Appunti per la lezione su

Funzioni trigonometriche e consonanza/dissonanza di un accordo musicale

Povo, 23 febbraio 2007

Sisto Baldo, Dip. Mat. UniTN

Cosa c'entrano le funzioni trigonometriche con la consonanza di un accordo?

Confesso che, in parte, ho scelto questo argomento perché mi intrigava il legame tra una delle parti... meno eccitanti del programma di matematica ed una disciplina che, come la musica, esercita un fascino notevole su chiunque! Ma, ancora, *che c'entrano il seno e il coseno con la consonanza?*

Probabilmente non vi sarà difficile dare una risposta a due domande che si ottengono "allargando" quella che ho appena fatto:

(i) *Cosa c'entrano le funzioni trigonometriche con la musica?*

(ii) *Cosa c'entra la matematica con la consonanza di un intervallo o di un accordo musicale?*

Potremmo rispondere alla domandina (i) ricordando che l'esempio più semplice di "nota musicale" è dato da una funzione periodica del tempo che varia con legge sinusoidale: per esempio, il diapason che usiamo per accordare la chitarra emette un segnale che è una sinusoide smorzata[1].

Che poi la matematica, e più precisamente l'aritmetica, abbia a che fare con la consonanza di un intervallo musicale, è una cosa nota fin dagli albori di entrambe le discipline: la leggenda vuole che sia stato Pitagora, facendo esperimenti con una corda tesa pizzicata, a "inventare" la prima scala musicale. Precisamente, i Pitagorici si resero conto che due note suonate in successione "suonano bene insieme" se le lunghezze delle corde usate per ottenerle stanno tra loro in un rapporto dato da *numeri interi piccoli.* Con linguaggio moderno e "reciproco", due note danno un intervallo *consonante* se il rapporto tra le loro frequenze è dato da piccoli numeri naturali: l'intervallo più consonante dopo l'unisono è quello di ottava (DO-DO, rapporto tra le frequenze uguale a 2), seguito da quello di quinta (DO-SOL, rapporto 3/2), etc.

Per esempio, potrei chiedervi di mettere in "ordine di consonanza" i seguenti quattro "bicordi" (accordi costituiti da due note suonate contemporaneamente):

[1] *La legge con cui vibra il nostro timpano, colpito dalla nota emessa da un diapason, è approssimativamente*

$$f(t) = Ce^{-\alpha t} \sin(2\pi \cdot 440t),$$

dove $440Hz$ *rappresenta la frequenza del LA centrale del pianoforte e* t *è il tempo misurato in secondi...*

- Do-Sol (intervallo di quinta, rapporto 3/2)
- Do-Si (intervallo di settima, rapporto 15/8 nella scala naturale)
- Do4-Do5 (intervallo d'ottava, rapporto 2)
- Do-Do# (semitono)

Vale forse la pena di ricordare che la musica occidentale ha definitivamente rinunciato, fin dal XVIII secolo, a questi rapporti interi "esatti", per sostituirli con le approssimazioni irrazionali che formano la *scala temperata*[2]. D'altra parte, all'atto pratico occorre un certo allenamento per distinguere i toni della scala temperata da quelli della scala naturale: in prima approssimazione possiamo continuare a ritenere che un intervallo sia più consonante se corrisponde ad un rapporto "semplice" tra le frequenze!

Benissimo, ma da dove viene questa strana predilezione del nostro istinto musicale per i rapporti tra interi piccoli? Se i Pitagorici si accontentavano di attribuire ai numeri interi delle valenze magico-esoteriche, la nostra mentalità scientifica di uomini del XXI secolo aspira a motivazioni più convincenti!

Certo, non dobbiamo pretendere troppo: la percezione soggettiva della consonanza o della dissonanza di un accordo non è putroppo una quantità che si possa misurare con uno strumento come la lunghezza, la massa o la carica elettrica... Potremo solo cercare un modello che riesca a "mettere d'accordo" il maggior numero possibile di ascoltatori!

Il modello di consonanza/dissonanza che voglio presentare oggi è stato sostanzialmente proposto da Hermann Helmholtz nella seconda metà dell'ottocento, per poi essere ripreso e reso quantitativo da vari studiosi a partire dal 1960[3]

L'intuizione di Helmholtz si basa essenzialmente su due fatti:

[2]Allo scopo di permettere la modulazione tra diverse tonalità, si è scelto di utilizzare una scala in cui tutti e dodici i semitoni sono uguali, e corrispondono ad un rapporto di frequenze uguale a $\sqrt[12]{2}$. Questa scelta ha avuto un certo prezzo (per esempio il rapporto di quinta diventa $1,498$ anziché 1.5...), ma la libertà per il compositore si è moltiplicata!

[3]Indagini statistiche di Plomp e Levelt nel 1965, modello matematico di Sethares ed altri negli anni '90...

