5 - La matematica del linguaggio musicale

Il linguaggio musicale scritto nasce quando comincia a farsi impellente la necessità di *conservare* la musica, sia perché la si vuole trasmettere, sia perché ce n'è troppa perché possa essere semplicemente memorizzata. Le informazioni che in qualche modo devono essere codificate riguardano l'altezza (pitch) dei suoni, la loro durata, l'intensità, il timbro (vocale o strumentale) e eventuali altre caratteristiche legate ai vari meccanismi di produzione.

Prima di analizzare il linguaggio musicale che tutti (si spera) conoscono, esaminiamo le caratteristiche dei vari linguaggi che si sono presentati nel corso della storia della musica: vi sono linguaggi di tipo approssimato, come quello neumatico del canto gregoriano, rappresentati da una sorta di stenografia, utile a ricordare una linea melodica e ritmica già conosciuta, ma difficilmente utilizzabile per proporne una nuova, se non all'interno di un *corpus* ben noto. La notazione neumatica rappresenta graficamente in modo qualitativo l'andamento della voce, sia per quanto riguarda l'altezza che la durata dei suoni; occorre tener presente che nel canto gregoriano, quasi esclusivamente monodico, l'estensione della voce non supera un'ottava.

Esistono poi dei linguaggi, utilizzabili solo su uno strumento, che rappresentano graficamente le posizioni delle dita sulla tastiera del cembalo o dell'organo, sulla tastiera del liuto, sui buchi del flauto, le cosiddette *intavolature*, che ancora oggi vengono presentate con lo slogan **Imparate a suonare (il flauto dolce, la chitarra, il pianoforte) senza conoscere la musica**: è inutile dire che se non si conosce esattamente l'accordatura dello strumento è impossibile capire di che musica si tratti.

Ci sono anche linguaggi, musicalmente molto più colti, che si potrebbero definire linguaggi armonici, nel senso che forniscono al suonatore di piano, chitarra o fisarmonica, le informazioni necessarie per trovare gli accordi da utilizzare per l'accompagnamento (l'armonizzazione) di una melodia, talvolta neppure rappresentata perché si suppone che l'accompagnatore la conosca; l'invenzione di questo linguaggio, il cosiddetto *basso numerato* viene attribuita ad Adriano Banchieri (Bologna 1567-1634).

Nelle intavolature è relativamente facile indicare con precisione l'altezza del suono (o dei suoni) da produrre, più difficile indicarne la durata: si possono inventare dei codici grafici, o numerici, che però non sono di immediata lettura: un linguaggio musicale dovrebbe avere la capacità di permettere l'esecuzione *in tempo reale*, cioè la lettura dei simboli dovrebbe richiedere un tempo molto più breve della loro durata; e dovrebbe essere indipendente dallo strumento utilizzato. Occorre anche distinguere tra un linguaggio dedicato ad un unico esecutore e uno valido per un gruppo di esecutori ai quali vengono affidate musiche diverse: in questo caso diventa fondamentale una rappresentazione precisa delle durate dei suoni che permetta al gruppo di andare "a tempo". Proviamo ad immaginare un linguaggio: supponendo di avere un numero limitato di suoni dobbiamo associare ad ogni altezza un simbolo, numerico, alfabetico o iconico, ed un altro simbolo, per esempio numerico, che ne indichi la

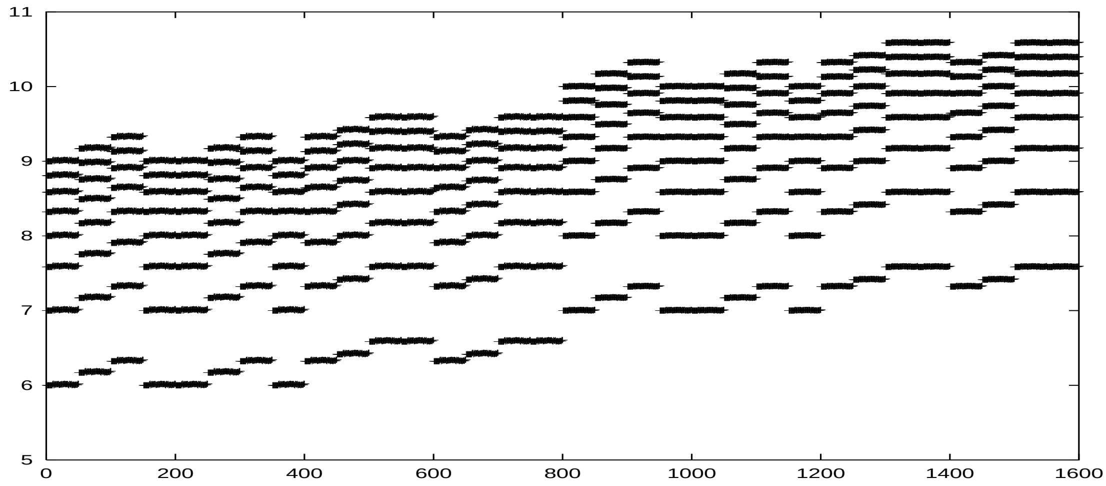

Fig. 1: Pseudosonogramma di "Fra Martino campanaro" con frequenze log.

durata: C1,D1,E1,C1,C1,D1,E1,C1, E1,F1,G2,ecc sono le prime tre battute di "Fra Martino campanaro", in cui si è adottata la notazione letterale anglosassone ed un numero proporzionale alla durata di ciascun suono; certamente non è comodissimo e diventa assai complicato se vi sono suoni simultanei da eseguire sulla tastiera, e se le durate hanno valori molto diversi. La Fig.1 ci mostra lo stesso pseudosonogramma del capitolo precedente con le frequenze in scala logaritmica, la quale evidenzia molto bene l'invarianza traslazionale (in senso verticale) della distribuzione spettrale dei suoni periodici; l'informazione sull'andamento melodico è tutta contenuta nel suono fondamentale, e se si collegano i diversi trattini con una linea continua si ottiene qualcosa di molto simile ad una successione di neumi. La scrittura neumatica ci dà un suggerimento: rappresentare le altezze in modo posizionale, però con riferimento ad una posizione fissa, un'origine, fissata per esempio da un rigo orizzontale: tuttavia i neumi sono poco precisi, ma anche se lo fossero l'occhio che li legge è anch'esso impreciso, e da queste due circostanze (oltre che dalla pratica musicale) nasce l'esigenza di *discretizzare* le altezze, introducendo altre righe orizzontali (non troppe né troppo poche!) e stabilendo la loro *distanza* in altezza. In questo modo si potrebbero scrivere sulle varie righe, o negli spazi tra una riga e l'altra, i numeri corrispondenti alle durate: chi legge questa notazione musicale ha una componente posizionale (l'altezza) ed una numerica (la durata); per entrambe occorre stabilire un punto di riferimento, cioè l'altezza corrispondente ad una certa riga (ad esempio quella centrale, o quella rossa) e l'unità di misura delle durate: ma che cosa significa 1? Un secondo, un respiro, un battito del cuore? Occorre cioè una convenzione che le stabilisca, o una volta per tutte o di volta in volta; vedremo che si presentano entrambi i casi.

Cerchiamo per prima cosa di organizzare le altezze, cioè di definire con un qualche criterio il significato delle varie righe: se dovessimo procedere secondo un criterio fisico-matematico potremmo decidere che le righe definiscono una coordinata in uno spazio definito dalla frequenza o dal logaritmo della frequenza e che la distanza tra due righe è una costante corrispondente, per esempio, alla più piccola distanza utilizzata musicalmente. Successivamente, essendoci accorti che la musica è Doppler-invariante, sceglieremmo la scala logaritmica, e utilizzeremmo il semitono come distanza tra due righe. Le cose sono andate in modo leggermente diverso: la scala è logaritmica, ma la distanza tra due righe non è costante, e per di più si utilizzano sia le righe che gli spazi tra una riga e l'altra; quest'ultimo fatto rende più agevole la lettura perché permette di dimezzare il numero di righe necessario per coprire un certo intervallo. Invece la distanza non costante dipende dal fatto che le melodie dell'antichità utilizzavano una divisione dell'ottava in parti non uguali, chiamate toni e semitoni, e le righe e gli spazi permettono la collocazione ordinata delle note della scala, indipendentemente dalla posizione dei semitoni, a patto che si accettino le *alterazioni*, quei segni speciali, diesis $\sharp$ e bemolle $\flat$, che innalzano e abbassano di un semitono l'altezza della nota che precedono.

