SENTIRE LA MATEMATICA

STAGE AL DIMA, 20 APRILE 2007

1. Il suono: propagazione e ricezione

1.1. **Cosa vuol dire suono?** Da un punto di vista fisico per suono in un certo punto dello spazio si intende una rapida variazione di pressione (compressione e rarefazione) intorno al valore assunto dalla pressione atmosferica in quel punto. Si definisce **sorgente sonora** qualsiasi dispositivo che provochi direttamente o indirettamente (ad esempio per percussione o per vibrazione) dette variazioni di pressione: in natura le sorgenti sonore sono quindi praticamente infinite come ognuno può constatare; affinché il suono si propaghi occorre poi che il mezzo che circonda la sorgente sia dotato di elasticità.

Le condizioni essenziali per la generazione e propagazione del suono così come definito sono due:

- una variazione di pressione nel mezzo intorno ad un valore di equilibrio (ad esempio la pressione atmosferica);
- la presenza di un mezzo elastico (nel vuoto non c'è propagazione sonora).

1.2. **Le onde piane armoniche.** Un esempio di segnale sonoro è dato da funzioni del tipo

$$W(x,t) = A\,\sin(k(x - vt)) \qquad \text{o anche} \qquad W_1(x,t) = A\,\cos(k(x - vt)).$$

dove x è la variabile spaziale e t è la variabile temporale. Queste rappresentano un segnale sonoro particolare: un'onda che si propaga lungo una direzione fissata (e misurata da x) con velocità v. Le due rappresentazioni sono equivalenti (si tratta della stessa funzione spostata di $\pi/2$ e quindi consideriamo la W). A è chiamata anche ampiezza dell'onda e è in relazione con l'intensità del segnale sonoro; k è anche detto numero d'onda e è in relazione con l'altezza del suono.

A t fissato, $W(x,t)$ è una funzione di x, che indica l'intensità del segnale nel punto x (pensiamo fissata una direzione di propagazione: onda piana significa che ha un certo valore uguale in tutti i punti di un piano ortogonale alla direzaione di propagazione). È una funzione periodica di periodo $\ell = 2\pi/k$. Questo numero ℓ è anche chiamato lunghezza d'onda.

Fissato ora x, queste sono funzioni di t, che indicano l'intensità del segnale a un certo istante t; il periodo temporale dopo il quale l'onda si ripete uguale è $T = 2\pi/(kv)$ e il numero di oscillazioni compiute in un secondo è la frequenza $\omega = 1/T = (kv)/(2\pi)$. Quindi

$$\text{velocità} \quad v = \omega\,\ell \quad \text{frequenza} \times \text{lunghezza}.$$

Il suono di un'onda piana armonica non è particolarmente interessante: ascoltiamo il suono generato (con software audacity sine wave 440Hz).

1.3. **Come si misura?** Consideriamo un'onda piana armonica

$$W(x,t) = A\sin(k(x-vt)).$$

Due sono le sue caratteristiche fondamentali che abbiamo notato: l'ampiezza A e la frequenza $\omega = (kv)/(2\pi)$.

La frequenza si misura in Hertz 1Hz=1oscillazione al secondo.

L'ampiezza è il termine fisico con cui viene definita l'intensità di un suono, ossia l'energia trasportata da un'onda.

Solitamente viene misurata in decibel, una misura logaritmica che descrive un rapporto con un valore preso come riferimento. Solitamente si misura una variazione di pressione atmosferica e di solito si sceglie il valore di riferimento di 20 micropascals, perché è al limite della sensibilità dell'orecchio umano. (È un valore molto piccolo: ricordiamo che la pressione atmosferica standard è di 1013 mbar e 1mbar=100 Pa).

$$20\log_{10}(p/p_0) \qquad p_0 = 20\mu Pa.$$

Ad esempio a un aumento di circa 6 dB corrisponde un raddoppio dell'intensità sonora, cioè un raddoppio del volume.

1.4. **Ricezione del suono.** I suoni "interessanti" sono quelli che possiamo ascoltare e possiamo ascoltare suoni che hanno frequenze comprese tra 20 e 20 000 Hz.

Un fatto interessante è che la nostra percezione del suono è differente alle varie frequenze; quando prima ragionavamo sul valore di riferimento p_0 si intendeva che il suono prodotto aveva una frequenza di $1kHz$, una frequenza "centrale".

Nella Tabella 1 ci sono le intensità medie di alcuni suoni noti.

Nella Figura 1 vediamo come si distribuiscono alcuni suoni in un grafico frequenza/ampiezza.

Infine nella Figura 2 vediamo quali sono le curve che sembrano essere percepite allo stesso volume al variare di frequenza e ampiezza. In questa figura compare il "phon", che non è una vera unità di misura oggettiva di qualche fenomeno fisico, ma una misura "psichica" della risposta dell'orecchio. Un suono di 80 phon corrisponde a un suono di 80 dB a 1 kHz, oppure a un suono di 90 dB a 60 Hz.

Si usano i logaritmi sia nelle frequenze sia nelle ampiezze perché la relazione tra proprietà percepite e proprietà fisiche non è una relazione lineare: per esempio, aumentando l'ampiezza di una forma d'onda di una uguale grandezza, non si ottengono uguali incrementi di volume (il volume sembra aumentare di meno via via che diventa più elevato). Analoga caratteristica vale per la frequenza: ad aumenti uguali di frequenza non corrispondono uguali incrementi di tono (l'incremento di tono sembra sempre più piccolo col crescere della frequenza).