(i) Ogni segnale periodico di frequenza f (la nota emessa da uno strumento musicale!), per complicato che sia il suo grafico, può sempre essere decomposto come somma di (infinite) funzioni di andamento sinusoidale alle frequenze f, $2f$, $3f$, $4f$, $5f, \ldots$: si tratta della cosiddetta *decomposizione in serie di Fourier* di una funzione periodica. Le componenti sinusoidali si chiamano *armonici* del segnale;

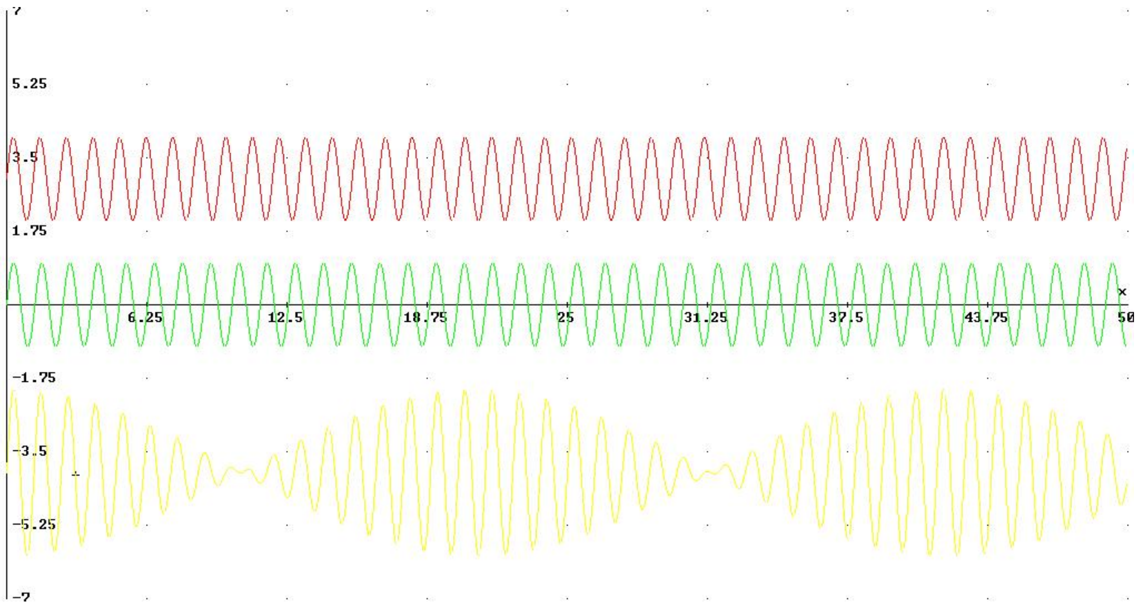
(ii) Tra due note sinusiodali di frequenza abbastanza vicina, si verifica il cosiddetto *fenomeno dei battimenti*, che viene di norma percepito dall'orecchio come "aspro e sgradevole"...

Nel primo fatto entrano prepotentemente le funzioni trigonometriche (come promesso!): potremmo dire in altre parole che qualunque nota può essere scimmiottata da un'*orchestra di diapason*, ciascuno con una frequenza multipla di quella della nota e con ampiezze opportune. Cercherò di convincervi tra poco che questo fatto, ancorché strano, è plausibile!

Il fenomeno dei battimenti, invece, è acusticamente e matematicamente più semplice: per questo motivo, lo vedremo per primo. Cosa succede se sommo due segnali sinusoidali con frequenze vicine tra loro ma non identiche? Per esempio, cosa succede se suono contemporaneamente due diapason, uno accordato sul La4 "canonico" a 440Hz, il secondo sul La4 a 415Hz usato talvolta dalle orchestre barocche che suonano con strumenti originali?

- Questo è un diapason normale...
- E questo un diapason barocco...
- ... mentre suonandoli assieme succede questo!

Quello che avete sentito è il fenomeno dei battimenti! Cerchiamo di capirlo con un grafico:



La figura mostra i grafici delle funzioni $\sin(5x)$, $\sin(5.3x)$ e $\sin(5x)+\sin(5.3x)$: gli ultimi due sono stati opportunamente traslati sull'asse verticale per vederli meglio. Le due sinusoidi "singole" partono più o meno in sincrono, per cui all'inizio la loro somma assomiglia ad una sinusoide di ampiezza doppia... ma poi la sinusoide con frequenza maggiore anticipa sempre più l'altra, fino al punto in cui ad una cresta della prima sinusoide corrisponde una valle della seconda! In quel punto, la somma delle due sinusoidi è nulla. L'effetto totale è un segnale ancora di andamento sinusoidale, ma con un'ampiezza che aumenta e diminuisce periodicamente con una frequenza più bassa.