Che cosa disegnare sulle righe o negli spazi? Dei numeri o dei simboli equivalenti alle durate? Quali sono le durate che vengono effettivamente utilizzate? Nella pura recitazione la durata dei suoni, ovvero delle sillabe, è relativamente libera e difficilmente può essere costretta in uno schema numerico semplice; nella recitazione poetica, metrica o ritmica, e a maggior ragione nella musica di accompagnamento della danza, le durate sono più precise e spesso sono multiple intere di una durata elementare, anzi addirittura multiple secondo una potenza di 2 o di 3 o un loro prodotto. Questo deriva probabilmente dal fatto che molte nostre azioni ripetute, il passo, il respiro, il nuoto, il lavoro (martellare, spalare, vangare, falciare, strofinare) comportano una divisione del periodo in due parti uguali, oppure in due parti diverse che si possono considerare con buona approssimazione una doppia dell'altra, in modo che il periodo viene diviso stavolta in modo ternario. Viene quasi naturale (a posteriori tutti son buoni a dire che era facile!) inventare una serie i simboli ognuno dei quali corrisponda ad una frazione, binaria o ternaria, del simbolo precedente.

Il risultato è la *notazione mensurale* in cui i simboli sono *maxima, longa, breve*, ognuno dei quali vale la metà o un terzo del precedente; successivamente, per permettere una più facile scrittura della musica, sono stati inventati nuovi simboli *semibreve, minima*, e ancora dopo sono stati introdotti *semiminima, croma, semicroma, biscroma, semibiscroma* ma sono stati abbandonati i primi, visto che una maxima equivale a 512 semibiscrome e non c'è musica che richieda una tale abbondanza di simboli; in pratica se ne usavano quattro, dalla maxima alla semibreve, e quindi occorreva definire tre rapporti, *maxima:longa, longa:breve, breve:semibreve*, chiamati *modo, tempo, prolazione* ognuno dei quali poteva valere 2 o 3. Ci sono in teoria 8 possibilità: 3:3:3 - 3:3:2 - 3:2:3 - 2:3:3 - 3:2:2 - 2:3:2 - 2:2:3 - 2:2:2, a ciascuna delle quali corrispondeva un simbolo collocato all'inizio del rigo musicale. Motivi teologici facevano considerare *perfetta* la divisione ternaria (la Santissima Trinità), *imperfetta* quella binaria, anche se quest'ultima veniva considerata più naturale dai teorici medioevali (mentre quella perfetta era forse "soprannaturale").

Un residuo di questa notazione lo troviamo oggi nelle indicazioni che stabiliscono la divisione della battuta e la collocazione degli accenti forti e deboli: 2/2, 4/4, 3/4, 6/8, 9/8, 12/8. Se ci limitiamo alle situazioni omogenee (3:3:3 e 2:2:2) i simboli indicano una scala temporale di tipo logaritmico, in base 3 o in base 2.

Vediamo allora che la notazione mensurale equivale ad un doppio codice logaritmico, nella sua versione più moderna in base 2 per le durate, in base 2 anche per le altezze, ma con una ulteriore divisione disomogenea; come tutti i codici logaritmici richiede la definizione dell'unità di misura, cioè l'altezza di una nota di riferimento e la durata di riferimento. Per la prima occorre stabilire il **La**, che è il suono prodotto da una certa canna dell'organo della chiesa del villaggio, il che produce una notevole varietà di La tra i diversi villaggi e le varie epoche, finché non si arriva a definire il La per mezzo di una convenzione internazionale, anch'essa suscettibile di variazioni, che lo fissa a 440Hz. Per le durate invece si decide che uno dei simboli corrisponde al *tactus*, il gesto più ampio del maestro di cappella, forse collegato al suo battito cardiaco; oggi usiamo uno strumento preciso, qual'è il metronomo, scrivendo che un certo simbolo dura una certa frazione di minuto.

Non so quanto i musicisti siano consapevoli del fatto che la notazione mensurale rappresenta per la musica una conquista analoga a quella che lo zero e i codici numerici posizionali (decimale, binario, o altro) hanno rappresentato per la matematica. La notazione mensurale, per mezzo di un numero limitato di simboli, permette la lettura (e talvolta l'esecuzione) in tempo reale anche di un numero relativamente elevato di suoni simultanei, ha incorporate le caratteristiche di Doppler-invarianza della musica, è universale, nel senso che non è legata ad un particolare strumento, ma non è capace di andare oltre alla divisione dell'ottava in dodici semitoni, e non fornisce informazioni sulla dinamica né sul timbro. Solo per la dinamica esiste una notazione simbolica aggiuntiva, anch'essa di carattere logaritmico, rappresentata dai simboli *...ffff, fff, ff, f, mf, mp, p, pp, ppp, pppp,...*, ognuno dei quali suggerisce un'intensità forse doppia del successivo.

Che la notazione delle durate segua un codice logaritmico in base due è comprensibile per quanto detto in precedenza; più difficile è capire perché anche le altezze seguano la stessa base, cioè perché due suoni distanti una o più ottave abbiano lo stesso nome, e perché l'ottava venga divisa in 12 parti. Un codice in cui la distanza (costante) fra le righe sia diversa dal semitono, e le noti distanti un'ottava e mezzo abbiano lo stesso nome è altrettanto semplice ed elegante, ma non è adatto per scrivere la musica quale si è sviluppata, almeno fino alla metà del '900. Perché? Questa domanda inevitabilmente ne trascina con sé molte altre: perché certi intervalli sono consonanti ed altri no, in che senso l'ottava è più privilegiata di altri intervalli consonanti? Qual'è il livello di tollerabilità del sistema temperato? Cercherò di rispondere ad alcune di queste domande ricordando, per cominciare, che la teoria pitagorica della consonanza è in realtà una giustificazione puramente numerologica di un fatto acquisito, cioè che il sistema musicale greco utilizzava i tetracordi, costruiti con toni e semitoni, e che l'intervallo di ottava e di quinta risultavano particolarmente gradevoli: scoprire che, a parità di altre condizioni, il rapporto 2:1 tra la lunghezza di due corde (o di due flauti) dà luogo all'ottava e il rapporto 3:2 dà luogo alla quinta giusta, fornisce un sostegno ad una teoria secondo la quale i rapporti semplici (costruiti con numeri interi piccoli) tra le lunghezze di corde e tubi sono alla base del criterio di consonanza. Secondo me la numerologia pitagorica descrive ma non spiega la consonanza, a meno che non venga sostenuta da altre considerazioni, che riguardano la natura dei suoni e la struttura del nostro apparato uditivo.

Gli antichi greci non sapevano che i rapporti tra le lunghezze erano uguali al rapporto tra i periodi di vibrazione: non avevano strumenti di misura abbastanza precisi e potevano forse avere un'idea solo qualitativa di questa relazione. Una relazione numerica precisa tra frequenze è certamente importante ma non giustifica un criterio di consonanza: se, come spesso si sostiene, il nostro orecchio è un banco di risuonatori in buona misura indipendenti, due suoni puri simultanei stimolano due neuroni e non si capisce perché certe coppie di neuroni siano *più uguali* di altre. Noi però sappiamo che i suoni sono quasi sempre accompagnati dal loro seguito di armonici, e quindi ogni suono eccita numerosi neuroni, e due suoni simultanei eccitano un numero doppio di neuroni, a meno che alcuni degli armonici non siamo comuni ad entrambi i suoni. Questo fatto suggerisce una giustificazione al criterio di consonanza numerologico: prendiamo ad esempio l'unisono, con tutti gli armonici in comune, e un intervallo irrazionale, con nessun armonico in comune; tra questi due estremi possiamo esaminare il comportamento di alcuni intervalli, e vedere se il conteggio degli armonici in comune fornisce un risultato soddisfacente. Ci accorgiamo subito di una difficoltà: gli armonici sono infiniti, ma non hanno lo stesso peso e solo un numero finito, diciamo n, supera una certa soglia; a questo punto possiamo valutare la frazione di armonici comuni (da tutti a nessuno) o il numero di neuroni coinvolti (da n a $2n$), e eventualmente tener conto anche delle ampiezze dei vari armonici. Se il secondo suono è il p-esimo armonico del primo, con $p \leq n$, la frazione di armonici in comune è $1/p$, mentre i neuroni eccitati sono $n(2 - 1/p)$, due quantità strettamente correlate. Se invece il secondo suono ha una frequenza multipla razionale del primo p/q con $p > q$ e primi tra loro il risultato è lo stesso; beninteso se n è relativamente piccolo, di n/p bisogna prendere la parte intera. Secondo questo criterio gli intervalli più piccoli dell'ottava in ordine di consonanza decrescente sono: quinta giusta $p/q = 3/2$, quarta giusta $4/3$, sesta maggiore $(Sol - Mi)5/3$, terza maggiore $(Do - Mi)5/4$, terza minore $(Mi - Sol)6/5$, settima minore $(Do - Si\flat)7/4$, quinta diminuita $(Mi - Si\flat)7/5$, terza minore $(Sol - Si\flat)7/6$, sesta minore $(Mi - Do)8/5$, seconda maggiore $(Si\flat - Do)8/7$, settima minore $(Mi - Re)9/5$, terza maggiore $(Si\flat - Re)9/7$, seconda maggiore $(Do - Re)9/8$, ecc.; da notare che alcuni intervalli compaiono più di una volta, per esempio la terza minore, ma si tratta di intervalli leggermente diversi, che diventano uguali solo nel sistema temperato. Inoltre il livello di consonanza dipende solo da p quando non si tiene conto dell'intensità degli armonici, mentre quando questa viene presa in considerazione c'è anche una dipendenza da q. Un risultato simile a quello che si ottiene in media con le valutazioni soggettive è stato presentato in una figura del capitolo precedente: i suoni che venivano confrontati avevano 12 armonici con ampiezze $1/p$ e una larghezza di

