Appunti di Teoria dei Segnali
Capitolo 14 - La modulazione (II)

Modulazione di frequenza e di fase (cenni)

INTRODUZIONE

Abbiamo osservato in precedenza che la *modulazione di ampiezza* di un segnale analogico s(t) consiste fondamentalmente nel trasmettere il segnale portante

$$c(t) = A_C \cos\left(2\pi f_C t + \varphi(t)\right)$$

facendone però variare, in ciascun istante, l'ampiezza A_C in modo proporzionale al valore assunto dal segnale s(t) in quell'istante: il segnale modulato risulta perciò essere del tipo

$$\underbrace{s_t(t) = A_C s(t) \cos\left(2\pi f_C t + \varphi(t)\right)}_{\text{modulazione AM}}$$

Viceversa, la **modulazione di fase** (PM) si effettua sempre trasmettendo la portante, ma facendone variare, istante per istante, la fase. Il segnale modulato in fase è dunque qualcosa del tipo seguente:

$$\underbrace{s_t(t) = A_C \cos\left(2\pi f_C t + K_P s(t)\right)}_{\text{modulazione PM}}$$

Per quanto riguarda, invece, la **modulazione di frequenza** (FM), il concetto è ovviamente quello di far variare, in ciascun istante, la frequenza della portante in modo proporzionale al valore assunto in quell'istante dal segnale da trasmettere s(t). Tuttavia, in questo caso l'espressione del segnale modulato è un po' più complicata da ottenere, in quanto non è SOLO la frequenza che subisce variazioni. Vediamo di spiegarci meglio.

Consideriamo intanto la generica portante sinusoidale, supponendo che la sua fase abbia un generico andamento temporale $\varphi(t)$:

$$c(t) = A_C \cos\left(2\pi f_C t + \varphi(t)\right)$$

Indichiamo con $\theta(t)$ l'argomento del Coseno:

$$\theta(t) = 2\pi f_C t + \varphi(t)$$

Si definisce "**frequenza istantanea**" del segnale c(t) la quantità

$$\boxed{f_i(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt}}$$

Evidentemente, questa è una grandezza associata alla portante, proprio in base a come abbiamo definito $\theta(t)$. <u>La caratteristica della modulazione di frequenza è quella di far variare, istante per istante, questa grandezza in modo proporzionale al valore assunto, in ciascun istante, dal segnale s(t)</u>. Vediamo allora quale legame intercorre tra questa grandezza e s(t) stesso: si fa in modo che risulti

$$f_i(t) = f_C + k_F s(t)$$

dove f_C è la frequenza della portante e $\mathbf{k_F}$ una costante da scegliere in modo opportuno. Evidentemente, ad ogni valore di s(t) corrisponderà un preciso valore di $f_i(t)$ e quindi anche un preciso valore dell'argomento $\theta(t)$ del coseno della portante. Vediamo allora quale relazione intercorre tra s(t) e $\theta(t)$: in base a come abbiamo definito $f_i(t)$ e in base all'ultima relazione scritta, è evidente che

$$\frac{1}{2\pi}\frac{d\theta(t)}{dt} = f_C + k_F s(t)$$

Questa è una equazione differenziale nella incognita $\theta(t)$: risolvendo si ottiene evidentemente

$$\theta(t) = 2\pi f_C t + 2\pi k_F \int_{-\infty}^{t} s(\tau)d\tau$$

Questo è dunque l'argomento del segnale modulato, il quale quindi risulta avere la seguente espressione generale:

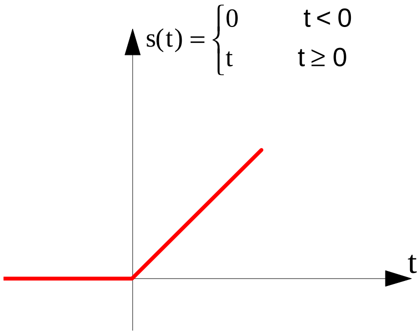
$$\underbrace{s_t(t) = A_C \cos\left(2\pi f_C t + 2\pi k_F \int_{-\infty}^{t} s(\tau)d\tau\right)}_{\text{modulazione FM}}$$

