

Matematica, musica, e computer

Introduzione a signal processing e serie di Fourier

Leonardo Robol

Settimana Matematica 2026

Dipartimento di Matematica, Pisa

Gli obiettivi di questo laboratorio sono:

- Capire come **modellizzare matematicamente** il suono.
- Sviluppare la teoria matematica per **rappresentare** suoni e musica su un **computer**.
- Simulare (i.e., scrivere un programma che generi suoni di) **strumenti musicali**.
- Applicare **effetti** (riverbero, pulire il rumore, ...).

Gli strumenti a nostra disposizione

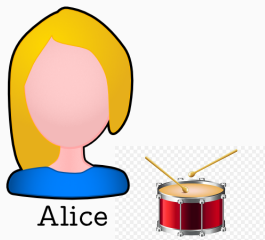
- Un **linguaggio di programmazione**: Python.
- Un po' di intuizione **fisica**.
- Un po' di **trigonometria** e qualche **integrale**.

Disclaimer: Gli integrali ci serviranno, ma non daremo quasi niente per scontato.

Ecco i veri protagonisti di questi 2 giorni:



Cos'è un suono? Supponiamo di fare un esperimento mentale, in cui Alice e Bob sono a qualche metro di distanza, e Alice percuote il tamburo.



Cosa succede fisicamente? Tutti immaginiamo che Bob sentirà il suono del tamburo, ma come viene generato e trasmesso?

Descriviamo la successione di eventi:

- Prima che il tamburo venga percosso, la pressione dell'aria è costante nella zona circostante.

Descriviamo la successione di eventi:

- Prima che il tamburo venga percosso, la pressione dell'aria è costante nella zona circostante.
- Quando la membrana del tamburo vibra, questo causa delle variazioni di pressione locale.

Descriviamo la successione di eventi:

- Prima che il tamburo venga percosso, la pressione dell'aria è costante nella zona circostante.
- Quando la membrana del tamburo vibra, questo causa delle variazioni di pressione locale.
- Queste vengono trasmesse attraverso l'aria.

Descriviamo la successione di eventi:

- Prima che il tamburo venga percosso, la pressione dell'aria è costante nella zona circostante.
- Quando la membrana del tamburo vibra, questo causa delle variazioni di pressione locale.
- Queste vengono trasmesse attraverso l'aria.
- Ad un certo punto, queste variazioni di pressione raggiungono le orecchie di Bob.

Descriviamo la successione di eventi:

- Prima che il tamburo venga percosso, la pressione dell'aria è costante nella zona circostante.
- Quando la membrana del tamburo vibra, questo causa delle variazioni di pressione locale.
- Queste vengono trasmesse attraverso l'aria.
- Ad un certo punto, queste variazioni di pressione raggiungono le orecchie di Bob.
- Questo manda in risonanza alcune parti dell'orecchio, che poi trasmettono un segnale elettrico ai nervi e da lì fino al cervello.

Descriviamo la successione di eventi:

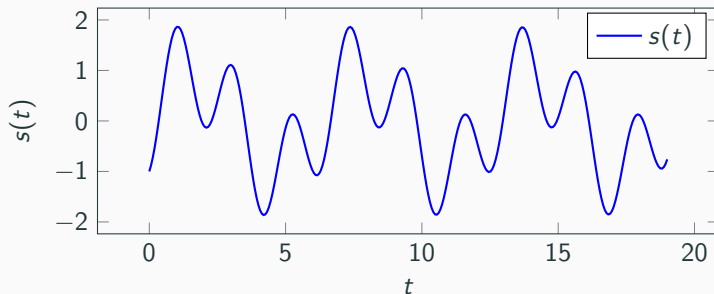
- Prima che il tamburo venga percosso, la pressione dell'aria è costante nella zona circostante.
- Quando la membrana del tamburo vibra, questo causa delle variazioni di pressione locale.
- Queste vengono trasmesse attraverso l'aria.
- Ad un certo punto, queste variazioni di pressione raggiungono le orecchie di Bob.
- Questo manda in risonanza alcune parti dell'orecchio, che poi trasmettono un segnale elettrico ai nervi e da lì fino al cervello.
- Bob sente il tamburo.

