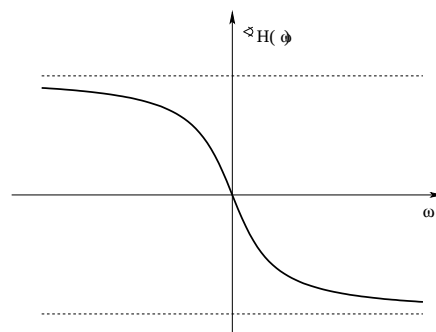
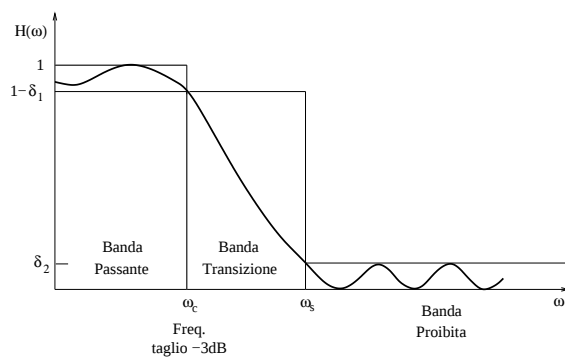

Introduzione all'elaborazione dei segnali

Alberto Bertoni Paola Campadelli Giuliano Grossi



Anno Accademico 2010-2011

Indice

1	Segnali e Sistemi	1
1.1	Segnali e Informazione	2
1.2	Classificazione dei Segnali	4
1.3	Esempi di segnali	8
1.4	Sistemi per l'Elaborazione dei Segnali Deterministici	13
1.4.1	Composizionalità nei Sistemi	16
1.5	Spazi vettoriali di segnali	18
1.5.1	Basi ortonormali	22
1.6	Sistemi Lineari	24
1.6.1	Sistemi Tempo-Invarianti	26
1.6.2	Sistemi Lineari Tempo-Invarianti: filtri lineari	28
1.6.3	Sistemi Causali e Stabili	30
1.6.4	Filtri FIR e IIR per segnali a tempo discreto	33
2	Analisi in Frequenza di Segnali Analogici	41
2.1	Numeri Complessi	43
2.2	Segnali Periodici	45
2.3	Segnali a spettro finito.	47
2.4	Segnali a spettro discreto: serie di Fourier.	50
2.5	Serie di Fourier: forme alternative	53
2.6	Serie di Fourier: problemi di convergenza	54
2.7	Segnali a spettro continuo: la trasformata di Fourier	56
2.7.1	Esistenza della Trasformata di Fourier	60
2.7.2	Trasformata di Fourier di Funzioni a valori Reali	62
2.7.3	Proprietà della Trasformata di Fourier	63
2.7.4	Energia e Potenza	65
2.8	Coppie Base di Trasformate	68
3	Filtri Analogici e Modulazione di Ampiezza	71
3.1	Risposta in Frequenza dei Sistemi Lineari Tempo-Invarianti	72
3.1.1	Filtri Ideali	73
3.2	Caratteristiche dei Filtri Analogici	75
3.3	Famiglie di Filtri Causali	78

3.3.1	Filtri di Butterworth	79
3.3.2	Realizzazione di Filtri Analogici	80
3.4	Modulazione e Demodulazione di Ampiezza (AM)	81
4	Conversione Analogico-Digitale	87
4.1	Campionamento	88
4.1.1	Spettro del segnale campionato	91
4.1.2	Aliasing	93
4.2	Quantizzazione	95
4.2.1	Quantizzatore Uniforme e Rumore di Quantizzazione	95
4.2.2	Quantizzatore Non Uniforme	99
4.3	Convertitore Analogico-Digitale (ADC)	100
5	Analisi in frequenza di sistemi a tempo discreto	103
5.1	Frequenza e Tempo Normalizzati	104
5.2	Trasformata Zeta	106
5.2.1	Proprietà della trasformata Zeta	107
5.3	Sistemi LTI a Tempo Discreto: filtri FIR e IIR	108
5.3.1	Filtri FIR	109
5.4	Filtri IIR	111
5.4.1	Implementazione di filtri come Reti	112

CAPITOLO 1

Segnali e Sistemi

I *segnali* sono oggetti fisici, quali le onde sonore o elettromagnetiche, atti a contenere o veicolare informazioni. Ne sono esempi i segnali del parlato, i segnali audio e video, i segnali sismici o i segnali biomedici. I segnali possono essere memorizzati, elaborati o trasmessi mediante processi fisici: chiameremo *sistema* un qualsiasi apparato in grado di memorizzare, elaborare o trasmettere segnali. In generale, un sistema *opera* ricevendo in ingresso segnali e producendo in uscita nuovi segnali. Le tensioni o le correnti (funzioni del tempo) in ingresso e in uscita a un circuito elettrico sono esempi di segnali, mentre il circuito stesso è un esempio di sistema; analogamente, è un sistema la telecamera che cattura la luce riflessa dagli oggetti inquadrati e impressiona la pellicola o una matrice di sensori.

I concetti di segnale e di sistema giocano oggi un ruolo di rilievo in diverse aree scientifiche e tecnologiche, dalle telecomunicazioni al controllo dei processi industriali, dalla multimedialità alla bioinformatica. Questi concetti meritano dunque di essere inquadrati e analizzati con la necessaria precisione, così che il linguaggio matematico risulta indispensabile per la loro descrizione e comprensione.

Le principali problematiche che si affrontano nello studio dei sistemi sono quelle di *analisi* e di *disegno* o *progettazione*. L'analisi riguarda lo studio di caratteristiche del comportamento di un dato sistema, il disegno si propone di identificare e realizzare, se esiste, un sistema in grado di esibire un comportamento specificato. Un esempio è dato dalla progettazione di sistemi per estrarre o ripristinare informazione danneggiata: il caso tipico è quello della comunicazione affetta da rumore in cui è richiesto lo sviluppo di sistemi in grado di correggere eventuali errori.

In questo capitolo vengono introdotte le nozioni di base riguardanti segnali e sistemi. Vengono presentate alcune semplici classificazioni dei tipi di segnale (analogici/digitali, deterministici/probabilistici); limitatamente ai segnali deterministici, sono richiamate alcune famiglie di segnali temporali che ricorrono frequentemente (sinusoidi, seni cardinali, rettangoli, impulsi). Viene introdotto poi il concetto di sistema, visto come un apparato in grado di trasformare un segnale in ingresso in un segnale in uscita; il comportamento del sistema viene descritto proprio dalla relazione ingresso-uscita.

Si osserva che la somma di segnali è ancora un segnale, così come la moltiplicazione di un segnale per uno scalare. Questo porta a studiare gli spazi di segnali come spazi vettoriali; viene introdotta l'importante nozione di *base* di uno spazio e, attraverso la nozione di *prodotto interno*, quella di *base ortonormale*. Queste nozioni sono centrali per lo studio della rappresentazione in frequenza dei segnali, oggetto di analisi nel capitolo 2.

Vengono poi introdotte due importanti classi di sistemi: i *sistemi lineari* e quelli *tempo-invarianti*. In particolare, si mostra che il comportamento di un sistema che è contemporaneamente lineare e tempo-invariante (LTI), è univocamente individuato dalla risposta del sistema al segnale impulsivo. Limitatamente ai sistemi LTI, chiamati anche *filtri lineari*, sono discusse le nozioni di *causalità* e *stabilità*. Vengono infine introdotte due sottoclassi di filtri per segnali digitali, i filtri *con risposta all'impulso finita* (FIR) e i filtri *ricorsivi*, che vengono anche discussi in relazione agli *automi a stati*.

1.1 Segnali e Informazione

Gli esseri viventi non sono entità isolate, potendo interagire attraverso l'invio e la ricezione di informazioni. L'informazione viene trasmessa attraverso un supporto fisico-chimico: tale supporto induce negli organi sensoriali del ricevente particolari reazioni che corrispondono all'estrazione dell'informazione stessa. Così gli insetti possono comunicare producendo sostanze chimiche, i ferormoni, capaci di suscitare specifiche reazioni comportamentali in individui che ne vengano in contatto. Ad esempio, i ferormoni di allarme trasmettono a quanti li captano l'informazione di porsi in un particolare stato (stato di allarme). Distinguiamo qui due aspetti: il supporto fisico dell'informazione (concentrazione di ferormoni) e l'informazione veicolata (allarme sì/no).

Questo è un esempio di segnale. Un *segnale* è costituito quindi da una entità fisica contenente qualche tipo di informazione che può essere estratta, trasmessa o elaborata.

Gli elementi costituenti un segnale sono allora il supporto fisico e l'informazione veicolata. Ad esempio, è noto che il gatto usa marcare il territorio con particolari secrezioni: il supporto fisico in questo caso è costituito da molecole, e l'informazione che si vuol comunicare (il possesso del territorio) è particolarmente semplice, riducendosi a 1 bit (sì/no). Molto più complesse sono le informazioni trasmesse dall'uomo tramite segnali acustici: si pensi a un docente che enuncia oralmente in classe il teorema di Pitagora, o a un pianista che esegue in pubblico un improvviso di Scubert!

Le organizzazioni umane sono caratterizzate da un crescente aumento di scambi di informazioni attraverso segnali. Segnali di cui abbiamo esperienza nella nostra vita quotidiana sono ad esempio:

- segnali del parlato, che si trasmettono via telefono o radio;
- segnali video e immagini;
- suoni e musica, riproducibili ad esempio da un lettore di DVD;
- segnali biomedici, come elettrocardiogrammi, radiografie, livelli di espressione genica.

La capacità di elaborare o trasmettere informazioni per via automatica è quindi una caratteristica rilevante nell'attuale società. Di conseguenza, i metodi per l'elaborazione dei segnali assumono oggi grande rilievo e trovano feconde applicazioni nelle seguenti aree:

1. Trasmissione, ricezione e memorizzazione efficiente ed affidabile dei segnali nelle telecomunicazioni; si pensi, ad esempio, al complicato sistema che ci permette di vedere la ripresa di una partita di calcio con un televisore digitale.
2. Estrazione di informazione da segnali rumorosi con tecniche di riconoscimento di forme, per previsione e controllo; si pensi alle tecniche diagnostiche per stimare la presenza di una malattia integrando dati provenienti da radiografie, NMR e dati di espressione genica dello stesso paziente.

Per progettare e analizzare metodi e sistemi dedicati all'elaborazione dei segnali è necessario poter formulare i concetti in forma matematicamente precisa.

Consideriamo un insieme C di valori di coordinate spazio-temporali (si pensi ai possibili valori del tempo) e un insieme R dei possibili valori di una grandezza fisica (si pensi ai valori di pressione); un *segnale* è una funzione

$$f : C \rightarrow R$$

che rappresenta o codifica informazione. Esso sarà denotato da $g = f(c)$, dove:

- c denota una o più variabili, che identificano le coordinate spaziali o temporali.
- g denota una variabile a valori in R , possibili risultati di misure di una grandezza fisica o chimica (corrente, tensione, pressione, concentrazione di molecole e così via).
- f denota la relazione funzionale che associa ad ogni valore di c il corrispondente valore della grandezza fisica g .

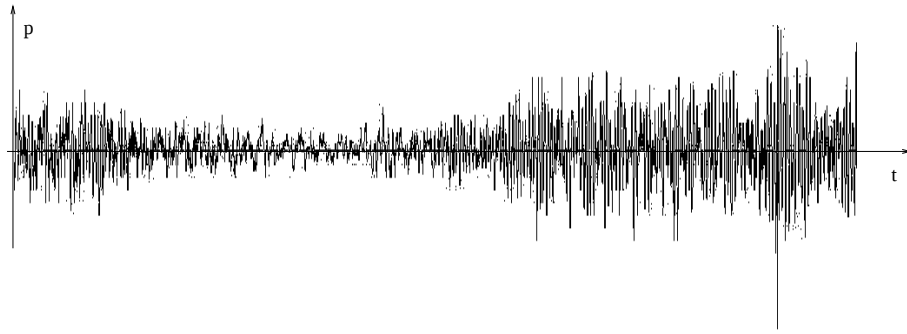


Figura 1.1 Segnale acustico.

In Figura 1.1 è mostrato un esempio di *segnale acustico*. Esso è identificato dalla complicata funzione $p(t)$ il cui grafico dà la pressione acustica, in un punto dello spazio, in funzione del tempo t . Segnali che hanno il tempo come variabile indipendente si dicono *segnali temporali*, e ad essi è dedicata gran parte di questo corso.

Un secondo esempio ci viene dato dall'immagine monocromatica di Figura 1.2, identificata dalla funzione $L(x, y)$, che dà la luminosità in funzione delle due coordinate spaziali (x, y) .



Figura 1.2 Immagine monocromatica.

1.2 Classificazione dei Segnali

Il mondo che ci circonda è in gran parte basato su segnali analogici: il tempo, le coordinate spaziali, grandezze fisiche come l'intensità della luce e del suono assumono valori continui. I nostri organi sensoriali traducono tali segnali in treni di impulsi elettrici; questi vengono analizzati dal cervello che prende decisioni estraendo le informazioni veicolate.

D'altro lato, la tecnologia digitale assume oggi grande rilievo. Tra i vantaggi delle macchine digitali su quelle analogiche ricordiamo il miglior controllo degli errori in trasmissione e la grande flessibilità nelle applicazioni; queste macchine dispongono oggi di una grande potenza di calcolo e sono collegate in rete. Non stupisce che si sia pensato di utilizzare questa tecnologia anche per elaborare i segnali.

Il mondo analogico è tuttavia distante da quello digitale: un elaboratore lavora su una scala di tempi discreta, scandita dal clock, ed inoltre i valori che può assumere una grandezza trattata digitalmente sono finiti. La trattazione digitale dei segnali richiede dunque la soluzione di due problemi di rilievo:

1. interfacciare il mondo analogico con quello digitale;
2. dotarsi di strumenti per trattare numericamente segnali digitali che, a causa delle approssimazioni, sono intrinsecamente rumorosi.

A questo riguardo, è utile per prima cosa introdurre alcune classificazioni dei segnali: segnali analogici/digitali, segnali deterministici/probabilistici.

Segnali Analogici e Digitali

Ricordiamo che un segnale temporale è rappresentato da una funzione $y = f(t)$, dove t è una variabile che identifica il tempo, y il valore di una opportuna grandezza fisica, quale la pressione, tensione, corrente o temperatura. Prendiamo in considerazione i seguenti casi: tempo continuo/discreto, valori continui/finiti.

1. Il tempo t , come grandezza fisica, fissata una unità di misura (il secondo, ad esempio) può assumere come valore un qualsiasi numero reale. Parleremo in questo caso di *tempo continuo*, poichè i numeri reali formano un continuo: dati due numeri $x < y$, esiste sempre un nuovo numero z compreso tra i due.

In molti contesti, siamo tuttavia interessati a considerare tempi presi in un sottoinsieme discreto di numeri reali. Un caso particolarmente interessante è il seguente: fissato un intervallo temporale τ , si considerano i multipli interi $n\tau$ di τ . Parleremo in questo caso di *tempo discreto*, poichè essenzialmente, fissato τ come unità di misura, il tempo assume valori interi, che rappresentano un insieme discreto: tra l'intero n e il suo successivo $n + 1$ non esiste alcun intero.

Il segnale a tempo discreto $f(n\tau)$ è ottenuto mediante campionamento regolare del segnale a tempo continuo, con passo di campionamento τ fissato: i campioni vengono ricavati ad istanti che sono multipli interi $n\tau$ di τ , con $-\infty < n < \infty$. Se chiamiamo *tempo normalizzato* la misura t' del tempo in cui τ viene preso come unità, cioè $t' = \frac{t}{\tau}$, allora agli istanti di campionamento $t = n\tau$ si ha che il tempo normalizzato è $\frac{n\tau}{\tau} = n$.

2. La variabile y rappresenta generalmente grandezze fisiche quali l'energia, la tensione o la temperatura, e quindi può assumere valori reali. Parleremo in questo caso di segnale a *valori continui*.

Se tuttavia vogliamo memorizzare i valori di un segnale in un calcolatore, ci scontriamo col fatto che un numero reale, essendo composto da infinite cifre, richiede infiniti bit per essere memorizzato. Avendo a disposizione al più un numero finito m di bit, che codificano solo 2^m valori, dovremo approssimare ogni numero reale con uno dei 2^m valori codificati. Il segnale risultante da questa approssimazione, che assumerà al più 2^m valori, sarà detto a *valori finiti*.

Consideriamo il seguente esempio. Indicando con $\text{sgn}(x)$ la funzione segno, definita da $\text{sgn}(x) = 1$ se $x > 0$, $\text{sgn}(x) = -1$ se $x < 0$, il segnale $\text{sgn}(f(t))$ risulta essere a 2 soli valori.

Come conseguenza della classificazione precedente (tempo continuo/discreto, valori continui/finiti) possiamo classificare i segnali in 4 classi: segnali a valori continui e tempo continuo, segnali a valori continui e tempo discreto, segnali a valori finiti e tempo continuo, segnali a valori finiti e tempo discreto. In Figura 1.3 sono illustrati i vari tipi di segnale.

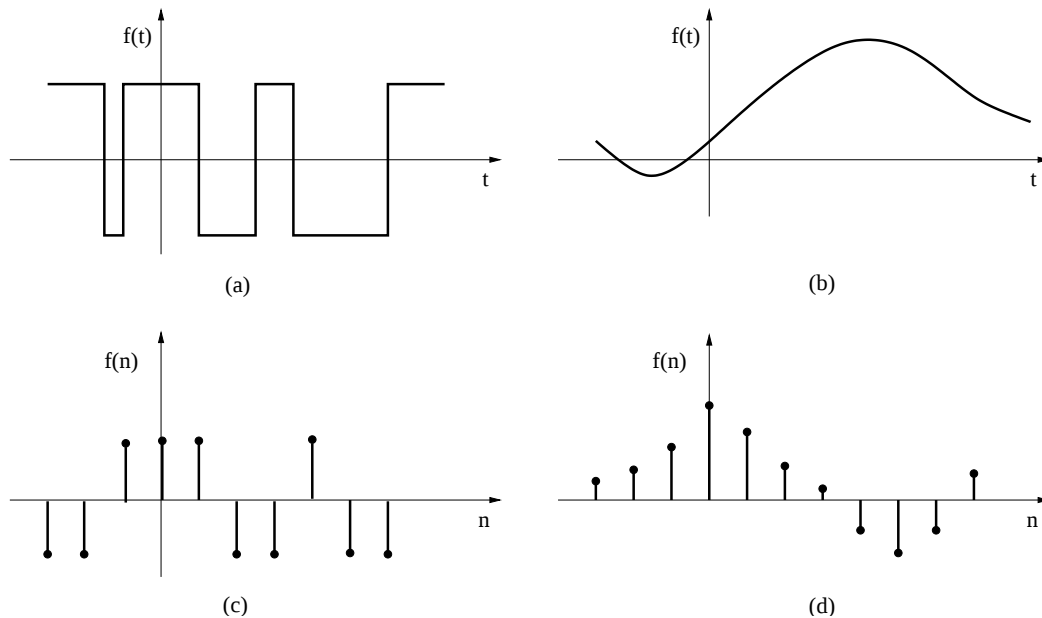


Figura 1.3 (a) Segnale a valori finiti e tempo continuo. (b) Segnale continuo a tempo continuo. (c) Segnale a valori finiti e tempo discreto. (d) Segnale continuo a tempo discreto.

I segnali a valori finiti e a tempo discreto sono detti anche *segnali digitali*, come opposto ai segnali continui a tempo continuo detti *segnali analogici*.

Osserviamo che un segnale digitale di durata temporale limitata può essere memorizzato su un supporto digitale, potendosi codificare con una quantità finita di bit. Sia infatti $f(t)$ un segnale tale che $f(t) = 0$ per $t \leq 0$ e per $t > T$, e quindi di durata T . Supponiamo

che $f(t)$ sia a valori finiti, diciamo a N valori, e sia campionato ai tempi $n\tau$. Il segnale digitale $f(n\tau)$ è perfettamente descritto dal vettore:

$$\left[f(\tau), f(2\tau), f(3\tau), \dots, f\left(\frac{T}{\tau}\tau\right) \right]$$

Ogni componente del vettore è codificabile con un numero di bit pari al primo intero maggiore o uguale di $\log_2(N)$, quindi il segnale digitale può essere memorizzato con M bit, dove:

$$M = \frac{T}{\tau} \cdot \lceil \log_2(N) \rceil = \nu T \cdot \lceil \log_2(N) \rceil$$

$\nu = \frac{1}{\tau}$, detta *frequenza di campionamento*, identifica il numero di campioni al secondo ed ha lo Hertz (Hz) come unità di misura.

Esempio 1.2.1

Determinare quanti byte sono richiesti per memorizzare un segnale acustico della durata di 1 minuto, campionato ad una frequenza di 8 KHz e quantizzato a 256 valori.

Per codificare 256 valori si richiedono $\log_2 256 = 8$ bit, equivalenti a 1 byte. Poiché i campioni sono $60 \cdot 8.000 = 480.000$, sono richiesti 480 Kbyte.

Segnali Deterministici e Probabilistici

Supponiamo di avere una sorgente, cioè una scatola nera che, opportunamente sollecitata, produca un segnale $f(t)$. Questo ci permette di riprodurre il segnale $f(t)$ in modo deterministico: una sollecitazione alla sorgente produrrà sempre lo stesso segnale.

Esistono tuttavia situazioni in cui il precedente modello non risulta adeguato. Per esempio in molti casi non c'è un modo univoco per rappresentare una informazione con un segnale: la stessa parola sarà realizzata in diversi modi da diverse persone o addirittura dalla stessa persona in diverse circostanze. Analogamente, il rumore prodotto dal passaggio di automobili su una strada pubblica non è modellabile da una sorgente deterministica, essendo soggetto a cambiamenti imprevedibili.

In questi casi potremo solo costruire una sorgente che, sollecitata, produce una realizzazione r scelta arbitrariamente in un insieme di realizzazioni possibili R , dando luogo al segnale $f(t, r)$. La sorgente si comporta in modo non deterministico perchè, sollecitata, può produrre segnali diversi. Va da sé che il trattamento di segnali prodotti da sorgenti non-deterministiche è soggetto a pesanti margini di incertezza. Fortunatamente, è spesso possibile assegnare la probabilità che la scatola, sollecitata, produca una data realizzazione: parleremo in questi casi di segnali probabilistici. In questa situazione, i margini di incertezza permangono, ma almeno è possibile quantificarli.

I segnali probabilistici, di grande importanza nella modellazione del rumore, non saranno trattati in questo testo.

1.3 Esempi di segnali

Un segnale temporale è descritto da una funzione $y = f(t)$. Qui di seguito introduciamo alcune famiglie di funzioni capaci di descrivere importanti tipi di segnali.

Segnali sinusoidali

Si consideri il cerchio goniometrico, cioè il cerchio di raggio 1 centrato nell'origine del piano cartesiano (figura 1.4).

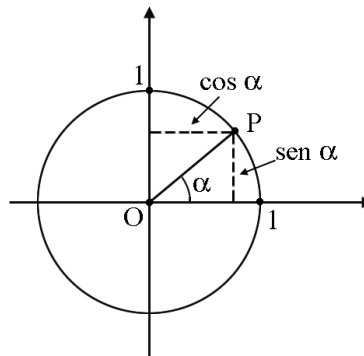


Figura 1.4 Definizione di seno e coseno di un angolo.

Consideriamo un angolo con vertice nell'origine e il cui primo lato coincida con il semi-asse positivo delle ascisse; nella figura con P si denota l'intersezione tra la circonferenza e il secondo lato dell'angolo. L'ampiezza α dell'angolo, in radianti, è la lunghezza dell'arco di cerchio intercettato dall'angolo, mentre il *coseno* di α e il *seno* di α sono rispettivamente la ascissa e l'ordinata di P .

Consideriamo la funzione $\sin \alpha$ che associa ad ogni angolo α il valore del proprio seno. Tale funzione è definita su tutto l'asse reale, è a valori nell'intervallo $[-1, 1]$ ed è continua. Essa è inoltre una funzione dispari, cioè $\sin(-\alpha) = -\sin \alpha$ e, per angoli α piccoli, vale l'approssimazione $\sin \alpha \approx \alpha$. Infine, tale funzione è periodica di periodo 2π , cioè per ogni α vale che $\sin \alpha = \sin(\alpha + 2\pi)$, ed ha un comportamento tipicamente oscillante (figura 1.5).

Considerazioni analoghe valgono per la funzione $\cos \alpha$, che associa ad ogni angolo α il valore del proprio coseno. Basta osservare che

$$\cos(\alpha) = \sin(\alpha + \pi/2)$$

La funzione $\cos$ non è altro che la funzione $\sin$ traslata. Essa risulta una funzione pari, cioè $\cos(-\alpha) = \cos \alpha$

Consideriamo ora la famiglia dei *segnali sinusoidali*, data dalla seguente espressione:

$$A \sin\left(2\pi \frac{t}{T} + \theta\right)$$

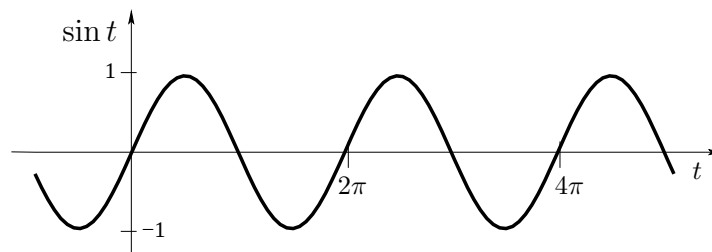


Figura 1.5 Funzione seno.

dove i parametri A, T e θ hanno il seguente significato:

- A è l'*ampiezza* e indica il modulo del valore massimo e del valore minimo assunto dal segnale.
- T è il periodo; la funzione assume infatti lo stesso valore dopo un tempo T , come si verifica osservando che, per $t = T$, $2\pi(t/T)$ vale 2π e ricordando che la funzione $\sin \alpha$ è periodica proprio di periodo 2π .
- θ è la *fase iniziale* e indica la fase al tempo $t = 0$.

Nella figura 1.6 il rappresentato il segnale sinusoidale $f(t) = A \sin(2\pi(t/T) + \theta)$ evidenziando il significato dei parametri A, T e θ .

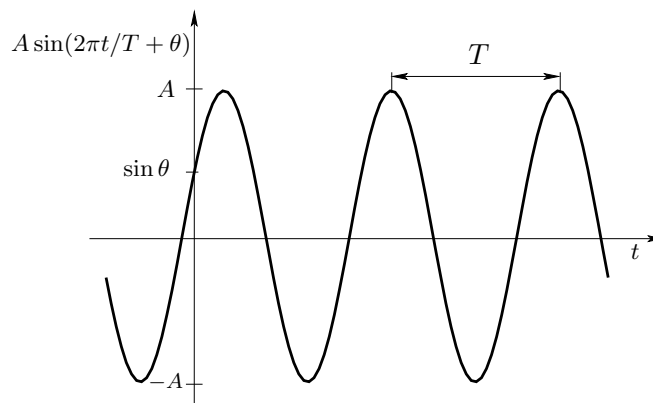


Figura 1.6 Funzione sinusoidale di periodo T , ampiezza A e fase iniziale θ

Ricordiamo infine che la *frequenza* è l'inverso del periodo, e denota il numero di oscillazioni effettuate in una unità di tempo. Detta ν la frequenza, vale che $\nu = 1/T$; la famiglia di funzioni può essere allora riscritta come

$$A \sin(2\pi\nu t + \theta)$$

Segnali seno cardinale

La funzione seno cardinale $\text{sinc } t$ è definita come segue:

$$\text{sinc}(t) = \frac{\sin(\pi t)}{\pi t}.$$

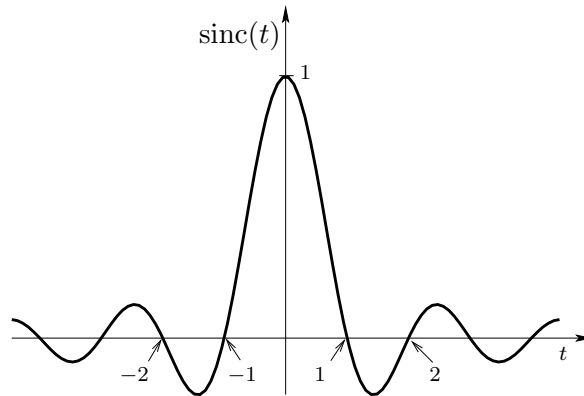


Figura 1.7 Funzione sinc.

Ponendo $\text{sinc } 0 = 1$, essa risulta definita su tutto l'asse reale e continua; si osservi infatti che, per valori piccoli di t , risulta $\pi t \approx \sin \pi t$, e quindi $\text{sinc } t \approx 1$. Essa è una funzione pari, cioè $\text{sinc}(-t) = \text{sinc}(t)$, e, pur non essendo periodica, ha un comportamento oscillante (vedi figura 1.7).

La famiglia di segnali sinc di ampiezza A e frequenza k è data da $A \text{sinc}(kt)$.

Nella figura 1.8 viene mostrato il comportamento di un segnale sinc al variare del parametro K .

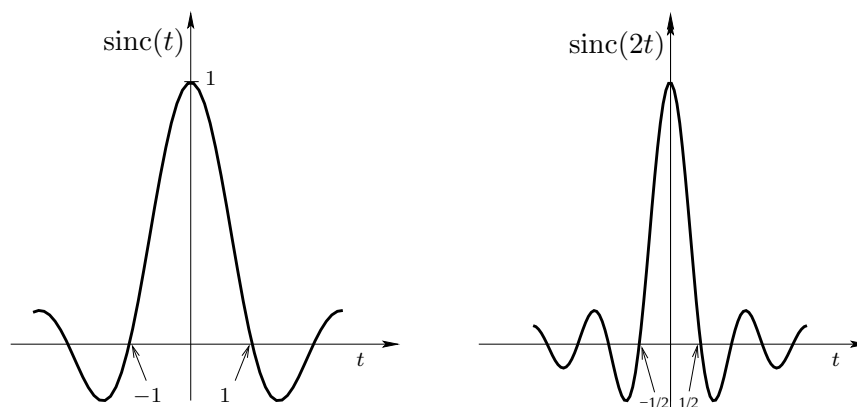


Figura 1.8 Funzioni sinc con $K = 1$ e $K = 2$.

Segnali rettangolari

La funzione *Rettangolo* è definita come segue:

$$\text{rect}(t) = \begin{cases} 1 & |t| \leq 1/2 \\ 0 & |t| > 1/2 \end{cases}$$

e il suo grafico è mostrato in figura 1.9(a).

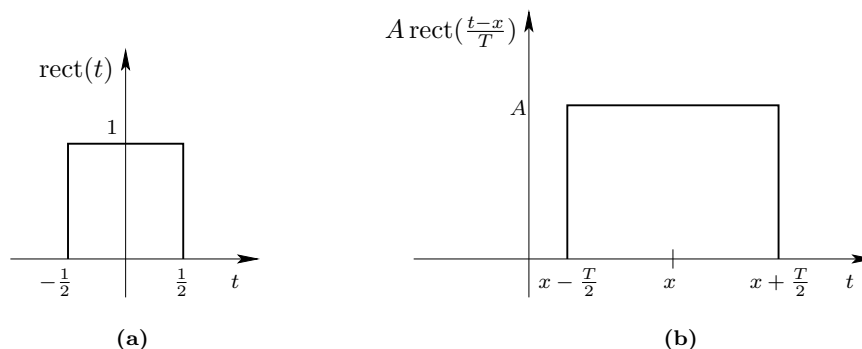


Figura 1.9 Segnali rettangolari di diversa ampiezza e durata.

Il nome è dovuto al fatto che il suo grafico è identificabile dal rettangolo di base 1 e di altezza 1, centrato intorno allo 0. Esso rappresenta lo scalino unitario ed ha due punti di discontinuità in $t = \pm \frac{1}{2}$.

La famiglia di *segnali rettangolari* è data da

$$A \cdot \text{rect}\left(\frac{t-x}{T}\right)$$

Fissati i parametri A , T e x , il grafico della funzione $f(t) = A \text{rect}(\frac{t-x}{T})$ risulta essere un rettangolo di base T , di altezza A , centrato attorno a x (figura 1.9(b)).

Si osservi che la somma di funzioni rettangolari dà luogo a *funzioni a gradino*, e che ogni funzione a gradino può essere ottenuta come somma di funzioni rettangolari.

Segnali impulsivi

La *funzione impulso* (o delta di Dirac) $\delta(t)$ può essere intuitivamente pensata come un rettangolo di base ‘infinitesima’ Δ e di altezza ‘infinita’ $\frac{1}{\Delta}$, in modo che l’area sottostante $\int_{-\infty}^{+\infty} \delta(t) dt$ sia 1. Essa non è quindi una funzione (nel senso di Dirichelet), e sarà vista come funzione generalizzata: infatti, mentre ha senso dire che $\delta(t) = 0$ se $t \neq 0$, non ha senso dire che $\delta(0) = +\infty$ poiché ∞ non è un numero.

La funzione impulso $\delta(t)$ può essere pensata come il seguente limite di rettangoli (vedi figura 1.10(a)):

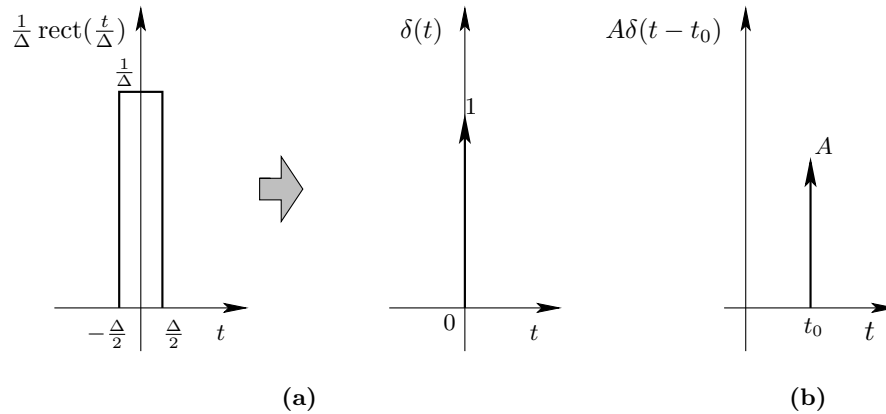


Figura 1.10 Impulso per la generazione della funzione delta di Dirac.

$$\delta(t) = \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} \cdot \text{rect}\left(\frac{t}{\Delta}\right).$$

Ne risulta che per la funzione $\delta(t)$ valgono le seguenti proprietà:

$$\delta(t) = 0 \text{ se } t \neq 0,$$

$$\int_{-\infty}^{\infty} \delta(t) dt = 1.$$

Fissato un reale t_0 , la funzione generalizzata $\delta(t - t_0)$ è interpretabile come impulso al tempo t_0 (vedi figura 1.10(b)) e valgono le seguenti proprietà:

$$\delta(t - t_0) = 0 \text{ se } t \neq t_0,$$

$$\int_{-\infty}^{\infty} \delta(t - t_0) dt = 1.$$

Si osservi infine che, fissato un segnale a tempo continuo $f(t)$, il segnale a tempo discreto, ottenuto fissando un intervallo temporale τ e considerando i valori di $f(t)$ sui multipli interi $n\tau$ di τ , può essere descritta da:

$$\sum_n f(n\tau) \delta(t - n\tau)$$

Esempi di segnali a tempo discreto

Si fissi un intervallo τ . Dato un segnale $f(t)$, il segnale a tempo discreto $f(n\tau)$ è rappresentato dalla funzione $x(n) = f(n\tau)$, dove n è il tempo normalizzato; fissato l'intervallo

τ , i segnali a tempo discreto sono quindi identificati da successioni che assegnano ad ogni intero n un valore $x(n)$.

Un importante segnale è l'*impulso unitario* $\delta(n)$, che rappresenta nel tempo discreto ciò che l'impulso rappresenta nel tempo continuo. Tale segnale vale 1 per $n = 0$ e 0 altrimenti, cioè:

$$\delta(n) = \begin{cases} 1, & n=0; \\ 0, & n \neq 0. \end{cases}$$

La famiglia

$$A \cdot \delta(n - a)$$

è formata dagli impulsi unitari di ampiezza A traslati al tempo a . Una importante proprietà dell'impulso unitario traslato è la seguente:

$$x(n)\delta(n - a) = x(a)\delta(n - a)$$

Questa proprietà discende dal fatto che, per $n \neq a$, risulta $x(n)\delta(n - a) = 0 = x(a)\delta(n - a)$.

Ogni segnale può essere espresso come somma di impulsi unitari traslati con opportune ampiezze: ad esempio, si consideri il segnale $y(n)$ tale che $y(2) = 3$ e $y(5) = -1$, $y(n) = 0$ per $n \neq 2$ e $n \neq 5$. Tale segnale è ottenibile come:

$$y(n) = 3\delta(n - 2) - \delta(n - 5)$$

Un altro interessante segnale a tempo discreto è il gradino unitario $u(n)$, definito da:

$$u(n) = \begin{cases} 1, & n \geq 0; \\ 0, & n < 0. \end{cases}$$

Tale segnale, espresso come somma di impulsi unitari, vale:

$$u(n) = \sum_{a=0}^{\infty} \delta(n - a)$$

1.4 Sistemi per l'Elaborazione dei Segnali Deterministici

Consideriamo ora due famiglie $\mathcal{F}_1$ ed $\mathcal{F}_2$ di segnali. Con *sistema* intendiamo un qualsiasi apparato in grado di trasformare segnali di $\mathcal{F}_1$ in segnali di $\mathcal{F}_2$. Il sistema può allora essere visto come una scatola nera il cui comportamento è la legge di trasformazione S che ad ogni segnale di $\mathcal{F}_1$ associa il corrispondente segnale di $\mathcal{F}_2$:

$$S : \mathcal{F}_1 \rightarrow \mathcal{F}_2.$$

In Figura 1.11 è data la rappresentazione grafica di un sistema che, avendo in ingresso il segnale $f(t)$, dà in uscita il segnale $g(t)$; si dice anche che $g(t)$ è la *risposta del sistema* S all'ingresso $f(t)$ e si scrive $g(t) = S(f(t))$.

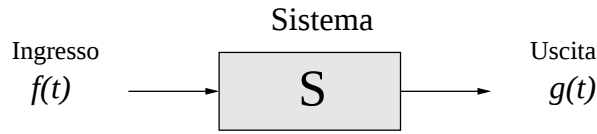


Figura 1.11 Rappresentazione grafica di un sistema.

Esempio 1.4.1

Semplici esempi di sistemi sono quelli che permettono di ottenere la amplificazione, la traslazione temporale (ritardo), ed il cambio di scala di segnali dati (Figura 1.12).

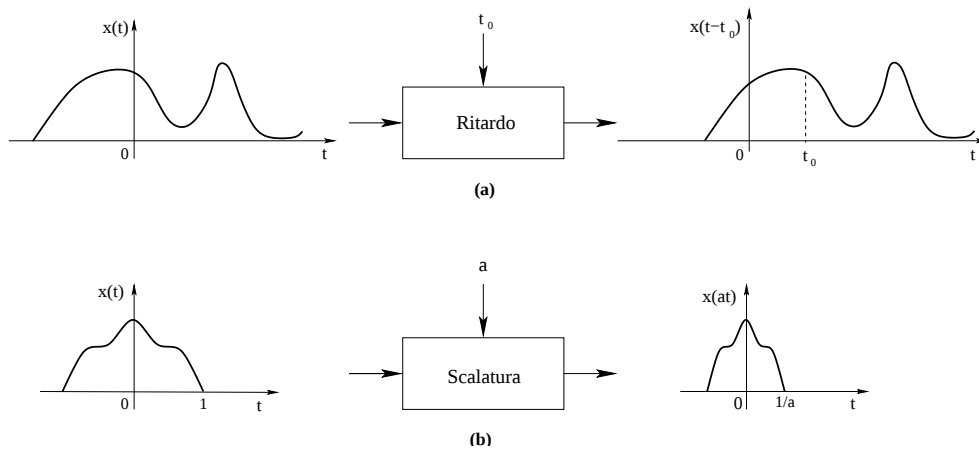


Figura 1.12 (a) Ritardo. (b) Scalatura.