La cosa può essere facilmente verificata con un conticino: date due sinusoidi di frequenza diversa ma vicina e uguale ampiezza, $\sin(\omega_1 t)$ e $\sin(\omega_2 t)$ con $\omega_1 < \omega_2$, sia $\overline{\omega} = \frac{\omega_1+\omega_2}{2}$ la media delle due pulsazioni, $\delta = \frac{\omega_2-\omega_1}{2}$ la metà della loro differenza. Per ipotesi, δ sarà piccolo. Si ha ovviamente $\omega_1 = \overline{\omega} - \delta$, $\omega_2 = \overline{\omega} + \delta$ da cui, grazie alle formule di addizione[4] si ottiene

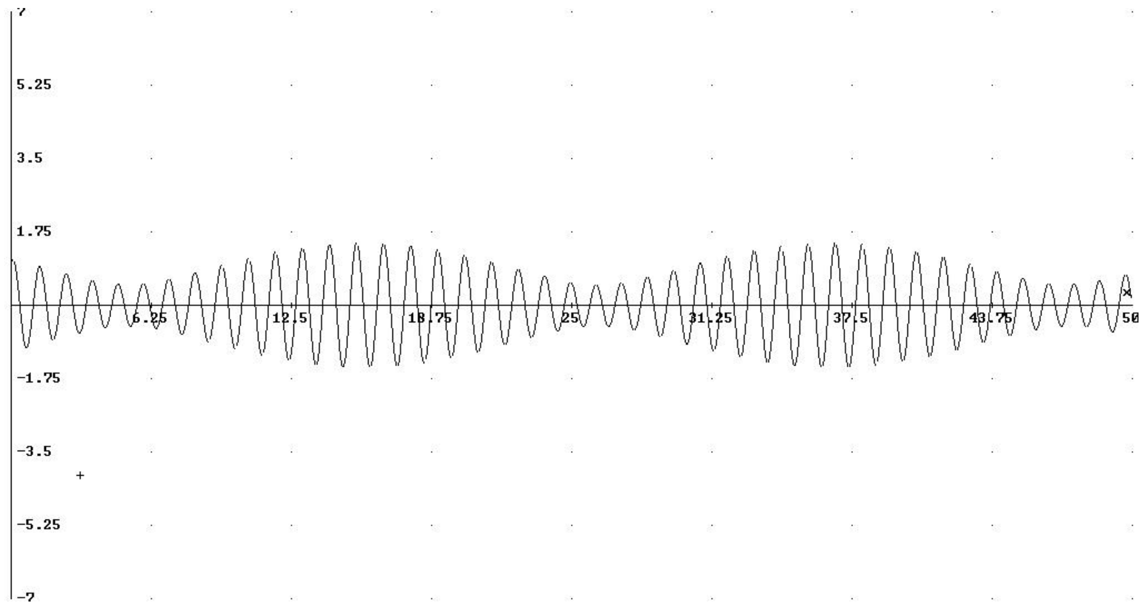
$$\begin{aligned}&\sin\omega_1 t + \sin\omega_2 t = \\ &\sin(\overline{\omega}-\delta)t + \sin(\overline{\omega}+\delta)t = 2\cos\delta t \sin\overline{\omega} t\end{aligned}$$

L'ultima espressione può essere facilmente interpretata come una sinusoide con pulsazione $\overline{\omega}$, la cui "ampiezza" varia lentamente secondo la legge $2\cos\delta t\ldots$[5]

[4] Chiamatele formule di prostaferesi, se vi piace di più...

[5] Abbiamo messo tra virgolette la parola ampiezza perché $2\cos\delta t$ assume anche valori negativi: in quegli intervalli, l'ampiezza "vera" sarà $-2\cos\delta t$, ma la nostra sinusoide sarà riflessa rispetto all'asse dei tempi!

Il fenomeno dei battimenti si ha anche tra segnali sinusoidali sfasati, o con ampiezze diverse. Per esempio, ecco il grafico di $1/2 \sin(5t) + \cos(5.3t)$:



In questo caso, però, l'analisi matematica del fenomeno sarebbe un po' più complicata...

Avendo capito il fenomeno dei battimenti (e avendolo sentito!), passiamo ad esaminare l'altra affermazione che abbiamo fatto: ogni funzione periodica "appena decente" si può scrivere come somma di seni e coseni con frequenze multiple di quella del segnale originale.

Con opportuna scelta dell'unità di tempo, possiamo sempre supporre che il periodo sia 2π. L'affermazione diventa allora la seguente: sia data una funzione $f(t)$, 2π-periodica e non troppo patologica. Vogliamo approssimarla con funzioni sinusoidali nel modo seguente:

$$f(t) \simeq A_0 + A_1 \cos(t + \phi_1) + A_2 \cos(2t + \phi_2) + A_3 \cos(3t + \phi_3) + \ldots + A_n \cos(nt + \phi_n).$$

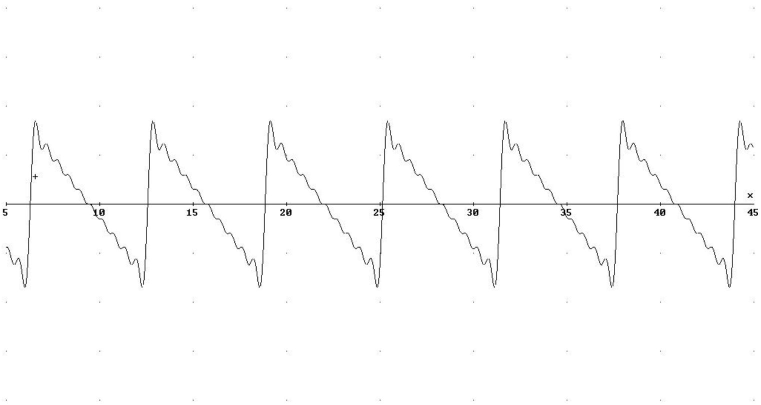
Le sinusoidi che compaiono in questa somma sono i cosiddetti "armonici" del segnale $f(t)$.

