}$$

Sostituendo dunque questa espressione di R nell'espressione del rapporto S/N, otteniamo

$$\frac{S}{N} = \frac{I^2 T}{\left(2qI + 4kT_0 F\left(\frac{2\pi C\xi}{T}\right)\right)\frac{1}{2}} \geq 4\gamma^2 10^{0.05}$$

Dividendo numeratore e denominatore per I, otteniamo

$$\frac{S}{N} = \frac{IT}{\left(2q + 4kT_0 F\left(\frac{2\pi C\xi}{IT}\right)\right)\frac{1}{2}} \geq 4\gamma^2 10^{0.05}$$

Possiamo a questo punto risolvere rispetto a IT:

$$IT \geq 399\left(2q + 8\pi C\xi kT_0 F\frac{1}{IT}\right)\frac{1}{2} \cong 4\gamma^2 10^{0.05}\left(2q + 8\pi C\xi kT_0 F\frac{1}{IT}\right) = 400\left(2q + 8\pi C\xi kT_0 F\frac{1}{IT}\right)$$

Valori numerici tipici, che possiamo utilizzare in questo conto, sono i seguenti: $q = 1.6\cdot 10^{-19}\,C$, $k = 1.38\cdot 10^{-23}\,J/°K$, $T_0 = 293°K$ $\xi = 0.5$, $C = 2pF$, $F = 2$. Con questi valori numerici, si ottiene

$$IT \geq 1.3\cdot 10^{-16} + 8\cdot 10^{-29}\cdot\frac{1}{IT} \quad \longrightarrow \quad I^2T^2 - 1.3\cdot 10^{-16} IT - 8\cdot 10^{-29} \geq 0$$

Abbiamo cioè una equazione di 2° grado nell'incognita IT: risolvendo e scartando la soluzione negativa, si trova $IT = 9\cdot 10^{-15}$. Da qui ricaviamo che il numero medio di fotoni per periodo di cifra è

$$N = \frac{IT}{q} = 56250$$

Questo valore, come ci aspettavamo, è molto maggiore dei 20 fotoni, per periodo di cifra, trovato nel caso ideale.

Tra l'altro, un valore così alto di N convalida l'ipotesi di considerare il rumore quantico come un rumore gaussiano a valle del filtro di ricezione (considerando dunque valido il teorema del limite centrale): infatti, 56250 fotoni in ogni periodo T garantiscono la validità della relazione $1/\chi << T$,

ossia il fatto che la distanza temporale media tra due impulsi sia molto piccola rispetto al periodo di cifra.

RIDUZIONE DEL RUMORE TERMICO

Consideriamo dunque il valore $IT = 9 \cdot 10^{-15}$ trovato nel calcolo precedente: sostituendo questo valore nell'espressione della densità di potenza di rumore termico, otteniamo

$$4kT_0F\left(\frac{2\pi C\xi}{IT}\right) = \frac{2 \cdot 10^{-31}}{IT} = 2.26 \cdot 10^{-17}$$

Questo valore è maggiore rispetto a quello della potenza di rumore quantico, dal che deduciamo, come in effetti accade in generale, che *il rumore termico è dominante rispetto al rumore quantico*.

Dato allora che il rumore quantico non è eliminabile nei sistemi in fibra ottica, perché è intrinseco nel funzionamento del fotorivelatore, possiamo chiederci se e come sia possibile ridurre il rumore termico, legato essenzialmente alla qualità dell'implementazione pratica.

