

INTRODUZIONE ALLA SINTESI DEL SUONO

Fondamenti di Teoria dei Segnali *

Francesca Ortolani

1 La Sintesi del Suono

Nei vari corsi di Fonia, Music Technology, Music Production, Musica Elettronica, etc, una larga area degli insegnamenti impartiti si concentra sullo studio dei sintetizzatori (analogici e digitali), dei Virtual Instruments e della computer music in generale. Prima di entrare in dettaglio nella descrizione di un sintetizzatore e dei suoi blocchi componenti, cerchiamo di capire cosa si intende per "Sintesi del Suono" e riassumiamo alcuni concetti fondamentali dell'Acustica e della Teoria dei Segnali. Quando si parla di sintesi del suono si intende solitamente un processo elettronico di generazione di suoni. Per capire il significato generale di sintesi e riportarlo al caso della generazione del suono, supponiamo di voler studiare un sistema più o meno complesso, ad esempio uno strumento musicale acustico. Osserviamo una chitarra e proviamo a scomporla in tanti blocchi, cercando di capire il funzionamento di ciascuno di essi. Come prima cosa vediamo che una chitarra è formata da corde che vibrano e una cassa acustica, di cui vogliamo studiare il comportamento fisico. Osserviamo che il timbro dello strumento può cambiare a seconda dei materiali utilizzati nella sua realizzazione, l'altezza (*pitch*) delle note emesse pizzicando le corde cambia a seconda della lunghezza delle corde stesse, della tensione, della sezione, etc. Ciò che stiamo facendo altro non è che un'operazione di analisi, ovvero stiamo scomponendo il nostro oggetto di studio negli elementi che lo costituiscono. La sintesi è l'operazione inversa, ovvero rimettere insieme questi blocchi. Esistono varie tecniche finalizzate alla produzione del suono per sintesi. Ne vedremo diverse in questo libro con lo scopo di capire i principi su cui si basano, ma senza addentrarci troppo negli aspetti progettuali e ingegneristici che richiederebbero da parte del lettore conoscenze in elettronica ed elaborazione numerica dei segnali troppo avanzate.

*Estratto da "Sintesi del Suono - Appunti e approfondimenti per i corsi di Music Technology" - Volume 1 - 1st Edition. Rev. 1.

2 Il Suono e la sua percezione

Per cominciare è bene distinguere quelli che sono i parametri fisici del suono dai parametri percettivi, ovvero rispettivamente i parametri oggettivi e quelli soggettivi. In termini fisici il suono è generato dalla variazione della pressione in un mezzo (ad esempio l'aria), che si propaga senza trasporto di materia, ovvero si ha un moto locale delle particelle che compongono il mezzo. Affinchè il suono si propaghi è necessario che il mezzo sia elastico. Quello che avviene nel processo di propagazione dell'energia acustica è che il mezzo si comprime e si espande (per intenderci, stiamo parlando di piccolissime variazioni di pressione, meno di 1 Pa). È importante afferrare il concetto che si tratta di un moto locale: le particelle oscillano avanti e indietro attorno ad un punto di riposo e l'oscillazione viene trasmessa alle particelle adiacenti senza che queste si spostino dalla loro posizione. Questo fenomeno ondulatorio è descritto da un moto armonico e le particelle che oscillano si comportano come un sistema massa-molla [2]. Per ora ipotizziamo che non ci siano fenomeni di attrito che smorzino il moto e consideriamo il moto periodico di periodo T : chiamando $x(t)$ il segnale acustico in questione, si avrà che $x(t) = x(t + T)$ se il segnale è periodico.

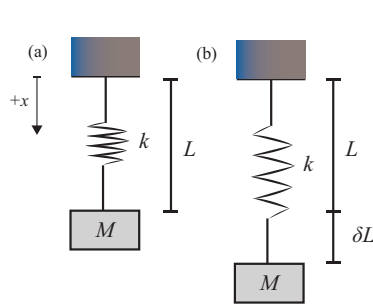


Figure 1: Sistema massa-molla. (a) Posizione di riposo. (b) Sistema in fase di espansione: la massa si è spostata di una quantità δL rispetto alla sua posizione di riposo.

La Fig. 1 mostra una massa M collegata ad una molla fissata ad un estremo. Essendo il moto oscillatorio si dovrà avere una forza di richiamo che riporti il sistema verso la sua posizione di equilibrio, ovvero quella posizione in cui il peso della massa bilancia la forza di richiamo F della molla. In figura k si riferisce alla costante di richiamo della molla, misurata in [Newton/metro]. Operiamo quindi il bilancio delle forze che agiscono sul sistema:

$$F = -kx = Ma \quad (1)$$

dove a è l'accelerazione delle particelle e $x(t) = A \sin(\omega t + \varphi)$ è il segnale acustico generato dal moto armonico, ovvero una funzione sinusoidale di pulsazione (o frequenza angolare) $\omega = 2\pi f$ [radianti/secondo], frequenza $f = 1/T$ [Hertz], [cicli/secondo], ampiezza A [dB], [Pascal], [Watt/mq] e fase iniziale φ [radianti].

Torneremo a breve nella descrizione in dettaglio di ciascuno di questi parametri.

Plottando l'elongazione del sistema massa-molla, o in modo del tutto equivalente l'ampiezza dell'oscillazione delle particelle in funzione del tempo, si ottiene un grafico del tipo di Fig. 2.

Graficamente si osserva che:

- al crescere dell'ampiezza corrisponde un'espansione dell'onda in verticale;

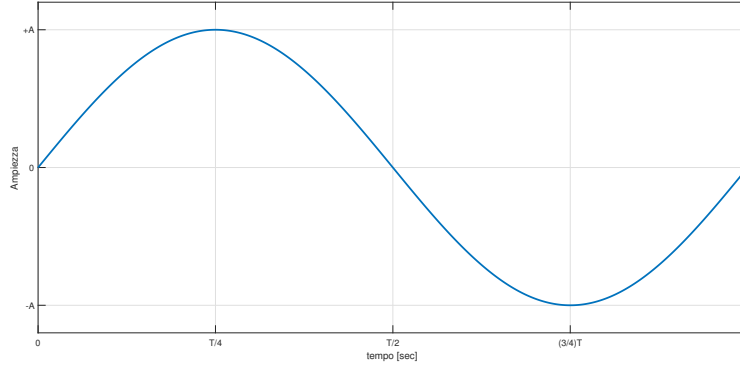


Figure 2: Sinusoide

- al crescere della frequenza si osserva una compressione dell'onda in orizzontale (e in modo reciproco al crescere del periodo corrisponde un'espansione dell'onda in orizzontale);
- al variare della fase si ha una traslazione dell'onda verso sinistra per fase crescente e verso destra per fase decrescente.

Le proprietà elastiche e la massa del mezzo elastico stabiliscono la velocità con cui la perturbazione si trasmette e la quantità di energia meccanica trasferita dalla sorgente sonora nell'unità di tempo. Calcoliamo la velocità e l'accelerazione del segnale rappresentato dal moto armonico derivando¹ il segnale rispettivamente una volta e due volte rispetto al tempo t :

$$v(t) = \frac{dx(t)}{dt} = A\omega \cos(\omega t + \varphi) \quad (2)$$

$$a(t) = \frac{d^2x(t)}{dt^2} = -A\omega^2 \sin(\omega t + \varphi) \quad (3)$$

Dato che $a(t) = \frac{-kx(t)}{M}$, si ha che $\omega = \sqrt{\frac{k}{M}}$, quindi maggiore è k , ovvero si ha una molla più rigida, e maggiore sarà la pulsazione ω . Questo esempio ci suggerisce che volendo creare un modello - nel processo di sintesi - che tenga conto del legame rigidità-frequenza, si può agire sul parametro k del sistema equivalente massa-molla.

¹Per chi non la conoscesse, in analisi matematica la derivata di una funzione calcolata nel punto x_0 , definita come il limite del rapporto incrementale per l'incremento Δh tendente a 0, ovvero:

$$f'(x_0) = \lim_{\Delta h \rightarrow 0} \frac{f(x_0 + \Delta h) - f(x_0)}{\Delta h},$$

altro non è che una misura della variazione del valore assunto dalla funzione in un punto x_0 rispetto ad una variabile, che nel nostro caso è il tempo, se ci spostiamo da questo di un incremento Δh molto piccolo.

Corrispondentemente al fenomeno fisico, il sistema uditivo umano risponde con un'eccitazione dell'orecchio e degli organi al suo interno [5].

- L'ampiezza viene percepita come *intensità* e dipende dalla quantità di energia trasportata dall'onda acustica ovvero $E = \frac{1}{2}kx^2 + \frac{1}{2}Mv^2$, dove il primo addendo rappresenta l'energia potenziale elastica della deformazione della molla rispetto alla posizione di equilibrio e il secondo l'energia cinetica della massa in movimento;
- La frequenza viene percepita come *altezza* (pitch) del suono e dipende dalla velocità di oscillazione della particella. In altre parole indica quante oscillazioni (o quanti cicli) la particella riesce a compiere in un secondo. Si ricordi che lo spettro di frequenze udibili per gli esseri umani è compreso mediamente tra 20 Hz e 20 kHz.
- L'insieme dei singoli contributi di frequenze che compongono un suono formano il suo spettro armonico. Per ora anticipiamo che in natura un suono reale non è mai formato dalla sola frequenza fondamentale che ne determina il pitch (soltanto una sinusoide "matematica" è pura in questo senso), ma a questa si sommano un certo numero di contributi a frequenze più alte (dette armoniche², se multiple della frequenza fondamentale). Lo spettro armonico è percepito come *timbro* del suono.
- La fase si misura come un angolo e rappresenta la posizione del moto nell'istante considerato come iniziale. Dal punto di vista percettivo la fase interviene sulla localizzazione del suono (differenza di fase $\Delta\varphi$) e indirettamente se ne possono vedere gli effetti nel timbro del suono percepito³. L'effetto sul timbro è dovuto al fatto che nello spettro armonico possono aversi interferenze costruttive/distruttive tra i vari contributi che si sovrappongono nella composizione dello spettro.

