Pasquale Mainolfi

Appunti per il corso di Musica Elettronica

Rappresentazione dei segnali nel dominio della frequenza: la Trasformata di Fourier

Senza alcun dubbio, fino a qualche tempo fa, il parametro oggettivamente più complesso associato ai fenomeni ondulatori era la forma d'onda, comprenderne la natura. E, continuerebbe tutt'oggi ad essere così, se non fosse per la brillante intuizione avuta dal matematico francese J. B. J. Fourier nel 1822 che, contro la classe e il sapere scientifico della sua epoca, ipotizzò uno sviluppo possibile grazie ad un'ampia classe di funzioni. Una prospettiva affascinante, del tutto nuova: *la forma d'onda come risultato di una complessa stratificazione di forme d'onda elementari.*

"*. . . una funzione periodica a tempo continuo $v(t)$ di periodo T e frequenza fondamentale $f_0 = \frac{1}{T}$ può essere scomposta in una sommatoria infinita di un termine costante a_0 (valore medio della funzione in un periodo) e di infinite sinusoidi (armoniche) di frequenza multipla della fondamentale; l'ampiezza delle armoniche è decrescente e tendente a zero con il crescere della frequenza.*
Allo stesso modo, ogni segnale T-periodico può essere ottenuto come sovrapposizione di infiniti segnali trigonometrici, ognuno dei quali associato ad una frequenza multipla della frequenza fondamentale.".

J. B. J. Fourier (1882)

$$v(t) = a_0 + \sum_{k=1}^{\infty} A_k \mathring{u} \sin(k\omega_0 t + \phi_k)$$

Discrete Fourier Trasform DFT - IDFT

la trasformata di fourier FT è un operatore matematico che applicato ad un segnale nel dominio del tempo ne individua un altro nel dominio della frequenza.

$$X(f) = \int_{-\infty}^{+\infty} x(t)\ e^{-i2\pi ft} dt \qquad x(t) = \int_{-\infty}^{+\infty} X(f)\ e^{i2\pi ft} df$$

L'equivalente della FT per l'elaborazione dei segnali digitali, campionati, è la DFT (Discrete Fourier Trasform)

$$DFT\ X[k] = \sum_{n=0}^{N-1} x(t)\ e^{\frac{-i2\pi nk}{N}} \qquad IDFT\ x[n] = \sum_{k=0}^{N-1} X[k]\ e^{\frac{-i2\pi nk}{N}}$$

$e^{ix} = \cos x + i\ \sin x$
$i = \sqrt{-1}$
$X[f] =$ vettore frequuenza o dominio trasformato
$x[n] =$ vettore tempo

$$X[k] = \sum_{n=0}^{N-1} x(t)\ \cos\left(\frac{2\pi nk}{N}\right) - i\ \sin\left(\frac{2\pi nk}{N}\right)$$
$$x[n] = \sum_{k=0}^{N-1} X(k)\ \cos\left(\frac{2\pi nk}{N}\right) + i\ \sin\left(\frac{2\pi nk}{N}\right)$$

Una funzione periodica $f(t)$, dunque, può essere rappresentata come:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(2\pi nt) + b_n \sin(2\pi nt)$$

$a_0,\ a_n,\ b_n =\ coefficienti\ di\ fourier$
$a_0 = \frac{1}{\pi} \sum_{n=0}^{N-1} x[n]$
$a_n = \frac{1}{\pi} \sum_{n=0}^{N-1} x[n] \cos(\frac{2\pi nt}{N})$
$b_n = \frac{1}{\pi} \sum_{n=0}^{N-1} x[n] \sin(\frac{2\pi nt}{N})$

In altre parole, rappresentiamo un segnale nel dominio della frequenza con N campioni *complessi* o, meglio, con N numeri complessi c, con $c = a + ib$.

$a = Re$ parte reale
$\sqrt{-1}b = Im$ parte immaginaria

Semplicemente una coppia di valori, uno reale ed uno immaginario che è possibile rappresentare come un vettore $\overrightarrow{v}$ nel piano *complesso di Gauss*: sull'asse x la parte reale e, sull'asse y la parte immaginaria.

