FUNZIONI CONVESSE

G. DI MEGLIO

Indice

Introduzione

In questi fogli è raccolta, con numerose aggiunte, la presentazione della *convessità* come svolta a lezione.
L'idea è quella di presentare l'argomento partendo dalla "definizione analitica" di funzione convessa, data mediante la disuguaglianza di convessità, e dimostrare ogni altra caratterizzazione a partire da questa.
Particolare attenzione è posta nello studio delle proprietà dei rapporti incrementali e delle loro conseguenze a livello di derivabilità e di localizzazione dei punti di estremo assoluto di funzioni convesse o concave.
Nell'ultimo paragrafo sono illustrate alcune applicazioni "classiche" della convessità, cioè le dimostrazioni di disuguaglianze utili in diversi settori della Matematica pura ed applicata.

`Avvertenza`: Nei paragrafi e nelle osservazioni contrassegnati con * sono usate nozioni base del Calcolo Integrale.

Date: 18 luglio 2018.

1. Convessità e Concavità delle Funzioni Reali

1.1. **Funzioni Convesse e Concave.** La seguente definizione è fondamentale:

Definizione 1 (Definizione Analitica di Funzioni Convesse e Concave)
Siano I un intervallo non banale ed $f : I \to \mathbb{R}$.
La funzione f si dice *convessa in* I se e solo se l'implicazione:

$$(1) \qquad 0 < t < 1 \quad \Rightarrow \quad f\big((1-t)x_0 + tx_1\big) \leq (1-t)f(x_0) + tf(x_1)$$

vale per ogni scelta di $x_0 < x_1 \in I$.
Invece, la f si dice *concava in* I se e solo se la funzione $-f$ è convessa in I, cioè se l'implicazione:

$$(2) \qquad 0 < t < 1 \quad \Rightarrow \quad f\big((1-t)x_0 + tx_1\big) \geq (1-t)f(x_0) + tf(x_1)$$

vale per ogni scelta di $x_0 < x_1 \in I$.

Vale la pena fare alcune importanti osservazioni circa la Definizione 1.

Osservazione 1 (Significato Geometrico della Convessità)**:** Supponiamo che f soddisfi la (1).
Osserviamo innanzitutto che per ogni $t \in]0,1[$, il punto $x_t := (1-t)x_0 + tx_1$ appartiene all'intervallo $]x_0, x_1[$, poiché:

$$\begin{aligned} x_t &> (1-t)x_0 + tx_0 = x_0 \\ x_t &< (1-t)x_1 + tx_1 = x_1 \ ; \end{aligned}$$

viceversa, per ogni punto $x \in]x_0, x_1[$ è possibile determinare un unico $t \in]0,1[$ tale che $x_t = x$: infatti, risolvendo l'equazione $(1-t)x_0 + tx_1 = x$ rispetto a t otteniamo:

$$t = \frac{x - x_0}{x_1 - x_0}$$

ed è evidentissimo che tale t appartenga a $]0,1[$.
Da quanto appena detto segue che la (1) può riscriversi come segue:

$$x_0 < x < x_1 \quad \Rightarrow \quad f(x) \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0}\,(x - x_0) + f(x_0)\ ,$$

la quale ha un'immediata interpretazione geometrica: infatti, tale disuguaglianza asserisce che il generico punto $P = (x, f(x))$ del grafico di f ottenuto prendendo $x \in [x_0, x_1] \subseteq I$ non si trova mai "al disopra" del punto appartenente al segmento (di secante al grafico) congiungente i punti $P_0 = (x_0, f(x_0))$ e $P_1 = (x_1, f(x_1))$ che ha la stessa ascissa (vedi Figura 1).

Analogamente ragionando, se f è concava in I otteniamo che la (2) può riscriversi:

$$x_0 < x < x_1 \quad \Rightarrow \quad f(x) \geq \frac{f(x_1) - f(x_0)}{x_1 - x_0}\,(x - x_0) + f(x_0)\ ,$$

la quale ha un ovvia interpretazione geometrica, poiché asserisce che il generico punto $P = (x, f(x))$ del grafico di f ottenuto prendendo $x \in [x_0, x_1] \subseteq I$ non si trova mai "al disotto" del punto Q appartenente al segmento (di secante al grafico) congiungente i punti $P_0 = (x_0, f(x_0))$ e $P_1 = (x_1, f(x_1))$ che ha la stessa ascissa. ♦

Osservazione 2 (Sul Dominio delle Funzioni Convesse e Concave)**:** Quanto osservato sopra ci aiuta a mostrare che l'insieme di definizione X di una funzione convessa o concava f è necessariamente un intervallo.
Fissati due punti $x_0 < x_1$ in X, affinché sia possibile calcolare il valore di f in

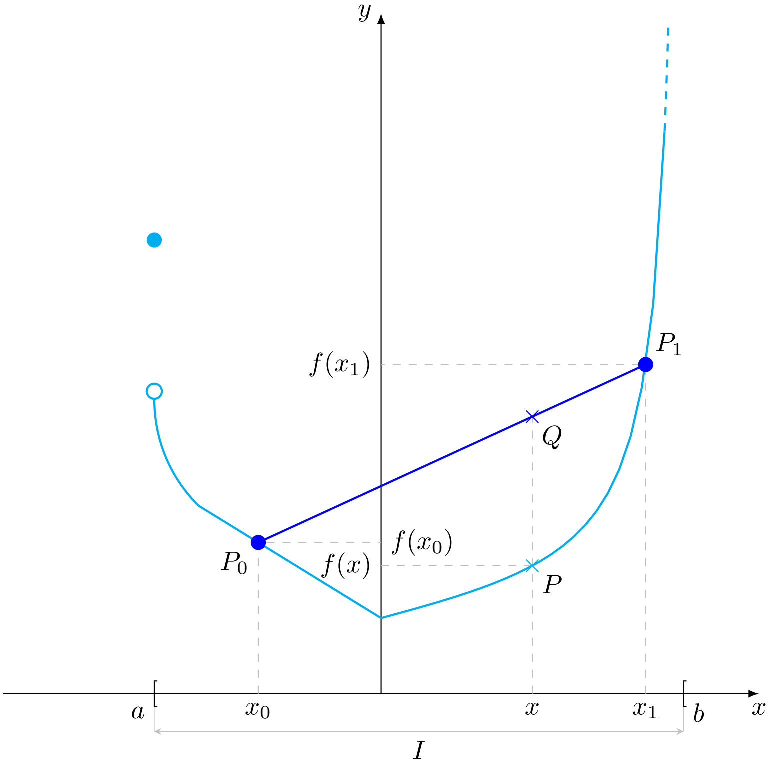

FIGURA 1. Interpretazione geometrica della DEFINIZIONE 1.

$x_t := (1-t)x_0 + tx_1$ per ogni $t \in]0,1[$ come indicato dalle (1) o (2) è necessario che risulti $x_t \in X$.
Poiché, come già osservato, i punti del tipo x_t descrivono al variare di $t \in]0,1[$ tutti i punti dell'intervallo $]x_0, x_1[$, da quanto detto sopra segue che $[x_0, x_1] \subseteq X$.
Dall'arbitrarietà nella scelta di $x_0 < x_1 \in X$ segue che il dominio X di f deve necessariamente soddisfare la *proprietà di connessione*:

$$\forall x_0 < x_1 \in X, \quad [x_0, x_1] \subseteq X \ ,$$

la quale individua tutti e soli gli intervalli della retta reale; pertanto X è un intervallo. ⧫

Osservazione 3 (Disuguaglianza di Convessità di Jensen[1])**:** Notiamo esplicitamente che le disuguaglianze (1) ed (2) possono essere estese ad un numero qualsiasi di punti di I usando il *Principio d'Induzione*.
Il lettore non avrà difficoltà a dimostrare che, in generale, una funzione $f : I \to \mathbb{R}$

[1]Johan Ludwig William Valdemar Jensen (1859 – 1925), matematico ed ingegnere danese.

è convessa [risp. concava] in I se e solo se:

$$0 < t_0, \ldots, t_n < 1,\ \sum_{k=0}^{n} t_k = 1 \quad \Rightarrow \tag{1'}$$

$$\Rightarrow \quad f\left(\sum_{k=0}^{n} t_k x_k\right) \leq \sum_{k=0}^{n} t_k f(x_k)$$

$$[\text{risp. } 0 < t_0, \ldots, t_n < 1,\ \sum_{k=0}^{n} t_k = 1 \quad \Rightarrow \tag{2'}$$

$$\Rightarrow \quad f\left(\sum_{k=0}^{n} t_k x_k\right) \geq \sum_{k=0}^{n} t_k f(x_k)]$$

vale per ogni scelta di $x_0 < \ldots < x_n \in I$ (con $n \geq 1$).
La disuguaglianza (2'), che generalizza la (1), è usualmente detta *disuguaglianza di convessità di Jensen.* $\diamondsuit$

Prima di enunciare alcune notevoli proprietà delle funzioni convesse o concave, riportiamo alcuni esempi.

`Esempio 1 (Funzioni Affini):` Scelti $m, q \in \mathbb{R}$, la funzione $f : \mathbb{R} \to \mathbb{R}$ definita ponendo $f(x) := mx + q$ è sia convessa sia concava in $\mathbb{R}$.
Infatti, abbiamo:

$$\begin{aligned} f\big((1-t)x_0 + tx_1\big) &= m\big((1-t)x_0 + tx_1\big) + q \\ &= m\big((1-t)x_0 + tx_1\big) + \big((1-t) + t\big)q \\ &= (1-t)(mx_0 + q) + t(mx_1 + q) \\ &= (1-t)f(x_0) + tf(x_1)\ . \end{aligned}$$

per ogni scelta di $x_0 < x_1 \in \mathbb{R}$ e di $t \in]0,1[$, cosicché sono contemporaneamente soddisfatte sia la (1) sia la (2). $\diamondsuit$

`Esempio 2 (Valore Assoluto):` La funzione $f : \mathbb{R} \to \mathbb{R}$ definita da $f(x) := |x|$ è convessa.
Infatti, per disuguaglianza triangolare abbiamo:

$$\begin{aligned} f\big((1-t)x_0 + tx_1\big) &= \big|(1-t)x_0 + tx_1\big| \\ &\leq \big|(1-t)x_0\big| - \big|tx_1\big| \\ &= (1-t)|x_0| + t - t|x_1| \\ &= (1-t)f(x_0) + tf(x_1) \end{aligned}$$

per ogni $x_0 < x_1 \in \mathbb{R}$ e $t \in]0,1[$, come volevamo. $\diamondsuit$

`Esempio 3:` La funzione $f : \mathbb{R} \to \mathbb{R}$ definita da $f(x) := 1 - |x|$ è concava.
Infatti, per la disuguaglianza triangolare abbiamo:

$$\begin{aligned} f\big((1-t)x_0 + tx_1\big) &= 1 - \underbrace{\big|(1-t)x_0 + tx_1\big|}_{\leq \left|(1-t)x_0\right| + \left|tx_1\right|} \\ &\geq 1 - \big|(1-t)x_0\big| - \big|tx_1\big| \\ &= (1-t) - (1-t)|x_0| + t - t|x_1| \\ &= (1-t)\big(1 - |x_0|\big) + t\big(1 - |x_1|\big) \\ &= (1-t)f(x_0) + tf(x_1) \end{aligned}$$

per ogni $x_0 < x_1 \in \mathbb{R}$ e $t \in]0,1[$, come volevamo. $\diamondsuit$

Osservazione 4 (Presenza di Tratti Rettilinei nei Grafici di Funzioni Convesse)**:** Notiamo esplicitamente che nei grafici delle funzioni proposte negli `Esempi` precedenti sono presenti dei tratti rettilinei.
Il fatto che il grafico di una funzione convessa o concava non sia sempre "ben curvo" è, in un certo senso, nella natura delle cose: infatti, non è difficile provare che se esistono $x_0 < x_1 \in I$ tali che nelle disuguaglianze (1) e (2) è soddisfatta l'uguaglianza per almeno un valore $\vartheta \in]0,1[$, ossia tali che:

$$f\big((1-\vartheta)x_0 + \vartheta x_1\big) = (1-\vartheta)f(x_0) + \vartheta f(x_1)\ ,$$

allora l'uguaglianza:

$$f\big((1-t)x_0 + tx_1\big) = (1-t)f(x_0) + tf(x_1)$$

è soddisfatta per ogni $t \in]0,1[$; ciò, geometricamente, equivale a dire che il tratto del grafico di f che si ottiene in corrispondenza dell'intervallo $[x_0, x_1]$ coincide con il segmento (di secante al grafico) di estremi P_0 e P_1. $\blacklozenge$

`Esempio 4:` La funzione $f : \mathbb{R} \to \mathbb{R}$ definita ponendo $f(x) := x^2$ è una funzione convessa.
Invero, scelti $x_0 < x_1 \in \mathbb{R}$, comunque si fissi $t \in]0,1[$ abbiamo:

$$\begin{aligned}
f\big((1-t)x_0 + tx_1\big) - (1-t)f(x_0) - tf(x_1)& \\
&= \big((1-t)x_0 + tx_1\big)^2 - (1-t)x_0^2 - tx_1^2 \\
&= \big((1-t)^2 - (1-t)\big)x_0^2 + 2t(1-t)x_0x_1 + \big(t^2 - t\big)x_1^2 \\
&= -t(1-t)x_0^2 + 2t(1-t)x_0x_1 - t(1-t)x_1^2 \\
&= -\underbrace{t(1-t)}_{>0}\ \underbrace{(x_0 - x_1)^2}_{>0} \\
&< 0
\end{aligned}$$

cosicché risulta addirittura $f\big((1-t)x_0 + tx_1\big) < (1-t)f(x_0) + tf(x_1)$. $\diamondsuit$

1.2. Funzioni Strettamente Convesse o Concave.

La definizione che segue individua due sottoclassi delle funzioni convesse e delle funzioni concave, le quali sono particolarmente interessanti dal punto di vista delle applicazioni:

Definizione 2 (Funzioni Strettamente Convesse o Concave)
Siano I un intervallo non banale ed $f : I \to \mathbb{R}$.
La f è detta *strettamente convessa in* I se e solo se la disuguaglianza:

$$(3) \qquad 0 < t < 1 \quad \Rightarrow \quad f\big((1-t)x_0 + tx_1\big) < (1-t)f(x_0) + tf(x_1)$$

vale per ogni $x_0 < x_1 \in I$.
Invece, la f è detta *strettamente convessa in* I se solo se la disuguaglianza:

$$(4) \qquad 0 < t < 1 \quad \Rightarrow \quad f\big((1-t)x_0 + tx_1\big) > (1-t)f(x_0) + tf(x_1)$$

vale per ogni $x_0 < x_1 \in I$.

Osservazione 5: L'unica differenza tra (1) e (3) e tra (2) e (4) è la disuguaglianza stretta. $\blacklozenge$

Osservazione 6 (Interpretazione Geometrica della Stretta Convessità)**:** Ragionando come nell'**Osservazione 1**, possiamo capire che anche la Definizione 2 ha un importante contenuto geometrico.

Invero, se f è strettamente convessa, il generico punto $P = (x, f(x))$ del grafico di f ottenuto prendendo $x \in]x_0, x_1[\subseteq I$ si trova sempre "al disotto" del punto appartenente al segmento (di secante al grafico) congiungente i punti $P_0 = (x_0, f(x_0))$ e $P_1 = (x_1, f(x_1))$ avente la stessa ascissa.

