

Audio Processing

Stefano Borzì

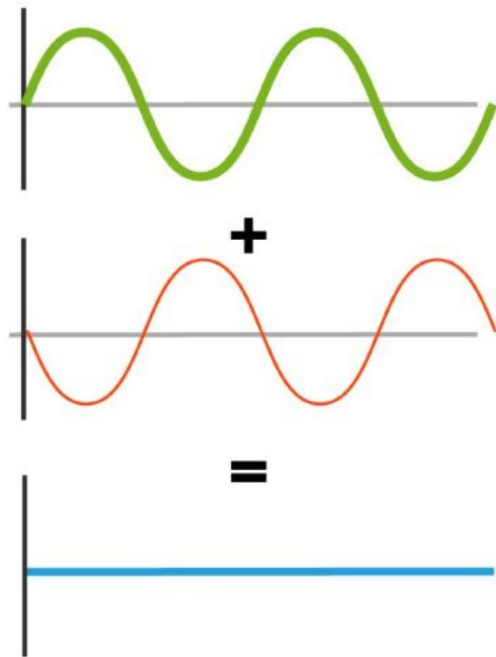
A dark blue diagonal gradient bar that starts from the bottom left corner and extends towards the top right corner, covering the lower half of the slide.

Invertire la fase

- 1.8:2 – Nozione di fase

In un editor audio generare due onde sinusoidali identiche.

- Invertire la fase di una delle due.
- Mixare le due tracce verificando di aver ottenuto il silenzio.