E' possibile ridurre il rumore termico andando a modificare direttamente la tecnologia del fotorivelatore; si può infatti pensare di migliorare l'efficienza del fotorivelatore sfruttando il cosiddetto **effetto valanga** (fotorivelatore ad effetto valanga, **APD**): abbiamo detto che il funzionamento del fotorivelatore PIN consiste, essenzialmente, nella ionizzazione degli atomi presenti nella RCS ad opera dei fotoni che provengono dalla fibra ottica; supponiamo di applicare al diodo una polarizzazione inversa sufficientemente elevata, in modo da avere un campo elettrico elevato (o, ciò che è lo stesso, un elevato gradiente di potenziale) nella regione di svuotamento; consideriamo quindi un elettrone che provenga dalla ionizzazione di un atomo ad opera di un fotone: grazie proprio all'elevato campo elettrico, l'elettrone viene notevolmente accelerato, acquistando una energia cinetica che potrebbe essere sufficiente a ionizzare un ulteriore atomo col quale l'elettrone va a collidere; si ha cioè una doppia ionizzazione, dovuta allo stesso fotone, ed è possibile che si abbiano ulteriori ionizzazioni se gli elettroni prodotti acquistano a loro volta sufficiente energia. *Si innesca così un **meccanismo a valanga**, col quale il singolo fotone non produce più 1 solo portatore di carica, ma **G** portatori di carica, con G>1.* Questo significa anche che vengono generati G impulsi di corrente e quindi anche che il quadrato della corrente media in uscita da un diodo a valanga risulta essere pari a G volte il quadrato della corrente media in uscita da un normale diodo fotorivelatore[1].

Allora, se torniamo all'espressione del rapporto S/N all'uscita del fotorivelatore, il numeratore, cioè la potenza di picco del segnale, non sarà più I^2, ma GI^2:

$$\frac{S}{N} = \frac{GI^2T}{\left(2qI + 4kT_0F\left(\frac{2\pi C\xi}{T}\right)\right)\frac{1}{2}} \geq 4\gamma^2 10^{0.05}$$

In realtà, però, questa formula non è completa, in quanto il meccanismo di funzionamento appena descritto ha anche influenza sul rumore quantico in uscita dal fotorivelatore: *si stima,*

[1] Dato che il fotorivelatore è un dispositivo con caratteristica quadratica (quindi non lineare), G non ha il significato di guadagno di corrente, bensì di guadagno di potenza.

infatti, sperimentalmente che il rumore quantico in uscita aumenta più di quanto aumenti la corrente. Possiamo allora modellare questo aumento mediante un guadagno, relativo alla densità spettrale di potenza del rumore quantico, pari a $\mathbf{G^{2\alpha}}$, dove ovviamente $\alpha>0.5$ proprio perché deve risultare $G^{2\alpha}>G$.

L'espressione da considerare diventa dunque

$$\frac{S}{N}=\frac{GI^2T}{\left(2qG^{2\alpha}I+4kT_0F\left(\frac{2\pi C\xi}{T}\right)\right)\frac{1}{2}}\geq 4\gamma^2 10^{0.05}$$

Se dividiamo numeratore e denominatore della frazione per G, otteniamo

$$\frac{S}{N}=\frac{I^2T}{\left(2qG^{2\alpha-1}I+4kT_0F\left(\frac{2\pi C\xi}{GT}\right)\right)\frac{1}{2}}\geq 4\gamma^2 10^{0.05}$$

Questa formula ci conferisce, per il dimensionamento, un grado di libertà in più rispetto al caso del fotodiodo PIN: infatti, il parametro G (detto **guadagno di valanga**) dipende da come è realizzato il diodo e dall'intensità del campo elettrico che instauriamo all'interno della sua zona di svuotamento, per cui è sotto il nostro controllo. Esso ci permette di stabilire di quanto il rumore quantico deve prevalere rispetto al rumore termico.

Naturalmente, a noi interessa ottenere le stesse prestazioni con il valore minimo possibile di corrente media I in uscita dal fotorivelatore, ossia quindi col minimo di potenza ottica in ricezione (e quindi anche in trasmissione). *Si può allora procedere determinando il valore di G che minimizza i requisiti in termini di potenza ottica in ricezione, per una prefissata probabilità di errore p(e) (e cioè per un prefissato valore di g^2):* in pratica, si risolve l'ultima equazione scritta rispetto alla corrente I, mantenendo G come parametro, e poi si minimizza il valore di I variando opportunamente G.