La tabella di seguito riassume quanto detto:

PARAMETRI FISICI	PARAMETRI PERCEPITI
Ampiezza	Intensità (Volume)
Frequenza	Altezza
Spettro Armonico	Timbro
Fase	Localizzazione, Spettro

²Artificialmente si possono aggiungere contributi a frequenze più basse della fondamentale.

³Il discorso sulla localizzazione è più complesso e coinvolge ovviamente anche l'intensità del suono ricevuto alle orecchie e altro.

3 Teorema di Fourier

Se volessimo far suonare un'onda sinusoidale, questa assomiglierebbe ad un fischio a frequenza $f = \frac{\omega}{2\pi}$. Il mondo che ci circonda emette suoni ben più complessi di una semplice sinusoide e anche quello che si avvicina maggiormente al suono di un fischio non sarà mai una sinusoide pura. Infatti la stessa idea di un suono puramente sinusoidale è un'astrazione matematica che nella realtà è difficoltoso realizzare. Anche riuscendo a suonare una sinusoide via software, ad esempio in Matlab o Octave, il suono emesso da un sistema di diffusione acustica il più lineare possibile avrà una minima percentuale di distorsione armonica aggiunta, ovvero contributi a frequenze multiple della fondamentale che colorano il segnale. Le non linearità sono termini additivi non voluti, solitamente; nel caso di sorgenti reali invece il fatto che i suoni non siano sinusoidi pure è ciò che caratterizza una voce da un'altra e che ci permette di distinguere, ad esempio, un violino da un flauto. A questo proposito il matematico e fisico Jean Baptiste Fourier ci insegna che un qualsiasi segnale periodico (e aperiodico per estensione del teorema) può essere ottenuto mediante la somma di infinite oscillazioni sinusoidali di ampiezza, frequenza e fase opportune. In termini matematici possiamo esprimere questa somma di sinusoidi con uno sviluppo in serie (SVILUPPO IN SERIE DI FOURIER) [6]:

$$x(t) = A_{DC} + 2 \sum_{k=1}^{\infty} A_k \sin(2\pi k f_0 t + \varphi_k) \quad (4)$$

con $k \geq 1$. Il termine A_{DC} rappresenta l'ampiezza in continua (frequenza nulla). I termini della sommatoria si riferiscono ai contributi della frequenza fondamentale f_0 e delle frequenze sue multiple, cioè alle frequenze $k f_0$, per k che va da 1 a infinito. Il "2" che moltiplica la sommatoria sta ad indicare che si devono considerare sia le frequenze positive che quelle negative. Il fatto che esistano frequenze negative può sembrare inizialmente qualcosa di strano, ma si capirà il significato fisico di tali frequenze quando parleremo di modulazioni e si dimostrerà che i conti non tornano se non le si prendono in esame. Comunque uno strumento come un analizzatore di spettro solitamente mostra solo le frequenze positive.

La composizione dei contributi frequenziali che formano un determinato suono è detta SPETTRO del segnale emesso. Abbiamo già detto che il *timbro* è quella qualità del suono che ci permette di distinguere uno strumento musicale da un altro e questo è dovuto proprio alle diverse composizioni spettrali delle sorgenti acustiche. Lo spettro dei diversi strumenti musicali (o anche di uno stesso strumento suonato in condizioni diverse di temperatura, umidità, etc.) differisce per la diversa distribuzione di energia fra la nota fondamentale e le armoniche superiori e/o frequenze non multiple della fondamentale, chiamate *parziali non armoniche*.

Rappresentare un segnale nel dominio del tempo o in quello della frequenza è del tutto equivalente, perché possiamo passare da uno all'altro dominio attraverso la *trasformata* e l'*antitrasformata di Fourier*. La trasformata di Fourier

è un'operazione che permette proprio di scomporre un segnale in una somma infinita di sinusoidi. Attraverso l'operazione inversa di antitrasformata ricombiniamo insieme il segnale decomposto.

$$x(t) \Leftrightarrow X(f) \quad (5)$$

Possiamo quindi dividere la teoria di Fourier in due fasi e per un segnale periodico si ha:

$$\text{ANALISI} \quad X_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-j2\pi k f_0 t} dt$$

$$\text{SINTESI} \quad x(t) = \sum_{k=-\infty}^{\infty} X_k e^{j2\pi k f_0 t}$$

dove $T_0 = \frac{1}{f_0}$ è il periodo del segnale.

Per chi non conoscesse l'operazione di integrale \int , questo altro non è che una somma continua su passi infinitesimi, cioè infinitamente piccoli.

Nota: È importante sottolineare una differenza tra forma d'onda (rappresentazione nel tempo) e spettro (rappresentazione in frequenza). La prima è una funzione reale, ovvero opera su valori reali (numeri reali) e li trasforma in altri valori reali. Lo spettro invece è una funzione complessa, ovvero scomponibile in uno spettro di fase e uno spettro di ampiezza. Possiamo infatti trasportare entrambe le informazioni relative alla fase e all'ampiezza in un unico numero complesso del tipo (in forma polare) $\rho e^{j\theta}$ con ρ modulo del numero complesso (cioè l'ampiezza) e θ fase del numero complesso. La j si chiama *unità immaginaria* ed è tale che $\sqrt{-1} = j$.

Per ulteriori approfondimenti sul campo complesso si suggerisce di consultare un libro di analisi matematica o Wikipedia.

Come esempio rappresentiamo una sinusoidale a frequenza 1 kHz, ampiezza 1 e fase iniziale nulla nel dominio del tempo e nel dominio della frequenza. Il codice che genera i grafici (Fig. 3 e Fig. 4) è stato implementato in ambiente MATLAB. Una descrizione autoesplicativa del codice MATLAB è fornita dai commenti in verde preceduti dal simbolo %.

```
Fs = 44100;      % Frequenza di campionamento
T = 1/Fs;        % Periodo di campionamento
L = 44100;       % Lunghezza del segnale in campioni
t = (0:L-1)*T;   % Vettore tempo
```

```

% Segnale sinusoidale a 1000 Hz.
x = 1*sin(2*pi*1000*t);

% Plot nel dominio del tempo
plot(t, x);
title('Sinusoide a 1000 Hz')
xlabel('t [sec]')
ylabel('A')

```

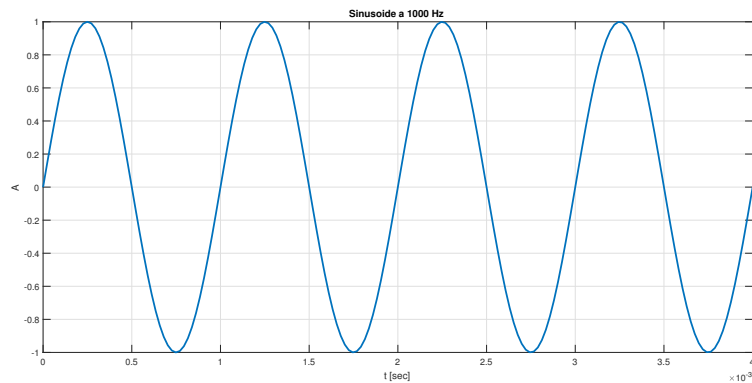


Figure 3: Onda sinusoidale a 1000 Hz.

Attraverso un'operazione di *Fast Fourier Transform* (FFT) rappresentiamo la stessa sinusoide nel dominio della frequenza. L'FFT è un algoritmo che restituisce la trasformata di Fourier discreta per un segnale campionato (discreto). Continuando con il codice, generiamo lo spettro di ampiezza e di fase:

```

X = fft(x)/L; % FFT del segnale x (con ampiezze normalizzate)
X = fftshift(X);

% Dato che l'FFT è periodica con fftshift si centra
% attorno allo zero il vettore X ottenuto
f = -(Fs/2):(Fs/2)-1); % Vettore frequenze

% Plot spettro di ampiezza di x(t)
plot(f, abs(X))
title('Spettro di Ampiezza di x(t)')
xlabel('Frequenza [Hz]')
ylabel('|X(f)|')

```

Come si vede in figura lo spettro di ampiezza di un'onda sinusoidale ideale a 1000 Hz è composto da due impulsi centrati in $-f_0 = -1000$ Hz e $f_0 = 1000$ Hz, ovvero:

$$X(f) = \frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0) \quad (6)$$

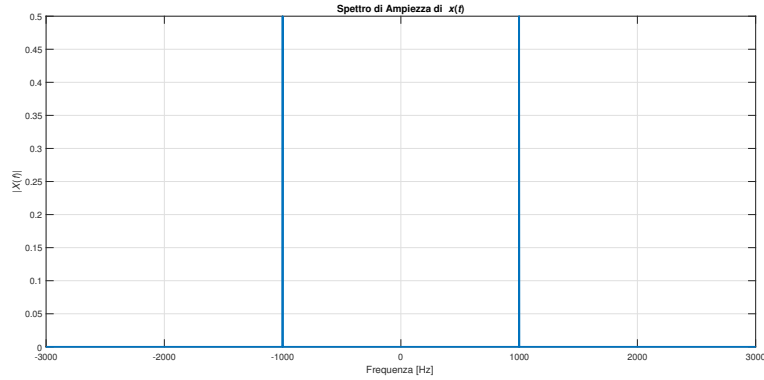


Figure 4: Spettro di un'onda sinusoidale a 1000 Hz.

dove si intende con $\delta(f - f_0)$ l'impulso matematico di Dirac centrato in f_0 (e con $\delta(f + f_0)$ quello centrato in $-f_0$).

