BREVE NOTA SU ALCUNI ARGOMENTI DEL CORSO ***MATEMATICA CORSO BASE***

Le dimostrazioni riportate sul testo di riferimento non verranno ripetute in questa nota

Ottobre 2020

A cura di

Stefano Patrì
Facoltà di Economia
Dipartimento MEMOTEF

Anno accademico 2020/2021

Indice

Capitolo 1

Cenni sugli insiemi numerici

Il primo insieme numerico che ogni individuo da bambino impara ad utilizzare è l'insieme dei numeri $0, 1, 2, 3, \ldots$ chiamato *insieme dei numeri naturali* ed indicato con il simbolo $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

1.1 L'insieme $\mathbb{N}$ dei numeri naturali

L'insieme dei numeri naturali è costituito dagli elementi $0, 1, 2, 3, \ldots$ il cui significato è quello "naturale" che ogni individuo apprende sin da bambino e che quindi non richiede ulteriori spiegazioni. Data una qualsiasi coppia a, b di numeri naturali, si può eseguire, com'è noto, l'*addizione* $a + b$ e la *moltiplicazione* $a \cdot b$. L'addizione gode delle tre proprietà

- *associativa*, ovvero per ogni tripla a, b, c di numeri naturali risulta
$$(a + b) + c = a + (b + c);$$
- *commutativa*, ovvero per ogni coppia a, b di numeri naturali risulta
$$a + b = b + a;$$
- *esistenza dell'elemento neutro*, ovvero per ogni numero naturale a risulta
$$a + 0 = 0 + a = a,$$
in modo tale che il numero naturale 0 prende il nome di *elemento neutro dell'addizione.*

La moltiplicazione gode delle tre proprietà

- *associativa*, ovvero per ogni tripla a, b, c di numeri naturali risulta
$$(a \cdot b) \cdot c = a \cdot (b \cdot c);$$
- *commutativa*, ovvero per ogni coppia a, b di numeri naturali risulta
$$a \cdot b = b \cdot a;$$
- *esistenza dell'elemento neutro*, ovvero per ogni numero naturale a risulta
$$a \cdot 1 = 1 \cdot a = a,$$
in modo tale che il numero naturale 1 prende il nome di *elemento neutro della moltiplicazione.*

Osserviamo infine che le due operazioni di *addizione* e *moltiplicazione* sono connesse da quella che prende il nome di *proprietà distributiva*, ovvero per ogni quadrupla di numeri naturali a, b, c, d vale la relazione

$$(a + b) \cdot (c + d) = (a \cdot c) + (a \cdot d) + (b \cdot c) + (b \cdot d).$$

Se però a tali proprietà aggiungessimo le richieste che

- per ogni numero naturale a esista un unico numero naturale b, chiamato *elemento opposto* di a rispetto all'operazione di *addizione*, tale che risulti $a + b = 0$,
- per ogni numero naturale $a \neq 0$ esista un unico numero naturale b, chiamato *elemento inverso* di a rispetto all'operazione di *moltiplicazione*, tale che valga la relazione $a \cdot b = 1$,

dobbiamo concludere che nell'insieme $\mathbb{N}$ dei numeri naturali non esistono gli *elementi opposti* dei numeri $a > 0$ e non esistono gli *elementi inversi* dei numeri $a > 1$.

Per fare in modo che ogni numero abbia il proprio *elemento opposto* rispetto all'*addizione* e che ogni numero diverso da zero abbia il proprio *elemento inverso* rispetto alla *moltiplicazione*, occorre dunque ampliare l'insieme dei numeri naturali aggiungendo ai numeri naturali dei "nuovi" numeri che siano gli elementi opposti e gli elementi inversi.

E' importante sottolineare che quando un insieme numerico "più piccolo" $\mathcal{I}$ viene ampliato in un insieme "più grande" $\mathcal{J}$, l'insieme "più grande" $\mathcal{J}$ deve contenere tutti i numeri dell'insieme "più piccolo" $\mathcal{I}$ e i "nuovi" numeri in modo tale che tutte le proprietà dell'insieme "più piccolo" $\mathcal{I}$ continuino a valere ed inoltre sia soddisfatta una certa proprietà "nuova" che i numeri dell'insieme "più piccolo" $\mathcal{I}$ "da soli" non soddisfano.

1.2 L'insieme $\mathbb{Z}$ dei numeri interi

Il primo ampliamento dell'insieme $\mathbb{N}$ che eseguiamo è quello necessario affinché ogni numero abbia il proprio elemento opposto rispetto all'addizione. Se per ogni coppia di numeri naturali $a, b > 0$ creiamo i "nuovi" numeri indicati con $-a, -b$ tale che risulti

- $a + (-a) = 0$,
- $-a > -b, \quad \text{se } a < b$,

otteniamo l'*insieme dei numeri interi* indicato con $\mathbb{Z}$ i cui elementi in ordine crescente sono

$$\mathbb{Z} = \{\dots, -4, -3, -2, -1, 0, 1, 2, 3, 4 \dots\}.$$

L'insieme $\mathbb{Z}$ dei *numeri interi* contiene i numeri naturali $0, 1, 2, 3, \dots$ e i "nuovi" numeri che sono gli opposti dei numeri interi positivi. Nell'insieme $\mathbb{Z}$ dei *numeri interi* l'operazione di *addizione* gode delle tre proprietà riportate nel paragrafo precedente e della quarta proprietà che è l'esistenza dell'elemento opposto di ogni elemento $a \in \mathbb{Z}$.

E' immediato rendersi conto che però rispetto all'operazione di *moltiplicazione* nemmeno i numeri dell'insieme $\mathbb{Z}$ possiedono i propri inversi. Affinché i numeri diversi da zero abbiano l'inverso rispetto alla moltiplicazione, occorre dunque ampliare l'insieme $\mathbb{Z}$ dei numeri interi in un insieme "più grande" che è l'insieme dei *numeri razionali*.

1.3 L'insieme $\mathbb{Q}$ dei numeri razionali

L'insieme $\mathbb{Q}$ dei *numeri razionali* è l'insieme contenente tutte le possibili *frazioni* con numeratore e denominatore numeri interi tali che il denominatore sia diverso da zero, ovvero abbiamo

$$\mathbb{Q} = \left\{ \frac{m}{n}, \ \text{tale che } m, n \in \mathbb{Z} \ \text{e} \ n \neq 0 \right\},$$

dove il numero intero m viene denominato *numeratore* e il numero intero n viene denominato *denominatore* della *frazione*.

Poiché le frazioni sono dunque il rapporto tra due numeri interi e la parola *rapporto*, in latino, veniva detta *ratio*, l'insieme delle *frazioni* prende il nome di insieme dei *numeri razionali*, ovvero *frazioni* e *numeri razionali* sono il medesimo concetto. Tutte le "usuali" operazioni con le frazioni discendono dal significato che, com'è noto, possiede il simbolo

$$\frac{m}{n},$$

che spesso, per motivi puramente estetici, indicheremo anche con la scrittura m/n.

In generale il significato di un "nuovo" simbolo deve essere spiegato sempre tramite un'operazione già nota che lo lega ad un numero già noto. In particolare il significato del simbolo m/n dei numeri dell'insieme $\mathbb{Q}$ è stabilito dalla sua moltiplicazione, operazione già nota, per un numero intero già noto. Il simbolo $m/n \in \mathbb{Q}$ rappresenta dunque quel numero che soddisfa la relazione

$$\frac{m}{n} \cdot n = n \cdot \frac{m}{n} = m. \tag{1.1}$$

Dalla definizione (1.1) discendono le "regole" dell'addizione e della moltiplicazione tra numeri razionali, le quali si ottengono richiedendo che tali operazioni siano "coerenti" con le proprietà dell'addizone e della moltiplicazione illustrate nel paragrafo 1.1.

Introducendo prima quella della moltiplicazione", abbiamo che la "regola" della moltiplicazione di due generiche frazioni $a/b, c/d \in \mathbb{Q}$ è data dalla relazione

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d} \tag{1.2a}$$

e dimostriamo che essa risulta coerente con le proprietà della moltiplicazione.

Dimostrare che la "regola" (1.2a) è coerente con le proprietà della moltiplicazione significa dimostrare che se applichiamo la definizione (1.1) ad ambo i membri della relazione (1.2a) con la moltiplicazione per il medesimo numero intero, otteniamo il medesimo risultato, ovvero i due membri della (1.2a) sono appunto coerenti rispetto alla definizione (1.1). Se moltiplichiamo per il numero intero $b \cdot d$ il primo membro della (1.2a) applicando le proprietà commutativa e associativa della moltiplicazione ed utilizzando la relazione (1.1) che definisce il significato del simbolo di frazione, ricaviamo

$$\left(\frac{a}{b} \cdot \frac{c}{d} \right) \cdot (b \cdot d) = \left(\frac{a}{b} \cdot b \right) \cdot \left(\frac{c}{d} \cdot d \right) = a \cdot c.$$

Se quindi moltiplichiamo per il numero intero $b \cdot d$ anche il secondo membro della (1.2a) utilizzando la relazione (1.1) che definisce il significato del simbolo di frazione, ricaviamo

$$\frac{a \cdot c}{b \cdot d} \cdot (b \cdot d) = a \cdot c.$$

Poiché dunque la moltiplicazione dei due membri della (1.2a) per il medesimo numero intero $b \cdot d$ fornisce il medesimo risultato, concludiamo che la relazione (1.2a) è la "regola" per la moltiplicazione di due frazioni, ovvero di due numeri razionali.

La "regola" per l'addizione di due numeri razionali, o frazioni, è data dalla relazione

$$\frac{a}{b} + \frac{c}{d} = \frac{(a \cdot d) + (b \cdot c)}{b \cdot d}, \tag{1.2b}$$

che si riconosce immediatamente essere l'usuale *regola del minimo comune denominatore*.

Per dimostrare che tale "regola" risulta coerente con le proprietà dell'addizione e della moltiplicazione, moltiplichiamo ambo i membri della (1.2b) per il numero intero $b \cdot d$.

Con l'applicazione delle proprietà distributiva, associativa e commutativa ed utilizzando la definizione (1.1) del simbolo di frazione, otteniamo dal primo membro

$$\left(\frac{a}{b} + \frac{c}{d}\right) \cdot (b \cdot d) = \left(\frac{a}{b} \cdot b \cdot d\right) + \left(\frac{c}{d} \cdot d \cdot b\right) = \left[\left(\frac{a}{b} \cdot b\right) \cdot d\right] + \left[\left(\frac{c}{d} \cdot d\right) \cdot b\right] =$$

$$= (a \cdot d) + (b \cdot c)$$

e dal secondo membro

$$\frac{(a \cdot d) + (b \cdot c)}{b \cdot d} \cdot (b \cdot d) = (a \cdot d) + (b \cdot c).$$

Poiché dunque il prodotto di ambo i membri della relazione (1.2b) per il medesimo numero intero $b \cdot d$ fornisce il medesimo risultato, concludiamo che la (1.2b) è la regola per l'addizione di due frazioni, ovvero di due numeri razionali.

Tutte le frazioni del tipo

$$\frac{0}{n}$$

con n qualsiasi diverso da zero, sono uguali a zero e tutte le frazioni del tipo

$$\frac{n}{n}$$

con n qualsiasi diverso da zero, sono uguali a 1.

Infine osserviamo che con i numeri razionali, ovvero con le frazioni, anche l'operazione di moltiplicazione è dotata della proprietà di esistenza dei numeri inversi di tutti i numeri diversi da zero, perché ogni frazione m/n diversa da zero ha come inversa la frazione diversa da zero n/m tale che risulti

$$\frac{m}{n} \cdot \frac{n}{m} = \frac{m \cdot n}{m \cdot n} = 1.$$

L'insieme dei numeri razionali contiene ovviamente l'insieme dei numeri interi perché per ogni numero intero $m \in \mathbb{Z}$, risulta

$$m = \frac{m}{1}.$$

1.3.1 Periodicità di un numero decimale

Ad ogni frazione m/n può essere associato il numero decimale che si ottiene dividendo il numeratore m per il denominatore n e si dice che la *frazione genera* il numero decimale, oppure che il numero decimale *è generato* dalla *frazione*. Ovviamente se il numeratore è multiplo del denominatore, il numero associato alla frazione risulta intero. Un numero decimale della forma

$$a_1a_2a_3\cdots a_h\,,\,(b_1b_2b_3\cdots b_k)\,[p_1p_2p_3\cdots p_q]\,[p_1p_2p_3\cdots p_q]\,[p_1p_2p_3\cdots p_q]\cdots$$

prende il nome di *numero decimale periodico* oppure semplicemente di *numero periodico*, dove $a_1a_2a_3\cdots a_h$ è la *parte intera* costituita da h cifre, le k cifre $b_1b_2b_3\cdots b_k$ costituiscono quello che si chiama *antiperiodo* e il “blocco” $p_1p_2p_3\cdots p_q$ con q cifre si ripete all'infinito prendendo il nome di *periodo*. In genere un *numero periodico* viene indicato tramite la seguente notazione

$$a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_k\,[p_1p_2p_3\cdots p_q]\,,$$

in cui il *periodo* è racchiuso in parentesi quadra ed è sottinteso che tale “blocco” si ripete identico all'infinito. Nel caso in cui il *periodo* sia costituito dalla sola cifra zero oppure dalla sola cifra 9, abbiamo le seguenti identificazioni

$$\begin{aligned} a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_k[0] = a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_k\,00000\cdots \equiv \\ \equiv a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_k \end{aligned} \tag{1.3a}$$

e

$$\begin{aligned} a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_{k-1}b_k[9] = a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_{k-1}b_k\,99999\cdots \equiv \\ \equiv a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_{k-1}(b_k+1). \end{aligned} \tag{1.3b}$$

Ad esempio abbiamo $27,358000000\cdots \equiv 27,358$ e $34,7651999999\cdots \equiv 34,7652$ ed il motivo di tali identificazioni è che, com'è immediato rendersi conto, non esiste nessun numero compreso tra i due numeri

$$27,358000000\cdots \qquad \text{e} \qquad 27,358$$

né tra i due numeri

$$34,7651999999\cdots \qquad \text{e} \qquad 34,7652.$$

Nel caso in cui un numero sia *decimale periodico* con la sola cifra periodica zero oppure con la sola cifra periodica 9, le cifre periodiche vengono eliminate mediante le prescrizioni espresse dalle identificazioni (1.3) e il *numero decimale periodico*, ritrovandosi così ad avere un numero finito di cifre decimali, viene denominato *numero decimale limitato*.

Osserviamo che una *frazione genera* un *numero decimale limitato* se e solo se nella scomposizione in fattori primi del denominatore vi sono solo i *fattori primi* 2 e/o 5.

A questo punto ci si rende conto “facilmente” che tutte le *frazioni*, tramite divisione del numeratore per il denominatore, forniscono sempre, in virtù della proprietà del *resto* di essere minore del divisore, un numero appartenente ad una delle seguenti tre “categorie”

- *numero intero*, se il numeratore è multiplo del denominatore;
- *numero decimale limitato*, se la scomposizione in fattori primi del denominatore contiene solo i fattori primi 2 e 5, oppure uno solo di essi;
- *numero decimale illimitato periodico*, se la scomposizione in fattori primi del denominatore contiene almeno un fattore primo diverso da 2 e da 5.

Viceversa, dato un *numero decimale periodico*, la frazione, detta *frazione generatrice*, che, mediante divisione del numeratore per il denominatore, genera tale numero periodico si ottiene dalla relazione (di cui omettiamo le semplice spiegazione illustrata in aula)

$$a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_k\,[p_1p_2p_3\cdots p_q] =$$

$$= \frac{(a_1a_2a_3\cdots a_h\,b_1b_2b_3\cdots b_k\,p_1p_2p_3\cdots p_q) - (a_1a_2a_3\cdots a_h\,b_1b_2b_3\cdots b_k)}{(999\cdots q\,\text{volte})(000\cdots k\,\text{volte})}\,.$$

Ad esempio il numero decimale $7,52144444\cdots \equiv 7,521[4]$ possiede $k = 3$ cifre di *antiperiodo* ed una sola cifra ($q = 1$) periodica. Pertanto al denominatore scriveremo un solo 9, perché appunto $q = 1$, e tre zeri, perché appunto $k = 3$, in modo tale che il numero periodico è generato dalla frazione (eventualmente semplificabile)

$$7,52144444\cdots \equiv 7,521[4] = \frac{75214 - 7521}{9000} = \frac{67693}{9000}\,.$$

Il numero decimale $23,49\underbrace{871}\underbrace{871}\underbrace{871}\cdots \equiv 23,49[871]$ possiede $k = 2$ cifre di *antiperiodo* e $q = 3$ cifre periodiche. Pertanto al denominatore scriveremo tre cifre 9, perché appunto $q = 3$, e due zeri, perché appunto $k = 2$, in modo tale che il numero periodico è generato dalla frazione (eventualmente semplificabile)

$$23,49\underbrace{871}\underbrace{871}\underbrace{871}\cdots \equiv 23,49[871] = \frac{2349871 - 2349}{99900} = \frac{2347522}{99900}\,.$$

Nel caso di un *numero decimale limitato*, associato tramite le identificazioni (1.3) ad un *numero periodico* avente la sola cifra periodica zero oppure la sola cifra periodica 9, scriviamo il *numero decimale limitato* nella forma $a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_{k-1}b_k$ e la sua *frazione generatrice* è

$$a_1a_2a_3\cdots a_h\,,\,b_1b_2b_3\cdots b_{k-1}b_k = \frac{a_1a_2a_3\cdots a_h\,b_1b_2b_3\cdots b_{k-1}b_k}{10^k}\,.$$

Ad esempio i *numeri decimali limitati* $3,14$ e $23,457$ si ottengono dalle *frazioni generatrici*

$$3,14 = \frac{314}{100} \qquad \text{e} \qquad 23,457 = \frac{23457}{1000}\,.$$

In conclusione possiamo dire che *numeri razionali*, *frazioni* e *numeri decimali periodici* sono nomi equivalenti per denominare gli elementi dell'insieme $\mathbb{Q}$.

1.3.2 Lacuna dei numeri razionali

Nell'insieme dei numeri razionali sono definite dunque le due operazioni di addizione e moltiplicazione che godono ciascuna delle quattro proprietà: commutativa, associativa, esistenza dell'elemento neutro ed esistenza degli opposti e degli inversi (l'inverso rispetto alla moltiplicazione esiste solo per i numeri diversi da zero). Inoltre le due operazioni sono connesse dalla proprietà distributiva. Per costruire una teoria matematica qual è, ad esempio, l'*Analisi Matematica*, occorre che l'insieme numerico che si utilizza possieda un'ulteriore proprietà che però l'insieme $\mathbb{Q}$ dei numeri razionali non possiede e che viene denominata *proprietà di continuità*. Dato un generico insieme numerico I, due suoi sottoinsiemi $\mathcal{S}_1$ e $\mathcal{S}_2$ si dicono *separati* se ogni numero di $\mathcal{S}_1$ è minore di tutti i numeri dell'insieme $\mathcal{S}_2$ e ogni numero di $\mathcal{S}_2$ è maggiore di tutti i numeri dell'insieme $\mathcal{S}_1$.

Un generico insieme numerico I possiede la *proprietà di continuità* se, considerati due suoi generici sottoinsiemi $\mathcal{S}_1$ e $\mathcal{S}_2$ che siano *separati*, esiste sempre un numero di I, indicato con s e denominato *numero separatore*, oppure semplicemente *separatore*, di $\mathcal{S}_1$ e $\mathcal{S}_2$, tale che nessun numero di $\mathcal{S}_1$ sia maggiore di s e nessun numero di $\mathcal{S}_2$ sia minore di s.

L'insieme $\mathbb{Q}$ dei numeri razionali non possiede la *proprietà di continuità* perché, considerati due sottoinsiemi *separati* $\mathcal{S}_1$ e $\mathcal{S}_2$ dell'insieme $\mathbb{Q}$, non esiste sempre un numero razionale che sia *separatore* di $\mathcal{S}_1$ e $\mathcal{S}_2$. Se consideriamo i due sottoinsiemi dell'insieme $\mathbb{Q}$

$$\mathcal{S}_1 = \left\{\frac{m}{n} \in \mathbb{Q},\ \text{tali che } \frac{m}{n} < 1\right\} \qquad \text{e} \qquad \mathcal{S}_2 = \left\{\frac{m}{n} \in \mathbb{Q},\ \text{tali che } \frac{m}{n} > 4\right\},$$

abbiamo che $\mathcal{S}_1$ e $\mathcal{S}_2$ sono "ovviamente" *separati* ed un loro *separatore* è, tra gli altri possibili, il numero razionale $s = 2$, perché $s = 2$ risulta simultaneamente maggiore di ciascun numero di $\mathcal{S}_1$ e minore di ciascun numero di $\mathcal{S}_2$.

Se però consideriamo i due sottoinsiemi dell'insieme $\mathbb{Q}$

$$\mathcal{S}_1 = \left\{\frac{m}{n} > 0,\ \text{tali che } \left(\frac{m}{n}\right)^2 < 2\right\} \qquad \text{e} \qquad \mathcal{S}_2 = \left\{\frac{m}{n} > 0,\ \text{tali che } \left(\frac{m}{n}\right)^2 > 2\right\}, \tag{1.4}$$

abbiamo che $\mathcal{S}_1$ e $\mathcal{S}_2$ sono "ovviamente" *separati* e dimostriamo che nell'insieme $\mathbb{Q}$ dei numeri razionali non esiste nessun *separatore* s dei due sottoinsiemi $\mathcal{S}_1, \mathcal{S}_2$.

A tale scopo dimostriamo innanzitutto che nessuna frazione positiva m/n ridotta ai minimi termini ha il quadrato uguale a 2. Dopo aver ricordato che una frazione ridotta ai minimi termini può avere solo numeratore e denominatore entrambi dispari oppure uno pari e l'altro dispari (perché una frazione con numeratore e denominatore entrambi pari non è ridotta ai minimi termini), dimostriamo che l'uguaglianza

$$\left(\frac{m}{n}\right)^2 = 2 \tag{1.5}$$

non è mai verificata per nessuna frazione m/n ridotta ai minimi termini.

Se infatti riscriviamo l'uguaglianza (1.5) nella forma $m^2 = 2n^2$, deduciamo che il quadrato m^2 è pari, in quanto doppio di un altro numero, da cui segue quindi che anche il numeratore m della frazione m/n è un numero pari e poiché la frazione m/n è ridotta ai minimi termini, ricaviamo che il denominatore n deve essere un numero dispari. Avremo allora dimostrato che nessuna frazione positiva m/n ridotta ai minimi termini può avere quadrato uguale a 2 se mostriamo che dal numeratore m pari discende che anche il denominatore n è pari. Se mostriamo che dal numeratore m pari discende che anche il denominatore n è pari, possiamo concludere dunque che non esiste nessuna frazione ridotta ai minimi il cui quadrato sia uguale a 2 e che quindi nessun'altra frazione può avere quadrato uguale a 2, altrimenti questa sarebbe poi riducibile ai minimi termini.

Per dimostrare che dal numeratore m pari discende che anche il denominatore n è pari, poniamo $m = 2k$, con k numero intero positivo, da cui segue dunque che l'uguaglianza $m^2 = 2n^2$ assume la forma $4k^2 = 2n^2$, ovvero, se semplifichiamo, $n^2 = 2k^2$, che mostra che il numero n^2 è pari e che quindi anche n è appunto un numero pari.

Poiché nessuna frazione possiede il quadrato uguale a 2, segue che il quadrato di ogni frazione positiva risulta minore di 2 oppure maggiore di 2, ovvero che ogni frazione positiva appartiene all'insieme $\mathcal{S}_1$ oppure all'insieme $\mathcal{S}_2$. A questo punto, per dimostrare che nell'insieme dei numeri razionali non esiste nessun numero s *separatore* dei due sottoinsiemi $\mathcal{S}_1$ e $\mathcal{S}_2$, basta dimostrare che qualunque numero razionale s si prenda, esiste sempre una frazione di $\mathcal{S}_1$ maggiore di s se $s \in \mathcal{S}_1$ ed esiste sempre una frazione di $\mathcal{S}_2$ minore di s se $s \in \mathcal{S}_2$. Infatti, se la frazione s considerata ha il quadrato minore di 2, ovvero $s^2 < 2$ e quindi s appartiene al sottoinsieme $\mathcal{S}_1$, segue che anche la frazione

$$s + \frac{2 - s^2}{2 + s},$$

che è maggiore di s in quanto $2 - s^2 > 0$, ha il quadrato minore di 2 ed appartiene pertanto anch'essa al sottoinsieme $\mathcal{S}_1$, come si riconosce sviluppando il quadrato per $s^2 < 2$

$$\left(s + \frac{2 - s^2}{2 + s}\right)^2 = \left(\frac{2s + 2}{s + 2}\right)^2 = 2\left[\frac{2s^2 + 4s + 2}{s^2 + 4s + 4}\right] < 2,$$

dove la frazione in parentesi quadra è minore di 1 perché la differenza negativa

$$(2s^2 + 4s + 2) - (s^2 + 4s + 4) = s^2 - 2 < 0$$

mostra che nella frazione in parentesi quadra il denominatore è maggiore del numeratore.

Se invece la frazione s considerata ha il quadrato maggiore di 2, ovvero $s^2 > 2$ e quindi la frazione s appartiene al sottoinsieme $\mathcal{S}_2$, segue che anche la frazione

$$s - \frac{s^2 - 2}{2 + s},$$

che è minore di s in quanto $s^2 - 2 > 0$, ha il quadrato maggiore di 2 ed appartiene pertanto anch'essa al sottoinsieme $\mathcal{S}_2$, come si riconosce sviluppando il quadrato per $s^2 > 2$

$$\left(s - \frac{s^2 - 2}{2 + s}\right)^2 = \left(\frac{2s + 2}{s + 2}\right)^2 = 2\left[\frac{2s^2 + 4s + 2}{s^2 + 4s + 4}\right] > 2,$$

dove stavolta la frazione in parentesi quadra è maggiore di 1 perché la differenza positiva

$$(2s^2 + 4s + 2) - (s^2 + 4s + 4) = s^2 - 2 > 0$$

mostra che nella frazione in parentesi quadra il denominatore è minore del numeratore.

Poiché dunque per ogni *frazione* $s \in \mathcal{S}_1$ vi sono numeri di $\mathcal{S}_1$ a sinistra e a destra del numero s e analogamente per ogni *frazione* $s \in \mathcal{S}_2$ vi sono numeri di $\mathcal{S}_2$ a sinistra e a destra del numero s, concludiamo che per i due sottoinsiemi *separati* (1.4) dell'*insieme dei numeri razionali* non esiste un numero razionale *separatore*, ovvero che nell'*insieme dei numeri razionali* non esiste sempre il *separatore* tra due suoi sottoinsiemi *separati*.

1.4 L'insieme $\mathbb{R}$ dei numeri reali

Per ottenere un insieme numerico che contenga sempre il numero *separatore* tra due suoi qualsiasi sottoinsiemi *separati* e con gli elementi del quale si possano eseguire l'addizione e la moltiplicazione che soddisfino tutte le proprietà già più volte citate, dobbiamo effettuare un ampliamento dell'insieme dei numeri razionali e costruire l'insieme numerico denominato *insieme dei numeri reali*, indicato con il simbolo $\mathbb{R}$, a cui appartengano i numeri razionali e "nuovi" numeri che non sono razionali e che, in virtù di quanto illustrato riguardo alla *lacuna dei numeri razionali*, costituiscano dunque i *separatori* di qualsiasi coppia di sottoinsiemi *separati* che contengano solo numeri razionali, ma che siano privi, nell'insieme dei numeri razionali stessi, appunto del numero *separatore*.

Poiché per ottenere l'*insieme dei numeri reali* dobbiamo aggiungere ai numeri razionali "nuovi" numeri che non sono razionali, questi "nuovi" numeri non razionali vengono denominati *numeri irrazionali* e il problema che si pone a questo punto è quindi quello del simbolo con cui si possano rappresentare i *numeri irrazionali.*

Innanzitutto possiamo dire che la rappresentazione decimale di un *numero irrazionale* è diversa dalle tre "categorie" riportate nel paragrafo 1.3.1 associate alle sole *frazioni*, ovvero un *numero irrazionale* non è un *numero intero*, né un *numero decimale limitato*, né un *numero illimitato periodico*, e quindi abbiamo che la rappresentazione decimale di un *numero irrazionale* è costituita da un numero *decimale, illimitato e non periodico.*

Poiché però non ha alcun senso scrivere un numero con infinite cifre decimali, né d'altra parte è possibile rappresentare come *frazione* un numero che *frazione* non è, segue che in generale un *numero irrazionale* non può essere rappresentato mediante simboli "già noti" con le cifre "usuali". In virtù di tali considerazioni, il problema della rappresentazione di un *numero irrazionale* si risolve dunque semplicemente dicendo che non tutti i *numeri irrazionali* possiedono una rappresentazione e che ad un *numero irrazionale* si attribuisce un simbolo solo se tale *numero irrazionale* "riveste un'importanza" in Matematica oppure nelle varie Scienze che utilizzano la Matematica come strumento di studio. Ad esempio il *numero irrazionale* che *separa* i due sottoinsiemi $\mathcal{S}_1$ e $\mathcal{S}_2$ riportati nella (1.4) è "molto importante" in Matematica e allora gli si attribuisce un simbolo, ovvero il simbolo $\sqrt{2}$.

Il numero che si ottiene dividendo la lunghezza di una qualsiasi circonferenza per la lunghezza del proprio diametro è un numero che è stato dimostrato essere un *numero irrazionale* e poiché esso è "importantissimo" in Matematica, segue che gli si attribuisce un simbolo ed in particolare la lettera π, detta *pi greco*, appunto dell'alfabeto greco.

Il numero decimale illimitato costruito scrivendo un qualsiasi numero intero, quale ad esempio il numero 5, seguito da infinite cifre decimali secondo la seguente sequenza: la cifra 0 una sola volta, poi la cifra 1 due volte, poi la cifra zero tre volte, poi la cifra 1 quattro volte, poi la cifra zero cinque volte, poi la cifra 1 sei volte e cosí via all'infinito sempre con l'alternanza dei due "blocchi" di zeri e di 1 via via "crescenti", ovvero

$$5,\underbrace{0}\underbrace{11}\underbrace{000}\underbrace{1111}\underbrace{00000}\underbrace{111111}....$$

è ovviamente un *numero irrazionale* perché il continuo aumentare alternato del numero delle cifre zero e 1 fa sì che nella parte decimale di tale numero non vi sia nessun "blocco" di cifre che si ripete immutato appunto periodicamente all'infinito.

Poiché però tale *numero irrazionale* non ha "nessun significato" particolare e "nessuna importanza" in Matematica, segue che a tale numero non si attribuisce nessun simbolo.

E' ovvio, ma tuttavia utile, sottolineare infine che, quando per indicare un *numero irrazionale* si scrive la sua parte intera seguita da un certo numero "finito" di cifre decimali, il numero che si scrive in questo modo al posto del *numero irrazionale* è un *numero decimale limitato*, ovvero una *frazione*, che può essere utilizzato come *approssimazione* del *numero irrazionale*, ma non coincide con il *numero irrazionale.*

Poiché nella "nostra" Matematica "non sono ammesse" le approssimazioni, segue che i *numeri irrazionali* debbono essere indicati esclusivamente con il loro simbolo e mai mediante la loro approssimazione con un numero "finito" di cifre decimali.

Capitolo 2

Elementi di logica delle proposizioni

Una teoria matematica è costruita in generale a partire dall'introduzione di un insieme di oggetti di cui si vogliono studiare le proprietà. Ad esempio, nella teoria della Geometria di Euclide (III sec. a.C.) viene introdotto l'insieme delle figure geometriche, quali angoli, poligoni, circonferenze, cerchi, e si studiano le proprietà, tra le quali ad esempio il *teorema di Pitagora*, di cui appunto le figure geometriche godono.

Una proprietà degli oggetti appartenenti all'insieme introdotto per la teoria che si vuole studiare è stabilita da quello che prende appunto il nome di *teorema*.

Un *teorema* stabilisce, per così dire, una *connessione logica* tra due *proposizioni matematiche*, indicate con $\mathcal{P}$ e $\mathcal{Q}$, dove con *connessione logica* si intende che dalla validità di una *proposizione*, ad esempio $\mathcal{P}$, segue la validità dell'altra *proposizione* che è la $\mathcal{Q}$.

Un *teorema* afferma, in altre parole, che se la *proposizione* $\mathcal{P}$ è vera, segue, come *conseguenza logica*, che anche la *proposizione* $\mathcal{Q}$ risulta vera.

A questo punto occorre quindi illustrare alcuni concetti basilari della *logica* delle *proposizioni*. In generale, una *proposizione* è un'affermazione della quale si possa stabilire univocamente e oggettivamente se essa sia vera oppure sia falsa, senza che il giudizio dipenda da un'opinione personale.

2.1 Esempi di proposizione

Per illustrare meglio il concetto di *proposizione*, consideriamo le seguenti *affermazioni*

- $\mathcal{P}_1$: "*il numero 4 è divisibile per* 2",
- $\mathcal{P}_2$: "*il numero 6 non è divisibile per* 2 ",
- $\mathcal{P}_3$: "*il coccodrillo è un mammifero*",
- $\mathcal{P}_4$: "*tutti i gatti sono capaci di arrampicarsi sugli alberi*"

e osserviamo che esse sono *proposizioni* perché è possibile stabilire univocamente e oggettivamente se ciascuna di esse sia vera oppure sia falsa, senza che il giudizio dipenda dall'opinione personale di qualcuno. In particolare abbiamo che la *proposizione* $\mathcal{P}_1$ è vera, le *proposizioni* $\mathcal{P}_2, \mathcal{P}_3$ sono false, la *proposizione* $\mathcal{P}_4$ è vera e tali giudizi sono i medesimi "per tutti". Data una *proposizione* $\mathcal{P}$, chiamiamo *proposizione contraria* della $\mathcal{P}$ quella *proposizione*, indicata con il simbolo $\overline{\mathcal{P}}$, che esprime l'affermazione contraria della $\mathcal{P}$.

E' immediato rendersi conto che le *proposizioni contrarie* delle quattro *proposizioni* precedenti $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_4$ sono rispettivamente

- $\overline{\mathcal{P}}_1$: "*il numero 4 non è divisibile per* 2",
- $\overline{\mathcal{P}}_2$: "*il numero 6 è divisibile per* 2",
- $\overline{\mathcal{P}}_3$: "*il coccodrillo non è un mammifero*",
- $\overline{\mathcal{P}}_4$: "*esiste un gatto che non è capace di arrampicarsi sugli alberi*".

E' immediato rendersi conto che se una *proposizione* $\mathcal{P}$ è vera, la sua *proposizione contraria* $\overline{\mathcal{P}}$ è falsa, mentre, viceversa, se una *proposizione* $\mathcal{P}$ è falsa, la sua *proposizione contraria* $\overline{\mathcal{P}}$ è vera e possiamo verificare infatti che la $\overline{\mathcal{P}}_1$ è falsa, la $\overline{\mathcal{P}}_2$ e la $\overline{\mathcal{P}}_3$ sono vere, la $\overline{\mathcal{P}}_4$ è falsa. Dal confronto tra la *proposizione* $\mathcal{P}_4$ e la sua *proposizione contraria* $\overline{\mathcal{P}}_4$ emerge una caratteristica che è di fondamentale importanza puntualizzare, ovvero che la *proposizione contraria* $\overline{\mathcal{P}}$ di una *proposizione* $\mathcal{P}$ del tipo

$$\mathcal{P} : \text{“}\textit{tutti gli elementi di un certo insieme possiedono una certa proprietà}''$$

è data dall'enunciato

$$\overline{\mathcal{P}} : \text{“}\textit{esiste (almeno) un elemento dell'insieme che non possiede quella certa proprietà}''.$$

2.1.1 Proposizioni con variabile

Se una *proposizione* $\mathcal{P}$ dipende da una variabile che rappresenta i diversi elementi di un certo insieme, non possiamo stabilire se la *proposizione* sia vera o se sia falsa perché non è fissato l'elemento a cui la *proposizione* con variabile è riferita.

Un esempio di *proposizione* con variabile è la seguente *proposizione*

$$\mathcal{P}_1 : \text{“}x \textit{ è un numero naturale divisibile per } 2'',$$

dove la variabile è la x che rappresenta, in questo caso, tutti i *numeri naturali.*

E' quindi evidente che quelli che dobbiamo distinguere, relativamente ad una *proposizione* con variabile, sono gli elementi per i quali essa è vera e quelli per i quali essa è falsa.

Una *proposizione* con variabile permette dunque di suddividere l'insieme $\mathcal{I}$ degli elementi rappresentati dalla variabile in due sottoinsiemi tali che uno dei due contenga tutti e soli gli elementi dell'insieme $\mathcal{I}$ che rendono vera la *proposizione* e l'altro contenga tutti e soli gli elementi dell'insieme $\mathcal{I}$ che rendono falsa la *proposizione.* Data una generica *proposizione* $\mathcal{P}$ con variabile, il sottoinsieme degli elementi per i quali la proposizione $\mathcal{P}$ risulta vera viene chiamato *insieme di validità della proposizione* $\mathcal{P}$ e viene indicato con il simbolo $\mathcal{I}_{\mathcal{P}}$. Poiché, come illustrato in precedenza, un elemento rende vera una *proposizione* e falsa la *proposizione contraria*, segue che l'*insieme di validità* $\mathcal{I}_{\mathcal{P}}$ di una *proposizione* $\mathcal{P}$ e l'*insieme di validità* $\mathcal{I}_{\overline{\mathcal{P}}}$ della *proposizione contraria* $\overline{\mathcal{P}}$ sono uno il complementare dell'altro all'interno dell'insieme di tutti gli elementi rappresentati dalla variabile. L'*insieme di validità* della *proposizione* $\mathcal{P}_1$ con variabile riportata nell'esempio

$$\mathcal{P}_1 : \text{“}x \textit{ è un numero naturale divisibile per } 2''$$

è ovviamente il sottoinsieme dei numeri naturali $\mathcal{I}_{\mathcal{P}_1} = \{0, 2, 4, 6, 8, \dots\}$ e poiché la *proposizione contraria* $\overline{\mathcal{P}}_1$ della *proposizione* $\mathcal{P}_1$ è la *proposizione*

$$\overline{\mathcal{P}}_1 : \text{“}x \textit{ è un numero naturale non divisibile per } 2\text{''},$$

possiamo osservare che l'*insieme di validità* $\mathcal{I}_{\overline{\mathcal{P}}_1}$ della *proposizione contraria* $\overline{\mathcal{P}}_1$ è il sottoinsieme dei numeri naturali

$$\mathcal{I}_{\overline{\mathcal{P}}_1} = \{1, 3, 5, 7, 9, \dots\}$$

e che i sottoinsiemi $\mathcal{I}_{\mathcal{P}_1}, \mathcal{I}_{\overline{\mathcal{P}}_1}$ sono effettivamente uno il complementare dell'altro all'interno dell'insieme dei numeri naturali.

Consideriamo ora la seguente *proposizione* $\mathcal{P}_2$ la cui variabile è il simbolo $f(x)$ che rappresenta una generica funzione $f(x)$, ovvero

$\mathcal{P}_2$: “$f(x)$ *è una funzione tale che per ogni numero reale* ε *positivo, troviamo un intervallo di tipo* $4 - \delta_\varepsilon < x < 4 + \delta_\varepsilon$, *i cui elementi* $x \neq 4$ *soddisfano la disequazione* $1 - \varepsilon < f(x) < 1 + \varepsilon$''

e stabiliamo se una data funzione, in particolare la funzione

$$f(x) = \frac{2x + 3}{3x - 1},$$

rende vera o rende falsa la *proposizione* $\mathcal{P}_2$ con variabile. Per stabilire se la funzione

$$f(x) = \frac{2x + 3}{3x - 1} \tag{2.1}$$

rende vera o falsa la *proposizione* $\mathcal{P}_2$, proviamo a determinare, per ogni fissato $\varepsilon > 0$, un intervallo di tipo $4 - \delta_\varepsilon < x < 4 + \delta_\varepsilon$, tutti gli elementi $x \neq 4$ del quale siano soluzione della disequazione

$$1 - \varepsilon < \frac{2x + 3}{3x - 1} < 1 + \varepsilon, \tag{2.2}$$

in modo tale che se troviamo tale intervallo, concludiamo che la funzione (2.1) rende vera la *proposizione* $\mathcal{P}_2$, mentre se non troviamo tale intervallo, concludiamo che la funzione (2.1) rende falsa la *proposizione* $\mathcal{P}_2$. Sebbene ε possa essere un qualsiasi numero reale positivo, consideriamo solo le disequazioni con i numeri ε positivi minori di 1 e le $x > 1/3$, affinché i tre membri della disequazione (2.2) siano tutti positivi. In virtù dell'identità

$$\frac{2x + 3}{3x - 1} = \frac{2}{3} + \frac{11}{9x - 3},$$

possiamo riscrivere la disequazione (2.2) nella forma (con $0 < \varepsilon < 1$ e $x > 1/3$)

$$1 - \varepsilon < \frac{2}{3} + \frac{11}{9x - 3} < 1 + \varepsilon,$$

da cui, se svolgiamo i passaggi utilizzando le proprietà delle disuguaglianze, segue

$$1 - \varepsilon < \frac{2}{3} + \frac{11}{9x - 3} < 1 + \varepsilon \qquad \Longrightarrow \qquad \frac{1}{3} - \varepsilon < \frac{11}{3(3x - 1)} < \frac{1}{3} + \varepsilon \qquad \Longrightarrow$$

$$\Longrightarrow \quad 1-3\varepsilon < \frac{11}{3x-1} < 1+3\varepsilon \quad \Longrightarrow \quad \frac{1-3\varepsilon}{11} < \frac{1}{3x-1} < \frac{1+3\varepsilon}{11} \quad \Longrightarrow$$
$$\Longrightarrow \quad \frac{11}{1+3\varepsilon} < 3x-1 < \frac{11}{1-3\varepsilon} \quad \Longrightarrow \quad \frac{12+3\varepsilon}{1+3\varepsilon} < 3x < \frac{12-3\varepsilon}{1-3\varepsilon} \quad \Longrightarrow$$
$$\Longrightarrow \quad \frac{4+\varepsilon}{1+3\varepsilon} < x < \frac{4-\varepsilon}{1-3\varepsilon} \quad \Longrightarrow \quad 4-\frac{11\varepsilon}{1+3\varepsilon} < x < 4+\frac{11\varepsilon}{1-3\varepsilon} \quad \Longrightarrow$$
$$\Longrightarrow \quad 4-\frac{11\varepsilon}{1+3\varepsilon} < x < 4+\frac{11\varepsilon}{1+3\varepsilon},$$

dove abbiamo sostituito le identità

$$\frac{4+\varepsilon}{1+3\varepsilon} = 4-\frac{11\varepsilon}{1+3\varepsilon}, \qquad \frac{4-\varepsilon}{1-3\varepsilon} = 4+\frac{11\varepsilon}{1-3\varepsilon}$$

e abbiamo quindi sostituito

$$4+\frac{11\varepsilon}{1+3\varepsilon} \qquad \text{al posto di} \qquad 4+\frac{11\varepsilon}{1-3\varepsilon}$$

perché risulta

$$4+\frac{11\varepsilon}{1+3\varepsilon} < 4+\frac{11\varepsilon}{1-3\varepsilon},$$

ovvero abbiamo scritto alla fine l'intervallo "più stretto"

$$4-\left(\frac{11\varepsilon}{1+3\varepsilon}\right) < x < 4+\left(\frac{11\varepsilon}{1+3\varepsilon}\right) \tag{2.3}$$

al posto dell'intervallo "più largo"

$$4-\frac{11\varepsilon}{1+3\varepsilon} < x < 4+\frac{11\varepsilon}{1-3\varepsilon},$$

in modo che l'intervallo finale (2.3) sia simmetrico con il medesimo

$$\delta_\varepsilon = \frac{11\varepsilon}{1+3\varepsilon}$$

in parentesi tonda nell'estremo sinistro e nell'estremo destro dell'intervallo.

Osserviamo quindi che l'intervallo (2.3), contenente le $x \neq 4$ che soddisfano la disequazione (2.2), non contiene in realtà tutte le soluzioni

$$4-\frac{11\varepsilon}{1+3\varepsilon} < x < 4+\frac{11\varepsilon}{1-3\varepsilon}$$

della disequazione (2.2) perché le soluzioni

$$4+\frac{11\varepsilon}{1+3\varepsilon} < x < 4+\frac{11\varepsilon}{1-3\varepsilon}$$

sono state trascurate quando abbiamo scritto l'intervallo "più stretto" (2.3).

Ma sebbene "alcune" soluzioni x siano state trascurate, tuttavia la funzione

$$f(x) = \frac{2x+3}{3x-1}$$

è tale che per ogni $\varepsilon > 0$ vi sia sempre un intervallo di tipo $4 - \delta_\varepsilon < x < 4 + \delta_\varepsilon$, qual è appunto quello che abbiamo trovato e riportato nella (2.3), tutti gli elementi $x \neq 4$ del quale soddisfano la disequazione $1 - \varepsilon < f(x) < 1 + \varepsilon$ e le soluzioni x della disequazione appartenenti all'intervallo trascurato

$$4 + \frac{11\varepsilon}{1 + 3\varepsilon} < x < 4 + \frac{11\varepsilon}{1 - 3\varepsilon}$$

non hanno "nessuna rilevanza". In virtù di tale verifica, poiché abbiamo determinato l'intervallo che dovevamo determinare, possiamo concludere quindi che la funzione

$$f(x) = \frac{2x + 3}{3x - 1}$$

rende vera la *proposizione* $\mathcal{P}_2$ con variabile e che tale funzione pertanto appartiene all'*insieme di validità* della *proposizione* $\mathcal{P}_2$.

Stabiliamo ora se la funzione

$$f(x) = \frac{x^2 - 16}{8x - 32}$$

rende vera o rende falsa la medesima *proposizione* $\mathcal{P}_2$ con variabile, ovvero proviamo a determinare, per ogni $\varepsilon > 0$, un intervallo di tipo $4 - \delta_\varepsilon < x < 4 + \delta_\varepsilon$, tutti gli elementi $x \neq 4$ del quale siano soluzione della disequazione

$$1 - \varepsilon < \frac{x^2 - 16}{8x - 32} < 1 + \varepsilon. \tag{2.4}$$

Poiché, come dichiarato nell'enunciato della *proposizione* $\mathcal{P}_2$, le x che debbono soddisfare la disequazione (2.4) sono le $x \neq 4$, possiamo utilizzare l'identità (che si ottiene scomponendo i polinomi e semplificando per $x \neq 4$)

$$\frac{x^2 - 16}{8x - 32} = \frac{x + 4}{8}$$

e riscrivere la disequazione (2.4) nella forma (con $0 < \varepsilon < 1$)

$$1 - \varepsilon < \frac{x + 4}{8} < 1 + \varepsilon,$$

da cui, se svolgiamo i passaggi utilizzando le proprietà delle disuguaglianze, segue

$$1 - \varepsilon < \frac{x + 4}{8} < 1 + \varepsilon \qquad \Longrightarrow \qquad 8 - 8\varepsilon < x + 4 < 8 + 8\varepsilon \qquad \Longrightarrow \qquad 4 - 8\varepsilon < x < 4 + 8\varepsilon.$$

Poiché per ogni $\varepsilon > 0$ abbiamo trovato l'intervallo $4 - 8\varepsilon < x < 4 + 8\varepsilon$, della forma $4 - \delta_\varepsilon < x < 4 + \delta_\varepsilon$ con $\delta_\varepsilon = 8\varepsilon$, tutti i numeri $x \neq 4$ del quale soddisfano la disequazione $1 - \varepsilon < f(x) < 1 + \varepsilon$ data nella (2.4), possiamo concludere che anche la funzione

$$f(x) = \frac{x^2 - 16}{8x - 32}$$

rende vera la *proposizione* $\mathcal{P}_2$ con variabile.

Per dimostrare invece che una data funzione $f(x)$ rende falsa la *proposizione* $\mathcal{P}_2$, dobbiamo dimostrare che la funzione $f(x)$ rende vera la *proposizione contraria* $\overline{\mathcal{P}}_2$ della $\mathcal{P}_2$, il cui enunciato si ottiene utilizzando la "regola" puntualizzata subito prima dell'inizio di questo paragrafo.

L'enunciato della *proposizione contraria* $\overline{\mathcal{P}}_2$ è quindi

$$\overline{\mathcal{P}}_2 : \text{“}f(x) \textit{ è una funzione tale che esiste un particolare numero reale } \varepsilon_0 \textit{ positivo per il quale non si trova nessun intervallo di tipo } 4-\delta_\varepsilon < x < 4+\delta_\varepsilon, \textit{ tutti gli elementi } x \neq 4 \textit{ del quale soddisfino la disequazione } 1-\varepsilon_0 < f(x) < 1+\varepsilon_0\text{''}$$

e verifichiamo che la funzione $f(x) = 2x$ rende falsa la *proposizione* $\mathcal{P}_2$ perché rende vera la *proposizione contraria* $\overline{\mathcal{P}}_2$. Infatti, se consideriamo il numero $\varepsilon_0 = 1/2$, da cui seguono i valori $1-\varepsilon_0 = 1-1/2 = 1/2$ e $1+\varepsilon_0 = 1+1/2 = 3/2$, concludiamo che non esiste nessun intervallo di tipo $4-\delta_\varepsilon < x < 4+\delta_\varepsilon$, tutti gli elementi $x \neq 4$ del quale possano soddisfare la disequazione $1/2 < 2x < 3/2$ perché ogni intervallo di tipo $4-\delta_\varepsilon < x < 4+\delta_\varepsilon$ contiene, ad esempio tra gli altri, l'elemento (scritto in forma decimale) $x = 3.9$ e tutti quelli compresi tra 3.9 e 4, i quali, però, com'è “facile” riconoscere, non soddisfano la disequazione $1/2 < 2x < 3/2$.

2.1.2 Proposizioni congiunte

Quando si congiungono due oppure più proposizioni tramite le congiunzioni “*e,o*” (ricordiamo che in lingua italiana la congiunzione “*o*” è equivalente a “*oppure*”), si ottiene ancora una proposizione, che potremmo chiamare *proposizione congiunta.*

Se ad esempio consideriamo le due proposizioni

$$\mathcal{P}_1 : \text{“}x \textit{ è un numero naturale divisibile per } 2\text{''},$$
$$\mathcal{P}_2 : \text{“}x \textit{ è un numero naturale divisibile per } 3\text{''},$$

costruiamo le due *proposizioni congiunte* $\mathcal{Q}_1$ e $\mathcal{Q}_2$ congiungendo $\mathcal{P}_1, \mathcal{P}_2$ con “*e*” e con “*o*”

- $\mathcal{Q}_1$: “$\mathcal{P}_1$ *e* $\mathcal{P}_2$”, ovvero $\mathcal{Q}_1$: “*x è un numero naturale divisibile per 2 e anche divisibile per 3*”, avente *insieme di validità*

$$\mathcal{I}_{\mathcal{Q}_1} = \{0, 6, 12, 18, 24, 30, 36, 42, 48, \dots\}, \tag{2.5a}$$

- $\mathcal{Q}_2$: “$\mathcal{P}_1$ *oppure* $\mathcal{P}_2$”, ovvero $\mathcal{Q}_2$: “*x è un numero naturale divisibile per 2 oppure divisibile per 3*”, avente *insieme di validità*

$$\mathcal{I}_{\mathcal{Q}_2} = \{0, 2, 3, 4, 6, 8, 9, 10, 12, 14, 15, 16, \dots\}. \tag{2.5b}$$

E' immediato rendersi conto che l'*insieme di validità* $\mathcal{I}_{\mathcal{Q}_1}$ di una *proposizione congiunta* tramite la congiunzione “*e*” è dato dall'*intersezione insiemistica* degli *insiemi di validità* delle due proposizioni $\mathcal{P}_1$ e $\mathcal{P}_2$ che vengono congiunte, mentre l'*insieme di validità* $\mathcal{I}_{\mathcal{Q}_2}$ di una *proposizione congiunta* tramite la congiunzione “*oppure*” è dato dall'*unione insiemistica* degli *insiemi di validità* delle due proposizioni $\mathcal{P}_1$ e $\mathcal{P}_2$ che vengono congiunte.

Se consideriamo la *proposizione congiunta* $\mathcal{Q}_1$ scritta in precedenza, ovvero

$$\mathcal{Q}_1 : \text{“}x \textit{ è un numero naturale divisibile per } 2 \textit{ e anche divisibile per } 3\text{''},$$

la sua *proposizione contraria* è la proposizione

$$\overline{\mathcal{Q}}_1 : \text{“}x \textit{ è un numero naturale NON divisibile per } 2 \textit{ oppure NON divisibile per } 3\text{''},$$

avente *insieme di validità*

$$\mathcal{I}_{\overline{\mathcal{Q}}_1} = \{1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 19, \dots\},$$

che, come si riconosce immediatamente, è l'insieme complementare, "dentro" l'insieme dei numeri naturali, dell'*insieme di validità* $\mathcal{I}_{\mathcal{Q}_1}$, riportato nella (2.5a), relativo alla proposizione $\mathcal{Q}_1$. Se a questo punto confrontiamo la *proposizione* $\mathcal{Q}_1$, ottenuta congiungendo due proposizioni $\mathcal{P}_1, \mathcal{P}_2$ tramite la congiunzione "*e*", con la sua *proposizione contraria* $\overline{\mathcal{Q}}_1$, deduciamo la "regola generale" per cui la *proposizione contraria* di una *proposizione* data dalla congiunzione di due proposizioni $\mathcal{P}_1, \mathcal{P}_2$ tramite la "*e*", si ricava congiungendo le due proposizioni contrarie $\overline{\mathcal{P}}_1, \overline{\mathcal{P}}_2$ tramite la congiunzione "*oppure*".

Se consideriamo la *proposizione congiunta* $\mathcal{Q}_2$ scritta in precedenza, ovvero

$$\mathcal{Q}_2 : \text{“}x \textit{ è un numero naturale divisibile per } 2 \textit{ oppure divisibile per } 3'' ,$$

la sua *proposizione contraria* è la proposizione

$$\overline{\mathcal{Q}}_2 : \text{“}x \textit{ è un numero naturale NON divisibile per } 2 \textit{ e NEMMENO divisibile per } 3'' ,$$

avente *insieme di validità*

$$\mathcal{I}_{\overline{\mathcal{Q}}_2} = \{1, 5, 7, 11, 13, 17, 19, 23, 25, \dots\},$$

che, come si riconosce immediatamente, è l'insieme complementare, "dentro" l'insieme dei numeri naturali, dell'*insieme di validità* $\mathcal{I}_{\mathcal{Q}_2}$, riportato nella (2.5b), relativo alla proposizione $\mathcal{Q}_2$. Se a questo punto confrontiamo la *proposizione* $\mathcal{Q}_2$, ottenuta congiungendo due proposizioni $\mathcal{P}_1, \mathcal{P}_2$ tramite la congiunzione "*oppure*", con la sua *proposizione contraria* $\overline{\mathcal{Q}}_2$, deduciamo la "regola generale" per cui la *proposizione contraria* di una *proposizione* data dalla congiunzione di due proposizioni $\mathcal{P}_1, \mathcal{P}_2$ tramite "*oppure*", si ricava congiungendo le due proposizioni contrarie $\overline{\mathcal{P}}_1, \overline{\mathcal{P}}_2$ tramite la congiunzione "*e*".

Simbolicamente possiamo quindi scrivere che la *proposizione contraria* della *proposizione congiunta*

$$\mathcal{Q}_1 : \text{“}\mathcal{P}_1 \textit{ e } \mathcal{P}_2'' \quad \text{è la proposizione} \quad \overline{\mathcal{Q}}_1 : \text{“}\overline{\mathcal{P}}_1 \textit{ oppure } \overline{\mathcal{P}}_2'' ,$$

mentre la *proposizione contraria* della *proposizione congiunta*

$$\mathcal{Q}_2 : \text{“}\mathcal{P}_1 \textit{ oppure } \mathcal{P}_2'' \quad \text{è la proposizione} \quad \overline{\mathcal{Q}}_2 : \text{“}\overline{\mathcal{P}}_1 \textit{ e } \overline{\mathcal{P}}_2'' .$$

Per concludere questa breve trattazione delle *proposizioni*, completiamo il discorso aggiungendo che *definire un concetto* significa assegnare un *nome* agli elementi dell'*insieme di validità* di una data *proposizione*, ovvero la *definizione di un concetto* non è altro che l'assegnazione di un *nome* agli elementi che rendono vera una data *proposizione.*

Ad esempio, i numeri dell'*insieme di validità* della *proposizione*

$$\mathcal{P}_1 : \text{“}x \textit{ è un numero naturale divisibile per } 2''$$

vengono chiamati *numeri naturali pari*, oppure, se non sorge equivoco sull'insieme dei numeri, vengono chiamati semplicemente *numeri pari* e l'enunciato della *definizione* assume, in genere, la forma

definizione: "un numero naturale divisibile per 2 viene denominato *numero pari*″.

Se, come altro esempio, consideriamo la *proposizione*, applicata all'insieme dei triangoli

$$\mathcal{P}_2 : \text{“}T \textit{ è un triangolo avente due lati uguali}'' ,$$

assegnamo a qualsiasi triangolo T che la rende vera il nome di *triangolo isoscele*, ovvero enunciamo la *definizione*, nota sin dai tempi della scuola elementare

definizione: "un triangolo avente due lati uguali viene denominato *triangolo isoscele*″.

2.2 Struttura di un teorema

Date due generiche *proposizioni* con variabile

$$\mathcal{P}: \text{“}x \textit{ soddisfa la proprietà } A\text{''},$$
$$\mathcal{Q}: \text{“}x \textit{ soddisfa la proprietà } B\text{''},$$

dove le due proprietà A, B siano distinte tra loro, un *teorema* è un *enunciato* che stabilisce una connessione logica tra le due *proposizioni* $\mathcal{P}$ e $\mathcal{Q}$, secondo uno dei due possibili schemi

“se è vera la proposizione $\mathcal{P}$, segue che è vera la proposizione $\mathcal{Q}$”,

riscrivibile nella forma “simbolica” $\mathcal{P} \Longrightarrow \mathcal{Q}$, oppure

“se è vera la proposizione $\mathcal{Q}$, segue che è vera la proposizione $\mathcal{P}$”,

riscrivibile nella forma “simbolica” $\mathcal{Q} \Longrightarrow \mathcal{P}$.

Il *teorema* $\mathcal{P} \Longrightarrow \mathcal{Q}$ può essere espresso anche dicendo che

“la proposizione $\mathcal{P}$ implica la proposizione $\mathcal{Q}$”,

mentre il *teorema* $\mathcal{Q} \Longrightarrow \mathcal{P}$ può essere espresso anche dicendo che

“la proposizione $\mathcal{Q}$ implica la proposizione $\mathcal{P}$”.

In un teorema la *proposizione* introdotta dal “*se*” è chiamata *ipotesi* del *teorema* ed è quella verso la quale non punta la freccia, mentre la *proposizione* introdotta dal “*segue che*” è chiamata *tesi* del *teorema* ed è quella verso cui punta la freccia.

Con la terminologia delle *implicazioni*, possiamo dunque dire che in un teorema l'ipotesi *implica* la tesi (verbo attivo) oppure, equivalentemente, la tesi *è implicata* dall'ipotesi (verbo passivo). Il *teorema* $\mathcal{P} \Longrightarrow \mathcal{Q}$ viene anche denominato *teorema scambiato* del teorema $\mathcal{Q} \Longrightarrow \mathcal{P}$ e, viceversa, il *teorema* $\mathcal{Q} \Longrightarrow \mathcal{P}$ viene anche denominato *teorema scambiato* del teorema $\mathcal{P} \Longrightarrow \mathcal{Q}$, ovvero il *teorema scambiato* di un dato *teorema* si ottiene scambiando semplicemente l'ipotesi e la tesi del teorema dato.

Osserviamo che date due *proposizioni* $\mathcal{P}, \mathcal{Q}$, se vale uno dei due *teoremi* $\mathcal{P} \Longrightarrow \mathcal{Q}$ oppure $\mathcal{Q} \Longrightarrow \mathcal{P}$, non è detto che valga anche l'altro, come mostra il seguente esempio.

Se consideriamo le due proposizioni

$$\begin{aligned} &\mathcal{P}: \text{“}x \text{ è un numero naturale divisibile per } 2\text{''}, \\ &\mathcal{Q}: \text{“}x \text{ è un numero naturale divisibile per } 4\text{''}, \end{aligned} \tag{2.6}$$

segue ovviamente che il *teorema* valido è $\mathcal{Q} \Longrightarrow \mathcal{P}$, perché tutti i numeri $4, 8, 12, 16, \ldots$ per i quali è vera l'*ipotesi* $\mathcal{Q}$, risultano evidentemente divisibili per 2, ovvero soddisfano anche di conseguenza la *tesi* $\mathcal{P}$, ma il *teorema contrario* della forma $\mathcal{P} \Longrightarrow \mathcal{Q}$ non è valido perché se un numero è divisibile per 2, ovvero soddisfa l'*ipotesi* $\mathcal{P}$, non segue necessariamente che tale numero soddisfi anche la *tesi* $\mathcal{Q}$ in quanto divisibile per 4, come accade per il numero 10, che soddisfa l'*ipotesi* $\mathcal{P}$ in quanto divisibile per 2, ma non soddisfa la *tesi* $\mathcal{Q}$ in quanto non divisibile per 4. E' importante a questo punto puntualizzare dunque che un *teorema* è valido se l'*insieme di validità* della *proposizione ipotesi* è un sottoinsieme dell'*insieme di validità* della *proposizione tesi* e infatti il *teorema* $\mathcal{Q} \Longrightarrow \mathcal{P}$, costruito con le due *proposizioni* (2.6), è valido perché se $\mathcal{I}_{\mathcal{P}}$ indica l'*insieme di validità* della $\mathcal{P}$ e $\mathcal{I}_{\mathcal{Q}}$ indica l'*insieme di validità* della $\mathcal{Q}$, risulta $\mathcal{I}_{\mathcal{Q}} \subset \mathcal{I}_{\mathcal{P}}$, ovvero $\mathcal{I}_{\mathcal{Q}}$ è un sottoinsieme di $\mathcal{I}_{\mathcal{P}}$.

Ricordiamo che in *teoria degli insiemi* un insieme A è contenuto "dentro" un insieme B, e si scrive $A \subset B$, se ogni elemento di A appartiene anche a B. Viceversa, se ogni elemento dell'insieme A appartiene anche all'insieme B, concludiamo che A "sta dentro" B, ovvero che A è un sottoinsieme di B. Se un insieme A è contenuto "dentro" un insieme B, segue che ogni elemento "fuori" dall'insieme B, ovvero appartenente al *complementare* dell'insieme B indicato con $\overline{B}$, sta automaticamente "fuori" dall'insieme A, ovvero appartiene al *complementare* dell'insieme A indicato con $\overline{A}$. Pertanto, se un insieme A è contenuto "dentro" un insieme B, ovvero $A \subset B$, abbiamo di conseguenza $\overline{B} \subset \overline{A}$.

Se applichiamo queste considerazioni insiemistiche, in particolare, agli *insiemi di validità* delle *proposizioni* ricordando che il *complementare* dell'*insieme di validità* di una *proposizione* coincide con l'*insieme di validità* della *proposizione contraria*, abbiamo che l'*inclusione insiemistica* $\mathcal{I}_{\mathcal{Q}} \subset \mathcal{I}_{\mathcal{P}}$, che dimostra la validità del *teorema* $\mathcal{Q} \Longrightarrow \mathcal{P}$, è equivalente all'*inclusione insiemistica* $\mathcal{I}_{\overline{\mathcal{P}}} \subset \mathcal{I}_{\overline{\mathcal{Q}}}$ che dimostra la validità del *teorema* $\overline{\mathcal{P}} \Longrightarrow \overline{\mathcal{Q}}$.

Pertanto possiamo concludere che ogni *teorema* enunciato nella forma $\mathcal{Q} \Longrightarrow \mathcal{P}$ può sempre essere enunciato anche nella forma equivalente in cui la proposizione contraria della *tesi* sia scritta all'*ipotesi* e la *proposizione contraria* dell'*ipotesi* sia scritta alla *tesi*, ovvero il *teorema* $\mathcal{Q} \Longrightarrow \mathcal{P}$ può sempre essere espresso nella forma equivalente $\overline{\mathcal{P}} \Longrightarrow \overline{\mathcal{Q}}$.

Come esempio abbiamo infatti che il teorema enunciato nella forma

$$\text{“se un numero è divisibile per 4, segue che esso è divisibile per 2''}$$

può essere enunciato anche nella forma equivalente

$$\text{“se un numero NON è divisibile per 2, segue che esso NON è divisibile per 4''}.$$

Dato dunque un *teorema* valido $\mathcal{Q} \Longrightarrow \mathcal{P}$, mentre il *teorema contrario* $\mathcal{P} \Longrightarrow \mathcal{Q}$ può non essere valido, invece il *teorema* $\overline{\mathcal{P}} \Longrightarrow \overline{\mathcal{Q}}$ è sempre valido ed è sempre equivalente al *teorema* valido dato $\mathcal{Q} \Longrightarrow \mathcal{P}$. Quando un teorema della forma $\mathcal{Q} \Longrightarrow \mathcal{P}$ viene dimostrato nella sua forma equivalente $\overline{\mathcal{P}} \Longrightarrow \overline{\mathcal{Q}}$, si dice che il *teorema* viene dimostrato *per assurdo*.

Per illustrare come si dimostra "*per assurdo*" un *teorema*, può essere utile ricordare, tra i tanti possibili, un noto teorema della Geometria Greca che viene appunto dimostrato *per assurdo*. L'enunciato di tale teorema è:

“date tre rette r, s, t su un medesimo piano, se r è parallela a t e s è parallela a t, segue che r è parallela a s'',

dove l'*ipotesi* è la *proposizione congiunta*

$$\mathcal{P} : \text{“}r \text{ è parallela a } t \text{ e } s \text{ è parallela a } t'',$$

ottenuta congiungendo tramite la congiunzione "*e*" le due proposizioni

$$\mathcal{P}_1 : \text{“}r \text{ è parallela a } t'', \qquad \text{e} \qquad \mathcal{P}_2 : \text{“}s \text{ è parallela a } t'',$$

e la *tesi* è la proposizione

$$\mathcal{Q} : \text{“}r \text{ è parallela a } s''.$$

La forma *per assurdo* di questo *teorema* si ottiene dunque scrivendo la *proposizione contraria* della *tesi* come *ipotesi* e la *proposizione contraria* dell'ipotesi come *tesi*.

La *proposizione contraria* della *tesi* è la *proposizione*

$$\overline{\mathcal{Q}} : \text{“}r \text{ NON è parallela a } s'',$$

mentre la *proposizione contraria* dell'*ipotesi* si ottiene congiungendo con "*oppure*" le due *proposizioni contrarie* di $\mathcal{P}_1$ e $\mathcal{P}_2$, ovvero

$$\overline{\mathcal{P}}: \text{“}r \text{ NON è parallela a } t \text{ oppure } s \text{ NON è parallela a } t''.$$

La forma *per assurdo* del *teorema* è pertanto

"date tre rette r, s, t su un medesimo piano, se r NON è parallela a s,
segue che r NON è parallela a t oppure s NON è parallela a t''.

Ci sono tuttavia dei casi in cui il *teorema contrario* di un *teorema* valido è ancora valido. Se consideriamo le due proposizioni

$$\mathcal{P}: \text{“}T \text{ è un triangolo con due lati uguali}'',$$
$$\mathcal{Q}: \text{“}T \text{ è un triangolo con due angoli uguali}'',$$

abbiamo che, com'è noto in Geometria, sono validi entrambi i *teoremi* $\mathcal{P} \Longrightarrow \mathcal{Q}$ e $\mathcal{Q} \Longrightarrow \mathcal{P}$, perché si può dimostrare che vale sia il *teorema*

"se in un triangolo due lati sono uguali, segue che anche due angoli sono uguali$''$

che il *teorema contrario*

"se in un triangolo due angoli sono uguali, segue che anche due lati sono uguali$''$.

Quando con due proposizioni $\mathcal{P}, \mathcal{Q}$ risultano validi entrambi i *teoremi*

$$\mathcal{P} \Longrightarrow \mathcal{Q} \qquad \text{e} \qquad \mathcal{Q} \Longrightarrow \mathcal{P},$$

i due *teoremi* distinti vengono rappresentati simbolicamente con l'unica scrittura

$$\mathcal{P} \Longleftrightarrow \mathcal{Q} \qquad \text{oppure con} \qquad \mathcal{Q} \Longleftrightarrow \mathcal{P}$$

e possono essere enunciati in una qualsiasi delle due forme unificate

"la *proposizione* $\mathcal{P}$ è valida *se e solo se* è valida la *proposizione* $\mathcal{Q}''$,
"la *proposizione* $\mathcal{Q}$ è valida *se e solo se* è valida la *proposizione* $\mathcal{P}''$,

dove l'espressione linguistica "*se e solo se*" sta a significare che valgono appunto entrambi i teoremi, ovvero sia il teorema $\mathcal{P} \Longrightarrow \mathcal{Q}$ che il *teorema contrario* $\mathcal{Q} \Longrightarrow \mathcal{P}$.