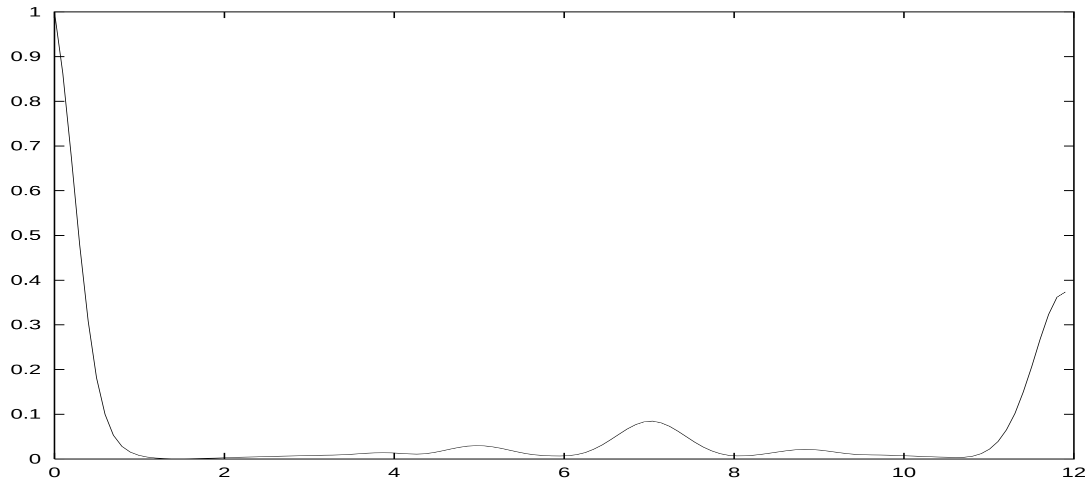

Fig. 2: Livello di consonanza, con ampiezze $1/p^3$ e larghezza 1/4 di tono.

circa un quarto di tono. Nelle figure che seguono si può apprezzare il ruolo degli armonici e del potere risolutivo dell'orecchio nella valutazione della consonanza: la Fig.2 mostra infatti il livello di consonanza all'interno dell'ottava calcolato con suoni dotati di 12 armonici, ampiezze degli armonici $1/p^3$, larghezza un quarto di tono, mentre la Fig.3 mostra il livello di consonanza con ampiezze $1/p$ ma larghezza sei volte maggiore.

Questo degli armonici in comune è, più o meno, il criterio di Helmholtz, il quale rende conto abbastanza bene di quelle che sono, in media, le valutazioni soggettive sulla consonanza dei vari intervalli: anche questo criterio tuttavia non individua le cause della consonanza, a meno che non venga utilizzato per valutare il numero dei neuroni coinvolti; cè tuttavia un punto non spiegato, e cioè il fatto che due suoni puri, privi di armonici, o sono coincidenti e hanno l'unico armonico in comune, o non lo sono e non ne hanno nessuno, eppure la valutazione soggettiva della consonanza conduce agli stessi risultati esposti sopra, anche se in misura molto meno marcata, come se certe dissonanze fossero meno fastidiose quando non ci sono armonici di mezzo. Manca ancora qualche cosa, che probabilmente riguarda il nostro sistema uditivo e l'influenza che su questo hanno le esperienze uditive, ma certamente la presenza degli armonici è un fatto assai importante, dal quale cercheremo di ricavare tutte le conseguenze possibili.

Diamo per scontato che l'ottava è un intervallo privilegiato, e consideriamo solo gli intervalli più piccoli dell'ottava ottenuti prendendo tutte le possibili coppie di armonici di un dato suono, dotato di un numero fissato di armonici com ampiezze decrescenti con la legge $1/n$; prendiamo cioè una coppia di interi $p > q$ e grafichiamo il prodotto delle ampiezze $1/pq$ in funzione della parte dopo la virgola del logaritmo in base 2 del rapporto p/q; il risultato si osserva nella serie di figure che seguono: la Fig.4 mostra cosa accade quando si aumenta il numero di armonici:

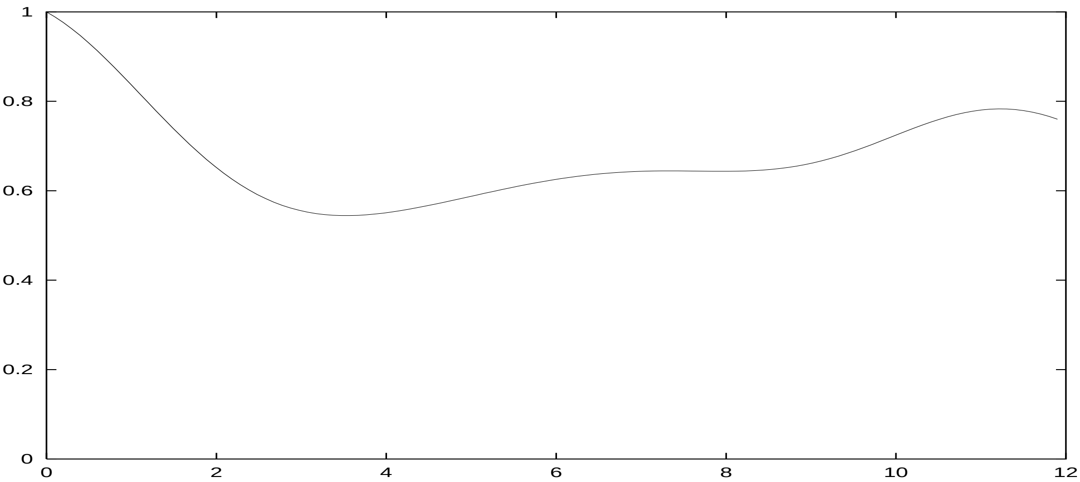

Fig. 3: Livello di consonanza, con ampiezze $1/p$ e larghezza un tono e mezzo

si vede che le righe principali vanno a collocarsi in posizioni discretamente approssimate da una scala che contiene, in ordine di importanza, i seguenti suoni: $Do, Sol, Fa, Mi, La\flat, Re, Si\flat$, e via via tanti altri, in particolare $Do\sharp$ e Si. Limitandoci ai sette più importanti si osserva che, mettendoli in ordine si ottiene una scala simmetrica per la disposizione dei semitoni che nella prima metà coincide con la scala maggiore e nella seconda metà con la scala minore naturale. Le Figg.5,6 presentano, sotto le righe ottenute con 64 armonici, la divisione dell'ottava in 12 o in 53 intervalli uguali. È chiaramente visibile la discrepanza tra il sistema temperato e quello naturale, soprattutto per quanto riguarda la terza maggiore; l'altra divisione dell'ottava non è stata scelta a caso ma risponde alla domanda:" Se 12 quinte approssimano 7 ottave con un eccesso del 2/100 , quale numero di quinte approssima un numero intero di ottave meglio di così?" La risposta è che 41 quinte approssimano per difetto 24 ottave con un errore di poco inferiore, ma 53 quinte approssimano 31 ottave con un eccesso del 3/1000; proseguendo nell'analisi si scopre che 306 e 359 quinte approssimano rispettivamente 179 e 210 ottave con un errore di 1.5/1000, e infine che 665 quinte approssimano 389 ottave con un errore del 3/100000. Si può certamente proseguire e trovare ancora di meglio, ma questi numeri sono pure curiosità, in quanto non è praticabile l'idea di costruire uno strumento in cui l'ottava è divisa in 300 o più parti, anche se, a ben pensarci, il nostro orecchio ha proprio 300 recettori per ottava (circa, perché nessuno li ha contati con precisione!): non sembra tuttavia che a questa ricchezza di tasti corrisponda un effettivo potere risolutivo in frequenza. Contentiamoci del risultato ottenuto con 53 tasti, poiché già si vede che la corrispondenza è molto buona per quasi tutte le note rilevanti della scala. La tabella numerica allegata mostra i rapporti di frequenza con il fondamentale di una scala di 53 intervalli uguali, delle scale temperata, naturale, pitagorica, la coppia di numeri interi il cui rapporto meglio approssima i numeri della prima colonna, il