La cosa importante da sottolineare è la seguente: l'argomento del coseno in quest'ultima relazione dipende evidentemente da come è fatto il segnale s(t); è possibile che questo segnale abbia una struttura tale che, una volta effettuata l'integrazione, vengano modificate sia la fase sia la frequenza della portante. Questo per dire che *la modulazione di frequenza NON comporta* *necessariamente solo la variazione della frequenza della portante, ma può comportare anche la variazione della fase.*

Questo concetti risulterà più chiaro tramite gli esempi seguenti.

ESEMPIO: MODULAZIONE FM E PM DELLA RAMPA

Supponiamo che il segnale analogico da trasmettere sia la classica **rampa**:



Le espressioni generali dei segnali modulati in frequenza e in fase sono le seguenti:

$$\underbrace{s_t(t) = A_C \cos(2\pi f_C t + K_P s(t))}_{\text{modulazione PM}}$$

$$\underbrace{s_t(t) = A_C \cos\left(2\pi f_C t + 2\pi k_F \int_{-\infty}^{t} s(\tau)d\tau\right)}_{\text{modulazione FM}}$$

E' evidente che, per t<0, in entrambi i casi il segnale modulato corrisponde alla portante in quanto s(t) vale 0. Vediamo perciò cosa succede per t≥0.

Per la modulazione di fase abbiamo quanto segue:

$$s_t(t) = A_C \cos(2\pi f_C t + K_P t) = A_C \cos((2\pi f_C + K_P)t) = A_C \cos\left(2\pi\left(f_C + \frac{K_P}{2\pi}\right)t\right)$$

Ciò che si ottiene non è dunque una variazione istantanea della fase rispetto a quella della portante, che è sempre 0, bensì una variazione istantanea della frequenza, che risulta essere $f_C + \frac{K_P}{2\pi}$.

La cosa interessante è che la variazione della frequenza è costante nel tempo, nel senso che la frequenza assume sempre lo stesso valore in ciascun istante.

Vediamo invece cosa succede modulando in frequenza:

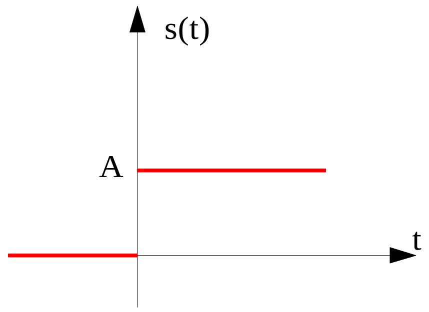
$$s_t(t) = A_C \cos\left(2\pi f_C t + 2\pi k_F \int_0^t \tau d\tau\right) = A_C \cos(2\pi f_C t + \pi k_F t^2) = A_C \cos\left(2\pi\left(f_C + k_F \frac{t}{2}\right)t\right)$$

Anche in questo caso, non ci sono variazioni di fase, mentre varia la frequenza della portante, che risulta essere $f_C + k_F \frac{t}{2}$.

La differenza con la modulazione di fase è però nel fatto che la frequenza varia istante per istante; anzi, si nota come essa aumenti linearmente all’aumentare del tempo, esattamente come la rampa.

ESEMPIO: MODULAZIONE FM E PM DEL GRADINO

Supponiamo adesso che il segnale analogico da trasmettere sia un gradino di altezza A:



Partiamo sempre dalle espressioni generali dei segnali modulati in frequenza e in fase:

$$\underbrace{s_t(t) = A_C \cos\left(2\pi f_C t + K_P s(t)\right)}_{\text{modulazione PM}}$$

$$\underbrace{s_t(t) = A_C \cos\left(2\pi f_C t + 2\pi k_F \int_{-\infty}^{t} s(\tau)d\tau\right)}_{\text{modulazione FM}}$$

Anche in questo caso, per t<0, il segnale modulato coincide con la portante. Vediamo invece per t≥0.

