



Laboratorio di elettronica e tecniche d'acquisizione dati

Esperienza V: Moises

Gentili Irene, Pace Lorenzo, Tintori Filippo

13 dicembre 2024

Indice

1	Introduzione	2
2	Aspetti teorici	2
3	Strumentazione	3
4	Funzioni implementate	3
5	Risintetizzazione di segnali testuali (A)	4
6	Risintetizzazione di segnali wav	5
6.1	Diapason (B1)	6
6.2	Altri file (B1)	8
6.3	Ultimi file (B2)	12
7	Risintetizzazione di segnali musicali (B3)	14
8	Conclusioni	15

1 Introduzione

Lo scopo della presente relazione è mostrare la metodologia e i risultati relativi all'analisi di Fourier di segnali audio. In linea di massima avviene il passaggio al dominio delle frequenze tramite una trasformata, per poi filtrare i segnali adeguatamente e restituire un audio coerente attraverso un'antitrasformata.

2 Aspetti teorici

Lo strumento matematico alla base dell'esperienza è la trasformata di Fourier che permette il passaggio dal dominio dei tempi a quello delle frequenze nell'analisi di un segnale dipendente in un certo modo da t . La teoria di fondo è che ogni segnale (sia periodico che aperiodico, sia continuo che discreto) si può interpretare come sovrapposizione di componenti armoniche, ossia come somma di funzioni sinusoidali e cosinusoidali con determinate frequenze, o equivalentemente come somma di esponenziali complessi con determinate fasi. Per frequenze discrete e segnali continui nel tempo, per un segnale periodico di frequenza $f_0 = 1/T_0$ lo sviluppo in serie di Fourier si scrive come serie bilatera di contributi armonici complessi:

$$x(t) = \sum_{k=-\infty}^{\infty} X_k e^{2ik\pi f_0 t} \quad (1)$$

Dove i coefficienti delle componenti (X_k) si ottengono tramite prodotti scalari con l'antitrasformata:

$$X_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-2ik\pi f_0 t} dt \quad (2)$$

La (1) è detta *equazione di analisi* perché studia il contenuto in frequenza del segnale. La (2) è un'*equazione di sintesi* perché ricostruisce il segnale a partire dalle sue armoniche. La conoscenza del segnale $x(t)$ nel dominio temporale è equivalente alla conoscenza della successione dei termini di Fourier nel dominio delle frequenze. Ovviamente, anche se la serie è infinita dal punto di vista fisico i segnali si ricostruiscono a partire da un numero finito di armoniche. Ciò è tanto più evidente nel caso comune di segnali campionati, uscenti per esempio da un sistema ADC reale che estrae come sequenza di impulsi il valore $x(t)$ digitalizzato al tempo di campionamento. Con una sequenza finita l'intervallo di campionamento è un T_0 finito che definisce salti in frequenza $\Delta f = 1/T_0$; nel caso di un segnale campionato con N_0 campioni con frequenza *di campionamento* $f_c = 1/T$, si identifica la sequenza periodica (di periodo N_0) con un insieme di N_0 reali $[x_0, x_1, \dots x_{N-1}] \subset \mathbb{R}$. Il contenuto in frequenza del segnale è reso nell'equazione di analisi (trasformata per segnali campionati):

$$X_{s,k} = \frac{1}{N_0} \sum_{n=0}^{N_0-1} x[n] e^{-2ink\pi/N_0} \quad (3)$$

Mentre l'antitrasformata (equazione di sintesi):

$$x[n] = \sum_{k=0}^{N_0-1} X_{s,k} e^{2ink\pi/N_0} \quad (4)$$

Questa rappresentazione del segnale prende il nome di trasformata *discreta* di Fourier (DFT); per sequenze periodiche l'antitrasformata discreta consiste in una somma con un numero finito di addendi.

In via del tutto generale si definisce poi *potenza* $P(x)$ di un segnale $x(t)$ d'una certa energia E il limite

$$P(x) = \lim_{T \rightarrow \infty} \frac{E}{T} = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} |x(t)|^2 dt$$

Com'è evidente dal confronto con la (2) il significato fisico del modulo quadro delle X_k è allora quello di una potenza. In particolare il modulo (in generale $X_k \in \mathbb{C}$) quadro di ogni coefficiente X_k rappresenta il contributo di quel termine alla potenza del segnale e la somma di tutti i moduli quadri dei coefficienti è la potenza totale del segnale. Andando a ricostruire un segnale a partire dalle componenti si considereranno i termini dominanti in potenza, coscienti sempre però che perché il segnale ricostruito sia più affine possibile all'originale va considerato il maggior numero possibile di termini. Questo, ancora, perché i segnali reali non sono monocromatici ma presentano uno spettro di frequenze più o meno ampio. La precisa corrispondenza fra i metodi computazionali impiegati e quanto trattato in questa sezione teorica sarà discussa nel seguito.

3 Strumentazione

Gli strumenti e il materiale di cui si dispone per lo svolgimento dell'esperienza sono essenzialmente il linguaggio di programmazione Python e i file `.wav` e `.txt` forniti per l'elaborazione:

- `data1.txt` (A);
- `data2.txt` (A);
- `data3.txt` (A);
- `diapason.wav` (B1);
- `pulita_semplice.wav` (B1);
- `pulita_media.wav` (B1);
- `pulita_difficile.wav` (B1);
- `distorta.wav` (B1);
- `pulita_pezzo.wav` (B2);
- `distorta_pezzo.wav`; (B2)
- `primo.wav` (B3);
- `secondo.wav` (B3).

Le sigle tra parentesi indicano la sezione dell'esperimento dove i rispettivi file sono stati impiegati. Si è fatto uso delle librerie scientifiche `scipy` (particolarmente adatta all'analisi dei segnali) e `numpy`. Congiuntamente alla relazione si riporta in allegato il codice implementato per l'analisi dati.

4 Funzioni implementate

Le procedure analizzate al punto 2 per trovare i coefficienti di Fourier di un dato segnale e riscriverlo nel dominio delle frequenze sono integrate in determinate funzioni delle librerie scientifiche di Python. In particolare `numpy` contiene il modulo `fft.fft` che calcola la DFT unidimensionale (discreta in n termini) con un algoritmo basato sulla Fast Fourier Transform. La funzione ha per argomento in ingresso l'array dei dati e restituisce l'array di numeri complessi che rappresentano i coefficienti X_k dei termini di frequenza k . S'è fatto spesso uso di `np.real`, `np.imag` per estrarre le parti reali e immaginarie di questi coefficienti $X_k \in \mathbb{C}$.

Altrettanto importante è stata la funzione inversa `np.fft.ifft` che computa l'anttrasformata sempre in regime discreto in modo che `ifft(fft(a))==a` fino all'accuratezza della rappresentazione numerica: senza questa funzione non sarebbe infatti stato possibile riascoltare il segnale una volta mascherato in frequenza. Per il resto, l'unica parte di codice con una rilevanza meno tecnica e più 'fisica' è la funzione `find_peaks` importata da `scipy.signal` usata per trovare i picchi in potenza (gli indici delle frequenze ad essi relative).