Nel caso dei due citati *teoremi* sul triangolo abbiamo che essi possono essere enunciati in una qualsiasi delle due forme unificate

"in un triangolo due lati sono uguali *se e solo se* anche due angoli sono uguali$''$,
"in un triangolo due angoli sono uguali *se e solo se* anche due lati sono uguali$''$.

Osserviamo che quando con due *proposizioni* $\mathcal{P}, \mathcal{Q}$ risultano validi entrambi i *teoremi* $\mathcal{P} \Longrightarrow \mathcal{Q}$ e $\mathcal{Q} \Longrightarrow \mathcal{P}$, abbiamo $\mathcal{I}_{\mathcal{P}} \subset \mathcal{I}_{\mathcal{Q}}$ e $\mathcal{I}_{\mathcal{Q}} \subset \mathcal{I}_{\mathcal{P}}$, ovvero sussiste l'uguaglianza $\mathcal{I}_{\mathcal{P}} = \mathcal{I}_{\mathcal{Q}}$.

Inoltre, per ogni teorema si dice che la proposizione *ipotesi* è *condizione sufficiente* per la proposizione *tesi* e che la proposizione *tesi* è *condizione necessaria* per la proposizione *ipotesi*. Date quindi due proposizioni $\mathcal{P}$ e $\mathcal{Q}$, per capire quale è sufficiente e quale è necessaria per l'altra, occorre prima scrivere scrivere l'implicazione corretta che lega logicamente le due proposizioni. Ancora con riferimento alle due *proposizioni* (2.6), poiché, come già detto, vale l'implicazione $\mathcal{Q} \Longrightarrow \mathcal{P}$, in cui $\mathcal{Q}$ è l'*ipotesi* e $\mathcal{P}$ è la *tesi*, concludiamo che la *proposizione* $\mathcal{Q}$ è *sufficiente* per la *proposizione* $\mathcal{P}$, ovvero "*essere multiplo di* 4" è *sufficiente* per "*essere multiplo di* 2", e che, equivalentemente, la *proposizione* $\mathcal{P}$ è *necessaria* per la *proposizione* $\mathcal{Q}$, ovvero "*essere multiplo di* 2" è *necessario* per "*essere multiplo di* 4".

Capitolo 3

Argomenti di algebra lineare

Chiamiamo *vettore* un elemento di uno *spazio vettoriale* e gli *spazi vettoriali* che consideriamo sono gli spazi S_n, definiti nel testo di riferimento, i cui elementi sono i vettori $\boldsymbol{v} = (x_1, x_2, x_3, \ldots, x_{n-1}, x_n)$ aventi n componenti $x_1, x_2, x_3, \ldots, x_{n-1}, x_n \in \mathbb{R}$.

Ricordiamo che con i *vettori* di uno *spazio vettoriale* S_n si possono eseguire le due *operazioni fondamentali* definite, per ogni coppia di *vettori* ad n componenti

$$\boldsymbol{v}_1 = (x_1, x_2, x_3, \ldots, x_{n-1}, x_n) \qquad \text{e} \qquad \boldsymbol{v}_2 = (y_1, y_2, y_3, \ldots, y_{n-1}, y_n)$$

e per ogni numero reale λ (λ è la lettera, chiamata *lambda*, dell'alfabeto greco), tramite le due uguaglianze, denominate rispettivamente *addizione di due vettori* e *moltiplicazione di un vettore per un numero*

$$\begin{aligned} \boldsymbol{v}_1 + \boldsymbol{v}_2 &= (x_1, x_2, x_3, \ldots, x_n) + (y_1, y_2, y_3, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \ldots, x_n + y_n), \\ k\boldsymbol{v}_1 &= k(x_1, x_2, x_3, \ldots, x_n) = (kx_1, kx_2, kx_3, \ldots, kx_n). \end{aligned}$$

Dati k *vettori* $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k$ del medesimo *spazio vettoriale* S_n e k numeri reali indicati con $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_k$ (α è la lettera, chiamata *alfa*, dell'alfabeto greco), l'operazione che consiste nella sequenza delle due operazioni dello *spazio vettoriale*

$$\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3\boldsymbol{v}_3 + \cdots + \alpha_{k-1}\boldsymbol{v}_{k-1} + \alpha_k\boldsymbol{v}_k \tag{3.1a}$$

si chiama *combinazione lineare* dei k *vettori* dati mediante i k *coefficienti* $\alpha_1, \alpha_2, \ldots, \alpha_k$ e denoteremo sempre i *vettori* con un simbolo in "grassetto" mentre i singoli numeri reali, come ad esempio i *coefficienti* dei *vettori*, con un simbolo senza "grassetto".

Il vettore $\boldsymbol{w}$ che si ottiene come risultato della combinazione lineare (3.1a), ovvero

$$\boldsymbol{w} = \alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3\boldsymbol{v}_3 + \cdots + \alpha_{k-1}\boldsymbol{v}_{k-1} + \alpha_k\boldsymbol{v}_k, \tag{3.1b}$$

viene chiamato *vettore linearmente dipendente* dai k vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_k$ oppure anche *vettore generato* dai k vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_k$. Viceversa, quando abbiamo un'uguaglianza della forma (3.1b), diciamo che i *vettori* $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_k$ *generano il vettore* $\boldsymbol{w}$.

Se chiamiamo *vettore nullo* di S_n il vettore indicato con $\mathbf{0}$ e avente componenti

$$\mathbf{0} = (\overbrace{0, 0, 0, \ldots, 0, 0}^{n \text{ zeri}}),$$

è immediato rendersi conto che risulta $\mathbf{0}+\mathbf{0}=\mathbf{0}$, $k\mathbf{0}=\mathbf{0}$ e $0\boldsymbol{v}=\mathbf{0}$, da cui segue che il *vettore nullo* è l'*elemento neutro* dell'operazione di *addizione* di due vettori e che il risultato della combinazione lineare (3.1a) in cui i coefficienti siano tutti zeri, è sempre, in virtù di come sono definite le due *operazioni fondamentali* sui vettori, "banalmente" il *vettore nullo*, ovvero per ogni insieme di k vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k$ vale sempre l'uguaglianza

$$0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 + 0\boldsymbol{v}_3 + \cdots + 0\boldsymbol{v}_{k-1} + 0\boldsymbol{v}_k = \mathbf{0} \tag{3.2}$$

e la combinazione lineare che *genera il vettore nullo* con coefficienti tutti nulli al primo membro della (3.2) viene chiamata *combinazione lineare banale*.

3.1 Insiemi di vettori e loro proprietà

Quella che vogliamo ora studiare è la proprietà di un dato insieme di k vettori appartenenti ad uno spazio vettoriale S_n rispetto all'uguaglianza

$$\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3\boldsymbol{v}_3 + \cdots + \alpha_{k-1}\boldsymbol{v}_{k-1} + \alpha_k\boldsymbol{v}_k = \mathbf{0},$$

la quale, per $\alpha_1 = \alpha_2 = \alpha_3 = \cdots = \alpha_{k-1} = \alpha_k = 0$, assume la forma (3.2) e per illustrare "meglio" il significato del nostro studio, analizziamo i seguenti esempi.

Se consideriamo l'insieme dei due vettori

$$\mathcal{I}_1 = \{\boldsymbol{u}_1 = (1,2,-1)\,,\ \boldsymbol{u}_2 = (3,1,-2)\,,\ \boldsymbol{u}_3 = (-7,1,4)\} \tag{3.3a}$$

appartenenti allo spazio vettoriale S_3, abbiamo che i vettori dell'insieme $\mathcal{I}_1$ *generano* il *vettore nullo* non solo ovviamente con i coefficienti tutti nulli, come nell'uguaglianza (3.2), ovvero $0\boldsymbol{u}_1+0\boldsymbol{u}_2+0\boldsymbol{u}_3=\mathbf{0}$, ma anche con altre scelte dei coefficienti non tutti nulli, ovvero

$$\begin{aligned}
&2\boldsymbol{u}_1 + (-3)\boldsymbol{u}_2 + (-1)\boldsymbol{u}_3 = 2(1,2,-1) + (-3)(3,1,-2) + (-1)(-7,1,4) = (0,0,0) = \mathbf{0},\\
&4\boldsymbol{u}_1 + (-6)\boldsymbol{u}_2 + (-2)\boldsymbol{u}_3 = \mathbf{0},\\
&6\boldsymbol{u}_1 + (-9)\boldsymbol{u}_2 + (-3)\boldsymbol{u}_3 = \mathbf{0}
\end{aligned}$$

e così via con infinite scelte dei coefficienti che quindi non sono tutti nulli.

Se quindi scriviamo un'uguaglianza $\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3\boldsymbol{v}_3 = \mathbf{0}$ con coefficienti $\alpha_1, \alpha_2, \alpha_3$ non tutti nulli, quale ad esempio l'equazione $2\boldsymbol{u}_1 + (-3)\boldsymbol{u}_2 + (-1)\boldsymbol{u}_3 = \mathbf{0}$, otteniamo infine le tre relazioni

$$\boldsymbol{u}_1 = \frac{3}{2}\,\boldsymbol{u}_2 + \frac{1}{2}\,\boldsymbol{u}_3\,, \qquad \boldsymbol{u}_2 = \frac{2}{3}\,\boldsymbol{u}_1 - \frac{1}{3}\,\boldsymbol{u}_3\,, \qquad \boldsymbol{u}_3 = 2\boldsymbol{u}_1 + (-3)\boldsymbol{u}_2\,,$$

il cui importante significato è che ogni vettore che nell'equazione $2\boldsymbol{u}_1+(-3)\boldsymbol{u}_2+(-1)\boldsymbol{u}_3=\mathbf{0}$ abbia avente coefficiente diverso da zero, può essere scritto (esplicitato al primo membro) come *combinazione lineare* dei rimanenti due vettori ("trasferiti" al secondo membro).

Se invece consideriamo l'insieme dei due vettori dello spazio vettoriale S_2

$$\mathcal{I}_2 = \{\boldsymbol{v}_1 = (1,1)\,,\ \boldsymbol{v}_2 = (2,3)\}, \tag{3.3b}$$

abbiamo, senza "preoccuparci" della dimostrazione, che i vettori dell'insieme $\mathcal{I}_2$ *generano* il *vettore nullo* solo mediante i coefficienti tutti nulli, ovvero solo tramite la combinazione lineare della forma (3.2) con *coefficienti banali* tutti nulli $0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 = \mathbf{0}$.

Poiché dunque l'insieme dei vettori $\mathcal{I}_2$ *genera il vettore nullo* solo con i *coefficienti banali*, ovvero con coefficienti tutti nulli, segue quindi che dalla relazione $0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 = \mathbf{0}$, che *genera* il *vettore nullo*, non è possibile esplicitare nessun vettore, scritto al primo membro, come combinazione lineare dell'altro vettore scritto al secondo membro, perché se a partire dall'uguaglianza $0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 = \mathbf{0}$ si "lascia" un vettore con il proprio coefficiente nullo al primo membro e si "trasferisce" l'altro al secondo membro, non è possibile dividere ambo i membri per il coefficiente nullo del vettore scritto al primo membro.

3.1.1 Insieme di vettori linearmente dipendente

Per estendere e generalizzare le proprietà che abbiamo illustrato relativamente agli insiemi di vettori $\mathcal{I}_1$ e $\mathcal{I}_2$ nelle (3.3), consideriamo un certo spazio vettoriale S_n dal quale estraiamo un insieme $\mathcal{I}$ contenente k vettori, ovvero

$$\mathcal{I} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k\}, \tag{3.4}$$

con $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k \in S_n$, ed enunciamo le due seguenti *proposizioni*

P_1: "*esiste una combinazione lineare dei vettori dell'insieme $\mathcal{I}$ con coefficienti non tutti nulli che dà come risultato il vettore nullo*",

P_2: "*almeno un vettore dell'insieme $\mathcal{I}$ è esprimibile come combinazione lineare dei rimanenti $k-1$ vettori*".

Una combinazione lineare di vettori in cui almeno un vettore abbia coefficiente diverso da zero, viene chiamata *combinazione lineare non banale.*

Dopo aver osservato "facilmente" che l'insieme dei vettori $\mathcal{I}_1$ in (3.3a) rende vere entrambe le proposizioni P_1 e P_2, mentre l'insieme dei vettori $\mathcal{I}_2$ in (3.3b) le rende false entrambe, dimostriamo l'equivalenza delle due proposizioni P_1 e P_2.

Dimostrare che le due proposizioni P_1 e P_2 sono equivalenti significa dimostrare che se una delle due proposizioni è valida, segue che è valida anche l'altra.

Teorema 3.1.1. *Se la proposizione P_1 è valida, segue che anche la proposizione P_2 è valida, ovvero simbolicamente abbiamo $P_1 \Longrightarrow P_2$.*

Dimostrazione. Per dimostrare $P_1 \Longrightarrow P_2$ esplicitiamo l'ipotesi P_1, ovvero, poiché esiste una combinazione lineare dei vettori dell'insieme $\mathcal{I}$ in (3.4) con coefficienti $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_k$ non tutti nulli che dà come risultato il *vettore nullo*, scriviamo l'uguaglianza

$$\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3\boldsymbol{v}_3 + \cdots + \alpha_{k-1}\boldsymbol{v}_{k-1} + \alpha_k\boldsymbol{v}_k = \mathbf{0}, \tag{3.5}$$

in cui supponiamo, senza perdita di generalità, che il coefficiente diverso da zero sia, ad esempio, il coefficiente α_2. Se ora riscriviamo l'uguaglianza (3.5) nella forma in cui il vettore $\boldsymbol{v}_2$ con coefficiente diverso da zero α_2 sta "da solo" al primo membro, otteniamo

$$\alpha_2\boldsymbol{v}_2 = -\alpha_1\boldsymbol{v}_1 - \alpha_3\boldsymbol{v}_3 - \cdots - \alpha_{k-1}\boldsymbol{v}_{k-1} - \alpha_k\boldsymbol{v}_k,$$

da cui, poiché α_2 è diverso da zero, segue la relazione finale

$$\boldsymbol{v}_2 = -\frac{\alpha_1}{\alpha_2}\boldsymbol{v}_1 - \frac{\alpha_3}{\alpha_2}\boldsymbol{v}_3 - \frac{\alpha_4}{\alpha_2}\boldsymbol{v}_4 - \cdots - \frac{\alpha_{k-1}}{\alpha_2}\boldsymbol{v}_{k-1} - \frac{\alpha_k}{\alpha_2}\boldsymbol{v}_k,$$

che mostra che l'insieme $\mathcal{I}$ soddisfa la proposizione P_2. *C.d.d*

Teorema 3.1.2. *Se la proposizione P_2 è valida, segue che anche la proposizione P_1 è valida, ovvero simbolicamente abbiamo $P_2 \Longrightarrow P_1$.*

Dimostrazione. Per dimostrare $P_2 \Longrightarrow P_1$ esplicitiamo l'ipotesi P_2, ovvero che almeno un vettore dell'insieme $\mathcal{I}$ in (3.4) sia esprimibile come combinazione lineare dei rimanenti $k-1$ vettori dell'insieme. Supponiamo, senza perdita di generalità, che il vettore esprimibile come combinazione lineare degli altri vettori sia, ad esempio, il vettore $\boldsymbol{v}_3$ e scriviamo dunque, appunto per ipotesi, la relazione

$$\boldsymbol{v}_3 = \beta_1\boldsymbol{v}_1 + \beta_2\boldsymbol{v}_2 + \beta_4\boldsymbol{v}_4 + \cdots + \beta_{k-1}\boldsymbol{v}_{k-1} + \beta_k\boldsymbol{v}_k\,.$$

Se a questo punto "portiamo" tutti i vettori al primo membro, otteniamo l'uguaglianza

$$-\beta_1\boldsymbol{v}_1 - \beta_2\boldsymbol{v}_2 + \boldsymbol{v}_3 - \beta_4\boldsymbol{v}_4 - \cdots - \beta_{k-1}\boldsymbol{v}_{k-1} - \beta_k\boldsymbol{v}_k = \boldsymbol{0} \tag{3.6}$$

che mostra che l'insieme $\mathcal{I}$ soddisfa la proposizione P_1 perché, sebbene non vi siano informazioni relative ai coefficienti β, se siano nulli oppure non nulli, tuttavia il coefficiente del vettore $\boldsymbol{v}_3$, come si riconosce, è uguale ad 1, ovvero esiste una combinazione lineare dei vettori dell'insieme $\mathcal{I}$ in (3.4) con coefficienti non tutti nulli, appunto la combinazione lineare al primo membro nella (3.6) in cui il coefficiente di $\boldsymbol{v}_3$ è diverso da zero, che dà come risultato il *vettore nullo* al secondo membro della (3.6). *C.d.d.*

In base a quanto discusso nel capitolo 2 sulla logica delle proposizioni, i due teoremi distinti 3.1.1, rappresentato anche con il simbolo $P_1 \Longrightarrow P_2$, e 3.1.2, rappresentato anche con il simbolo $P_2 \Longrightarrow P_1$, possono essere rappresentati dall'unico simbolo $P_1 \Longleftrightarrow P_2$ e possono essere enunciati tramite le seguenti due formulazioni unificate.

Prima formulazione del teorema $P_1 \Longleftrightarrow P_2$. *Dato un insieme di vettori, esiste una loro combinazione lineare con coefficienti non tutti nulli che dà come risultato il vettore nullo se e solo se almeno uno dei vettori dell'insieme è esprimibile come combinazione lineare dei rimanenti vettori dell'insieme.*

oppure, se invertiamo le proposizioni

Prima formulazione del teorema $P_1 \Longleftrightarrow P_2$. *Dato un insieme di vettori, almeno uno di essi è esprimibile come combinazione lineare dei rimanenti vettori dell'insieme se e solo se esiste una loro combinazione lineare con coefficienti non tutti nulli che dà come risultato il vettore nullo.*

Come abbiamo spiegato nel capitolo sulla logica delle proposizioni, definire un concetto significa attribuire un nome agli elementi appartenenti all'*insieme di validità* di una proposizione. Possiamo introdurre dunque la seguente

Definizione 3.1.1. *Se un insieme $\mathcal{I}$ di vettori soddisfa una delle due proposizioni P_1, P_2, e quindi, in virtù dei due teoremi 3.1.1 e 3.1.2 che abbiamo dimostrato, di conseguenza anche l'altra, l'insieme $\mathcal{I}$ viene chiamato "insieme linearmente dipendente".*

Come si riconosce, la *definizione* 3.1.1 di insieme *linearmente dipendente* è stata introdotta dopo la discussione dei due teoremi 3.1.1 e 3.1.2 ed è stata attribuita a quegli insiemi che rendono vere simultaneamente le due proposizioni P_1, P_2.

Generalmente, però, nei testi che trattano l'Algebra Lineare la *definizione* 3.1.1 di insieme *linearmente dipendente* viene attribuita inizialmente a quegli insiemi che rendono vera solo una delle due proposizioni P_1 oppure P_2. Supponiamo che la *definizione* di insieme *linearmente dipendente* sia introdotta mediante la validità della proposizione P_1, ovvero supponiamo di scrivere

Prima *definizione alternativa* di insieme *linearmente dipendente*. *Un insieme di vettori viene chiamato "linearmente dipendente" se rende vera la proposizione P_1.*

A partire da questa definizione, il teorema 3.1.1 può essere riscritto con la formulazione

Prima nuova formulazione del teorema 3.1.1. *Se un insieme di vettori è "linearmente dipendente", segue che almeno un vettore dell'insieme è esprimibile come combinazione lineare degli altri vettori dell'insieme.*

Il teorema 3.1.2 può essere riscritto con la formulazione

Prima nuova formulazione del teorema 3.1.2. *Se almeno un vettore di un insieme di vettori è esprimibile come combinazione lineare degli altri vettori dell'insieme, segue che l'insieme di vettori è "linearmente dipendente".*

Tali prime nuove formulazioni dei teoremi 3.1.1 e 3.1.2 possono quindi essere unificate nell'unica formulazione

Teorema. *Un insieme di vettori è "linearmente dipendente" se e solo se almeno un vettore dell'insieme è esprimibile come combinazione lineare degli altri vettori dell'insieme.*

Supponiamo ora invece che la *definizione* di insieme *linearmente dipendente* sia introdotta mediante la validità della proposizione P_2, ovvero supponiamo di scrivere

Seconda *definizione alternativa* di insieme *linearmente dipendente.* *Un insieme di vettori è chiamato "linearmente dipendente" se rende vera la proposizione P_2.*

A partire da questa definizione, il teorema 3.1.1 può essere riscritto con la formulazione

Seconda nuova formulazione del teorema 3.1.1. *Dato un insieme di vettori, se esiste una loro combinazione lineare con coefficienti non tutti nulli che dà come risultato il vettore nullo, segue che l'insieme di vettori è "linearmente dipendente".*

Il teorema 3.1.2 può essere riscritto con la formulazione

Seconda nuova formulazione del teorema 3.1.2. *Se un insieme di vettori è "linearmente dipendente", segue che esiste una loro combinazione lineare con coefficienti non tutti nulli che dà come risultato il vettore nullo.*

Tali seconde nuove formulazioni dei teoremi 3.1.1 e 3.1.2 possono quindi essere unificate nell'unica formulazione

Teorema. *Un insieme di vettori è "linearmente dipendente" se e solo se esiste una loro combinazione lineare con coefficienti non tutti nulli che dà come risultato il vettore nullo.*

E' molto importante sottolineare quindi che per evitare confusioni su tutta questa terminologia e su tutte queste formulazioni, è fondamentale distinguere le *definizioni* dai *teoremi* e puntualizzare all'inizio la scelta della formulazione della *definizione*, da cui discende poi la corrispondente formulazione dei *teoremi* in cui debbono essere riconosciute e distinte la proposizione *ipotesi* e la proposizione *tesi*.

La denominazione *linearmente dipendente* attribuita ad un insieme di vettori che soddisfa le due proposizioni P_1, P_2 deriva, in particolare, dalla proprietà contenuta nella proposizione P_2 secondo la quale un vettore esprimibile come combinazione lineare di altri vettori viene chiamato *vettore linearmente dipendente* da quelli di cui esso è appunto combinazione lineare. Quindi un insieme di vettori viene chiamato *linearmente dipendente*, come dichiarato nella *definizione 3.1.1*, se anche solo un vettore di tale insieme risulta esprimibile come combinazione lineare degli altri vettori dell'insieme, ovvero se anche solo un vettore dell'insieme di vettori risulta *linearmente dipendente* dagli altri rimanenti vettori dell'insieme. E' importante sottolineare quindi che se un insieme di vettori è *linearmente dipendente*, questo non significa che tutti i vettori dell'insieme siano *linearmente dipendenti* dai rimanenti vettori dell'insieme, ovvero, in altre parole, un insieme di vettori può essere *linearmente dipendente* anche se qualche vettore dell'insieme non fosse *linearmente dipendente* dagli altri vettori dell'insieme, come mostra il seguente esempio. Se consideriamo l'insieme $\mathcal{I}$ dei tre vettori

$$\mathcal{I} = \{\boldsymbol{v}_1 = (2,3), \quad \boldsymbol{v}_2 = (4,6), \quad \boldsymbol{v}_3 = (1,-1)\}, \tag{3.7}$$

abbiamo che il vettore $\boldsymbol{v}_1$ è combinazione lineare dei due rimanenti vettori $\boldsymbol{v}_2, \boldsymbol{v}_3$, ovvero $\boldsymbol{v}_1$ è *linearmente dipendente* dai vettori $\boldsymbol{v}_2, \boldsymbol{v}_3$, perché risulta

$$\boldsymbol{v}_1 = \frac{1}{2}\boldsymbol{v}_2 + 0\boldsymbol{v}_3\,; \tag{3.8a}$$

il vettore $\boldsymbol{v}_2$ è *linearmente dipendente* dai vettori $\boldsymbol{v}_1, \boldsymbol{v}_3$, perché risulta

$$\boldsymbol{v}_2 = 2\boldsymbol{v}_1 + 0\boldsymbol{v}_3\,, \tag{3.8b}$$

ma il vettore $\boldsymbol{v}_3$ non è *linearmente dipendente* dai vettori $\boldsymbol{v}_1, \boldsymbol{v}_2$ perché nessuna combinazione lineare dei due vettori $\boldsymbol{v}_1, \boldsymbol{v}_2$ può dare come risultato il vettore $\boldsymbol{v}_3$.

Tuttavia l'insieme $\mathcal{I}$ dei tre vettori $\mathcal{I} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ è *linearmente dipendente* perché in effetti almeno un vettore dell'insieme $\mathcal{I}$ è combinazione lineare degli altri due, come mostrano in particolare le due combinazioni lineari (3.8) relative ai vettori $\boldsymbol{v}_1$ e $\boldsymbol{v}_2$.

In ogni caso non è sorprendente che il vettore $\boldsymbol{v}_3$ dell'insieme pur *linearmente dipendente* $\mathcal{I} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ non sia esprimibile come combinazione lineare degli altri due vettori dell'insieme $\mathcal{I}$ perché in tutte le infinite combinazioni lineari dei tre vettori con coefficienti non tutti nulli che dànno il *vettore nullo* come risultato, il vettore $\boldsymbol{v}_3$ ha sempre coefficiente zero. Alcune di tali combinazioni lineari dei tre vettori con coefficienti non tutti nulli che dànno come risultato il *vettore nullo* sono

$$2\boldsymbol{v}_1 + (-1)\boldsymbol{v}_2 + 0\boldsymbol{v}_3 = \boldsymbol{0}, \qquad 4\boldsymbol{v}_1 + (-2)\boldsymbol{v}_2 + 0\boldsymbol{v}_3 = \boldsymbol{0}, \qquad 6\boldsymbol{v}_1 + (-3)\boldsymbol{v}_2 + 0\boldsymbol{v}_3 = \boldsymbol{0}$$

e così via, e poiché, come già sappiamo, da tali equazioni possono essere esplicitati solo i vettori che hanno coefficiente diverso da zero, segue che il vettore $\boldsymbol{v}_3$ non può essere combinazione lineare dei due vettori $\boldsymbol{v}_1, \boldsymbol{v}_2$ perché $\boldsymbol{v}_3$ ha appunto sempre coefficiente nullo.

Infine vogliamo studiare cosa accade quando i vettori di un insieme *linearmente dipendente* dànno come risultato un vettore diverso dal *vettore nullo*.

Se consideriamo i vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ dell'insieme *linearmente dipendente* (3.7), abbiamo che essi *generano* tra gli altri, ad esempio, il vettore $\boldsymbol{w} = (6,9)$ con infinite scelte dei coefficienti attribuibili ai tre vettori, perché risulta

$$\boldsymbol{w} = 3\boldsymbol{v}_1 + 0\boldsymbol{v}_2 + 0\boldsymbol{v}_3\,, \qquad \boldsymbol{w} = 1\boldsymbol{v}_1 + 1\boldsymbol{v}_2 + 0\boldsymbol{v}_3\,, \qquad \boldsymbol{w} = 2\boldsymbol{v}_1 + \frac{1}{2}\boldsymbol{v}_2 + 0\boldsymbol{v}_3$$

e così via. Enunciamo quindi il seguente teorema che possiamo chiamare *teorema dei vettori linearmente dipendenti*.

Teorema 3.1.3. *(Teorema dei vettori linearmente dipendenti) Se un insieme di vettori* $\mathcal{I} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k\}$ *è linearmente dipendente e risulta*

$$\boldsymbol{w} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \alpha_3 \boldsymbol{v}_3 + \cdots + \alpha_{k-1} \boldsymbol{v}_{k-1} + \alpha_k \boldsymbol{v}_k, \tag{3.9}$$

segue che i coefficienti $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k$ *della combinazione lineare non sono unici, ovvero i vettori dell'insieme* $\mathcal{I}$ *generano il vettore* $\boldsymbol{w}$ *in infiniti modi.*

Dimostrazione. Poiché per ipotesi l'insieme $\mathcal{I}$ è *linearmente dipendente*, segue che essi soddisfano la proposizione P_1, ovvero esiste una combinazione lineare con coefficienti non tutti nulli che dà come risultato il *vettore nullo* e quindi possiamo scrivere

$$\beta_1 \boldsymbol{v}_1 + \beta_2 \boldsymbol{v}_2 + \beta_3 \boldsymbol{v}_3 + \cdots + \beta_{k-1} \boldsymbol{v}_{k-1} + \beta_k \boldsymbol{v}_k = \boldsymbol{0}, \tag{3.10}$$

dove almeno un coefficiente β_i è appunto diverso da zero. Se adesso ricordiamo che il *vettore nullo* è l'elemento neutro dell'addizione di vettori e sostituiamo il primo membro della (3.10) al posto del *vettore nullo*, otteniamo

$$\begin{aligned} \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{0} &= (\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \alpha_3 \boldsymbol{v}_3 + \cdots + \alpha_k \boldsymbol{v}_k) + (\beta_1 \boldsymbol{v}_1 + \beta_2 \boldsymbol{v}_2 + \beta_3 \boldsymbol{v}_3 + \cdots + \beta_k \boldsymbol{v}_k) = \\ &= (\alpha_1 + \beta_1)\boldsymbol{v}_1 + (\alpha_2 + \beta_2)\boldsymbol{v}_2 + (\alpha_3 + \beta_3)\boldsymbol{v}_3 + \cdots + (\alpha_k + \beta_k)\boldsymbol{v}_k \,, \end{aligned}$$

ovvero, se mettiamo in evidenza i vettori, otteniamo la relazione

$$\boldsymbol{w} = (\alpha_1 + \beta_1)\boldsymbol{v}_1 + (\alpha_2 + \beta_2)\boldsymbol{v}_2 + (\alpha_3 + \beta_3)\boldsymbol{v}_3 + \cdots + (\alpha_k + \beta_k)\boldsymbol{v}_k \,, \tag{3.11}$$

che mostra che il vettore $\boldsymbol{w}$, espresso per ipotesi tramite la combinazione lineare (3.9), può essere espresso anche tramite la seconda combinazione lineare (3.11) in cui il coefficiente di almeno un vettore al secondo membro è diverso dal coefficiente del corrispondente vettore al secondo membro nella combinazione lineare (3.9) perché almeno un coefficiente β_i nella (3.10) è diverso da zero e quindi almeno un coeffciente $\alpha_i + \beta_i$ è diverso da α_i. *C.d.d.*

Poiché un insieme di vettori si dice *linearmente dipendente* se i vettori dell'insieme *generano* il *vettore nullo* con infinite scelte dei coefficienti, possiamo esprimere il significato del teorema 3.1.3 dicendo dunque che se i vettori di un insieme *generano* il *vettore nullo* con infinite scelte dei coefficienti, segue che essi *generano* con infinite scelte dei coefficienti anche qualsiasi altro vettore che essi possono *generare.*

3.1.2 Insieme di vettori linearmente indipendente

Se indichiamo con $\overline{P}_1$ la proposizione contraria della P_1 e con $\overline{P}_2$ la proposizione contraria della P_2, dove le proposizioni P_1, P_2 sono quelle introdotte all'inizio del par. 3.1.1, abbiamo, per quanto discusso sulla logica delle proposizioni, che il teorema $P_1 \Longrightarrow P_2$ è equivalente al teorema $\overline{P}_2 \Longrightarrow \overline{P}_1$ e che il teorema $P_2 \Longrightarrow P_1$ risulta equivalente al teorema $\overline{P}_1 \Longrightarrow \overline{P}_2$. Se, per evitare la *barra* sopra le lettere, denotiamo la proposizione $\overline{P}_1$ con Q_1 e la proposizione $\overline{P}_2$ con Q_2, le due proposizioni $\overline{P}_1 \equiv Q_1$ e $\overline{P}_2 \equiv Q_2$ sono

- $Q_1 \equiv \overline{P}_1$: "*l'unica combinazione lineare dei vettori dell'insieme* $\mathcal{I}$ *che dà come risultato il vettore nullo è quella con coefficienti tutti nulli*";
- $Q_2 \equiv \overline{P}_2$: "*nessun vettore dell'insieme* $\mathcal{I}$ *è esprimibile come combinazione lineare dei rimanenti* $k-1$ *vettori*".

A questo punto la formulazione del teorema $\overline{P}_1 \Longrightarrow \overline{P}_2$ (scrivibile come $Q_1 \Longrightarrow Q_2$) è

Teorema 3.1.4. *Se l'unica combinazione lineare dei vettori di un insieme che dà come risultato il vettore nullo è la "combinazione lineare banale", ovvero quella con coefficienti tutti nulli, segue che nessun vettore dell'insieme è esprimibile come combinazione lineare dei rimanenti vettori dell'insieme.*

Tale teorema, rappresentato mediante il simbolo $\overline{P}_1 \Longrightarrow \overline{P}_2$, risulta già dimostrato perché esso è equivalente, per la logica delle proposizioni, al teorema $P_2 \Longrightarrow P_1$, appunto dimostrato nel paragrafo precedente.

La formulazione del teorema $\overline{P}_2 \Longrightarrow \overline{P}_1$ (scrivibile come $Q_2 \Longrightarrow Q_1$) è

Teorema 3.1.5. *Se nessun vettore di un insieme di vettori è esprimibile come combinazione lineare dei rimanenti vettori dell'insieme, segue che l'unica combinazione lineare dei vettori dell'insieme che dà come risultato il vettore nullo è la "combinazione lineare banale", ovvero quella con coefficienti tutti nulli.*

Tale teorema, rappresentato mediante il simbolo $\overline{P}_2 \Longrightarrow \overline{P}_1$, risulta già dimostrato perché esso è equivalente, per la logica delle proposizioni, al teorema $P_1 \Longrightarrow P_2$, appunto dimostrato nel paragrafo precedente.

I due teoremi 3.1.4 e 3.1.5 mostrano dunque che le due proposizioni Q_1, Q_2 risultano equivalenti e tale equivalenza non è sorprendente perché corrisponde ovviamente all'equivalenza delle due proposizioni P_1, P_2 delle quali Q_1, Q_2 sono rispettivamente le contrarie. Possiamo pertanto rappresentare i due teoremi con l'usuale simbolo $Q_1 \Longleftrightarrow Q_2$ oppure con $Q_2 \Longleftrightarrow Q_1$ e riformulare i due teoremi 3.1.4 e 3.1.5 in forma unificata tramite i due seguenti enunciati.

Teorema $Q_1 \Longleftrightarrow Q_2$. *Dato un insieme di vettori, l'unica loro combinazione lineare che dà come risultato il vettore nullo è la "combinazione lineare banale", ovvero quella con coefficienti tutti nulli, se e solo se nessun vettore dell'insieme è esprimibile come combinazione lineare dei rimanenti vettori dell'insieme.*

Teorema $Q_2 \Longleftrightarrow Q_1$. *Dato un insieme di vettori, nessun vettore dell'insieme è esprimibile come combinazione lineare dei rimanenti vettori dell'insieme se e solo se l'unica loro combinazione lineare che dà come risultato il vettore nullo è la "combinazione lineare banale", ovvero quella con coefficienti tutti nulli.*

Possiamo a questo punto introdurre la seguente

Definizione 3.1.2. *Se un insieme $\mathcal{I}$ di vettori soddisfa una delle due proposizioni Q_1, Q_2, e quindi, in virtù dei due teoremi $Q_1 \Longrightarrow Q_2$ e $Q_2 \Longrightarrow Q_1$, di conseguenza anche l'altra, l'insieme $\mathcal{I}$ viene chiamato "insieme linearmente indipendente".*

Se, come abbiamo svolto per un insieme di vettori *linearmente dipendente*, enunciamo la *definizione* di insieme di vettori *linearmente indipendente* riferendola alla validità della sola proposizione Q_1 oppure alla validità della sola proposizione Q_2, possiamo enunciare anche i teoremi 3.1.4 e 3.1.5 in forma diversa, scrivendo la *definizione* di insieme di vettori *linearmente indipendente* al posto della proprietà a cui si è scelto di attribuire tale *definizione*.

Possiamo dimostrare quindi il seguente teorema sulla combinazione lineare di vettori appartenenti ad un insieme linearmente indipendente.

Teorema 3.1.6. ***(Teorema dei vettori linearmente indipendenti)*** *Se un insieme di vettori* $\mathcal{I} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k\}$ *è linearmente indipendente e risulta*

$$\boldsymbol{w} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \alpha_3 \boldsymbol{v}_3 + \dots + \alpha_{k-1} \boldsymbol{v}_{k-1} + \alpha_k \boldsymbol{v}_k, \tag{3.12}$$

segue che i coefficienti $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k$ *della combinazione lineare sono unici.*

Dimostriamo tale teorema *per assurdo*, ovvero dimostriamo la sua formulazione equivalente che, in base a quanto illustrato nel capitolo sulla logica delle proposizioni, si ottiene scrivendo il *contrario della tesi* come *ipotesi* e il *contrario dell'ipotesi* come *tesi.*

L'enunciato equivalente del teorema 3.1.6 è dunque

Enunciato equivalente del teorema 3.1.6. *Se i coefficienti* $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k$ *della combinazione lineare al secondo membro nella relazione*

$$\boldsymbol{w} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \alpha_3 \boldsymbol{v}_3 + \dots + \alpha_{k-1} \boldsymbol{v}_{k-1} + \alpha_k \boldsymbol{v}_k$$

non sono unici, segue che i vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k$ *risultano "linearmente dipendenti".*

Dimostrazione. L'esplicitazione dell'*ipotesi* è che vi sono due combinazioni lineari distinte (ovvero con coefficienti non unici) che dà il vettore $\boldsymbol{w}$ come risultato.

Da tale ipotesi segue che se scriviamo dunque il vettore $\boldsymbol{w}$ tramite due distinte combinazioni lineari

$$\begin{cases} \boldsymbol{v} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \alpha_3 \boldsymbol{v}_3 + \dots + \alpha_{k-1} \boldsymbol{v}_{k-1} + \alpha_k \boldsymbol{v}_k \\ \boldsymbol{v} = \beta_1 \boldsymbol{v}_1 + \beta_2 \boldsymbol{v}_2 + \beta_3 \boldsymbol{v}_3 + \dots + \beta_{k-1} \boldsymbol{v}_{k-1} + \beta_k \boldsymbol{v}_k \, , \end{cases}$$

in cui appunto almeno un coefficiente β_i sia distinto dal corrispondente coefficiente α_i, otteniamo, sottraendo membro a membro, l'uguaglianza

$$(\alpha_1 - \beta_1)\boldsymbol{v}_1 + (\alpha_2 - \beta_2)\boldsymbol{v}_2 + (\alpha_3 - \beta_3)\boldsymbol{v}_3 + \dots + (\alpha_{k-1} - \beta_{k-1})\boldsymbol{v}_{k-1} + (\alpha_k - \beta_k)\boldsymbol{v}_k = \boldsymbol{0} \tag{3.13}$$

in cui almeno un coefficiente $\alpha_i - \beta_i$ risulta diverso da zero.

A questo punto osserviamo che la relazione (3.13) ha al primo membro una combinazione lineare dei vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k$ con coefficienti non tutti nulli che dà come risultato il *vettore nullo* al secondo membro e quindi concludiamo che l'insieme dei vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k$ risulta *linearmente dipendente*, ovvero vale la *tesi. C.d.d.*

Poiché un insieme di vettori si dice *linearmente indipendente* se i vettori dell'insieme *generano* il *vettore nullo* solo con coefficienti tutti nulli, ovvero in modo unico, possiamo esprimere il significato del teorema 3.1.6 dicendo dunque che se i vettori di un insieme *generano* il *vettore nullo* unicamente con coefficienti tutti nulli, segue che essi *generano* con coefficienti unici anche qualsiasi altro vettore che essi possono *generare.*

3.2 Sistemi di equazioni lineari e matrici

Consideriamo lo spazio vettoriale S_m, ovvero lo spazio vettoriale dei vettori aventi m componenti, e scegliamo in esso il vettore $\boldsymbol{w} = (b_1, b_2, \dots, b_m)$ e n vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_n$ aventi componenti a_{ij}, dove a_{ij} rappresenta la i-esima componente del j-esimo vettore $\boldsymbol{v}_j$

$$\boldsymbol{v}_1 = (a_{11}, a_{21}, \dots, a_{m1}), \quad \boldsymbol{v}_2 = (a_{12}, a_{22}, \dots, a_{m2}), \quad \dots, \quad \boldsymbol{v}_n = (a_{1n}, a_{2n}, \dots, a_{mn}).$$

Vogliamo mostrare quindi come l'*equazione vettoriale* $x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + x_3\boldsymbol{v}_3 + \dots + x_n\boldsymbol{v}_n = \boldsymbol{w}$ nelle n incognite $x_1, x_2, x_3, \dots, x_n$ possa essere riscritta nella forma di un *sistema di equazioni lineari*. Se sostituiamo le componenti ai vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{w}$, l'*equazione vettoriale* $x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + x_3\boldsymbol{v}_3 + \dots + x_{n-1}\boldsymbol{v}_{n-1} + x_n\boldsymbol{v}_n = \boldsymbol{w}$ assume la forma

$$x_1(a_{11}, a_{21}, \dots, a_{m1}) + x_2(a_{12}, a_{22}, \dots, a_{m2}) + \dots + x_n(a_{1n}, a_{2n}, \dots, a_{mn}) = (b_1, b_2, \dots, b_m),$$

la quale, se riscriviamo in colonna le componenti dei vettori, diventa

$$x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + x_3\begin{pmatrix} a_{13} \\ a_{23} \\ \vdots \\ a_{m3} \end{pmatrix} + \dots + x_{n-1}\begin{pmatrix} a_{1,n-1} \\ a_{2,n-1} \\ \vdots \\ a_{m,n-1} \end{pmatrix} + x_n\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_n \\ \vdots \\ b_m \end{pmatrix},$$

da cui, se uguagliamo le componenti corrispondenti nei due membri, ricaviamo appunto il *sistema lineare* avente tante equazioni, ovvero m, quante sono le componenti di ciascun vettore e tante incognite, ovvero n, quanti sono i vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1,n-1}x_{n-1} + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2,n-1}x_{n-1} + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \dots + a_{3,n-1}x_{n-1} + a_{3n}x_n = b_3 \\ \qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \dots + a_{m,n-1}x_{n-1} + a_{mn}x_n = b_m\,. \end{cases} \tag{3.14}$$

Ad un *sistema di equazioni lineari* si associano la *matrice*, che indicheremo con A, contenente i soli coefficienti delle incognite e chiamata *matrice incompleta*, ovvero

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{m,n-1} & a_{mn} \end{pmatrix}, \tag{3.15a}$$

e la *matrice*, indicata con B e denominata *matrice completa*, contenente i coefficienti delle incognite e la colonna dei termini noti, ovvero

$$B = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3,n-1} & a_{3n} & b_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{m,n-1} & a_{mn} & b_m \end{pmatrix}. \tag{3.15b}$$

Una *matrice* è dunque una *tabella* di numeri disposti in righe e in colonne e quando il numero delle righe è uguale al numero delle colonne, diciamo che la matrice è *quadrata*, mentre se il numero delle righe è diverso dal numero delle colonne, diciamo che la matrice è *rettangolare*. Una *matrice* che possiede m righe e n colonne viene chiamata *matrice* $m \times n$ e il suo elemento sulla i-esima riga e j-esima colonna è chiamato *elemento* i, j e si denota con una lettera, ad esempio con la lettera a, avente gli indici i, j, ovvero con a_{ij}.

Per una *matrice quadrata* chiamiamo *ordine* il numero delle sue righe e delle sue colonne: quindi una matrice quadrata di ordine 1 possiede una sola riga ed una sola colonna, ovvero possiede un solo elemento ed è una matrice 1×1; una matrice quadrata di ordine 2 possiede due righe e due colonne ed è una matrice 2×2; una matrice quadrata di ordine 3 possiede tre righe e tre colonne ed è una matrice 3×3, e così via.

Quando si parla di matrice avente un certo *ordine*, si può anche omettere di scrivere *quadrata* perché è sottinteso che una *matrice di ordine* n sia una *matrice quadrata* $n \times n$.

3.2.1 Determinante delle matrici quadrate

Per una generica *matrice quadrata* $\mathcal{A}$ di ordine n si introduce il concetto di *determinante*, indicato con il simbolo $\det A$, che è un numero associato alla matrice attarverso la "regola" che ora illustreremo per ogni ordine n.

Se indichiamo con $\mathcal{A} = (a)$ la matrice di ordine $n = 1$, avente una sola riga e una sola colonna, ovvero un solo elemento denotato con il simbolo a, il suo *determinante* è dato dall'elemento stesso della matrice, ovvero $\det \mathcal{A} = \det(a) = a$.

Prima di illustrare la "regola" del *determinante* di una matrice di ordine $n \geqslant 2$, introduciamo, riferendoci per i nostri scopi in particolare alle matrici quadrate, il concetto di *matrice minore complementare* di un elemento di una matrice.

Definizione 3.2.1. *(concetto di "matrice minore complementare") Data una matrice quadrata, la "matrice minore complementare" di un elemento della matrice è la matrice che si ottiene dalla matrice data togliendo la riga e la colonna dell'elemento.*

Vediamo degli esempi. Data la matrice

$$\mathcal{A} = \begin{pmatrix} 3 & -5 \\ 7 & 2 \end{pmatrix}, \tag{3.16}$$

la *matrice minore complementare* dell'elemento -5 sulla prima riga e seconda colonna è costituita dal solo elemento 7, evidenziato in carattere "grassetto", che si ottiene togliendo appunto tutta la riga e tutta la colonna dell'elemento -5, ovvero

$$\mathcal{A} = \begin{pmatrix} 3 & -5 \\ \mathbf{7} & 2 \end{pmatrix}.$$

Sempre nella matrice (3.16) la *matrice minore complementare* dell'elemento 3 sulla prima riga e prima colonna è costituita dal solo elemento 2, evidenziato in carattere "grassetto", che si ottiene togliendo appunto tutta la riga e tutta la colonna dell'elemento 3, ovvero

$$\mathcal{A} = \begin{pmatrix} 3 & -5 \\ 7 & \mathbf{2} \end{pmatrix}.$$

Data la matrice

$$\mathcal{A} = \begin{pmatrix} -2 & 4 & -7 \\ 9 & -5 & 8 \\ 3 & -1 & -6 \end{pmatrix}, \tag{3.17}$$

la *matrice minore complementare* del suo elemento 8 sulla seconda riga e terza colonna è la matrice indicata con $\mathcal{M}(8)$

$$\mathcal{M}(8) = \begin{pmatrix} -2 & 4 \\ 3 & -1 \end{pmatrix},$$

corrispondente agli elementi evidenziati in carattere "grassetto" che si ottengono togliendo nella matrice (3.17) tutta la riga e tutta la colonna dell'elemento 8, ovvero

$$\mathcal{A} = \begin{pmatrix} \mathbf{-2} & \mathbf{4} & -7 \\ 9 & -5 & 8 \\ \mathbf{3} & \mathbf{-1} & -6 \end{pmatrix}.$$

Sempre nella matrice (3.17) la *matrice minore complementare* del suo elemento 4 sulla prima riga e seconda colonna è la matrice indicata con $\mathcal{M}(4)$

$$\mathcal{M}(4) = \begin{pmatrix} 9 & 8 \\ 3 & -6 \end{pmatrix},$$

corrispondente agli elementi evidenziati in carattere "grassetto" che si ottengono togliendo nella matrice (3.17) tutta la riga e tutta la colonna dell'elemento 4, ovvero

$$\mathcal{A} = \begin{pmatrix} -2 & 4 & -7 \\ \mathbf{9} & -5 & \mathbf{8} \\ \mathbf{3} & -1 & \mathbf{-6} \end{pmatrix}.$$

Infine, sempre ancora nella matrice (3.17) la *matrice minore complementare* del suo elemento -5 sulla seconda riga e seconda colonna è la matrice indicata con $\mathcal{M}(-5)$

$$\mathcal{M}(-5) = \begin{pmatrix} -2 & -7 \\ 3 & -6 \end{pmatrix},$$

corrispondente agli elementi evidenziati in carattere "grassetto" che si ottengono togliendo nella matrice (3.17) tutta la riga e tutta la colonna dell'elemento -5, ovvero

$$\mathcal{A} = \begin{pmatrix} \mathbf{-2} & 4 & \mathbf{-7} \\ 9 & -5 & 8 \\ \mathbf{3} & -1 & \mathbf{-6} \end{pmatrix}.$$

Indicando dunque con il simbolo $\mathcal{M}(a_{ij})$ la *matrice minore complementare* dell'elemento a_{ij} sulla riga i-esima e sulla colonna j-esima di una data matrice quadrata $\mathcal{A}$ di ordine $n \geqslant 2$, il suo *determinante* si può ottenere in due modi. Se scegliamo una qualsiasi riga, chiamata riga h-esima, il *determinante* della matrice $\mathcal{A}$ si ottiene dallo sviluppo

$$\det \mathcal{A} = \sum_{k=1}^{n} (-1)^{h+k} a_{hk} \det \mathcal{M}(a_{hk}),$$

il cui significato è la somma sull'indice k di colonna, ovvero

$$\begin{aligned} \det \mathcal{A} &= \sum_{k=1}^{n} (-1)^{h+k} a_{hk} \det \mathcal{M}(a_{hk}) = \\ &= (-1)^{h+1} a_{h1} \det \mathcal{M}(a_{h1}) + (-1)^{h+2} a_{h2} \det \mathcal{M}(a_{h2}) + \dots + (-1)^{h+n} a_{hn} \det \mathcal{M}(a_{hn}). \end{aligned} \tag{3.18a}$$

Se scegliamo invece una qualsiasi colonna, chiamata colonna k-esima, il *determinante* della matrice $\mathcal{A}$ si ottiene anche dallo sviluppo

$$\det \mathcal{A} = \sum_{h=1}^{n} (-1)^{h+k} a_{hk} \det \mathcal{M}(a_{hk}),$$

il cui significato è la somma sull'indice h di riga, ovvero

$$\begin{aligned}\det \mathcal{A} &= \sum_{h=1}^{n} (-1)^{h+k} a_{hk} \det \mathcal{M}(a_{hk}) = \\ &= (-1)^{1+k} a_{1k} \det \mathcal{M}(a_{1k}) + (-1)^{2+k} a_{2k} \det \mathcal{M}(a_{2k}) + \cdots + (-1)^{n+k} a_{nk} \det \mathcal{M}(a_{nk}).\end{aligned} \tag{3.18b}$$

E' immediato verificare che il risultato del *determinante* di $\mathcal{A}$ ottenuto mediante lo sviluppo (3.18a) con qualsiasi riga oppure mediante lo sviluppo (3.18b) con qualsiasi colonna è sempre il medesimo e possiamo concludere quindi che il *determinante* di una data matrice quadrata $\mathcal{A}$ di ordine $n \geqslant 2$ non dipende dalla scelta della riga oppure della colonna mediante la quale esso viene calcolato.

Dalle espressioni (3.18) si riconosce immediatamente che la "scelta migliore" della riga oppure della colonna mediante la quale calcolare il determinante di una matrice quadrata è quella riga oppure quella colonna che abbia il maggiorn numero di elementi zero, ai quali, nelle somme (3.18) corrisponde un addendo nullo. E' "comodo" osservare che il *determinante* della generica matrice quadrata $\mathcal{A}$ di ordine 2

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

ottenuto, ad esempio, mediante lo sviluppo secondo la prima riga, è

$$\begin{aligned}\det \mathcal{A} &= (-1)^{1+1} a_{11} \det \mathcal{M}(a_{11}) + (-1)^{1+2} a_{12} \det \mathcal{M}(a_{12}) = \\ &= a_{11} \det(a_{22}) - a_{12} \det(a_{21}) = a_{11}a_{22} - a_{12}a_{21}\,,\end{aligned}$$

ovvero il *determinante* della generica matrice quadrata di ordine 2 è dato dalla differenza tra il prodotto dei due elementi sulla diagonale "che scende" e il prodotto dei due elementi sulla diagonale "che sale". Presentiamo ora degli esempi di calcolo del *determinante* di alcune matrici quadrate di ordine 2 e di ordine 3, utilizzando, in particolare per le matrici di *ordine* 3, uno dei due sviluppi equivalenti (3.18a) oppure (3.18b).

Data la matrice

$$\mathcal{A} = \begin{pmatrix} 2 & -4 \\ 5 & 3 \end{pmatrix},$$

il suo *determinante* è $\det \mathcal{A} = (2)(3) - (5)(-4) = 6 - (-20) = 6 + 20 = 26$.

Data la matrice

$$\mathcal{A} = \begin{pmatrix} -7 & -5 \\ -3 & 2 \end{pmatrix},$$

il suo *determinante* è $\det \mathcal{A} = (2)(-7) - (-5)(-3) = -14 - (15) = -29$.

Data la matrice

$$\mathcal{A} = \begin{pmatrix} 5 & -2 \\ 10 & 4 \end{pmatrix},$$

il suo *determinante* è $\det \mathcal{A} = (5)(4) - (10)(-2) = 20 - (-20) = 20 + 20 = 40$.

Data la matrice

$$\mathcal{A} = \begin{pmatrix} 5 & -2 \\ 10 & -4 \end{pmatrix},$$

il suo *determinante* è $\det \mathcal{A} = (5)(-4) - (10)(-2) = -20 - (-20) = -20 + 20 = 0$.

Dal confronto tra gli ultimi due esempi ricaviamo “facilmente” la proprietà per cui il *determinante* di una matrice quadrata di ordine 2 è zero *se e solo se* le due righe oppure le due colonne formano una proporzione (includendo ovviamente anche i segni).

Data la matrice

$$\mathcal{A} = \begin{pmatrix} 7 & -4 & -6 \\ -2 & 3 & 9 \\ 1 & -8 & 5 \end{pmatrix},$$

calcoliamo il suo *determinante* utilizzando, ad esempio, la prima riga e ottenendo dunque

$$\det \mathcal{A} = (-1)^{1+1} \cdot 7 \cdot \det \mathcal{M}(7) + (-1)^{1+2} \cdot (-4) \cdot \det \mathcal{M}(-4) + (-1)^{1+3} \cdot (-6) \cdot \det \mathcal{M}(-6) =$$

$$= 7 \det \begin{pmatrix} 3 & 9 \\ -8 & 5 \end{pmatrix} + 4 \det \begin{pmatrix} -2 & 9 \\ 1 & 5 \end{pmatrix} - 6 \det \begin{pmatrix} -2 & 3 \\ 1 & -8 \end{pmatrix} = (7)(87) + (4)(-19) - (6)(13) = 455.$$

Data la matrice

$$\mathcal{A} = \begin{pmatrix} -5 & 4 & -3 \\ 2 & -1 & -8 \\ -7 & 0 & 9 \end{pmatrix},$$

calcoliamo in questo caso il suo *determinante* utilizzando la terza riga, oppure anche la seconda colonna, perché nella terza riga e nella seconda colonna vi è un elemento nullo che quindi non dà contributo nello sviluppo del *determinante*.

Scegliendo in particolare la seconda colonna, otteniamo dunque il *determinante* di $\mathcal{A}$

$$\det \mathcal{A} = \left[(-1)^{1+2} \cdot 4 \cdot \det \mathcal{M}(4)\right] + \left[(-1)^{2+2} \cdot (-1) \cdot \det \mathcal{M}(-1)\right] + \left[(-1)^{3+2} \cdot 0 \cdot \det \mathcal{M}(0)\right] =$$

$$= -4 \det \begin{pmatrix} 2 & -8 \\ -7 & 9 \end{pmatrix} - 1 \det \begin{pmatrix} -5 & -3 \\ -7 & 9 \end{pmatrix} + 0 \det \mathcal{M}(0) = -4(18 - 56) - 1(-45 - 21) + 0 = 218.$$

Data la matrice

$$\mathcal{A} = \begin{pmatrix} -7 & 6 & -3 \\ 0 & 0 & -4 \\ -5 & 2 & 1 \end{pmatrix},$$

calcoliamo in questo caso il suo determinante utilizzando la seconda riga che possiede due elementi nulli in modo da ottenere dunque il *determinante* di $\mathcal{A}$

$$\det \mathcal{A} = \left[(-1)^{2+1} \cdot 0 \cdot \det \mathcal{M}(0)\right] + \left[(-1)^{2+2} \cdot 0 \cdot \det \mathcal{M}(0)\right] + \left[(-1)^{2+3} \cdot (-4) \cdot \det \mathcal{M}(-4)\right] =$$

$$= 4 \det \begin{pmatrix} -7 & 6 \\ -5 & 2 \end{pmatrix} = 4[(-7)(2) - (-5)(6)] = -48.$$

3.2.2 Rango di matrici

Per illustrare il concetto di *rango* di una qualsiasi matrice quadrata oppure rettangolare, dobbiamo introdurre preliminarmente il concetto di *matrice minore* contenuta in una data matrice. Data una matrice $\mathcal{A}$ e scelto in essa un ugual numero ammissibile di righe e di colonne, la matrice avente come elementi gli elementi delle righe e delle colonne scelte nella matrice $\mathcal{A}$ prende il nome di *matrice minore* contenuta in $\mathcal{A}$ corrispondente alle righe e alle colonne che sono state scelte.

E' immediato rendersi conto che la scelta di un ugual numero di righe e colonne all'interno di una data matrice $\mathcal{A}$ quadrata oppure rettangolare è una *scelta ammissibile* se nella matrice $\mathcal{A}$ esistono le righe e colonne scelte. In altre parole non si possono scegliere in una matrice $\mathcal{A}$ più righe né più colonne di quelle che la matrice $\mathcal{A}$ possieda.

Inoltre, poiché le *matrici minori* contenute all'interno di una data matrice $\mathcal{A}$ quadrata oppure rettangolare si ottengono sempre tramite la scelta di un ugual numero di righe e di colonne nella matrice $\mathcal{A}$, segue che, per loro definizione, le *matrici minori* contenute all'interno di una data matrice $\mathcal{A}$ quadrata oppure rettangolare sono sempre quadrate.

In particolare, una *matrice minore* contenuta all'interno di una data matrice $\mathcal{A}$ è di ordine n se essa viene "costruita" mediante la scelta di n righe e di n colonne.

Ad esempio, nella matrice

$$\mathcal{A} = \begin{pmatrix} -2 & 8 & -7 & -9 \\ 9 & -6 & -4 & 3 \\ -3 & -1 & -1 & 5 \end{pmatrix} \tag{3.19}$$

avente tre righe e quattro colonne, le *matrici minori* "più grandi" che si possono ottenere sono di ordine 3 perché nella matrice $\mathcal{A}$ si possono scegliere "al massimo" tre righe e tre colonne. Se scegliamo tre righe e tre colonne, non abbiamo nessuna "ambiguità" sulle tre righe e dobbiamo invece specificare quali tre righe scegliamo delle possibili quattro.

Se nella matrice (3.19) scegliamo le tre righe e le colonne prima, seconda e terza, otteniamo il minore di ordine 3 che indichiamo con il simbolo $\mathcal{M}_1^{(3\times3)}$

$$\mathcal{M}_1^{(3\times3)} = \begin{pmatrix} -2 & 8 & -7 \\ 9 & -6 & -4 \\ -3 & -1 & -1 \end{pmatrix},$$

il cui determinante è $\det \mathcal{M}_1^{(3\times3)} = 353$; se scegliamo le tre righe e le colonne prima, seconda e quarta, otteniamo il minore di ordine 3 che indichiamo con il simbolo $\mathcal{M}_2^{(3\times3)}$

$$\mathcal{M}_2^{(3\times3)} = \begin{pmatrix} -2 & 8 & -9 \\ 9 & -6 & 3 \\ -3 & -1 & 5 \end{pmatrix},$$

il cui determinante è $\det \mathcal{M}_2^{(3\times3)} = -135$, e così via per le altre *matrici minori* di ordine 3.

Se nella matrice (3.19) scegliamo due righe e due colonne, dobbiamo invece specificare sia la coppia di righe che la coppia di colonne. Se scegliamo le righe prima e terza e le colonne prima e quarta, otteniamo il minore di ordine 2 avente determinante -37

$$\begin{pmatrix} -2 & -9 \\ -3 & 5 \end{pmatrix};$$

se scegliamo le righe seconda e terza e le colonne seconda e quarta, otteniamo il minore di ordine 2 avente determinante -27

$$\begin{pmatrix} -6 & 3 \\ -1 & 5 \end{pmatrix}$$

e così via per le altre *matrici minori* di ordine 2. Infine le *matrici minori* di ordine 1 si ottengono scegliendo all'interno della matrice (3.19) una sola riga e una sola colonna.

Se scegliamo la seconda riga e la terza colonna, otteniamo il minore di ordine 1, che indichiamo con il simbolo $\mathcal{M}^{(1\times1)}$, avente il solo elemento -4, ovvero $\mathcal{M}^{(1\times1)} = (-4)$, il cui determinante è $\det \mathcal{M}^{(1\times1)} = -4$, e così via per le altre *matrici minori* di ordine 1.

Data una matrice $\mathcal{A}$ quadrata oppure rettangolare, la procedura per stabilire il *rango* della matrice $\mathcal{A}$ consiste nel calcolare il *determinante* delle *matrici minori* contenute nella matrice $\mathcal{A}$ partendo dalle *matrici minori* di ordine "più grande" fino a quando non si trovi la prima *matrice minore* avente determinante diverso da zero.

Definizione 3.2.2. *(concetto di "rango di matrice") Data una matrice $\mathcal{A}$ quadrata oppure rettangolare, l'ordine della prima matrice minore con determinante diverso da zero che si trovi nella matrice $\mathcal{A}$ partendo dalle sue matrici minori di ordine "più grande", prende il nome di "rango" della matrice $\mathcal{A}$.*

E' molto importante puntualizzare che nella definizione 3.2.2 del concetto di *rango di matrice* il termine *prima matrice minore* dà luogo ad un'ambiguità perché quale sia la *prima matrice minore* con determinante diverso da zero che si trovi nella matrice $\mathcal{A}$ partendo dalle sue *matrici minori* di ordine "più grande", dipende dalla "soggettività" di chi calcola il *rango* della matrice assegnata. Tuttavia tale "ambiguità" è del tutto irrilevante perché tutte le possibili *matrici minori* in virtù delle quali si può stabilire il *rango* di una matrice assegnata, hanno il medesimo ordine e pertanto forniscono il medesimo *rango*.

Quando risolveremo i sistemi di equazioni lineari e dovremo calcolare i *ranghi* di particolari matrici, vedremo che, per una data matrice, il calcolo del suo *rango* basato su una possibile *matrice minore* $\mathcal{M}_1$ oppure basato su un'altra possibile *matrice minore* $\mathcal{M}_2$ diversa da $\mathcal{M}_1$, è del tutto equivalente perché entrambe le possibili scelte delle *matrici minori* $\mathcal{M}_1, \mathcal{M}_2$ forniscono le medesime soluzioni del sistema assegnato.

Talvolta il termine *rango* viene sostituito dal termine *caratteristica* e quindi il concetto di *rango di una matrice* può anche essere chiamato, come nel testo di riferimento, *caratteristica di una matrice.* Utilizzando in questa nota la terminologia *rango di matrice*, presentiamo ora degli esempi di calcolo del *rango* di alcune matrici.

Esempio 1. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} 0 & -5 \end{pmatrix}$$

avente una riga e due colonne, le cui *matrici minori* "più grandi" sono quelle di ordine 1.

Il determinante della *matrice minore* corrispondente alla scelta dell'unica riga e della prima colonna è zero perché tale *matrice minore* di ordine 1 ha il solo elemento 0.

Poiché la prima *matrice minore* che troviamo con determinante diverso da zero all'interno della matrice $\mathcal{A}$ è la *matrice minore* di ordine 1 avente il solo elemento -5 e quindi determinante $-5 \neq 0$, corrispondente alla scelta dell'unica riga e della seconda colonna, concludiamo che la matrice $\mathcal{A}$ possiede *rango* 1.

Esempio 2. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} 3 & -2 \\ 6 & -4 \end{pmatrix}$$

avente due righe e due colonne e quindi un'unica *matrice minore* di ordine 2 coincidente con la matrice $\mathcal{A}$ stessa, che è anche ovviamente la *matrice minore* "più grande" contenuta nella matrice $\mathcal{A}$. Partendo dalla *matrice minore* "più grande", ovvero da $\mathcal{A}$ stessa, abbiamo che il suo determinante è $\det \mathcal{A} = 0$ e poiché dunque la prima *matrice minore* che troviamo con determinante diverso da zero all'interno della matrice $\mathcal{A}$ è la *matrice minore* di ordine 1 avente il solo elemento 3 e quindi determinante $3 \neq 0$, corrispondente alla scelta della prima riga e prima colonna, concludiamo che la matrice $\mathcal{A}$ possiede *rango* 1.

Osserviamo che se avessimo calcolato il *rango* prendendo come *prima matrice minore* con determinante diverso da zero che si trova nella matrice $\mathcal{A}$, la *matrice minore* corrispondente alla scelta della prima riga e seconda colonna, avremmo utilizzato la *matrice minore* contenente il solo elemento -2 e avente quindi sempre ordine 1, da cui avremmo concluso ugualmente che la matrice $\mathcal{A}$ possiede *rango* 1.

Esempio 3. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} -4 & 5 \\ 3 & 7 \end{pmatrix}$$

avente due righe e due colonne e quindi un'unica *matrice minore* di ordine 2 coincidente con la matrice $\mathcal{A}$ stessa, che è anche ovviamente la *matrice minore* "più grande" contenuta nella matrice $\mathcal{A}$. Poiché dunque, partendo dalla *matrice minore* "più grande", ovvero da $\mathcal{A}$ stessa, la prima *matrice minore* che troviamo con determinante diverso da zero all'interno della matrice $\mathcal{A}$ è la *matrice minore* di ordine 2 coincidente con la matrice $\mathcal{A}$ stessa, il cui determinante è appunto $\det \mathcal{A} = -43 \neq 0$, concludiamo che la matrice $\mathcal{A}$ possiede *rango* 2.

Esempio 4. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} 3 & -2 & -1 \\ 6 & -4 & 5 \end{pmatrix}$$

avente due righe e tre colonne, le cui *matrici minori* "più grandi" sono di ordine 2.

Partendo dalle *matrici minori* "più grandi", abbiamo che il determinante della *matrice minore* corrispondente alla scelta delle due righe e delle prime due colonne è

$$\det \begin{pmatrix} 3 & -2 \\ 6 & -4 \end{pmatrix} = 0$$

e abbiamo che il determinante della *matrice minore* corrispondente alla scelta delle due righe e delle colonne seconda e terza è

$$\det \begin{pmatrix} -2 & -1 \\ -4 & 5 \end{pmatrix} = -14 \neq 0,$$

da cui segue quindi che la matrice $\mathcal{A}$ possiede *rango* 2.

Come per l'esempio 2, osserviamo che se avessimo calcolato il *rango* prendendo come *prima matrice minore* con determinante diverso da zero che si trova nella matrice $\mathcal{A}$, la *matrice minore* corrispondente alla scelta delle due righe e delle colonne prima e terza, avremmo utilizzato la *matrice minore* avente determinante $21 \neq 0$ data da

$$\begin{pmatrix} 3 & -1 \\ 6 & 5 \end{pmatrix},$$

il cui ordine è sempre 2 ed in virtù della quale avremmo concluso quindi ugualmente che la matrice $\mathcal{A}$ possiede *rango* 2.

Esempio 5. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} -2 & 4 \\ 3 & -6 \\ 8 & -7 \end{pmatrix}$$

avente tre righe e due colonne, le cui *matrici minori* "più grandi" sono di ordine 2.

Partendo dalle *matrici minori* "più grandi", abbiamo che il determinante della *matrice minore* corrispondente alla scelta delle prime due righe e delle due colonne è

$$\det\begin{pmatrix} -2 & 4 \\ 3 & -6 \end{pmatrix} = 0$$

e abbiamo che il determinante della *matrice minore* corrispondente alla scelta delle righe prima e terza e delle due colonne è

$$\det\begin{pmatrix} -2 & 4 \\ 8 & -7 \end{pmatrix} = 14 - 32 = -18 \neq 0,$$

da cui segue quindi che la matrice $\mathcal{A}$ possiede *rango* 2.

Esempio 6. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} -8 & 6 & -4 \\ 4 & -3 & 2 \end{pmatrix}$$

avente due righe e tre colonne, le cui *matrici minori* "più grandi" sono di ordine 2.

Partendo dalle *matrici minori* "più grandi", abbiamo che tutte le *matrici minori* di ordine 2 hanno determinante nullo perché risulta

$$\det\begin{pmatrix} -8 & 6 \\ 4 & -3 \end{pmatrix} = 0, \qquad \det\begin{pmatrix} -8 & -4 \\ 4 & 2 \end{pmatrix} = 0, \qquad \det\begin{pmatrix} 6 & -4 \\ -3 & 2 \end{pmatrix} = 0$$

e poiché la prima *matrice minore* che troviamo con determinante diverso da zero all'interno della matrice $\mathcal{A}$ è la *matrice minore* di ordine 1 avente il solo elemento -8 e quindi determinante $-8 \neq 0$, corrispondente alla scelta della prima riga e prima colonna, possiamo concludere che la matrice $\mathcal{A}$ possiede *rango* 1.

Esempio 7. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} 3 & -4 & 2 & 5 \\ -6 & 8 & -4 & 7 \end{pmatrix}$$

avente due righe e quattro colonne, le cui *matrici minori* "più grandi" sono di ordine 2.