1. L'amplificazione trasforma il segnale $f(t)$ nel segnale $Af(t)$, che è lo stesso segnale amplificato di un fattore A .
2. La traslazione (o ritardo) trasforma il segnale $f(t)$ nel segnale $f(t - t_0)$, che è lo stesso segnale ritardato di un tempo t_0 .
3. Il cambio di scala nel tempo (o scalatura) trasforma il segnale $f(t)$ nel segnale $f(at)$. L'effetto che si ottiene in questo caso è quello di una compressione lineare (se $|a| > 1$) o di un allungamento o rilassamento lineare (se $|a| < 1$).

Esempio 1.4.2

Un importante esempio di sistema è il Campionatore a periodo τ (o equivalentemente a frequenza di campionamento $F_s = 1/\tau$), illustrato in Figura 1.13.

Esso trasforma il segnale a tempo continuo $f(t)$ nel segnale a tempo discreto $f(n\tau)$, che può essere descritto da:

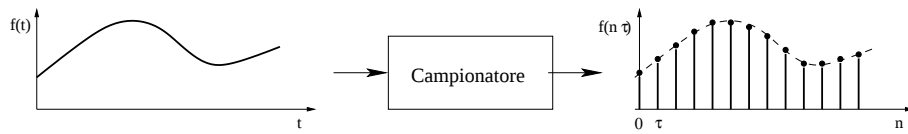


Figura 1.13 Campionatore.

$$\sum_{n=-\infty}^{+\infty} f(n\tau)\delta(t - n\tau)$$

In seguito sarà affrontato l'importante problema di determinare condizioni per cui il segnale $f(t)$ è ricostruibile a partire dal segnale campionato $f(n\tau)$.

Esempio 1.4.3

Dato un insieme finito $V = \{x_1, \dots, x_m\}$ di numeri, consideriamo la funzione Q che ad ogni reale x associa l'elemento in V più vicino ad x , cioè:

$$Q(x) = \arg \min_{x_k \in V} |x - x_k|$$

Il sistema quantizzatore associa al segnale $f(t)$ il segnale $Q(f(t))$. Tale sistema riceve in ingresso un segnale a valori continui e restituisce un segnale a valori finiti. Un esempio di quantizzatore a 2 valori $\{1, -1\}$ è dato in Figura 1.14.

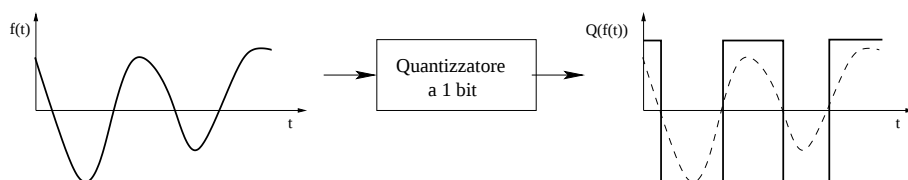


Figura 1.14 Quantizzatore.

In questo caso:

$$Q(x) = \begin{cases} 1 & \text{se } x > 0 \\ -1 & \text{se } x \leq 0 \end{cases}$$

La funzione Q coincide quindi con la funzione $\text{Signum}(x)$.

1.4.1 Composizionalità nei Sistemi

Un sistema complesso può essere ottenuto *combinando assieme* vari sistemi semplici, che ne costituiscono le componenti; per poter trattare con sistemi complessi, è di grande importanza riuscirne a comprenderne il comportamento a partire da quello delle loro componenti. Un approccio che rende agevole questo obbiettivo è quello modulare: si introducono esplicitamente operazioni che trasformano due o più sistemi in un nuovo sistema, e la costruzione di un sistema complesso viene ottenuta applicando queste operazioni a partire da sistemi semplici.

Senza pretesa di voler essere nemmeno lontanamente esaurienti, introduciamo in questa sezione le operazioni di *composizione sequenziale* (o *cascata*), *composizione parallela*, *retroazione* e *inversione*, mostrando su semplici esempi la capacità di astrazione fornita dall'approccio modulare.

Composizione sequenziale o cascata. Dati due sistemi $S_1 : \mathcal{F}_1 \rightarrow \mathcal{F}_2$ e $S_2 : \mathcal{F}_2 \rightarrow \mathcal{F}_3$, la loro composizione è il sistema $S_3 : \mathcal{F}_1 \rightarrow \mathcal{F}_3$, tale che $S_3(f(t)) = S_2(S_1(f(t)))$ ottenuta ponendo in ingresso a S_2 la risposta di S_1 .

Composizione parallela. Dati due sistemi $S_1 : \mathcal{F}_1 \rightarrow \mathcal{F}_2$ e $S_2 : \mathcal{F}_1 \rightarrow \mathcal{F}_2$, la loro composizione parallela è il sistema che ha come risposta la somma delle risposte di S_1 e S_2 .

Retroazione. Dati due sistemi $S_1 : \mathcal{F}_1 \rightarrow \mathcal{F}_2$ e $S_2 : \mathcal{F}_2 \rightarrow \mathcal{F}_3$, il sistema ottenuto per retroazione è il sistema $\S_3$ che ha su ingresso f una uscita g ottenuta ponendo in ingresso a S_1 la differenza tra f e la risposta di S_2 a g .

Le operazioni introdotte sono illustrate in Figura 1.15.

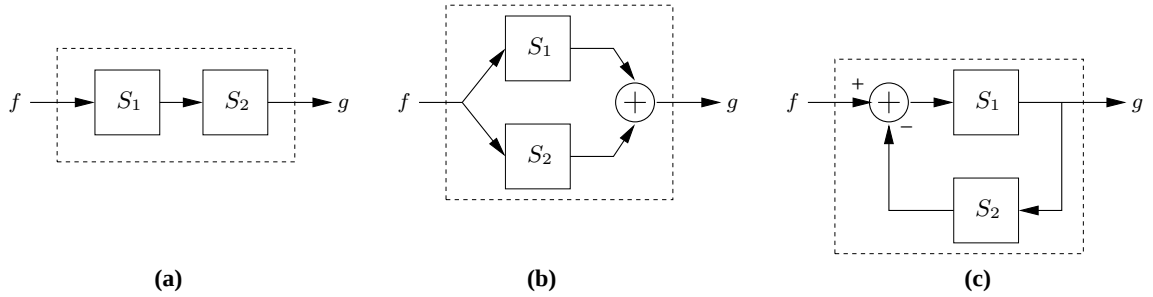


Figura 1.15 (a) Cascata. (b) Composizione parallela. (c) Retroazione.

Un'ulteriore operazione generale per sistemi è la *inversione*. Essa si applica a sistemi $S : \mathcal{F}_1 \rightarrow \mathcal{F}_2$ per cui la relazione ingresso uscita è una corrispondenza biunivoca.

Inversione. Dato un sistema $S : \mathcal{F}_1 \rightarrow \mathcal{F}_2$, il sistema inverso è il sistema $\bar{S} : \mathcal{F}_2 \rightarrow \mathcal{F}_1$ tale che, se $g(t) = S(f(t))$, allora per ogni $g(t)$, vale che $\bar{S}(g(t)) = f(t)$, o almeno $\bar{S}(g(t)) \approx f(t)$.

Esempio 1.4.4

Conversione analogico-digitale. Dalla composizione sequenziale di un campionatore e di un quantizzatore si ottiene un convertitore analogico digitale (ADC), che trasforma un segnale analogico in un segnale digitale (vedi Figura 1.16).

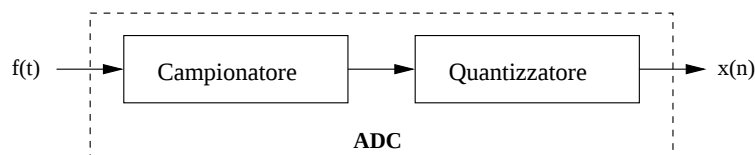


Figura 1.16 Conversione analogico-digitale.

Esempio 1.4.5

Elaborazione digitale di segnali analogici e trasmissione di segnali digitali. Un convertitore analogico-digitale (ADC) è un sistema che trasforma, in modo approssimativamente fedele, segnali analogici in segnali digitali.

L'operazione inversa è compiuta dal convertitore digitale-analogico (DAC): esso trasforma un segnale digitale in un segnale analogico in modo tale che, se $y(n) = \text{ADC}(f(t))$, allora $\text{DAC}(y(n)) \approx f(t)$. Gli ADC e i DAC realizzano dunque l'interfaccia tra il mondo analogico e quello digitale. La composizione sequenziale presentata in Figura 1.17 modella il trattamento di segnali analogici con programmi eseguiti da un elaboratore digitale, mentre quella presentata in Figura 1.18 modella la trasmissione a distanza di segnali digitali utilizzando un supporto analogico.

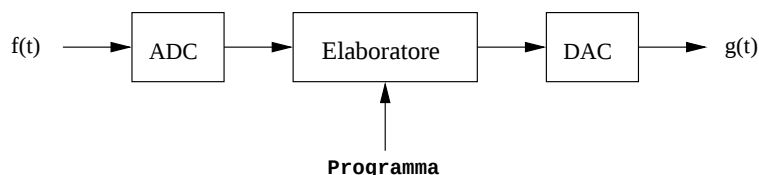


Figura 1.17 Elaborazione digitale dei segnali analogici.



Figura 1.18 Trasmissione di segnali digitali su supporto analogico.

Esempio 1.4.6

Il controllo digitale di un dispositivo analogico richiede, dato un sistema analogico da controllare (controllato), di determinare un opportuno sistema digitale (controllore) tale che l'interazione tra i due sistemi permetta di ottenere una relazione ingresso-uscita desiderata. Nella Figura 1.19 viene rappresentato un semplice modello per il controllo digitale di un dispositivo analogico; una tipica applicazione è quella del controllo motore. L'algoritmo di controllo riceve in ingresso un segnale digitale di errore; il segnale d'uscita viene trasformato in analogico e processato dal sistema controllato. L'uscita del sistema controllato, convertita in digitale, viene posta in retroazione concorrendo alla formazione del segnale di errore.

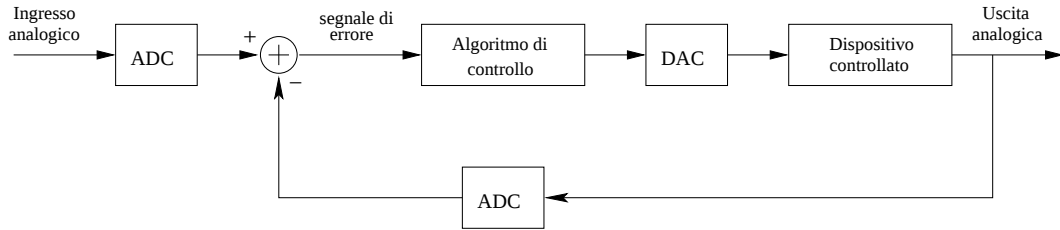


Figura 1.19 Modello di controllo digitale di dispositivi analogici.

1.5 Spazi vettoriali di segnali

Due importanti operazioni tra segnali sono la sovrapposizione e l'amplificazione. In questa sezione prendiamo in considerazione l'algebra di queste operazioni (spazi vettoriali) ricavando alcuni risultati, di grande interesse sia per problemi di rappresentazione dei segnali, quali la loro rappresentazione in frequenza, sia per i problemi di analisi e progettazione dei filtri, che rappresentano una rilevante classe di sistemi.

Dati due segnali $f_1(t)$ e $f_2(t)$, la loro *somma* è il segnale $f(t)$ ottenuto dalla sovrapposizione di $f_1(t)$ e $f_2(t)$ così definito:

$$f(t) = f_1(t) + f_2(t)$$

Dato il segnale $f_1(t)$ e un numero reale a , il prodotto di $f_1(t)$ per lo *scalare* a è il segnale $f(t)$ ottenuto amplificando $f_1(t)$ di un fattore a , cioè:

$$f(t) = af_1(t)$$

Studiamo ora alcune proprietà della struttura algebrica ottenuta considerando queste operazioni sui segnali.

L'operazione di somma è commutativa e associativa; questo significa che per tutti i segnali $f(t)$, $g(t)$ e $h(t)$, vale:

$$\begin{aligned} f(t) + g(t) &= g(t) + f(t) \\ (f(t) + g(t)) + h(t) &= f(t) + (g(t) + h(t)) \end{aligned}$$

Esiste inoltre un segnale 0 (il segnale ovunque nullo) tale che, per ogni segnale $f(t)$:

$$f(t) + 0 = f(t)$$

e, per ogni segnale $f(t)$, esiste il segnale opposto $-f(t)$ tale che

$$f(t) - f(t) = 0$$

Vale infine che

$$\begin{aligned} a(bf(t)) &= abf(t) \text{ (} a \text{ e } b \text{ sono scalari)} \\ a(f_1(t) + f_2(t)) &= af_1(t) + af_2(t) \end{aligned}$$

La struttura algebrica, composta dall'insieme dei segnali con le operazioni di somma e prodotto per uno scalare, è quella di *spazio vettoriale*, in quanto verifica le proprietà precedentemente esposte. Combinando le operazioni di prodotto per uno scalare e di somma si ottiene la cosiddetta *combinazione lineare*:

Combinazione lineare. Data una famiglia $\{f_1(t), \dots, f_r(t)\}$ di segnali, la somma pesata

$$f(t) = a_1 f_1(t) + a_2 f_2(t) + \dots + a_r f_r(t) = \sum_{k=1}^r a_k f_k(t)$$

è detta *combinazione lineare* dei segnali $f_1(t), \dots, f_r(t)$.

Per ottenere la combinazione lineare dei segnali $f_1(t), \dots, f_r(t)$ basta dunque amplificare il generico segnale $f_k(t)$ di un fattore a_k e sovrapporre i segnali ottenuti. Il coefficiente a_k è detto anche *peso* di $f_k(t)$ nella combinazione lineare.

Consideriamo un sottospazio vettoriale $\mathcal{F}$ di segnali, contenente la famiglia di segnali $\{f_1(t), \dots, f_r(t)\}$. Se ogni segnale $f(t) \in \mathcal{F}$ può essere ottenuto come combinazione lineare di segnali della famiglia $\{f_1(t), \dots, f_r(t)\}$, cioè se per ogni $f(t)$ esistono coefficienti $a_1, \dots, a_r$ per cui:

$$f(t) = \sum_{k=1}^r a_k f_k(t)$$

allora $\{f_1(t), \dots, f_r(t)\}$ è detto *insieme di generatori* di $\mathcal{F}$. I segnali $f_1(t), \dots, f_r(t)$ sono invece detti *linearmente indipendenti* se nessuno di essi può essere ottenuto come combinazione lineare dei rimanenti.

Esempio 1.5.1

In questo esempio prendiamo in considerazione i segnali a tempo discreto n . Dati gli interi $a_1 < a_2 < \dots < a_k$, gli impulsi unitari traslati $\delta(n - a_1), \dots, \delta(n - a_k)$ sono segnali

linearmente indipendenti. Consideriamo infatti un generico segnale $\delta(n - a_j)$. Se per assurdo fosse $\delta(n - a_j) = \sum_{s \neq j} A_s \delta(n - a_s)$, ponendo $n = a_j$ sarebbe:

$$1 = \delta(0) = \sum_{s \neq j} A_s \delta(a_j - a_s) = \sum_{s \neq j} A_s \cdot 0 = 0$$

Quindi $\delta(n - a_j)$ non può essere espresso come combinazione lineare dei rimanenti.

Esempio 1.5.2

In questo esempio prendiamo in considerazione i segnali a tempo continuo t . Dati i reali positivi $0 < \nu_1 < \nu_2 < \dots < \nu_k$, si può provare che la famiglia

$$\{\sin 2\pi\nu_1 t, \dots, \sin 2\pi\nu_k t, \cos 2\pi\nu_1 t, \dots, \cos 2\pi\nu_k t\}$$

di segnali sinusoidali di frequenza $\nu_1, \nu_2, \dots, \nu_k$ è formata da segnali linearmente indipendenti.

I generatori linearmente indipendenti portano all'importante concetto di *base*:

Base di segnali. Dato un sottospazio vettoriale $\mathcal{F}$ di segnali, una famiglia di segnali $\{f_1(t), \dots, f_r(t)\}$ di $\mathcal{F}$ è una *base* se è un insieme di generatori di $\mathcal{F}$ formata da segnali linearmente indipendenti.

Una base è essenzialmente una famiglia $\{f_1(t), \dots, f_r(t)\}$ di generatori di $\mathcal{F}$ che risulta essere minimale. Questo significa che, togliendo alla famiglia $\{f_1(t), \dots, f_r(t)\}$ un qualsiasi suo elemento $f_k(t)$, il nuovo insieme $\{f_1(t), \dots, f_{k-1}(t), f_{k+1}(t), \dots, f_r(t)\}$ non è più in grado di generare $\mathcal{F}$.

Per le basi, cioè per gli insiemi minimali di generatori, valgono le seguenti proprietà:

1. Uno spazio vettoriale $\mathcal{F}$ può avere più basi, ma tutte le basi contengono lo stesso numero di elementi; tale numero è detto *dimensione* dello spazio.
2. Data una base $\{f_1(t), \dots, f_r(t)\}$, per ogni $f(t) \in \mathcal{F}$ esiste una sola sequenza di coefficienti F_k tali che:

$$f(t) = \sum_{k=1}^r F_k f_k(t)$$

Risulta allora che $f(t)$ è biunivocamente individuato dalla sequenza dei coefficienti F_k . La funzione $f(t)$ e la sequenza F_k contengono allora le stesse informazioni perchè, data una delle due, è possibile ricostruire l'altra: da questo punto di vista, $F_1, \dots, F_r$ è una *rappresentazione* alternativa di $f(t)$.

In seguito, supporremo che queste nozioni siano estese al caso di basi con infiniti elementi: fissata una base $\{f_1(t), \dots, f_k(t), \dots\}$ di $\mathcal{F}$ per ogni $f(t) \in \mathcal{F}$ è possibile individuare una e una sola sequenza di coefficienti F_k tali che:

$$f(t) = \sum_{k=1}^{\infty} F_k f_k(t)$$

Esempio 1.5.3

In questo esempio prendiamo in considerazione i segnali a tempo discreto n . Una base per tali segnali è ottenuta considerando tutti gli impulsi traslati

$$\{\delta(n-a) \mid -\infty < a < +\infty\}$$

Abbiamo infatti già visto che tali segnali sono linearmente indipendenti; per mostrare che ogni segnale è ottenibile come combinazione lineare di impulsi traslati, basta osservare che, fissato un qualsiasi segnale $x(n)$, vale:

$$x(n) = \sum_{a=-\infty}^{\infty} x(a) \delta(n-a)$$

Per dimostrare la precedente identità, basta osservare che, per $n = a$, vale $\delta(n-a) = 1$, mentre, per $n \neq a$, vale $\delta(n-a) = 0$

Esempio 1.5.4

In analogia al caso precedente, nello spazio dei segnali a tempo continuo, l'insieme delle funzioni impulsive traslate (delta di Dirac) $\{\delta(t-x) : x \in R\}$, presentata nella sezione 1.3, identifica una base. Infatti ogni segnale $f(t)$ può essere espresso come combinazione lineare generalizzata di impulsi traslati, e il coefficiente moltiplicativo di $\delta(t-x)$ è proprio $f(x)$, ovvero:

$$f(t) = \int_{-\infty}^{\infty} f(x) \delta(t-x) dx.$$

Questo può essere dimostrato osservando per prima cosa che $f(t)\delta(t-x) = f(x)\delta(t-x)$, poiché $\delta(t-x) = 0$ per $x \neq t$. Allora:

$$\int_{-\infty}^{+\infty} f(x) \delta(t-x) dx = \int_{-\infty}^{+\infty} f(t) \delta(t-x) dx = f(t) \int_{-\infty}^{+\infty} \delta(t-x) dx = f(t).$$

1.5.1 Basi ortonormali

Abbiamo osservato che, data una base $\{f_1(t), \dots, f_k(t), \dots\}$, ogni segnale $f(t)$ è ottenibile come

$$f(t) = \sum_k F_k f_k(t)$$

dove i coefficienti F_k sono univocamente determinati, e quindi identificano una rappresentazione alternativa del segnale $f(t)$. In questo ambito, è di grande rilievo il seguente problema:

Problema di rappresentazione: Data una base $\{f_1(t), \dots, f_k(t), \dots\}$ per uno spazio di segnali $\mathcal{F}$, e un segnale $f(t) \in \mathcal{F}$, determinare i coefficienti $F_1, \dots, F_k, \dots$ tali che:

$$f(t) = \sum_k F_k f_k(t)$$

Questo problema ammette una soluzione semplice ed efficace quando le basi sono *ortonormali*.

Per introdurre la nozione di ortonormalità di una base, consideriamo nello spazio di segnali una operazione $\circ$, che chiameremo *prodotto interno*. Il prodotto interno è una operazione binaria, che associa ad ogni coppia di segnali $f(t)$ e $g(t)$ il numero reale $f(t) \circ g(t)$. Si suppone inoltre che l'operazione $\circ$ verifichi le seguenti proprietà:

1. se $f(t)$ non è il segnale nullo, allora $f(t) \circ f(t) > 0$
2. $(af(t) + bg(t)) \circ h(t) = a(f(t) \circ h(t)) + b(g(t) \circ h(t))$

Esempio 1.5.5

Consideriamo lo spazio $\mathbb{R}^n$ formato da n -uple di numeri reali, cioè $\mathbb{R}^n = \{(x_1, \dots, x_n), x_k \in \mathbb{R}\}$, con le operazioni:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

$$a(x_1, \dots, x_n) = (ax_1, \dots, ax_n)$$

Questo spazio è uno spazio vettoriale, e, posto $X = (x_1, \dots, x_n)$ e $Y = (y_1, \dots, y_n)$, l'operazione

$$X \circ Y = \sum_{i=1}^n x_i y_i$$

è un prodotto interno.

Osserviamo che $\|X\| = \sqrt{X \circ X}$ è la lunghezza del vettore X ed è possibile mostrare che:

$$X \circ Y = \|X\| \cdot \|Y\| \cdot \cos \theta$$

dove θ è l'angolo tra il vettore X e Y .

Esempio 1.5.6

Per segnali continui definiti sull'intervallo $[0, 1]$ la seguente operazione è un prodotto interno:

$$f(t) \circ g(t) = \int_0^1 f(t)g(t)dt$$

Esempio 1.5.7

Per i segnali a tempo discreto (con opportune condizioni che garantiscono la convergenza) la seguente operazione è un prodotto interno

$$f(n) \circ g(n) = \sum_{k=-\infty}^{+\infty} f(k)g(k)$$

Consideriamo ora uno spazio di segnali $\mathcal{F}$, dotato dell'operazione di prodotto interno $g(t) \circ h(t)$, ed una sua base $\{f_1(t), \dots, f_k(t), \dots\}$. La base $\{f_1(t), \dots, f_k(t), \dots\}$ è detta *ortonormale* se vale:

1. $f_i(t) \circ f_j(t) = 0$ se $i \neq j$
2. $f_i(t) \circ f_i(t) = 1$ per ogni i

A giustificazione della precedente definizione, si osservi che in esempio 1.5.5, il prodotto interno $X \circ Y$ tra vettori X e Y è 0 se l'angolo θ tra X e Y è tale che $\cos \theta = 0$; questo succede quando i vettori X e Y sono ortogonali. Usando questa analogia, possiamo dire che gli elementi di una base ortonormale hanno tutti lunghezza 1 e ognuno è ortogonale agli altri.

Possiamo a questo punto enunciare l'importante risultato, che risolve il problema di rappresentazione quando la base è ortonormale:

Fatto 1: se $\{f_1(t), \dots, f_k(t), \dots\}$ è una base ortonormale per uno spazio di segnali $\mathcal{F}$, allora ogni segnale $f(t) \in \mathcal{F}$ è tale che:

$$f(t) = \sum_k F_k f_k(t)$$

dove

$$F_j = f(t) \circ f_j(t)$$

Il risultato precedente può essere giustificato come segue. Se $\{f_1(t), \dots, f_k(t), \dots\}$ è una base per uno spazio di segnali $\mathcal{F}$, allora esistono coefficienti $\{F_1, \dots, F_k, \dots\}$ tali che $f(t) = \sum_k F_k f_k(t)$.

Consideriamo il prodotto interno $f(t) \circ f_j(t)$. Per la proprietà di linearità del prodotto interno, vale:

$$f(t) \circ f_j(t) = \left(\sum_k F_k f_k(t) \right) \circ f_j(t) = \sum_k F_k (f_k(t) \circ f_j(t))$$

Ricordando che, se $k \neq j$ vale $f_k(t) \circ f_j(t) = 0$ e che, se $k = j$, è $f_j(t) \circ f_j(t) = 1$, possiamo concludere che:

$$\sum_k F_k (f_k(t) \circ f_j(t)) = F_j (f_j(t) \circ f_j(t)) = F_j$$

Concludiamo allora che $f(t) \circ f_j(t) = F_j$.

Il risultato precedente trova una importante applicazione nello sviluppo di Fourier, che sarà discusso in seguito. Presentiamo qui un semplice esempio di base ortonormale.

Esempio 1.5.8

Consideriamo lo spazio dei segnali $x(n)$ a tempo discreto (con la condizione aggiuntiva che $\sum_{n=-\infty}^{+\infty} x^2(n) < \infty$).

Per i segnali a tempo discreto una nozione di prodotto interno è la seguente:

$$x(n) \circ y(n) = \sum_{n=-\infty}^{\infty} x(n)y(n)$$

Si osserva immediatamente che la base $\{\delta(n-k) | k \in Z\}$ è una base ortonormale. Infatti:

$$\delta(n-k) \circ \delta(n-k) = \sum_{n=-\infty}^{\infty} \delta^2(n-k) = 1$$

se invece $k \neq j$, vale che:

$$\delta(n-k) \circ \delta(n-j) = \sum_{n=-\infty}^{\infty} (n-k)\delta(n-j) = 0$$

1.6 Sistemi Lineari

Una classe di sistemi particolarmente importante è quella dei sistemi lineari, che introduciamo in questa sezione. L'interesse per questi sistemi è duplice: da un lato esistono potenti strumenti per l'analisi e la progettazione dei sistemi lineari, dall'altro molti sistemi possono essere approssimati con sistemi lineari.

Consideriamo due spazi vettoriali di segnali $\mathcal{F}_1$ e $\mathcal{F}_2$ e un sistema S , il cui comportamento è dato da

$$S : \mathcal{F}_1 \rightarrow \mathcal{F}_2$$

Se succede che la risposta a una somma di segnali è la somma delle risposte, e che la risposta alla moltiplicazione di un segnale per uno scalare è la moltiplicazione della risposta per lo scalare, allora il sistema è detto *lineare*.

Formalmente:

Definizione 1.1 Siano $\mathcal{F}_1$ e $\mathcal{F}_2$ due spazi lineari di segnali. Un sistema $S : \mathcal{F}_1 \rightarrow \mathcal{F}_2$ è lineare se

$$\begin{aligned} S(f(t) + g(t)) &= S(f(t)) + S(g(t)), \\ S(af(t)) &= aS(f(t)). \end{aligned}$$

Esempio 1.6.1

La modulazione di ampiezza MA di un segnale $f(t)$ è realizzata moltiplicando $f(t)$ per $A \cos \omega t$, cioè:

$$\text{MA}(f(t)) = A \cos \omega t \cdot f(t).$$

Essa è descritta da un sistema lineare, infatti:

$$\begin{aligned} \text{MA}(af(t) + bg(t)) &= A \cos \omega t \cdot (af(t) + bg(t)) \\ &= aA \cos \omega t \cdot f(t) + bA \cos \omega t \cdot g(t) \\ &= a \cdot \text{MA}(f(t)) + b \cdot \text{MA}(g(t)) \end{aligned}$$

Esempio 1.6.2

La modulazione di fase MF di un segnale $f(t)$ è descritta dal sistema

$$\text{MF}(f(t)) = \rho \cos(\omega t + f(t)).$$

Poichè in generale $\cos(\omega t + (f + g)) \neq \cos(\omega t + f) + \cos(\omega t + g)$, MF non è un sistema lineare.

Consideriamo ora nello spazio dei segnali una base $\{f_1(t), \dots, f_n(t)\}$, e ricordiamo che ogni altro segnale $f(t)$ può essere ottenuto in un unico modo come combinazione lineare $f(t) = \sum_i F_i f_i(t)$ di elementi della base.

Ogni sistema lineare S possiede la proprietà di *sovrapposizione*: se l'ingresso $f(t) = \sum_{k=1}^n F_k f_k(t)$ è una combinazione lineare dei segnali $f_1(t), \dots, f_n(t)$, allora la risposta $g(t)$ è la combinazione lineare delle risposte $S(f_1(t)), \dots, S(f_n(t))$, vale a dire:

$$S(f(t)) = S\left(\sum_{k=1}^n F_k f_k(t)\right) = \sum_{k=1}^n F_k S(f_k(t))$$

Conoscendo quindi la risposta $S(f_i(t))$ del sistema ad ogni elemento $f_i(t)$ della base, riusciamo a determinare la risposta $S(f(t))$ del sistema per tutti i segnali $f(t)$ dello spazio.

Queste considerazioni, introdotte per spazi a base finita, possono essere estese (con qualche precauzione) a spazi più generali. I seguenti esempi si riferiscono alla base formata dagli impulsi unitari traslati per segnali a tempo discreto, e alla base indicata su un continuo, data dagli impulsi traslati, per segnali a tempo continuo.

Esempio 1.6.3

Ricordiamo che una base per i segnali a tempo discreto è data da tutte le traslazioni dell'impulso unitario $\{\delta(n-a) : a \in \mathbf{Z}\}$. In particolare, per ogni segnale $x(n)$ risulta:

$$x(n) = \sum_{k=-\infty}^{+\infty} x(k)\delta(n-k)$$

Se S è un sistema lineare, possiamo concludere che:

$$S(x(n)) = S\left(\sum_{k=-\infty}^{+\infty} x(k)\delta(n-k)\right) = \sum_{k=-\infty}^{+\infty} x(k)S(\delta(n-k))$$

Se chiamiamo $m(n, k)$ la risposta di S all'impulso traslato $\delta(n-k)$, vale:

$$S(x(n)) = S\left(\sum_{k=-\infty}^{+\infty} x(k)m(n, k)\right)$$

Esempio 1.6.4

Ricordiamo che, in analogia col caso precedentemente trattato, una base per i segnali a tempo discreto è data da tutte le traslazioni $\{\delta(t-x) : x \in R\}$ della funzione impulsiva $\delta(t)$. In particolare, per ogni segnale $f(t)$ risulta:

$$f(t) = \int_{-\infty}^{+\infty} f(x)\delta(t-x)dx.$$

Se S è un sistema lineare, possiamo concludere che:

$$S(f(t)) = S\left(\int_{-\infty}^{+\infty} f(x)\delta(t-x)dx\right) = \int_{-\infty}^{+\infty} f(x)S(\delta(t-x))dx.$$

Se chiamiamo $m(t, x)$ la risposta di S alla funzione impulsiva traslata $\delta(t-x)$, vale:

$$S(f(t)) = \int_{-\infty}^{+\infty} f(x)m(t, x)dx.$$

1.6.1 Sistemi Tempo-Invarianti

Per molti sistemi è importante che le caratteristiche del sistema stesso non si modifichino nel corso del tempo: ruotare il volante di un'automobile in senso orario deve produrre l'effetto di far girare la macchina a destra, in qualsiasi istante la rotazione venga applicata. Intuitivamente, questo significa che la risposta del sistema a un segnale ritardato è la risposta del sistema al segnale, a sua volta ritardata. Questa nozione è formalizzata dal concetto di sistema tempo invariante:

Definizione 1.2 *Un sistema $S : \mathcal{F}_1 \rightarrow \mathcal{F}_2$ è tempo-invariante quando per ogni ingresso $f(t)$, se $g(t) = S(f(t))$ allora $g(t - t_0) = S(f(t - t_0))$, per ogni t_0 .*

Ad esempio, il pianoforte può essere visto come un sistema che trasforma la pressione delle dita in suono. Medesime esecuzioni in tempi diversi portano agli stessi suoni, naturalmente in tempi diversi. In questo senso il pianoforte è un sistema tempo-invariante. Tutto ciò è subordinato al fatto che le caratteristiche meccaniche del pianoforte non si modifichino nel tempo: quando questo accade, il pianoforte diventa un sistema tempo-variante ed è necessario l'intervento dell'accordatore.

Una classe di sistemi tempo-invarianti può essere ottenuta come segue. Fissata una funzione $G(x)$ a variabile reale, si consideri il sistema il cui comportamento è descritto da:

$$g(t) = G(f(t))$$

Si osserva che tale sistema è tempo-invariante. Infatti:

$$g(t - t_0) = g(t)|_{t-t_0} = G(f(t))|_{t-t_0} = G(f(t - t_0))$$

La linearità e l'invarianza temporale sono concetti indipendenti, come mostrano gli esempi che seguono.

Esempio 1.6.5

Il sistema $g(t) = a f(t)$, che amplifica l'ingresso di un fattore a , è tempo-invariante. Infatti la risposta al segnale ritardato $f(t - t_0)$ è $af(t - t_0)$, che coincide con $g(t - t_0)$. Si osservi che tale sistema è sia lineare che tempo-invariante.

Esempio 1.6.6

La modulazione di ampiezza MA data da:

$$g(t) = \text{MA}(f(t)) = \rho \cos(\omega t) \cdot f(t).$$

risulta essere un sistema lineare ma non tempo invariante. Infatti la risposta al segnale ritardato $f(t - t_0)$ è $\rho \cos(\omega t) \cdot f(t - t_0)$, diversa da $g(t - t_0) = \rho \cos(\omega(t - t_0)) \cdot f(t - t_0)$.

Esempio 1.6.7

Si consideri la funzione gradino unitario $u(x)$ definita da

$$u(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ 0 & \text{altrimenti.} \end{cases}$$

Il circuito sogliatore è descritto dal sistema $S(f(t)) = u(f(t))$. Tale sistema è tempo-invariante ma non lineare, come si verifica immediatamente.

Esempio 1.6.8

La modulazione di fase MF di un segnale $f(t)$ è descritta dal sistema

$$\text{MF}(f(t)) = \rho \cos(\omega t + f(t)).$$

Posto $g(t) = mf(f(t))$, risulta che $g(t - t_0) = \rho \cos(\omega t - t_0 + f(t - t_0))$. Osservando che $mf(f(t - t_0)) = \rho \cos(\omega t + f(t - t_0))$, segue :

$$g(t - t_0) \neq mf(f(t - t_0))$$

La modulazione di fase non è un sistema tempo invariante, oltre a non essere un sistema lineare.

1.6.2 Sistemi Lineari Tempo-Invarianti: filtri lineari

Di grande interesse risultano i sistemi che sono contemporaneamente lineari e tempo-invarianti, per i quali useremo l'acronimo *LTI*. Questi sistemi, per motivi che risulteranno chiari in seguito, sono anche detti *filtri lineari*. Per essi è disponibile una buona teoria, di grande aiuto all'analisi e alla progettazione; molti sistemi reali, inoltre, possono in prima approssimazione venir modellati come sistemi LTI.

Introduciamo ora una importante proprietà dei sistemi lineari tempo-invarianti (LTI), incominciando dai sistemi che operano sui segnali a tempo discreto.

Consideriamo quindi un sistema S lineare tempo-invariante (LTI) per segnali a tempo discreto. Dando in ingresso a S l'impulso unitario $\delta(n)$, S dà in uscita un certo segnale $h(n)$, che per definizione è la *risposta del sistema all'impulso*. Ebbene, il comportamento di S è completamente individuato dalla sua risposta $h(n)$ all'impulso, nel senso che, conoscendo $h(n) = S(\delta(n))$, siamo in grado di calcolare l'uscita $S(x(n))$ di S su qualsiasi ingresso $x(n)$.

Fatto 1.1 Se S è un sistema lineare tempo-invariante, allora

$$S(x(n)) = \sum_{j=-\infty}^{\infty} h(j)x(n-j)$$

dove $h(n)$ è la risposta $S(\delta(n))$ del sistema all'impulso unitario.

Per provare questo risultato, cominciamo con l'osservare che il generico segnale $x(n)$ può essere ottenuto combinando linearmente gli impulsi traslati:

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k)$$

Poichè il sistema S è lineare, vale:

$$S(x(n)) = S\left(\sum_{k=-\infty}^{\infty} x(k)\delta(n-k)\right) = \sum_{k=-\infty}^{\infty} x(k)S(\delta(n-k))$$

Poichè il sistema S è anche tempo-invariante e $h(n) = S(\delta(n))$, segue:

$$S(\delta(n-k)) = h(n-k)$$

Questo comporta:

$$S(x(n)) = \sum_{k=-\infty}^{\infty} x(k)h(n-k)$$

Ponendo infine $j = n - k$, e quindi $k = n - j$, si ottiene

$$S(x(n)) = \sum_{j=-\infty}^{\infty} h(j)x(n-j)$$

Analogo risultato vale per sistemi che operano su segnali analogici: se S è un sistema lineare tempo-invariante (LTI) per segnali a tempo continuo, il suo comportamento è completamente individuato dalla sua risposta $h(t)$ alla funzione impulsiva $\delta(t)$.

Infatti:

Fatto 1.2 *Se S è un sistema lineare tempo-invariante, allora*

$$S(f(t)) = \int_{-\infty}^{\infty} f(x)h(t-x)dx = \int_{-\infty}^{\infty} h(y)f(t-y)dy,$$

dove $h(t)$ è la risposta $S(\delta(t))$ del sistema all'impulso $\delta(t)$.

Poichè il sistema è lineare vale infatti:

$$S(f(t)) = S\left(\int_{-\infty}^{\infty} f(x)\delta(t-x)dx\right) = \int_{-\infty}^{\infty} f(x)S(\delta(t-x))dx.$$

Essendo inoltre tempo-invariante, se $h(t) = S(\delta(t))$ allora $h(t-x) = S(\delta(t-x))$, e quindi

$$S(f(t)) = \int_{-\infty}^{\infty} f(x)h(t-x)dx.$$

Col cambio di variabile $y = t - x$, e quindi $x = t - y$, si ottiene

$$\int_{-\infty}^{\infty} f(x)h(t-x)dx = \int_{-\infty}^{\infty} h(y)f(t-y)dy$$

La legge che associa a due segnali a tempo discreto $x(n)$ e $y(n)$ il segnale

$$z(n) = \sum_{k=-\infty}^{\infty} x(k)y(n-k)$$

è detta *prodotto di convoluzione* di x e y , e si denota $z = x * y$

Analogamente, per segnali a tempo continuo f e g il prodotto di convoluzione $f * g$ è dato da

$$(f * g)(t) = \int_{-\infty}^{\infty} f(x)h(t-x)dx = \int_{-\infty}^{\infty} h(y)f(t-y)dy$$

I precedenti risultati (Fatto 1.2 e Fatto 1.1), espressi in termini di convoluzione, asseriscono che la risposta di un sistema LTI a un dato segnale d'ingresso è ottenuta dalla convoluzione del segnale d'ingresso con la risposta del sistema all'impulso.

Esempio 1.6.9

Sia dato un sistema LTI per segnali a tempo discreto S la cui risposta all'impulso è la funzione $h(n) = \delta(n) - \delta(n-1)$. Si vuole determinare la risposta $y(n)$ del sistema al segnale $x(n) = \delta(n) + \delta(n-1)$.

Il problema è risolto ricordando che l'uscita di un sistema LTI è ottenuta come convoluzione tra l'ingresso e la risposta all'impulso:

$$y(n) = \sum_{j=-\infty}^{\infty} h(j)x(n-j)$$

Gli unici valori di j che contribuiscono alla somma sono 0 e 1, poichè per ogni altro valore è $h(j) = 0$. Ne segue che $y(n) = h(0)x(n) + h(1)x(n-1)$, da cui otteniamo:

$$y(n) = x(n) - x(n-1)$$

Sostituendo i valori di $x(n)$ per $n = \dots, 0, 1, 2, 3, \dots$, si ottiene:

$$y(0) = 1, \quad y(2) = -1, \quad y(n) = 0 \quad \text{per } n \neq 2 \text{ e } n \neq 0$$

1.6.3 Sistemi Causali e Stabili

Nel mondo fisico vige, fino a prova contraria, il cosiddetto principio di causalità: l'effetto non può precedere la causa che lo ha generato. Una condizione per la realizzabilità fisica di un sistema per il trattamento di segnali temporali è quindi che esso sia *causale*:

Definizione 1.3 *Un sistema si dice causale se ad ogni istante di tempo l'uscita dipende unicamente dai valori dell'ingresso al tempo presente e ai tempi passati.*

Per questo motivo un sistema causale viene anche detto nonanticipativo, nel senso che il valore della risposta al tempo t non richiede la conoscenza dei valori dell'ingresso a tempi $t' > t$. Come conseguenza, se due ingressi del sistema sono identici fino ad un certo istante t_0 , le corrispondenti uscite devono essere uguali fino a quell'istante.