C'è un teorema che dice che c'è un solo modo di fare questa approssimazione in modo ottimale: data una funzione periodica $f(t)$ limitata ed integrabile[6], c'è un unico modo di scegliere le ampiezze $A_0, A_1, A_2, A_3, \ldots, A_n$ e gli angoli di traslazione $\phi_1, \phi_2, \ldots, \phi_n$ in modo che la "distanza euclidea" tra $f(t)$ e la somma di sinusoidi

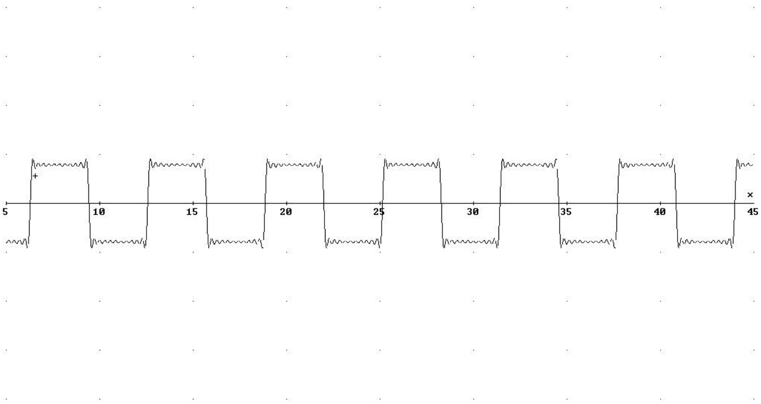
[6]Non sapete cosa vuol dire? Non preoccupatevi... basta che f sia decente!

sia la più piccola possibile. Inoltre, tale distanza diventa arbitrariamente piccola se si prende n abbastanza grande[7]!

Tanto per rendere plausibile questo teorema, diamo due esempi che mostrano come sommando seni e coseni si possano ottenere delle "forme d'onda" che non hanno proprio nulla di sinusoidale: la prima figura mostra il grafico della funzione $\sin t + \frac{1}{2}\sin 2t + \frac{1}{3}\sin 3t + \ldots + \frac{1}{10}\sin 10t$,



la seconda il grafico di $\sin t + \frac{1}{3}\sin 3t + \frac{1}{5}\sin 5t + \ldots + \frac{1}{21}\sin 21t$

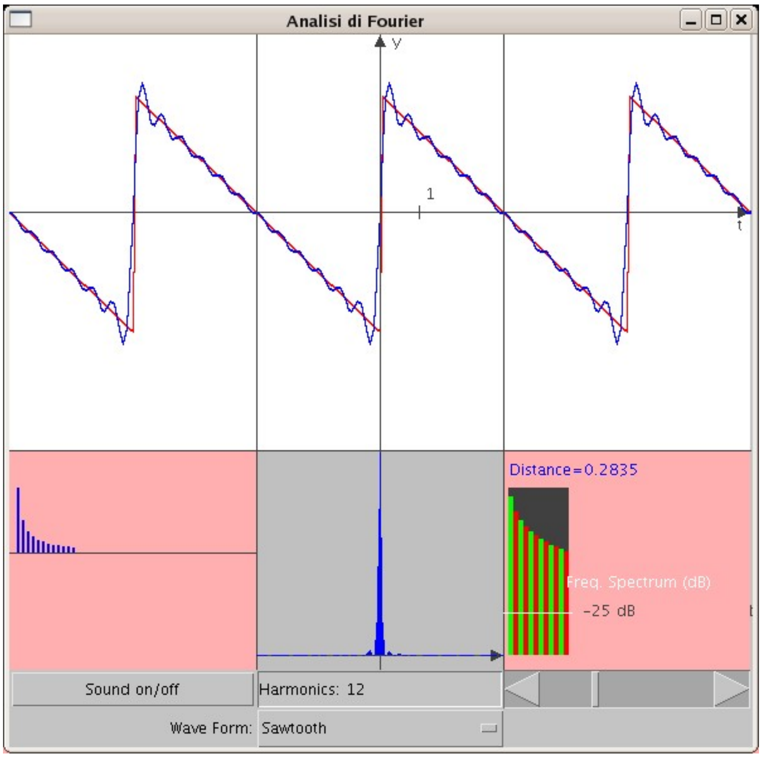


I due grafici assomigliano in modo perturbante ad un'onda a dente di sega e ad un'onda quadra!