Realizzare un modello matematico

Per progettare il modello, facciamo alcune osservazioni:

- Il suono, secondo la nostra descrizione, corrisponde a **variazioni di pressione**.
- Per l'ascoltatore (in questo caso Bob) l'unica cosa rilevante sono le variazioni di pressioni **locali** nel suo orecchio.

Dunque, per rappresentare un suono possiamo modellare la variazione di pressione nel tempo, all'orecchio dell'ascoltatore; chiameremo questa funzione $s(t)$. Ad esempio, potrebbe assomigliare a questa:



Il nostro orecchio non rileva la pressione sonora istantanea, ma solo le variazioni che si ripetono nel tempo.

L'orecchio riconosce naturalmente segnali periodici di periodo T , ovvero funzioni tali che

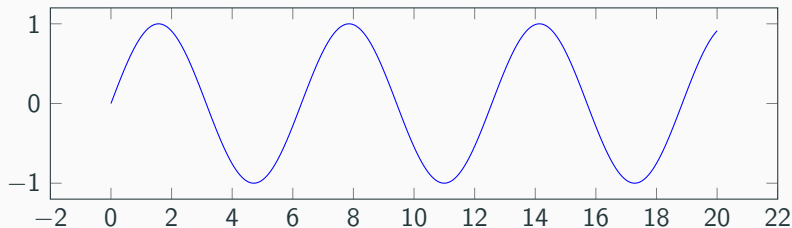
$$s(t + kT) = s(t), \quad \forall k \in \mathbb{Z}.$$

- Se una funzione ha **periodo** T , allora ha **frequenza** $\frac{1}{T}$. La frequenza si misura in Hz, definiti come $\frac{1}{s}$.
- Il nostro orecchio è in grado di riconoscere ampiezza, frequenza e forma della funzione. Queste vengono percepite rispettivamente come volume, altezza, e timbro.

Un esempio

Consideriamo la funzione periodica $s(t) = A \sin(2\pi\omega t)$.

- Il periodo è $\frac{1}{\omega}$, e la frequenza ω .
- La costante A regola l'intensità del volume.
- La forma è quella della funzione seno.



Che “suono” avrebbe questa funzione? Proviamo ad ascoltarlo.

Il modello di suono che abbiamo fino ad ora considerato utilizza funzioni su un continuo, i.e., definite su un intervallo $[0, T)$. Noi, però, vorremmo rappresentare l'audio su un computer.

- Un computer ha una quantità di memoria finita, mentre il numero di valori della funzione fra 0 e T è un numero **infinito** — ed anche piuttosto grande!
- Possiamo ovviare al problema considerando i valori della funzione solamente in un numero discreto di punti, ad esempio:

$$0, \Delta t, 2\Delta t, 3\Delta t, \dots, N\Delta t = T.$$

- Quanta informazione perdiamo in questo passaggio? Possiamo ricostruire il suono originale dalla sua misurazione in un numero finito di punti?

Cerchiamo di formalizzare questa riduzione.

- Consideriamo $s(t)$ una funzione $s(t) : [0, T] \rightarrow \mathbb{R}$.
- Fissiamo un intero $N > 0$, e definiamo una sequenza di punti x_j

$$x_j := s(j\Delta t), \quad \Delta t := \frac{T}{N}, \quad j = 0, \dots, N-1.$$

Osservazioni: data $s(t)$ e fissato N , il vettore x_j è **completamente determinato**. Il processo inverso, però, non è univoco!

Perché?

Scegliamo $T = 1$, e $\Delta t = \frac{1}{12}$.

