Capitolo 18

Caratterizzazione circuitale, rumore ed equalizzazione dati

Dedichiamo questo capitolo al completamento della discussione di tre aspetti associati al contesto reale in cui avviene la trasmissione dei segnali, che sono propedeutici allo studio dei mezzi trasmissivi del cap. 19, e profondamente collegati all'aspetto fisico che caratterizza, appunto, la trasmissione. Può essere di aiuto far precedere la lettura di questo capitolo da un ripasso di quanto esposto al cap. 8; viceversa, è possibile saltare direttamente al capitolo successivo, e tornare qui seguendo qualche riferimento ad uno dei temi seguenti.

Il primo aspetto è la descrizione dell'attraversamento di un circuito elettrico da parte di un segnale, di immediata applicabilità nel caso dei collegamenti in cavo, ma che rappresenta un passaggio obbligato per caratterizzare qualunque interfaccia elettrica, inclusi quindi i collegamenti radio. Viene poi ripreso il discorso relativo al rumore termico (§ 8.4.2.1) allargandone le cause anche alla rumorosità introdotta dai circuiti attraversati, individuando come tenere conto dei diversi contributi al rumore in maniera unitaria, compreso il caso dei collegamenti composti da più tratte collegate da ripetitori, che come vedremo segnano un nuovo punto a favore delle trasmissioni numeriche. L'ultimo aspetto affrontato è relativo alle tecniche di equalizzazione attuabili per il caso di una modulazione numerica trasmessa su di un canale che presenta distorsione lineare (§ 8.2).

18.1 Modello circuitale dei segnali

Fino ad ora un *segnale* è stato trattato nelle diverse forme di espressione analitica, sequenza simbolica, grandezza aleatoria, forma d'onda nel tempo, e densità spettrale. E' giunto il tempo di confrontarci con la corrispettiva grandezza elettrica (o elettromagnetica) che lo veicola.

Potenza di segnale e grandezze elettriche La caratterizzazione energetica dei segnali si è finora svolta *a prescindere* dalla natura fisica degli stessi, in quanto non è mai stato specificato se si trattasse di tensioni o correnti, né si sono indicate le

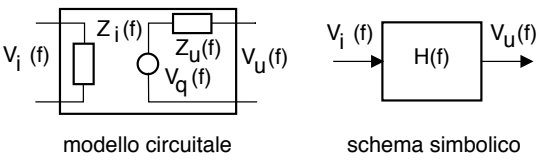
impedenze in gioco. Trattando ora di grandezze elettriche le potenze di segnale, di tensione o di corrente, saranno misurate in $(Volt)^2$ o in $(Ampere)^2$ rispettivamente.

Esempio Sia $x(t)$ un segnale di tensione. La sua potenza $\mathcal{P}_x$ ha unità di misura $\left[V^2\right]$, mentre la relativa densità di potenza $\mathcal{P}_x(f)$ si esprime in $\left[V^2/Hz\right]$.

Eseguiamo quindi una distinzione relativa al ruolo che il circuito ha nei confronti del segnale, tradizionalmente basata sul *numero di porte* del circuito.

Numero di porte Le coppie di morsetti a cui applicare o da cui prelevare un segnale vengono denominate *porte*. In questo senso un *generatore* che appunto produce il segnale, ed una *impedenza* di carico che ne assorbe la potenza, costituiscono reti ad *una* porta. Al contrario, l'oggetto che abbiamo fin qui indicato come filtro, o canale, da un punto di vista circuitale è un sistema fisico dotato di una relazione ingresso-uscita, e per questo indicato come rete *due porte*.

Modello di rappresentazione Un circuito può essere rappresentato mediante il suo *modello circuitale*, in cui sono evidenziati generatori, resistenze, impedenze, generatori controllati..., oppure il suo *schema simbolico*, in cui sono mostrate solamente le relazioni funzionali tra i segnali in transito.



Proprietà delle reti due porte Le proprietà di *linearità, permanenza, realizzabilità ideale e fisica, stabilità*, già definite al § 1.6 per i sistemi fisici, possono essere verificate o meno nelle reti due porte. D'altra parte, alcune relazioni e grandezze che nella *teoria dei circuiti* sono definite per segnali puramente sinusoidali, come per la *corrente alternata*, nella *teoria dei segnali* devono essere ridefinite in modo da tenere nel giusto conto dell'intera *densità spettrale* dei segnali con un contenuto informativo.

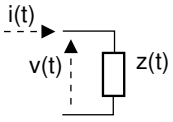
18.1.1 Bipoli

Distinguiamo tra quelli tipo attivo o di tipo passivo.

Passivi Non contengono generatori, e sono caratterizzati dalle relazioni esistenti tra la tensione ai loro capi e la corrente che vi scorre (entrante). Il legame tra le due grandezze è una *convoluzione*

$$v(t) = i(t) * z(t)$$

in cui si suppone $i(t)$ la causa e $v(t)$ l'effetto. La trasformata di Fourier fornisce $V(f) = I(f) \cdot Z(f)$ in cui $Z(f)$ prende il nome di *impedenza*, e può scriversi nei termini di parte reale ed immaginaria:



$$Z(f) = R(f) + jX(f)$$

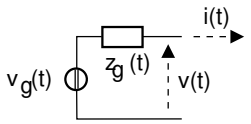
in cui $R(f)$ (*resistenza*) è una funzione *pari* di f (e sempre positiva), mentre $X(f)$ (*reattanza*) è dispari: pertanto, $Z(f) = Z^*(-f)$ e quindi $z(t)$ è reale. Allo stesso tempo

è definita *l'ammettenza*

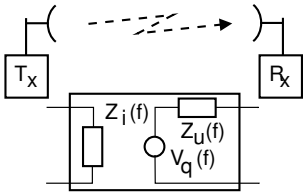
$$Y(f) = \frac{1}{Z(f)} = \frac{1}{R(f) + jX(f)} = \frac{R(f) - jX(f)}{|Z(f)|^2}$$

e la corrispondente $y(t) = \mathcal{F}^{-1}\{Y(f)\}$, permettendo di scrivere $i(t) = v(t) * y(t)$, in cui il ruolo di causa ed effetto per $i(t)$ e $v(t)$ è invertito.

Attivi Sono bipoli al cui interno è presente un generatore. Per il teorema di Thévenin,[1] qualunque circuito può essere ridotto ad un generatore di tensione con in serie una impedenza (vedi figura), in cui $V_g(f)$ rappresenta la tensione a vuoto, ossia quando $I(f) = 0$ (considerata uscente nei bipoli attivi).

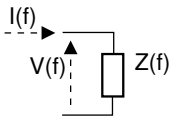


Esempio Una antenna trasmittente (§ 20.1) è schematizzabile come un bipolo passivo di impedenza pari all'impedenza di ingresso dell'antenna che assorbe la potenza erogata dal trasmettitore. Una antenna ricevente è schematizzabile come un generatore di tensione con in serie la propria impedenza di uscita, e trasferisce allo stadio di ingresso del ricevitore la potenza ricevuta per via elettromagnetica.



18.1.1.1 Potenza assorbita da un bipolo

Se ad un bipolo passivo di impedenza $Z(f)$ è applicato un segnale di tensione con spettro di densità di potenza $\mathcal{P}_v(f)$ la potenza *dissipata* sul bipolo (o assorbita), indicata come $\mathcal{W}_z$ per distinguerla da quella *di segnale* $\mathcal{P}_v$, ha densità



$$\mathcal{W}_z(f) = \mathcal{P}_v(f) \cdot \Re\{Y(f)\} = \mathcal{P}_v(f) \frac{R(f)}{|Z(f)|^2} \qquad \left[\frac{V^2}{\Omega \cdot Hz}\right] = \left[\frac{Watt}{Hz}\right] \tag{18.1}$$

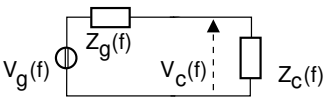
La dimostrazione di tale relazione è fornita al § 18.5.1. Osserviamo come la dipendenza di $Y(f)$ dalla frequenza svolga *una azione filtrante*, e la potenza totale assorbita (o dissipata) su $Z(f)$ vale

$$\mathcal{W}_z = \int_{-\infty}^{\infty} \mathcal{P}_v(f) \frac{R(f)}{|Z(f)|^2} df \qquad [Watt]$$

18.1.1.2 Connessione tra generatore e carico

La tensione ai capi del carico è valutabile applicando la *regola del partitore*[2]:

$$V_c(f) = V_g(f) \frac{Z_c(f)}{Z_c(f) + Z_g(f)}$$



ossia $V_c(f) = V_g(f) H(f)$ con $H(f) = \frac{Z_c(f)}{Z_c(f)+Z_g(f)}$. La densità di potenza di segnale ai

[1] Vedi ad es. `http://it.wikipedia.org/wiki/Teorema_di_Thévenin`

[2] `http://it.wikipedia.org/wiki/Partitore_di_tensione`

capi del carico vale quindi $\mathcal{P}_{v_c}(f) = \mathcal{P}_{v_g}(f)\,|H(f)|^2$, e la densità di potenza *dissipata* su $Z_c(f)$ risulta

$$\begin{aligned}\mathcal{W}_{z_c}(f) &= \mathcal{P}_{v_c}(f)\,\frac{R_c(f)}{|Z_c(f)|^2} = \mathcal{P}_{v_g}(f)\left|\frac{Z_c(f)}{Z_c(f)+Z_g(f)}\right|^2 \frac{R_c(f)}{|Z_c(f)|^2} = \\ &= \mathcal{P}_{v_g}(f)\,\frac{R_c(f)}{\left|Z_c(f)+Z_g(f)\right|^2} \qquad (18.2)\end{aligned}$$

Osserviamo dunque che la densità di potenza assorbita dal carico dipende da $Z_c(f)$, che compare sia a denominatore che a numeratore con $R_c(f)$. Ci chiediamo allora quale sia il valore di Z_c che realizza il *massimo trasferimento di potenza* tra generatore e carico, sfruttando così appieno la *potenzialità* del generatore, indicata come *potenza disponibile*.

18.1.1.3 Potenza disponibile e massimo trasferimento di potenza

La $\mathcal{W}_{z_c}(f)$ espressa da (18.2) risulta massima quando il denominatore viene reso minimo, ed al § 18.5.2 si mostra che ciò avviene qualora risulti $R_c(f) = R_g(f)$ e $X_c(f) = -X_g(f)$, ovvero $Z_c(f) = Z_g^*(f)$, in modo da poter enunciare

$$\begin{aligned}&se \qquad Z_c(f) = Z_g^*(f) \qquad (18.3)\\ &allora \qquad \mathcal{W}_{z_c}(f) = \max_{Z_c(f)}\left\{\mathcal{W}_{z_c}(f)\right\} = \frac{\mathcal{P}_{v_g}(f)}{4R_g(f)} = \mathcal{W}_{d_g}(f)\end{aligned}$$

Il valore $\mathcal{W}_{d_g}(f) = \frac{\mathcal{P}_{v_g}(f)}{4R_g(f)}$ prende il nome di spettro di potenza *disponibile* del generatore, dipende solo dai suoi parametri $\mathcal{P}_{v_g}(f)$ e $R_g(f)$, e rappresenta la *massima* potenza ceduta ad un carico che è *adattato* per il *massimo trasferimento di potenza*[3].

La potenza disponibile $\mathcal{W}_{d_g}(f)$ è pertanto *una grandezza caratteristica* del generatore; la potenza effettivamente ceduta ad un carico generico $Z_c(f) \neq Z_g^*(f)$, risulta inferiore a $\mathcal{W}_{d_g}(f)$ di una quantità

$$\alpha(f) = \frac{4R_g(f)\,R_c(f)}{\left|Z_g(f)+Z_c(f)\right|^2}$$

(vedi § 18.5.3) e quindi in generale si ha $\mathcal{W}_{z_c}(f) = \alpha(f)\,\mathcal{W}_{d_g}(f)$.

18.1.1.4 Adattamento di impedenza per assenza di distorsione lineare

Abbiamo già osservato come la tensione ai capi del carico abbia valore $V_c(f) = V_g(f) \cdot \frac{Z_c(f)}{Z_c(f)+Z_g(f)} = V_g(f)\,H(f)$. Ci chiediamo ora quali condizioni debbano sussistere affinché $H(f)$ si comporti come un *canale perfetto* (pag. 231), ovvero risulti $|H(f)| = cost$ e $\arg\{H(f)\} = -2\pi f\tau$: tali condizioni sono anche indicate come *assenza di distorsione lineare*. Il risultato cercato si ottiene qualora si ponga

$$Z_c(f) = \alpha Z_g(f) \quad \text{con } \alpha \text{ reale}$$

[3] E' bene notare esplicitamente che questo massimo è valido solo nel caso in cui non sia possibile modificare la $Z_g(f)$. Altrimenti, per un qualunque valore fissato di $Z_c(f)$, il massimo di $\mathcal{W}_{z_c}(f)$ si ottiene quando $Z_g(f) \to 0$.

infatti in tal caso risulta

$$H(f) = \frac{\alpha Z_g(f)}{(1+\alpha) Z_g(f)} = \frac{\alpha}{1+\alpha}$$

ossia $H(f)$ costante. La condizione $Z_c(f) = \alpha Z_g(f)$ prende il nome di *adattamento di impedenza*, a volte ristretta al caso in cui $\alpha = 1$.

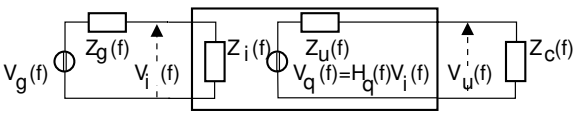
$Z_g(f)$ reale Notiamo che massimo trasferimento di potenza ed assenza di distorsione lineare possono sussistere *congiuntamente,* a patto che $Z_g(f) = R_g$, ovvero quando sia il generatore che il carico sono caratterizzati da una impedenza reale.

18.1.2 Reti due porte

Come anticipato un circuito accessibile mediante due coppie di morsetti è detto rete due porte, e può essere rappresentata secondo almeno due diversi formalismi: il *modello circuitale* e lo *schema simbolico*.

18.1.2.1 Modello circuitale

In figura è mostrato un possibile modello circuitale[4] per una rete due porte, caratterizzata in termini di impedenza di ingresso $Z_i(f)$, di uscita $Z_u(f)$, e di un generatore controllato con tensione a vuoto



$$V_q(f) = H_q(f) V_i(f)$$

Le condizioni di chiusura sono quelle di un generatore $V_g(f)$ con impedenza $Z_g(f)$ in ingresso, e di una impedenza di carico $Z_c(f)$ in uscita.

La tensione $V_i(f)$ all'ingresso della rete

$$V_i(f) = V_g(f) H_i(f)$$

dipende da quella del generatore $V_g(f)$ mediante il rapporto di partizione $H_i(f) = \frac{Z_i(f)}{Z_g(f)+Z_i(f)}$, così come la tensione in uscita

$$V_u(f) = V_q(f) H_u(f)$$

dipende da quella del generatore controllato $V_q(f)$ mediante il rapporto di partizione $H_u(f) = \frac{Z_c(f)}{Z_u(f)+Z_c(f)}$. Combinando queste relazioni, si ottiene che la risposta in frequenza complessiva $H(f)$ risulta:

$$V_u(f) = V_g(f) H_i(f) H_q(f) H_u(f) = V_g(f) H(f) \tag{18.4}$$

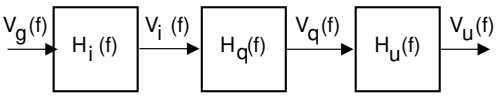
[4]Sono chiaramente possibili modelli diversi, basati su topologie e relazioni differenti. Esistono infatti circuiti a T, ad L, a scala, a traliccio, a pigreco; le relazioni tra le grandezze di ingresso ed uscita possono essere espresse mediante modelli definiti in termini di impedenze, ammettenze, e parametri ibridi.

Il caso qui trattato è quello di un modello ibrido, con la particolarità di non presentare influenze esplicite dell'uscita sull'ingresso. Qualora il circuito che si descrive presenti una dipendenza, ad esempio di Z_i da Z_c, o Z_u da Z_g, questo deve risultare nell'espressione della grandezza dipendente. Viceversa, qualora il circuito presenti in ingresso un generatore controllato da una grandezza di uscita, il modello non è più applicabile.

La relazione mostra come $H(f)$ dipenda, oltre che dalla risposta in frequenza intrinseca della rete $H_q(f)$, anche dalle *condizioni di adattamento* che si realizzano in ingresso ed in uscita.

18.1.2.2 Schema simbolico

Lo stesso modello circuitale descritto può essere rappresentato equivalentemente mediante lo schema simbolico rappresentato a lato, in cui sono evidenziate le tre funzioni di trasferimento sopra ricavate, e che operano sui segnali indicati. Lo schema simbolico ha il vantaggio di trascendere dal modello circuitale soggiacente, e di rendere del tutto evidente come la risposta in frequenza complessiva abbia origine dal prodotto di tre termini, di cui solo uno ($H_q(f)$) rappresenta la rete due porte in senso stretto.



18.1.2.3 Trasferimento energetico

Applicando ora la (18.1) alla potenza ceduta dal generatore controllato $V_q(f)$ al carico $Z_c(f)$, e tenendo conto della (18.4), si ottiene

$$\mathcal{W}_c(f) = \mathcal{P}_{v_u}(f)\,\frac{R_c(f)}{|Z_c(f)|^2} = \mathcal{P}_{v_g}(f)\,|H(f)|^2\,\frac{R_c(f)}{|Z_c(f)|^2}$$

che dipende anche da $Z_c(f)$. Proseguiamo l'analisi nel tentativo di individuare una relazione di trasferimento energetico che rappresenti caratteristiche *della sola rete*.

Guadagno di tensione E' definito come il rapporto tra tensione di uscita e di ingresso:

$$G_v(f) = \frac{V_u(f)}{V_i(f)} = H_q(f)\,H_u(f)$$

Evidentemente, dipende dalle condizioni di chiusura all'uscita della rete.