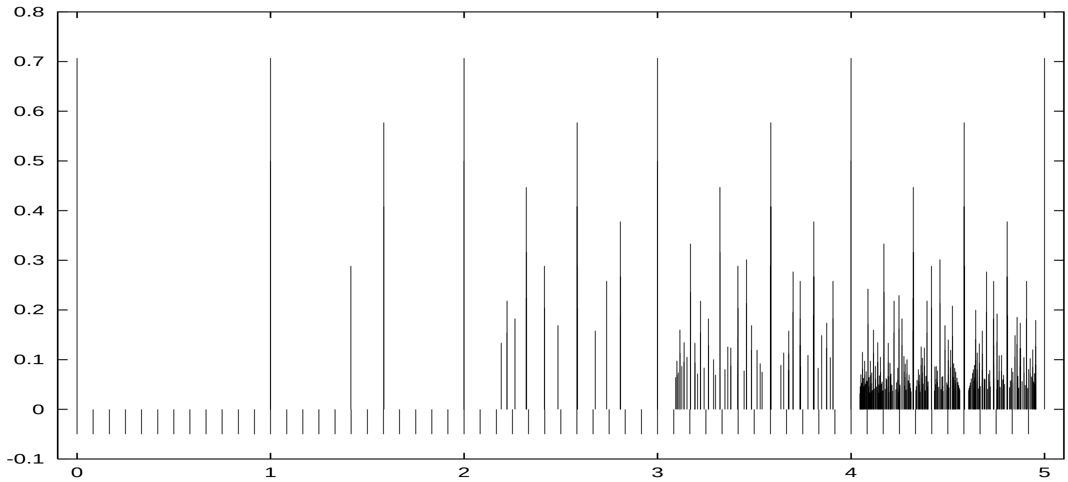

Fig. 4: Intervalli tra gli armonici rapportati all'ottava: da 2 a 32 armonici.

valore in cent (centesimi di semitono temperato).

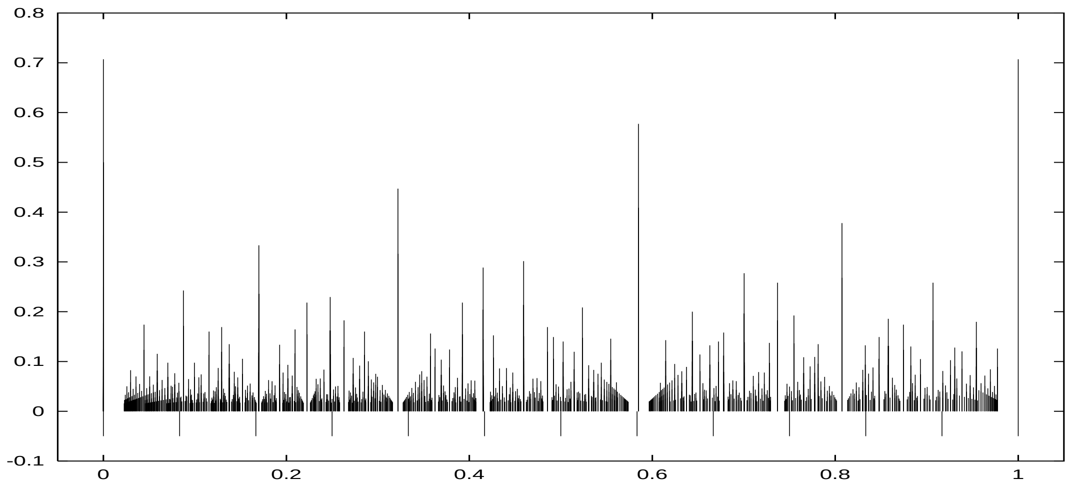

Fig. 5: Intervalli tra 64 armonici rapportati all'ottava divisa in 12 semitoni.

Questa divisione dell'ottava in 53 parti uguali permette, in linea di principio, di risolvere un problema antico quasi quanto la stessa musica, quello del temperamento: è ben noto che per poter eseguire su uno strumento dall'intonazione non modificabile, come potrebbe essere un'organo o anche una qualsiasi tastiera, un brano in qualsiasi tonalità, occorre aumentare il numero dei tasti o rinunciare a certe consonanze perfette. Innumerevoli sono stati i tentativi in entrambe le direzioni: strumenti con più di 12 tasti per ottava sono stati effettivamente progettati e costruiti, ma suonarli non doveva certo essere facile, specialmente in brani che richiedessero una

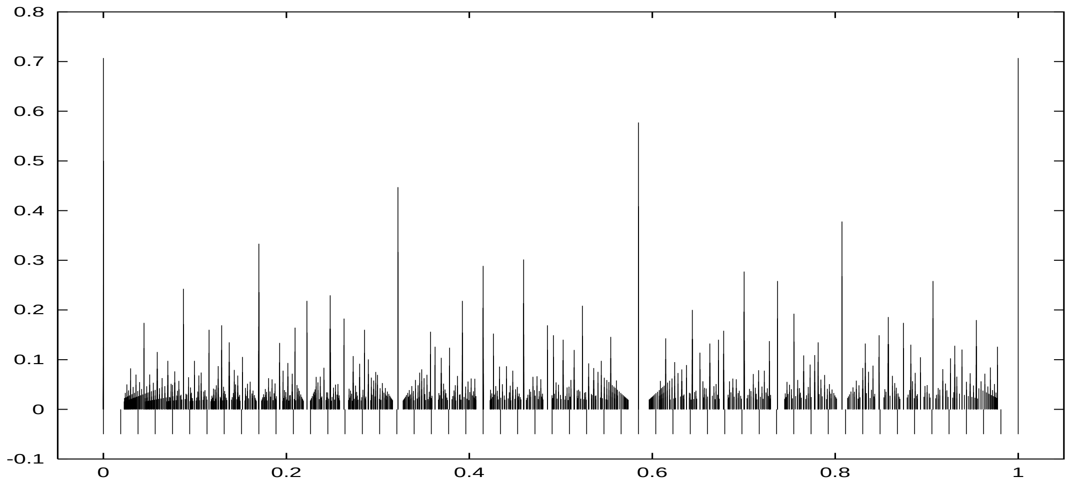

Fig. 6: Intervalli tra 64 armonici rapportati all'ottava divisa in 53 parti uguali.

certa agilità; il problema del temperamento è stato definitivamente risolto con la divisione dell'ottava in 12 semitoni, anche se gli strumenti ad arco, certi strumenti a fiato e la voce umana, conservano la libertà di intonazione nel continuo e possono adattarsi a qualsiasi esigenza. Se si confronta il sistema di 53 intervalli uguali con quello naturale e con quello pitagorico ci si accorge che, scegliendo opportunamente i "tasti" da utilizzare, è possibile approssimare entrambi i sistemi con una precisione ben maggiore di quella con cui i tanti sistemi approssimano la scala naturale: infatti la scala di Do maggiore viene approssimata dai vari sistemi con uno scarto quadratico medio che va dai 10.2 cents del temperameno equabile ai 5.5 cents di quello di Kimberger, mentre con 7 tasti dei 53 possibili lo scarto è di 0.75 cents! Se poi si esaminano le scale cromatiche con una distinzione tra i diesis e i bemolle si vede che con 21 tasti lo scarto è di 2,26 cents per la scala naturale e addirittura di 0.56 cents per quella pitagorica: la ragione di questo successo è dovuta al fatto che il 53-esimo di ottava, pari a 1.013164 è quasi esattamente il comma pitagorico, pari a 1.0136, che distingue il semitono cromatico dal semitono diatonico. Naturalmente una tastiera con 53 tasti per ottava non sarebbe suonabile, ma la tecnologia odierna permette la realizzazione di uno strumento dotato di 53 note (corde, tubi o circuiti elettronici) per ottava, controllata da una tastiera del tutto normale di 12 tasti, con l'aggiunta di un dispositivo che collega i 12 tasti alle 12 note (tra le 53 disponibili) più adeguate al brano che si sta eseguendo: sarebbe in fondo qualcosa di analogo al pannello che sull'organo seleziona i registri accoppiati a ciascuna tastiera e che negli strumenti moderni è addirittura preprogrammabile dall'esecutore che, al momento opportuno, non deve far altro che premere un pulsante per cambiare registro.

Possiamo domandarci:"Se nel 1500 fosse stata disponibile l'odierna tecnologia la musica avrebbe avuto uno sviluppo diverso?" Si possono dare infinite risposte, ma una mi sembra interessante: la disponibilità di uno strumento semplice e poco costoso in cui l'accordatura possa essere variata in corso di esecuzione (la voce umana, se associata ad un apparato uditivo ben addestrato, è capace di questo!) forse non avrebbe stimolato i teorici a escogitare i tanti temperamenti fino a produrre quello equabile, le cui conseguenze sullo sviluppo della musica sono note a tutti: basta pensare che la pratica dell'enarmonia, spesso adottata dai madrigalisti del '500 (in particolare da Gesualdo da Venosa) in ambito vocale, ha dovuto aspettare almeno due secoli per affermarsi anche in ambito strumentale. Spesso le situazioni difficili aguzzano l'ingegno e permettono di trovare soluzioni rivoluzionarie.....