Nei sistemi LTI il comportamento del sistema è univocamente individuato dalla sua risposta all'impulso $\delta(t)$. Poichè l'impulso $\delta(t)$ arriva all'istante 0, se il sistema è causale la risposta all'impulso dovrà essere nulla per $t < 0$. Per i sistemi LTI la nozione di causalità può essere quindi caratterizzata da una semplice proprietà della risposta all'impulso:

Definizione 1.4 *Un sistema S lineare tempo-invariante è causale se la sua risposta $h(t)$ all'impulso è nulla per valori di t negativi, cioè quando $h(t) = 0$ per $t < 0$.*

Se un sistema LTI è causale, allora si può porre:

1. *caso continuo:*

$$S(f(t)) = \int_0^{\infty} h(x)f(t-x)dx.$$

2. *caso discreto:*

$$S(x(n)) = \sum_{k=0}^{\infty} h(k)x(n-k).$$

Esempio 1.6.10

Sia dato un sistema LTI per segnali a tempo discreto S la cui risposta all'impulso è la funzione $h(n) = \delta(n+2) - \delta(n-2)$. Si vuole determinare di quanto deve essere traslata la risposta all'impulso $h(n)$ perchè il sistema diventi causale.

Per prima cosa, si osserva che S non è causale. Infatti al tempo negativo $n = -2$ vale che $h(-2) \neq 0$: l'impulso $\delta(n)$ arriva all'istante 0, mentre il sistema ha già risposto al tempo -2 . Traslando $h(n)$ di 2 si ottiene $h(n-2) = \delta(n) - \delta(n-4)$; poichè per $n < 0$ vale $h(n-2) = 0$, il sistema che ha $h(n-2)$ come risposta all'impulso è causale.

Un'altra nozione di interesse pratico è quella di *stabilità*. Ricordiamo che un segnale a tempo continuo $f(t)$ è detto limitato se esiste $M > 0$ tale che $|f(t)| < M$ per ogni t . Analogamente, un segnale a tempo discreto $x(n)$ è detto limitato se esiste $M > 0$ tale che $|x(n)| < M$ per ogni n .

Esempio 1.6.11

Il segnale a tempo discreto $\delta(n-3)$ è limitato: il suo valore massimo è 1. Anche il segnale $\cos\pi n$ risulta limitato tra -1 e 1 . Il segnale $n \cdot \cos\pi n$ non è invece limitato, ma diverge oscillando.

Sistemi che trasformano segnali limitati in segnali limitati sono detti stabili:

Definizione 1.5 Un sistema S è stabile (o BIBO, cioè Bounded Input Bounded Output) se trasforma segnali limitati in segnali limitati.

Un sistema che non è stabile è detto *instabile*. Un sistema risulta instabile se esiste in segnale limitato che, posto in ingresso al sistema, dà luogo ad una risposta non limitata.

La stabilità è una specifica richiesta in molte applicazioni, è quindi utile determinare condizioni che garantiscano la stabilità di un sistema. Per i sistemi lineari tempo-invarianti la stabilità può essere caratterizzata da una semplice proprietà della risposta all'impulso:

Fatto 1.3 Un sistema lineare tempo-invariante S è stabile se e solo se, detta $h(t)$ la risposta di S all'impulso, vale che:

$$\int_{-\infty}^{+\infty} |h(t)| dt < +\infty.$$

Dimostrazione. Supponiamo, per prima cosa, che $\int_{-\infty}^{+\infty} |h(t)| dt < +\infty$ e dimostriamo che il sistema è stabile. Supponiamo che l'ingresso $f(t)$ sia limitato, cioè esiste $M > 0$ tale che $|f(t)| < M$ per ogni t . Sia $g(t)$ la risposta del sistema su ingresso $f(t)$; vale:

$$|g(t)| = \left| \int_{-\infty}^{+\infty} h(\tau) f(t - \tau) d\tau \right| \leq \int_{-\infty}^{+\infty} |h(\tau)| |f(t - \tau)| d\tau \leq M \int_{-\infty}^{+\infty} |h(\tau)| d\tau < +\infty.$$

Questo implica che $g(t)$ è limitato e quindi il sistema è stabile.

Supponiamo ora che $\int_{-\infty}^{+\infty} |h(t)| dt = +\infty$ e dimostriamo che il sistema non è stabile.

Se poniamo infatti in ingresso al sistema il segnale $f(t) = \text{sgn}(h(-t))$ chiaramente limitato, l'uscita $g(t)$ è tale che:

$$g(0) = \int_{-\infty}^{+\infty} h(\tau) f(-\tau) d\tau = \int_{-\infty}^{+\infty} h(\tau) \text{sgn}(h(\tau)) d\tau = \int_{-\infty}^{+\infty} |h(\tau)| d\tau = +\infty.$$

Il sistema non risulta quindi stabile. ■

Per sistemi LTI che operano su segnali a tempo discreto vale un risultato analogo:

Fatto 1.4 Un sistema LTI è stabile se e solo se la risposta $h(n)$ all'impulso unitario è tale che $\sum_{n=-\infty}^{+\infty} |h(n)| < +\infty$.

Esempio 1.6.12

Il sistema la cui risposta all'impulso è $h(n) = n \cdot \cos \pi n$ è instabile. Infatti esso trasforma l'impulso $\delta(n)$, che è un segnale limitato, nel segnale $n \cdot \cos \pi n$, che non è limitato.

Anche il sistema la cui risposta all'impulso è $h(n) = u(n)$ è instabile. Ricordiamo che $u(n) = 1$ per $n \geq 0$, mentre $u(n) = 0$ per $n < 0$. In questo caso la risposta all'impulso è limitata, ma vale che $\sum_{n=-\infty}^{+\infty} |u(n)| = \sum_{n=0}^{\infty} 1 = \infty$

Il sistema la cui risposta all'impulso è $\frac{1}{2^n} \cdot u(n)$ è stabile. Infatti $\sum_{n=-\infty}^{+\infty} |\frac{1}{2^n} \cdot u(n)| = \sum_{n=0}^{\infty} \frac{1}{2^n} = 2 < \infty$.

1.6.4 Filtri FIR e IIR per segnali a tempo discreto

In questa sezione ci concentriamo su particolari sistemi LTI che trasformano segnali a tempo discreto. In un sistema LTI, sappiamo che l'uscita $y(n)$ viene determinata dalla convoluzione dell'ingresso $x(n)$ con la risposta all'impulso $h(n)$:

$$y(n) = \sum_{j=-\infty}^{\infty} h(j)x(n-j)$$

E' di particolare interesse il caso in cui la risposta all'impulso $h(n)$ è finita, cioè esistono interi $a < b$ tali che $h(n) = 0$ per $n > a$ oppure $n < b$. In questo caso, infatti, la convoluzione si riduce ad una combinazione lineare finita, cosa che permette di ottenere un semplice algoritmo di calcolo dell'uscita $y(n)$:

$$y(n) = \sum_{j=a}^b h(j)x(n-j)$$

Filtri in cui la risposta all'impulso è finita sono detti *Filtri FIR* (Finite Impulse Response); di grande interesse sono i filtri FIR causali, in cui la relazione ingresso-uscita è del tipo:

$$y(n) = \sum_{j=0}^M h(j)x(n-j)$$

Si osservi che l'uscita $y(n)$ al tempo n , dato l'ingresso $x(n)$, è la combinazione lineare di $x(n)$ e degli M precedenti ingressi, con coefficienti dati dalla risposta all'impulso. Per poter calcolare l'uscita $y(n)$ al tempo n , dato l'ingresso $x(n)$, si devono quindi eseguire $M + 1$ moltiplicazioni e M addizioni, utilizzando gli M elementi $x(n-1), x(n-2), \dots, x(n-M)$ precedentemente memorizzati: si dice che il sistema ha *memoria* M .

Una importante caratteristica dei filtri FIR è la loro stabilità. Dato un filtro FIR, infatti, la sua risposta $h(n)$ all'impulso è finita, e quindi risulta

$$\sum_{n=-\infty}^{\infty} |h(n)| < \infty$$

La precedente condizione implica che il filtro è stabile.

Esempio 1.6.13

Sia dato il filtro con risposta all'impulso $h(n) = \delta(n) - \delta(n-1)$. Si vuole determinare la memoria del filtro e la risposta $y(n)$ al segnale $x(n) = \delta(n) + \delta(n-1) + \delta(n-2) + \delta(n-3)$.

Il filtro è un filtro FIR causale, poiché $h(n) = 0$ per $n < 0$ oppure $n > 1$. In particolare, la relazione ingresso-uscita è data da:

$$y(n) = x(n) - x(n-1)$$

Questo implica che il filtro ha memoria 1.

Per $x(n) = \delta(n) + \delta(n-1) + \delta(n-2) + \delta(n-3)$ risulta:

$$y(n) = \delta(n) + \delta(n-1) + \delta(n-2) + \delta(n-3) - (\delta(n-1) + \delta(n-2) + \delta(n-3) + \delta(n-4)) = \delta(n) - \delta(n-4)$$

Esempio 1.6.14

Sia dato un sistema LTI per segnali a tempo discreto S la cui risposta all'impulso è la funzione $h(n) = \delta(n+2) - \delta(n-2)$. Si vuole determinare di quanto deve essere traslata la risposta all'impulso $h(n)$ perchè il sistema diventi causale.

Per prima cosa, si osserva che S non è causale. Infatti al tempo negativo $n = -2$ vale che $h(-2) \neq 0$: l'impulso $\delta(n)$ arriva all'istante 0, mentre il sistema ha già risposto al tempo -2 . Traslando $h(n)$ di 2 si ottiene $h(n-2) = \delta(n) - \delta(n-4)$; poichè per $n < 0$ vale $h(n-2) = 0$, il sistema che ha $h(n-2)$ come risposta all'impulso è causale.

Molte caratteristiche dei filtri FIR possono essere estese al caso di sistemi, generalmente non lineari, caratterizzati da una relazione ingresso-uscita del tipo:

$$y(n) = F(x(n), x(n-1), \dots, x(n-M))$$

dove F è una funzione in $M+1$ variabili. Si osservi che tali sistemi sono sempre tempo invarianti, mentre risultano lineari solo se la funzione F è una combinazione lineare. Come nel caso lineare, per determinare l'uscita $y(n)$ al tempo n , dato l'ingresso $x(n)$, è necessario aver memorizzato gli M elementi $x(n-1), x(n-2), \dots, x(n-M)$: il sistema ha memoria M .

Esempio 1.6.15

Si consideri il seguente filtro non lineare, detto filtro mediano, con relazione ingresso-uscita:

$$z(n) = \text{mediana}(x(n), x(n-1), x(n-2))$$

Si determini la risposta del filtro mediano all'ingresso

$$x(n) = \delta(n) + 20\delta(n-1) + \delta(n-2) + \delta(n-3)$$

in cui si osserva un anomalo coefficiente 20, e la si confronti con quella del filtro lineare FIR con relazione ingresso-uscita data da:

$$y(n) = \frac{1}{3}(x(n) + x(n-1) + x(n-2))$$

Le risposte dei due filtri all'ingresso $x(n)$ sono:

$$z(n) = \delta(n) + \delta(n-1) + \delta(n-2) + \delta(n-3)$$

$$y(n) = 0.33\delta(n) + 7\delta(n-1) + 7.33\delta(n-2) + 7.33\delta(n-3) + 0.66\delta(n-4) + 0.33\delta(n-5)$$

In conclusione, si osserva che solo il filtro mediano riesce a eliminare l'anomalo coefficiente 20.

Un filtro lineare in cui la risposta all'impulso $h(n)$ non è finita è detto *Filtro IIR* (Infinite Impulse Reponse). In seguito tratteremo il caso di filtri IIR causali. In questo caso, detta $h(n)$ la risposta all'impulso, la relazione ingresso-uscita è data da:

$$y(n) = \sum_{k=0}^{\infty} h(k) \cdot x(n-k)$$

La formula precedente non porta direttamente ad un algoritmo di calcolo, in quanto viene richiesta l'esecuzione di infinite operazioni. Tuttavia, se il segnale di ingresso $x(n)$ è nullo prima di un tempo a , cioè se $x(n) = 0$ per $n < a$, allora è possibile calcolare l'uscita $y(n)$. Essa risulta nulla per $n < a$; per $n \geq a$ vale invece:

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1) + \dots + h(n-a) \cdot x(a)$$

Si osservi che il numero di operazioni necessarie a calcolare l'uscita $y(n)$ al tempo n , dato l'ingresso $x(n)$, è proporzionale ad n , mentre nei filtri FIR risulta indipendente da n . In particolare, il calcolo di $y(a), y(a+1), \dots, y(n)$ richiede $O(n^2)$ operazioni di somma e prodotto con una memoria $O(n)$.

Introduciamo ora una sottoclasse di filtri IIR, i cosiddetti *Filtri ricorsivi*, per i quali per il calcolo di $y(a), y(a+1), \dots, y(n)$ richiede solo $O(n)$ operazioni di somma e prodotto ed una memoria $O(1)$, come nel caso dei filtri FIR.

Un filtro lineare è detto *ricorsivo* se ammette una relazione ingresso-uscita del tipo:

$$y(n) = \sum_{k=1}^D a(k) \cdot y(n-k) + \sum_{k=0}^M b(k) \cdot x(n-k)$$

Si osservi che nella precedente equazione $y(n)$ è definito ricorsivamente, da cui il nome filtro ricorsivo; i filtri ricorsivi sono anche chiamati sistemi auto-regressivi a media mobile (ARMA).

Analizzando la relazione ingresso-uscita, si osserva che, per calcolare l'uscita $y(n)$ al tempo n , dato l'ingresso $x(n)$, è necessario conoscere i valori $y(n-1), y(n-2), \dots, y(n-D)$ e $x(n-1), x(n-2), \dots, x(n-M)$, per un totale di $N+M$ elementi; $y(n)$ può essere calcolato mediante $D+M$ moltiplicazioni e $D+M-1$ addizioni. Per segnali di ingresso $x(n)$ nulli prima di un tempo a , è possibile calcolare l'uscita $y(n)$ osservando che, per la causalità del

filtro, risulta $y(a-1) = \dots = y(a-D) = 0$; in particolare, il calcolo di $y(a), y(a+1), \dots, y(n)$ richiede solo $O(n)$ operazioni di somma e prodotto e una memoria $D + M$.

Vediamo ora come sia possibile determinare la risposta all'impulso di un filtro ricorsivo a partire sua relazione ingresso-uscita. Sia fissato un filtro con relazione ingresso-uscita data da:

$$y(n) = \sum_{k=1}^D a(k) \cdot y(n-k) + \sum_{k=0}^M b(k) \cdot x(n-k)$$

Detta $h(n)$ la risposta del filtro all'impulso $\delta(n)$, essa verifica quindi:

$$h(n) = \sum_{k=1}^D a(k) \cdot h(n-k) + \sum_{k=0}^M b(k) \cdot \delta(n-k)$$

La precedente equazione permette di calcolare $h(k)$ noti i valori $h(k-1), \dots, h(k-N+1)$. Poiché il filtro è causale, risulta che $h(-1) = h(-2) = \dots = h(-N+1) = 0$ ed è quindi possibile calcolare $h(n)$ per ogni $n \geq 0$. In conclusione, se il filtro è causale, la relazione ingresso-uscita permette di individuare univocamente la risposta del filtro all'impulso.

Possiamo calcolare la forma esplicita di $h(n)$ come segue. Si osservi per prima cosa che, per ogni $n > M$, vale l'equazione:

$$h(n) = \sum_{k=1}^D a(k) \cdot h(n-k)$$

Si tratta di una equazione alle differenze lineare omogenea a coefficienti costanti che viene risolta con i seguenti due passi principali.

1. Assumendo che la soluzione sia il segnale λ^n , sostituendo $h(n)$ si ottiene l'equazione $\lambda^n - \sum_{k=1}^D a(k) \cdot \lambda^{n-k} = 0$, e questo implica che vale, se $\lambda \neq 0$:

$$\lambda^D + \sum_{k=1}^D a(k) \cdot \lambda^{D-k} = 0$$

Il polinomio $\lambda^D + \sum_{k=1}^D a(k) \cdot \lambda^{D-k}$ è detto polinomio caratteristico. Assumendo per semplicità che le radici che le radici $\lambda_1, \dots, \lambda_D$ del polinomio caratteristico siano tutte distinte, i segnali $\lambda_1^n, \dots, \lambda_D^n$ sono potenziali soluzioni della relazione ingresso-uscita del filtro, così come ogni loro combinazione lineare $\sum_{k=1}^D c(k) \cdot \lambda_k^n$.

2. Tentativamente si pone $h(n) = \sum_{k=1}^D c(k) \cdot \lambda_k^n$, in cui i coefficienti $c(1), \dots, c(D)$ sono incogniti. Allo scopo di determinare tali coefficienti, si calcolano per prima cosa i valori $h(M+1), \dots, h(M+D)$. I coefficienti $c(1), \dots, c(D)$ sono ottenuti risolvendo il sistema lineare:

$$\sum_{k=1}^D c(k) \lambda_k^{M+1} = h(M+1)$$

$$\dots$$

$$\sum_{k=1}^D c(k) \lambda_k^{M+D} = h(M+D)$$

In conclusione, la risposta all'impulso di un filtro ricorsivo causale in cui le radici del polinomio caratteristico sono distinte è, da un certo tempo in poi, una combinazione lineare di segnali del tipo λ^n ; se il polinomio caratteristico ha radici multiple, si può provare che la risposta all'impulso è una combinazione lineare di segnali del tipo $n^m \cdot \lambda^n$ con m intero non negativo.

Esempio 1.6.16

Si consideri il filtro IIR S la cui risposta all'impulso è $h(n) = \frac{1}{n} \cdot u(n-1)$. È S un filtro ricorsivo? Determinare la risposta del filtro al segnale di ingresso $u(n)$. È S stabile?

Per $n > 0$, la risposta all'impulso è $\frac{1}{n}$. Poiché tale segnale non può essere ottenuto come combinazione lineare di segnali del tipo $n^m \cdot \lambda^n$ con m intero non negativo, questo filtro lineare è IIR ma non ricorsivo.

Se il segnale di ingresso è $u(n)$, l'uscita $y(n)$ è ottenibile dalla convoluzione:

$$y(n) = \sum_{k=1}^{\infty} \frac{1}{k} \cdot u(n-k) = \sum_{k=1}^n \frac{1}{k} \approx \ln n$$

Questo filtro trasforma il segnale $u(n)$, che è limitato, in un segnale che, per grandi n , è approssimativamente $\ln n$, quindi non limitato: il filtro è instabile

Esempio 1.6.17

Determinare la risposta $h(n)$ all'impulso del sistema descritto dalla seguente relazione ingresso-uscita:

$$y(n) = -0.5 \cdot y(n-1) + 0.24 \cdot y(n-2) + x(n)$$

La risposta $h(n)$ all'impulso verifica l'equazione:

$$h(n) = -0.5 \cdot h(n-1) + 0.24 \cdot h(n-2) + \delta(n)$$

Per $n < 0$ vale $h(n) = 0$ per la causalità, ed è quindi possibile calcolare $h(0) = 1$, $h(1) = -0.5$, $h(2) = 0.49$. Per $n > 2$ vale l'equazione

$$y(n) = -0.5 \cdot y(n-1) + 0.24 \cdot y(n-2)$$

il cui polinomio caratteristico risulta $\lambda^2 + 0.5 \cdot \lambda - 0.49$. Le radici del polinomio caratteristico sono $\lambda_1 = -0.8$ e $\lambda_2 = 0.3$; ricordando che $h(1) = -0.5$ e $h(2) = 0.49$, si ottiene il sistema:

$$\begin{aligned} c(1) \cdot (-0.8)^1 + c(2) \cdot (0.3)^1 &= -0.5 \\ c(1) \cdot (-0.8)^2 + c(2) \cdot (0.3)^2 &= 0.49 \end{aligned}$$

la cui soluzione è $c(1) = \frac{8}{11}$ e $c(2) = \frac{3}{11}$. In conclusione:

- per $n < 0$ vale $h(n) = 0$
- per $n = 0$ vale $h(n) = 1$
- per $n > 0$ vale $h(n) = \frac{8}{11} \cdot (-0.8)^n + \frac{3}{11} \cdot 0.3^n$

Mentre i filtri FIR sono sempre stabili, i filtri IIR possono essere instabili. Poiché la stabilità di un filtro è una specifica in molte applicazioni, è importante l'individuazione di metodi efficienti per determinare se un filtro è stabile. Per filtri IIR ricorsivi esiste un semplice criterio di stabilità:

Fatto: Un filtro causale ricorsivo S è stabile se e solo se, per ogni radice λ_k del polinomio caratteristico, vale $|\lambda_k| < 1$.

Questo risultato può essere intuito sulla base delle seguenti osservazioni. La risposta $h(n)$ del filtro all'impulso è una combinazione lineare di segnali del tipo $n^{m_k} \cdot \lambda_k^n$, e condizione necessaria e sufficiente per la stabilità per il filtro è che $\sum_{n=0}^{\infty} |h(n)| < \infty$. Ricordando che $\sum_{n=0}^{\infty} |n^{m_k} \cdot \lambda_k^n| < \infty$ se e solo se $|\lambda_k| < 1$, si ottiene la condizione di stabilità enunciata.

Esempio 1.6.18

Dato il filtro S con relazione ingresso-uscita $y(n) = -0.5 \cdot y(n-1) + 0.24 \cdot y(n-2) + x(n)$, studiarne la stabilità.

Il polinomio caratteristico di S è $\lambda^2 + 0.5 \cdot \lambda + 0.24$, le cui radici sono $\lambda_1 = -0.8$ e $\lambda_2 = 0.3$. Poiché $|-0.8| < 1$ e $|0.3| < 1$, il filtro è stabile.

Mentre i filtri FIR sono sempre stabili, i filtri IIR possono essere instabili. Poiché la stabilità di un filtro è una specifica in molte applicazioni, è importante l'individuazione di metodi efficienti per determinare se un filtro è stabile. Per filtri IIR ricorsivi esiste un semplice criterio di stabilità:

Fatto: Sia assegnato un filtro ricorsivo S che verifica la relazione ingresso-uscita

$$y(n) = \sum_{k=1}^D a(k) \cdot y(n-k) + \sum_{k=0}^M b(k) \cdot x(n-k)$$

Detto *polinomio caratteristico* del filtro il polinomio $P(\lambda) = \lambda^D - \sum_{k=1}^D a(k) \cdot \lambda^{D-k}$, allora S è stabile se e solo se ogni soluzione λ_k dell'equazione $P(\lambda) = 0$ verifica $|\lambda_k| < 1$.

Esempio 1.6.19

Dato il filtro S con relazione ingresso-uscita $y(n) = -0.5 \cdot y(n-1) + 0.24 \cdot y(n-2) + x(n)$, studiarne la stabilità.

Il polinomio caratteristico di S è $\lambda^2 + 0.5 \cdot \lambda + 0.24$, le cui radici sono $\lambda_1 = -0.8$ e $\lambda_2 = 0.3$. Poichè $|-0.8| < 1$ e $|0.3| < 1$, il filtro è stabile.

CAPITOLO 2

Analisi in Frequenza di Segnali Analogici

Osservare lo stesso oggetto da più punti di vista può essere vantaggioso, perchè ci permette di adottare, nelle varie circostanze, il punto di vista che risulta più significativo. Questo atteggiamento, trasportato nell'ambito dei segnali, ci porta a prendere in considerazioni più rappresentazioni alternative dello stesso segnale.

Rappresentazioni dei segnali diversa da quella temporale possono essere ottenute come segue: invece di esprimere il segnale $f(t)$ come sovrapposizione di impulsi $\delta(t - x)$, lo si rappresenta come sovrapposizione di un'altra famiglia $H(t, x)$ di segnali, a patto che tali segnali formino una base. Nel primo caso, abbiamo osservato che:

$$f(t) = \int_{-\infty}^{+\infty} f(x)\delta(t - x)dx$$

Il segnale è dunque individuato dai coefficienti $f(x)$, che esprimono la sua rappresentazione

temporale. Nel secondo caso, risulterà:

$$f(t) = \int_{-\infty}^{+\infty} a(x)H(t, x)dx$$

Come sappiamo, infatti, i coefficienti $a(x)$ individuano biunivocamente il segnale $f(t)$ e ne esprimono una nuova rappresentazione.

Una base particolarmente interessante è quella introdotta e studiata da Jean Baptiste Joseph Baron de Fourier, governatore del Basso Egitto sotto Napoleone e autore di un monumentale trattato sulla propagazione del calore pubblicato nel 1822. In questo lavoro veniva introdotto il concetto di *serie di Fourier*; la *trasformata di Fourier* (o integrale di Fourier) è una naturale estensione della serie che si applica a segnali arbitrari, mentre la serie si riferisce solo a segnali periodici. Obiettivo di questo capitolo è la presentazione di risultati sulla rappresentazione di Fourier di un segnale, detta anche *rappresentazione in frequenza* o *spettro* del segnale.

Il termine *spettro* compare in un articolo sottoposto nel 1792 alla Royal Society da Isaac Newton, e denota la banda continua di colori prodotta da un prisma che decompone la luce solare. Più tardi, nella metà del 1800, Kirchhoff e Bunsen scoprirono che ogni elemento chimico, portato all'incandescenza, irradia una sua propria luce colorata che ne identifica lo spettro. Quindi il lavoro iniziato da Newton di analisi della luce nelle sue componenti di colore è a tutti gli effetti una *analisi in frequenza*, i cui principi sono richiamati in questo capitolo.

Il capitolo è suddiviso come segue. Viene richiamata brevemente la nozione di numero complesso e le principali operazioni su tali numeri; viene in particolare evidenziata la relazione di Eulero

$$e^{i\theta} = \cos \theta + i \sin \theta$$

Tale relazione permette di esprimere i segnali sinusoidali $\sin 2\pi\nu t$ e $\cos 2\pi\nu t$ in termini di esponenziali immaginari $e^{2\pi i\nu t}$ e $e^{-2\pi i\nu t}$. L'apparente complicazione legata al fatto che prendiamo in considerazione segnali a valori complessi ha il vantaggio di semplificare le manipolazioni algebriche richieste per trattare la trigonometria. I segnali effettivi, che assumono valori reali, possono essere trattati come una sottoclasse.

Il resto del capitolo è dedicato alla discussione dell'importante concetto di *spettro* di un segnale. Tale nozione viene dapprima introdotta nel semplice caso di segnali a *spettro finito*. Si discute poi la espansione in serie di Fourier per segnali periodici, cosa che permette di ottenerne lo *spettro discreto*. Questo approccio viene infine esteso a segnali non periodici, introducendo la trasformata di Fourier che permette di trattare segnali a *spettro continuo*. Vengono poi studiate alcune proprietà della trasformata di Fourier, che giustificano l'uso di semplici manipolazioni simboliche, e vengono discusse alcune caratteristiche degli spettri dei segnali a valori reali. Vengono infine introdotte le nozioni di *energia* e *potenza* di segnali, e i loro corrispettivi spettrali, la *densità di energia* e la *densità di potenza* spettrali.

2.1 Numeri Complessi

Richiamiamo qui brevemente la nozione di numero complesso. Un numero complesso z è identificato da una coppia di numeri reali (a, b) ; si scrive solitamente $z = a + ib$, chiamando $i = \sqrt{-1}$ *unità immaginaria*, $a = \operatorname{Re}\{z\}$ *parte reale* e $b = \operatorname{Im}\{z\}$ *parte immaginaria* di z .

Le operazioni di addizione e moltiplicazione di due numeri complessi z e w sono definite nel modo seguente:

$$\text{somma: } (a + ib) + (c + id) = (a + c) + i(b + d),$$

$$\text{prodotto: } (a + ib)(c + id) = (ac - bd) + i(ad + bc).$$

Rispetto a tali operazioni i numeri complessi formano un campo; in particolare si verifica direttamente che l'inverso rispetto alla moltiplicazione è dato da:

$$(a + ib)^{-1} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2}.$$

Il sottocampo di numeri complessi della forma $a + i \cdot 0$ (cioè tutti quelli che hanno parte immaginaria 0) è isomorfo al campo reale $\mathbf{R}$.

Poichè un numero complesso $z = a + ib$ identifica una coppia ordinata di numeri reali, esso può essere rappresentato geometricamente come un punto del piano cartesiano: la parte reale è l'ascissa, la parte immaginaria è l'ordinata. Descrivendo tale punto in coordinate polari (r, θ) , come mostrato in Figura 2.1, si ricava la forma *trigonometrica* di un numero complesso:

$$z = r(\cos \theta + i \sin \theta).$$

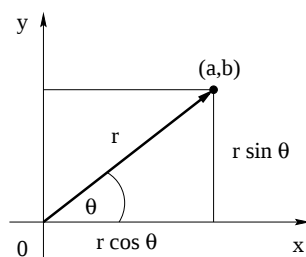


Figura 2.1 Rappresentazione geometrica del numero complesso $a + ib$.

La seguente coppia di trasformazioni permette di determinare le coordinate polari (r, θ) note le coordinate cartesiane (a, b) e viceversa:

$$r = \sqrt{a^2 + b^2} = \sqrt{(\operatorname{Re}\{z\})^2 + (\operatorname{Im}\{z\})^2}, \quad \theta = \arctan\left(\frac{b}{a}\right) = \arctan\left(\frac{\operatorname{Im}\{z\}}{\operatorname{Re}\{z\}}\right),$$

r è chiamato *modulo* di z mentre l'angolo θ è la sua *fase*; scriveremo $r = |z|$ e $\theta = \text{ph } z$.

Un'altra utile rappresentazione di un numero complesso è la cosiddetta forma *esponenziale* o *polare*:

$$z = re^{i\theta},$$

che si ottiene applicando alla forma trigonometrica la relazione di Eulero:

$$e^{i\theta} = \cos \theta + i \sin \theta.$$

Quest'ultima rappresentazione risulta particolarmente semplice quando si debbano effettuare operazioni di moltiplicazione e divisione, poichè il prodotto di due numeri complessi è il numero complesso che ha come modulo il prodotto dei due moduli e come fase la somma delle due fasi:

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|; \quad \text{ph}(z_1 \cdot z_2) = \text{ph}(z_1) + \text{ph}(z_2).$$

Infatti, se $z_1 = r_1 e^{i\theta}$ e $z_2 = r_2 e^{i\phi}$, si ha:

$$\begin{aligned} z_1 z_2 &= r_1 (\cos \theta + i \sin \theta) \cdot r_2 (\cos \phi + i \sin \phi) \\ &= r_1 r_2 [\cos \theta \cos \phi - \sin \theta \sin \phi + i(\sin \theta \cos \phi + \cos \theta \sin \phi)] \\ &= r_1 r_2 [\cos(\theta + \phi) + i \sin(\theta + \phi)] = r_1 r_2 e^{i(\theta + \phi)} \end{aligned}$$

La potenza ennesima di $z = re^{i\theta}$ si ottiene facilmente applicando ripetutamente la formula precedente:

$$z^n = r^n e^{in\theta} = r^n (\cos n\theta + i \sin n\theta).$$

Una radice n -esima di un numero complesso z è un numero x tale che $x^n = z$. Se $z = \rho e^{i\theta}$ e $x = r e^{i\psi}$, vale allora:

$$\rho e^{i\theta} = r^n e^{in\psi}$$

Osservando che due numeri complessi sono uguali se hanno lo stesso modulo e la differenza fra le fasi è multipla intera di un angolo giro, otteniamo:

$$r^n = \rho \quad n\psi - \theta = 2k\pi \quad (k \in \mathbf{Z}).$$

Si conclude allora che un numero complesso $\rho e^{i\theta}$ ha n radici complesse tutte con lo stesso modulo $r = \rho^{1/n}$ e fasi $\psi = \frac{\theta}{n} + \frac{2k\pi}{n}$ ($k = 0, \dots, n-1$).

Se $z = a + ib$, il *complesso coniugato* di z è il numero complesso $\bar{z} = a - ib$. Geometricamente $\bar{z}$ rappresenta il simmetrico di z rispetto all'asse reale. La definizione di coniugato implica che:

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2, \quad \overline{z_1 / z_2} = \bar{z}_1 / \bar{z}_2.$$

Risulta inoltre che la somma e il prodotto di numeri complessi coniugati sono sempre numeri reali; in particolare:

$$z + \bar{z} = 2 \operatorname{Re} \{z\}, \quad z \bar{z} = |z|^2.$$

La notazione esponenziale $z = re^{i\theta}$ evidenzia con immediatezza il modulo r e la fase θ del numero complesso z . Di particolare interesse sono i numeri di modulo 1, esprimibili come $e^{i\theta}$, che si trovano sulla circonferenza di raggio 1. $e^{i\theta}$ può essere vista come una funzione della variabile reale θ ; la relazione di Eulero permette di ottenere:

$$e^{i\theta} = \cos \theta + i \sin \theta, \quad e^{-i\theta} = \cos \theta - i \sin \theta$$

Poichè $\cos \theta$ e $\sin \theta$ sono funzioni a valori reali periodiche di periodo 2π , lo stesso periodo sarà posseduto dalle funzioni a valori complessi $e^{i\theta}$ e $e^{-i\theta}$. Risolvendo le precedenti equazioni rispetto a $\cos \theta$ e $\sin \theta$, si ottiene:

$$\cos \theta = \frac{1}{2} (e^{i\theta} + e^{-i\theta}), \quad \sin \theta = \frac{1}{2i} (e^{i\theta} - e^{-i\theta}).$$

Queste relazioni sono particolarmente utili, in quanto riducono le complesse regole di manipolazione della trigonometria alle più semplici regole di manipolazione algebrica degli esponenziali.

2.2 Segnali Periodici

In natura esistono fenomeni che avvengono una volta e poi non si verificano più, ma esistono anche molti fenomeni che si ripetono con regolarità a intervalli uguali di tempo. Esempi sono l'alternanza del giorno e della notte, le vibrazioni delle corde di un violino, il battito cardiaco. Segnali collegati a tali fenomeni sono quelli che si ripetono periodicamente nel tempo e vengono chiamati *segnali periodici*.

Più precisamente, diremo che un segnale $f(t)$ è *periodico* di periodo T se, per ogni t , vale che $f(t) = f(t + T)$.

Se un segnale $f(t)$ è periodico di periodo T allora, per ogni k intero, $f(t) = f(t + kT)$; la funzione $f(t)$ periodica di periodo T viene pertanto univocamente individuata dalla sua restrizione all'intervallo $-\frac{T}{2} \leq t \leq \frac{T}{2}$, come mostrato in Figura 2.2.

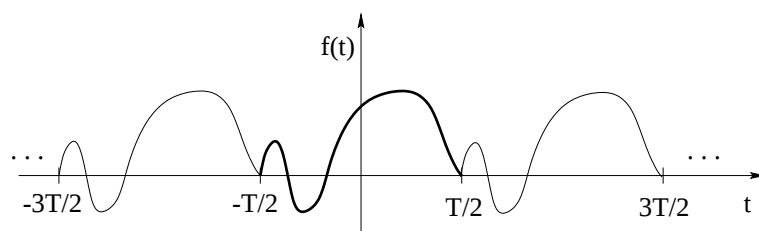


Figura 2.2 Funzione periodica di periodo T .

Esempio 2.2.1

Poichè $\sin(t + 2\pi) = \sin t$ e $\cos(t + 2\pi) = \cos t$ per ogni t , le funzioni sinusoidali $\sin t$ e $\cos t$ sono funzioni periodiche di periodo 2π . Le funzioni $\cos(2\pi i \frac{t}{T})$ e $\sin(2\pi i \frac{t}{T})$ sono periodiche di periodo T .

E' facile verificare che la combinazione lineare di funzioni periodiche di periodo T è una funzione periodica dello stesso periodo; di conseguenza, la funzione esponenziale complessa $e^{i\theta} = \cos \theta + i \sin \theta$ è periodica di periodo 2π , e quindi la funzione $e^{2\pi i \frac{t}{T}}$ risulta periodica di periodo T .

Dato un segnale periodico $f(t)$, di chiama *frequenza* del segnale il numero di ripetizioni del periodo T nell'unità di tempo. Detta ν la frequenza del segnale $f(t)$ periodico di periodo T , risulta quindi:

$$\nu = \frac{1}{T}$$

Se l'unità di misura del tempo è il secondo (sec), la frequenza è misurata in sec^{-1} o Hertz (Hz). Multipli sono il kHz (1.000 Hz), MHz (1.000.000 Hz), GHz (1.000.000.000 Hz). Nota la frequenza ν di un segnale periodico, il periodo T può essere determinato da:

$$T = \frac{1}{\nu}$$

L'importante segnale $\sin t$ ha un periodo di 2π sec, e quindi una frequenza di $\frac{1}{2\pi}$ Hz; introducendo una nuova unità di misura della frequenza, il rad/sec, che equivale a $\frac{1}{2\pi}$ Hz, il segnale $\sin t$ risulta di 1 rad/sec.

Esempio 2.2.2

Il segnale $\cos \omega t$ ha una frequenza di ω rad/sec oppure di $\frac{\omega}{2\pi}$ Hz. I segnali $\cos 2\pi \nu t$, $\sin 2\pi \nu t$, $e^{2\pi i \nu t}$ hanno frequenza ν Hz oppure di $2\pi \nu$ rad/sec.

Il segnale $C(t) = \sum_{n=-\infty, +\infty} \delta(t - nT)$ ha periodo T sec, come si può verificare; questo significa che ha frequenza di $\frac{1}{T}$ Hz oppure di $\frac{2\pi}{T}$ rad/sec.

Il segnale $\cos(6 \cdot 10^3 \pi \nu t)$ ha una frequenza di 3 kHz, o equivalentemente ha una frequenza di circa 19.000 rad/sec. Un clock del tipo $\sum_{n=-\infty, +\infty} \delta(t - n10^{-9})$ ha un periodo di 1 nanosec, o equivalentemente una frequenza di 1 GHz.

Esempio 2.2.3

Nei sistemi di comunicazione è usuale dividere le frequenze in bande. Una convenzione spesso usata è quella di identificare le bande con numeri: la banda N è data da tutte le frequenze:

$$0.3 \times 10^N \text{ Hz} < \text{banda } N < 3 \times 10^N.$$

Ad esempio, la banda 6 va da 300 kHz a 3 MHz. Tipicamente, nelle telecomunicazioni:

- le frequenze audio occupano le bande 2,3,4 (30 Hz ÷ 30 kHz);
- le frequenze video occupano le bande 1,2,3,4,5,6,7 (fino a 30 MHz);
- le microonde occupano le bande 8,9,10,11 (30 MHz ÷ 300 GHz).

La banda 5 è anche detta *LF* (bassa frequenza), la banda 6 è detta *MF* (media frequenza), la banda 7 è detta *HF* (alta frequenza), la banda 8 è detta *VHF* e la banda 9 è detta *UHF*. Le usuali trasmissioni radio in modulazione di ampiezza avvengono nella banda *MF*(6), quelle in modulazione di frequenza nella banda *VHF* (8).

Esempio 2.2.4

Un pianoforte è composto da 88 tasti, suddivisi in ottave composte da 12 tasti consecutivi. Fissato il tasto *La4* (cioè il *La* della 4° ottava) come origine, ogni tasto è individuato dall'intero c , con segno, che esprime lo scostamento da *La4*. Per esempio, il tasto *Do4* ha scostamento $c = -9$.

Ogni tasto genera un suono di data frequenza, crescente da sinistra a destra. Qual'è la legge che associa ad ogni tasto la sua frequenza?

La soluzione moderna, detta temperamento equabile, stabilisce che il tasto di scostamento c abbia una frequenza ν_c data da:

$$\nu_c = 440 \cdot 2^{\frac{c}{12}} \text{ Hz}$$

In questo modo, ad ogni ottava la frequenza raddoppia ed il rapporto tra la frequenza di un tasto e quella del suo precedente è la costante $2^{\frac{1}{12}} \approx 1.0595$.

Così la frequenza del tasto *La4* è di 440 Hz, del tasto *La5* di 880 Hz, del tasto *Do4* di $440 \cdot 2^{\frac{-9}{12}} \approx 262$ Hz.

