1

Introduzione all’Audio Digitale

La catena Naturale

Il **suono** consiste in vibrazioni (tra **20** e **20000** cicli al secondo) si propaga attraverso le vibrazioni di un mezzo (tipicamente l’aria) e viene percepito attraverso le **orecchie**. Al di fuori dei quella gamma di frequenze gli orecchie umane non sono sensibili e non si parla di **suono** ma semplicemente di **vibrazioni**. Molti animali (come i cani e i pipistrelli) sono sensibili a frequenze molto superiori (gli **ultrasuoni**). A partire dai primi anni del ’900 si è imparato a trasmettere e memorizzare i suoni con sempre migliori caratteristiche di fedeltà.

La catena Analogica

Vediamo una costosa catena **analogica** del periodo di massimo splendore di questo tipo di tecnologia (gli anni ’70).

La musica viene venduta sui dischi in vinile, da suonare su un giradischi, amplificata da un preamplificatore e da un amplificatore e poi inviata alle casse acustiche.

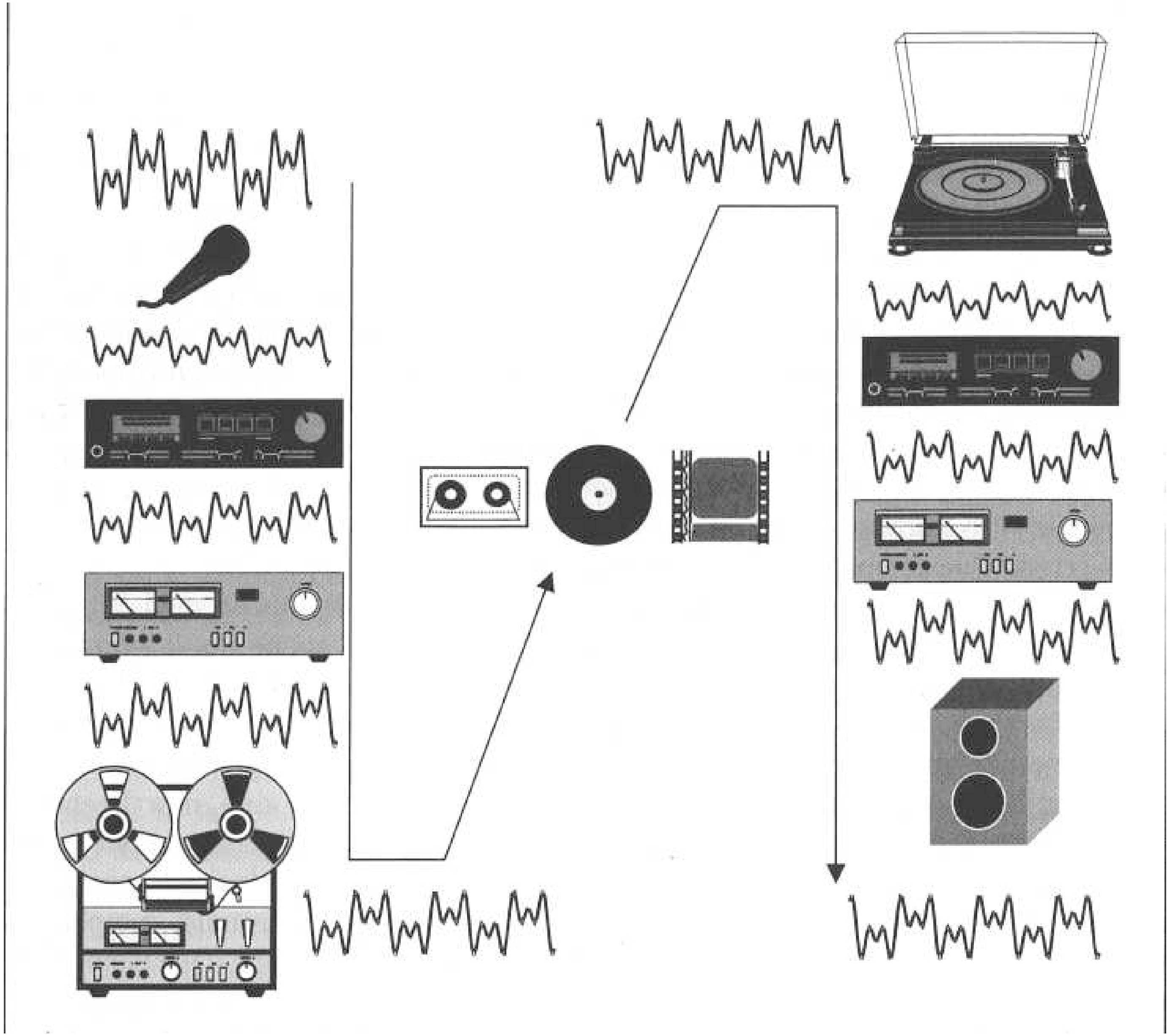


Figura tratta da Lombardo, Valle, "Audio e Multimedia, II ed." Apogeo, 2005

Segnali finiti ed infiniti, spettro di frequenza

La prima idealizzazione che serve per poter trattare matematicamente i segnali audio è quella di schematizzarli con una funzione **s(t)** che associa ad ogni istante di tempo **t** un valore finito reale (positivo, negativo o nullo); tipicamente questo valore è la tensione misurata in un punto del circuito elettrico ma si può usare anche la corrente oppure un'altra quantità significativa.