5 Risintetizzazione di segnali testuali (A)

Nella prima parte dell'esperimento si è fatto uso dei tre file .txt a due colonne (tempo in s e ampiezza in u.a.) corrispondenti a 10 s di audio campionato a 44100 Hz. Si è iniziato rappresentando la *waveform* del segnale, ossia il grafico dell'evoluzione temporale dell'ampiezza. Dopo aver applicato la funzione `fft` al segnale si sono visualizzati i coefficienti di Fourier graficandone parte reale, parte immaginaria e modulo quadrato.

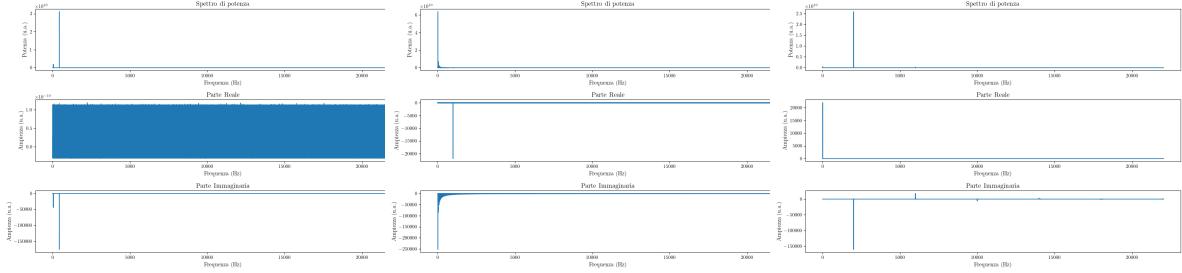


Figura 1: Spettro data1.txt

Figura 2: Spettro data2.txt

Figura 3: Spettro data3.txt

In questo modo è stato possibile individuare i picchi in potenza del segnale dato, ossia i punti che costituiscono dei massimi relativi della potenza come funzione della frequenza e che sono sopra un certo valore di soglia. Il senso computazionale dello stabilire manualmente il valore di questa soglia è che mascherando l'array (ossia ponendo a zero i coefficienti a potenza inferiore) si possono considerare solo gli indici relativi alle frequenze significative, risparmiando tra l'altro sensibilmente in termini di tempi di calcolo. La *risintetizzazione* è stata operata in due modi diversi: al fine di controllare il corretto comportamento della funzione `fft`, se ne sono confrontati i risultati con quelli ottenuti in termini di seni e coseni. Ossia, di quelli ottenuti implementando manualmente l'antitrasformata in termini di componenti armoniche:

$$x[n] = \frac{1}{N_0} \sum_{k=0}^{[N_0/2-1]} \left(\operatorname{Re}(X_k) \cdot \cos\left(\frac{2\pi kn}{N_0}\right) - \operatorname{Im}(X_k) \cdot \sin\left(\frac{2\pi kn}{N_0}\right) \right)$$

La divisione per N_0 è fatta per adattarsi alla definizione di `numpy` della trasformata inversa. Di seguito sono riportati i grafici relativi alla risintetizzazione. Nel primo segnale s'è preso il picco con potenza minore, distinguendo fra due sinusoidi il segnale e il rumore. Nel secondo file si sono considerati i primi 12 picchi in potenza: avendo notato uno spettro simile a quello dell'onda quadra, si è scelto di tagliare il picco che si discosta maggiormente dall'andamento $\propto k^{-1}$ tipico dei segnali quadrati. Nel terzo caso l'onda è triangolare; in questa circostanza l'andamento è $\propto k^{-2}$ e sulla base di questo si è tolto il primo picco, essendo l'unico non equidistante dagli altri.

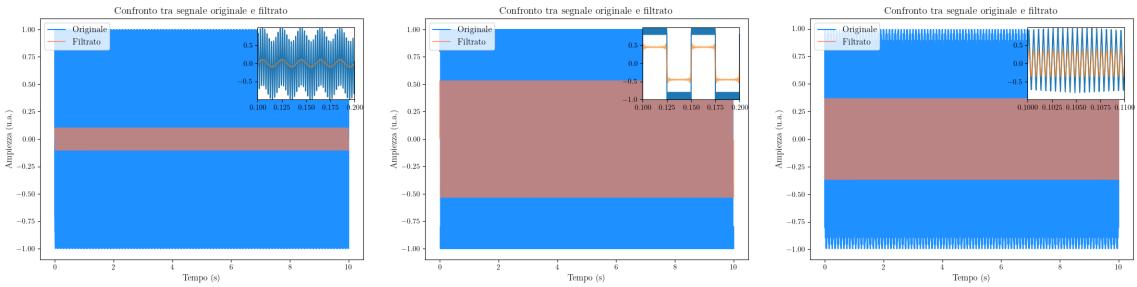


Figura 4: Segnale risintetiz-
zato con funzioni integrate
(data1.txt)

Figura 5: Segnale risintetiz-
zato con funzioni integrate
(data2.txt)

Figura 6: Segnale risintetiz-
zato con funzioni integrate
(data3.txt)

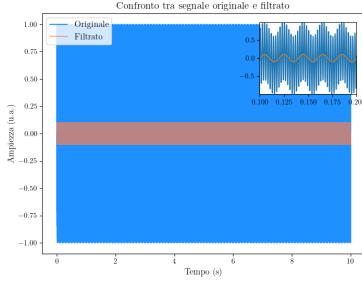


Figura 7: Segnale risintetizzato con seni e coseni (data1.txt)

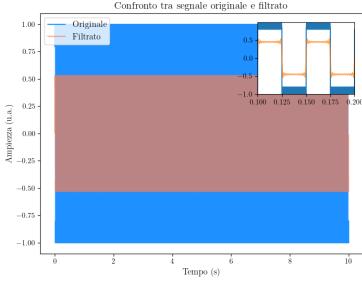


Figura 8: Segnale risintetizzato con seni e coseni (data2.txt)

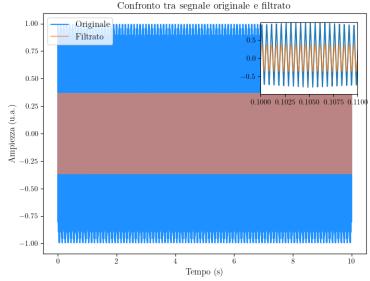


Figura 9: Segnale risintetizzato con seni e coseni (data3.txt)