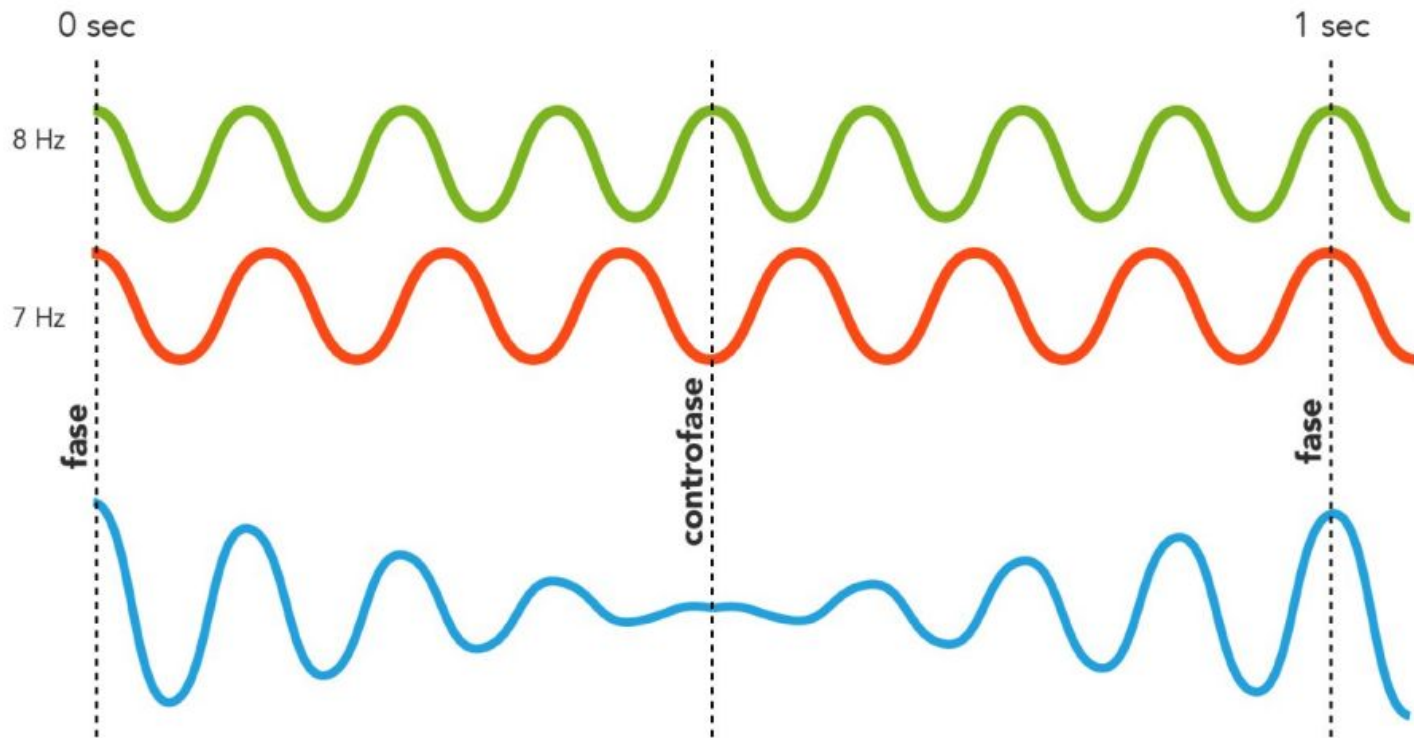
I battimenti

- 1.8.6 – Forma d'onda: i battimenti

In un editor audio generare due onde sinusoidali con frequenze vicine

- 300 Hz e 302 Hz, con ampiezza 0,5
- Mixare le due tracce

I battimenti



Velocità del suono

■ Velocità del suono

Calcolare la velocità del suono nell'aria alle seguenti temperature

- $T_1 = 0^\circ\text{C}$
- $T_2 = 20^\circ\text{C}$
- $T_3 = -20^\circ\text{C}$
- $T_4 = 35^\circ\text{C}$

Cosa succede
per $T =$
 $-273,15^\circ\text{C}$?

$$v_{m,T} = v_{m,0} + \alpha_m(T)$$

$$v_{aria,0} = 331,45 \frac{\text{m}}{\text{s}}$$
$$\alpha_{aria}(T) = 0,62 T$$

$$v = \sqrt{C_s \cdot R \cdot T} = \sqrt{1.4 \cdot 287.05 \cdot T} = \sqrt{401,8 \cdot T}$$

$$T_K = T_C + 273.15$$

$$v = \sqrt{401,8 \cdot (T_C + 273.15)}$$

$$v = \sqrt{401,8 \cdot (0 + 273.15)} = 331.28$$

$$v = \sqrt{401,8 \cdot (20 + 273.15)} = 343.20$$

$$v = \sqrt{401,8 \cdot (-20 + 273.15)} = 318.92$$

$$v = \sqrt{401,8 \cdot (35 + 273.15)} = 351.87$$

- $T_1 = 0^\circ\text{C}$
- $T_2 = 20^\circ\text{C}$
- $T_3 = -20^\circ\text{C}$
- $T_4 = 35^\circ\text{C}$

Cosa succede
per $T =$
 $-273,15^\circ\text{C}$?

$$v = \sqrt{401,8 \cdot (-273.15 + 273.15)} = \sqrt{401,8 \cdot 0} = 0$$

Formanti delle vocali

- 2.6.3 – Registrare una vocale e individuare le formanti

In un editor audio registrare in successione le vocali usando un microfono

- Visualizzare la traccia come sonogramma
- Osservare le principali regioni delle frequenze formanti:
 - i. A: 800-1200 Hz
 - ii. E: 400-600 Hz e 2200-2600 Hz
 - iii. I: 200-400 Hz e 3000-3500 Hz
 - iv. O: 400-600 Hz
 - v. U: 200-400 Hz

Formanti delle vocali



```
$ sudo apt install espeak
```



```
$ brew install espeak
```



```
$ winget install -e --id espeak.espeak
```

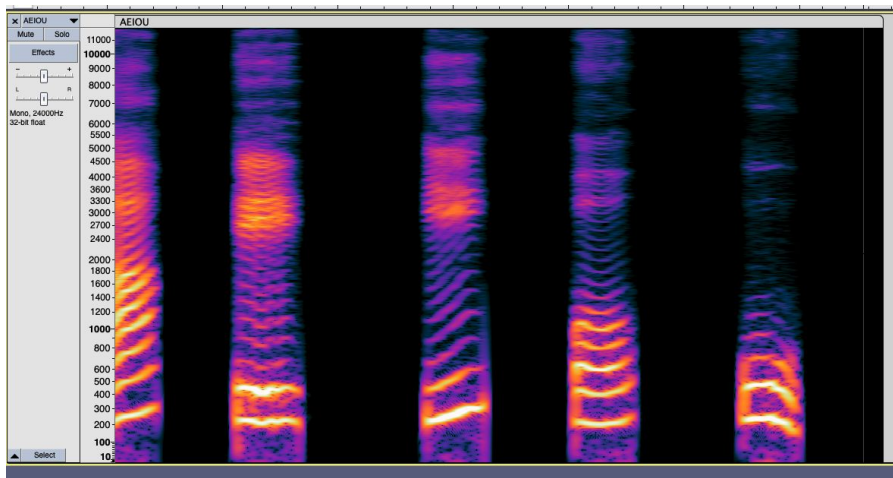
<https://espeak.sourceforge.net/download.html>