Nel prossimo paragrafo analizzeremo alcune delle forme d'onda fondamentali che saranno utilizzate come blocchi base della sintesi additiva e sottrattiva.

Nota: Lo spettro di ampiezza di un segnale periodico è discreto, cioè è formato da righe. Infine, dato che abbiamo posto la fase iniziale della senoide a zero, lo spettro di fase è nullo su tutto il range delle frequenze.

Il teorema di Fourier si può estendere al caso di segnali aperiodici. Si opera una periodicizzazione del periodo T_0 facendolo tendere all'infinito ($T_0 \rightarrow \infty$), ovvero si costruisce un segnale periodico $x_p(t)$ tale che il segnale aperiodico che vogliamo analizzare sia $x(t) = \lim_{T_0 \rightarrow \infty} x_p(t)$.

Nota: In questo modo, ricostruendo il segnale con l'antitrasformata di Fourier, si ha:

$$x(t) = \int_{-\infty}^{+\infty} X(f) e^{j2\pi k f_0 t} df \quad (7)$$

e dove prima avevo una sommatoria (quindi una somma discreta) adesso ho un integrale. Questo mi suggerisce che facendo tendere T_0 all'infinito si riduce la distanza tra armoniche consecutive (in quanto f_0 diminuisce) e lo spettro di ampiezza si infittisce fino a diventare continuo ($k f_0 - (k - 1) f_0 = f_0$).

Riassumendo:

Segnali periodici	\rightarrow	Spettro a righe
Segnali aperiodici	\rightarrow	Spettro continuo

A titolo di esempio mostriamo gli spettri di ampiezza del suono di una corda di chitarra pizzicata e un colpo di cassa di batteria (Fig. 5 e Fig. 6).

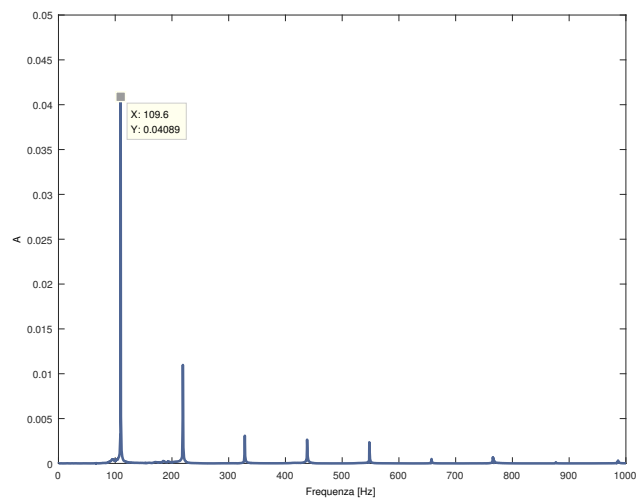


Figure 5: (Spettro di ampiezza) Corda di chitarra acustica pizzicata (A2: 110 Hz).

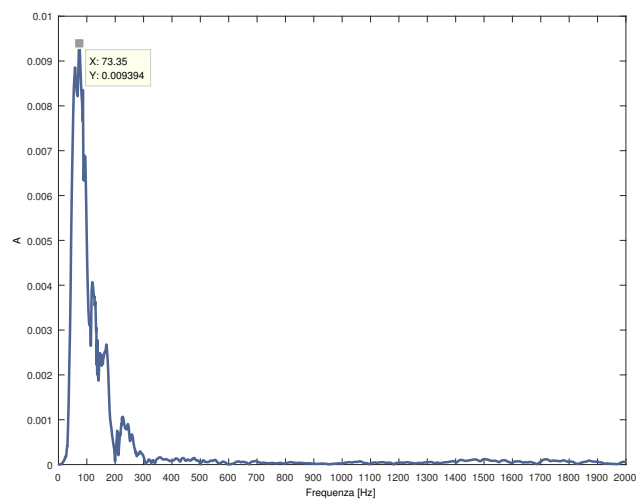


Figure 6: (Spettro di ampiezza) Kick Drum.

4 Forme d'onda di base

Vediamo ora alcune forme d'onda interessanti nel campo della sintesi del suono, perché le ritroveremo spesso nei blocchi base che compongono un sintetizzatore [6], [4]. Pur essendo molto semplici, esse sono la composizione di diverse sinusoidi di ampiezze e frequenze opportune, secondo quanto detto nel Teorema di Fourier. La somma di due o più onde è comunemente chiamata *sintesi additiva*. In particolare ci soffermeremo su onde triangolari, onde a dente di sega e onde quadre.

Premessa importante: indichiamo con f_0 la frequenza fondamentale a cui corrisponde l'ordine $k = 1$ e con $f_2, f_3, \dots, f_k, \dots$ tutte le armoniche a cui corrisponde l'ordine $k = 2, 3, \dots, k, \dots$

Se prendo una senoide a 220 Hz (A3) e sommo ad essa le sinusoidi a 660 Hz e 1100 Hz (ovvero la 3^a e la 5^a armonica: $\frac{220}{660} = \frac{1}{3}$ e $\frac{220}{1100} = \frac{1}{5}$), opportunamente scalate in ampiezza rispettivamente di $1/3$ e $1/5$ dell'ampiezza della senoide a 220 Hz, ottengo l'onda mostrata in Fig. 7.

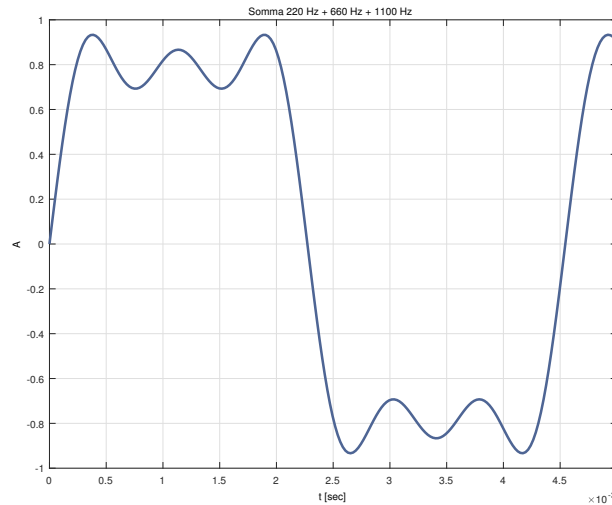


Figure 7: Somma delle sinusoidi a 220 Hz, 660 Hz (3^a armonica) e 1100 Hz (5^a armonica).

Se adesso aggiungo anche la 7^a armonica scalata in ampiezza di $1/7$ dell'ampiezza della senoide a 220 Hz ottengo l'onda in Fig. 8.

In definitiva se continuo ad aggiungere armoniche dispari ottengo una forma d'onda che si avvicina sempre di più all'onda quadra (Fig. 9).

Per ottenere altre forme d'onda si dovranno sommare tra loro armoniche differenti con ampiezze opportune rispetto alla frequenza fondamentale.

Si osservi dalla Fig. 9, in cui è ben visibile, il fenomeno per cui in prossimità delle discontinuità della funzione (cioè dove si ha il passaggio brusco dal valore massimo al valore minimo e viceversa) si ha una sovralongazione del valore

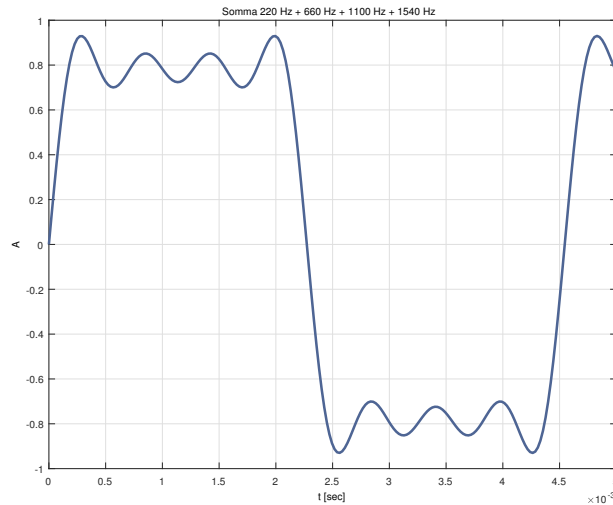


Figure 8: Somma delle sinusoidi a 220 Hz, 660 Hz (3^a armonica), 1100 Hz (5^a armonica) e 1540 Hz (7^a armonica).

della funzione stessa. Data una funzione periodica, la sua rappresentazione tramite serie di Fourier ha infiniti termini. Quando si ricostruisce il segnale, se si tronca la serie ad un numero finito di termini e se la funzione di partenza presenta discontinuità di prima specie (la funzione presenta un salto finito nel punto di discontinuità), come nel caso dell'onda quadra, si ha il fenomeno, chiamato *Fenomeno di Gibbs*, che genera queste sovraelongazioni. Aumentando il numero di componenti (le armoniche) nella somma, il valore di picco della sovraelongazione rimane costante e le oscillazioni sovraelongate si distribuiscono in prossimità delle discontinuità.