Analogamente, se f è strettamente concava, il generico punto $P = (x, f(x))$ del grafico di f ottenuto prendendo $x \in]x_0, x_1[\subseteq I$ si trova sempre "al disopra" del punto appartenente al segmento (di secante al grafico) congiungente i punti $P_0 = (x_0, f(x_0))$ e $P_1 = (x_1, f(x_1))$ avente la stessa ascissa. ♦

Osservazione 7 (Assenza di Tratti Rettilinei nel Grafico di una Funzione Strettamente Convessa)**:** Notiamo esplicitamente che la disuguaglianza (3) [risp. (4)] si può enunciare anche dicendo che vale la disuguaglianza di convessità (1) [risp. disuguaglianza di concavità (2)] e che non esistono né punti $x_0 < x_1$ né valori di $t \in]0, 1[$ tali che risulti $f\big((1-t)x_0 + tx_1\big) = (1-t)f(x_0) + tf(x_1)$.
Ciò implica che il grafico di una funzione strettamente convessa, così come quello di una funzione strettamente concava, non può contenere segmenti ossia che esso è sempre "ben curvo". ♦

`Esempio 5:` La funzione $f : \mathbb{R} \to \mathbb{R}$ definita ponendo $f(x) := x^2$ è strettamente convessa poiché, come visto nell'`Esempio 4`, essa soddisfa la (1) sempre col segno di disuguaglianza stretta. ♢

`Esempio 6:` La funzione $f :]0, +\infty[\to \mathbb{R}$ definita ponendo $f(x) := 1/x$ è strettamente convessa in $]0, +\infty[$.
Invero, scelti $0 < x_0 < x_1$ e $0 < t < 1$ abbiamo:

$$\begin{aligned} f\big((1-t)x_0 + tx_1\big) - (1-t)f(x_0) - tf(x_1) &= \frac{1}{(1-t)x_0 + tx_1} - \frac{1-t}{x_0} - \frac{t}{x_1} \\ &= \underbrace{(t^2 - t)}_{<0} \cdot \underbrace{\frac{(x_1 - x_0)^2}{\big((1-t)x_0 + tx_1\big)x_0 x_1}}_{>0} \\ &< 0 \end{aligned}$$

cosicché è soddisfatta la (3).
Analogamente si prova che la funzione $g :]-\infty, 0[\to \mathbb{R}$ definita da $g(x) := 1/x$ è strettamente concava in $]-\infty, 0[$. ♢

Altri esempi di funzioni strettamente convesse o strettamente concave li incontreremo più avanti.

Osservazione 8: La *disuguaglianza di convessità di Jensen* (1') [risp. l'analoga disuguaglianza di concavità (2')] vale col segno di disuguaglianza stretta per le funzioni strettamente convesse [risp. strettamente concave]. ♦

1.3. **Alcune Proprietà Algebriche delle Funzioni Convesse.** Concludiamo questo paragrafo iniziale con alcune proprietà di tipo algebrico, cioè riguardanti il comportamento delle funzioni convesse rispetto alle usuali operazioni.
Per semplicità, gli enunciati riguardano solamente le funzioni convesse e strettamente convesse; lasciamo al lettore il compito di convincersi che essi sussistono, con le modifiche del caso, anche per le funzioni concave.

Proposizione 1

Siano I un intervallo non banale ed $f, g : I \to \mathbb{R}$.

i. *Se f e g sono entrambe convesse in I, allora $\alpha f + \beta g$ è convessa in I per ogni $\alpha, \beta \geq 0$.*

ii. *Se f è strettamente convessa e g è convessa in I, allora $\alpha f + \beta g$ è strettamente convessa in I per ogni $\alpha > 0$ e $\beta \geq 0$.*

iii. *Se f è convessa [risp. strettamente convessa] in I, allora αf è concava [risp. strettamente concava] in I per ogni $\alpha < 0$.*

Dimostrazione. Le *i* - *iii* sono immediate conseguenze delle definizioni e delle proprietà della relazione d'ordine. □

PROPOSIZIONE 2
Siano I, J intervalli non banali, $f : I \to J$ e $g : J \to \mathbb{R}$.

i. *Se f è convessa in I e g è crescente e convessa [risp. strettamente convessa] in J, allora la funzione composta $g \circ f$ è convessa [risp. strettamente convessa] in I.*

ii. *Se f è convessa in I e g è decrescente e concava [risp. strettamente concava] in J, allora la funzione composta $g \circ f$ è concava in I.*

iii. *Se f è strettamente convessa in I e g è strettamente crescente e convessa in J, allora la funzione composta $g \circ f$ è strettamente convessa in I.*

iv. *Se f è strettamente convessa in I e g è strettamente decrescente e concava in J, allora la funzione composta $g \circ f$ è concava in I.*

Dimostrazione. Proviamo la *iii*, analogamente ragionandosi negli altri casi.
Fissati che siano $x_0 < x_1 \in I$ e $t \in]0,1[$ abbiamo:

$$\begin{aligned} f(x_t) < (1-t)f(x_0) + tf(x_1) \quad \Rightarrow \quad g\big(f(x_t)\big) &< g\big((1-t)f(x_0) + tf(x_1)\big) \\ &\leq (1-t)g\big(f(x_0)\big) + tg\big(f(x_1)\big) \end{aligned}$$

da cui segue la stretta convessità di $g \circ f$. □

PROPOSIZIONE 3
Siano I un intervallo non banale ed $f : I \to \mathbb{R}$ una funzione invertibile con inversa $f^{-1} : f(I) \to \mathbb{R}$.

i. *Se f è crescente e convessa [risp. strettamente convessa] in I, allora f^{-1} è crescente e concava [risp. strettamente concava] in $f(I)$.*

ii. *Se f è decrescente e convessa [risp. strettamente convessa] in I, allora f^{-1} è decrescente e convessa [risp. strettamente convessa] in $f(I)$.*

Dimostrazione. Dimostriamo la *ii*, analogamente ragionandosi per la *i*.
Supponiamo che f sia invertibile, decrescente e convessa; in tal caso, per noti fatti, f è strettamente decrescente e così pure f^{-1} e perciò dobbiamo provare solo che f^{-1} è convessa.
A tal uopo, fissiamo $y_0 < y_1 \in f(I)$ ed osserviamo che esistono due punti $x_0, x_1 \in I$

tali che $f(x_0) = y_0$ ed $f(x_1) = y_1$; inoltre, data la stretta monotònia di f, si ha $x_0 > x_1$. Scelto ora un $t \in]0,1[$ e posto $\tau = 1 - t \in]0,1[$ abbiamo:

$$\begin{aligned} f\big((1-\tau)x_1 + \tau x_0\big) \leq (1-\tau)f(x_1) + \tau f(x_0) \quad &\Rightarrow \\ &\Rightarrow \quad (1-\tau)x_1 + \tau x_0 \geq f^{-1}\Big((1-\tau)f(x_1) + \tau f(x_0)\Big) \\ &\Rightarrow \quad (1-\tau)f^{-1}(y_1) + \tau f^{-1}(y_0) \geq f^{-1}\Big((1-\tau)y_1 + \tau y_0\Big) \\ &\Rightarrow \quad f^{-1}\Big((1-t)y_0 + t y_1\Big) \leq (1-t)f^{-1}(y_0) + t f^{-1}(y_1) , \end{aligned}$$

come volevamo.
Osserviamo esplicitamente che se f è strettamente convessa, la prima delle disuguaglianze precedenti è stretta e tale rimane (seppure con verso opposto) anche applicando ad ambo i suoi membri la funzione strettamente decrescente f^{-1}; da ciò segue che f^{-1} è strettamente convessa. □

2. Rapporti Incrementali di Funzioni Convesse

Da qui in avanti, per pura semplicità, il discorso verterà sulle funzioni convesse, strettamente e no.
Avvisiamo il lettore che quanto diremo vale anche per le funzioni concave, a patto di invertire il verso di qualche disuguaglianza.

Come notato nell'**Osservazione 1**, una funzione $f : I \to \mathbb{R}$ è convessa se e solo se risulta:

$$x_0 < x < x_1 \quad \Rightarrow \quad f(x) \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0}\,(x - x_0) + f(x_0)$$

e ciò evidentemente importa:

$$x_0 < x < x_1 \quad \Rightarrow \quad \frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0}\ ;$$

dato che la prima implicazione può scriversi pure come:

$$x_0 < x < x_1 \quad \Rightarrow \quad f(x) \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0}\,(x - x_1) + f(x_1)$$

(poiché il secondo membro è una riscrittura del secondo membro dell'equazione della retta secante il grafico di f in P_0 e P_1) è chiaro che da essa segue anche:

$$x_0 < x < x_1 \quad \Rightarrow \quad \frac{f(x) - f(x_1)}{x - x_1} \geq \frac{f(x_1) - f(x_0)}{x_1 - x_0}\ .$$

Altrettanto evidente è il fatto che da ognuna delle implicazioni:

$$\begin{aligned} x_0 < x < x_1 \quad &\Rightarrow \quad \frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0} \\ x_0 < x < x_1 \quad &\Rightarrow \quad \frac{f(x) - f(x_1)}{x - x_1} \geq \frac{f(x_1) - f(x_0)}{x_1 - x_0} \end{aligned}$$

si può ricavare la disuguaglianza di convessità (1).
Abbiamo così dimostrato un'importante caratterizzazione della convessità, espressa di solito nel seguente:

Lemma 1 (Convessità e Rapporti Incrementali)
Siano I un intervallo non banale ed $f : I \to \mathbb{R}$.
La f è convessa in I se e solo se per ogni terna di punti $x_0 < x < x_1 \in I$ risulta:

$$\frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x) - f(x_1)}{x - x_1}\ . \tag{5}$$

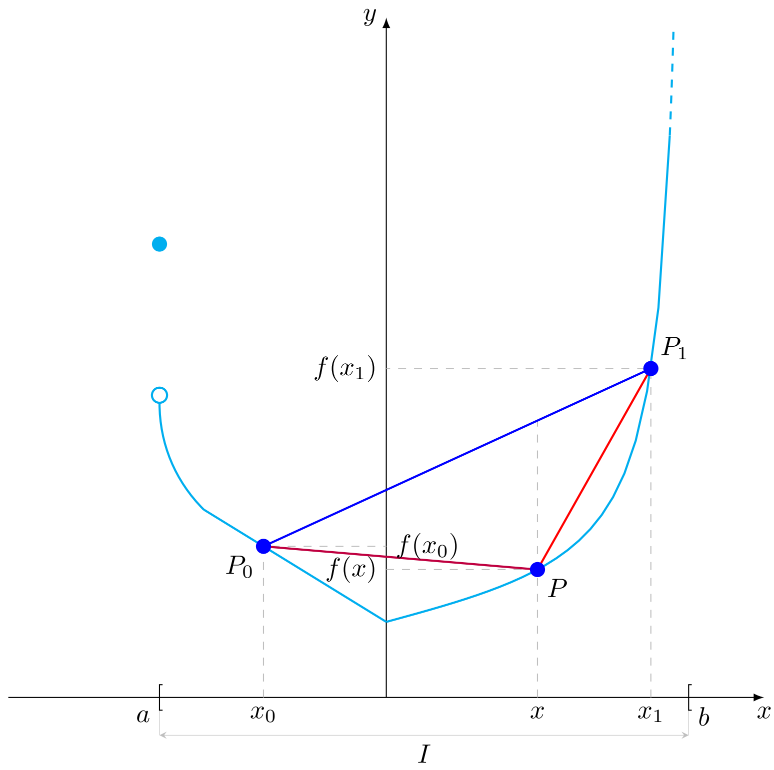

FIGURA 2. Interpretazione geometrica del LEMMA 1.

Osservazione 9 (Interpretazione Geometrica del Lemma)**:** Il LEMMA 1 asserisce una semplice proprietà geometrica del grafico di una funzione convessa, cioè il fatto che, comunque si fissino tre punti $P_0 = (x_0, f(x_0))$, $P = (x, f(x))$ e $P_1 = (x_1, f(x_1))$ con ascisse $x_0 < x < x_1$, le retta secante il grafico in P_0 e P ha pendenza non superiore di quella condotta per P_0 e P_1, la quale a sua volta ha pendenza non superiore a quella passante per P e P_1 (cfr. FIGURA 2). ♦

Osservazione 10: Ribadiamo una cosa già evidenziata nella dimostrazione del LEMMA, ossia che per dimostrare la convessità di f basta la sola disuguaglianza tra i membri esterni di (5), cioè basta che:

$$\frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x) - f(x_1)}{x - x_1}$$

per ogni terna di punti $x_0 < x < x_1 \in I$. ♦

Osservazione 11: Le (5) valgono col segno di disuguaglianza stretta se e solo se la f è strettamente convessa in I. ♦

Il LEMMA 1 ha numerose semplici conseguenze le quali consentono di stabilire utilissime proprietà delle funzioni convesse: qui di seguito ne proponiamo alcune molto importanti.
Abbiamo:

PROPOSIZIONE 4 (Monotònia del Rapporto Incrementale)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale ed $f : I \to \mathbb{R}$.

La f è convessa in I se e solo se, per ogni $x_0 \in I$, la funzione rapporto incrementale centrato in x_0, cioè $\mathrm{r}(x;x_0) := \frac{f(x)-f(x_0)}{x-x_0}$, è crescente in $I-\{x_0\}$.

Dimostrazione. $\Rightarrow$) *Se f è convessa, allora $\mathrm{r}(x;x_0)$ è crescente per ogni $x_0 \in I$.*
Fissiamo $x_1 < x_2 \in I - \{x_0\}$ e distinguiamo i casi:

se $x_0 < x_1 < x_2$: sfruttando la prima disuguaglianza delle (5) otteniamo:

$$\mathrm{r}(x_1;x_0) = \frac{f(x_1)-f(x_0)}{x_1-x_0} \leq \frac{f(x_2)-f(x_0)}{x_2-x_0} = \mathrm{r}(x_2;x_0)\ ;$$

se $x_1 < x_0 < x_2$: sfruttando entrambe le disuguaglianze (5) troviamo:

$$\mathrm{r}(x_1;x_0) = \frac{f(x_0)-f(x_1)}{x_0-x_1} \leq \frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_0)-f(x_2)}{x_0-x_2} = \mathrm{r}(x_2;x_0)\ ;$$

se $x_1 < x_2 x_0$: sfruttando la seconda disuguaglianza delle (5) otteniamo:

$$\mathrm{r}(x_1;x_0) = \frac{f(x_1)-f(x_0)}{x_1-x_0} \leq \frac{f(x_2)-f(x_0)}{x_2-x_0} = \mathrm{r}(x_2;x_0)\ ;$$

pertanto si ha sempre $\mathrm{r}(x_1;x_0) \leq \mathrm{r}(x_2;x_0)$.

$\Leftarrow$) *Se $\mathrm{r}(x;x_0)$ è crescente per ogni $x_0 \in I$, allora f è convessa.*
Scelti $x_0 < x < x_1 \in I$ per monotònia di $\mathrm{r}(\cdot;x_0)$ abbiamo:

$$\mathrm{r}(x;x_0) \leq \mathrm{r}(x_1;x_0)\ ,$$

mentre per monotònia di $\mathrm{r}(\cdot;x_1)$ abbiamo:

$$\mathrm{r}(x_1;x_0) = \mathrm{r}(x_0;x_1) \leq \mathrm{r}(x;x_1)\ ;$$

dunque è:

$$\frac{f(x)-f(x_0)}{x-x_0} \leq \frac{f(x_1)-f(x_0)}{x_1-x_0} \leq \frac{f(x)-f(x_1)}{x-x_1}$$

e per il Lemma 1 f è convessa. □

Osservazione 12: Se f è strettamente convessa, ogni rapporto incrementale $\mathrm{r}(\cdot;x_0)$ è strettamente crescente in $I - \{x_0\}$. ♦

Proposizione 5 (Esistenza delle Derivate Destra e Sinistra nei Punti Interni)
Siano $I \subseteq \mathbb{R}$ intervallo non banale ed $f : I \to \mathbb{R}$.
Se f è convessa in I, i limiti:

$$\lim_{x\to x_0^\pm} \frac{f(x)-f(x_0)}{x-x_0}$$

esistono per ogni scelta di $x_0 \in I$.[2]
In particolare, se x_0 è interno ad I, entrambi i limiti precedenti esistono finiti, nel senso che esistono due numeri $l_- \leq l_+ \in \mathbb{R}$ tali che:

$$\lim_{x\to x_0^-} \frac{f(x)-f(x_0)}{x-x_0} = l_- \qquad\qquad \lim_{x\to x_0^+} \frac{f(x)-f(x_0)}{x-x_0} = l_+\ ,$$

cosicché f è derivabile da sinistra e da destra in ogni punto interno al suo intervallo di definizione e $f'_\pm(x_0) = l_\pm$.

[2]Chiaramente, se x_0 è un estremo di I esiste l'unico tra i due limiti delle restrizioni che ha senso calcolare.