Modulando in fase, otteniamo

$$s_t(t) = A_C \cos\left(2\pi f_C t + K_P A\right)$$

per cui, questa volta, la frequenza rimane quella della portante, mentre c’è uno sfasamento rispetto alla portante pari a $K_P A$. Tale sfasamento è costante nel tempo.

Modulando invece in frequenza, abbiamo quanto segue:

$$s_t(t) = A_C \cos\left(2\pi f_C t + 2\pi k_F \int_0^t A d\tau\right) = A_C \cos\left(2\pi f_C t + 2\pi A k_F t\right) = A_C \cos\left(2\pi\left(f_C + A k_F\right)t\right)$$

Si osserva che non ci sono variazioni di fase, mentre cambia la frequenza, che diventa $f_C + A k_F$. Anche in questo caso, la nuova frequenza rimane costante nel tempo.

Si osserva in particolare che la frequenza risulta aumentata, così come accadeva anche nel caso della rampa: questo è un risultato generale, nel senso che *la frequenza, nella modulazione FM, aumenta quando s(t) è positivo mentre diminuisce quando s(t) è negativo.*

Modulazione FM di un singolo tono modulante

ESPRESSIONE DEL SEGNALE MODULATO

A questo punto, lasciamo perdere la modulazione di fase e concentriamoci sulla modulazione di frequenza. In particolare studiamo un caso particolare, ossia quello in cui il segnale analogico s(t) da trasmettere sia un **singolo tono modulante**:

$$s(t) = A_S \cos(2\pi f_S t)$$

Naturalmente, è superfluo sottolineare come la frequenza f_S di questo segnale non abbia niente a che fare con quella f_C della portante.

Vediamo qual è il risultato della modulazione di frequenza di tale segnale: abbiamo detto che l'espressione generale del segnale modulato FM è

$$s_t(t) = A_C \cos\left(2\pi f_C t + 2\pi k_F \int_{-\infty}^{t} s(\tau)d\tau\right)$$

per cui dobbiamo andare a sostituire l'espressione di s(t) e fare i calcoli. Al fine di semplificarci tali calcoli (ed in particolare il calcolo dell'integrale) supponiamo che s(t) valga 0 prima di t=0 e sia $s(t) = A_S \cos(2\pi f_S t)$ dopo t=0. In tal modo, il segnale modulato risulta essere il seguente:

$$s_t(t) = A_C \cos\left(2\pi f_C t + 2\pi k_F \int_{-\infty}^{t} s(\tau)d\tau\right) = A_C \cos\left(2\pi f_C t + 2\pi k_F \int_{0}^{t} A_S \cos(2\pi f_S \tau)d\tau\right) =$$
$$= A_C \cos\left(2\pi f_C t + 2\pi k_F A_S \left(\frac{1}{2\pi f_S}\sin(2\pi f_S t)\right)\right) = A_C \cos\left(2\pi f_C t + \frac{k_F A_S}{f_S}\sin(2\pi f_S t)\right)$$

A questo punto diamo due definizioni che ci consentano di semplificare quella espressione: in primo luogo, si definisce **deviazione di frequenza** la quantità

$$\Delta f = k_F A_S$$

Si può verificare come essa rappresenti la massima differenza di frequenza tra il segnale modulato ed il segnale portante.

Possiamo dunque riscrivere il segnale modulato nella forma

$$s_t(t) = A_C \cos\left(2\pi f_C t + \frac{\Delta f}{f_S}\sin(2\pi f_S t)\right)$$

Si chiama inoltre **indice di modulazione** la quantità

$$m = \frac{\Delta f}{f_S}$$

Con questa seconda posizione, l’espressione del segnale modulato diventa

$$\boxed{s_t(t) = A_C \cos(2\pi f_C t + m\sin(2\pi f_S t))}$$

Questa è l’espressione più generale possibile per il segnale modulato.