- Se consideriamo $s_1(t) = \sin(6\pi t)$, otteniamo

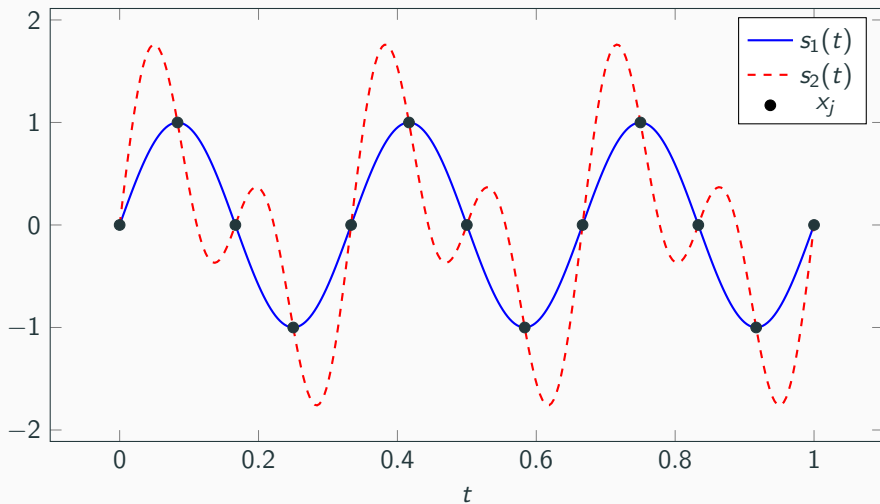
$$x_1 = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 \end{bmatrix}.$$

- Se invece prendiamo $s_2(t) = 2\sin(9\pi t)\cos(3\pi t)$, cosa succede? Non troppo sorprendentemente, i due vettori sono esattamente uguali:

$$x_2 = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 \end{bmatrix}.$$

Le funzioni $s_1(t)$, e $s_2(t)$ però sono distinte.

Un esempio (continua)



Questo fenomeno è noto come **aliasing**.

La situazione non è poi grave come sembra

- L'idea di misurare il segnale solo in un insieme finito è alla base di tutto l'audio digitale moderno — evidentemente non funziona così male.
- Abbiamo però visto che è possibile perdere molta informazione.

Fortunatamente, è possibile prendere degli accorgimenti perché questo non impatti sul nostro ascolto.

Teorema (Shannon-Nyquist)

Sia $s(t)$ un segnale ottenuto sovrapponendo segnali con frequenze comprese in un intervallo $[0, B]$, ovvero tale che

$$s(t) = s_1(t) + s_2(t) + \dots,$$

e $s_j(t)$ sia di periodo T con $\frac{1}{T} \in [0, B]$. Allora, la trasformazione da $s(t)$ al vettore discreto $x_j = s(j\Delta t)$ è univoca se $\Delta t < \frac{1}{2B}$.

Take-home message

Il teorema di Shannon ci dice: a patto di non dover rappresentare frequenze troppo alte, il suono è sempre recuperabile scegliendo un Δt sufficientemente piccolo.

Il teorema di Shannon ci dice: a patto di non dover rappresentare frequenze troppo alte, il suono è sempre recuperabile scegliendo un Δt sufficientemente piccolo.

- Quali frequenze sono di interesse per noi?
- ...e di conseguenza quanto va scelto questo Δt ?

- L'uomo può sentire solo frequenze fra 20 Hz e 20.000 Hz.
- Il limite superiore si abbassa notevolmente con l'età, fino a scendere anche sotto i 10.000 Hz.

Tutto quello che sta sotto i 20 Hz viene chiamato **infrasuono**, quello che sta sopra i 20.000 Hz **ultrasuono**.

- Nel rappresentare l'audio, è chiaro che questo ci dà un valore massimo per la frequenza B nel teorema di Shannon.
- Per rappresentare correttamente i 20.000 Hz, dobbiamo scegliere

$$\Delta t \leq \frac{1}{2B} \leq \frac{1}{40000} s,$$

che implica una frequenza di registrazione superiore a 40000 Hz.

Un viaggio negli anni '70

Immaginiamo di tornare a ~ 40 anni fa, dove alla Philips e alla Sony stavano progettando il primo CD audio.

- A quei tempi, erano ben consapevoli del teorma di Shannon, e dunque sapevano di doversi tenere sopra ai 40000 Hz.
- Tuttavia, nessuno voleva dare vantaggio al concorrente nel poter usare tool già a disposizione, e quindi litigarono parecchio fra 44000 Hz, 44100, 44056, e altre varianti.
- Alla fine vinse la Sony con 44100 Hz, che è diventato lo standard nei CD audio, e in molta musica digitale.

In alcune applicazioni si utilizzano altre frequenze, come 48 kHz, o 96 kHz. Alcuni tipi di trasmissioni utilizzando basse frequenze come 11025 Hz o 22050 Hz.