Dimostrazione. Visto che $\mathrm{r}(x;x_0)$ è crescente in $I-\{x_0\}$, per il *Teorema sulla Regolarità delle Funzioni Monotòne* troviamo:

$$\lim_{x\to x_0^-}\frac{f(x)-f(x_0)}{x-x_0}=$$
$$=\sup_{x<x_0}\mathrm{r}(x;x_0)=\begin{cases} l_-\in\mathbb{R} & \text{, se } \mathrm{r}(x;x_0) \text{ ha qualche maggiorante per} \\ & x<x_0 \text{ intorno a } x_0 \\ +\infty & \text{, altrimenti}\end{cases}$$
$$\lim_{x\to x_0^+}\frac{f(x)-f(x_0)}{x-x_0}=$$
$$=\inf_{x>x_0}\mathrm{r}(x;x_0)=\begin{cases} l_+\in\mathbb{R} & \text{, se } \mathrm{r}(x;x_0) \text{ ha qualche minorante per} \\ & x>x_0 \text{ intorno a } x_0 \\ -\infty & \text{, altrimenti}\end{cases},$$

cosicché i due limiti esistono sempre.
D'altra parte, se x_0 è interno ad I esistono certamente due punti $x_1,x_2\in I$ tali che $x_1<x_0<x_2$; per il Lemma 1, comunque si scelga $x<x_0$ abbiamo:

$$\mathrm{r}(x;x_0)\le\mathrm{r}(x_0;x_2)\ ,$$

di modo che $\mathrm{r}(x_0;x_2)$ è un maggiorante dei rapporti incrementali $\mathrm{r}(x;x_0)$ per $x<x_0$ e dunque:

$$\lim_{x\to x_0^-}\frac{f(x)-f(x_0)}{x-x_0}=l_-\in\mathbb{R}\ ;$$

del tutto analogamente, fissato $x_0<x$ dal Lemma 1 traiamo:

$$\mathrm{r}(x_1;x_0)\le\mathrm{r}(x;x_0)\ ,$$

dunque $\mathrm{r}(x_1;x_0)$ è un minorante dei rapporti incrementali $\mathrm{r}(x;x_0)$ per $x>x_0$ e dunque:

$$\lim_{x\to x_0^+}\frac{f(x)-f(x_0)}{x-x_0}=l_+\in\mathbb{R}\ .$$

Infine, sempre dal Lemma 1 si desume che gli insiemi:

$$L=\Big\{\mathrm{r}(x;x_0),\quad \text{con } x<x_0\Big\}$$
$$U=\Big\{\mathrm{r}(x;x_0),\quad \text{con } x>x_0\Big\}$$

sono separati e che U è l'insieme dei maggioranti; pertanto $l_-=\sup L\le\inf U=l_+$. □

Da quanto appena acquisito consegue immediatamente il:

Teorema 1 (Continuità nei Punti Interni)
Siano $I\subseteq\mathbb{R}$ non banale ed $f:I\to\mathbb{R}$.
Se f è convessa in I, allora f è continua in ogni punto x_0 interno ad I.

Dimostrazione. La dimostrazione ricalca quella fornita per provare che una funzione derivabile in un punto è ivi continua.
Fissiamo x_0 interno ad I. Visto che x_0 è di accumulazione sia da sinistra sia da destra per I, provare che f è continua in x_0 equivale a provare che:

$$\lim_{x\to x_0^-}f(x)=f(x_0)=\lim_{x\to x_0^+}f(x)\ .$$

Per quanto asserito nella PROPOSIZIONE 5, abbiamo:

$$\begin{aligned}\lim_{x\to x_0^\pm} f(x) &= \lim_{x\to x_0^\pm} \frac{f(x)-f(x_0)}{x-x_0}\cdot(x-x_0)+f(x_0)\\ &= l_\pm \cdot 0 + f(x_0)\\ &= f(x_0)\ .\end{aligned}$$

□

Osservazione 13 (Comportamento nei Punti Interni)**:** Geometricamente, la PROPOSIZIONE 5 ed il TEOREMA 1 assicurano che il grafico di una funzione convessa f presenta unicamente punti angolosi con ascisse interne all'intervallo di definizione e che le eventuali discontinuità di f (ovviamente di prima specie) possono essere localizzate unicamente negli estremi dell'intervallo di definizione. ♦

Osservazione 14 (Comportamento negli Estremi – I)**:** A norma della PROPOSIZIONE 5, gli unici punti in cui i limiti delle restrizioni $\lim_{x\to x_0^\pm} \frac{f(x)-f(x_0)}{x-x_0}$ possono uguagliare $\pm\infty$ sono gli estremi dell'intervallo I.
Detti $a < b$ tali estremi, osserviamo esplicitamente che affinché ciò avvenga c'è bisogno che $a \neq -\infty$ [risp. $b \neq +\infty$] e che f sia definita in a [risp. b], altrimenti i rapporti incrementali $\mathrm{r}(x;a)$ [risp. $\mathrm{r}(x;b)$] non possono essere calcolati.
In tal caso, il $\lim_{x\to a^+} \frac{f(x)-f(a)}{x-a}$ può essere uguale a $-\infty$ per due motivi:

f **non è continua in** a**:** in tal caso da:

$$\lim_{x\to a^+} \frac{f(x)-f(a)}{x-a} = -\infty$$

segue (per *permanenza del segno*) che $f(x) < f(a)$ intorno ad a ed inoltre si può dimostrare che deve necessariamente risultare $\lim_{x\to a^+} f(x) = l \in \mathbb{R}$, di modo che a è un punti di *discontinuità eliminabile*;

f **è continua in** a**:** allora il grafico di f ha tangente verticale in $(a, f(a))$.

Lo stesso tipo di distinzione si può fare per illustrare il comportamento di f nell'estremo superiore b, qualora f sia definita in b. ♦

Osservazione 15 (Comportamento negli Estremi – II)**:** Quando f non è definita in a ed $a \in \mathbb{R}$, si può comunque dimostrare che il limite $\lim_{x\to a^+} f(x)$ esiste *sempre* ed è un valore $l \in \mathbb{R} \cup \{+\infty\}$. Pertanto, quando f non è definita nell'estremo inferiore a di I, o il grafico di f presenta un asintoto verticale da destra verso l'alto d'equazione $x = a$ oppure f è prolungabile con continuità su a ponendo $f(a) = l$ (a seconda che $l = +\infty$ oppure $l \in \mathbb{R}$).
Infine, se $a = -\infty$ allora il limite $\lim_{x\to -\infty} f(x)$ esiste *sempre*, ma può prendere ogni valore appartenente ad $\widehat{\mathbb{R}}$. In tal caso, il grafico di f può avere un asintoto orizzontale a sinistra o meno (a seconda che $\lim_{x\to -\infty} f(x)$ sia finito o no), può avere un asintoto obliquo o no.
Lo stesso tipo di considerazioni si può fare per l'estremo b quando f non vi è definita e le casistiche che si ritrovano sono le medesime del caso precedente. ♦

Dal TEOREMA 1 e dalla PROPOSIZIONE 5 traiamo immediatamente il seguente risultato che illustra un proprietà geometrica interessante dei grafici delle funzioni convesse:

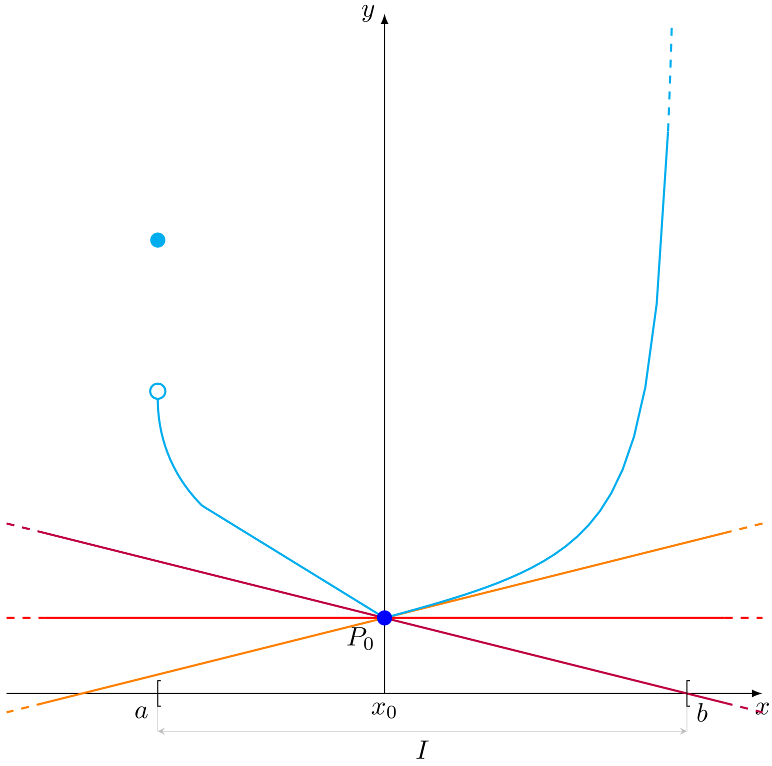

FIGURA 3. Interpretazione geometrica della PROPOSIZIONE 6.

PROPOSIZIONE 6 (Esistenza di Rette di Supporto al Grafico)
Siano $I = (a, b) \subseteq \mathbb{R}$ intervallo non banale, $f : I \to \mathbb{R}$ convessa in I ed $x_0 \in I$.
Se x_0 è un punto interno ad I, allora per ogni $m \in [f'_-(x_0), f'_+(x_0)]$ risulta:

$$\forall x \in I, \quad f(x) \geq m(x - x_0) + f(x_0) \ . \tag{6}$$

Se $x_0 = a$ [risp. $x_0 = b$] e se esiste (finita) $f'_+(a)$ [risp. $f'_-(b)$], allora la (6) vale per ogni $m \in]-\infty, f'_+(a)]$ [risp. $m \in [f'_-(b), +\infty[$].

Osservazione 16 (Interpretazione Geometrica della Proposizione 6)**:** La PROPOSIZIONE 6 asserisce che, fissato x_0 interno ad I, il grafico della funzione f si trova tutto nel semipiano "al disopra" della retta di equazione $y = m(x - x_0) + f(x_0)$, condotta per $P_0 = (x_0, f(x_0))$ ed avente per coefficiente angolare un qualsiasi $m \in [f'_-(x_0), f'_+(x_0)]$.
Analogamente, se è possibile prendere $x_0 = a$ [risp. $x_0 = b$] e se esiste la derivata $f'_+(a)$ [risp. $f'_-(x_0)$], allora il grafico di f si trova tutto nel semipiano "al disopra" della retta di equazione $y = m(x-a)+f(a)$ [risp. $y = m(x-b)+f(b)$], passante per $A = (a, f(a))$ [risp. $B = (b, f(b))$] ed avente come coefficiente angolare un qualsiasi $m \leq f'_+(a)$ [risp. $m \geq f'_-(b)$].
Geometricamente, accade ciò che è illustrato in FIGURA 3.
Le rette del tipo ora detto si chiamano usualmente *rette di supporto* (o *rette radenti*) *al grafico di f in P_0*. ♦

Osservazione 17 (Sottodifferenziale di una Funzione Convessa)**:** Fissato un punto x_0 nell'intervallo di definizione I, l'insieme degli $m \in \mathbb{R}$ per cui vale la (6) si chiama

sottodifferenziale di f in x_0 e si denota col simbolo $\partial f(x_0)$.
La Proposizione 6 assicura che:

- $\partial f(x_0) = [f'_-(x_0), f'_+(x_0)]$, se x_0 è interno all'intervallo I,
- $\partial f(a) =]-\infty, f'_+(a)]$, se $a \in I$ ed f è derivabile in a da destra,
- $\partial f(a) = [f'_-(b), +\infty[$, se $b \in I$ ed f è derivabile in b da sinistra. ♦

Osservazione 18: Evidentemente, se f è derivabile nel punto interno x_0, l'insieme $\partial f(x_0) = [f'_-(x_0), f'_+(x_0)]$ si riduce ad un solo punto, cioè risulta $\partial f(x_0) = \{f'(x_0)\}$; in tal caso, la famiglia delle rette di supporto si riduce ad un'unica retta, quella di equazione $y = f'(x_0)(x-x_0)+f(x_0)$ ossia la tangente al grafico in $P_0 = (x_0, f(x_0))$. Geometricamente, quando è possibile tracciare la retta tangente al grafico di una funzione convessa, il grafico stesso si trova sempre "al disopra" di tale retta. ♦

Dimostrazione. Facciamo la dimostrazione per x_0 interno, ragionandosi analogamente negli altri due casi.
Dato che:

$$m \leq f'_+(x_0) = \inf_{x>x_0} \frac{f(x) - f(x_0)}{x - x_0} \ ,$$

per ogni $x > x_0$ abbiamo:

$$f(x) \geq m(x - x_0) + f(x_0) \ ;$$

analogamente, visto che:

$$m \geq f'_-(x_0) = \inf_{x<x_0} \frac{f(x) - f(x_0)}{x - x_0} \ ,$$

per ogni $x < x_0$ abbiamo:

$$f(x) \geq m(x - x_0) + f(x_0) \ .$$

Quindi, la(6) vale per per $x \neq x_0$, ma ovviamente vale anche per $x = x_0$ e perciò vale in tutto I. □

Osservazione 19 (Rette di Supporto e Punti di Contatto)**:** Notiamo esplicitamente che, in generale, una retta di supporto al grafico in $P_0 = (x_0, f(x_0))$ può contenere tutto un tratto di grafico oltre al punto P_0: ciò accade, essenzialmente, quando f è convessa senza esserlo strettamente.
Infatti, come già osservato in precedenza, quando f è strettamente convessa il grafico non può contenere tre punti allineati, dunque nessun suo tratto può giacere su una retta. ♦

3. Derivate e Funzioni Convesse

Abbiamo visto che una funzione convessa non è necessariamente derivabile in ogni punto del suo intervallo di definizione (cfr. `Esempio` 2); tuttavia, si può dimostrare (cfr. Appendice A) che una funzione convessa è derivabile in "quasi ogni punto"[3] del proprio intervallo di definizione.
Pertanto ha qualche importanza lo studio delle derivate delle funzioni convesse che proponiamo in questa sezione.

[3]Questa dicitura viene resa precisa nell'ambito della *Teoria della Misura*; qui ci limitiamo a segnalare che il suo significato in questo contesto è quello fornito nell'enunciato del Teorema 9 in Appendice A.

3.1. **Convessità e Derivata Prima.** Il primo risultato utile in tal senso è il seguente:

TEOREMA 2 (Monotònia della Derivata Prima)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale ed $f : I \to \mathbb{R}$ una funzione derivabile internamente ad I.
La f è convessa in I se e solo se f' è una funzione crescente in I.

Dimostrazione. $\Rightarrow$) *Se f è convessa, allora f' è crescente.*
Scegliamo $x_0 < x_1$ interni ad I. Scelto un qualsiasi $x \in]x_0, x_1[$, per il LEMMA 1 e le PROPOSIZIONI 4 & 5 troviamo:

$$f'_+(x_0) \leq \frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x) - f(x_1)}{x - x_1} \leq f'_-(x_1)$$

e, visto che $f'_+(x_0) = f'(x_0)$ ed $f'_-(x_1) = f'(x_1)$, da ciò segue $f'(x_0) \leq f'(x_1)$, come volevamo.

$\Leftarrow$) *Se f' è crescente, allora f è convessa.*
A norma della LEMMA 1 e dell'**Osservazione 10** basta provare che, per ogni scelta di $x_0 < x <_1 \in I$, risulta $\mathrm{r}(x; x_0) \leq \mathrm{r}(x; x_1)$.
Fissiamo $x_0 < x < x_1 \in I$; per il *Teorema di Lagrange* esistono due punti $\xi \in]x_0, x[$ ed $\eta \in]x, x_1[$ tali che:

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi) \leq f'(\eta) = \frac{f(x_1) - f(x)}{x_1 - x} = \frac{f(x) - f(x_1)}{x - x_1} \ ,$$

come volevamo. □

Osservazione 20: Ovviamente, una funzione derivabile f è strettamente convessa in I se e solo se la funzione f' è strettamente crescente in I. ♦

Quanto finora detto ci consente di determinare facilmente la concavità o convessità di alcune funzioni elementari "di base".