Partendo dalle *matrici minori* "più grandi", abbiamo che le *matrici minori* di ordine 2

$$\begin{pmatrix} 3 & -4 \\ -6 & 8 \end{pmatrix}, \qquad \begin{pmatrix} 3 & 2 \\ -6 & -4 \end{pmatrix}, \qquad \begin{pmatrix} -4 & 2 \\ 8 & -4 \end{pmatrix}$$

hanno determinante nullo e che la prima *matrice minore* con determinante diverso da zero all'interno della matrice $\mathcal{A}$ che troviamo è la *matrice minore* di ordine 2

$$\begin{pmatrix} 3 & 5 \\ -6 & 7 \end{pmatrix},$$

corrispondente alla scelta delle due righe e delle due colonne prima e quarta, il cui determinante è

$$\det\begin{pmatrix} 3 & 5 \\ -6 & 7 \end{pmatrix} = 51 \neq 0,$$

da cui segue quindi che la matrice $\mathcal{A}$ possiede *rango* 2.

Esempio 8. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} 3 & -4 & 2 \\ 6 & 8 & -9 \\ 1 & 5 & -7 \end{pmatrix}$$

avente tre righe e tre colonne e quindi un'unica *matrice minore* di ordine 3 coincidente con la matrice $\mathcal{A}$ stessa, che è anche ovviamente la *matrice minore* "più grande" contenuta nella matrice $\mathcal{A}$. Poiché dunque, partendo dalla *matrice minore* "più grande", ovvero da $\mathcal{A}$ stessa, la prima *matrice minore* che troviamo con determinante diverso da zero all'interno della matrice $\mathcal{A}$ è la *matrice minore* di ordine 3 coincidente con la matrice $\mathcal{A}$ stessa, il cui determinante è $\det \mathcal{A} = -113 \neq 0$, concludiamo che la matrice $\mathcal{A}$ possiede *rango* 3.

Esempio 9. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} 3 & 6 & -4 \\ 4 & 8 & -7 \\ 2 & 4 & -1 \end{pmatrix}$$

avente tre righe e tre colonne e quindi un'unica *matrice minore* di ordine 3 coincidente con la matrice $\mathcal{A}$ stessa, che è anche ovviamente la *matrice minore* "più grande" contenuta nella matrice $\mathcal{A}$. Partendo dalla *matrice minore* "più grande", ovvero da $\mathcal{A}$ stessa, abbiamo che il suo determinante è $\det \mathcal{A} = 0$ e poiché la prima *matrice minore* che troviamo con determinante diverso da zero all'interno della matrice $\mathcal{A}$ è la *matrice minore* di ordine 2

$$\mathcal{M} = \begin{pmatrix} 6 & -4 \\ 8 & -7 \end{pmatrix},$$

corrispondente alla scelta delle righe prima e seconda e delle colonne seconda e terza, avente determinante $\det \mathcal{M} = -10 \neq 0$, concludiamo che la matrice $\mathcal{A}$ possiede *rango* 2.

Esempio 10. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} 4 & 8 & -7 & 3 \\ 3 & 6 & -4 & 5 \\ 2 & 4 & -1 & 7 \end{pmatrix}$$

avente tre righe e quattro colonne le cui *matrici minori* "più grandi" sono di ordine 3.

Partendo dalle *matrici minori* "più grandi", abbiamo che tutte le *matrici minori* di ordine 3 hanno determinante nullo perché risulta

$$\det \begin{pmatrix} 4 & 8 & -7 \\ 3 & 6 & -4 \\ 2 & 4 & -1 \end{pmatrix} = \det \begin{pmatrix} 4 & 8 & 3 \\ 3 & 6 & 5 \\ 2 & 4 & 7 \end{pmatrix} = \det \begin{pmatrix} 4 & -7 & 3 \\ 3 & -4 & 5 \\ 2 & -1 & 7 \end{pmatrix} = \det \begin{pmatrix} 8 & -7 & 3 \\ 6 & -4 & 5 \\ 4 & -1 & 7 \end{pmatrix} = 0$$

e poiché la prima *matrice minore* che troviamo con determinante diverso da zero all'interno della matrice $\mathcal{A}$ è la *matrice minore* di ordine 2

$$\mathcal{M} = \begin{pmatrix} 4 & -7 \\ 3 & -4 \end{pmatrix},$$

corrispondente alla scelta delle righe prima e seconda e delle colonne prima e terza, avente determinante $\det \mathcal{M} = 5 \neq 0$, concludiamo che la matrice $\mathcal{A}$ possiede *rango* 2.

Esempio 11. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} 4 & 8 & -7 & 3 \\ 3 & 6 & -4 & -5 \\ 2 & 4 & -1 & -7 \end{pmatrix}$$

avente tre righe e quattro colonne le cui *matrici minori* "più grandi" sono di ordine 3.

Partendo dalle *matrici minori* "più grandi", abbiamo che le *matrici minori* di ordine 3

$$\begin{pmatrix} 4 & 8 & -7 \\ 3 & 6 & -4 \\ 2 & 4 & -1 \end{pmatrix} \qquad \text{e} \qquad \begin{pmatrix} 4 & 8 & 3 \\ 3 & 6 & -5 \\ 2 & 4 & -7 \end{pmatrix}$$

hanno determinante nullo e che la *matrice minore* di ordine 3

$$\det \begin{pmatrix} 4 & -7 & 3 \\ 3 & -4 & -5 \\ 2 & -1 & -7 \end{pmatrix}$$

ha invece determinante $30 \neq 0$. Poiché la prima *matrice minore* che troviamo con determinante diverso da zero all'interno di $\mathcal{A}$ è la *matrice minore* di ordine 3, corrispondente alla scelta delle tre righe e delle colonne prima, terza e quarta, concludiamo dunque che la matrice $\mathcal{A}$ possiede *rango* 3.

Esempio 12. Consideriamo la matrice

$$\mathcal{A} = \begin{pmatrix} -2 & 3 & 4 & 2 \\ 4 & -6 & -8 & -4 \\ -6 & 9 & 12 & 6 \end{pmatrix}$$

avente tre righe e quattro colonne le cui *matrici minori* "più grandi" sono di ordine 3.

Poiché tutte le *matrici minori* di ordine 3 e tutte le *matrici minori* di ordine 2 hanno determinante nullo, concludiamo dunque che la matrice $\mathcal{A}$ possiede *rango* 1.

Esempio 13. Considerando le tre matrici

$$\mathcal{A}_1 = \begin{pmatrix} -2 & 3 & -1 \\ 5 & -6 & 4 \\ -6 & 9 & -3 \\ 7 & -8 & 6 \end{pmatrix}, \qquad \mathcal{A}_2 = \begin{pmatrix} -2 & 3 & -1 \\ -5 & 6 & -4 \\ -6 & 9 & -3 \\ 7 & -8 & 5 \end{pmatrix}, \qquad \mathcal{A}_3 = \begin{pmatrix} -2 & 4 & 6 \\ 5 & -10 & -15 \\ -1 & 2 & 3 \\ 3 & -6 & -9 \end{pmatrix},$$

si lascia come semplice esercizio per lo studente la verifica che la matrice $\mathcal{A}_1$ ha *rango* 2, la matrice $\mathcal{A}_2$ ha *rango* 3, la matrice $\mathcal{A}_3$ ha *rango* 1.

3.3 Sistemi lineari e teorema di Rouché-Capelli

Dato un sistema di equazioni lineari (3.14), avente n equazioni e k incognite e associate ad esso le due matrici A, B riportate nelle (3.15), dobbiamo stabilire prima di tutto se il sistema possiede soluzioni oppure no e stabilire in secondo luogo, nel caso in cui le soluzioni esistano, quante siano tali soluzioni. Quest'analisi preliminare di ogni sistema si esegue applicando il *teorema di Rouché-Capelli*, che possiamo considerare "idealmente" suddiviso in due parti, la prima delle quali stabilisce se il sistema di equazioni possiede soluzioni, mentre la seconda stabilisce, nel caso in cui il sistema abbia soluzioni, quante soluzioni il sistema possieda.

Teorema 3.3.1. ***(Teorema di Rouché-Capelli)*** *Un sistema di equazioni lineari* (3.14) *possiede soluzioni se e solo se le due matrici A,B ad esso associate e riportate nelle* (3.15) *possiedono il medesimo rango. Se i ranghi delle due matrici A,B riportate nelle* (3.15) *sono uguali tra loro e coincidono con il numero n delle incognite, segue che il sistema di equazioni lineari* (3.14) *possiede soluzione unica. Se i ranghi delle due matrici* (3.15) *sono il medesimo numero r che risulta minore del numero n delle incognite, segue invece che il sistema di equazioni lineari* (3.14) *possiede infinite soluzioni, ovvero* ∞^{n-r} *soluzioni.*

Il simbolo ∞^{n-r} che esprime il numero delle soluzioni del sistema (3.14) quando queste siano infinite, si legge *infinito alla n-r* ed il suo significato sarà illustrato tramite la procedura risolutiva di un sistema (3.14) che possieda appunto infinite soluzioni.

Un sistema (3.14) si studia dunque associando al sistema le due matrici A, B riportate nelle (3.15) e calcolando i *ranghi* di tali matrici che indicheremo con $rg(A)$ e $rg(B)$.

Se il *rango* della matrice A risulta diverso dal *rango* della matrice B, concludiamo quindi che il sistema, in virtù del *teorema di Rouché-Capelli*, non possiede soluzione.

Se risulta invece $rg(A) = rg(B) = r$, il sistema, in virtù del *teorema di Rouché-Capelli*, possiede soluzioni che vengono calcolate attraverso i seguenti passi:

- si considera la *matrice minore* comune $\mathcal{M}$, contenuta nelle due matrici A, B, avente ordine r e ovviamente determinante diverso da zero, dalla quale è stato ottenuto appunto il medesimo *rango* r delle due matrici A, B;

- a tale *matrice minore* comune $\mathcal{M}$ si associa la corrispondente "parte" del sistema costituita da r equazioni ed r incognite e si eliminano tutte quelle equazioni "al di fuori" della *matrice minore* comune $\mathcal{M}$, riportata corrispondentemente nel sistema;

- le equazioni che non sono state eliminate debbono essere riscritte nella forma in cui tutte le r incognite corrispondenti alla *matrice minore* comune $\mathcal{M}$ siano riportate al primo membro e tutte le rimanenti $k - r$ incognite "al di fuori" della *matrice minore* comune $\mathcal{M}$ siano "trasportate" al secondo membro;

- ad ognuna delle $n - r$ incognite "trasportate" al secondo membro si attribuisce un valore arbitrario diverso da incognita ad incognita e mediante la *regola di Cramer* si calcolano le r incognite "rimaste" al primo membro e corrispondenti alla *matrice minore* comune $\mathcal{M}$, riportata nel sistema.

Se i *ranghi* delle due matrici sono uguali tra loro, ovvero $rg(A) = rg(B) = r$, e coincidono con il numero delle incognite, ovvero $rg(A) = rg(B) = r = n$, segue che la *matrice minore* comune $\mathcal{M}$, contenuta nelle due matrici A, B e avente determinante diverso da zero, dalla quale è stato ottenuto appunto il medesimo *rango*, ha ordine $r = n$ e pertanto la "parte" del sistema corrispondente a tale *matrice minore* comune $\mathcal{M}$ contiene tutte le incognite. Se quindi le equazioni del sistema, "contenute" nella "parte" corrispondente alla *matrice minore* comune $\mathcal{M}$, vengono riscritte con tutte le $n = r$ incognite al primo membro, segue che nessuna incognita viene "trasportata" al secondo membro e che a nessuna incognita viene attribuito un valore arbitrario, in modo tale che dunque la soluzione del sistema, ottenuta mediante la *regola di Cramer*, sia unica.

Se invece risulta $rg(A) = rg(B) = r < n$, segue la *matrice minore* comune $\mathcal{M}$, contenuta nelle due matrici A, B e avente determinante diverso da zero, dalla quale è stato ottenuto appunto il medesimo *rango*, ha ordine $r < n$ e pertanto la "parte" del sistema corrispondente a tale *matrice minore* comune $\mathcal{M}$ non contiene tutte le incognite, ma contiene solo $r < n$ incognite in $n-r$ equazioni. Se quindi le $n-r$ equazioni del sistema, "contenute" nella "parte" del sistema corrispondente alla *matrice minore* comune $\mathcal{M}$, vengono riscritte con r incognite al primo membro e con le rimanenti $n-r$ incognite al secondo membro, segue che alle $n-r$ incognite "trasportate" al secondo membro, indicate con i simboli $x_{i_1}, x_{i_2}, \ldots, x_{i_{n-r}}$, si può attribuire un valore arbitrario, ovvero

$$x_{i_1} = \alpha_{i_1}, \qquad x_{i_2} = \alpha_{i_2}, \qquad x_{i_3} = \alpha_{i_3}, \qquad \ldots, \qquad x_{i_{n-r}} = \alpha_{i_{n-r}}$$

e che quindi da ognuna delle possibili infinite scelte $x_{i_1} = \alpha_{i_1}, x_{i_2} = \alpha_{i_2}, \ldots, x_{i_{n-r}} = \alpha_{i_{n-r}}$ discende una soluzione, in modo tale dunque che le soluzioni del sistema, ottenuta sempre mediante la *regola di Cramer*, siano infinite. Quando le soluzioni di un sistema sono infinite, si dice che il loro numero è ∞^{n-r} per indicare quante sono le $n-r$ incognite del sistema "trasportate" al secondo membro, le quali, tramite i valori arbitrari che possono essere attribuiti loro, generano appunto tutte le infinite soluzioni del sistema stesso.

Per illustrare concretamente come si studia e si risolve un sistema di equazioni lineari, presentiamo ora quindi alcuni esempi di calcolo.

Esempio 1. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y = -4 \\ 6x - 4y = 5, \end{cases}$$

con $m = 2$ equazioni ed $n = 2$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 \\ 6 & -4 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 \\ 6 & -4 & 5 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 1$ e $rg(B) = 2$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) \neq rg(B)$, il sistema non possiede soluzione.

Esempio 2. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y = -4 \\ 6x - 4y = -8, \end{cases} \tag{3.20}$$

con $m = 2$ equazioni ed $n = 2$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 \\ 6 & -4 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 \\ 6 & -4 & -8 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 1$ e $rg(B) = 1$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) = rg(B)$, il sistema possiede soluzioni e che, poiché il *rango* $r = 1$ delle due matrici è minore del numero $n = 2$ delle incognite, le soluzioni del sistema sono ∞^{2-1}, ovvero ∞^1. Per calcolare le ∞^1 soluzioni, osserviamo che la *matrice minore* comune $\mathcal{M}$ contenuta in entrambe le matrici A, B, in virtù della quale risulta $rg(A) = rg(B) = 1$, è quella avente ordine 1 evidenziata, ovvero

$$A = \begin{pmatrix} 3 & -2 \\ 6 & -4 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 \\ 6 & -4 & -8 \end{pmatrix}, \tag{3.21a}$$

la quale corrisponde alla parte del sistema evidenziata

$$\begin{cases} 3x - 2y = -4 \\ 6x - 4y = -8. \end{cases} \tag{3.21b}$$

Se "eliminiamo" la seconda equazione che sta "al di fuori" della *matrice minore* $\mathcal{M}$ comune evidenziata nel sistema e riscriviamo la prima equazione con l'incognita x al primo membro, in quanto contenuta nella *matrice minore* comune $\mathcal{M}$ evidenziata nel sistema, e con l'incognita y al secondo membro, in quanto "al di fuori" della *matrice minore* $\mathcal{M}$ evidenziata nel sistema, otteniamo il "nuovo" sistema "estratto" dal sistema iniziale

$$\left\{\ 3x = 2y - 4,\right. \tag{3.22}$$

nel quale attribuiamo un valore arbitrario $y = \alpha$ all'incognita y "trasportata" al secondo membro in quanto "al di fuori" della *matrice minore* comune $\mathcal{M}$. Il "sistema finale" (3.22) ha la soluzione $x = (2\alpha - 4)/3$, che, unita al valore di y, dà le ∞^1 soluzioni

$$\begin{cases} x = (2\alpha - 4)/3 \\ y = \alpha \end{cases} \tag{3.23a}$$

di tutto il sistema iniziale (3.20), tra le quali vi sono, ad esempio, le soluzioni generate dalla scelta dei vari valori di α

$$\begin{cases} x = -4/3 \\ y = 0 \end{cases}, \quad \begin{cases} x = -2/3 \\ y = 1 \end{cases}, \quad \begin{cases} x = 0 \\ y = 2 \end{cases}, \quad \begin{cases} x = 2/3 \\ y = 3 \end{cases}, \quad \begin{cases} x = 4/3 \\ y = 4 \end{cases}, \tag{3.23b}$$

generate, in particolare, dalla scelta dei valori rispettivamente $\alpha = 0, 1, 2, 3, 4$, e così via.

E' molto importante puntualizzare che la *matrice minore* $\mathcal{M}$ comune alle matrici A, B, evidenziata nelle due matrici (3.21a) e quindi riportata corrispondentemente nel sistema, come mostrato nella (3.21b), non è l'unica *matrice minore* comune alle due matrici A, B dalla quale si ricava l'uguaglianza dei due *ranghi* $rg(A) = rg(B) = 1$.

Com'è immediato rendersi conto, abbiamo infatti che l'uguaglianza dei due *ranghi* $rg(A) = rg(B) = 1$ è data anche, ad esempio, dalla *matrice minore* $\mathcal{M}$ comune alle due matrici A, B relativa alla prima riga e seconda colonna, come evidenziato qui di seguito

$$A = \begin{pmatrix} 3 & -2 \\ 6 & -4 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 \\ 6 & -4 & -8 \end{pmatrix}, \tag{3.24a}$$

che nel sistema corrisponde alla parte evidenziata relativa all'incognita y nella prima equazione

$$\begin{cases} 3x - 2y = -4 \\ 6x - 4y = -8. \end{cases} \tag{3.24b}$$

Se utilizziamo la *matrice minore* $\mathcal{M}$ comune alle matrici A, B evidenziata nelle (3.24), "eliminiamo" sempre la seconda equazione in quanto non contenuta nella *matrice minore* comune, ma nella prima equazione "estraiamo" l'incognita y, che "lasciamo" al primo membro, e "trasportiamo" al secondo membro l'incognita x, in quanto "al di fuori" della *matrice minore* comune, attribuendole valore arbitrario che, per evitare "confusioni", indichiamo con β, ovvero $x = \beta$. In questo caso il sistema "estratto" dal sistema (3.20) è

$$\left\{\ -2y = -3\beta - 4,\right.$$

la cui soluzione è $y = (3\beta + 4)/2$, la quale, unita al valore arbitrario $x = \beta$, dà le infinite soluzioni

$$\begin{cases} x = \beta \\ y = (3\beta + 4)/2 \end{cases} \tag{3.25}$$

di tutto il sistema iniziale (3.20). Lasciando come semplice esercizio la verifica per i valori $\beta = -4/3, -2/3, 0, 2/3, 4/3$ la seconda forma (3.25) delle soluzioni riproduce le medesime soluzioni (3.23b) ottenute dalla prima forma delle soluzioni (3.23a), ci "convinciamo" che tutte le ∞^1 soluzioni generate dai valori di α attribuiti nell'espressione (3.23a) coincidono con le corrispondenti ∞^1 soluzioni generate dai valori di β attribuiti nell'espressione (3.25).

Pertanto concludiamo che la scelta della *matrice minore* comune alle due matrici A, B, in virtù della quale si ottiene l'uguaglianza dei due *ranghi* $rg(A) = rg(B)$, è del tutto irrilevante perché se l'uguaglianza dei due *ranghi* $rg(A) = rg(B)$ può essere stabilita mediante due diverse *matrici minori* $\mathcal{M}_1, \mathcal{M}_2$ comuni alle due matrici A, B, segue che le soluzioni ottenute risolvendo il sistema "estratto" corrispondente alla *matrice minore* $\mathcal{M}_1$ hanno, sì, espressione diversa dall'espressione delle soluzioni ottenute risolvendo il sistema "estratto" in corrispondenza della scelta della *matrice minore* $\mathcal{M}_2$, ma tuttavia le n-uple generate da entrambe le espressioni coincidono.

Nel seguito, quindi, risolveremo i sistemi sempre effettuando una scelta particolare della *matrice minore* $\mathcal{M}$ comune alle due matrici A, B da cui segua l'uguaglianza dei due *ranghi* $rg(A) = rg(B)$, ma se si risolve il sistema utilizzando una diversa *matrice minore* $\mathcal{M}$ comune alle due matrici A, B da cui si ricavi la medesima uguaglianza $rg(A) = rg(B)$ dei due *ranghi*, si ottengono in ogni caso le medesime soluzioni.

Esempio 3. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y = -4 \\ 6x - 5y = 8, \end{cases}$$

con $m = 2$ equazioni ed $n = 2$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 \\ 6 & -5 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 \\ 6 & -5 & 8 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 2$ e $rg(B) = 2$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) = rg(B)$, il sistema possiede soluzioni e che, poiché il *rango* $r = 2$ delle due matrici coincide con il numero $n = 2$ delle incognite, la soluzione del sistema è unica. Per calcolare tale soluzione unica, osserviamo che la *matrice minore* comune $\mathcal{M}$ contenuta in entrambe le matrici A, B, in virtù della quale risulta $rg(A) = rg(B) = 2$, è quella avente ordine 2 evidenziata, ovvero

$$A = \left(\boxed{\begin{matrix} 3 & -2 \\ 6 & -5 \end{matrix}} \right) \qquad \text{e} \qquad B = \left(\boxed{\begin{matrix} 3 & -2 \\ 6 & -5 \end{matrix}} \begin{matrix} -4 \\ 8 \end{matrix} \right),$$

che nel sistema corrisponde alla "parte" evidenziata

$$\left\{ \boxed{\begin{matrix} 3x - 2y \\ 6x - 5y \end{matrix}} \begin{matrix} = -4 \\ = 8. \end{matrix} \right.$$

In questo caso la *matrice minore* comune $\mathcal{M}$, riportata corrispondentemente sul sistema, contiene tutte le incognite in tutte le equazioni, da cui segue che non "trasportiamo" nessuna incognita al secondo membro e che quindi la soluzione del sistema è unica. Poiché inoltre non "eliminiamo" nessuna equazione, calcoliamo la soluzione unica dal sistema iniziale

$$\begin{cases} 3x - 2y = -4 \\ 6x - 5y = 8, \end{cases} \tag{3.26}$$

che risolviamo mediante la *regola di Cramer*. La *regola di Cramer* si applica al sistema riscritto in cui le incognite relative alla *matrice minore* comune $\mathcal{M}$ sono "lasciate" al primo membro e le incognite "al di fuori" della *matrice minore* comune $\mathcal{M}$ sono "trasportate" al secondo membro. La *regola di Cramer* consiste nello scrivere ogni incognita "rimasta" al primo membro come frazione in cui al denominatore vi sia il determinante della matrice dei coefficienti (del sistema a cui si applica appunto la *regola di Cramer*) ed al numeratore il determinante della medesima matrice in cui la colonna corrispondente all'incognita calcolata sia sostituita dalla colonna dei termini noti.

Per il sistema (3.26) in oggetto otteniamo dunque la soluzione unica

$$x = \frac{\det\begin{pmatrix} -4 & -2 \\ 8 & -5 \end{pmatrix}}{\det\begin{pmatrix} 3 & -2 \\ 6 & -5 \end{pmatrix}} = -12 \qquad \text{e} \qquad y = \frac{\det\begin{pmatrix} 3 & -4 \\ 6 & 8 \end{pmatrix}}{\det\begin{pmatrix} 3 & -2 \\ 6 & -5 \end{pmatrix}} = -16,$$

ovvero $x = -12, y = -16$, che, come si riconosce immediatamente, verifica entrambe le equazioni del sistema perché risulta $3(-12) - 2(-16) = -4$ e $6(-12) - 5(-16) = 8$.

Esempio 4. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y - z = 5 \\ 6x - 4y + 3z = -7, \end{cases} \tag{3.27}$$

con $m = 2$ equazioni ed $n = 3$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 & -1 \\ 6 & -4 & 3 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -1 & 5 \\ 6 & -4 & 3 & -7 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 2$ e $rg(B) = 2$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) = rg(B)$, il sistema possiede soluzioni e che, poiché il *rango* $r = 2$ delle due matrici è minore del numero $n = 3$ delle incognite, le soluzioni del sistema sono $\infty^{3-2} = \infty^1$.

Per calcolare tali infinite soluzioni, osserviamo che la *matrice minore* comune $\mathcal{M}$ contenuta in entrambe le matrici A, B, in virtù della quale risulta $rg(A) = rg(B) = 2$, è quella avente ordine 2 evidenziata, ovvero

$$A = \begin{pmatrix} 3 & \boxed{\begin{matrix} -2 & -1 \\ -4 & 3 \end{matrix}} \\ 6 & \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & \boxed{\begin{matrix} -2 & -1 \\ -4 & 3 \end{matrix}} & 5 \\ 6 & & -7 \end{pmatrix},$$

che nel sistema corrisponde alla "parte" evidenziata

$$\begin{cases} 3x \boxed{-2y - z} = 5 \\ 6x \boxed{-4y + 3z} = -7. \end{cases}$$

In questo caso la *matrice minore* comune $\mathcal{M}$, riportata corrispondentemente sul sistema, contiene le incognite y, z in entrambe le equazioni, da cui segue che "trasportiamo" l'incognita x al secondo membro attribuendole il valore arbitrario $x = \alpha$ e che dunque il sistema possiede appunto ∞^1 soluzioni. Se riscriviamo le due equazioni "lasciando" le due incognite y, z al primo membro, in quanto contenute nella *matrice minore* comune $\mathcal{M}$

evidenziata nel sistema, e con l'incognita $x = \alpha$ al secondo membro, in quanto "al di fuori" della *matrice minore* $\mathcal{M}$ evidenziata nel sistema, otteniamo il "nuovo" sistema "estratto" dal sistema iniziale

$$\begin{cases} -2y - z = 5 - 3\alpha \\ -4y + 3z = -7 - 6\alpha, \end{cases} \tag{3.28}$$

che risolviamo mediante la *regola di Cramer*, illustrata per la soluzione dell'esercizio 3, per ricavare i valori delle incognite y, z. Dal sistema (3.28) otteniamo la soluzione

$$y = \frac{\det\begin{pmatrix} 5-3\alpha & -1 \\ -7-6\alpha & 3 \end{pmatrix}}{\det\begin{pmatrix} -2 & -1 \\ -4 & 3 \end{pmatrix}} = \frac{15\alpha - 8}{10} \qquad \text{e} \qquad z = \frac{\det\begin{pmatrix} -2 & 5-3\alpha \\ -4 & -7-6\alpha \end{pmatrix}}{\det\begin{pmatrix} -2 & -1 \\ -4 & 3 \end{pmatrix}} = -\frac{17}{5},$$

la quale, unita al valore arbitrario di x, dà le ∞^1 soluzioni di tutto il sistema iniziale (3.27)

$$\begin{cases} x = \alpha \\ y = (15\alpha - 8)/10 \\ z = -17/5 \end{cases},$$

tra le quali vi sono, ad esempio, le soluzioni generate dalla scelta dei vari valori di α

$$\begin{cases} x = 0 \\ y = -4/5 \\ z = -17/5 \end{cases}, \qquad \begin{cases} x = 1 \\ y = 7/10 \\ z = -17/5 \end{cases}, \qquad \begin{cases} x = 2 \\ y = 11/5 \\ z = -17/5 \end{cases}, \qquad \begin{cases} x = 3 \\ y = 37/10 \\ z = -17/5 \end{cases},$$

generate, in particolare, dalla scelta dei valori rispettivamente $\alpha = 0, 1, 2, 3$, e così via.

Esempio 5. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y - 4z = 5 \\ -6x + 4y + 8z = 5, \end{cases}$$

con $m = 2$ equazioni ed $n = 3$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 & -4 \\ -6 & 4 & 8 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 & 5 \\ -6 & 4 & 8 & 5 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 1$ e $rg(B) = 2$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) \neq rg(B)$, il sistema non possiede soluzioni.

Esempio 6. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y - 4z = 1 \\ -6x + 4y + 8z = -2, \end{cases} \tag{3.29}$$

con $m = 2$ equazioni ed $n = 3$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 & -4 \\ -6 & 4 & 8 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 & 1 \\ -6 & 4 & 8 & -2 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 1$ e $rg(B) = 1$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) = rg(B)$, il sistema possiede soluzioni e che, poiché il *rango* $r = 1$ delle due matrici è minore del numero $n = 3$ delle incognite, le soluzioni del sistema sono $\infty^{3-1} = \infty^2$. Per calcolare tali infinite soluzioni, osserviamo che la *matrice minore* comune $\mathcal{M}$ contenuta in entrambe le matrici A, B, in virtù della quale risulta $rg(A) = rg(B) = 1$, è quella avente ordine 1 evidenziata, ovvero

$$A = \begin{pmatrix} 3 & -2 & -4 \\ -6 & 4 & 8 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 & 1 \\ -6 & 4 & 8 & -2 \end{pmatrix},$$

che nel sistema corrisponde alla "parte" evidenziata

$$\begin{cases} 3x - 2y - 4z = 1 \\ -6x + 4y + 8z = -2. \end{cases}$$

In questo caso la *matrice minore* comune $\mathcal{M}$, riportata corrispondentemente sul sistema, contiene la sola incognita x nella prima equazione, da cui segue che "trasportiamo" le incognite y, z al secondo membro attribuendo loro il valore arbitrario $y = \alpha, z = \beta$ e che dunque il sistema possiede appunto ∞^2 soluzioni. Se riscriviamo la prima equazione "lasciando" l'incognita x al primo membro, in quanto contenuta nella *matrice minore* comune $\mathcal{M}$ evidenziata nel sistema, e "trasportando" le incognite $y = \alpha, z = \beta$ al secondo membro, in quanto "al di fuori" della *matrice minore* $\mathcal{M}$ evidenziata nel sistema, otteniamo il "nuovo" sistema "estratto" dal sistema iniziale

$$\left\{ \; 3x = 2\alpha + 4\beta + 1 \right.$$

la cui soluzione è $x = (2\alpha+4\beta+1)/3$, che, unita ai valori arbitrari di y, z, dà le ∞^2 soluzioni

$$\begin{cases} x = (2\alpha + 4\beta + 1)/3 \\ y = \alpha \\ z = \beta \end{cases}$$

del sistema iniziale (3.29).

Esempio 7. Dato il sistema di equazioni lineari

$$\begin{cases} 2x - 4y - 3z = 2 \\ 3x + 2y + z = -6 \\ 4x + 8y - 5z = 6 \end{cases} \tag{3.30}$$

con $m = 3$ equazioni ed $n = 3$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 2 & -4 & -3 \\ 3 & 2 & 1 \\ 4 & 8 & -5 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 2 & -4 & -3 & 2 \\ 3 & 2 & 1 & -6 \\ 4 & 8 & -5 & 6 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 3$ e $rg(B) = 3$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) = rg(B)$, il sistema possiede soluzioni e che, poiché il *rango* $r = 3$ delle due matrici coincide con il numero $n = 3$ delle incognite, il sistema possiede un'unica soluzione che ricaviamo, mediante la *regola di*

Cramer applicata a tutto il sistema iniziale (3.30), senza "trasportare" nessuna incognita al secondo membro e senza "eliminare" nessuna equazione perché la *matrice minore* comune contenuta in entrambe le matrici A, B, in virtù della quale risulta $rg(A) = rg(B) = 3$, coincide con la matrice A. La soluzione unica del sistema (3.30) è dunque

$$x = \frac{\det\begin{pmatrix} \boxed{\begin{matrix} 2 \\ -6 \\ 6 \end{matrix}} & \begin{matrix} -4 \\ 2 \\ 8 \end{matrix} & \begin{matrix} -3 \\ 1 \\ -5 \end{matrix} \end{pmatrix}}{\det\begin{pmatrix} 2 & -4 & -3 \\ 3 & 2 & 1 \\ 4 & 8 & -5 \end{pmatrix}} = -\frac{3}{2}, \qquad y = \frac{\det\begin{pmatrix} \begin{matrix} 2 \\ 3 \\ 4 \end{matrix} & \boxed{\begin{matrix} 2 \\ -6 \\ 6 \end{matrix}} & \begin{matrix} -3 \\ 1 \\ -5 \end{matrix} \end{pmatrix}}{\det\begin{pmatrix} 2 & -4 & -3 \\ 3 & 2 & 1 \\ 4 & 8 & -5 \end{pmatrix}} = \frac{1}{4},$$

$$z = \frac{\det\begin{pmatrix} \begin{matrix} 2 \\ 3 \\ 4 \end{matrix} & \begin{matrix} -4 \\ 2 \\ 8 \end{matrix} & \boxed{\begin{matrix} 2 \\ -6 \\ 6 \end{matrix}} \end{pmatrix}}{\det\begin{pmatrix} 2 & -4 & -3 \\ 3 & 2 & 1 \\ 4 & 8 & -5 \end{pmatrix}} = -2,$$

dove è stata evidenziata la colonna dei termini noti, ed è immediato verificare che la soluzione $x = -3/2, y = 1/4, z = -2$ soddisfa le tre equazioni del sistema iniziale (3.30).

Esempio 8. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y - 4z = 1 \\ -6x + 4y + 8z = -1 \\ 7x - 9y + 2z = -7 \end{cases}$$

con $m = 2$ equazioni ed $n = 3$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 & -4 \\ -6 & 4 & 8 \\ 7 & -9 & 2 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 & 1 \\ -6 & 4 & 8 & -1 \\ 7 & -9 & 2 & -7 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 2$ e $rg(B) = 3$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) \neq rg(B)$, il sistema non possiede soluzioni.

Esempio 9. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y - 4z = -1 \\ -6x + 4y + 8z = 2 \\ 7x - 9y + 2z = -5 \end{cases} \tag{3.31}$$

con $m = 3$ equazioni ed $n = 3$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 & -4 \\ -6 & 4 & 8 \\ 7 & -9 & 2 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 & -1 \\ -6 & 4 & 8 & 2 \\ 7 & -9 & 2 & -5 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 2$ e $rg(B) = 2$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) = rg(B)$, il sistema possiede soluzioni e che, poiché il *rango* $r = 2$ delle due matrici è minore del numero $n = 3$ delle incognite, le soluzioni del sistema sono infinite ed in particolare $\infty^{3-2} = \infty^1$.

Per calcolare tali infinite soluzioni, osserviamo che la *matrice minore* comune $\mathcal{M}$ contenuta in entrambe le matrici A, B, in virtù della quale risulta $rg(A) = rg(B) = 2$, è quella avente ordine 2 evidenziata, ovvero

$$A = \begin{pmatrix} 3 & -2 & -4 \\ -6 & 4 & 8 \\ 7 & -9 & 2 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -4 & -1 \\ -6 & 4 & 8 & 2 \\ 7 & -9 & 2 & -5 \end{pmatrix},$$

che nel sistema corrisponde alla "parte" evidenziata

$$\begin{cases} 3x - 2y - 4z = -1 \\ -6x + 4y + 8z = 2 \\ 7x - 9y + 2z = -5. \end{cases}$$

In questo caso la *matrice minore* comune $\mathcal{M}$, riportata corrispondentemente sul sistema, contiene le incognite x, z nella prima e terza equazione, da cui segue che "eliminiamo" la seconda equazione che è "al di fuori" della *matrice minore* comune $\mathcal{M}$.

A questo punto "lasciamo" al primo membro le due incognite x, z della prima e terza equazione, in quanto contenute nella *matrice minore* comune $\mathcal{M}$ evidenziata nel sistema, e "trasportiamo" l'incognita y al secondo membro attribuendole il valore arbitrario $y = \alpha$, in quanto "al di fuori" della *matrice minore* $\mathcal{M}$ evidenziata nel sistema.

In tal modo otteniamo il "nuovo" sistema "estratto" dal sistema iniziale

$$\begin{cases} 3x - 4z = 2\alpha - 1 \\ 7x + 2z = 9\alpha - 5, \end{cases} \tag{3.32}$$

che risolviamo mediante la *regola di Cramer*, illustrata per la soluzione dell'esercizio 3, per ricavare i valori delle incognite x, z. Dal sistema (3.28) otteniamo la soluzione

$$x = \frac{\det\begin{pmatrix} 2\alpha - 1 & -4 \\ 9\alpha - 5 & 2 \end{pmatrix}}{\det\begin{pmatrix} 3 & -4 \\ 7 & 2 \end{pmatrix}} = \frac{20\alpha - 11}{17} \qquad \text{e} \qquad z = \frac{\det\begin{pmatrix} 3 & 2\alpha - 1 \\ 7 & 9\alpha - 5 \end{pmatrix}}{\det\begin{pmatrix} 3 & -4 \\ 7 & 2 \end{pmatrix}} = \frac{13\alpha - 8}{34},$$

la quale, unita al valore arbitrario di y, dà le ∞^1 soluzioni di tutto il sistema (3.31)

$$\begin{cases} x = (20\alpha - 11)/17 \\ y = \alpha \\ z = (13\alpha - 8)/34 \end{cases},$$

tra le quali vi sono, ad esempio, le soluzioni generate dalla scelta dei vari valori di α

$$\begin{cases} x = -11/17 \\ y = 0 \\ z = -4/17 \end{cases}, \qquad \begin{cases} x = 9/17 \\ y = 1 \\ z = 5/34 \end{cases}, \qquad \begin{cases} x = 29/17 \\ y = 2 \\ z = 9/17 \end{cases}, \qquad \begin{cases} x = 49/17 \\ y = 3 \\ z = 31/34 \end{cases},$$

generate, in particolare, dalla scelta dei valori rispettivamente $\alpha = 0, 1, 2, 3$, e così via.

Esempio 10. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y = -1 \\ -6x + 4y = 2 \\ 5x + 2y = 9 \end{cases} \tag{3.33}$$

con $m = 3$ equazioni ed $n = 2$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 \\ -6 & 4 \\ 5 & 2 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -1 \\ -6 & 4 & 2 \\ 5 & 2 & 9 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 2$ e $rg(B) = 2$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) = rg(B)$, il sistema possiede soluzioni e che, poiché il *rango* $r = 2$ delle due matrici coincide con il numero $n = 2$ delle incognite, la soluzione del sistema è unica. Per calcolare tale soluzione unica, osserviamo che la *matrice minore* comune $\mathcal{M}$ contenuta in entrambe le matrici A, B, in virtù della quale risulta $rg(A) = rg(B) = 2$, è quella avente ordine 2 evidenziata, ovvero

$$A = \begin{pmatrix} 3 & -2 \\ \boxed{\begin{matrix} -6 & 4 \\ 5 & 2 \end{matrix}} \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -1 \\ \boxed{\begin{matrix} -6 & 4 \\ 5 & 2 \end{matrix}} & \begin{matrix} 2 \\ 9 \end{matrix} \end{pmatrix},$$

che nel sistema corrisponde alla "parte" evidenziata

$$\begin{cases} 3x - 2y = -1 \\ \boxed{\begin{matrix} -6x + 4y \\ 5x + 2y \end{matrix}} \begin{matrix} = 2 \\ = 9. \end{matrix} \end{cases}$$

In questo caso la *matrice minore* comune $\mathcal{M}$, riportata corrispondentemente sul sistema, contiene entrambe le incognite x, y nella seconda e terza equazione, da cui segue che nessuna incognita viene "trasportata" al secondo membro e che "eliminiamo" la prima equazione in quanto "al di fuori" della *matrice minore* comune $\mathcal{M}$.

Se dal sistema iniziale (3.33) "estraiamo" dunque la "parte" corrispondente alla *matrice minore* comune $\mathcal{M}$, "lasciando" al primo membro le incognite x, y della seconda e terza equazione ed "eliminando" la prima equazione, otteniamo il sistema

$$\begin{cases} -6x + 4y = 2 \\ 5x + 2y = 9, \end{cases}$$

che risolviamo mediante la *regola di Cramer* ricavando la soluzione unica

$$x = \frac{\det\begin{pmatrix} 2 & 4 \\ 9 & 2 \end{pmatrix}}{\det\begin{pmatrix} -6 & 4 \\ 5 & 2 \end{pmatrix}} = 1 \qquad \text{e} \qquad y = \frac{\det\begin{pmatrix} -6 & 2 \\ 5 & 9 \end{pmatrix}}{\det\begin{pmatrix} -6 & 4 \\ 5 & 2 \end{pmatrix}} = 2,$$

ovvero $x = 1, y = 2$ che soddisfa tutte e tre le equazioni del sistema iniziale (3.33).

Esempio 11. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y = -1 \\ -6x + 4y = 1 \\ 5x + 2y = 9 \end{cases}$$

con $m = 3$ equazioni ed $n = 2$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 \\ -6 & 4 \\ 5 & 2 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -1 \\ -6 & 4 & 1 \\ 5 & 2 & 9 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 2$ e $rg(B) = 3$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) \neq rg(B)$, il sistema non possiede soluzioni.

Esempio 12. Dato il sistema di equazioni lineari

$$\begin{cases} 3x - 2y = -1 \\ -6x + 4y = 2 \\ 9x - 6y = -3 \end{cases} \tag{3.34}$$

con $m = 3$ equazioni ed $n = 2$ incognite, scriviamo le due matrici A, B ad esso associate

$$A = \begin{pmatrix} 3 & -2 \\ -6 & 4 \\ 9 & -6 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -1 \\ -6 & 4 & 2 \\ 9 & -6 & -3 \end{pmatrix}$$

e osserviamo immediatamente che risulta $rg(A) = 1$ e $rg(B) = 1$, da cui segue che, in virtù del *teorema di Rouché-Capelli* applicato al caso $rg(A) = rg(B)$, il sistema possiede soluzioni e che, poiché il *rango* $r = 1$ delle due matrici è minore del numero $n = 2$ delle incognite, le soluzioni del sistema sono ∞^1. Per calcolare tali soluzioni, osserviamo che la *matrice minore* comune $\mathcal{M}$ contenuta in entrambe le matrici A, B, in virtù della quale risulta $rg(A) = rg(B) = 1$, è quella avente ordine 1 evidenziata, ovvero

$$A = \begin{pmatrix} 3 & -2 \\ -6 & 4 \\ 9 & -6 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 3 & -2 & -1 \\ -6 & 4 & 2 \\ 9 & -6 & -3 \end{pmatrix}$$

che nel sistema corrisponde alla "parte" evidenziata

$$\begin{cases} 3x - 2y = -1 \\ -6x + 4y = 2 \\ 9x - 6y = -3. \end{cases}$$

In questo caso la *matrice minore* comune $\mathcal{M}$, riportata corrispondentemente sul sistema, contiene solo l'incognita x nella prima equazione, da cui segue che "trasportiamo" l'incognita y al secondo membro attribuendole valore arbitrario $y = \alpha$ e che "eliminiamo" la seconda e terza equazione in quanto "al di fuori" della *matrice minore* comune $\mathcal{M}$.

Se dal sistema iniziale (3.34) "estraiamo" la "parte" corrispondente alla *matrice minore* comune, "lasciando" al primo membro l'incognita x della prima equazione ed "eliminando" la seconda e terza equazione, abbiamo il sistema con la sola equazione $3x = 2\alpha - 1$, la cui soluzione è $x = (2\alpha - 1)/3$, che, unita al valore arbitrario $y = \alpha$, dà le ∞^1 soluzioni di tutto il sistema iniziale (3.34)

$$\begin{cases} x = (2\alpha - 1)/3 \\ y = \alpha, \end{cases}$$

tra le quali vi sono, ad esempio, le soluzioni

$$\begin{cases} x = -1/3 \\ y = 0 \end{cases}, \quad \begin{cases} x = 1/3 \\ y = 1 \end{cases}, \quad \begin{cases} x = 1 \\ y = 2 \end{cases}, \quad \begin{cases} x = 5/3 \\ y = 3 \end{cases}, \quad \begin{cases} x = 7/3 \\ y = 4 \end{cases},$$

generate, in particolare, dai valori $\alpha = 0, 1, 2, 3, 4$.

Esempio 13. Dato il sistema di equazioni lineari

$$\begin{cases} -2x_1 + x_2 + 7x_3 - 4x_4 = 0 \\ 3x_1 + 2x_2 - 5x_3 + 6x_4 = 0 \\ -x_1 + 4x_2 + 9x_3 - 7x_4 = 0 \end{cases} \tag{3.35}$$

con $m = 3$ equazioni ed $n = 2$ incognite, scriviamo in questo caso la sola matrice A ad esso associata

$$A = \begin{pmatrix} -2 & 1 & 7 & -4 \\ 3 & 2 & -5 & 6 \\ -1 & 4 & 9 & -7 \end{pmatrix}.$$

Fino a questo punto abbiamo utilizzato il *teorema di Rouché-Capelli* solo nel "verso" per cui dall'ipotesi di uguaglianza dei due *ranghi* $rg(A) = rg(B)$ segue come tesi l'esistenza delle soluzioni del sistema.

Come è immediato riconoscere, nel caso del sistema iniziale (3.35) non occorre calcolare i *ranghi* delle due matrici A, B per stabilire se il sistema possiede soluzioni oppure no, perché il sistema (3.35) ha i termini noti tutti nulli nei secondi membri e quindi possiede almeno la soluzione, detta *soluzione banale*, $x_1 = x_2 = x_3 = x_4 = 0$.

Un sistema avente i termini noti tutti nulli nei secondi membri viene denominato *sistema omogeneo* e sottolineiamo quindi che qualsiasi *sistema omogeneo* possiede sempre almeno la *soluzione banale* costituita da tutte le incognite uguali a zero.

Se utilizziamo il *teorema di Rouché-Capelli* nel "verso" per cui se un sistema possiede soluzioni, segue che le due matrici A, B ad esso associate possiedono il medesimo *rango*, possiamo concludere dunque che per un qualsiasi *sistema omogeneo* vale sempre l'uguaglianza dei due *ranghi* $rg(A) = rg(B)$, da cui segue pertanto che per un qualsiasi *sistema omogeneo* non occorre calcolare il *rango* della matrice B perché in un *sistema omogeneo* il *rango* della matrice B è sempre coincidente con il *rango* della matrice A.

Poiché risulta $rg(A) = rg(B) = 3$, segue che, in virtù del *teorema di Rouché-Capelli*, il sistema possiede ∞^1 soluzioni, per calcolare le quali osserviamo che la *matrice minore* comune $\mathcal{M}$ contenuta in A (e quindi anche in B) che fornisce $rg(A) = rg(B) = 3$ è quella avente ordine 3 evidenziata, ovvero

$$A = \left(\begin{array}{c|ccc|} -2 & 1 & 7 & -4 \\ 3 & 2 & -5 & 6 \\ -1 & 4 & 9 & -7 \end{array}\right),$$

che nel sistema corrisponde alla "parte" evidenziata

$$\begin{cases} -2x_1 & \boxed{+ x_2 + 7x_3 - 4x_4} = 0 \\ 3x_1 & \boxed{+ 2x_2 - 5x_3 + 6x_4} = 0 \\ -x_1 & \boxed{+ 4x_2 + 9x_3 - 7x_4} = 0. \end{cases}$$

Utilizzando quindi tutte e tre le equazioni del sistema, "lasciando" le incognite x_2, x_3, x_4 al primo membro in quanto contenute nella *matrice minore* comune $\mathcal{M}$, "trasportiamo" l'incognita x_1 al secondo membro in quanto "al di fuori" della *matrice minore* comune $\mathcal{M}$ attribuendole il valore arbitrario $x_1 = \alpha$, ottenendo in tal modo dal sistema iniziale (3.35) il sistema "estratto"

$$\begin{cases} x_2 + 7x_3 - 4x_4 = 2\alpha \\ 2x_2 - 5x_3 + 6x_4 = -3\alpha \\ 4x_2 + 9x_3 - 7x_4 = \alpha. \end{cases}$$

Se risolviamo tale sistema con la *regola di Cramer*, ricaviamo la soluzione

$$x_2 = \frac{\det\begin{pmatrix} 2\alpha & 7 & -4 \\ -3\alpha & -5 & 6 \\ \alpha & 9 & -7 \end{pmatrix}}{\det\begin{pmatrix} 1 & 7 & -4 \\ 2 & -5 & 6 \\ 4 & 9 & -7 \end{pmatrix}}, \qquad x_3 = \frac{\det\begin{pmatrix} 1 & 2\alpha & -4 \\ 2 & -3\alpha & 6 \\ 4 & \alpha & -7 \end{pmatrix}}{\det\begin{pmatrix} 1 & 7 & -4 \\ 2 & -5 & 6 \\ 4 & 9 & -7 \end{pmatrix}}, \qquad x_4 = \frac{\det\begin{pmatrix} 1 & 7 & 2\alpha \\ 2 & -5 & -3\alpha \\ 4 & 9 & \alpha \end{pmatrix}}{\det\begin{pmatrix} 1 & 7 & -4 \\ 2 & -5 & 6 \\ 4 & 9 & -7 \end{pmatrix}},$$

che, unita al valore arbitrario $x_1 = \alpha$, dà le ∞^1 soluzioni di tutto il sistema iniziale (3.35)

$$\begin{cases} x_1 = \alpha \\ x_2 = -11\alpha/19 \\ x_3 = 7\alpha/19 \\ x_4 = 0. \end{cases}$$

3.3.1 Sistemi parametrici

Un sistema di equazioni lineari viene denominato *sistema parametrico* se almeno un coefficiente di un'incognita oppure un termine noto sono espressi tramite un parametro, indicato nel seguito con k, che può assumere qualsiasi valore reale. E' immediato rendersi conto quindi che la scrittura di un *sistema parametrico* comprende in realtà infiniti sistemi che corrispondono ognuno ad una particolare scelta del valore attribuito al parametro k.

Ad esempio il *sistema parametrico*

$$\begin{cases} 3x - ky = -1 \\ 2kx + 3y = k + 4 \end{cases}$$

contiene, tra gli altri infiniti, i sistemi in particolare

$$\begin{cases} 3x - y = -1 \\ 2x + 3y = 5 \end{cases}, \qquad \begin{cases} 3x - 2y = -1 \\ 4x + 3y = 6 \end{cases}, \qquad \begin{cases} 3x - 3y = -1 \\ 6x + 3y = 7 \end{cases}, \qquad \begin{cases} 3x - 4y = -1 \\ 8x + 3y = 8 \end{cases},$$

corrispondenti ai valori attribuiti al parametro k rispettivamente $k = 1, 2, 3, 4$.

Risolvere un *sistema parametrico* significa stabilire quante sono le soluzioni di ciascun sistema associato al corrispondente valore del parametro reale k e calcolare le soluzioni di ciascun sistema che appunto possiede le soluzioni. Presentiamo ora alcuni esempi.

Esempio 1. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} 3x + ky = 5 \\ 6x - 8y = -7, \end{cases} \tag{3.36a}$$

avente $m = 2$ equazioni ed $n = 2$ incognite, calcoliamo, al variare del parametro k, il *rango* della matrice associata A

$$A = \begin{pmatrix} 3 & k \\ 6 & -8 \end{pmatrix}.$$

La *matrice minore* "più grande" contenuta in tale matrice A è la matrice di ordine 2 coincidente con la matrice A stessa e osserviamo che il *rango* di A è 2 se il determinante della matrice A è diverso da zero, mentre è 1 se il determinante di A è uguale a zero.

Poiché il determinante di A è

$$\det A = \det \begin{pmatrix} 3 & k \\ 6 & -8 \end{pmatrix} = -6k - 24$$

e si annulla per $k = -4$, concludiamo che il *rango* di A è

$$rg(A) = \begin{cases} 2 & \text{se} \quad k \neq -4 \\ 1 & \text{se} \quad k = -4. \end{cases}$$

La matrice B associata al sistema è

$$B = \begin{pmatrix} 3 & k & 5 \\ 6 & -8 & -7 \end{pmatrix}$$

e il suo *rango* per $k \neq -4$ è 2 come quello della matrice A perché all'interno della matrice B la *matrice minore* coincidente con la matrice A, ovvero corrispondente alle due righe e alle prime due colonne, possiede, per $k \neq -4$, determinante ovviamente diverso da zero.

Pertanto dobbiamo calcolare il *rango* di B solo per il valore $k = -4$, ovvero dobbiamo calcolare il *rango* della matrice indicata con $B(-4)$ e ottenuta sostituendo $k = -4$ in B

$$B(-4) = \begin{pmatrix} 3 & -4 & 5 \\ 6 & -8 & -7 \end{pmatrix}.$$

Il *rango* di $B(-4)$ è uguale a 2 in virtù della *matrice minore* di ordine 2 avente determinante $68 \neq 0$ e corrispondente alle due righe e alle colonne seconda e terza.

Dopo aver scritto dunque

$$rg(B) = \begin{cases} 2 & \text{se} \quad k \neq -4 \\ 2 & \text{se} \quad k = -4 \end{cases}$$

ed aver confrontato i *ranghi* delle matrici A, B, concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede soluzione unica per $k \neq -4$, perché per $k \neq -4$ risulta $rg(A) = rg(B) = 2$;
- non possiede nessuna soluzione per $k = -4$, perché per $k = -4$ risulta $rg(A) \neq rg(B)$.

Per calcolare la soluzione unica di ogni sistema associato ad un valore $k \neq -4$ del parametro, osserviamo che la *matrice minore* di ordine 2 che dà *rango* 2 ad entrambe le matrici A, B comprende entrambe le incognite in entrambe le equazioni del sistema iniziale (3.36a) e quindi "lasciamo" al primo membro entrambe le incognite in entrambe le equazioni, senza "trasportare" nessuna incognita al secondo membro.

La soluzione unica di ogni sistema (3.36a) avente $k \neq -4$ si ottiene mediante la *regola di Cramer*, ovvero ricaviamo

$$x = \frac{\det\begin{pmatrix} 5 & k \\ -7 & -8 \end{pmatrix}}{\det\begin{pmatrix} 3 & k \\ 6 & -8 \end{pmatrix}} = \frac{40 - 7k}{6k + 24}, \qquad y = \frac{\det\begin{pmatrix} 3 & 5 \\ 6 & -7 \end{pmatrix}}{\det\begin{pmatrix} 3 & k \\ 6 & -8 \end{pmatrix}} = \frac{17}{2k + 8}. \tag{3.36b}$$

Il significato di tale soluzione unica è che, ad esempio, i sistemi (3.36a)

$$\begin{cases} 3x + y = 5 \\ 6x - 8y = -7 \end{cases}, \quad \begin{cases} 3x + 2y = 5 \\ 6x - 8y = -7 \end{cases}, \quad \begin{cases} 3x + 3y = 5 \\ 6x - 8y = -7 \end{cases}, \quad \begin{cases} 3x + 4y = 5 \\ 6x - 8y = -7 \end{cases},$$

ottenuti attribuendo i valori rispettivamente $k = 1, 2, 3, 4$ al parametro k appunto nel sistema iniziale (3.36a), hanno ognuno la soluzione unica rispettivamente

$$\begin{cases} x = 11/10 \\ y = 17/10 \end{cases}, \quad \begin{cases} x = 13/18 \\ y = 17/12 \end{cases}, \quad \begin{cases} x = 19/42 \\ y = 17/14 \end{cases}, \quad \begin{cases} x = 1/4 \\ y = 17/16 \end{cases},$$

che si ottiene mediante i corrispondenti valori rispettivamente $k = 1, 2, 3, 4$ sostituiti al parametro k nella soluzione (3.36b).

Esempio 2. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} -2x - 3y = k \\ -kx + 6y = 8, \end{cases} \tag{3.37}$$

avente $m = 2$ equazioni ed $n = 2$ incognite, calcoliamo, al variare del parametro k, il *rango* della matrice associata A

$$A = \begin{pmatrix} -2 & -3 \\ -k & 6 \end{pmatrix}.$$

La *matrice minore* "più grande" contenuta in tale matrice A è la matrice di ordine 2 coincidente con la matrice A stessa e osserviamo che il *rango* di A è 2 se il determinante della matrice A è diverso da zero, mentre è 1 se il determinante di A è uguale a zero.

Poiché il determinante di A è

$$\det A = \det\begin{pmatrix} -2 & -3 \\ -k & 6 \end{pmatrix} = -3k - 12$$

e si annulla per $k = -4$, concludiamo che il *rango* di A è

$$rg(A) = \begin{cases} 2 & \text{se} \quad k \neq -4 \\ 1 & \text{se} \quad k = -4. \end{cases}$$

La matrice B associata al sistema è

$$B = \begin{pmatrix} -2 & -3 & k \\ -k & 6 & 8 \end{pmatrix}$$

e il suo *rango* per $k \neq -4$ è 2 come quello della matrice A perché all'interno della matrice B la *matrice minore* coincidente con la matrice A, ovvero corrispondente alle due righe e alle prime due colonne, possiede, per $k \neq -4$, determinante ovviamente diverso da zero.

Pertanto dobbiamo calcolare il *rango* di B solo per il valore $k = -4$, ovvero dobbiamo calcolare il *rango* della matrice indicata con $B(-4)$ e ottenuta sostituendo $k = -4$ in B

$$B(-4) = \begin{pmatrix} -2 & -3 & -4 \\ 4 & 6 & 8 \end{pmatrix}.$$

Il *rango* di $B(-4)$ è uguale a 1 perché tutte le *matrici minori* di ordine 2 hanno determinante zero e la prima *matrice minore* avente determinante diverso da zero è, tra le altre, la *matrice mionore* appunto di ordine 1 corrispondente alla prima riga e prima colonna. Dopo aver scritto dunque

$$rg(B) = \begin{cases} 2 & \text{se} \quad k \neq -4 \\ 1 & \text{se} \quad k = -4 \end{cases}$$

ed aver confrontato i *ranghi* delle matrici A, B, concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede soluzione unica per $k \neq -4$, perché per $k \neq -4$ risulta $rg(A) = rg(B) = 2$;
- possiede ∞^1 soluzioni per $k = -4$, perché per $k = -4$ risulta $rg(A) = rg(B) = 1$.

Per calcolare la soluzione unica di ogni sistema associato ad un valore $k \neq -4$ del parametro, osserviamo che la *matrice minore* di ordine 2 che dà *rango* 2 ad entrambe le matrici A, B comprende entrambe le incognite in entrambe le equazioni del sistema iniziale (3.37) e quindi "lasciamo" al primo membro entrambe le incognite in entrambe le equazioni, senza "trasportare" nessuna incognita al secondo membro.

La soluzione unica di ogni sistema (3.37) avente $k \neq -4$ si ottiene mediante la *regola di Cramer*, ovvero ricaviamo

$$x = \frac{\det\begin{pmatrix} k & -3 \\ 8 & 6 \end{pmatrix}}{\det\begin{pmatrix} -2 & -3 \\ -k & 6 \end{pmatrix}} = -\frac{6k+24}{3k+12} = -2\,, \qquad y = \frac{\det\begin{pmatrix} -2 & k \\ -k & 8 \end{pmatrix}}{\det\begin{pmatrix} -2 & -3 \\ -k & 6 \end{pmatrix}} = \frac{16-k^2}{3k+12} = \frac{4-k}{3}\,.$$

Per calcolare le ∞^1 soluzioni del sistema associato al valore del parametro $k = -4$, sostituiamo nel sistema iniziale (3.37) il valore del parametro appunto $k = -4$, ovvero

$$\begin{cases} -2x - 3y = -4 \\ 4x + 6y = 8, \end{cases}$$

e osserviamo che la *matrice minore* di ordine 1 che dà *rango* 1 ad entrambe le matrici A, B è, tra le altre possibili, la *matrice minore* corrispondente alla prima riga e seconda colonna che comprende l'incognita y nella prima equazione. Se quindi riscriviamo il sistema (3.37) con la sola incognita y della prima equazione "lasciata" al primo membro e con l'attribuzione di un valore arbitrario $x = \alpha$ all'incognita x "trasportata" al secondo membro, abbiamo l'equazione $-3y = 2\alpha - 4$. Se uniamo la soluzione $y = (4-2\alpha)/3$ al valore arbitrario $x = \alpha$, otteniamo le ∞^1 soluzioni di tutto il sistema iniziale (3.37)

$$\begin{cases} x = \alpha \\ y = (4 - 2\alpha)/3. \end{cases}$$

Esempio 3. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} 8x + (k-9)y = -6 \\ (2k-6)x - 2y = 2-k, \end{cases} \tag{3.38}$$

avente $m = 2$ equazioni ed $n = 2$ incognite, calcoliamo, al variare del parametro k, il *rango* della matrice associata A

$$A = \begin{pmatrix} 8 & k-9 \\ 2k-6 & -2 \end{pmatrix}.$$

La *matrice minore* "più grande" contenuta in tale matrice A è la matrice di ordine 2 coincidente con la matrice A stessa e osserviamo che il *rango* di A è 2 se il determinante della matrice A è diverso da zero, mentre è 1 se il determinante di A è uguale a zero.

Poiché il determinante di A è

$$\det A = \det \begin{pmatrix} 8 & k-9 \\ 2k-6 & -2 \end{pmatrix} = -2\left(k^2 - 12k + 35\right)$$

e si annulla per $k = 5$ oppure per $k = 7$, concludiamo che il *rango* di A è

$$rg(A) = \begin{cases} 2 & \text{se} \quad k \neq 5 \text{ e } k \neq 7 \\ 1 & \text{se} \quad k = 5 \\ 1 & \text{se} \quad k = 7. \end{cases}$$

La matrice B associata al sistema è

$$B = \begin{pmatrix} 8 & k-9 & -6 \\ 2k-6 & -2 & 2-k \end{pmatrix}$$

e il suo *rango* per $k \neq 5$ e $k \neq 7$ è 2 come quello della matrice A perché nella matrice B la *matrice minore* coincidente con la matrice A, ovvero corrispondente alle due righe e alle prime due colonne, possiede, per $k \neq 5$ e $k \neq 7$, ovviamente determinante diverso da zero.

Pertanto dobbiamo calcolare il *rango* di B solo per i valori $k = 5$ e $k = 7$, ovvero dobbiamo calcolare il *rango* delle matrici

$$B(5) = \begin{pmatrix} 8 & -4 & -6 \\ 4 & -2 & -3 \end{pmatrix} \qquad \text{e} \qquad B(7) = \begin{pmatrix} 8 & -2 & -6 \\ 8 & -2 & -5 \end{pmatrix}.$$

Il *rango* della matrice $B(5)$ è uguale a 1 perché tutte le sue *matrici minori* di ordine 2 hanno determinante zero e la prima *matrice minore* avente determinante diverso da zero è, tra le altre, la *matrice minore* appunto di ordine 1 corrispondente alla prima riga e prima colonna. Invece il *rango* della matrice $B(7)$ è uguale a 2 perché la sua *matrice minore* di ordine 2 corrispondente alle due righe e alle colonne seconda e terza ha determinante diverso da zero. Dopo aver scritto dunque il *rango* della matrice B

$$rg(B) = \begin{cases} 2 & \text{se} \quad k \neq 5 \text{ e } k \neq 7 \\ 1 & \text{se} \quad k = 5 \\ 2 & \text{se} \quad k = 7 \end{cases}$$

ed aver confrontato i *ranghi* delle matrici A, B, concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede soluzione unica per $k \neq 5$ e $k \neq 7$, perché per tali k risulta $rg(A) = rg(B) = 2$;
- possiede ∞^1 soluzioni per $k = 5$, perché per $k = 5$ risulta $rg(A) = rg(B) = 1$;
- non possiede nessuna soluzione per $k = 7$, perché per $k = 7$ risulta $rg(A) \neq rg(B)$.

Per calcolare la soluzione unica di ogni sistema associato ad un valore $k \neq 5$ e $k \neq 7$ del parametro, osserviamo che la *matrice minore* di ordine 2 che dà *rango* 2 ad entrambe le matrici A, B comprende entrambe le incognite in entrambe le equazioni del sistema iniziale (3.38) e quindi "lasciamo" al primo membro entrambe le incognite in entrambe le equazioni, senza "trasportare" nessuna incognita al secondo membro.

La soluzione unica di ogni sistema (3.38) avente $k \neq 5$ e $k \neq 7$ si ottiene mediante la *regola di Cramer* e ricaviamo $x = (k-6)/(14-2k), y = 2/(7-k)$ perché risulta

$$x = \frac{\det\begin{pmatrix} -6 & k-9 \\ 2-k & -2 \end{pmatrix}}{\det\begin{pmatrix} 8 & k-9 \\ 2k-6 & -2 \end{pmatrix}} = \frac{k-6}{14-2k}, \qquad y = \frac{\det\begin{pmatrix} 8 & -6 \\ 2k-6 & 2-k \end{pmatrix}}{\det\begin{pmatrix} 8 & k-9 \\ 2k-6 & -2 \end{pmatrix}} = \frac{2}{7-k}.$$

Per calcolare le ∞^1 soluzioni del sistema associato al valore del parametro $k = 5$, sostituiamo nel sistema iniziale (3.38) il valore del parametro appunto $k = 5$, ovvero

$$\begin{cases} 8x - 4y = -6 \\ 4x - 2y = -3, \end{cases}$$

e osserviamo che la *matrice minore* di ordine 1 che dà *rango* 1 ad entrambe le matrici A, B è, tra le altre possibili, la *matrice minore* corrispondente alla prima riga e seconda colonna che comprende l'incognita y nella prima equazione. Se quindi riscriviamo il sistema (3.38) con la sola incognita y della prima equazione "lasciata" al primo membro e con l'attribuzione di un valore arbitrario $x = \alpha$ all'incognita x "trasportata" al secondo membro, abbiamo l'equazione $-4y = -8\alpha - 6$. Se uniamo la soluzione $y = (4\alpha + 3)/2$ al valore arbitrario $x = \alpha$, otteniamo le ∞^1 soluzioni di tutto il sistema iniziale (3.38)

$$\begin{cases} x = \alpha \\ y = (4\alpha + 3)/2. \end{cases}$$

Esempio 4. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} (2-k)x - 6y = 0 \\ -6x - (k+3)y = 0, \end{cases} \tag{3.39}$$

avente $m = 2$ equazioni ed $n = 2$ incognite, calcoliamo, al variare del parametro k, il *rango* della matrice associata A

$$A = \begin{pmatrix} 2-k & -6 \\ -6 & -k-3 \end{pmatrix}.$$

La *matrice minore* "più grande" contenuta in tale matrice A è la matrice di ordine 2 coincidente con la matrice A stessa e osserviamo che il *rango* di A è 2 se il determinante della matrice A è diverso da zero, mentre è 1 se il determinante di A è uguale a zero.

Poiché il determinante di A è

$$\det A = \det\begin{pmatrix} 2-k & -6 \\ -6 & -k-3 \end{pmatrix} = k^2 + k - 42$$

e si annulla per $k = 6$ oppure per $k = -7$, concludiamo che il *rango* di A è

$$rg(A) = \begin{cases} 2 & \text{se} \quad k \neq 6 \text{ e } k \neq -7 \\ 1 & \text{se} \quad k = 6 \\ 1 & \text{se} \quad k = -7. \end{cases}$$

Poiché il sistema (3.39) è *omogeneo*, segue che i *ranghi* delle due matrici A, B sono sempre uguali tra loro per ogni valore del parametro k, ovvero abbiamo sempre l'uguaglianza $rg(A) = rg(B)$ per ogni valore del parametro k. Dal confronto tra i *ranghi* delle matrici A, B, concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede soluzione unica per $k \neq 6$ e $k \neq -7$, perché per tali k si ha $rg(A) = rg(B) = 2$;
- possiede ∞^1 soluzioni per $k = 6$, perché per $k = 6$ risulta $rg(A) = rg(B) = 1$;
- possiede ∞^1 soluzioni per $k = -7$, perché per $k = -7$ risulta $rg(A) = rg(B) = 1$.

Poiché il sistema (3.39) è *omogeneo*, segue che la coppia $x = 0, y = 0$ è una soluzione e allora possiamo concludere che la soluzione unica di tutti i sistemi corrispondenti a tutti i valori del parametro $k \neq 6$ e $k \neq -7$ è sempre $x = 0, y = 0$, ovvero gli infiniti sistemi associati ai valori del parametro $k \neq 6$ e $k \neq -7$ hanno tutti soluzione unica e in aggiunta tutti la medesima soluzione unica $x = 0, y = 0$.