2.3 Segnali a spettro finito.

Un segnale temporale $f(t)$ assume valori reali. Estendiamo ora tale nozione andando a considerare segnali $f(t)$ a tempo continuo e a valori complessi; questo significa che t assume valori reali mentre $f(t)$ può assumere valori complessi. Considerare segnali a valori complessi permette una trattazione più semplice ed elegante, potendosi sostituire le funzioni trigonometriche $\sin\theta$ e $\cos\theta$ con la più semplice esponenziale immaginaria $e^{i\theta}$; ricordiamo tuttavia che, nel nostro contesto, i segnali fisicamente significativi sono quelli che assumono valori reali. Essi potranno essere trattati come casi particolari di segnali a valori complessi.

Un importante segnale a valori complessi è il segnale sinusoidale

$$e^{2\pi i \nu t}$$

A causa della relazione di Eulero, vale $e^{2\pi i \nu t} = \cos 2\pi \nu t + i \sin 2\pi \nu t$. Il segnale $e^{2\pi i \nu t}$ ha quindi frequenza ν , con $-\infty < \nu < +\infty$.

Esempio 2.3.1

Il segnale $e^{400\pi it}$ ha una frequenza di 2 KHz, mentre il segnale $e^{-700\pi it}$ ha una frequenza di -350 Hz.

Fissiamo ora un insieme di frequenze $\nu_1 < \nu_2 < \dots < \nu_m$ e l'insieme dei corrispondenti segnali sinusoidali $e^{2\pi i \nu_1 t}, e^{2\pi i \nu_2 t}, \dots, e^{2\pi i \nu_m t}$. Consideriamo un generico segnale $f(t)$ ottenuto combinando linearmente i precedenti segnali sinusoidali con coefficienti $F(\nu_1), \dots, F(\nu_m)$:

$$f(t) = \sum_{n=1}^m F(\nu_n) e^{2\pi i \nu_n t}$$

Ricordando che i segnali $e^{2\pi i \nu_1 t}, e^{2\pi i \nu_2 t}, \dots, e^{2\pi i \nu_m t}$ sono linearmente indipendenti, i pesi $F(\nu_1), F(\nu_2), \dots, F(\nu_m)$ sono univocamente individuati da $f(t)$. Diremo che il segnale $f(t)$ contiene le frequenze $\nu_1, \nu_2, \dots, \nu_m$, mentre i coefficienti $F(\nu_1), F(\nu_2), \dots, F(\nu_m)$ saranno detti *spettro* del segnale. Si osservi che lo spettro è una diversa rappresentazione del segnale e che, in questo caso, lo spettro è finito. Ricordiamo infine che, essendo lo spettro una funzione a valori complessi, saranno necessari due grafici per rappresentarla: il grafico del modulo dello spettro e quello della sua fase.

Esempio 2.3.2

Dato il segnale $f(t) = 4\cos 600\pi t \cos 200\pi t + 2\sin 400\pi t$, vogliamo determinare le frequenze presenti nel segnale ed il suo spettro.

Come conseguenza della relazione di Eulero, vale $\cos 2\pi \nu t = \frac{1}{2} (e^{2\pi i \nu t} + e^{-2\pi i \nu t})$ e $\sin 2\pi \nu t = \frac{1}{2i} (e^{2\pi i \nu t} - e^{-2\pi i \nu t})$. Risulta quindi:

$$f(t) = 4 \cdot \frac{1}{2} \cdot (e^{600\pi it} + e^{-600\pi it}) \cdot \frac{1}{2} \cdot (e^{200\pi it} + e^{-200\pi it}) + 2 \cdot \frac{1}{2i} \cdot (e^{400\pi it} - e^{-400\pi it})$$

Sviluppando le moltiplicazioni, risulta infine:

$$f(t) = e^{800\pi it} + (1-i)e^{400\pi it} + (1+i)e^{-400\pi it} + e^{-800\pi it}$$

Le frequenze contenute in $f(t)$ sono dunque -400 Hz, -200 Hz, 200 Hz e 400 Hz. Lo spettro è $F(-400) = 1$, $F(-200) = 1-i$, $F(200) = 1+i$, $F(400) = 1$. Il modulo dello spettro è $|F(-400)| = 1$, $|F(-200)| = \sqrt{2}$, $|F(200)| = \sqrt{2}$, $|F(400)| = 1$, mentre la fase è $\text{ph}(F(-400)) = 0$, $\text{ph}(F(-200)) = -\frac{\pi}{4}$, $\text{ph}(F(200)) = \frac{\pi}{4}$, $\text{ph}(F(400)) = 0$. Come si può notare, le frequenze a 200 Hz hanno più potenza di quelle a 400 Hz; si osservi inoltre che, cambiando segno alla frequenza, il modulo dello spettro resta inalterato, mentre la fase cambia segno.

Il metodo usato nell'esempio precedente può essere esteso a determinare le frequenze contenute e lo spettro di segnali del tipo:

$$P(\cos 2\pi \nu_1 t, \dots, \cos 2\pi \nu_m t, \sin 2\pi \nu_1 t, \dots, \sin 2\pi \nu_m t)$$

per un dato polinomio $P(x_1, \dots, x_m, y_1, \dots, y_m)$. Basta porre:

$$\begin{aligned}\cos 2\pi\nu_k t &= \frac{1}{2}(e^{2\pi i\nu_k t} + e^{-2\pi i\nu_k t}) \\ \sin 2\pi\nu_k t &= \frac{1}{2i}(e^{2\pi i\nu_k t} - e^{-2\pi i\nu_k t})\end{aligned}$$

sviluppando poi i prodotti, fino ad ottenere una combinazione lineare di funzioni sinusoidali del tipo $e^{2\pi i\nu t}$. Le frequenze di tali funzioni sinusoidali determinano le frequenze contenute nel segnale, mentre i coefficienti della combinazione lineare individuano lo spettro del segnale. Si osservi che, se il segnale $f(t)$ è a valori reali, contenendo una frequenza ν contiene anche la frequenza $-\nu$.

Risulta inoltre che il modulo dello spettro è una funzione pari (cioè $|F(-\nu)| = |F(\nu)|$), mentre la fase dello spettro è una funzione dispari (cioè $ph(F(-\nu)) = -ph(F(\nu))$).

Esempio 2.3.3

Consideriamo il segnale $f(t) = \cos(2\pi\Delta t) \cdot \cos(2\pi\Omega t)$, dove le frequenze Δ e Ω sono tali che $0 \approx \Delta$ e $\Delta \ll \Omega$. Come si può notare, in Figura 2.3 è rappresentato il tipico grafico di $f(t)$.

Si distingue bene la sinusoide di frequenza Ω , la cui intensità varia periodicamente nel tempo, e l'involuppo di frequenza Δ . Ciò dà luogo al fenomeno dei battimenti: il segnale a frequenza Ω diminuisce di intensità, cresce, diminuisce, cresce, ... con un periodo pari a $\frac{1}{\Delta}$.

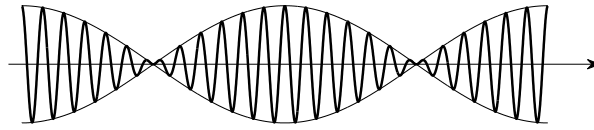


Figura 2.3 Esempio di battimenti.

Calcoliamo lo spettro di $f(t)$ applicando la formula di Eulero:

$$f(t) = \frac{1}{4}[e^{-2\pi(\Omega+\Delta)t} + e^{-2\pi(\Omega-\Delta)t} + e^{2\pi(\Omega-\Delta)t} + e^{2\pi(\Omega+\Delta)t}]$$

$f(t)$ contiene dunque le due frequenze positive $\Omega + \Delta$ e $\Omega - \Delta$, che sono quasi uguali perchè $\Delta \approx 0$: segnali che contengono due frequenze molto vicine generano dei battimenti, che tendono a sparire quando le due frequenze coincidono. Infatti il periodo con cui si alterna l'alta e la bassa intensità è proporzionale a $\frac{1}{\Delta}$, che tende a ∞ quando Δ tende a 0.

I battimenti si possono sentire chiaramente durante l'accordatura dei vari strumenti di un'orchestra a partire dal La del violino solista. Il battimento è il modo con cui avviene l'accordatura: quando non lo si sente più allora le frequenze del La dei vari strumenti sono identiche.

2.4 Segnali a spettro discreto: serie di Fourier.

Abbiamo visto che alcuni segnali ammettono uno spettro finito, potendosi rappresentare come combinazione lineare di un numero finito di funzioni sinusoidali del tipo $e^{2\pi i \nu t}$. In questo paragrafo mostriamo che una vasta classe di segnali, i segnali periodici di dato periodo T , possono essere rappresentati come combinazioni lineari di segnali sinusoidali del tipo $e^{2\pi i \frac{n}{T}t}$, con n intero, $-\infty < n < +\infty$. Lo spettro corrispondente risulta essere generalmente infinito, ma discreto.

Consideriamo tutti i segnali a valori complessi periodici di periodo T , cioè tutti quei segnali $f(t)$ tale che $f(t) = f(t + T)$ per ogni t . Questi segnali formano uno spazio vettoriale, poichè la combinazione lineare di segnali periodici di periodo T risulta essere ancora un segnale periodico di periodo T .

In questo spazio è possibile introdurre il seguente prodotto interno:

$$f(t) \circ g(t) = \frac{1}{T} \int_0^T f(t) \bar{g}(t) dt$$

Le due proprietà del prodotto interno sono infatti verificate:

Proprietà 1 Se $f(t) \neq 0$, allora

$$f(t) \circ f(t) = \frac{1}{T} \int_0^T f(t) \bar{f}(t) dt = \frac{1}{T} \int_0^T |f(t)|^2 dt > 0$$

Proprietà 2

$$\begin{aligned} (c_1 f_1(t) + c_2 f_2(t)) \circ g(t) &= \frac{1}{T} \int_0^T (c_1 f_1(t) + c_2 f_2(t)) \bar{g}(t) dt = \\ &= c_1 \frac{1}{T} \int_0^T f_1(t) \bar{g}(t) dt + c_2 \frac{1}{T} \int_0^T f_2(t) \bar{g}(t) dt = \\ &= c_1 (f_1(t) \circ g(t)) + c_2 (f_2(t) \circ g(t)) \end{aligned}$$

Il nucleo dei risultati di Fourier è basato sul fatto che i segnali sinusoidali di periodo T

$$\{e^{2\pi i \frac{n}{T}t} | -\infty < n < +\infty\}$$

sono una base ortonormale nello spazio dei segnali periodici di periodo T , detta *base di Fourier*. Questo significa che:

1. $e^{2\pi i \frac{n}{T}t} \circ e^{2\pi i \frac{n}{T}t} = 1$ per ogni n
2. $e^{2\pi i \frac{n}{T}t} \circ e^{2\pi i \frac{k}{T}t} = 0$ se $n \neq k$

Per provare queste proprietà, distinguiamo due casi:

Caso 1: Se $n = k$, allora

$$e^{\frac{2\pi i}{T}nt} \circ e^{\frac{2\pi i}{T}nt} = \frac{1}{T} \int_0^T e^{\frac{2\pi i}{T}nt} \cdot e^{-\frac{2\pi i}{T}nt} dt = \frac{1}{T} \int_0^T 1 dt = 1$$

Caso 2: Se $n \neq k$, allora

$$\begin{aligned} e^{2\pi i \frac{n}{T}t} \circ e^{2\pi i \frac{k}{T}t} &= \frac{1}{T} \int_0^T e^{2\pi i \frac{n-k}{T}t} dt = \\ &= \frac{1}{T} \left[\frac{T}{2\pi i(n-k)} e^{2\pi i \frac{n-k}{T}t} \right]_0^T = \\ &= \frac{1}{2\pi i(n-k)} (e^{2\pi i(n-k)} - 1) = 0 \end{aligned}$$

Essendo $\{e^{2\pi i \frac{n}{T}t} | -\infty < n < +\infty\}$ una base ortonormale, sappiamo che ogni segnale periodico $f(t)$ è ottenibile come combinazione lineare $f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i \frac{n}{T}t}$ degli elementi della base, e i coefficienti c_n sono ottenuti da $c_n = f(t) \circ e^{2\pi i \frac{n}{T}t}$, poichè la base è ortonormale. In conclusione, otteniamo per i segnali periodici la seguente rappresentazione di Fourier:

1. Ogni segnale $f(t)$ di periodo T può essere espresso come sovrapposizione dei segnali $e^{2\pi i \frac{n}{T}t} (-\infty < n < +\infty)$ con coefficienti c_n , cioè:

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i \frac{n}{T}t}$$

2. Il generico coefficiente c_n è ottenuto da:

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) e^{-2\pi i \frac{n}{T}t} dt$$

Fissata quindi la base di Fourier, dato il segnale $f(t)$ si possono ottenere i coefficienti c_n , e viceversa dati i coefficienti c_n si può ricostruire il segnale $f(t)$: i coefficienti c_n individuano quindi perfettamente il segnale e ne danno una rappresentazione alternativa, detta *rappresentazione in frequenza* o *spettro* del segnale. I coefficienti c_n sono generalmente infiniti ma, essendo individuati da numeri interi, formano un insieme discreto.

Esempio 2.4.1

Determinare lo spettro del segnale $C(t) = \sum_{k=-\infty}^{+\infty} \delta(t - kT)$.

Si verifica facilmente che tale segnale è periodico di periodo T . Il suo spettro è:

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} C(t) e^{-2\pi i \frac{n}{T}t} dt = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \sum_{k=-\infty}^{+\infty} \delta(t - kT) e^{-2\pi i \frac{n}{T}t} dt = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \delta(t) e^{-2\pi i \frac{n}{T}t} dt = \frac{1}{T}$$

Il segnale contiene quindi tutte le frequenze multiple intere di $\frac{1}{T}$ con la stessa potenza.

Esempio 2.4.2

Determinare lo spettro del segnale $R(t) = \sum_{k=-\infty}^{+\infty} \text{rect}(\frac{t-kT}{B})$, quando $B < T$.

Si verifica facilmente che tale segnale è periodico di periodo T . Il suo spettro è:

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \text{rect}(\frac{t}{B}) e^{-2\pi i \frac{n}{T} t} dt = \frac{1}{T} \int_{-\frac{B}{2}}^{\frac{B}{2}} e^{-2\pi i \frac{n}{T} t} dt$$

Calcolando l'integrale definito si ha:

$$\frac{1}{T} \int_{-\frac{B}{2}}^{\frac{B}{2}} e^{-2\pi i \frac{n}{T} t} dt = \frac{1}{2\pi i n} (e^{\pi i n \frac{B}{T}} - e^{-\pi i n \frac{B}{T}})$$

Ricordando che $\sin\theta = \frac{1}{2i}(e^{i\theta} - e^{-i\theta})$ si conclude:

$$c_n = \frac{1}{\pi n} \text{sen}(\pi n \frac{B}{T}) = \frac{B}{T} \text{Sinc}(\pi n \frac{B}{T})$$

Come si deduce dal grafico della funzione seno cardinale, la potenza delle componenti ad alta frequenza va a 0.

Prendiamo ora in considerazione le componenti $e^{2\pi i \frac{n}{T} t}$ della base, talvolta chiamate *componenti armoniche*.

Per $n = 0$, il segnale $e^{2\pi i \frac{n}{T} t}$ è il segnale ovunque 1 ed è detto *componente continua*. Se invece $n \neq 0$, il segnale $e^{2\pi i \frac{n}{T} t}$ è un segnale sinusoidale periodico con frequenza $\frac{n}{T}$. Tale segnale è detto *ennesima armonica*; la prima armonica ($n = 1$) ha frequenza $\frac{1}{T}$, e quindi ha lo stesso periodo T del segnale $f(t)$. Per questo motivo la prima armonica è detta *armonica fondamentale*, e le frequenze di tutte le altre armoniche sono multipli interi di quella della prima armonica: $\frac{n}{T} = n \frac{1}{T}$.

Il coefficiente c_n rappresenta il peso dell'armonica $e^{\frac{2\pi i}{T} n t}$ nel segnale $f(t)$. In particolare il coefficiente c_0 è pari al valor medio del segnale $f(t)$, poichè:

$$c_0 = \frac{1}{T} \int_0^T f(t) dt$$

I coefficienti c_n sono usualmente numeri complessi, e quindi sono individuati dal loro modulo $|c_n|$ e dalla loro fase $ph(c_n)$, dove:

1. $|c_n|$ = Modulo del peso dell'armonica di frequenza $\frac{n}{T}$
2. $ph(c_n)$ = Fase del peso dell'armonica di frequenza $\frac{n}{T}$.

Osserviamo che, dato un segnale $f(t)$ a valori reali, il suo spettro c_n è generalmente a valori complessi. In questo caso il modulo $|c_n|$ è tuttavia una funzione pari, cioè $|c_{-n}| = |c_n|$, mentre la sua fase è una funzione dispari, cioè $ph(c_{-n}) = -ph(c_n)$. Questo implica che il coefficiente c_n dello spettro è complesso ed è uguale al coniugato di c_{-n} , cioè $c_n = \bar{c}_{-n}$. Questa proprietà sarà dimostrata nel caso generale, dopo aver introdotto la trasformata di Fourier.

2.5 Serie di Fourier: forme alternative

Presentiamo ora alcune forme alternative per la serie di Fourier, utili soprattutto per segnali $f(t)$ a valori reali. Se $f(t)$ è a valori reali, abbiamo osservato che il coefficiente c_n dello spettro è complesso ed è uguale al coniugato di c_{-n} , cioè $c_n = \bar{c}_{-n}$. Questo fatto ci consente di derivare una forma alternativa della serie di Fourier. Infatti, consideriamo lo sviluppo in serie di Fourier:

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i \frac{n}{T} t}, \quad \text{dove} \quad c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) e^{2\pi i \frac{n}{T} t} dt$$

Sostituendo c_{-n} con $\bar{c}_n$ otteniamo:

$$\begin{aligned} f(t) &= c_0 + \sum_{n=1}^{\infty} \left[c_n e^{2\pi i \frac{n}{T} t} + c_{-n} e^{-2\pi i \frac{n}{T} t} \right] \\ &= c_0 + \sum_{n=1}^{\infty} \left[c_n e^{2\pi i \frac{n}{T} t} + \bar{c}_n e^{-2\pi i \frac{n}{T} t} \right] \\ &= c_0 + \sum_{n=1}^{\infty} \left[c_n e^{2\pi i \frac{n}{T} t} + \overline{c_n e^{2\pi i \frac{n}{T} t}} \right] \\ &= c_0 + \sum_{n=1}^{\infty} 2 \operatorname{Re} \left\{ c_n e^{2\pi i \frac{n}{T} t} \right\}. \end{aligned}$$

Se esprimiamo c_n come $c_n = r_n e^{i\theta_n}$, la precedente equazione diventa:

$$f(t) = c_0 + \sum_{n=1}^{\infty} 2 \operatorname{Re} \left\{ r_n e^{2\pi i \frac{n}{T} t + i\theta_n} \right\} = c_0 + 2 \sum_{n=1}^{\infty} r_n \cos(2\pi i \frac{n}{T} t + \theta_n). \quad (2.1)$$

La precedente equazione è nota come *forma trigonometrica combinata* della serie di Fourier.

E' possibile derivare un'ulteriore forma riscrivendo c_n in forma algebrica $c_n = a_n - ib_n$, con a_n e b_n reali. Ricordando che $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$, otteniamo:

$$f(t) = c_0 + 2 \sum_{n=1}^{\infty} \left[a_n \cos(2\pi i \frac{n}{T} t) + b_n \sin(2\pi i \frac{n}{T} t) \right]. \quad (2.2)$$

L'equazione (2.2) rappresenta la serie di Fourier in *forma trigonometrica*; essa coincide con la formulazione originale di Fourier.

Una funzione $f(t)$ è detta *pari* se $f(-t) = f(t)$, mentre se $f(-t) = -f(t)$ è detta *dispari*. Il segnale $\cos 2\pi \nu t$ è un esempio di funzione pari, mentre $\sin 2\pi \nu t$ è un esempio di funzione dispari; si osservi che la combinazione lineare di funzioni pari è pari e la combinazione lineare di funzioni dispari è dispari.

Funzioni pari (o dispari) ammettono una rappresentazione in forma trigonometrica particolarmente semplice; se infatti $f(t)$ è pari, vale:

$$f(t) = c_0 + 2 \sum_{n=1}^{\infty} a_n \cos(2\pi i \frac{n}{T} t)$$

Se invece $f(t)$ è dispari, vale:

$$f(t) = 2 \sum_{n=1}^{\infty} b_n \sin(2\pi i \frac{n}{T} t)$$

Esempio 2.5.1

Si consideri il seguente segnale periodico di periodo $T = \frac{2\pi}{\omega_0}$:

$$f(t) = 10 + 3 \cos \omega_0 t + 5 \cos(2\omega_0 t + \frac{\pi}{6}) + 4 \sin 3\omega_0 t.$$

Sostituendo i seni e i coseni con le esponenziali complesse:

$$f(t) = 10 + \frac{3}{2} [e^{i\omega_0 t} + e^{-i\omega_0 t}] + \frac{5}{2} [e^{i(2\omega_0 t + \frac{\pi}{6})} + e^{-i(2\omega_0 t + \frac{\pi}{6})}] + \frac{4}{2i} [e^{i3\omega_0 t} - e^{-i3\omega_0 t}]$$

o, equivalentemente:

$$f(t) = 10 + \frac{3}{2} e^{i\omega_0 t} + \frac{3}{2} e^{-i\omega_0 t} + \frac{5}{2} e^{i\frac{\pi}{6}} e^{i2\omega_0 t} + \frac{5}{2} e^{-i\frac{\pi}{6}} e^{-i2\omega_0 t} - 2ie^{i3\omega_0 t} + 2ie^{-i3\omega_0 t}$$

I coefficienti della serie di Fourier per il segnale $f(t)$ sono dunque:

$$\begin{aligned} c_0 &= 10, & c_1 &= c_{-1} = \frac{3}{2}, & c_2 &= \frac{5}{2} \left(\frac{\sqrt{3}}{2} + i\frac{1}{2} \right), \\ c_{-2} &= \frac{5}{2} \left(\frac{\sqrt{3}}{2} - i\frac{1}{2} \right), & c_3 &= -c_{-3} = -2i. \end{aligned}$$

Tutti gli altri coefficienti risultano nulli.

2.6 Serie di Fourier: problemi di convergenza

Abbiamo visto che un segnale periodico $f(t)$ “sufficientemente regolare” può essere ottenuto come somma infinita di componenti sinusoidali. Nella pratica tuttavia solo un numero finito di tali componenti potranno essere considerate; il problema diventa quello di approssimare una funzione periodica $f(t)$ con la somma parziale di Fourier

$$S_N(t) = \sum_{n=-N}^N c_n e^{in\frac{2\pi}{T}t},$$

dove i coefficienti c_n sono quelli dati nell'espansione in serie di Fourier di $f(t)$. Discutiamo brevemente in questo paragrafo cosa significa “sufficientemente regolare” analizzando come le somme parziali di Fourier convergano.

Si dice che una funzione $f(t)$ periodica di periodo T soddisfa le *condizioni di Dirichlet* se:

1. $f(t)$ ha al più un numero finito di discontinuità nel periodo;
2. $f(t)$ ha al più un numero finito di massimi e minimi nel periodo;
3. $f(t)$ è limitata; per poter trattare le funzioni impulsive, tale richiesta viene estesa a quella di assoluta integrabilità per la funzione, ossia:

$$\int_T |f(t)| dt < +\infty.$$

Per ogni funzione $f(t)$ che rispetta le condizioni di Dirichlet, le seguenti asserzioni risultano vere.

- La successione delle somme parziali di Fourier converge a $f(p)$ in ogni punto p di continuità.
- Se la funzione ha una discontinuità isolata in un punto p , la serie converge al valor medio del limite sinistro e limite destro della funzione in quel punto, cioè:

$$\sum_{n=-\infty}^{+\infty} c_n e^{in \frac{2\pi}{T} p} = \frac{f(p^-) + f(p^+)}{2}.$$

Senza la pretesa di dare una trattazione matematica rigorosa, ci limitiamo infine a segnalare un fenomeno studiato in questa teoria (noto come fenomeno di Gibbs): in certi casi la successione delle somme parziali converge in ogni punto, ma non uniformemente. Ad esempio, nella Figura 2.4 viene illustrata la convergenza alla serie $S(t) = \sum_{n=-\infty}^{+\infty} r_n e^{int}$, dove $S(t)$ è una funzione periodica di periodo 2π definita come:

$$S(t) = \begin{cases} 1, & \text{se } |t| < 1 \\ 0, & \text{se } 1 < |t| < \pi. \end{cases}$$

Come si nota dalla figura, esiste un valore $\delta > 0$ tale che, per ogni N :

$$\sup_{-\pi < t < \pi} |S(t) - S_N(t)| \geq \delta.$$

Si osservi dalla stessa figura che i punti t_N , per cui $|S(t_N) - S_N(t_N)| \approx \delta$, per N grande sono prossimi ai punti di discontinuità di $S(t)$.

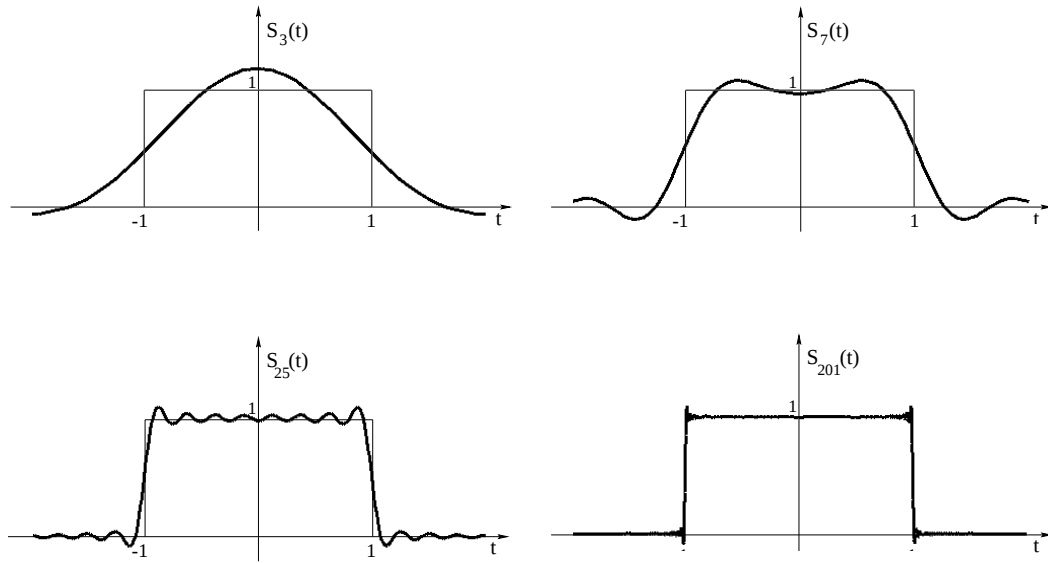


Figura 2.4 Convergenza delle somme parziali $S_N(t) = \sum_{n=-N}^N r_n e^{int}$.

2.7 Segnali a spettro continuo: la trasformata di Fourier

Abbiamo visto che i segnali periodici di periodo T sono ottenibili come combinazione lineare della componente continua 1 e delle armoniche $e^{i2\pi \frac{n}{T}t}$, le cui frequenze sono $\pm \frac{1}{T}, \pm \frac{2}{T}, \pm \frac{n}{T}, \pm \frac{n+1}{T}, \dots$

Tuttavia molti segnali di interesse non sono periodici e quindi la teoria precedentemente esposta non si può applicare direttamente. Questa sezione è dedicata coprire questa lacuna, attraverso l'introduzione della *trasformata di Fourier*, che permette di ottenere lo spettro di un qualsiasi segnale.

Giustificiamo l'introduzione della trasformata a partire dalla serie di Fourier. L'idea principale è la seguente: dato un segnale non periodico $f(t)$, fissato T , si considera il segnale, $f_T(t)$, periodico di periodo T e coincidente con $f(t)$ nell'intervallo $[-\frac{T}{2}, \frac{T}{2}]$.

Possiamo allora pensare che:

$$f(t) = \lim_{T \rightarrow \infty} f_T(t)$$

Poichè $f_T(t)$ è di periodo T , tale segnale può essere sviluppata in serie di Fourier:

$$f_T(t) = \sum_{n=-\infty}^{\infty} c_n e^{i \frac{2\pi}{T} n t},$$

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f_T(t) e^{-i \frac{2\pi}{T} n t} dt.$$

Chiamiamo ora $\nu_n = \frac{n}{T}$ la frequenza dell'ennesima armonica e $\Delta\nu = \nu_{n+1} - \nu_n = \frac{1}{T}$ l'incremento tra una frequenza e la successiva; ponendo infine $F(\nu_n) = c_n T$, le precedenti espressioni si scrivono:

$$f_T(t) = \sum_{n=-\infty}^{\infty} F(\nu_n) e^{i2\pi\nu_n t} \Delta\nu,$$

$$F(\nu_n) = \int_{-\frac{T}{2}}^{\frac{T}{2}} f_T(t) e^{-i2\pi\nu_n t} dt.$$

Per $T \rightarrow \infty$ vale che $f_T(t)$ coincide con $f(t)$ e $\Delta\nu \rightarrow 0$ e quindi la serie precedente si riduce ad un integrale definito, ottenendo:

$$f(t) = \int_{-\infty}^{\infty} F(\nu) e^{i2\pi\nu t} d\nu,$$

$$F(\nu) = \int_{-\infty}^{\infty} f(t) e^{-i2\pi\nu t} dt.$$

La funzione $F(\nu)$ è detta *trasformata di Fourier* di $f(t)$ e può essere vista come la rappresentazione in frequenza del segnale $f(t)$. La trasformata di Fourier $F(\nu)$ identifica quindi lo *spettro* del segnale $f(t)$; $F(\nu)$ è essenzialmente il peso della componente armonica $e^{i2\pi\nu t}$, e generalmente un segnale non periodico è ottenuto come la sovrapposizione di componenti armoniche $e^{i2\pi\nu t}$, per valori reali di ν : un segnale non periodico può contenere tutte le frequenze.

Si può notare che la corrispondenza $f(t) \leftrightarrow F(\nu)$ è una corrispondenza biunivoca e lineare. Tale corrispondenza individua un operatore $\mathcal{F}$, detto *trasformata di Fourier*, definito da $F(\nu) = \mathcal{F}\{f(t)\}$, unitamente all'operatore inverso $\mathcal{F}^{-1}$, detto *antitrasformata di Fourier*, definito da $f(t) = \mathcal{F}^{-1}\{F(\nu)\}$.

Il *supporto* dello spettro $F(\nu)$ è dato dall'insieme $\{\nu : F(\nu) \neq 0\}$. Risultano di particolare interesse i segnali a supporto limitato, detti anche a *banda limitata*: un segnale $f(t)$ è detto a *banda limitata* dalla frequenza B se per $|\nu| > B$ risulta $F(\nu) = 0$. In Figura 2.5 è mostrato il modulo di un tipico spettro $V(\nu)$ della voce umana.

L'industria telefonica si avvantaggia della forma dello spettro della voce per ridurre la dimensione della banda di frequenza necessaria alla trasmissione troncandone le componenti oltre i 3000 Hz; come si vede in Figura 2.6, per molti aspetti il "peso" delle componenti a frequenza superiore a $3K\text{Hertz}$ può essere considerato trascurabile.

La trasformata di Fourier gioca per i segnali aperiodici un ruolo analogo a quello della serie di Fourier per i segnali periodici, poichè entrambe esprimono il segnale come

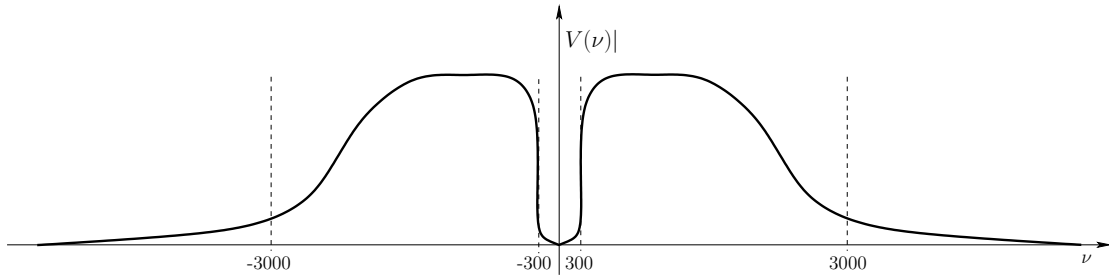


Figura 2.5 Tipico spettro (modulo) della voce umana.

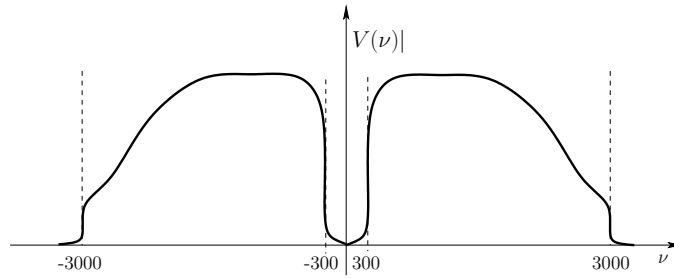


Figura 2.6 Tipico spettro troncato della voce umana.

combinazione lineare di funzioni sinusoidali di data frequenza, espresse come esponenziali complessi.

Per i segnali periodici, questi esponenziali complessi sono definiti per frequenze discrete $n\frac{1}{T}$ ($-\infty < n < +\infty$) e sono pesati con ampiezze c_n .

Per segnali non periodici, gli esponenziali complessi sono definiti su un continuo di frequenze ν ($-\infty < \nu < +\infty$) e sono pesati con ampiezze pari a $F(\nu)d\nu$.

Esempio 2.7.1

Si consideri il segnale

$$f(t) = e^{-at}u(t), \quad a > 0.$$

La trasformata di Fourier di $f(t)$ è

$$F(\nu) = \int_0^{+\infty} e^{-at} e^{-i2\pi\nu t} dt = -\frac{1}{a + i2\pi\nu} e^{-(a+i2\pi\nu)t} \Big|_0^{\infty} = \frac{1}{a + i2\pi\nu}, \quad a > 0.$$

Poiché questa trasformata è a valori complessi, $F(\nu)$ può essere rappresentata mediante due grafici, rispettivamente del suo modulo e della sua fase, (Figura 2.7), in cui modulo e fase sono dati da:

$$|F(\nu)| = \frac{1}{\sqrt{a^2 + (2\pi\nu)^2}}, \quad ph(F(\nu)) = -\arctan \frac{2\pi\nu}{a}.$$

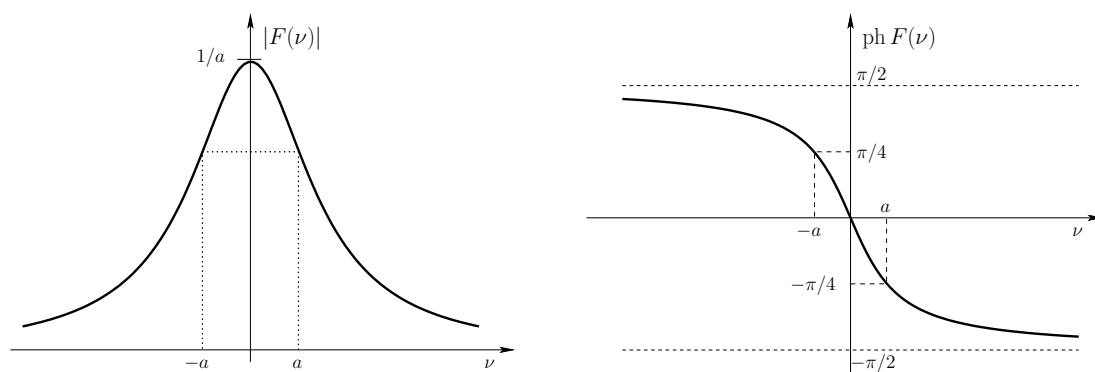


Figura 2.7 Modulo e fase della trasformata di Fourier di $f(t) = e^{-at}u(t)$, $a > 0$.

Esempio 2.7.2

Si consideri il segnale impulso rettangolare

$$\text{rect}(t) = \begin{cases} 1 & \text{se } |t| \leq \frac{1}{2} \\ 0 & \text{se } |t| > \frac{1}{2} \end{cases}$$

La trasformata di Fourier di questo segnale è la funzione reale

$$F(\nu) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-i2\pi\nu t} dt = \text{sinc}(\nu).$$

In Figura 2.8 è rappresentato il segnale impulso rettangolare e la sua trasformata di Fourier.

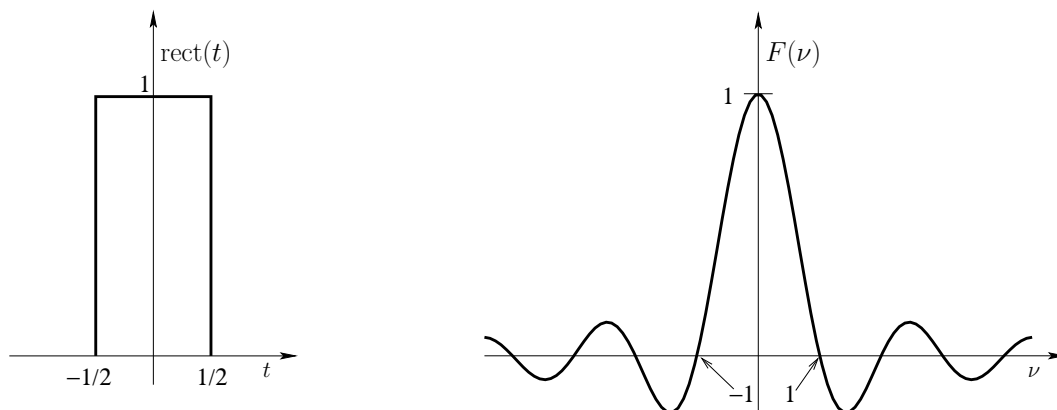


Figura 2.8 Il segnale impulso rettangolare e la sua trasformata di Fourier.

2.7.1 Esistenza della Trasformata di Fourier

Nella sezione precedente abbiamo ottenuto la forma generale della trasformata di Fourier e della sua inversa; qui ci proponiamo di discutere brevemente sotto quali condizioni tale trasformata esista.

In generale, perchè la trasformata di una data funzione $f(t)$ esista, deve esistere ed essere finito per ogni t il valore:

$$\int_{-\infty}^{+\infty} f(t)e^{-i2\pi\nu t} dt.$$

Un insieme di condizioni sufficienti per l'esistenza della trasformata di Fourier è dato dalle cosiddette condizioni di Dirichlet, elencate di seguito:

1. $f(t)$ deve essere assolutamente integrabile, cioè

$$\int_{-\infty}^{+\infty} |f(t)| dt < \infty;$$

2. $f(t)$ deve avere un numero finito di minimi e di massimi in ogni intervallo finito;
3. $f(t)$ deve avere un numero finito di discontinuità in ogni intervallo finito.

In particolare, per quanto riguarda l'ultima condizione si deve aggiungere che nei punti di discontinuità in cui il limite destro ed il limite sinistro della $f(t)$ sono diversi, la trasformata inversa converge al valore medio dei due limiti stessi.

Tutte le funzioni di interesse pratico che considereremo di seguito soddisfano le condizioni 2. e 3. sopra citate. Utilizzeremo invece spesso funzioni (o funzioni generalizzate) che non verificano la prima condizione, pur ammettendo trasformata di Fourier. Alcuni importanti esempi sono elencati di seguito.

Esempio 2.7.3

In questo esempio calcoliamo la trasformata di Fourier della delta di Dirac $\delta(t)$ e la trasformata di Fourier della funzione costante 1. Si osservi che in quest'ultimo caso la prima condizione di Dirichlet non è verificata.

Ricordando che, se $f(t)$ è una funzione continua in 0, allora vale:

$$\int_{-\infty}^{+\infty} f(t)\delta(t)dt = f(0),$$

possiamo concludere che:

$$\mathcal{F}\{\delta(t)\} = \int_{-\infty}^{+\infty} \delta(t)e^{-i2\pi\nu t} dt = e^{-i2\pi\nu t}\big|_{t=0} = 1.$$

Usando il fatto che la trasformata di Fourier è unica, dalla precedente espressione segue immediatamente che:

$$\mathcal{F}^{-1}\{1\} = \int_{-\infty}^{+\infty} 1 \cdot e^{i2\pi\nu t} d\nu = \delta(t). \quad (2.3)$$

Possiamo quindi concludere che la trasformata di Fourier della funzione costante $f(t) = 1$ è $\delta(t)$.