`Esempio 7 (Potenze ad Esponente Reale):` Fissato $\alpha \in \mathbb{R}$ diverso da 0, la funzione potenza $f(x) := x^\alpha$ è derivabile in $]0, +\infty[$ ed ha derivata:

$$f'(x) = \alpha x^{\alpha - 1} \ .$$

Se $\alpha > 1$, f' è strettamente crescente ed f è strettamente convessa; se $0 < \alpha < 1$, f' è strettamente decrescente ed f è strettamente concava; se $\alpha < 0$, f' è strettamente crescente e dunque f è strettamente convessa; infine, se $\alpha = 1$, la f è convessa ma non strettamente. ♢

Osservazione 21 (Potenze ad Esponente Naturale ed Intero)**:** Lo stesso tipo di considerazioni ora svolte si può fare per studiare la concavità delle funzioni potenza ad esponente naturale ed intero.
In particolare, si trova che $f(x) := x^n$ (con $n \in \mathbb{N}$ ed $n \geq 2$) è strettamente convessa in $\mathbb{R}$ se n è pari, mentre è strettamente convessa in $[0, +\infty[$ e strettamente concava in $]-\infty, 0]$ se n è dispari.
Analogamente, $f(x) := x^{-n} = \frac{1}{x^n}$ (con $n \in \mathbb{N}$ ed $n \geq 1$) è strettamente convessa in $]-\infty, 0[$ ed in $]0, +\infty[$ se n è pari ed è strettamente convessa in $]0, +\infty[$ e strettamente concava in $]-\infty, 0[$ se n è dispari. ♦

`Esempio 8 (L'Esponenziale è Strettamente Convesso):` Fissato che sia $a > 0$ e $\neq 1$, la funzione esponenziale $f(x) := a^x$ è derivabile in tutto $\mathbb{R}$ ed ha derivata:

$$f'(x) = \log a \cdot a^x = \log a \cdot f(x) \ .$$

Se $a > 1$, $\log a > 0$ ed f è strettamente crescente, dunque f' è strettamente crescente; invece, se $0 < a < 1$, $\log a < 0$ ed f è strettamente decrescente, dunque f' è strettamente crescente.
Ne consegue che la funzione esponenziale è sempre strettamente convessa. $\diamondsuit$

`Esempio 9 (Logaritmi):` Fissato $a > 0$ e $\neq 1$, la funzione $f(x) := \log_a x$ è derivabile in $]0, +\infty[$ ed ha derivata:

$$f'(x) = \frac{1}{x \cdot \log a} \ .$$

Se $a > 1$ risulta $\log a > 0$, dunque f' è strettamente decrescente ed f è strettamente concava; mentre se $0 < a < 1$ risulta $\log a < 0$, dunque f' è strettamente crescente ed f è strettamente convessa.
Conseguentemente, il logaritmo è strettamente concavo se ha base > 1 ed è strettamente convesso se ha base > 1. $\diamondsuit$

`Esempio 10 (Coseno e Seno):` La funzione $f(x) := \cos x$ è derivabile in $\mathbb{R}$ ed ha $f'(x) = -\sin x$; pertanto essa risulta strettamente convessa in ogni intervallo in cui il seno è funzione decrescente, i.e. $[\pi/2 + 2k\pi, 3/2\pi + 2k\pi]$ ($k \in \mathbb{Z}$), e strettamente concava in ogni intervallo in cui il seno è funzione crescente, cioè $[-\pi/2 + 2k\pi, \pi/2 + 2k\pi]$ ($k \in \mathbb{Z}$).
Analogamente, $f(x) := \sin x$ è derivabile in $\mathbb{R}$ con $f'(x) = \cos x$; dunque essa è strettamente convessa in ogni intervallo $[-\pi + 2k\pi, 2k\pi]$ ($k \in \mathbb{Z}$), in cui il coseno è strettamente crescente, ed è strettamente concava in ogni intervallo $[2k\pi, \pi + 2k\pi]$ ($k \in \mathbb{Z}$), in cui il coseno è strettamente decrescente. $\diamondsuit$

Come già osservato in precedenza (cfr. **Osservazione 18**), se f è una funzione convessa in I, allora in ogni punto P_0 del grafico di f con ascissa x_0 interna all'intervallo di definizione I esiste un'unica retta di supporto al grafico, che coincide con la retta tangente in P_0. Abbastanza sorprendentemente, questa proprietà si inverte e perciò sussiste il seguente:

Teorema 3
Siano I un intervallo non banale ed $f : I \to \mathbb{R}$ derivabile internamente ad I.
La f è convessa in I se e solo se, comunque si scelga un punto x_0 interno ad I, risulta:

$$\forall x \in I, \quad f(x) \geq f'(x_0) \cdot (x - x_0) + f(x_0) \ . \tag{7}$$

Dimostrazione. Visto che l'implicazione $\Rightarrow$ è già stata acquisita nell'**Osservazione 18**, dimostriamo unicamente l'implicazione inversa.
Supponiamo che la (7) valga per ogni punto interno ad I e scegliamo due punti $x_0 < x_1$ interni ad I.
Abbiamo:

$$\begin{aligned} f(x_1) &\geq f'(x_0)(x_1 - x_0) + f(x_0) \\ f(x_0) &\geq f'(x_1)(x_0 - x_1) + f(x_1) \ , \end{aligned}$$

da cui:

$$f(x_1) + f(x_0) \geq f'(x_0)(x_1 - x_0) + f'(x_1)(x_0 - x_1) + f(x_0) + f(x_1)$$

cioè:

$$\big(f'(x_0) - f'(x_1)\big)(x_1 - x_0) \leq 0\ ;$$

dato che $x_0 < x_1$, la precedente implica $f'(x_0) - f'(x_1) \leq 0$, ossia $f'(x_0) \leq f'(x_1)$. Per l'arbitrarietà nella scelta di x_0 ed x_1 la funzione f' è crescente nell'interno di I, dunque f è convessa per il TEOREMA 2. □

Osservazione 22: Non è difficile intuire che se nella (7) l'uguaglianza vale solo per $x = x_0$, allora f è strettamente convessa in I.
Infatti, lo stesso ragionamento usato nella dimostrazione mostra che in tal caso si ha $f'(x_0) < f'(x_1)$ quando $x_0 < x_1$, sicché f' è strettamente crescente in I. ♦

Osservazione 23: Limitatamente alle funzioni derivabili, la disuguaglianza (7) è del tutto equivalente alla definizione di convessità.
Per questo motivo, nei manuali in cui viene posta maggiore attenzione alle proprietà delle funzioni derivabili, la (7) viene usata al posto della (1) per *definire* la nozione di funzione convessa (cfr. [MS]). ♦

3.2. **Convessità e Derivata Seconda.** Dal TEOREMA 2 e dai *Criteri di Monotònia* e *di Stretta Monotònia* traiamo in maniera diretta il seguente:

TEOREMA 4 (Convessità e Segno della Derivata Seconda)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale, $f : I \to \mathbb{R}$ derivabile due volte internamente ad I.
La f è convessa in I se e solo se $f''(x) \geq 0$ per ogni x interno ad I.
La f è strettamente convessa in I se e solo se $f''(x) \geq 0$ per ogni x interno ad I e se f'' non si annulla su alcun intervallo $[a, b] \subseteq I$.

Dimostrazione. Abbiamo f convessa in I se e solo se f' è crescente in I e ciò, stante l'ipotesi di derivabilità posta su f, si verifica se e solo se $f''(x) \geq 0$ internamente ad I.
Analogamente f è strettamente convessa se e solo se f' è strettamente crescente, cioè se e solo se f'' è non negativa nell'interno di I e non si annulla su alcun sottointervallo contenuto in I. □

4. ESTREMI DI FUNZIONI CONVESSE

In questo paragrafo mostriamo che gli estremi delle funzioni convesse, quando esistono, sono "facilmente" individuabili.

4.1. **Minimi Assoluti e Relativi.** Iniziamo il discorso osservando esplicitamente che una funzione convessa non è tenuta a presentare punti di minimo né assoluto né relativo: per lumeggiare tale circostanza forniamo il seguente:

`Esempio 11:` La funzione $f : [0, 1] \to \mathbb{R}$ definita ponendo:

$$f(x) := \begin{cases} 1 & ,\text{ se } x = 0 \\ x & ,\text{ se } 0 < x \leq 1 \end{cases}$$

è convessa (non strettamente) in $[0, 1]$, in quanto ha $f'(x) = 1$ crescente (non strettamente) nei punti di $]0, 1[$.
La f non è dotata di minimo in $[0, 1]$, in quanto $\inf_I f = 0$ e però in ogni punto $x \in I$ risulta $f(x) > 0$. ♦

Il seguente risultato è molto semplice da dimostrare e fornisce informazioni preziose sui punti di minimo assoluto di una funzione convessa:

PROPOSIZIONE 7 (Minimi di Funzioni Convesse)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale, $f : I \to \mathbb{R}$ ed $x_0 < x_1 \in I$.
Se f è convessa ed x_0 ed x_1 sono punti di minimo assoluto, allora ogni altro punto x appartenente all'intervallo di estremi $]x_0, x_1[$ è un punto di minimo assoluto per f.

Osservazione 24: Il succo di questa Proposizione si può esprimere come segue:

> "Se una funzione convessa prende il suo minimo, lo prende o in un solo punto oppure in infiniti punti". ♦

Dimostrazione. Per comodità poniamo $m := \min_{x \in I} f(x)$, di modo che $f(x_0) = m = f(x_1)$.
Come sappiamo, scelto $x \in]x_0, x_1[$, possiamo determinare un unico $t \in]0, 1[$ tale che $x = (1-t)x_0 + tx_1$ (cioè $t = \frac{x - x_0}{x_1 - x_0}$); usando la disuguaglianza di convessità (1) troviamo:

$$\begin{aligned} f(x) &\leq (1-t)f(x_0) + tf(x_1) \\ &= (1-t)m + tm \\ &= m \ , \end{aligned}$$

e, confrontando tale disuguaglianza con la più evidente $m \leq f(x)$, otteniamo $f(x) = m$. Dall'arbitrarietà nella scelta di $x \in]x_0, x_1[$ segue la tesi. □

PROPOSIZIONE 8 (Unicità del Minimo per Funzioni Strettamente Convesse)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale, $f : I \to \mathbb{R}$ ed $x_0 \in I$.
Se f è strettamente convessa e prende minimo in x_0, non esiste alcun altro punto $x_1 \neq x_0$ in cui f prende il suo minimo.

Osservazione 25: Il succo di questa Proposizione si può esprimere come segue:

> "Se una funzione strettamente convessa prende il suo minimo, lo prende in un solo punto". ♦

Dimostrazione. Per comodità, poniamo $m := \min_{x \in I} f(x)$.
Per assurdo, supponiamo che f abbia due punti di minimo distinti $x_0 < x_1 \in I$, di modo che $f(x_0) = m = f(x_1)$.
Comunque fissiamo $x \in]x_0, x_1[$ riusciamo a determinare un $t \in]0, 1[$ tale che $x = (1-t)x_0 + tx_1$, sicché per stretta convessità avremmo:

$$f(x) < (1-t)f(x_0) + tf(x_1) = (1-t)m + tm \ ,$$

cioè $f(x) < m$. **Ma ciò è assurdo**, in quanto per ogni $x \in I$ risulta $m \leq f(x)$ e dunque sarebbe violato il *Principio di Tricotomia*. □

Osserviamo esplicitamente che nelle due PROPOSIZIONI precedenti non si è esclusa la possibilità che il punto di minimo x_0 possa cadere negli estremi dell'intervallo I: ciò è dovuto al fatto che tale caso si presenta piuttosto naturalmente.

`Esempio 12:` La funzione $f(x) := x^2$ è strettamente convessa in $I = [0, 1]$ e prende il suo minimo nel punto $x_0 = 0$, estremo di I. ♦

Occupiamoci ora della localizzazione dei punti di estremo interni all'intervallo di definizione.
Una conseguenza notevole dell'esistenza di rette di supporto è la seguente:

Teorema 5 (Localizzazione dei Minimi di Funzioni Convesse)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale, $x_0 \in I$ ed $f : I \to \mathbb{R}$ una funzione convessa. Il punto x_0 è di minimo assoluto per f in I se e solo se $0 \in \partial f(x_0)$.

Osservazione 26: Per quanto detto nella **Osservazione 17**, la condizione $0 \in \partial f(x_0)$ equivale a $f'_-(x_0) \leq 0 \leq f'_+(x_0)$ se x_0 è un punto interno ad I, o ad $f'_+(a) \geq 0$ ovvero a $f'_-(b) \leq 0$ nel caso in cui x_0 coincida con uno degli estremi di I (in cui f è definita e derivabile da destra o da sinistra). ♦

Dimostrazione. $\Rightarrow$) Supponiamo che x_0 sia un punto di minimo assoluto per f in I, i.e. che risulti:

$$f(x) \geq f(x_0)$$

per ogni $x \in I$. Scelto $x \neq x_0$, allora si ha:

$$\begin{aligned} x < x_0 \quad &\Rightarrow \quad \mathrm{r}(x; x_0) \leq 0 \\ x > x_0 \quad &\Rightarrow \quad \mathrm{r}(x; x_0) \geq 0 \end{aligned}$$

cosicché, per il *Teorema di Permanenza del Segno Inverso*, risulta:

$$f'_-(x) = \lim_{x \to x_0^-} \mathrm{r}(x; x_0) \leq 0 \leq \lim_{x \to x_0^+} \mathrm{r}(x; x_0) = f'_+(x_0) \; ,$$

cioè $0 \in \partial f(x_0)$.
Nel caso in cui x_0 coincida con uno degli estremi di I si ragiona in maniera analoga.

$\Leftarrow$) Poiché $0 \in \partial f(x_0)$, la retta orizzontale di equazione $y = f(x_0)$ è una retta di supporto al grafico di f in $P_0 = (x_0, f(x_0))$; da ciò segue immediatamente che:

$$f(x) \geq f(x_0)$$

per ogni $x \in I$, per cui x_0 è un punto di minimo assoluto. □

Osservazione 27: Se f è convessa e se valgono le disuguaglianze strette $f'_-(x_0) < 0 < f'_+(x_0)$ (nel caso $x_0 \in \operatorname{int} I$), $f'_+(a) > 0$ (se $x_0 = a$) ovvero $f'_-(b) < 0$ (qualora $x_0 = b$), allora il punto x_0 è l'*unico* punto di minimo.
Infatti, in tali ipotesi, si possono fissare $f'_-(x_0) \leq m < 0$ e $0 < M \leq f'_+(x_0)$ ed ottenere:

$$\begin{aligned} x < x_0 \quad &\Rightarrow \quad f(x) - f(x_0) \geq m(x - x_0) > 0 \\ x > x_0 \quad &\Rightarrow \quad f(x) - f(x_0) \geq M(x - x_0) > 0 \; , \end{aligned}$$

cosicché $f(x) > f(x_0)$ per ogni $x \neq x_0$. ♦

Più interessante, ancorché più restrittivo del Teorema 5, è il fatto che per le funzioni convesse vale un inverso del *Teorema di Fermat*:

Proposizione 9 (Inverso del Teorema di Fermat per Funzioni Convesse)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale, $f : I \to \mathbb{R}$ ed $x_0 \in I$ un punto interno. Se f è convessa in I e se f è derivabile in x_0 e risulta $f'(x_0) = 0$, allora x_0 è un punto di minimo assoluto per f.

La dimostrazione è essenzialmente identica a quella del Teorema 5 ed è lasciata al lettore.

Osservazione 28: Se f è strettamente convessa e se $f'(x_0) = 0$, allora x_0 è l'*unico* punto di minimo per f in I. ♦

Osserviamo, invece, esplicitamente che ogni punto di minimo relativo di una funzione convessa è *ipso facto* un punto di minimo assoluto: vale infatti la:

Proposizione 10 (I Minimi Relativi di Funzioni Convesse sono Minimi Assoluti)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale ed $f : I \to \mathbb{R}$ convessa.
Se $x_0 \in I$ è un punto di minimo relativo per f, allora x_0 è un punto di minimo assoluto.