6 Risintetizzazione di segnali wav

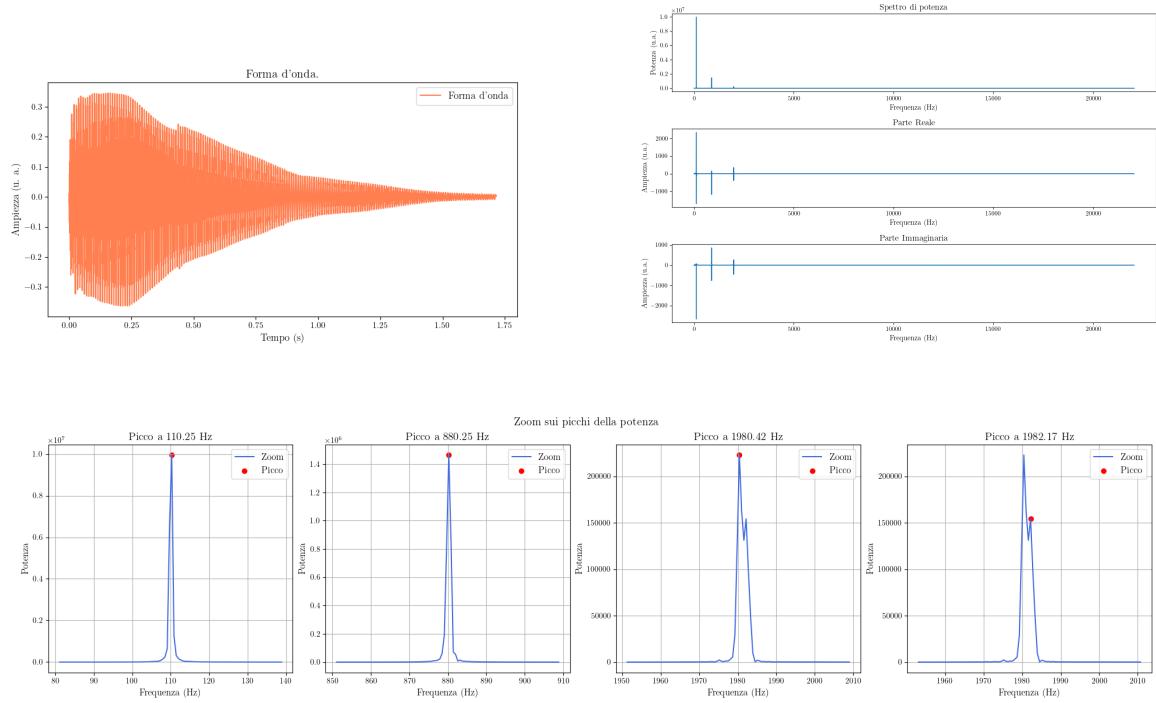
Nelle parti dell'esperimento successive alla A i file da analizzare sono .wav; per questo tipo di file il programma acquisisce i dati come matrice, o meglio array bidimensionale, separato in canali diversi (1, 2) per la tipologia (*stereo*) del sistema di input. Nel seguito si considererà solo il primo dei due canali dei file .wav. Per ogni file si è ricostruito il segnale tramite quanto già descritto al punto 5, risintetizzando con un numero sempre maggiore di armoniche. Nell'ordine si è considerato:

- il solo picco principale;
- il picco principale e il picco secondario;
- tutti i picchi principali, ma solo il termine centrale;
- tutti i picchi principali con in più i due termini immediatamente precedenti e immediatamente successivi.

Nella serie di antitrasformate fatte il risultato atteso è che la qualità del segnale migliori progressivamente aggiungendo le armoniche.

6.1 Diapason (B1)

Di seguito la waveform, lo spettro relativo al segnale del diapason e i picchi di quest'ultimo:



La soglia scelta per individuare i picchi della potenza è stata fissata a 10^4 u.a. in ampiezza. A questi picchi corrispondono le note La_2 ($\nu = 110,25 \text{ Hz}$), La_5 ($\nu = 880,25 \text{ Hz}$) e Si_6 ($\nu = 1980,42 \text{ Hz}$). Si nota che il terzo picco (quello del Si_6) ha un altro picco a potenza minore poco distante in frequenza; nella trattazione si è considerato il solo termine principale. Nel codice allegato sia il segnale del diapason che gli altri 4 relativi alla parte (B1) dell'esperimento sono stati analizzati sia in termini di funzioni `np.fft` sia tramite seni e coseni; si è scelto di riportare il confronto solo per il diapason, poiché per tutte le altre il risultato è esattamente lo stesso a conferma del corretto comportamento delle funzioni integrate. Infatti lo scarto fra le due rappresentazioni è in tutti i casi dell'ordine delle $\sim 10^{-14} - 10^{-15}$ u.a., cioè ai limiti delle approssimazioni numeriche. Si riportano le waveform risintetizzate, in ordine, attraverso l'antitrasformata e con seni e coseni:

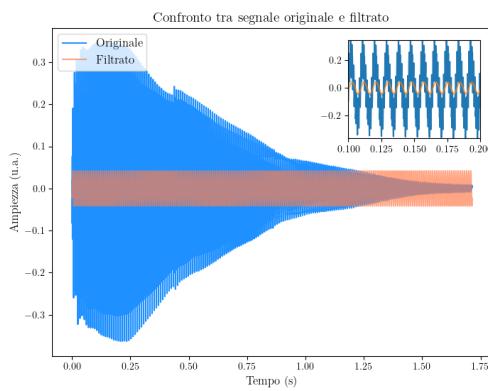


Figura 10: Tramite funzioni integrate (1).

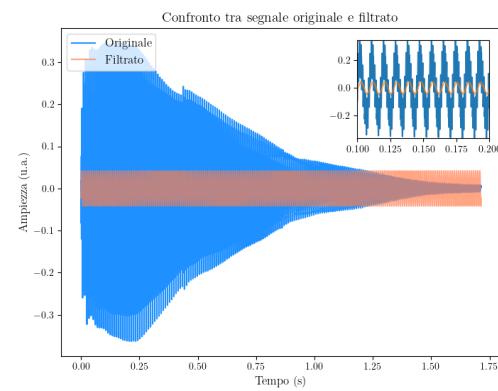


Figura 11: Tramite seni e coseni (1).

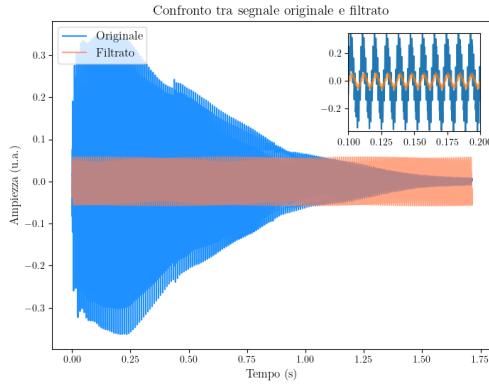


Figura 12: Tramite funzioni integrate (2).

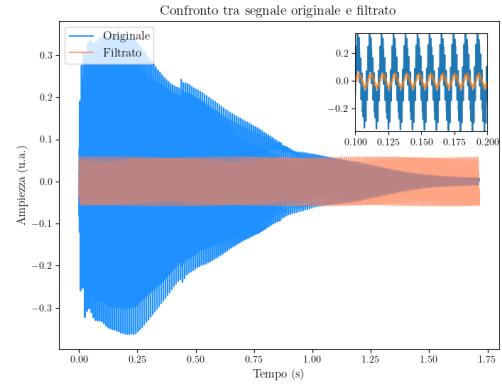


Figura 13: Tramite seni e coseni (2).

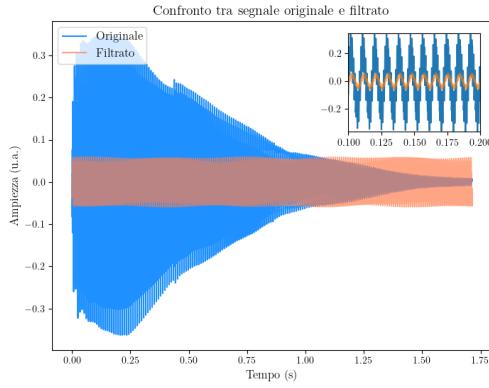


Figura 14: Tramite funzioni integrate (3).