Potremmo sintetizzare questi concetti dicendo che l'orecchio è "logaritmico". Inoltre le frequenze e le intensità relative al parlato sono quelle che l'orecchio riesce a distinguere meglio.

Tabella 1. Esempi di Livelli di intensità del suono

Sorgente	**Livello (dB)**
Soglia dell'udito (a 1000 Hz)	0
Zanzara vicino all'orecchio	10
Bisbiglio	15-20
Rumore di fondo notturno in città	30
Ufficio silenzioso	50
Conversazione tra 2 persone a 1 m	55
Ristorante affollato	60-65
Traffico cittadino diurno	70-80
Martello pneumatico (a 3 m)	90
Metropolitana	100
Complesso rock in locale chiuso	110
Soglia del dolore	120
Jet al decollo (a 50 m)	130
Rottura del timpano	160
Lancio di un missile (a 50 m)	200
Massimo rumore prodotto in laboratorio	210

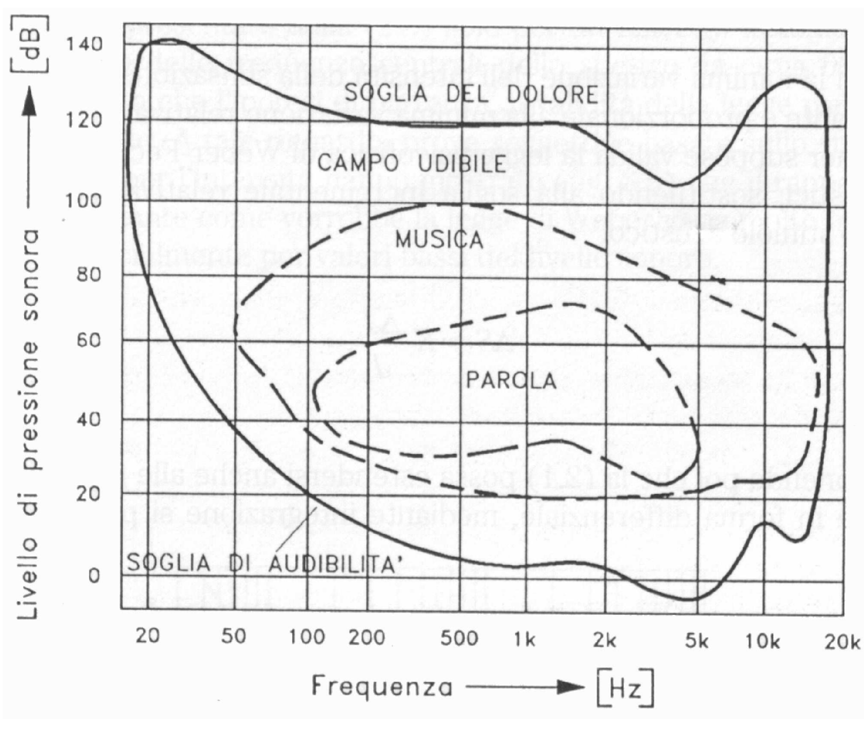

Figura 1. Frequenza e intensità dei suoni

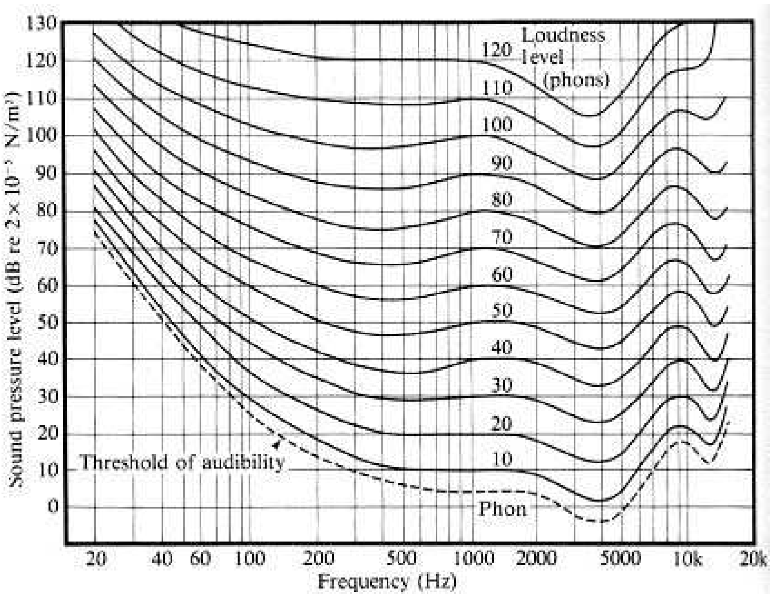

FIGURA 2. Curve di ugual percezione

2. Sovrapposizione di onde sonore: le serie di Fourier

Pensiamo di essere un ascoltatore, fisso in una certa posizione dello spazio. Ci interessa conoscere un dato segnale sonoro al variare del tempo. Quindi dimentichiamo per il momento la variabile spaziale (ad esempio pensiamo che sia $x = 0$) e concentriamoci su quella temporale.

Esercizio: disegnare i grafici di $2\sin(t)$; $\sin(2t)$; $\sin(t) + \cos(t)$; $\sin(t) + \sin(2t)$.