Esempio 2.7.4

Un importante esempio di funzione usata nei sistemi di comunicazione è la funzione signum o sgn , definita come segue:

$$\text{sgn}(t) = \begin{cases} 1, & \text{per } t > 0 \\ 0, & \text{per } t = 0 \\ -1, & \text{per } t < 0 \end{cases}.$$

Anche in questo caso la funzione non è assolutamente integrabile, quindi occorre trovare un diverso metodo di integrazione per calcolarne la trasformata. A tal riguardo si consideri la seguente uguaglianza di facile verifica:

$$\text{sgn}(t) = \lim_{a \rightarrow 0} e^{-a|t|} \text{sgn}(t).$$

La sua trasformata di Fourier risulta pertanto essere:

$$\mathcal{F}\{\text{sgn}(t)\} = \mathcal{F}\left\{\lim_{a \rightarrow 0} e^{-a|t|} \text{sgn}(t)\right\} = \lim_{a \rightarrow 0} \mathcal{F}\left\{e^{-a|t|} \text{sgn}(t)\right\},$$

in cui l'operazione di limite e di integrazione sono stati scambiati senza dare una giustificazione matematica precisa. Poichè:

$$\mathcal{F}\left\{e^{-a|t|} \text{sgn}(t)\right\} = - \int_{-\infty}^0 e^{(a-i2\pi\nu)t} dt + \int_0^{\infty} e^{-(a+i2\pi\nu)t} dt = -\frac{1}{a-i2\pi\nu} + \frac{1}{a+i2\pi\nu},$$

possiamo concludere che:

$$\mathcal{F}\{\text{sgn}(t)\} = \lim_{a \rightarrow 0} -\frac{1}{a-i2\pi\nu} + \frac{1}{a+i2\pi\nu} = \frac{1}{i\pi\nu}. \quad (2.4)$$

Esempio 2.7.5

Una funzione, legata alla funzione signum da una semplice trasformazione, è la funzione gradino unitario $u(t)$, definita nell'Esempio 1.6.7. Infatti essa può essere scritta come:

$$u(t) = \frac{1}{2} + \frac{1}{2} \text{sgn}(t),$$

da cui otteniamo la sua trasformata di Fourier:

$$\begin{aligned} \mathcal{F}\{u(t)\} &= \frac{1}{2} \mathcal{F}\{1\} + \frac{1}{2} \mathcal{F}\{\text{sgn}(t)\} && (\text{per la linearità}) \\ &= \frac{1}{2} \delta(\nu) + \frac{1}{i2\pi\nu}. \end{aligned}$$

2.7.2 Trasformata di Fourier di Funzioni a valori Reali

La trasformata di Fourier $F(\nu)$ di un segnale reale $f(t)$ è in generale una funzione a valori complessi. Per visualizzare graficamente $F(\nu)$ è allora necessario considerare separatamente il modulo $|F(\nu)|$ e la fase $phF(\nu)$ della trasformata, le cui espressioni analitiche sono:

$$|F(\nu)| = \sqrt{(\operatorname{Re}\{F(\nu)\})^2 + (\operatorname{Im}\{F(\nu)\})^2},$$

$$phF(\nu) = \arctan\left(\frac{\operatorname{Im}\{F(\nu)\}}{\operatorname{Re}\{F(\nu)\}}\right).$$

Dato un segnale reale $f(t)$ che ammette $F(\nu)$ come trasformata, a causa dell'identità $e^{i2\pi\nu t} = \cos(2\pi\nu t) + i\sin(2\pi\nu t)$, la parte reale $\operatorname{Re}\{F(\nu)\}$ e la parte immaginaria $\operatorname{Im}\{F(\nu)\}$ sono rispettivamente:

$$\operatorname{Re}\{F(\nu)\} = \int_{-\infty}^{+\infty} f(t) \cos(2\pi\nu t) dt, \quad \operatorname{Im}\{F(\nu)\} = \int_{-\infty}^{+\infty} f(t) \sin(2\pi\nu t) dt.$$

Poiché la funzione coseno è pari e la funzione seno è dispari, si ha che:

$$\operatorname{Re}\{F(\nu)\} = \operatorname{Re}\{F(-\nu)\}, \quad \operatorname{Im}\{F(\nu)\} = -\operatorname{Im}\{F(-\nu)\}.$$

Vale di conseguenza che:

$$|F(-\nu)| = |F(\nu)|, \quad phF(-\nu) = -phF(\nu).$$

Ne consegue che per funzioni reali il modulo della trasformata $|F(\nu)|$ è una funzione pari e la fase $phF(\nu)$ è una funzione dispari. In Figura 2.9 è mostrato un tipico esempio del modulo e della fase della trasformata di una funzione reale (si veda anche l'Esempio 2.7.1).

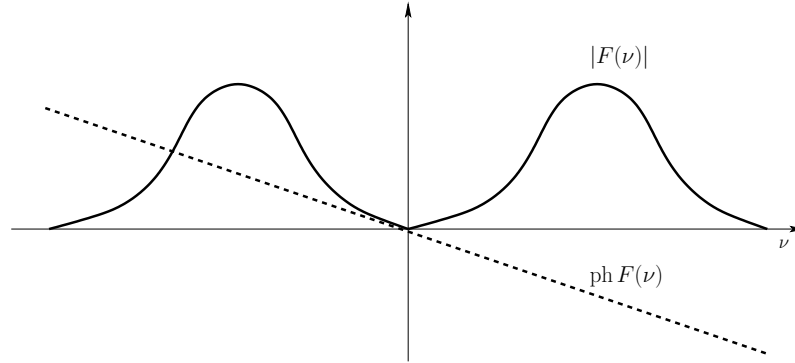


Figura 2.9 Modulo e fase della trasformata di Fourier di un segnale $f(t)$ reale.

$|F(\nu)|$ e $phF(\nu)$ (e quindi $F(\nu)$) sono quindi ricostruibili a partire da valori di $|F(\nu)|$ e $phF(\nu)$ per $\nu \geq 0$. In questo senso l'informazione dello spettro di un segnale reale è portata dai valori positivi di ν .

Tabella 2.1 Proprietà delle Trasformate di Fourier.

Proprietà	$\mathbf{f(t)}$	$\leftrightarrow$	$\mathbf{F(\nu)}$
Linearità	$af(t) + bg(t)$		$aF(\nu) + bG(\nu)$
Dualità (simmetria)	$F(t)$		$f(-\nu)$
Traslazione (tempo)	$f(t - t_0)$		$e^{-i2\pi\nu t_0} F(\nu)$
Traslazione (frequenza)	$e^{i2\pi\nu_0 t} f(t)$		$F(\nu - \nu_0)$
Convoluzione (tempo)	$\int_{-\infty}^{+\infty} f(x)g(t - x)dx$		$F(\nu)G(\nu)$
Convoluzione (frequenza)	$f(t)g(t)$		$\int_{-\infty}^{+\infty} F(y)G(\nu - y)dy$
Modulazione	$f(t) \cos(2\pi\nu_0 t)$		$\frac{1}{2}[F(\nu + \nu_0) + F(\nu - \nu_0)]$
Scalatura	$f(at)$		$\frac{1}{ a } F\left(\frac{\nu}{a}\right)$
Differenziazione	$\frac{d^n}{dt^n} f(t)$		$(i2\pi\nu)^n F(\nu)$

2.7.3 Proprietà della Trasformata di Fourier

Nella Tabella 2.1 vengono riportate le principali proprietà della trasformata di Fourier, il cui contributo principale è la comprensione della relazione tra dominio del tempo e dominio delle frequenze dei segnali. Pur rimandando alla letteratura specializzata le dimostrazioni matematiche formali, si ritiene utile dare una giustificazione delle principali proprietà.

La familiarità all'uso di queste proprietà agevola di molto lo studio e la comprensione dei problemi relativi al campionamento, ai filtri e alla modulazione di ampiezza.

Linearità

$$\begin{aligned}
 \mathcal{F}\{af(t) + bg(t)\} &= \int_{-\infty}^{+\infty} (af(t) + bg(t))e^{-i2\pi\nu t} dt \\
 &= a \int_{-\infty}^{+\infty} f(t)e^{-i2\pi\nu t} dt + b \int_{-\infty}^{+\infty} g(t)e^{-i2\pi\nu t} dt \\
 &= a\mathcal{F}\{f(t)\} + b\mathcal{F}\{g(t)\}
 \end{aligned}$$

Dualità (simmetria)

Ricordando che:

$$\mathcal{F}\{f(t)\} = F(\nu) = \int_{-\infty}^{+\infty} f(t)e^{-i2\pi\nu t} dt, \quad (2.5)$$

$$\mathcal{F}^{-1}\{F(\nu)\} = f(t) = \int_{-\infty}^{+\infty} F(\nu)e^{i2\pi\nu t} d\nu, \quad (2.6)$$

ed osservando la simmetria esistente tra la trasformata (equazione (2.5)) e la sua inversa (equazione (2.6)) rispettivamente nella variabile t e nella variabile ν , vale allora:

$$\begin{aligned} \mathcal{F}\{F(t)\} &= \int_{-\infty}^{+\infty} F(t)e^{-i2\pi\nu t} dt \\ &= \int_{-\infty}^{+\infty} F(t)e^{i2\pi(-\nu)t} dt \\ &= f(-\nu). \end{aligned}$$

Traslazione nel tempo e in frequenza

Riguardo la traslazione nel tempo si ha che:

$$\begin{aligned} \mathcal{F}\{f(t - t_0)\} &= \int_{-\infty}^{+\infty} f(t - t_0)e^{-i2\pi\nu t} dt \\ &= \int_{-\infty}^{+\infty} f(\tau)e^{-i2\pi\nu(\tau+t_0)} d\tau \quad (\text{con } \tau = t - t_0, \text{ da cui } dt = d\tau) \\ &= e^{-i2\pi\nu t_0} F(\nu). \end{aligned} \quad (2.7)$$

La proprietà di traslazione in frequenza si verifica facilmente combinando la proprietà di dualità e quella di traslazione nel tempo.

Convoluzione nel tempo e in frequenza

Per la proprietà di convoluzione nel tempo si ha:

$$\begin{aligned} \mathcal{F}\left\{\int_{-\infty}^{+\infty} f(\tau)g(t - \tau)d\tau\right\} &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(\tau)g(t - \tau)d\tau\right) e^{-i2\pi\nu t} dt \quad (\text{per la (??)}) \\ &= \int_{-\infty}^{+\infty} f(\tau) \left(\int_{-\infty}^{+\infty} g(t - \tau)e^{-i2\pi\nu t} dt\right) d\tau \\ &= \int_{-\infty}^{+\infty} f(\tau)e^{-i2\pi\nu\tau} G(\nu)d\tau \quad (\text{per la traslaz. nel tempo (2.7)}) \\ &= G(\nu) \int_{-\infty}^{+\infty} f(\tau)e^{-i2\pi\nu\tau} d\tau \\ &= G(\nu)F(\nu) \end{aligned}$$

La convoluzione in frequenza si giustifica applicando la proprietà di dualità a quella di convoluzione nel tempo.

Modulazione

$$\begin{aligned}
 \mathcal{F}\{f(t)\cos(2\pi\nu_0 t)\} &= \int_{-\infty}^{+\infty} f(t)\cos(2\pi\nu_0 t)e^{-i2\pi\nu t} dt \\
 &= \int_{-\infty}^{+\infty} f(t)\frac{e^{i2\pi\nu_0 t} + e^{-i2\pi\nu_0 t}}{2}e^{-i2\pi\nu t} dt \\
 &= \frac{1}{2}\int_{-\infty}^{+\infty} f(t)e^{-i2\pi(\nu+\nu_0)t} dt + \frac{1}{2}\int_{-\infty}^{+\infty} f(t)e^{-i2\pi(\nu-\nu_0)t} dt \\
 &= \frac{1}{2}[F(\nu+\nu_0) + F(\nu-\nu_0)].
 \end{aligned}$$

Scalatura

Se a è un reale positivo si ha che:

$$\begin{aligned}
 \mathcal{F}\{f(at)\} &= \int_{-\infty}^{+\infty} f(at)e^{-i2\pi\nu t} dt \\
 &= \frac{1}{a}\int_{-\infty}^{+\infty} f(\tau)e^{-i\frac{2\pi\nu}{a}\tau} d\tau \quad (\text{con } \tau = at, \text{ da cui } dt = d\tau/a) \\
 &= \frac{1}{a}F\left(\frac{\nu}{a}\right).
 \end{aligned}$$

Analogamente, se a è un reale negativo si ha che $\mathcal{F}\{f(at)\} = -\frac{1}{a}F\left(\frac{\nu}{a}\right)$.

Differenziazione

Differenziando entrambi i membri della (2.6) si ha:

$$\frac{d}{dt}f(t) = \int_{-\infty}^{+\infty} i2\pi\nu F(\nu)e^{i2\pi\nu t} d\nu,$$

da cui

$$\mathcal{F}\left\{\frac{d}{dt}f(t)\right\} = i2\pi\nu F(\nu).$$

2.7.4 Energia e Potenza

Consideriamo una resistenza R , ai cui capi è applicata una tensione $V(t)$; essa sarà attraversata da una corrente $i(t) = V(t)/R$ così che la potenza $P(t)$ risulta $P(t) = V(t) \cdot i(t) = \frac{V^2(t)}{R}$: la potenza è quindi proporzionale al quadrato del segnale $V(t)$.

Generalizzando, diremo che la *potenza istantanea* di un segnale $f(t)$ è $|f(t)|^2$. Ne segue che l'energia $E(f)$ di un segnale $f(t)$ su un intervallo finito $t_1 \leq t \leq t_2$ può essere definita come

$$E(f) = \int_{t_1}^{t_2} |f(t)|^2 dt,$$

Nel caso in cui $\lim_{T \rightarrow +\infty} \int_{-T}^T |f(t)|^2 dt < +\infty$, possiamo concludere che l'energia complessiva del segnale è

$$E_{\infty}(f) = \int_{-\infty}^{+\infty} |f(t)|^2 dt.$$

$E_{\infty}(f)$ è detta *energia totale* (o più semplicemente *energia*) del segnale $f(t)$.

Un segnale $f(t)$ che soddisfa la condizione $E_{\infty} < \infty$ viene detto *a energia finita*. Esempi di segnali a energia finita sono i segnali a supporto limitato, come il rettangolo, o segnali che tendono a zero in modo sufficientemente veloce quando t tende a infinito.

Esempio 2.7.6

Si vuol calcolare l'energia del segnale $f(t) = Ae^{-at}u(t)$, con $a > 0$. Vale che:

$$E_{\infty}(f) = \int_{-\infty}^{+\infty} |Ae^{-at}u(t)|^2 dt = \int_0^{\infty} A^2 e^{-2at} dt = \frac{A^2}{2a}$$

Ci sono segnali $f(t)$ per cui $\lim_{T \rightarrow +\infty} \int_{-T}^T |f(t)|^2 dt = +\infty$, cioè l'energia del segnale è infinita. Questa è la situazione, ad esempio, di tutti i segnali periodici; in questo caso è utile introdurre il concetto di potenza media. Si osservi infatti che la potenza media P_T di un segnale $f(t)$ nell'intervallo $-T \leq t \leq T$ può essere definita come

$$P_T(f) = \frac{1}{2T} \int_{-T}^T |f(t)|^2 dt,$$

Possiamo chiamare *potenza media* (o più semplicemente *potenza*) del segnale $f(t)$ il seguente valore:

$$P_{\infty}(f) = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T |f(t)|^2 dt,$$

Il concetto di potenza media non è significativo per segnali a energia finita: se infatti $E_{\infty}(f) < \infty$, allora $P_{\infty}(f) = 0$. Esistono invece segnali per cui $E_{\infty}(f) = \infty$ ma $P_{\infty}(f) < \infty$: è il caso, ad esempio, di tutti i segnali periodici. In questo caso è il concetto di potenza ad essere significativo e chiameremo *a potenza finita* tali segnali.

Esempio 2.7.7

Si vuol calcolare la potenza media dei segnali $f(t) = Au(t)$ e $g(t) = A \cos 2\pi \nu t$.

Si verifica facilmente che per tali segnali l'energia è infinita. Per quanto riguarda la potenza, vale:

$$P_{\infty}(f) = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T |Au(t)|^2 dt = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_0^T A^2 dt = \lim_{T \rightarrow +\infty} \frac{A^2 T}{2T} = \frac{A^2}{2}$$

$$P_{\infty}(g) = \lim_{T \rightarrow +\infty} \frac{A^2}{2T} \int_{-T}^T \cos^2 2\pi \nu t \cdot dt = \lim_{T \rightarrow +\infty} \frac{A^2}{2T} \int_{-T}^T \frac{1 + \cos 4\pi \nu t}{2} \cdot dt =$$

$$= \lim_{T \rightarrow +\infty} \frac{A^2}{2T} \left(T + \frac{1}{2\pi\nu} \cos 4\pi\nu t \right) = \frac{A^2}{2}$$

Come suggerito dall'analisi del segnale $g(t)$, i segnali periodici hanno tutti energia infinita, ma potenza finita.

In molte situazioni è noto lo spettro di un segnale, e non la sua rappresentazione temporale; il calcolo dell'energia richiederebbe, in questo caso, di calcolare preliminarmente la antitrasformata dello spettro. Fortunatamente questo si può evitare, in quanto è possibile esprimere l'energia di un segnale direttamente a partire dal suo spettro, utilizzando un'importante proprietà della trasformata di Fourier data dal seguente:

Fatto 2.1 (Teorema di Parseval) *Se $f(t)$ è un segnale continuo e $F(\nu)$ la sua trasformata di Fourier, allora:*

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-\infty}^{+\infty} |F(\nu)|^2 d\nu. \quad (2.8)$$

Dimostrazione. La relazione (2.8) segue dalla diretta applicazione della definizione di trasformata di Fourier. Infatti:

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-\infty}^{+\infty} f(t) \bar{f}(t) dt = \int_{-\infty}^{+\infty} f(t) \left[\int_{-\infty}^{+\infty} \bar{F}(\nu) e^{-2\pi i \nu t} d\nu \right] dt$$

Invertendo l'ordine di integrazione si ha

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-\infty}^{+\infty} \bar{F}(\nu) d\nu \left[\int_{-\infty}^{+\infty} f(t) e^{-2\pi i \nu t} dt \right]$$

Il termine tra parentesi quadre è la trasformata di Fourier di $f(t)$, quindi

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-\infty}^{+\infty} |F(\nu)|^2 d\nu.$$

■

A causa del teorema di Parseval risulta dunque:

$$E_{\infty}(f) = \int_{-\infty}^{+\infty} |F(\nu)|^2 d\nu.$$

Dato un segnale a energia finita $f(t)$, possiamo allora attribuire al quadrato del modulo della trasformata di Fourier $|F(\nu)|^2$ il seguente significato: $|F(\nu)|^2 d\nu$ è il contributo all'energia del segnale dato dalle componenti sinusoidali con frequenza compresa tra ν e $\nu + d\nu$. La funzione $|F(\nu)|^2$ è usualmente chiamata *densità spettrale di energia* di $f(t)$.

Per segnali a potenza finita si può ottenere un risultato analogo. Sia infatti $f(t)$ un segnale a energia infinita ma con $P_{\infty}(f) < \infty$. Consideriamo il segnale $f_T(t)$ che coincide

con $f(t)$ nell'intervallo $[-T, T]$, ma che risulta nullo esternamente a tale intervallo; dalla precedente definizione segue che:

$$f(t) = \lim_{T \rightarrow +\infty} f_T(t)$$

Il segnale $f_T(t)$ è a energia finita e, detta $F_T(\nu)$ la sua trasformata di Fourier, a causa del teorema di Parseval risulta $\int_{-\infty}^{+\infty} |f_T(t)|^2 dt = \int_{-\infty}^{+\infty} |F_T(\nu)|^2 d\nu$. Possiamo concludere:

$$P_\infty(f) = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-\infty}^{+\infty} |f_T(t)|^2 dt = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-\infty}^{+\infty} |F_T(\nu)|^2 d\nu = \int_{-\infty}^{+\infty} \lim_{T \rightarrow +\infty} \frac{1}{2T} |F_T(\nu)|^2 d\nu$$

La quantità $\lim_{T \rightarrow +\infty} \frac{1}{2T} |F_T(\nu)|^2$, quando il limite esiste, è chiamata *densità spettrale di potenza* ed è usualmente denotata con $S_f(\nu)$. Vale dunque:

$$P_\infty(f) = \int_{-\infty}^{+\infty} S_f(\nu) d\nu.$$

Esempio 2.7.8

Si vuol determinare la densità spettrale di potenza del gradino unitario $u(t)$.

Chiamiamo $S_u(\nu)$ la densità di potenza spettrale del gradino unitario $u(t)$. La restrizione di $u(t)$ all'intervallo $[-T, T]$ è la funzione $\text{rect} \frac{t-T/2}{T}$, la cui trasformata $U_T(t)$ è calcolabile mediante le proprietà di traslazione e scalatura, e verifica:

$$|U_T(t)|^2 = |T \cdot \text{sinc} T\nu|^2 = \frac{\sin^2 \pi T\nu}{\pi^2 \nu^2}$$

Se $\nu \neq 0$ vale che $S_u(\nu) = \lim_{T \rightarrow +\infty} \frac{1}{2T} |U_T(\nu)|^2 = 0$; vale inoltre che $\int_{-\infty}^{+\infty} S_u(\nu) d\nu = P_\infty(u) = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-\infty}^{+\infty} |u_T(t)|^2 dt = \frac{1}{2}$. La densità di potenza spettrale $S_u(\nu)$ verifica quindi:

- per $\nu \neq 0$ vale $S_u(\nu) = 0$
- $\int_{-\infty}^{+\infty} S_u(\nu) d\nu = \frac{1}{2}$

Ne consegue che $S_u(\nu) = \frac{1}{2} \delta(\nu)$.

2.8 Coppie Base di Trasformate

Nella Tabella 2.2 vengono riportate alcune coppie trasformata-antitrasformata di Fourier che spesso si incontrano nella pratica. Spesso il calcolo della trasformata di un segnale può essere ridotto all'applicazione a queste coppie di proprietà della trasformata. In seguito proviamo la correttezza di alcune coppie elencate in tabella.

Abbiamo già visto che $\mathcal{F}\{rect(t)\} = sinc$ e che $\mathcal{F}\{\delta(t)\} = 1$; applicando la proprietà di dualità segue che $\mathcal{F}\{sinc(t)\} = rect(\nu)$ e $\mathcal{F}\{1\} = \delta(\nu)$, mentre applicando la proprietà di traslazione si ottiene $\mathcal{F}\{\delta(t - t_0)\} = e^{2\pi i t_0 \nu}$.

Applicando di nuovo la dualità, si ottiene che $\mathcal{F}\{e^{2\pi i \nu_0 t}\} = \delta(\nu - \nu_0)$; ricordando che $\cos(2\pi \nu_0 t) = \frac{1}{2}(e^{2\pi i \nu_0 t} + e^{-2\pi i \nu_0 t})$, si ottiene $\mathcal{F}\{\cos(2\pi \nu_0 t)\} = \mathcal{F}\{\frac{1}{2}(e^{2\pi i \nu_0 t} + e^{-2\pi i \nu_0 t})\}$, e quindi, per la linearità, $\mathcal{F}\{\cos(2\pi \nu_0 t)\} = \frac{1}{2}(\delta(\nu - \nu_0) + \delta(\nu + \nu_0))$.

Ricordando infine che $\sin(2\pi \nu_0 t) = \frac{1}{2i}(e^{2\pi i \nu_0 t} - e^{-2\pi i \nu_0 t})$, si ottiene $\mathcal{F}\{\sin(2\pi \nu_0 t)\} = \frac{1}{2i}(\delta(\nu - \nu_0) - \delta(\nu + \nu_0))$.

Tabella 2.2 Alcune coppie base di trasformata-antitrasformata (a , dove compare, è un numero reale positivo).

$\mathbf{f(t)}$	$\mathbf{F(\nu)}$
$rect(t) = \begin{cases} 1 & t \leq 1/2 \\ 0 & t > 1/2 \end{cases}$	$sinc(\nu)$
$sinc(t)$	$rect(\nu)$
$\delta(t)$	1
$\delta(t - t_0)$	$e^{-i2\pi \nu t_0}$
1	$\delta(\nu)$
$e^{\pm i2\pi \nu_0 t}$	$\delta(\nu \mp \nu_0)$
$\cos(2\pi \nu_0 t)$	$\frac{1}{2}[\delta(\nu - \nu_0) + \delta(\nu + \nu_0)]$
$\sin(2\pi \nu_0 t)$	$\frac{i}{2}[\delta(\nu + \nu_0) - \delta(\nu - \nu_0)]$
$e^{-at}u(t)$	$\frac{1}{a + i2\pi \nu}$
$sgn(t)$	$\frac{1}{i\pi \nu}$
$u(t)$	$\frac{1}{2}\pi\delta(\nu) + \frac{1}{i2\pi \nu}$

Esempio 2.8.1

Si voglia calcolare $\mathcal{F}\{\frac{1}{100}sinc\frac{t}{100} - \delta(t)\}$.

Applicando le proprietà di linearità e scalatura si ottiene:

$$\mathcal{F}\{\frac{1}{100}sinc\frac{t}{100} - \delta(t)\} = \frac{1}{100}\mathcal{F}\{sinc\frac{t}{100}\} - \mathcal{F}\{\delta(t)\} = rect(100\nu) - 1$$

Esempio 2.8.2

Determinare il limite di banda del segnale $100\text{sinc}(700t) + 3\cos(600\pi t)$.

Calcolando la trasformata del segnale, si ottiene:

$$\begin{aligned}\mathcal{F}\{100\text{sinc}(700t) + 3\cos(600\pi t)\} &= 100\mathcal{F}\{\text{sinc}(700t)\} + 3\mathcal{F}\{\cos(2\pi 300t)\} \\ &= \frac{10}{7}\text{rect}\frac{\nu}{700} + \frac{3}{2}(\delta(\nu - 300) + \delta(\nu + 300))\end{aligned}$$

Il segnale contiene tutte le frequenze ν comprese tra -350 Hz e 350 Hz, quindi il limite di banda è di 350 Hz.

Esempio 2.8.3

Il calcolo della trasformata di Fourier di funzioni periodiche che può essere effettuato a partire dalla loro rappresentazione in serie di Fourier. Se $f(t)$ è periodica di periodo T , allora

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i \frac{n}{T} t}, \quad \text{con} \quad c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) e^{-2\pi i \frac{n}{T} t} dt.$$

Usando l'espansione in serie e applicando la proprietà di linearità, la trasformata di Fourier di $f(t)$ è

$$F(\nu) = \int_{-\infty}^{+\infty} \left(\sum_{n=-\infty}^{+\infty} c_n e^{2\pi i \frac{n}{T} t} \right) e^{2\pi i \nu t} dt = \sum_{n=-\infty}^{+\infty} c_n \int_{-\infty}^{+\infty} e^{2\pi i \frac{n}{T} t} e^{2\pi i \nu t} dt$$

Ricordando che la trasformata dell'esponenziale immaginaria è l'impulsiva traslata, si ottiene infine:

$$\sum_{n=-\infty}^{+\infty} c_n e^{2\pi i \frac{n}{T} t} \xleftrightarrow{\mathcal{F}} \sum_{n=-\infty}^{+\infty} c_n \delta\left(\nu - \frac{n}{T}\right)$$

Lo spettro di un segnale periodico risulta essere una sovrapposizione di funzioni impulsive traslate alle varie armoniche del segnale. Il peso di ogni impulso è il corrispondente coefficiente della serie di Fourier.

CAPITOLO 3

Filtri Analogici e Modulazione di Ampiezza

In questo capitolo viene esemplificata la potenza dell'analisi in rappresentazione di frequenza su due solidi esempi: i filtri analogici e la modulazione di ampiezza.

Viene ripreso lo studio dei sistemi lineari tempo-invarianti a tempo continuo. In questi sistemi l'uscita è ottenuta dalla convoluzione dell'ingresso con la risposta del sistema all'impulso. Prendendo in considerazione la rappresentazione in frequenza, si osserva che lo spettro dell'uscita è il prodotto dello spettro dell'ingresso per lo spettro della risposta all'impulso, detto *funzione di trasferimento*. I sistemi lineari tempo-invarianti esibiscono quindi capacità di filtrare componenti in frequenza, e per questo motivo sono anche chiamati *filtri lineari*.

Viene introdotta l'importante classe dei *filtri ideali* (passa basso/alto/banda), come filtri caratterizzati da funzioni di trasferimento a modulo costante in banda passante, nullo in banda proibita e aventi fase lineare. Si osserva che tali filtri non sono causali, quindi possono essere soltanto approssimati da filtri fisicamente realizzabili; si accenna

alle principali caratteristiche dei filtri realizzabili, introducendo in particolare i *filtri di Butterworth*.

Viene motivata la necessità di modulare i segnali per la loro trasmissione, ed in particolare si mostra come la *modulazione di ampiezza* sia un sistema lineare in grado di rilocalizzare il segnale su una diversa banda di frequenza, senza perdita di informazione. Si discute la realizzazione del sistema inverso, il *demodulatore* e si accenna ad un diverso tipo di modulazione, la *modulazione di ampiezza convenzionale*, utilizzata nelle trasmissioni radio in AM.

3.1 Risposta in Frequenza dei Sistemi Lineari Tempo-Invarianti

I sistemi lineari tempo invarianti (LTI), studiati in Capitolo 1, sono sistemi le cui caratteristiche non cambiano nel tempo e in cui vale il principio di sovrapposizione: la risposta a una combinazione lineare di segnali è la combinazione lineare delle risposte ai singoli segnali. Come abbiamo visto, il comportamento di tali sistemi è completamente individuato dalla risposta del sistema alla funzione impulsiva $\delta(t)$. Infatti se S è un sistema lineare tempo-invariante e $h(t)$ è la risposta $h(t) = S(\delta(t))$ del sistema all'impulso $\delta(t)$, allora la risposta $g(t) = S(f(t))$ del sistema ad un qualsiasi ingresso $f(t)$ è data da:

$$g(t) = \int_{-\infty}^{\infty} f(x)h(t-x)dx.$$

In altri termini, l'uscita $g(t)$ del sistema S è la convoluzione dell'ingresso $f(t)$ con la risposta all'impulso $h(t)$.

Studiamo ora come può essere descritto il rapporto ingresso-uscita del sistema se utilizziamo per i segnali la rappresentazione in frequenza. La risposta all'impulso $h(t)$ sarà rappresentata attraverso la sua trasformata di Fourier $H(\nu)$, che chiameremo *funzione di trasferimento* del sistema: la funzione di trasferimento di un sistema lineare tempo invariante S è lo spettro della sua risposte all'impulso.

Ci chiediamo ora quale è lo spettro $G(\nu)$ dell'uscita $g(t)$, conoscendo lo spettro $F(\nu)$ dell'ingresso $f(t)$ e la funzione di trasferimento $H(\nu)$ del sistema. La semplice risposta è data dal seguente:

Fatto 3.1 *Lo spettro della risposta $G(\nu)$ di un sistema LTI con funzione di trasferimento $H(\nu)$ è il prodotto dello spettro $F(\nu)$ dell'ingresso per la funzione di trasferimento $H(\nu)$, cioè $G(\nu) = H(\nu)F(\nu)$.*

Per provare questo fatto, ricordiamo che $F(\nu)$, $H(\nu)$ e $G(\nu)$ sono le trasformate di Fourier rispettivamente di $f(t)$, $h(t)$ e $g(t)$. Poichè $g(t)$ è la convoluzione di $h(t)$ e $f(t)$, dalla proprietà di convoluzione riportata in Tabella 2.1 si ottiene $G(\nu) = H(\nu)F(\nu)$.

Consideriamo ora un sistema LTI S che ha una funzione di trasferimento $H(\nu)$, e supponiamo che per una frequenza ν_1 sia $H(\nu_1) = 1$ mentre per una frequenza ν_2 sia $H(\nu_2) = 0$. Sia $G(\nu)$ lo spettro della risposta di S su un ingresso il cui spettro è $F(\nu)$. Poichè $G(\nu) = H(\nu)F(\nu)$, risulterà $G(\nu_1) = 1 \cdot F(\nu_1) = F(\nu_1)$, mentre $G(\nu_2) = 0 \cdot F(\nu_2) = 0$: a livello di frequenza ν_1 il sistema dà in uscita lo stesso peso $F(\nu_1)$ dell'ingresso, a livello di frequenza ν_2 il sistema dà in uscita 0, cancellando il peso $F(\nu_1)$ del segnale di ingresso.

I sistemi lineari tempo-invarianti sono quindi in grado di operare una certa selezione sulle frequenze, ampliando o attenuando in uscita le componenti armoniche dell'ingresso. Per questa attitudine a filtrare componenti in frequenza, i sistemi LTI sono anche detti *filtri lineari*.

Esempio 3.1.1

Osserviamo $|H(\nu)|^2$ è il guadagno della potenza dell'uscita sull'ingresso, a livello di frequenza ν . Infatti, da $G(\nu) = H(\nu)F(\nu)$ possiamo derivare:

$$|H(\nu)|^2 = \frac{|G(\nu)|^2}{|F(\nu)|^2} = \frac{|G(\nu)|^2 d\nu}{|F(\nu)|^2 d\nu}.$$

Ricordiamo che $|G(\nu)|^2 d\nu$ e $|F(\nu)|^2 d\nu$ rappresentano la potenza (o l'energia) delle componenti a frequenza tra ν e $\nu + d\nu$ rispettivamente nel segnale di uscita e quello di ingresso; $|H(\nu)|^2$ individua allora il guadagno, cioè il rapporto tra la potenza (o l'energia) del segnale in uscita e quello in ingresso alle varie frequenze.

3.1.1 Filtri Ideali

Un sistema senza distorsione è un sistema che riproduce in uscita la stessa forma del segnale d'ingresso, a meno di un eventuale fattore amplificativo e di un eventuale ritardo temporale. Un tale sistema può essere quindi descritto dalla trasformazione:

$$g(t) = Af(t - t_0).$$

Passando alle trasformate di Fourier e applicando la proprietà di traslazione temporale, si ha:

$$G(\nu) = Ae^{-2\pi i \nu t_0} F(\nu).$$

La funzione di trasferimento $H(\nu)$ del sistema è quindi data da:

$$H(\nu) = \frac{G(\nu)}{F(\nu)} = Ae^{-2\pi i \nu t_0}.$$

Si noti che il modulo della funzione di trasferimento è costante, vale cioè che $|H(\nu)| = A$, mentre la fase è lineare nella frequenza ν , cioè $\text{ph } H(\nu) = -2\pi \nu t_0$.

Un sistema che annulla le componenti armoniche in determinati intervalli di frequenza e che si comporta come un sistema senza distorsione sulle frequenze rimanenti è detto *filtro*

ideale. Un esempio di filtro ideale è il filtro *passa-basso*, che passa (riproduce in uscita con guadagno A e fase lineare) le componenti con frequenza non superiore a una certa frequenza ν_c detta *frequenza di taglio*, ed elimina quelle con frequenza superiore a questa soglia. Tale filtro ha quindi la seguente funzione di trasferimento:

$$H(\nu) = Ae^{-2\pi i \nu t_0} \text{rect}\left(\frac{\nu}{\nu_c}\right)$$

il cui modulo $\text{rect}(\frac{\nu}{\nu_c})$ e la cui fase $-2\pi i t_0 \nu$ sono mostrati in Figura 3.1.

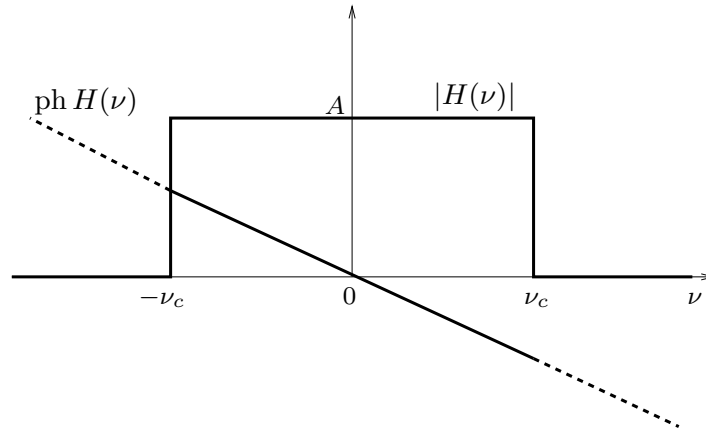


Figura 3.1 Risposta in frequenza di un filtro passa-basso ideale.

Un filtro *passa-alto*, viceversa, elimina le componenti in frequenza basse e passa quelle alte (superiori a ν_c); il filtro *passa-banda* infine, passa una *banda* o *intervallo* di componenti in frequenza ($\nu_a < \nu < \nu_b$) ed elimina quelle inferiori o superiori ad essa. Le bande evidenziate in questi esempi sono quindi due: la banda che interessa preservare effettivamente (*banda passante*) e la banda nella quale si richiede l'eliminazione (*banda proibita*). Modulo e fase di filtri passa-alto e passa-banda sono mostrati in Figura 3.2.

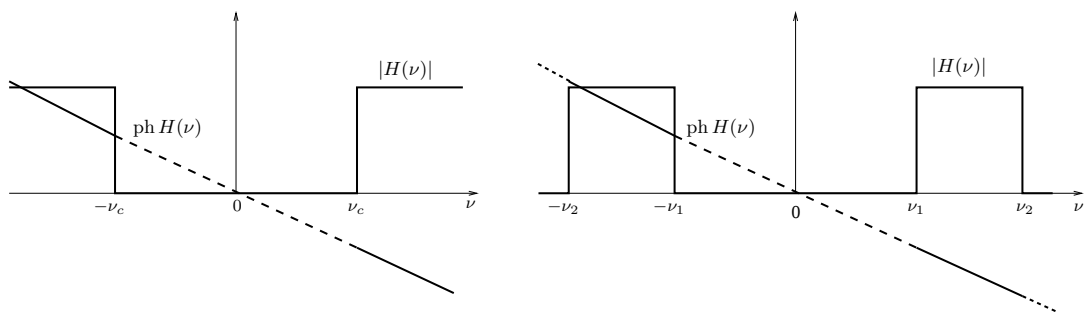


Figura 3.2 Risposta in frequenza di filtri passa-alto e passa-banda.

Esempio 3.1.2

Supponiamo che il segnale $f(t)$ venga corrotto da un segnale-disturbo $n(t)$, trasformandosi nel segnale $g(t) = f(t) + n(t)$. Siamo interessati a ricostruire il segnale $f(t)$ a partire dal segnale corrotto $g(t)$. Questo non è in generale possibile, poichè la conoscenza della somma $f(t) + n(t)$ non permette di individuare univocamente il primo addendo. La ricostruzione è tuttavia possibile quando i supporti degli spettri $F(\nu)$ e $N(\nu)$ dei segnali $f(t)$ e $n(t)$ sono disgiunti.

Per esempio, consideriamo il caso in cui $F(\nu) = 0$ per $|\nu| > \nu_1$ e $N(\nu) = 0$ per $|\nu| < \nu_2$, con $\nu_1 < \nu_2$. Applicando al segnale somma $g(t)$ un filtro ideale passa-basso con frequenza di taglio ν_1 , la cui funzione di trasferimento è quindi $H(\nu) = \text{rect}(\frac{\nu}{\nu_1})$, si ottiene in uscita il segnale $f(t)$. Detto infatti $M(\nu)$ lo spettro dell'uscita, vale:

$$M(\nu) = G(\nu)H(\nu) = F(\nu)\text{rect}(\frac{\nu}{\nu_1}) + N(\nu)\text{rect}(\frac{\nu}{\nu_1}) = F(\nu)$$

Osservando che $\text{rect}(\frac{\nu}{\nu_1}) = 0$ per $|\nu| > \nu_1$ mentre $F(\nu) = 0$ per $|\nu| < \nu_2$, risulta $N(\nu)\text{rect}(\frac{\nu}{\nu_1}) = 0$. Antitrasformando, concludiamo che il filtro, avendo in ingresso il segnale corrotto $f(t) + n(t)$, dà in uscita il segnale originale $f(t)$.

Esempio 3.1.3

Supponiamo che il segnale di una conversazione $f(t)$, che ha un limite di banda di 3400 Hz, venga campionato con una frequenza di campionamento di 8000 Hz, ottenendo il segnale campionato $f_c(t)$. Si vuole ricostruire il segnale $f(t)$ a partire da $f_c(t)$.