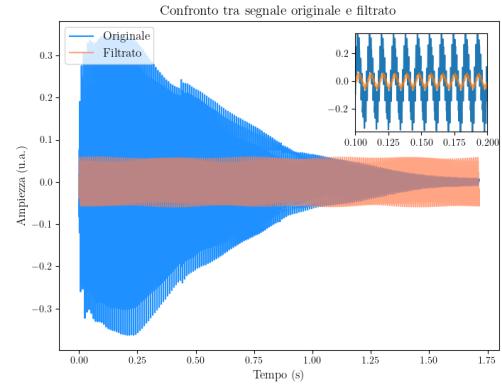


Figura 15: Tramite seni e coseni (3).

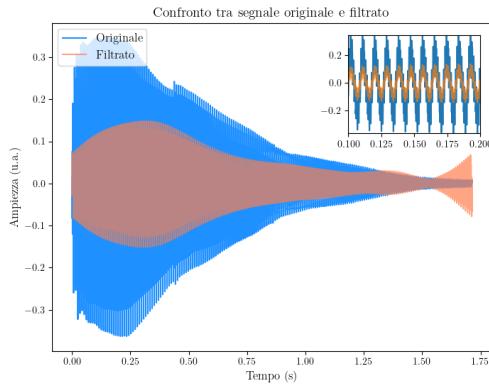


Figura 16: Tramite funzioni integrate (4).

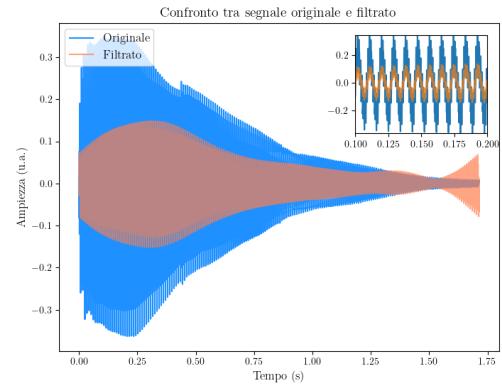
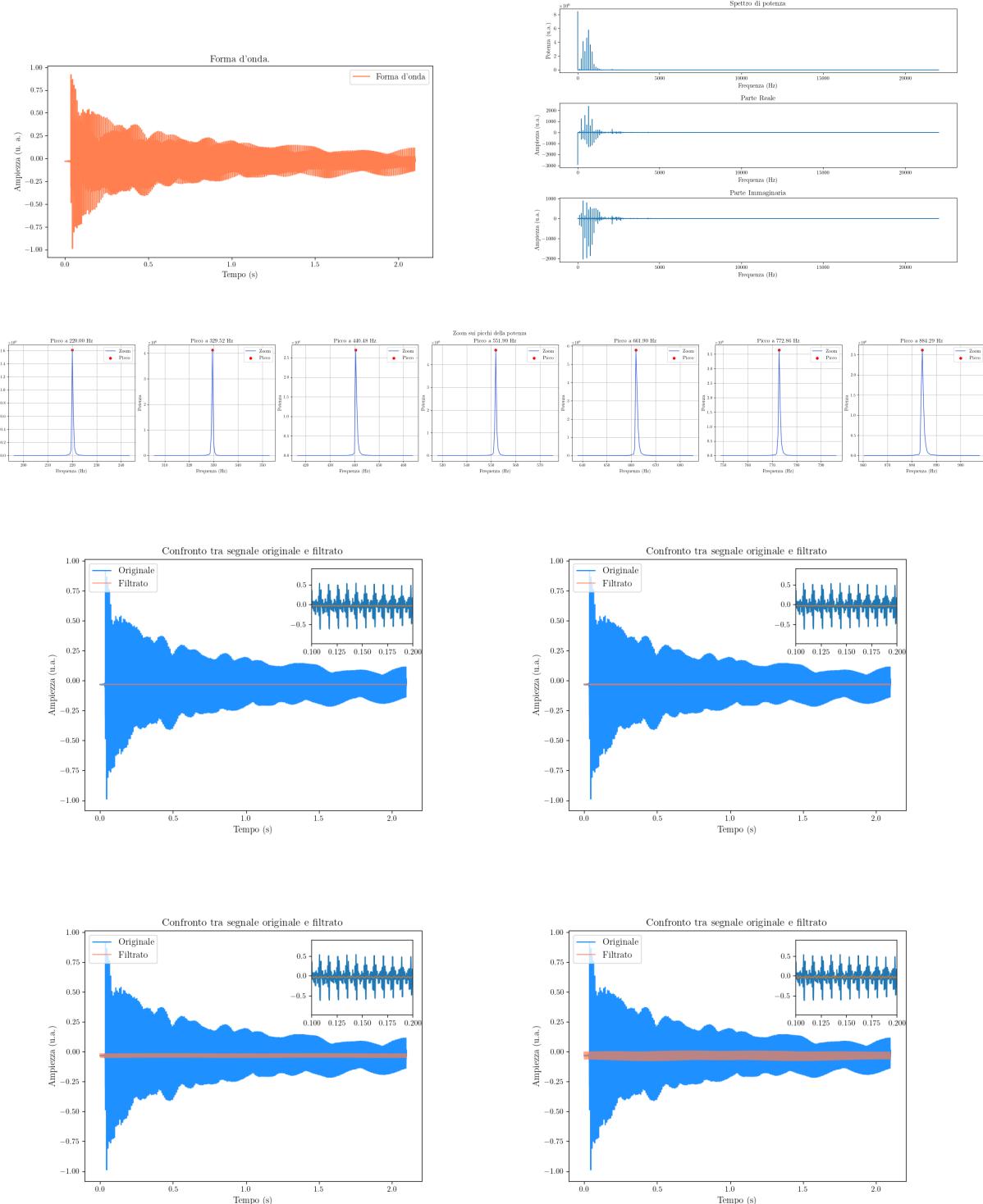


Figura 17: Tramite seni e coseni (4).