Riassumiamo la struttura del nostro modello:

- Un suono si può rappresentare matematicamente come una funzione $s(t)$, definita sull'intervallo di tempo $[0, t_{\max}]$ di interesse.
- Se questo suono corrisponde ad un'unica nota prodotta da uno strumento musicale, la funzione $s(t)$ è periodica di periodi $T = \frac{1}{f}$, con f uguale alla frequenza della nota scelta.
- La *forma* di questa funzione periodica determina il timbro dello strumento che ascoltiamo.

Ma come si ricava la frequenza di una determinata nota?

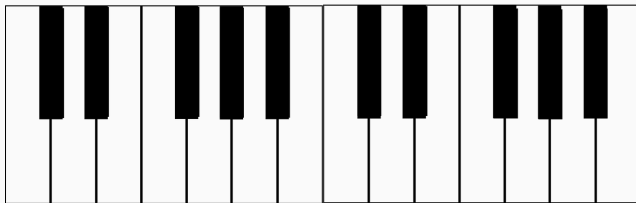
Il La a 440 Hz, e le altre note

Prendiamo un riferimento iniziale: il la sopra il Do centrale nella tastiera del pianoforte ha — per convenzione — la frequenza di 440 Hz.

Tutte le altre note si possono ricavare con le seguenti osservazioni:

- Per salire di un'ottava, basta **raddoppiare** la frequenza, per scendere bisogna **dimezzarla**.
- Le altre note sono distribuite uniformemente in scala **logaritmica**.

Possiamo chiarirci le idee guardando la tastiera di un pianoforte.



Chiamiamo f_{Do} la frequenza del Do centrale.

Per salire di un semitono (e raggiungere il Do $\sharp$), abbiamo bisogno di moltiplicare la frequenza per un numero positivo $\lambda_s > 1$.

Dunque:

$$f_{Do\sharp} = f_{Do}\lambda_s, \quad f_{Re} = f_{Do\sharp}\lambda_s = f_{Do}\lambda_s^2, \quad \dots$$

La frequenza del Do sopra il Do centrale (chiamiamolo temporaneamente Do2), soddisferà dunque:

$$f_{Do2} = 2f_{Do} = f_{Do}\lambda_s^{12},$$

perché ci sono esattamente 12 semitoni in un'ottava. Dunque, $\lambda_s = \sqrt[12]{2}$.

La scala logaritmica

- Il nostro orecchio percepisce l'altezza dei suoni in modo naturale secondo questa scala.
- Questo non vale solo per l'altezza, ma anche per il volume! In effetti, anche questo si misura in decibel che è un modo implicito per tenere traccia di $\log(A)$ invece che di A , dove A è la massima ampiezza del segnale.

Come possiamo calcolare efficacemente la frequenza di una nota?

- Supponiamo di sapere la frequenza f_{Do} del Do centrale.
- Determiniamo i j semitoni di differenza.
- Allora

$$f = f_{Do} \cdot \sqrt[12]{2^j} \iff f = 2^{\log_2(f_{Do}) + \frac{j}{12}}$$

Un prototipo di sintetizzatore

Supponiamo di voler scrivere un programma che suoni un Do per un secondo, poi un Re, e poi un Mi. Secondo la nostra regola:

$$f_{Do} \approx 523 \text{ Hz}, \quad f_{Re} \approx 587 \text{ Hz}, \quad f_{Mi} \approx 659 \text{ Hz}$$

Dunque, possiamo definire:

$$s(t) = \begin{cases} \sin(2\pi f_{Do} t) & 0 \leq t < 1 \\ \sin(2\pi f_{Re} t) & 1 \leq t < 2 \\ \sin(2\pi f_{Mi} t) & 2 \leq t \leq 3 \end{cases}.$$

Dovendola assemblare su un computer, valuteremo $s(t)$ solo nei punti $t = \Delta t \cdot j$, con $\Delta t = \frac{1}{44100}$.

Ci manca però un fattore fondamentale: questa funzione avrà la “forma” giusta?