A questo riguardo, ricordiamo che lo spettro $F_c(\nu)$ è ottenibile, a meno di una costante moltiplicativa, dalla somma delle repliche dello spettro $F(\nu)$ del segnale $f(t)$, traslate di $8000 \cdot n$ Hz, per qualsiasi intero n . Poichè $8000 > 2 \cdot 3400$, tutte le repliche sono disgiunte; per ricostruire $F(\nu)$ basterà allora eliminare tutte le repliche, ad esclusione dello spettro del segnale, con un filtro passa-basso con frequenza di taglio di 4000 Hz.

3.2 Caratteristiche dei Filtri Analogici

Nella sezione precedente, abbiamo introdotto la nozione di filtro ideale e delineato le principali tipologie: filtro passa-basso, passa-alto e passa-banda. In questo paragrafo studiamo la realizzazione pratica di filtri ideali; faremo riferimento esclusivamente a filtri passa-basso, poichè le considerazioni riportate di seguito possono essere estese senza difficoltà agli altri casi.

La funzione di trasferimento tipica di un filtro ideale passa-basso è $\text{rect}(\frac{\nu}{2\nu_c})$. Poichè la funzione di trasferimento di un filtro è la trasformata di Fourier della sua risposta all'impulso, la risposta all'impulso $h(t)$ di un filtro passa-basso ideale è l'antitrasformata di Fourier della funzione $\text{rect}(\frac{\nu}{2\nu_c})$. Sarà allora:

$$h(t) = \mathcal{F}^{-1} \left\{ \text{rect}(\frac{\nu}{2\nu_c}) \right\} = 2\nu_c \text{sinc}(2\nu_c t).$$

Si osserva che $h(t)$ è in generale diversa da zero per $t < 0$ e quindi il filtro ideale è un sistema lineare tempo-invariante ma non causale. Questo significa che tale filtro non può essere realizzato da un sistema fisico che verifichi il principio di causalità. Ricordiamo ora che la funzione di trasferimento H di un filtro ideale passa-basso possiede le seguenti caratteristiche:

1. $|H|$ è costante nella banda passante ed è identicamente nulla nella banda proibita;
2. la banda passante e la banda proibita sono confinanti (separate dalla frequenza di taglio);
3. la fase $\text{ph } H$ è lineare; questo significa che le varie componenti armoniche nella banda passante hanno tutte lo stesso ritardo temporale.

Come conseguenza del fatto che il filtro ideale non è realizzabile, possiamo osservare che ogni sistema realizzabile non potrà mai verificare contemporaneamente le caratteristiche 1., 2. e 3.; un filtro ideale potrà solo essere approssimato. Introduciamo qui alcuni parametri che descrivono il grado di approssimazione di un filtro ideale mediante un filtro fisicamente realizzabile.

Indicando con $H(\nu)$ la funzione di trasferimento di un eventuale filtro passa-basso realizzabile, sappiamo che $H(\nu)$ è completamente specificata dal suo modulo $|H(\nu)|$ e dalla sua fase $\text{ph } H(\nu)$. La Figura 3.3 mostra la tipica forma di $|H(\nu)|$ e $\text{ph } H(\nu)$ per un filtro passa-basso realizzabile.

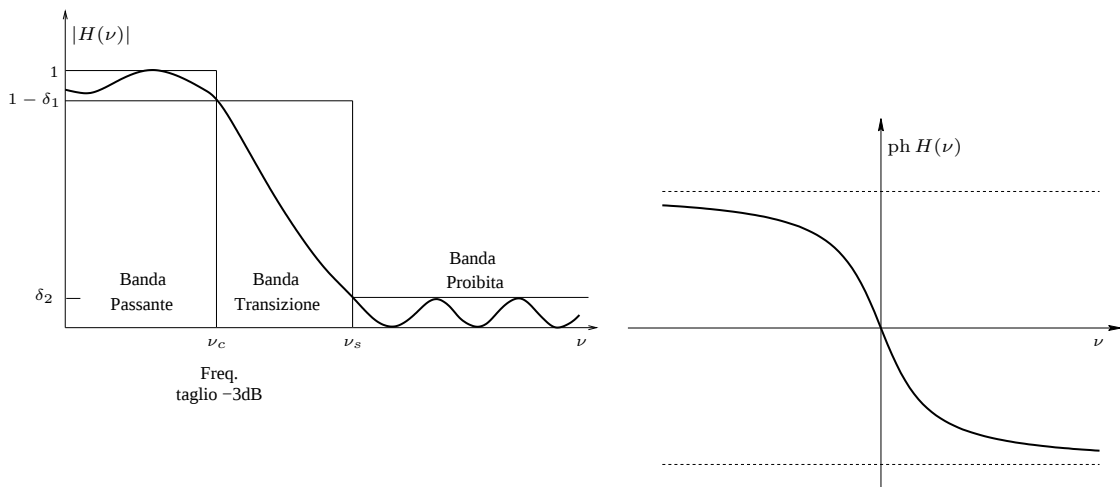


Figura 3.3 Modulo e fase di un filtro passa-basso realizzabile.

Rispetto a un filtro ideale possiamo rilevare le seguenti differenze:

1. $|H(\nu)|$ non è costante nella banda passante e non è identicamente nullo nella banda proibita; si possono rilevare inoltre oscillazioni (ripple) di ampiezza non trascurabile

sia nella banda passante che in quella proibita. Un parametro importante è l'ampiezza della massima oscillazione in banda proibita δ_2 , il cui logaritmo $-20 \log_{10} \delta_2$ dB è detto *attenuazione*.

2. la banda passante e la banda proibita non confinano, ma sono separate da una banda detta *banda di transizione*. Tale banda inizia dalla frequenza di taglio ν_c e termina alla frequenza di stop ν_s ; la dimensione della banda di transizione è data da $\nu_s - \nu_c$.
3. la fase $\text{ph } H(\nu)$ non risulta essere lineare.

Analizziamo ora separatamente il significato fisico del modulo e della fase della funzione di trasferimento di un filtro realizzabile. Gli elementi principali che specificano un filtro (passa-basso) sono:

1. la frequenza di taglio ν_c che identifica la fine della banda passante, e la frequenza di stop ν_s che identifica l'inizio della banda proibita;
2. le dimensioni massime δ_1 e δ_2 permesse alle oscillazioni rispettivamente in banda passante e in banda proibita.

Con frequenza di taglio ν_c s'intende usualmente la frequenza per la quale $|H(\nu)| \approx 1$ in banda passante, ν_c è la frequenza per cui

$$H^2(\nu_c) = \frac{1}{2}.$$

La frequenza ν_c è detta anche *frequenza di taglio a 3 dB* poiché $10 \log(\frac{1}{2}) \approx -3$ dB.

Esempio 3.2.1

Determinare la frequenza di taglio a 3 dB del filtro passa-basso con guadagno $|H(\nu)|^2 = \frac{1}{(1+\nu^2/100)}$.

Il guadagno in banda passante è 1 e la frequenza di taglio ν_c è quella per cui $\frac{1}{(1+\nu^2/100)} = 1/2$. Risolvendo l'equazione si ottiene $\nu_c = 10$ Hz.

Usualmente la massima oscillazione δ_2 permessa in banda proibita viene espressa in dB attraverso l'attenuazione R_p , dove

$$R_p = -10 \log(\delta_2^2).$$

La banda proibita è data dall'insieme delle frequenze per le quali il guadagno è inferiore a una opportuna soglia di attenuazione che normalmente viene stabilita in funzione della particolare applicazione. Se indichiamo con ν_s , come in Figura 3.3, la frequenza di stop, cioè la frequenza di inizio della banda proibita, le frequenze tra ν_c ed ν_s costituiscono la banda di transizione.

Esempio 3.2.2

Si consideri il filtro con funzione di trasferimento $H(\nu)$ tale che $|H(\nu)|^2 = \frac{1}{1+(\nu/100)^8}$. Determinare l'ampiezza della banda di transizione sapendo che la frequenza di stop corrisponde ad un'attenuazione di 40 dB.

La frequenza di taglio a 3 dB è di 100 Hz. La frequenza di stop ν_s è tale che $40 = -20 \log_{10} |H(\nu_s)|$. Risolvendo tale equazione si ottiene $\nu_s \approx 316$ Hz.

La dimensione della banda di transizione risulta $316 - 100 = 216$ Hz.

Discutiamo ora l'effetto prodotto dalla non linearità della fase della funzione di trasferimento del filtro. A tal riguardo, consideriamo per semplicità un sistema che ammette una funzione di trasferimento con modulo $|H(\nu)| = G(\nu)$ e fase $\text{ph } H(\nu) = \phi(\nu)$, così che:

$$H(\nu) = G(\nu)e^{i\phi(\nu)}.$$

Il segnale (complesso) $e^{2\pi i \nu_1 t}$ di frequenza ν_1 viene trasformato dal sistema nel segnale $G(\nu_1)e^{i(2\pi \nu_1 t + \phi(\nu_1))}$. Se la fase è lineare, cioè $\phi(\nu) = -2\pi t_0 \nu$ per un'opportuna costante t_0 , il segnale di uscita è

$$G(\nu_1)e^{2\pi i \nu_1 (t-t_0)}.$$

Il segnale di uscita risulta allora essere lo stesso segnale di ingresso ritardato di un tempo t_0 , qualsiasi sia la frequenza ν_1 : una fase lineare comporta un ritardo temporale uguale per tutte le componenti armoniche.

Una fase non lineare crea invece ritardi differenti per le componenti a diversa frequenza, creando una distorsione complessiva del segnale. Per certe applicazioni (ad esempio nei modem) le distorsioni create dalla non linearità della fase devono essere il più possibile eliminate; in altre applicazioni la non linearità può essere utilizzata per dar luogo ad effetti speciali.

3.3 Famiglie di Filtri Causali

Abbiamo visto che un filtro ideale non è causale e quindi può essere soltanto approssimato con filtri realizzabili fisicamente. A questo riguardo abbiamo introdotto parametri che denotano la bontà nell'approssimarne il guadagno (dimensione della banda di transizione, attenuazione, oscillazioni) o la fase (linearità).

La progettazione di un filtro è fortemente dipendente dall'applicazione; in certi casi (per esempio nei sistemi audio) è richiesta un'ottima risposta in fase. In altre applicazioni la linearità della fase è di scarso rilievo, mentre è critica l'accuratezza nell'approssimare il guadagno, e così via.

In aiuto al progettista, sono state introdotte e analizzate varie classi di filtri usualmente disponibili in sistemi di calcolo automatico per la progettazione, l'implementazione e la simulazione di filtri, quali ad esempio SCILAB. Le principali famiglie sono quella dei filtri

di Butterworth, di Chebyshev, di Cauer (o ellittici) e di Bessel. In Tabella 3.1 è mostrata una grossolana valutazione comparativa della bontà di questi filtri (a parità di ordine); analizzeremo poi più in dettaglio la classe dei filtri di Butterworth.

Tabella 3.1 Caratteristiche qualitative di famiglie di filtri.

Filtro	Accuratezza appross. guadagno	Linearità fase
Butterworth	media	media
Chebyshev	buona	cattiva
Ellittico	ottima	pessima
Bessel	cattiva	buona

3.3.1 Filtri di Butterworth

I filtri di Butterworth costituiscono una famiglia di filtri che soddisfa bene i requisiti sul guadagno in banda passante e meno bene in banda di transizione. Sebbene non esibiscano una fase lineare in banda passante, l'approssimazione non è troppo cattiva.

Un filtro di Butterworth è caratterizzato da 2 parametri: l'ordine N e la frequenza di taglio ν_c . La forma generale del modulo della funzione di trasferimento di un filtro di Butterworth di ordine N e frequenza di taglio ν_c è:

$$\frac{1}{|B_N(i\frac{\nu}{\nu_c})|} = \frac{1}{\sqrt{1 + \left(\frac{\nu}{\nu_c}\right)^{2N}}},$$

dove $B_N(s)$ è un opportuno polinomio detto N -esimo polinomio di Butterworth.

La risposta in frequenza di alcuni filtri di Butterworth è riportata in Figura 3.4.

Si osservi:

- La frequenza di taglio a 3 dB è ν_c , indipendentemente dall'ordine N del filtro.
- L'attenuazione nella banda proibita dipende da N in modo critico; risulta infatti $R_p \approx 20N \log(\frac{\nu_s}{\nu_c})$.
- Non sono presenti oscillazioni nè in banda passante nè in banda proibita: il filtro di Butterworth è quello che ha la caratteristica di essere piatto in banda passante.

In una tipica situazione di progetto, il parametro ν_c risulta essere la frequenza di taglio desiderata e l'ordine N viene scelto in modo tale da soddisfare la richiesta di attenuazione in banda proibita.

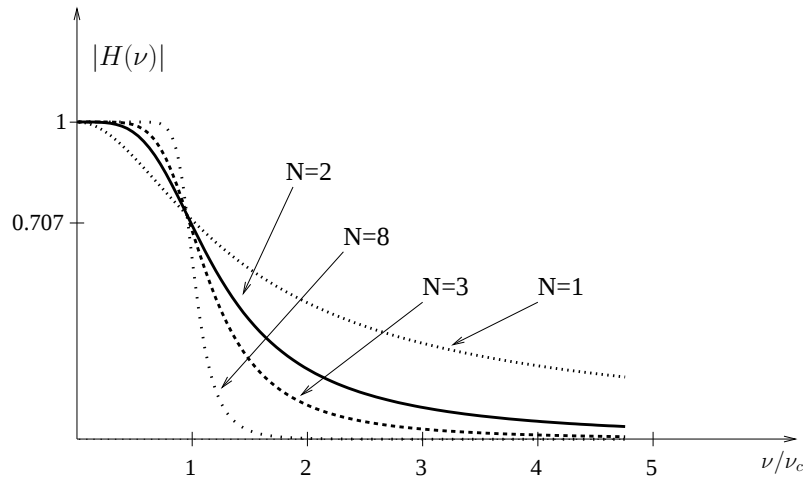


Figura 3.4 Risposta in frequenza di filtri di Butterworth.

Esempio 3.3.1

Determinare l'ordine di un filtro di Butterworth con frequenza di taglio di 100 Hz e frequenza di stop di 150 Hz con attenuazione a 40 dB.

Il guadagno di un filtro di Butterworth di ordine N , con frequenza di taglio pari a 100 Hz è:

$$|H(\nu)|^2 = \frac{1}{1 + \left(\frac{\nu}{100}\right)^{2N}}.$$

La frequenza di stop è per ipotesi la frequenza ν_s che produce un'attenuazione di 40 dB, cioè:

$$40 = 20N \log \frac{150}{100}.$$

Da questa equazione si ricava che $N \approx 11.36$. Poiché l'ordine del filtro deve essere un intero, concludiamo che il filtro di ordine 12 soddisfa la specifica.

3.3.2 Realizzazione di Filtri Analogici

Dato un filtro, descritto ad esempio dalla sua funzione di trasferimento $H(\nu)$, la sua realizzazione consiste nel progettare un circuito elettrico che, visto come sistema, ha $H(\nu)$ come funzione di trasferimento. Questo circuito, dopo un accurato esame di sue eventuali imperfezioni come capacità parassite o altro, potrà essere implementato ottenendo la realizzazione fisica del filtro.

Le realizzazioni di filtri analogici possono essere classificate sulla base delle componenti costituenti; elenchiamo qui brevemente le principali classi con alcune caratteristiche.

Filtri RLC passivi. Essi consistono di componenti passive come induttanze, condensatori e resistenze. Per il loro funzionamento non richiedono potenza aggiuntiva, ma spesso esibiscono perdite significative in banda passante; la presenza di induttanze, inoltre, pone seri problemi alla loro miniaturizzazione. Lavorano bene fino a 500 Hz.

Filtri RC attivi. Essi non contengono induttanze, ma resistenze, condensatori e amplificatori operazionali. Si prestano bene alla miniaturizzazione, ma richiedono potenza aggiuntiva. Lavorano bene da 1Hz fino a 500 Hz.

Filtri a microonde. Essi consistono di componenti passive come linee di trasmissione, linee di trasmissione accoppiate e cavità risonanti. Non richiedono potenza aggiuntiva e sono utilizzate per filtri che lavorano sopra i 300 MHz.

Filtri a cristallo. Sono costituiti da risuonatori piezoelettrici ed lavorano dai 10 KHz ai 200 MHz. Con risuonatori di quarzo si possono realizzare filtri passa-banda con ampiezza di banda veramente ristretta.

Filtri elettromeccanici. Essi sono costituiti da risuonatori meccanici: per prima cosa il segnale elettrico viene trasformato in vibrazioni meccaniche, viene poi applicato il filtro e infine il segnale è riconvertito in segnale elettrico. Questi filtri lavorano fino a 200 MHz.

3.4 Modulazione e Demodulazione di Ampiezza (AM)

Per trasmettere un segnale, lo si può trasformare in un segnale elettrico; è poi opportuno rilocalizzare tale segnale, che occupa una certa banda di frequenza, in un'altra banda al fine di poterlo trasmettere. La modulazione è un sistema che permette di rilocalizzare un segnale in una diversa banda di frequenza, mantenendo l'informazione codificata. Si applica la modulazione per tre ragioni principali:

1. La presenza di più segnali sulla stessa banda di frequenza può creare problemi di sovrapposizione. Per esempio, i segnali del parlato variano su un range di frequenze da 0 a 4000 Hz, la musica da 0 a 20 kHz, i segnali video originali da 0 a 5 MHz. Se più segnali vengono trasmessi contemporaneamente sulla stessa frequenza, in ricezione si hanno pesanti interferenze: è allora necessario differenziarli rilocalizzandoli su bande di frequenza diverse, in modo da non creare interferenze.
2. Ci sono numerosi disturbi alle basse frequenze (luce elettrica, motori elettrici), che possono essere superati utilizzando frequenze molto più alte.
3. In un'onda elettromagnetica, la lunghezza d'onda λ è inversamente proporzionale alla frequenza: $\lambda = \frac{c}{\nu}$, dove $c = 3 \cdot 10^8 m/sec$ è la velocità della luce, mentre ν è la frequenza. Poiché la lunghezza dell'antenna ricevente deve essere circa la metà della lunghezza d'onda, trasmissioni a bassa frequenza richiederebbero antenne eccessivamente lunghe.

Esempio 3.4.1

Determinare la lunghezza di un'antenna per la ricezione di segnali in banda 7. Ricordiamo che la banda 7 è data dalle frequenze comprese tra 0.3×10^7 Hz e 3×10^7 Hz; una tipica frequenza in tale banda è 10^7 Hz. Per tale frequenza la lunghezza L dell'antenna deve essere la metà della lunghezza d'onda λ , data da $\lambda = \frac{c}{\nu} = \frac{3 \cdot 10^8}{10^7} = 30m$. Di conseguenza $L = \frac{\lambda}{2} = 15m$.

Esistono vari sistemi che permettono di rilocalizzare un segnale su una diversa banda di frequenza, tra i quali ricordiamo la modulazione di ampiezza, frequenza o fase. In questa sezione studiamo il caso della modulazione di ampiezza e della corrispondente operazione inversa di demodulazione.

La *modulazione di ampiezza* (AM) è ottenuta moltiplicando un segnale $f(t)$ per una sinusoidale di data frequenza ν_0 . Essa è descritta dal sistema MA dove:

$$MA(f(t)) = \cos 2\pi\nu_0 t \cdot f(t)$$

La modulazione di ampiezza è anche detta *modulazione lineare*, poichè il sistema MA è lineare. Tale sistema non è tuttavia tempo invariante.

Analizziamo ora l'effetto della modulazione di ampiezza sugli spettri: dato $g(t) = MA(f(t))$, si vuol determinare la relazione tra gli spettri $G(\nu)$ e $F(\nu)$ di $g(t)$ e $f(t)$. Ricordando che $\cos 2\pi\nu_0 t = \frac{1}{2}(e^{2\pi i\nu_0 t} + e^{-2\pi i\nu_0 t})$, il problema è ridotto al calcolo della trasformata di Fourier di $\frac{1}{2}e^{2\pi i\nu_0 t} \cdot f(t) + \frac{1}{2}e^{-2\pi i\nu_0 t} \cdot f(t)$. Lo spettro di $G(\nu)$ è ottenuto traslando $F(\nu)$ di $+\nu_0$ e di $-\nu_0$ (vedi anche proprietà di modulazione in paragrafo 2.7.3). Si ottiene:

$$G(\nu) = \frac{1}{2} (F(\nu + \nu_0) + (F(\nu - \nu_0))).$$

Se succede che i supporti di $F(\nu + \nu_0)$ e $F(\nu - \nu_0)$ sono disgiunti, l'effetto della modulazione è quello di rilocalizzare il segnale $f(t)$ su una diversa banda di frequenze, preservandone l'informazione. Supponiamo che $F(\nu)$ sia a banda limitata da B , cioè che $F(\nu)$ sia nulla per $|\nu| \geq B$ come mostrato nel grafico di Figura 3.5.

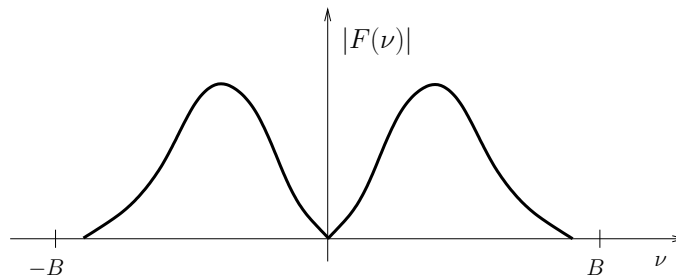


Figura 3.5 Modulo della trasformata di un segnale reale f .

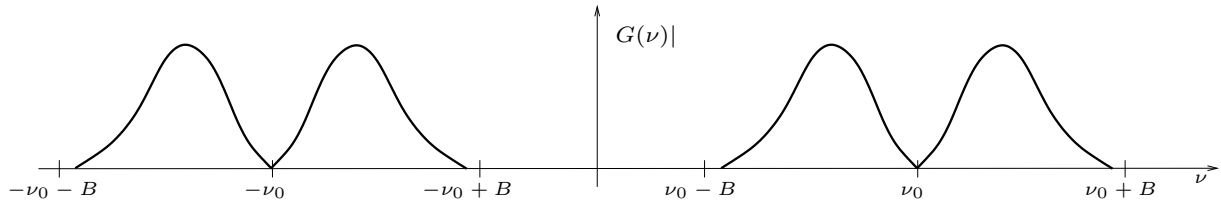


Figura 3.6 Modulo della trasformata del segnale modulato g .

Come mostra il grafico di Figura 3.6, i supporti di $F(\nu + \nu_0)$ e $F(\nu - \nu_0)$ sono disgiunti se $\nu_0 > B$.

Se i supporti sono disgiunti, la conoscenza di $G(\nu)$ permette di ricostruire $F(\nu)$, e quindi $f(t)$. Possiamo allora concludere:

Fatto 3.2 *Un segnale $f(t)$ a banda limitata da B può essere completamente ricostruito dal segnale modulato $g(t) = \cos 2\pi\nu_0 t \cdot f(t)$, se $\nu_0 > B$.*

Abbiamo visto sotto quali condizioni un segnale $f(t)$ possa essere ricostruito a partire dal segnale modulato $g(t)$. Descriviamo ora come si può realizzare effettivamente un sistema per la demodulazione, in grado di realizzare la ricostruzione.

Per prima cosa si applica al segnale $g(t)$ una nuova modulazione con la stessa frequenza ν_0 , ottenendo $z(t) = MA(g(t))$. La relazione tra gli spettri del segnale $g(t)$ e del segnale modulato $z(t)$ è data da $Z(\nu) = \frac{1}{2}(G(\nu + \nu_0) + G(\nu - \nu_0))$. Ricordando che $G(\nu) = \frac{1}{2}(F(\nu + \nu_0) + F(\nu - \nu_0))$, si conclude:

$$Z(\nu) = \frac{1}{2}F(\nu) + \frac{1}{4}(F(\nu - 2\nu_0) + F(\nu + 2\nu_0))$$

Osserviamo che se $f(t)$ è a banda limitata da B e se $\nu_0 > B$, i supporti di $F(\nu)$, $F(\nu - 2\nu_0)$ e $+F(\nu + 2\nu_0)$ sono disgiunti: applicando un filtro passa-basso ideale con frequenza di taglio B si cancellano le componenti $F(\nu - 2\nu_0)$ e $+F(\nu + 2\nu_0)$ ottenendo $\frac{1}{2}F(\nu)$.

Il sistema complessivo di modulazione, trasmissione e demodulazione è rappresentato nella Figura 3.7.

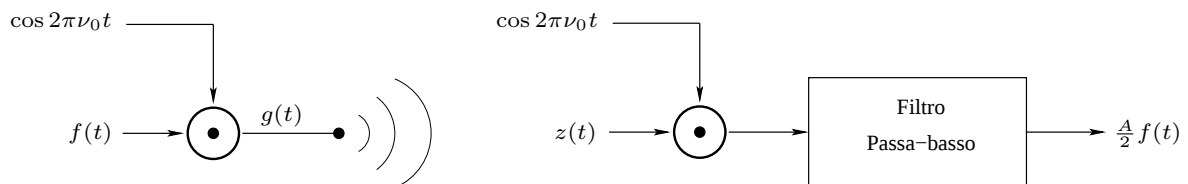


Figura 3.7 Modulazione AM, trasmissione e demodulazione AM.

Ricordiamo che un segnale a valori reali $f(t)$ ha uno spettro $F(\nu)$ tale che $|F(\nu)|$ è una funzione pari mentre $ph(F(\nu))$ è una funzione dispari: la conoscenza dello spettro sulle frequenze positive permette la ricostruzione di tutto lo spettro. Se in particolare $F(\nu)$ è a banda limitata da B , $F(\nu)$ può essere ricostruito a partire dai valori di $F(\nu)$ con $0 \leq \nu \leq B$.

Il segnale modulato $g(t)$ è quindi ridondante ai fini della ricostruzione di $f(t)$, poichè lo spettro $F(\nu)$ può essere ricostruito sulla base della conoscenza di $G(\nu)$ nell'intervallo (side) $\nu_0 \leq \nu \leq \nu_0 + B$ anzichè sull'intero intervallo (double-side) $\nu_0 - B \leq \nu \leq \nu_0 + B$.

Per questo motivo la modulazione di ampiezza prima introdotta è chiamata AMDSB-SC (Amplitude Modulation Double-Sideband Suppressed Carrier); essa comporta una richiesta di banda per la trasmissione di dimensione almeno $2B$.

Esempio 3.4.2

Partendo dall'osservazione che è possibile ricostruire $F(\nu)$, e quindi $f(t)$, sulla base della conoscenza di $G(\nu)$ nel solo intervallo $\nu_0 \leq \nu \leq \nu_0 + B$, l'operazione di modulazione può essere perfezionata per poter risparmiare banda per la trasmissione (riducendola del 50%). Uno dei metodi adottati consiste nell'applicare un filtro passa-banda per eliminare le frequenze indesiderate nell'intervallo $\nu_0 - B \leq \nu \leq \nu_0$.

Le stazioni per trasmissione delle radio commerciali in AM utilizzano una modulazione leggermente diversa da quella definita precedentemente. Il sistema, detto AM convenzionale, è descritto da:

$$g(t) = A(1 + af(t)) \cos 2\pi\nu_0 t.$$

dove $f(t)$ è un segnale normalizzato, in modo che $|f(t)| \leq 1$, ed il parametro a è un numero reale compreso tra 0 ed 1, detto indice di modulazione. Mentre $f(t)$ oscilla tra valori positivi e negativi, il termine $1 + af(t)$ risulta sempre positivo: il segnale $g(t)$ può quindi essere demodulato con un semplice *rivelatore di inviluppo*. Il rivelatore di inviluppo è basato sul fatto che, per funzioni $s(t)$ a valori positivi, l'inviluppo di $s(t) \cos 2\pi\nu_0 t$ coincide proprio con $s(t)$, come si rileva in Figura 3.8.

Un semplice rivelatore di inviluppo è dato dal circuito mostrato in Figura 3.9. Esso è composto da un sistema che, avendo in ingresso il segnale $m(t)$, dà in uscita il segnale $m(t)u(m(t))$ in cui le componenti negative del segnale sono eliminate, e da un filtro passa-basso, ottenuto con un circuito RC, che elimina le componenti in alta frequenza.

La modulazione AM convenzionale richiede un apparato di trasmissione più complesso rispetto alla modulazione precedentemente presentata, mentre il ricevitore viene semplificato. Il sistema risultante è quindi particolarmente indicato per le trasmissioni broadcast, in cui il traffico va da un trasmettitore a molti ricevitori. Tale modulazione viene effettivamente utilizzata nelle trasmissioni radio in AM. La banda disponibile per tali trasmissioni varia da 535 a 1605 kHz; la banda viene divisa in identici blocchi di 10 kHz e ad ogni stazione viene assegnato un blocco: la stazione può allora trasmettere senza interferire messaggi con limite di banda di 5 kHz.

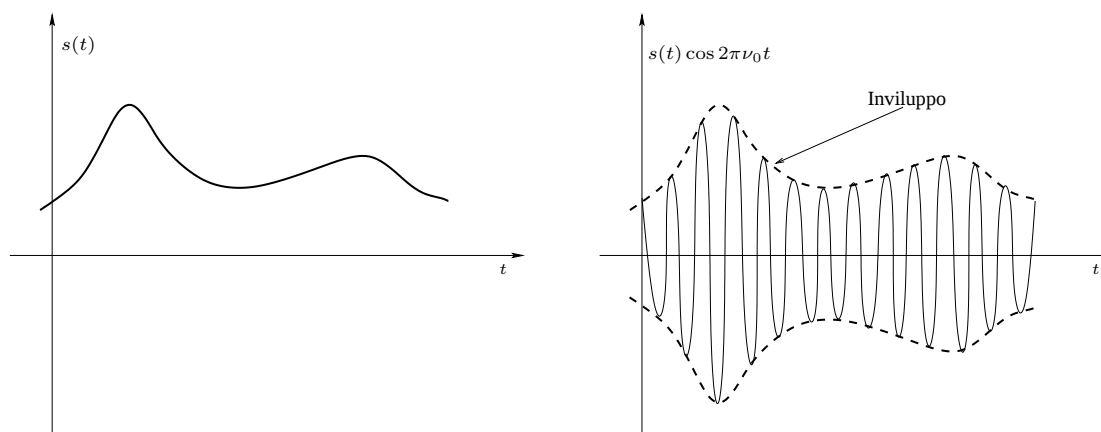


Figura 3.8 Inviluppo del segnale $s(t) \cos 2\pi\nu_0 t$, con $s(t) \geq 0$.

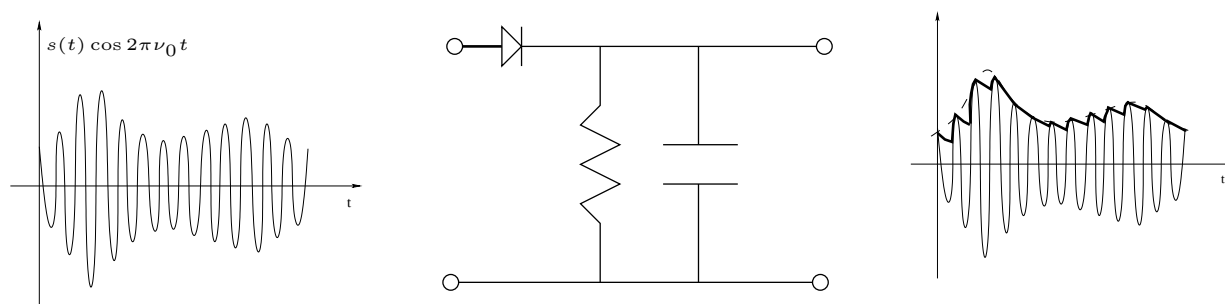


Figura 3.9 Rivelatore di inviluppo.

Esempio 3.4.3

Per trasmettere due segnali $f_1(t)$ e $f_2(t)$ a banda limitata da B sullo stesso canale, possiamo procedere come segue:

1. si modulano $f_1(t)$ e $f_2(t)$ con due diverse frequenze ν_1 e ν_2 ottenendo $g_1(t) = A \cos 2\pi\nu_1 t \cdot f_1(t)$ e $g_2(t) = A \cos 2\pi\nu_2 t \cdot f_2(t)$;
2. si sommano i due segnali $g_1(t)$ e $g_2(t)$ trasmettendo il segnale $g(t)$ risultante. Se $\nu_1, \nu_2 > B$ e $|\nu_1 - \nu_2| > 2B$, è facile osservare che i supporti degli spettri $G_1(\nu)$ e $G_2(\nu)$ di $g_1(t)$ e $g_2(t)$ risultano disgiunti, e quindi da $G(\nu) = G_1(\nu) + G_2(\nu)$ è possibile ricostruire $G_1(\nu)$ e $G_2(\nu)$ con l'applicazione di opportuni filtri; un esempio è mostrato in Figura 3.10.

Il metodo precedente può essere facilmente esteso alla trasmissione sullo stesso canale di N segnali a banda limitata da B : in questo caso la larghezza della banda utilizzata dovrà essere almeno $2NB$.

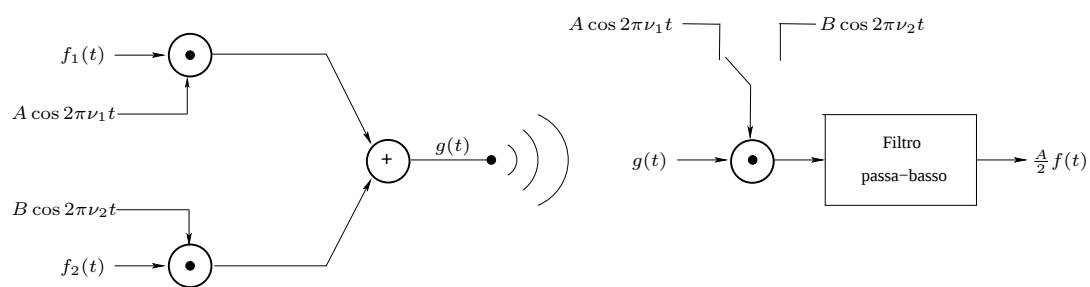


Figura 3.10 Trasmissione simultanea di 2 segnali modulati (AM) e ricezione.

CAPITOLO 4

Conversione Analogico-Digitale

I segnali del mondo reale sono analogici, mentre un elaboratore digitale è in grado di memorizzare e trattare esclusivamente sequenze finite di bit. Per trattare con tecniche digitali i segnali analogici è allora necessario, in via preliminare, approssimare quest'ultimi con segnali digitali. I sistemi che trasformano un segnale analogico nel corrispondente digitale sono detti *convertitori analogico-digitali* (ADC), mentre quelli che realizzano l'operazione inversa di trasformare un segnale digitale in un segnale analogico sono detti *convertitori digitali-analogici* (DAC).

I principi di base per la conversione sono quelli relativi alle operazioni di *campionamento* e di *quantizzazione*.

Il primo paragrafo è dedicato allo studio dell'operazione di campionamento. Campionare un segnale a tempo continuo significa rilevare le ampiezze del segnale su un insieme discreto di tempi: i frames catturati da una telecamera che inquadra una scena reale ne costituiscono un esempio. Viene qui discusso il fondamentale teorema del campionamento: un segnale a banda limitata da B Hz può essere perfettamente ricostruito dai suoi campioni presi con una frequenza di almeno $2B$ Hz. Viene poi descritto come lo spettro del segnale

campionato possa essere ottenuto da quello del segnale, e viene richiamato il fenomeno dell'*equivocazione* (*aliasing*), che si verifica quando la frequenza di campionamento è inferiore a $2B$.

La quantizzazione permette di trasformare un segnale a valori continui in un segnale a valori su un insieme finito. Questa operazione in generale introduce un errore irreversibile nel segnale quantizzato: dato il segnale quantizzato, non è in generale possibile ricostruire il segnale originale. È tuttavia possibile controllare la qualità con cui il segnale quantizzato approssima quello analogico: un tipico indice di qualità è il *rapporto segnale-rumore* SQNR. I sistemi che realizzano l'operazione di quantizzazione sono chiamati *quantizzatori*: nel secondo paragrafo viene introdotto e studiato il quantizzatore uniforme a m bit, mostrando in particolare che ogni bit aggiunto al quantizzatore comporta un miglioramento di 6 dB nel rapporto segnale-rumore. Successivamente si mostra come si possano realizzare semplici modelli di quantizzatori non uniformi, ed infine si discutono alcune problematiche sul convertitore analogico-digitale.

4.1 Campionamento

Campionare un segnale a tempo continuo significa rilevare le ampiezze del segnale su un insieme discreto di tempi. Ad esempio, fissato un intervallo di tempo di ampiezza τ , un campionamento uniforme con periodo τ di un segnale $f(t)$ corrisponde all'osservazione del segnale ai tempi $n\tau$ ($-\infty < n < \infty$); il segnale campionato può essere interpretato come il segnale a tempo discreto $f(n\tau)$.

Dato un segnale $f(t)$, è anche utile interpretare il segnale campionato con periodo τ come il segnale $f_s(t)$ ottenuto sommando gli impulsi traslati $\delta(t - n\tau)$ ($-\infty < n < \infty$) ognuno dei quali è pesato col valore $f(n\tau)$:

$$f_s(t) = \sum_{n=-\infty}^{+\infty} f(n\tau)\delta(t - n\tau)$$

Il sistema *campionatore uniforme* con periodo τ , o equivalentemente con frequenza di campionamento $A = 1/\tau$, trasforma un segnale a tempo continuo $f(t)$ nel segnale a tempo discreto $f(n\tau)$, come mostrato in Figura 4.1.

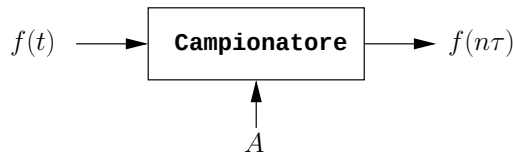


Figura 4.1 Campionatore.

Si osservi che il sistema campionatore $Camp$ è un sistema lineare:

$$Camp(a.f(t) + b.g(t)) = a.f(n\tau) + b.g(n\tau) = a.Camp(f(t)) + b.Camp(g(t))$$

Camp non è invece un sistema tempo invariante.

Dato un segnale $f(t)$ e un periodo di campionamento τ , è sempre univocamente definibile il segnale campionato $f(n\tau)$. Ci poniamo qui il problema inverso: sotto che ipotesi possiamo ricostruire il segnale $f(t)$ conoscendo solo i suoi campioni $f(n\tau)$? La soluzione a questo problema è cruciale se si desidera interfacciare il mondo analogico con quello digitale. Il campionamento è un passo per la trasformazione di un segnale analogico in un segnale digitale, quindi l'operazione inversa sarà un passo per la trasformazione di un segnale digitale in uno analogico: ci chiediamo quando questo possa essere effettuato senza perdita di informazione.

Una precisa risposta al problema in questione è enunciata nel teorema del campionamento: un segnale $f(t)$ è ricostruibile da $f(n\tau)$ se le componenti armoniche contenute nel segnale hanno frequenze inferiori a $A/2$, dove $A = \frac{1}{\tau}$ è la frequenza di campionamento.

Teorema 4.1 (Teorema del campionamento) *Un segnale $f(t)$ a banda limitata da B Hz, la cui trasformata di Fourier $F(\nu)$ è nulla per $|\nu| \geq B$, può essere univocamente ricostruito dai suoi campioni $f(n\tau)$ ($-\infty < n < \infty$) presi a frequenza $A = \frac{1}{\tau}$, se $A \geq 2B$. La frequenza $2B$ è detta frequenza di Nyquist e identifica la minima frequenza di campionamento che permette la ricostruzione univoca del segnale.*

Dimostrazione. Sia $f(t)$ un segnale a banda limitata da B , la cui trasformata di Fourier $F(\nu)$ sia tale che $F(\nu) = 0$ per $|\nu| \geq B$.

Fissiamo una frequenza di campionamento A tale che $A \geq 2B$ e indichiamo con $\tau = \frac{1}{A}$ l'intervallo di campionamento. Poiché il segnale è a banda limitata da B , e $B \leq \frac{A}{2}$, la funzione $F(\nu)$ è nulla all'esterno dell'intervallo $[-\frac{A}{2}, \frac{A}{2}]$; un esempio è riportato nel grafico in Figura 4.2.

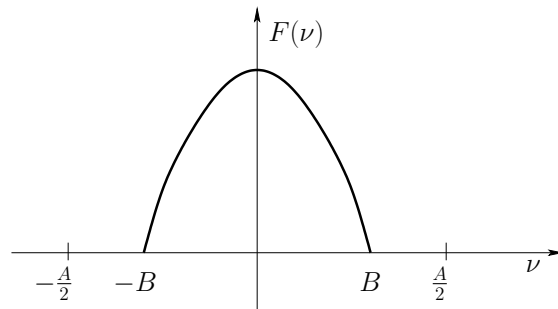


Figura 4.2 Segnale a banda limitata.