Sia data una funzione $f(t)$, periodica di periodo T. L'idea fondamentale che ha avuto Fourier è che $f(t)$ può sempre essere rappresentata come una sovrapposizione di onde armoniche (piane) tutte di periodo T del tipo

$$\begin{aligned} f(t) \simeq{} & a_0 + a_1 \cos(2\pi t/T) + b_1 \sin(2\pi t/T) + a_2 \cos(4\pi t/T) + b_2 \sin(4\pi t/T) \\ & + a_3 \cos(6\pi t/T) + b_3 \sin(6\pi t/T) + \cdots \end{aligned}$$

Notiamo che

$$\int_0^T \sin(2p\pi t/T)\, \cos(2q\pi t/T)\, dt = 0 \qquad \forall p, q \in \mathbb{N}$$

e inoltre se $n \neq m$ valgono

$$\int_0^T \sin(2n\pi t/T)\, \sin(2m\pi t/T)\, dt = 0 \qquad \int_0^T \sin^2(2n\pi t/T)\, dt = T/2$$

e relazioni analoghe per i coseni.

Allora i coefficienti a_n e b_n si calcolano integrando sull'intervallo $[0, T]$:

$$a_0 = \frac{1}{T}\int_0^T f(t)\,dt$$
$$a_n = \frac{2}{T}\int_0^T f(t)\,\cos(2n\pi\,t/T)\,dt$$
$$b_n = \frac{2}{T}\int_0^T f(t)\,\sin(2n\pi\,t/T)\,dt$$

Ad esempio, consideriamo la funzione periodica (Figura 3) definita da: $f(t) = a$ per $0 \le t \le T/2$ e $f(t) = -a$ per $T/2 < t \le T$. Questa corrisponde alla cosiddetta onda quadra, ossia

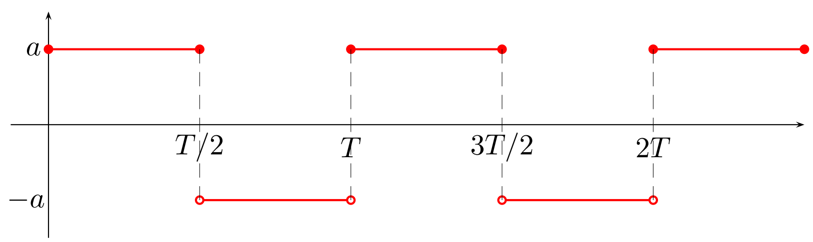

Figura 3. L'onda quadra

È facile controllare che

$$f(t) \simeq \frac{4a}{\pi}\left(\sin(2\pi t/T) + \tfrac{1}{3}\sin(6\pi t/T) + \tfrac{1}{5}\sin(10\pi t/T) + \cdots\right)$$

e nella figura sottostante (disponibile anche una animazione) vediamo come al crescere dei termini considerati otteniamo una buona somiglianza con la curva di partenza, almeno lontano dalle discontinuità (questo fenomeno è noto come fenomeno di Gibbs, ossia oscillazioni ampie attorno ai punti di discontinuità).

Anche le funzioni non periodiche (quali quelle che rappresentano un fenomeno impulsivo) possono essere adeguatamente rappresentate da una somma di infiniti termini armonici. Infatti, per un processo non periodico che si esaurisce in un intervallo di tempo limitato possiamo sempre immaginare di sostituire alla funzione impulsiva $f(t)$, che rappresenta analiticamente il processo, una funzione periodica ottenuta ripetendo la $f(t)$ ad intervalli opportuni di lunghezza T, se T è la lunghezza di un intervallo di tempo contenente l'intero processo. Abbiamo così costruito una funzione periodica di periodo T, che può essere espressa con una serie di Fourier, con T molto grande o, equivalentemente $\omega = 1/T$ molto piccolo:

$$\begin{aligned} f(t) \simeq\ & a_0(\omega) + a_1(\omega)\cos(2\pi\omega t) + b_1(\omega)\sin(2\pi\omega t) + a_2(\omega)\cos(4\pi\omega t) + b_2(\omega)\sin(4\pi\omega t) \\ & + a_3(\omega)\cos(6\pi\omega t) + b_3(\omega)\sin(6\pi\omega t) + \cdots \end{aligned}$$

Osserviamo anche che al crescere dell'intervallo T il periodo della funzione cresce e la frequenza corrispondentemente decresce; inoltre i singoli addendi, sono molto simili se ω è piccolo. Al limite,

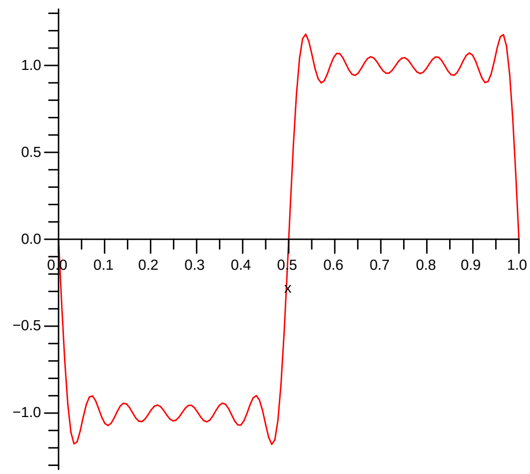

FIGURA 4. L'onda quadra ($a = 1$, $T = 1$) approssimata con i primi 20 termini della sua serie di Fourier.

facendo tendere T all'infinito, si intuisce che la rappresentazione corretta della funzione impulsiva sarà data da un integrale, al posto della serie, con la frequenza come variabile continua:

$$f(t) = \int_0^{\infty} (a(\omega)\sin(2\pi\omega t) + b(\omega)\ \cos(2\pi\omega t))\, d\omega.$$

Da questo discende l'importanza assunta nella trattazione dalle onde armoniche: tramite esse, infatti, qualunque fenomeno ondulatorio, periodico o impulsivo, può essere "scomposto" in componenti elementari, le onde armoniche. Per questo motivo d'ora in poi ci riferiremo quasi esclusivamente ad onde armoniche e a loro sovrapposizioni.