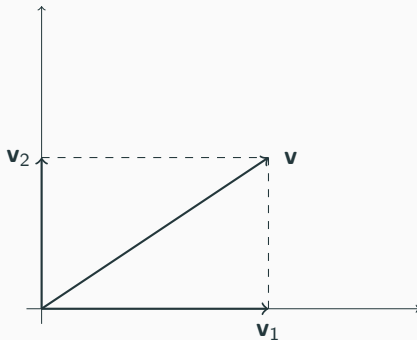
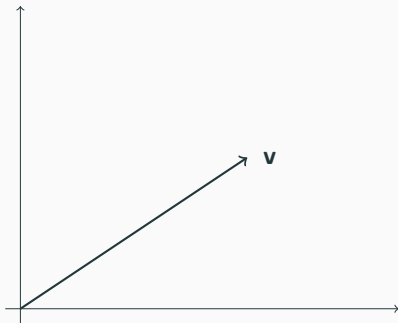
Se vogliamo davvero realizzare un sintetizzatore, ci rimane da risolvere il problema più annoso: come realizziamo una funzione $s(t)$ periodica ma con la “forma” giusta?

Per rispondere a questa domanda, abbiamo bisogno di un detour per argomenti apparentemente scorrelati, ovvero **vettori** e **prodotti scalari**.

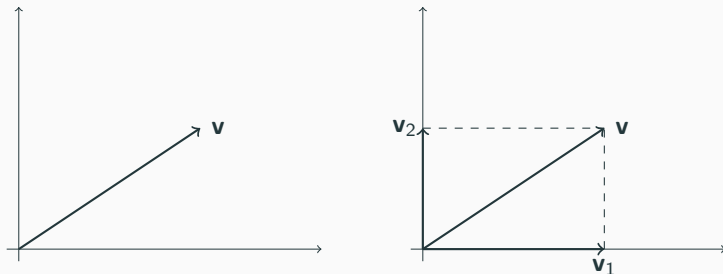
Vettori

Cos'è un vettore? Per il momento non ci preoccupiamo di dare una definizione precisa, e partiamo da un esempio: i vettori all'interno del piano cartesiano $\mathbb{R}^2$.

Graficamente, possiamo rappresentare un vettore $\mathbf{v}$ e leggere le sue componenti lungo l'asse delle x e delle y .



Sistemi di riferimento ortogonali



Le componenti $\mathbf{v}_1$ e $\mathbf{v}_2$ sono semplicemente una delle tante decomposizioni additive di $\mathbf{v}$, ovvero una scrittura di $\mathbf{v}$ come $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$.

- $\mathbf{v}_1$ e $\mathbf{v}_2$ sono multipli di due vettori di lunghezza unitaria lungo l'asse x ed y , che chiamiamo $\mathbf{e}_1$ ed $\mathbf{e}_2$. In particolare,

$$\mathbf{v}_1 = v_1 \mathbf{e}_1, \quad \mathbf{v}_2 = v_2 \mathbf{e}_2$$

Decomposizione ortogonale

Questa procedura di scomposizione ci ha fornito la scrittura:

$$\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2$$

Per un vettore nel piano, ci bastano sempre due vettori $\mathbf{e}_1$ ed $\mathbf{e}_2$; possiamo usare la stessa idea per “decomporre” una funzione periodica? Quante componenti ci servono?

Obiettivo, scrivere:

$$s(t) = a_1 s_1(t) + a_2 s_2(t) + \dots$$

Decomposizione di funzioni periodiche

Possiamo osservare che esiste un modo di comporre seni e coseni per costruire funzioni periodiche di periodo 2π .

Infatti, per ogni scelta di coefficienti a_i per $i = 0, \dots, k$ e b_i per $i = 1, \dots, k$, la funzione

$$s(t) = \frac{a_0}{2} + a_1 \cos(t) + b_1 \sin(t) + a_2 \cos(2t) + b_2 \sin(2t) + \dots + a_k \cos(kt) + b_k \sin(kt)$$

sia una **funzione periodica** su $[0, 2\pi]$.

Decomposizione di funzioni periodiche

Possiamo osservare che esiste un modo di comporre seni e coseni per costruire funzioni periodiche di periodo 2π .