Tornando al discorso principale vediamo che tutte queste considerazioni, per quanto suggestive, ancora non coinvolgono il sistema uditivo e soprattutto non rispondono alla domanda:"Perché l'ottava?" Vediamo allora in che modo possiamo coinvolgere il sistema uditivo in questa analisi: dobbiamo per forza prendere in esame la corteccia uditiva, cioè quella parte del sistema percettivo che può modificarsi sotto l'azione degli stimoli: l'idea è che la stimolazione simultanea di una coppia di neuroni contribuisca a rinforzare le loro mutue connessioni sinaptiche, e quindi, se la cosa avviene sistematicamente, la stimolazione di un solo membro della coppia possa eccitare anche l'altro; è il principio del riflesso condizionato.

Orbene, la Fig.5 non è altro che una riga della matrice di correlazione prodotta da suoni dotati di armonici con pesi $1/n$, e possiamo immaginare che i pesi sinaptici siano organizzati in modo simile: allora l'ottava si presenta come la correlazione più forte; se un sistema uditivo si sviluppa in un ambiente dominato da suoni periodici la sua struttura sinaptica potrebbe essere simile a questa; se invece si sviluppa in un ambiente dominato da rumori la riga della matrice di correlazione presenterà un andamento molto più regolare che non evidenzia intervalli privilegiati. È mia opinione che il nostro ambiente, ma anche quello di molti animali, presenti una prevalenza di suoni periodici (le voci umane ed animali) sui rumori (fruscii ecc.), e quindi che i picchi della matrice emergano in modo abbastanza pronunciato rispetto ad un fondo continuo.

È possibile simulare, per mezzo di una rete neurale, un processo di apprendimento in seguito al quale si attribuiscono a classi di stimoli delle etichette e si collocano queste classi in modo da evidenziare una loro relazione di vicinato: le reti in questione prendono il nome dal loro inventore, Kohonen, e vengono spesso definite *reti topologiche autoorganizzate*; la loro struttura generale è costituita da un primo strato di ingresso di n unità che ricevono gli stimoli, e da un secondo strato di uscita, di $m >> n$ unità, in genere disposte su di un supporto bidimensionale con topologia toroidale per evitare effetti di bordo; ogni unità d'ingresso è connessa con tutte le unità di uscita e i pesi sinaptici vengono scelti inizialmente a caso, nell'intervallo -1,+1. Vengono poi inviati gli stimoli, in successione casuale e in numero molto elevato, anche con ripetizioni, e ogni volta viene individuata sul secondo strato l'unità vincente, quella cioè che dà la risposta più elevata: l'aggiornamento dei pesi sinaptici consiste nel rinforzare i pesi dell' unità vincente e, in misura via via decrescente, quelli delle unità adiacenti, fino a diminuire quelli delle unità lontane; se gli stimoli, come spesso accade, sono raggruppabili in categorie, al termine dell'apprendimento l'organizzazione dei pesi sinaptici darà luogo, sul secondo strato, a una serie di regioni, chiamate *bolle*, che rispondono preferibilmente a stimoli appartenenti a una determinata categoria; inoltre bolle adiacenti corrispondono a categorie somiglianti di stimoli: in questo modo si configura in modo del tutto spontaneo una *topologia degli stimoli*.

Marco Beato, nella sua tesi di laurea (1995), si è posto il problema di classificare per mezzo di una rete di Kohonen gli stimoli musicali, tentando di simulare in questo modo il comportamento del sistema uditivo: il primo strato, che voleva rappresentare l'orecchio, era costituito da 60 unità risonanti con frequenze distribuite logaritmicamente su 5 ottave, mentre il secondo strato era costituito da 30x30 unità; gli stimoli erano la rappresentazione spettrale di suoni dotati di 16 armonici, con ampiezze decrescenti secondo la legge $1/n$, e frequenza fondamentale variabile sulle prime due ottave dell'orecchio. Il solo parametro variabile della rete (a parte alcuni parametri a carattere tecnico utili a controllare la velocità di apprendimento e la stabilità della rete) era il potere risolutivo delle unità del primo strato, che poteva andare da un valore elevato, in cui un suono privo di armonici eccitava solo l'unità corrispondente e le due immediatamente adiacenti, ad un valore basso, in cui l'eccitazione era distribuita su numerose unità vicine a quella centrale.

Sono state effettuate numerose simulazioni, variando gli stimoli e variando il potere risolutivo dell'orecchio. In una prima fase sono stati scelti 25 suoni, a distanza di semitono, che coprivano quindi 2 ottave, con l'ultimo suono a due ottave di distanza dal primo: a basso potere risolutivo le bolle si organizzavano secondo una scala cromatica aperta, mentre aumentando il potere risolutivo le bolle distanti un'ottava l'una dall'altra risultavano sistematicamente adiacenti, e la scala cromatica risultava chiusa: il primo risultato interessante è quindi la riproduzione del fatto che *note che distano un'ottava hanno lo stesso nome*. Aumentando ancora il potere risolutivo la continuità della scala cromatica si spezza e contemporaneamente appare un legame (adiacenza di bolle) che corrisponde all'intervallo di quinta: il circolo delle quinte non è apparso tutto intero, ma spezzato in due parti, che è già comunque un risultato notevole. Ripetendo la simulazione con 25 suoni, separati stavolta da quarti di tono, abbiamo ottenuto dapprima la scala aperta, poi la scala chiusa sull'ottava, poi la compresenza di adiacenze di quarti di tono e di quinta, ma anche un interessante risultato nuovo: i 25 suoni, anche se non completamente organizzati secondo le quinte, si disponevano su due grandi bolle, ognuna delle quali conteneva i suoni appartenenti ad una delle due scale cromatiche distanti un quarto di tomo l'una dall'altra.

Da queste simulazioni si possono ricavare numerose conclusioni: l'organizzazione topologica dello spazio dei suoni non è univoca, ma dipende dal potere risolutivo delle unità sensibili, e per certi valori abbiamo delle *transizioni di fase topologiche* passando attraverso le quali l'ordinamento dei suoni viene sconvolto e poi ricostituito secondo un diverso schema; le correlazioni responsabili di questo comportamento sono quelle di adiacenza lungo l'asse delle frequenze, quelle di ottava, quelle di quinta, e probabilmente si potrebbe continuare, se solo si potresse aumentare il numero di unità e diminuire la distanza tra i suoni senza dover aspettare tempi proibitivi per la conclusione dell'apprendimento; infine che la scala cromatica di 12 semitoni, anche non ordinati, rappresenta un certo livello di organizzazione dello spazio sonoro.

Se volessimo rappresentare in modo grafico le successive organizzazioni topologiche dello spazio sonoro potremmo cominciare con la topologia naturale dei reali, rappresentata dall'asse delle frequenze continue; successivamente, per mettere in evidenza l'invarianza Doppler e la struttura tonotopica della coclea e della corteccia uditive potremmo cambiare la metrica, passando dall'asse delle frequenze a quello dei logaritmi delle frequenze; poi, per evidenziare il ruolo dell'ottava, scegliere 2 come base dei logaritmi, e *arrotolare* l'asse sulla superficie di un cilindro, in modo che ogni giro dell'elica cilindrica rappresenti un'ottava; possiamo poi *schiacciare* l'elica in una circonferenza e discretizzare i suoi punti, selezionandone 12; possiamo poi cambiare l'ordine di questi punti ottenendo il circolo delle quinte (possibili varianti sono il circolo chiuso del sistema temperato e quello aperto del sistema armonico); possiamo tentare di recuperare il *glissando*, avvolgendo i 12 punti sulla superficie di un toro e collegandoli in modo continuo, oppure in modo discretizzato, a salti di quinta; infine possiamo recuperare l'ottava, *aprendo* il toro variando con continuità uno dei suoi raggi e trasformandolo in una superficie aperta. Tutto questo supponendo che un suono possa essere sempre rappresentato da un punto su una linea o su una varietà a pochissime dimensioni.