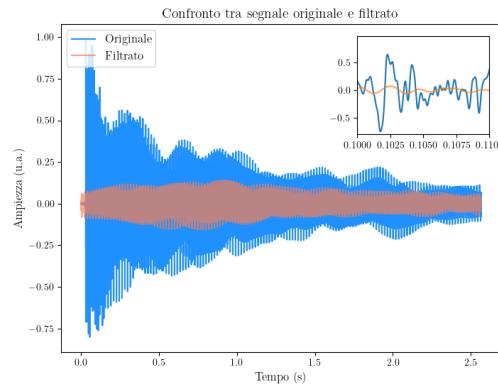
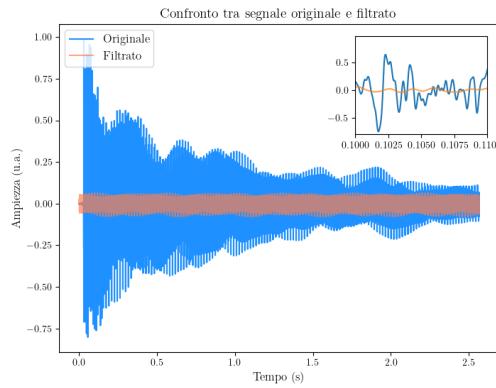
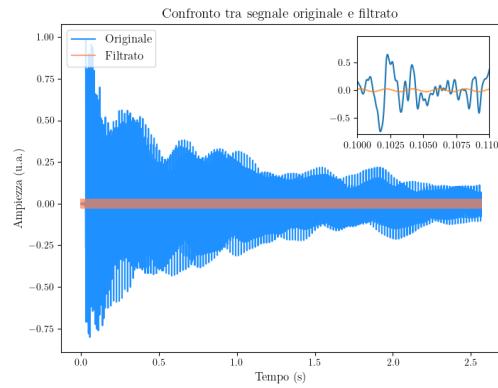
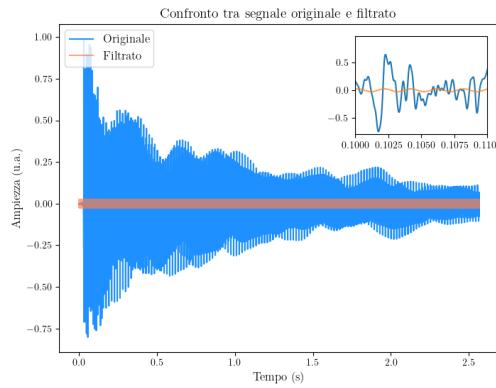
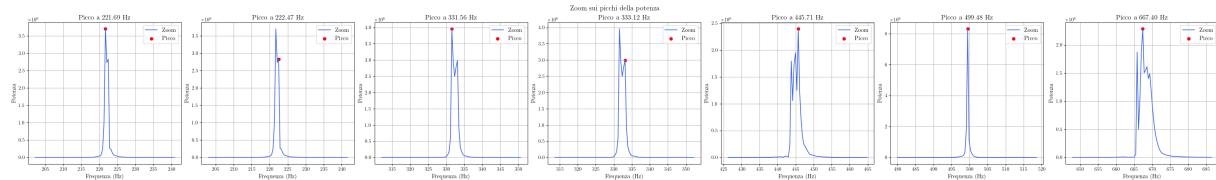
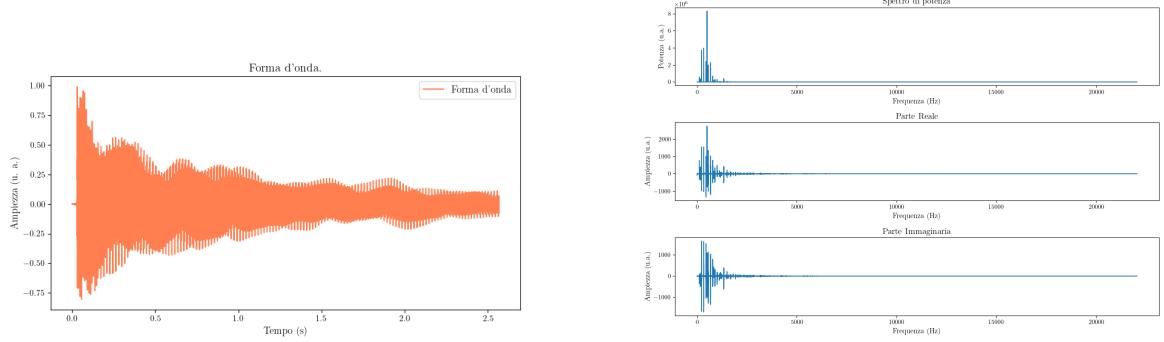
6.2 Altri file (B1)

Di seguito sono riportati il grafico d'onda, lo spettro, i picchi scelti con le apposite soglie, i grafici del segnale risintetizzato (nelle 4 forme indicate all'elenco nel paragrafo precedente).

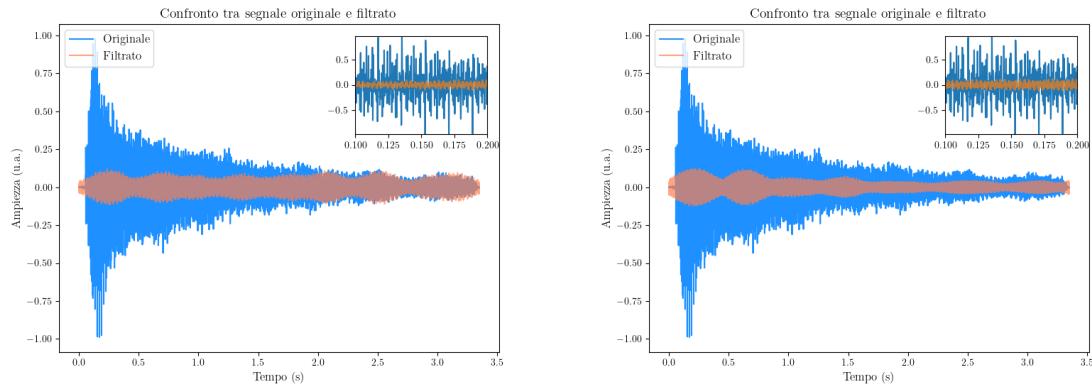
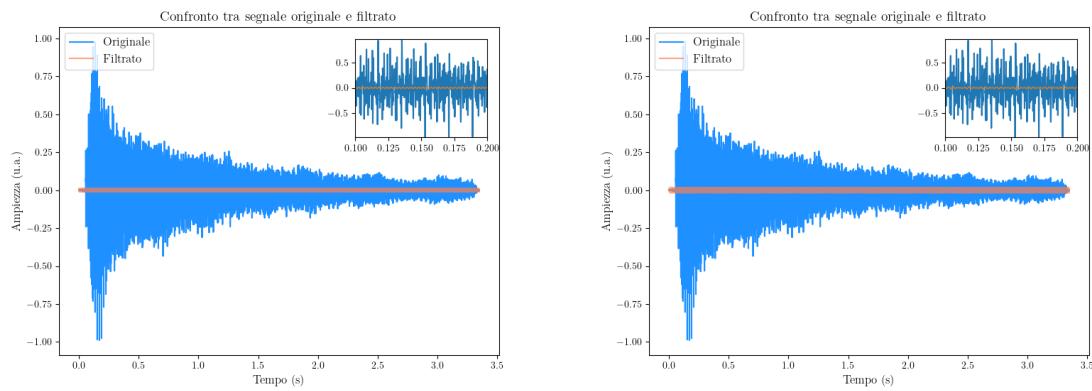
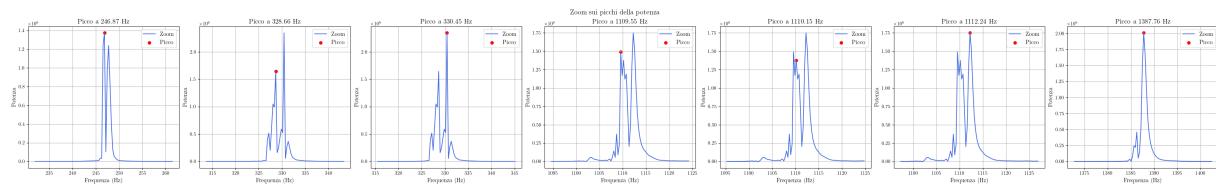
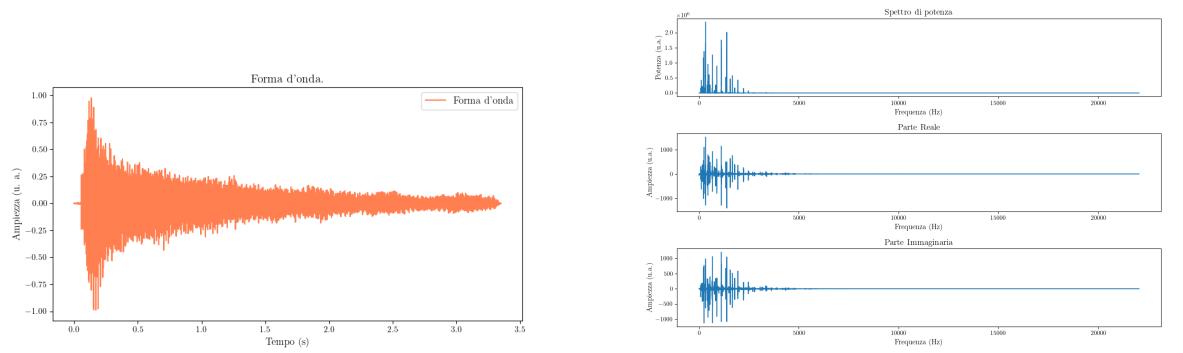
`pulita_semplice.wav`: soglia a 10^4 u.a. ;