Infatti, per ogni scelta di coefficienti a_i per $i = 0, \dots, k$ e b_i per $i = 1, \dots, k$, la funzione

$$s(t) = \frac{a_0}{2} + a_1 \cos(t) + b_1 \sin(t) + a_2 \cos(2t) + b_2 \sin(2t) + \dots + a_k \cos(kt) + b_k \sin(kt)$$

sia una **funzione periodica** su $[0, 2\pi]$.

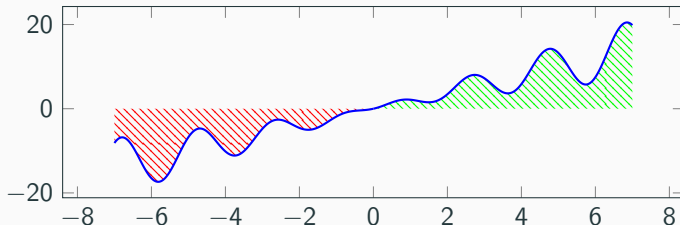
Si può fare anche il percorso inverso? Ovvero, possiamo, data $s(t)$, trovare le sue componenti a_i, b_i “lungo le funzioni”

$$1, \sin(t), \cos(t), \sin(2t), \cos(2t), \dots$$

Proiezioni

I coefficienti a_i e b_i sono le **proiezioni** di $s(t)$ sulle funzioni di base che abbiamo scelto, e possono essere “misurati” usando degli integrali.

$$\int_a^b f(t) dt = \text{“Area sottesa al grafico di } f(t) \text{ fra } a \text{ e } b\text{”},$$



il che in particolare permette di mostrare abbastanza facilmente che:

$$\int_0^{2\pi} f_1(t) + f_2(t) dt = \int_0^{2\pi} f_1(t) dt + \int_0^{2\pi} f_2(t) dt.$$

Passo 0:

Cosa succede se calcoliamo l'integrale

$$\int_0^{2\pi} s(t) dt?$$

Passo 0:

Cosa succede se calcoliamo l'integrale

$$\int_0^{2\pi} s(t) dt?$$

Usando la linearità degli integrali, otteniamo:

$$\int_0^{2\pi} s(t) dt = \int_0^{2\pi} \frac{a_0}{2} dt + \sum_{i=1}^k \left(a_i \int_0^{2\pi} \cos(it) dt + b_i \int_0^{2\pi} \sin(it) dt \right).$$

Osserviamo che:

$$\int_0^{2\pi} \cos(it) dt = 0, \quad \int_0^{2\pi} \sin(it) dt = 0,$$

per ogni $i \geq 1$. Dunque:

$$\int_0^{2\pi} s(t) dt = \frac{a_0}{2} \cdot 2\pi = a_0\pi \implies a_0 = \frac{1}{\pi} \int_0^{2\pi} s(t) dt.$$

Un trucco simile si può usare anche per gli altri coefficienti (come?).

Tramite un computer è facile approssimare questi integrali, e trovare i coefficienti

$$a_0, a_1, b_1, \dots$$

- Possiamo usare questa idea per costruire funzioni periodiche con varie forme;
- Data una registrazione di uno strumento, possiamo identificare i coefficienti a_i, b_i , e usarli per riprodurre lo stesso timbro;
- Cambiando la frequenza, possiamo ottenere note diverse.

Abbiamo tutti gli ingredienti per costruire un sintetizzatore.

Convergenza

In effetti, nessuno ci obbliga a fermarci a sommare un numero finito di funzioni seno e coseno, e possiamo estendere l'interesse a funzioni del tipo:

$$\begin{aligned} s(t) &= \frac{a_0}{2} + a_1 \cos(t) + b_1 \sin(t) + a_2 \cos(2t) + b_2 \sin(2t) + \dots \\ &= \frac{a_0}{2} + \sum_{i=1}^{\infty} (a_i \cos(it) + b_i \sin(it)). \end{aligned} \tag{1}$$

È naturale chiedersi se ogni funzione periodica su $[0, 2\pi]$ debba per forza essere del tipo sopra, per qualche opportuna scelta di a_i, b_i .