Per calcolare le ∞^1 soluzioni del sistema associato al valore del parametro $k = 6$, sostituiamo nel sistema iniziale (3.39) il valore del parametro appunto $k = 6$, ovvero

$$\begin{cases} -4x - 6y = 0 \\ -6x - 9y = 0, \end{cases}$$

e osserviamo che la *matrice minore* di ordine 1 che dà *rango* 1 ad entrambe le matrici A, B è, tra le altre possibili, la *matrice minore* corrispondente alla prima riga e seconda colonna che comprende l'incognita y nella prima equazione. Se quindi riscriviamo il sistema (3.39) con la sola incognita y della prima equazione "lasciata" al primo membro e con l'attribuzione di un valore arbitrario $x = \alpha$ all'incognita x "trasportata" al secondo membro, abbiamo l'equazione $-6y = 4\alpha$. Se uniamo la soluzione $y = -2\alpha/3$ al valore arbitrario $x = \alpha$, otteniamo le ∞^1 soluzioni di tutto il sistema iniziale (3.39)

$$\begin{cases} x = \alpha \\ y = -2\alpha/3. \end{cases}$$

Per calcolare le ∞^1 soluzioni del sistema associato al valore del parametro $k = -7$, sostituiamo nel sistema iniziale (3.39) il valore del parametro appunto $k = -7$, ovvero

$$\begin{cases} 9x - 6y = 0 \\ -6x + 4y = 0, \end{cases}$$

e osserviamo che la *matrice minore* di ordine 1 che dà *rango* 1 ad entrambe le matrici A, B è, tra le altre possibili, la *matrice minore* corrispondente alla prima riga e seconda colonna che comprende l'incognita y nella prima equazione. Se quindi riscriviamo il sistema (3.39) con la sola incognita y della prima equazione "lasciata" al primo membro e con l'attribuzione di un valore arbitrario $x = \alpha$ all'incognita x "trasportata" al secondo membro, abbiamo l'equazione $-6y = -9\alpha$. Se uniamo la soluzione $y = 3\alpha/2$ al valore arbitrario $x = \alpha$, otteniamo le ∞^1 soluzioni di tutto il sistema iniziale (3.39)

$$\begin{cases} x = \alpha \\ y = 3\alpha/2. \end{cases}$$

Esempio 5. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} -9x - ky + 6z = 2 \\ kx + y + 2z = -4, \end{cases} \tag{3.40}$$

avente $m = 2$ equazioni ed $n = 3$ incognite, calcoliamo, al variare del parametro k, il *rango* della matrice associata A

$$A = \begin{pmatrix} -9 & -k & 6 \\ k & 1 & 2 \end{pmatrix}.$$

Le *matrici minori* "più grandi" contenute in tale matrice A hanno ordine 2 e poiché i loro determinanti hanno espressione

$$\det\begin{pmatrix} -9 & -k \\ k & 1 \end{pmatrix} = k^2 - 9, \qquad \det\begin{pmatrix} -9 & 6 \\ k & 2 \end{pmatrix} = -6k - 18, \qquad \det\begin{pmatrix} -k & 6 \\ 1 & 2 \end{pmatrix} = -2k - 6,$$

concludiamo che per $k = -3$ tutte le *matrici minori* di ordine 2 hanno determinante zero, da cui segue dunque il *rango* di A dato da

$$rg(A) = \begin{cases} 2 & \text{se} \quad k \neq -3 \\ 1 & \text{se} \quad k = -3. \end{cases}$$

Puntualizziamo che per $k = 3$ il *rango* di A è 2 perché la matrice A con $k = 3$ è

$$A(3) = \begin{pmatrix} -9 & -3 & 6 \\ 3 & 1 & 2 \end{pmatrix}$$

e possiamo osservare che sebbene si annulli il determinante della *matrice minore* relativa alle due righe e alle prime due colonne, tuttavia abbiamo che nella matrice $A(3)$ la prima *matrice minore* con determinante diverso da zero è di ordine 2 ed in particolare possiamo considerare quella relativa alle due righe e alle colonne seconda e terza.

La matrice B associata al sistema è

$$B = \begin{pmatrix} -9 & -k & 6 & 2 \\ k & 1 & 2 & -4 \end{pmatrix}$$

e il suo *rango* per $k \neq -3$ è 2 come quello della matrice A perché nella matrice B vi è, per $k \neq -3$, almeno una *matrice minore* di ordine 2 con determinante diverso da zero contenuta in A. Pertanto dobbiamo calcolare il *rango* di B solo per il valore $k = -3$, ovvero dobbiamo calcolare il *rango* della matrice

$$B(-3) = \begin{pmatrix} -9 & 3 & 6 & 2 \\ -3 & 1 & 2 & -4 \end{pmatrix}.$$

Il *rango* della matrice $B(-3)$ è uguale a 2 perché la prima *matrice minore* avente determinante diverso da zero è, tra le altre, la *matrice minore*, appunto di ordine 2, corrispondente alle due righe e alle due colonne terza e quarta.

Dopo aver scritto dunque il *rango* della matrice B

$$rg(B) = \begin{cases} 2 & \text{se} \quad k \neq -3 \\ 2 & \text{se} \quad k = -3 \end{cases}$$

ed aver confrontato i *ranghi* delle matrici A, B, concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede ∞^1 soluzioni per $k \neq -3$, perché per $k \neq -3$ risulta $rg(A) = rg(B) = 2$;
- non possiede nessuna soluzione per $k = -3$, perché per $k = -3$ risulta $rg(A) \neq rg(B)$.

Per calcolare le ∞^1 soluzioni di tutti i sistemi associati ai valori del parametro $k \neq -3$, consideriamo come *matrice minore* comune alle due matrici A, B che dà *rango* 2 ad entrambe le matrici, la *matrice minore* di ordine 2 relativa alle due righe e alle colonne seconda e terza, la quale comprende le incognite y, z in entrambe le equazioni del sistema iniziale (3.40). Se quindi per $k \neq -3$ "lasciamo" al primo membro le incognite y, z nelle due equazioni, in quanto contenute nella *matrice minore* comune alle due matrici A, B, e "trasportiamo" l'incognita x, in quanto "al di fuori" della *matrice minore* comune alle due matrici A, B, al secondo membro attribuendole valore arbitrario $x = \alpha$, otteniamo il sistema "estratto" dal sistema iniziale (3.40) nella forma

$$\begin{cases} -ky + 6z = 9\alpha + 2 \\ y + 2z = -k\alpha - 4, \end{cases}$$

che risolviamo mediante la *regola di Cramer* per ricavare

$$y = \frac{\det\begin{pmatrix} 9\alpha+2 & 6 \\ -k\alpha-4 & 2 \end{pmatrix}}{\det\begin{pmatrix} -k & 6 \\ 1 & 2 \end{pmatrix}} = -3\alpha - \frac{14}{k+3}, \qquad z = \frac{\det\begin{pmatrix} -k & 9\alpha+2 \\ 1 & -k\alpha-4 \end{pmatrix}}{\det\begin{pmatrix} -k & 6 \\ 1 & 2 \end{pmatrix}} = \frac{3-k}{2}\,\alpha - \frac{2k-1}{k+3}.$$

Se uniamo i risultati di y, z al valore arbitrario $x = \alpha$, possiamo scrivere le ∞^1 soluzioni di tutto il sistema iniziale (3.40)

$$x = \alpha, \qquad y = -3\alpha - \frac{14}{k+3}, \qquad z = \frac{3-k}{2}\,\alpha - \frac{2k-1}{k+3}.$$

Esempio 6. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} -9x - ky + 6z = -3 \\ kx + y + 2z = -1, \end{cases} \tag{3.41}$$

avente $m = 2$ equazioni ed $n = 3$ incognite, riconosciamo che la matrice A associata al sistema (3.41) coincide con la matrice A dell'esercizio precedente e quindi riscriviamo il suo *rango* come nell'esercizio precedente, ovvero il *rango* di A è

$$rg(A) = \begin{cases} 2 & \text{se} \quad k \neq -3 \\ 1 & \text{se} \quad k = -3. \end{cases}$$

La matrice B associata al sistema (3.41) è

$$B = \begin{pmatrix} -9 & -k & 6 & -3 \\ k & 1 & 2 & -1 \end{pmatrix}$$

e il suo *rango* per $k \neq -3$ coincide con il *rango* 2 della matrice A. E' immediato rendersi conto che il *rango* della matrice B per il valore $k = -3$, ovvero il *rango* della matrice

$$B(-3) = \begin{pmatrix} -9 & 3 & 6 & -3 \\ -3 & 1 & 2 & -1 \end{pmatrix},$$

è uguale ad 1 e dunque il *rango* della matrice B è

$$rg(B) = \begin{cases} 2 & \text{se} \quad k \neq -3 \\ 1 & \text{se} \quad k = -3. \end{cases}$$

Dopo aver confrontato i *ranghi* delle matrici A, B, concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede ∞^1 soluzioni per $k \neq -3$, perché per $k \neq -3$ risulta $rg(A) = rg(B) = 2$;
- possiede ∞^2 soluzioni per $k \neq -3$, perché per $k \neq -3$ risulta $rg(A) = rg(B) = 1$.

Per calcolare le ∞^1 soluzioni di tutti i sistemi associati ai valori del parametro $k \neq -3$, procediamo esattamente come nell'esercizio precedente, "lasciando" al primo membro le incognite y, z nelle due equazioni e "trasportando" l'incognita x al secondo membro attribuendole valore arbitrario $x = \alpha$.

In questo modo il sistema "estratto" dal sistema iniziale (3.41) è

$$\begin{cases} -ky + 6z = 9\alpha - 3 \\ y + 2z = -k\alpha - 1 \end{cases}$$

e lascio agli studenti il semplice esercizio della determinazione, mediante la *regola di Cramer*, della sua soluzione y, z, da unire infine al valore arbitrario $x = \alpha$ per ricavare le ∞^1 soluzioni x, y, z di tutto il sistema iniziale (3.41).

Per calcolare le ∞^2 soluzioni del sistema associato al valore del parametro $k = -3$, sostituiamo nel sistema iniziale (3.41) il valore del parametro appunto $k = -3$, ovvero

$$\begin{cases} -9x + 3y + 6z = -3 \\ -3x + y + 2z = -1, \end{cases}$$

e osserviamo che la *matrice minore* di ordine 1 che dà *rango* comune 1 ad entrambe le matrici A, B scritte con $k = -3$, è, tra le altre possibili, la *matrice minore* corrispondente alla seconda riga e seconda colonna che comprende l'incognita y nella seconda equazione.

Se quindi riscriviamo il sistema (3.41) con la sola incognita y della seconda equazione "lasciata" al primo membro, in quanto contenuta nella *matrice minore* comune alle due matrici A, B, e con l'attribuzione di un valore arbitrario $x = \alpha$ e $z = \beta$ alle due incognite x, z "trasportate" al secondo membro in quanto "al di fuori" della *matrice minore* comune alle due matrici A, B, abbiamo l'equazione $y = 3\alpha - 2\beta - 1$, la cui soluzione, unita ai valori arbitrari $x = \alpha, z = \beta$, fornisce le ∞^2 soluzioni di tutto il sistema iniziale (3.41)

$$\begin{cases} x = \alpha \\ y = 3\alpha - 2\beta - 1 \\ z = \beta. \end{cases}$$

Esempio 7. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} 4x + ky + z = 2 \\ kx + y + 5z = -1, \end{cases} \tag{3.42}$$

avente $m = 2$ equazioni ed $n = 3$ incognite, calcoliamo, al variare del parametro k, il *rango* della matrice associata A

$$A = \begin{pmatrix} 4 & k & 1 \\ k & 1 & 5 \end{pmatrix}.$$

Le *matrici minori* "più grandi" contenute in tale matrice A hanno ordine 2 e i loro determinanti hanno espressione

$$\det\begin{pmatrix}4 & k\\ k & 1\end{pmatrix} = 4 - k^2, \qquad \det\begin{pmatrix}4 & 1\\ k & 5\end{pmatrix} = 20 - k, \qquad \det\begin{pmatrix}k & 1\\ 1 & 5\end{pmatrix} = 5k - 1. \tag{3.43}$$

Ricordando che la matrice A possiede *rango* 2 se essa contiene almeno una *matrice minore* di ordine 2 con determinante diverso da zero e che invece il *rango* è 1 se tutte le tre *matrici minori* di ordine 2 contenute in A hanno determinante zero, riconosciamo immediatamente che, per ogni valore attribuito al parametro k, c'è sempre nella matrice A almeno una *matrice minore* di ordine 2 con determinante diverso da zero.

Poiché infatti i determinanti delle *matrici minori* (3.43) si annullano il primo per i valori $k = \pm 2$, il secondo per $k = 20$ e il terzo per $k = 1/5$, segue che per qualsiasi valore attribuito al parametro k tutti i determinanti (3.43) sono tutti e tre diversi da zero oppure, se uno dei tre è zero, gli altri due sono diversi da zero perché non cé nessun valore di k per il quale si annullano tutti e tre simultaneamente. Poiché dunque la matrice A possiede sempre almeno una *matrice minore* di ordine 2 con determinante diverso da zero, segue che anche la matrice B possiede sempre almeno una *matrice minore* di ordine 2 con determinante diverso da zero.

Pertanto concludiamo che le matrici A, B hanno sempre il medesimo *rango* 2 e che quindi i sistemi (3.42) possiedono sempre ∞^1 soluzioni per ogni valore del parametro k.

Per calcolare le ∞^1 soluzioni dei sistemi (3.42), dobbiamo prestare attenzione alla *matrice minore* di ordine 2 comune alle due matrici A, B con determinante diverso da zero che dà *rango* 2 alle due matrici A, B. Per $k = -2$ oppure $k = 2$, non possiamo "estrarre" dal sistema (3.43) la *matrice minore* di ordine 2 corrispondente alle incognite x, y, ma solo le *matrici minori* di ordine 2 corrispondenti alle incognite x, z oppure alle incognite y, z; per $k = 20$ non possiamo "estrarre" dal sistema (3.43) la *matrice minore* di ordine 2 corrispondente alle incognite x, z, ma solo le *matrici minori* di ordine 2 corrispondenti alle incognite x, y oppure alle incognite y, z; infine per $k = 1/5$ non possiamo "estrarre" dal sistema (3.43) la *matrice minore* di ordine 2 corrispondente alle incognite y, z, ma solo le *matrici minori* di ordine 2 corrispondenti alle incognite x, y oppure alle incognite x, z.

Esempio 8. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} 2x - y = k \\ 3x + ky = -6 \\ 7x + 3y = 4 \end{cases} \tag{3.44a}$$

avente $m = 3$ equazioni ed $n = 2$ incognite, abbiamo che il *rango* della matrice associata

$$A = \begin{pmatrix}2 & -1\\ 3 & k\\ 7 & 3\end{pmatrix}$$

è sempre 2 per ogni valore del parametro k perché la prima *matrice minore* con determinante diverso da zero che troviamo in A è quella di ordine 2 relativa alle righe prima e terza e alle due colonne.

La matrice associata al sistema

$$B = \begin{pmatrix}2 & -1 & k\\ 3 & k & -6\\ 7 & 3 & 4\end{pmatrix}$$

possiede *rango* 3 se il determinante di B stessa risulta diverso da zero, mentre è 2 se il determinante di B risulta uguale a zero. Poiché abbiamo $\det B = -7k^2 + 17k + 90$ e $\det B = 0$ per $k = 5$ oppure per $k = -18/7$, il *rango* di B è

$$rg(B) = \begin{cases} 3 & \text{se} \quad k \neq 5 \text{ e } k \neq -18/7 \\ 2 & \text{se} \quad k = 5 \\ 2 & \text{se} \quad k = -18/7 \end{cases}$$

e dal confronto tra i *ranghi* delle due matrici A, B, concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede soluzione unica per $k = 5$ oppure per $k = -18/7$, perché $rg(A) = rg(B) = 2$;
- non possiede soluzioni per $k \neq -3$ e $k \neq -18/7$ perché risulta $rg(A) \neq rg(B)$.

Per calcolare la soluzione unica del sistema con $k = 5$, risolviamo il sistema

$$\begin{cases} 2x - y = 5 \\ 7x + 3y = 4, \end{cases} \tag{3.44b}$$

"estratto" dal sistema iniziale (3.44a) relativamente alla *matrice minore* di ordine 2 corrispondente alle equazioni prima e terza, in cui abbiamo posto appunto $k = 5$.

La soluzione unica del sistema iniziale (3.44a), che si ottiene risolvendo mediante la *regola di Cramer* il sistema "estratto" (3.44b), è

$$x = \frac{\det\begin{pmatrix} 5 & -1 \\ 4 & 3 \end{pmatrix}}{\det\begin{pmatrix} 2 & -1 \\ 7 & 3 \end{pmatrix}} = \frac{19}{13}, \qquad y = \frac{\det\begin{pmatrix} 2 & 5 \\ 7 & 4 \end{pmatrix}}{\det\begin{pmatrix} 2 & -1 \\ 7 & 3 \end{pmatrix}} = -\frac{27}{13}$$

e osserviamo che essa soddisfa anche la seconda equazione del sistema iniziale (3.44a).

Per calcolare la soluzione unica del sistema con $k = -18/7$, risolviamo il sistema

$$\begin{cases} 2x - y = -18/7 \\ 7x + 3y = 4, \end{cases} \tag{3.44c}$$

"estratto" dal sistema iniziale (3.44a) relativamente alla *matrice minore* di ordine 2 corrispondente alle equazioni prima e terza, in cui abbiamo posto appunto $k = -18/7$.

La soluzione unica del sistema iniziale (3.44a), che si ottiene risolvendo mediante la *regola di Cramer* il sistema "estratto" (3.44c), è

$$x = \frac{\det\begin{pmatrix} -18/7 & -1 \\ 4 & 3 \end{pmatrix}}{\det\begin{pmatrix} 2 & -1 \\ 7 & 3 \end{pmatrix}} = -\frac{2}{7}, \qquad y = \frac{\det\begin{pmatrix} 2 & -18/7 \\ 7 & 4 \end{pmatrix}}{\det\begin{pmatrix} 2 & -1 \\ 7 & 3 \end{pmatrix}} = 2$$

e osserviamo che essa soddisfa anche la seconda equazione del sistema iniziale (3.44a).

Esempio 9. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} x + ky = 3 \\ kx - 2ky = -6 \\ (1-k)x - 6y = 9 \end{cases} \tag{3.45a}$$

con $m = 3$ equazioni ed $n = 2$ incognite, abbiamo che i determinanti delle tre *matrici minori* di ordine 2 contenute nella matrice associata

$$A = \begin{pmatrix} 1 & k \\ k & -2k \\ 1-k & -6 \end{pmatrix}$$

sono

$$\det\begin{pmatrix} 1 & k \\ k & -2k \end{pmatrix} = -k(k+2), \qquad \det\begin{pmatrix} 1 & k \\ 1-k & -6 \end{pmatrix} = k^2-k-6, \qquad \det\begin{pmatrix} k & -2k \\ 1-k & -6 \end{pmatrix} = -2k(k+2)$$

e, come si vede, si annullano tutti per $k = -2$, da cui segue che il *rango* della matrice A è

$$rg(A) = \begin{cases} 2 & \text{se} \quad k \neq -2 \\ 1 & \text{se} \quad k = -2. \end{cases}$$

La matrice B associata al sistema

$$B = \begin{pmatrix} 1 & k & 3 \\ k & -2k & -6 \\ 1-k & -6 & 9 \end{pmatrix}$$

ha determinante $\det B = -9(k+2)^2$ e poiché, come si riconosce immediatamente, per il valore del parametro $k = -2$ si annulla non solo il determinante dell'unica *matrice minore* di ordine 3 coincidente con B stessa, ma si annullano anche i determinanti di tutte le *matrici minori* di ordine 2, segue che il *rango* della matrice B è

$$rg(B) = \begin{cases} 3 & \text{se} \quad k \neq -2 \\ 1 & \text{se} \quad k = -2. \end{cases}$$

Dal confronto tra i *ranghi* delle due matrici A, B, concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede ∞^1 soluzioni per $k = -2$ perché per $k = -2$ risulta $rg(A) = rg(B) = 1$;
- non possiede soluzioni per $k \neq -2$ perché per $k \neq -2$ risulta $rg(A) \neq rg(B)$.

Per calcolare le ∞^1 soluzioni del sistema (3.45a) con $k = -2$, sostituiamo il valore $k = -2$ nel sistema (3.45a) ottenendo così il sistema numerico

$$\begin{cases} x - 2y = 3 \\ -2x + 4y = -6 \\ 3x - 6y = 9 \end{cases} \tag{3.45b}$$

e riconosciamo che le matrici A, B associate al sistema (3.45b) possiedono *rango* 1 in virtù della *matrice minore* comune di ordine 1 con determinante diverso da zero, relativa all'incognita x nella prima equazione. Se nel sistema (3.45b) "estraiamo" l'incognita x nella prima equazione "trasportando" l'incognita y al secondo membro uguagliata al valore arbitrario $y = \alpha$, seguono il sistema $x = 2\alpha + 3$ e le ∞^1 soluzioni di tutto il sistema (3.45a)

$$\begin{cases} x = 2\alpha + 3 \\ y = \alpha. \end{cases}$$

Esempio 10. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} -5x - ky + 4z = -2 \\ 7x - 9y + kz = -k \\ -3x + 5y + z = 3 \end{cases} \tag{3.46a}$$

avente $m = 3$ equazioni ed $n = 3$ incognite, calcoliamo, al variare del parametro k, il *rango* della matrice associata A

$$A = \begin{pmatrix} -5 & -k & 4 \\ 7 & -9 & k \\ -3 & 5 & 1 \end{pmatrix}.$$

Poiché il determinante della matrice A è $\det A = 3k^2 + 32k + 77$ e si annulla per i valori del parametro $k = -11/3$ e $k = -7$, segue che il *rango* della matrice A è

$$rg(A) = \begin{cases} 3 & \text{se} \quad k \neq -11/3 \text{ e } k \neq -7 \\ 2 & \text{se} \quad k = -11/3 \\ 2 & \text{se} \quad k = -7 \end{cases}$$

e che anche il *rango* della matrice B per tutti i valori del parametro $k \neq -11/3$ e $k \neq -7$ è uguale a 3. Dobbiamo calcolare ancora quindi solo il *rango* della matrice B per il valore del parametro $k = -11/3$, ovvero della matrice indicata con $B(-11/3)$, e il *rango* della matrice B per il valore del parametro $k = -7$, ovvero della matrice indicata con $B(-7)$.

La matrice $B(-11/3)$ è

$$B(-11/3) = \begin{pmatrix} -5 & 11/3 & 4 & -2 \\ 7 & -9 & -11/3 & 11/3 \\ -3 & 5 & 1 & 3 \end{pmatrix}$$

ed il suo *rango* è 3 perché la prima *matrice minore* che troviamo in essa con determinante diverso da zero è quella di ordine 3 data dalle tre righe e dalle tre colonne seconda, terza e quarta. La matrice $B(-7)$ è

$$B(-11/3) = \begin{pmatrix} -5 & 7 & 4 & -2 \\ 7 & -9 & -7 & 7 \\ -3 & 5 & 1 & 3 \end{pmatrix}$$

ed il suo *rango* è 2 perché tutte le *matrici minori* di ordine 3 hanno determinante zero e la prima *matrice minore* che troviamo in essa con determinante diverso da zero è quella di ordine 2 data dalle prime due righe e prime due colonne.

Dope aver scritto il *rango* di B

$$rg(B) = \begin{cases} 3 & \text{se} \quad k \neq -11/3 \text{ e } k \neq -7 \\ 3 & \text{se} \quad k = -11/3 \\ 2 & \text{se} \quad k = -7 \end{cases}$$

ed aver confrontato i *ranghi* delle due matrici A, B, concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede soluzione unica per $k \neq -11/3$ e $k \neq -7$, perché risulta $rg(A) = rg(B) = 3$;
- possiede ∞^1 soluzioni per $k = -7$ perché per $k = -7$ risulta $rg(A) = rg(B) = 2$;
- non possiede soluzioni per $k \neq -11/3$ perché per $k \neq -11/3$ risulta $rg(A) \neq rg(B)$.

La soluzione unica del sistema (3.46a), corrispondente al caso di matrici A, B con *rango* 3 per i valori del parametro $k \neq -11/3$ e $k \neq -7$, si calcola utilizzando come *matrice minore* di ordine 3 quella coincidente con A stessa nel sistema iniziale (3.46a) che contiene le tre incognite nelle tre equazioni e quindi, mediante la *regola di Cramer*, si ottiene

$$x = \frac{\det\begin{pmatrix} -2 & -k & 4 \\ -k & -9 & k \\ 3 & 5 & 1 \end{pmatrix}}{\det\begin{pmatrix} -5 & -k & 4 \\ 7 & -9 & k \\ -3 & 5 & 1 \end{pmatrix}}, \qquad y = \frac{\det\begin{pmatrix} -5 & -2 & 4 \\ 7 & -k & k \\ -3 & 3 & 1 \end{pmatrix}}{\det\begin{pmatrix} -5 & -k & 4 \\ 7 & -9 & k \\ -3 & 5 & 1 \end{pmatrix}}, \qquad z = \frac{\det\begin{pmatrix} -5 & -k & -2 \\ 7 & -9 & -k \\ -3 & 5 & 3 \end{pmatrix}}{\det\begin{pmatrix} -5 & -k & 4 \\ 7 & -9 & k \\ -3 & 5 & 1 \end{pmatrix}},$$

dove il calcolo dei determinanti viene lasciato come semplice esercizio agli studenti.

Per calcolare le ∞^1 soluzioni del sistema (3.46a) con $k = -7$, sostituiamo appunto il valore del parametro $k = -7$ nel sistema (3.46a) ottenendo così il sistema numerico

$$\begin{cases} -5x + 7y + 4z = -2 \\ 7x - 9y - 7z = 7 \\ -3x + 5y + z = 3 \end{cases} \tag{3.46b}$$

ed utilizzando come *matrice minore* di ordine 2 che dà *rango* 2 alle matrici A, B quella relativa alle prime due righe e prime due colonne.

Se dal sistema (3.46b) “estraiamo” dunque il sistema

$$\begin{cases} -5x + 7y = -4\alpha - 2 \\ 7x - 9y = 7\alpha + 7 \end{cases}$$

in cui abbiamo “trasportato” l'incognita z al secondo membro attribuendole valore arbitrario $z = \alpha$, in quanto “al di fuori” della *matrice minore* di ordine 2 che dà *rango* 2 alle matrici A, B, otteniamo la soluzione

$$x = \frac{\det\begin{pmatrix} -4\alpha - 2 & 7 \\ 7\alpha + 7 & -9 \end{pmatrix}}{\det\begin{pmatrix} -5 & 7 \\ 7 & -9 \end{pmatrix}} = \frac{13\alpha + 31}{4}, \qquad y = \frac{\det\begin{pmatrix} -5 & -4\alpha - 2 \\ 7 & 7\alpha + 7 \end{pmatrix}}{\det\begin{pmatrix} -5 & 7 \\ 7 & -9 \end{pmatrix}} = \frac{7\alpha + 21}{4},$$

che, unita al valore arbitrario $z = \alpha$, dà le ∞^1 soluzioni di tutto il sistema (3.46b)

$$\begin{cases} x = (13\alpha + 31)/4 \\ y = (7\alpha + 21)/4 \\ z = \alpha \end{cases}$$

che verificano, in particolare, anche la terza equazione del sistema (3.46b).

Esempio 11. Dato il sistema parametrico di equazioni lineari

$$\begin{cases} (9 - k)x - 4z = 0 \\ (-k - 5)y = 0 \\ -4x - (k + 6)z = 0 \end{cases} \tag{3.47a}$$

avente $m = 3$ equazioni ed $n = 3$ incognite, calcoliamo, al variare del parametro k, il *rango* della matrice associata A

$$A = \begin{pmatrix} 9-k & 0 & -4 \\ 0 & -k-5 & 0 \\ -4 & 0 & -k-6 \end{pmatrix}.$$

Se sviluppiamo il determinante di A secondo la seconda riga, otteniamo

$$\det A = (-k-5)(k^2 - 3k - 70)$$

che si annulla per i tre valori del parametro $k = -5, -7, 10$, da cui segue il *rango* di A

$$rg(A) = \begin{cases} 3 & \text{se} \quad k \neq -5 \text{ e } k \neq -7 \text{ e } k \neq 10 \\ 2 & \text{se} \quad k = -5 \\ 2 & \text{se} \quad k = -7 \\ 2 & \text{se} \quad k = 10. \end{cases}$$

Poiché il sistema è omogeneo, segue che i *ranghi* delle due matrici A, B sono sempre uguali e quindi concludiamo che, in virtù del *teorema di Rouché-Capelli*, il sistema

- possiede soluzione unica per $k \neq -5$ e $k \neq -7$ e $k \neq 10$
- possiede ∞^1 soluzioni per $k = -5$ perché per $k = -5$ risulta $rg(A) = rg(B) = 2$;
- possiede ∞^1 soluzioni per $k = -7$ perché per $k = -7$ risulta $rg(A) = rg(B) = 2$;
- possiede ∞^1 soluzioni per $k = 10$ perché per $k = 10$ risulta $rg(A) = rg(B) = 2$.

Tutti i sistemi con valori del parametro k diverso dai tre valori $-5, -7, 10$ hanno la medesima soluzione unica $x = 0, y = 0, z = 0$, ovvero possiedono come soluzione unica la *soluzione banale*. Per calcolare le ∞^1 soluzioni del sistema con $k = -5$, sostituiamo appunto il valore $k = -5$ nel sistema (3.47a) ottenendo così il sistema

$$\begin{cases} 14x - 4z = 0 \\ 0y = 0 \\ -4x - z = 0, \end{cases} \tag{3.47b}$$

la cui matrice associata A è

$$A = \begin{pmatrix} 14 & 0 & -4 \\ 0 & 0 & 0 \\ -4 & 0 & -1 \end{pmatrix}$$

e possiede *rango* 2. Se da tale sistema "estraiamo" la "parte" relativa alle incognite x, z nella prima e terza equazione "trasportando" l'incognita y al secondo membro con l'attribuzione del valore arbitrario $y = \alpha$, otteniamo il sistema

$$\begin{cases} 14x - 4z = 0 \\ -4x - z = 0, \end{cases}$$

da cui ricaviamo le ∞^1 soluzioni

$$\begin{cases} x = 0 \\ y = \alpha \\ z = 0 \end{cases}$$

di tutto il sistema (3.47b). Per calcolare le ∞^1 soluzioni del sistema con $k = -7$, sostituiamo appunto il valore $k = -7$ nel sistema (3.47a) ottenendo così il sistema

$$\begin{cases} 16x - 4z = 0 \\ 2y = 0 \\ -4x + z = 0, \end{cases} \tag{3.47c}$$

la cui matrice associata A è

$$A = \begin{pmatrix} 16 & 0 & -4 \\ 0 & 2 & 0 \\ -4 & 0 & 1 \end{pmatrix}$$

e possiede *rango* 2. Se da tale sistema "estraiamo" la "parte" relativa alle incognite x, y nelle prime due equazioni "trasportando" l'incognita z al secondo membro con l'attribuzione del valore arbitrario $z = \alpha$, otteniamo il sistema

$$\begin{cases} 16x = 4\alpha \\ 2y = 0 \end{cases}$$

da cui ricaviamo le ∞^1 soluzioni

$$\begin{cases} x = \alpha/4 \\ y = 0 \\ z = \alpha \end{cases}$$

di tutto il sistema (3.47c). Per calcolare le ∞^1 soluzioni del sistema con $k = 10$, sostituiamo appunto il valore $k = 10$ nel sistema (3.47a) ottenendo così il sistema

$$\begin{cases} -x - 4z = 0 \\ -15y = 0 \\ -4x - 16z = 0, \end{cases} \tag{3.47d}$$

la cui matrice associata A è

$$A = \begin{pmatrix} -1 & 0 & -4 \\ 0 & -15 & 0 \\ -4 & 0 & -16 \end{pmatrix}$$

e possiede *rango* 2. Se da tale sistema "estraiamo" la "parte" relativa alle incognite x, y nelle prime due equazioni "trasportando" l'incognita z al secondo membro con l'attribuzione del valore arbitrario $z = \alpha$, otteniamo il sistema

$$\begin{cases} -x = 4\alpha \\ -15y = 0 \end{cases}$$

da cui ricaviamo le ∞^1 soluzioni

$$\begin{cases} x = -4\alpha \\ y = 0 \\ z = \alpha \end{cases}$$

di tutto il sistema (3.47d).

Capitolo 4

Introduzione allo studio delle funzioni

Per rendere l'esposizione "meno complicata" possibile, possiamo considerare la Matematica, appunto per semplicità e senza pretesa di eccessiva precisione storico-cronologica, suddivisa in due parti, la prima delle quali è l'Algebra, sviluppatasi "più o meno" tra la fine del Quattrocento e poco dopo la metà del Cinquecento, e la seconda è la Teoria delle Funzioni, nata intorno agli anni Trenta del Seicento con l'introduzione del concetto di piano e di riferimento cartesiano da parte del grande Filosofo e Matematico francese Cartesio (1596-1650). L'Algebra ha come oggetto di studio la teoria delle equazioni, ovvero la ricerca dei numeri da attribuire all'incognita x per i quali si verifica l'uguaglianza $E(x) = 0$, dove $E(x)$ indica un'espressione contenente appunto l'incognita x.

Le espressioni $E(x)$ consistono in una sequenza di operazioni sull'incognita x e le operazioni, chiamate *funzioni elementari*, che vengono utilizzate per costruire una qualsiasi espressione sono la *potenza* x^α, l'*esponenziale* a^x, il *logaritmo* $\log_a x$ e le cosiddette *funzioni circolari*, quali $\sin x, \cos x, \tan x$ con le rispettive inverse $\arcsin x, \arccos x, \arctan x$.

Con le sole *potenze* x^n aventi esponenti $\alpha \equiv n$ interi positivi oppure zero si costruiscono tutti i possibili polinomi, indicati con $\mathcal{P}_n(x)$, aventi grado n e variabile x.

Se l'espressione $E(x)$ al primo membro dell'equazione $E(x) = 0$ è un polinomio di grado n nella variabile x, l'equazione $\mathcal{P}_n(x) = 0$ è chiamata *equazione algebrica* di grado n e i metodi per risolvere oppure per "tentare" di risolvere un'*equazione algebrica* sono solo

- l'*identità di secondo grado*,
- il *teorema di Ruffini*,
- un cambiamento di incognita per particolari equazioni di grado pari.

L'*identità di secondo grado* è un'identità che permette di scrivere ogni polinomio di secondo grado nella variabile x con coefficienti a, b, c, avente la nota forma $ax^2 + bx + c$, in una forma in cui la variabile x compare una sola volta, ovvero

$$ax^2 + bx + c \equiv a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right]. \tag{4.1}$$

Tra i "tanti pregi" che l'*identità di secondo grado* (4.1) possiede, vi è senz'altro quello di permettere di calcolare le due soluzioni dell'equazione di secondo grado perché se riscriviamo l'equazione di secondo grado $ax^2 + bx + c = 0$ nella forma

$$a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right] = 0$$

ed uguagliamo a zero il contenuto della parentesi quadra, ricaviamo immediatamente le due "famose" soluzioni

$$x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \qquad \text{e} \qquad x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \tag{4.2}$$

le quali dànno come somma

$$x_1 + x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} + \frac{-b + \sqrt{b^2 - 4ac}}{2a} = -\frac{b}{a}$$

e come prodotto

$$x_1 \cdot x_2 = \left(\frac{-b - \sqrt{b^2 - 4ac}}{2a}\right)\left(\frac{-b + \sqrt{b^2 - 4ac}}{2a}\right) = \frac{b^2 - (b^2 - 4ac)}{4a^2} = \frac{c}{a}.$$

Il *teorema di Ruffini* stabilisce che se un polinomio $\mathcal{P}_n(x)$ di grado n nella variabile x si annulla con la sostituzione alla x del numero $x = a$, allora segue che nella scomposizione del polinomio $\mathcal{P}_n(x)$ vi sarà il binomio di primo grado $x - a$ che si annulla per il medesimo numero a per il quale si annulla il polinomio $\mathcal{P}_n(x)$. E' importante sottolineare che il *teorema di Ruffini* non è una *formula* per risolvere le equazioni algebriche di grado n perché l'ipotesi del teorema è che un certo numero a attribuito alla x del polinomio annulli il polinomio stesso, ma il teorema non fornisce nessuna regola per trovare tale numero a.

Se applichiamo il *teorema di Ruffini* al polinomio di secondo grado $\mathcal{P}_2(x)$ che si annulla con la x uguale ai due numeri x_1, x_2 dati in (4.2), segue che la scomposizione del polinomio contiene i due binomi di primo grado $x - x_1$ e $x - x_2$, ovvero vale l'uguaglianza

$$ax^2 + bx + c = a(x - x_1)(x - x_2),$$

dove abbiamo riportato nel secondo membro il coefficiente a perché nella moltiplicazione tra i due binomi al secondo membro, il prodotto della x del primo binomio per la x del secondo binomio deve dare come risultato il termine ax^2 del primo membro.

In generale la scomposizone di un polinomio $\mathcal{Q}_n(x)$ di grado n avente a_n come coefficiente della potenza x^n e i k numeri $x_1, x_2, x_3, \ldots, x_k$ come *zeri*, è data da

$$\mathcal{Q}_n(x) = a_n(x - x_1)^{\alpha_1} \cdot (x - x_2)^{\alpha_2} \cdot (x - x_3)^{\alpha_3} \cdot \cdots \cdot (x - x_k)^{\alpha_k}, \tag{4.3}$$

con la condizione $\alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_k = n$. L'esponente α_i del binomio $x - x_i$ nella scomposizone (4.3) viene chiamato *molteplicità algebrica* del corrispondente *zero* x_i.

Nel caso in cui un binomio $x - x_i$ nella scomposizone (4.3) abbia esponente 1, ovvero nel caso in cui lo *zero* x_i del binomio $x - x_i$ possiede *molteplicità algebrica* 1 come *zero* del polinomio $\mathcal{Q}_n(x)$, chiamiamo lo *zero* x_i anche *zero semplice* del polinomio $\mathcal{Q}_n(x)$.

Poiché le soluzioni di ogni equazione di secondo grado sono date dai due numeri (4.2), segue che le equazioni di grado pari per le quali, nel caso in cui l'equazione abbia particolari proprietà, cerchiamo un possibile metodo risolutivo, sono le equazioni di grado pari uguale ad almeno 4. In particolare illustriamo le proprietà di due tipi di equazione avente grado pari, ovvero l'equazione, detta *equazione trinomia*, avente la forma $ax^{2n} + bx^n + c = 0$, e l'equazione di grado n pari, chiamata *equazione reciproca*, avente il medesimo coefficiente davanti ad ogni coppia di potenze (x^k, x^{n-k}), dove x^0 è la potenza sottintesa $x^0 \equiv 1$ avente come coefficiente il *termine noto* del polinomio al primo membro dell'equazione.

L'*equazione trinomia* si risolve mediante il cambiamento di incognita $x^n = y$, da cui segue che l'*equazione trinomia* diventa $ay^2 + by + c = 0$, le cui soluzioni y_1, y_2 forniscono le due equazioni, dette *equazioni binomie*, $x^n = y_1$ e $x^n = y_2$, che si risolvono mediante i metodi dell'*Analisi delle Variabili Complesse* e che forniscono n soluzioni ciascuna.

Per quanto riguarda le *equazioni reciproche*, abbiamo che quella di quarto grado ha il medesimo coefficiente a davanti alla coppia di potenze (x^4, x^0) e il medesimo coefficiente b davanti alla coppia di potenze (x^3, x), ovvero è della forma $ax^4 + bx^3 + cx^2 + bx + a = 0$, che possiamo riscrivere nella forma in cui ambo i membri sono divisi per x^2, ovvero

$$a\left(x^2 + \frac{1}{x^2}\right) + b\left(x + \frac{1}{x}\right) + c = 0. \tag{4.4a}$$

L'*equazione reciproca* di sesto grado ha il medesimo coefficiente a davanti alla coppia di potenze (x^6, x^0), il medesimo coefficiente b davanti alla coppia di potenze (x^5, x) e il medesimo coefficiente c davanti alla coppia di potenze (x^4, x^2), ovvero possiede la forma generale $ax^6 + bx^5 + cx^4 + dx^3 + cx^2 + bx + a = 0$, che possiamo riscrivere nella forma in cui ambo i membri sono divisi per x^3, ovvero

$$a\left(x^3 + \frac{1}{x^3}\right) + b\left(x^2 + \frac{1}{x^2}\right) + c\left(x + \frac{1}{x}\right) + d = 0. \tag{4.4b}$$

Infine l'*equazione reciproca* di ottavo grado ha il medesimo coefficiente a davanti alla coppia di potenze (x^8, x^0), il medesimo coefficiente b davanti alla coppia di potenze (x^7, x), il medesimo coefficiente c davanti alla coppia di potenze (x^6, x^2), il medesimo coefficiente d davanti alla coppia di potenze (x^5, x^3), ovvero possiede la forma generale

$$ax^8 + bx^7 + cx^6 + dx^5 + gx^4 + dx^3 + cx^2 + bx + a = 0,$$

che possiamo riscrivere nella forma in cui ambo i membri sono divisi per x^4, ovvero

$$a\left(x^4 + \frac{1}{x^4}\right) + b\left(x^3 + \frac{1}{x^3}\right) + c\left(x^2 + \frac{1}{x^2}\right) + d\left(x + \frac{1}{x}\right) + g = 0, \tag{4.4c}$$

e così via per ogni *equazione reciproca* di grado pari maggiore di 8.

Se effettuiamo il cambiamento di incognita

$$x + \frac{1}{x} = y, \tag{4.5}$$

otteniamo

$$x^2 + \frac{1}{x^2} = \left(x + \frac{1}{x}\right)^2 - 2 = y^2 - 2, \tag{4.6a}$$

$$x^3 + \frac{1}{x^3} = \left(x + \frac{1}{x}\right)^3 - 3\left(x + \frac{1}{x}\right) = y^3 - 3y, \tag{4.6b}$$

$$x^4 + \frac{1}{x^4} = \left(x + \frac{1}{x}\right)^4 - 4\left(x + \frac{1}{x}\right)^2 + 2 = y^4 - 4y^2 + 2, \tag{4.6c}$$

da cui, se sostituiamo nelle equazioni (4.4), segue che le tre equazioni (4.4) assumono rispettivamente le forme

$$
\begin{aligned}
&ay^2 + by + (c - 2a) = 0,\\
&ay^3 + by^2 + (c - 3a)y + (d - 2b) = 0,\\
&ay^4 + by^3 + (c - 4a)y^2 + (d - 3b)y + (2a - 2c + g) = 0,
\end{aligned}
$$

ovvero, a seguito del cambiamento di incognita (4.5) e degli sviluppi di tipo (4.6), ogni *equazione reciproca* di grado n pari si trasforma in una equazione avente grado $n/2$.

Può essere utile ricordare che lo sviluppo generale della potenza $(a+b)^n$, applicato nelle espressioni (4.6) al caso particolare della potenza

$$\left(x + \frac{1}{x}\right)^n,$$

è dato dalla somma degli $n+1$ monomi ordinati secondo la sequenza

$$a^n, \quad a^{n-1}b, \quad a^{n-2}b^2, \quad a^{n-3}b^3, \quad \ldots, \quad a^3b^{n-3}, \quad a^2b^{n-2}, \quad ab^{n-1}, \quad b^n,$$

ciascuno moltiplicato per il relativo coefficiente che si trova nella riga n-esima (riga n-esima corrispondente all'esponente n) del triangolo di Tartaglia avente la seguente forma

$$
\begin{array}{ll}
\text{riga } 1 & \qquad\qquad 1 \quad 1\\
\text{riga } 2 & \qquad\quad 1 \quad 2 \quad 1\\
\text{riga } 3 & \qquad 1 \quad 3 \quad 3 \quad 1\\
\text{riga } 4 & \quad 1 \quad 4 \quad 6 \quad 4 \quad 1\\
\text{riga } 5 & 1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1
\end{array}
\tag{4.7}
$$

e così via, in cui, ad eccezione degli elementi 1 estremi di ciascuna riga, ogni elemento a è la somma dei due elementi b, c nella riga immediatamente "al di sopra", sotto ai quali si trova l'elemento a nella riga immediatamente "al di sotto". Per esempio, la potenza $(a+b)^5$ ha come sviluppo il polinomio con sei monomi $(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$.

Per concludere questa sintesi dell'Algebra, dobbiamo tuttavia puntualizzare che in realtà anche per le equazioni algebriche di terzo e quarto grado esiste una formula risolutiva che consiste in una sequenza di operazioni con sole frazioni e radici, ma tali formule per risolvere le equazioni algebriche di terzo e quarto grado sono "così scomode" che "è come se non esistessero" e quindi in Algebra queste formule non vengono mai utilizzate.

Le formule per risolvere le equazioni algebriche di grado superiore al quarto non vennero mai trovate, ma i Matematici del Cinquecento non riuscirono a capire se essi non erano capaci di trovarle oppure se tali formule non esistevano.

Si dovettero attendere gli anni Trenta dell'Ottocento per veder risolta definitivamente la questione delle formule risolutive delle equazioni algebriche e si raggiunse un quadro completo della teoria delle equazioni algebriche allorquando il grande Matematico francese Évariste Galois (1811-1832) dimostrò che per le equazioni algebriche di grado superiore al quarto non può esistere nessuna formula risolutiva che consista in una sequenza di operazioni con sole frazioni e radici.

4.1 Concetto di funzione

Nel corso della seconda metà del Cinquecento possiamo dire che lo studio delle equazioni algebriche cominciò quindi a poco a poco a suscitare sempre minore interesse nelle ricerche dei Matematici di quel tempo e quando Cartesio, intorno agli anni Trenta del Seicento, introdusse sul *piano geometrico* il concetto di *sistema di riferimento*, detto appunto *sistema di riferimento cartesiano*, basato sulle *ascisse* e sulle *ordinate* riportate su due *assi* ortogonali, si aprì dunque una nuova epoca nello sviluppo della Matematica.

Un *piano geometrico* sul quale sia fissato un *sistema di riferimento cartesiano* individuato da due *assi* ortogonali, viene denominato semplicemente *piano cartesiano.*

Le espressioni $E(x)$, utilizzate fino a quel momento in Algebra come primo membro dell'equazione $E(x) = 0$, cominciarono, dopo l'introduzione appunto del *sistema di riferimento cartesiano*, ad essere utilizzate per *generare coppie di numeri*, il primo dei quali sia il valore attribuito alla variabile x e il secondo sia il risultato fornito dall'espressione considerata mediante le operazioni contenute nell'espressione stessa. Per il nostro studio, i numeri attribuiti alla variabile x apparterranno sempre ad un sottoinsieme $\mathcal{A}$ dell'insieme dei *numeri reali* $\mathbb{R}$ oppure anche a tutto l'insieme $\mathbb{R}$ dei *numeri reali* e saremo sempre interessati al solo caso in cui i risultati forniti dall'espressione siano *numeri reali.*

Per sottolineare questo nuovo e diverso modo di utilizzare un'espressione $E(x)$ già nota in Algebra, l'espressione $E(x)$ stessa venne denotata con il nuovo simbolo $f(x)$ e venne chiamata con il nuovo nome di *funzione.* Per indicare che $\mathcal{A}$ può essere sottoinsieme dell'insieme $\mathbb{R}$ dei *numeri reali* oppure può anche coincidere con tutto l'insieme $\mathbb{R}$ dei *numeri reali*, si utilizza la notazione $\mathcal{A} \subseteq \mathbb{R}$. Una *funzione* è quindi una *regola*, indicata usualmente nei testi con il simbolo $f : \mathcal{A} \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, che ad un numero appartenente ad un sottoinsieme $\mathcal{A} \subseteq \mathbb{R}$ associa un risultato che appartiene all'insieme $\mathbb{R}$ dei *numeri reali*, ovvero una funzione *genera coppie di numeri reali* secondo il seguente schema

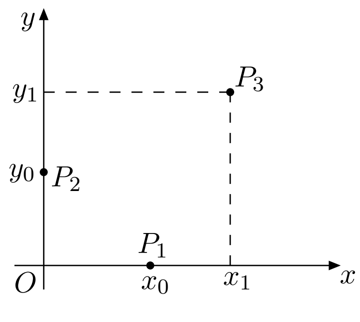
$$\begin{array}{c|c} x & f(x) \\ \hline x_1 & f(x_1) \\ x_2 & f(x_2) \\ x_3 & f(x_3) \\ \vdots & \vdots \end{array}$$

in cui sulla colonna di sinistra scriviamo i numeri $x_1, x_2, x_3, \ldots$ attribuiti alla variabile x della funzione $f(x)$ e sulla colonna di destra riportiamo i corrispondenti risultati che si ottengono eseguendo le operazioni contenute nella $f(x)$, ovvero $f(x_1)$ è il risultato che otteniamo attribuendo il numero x_1 alla variabile x della funzione, $f(x_2)$ è il risultato che otteniamo attribuendo il numero x_2 alla variabile x della funzione e così via.

L'insieme $\{(x_1, f(x_1)), (x_2, f(x_2)), (x_3, f(x_3)), \ldots\}$ delle *coppie* di numeri reali *generate* da una *funzione* $f(x)$ secondo tale schema è chiamato *grafico* della *funzione* $f(x)$.

E' utile ricordare la rappresentazione dei punti in un *sistema di riferimento cartesiano* avente origine $O(0, 0)$, dove un generico punto P_1 sull'asse x ha coordinate $P_1(x_0, 0)$, un generico punto P_2 sull'asse y ha coordinate $P_2(0, y_0)$ ed un generico punto P_3 non appartenente agli assi ha coordinate $P_3(x_1, y_1)$, come mostrato nella seguente figura 4.1.

Fig. 4.1



Poiché qualche operazione contenuta nella *funzione* può, in corrispondenza di qualche numero reale attribuito alla x, non dare risultato, chiamiamo *dominio* della *funzione* $f(x)$ l'insieme, indicato con $\mathfrak{D}$ di tutti e soli i numeri reali che, attribuiti alla variabile x della *funzione* stessa, forniscono come risultato un numero reale.

Il *dominio* di una *funzione* può essere chiamato anche *campo di esistenza* oppure *campo di definizione* della *funzione.* Ad esempio il *dominio* della funzione

$$f(x) = \frac{1}{x}$$

è l'insieme di tutti i numeri reali diversi da zero perché al denominatore x non si può attribuire valore zero ed indichiamo tale *dominio* con la scrittura

$$\mathfrak{D} = (-\infty, 0) \cup (0, \infty).$$

Il *dominio* della funzione

$$f(x) = \sqrt{\frac{4 - x^2}{x - 5}}$$

è l'insieme di tutti i numeri reali per i quali l'espressione sotto radice quadrata, chiamata *radicando*, sia positiva oppure zero ed indichiamo tale *dominio* con la scrittura

$$\mathfrak{D} = (-\infty, -2] \cup [2, 5).$$

Il *dominio* della funzione

$$f(x) = \frac{x - 1}{x^2 - 9}$$

è l'insieme di tutti i numeri reali diversi dai due numeri ± 3 perché il denominatore non può risultare zero ed indichiamo tale *dominio* con la scrittura

$$\mathfrak{D} = (-\infty, -3) \cup (-3, 3) \cup (3, \infty).$$

Il *dominio* della funzione

$$f(x) = \frac{x - 1}{\log(x^2 - 4)}$$

è l'insieme di tutti i numeri reali per i quali l'*argomento* del logaritmo risulta positivo e diverso da 1 perché il logaritmo con *argomento* uguale a 1 dà risultato zero al denominatore ed indichiamo tale *dominio* con la scrittura

$$\mathfrak{D} = (-\infty, -\sqrt{5}) \cup (-\sqrt{5}, -2) \cup (2, \sqrt{5}) \cup (\sqrt{5}, \infty).$$

Il *dominio* della funzione

$$f(x) = \frac{\sqrt{x^2 - 1}}{\log(2x + 5)}$$

è l'insieme di tutti i numeri reali per i quali l'*argomento* del logaritmo risulta positivo e diverso da 1 ed il radicando della radice quadrata risulta non negativo.

Indichiamo tale *dominio* con la scrittura

$$\mathfrak{D} = \left(-\frac{5}{2}, -2\right) \cup (-2, -1] \cup [1, \infty).$$

L'obiettivo dello *studio di una funzione* $f(x)$ è quello di determinare la *forma della curva* "costituita" da tutti i punti appartenenti al *grafico* della funzione, ovvero *generati* dalla funzione stessa secondo lo schema che abbiamo descritto per cui da un valore x_0 nel *dominio* attribuito alla x si ottiene dalla *funzione* il corrispondente risultato $f(x_0)$.

Con il termine *determinare la forma della curva* intendiamo che della *curva* generata dalla *funzione* $f(x)$ dobbiamo "riconoscere", in primo luogo, i "tratti" aventi quello che viene chiamato *andamento crescente* e i "tratti" aventi quello che viene chiamato *andamento decrescente*. Poichè abbiamo inoltre che ciascuno dei due *andamenti*, *crescente* e *decrescente*, può presentarsi in due "forme" denominate *convessa* e *concava*, dobbiamo "riconoscere", in secondo luogo, se l'*andamento crescente* risulta *convesso* oppure *concavo* e, analogamente, se l'*andamento decrescente* risulta *convesso* oppure *concavo*.

Talvolta, ove occorra, indicheremo con y il risultato *generato* da una funzione $f(x)$ con l'assegnazione di un numero, appartenente al *dominio*, alla variabile x, ovvero $y = f(x)$.

4.1.1 Polinomio di primo grado e retta

La prima *funzione* $f(x)$ che studiamo ha l'espressione "più semplice con la x", ovvero ha come espressione il polinomio di primo grado e quindi scriviamo $f(x) = mx + n$.

Per studiare tale *funzione*, consideriamo lo schema delle coppie di punti generate dalla funzione stessa

$$\begin{array}{c|c} x & f(x) = mx+n \\ \hline x_1 & f(x_1) = mx_1 + n \\ x_2 & f(x_2) = mx_2 + n \\ x_3 & f(x_3) = mx_3 + n \\ \vdots & \vdots \end{array}$$

e sviluppiamo l'espressione, chiamata *rapporto incrementale*, per una generica coppia di numeri reali distinti tra loro x_1, x_2 attribuiti alla variabile x

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{(mx_1 + n) - (mx_2 + n)}{x_1 - x_2} = \frac{mx_1 - mx_2}{x_1 - x_2} = \frac{m(x_1 - x_2)}{x_1 - x_2} = m.$$

Se dei punti consideriamo non solo l'ascissa, ovvero x_1, x_2, ma anche le corrispondenti coordinate, ovvero $(x_1, f(x_1))$ e $(x_2, f(x_2))$, possiamo dire che il *rapporto incrementale*

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} \tag{4.8}$$

di una funzione $f(x)$ in corrispondenza di due suoi generici punti $(x_1, f(x_1))$ e $(x_2, f(x_2))$ rappresenta il rapporto tra la *lunghezza con segno* $f(x_1) - f(x_2)$ del cateto verticale e la *lunghezza con segno* $x_1 - x_2$ del cateto orizzontale nel triangolo rettangolo "individuato" dai due punti, come mostrato nella seguente fig. 4.2

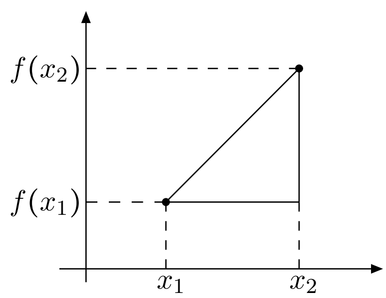


Fig. 4.2

dove, sebbene la *lunghezza* di un segmento sia un numero sempre positivo oppure zero, tuttavia con il termine *lunghezza con segno* si intende che le differenze $f(x_1)-f(x_2)$ e x_1-x_2 possono essere anche negative, purché in ogni caso l'ordine degli indici x_1, x_2 oppure x_2, x_1 sia sempre il medesimo nel numeratore e nel denominatore dell'espressione (4.8).

Poiché il risultato m del *rapporto incrementale* ottenuto dalla funzione $f(x) = mx+n$ è dunque il medesimo per ogni coppia di numeri x_i, x_j attribuiti alla x distinti tra loro, segue che se consideriamo tre generici punti $P_1(x_1, f(x_1)), P_2(x_2, f(x_2)), P_3(x_3, f(x_3))$ generati dalla funzione e scegliamo quindi due coppie qualsiasi di tali punti, ad esempio e senza perdita di generalità soltanto le due coppie $(P_1, P_2), (P_2, P_3)$, tra le tre possibili coppie totali $(P_1, P_2), (P_2, P_3), (P_1, P_3)$, ricaviamo l'informazione che i triangoli rettangoli "individuati" dai due punti (P_1, P_2) e dai due punti (P_2, P_3) hanno il medesimo rapporto m tra le *lunghezze con segno* del cateto verticale e del cateto orizzontale. Possiamo pertanto concludere che i tre punti $P_1(x_1, f(x_1)), P_2(x_2, f(x_2)), P_3(x_3, f(x_3))$ appartengono alla medesima retta perché, come mostrato nella seguente fig. 4.3, il triangolo rettangolo individuato dai due punti P_1, P_2 e il triangolo rettangolo individuato dai due punti P_2, P_3, avendo il medesimo rapporto m tra le *lunghezze con segno* del cateto verticale e del cateto orizzontale, risultano *simili* tra loro e possiedono quindi il medesimo angolo α.

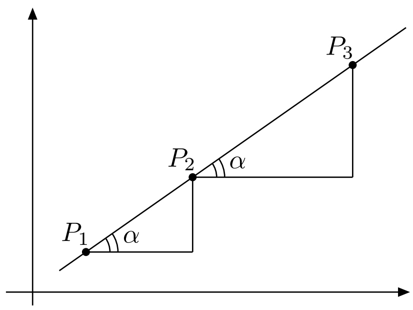


Fig. 4.3

Quindi, se il punto P_3 appartiene alla medesima retta dei due punti P_1, P_2, segue che ogni punto generato dalla funzione $f(x) = mx + n$ appartiene alla medesima retta dei due punti P_1, P_2, ovvero che la funzione *polinomio di primo grado* $f(x) = mx + n$, chiamata anche *funzione affine*, genera una retta. Può essere utile ricordare che, dati due punti sul *piano cartesiano* $P(x_1, y_1)$ e $Q(x_2, y_2)$, con $x_1 \neq x_2$ e $y_1 \neq y_2$, la retta passante per P, Q ha equazione cartesiana che può essere scritta nelle due forme identiche tra loro

$$\frac{y-y_1}{y_2-y_1} = \frac{x-x_1}{x_2-x_1} \qquad \text{oppure} \qquad \frac{y-y_2}{y_2-y_1} = \frac{x-x_2}{x_2-x_1},$$

da cui seguono corrispondentemente le due forme esplicite identiche tra loro

$$y = \frac{y_1-y_2}{x_1-x_2}\,x + \left[y_1 - \frac{y_1-y_2}{x_1-x_2}\,x_1\right] \tag{4.9a}$$

oppure

$$y = \frac{y_1-y_2}{x_1-x_2}\,x + \left[y_2 - \frac{y_1-y_2}{x_1-x_2}\,x_2\right] \tag{4.9b}$$

e lasciamo come semplice esercizio la verifica che entrambe le equazioni (4.9) soddisfano la condizione di passaggio per i due punti $P(x_1, y_1)$ e $Q(x_2, y_2)$, con $x_1 \neq x_2$ e $y_1 \neq y_2$.

Se due punti distinti hanno la medesima ascissa, che indichiamo con x_0, segue che essi appartengono ad una retta verticale ed in particolare alla retta verticale avente equazione cartesiana $x = x_0$. Se invece due punti distinti hanno la medesima ordinata, che indichiamo con y_0, segue che essi appartengono ad una retta orizzontale ed in particolare alla retta orizzontale avente equazione cartesiana $y = y_0$. Può essere utile osservare infine che l'equazione $y = y_0$ della retta orizzontale si ottiene dalle equazioni (4.9) per $y_1 = y_2$.

4.1.2 Proprietà della funzione polinomio di primo grado

La *funzione polinomio di primo grado* con termine noto uguale a zero viene denominata *funzione lineare*, mentre la *funzione polinomio di primo grado* con termine noto diverso da zero viene denominata *funzione affine*. Se una *funzione* $f(x)$ verifica l'uguaglianza

$$f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2),$$

per ogni valore reale $\alpha_1, \alpha_2, x_1, x_2$, dove il numero $\alpha_1 x_1 + \alpha_2 x_2$ è, come si riconosce, una combinazione lineare dei due numeri x_1, x_2 con coefficienti rispettivamente α_1, α_2, si dice che la *funzione* $f(x)$ *conserva tutte le combinazioni lineari* e tale terminologia significa che il risultato $f(\alpha_1 x_1 + \alpha_2 x_2)$ che la *funzione* $f(x)$ fornisce quando alla sua variabile x si attribuisce la combinazione lineare $x = \alpha_1 x_1 + \alpha_2 x_2$, coincide con la combinazione lineare $\alpha_1 f(x_1) + \alpha_2 f(x_2)$ dei due risultati $f(x_1), f(x_2)$. Dimostriamo quindi l'importante proprietà della *funzione lineare* $f(x) = mx$ di *conservare* tutte le combinazioni lineari.

Abbiamo infatti

$$f(\alpha_1 x_1 + \alpha_2 x_2) = m(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1(m x_1) + \alpha_2(m x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2),$$

perché risulta $mx_1 = f(x_1)$ e $mx_2 = f(x_2)$. Dimostriamo ora quindi la proprietà della *funzione affine* $f(x) = mx + n$ di *conservare* tutte e sole quelle combinazioni lineari i cui coefficienti α_1, α_2 soddisfino la condizione che la loro somma sia 1. Infatti poiché per la *funzione affine* risulta $mx_1 = f(x_1) - n$ e $mx_2 = f(x_2) - n$, otteniamo

$$\begin{gathered} f(\alpha_1 x_1 + \alpha_2 x_2) = m(\alpha_1 x_1 + \alpha_2 x_2) + n = \alpha_1(m x_1) + \alpha_2(m x_2) + n = \\ = \alpha_1 [f(x_1) - n] + \alpha_2 [f(x_2) - n] + n = \alpha_1 f(x_1) + \alpha_2 f(x_2) + [1 - (\alpha_1 + \alpha_2)] n \end{gathered}$$

ed è immediato rendersi conto che se la somma $\alpha_1 + \alpha_2$ dei coefficienti è uguale a 1, segue che il termine nell'ultima parentesi quadra è zero e quindi la *funzione* $f(x) = mx + n$ verifica la condizione di conservazione delle combinazioni lineari $f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2)$.

4.1.3 Andamenti crescenti, decrescenti, convessi e concavi

Poiché dunque una *funzione* $f(x)$ *genera*, nei casi di interesse del "nostro" studio, infiniti punti che costituiscono una *curva* sul *piano cartesiano*, nel seguito utilizzeremo i termini *funzione* e *curva* come "sinonimi". Per poter "riconoscere" l'*andamento crescente* e l'*andamento decrescente* di un "tratto" di una *funzione*, occorre trovare preliminarmente una loro precisa definizione matematica che sia in accordo con l'*idea* "intuitiva" e "visiva" di *andamento crescente* caratterizzato dal "comportamento" da sinistra a destra "verso l'alto" e con l'*idea* "intuitiva" e "visiva" di *andamento decrescente* caratterizzato dal "comportamento" da sinistra a destra "verso il basso", come mostrato rispettivamente nelle seguenti figg. 4.4(a) e 4.4(b).

Fig. 4.4(a)

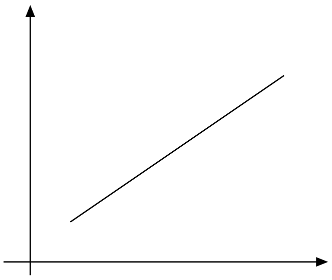
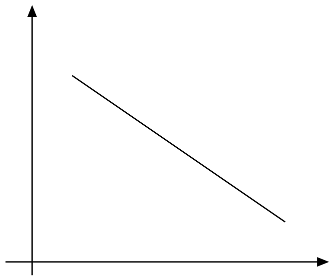


Fig. 4.4(b)



Per definire in maniera rigorosa e formale l'*andamento crescente* e l'*andamento decrescente* di una funzione $f(x)$ in un intervallo $[a,b]$, consideriamo le seguenti figg. 4.4(c) e 4.4(d) e osserviamo che l'*idea* "visiva" di *andamento crescente* che "abbiamo nella nostra mente" di un "tratto" di *funzione* $f(x)$, compreso tra due estremi aventi ascisse a,b, corrisponde alla proprietà secondo cui da ogni coppia di punti x_1, x_2 compresi nell'intervallo $[a,b]$, segue che l'ascissa maggiore, ovvero il valore x_2 nella fig. 4.4(c), dà risultato $f(x_2)$ maggiore e l'ascissa minore, ovvero il valore x_1 nella fig. 4.4(c), dà risultato $f(x_1)$ minore.

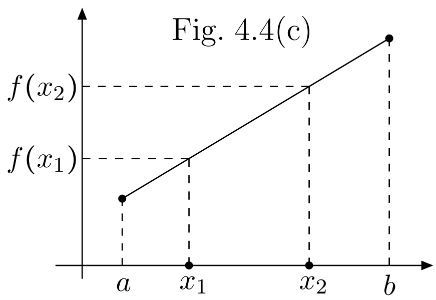


Fig. 4.4(c)

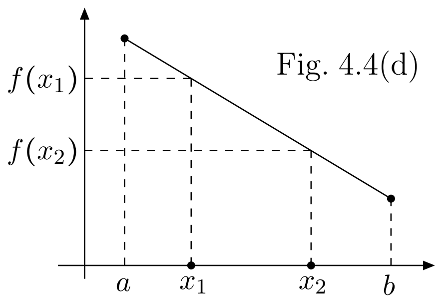


Fig. 4.4(d)

Possiamo dire quindi che in un "tratto" *crescente*, compreso in un intervallo $[a,b]$, di una *funzione* $f(x)$, la relazione di disuguaglianza che intercorre tra i risultati $f(x_1), f(x_2)$ *generati* da qualsiasi coppia di ascisse, rispettivamente x_1, x_2, appartenenti all'intervallo, coincide con la relazione di disuguaglianza che intercorre tra le ascisse medesime.

Questa proprietà può essere espressa pertanto matematicamente mediante la seguente

Definizione 4.1.1. *(Funzione crescente in un intervallo) Una funzione f(x) definita in un intervallo $[a,b]$ si dice "crescente nell'intervallo $[a,b]$" se vale la seguente condizione*

per ogni coppia di valori x_1, x_2 compresi nell'intervallo $[a,b]$ risulta

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \geqslant 0. \tag{4.10a}$$

Analogamente osserviamo che l'*idea* "visiva" di *andamento decrescente* di un "tratto" di *funzione* $f(x)$, compreso tra due estremi aventi ascisse a,b, corrisponde alla proprietà secondo cui da ogni coppia di punti x_1, x_2 compresi nell'intervallo $[a,b]$, segue che l'ascissa maggiore, ovvero il valore x_2 nella fig. 4.4(d), dà risultato $f(x_2)$ minore e l'ascissa minore, ovvero il valore x_1 nella fig. 4.4(d), dà risultato $f(x_1)$ maggiore.

Possiamo dire quindi che in un "tratto" *decrescente*, compreso in un intervallo $[a,b]$, di una *funzione* $f(x)$, la relazione di disuguaglianza che intercorre tra i risultati $f(x_1), f(x_2)$ *generati* da qualsiasi coppia di ascisse, rispettivamente x_1, x_2, appartenenti all'intervallo, è "scambiata" rispetto alla relazione di disuguaglianza che intercorre tra le ascisse medesime.

Questa proprietà può essere espressa pertanto matematicamente mediante la seguente

Definizione 4.1.2. *(Funzione decrescente in un intervallo) Una funzione f(x) definita in un intervallo* $[a, b]$ *si dice "decrescente nell'intervallo* $[a, b]$*" se vale la seguente condizione*

per ogni coppia di valori x_1, x_2 *compresi nell'intervallo* $[a, b]$ *risulta*

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leqslant 0. \tag{4.10b}$$

L'*idea* "visiva" dell'*andamento convesso* e dell'*andamento concavo* di una funzione $f(x)$ in un intervallo $[a, b]$ è riportata nelle seguenti figg. 4.5(a),4.5(b)

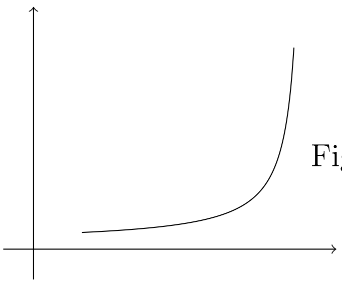


Fig. 4.5(a)

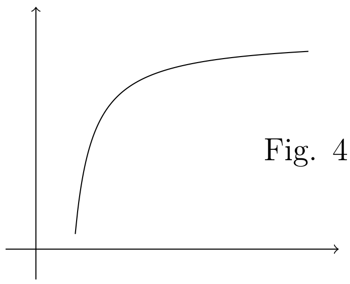


Fig. 4.5(b)

per il caso di *comportamento crescente* e nelle seguenti figg. 4.5(c),4.5(d)

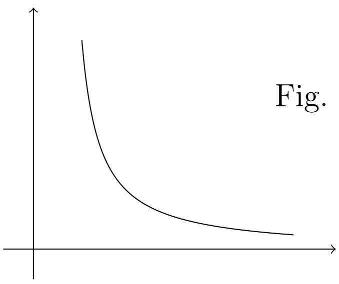


Fig. 4.5(c)

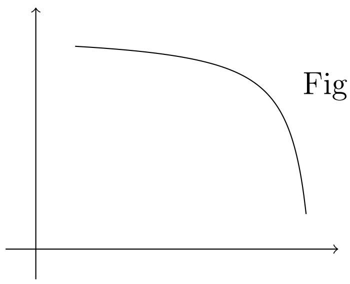


Fig. 4.5(d)

per il caso di *comportamento decrescente.*

Poiché dunque l'*andamento crescente* e l'*andamento decrescente* possono avere separatamente sia il *comportamento convesso* che il *comportamento concavo*, segue che per definire in modo formale il *comportamento convesso* e il *comportamento concavo* di un "tratto" di curva, non occorre che il "tratto" abbia solo l'*andamento crescente* oppure solo l'*andamento decrescente* e possiamo fissare il concetto di *comportamento convesso* e di *comportamento concavo* riferendoci quindi rispettivamente alla fig. 4.6(a) e alla fig. 4.6(b).