4.1 Onda quadra (square wave)

Calcoliamo ora i coefficienti di Fourier di un'onda quadra di ampiezza A , frequenza f_0 e periodo T_0 (non è lo scopo dei corsi a cui sono rivolti questi appunti essere in grado di risolvere integrali complicati, ma riportiamo il calcolo intero per completezza e per interpretarne il risultato):

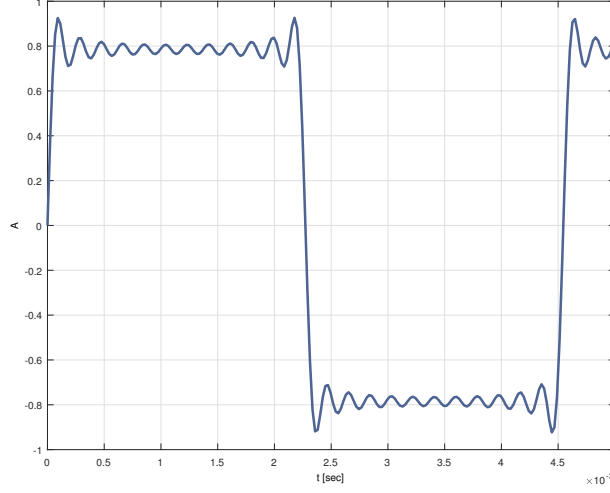


Figure 9: Aggiungendo armoniche dispari mi avvicino sempre di più ad un'onda quadra.

$$\begin{aligned}
 X_k &= -\frac{2j}{T_0} \int_0^{T_0/2} x(t) \sin(2\pi k f_0 t) dt = -j \frac{2A}{T_0} \int_0^{T_0/2} \sin(2\pi k f_0 t) dt \\
 &= j \frac{2A}{2\pi k f_0 T_0} \cos(2\pi k f_0 t) \Big|_{t=0}^{t=T_0/2} = j \frac{A}{\pi k} [\cos(\pi k) - 1] \\
 &= j \frac{A}{\pi k} [(-1)^k - 1].
 \end{aligned} \tag{8}$$

Questo calcolo ci serve al fine di capire che si ha:

$$X_k = \begin{cases} 0 & k \text{ pari} \\ \frac{2A}{j\pi k} & k \text{ dispari} \end{cases}$$

k è l'ordine del coefficiente, o ugualmente è l'armonica k -esima che vogliamo considerare.

Si ottiene lo sviluppo in serie di Fourier di $x(t)$:

$$x(t) = 2j \sum_{k=1}^{\infty} X_k \sin(2\pi k f_0 t) = \frac{4A}{\pi} \sum_{k=1}^{\infty} \frac{\sin(2\pi k f_0 t)}{k} \tag{9}$$

La sequenza $\{X_k\}$ dei coefficienti dell'onda quadra è immaginaria pura (notare la presenza dell'unità immaginaria j : la frazione $\frac{2A}{j\pi k}$ ha solo parte immaginaria) e dispari (i coefficienti di ordine pari sono tutti nulli). Si osserva dalla

sequenza dei coefficienti che l'ampiezza delle righe spettrali tende a zero come $1/k$ quando $k \rightarrow \infty$. Risulta quindi che le componenti armoniche di ordine elevato hanno ampiezza ancora relativamente grande rispetto a quella delle prime armoniche. Ascoltando un'onda quadra infatti si ha la sensazione di un suono molto ricco e "massiccio" nel timbro.

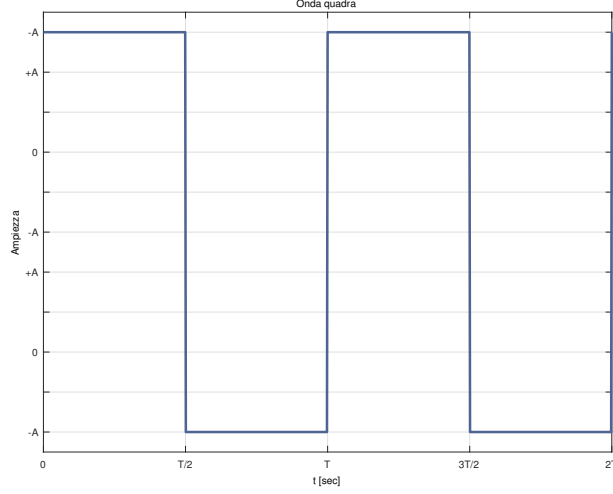


Figure 10: Onda quadra nel dominio del tempo.

4.2 Onda triangolare (triangular wave)

Vediamo come si comportano diversamente le componenti spettrali di un'onda triangolare di ampiezza A e periodo T_0 . La funzione che descrive un'onda triangolare è la seguente:

$$x(t) = A \left(1 - \frac{4t}{T_0} \right) \quad (10)$$

per $0 \leq t \leq T_0$.

Calcoliamo i coefficienti di Fourier:

$$\begin{aligned} X_k &= \frac{2}{T_0} \int_0^{T_0/2} A \left(1 - \frac{4t}{T_0} \right) \cos(2\pi k f_0 T_0) dt \\ &= \frac{2A}{T_0} \int_0^{T_0/2} \cos(2\pi k f_0 T_0) dt - \frac{8A}{T_0} \int_0^{T_0/2} \frac{t}{T_0} \cos(2\pi k f_0 T_0) dt \\ &= 0 - \frac{8A}{T_0} \frac{T_0}{(2\pi k)^2} [(-1)^k - 1]. \end{aligned} \quad (11)$$

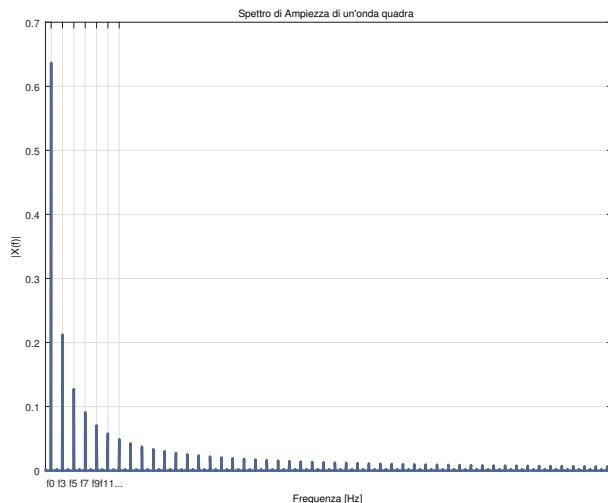


Figure 11: Onda quadra nel dominio della frequenza. Le armoniche sono tutte dispari e decrescono come $1/k$.

Si ottiene quindi:

$$X_k = \frac{2A}{(\pi k)^2} [1 - (-1)^k] = \begin{cases} 0 & k \text{ pari} \\ \frac{4A}{(\pi k)^2} & k \text{ dispari} \end{cases}$$

e sviluppando il segnale ad onda triangolare in serie di Fourier si ha:

$$x(t) = \frac{8A}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^{(k-1)/2}}{k^2} \sin(2\pi k f_0 t) \quad (12)$$

Anche in questo caso tutti i coefficienti di ordine pari sono nulli e l'onda ha coefficienti reali (non è presente l'unità j infatti). Bisogna notare però che le armoniche presenti nell'onda triangolare tendono a zero molto più velocemente rispetto al caso dell'onda quadra. Esse infatti per $k \rightarrow \infty$ tendono a zero come $1/k^2$. Il suono di un'onda triangolare appare quindi più "morbido" e meno ricco. Questa è una conseguenza del fatto che l'onda triangolare non presenta discontinuità (ma la sua derivata sì), a differenza dell'onda quadra, le cui brusche variazioni temporali del valore del segnale comportano la presenza di armoniche di ordine più elevato.

Introduciamo brevemente altri tipi di onde fondamentali nella sintesi del suono.

4.3 Onda a dente di sega (sawtooth wave)

L'onda a dente di sega di periodo T_0 è descritta dalla funzione:

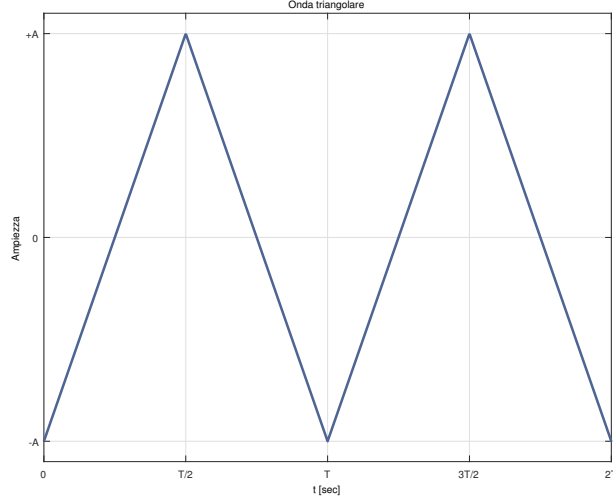


Figure 12: Onda triangolare nel dominio del tempo.

$$x(t) = \left(\frac{t}{T_0} - \left\lfloor \frac{t}{T_0} - \frac{1}{2} \right\rfloor \right) \quad (13)$$

dove il simbolo $\lfloor \cdot \rfloor$ denota la parte intera inferiore. Esprimendo il segnale generico in serie di Fourier come:

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(2\pi k f_0 t) + b_k \sin(2\pi k f_0 t)] \quad (14)$$

possiamo calcolare i coefficienti:

$$\begin{cases} a_k = \frac{2}{T_0} \int_0^{T_0} x(t) \cos(2\pi k f_0 t) dt \\ b_k = \frac{2}{T_0} \int_0^{T_0} x(t) \sin(2\pi k f_0 t) dt \end{cases} \quad (15)$$

con $k = 0, 1, 2, \dots$ e risulta in definitiva dopo un po' di calcoli:

$$\begin{cases} a_k = 0 \\ b_k = -\frac{1}{\pi k} \end{cases} \quad (16)$$

Si ottiene quindi lo sviluppo in serie di Fourier:

$$x(t) = \frac{1}{2} - \sum_{k=1}^{\infty} \frac{1}{\pi k} \sin(2\pi k t) \quad (17)$$

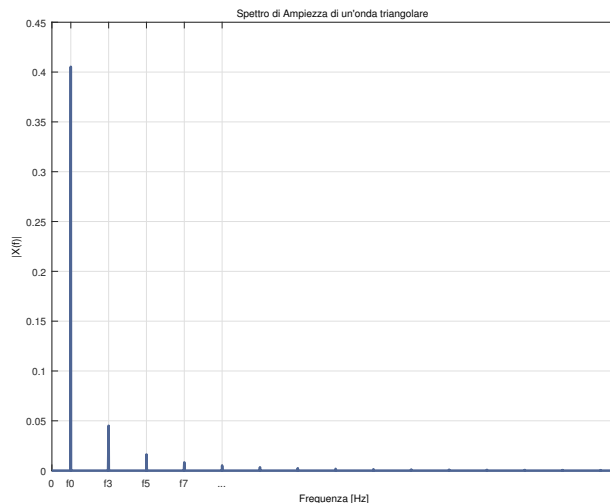


Figure 13: Onda triangolare nel dominio della frequenza.