Così facendo, si trova il valore ottimale di G: valori alti di G corrispondono ad un rumore quantico molto forte, mentre il valore minimo di G, cioè 1, ci riporta al fotodiodo PIN esaminato in precedenza, col rumore termico che limita le prestazioni.

Esiste una situazione ottimale che fissa il punto di funzionamento del sistema in modo che i pesi del rumore termico e del rumore quantico siano confrontabili. In questo caso, il disturbo non è più del tutto gaussiano. Si ottiene così una riduzione del numero di fotoni necessari in trasmissione, anche se siamo ancora su valori molto maggiori di 10 (che è il limite teorico massimo).

SISTEMA ETERODINA E OMODINA

Il sistema di trasmissione, così come è stato descritto nei precedenti paragrafi, non offre altri margini di miglioramento, fin quando la modulazione resta di tipo OOK e, in ricezione, si effettua una rivelazione diretta del segnale. Al contrario, *si possono migliorare le prestazioni usando un sistema di modulazione più efficiente rispetto alla modulazione OOK:* ad esempio, si può pensare ad una modulazione PSK, realizzabile mediante una codifica di linea di tipo antipodale, ossia utilizzando forme d'onda modulanti di tipo rettangolare, con ampiezze positive o negative..

C'è però da fare una osservazione: mentre prima abbiamo effettuato una trasmissione incoerente sulla fibra, la demodulazione di una forma d'onda PSK richiede, in ricezione, una demodulazione coerente. Vediamo allora come si procede.

Cominciamo dalla modulazione, che si effettua mediante un cosiddetto **modulatore elettro-ottico**: esso è costituito da materiale piezoelettrico, le cui caratteristiche variano in funzione del campo elettrico applicato. Controllando, per mezzo proprio del campo elettrico, la costante dielettrica del materiale, è possibile controllare la velocità di propagazione della radiazione. Ciò significa far variare il ritardo di una mezza lunghezza d'onda a seconda che si voglia trasmettere 1 o 0, ottenendo appunto una modulazione PSK.

Questo segnale modulato PSK (quindi un segnale passa-banda centrato sulla frequenza della portante ottica) giunge dunque in ricezione, dove va demodulato. Il fotorivelatore fornisce in uscita una corrente proporzionale alla potenza ottica incidente, cioè una corrente proporzionale al quadrato dell'ampiezza della radiazione incidente. Il rivelatore, dunque, come sappiamo, è di tipo quadratico.

Sia s(t) il segnale modulante che intendiamo trasmettere, cioè la sequenza binaria di 1 e 0 che rappresenta le informazioni: sarà quindi un segnale con due soli possibili valori, del tipo $s(t) = \pm A$. Questo segnale va a modulare in ampiezza[1] una portante sinusoidale $\cos(2\pi f_P t + \varphi_P)$ a frequenza ottica: il segnale trasmesso sulla fibra risulta essere quindi

$$s_t(t) = s(t)\cos(2\pi f_P t + \varphi_P)$$

Questo segnale va in ingresso al fotorivelatore, sul quale perciò incide una potenza ottica

$$P(t) = s_t^2(t) = s^2(t)\cos^2(2\pi f_P t + \varphi_P) = s^2(t)\frac{1+\cos(2\pi(2f_P)t + 2\varphi_P)}{2} = \frac{s^2(t)}{2} + \frac{s^2(t)}{2}\cos(2\pi(2f_P)t + 2\varphi_P)$$

Il valore medio statistico della corrente prodotta dal fotorivelatore è allora pari, in base alla stessa formula ricavata in precedenza, a

$$I(t) = q\chi(t) = \frac{\eta q}{hf}P(t) = \rho P(t) = \rho\frac{s^2(t)}{2} + \rho\frac{s^2(t)}{2}\cos(2\pi(2f_P)t + 2\varphi_P)$$