Dimostrazione. Supponiamo che x_0 sia un punto interno, ragionandosi in maniera analoga quando esso coincida con uno degli estremi di I.
Se $x_0 \in I$ è di minimo relativo si ha:

$$f(x) \geq f(x_0)$$

in un conveniente intorno $[x_0 - \delta, x_0 + \delta] \subseteq I$; ragionando come nella dimostrazione del Teorema 5 otteniamo nuovamente:

$$f'_-(x_0) = \lim_{x \to x_0^-} \mathrm{r}(x; x_0) \leq 0 \leq \lim_{x \to x_0^+} \mathrm{r}(x; x_0) = f'_+(x_0)$$

ossia $0 \in \partial f(x_0)$ e ciò, per il Teorema 5, equivale a dire che x_0 è un punto di minimo assoluto per f. □

4.2. **Massimi Assoluti e Relativi.** Iniziamo il paragrafo osservando che una funzione convessa non è tenuta a prendere massimo assoluto nel proprio intervallo di definizione, come mostrano i seguenti esempi:

`Esempio 13:` La funzione $f :]-1, 1[\to \mathbb{R}$ definita ponendo $f(x) := \frac{1}{1-x^2}$ è pari, positiva, continua e derivabile quante volte si vuole in $]-1, 1[$; inoltre, risultando:

$$\begin{aligned} f'(x) &= \frac{2x}{(1-x^2)^2} \\ &= 2x\ f^2(x) \\ f''(x) &= 2f^2(x) + 4x\ f(x) \cdot f'(x) \\ &= 2f^2(x) \cdot \Big(4x^2\ f(x) + 1\Big) \\ &> 0 \end{aligned}$$

per ogni $x \in]-1, 1[$, la f è anche strettamente convessa in $]-1, 1[$.
Evidentemente, però, f non può avere massimo assoluto in $]-1, 1[$ poiché, essendo:

$$\lim_{x \to -1^+} f(x) = +\infty = \lim_{x \to 1^-} f(x)\ ,$$

f non è limitata superiormente. ♢

`Esempio 14:` La funzione convessa $f(x) := |x|$ non ha massimo assoluto in $]-1, 1[$, pur essendo limitata superiormente. ♢

Per quanto riguarda i massimi delle funzioni convesse (quando esistono) possiamo dire certamente che essi non possono essere assunti in punti interni, se non nei casi banali.
Infatti, la caratterizzazione della convessità in termini di rapporti incrementali implica il seguente:

Teorema 6 (Principio Forte del Massimo)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale ed $f : I \to \mathbb{R}$ una funzione convessa.
La f prende massimo assoluto in un punto interno ad I se e solo se f è costante in I.

Dimostrazione. La dimostrazione dell'implicazione $\Leftarrow$ è banalissima, perciò proveremo solo la $\Rightarrow$.

$\Rightarrow$) Supponiamo che f prenda massimo assoluto in un punto x_0 interno ad I. Per definizione di massimo, abbiamo:

$$f(x) - f(x_0) \leq 0$$

ovunque in I; pertanto, scelti arbitrariamente $x, y \in I - \{x_0\}$ in modo che sia $x < x_0 < y$ troviamo:

$$\mathrm{r}(y; x_0) = \frac{f(y) - f(x_0)}{y - x_0} \leq 0 \leq \frac{f(x) - f(x_0)}{x - x_0} = \mathrm{r}(x; x_0) \; .$$

La convessità di f implica che $\mathrm{r}(\cdot; x_0)$ è crescente in $I - \{x_0\}$, cosicché $\mathrm{r}(x; x_0) \leq \mathrm{r}(y; x_0)$; confrontando questa relazione con la precedente troviamo:

$$\mathrm{r}(y; x_0) \leq 0 \leq \mathrm{r}(x; x_0) \leq \mathrm{r}(y; x_0)$$

che implica $\mathrm{r}(x; x_0) = 0 = \mathrm{r}(y; x_0)$, ossia $f(x) = f(x_0) = f(y)$. Ciò, stante l'arbitrarietà nella scelta di x ed y in $I - \{x_0\}$, significa che f è costante in I. □

Osservazione 29: Notiamo esplicitamente che la presenza di un massimo assoluto interno implica la costanza di f in *tutto* I, estremi inclusi (se I non è aperto).
Perciò la presenza di un massimo interno assicura che f è continua in *tutto* I, contro il fatto che, di norma, la continuità di una funzione convessa è assicurata solo nei punti interni. ⧫

Osservazione 30: Il Principio Forte del Massimo non vale per i massimi relativi. Infatti, nel caso f prenda un massimo relativo interno, ragionando come nella dimostrazione precedente riusciamo solo a concludere che f è costante in un opportuno intorno di x_0... Ma su ciò che accade fuori da tale intorno non riusciamo a dir nulla!
Ad esempio, la funzione $f : [-2, 2] \to \mathbb{R}$ definita ponendo:

$$f(x) := \begin{cases} -x & , \text{se } -2 \leq x \leq -1 \\ 1 & , \text{se } -1 \leq x \leq 1 \\ x & , \text{se } 1 \leq x \leq 2 \end{cases}$$

è convessa in $[-2, 2]$, ha un massimo relativo in $x_0 = 0$ (poiché f è costante intorno a 0) ma non è ovunque costante in $[-2, 2]$.
Questo esempio, tuttavia, illustra molto bene ciò che in realtà accade: i punti di massimo relativo interni all'intervallo di definizione di una funzione convessa sono "falsi" massimi, poiché essi riescono ad essere punti di estremo solo su sottointervalli in cui la funzione assume sempre lo stesso valore (i.e., è costante). ⧫

Per il Principio Forte del Massimo, gli eventuali punti di massimo assoluto di una funzione convessa non costante non possono andar ricercati nell'interno dell'intervallo di definizione, perciò possiamo affermare il:

Teorema 7 (Principio del Massimo)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale ed $f : I \to \mathbb{R}$ convessa in I.
Se f prende massimo in I, allora o f è costante in I oppure il massimo è preso negli estremi di I.

Dimostrazione. Se I è aperto, allora il massimo è preso necessariamente in un punto interno ad I e, per il *Principio Forte del Massimo*, f è costante in I.
Se I non è aperto, allora o il massimo è preso sulla frontiera oppure è preso in un

punto interno. Nel primo caso siamo a posto, perché la frontiera di I è costituita dagli estremi di I che appartengono ad I; mentre nel secondo caso f è costante in tutto I, sempre per il *Principio Forte del Massimo*. □

Osservazione 31: Per le funzioni concave valgono analoghi risultati, detti *Principio Forte del Minimo* e *Principio del Minimo*: lasciamo al lettore il compito di enunciarli e di dimostrarli basandosi sui risultati appena acquisiti. ♦

5. Disuguaglianze Convesse

5.1. **Disuguaglianze tra Medie.** Abbiamo già osservato che la funzione esponenziale è strettamente convessa su $\mathbb{R}$. Pertanto, la disuguaglianza di Jensen scritta per la funzione $f(x) := \mathbf{e}^x$ in corrispondenza di $n+1$ numeri reali $x_0, \ldots, x_n$ ed altrettanti scalari $t_0, \ldots, t_n \in [0,1]$ tali che $\sum_{k=0}^n t_k = 1$ fornisce:

$$\mathbf{e}^{\sum_{k=0}^n t_k x_k} \leq \sum_{k=0}^n t_k \mathbf{e}^{x_k} ,$$

con uguaglianza se e solo se $x_0 = x_1 = \cdots = x_n$.
Ponendo $a_k := \mathbf{e}^{x_k}$, dalla precedente segue la disuguaglianza:

$$a_0^{t_0} \cdot a_1^{t_1} \cdots a_n^{t_n} \leq \sum_{k=0}^n t_k a_k , \tag{8}$$

nota come *disuguaglianza tra media geometrica e media aritmetica pesate*, in cui vale l'uguaglianza solo se $a_0 = a_1 = \cdots = a_n$.
La (8) generalizza l'elementare *disuguaglianza tra media geometrica e media aritmetica* di due numeri reali positivi, cioè:

$$\sqrt{a_0 a_1} \leq \frac{a_0 + a_1}{2} ,$$

che dovrebbe esser già nota al lettore.[4]

Analogamente, la funzione potenza con esponente $\alpha > 1$ è strettamente convessa in $[0, +\infty[$. Pertanto la disuguaglianza di Jensen scritta per la funzione $f(x) := x^\alpha$ in corrispondenza di $n+1$ numeri positivi $x_0, \ldots, x_n$ ed altrettanti scalari $t_0, \ldots, t_n \in [0,1]$ tali che $\sum_{k=0}^n t_k = 1$ fornisce:

$$\left(\sum_{k=0}^n t_k x_k \right)^\alpha \leq \sum_{k=0}^n t_k x_k^\alpha ,$$

con uguaglianza se e solo se $x_0 = x_1 = \cdots = x_n$.
Prendendo $\alpha = \frac{q}{p}$, con $q > p > 0$, ed $x_k = a_k^p$, con $a_k \geq 0$, dalla precedente otteniamo la disuguaglianza:

$$\left(\sum_{k=0}^n t_k a_k^p \right)^{\frac{1}{p}} \leq \left(\sum_{k=0}^n t_k a_k^q \right)^{\frac{1}{q}} , \tag{9}$$

detta *disuguaglianza tra p-media e q-media pesate*, in cui vale l'uguaglianza se e solo se $a_0 = a_1 = \cdots = a_n$.
La (9) generalizza l'elementare *disuguaglianza tra media aritmetica e media quadratica* di due numeri reali positivi, ossia:

$$\frac{a_0 + a_1}{2} \leq \sqrt{\frac{a_0^2 + a_1^2}{2}} ,$$

di semplice dimostrazione.

[4] Se il lettore non la conosce, si diletti a dimostrarla usando il calcolo letterale.

5.2. **Disuguaglianza di Cauchy con ε.** Vogliamo provare che per ogni valore $\varepsilon > 0$ esistono delle costanti $C > 0$ tali che la disuguaglianza:

$$a_0 a_1 \leq \frac{\varepsilon}{2}\, a_0^2 + C\, a_1^2 \tag{10}$$

valga per ogni paio di numeri reali a_0 ed a_1.

In realtà è evidente che se a_0 ed a_1 hanno segno opposto o se uno dei due numeri è uguale a 0, la disuguaglianza (10) è certamente valida; d'altra parte, se $a_0, a_1 < 0$, la disuguaglianza (10) vale se e solo se essa è valida per i numeri positivi $-a_0$ e $-a_1$.
Da quanto appena osservato discende che per acquisire la tesi basta dimostrare l'esistenza di costanti $C > 0$ tali che la (10) valga per ogni paio di numeri $a_0, a_1 > 0$.

Dividendo membro a membro la (10) per a_1 ed introducendo la variabile ausiliaria $x = \frac{a_0}{a_1} > 0$ otteniamo:

$$x \leq \frac{\varepsilon}{2}\, x^2 + C\ ;$$

dunque dimostrare l'esistenza di $C > 0$ per cui vale la (10) equivale a provare l'esistenza di costanti $C > 0$ tali che la funzione $\varphi :]0, +\infty[\to \mathbb{R}$ definita ponendo $\varphi(x) := \frac{\varepsilon}{2} x^2 - x + C$ ha estremo inferiore ≥ 0.
La φ è continua e derivabile quanto si vuole in $]0, +\infty[$ ed in più risulta:

$$\varphi'(x) = \varepsilon\ x - 1$$
$$\varphi'' = \varepsilon > 0\ ,$$

cosicché φ è strettamente convessa in $]0, +\infty[$ e, per la Proposizione 9, essa prende minimo assoluto nell'unico punto in cui è nulla la derivata prima, i.e. $x_\varepsilon = \frac{1}{\varepsilon}$; conseguentemente, abbiamo:

$$\inf_{x>0} \varphi(x) = \min_{x>0} \varphi(x) = \varphi(x_\varepsilon) = C - \frac{1}{2\varepsilon}$$

e tale estremo inferiore è non negativo non appena si scelga $C \geq \frac{1}{2\varepsilon} =: c(\varepsilon)$. Il numero $c(\varepsilon)$ è anche detto *costante ottimale* per la (10).
Chiaramente, se scegliamo di scrivere la (10) con $C > c(\varepsilon)$, essa sarà sempre soddisfatta col segno di disuguaglianza stretta a meno del caso banale, cioè avremo:

$$a_0 a_1 < \frac{\varepsilon}{2}\, a_0^2 + C\, a_1^2$$

per ogni $a_0, a_1 \in \mathbb{R}$ non contemporaneamente nulli; invece, scegliendo $C = c(\varepsilon)$, l'uguaglianza in (10) sarà verificata, oltre al caso banale $a_0 = 0 = a_1$, anche quando $\frac{a_0}{a_1} = x_\varepsilon = \frac{1}{\varepsilon}$, ossia per $a_1 = \varepsilon a_0$.

Abbiamo così dimostrato la:

Proposizione 11 (Disuguaglianza di Cauchy con ε)
Per ogni $\varepsilon > 0$ esistono costanti $C > 0$ tali che la disuguaglianza (10) *è valida per ogni $a_0, a_1 \in \mathbb{R}$.*
Per la precisione, le costanti C che soddisfano la precedente sono tutte e sole quelle maggiori od uguali a $c(\varepsilon) = \frac{1}{2\varepsilon}$.
L'uguaglianza è soddisfatta in (10) *se e solo se $a_0 = 0 = a_1$ quando $C > c(\varepsilon)$ oppure se $a_1 = \varepsilon a_0$ quando $C = c(\varepsilon)$.*

Osservazione 32 (Disuguaglianza Integrale di Cauchy *)**:** La (10) può essere applicata per ottenere un'analoga disuguaglianza tra integrali definiti.

Prese due funzioni continue $f, g : [a, b] \to \mathbb{R}$, per ogni $x \in [a, b]$ possiamo porre $a_0 = |f(x)|$ e $a_1 = |g(x)|$ in (10) e ricavare:

$$|f(x)|\ |g(x)| \leq \frac{\varepsilon}{2}\ f^2(x) + C\ g^2(x)$$

da cui, integrando membro a membro, ricaviamo la relazione:

$$\int_a^b \big|f(x)g(x)\big|\ \mathrm{d}\,x \leq \frac{\varepsilon}{2}\ \int_a^b f^2(x)\ \mathrm{d}\,x + C\ \int_a^b g^2(x)\ \mathrm{d}\,x$$

detta *disuguaglianza integrale di Cauchy con* ε.
L'uguaglianza, evidentemente, si verifica solo se $f(x) = 0 = g(x)$ quando la costante C non è ottimale, oppure se $g(x) = \varepsilon f(x)$ nel caso $C = c(\varepsilon)$. ⧫

5.3. **Disuguaglianza di Young.** Scelti $x_0, x_1 \in \mathbb{R}$ e $t \in]0, 1[$, la stretta convessità dell'esponenziale implica:

$$\mathbf{e}^{(1-t)x_0 + tx_1} \leq (1-t)\mathbf{e}^{x_0} + t\mathbf{e}^{x_1}\ ,$$

con uguaglianza solo se $x_0 = x_1$.
Scegliendo $x_0 = \log a_0^p$, $x_1 = \log a_1^{\frac{p}{p-1}}$ e $t = \frac{p-1}{p}$ con $a_0, a_1 > 0$ e $p > 1$, dalla precedente traiamo:

$$\mathbf{e}^{\frac{1}{p}\log a_0^p + \frac{p-1}{p}\log a_1^{\frac{p}{p-1}}} \leq \frac{1}{p}\mathbf{e}^{\log a_0^p} + \frac{p-1}{p}\mathbf{e}^{\log a_1^{\frac{p}{p-1}}}$$

cioè:

$$a_0 a_1 \leq \frac{1}{p} a_0^p + \frac{p-1}{p} a_1^{\frac{p}{p-1}}\ , \tag{dY}$$

che è detta *disuguaglianza di Young*[5]. Notiamo che essa rimane valida anche se $a_0, a_1 = 0$ e che l'uguaglianza è soddisfatta solo se $a_1 = a_0^{p-1}$.