Guadagno di potenza E' il rapporto tra la potenza ceduta al carico e quella assorbita all'ingresso della rete:

$$\begin{aligned} G_{\mathcal{W}}(f) &= \frac{\mathcal{W}_c(f)}{\mathcal{W}_i(f)} = \mathcal{P}_{v_g}(f)\,|H(f)|^2\,\frac{R_c(f)}{|Z_c(f)|^2}\cdot\frac{1}{\mathcal{P}_{v_g}(f)}\,\frac{\left|Z_g(f)+Z_i(f)\right|^2}{R_i(f)} = \\ &= |H(f)|^2\,\frac{R_c(f)}{R_i(f)}\cdot\left|\frac{Z_g(f)+Z_i(f)}{Z_c(f)}\right|^2 = \\ &= \left|H_q(f)\right|^2\cdot\frac{R_c(f)}{R_i(f)}\cdot\left|\frac{Z_i(f)}{Z_u(f)+Z_c(f)}\right|^2 \end{aligned}$$

ed evidentemente è ancora funzione di $Z_c(f)$ ([5]). Notiamo ora che, qualora il carico sia adattato per il massimo trasferimento di potenza ($Z_c(f) = Z_u^*(f)$), la potenza ceduta

[5] L'ultimo passaggio tiene conto che (omettendo la dipendenza da f):

$$|H|^2\cdot\left|\frac{Z_g+Z_i}{Z_c}\right|^2 = \left|\frac{Z_i}{Z_i+Z_g}H_q\frac{Z_c}{Z_c+Z_u}\right|^2\left|\frac{Z_g+Z_i}{Z_c}\right|^2 = \left|H_q\right|^2\cdot\left|\frac{Z_i}{Z_u+Z_c}\right|^2$$

a $Z_c(f)$ (e quindi $G_{\mathcal{W}}(f)$) è massima, e la dipendenza di $G_{\mathcal{W}}(f)$ da $Z_c(f)$ decade, risultando

$$G_{\mathcal{W}_{Max}}(f) = \left|H_q(f)\right|^2 \cdot \frac{|Z_i(f)|^2}{4R_i(f)\,R_u(f)} \tag{18.5}$$

Guadagno disponibile Il rapporto tra la potenza disponibile di uscita e quella disponibile del generatore posto in ingresso della rete, indipendentemente dal fatto se l'ingresso della rete presenti o meno le condizioni per il massimo trasferimento di potenza, è detto *guadagno disponibile* e vale

$$\begin{aligned} G_d(f) &= \frac{\mathcal{W}_{d_u}(f)}{\mathcal{W}_{d_g}(f)} = \frac{\mathcal{P}_{v_q}(f)}{4R_u(f)} \cdot \frac{4R_g(f)}{\mathcal{P}_{v_g}(f)} = \\ &= \mathcal{P}_{v_g}(f)\,|H_i(f)|^2 \left|H_q(f)\right|^2 \cdot \frac{R_g(f)}{R_u(f)} \cdot \frac{1}{\mathcal{P}_{v_g}(f)} = \\ &= |H_i(f)|^2 \left|H_q(f)\right|^2 \cdot \frac{R_g(f)}{R_u(f)} \end{aligned} \tag{18.6}$$

La relazione trovata mostra la dipendenza di $G_d(f)$ dalle condizioni di chiusura in ingresso. Quando l'impedenza di ingresso $Z_i(f)$ è tale da permettere il conseguimento del massimo trasferimento di potenza, ovvero $Z_g(f) = Z_i^*(f)$, la dipendenza decade ed $|H_i(f)|^2 = \left|\frac{Z_i(f)}{Z_i(f)+Z_i^*(f)}\right|^2 = \frac{|Z_i(f)|^2}{4R_i^2(f)}$; considerando inoltre che $R_g(f) = R_i(f)$, la (18.6) diviene

$$G_{d_{Max}}(f) = \left|H_q(f)\right|^2 \cdot \frac{|Z_i(f)|^2}{4R_u(f)\,R_i(f)} \tag{18.7}$$

Quest'ultima quantità è chiamata *guadagno disponibile* DELLA RETE DUE PORTE ed è quella che appunto dipende solo dai parametri della rete stessa. Confrontando (18.7) con (18.5) notiamo che $G_{d_{Max}}(f)$ coincide con $G_{\mathcal{W}_{Max}}(f)$. Confrontando (18.7) con (18.6), troviamo che $G_d(f) = |H_i(f)|^2\, G_{d_{Max}}(f)\, \frac{4R_g(f)R_i(f)}{|Z_i(f)|^2}$. Considerando ora che $|H_i(f)|^2 \frac{1}{|Z_i(f)|^2} = \left|\frac{Z_i(f)}{Z_i(f)+Z_g(f)}\right|^2 \frac{1}{|Z_i(f)|^2} = \frac{1}{\left|Z_i(f)+Z_g(f)\right|^2}$, otteniamo

$$G_d(f) = \frac{4R_g(f)\,R_i(f)}{\left|Z_g(f)+Z_i(f)\right|^2} \cdot G_{d_{Max}}(f)$$

che ci consente di valutare $G_d(f)$ nelle reali condizioni di chiusura in ingresso, a partire da $G_{d_{Max}}(f) = G_{\mathcal{W}_{Max}}(f)$ che dipende solo dalla rete.

Collegamento generatore-carico mediante rete due porte

- Considerando generatore e porta di ingresso della rete adattati per il massimo trasferimento di potenza, la densità di potenza disponibile in uscita risulta

  $$\mathcal{W}_{d_u}(f) = G_{d_{Max}}(f)\,\mathcal{W}_{d_g}(f)$$

  e dunque l'uscita della rete due porte si comporta come un generatore equivalente, caratterizzato da una nuova $\mathcal{W}_{d_u}(f)$ ed una diversa impedenza interna $Z_u(f)$;
- se in ingresso non si verifica massimo trasferimento di potenza $G_d(f)$ si riduce di

un fattore $\beta(f) = \frac{4R_g(f)R_i(f)}{|Z_g(f)+Z_i(f)|^2}$, e dunque la nuova potenza disponibile in uscita risulta

$$\mathcal{W}_{d_u}(f) = \beta(f) \cdot G_{d_{Max}}(f)\, \mathcal{W}_{d_g}(f) = \frac{4R_g(f)\, R_i(f)}{\left|Z_g(f) + Z_i(f)\right|^2} \cdot G_{d_{Max}}(f)\, \mathcal{W}_{d_g}(f)$$

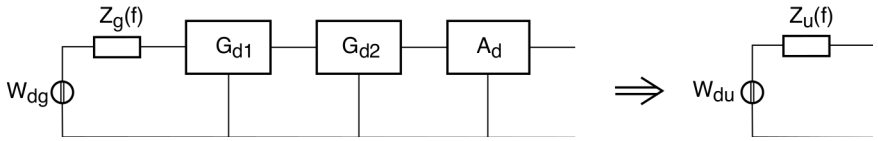
- se infine il carico $Z_c(f)$ in uscita alla rete non è adattato, assorbe una potenza inferiore a $\mathcal{W}_{d_u}(f)$ e pari a (vedi Appendice 18.5.3)

$$\mathcal{W}_c(f) = \alpha(f) \cdot \mathcal{W}_{d_u}(f) = \frac{4R_u(f)\, R_c(f)}{|Z_u(f) + Z_c(f)|^2} \cdot \mathcal{W}_{d_u}(f)$$

Reti passive Se una rete non contiene elementi attivi allora $G_{d_{Max}}(f) \leq 1$ per qualunque f. In questo caso si parla più propriamente di *attenuazione disponibile* $A_d(f) = \frac{1}{G_d(f)}$ ovvero $A_d(f)\,[\text{dB}] = -G_d(f)\,[\text{dB}]$.

Reti in cascata Se più reti sono connesse tra loro l'una di seguito all'altra, e si verificano per ciascuna coppia le condizioni di massimo trasferimento di potenza tra lo stadio di uscita di una e quello di ingresso della successiva, il guadagno disponibile complessivo è il prodotto dei singoli guadagni disponibili: $G_{d_{Tot}} = G_{d1} \cdot G_{d2} \cdot \ldots \cdot G_{dN}$.

Esempio In figura è mostrato un generatore con potenza disponibile W_{dg} collegato ad una serie di tre reti due porte; l'effetto complessivo è quello di un nuovo generatore



di uscita con potenza disponibile W_{du} pari al prodotto di quella del generatore originario, moltiplicata per i guadagni disponibili delle reti attraversate, tenendo anche eventualmente conto delle attenuazioni supplementari[6]:

$$W_{du} = W_{dg} \cdot G_{d1} \cdot G_{d2} \cdot \frac{1}{A_d} \cdot \frac{1}{A_s}$$

che può essere egualmente valutato operando in decibel, come

$$W_{du}\,[dBW] = W_{dg}\,[dBW] + G_{d1}\,[dB] + G_{d2}\,[dB] - A_d\,[dB] - A_s\,[dB]$$

in cui ovviamente, qualora W_{dg} fosse espresso in dBm anziché dBW, lo stesso accadrebbe per W_{du}.

Collegamento radio Con riferimento al circuito equivalente per una coppia di antenne di pag. 601, precisiamo che la potenza trasmessa è quella assorbita dall'impedenza di ingresso dell'antenna trasmittente, mentre quella ricevuta è quella ceduta dal generatore equivalente dell'antenna ricevente, all'impedenza di ingresso del ricevitore.

[6] L'attenuazione *supplementare* (pag. 639) può esprimere il peggioramento dovuto al mancato verificarsi delle condizioni di massimo trasferimento tra le impedenze di uscita e di ingresso delle reti affacciate.

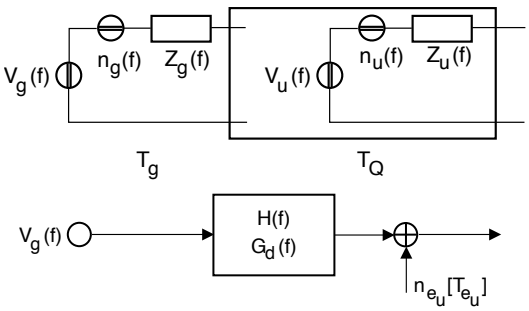
18.2 Rumore nelle reti due porte

Al § 8.4.2.1 abbiamo mostrato come in un bipolo attivo, una volta ridotto[7] nella sua forma canonica come un generatore di segnale (a vuoto) $V_g(f)$ con in serie una impedenza interna $Z_g(f)$ a temperatura T_g, quest'ultima contenga al suo interno *anche* un generatore di rumore *termico*, con densità di potenza di segnale $\mathcal{P}_n(f) = 2kT_gR(f)$ Volt^2, generato dalla sola parte reale $R(f) = \mathcal{R}\{Z_g(f)\}$ dell'impedenza. Qualora tale generatore sia chiuso su di un carico $Z_c(f) = Z_g^*(f)$ adattato per il massimo trasferimento di potenza, su $Z_c(f)$ si dissipa sia la potenza *disponibile* di segnale $\mathcal{W}_{dg}(f) = \frac{\mathcal{P}_g(f)}{4R(f)}$, sia quella (sempre disponibile) di rumore $\mathcal{W}_{d_n}(f) = \frac{\mathcal{P}_n(f)}{4R(f)} = \frac{1}{2}kT_g \left[\frac{\text{Watt}}{\text{Hz}}\right]$: pertanto, il generatore nasce *già di per se* rumoroso, con un

$$SNR_g(f) = \frac{\mathcal{W}_{dg}(f)}{\frac{1}{2}kT_g}$$

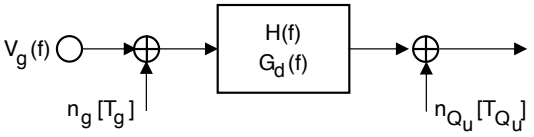
Qualora tra il generatore ed il carico siano invece interposte una (o più) reti due porte, occorre investigare su come tenere conto dell'ulteriore rumore introdotto da queste ultime.

Temperatura equivalente di uscita Collegando un generatore rumoroso a temperatura T_g all'ingresso di una rete due porte a temperatura T_Q (vedi fig. a lato), in uscita della rete troviamo un processo di rumore dipendente sia dal generatore che dalla rete, e la cui potenza disponibile $\mathcal{W}_{dn_u}(f)$ può essere espressa in funzione di una temperatura *equivalente* di uscita $T_{e_u}(f)$ (seconda figura), tale che



$$\mathcal{W}_{dn_u}(f) = \frac{1}{2}kT_{e_u}(f)$$

D'altra parte a $T_{e_u}(f)$ concorrono sia la temperatura del generatore $T_g(f)$, che la rete con una propria $T_{Q_u}(f)$ *equivalente di uscita* (fig. a lato); scriviamo dunque

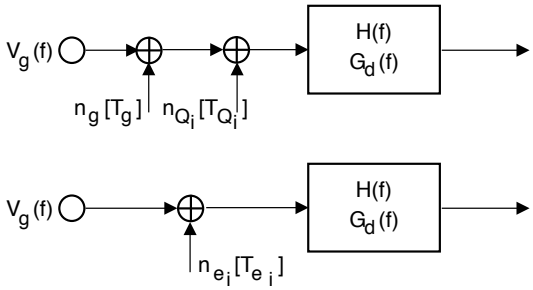


$$\mathcal{W}_{dn_u}(f) = \frac{1}{2}k \cdot \left[T_g(f)\, G_d(f) + T_{Q_u}(f)\right]$$

in cui la potenza disponibile in ingresso alla rete (che ha guadagno disponibile $G_d(f)$) è riportata in uscita, moltiplicata per $G_d(f)$.

Temperatura di sistema Se ora riportiamo in ingresso alla rete il contributo di

[7] Applicando il *teorema di Thévenin*, vedi `https://it.wikipedia.org/wiki/Teorema_di_Thévenin`



rumore dovuto a T_{Q_u} otteniamo

$$\mathcal{W}_{dn_u}(f) = \frac{1}{2}kG_d(f)\cdot\left[T_g(f) + T_{Q_i}(f)\right]$$

(in cui $T_{Q_i}(f) = \frac{T_{Q_u}(f)}{G_d}$), ovvero

$$\mathcal{W}_{dn_u}(f) = \frac{1}{2}kG_d(f)\,T_{e_i}(f)$$

dove

$$T_{e_i}(f) = T_g(f) + T_{Q_i}(f) = T_g(f) + \frac{T_{Q_u}(f)}{G_d(f)} \tag{18.8}$$

è detta anche *temperatura di sistema* $T_s = T_{e_i}$, poiché riporta in ingresso alla rete tutti i contributi al rumore di uscita, dovuti sia al generatore che alla rete. Siamo però rimasti con un problema irrisolto: che dire a riguardo di T_{Q_i} *e* T_{Q_u}?

18.2.1 Reti passive

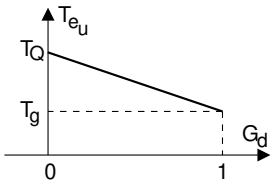
Nel caso in cui la rete due porte non contenga elementi attivi, e considerando tutti i componenti passivi della rete alla stessa temperatura T_Q, si può mostrare che risulta

$$\begin{cases} T_{Q_u}(f) &= [1 - G_d(f)]\,T_Q \\ T_{Q_i}(f) &= \frac{T_{Q_u}(f)}{G_d(f)} = [A_d(f) - 1]\,T_Q \end{cases}$$

in modo da poter scrivere:

$$\begin{cases} T_{e_u}(f) &= G_d(f)\,T_g(f) + T_{Q_u}(f) = G_d(f)\,T_g(f) + [1 - G_d(f)]\,T_Q \\ T_{e_i}(f) &= \frac{T_{e_u}(f)}{G_d(f)} = T_g(f) + [A_d(f) - 1]\,T_Q \end{cases}$$

Questo risultato evidenzia come per una rete passiva (con $0 \le G_d \le 1$), la temperatura di rumore equivalente in uscita sia una *media pesata* delle temperature del generatore e della rete. Nei casi limite in cui $G_d = 0$ oppure 1, la $T_{e_u}(f)$ è pari rispettivamente a T_Q e $T_g(f)$; infatti i due casi corrispondono ad una "assenza" della rete oppure ad una rete che non attenua.



18.2.1.1 Rapporto SNR in uscita

La valutazione del rapporto segnale rumore in uscita alla rete porta a

$$SNR_u(f) = \frac{\mathcal{W}_{dg}(f)\,G_d(f)}{\frac{1}{2}kT_{e_i}(f)\,G_d(f)} = \frac{\mathcal{W}_{dg}(f)}{\frac{1}{2}k\cdot\left[T_g(f) + [A_d(f) - 1]\,T_Q\right]}$$

Ricordando che il generatore in ingresso presenta un $SNR_i(f) = \frac{\mathcal{W}_{dg}(f)}{\frac{1}{2}kT_g(f)}$ possiamo valutare il *peggioramento* prodotto dalla presenza della rete come

$$\begin{aligned} \frac{SNR_i(f)}{SNR_u(f)} &= \frac{\mathcal{W}_{dg}(f)}{\frac{1}{2}kT_g(f)} \cdot \frac{\frac{1}{2}k\cdot\left[T_g(f) + [A_d(f) - 1]\,T_Q\right]}{\mathcal{W}_{dg}(f)} = \\ &= 1 + \frac{T_Q}{T_g(f)}\cdot[A_d(f) - 1] \end{aligned} \tag{18.9}$$

18.2.1.2 Fattore di rumore per reti passive

Il rapporto (18.9) $F(f) = 1 + \frac{T_Q}{T_g(f)} \cdot [A_d(f) - 1] \geq 1$ è chiamato *fattore di rumore*[8] della rete passiva, e rappresenta il peggioramento dell'SNR dovuto alla sua presenza, potendo infatti scrivere

$$SNR_u(f) = \frac{SNR_i(f)}{F(f)} \leq SNR_i(f)$$

Notiamo subito che se $T_g(f) = T_Q$, allora $F = A_d$: pertanto una rete passiva che si trova alla stessa temperatura del generatore presenta un fattore di rumore pari all'attenuazione. Infatti, mentre la potenza disponibile di rumore è la stessa (essendo generatore e rete alla stessa temperatura), il segnale si attenua di un fattore A_d.

18.2.2 Reti attive

In questo caso il rumore introdotto dalla rete non ha origine *solo dai resistori*, e dunque *non è più vero* che $T_{Q_u}(f) = [1 - G_d(f)]\, T_Q$. Inoltre, il guadagno disponibile può assumere valori $G_d > 1$. Per le reti attive si può quindi esprimere l'SNR in uscita come

$$SNR_u(f) = \frac{\mathcal{W}_{dg}(f)\, G_d(f)}{\frac{1}{2}k\left[G_d(f)\, T_g(f) + T_{Q_u}(f)\right]} = \frac{\mathcal{W}_{dg}(f)}{\frac{1}{2}k \cdot \left[T_g(f) + T_{Q_i}(f)\right]}$$

ed il peggioramento individuato in (18.9) come

$$\frac{SNR_i(f)}{SNR_u(f)} = \frac{\mathcal{W}_{dg}(f)}{\frac{1}{2}kT_g(f)} \cdot \frac{\frac{1}{2}k \cdot \left[T_g(f) + T_{Q_i}(f)\right]}{\mathcal{W}_{dg}(f)} \tag{18.10}$$

$$= 1 + \frac{T_{Q_i}(f)}{T_g(f)} = F\left(f, T_g\right) \tag{18.11}$$

Quest'ultima espressione dipende ancora da T_g. Allo scopo di ottenere una grandezza che dipenda solamente dalla rete due porte, si definisce quindi il

18.2.2.1 Fattore di rumore per reti attive

Viene posto pari a

$$F(f) = 1 + \frac{T_{Q_i}(f)}{T_0}$$

e rappresenta il peggioramento di SNR causato dalla rete quando il generatore è a temperatura ambiente $T_0 = 290\ ^oK = 17\ ^oC$. In realtà non ci è dato di conoscere $T_{Q_i}(f)$, mentre invece $F(f)$ può essere misurato a partire dal *rapporto dei rapporti* SNR, ed è proprio ciò che fa il costruttore della rete due porte. Il valore di $F(f)$ misurato ci permette dunque il calcolo di $T_{Q_i}(f) = T_0 [F(f) - 1]$ che, sostituito nella (18.8), permette finalmente di valutare la temperatura di sistema come

$$T_{e_i}(f) = T_g(f) + T_0 [F(f) - 1] \tag{18.12}$$

[8] A volte si incontra anche il termine *figura di rumore*, derivato dall'inglese NOISE FIGURE (che in realtà si traduce *cifra di rumore*), e che si riferisce alla misura di F in decibel.