$\overrightarrow{v} = a + ib$

con modulo:
$||\overrightarrow{v}|| = \sqrt{a^2 + b^2}$

e, fase:
$\theta = \tan^{-1}(\frac{b}{a})$

Nel dominio trasformato, le frequenze $[0 \leq k \leq \frac{N}{2} - 1]$ corrispondono alla contenuto reale, ovvero, quelle componenti che generalmente si considerano ai fini dell'analisi dei risultati, viceversa, le componenti $[\frac{N}{2} \leq k \leq N - 1]$ si considerano *frequenze negative*. Tali componenti non rappresentano fisicamente frequenze negative, ma piuttosto un aspetto matematico delle componenti di

frequenza in termini di fase e rotazione nel piano complesso.
Uno spettro, dunque, completamente determinato da $\frac{N}{2}+1$ componenti con N pari e $\frac{N+1}{2}$ con N dispari e, $X[N/2] =$ soglia di Nyquist.

Nota: è possibile pensare la DFT come come un semplice prodotto tra matrici.

$$X[k] = \sum_{n=0}^{N-1} x[t]\; e^{\frac{-i2\pi nk}{N}} = \sum_{n=0}^{N-1} x[t]\; W_N^{-nk}$$

se poniamo:
$W_N = e^{\frac{-i2\pi}{N}}$

allora:

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & W_N & W_N^2 & \dots & W_N^{N-1} \\ 1 & W_N^2 & W_N^4 & \dots & W_N^{2(N-1)} \\ 1 & W_N^3 & W_N^6 & \dots & W_N^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \dots & W_N^{(N-1)^2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{N-1} \end{bmatrix}$$

Short-Time Fourier Trasform STFT - ISTFT

la DFT fornisce una misura del peso delle diverse componenti in un segnale **periodico**. In natura, però, è molto difficile trovare segnali periodici. Per catturare le variazioni spettrali tempo-dipendenti di un segnale è necessario un campionamento ad intervalli regolari, analizzare di volta in volta brevi finestre di campioni in cui si ipotizza che il segnale possa essere effettivamente periodico: tante piccole DFT.
Tale procedura prende il nome di *Short Time Fourier Trasform* o *STFT*.

$$X[m,k] = \sum_{n=0}^{N-1} x[n]w[n-mR]e^{-\frac{i2\pi nk}{N}}$$

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[m,k]e^{\frac{i2\pi nk}{N}}$$

$x[n] =$ segnale in input
$w[n] =$ finestra di analisi
$X[m,k] =$ DFT della finestra analizzata
$R =$ hop size

Una STFT si presenta, dunque, come una matrice $n \times m$, con n righe pari alla lunghezza in campioni della finestra (*bins*)

$$n = \begin{cases} \frac{DFT_{length}}{2} + 1 & \text{se il segnale è reale} \\ DFT_{length} & \text{per segnali non reali} \end{cases}$$

ed m colonne (*frames*) pari alla lunghezza totale del segnale diviso R (*fattore di salto*, definisce il tasso di campionamento della STFT).

$$m = \left\lfloor \frac{x_{length} - w_{length}}{w_{length} - R} \right\rfloor + 1$$

Il fattore R definisce la quantità di sovrapposizioni tra le finestre di analisi (*overlap*). Ciò permette di ridurre in modo significativo la perdita di informazioni (*frequency leakage*) dovuta dall'attenuazione agli estremi della finestra di analisi. Nota bene che, per assicurarsi una valida ricostruzione del segnale:

$$\sum_{m=-\infty}^{\infty} w[n - mR] = 1$$

Detto ciò, ogni colonna della matrice conterrà le informazioni di modulo e fase di ciascuna componente analizzata (*bin*). Se ipotizziamo un segnale di $N = 16$ campioni ed una finestra w di lunghezza $w_{length} = 4$, $R = 4$ ed $sr = 44100$, la matrice risultante avrà quattro righe ed altrettante colonne. È come se facessimo passare ogni frame attraverso un banco di quattro filtri passa-banda con frequenza centrale fissata a:

$$f_{(centrale,\ \mathrm{bin})} = \frac{n}{w_{length}} \text{ů} sr \qquad n = [0, 1, 2, 3, ..., w_{length} - 1]$$

da ciò consegue che: **la definizione frequenziale è direttamente proporzionale al valore che assume w_{length},, la definizione temporale è, invece, inversamente proporzionale**.
Sono vani gli sforzi fatti per inseguire una fedele rappresentazione tempo-frequenza di un segnale, non c'è verso di sapere con esattezza quali sono le componenti spettrali nei singoli istanti di tempo: quanto più una funzione è localizzata nel dominio temporale, tanto più la sua trasformata è dispersa nel dominio spettrale e, viceversa.

Fast Fourier Trasform FFT - IFFT

La DFT nella sua forma classica

$$X[k] = \sum_{n=0}^{N-1} x[n] \exp^{-jk\frac{2\pi}{N}n}$$

è estremamente lenta e dispendiosa a causa dell'eccessivo carico computazionale: passare dal dominio del tempo a quello della frequenza significa dover eseguire N addizioni complesse ed N moltiplicazioni complesse, con una complessità diretta pari a $O(n^2)$. Un carico computazionale tipico di un prodotto matrice vettore.

Un problema importante, fino al 1965, anno in cui J. Cooley e J. Tukey propongono un algoritmo per la semplificazione dei calcoli. Decomporre il problema, una DFT di lunghezza N, in piccoli sottoproblemi (*divide et impera*) in modo tale ridurre il carico computazionale e, nel contempo eliminare buona parte di operazioni inutili e ridondanti. Alla base l'idea, dato un vettore di lunghezza $N = n \times m$, di effettuare m DFT di lunghezza n o viceversa. Parliamo dell'algoritmo, o meglio degli algoritmi di calcolo veloci noti con il nome di FFT (Fast Fourier Trasform). Radix-2 FFT, è sicuramente quello più conosciuto. Esso consente di ridurre le operazioni e passare da $O(n^2)$ a $N \log_2 N$ (significa che, per eseguire una DFT a 512 punti l'algoritmo necessita di 4608 operazioni invece che 262144).
Per capire come questo sia possibile dobbiamo considerare il lemma di Danielson-Lanczos: *una trasformata di Fourier discreta di lunghezza N può essere riscritta come la somma di due trasformate di Fourier discrete di lunghezza N/2, una formata dai termini ad indice pari e una formata dai termini ad indice dispari.*

$$
\begin{aligned}
X[k] &= \sum_{n=0}^{N-1} x[n] W_N^{-nk} \\
&= \sum_{n=0}^{\frac{N}{2}-1} x[2n] e^{-\frac{i2\pi(2n)k}{N}} + \sum_{n=0}^{\frac{N}{2}-1} x[2n+1] e^{-\frac{i2\pi(2n+1)k}{N}} \\
&= \sum_{n=0}^{\frac{N}{2}-1} x[2n] e^{-\frac{i2\pi nk}{N/2}} + \sum_{n=0}^{\frac{N}{2}-1} x[2n+1] e^{-\frac{i2\pi nk}{N/2} \text{ů} \frac{i\pi k}{N/2}} \\
&= \sum_{n=0}^{\frac{N}{2}-1} x[2n] e^{-\frac{i2\pi nk}{N/2}} + C_k \sum_{n=0}^{\frac{N}{2}-1} x[2n+1] e^{-\frac{i2\pi nk}{N/2}}
\end{aligned}
$$

con il termine $C_k = e^{-\frac{i2\pi k}{N}}$

Se tutto restasse così, il risparmio sarebbe minimo: due DFT di lunghezza dimezzata. Se, però, consideriamo di iterare il procedimento fino ad arrivare ad un numero N di DFT unitarie, il gioco è fatto! Infatti, ad una DFT unitaria corrisponde copiare l'input in output e, tutto si riduce ad una combinazione lineare degli input pesati dal termine C_k.