I segnali di durata infinita sono ovviamente una idealizzazione ma evitano la complicazione di dover definire l'inizio e la fine del segnale. Da un punto di vista pratico, per noi che ci interessiamo alle frequenze audio, ogni segnale più lungo di un minuto si può considerare infinito.

Il primo esempio di segnale infinito che viene in mente è l'alimentazione a **50Hz** della rete elettrica. Dato un segnale si può definire la sua **energia** (nel caso della rete la bolletta che si paga per avere quel segnale), e la sua **potenza** (quanta luce fa una lampadina attaccata a quella spina). Nel caso del segnale a **50Hz** (a parte i disturbi) tutta la potenza è concentrata su di una sola frequenza; nel caso di un segnale musicale invece, la potenza è suddivisa tra varie frequenze: una **nota** di uno strumento ha una frequenza fondamentale (che dà il nome alla nota stessa) e le **armoniche** (che determinano il timbro dello strumento). Se il segnale è più complicato (un pezzo d'opera con accompagnamento e rumori ambientali) pressoché tutte le frequenze sono presenti, magari in minima parte, ma dal punto di vista della riproduzione audio tradizionalmente viene considerato solo l'intervallo **20Hz-20Khz** che rappresenta le frequenze udibili da esseri umani giovani e sani.

Dominio del tempo e dominio della frequenza

Dato un segnale **infinito** il suo contenuto in frequenza (il suo **spettro**) su può ottenere con una operazione matematica detta **Integrale di Fourier**.

Il risultato dell'Integrale di Fourier di **s(t)** è una funzione a valori complessi **S(f)** definita tra meno infinito e più infinito. I valori complessi portano con sé l'informazione relativa alla fase del segnale. Per esempio se si considera il segnale emesso da un sistema di altoparlanti multivia, spostando la posizione del tweeter cambia la fase del segnale e non il suo contenuto in frequenza, in altri termini il valore assoluto dell'Integrale di Fourier resta lo stesso.

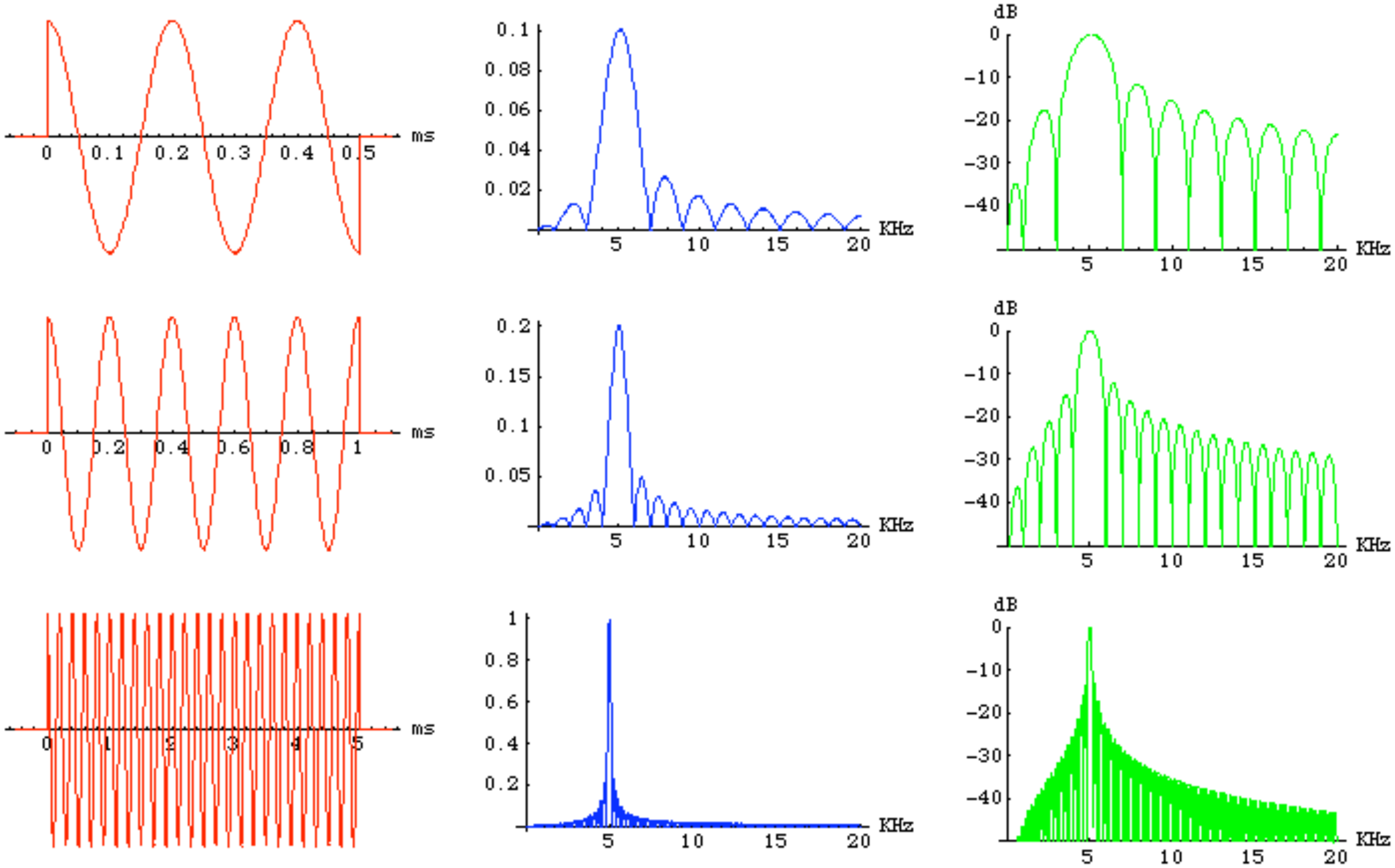
L'integrale di Fourier è reversibile, nel senso che a partire dal contenuto in frequenza si può costruire **esattamente** il segnale originale. Ricordo che stiamo parlando di operazioni **matematiche** svolte su **segnali infiniti** lavorando con quantità **senza errori**, in altra parola **segni di lapis su di un pezzo di carta** ed **esattamente** va inteso in questo senso.