Se f_P è una frequenza ottica, $2f_P$ sarà una frequenza estremamente elevata, per cui possiamo senz'altro ritenere che la componente $\rho\frac{s^2(t)}{2}\cos(2\pi(2f_P)t + 2\varphi_P)$ dia un contributo assolutamente trascurabile alla corrente di uscita. Quindi, tale corrente vale

$$I(t) \cong \rho\frac{s^2(t)}{2}$$

Il valore medio di questa corrente è $I = \rho\frac{A^2}{2}$, dove ricordiamo che ±A sono gli unici valori assunti dal segnale modulante s(t). Andiamo allora a calcolare la potenza media di segnale in uscita e la densità spettrale di potenza del rumore quantico:

[1] Ricordiamo che la modulazione PSK coincide con la modulazione ASK quando i rettangoli modulanti sono di tipo antipodale, cioè con ampiezze uguali in modulo e opposte in segno.

$$h_{nq} = 2qI = 2q\rho\frac{A^2}{2} = 2q\rho P_R$$

$$P_{S,media} = E\left[\left[\rho s(t)\cos\left(2\pi(2f_P)t + (2\varphi_P)\right)\right]^2\right] = \rho^2 E\left[s^2(t)\cos^2\left(2\pi(2f_P)t + (2\varphi_P)\right)\right] = \frac{\rho^2 A^2}{2} = \rho^2 P_R$$

dove abbiamo evidentemente posto $P_R = \frac{A^2}{2}$ in quanto si tratta della potenza media ricevuta dalla fibra (cioè la potenza media del segnale modulato).

Questo dunque avviene nel caso di **rivelazione diretta**.

Supponiamo adesso di procedere in altro modo, in sede di ricezione. Supponiamo infatti che il segnale portato ad incidere sul fotodiodo non sia più solo il segnale modulato $s_t(t) = s(t)\cos\left(2\pi f_P t + \varphi_P\right)$, ma la somma di questo segnale con una oscillazione locale del tipo $A_L\cos\left(2\pi f_L t + \varphi_L\right)$. La potenza ottica che incide sul fotodiodo è quindi

$$\begin{aligned}
&P(t) = \left[s(t)\cos\left(2\pi f_P t + \varphi_P\right) + A_L\cos\left(2\pi f_L t + \varphi_L\right)\right]^2 = \\
&= s^2(t)\cos^2\left(2\pi f_P t + \varphi_P\right) + A_L^2\cos^2\left(2\pi f_L t + \varphi_L\right) + A_L s(t)\cos\left(2\pi f_P t + \varphi_P\right)\cos\left(2\pi f_L t + \varphi_L\right) = \\
&= \left[\frac{s^2(t)}{2} + \frac{s^2(t)}{2}\cos\left(2\pi(2f_P)t + 2\varphi_P\right)\right] + \left[\frac{A_L^2}{2} + \frac{A_L^2}{2}\cos\left(2\pi(2f_L)t + 2\varphi_L\right)\right] + \\
&+ \left[A_L s(t)\cos\left(2\pi(f_L - f_P)t + (\varphi_L - \varphi_P)\right) + A_L s(t)\cos\left(2\pi(f_L + f_P)t + (\varphi_L + \varphi_P)\right)\right]
\end{aligned}$$

Abbiamo dunque 6 diversi contributi: possiamo subito trascurare, per le stesse considerazioni fatte prima, i termini ad alta frequenza, che sono evidentemente quelli attorno alle frequenze $2f_P$, $2f_L$ e f_P+f_L: abbiamo perciò che

$$P(t) \cong \frac{s^2(t)}{2} + \frac{A_L^2}{2} + A_L s(t)\cos\left(2\pi(f_L - f_P)t + (\varphi_L - \varphi_P)\right)$$