Abbiamo visto quale importanza abbiano gli armonici nel determinare la struttura topologica dello spazio uditivo: se vi fossero soltanto suoni puri l'unica topologia ottenibile con la rete di Kohonen sarebbe quella unidimensionale della retta. Tuttavia i pesi degli armonici possono variare moltissimo, nell'ambito della modulazione timbrica, e questo spiega perché una rappresentazione completa abbia bisogno di un numero di dimensioni molto più grande, pari, per ogni frequenza del fondamentale, al numero degli armonici presenti, o meglio percepibili, nello spettro. È possibile approfittare di questa variabilità per dare un senso ai punti della superficie del cilindro che non si trovano sull'elica: secondo quanto detto i punti dell'elica che si trovano sulla stessa verticale corrispondono a suoni dello stesso nome (croma) distanti una o più ottave; i punti intermedi hanno la stessa croma ma corrispondono a percezioni possibili? La risposta è affermativa ed è possibile rendersene conto creando una successione di suoni ottenuti in questo modo: siano a_n le ampiezze degli armonici di un dato fondamentale, per esempio $a_n = 1/n$ e $a_{2n} = 1/n$ quelle del suono di ugual timbro ma un'ottava sopra; consideriamo ora il suono con ampiezze $c_n = (1 - x)a_n + xa_{2n}$, dove il coefficiente x varia tra 0 e 1 in modo continuo o discreto: il suono percepito avrà la stessa croma degli estremi, ed un'altezza che varia di un'ottava, manifestandosi come una variazione continua di timbro.

In modo analogo è possibile costruire delle scale *assurde* che percorrono una traiettoria circolare sul cilindro e dopo un giro si ritrovano al punto di partenza, oppure scale ascendenti (discendenti) che scendono (salgono) di un'ottava.

E se gli armonici non ci fossero? Tutto ciò che stato detto finora crollerebbe?

Apparentemente si, a meno che non si ricordi che il sistema uditivo è molto più complesso del modello che ne abbiamo fornito; in particolare, come tutti i sistemi reali, ha certamente qualche aspetto non lineare nel suo comportamento, facilmente localizzabile nella sua parte meccanica (timpano, catena degli ossicini e muscoli annessi, finestra ovale). La caratteristica di un sistema non lineare, che si mantiene anche nel caso elementare dell'oscillatore unidimensionale, è che la frequenza delle sue oscillazioni libere dipende dall'ampiezza, e che una forza di andamento sinusoidale produce un comportamento tanto più distorto quanto più intensa è la forza. Se si esamina con maggiore attenzione la forma delle oscillazioni libere si nota che le piccole oscillazioni sono isocrone e puramente sinusoidali, mentre quelle grandi presentano delle componenti armoniche. Questo fatto già è sufficiente per recuperare gli armonici, quelli nella risposta, se l'ampiezza dello stimolo è abbastanza grande. Se poi esaminiamo le cose più a fondo troviamo anche altri elementi a sostegno di una teoria della consonanza.

Il modo più semplice per descrivere un oscillatore (unidimensionale) non lineare è quello di riprendere l'equazione del moto e considerare la costante di richiamo elastica k non più come una costante, ma come una funzione dello spostamento $k(x)$: se la funzione è pari l'oscillatore si comporta in modo simmetrico, ma nel caso più generale la funzione può non avere parità definita, come accade nella maggior parte dei sistemi naturali. Per semplicità consideriamo tuttavia funzioni pari, oppure poniamo nell'equazione del moto la costante k moltiplicata per una funzione dispari di x, come per esempio $\tanh x, \sin x, x + \epsilon x^3, \sinh x, \tan x$, la quale si riduce a x per piccole oscillazioni; le prime due funzioni propongono un comportamento più *soffice* di quello dell'oscillatore lineare, che si manifesta nella diminuzione della frequenza in funzione dell'ampiezza, mentre le successive sono più *dure* e la frequenza aumenta con l'ampiezza. L'andamento di queste forze è mostrato nella fig.7. Le figg.8 e 9

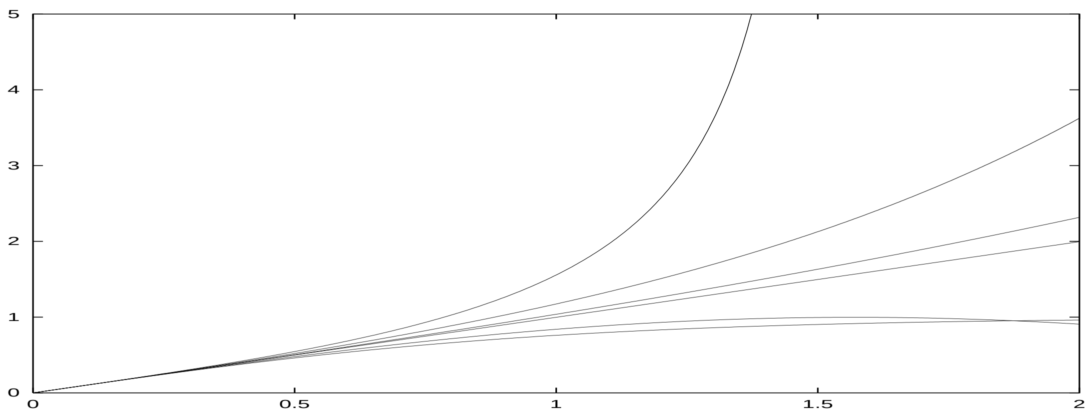

Fig. 7: Esempi di forze di richiamo non lineari. Dall'alto in basso: $f(x) = \tan x, \sinh x, x + \epsilon x^3, x, \sin x, \tanh x$.

mostrano il comportamento di un oscillatore non lineare ($f(x) = \sinh x$), ottenuto integrando numericamente l'equazione del moto.

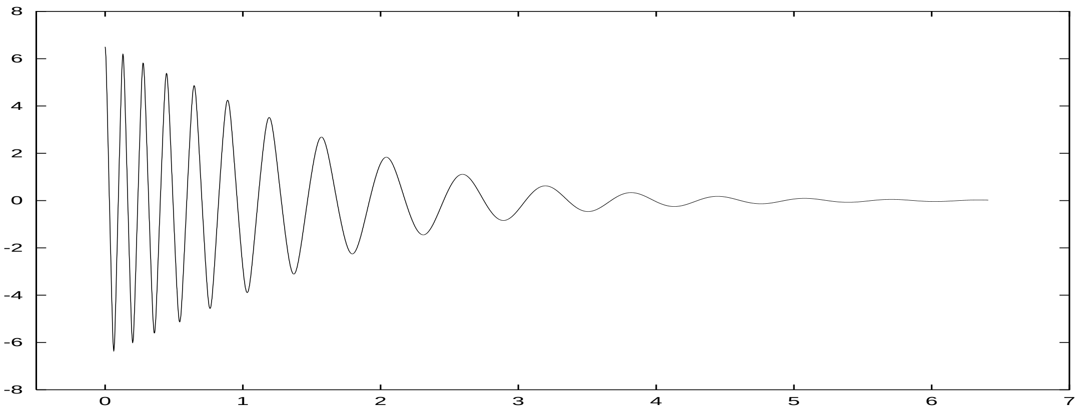

Fig. 8: Andamento temporale dell'oscillatore non lineare smorzato. $f(x) = \sinh x$; $\omega_0 = 10.$, $s = 2$. Grande ampiezza iniziale.

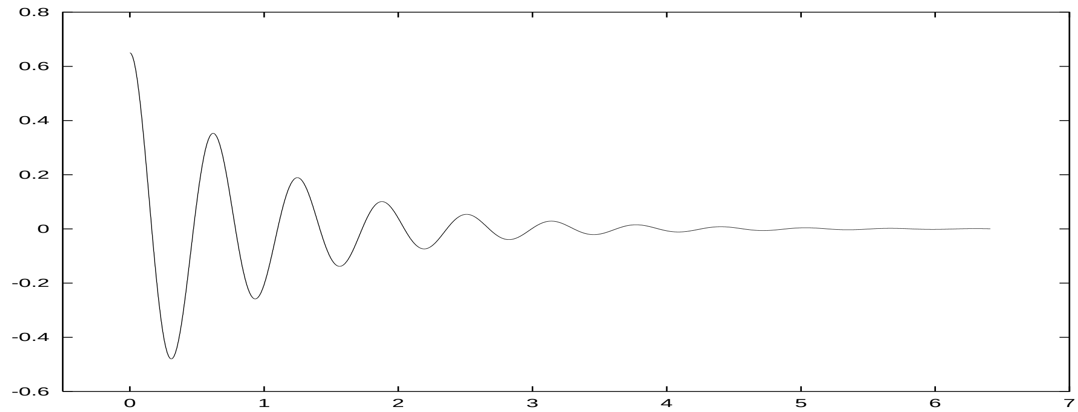

Fig. 9: Andamento temporale dell'oscillatore non lineare smorzato. $f(x) = \sinh x$; $\omega_0 = 10.$, $s = 2$. Piccola ampiezza iniziale.