Il teorema è reso ancora più plausibile dall'uso di un programmino java che ho scritto per il Progetto Lauree Scientifiche (o di uno degli altri programmini java che illustrano le serie di Fourier e si trovano facilmente in rete):

[7]La "distanza euclidea" tra funzioni è un modo ragionevole per misurare "quanto sono diverse": se per caso conoscete il concetto di integrale, posso dirvi che per definizione si pone

$$dist(f,g) = \sqrt{\int_{-\pi}^{\pi} (f(t) - g(t))^2 \, dt}.$$



l'applet java si trova all'indirizzo
`http://www.science.unitn.it/~baldo/divulgazione/Fourier_bis/fourier.html`
e mostra come qualunque forma d'onda noi disegniamo possa essere approssimata con somme di sinusoidi.

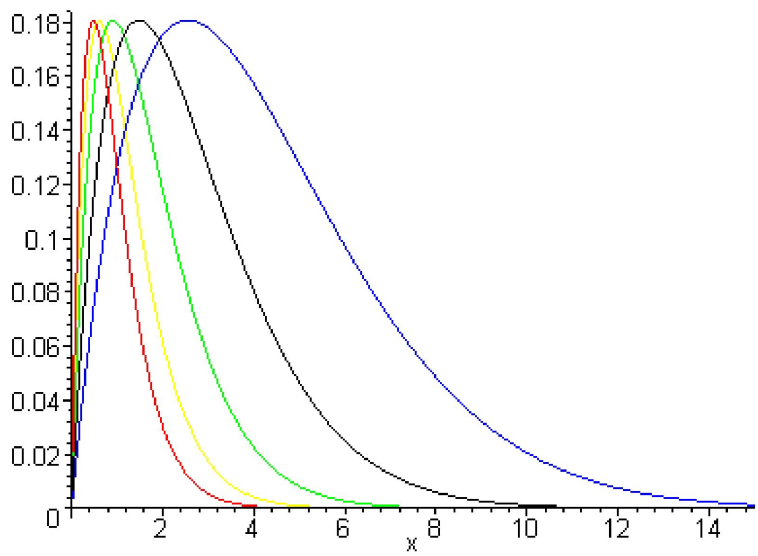
C'è anche il "sonoro", che permette di convincerci della rilevanza acustica di questa bizzarra scomposizione: come già osservava il vecchio Helmholtz, il fatto che una funzione possa essere decomposta in un certo modo non è detto a priori che sia significativo! Per esempio, possiamo certamente scrivere $2 = (1-\pi)+(1+\pi)$, ma questa decomposizione non ci illumina granché sulla "natura" del numero intero 2... D'altra parte, ci sono notevoli evidenze sperimentali del fatto che il nostro orecchio interno faccia *proprio* una decomposizione in componenti sinusoidali dei segnali che gli arrivano: uno degli indizi risiede nel fatto che siamo incapaci di "sentire" gli angoli di fase ϕ_i nella decomposizione di un segnale in sinusoidi. Cambiando gli angoli di fase, la forma d'onda cambia completamente, ma il nostro orecchio sente lo stesso suono[8]!

Cerchiamo ora di mettere insieme le due cose che abbiamo scoperto (i battimenti e la decomposizione di un segnale periodico come somma di sinusoidi) per

[8] L'applet consente di sentire... e vedere un'onda a dente di sega (Sawtooth), e un'onda a dente di sega con le fasi cambiate in modo random (Sawtooth???). Stesso suono con forme d'onda diversissime!

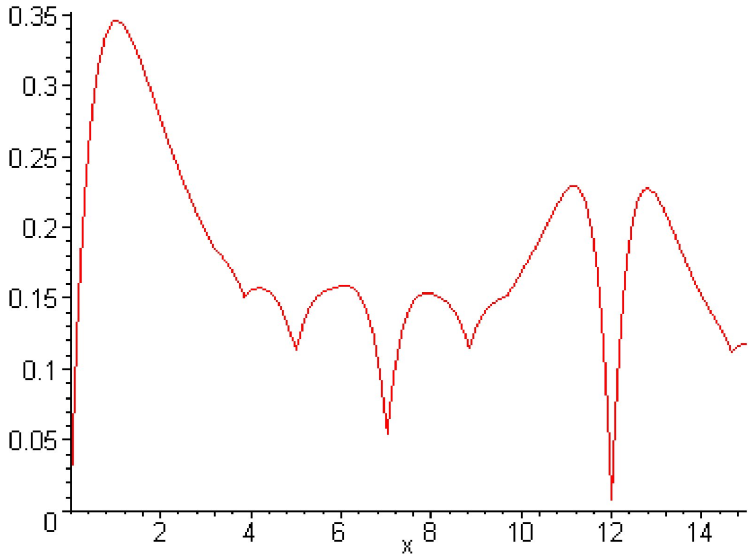
trovare il modo di "misurare" la dissonanza di un bicordo!