Convergenza

In effetti, nessuno ci obbliga a fermarci a sommare un numero finito di funzioni seno e coseno, e possiamo estendere l'interesse a funzioni del tipo:

$$\begin{aligned} s(t) &= \frac{a_0}{2} + a_1 \cos(t) + b_1 \sin(t) + a_2 \cos(2t) + b_2 \sin(2t) + \dots \\ &= \frac{a_0}{2} + \sum_{i=1}^{\infty} (a_i \cos(it) + b_i \sin(it)). \end{aligned} \tag{1}$$

È naturale chiedersi se ogni funzione periodica su $[0, 2\pi]$ debba per forza essere del tipo sopra, per qualche opportuna scelta di a_i, b_i .

Determinare esattamente quando questo succede (e come succede) ha impegnato molto i matematici alla fine del XIX secolo. Come regola generale, più la funzione è “regolare”, più è veloce la sua convergenza.

Chiamiamo $s_k(t)$ per $k = 0, 1, \dots$ la somma parziale dei coefficienti di Fourier ottenuta sommando fino all'indice k . Per cui:

$$s_0(t) = \frac{a_0}{2}, \quad s_1(t) = \frac{a_0}{2} + a_1 \cos(t) + b_1 \sin(t), \quad \dots$$

Esempi di convergenza

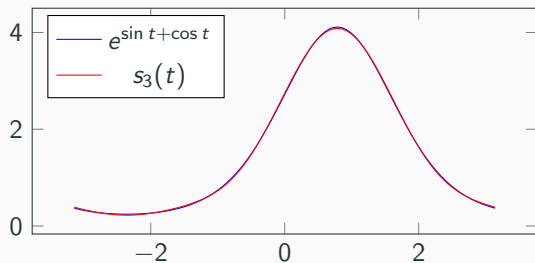
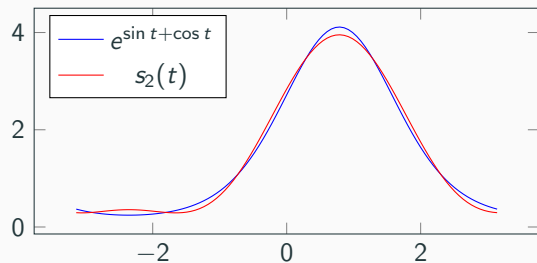
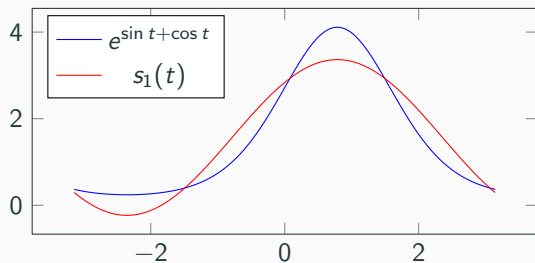
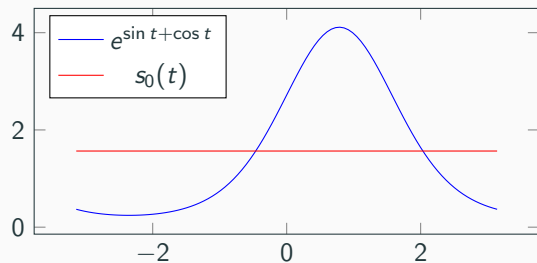
Facciamo qualche esempio per chiarirci le idee. Consideriamo la funzione $s(t) = e^{\sin t + \cos t}$, definita sull'intervallo $[-\pi, \pi]$, e osserviamo che:

- La funzione $s(t)$ è **periodica di periodo 2π** , i.e., se la estendessimo su tutto $\mathbb{R}$ avremmo $s(t + 2k\pi) = s(t)$ per ogni $k \in \mathbb{Z}$.
- La funzione è C^m per ogni $m \geq 0$, ovvero è C^∞ : possiamo **derivarla un numero arbitrario di volte**, e otterremo sempre una funzione continua (e periodica!).
- La funzione non è data da una somma finita di seni e coseni, i.e.,

$$s(t) \neq \sum_{i=0}^k a_i \cos(it) + b_i \sin(it)$$

per ogni scelta di a_i, b_i e k finito.