3. Una sorgente sonora: la chitarra

Consideriamo una corda tenuta fissa agli estremi che viene spostata leggermente dalla sua posizione di riposo. Sia $S_t(x)$ la funzione che descrive lo spostamento dalla posizione di riposo della corda nel punto x della corda e al tempo t. In Figura 5 vediamo la situazione al tempo t fissato: la corda tenuta fissa agli estremi — l'origine e il punto $(L, 0)$ — è disegnata in rosso; un punto sulla corda ha coordinate $(x, S_t(x))$.

La corda è sottoposta alla tensione T (perché è fissata agli estremi) e questa è l'unica forza in gioco. Cerchiamo di scrivere un'equazione che coinvolga $S_t(x)$. Consideriamo un piccolo elemento di corda sottoposto ai suoi estremi alla tensione T, la quale forma con l'asse x gli angoli α e α', come nella Figura 6. Ovviamente $\operatorname{tg}\alpha$ è la derivata rispetto a x della funzione $S_t(x)$.

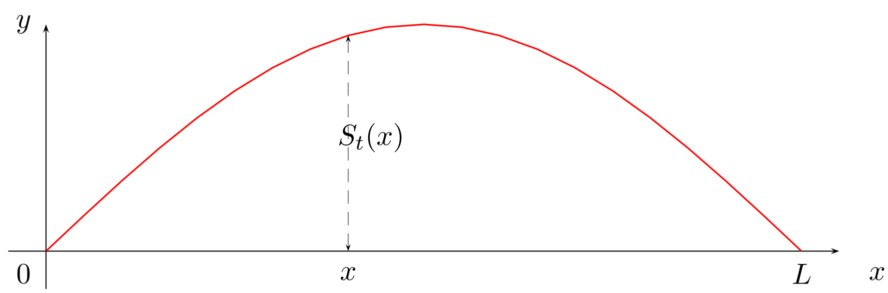

FIGURA 5. La corda al tempo t

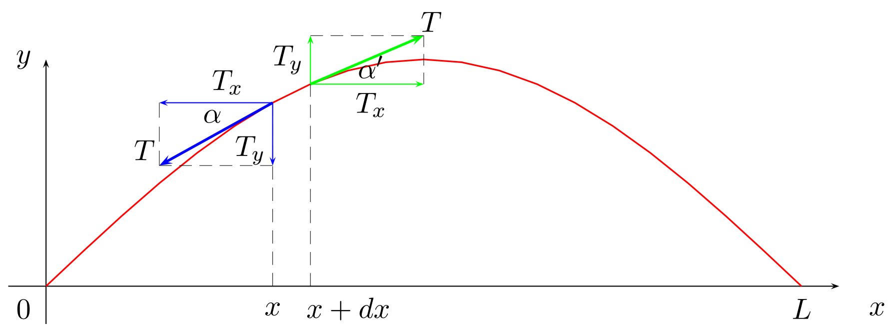

FIGURA 6. La tensione

Scriviamo l'equazione di Newton $F = ma$ (forza=massa $\times$ accelerazione). Le componenti della forza sono:

$$F_x = T(\cos\alpha' - \cos\alpha)$$
$$F_y = T(\sin\alpha' - \sin\alpha).$$

Se supponiamo che α e α' siano piccoli (altrimenti la corda si romperebbe), allora $\cos\alpha \simeq \cos\alpha' \simeq 1$ e $\sin\alpha' - \sin\alpha \simeq \operatorname{tg}\alpha' - \operatorname{tg}\alpha \simeq (\operatorname{tg}\alpha)'\,dx$. Quindi $F_x = 0$ mentre:

$$F_y = T\,(\operatorname{tg}\alpha)'\,dx.$$

Ricordiamo che $\operatorname{tg}\alpha$ è la derivata rispetto a x della funzione $S_t(x)$, quindi $F_y = T\,\frac{\partial^2}{\partial x^2} S_t(x)\,dx$.

Uguagliamo la forza al prodotto della massa del trattino di corda di lunghezza circa dx per l'accelerazione; la massa è $\lambda\,dx$, dove λ è la densità lineare della corda ($\lambda = \rho\,S$ se S è la sezione della corda e ρ la sua densità di volume); mentre l'accelerazione è la derivata seconda rispetto al tempo $\frac{\partial^2}{\partial t^2} S_t(x)$. In conclusione si ottiene:

$$\frac{\partial^2}{\partial x^2} S_t(x) = \frac{\lambda}{T}\,\frac{\partial^2}{\partial t^2} S_t(x).$$

Questa equazione si chiama equazione delle onde e va appaiata con i dati

$$S_t(0) = S_t(L) = 0 \qquad \forall t \tag{1}$$

$$S_0(x) = f(x) \qquad \forall x \tag{2}$$

$$\tfrac{\partial}{\partial t} S_t(x)|_{t=0} = 0 \qquad \forall x \tag{3}$$

che ci dicono che la corda rimane sempre fissata agli estremi, che la corda viene rilasciata da ferma a partire da una certa posizione descritta dalla funzione f.