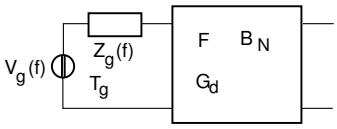
mentre dalla (18.11) si determina il peggioramento dell'SNR come

$$\frac{SNR_i(f)}{SNR_u(f)} = 1 + \frac{T_0}{T_g(f)}\left[F(f) - 1\right]$$

Riassunto

- il fattore di rumore è definito come il peggioramento di SNR dovuto alla presenza della rete tra generatore e carico, quando il generatore è a temperatura $T_0 = 290$ ^{o}K = 17 ^{o}C;
- dal fattore di rumore si deriva la temperatura di sistema $T_{e_i}(f) = T_g(f) + T_0\left[F(f) - 1\right]$;
- Se $T_g = T_0$ allora $T_{e_i}(f) = F(f)\,T_0$, e dunque la temperatura di sistema T_{e_i} è $F(f)$ volte quella del generatore;
- Se la rete non è rumorosa si ottiene $F = 1$ (pari a 0 dB);
- Se la rete è passiva allora $F(f) = \left[A_d(f) - 1\right]\frac{T_Q}{T_0} + 1$, e se è a temperatura $T_Q = T_0$ allora $F = A_d$.

Esempio Sia data una rete due porte con assegnati guadagno disponibile G_d, banda di rumore B_N e fattore di rumore F. Valutare il rapporto segnale rumore disponibile in uscita nei due casi in cui il generatore si trovi ad una generica temperatura T_g oppure a T_0.



Soluzione Sappiamo che la densità di potenza disponibile di rumore in uscita vale

$$\mathcal{W}_{dn_u}(f) = \frac{1}{2}kT_{e_i}G_d = \frac{1}{2}k \cdot \left[T_g + T_{Q_i}\right] \cdot G_d$$

in generale $F = 1 + \frac{T_{Q_i}}{T_0}$ e quindi $T_{Q_i} = T_0(F-1)$, dunque

$$\mathcal{W}_{dn_u}(f) = \frac{1}{2}k \cdot \left[T_g + T_0(F-1)\right] \cdot G_d$$

Pertanto, la potenza disponibile di rumore si ottiene integrando la densità sulla banda di rumore

$$\mathcal{W}_{dn_u} = k \cdot \left[T_g + T_0(F-1)\right] \cdot G_d B_N$$

che, nel caso in cui $T_g = T_0$, si riduce a $\mathcal{W}_{dn_u} = kT_0FG_dB_N$. Per la potenza di segnale, si ha invece $\mathcal{W}_{ds_u} = \mathcal{W}_{dg}G_d$, e pertanto se $T_g = T_0$, risulta

$$SNR_u = \frac{SNR_i}{F} = \frac{\mathcal{W}_{dg}}{kT_0FB_N} = \frac{\mathcal{W}_{dg}}{kT_{e_i}B_N}$$

ottenendo quindi lo stesso SNR in ingresso, ma con un rumore F volte più potente. Nel caso in cui T_g sia generico, considerando un fattore di rumore costante nella banda di rumore B_N, otteniamo:

$$\begin{aligned} SNR_u &= \frac{SNR_i}{F\left(T_g\right)} = \frac{\mathcal{W}_{dg}}{kT_gB_N} \cdot \frac{1}{1 + \frac{T_{Q_i}}{T_g}} = \frac{\mathcal{W}_{dg}}{kT_gB_N} \cdot \frac{1}{1 + \frac{T_0(F-1)}{T_g}} = \\ &= \frac{\mathcal{W}_{dg}}{k\left[T_g + T_0(F-1)\right]B_N} = \frac{\mathcal{W}_{dg}}{kT_{e_i}B_N} \end{aligned}$$

Esercizio Un trasmettitore con potenza di 50 mW e portante 30 MHz modula AM BLD PS un segnale con banda $\pm B = \pm$ 10 KHz, prodotto da un generatore a temperatura T_0. Qualora si desideri mantenere un SNR in ricezione di almeno 25 dB, determinare la distanza che è possibile coprire adottando antenne isotrope, ed un ricevitore caratterizzato da un fattore di rumore F = 10 dB. *N.B.: l'esercizio può essere affrontato con le ulteriori conoscenze dei §§ 19.1 e 20.2.*

Svolgimento Assumendo che si verifichino le condizioni di massimo trasferimento di potenza, il valore desiderato $SNR = SNR_0 = \frac{W_R}{W_N}$ può essere ottenuto se $W_R = W_N \cdot SNR = 2B \cdot \mathcal{W}_{dN}(f) \cdot SNR = B \cdot kT_0F \cdot SNR$, e quindi occorre ricevere un potenza

$$W_{R_{min}}(\text{dBm}) = 10\log_{10}10^4(\text{Hz}) - 174(\text{dBm/Hz}) + F_{dB} + SNR_{dB} =$$
$$= 40 - 174 + 10 + 25 = -99 \text{ dBm}.$$

Il guadagno di sistema (pag. 639) risulta allora pari a

$$G_s(\text{dB}) = W_T(\text{dBm}) - W_{R_{min}}(\text{dBm}) = 10\log_{10}50 + 99 = 17 + 99 = 116 \text{ dB}$$

Non prevedendo nessun margine, l'attenuazione dovuta alla distanza può essere pari al guadagno di sistema, e pertanto applicando la (20.6) di pag. 668 scriviamo

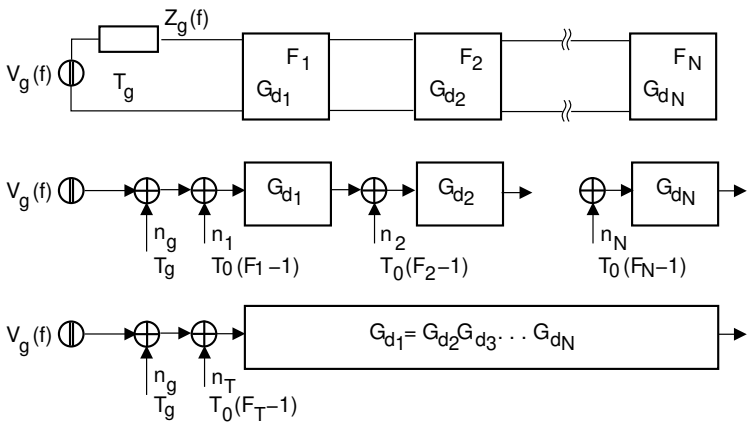
$$A_d = 116 = 32.4 + 20\log_{10} f(\text{MHz}) + 20\log_{10} d(\text{Km}) =$$
$$= 32.4 + 29.5 + 20\log_{10} d(\text{Km})$$

e quindi $2.7 = \log_{10} d(\text{Km})$, da cui $d = 10^{2.7} = 501$ Km. Svolgendo nuovamente i calcoli nel caso in cui il fattore di rumore del ricevitore sia pari a 20 dB e 100 dB, si ottiene che la nuova massima distanza risulta rispettivamente di 158 Km e di 15 metri.

18.2.3 Fattore di rumore per reti in cascata

Sappiamo che il guadagno disponibile dell'unica rete due porte equivalente alle N reti poste in cascata è pari al prodotto dei singoli guadagni, ovvero $G_d = \Pi_{n=1}^{N} G_{d_n}$. Come determinare invece il fattore di rumore equivalente complessivo ?

Con riferimento alla figura mostrata a lato, il singolo contributo di rumore dovuto a ciascuna rete può essere riportato *all'ingresso* della rete stessa, individuando così una temperatura $T_{Q_i}^{(n)} = T_0\left(F^{(n)} - 1\right)$



I singoli contributi possono quindi essere riportati *a monte* delle reti che li precedono, dividendo la potenza (ovvero la temperatura) per il guadagno disponibile delle reti *scavalcate*. Dato che i contributi di rumore sono indipendenti, le loro potenze si

sommano, e dunque è lecito sommare le singole temperature $T_{Q_i}^{(n)}$ riportate all'ingresso, in modo da ottenere un unico contributo complessivo di valore

$$T_{Q_i}^{(T)} = T_{Q_i}^{(1)} + T_{Q_i}^{(2)} \frac{1}{G_{d_1}} + T_{Q_i}^{(3)} \frac{1}{G_{d_1} G_{d_2}} + \cdots + T_{Q_i}^{(N)} \frac{1}{\prod_{n=1}^{N-1} G_{d_n}}$$

in cui, sostituendo le espressioni per i $T_{Q_i}^{(n)}$ si ottiene

$$T_{Q_i}^{(T)} = T_0 \cdot \left[F_1 - 1 + \frac{F_2 - 1}{G_{d_1}} + \frac{F_3 - 1}{G_{d_1} G_{d_2}} + \cdots + \frac{F_N - 1}{\prod_{n=1}^{N-1} G_{d_n}} \right]$$

Applicando la definizione $F^{(T)} = 1 + \frac{T_{Q_i}^{(T)}}{T_0}$, si ottiene

$$F^{(T)} = F_1 + \frac{F_2 - 1}{G_{d_1}} + \frac{F_3 - 1}{G_{d_1} G_{d_2}} + \cdots + \frac{F_N - 1}{\prod_{n=1}^{N-1} G_{d_n}}$$

che costituisce proprio l'espressione cercata:

$$F^{(T)} = F_1 + \sum_{i=2}^{N} \frac{F_i - 1}{\prod_{j=1}^{i-1} G_{d_j}}$$

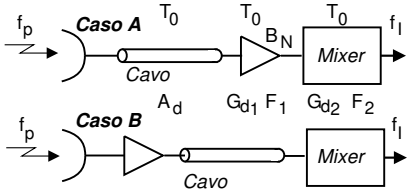
Il risultato, noto come *formula di Friis*[9], si presta alle seguenti considerazioni:

- la *prima* rete due porte deve avere F *più piccolo possibile*, dato che quest'ultimo non può essere ridotto in alcun modo e contribuisce per intero ad $F^{(T)}$;
- la *prima* rete due porte deve avere G_d *più elevato possibile*, dato che quest'ultimo divide tutti i contributi di rumore delle reti seguenti.

Pertanto l'elemento che determina in modo preponderante il rumore prodotto da una cascata di reti due porte è la *prima* rete della serie, ed il suo progetto deve essere eseguito con cura particolare, anche tenendo conto del fatto che le due esigenze sopra riportate sono spesso in contrasto tra loro. E' inoltre appena il caso di ricordare che l'espressione ottenuta *non è* in dB, mentre spesso F è fornito appunto in dB; pertanto per il calcolo di $F^{(T)}$ occorre prima esprimere tutti gli F_i in unità lineari.

Esercizio Una trasmissione video modulata AM-BLU con portante f_p = 2 GHz viene ricevuta secondo uno dei due schemi in figura, indicati come caso **A** e **B**.

E' presente una discesa in cavo coassiale con ϕ = 1.2/4.4 mm lunga 50 metri, un filtro-amplificatore con guadagno disponibile G_{d_1} = 20 dB, fattore di rumore F_1 = .4 dB e banda di rumore B_N = 7 MHz, ed un mixer che converte il segnale a frequenza intermedia f_I, e che esibisce G_{d_2} = 0 dB e F_2 = 10 dB. Tutti i componenti a valle dell'antenna si trovano alla stessa temperatura T_0 = 290 ^{o}K. Calcolare:



[9] Vedi `https://it.wikipedia.org/wiki/Formula_di_Friis`, ma da non confondere con la (20.6), anche se si tratta... della stessa persona!

1) La minima potenza disponibile W_{d_R} che occorre ricevere per ottenere SNR_0 = 50 dB nei due casi. Ripetere il calcolo supponendo l'antenna ricevente a temperatura T_a = 10 ^{o}K anziché T_0.

2) La minima potenza che è necessario trasmettere per superare un collegamento terrestre lungo 50 Km, con antenne di guadagno $G_T = G_R$ = 30 dB. Ripetere il calcolo per un down link satellitare in orbita geostazionaria, con $G_T = G_R$ = 40 dB.

3) Il valore efficace della tensione ai capi del generatore equivalente di uscita dell'amplificatore di potenza del trasmettitore, per il caso migliore (tra **A** e **B**) del collegamento terrestre, nel caso di massimo trasferimento di potenza con $Z_u = Z_a$ = 50 Ω, oppure con Z_u = 50 Ω e Z_a = 50 - j 50 Ω.

Svolgimento Determiniamo innanzitutto l'attenuazione del cavo coassiale, che risulta $A_d\,(f) = A_0\sqrt{f\,(MHz)}$ dB/Km. Per il diametro indicato risulta A_0 = 5.3 dB/Km, ed alla frequenza di 2 GHz si ottiene $A_d\,(f)_{dB} = 5.3\sqrt{2\cdot 10^3} = 237$ dB/Km; e quindi in 50 metri si hanno 11.85 ≃ 12 dB. Dato che il cavo è a temperatura T_0, risulta anche $F_{cavo} = A_d$ = 12 dB. Riassumendo:

	$A_d = F_{cavo}$	F_1	G_{d_1}	F_2	G_{d_2}
dB	11.85	.4	20	10	0
lineare	15.3	1.1	100	10	1

1) -

A) Il fattore di rumore complessivo risulta

$$F^A = F_{cavo} + A_d\,(F_1 - 1) + \frac{A_d}{G_{d_1}}\,(F_2 - 1) = 15.3 + 15.3\cdot(.1) + \frac{15.3}{100}\,(9) = 18.2$$

ovvero pari a 12.6 dB. Dato che per la trasmissione televisiva AM-BLU si ha $SNR = SNR_0$, scriviamo

$$W_{d_R} = SNR\cdot W_{d_N} = SNR_0\cdot F^A\cdot B_N\cdot kT_0 \text{ e quindi}$$

$$W_{d_R}(\text{dBm}) = SNR_0(\text{dB}) + F^A(\text{dB}) + B_N(\text{dBMHz}) + KT_0(\text{dBm/MHz}) =$$

$$= 50 + 12.6 + 8.45 - 114 = -43 \text{ dBm}$$

B) Il fattore di rumore complessivo risulta ora

$$F^B = F_1 + \frac{(F_{cavo}-1)}{G_{d_1}} + \frac{A_d}{G_{d_1}}\,(F_2 - 1) = 1.1 + \frac{14.3}{100} + \frac{15.3}{100}\,(9) = 2.26$$

ovvero pari a 3.5 dB. La differenza con il caso **A** è di 9.1 dB, e la potenza disponibile che occorre ricevere diminuisce pertanto della stessa quantità, e quindi ora risulta W_{d_R} = -52.1 dBm.

Nel caso in cui T_a = 10 ^{o}K ≠ T_0, non si ottiene più $T_{e_i} = FT_0$, ma occorre introdurre la T_{Q_i} della rete riportata al suo ingresso, e considerare la rete non rumorosa in modo da scrivere $T_{e_i} = T_g + T_{Q_i} = T_A + T_0\,(F-1)$. Ripetiamo i calcoli per i due casi **A** e **B**:

A) $W_{dR_W} = SNR\cdot W_{d_N} = SNR\cdot B_N\cdot k\cdot\left(T_a + T_{Q_i}\right) =$

$$= SNR\cdot B_N\cdot k\cdot\left(T_a + T_0\left(F^A - 1\right)\right), \text{ che espresso in dB fornisce}$$

$$W_{dR_{dBW}} = SNR_{dB} + 10\log_{10} 7\cdot 10^6 + 10\log_{10}\left(1.38\cdot 10^{-23}\,(10 + 290\cdot 17.2)\right)$$

$$= 50 + 68.5 - 191.61 = -73.11 \text{ dBW} = -43.11 \text{ dBm}$$

B) $W_{d_R}(\text{dBW}) = 50 + 68.5 + 10\log_{10}\left(1.38\cdot 10^{-23}\,(10 + 290\cdot 1.26)\right) = -84.3$ dBW = -54.3 dBm

Notiamo che se la T_a è ridotta, le prestazioni per la configurazione **A** migliorano di soli 0.11 dB, mentre nel caso **B** il miglioramento è di circa 2.2 dB. Questo risultato trova spiegazione con il fatto che in **A** predomina comunque il T_{Q_i} prodotto dal cavo.

2) In un collegamento radio terrestre si assume $T_a = 290\ ^oK$. Inoltre, per il caso in esame si trova una attenuazione disponibile pari a

$$A_d = 32.4 + 20\log_{10} f(\text{MHz}) + 20\log_{10} d(\text{Km}) - G_T - G_R =$$
$$= 32.4 + 66 + 34 - 60 = 72.4 \text{ dB}$$

A) $W_{d_T} = W_{d_R} + A_d = -43.11 + 72.4 = 29{,}29$ dBm = 850 mW

B) $W_{dT} = W_{d_R} + A_d = -54.3 + 72.4 = 18.1$ dBm = 66 mWatt

Per il downlink si ha $d = 36.000$ Km, mentre $T_a = 10\ ^oK$. Pertanto:

$$A_d = 32.4 + 20\log_{10} f(\text{MHz}) + 20\log_{10} d(\text{Km}) - G_T - G_R =$$
$$= 32.4 + 66 + 91.12 - 80 = 109.5 \text{ dB}$$

e quindi, utilizzando il valore W_{d_R} ottenuto per il caso **B**, otteniamo

$$W_{d_T} = W_{d_R} + A_d = -54.3 + 109.5 = 55.2 \text{ dBm} = 25.2 \text{ dBW} \rightarrow 331 \text{ Watt}$$

3) Nel caso di adattamento, la potenza ceduta all'antenna T_x è proprio quella disponibile del generatore, e quindi si ha $W_{d_T} = \frac{\sigma_g^2}{4R}$, da cui

$$\sigma_g = \sqrt{W_{d_T} 4R} = \sqrt{66 \cdot 10^{-3} \cdot 4 \cdot 50} = 3.63 \text{ Volt.}$$

In caso di disadattamento, desiderando che la potenza ceduta all'antenna trasmittente rimanga la stessa, e supponendo le impedenze indipendenti dalla frequenza, scriviamo (in accordo alla relazione (18.1))

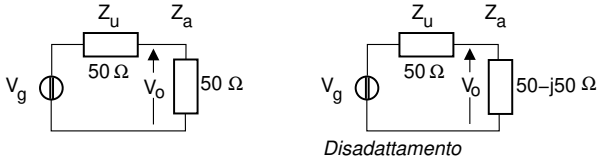
$$\mathcal{W}_T = \mathcal{P}_{v_o} \frac{R_a}{|Z_a|^2} = P_{v_o} \frac{50}{50^2+50^2} = P_{v_o} \cdot 10^{-2}$$

e quindi $P_{v_o} \simeq 6.6$ (Volt^2). Applicando ora la regola del partitore, si ottiene

$$P_{v_o} = P_{v_g} \left| \frac{Z_a}{Z_a+Z_u} \right|^2 = P_{v_g} \left| \frac{50-j50}{50+50-j50} \right|^2 = P_{v_g} \frac{50^2+50^2}{100^2+50^2} = P_{v_g} \cdot 0.4.$$

Dunque, $P_{v_g} = \frac{P_{v_o}}{0.4} = \frac{6.6}{0.4} = 16.5$ Volt^2, ovvero $V_{g_{eff}} = \sqrt{16.5} \simeq 4$ Volt.

Evidentemente, il disadattamento produce un innalzamento del valore efficace, se si vuol mantenere la stessa potenza di uscita.



Disadattamento

18.3 Rumore nei ripetitori

Come illustreremo al cap. 19, la propagazione del segnale attraverso un mezzo trasmissivo ne determina una *attenuazione* la cui entità aumenta con la distanza. La realizzazione di un collegamento molto lungo mediante un'unica *tratta* è pertanto

praticamente impossibile, sia a causa del livello troppo ridotto del segnale che sarebbe ricevuto, sia (per un collegamento radio) per la mancanza di condizioni di visibilità. Occorre pertanto suddividere il collegamento in più *tratte*, intervallate da stadi di amplificazione (o *ripetitori*) progettati in modo da compensare l'attenuazione del segmento appena attraversato.