Convergenza per una funzione C^∞

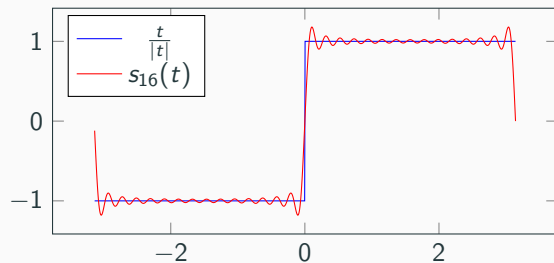
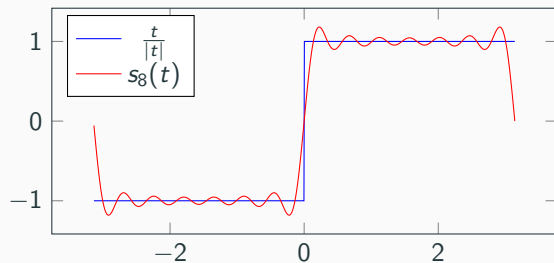
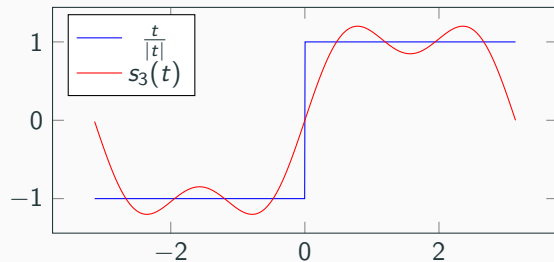
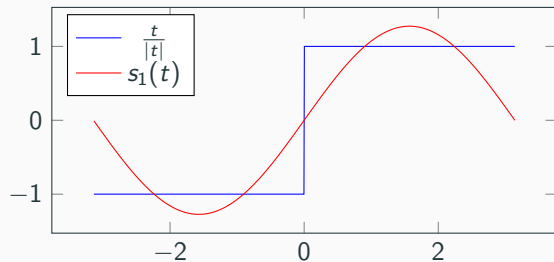


Consideriamo invece un esempio di funzione discontinua, come $s(t) = \frac{t}{|t|}$, definita sempre su $[-\pi, \pi]$, e posta uguale a 0 in 0. Questa funzione vale -1 sui numeri negativi, e 1 su quelli positivi.

Dunque, il suo grafico fa un “salto” di altezza 2 ad $t = 0$. Seni e coseni sono funzioni continue, e lo stesso vale per le loro somme.

Cosa ci aspettiamo dalla convergenza?

Un esempio più critico (con fenomeno di Gibbs)



- Abbiamo trovato un modo di decomporre una somma di seni e coseni usando dei prodotti scalari, e misurare quanto di una certa frequenza c'è all'interno di un segnale.
- Questo ci permette di **rappresentare molto forme d'onda** come combinazioni di seni e coseni, che sono semplici da costruire!
- Come si **ottengono i coefficienti** a_i, b_i ? Per ora, ci fideremo del computer che sa approssimare molto bene quegli integrali.
- Problemi aperti: quante forme d'onda si possono rappresentare? Cosa succede se il numero di componenti sono infinite?

È il momento di passare alla pratica:

<https://web.dm.unipi.it/robo1/teaching/matematica-musica-2026/>

Lo stesso fenomeno si riscontra, anche se in maniera più debole, con un rallentamento della convergenza per funzioni la cui derivata (oppure la derivata m -esima, per qualche m) non è continua.

Commenti finali sulla convergenza

Lo stesso fenomeno si riscontra, anche se in maniera più debole, con un rallentamento della convergenza per funzioni la cui derivata (oppure la derivata m -esima, per qualche m) non è continua.

Data una funzione con m derivate continue, le somme di Fourier convergeranno a questa funzione in modo polinomiale:

$$|s(t) - s_j(t)| \leq Cj^{-m},$$

Se invece la funzione ha un numero infinito di derivate continue, come ad esempio $s(t) = e^{\sin t + \cos t}$ che abbiamo considerato in precedenza, allora la convergenza sarà più veloce di ogni polinomio, e avremo

$$|s(t) - s_j(t)| \leq C\rho^j$$

per un'opportuna costante C ed un numero reale $0 < \rho < 1$.