Osservazione 33 (Disuguaglianza integrale di Young *)**:** Prese due funzioni continue $f, g : [a, b] \to \mathbb{R}$, possiamo porre $a_0 = |f(x)|$ e $a_1 = |g(x)|$ nella disuguaglianza (dY); integrando ambo i membri otteniamo la relazione:

$$\int_a^b \big|f(x)g(x)\big|\ \mathrm{d}\,x \leq \frac{1}{p}\ \int_a |f(x)|^p\ \mathrm{d}\,x + \frac{p}{p-1}\ \int_a^b |g(x)|^{\frac{p}{p-1}}\ \mathrm{d}\,x$$

detta *disuguaglianza integrale di Young.*
L'uguaglianza vale, ovviamente, quando $|g(x)| = |f(x)|^{p-1}$ ovunque in $[a, b]$. ⧫

5.4. **Il Problema dello N-gono Inscritto con Perimetro Massimo.** Fissiamo $N \geq 3$.
Chiamiamo *N-gono* un poligono convesso avente N lati, N vertici ed N angoli interni.[6]
Diciamo che un N-gono è *inscritto in una circonferenza di raggio r* se esso ha come vertici N punti di tale circonferenza.
Con semplici considerazioni geometriche si può vedere che un N-gono inscritto in una circonferenza individua N angoli al centro $\vartheta_0, \ldots, \vartheta_{N-1}$, le cui ampiezze sommano a 2π, rispetto ai quali è molto semplice descrivere il perimetro dello N-gono.

---

[5]William Henry Young (1863 – 1942), matematico inglese.

[6]Per $N = 3$, otteniamo un triangolo; per $N = 4$, un quadrilatero; per $N = 5$, un pentagono; per $N = 6$, un esagono; per $N = 7$, un eptagono; per $N = 8$, un ottagono; per $N = 9$, un ennagono; per $N = 10$, un decagono; etc. . .

Infatti, il lato che sottende l'angolo al centro ϑ_k ha lunghezza:

$$l_k = r \sin \frac{\vartheta_k}{2}$$

cosicché il perimetro dello N-gono si può esprimere come funzione degli angoli al centro $\vartheta_0, \ldots \vartheta_{N-1}$ come:

$$P(\vartheta_0, \ldots, \vartheta_{N-1}) = \sum_{k=0}^{N-1} l_k = r \sum_{k=0}^{N-1} \sin \frac{\vartheta_k}{2} \; .$$

Ci domandiamo se tra tutti gli N-goni inscritti in una circonferenza di raggio r ce n'è qualcuno con perimetro massimo.

Per rispondere a tale domanda, notiamo che l'essere $\vartheta_k \in [0, 2\pi]$ implica $\frac{\vartheta_k}{2} \in [0, \pi]$ e che la funzione $f(x) := \sin x$ è strettamente concava in $[0, \pi]$ (cfr. `Esempio 10`); pertanto la disuguaglianza di Jensen con $x_k = \frac{\vartheta_k}{2}$ e $t_k = \frac{1}{N}$ per ogni $k = 0, \ldots, N-1$ implica:

$$\sin\left(\sum_{k=0}^{N-1} \frac{1}{N} \frac{\vartheta_k}{2}\right) \geq \sum_{k=0}^{N-1} \frac{1}{N} \sin \frac{\vartheta_k}{2}$$

ossia:

$$N \sin \frac{2\pi}{2N} \geq \sum_{k=0}^{N-1} \sin \frac{\vartheta_k}{2} \; ,$$

con uguaglianza solo se $\vartheta_0 = \cdots = \vartheta_{N-1} = \frac{2\pi}{N}$.
Da ciò segue immediatamente che:

$$P(\vartheta_0, \ldots, \vartheta_{N-1}) \leq rN \sin \frac{2\pi}{2N} = P\left(\frac{2\pi}{N}, \ldots, \frac{2\pi}{N}\right) ,$$

con uguaglianza solo se $\vartheta_0 = \cdots = \vartheta_{N-1} = \frac{2\pi}{N}$.
Ciò significa che gli N-goni regolari[7] hanno perimetro strettamente più grande di tutti gli altri N-goni inscritti nella medesima circonferenza.

5.5. **Disuguaglianza Integrale di Jensen *.** Consideriamo una funzione continua $\varphi : [a, b] \to]A, B[$ ed una funzione convessa $f :]A, B[\to \mathbb{R}$.
Per il *Teorema della Media Integrale* abbiamo:

$$A < \frac{1}{b-a} \int_a^b \varphi(t)\, \mathrm{d}\, t < B \; ,$$

cosicché il punto $x_0 = \frac{1}{b-a} \int \varphi(t)\, \mathrm{d}\, t$ è interno all'intervallo di definizione di f.
Scelto un $m \in [f'_-(x_0), f'_+(x_0)]$, la retta d'equazione $y = m(x - x_0) + f(x_0)$ è una retta di supporto per il grafico di f, sicché risulta:

$$f(x) \geq m(x - x_0) + f(x_0)$$

per ogni $x \in]A, B[$; in particolare la disuguaglianza vale per $x = \varphi(t)$ ed in tal caso si riscrive:

$$f(\varphi(t)) \geq m\,(\varphi(t) - x_0) + f(x_0) \; .$$

[7]Un N-gono è detto *regolare* se esso ha tutti i lati congruenti. In tal caso gli angoli al centro sottesi dai suoi lati hanno tutti ugual misura, pari a $\frac{2\pi}{N}$.

Integrando ambo i membri su $[a,b]$ e dividendo per $b-a$ otteniamo:

$$\begin{aligned}\frac{1}{b-a}\int_a^b f\left(\varphi(t)\right)\mathrm{d}\,t &\geq m\Big(\underbrace{\frac{1}{b-a}\int_a^b \varphi(t)\,\mathrm{d}\,t}_{=x_0} - x_0\Big) + f(x_0)\\ &= m(x_0 - x_0) + f(x_0)\\ &= f\left(\frac{1}{b-a}\int_a^b \varphi(t)\,\mathrm{d}\,t\right)\ .\end{aligned}$$

Abbiamo così provato la disuguaglianza:

$$f\left(\frac{1}{b-a}\int_a^b \varphi(t)\,\mathrm{d}\,t\right) \leq \frac{1}{b-a}\int_a^b f\left(\varphi(t)\right)\mathrm{d}\,t \tag{dJ}$$

nota come *disuguaglianza integrale di Jensen*.
Ovviamente, se φ è concava, la disuguaglianza di Jensen ha verso opposto.

Si può inoltre dimostrare una disuguaglianza integrale di Jensen "pesata". Per fare ciò, notiamo che la (dJ) può essere riscritta come:

$$f\left(\frac{1}{\int_a^b 1\,\mathrm{d}\,t}\int_a^b \varphi(t)\cdot 1\,\mathrm{d}\,t\right) \leq \frac{1}{\int_a^b 1\,\mathrm{d}\,t}\int_a^b f\left(\varphi(t)\right)\cdot 1\,\mathrm{d}\,t$$

e che, sostituendo la funzione identicamente uguale ad 1 con un'arbitraria funzione continua $w:[a,b]\to[0,+\infty[$ e non identicamente nulla, la precedente assume la forma:

$$f\left(\frac{1}{\int_a^b w(t)\,\mathrm{d}\,t}\int_a^b \varphi(t)\cdot w(t)\,\mathrm{d}\,t\right) \leq \frac{1}{\int_a^b w(t)\,\mathrm{d}\,t}\int_a^b f\left(\varphi(t)\right)\cdot w(t)\,\mathrm{d}\,t\ . \tag{dJw}$$

La (dJw) è detta *disuguaglianza integrale di Jensen pesata* e la sua dimostrazione, analoga a quella proposta più sopra, è lasciata come esercizio al lettore.

Osservazione 34: Notiamo che, posto $W := \int_a^b w(\tau)\,\mathrm{d}\,\tau$ e $t(x) = \frac{w(x)}{W}$, la disuguaglianza integrale di Jensen può essere riscritta:

$$f\left(\int_a^b \varphi(x)\ t(x)\,\mathrm{d}\,x\right) \leq \int_a^b f\left(\varphi(x)\right)\ t(x)\,\mathrm{d}\,x$$

e che la funzione $t(\cdot)$ è continua e gode della proprietà:

$$\int_a^b t(x)\,\mathrm{d}\,x = 1\ ,$$

del tutto analoga a quella dei coefficienti $t_1,\ldots,t_n$ nella disuguaglianza di convessità di Jensen.
Posto $\varphi(t) = t$, la precedente fornisce:

$$f\left(\int_a^b t(x)x\,\mathrm{d}\,x\right) \leq \int_a^b t(x)f(x)\,\mathrm{d}\,x$$

la quale può essere euristicamente interpretata come una "versione continua" della disuguaglianza di convessità di Jensen (2'), nella quale le somme $\sum_{k=1}^n t_k x_k$ e $\sum_{k=1}^n t_k f(x_k)$ sono sostituite dagli integrali $\int_a^b t(x)x\,\mathrm{d}\,x$ e $\int_a^b t(x)f(x)\,\mathrm{d}\,x$. ♦

Osservazione 35 (Caso d'Uguaglianza)**:** Se f è strettamente convessa, l'uguaglianza in (dJ) vale se e solo se φ è costante.
Ripercorrendo a ritroso la dimostrazione, ci accorgiamo che, se $f\left(\frac{1}{b-a}\int_a^b \varphi(t)\,\mathrm{d}\,t\right) = \frac{1}{b-a}\int_a^b f\left(\varphi(t)\right)\mathrm{d}\,t$, si ha uguaglianza in $f(\varphi(t)) \geq m(\varphi(t)-x_0)+f(x_0)$ per ogni possibile $t \in [a,b]$; ciò significa che $\varphi(t)$ coincide con l'ascissa x_0 dell'unico punto di contatto tra il grafico della funzione strettamente convessa f e la sua retta di supporto $y = m(x-x_0)+f(x_0)$. Dunque $\varphi(t) = x_0 = \frac{1}{b-a}\int_a^b \varphi(t)\,\mathrm{d}\,t$ ovunque in $[a,b]$. ♦

5.6. **Perimetro e Simmetrizzazione di Steiner di Alcuni Insiemi Piani *.** Informalmente parlando, la cosiddetta *simmetrizzazione di Steiner*[8] è un metodo che consente di modificare una figura geometrica piana $\mathcal{F}$ in modo da ottenere una nuova figura $\mathcal{F}^s$ (detta *simmetrizzata di Steiner di* $\mathcal{F}$) dotata di simmetria assiale ed avente la stessa area di $\mathcal{F}$.
In questo paragrafo ci occuperemo di mostrare come ottenere la simmetrizzata di Steiner di alcune particolari figure piane, i.e. quelle delimitate da grafici di funzioni di classe C^1, e mostreremo che il perimetro di tali figure decresce quando esse siano sottoposte a tale operazione. Tale risultato può essere ampiamente generalizzato e le sue generalizzazioni sono molto utili quando si vogliano analizzare i rapporti tra quantità ingegneristicamente rilevanti e caratteristiche geometriche di varie tipologie di sistemi.[9]

Consideriamo due funzioni $f_1, f_2 : [a,b] \to \mathbb{R}$ continue in $[a,b]$, derivabili in $]a,b[$ con derivata continua fin negli estremi e tali che $f_1(x) \leq f_2(x)$ per ogni $x \in [a,b]$. Chiamiamo $\mathcal{F}$ la porzione di piano compresa tra le rette verticali di equazioni $x = a$ ed $x = b$ ed i diagrammi dei grafici di f_1 ed f_2, i.e. poniamo:

$$\mathcal{F} := \left\{(x,y) \in \mathbb{R}^2 :\ a \leq x \leq b \text{ e } f_1(x) \leq y \leq f_2(x)\right\}.$$

Inoltre, fissiamo come asse di simmetrizzazione l'asse delle x.
La simmetrizzata di Steiner di $\mathcal{F}$ è la figura $\mathcal{F}^s$ che si ottiene col procedimento descritto qui di seguito:

(1) si sceglie un punto $a \leq \xi \leq b$ e si traccia la retta verticale di equazione $x = \xi$;

(2) dalla definizione di $\mathcal{F}$ segue che l'intersezione $\mathcal{F}_\xi$ tra la retta $x = \xi$ e la figura $\mathcal{F}$ è un segmento parallelo all'asse y (o, eventualmente, un punto), precisamente $\mathcal{F}_\xi = \{\xi\} \times [f_1(\xi), f_2(\xi)]$;

(3) si muove il segmento $\mathcal{F}_\xi$ lungo la retta su cui esso giace fino a far coincidere il suo punto medio con il punto $(\xi, 0)$ (d'intersezione tra la retta $x = \xi$ e l'asse delle x);

(4) detto $\mathcal{F}^s_\xi$ il segmento ottenuto al passo precedente, la simmetrizzata di Steiner di $\mathcal{F}$ coincide con l'unione di tutti i segmenti $\mathcal{F}^s_\xi$ ottenuti al variare di

---

[8] Jakob Steiner (1796 – 1863), matematico svizzero, noto per i suoi numerosi contributi alla Geometria.

[9] Ad esempio, usando la simmetrizzazione di Steiner è possibile mostrare che, a parità di caratteristiche fisiche, tra le membrane aventi la stessa area quella che emette la nota principale più grave è quella circolare. Analogamente, usando tecniche simili a quella di Steiner si può dimostrare che, a parità di caratteristiche fisiche, le travi che presentano la maggiore rigidità alla torsione sono quelle a sezione circolare.

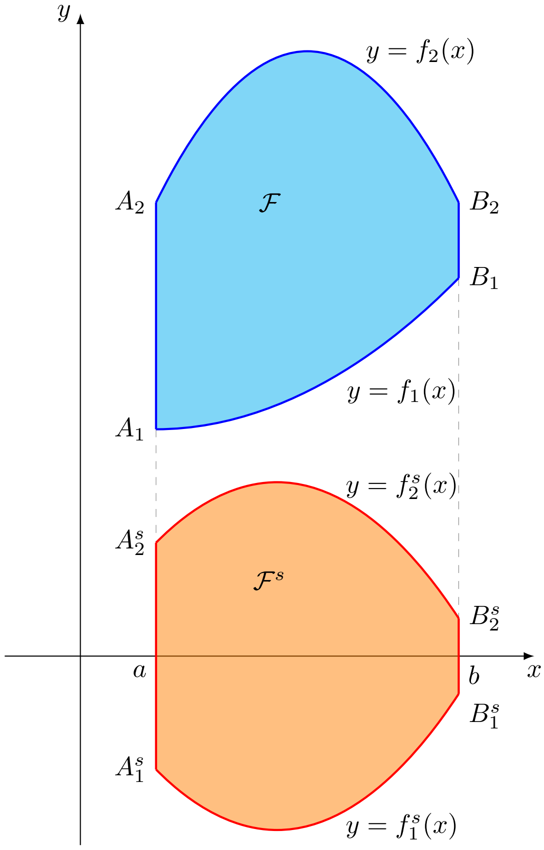

FIGURA 4. Simmetrizzazione di Steiner di una figura $\mathcal{F}$ delimitata dai grafici di due funzioni derivabili. Si noti che i segmenti $\mathcal{F}_a^s = \overline{A_1^s A_2^s}$ ed $\mathcal{F}_b^s = \overline{B_1^s B_2^s}$ sono ottenuti dai segmenti $\mathcal{F}_a = \overline{A_1 A_2}$ ed $\mathcal{F}_b = \overline{B_1 B_2}$ come indicato nella costruzione.

$\xi \in [a, b]$, i.e. $\mathcal{F}^s := \cup_{a \le \xi \le b} \mathcal{F}_\xi^s$.

La FIGURA 4 mostra la simmetrizzata di Steiner della figura $\mathcal{F}$ costituita dai punti compresi tra le rette di equazione $x = 1$ ed $x = 5$ ed i grafici delle funzioni $f_1(x) := \frac{1}{8}(x-1)^2 + 3$ ed $f_2(x) := 8 - \frac{1}{2}(x-1)^2$.