Lo spettro armonico comprende sia armoniche pari sia armoniche dispari, che tendono a zero come $1/k$ per $k \rightarrow \infty$.

4.4 Onda a treno di impulsi (pulse wave)

L'onda a treno di impulsi deriva dall'onda quadra (l'onda quadra è un particolare treno di impulsi). La caratteristica di questo tipo di onda è la larghezza dell'impulso (*duty cycle*), che in alcuni sintetizzatori si può controllare. La larghezza di impulso è la percentuale di tempo in cui il segnale si trova al suo valore "alto". Ad esempio un'onda quadra avrà una larghezza di impulso del 50%, ovvero per metà periodo il segnale sarà alto e per l'altra metà sarà basso. Anche per il treno di impulsi le armoniche, sia dispari che pari, tendono a zero come $1/k$ per $k \rightarrow \infty$. La peculiarità sta nel fatto che un'onda di questo tipo con una larghezza di impulso proprio pari a $1/k$ mancherà delle armoniche multiple di k .

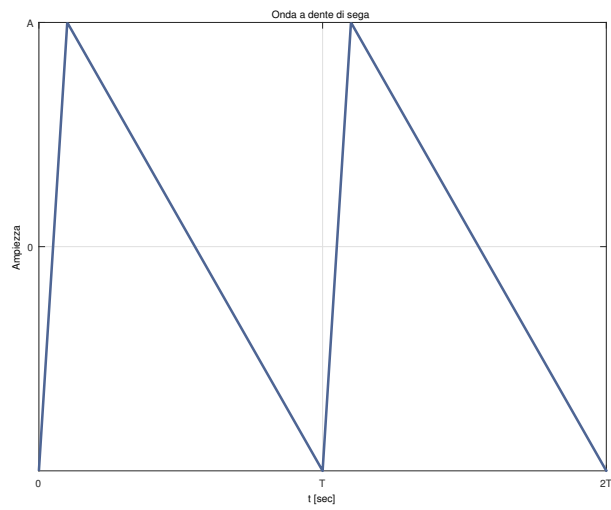


Figure 14: Onda a dente di sega nel dominio del tempo.

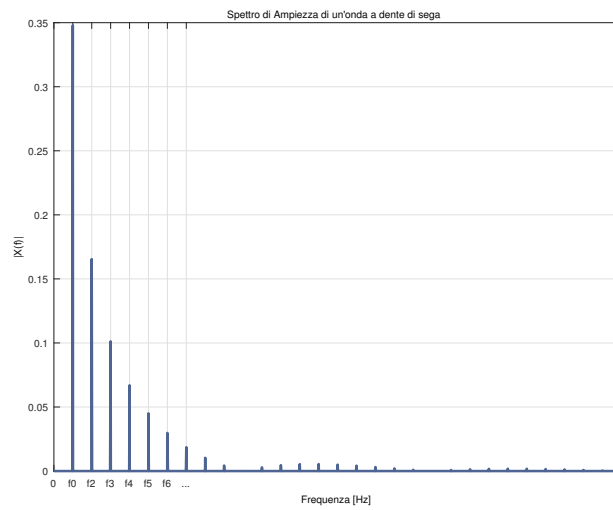


Figure 15: Onda a dente di sega nel dominio della frequenza.

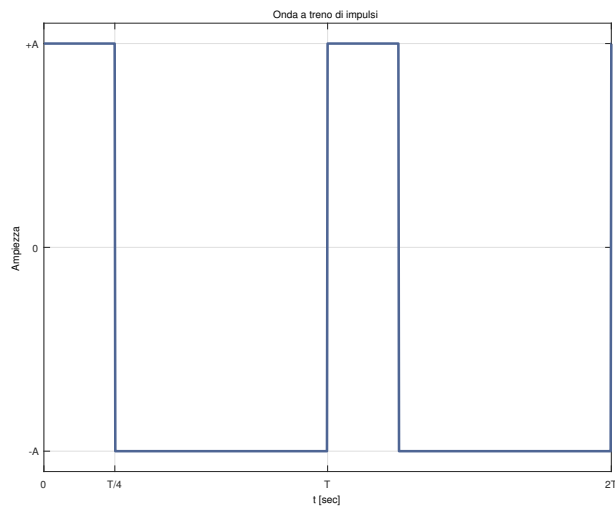


Figure 16: Onda a treno di impulsi con larghezza di impulso del 25% nel dominio del tempo.

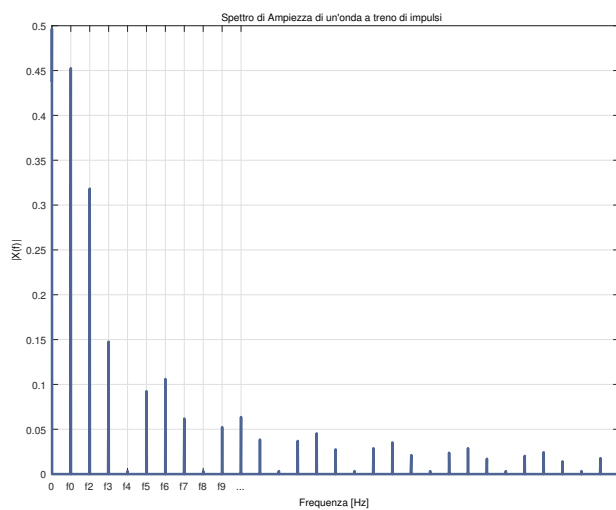


Figure 17: Onda a treno di impulsi con larghezza di impulso del 25% nel dominio della frequenza. Come si può notare mancano le armoniche di ordine multiplo di $k = 4$ ($25\% = 1/4$), quindi f_4 , f_8 , f_{12} , ...

5 Rumore

Parlando di *rumore* è bene darne una definizione generale prima di introdurlo nella sintesi del suono. Generalmente il rumore è qualcosa di non voluto che si somma ad un segnale che stiamo elaborando o trasmettendo, compromettendone la comprensione a seconda della potenza che il rumore presenta rispetto a quella del segnale. Esistono diversi tipi di rumore. Operiamo però preliminarmente una distinzione tra segnali disturbanti (tipo voci di ambiente che si sovrappongono ad un suono su cui ci focalizziamo) e quello che viene chiamato comunemente "rumore" (*noise*, tipicamente indesiderato) che presenta delle caratteristiche spettrali tipiche (ad esempio il *rumore termico*). Proprio per le sue caratteristiche spettrali interessanti una sorgente di rumore può diventare un utilissimo blocco del sintetizzatore additivo. Si nota infatti che in realtà molte cose in natura hanno un suono che assomiglia al rumore. Se volessimo sintetizzare il suono emesso dal mare, useremmo proprio una sorgente di rumore. Vediamo ora quali sono queste caratteristiche e che tipi di sorgenti di rumore possiamo trovare in un sintetizzatore.

5.1 Rumore bianco (White noise)

Il rumore bianco, detto con parole davvero difficili, è un *processo aleatorio ergodico a media nulla con spettro di densità di potenza costante*. Tradotto in parole poco più semplici, si ha una famiglia di variabili casuali (aleatorie, appunto) che assumono determinati valori secondo una famiglia di funzioni (processo). A ciascuna di queste funzioni è associato un valore appartenente ad uno spazio di valori campione. Istanziando il processo (casuale) sto di fatto estraendo una di queste funzioni e di conseguenza ottengo per la variabile aleatoria un certo valore. Ergodico significa che le medie statistiche coincidono con quelle temporali, ma qui stiamo uscendo dalla nostra materia. A noi che studiamo Music Technology non importa poi moltissimo avere queste nozioni di teoria dei segnali così avanzate, ma ci interessa capire come si presenta il rumore bianco alle nostre orecchie e alla nostra strumentazione.

Ciò che è più importante per noi è sapere che:

- **il rumore bianco ha valore medio nullo:** facendo la media di tutti i valori assunti dalla variabile aleatoria ottengo zero;
- **lo spettro di densità di potenza è costante:** il rumore bianco è caratterizzato da uno spettro armonico continuo, comprendente cioè tutte le frequenze nel range di interesse (20 Hz - 20 kHz) e di ampiezza casuale, e se ne calcolo lo spettro di densità di potenza⁴, definito come:

$$S_N(f) = \lim_{\Delta T \rightarrow \infty} \frac{|N_{\Delta T}(f)|^2}{\Delta T} \quad [\text{W/Hz}]$$

⁴Attenzione a non confondere la densità di potenza con la potenza!! La potenza è l'integrale: $P_N(f) = \int_{-\infty}^{+\infty} S_N(f) df$ [W] e assume valore infinito se la banda non è limitata.

dove $N_{\Delta}(f)$ è la trasformata di Fourier del rumore considerato in un certo intervallo ΔT , si ha che

$$S_N(f) = N_0 = \text{costante} \quad \forall f$$

Il nome "bianco" deriva dall'analogia con la luce bianca, la quale contiene tutte le lunghezze d'onda dello spettro visibile.