Costruiamo ora la funzione $Q(\nu)$ periodica di periodo A e coincidente con $F(\nu)$ nell'intervallo $[-\frac{A}{2}, \frac{A}{2}]$; come si vede dal grafico di $Q(\nu)$ riportato in Figura 4.3, $Q(\nu)$ è ottenuto sommando le repliche di $F(\nu)$ traslate di nA , ($-\infty < n < \infty$):

$$Q(\nu) = \sum_{n=-\infty}^{+\infty} F(\nu - nA)$$

Poichè $A \geq 2B$, tali repliche non si sovrappongono.

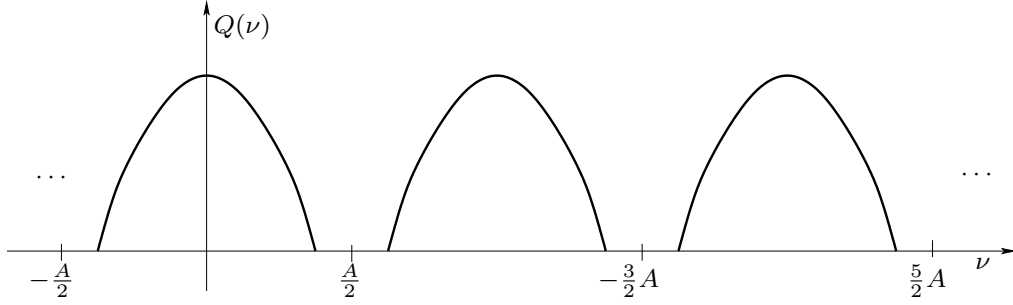


Figura 4.3 Segnale periodico di periodo A .

Conoscendo $Q(\nu)$, risulta possibile ricostruire univocamente $F(\nu)$; in particolare:

$$F(\nu) = \text{rect}\left(\frac{\nu}{A}\right) Q(\nu).$$

Infatti, la funzione $\text{rect}(\frac{\nu}{A})$ vale 1 se $-\frac{A}{2} \leq \nu \leq \frac{A}{2}$, vale 0 altrove. Osserviamo ora che la funzione $Q(\nu)$ è una funzione periodica di periodo A , quindi è sviluppabile in serie di Fourier:

$$Q(\nu) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i \frac{n}{A} \nu}, \quad \text{dove} \quad c_n = \frac{1}{A} \int_{-\frac{A}{2}}^{\frac{A}{2}} Q(\nu) e^{-2\pi i \frac{n}{A} \nu} d\nu. \quad (4.1)$$

Poiché $Q(\nu)$ coincide con $F(\nu)$ nell'intervallo $[-\frac{A}{2}, \frac{A}{2}]$, mentre al di fuori di tale intervallo vale 0, segue:

$$\int_{-\frac{A}{2}}^{\frac{A}{2}} Q(\nu) e^{-2\pi i \frac{n}{A} \nu} d\nu = \int_{-\frac{A}{2}}^{\frac{A}{2}} F(\nu) e^{-2\pi i \frac{n}{A} \nu} d\nu = \int_{-\infty}^{+\infty} F(\nu) e^{2\pi i \frac{-n}{A} \nu} d\nu = f\left(\frac{-n}{A}\right). \quad (4.2)$$

L'ultimo passaggio deriva dal fatto che $f(t)$ può essere ottenuta come antitrasformata di Fourier di $F(\nu)$, e quindi, per ogni t :

$$f(t) = \int_{-\infty}^{+\infty} F(\nu) e^{2\pi i \nu t} d\nu.$$

Scegliendo $t = \frac{-n}{A}$ si ottiene il risultato di cui sopra. Da (4.1) e da (4.2) e ricordando che $\frac{1}{A} = \tau$, si ottiene:

$$c_n = \frac{1}{A} \int_{-\frac{A}{2}}^{\frac{A}{2}} Q(\nu) e^{-2\pi i \frac{n}{A} \nu} d\nu = \frac{1}{A} f\left(-\frac{n}{A}\right) = \tau f(-n\tau). \quad (4.3)$$

Infine, da (4.1), (4.2) e (4.3) si conclude:

$$\begin{aligned}
 F(\nu) &= \text{rect}\left(\frac{\nu}{A}\right) Q(\nu) \\
 &= \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i \frac{n}{A} \nu} \text{rect}\left(\frac{\nu}{A}\right) \\
 &= \tau \sum_{n=-\infty}^{+\infty} f(-n\tau) e^{2\pi i n \tau \nu} \text{rect}(\tau \nu). \tag{4.4}
 \end{aligned}$$

La (4.4) mostra che $F(\nu)$ (e quindi la sua antitrasformata $f(t)$) può essere ricostruita sulla base della conoscenza dei valori $f(n\tau)$ ($-\infty < n < \infty$), ottenuti campionando $f(t)$ con frequenza $\frac{1}{\tau} = A$.

Più precisamente, si ha:

$$\begin{aligned}
 f(t) &= \mathcal{F}^{-1}\{F(\nu)\} \\
 &= \tau \sum_{n=-\infty}^{+\infty} f(-n\tau) \mathcal{F}^{-1}\{e^{2\pi i n \tau \nu} \text{rect}(\tau \nu)\} \quad (\text{per la linearità e per la (4.4)}) \\
 &= \tau \sum_{k=-\infty}^{+\infty} f(k\tau) \mathcal{F}^{-1}\{e^{-2\pi i k \tau \nu} \text{rect}(\tau \nu)\} \quad (\text{ponendo } k = -n) \\
 &= \tau \sum_{k=-\infty}^{+\infty} f(k\tau) \text{sinc}\left(\frac{t}{\tau} - k\right).
 \end{aligned}$$

L'ultimo passaggio è ottenuto applicando la proprietà di traslazione temporale e ricordando che l'antitrasformata della funzione rettangolo $\text{rect}(\tau \nu)$ è $\frac{1}{\tau} \text{sinc}\left(\frac{t}{\tau}\right)$. ■

Esempio 4.1.1

Il segnale di una tipica conversazione telefonica non contiene frequenze significative superiori a 3400 Hz; esso ha quindi una frequenza di Nyquist di 6800 Hz e può quindi essere ricostruito da un suo campionamento a 8 kHz, che è la frequenza di campionamento utilizzata nella telefonia cellulare.

4.1.1 Spettro del segnale campionato

Prendiamo in considerazione un segnale a tempo continuo $f(t)$; il suo campionamento ad intervalli τ può essere interpretato come un segnale nullo per tutti i tempi diversi da $n\tau$ ($-\infty < n < \infty$), mentre ai tempi $n\tau$ si concentra la potenza del segnale. Il segnale campionato può essere visto come un segnale a tempo continuo $f_s(t)$ dove:

$$f_s(t) = \sum_{n=-\infty}^{+\infty} f(n\tau) \delta(t - n\tau).$$

Il segnale campionato risulta quindi una combinazione lineare di impulsi $\delta(t - n\tau)$, in posizioni $n\tau$, pesati coi valori dei campioni $f(n\tau)$ (vedi Figura (4.4)).

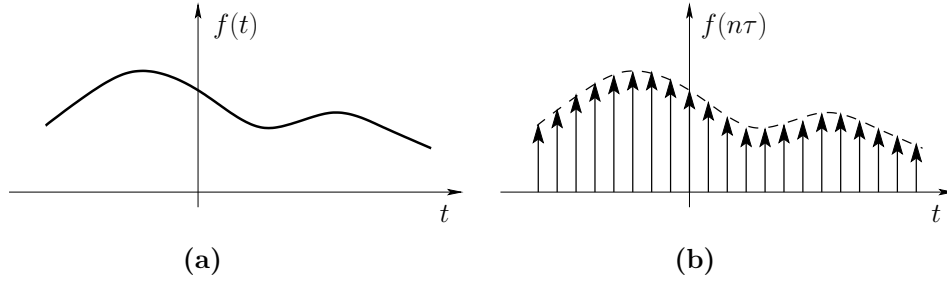


Figura 4.4 (a) Segnale $f(t)$. (b) Segnale campionato $f(n\tau)$.

Ci proponiamo ora di studiare la relazione tra lo spettro $F(\nu)$ del segnale $f(t)$ e lo spettro $F_s(\nu)$ del segnale campionato $f_s(t)$. E' possibile mostrare che, a meno del fattore $\frac{1}{\tau}$, lo spettro $F_s(\nu)$ è la somma delle repliche traslate di nA dello spettro $F(\nu)$ del segnale $f(t)$:

Teorema 4.2 *Il segnale $f_s(t) = \sum_{n=-\infty}^{+\infty} f(n\tau)\delta(t - n\tau)$, ottenuto campionando $f(t)$ con frequenza $A = \frac{1}{\tau}$, ha spettro $F_s(\nu) = \frac{1}{\tau} \sum_{n=-\infty}^{+\infty} F(\nu - nA)$*

Dimostrazione. Osserviamo per prima cosa che $f(t)\delta(t - n\tau) = f(n\tau)\delta(t - n\tau)$, e quindi vale:

$$f_s(t) = \sum_{n=-\infty}^{+\infty} f(n\tau)\delta(t - n\tau) = f(t) \sum_{n=-\infty}^{+\infty} \delta(t - n\tau)$$

La funzione $\sum_{n=-\infty}^{+\infty} \delta(t - n\tau)$ è una funzione periodica con periodo τ e coincide con $\delta(t)$ nell'intervallo $[-\frac{\tau}{2}, \frac{\tau}{2}]$. Essa è quindi sviluppabile in serie di Fourier:

$$\sum_{n=-\infty}^{+\infty} \delta(t - n\tau) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i n \frac{t}{\tau}}$$

$$c_n = \frac{1}{\tau} \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} \delta(t) e^{-2\pi i n \frac{t}{\tau}} dt = \frac{1}{\tau}$$

Possiamo quindi riscrivere $f_s(t) = \frac{1}{\tau} \sum_{n=-\infty}^{+\infty} f(t) e^{2\pi i n \frac{t}{\tau}} = \frac{1}{\tau} \sum_{n=-\infty}^{+\infty} f(t) e^{2\pi i n A t}$, poichè $A = \frac{1}{\tau}$. Calcoliamo infine la trasformata di Fourier di $f_s(t)$ utilizzando l'ultima espressione e applicando le proprietà di linearità e traslazione. Si ottiene:

$$F_s(\nu) = \frac{1}{\tau} \sum_{n=-\infty}^{+\infty} F(\nu - nA)$$

Abbiamo visto che, a meno del fattore $\frac{1}{\tau}$, lo spettro $F_s(\nu)$ del segnale $f_s(t)$ ottenuto campionando $f(t)$ con frequenza $A = \frac{1}{\tau}$ è la somma delle repliche traslate di nA dello spettro $F(\nu)$ del segnale $f(t)$. In Figura 4.5 è mostrato il tipico spettro di musica udibile e lo spettro della stessa musica digitalizzata a 44.1 kHz e a 88.2 kHz.

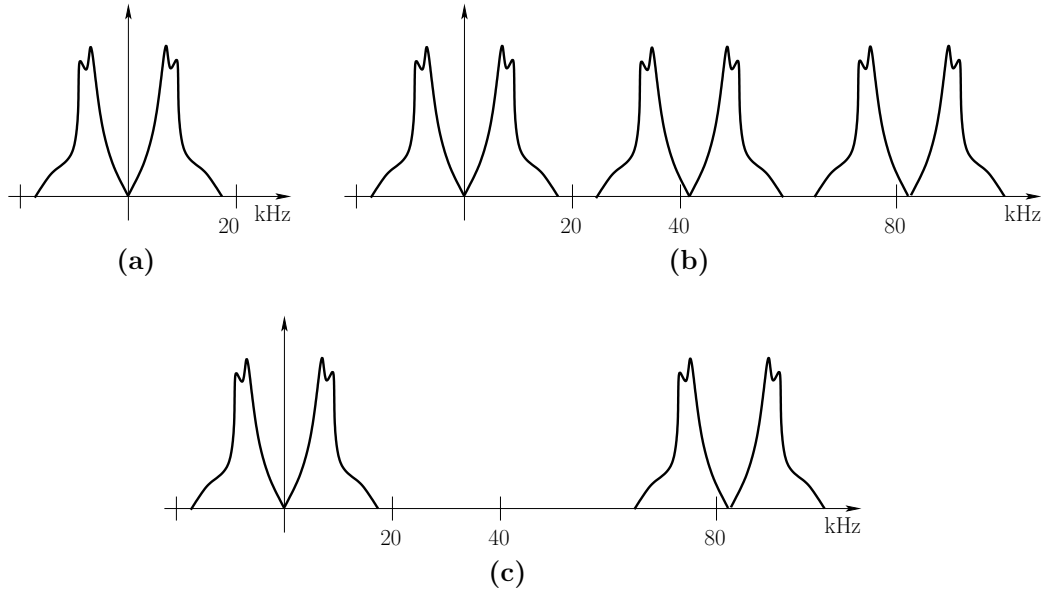


Figura 4.5 (a) Musica udibile. (b) Musica digitalizzata a 44.1 kHz. (c) Musica digitalizzata a 88.2 kHz.

In questo caso la ricostruzione del segnale udibile a partire dalla musica digitalizzata può essere fatta applicando un sistema che elimina tutte le repliche. Come vedremo, sistemi che realizzano questa operazione sono i filtri passa basso, detti *filtri di ricostruzione* in questa applicazione.

4.1.2 Aliasing

Se campioniamo un segnale $f(t)$ a banda limitata da B con una frequenza $A < 2B$, le repliche dello spettro $F(\nu)$ traslate di nA si possono sovrapporre (almeno in parte). Il campionamento in questo caso dà luogo a un segnale il cui spettro è la somma delle repliche sovrappoventesi, come mostrato in Figura 4.6.

Si ottiene quindi un effetto, chiamato *aliasing* da alias (clone, copia) parola presa a prestito dall'inglese (a sua volta presa a prestito dal latino), per cui non è in generale possibile ricostruire il segnale di partenza $f(t)$ sulla base del segnale campionato $f_s(t)$, ed in particolare il filtro di ricostruzione non è in grado di ricostruire lo spettro $F(\nu)$: il segnale ricostruito è quindi generalmente distorto.

Per evitare che un campionatore con frequenza di campionamento A abbia in ingresso un segnale con limite di banda maggiore di $\frac{A}{2}$, si può far precedere il campionatore da un

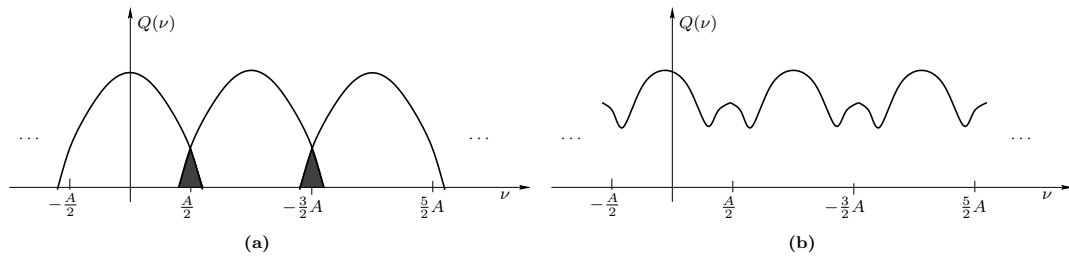


Figura 4.6 (a) Aliasing: sovrapposizione delle repliche. (b) Aliasing: spettro del segnale sottocampionato.

sistema, detto *filtro anti-aliasing*, che elimina le frequenze maggiori di $\frac{A}{2}$. Un campionatore preceduto dal filtro è mostrato in Figura 4.7.

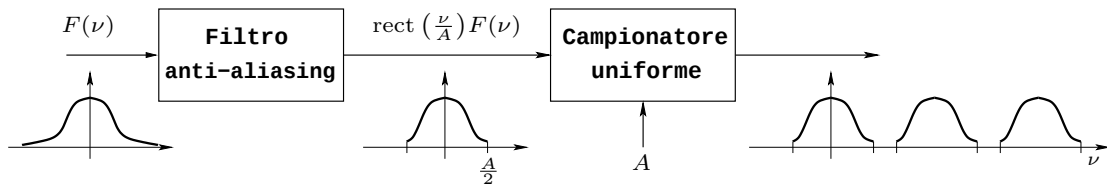


Figura 4.7 Sistema campionatore preceduto da filtro anti-aliasing.

Un segnale campionato ad una frequenza inferiore alla sua frequenza di Nyquist è detto *sottocampionato*; analogamente, un segnale campionato ad una frequenza superiore è detto *sovracampionato*. Un segnale sovracampionato può essere perfettamente ricostruito dai suoi campioni. Il sottocampionamento di un segnale induce invece il fenomeno dell'aliasing, cioè la sovrapposizione delle repliche dello spettro, cosa che non ne consente in generale la ricostruzione. Tuttavia esistono situazioni in cui il sottocampionamento è utile, come nel caso descritto nel seguente esempio.

Esempio 4.1.2

L'orecchio umano non percepisce segnali aventi spettro localizzato sopra i 20 kHz; questi tipi di segnali sono chiamati ultrasuoni. Se si suppone di ascoltare un ultrasuono con spettro tra i 23 e 24 kHz (vedi Figura 4.8) l'orecchio umano non percepisce alcunché e la frequenza di Nyquist del segnale è 48 kHz.

Sottocampioniamo il segnale con una frequenza di 22kHz: si formano in questo caso, tra le altre, repliche del segnale aventi uno spettro tra 1 e 2 kHz (e simmetrico), che mantengono la forma dello spettro originale. Lo spettro risulta pertanto rilocato su frequenze inferiori a 20 kHz rendendolo quindi udibile (Figura (4.8)).

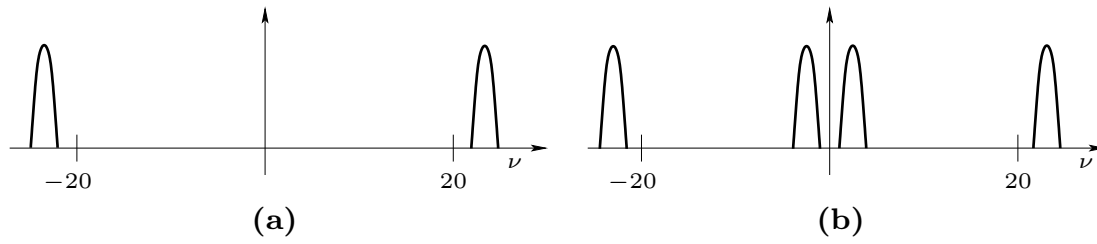


Figura 4.8 (a) Spettro di un ultrasuono. (b) Spettro del segnale sottocampionato.

4.2 Quantizzazione

La quantizzazione è il processo che permette di trasformare un segnale a valori continui in un segnale che assume un numero finito di valori. Un modo semplice di quantizzare consiste nel prefissare un insieme finito di l valori numerici $\{x_1, \dots, x_l\}$ e di associare ad ogni numero x il valore numerico x_k che è più vicino a x .

Se i segnali che prendiamo in considerazione hanno ampiezze comprese tra $-\frac{V}{2}$ e $\frac{V}{2}$, questo può essere ottenuto dividendo l'insieme $[-\frac{V}{2}, \frac{V}{2}]$ in l intervalli, detti *livelli*, ed attribuendo ad un punto $x \in [-\frac{V}{2}, \frac{V}{2}]$ il centro del livello in cui x cade. Detti $\{x_1, \dots, x_l\}$ i centri dei vari livelli, l'operazione di quantizzazione può essere allora descritta dalla funzione Q che ad ogni x associa il centro più vicino:

$$Q(x) = \arg \min_{x_i \in \{x_1, \dots, x_l\}} |x - x_i|.$$

Il sistema che realizza l'operazione di quantizzazione è detto *quantizzatore*. Poiché $\{x_1, \dots, x_l\}$ non è uno spazio vettoriale, il quantizzatore non è in generale un sistema lineare, pur essendo tempo-invariante.

La quantizzazione Q è una funzione multi-uno, quindi essa introduce un errore irreversibile: dato il segnale quantizzato $Q(f(t))$, non è possibile ricostruire in modo esatto il segnale d'origine $f(t)$. Ogni tentativo di ricostruire il segnale comporterà quindi un errore; l'idea intuitiva è che si possa diminuire l'errore aumentando il numero di livelli. Nel prossimo paragrafo accenniamo ad un'analisi quantitativa di questo fenomeno.

4.2.1 Quantizzatore Uniforme e Rumore di Quantizzazione

Un quantizzatore in cui l'intervallo $[-\frac{V}{2}, \frac{V}{2}]$ è suddiviso in l livelli di uguale ampiezza $\frac{V}{l}$ è detto *quantizzatore uniforme*; il valore V è detto *range* del quantizzatore e il numero $\Delta = \frac{V}{l}$ è chiamato *passo di quantizzazione*.

Se $l = 2^m$, gli elementi $\{x_1, \dots, x_l\}$ possono essere codificati con le 2^m parole di m bit:

$$x_i = b_{i1} \cdots b_{im}, \quad \text{con } b_{ik} \in \{0, 1\} \quad (1 \leq i \leq l).$$

Il sistema in questo caso è detto *quantizzatore uniforme* a m bit.

La Figura 4.9 mostra il risultato del campionamento (pallino bianco) e campionamento più quantizzazione uniforme a quattro livelli (pallino nero) di un segnale $f(t)$.

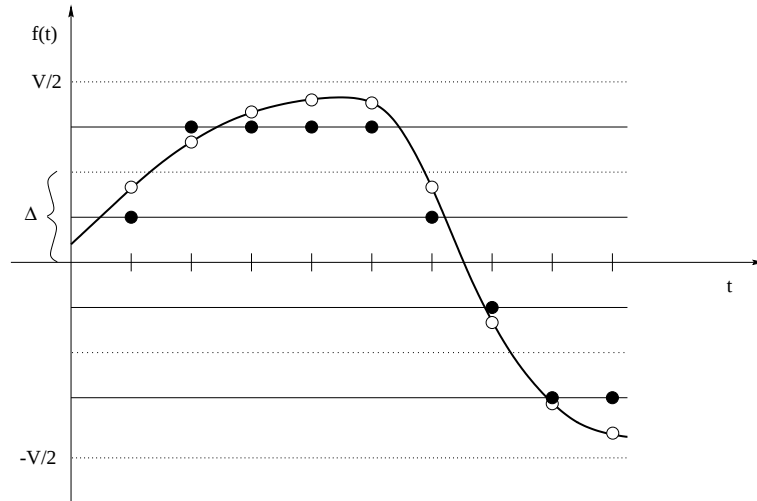


Figura 4.9 Campionamento più quantizzazione uniforme a quattro livelli di un segnale $f(t)$.

Come ben evidenziato dalla Figura, la quantizzazione Q è una funzione multi-uno che introduce un errore irreversibile nel segnale quantizzato. Poichè il numero x viene approssimato dal quantizzatore con $Q(x)$, una naturale misura dell'errore sul numero x è la seguente:

$$e(x) = Q(x) - x.$$

La funzione $e(x)$ viene detta *errore di quantizzazione*; la Figura 4.10 mostra il grafico dell'errore di quantizzazione $e(x)$ per un quantizzatore uniforme di sei livelli.

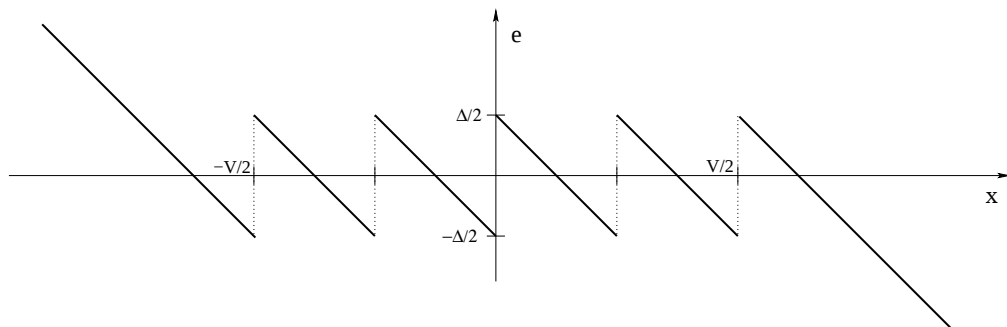


Figura 4.10 Errore di quantizzazione introdotto dal quantizzatore uniforme a sei livelli.

L'errore di quantizzazione ha un comportamento ben differenziato in due zone:

1. Se $x < -\frac{V}{2}$ oppure $x > \frac{V}{2}$, l'errore può essere arbitrariamente grande: in questo caso l'errore è detto *distorsione da overload* e lo si controlla cercando di garantire che i valori del segnale $f(t)$ in ingresso al quantizzatore rientrino nel *range* del quantizzatore, richiedendo cioè che $-\frac{V}{2} \leq f(t) \leq \frac{V}{2}$.
2. Se x è invece interno all'intervallo $-\frac{V}{2} \leq x \leq \frac{V}{2}$, l'errore $e(x)$ si mantiene in valore assoluto minore o uguale a $\frac{\Delta}{2}$, cioè la metà del passo di quantizzazione; tale errore è detto *rumore granulare* e non può essere eliminato.

In seguito considereremo segnali $f(t)$ per cui $-\frac{V}{2} \leq f(t) \leq \frac{V}{2}$, in modo che l'unica sorgente di errore sia il rumore granulare. Supporremo inoltre che i nostri segnali oscillino intorno a 0, in modo che la media dei valori sia 0, valga cioè $\lim_{T \rightarrow \infty} \frac{\int_{-T}^T f(t) dt}{2T} = 0$.

Una naturale misura della dispersione, o potenza, di un segnale $f(t)$ è la sua *varianza* $\text{var} < f >$, data da:

$$\text{var} < f > = \lim_{T \rightarrow \infty} \frac{\int_{-T}^T f^2(t) dt}{2T} = 0$$

Esempio 4.2.1

Consideriamo il segnale sinusoidale di ampiezza $\frac{V}{2}$ dato da $\frac{V}{2} \sin t$. La media di tale segnale è 0, poiché:

$$\lim_{T \rightarrow \infty} \frac{\int_{-T}^T \sin t dt}{2T} = 0.$$

La varianza σ^2 di tale segnale è invece $\frac{V^2}{8}$. Infatti:

$$\sigma^2 = \lim_{T \rightarrow \infty} \frac{\int_{-T}^T (\frac{V}{2} \sin t - 0)^2 dt}{2T} = \lim_{T \rightarrow \infty} \frac{\int_{-T}^T \frac{V^2}{4} \left(\frac{1 - \cos 2t}{2} \right) dt}{2T} = \frac{V^2}{8}.$$

Esempio 4.2.2

Consideriamo la risposta $Q(f(t))$ di un quantizzatore al segnale $f(t)$, in modo che l'errore tra il segnale quantizzato e il segnale sia $e(f(t)) = Q(f(t)) - f(t)$. Sappiamo che, se Δ è il passo di quantizzazione, risulta $-\frac{\Delta}{2} \leq e(f(t)) \leq \frac{\Delta}{2}$.

Per molti segnali deterministici l'errore è uniformemente distribuito tra $-\frac{\Delta}{2}$ e $\frac{\Delta}{2}$. Questo significa che la probabilità che l'errore sia compreso fra e ed $e + de$ è $\frac{de}{\Delta}$. La varianza dell'errore è allora:

$$\text{var} < e(f) > = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} e^2 \frac{de}{\Delta} = \frac{\Delta^2}{12}.$$

Allo scopo di definire una misura di prestazione di un quantizzatore, ricordiamo che il quantizzatore, ricevendo in ingresso un segnale $f(t)$, dà in uscita il segnale $Q(f(t))$ diverso dall'originale, con un errore pari a $e(f(t)) = Q(f(t)) - f(t)$. La potenza del segnale è $\text{var} < f >$, quella dell'errore (o rumore) di quantizzazione è $\text{var} < e(f) >$; un indice di quanto la potenza del segnale superi quella dell'errore è dato dal rapporto $\frac{\text{var} < f >}{\text{var} < e(f) >}$, o, passando al logaritmo di tale rapporto, dal rapporto segnale-rumore di quantizzazione SQNR (Signal Quantization to Noise Ratio), misurato in deciBell (dB):

$$\text{SQNR} = 10 \log_{10} \frac{\text{var} < f >}{\text{var} < e(f) >} \text{ dB},$$

Ci si può chiedere perchè, nella definizione che abbiamo dato di SQNR, non si è preso direttamente in considerazione il rapporto segnale/rumore, bensì il suo logaritmo.

Senza voler essere minimamente esaurienti, osserviamo che spesso la *percezione* I che abbiamo di una grandezza fisica G non è, neanche approssimativamente, proporzionale alla grandezza stessa, bensì a $\log_a \frac{G}{G_0}$, dove G_0 è il più piccolo valore di G che riusciamo a percepire (legge di Weber-Fechner). Spesso viene fissata a 10 la base del logaritmo ($a = 10$), ottenendo la percezione della grandezza pari a $\log_{10} \frac{G}{G_0}$ Bell, o, utilizzando il sottomultiplo deciBell del Bell (abbreviato in dB), pari a $10 \log_{10} G$ dB.

Esempio 4.2.3

Dato un segnale acustico $p(t)$, la pressione sonora efficace p_{eff} è la radice della varianza di $p(t)$, cioè:

$$p_{eff} = \sqrt{\text{var} < p >}$$

La nostra percezione dell'intensità di un suono non è tuttavia proporzionale a p_{eff} , ma può essere approssimativamente descritta dal livello sonoro L_p , dove:

$$L_p = 10 \log_{10} \frac{\text{var} < p >}{\text{var} < p_0 >} \text{ dB}$$

Con p_0 denotiamo qui la pressione sonora della soglia uditiva.

La misura SQNR di prestazione di un quantizzatore con range V è data dal rapporto segnale-rumore di quantizzazione SQNR (Signal Quantization to Noise Ratio) di un segnale in ingresso sinusoidale del tipo $\frac{V}{2} \sin t$, ipotizzando che l'errore sia uniformemente distribuito tra $-\frac{\Delta}{2}$ e $\frac{\Delta}{2}$. Vale il seguente importante:

Teorema 4.3 *Se un segnale sfrutta tutto il range V di un quantizzatore a m bit, allora $\text{SQNR} \approx 6.02m + 1.76$ dB. In particolare, ogni bit aggiunto ad un quantizzatore comporta un incremento di 6.02 dB al rapporto segnale rumore.*

Dimostrazione. Dagli esempi 4.2.1, 4.2.2 precedentemente studiati, sappiamo che il segnale $f(t) = \frac{V}{2} \sin t$ ha varianza $\text{var} < f > = \frac{V^2}{8}$ e il corrispondente errore di quantizzazione $e(f)$, uniformemente distribuito, ha varianza $\text{var} < e(f) > = \frac{\Delta^2}{12}$. Segue:

$$\text{SQNR} = 10 \log_{10} \frac{\text{var} < f >}{\text{var} < e(f) >} = 10 \log_{10} \frac{V^2/8}{\Delta^2/12}$$

Osserviamo che $\frac{V^2/8}{\Delta^2/12} = (\frac{V}{\Delta})^2 \cdot \frac{12}{8} = l^2 \cdot \frac{12}{8}$, dove $l = 2^m$ è il numero di livelli. Applicando le proprietà della funzione log, si ottiene infine:

$$SQNR = 20 \log_{10} l + 10 \log_{10} \frac{12}{8} = (20 \log_{10} 2) \cdot m + 1.76 = 6.02m + 1.76$$

I

Esempio 4.2.4

Determinare il numero di bit da aggiungere a un quantizzatore per migliorare il rapporto segnale-rumore da 40 dB a 68 dB.

Osservando che la differenza tra le prestazioni richieste è di 18 dB e che ogni bit aggiunto al quantizzatore migliora SQNR di 6.02 dB, concludiamo che basta aggiungere $3 \approx \frac{18}{6.02}$ bit.

Esempio 4.2.5

Determinare lo SQNR di un quantizzatore tipico usato in telefonia digitale, sapendo che il numero di livelli è 256.

Poichè 256 livelli possono essere codificati con 8 bit, il rapporto segnale-rumore è $6.02 \cdot 8 + 1.76 \approx 49.76$ dB.

4.2.2 Quantizzatore Non Uniforme

Spesso per segnali reali la probabilità che un segnale abbia valore tra y e $y + dy$ viene a dipendere da y . La Figura 4.11 mostra come in un classico segnale del parlato ampiezze elevate siano meno probabili di ampiezze piccole:

È intuitivo che in questo caso una quantizzazione più fine per ampiezze piccole migliori la qualità del segnale quantizzato, diminuendo l'errore quadratico medio.

Questo risultato può essere ottenuto come segue:

1. Si applica al segnale (che per semplicità consideriamo normalizzato a valori in $[0, 1]$) un funzione F invertibile che comprime le ampiezze vicine a 1 (vedi Figura 4.12).
2. Si applica al segnale “compresso” un quantizzatore uniforme.
3. Il segnale quantizzato viene “decompresso” applicando la funzione F^{-1} inversa di F .

Questo processo, detto *companding* (COMPressing and exPANDING) permette di realizzare un quantizzatore non uniforme, come schematizzato nella Figura 4.13.

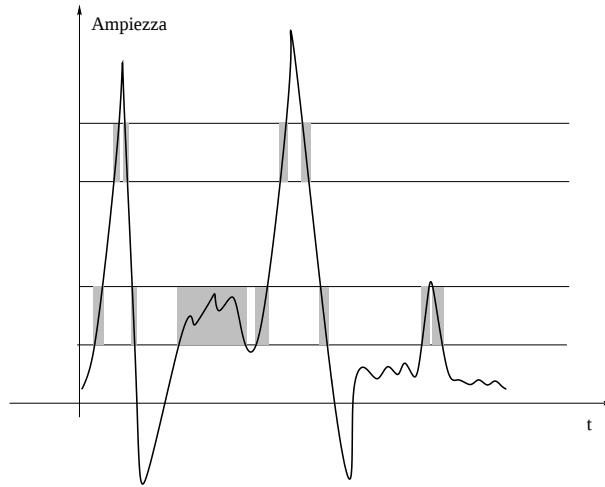


Figura 4.11 Probabilità di varie ampiezze (in grigio).

Esempio 4.2.6

Nelle applicazioni in telefonia viene usata la famiglia di funzioni μ -law:

$$F_{\mu}(f) = \frac{\ln(1 + \mu|f|)}{\ln(1 + \mu)} \operatorname{sgn}(f), \quad \text{con } -1 \leq f \leq 1,$$

dove f è il segnale normalizzato e μ è un parametro opportuno (usualmente posto a 100 o più recentemente a 255). La funzione μ -law inversa $F_{\mu}^{-1}(y)$ è data da:

$$F_{\mu}^{-1}(y) = \frac{1}{\mu} \left((1 + \mu)^{|y|} - 1 \right) \operatorname{sgn}(y).$$

4.3 Convertitore Analogico-Digitale (ADC)

Applicando un campionatore con frequenza di campionamento A e consecutivamente un quantizzatore a m bit, il segnale $f(t)$ osservato per un tempo T può essere trasformato in un vettore di $T \cdot A$ componenti, ognuna delle quali rappresentabile con m bit: esso può quindi essere memorizzato in forma digitale usando $m \cdot T \cdot A$ bit ed eventualmente modificato.

Il sistema che realizza questa trasformazione è detto *convertitore analogico-digitale* (ADC) e può essere descritto come in Figura 4.14.

Il filtro antialiasing in figura è un filtro passa-basso che cancella tutte le frequenze superiori alla metà della frequenza di campionamento; l'effetto è quello di porre in ingresso al campionatore un segnale a banda limitata, la cui frequenza di Nyquist non superi la frequenza di campionamento.

Esistono essenzialmente due differenti tipologie di convertitori analogico-digitale.

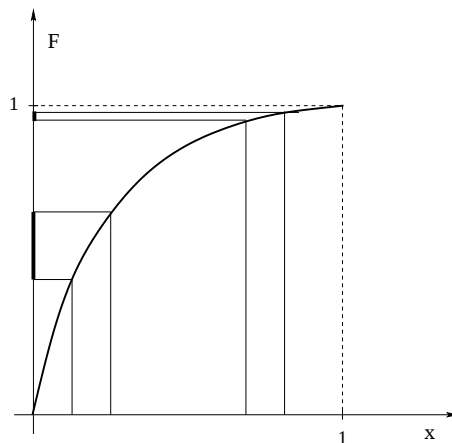
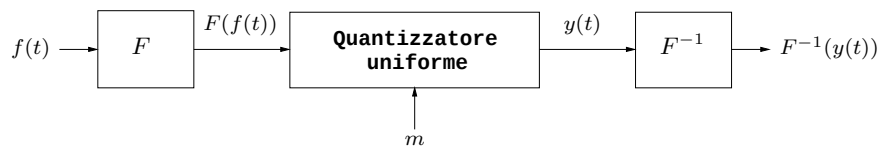
Figura 4.12 Funzione di compressione F .

Figura 4.13 Quantizzatore non uniforme.

- Nel primo tipo, rappresentato nella figura precedente, il campionatore opera vicino alla frequenza di Nyquist del segnale e il quantizzatore è un quantizzatore ad m bit ($m \gg 1$).
- Nel secondo tipo si usa un campionatore a frequenza molto superiore alla tasso di Nyquist e si sfrutta il sovracampionamento del segnale per ridurre i bit del quantizzatore, con l'aiuto di filtri passa-basso.

Nei convertitori analogico-digitali del primo tipo l'elemento critico è il quantizzatore. Fra i vari modelli disponibili, presentiamo qui il Flash ADC di cui riportiamo in Figura 4.15 la realizzazione del quantizzatore a 2 bit.

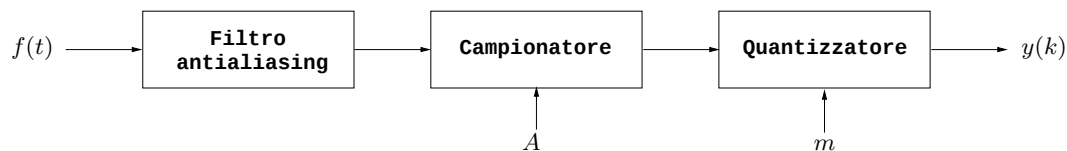


Figura 4.14 Convertitore analogico-digitale

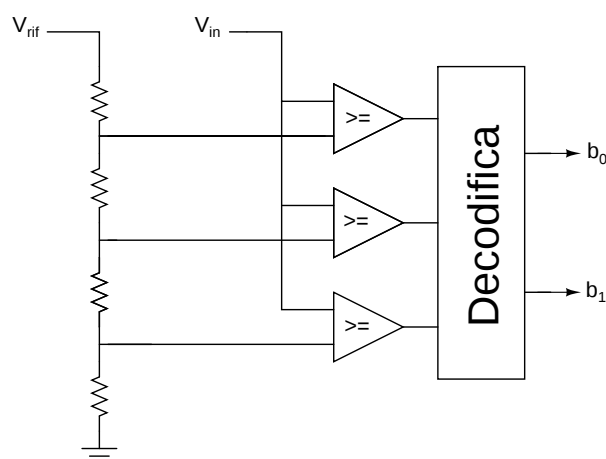


Figura 4.15 Flash ADC (2 bit).

Nel caso generale di un quantizzatore a m bit, la tensione del segnale di ingresso V_{in} viene confrontata con $2^m - 1$ tensioni di riferimento ottenute con un sistema di 2^m resistenze uguali poste in serie. Le uscite binarie dei comparatori (indicati col simbolo $>=$ in figura) vengono poi trasformate negli m bit di codifica da un opportuno circuito booleano (chiamato Decodifica in figura).

Il flash ADC è sicuramente il più veloce convertitore disponibile, ma è costituito da un numero elevato (2^m) di resistenze che devono essere molto accurate: questo rende difficile e costosa la realizzazione per alti valori di m ($m \gg 8$).

Questo fatto è una caratteristica dei convertitori analogico-digitali di questo tipo, che richiedono un'estrema accuratezza nella costruzione del sistema quantizzatore. Risulta allora conveniente un diverso tipo di convertitore, basato sull'idea di operare con frequenze di campionamento molto superiori al tasso di Nyquist utilizzando però un quantizzatore con un minor numero di bit.