SPETTRO E OCCUPAZIONE DI BANDA DEL SEGNALE MODULATO

Una caratteristica importante del segnale modulato è, come sappiamo, l’occupazione di banda, ossia l’intervallo di frequenze da esso coperto: infatti, è anche in base a questa caratteristica che dobbiamo progettare i mezzi trasmissivi ed i vari filtri usati negli apparati di modulazione/demodulazione. Al fine di valutare l’occupazione di banda del segnale $s_t(t)$, ci mettiamo inizialmente in una condizione particolare: supponiamo infatti che sia $m << 1$ e vediamo cosa succede.

Intanto, a prescindere da quella condizione particolare, possiamo usare le formule di duplicazione del coseno per riscrivere il segnale modulato

$$s_t(t) = A_C \cos(2\pi f_C t + m\sin(2\pi f_S t))$$

nella forma equivalente

$$s_t(t) = A_C \cos(2\pi f_C t)\cos(m\sin(2\pi f_S t)) - A_C \sin(2\pi f_C t)\sin(m\sin(2\pi f_S t))$$

A questo punto, il fatto che m<<1 ci consente di fare le seguenti due approssimazioni:

$$\cos(m\sin(2\pi f_S t)) \cong 1$$
$$\sin(m\sin(2\pi f_S t)) \cong m\sin(2\pi f_S t)$$

Sulla base di queste approssimazioni, il segnale modulato diventa il seguente:

$$s_t(t) = A_C \cos(2\pi f_C t) - mA_C \sin(2\pi f_C t)\sin(2\pi f_S t)$$

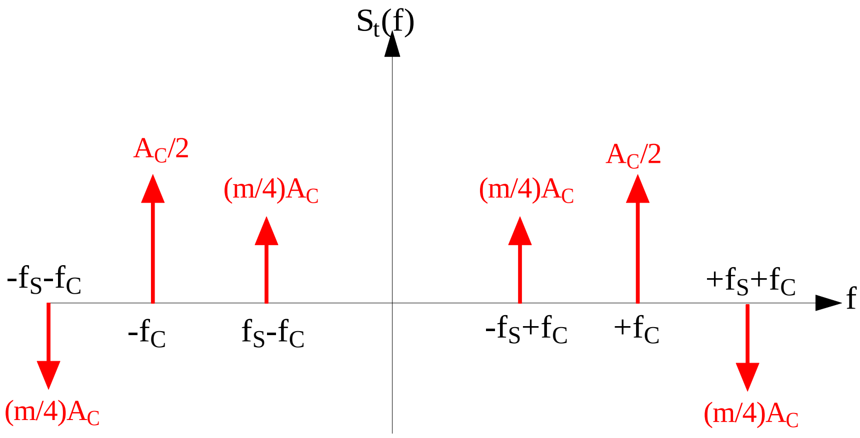
Sfruttando le opportune formule trigonometriche, possiamo inoltre sviluppare il secondo termine a secondo membro, ottenendo quanto segue:

$$s_t(t) = A_C \cos(2\pi f_C t) - \frac{mA_C}{2}\cos(2\pi(f_C + f_S)t) + \frac{mA_C}{2}\cos(2\pi(f_C - f_S)t)$$

Passiamo adesso al dominio della frequenza: lo spettro di questo segnale è il seguente:

$$S_t(f) = A_C\left[\frac{1}{2}\delta(f - f_C) + \frac{1}{2}\delta(f + f_C)\right] - \frac{mA_C}{2}\left[\frac{1}{2}\delta(f - f_C - f_S) + \frac{1}{2}\delta(f + f_C + f_S)\right] + \frac{mA_C}{2}\left[\frac{1}{2}\delta(f - f_C + f_S) + \frac{1}{2}\delta(f + f_C - f_S)\right]$$

Abbiamo dunque 6 impulsi posizionati in posizioni simmetriche rispetto all'origine delle frequenze:



Ciò che si deduce è dunque che la banda del segnale modulato è $f_S + f_C$, ossia corrisponde alla somma della banda (che poi è la frequenza) della portante e del segnale modulante. Quindi, i valori non nulli dello spettro si hanno nei due intervalli $[-f_S - f_C, f_S - f_C]$ e $[-f_S + f_C, f_S + f_C]$, mentre al di fuori di tali intervalli si hanno solo valori nulli.