Il valore medio statistico della corrente prodotta dal fotorivelatore è allora pari, in base alla stessa formula ricavata in precedenza, a

$$I(t) = \rho P(t) = \rho\frac{s^2(t)}{2} + \rho\frac{A_L^2}{2} + \rho A_L s(t)\cos\left(2\pi(f_L - f_P)t + (\varphi_L - \varphi_P)\right)$$

Questa espressione si presta ad alcune interessanti osservazioni: supponendo che l'oscillazione locale abbia una ampiezza A_L molto maggiore dell'ampiezza del segnale modulante, risulta chiaramente che $\frac{s^2(t)}{2} << \frac{A_L^2}{2}$, per cui la componente predominante della corrente uscente è $\rho\frac{A_L^2}{2}$ ed è perciò questa componente la principale responsabile della generazione di rumore quantico; la componente di segnale è invece $\rho A_L s(t)\cos\left(2\pi(f_L - f_P)t + (\varphi_L - \varphi_P)\right)$. Possiamo allora andare a valutare la potenza media di segnale in uscita e la densità spettrale di potenza del rumore quantico: supponendo $\varphi_L=\varphi_P$ (su questo aspetto torneremo dopo), abbiamo

$$h_{nq} = 2qI \cong 2q\rho\frac{A_L^2}{2}$$

$$P_{S,media} = E\left[\left[\rho A_L s(t)\cos\left(2\pi(f_L - f_P)t\right)\right]^2\right] = \rho^2 A_L^2 E\left[s^2(t)\cos^2\left(2\pi(f_L - f_P)t\right)\right] = \frac{\rho^2 A_L^2 A^2}{2} = \rho^2 A_L^2 P_R$$

La potenza $P_{S,media}$ può essere resa molto più grande di quella che si ricava con rivelazione diretta (basta prendere A_L sufficientemente elevato): così facendo, riusciamo a rendere trascurabile l'effetto del rumore termico introdotto dal successivo amplificatore, per cui possiamo affermare che ***la densità spettrale di potenza di rumore da considerare nei calcoli è solo quella relativa al rumore quantico:***

$$h_{n,tot} = h_{nq} = 2q\rho\frac{A_L^2}{2}$$

Siamo dunque in un modo di funzionamento detto **quantum limited**, nel quale cioè le prestazioni non sono più limitate dal rumore termico, ma dal rumore quantico.

Calcoliamo allora il rapporto segnale/rumore all'uscita del fotodiodo, definendolo come rapporto tra l'energia della forma d'onda e la densità spettrale di rumore unilatera:

$$\frac{S}{N} = \frac{\frac{\rho^2 A_L^2 A^2}{2}T}{\left(2q\rho\frac{A_L^2}{2}\right)\frac{1}{2}}$$

Dato che stiamo considerando una codifica di tipo antipodale, dobbiamo imporre che questo rapporto sia maggiore o al più uguale a $\gamma^2 10^{0.05}$, dove abbiamo cioè eliminato il fattore 4 che invece compariva per la codifica ortogonale:

$$\frac{\frac{\rho^2 A_L^2 A^2}{2}T}{\left(2q\rho\frac{A_L^2}{2}\right)\frac{1}{2}} \geq \gamma^2 10^{0.05}$$

Facendo le opportune semplificazioni, otteniamo

$$\frac{\rho A^2 T}{q} \geq \gamma^2 10^{0.05}$$

Come si vede, questo rapporto S/N non dipende da A_L: questo risultato non deve sorprendere, in quanto abbiamo fatto l'ipotesi che A_L stesso fosse sufficientemente grande da rendere trascurabili gli altri termini e, in particolare, il rumore termico.