Il vantaggio ottenuto è duplice:

1. operando a frequenze molto superiori al tasso di Nyquist, il filtro antialiasing diventa meno critico di quanto non lo sia nei convertitori del primo tipo, e può essere progettato con maggior facilità;
2. l'aumento della frequenza di campionamento, con l'ausilio di opportuni filtri, si può tradurre in un miglioramento del SQNR del convertitore.

CAPITOLO 5

Analisi in frequenza di sistemi a tempo discreto

Questo capitolo è dedicato all'introduzione delle caratteristiche (in tempo e frequenza) dei filtri digitali, unitamente ai loro strumenti di analisi.

Per prima cosa vengono introdotte le nozioni di *tempo e frequenza normalizzati*. Un segnale a tempo discreto è infatti ottenuto campionando un segnale analogico a un dato intervallo τ (e quindi ad una frequenza di campionamento $A = \frac{1}{\tau}$): è quindi utile rappresentare sia il tempo che la frequenza come numeri puri, normalizzandoli sia rispetto a τ che ad A .

Vien poi studiato il concetto di *trasformata zeta* di un segnale. Essenzialmente, l'idea è quella di associare in maniera biunivoca ad ogni segnale a tempo discreto $x(n)$ una funzione $X(z)$ a variabile complessa. Storicamente, i fondamenti di questa nozione sono stati introdotti da De Moivre nella prima metà del 1700, con applicazioni al calcolo delle probabilità. La trasformata zeta permette di ottenere agevolmente lo spettro del segnale $x(n)$, che risulta essere $X(e^{2\pi i\Omega})$, dove Ω è la frequenza normalizzata.

Dopo aver introdotto le trasformate zeta di alcuni segnali, vengono presentate alcune proprietà che portano a utili regole di manipolazione simbolica. Si osserva l'importante fatto che la trasformata $Y(z)$ della risposta di un sistema LTI S a un ingresso $x(n)$ è $H(z)X(z)$, dove $H(z)$ è la funzione di trasferimento del sistema, cioè la trasformata zeta della risposta di S all'impulso. La funzione di trasferimento $H(z)$ gioca un ruolo cruciale: da un lato la risposta in frequenza di un sistema è descritta da $H(e^{2\pi i\Omega})$, dall'altro da semplici proprietà di $H(z)$ è possibile determinare la causalità e stabilità del sistema.

Vengono quindi studiati i filtri con risposta all'impulso finita (FIR) e particolari filtri con risposta all'impulso infinita (IIR), i cosiddetti filtri ricorsivi. Si osserva che i filtri FIR sono sempre stabili e possono avere fase lineare e si propone un semplice criterio per la stabilità di filtri IIR. Si discute infine la rappresentazione di filtri FIR e IIR ricorsivi come *reti* ottenute componendo 3 tipi di sistemi-base (somma, moltiplicazione per costante, ritardo): questa rappresentazione è particolarmente indicata come punto di partenza per la realizzazione via s/w o h/w del filtro.

5.1 Frequenza e Tempo Normalizzati

Il segnale $f(t)$, campionato con intervallo di campionamento τ e quindi con frequenza di campionamento $A = 1/\tau$, può essere descritto da:

$$f_s(t) = \sum_{n=-\infty}^{\infty} f(n\tau) \delta(t - n\tau)$$

Lo spettro del segnale campionato risulta essere:

$$F_s(\nu) = \sum_{n=-\infty}^{\infty} f(n\tau) \cdot e^{-2\pi i n \tau \nu} = \frac{1}{\tau} \sum_{n=-\infty}^{+\infty} F(\nu - nA)$$

Come osservato in 4.1.1, $F_s(\nu)$ è una funzione periodica di periodo A in quanto somma di repliche periodiche dello spettro $F(\nu)$ di $f(t)$. Osserviamo ora che le formule precedenti possono essere semplificate, se si fissa come unità di misura del tempo l'intervallo di campionamento τ anziché, come usualmente, il secondo. Questo equivale a considerare come misura del tempo il cosiddetto *tempo normalizzato* t' , cioè il rapporto tra il tempo t e l'intervallo τ :

$$t' = \frac{t}{\tau}, \quad t = t'\tau$$

Il segnale $f(t)$, descritto attraverso il tempo normalizzato, diventa il segnale $x(t')$ dove naturalmente:

$$x(t') = f(t'\tau)$$

Si osservi che il tempo normalizzato corrispondente ai multipli interi $n\tau$ di τ è $\frac{n\tau}{\tau} = n$: utilizzando il tempo normalizzato, il segnale a tempo discreto ottenuto campionando un segnale a tempo continuo con intervallo di campionamento τ , è descritto dalla sequenza $x(n)$, che non contiene esplicitamente il termine τ .

Esempio 5.1.1

Se l'intervallo di campionamento τ è 0.001 sec, il tempo normalizzato t' corrispondente a un tempo di 3 sec è $\frac{3}{0.001} = 3000$.

In modo analogo chiamiamo *frequenza normalizzata* Ω la misura di frequenza ottenuta avendo la frequenza di campionamento $A = 1/\tau$ come unità. La frequenza normalizzata è quindi data dal rapporto tra la frequenza e la frequenza di campionamento:

$$\Omega = \frac{\nu}{A} = \nu\tau \quad , \quad \nu = \frac{\Omega}{\tau}$$

Esempio 5.1.2

Se la frequenza di campionamento è 1000 Hz, la frequenza normalizzata Ω corrispondente a una frequenza di 410 Hz è $\frac{410}{1000} = 0.41$.

La frequenza normalizzata $\Omega = 0.25$, se la frequenza di campionamento è di 800 Hz, corrisponde ad una frequenza di $800 \cdot 0.25 = 200$ Hz.

Ricordiamo che lo spettro $F_s(\nu)$ del segnale campionato $f_s(t)$ è

$$F_s(\nu) = \sum_{n=-\infty}^{\infty} f(n\tau) \cdot e^{-2\pi i n \tau \nu}$$

Poichè la relazione tra frequenza ν e frequenza normalizzata Ω è $\nu = \frac{\Omega}{\tau}$, detto $X_s(\Omega)$ lo spettro del segnale campionato espresso in funzione della frequenza normalizzata Ω , vale:

$$X_s(\Omega) = F_s\left(\frac{\Omega}{\tau}\right) = \sum_{n=-\infty}^{\infty} f(n\tau) e^{-2\pi i n \Omega} = \sum_{n=-\infty}^{\infty} x(n) e^{-2\pi i n \Omega}$$

Concludendo, l'uso del tempo normalizzato $t' = \frac{t}{\tau}$ e della frequenza normalizzata $\Omega = \frac{\nu}{\nu_s}$ permette da un lato di descrivere un segnale campionato attraverso una sequenza $x(n)$, dall'altro di semplificare il suo spettro $X_s(\Omega)$:

$$X_s(\Omega) = \sum_{n=-\infty}^{\infty} x(n) e^{-2\pi i n \Omega}$$

Si osservi che $X_s(\Omega)$ è una funzione periodica di periodo 1. Essa è univocamente determinata da $|X_s(\Omega)|$ e da $\text{ph}(X_s(\Omega))$ con $-\frac{1}{2} \leq \Omega \leq \frac{1}{2}$.

Si osservi che $\Omega = 0$ corrisponde alla frequenza $\nu = 0$, mentre $\Omega = \frac{1}{2}$ corrisponde alla frequenza di Nyquist, metà della frequenza di campionamento.

5.2 Trasformata Zeta

Per motivare l'introduzione di questa nuova trasformata, richiamiamo che il segnale a tempo discreto $x(n)$ ha lo spettro $X_s(\Omega)$ dove:

$$X_s(\Omega) = \sum_{n=-\infty}^{\infty} x(n) e^{-2\pi i n \Omega} = \sum_{n=-\infty}^{\infty} x(n) (e^{2\pi i \Omega})^{-n}$$

Posto $z = e^{2\pi i \Omega}$, l'espressione precedente diventa: $\sum_{n=-\infty}^{\infty} x(n) z^{-n}$. Questo motiva la seguente importante:

Definizione 5.1 *Dato il segnale a tempo discreto $x(n)$, la sua trasformata zeta è la funzione a variabile complessa*

$$X(z) = \sum_{n=-\infty}^{\infty} x(n) z^{-n}$$

Si osservi che $X(z)$ non è generalmente definita su tutto il piano complesso, poichè per certi valori di z la serie $\sum_{n=-\infty}^{\infty} x(n) z^{-n}$ può non convergere; si può dimostrare che l'insieme di definizione di $X(z)$ è una corona circolare centrata intorno all'origine.

Conoscendo la trasformata zeta $X(z)$ di un segnale $x(n)$, è possibile costruire lo spettro $X_s(\Omega)$ del segnale:

Fatto 5.1 *Lo spettro $X_s(\Omega)$ del segnale $x(n)$, la cui trasformata zeta è $X(z)$, è dato da*

$$X_s(\Omega) = \sum_{n=-\infty}^{\infty} x(n) e^{-2\pi i n \Omega} = X(e^{2\pi i \Omega}).$$

Poichè $e^{2\pi i \Omega}$ è una funzione periodica di periodo 1, anche $X_s(\Omega)$ lo è.

Si può ulteriormente dimostrare che da $X(z)$ è possibile ricostruire i valori $x(n)$; ci limiteremo qui a dare la formula di ricostruzione nell'ipotesi che $X(z)$ sia definita sulla circonferenza di raggio 1. Da Fatto 5.1 vale infatti:

$$X_s(\Omega) = \sum_{n=-\infty}^{\infty} x(n) e^{-2\pi i n \Omega}$$

Osservando che l'espressione precedente è lo sviluppo in serie di Fourier della funzione periodica $X_s(\Omega)$, il coefficiente $x(n)$ della combinazione lineare è dato da:

$$x(n) = \int_{-\frac{1}{2}}^{\frac{1}{2}} X_s(\Omega) e^{2\pi i n \Omega} d\Omega = \int_{-\frac{1}{2}}^{\frac{1}{2}} X(e^{2\pi i \Omega}) e^{2\pi i n \Omega} d\Omega.$$

5.2.1 Proprietà della trasformata Zeta

La trasformata zeta di un segnale a tempo discreto $x(n)$ è la funzione $X(z)$ a variabile complessa tale che:

$$X(z) = \sum_{n=-\infty}^{+\infty} x(n)z^{-n}.$$

Dato il segnale, possiamo costruire la trasformata; data la trasformata, possiamo ricostruire il segnale: la trasformata è una nuova rappresentazione del segnale.

Le trasformate di alcuni segnali (tra cui l'impulso $\delta(n)$ e il gradino unitario $u(n)$) si possono calcolare direttamente e sono date in Tabella 5.1.

Tabella 5.1 Alcune coppie note di trasformate.

Segnale	Trasformata
$\delta(n)$	1
$\delta(n - m)$	z^{-m}
$a^n u(n)$	$\frac{1}{1 - az^{-1}}$

Ad esempio la trasformata zeta di $a^n u(n)$ è:

$$\sum_{n=-\infty}^{\infty} a^n u(n) z^{-n} = \sum_{n=0}^{\infty} a^n z^{-n} = \sum_{n=0}^{\infty} (az^{-1})^n = \frac{1}{1 - az^{-1}}.$$

La trasformata zeta possiede un certo numero di proprietà che la rendono uno strumento flessibile ed efficace per lo studio dei segnali a tempo discreto. Le principali proprietà sono presentate in Tabella 5.2.

Tabella 5.2 Proprietà della trasformata z .

Proprietà	$\mathbf{x(n)}$	$\leftrightarrow$	$\mathbf{X(z)}$
Linearità	$ax(n) + by(n)$		$aX(z) + bY(z)$
Traslazione	$x(n - n_0)$		$z^{-n_0} X(z)$
Convoluzione	$\sum_{k=-\infty}^{+\infty} x(k)y(n - k)$		$X(z)Y(z)$

La linearità è provata come segue:

$$\sum_{n=-\infty}^{+\infty} (ax(n) + by(n))z^{-n} = a \sum_{n=-\infty}^{+\infty} x(n)z^{-n} + b \sum_{n=-\infty}^{+\infty} y(n)z^{-n} = aX(z) + bY(z).$$

Per quanto riguarda la proprietà di traslazione, si consideri $\sum_{n=-\infty}^{+\infty} x(n-a)z^{-n}$; ponendo $k = n - a$, quindi $n = a + k$, si ha:

$$\sum_{n=-\infty}^{+\infty} x(n-a)z^{-n} = \sum_{k=-\infty}^{+\infty} x(k)z^{-a-k} = z^{-a} \sum_{k=-\infty}^{+\infty} x(k)z^{-k} = z^{-a} X(z).$$

Proviamo infine la proprietà di convoluzione:

$$\begin{aligned} X(z)Y(z) &= \sum_{k=-\infty}^{+\infty} x(k)z^{-k} \sum_{j=-\infty}^{+\infty} y(j)z^{-j} = \sum_{k=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} x(k)y(j)z^{-k-j} \\ &= \sum_{n=-\infty}^{+\infty} z^{-n} \sum_{k+j=n} x(k)y(j) = \sum_{n=-\infty}^{+\infty} \left(\sum_{k=-\infty}^{+\infty} x(k)y(n-k) \right) z^{-n}. \end{aligned}$$

Applicando le proprietà presentate in Tabella 5.2 alle coppie di trasformate elencate in Tabella 5.1 è possibile determinare le trasformate zeta di una vasta classe di segnali di interesse pratico.

Esempio 5.2.1

Determinare la trasformata z del segnale $2^{-n}u(n) + u(n-3)$.

$2^{-n}u(n)$	$\longleftrightarrow$	$\frac{1}{1 - (2z)^{-1}}$	da Tabella 5.1
$u(n)$	$\longleftrightarrow$	$\frac{1}{1 - z^{-1}}$	da Tabella 5.1
$u(n-3)$	$\longleftrightarrow$	$\frac{z^{-3}}{1 - z^{-1}}$	Prop. di traslaz.
$2^{-n}u(n) + u(n-3)$	$\longleftrightarrow$	$\frac{1}{1 - z^{-1}} + \frac{z^{-3}}{1 - z^{-1}}$	Prop. di linear.

5.3 Sistemi LTI a Tempo Discreto: filtri FIR e IIR

Sia dato un sistema LTI S a tempo discreto; sappiamo dal Capitolo 1 che l'uscita $y(n)$ del sistema S avente come ingresso il segnale $x(n)$ è ottenuto dalla convoluzione di $x(n)$ con $h(n)$, dove $h(n)$ è la risposta del sistema all'impulso unitario $\delta(n)$:

$$y(n) = h(n) * x(n) = \sum_{k=-\infty}^{+\infty} h(k)x(n-k).$$

Dalla proprietà di convoluzione (riportata in Tabella 5.2) si ottiene che:

$$Y(z) = H(z)X(z),$$

dove $Y(z)$, $H(z)$ e $X(z)$ sono rispettivamente la trasformata z di $y(n)$, $h(n)$ e $x(n)$. La funzione $H(z)$, trasformata z della risposta all'impulso $h(n)$, è detta *funzione di trasferimento* del sistema (vedi Figura 5.1).

Riassumendo si ha quindi l'importante risultato:

Fatto 5.2 *Dato un sistema LTI la cui funzione di trasferimento è $H(z)$, detta $Y(z)$ e $X(z)$ rispettivamente la trasformata z dell'uscita $y(n)$ e dell'ingresso $x(n)$, allora vale:*

$$Y(z) = H(z)X(z).$$

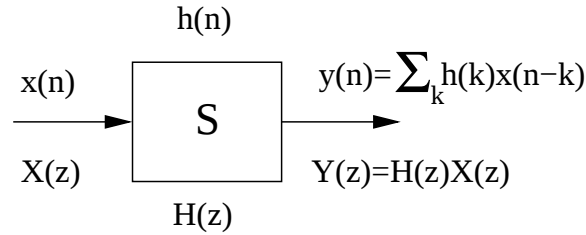


Figura 5.1 Funzione di trasferimento del sistema S .

Si osservi che la funzione di trasferimento può essere calcolata come $H(z) = \frac{Y(z)}{X(z)}$. La conoscenza della funzione di trasferimento di un sistema LTI permette di determinare molte proprietà del sistema, tra le quali la stabilità e la risposta in frequenza, come vedremo in seguito. E' allora importante sviluppare metodi per il calcolo della funzione di trasferimento: le prossime sezioni presentano tecniche per determinare la funzione di trasferimento dei filtri FIR e IIR.

5.3.1 Filtri FIR

Un sistema LTI causale a tempo discreto è detto filtro FIR (Finite Impulse Response) se la risposta $h(n)$ all'impulso unitario è finita nel senso che $h(n) = 0$ per $n < 0$ e per $n \geq M$ per un opportuno $M > 0$. Ricordando che l'uscita è ottenuta dalla convoluzione dell'ingresso con la risposta all'impulso, la relazione ingresso-uscita è data dalla seguente equazione alle differenze finite:

$$y(n) = \sum_{k=0}^{M-1} h(k)x(n-k). \quad (5.1)$$

Passando alle trasformate z e applicando la proprietà della traslazione temporale, si ottiene:

$$Y(z) = \sum_{k=0}^{M-1} h(k)z^{-k}X(z) = \left(\sum_{k=0}^{M-1} h(k)z^{-k} \right) X(z)$$

Si può a questo punto calcolare la funzione di trasferimento $H(z)$, che risulta essere un polinomio in z^{-1} :

$$H(z) = \frac{Y(z)}{X(z)} = \sum_{k=0}^{M-1} h(k)z^{-k}$$

Le caratteristiche più interessanti dei filtri FIR sono le seguenti:

1. Un filtro FIR è sempre causale e stabile.
2. Un filtro FIR può avere fase lineare: se la funzione $h(n)$ è simmetrica o antisimmetrica rispetto a $(M-1)/2$ (cioè $h(k) = h(M-1-k)$ oppure $h(k) = -h(M-1-k)$), allora la fase $\text{ph } H(e^{2\pi i\Omega})$ è lineare.

In Figura 5.2 sono rappresentate le risposte all'impulso di due filtri FIR a fase lineare, uno antisimmetrico (a), l'altro simmetrico (b).

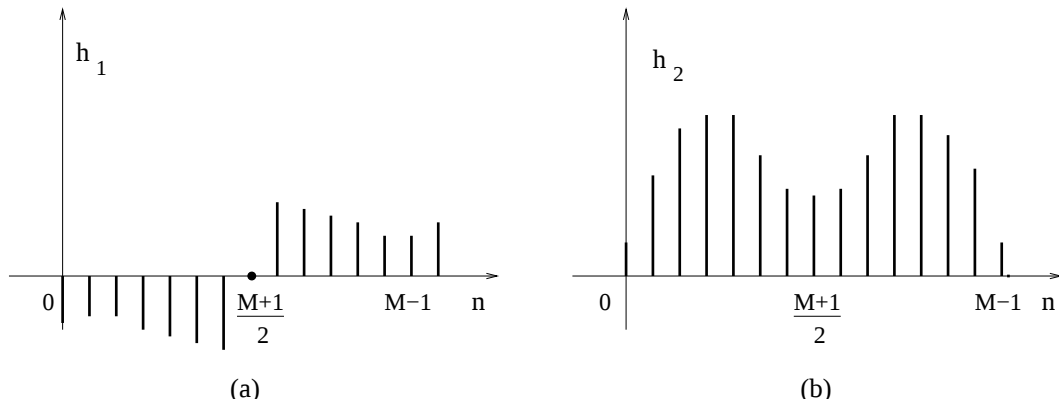


Figura 5.2 Risposta all'impulso di un filtro a fase lineare (a) antisimmetrico (b) simmetrico.

Esempio 5.3.1

Determinare la risposta in frequenza (guadagno e fase) del seguente filtro FIR in cui la relazione ingresso-uscita è data da:

$$y(n) = \frac{1}{2}x(n) + \frac{1}{2}x(n-1)$$

Passando alle trasformate zeta si ottiene $Y(z) = \frac{1}{2}X(z) + \frac{1}{2}z^{-1}X(z)$, da cui possiamo determinare la funzione di trasferimento del filtro:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{2} + \frac{1}{2}z^{-1}.$$

La risposta in frequenza è $H(e^{2\pi i\Omega})$, cioè:

$$H(e^{2\pi i\Omega}) = \frac{1}{2} + \frac{1}{2}e^{-i2\pi\Omega} = \frac{e^{i\pi\Omega} + e^{-i\pi\Omega}}{2}e^{-i\pi\Omega} = \cos \pi\Omega \cdot e^{-i\pi\Omega}$$

Il guadagno risulta quindi $\cos^2 \pi\Omega$ e la fase $-\pi\Omega$.

Si osserva che la fase è lineare e che il filtro alle basse frequenze ($\Omega \approx 0$) ha guadagno 1, mentre alle alte frequenze ($\Omega \approx \frac{1}{2}$) ha guadagno 0: si tratta di un filtro passa basso.

5.4 Filtri IIR

Consideriamo ora un sistema LTI in cui la relazione ingresso-uscita verifica la seguente equazione alle differenze finite:

$$y(n) = \sum_{k=0}^{L-1} a_k x(n-k) + b_1 y(n-1) + b_2 y(n-2) + \dots + b_M y(n-M)$$

Filtri causali che verificano l'equazione precedente sono detti *filtri ricorsivi* poichè l'uscita al passo n è definita ricorsivamente in funzione delle uscite ai passi $n-1, n-2, \dots, n-M$. Si osservi che se $b_1 = b_2 = \dots = b_M = 0$ il filtro è un filtro FIR.

Per determinare la funzione di trasferimento si passa alla trasformata zeta. Ricordando le proprietà di linearità e traslazione si ottiene:

$$Y(z) = \sum_{k=0}^{L-1} a_k z^{-k} X(z) + b_1 z^{-1} Y(z) + b_2 z^{-2} Y(z) + \dots + b_M z^{-M} Y(z)$$

Risolvendo questa equazione di primo grado rispetto a $Y(z)$ si ha:

$$Y(z) = \frac{\sum_{k=0}^{L-1} a_k z^{-k}}{1 - b_1 z^{-1} - \dots - b_M z^{-M}} X(z)$$

Da questo si ottiene la funzione di trasferimento, $H(z)$, del filtro:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{L-1} a_k z^{-k}}{1 - b_1 z^{-1} - \dots - b_M z^{-M}}$$

In questo caso la funzione di trasferimento è una funzione razionale, riducibile cioè al rapporto di due polinomi:

$$H(z) = \frac{P(z)}{Q(z)}, \quad \text{dove } P(z) \text{ e } Q(z) \text{ sono polinomi}$$

Contrariamente ai filtri FIR, i filtri IIR causali non hanno mai fase lineare e possono essere non stabili. Poichè la stabilità è una condizione importante per le molte applicazioni, diamo senza dimostrazione una semplice caratterizzazione dei filtri IIR stabili in termini di proprietà della loro funzione di trasferimento. Ricordiamo che, data una funzione razionale $\frac{P(z)}{Q(z)}$ gli zeri di $P(z)/Q(z)$ sono le soluzioni dell'equazione $P(z) = 0$, mentre i poli di $P(z)/Q(z)$ sono le soluzioni dell'equazione $Q(z) = 0$.

Fatto 5.3 *Un filtro è causale e stabile se i poli della sua funzione di trasferimento $H(z)$ sono contenuti nel cerchio di raggio 1, cioè se ogni polo ha modulo strettamente minore di 1.*

Esempio 5.4.1

Determinare se il filtro seguente è stabile.

$$y(n) = x(n) + 2y(n-1) - \frac{3}{4}y(n-2)$$

Passando alla trasformata z , si ottiene $Y(z) = X(z) + 2z^{-1}Y(z) - \frac{3}{4}z^{-2}Y(z)$. Risolvendo rispetto a $Y(z)$ si determina la funzione di trasferimento

$$H(z) = \frac{1}{1 - 2z^{-1} + \frac{3}{4}z^{-2}} = \frac{z^2}{z^2 - 2z + \frac{3}{4}}$$

I poli di $H(z)$ sono le soluzioni dell'equazione $z^2 - 2z + \frac{3}{4} = 0$, cioè $z_1 = \frac{1}{2}$ e $z_2 = \frac{3}{2}$. Poichè $|\frac{3}{2}| > 1$, un polo ha modulo maggiore di 1 e quindi il filtro è instabile.

Esempio 5.4.2

Determinare se il filtro seguente è stabile.

$$y(n) = x(n) - \frac{3}{4}y(n-2).$$

La funzione di trasferimento risulta essere:

$$H(z) = \frac{1}{1 + \frac{3}{4}z^{-2}} = \frac{z^2}{z^2 + \frac{3}{4}}$$

I poli di $H(z)$ sono le soluzioni dell'equazione $z^2 + \frac{3}{4} = 0$, cioè $z_1 = \frac{\sqrt{3}}{2}i$ e $z_2 = -\frac{\sqrt{3}}{2}i$. Poichè $|z_1| = |z_2| = \frac{\sqrt{3}}{2} < 1$, tutti i poli sono contenuti nel cerchio di raggio 1, quindi il filtro è stabile.

5.4.1 Implementazione di filtri come Reti

Nelle scorse sezioni abbiamo studiato due classi di filtri digitali (FIR e IIR ricorsivi), la cui relazione ingresso-uscita viene data attraverso semplici equazioni alle differenze. In questa sezione rappresentiamo invece tali sistemi come *reti* ottenute componendo 3 tipi di sistemi-base (somma, moltiplicazione per costante, ritardo): questa rappresentazione è particolarmente indicata come punto di partenza per la realizzazione via s/w o h/w del filtro.

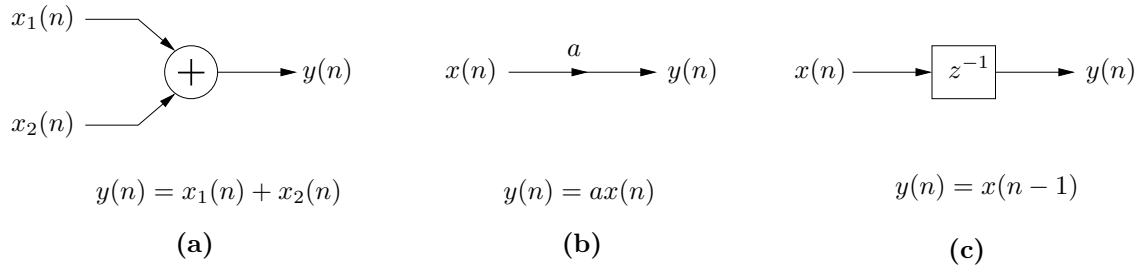


Figura 5.3 (a) Sistema sommatore. (b) Moltiplicatore per costante. (c) Ritardo unitario.

Consideriamo 3 semplici sistemi-base: il *sommatore*, il *moltiplicatore per costante*, il *ritardo unitario*. Il *sommatore* riceve in ingresso due segnali $x_1(n)$ e $x_2(n)$ dando in uscita il segnale somma $y(n) = x_1(n) + x_2(n)$; dato un numero a , il *moltiplicatore per a* riceve in ingresso il segnale $x(n)$ e dà in uscita il segnale $y(n) = a \cdot x(n)$; il *ritardo unitario* riceve in ingresso il segnale $x(n)$ e dà in uscita $y(n) = x(n-1)$, cioè lo stesso segnale ritardato. L'illustrazione di tali sistemi mediante diagrammi a blocchi è data in Figura 5.3.

Componendo sistemi dei 3 tipi descritti si ottiene una rete; una *rete* è un grafo con un nodo-sorgente (il nodo di ingresso), un nodo-pozzo (il nodo di uscita), mentre gli altri nodi sono etichettati con una delle 3 operazioni elementari: ogni nodo etichettato con $+$ ha due archi entranti, ogni nodo etichettato con una costante a ha un unico arco entrante, così come ogni nodo-ritardo, etichettato con z^{-1} . Ponendo in ingresso a una rete il segnale $x(n)$, si ottiene in uscita un preciso segnale $y(n)$: la rete realizza dunque un sistema che, essendo costituito da semplici sistemi LTI, sarà a sua volta LTI.

Esempio 5.4.3

La Figura 5.4 mostra un esempio di rete, contenente due nodi sommatore, un nodo ritardo e tre nodi moltiplicatori.

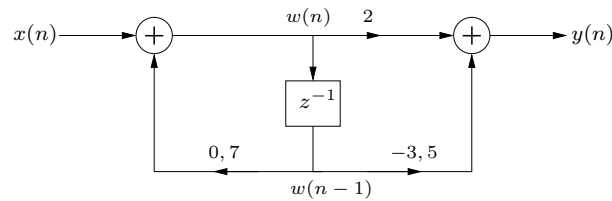


Figura 5.4 Diagrammi a blocchi con ritardi “compattati”.

Indicando con $x(n)$ l'ingresso della rete, $y(n)$ la sua uscita e $w(n)$ l'uscita del primo sommatore, si osserva che gli ingressi al primo sommatore sono $x(n)$ e $0,7w(n-1)$, mentre il secondo sommatore ha come uscita $y(n)$ e come ingressi $2w(n)$ e $-3,5w(n-1)$.

1). Si ottiene quindi:

$$\begin{aligned} w(n) &= x(n) + 0,7w(n-1) \\ y(n) &= 2w(n) - 3,5w(n-1). \end{aligned}$$

Eliminando dalle precedenti due equazioni $w(n-1)$ si ottiene $w(n) = \frac{5}{3}x(n) - \frac{1}{3}y(n)$ e quindi $w(n-1) = \frac{5}{3}x(n-1) - \frac{1}{3}y(n-1)$. Sostituendo i valori ottenuti di $w(n)$ e $w(n-1)$ nell'equazione $y(n) = 2w(n) - 3,5w(n-1)$, si ottiene infine:

$$y(n) = 0,7y(n-1) + 2x(n) - 3,5x(n-1).$$

La rete realizza dunque il filtro ricorsivo IIR la cui relazione ingresso-uscita è data dalla precedente equazione.

Una importante osservazione è che lo stesso filtro, descritto dalla relazione ingresso-uscita, può essere realizzato da reti molto diverse. Il seguente esempio mostra una rete, molto diversa da quella studiata nell'esempio precedente, che realizza lo stesso sistema.

Esempio 5.4.4

Si vuol realizzare mediante una rete il filtro ricorsivo IIR descritto dalla seguente equazione alle differenze:

$$y(n) = 0,7y(n-1) + 2x(n) - 3,5x(n-1), \quad (5.2)$$

Una possibile soluzione, che ricalca l'equazione alle differenze, è data in Figura 5.5. La prima parte della rete si occupa di realizzare la combinazione lineare degli ingressi $x(n)$ e $x(n-1)$. La seconda parte somma tale combinazione con l'uscita ritardata e moltiplicata per 0.7.

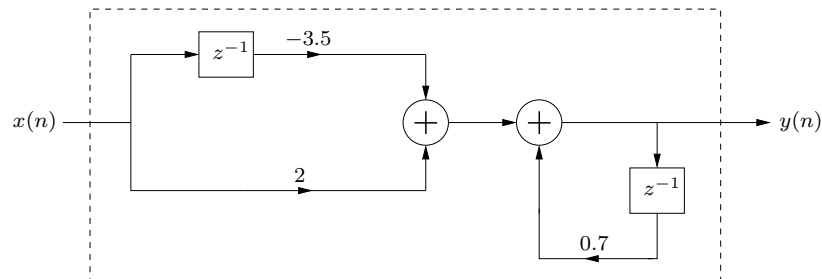


Figura 5.5 Realizzazione di un sistema LTI con diagrammi a blocchi.

Le reti che abbiamo introdotte realizzano filtri FIR o IIR ricorsivi. Ad ogni rete si potrà dunque associare la funzione di trasferimento del filtro realizzato, che permette di derivare tutte le proprietà del sistema. Lo stesso sistema può essere tuttavia realizzato da differenti reti, che ne rappresentano differenti implementazioni. Si possono quindi considerare problematiche di due tipi:

Analisi: data una rete, trovare la funzione di trasferimento del filtro digitale realizzato dalla rete stessa; la funzione di trasferimento contiene infatti tutte le informazioni sul sistema realizzato.

Sintesi: dato un filtro, specificato ad esempio dalla relazione ingresso-uscita o equivalentemente dalla sua funzione di trasferimento, determinare una rete che lo realizzi, eventualmente ottimizzando rispetto a particolari richieste implementative (ad esempio, richiedendo che le componenti di ritardo siano ridotte al minimo).

Affrontiamo qui solo il primo problema: data la rete, determinare la funzione di trasferimento del sistema che la rete realizza.

Risulta conveniente considerare le operazioni dei nodi delle reti come operazioni sulle trasformate z dei segnali di ingresso. Poichè la trasformata di una combinazione lineare è la combinazione lineare delle trasformate, i nodi somma e moltiplicazione per costante si comportano, nello spazio delle trasformate, sommando e moltiplicando per costante gli ingressi. Il nodo ritardo, invece, si comporta moltiplicando l'ingresso per z^{-1} , come conseguenza della regola della trasformata del segnale ritardato. In Figura 5.6 è rappresentato il comportamento del sommatore, moltiplicazione per costante, ritardo sia nello spazio dei segnali che in quello delle loro trasformate.

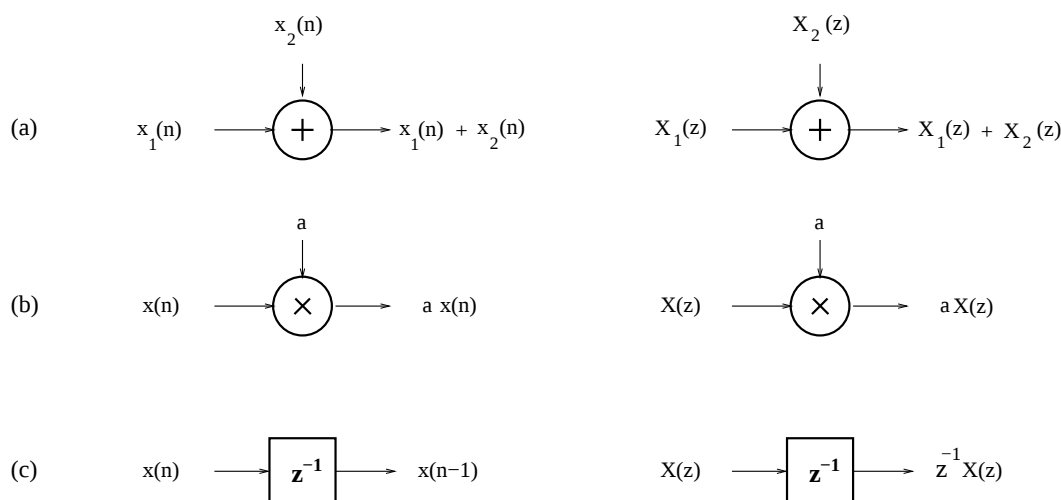


Figura 5.6 Operatore di somma (a), moltiplicazione per costante (b) e ritardo temporale (c), rispettivamente nel dominio del tempo n e nella trasformata z .

La seguente procedura riduce il calcolo della funzione di trasferimento del sistema realizzato da una rete a un problema di eliminazione di variabili:

Input: Una rete R con un unico ingresso e un'unica uscita;

1. Associa all'arco di ingresso la funzione $X(z)$, all'arco di uscita la funzione $Y(z)$, all'arco di uscita di ogni nodo interno k la funzione ausiliaria $W_k(z)$.
2. Per ogni nodo della rete, ad esclusione del nodo di ingresso, costruisci un'equazione come segue:
 - a. Se il nodo k è un nodo di ritardo temporale e la funzione associata all'arco di ingresso è $W(z)$, poni $W_k(z) = z^{-1}W(z)$.
 - b. Se il nodo k è un nodo di somma e le funzioni associate ai due archi di ingresso sono $W(z)$ e $S(z)$, poni $W_k(z) = W(z) + S(z)$.
 - c. Se il nodo k è un nodo di moltiplicazione per la costante a e la funzione associata all'arco di ingresso è $W(z)$, poni $W_k(z) = aW(z)$.
3. Elimina dalle equazioni ottenute le funzioni ausiliarie $W_k(z)$ associate ai nodi interni, ottenendo una relazione del tipo $Y(z) = H_R(z)X(z)$.

Output: La funzione di trasferimento $H_R(z) = \frac{Y(z)}{X(z)}$

Esempio 5.4.5

Si consideri la rete R1 indicata in Figura 5.7. Si determini la funzione di trasferimento del filtro realizzato dalla rete, la sua relazione ingresso-uscita e si discuta la stabilità.

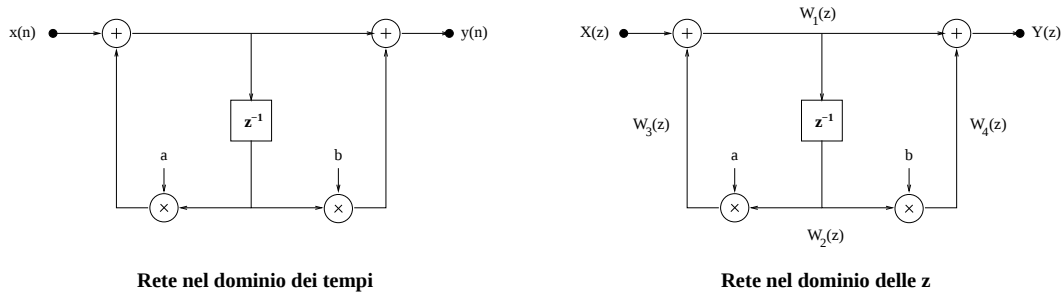


Figura 5.7 Esempio di rete.

Per ottenere la relazione ingresso-uscita $Y(z) = H_{R1}(z)X(z)$ e quindi la funzione di trasferimento $H_{R1}(z)$ del filtro da essa realizzato, è sufficiente scrivere le equazioni ai nodi ed eliminare successivamente le variabili ausiliarie.

Le equazioni sono le seguenti:

$$W_1(z) = X(z) + W_3(z)$$

$$W_2(z) = z^{-1}W_1(z)$$

$$W_3(z) = aW_2(z)$$

$$W_4(z) = bW_2(z)$$

$$Y(z) = W_1(z) + W_4(z),$$

da cui si ricava che la funzione di trasferimento $H_{R1}(z)$ è:

$$H_{R1}(z) = \frac{Y(z)}{X(z)} = \frac{1 + bz^{-1}}{1 - az^{-1}}.$$

Dalla precedente equazione si deriva facilmente:

$$Y(z) = az^{-1}Y(z) + X(z) + bz^{-1}X(z).$$

Antitrasformando, otteniamo la relazione ingresso-uscita del filtro IIR realizzato dalla rete:

$$y(n) = ay(n-1) + x(n) + bx(n-1).$$

La funzione di trasferimento $H_{R1}(z)$ ha un unico polo $z = a$: se $|a| < 1$ il filtro è stabile, altrimenti è instabile.

Riferimenti

Come si evince dalla lettura di queste dispense introduttive sull'elaborazione dei segnali, la comprensione dei concetti di base (spettro, campionamento, quantizzazione, filtri lineari analogici e digitali, modulazione lineare) è agevolata dall'utilizzo di strumenti matematici avanzati. Per questo motivo i libri che espongono questi concetti in maniera rigorosa introducono con un minimo di profondità le nozioni matematiche di riferimento: in quest'area bisogna diffidare da soluzioni semplicistiche, spesso portatrici di ambiguità e confusione.

Vogliamo qui segnalare alcuni riferimenti bibliografici agli studenti che desiderano approfondire gli argomenti trattati. In questi libri, che generalmente raccolgono materiale sviluppato in corsi universitari, la ricerca della semplicità non va a scapito della serietà della trattazione.

Il volume [1] dà evidenza ai metodi matematici collegati alle diverse rappresentazione dei segnali, esemplificandone l'interesse e la potenza per lo studio di sistemi lineari quali i filtri e la modulazione di ampiezza. Una introduzione alle tecniche di elaborazione di segnali digitali si trova in [3], semplificazione della classica trattazione [2]. [4] è un libro più avanzato, ma di grande chiarezza, dedicato allo studio dei principi della comunicazione analogica e digitale.

[1] C.L.Phillips, J.M.Parr, E.A.Riskin, Signals, Systems and Transforms, Prentice Hall, 2003

[2] A.V.Oppenheim, R.W.Schafer, Discrete Time Signal Processing, Prentice Hall, 1989

[3] J.H.McClellan, R.W.Schafer, M.A.Joder, DSP first: a multimedia approach, Prentice Hall International, 1998

[4] Y.D.Gibson, Digital and Analog Communications, Macmillan Publishing Company, 1993