18.3.1 Ripetitore trasparente

L'aggettivo *trasparente* si riferisce al fatto che, oltre al segnale, viene amplificato anche il rumore presente in ingresso, e nel caso di una trasmissione analogica, questo è l'unico modo di procedere. Le trasmissioni numeriche invece adottano un ripetitore *rigenerativo* che produce un *nuovo* segnale, privo di rumore ma con qualche bit errato in più: affronteremo questo caso al § 18.3.2.

Analizziamo ora la questione con riferimento ad un collegamento radio, anche se la trattazione può essere estesa ad altre tecniche trasmissive, come il cavo o le fibre ottiche, e consideriamo una successione di M tratte come mostrato in fig. 18.1. Il ripetitore (il *triangolo*) interposto tra ogni coppia di tratte (le *saette*) amplifica il segnale (ed il rumore) di una quantità pari al proprio guadagno disponibile G_{d_i}, reso uguale all'inverso dell'attenuazione disponibile della tratta precedente, ovvero $G_{d_i} = 1/A_{d_i}$. Il rumore termico accumulato alla fine del collegamento può essere calcolato con i metodi già discussi, ma considerando che il livello di segnale di uscita è lo stesso per tutti i ripetitori si ritrova il risultato ottenuto al § 8.4.1, come andiamo ad illustrare. Al § 18.3.1.2 valuteremo poi come le distorsioni *di non linearità* (§ 8.3) degli amplificatori possano intervenire nel progetto.

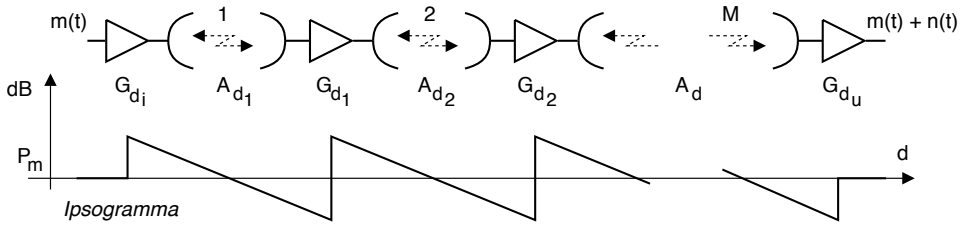


Figura 18.1: Suddivisione di un collegamento in M tratte, e relativo diagramma del livello di segnale in dB o *ipsogramma*

18.3.1.1 Rumore termico accumulato

In base all'uscita dell'ultimo ripetitore indicata come $m(t)+n(t)$ si può definire un SNR complessivo come $SNR_T = \mathcal{P}_m/\mathcal{P}_n$. D'altra parte il rumore $n(t)$ è dovuto ai contributi di rumore $n_i(t)$ introdotti nelle singole tratte; essendo tali contributi statisticamente indipendenti tra loro la potenza di rumore accumulata è *la somma* delle singole potenze:

$$\mathcal{P}_n = \sigma_n^2 = E\left\{n^2(t)\right\} = E\left\{\left(\sum_i n_i(t)\right)^2\right\} = \sum_i E\left\{n_i^2(t)\right\} = \sum_{i=1}^{M} \mathcal{P}_{n_i}$$

TeoriadeiSegnali.it

Deriviamo i singoli termini $\mathcal{P}_{n_i}$ osservando che all'uscita di ogni ripetitore è definito un $SNR_i = \mathcal{P}_{m_i}/\mathcal{P}_{n_i}$ *locale* da cui ottenere $\mathcal{P}_{n_i} = \mathcal{P}_{m_i}/SNR_i$, consentendoci di scrivere

$$SNR_T = \frac{\mathcal{P}_m}{\sum_i \mathcal{P}_{n_i}} = \frac{\mathcal{P}_m}{\sum_i \mathcal{P}_{m_i}/SNR_i}$$

A questo punto notiamo che, essendo il livello di segnale $\mathcal{P}_{m_i}$ in ingresso a ciascun ripetitore lo stesso, ovvero $\mathcal{P}_{m_i} = \mathcal{P}_m$ per $\forall i$, i singoli contributi $\mathcal{P}_{n_i}$ al rumore complessivo possono essere espressi nei termini di uno stesso livello di segnale, ovvero $\mathcal{P}_{n_i} = \frac{\mathcal{P}_m}{SNR_i}$, e dunque per l'$SNR$ complessivo si ottiene

$$SNR_T = \frac{\mathcal{P}_m}{\mathcal{P}_m \sum_i \frac{1}{SNR_i}} = \frac{1}{\sum_i \frac{1}{SNR_i}}$$

Questo risultato può essere espresso con la frase

> *l'SNR prodotto da più cause indipendenti è il parallelo degli SNR dovuti alle diverse cause di rumore*

per via della analogia formale con l'espressione della resistenza equivalente di un parallelo di resistenze; l'analogia evidenzia, tra l'altro, che se una tratta è considerevolmente peggiore delle altre, SNR_T dipenderà essenzialmente da questa.

Il risultato a cui siamo giunti ha validità più generale del caso illustrato, e può essere invocato ogni volta che un sistema di comunicazione è affetto da più cause di disturbo additivo indipendenti tra loro, per ognuna delle quali si sia separatamente in grado di giungere ad una espressione di SNR, come illustrato anche al § 8.4.1.

Proseguiamo l'analisi ipotizzando ora che tutte le tratte siano *uguali tra loro*, ovvero con eguali A_d e G_d, uguali temperature di rumore, ed uguali SNR_i. In tal caso si ottiene

$$SNR_T = \frac{1}{\frac{M}{SNR_i}} = \frac{SNR_i}{M}$$

ovvero un $M-$esimo dell'SNR_i locale di ciascun ripetitore[10]. Sembrerebbe dunque che per migliorare l'SNR complessivo sia sufficiente aumentare la potenza di trasmissione di tutti gli stadi, in modo da elevare la potenza ricevuta. In realtà la potenza trasmessa non può aumentare a piacere, in quanto intervengono fenomeni di non-linearità.

18.3.1.2 Compromesso tra rumore termico e distorsione

Ai §§ 8.3 e 13.3 si è studiato come per un segnale modulato la presenza di un elemento a comportamento non lineare (tipicamente l'amplificatore di potenza del ripetitore) produce *interferenza in banda*, la cui potenza $\mathcal{P}_{NL}$ dipende con legge cubica dalla potenza del segnale trasmesso. Indichiamo quindi con $SNR_{NL} = \frac{\mathcal{P}_m}{\mathcal{P}_{NL}}$ il rapporto SNR complessivo del collegamento dovuto a cause di non linearità, ed osserviamo che questo *diminuisce* all'aumentare della potenza trasmessa da ogni ripetitore. L'SNR complessivo deve tener conto sia del rumore termico che della distorsione non lineare, e

[10] Con $SNR_i = \alpha SNR_0 = \alpha \frac{\mathcal{P}_R}{\mathcal{P}_n}$, dove $\mathcal{P}_n = kT_{e_i}W$ è la potenza di rumore nella banda di messaggio W, $\mathcal{P}_R$ è la potenza ricevuta da un ripetitore (uguale per tutti se le tratte sono uguali), e α è un fattore che dipende dal tipo di modulazione (cap. 14).

dato che questi sono *statisticamente indipendenti*, può essere espresso come "il parallelo" di entrambi, ossia

$$SNR = \frac{1}{\frac{1}{SNR_T} + \frac{1}{SNR_{NL}}}$$

Dato che all'aumentare della potenza di trasmissione SNR_T diminuisce mentre SNR_{NL} aumenta, l'SNR complessivo presenta *un massimo* per un certo valore di potenza trasmessa, ovvero esiste un dimensionamento *ottimo* in grado di fornire il miglior SNR complessivo.

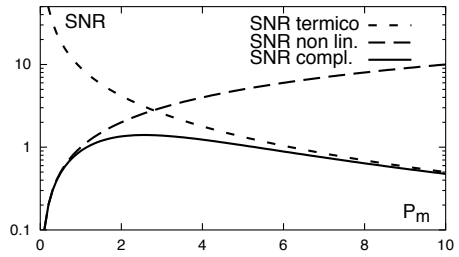
Esempio La figura a lato mostra l'andamento di

$$SNR = \frac{1}{\frac{1}{SNR_T} + \frac{1}{SNR_I}}$$

dovuto ai due termini

. $SNR_T = \mathcal{P}_m$ e $SNR_{NL} = \frac{\mathcal{P}_m}{.1 \cdot \mathcal{P}_m^2 + .01 \cdot \mathcal{P}_m^3}$

Come si vede, con questi valori SNR presenta un massimo per $\mathcal{P}_m \simeq 2.5$.



18.3.2 Ripetitore rigenerativo

Affrontiamo ora il caso in cui la trasmissione sia di natura numerica (§ 16): in tal caso i dispositivi intermedi di amplificazione anziché essere di tipo *trasparente*, vengono detti *rigenerativi*, ovvero mentre dal lato ricevente svolgono l'intero processo di decisione (§ 15.4) sui simboli del messaggio, dal lato di trasmissione generano un *nuovo* segnale in cui è *assente* il rumore presente in ricezione.

La valutazione delle prestazioni *complessive* nel caso si adottino M ripetitori rigenerativi posti in serie lungo un medesimo collegamento è piuttosto semplice, qualora siano equispaziati, ed ognuno caratterizzato dalla stessa $p = P_e^{bit}$. In tale circostanza si ricade infatti nel caso delle *prove ripetute*, permettendo di esprimere la probabilità che lo stesso bit sia sbagliato da n ripetitori (su M) mediante la distribuzione di Bernoulli (§ 22.1) $p(n) = \binom{M}{n} p^n (1-p)^{M-n}$. Ovviamente il bit risulterà sbagliato solo se avrà subito un numero *dispari* di errori, e dunque complessivamente possiamo scrivere

$$P_e^T = \sum_{\substack{n=1 \\ n\ dispari}}^{M} p(n) = Mp(1-p)^{M-1} + \frac{M(M-1)(M-2)}{3!} p^3 (1-p)^{M-3} + \cdots$$

Nel caso, come spesso accade, che $p \ll 1$, l'espressione indicata è bene approssimata dal solo primo termine, e dunque si può affermare che $P_e^T \simeq MP_e^{bit}$, ovvero che la prob. di errore sul bit è proporzionale al numero di tratte.

Esempio con un valore di $p = 10^{-6}$ occorrono 10 tratte per arrivare a $P_e^T = 10^{-5}$.

Ripetitore trasparente nelle trasmissioni numeriche In questo caso ogni ripetitore *trasparente* amplifica oltre al segnale anche il rumore già presente, e ne aggiunge di

suo: come mostrato al § 18.3.1.1 nel caso di tratte identiche all'estremo di destinazione del collegamento si osserva un $SNR_T = \frac{SNR_i}{M}$, pari cioè a quello di ogni singola tratta, diviso per il numero delle tratte. In tal caso l'E_b/N_0 del ricevitore a destinazione (l'unico ora che esegue il processo di decisione) risulta ridotto dello stesso fattore M (vedi eq. (15.16) pag. 455), producendo un peggioramento delle sue condizioni operative di $10 \log_{10} M$ dB rispetto a quelle delle singole tratte.

Esercizio Una trasmissione numerica 16-QAM (§ 16.3) è realizzata mediante un collegamento suddiviso in $M = 16$ tratte uguali. Qualora si desideri una P_e^T complessiva pari a 10^{-5}, valutare il valore di E_b/N_0 necessario in ingresso a ciascuna tratta, nei due casi **a)** ripetitori rigenerativi oppure **b)** ripetitori trasparenti. Indicare quindi l'entità del rapporto tra le potenze di segnale in ingresso ai ripetitori nei due casi.

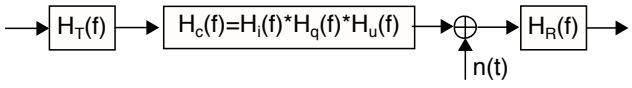
Svolgimento Nel caso **a)** risulta $P_e^T \simeq M P_e$, dunque per ogni tratta si deve ottenere una $P_e = P_e^T/M = 10^{-5}/16 = 6.25 \cdot 10^{-7}$, a cui (per un 16-QAM) corrisponde (fig. 16.14) un valore $E_b/N_0 \simeq 14.5$ dB. Nel caso **b)** la decisione ha luogo solamente nell'ultima tratta, e per ottenere una $P_e^T = 10^{-5}$, la modulazione 16-QAM necessita di un valore $E_b/N_0 \simeq 13.5$ dB. D'altra parte, per ottenere un tale E_b/N_0 in ingresso al decisore, *ogni* tratta che lo precede deve presentare un E_b/N_0 in ingresso allo stadio di amplificazione *maggiore* (rispetto a quello del decisore) di una quantità pari a $\log_2 M = 4$ dB, ossia $E_b/N_0 = 13.5 + 4 = 17.5$ dB, ovvero 3 dB *in più* rispetto al caso **a)**.

Essendo le tratte identiche per ognuna di esse il valore $N_0 = 1/2 k T_{ei}$ è lo stesso, in quanto (vedi eq. 18.12) $T_{ei} = T_g + T_0 (F - 1)$, e sia T_g che F sono gli stessi per tutte le tratte. Pertanto la differenza di 3 dB tra i valori di E_b/N_0 nei casi **a)** e **b)** si riflette in un *raddoppio* della potenza in ingresso ad ogni ripetitore trasparente, in confronto a quello rigenerativo.

18.4 Equalizzazione numerica

Occupiamoci ora di un argomento del tutto differente, ovvero come *correggere* una risposta in frequenza complessiva del mezzo trasmissivo e delle sue chiusure $H_c(f) = H_i(f) H_q(f) H_u(f)$ che presenta distorsione lineare (§ 8.2) e dunque non soddisfa la condizione di canale perfetto (pag. 231), specializzando il discorso al caso di una trasmissione dati[11] per la quale $H_c(f)$ può anche non essere nota a priori.

Andando con ordine, al § 15.3 abbiamo già osservato come l'equalizzazione può essere perseguita inserendo elementi filtranti sia presso il trasmettitore ($H_T(f)$) che al ricevitore ($H_R(f)$), realizzati in modo da ottenere



$$H_T(f) H_c(f) H_R(f) = a e^{-j2\pi f \tau}$$

annullando così l'effetto di $H_c(f)$. Scegliendo di utilizzare uno solo dei due filtri si può delegare tutto il compito dell'equalizzazione a $H_T(f)$, riservando ad $H_R(f)$ il ruolo di filtrare il rumore esterno alla banda di segnale, come fatto ai cap. 14 e 19.

[11] I casi di segnale di banda base e di segnale modulato sono trattati nel seguito in modo unitario, facendo riferimento solo ai primi, ovvero alle c.a. di b.f. nel caso di segnale modulato.

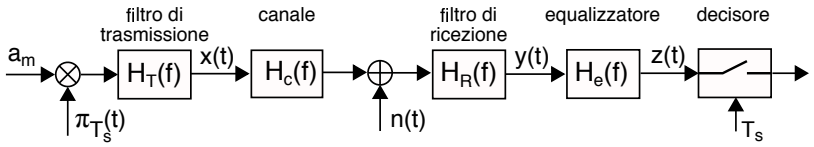


Figura 18.2: Equalizzazione di un canale con distorsione lineare

Quando la $H_c(f)$ del canale è *nota a priori* la sua inversione[12] può essere approssimata sintetizzando $H_T(f)$ mediante un metodo di progetto filtri analogici (§ 5.1).

D'altra parte se l'equalizzazione è svolta solamente *al lato ricevente* si possono applicare tecniche volte a *stimare* la $H_c(f)$ qualora non sia nota[13], e quindi *sintetizzare* una realizzazione numerica di $H_R(f)$ tale da invertire $H_c(f)$. In tal caso si determina la *colorazione* del rumore in ingresso al ricevitore (vedi pag. 621), che viene esaltato anche di molto nelle bande per le quali $|H_c(f)|$ è particolarmente piccolo, come ad es. per un canale con echi di ampiezza simile, vedi pag. 238.

Nel caso di una *trasmissione numerica* è già prevista la presenza di un filtro $H_T(f)$ con la funzione di sagomatore di impulso (come componente del codificatore di linea, § 15.1.2), mentre la teoria del filtro adattato (§ 7.6) stabilisce che realizzando[14] $H_R(f) = H_T^*(f)$ si ottiene il massimo SNR all'istante di decisione. Lo schema di elaborazione complessivo è dunque quello rappresentato in fig. 18.2, che adottiamo come riferimento, ed in cui $H_R(f) H_T^*(f) = G(f)$, in modo che *assieme* realizzino un filtro di Nyquist[15], e l'equalizzazione di $H_c(f)$ sia demandata per intero al ricevitore, per mezzo di $H_e(f)$. In ingresso ad $H_e(f)$ si presenta dunque il segnale

$$y(t) = \sum_{m=-\infty}^{\infty} a_m \tilde{g}(t - mT_s) + \nu(t) \tag{18.13}$$

che dipende oltre che dai simboli a_m della sequenza informativa, anche dalla risposta impulsiva complessiva $\tilde{g}(t) = h_T(t) * h_c(t) * h_R(t)$[16], mentre $\nu(t)$ tiene conto dell'effetto di $H_R(f)$ su $n(t)$. Scegliamo di realizzare $H_e(f)$ nella forma di un filtro FIR (§ 5.2.1, vedi fig. 18.3) con $2N+1$ coefficienti c_n, in modo che alla sua uscita si trovi il segnale[17]

$$z(t) = \sum_{n=-N}^{N} c_n y(t - n\tau) \tag{18.14}$$

[12] Detta anche operazione di *deconvoluzione*, vedi `https://en.wikipedia.org/wiki/Deconvolution`.

[13] Operazione necessaria se il canale *cambia* da un collegamento all'altro, come nel transito in una rete commutata, o nel caso di una comunicazione radiomobile.

[14] Si veda l'eq. (7.23) a pag. 213, tralasciando il termine di fase lineare necessario a rendere $h_R(t)$ causale.

[15] Sia che si abbia $H_R(f) = H_T(f) = \sqrt{G(f)}$ come per il caso del ricevitore ottimo (§ 15.5), sia nel caso in cui $H_T(f) = G(f)$, e $H_R(f)$ ha il solo scopo di filtrare il rumore. C'è poi l'aspetto trattato al § 15.5.1 relativo all'equalizzazione *distribuita*, che prevede di ripartire il compito in parti uguali tra i due lati, e che necessita di un canale *di ritorno* per comunicare al trasmettitore le stime operate a destinazione.

[16] Che equivale all'impulso $\tilde{g}(t)$ di eq. (15.1) a pag. 437

[17] La (18.14) rappresenta l'uscita di un filtro FIR *anti-causale*, dato che $z(t)$ dipende da valori *futuri* di ingresso, fino a $y(t + N\tau)$, mentre l'espressione *causale* che rispecchia la notazione usata nella figura dovrebbe essere $z(t) = \sum_{n=-N}^{N} c_n y(t - n\tau - N\tau)$, che corrisponde ad un semplice ritardo dell'ingresso pari a $N\tau$, trascurato per semplicità di notazione nel seguito.