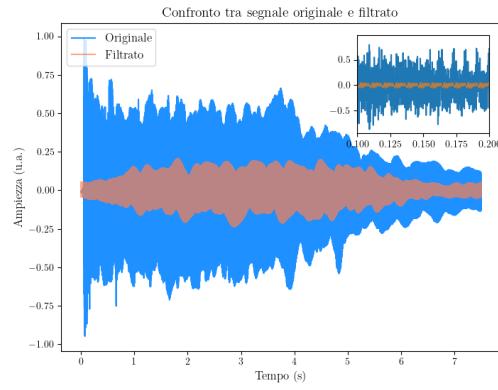
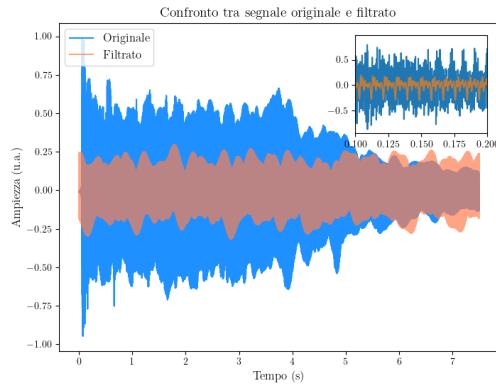
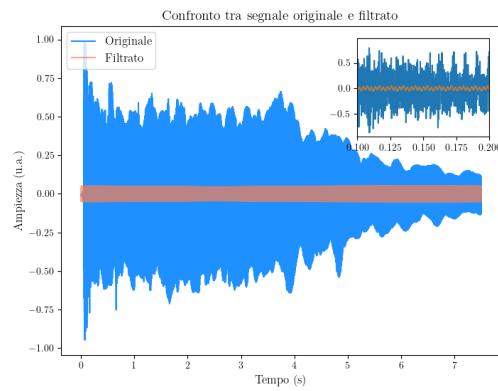
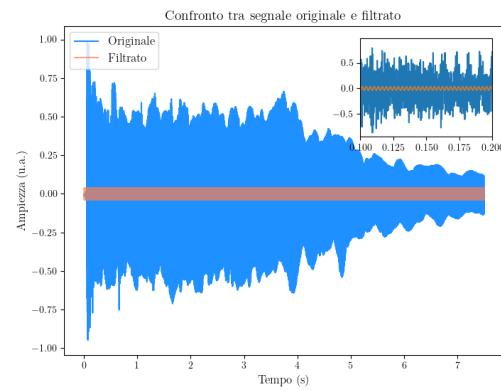
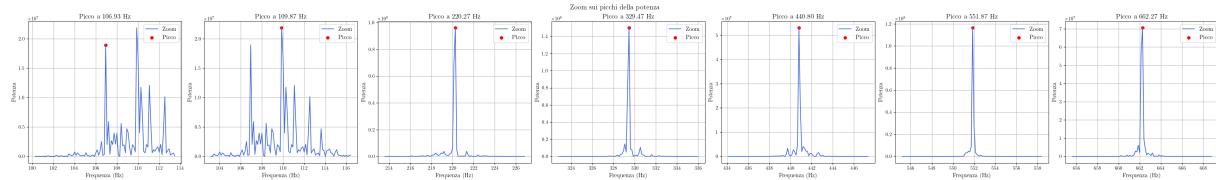
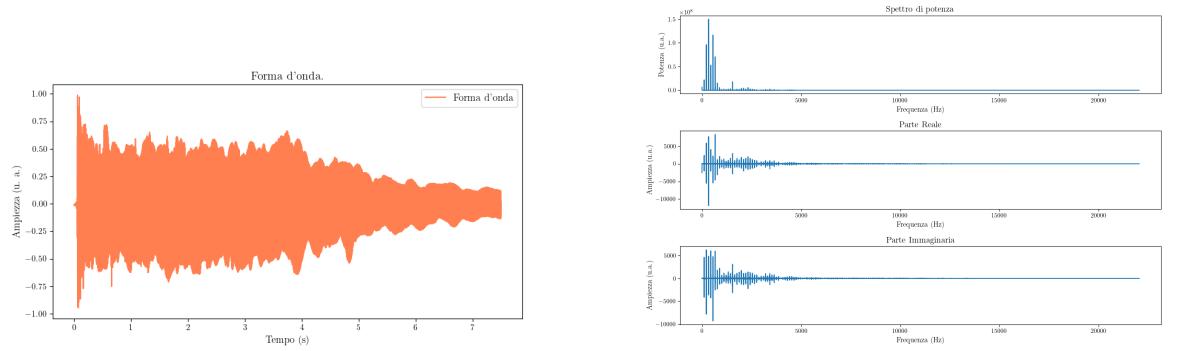
pulita_media.wav: soglia a $2 \cdot 10^6$ u.a. ;



pulita_difficile.wav: soglia a $2,3 \cdot 10^6$ u.a. ;