Di solito quando si parla di spettro di un segnale si trascura la fase e si considera **|S(t)|** (oppure se ci interessa la potenza espressa in decibel la quantità **20 Log_{10} |S(t)|**). In entrambi i casi si può dimostrare che la parte positiva e la parte negativa delle funzioni risultanti sono specularmente uguali e quindi di solito se ne disegna solo la parte positiva. Con questa semplificazione però si perde la reversibilità e quello che resta è solo una misura effettuata sul segnale, priva dell'informazione completa.

Vediamo adesso che cosa succede se il segnale non è infinito ma di durata **T**. Questa situazione si può rappresentare considerando il segnale ancora infinito ma moltiplicato per una finestra rettangolare, ovvero una funzione anch'essa infinita che vale sempre **0** tranne in una porzione di durata **T** in cui vale **1**. Con questo trucco un segnale finito viene visto come il prodotto di due segnali infiniti. In pratica se la finestra è molto ampia tutto resta più o meno come prima, se la finestra è stretta, invece, la risposta cambia decisamente.

Tanto per avere un'idea di questo fenomeno vediamo un segnale a **5KHz** di durata rispettivamente di **0.5ms**, **1ms** e **5ms**, ovvero una sinusoide infinita vista attraverso un finestra di durata variabile.

La trasformata di Fourier dei segnali in rosso (effettuata per via matematica) dà il vero spettro di frequenza, poiché un segnale complesso di durata infinita non si può disegnare, prendiamo il valore assoluto della trasformata (grafici in blu) ottenendo un segnale reale e perdendo le informazioni sulla fase. Se invece del valore assoluto plottiamo il suo logaritmo in base **10** moltiplicato per **20** si ottiene la potenza del segnale alle varie frequenze espressa in **dB**. È questa la rappresentazione che si trova nelle riviste (di solito con anche la scala delle ascisse in forma logaritmica). Si nota che mano a mano che l'impulso diventa più lungo il suo contenuto in frequenza si concentra intorno ai **5KHz** e l'influenza del taglio diviene meno avvertibile.



I grafici rossi rappresentano un segnale nel dominio del tempo (un burst a 5Khz di durata variabile) I grafici blu il valore del valore assoluto dell'integrale di Fourier, i grafici verdi la potenza rispetto al valore massimo

Sistemi lineari, risposta in frequenza, filtri

Un fenomeno che accade in continuazione ai segnali audio (musicali e non) è l'avventura di essere modificati dai dispositivi che incontrano nel loro passaggio. Se si ascolta una conversazione da dietro una porta chiusa il segnale che passa, oltre che essere attenuato, è anche deformato al punto che a volte non è più intelligibile. Una operazione del genere in teoria dei segnali viene chiamata **filtraggio** e il dispositivo che la compie **filtro.** Due importanti proprietà che devono avere i filtri per poter essere studiati facilmente sono la **linearità** e l'**invarianza temporale**.

Un filtro è lineare se non introduce distorsione qualunque sia il livello del segnale di ingresso, matematicamente, detta f(t) la trasformazione indotta dal filtro si ha

$$f(a\ g(t) + b\ h(t)) = a\ f(g(t)) + b\ f(h(t))$$

Un amplificatore non sarà mai perfettamente lineare: se gli date abbastanza segnale va in *clipping* e se gli date un milione di volt brucia. Anche un pezzo di filo non è lineare (con abbastanza corrente fonde).

Nota bene: linearità non vuole dire **risposta piatta**, un controllo di tono ideale è **lineare** anche se altera volutamente la risposta in frequenza mentre un compressore di dinamica o un codificatore **MP3** sono **non lineari** già in sede di progetto.

L'**invarianza temporale** significa che il comportamento del sistema è lo stesso un minuto, un microsecondo, un secolo prima o dopo un istante dato. Anche in questo caso è ovvio che un dispositivo reale non può essere invariante nel tempo (visto che ha una data di costruzione e prima o poi si rompe o viene spento e/o buttato via) ma anche durante il suo normale funzionamento le variazioni di temperatura, di tensione di alimentazione o le interferenze possono alterare il suo comportamento.

La comodità di supporre vere la linearità e l'invarianza temporale sta nel fatto che il comportamento di un filtro lineare e invariante del tempo è completamente determinato da una funzione complessa detta funzione di trasferimento. Questa non è altro che l'**Integrale di Fourier** del segnale di uscita quando all'ingresso viene presentato un singolo impulso ideale (un oggetto matematico troppo complicato per essere definito in meno di 5 pagine). Anche questa è un operazione matematica (segni di lapis su un pezzo di carta). In pratica, visto che l'impulso ideale non ha realtà fisica, si preferisce sollecitare il sistema con altri tipi di segnali realizzabili (rumore rosa, segnali sinusoidali, onde quadre, gradini etc.) ottenendo i soliti grafici di risposta in frequenza e/o in fase (si vedano per esempio le prove classiche e non dei sistemi di altoparlanti).