Generare “a, e, i, o, u”:

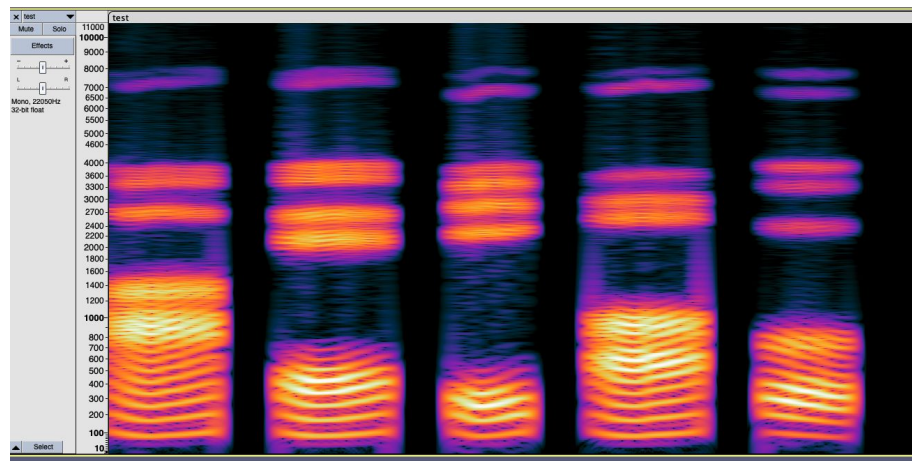
```
$ echo "a, e, i, o, u" | espeak -v it --stdout > aeiou.wav
```



Formanti delle vocali



eSpeak



Bitrate

- Cos'è il bitrate nell'audio digitale? Da quali valori dipende? Cosa rappresentano tali valori?
- Discutere delle operazioni di sovracampionamento e sottocampionamento. Cosa sono? Come si possono effettuare? Che effetti hanno?
- Un segnale audio non compresso con bitrate pari a 1200 kbps e frequenza di campionamento pari a 44000Hz viene sottocampionato 2 volte. La prima volta il bitrate viene portato a 600 kbps e la seconda volta a 300kbps. Dopo il primo sottocampionamento il segnale non presenta distorsioni. Dopo il secondo sottocampionamento si rileva invece una distorsione. È possibile una cosa del genere? Motivare.
- È possibile stabilire un range di valori per la frequenza di Nyquist di tale segnale? Se sì qual è?

Bitrate

- **Cos'è il bitrate nell'audio digitale? Da quali valori dipende? Cosa rappresentano tali valori?**

Bitrate è il termine utilizzato per descrivere la quantità di dati trasformati in audio. A un bitrate superiore corrisponde generalmente una migliore qualità dell'audio.

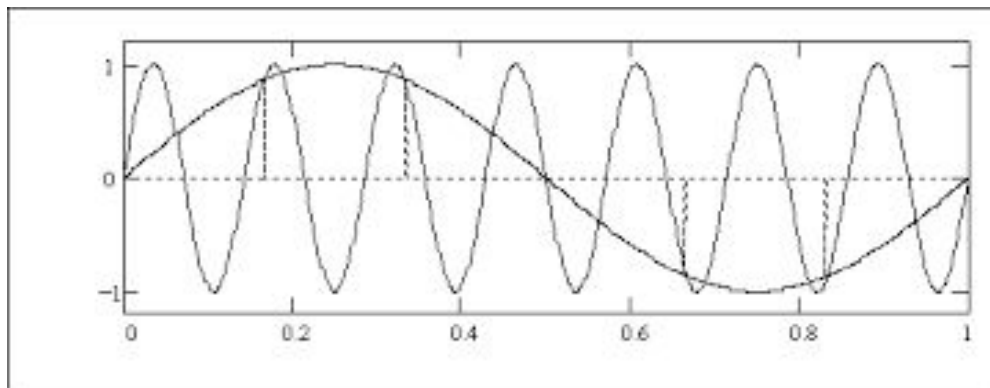
Bitrate è la quantità di informazioni rispetto ad un determinato intervallo di tempo.