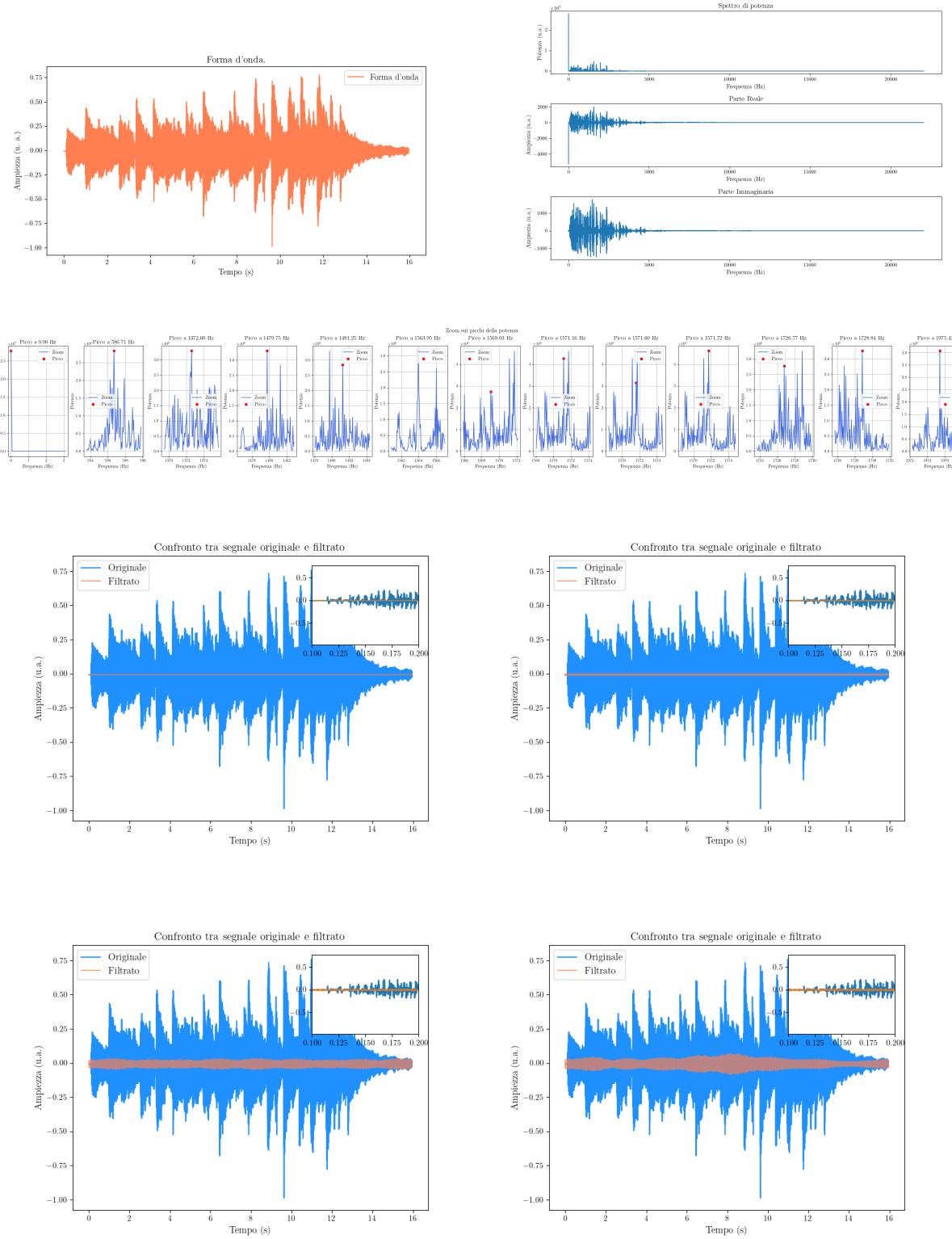
`distorta.wav`: soglia a $1,8 \cdot 10^7$ u.a.



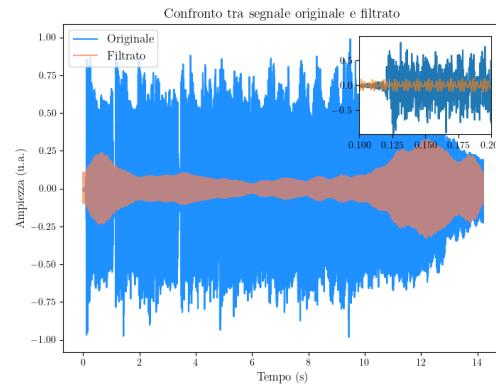
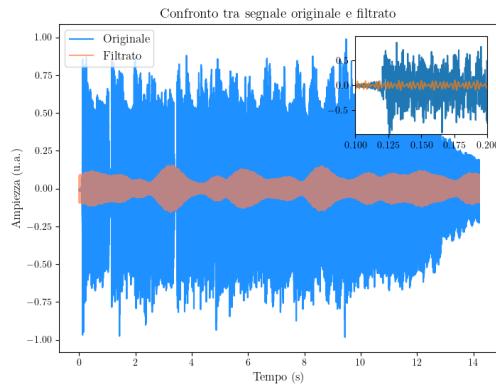
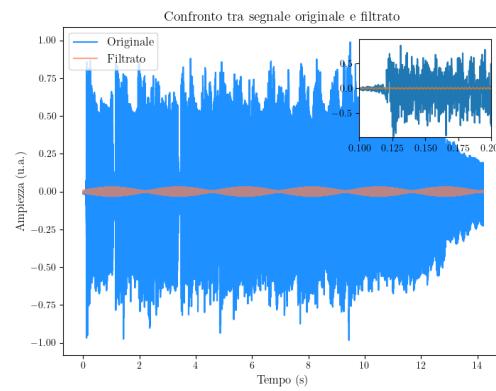
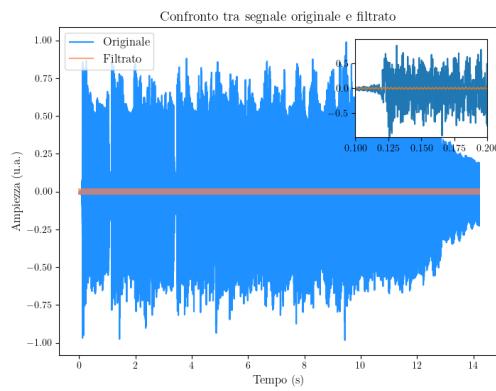
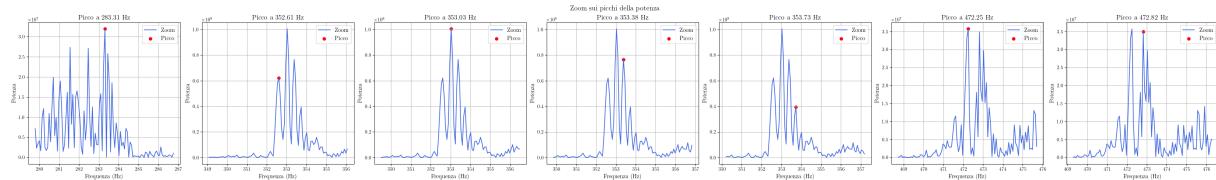
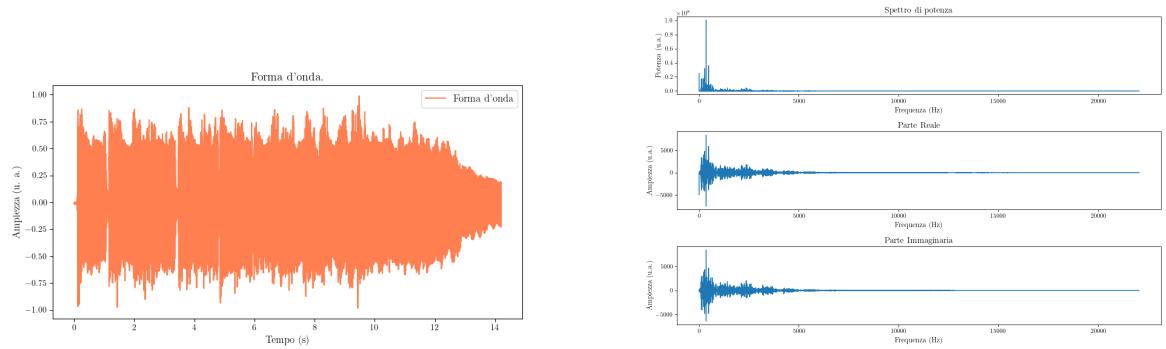
6.3 Ultimi file (B2)

L'analisi dei due file `pulita_pezzo.wav` e `distorta_pezzo.wav` ha seguito le stesse metodologie.