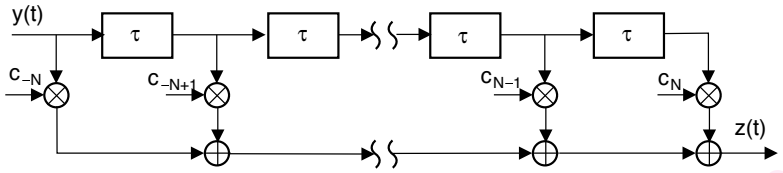


Figura 18.3: Filtro FIR di equalizzazione

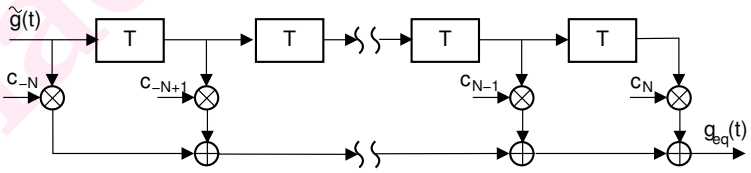
Osserviamo ora che se $H_e(f)$ riesce nel suo intento l'ISI introdotta dal canale è completamente rimossa e campionando l'uscita dell'equalizzatore (18.14) in corrispondenza degli istanti di simbolo si ottiene un valore $z(mT_s)$ che dipende da un solo a_m, oltre che dal rumore filtrato via $H_R(f)\,H_e(f)$; pertanto i coefficienti c_n possono essere scelti direttamente con l'obbiettivo di annullare l'ISI, anziché con quello di invertire $H_c(f)$.

Illustriamo per prima una tecnica che non tiene in particolare conto la presenza del rumore, che viene invece considerato al § 18.4.2 e seguenti.

18.4.1 Equalizzatore zero forcing

Questa tecnica prevede che gli impulsi $\tilde{g}(t)$ ricevuti distorti attraversino un filtro trasversale i cui coefficienti sono calcolati in modo da (quasi) ripristinare le condizioni di Nyquist (nel dominio del tempo) per l'impulso $g_{eq}(t)$ in uscita dal filtro.

Per poter funzionare, si fa precedere il messaggio informativo dalla trasmissione di un impulso isolato *di apprendimento*, la cui forma d'onda $\tilde{g}(t)$ ricevuta distorta presenta un picco a $t = 0$ ed ISI su entrambi i lati (Fig. 18.4a). Tale impulso $\tilde{g}(t)$ è posto in ingresso al filtro trasversale $h_{eq}(t)$ con $2N + 1$ coefficienti che ridisegnamo sotto con il ritardo τ tra gli ingressi pari a T, per un ritardo totale di attraversamento $2NT$. In uscita da $h_{eq}(t)$ si ottiene l'impulso



$$g_{eq}(t) = \sum_{n=-N}^{N} c_n \tilde{g}(t - nT - NT) \tag{18.15}$$

che campionato agli istanti $t_k = kT + NT$ fornisce

$$g_{eq}(t_k) = \sum_{n=-N}^{N} c_n \tilde{g}(kT - nT) = \sum_{n=-N}^{N} c_n \tilde{g}_{k-n} \tag{18.16}$$

avendo adottato l'abbreviazione $\tilde{g}_{k-n} = \tilde{g}(kT - nT)$; il risultato prende quindi la forma di una *convoluzione discreta*. L'operazione di *equalizzazione Zero Forcing* consiste nel calcolare i coefficienti c_n in modo che $g_{eq}(t)$ soddisfi le condizioni di Nyquist almeno sui $2N + 1$ termini centrati sull'origine, ovvero

$$g_{eq}(t_k) = \begin{cases} 1 & k = 0 \\ 0 & k = \pm 1, \pm 2, \cdots, \pm N \end{cases} \tag{18.17}$$

assicurando così l'assenza di ISI *almeno* nei confronti degli N simboli precedenti

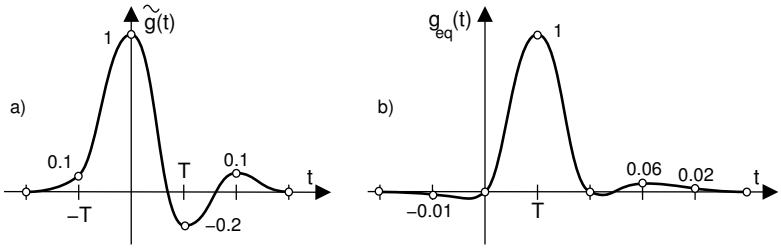


Figura 18.4: Impulso con ISI prima e dopo equalizzazione, caso di $N = 1$

e successivi. Il valore dei coefficienti c_n in grado di soddisfare queste condizioni si ottengono risolvendo un sistema di $2N+1$ equazioni nelle $2N+1$ incognite c_n, impostate a partire dalle relazioni (18.16) e (18.17), e rappresentato in forma matriciale come

$$\begin{bmatrix} \tilde{g}_0 & \cdots & \tilde{g}_{-2N} \\ \vdots & & \vdots \\ \tilde{g}_{N-1} & \cdots & \tilde{g}_{-N-1} \\ \tilde{g}_N & \cdots & \tilde{g}_{-N} \\ \tilde{g}_{N+1} & \cdots & \tilde{g}_{-N+1} \\ \vdots & & \vdots \\ \tilde{g}_{2N} & \cdots & \tilde{g}_0 \end{bmatrix} \begin{bmatrix} c_{-N} \\ \vdots \\ c_{-1} \\ c_0 \\ c_1 \\ \vdots \\ c_N \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{18.18}$$

in cui la matrice dei termini noti è costituita dai $4N + 1$ campioni di $\tilde{g}(t)$ prelevati agli istanti di simbolo $-2N, \ldots, 2N$ posti in modo simmetrico rispetto allo zero. Anche se il metodo non garantisce nulla per istanti esterni a $(-NT, NT)$, è *ottimo* nel senso che minimizza l'ISI di picco, ed è semplice da realizzare. L'operazione di combinazione di più campioni ricevuti eseguita dal filtro di equalizzazione può avere l'effetto di aumentare la potenza di rumore, ma se questa non è eccessiva tale peggioramento è più che compensato dal miglioramento dell'ISI.

Esempio Vogliamo realizzare un equalizzatore del terzo ordine, applicando quanto esposto all'impulso con ISI mostrato alla fig. 18.4a. Inserendo i valori di $\tilde{g}_k$ mostrati nella matrice dei coefficienti (18.18) otteniamo

$$\begin{bmatrix} 1.0 & 0.1 & 0.0 \\ -0.2 & 1.0 & 0.1 \\ 0.1 & -0.2 & 1.0 \end{bmatrix} \begin{bmatrix} c_{-1} \\ c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

da cui è possibile calcolare $c_{-1} = -0.096 \quad c_0 = 0.96 \quad c_1 = 0.2$. Inserendo ora uno ad uno i campioni di $\tilde{g}(t)$ in un filtro trasversale con questi coefficienti, si ottengono i valori di $g_{eq}(t)$ mostrati in fig. 18.4b assieme ad una curva interpolata. Notiamo che sebbene il risultato desiderato (di azzerare $g_{eq}(t)$ ai due istanti di simbolo ai lati del picco) sia stato ottenuto, risulta essere ancora $g_{eq} \neq 0$ ad istanti più remoti.

Effetto del rumore dopo equalizzazione Come anticipato l'approccio dello *zero forcing* trascura la presenza del rumore in ricezione. Poniamo che riesca nel suo intento di realizzare una $H_e(f) = 1/H_c(f)$ come desiderato: la densità di potenza del rumore in uscita sarà allora pari a

$$\mathcal{P}_n(f) = \frac{N_0}{2} |H_R(f)|^2 |H_e(f)|^2 = \frac{N_0}{2} \frac{|H_R(f)|^2}{|H_c(f)|^2}$$

che dunque subisce una maggiore amplificazione proprio alle frequenze per le quali $|H_c(f)|^2$ è particolarmente ridotta. Illustriamo dunque nel seguito alcune tecniche che invece tengono conto della presenza del rumore.

18.4.2 Equalizzatore MMSE e filtro di Wiener

Con questa tecnica si sceglie di realizzare il filtro FIR $H_e(f)$ in forma totalmente numerica (§ 4.6.1) e di campionare $y(t)$ con periodo $\tau = T_s/2$[18], ovvero facendo lavorare il filtro alla velocità di due campioni per simbolo. I coefficienti $\mathbf{c}$ del filtro FIR $H_e(f)$ di fig. 18.3 sono individuati come

$$\tilde{\mathbf{c}} = \arg\min \left\{ \sigma_e^2(\mathbf{c}) \right\}$$

ossia tali da rendere *minimo*[19] l'*errore quadratico medio* $\sigma_e^2 = E\left\{e_m^2\right\}$, in cui

$$e_m = z(mT_s) - a_m \tag{18.19}$$

è l'errore tra i campioni $z(mT_s)$ della (18.14) in uscita da $H_e(f)$, ed i corrispondenti valori a_m della sequenza informativa[20] che è stata trasmessa. Tale criterio è noto come MMSE (MINIMUM MEAN SQUARE ERROR o *minimo errore quadratico medio*), e determina un risultato noto anche come *filtro di* WIENER[21]. In presenza di processi ergodici a media nulla l'*errore quadratico medio* σ_e^2 corrisponde alla *potenza* dell'errore, che (applicando la (18.14) alla (18.19)) può essere espressa come

$$\sigma_e^2 = E\left\{ (z(mT_s) - a_m)^2 \right\} = E\left\{ \left(\sum_{n=-N}^{N} c_n y(mT_s - n\tau) - a_m \right)^2 \right\} \tag{18.20}$$

evidenziandone la dipendenza dai coefficienti c_n. La minimizzazione di σ_e^2 si ottiene calcolando le sue derivate rispetto ai coefficienti c_n, ed eguagliando l'espressione delle stesse a zero[22], ovvero

$$\begin{aligned} \frac{\partial}{\partial c_k} \sigma_e^2 &= 2E\left\{ \left[\sum_{n=-N}^{N} c_n y(mT_s - n\tau) - a_m \right] y(mT_s - k\tau) \right\} = \\ &= 2\left(\sum_{n=-N}^{N} c_n E\left\{ y(mT_s - n\tau)\, y(mT_s - k\tau) \right\} - E\left\{ a_m y(mT_s - k\tau) \right\} \right) = 0 \end{aligned} \tag{18.21}$$

nella quale riconosciamo che i *valori attesi* che vi compaiono corrispondono rispettiva-

[18] Dato che per evitare aliasing (§ 4.1.1) la frequenza di campionamento $f_c = \frac{1}{\tau}$ deve risultare maggiore del doppio della massima frequenza presente in $y(t)$, scegliamo $f_c = 2f_s$, ovvero $\tau = T_s/2$. Infatti, un segnale dati di banda base a coseno rialzato occupa una banda a frequenze positive pari a $B = \frac{f_s}{2}(1+\gamma)$, e dunque scegliendo $f_c = 2f_s$ ci cauteliamo anche nel caso di $\gamma = 1$. La scelta di porre $\tau < T_s$ viene indicata anche con il termine di *fractionally spaced equalizer* o FSE.

[19] Piuttosto che *l'azzeramento* dell'errore, come nell'approccio *zero forcing*.

[20] Tale errore dipende sia dal rumore presente in ingresso che dall'ISI introdotta dal canale, le cui potenze sono quindi minimizzate in forma congiunta.

[21] Vedi `http://en.wikipedia.org/wiki/Wiener_filter`.

[22] In realtà la minimizzazione ha successo solamente se σ_e^2 è una funzione *convessa* (vedi `https://it.wikipedia.org/wiki/Funzione_convessa`) dei c_n, ma più avanti (pag. 625) troveremo che questo è proprio il nostro caso.

mente all'*autocorrelazione* (§ 7.1.4)

$$\mathcal{R}_Y(n-k) = E\{y(mT_s - n\tau)\, y(mT_s - k\tau)\} \tag{18.22}$$

tra coppie di campioni del segnale ricevuto distanti $|n-k|\,\tau$, ed all'*intercorrelazione*

$$\mathcal{R}_{YA}(k) = E\{y(mT_s - k\tau)\, a_m\} \tag{18.23}$$

tra i campioni di $y(t)$ distanti $|k|\,\tau$ dal generico simbolo a_m ed il valore del simbolo stesso, dove il valore atteso è calcolato rispetto alla variabilità della sequenza a_m e del rumore. Dato che entrambi gli indici n e k variano tra $-N$ e N, le eq. (18.21) si riscrivono come

$$\sum_{n=-N}^{N} c_n \mathcal{R}_Y(n-k) = \mathcal{R}_{YA}(k) \qquad k = -N, -N+1, \cdots, 0, \cdots, N-1, N \tag{18.24}$$

note come equazioni di *Wiener-Hopf*, ovvero un sistema di $2N+1$ equazioni nelle altrettante incognite c_n, la cui matrice dei coefficienti e vettore dei termini noti sono costituiti rispettivamente dalle (18.22) e (18.23).

Soluzione delle equazioni di Wiener-Hopf Adottiamo innanzitutto per la (18.24) una notazione matriciale del tipo

$$\mathbf{Bc} = \mathbf{d} \tag{18.25}$$

dove $\mathbf{B} = [b_{nk}] = [\mathcal{R}_Y(n-k)]$ è una matrice $(2N+1)\times(2N+1)$, $\mathbf{c} = [c_n]$ è un vettore colonna i cui elementi sono i coefficienti del filtro, e $\mathbf{d}$ è il vettore dei termini noti $\mathbf{d} = [d_k] = [\mathcal{R}_{YA}(k)]$[23]. Con tale notazione i coefficienti *ottimi* $\tilde{\mathbf{c}}$ si ottengono come

$$\tilde{\mathbf{c}} = \mathbf{B}^{-1}\mathbf{d}$$

Osserviamo che la *parità* di $\mathcal{R}_Y(n)$ rispetto ad n (vedi schema a lato) rende $\mathbf{B}$ una matrice di *Toepliz*[24], per la quale esistono metodi che semplificano in parte la soluzione diretta del sistema di equazioni, vedi anche la nota 14 a pag. 293.

$$\mathbf{B} = \begin{bmatrix} \mathcal{R}_Y(0) & \mathcal{R}_Y(1) & \mathcal{R}_Y(2) & \cdots & \mathcal{R}_Y(2N) \\ \mathcal{R}_Y(1) & \mathcal{R}_Y(0) & \mathcal{R}_Y(1) & \cdots & \mathcal{R}_Y(2N-1) \\ \mathcal{R}_Y(2) & \mathcal{R}_Y(1) & \mathcal{R}_Y(0) & \cdots & \mathcal{R}_Y(2N-2) \\ \mathcal{R}_Y(3) & \mathcal{R}_Y(2) & \mathcal{R}_Y(1) & \cdots & \mathcal{R}_Y(2N-3) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathcal{R}_Y(2N) & \mathcal{R}_Y(2N-1) & \cdots & \cdots & \mathcal{R}_Y(0) \end{bmatrix}$$

Risposta in frequenza del filtro di Wiener Affrontiamo una prima digressione per derivare l'espressione assunta da $H_e(f)$ nel caso ottimo, ovvero quando $\mathbf{c} = \tilde{\mathbf{c}}$ e σ_e^2 è minimo. Conviene procedere direttamente nel dominio delle sequenze, ed osservare che la risposta impulsiva campionata (con periodo τ) del filtro FIR $H_e(f)$ di

[23] In effetti, $\mathcal{R}_Y(n)$ e $\mathcal{R}_{YA}(n)$ *non sono note* al ricevitore, ma nell'ipotesi di stazionarietà ed ergodicità per $y(t)$ ed a_n, possono essere *stimate* a partire da una fase iniziale *di apprendimento* durante la quale viene trasmesso un segnale di test associato ad una sequenza $\{a_k\}$ nota al ricevitore, in base alla quale quest'ultimo calcola le medie temporali

$$\hat{\mathcal{R}}_Y(n) = \tfrac{1}{K}\textstyle\sum_{k=1}^{K} y(kT_s - n\tau)\, y(kT_s) \qquad \text{e} \qquad \hat{\mathcal{R}}_{YA}(n) = \tfrac{1}{K}\textstyle\sum_{k=1}^{K} y(kT_s - n\tau)\, a_k$$

che quindi utilizza al posto delle medie di insieme nella (18.24).

[24] `http://en.wikipedia.org/wiki/Toeplitz_matrix`

fig. 18.3 è esattamente pari ai coefficienti **c**, ovvero $h_e(n) = c_{n-N}$ con $n = 0, 1, 2 \cdots, 2N$. Pertanto la (18.24) può essere letta come l'espressione della sequenza $\mathcal{R}_{YA}(k)$ (per $k = 0, \pm 1, \pm 2, \cdots, \pm N$) nei termini di una *convoluzione discreta* tra $h_e(n)$ ed $\mathcal{R}_Y(n)$ anticipata di N:

$$\mathcal{R}_{YA}(k) = \sum_{n=0}^{2N} h_e(n)\, \mathcal{R}_Y(k - n + N) \tag{18.26}$$

Effettuando la DTFT (§ 4.4) di entrambi i membri di (18.26) si ottiene[25] $\mathcal{P}_{YA}(f) = H_e(f)\, \mathcal{P}_Y(f)\, e^{j2\pi f N\tau}$ in cui $\mathcal{P}_Y(f)$ è la densità di potenza della sequenza $y(n)$ in ingresso ad $h_e(t)$, e $\mathcal{P}_{YA}(f)$ è la densità di potenza *mutua*[26] tra $y(n)$ e la sequenza informativa $a(n)$. Otteniamo pertanto

$$H_e(f) = \frac{\mathcal{P}_{YA}(f)}{\mathcal{P}_Y(f)} \mathrm{e}^{-j2\pi f N\tau} \tag{18.27}$$

dove il termine $\mathrm{e}^{-j2\pi f N\tau}$ esprime il ritardo necessario a realizzare un filtro causale, e viene trascurato nel seguito. Le densità di potenza $\mathcal{P}_Y(f)$ e $\mathcal{P}_{YA}(f)$ possono essere valutate eseguendo la DTFT delle corrispondenti correlazioni che figurano nella (18.26), che (per simboli a_m *incorrelati*) risultano[27]

$$\mathcal{R}_Y(n) = \sigma_a^2 \mathcal{R}_H(n) + \mathcal{R}_\nu(n) \quad \mathrm{e}^{28} \quad \mathcal{R}_{YA}(n) = \sigma_a^2 h(-n)$$

in cui $\mathcal{R}_H(n)$ è l'autocorrelazione dei campioni (con periodo τ) delle risposta impulsiva complessiva $h(t) = h_T(t) * h_c(t) * h_R(t)$, $\mathcal{R}_\nu(n)$ è la correlazione tra campioni di $\nu(t)$ a distanza $n\tau$, e $\sigma_a^2 = E\left\{a_k^2\right\}$, sottintendendo gli a_k a media nulla. La (18.27) si riscrive quindi come

$$H_e(f) = \frac{\sigma_a^2 H^*(f)}{\sigma_a^2 |H(f)|^2 + \mathcal{P}_\nu(f)} \tag{18.28}$$

[25] Viene applicata l'equivalenza tra convoluzione nel tempo e prodotto in frequenza, nonché quella tra trasformata della correlazione e densità di potenza.

[26] Analogamente all'energia mutua (§ 3.2), $\mathcal{P}_{YA}(f)$ esprime una *similitudine* tra due segnali e/o processi, frequenza per frequenza. Laddove risulti $\mathcal{P}_{YA}(f) = 0$, i segnali sono *ortogonali* in tale regione di frequenze.