Questo fatto è ovviamente strettamente legato all'ipotesi di partenza di un indice di modulazione m molto piccolo. Si può già intuire che, <u>per m generico, la banda del segnale modulato aumenta</u>: il motivo sta nel modo in cui varia il coseno presente nella espressione

$$s_t(t) = A_C \cos(2\pi f_C t + m\sin(2\pi f_S t))$$

al variare di m.

Vediamo ad ogni modo che cosa accade dal punto di vista matematico. Ricordando dalla teoria dei numeri complessi che vale la relazione

$$e^{j\vartheta} = \cos\vartheta + j\sin\vartheta$$

possiamo riscrivere il segnale modulato nel modo seguente:

$$s_t(t) = A_C \operatorname{Re}\left\{e^{j2\pi f_C t} e^{jm\sin(2\pi f_S t)}\right\}$$

A questo punto, è evidente che il segnale $e^{jm\sin(2\pi f_S t)}$ è un segnale periodico di periodo $1/f_S$, in quanto è tale il segnale presente all'esponente. Sappiamo allora che un segnale periodico è esprimibile mediante uno sviluppo in serie di Fourier: se poniamo

$$\tilde{s}(t) = e^{jm\sin(2\pi f_S t)}$$

il suo sviluppo in serie sarà

$$\tilde{s}(t) = \sum_{k=-\infty}^{+\infty} c_K e^{j2\pi k f_S t}$$

Per determinare questo sviluppo, dobbiamo evidentemente calcolarne i coefficienti: usando la normale definizione abbiamo che

$$c_K = \frac{1}{T}\int_{-T/2}^{+T/2} \tilde{s}(t) e^{-j2\pi k f_S t} dt$$

dove $T=1/f_S$.

Sostituendo l'espressione del segnale, abbiamo perciò che

$$c_K = \frac{1}{T}\int_{-T/2}^{+T/2} e^{jm\sin(2\pi f_S t)} e^{-j2\pi k f_S t} dt = \frac{1}{T}\int_{-T/2}^{+T/2} e^{jm\sin(2\pi f_S t) - j2\pi k f_S t} dt = f_S \int_{-1/2f_S}^{+1/2f_S} e^{jm\sin(2\pi f_S t) - j2\pi k f_S t} dt$$

Facendo adesso il cambio di variabile $x=2\pi f_S t$, abbiamo inoltre che

$$c_K = f_S \int_{-\pi}^{+\pi} e^{jm\sin x - jkx} \frac{1}{2\pi f_S} dx = \frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{jm\sin x - jkx} dx$$

A questo punto, la risoluzione di quell'integrale è possibile ma è tutt'altro che agevole. Si verifica comunque che il risultato è comunque un numero reale per ogni k: poniamo allora

$$c_K = \frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{jm\sin x - jkx} dx = J_K(m)$$

per cui il nostro sviluppo in serie risulta essere

$$\tilde{s}(t) = \sum_{k=-\infty}^{+\infty} J_K(m) e^{j2\pi k f_S t}$$

Tornando allora all'espressione del segnale modulato $s_t(t)$, esso diventa

$$s_t(t) = A_C \operatorname{Re}\left\{ e^{j2\pi f_C t} \sum_{k=-\infty}^{+\infty} J_K(m) e^{j2\pi k f_S t} \right\}$$

Il termine esponenziale $e^{j2\pi f_C t}$ non dipende chiaramente da k, per cui lo portiamo dentro la sommatoria:

$$s_t(t) = A_C \operatorname{Re}\left\{ \sum_{k=-\infty}^{+\infty} e^{j2\pi f_C t} J_K(m) e^{j2\pi k f_S t} \right\}$$