Ripetendo allora gli stessi calcoli fatti in precedenza, si ottengono 21 fotoni per simbolo (prendendo $\eta=1$ [1]), cioè un numero molto inferiore al valore di 56250/2=28125 fotoni per bit ottenuto prima. Non solo, ma si tratta di un valore estremamente vicino al valore limite ideale, che abbiamo stimato essere di 10 fotoni per simbolo. Confrontando il valore ottenuto in questo caso con

[1] Se prendessimo $\eta=0.8$, che è un valore sicuramente più ragionevole di $\eta=1$ si ottengono 26 fotoni per simbolo

quello ottenuto prima nel *sistema incoerente con APD* (che sta per *avalance photodetector*, ossia fotodiodo a valanga), si ricava un guadagno di circa 20 dB.

Questo sistema prende il nome di **struttura eterodina**, per indicare il fatto che, mediante l'oscillazione locale $A_L \cos(2\pi f_L t + \varphi_L)$, abbiamo traslato il segnale passa-banda in uscita dalla fibra a frequenza più bassa.

Esiste invece un altro sistema, detto a **struttura omodina**, in cui si effettua la stessa traslazione, ma questa volta fino a frequenza nulla: questo si ottiene, evidentemente, ponendo $f_P=f_L$, in modo che la corrente media in uscita dal fotodiodo sia

$$I(t) = \rho \frac{s^2(t)}{2} + \rho \frac{A_L^2}{2}$$

Così facendo, abbiamo in pratica effettuato già la demodulazione, visto che disponiamo direttamente del segnale s(t): questo è un vantaggio in quanto sappiamo che la demodulazione coerente presenta in uscita un rapporto S/N di 3dB superiore all'ingresso[1]. Se allora questo aumento di 3dB non è presente, è evidente che, *con un sistema omodina, possiamo trasmettere 3 dB in meno di potenza.*

Osservazioni varie

Facciamo adesso una serie di osservazioni circa i conti effettuati nel paragrafo precedente.

Intanto, riprendiamo l'espressione della corrente media in uscita dal fotodiodo, ottenuta trascurando i termini ad altissima frequenza:

$$I(t) = \rho P(t) = \rho \frac{s^2(t)}{2} + \rho \frac{A_L^2}{2} + \rho A_L s(t) \cos\left(2\pi(f_L - f_P)t + (\varphi_L - \varphi_P)\right)$$

Supponiamo di riuscire ad ottenere che la portante e l'oscillazione locale siano in fase, ossia $\varphi_L=\varphi_P$: in questo caso, la corrente in uscita dal fotodiodo è

$$I(t) = \rho \frac{s^2(t)}{2} + \rho \frac{A_L^2}{2} + \rho A_L s(t) \cos\left(2\pi(f_L - f_P)t\right) \cong \rho \frac{A_L^2}{2} + \rho A_L s(t) \cos\left(2\pi(f_L - f_P)t\right)$$

Se dimensioniamo opportunamente le frequenze f_L ed f_P, la componente di segnale $\rho A_L s(t) \cos\left(2\pi(f_L - f_P)t\right)$ si trova centrata nell'intervallo delle microonde, il che significa che, a valle del fotodiodo, possiamo utilizzare un normale **demodulatore a microonde**, effettuando così le note tecniche di demodulazione coerente.

C'è adesso da considerare un fatto: il termine $\rho A_L s(t) \cos\left(2\pi(f_L - f_P)t + (\varphi_L - \varphi_P)\right)$ si trova ad una frequenza più bassa di f_P solo nell'ipotesi in cui le fasi φ_L e φ_P siano costanti nel tempo; se anche una sola di esse fosse invece variabile nel tempo, il segnale sarebbe una sinusoide modulata in ampiezza dal segnale modulante, ma modulata anche in fase:

$$\rho A_L s(t) \cos\left(2\pi(f_L - f_P)t + \varphi(t)\right)$$

dove evidentemente $\varphi(t) = \varphi_L(t) - \varphi_P(t)$.