Per definire il concetto di *comportamento convesso* e di *comportamento concavo* di un "tratto" di curva associata alla funzione $f(x)$ all'interno di un intervallo $[a, b]$, utilizziamo quindi le due figg. 4.6(a) e 4.6(b) in cui all'interno dell'intervallo $[a, b]$ stesso sono state scelte due generiche ascisse x_1, x_2 da cui si ottengono i corrispondenti due punti P, Q sulla curva aventi coordinate rispettivamente $P(x_1, f(x_1))$ e $Q(x_2, f(x_2))$.

Se consideriamo quindi la retta s secante la curva nei due punti P, Q, diciamo che il "tratto" di curva $f(x)$ all'interno dell'intervallo $[a, b]$ è *convesso* se l'*arco* P, Q sulla curva si trova "tutto al di sotto" rispetto al segmento P, Q sulla retta secante s, come si vede riportato in fig. 4.6(a). Se invece l'*arco* P, Q sulla curva si trova "tutto al di sopra" rispetto al segmento P, Q sulla retta secante s, diciamo che il "tratto" di curva $f(x)$ all'interno dell'intervallo $[a, b]$ è *concavo*, come si vede riportato in fig. 4.6(b).

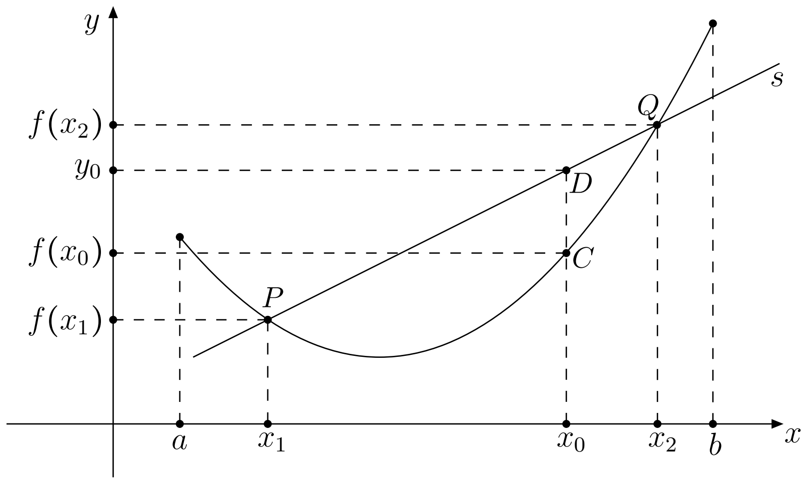


fig. 4.6(a)

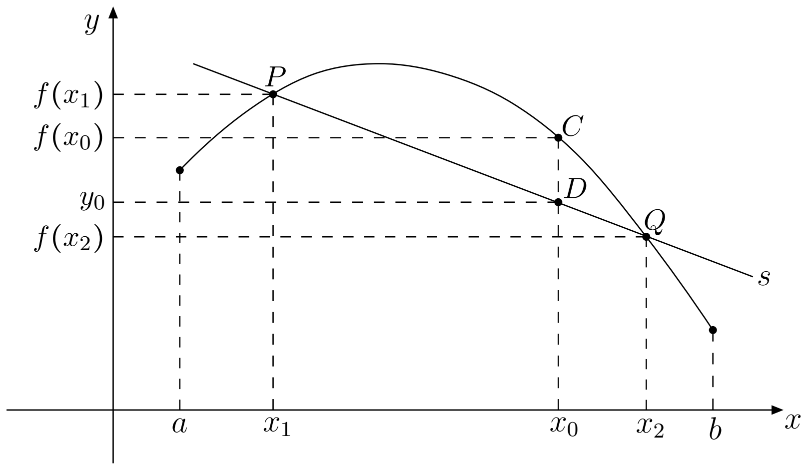


fig. 4.6(b)

Per stabilire se l'*arco* P, Q sulla curva si trova "tutto al di sopra" oppure "tutto al di sotto" rispetto al segmento P, Q sulla retta secante s, consideriamo una generica ascissa x_0 appartenente all'intervallo $[x_1, x_2]$ e confrontiamo la corrispondente ordinata $f(x_0)$ del punto C sulla curva con la corrispondente ordinata y_0 del punto D sulla retta s secante, in modo tale da poter concludere che la funzione $f(x)$ è *convessa* in tutto l'intervallo $[a, b]$ se l'ordinata $f(x_0)$ del punto C sulla curva è minore dell'ordinata y_0 del punto D sulla retta secante s, mentre la funzione $f(x)$ è *concava* in tutto l'intervallo $[a, b]$ se l'ordinata $f(x_0)$ del punto C sulla curva è maggiore dell'ordinata y_0 del punto D sulla retta s secante.

Se utilizziamo, in particolare, la prima delle due forme (4.9) per l'equazione di una retta passante per i due punti aventi coordinate $(x_1, f(x_1))$ e $(x_2, f(x_2))$, abbiamo che l'equazione cartesiana della retta s secante nelle due figg. 4.6 è

$$y = \left[\frac{f(x_2) - f(x_1)}{x_2 - x_1}\right] x + \left[f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}\, x_1\right], \tag{4.11}$$

e possiamo enunciare dunque le definizioni di funzione *convessa* e *concava* in $[a, b]$.

Definizione 4.1.3. *Una funzione $f(x)$ si dice **convessa** nell'intervallo $[a, b]$ se per ogni tripla di numeri x_1, x_0, x_2, con $a \leqslant x_1 < x_0 < x_2 \leqslant b$, risulta*

$$f(x_0) \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1}\, x_0 + f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}\, x_1 \,. \tag{4.12a}$$

Definizione 4.1.4. *Una funzione $f(x)$ si dice* **concava** *nell'intervallo $[a,b]$ se per ogni tripla di numeri x_1, x_0, x_2, con $a \leqslant x_1 < x_0 < x_2 \leqslant b$, risulta*

$$f(x_0) \geqslant \frac{f(x_2)-f(x_1)}{x_2-x_1}\,x_0 + f(x_1) - \frac{f(x_2)-f(x_1)}{x_2-x_1}\,x_1\,. \tag{4.12b}$$

Dopo aver enunciato le definizioni di funzione *crescente, decrescente, convessa* e *concava* in un intervallo, potremmo, in linea di principio, studiare tutte le funzioni, ma si pongono due problemi, ovvero in primo luogo che, data una funzione $f(x)$, non conosciamo "a priori" un intervallo $[a,b]$ in cui essa ha uno dei possibili diversi *comportamenti* e, in secondo luogo che, anche ove fosse assegnato un intervallo $[a,b]$, le disuguaglianze (4.10) per ogni $a \leqslant x_1, x_2 \leqslant b$ e le disuguaglianze (4.12) per ogni $a \leqslant x_1 < x_0 < x_2 \leqslant b$ non permettono, in generale per "funzioni complicate", di ottenere risultati.

Chiamiamo dunque *funzioni della geometria analitica* quelle "poche" funzioni per le quali il *comportamento crescente, decrescente, convesso* e *concavo* può essere stabilito mediante le disuguaglianze (4.10) e (4.12) in modo "puramente" algebrico.

4.2 Funzioni della geometria analitica

Le *funzioni della geometria analitica* che studieremo sono la *funzione polinomio di secondo grado $f(x) = ax^2 + bx + c$, la funzione potenza $f(x) = x^n$* con esponente intero positivo, *la funzione potenza $f(x) = x^{-n} = 1/x^n$* con esponente intero negativo, *la funzione esponenziale $f(x) = a^x$*, con base $0 < a < 1$ oppure $a > 1$, *la funzione logaritmo $f(x) = \log_a x$*, con base $0 < a < 1$ oppure $a > 1$. Come vedremo, la *forma* di tali curve sarà ottenuta mediante le sole proprietà algebriche delle espressioni stesse che compongono ciascuna *funzione.*

4.2.1 La funzione polinomio di secondo grado

Com'è ben noto, laa *funzione polinomio di secondo grado* $f(x) = ax^2 + bx + c$ genera punti che formano quella curva che viene denominata *parabola.*

Osservando che il *dominio* della *funzione* è tutto l'insieme dei numeri reali, studiamo l'espressione (4.8) del *rapporto incrementale* per stabilire in quali intervalli la $f(x)$ è crescente oppure decrescente secondo le disuguaglianze (4.10). Per $x_1 \neq x_2$ abbiamo

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} = \frac{(ax_1^2+bx_1+c)-(ax_2^2+bx_2+c)}{x_1-x_2} = \frac{a(x_1^2-x_2^2)+b(x_1-x_2)}{x_1-x_2} =$$

$$= \frac{a(x_1-x_2)(x_1+x_2)+b(x_1-x_2)}{x_1-x_2} = \frac{(x_1-x_2)[a(x_1+x_2)+b]}{x_1-x_2} = a(x_1+x_2)+b$$

e distinguiamo allora i due casi $a > 0$ e $a < 0$ ricordando l'ovvia proprietà della somma di due numeri secondo cui la somma di due numeri è maggiore di un numero N se e solo se ciascuno dei due numeri è maggiore di $N/2$, mentre la somma di due numeri è minore di un numero N se e solo se ciascuno dei due numeri è minore di $N/2$.

Poiché abbiamo dunque lo sviluppo

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} = a(x_1+x_2)+b = a\left(x_1+x_2+\frac{b}{a}\right) = a\left[x_1+x_2-\left(-\frac{b}{a}\right)\right],$$

segue che nel caso in cui il coefficiente a sia positivo, il segno del *rapporto incrementale* al primo membro coincide con il segno della "parentesi quadra" al terzo membro.

Applicando la proprietà della somma di due numeri ricordata in precedenza per la quale si consideri $N = -b/a$, possiamo concludere quindi che per tutte le coppie di numeri x_1, x_2 minori di $N/2 = -b/(2a)$, il *rapporto incrementale* è negativo, mentre per tutte le coppie di numeri x_1, x_2 maggiori di $N/2 = -b/(2a)$, il *rapporto incrementale* è positivo, da cui segue che la *parabola* con coefficiente $a > 0$ risulta decrescente per $x < -b/(2a)$ e risulta crescente per $x > -b/(2a)$, essendo quindi $x = -b/(2a)$ la nota ascissa del *vertice*.

Se invece il coefficiente a è negativo, il segno del *rapporto incrementale* al primo membro è opposto rispetto al segno della "parentesi quadra" al terzo membro.

Pertanto possiamo concludere che per tutte le coppie di numeri x_1, x_2 minori del numero $N/2 = -b/(2a)$, il *rapporto incrementale* risulta positivo, mentre per tutte le coppie di numeri x_1, x_2 maggiori di $N/2 = -b/(2a)$, il *rapporto incrementale* risulta negativo, da cui segue che la *parabola* con coefficiente $a < 0$ risulta crescente per $x < -b/(2a)$ e risulta decrescente per $x > -b/(2a)$, essendo $x = -b/(2a)$ sempre l'ascissa del *vertice*.

Per stabilire in quali intervalli la $f(x)$ sia *convessa* oppure *concava*, scriviamo la retta secante che passa per i punti della *parabola* $P(x_1, ax_1^2 + bx_1 + c)$ e $Q(x_2, ax_2^2 + bx_2 + c)$, la cui equazione cartesiana ha la forma dell'espressione (4.11) ed è pertanto

$$y = [a(x_1 + x_2) + b]\, x + \left\{ax_1^2 + bx_1 + c - [a(x_1 + x_2) + b]\, x_1\right\},$$

da cui segue l'ordinata $y_0 = [a(x_1 + x_2) + b]\, x_0 + \{ax_1^2 + bx_1 + c - [a(x_1 + x_2) + b]\, x_1\}$ del punto D sulla retta secante avente ascissa x_0, com'è mostrato nelle figg. 4.6, e quindi lo sviluppo della differenza tra l'ordinata y_0 del punto D e l'ordinata $f(x_0)$ del punto C, corrispondenti all'ascissa x_0, è

$$\begin{aligned} y_0 - f(x_0) &= [a(x_1 + x_2) + b]\, x_0 + \{ax_1^2 + bx_1 + c - [a(x_1 + x_2) + b]\, x_1\} - (ax_0^2 + bx_0 + c) = \\ &= ax_0x_1 + ax_0x_2 - ax_1x_2 - ax_0^2 = a(x_1 - x_0)(x_0 - x_2). \end{aligned}$$

Poiché, com'è mostrato sempre nelle figg. 4.6, l'ascissa x_0 è un'ascissa generica compresa tra le due generiche ascisse x_1, x_2 a loro volta comprese nell'intervallo $[a, b]$, ovvero le tre ascisse x_1, x_0, x_2 sono ordinati nella sequenza $x_1 < x_0 < x_2$, segue che entrambi i binomi $x_0 - x_2$ e $x_1 - x_0$ sono negativi e che quindi il loro prodotto è sempre positivo.

Pertanto, se il coefficiente a della *funzione* è positivo, segue che la differenza $y_0 - f(x_0)$ è positiva, ovvero che il segmento CD sulla retta secante si trova "al di sopra" dell'*arco* di curva tra i medesimi punti C, D e che quindi, data l'arbitrarietà dei tre punti x_1, x_0, x_2 su tutto l'insieme dei numeri reali, tutta la *parabola* è *convessa*. Se invece il coefficiente a è negativo, segue che la differenza $y_0 - f(x_0)$ è negativa, ovvero che il segmento CD sulla retta secante si trova "al di sotto" dell'*arco* di curva tra i medesimi punti C, D e che quindi, data l'arbitrarietà dei tre punti x_1, x_0, x_2 su tutto l'insieme dei numeri reali, tutta la *parabola* è *concava*. Utilizzando infine l'*identità di secondo grado* (4.1) per riscrivere la funzione $f(x) = ax^2 + bx + c$ nella forma equivalente

$$f(x) = a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right], \tag{4.13}$$

otteniamo inoltre l'informazione che da ogni coppia $-b\,/\,(2a) - \delta$ e $-b\,/\,(2a) + \delta$ di ascisse simmetriche rispetto al vertice $x_V = -b/(2a)$, sostituite alla variabile x, si ricava sempre il medesimo risultato perché si ha

$$a\left[\left(-\frac{b}{2a}\pm\delta+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a^2}\right]=a\left(\delta^2-\frac{b^2-4ac}{4a^2}\right),$$

da cui si conclude che in ogni *parabola* il “tratto” decrescente e quello crescente sono sempre l'uno il simmetrico dell'altro rispetto alla retta verticale, denominata *asse della parabola*, passante per il *vertice* ed avente equazione cartesiana $x = -b/(2a)$. Per rappresentare graficamente una *parabola* basta dunque rappresentare il suo *vertice* V avente coordinate

$$V\left(-\frac{b}{2a},-\frac{b^2-4ac}{4a}\right),$$

dove l'ordinata si ottiene sostituendo l'ascissa $-b/(2a)$ alla x nell'espressione (4.13), e tracciare quindi la curva in modo che essa sia, per $a > 0$, *decrescente* e *crescente* rispettivamente a sinistra e a destra del *vertice* e tutta *convessa*, come mostrato nella figura 4.7(a), mentre sia, per $a < 0$, *crescente* e *decrescente* rispettivamente a sinistra e a destra del *vertice* e tutta *concava*, come mostrato nella figura 4.7(b).

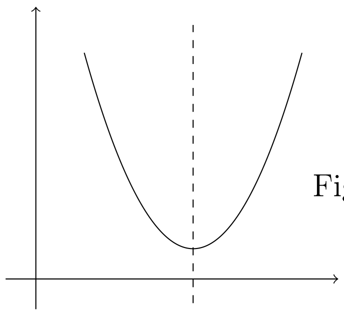


Fig. 4.7(a)

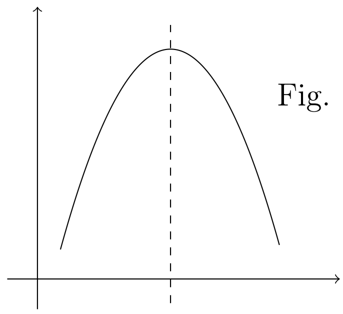


Fig. 4.7(b)

4.2.2 La funzione potenza con esponente intero positivo

La seconda *funzione* che studiamo è la *funzione potenza* $f(x) = x^k$ con esponente k intero positivo maggiore di 2, per la quale distinguiamo il caso degli esponenti pari $2n{+}2$ e il caso degli esponenti dispari $2n{+}1$, con $n = 1, 2, 3, \ldots$ Nel caso degli esponenti interi positivi pari, studiamo le *funzioni* $f(x) = x^4, f(x) = x^6, f(x) = x^8$ per ricavare la “regola ricorsiva” da “estendere” a tutti i successivi esponenti pari. In virtù del *teorema di Ruffini* le scomposizioni dei polinomi $x_1^4 - x_2^4, x_1^6 - x_2^6, x_1^8 - x_2^8$ sono

$$\begin{aligned}&x_1^4-x_2^4=(x_1-x_2)(x_1+x_2)\left(x_1^2+x_2^2\right),\quad x_1^6-x_2^6=(x_1-x_2)(x_1+x_2)\left(x_1^4+x_1^2x_2^2+x_2^4\right),\\&x_1^8-x_2^8=(x_1-x_2)(x_1+x_2)\left(x_1^6+x_1^4x_2^2+x_1^2x_2^4+x_2^6\right),\end{aligned}$$

da cui segue che la scomposizione di ogni polinomio $x_1^{2n} - x_2^{2n}$, dove $2n$ indica un numero pari per qualsiasi n intero positivo, è sempre data dal prodotto dei tre fattori $x_1{-}x_2, x_1{+}x_2$ e di un terzo polinomio che è sempre positivo in quanto somma di tutti quadrati.

Poiché ogni *funzione* $f(x) = x^{2n}$ ha come *dominio* tutto l'insieme dei numeri reali e *genera* il punto $(0, 0)$ in quanto $f(0) = 0$, possiamo studiare il *rapporto incrementale* solo per coppie di numeri distinti x_1, x_2 entrambi diversi da zero.

Poiché i *rapporti incrementali* delle *funzioni* $f(x) = x^4, f(x) = x^6$ e $f(x) = x^8$ sono

$$\frac{x_1^4-x_2^4}{x_1-x_2}=\frac{(x_1-x_2)(x_1+x_2)\left(x_1^2+x_2^2\right)}{x_1-x_2}=(x_1+x_2)\left(x_1^2+x_2^2\right),$$

$$\frac{x_1^6 - x_2^6}{x_1 - x_2} = \frac{(x_1 - x_2)(x_1 + x_2)\left(x_1^4 + x_1^2 x_2^2 + x_2^4\right)}{x_1 - x_2} = (x_1 + x_2)\left(x_1^4 + x_1^2 x_2^2 + x_2^4\right)$$

e così via per tutte le successive *funzioni* $f(x) = x^{2n}$ con esponente intero positivo pari, possiamo riconoscere che essi sono sempre negativi per ogni coppia di numeri x_1, x_2 entrambi negativi e sempre positivi per ogni coppia di numeri x_1, x_2 entrambi positivi, da cui segue che le *funzioni* $f(x) = x^{2n}$ risultano tutte *decrescenti* sulla semiretta delle x negative e *crescenti* sulla semiretta delle x positive. Per studiare il *comportamento convesso* oppure *concavo* di tali *funzioni*, consideriamo la *funzione* $f(x) = x^4$ e dopo aver scritto l'equazione della retta ad essa secante nei suoi due generici punti $(x_1, x_1^4), (x_2, x_2^4)$, con x_1, x_2 distinti ed entrambi diversi da zero, secondo l'espressione (4.11), esaminiamo la differenza $y_0 - f(x_0)$ tra l'ordinata y_0 del punto D e l'ordinata $f(x_0) = x_0^4$ del punto C riportati nelle figg. 4.6. Lo sviluppo di tale differenza e la relativa scomposizione dànno

$$\begin{aligned} y_0 - f(x_0) &= (x_1 + x_2)\left(x_1^2 + x_2^2\right)x_0 + x_1^4 - (x_1 + x_2)\left(x_1^2 + x_2^2\right)x_1 - x_0^4 = \\ &= (x_1 - x_0)(x_0 - x_2)\left(x_0^2 + x_1^2 + x_2^2 + x_0 x_1 + x_0 x_2 + x_1 x_2\right) \end{aligned}$$

e poiché per $x_1 < x_0 < x_2$ il prodotto $(x_1 - x_0)(x_0 - x_2)$ risulta sempre positivo, segue che per $x_1 < x_0 < x_2 < 0$ e per $0 < x_1 < x_0 < x_2$ la differenza $y_0 - f(x_0)$ risulta sempre positiva perché per x_1, x_0, x_2 concordi negativi oppure positivi anche il terzo fattore $x_0^2 + x_1^2 + x_2^2 + x_0 x_1 + x_0 x_2 + x_1 x_2$ è positivo in quanto somma di addendi tutti positivi.

Se invece due numeri dei tre x_1, x_0, x_2 hanno un segno e il terzo numero ha l'altro segno, ovvero x_1, x_0 sono negativi e $x_2 > 0$, oppure x_1 è negativo e x_0, x_2 sono positivi, segue in ogni caso che nel fattore $x_0^2 + x_1^2 + x_2^2 + x_0 x_1 + x_0 x_2 + x_1 x_2$ due monomi dei tre $x_0 x_1, x_0 x_2, x_1 x_2$ sono negativi ed il terzo è positivo. Facendo una scelta, che tuttavia è del tutto generale, possiamo considerare che siano x_1 negativo e x_0, x_2 positivi, da cui otteniamo che i due monomi $x_0 x_1, x_1 x_2$ sono negativi e $x_0 x_2$ è positivo. Si deduce quindi "facilmente" che anche in questo caso il terzo fattore $x_0^2 + x_1^2 + x_2^2 + x_0 x_1 + x_0 x_2 + x_1 x_2$ risulta positivo perché, tramite l'*identità di secondo grado* (4.1) relativa al polinomio con due variabili, esso può essere riscritto nella forma

$$\begin{aligned} x_0^2 + x_1^2 + x_2^2 + x_0 x_1 + x_0 x_2 + x_1 x_2 &= x_1^2 + \left[\left(x_0 + \frac{x_1}{2}\right)^2 - \frac{x_1^2}{4}\right] + \left[\left(x_2 + \frac{x_1}{2}\right)^2 - \frac{x_1^2}{4}\right] + x_0 x_2 = \\ &= \frac{x_1^2}{2} + \left(x_0 + \frac{x_1}{2}\right)^2 + \left(x_2 + \frac{x_1}{2}\right)^2 + x_0 x_2 \,, \end{aligned}$$

ovvero nella somma di addendi tutti positivi. Per la *funzione* $f(x) = x^6$ lo studio del segno della differenza $y_0 - f(x_0)$ è "molto più complicato" e quindi lo omettiamo concludendo che ogni *funzione potenza* $f(x) = x^{2n}$ è tutta *convessa* e *genera* i due punti $(\pm 1, 1)$.

Pertanto la *forma* del grafico di tutte le *funzioni potenza* $f(x) = x^{2n}$ con esponente pari è quello riportato nella seguente fig. 4.8.

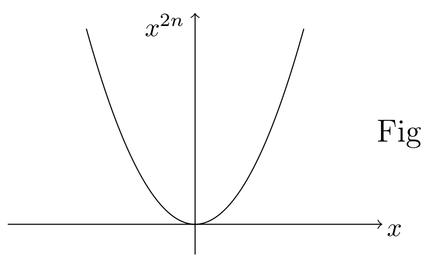


Fig. 4.8

Per analizzare il caso degli esponenti dispari, dimostriamo preliminarmente che la somma $1+z+z^2+z^3+\cdots+z^{2n-1}+z^{2n}$, in cui l'ultimo addendo abbia un qualsiasi esponente pari fissato, è sempre positiva per ogni numero reale z. Illustreremo solo il caso della somma fino all'addendo z^4 e il caso fino all'addendo z^6, in modo da "capire" quale sia la "regola ricorsiva". A tale scopo dimostriamo l'uguaglianza

$$1+z+z^2+z^3+\cdots+z^{n-1}+z^n=\frac{1-z^{n+1}}{1-z}\,, \tag{4.14}$$

per la quale consideriamo i prodotti di polinomi

$$\begin{aligned}&(1-z)(1+z)=1-z^2, \qquad (1-z)(1+z+z^2)=1-z^3,\\ &(1-z)(1+z+z^2+z^3)=1-z^4, \qquad (1-z)(1+z+z^2+z^3+z^4)=1-z^5\end{aligned}$$

e così via, da cui per induzione segue l'identità valida per ogni numero intero positivo n

$$(1-z)(1+z+z^2+z^3+\cdots+z^{n-1}+z^n)=1-z^{n+1},$$

la quale, per ogni $z\neq 1$, può essere riscritta nella forma

$$1+z+z^2+z^3+\cdots+z^{n-1}+z^n=\frac{1-z^{n+1}}{1-z}\,. \tag{4.15a}$$

Poiché per ogni z tale che $-1<z<1$ segue anche $-1<z^{n+1}<1$, possiamo ricavare dall'identità (4.15a) la proprietà per cui la somma $1+z+z^2+z^3+\cdots+z^{n-1}+z^n$, con $-1<z<1$, risulta sempre positiva per ogni n perché per valori di z tali che $-1<z<1$ si ha appunto

$$1+z+z^2+z^3+\cdots+z^{n-1}+z^n=\frac{1-z^{n+1}}{1-z}>0. \tag{4.15b}$$

Se a questo punto consideriamo la somma $1+z+z^2+z^3+z^4$ fino all'esponente pari 4 e la riscriviamo separando gli addendi con esponente pari e gli addendi con esponente dispari, otteniamo

$$\begin{aligned}1+z+z^2+z^3+z^4&=1+(z+z^3)+(z^2+z^4)=1+z(1+z^2)+z^2(1+z^2)=\\&=1+(1+z^2)(z^2+z),\end{aligned}$$

ovvero $1+z+z^2+z^3+z^4=1+(1+z^2)(z^2+z)$. Poiché il polinomio di secondo grado $1+z^2$ è sempre positivo ed il fattore z^2+z è sempre positivo "al di fuori" dell'intervallo $[-1,0]$, concludiamo che la somma $1+z+z^2+z^3+z^4$ è positiva per ogni valore reale di z perché per ogni z appartenente all'intervallo $(-1,0)$ si applica la disuguaglianza (4.15b).

Poiché per la somma $1+z+z^2+z^3+z^4+z^5+z^6$ fino all'ultimo addendo con l'esponente pari 6 risulta

$$\begin{aligned}1+z+z^2+z^3+z^4+z^5+z^6&=1+(z+z^3+z^5)+(z^2+z^4+z^6)=\\&=1+z(1+z^2+z^4)+z^2(1+z^2+z^4)=1+(1+z^2+z^4)(z^2+z),\end{aligned}$$

ci si rende conto facilmente che tutte le somme $1+z+z^2+z^3+\cdots+z^{2n-1}+z^{2n}$ fino a qualsiasi potenza con esponente pari risultano sempre positive per ogni valore reale di z.

Per studiare il *rapporto incrementale* delle *funzioni potenza* con esponente dispari, consideriamo i casi delle *funzioni* $f(x)=x^3$ e $f(x)=x^5$ per ricavare la "regola generale".

Poiché ogni *funzione* $f(x) = x^{2n+1}$ ha come *dominio* tutto l'insieme dei numeri reali e *genera* il punto $(0,0)$ in quanto $f(0) = 0$, possiamo studiare il *rapporto incrementale* solo per coppie di numeri distinti x_1, x_2 entrambi diversi da zero. Per la *funzione* $f(x) = x^3$ abbiamo, in virtù del *teorema di Ruffini*, che il *rapporto incrementale*

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{x_1^3 - x_2^3}{x_1 - x_2} = \frac{(x_1 - x_2)(x_1^2 + x_1x_2 + x_2^2)}{x_1 - x_2} = x_1^2 + x_1x_2 + x_2^2 = \left(x_1 + \frac{x_2}{2}\right)^2 + \frac{3x_2^2}{4}$$

risulta sempre positivo in quanto somma di addendi tutti positivi, da cui segue che la *funzione* $f(x) = x^3$ è *crescente* su tutto l'insieme dei numeri reali.

Per la *funzione* $f(x) = x^5$ abbiamo, in virtù del *teorema di Ruffini* e ponendo $x_2/x_1 = z$, che il *rapporto incrementale*

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{x_1^5 - x_2^5}{x_1 - x_2} = \frac{(x_1 - x_2)(x_1^4 + x_1^3x_2 + x_1^2x_2^2 + x_1x_2^3 + x_2^4)}{x_1 - x_2} =$$

$$= x_1^4 + x_1^3x_2 + x_1^2x_2^2 + x_1x_2^3 + x_2^4 = x_1^4\left[1 + \left(\frac{x_2}{x_1}\right) + \left(\frac{x_2}{x_1}\right)^2 + \left(\frac{x_2}{x_1}\right)^3 + \left(\frac{x_2}{x_1}\right)^4\right] =$$

$$= x_1^4\left[1 + z + z^2 + z^3 + z^4\right]$$

risulta sempre positivo in quanto somma di addendi tutti positivi, da cui segue che la *funzione* $f(x) = x^5$ è *crescente* su tutto l'insieme dei numeri reali e che, analogamente, tutte le *funzioni potenza* $f(x) = x^{2n-1}$ con esponenti dispari sono *crescenti* su tutto l'insieme dei numeri reali. Per studiare il *comportamento convesso* oppure *concavo* di tali *funzioni*, consideriamo la *funzione* $f(x) = x^3$ e dopo aver scritto l'equazione della retta ad essa secante nei suoi due generici punti $(x_1, x_1^4), (x_2, x_2^4)$, con x_1, x_2 distinti ed entrambi diversi da zero, secondo l'espressione (4.11), esaminiamo la differenza $y_0 - f(x_0)$ tra l'ordinata y_0 del punto D e l'ordinata $f(x_0) = x_0^3$ del punto C riportati nelle figg. 4.6.

Lo sviluppo di tale differenza e la relativa scomposizione dànno

$$\begin{aligned} y_0 - f(x_0) &= (x_1^2 + x_1x_2 + x_2^2)\,x_0 + x_1^3 - (x_1^2 + x_1x_2 + x_2^2)\,x_1 - x_0^3 = \\ &= x_1^2(x_0 - x_2) + x_1x_2(x_0 - x_2) - x_0(x_0 + x_2)(x_0 - x_2) = (x_0 - x_2)\,(x_1^2 - x_0^2 + x_1x_2 - x_0x_2) = \\ &= (x_0 - x_2)[(x_1 - x_0)(x_1 + x_0) + x_2(x_1 - x_0)] = (x_1 - x_0)(x_0 - x_2)(x_1 + x_0 + x_2) \end{aligned}$$

e poiché per $x_1 < x_0 < x_2$ il prodotto $(x_1 - x_0)(x_0 - x_2)$ risulta sempre positivo, segue che per $x_1 < x_0 < x_2 < 0$ il terzo fattore è negativo e la differenza $y_0 - f(x_0)$ risulta negativa, mentre per $0 < x_1 < x_0 < x_2$ il terzo fattore è positivo e la differenza $y_0 - f(x_0)$ risulta positiva. Pertanto possiamo concludere che la *funzione* $f(x) = x^3$ è *concava* sulla semiretta delle x negative, mentre è *convessa* sulla semiretta delle x positive.

Se anche per la *funzione* $f(x) = x^5$ scriviamo l'equazione della retta ad essa secante nei suoi due generici punti $(x_1, x_1^4), (x_2, x_2^4)$, con x_1, x_2 distinti ed entrambi diversi da zero, secondo l'espressione (4.11), e sviluppiamo la differenza $y_0 - f(x_0)$ tra l'ordinata y_0 del punto D e l'ordinata $f(x_0) = x_0^5$ del punto C riportati nelle figg. 4.6, otteniamo

$$\begin{aligned} y_0 - x_0^5 &= (x_1^4 + x_1^3x_2 + x_1^2x_2^2 + x_1x_2^3 + x_2^4)\,x_0 + x_1^4 - (x_1^4 + x_1^3x_2 + x_1^2x_2^2 + x_1x_2^3 + x_2^4)\,x_1 - x_0^5 = \\ &= (x_1 - x_0)(x_0 - x_2)\left[(x_0^2 + x_2^2)\,(x_1 + x_0 + x_2) + x_1(x_0 + x_1)(x_1 + x_2)\right] \end{aligned}$$

e poiché per $x_1 < x_0 < x_2$ il prodotto $(x_1 - x_0)(x_0 - x_2)$ risulta sempre positivo, segue che per $x_1 < x_0 < x_2 < 0$ il terzo fattore è negativo e la differenza $y_0 - f(x_0)$ risulta negativa, mentre per $0 < x_1 < x_0 < x_2$ il terzo fattore è positivo e la differenza $y_0 - f(x_0)$ risulta positiva. Pertanto possiamo concludere che la *funzione* $f(x) = x^5$ è *concava* sulla semiretta delle x negative, mentre è *convessa* sulla semiretta delle x positive.

Poiché dunque tutte le *funzioni potenza* $f(x) = x^{2n+1}$ con esponenti dispari sono *crescenti* su tutto l'insieme dei numeri reali, *concave* sulla semiretta delle x negative e *convesse* sulla semiretta delle x positive, il loro grafico qualitativo è quello rappresentato nella seguente fig. 4.9.

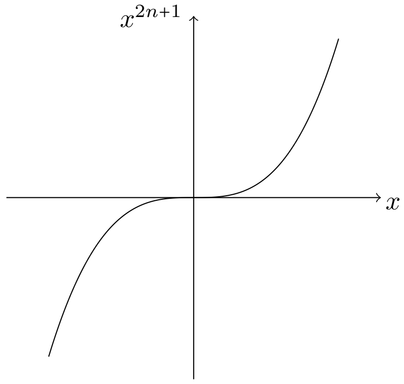


Fig. 4.9

4.2.3 La funzione potenza con esponente intero negativo

La *funzione potenza* con esponente intero negativo $k = -1, -2, -3, \ldots$ è la *funzione* $f(x) = 1/x^{-k}$, per la quale distinguiamo il caso degli esponenti pari, ovvero $f(x) = 1/x^2$, $f(x) = 1/x^4$ e cosi via, e il caso degli esponenti dispari $f(x) = 1/x$, $f(x) = 1/x^3$ e cosi via.

Riconoscendo che il *dominio* di tutte queste funzioni è l'insieme dei numeri reali diversi da zero, ovvero negativi e positivi, cominciamo con il caso degli esponenti pari e studiamo il *rapporto incrementale* della funzione $f(x) = 1/x^2$ in un intervallo contenuto nella semiretta delle x negative oppure nella semiretta delle x positive, ovvero per x_1, x_2 distinti e concordi in segno. Il *rapporto incrementale* della funzione $f(x) = 1/x^2$ è

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{\dfrac{1}{x_1^2} - \dfrac{1}{x_2^2}}{x_1 - x_2} = -\frac{x_1^2 - x_2^2}{x_1^2 x_2^2 (x_1 - x_2)} = -\frac{x_1 + x_2}{x_1^2 x_2^2}$$

e si riconosce che esso è positivo sulla semiretta delle x negative e negativo sulla semiretta delle x positive, da cui segue che la funzione $f(x) = 1/x^2$ è *crescente* sulle x negative e *decrescente* sulle x positive. Per stabilire dove la funzione abbia *comportamento convesso* oppure *comportamento concavo*, sudiamo la differenza $y_0 - f(x_0)$ tra le ordinate dei punti mostrati nelle figg. 4.6, ovvero

$$y_0 - f(x_0) = -\frac{x_1 + x_2}{x_1^2 x_2^2} x_0 + \frac{1}{x_1^2} + \frac{x_1 + x_2}{x_1^2 x_2^2} x_1 - \frac{1}{x_0^2} = \frac{-(x_1 + x_2)(x_0^3 - x_0^2 x_1) + x_0^2 x_2^2 - x_1^2 x_2^2}{x_1^2 x_0^2 x_2^2} =$$

$$= \frac{(x_0 - x_1)(x_2 - x_0)(x_0 x_1 + x_0 x_2 + x_1 x_2)}{x_1^2 x_0^2 x_2^2}$$

e si riconosce che tale differenza, sia per $x_1 < x_0 < x_2 < 0$ che per $0 < x_1 < x_0 < x_2$, è sempre positiva, da cui segue che la funzione $f(x) = 1/x^2$ è *convessa* su tutto il suo *dominio*.

Se studiamo quindi il *rapporto incrementale* della funzione $f(x) = 1/x^4$ in un intervallo contenuto nella semiretta delle x negative oppure nella semiretta delle x positive, ovvero per x_1, x_2 distinti e concordi in segno, otteniamo

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{\frac{1}{x_1^4} - \frac{1}{x_2^4}}{x_1 - x_2} = -\frac{x_1^4 - x_2^4}{x_1^2 x_2^2 (x_1 - x_2)} = -\frac{(x_1 - x_2)(x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_2^3)}{x_1^4 x_2^4 (x_1 - x_2)} =$$

$$= -\frac{x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_2^3}{x_1^4 x_2^4} = -\frac{(x_1 + x_2)(x_1^2 + x_2^2)}{x_1^4 x_2^4}$$

e si riconosce che esso è positivo sulla semiretta delle x negative e negativo sulla semiretta delle x positive, da cui segue che anche la funzione $f(x) = 1/x^4$ è *crescente* sulle x negative e *decrescente* sulle x positive. Per stabilire dove la funzione abbia *comportamento convesso* oppure *comportamento concavo*, sudiamo la differenza $y_0 - f(x_0)$ tra le ordinate dei punti mostrati nelle figg. 4.6, ovvero

$$y_0 - f(x_0) = -\frac{x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_2^3}{x_1^4 x_2^4} x_0 + \frac{1}{x_1^4} + \frac{x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_2^3}{x_1^4 x_2^4} x_1 - \frac{1}{x_0^4} =$$

$$= \frac{(x_0 - x_1)(x_2 - x_0)\,\mathcal{P}_6(x_0, x_1, x_2)}{x_1^4 x_0^4 x_2^4},$$

dove $\mathcal{P}_6(x_0, x_1, x_2)$ indica il polinomio di sesto grado nelle tre variabili x_0, x_1, x_2

$$\mathcal{P}_6(x_0, x_1, x_2) = x_0^3 x_1^3 + x_0^3 x_1^2 x_2 + x_0^2 x_1^3 x_2 + x_0^3 x_2^3 + x_0^2 x_1 x_2^3 + x_0 x_1^2 x_2^3 + x_0^3 x_1 x_2^2 + x_0^2 x_1^2 x_2^2 + x_0 x_1^3 x_2^2,$$

e si riconosce che tale differenza, sia per $x_1 < x_0 < x_2 < 0$ che per $0 < x_1 < x_0 < x_2$, è sempre positiva, da cui segue che la funzione $f(x) = 1/x^4$ è *convessa* su tutto il suo *dominio*.

Concludiamo quindi che tutte le funzioni $f(x) = 1/x^{2n}$ hanno il grafico qualitativo riportato in fig. 4.10(a).

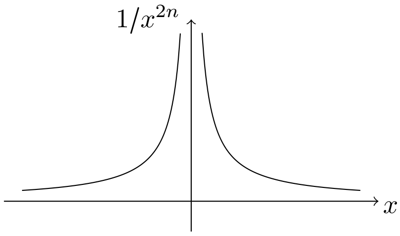


Fig. 4.10(a)

Per studiare il caso degli esponenti dispari, studiamo, in particolare, dove sono *crescenti, decrescenti, convesse, concave* le due funzioni $f(x) = 1/x$ e $f(x) = 1/x^3$, considerando un intervallo contenuto nella semiretta delle x negative oppure nella semiretta delle x positive, ovvero per x_1, x_2 distinti e concordi in segno.

Il *rapporto incrementale* della funzione $f(x) = 1/x$ è

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{\frac{1}{x_1} - \frac{1}{x_2}}{x_1 - x_2} = -\frac{x_1 - x_2}{x_1 x_2 (x_1 - x_2)} = -\frac{1}{x_1 x_2}$$

e si riconosce che esso è sempre negativo sia sulla semiretta delle x negative che sulla semiretta delle x positive, da cui segue che la funzione $f(x) = 1/x$ è *crescente* sia sulle x

negative che sulle x positive. Per stabilire dove la funzione abbia *comportamento convesso* oppure *comportamento concavo*, sudiamo la differenza $y_0 - f(x_0)$ tra le ordinate dei punti mostrati nelle figg. 4.6, ovvero

$$y_0 - f(x_0) = -\frac{1}{x_1 x_2}\, x_0 + \frac{1}{x_1} + \frac{1}{x_1 x_2}\, x_1 - \frac{1}{x_0} = \frac{x_0 x_2 - x_0^2 + x_0 x_1 - x_1 x_2}{x_1 x_0 x_2} = \frac{(x_0 - x_1)(x_2 - x_0)}{x_1 x_0 x_2}$$

e si riconosce che tale differenza risulta negativa per ogni tripla $x_1 < x_0 < x_2 < 0$ e risulta positiva per ogni tripla $0 < x_1 < x_0 < x_2$, da cui segue che la funzione $f(x) = 1/x$ è *concava* sulla semiretta delle x negative e *convessa* sulla semiretta delle x positive.

Il *rapporto incrementale* della funzione $f(x) = 1/x^3$ è

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{\dfrac{1}{x_1^3} - \dfrac{1}{x_2^3}}{x_1 - x_2} = -\frac{(x_1 - x_2)(x_1^2 + x_1 x_2 + x_2^2)}{x_1^3 x_2^3\,(x_1 - x_2)} = -\frac{x_1^2 + x_1 x_2 + x_2^2}{x_1^3 x_2^3}$$

e si riconosce che esso è sempre negativo sia sulla semiretta delle x negative che sulla semiretta delle x positive, da cui segue che la funzione $f(x) = 1/x$ è *crescente* sia sulle x negative che sulle x positive. Per stabilire dove la funzione abbia *comportamento convesso* oppure *comportamento concavo*, sudiamo la differenza $y_0 - f(x_0)$ tra le ordinate dei punti mostrati nelle figg. 4.6, ovvero

$$y_0 - f(x_0) = -\frac{x_1^2 + x_1 x_2 + x_2^2}{x_1^3 x_2^3}\, x_0 + \frac{1}{x_1^3} + \frac{x_1^2 + x_1 x_2 + x_2^2}{x_1^3 x_2^3}\, x_1 - \frac{1}{x_0^3} =$$
$$= \frac{(x_0 - x_1)(x_2 - x_0)\left(x_0^2 x_2^2 + x_0 x_1 x_2^2 + x_0^2 x_1 x_2 + x_1^2 x_2^2 + x_1^3 x_2 + x_1^2 x_0^2\right)}{x_1^3 x_0^3 x_2^3}$$

e si riconosce che tale differenza risulta negativa per ogni tripla $x_1 < x_0 < x_2 < 0$ e risulta positiva per ogni tripla $0 < x_1 < x_0 < x_2$, da cui segue che anche la funzione $f(x) = 1/x^3$ è *concava* sulla semiretta delle x negative e *convessa* sulla semiretta delle x positive.

Concludiamo che tutte le *funzioni potenza* con esponente intero negativo dispari *generano* i due punti $(-1,-1), (1,1)$, risultando sempre *decrescenti* in tutto il loro *dominio*, *concave* sulla semiretta delle x negative e *convesse* sulla semiretta delle x positive, da cui segue che il loro grafico qualitativo è quello riportato nella seguente fig. 4.10(b).

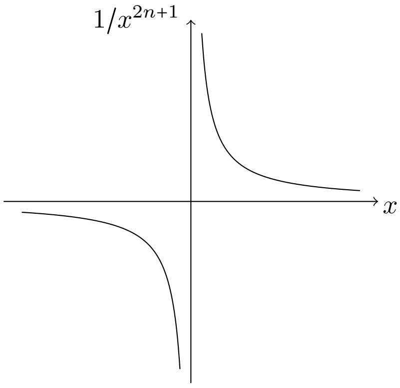


Fig. 4.10(b)

4.2.4 La funzione esponenziale e la funzione logaritmo

La *funzione esponenziale* ha espressione $f(x) = a^x$ e distinguiamo il caso di base a tale che $0 < a < 1$ e il caso di base a tale che $a > 1$. Utilizzando solo le proprietà delle potenze senza studiare il *comportamento concavo* oppure *convesso*, abbiamo che la *funzione esponenziale* passa sempre per il punto $(0,1)$ risultando sempre positiva, tutta *decrescente convessa* per $0 < a < 1$ e tutta *crescente convessa* per $a > 1$. I grafici qualitativi della *funzione esponenziale* nei due casi sono pertanto quelli riportati nelle seguenti figg. 4.11.

La *funzione logaritmo* ha espressione $f(x) = \log_a x$ e distinguiamo il caso di base a tale che $0 < a < 1$ e il caso di base a tale che $a > 1$. Utilizzando solo le proprietà delle potenze senza studiare il *comportamento concavo* oppure *convesso*, abbiamo che la *funzione logaritmo* passa sempre per il punto $(1,0)$ risultando negativa per $0 < x < 1$, positiva per $x > 1$, tutta *decrescente convessa* per $0 < a < 1$ e tutta *crescente concava* per $a > 1$.

I grafici qualitativi della *funzione logaritmo* nei due casi sono pertanto quelli riportati nelle seguenti figg. 4.12.

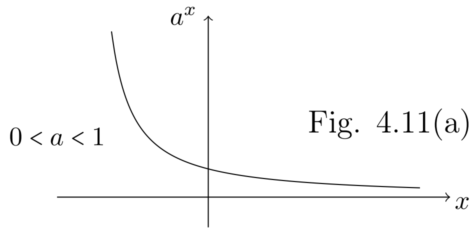


Fig. 4.11(a)

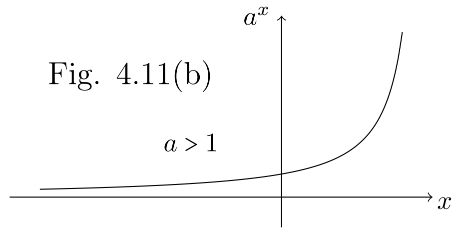


Fig. 4.11(b)

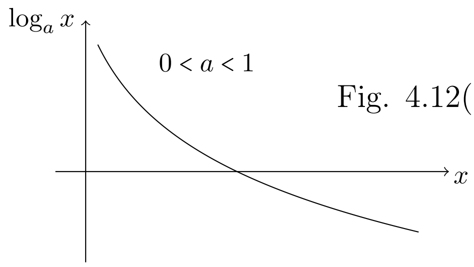


Fig. 4.12(a)

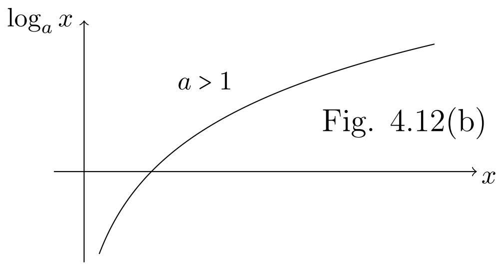


Fig. 4.12(b)

Abbiamo visto che la determinazione del *comportamento crescente, decrescente, concavo, convesso* di una *funzione* mediante lo studio delle disuguaglianze (4.10) e (4.12) risulta "molto difficoltoso" e, in particolare per le *funzioni potenza*, la "difficoltà" dello studio delle disuguaglianze (4.10) e (4.12) "aumenta" all'aumentare degli esponenti.

Inoltre abbiamo che "in pratica" per *funzioni* di "tipo diverso" rispetto a quelle che abbiamo analizzato, le disuguaglianze (4.10) e (4.12) "non forniscono più" nessun risultato.

Per studiare dunque tutte le altre *funzioni* "diverse" da quelle di questo paragrafo, occorre introdurre nuovi concetti matematici e più precisamente gli strumenti del Calcolo Infinitesimale e del Calcolo Differenziale che fanno parte di quella branca della Matematica denominata *Analisi Matematica.* L'*Analisi Matematica* venne introdotta da Newton verso la fine degli anni Ottanta del Seicento per formulare la "famosa" legge del moto in Fisica secondo cui la *forza* esercitata su un *corpo* è proporzionale all'*accelerazione* che il *corpo* stesso subisce e Leibnitz si rese conto che questa teoria di Newton permetteva anche di studiare le disuguaglianze (4.10) e (4.12), ovvero di determinare appunto il *comportamento crescente, decrescente, concavo, convesso* di una "qualsiasi" *funzione.*

4.3 Concetto di limite

Per illustrare il concetto di *limite*, introduciamo delle *proposizioni* ognuna delle quali esprima una certa proprietà delle *funzioni* e attribuiamo dei *nomi* alle *funzioni* che soddisfano tali *proposizioni*, nelle quali i simboli x_0, ℓ, ove vengano utilizzati, indicano sempre due numeri reali fissati e i parametri $\varepsilon, \mathcal{M}, \delta, \mathcal{N}$ sono numeri reali positivi.

4.3.1 Convergenza al finito

Per il concetto di *convergenza al finito* distingueremo, dato un punto x_0, tre casi corrispondenti ai tre tipi di insiemi delle x rispettivamente

- insieme $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$, denominato *intorno completo del punto* x_0;
- insieme $(x_0 - \delta, x_0)$, denominato *intorno sinistro del punto* x_0;
- insieme $(x_0, x_0 + \delta)$, denominato *intorno destro del punto* x_0.

E' importante puntualizzare ora che in realtà l'insieme delle x denominato *intorno completo del punto* x_0 è un intervallo comprendente anche il punto x_0 stesso, ovvero $(x_0 - \delta_\varepsilon, x_0 + \delta_\varepsilon)$, ma poiché, come vedremo, per il concetto di *tendenza della* x *ad* x_0, non occorre attribuire il valore x_0 alla variabile x di una funzione, segue che, per illustrare il concetto di *limite per* x *tendente ad* x_0, ovvero di *limite per* x *tendente al finito*, possiamo escludere il punto x_0 dall'*intorno completo del punto* x_0 stesso.

PRIMO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_1$: "*$f(x)$ è una funzione tale che per ogni numero reale ε positivo, troviamo un intervallo avente la forma $(x_0 - \delta_\varepsilon, x_0) \cup (x_0, x_0 + \delta_\varepsilon)$, tutti gli elementi del quale soddisfano la disequazione $\ell - \varepsilon < f(x) < \ell + \varepsilon$*",

possiamo enunciare la seguente definizione.

Definizione 4.3.1. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_1$, si dice che*

la funzione $f(x)$ tende al "limite" ℓ per "x tendente a x_0 sia da sinistra che da destra"

oppure anche

il "limite" della funzione $f(x)$ per "x tendente a x_0 sia da sinistra che da destra" vale ℓ

e si scrive

$$\lim_{x \to x_0} f(x) = \ell.$$

E' di fondamentale importanza sottolineare che l'insieme completo delle soluzioni della disequazione $\ell - \varepsilon < f(x) < \ell + \varepsilon$ non ha "grande rilevanza" ai fini del *limite* ℓ della funzione $f(x)$ per x tendente a x_0. La proprietà cruciale per il *limite* è solo che l'insieme di tutte le soluzioni della disequazione $\ell - \varepsilon < f(x) < \ell + \varepsilon$ possieda sempre, per ogni valore del parametro $\varepsilon > 0$, un sottoinsieme avente la forma $(x_0 - \delta_\varepsilon, x_0) \cup (x_0, x_0 + \delta_\varepsilon)$, indipendentemente se la disequazione abbia altre soluzioni oppure no e quali siano.

Esempio 1. Dimostriamo che la funzione $f(x) = 3x - 5$ soddisfa la *proposizione* $\mathcal{P}_1$ con $x_0 = 4$ e $\ell = 7$, ovvero verifichiamo che l'insieme delle soluzioni di ogni disequazione

$$7 - \varepsilon < 3x - 5 < 7 + \varepsilon \tag{4.16a}$$

possiede sempre un sottoinsieme della forma $(4-\delta_\varepsilon,4)\cup(4,4+\delta_\varepsilon)$. Sviluppiamo quindi

$$7-\varepsilon<3x-5<7+\varepsilon \quad\Longrightarrow\quad 12-\varepsilon<3x<12+\varepsilon \quad\Longrightarrow\quad 4-\frac{\varepsilon}{3}<x<4+\frac{\varepsilon}{3}$$

e osserviamo che per ogni ε positivo, esiste sempre un intervallo avente la forma

$$(4-\delta_\varepsilon,4)\cup(4,4+\delta_\varepsilon), \tag{4.16b}$$

con $\delta_\varepsilon=\varepsilon/3$, il quale, essendo un sottoinsieme di tutto l'insieme $(4-\delta_\varepsilon,4+\delta_\varepsilon)$ delle soluzioni della disequazione $7-\varepsilon<f(x)<7+\varepsilon$, è costituito anch'esso, in quanto sottoinsieme, di elementi che sono tutti soluzione della disequazione $7-\varepsilon<f(x)<7+\varepsilon$.

Poiché dunque l'insieme di tutte le soluzioni di ogni disequazione (4.16a) possiede sempre un sottoinsieme della forma (4.16b), concludiamo che la funzione $f(x)=3x-5$ rende vera la *proposizione* $\mathcal{P}_1$, con $x_0=4$ e $\ell=7$, e che tale funzione pertanto appartiene all'*insieme di validità* della *proposizione* $\mathcal{P}_1$, in modo che valga quindi

$$\lim_{x\to 4}(3x-5)=7.$$

Esempio 2. Dimostriamo che la funzione

$$f(x)=\frac{2x+9}{3x+7}$$

soddisfa la *proposizione* $\mathcal{P}_1$ con $x_0=-2$ e $\ell=5$, ovvero verifichiamo che l'insieme delle soluzioni di ogni disequazione

$$5-\varepsilon<\frac{2x+9}{3x+7}<5+\varepsilon \tag{4.17a}$$

possiede sempre un sottoinsieme della forma $(-2-\delta_\varepsilon,-2)\cup(-2,-2+\delta_\varepsilon)$.

In virtù dell'identità

$$\frac{2x+9}{3x+7}=\frac{2}{3}+\frac{13}{9x+21},$$

possiamo riscrivere la disequazione (4.17a) nella forma (con $0<\varepsilon<5$ e $x>0$)

$$5-\varepsilon<\frac{2}{3}+\frac{13}{9x+21}<5+\varepsilon,$$

da cui, se svolgiamo i passaggi utilizzando le proprietà delle disuguaglianze, segue

$$5-\varepsilon<\frac{2}{3}+\frac{13}{9x+21}<5+\varepsilon \quad\Longrightarrow\quad \frac{13}{3}-\varepsilon<\frac{13}{9x+21}<\frac{13}{3}+\varepsilon \quad\Longrightarrow$$

$$\Longrightarrow\quad 13-3\varepsilon<\frac{13}{3x+7}<13+3\varepsilon \quad\Longrightarrow\quad \frac{13-3\varepsilon}{13}<\frac{1}{3x+7}<\frac{13+3\varepsilon}{13} \quad\Longrightarrow$$

$$\Longrightarrow\quad \frac{13}{13+3\varepsilon}<3x+7<\frac{13}{13-3\varepsilon} \quad\Longrightarrow\quad \frac{-78-21\varepsilon}{13+3\varepsilon}<3x<\frac{-78+21\varepsilon}{13-3\varepsilon} \quad\Longrightarrow$$

$$\Longrightarrow\quad \frac{-26-7\varepsilon}{13+3\varepsilon}<x<\frac{-26+7\varepsilon}{13-3\varepsilon} \quad\Longrightarrow\quad -2-\frac{\varepsilon}{13+3\varepsilon}<x<-2+\frac{\varepsilon}{13-3\varepsilon} \quad\Longrightarrow$$

$$\Longrightarrow\quad -2-\frac{\varepsilon}{13+3\varepsilon}<x<-2+\frac{\varepsilon}{13+3\varepsilon},$$

dove abbiamo sostituito le identità

$$\frac{-26-7\varepsilon}{13+3\varepsilon}=-2-\frac{\varepsilon}{13+3\varepsilon}\,,\qquad\qquad \frac{-26+7\varepsilon}{13-3\varepsilon}=-2+\frac{\varepsilon}{13-3\varepsilon}$$

e abbiamo quindi sostituito

$$-2+\frac{\varepsilon}{13+3\varepsilon}\qquad\text{al posto di}\qquad -2+\frac{\varepsilon}{13-3\varepsilon}$$

perché risulta

$$-2+\frac{\varepsilon}{13+3\varepsilon}<-2+\frac{\varepsilon}{13-3\varepsilon}\,,$$

ovvero abbiamo scritto alla fine l'intervallo "più stretto"

$$-2-\left(\frac{\varepsilon}{13+3\varepsilon}\right)<x<-2+\left(\frac{\varepsilon}{13+3\varepsilon}\right) \tag{4.17b}$$

al posto dell'intervallo "più largo", comprendente tutte le soluzioni

$$-2-\frac{\varepsilon}{13+3\varepsilon}<x<-2+\frac{\varepsilon}{13-3\varepsilon}\,,$$

in modo che l'intervallo finale (4.17b) sia simmetrico rispetto ad $x_0=-2$ con il medesimo

$$\delta_\varepsilon=\frac{\varepsilon}{13+3\varepsilon}$$

in parentesi tonda nell'estremo sinistro e nell'estremo destro dell'intervallo.

Osserviamo quindi che tutte le x dell'intervallo (4.17b), ovvero tutte le x che appartengono all'intervallo $x\in(-2-\delta,-2+\delta)$ non sono in realtà tutte le soluzioni

$$-2-\frac{\varepsilon}{13+3\varepsilon}<x<-2+\frac{\varepsilon}{13-3\varepsilon}$$

della disequazione (4.17a) perché le soluzioni

$$-2+\frac{\varepsilon}{13+3\varepsilon}<x<-2+\frac{\varepsilon}{13-3\varepsilon}$$

sono state trascurate quando abbiamo scritto l'intervallo "più stretto" (4.17b).

Ma sebbene "alcune" soluzioni x siano state trascurate, tuttavia la funzione

$$f(x)=\frac{2x+9}{3x+7}$$

è tale che per ogni $\varepsilon>0$ vi sia sempre un intervallo di tipo

$$(-2-\delta_\varepsilon,-2)\cup(-2,-2+\delta_\varepsilon),$$

riportato nella (4.17b) anche senza il valore $x_0=-2$, tutti gli elementi del quale soddisfano la disequazione $5-\varepsilon<f(x)<5+\varepsilon$ e le soluzioni x della disequazione appartenenti all'intervallo trascurato

$$-2+\frac{\varepsilon}{13+3\varepsilon}<x<-2+\frac{\varepsilon}{13-3\varepsilon}$$

non hanno "nessuna rilevanza". In virtù di tale verifica, poiché abbiamo determinato l'intervallo che dovevamo determinare, possiamo concludere quindi che la funzione

$$f(x) = \frac{2x+9}{3x+7}$$

rende vera la *proposizione* $\mathcal{P}_1$, con $x_0 = -2$ e $\ell = 5$, e che tale funzione pertanto appartiene all'*insieme di validità* della *proposizione* $\mathcal{P}_1$, in modo che valga dunque

$$\lim_{x\to-2} \frac{2x+9}{3x+7} = 5.$$

Esempio 3. Dimostriamo che la funzione

$$f(x) = \frac{x-5}{x-5} \tag{4.18a}$$

soddisfa la *proposizione* $\mathcal{P}_1$ con $x_0 = 5$ e $\ell = 1$. Poiché, com'è immediato rendersi conto, le soluzioni di ogni disequazione

$$1 - \varepsilon < \frac{x-5}{x-5} < 1 + \varepsilon \tag{4.18b}$$

sono tutte le x appartenenti all'insieme

$$(-\infty, 5) \cup (5, +\infty), \tag{4.18c}$$

avente ovviamente il sottoinsieme $(5 - \delta_\varepsilon, 5) \cup (5, 5 + \delta_\varepsilon)$, possiamo quindi concludere che la funzione (4.18a) rende vera la *proposizione* $\mathcal{P}_1$ con $x_0 = 5$ e $\ell = 1$. Possiamo inoltre osservare che se nella funzione (4.18a) consideriamo il generico binomio $x - x_0$ al posto del binomio $x - 5$, segue che le soluzioni di ogni disequazione (4.18b) in cui si sostituisca x_0 al posto del numero 5 sono sempre tutte le x appartenenti all'insieme (4.18c) in cui si sostituisca x_0 al posto del numero 5. Abbiamo pertanto il *limite* generale

$$\lim_{x\to x_0} \frac{x-x_0}{x-x_0} = 1 \tag{4.19}$$

per ogni valore di x_0, che possiamo considerare come il *limite notevole* per risolvere tutti i *limiti indeterminati* di rapporti di polinomi. Se indichiamo la funzione $(x - x_0)/(x - x_0)$ ponendo

$$f(x) = \frac{x-x_0}{x-x_0},$$

emerge l'importantissima proprietà per cui può esistere il *limite* di una funzione, ovvero

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} \frac{x-x_0}{x-x_0} = 1,$$

anche se non esiste il risultato $f(x_0)$ della sostituzione di x_0 alla variabile x della medesima funzione. Abbiamo in effetti che $f(x_0)$ non esiste perché x_0 non appartiene al *dominio* della funzione $f(x)$. Possiamo dire quindi che il risultato $f(x_0)$ della sostituzione di x_0 alla variabile x di una funzione esiste solo se x_0 appartiene al *dominio* della funzione, mentre affinché possa esistere il *limite* della $f(x)$ per x tendente ad x_0, basta che il valore x_0 sia quello che viene denominato *punto di accumulazione* del *dominio* della $f(x)$.

In effetti, sebbene x_0 non appartenga *dominio* $(-\infty, x_0) \cup (x_0, +\infty)$ della funzione

$$f(x) = \frac{x - x_0}{x - x_0},$$

tuttavia x_0 è in ogni caso un *punto di accumulazione* del suo *dominio.*

Come esempio di applicazione del *limite notevole* (4.19), risolviamo ad esempio il seguente *limite*

$$\lim_{x \to -4} \frac{x^3 + 64}{2x^3 + 7x^2 - 5x - 4}.$$

Poiché il polinomio al numeratore e il polinomio al denominatore si annullano entrambi per il valore $x = -4$, segue che nella loro scomposizione è presente, in virtù del *teorema di Ruffini*, il binomio $x + 4$, che si annulla appunto per il valore $x = -4$.

Pertanto scomponendo i due polinomi otteniamo

$$\lim_{x \to -4} \frac{x^3 + 64}{2x^3 + 7x^2 - 5x - 4} = \lim_{x \to -4} \frac{(x + 4)(x^2 - 4x + 16)}{(x + 4)(2x^2 - x - 1)} =$$

$$= \lim_{x \to -4} \frac{x + 4}{x + 4} \cdot \lim_{x \to -4} \frac{x^2 - 4x + 16}{2x^2 - x - 1} = 1 \cdot \frac{48}{35} = \frac{48}{35}$$

perché il primo *limite* coincide con il *limite* (4.19) e il secondo dà 48/35.

Abbiamo ottenuto quindi che la funzione

$$f(x) = \frac{x^3 + 64}{2x^3 + 7x^2 - 5x - 4}$$

possiede il *limite* per x tendente al valore -4, ma $f(-4)$ non esiste perché $x = -4$ non appartiene al dominio della $f(x)$. Quando vale dunque la *convergenza al finito*

$$\lim_{x \to x_0} f(x) = \ell \in \mathbb{R},$$

possiamo distinguere il caso in cui esista anche $f(x_0)$ dal caso in cui $f(x_0)$ non esista introducendo il concetto di *funzione continua* mediante la seguente definizione.

Definizione 4.3.2. *(Definizione di "continuità") Se per una funzione f(x) in un punto x_0 risulta*

$$\lim_{x \to x_0} f(x) = \ell$$

e il valore ℓ coincide con il risultato $f(x_0)$ della sostituzione del valore x_0 alla variabile x, allora si dice che "la funzione $f(x)$ è continua nel punto x_0".

SECONDO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_1^{(a)}$: "*$f(x)$ è una funzione tale che per ogni numero reale ε positivo, troviamo un intervallo $x_0 - \delta_\varepsilon < x < x_0$, tutti gli elementi del quale soddisfano la disequazione* $\ell - \varepsilon < f(x) < \ell + \varepsilon$",

possiamo enunciare la seguente definizione.

Definizione 4.3.3. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_1^{(a)}$, si dice che*

la funzione $f(x)$ tende al "limite" ℓ per "x tendente a x_0 da sinistra"

oppure anche

il "limite" della funzione $f(x)$ per "x tendente a x_0 da sinistra" vale ℓ

e si scrive

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

TERZO CASO. Se consideriamo quindi la seguente *proposizione*

$\mathcal{P}_1^{(b)}$: "*$f(x)$ è una funzione tale che per ogni numero reale ε positivo, troviamo un intervallo $x_0 < x < x_0 + \delta_\varepsilon$, tutti gli elementi del quale soddisfano la disequazione $\ell - \varepsilon < f(x) < \ell + \varepsilon$*",

possiamo enunciare la seguente definizione.

Definizione 4.3.4. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_1^{(b)}$, si dice che*

la funzione $f(x)$ tende al "limite" ℓ per "x tendente a x_0 da destra"

oppure anche

il "limite" della funzione $f(x)$ per "x tendente a x_0 da destra" vale ℓ

e si scrive

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

Esempio 4. Dimostriamo che la funzione

$$f(x) = |x-2| + 3\,\frac{|x-2|}{x-2}$$

soddisfa per $x_0 = 2$ la *proposizione* $\mathcal{P}_1^{(a)}$ con $\ell = -3$ e la *proposizione* $\mathcal{P}_1^{(b)}$ con $\ell = 3$.

Ricordando che il modulo $|x-2|$ equivale a scrivere

$$|x-2| = \begin{cases} 2-x & \text{se} \quad x < 2 \\ x-2 & \text{se} \quad x > 2, \end{cases}$$

riscriviamo la funzione nella forma

$$f(x) = |x-2| + 3\,\frac{|x-2|}{x-2} = \begin{cases} -x-1 & \text{se} \quad x < 2 \\ x+1 & \text{se} \quad x > 2, \end{cases}$$

da cui segue che i risultati dei valori $x < 2$ si ottengono solo dall'espressione $-x-1$ e i risultati dei valori $x > 2$ si ottengono solo dall'espressione $x+1$.

A questo punto la funzione rende vera la *proposizione* $\mathcal{P}_1^{(a)}$ con $x_0 = 2$ e $\ell = -3$ se per ogni disequazione $-3 - \varepsilon < f(x) < -3 + \varepsilon$ troviamo un intervallo di soluzioni avente la forma $(2-\delta_\varepsilon, 2)$ e rende vera la *proposizione* $\mathcal{P}_1^{(b)}$ con $x_0 = 2$ e $\ell = 3$ se per ogni disequazione $3-\varepsilon < f(x) < 3+\varepsilon$ troviamo un intervallo di soluzioni della forma $(2, 2+\delta_\varepsilon)$.

Poiché l'insieme delle soluzioni di tutte le disequazioni

$$-3-\varepsilon < -x-1 < -3+\varepsilon$$

si ottengono risolvendo

$$-3-\varepsilon < -x-1 < -3+\varepsilon \quad\Longrightarrow\quad -2-\varepsilon < -x < -2+\varepsilon \quad\Longrightarrow\quad 2-\varepsilon < x < 2,$$

ovvero per ogni ε positivo, esiste sempre un intervallo avente la forma

$$(2-\delta_\varepsilon, 2),$$

con $\delta_\varepsilon = \varepsilon$, concludiamo che la $f(x)$ rende vera la *proposizione* $\mathcal{P}_1^{(a)}$ con $x_0 = 2$ e $\ell = -3$ e scriviamo

$$\lim_{x\to 2^-} |x-2| + 3\,\frac{|x-2|}{x-2} = -3.$$

Analogamente, poiché l'insieme delle soluzioni di tutte le disequazioni

$$3-\varepsilon < x+1 < 3+\varepsilon$$

si ottengono risolvendo

$$3-\varepsilon < x+1 < 3+\varepsilon \quad\Longrightarrow\quad 2 < x < 2+\varepsilon,$$

ovvero per ogni ε positivo, esiste sempre un intervallo avente la forma

$$(2, 2+\delta_\varepsilon),$$

con $\delta_\varepsilon = \varepsilon$, concludiamo che la $f(x)$ rende vera la *proposizione* $\mathcal{P}_1^{(b)}$ con $x_0 = 2$ e $\ell = 3$ e scriviamo

$$\lim_{x\to 2^+} |x-2| + 3\,\frac{|x-2|}{x-2} = 3.$$

Osserviamo che se una *funzione* $f(x)$ soddisfa la *proposizione* $\mathcal{P}_1$, segue che il *limite* della $f(x)$ vale ℓ per x tendente a x_0 sia da sinistra che da destra, ovvero vale l'uguaglianza

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

4.3.2 Convergenza all'infinito

PRIMO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_2^{(a)}$: "*$f(x)$ è una funzione tale che per ogni numero reale ε positivo, troviamo una semiretta avente la forma $(-\infty, -\mathcal{N}_\varepsilon)$, tutti gli elementi della quale soddisfano la disequazione $\ell-\varepsilon < f(x) < \ell+\varepsilon$*",

possiamo enunciare la seguente definizione.