Helmholtz ebbe un'intuizione geniale: se suoniamo due note *perfettamente sinusoidali*, la consonanza del bicordo non segue le leggi cui sono abituati i musicisti, in quanto la dissonanza è data unicamente dalla presenza o assenza di battimenti. Precisamente, se fissiamo la frequenza della prima nota e facciamo aumentare lentamente la seconda, si produrranno dei battimenti che daranno quasi subito una sensazione di "ruvidezza" o dissonanza. Questo rimane vero, per un po', anche quando la differenza tra le due frequenze è troppo grande per poter percepire i "singoli" battimenti, ma è ancora abbastanza piccola. Aumentando ancora di più la frequenza della seconda nota, cominciamo a percepire due toni nettamente distinti, e la sensazione di dissonanza scompare. Quanto detto è rappresentato nel seguente grafico, che reca in ascissa il rapporto tra le due frequenze (misurato in semitoni), e in ordinata la dissonaza percepita dagli ascoltatori (misurata in unità arbitrarie): le diverse curve corrispondono a diverse frequenze della nota più bassa (si è trovato sperimentalmente che l'intervallo di frequenze in cui si percepisce la dissonanza diventa più stretto all'aumentare dell'altezza delle note coinvolte: questo è ragionevole se si pensa che la scala in semitoni è di tipo logaritmico, mentre il fenomeno dei battimenti dipende dalla differenza tra le frequenze):[9]



[9] Le curve rappresentate nel grafico sono state ottenute grazie all'apporto di più studiosi: nel 1965 R. Plomp e M. Levelt hanno pubblicato i risultati di uno studio sulla consonanza tra sinusoidi, effettuato su vari gruppi di ascoltatori e con diverse altezze delle note, allo scopo di verificare sperimentalmente le affermazioni di Helmholtz. Le curve in figura rappresentano delle espressioni matematiche proposte nel 1993 da W. Sethares per rappresentare i risultati dello studio di Plomp e Levelt.

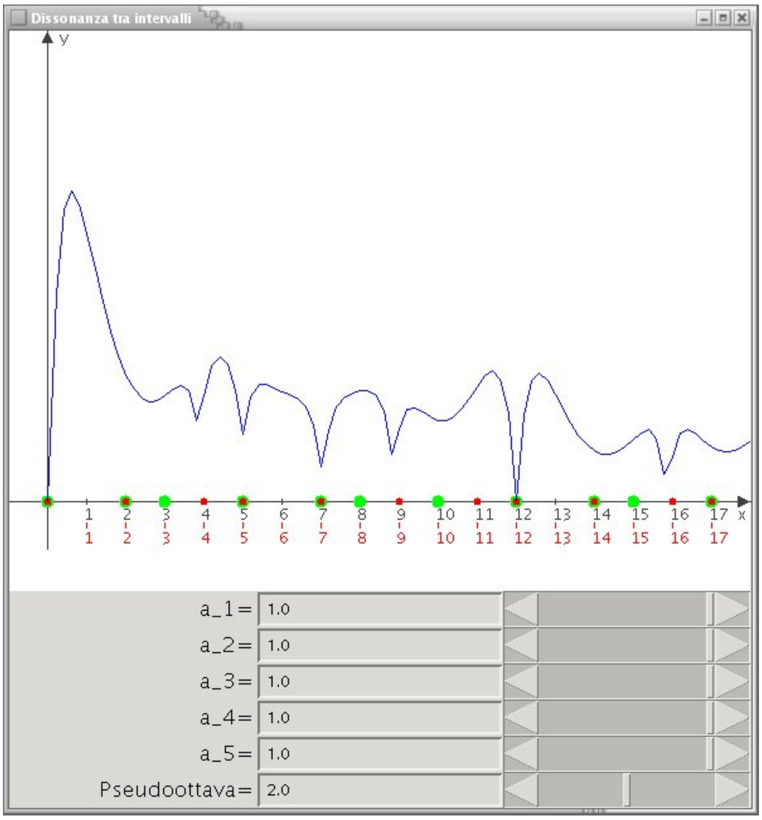
Supponiamo ora di suonare contemporaneamente due note "reali", prodotte da uno strumento musicale: abbiamo visto che ciascuna di queste si decompone come somma di sinusoidi con frequenza multipla dell'altezza della nota! È allora naturale ipotizzare che la *dissonanza totale* del bicordo sarà data dalla somma delle dissonanze tra tutte le coppie di suoni sinusoidali che compaiono nella decomposizione[10]. Ecco quel che si ottiene rappresentando graficamente il grafico di un bicordo tra due note, ciascuna costituita da 5 armoniche di ampiezza uguale: si tratta di un suono un po' caricaturale che però mostra molto bene quel che succede!



In ascissa abbiamo ancora il rapporto di frequenza tra le due note in semitoni, in ordinata la dissonanza. I minimi della curva rappresentano quindi *i bicordi relativamente più consonanti* secondo questa teoria. E si trovano esattamente dove ce li aspettiamo: abbiamo minimi profondissimi in corrispondenza dell'unisono, dell'ottava e della quinta, minimi un po' meno pronunciati in corrispondenza della terza e della quarta maggiore. La settima, il tono e il semitono sono invece previsti correttamente come dissonanti!