Occorre sottolineare che il comportamento del filtro non dipende assolutamente dalla sua natura fisica, uno stesso effetto (per esempio una variazione di fase) può essere ottenuto spostando in avanti o indietro il tweeter oppure con una rete di capacità, resistenze ed induttanze. Questo accorgimento viene usato nei concerti all'aperto dove gli altoparlanti delle varie bande di frequenza vengono messi dove si può e poi la fase viene aggiustata per via elettronica (è più semplice ruotare un potenziometro che noleggiare e posizionare una gru o una torre di tubi per mettere le casse in posizioni particolari).

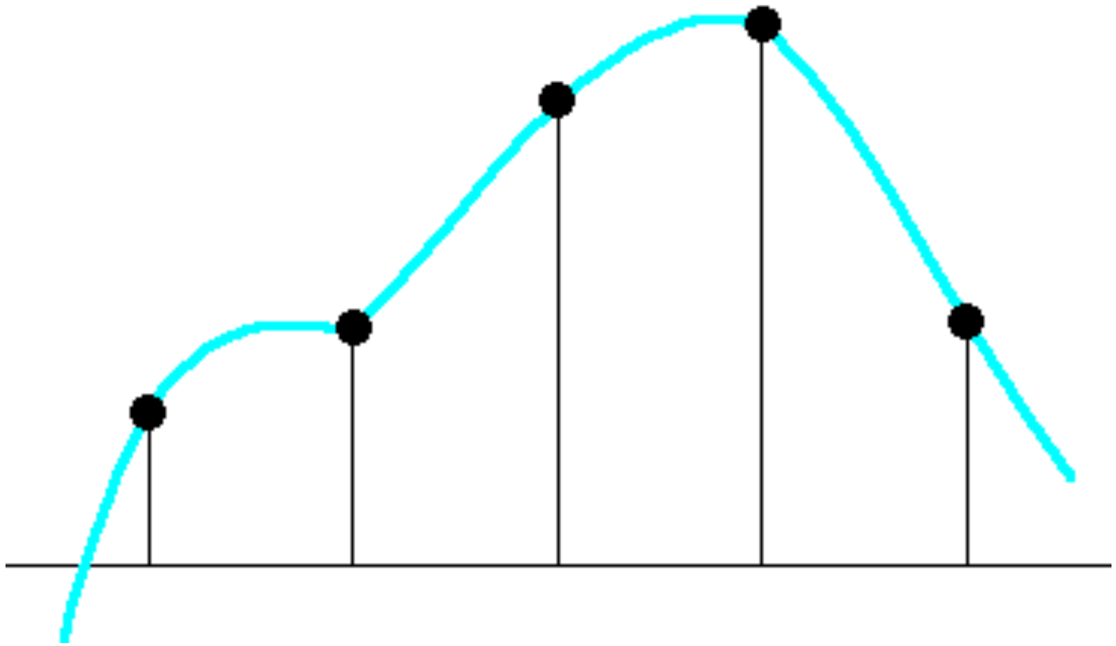
Introduzione all'Audio analogico

Prima di parlare di segnali digitali diciamo due parole sui segnali analogici. **In natura non esistono segnali veramente analogici**: le varie forme di trasmissione comunemente chiamate "analogiche" in realtà si basano su una discretizzazione più o meno fine realizzata dalle varie soluzioni fisiche adottate. La risoluzione di un **LP** è limitata dalle dimensioni delle asperità si cui striscia la puntina e, siccome queste vengono erose durante l'ascolto, un **LP** non suona mai due volte nello stesso modo. Il nastro magnetico possiede dei domini di magnetizzazione di dimensione finita e la loro struttura determina la risposta in frequenza del nastro e la quantità di rumore. In un filo elettrico il segnale è portato da un flusso discreto di elettroni e, infine, il suono nell'aria è veicolato dal moto di molecole di gas. La caratteristica comune di queste discretizzazioni è che tipicamente sono completamente scorrelate dal segnale dando origine ad un rumore di fondo che in alcuni casi (il filo, l'aria) è quasi completamente inavvertibile, in altri casi (gli **LP** e i **nastri analogici**) è avvertibile ma (a parere di molti) meno fastidioso delle alterazioni introdotte dal digitale.

Anche in campo analogico può essere presente rumore correlato (per esempio la distorsione dell'amplificatore o le interferenze tra i cavi di alimentazione e quelli di segnale).

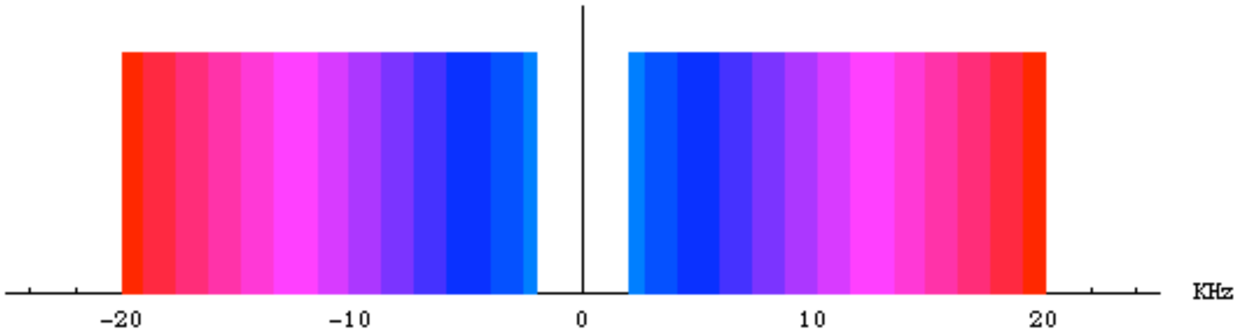
Campionamento

A questo punto siamo finalmente pronti per parlare di **segnali digitali**. Un segnale discreto si ottiene da un segnale continuo con una operazione di **campionamento temporale (Figura 2)**. Si tratta di valutare il segnale in istanti temporali spaziati di un intervallo costante **Δ** trasformando la funzione continua **s(t)** in una sequenza infinita di numeri reali s_i. La quantità $f_c=1/\Delta$ viene detta frequenza di campionamento e misurata in **Hz**.