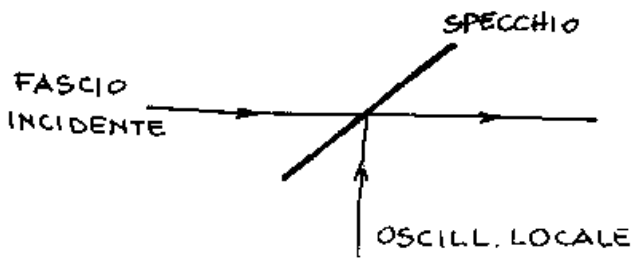
[1] Infatti, in un demodulatore coerente, la potenza di rumore rimane invariata dall'ingresso all'uscita (in quanto la densità spettrale di rumore raddoppia, ma la banda in cui integrarla si dimezza), mentre la potenza di segnale raddoppia.

In questo caso, non si avrebbe più una traslazione del segnale in bassa frequenza, ma una modifica dello spettro del segnale. Vediamo perché: sappiamo che, in uno schema di demodulazione coerente, un eventuale sfasamento φ costante tra portante e oscillazione locale provoca semplicemente una attenuazione del segnale di un fattore $\cos\varphi$; se, invece, lo sfasamento è funzione del tempo, allora, all'uscita del demodulatore, si ottiene un segnale modulato per $\cos\varphi(t)$. Allora, sarebbe comunque possibile effettuare una demodulazione coerente, a patto però di poter generare localmente la stessa modulazione di fase, ossia a patto di avere $\varphi_L(t) = \varphi_P(t)$ istante per istante.

Dovremmo cioè generare, in ricezione, una oscillazione locale del tipo $A_L \cos(2\pi f_L t + \varphi_L(t))$, con appunto $\varphi_L(t) = \varphi_P(t)$. Ma questo non è chiaramente ottenibile, in quanto la modulazione di fase (cioè il fatto che φ_P sia funzione del tempo) è il risultato della realizzazione di un processo casuale, per cui non è in alcun modo riproducibile.

Un'altra osservazione importante è la seguente: se la sorgente in trasmissione è a spettro largo, come ad esempio un LED, non è pensabile effettuare una traslazione in frequenza del segnale ottico, allo scopo di modularlo o demodularlo: questo a causa dei limiti dei **filtri ottici** (ossia filtri che lavorano a frequenza ottica), i quali hanno generalmente una banda passante piuttosto ampia, per cui ... Al contrario, per effettuare la traslazione in frequenza del segnale ottico, è necessario avere una sorgente che emetta una riga spettrale molto stretta rispetto allo spettro del segnale modulante.

Infine, ci chiediamo come si possa realizzare, nella pratica, la composizione di due segnali ottici come quella del segnale modulato $s(t)\cos(2\pi f_P t + \varphi_P)$ con l'oscillazione locale $A_L \cos(2\pi f_L t + \varphi_L)$. Basta utilizzare uno **specchio riflettente**, secondo lo schema della figura seguente:



Lo specchio, e in particolare la sua inclinazione, è tale che il segnale modulato $s(t)\cos(2\pi f_P t + \varphi_P)$ passi indisturbato, mentre l'oscillazione locale $A_L \cos(2\pi f_L t + \varphi_L)$ venga riflessa e si sovrapponga al segnale modulato.

Ci sono però due aspetti da considerare:

- in primo luogo, abbiamo capito, dai discorsi precedenti, che la composizione deve essere fatta in modo tale che le fasi dei due segnali siano uguali (oltre che costanti nel tempo);
- in secondo luogo, nel caso di fasci luminosi bisogna anche tener conto dei fronti d'onda, che possono combinarsi costruttivamente o meno: ciò significa che le due oscillazioni devono arrivare, sulla superficie dello specchio, non solo con la stessa fase, ma anche con fronti d'onda allineati; in altre parole, *non basta la coerenza temporale (cioè appunto le fasi uguali), ma serve anche la coerenza spaziale (cioè appunto fronti d'onda perfettamente sovrapposti).*

Autore: **SANDRO PETRIZZELLI**
e-mail: sandry@iol.it
sito personale: http://users.iol.it/sandry
succursale: http://digilander.iol.it/sandry1