[27] Ricordando che $y(t) = \sum_k a_k h(t - kT_s) + \nu(t)$ in cui $h(t)$ è la risposta impulsiva complessiva $h(t) = h_T(t) * h_c(t) * h_R(t)$ (vedi fig. 18.2) e $\nu(t)$ rappresenta l'effetto di $H_R(f)$ su $n(t)$, dalla (18.22) scriviamo $\mathcal{R}_Y(n) = E\{y(mT_s)\, y(mT_s + n\tau)\}$ e quindi esprimendo $h(mT_s - kT_s)$ come $h((m-k)T_s)$ si ottiene

$$\begin{aligned}\mathcal{R}_Y(n) = \quad & E\{(\textstyle\sum_k a_k h((m-k)T_s) + \nu(mT_s))\,(\sum_i a_h h((m-i)T_s + n\tau) + \nu(mT_s + n\tau))\} = \\ & E\{\textstyle\sum_k \sum_i a_k a_i h((m-k)T_s)\, h((m-i)T_s + n\tau) + \nu(mT_s)\,\nu(mT_s + n\tau) + \\ & + \nu(mT_s + n\tau) \textstyle\sum_k a_k h((m-k)T_s) + \nu(mT_s) \sum_h a_h h((m-i)T_s + n\tau)\}\end{aligned}$$

i cui termini dell'ultima riga risultano nulli se i simboli a_m sono statisticamente indipendenti dai campioni di rumore, ed almeno uno dei due processi è a media nulla. Essendo inoltre $E\{a_k a_i\} = 0$ per $i \neq k$, il termine con la doppia sommatoria si riduce a

$$E\left\{a_k^2\right\} \textstyle\sum_k h((m-k)T_s)\, h((m-k)T_s + n\tau) = \mathcal{R}_A(0)\, \mathcal{R}_H(n) = \sigma_a^2 \mathcal{R}_H(n)$$

avendo considerato gli a_k a media nulla. Infine, il termine $E\{\nu(mT_s)\,\nu(mT_s + n\tau)\}$ risulta pari alla correlazione $\mathcal{R}_{\nu\nu}(n)$ del processo di rumore.

[28] Dalla (18.23) scriviamo $\mathcal{R}_{YA}(n) = E\{y(mT_s - n\tau)\, a_m\}$ e quindi

$$\begin{aligned}\mathcal{R}_{YA}(n) = \quad & E\{(\textstyle\sum_k a_k h((m-k)T_s - n\tau) + \nu(mT_s - n\tau))\, a_m\} = \\ = \quad & \textstyle\sum_k E\{a_k a_m\}\, h((m-k)T_s - n\tau) + E\{\nu(mT_s - n\tau)\, a_m\}\end{aligned}$$

e come prima troviamo $E\{\nu(mT_s - n\tau)\, a_m\} = 0$, mentre della sommatoria rimane il solo termine $k = m$.

Possiamo osservare che qualora $\mathcal{P}_\nu(f) = 0$ si ottiene $H_e(f) = \frac{1}{H(f)}$, che pertanto assolve in pieno il compito di equalizzazione, come avviene per l'approccio *zero forcing* (§ 18.4.1). D'altra parte, dividendo numeratore e denominatore di $H_e(f)$ per $\sigma_a^2 |H(f)|^2$ si ottiene

$$H_e(f) = \frac{1}{H(f)} \frac{1}{1 + \frac{1}{SNR(f)}}$$

che evidenzia come l'effetto del rumore sia quello *attenuare* la risposta in frequenza $H_e(f)$ nelle regioni in cui $SNR(f)$ (non in dB) è più piccolo, ossia dove c'è meno segnale, e/o più rumore. Infine, nel caso di rumore elevato la (18.28) diviene $H_e(f) = H^*(f) \frac{\sigma_a^2}{\mathcal{P}_\nu(f)}$ (o, in caso di rumore *bianco*, $H_e(f) = H^*(f) \frac{\sigma_a^2}{\sigma_\upsilon^2}$) come per un *filtro adattato* (§ 7.6); essendo il termine $H^*(f)$ comunque presente, si può scegliere di realizzare $H_R(f)$ (vedi fig. 18.2) con il solo scopo di limitare la potenza di rumore[29].

Qualora i simboli informativi a_m siano tra loro *correlati* la (18.28) diviene[30]

$$H_e(f) = \frac{\mathcal{P}_A(f) H^*(f)}{\mathcal{P}_A(f) |H(f)|^2 + \mathcal{P}_\nu(f)}$$

in cui $\mathcal{P}_A(f)$ è la DTFT della correlazione dei simboli $\mathcal{R}_A(n) = E\{a_m a_{m+n}\}$ o *spettro del codice* (pag. 226). La stessa espressione è valida per un filtro di Wiener di tipo più generale, che esegue ad esempio la *deriverberazione* di un segnale audio con densità di potenza $\mathcal{P}_A(f)$.

Analisi della potenza di errore Sviluppiamo ora una seconda digressione sul significato *geometrico* della minimizzazione di σ_e^2, che può essere considerato come una funzione dei coefficienti $\mathbf{c}$, e di cui si cerca il minimo (ovvero $\tilde{\mathbf{c}} = \arg\min\left\{\sigma_e^2(\mathbf{c})\right\}$), che mostriamo essere unico. Infatti, l'espressione di $\sigma_e^2(\mathbf{c}) = E\left\{e_m^2\right\}$ si ottiene dalla (18.20) come

$$\begin{aligned}\sigma_e^2(\mathbf{c}) &= E\left\{\left(\sum_{n=-N}^{N} c_n y(mT_s - n\tau) - a_m\right)\left(\sum_{k=-N}^{N} c_k y(mT_s - k\tau) - a_m\right)\right\} = \\ &= \sum_{n=-N}^{N}\sum_{k=-N}^{N} c_n c_k \mathcal{R}_Y(n-k) - 2\sum_{k=-N}^{N} c_k \mathcal{R}_{YA}(k) + E\left\{a_m^2\right\}\end{aligned} \tag{18.29}$$

[29] In particolare, scegliendo $H_R(f) = rect_{2f_s}(f)$ si ottiene che i campioni di rumore presi con intervallo $\tau = 1/2T_s$ sono incorrelati e dunque statisticamente indipendenti perché gaussiani. Infatti $\mathcal{P}_\nu(f) = \frac{N_0}{2}|H_R(f)|^2$ e dunque $\mathcal{R}_\nu(\tau) = \frac{N_0}{2} 2f_s sinc(2f_s\tau)$, che quindi si azzera per $\tau = 1/2T_s$, vedi § 7.2.4. In tal caso la (18.28) diviene $H_e(f) = \frac{\sigma_a^2 H^*(f)}{\sigma_a^2 |H(f)|^2 + \sigma_\upsilon^2}$.

[30] Per quanto il risultato sembri banale, la dimostrazione non è troppo diretta, e si basa sullo scomporre il termine

$$E\{\sum_k \sum_i a_k a_i h((m-k)T_s)\, h((m-i)T_s + n\tau)\} = \sum_k \sum_i \mathcal{R}_A(i-k)\, h((m-k)T_s)\, h((m-i)T_s + n\tau)$$

tenendo conto della stazionarietà di a_m e dunque della simmetria di $\mathcal{R}_A(n)$, in una somma di termini

$$\sum_p \mathcal{R}_A(p) \sum_i h((m-i)T_s)\, h((m-i-p)T_s + n\tau) = \sum_p \mathcal{R}_A(p)\, \mathcal{R}_H(n+2p)$$

in cui si tiene conto che $T_s = 2\tau$, mentre gli a_m sono prodotti uno per simbolo; introducendo $\mathcal{R}'_A(n) = \begin{cases}\mathcal{R}_A\left(\frac{n}{2}\right) & n\ pari \\ 0 & altrimenti\end{cases}$ e con qualche cambio di variabile si ottiene infine $\sum_q \mathcal{R}'_A(q)\, \mathcal{R}_H(n-q)$ che ha l'aspetto rassicurante della convoluzione, da cui scaturisce il risultato.

che è un *polinomio completo* di secondo grado nelle $2N+1$ variabili c_n, ed essendo $[\mathcal{R}_Y(n-k)]$ una matrice *definita positiva*, $\sigma_e^2(\mathbf{c})$ presenta *un unico* minimo globale[31], e la sua espressione può essere riscritta in forma matriciale come[32]

$$\sigma_e^2(\mathbf{c}) = \mathbf{c}^t\mathbf{B}\mathbf{c} - 2\mathbf{c}^t\mathbf{d} + \sigma_a^2 \tag{18.30}$$

In corrispondenza della soluzione $\tilde{\mathbf{c}} = \mathbf{B}^{-1}\mathbf{d}$ della (18.25) si ottiene la *minima* varianza di errore, pari a[33]

$$\sigma_{e_{min}}^2 = \sigma_e^2(\tilde{\mathbf{c}}) = \sigma_a^2 - \mathbf{d}^t\mathbf{B}^{-1}\mathbf{d} = \sigma_a^2 - \tilde{\mathbf{c}}^t\mathbf{B}\tilde{\mathbf{c}} = \sigma_a^2 - \sigma_z^2 \tag{18.31}$$

avendo notato che il termine $\sigma_z^2 = \tilde{\mathbf{c}}^t\mathbf{B}\tilde{\mathbf{c}}$ corrisponde[34] alla potenza dell'uscita $z = z_m$, ovvero $\sigma_z^2 = E\left\{z^2\right\}$. Accade ora che la (18.30) può essere espressa anche come[35] $\sigma_e^2(\mathbf{c}) = (\mathbf{B}\mathbf{c}-\mathbf{d})^t\mathbf{B}^{-1}(\mathbf{B}\mathbf{c}-\mathbf{d}) - \mathbf{d}^t\mathbf{B}^{-1}\mathbf{d} + \sigma_a^2$, e tenendo conto della prima delle (18.31) si ottiene

$$\sigma_e^2(\mathbf{c}) = \sigma_{e_{min}}^2 + (\mathbf{B}\mathbf{c}-\mathbf{d})^t\mathbf{B}^{-1}(\mathbf{B}\mathbf{c}-\mathbf{d}) \tag{18.32}$$

evidentemente pari a $\sigma_{e_{min}}^2$ qualora $\mathbf{c} = \tilde{\mathbf{c}}$ e dunque $\mathbf{B}\tilde{\mathbf{c}} = \mathbf{d}$. Essendo $\mathbf{B}$ una matrice definita positiva lo è anche la sua inversa $\mathbf{B}^{-1}$, ed esiste (vedi § 6.7.3) una matrice unitaria $\mathbf{\Gamma}$ con colonne pari agli autovettori di $\mathbf{B}^{-1}$ tale da poter scrivere $\mathbf{B}^{-1} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}^t$, dove $\mathbf{\Lambda}$ è una matrice diagonale con elementi pari all'inverso degli autovalori λ_i di $\mathbf{B}$; indicando allora con $\mathbf{c}_\delta$ il vettore $\mathbf{B}\mathbf{c}-\mathbf{d} = \mathbf{B}(\mathbf{c}-\tilde{\mathbf{c}})$ legato alla *differenza* tra i coefficienti generici e quelli ottimi, e con $\mathbf{u} = \mathbf{\Gamma}^t\mathbf{c}_\delta$ la sua versione trasformata nella base individuata dagli autovettori, la (18.32) diviene

$$\sigma_e^2(\mathbf{u}) = \sigma_{e_{min}}^2 + \mathbf{u}^t\mathbf{\Lambda}\mathbf{u} = \sigma_{e_{min}}^2 + \sum_{i=0}^{2N}\frac{u_i^2}{\lambda_i}$$

che rappresenta l'equazione di un iperparaboloide ellittico, con assi principali pari agli autovettori di $\mathbf{B}^{-1}$.

---

[31] Ciò deriva dal fatto che $[\mathcal{R}_Y(n-k)] = \mathbf{B}$ è legata alla *matrice di covarianza* $\mathbf{\Sigma}_Y$ (§ 6.7.3) dalla relazione $[\mathcal{R}_Y(n-k)] = \mathbf{\Sigma}_Y + m_Y^2$ (vedi eq. (7.3) a pag. 192 in condizioni stazionarie), e dato che $\mathbf{\Sigma}_Y$ è una *matrice definita positiva* lo è anche $[\mathcal{R}_Y(n-k)]$, non decadendo la proprietà in seguito alla somma per una quantità positiva (m_Y^2). Pertanto la *forma quadratica* $\sum_{n=-N}^{N}\sum_{k=-N}^{N} c_n c_k \mathcal{R}_Y(n-k)$ è ovunque convessa (vedi § 6.7.3), e dotata di un unico minimo globale. Questa proprietà permane anche considerando gli altri due termini della (18.29) che concorrono al valore di $\sigma_e^2(\mathbf{c})$, essendo questi rispettivamente lineari in $\mathbf{c}$ e costanti, e che dunque non ne modificano la convessità.

[32] Considerando la sequenza informativa a_m a media nulla, si ha $E\left\{a_m^2\right\} = \sigma_a^2$.

[33]
$$\begin{aligned}\sigma_e^2(\tilde{\mathbf{c}}) &= \tilde{\mathbf{c}}^t\mathbf{B}\tilde{\mathbf{c}} - 2\tilde{\mathbf{c}}^t\mathbf{d} + \sigma_a^2 = \tilde{\mathbf{c}}^t\mathbf{B}\tilde{\mathbf{c}} - 2\tilde{\mathbf{c}}^t\mathbf{B}\tilde{\mathbf{c}} + \sigma_a^2 = \\ &= \sigma_a^2 - \tilde{\mathbf{c}}^t\mathbf{B}\tilde{\mathbf{c}} = \sigma_a^2 - \tilde{\mathbf{c}}^t\mathbf{d} = \sigma_a^2 - \mathbf{d}^t\mathbf{B}^{-1}\mathbf{d}\end{aligned}$$

avendo prima sostituito $\mathbf{d} = \mathbf{B}\tilde{\mathbf{c}}$ e poi il contrario, e valutato $\tilde{\mathbf{c}}^t = \left(\mathbf{B}^{-1}\mathbf{d}\right)^t = \mathbf{d}^t\left(\mathbf{B}^{-1}\right)^t = \mathbf{d}^t\mathbf{B}^{-1}$ in quanto $\mathbf{B}$ è una matrice simmetrica, così come la sua inversa.

[34] Considerando z a media nulla, si ha $\sigma_z^2 = E\left\{z^2\right\} = E\left\{\tilde{\mathbf{c}}^t\mathbf{y}\mathbf{y}^t\tilde{\mathbf{c}}\right\} = \tilde{\mathbf{c}}^t E\left\{\mathbf{y}\mathbf{y}^t\right\}\tilde{\mathbf{c}} = \tilde{\mathbf{c}}^t\mathbf{B}\tilde{\mathbf{c}}$.

[35] Infatti

$$\begin{aligned}(\mathbf{B}\mathbf{c}-\mathbf{d})^t\mathbf{B}^{-1}(\mathbf{B}\mathbf{c}-\mathbf{d}) &= \mathbf{c}^t\mathbf{B}^t\mathbf{B}^{-1}\mathbf{B}\mathbf{c} - \mathbf{c}^t\mathbf{B}^t\mathbf{B}^{-1}\mathbf{d} - \mathbf{d}^t\mathbf{B}^{-1}\mathbf{B}\mathbf{c} + \mathbf{d}^t\mathbf{B}^{-1}\mathbf{d} = \\ &= \mathbf{c}^t\mathbf{B}\mathbf{c} - \mathbf{c}^t\mathbf{d} - \mathbf{d}^t\mathbf{c} + \mathbf{d}^t\mathbf{B}^{-1}\mathbf{d} = \mathbf{c}^t\mathbf{B}\mathbf{c} - 2\mathbf{c}^t\mathbf{d} + \mathbf{d}^t\mathbf{B}^{-1}\mathbf{d}\end{aligned}$$

18.4.3 Metodo del gradiente

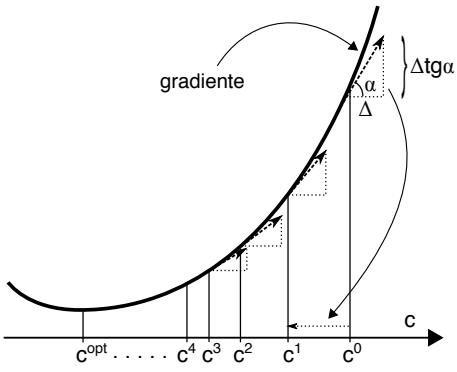
Benché si sia pervenuti alla (18.32) che esprime $\sigma_e^2(\mathbf{c})$ in funzione del vettore di coefficienti $\mathbf{c}$ di $H_e(f)$, la sua minimizzazione *diretta* comporta la soluzione del sistema (18.25). Ciò non è necessario adottando un metodo *iterativo* noto come *discesa del gradiente*[36] che si basa su di un approccio per così dire *altimetrico* e che considera $\sigma_e^2(\mathbf{c})$ come una (iper)superficie in uno spazio $2N+1$ dimensionale, e partendo da un vettore iniziale $\mathbf{c}^0$ di coefficienti scelti casualmente e dall'associata *quota* $\sigma_e^2(\mathbf{c}^0)$, ad ogni iterazione calcola nuovi coefficienti $\mathbf{c}^i$ tali che $\sigma_e^2(\mathbf{c}^{i+1}) \leq \sigma_e^2(\mathbf{c}^i)$, di fatto *scendendo* lungo il crinale[37] della superficie $\sigma_e^2(\mathbf{c})$.

A questo scopo ci si basa sulla conoscenza del *gradiente* di $\sigma_e^2(\mathbf{c})$ che indichiamo come $\mathbf{g}(\mathbf{c})$, cioè (pag. 9.6.1) del vettore le cui componenti g_n sono le derivate parziali di σ_e^2 rispetto ai coefficienti che compongono $\mathbf{c}$, ovvero $g_n = \frac{\partial \sigma_e^2(\mathbf{c})}{\partial c_n}$: esse sono tanto maggiori quanto più è ripida la pendenza di σ_e^2 nella direzione di c_n, e si azzerano tutte nel punto di minimo $\tilde{\mathbf{c}}$. Pertanto *l'orientamento* del vettore $\mathbf{g}(\mathring{\mathbf{c}})$ per ogni vettore $\mathring{\mathbf{c}} \neq \tilde{\mathbf{c}}$ indica *la direzione* verso la quale $\sigma_e^2(\mathring{\mathbf{c}})$ cresce di più.

Essendo anche il gradiente una funzione di $\mathbf{c}$ il suo valore può essere calcolato a partire dalla (18.30) ottenendo $\mathbf{g}(\mathbf{c}) = 2\mathbf{B}\mathbf{c} - 2\mathbf{d}$, che infatti si azzera per $\tilde{\mathbf{c}} = \mathbf{B}^{-1}\mathbf{d}$. Scegliendo casualmente un vettore iniziale di coefficienti $\mathbf{c}^0$, ad ogni iterazione i nuovi valori sono calcolati come

$$\mathbf{c}^{i+1} = \mathbf{c}^i - \Delta \mathbf{g}(\mathbf{c}^i) \tag{18.33}$$

ovvero *spostando* ogni coefficiente c_n in direzione *opposta* alla relativa componente del gradiente $\mathbf{g}^i = \mathbf{g}(\mathbf{c}^i)$ calcolato in quel punto, di una quantità ad esso proporzionata dal parametro Δ, quest'ultimo pari ad piccolo numero positivo che determina la velocità di convergenza dell'algoritmo[38]. Viene quindi calcolato un nuovo valore $\mathbf{g}^{i+1} = \mathbf{g}(\mathbf{c}^{i+1})$, ed il procedimento iterato fino a quando non si verifica che $|\mathbf{g}^i| < \varepsilon$, nel qual caso si ritiene raggiunto l'ottimo $\tilde{\mathbf{c}}$. La figura precedente rappresenta un esempio unidimensionale di applicazione del metodo, in cui il gradiente è pari alla derivata dy/dx, ovvero alla tangente dell'angolo α tra la funzione e l'asse orizzontale, e mostra i diversi valori c^i ottenuti nelle iterazioni.