Bitrate

- Discutere delle operazioni di sovracampionamento e sottocampionamento. Cosa sono? Come si possono effettuare? Che effetti hanno?

Il sovracampionamento è il processo di campionamento di un segnale a una frequenza di campionamento significativamente superiore al **Nyquist rate**.

Il sottocampionamento è un processo analogo dove la frequenza di campionamento è inferiore *rispetto alla precedente*.



Bitrate

LAME Bitrate Overview			
lame option	Average kbit/s	Bitrate range kbit/s	ffmpeg option
<code>-b 320</code>	320	320 CBR (non VBR) example	<code>-b:a 320k</code> (NB this is 32KB/s, or its max)
<code>-V 0</code>	245	220-260	<code>-q:a 0</code> (NB this is VBR from 220 to 260 KB/s)
<code>-V 1</code>	225	190-250	<code>-q:a 1</code>
<code>-V 2</code>	190	170-210	<code>-q:a 2</code>
<code>-V 3</code>	175	150-195	<code>-q:a 3</code>
<code>-V 4</code>	165	140-185	<code>-q:a 4</code>
<code>-V 5</code>	130	120-150	<code>-q:a 5</code>
<code>-V 6</code>	115	100-130	<code>-q:a 6</code>
<code>-V 7</code>	100	80-120	<code>-q:a 7</code>
<code>-V 8</code>	85	70-105	<code>-q:a 8</code>
<code>-V 9</code>	65	45-85	<code>-q:a 9</code>

Bitrate

quality scale

```
$ ffmpeg -i input.mp3 -q:a 0 output.mp3
```

0 — alta qualità

5 — medio

9 — bassa qualità

```
$ ffprobe -v error -print_format json -show_entries stream=bit_rate output-0.mp3 > output.json
```

Bitrate

- Un segnale audio non compresso con bitrate pari a 1200 kbps e frequenza di campionamento pari a 44000Hz viene sottocampionato 2 volte. La prima volta il bitrate viene portato a 600 kbps e la seconda volta a 300kbps. Dopo il primo sottocampionamento il segnale non presenta distorsioni. Dopo il secondo sottocampionamento si rileva invece una distorsione. È possibile una cosa del genere? Motivare.
- È possibile stabilire un range di valori per la frequenza di Nyquist di tale segnale? Se sì qual è?

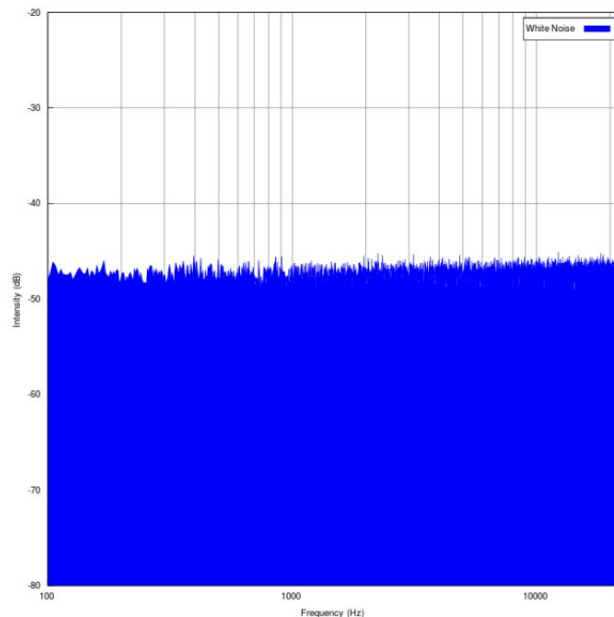
Rumore nei segnali audio

Quando un segnale sonoro si definisce “rumore bianco”?

Un segnale sonoro si definisce "rumore bianco" quando ha una densità spettrale di potenza costante su tutte le frequenze udibili nell'intervallo di udibilità umana

Per cosa può essere utilizzato tale rumore? Discutere.

Viene utilizzato per il mascheramento acustico, misurazione, filtri ecc.



Rumore nei segnali audio

Quale, tra i possibili tipi di rumore sonoro, viene impiegato per equalizzare un segnale per far sì che tutte le frequenze siano percepite allo stesso volume? Motivare la risposta.

Il rumore rosa ha una densità spettrale di potenza che diminuisce all'aumentare della frequenza per questo è utilizzato per equalizzare un segnale audio in modo che tutte le frequenze siano percepite allo stesso volume