`pulita_pezzo.wav`: soglia a $3 \cdot 10^6$ u.a.;



`distorta_pezzo.wav`: soglia a $3 \cdot 10^7$ u.a.;



7 Risintetizzazione di segnali musicali (B3)

Nell'ultima parte dell'esperimento si è provato ad analizzare due segnali reali corrispondenti a 10 s delle intro di due canzoni. I file forniti contengono l'evidente sovrapposizione di due strumenti, di cui uno più acuto e uno più grave in tono. Nel seguito si sono graficati, al solito, la waveform e lo spettro dei coefficienti di Fourier. Al fine di distinguere i due strumenti si è implementato uno *spettrogramma*. Come si vede in figura, esso è essenzialmente un grafico della frequenza in funzione del tempo con una terza dimensione data dalla scala di colore; il colore identifica la potenza associata alle determinate frequenze evidenziando dove i due strumenti "piccano" nei due pezzi di canzoni. Confrontando questi dati con il pezzo di canzone si è stati in grado di identificare i due strumenti. Uno dei due (probabilmente un basso) ha picchi in potenza distribuiti sulle frequenze basse (150-1500 Hz per la prima canzone, 100-500 Hz per la seconda), mentre l'altro (chitarra) picca su frequenze più alte (2000-10000 Hz per la prima canzone, 800-10000 Hz per la seconda). Si sono salvati distintamente (su file .wav diversi) i due strumenti per riascoltarli e determinare l'accordo col canale unico originario. In alto la canzone 1, in basso la 2.

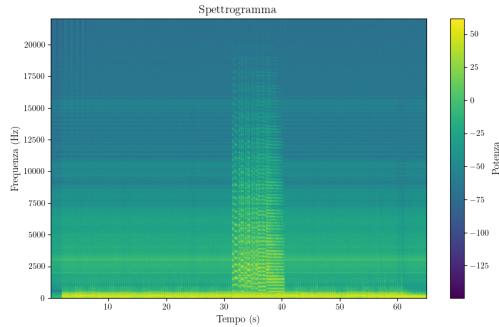
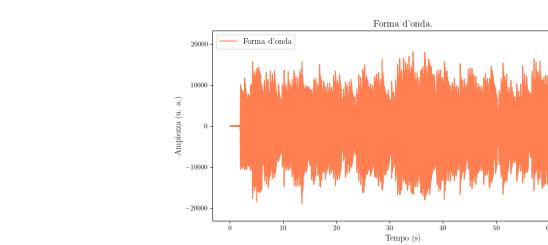


Figura 18: 1 (*Zombie*)

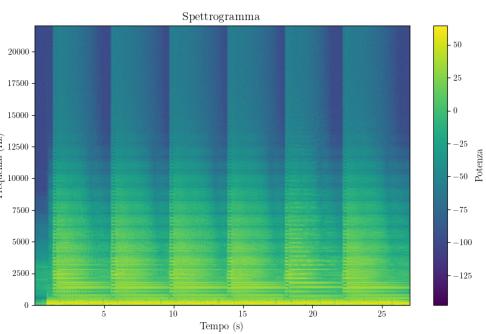
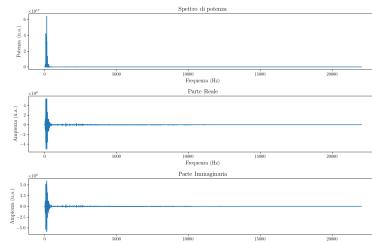
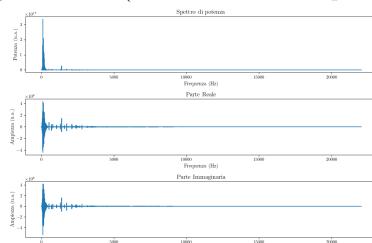
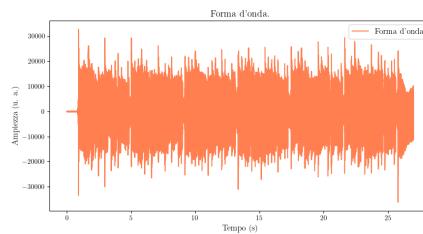


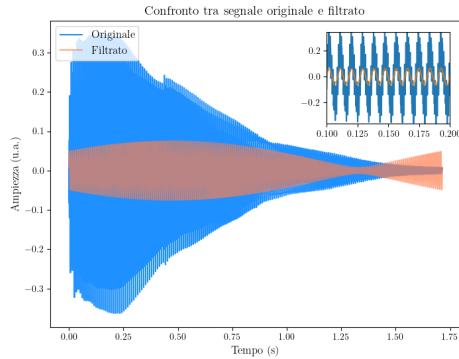
Figura 19: 2 (*Smells like teen spirits*)



8 Conclusioni

L'analisi portata avanti ha soddisfatto gli obiettivi fissati ed è in accordo con quanto trattato nella teoria. La parte di risintetizzazione di segnali da file .txt è servita a controllare il corretto funzionamento delle funzioni integrate, perché i grafici e i valori numerici delle trasformate mediante `np.fft` e tramite seni e coseni sono indistinguibili ai limiti della rappresentazione numerica. I rumori (singole sinusoidi) sono stati rimossi dai tre segnali (sinusoidale, quadro, triangolare). Della parte (B1) relativa alla risintetizzazione di .wav, vale la pena notare che nel diapason è stato ripetuto il confronto con la trasformata in termini di seni e coseni ed essa ha avuto esito positivo.

Avendo aumentato la risoluzione intorno ai picchi, è evidente che questi non sono singoli impulsi ma hanno una determinata larghezza. Dal punto di vista fisico questo è rilevante, ed è ciò che dà al *suono* inteso come segnale la sua particolarità. In molte delle figure riportate per i picchi in potenza sono presenti più picchi secondari intorno al massimo, tanto che in alcuni casi è stato necessario adattare più volte la soglia per ridurre sensibilmente il tempo computazionale. Quindi, associato a un contributo armonico (a una data frequenza) ci sono vari contributi secondari (a frequenze vicine). L'interferenza di queste armoniche prossime fra loro è il fenomeno dei *battimenti*, comune in tutta la fisica, che in generale determina un inviluppo del segnale che ne modula l'ampiezza. Se si considerano ad esempio due sole frequenze vicine f , $f + df$, il risultato della loro sovrapposizione è una sinusoide di frequenza pari a $f + df/2$ e di ampiezza modulata secondo un inviluppo sinusoidale di frequenza $df/2$. Ovviamente il caso può essere più complicato, possono essere presenti molti più picchi e anche nei casi considerati è stato verificato che aggiungendo un numero crescente di contributi armonici il segnale si fa sempre più vicino all'originale. Il fenomeno dei battimenti può intravedersi per esempio nella figura seguente che rappresenta il segnale filtrato per il diapason considerando l'intera larghezza del picco principale e non il solo valore centrale:



Infine nella parte conclusiva (B3) si è cercato di studiare un segnale progressivamente più complicato e si ritiene di aver individuato i due strumenti in maniera efficiente tramite lo spettrogramma.