Oltre alla variazione di frequenza con l'ampiezza la non linearità produce una distorsione delle oscillazioni spontanee che non sono più puramente sinusoidali, come si può vedere nelle figg.8,9: le oscillazioni di grande ampiezza assomigliano ad un'onda triangolare, dotata di numerosi armonici, mentre quelle piccole si ammorbidiscono e diventano puramente sinusoidali

L'analisi di Fourier mostra la comparsa degli armonici dispari del fondamentale, con ampiezze decrescenti con l'ordine dell'armonico, crescenti con l'ampiezza dell'oscillazione: possiamo chiamarli *armonici fantasma*.

Quando l'oscillatore non lineare viene forzato, e si trova in condizioni di smorzamento critico per le piccole oscillazioni, il suo comportamento può essere complicato, se la forza non è puramente sinusoidale. Nel caso sinusoidale puro è abbastanza facile dimostrare che il moto ha lo stesso periodo della forza $f(t) = F\sin(\omega t)$ (figg. 10 e 12) ma al suo fondamentale si aggiungono i soliti fantasmi degli armonici dispari, come si vede nelle figg.11 e 13.

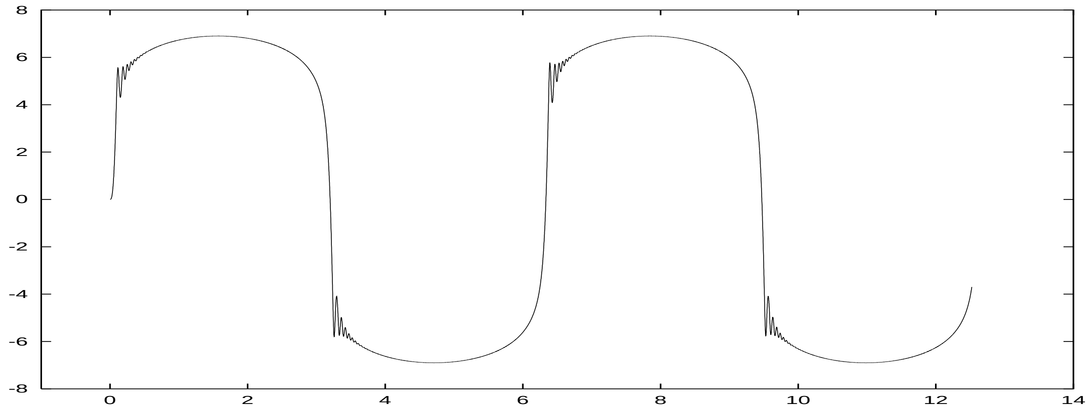

Fig. 10: Oscillatore non lineare forzato: $\omega_0 = 10.\omega = 1.F = 50000, s = 20.$

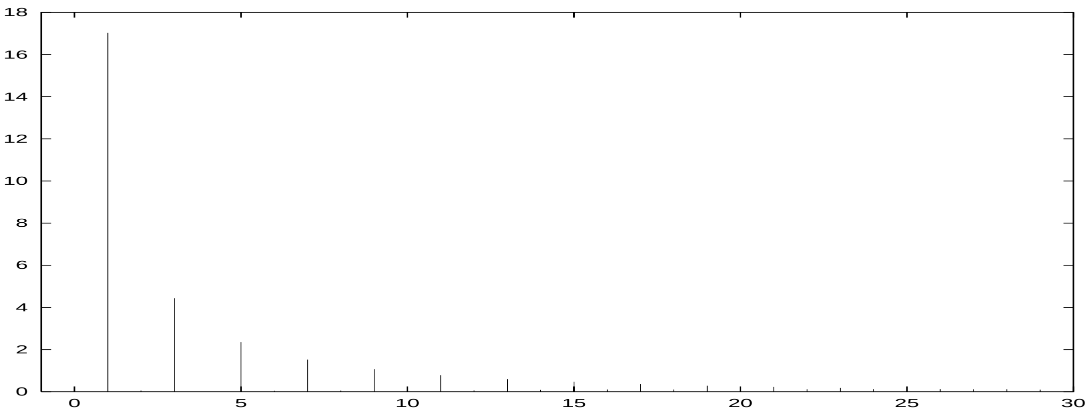

Fig. 11: Evidenza degli armonici dispari nello spettro di Fourier del caso precedente.

Se invece la forza ha un andamento complesso è sempre possibile rappresentarlo come una sovrapposizione di moti puramente sinusoidali, ma per il sistema non lineare non vale il principio di sovrapposizionee quindi non è più vero che il moto complessivo risulta dalla sovrapposizione dei moti semplici, anche se distorti. Per rendercene conto prendiamo in esame una forza che sia la sovrapposizione di due forze di uguale ampiezza ma di periodi diversi, in rapporto razionale tra di loro: la forza è ancora periodica con un periodo pari al minimo comune multiplo dei due periodi, e anche il moto dell'oscillatore, che però presenta, con varie ampiezze, tutti gli armonici della frequenza fondamentale, che è l'inverso del periodo globale.

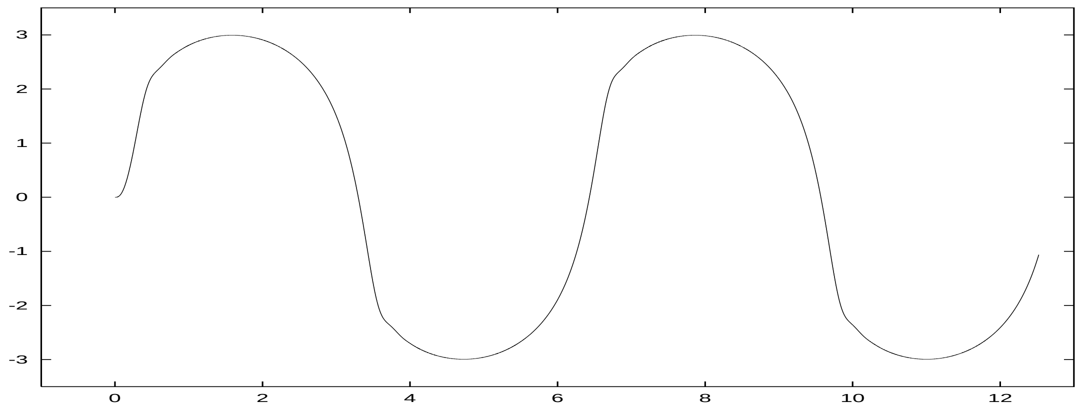

Fig. 12: Oscillatore non lineare forzato: $\omega_0 = 10.\omega = 1.F = 1000, s = 20.$

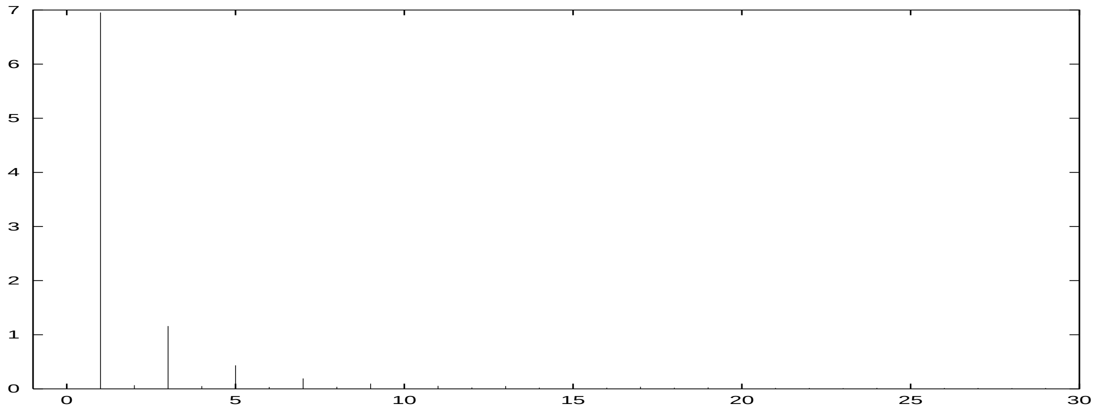

Fig. 13: Oscillatore forzato: spettro di Fourier del caso precedente.

Le figure seguenti mostrano varie situazioni, in cui le due frequenze forzanti sono via via più *dissonanti* l'una rispetto all'altra: si va da un rapporto 3/2 (quinta giusta) fig.14, a 5/3 (sesta maggiore)fig.15, a 9/8 (tono grande) fig.16, e si vedono chiaramente le distribuzioni spettrali che partono tutte dal cosiddetto *fondamentale mancante* con i suoi armonici che diventano più importanti con l'aumentare della dissonanza; la fig.17 mostra invece la traiettoria nello spazio delle fasi per il caso di fig.16.