L'idea fondamentale che ha avuto Fourier, di cui parlavamo anche prima, è che una tale soluzione possa sempre essere scritta come sovrapposizione di soluzioni "molto semplici", che sono quelle che si trovano cercando una soluzione della forma

$$S_t(x) = U(x)\,V(t).$$

Sostituendo nell'equazione otteniamo

$$U''(x)\,V(t) = \frac{\lambda}{T}\,U(x)\,V''(t)$$

e quindi le funzioni U''/U e V''/V sono indipendenti dalle variabili x e t, cioè sono costanti. In particolare, per una certa costante k si ha

$$\frac{U''(x)}{U(x)} = \frac{\lambda}{T}\,\frac{V''(t)}{V(t)} = k,$$

da accoppiarsi coi dati (al momento dimentichiamo i (2), (3))

$$U(0) = U(L) = 0.$$

Equazioni di questo tipo sono state molto studiate e è noto che hanno soluzioni non banali solo se $k < 0$ e $\sqrt{-k}\,L$ è un multiplo intero di π. Tali soluzioni sono multipli costanti di

$$U(x) = \sin(\sqrt{-k}\,x).$$

Analogamente, usando il dato iniziale (3), otteniamo soluzioni tali che $V'(0) = 0$, che sono multipli costanti di

$$V(t) = \cos(\sqrt{-kT/\lambda}\,t).$$

Quindi tutte le soluzioni elementari sono in dipendenza da un intero non negativo $n = \sqrt{-k}\,L/\pi$ e sono multipli (costanti) di

$$2\sin(n\pi/L\,x)\cos(n\pi/L\sqrt{T/\lambda}\,t) = \sin(n\pi/L(x - v\,t)) + \sin(n\pi/L(x + v\,t)),$$

dove $v = \sqrt{T/\lambda}$.

La (2) è soddisfatta se f è una funzione abbastanza regolare prendendo una opportuna combinazione lineare infinita di armoniche del tipo seno. Come prima, se immaginiamo che debba risultare, per $t = 0$,

$$f(x) = a_1\,\sin(x\pi/L) + a_2\,\sin(2\pi/L\,x) + \cdots$$

allora, siccome se $n \neq m$ vale $\int_0^L \sin(n\pi x/L)\,\sin(m\pi x/L)\,dx = 0$, deve essere

$$\begin{aligned}\int_0^L f(x)\,\sin(x\pi/L)\,dx &= \int_0^L (a_1\,\sin(x\pi/L) + a_2\,\sin(2\pi/L\,x) + \cdots)\sin(x\pi/L)\,dx \\ &= a_1 \int_0^L \sin^2(x\pi/L)\,dx + a_2 \int_0^L \sin(2\pi/L\,x)\,\sin(x\pi/L)\,dx + \cdots \\ &= a_1 \int_0^L \sin^2(x\pi/L)\,dx \\ &= a_1\,L/2.\end{aligned}$$

In generale occorre scegliere

$$a_n = \frac{2}{L}\int_0^L f(x)\,\sin(nx\pi/L)\,dx.$$

Le funzioni $\sin(n\pi/L(x - vt))$ e $\sin(n\pi/L(x + vt))$ sono funzioni sinusoidali che si propagano con velocità v nella direzione x. Ricordiamo che la velocità $v = \sqrt{T/\lambda}$, quindi cresce al crescere della tensione della corda e decresce al crescere della sua densità lineare (ad esempio al crescere della sua sezione). Lo spostamento dalla posizione di equilibrio è perpendicolare alla direzione di propagazione x.

Filmino delle funzioni e di qualcuna delle loro sovrapposizioni, che mima il comportamento di una corda pizzicata a metà e a un terzo.

Le particelle dell'atmosfera vicine alla corda che si muove entrano in vibrazione e trasmettono la perturbazione (compressione e rarefazione) a quelle vicine oscillando intorno alla loro posizione di equilibrio. Si creano quindi nel tempo corrispondentemente onde sonore che sono sovrapposizioni di "onde sonore semplici".

Esempi: chitarra e violino.

Dopo useremo anche diapason e flauto non sono di questo tipo, (anche se l'equazione cui si perviene è sostanzialmente uguale), voce umana.

4. Caratteristiche del suono

Le caratteristiche fondamentali di ogni suono da un punto di vista della percezione sono tre: l'Altezza, l'Intensità ed il Timbro. Delle prime due abbiamo già parlato: si possono identificare, da un punto di vista matematico e fisico, con l'ampiezza e la frequenza.

Tabella 2. Parametri del suono

Parametro Percettivo	Parametro Fisico	Rappresenta
Altezza	Frequenza	Tonalità Audio (bassi, medi, alti)
Intensità	Ampiezza	Volume
Timbro	Spettro	Tipologia di strumento

Ben poche sorgenti sonore danno un suono puro (crearlo con Audacity oppure col diapason), in generale si ha a che fare con suoni complessi non rappresentabili con una semplice sinusoide; tuttavia i suoni sono sempre composti da un numero variabile di suoni sinusoidali; se i componenti del suono sono costituiti da una frequenza "fondamentale" e da un numero finito (o infinito) di frequenze che sono multipli della frequenza fondamentale il suono risultante è un suono periodico o armonico (in generale piacevole). Un rumore (generare con Audacity rumore bianco) ha invece una distribuzione di frequenze "casuale" e l'onda che ne risulta è aperiodica.

4.1. **Intensità o Ampiezza.** Nella Figura 7 sono rappresentati tre suoni armonici puri aventi uguale frequenza ma ampiezza differente. In pratica, essi presentano la stessa altezza/nota ma hanno un volume differente.