18.4.3.1 Equalizzazione adattiva LMS

Il metodo ora esposto si basa sul calcolo ripetuto di $\mathbf{g}^i = 2\left(\mathbf{B}\mathbf{c}^i - \mathbf{d}\right)$, ma... se la $H(f)$ del canale varia nel tempo, lo stesso avviene per $\mathbf{B}$ e per $\mathbf{d}$, così come dovrebbe avvenire

[36] Vedi `http://it.wikipedia.org/wiki/Discesa_del_gradiente`

[37] Coma abbiamo discusso, il minimo è unico, dunque raggiungibile da dovunque si parta.

[38] Una regola pratica suggerisce di porre $\Delta = \frac{1}{5(2N+1)\mathcal{P}_Y}$ in cui $\mathcal{P}_Y$ è la potenza ricevuta di segnale più rumore. Valori maggiori possono accelerare la convergenza, ma dare anche luogo ad instabilità della soluzione, mentre valori più piccoli rallentano la convergenza, ma possono produrre errori finali minori.

per il valore dei coefficienti $\tilde{\mathbf{c}}$! Per fronteggiare la contingenza ed allo stesso tempo evitare di stimare $\mathcal{R}_Y(n)$ e $\mathcal{R}_{YA}(n)$ si può adottare un diverso metodo di discesa del gradiente, detto ora *stocastico* (o LMS[39]) perché opera a partire da una *stima* del gradiente, aggiornata ad ogni periodo di simbolo, così come ad ogni simbolo sono aggiornati i coefficienti $\mathbf{c}$, utilizzati *allo stesso tempo* per eseguire l'equalizzazione.

Tutto inizia considerando che in base alla (18.21) il gradiente può essere valutato (a meno di un fattore due) per i valori $\mathbf{c}_m$ utilizzati all'istante $t = mT_s$ anche come[40]

$$\mathbf{g}_m = E\{e_m \mathbf{y}_m\} \tag{18.34}$$

in cui $e_m = z(mT_s) - a_m$ è l'errore commesso dall'equalizzatore rispetto al simbolo trasmesso (e, nella fase di *apprendimento*, noto in ricezione[41]) ed $\mathbf{y}_m$ è il vettore dei $2N+1$ campioni $y_m(n)$ memorizzati *nello stato* del filtro $h_e(t)$ al momento di produrre l'uscita $z(mT_s)$. La (18.34) viene poi *approssimata* tralasciando di effettuare la media di insieme, ossia valutando la stima del gradiente come[42] $\hat{\mathbf{g}}_m = e_m \mathbf{y}_m$, da usare poi nella formula di aggiornamento dei coefficienti c_n derivata dalla (18.33):

$$\mathbf{c}_{m+1} = \mathbf{c}_m - \Delta e_m \mathbf{y}_m \tag{18.35}$$

La (18.35) è improntata ad un principio di *correzione dai propri errori*, ovvero i coefficienti ricevono *un colpetto* in su o in giù, in modo da ridurre l'errore commesso[43]. Il

[39] Ossia LEAST MEAN SQUARE, vedi anche `http://en.wikipedia.org/wiki/Least_mean_squares_filter`

[40] Un interessante modo di interpretare la (18.34) si basa sull'osservare che la minimizzazione di (18.20) ovvero l'azzeramento del gradiente (18.21) $\mathbf{g}_m = \mathbf{0}$ comporta che ciascun valore di errore e_m deve essere *ortogonale* (in senso statistico, vedi § 7.7.2) a tutte le osservazioni $y_{m-N}, \cdots, y_{m+N}$ da cui dipende la stima $z(mT)$, ossia le $2N+1$ coppie di sequenze devono essere *incorrelate*, e quindi prive di legami di tipo lineare.

[41] La sequenza dei simboli trasmessi *è nota* all'inizio della trasmissione, e spesso generata nella forma di una sequenza *pseudo-noise*. Durante la fase *di apprendimento* l'algoritmo viene lasciato iterare arrivando alla stima dei coefficienti $\mathbf{c}$; terminata questa fase ha inizio la trasmissione vera e propria, e l'equalizzatore commuta su di una modalità *dipendente dalle decisioni*, in cui l'errore e_m è valutato rispetto ai valori $\widetilde{a}_m$ emessi dal decisore: se la probabilità di errore per simbolo non è troppo elevata, la maggior parte delle volte la decisione è esatta, ed il metodo continua a funzionare correttamente.

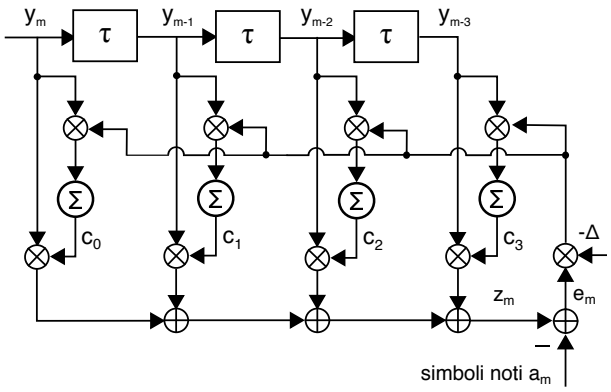
[42] Poniamo di eseguire l'aggiornamento (18.35) dei coefficienti $\mathbf{c}$ ogni K periodi di simbolo: otterremmo $\mathbf{c}_{m+K} = \mathbf{c}_m - \Delta \sum_{k=0}^{K-1} e_{m+k} \mathbf{y}_{m+k}$, in cui il termine sommatoria risulta in effetti proporzionale alla stima di $E\{e_m \mathbf{y}_m\}$, almeno più di quanto non appaia il termine $e_m \mathbf{y}_m$ che compare isolato nella (18.35).

[43] Possiamo infatti notare che alla $m-$esima iterazione della (18.35) il coefficiente c_n viene *aumentato* se e_m e y_n sono di segno opposto, e *diminuito* se concorde. Ovvero: se l'uscita è *piccola* allora $e_m = z(mT_s) - a_m < 0$, e il contributo c_n dell'n-esima memoria viene aumentato se $y_m(n)$ è positivo, o diminuito se $y_m(n)$ è negativo; viceversa, se $e_m = z(mT_s) - a_m > 0$ (uscita *troppo grande*) il valore di c_n viene aumentato se $y_m(n)$ è negativo, e diminuito nel caso opposto. Inoltre, se c_n e $y_m(n)$ hanno lo stesso segno contribuiscono a z_m come un termine positivo, oppure come un termine negativo se di segno diverso. Infine, la colonna δ esprime la variazione del contributo di $c_n y_m(n)$ a z_m, frutto della variazione di c_n e dei segni di c_n e di $y_m(n)$. Come si può verificare, δ è sempre tale da contribuire alla riduzione di e_m.

e_m	$y_m(n)$	c_n	c_n	$c_n y_m(n)$	δ
+	+	↓	+	+	↓
+	+	↓	−	−	↓
+	−	↑	+	−	↓
+	−	↑	−	+	↓
−	+	↑	+	+	↑
−	+	↑	−	−	↑
−	−	↓	+	−	↑
−	−	↓	−	+	↑

metodo del gradiente stocastico trova applicazione in molteplici ambiti, ivi comprese le reti neurali.

La figura a lato mostra uno schema di architettura per il filtro adattivo, che opera come equalizzatore con periodo di campionamento $\tau = {}^{T_s}/_{2}$, e ad ogni periodo di simbolo esegue l'aggiornamento dei propri coefficienti, in accordo alla (18.35), mediante l'operatore di *accumulo* simbolizzato dalla $\sum$ cerchiata.



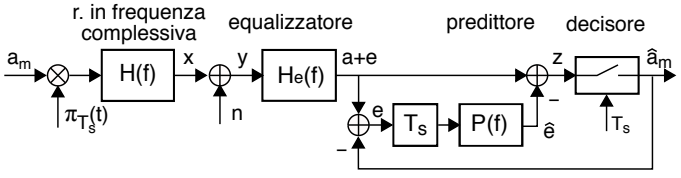
18.4.4 Equalizzatore a reazione

Questo approccio realizza una sorta di *equalizzazione dell'equalizzazione*, vediamo in che senso. Intanto osserviamo che i valori e_m che compongono la sequenza di errore $e_m = z(mT_s) - a_m$ (la differenza tra l'uscita z_m di $H_e(f)$ ed il simbolo trasmesso a_m) sono tra loro *correlati*, dato che

- il rumore bianco $n(t)$ (da cui $z(mT_s)$ anche dipende) si è *colorato* attraversando $h_R(t) * h_e(t)$;
- se l'ISI non è completamente rimossa significa che $H_e(f) \neq {}^{1}/_{H(f)}$ e quindi non si verificano le condizioni di Nyquist, ed i simboli *ricevuti* a_m divengono statisticamente dipendenti.

La tecnica del *decision feedback equalizer* o DFE sfrutta questa correlazione per *predire* i valori e_m a partire dagli errori agli istanti di simbolo precedenti, mediante un secondo filtro FIR che stima $\hat{e}_m = \sum_{n=1}^{N} p_n e_{m-n}$, e contribuisce ad una grandezza di decisione $z_m = a_m + e_m - \hat{e}_m$ affetta da un errore ancora più piccolo.

Il filtro di predizione è indicato come $P(f)$ nello schema di figura, nel quale si suppone assenza di errori *di decisione*, ovvero $\hat{a}_m = a_m$, ed i coefficienti p_n *ottimizzati* applicando ad es. il criterio MMSE all'errore sull'errore, ossia $E\left\{(e_m - \hat{e}_m)^2\right\} = \min$. Ma lo stesso schema può essere *ricablato* come appare in fig. 18.5-a), in cui $P(f)$ ed il sommatore in uscita sono *sdoppiati*: ciò consente di riunificare $H_e(f)$ e $P(f)$ in un unico *filtro in avanti* o di *smoothing*, la cui uscita viene ora fatta dipendere dal segnale ricevuto nei soli istanti di simbolo *passati*[44].



Lo schema di equalizzazione diviene quindi quello di fig. 18.5-b), in cui il filtro all'indietro *o di controreazione* è indicato come *di feedback,* opera alla frequenza di

[44] Ovvero il filtro FIR ha solo i ritardi relativi ai coefficienti c_n con $n \geq 0$, così come la sommatoria (18.14).

simbolo f_s, e la sua uscita $\hat{e}_m$ dipende solamente dalle decisioni precedenti $\hat{a}_m$[45], una cui combinazione lineare è usata per *correggere* i valori emessi dal primo filtro: la grandezza in ingresso al decisore può quindi essere ora espressa come

$$z_m = \sum_{n=0}^{N_1} c_n y\left(mT_s - n\tau\right) - \sum_{n=1}^{N_2} d_n \hat{a}_{m-n} \tag{18.36}$$

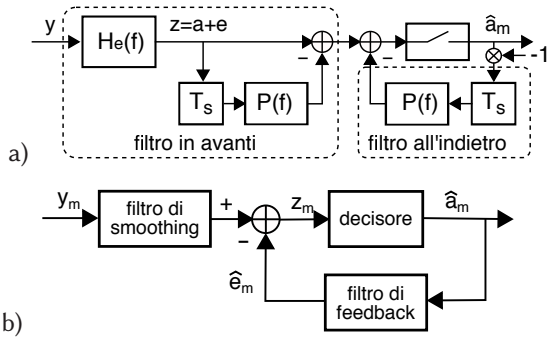


Figura 18.5: Schema simbolico di un decision feedback equalizer

I coefficienti d_n del filtro di feedback possono essere anch'essi stimati in maniera adattiva mediante il metodo del gradiente stocastico, dando così luogo alla architettura computazionale di figura 18.6, in cui la grandezza che *guida* l'aggiornamento dei d_n è la stessa misura dell'errore e_m usata per l'aggiornamento dei c_n del filtro in avanti, portando ad adattare la (18.35) come $\mathbf{d}_{m+1} = \mathbf{d}_m - \Delta e_m \mathbf{a}_m$, in cui $\mathbf{a}_m$ è il vettore dei simboli precedentemente decisi $a_{m-1}, a_{m-2}, \cdots, a_{m-N_2}$ e presenti nella memoria del filtro di feedback.

Dato che il filtro di feedback opera sui simboli già decisi e per questo privi di rumore, non introduce nuovo rumore, che anzi è in generale ridotto grazie alla minore lunghezza del filtro in avanti. D'altra parte la presenza del decisore nell'anello di reazione rende il metodo *non lineare*, e per questo l'effetto *moltiplicativo* di eventuali errori di decisione non può essere analizzato dal punto di vista teorico, ma solo essere studiato mediante simulazione.

[45] Anche qui, nella fase di apprendimento sono usati i valori a_m noti a priori, mentre in quella successiva si usano quelli $\tilde{a}_m$ emessi dal decisore, supposti per la maggior parte esatti.

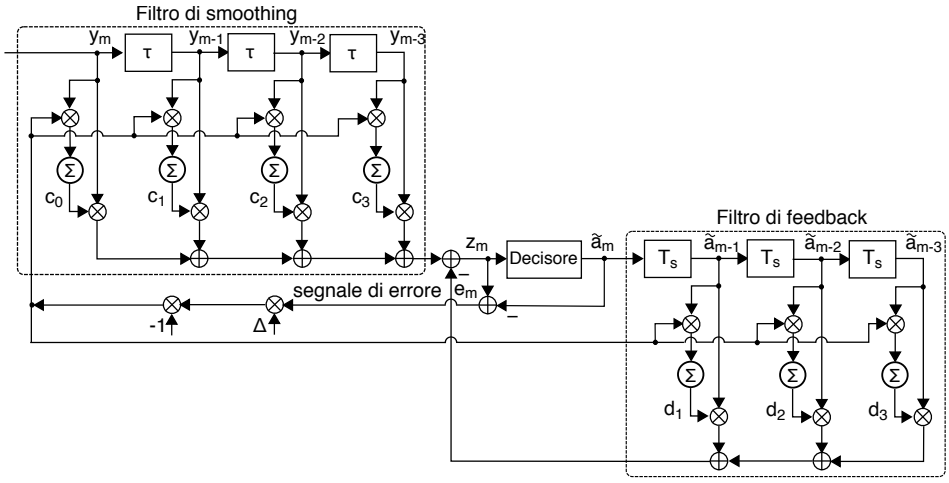


Figura 18.6: Architettura di un equalizzatore DFE basato sul gradiente stocastico

18.4.5 Equalizzazione come sequenza di massima verosimiglianza

Le tecniche finora illustrate effettuano la decisione simbolo per simbolo, mentre il metodo che stiamo per analizzare produce una stima di quanto ricevuto solo dopo aver preso in esame una *intera sequenza* di N simboli, operando una *maximum likelihood sequence detection* o MLSD. In questo caso non viene utilizzato nessun filtro aggiuntivo, e la funzione di equalizzazione è assolta dal decisore che opera in base all'algoritmo *di Viterbi* (pag. 577), con costi dei percorsi calcolati in modo simile a quelli del *soft decision decoding* di pag. 581.

Per illustrare il funzionamento di MLSD iniziamo con il *riconsiderare* i campioni $y_m = y(mT_s)$ del segnale ricevuto (18.13) come

$$y_m = \sum_{n=0}^{N} a_n h(mT_s - nT_s) + \nu_m = h_0 a_m + \sum_{n=0,\, n\neq m}^{N} a_n h_{m-n} + \nu_m \tag{18.37}$$

in cui $h(t) = h_T(t) * h_c(t) * h_R(t)$, evidenziando così le componenti di errore dovuto all'ISI (la $\sum_{n\neq m}$) ed al rumore ν_m. Il problema viene quindi impostato come quello della stima della sequenza $\mathbf{a} = \{a_m\}$ a partire dalla conoscenza della sola sequenza di osservazione $\mathbf{y} = \{y_m\}$, e per il quale la soluzione *di massima verosimiglianza* individua $\hat{\mathbf{a}}$ in modo tale che $p(\mathbf{y}/\mathbf{a})|_{\mathbf{a}=\hat{\mathbf{a}}}$ è massima.

La (18.37) evidenzia inoltre come la componente aleatoria di $\mathbf{y}$ sia tutta riconducibile a quella della sequenza di rumore ν, dato che l'effetto del canale si riduce ad una *convoluzione discreta* e dunque è deterministico; pertanto y_m è una v.a. con media $\bar{y}_m(\mathbf{a}) = \sum_{n=0}^{N} a_n h_{m-n}$ e d.d.p. gaussiana $p_\nu(y_m/\mathbf{a})$ condizionata ad $\mathbf{a}$, che ne determina la sequenza dei valori medi. Se le realizzazioni ν_m sono staticamente indipendenti[46] la $p(\mathbf{y}/\mathbf{a})$ è il prodotto dei singoli contributi $p_\nu(y_m/\mathbf{a})$, da valutare per tutti gli $_m$ e per tutti i modi di scegliere $\mathbf{a}$, e quindi trovare $\hat{\mathbf{a}} = \arg\max \prod_{m=0}^{M} p_\nu(y_m/\mathbf{a})$ o, equivalentemente

$$\hat{\mathbf{a}} = \arg\min \sum_{m=0}^{M} -\ln(p_\nu(y_m/\mathbf{a})) \tag{18.38}$$

Quest'ultimo passaggio trova motivazione nella espressione della d.d.p. gaussiana $p_\nu(y) = \frac{1}{\sqrt{2\pi}\sigma_\nu} \exp\left\{-\frac{(y-\bar{y})^2}{2\sigma_\nu^2}\right\}$ per cui $\ln(p_\nu(y)) = -\ln\left(\sqrt{2\pi}\sigma_\nu\right) - \frac{(y-\bar{y})^2}{2\sigma_\nu^2}$, e dato che σ_ν è la stessa per tutti gli istanti m, il suo valore non contribuisce alla minimizzazione (18.38), che diviene pertanto

$$\hat{\mathbf{a}} = \arg\min \sum_{m=0}^{M} (y_m - \bar{y}_m(\mathbf{a}))^2$$

La sequenza ottima $\hat{\mathbf{a}}$ è quindi quella che minimizza la somma degli scarti quadratici tra le osservazioni y_m ed il corrispondente valore atteso, e viene trovata utilizzando l'algoritmo di *Viterbi* (pag. 577) riconducendo il problema a quello di individuare un *percorso di minimo costo* nell'ambito del *traliccio* che rappresenta tutte le possibili sequenze $\mathbf{a}$.

[46] A questo proposito, è opportuno che $H_R(f)$ sia realizzato come un passa basso e non come un filtro adattato, o nel caso di rumore colorato in ingresso, che si effettui una operazione di *sbiancamento*.