Non è difficile dimostrare che $\mathcal{F}^s$ è ancora una figura delimitata da grafici di funzioni definite in $[a, b]$ e che, precisamente, si ha:

$$\mathcal{F}^s = \left\{ (x, y) \in \mathbb{R}^2 : \ a \le x \le b \text{ e } -\frac{f_2(x) - f_1(x)}{2} \le y \le \frac{f_2(x) - f_1(x)}{2} \right\} .$$

Conseguentemente, dato che f_1 ed f_2 sono continue in $[a, b]$ e derivabili in $]a, b[$, anche le funzioni:

$$\begin{aligned} f_2^s(x) &:= \frac{f_2(x) - f_1(x)}{2} \\ f_1^s(x) &:= -\frac{f_2(x) - f_1(x)}{2} \\ &= -f_2^s(x) , \end{aligned}$$

i cui grafici delimitano $\mathcal{F}^s$, godono delle stesse proprietà.

Osservazione 36 (La Simmetrizzazione non Cambia l'Area)**:** Osserviamo esplicitamente che, come detto sopra, le figure $\mathcal{F}$ e la sua simmetrizzata di Steiner $\mathcal{F}^s$ hanno la stessa area: infatti, per noti risultati sull'integrazione definita [DM14, § 3, Osservazione 6], abbiamo:

$$\begin{aligned}\operatorname{area}(\mathcal{F}^s) &= \int_a^b \left| f_2^s(x) - f_1^s(x) \right| \,\mathrm{d}\,x \\ &= \int_a^b 2\, f_2^s(x) \,\mathrm{d}\,x \\ &= \int_a^b \Big(f_2(x) - f_1(x) \Big) \,\mathrm{d}\,x \\ &= \operatorname{area}(\mathcal{F}) \ .\end{aligned}$$

♦

Il perimetro della figura $\mathcal{F}$ è la somma delle lunghezze dei segmenti verticali di estremi $A_1 := (a, f_1(a))$, $A_2 := (a, f_2(a))$ e $B_1 := (b, f_1(b))$, $B_2 := (b, f_2(b))$ e delle lunghezze delle curve grafico di f_1 ed f_2: per quanto provato in [DM14, § 2], dunque, risulta:

$$\begin{aligned}\operatorname{per}(\mathcal{F}) = \Big(f_2(a) - f_1(a)\Big) + \Big(f_2(b) - f_1(b)\Big) + \int_a^b \sqrt{1 + \Big(f_1'(x)\Big)^2} \,\mathrm{d}\,x \\ + \int_a^b \sqrt{1 + \Big(f_2'(x)\Big)^2} \,\mathrm{d}\,x \ ;\end{aligned}$$

analogamente il perimetro di $\mathcal{F}^s$ è dato da:

$$\begin{aligned}\operatorname{per}(\mathcal{F}^s) &= \Big(f_2^s(a) - f_1^s(a)\Big) + \Big(f_2^s(b) - f_1^s(b)\Big) + \int_a^b \sqrt{1 + \Big((f_1^s)'(x)\Big)^2} \,\mathrm{d}\,x \\ &\quad + \int_a^b \sqrt{1 + \Big((f_2^s)'(x)\Big)^2} \,\mathrm{d}\,x \\ &= 2f_2^s(a) + 2f_2^s(b) + 2\int_a^b \sqrt{1 + \Big((f_2^s)'(x)\Big)^2} \,\mathrm{d}\,x \ .\end{aligned}$$

Vogliamo provare il:

TEOREMA 8 (Il Perimetro Decresce per Simmetrizzazione)
Sia $\mathcal{F}$ una figura delimitata dai grafici di due funzioni $f_1, f_2 : [a, b] \to \mathbb{R}$ continue e derivabili in $[a, b]$ con derivate continue fin negli estremi.
Detta $\mathcal{F}^s$ la simmetrizzata di Steiner di $\mathcal{F}$, risulta:

$$\operatorname{per}(\mathcal{F}^s) \le \operatorname{per}(\mathcal{F}) \ . \tag{11}$$

Dimostrazione. Dato che:

$$\begin{aligned}2f_2^s(a) &= f_2(a) - f_1(a) \\ 2f_2^s(b) &= f_2(b) - f_1(b) \ ,\end{aligned}$$

è chiaro che per provare la disuguaglianza (11) occorre e basta provare che:

$$
\begin{aligned}
2\int_a^b \sqrt{1+\Big((f_2^s)'(x)\Big)^2}\ \mathrm{d}\,x \leq & \int_a^b \sqrt{1+\Big(f_1'(x)\Big)^2}\ \mathrm{d}\,x \\
& + \int_a^b \sqrt{1+\Big(f_2'(x)\Big)^2}\ \mathrm{d}\,x\ .
\end{aligned} \tag{12}
$$

Consideriamo la funzione $\varphi : \mathbb{R} \to \mathbb{R}$ definita ponendo $\varphi(u) := \sqrt{1+u^2}$: evidentemente φ è pari, positiva, continua e derivabile quante volte si vuole in $\mathbb{R}$ ed ha:

$$
\begin{aligned}
\varphi'(u) &= \frac{u}{\sqrt{1+u^2}} \\
&= \frac{u}{\varphi(u)} \\
\varphi''(u) &= \frac{\varphi(u) - u\varphi'(u)}{\varphi^2(u)} \\
&= \frac{\varphi(u) - \frac{u^2}{\varphi(u)}}{\varphi^2(u)} \\
&= \frac{\varphi^2(u) - u^2}{\varphi^3(u)} \\
&= \frac{1}{\varphi^3(u)} > 0\ ,
\end{aligned}
$$

cosicché φ è strettamente convessa in $\mathbb{R}$. Conseguentemente, per ogni $x \in [a,b]$ si ha:

$$
\begin{aligned}
\varphi\Big(f_2^s(x)\Big) &= \varphi\left(\frac{f_2(x) - f_1(x)}{2}\right) \\
&= \varphi\left(\frac{1}{2}\ f_2(x) + \frac{1}{2}\ \Big(-f_1(x)\Big)\right) \\
&\overset{(1)}{\leq} \frac{1}{2}\ \varphi\Big(f_2(x)\Big) + \frac{1}{2}\ \varphi\Big(-f_1(x)\Big) \\
&= \frac{1}{2}\ \varphi\Big(f_2(x)\Big) + \frac{1}{2}\ \varphi\Big(f_1(x)\Big)
\end{aligned}
$$

e perciò integrando il primo e l'ultimo membro su $[a,b]$ si trova:

$$
\int_a^b \sqrt{1+\Big((f_2^s)'(x)\Big)^2}\ \mathrm{d}\,x \leq \frac{1}{2}\int_a^b \sqrt{1+\Big(f_1'(x)\Big)^2}\ \mathrm{d}\,x + \frac{1}{2}\int_a^b \sqrt{1+\Big(f_2'(x)\Big)^2}\ \mathrm{d}\,x\ ,
$$

che è la (12). □

Esercizi

`Esercizio 1 (Disuguaglianza di Convessità Ciclica):` Siano $I \subseteq \mathbb{R}$ un intervallo non banale ed $f : I \to \mathbb{R}$.
Provare che f è convessa in I se e solo se per ogni terna di punti ordinati $x_0 < x_1 < x_2 \in I$ risulta:

$$
f(x_0)\cdot(x_1 - x_2) + f(x_1)\cdot(x_2 - x_0) + f(x_2)\cdot(x_0 - x_1) \leq 0\ .
$$

`Esercizio 2 (1/2-Convessità):` Siano $I \subseteq \mathbb{R}$ un intervallo non banale ed $f : I \to \mathbb{R}$.

Si dice che f è $1/2$-*convessa* (o *convessa nel punto medio*) *in* I se e solo se risulta:

$$f\left(\frac{y_1+y_2}{2}\right) \leq \frac{f(y_1)+f(y_2)}{2}$$

per ogni $y_1 < y_2 \in I$.
Vogliamo dimostrare che, se f è continua ed $1/2$-convessa in I, allora f è convessa in I.

0. Dimostrare che i *numeri diadici*, cioè i numeri razionali del tipo $\frac{m}{2^n}$ ($m, n \in \mathbb{N}$ ed $m \leq 2^n$), sono densi in $[0,1]$, i.e. che per ogni $t \in [0,1]$ è vera la seguente proposizione:

$$\forall \varepsilon > 0,\ \exists m, n \in \mathbb{N}: \quad m \leq 2^n \text{ e } \left|t - \frac{m}{2^n}\right| < \varepsilon\ .$$

Mostrare che la precedente equivale ad affermare che per ogni $t \in [0,1]$ esiste una successione (t_k) di numeri diadici che converge verso λ.

[`Suggerimento`: Mostrare che, fissato $\varepsilon > 0$, si può scegliere $n \in \mathbb{N}$ in modo che $1/2^n < \varepsilon$; scelto $t \in [0,1]$, si può porre $m := \lfloor 2^n t \rfloor$ (in cui $\lfloor\cdot\rfloor$ denota la parte intera); osservare che, per le proprietà della parte intera, risulta $m \leq 2^n t < m+1$ e concludere.]

1. Provare che la disuguaglianza di $1/2$-convessità implica:

$$f\left(\frac{y_1+y_2+\cdots+y_{2^n}}{2^n}\right) \leq \frac{f(y_1)+f(y_2)+\cdots+f(y_{2^n})}{2^n}$$

per ogni $n \in \mathbb{N}$ ed ogni scelta di 2^n punti $y_1 < y_2 < \cdots < y_{2^n} \in I$.

[`Suggerimento`: Fare induzione su n.]

2. Mostrare che la disuguaglianza precedente implica che risulta:

$$f\left(\left(1-\frac{m}{2^n}\right)\ x_0 + \frac{m}{2^n}\ x_1\right) \leq \left(1-\frac{m}{2^n}\right)\ f(x_0) + \frac{m}{2^n}\ f(x_1)$$

per ogni $x_0 < x_1 \in I$ ed ogni numero diadico $\frac{m}{2^n}$ con $m \leq 2^n$, cosicché f soddisfa la (1) quando $t = \frac{m}{2^n} \in [0,1]$.

[`Suggerimento`: Scelto $n \in \mathbb{N}$, se $m = 0, 2^n$, la cosa è banale; se $0 < m < 2^n$, scrivere la disuguaglianza del punto **1** per $y_1 = \cdots y_m = x_1$ e $y_{m+1} = \cdots = y_{2^n} = x_0$.]

3. Usare la continuità di f e la densità dei numeri diadici in $[0,1]$ per mostrare che la (1) è soddisfatta per ogni $t \in [0,1]$.

[`Suggerimento`: Fissato $t \in [0,1]$, è possibile scegliere un successione di numeri diadici $(t_k) \subseteq [0,1]$ tale che $t_k \to t$; scelti $x_0 < x_1 \in I$, si ha $f((1-t)x_0 + tx_1) = \lim_k f((1-t_k)x_0 + t_k x_1) \leq \ldots$]

`Esercizio 3:` Sia $I \subseteq \mathbb{R}$ un intervallo non banale.

1. Provare che se $f, g : I \to \mathbb{R}$ sono entrambe convesse in I, allora pure la funzione $h : I \to \mathbb{R}$ definita ponendo:

$$h(x) := \max\left\{f(x), g(x)\right\}$$

è convessa in I.

2. Più in generale, provare che se le funzioni $f_1, \dots, f_n : I \to \mathbb{R}$ sono convesse in I, allora pure la funzione $h : I \to \mathbb{R}$ definita ponendo:

$$h(x) := \max\Big\{f_1(x), \dots, f_n(x)\Big\}$$

è convessa in I.

`Esercizio 4 (Entropia):` Sia $f : [0, +\infty[\to \mathbb{R}$ la funzione definita ponendo:

$$f(x) := \begin{cases} x \log x & , \text{ se } x > 0 \\ 0 & , \text{ se } x = 0 \end{cases} .$$

1. Provare che f è continua e strettamente convessa in $[0, +\infty[$.

2. Per ogni fissata N-upla $p = (p_1, \dots, p_N)$ di numeri tali che $p_1, \dots, p_N \geq 0$ e $\displaystyle\sum_{n=1}^{N} p_n = 1$, si chiama *entropia di* p la quantità:

$$H(p) := -\sum_{n=1}^{N} f(p_n) = -\sum_{n=1}^{N} p_n \ \log p_n \ .$$

Dimostrare che $H(p) \geq 0$ e che $H(p) = 0$ se e solo se esiste un indice $\nu \in \{1, \dots, N\}$ tale che $p_\nu = 1$ e $p_n = 0$ per $n \neq \nu$.

[**Suggerimento**: Osservare preliminarmente che le condizioni $p_n \geq 0$ e $\sum_{n=1}^{N} p_n = 1$ implicano $0 \leq p_n \leq 1$. L'implicazione $\Leftarrow$ è banale; per provare la $\Rightarrow$, ragionare per assurdo.]

3. Mostrare che $H(p)$ è massima se e solo se $p_n = \frac{1}{N}$ per ogni $n = 1, \dots, N$.

[**Suggerimento**: Moltiplicare e dividere $\sum_{n=1}^{N} p_n \ \log p_n$ per N ed usare la disuguaglianza di convessità di Jensen (1').]

`Esercizio 5 (Unicità degli Zeri di Funzioni Convesse):` Siano $a < b \in \mathbb{R}$ ed $f : [a, b] \to \mathbb{R}$ una funzione continua e convessa in $[a, b]$.
Provare che se $f(a) \cdot f(b) < 0$, allora esiste un unico punto $\xi \in]a, b[$ tale che $f(\xi) = 0$.

`Esercizio 6 (Teorema di Lagrange Approssimato):` Siano $a < b \in \mathbb{R}$ ed $f : [a, b] \to \mathbb{R}$ convessa in $[a, b]$.
Provare che esiste un punto $\xi \in]a, b[$ tale che:

$$f'_-(\xi) \leq \frac{f(b) - f(a)}{b - a} \leq f'_+(\xi) \ .$$

`Esercizio 7:` Siano $I \subseteq \mathbb{R}$ un intervallo non banale, $\varphi : I \to \mathbb{R}$ una funzione crescente ed $c \in I$.
Dimostrare che la funzione $f : I \to \mathbb{R}$ definita ponendo:

$$f(x) := \int_c^x \varphi(t) \ \mathrm{d}\, t$$

è convessa in I (strettamente se φ è strettamente crescente).

[`Suggerimento`: Se φ è continua, la dimostrazione è semplice (basta usare il T.F.C.I.); altrimenti, si dimostri che f è 1/2-convessa (cfr. `Esercizio` 2) valutando le differenze $f(x_1) - f(\frac{x_0 + x_1}{2})$ ed $f(\frac{x_0 + x_1}{2}) - f(x_0)$; si usi la continuità di f per concludere.]

`Esercizio 8 (Assoluta Continuità delle Funzioni Convesse):` Siano $I \subseteq \mathbb{R}$ un intervallo non banale ed $f : I \to \mathbb{R}$.
Si dice che f è *assolutamente continua in* I se e solo se in corrispondenza di ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che per ogni $N \in \mathbb{N}$ e per famiglia finita d'intervalli a due a due disgiunti $[a_1, b_1], [a_2, b_2], \ldots, [a_N, b_N] \subseteq I$ risulta:

$$\text{(AC)} \qquad \sum_{n=1}^{N} b_n - a_n < \delta \quad \Rightarrow \quad \sum_{n=1}^{N} \big| f(b_n) - f(a_n) \big| < \varepsilon \; .$$

Dimostrare che se $f : I \to \mathbb{R}$ è convessa (o concava) in I, allora f è assolutamente continua in I.

`Esercizio 9:` Sia $f : \mathbb{R} \to \mathbb{R}$ la funzione esponenziale $f(x) := a^x$ con $a > 1$.
Provare che per ogni $m \geq 0$ esiste un unico valore $f^*(m) \in \mathbb{R}$ tale che:

$$(13) \qquad \forall x \in \mathbb{R}, \quad mx \leq f(x) + f^*(m) \; .$$

Cosa succede per $0 < a < 1$?

[`Suggerimento`: Chiaramente $f^*(m)$ coincide con l'estremo superiore della funzione $mx - f(x)$; determinare tale valore con le tecniche del Calcolo Differenziale.]