Una funzione continua (la curva celeste) e i suoi valori campionati, (i punti neri).

Che cosa accade al segnale durante questa operazione? Vediamo nel dominio della frequenza lo spettro di un ipotetico segnale di prova composto da tutte le frequenze tra **2KHz** e **20KHz**, tutte allo stesso livello. Per poterle distinguere nelle successive manipolazioni le frequenze basse sono in blu e quelle alte in rosso. Per capire cosa succede con il campionamento è fondamentale raffigurare il segnale completo (ovvero entrambe le porzioni positiva e negativa dell'asse delle frequenze).

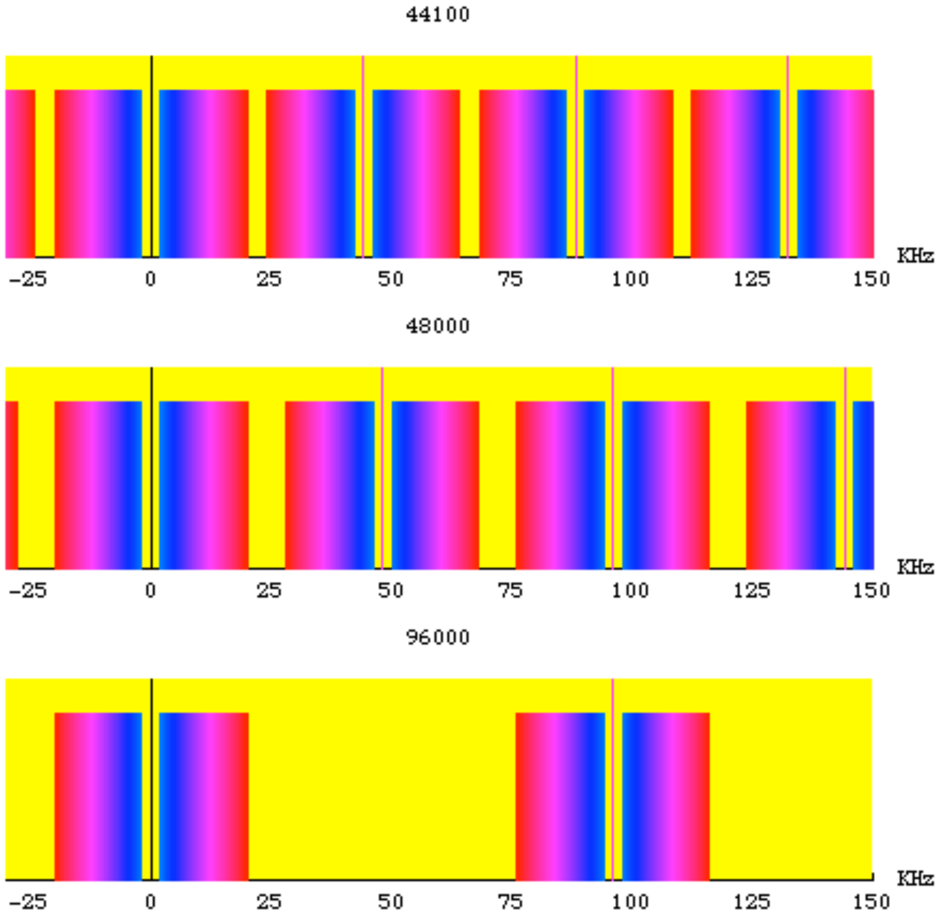


Segnale 2KHz, 20KHz nel dominio della frequenza

Nella pagina seguente si vede lo spettro del nostro segnale analogico una volta campionato con **f_c=44.1KHz**, **f_c=48KHz** o **f_c=96KHz**.

Si può dimostrare che l'effetto del campionamento nel dominio della frequenza è la ripetizione infinita dello spettro di frequenza del segnale originale spostato di un intervallo f_c.

NotaBene: gli spettri di frequenza di segnali campionati sono su sfondo giallo per ricordare che questi spettri in realtà si estendono periodicamente all'infinito). Si nota come entrambe le porzioni dello spettro siano interessate alla ripetizione periodica, ma anche la frequenza di campionamento più bassa garantisce una assenza di sovrapposizioni per il segnale di prova.