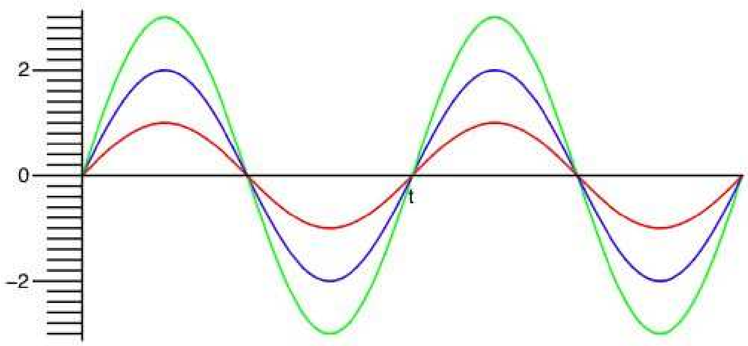

Figura 7. Stessa altezza, ma diverso volume

Tabella 3. Tabella frequenze

Frequenza	Tipologia toni
20-500 Hz	Bassi
500-8000 Hz	Medi
8000-20000 Hz	Acuti

4.2. **Altezza o frequenza.** Un suono (complesso) qualsiasi contiene molte frequenze. Affinché in un suono si possa individuare una frequenza speciale, che caratterizza la sensazione globale di gravità/acutezza trasmessa dal suono, occorre che il segnale sia periodico. I suoni prodotti da strumenti musicali hanno questa proprietà, come abbiamo visto nell'esempio della chitarra o del violino.

Si definisce ottava l'intervallo musicale tra due "do" consecutivi. Dalla metà del settecento tutti gli strumenti ad intonazione fissa (pianoforte, organo,..) sono accordati secondo la scala temperata. Questa scala divide l'ottava in 12 intervalli ognuno dei quali vale un semitono. Ogni semitono corrisponde alla frequenza del semitono precedente moltiplicata per $2^{1/12}$, in modo tale che la prima nota dell'ottava seguente abbia frequenza doppia della nota corrispondente nell'ottava precedente.

TABELLA 4. Le frequenze e le note dell'ottava fondamentale

Nota		Frequenza (in Hz)
la	A	440 $(= 440 \times 2^{0/12})$
la#	A#	466.2$(= 440 \times 2^{1/12})$
si	B(o H)	493.8$(= 440 \times 2^{2/12})$
do	C	523.2$(= 440 \times 2^{3/12})$
do#	C#	554.4$(= 440 \times 2^{4/12})$
re	D	587.3$(= 440 \times 2^{5/12})$
re#	D#	622.2$(= 440 \times 2^{6/12})$
mi	E	659.2$(= 440 \times 2^{7/12})$
fa	F	698.4$(= 440 \times 2^{8/12})$
fa#	F#	740.0$(= 440 \times 2^{9/12})$
sol	G	784.0$(= 440 \times 2^{10/12})$
sol#	G#	830.6$(= 440 \times 2^{11/12})$

Il diapason suona appunto il la dell'ottava più importante (LA3); la corda "la" della chitarra è invece un LA1 (frequenza 110Hz), nel senso che la sua armonica fondamentale ha proprio (o dovrebbe se mal accordata) frequenza fondamentale 110Hz; ma oltre a questa si vedono altre armoniche (che hanno intensità più basse).

4.3. **Timbro.** Gli strumenti musicali non emettono toni puri (ovvero sinusoidi perfette, che sono anche sgradevoli) ma sono caratterizzati da suoni con forme d'onda molto differenti tra loro anche quando suonano la stessa nota. Un suono complesso è il risultato della sovrapposizione di più onde armoniche con frequenza, ampiezza e fase diverse. Il timbro, dal punto di vista percettivo, è il suono caratteristico di uno strumento. Una nota suonata da una pianoforte avrà un timbro differente rispetto alla stessa nota prodotta da un violino o da un flauto. È determinato dalle caratteristiche fisiche dello strumento, quali il mezzo utilizzato per produrre il suono (corde, pelle, ecc.).

Il timbro, dal punto di vista fisico-matematico, è legato alla forma d'onda, ossia dalle armoniche presenti che la descrivono. Generalmente il timbro viene riconosciuto più facilmente solo all'inizio della nota (attacco), mentre è più difficile da distinguere se il suono viene prolungato, perché compaiono altri fattori.

Abbiamo visto nel caso della chitarra, o di uno strumento a corda, che vengono generate insieme alla frequenza fondamentale altri modi semplici di vibrazione della corda, che però hanno necessariamente frequenza multipla di quella fondamentale.

Le armoniche di un suono puro (ovvero di un segnale sinusoidale ad una certa frequenza che chiameremo fondamentale) sono i suoni (segnali) di frequenza multipla di quella fondamentale.

Ad esempio, il “do3” astratto avrebbe frequenza 261.6. Verifichiamo a cosa corrispondono le sue armoniche principali.

- seconda armonica 261.6*2 = 523.2 Do dell’ottava successiva
- terza armonica 261.6*3 = 784.8 Sol dell’ottava successiva
- quarta armonica 261.6*4 = 1046.4 Do di due ottave sopra
- quinta armonica 261.6*5 = 1318 Mi di due ottave sopra

Questo spiega perché le tre note “suonano bene assieme” (e formano l’accordo di “do-maggiore” do-mi-sol-do): suonando un do in realtà si emettono frequenze (di intensità minore) corrispondenti a un mi e a un sol. La forma d’onda dell’accordo che ne risulta è regolare, come si vede nella Figura 8.