Possiamo sperimentare la teoria grazie ad un altro

[10]Tali dissonanze vanno pesate opportunamente con l'ampiezza delle varie componenti sinusoidali: per farlo, ci sono proposte leggermente diverse dovute a W. Sethares (1993) e P. Vassilakis (2001). Ho scelto di seguire Vassilakis perché mi sembra dia un modello più realistico...



programmino java disponibile all'indirizzo
`http://www.science.unitn.it/~baldo/divulgazione/appletdissonanza/appletdissonanza`
che ho realizzato l'estate scorsa mentre collaboravo alle ricerche di Lorenzo Valdan per la sua tesi di Laurea[11], della quale sono stato relatore assieme a Stefano Oss.

Bellissimo, vero? Abbiamo una teoria relativamente semplice che rende perfettamente conto del fenomeno della consonanza musicale. Dunque, la percezione della musica si presta ad una precisa analisi matematica!

Ma sarà *proprio* così?

Non completamente: questa teoria spiega molte cose, ma non è del tutto soddisfacente da altri punti di vista. Un bellissimo esperimento che è stato proposto per..confutarla parzialmente è il seguente: cosa succede se decidiamo di "allargare" l'intervallo di ottava, sostituendo il canonico rapporto 2 tra le frequenze con un valore più grande?

Secondo la teoria di Helmholtz, Plomp e Levelt, Sethares, se abbiamo cura di allargare anche, in proporzione, gli intervalli di frequenza tra gli armonici delle nostre note, tutto quanto visto sopra dovrebbe rimanere vero e otteniamo ancora

[11] L. Valdan: *Applicazioni dell'analisi di Fourier allo studio della consonanza in musica.* Tesi di Laurea Specialistica in Matematica, Università di Trento, Luglio 2006

una scala di valori relativamente consonanti. Anzi, la consonanza dovrebbe migliorare perché se allontaniamo gli armonici riduciamo i battimenti! L'applet visto prima consente anche di espandere l'ottava e di sperimentare quel che succede: possiamo convincerci che ci sono ancora dei bicordi consonanti, ma certamente il tutto suona assai peggio che con l'ottava normale!

Se avete uno stomaco forte, posso farvi sentire come io e Lorenzo Valdan ci siamo divertiti a storpiare un canone di Bach in ottava espansa: ascoltate i due pezzi seguenti, e ditemi quale vi sembra meno schifoso.

- Primo pezzo...
- Secondo pezzo...

Di solito ottengo risposte diverse: i non musicisti preferiscono di norma il primo, i musicisti il secondo... La ragione è che il primo canone è suonato in ottava espansa e con timbro espanso, come spiegato sopra. Nel secondo canone, abbiamo allontanato gli armonici del timbro... ma abbiamo suonato le note come le ha scritte Bach senza espandere l'ottava! La teoria di Sethares dice che dovrebbe essere preferibile il primo metodo, perché i toni dell'ottava espansa sono quelli che minimizzano la dissonanza quando le note hanno gli armonici nella strana posizione che abbiamo scelto. Ma i musicisti son sempre sono d'accordo: difficile capire se la ragione è unicamente culturale, o se ci sono motivazioni più profonde!

Altra ragione per non credere completamente a Helmholtz e a Sethares: sentite come una

- cadenza "normale"
- si trasforma se suonata in ottava espansa...

Il carattere "conclusivo" della cadenza è completamente perso!

A parziale consolazione, permettetemi di osservare che la ricerca sulla consonanza musicale è tuttora fiorente... e nessuna delle spiegazioni proposte sino a questo momento sembra spiegare tutto!

Funzioni trigonometriche e consonanza/dissonanza di un accordo musicale

Povo, 23 febbraio 2007

Sisto Baldo, Dip. Mat. UniTN

Qualche esercizio conclusivo...

1. Disegnate la forma d'onda del diapason. In altre parole, disegnate il grafico della funzione $f(t) = e^{-\alpha t} \sin \omega t$.

2. Qual è la differenza, misurata in semitoni, tra la nota emessa dal diapason normale a $440Hz$ e quella del diapason barocco a $415Hz$? Ricordo che un semitono corrisponde ad un rapporto tra le frequenze pari a $\sqrt[12]{2}$, per cui k semitoni corrispondono ad un rapporto tra le frequenze pari a $(\sqrt[12]{2})^k$.

3. Se avete avuto modo di maneggiare i vecchi dischi in vinile, vi sarà probabilmente capitato di suonare per errore un disco a 33 giri con la velocità del piatto impostata su 45 giri/minuto. Ovviamente, la musica risulta accelerata... ma anche più acuta! Di quanti semitoni si alza?