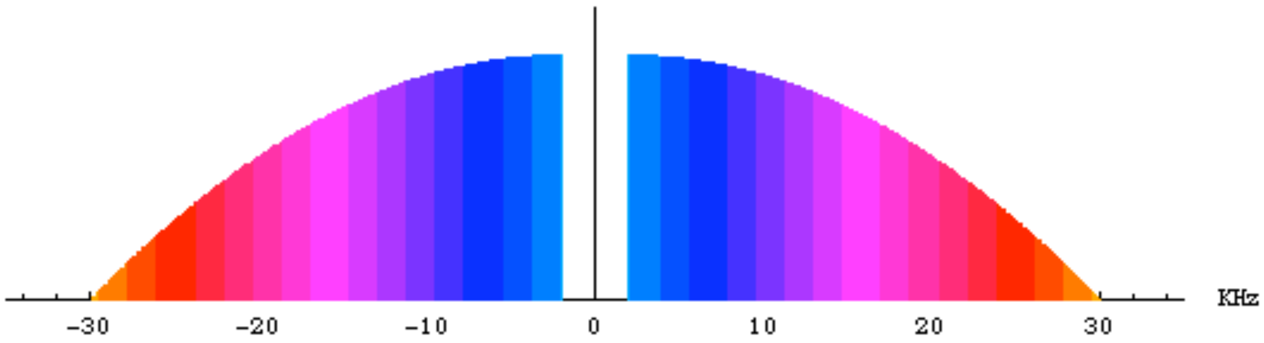


Porzione di spettro di un segnale 2KHz-20KHz campionato con f_C=44.1KHz, f_C=48KHz e f_C=96KHz, lo spettro completo è periodico e si estende all'infinito.

Aliasing

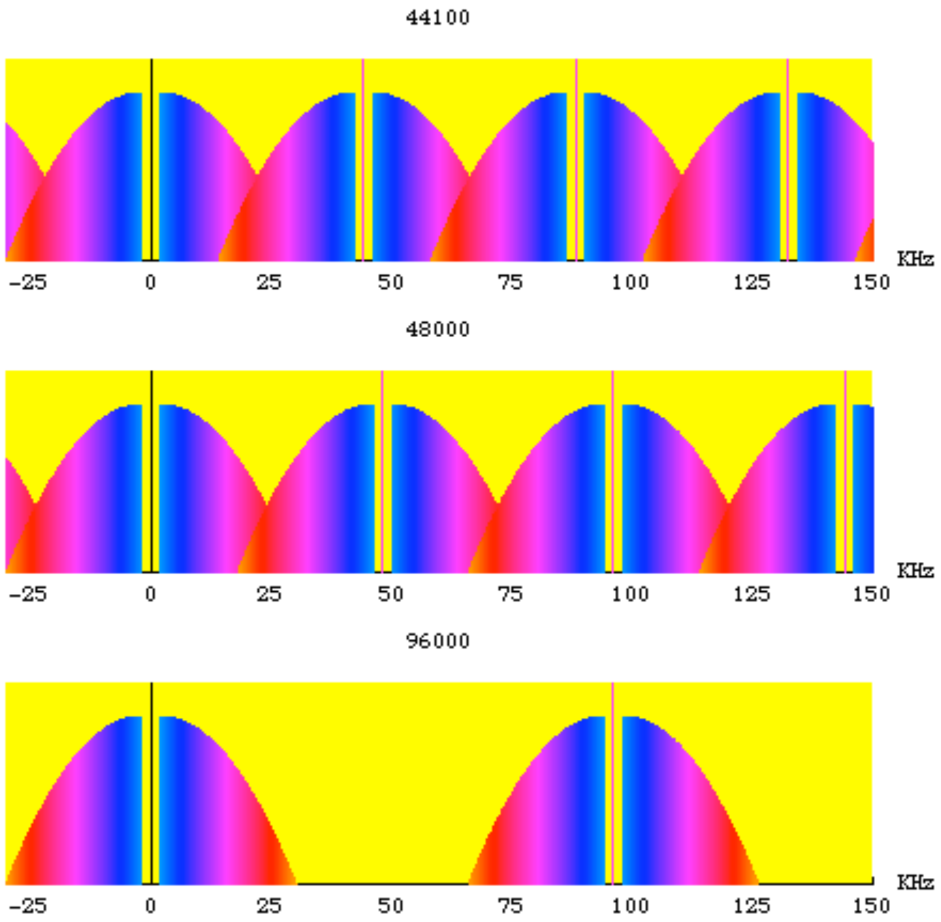
Se la frequenza di campionamento è f_c Il valore $f_c/2$ è detto **frequenza di Nyquist** e rappresenta il limite teorico utile per rappresentare segnali analogici in forma numerica. Infatti se il segnale continuo ha componenti superiori alla frequenza di Nyquist si può "perdere roba" durante la fase di campionamento.

Vediamo lo spettro di un ipotetico segnale con frequenze comprese tra **2KHz** e **30KHz**