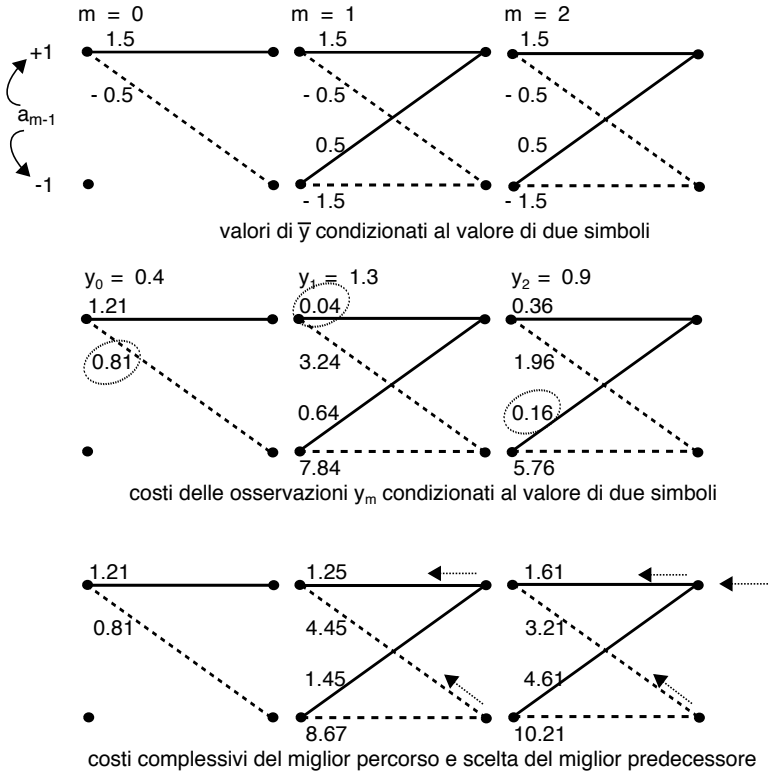


Figura 18.7: Costruzione del traliccio per un decodificatore MLSD

Esempio Ipotizziamo una $h_n = \begin{cases} 1 & con\ n = 0 \\ 0.5 & con\ n = 1 \end{cases}$ ed una trasmissione antipodale binaria con $a = \{1, -1\}$; la (18.37) si scrive quindi come

$a_m a_{m-1}$	$\bar{y}_m$
1,1	1.5
1,-1	0.5
-1,1	-0.5
-1,-1	-1.5

$$y_m = a_m h_0 + a_{m-1} h_1 + \nu_m$$

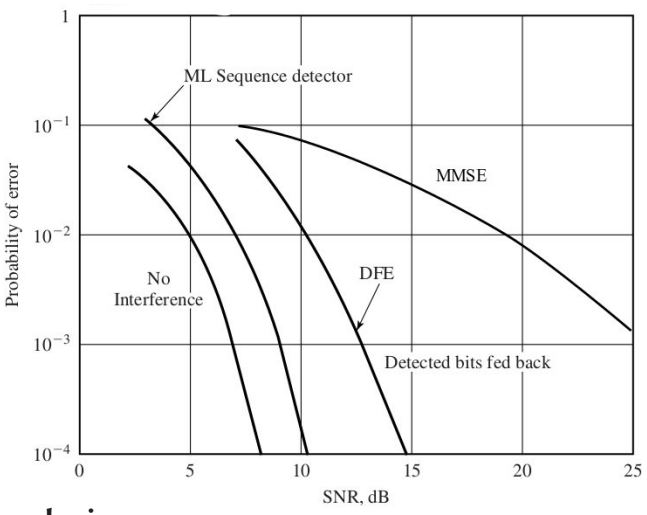
che rappresenta un filtro FIR di primo ordine, la cui uscita in assenza di rumore assume uno dei 4 possibili valori mostrati in tabella in funzione di $a_m a_{m-1}$. Il traliccio di ricerca può quindi essere tracciato come appare in alto in fig. 18.7, identificando i valori di a_{m-1} con le righe +1 e −1, e quelli di a_m con il tratteggio (−1) o la linea continua (1) degli archi che uniscono le colonne, considerando $a_{-1} = a_{-2} = 1$; sopra ogni arco è quindi scritto il valore di $\bar{y}_m$ corrispondente. Nella parte centrale della figura gli archi sono invece etichettati con i costi $(y_m - \bar{y}_m(\mathbf{a}))^2$ associati all'osservazione della sequenza $\mathbf{y} = 0.4, 1.3, 0.9$, e sono evidenziati i costi *minimi* simbolo per simbolo, che porterebbero a decidere per $-1, 1, 1$. Al contrario, la figura in basso mostra il costo totale del *miglior percorso* che raggiunge le righe del traliccio per ogni istante di simbolo, assieme alle frecce che indicano il miglior predecessore. L'applicazione del criterio MLSD determina quindi la decisione per $1, 1, 1$, con un bit di differenza.

Il metodo con cui l'algoritmo di Viterbi esplora il traliccio per individuare il percorso di minimo costo, e l'associata sequenza $\hat{\mathbf{a}}$, è illustrato al § 17.4.2.1 e non viene qui ripetuto.

18.4.6 Confronto delle prestazioni di equalizzazione

La figura che segue[47] confronta P_e vs. SNR per una modulazione binaria antipodale nei casi di assenza di ISI oppure di un canale con $h_0 h_1 h_2 = 0.4, 0.8, 0.4$, i cui effetti sono compensati mediante equalizzazione realizzata in base alle tecniche fin qui illustrate.

Come si vede il MLSD si comporta meglio di tutti, e determina il valore delle *prestazioni ottime*. Al contrario, le prestazioni di MMSE appaiono come le peggiori, essenzialmente a causa della amplificazione del rumore nelle bande in cui il canale ha guadagno nullo[48]; d'altro canto, il DFE ne migliora sensibilmente le prestazioni, pur non raggiungendo quelle di MLSD. Ciò che rende poco diffuso l'MLSD però è l'elevata complessità computazionale, oltre al ritardo da attendere per una decisione su intere sequenze, nonché le difficoltà a rendere adattativo l'algoritmo nel caso di condizioni non stazionarie.



18.4.7 Considerazioni conclusive

A volte l'operazione di equalizzazione non si svolge mediante un filtro, ma in un modo peculiare a seconda del tipo di modulazione. Abbiamo osservato al § 16.8.7 come nel caso dell'OFDM sia sufficiente moltiplicare il simbolo $\underline{a}_n$ ricevuto su ogni portante per l'inverso della risposta in frequenza $\underline{H}_n$ del canale equivalente passa basso, ovvero $\tilde{\underline{a}}_n = \underline{a}_n \cdot \frac{1}{\underline{H}_n}$. Nel caso di una trasmissione a spettro espanso, la stima delle caratteristiche dei cammini multipli viene usata da un ricevitore Rake (§ 20.5.2) per attuare l'equalizzazione di canale mediante ritardi e prodotti complessi; qualora però il rumore additivo sia colorato (o generato da interferenti) il ricevitore RAKE può essere *generalizzato* per tenerne conto.

Citiamo infine la soluzione dell'equalizzazione *turbo*[49], in cui lo stadio di equalizzazione e quello di decodifica si scambiano *decisioni soffici* ovvero verosimiglianze logaritmiche che rappresentano l'*informazione estrinseca* (pag. 587) relativa ai simboli ricevuti, iterando i rispettivi compiti finché non pervengono alle stesse decisioni.

[47] Tratta da J.G. Proakis, M. Salehi, *Communication systems engineering*, 2^{nd} Ed., 2002 Prentice-Hall

[48] Nonostante MMSE tenti di minimizzare σ_e^2 tenendo conto sia dell'ISI che del rumore, le sue prestazioni degradano in presenza di canali che manifestano profonde attenuazioni per alcune frequenze. D'altra parte per canali meno problematici, le sue prestazioni possono avvicinarsi a quelle dell'MLSD.

[49] Vedi ad es. `https://en.wikipedia.org/wiki/Turbo_equalizer`

18.5 Appendici

18.5.1 Potenza assorbita da un bipolo

La dimostrazione inizia definendo una potenza *istantanea* assorbita dal bipolo come $w(t) = v(t)\, i(t) = v(t) \cdot (v(t) * y(t))$. La potenza *media* (nel tempo) allora risulta

$$\begin{aligned}
\mathcal{W}_z &= \lim_{\Delta t \to \infty} \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} w(t)\, dt = \\
&= \lim_{\Delta t \to \infty} \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} v(t) \left[\int_{-\infty}^{\infty} v(t-\tau)\, y(\tau)\, d\tau \right] dt = \\
&= \int_{-\infty}^{\infty} y(\tau) \left[\lim_{\Delta t \to \infty} \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} v(t)\, v(t-\tau)\, dt \right] d\tau \\
&= \int_{-\infty}^{\infty} y(\tau)\, \mathcal{R}_v(-\tau)\, d\tau = \int_{-\infty}^{\infty} Y(f)\, \mathcal{P}_v(f)\, df = \\
&= \int_{-\infty}^{\infty} [\Re\{Y(f)\} + j\Im\{Y(f)\}]\, \mathcal{P}_v(f)\, df = \\
&= \int_{-\infty}^{\infty} \mathcal{P}_v(f)\, \frac{R(f)}{|Z(f)|^2} df
\end{aligned}$$

Nel terzultimo passaggio si è fatto uso del teorema di Parseval, e del fatto che $\mathcal{R}_v(\tau)$ è pari; nell'ultimo, si è tenuto conto che $\mathcal{P}_v(f)$, $R(f)$ e $|Z(f)|^2 = R^2(f) + X^2(f)$ sono funzioni pari di f, mentre $X(f)$ è dispari: pertanto il termine

$$\int_{-\infty}^{\infty} \Im\{Y(f)\}\, \mathcal{P}_v(f)\, df = \int_{-\infty}^{\infty} \mathcal{P}_v(f)\, \frac{X(f)}{|Z(f)|^2} df$$

è nullo. Notiamo che quest'ultimo termine rappresenta la *potenza reattiva*, che non è trasformata in altre forme di energia, e viene accumulata e restituita dalla componente reattiva del carico. Al contrario, il termine relativo a $\Re\{Y(f)\}$ rappresenta la potenza assorbita dalla componente resistiva, nota come *potenza attiva*, che viene completamente dissipata.

Avendo espresso la potenza assorbita $\mathcal{W}_z$ nella forma di un integrale in f, la funzione integranda è intuitivamente associabile allo spettro di densità di potenza: $\mathcal{W}_z(f) = \mathcal{P}_v(f)\, \frac{R(f)}{|Z(f)|^2}$. Lo stesso risultato può essere confermato svolgendo il seguente calcolo più diretto, pensando al bipolo come ad un filtro la cui grandezza di ingresso è $v(t)$ e quella di uscita $i(t)$.

La definizione di potenza media $\mathcal{W}_z = \lim_{\Delta t \to \infty} \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} w(t)\, dt$, in cui $w(t) = v(t)\, i(t)$, mostra come questa sia equivalente alla funzione di intercorrelazione tra i e v calcolata in $\tau = 0$: $\mathcal{W}_z = \mathcal{R}_{vi}(0)$. Allora, è ragionevole assumere che $\mathcal{W}_z(f) = \mathcal{F}\{\mathcal{R}_{vi}(\tau)\}$. Indicando infatti con $\otimes$ l'integrale di intercorrelazione, e ricordando che gli operatori di convoluzione e correlazione godono della proprietà commutativa, possiamo scrivere

$$\mathcal{R}_{vi}(\tau) = v(t) \otimes i(t) = v(t) \otimes (v(t) * y(t)) = (v(t) \otimes v(t)) * y(t) = \mathcal{R}_v(\tau) * y(t)$$

quindi, risulta che

$$\mathcal{W}_z(f) = \mathcal{F}\{\mathcal{R}_{vi}(\tau)\} = \mathcal{P}_v(f)\, Y(f) = \mathcal{P}_v(f)\, \frac{R(f) - jX(f)}{|Z(f)|^2}$$

In base alle stesse considerazioni già svolte, si verifica come il termine immaginario non contribuisce alla potenza media assorbita, e quindi può essere omesso dalla definizione di *potenza attiva* $\mathcal{W}_z(f)$.

18.5.2 Condizioni per il massimo trasferimento di potenza

Svolgiamo per intero la dimostrazione delle (18.3). Verifichiamo subito che, mantenendo $Z_g(f)$ fisso e per qualunque valore di $R_c(f)$, la potenza ceduta ad un carico espressa dalla (18.2): $\mathcal{W}_{z_c}(f) = \mathcal{P}_{v_g}(f)\, \frac{R_c(f)}{|Z_c(f)+Z_g(f)|^2}$ risulta massima se il suo denominatore è il più piccolo possibile, e ciò si verifica quando $X_c(f) = -X_g(f)$, ed in tal caso risulta

$$\mathcal{W}_{z_c}(f) = \mathcal{P}_{v_g}(f)\, \frac{R_c(f)}{\big(R_c(f) + R_g(f)\big)^2} = \frac{\mathcal{P}_{v_g}(f)}{R_c(f) + 2R_g(f) + \big(R_g(f)\big)^2/R_c(f)} \tag{18.39}$$

Per individuare ora la condizione su $R_c(f)$ che rende minimo il denominatore (e dunque $\mathcal{W}_{z_c}(f)$ massima), eseguiamone la derivata rispetto ad R_c (omettendo per brevità la dipendenza da f) ed eguagliamola a zero:

$$\frac{d}{dR_c}\left(R_c + 2R_g + \frac{R_g^2}{R_c}\right) = 1 - \left(\frac{R_g}{R_c}\right)^2 = 0 \tag{18.40}$$

che fornisce la condizione $R_c(f) = \pm R_g(f)$ in cui il valore negativo viene scartato mentre quello positivo, assieme alla condizione $X_c(f) = -X_g(f)$ determina la condizione $Z_c(f) = Z_g^*(f)$ espressa alla (18.3). Volendo verificare che la (18.40) individui effettivamente un minimo e non un massimo del denominatore di (18.39), se ne può eseguire la derivata seconda, ottenendo

$$\frac{d^2}{dR_c^2}\left(R_c + 2R_g + \frac{R_g^2}{R_c}\right) = \frac{d}{dR_c}\left[1 - \left(\frac{R_g}{R_c}\right)^2\right] = 2\frac{R_g^2}{R_c^3}$$

che verifichiamo immediatamente essere sempre positiva.

18.5.3 Potenza ceduta ad un carico $Z_c(f) \neq Z_g^*(f)$

Avendo a disposizione un generatore di segnale di potenza disponibile $\mathcal{W}_d(f)$ ed impedenza interna $Z_g(f)$ assegnate, la tensione a vuoto ai suoi capi ha densità di potenza (di segnale) pari a $\mathcal{P}_v(f) = \mathcal{W}_d(f)\, 4R_g(f)$. Collegando al generatore un carico generico $Z_c(f)$, la potenza dissipata da quest'ultimo risulta pari a

$$\mathcal{W}_{z_c}(f) = \mathcal{P}_v(f)\, \frac{R_c(f)}{\big|Z_g(f) + Z_c(f)\big|^2}$$

Il rapporto tra la densità di potenza effettivamente ceduta a $Z_c(f)$, e quella che le sarebbe ceduta se $Z_c(f)$ fosse adattato per il massimo trasferimento di potenza, fornisce

la perdita di potenza subita:

$$\frac{\mathcal{W}_{z_c}(f)}{\mathcal{W}_d(f)} = \mathcal{W}_d(f)\, 4R_g \frac{R_c(f)}{\left|Z_g(f) + Z_c(f)\right|^2} \cdot \frac{1}{\mathcal{W}_d(f)} = \frac{4R_g(f)\, R_c(f)}{\left|Z_g(f) + Z_c(f)\right|^2} = \alpha(f)$$

Pertanto, se $Z_c(f) \neq Z_g^*(f)$, su $Z_c(f)$ si dissipa una potenza pari a $\mathcal{W}_{z_c}(f) = \alpha(f)\,\mathcal{W}_d(f)$. Il medesimo risultato è valido anche per l'analisi dell'accoppiamento tra il generatore equivalente di uscita di una rete due porte ed un carico.

Esempio Consideriamo un generatore con $Z_g(f)$ resistiva e pari a 50 Ω, e con densità di potenza disponibile (a frequenze positive)

$$\mathcal{W}_d^+(f) = \frac{\mathcal{W}_d}{4W} rect_{2W}(f - f_0)$$

in cui $\mathcal{W}_d$ = 1 Watt è la potenza disponibile totale, distribuita uniformemente in una banda $2W$ = 10 KHz centrata a frequenza $f_0 = 1MHz$. Il generatore è collegato ad un carico

$$Z_c(f) = R_c(f) + jX_c(f)$$

con $R_c(f)$ = 50 Ω ed $X_c(f) = 2\pi f L$ = 50 Ω per $f = f_0$ (da cui $L = \frac{50}{2\pi 10^6} = 7.96\,\mu H$).

Essendo la banda di segnale $2W \ll f_0$, approssimiamo la dipendenza da f di $X_c(f)$ come una costante. In queste ipotesi la potenza effettivamente ceduta al carico risulta $\mathcal{W}_{z_c} = \alpha \mathcal{W}_d$, con

$$\alpha = \frac{4R_g R_c}{\left|Z_g + Z_c\right|^2} = \frac{4 \cdot 50 \cdot 50}{\left|50 + 50 + j50\right|^2} = \frac{10000}{12500} = 0.8$$

e quindi $\mathcal{W}_{z_c}$ = 0,8 Watt ovvero, in dBm: $10 \log_{10} 0.8$ = -0.97 dbW = 29.03 dBm.

Il valore $\alpha_{dB} = 10 \log_{10} \alpha$ = 0,97 dB rappresenta il valore della perdita di potenza causata dal mancato verificarsi delle condizioni di massimo trasferimento di potenza, e può essere tenuto in conto come una attenuazione supplementare al collegamento, in fase di valutazione del *link budget* (vedi capitolo seguente).

L'opera

Trasmissione dei Segnali e Sistemi di Telecomunicazione

è il risultato di un progetto ventennale di cultura libera, aggiornato di continuo ed evolutosi fino alla forma attuale. La sua disponibilità pubblica è regolata dalle norme di licenza Creative Commons

Attribuzione - Non commerciale - Condividi allo stesso modo

`https://creativecommons.org/licenses/by-nc-sa/4.0/deed.it`

e tutte le risorse relative al testo sono accessibili presso

`https://teoriadeisegnali.it/libro/`

Puoi contribuire al suo successo promuovendone la diffusione e supportarne lo sviluppo attraverso una donazione, in buona parte devoluta ai progetti *open source*[1] che ne hanno resa possibile realizzazione e divulgazione. Ai donatori viene accordato un accesso *vitalizio* al formato PDF *navigabile* di tutte le edizioni presenti *e future*.

[1]
- Lyx - `http://www.lyx.org/`
- LaTeX- `https://www.latex-project.org/`
- TeX Users Group - `https://tug.org/`
- Inkscape - `http://www.inkscape.org/`
- Gnuplot - `http://www.gnuplot.info/`
- Octave - `http://www.gnu.org/software/octave/`
- Geany - `https://www.geany.org/`
- Linux - `https://www.linux.it/`
- Free Software Foundation - `https://shop.fsf.org/`
- GNOME Foundation - `https://www.gnome.org/`
- Mozilla Foundation - `https://www.mozilla.org/it/`
- Wikipedia - `https://it.wikipedia.org`
- Internet Archive - `https://archive.org/about/`
- Creative Commons - `https://creativecommons.it/chapterIT/`
- WordPress - `https://it.wordpress.org/`
- Phplist - `https://www.phplist.org/`