Definizione 4.3.5. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_2^{(a)}$, si dice che*

$$\textit{la funzione } f(x) \textit{ tende al "limite" } \ell \textit{ per "} x \textit{ tendente a } -\infty''$$

oppure anche

$$\textit{il "limite" della funzione } f(x) \textit{ per "} x \textit{ tendente a } -\infty'' \textit{ vale } \ell$$

e si scrive

$$\lim_{x\to-\infty} f(x) = \ell.$$

SECONDO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_2^{(b)}$: "*$f(x)$ è una funzione tale che per ogni numero reale ε positivo, troviamo una semiretta avente la forma $(\mathcal{N}_\varepsilon, \infty)$, tutti gli elementi della quale soddisfano la disequazione $\ell - \varepsilon < f(x) < \ell + \varepsilon''$*,

possiamo enunciare la seguente definizione.

Definizione 4.3.6. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_2^{(b)}$, si dice che*

$$\textit{la funzione } f(x) \textit{ tende al "limite" } \ell \textit{ per "} x \textit{ tendente a } +\infty''$$

oppure anche

$$\textit{il "limite" della funzione } f(x) \textit{ per "} x \textit{ tendente a } +\infty'' \textit{ vale } \ell$$

e si scrive

$$\lim_{x\to+\infty} f(x) = \ell.$$

Anche in questo caso è "molto importante" sottolineare che l'insieme completo delle soluzioni della disequazione $\ell-\varepsilon < f(x) < \ell+\varepsilon$ non ha "grande rilevanza" ai fini del *limite* ℓ della funzione $f(x)$ per x tendente a x_0. La proprietà cruciale per il *limite* è solo che l'insieme di tutte le soluzioni della disequazione $\ell - \varepsilon < f(x) < \ell + \varepsilon$ possieda sempre, per ogni valore del parametro $\varepsilon > 0$, un sottoinsieme avente la forma di semiretta $(-\infty, -\mathcal{N}_\varepsilon)$, per $x \to -\infty$, oppure $(\mathcal{N}_\varepsilon, \infty)$, per $x \to +\infty$, indipendentemente se la disequazione abbia altre soluzioni oppure no e quali siano.

4.3.3 Divergenza al finito

PRIMO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_3^{(a)}$: "*$f(x)$ è una funzione tale che per ogni numero reale $\mathcal{M}$ positivo, troviamo un intervallo $x_0 - \delta_\varepsilon < x < x_0$, tutti gli elementi del quale soddisfano la disequazione $f(x) > \mathcal{M}''$*,

possiamo enunciare la seguente definizione.

Definizione 4.3.7. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_3^{(a)}$, si dice che*

$$la\ funzione\ f(x)\ tende\ a\ +\infty\ per\ \text{“}x\ tendente\ a\ x_0\ da\ sinistra\text{”}$$

oppure anche

$$il\ \text{“}limite\text{”}\ della\ funzione\ f(x)\ per\ \text{“}x\ tendente\ a\ x_0\ da\ sinistra\text{”}\ \grave{e}\ +\infty$$

e si scrive

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

SECONDO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_3^{(b)}$: “*$f(x)$ è una funzione tale che per ogni numero reale $\mathcal{M}$ positivo, troviamo un intervallo $x_0 < x < x_0 + \delta_\varepsilon$, tutti gli elementi del quale soddisfano la disequazione $f(x) > \mathcal{M}$*”,

possiamo enunciare la seguente definizione.

Definizione 4.3.8. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_3^{(b)}$, si dice che*

$$la\ funzione\ f(x)\ tende\ a\ +\infty\ per\ \text{“}x\ tendente\ a\ x_0\ da\ destra\text{”}$$

oppure anche

$$il\ \text{“}limite\text{”}\ della\ funzione\ f(x)\ per\ \text{“}x\ tendente\ a\ x_0\ da\ destra\text{”}\ \grave{e}\ +\infty$$

e si scrive

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

TERZO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_3^{(c)}$: “*$f(x)$ è una funzione tale che per ogni numero reale $\mathcal{M}$ positivo, troviamo un intervallo $x_0 - \delta_\varepsilon < x < x_0$, tutti gli elementi del quale soddisfano la disequazione $f(x) < -\mathcal{M}$*”,

possiamo enunciare la seguente definizione.

Definizione 4.3.9. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_3^{(c)}$, si dice che*

$$la\ funzione\ f(x)\ tende\ a\ -\infty\ per\ \text{“}x\ tendente\ a\ x_0\ da\ sinistra\text{”}$$

oppure anche

$$il\ \text{“}limite\text{”}\ della\ funzione\ f(x)\ per\ \text{“}x\ tendente\ a\ x_0\ da\ sinistra\text{”}\ \grave{e}\ -\infty$$

e si scrive

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

QUARTO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_3^{(d)}$: "*$f(x)$ è una funzione tale che per ogni numero reale $\mathcal{M}$ positivo, troviamo un intervallo $x_0 < x < x_0 + \delta_\varepsilon$, tutti gli elementi del quale soddisfano la disequazione $f(x) < -\mathcal{M}$*",

possiamo enunciare la seguente definizione.

Definizione 4.3.10. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_3^{(d)}$, si dice che*

la funzione $f(x)$ tende a $-\infty$ per "x tendente a x_0 da destra"

oppure anche

il "limite" della funzione $f(x)$ per "x tendente a x_0 da destra" è $-\infty$

e si scrive

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

4.3.4 Divergenza all'infinito

PRIMO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_3^{(a)}$: "*$f(x)$ è una funzione tale che per ogni numero reale $\mathcal{M}$ positivo, troviamo una semiretta $(-\infty, -\mathcal{N}_\mathcal{M})$, tutti gli elementi della quale soddisfano la disequazione $f(x) > \mathcal{M}$*",

possiamo enunciare la seguente definizione.

Definizione 4.3.11. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_3^{(a)}$, si dice che*

la funzione $f(x)$ tende a $+\infty$ per "x tendente a $-\infty$"

oppure anche

il "limite" della funzione $f(x)$ per "x tendente a $-\infty$" vale $+\infty$

e si scrive

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

SECONDO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_3^{(b)}$: "*$f(x)$ è una funzione tale che per ogni numero reale $\mathcal{M}$ positivo, troviamo una semiretta $(\mathcal{N}_\mathcal{M}, +\infty)$, tutti gli elementi della quale soddisfano la disequazione $f(x) > \mathcal{M}$*",

possiamo enunciare la seguente definizione.

Definizione 4.3.12. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_3^{(b)}$, si dice che*

$$la\ funzione\ f(x)\ tende\ a\ +\infty\ per\ \text{“}x\ tendente\ a\ +\infty''$$

oppure anche

$$il\ \text{“}limite\text{”}\ della\ funzione\ f(x)\ per\ \text{“}x\ tendente\ a\ +\infty''\ vale\ +\infty$$

e si scrive

$$\lim_{x\to+\infty} f(x) = +\infty.$$

TERZO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_3^{(c)}$: *"$f(x)$ è una funzione tale che per ogni numero reale $\mathcal{M}$ positivo, troviamo una semiretta $(-\infty, \mathcal{N}_{\mathcal{M}})$, tutti gli elementi della quale soddisfano la disequazione $f(x) < -\mathcal{M}''$,*

possiamo enunciare la seguente definizione.

Definizione 4.3.13. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_3^{(c)}$, si dice che*

$$la\ funzione\ f(x)\ tende\ a\ -\infty\ per\ \text{“}x\ tendente\ a\ -\infty''$$

oppure anche

$$il\ \text{“}limite\text{”}\ della\ funzione\ f(x)\ per\ \text{“}x\ tendente\ a\ -\infty''\ vale\ -\infty$$

e si scrive

$$\lim_{x\to-\infty} f(x) = -\infty.$$

QUARTO CASO. Se consideriamo la seguente *proposizione*

$\mathcal{P}_3^{(d)}$: *"$f(x)$ è una funzione tale che per ogni numero reale $\mathcal{M}$ positivo, troviamo una semiretta $(\mathcal{N}_{\mathcal{M}}, +\infty)$, tutti gli elementi della quale soddisfano la disequazione $f(x) < -\mathcal{M}''$,*

possiamo enunciare la seguente definizione.

Definizione 4.3.14. *Se una funzione $f(x)$ soddisfa la proposizione $\mathcal{P}_3^{(d)}$, si dice che*

$$la\ funzione\ f(x)\ tende\ a\ -\infty\ per\ \text{“}x\ tendente\ a\ +\infty''$$

oppure anche

$$il\ \text{“}limite\text{”}\ della\ funzione\ f(x)\ per\ \text{“}x\ tendente\ a\ +\infty''\ vale\ -\infty$$

e si scrive

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

Capitolo 5

Successioni e serie numeriche

Prima di introdurre il concetto di successione, indichiamo con $\mathbb{N}$ l'insieme dei numeri interi positivi, chiamati anche *numeri naturali*, e con $\mathbb{R}$ l'insieme dei numeri reali.

Precisiamo però che nella teoria delle successioni, l'insieme dei numeri naturali in alcuni casi ha come primo elemento il numero 0, mentre in altri casi ha come primo elemento il numero 1. Tuttavia non sarà difficile rendersi conto che è del tutto irrilevante per la teoria considerare l'insieme dei numeri naturali con primo elemento 0 oppure 1.

5.1 Successioni numeriche

Definizione 5.1.1. *Una "successione numerica" (talvolta chiamata anche semplicemente successione) è un'applicazione che ad un numero naturale n associa, come risultato, un numero reale, indicato con il simbolo a_n, ovvero $a_n : \mathbb{N} \longrightarrow \mathbb{R}$.*

Esempio 1. Consideriamo la successione

$$a_n = \frac{n-1}{n}$$

essa associa ad $n = 1, 2, 3, \ldots$ i risultati

$$a_1 = 0, \qquad a_2 = \frac{1}{2}, \qquad a_3 = \frac{2}{3}, \qquad a_4 = \frac{3}{4}, \qquad \ldots$$

In particolare si ha che tale successione è monotòna crescente perché vale $a_{n+1} > a_n$, come si deduce eseguendo

$$a_{n+1} - a_n = \frac{n}{n+1} - \frac{n-1}{n} = \frac{1}{n(n+1)} > 0, \quad \forall\, n \in \mathbb{N}.$$

Esempio 2. Consideriamo la successione

$$a_n = \frac{n+1}{n}$$

essa associa ad $n = 1, 2, 3, \ldots$ i risultati

$$a_1 = 2, \qquad a_2 = \frac{3}{2}, \qquad a_3 = \frac{4}{3}, \qquad a_4 = \frac{5}{4}, \qquad \ldots$$

In particolare si ha che tale successione è monotòna decrescente perché vale $a_{n+1} < a_n$, come si deduce eseguendo

$$a_{n+1} - a_n = \frac{n+2}{n+1} - \frac{n+1}{n} = -\frac{1}{n(n+1)} < 0, \quad \forall\, n \in \mathbb{N}.$$

Esempio 3. Consideriamo la successione

$$a_n = n^2$$

essa associa ad $n = 1, 2, 3, ...$ i risultati $a_1 = 1$, $a_2 = 4$, $a_3 = 9$, $a_4 = 16, \ldots$

In particolare si ha che tale successione è monotòna crescente perché vale $a_{n+1} > a_n$, come si deduce eseguendo $a_{n+1} - a_n = (n+1)^2 - n^2 = 2n + 1 > 0, \quad \forall\, n \in \mathbb{N}$.

Definizione 5.1.2. *Una successione a_n si dice "limitata" se esistono due numeri reali L, M tali che valga $L \leqslant a_n \leqslant M$, per ogni $n \in \mathbb{N}$; si dice "limitata superiormente" se esiste un numero reale positivo M tale che valga $a_n \leqslant M$, per ogni $n \in \mathbb{N}$; si dice "limitata inferiormente" se esiste un numero reale positivo L tale che valga $L \leqslant a_n$, per ogni $n \in \mathbb{N}$.*

Ovviamente una successione è limitata se risulta limitata sia superiormente che inferiormente. In particolare si ha che le successioni degli esempi 1 e 2 sono limitate: infatti per quella del primo esempio risulta $0 \leqslant a_n \leqslant 1$, per ogni $n \in \mathbb{N}$; mentre per quella del secondo esempio risulta $1 \leqslant a_n \leqslant 2$, per ogni $n \in \mathbb{N}$. La successione del terzo esempio, invece, è limitata solo inferiormente perché risulta $0 \leqslant a_n = n^2$, ma per qualsiasi M che si scelga, non si potrà avere $a_n \leqslant M$, per ogni $n \in \mathbb{N}$, in quanto, non appena $n > \sqrt{M}$, si ha $a_n > M$.

Definizione 5.1.3. *Una successione a_{n_k} prende il nome di "sottosuccessione" della successione a_n se essa è definita come $a_{n_k} : \mathcal{S} \longrightarrow \mathbb{R}$, dove $\mathcal{S}$ è un sottoinsieme infinito dell'insieme $\mathbb{N}$ dei numeri naturali.*

Esempio. Data la successione

$$a_n = (-1)^n \frac{3n}{n+5}$$

e il sottoinsieme dei naturali dato dai pari, segue che la sottosuccessione definita su tale sottoinsieme assume la forma

$$a_{2k} = \frac{6k}{2k+5},$$

dove si è posto $n_k = 2k$; se si considera invece il sottoinsieme dei naturali dato dai dispari, allora la sottosuccessione definita su tale sottoinsieme assume la forma

$$a_{2k+1} = -\frac{6k+3}{2k+6},$$

dove si è posto $n_k = 2k + 1$.

Viene lasciata come esercizio la verifica che la sottosuccessione sui pari è monotòna crescente, mentre quella sui dispari è monotòna decrescente da cui si deduce quindi che esiste $M = 3$ con

$$-3 \leqslant a_n \leqslant 3$$

ovvero che la successione a_n è limitata.

5.2 Limiti di successioni

Definizione 5.2.1. *Data una successione* $a_n : \mathbb{N} \longrightarrow \mathbb{R}$*, si dice che essa, per* n *tendente all'infinito, possiede limite* L*, oppure "converge" a* L*, con* $L \in \mathbb{R}$*, se si verifica la seguente condizione: per ogni* $\varepsilon > 0$*, esiste* $n_\varepsilon \in \mathbb{N}$ *tale che valga* $L - \varepsilon \leqslant a_n \leqslant L + \epsilon$*,* $\forall\, n > n_\epsilon$*.*

Ricordiamo che la disuguaglianza $L-\varepsilon \leqslant a_n \leqslant L+\epsilon$ può essere anche scritta nella forma $|a_n - L| \leqslant \varepsilon$. Quando una successione $a_n : \mathbb{N} \longrightarrow \mathbb{R}$ converge a L per n tendente all'infinito, si scrive

$$\lim_{n\to+\infty} a_n = L.$$

Quando si dice soltanto che una successione è *convergente*, si intende che essa possiede un certo limite numerico L senza che sia importante specificare quale numero $L \in \mathbb{R}$ sia effettivamente il suo limite.

Definizione 5.2.2. *Si dice che una successione* $a_n : \mathbb{N} \longrightarrow \mathbb{R}$ *"diverge" per* n *tendente all'infinito, se*

Ricordiamo che le disuguaglianze $a_n > M$ oppure $a_n < -M$ possono essere espresse con l'unica scrittura $|a_n| > M$. Quando una successione $a_n : \mathbb{N} \longrightarrow \mathbb{R}$ diverge all'infinito per n tendente all'infinito, si scrive

$$\lim_{n\to+\infty} a_n = +\infty.$$

Quando una successione $a_n : \mathbb{N} \longrightarrow \mathbb{R}$ diverge all'infinito negativo, per n tendente all'infinito, si scrive

$$\lim_{n\to+\infty} a_n = -\infty.$$

Si riconosce immediatamente che la successione $a_n = (-1)^n$, la quale che assume gli unici due valori ± 1, non possiede limite. Quindi concludiamo che una successione può avere dunque tre comportamenti: può essere convergente ad un numero L, può essere divergente a $+\infty$ o a $-\infty$, può essere priva di limite, come la successione $a_n = (-1)^n$.

Teorema 5.2.1. *Se una successione* a_n *è convergente, allora essa è limitata.*

Dimostrazione. Detto L il limite di tale successione, avremo, per definizione di successione convergente, che per ogni $n > n_\varepsilon$ essa verifica la relazione di disuguaglianza $L-\varepsilon \leqslant a_n \leqslant L+\varepsilon$; mentre gli elementi a_n, con $1 \leqslant n \leqslant n_\varepsilon$, sono in quantità finita e pertanto sono tutti compresi fra il loro minimo M_1 e il loro massimo M_2. Confrontando quindi $L-\varepsilon$ con M_1 e poi $L+\varepsilon$ con M_2, si conclude allora che esiste un numero M tale che tutti gli elementi della successione siano compresi nell'intervallo $[-M, M]$, cioè che la successione è limitata, sia superiormente che inferiormente, *c.d.d.*

Osserviamo infine che scambiando ipotesi e tesi del teorema 5.2.1 non si ottiene un teorema valido perché se una successione è limitata, non è detto che sia convergente, come dimostra il contresempio dato dalla successione, già discussa, $a_n = (-1)^n$, che è limitata perché assume i soli valori ± 1, ma non risulta convergente. Tale successione contiene le due sottosuccessioni, sui pari e sui dispari, date da $a_{2k} = 1$ e $a_{2k+1} = -1$ che convergono rispettivamente a 1 e a -1. Riguardo ad una successione limitata, abbiamo allora un teorema collegato solo alle sue sottosuccessioni, di cui omettiamo la dimostrazione.

Teorema 5.2.2. *Una successione limitata possiede sempre almeno una sottosuccessione convergente.*

Poiché dunque la limitatezza di una successione non è sufficiente ad assicurarne anche la convergenza, allora ci chiediamo quale sia una condizione per la successione che implichi automaticamente la convergenza della successione stessa.

Teorema 5.2.3. *Una successione monotòna ammette sempre un limite uguale ad un numero $L \in \mathbb{R}$ oppure uguale all'infinito.*

Dimostrazione. Consideriamo solo il caso che la successione sia monotòna crescente. Si possono verificare allora due sole situazioni: la successione è limitata oppure non limitata. Nel caso in cui la successione sia monotòna crescente e limitata, allora essa possiede l'estremo superiore la cui definizione coincide con quella di limite che pertanto esiste ed è un numero. Se la successione è monotòna crescente e non limitata, allora la successione crescente assume valori maggiori di qualunque valore $M > 0$ scelto arbitrariamente e pertanto diverge a $+\infty$. E' immediata l'estensione del teorema ai casi di successione monotòna decrescente e di successione monotòna (crescente o decrescente) a partire da un certo elemento $a_{\bar{n}}$. La proprietà di monotonia della successione fa sì che il comportamento indeterminato (associato ad una sorta di "*oscillazione*") sia escluso, *c.d.d.*

Per il calcolo dei limiti di successione, risulta spesso utile un teorema che collega il comportamento di una successione al comportamento di altre successioni.

Teorema 5.2.4. *(del confronto)*

Siano a_n, b_n, c_n tre successioni ad elementi non negativi e tali che per ogni $n \in \mathbb{N}$ valga la disuguaglianza $a_n \leqslant b_n \leqslant c_n$. Se le successioni a_n e c_n possiedono il medesimo limite

$$\lim_{n\to+\infty} a_n = \lim_{n\to+\infty} c_n = L,$$

allora si ha anche

$$\lim_{n\to+\infty} b_n = L.$$

Se invece la successione a_n diverge a $+\infty$, allora segue anche entrambe le successioni b_n e c_n divergono a $+\infty$.

Dimostrazione. Per definizione di convergenza, se a_n e c_n convergono a L, si hanno, per ogni $\varepsilon > 0$ scelto, le disuguaglianze

$$L - \varepsilon \leqslant a_n \leqslant L + \varepsilon, \quad \forall\, n > n_{1\varepsilon} \qquad \text{e} \qquad L - \varepsilon \leqslant c_n \leqslant L + \varepsilon, \quad \forall\, n > n_{2\varepsilon}\,. \tag{5.1}$$

Se indichiamo con n_ε il numero più grande tra i due numeri $n_{1\varepsilon}, n_{2\varepsilon}$, segue che per $n > n_\varepsilon$ valgono entrambe le disuguaglianze (5.1) che possiamo riunire nella sequenza di disuguaglianze $L - \varepsilon \leqslant a_n \leqslant b_n \leqslant c_n \leqslant L + \varepsilon$, da cui estraiamo la disuguaglianza per la successione b_n data da $L - \varepsilon \leqslant b_n \leqslant L + \varepsilon$, per ogni $n > n_\varepsilon$, che definisce il limite

$$\lim_{n\to+\infty} b_n = L.$$

Si lascia come semplice esercizio di dimostrare la parte riguardante il caso di successioni divergenti, *c.d.d.*

Riassumiamo in conclusione l'essenza del teorema 5.2.4 dicendo che nel caso di elementi non negativi, una successione con elementi non maggiori degli elementi di una successione convergente, converge anch'essa; mentre una successione con elementi non minori degli elementi di una successione divergente, diverge anch'essa. Non si può dire nulla invece di una successione ad elementi non maggiori degli elementi di una successione divergente; né di una successione ad elementi non minori degli elementi di una successione convergente.

5.3 Il numero e

Il numero e viene chiamato anche numero di Nepero, dal matematico scozzese di nome John Napier latinizzato come Giovanni Nepero. Consideriamo dunque la seguente successione numerica $a_n : \mathbb{N} \longrightarrow \mathbb{R}$

$$a_n = \left(1 + \frac{1}{n}\right)^n, \tag{5.2}$$

che per $n = 1, 2, 3, \ldots$ genera i risultati rispettivamente 2, 9/4, 64/27, ...

Dimostriamo che tale successione a_n verifica due proprietà che possiamo esprimere tramite le due disuguaglianze $a_{n+1} > a_n$ e $2 \leqslant a_n < 3$, valide per ogni numero $n \in \mathbb{N}$, le quali mostrano che la successione dei numeri a_n è, rispettivamente, crescente e limitata.

Per dimostrare che i valori generati dalla a_n sono crescenti, utilizziamo la disuguaglianza di Bernouilli $(1-x)^n > 1 - nx$, valida per ogni numero intero $n > 1$ e per ogni x nell'intervallo $x \in (0,1)$. La disuguaglianza di Bernouilli si dimostra con le disuguaglianze

$$\begin{aligned}
&(1-x)^2 = 1 - 2x + x^2 > 1 - 2x,\\
&(1-x)^3 = (1-x)^2(1-x) > (1-2x)(1-x) = 1 - 3x + 2x^2 > 1 - 3x,\\
&(1-x)^4 = (1-x)^3(1-x) > (1-3x)(1-x) = 1 - 4x + 3x^2 > 1 - 4x,\\
&(1-x)^5 = (1-x)^4(1-x) > (1-4x)(1-x) = 1 - 5x + 3x^2 > 1 - 5x
\end{aligned}$$

e così via, da cui, "per induzione", segue che per ogni valore $n \in \mathbb{N}$ e per ogni $x \in (0,1)$ risulta appunto valida "in generale" la disuguaglianza $(1-x)^n > 1 - nx$.

Possiamo a questo punto dimostrare che per ogni n numero naturale vale la disuguaglianza $a_{n+1} > a_n$, scritta nella forma equivalente $a_{n+1}/a_n > 1$. Abbiamo dunque

$$\frac{a_{n+1}}{a_n} = \frac{\left(1 + \dfrac{1}{n+1}\right)^{n+1}}{\left(1 + \dfrac{1}{n}\right)^n} = \frac{\left(\dfrac{n+2}{n+1}\right)^{n+1}}{\left(\dfrac{n+1}{n}\right)^n} = \frac{\left(\dfrac{n+2}{n+1}\right)\left(\dfrac{n+2}{n+1}\right)^n}{\left(\dfrac{n+1}{n}\right)^n} =$$

$$= \left(\frac{n+2}{n+1}\right)\left[\left(\frac{n+2}{n+1}\right)\left(\frac{n}{n+1}\right)\right]^n = \left(\frac{n+2}{n+1}\right)\left[\frac{n^2+2n}{n^2+2n+1}\right]^n =$$

$$= \left(\frac{n+2}{n+1}\right)\left[1 - \frac{1}{n^2+2n+1}\right]^n > \left(\frac{n+2}{n+1}\right)\left[1 - \frac{n}{n^2+2n+1}\right] = \frac{n^3+3n^2+3n+2}{n^3+3n^2+3n+1} > 1,$$

dove nell'ultima disuguaglianza il numeratore è maggiore del denominatore e abbiamo utilizzato la disuguaglianza di Bernouilli con $x = 1/(n^2 + 2n + 1) \in (0,1)$.

Poiché se consideriamo primo e ultimo membro della sequenza risulta quindi

$$a_{n+1}/a_n > 1, \quad \text{per ogni } n \in \mathbb{N},$$

ovvero $a_{n+1} > a_n$, concludiamo che ogni termine successivo a_{n+1} risulta sempre maggiore del proprio termine precedente a_n, ovvero appunto la successione a_n è monotòna crescente ed in particolare tutti i numeri a_n sono maggiori di 2 perché 2 è il primo elemento a_1.

Dimostriamo ora la seconda proprietà fondamentale della successione a_n che stabilisce la limitatezza della successione a_n e che si esprime tramite la disuguaglianza

$$2 \leqslant \left(1 + \frac{1}{n}\right)^n < 3, \quad \text{per ogni } n \in \mathbb{N}, \tag{5.3}$$

dove la parte di sinistra è stata già dimostrata perché risulta $2 \leqslant a_n$, per ogni $n \in \mathbb{N}$.

Per dimostrare la parte di destra della disuguaglianza (5.3), ricordiamo la struttura della potenza n-esima del binomio $1 + x$ e utilizziamo l'espressione della somma (4.15a).

Lo sviluppo di $(1+x)^n$ è dato dalla somma dei termini $1, x, x^2, x^3, \ldots, x^{n-1}, x^n$, ciascuno moltiplicato per il relativo coefficiente che si trova nella riga n-esima (riga n-esima corrispondente all'esponente n) del *triangolo di Tartaglia* (4.7). Poiché per $n = 1$ si ha ovviamente $(1 + x)^1 = 1 + x$, scriviamo alcune potenze con esponenti $n > 1$, ovvero

$$\begin{aligned}(1+x)^2 &= 1 + 2x + x^2\\ (1+x)^3 &= 1 + 3x + 3x^2 + x^3\\ (1+x)^4 &= 1 + 4x + 6x^2 + 4x^3 + x^4\\ (1+x)^5 &= 1 + 5x + 10x^2 + 10x^3 + 5x^4 + x^5\end{aligned} \tag{5.4}$$

e così via. Dopo aver osservato che per ogni n il primo e l'ultimo addendo dello sviluppo sono sempre rispettivamente 1 e x^n, abbiamo inoltre che per ogni n tutti i coefficienti dei termini intermedi dello sviluppo da x a x^{n-1}, dati dalla corrispondente riga del *triangolo di Tartaglia*, sono nell'ordine

$$n, \quad \frac{n(n-1)}{2!}, \quad \frac{n(n-1)(n-2)}{3!}, \quad \frac{n(n-1)(n-2)(n-3)}{4!}, \quad \text{e così via,} \tag{5.5}$$

i quali vengono denominati *fattori binomiali.* A questo punto sviluppiamo l'espressione della a_n nella (5.2) utilizzando prima l'espressione dei coefficienti (5.5) con $x = 1/n$ e quindi la disuguaglianza, valida per ogni $0 < z < 1$ ed in particolare per $z = 1/2$, che si ottiene dalla relazione (4.15a), ovvero

$$1 + z + z^2 + z^3 + z^4 + \cdots + z^{n-1} + z^n = \frac{1 - z^{n+1}}{1 - z} < \frac{1}{1 - z}\,.$$

Abbiamo dunque

$$\begin{aligned}\left(1 + \frac{1}{n}\right)^n &= 1 + 1 + \frac{n(n-1)}{2}\,\frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!}\,\frac{1}{n^3} + \frac{n(n-1)(n-2)(n-3)}{4!}\,\frac{1}{n^4} + \cdots =\\ &= 1 + 1 + \frac{n(n-1)}{n^2}\,\frac{1}{2} + \frac{n(n-1)(n-2)}{n^3}\,\frac{1}{3!} + \frac{n(n-1)(n-2)(n-3)}{n^4}\,\frac{1}{4!} + \cdots <\\ &< 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \cdots < 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} \cdots =\\ &= 1 + \left[1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} \ldots\right] < 1 + \frac{1}{1 - \frac{1}{2}} = 3,\end{aligned}$$

dove sono state utilizzate inoltre le evidenti disuguaglianze $k! > 2^{k-1}$, per ogni $k \in \mathbb{N}$, e

$$\frac{n(n-1)}{n^2} < 1, \qquad \frac{n(n-1)(n-2)}{n^3} < 1, \qquad \frac{n(n-1)(n-2)(n-3)}{n^4} < 1 \quad \text{e così via.}$$

Poiché dunque l'insieme dei numeri razionali (e quindi anche reali) $\{a_n\}$ risulta limitato, in particolare tra i numeri 2 e 3, segue, per una proprietà dei numeri reali, che tale insieme possiede quelli che vengono denominati *estremo inferiore*, indicato con il simbolo *inf*, ed *estremo superiore*, indicato con il simbolo *sup*.

L'*estremo inferiore* coincide con il primo elemento della successione che si ottiene per $n = 1$ e vale 2 in quanto $(1 + 1)^1 = 2$. Tale *estremo inferiore* è pertanto anche il *minimo* dell'insieme dei numeri $\{a_n\}$. Poiché gli elementi della successione sono infiniti ed ogni elemento, come abbiamo dimostrato, risulta sempre maggiore dell'elemento precedente, concludiamo che l'elemento *massimo* di tutto l'insieme dei numeri $\{a_n\}$ non esiste, ma di tale insieme esiste solo l'*estremo superiore*, denotato con il simbolo e, che risulta ovviamente maggiore del primo elemento della successione, ovvero del numero 2, e risulta minore del numero 3 oppure può arrivare a valere anche 3 stesso.

Abbiamo quindi

$$e = \sup_{n\in\mathbb{N}} \left\{ \left(1 + \frac{1}{n}\right)^n \right\} = \sup \left\{ 2\,, \frac{9}{4}\,, \frac{64}{27}\,, \frac{625}{256}\,, \dots \right\} \leqslant 3$$

e se rappresentiamo le coppie $(1,2)$, $(2,9/4)$, $(3,64/27)$, $(4,625/256),\dots$, generate dalla successione a_n, nel riferimento cartesiano in fig. 5.1, è immediato rendersi conto che, in virtù del comportamento monotòno crescente della successione a_n e della definizione della convergenza all'infinito, l'*estremo superiore* dell'insieme $\{a_n\}$ coincide con il limite della successione a_n stessa per n tendente all'infinito, ovvero risulta

$$e = \sup_{n\in\mathbb{N}} \left\{ \left(1 + \frac{1}{n}\right)^n \right\} = \sup \left\{ 2\,, \frac{9}{4}\,, \frac{64}{27}\,, \frac{625}{256}\,, \dots \right\} = \lim_{n\longrightarrow\infty} \left(1 + \frac{1}{n}\right)^n .$$

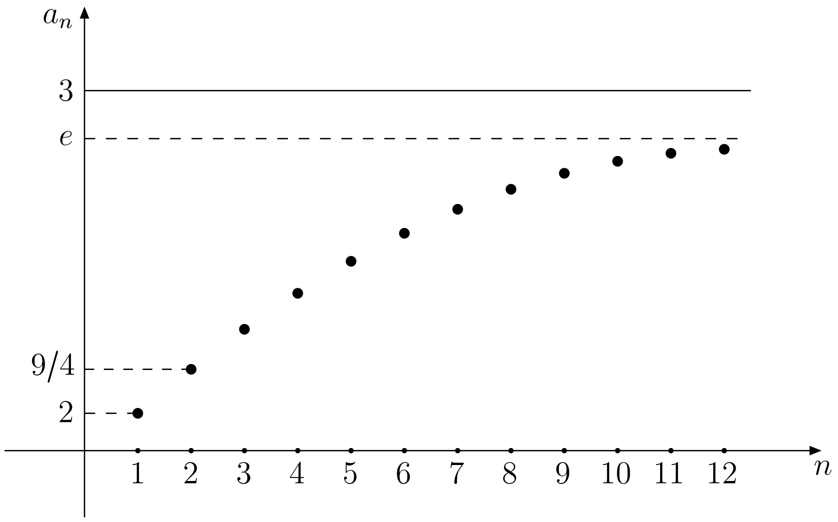


Fig. 5.1

A questo punto, senza affrontare in modo rigoroso il delicato problema del passaggio dalla variabile intera positiva discreta n alla variabile reale continua x, possiamo cambiare il tipo di variabile e scrivere in conclusione

$$\lim_{x\longrightarrow\pm\infty} \left(1 + \frac{1}{x}\right)^x = \lim_{n\longrightarrow\infty} \left(1 + \frac{1}{n}\right)^n = e.$$

Se ora prendiamo il logaritmo in base e della a_n, indicato semplicemente con log senza che se ne scriva la base, abbiamo la sequenza

$$\lim_{x\longrightarrow\pm\infty} \log \left(1 + \frac{1}{x}\right)^x = \lim_{n\longrightarrow\infty} \log \left(1 + \frac{1}{n}\right)^n = \log e = 1. \tag{5.6}$$

Per le proprietà del logaritmo possiamo riscrivere le relazioni (5.6) nella forma (riferendoci alla sola variabile reale x)

$$\lim_{x\longrightarrow\pm\infty} \frac{\log\left[1 + (1/x)\right]}{1/x} = 1. \tag{5.7}$$

Se ora poniamo $z = 1/x$, abbiamo che per x tendente all'infinito (sia positivo che negativo), la nuova variabile z tende a zero e quindi possiamo scrivere il limite (5.7) nella nuova forma (*limite notevole del logaritmo*)

$$\lim_{z\longrightarrow 0} \frac{\log(1 + z)}{z} = 1. \tag{5.8}$$

Se infine poniamo $w = \log(1+z)$, abbiamo che per z tendente a zero, anche la nuova variabile w tende a zero e quindi dal limite notevole del logaritmo (5.8) otteniamo il *limite notevole dell'esponenziale*

$$\lim_{w\longrightarrow 0} \frac{w}{e^w - 1} = \lim_{w\longrightarrow 0} \frac{e^w - 1}{w} = 1. \tag{5.9}$$

Vogliamo dimostrare ora il seguente limite con $n \in \mathbb{N}$

$$\lim_{n\to\infty} \frac{\log n}{n} = 0 \tag{5.10}$$

e a tale scopo basta dimostrare la sola parte destra della disuguaglianza

$$1 \leqslant \sqrt[n]{n} < 1 + \frac{2}{\sqrt{n}}, \tag{5.11a}$$

la cui parte sinistra è evidente perché per ogni $n \in \mathbb{N}$ si ha ovviamente $\sqrt[n]{n} \geqslant 1$.

Prima di dimostrare la disuguaglianza (5.11a) vogliamo dedurne un'importante conseguenza. Dal teorema del confronto applicato alla disuguaglianza (5.11a) segue

$$1 \leqslant \lim_{n\longrightarrow\infty} \sqrt[n]{n} \leqslant \lim_{n\longrightarrow\infty} \left(1 + \frac{2}{\sqrt{n}}\right) = 1,$$

da cui ricaviamo

$$\lim_{n\longrightarrow\infty} \sqrt[n]{n} = 1. \tag{5.11b}$$

Se poi nel limite (5.11b) applichiamo il logaritmo ad ambo i membri, otteniamo

$$\lim_{n\longrightarrow\infty} \log \sqrt[n]{n} = \lim_{n\longrightarrow\infty} \log n^{1/n} = \lim_{n\longrightarrow\infty} \frac{\log n}{n} = \log 1 = 0, \tag{5.11c}$$

ovvero il limite (5.10) dato dall'uguaglianza tra terzo e ultimo membro della (5.11c).

Per dimostrare dunque la parte destra della disuguaglianza (5.11a), eleviamo alla potenza con esponente n il secondo e il terzo membro della disuguaglianza (5.11a) stessa.

Abbiamo dunque $\left(\sqrt[n]{n}\right)^n = n$ e utilizzando i coefficienti (5.5) con $x = 2/\sqrt{n}$

$$\left(1 + \frac{2}{\sqrt{n}}\right)^n = 1 + \frac{2n}{\sqrt{n}} + \frac{n(n-1)}{2}\frac{4}{n} + \ldots > 1 + \frac{2n}{\sqrt{n}} + \frac{n(n-1)}{2}\frac{4}{n} = 2\left(n + \sqrt{n}\right) - 1 > n,$$

ovvero

$$\left(1 + \frac{2}{\sqrt{n}}\right)^n > n = \left(\sqrt[n]{n}\right)^n.$$

Poiché quindi la potenza n-esima del secondo membro della (5.11a) risulta minore della potenza n-esima del terzo membro della (5.11a) stessa, segue in conclusione la parte destra della disuguaglianza (5.11a), da cui abbiamo già ricavato il limite (5.10).

In virtù del limite (5.10), segue che il limite di ogni frazione avente un logaritmo al numeratore e l'argomento del logaritmo stesso al denominatore è zero se l'argomento del logaritmo e il denominatore, oltre ad essere coincidenti, tendono ad infinito.

Dalla sequenza di identità con α, β coppia di numeri reali positivi

$$\frac{\log^\alpha n}{n^\beta} = \left(\frac{\log n}{n^{\beta/\alpha}}\right)^\alpha = \left(\frac{\alpha}{\beta}\right)^\alpha \left[\frac{\log\left(n^{\beta/\alpha}\right)}{n^{\beta/\alpha}}\right]^\alpha$$

e con la sostituzione della variabile naturale n con la variabile reale x, otteniamo infine il limite per ogni coppia di numeri α, β reali positivi

$$\lim_{x \longrightarrow +\infty} \frac{\log^\alpha x}{x^\beta} = \left(\frac{\alpha}{\beta}\right)^\alpha \left[\lim_{x \longrightarrow +\infty} \frac{\log\left(x^{\beta/\alpha}\right)}{\left(x^{\beta/\alpha}\right)}\right]^\alpha = 0. \tag{5.12}$$

Utilizzando il risultato (5.12) e indicando con $\mathcal{P}_n(x), \mathcal{Q}_m(x)$ due polinomi di grado rispettivamente n, m, possiamo dimostrare che il limite della funzione rapporto tra $\log \mathcal{P}_n(x)$ e $\mathcal{Q}_m(x)$ è sempre zero per x tendente ad ogni infinito, positivo o negativo, compatibile con il dominio della funzione. Supponendo infatti, senza perdita di generalità e riconoscendo che la dimostrazione con l'altro infinito sarebbe identica, che $+\infty$ appartenga al dominio della funzione, abbiamo

$$\lim_{x \longrightarrow +\infty} \frac{\log \mathcal{P}_n(x)}{\mathcal{Q}_m(x)} = \lim_{x \longrightarrow +\infty} \frac{\log(a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots)}{b_m x^m + b_{m-1}x^{m-1} + b_{m-2}x^{m-2} + \dots} =$$

$$= \lim_{x \longrightarrow +\infty} \frac{\log\left[x^n\left(a_n + \dfrac{a_{n-1}}{x} + \dfrac{a_{n-2}}{x^2} + \dots\right)\right]}{x^m\left(b_m + \dfrac{b_{m-1}}{x} + \dfrac{b_{m-2}}{x^2} + \dots\right)} =$$

$$= \lim_{x \longrightarrow \infty} \frac{n \log x + \log\left(a_n + \dfrac{a_{n-1}}{x} + \dfrac{a_{n-2}}{x^2} + \dots\right)}{x^m\left(b_m + \dfrac{b_{m-1}}{x} + \dfrac{b_{m-2}}{x^2} + \dots\right)} = \lim_{x \longrightarrow \infty}\left(\frac{\log x}{x^m}\right)\left[\frac{n}{b_m + \dfrac{b_{m-1}}{x} + \dfrac{b_{m-2}}{x^2} + \dots}\right] +$$

$$+ \lim_{x \longrightarrow +\infty}\left[\frac{\log\left(a_n + \dfrac{a_{n-1}}{x} + \dfrac{a_{n-2}}{x^2} + \dots\right)}{b_m + \dfrac{b_{m-1}}{x} + \dfrac{b_{m-2}}{x^2} + \dots}\right]\left(\frac{1}{x^m}\right) = 0$$

perché ciascuno dei due limiti finali vale zero in quanto le parentesi quadre di entrambi tendono ad un numero, mentre le parentesi tonde tendono a zero.

Vogliamo in ultimo dimostrare il limite con $n \in \mathbb{N}$ e α reale

$$\lim_{n \longrightarrow \infty} \frac{n^\alpha}{e^n} = 0. \tag{5.13}$$

Nel caso α reale negativo, il limite (5.13) è ovvio. Pertanto dimostriamo tale limite solo per α positivo e in particolare per α intero, essendo immediata l'estensione ad ogni α reale. Poiché la procedura della dimostrazione, come ci si renderà facilmente conto, è la medesima per ogni esponente α intero positivo, scegliamo, a mo' di esempio, il caso con esponente $\alpha = 2$. Utilizziamo la disuguaglianza di Bernouilli nella forma $(1+x)^n \geqslant 1 + nx$, con x positivo, la cui dimostrazione è immediata perché nelle (5.4) si riconosce che $1 + nx$ al secondo membro, con qualsiasi n fissato, è minore di tutto lo sviluppo di $(1+x)^n$.

Poiché si ha $e > 1$, anche tutte le sue radici con qualunque indice risultano maggiori di 1 e allora, per l'esponente $\alpha = 2$ nel limite, consideriamo la radice cubica di e, ovvero la radice avente indice uguale all'esponente scelto più 1 (se il limite avesse $\alpha = 3$, prenderemmo la radice quarta di e e così via). Poiché $\sqrt[3]{e} > 1$, poniamo $\sqrt[3]{e} = 1 + h$, con $h > 0$, da cui segue

$$\sqrt[3]{e^n} = \left(\sqrt[3]{e}\right)^n = (1+h)^n > 1 + nh > nh \qquad \Longrightarrow \qquad e^n > (hn)^3 = h^3 n^3.$$

Tenendo presente che $1/h^3$ è un numero che chiamiamo c e applicando il teorema del confronto, ricaviamo dunque

$$0 < \frac{n^2}{e^n} < \frac{n^2}{h^3 n^3} = \frac{(1/h^3)}{n} = \frac{c}{n} \qquad \Longrightarrow \qquad 0 < \lim_{n \longrightarrow \infty} \frac{n^2}{e^n} < \lim_{n \longrightarrow \infty} \frac{c}{n} = 0,$$

ovvero abbiamo dimostrato in conclusione il risultato del limite (5.13) con esponente $\alpha = 2$ e in generale con ogni esponente α reale. Sostituendo, come in precedenza, la variabile naturale n con la variabile reale x, otteniamo infine il limite

$$\lim_{x \longrightarrow +\infty} \frac{\mathcal{P}_n(x)}{e^x} = \lim_{x \longrightarrow +\infty} \frac{a_n x^n + a_{n-1} x^{n-1} + \dots}{e^x} = \lim_{x \longrightarrow +\infty} \left(\frac{x^n}{e^x}\right) \left[a_n + \frac{a_{n-1}}{x} + \dots\right] = 0$$

perché la parentesi tonda tende a zero in virtù del limite (5.13) e la parentesi quadra tende alla costante a_n.

5.4 Serie numeriche

Per cominciare consideriamo soltanto le serie ad addendi non negativi perché le serie con addendi a segno alterno possono essere studiate in un secondo momento.

Innanzitutto introduciamo la notazione

$$a_1 + a_2 + a_3 + \dots + a_n \equiv \sum_{k=1}^{n} a_k$$

per indicare la somma degli n addendi $a_1, a_2, a_3, \dots, a_n$ senza doverli esplicitare sempre ogni volta.

Definizione 5.4.1. *Data una successione a_k di elementi non negativi, ovvero $a_k \geqslant 0$, la successione S_n di espressione*

$$S_n = \sum_{k=1}^{n} a_k$$

viene denominata "successione delle somme parziali" (degli addendi a_k) e il limite

$$\lim_{n \to +\infty} S_n = \lim_{n \to +\infty} \sum_{k=1}^{n} a_k = \sum_{k=1}^{+\infty} a_k \tag{5.14}$$

prende il nome di "serie numerica di addendo generico a_k" e viene quindi indicata semplicemente con la scrittura

$$\sum_{k=1}^{+\infty} a_k.$$

La successione a_k viene denominata "successione dell'addendo generico" o, semplicemente, "addendo generico".

Sottolineiamo dunque che gli elementi della successione S_n non sono $a_1, a_2, a_3, \ldots$, bensì gli elementi

$$S_1 = a_1, \quad S_2 = a_1 + a_2, \quad S_3 = a_1 + a_2 + a_3, \qquad S_4 = a_1 + a_2 + a_3 + a_4, \quad \ldots.$$

e inoltre, poiché vale $a_k \geqslant 0$ per ogni k, la successione S_n è monotòna crescente perché si aggiunge ogni volta un addendo non negativo in più e dunque segue, per il teorema 5.2.4, che la successione S_n possiede sicuramente limite.

Definizione 5.4.2. *Data una serie numerica*

$$\sum_{k=1}^{+\infty} a_k = \lim_{n\to+\infty} S_n \equiv \lim_{n\to+\infty} \sum_{k=1}^{n} a_k\,,$$

con $a_k \geqslant 0$, per ogni k, se il limite per n tendente all'infinito della sucessione S_n delle sue somme parziali è un numero reale (positivo), allora la serie si dice "convergente"; mentre se per n tendente all'infinito la sucessione S_n è divergente, allora la serie si dice "divergente".

5.4.1 Esempio di *serie telescopica*

Consideriamo l'addendo generico di espressione $a_k = 1/(k^2+k)$ e studiamo la serie associata

$$\sum_{k=1}^{\infty} \frac{1}{k^2+k}\,, \tag{5.15}$$

che appartiene a quella categoria di serie chiamate *serie telescopiche.* Per determinare il "comportamento" di tale serie, studiamo la corrispondente successione delle somme parziali

$$S_n = \sum_{k=1}^{n} \frac{1}{k^2+k}$$

il cui limite è la serie (5.15). Abbiamo

$$S_n = \sum_{k=1}^{n} \frac{1}{k^2+k} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) =$$

$$= \left(1-\frac{1}{2}\right) + \left(\frac{1}{2}-\frac{1}{3}\right) + \left(\frac{1}{3}-\frac{1}{4}\right) + \cdots + \left(\frac{1}{n-1}-\frac{1}{n}\right) + \left(\frac{1}{n}-\frac{1}{n+1}\right) = 1 - \frac{1}{n+1}\,,$$

perché tutti gli addendi, tranne il primo e l'ultimo, si cancellano a due a due.

Segue dunque per la definizione (5.14) di serie

$$\sum_{k=1}^{\infty} \frac{1}{k^2+k} = \lim_{n\to+\infty} \sum_{k=1}^{n} \frac{1}{k^2+k} = \lim_{n\to+\infty} S_n = \lim_{n\to+\infty} \left(1 - \frac{1}{n+1}\right) = 1$$

e quindi concludiamo che la serie telescopica (5.15) risulta convergente, in particolare a 1.

Vediamo ora una condizione necessaria per la convergenza di una serie: poiché tale condizione non è sufficiente per la convergenza di una serie, si può soltanto dedurre la divergenza della serie quando tale condizione non è soddisfatta.

Teorema 5.4.1. *(condizione necessaria per la convergenza di una serie)*
Se una serie ad addendi non negativi risulta convergente, segue che il suo addendo generico tende a zero.

Dimostrazione. Dimostriamo il teorema per assurdo, ovvero dimostriamo che se l'addendo generico a_k di una una serie non tende a zero, segue che la serie risulta divergente. Se dunque, per assurdo, valesse

$$\lim_{k\to+\infty} a_k = L > 0,$$

possiamo trovare un $\varepsilon > 0$ “sufficientemente piccolo” tale che valga $L - \varepsilon > 0$ e, per la definizione di limite, valga la disuguaglianza $0 < L - \varepsilon < a_k < L + \varepsilon$, per ogni k maggiore di un certo k_ε. Sfruttando la disuguaglianza di sinistra, si ricava dunque

$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{k_\varepsilon} a_k + \sum_{k=k_\varepsilon+1}^{\infty} a_k > \sum_{k=1}^{k_\varepsilon} a_k + \sum_{k=k_\varepsilon+1}^{\infty} (L - \varepsilon) = +\infty$$

perché la somma di infiniti addendi tutti uguali a $L - \varepsilon > 0$ è appunto infinita, *c.d.d.*

Poiché una serie è il limite della successione delle sue somme parziali, segue che per stabilire se una serie è convergente o divergente si può applicare ad essa il criterio del confronto delle successioni (teorema 5.2.4), che per le serie viene espresso tramite il seguente teorema la cui dimostrazione è identica alla dimostrazione del teorema 5.2.4.

Teorema 5.4.2. *Date due serie*

$$\sum_{k=0}^{\infty} a_k \qquad e \qquad \sum_{k=0}^{\infty} b_k \,,$$

se da un certo k_0 in poi, cioè per ogni $k \geqslant k_0$, si ha la disuguaglianza $a_k < b_k$ e la serie di addendo generico b_k è convergente, segue che anche la serie di addendo generico a_k risulta convergente; se invece per ogni $k \geqslant k_0$ si ha $a_k > b_k$ e la serie di addendo generico b_k è divergente, segue che anche la serie di addendo generico a_k risulta divergente.

5.4.2 Serie geometrica

La serie avente addendo generico x^k, con $x \in \mathbb{R}$, prende il nome di *serie geometrica di ragione x*. Pertanto la serie geometrica si scrive

$$\sum_{k=0}^{\infty} x^k$$

Nel caso $x = 1$, la serie geometrica è divergente perché si ha

$$\sum_{k=0}^{\infty} 1 = \lim_{n\to+\infty} \sum_{k=1}^{n} 1 = \lim_{n\to+\infty} \overbrace{1 + 1 + 1 + \ldots + 1}^{n \text{ volte}} = \lim_{n\to+\infty} n = \infty.$$

Per studiare la serie geometrica nel caso $x \neq 1$, utilizziamo l'identità (4.15a), dalla quale ricaviamo

$$\sum_{k=0}^{\infty} x^k = \lim_{n\to+\infty} \sum_{k=1}^{n} x^k = \lim_{n\to+\infty} \frac{1 - x^{n+1}}{1 - x} = \begin{cases} \dfrac{1}{1-x}, & \text{se } -1 < x < 1 \\ \infty, & \text{se } x \geqslant 1 \\ \text{indeterminata}, & \text{se } x \leqslant 1 \end{cases}$$

perché per $-1 < x < 1$ risulta

$$\lim_{n\to+\infty} x^{n+1} = 0;$$

per $x \geqslant 1$ risulta

$$\lim_{n\to+\infty} x^{n+1} = \infty;$$

per $x \leqslant -1$ risulta che il limite

$$\lim_{n\to+\infty} x^{n+1}$$

non esiste.

5.4.3 Serie armonica

La serie avente addendo generico $1/k^\alpha$, con $\alpha > 0$, prende il nome di *serie armonica di esponente* α. Pertanto la serie armonica di esponente α si scrive

$$\sum_{k=0}^{\infty} \frac{1}{k^\alpha}$$

La serie armonica di esponente $\alpha = 1$ viene chiamata semplicemente *serie armonica.*

Dimostriamo innanzitutto che la serie armonica, sebbene per essa valga la condizione necessaria per la convergenza di una serie, è divergente.

Sviluppando le somme parziali della serie armonica, si ha

$$S_2 = 1 + \frac{1}{2}, \qquad S_4 = \left(1 + \frac{1}{2}\right) + \frac{1}{3} + \frac{1}{4} > \left(1 + \frac{1}{2}\right) + 2 \cdot \frac{1}{4} = 1 + 2 \cdot \frac{1}{2},$$

$$S_{2^3} = S_8 = S_4 + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > 1 + 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{8} = 1 + 3 \cdot \frac{1}{2},$$

dove sono state sfruttate le disuguaglianze

$$\frac{1}{3} + \frac{1}{4} > \frac{1}{4} + \frac{1}{4} > 2 \cdot \frac{1}{4} = \frac{1}{2} \qquad \text{e} \qquad \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > 4 \cdot \frac{1}{8} = \frac{1}{2}.$$

In generale quindi si ha per ogni n

$$S_{2^n} > 1 + n \cdot \frac{1}{2} = \frac{n+2}{2},$$

da cui deduciamo la divergenza della serie armonica perché con il teorema del confronto ricaviamo

$$\sum_{k=0}^{\infty} \frac{1}{k} = \lim_{n\to+\infty} \sum_{k=0}^{2^n} \frac{1}{k} > \lim_{n\to+\infty} \frac{n+2}{2} = +\infty.$$

Sempre per confronto si ha poi che diverge anche ogni serie armonica di esponente $\alpha \in (0,1)$ perché per $0 < \alpha < 1$ si ha

$$\lim_{n\to+\infty} \sum_{k=0}^{n} \frac{1}{k^\alpha} > \lim_{n\to+\infty} \sum_{k=0}^{n} \frac{1}{k} = +\infty,$$

dove si è sfruttata la disuguaglianza $k^\alpha < k$, ovvero

$$\frac{1}{k^\alpha} > \frac{1}{k}.$$

Dimostriamo ora che la serie armonica di esponente $\alpha > 1$ è convergente. Per $\alpha > 1$ abbiamo le seguenti disuguaglianze

$$S_3 = 1 + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} < 1 + 2 \cdot \frac{1}{2^\alpha} = 1 + \frac{1}{2^{\alpha-1}}$$

$$S_7 = S_3 + \frac{1}{4^\alpha} + \frac{1}{5^\alpha} + \frac{1}{6^\alpha} + \frac{1}{7^\alpha} < 1 + \frac{1}{2^{\alpha-1}} + 4 \cdot \frac{1}{4^\alpha} = 1 + \frac{1}{2^{\alpha-1}} + \left(\frac{1}{2^{\alpha-1}}\right)^2$$

$$S_{15} = S_7 + \frac{1}{8^\alpha} + \dots + \frac{1}{15^\alpha} < 1 + \frac{1}{2^{\alpha-1}} + \left(\frac{1}{2^{\alpha-1}}\right)^2 + 8 \cdot \frac{1}{8^\alpha} = \sum_{k=0}^{3} \left(\frac{1}{2^{\alpha-1}}\right)^k$$

e quindi in generale, facendo attenzione agli indici

$$\sum_{k=1}^{2^n-1} \frac{1}{k^\alpha} = S_{2^n-1} < \sum_{h=0}^{n-1} \left(\frac{1}{2^{\alpha-1}}\right)^h$$

Passando al limite in ambo i membri di quest'ultima disuguaglianza, si ha

$$\sum_{k=1}^{+\infty} \frac{1}{k^\alpha} = \lim_{n\to+\infty} \sum_{k=0}^{2^n-1} \frac{1}{k^\alpha} < \lim_{n\to+\infty} \sum_{h=0}^{n-1} \left(\frac{1}{2^{\alpha-1}}\right)^h = \text{convergente}$$

perché per $\alpha > 1$

$$\sum_{h=0}^{+\infty} \left(\frac{1}{2^{\alpha-1}}\right)^h = \lim_{n\to+\infty} \sum_{h=0}^{n-1} \left(\frac{1}{2^{\alpha-1}}\right)^h$$

è una serie geometrica di ragione $1/(2^{\alpha-1}) < 1$ e quindi convergente. Per il teorema del confronto si conclude che la serie armonica di esponente $\alpha > 1$ è convergente.

Riassumendo abbiamo dunque

$$\sum_{k=0}^{\infty} \frac{1}{k^\alpha} = \lim_{n\to+\infty} \sum_{k=1}^{n} \frac{1}{k^\alpha} = \begin{cases} \text{divergente,} & \text{se} \quad \alpha \leqslant 1 \\ \text{convergente,} & \text{se} \quad \alpha > 1. \end{cases}$$

5.4.4 Due ulteriori serie particolari

Esempio 1. Consideriamo la serie

$$\sum_{k=2}^{+\infty} \frac{1}{k\log k} \tag{5.16a}$$

e studiamone la successione S_n delle somme parziali con la stessa procedura con cui è stata studiata la serie armonica. Tenendo conto che gli addendi di tale serie sono decrescenti e ricordando la proprietà del logaritmo $\log(2^m) = m\log 2$, abbiamo

$$S_4 = \sum_{k=2}^{4} \frac{1}{k\log k} = \frac{1}{2\log 2} + \frac{1}{3\log 3} + \frac{1}{4\log 4} > \frac{1}{2\log 2} + \frac{2}{4\log 4} =$$
$$= \frac{1}{2\log 2} + \frac{1}{4\log 2} = \frac{1}{2\log 2}\left(1 + \frac{1}{2}\right);$$
$$S_8 = S_4 + \frac{1}{5\log 5} + \dots + \frac{1}{8\log 8} > \frac{1}{2\log 2} + \frac{1}{4\log 2} + \frac{4}{8\log 8} =$$
$$= \frac{1}{2\log 2} + \frac{1}{4\log 2} + \frac{1}{6\log 2} = \frac{1}{2\log 2}\left(1 + \frac{1}{2} + \frac{1}{3}\right);$$
$$S_{16} = S_8 + \frac{1}{9\log 9} + \dots + \frac{1}{16\log 16} > \frac{1}{2\log 2} + \frac{1}{4\log 2} + \frac{1}{6\log 2} + \frac{8}{16\log 16} =$$
$$= \frac{1}{2\log 2} + \frac{1}{4\log 2} + \frac{1}{6\log 2} + \frac{1}{8\log 2} = \frac{1}{2\log 2}\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right)$$

e in generale

$$S_{2^n} > \frac{1}{2\log 2}\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n}\right) = \frac{1}{2\log 2}\sum_{k=1}^{n} \frac{1}{k},$$

da cui si ricava che la serie considerata è divergente perché la successione S_{2^n} delle sue somme parziali è maggiore della successione delle somme parziali della serie armonica che è appunto divergente

$$\sum_{k=2}^{+\infty} \frac{1}{k \log k} = \lim_{n \to +\infty} S_{2^n} = \lim_{n \to +\infty} \sum_{k=2}^{2^n} \frac{1}{k \log k} > \lim_{n \to +\infty} \frac{1}{2 \log 2} \sum_{k=1}^{n} \frac{1}{k} = \infty.$$

Esempio 2. Consideriamo la serie

$$\sum_{k=2}^{+\infty} \frac{1}{k (\log k)^2} \tag{5.16b}$$

e studiamone la successione S_n delle somme parziali come per la serie armonica tenendo conto che anche in questa serie gli addendi sono decrescenti

$$S_3 = \frac{1}{2(\log 2)^2} + \frac{1}{3(\log 3)^2} < 2 \cdot \frac{1}{2(\log 2)^2} = \frac{1}{(\log 2)^2}\,;$$

$$S_7 = S_3 + \frac{1}{4(\log 4)^2} + \ldots + \frac{1}{7(\log 7)^2} < \frac{1}{(\log 2)^2} + \frac{4}{4(\log 4)^2} = \frac{1}{(\log 2)^2}\left(1 + \frac{1}{4}\right);$$

$$\begin{aligned} S_{15} = S_7 + \frac{1}{8(\log 8)^2} + \ldots + \frac{1}{15(\log 15)^2} &< \frac{1}{(\log 2)^2} + \frac{1}{4(\log 2)^2} + \frac{8}{8(\log 8)^2} = \\ &= \frac{1}{(\log 2)^2} + \frac{1}{4(\log 2)^2} + \frac{1}{9(\log 2)^2} = \frac{1}{(\log 2)^2}\left(1 + \frac{1}{4} + \frac{1}{9}\right) \end{aligned}$$

e in generale

$$S_{2^{n+1}-1} < \frac{1}{(\log 2)^2}\left(1 + \frac{1}{4} + \frac{1}{9} + \ldots + \frac{1}{n^2}\right) = \frac{1}{(\log 2)^2} \sum_{k=1}^{n} \frac{1}{k^2}\,,$$

da cui si ricava che la serie considerata è convergente perché

$$\sum_{k=2}^{+\infty} \frac{1}{k(\log k)^2} = \lim_{n \to +\infty} \sum_{k=2}^{2^{n+1}-1} \frac{1}{k(\log k)^2} < \lim_{n \to +\infty} \frac{1}{(\log 2)^2} \sum_{k=1}^{n} \frac{1}{k^2} = T \in \mathbb{R}.$$

5.4.5 Criteri di convergenza per serie ad addendi positivi

Illustriamo ora alcuni criteri per stabilire immediatamente quale sia il comportamento della serie assegnata. Tali criteri rappresentano un metodo per creare un confronto fra la serie data e una serie avente comportamento noto.

Teorema 5.4.3. *(criterio del confronto asintotico)*
Date due serie di addendo generico a_k e b_k, se vale

$$\lim_{k \to +\infty} \frac{a_k}{b_k} = L > 0,$$

segue che le due serie

$$\sum_{k=0}^{\infty} a_k \qquad e \qquad \sum_{k=0}^{\infty} b_k$$

hanno il medesimo comportamento.

Dimostrazione. In virtù del limite, abbiamo che per un certo $\varepsilon > 0$ tale che valga $L - \varepsilon > 0$, esiste di conseguenza un k_ε tale che per ogni $k \geqslant k_\varepsilon$ si abbia

$$0 < L - \varepsilon < \frac{a_k}{b_k} < L + \varepsilon.$$

Se la serie di addendo b_k è convergente, allora utilizzando la disuguaglianza di destra $a_k < (L + \varepsilon)\, b_k$ si ottiene per $n > k_\varepsilon$

$$\lim_{n\to+\infty} \sum_{k=0}^{n} a_k = \sum_{k=0}^{k_\varepsilon - 1} a_k + \lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} a_k < \sum_{k=0}^{k_\varepsilon - 1} a_k + (L + \varepsilon) \lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} b_k\,,$$

da cui si ricava che la serie di addendo a_k è convergente perché, per la convergenza della serie di addendo b_k, si ha che è finito il limite

$$\lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} b_k\,.$$

Se invece la serie di addendo b_k è divergente, allora utilizzando la disuguaglianza di sinistra $a_k > (L - \varepsilon)\, b_k$ si ottiene per $n > k_\varepsilon$

$$\lim_{n\to+\infty} \sum_{k=0}^{n} a_k = \sum_{k=0}^{k_\varepsilon - 1} a_k + \lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} a_k > \sum_{k=0}^{k_\varepsilon - 1} a_k + (L - \varepsilon) \lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} b_k\,,$$

da cui si ricava che la serie di addendo a_k è divergente perché per la divergenza della serie di addendo b_k si ha che è infinito il limite

$$\lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} b_k, \quad \textit{c.d.d.}$$

Teorema 5.4.4. *(criterio della radice)*
Data una serie di addendo generico a_k, se vale

$$\lim_{k\to+\infty} \sqrt[k]{a_k} = L,$$

segue che la serie

$$\sum_{k=0}^{+\infty} a_k$$

è convergente se $L < 1$ ed è divergente se $L > 1$.

Dimostrazione. Se $L < 1$, allora scegliamo un $\varepsilon > 0$ che dia $L+\varepsilon < 1$ in corrispondenza del quale esisterà k_ε tale che per ogni $k \geqslant k_\varepsilon$ si abbia $\sqrt[k]{a_k} < L + \varepsilon$, ovvero $a_k < (L + \varepsilon)^k$. Di conseguenza si ottiene per $n > k_\varepsilon$

$$\lim_{n\to+\infty} \sum_{k=0}^{n} a_k = \sum_{k=0}^{k_\varepsilon - 1} a_k + \lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} a_k < \sum_{k=0}^{k_\varepsilon - 1} a_k + \lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} (L + \varepsilon)^k,$$

da cui si ricava che la serie di addendo a_k è convergente perché il limite

$$\lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} (L + \varepsilon)^k$$

risulta un numero finito essendo $L + \varepsilon < 1$.

Se invece vale $L > 1$, allora scegliamo un $\varepsilon > 0$ che dia $L - \varepsilon > 1$ in corrispondenza del quale esisterà k_ε tale che per ogni $k \geqslant k_\varepsilon$ si abbia $\sqrt[k]{a_k} > L - \varepsilon$, ovvero $a_k > (L-\varepsilon)^k$.

Di conseguenza si ottiene per $n > k_\varepsilon$

$$\lim_{n\to+\infty} \sum_{k=0}^{n} a_k = \sum_{k=0}^{k_\varepsilon - 1} a_k + \lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} a_k > \sum_{k=0}^{k_\varepsilon-1} a_k + \lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} (L-\varepsilon)^k,$$

da cui si ricava che la serie di addendo a_k è divergente perché il limite

$$\lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} (L-\varepsilon)^k$$

risulta infinito essendo $L - \varepsilon > 1$, *c.d.d.*

Teorema 5.4.5. *(criterio del rapporto)*

Data una serie di addendo generico a_k, se vale

$$\lim_{k\to+\infty} \frac{a_{k+1}}{a_k} = L,$$

segue che la serie

$$\sum_{k=0}^{+\infty} a_k$$

è convergente se $L < 1$ ed è divergente se $L > 1$.

Dimostrazione. Se $L < 1$, allora scegliamo un $\varepsilon > 0$ che dia $L+\varepsilon < 1$ in corrispondenza del quale esisterà k_ε tale che per ogni $k \geqslant k_\varepsilon$ si abbia

$$\frac{a_{k+1}}{a_k} < L + \varepsilon$$

ovvero $a_{k+1} < (L+\varepsilon)\, a_k$, da cui segue

$$a_k < (L+\varepsilon)\, a_{k-1} < (L+\varepsilon)^2\, a_{k-2} < \ldots < (L+\varepsilon)^{k-k_\varepsilon}\, a_{k_\varepsilon} = \frac{(L+\varepsilon)^k}{(L+\varepsilon)^{k_\varepsilon} a_{k_\varepsilon}}\,.$$

Di conseguenza si ottiene per $n > k_\varepsilon$

$$\lim_{n\to+\infty} \sum_{k=0}^{n} a_k = \sum_{k=0}^{k_\varepsilon-1} a_k + \lim_{n\to\infty} \sum_{k=k_\varepsilon}^{n} a_k < \sum_{k=0}^{k_\epsilon-1} a_k + \frac{1}{(L+\varepsilon)^{k_\varepsilon} a_{k_\varepsilon}} \lim_{n\to\infty} \sum_{k=k_\varepsilon}^{n} (L+\varepsilon)^k,$$

da cui si ricava che la serie di addendo a_k è convergente perché il limite

$$\lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} (L+\varepsilon)^k$$

risulta un numero finito essendo $L + \varepsilon < 1$.

Se invece risulta $L > 1$, allora scegliamo un $\varepsilon > 0$ che dia $L - \varepsilon > 1$ in corrispondenza del quale esisterà k_ε tale che per ogni $k \geqslant k_\varepsilon$ si abbia

$$\frac{a_{k+1}}{a_k} > L - \varepsilon,$$

ovvero $a_{k+1} > (L-\varepsilon)\, a_k$, da cui segue

$$a_k > (L-\varepsilon)\, a_{k-1} > (L-\varepsilon)^2\, a_{k-2} > \ldots > (L-\varepsilon)^{k-k_\varepsilon}\, a_{k_\varepsilon} = \frac{(L-\varepsilon)^k}{(L-\varepsilon)^{k_\varepsilon} a_{k_\varepsilon}}\,.$$

Di conseguenza si ottiene per $n > k_\varepsilon$

$$\lim_{n\to+\infty} \sum_{k=0}^{n} a_k = \sum_{k=0}^{k_\varepsilon -1} a_k + \lim_{n\to\infty} \sum_{k=k_\varepsilon}^{n} a_k > \sum_{k=0}^{k_\varepsilon -1} a_k + \frac{1}{(L-\varepsilon)^{k_\varepsilon} a_{k_\varepsilon}} \lim_{n\to\infty} \sum_{k=k_\varepsilon}^{n} (L-\varepsilon)^k,$$

da cui si ricava che la serie di addendo a_k è divergente perché il limite

$$\lim_{n\to+\infty} \sum_{k=k_\varepsilon}^{n} (L-\varepsilon)^k$$

risulta infinito essendo $L-\varepsilon > 1$, *c.d.d.*

Osserviamo che se nel criterio della radice e nel criterio del rapporto si ottiene un limite $L = 1$, allora tali criteri non permettono di concludere nulla sul comportamento della serie in oggetto. Sottolineiamo poi che non va fatta confusione fra il criterio del rapporto e il criterio del confronto asintotico: nel primo caso si deve studiare il rapporto fra l'addendo generico successivo a_{k+1} e l'addendo generico precedente a_k di una data serie; nel secondo caso si studia il rapporto fra l'addendo generico di una serie e l'addendo generico di un'altra serie. Concludiamo che per le due serie (5.16) i criteri della radice e del rapporto non forniscono nessuna informazione circa il comportamento della serie e, come va ripetuto in tutti i casi analoghi, abbiamo studiato direttamente la successione delle somme parziali.