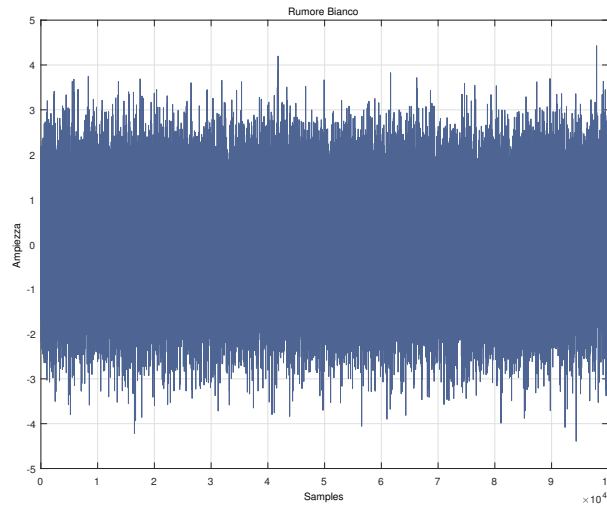


Figure 18: Campioni di rumore bianco.

5.2 Rumore rosa (pink noise)

Mentre il rumore bianco presenta la stessa distribuzione di potenza per tutte le frequenze, il rumore rosa ha la stessa distribuzione di potenza per ottave. La densità spettrale di potenza nel rumore rosa scende di 3 dB per ottava, quindi è inversamente proporzionale alla frequenza.

5.3 Rumore blu e altri "colori"

In modo complementare al rumore rosa, nel rumore blu la densità spettrale di potenza nel rumore rosa cresce di 3 dB per ottava.

Si definiscono anche i rumori rosso (-6 dB per ottava), violetto (+6 dB per ottava), grigio (si ottiene dal rumore bianco applicando un filtro pari all'inverso delle curve di sensibilità di Fletcher-Munson).

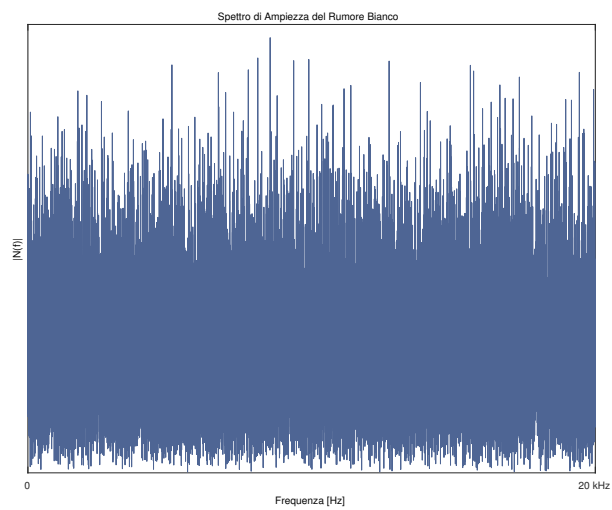


Figure 19: Spettro di ampiezza del rumore bianco.

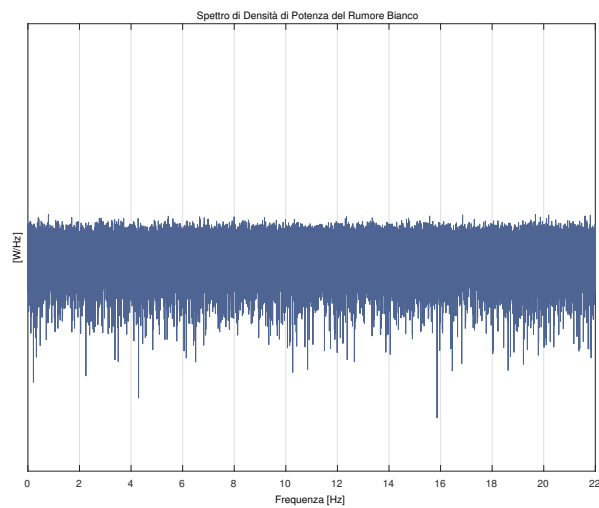


Figure 20: Spettro di densità di potenza del rumore bianco.

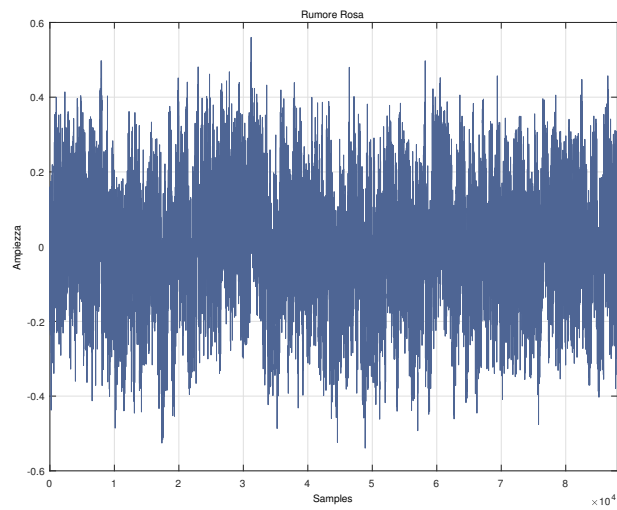


Figure 21: Campioni di rumore rosa.

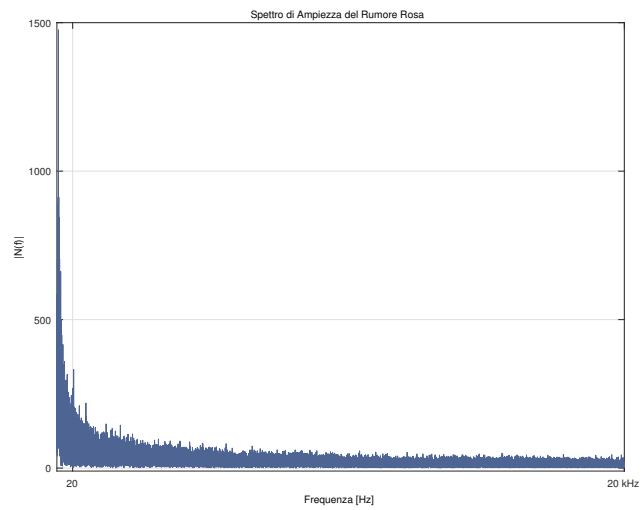


Figure 22: Spettro di ampiezza del rumore rosa.

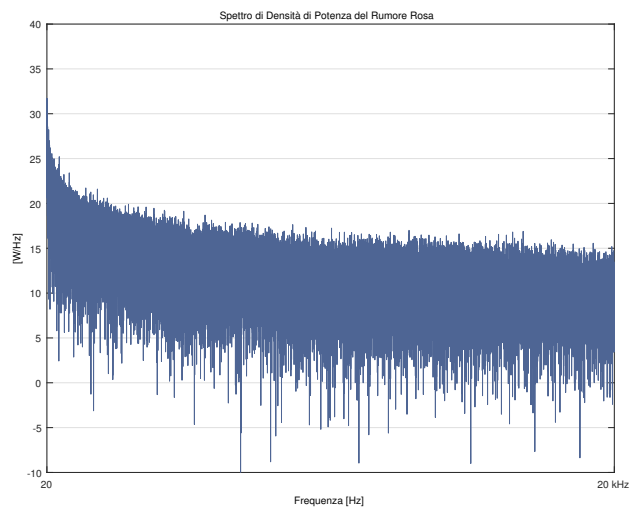


Figure 23: Spettro di densità di potenza del rumore rosa.

6 Involuppi

Il contenuto spettrale di un'onda acustica è di fondamentale importanza, ma non è tutto. Un suono può subire un'evoluzione nel tempo. Questa evoluzione può coinvolgere ad esempio l'ampiezza o il contenuto spettrale stesso. Nella descrizione di un suono possiamo modellare queste evoluzioni attraverso degli involuppi. Per cominciare, per capire cosa si intende con "involuppo", parliamo di quello che viene chiamato *involuppo di ampiezza*, ovvero come l'ampiezza di un suono evolve nel tempo. È quasi impossibile trovare in natura suoni che attaccano e decadono istantaneamente, invece si avranno più verosimilmente attacchi e decadimenti più o meno rapidi. Generalizzando anche al caso di strumenti sintetizzati e non solo suoni che si trovano in natura, si può descrivere l'involuppo di un suono attraverso alcuni *stadi* o *fasi* di cui esso è formato:

- **Attack:** intervallo di tempo in cui l'ampiezza del suono passa da zero al valore massimo;
- **Decay:** intervallo di tempo in cui l'ampiezza del suono scende dal suo picco massimo al livello del sustain;
- **Sustain:** intervallo di tempo in cui l'ampiezza del suono rimane su un livello costante (es. tenendo premuto un tasto di un organo, il livello del suono rimane costante finché il tasto rimane premuto);
- **Release:** intervallo di tempo in cui l'ampiezza del suono cade dal livello del sustain a zero quando il suono è terminato (es. quanto viene rilasciato il tasto della nota suonata).

Sui sintetizzatori troveremo spesso il controllo di ADSR, di cui parleremo in seguito, che servirà proprio a regolare l'involuppo del suono che vogliamo sintetizzare.

Esistono ovviamente delle varianti a questo schema. Se ci troviamo di fronte ad un AHDSR, la "H" sta per *hold*, ovvero un intervallo di tempo in cui l'ampiezza del suono è mantenuta al livello di picco massimo. Lasciamo ai prossimi capitoli gli approfondimenti su questo argomento.