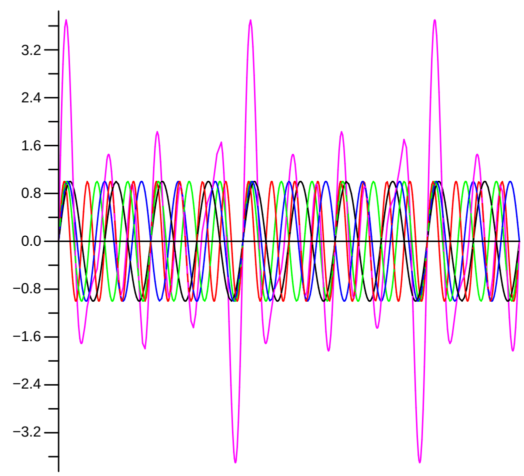

FIGURA 8. In nero, rosso, blu e verde i toni puri corrispondenti a do-mi-sol-do; in magenta l’onda risultante

Un accordo invece “dissonante” è invece ad esempio: do-fa#-si-fa-si♭ la sua forma d’onda è molto più irregolare, come si nota nella Figura 9 e il suono più aspro.

5. Applicazioni

Molte sono le applicazioni di quanto abbiamo visto, fra le più note citiamo gli algoritmi di compressione per file contenenti voce, musica, . . . (MP3, VoIP,. . .), lo studio di sintetizzatori per la composizione di brani musicali, il riconoscimento della qualità di strumenti musicali.

Alla base di queste applicazioni c’è la digitalizzazione del segnale sonoro.

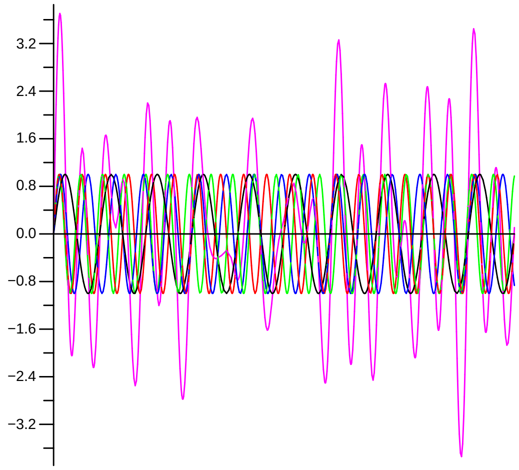

FIGURA 9. In nero, rosso, blu e verde i toni puri corrispondenti a do-fa♯-si-fa-si♭; in magenta l'onda risultante

5.1. **Digitalizzazione dei segnali.** Per passare da un segnale analogico a un segnale digitale (ovvero a un file contenente una schiera di zeri e uni) occorrono due operazioni. La prima consiste nel trasformare l'onda (continua) in una sequenza di "fotografie" del segnale, valutato a istanti distinti. Questa operazione si chiama **campionamento**.

La seconda operazione consiste nel trasformare ciascuna "fotografia" (ciascun campione) del segnale, che in linea di principio può assumere infiniti valori, in una "fotografia" che possa avere solo un numero finito di possibili valori. In questo modo siamo in grado di associare una sequenza binaria finita a ciascuno di questi valori. Questa operazione si chiama **quantizzazione**.

Ovviamente il segnale digitale è un'approssimazione di quello analogico di origine, ossia, digitalizzando, si perdono informazioni/qualità. D'altra parte, esistono limiti fisici anche per i segnali analogici, per cui è inutile andare oltre un certo grado di fedeltà all'originale. Quindi, oltre un certo livello, posso assumere che il segnale digitalizzato sia sempre fedele a quello originale.

In particolare, esistono criteri per stabilire i parametri di digitalizzazione. Per il campionamento, ci interessa stabilire quanti campioni vanno presi nell'unità di tempo di 1 secondo (frequenza di campionamento, misurata in Hz); per la quantizzazione: ci interessa stabilire quale sia il range di valori per ogni campione (dinamica), da cui calcolare quanti possibili valori desideriamo poter distinguere, ossia la risoluzione o profondità.

Ovviamente più è alta la frequenza di campionamento, migliore è l'approssimazione; più alta è la dinamica, migliore è l'approssimazione. Tuttavia questo si traduce anche in molti più dati da memorizzare e da trattare, quindi file molto grossi. Conviene allora usare alcuni criteri che ci aiutino a stabilire qual è il minimo teorico per non avere perdita di qualità, e non campionare o esagerare nel numero di bit per la dinamica.

Per la frequenza di campionamento, si devono considerare le frequenze presenti nel segnale (=banda). Per la musica ad alta fedeltà occorre conservare le frequenze almeno fino a 20 KHz. Per la comprensibilità della voce, basta limitarsi a 8 KHz.

C'è il Teorema di incertezza di Heisenberg che ci dice che un segnale non può essere localizzato sia in tempo sia in frequenza (cioè se un suono si esaurisce nel tempo, esso sarà la sovrapposizione di un numero infinito di frequenze). Tuttavia, ci basta ottenere una buona approssimazione, almeno per il nostro orecchio, del suono analogico. Ci viene in aiuto il Teorema del Campionamento o di Nyquist-Shannon che dice che la frequenza del campionamento deve essere almeno il doppio della banda, e quindi:

la musica va campionata a 40 KHz (lo standard CD campiona a 44.1 KHz)

il parlato va campionato a 16 KHz.

Per quanto riguarda la dinamica, non c'è un criterio così semplice. La dinamica dipende da due fattori: la massima estensione possibile del segnale (valore massimo valore minimo), e la minima differenza di segnali che considero significativa. Per esempio, nel digitalizzare una voce, occorre considerare l'intensità corrispondente all'urlo massimo (che sono disposto a sentire) e l'entità del minimo sussurro (che voglio rilevare). Quest'ultima grandezza non può essere piccola a piacere, perché per esempio sotto una certa entità sarà più forte il rumore ambientale del mio segnale. Note queste due grandezze, che possiamo chiamare rispettivamente U e S , la dinamica si calcola come segue: dinamica $= 20 \log_{10}(U/S)$ e, come abbiamo detto, si esprime in decibel. Sappiamo che ogni bit rappresenta un raddoppio nel numero di parole rappresentabili. Infatti in base 2 ogni cifra in più raddoppia la gamma di valori rappresentabili, come in base 10 ogni cifra in più decuplica la gamma di valori rappresentabili.