La parte reale è un operatore lineare, per cui possiamo riscrivere il segnale modulato nella forma

$$s_t(t) = A_C \sum_{k=-\infty}^{+\infty} \operatorname{Re}\left\{ e^{j2\pi f_C t} J_K(m) e^{j2\pi k f_S t} \right\}$$

Avendo detto poi che i termini $J_K(m)$ risultano essere reali, abbiamo anche che

$$s_t(t) = A_C \sum_{k=-\infty}^{+\infty} J_K(m) \operatorname{Re}\left\{ e^{j2\pi f_C t} e^{j2\pi k f_S t} \right\}$$

Applicando infine la relazione $e^{j\vartheta} = \cos\vartheta + j\sin\vartheta$, possiamo concludere che

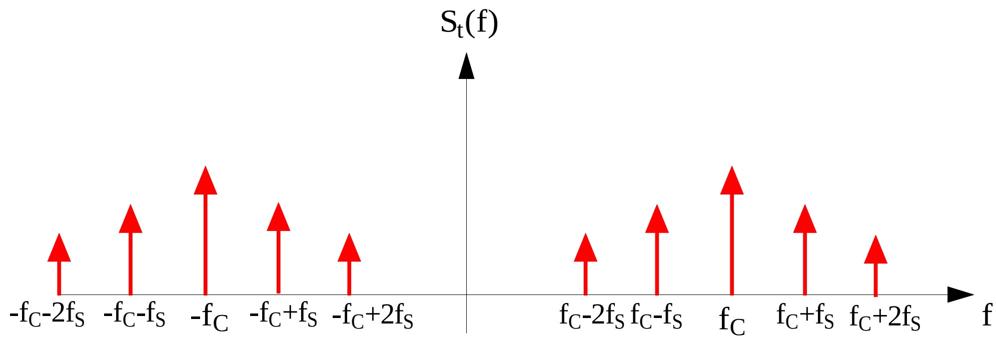
$$\boxed{s_t(t) = A_C \sum_{k=-\infty}^{+\infty} J_K(m) \cos\left(2\pi\left(f_C + k f_S\right)t\right)}$$

Questa è dunque l'espressione più generale possibile, nel dominio del tempo, per il singolo tono modulato in FM. Dato che a noi interessa l'occupazione di banda di questo segnale, effettuiamo la trasformata di Fourier per passare al dominio della frequenza: è immediato verificare che lo spettro di $s_t(t)$ risulta essere

$$\boxed{S_t(f) = A_C \sum_{k=-\infty}^{+\infty} \frac{J_K(m)}{2} \left[\delta\left(f - \left(f_C + k f_S\right)\right) + \delta\left(f + \left(f_C + k f_S\right)\right) \right]}$$

Abbiamo dunque una successione di infiniti impulsi: questo significa che il segnale modulato presenta teoricamente tutte le infinite componenti in frequenza ed implica, quindi, che, se vogliamo trasmettere l'intero contenuto energetico (o di potenza) del segnale, possiamo usare il nostro mezzo trasmissivo per trasmettere 1 solo segnale.

C'è però da osservare un fatto: nel valutare, con appositi metodi numerici, i valori dei coefficienti $J_K(m)$, si trova che essi assumono valori (reali) sempre più piccoli al crescere del valore assoluto di k. Questo significa che gli impulsi di cui è composto $S_t(f)$ non hanno tutti la stessa area: al contrario, fissato l'indice di modulazione m, essi presentano il massimo valore in corrispondenza di k=0, mentre poi presentano valori via via decrescenti all'aumentare di k in valore assoluto. Possiamo cioè schematizzare la cosa nel modo seguente:



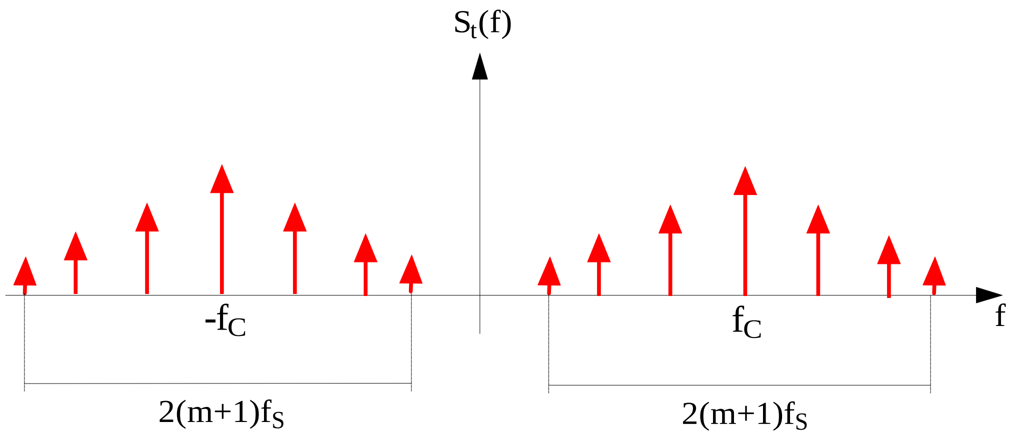
Si nota che il valore massimo si ha in corrispondenza della frequenza f_C della portante (e ovviamente del suo valore simmetrico $-f_C$), mentre, allontanandosi via via de essa, l'area degli impulsi decresce.

Questo fatto è molto importante in quanto consente di fare il seguente discorso: se l'area degli impulsi va via via decrescendo, possiamo ritenere che esisterà un valore di |k| oltre il quale l'area degli impulsi sarà talmente piccola da poterla ritenere nulla. Detto in altri termini, esisterà un valore di |k| oltre il quale noi potremo trascurare il contributo energetico (o di potenza) del segnale.

Detto ancora in altre parole, possiamo ritenere che il contenuto energetico (o di potenza) del segnale sia concentrato in un certo intorno della frequenza della portante: in particolare, lo standard internazionale considera la cosiddetta **approssimazione di Carson**, secondo la quale l'ampiezza dell'intorno, centrato in f_C, entro il quale è concentrata l'energia del segnale vale

$$\boxed{B_{RF} = 2(m+1)f_S}$$

dove ovviamente m è sempre l'indice di modulazione e f_S la frequenza del singolo tono modulante che stiamo trasmettendo.



Allora, utilizzando questa approssimazione, è evidente che il nostro segnale modulato ha una occupazione di banda limitata, il che significa che possiamo usare lo stesso mezzo trasmissivo per trasmettere più di un segnale. Naturalmente, dobbiamo fare in modo che gli spettri dei vari segnali non si sovrappongano: è una cosa facile da fare in quanto, nota la frequenza di ciascun segnale, è sufficiente spaziare in modo opportuno la frequenza della sua portante rispetto alle altre. Per esempio, supponiamo di voler trasmettere due singoli toni modulanti, ossia due segnali del tipo

$$s(t) = A_S \cos(2\pi f_S t)$$
$$x(t) = A_X \cos(2\pi f_X t)$$

Supponiamo che siano note le rispettive frequenze f_S e f_X; ciò che dobbiamo fare è scegliere in modo opportuno le frequenze delle due portanti

$$c_S(t) = A_{CS}\cos\left(2\pi f_{CS}t + \varphi_S\right)$$
$$c_X(t) = A_{CX}\cos\left(2\pi f_{CX}t + \varphi_X\right)$$

Trasmettendo s(t) modulato in FM con frequenza portante f_{CS} e con indice di modulazione m, abbiamo visto prima che l'intervallo entro il quale si ritiene concentrata l'energia di s(t) stesso è

$$\left[f_{CS} - (m+1)f_S, f_{CS} + (m+1)f_S\right]$$

In modo analogo, trasmettendo invece x(t) modulato in FM con frequenza portante f_{CX} e indice di modulazione ancora m, l'intervallo entro il quale si ritiene concentrata l'energia di x(t) è

$$\left[f_{CX} - (m+1)f_X, f_{CX} + (m+1)f_X\right]$$

Allora, dobbiamo fare in modo che questi due intervalli non abbiano alcun valore in comune.

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