Prendiamo in considerazione come esempio il suono di uno strumento acustico come una chitarra classica. Suonando una corda notiamo come non solo l'ampiezza vari nel tempo, ma anche il contenuto spettrale si trasforma e se inizialmente sembra più ricco, con il passare dei secondi il suono perde una parte delle sue armoniche più alte. Riuscire a rendere questo effetto nella sintesi del suono significa essere in grado di avvicinarsi alla realtà dello strumento acustico. Per realizzare questo effetto, possiamo ad esempio controllare l'azione di un filtro (in questo esempio ci farebbe comodo un filtro passabasso) attraverso un ADSR.

Esiste uno strumento che ci permette di visualizzare l'evoluzione spettrale nel tempo: lo *spettrogramma*, che ha sull'asse delle ascisse il tempo e sull'asse delle ordinate la frequenza.

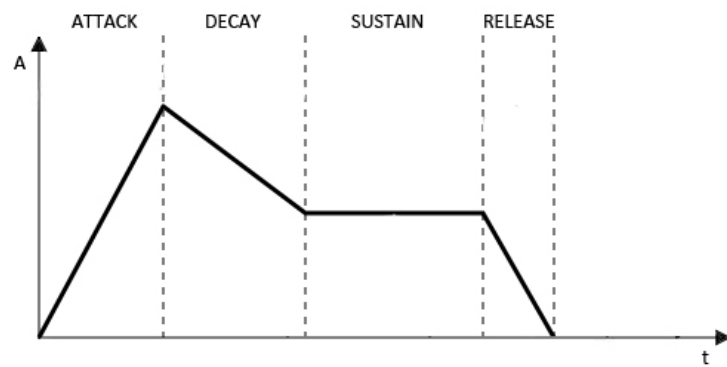


Figure 24: ADSR

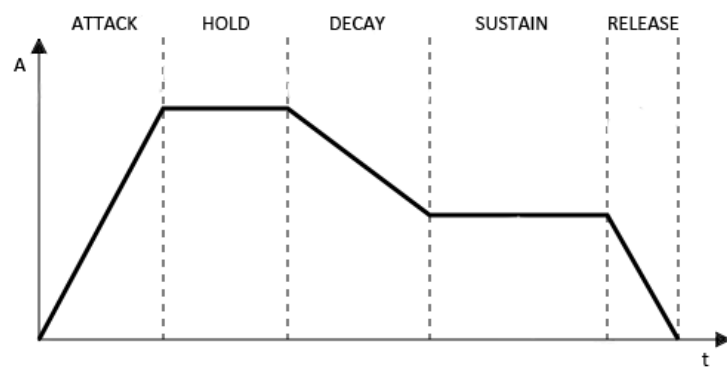


Figure 25: AHDSR

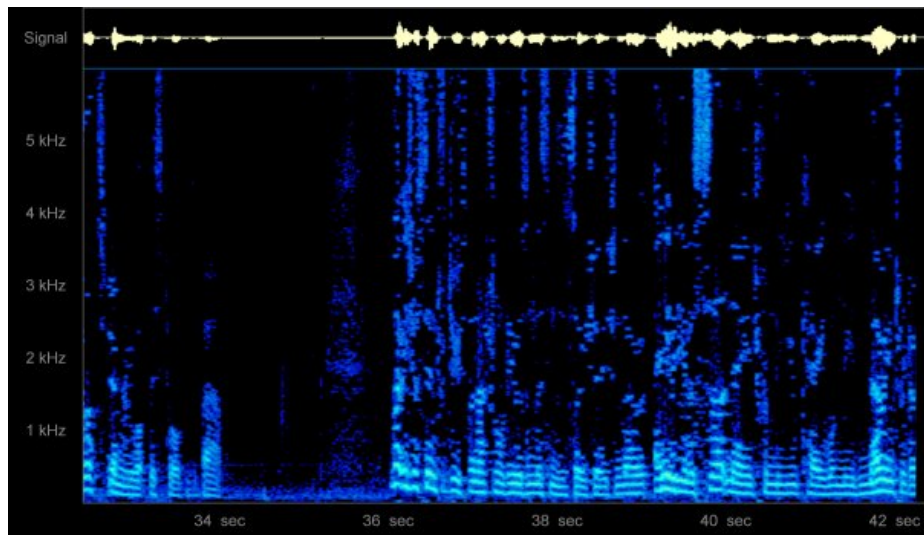


Figure 26: Esempio di spettrogramma. In alto nella figura è rappresentata anche la forma d'onda del segnale di cui è stato calcolato lo spettrogramma.

7 Modulazioni

7.1 La teoria delle modulazioni

L'utilizzo delle modulazioni di ampiezza, frequenza, fase etc. è sorto dapprima nell'ambito delle trasmissioni via radio e via cavo con lo scopo di inviare facilmente un messaggio adattandolo al mezzo di trasmissione [1, 3]. Per comprendere meglio gli effetti di una modulazione, facciamo un esempio pratico: volendo trasmettere un segnale a 20 kHz con un'antenna a dipolo (il risultato viene dalla teoria delle antenne, si chiede al lettore di accettarlo così), avremmo bisogno di un'antenna lunga 375 m (ovvero $\lambda/4$)!! Come potremmo fare per trasmettere lo stesso segnale e utilizzare un'antenna più corta? Potremmo utilizzare una sinusoide a frequenza molto alta (detta PORTANTE, *carrier*), *modulata* - ad esempio in ampiezza - dal segnale che vogliamo trasmettere (SEGNALE MODULANTE). Dato che $\lambda = \frac{c}{f}$, aumentando la frequenza diminuisce la lunghezza d'onda. In questo modo si può utilizzare un'antenna di lunghezza minore.

Nel campo della sintesi del suono gli scopi e le possibilità sono differenti: è possibile infatti modulare un segnale di sorgente utilizzando un segnale di controllo. Come prima opzione possiamo utilizzare un oscillatore a bassa frequenza (*Low Frequency Oscillator*, LFO) per modulare, ad esempio, la frequenza di un oscillatore in banda audio ottenendo così un effetto VIBRATO oppure utilizzare lo stesso LFO per modulare l'ampiezza del segnale sorgente ottenendo un TREMOLO.

Nelle modulazioni dette per *distorsione* (AM, FM, Ring Modulator), andando a modulare il segnale sorgente, osserviamo che nello spettro del segnale in uscita compaiono delle BANDE LATERALI (*side bands*) attorno al segnale sorgente di partenza. Come abbiamo visto, una sinusoide pura appare nel dominio della frequenza come una riga sull'analizzatore di spettro. Tutte le modulazioni influiscono sul contenuto spettrale del segnale sommando alla riga che rappresenta la portante non modulata le bande laterali. Le bande laterali prendono il nome di BANDA LATERALE INFERIORE (LSB *Lower Side Band*) e BANDA LATERALE SUPERIORE (USB *Upper Side Band*) e si trovano, rispettivamente, alla sinistra e alla destra della riga della portante per le frequenze positive e contengono la stessa informazione (per le frequenze negative invece si trovano rispettivamente alla destra e alla sinistra della portante).

7.2 Modulazione di Ampiezza (Amplitude Modulation, AM)

Un segnale modulante a frequenza f_m modula in ampiezza una portante sinusoidale a frequenza f_p , producendo in uscita il segnale:

$$\begin{aligned} v(t) &= V_p (1 + m_a \cos \omega_m t) \cos \omega_p t \\ &= V_p \cos \omega_p t + \frac{m_a V_p}{2} \cos [(\omega_p + \omega_m) t] + \frac{m_a V_p}{2} \cos [(\omega_p - \omega_m) t] \end{aligned} \quad (18)$$

dove m_a è detto *indice di modulazione di ampiezza* ($0 < m_a < 1$) e $\omega_{m,p} = 2\pi f_{m,p}$.

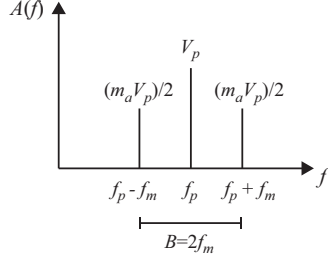


Figure 27: Spettro risultante della modulazione AM.

Come si può notare (vedi Fig. 27), il risultato consiste di 3 sinusoidi: il segnale portante inalterato, un segnale sinusoidale a frequenza somma ($f_p + f_m$) e un segnale sinusoidale a frequenza differenza ($f_p - f_m$). Se l'oscillatore modulante non è un LFO (frequenze sub-audio), l'effetto della modulazione non è un tremolo (come ci aspetteremmo da una modulazione LFO), ma abbiamo un suono più ricco. Questo perché le bande laterali sono ad una distanza tale da poter essere percepite individualmente. Si

vedrà di seguito che con la modulazione FM si riuscirà ad ottenere risultati ancora più interessanti.

Facendo un bilancio, considerando la potenza media di ciascun contributo spettrale come

$$P_x = \lim_{T \rightarrow \infty} \frac{E_{xT}}{T} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 dt \quad (19)$$

e sommando gli integrali ottenuti, se inizialmente si aveva in ingresso potenza totale $P_{IN} = V_p^2 + m_a^2 V_p^2$, in uscita si ottiene:

$$P_{OUT} = V_p^2 + \frac{m_a^2 V_p^2}{4} + \frac{m_a^2 V_p^2}{4} \quad (20)$$

Nel caso generale in cui la portante sia costituita da k parziali, lo spettro risultante presenterà $2k + 1$ parziali (k somme + k differenze + 1 portante).

7.3 Modulazione di Fase (Phase Modulation, PM)

Il segnale modulante agisce sulla fase della portante. Il risultato ottenuto è il seguente:

$$v(t) = V_p \cos(\omega_p t + k_f V_m \sin \omega_m t) \quad (21)$$

dove k_f è detto *indice di modulazione di fase* e $k_f V_m$ indica il massimo scarto di fase.