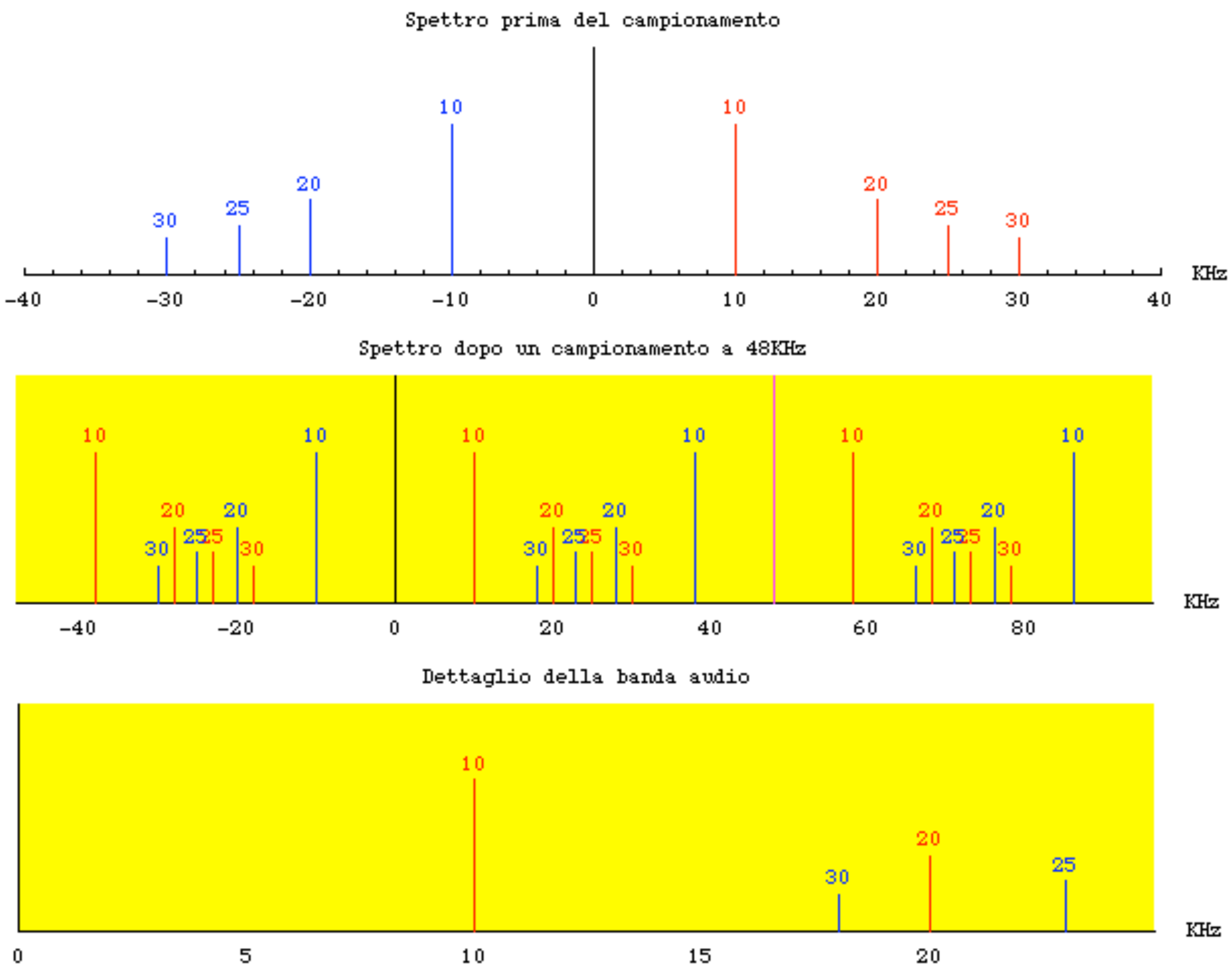
Segnale 2KHz-30KHz

Se campioniamo questo segnale con **f_c=44.1KHz**, **f_c=48KHz** e **f_c=96KHz** nei primi due casi si verifica il fenomeno dell'**aliasing** mentre nel terzo va ancora tutto bene



Porzione di spettro di un segnale 2KHz, 30KHz campionato con **f_c=44.1KHz**, **f_c=48KHz** e **f_c=96KHz**. Si nota la sovrapposizione che si verifica nei primi due casi.

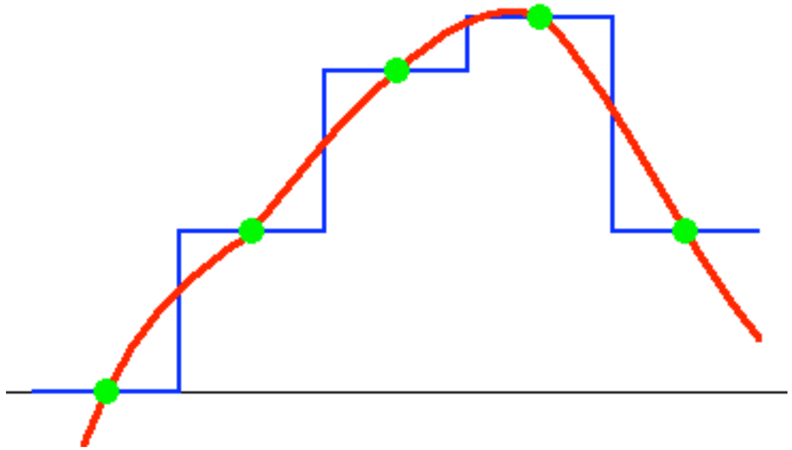
Vale la pena di vedere in dettaglio cosa significa “perdere roba”. Se la frequenza di campionamento è **48 KHz** la frequenza di Nyquist è **24KHz.** A causa delle sovrapposizioni tutti i segnali superiori a **24KHz** vengono tagliati e ribattuti a bassa frequenza comparendo in banda audio come segnali “fantasma”, per esempio un segnale a **25KHz** si comporta esattamente come se fosse a **23KHz**, e i **30 KHz** divengono **18KHz**.



Quattro toni tra **10KHz** e **30KHz** campionati con **f_c=48KHz**. Si noti l’aliasing per i due toni a frequenza più alta.