Si possono fare alcune considerazioni: se nel caso 5/3 pensiamo alle due note come a un Sol2 e un Mi3 il fondamentale mancante è un Do1, mentre molti sostengono che il suono percepito è il suono-differenza, che è invece un Do2; l'equivoco può nascere se si pensa che l'esperimento, fatto per esempio con un violino (il terzo suono di Tartini), mette in giuoco coppie di suoni dotati a loro volta di armonici, per cui la situazione è molto più complicata. La seconda considerazione prende lo spunto dal fatto che all'aumentare della dissonanza aumenta la densità degli armonici in un determinato intervallo, suggerendo in questo modo che la dissonanza sia il risultato di un maggior impegno del sistema uditivo: il caso limite è quello del rapporto irrazionale tra le due frequenze forzanti che dà luogo addirittura ad una distribuzione spettrale continua la quale, anche se discretizzata dal sistema uditivo, gli richiede il massimo impegno.

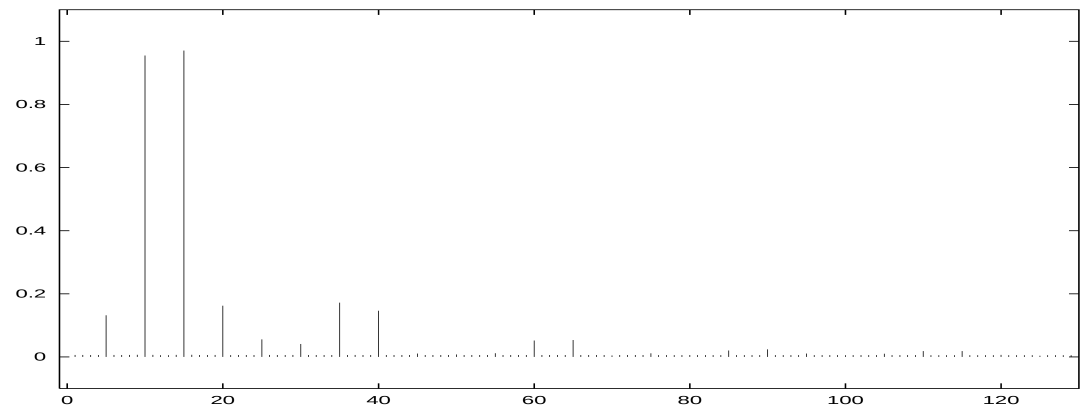

Fig. 14: Oscillatore non lineare forzato da due frequenze: $\omega_1 = 10, \omega_2 = 15, F = 50000$. Il fondamentale mancante ha frequenza 5.

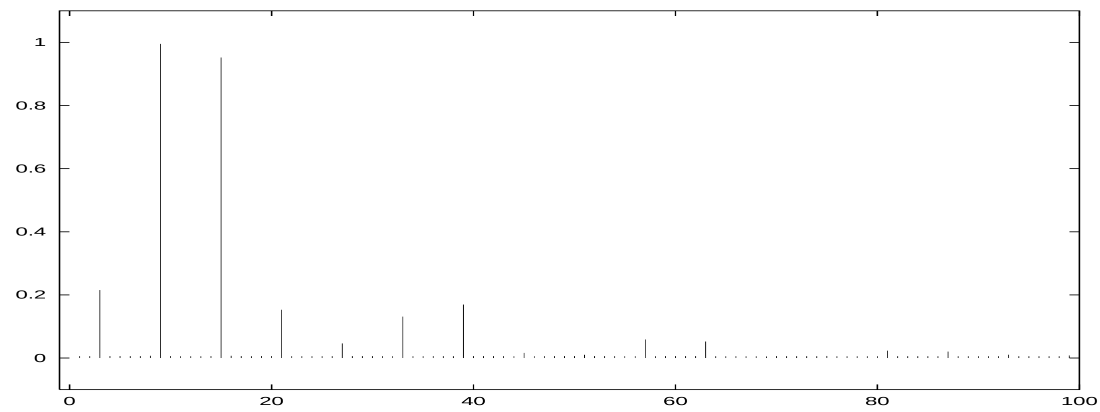

Fig. 15: Oscillatore non lineare forzato da due frequenze: $\omega_1 = 9, \omega_2 = 15, \quad F = 50000$. Il fondamentale mancante ha frequenza 3.

Credo di poter concludere questo capitolo sottolineando ancora una volta l'importanza degli armonici in qualunque criterio di consonanza, anche quello numerologico, aggiungendo però che le esperienze uditive che hanno contribuito allo sviluppo sinaptico e la plasticità del cervello hanno utilizzato gli armonici per giungere al risultato che tutti conosciamo; un'ipotetica specie che si fosse sviluppata in un ambiente

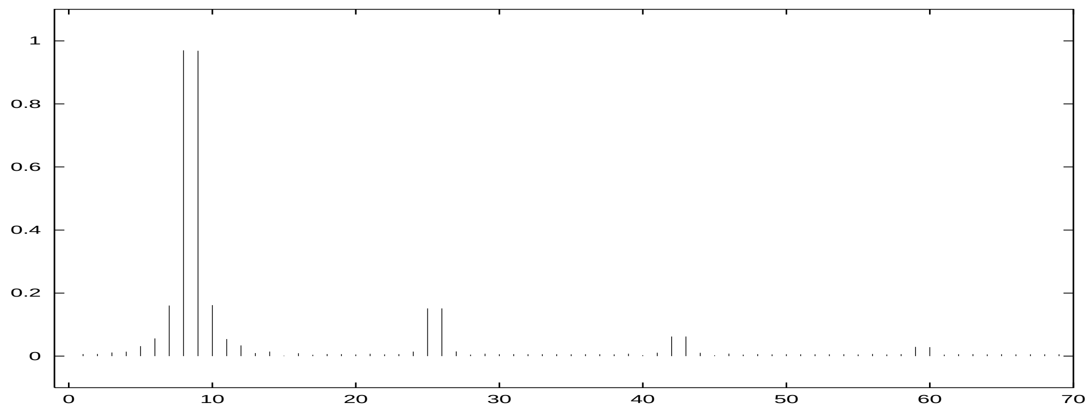

Fig. 16: Oscillatore non lineare forzato da due frequenze: $\omega_1 = 8, \omega_2 = 9, \quad F = 50000$. Il fondamentale mancante ha frequenza 1.

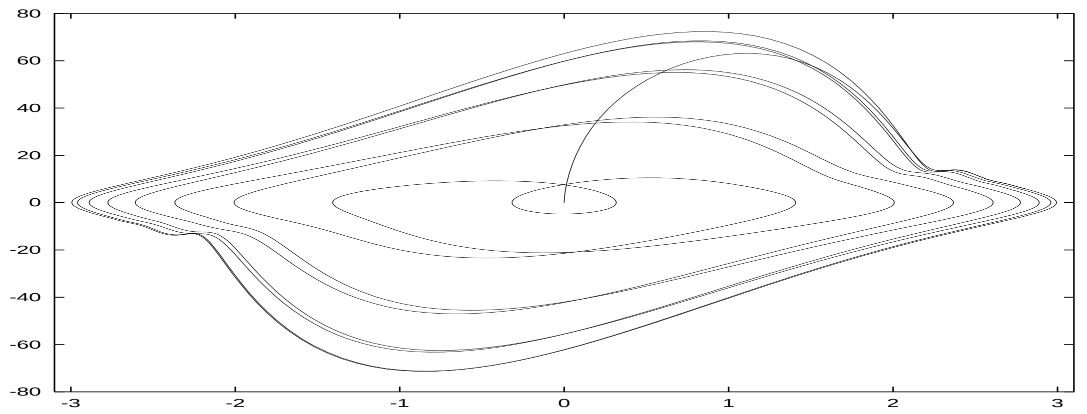

Fig. 17: Traiettoria nello spazio delle fasi x, v dell'oscillatore di fig.15.

dominato da suoni aperiodici, prodotti per esempio da membrane, avrebbe probabilmente sviluppato un'estetica musicale completamente diversa. E visto che ho scritto la parola *estetica* voglio ricordare che anch'essa si evolve nel corso del tempo, insieme al concetto di consonanza, e che molti grandi compositori si sono collocati al margine dell'estetica ufficiale del loro tempo, venendone in un primo momento rifiutati, per essere magari accettati entusiasticamente solo dopo la loro scomparsa.

Un'ultima considerazione di carattere fisiologico: nell'esplorare lo spazio uditivo mi sono mosso soprattutto nel dominio della frequenza, come se il tempo non fosse rilevante. Certamente lo è, ma le nostre conoscenze su quel che accade nell'orecchio interno durante l'ascolto di suoni complessi sono ancora troppo limitate per basare su di esse una teoria della consonanza: è lecito pensare che il meccanismo di trasduzione analogico-digitale, che trasforma l'onda di pressione del fluido cocleare nel segnale elettrico lungo le fibre del nervo acustico, abbia un comportamento non lineare molto più complesso del modellino elementare presentato poco sopra, e che fenomeni di interferenza e di inibizione concorrano a determinare il livello di gradevolezza di un insieme di suoni, come o più dei rapporti tra le corrispondenti frequenze.