Capitolo 6

Teoremi sugli estremi locali

Una funzione continua in un intervallo (a,b) si dice *derivabile* in un punto $x_0 \in (a,b)$ se esiste il limite

$$\lim_{x \longrightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{6.1}$$

e tale limite è un numero $\ell \in \mathbb{R}$. In tal caso il numero ℓ prende il nome di *derivata* della funzione $f(x)$ nel punto x_0 e la funzione $f(x)$ si dice *derivabile* nel punto x_0.

Sottolineiamo che quando la funzione $f(x)$ è derivabile nel punto x_0, il limite (6.1) risulta uguale al medesimo valore ℓ per $x \longrightarrow x_0$ sia da sinistra che da destra.

La derivata della funzione $f(x)$ viene indicata con $f'(x)$ e il valore della derivata nel punto x_0 viene indicato con $f'(x_0)$. Indicheremo quindi con $f''(x_0), f'''(x_0), f^{(iv)}(x_0)$, ecc. le derivate successive della $f(x)$ calcolate nel punto x_0. Un punto x_0 viene denominato *punto stazionario* se in x_0 risulta $f'(x_0) = 0$. Una funzione si dice *continua in un intervallo* se è continua in tutti i punti di un intervallo e analogamente si dice *derivabile in un intervallo* se è derivabile in tutti i punti dell'intervallo considerato.

Come esempio di applicazione del limite (6.1), scriviamo il *rapporto incrementale* della funzione $f(x) = x^2 e^{-2x^2}$ nel generico punto x_0 e calcoliamo il suo limite per $x \to x_0$.

Abbiamo

$$\lim_{x \longrightarrow x_0} \frac{x^2 e^{-2x^2} - x_0^2 e^{-2x_0^2}}{x - x_0} = \lim_{x \longrightarrow x_0} \frac{\left(x^2 e^{-2x^2} - x_0^2 e^{-2x^2}\right) + \left(x_0^2 e^{-2x^2} - x_0^2 e^{-2x_0^2}\right)}{x - x_0} =$$

$$= \lim_{x \longrightarrow x_0} \left[\frac{x^2 e^{-2x^2} - x_0^2 e^{-2x^2}}{x - x_0} + \frac{x_0^2 e^{-2x^2} - x_0^2 e^{-2x_0^2}}{x - x_0} \right] = \lim_{x \longrightarrow x_0} \frac{x^2 - x_0^2}{x - x_0} e^{-2x^2} + x_0^2 \lim_{x \longrightarrow x_0} \frac{e^{-2x^2} - e^{-2x_0^2}}{x - x_0} =$$

$$= \lim_{x \longrightarrow x_0} \frac{(x - x_0)(x + x_0)}{x - x_0} e^{-2x^2} + x_0^2 e^{-2x_0^2} \lim_{x \longrightarrow x_0} \frac{e^{-2(x^2 - x_0^2)} - 1}{-2(x^2 - x_0^2)} \left[-2(x + x_0)\right] =$$

$$= 2x_0 e^{-2x_0^2} - 4x_0^3 e^{-2x_0^2}.$$

6.1 Retta tangente e differenziale

Se la $f(x)$ è derivabile in $x_0 \in (a,b)$, la retta avente equazione cartesiana

$$y = f(x_0) + f'(x_0)(x - x_0)$$

prende il nome di *retta tangente* alla $f(x)$ nel punto $P(x_0, f(x_0))$ ed è immediato rendersi conto che per il punto P passano sia il grafico della funzione che quello della retta tangente.

Indichiamo ora con $\Delta f(x_0)$ la variazione $\Delta f(x_0) = f(x) - f(x_0)$ che la funzione subisce quando si passa dal punto x_0 al punto x e indichiamo con $df(x_0)$ il differenziale della $f(x)$ in x_0 coincidente con la variazione che l'ordinata della retta tangente subisce quando si passa dal punto x_0 al punto x. L'espressione del differenziale è dunque

$$df(x_0) = \Big[f(x_0) + f'(x_0)(x - x_0) \Big] - f(x_0) = f'(x_0)(x - x_0) \tag{6.2}$$

e verifica la proprietà fondamentale stabilita dal seguente limite

$$\begin{aligned} \lim_{x \to x_0} \frac{\Delta f(x_0) - df(x_0)}{x - x_0} &= \lim_{x \to x_0} \frac{[f(x) - f(x_0)] - f'(x_0)\,(x - x_0)}{x - x_0} = \\ &= \lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} - \frac{f'(x_0) \cdot (x - x_0)}{x - x_0} \right] = \\ &= \left[\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \right] - f'(x_0) = f'(x_0) - f'(x_0) = 0, \end{aligned}$$

ovvero in un punto x_0 il differenziale $df(x_0)$ differisce dalla variazione $\Delta f(x_0)$ per un infinitesimo di ordine superiore a 1, perché al denominatore vi è $x - x_0$ elevato a 1.

Se poniamo $R_1(x) = f(x) - [f(x_0) + f'(x_0)(x - x_0)]$, per cui $R_1(x)$ è la differenza (o *resto di ordine 1*), in un punto x di un intorno di x_0, tra il valore di $f(x)$ e il corrispondente valore calcolato sulla tangente, possiamo anche scrivere

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + R_1(x) = \mathcal{P}_1(x) + R_1(x), \tag{6.3a}$$

avendo chiamato $\mathcal{P}_1(x)$ il polinomio di primo grado

$$\mathcal{P}_1(x) = f(x_0) + f'(x_0)(x - x_0) \tag{6.3b}$$

che esprime l'equazione della tangente. La relazione (6.3a) può essere utilizzata per approssimare, nell'intorno del punto x_0, la funzione $f(x)$ con la sua tangente, ovvero

$$f(x) \simeq \mathcal{P}_1(x) = f(x_0) + f'(x_0)(x - x_0), \tag{6.4}$$

perché con l'approssimazione (6.4) si commette un errore $R_1(x)$ che risulta infinitesimo di ordine superiore a 1. Infatti, se esplicitiamo $R_1(x)$ tramite le (6.3) e calcoliamo

$$\lim_{x \to x_0} \frac{R_1(x)}{x - x_0} = \lim_{x \to x_0} \frac{f(x) - \mathcal{P}_1(x)}{x - x_0} = \lim_{x \to x_0} \frac{\Delta f(x_0) - df(x_0)}{x - x_0} = 0, \tag{6.5}$$

otteniamo appunto che al tendere di x ad x_0 l'errore $R_1(x)$ che si commette con l'approssimazione (6.4), cioè utilizzando il polinomio $\mathcal{P}_1(x)$ al posto della $f(x)$, tende a zero *più velocemente* di $x - x_0$ e quindi per x abbastanza vicino ad x_0 esso è trascurabile.

In modo del tutto analogo, il limite al terzo membro della (6.5) permette di effettuare anche un'altra approssimazione, ovvero si può utilizzare, come si fa "spesso" nella teoria economica, il differenziale df al posto della variazione Δf, o viceversa.

6.2 Estremi locali o relativi

Data una funzione continua in un intervallo (a,b), un punto $x_0 \in (a,b)$ viene denominato *minimo relativo* oppure *minimo locale* della funzione $f(x)$ se esiste un intorno I del punto x_0 ("opportunamente piccolo") tale che risulti $f(x) - f(x_0) \geqslant 0$, per ogni $x \in I$.

Un punto $x_0 \in (a,b)$ viene chiamato *massimo relativo* oppure *massimo locale* della funzione $f(x)$ se esiste un intorno I di x_0 tale che risulti $f(x) - f(x_0) \leqslant 0$, per ogni $x \in I$.

I valori (cioè i risultati) massimi e minimi (relativi o assoluti) della $f(x)$ vengono anche chiamati *estremi* della $f(x)$. Un punto x_0 si dice *estremante* della $f(x)$ se $f(x_0)$ è un estremo della $f(x)$. Quindi, riassumendo abbiamo che se un punto $P(x_0, f(x_0))$ è un punto di massimo o minimo (locale o assoluto) della $f(x)$, l'ordinata $f(x_0)$ e l'ascissa x_0 del punto P vengono chiamate rispettivamente *estremo* ed *estremante* della $f(x)$.

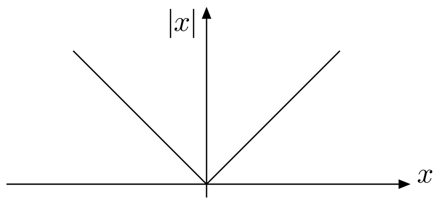
Teorema 6.2.1. *(condizione necessaria per l'estremo)*

Se una funzione $f(x)$ è continua e derivabile in un intorno di un punto x_0 e il punto x_0 è un estremante locale per la $f(x)$, segue che risulta $f'(x_0) = 0$ (ovvero un punto estremante locale per una funzione derivabile in un intorno del punto stesso risulta necessariamente stazionario per la medesima funzione).

Questo teorema può essere equivalentemente espresso dicendo che se in un punto x_0 risulta $f'(x_0) \neq 0$, allora il punto x_0 non è un punto estremante per la $f(x)$.

Riguardo al teorema 6.2.1 è importante sottolineare il ruolo cruciale dell'ipotesi di derivabilità della $f(x)$ in un intorno del punto estremante x_0 perché se un punto x_0 è estremante per una $f(x)$, ma la $f(x)$ non fosse derivabile in x_0, segue che non avrebbe senso parlare di derivata della $f(x)$ in x_0, né quindi del teorema 6.2.1, come mostra, ad esempio, la funzione $f(x) = |x|$ riportata nella seguente figura

$$f(x) = |x| = \begin{cases} -x & \text{se} \quad x \leqslant 0 \\ x & \text{se} \quad x \geqslant 0 \end{cases}$$



che ha un minimo nel punto $x_0 = 0$ in cui però la $f(x)$ non è derivabile e in cui quindi non si annulla la derivata (che in effetti non esiste!).

Se invertiamo ipotesi e tesi del teorema 6.2.1 non otteniamo un altro teorema valido perché la funzione potenza $f(x) = x^3$ è un *controesempio* che mostra che sebbene nel punto $x_0 = 0$ si annulli la derivata prima $f'(x) = 3x^2$ della funzione $f(x)$, tuttavia il punto $x_0 = 0$ non è un punto estremante, come si riconosce nella fig. 4.9 alla fine del par. 4.2.2.

Teorema 6.2.2. *(di Rôlle)*

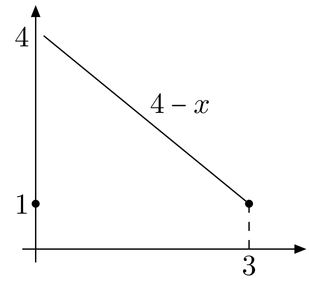
Se una funzione $f(x)$ è continua nell'intervallo chiuso $[a,b]$ e derivabile nell'intervallo aperto (a,b) e se inoltre risulta $f(a) = f(b)$, allora segue che esiste un punto $c \in (a,b)$ tale che $f'(c) = 0$.

Per le dimostrazioni dei teoremi 6.2.1 e 6.2.2, rimandiamo al testo di riferimento.

Riguardo al teorema di Rôlle, è istruttivo discutere brevemente il ruolo dell'ipotesi di continuità della funzione $f(x)$ nei punti a, b di frontiera dell'intervallo $[a,b]$.

Presentiamo un esempio dal quale si deduce che senza tale ipotesi, ovvero se la funzione $f(x)$ non fosse continua in uno dei due punti di bordo dell'intervallo $[a,b]$, non si otterrebbe necessariamente la tesi, ovvero il teorema di Rôlle non sarebbe valido.

Se consideriamo la funzione



$$f(x) = \begin{cases} 4 - x & \text{per} \quad x \in (0,3] \\ 1 & \text{per} \quad x = 0 \end{cases}$$

nell'intervallo $[0,3]$, il cui grafico è quello riportato nella figura qui accanto, abbiamo che essa risulta continua in $(0,3]$, ma non nell'estremo sinistro $x = 0$ perché il limite della funzione per la variabile x tendente a zero da destra è 4 che è diverso da $f(0) = 1$.

La funzione è inoltre derivabile nell'intervallo $(0,3)$ e verifica l'ipotesi $f(0) = f(3) = 1$, ma poiché la sua derivata è $f'(x) = -1$ per ogni $x \in (0,3)$, concludiamo che non esiste nessun punto $c \in (0,3)$ in cui valga $f'(c) = 0$, ovvero, come si vede anche chiaramente nella figura, non esiste nessun punto della funzione con tangente orizzontale.

Questa funzione non costituisce una violazione del teorema di Rôlle perché a tale funzione, che non è continua in $x = 0$, non soddisfa tutte le ipotesi del teorema di Rôlle.

Tale $f(x)$ mostra dunque l'importanza cruciale dell'ipotesi di continuità della funzione al bordo ai fini del teorema di Rôlle, perché quand'anche valga l'uguaglianza $f(a) = f(b)$, la funzione considerata in questo esempio mostra che senza continuità in un estremo dell'intervallo, può non esistere il punto $c \in (a,b)$ in cui risulti $f'(c) = 0$.

Se consideriamo solo funzioni continue nell'intervallo chiuso $[a,b]$ e derivabili nell'intervallo aperto (a,b), possiamo enunciare la formulazione equivalente del teorema di Rôlle scambiando ipotesi e tesi e scrivendo la proposizione contraria di entrambe, ovvero

Teorema 6.2.3. *(di Rôlle equivalente)*

Data una funzione $f(x)$ continua nell'intervallo chiuso $[a,b]$ e derivabile nell'intervallo aperto (a,b), se non esiste nessun punto $c \in (a,b)$ tale che $f'(c) = 0$, ovvero se vale la condizione $f'(x) \neq 0$ per ogni $x \in (a,b)$, allora segue che risulta $f(a) \neq f(b)$.

Teorema 6.2.4. *(di Cauchy)*

Se le due funzioni $f(x)$ e $g(x)$ sono continue in $[a,b]$ e derivabili in (a,b), con $g'(x) \neq 0$ per ogni $x \in (a,b)$, segue che esiste almeno un punto $c \in (a,b)$ tale che

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)} . \tag{6.6}$$

Dimostrazione. Osserviamo innanzitutto che dall'ipotesi $g'(x) \neq 0$ per ogni x appartenente all'intervallo (a,b) segue, in virtù del teorema di Rôlle equivalente 6.2.3, la disuguaglianza $g(a) \neq g(b)$. Se consideriamo la funzione "ausiliaria"

$$H(x) = \Big[f(b) - f(a) \Big] g(x) - \Big[g(b) - g(a) \Big] f(x),$$

abbiamo che essa è continua in $[a,b]$ e derivabile in (a,b) in quanto combinazione lineare delle due funzioni $f(x)$ e $g(x)$ appunto continue in $[a,b]$ e derivabili in (a,b); inoltre si riconosce che la funzione $H(x)$ soddisfa la condizione $H(a) = H(b) = f(b)g(a) - f(a)g(b)$.

Poiché dunque la funzione $H(x)$ verifica le ipotesi del teorema di Rôlle, esiste un punto $c \in (a,b)$ in cui si annulla la sua derivata $H'(x) = [f(b) - f(a)]\, g'(x) - [g(b) - g(a)]\, f'(x)$, ovvero

$$H'(c) = [f(b) - f(a)]\, g'(c) - [g(b) - g(a)]\, f'(c) = 0$$

da cui, poiché risulta $g(a) \neq g(b)$, segue appunto la tesi

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)},$$

come dovevasi dimostrare (cdd).

Esercizio. Date le due funzioni

$$f(x) = \log(2x+1) \qquad \text{e} \qquad g(x) = \frac{2}{x},$$

verificare se esse soddisfano le ipotesi del *teorema di Cauchy* nell'intervallo $[1,2]$ e, in caso affermativo, determinare un punto $c \in (1,2)$ la cui esistenza è assicurata dal teorema.

Soluzione. Si riconosce immediatamente che le due funzioni $f(x), g(x)$ sono continue e derivabili in tutto l'intervallo $[1,2]$ con $g'(x) \neq 0$ per ogni $x \in (1,2)$ e allora scriviamo l'equazione della tesi del *teorema di Cauchy* con

$$f(1) = \log 3, \quad f(2) = \log 5, \quad g(1) = 2, \quad g(2) = 1, \quad f'(x) = \frac{2}{2x+1}, \quad g'(x) = -\frac{2}{x^2},$$

da cui segue l'equazione della tesi del *teorema di Cauchy* nella forma

$$-\frac{c^2}{2c+1} = \frac{\log 5 - \log 3}{1-2}.$$

Pertanto, se poniamo $\log(5/3) \equiv A > 0$, otteniamo il punto c risolvendo l'equazione

$$\frac{c^2}{2c+1} = \log(5/3) \quad \Longrightarrow \quad c^2 - 2Ac - A = 0 \quad \Longrightarrow$$

$$\Longrightarrow \quad c_+ = A + \sqrt{A^2 + A} = \log(5/3) + \sqrt{\log^2(5/3) + \log(5/3)},$$

dove la seconda soluzione $c_- = A - \sqrt{A^2+A} = \log(5/3) - \sqrt{\log^2(5/3) + \log(5/3)}$ non è un punto c del *teorema di Cauchy* perché il numero $c_- = \log(5/3) - \sqrt{\log^2(5/3) + \log(5/3)}$ è negativo e quindi non appartiene all'intervallo $(1,2)$.

Teorema 6.2.5. *(di Lagrange)*
Se $f(x)$ è continua in $[a,b]$ e derivabile in (a,b), allora esiste un punto $c \in (a,b)$ tale che:

$$\frac{f(b)-f(a)}{b-a} = f'(c). \tag{6.7}$$

Dimostrazione del teorema di Lagrange. Le due funzioni $f(x)$ e $g(x) = x$ verificano le ipotesi del teorema di Cauchy perché $g'(x) = 1 \neq 0$ per ogni x e quindi, poiché $g(a) = a$ e $g(b) = b$, abbiamo immediatamente che la relazione (6.6) del teorema di Cauchy diventa la relazione (6.7) del teorema di Lagrange, *cdd.*

Osserviamo che la relazione (6.7) del teorema di Lagrange ha una semplice interpretazione geometrica: il primo membro della relazione (6.7) rappresenta la pendenza della retta secante alla curva $y = f(x)$ nei punti di ascissa a e b, mentre $f'(c)$ rappresenta la pendenza della retta tangente nel punto di ascissa c. Quindi la relazione (6.7) afferma che per una funzione $f(x)$ continua nell'intervallo chiuso $[a,b]$ e derivabile nell'intervallo aperto (a,b), esiste almeno un punto della curva, interno all'intervallo, in cui la retta tangente alla curva è parallela alla retta secante che passa per gli estremi dell'intervallo.

Utilizzando il teorema di Lagrange possiamo dimostrare il teorema che mette in relazione il comportamento crescente o decrescente di una funzione in un intervallo con il segno della derivata prima della funzione nello stesso intervallo.

Esercizio. Data la funzione

$$f(x) = -2x + \log(2x+1)\,,$$

verificare se essa soddisfa le ipotesi del *teorema di Lagrange* nell'intervallo $[0,1]$ e, in caso affermativo, determinare un punto $c \in (0,1)$ la cui esistenza è assicurata dal teorema.

Soluzione. Si riconosce immediatamente che la funzione $f(x)$ è continua e derivabile in tutto l'intervallo $[0,1]$ e allora scriviamo l'equazione della tesi del *teorema di Lagrange* con

$$f(0) = 0, \quad f(1) = -2 + \log 3, \quad f'(x) = -2 + \frac{2}{2x+1} = -\frac{4x}{2x+1}$$

da cui segue l'equazione della tesi del *teorema di Lagrange* nella forma

$$-\frac{4c}{2c+1} = -2 + \log 3.$$

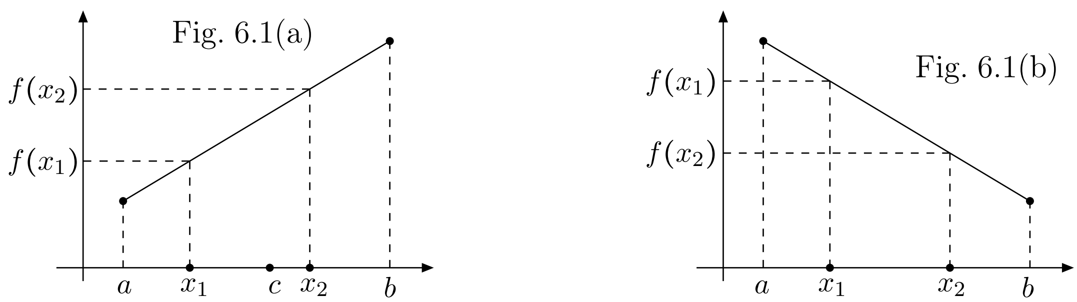
Pertanto, se poniamo $A \equiv 1/(2 - \log 3) > 0$, otteniamo il punto c risolvendo l'equazione

$$\frac{2c+1}{4c} = A \implies 2c + 1 = 4Ac \implies (4A-2)c = 1 \implies$$

$$\implies c = \frac{1}{4A-2} \implies c = \frac{2 - \log 3}{\log 9} \in (0,1).$$

Teorema 6.2.6. *Data una funzione contiua e derivabile in (a,b), se risulta $f'(x) > 0$ per ogni $x \in (a,b)$, segue che la funzione è crescente in (a,b).*

Dimostrazione. Per dimostrare che la $f(x)$ è crescente in (a,b), occorre dimostrare che per ogni $x_1, x_2 \in (a,b)$ si ottiene la prima delle due disuguaglianze (4.10).



Poiché la $f(x)$ è continua e derivabile nell'intervallo aperto (a,b), se dunque x_1, x_2 sono due generici punti appartenenti all'intervallo (a,b), la funzione $f(x)$ è continua e derivabile nell'intervallo chiuso $[x_1, x_2]$ e quindi per il teorema di Lagrange segue che esiste un punto c compreso tra x_1 e x_2, mostrato in fig. 6.1(a), ovvero tra a e b, tale che

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c). \tag{6.8}$$

Ma se la derivata prima $f'(x) > 0$ è positiva in tutti i punti $x \in (a, b)$, segue che la derivata $f'(c)$ è positiva e quindi dal confronto tra la (4.8) e la (6.8) si deduce che il *rapporto incrementale* (4.8) è positivo come $f'(c)$ al secondo membro della (6.8), da cui si conclude che vale la disuguaglianza (4.10a), ovvero che la $f(x)$ verifica la definizione di funzione crescente in (a, b), *cdd.*

In modo identico si dimostra che se risulta $f'(x) < 0$ per ogni $x \in (a, b)$, allora la funzione $f(x)$ è decrescente nell'intervallo (a, b), perché dall'uguaglianza (6.8) si deduce che il *rapporto incrementale* (4.8) è negativo, ovvero che la $f(x)$ verifica la definizione di funzione decrescente in (a, b), espressa mediante la disuguaglianza (4.10b).

Utilizzando il teorema di Cauchy si può dimostrare anche il seguente teorema che possiamo chiamare *teorema di Cauchy con le derivate successive.*

Teorema 6.2.7. *(teorema di Cauchy con le derivate successive)*

Se le due funzioni $F(x)$ e $G(x)$ sono continue e derivabili $n + 1$ volte con derivate continue in un intorno I di x_0 e inoltre verificano le relazioni con le derivate successive fino all'ordine n

$$F(x_0) = F'(x_0) = F''(x_0) = \cdots = F^{(n-2)}(x_0) = F^{(n-1)}(x_0) = F^{(n)}(x_0) = 0,$$

$$G(x_0) = G'(x_0) = G''(x_0) = \cdots = G^{(n-2)}(x_0) = G^{(n-1)}(x_0) = G^{(n)}(x_0) = 0,$$

con $G(x), G'(x), G''(x), \ldots, G^{(n-1)}, G^{(n)}(x), G^{(n+1)}(x)$ diverse da zero per ogni $x \in I$ diverso da x_0, allora per $x \in I$ esiste un punto c compreso tra x e x_0 tale che valga

$$\frac{F(x)}{G(x)} = \frac{F^{(n+1)}(c)}{G^{(n+1)}(c)}. \tag{6.9}$$

Dimostrazione. In virtù del teorema di Cauchy possiamo scrivere

$$\frac{F(x)}{G(x)} = \frac{F(x) - F(x_0)}{G(x) - G(x_0)} = \frac{F'(c_1)}{G'(c_1)}, \tag{6.10a}$$

con c_1 compreso tra x e x_0 ed essendo state effettuate al secondo membro le due sostituzioni $F(x) = F(x) - F(x_0)$ e $G(x) = G(x) - G(x_0)$ perché per ipotesi $F(x_0) = G(x_0) = 0$.

Applicando ripetutamente il teorema di Cauchy, otteniamo analogamente

$$\frac{F'(c_1)}{G'(c_1)} = \frac{F'(c_1) - F'(x_0)}{G'(c_1) - G'(x_0)} = \frac{F''(c_2)}{G''(c_2)}, \tag{6.10b}$$

con c_2 compreso tra c_1 e x_0 e quindi

$$\frac{F''(c_2)}{G''(c_2)} = \frac{F''(c_2) - F''(x_0)}{G''(c_2) - G''(x_0)} = \frac{F'''(c_3)}{G'''(c_3)}, \tag{6.10c}$$

con c_3 compreso tra c_2 e x_0, fino a

$$\frac{F^{(n)}(c_n)}{G^{(n)}(c_n)} = \frac{F^{(n)}(c_n) - F^{(n)}(x_0)}{G^{(n)}(c_n) - G^{(n)}(x_0)} = \frac{F^{(n+1)}(c)}{G^{(n+1)}(c)}, \tag{6.10d}$$

con c_n, c compresi tra x e x_0.

Sostituendo "*a ritroso*" la (6.10d) nella precedente e così via fino a sostituire la (6.10c) nella (6.10b), la (6.10b) nella (6.10a), otteniamo infine la sequenza di uguaglianze

$$\frac{F(x)}{G(x)} = \frac{F'(c_1)}{G'(c_1)} = \frac{F''(c_2)}{G''(c_2)} = \frac{F'''(c_3)}{G'''(c_3)} = \cdots = \frac{F^{(n-1)}(c_{n-1})}{G^{(n-1)}(c_{n-1})} = \frac{F^{(n)}(c_n)}{G^{(n)}(c_n)} = \frac{F^{(n+1)}(c)}{G^{(n+1)}(c)},$$

con tutti i punti $c_1, c_2, c_3, \ldots, c_{n-1}, c_n, c$ compresi tra x e x_0, da cui, prendendo primo e ultimo termine, ricaviamo l'uguaglianza (6.9), con c compreso tra x e x_0, *cdd.*

6.3 Approssimazioni di ordine superiore

La (6.3a) può essere generalizzata in modo da utilizzare, per approssimare la $f(x)$, invece della retta tangente (polinomio $\mathcal{P}_1(x)$ di primo grado), una parabola (polinomio di secondo grado), una cubica (polinomio di terzo grado), e così via, cioè in generale un polinomio $\mathcal{P}_n(x)$ di grado n. Se $f(x)$ è derivabile n volte con derivate continue in un intorno I di x_0, definiamo un polinomio $\mathcal{P}_n(x)$ di grado (o ordine) n relativo alla $f(x)$ nel punto x_0 tale che, in analogia con la (6.5), valga

$$\lim_{x\to x_0} \frac{f(x)-\mathcal{P}_n(x)}{(x-x_0)^n} = 0, \tag{6.11}$$

ovvero differisca dalla $f(x)$ nel punto x_0 per un infinitesimo di ordine superiore a n.

In questo caso generalizziamo la relazione (6.3a) ponendo

$$f(x) = \mathcal{P}_n(x) + R_n(x), \tag{6.12}$$

ovvero $R_n(x) = f(x) - \mathcal{P}_n(x)$, in modo che la (6.11) possa essere riscritta nella forma

$$\lim_{x\to x_0} \frac{R_n(x)}{(x-x_0)^n} = 0 \tag{6.13}$$

che mostra, appunto, che l'errore che si commette approssimando la $f(x)$ tramite il polinomio $\mathcal{P}_n(x)$ è infinitesimo di ordine superiore ad n, cioè superiore al grado del polinomio. La (6.12), che per $n = 1$ coincide ovviamente con la (6.3a), prende il nome di **formula di Taylor** in cui il polinomio $\mathcal{P}_n(x)$ viene denominato *polinomio di Taylor* di ordine[1] n relativo alla $f(x)$ nel punto x_0 e $R_n(x)$ prende invece il nome di *resto* $n-$*esimo* della formula di Taylor.

Teorema 6.3.1. *(Formula di Taylor)*

Scelto un qualsiasi numero naturale $n = 0, 1, 2, 3, \dots,$ *se* $f(x)$ *è derivabile* $n+1$ *volte con derivate continue in un intorno* I *di* x_0 *e* $x \in I$, *allora esiste un punto* c *compreso nell'intervallo* (x, x_0) *oppure nell'intervallo* (x_0, x) *tale che valga l'uguaglianza*

$$f(x) = \mathcal{P}_n(x) + R_n(x),$$

denominata "formula di Taylor", dove $\mathcal{P}_n(x)$ *indica il cosiddetto "polinomio di Taylor" e* $R_n(x)$ *indica il cosiddetto "resto* $n-$*esimo" che assumono rispettivamente la forma*

$$\begin{aligned}\mathcal{P}_n(x) = {} & f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \frac{f'''(x_0)}{3!}(x-x_0)^3 + \\ & + \frac{f^{(IV)}(x_0)}{4!}(x-x_0)^4 + \dots + \frac{f^{(n-1)}(x_0)}{(n-1)!}(x-x_0)^{n-1} + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\end{aligned} \tag{6.14a}$$

e

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}. \tag{6.14b}$$

[1]Non è detto che il grado di $\mathcal{P}_n(x)$ sia proprio n. Infatti $f^{(n)}(x_0)$ (ed altre derivate di ordine inferiore) potrebbe essere uguale a zero, per cui $\mathcal{P}_n(x)$ avrebbe grado inferiore ad n. Per questo parliamo di *ordine* e non di grado del polinomio di Taylor.

Prima di dare la dimostrazione di questo teorema, osserviamo che in genere non si conosce il valore del punto c, ma si sa solo che esso esiste, e quindi la (6.14b) non serve per calcolare il valore esatto di $R_n(x)$ ma permette "spesso" di stimarne l'ordine di grandezza, come mostrato più avanti nell'Esempio 2.

Dimostrazione (del teorema sulla formula di Taylor). Consideriamo le due funzioni

$$F(x) = f(x) - f(x_0) - f'(x_0)(x - x_0) - \frac{f''(x_0)}{2!}(x - x_0)^2 - \frac{f'''(x_0)}{3!}(x - x_0)^3 +$$

$$- \frac{f^{(IV)}(x_0)}{4!}(x - x_0)^4 - \ldots - \frac{f^{(n-1)}(x_0)}{(n-1)!}(x - x_0)^{n-1} - \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

e $G(x) = (x - x_0)^{n+1}$ e dimostriamo che esse verificano le ipotesi del teorema di Cauchy con le derivate successive. La verifica per la $G(x)$ è immediata dal momento che la funzione soddisfa $G(x_0) = 0$ e inoltre le derivate fino all'ordine n

$$G'(x) = (n+1)\,(x - x_0)^n, \qquad G''(x) = (n+1)n\,(x - x_0)^{n-1},$$

$$G'''(x) = (n+1)n(n-1)\,(x - x_0)^{n-2}, \ldots, G^{(n)}(x) = (n+1)!\,(x - x_0)$$

si annullano tutte per $x = x_0$ perché si annullano tutti i binomi $x - x_0$.

Vale inoltre la relazione con la derivata $(n+1) - esima$

$$G^{(n+1)}(x) = (n+1)! = G^{(n+1)}(c).$$

Per la $F(x)$ abbiamo $F(x_0) = 0$ e che anche le derivate fino all'ordine n, come si vede, si annullano tutte per $x = x_0$ perché si ha

$$F'(x) = f'(x) - f'(x_0) - f''(x_0)\,(x - x_0) - \frac{f'''(x_0)}{2!}(x - x_0)^2 - \frac{f^{(IV)}(x_0)}{3!}(x - x_0)^3 +$$

$$- \frac{f^{(V)}(x_0)}{4!}(x - x_0)^4 - \cdots - \frac{f^{(n-1)}(x_0)}{(n-2)!}(x - x_0)^{n-2} - \frac{f^{(n)}(x_0)}{(n-1)!}(x - x_0)^{n-1},$$

$$F''(x) = f''(x) - f''(x_0) - f'''(x_0)\,(x - x_0) - \frac{f^{(IV)}(x_0)}{2!}(x - x_0)^2 - \frac{f^{(V)}(x_0)}{3!}(x - x_0)^3 +$$

$$- \cdots - \frac{f^{(n-1)}(x_0)}{(n-3)!}(x - x_0)^{n-3} - \frac{f^{(n)}(x_0)}{(n-2)!}(x - x_0)^{n-2},$$

$$F'''(x) = f'''(x) - f'''(x_0) - f^{(IV)}(x_0)\,(x - x_0) - \frac{f^{(V)}(x_0)}{2!}(x - x_0)^2 +$$

$$- \cdots - \frac{f^{(n-1)}(x_0)}{(n-4)!}(x - x_0)^{n-4} - \frac{f^{(n)}(x_0)}{(n-3)!}(x - x_0)^{n-3}, \quad \cdots$$

$$F^{(n-1)}(x) = f^{(n-1)}(x) - f^{(n-1)}(x_0) - f^{(n)}(x_0)\,(x - x_0),$$

$$F^{(n)}(x) = f^{(n)}(x) - f^{(n)}(x_0).$$

Infine vale la relazione relativa alla derivata $(n+1) - esima$

$$F^{(n+1)}(c) = f^{(n+1)}(c).$$

Poiché dunque la $F(x)$ e la $G(x)$ verificano le ipotesi del teorema di Cauchy con le derivate $(n+1)-esime$, esiste un punto c compreso tra x e x_0 tale che valga la (6.9) con le derivate $(n+1)-esime$, ovvero

$$F(x) = \frac{F^{(n+1)}(c)}{G^{(n+1)}(c)} G(x), \tag{6.15}$$

che con le $F(x), G(x)$ considerate e con le sostituzioni

$$G^{(n+1)}(c) = (n+1)! \qquad \text{e} \qquad F^{(n+1)}(c) = f^{(n+1)}(c)$$

fornisce appunto la formula di Taylor $f(x) = \mathcal{P}_n(x) + R_n(x)$, con $\mathcal{P}_n(x)$ dato dalla (6.14a) e $R_n(x)$ dato dalla (6.14b), *cdd.*

E' istruttivo osservare l'analogia tra il teorema di Lagrange e il teorema sulla formula di Taylor in modo che appaia più evidente in che senso il teorema sulla formula di Taylor è la generalizzazione del teorema di Lagrange. Poiché la relazione (6.7) del teorema di Lagrange può essere riscritta nella forma $f(b) = f(a) + f'(c)(b-a)$, ovvero nella forma

$$f(x) = f(x_0) + f'(c)(x - x_0), \tag{6.16}$$

in cui siano identificati $x = b$ e $x_0 = a$, possiamo enunciare il teorema di Lagrange dicendo che se una funzione è continua e derivabile in un intorno I di x_0 e $x \in I$, allora esiste un punto c compreso tra x e x_0 in cui vale $f(x) = f(x_0) + f'(c)(x - x_0)$, ovvero in cui vale la relazione (6.12) con $n = 0$, perché il polinomio di Taylor $\mathcal{P}_0(x)$ di grado zero è rappresentato dalla sola costante $f(x_0)$. A questo punto è immediato rendersi conto che il polinomio di Taylor (6.14a) e il resto n-esimo (6.14b) sono la generalizzazione dei termini rispettivamente $f(x_0)$ e $f'(c)(x - x_0)$ del teorema di Lagrange nella forma (6.16).

Può essere utile sottolineare inoltre l'analogia tra il teorema di Lagrange e il teorema sulla formula di Taylor anche dal punto di vista delle loro rispettive dimostrazioni: così come il teorema di Lagrange si ottiene dal teorema di Cauchy, analogamente e *mutatis mutandis* il teorema sulla formula di Taylor si ottiene dal teorema di Cauchy con le derivate successive. E' immediato verificare *a posteriori* che il polinomio $\mathcal{P}_n(x)$ e il resto $R_n(x)$ della formula di Taylor soddisfano per ogni n le relazioni (6.11), ovvero le relazioni (6.13), che sono una generalizzazione dei primi due limiti (6.5). Si ha infatti

$$\begin{aligned} &\lim_{x \longrightarrow x_0} \frac{f(x) - \mathcal{P}_n(x)}{(x - x_0)^n} = \lim_{x \longrightarrow x_0} \frac{R_n(x)}{(x - x_0)^n} = \\ &= \lim_{x \longrightarrow x_0} \frac{f^{(n+1)}(c)}{(n+1)!} \cdot \frac{(x - x_0)^{n+1}}{(x - x_0)^n} = \lim_{x \longrightarrow x_0} \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x - x_0) = 0. \end{aligned} \tag{6.17}$$

Utilizzando il limite (6.17) possiamo dimostrare immediatamente il seguente

Teorema 6.3.2. *Esiste un intorno del punto x_0 in cui il segno della somma dei due termini*

$$\frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + R_n(x) \tag{6.18}$$

coincide con il segno del solo primo termine.

Dimostrazione. Infatti, se esprimiamo la somma (6.18) tramite l'identità

$$\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_n(x) = \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\left[1+\frac{n!}{f^{(n)}(x_0)}\,\frac{R_n(x)}{(x-x_0)^n}\right], \tag{6.19}$$

abbiamo che per x tendente a x_0, in virtù del limite (6.17), la parentesi quadra al secondo membro della (6.19) tende al numero positivo 1. Allora per il teorema della *permanenza del segno*, esiste un intorno di x_0 in cui tutta la parentesi quadra risulta positiva, da cui segue che il segno del primo membro della (6.19) dipende solo dal termine davanti alla parentesi quadra, ovvero solo dal primo termine della somma (6.18), *cdd.*

Nel caso $x_0 = 0$, la (6.14a) diventa:

$$f(x) = f(0) + f'(0)\,x + \frac{f''(0)}{2!}\,x^2 + \ldots + \frac{f^{(n)}(0)}{n!}\,x^n + R_n(x), \tag{6.20}$$

denominata anche *formula di Mc Laurin*. Affinché possa essere più chiaro il significato della *formula di Taylor*, riportiamo le sua espressione in cui il polinomio scritto nella prima parentesi quadra abbia grado rispettivamente per $n = 0, 1, 2$

$$f(x) = [f(x_0)] + [f'(c)(x-x_0)] \qquad \text{(coincidente con il teorema di lagrange)},$$

$$f(x) = [f(x_0) + f'(x_0)(x-x_0)] + \left[\frac{f''(c)}{2!}(x-x_0)^2\right], \tag{6.21a}$$

$$f(x) = \left[f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2\right] + \left[\frac{f'''(c)}{3!}(x-x_0)^3\right]. \tag{6.21b}$$

Esercizio 1. Data la funzione $f(x) = \log(3x+2)$, verificare se essa soddisfa le ipotesi del teorema della formula di Taylor relativamente ai punti $x_0 = -1/3$ e $x = 0$.

In caso affermativo scrivere la formula di Taylor associata alla funzione $f(x)$ relativamente ai punti dati e contenente il polinomio di Taylor sviluppato fino all'ordine $n = 2$, determinando quindi i punti c la cui esistenza è assicurata dalla tesi del teorema.

Soluzione. Poiché il dominio della funzione data è l'insieme $(-2/3, \infty)$, abbiamo che esiste un intorno di $x_0 = -1/3$, contenente anche il punto $x = 0$, in cui la $f(x)$ data risulta continua e derivabile infinite volte. La formula di Taylor di una funzione $f(x)$ contenente il polinomio di Taylor sviluppato fino all'ordine 2 è quella riportata nella (6.21b) e poiché risulta

$$f'(x) = \frac{3}{3x+2}, \qquad f''(x) = -\frac{9}{(3x+2)^2}, \qquad f'''(x) = \frac{54}{(3x+2)^3},$$

scriviamo la formula di Taylor sostituendo i valori $x_0 = -1/3$ e $x = 0$ nell'espressione (6.21b) e ricaviamo i punti c risolvendo l'equazione

$$f(0) = \left[f\left(-\frac{1}{3}\right) + f'\left(-\frac{1}{3}\right)\left(\frac{1}{3}\right) + \frac{1}{2!}f''\left(-\frac{1}{3}\right)\left(\frac{1}{3}\right)^2\right] + \frac{1}{3!}f'''(c)\left(\frac{1}{3}\right)^3 \qquad \Longrightarrow$$

$$\Longrightarrow \qquad \log 2 = 0 + 1 - \frac{9}{2}\left(\frac{1}{3}\right)^2 + \frac{1}{6}\,\frac{54}{(3c+2)^3}\left(\frac{1}{3}\right)^3 \qquad \Longrightarrow$$

$$\Longrightarrow \qquad -\frac{1}{2}+\log 2=\frac{1}{3(3c+2)^3} \qquad \Longrightarrow \qquad \frac{1}{(3c+2)^3}=\frac{6\log 2-3}{2} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad (3c+2)^3=\frac{2}{6\log 2-3} \qquad \Longrightarrow \qquad 3c+2=\sqrt[3]{\frac{2}{6\log 2-3}} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad 3c=-2+\sqrt[3]{\frac{2}{6\log 2-3}} \qquad \Longrightarrow \qquad c=\frac{1}{3}\left[\sqrt[3]{\frac{2}{6\log 2-3}}-2\right]\in\left(-\frac{1}{3},0\right).$$

Esercizio 2. Risolvere il medesimo esercizio precedente sviluppando il polinomio di Taylor fino all'ordine $n=1$.

Soluzione. La formula di Taylor di una funzione $f(x)$ contenente il polinomio di Taylor sviluppato fino all'ordine $n=1$ è quella riportata nella (6.21a) e dobbiamo risolvere quindi l'equazione di secondo grado nell'incognita c

$$f(0)=\left[f\left(-\frac{1}{3}\right)+f'\left(-\frac{1}{3}\right)\left(\frac{1}{3}\right)\right]+\frac{1}{2!}f''(c)\left(\frac{1}{3}\right)^2 \qquad \Longrightarrow$$

$$\Longrightarrow \qquad \log 2=0+1-\frac{1}{2}\frac{9}{(3c+2)^2}\left(\frac{1}{3}\right)^2 \qquad \Longrightarrow \qquad -1+\log 2=-\frac{1}{2(3c+2)^2} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad (3c+2)^2=\frac{1}{2-2\log 2} \qquad \Longrightarrow \qquad 3c+2=\pm\frac{1}{\sqrt{2-2\log 2}} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad 3c=-2\pm\frac{1}{\sqrt{2-2\log 2}} \qquad \Longrightarrow \qquad c_\pm=\frac{1}{3}\left[\pm\frac{1}{\sqrt{2-2\log 2}}-2\right].$$

Poiché i punti c della tesi del teorema sulla formula di Taylor appartengono all'intervallo (x_0,x) oppure (x,x_0), concludiamo che il punto c richiesto è solo quello con il segno positivo davanti alla radice quadrata perché dei due punti $c_\pm$ solo

$$c_+=\frac{1}{3}\left[\frac{1}{\sqrt{2-2\log 2}}-2\right]$$

appartiene all'intervallo $(-1/3,0)$. Dal confronto tra le soluzioni di questi due esercizi sulla formula di Taylor emerge che il valore del punto c della formula di Taylor di una data funzione dipende dall'ordine n fino al quale viene sviluppato il polinomio di Taylor della funzione stessa, ovvero, come abbiamo visto, il punto c di una funzione cambia se sviluppiamo il suo polinomio di Taylor fino all'ordine $n=2$ oppure fino all'ordine $n=1$.

6.3.1 Polinomio e formula di Taylor di funzioni elementari

Determiniamo quindi il polinomio di Taylor delle funzioni elementari

1) $f(x)=e^x$, 2) $f(x)=\log(1+x)$, 3) $f(x)=\sin x$, 4) $f(x)=\cos x$

nel punto $x_0=0$, ovvero il polinomio di Mc Laurin.

1) La funzione $f(x)=e^x$ è dotata di derivate di qualsiasi ordine per ogni x e ogni derivata è sempre $f^{(k)}(x)=e^x$ per ogni k intero positivo. Se quindi scegliamo $x_0=0$ e

applichiamo alla $f(x) = e^x$ la formula di Mc Laurin (6.20), abbiamo $f^{(k)}(0) = 1$ per ogni indice intero $k = 1, 2, 3, \ldots, n$ e $f^{(n+1)}(c) = e^c$, da cui segue

$$e^x = \left[1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots + \frac{x^{n-1}}{(n-1)!} + \frac{x^n}{n!}\right] + \left[\frac{e^c}{(n+1)!}\, x^{n+1}\right], \tag{6.22}$$

perché il *resto* $R_n(x)$ della funzione $f(x) = e^x$, dato dalla (6.14b), ha espressione

$$R_n(x) = \frac{e^c}{(n+1)!}\, x^{n+1},$$

che nel caso particolare $x = 1$ diventa

$$R_n(1) = \frac{e^c}{(n+1)!} > 0, \tag{6.23}$$

in cui c è un opportuno numero del quale sappiamo solo che è compreso tra 0 e 1.

Vediamo come la (6.22) può essere utilizzata per stimare l'ordine di grandezza dell'errore che si commette nell'approssimazione. Supponiamo di voler calcolare un valore approssimato di e. Il valore esatto di e coincide con il risultato della funzione e^x in cui si sostituisca $x = 1$. Poiché però non siamo grado di calcolare tale risultato, utilizziamo la formula di McLaurin (6.20) con $x = 1$ per ottenere un valore approssimato, ovvero

$$e = \left[1 + 1 + \frac{1}{2} + \frac{1}{3!} + \ldots + \frac{1}{(n-1)!} + \frac{1}{n!}\right] + R_n(1), \tag{6.24}$$

che riscriviamo nella forma

$$0 < e - \left[1 + 1 + \frac{1}{2} + \frac{1}{3!} + \ldots + \frac{1}{(n-1)!} + \frac{1}{n!}\right] = R_n(1) = \frac{e^c}{(n+1)!}, \tag{6.25a}$$

il cui importante significato è che il valore $R_n(1)$ rappresenta l'*errore* che "si commette" nell'approssimare il numero e con il numero ottenuto calcolando in $x = 1$ il polinomio di Taylor della funzione e^x. Il secondo membro nella relazione (6.25a) è ovviamente positivo perché nell'uguaglianza (6.24) il numero e è la somma della parentesi quadra positiva e del resto $R_n(1)$ che, per la (6.23), è positivo, da cui segue che nella (6.24) il numero e risulta maggiore sia del termine in parentesi quadra che del *resto* $R_n(1)$, ovvero la differenza al secondo membro nella (6.25a) risulta appunto positiva.

Poiché, come sappiamo, si ha $2 < e \leqslant 3$, da cui per $0 < c < 1$ segue $1 < e^c < e \leqslant 3$, otteniamo quindi

$$R_n(1) = \frac{e^c}{(n+1)!} < \frac{3}{(n+1)!}\,. \tag{6.25b}$$

Pertanto, approssimando il numero e con la (6.24), per i valori di n che vanno ad esempio da 1 a 7 abbiamo le seguenti approssimazioni

$$\begin{array}{ll}
\mathcal{P}_1(1) = 2 & R_1(1) < 3/2 = 1,5 \\
\mathcal{P}_2(1) = 2,5 & R_2(1) < 3/3! = 0,5 \\
\mathcal{P}_3(1) = 2,\overline{6} & R_3(1) < 3/4! = 0,125 \\
\mathcal{P}_4(1) = 2,708 & R_4(1) < 3/5! = 0,025 \\
\mathcal{P}_5(1) = 2,716 & R_5(1) < 3/6! = 0,0042 \\
\mathcal{P}_6(1) = 2,7180 & R_6(1) < 3/7! = 0,0006 \\
\mathcal{P}_7(1) = 2,71825 & R_7(1) < 3/8! = 0,00007
\end{array}$$

Calcolando il valore del numero e con una calcolatrice (che dia almeno 5 cifre decimali esatte) si ha $e \simeq 2,71828\ldots$ che come si vede coincide con $\mathcal{P}_7(x)$ fino alla quarta cifra decimale, d'accordo con il fatto che l'errore $R_7(x) < 0,00007 < 10^{-4}$.

2) La funzione $f(x) = \log(1+x)$ è dotata di derivate di qualsiasi ordine per ogni x e si ha in particolare

$$f(0) = \log 1 = 0, \qquad f'(0) = \left.\frac{1}{1+x}\right|_{x=0} = 1, \qquad f''(0) = -\left.\frac{1}{(1+x)^2}\right|_{x=0} = -1,$$

$$f'''(0) = \left.\frac{2}{(1+x)^3}\right|_0 = 2, \qquad f^{(IV)}(0) = -\left.\frac{6}{(1+x)^4}\right|_0 = -6, \quad \ldots \quad , f^{(k)}(0) = \frac{(-1)^{k+1}}{k!},$$

da cui, se sostituiamo nella (6.20), otteniamo

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + \frac{(-1)^{n+1}x^n}{n} + R_n(x).$$

3) La funzione $f(x) = \sin x$ è dotata di derivate di qualsiasi ordine per ogni x e si ha in particolare

$$f(0) = \sin 0 = 0, \qquad f'(0) = \left.\cos x\right|_0 = 1, \qquad f''(0) = \left.-\sin x\right|_0 = 0,$$

$$f'''(0) = \left.-\cos x\right|_0 = -1, \qquad f^{(IV)}(0) = \left.\sin x\right|_0 = 0, \qquad f^{(v)}(0) = \left.\cos x\right|_0 = 1,$$

ovvero $f^{(2k)}(0) = 0$ e $f^{(2k+1)}(0) = (-1)^k$, da cui, se sostituiamo nella (6.20), otteniamo

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + \frac{(-1)^n x^{2n+1}}{(2n+1)!} + R_{2n+1}(x),$$

cioè il polinomio di Mc Laurin della funzione $\sin x$ contiene solo le potenze di esponente dispari perché la funzione $\sin x$ è una funzione dispari.

4) Anche la funzione $f(x) = \cos x$ è dotata di derivate di qualsiasi ordine per ogni x e abbiamo in particolare

$$f(0) = \cos 0 = 1, \qquad f'(0) = \left.-\sin x\right|_0 = 0, \qquad f''(0) = \left.-\cos x\right|_0 = -1,$$

$$f'''(0) = \left.\sin x\right|_0 = 0, \qquad f^{(IV)}(0) = \left.\cos x\right|_0 = 1, \qquad f^{(v)}(0) = \left.-\sin x\right|_0 = 0,$$

ovvero $f^{(2k)}(0) = (-1)^k$ e $f^{(2k+1)}(0) = 0$, da cui, se sostituiamo nella (6.20), otteniamo

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + \frac{(-1)^n x^{2n}}{(2n)!} + R_{2n}(x),$$

cioè il polinomio di Mc Laurin della funzione $\cos x$ contiene solo le potenze di esponente pari perché la funzione $\cos x$ è una funzione pari.

6.3.2 Irrazionalità del numero e (prima dimostrazione)

Dalle due relazioni (6.25) segue la disuguaglianza valida per ogni numero naturale $n = 0, 1, 2, 3, \ldots$

$$0 < e - \left[1 + 1 + \frac{1}{2} + \frac{1}{3!} + \ldots + \frac{1}{(n-1)!} + \frac{1}{n!}\right] < \frac{3}{(n+1)!} \tag{6.26}$$

e possiamo dunque dimostrare che il numero e risulta essere un *numero irrazionale*, ovvero un numero non scrivibile come frazione, perché se, per assurdo, il numero e fosse il *numero razionale* frazionario $e = h/k$, con h, k due numeri interi positivi, si otterrebbe una conclusione "assurda" dalla disuguaglianza (6.26).

Per illustrare quale sia la conclusione "assurda" che si otterrebbe sostituendo il *numero razionale* $e = h/k$ nella disuguaglianza (6.26), ricordiamo brevemente la "semplice" proprietà di semplificazione dei *fattoriali*, ovvero l'ovvia proprietà per cui una frazione avente al numeratore un *fattoriale* maggiore, oppure uguale, del *fattoriale* al denominatore, risulta sempre essere un numero intero, come mostriamo tramite l'esempio delle frazioni

$$\frac{7!}{4!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7}{1 \cdot 2 \cdot 3 \cdot 4} = \frac{\cancel{1} \cdot \cancel{2} \cdot \cancel{3} \cdot \cancel{4} \cdot 5 \cdot 6 \cdot 7}{\cancel{1} \cdot \cancel{2} \cdot \cancel{3} \cdot \cancel{4}} = 5 \cdot 6 \cdot 7 = 210 \in \mathbb{N}.$$

A questo punto, se "per assurdo", come detto, il numero e fosse la frazione *numero razionale* $e = h/k$, sostituiamo il numero $e = h/k$ nella disuguaglianza (6.26) in cui scegliamo $n = k + 2$, ovvero n uguale al numero che si ottiene aumentando di 2 unità il denominatore della frazione h/k che dovrebbe rappresentare il numero e.

Se nella disuguaglianza (6.26) sostituiamo $e = h/k$ e poniamo appunto $n = k + 2$, dove il denominatore k può essere un qualsiasi numero naturale $k = 1, 2, 3, \ldots$, otteniamo la disuguaglianza

$$0 < \frac{h}{k} - \left[1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \ldots + \frac{1}{(k-1)!} + \frac{1}{k!} + \frac{1}{(k+1)!} + \frac{1}{(k+2)!}\right] < \frac{3}{(k+3)!} \tag{6.27a}$$

che risulta assurda per il seguente motivo. Infatti se moltiplichiamo i tre membri per il numero $(k+2)!$, otteniamo la disuguaglianza

$$0 < \frac{(k+2)!\,h}{k} - (k+2)!\left[1 + 1 + \frac{1}{2} + \frac{1}{3!} + \ldots + \frac{1}{(k+1)!} + \frac{1}{(k+2)!}\right] < \frac{3[(k+2)!]}{(k+3)!}, \tag{6.27b}$$

nella quale il secondo membro è un numero intero positivo N dato dalla somma di frazioni

$$\frac{(k+2)!}{k}, \quad \frac{(k+2)!}{2}, \quad \frac{(k+2)!}{3!}, \quad \frac{(k+2)!}{4!}, \quad \ldots, \quad \frac{(k+2)!}{(k+1)!}, \quad \frac{(k+2)!}{(k+2)!}$$

che sono tutti numeri naturali e inoltre il terzo membro è il numero

$$\frac{3[(k+2)!]}{(k+3)!} = \frac{3}{k+3},$$

il quale risulta minore di 1 perché per ogni possibile valore $k = 1, 2, 3, \ldots$ si ha $k + 3 > 3$.

Possiamo concludere quindi che la disuguaglianza (6.27a) è assurda perché da essa discende la disuguaglianza (6.27b) che è assurda perché ovviamente nessun numero intero positivo N può essere minore di 1. Se quindi la disuguaglianza (6.27a), nella quale si ponga $e = h/k$, è assurda, risulta dimostrato che il numero e non può essere uguale a nessuna frazione h/k, ovvero che il numero e appartiene all'insieme dei *numeri irrazionali*.

6.3.3 Serie di Taylor

Se la funzione $f(x)$ è dotata di derivate di qualsiasi ordine, segue che la (6.14a) ha significato per n intero comunque grande. Si può dimostrare inoltre che per ogni x per il quale valga il limite

$$\lim_{n\to\infty} R_n(x) = 0,$$

risulta quindi

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \frac{f^{(3)}(x_0)}{3!}(x - x_0)^3 + \cdots\cdots \tag{6.28}$$

cioè la $f(x)$ è la somma di una serie di funzioni, che prende il nome di **serie di Taylor**.

Si dice allora che la $f(x)$ è *sviluppabile* in serie di Taylor. Poichè di gran parte delle funzioni si conosce il valore esatto per $x_0 = 0$, la (6.28) viene quasi sempre utilizzata per sviluppare la funzione in un intorno dell'origine ed in tal caso è nota anche con il nome di serie di **Mc Laurin**:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \cdots\cdots\cdots \tag{6.29}$$

Lo sviluppo in serie di McLaurin della funzione e^x è dato, per la (6.22), da

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\cdots \tag{6.30}$$

In particolare, per $x = 1$ si ha:

$$e = 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \cdots\cdots, \tag{6.31}$$

ovvero il numero e non è solo l'estremo superiore e il limite della successione (5.2), ma abbiamo ottenuto che esso è anche la somma della serie (6.31).

6.4 Condizioni sufficienti per l'estremo

Il teorema 6.2.1 è, come già discusso, una condizione solo necessaria per l'estremo ed è chiamata anche *condizione del primo ordine* perché in tale teorema l'ordine più alto delle derivate che vi compaiono è appunto il primo ordine. Se nell'intervallo aperto (a, b) la funzione $f(x)$ è dotata anche di derivata seconda continua, il teorema che segue, nel caso in cui x_0 sia un punto stazionario e $f''(x_0) \neq 0$, fornisce una semplice condizione affinchè il punto x_0 sia anche un estremante.

Teorema 6.4.1. *(condizione sufficiente per l'estremo)*

Se $f'(x_0) = 0$ e $f''(x_0) \neq 0$, allora x_0 è un punto estremante: in particolare se risulta $f''(x_0) > 0$, in x_0 c'è un minimo relativo, mentre se risulta $f''(x_0) < 0$, c'è un massimo relativo.

Dimostrazione. Per le ipotesi fatte possiamo utilizzare la formula di Taylor fino all'ordine $n = 2$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + R_2(x) \tag{6.32}$$

la quale, tenuto conto dell'ipotesi $f'(x_0) = 0$, può essere riscritta nella forma

$$f(x) - f(x_0) = \frac{f''(x_0)}{2}(x - x_0)^2 + R_2(x), \tag{6.33}$$

in modo che il segno del primo membro $f(x) - f(x_0)$ possa essere ottenuto dal segno del secondo membro. In virtù del teorema 6.3.2, sarà possibile trovare un intorno completo di x_0 nel quale il segno del secondo membro della (6.33), ovvero il segno di $f(x) - f(x_0)$, coincide con il segno del solo termine

$$\frac{f''(x_0)}{2}(x - x_0)^2,$$

che geometricamente rappresenta una parabola.

Quindi se risulta $f''(x_0) > 0$, si avrà $f(x) - f(x_0) > 0$ in tutto un intorno di x_0, cioè $f(x) > f(x_0)$, e quindi in x_0 c'è un minimo proprio. Viceversa, se $f''(x_0) < 0$, si avrà $f(x) < f(x_0)$ e quindi in x_0 c'è un massimo proprio, *cdd.*

La condizione sufficiente appena dimostrata può essere generalizzata al caso in cui la funzione $f(x)$ sia derivabile n volte in x_0 con derivate continue in un intorno di x_0.

Se in queste ipotesi la funzione $f(x)$ verifica

$$f'(x_0) = f''(x_0) = f'''(x_0) = f^{(IV)}(x_0) = \cdots = f^{(n-1)}(x_0) = 0 \tag{6.34a}$$

e $f^{(n)}(x_0) \neq 0$ con n pari, segue che x_0 è un estremante: in particolare se $f^{(n)}(x_0) > 0$, in x_0 c'è un **minimo**, se $f^{(n)}(x_0) < 0$, in x_0 c'è un **massimo**. Abbiamo così ottenuto una condizione sufficiente di ordine n per l'estremo.

Infatti, in virtù delle uguaglianze (6.34a), il polinomio di Taylor di ordine n relativo alla $f(x)$ nel punto x_0 è

$$\mathcal{P}(x) = f(x_0) + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n,$$

da cui segue

$$f(x) - f(x_0) \approx \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + R_n(x), \tag{6.34b}$$

dove il segno di $(x - x_0)^n$ è sempre positivo perché n è pari.

Ora, se valgono le (6.34a) con n pari e $f^{(n)}(x_0) > 0$, dalla (6.34b) discende $f(x) - f(x_0) \geqslant 0$ per ogni x appartenente ad un intorno di x_0, cioè $f(x) \geqslant f(x_0)$ e il punto x_0 è un minimo; mentre se valgono le (6.34a) con n pari e $f^{(n)}(x_0) < 0$, dalla (6.34b) discende $f(x) - f(x_0) \leqslant 0$ per ogni x appartenente ad un intorno di x_0, cioè $f(x) \leqslant f(x_0)$ e il punto x_0 è un massimo.

Capitolo 7

Teoremi sui flessi e sulla concavità

Nel paragrafo 4.1.3 abbiamo introdotto i concetti di *comportamento convesso* e *comportamento concavo* di una funzione $f(x)$ in un intervallo utilizzando il confronto tra la funzione stessa e le sue rette secanti in due punti appartenenti all'intervallo stesso.

La definizione del *comportamento convesso* e del *comportamento concavo* formulata nel paragrafo 4.1.3 per una $f(x)$ in un intervallo, prescinde dal fatto che la funzione sia derivabile oppure non lo sia e vale ugualmente per esempio anche per funzioni, sì, continue, ma non derivabili che abbiano punti angolosi.

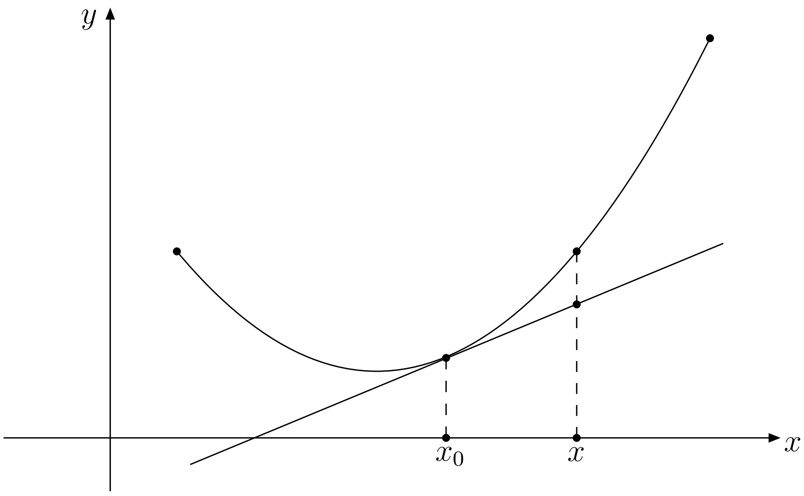


Fig. 7.1(a)

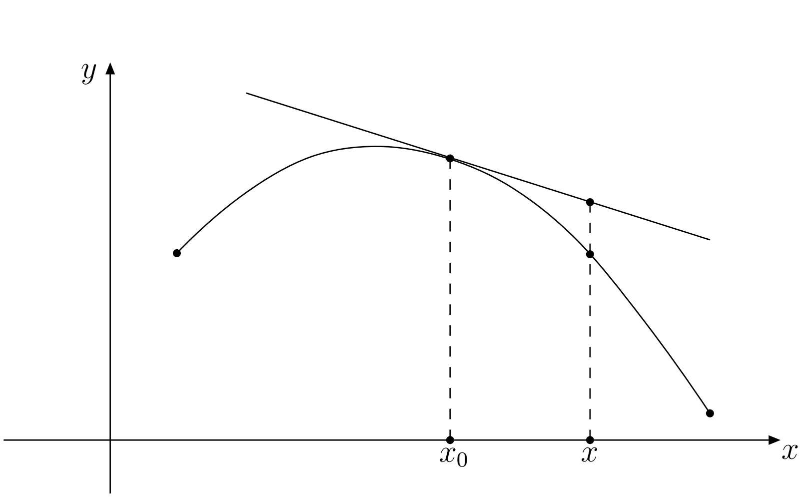


fig. 7.1(b)

Poiché vogliamo collegare i concetti di *comportamento convesso* e *comportamento concavo* di una funzione $f(x)$ alle derivate della $f(x)$ stessa, dobbiamo modificare preliminarmente la definizione del *comportamento convesso* e del *comportamento concavo* per esprimerla sulla base del confronto tra la $f(x)$ e le rette tangenti nei suoi punti. Se una funzione $f(x)$ è derivabile nell'intervallo $[a,b]$, si ha che esiste appunto la retta tangente al suo grafico in ogni punto $x_0 \in [a,b]$ e dimostriamo allora che la funzione è *convessa* se e solo se essa non si trova mai "al di sotto" della retta tangente. In altre parole, dimostriamo che se in un intervallo $[a,b]$ il grafico di una funzione è tale che ogni corda di secante non si trova mai "al di sotto" del corrispondente arco del grafico (fig. 4.6(a)), allora in $[a,b]$ il grafico della funzione non si trova mai "al di sotto" (fig. 7.1) della retta tangente condotta al grafico in un generico punto $x_0 \in [a,b]$. Invertendo le disuguaglianze si ottiene l'analogo per le funzioni *concave* che restano invece sempre "al di sotto" della tangente.

7.1 Concavità, convessità e rette tangenti

Se una funzione $f(x)$ è *convessa* in un intervallo $[a,b]$, ovvero risulta soddisfatta la disuguaglianza (4.12a), possiamo riscrivere tale disuguaglianza nella forma

$$f(x_0) - f(x_1) \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x_0 - x_1) \quad \Longrightarrow \quad \frac{f(x_0) - f(x_1)}{x_0 - x_1} \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

da cui, se denotiamo x_0 come x e calcoliamo il limite per $x \to x_1$ da destra (perché $x > x_1$), segue

$$\lim_{x \to x_1^+} \frac{f(x) - f(x_1)}{x - x_1} = f'(x_1) \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1} \quad \Longrightarrow \quad f(x_2) \geqslant f(x_1) + f'(x_1)(x_2 - x_1).$$

A questo punto possiamo denotare x_2 come x e riscrivere quest'ultima disuguaglianza nella forma $f(x) \geqslant f(x_1) + f'(x_1)(x - x_1)$, per ogni $x > x_1$.

Se invece esprimiamo il *comportamento convesso* della $f(x)$ nell'intervallo $[a,b]$ tramite l'equazione della secante nella forma (4.9b), ovvero nella forma

$$y = \frac{f(x_1) - f(x_2)}{x_1 - x_2}\, x + f(x_2) - \frac{f(x_1) - f(x_2)}{x_1 - x_2}\, x_2\,,$$

otteniamo analogamente

$$f(x_0) - f(x_2) \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x_0 - x_2) \quad \Longrightarrow \quad \frac{f(x_0) - f(x_2)}{x_0 - x_2} \geqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

da cui, se denotiamo x_0 come x e calcoliamo il limite per $x \to x_2$ da sinistra (perché $x < x_2$), segue

$$\lim_{x \to x_2^-} \frac{f(x) - f(x_2)}{x - x_2} = f'(x_2) \geqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1} \quad \Longrightarrow \quad f(x_1) \geqslant f(x_2) + f'(x_2)(x_1 - x_2).$$

A questo punto possiamo denotare x_1 come x e riscrivere quest'ultima disuguaglianza nella forma $f(x) \geqslant f(x_2) + f'(x_2)(x - x_2)$, per ogni $x < x_2$. Le due disuguaglianze

$$f(x) \geqslant f(x_1) + f'(x_1)(x - x_1) \qquad \text{e} \qquad f(x) \geqslant f(x_2) + f'(x_2)(x - x_2),$$

rispettivamente per ogni $x > x_1$ e per ogni $x < x_2$, possono essere riunite, per la genericità dei punti x_1, x_2, nell'unica disuguaglianza

$$f(x) \geqslant f(x_0) + f'(x_0)(x - x_0), \tag{7.1a}$$

per ogni $x_0, x \in [a, b]$, che mostra appunto che se una funzione è *convessa* in $[a, b]$ secondo la definizione 4.1.3, allora, condotta la retta tangente per un generico punto $x_0 \in (a, b)$, la curva $f(x)$ non si trova mai, per $x \in (a, b)$, "al di sotto" della retta tangente (fig. 7.1(a)).

Analogamente, se una funzione è *concava* in $[a, b]$ secondo la definizione 4.1.4, allora, invertendo le disuguaglianze, otteniamo la disuguaglianza rispetto alla retta tangente

$$f(x) \leqslant f(x_0) + f'(x_0)(x - x_0), \tag{7.1b}$$

per ogni $x_0, x \in [a, b]$, che mostra appunto che se una funzione è *concava* in $[a, b]$ secondo la definizione 4.1.4, allora, condotta la retta tangente per un generico punto $x_0 \in (a, b)$, la curva $f(x)$ non si trova mai, per $x \in (a, b)$, "al di sopra" della retta tangente (fig. 7.1(b)).

Queste proprietà geometriche della retta tangente suggeriscono un "nuovo" modo di definire (per funzioni derivabili anche se non in tutto un intervallo) i concetti di *convessità* e di *concavità* in un punto che fa intervenire le rette tangenti invece delle rette secanti (corde). Poiché per lo studio della convessità e della concavità di una funzione $f(x)$ in un intervallo (a, b) dovremo utilizzare le rette tangenti al suo grafico, considereremo d'ora in poi solo funzioni continue in $[a, b]$ e derivabili almeno una volta in (a, b), perché se $f(x)$ è derivabile almeno una volta nell'intervallo (a, b), esiste la tangente al suo grafico in ogni punto $x_0 \in (a, b)$, avente equazione $y = f(x_0) + f'(x_0)(x - x_0)$.