7.4 Modulazione di Frequenza (Frequency Modulation, FM)

Il segnale modulante agisce sulla derivata della fase istantanea del segnale portante. Per realizzare una modulazione di frequenza si deve operare in questo

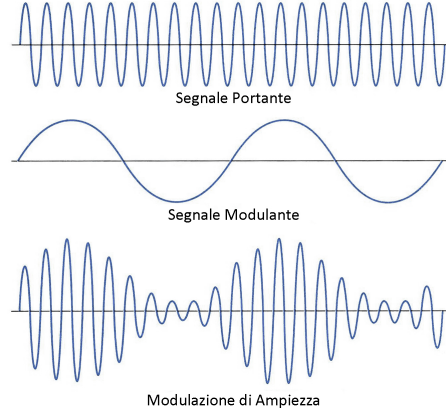


Figure 28: Modulazione di Ampiezza.

modo:

$$v(t) = V_p \cos \left(\omega_p t + \frac{k_f V_m}{\omega_m} \sin \omega_m t \right) = V_p \cos [\omega_p t + I_f(\omega_m) \sin \omega_m t] \quad (22)$$

dove $I_f(\omega_m)$ è detto *indice di modulazione di frequenza* ed è inversamente proporzionale a ω_m a parità di V_m . Nella terminologia del DX7 Yamaha (famoso per essere tra i primi synth a modulazione FM) $k_f V_m$ è detto DEVIAZIONE ed è fornita in termini della frequenza, mentre f_m è chiamata SPEED.

Quando la f_m era quella di un LFO si otteneva un effetto vibrato. Nel caso della modulazione FM, in cui la frequenza del segnale modulante non è in banda sub-audio, l'effetto è ben diverso. Le frequenze f_k risultanti dall'operazione di modulazione FM si ottengono come:

$$f_k = f_p \pm k f_m \text{ con } k = 0, 1, 2, \dots, \infty$$

A differenza della sintesi AM posso avere infiniti contributi spettrali come risultato della modulazione e questo ha decretato il successo della sintesi FM rispetto a quella a modulazione di ampiezza. Di fatto però si limita k ad un valore finito, anche perché avendo la ripartizione della potenza su tutte le parziali, all'infinito il suono sintetizzato risulterebbe "poco interessante". Tra l'altro oltre i 20 kHz non si riuscirebbero ad udire ulteriori contributi nello spettro.

Anche in questo caso quindi si può parlare di *Upper Side Band* e *Lower Side Band*. Lo spettro del segnale risultante può essere messo in relazione con le funzioni di Bessel J_k della prima specie di argomento I_f . Ricaviamo le ampiezze delle parziali risultanti [5] per le frequenze positive:

$$V_k = V_p \cdot \begin{cases} (-1)^k \cdot J_k(I_f) & \text{LSB} \\ J_k(I_f) & \text{USB} \end{cases}$$

(per le frequenze negative ciascuna componente è in controfase rispetto alla corrispondente frequenza positiva).

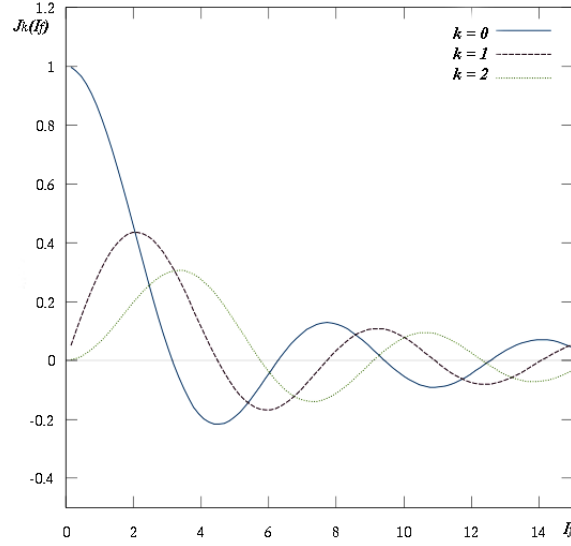


Figure 29: Funzioni di Bessel di prima specie.

Ricordando che $I_f = \frac{k_f}{\omega_m} = \frac{d}{\omega_m}$, se $d = 0$ la modulazione è "off". Aumentando d "stravolgo" lo spettro.

Per ricavare $J_k(I_f)$ usiamo tabelle apposite su cui sono riportati i valori assunti dalle funzioni di Bessel a seconda dei parametri k e I_f . Le tabelle si possono trovare facilmente su internet o su libri di Comunicazioni Elettriche, Telecomunicazioni, Radiotecnica, Analisi Matematica, etc.

Si osservi che i valori di $J_k(I_f)$ tendono a diminuire al crescere di k e dell'indice di modulazione I_f e ciò influisce direttamente sui valori delle ampiezze delle parziali.

Dato che è necessario lavorare con segnali limitati in banda fissiamo un valore massimo per k (derivato dalla REGOLA DI CARSON):

$$k_{MAX} = I_f + 1 \quad (23)$$

Dato che k_{MAX} deve risultare un numero intero e I_f generalmente non lo è, si può approssimare per eccesso k_{MAX} senza perdere porzioni di banda utile.

Infine per tutte le sintesi per distorsione vale la seguente equazione [5] (EQUAZIONE FONDAMENTALE DELLE TECNICHE DI SINTESI PER DISTORSIONE):

$$\frac{f_p}{f_m} = \frac{N_1}{N_2} \quad (24)$$

dalla cui analisi si ricava la *condizione di armonicità dello spettro* e informazioni sul contenuto spettrale.

N_1 e N_2 sono numeri naturali primi fra loro (cioè non semplificabili).

$$\frac{f_p}{N_1} = \frac{f_m}{N_2} = f_0$$

dove f_0 è la nota suonata sul synth.

Studiamo 3 casi di spettri armonici risultanti.

Esempio $N_2 = 1$:

$$\frac{f_p}{f_m} = \frac{2}{1} = \frac{200}{100} \quad \begin{cases} f_p = N_1 \cdot f_0 \\ f_m = f_0 \quad (N_2 = 1) \end{cases}$$

si avrà:

$$f_k = N_1 \cdot f_0 \pm k f_0 = (N_1 \pm k) f_0 \rightarrow \begin{cases} f_p = 2f_0 \\ f_m = f_0 \end{cases} \quad f_0 = \frac{200}{2} = \frac{100}{1} = 100 \text{ [Hz]} \\ f_k = (2 \pm k) 100$$

Sostituisco i diversi valori di k e ottengo lo spettro armonico completo della fondamentale f_0 .

Esempio $N_2 = 2$:

$$\begin{cases} f_p = N_1 \cdot f_0 \\ f_m = 2f_0 \end{cases} \quad \text{Attenzione: in questo caso } N_1 \text{ deve essere comunque DISPARI.} \\ f_k = N_1 \cdot f_0 \pm 2k f_0 = (N_1 \pm 2k) f_0$$

In questo caso si ottiene lo spettro armonico di f_0 mancante delle frequenze multiple di 2 (cioè si avranno solo armoniche dispari).

Esempio $N_2 = M_{piccolo}$:

In modo simile al caso precedente, in questo caso si ottiene lo spettro armonico di f_0 mancante delle frequenze multiple di M . Se M fosse grande non riuscirei a vedere le armoniche mancanti, perché sarebbero troppo lontane da f_0 .

7.5 Modulazione ad Anello (Ring Modulation)

La Ring Modulation è una variante della modulazione di ampiezza. La differenza sta nel fatto che nel segnale di uscita la portante non è presente. Il segnale risultante è quindi:

$$v(t) = \frac{m_a V_p}{2} \cos[(\omega_m + \omega_p)t] + \frac{m_a V_p}{2} \cos[(\omega_m - \omega_p)t] \quad (25)$$

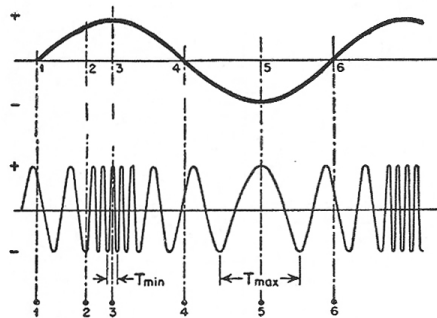


Figure 30: Modulatione di Frequenza.

7.6 Altre modulazioni

7.6.1 Modulatione della Larghezza di Impulso (Pulse Width Modulation PWM)

Per PWM si intende la variazione automatica della larghezza di impulso di un'onda impulsiva più comunemente realizzata utilizzando un LFO.

7.6.2 Cross-Modulation (CM)

Questo tipo di modulazione si ottiene collegando le uscite di due oscillatori ai reciproci ingressi di controllo della frequenza.

7.6.3 LFO Modulation

Per creare effetti di Vibrato o di Tremolo che non abbiano una frequenza fissa si può modulare un LFO con un altro LFO.

References

- [1] M.G. Di Benedetto. *Comunicazioni Elettriche - Fondamenti*. Pearson Prentice Hall.
- [2] L. L. Beranek. *Acoustics*.
- [3] L. Franchina and P. Marietti. *Elettronica dei Sistemi a Piccola Banda Frazionale*. Casa Editrice Ambrosiana.
- [4] P. Guaccero, L. Proietti, and L. Zaccheo. *Lezioni di Music Technology*. Appunti presi a lezione.

- [5] V. Lombardozzi. *Lezioni di Fisica Acustica ed Elettroacustica*. Appunti presi a lezione.
- [6] M. Luise and G. M. Vitetta. *Teoria dei Segnali*. McGraw-Hill.