`Esercizio 10:` Sia $f : \mathbb{R} \to \mathbb{R}$ la funzione definita ponendo $f(x) := \frac{\varepsilon}{2} x^2$ con $\varepsilon > 0$.
Provare che per ogni $m \in \mathbb{R}$ esiste un unico valore $f^*(m) \in \mathbb{R}$ per cui valga la (13).
Mostrare che la (13) coincide con la *disuguaglianza di Cauchy con* ε.

`Esercizio 11 (Disuguaglianze di Hermite[10]-Hadamard[11]):` Siano $a < b \in \mathbb{R}$ ed $f : [a, b] \to \mathbb{R}$ continua e convessa in $[a, b]$.
Provar che:

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a} \int_a^b f(x) \;\mathrm{d}\,x \leq \frac{f(a) + f(b)}{2} \; .$$

[`Suggerimento`: Per la prima disuguaglianza, usare la (dJ) con $\varphi(t) := t$; per la seconda, fare la sostituzione $t = a + (b - a)x$ nell'integrale ed usare la convessità di f.]

`Esercizio 12:` Siano $a > 0$ ed $f_1, f_2 : [-a, a] \to \mathbb{R}$ continue, pari e tali che $f_1(x) \leq f_2(x)$ in $[-a, a]$.

1. Provare che la figura $\mathcal{F}$ compresa tra le rette di equazione $x = \pm a$ ed i grafici di f_1 ed f_2 è simmetrica rispetto all'asse y.

[**Suggerimento**: Si deve dimostrare che $(x, y) \in \mathcal{F}$ implica $(-x, y) \in \mathcal{F}$. Sfruttare le proprietà di f_1 ed f_2.]

2. Dimostrare che la simmetrizzata di Steiner $\mathcal{F}^s$ è simmetrica rispetto ad entrambi gli assi, ossia che essa è simmetrica rispetto all'origine degli assi O.

[**Suggerimento**: Si deve provare che $(x, y) \in \mathcal{F}^s$ implica $(-x, y), (x, -y) \in \mathcal{F}^s$ oppure $(-x, -y) \in \mathcal{F}^s$.]

---

[10] Charles Hermite (1822–1901), matematico francese.
[11] Jacques Hadamard (1865–1963), matematico francese.

Appendice A. Derivabilità "Quasi Ovunque" delle Funzioni Convesse

Abbiamo visto che una funzione convessa non è necessariamente derivabile in ogni punto del suo intervallo di definizione (cfr. Esempio 2); tuttavia, in generale, la derivabilità è assicurata in tutto l'interno dell'intervallo di definizione a meno dei punti di una successione:

Teorema 9 (Derivabilità "Quasi Ovunque" delle Funzioni Convesse)
Siano $I \subseteq \mathbb{R}$ un intervallo non banale ed $f : I \to \mathbb{R}$.
Se f è convessa, allora essa è derivabile nell'interno di I ad eccezione al più di un'infinità numerabile di punti.

Alla dimostrazione di tale risultato fondamentale, premettiamo un lemma tecnico sulle funzioni derivata destra e sinistra di una funzione convessa:

Lemma 2 (Monotònia delle Derivate Destra e Sinistra)
Siano I un intervallo non banale ed $f : I \to \mathbb{R}$.
Se f è convessa in I, allora le funzioni f'_+ ed f'_- sono crescenti nell'interno di I.

Dimostrazione. Fissiamo $x_0 < x_1$ interni ad I e scegliamo tre punti x, y e z interni ad I e tali che $x < x_0 < y < x_1 < z$. La monotònia dei rapporti incrementali di f assicura che:

$$\begin{aligned} \mathrm{r}(x; x_0) \leq \mathrm{r}(y; x_0) = \mathrm{r}(x_0; y) \leq \mathrm{r}(x_1; y) = \mathrm{r}(y; x_1) \\ \mathrm{r}(y; x_0) = \mathrm{r}(x_0; y) \leq \mathrm{r}(x_1; y) = \mathrm{r}(y; x_1) \leq \mathrm{r}(z; x_1) \end{aligned}$$

dunque:

$$\begin{aligned} f'_-(x_0) &= \lim_{x \to x_0^-} \mathrm{r}(x; x_0) \leq \lim_{y \to x_1^-} \mathrm{r}(y; x_1) = f'_-(x_1) \\ f'_+(x_0) &= \lim_{y \to x_0^+} \mathrm{r}(y; x_0) \leq \lim_{z \to x_1^+} \mathrm{r}(z; x_1) = f'_+(x_1) , \end{aligned}$$

come volevamo. □

Dimostrazione del Teorema 9. Consideriamo la funzione Φ definita nell'interno di I ponendo:

$$\Phi(x) := f'_+(x) - f'_-(x) ;$$

tale funzione è ovunque ≥ 0 ed è uguale a zero in tutti e soli i punti interni in cui f è derivabile. Conseguentemente, l'asserto rimane provato se riusciamo a mostrare che l'insieme $\Delta := \{x \in \operatorname{int} I : \ \Phi(x) > 0\}$ è al più numerabile.

`Passo 1`. Fissati $x_0 < x_1$ interni ad I, dalla Proposizione 5 segue che:

$$f'_-(x_0) \leq f'_+(x_0) \leq f'_-(x_1) \leq f'_+(x_1) ,$$

e, ragionando induttivamente, si vede che comunque si scelgano $n + 1$ punti $x_0 < x_1 < x_2 < \cdots < x_n$ interni ad I risulta:

$$f'_-(x_0) \leq f'_+(x_0) \leq f'_-(x_1) \leq f'_+(x_1) \leq f'_-(x_2) \leq f'_+(x_2) \leq \cdots \leq f'_-(x_n) \leq f'_+(x_n) ;$$

conseguentemente, gli intervalli $[f'_-(x_k), f'_+(x_k)]$ con $k = 0, 1, 2, \ldots, n$ formano una famiglia di sottoinsiemi di $[f'_-(x_0), f'_+(x_n)]$ costituita da intervalli aventi al più un estremo in comune e ciò implica che la somma delle loro ampiezze non supera l'ampiezza dell'intervallo in cui essi sono contenuti, i.e. che:

$$\sum_{k=0}^{n} \Phi(x_k) = \sum_{k=0}^{n} f'_+(x_k) - f'_-(x_k) \leq f'_+(x_n) - f'_-(x_0). \tag{14}$$

`Passo 2`. Fissiamo un intervallo $[\alpha,\beta]\subset \operatorname{int} I$ e, scelti arbitrariamente $n+1$ punti $x_0<x_1<x_2<\cdots<x_n\in[\alpha,\beta]$, per la (14) e per il Lemma 2 abbiamo:

$$\sum_{k=0}^{n}\Phi(x_k)\leq f'_+(x_n)-f'_-(x_0)\leq f'_+(\beta)-f'_-(\alpha)\;; \tag{15}$$

evidentemente, l'arbitrarietà nella scelta di $x_0,\dots,x_n$ implica che la maggiorazione (15) vale per qualsiasi famiglia di $n+1$ punti appartenenti ad $[\alpha,\beta]$.
Proviamo che la (15) assicura che l'insieme $\Delta_{[\alpha,\beta]}:=\{x\in[\alpha,\beta]:\ \Phi(x)>0\}$ è al più numerabile. Invero, per la (15) esistono al più un numero finito di punti $x_0^{(0)}<x_1^{(0)}<\cdots<x_{n_0-1}^{(0)}$ in cui risulta $\Phi(x)\geq 1$ ed il loro numero n_0 non può superare la parte intera di $f'_+(\beta)-f'_-(\alpha)$, poiché altrimenti si avrebbe:

$$\sum_{k=0}^{n_0-1}\Phi(x_k^{(0)})\geq\sum_{k=0}^{n_0-1}1=n_0\geq\lfloor f'_+(\beta)-f'_-(\alpha)\rfloor+1>f'_+(\beta)-f'_-(\alpha)$$

contro la (15); analogamente, esistono al più un numero finito di punti $x_0^{(1)}<x_1^{(1)}<\cdots<x_{n_1-1}^{(1)}$ in cui risulta $\frac{1}{10}\leq\Phi(x)<1$ ed il loro numero n_1 non può superare la parte intera di $10\big(f'_+(\beta)-f'_-(\alpha)\big)$, poiché altrimenti si avrebbe[12]:

$$\sum_{k=0}^{n_1-1}\Phi(x_k^{(1)})\geq\sum_{k=0}^{n_1-1}\frac{1}{10}=\frac{1}{10}n_1\geq\frac{1}{10}\lfloor 10\left(f'_+(\beta)-f'_-(\alpha)\right)\rfloor+\frac{1}{10}>f'_+(\beta)-f'_-(\alpha)$$

contro la (15);... Ragionando per ricorrenza, fissato $h\in\mathbb{N}$ con $h\geq 1$, esistono al più un numero finito di punti $x_0^{(h)}<x_1^{(h)}<\cdots<x_{n_h-1}^{(h)}$ in cui risulta $\frac{1}{10^h}\leq\Phi(x)<\frac{1}{10^{h-1}}$ ed il loro numero n_h non può superare la parte intera di $10^h\big(f'_+(\beta)-f'_-(\alpha)\big)$, poiché altrimenti si avrebbe:

$$\sum_{k=0}^{n_h-1}\Phi(x_k^{(h)})\geq\sum_{k=0}^{n_h-1}\frac{1}{10^h}=\frac{1}{10^h}n_h\geq\frac{1}{10^h}\lfloor 10^h\left(f'_+(\beta)-f'_-(\alpha)\right)\rfloor+\frac{1}{10^h}>f'_+(\beta)-f'_-(\alpha)$$

contro la (15). Dato che ogni $x\in\Delta_{[\alpha,\beta]}$ appartiene ad uno ed uno soltanto degli insiemi $\{x\in[\alpha,\beta]:\ \Phi(x)\geq 1\}$ od $\left\{x\in[\alpha,\beta]:\ \frac{1}{10^h}\leq\Phi(x)<\frac{1}{10^{h-1}}\right\}$ con $h\in\mathbb{N}$ e che, viceversa, ogni elemento degli insiemi $\{x\in[\alpha,\beta]:\ \Phi(x)\geq 1\}$ od $\left\{x\in[\alpha,\beta]:\ \frac{1}{10^h}\leq\Phi(x)<\frac{1}{10^{h-1}}\right\}$ con $h\in\mathbb{N}$ è certamente in $\Delta_{[\alpha,\beta]}$, si ha:

$$\Delta_{[\alpha,\beta]}=\{x\in[\alpha,\beta]:\ \Phi(x)\geq 1\}\cup\left(\bigcup_{h=1}^{\infty}\left\{x\in[\alpha,\beta]:\ \frac{1}{10^h}\leq\Phi(x)<\frac{1}{10^{h-1}}\right\}\right)\;;$$

e, visto che gli insiemi $\{x\in[\alpha,\beta]:\ \Phi(x)\geq 1\}$ ed $\left\{x\in[\alpha,\beta]:\ \frac{1}{10^h}\leq\Phi(x)<\frac{1}{10^{h-1}}\right\}$ con $h\in\mathbb{N}$ sono finiti e visto che l'unione numerabile di insiemi finiti è al più numerabile, è evidente che $\Delta_{[\alpha,\beta]}$ è al più numerabile.

`Passo 3`. Detti $a<b$ gli estremi dell'intervallo I, abbiamo $\operatorname{int} I=]a,b[$. Per noti fatti, possiamo determinare due successioni (α_n) e (β_n) di punti interni ad I tali che:

(1) $\alpha_{n+1}<\alpha_n$ e $\beta_n<\beta_{n+1}$ per ogni indice $n\in\mathbb{N}$,

(2) $\alpha_n\to a$ e $\beta_n\to b$,

[12] Si tenga presente che per ogni $x>0$ risulta $\lfloor x\rfloor\leq x<\lfloor x\rfloor+1$, dunque si ha anche $\lfloor 10^h x\rfloor\leq 10^h x<\lfloor 10^h x\rfloor+1$ per ogni $h\in\mathbb{N}$.

e ciò implica che risulta:

$$]a,b[=\bigcup_{n=0}^{\infty}[\alpha_n,\beta_n]\ .$$

Infatti, è evidente che $\bigcup_{n=0}^{\infty}[\alpha_n,\beta_n]\subseteq]a,b[$; d'altra parte, fissato arbitrariamente $x\in]a,b[$ esiste certamente (per definizione di limite) un indice $\nu\in\mathbb{N}$ tale che $a<\alpha_n<x<\beta_n<b$ per ogni $n>\nu$, cosicché $x\in\bigcup_{n=0}^{\infty}[\alpha_n,\beta_n]$ e perciò vale anche l'inclusione inversa $]a,b[\subseteq\bigcup_{n=0}^{\infty}[\alpha_n,\beta_n]$.
Quanto appena trovato implica che:

$$\Delta=\bigcup_{n=0}^{\infty}\Delta_{[\alpha_n,\beta_n]}\ ,$$

cosicché Δ è al più numerabile, poiché unione numerabile di insiemi al più numerabili. □

Un'altra curiosa conseguenza del Lemma 2 è che le derivate sinistra e destra di una funzione convessa sono integrabili nel senso di Riemann su ogni intervallo compatto contenuto in I, in quanto monotòne.
Ciò ha una gradevolissima conseguenza, cioè che f si può "ricostruire" usando una qualsiasi funzione integrale di tali derivate, ovvero (nel caso f sia derivabile) usando una qualsiasi funzione integrale della derivata prima. Vale infatti il seguente:

Teorema 10 (Teorema Fondamentale del Calcolo Integrale per Funzioni Convesse)
Siano $I\subseteq\mathbb{R}$ non banale ed $f:I\to\mathbb{R}$ una funzione convessa in I.
Le funzioni $f'_\pm:\operatorname{int}I\to\mathbb{R}$ sono integrabili sugli intervalli compatti contenuti in I e, per ogni $x_0\in I$, risulta:

$$f(x)=f(x_0)+\int_{x_0}^{x}f'_+(t)\ \mathrm{d}\,t=f(x_0)+\int_{x_0}^{x}f'_-(t)\ \mathrm{d}\,t$$

per ogni $x\in I$.
Dunque, se f è derivabile internamente ad I, l'uguaglianza:

$$f(x)=f(x_0)+\int_{x_0}^{x}f'(t)\ \mathrm{d}\,t \tag{16}$$

vale per ogni $x\in I$.

La dimostrazione è reperibile in [R, § 24].

Osservazione 37 (Maggiore Generalità della (16))**:** L'uguaglianza (16):

$$f(x)=f(x_0)+\int_{x_0}^{x}f'(t)\ \mathrm{d}\,t$$

ha validità più generale, poiché essa si estende anche alle funzioni convesse che non sono ovunque derivabili nel proprio intervallo di definizione (pur essendolo "quasi ovunque", cioè nel senso del Teorema 9).
Tale estensione si basa su proprietà più o meno fini dell'integrale di Riemann e delle funzioni di variabile reale (come la (AC) dell'`Esercizio 8`) che non sono affrontate nei corsi di base. Pertanto, non potendo entrare nel dettaglio della dimostrazione, in questa sede dobbiamo accontentarci di avere accennato a tale proprietà.
Il lettore interessato allo studio della questione può fare riferimento a testi specialistici, come [R, W] (in cui la teoria di base della convessità è affrontata globalmente, anche sotto l'aspetto geometrico e computazionale) oppure [L]. ♦

Riferimenti bibliografici

[DM14] Di Meglio, G. (2017) *Alcune Applicazioni del Calcolo Integrale.*

[L] Leoni, G. (2009), **A First Course in Sobolev Spaces**, Graduate Studies in Mathematics, vol. 105, American Mathematical Society, Providence.

[MS] Marcellini, P. & Sbordone, C. (1998), **Analisi Matematica Uno**, Liguori, Napoli.

[R] Rockafellar, R. T. (1970), **Convex Analysis**, Princeton University Press, Princeton.

[W] Webster, R. (1994), **Convexity**, Oxford University Press, New York.

Guglielmo Di Meglio, PhD
Scuola Politecnica e delle Scienze di Base
Università degli Studi di Napoli "Federico II"
piazzale Tecchio 80
80126 Napoli – ITALY
email: guglielmo.dimeglio@unina.it