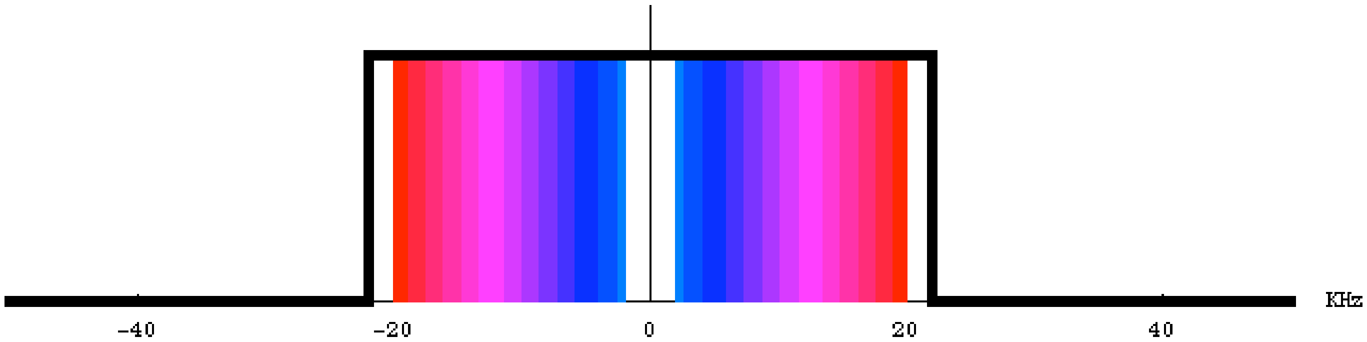
Interpolazione

Una volta ottenuto un segnale campionato si può tornare ad un segnale continuo attraverso una procedura di **interpolazione** che ricalcola il valore della funzione in tutti i punti dell'asse dei tempi a partire da quelli noti. Vediamo due possibili interpolatori, il **Sample and Hold** che mantiene un valore costante nell'intorno del campione e una **interpolazione cubica**, ovvero l'approssimazione del segnale con pezzi di polinomi di terzo grado.



Valori campionati (in verde), interpolazione Sample and Hold (in blu), interpolazione cubica (in rosso)

Da un punto di vista matematico un interpolatore è un filtro che trasforma una risposta in frequenza periodica (tipica dei segnali digitali) in una non periodica tipica dei segnali analogici. Tra gli infiniti interpolatori possibili ne esiste un ideale che uccide tutte le componenti introdotte dalla digitalizzazione restituendo il segnale originale.



Ricostruzione del segnale originale attraverso un interpolatore ideale, la linea nera rappresenta la risposta in frequenza dell'interpolatore ideale che restituisce il segnale originale.

Il **Teorema del Campionamento** assicura che questo è possibile per tutti i segnali che non hanno componenti di frequenza al di fuori della banda **0-f_c/2**. La scelta **f_c=44.1KHz** per lo standard **CD**, fissando la frequenza di Nyquist a **22050Hz** doveva, almeno nelle intenzioni assicurare, una buona riproducibilità della banda audio. In realta questa scelta troppo bassa è stata la causa principale della difficoltà di ottenere da un **CD** un audio di qualità. Gia con la frequenza di campionamento di **48Khz**, adottata successivamente per lo standard **DAT** la qualità media migliora sensibilmente.

Esiste un aneddoto in proposito, la moglie del presidente della Sony impose che su un **CD** ci doveva stare per intero la Nona di Beethoven e così la durata fu fissata in **74** minuti e la **f_c** abbassata di conseguenza.

In realtà la scelta di 44100 è legata al fatto che i primi registratori digitali usavano nastri video in standard PAL o NTSC e 44100 faceva tornare bene l'impaccamento dei dati. (i dettagli in una prossima lezione).

Conclusioni

Il Teorema del Campionamento è un risultato fondamentale di teoria dell'Informazione, insieme al teorema di Shannon (che interessa solo marginalmente l'Hi-FI) è la base di tutte le applicazioni dell'informatica alle telecomunicazioni. Naturalmente come tutti i risultati teorici abbastanza importanti da essere noti al di fuori della ristretta cerchia degli addetti ai lavori viene usato a sproposito e il suo significato viene spesso frainteso.

Il Teorema del Campionamento garantisce solo che un segnale **infinito**, di larghezza di banda **limitata**, campionato per tutta la sua durata **infinita** (producendo quindi una quantità **infinita** di numeri reali a precisione **infinita**) può essere ricostruito **esattamente** utilizzando un combinazione lineare dei suoi **infiniti** campioni con funzioni di lunghezza **infinita**.

È la sagra dell'**infinito** ed è chiaro che stiamo ancora parlando di segni di lapis su un pezzo di carta. Credere che il **Teorema del Campionamento** garantisca che tutti i **CD**-player suonino nello stesso modo perché ricostruiscono perfettamente il segnale di partenza è ridicolo ed ingenuo.

Nelle lezioni seguenti vedremo quali approssimazioni valgono nel caso del trattamento pratico dei segnali audio digitali, quali sono le tecniche usate per questo trattamento e le basi matematiche su cui si fondano.

Bibliografia

La bibliografia si intende come spunto di approfondimento, non è necessario consultarla tutta per superare l'esame

Audio Review, Rivista mensile, Technimedia

V. Lombardo, A. Valle. Audio e Multimedia, III ed. Apogeo 2008.

J. Maes, V. Mercammen, **Manuale di Tecnologia Audio Digitale**, Hoepli. *Il testo è scritto dai tecnici della Sony Europe e presenta informazioni dettagliate molto difficili da trovare in italiano*.

K. Pohlmann. **Principles of Digital Audio IV ed.** McGraw-Hill. *L'autore è uno dei massimi esperti di tecniche digitali*.

J. Watkinson, **The Art of Digital Audio III ed.**, Focal Press. *Altro testo di riferemeto in lingua inglese*.

U. Zölzer, **Digital Audio Signal Processing,** Wiley. *L'autore è un docente universitario di teoria dei segnali e il testo presenta argomenti dei suoi corsi*.