Inoltre, possiamo calcolare che un raddoppio, espresso in decibel, è circa 6 dB. Infatti $20 \log_{10} 2 = 6.0206$ circa. Allora la regola per calcolare la risoluzione o profondità a partire dalla dinamica si può esprimere come segue: le parole di codice necessarie per digitalizzare un segnale che ha dinamica X dB devono essere rappresentate con numeri binari a X / 6 bit.

Questa è la regola e sembra abbastanza semplice. Il problema risiede nel fatto che è spesso difficile stabilire la dinamica necessaria. Per l'audio ad alta fedeltà, convenzionalmente si decide che 100 dB di dinamica sono sufficienti (anche eccessivi), quindi si dovrebbero usare $100/6 = 16.6667 = 17$ bit. In realtà il CD si "accontenta" di 96 dB, con cui si arriva a 16 bit.

5.2. **La compressione.** Un modo tipico per trattare i dati multimediali è la compressione, per ridurre la dimensione dei file che li contengono. Comprimere un segnale significa rappresentarlo in una forma che consenta l'eliminazione di una parte dei dati. Esistono due tipi di compressione, corrispondenti a due tipi di dati eliminati:

compressione senza perdita: viene eliminata una parte dei dati, il cui contenuto informativo si può però ricavare dai dati restanti. Viene quindi ridotta la ridondanza dei dati, ma ciò che viene eliminato si può ricostruire;

compressione con perdita: viene eliminata una parte dei dati contenenti informazione che non si può ricavare dai dati restanti. In questo caso viene effettivamente persa informazione, e la bontà del metodo di compressione sta nel fatto di eliminare informazione poco percettibile.

Un esempio di compressione senza perdita è: se incontro 10 pixel con livello di luminosità 112, anziché scrivere 10 volte il numero 112 scrivo solo i numeri 10 e 112. Questo è possibile e vantaggioso per immagini create al calcolatore, cioè puramente artificiali, per le quali è possibile che ampie aree abbiano esattamente lo stesso colore.

Un esempio di compressione con perdita è: se ho un'immagine 500 x 500 e butto via una riga sì e una riga no, ottengo un'immagine più piccola; se replico le righe dispari e le metto in posizione pari ottengo ancora un'immagine 500 x 500, ma avendo memorizzato solo 250 x 500 valori. Ovviamente bisogna vedere se percettivamente il risultato è ancora buono.

Abbiamo distinto due categorie di suoni da trattare digitalmente: la voce e la musica.

Lo studio della comunicazione vocale è l'oggetto della telefonia. La telefonia attraverso il calcolatore si appoggia alla trasmissione su Internet e si chiama Voice Over IP (VOIP). Nella comunicazione vocale l'intelligibilità delle parole è il parametro più importante. Alcune frequenze vengono eliminate proprio per rendere massima l'intelligibilità, in quanto non sono utili. Per esempio, gli acuti oltre una certa frequenza (sopra i 3500 Hz) e i bassi (fino a 300 Hz) si possono e vanno eliminati. Questo significa che una frequenza di campionamento, per esempio, di 8000 Hz = 8 kHz è adeguata. Per quanto riguarda la dinamica, essa non è estrema come nelle applicazioni musicali, e spesso bastano 8 bit.

Le applicazioni sonore che riguardano la musica sono quelle che richiedono il massimo di qualità possibile. Questo significa banda passante almeno di 20 kHz (e quindi frequenze di campionamento minime di 40 kHz, ma ci sono standard a 48 kHz, 96 kHz e 192 kHz), e risoluzione di almeno 16 bit (ma ci sono sistemi a 24 bit). Inoltre i segnali di questa categoria sono sempre almeno stereofonici (ma spesso anche 5.1), il che significa che lo spazio occupato da un canale va poi moltiplicato per 2, 5 o 6 per calcolare l'occupazione effettiva. Per finire, essendo orientati all'intrattenimento, tali segnali non si limitano ad effetti sonori di pochi secondi o a singoli messaggi vocali, ma richiedono di memorizzare da minuti a ore di segnale. Queste caratteristiche hanno reso opportuna l'invenzione di una serie di tecniche di compressione con perdita efficiente, che mantenesse la qualità sonora pur riducendo significativamente la dimensione dei dati da memorizzare. L'esempio più di successo di tecniche di questa categoria si chiama MP3. Esso è un algoritmo che è stato inventato per codificare l'audio del video nel formato MPEG.

MP3 (o, più esattamente "MPEG-1/2 Audio Layer 3" (MPEG=Moving Picture Experts Group)) è un algoritmo di compressione audio in grado di ridurre drasticamente la quantità di dati richiesti per riprodurre un suono, rimanendo comunque una riproduzione fedele del file originale non compresso.

Compressione: si tralasciano alcune frequenze, effetto mascheramento delle frequenze "forti" rispetto a quelle vicine più deboli.

5.3. **Wavelets.** Nuove tecniche di manipolazione dei suoni sono comunque allo studio e è già uscito lo standard MPEG-4, che è basato sulla decomposizione del segnale sonoro in altre funzioni, chiamate “ondine”, anziché in serie (o integrale) di seni e coseni. Usano questa compressione le suonerie dei cellulari AAC.