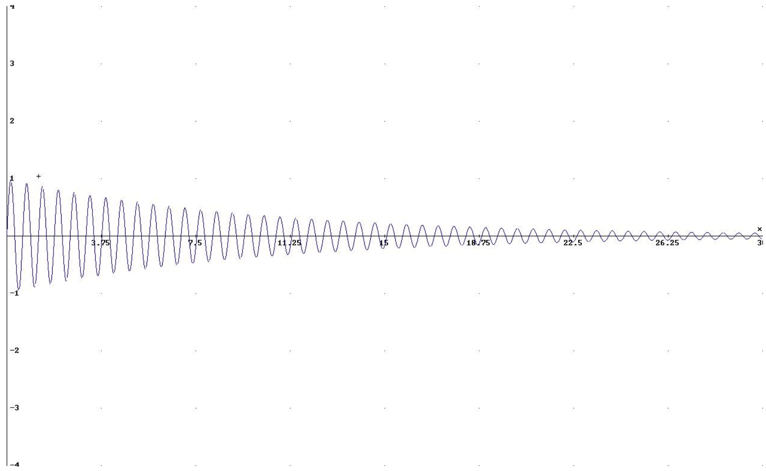
4. Un esercizio sull'ottava espansa! Supponiamo di espandere il rapporto d'ottava da 2 a un numero $A > 2$. Abbiamo accennato al fatto che occorre "allontanare" gli armonici. Nell'ottava normale, gli armonici di una nota di frequenza f hanno frequenze $f, 2f, 3f, 4f, \ldots$. Nell'ottava espansa gli armonici dovranno avere frequenze $f, F_2, F_3, F_4, \ldots$, dove le frequenze sono scelte in modo che il rapporto tra la frequenza del k-esimo armonico e la fondamentale f, misurata in *semitoni espansi*, sia uguale al rapporto tra la frequenza del k-esimo armonico standard e la fondamentale misurata in semitoni standard (un semitono espanso vale evidentemente $\sqrt[12]{A}$). Come dobbiamo scegliere le frequenze F_k? [*Si imposta l'equazione* $\log_A(F_k/f) = \log_2 k$, *da cui* $F_k = k^{\log_2 A} f$...]

5. Mostrate che ogni espressione del tipo $a \cos \omega t + b \sin \omega t$ può essere riscritta nella forma $A \cos(\omega t - \phi)$, ed è quindi una sinusoide opportunamente traslata lungo l'asse dei tempi. [*SUGG.: Si tracci, nel piano cartesiano, il punto* $P = (a, b)$. *Sia poi* A *la distanza del punto* P *dall'origine* O, ϕ *l'angolo formato dal semiasse delle* x *positive col segmento* $\overline{OP}$...]

6. Riscrivete in modo astuto l'espressione $\sin \omega_1 t + \cos \omega_2 t$ in una forma che metta in evidenza i battimenti che si hanno quando ω_1 e ω_2 sono vicini. [*Sugg.: Porre* $\overline{\omega} = (\omega_1 + \omega_2)/2$, $\delta = (\omega_2 - \omega_1)/2$. *Fare poi il contazzo e usare l'esercizio precedente*...]

SOLUZIONI:

1. La funzione data può essere vista come una sinusoide con ampiezza che decresce secondo la legge $e^{-\alpha t}$: il grafico che si ottiene sarà dunque



2. Detto k il numero di semitoni richiesto, deve essere $2^{k/12} = 440/415$, da cui prendendo i logaritmi $k = 12 \log(440/415)/\log 2 \simeq 1.01$: il rapporto tra le due frequenze è circa di un semitono.

3. La frequenza verrà moltiplicata per il rapporto tra le due velocità del piatto. Quindi, se ragioniamo come nell'esercizio precedente e k indica il numero di semitoni di cui si alza la musica, dovrà essere $2^{k/12} = 45/33$, da cui prendendo i logaritmi $k = 12 \log(45/33)/\log 2 \simeq 5.36$. Le frequenze si alzano di un po' più di 5 semitoni: per esempio, un DO diventa un FA (crescente).

4. Il suggerimento spiega già tutto!

5. Seguendo il suggerimento, vediamo che $a = A \cos \phi$, $b = A \sin \phi$. Sostituendo nell'espressione di partenza otteniamo $a \cos \omega t + b \sin \omega t = A(\cos \phi \cos \omega t + \sin \phi \sin \omega t) = A \cos(\omega t - \phi)$.

6. Seguendo il suggerimento troviamo, usando le formule di addizione e raccogliendo:

$$\sin \omega_1 t + \cos \omega_2 t = \sin(\overline{\omega} - \delta)t + \cos(\overline{\omega} + \delta)t =$$
$$(\sin \overline{\omega} t + \cos \overline{\omega} t)(\cos \delta t - \sin \delta t).$$

L'esercizio precedente ci suggerisce di riscrivere le due parentesi nell'ultima riga rispettivamente come $\sqrt{2} \cos(\overline{\omega} t - \pi/4)$ e $\sqrt{2} \cos(\delta t + \pi/4)$. Otteniamo allora quanto richiesto:

$$\sin \omega_1 t + \cos \omega_2 t = 2 \cos(\overline{\omega} t - \pi/4) \cos(\delta t + \pi/4).$$