Quindi, le definizioni 4.1.3 e 4.1.4, rispettivamente del *comportamento convesso* e del *comportamento concavo*, vengono dunque riformulate in termini delle rette tangenti, secondo quanto mostrato nelle figg. 7.1.

Abbiamo dunque la "nuova" definizione del *comportamento convesso*

Definizione 7.1.1. *Se è possibile trovare un intorno completo di x_0 nel quale il grafico della $f(x)$ si mantiene sempre (per $x \neq x_0$) "al di sopra" della retta tangente (fig. 7.1(a)), si dice che in x_0 la curva è **convessa**, ovvero "volge la concavità verso l'alto"*

e la "nuova" definizione del *comportamento concavo*

Definizione 7.1.2. *Se è possibile trovare un intorno completo di x_0 nel quale il grafico della $f(x)$ si mantiene "al di sotto" della retta tangente (fig. 7.1(b)), si dice che in x_0 la curva è **concava**, ovvero "volge la concavità verso il basso".*

Infine possiamo passare dal *comportamento* nell'intorno di un punto x_0 al *comportamento* in tutto un intervallo ponendo la seguente

Definizione 7.1.3. *Se una funzione $f(x)$ è convessa in tutti i punti di un intervallo $[a, b]$, la funzione $f(x)$ si dice convessa in (a, b). Se una $f(x)$ è concava in tutti i punti di un intervallo (a, b), la $f(x)$ si dice concava in (a, b).*

Per confrontare il grafico di una funzione $f(x)$ e le rette tangenti al suo grafico nei punti $x_0 \in (a, b)$, consideriamo l'espressione

$$f(x) - [f(x_0) + f'(x_0)(x - x_0)], \tag{7.2}$$

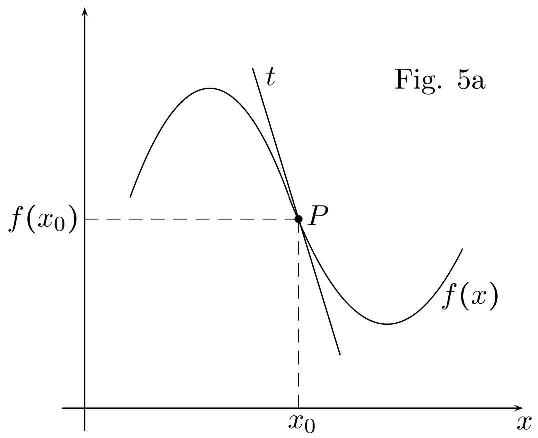
che rappresenta la differenza tra l'ordinata generata dalla funzione $f(x)$ e l'ordinata generata dalla retta tangente in corrispondenza di una medesima ascissa x.

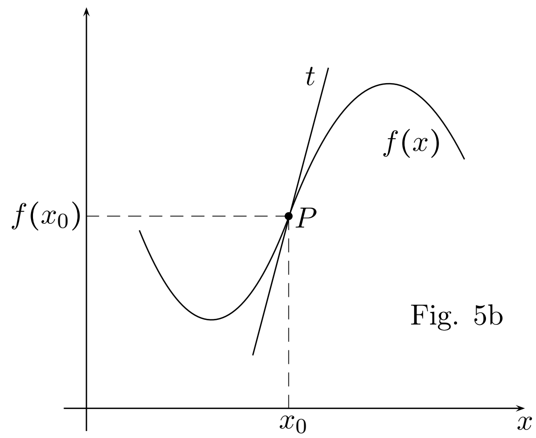
Possiamo dire quindi che in un intorno I di x_0 il grafico della funzione $f(x)$ si mantiene sempre (per $x \neq x_0$) al di sopra della retta tangente, ovvero la funzione $f(x)$ è *convessa* in I, se vale la disuguaglianza (7.1a), ovvero se l'espressione (7.2) risulta positiva per ogni $x \in I$, con $x \neq x_0$; mentre invece il grafico della funzione $f(x)$ si mantiene sempre (per $x \neq x_0$) al di sotto della retta tangente, ovvero la $f(x)$ è *concava* in I, se vale dunque la disuguaglianza (7.1b), ovvero se l'espressione (7.2) risulta negativa per ogni $x \in I$, con $x \neq x_0$.

E' importante ora caratterizzare il punto, che viene denominato *punto di flesso* oppure semplicemente *flesso*, in cui il *comportamento concavo* e il *comportamento convesso* si scambiano.

Definizione 7.1.4. *Data una funzione $f(x)$ definita in un intervallo $[a,b]$ e un punto $x_0 \in (a,b)$, se esiste un intorno I di x_0 in cui la $f(x)$ sia derivabile almeno una volta e in cui l'espressione* (7.2)*, ovvero $f(x) - [f(x_0) + f'(x_0)(x - x_0)]$, cambia segno nel passaggio dai valori $x < x_0$ ai valori $x > x_0$, dove $y = f(x_0) + f'(x_0)(x - x_0)$ denota l'equazione della retta tangente al grafico della $f(x)$ in x_0, allora si dice che in x_0 la curva ha un **flesso**, ovvero che x_0 è un punto di **flesso** per la $f(x)$.*

E' immediato rendersi conto che la condizione del cambiamento di segno dell'espressione $f(x) - [f(x_0) + f'(x_0)(x - x_0)]$ nel passaggio da $x < x_0$ a $x > x_0$, equivale a dire che le due parti di curva corrispondenti rispettivamente all'intorno sinistro e all'intorno destro di x_0 si trovano, come mostrato nelle seguenti figg. 5, in parti opposte rispetto alla retta t tangente in x_0 avente appunto equazione $y = f(x_0) + f'(x_0)(x - x_0)$.





In particolare abbiamo in fig. 5a che la (7.2) si mantiene negativa in un intorno sinistro e positiva in un intorno destro di x_0; in fig. 5b, invece, la (7.2) si mantiene positiva in un intorno sinistro e negativa in un intorno destro di x_0. Pertanto il punto P di ascissa x_0 è un *flesso* per la funzione $f(x)$ sia in fig. 5a che in fig. 5b.

Osserviamo che se $f'(x_0) = 0$ (la tangente è orizzontale) e $f(x)$ è convessa in x_0, allora essa ha un *minimo* locale in tale punto; se invece è concava ha un *massimo*.

Quindi sia per le funzioni *concave* che per quelle *convesse*, la condizione del primo ordine $f'(x_0) = 0$ per estremi locali è non solo *necessaria* ma anche *sufficiente*.

Poiché una funzione $f(x)$ è convessa nell'intervallo (a,b) se e solo se la $f(x)$ risulta convessa in tutti i punti $x_0 \in (a,b)$, segue che $f(x)$ è convessa nell'intervallo (a,b) se e solo se l'espressione (7.2) risulta positiva per ogni coppia di punti $x_0, x \in [a,b]$.

A questo punto osserviamo che se la $f(x)$ è dotata anche di derivate di ordine successivo al primo, è possibile studiare facilmente il segno della (7.2) tramite il polinomio di Taylor.

7.2 Teoremi su flessi, convessità e concavità

Teorema 7.2.1. *(Condizione necessaria per il flesso)*

Data $f(x)$ definita e derivabile due volte in un intorno I di x_0, se x_0 è un punto di flesso per la $f(x)$, segue $f''(x_0) = 0$.

Prima di dare la dimostrazione di questo teorema, osserviamo che se scambiamo l'ipotesi e la tesi di questo teorema otteniamo un "teorema falso".

Infatti se consideriamo come *controesempio* la funzione $g(x) = x^4$, abbiamo che la derivata seconda $g''(x) = 12x^2$ si annulla in $x_0 = 0$, ma il punto $x_0 = 0$ non è un flesso per la funzione $g(x) = x^4$.

Dimostrazione del teorema 7.2.1. Procediamo per assurdo, ovvero dimostriamo che se $f''(x_0) \neq 0$, segue che x_0 **non** è un flesso. Supponiamo ora che sia $f''(x_0) \neq 0$ e scriviamo quindi la formula di Taylor con $n = 2$ nella forma

$$f(x) - [f(x_0) + f'(x_0)(x - x_0)] = \frac{f''(x_0)}{2}(x - x_0)^2 + R_2(x), \tag{7.3}$$

dove, come sappiamo, il segno del secondo membro coincide, in un opportuno intorno del punto x_0, con il segno del suo primo addendo e quindi con il segno di $f''(x_0)$, essendo il quadrato $(x - x_0)^2$ sempre positivo. Pertanto abbiamo che se $f''(x_0) > 0$, la $f(x)$ è **convessa** in un intorno di x_0 (f sopra la tangente), mentre se $f''(x_0) < 0$, la $f(x)$ è **concava** in un intorno di x_0 (f sotto la tangente). Pertanto possiamo concludere che se risulta $f''(x_0) \neq 0$, segue che x_0 non può essere flesso, ovvero che se x_0 è un punto di flesso per la $f(x)$, in x_0 deve valere $f''(x_0) = 0$, *cdd.*

Teorema 7.2.2. *(condizione per la convessità in un intervallo)*

Se una funzione $f(x)$ è derivabile due volte in un intervallo $[a, b]$ e risulta $f''(x) \geqslant 0$, per ogni $x \in [a, b]$, segue che $f(x)$ è convessa in tutto $[a, b]$.

Dimostrazione del teorema 7.2.2. Scelto un generico punto $x_0 \in [a, b]$ e scritta la relazione (7.3), se $f''(x) \geqslant 0$ per ogni $x \in [a, b]$, abbiamo che il secondo membro della (7.3) risulta positivo. Ma se è positivo il secondo membro della (7.3), risulta positiva, per ogni coppia di punti $x_0, x \in [a, b]$, anche l'espressione (7.2), coincidente appunto con il primo membro della (7.3), *cdd.*

Supponiamo ora che in (a, b) la funzione $f(x)$ sia dotata anche di derivata terza continua (sia cioè, come si dice, di classe $C^3[(a, b)]$). Il teorema che segue, nel caso in cui valga $f''(x_0) = 0$, fornisce una semplice condizione affinchè x_0 sia un flesso.

Teorema 7.2.3. *(condizione sufficiente per il flesso)*

Se $f''(x_0) = 0$ e $f'''(x_0) \neq 0$, segue che x_0 è un punto di flesso.

Dimostrazione. Se abbiamo $f''(x_0) = 0$ e $f'''(x_0) \neq 0$, si può utilizzare il polinomio di Taylor di terzo grado, ovvero

$$f(x) - [f(x_0) + f'(x_0)(x - x_0)] = \frac{f'''(x_0)}{3!}(x - x_0)^3 + R_3(x), \tag{7.4}$$

dove, come sappiamo, il segno del secondo membro coincide, in un opportuno intorno del punto x_0, con il segno del primo addendo, ovvero anche in questo caso in un opportuno intorno di x_0 la differenza a primo membro ha lo stesso segno del termine

$$\frac{f'''(x_0)}{3!}(x-x_0)^3,$$

il quale però non dipende più solo dal segno di $f'''(x_0)$ in quanto $(x-x_0)^3$ è negativo per $x < x_0$ e positivo per $x > x_0$. Quindi la differenza a primo membro della (7.4) ha segno opposto a quello di $f'''(x_0)$ se x è a sinistra di x_0 ed ha invece lo stesso segno se è a destra, ovvero la concavità della funzione cambia dai punti a sinistra ai punti a destra di x_0 e quindi, secondo la definizione 7.1.4, il punto x_0 risulta essere un punto di flesso, *cdd.*

La condizione sufficiente appena dimostrata può essere generalizzata al caso in cui la funzione $f(x)$ sia derivabile n volte in x_0 con derivate continue in un intorno di x_0.

Se in queste ipotesi la funzione $f(x)$ verifica

$$f''(x_0) = f'''(x_0) = f^{(iv)}(x_0) = \cdots = f^{(n-1)}(x_0) = 0 \tag{7.5a}$$

e $f^{(n)}(x_0) \neq 0$ con n dispari, segue che x_0 è un flesso.

Infatti, in virtù delle (7.5a), il polinomio di Taylor di ordine n relativo alla $f(x)$ nel punto x_0 è

$$\mathcal{P}(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n,$$

da cui segue

$$f(x) - [f(x_0) + f'(x_0)(x-x_0)] = \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_n(x). \tag{7.5b}$$

Poiché n è dispari, segue che $(x-x_0)^n$, e quindi tutto il secondo membro della (7.5b), cambia segno in ogni intorno di x_0 se prendiamo valori a sinistra e valori a destra di x_0.

Poiché dunque anche il primo membro della (7.5b) cambia segno in ogni intorno di x_0 se prendiamo valori a sinistra e valori a destra di x_0, concludiamo che x_0 è un punto di flesso.

Esempio 1. *La $f(x) = x^2$ ha $f''(x) = 2 > 0$ per ogni x. Quindi la funzione è convessa in qualunque punto dell'asse reale.*

Esempio 2. *La $f(x) = x^3$ ha $f''(x) = 6x$* e *$f'''(x) = 6$. Per $x_0 < 0$ è $f''(x_0) < 0$ e quindi la funzione è concava. Per $x_0 > 0$ è $f''(x_0) > 0$ e la funzione è convessa. Per $x_0 = 0$ è $f''(0) = 0$ e $f'''(0) > 0$ per cui in tale punto c'è un flesso.*

Esempio 3. *La $f(x) = x^4$ ha $f''(x) = 12x^2$; $f'''(x) = 24x$* e *$f^{(IV)}(x) = 24$. Per $x \neq 0$ dal solo esame della derivata seconda si capisce che la funzione è convessa. Per $x_0 = 0$ si ha $f''(0) = f'''(0) = 0$ ed invece $f^{(IV)}(0) > 0$ per cui essendo la prima derivata non nulla e positiva la quarta di ordine pari si conclude che anche in 0 la funzione è convessa.*

Esempio 4. *La $f(x) = x^5$ ha $f''(x) = 20x^3$; $f'''(x) = 60x^2$; $f^{(IV)}(x) = 120x$* e *$f^{(V)}(x) = 120$. Anche in questo caso per gli $x \neq 0$ basta la derivata seconda per capire il comportamento della funzione. Invece per $x_0 = 0$ si ha $f''(0) = f'''(0) = f^{(IV)}(0) = 0$ ed invece $f^{(V)}(0) > 0$ ed essendo la prima derivata diversa da zero e positiva la quinta di ordine dispari in* 0, si conclude che *c'è un flesso.*

Esempio 5. *Se $f(x) = e^x$ è anche $f''(x) = e^x$ sempre positiva: e^x è sempre convessa. Analogamente se $f(x) = a^x$, allora $f'(x) = a^x \log a$ e $f''(x) = a^x (\log a)^2$, quindi anche in questo caso la derivata seconda è sempre positiva per cui anche a^x è sempre convessa.*

Capitolo 8

Elementi di calcolo integrale

Data una funzione $f(x)$ positiva e continua nell'intervallo $[a, b]$ chiuso e limitato, consideriamo la regione Ω del piano delimitata, come mostrato nella fig. 5.1(a), dal segmento sull'asse x tra gli estremi a e b, dai segmenti verticali compresi tra le coppie di punti rispettivamente $(a, 0), (a, f(a))$ e $(b, 0), (b, f(b))$, dall'arco della curva $f(x)$ tra a e b.

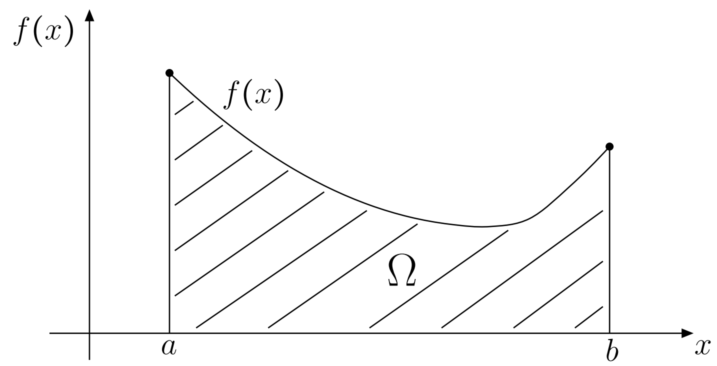
8.1 Area di regioni piane

Definizione 8.1.1. *Definiamo "integrale" della funzione $f(x)$ positiva e continua nell'intervallo $[a, b]$ chiuso e limitato quel numero positivo, indicato con il simbolo*

$$\int_a^b f(x)\, dx, \tag{8.1}$$

che esprime il valore dell'area della regione geometrica piana Ω rappresentata in fig. 5.1(a) e indicata con "$\mathcal{A}(\Omega)$". La funzione $f(x)$ viene chiamata "funzione integranda" e l'intervallo $[a, b]$ viene denominato "intervallo d'integrazione".

fig. 5.1(a)



Vogliamo far presente che tale definizione di *integrale*, risalente ai *Principia* di Newton, non è "molto corretta" nell'ambito della moderna Analisi Matematica perché il concetto di *integrale* viene introdotto sulla base del concetto di *area* che non è stato definito in precedenza e quindi non è rispettato il principio della logica aristotelica secondo il quale un concetto può essere definito solo sulla base di altri concetti che siano già stati definiti in precedenza. Tuttavia reputiamo che concettualmente non sia del tutto privo di fondamento definire l'integrale come area perché per definire un concetto si può pur sempre adottare come valido quel metodo che fa ricorso alle cosiddette *idee platoniche*, dette anche *archetipi platonici*, i quali, appunto secondo la *Metafisica Platonica*, si trovano nel cosiddetto *Iperuranio* e possono venir considerati con l'*intelletto* di individui razionali, anche in mancanza di una definizione formale. Pertanto, sebbene il concetto di area non sia stato definito in modo formale tramite una definizione precedente, tuttavia possiamo dire che l'area di una regione di piano è un *archetipo platonico* di cui non diamo definizione formale, ma del quale "tutti" abbiamo l'*idea* nel *nostro intelletto di individui razionali.*

Poiché dunque l'integrale (8.1) di funzioni positive sull'intervallo d'integrazione rappresenta l'area geometrica della regione Ω mostrata in fig. 5.1(a), ovvero poniamo

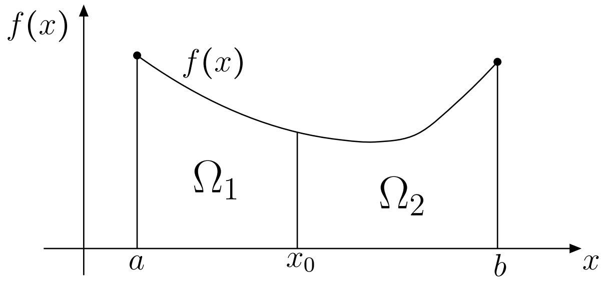
$$\mathcal{A}(\Omega) = \int_a^b f(x)\,dx, \tag{8.2}$$

dalla proprietà geometrica "intuitiva" di *additività* delle aree mostrata nella fig. 5.1(b), possiamo ricavare la proprietà evidente (sulla cui dimostrazione non ci soffermiamo) dell'integrale, espressa dalla seguente relazione valida per ogni $x_0 \in (a,b)$

$$\int_a^b f(x)\,dx = \mathcal{A}(\Omega_1) + \mathcal{A}(\Omega_2) = \int_a^{x_0} f(x)\,dx + \int_{x_0}^b f(x)\,dx \tag{8.3}$$

e possiamo dimostrare, in virtù delle proprietà generali di confronto tra aree di figure piane, la seguente proprietà dell'integrale denominata *teorema del valor medio.*

fig. 5.1(b)



Teorema 8.1.1. *(del valor medio)*

Data una funzione $f(x)$ positiva e continua nell'intervallo chiuso e limitato $[a,b]$, esiste un punto $c \in (a,b)$ tale che risulti

$$\frac{1}{b-a}\int_a^b f(x)\,dx = f(c) \qquad \textit{oppure} \qquad \int_a^b f(x)\,dx = f(c)\,(b-a). \tag{8.4}$$

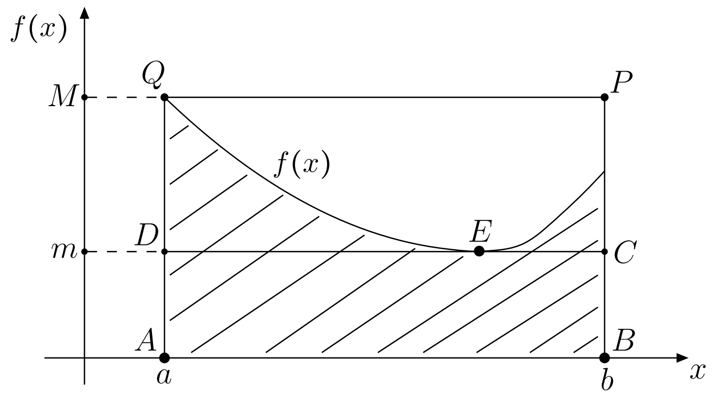
Dimostrazione. Poiché la funzione $f(x)$ è continua nell'intervallo $[a,b]$ chiuso e limitato, segue, per il teorema di Weierstrass, che la $f(x)$ possiede massimo e minimo assoluti in $[a,b]$, che, con riferimento alla fig. 5.2, coincidono rispettivamente con il punto Q, avente ascissa a e ordinata $f(a) = M$, e con il punto E, avente ordinata m. L'area della regione piana Ω (indicata con il tratteggio) risulta palesemente maggiore dell'area del rettangolo A,B,C,D e minore dell'area del rettangolo A,B,P,Q. Poiché l'area del rettangolo A,B,C,D vale $(b-a)m$, prodotto della base $\overline{AB} = b-a$ per l'altezza $\overline{AD} = m$ e, analogamente, l'area del rettangolo A,B,P,Q vale $(b-a)M$, possiamo scrivere la disuguaglianza tra aree $(b-a)m \leqslant area(\Omega) \leqslant (b-a)M$, ovvero sostituendo la (8.2)

$$(b-a)m \leqslant \int_a^b f(x)\,dx \leqslant (b-a)M,$$

da cui, se dividiamo i tre membri per il numero positivo $b-a$, segue la disuguaglianza per l'integrale

$$m \leqslant \frac{1}{b-a} \int_a^b f(x)\,dx \leqslant M. \tag{8.5}$$

fig. 5.2

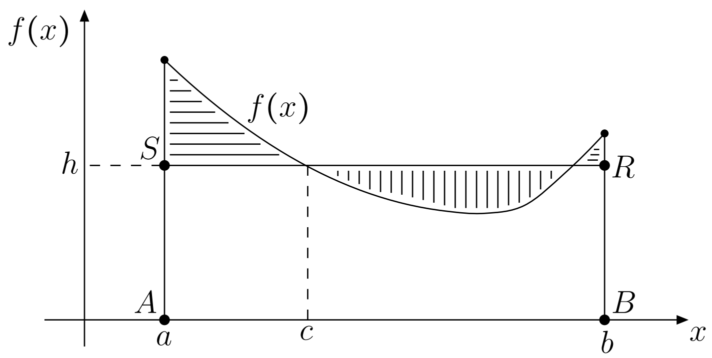


La disuguaglianza (8.5) mostra che il numero

$$\frac{1}{b-a} \int_a^b f(x)\,dx \tag{8.6}$$

è compreso tra il minimo m e il massimo M assoluti della funzione $f(x)$ continua sull'intervallo chiuso e limitato $[a,b]$. Il teorema dei valori intermedi di una funzione $f(x)$ continua sull'intervallo chiuso e limitato $[a,b]$ assicura che per ogni numero y_0 compreso tra il minimo m e il massimo M assoluti della $f(x)$, esiste sempre almeno un punto $c \in (a,b)$ tale che risulti $y_0 = f(c)$. Se come numero y_0 compreso tra m e M consideriamo l'integrale (8.6), segue che esiste almeno un punto $c \in (a,b)$ in cui vale la tesi (8.4), *c.d.d.*

fig. 5.3



Possiamo esprimere il teorema del valor medio dicendo che, data una funzione f positiva e continua in un intervallo chiuso e limitato, il suo integrale tra i due estremi dell'intervallo stesso è uguale al prodotto tra la differenza dei due estremi e il valore della stessa funzione f calcolato su un punto c compreso tra i due estremi.

Il significato geometrico del teorema del valor medio è mostrato in fig. 5.3. Poiché per una funzione $f(x)$ positiva e continua nell'intervallo $[a,b]$ chiuso e limitato, l'area della regione piana Ω individuata dalla $f(x)$ risulta compresa tra le aree dei due rettangoli aventi base $\overline{AB} = b-a$ e altezze m e M, segue che esiste sempre un rettangolo, indicato con A, B, R, S e avente altezza $h = f(c)$ intermedia tra il minimo m e il massimo M assoluti della $f(x)$ in $[a,b]$, la cui area $f(c)(b-a)$, data dal prodotto della base $\overline{AB} = b-a$ per l'altezza $\overline{AS} = h = f(c)$, coincide con l'area della regione piana Ω individuata dalla $f(x)$, ovvero verifica la la seconda forma della tesi (8.4) del teorema del valor medio.

8.2 Funzione primitiva e integrale

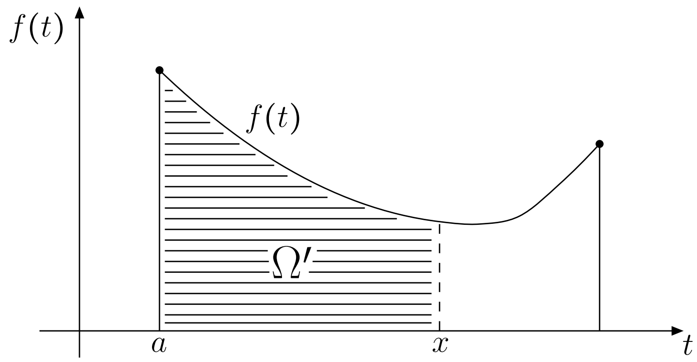
Il problema che vogliamo affrontare a questo punto è quello di determinare il valore dell'integrale (8.2) che esprime l'area della regione piana Ω riportata in fig. 5.1(a).

A questo scopo utilizziamo la fig. 5.4 in cui le ascisse del riferimento cartesiano e la funzione positiva e continua nell'intervallo $[a, b]$ chiuso e limitato sono denotate rispettivamente con la variabile t e con $f(t)$. E' immediato rendersi conto quindi che se sul piano cartesiano cambiamo denominazione della variabile per le ascisse da x a t, le curve associate rispettivamente alla funzione $f(x)$, se la variabile è x, e alla funzione $f(t)$, se la variabile è t, ovviamente coincidono e quindi individuano la stessa regione piana Ω.

Per ricavare la *formula generale* che collega l'integrale (8.2) alla funzione integranda f e agli estremi a, b, consideriamo non tutta la regione piana Ω tra gli estremi fissati a, b, ma una sua "porzione" (quella tratteggiata in fig. 5.4) tra l'estremo fissato a e l'estremo variabile $x \in [a, b]$. In questo modo, il valore dell'area della regione tratteggiata riportata in fig. 5.4 diventa a sua volta una funzione perché l'area individuata dalla curva $f(t)$ tra gli estremi a e x assume valori che variano al variare dell'estremo x.

Quindi, il motivo per cui abbiamo cambiato la denominazione della variabile sull'asse delle ascisse è che in questo modo non si genera confusione tra la variabile t da cui dipende la funzione integranda $f(t)$ e l'estremo x da cui dipende non la funzione integranda $f(t)$, bensì la funzione che esprime l'area della regione Ω' individuata dalla curva $f(t)$.

fig. 5.4



In virtù della definizione 8.1.1 di integrale e "per coerenza" con la relazione (8.2), abbiamo dunque che l'area della regione piana Ω' individuata dalla curva $f(t)$ tra gli estremi a e x, mostrata in fig. 5.4, si esprime dunque tramite l'integrale

$$\int_a^x f(t)\, dt \tag{8.7}$$

del quale possiamo subito dire che fornisce risultato zero per $x = a$ perché se il secondo estremo x coincide con il primo estremo a, la regione piana Ω' "collassa" sul segmento di ascissa a e dunque la sua area è "palesemente" nulla. Poiché l'area della regione piana Ω', ovvero l'integrale (8.7), dipende dall'estremo x fin dove si estende la regione Ω', possiamo porre l'integrale (8.7) uguale appunto ad una funzione $G(x)$, ovvero

$$G(x) = \int_a^x f(t)\, dt, \tag{8.8a}$$

denominata *funzione integrale* della funzione integranda $f(t)$.

Se a questo punto osserviamo la relazione (8.8a), ci rendiamo conto immediatamente che essa contiene un'incoerenza perché mentre la funzione scritta come integrale al secondo membro fornisce, come discusso, risultato zero per $x = a$, invece la funzione $G(x)$ al primo membro fornisce, per il medesimo valore $x = a$, risultato generico $G(a) \neq 0$. Per eliminare tale inconsistenza riscriviamo la funzione integrale $G(x)$ data dalla (8.8a) nella forma finale

$$G(x) = G(a) + \int_a^x f(t)\,dt, \tag{8.8b}$$

in modo tale che ambo i membri della (8.8b) diano, per $x = a$, il medesimo risultato $G(a)$ e dunque la funzione $G(x)$ sia consistente con le proprietà delle aree.

Se riusciamo a determinare l'espressione della funzione $G(x)$, otteniamo dunque dall'uguaglianza (8.8b) la *formula generale* dell'integrale (8.7), ovvero

$$\int_a^x f(t)\,dt = G(x) - G(a),$$

da cui, se poniamo in particolare $x = b$, ricaviamo la *formula generale* dell'integrale (8.2)

$$\int_a^b f(x)\,dx = \left. G(x)\right|_a^b = G(b) - G(a), \tag{8.9}$$

della quale dovremo verificare la coerenza con la definizione iniziale 8.1.1 di integrale.

Infatti, poiché l'area geometrica di una regione limitata di piano risulta sempre ovviamente positiva e l'integrale è stato definito appunto come area geometrica di una regione limitata di piano, verificheremo che l'integrale (8.9) risulta effettivamente positivo.

La relazione che esprime il legame tra la funzione integrale $G(x)$ e la corrispondente funzione integranda $f(t)$ viene stabilita dal seguente teorema che prende il nome di *teorema fondamentale del calcolo integrale* oppure anche *teorema di Torricelli-Barrow.*

Teorema 8.2.1. *(di Torricelli-Barrow)*

Data una funzione $f(x)$ positiva e continua sull'intervallo chiuso e limitato $[a, b]$, segue che la funzione $G(x)$ definita dalle uguaglianze (8.8) *risulta derivabile e soddisfa la relazione (fondamentale) $G'(x) = f(x)$.*

Per dimostrare questo teorema calcoleremo esplicitamente la derivata prima della funzione $G(x)$ e verificheremo che essa coincide appunto con la funzione integranda $f(x)$.

Poiché le due funzioni $G(x)$ definite nella (8.8a) e nella (8.8b) differiscono solo per una costante additiva, è ovvio che entrambe le $G(x)$ abbiano la medesima derivata prima.

Pertanto, per calcolare la derivata prima della $G(x)$ è indifferente utilizzare la sua espressione (8.8a) oppure (8.8b) e quindi per la $G(x)$ utilizzeremo, per comodità, l'espressione (8.8a), perché essa ha un addendo in meno, ovvero in essa manca l'addendo irrilevante $G(a)$, che appunto non gioca nessun ruolo ai fini del calcolo della derivata.

Dimostrazione del teorema di Torricelli-Barrow. Per calcolare la derivata prima della funzione $G(x)$ in un generico punto $x_0 \in (a, b)$, sviluppiamo il suo *rapporto incrementale* nel punto x_0 utilizzando la proprietà geometrica (8.3) di additività delle aree e il teorema 8.1.1 del valor medio. Con riferimento alla fig. 5.5, abbiamo la relazione

$$\int_a^x f(t)\,dt - \int_a^{x_0} f(t)\,dt = \int_{x_0}^x f(t)\,dt,$$

da cui segue lo sviluppo del rapporto incrementale della $G(x)$

$$\frac{G(x) - G(x_0)}{x - x_0} = \frac{1}{x - x_0}\left[\int_a^x f(t)\,dt - \int_a^{x_0} f(t)\,dt\right] = \frac{1}{x - x_0}\int_{x_0}^x f(t)\,dt = f(c),$$

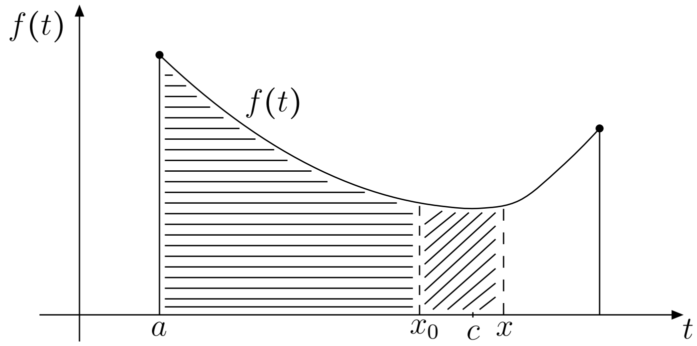
dove per l'ultima uguaglianza abbiamo applicato la tesi (8.4) del teorema del valor medio con le identificazioni $x_0 \equiv a$ e $x \equiv b$ e con $c \in (x_0, x)$. Poiché la derivata di una funzione in un punto è data dal limite del rapporto incrementale nel punto, segue che la derivata prima della funzione $G(x)$ nel generico punto $x_0 \in (a, b)$, indicata con $G'(x_0)$ è data da

$$G'(x_0) = \lim_{x \to x_0} \frac{G(x) - G(x_0)}{x - x_0} = \lim_{x \to x_0} f(c) = \lim_{c \to x_0} f(c) = f(x_0) \in \mathbb{R}, \tag{8.10}$$

dove il limite per $x \to x_0$ può essere sostituito dal limite per $c \to x_0$ perché, come si evince dalla fig. 5.5, se x tende a x_0, segue che anche c, essendo sempre compreso tra x_0 e x, tende sempre ugualmente a x_0. Infine, nell'ultima uguaglianza è stata applicata l'ipotesi di continuità della funzione f per cui il limite della funzione f continua quando la sua variabile c tende al valore x_0, coincide con il risultato $f(x_0)$ della funzione f nel punto x_0.

Poiché per la continuità della funzione f all'interno dell'intervallo $[a, b]$ si ha che $f(x_0)$ è un numero reale (finito), segue dunque che nel punto x_0 esiste il limite del rapporto incrementale della $G(x)$ e, per confronto tra primo e ultimo membro della (8.10), tale limite è il numero reale $f(x_0)$. Concludiamo pertanto che la $G(x)$ risulta derivabile e la sua derivata nel punto x_0 coincide con il risultato $f(x_0)$ del limite, ovvero abbiamo ottenuto la tesi $G'(x) = f(x)$, in cui è stato omesso l'indice in x_0, *c.d.d.*

fig. 5.5



E' immediato osservare che se una funzione $G(x)$ soddisfa la condizione $G'(x) = f(x)$, tale condizione risulta soddisfatta anche da tutte le funzioni $\widetilde{G}(x) = G(x) + k$, con k numero reale, perché si ha $\widetilde{G}'(x) = G'(x) = f(x)$. Quindi possiamo concludere che le funzioni $\widetilde{G}(x)$ che soddisfano la condizione $\widetilde{G}'(x) = f(x)$ sono infinite e che la differenza, indicata con $H(x)$, tra due qualsiasi di queste $\widetilde{G}(x)$, indicate con $\widetilde{G}_1(x)$ e $\widetilde{G}_2(x)$, ovvero $H(x) = \widetilde{G}_1(x) - \widetilde{G}_2(x)$, è sempre uguale ad un numero reale perché se vale $\widetilde{G}_1'(x) = \widetilde{G}_2'(x)$, segue $H'(x) = \widetilde{G}_1'(x) - \widetilde{G}_2'(x) = 0$ da cui si ottiene quindi che la $H(x)$, avendo derivata nulla, è appunto un numero reale.

Definizione 8.2.1. *Ogni funzione $\widetilde{G}(x)$ che soddisfa la condizione $\widetilde{G}'(x) = f(x)$ prende il nome di "funzione primitiva" o semplicemente "primitiva" della funzione $f(x)$.*

L'insieme di tutte le funzioni $\widetilde{G}(x)$ che soddisfano la condizione $\widetilde{G}'(x) = f(x)$ prende il nome di "integrale indefinito" della funzione $f(x)$ e si scrive

$$\widetilde{G}(x) = \int f(x)\, dx + k. \tag{8.11}$$

L'integrale al primo membro nella (8.9) *prende invece il nome di "integrale definito" della funzione $f(x)$ tra gli estremi a e b, a cui è associato l'integrale indefinito* (8.11).

Sebbene le funzioni primitive $\widetilde{G}(x)$ di una data funzione $f(x)$ siano infinite, tuttavia il risultato al secondo membro della formula generale (8.9) risulta indipendente dalla scelta della particolare primitiva $\widetilde{G}(x)$ perché se scegliamo un qualsiasi valore di $k \in \mathbb{R}$ e poniamo $\widetilde{G}(x) = G(x) + k$, dal secondo membro della formula generale (8.9) otteniamo sempre il medesimo risultato $\widetilde{G}(b) - \widetilde{G}(a) = [G(b) + k] - [G(a) + k] = G(b) - G(a)$.

Quindi, nel secondo membro della formula generale (8.9) del calcolo integrale è del tutto equivalente scrivere $\widetilde{G}(b) - \widetilde{G}(a)$ oppure $G(b) - G(a)$. Possiamo allora enunciare la formula generale (8.9) del calcolo integrale nel seguente modo. Per calcolare l'integrale definito riportato al primo membro della (8.9), occorre prima risolvere il corrispondente integrale indefinito (8.11), in modo tale che poi l'integrale definito (8.9) sia uguale alla differenza dei valori forniti da una qualsiasi delle primitive $\widetilde{G}(x)$ sui due estremi.

Riguardo al segno, è immediato rendersi conto che per funzioni $f(x)$ positive sull'intervallo $[a, b]$ l'integrale (8.9), ovvero l'integrale (8.1), risulta positivo perché dalla relazione $\widetilde{G}'(x) = G'(x) = f(x) > 0$ segue che ogni primitiva $\widetilde{G}(x)$ è crescente in $[a, b]$, ovvero $\widetilde{G}(b) - \widetilde{G}(a) = G(b) - G(a) > 0$. Quindi possiamo concludere che la definizione iniziale 8.1.1 di integrale è, come si dice in Matematica, *ben posta*, in quanto coerente con la proprietà geometrica per cui un'area deve essere espressa tramite un numero positivo.

Infine osserviamo che in virtù della tesi del teorema di Torricelli-Barrow, l'integrale indefinito (8.11) gode di quella che viene denominata *proprietà di linearità*, espressa dalla relazione valida per ogni coppia $\alpha_1, \alpha_2 \in \mathbb{R}$ e per ogni coppia di funzioni $f_1(x), f_2(x)$

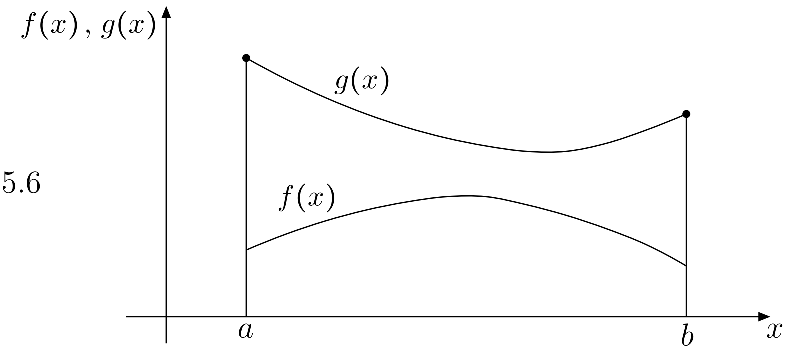
$$\int [\alpha_1 f_1(x) + \alpha_2 f_2(x)]\, dx = \alpha_1 \int f_1(x)\, dx + \alpha_2 \int f_2(x)\, dx. \tag{8.12}$$

Gli integrali di funzioni positive e continue nell'intervallo $[a, b]$ godono dell'ovvia proprietà secondo cui se risulta $0 < f(x) < g(x)$ per ogni $x \in (a, b)$, segue

$$\int_a^b f(x)\, dx < \int_a^b g(x)\, dx, \tag{8.13}$$

perché, come mostrato in fig. 5.6, l'area della regione piana delimitata dalla funzione $f(x)$ è "palesemente" minore dell'area della regione piana delimitata dalla funzione $g(x)$.

fig. 5.6



8.2.1 Area con segno

Se una funzione $f(x)$, continua sull'intervallo chiuso e limitato $[a, b]$, ha segno negativo sull'intervallo, segue che il suo integrale definito (8.9) risulta negativo perché la funzione primitiva $\widetilde{G}(x)$ è decrescente e si ha $\widetilde{G}(b) - \widetilde{G}(a) = G(b) - G(a) < 0$. Per generalizzare a funzioni negative la definizione 8.1.1 di integrale come area di una regione piana limitata, introduciamo il concetto di *area con segno*, riferendoci alla fig. 5.7(a).

Definizione 8.2.2. *Data una funzione $f(x)$ negativa e continua nell'intervallo chiuso e limitato $[a, b]$, il suo integrale sull'intervallo è definito dalla relazione*

$$\int_a^b f(x)\,dx \equiv -\int_a^b |f(x)|\,dx \tag{8.14}$$

e fornisce il valore $-\mathcal{A}(\Omega)$ uguale all'opposto dell'area geometrica positiva $\mathcal{A}(\Omega)$ della regione piana Ω rappresentata in fig. 5.7(a), ovvero

$$\int_a^b f(x)\,dx \equiv -\int_a^b |f(x)|\,dx = -\mathcal{A}(\Omega).$$

Poiché l'integrale (8.14) dà come risultato un numero negativo, chiamiamo *area con segno* il risultato, positivo per funzioni positive oppure negativo per funzioni negative, dell'integrale (8.9). Quindi, se la funzione $f(x)$ è positiva sull'intervallo chiuso $[a, b]$, il suo integrale (8.9) risulta positivo e diremo che l'area della regione piana Ω, riportata in fig. 5.1(a), è positiva e coincidente con l'area geometrica $\mathcal{A}(\Omega)$ della regione Ω; se invece la funzione $f(x)$ è negativa sull'intervallo chiuso $[a, b]$, il suo integrale (8.9) risulta negativo e diremo che l'area della regione piana Ω, riportata in fig. 5.7(a), è negativa e coincidente con $-\mathcal{A}(\Omega)$. Per definire l'integrale nel caso di funzione continua sia positiva che negativa sull'intervallo $[a, b]$, come mostrato in fig. 5.7(b), l'intervallo $[a, b]$ deve essere suddiviso in sottintervalli disgiunti su ciascuno dei quali la funzione integranda abbia un solo segno oppure sia nulla. In particolare, nel caso riportato in fig. 5.7(b) la funzione $f(x)$ risulta positiva o nulla in $[a, x_0]$ e negativa oppure nulla in $[x_0, b]$. Quindi il suo integrale definito (8.9) su tutto l'intervallo $[a, b]$ viene definito, sulla base della proprietà (8.3) di additività, come somma dell'integrale sul sottintervallo $[a, x_0]$ in cui la $f(x)$ è positiva e dell'integrale, scritto secondo la (8.14), sul sottintervallo $[x_0, b]$ in cui la $f(x)$ è negativa.

Abbiamo dunque

$$\int_a^b f(x)\,dx = \int_a^{x_0} f(x)\,dx + \int_{x_0}^b f(x)\,dx = \int_a^{x_0} f(x)\,dx - \int_{x_0}^b |f(x)|\,dx, \tag{8.15}$$

ovvero, se una funzione $f(x)$ continua in un intervallo $[a, b]$ cambia segno su tale intervallo, il suo integrale su tutto l'intervallo $[a, b]$, definito come estensione della (8.15), risulta uguale alla somma delle *aree con segno* positivo e delle *aree con segno* negativo.

In particolare, l'integrale (8.15) riferito alla $f(x)$ riportata in fig. 5.7(b) fornisce risultato $\mathcal{A}(\Omega_1) - \mathcal{A}(\Omega_2)$ che è diverso dall'area geometrica complessiva $\mathcal{A}(\Omega_1) + \mathcal{A}(\Omega_2)$ dell'*unione insiemistica* $\Omega_1 \cup \Omega_2$ delle due regioni piane Ω_1 e Ω_2.

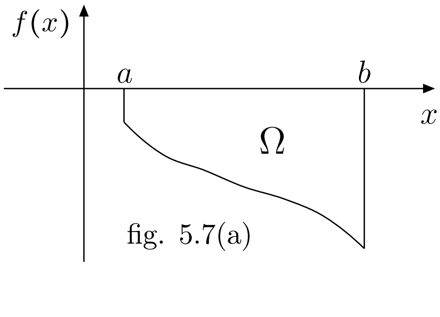


fig. 5.7(a)

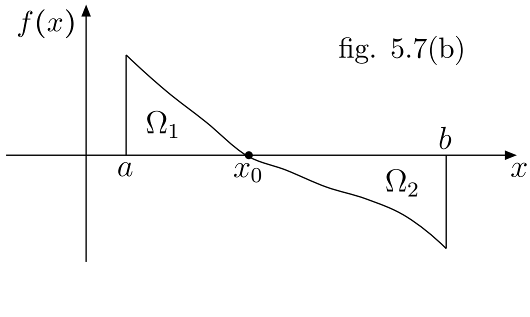


fig. 5.7(b)

8.3 Calcolo di funzioni primitive

Poiché il valore dell'integrale definito al primo membro della (8.9) si determina a partire da una funzione primitiva (8.11) tramite la quale si esegua $\widetilde{G}(b) - \widetilde{G}(a)$, segue che il calcolo degli integrali definiti consiste nel calcolo delle funzioni primitive delle funzioni integrande e quindi presentiamo alcuni metodi per ottenere gli integrali indefiniti (8.11).

La determinazione degli integrali indefiniti consiste nel ricondurre l'integrale indefinito da calcolare ai seguenti integrali indefiniti

$$1)\ \int x^n\,dx = \frac{x^{n+1}}{n+1} + k, \quad \forall n \neq -1; \qquad 2)\ \int \frac{dx}{ax+b} = \frac{1}{a}\log|ax+b| + k;$$

$$3)\ \int e^{ax}\,dx = \frac{e^{ax}}{a} + k; \qquad 4)\ \int \sin(ax+b)\,dx = -\frac{1}{a}\cos(ax+b) + k;$$

$$5)\ \int \cos(ax+b)\,dx = \frac{1}{a}\sin(ax+b) + k; \qquad 6)\ \int a\,dx = ax + k;$$

$$7)\ \text{se } b^2 - 4ac = \Delta < 0, \quad \int \frac{dx}{ax^2+bx+c} = \frac{2}{\sqrt{4ac-b^2}}\arctan\left(\frac{2ax+b}{\sqrt{4ac-b^2}}\right) + k,$$

che vengono denominati *integrali immediati* perché si verifica "immediatamente" che le derivate delle primitive coincidono effettivamente con le funzioni integrande corrispondenti. Può essere utile osservare che dal primo integrale immediato discendono gli altri due integrali immediati associati

$$1')\ \int (x-x_0)^n\,dx = \frac{(x-x_0)^{n+1}}{n+1} + k, \quad \forall n > 0;$$

$$1'')\ \int \frac{dx}{(x-x_0)^n} = -\frac{1}{(n-1)(x-x_0)^{n-1}} + k, \quad \forall n > 0, \text{ con } n \neq 1.$$

Può essere utile sottolineare inoltre che se il denominatore del settimo integrale immediato fosse in particolare $x^2 + a^2$, l'integrale immediato assumerebbe la forma

$$7')\ \int \frac{dx}{x^2+a^2} = \frac{1}{a}\arctan\left(\frac{x}{a}\right) + k.$$

Per ricondurre un integrale indefinito a quelli immediati che abbiamo riportato, presentiamo due metodi: il *metodo di sostituzione* e il *metodo per parti.*

8.3.1 Integrale di frazioni algebriche

Una *frazione algebrica* è una frazione avente un polinomio al numeratore e un polinomio al denominatore, indicati rispettivamente con $\mathcal{P}_m(x)$ e $\mathcal{Q}_n(x)$, dove i due indici m, n indicano il grado del polinomio. In particolare, ma in realtà senza alcuna perdita di generalità, considereremo solo frazioni algebriche non semplificabili, ovvero frazioni algebriche tali che i polinomi $\mathcal{P}_m(x)$ e $\mathcal{Q}_n(x)$ non abbiano nessuno *zero* in comune.

Per mostrare come si determina la funzione *primitiva* di una *frazione algebrica* avente dunque espressione $\mathcal{P}_m(x)/\mathcal{Q}_n(x)$, distinguiamo i due casi $m < n$ e $m \geqslant n$ nei quali, per semplicità, ci limitiamo a considerare al denominatore solo polinomi $\mathcal{Q}_n(x)$ i cui *zeri* abbiano tutti, nella scomposizione (4.3), *molteplicità algebrica* massima uguale a 2.

Caso dei gradi m < n e denominatore con tutti zeri semplici.

Dati due polinomi $\mathcal{P}_m(x)$ e $\mathcal{Q}_n(x)$, di gradi $m < n$ e con $\mathcal{Q}_n(x)$ avente *n zeri semplici*, indicati con $x_1, x_2, x_3 \ldots, x_n$, dimostriamo che vale la seguente identità

$$\frac{\mathcal{P}_m(x)}{\mathcal{Q}_n(x)} = \frac{a_1}{x - x_1} + \frac{a_2}{x - x_2} + \frac{a_3}{x - x_3} + \cdots + \frac{a_n}{x - x_n}, \tag{8.16}$$

in cui ciascun coefficiente a_i è dato dal rapporto

$$a_i = \frac{\mathcal{P}_m(x_i)}{\mathcal{Q}'_n(x_i)}, \tag{8.17}$$

dove $\mathcal{Q}'_n(x_i)$ indica la derivata del polinomio $\mathcal{Q}_n(x)$ calcolata nello *zero semplice* x_i.

Infatti, se riscriviamo l'identità (8.16)

$$\frac{\mathcal{P}_m(x)}{\mathcal{Q}_n(x)} = \frac{a_1}{x - x_1} + \frac{a_2}{x - x_2} + \cdots + \frac{a_i}{x - x_i} + \cdots + \frac{a_n}{x - x_n}$$

e moltiplichiamo ambo i membri per il binomio $x - x_i$, otteniamo

$$\mathcal{P}_m(x)\left[\frac{x - x_i}{\mathcal{Q}_n(x)}\right] = \frac{a_1(x - x_i)}{x - x_1} + \frac{a_2(x - x_i)}{x - x_2} + \cdots + a_i + \cdots + \frac{a_n(x - x_i)}{x - x_n}. \tag{8.18}$$

Poiché il polinomio $\mathcal{Q}_n(x)$ si annulla per $x = x_i$, segue che il limite per $x \to x_i$ della parentesi quadra al primo membro nella relazione (8.18) si ottiene tramite il teorema di De L'Hôspital e quindi il limite per $x \to x_i$ di tutto il primo membro è

$$\lim_{x \to x_i} \mathcal{P}_m(x)\left[\frac{x - x_i}{\mathcal{Q}_n(x)}\right] = \left[\lim_{x \to x_i} \mathcal{P}_m(x)\right]\left[\lim_{x \to x_i} \frac{x - x_i}{\mathcal{Q}_n(x)}\right] = \frac{\mathcal{P}_m(x_i)}{\mathcal{Q}'_n(x_i)}.$$

Poiché il limite per x tendente al valore x_i del secondo membro nella relazione (8.18) è il coefficiente a_i, segue, dall'uguaglianza dei due limiti, la relazione (8.17).

Vediamo quindi un'applicazione della formula (8.17) per il calcolo dei coefficienti costanti $a_1, a_2, a_3, \ldots, a_n$ ai numeratori del secondo membro nell'uguaglianza (8.16). Supponendo di voler determinare la decomposizione (8.16) nel caso della frazione algebrica

$$\frac{3x^2 + 5x - 4}{2x^3 + x^2 - 5x + 2},$$

il cui denominatore ha derivata $6x^2 + 2x - 5$, determiniamo gli zeri $x_1 = 1, x_2 = -2, x_3 = 1/2$ del denominatore e quindi poniamo

$$\frac{3x^2 + 5x - 4}{2x^3 + x^2 - 5x + 2} = \frac{a_1}{x - 1} + \frac{a_2}{x + 2} + \frac{a_3}{x - 1/2}, \tag{8.19}$$

da cui, in virtù della formula (8.17), seguono immediatamente i tre coefficienti

$$a_1 = \left.\frac{3x^2 + 5x - 4}{6x^2 + 2x - 5}\right|_{x=1} = \frac{4}{3}, \qquad a_2 = \left.\frac{3x^2 + 5x - 4}{6x^2 + 2x - 5}\right|_{x=-2} = -\frac{2}{15},$$

$$a_3 = \left.\frac{3x^2 + 5x - 4}{6x^2 + 2x - 5}\right|_{x=1/2} = \frac{3}{10}$$

e quindi

$$\frac{3x^2 + 5x - 4}{2x^3 - x^2 - 5x + 2} = \frac{4/3}{x - 1} - \frac{2/15}{x + 2} + \frac{3/10}{x - 1/2}.$$

Per calcolare infine l'integrale indefinito della frazione algebrica al primo membro nell'uguaglianza (8.19), basta integrare la somma di integrali immediati

$$\int \frac{3x^2+5x-4}{2x^3-x^2-5x+2}\,dx = \frac{4}{3}\int\frac{dx}{x-1} - \frac{2}{15}\int\frac{dx}{x+2} + \frac{3}{10}\int\frac{dx}{x-1/2}\,.$$

Caso dei gradi m < n e denominatore con *zeri* di *molteplicità algebrica* 2.
Dati due polinomi $\mathcal{P}_m(x)$ e $\mathcal{Q}_n(x)$ con $m<n$ e con $\mathcal{Q}_n(x)$ avente uno *zero*, indicato con x_1, di *molteplicità algebrica* 2, dimostriamo che vale la seguente identità

$$\frac{\mathcal{P}_m(x)}{\mathcal{Q}_n(x)} = \frac{a_1^{(2)}}{(x-x_1)^2} + \frac{a_1^{(1)}}{x-x_1} + f(x)\,, \tag{8.20}$$

dove $f(x)$ è una frazione algebrica che per x tendente a x_1 non tende a zero né tende ad infinito, e i coefficienti $a_1^{(2)}, a_1^{(1)}$ assumono i valori

$$a_1^{(2)} = \frac{2\mathcal{P}_m(x_1)}{\mathcal{Q}_n''(x_1)} \qquad \text{e} \qquad a_1^{(1)} = \frac{2\mathcal{P}_m'(x_1)}{\mathcal{Q}_n''(x_1)} - \frac{2\mathcal{P}_m(x_1)\,\mathcal{Q}_n'''(x_1)}{3[\mathcal{Q}_n''(x_1)]^2}\,. \tag{8.21}$$

Infatti, se riscriviamo l'identità (8.20) moltiplicando ambo i membri per $(x-x_1)^2$, otteniamo

$$\mathcal{P}_m(x)\left[\frac{(x-x_1)^2}{\mathcal{Q}_n(x)}\right] = a_1^{(2)} + a_1^{(1)}(x-x_1) + f(x)(x-x_1)^2, \tag{8.22}$$

dove $f(x)$ è una frazione algebrica che non si annulla né risulta *singolare* per $x=x_1$ e, per ipotesi, il polinomio $\mathcal{Q}_n(x)$ è infinitesimo di secondo ordine per $x=x_1$. Poiché dunque il limite per $x\to x_1$ della parentesi quadra, avente numeratore e denominatore infinitesimi di secondo ordine, al primo membro della relazione (8.22) si ottiene applicando il teorema di De L'Hôspital due volte, segue che il limite di tutto il primo membro per $x\to x_1$ è

$$\lim_{x\to x_1}\mathcal{P}_m(x)\left[\frac{(x-x_1)^2}{\mathcal{Q}_n(x)}\right] = \frac{2\mathcal{P}_m(x_1)}{\mathcal{Q}_n''(x_1)}\,.$$

Poiché il limite per $x\to x_1$ del secondo membro nella relazione (8.22) coincide con il coefficiente $a_1^{(2)}$, segue, dall'uguaglianza dei due limiti, la prima delle due relazioni (8.21).
Se a questo punto deriviamo ambo i membri dell'identità (8.22), otteniamo

$$\begin{aligned}\mathcal{P}_m'(x)\left[\frac{(x-x_1)^2}{\mathcal{Q}_n(x)}\right] + \mathcal{P}_m(x)\left[\frac{(x-x_1)^2}{\mathcal{Q}_n(x)}\right]\left[\frac{2\mathcal{Q}_n(x)-\mathcal{Q}_n'(x)(x-x_1)}{\mathcal{Q}_n(x)(x-x_1)}\right] &= \\ = a_1^{(1)} + f'(x)(x-x_1)^2 + 2f(x)(x-x_1).\end{aligned} \tag{8.23}$$

Poiché l'ultima parentesi quadra al primo membro della (8.23) contiene una frazione avente numeratore e denominatore infinitesimi di terzo ordine per $x=x_1$, segue che il limite per x tendente al valore x_1 appunto dell'ultima parentesi quadra al primo membro si ottiene applicando il teorema di De L'Hôspital tre volte. Poiché il limite per x tendente al valore x_1 del secondo membro della (8.23) è $a_1^{(1)}$, segue, dall'uguaglianza dei limiti dei due membri della (8.23) per $x\to x_1$, la seconda delle relazioni (8.21).
Vediamo quindi un'applicazione delle formule (8.21) per il calcolo dei coefficienti costanti $a_1^{(2)}, a_1^{(1)}$ ai numeratori del secondo membro nell'uguaglianza (8.20).

Per determinare la decomposizione (8.20) relativo alla frazione algebrica

$$\frac{\mathcal{P}_2(x)}{\mathcal{Q}_3(x)} = \frac{3x^2 - 9x - 1}{2x^3 - 11x^2 + 12x + 9}\,,$$

il cui denominatore possiede lo *zero semplice* $x_1 = -1/2$ e lo *zero* $x_2 = 3$ avente molteplicità algebrica 2, poniamo

$$\frac{3x^2 - 9x - 1}{2x^3 - 11x^2 + 12x + 9} = \frac{a_1}{x + 1/2} + \frac{a_2^{(2)}}{(x-3)^2} + \frac{a_2^{(1)}}{x-3}$$

e applichiamo la formula (8.17) per il coefficiente a_1 "sopra" il binomio $x + 1/2$ e le due formule (8.21) per i coefficienti $a_2^{(2)}, a_2^{(1)}$. Dalle derivate

$$\mathcal{P}_2'(x) = 6x - 9, \qquad \mathcal{Q}_3'(x) = 6x^2 - 22x + 12, \qquad \mathcal{Q}_3''(x) = 12x - 22, \qquad \mathcal{Q}_3'''(x) = 12,$$

otteniamo quindi

$$a_1 = \left.\frac{\mathcal{P}_2(x)}{\mathcal{Q}_3'(x)}\right|_{x=-1/2} = \frac{17}{98}\,, \qquad a_2^{(2)} = \left.\frac{2\mathcal{P}_2(x)}{\mathcal{Q}_3''(x)}\right|_{x=3} = -\frac{1}{7}\,,$$

$$a_2^{(1)} = \left[\frac{2\mathcal{P}_2'(x)}{\mathcal{Q}_3''(x)} - \frac{2\mathcal{P}_2(x)\,\mathcal{Q}_3'''(x)}{3[\mathcal{Q}_3''(x)]^2}\right]_{x=3} = \frac{65}{49}\,,$$

da cui segue infine

$$\frac{3x^2 - 9x - 1}{2x^3 - 11x^2 + 12x + 9} = \frac{17/98}{x + 1/2} - \frac{1/7}{(x-3)^2} + \frac{65/49}{x-3}$$

e quindi

$$\int \frac{3x^2 - 9x - 1}{2x^3 - 11x^2 + 12x + 9}\,dx = \frac{17}{98}\int \frac{dx}{x + 1/2} - \frac{1}{7}\int \frac{dx}{(x-3)^2} + \frac{65}{49}\int \frac{dx}{x-3} =$$
$$= \frac{17}{98}\log\left|x + \frac{1}{2}\right| + \frac{1}{7(x-3)} + \frac{65}{49}\log|x - 3| + k.$$

Caso dei gradi $\mathbf{m \geqslant n}$

Dati due polinomi $\mathcal{P}_m(x)$ e $\mathcal{Q}_n(x)$ con $m \geqslant n$ e con $\mathcal{Q}_n(x)$ avente *zeri* di *molteplicità algebrica* 1 oppure 2, abbiamo l'identità

$$\begin{aligned}\frac{\mathcal{P}_m(x)}{\mathcal{Q}_n(x)} &= B_{m-n}x^{m-n} + B_{m-n-1}x^{m-n-1} + B_{m-n-2}x^{m-n-2} + \cdots + B_1x + B_0 + \\ &\quad + \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2} + \cdots + \frac{A_3^{(2)}}{(x - x_3)^2} + \frac{A_3^{(1)}}{x - x_3} + \cdots,\end{aligned} \tag{8.24a}$$

dove il polinomio $B_{m-n}x^{m-n} + B_{m-n-1}x^{m-n-1} + B_{m-n-2}x^{m-n-2} + \cdots + B_1x + B_0$ ha grado uguale alla differenza tra i gradi m del numeratore ed n del denominatore, i coefficienti A si ottengono sempre tramite le relazioni (8.17) e (8.21), i coefficienti B si ottengono uguagliando i coefficienti delle $m - n + 1$ potenze "più alte" $x^m, x^{m-1}, x^{m-2}, \ldots, x^{n+1}, x^n$ nei due membri dell'equazione

$$\mathcal{P}_m(x) = \mathcal{Q}_n(x)\left[B_{m-n}x^{m-n} + B_{m-n-1}x^{m-n-1} + B_{m-n-2}x^{m-n-2} + \cdots + B_1x + B_0\right], \tag{8.24b}$$

osservando che il prodotto del polinomio $\mathcal{Q}_n(x)$ per ciascuna frazione avente un numeratore A_i è una frazione avente al numeratore un polinomio di grado sempre minore di n e ricordando che il generico polinomio di grado zero è dato da una costante diversa da zero.

Per illustrare qualche applicazione dell'identità (8.24a), consideriamo la frazione

$$\frac{\mathcal{P}_2(x)}{\mathcal{Q}_2(x)} = \frac{3x^2+1}{2x^2-x-6}\,, \tag{8.25}$$

nella quale la differenza tra il grado del numeratore e il grado del denominatore è $2-2=0$ e il denominatore $\mathcal{Q}_2(x)$ ha i due *zeri semplici* $x_1 = 2$ e $x_2 = -3/2$. In questo caso uguagliamo la frazione algebrica (8.25) alla somma

$$\frac{3x^2+1}{2x^2-x-1} = B + \frac{A_1}{x-2} + \frac{A_2}{x+3/2}\,,$$

dove il numero B rappresenta un polinomio di grado zero, ovvero di grado uguale alla differenza tra il grado del numeratore e il grado del denominatore nella frazione (8.25).

Applicando la formula (8.17) otteniamo i coefficienti

$$A_1 = \left.\frac{3x^2+1}{4x-1}\right|_{x=2} = \frac{13}{7} \qquad \text{e} \qquad A_2 = \left.\frac{3x^2+1}{4x-1}\right|_{x=-3/2} = -\frac{31}{28}$$

e dall'equazione (8.24b), la cui espressione relativa alla (8.25) è $3x^2+1 = (2x^2-x-1)\,B$, segue l'uguaglianza $3 = 2B$ tra i coefficienti delle potenze x^2 in ambo i membri, da cui si ricava il valore del polinomio di grado zero $B = 3/2$.

Pertanto la frazione (8.25) si decompone nella somma

$$\frac{3x^2+1}{2x^2-x-6} = \frac{3}{2} + \frac{13/7}{x-2} - \frac{31/28}{x+3/2}\,,$$

da cui segue dunque l'integrale indefinito

$$\int \frac{3x^2+1}{2x^2-x-6}\,dx = \frac{3}{2}\int dx + \frac{13}{7}\int \frac{dx}{x-2} - \frac{31}{28}\int \frac{dx}{x+3/2} =$$
$$= \frac{3}{2}\,x + \frac{13}{7}\log|x-2| - \frac{31}{28}\log\left|x+\frac{3}{2}\right| + k.$$

Se come secondo esempio consideriamo la frazione

$$\frac{\mathcal{P}_3(x)}{\mathcal{Q}_2(x)} = \frac{x^3-2x}{2x^2-x-6}\,, \tag{8.26}$$

nella quale la differenza tra il grado del numeratore e il grado del denominatore è $3-2=1$ e il denominatore $\mathcal{Q}_2(x)$ ha i due *zeri semplici* $x_1 = 2$ e $x_2 = -3/2$. In questo caso uguagliamo la frazione algebrica (8.26) alla somma

$$\frac{x^3-2x}{2x^2-x-1} = B_1x + B_0 + \frac{A_1}{x-2} + \frac{A_2}{x+3/2}\,,$$

dove il polinomio $B_1x + B_0$ rappresenta un polinomio di grado 1, ovvero di grado uguale alla differenza tra il grado del numeratore e il grado del denominatore nella frazione (8.26).

Applicando la formula (8.17) otteniamo i coefficienti

$$A_1 = \left.\frac{x^3-2x}{4x-1}\right|_{x=2} = \frac{4}{7} \qquad \text{e} \qquad A_2 = \left.\frac{x^3-2x}{4x-1}\right|_{x=-3/2} = -\frac{3}{56}$$

e dall'equazione (8.24b), avente espressione $x^3-2x = (2x^2-x-1)\,(B_1x+B_0)$ il cui secondo membro è della forma $x^3-2x = (2B_1)x^3 + (2B_0-B_1)x^2 + \alpha_1x + \alpha_0$, seguono dunque le due uguaglianze $1 = 2B_1$ tra i coefficienti delle potenze x^3 e $0 = -B_1 + 2B_0$, tra i coefficienti delle potenze x^2 in ambo i membri, da cui si ricavano i valori $B_1 = 1/2, B_0 = 1/4$.

Pertanto la frazione (8.26) si decompone nella somma

$$\frac{x^3 - 2x}{2x^2 - x - 6} = \frac{1}{2}\,x + \frac{1}{4} + \frac{4/7}{x-2} + \frac{3/56}{x+3/2}\,,$$

da cui segue dunque l'integrale indefinito

$$\int \frac{x^3 - 2x}{2x^2 - x - 6}\,dx = \int \left(\frac{1}{2}\,x + \frac{1}{4}\right) dx + \frac{4}{7}\int \frac{dx}{x-2} + \frac{3}{56}\int \frac{dx}{x+3/2} =$$

$$= \frac{1}{4}\,x^2 + \frac{1}{4}\,x + \frac{4}{7}\,\log|x-2| + \frac{3}{56}\,\log\left|x + \frac{3}{2}\right| + k.$$

Come ultimo esempio consideriamo la frazione

$$\frac{\mathcal{P}_5(x)}{\mathcal{Q}_3(x)} = \frac{x^5 - x^2/4}{2x^3 - 5x^2 + 4x - 1}\,, \tag{8.27}$$

nella quale la differenza tra il grado del numeratore e il grado del denominatore è $5-3 = 2$ e il denominatore $\mathcal{Q}_3(x)$ possiede lo *zero semplice* $x_1 = 1/2$ e lo *zero* $x_2 = 1$ avente *molteplicità algebrica* 2. In questo caso uguagliamo la frazione algebrica (8.27) alla somma

$$\frac{x^5 - x^2/4}{2x^3 - 5x^2 + 4x - 1} = B_2x^2 + B_1x + B_0 + \frac{A_1}{x - 1/2} + \frac{A_2^{(2)}}{(x-1)^2} + \frac{A_2^{(1)}}{x-1}\,,$$

dove il polinomio $B_2x^2 + B_1x + B_0$ rappresenta un polinomio di grado 2, ovvero di grado uguale alla differenza tra il grado del numeratore e il grado del denominatore nella (8.27).

Applicando la formula (8.17) otteniamo il coefficiente

$$A_1 = \left.\frac{x^5 - x^2/4}{6x^2 - 10x + 4}\right|_{x=1/2} = -\frac{1}{16}$$

e applicando le formule (8.21) otteniamo i coefficienti

$$A_2^{(2)} = \left.\frac{2\mathcal{P}_m(x)}{\mathcal{Q}_n''(x)}\right|_{x=1} = \left.\frac{2x^5 - x^2/2}{12x - 10}\right|_{x=1} = \frac{3}{4}\,,$$

$$A_2^{(1)} = \left.\frac{2\mathcal{P}_m'(x)}{\mathcal{Q}_n''(x)} - \frac{2\mathcal{P}_m(x)\,\mathcal{Q}_n'''(x)}{3[\mathcal{Q}_n''(x)]^2}\right|_{x=1} = \left.\frac{10x^4 - x}{12x - 10} - \frac{8x^5 - 2x^2}{(12x-10)^2}\right|_{x=1} = 3,$$

mentre dall'equazione (8.24b), la cui espressione relativa alla frazione algebrica (8.27) è

$$x^5 - x^2/4 = \left(2x^3 - 5x^2 + 4x - 1\right)(B_2x^2 + B_1x + B_0),$$

seguono l'uguaglianza $1 = 2B_2$ tra i coefficienti delle potenze x^5, l'uguaglianza $0 = 2B_1 - 5B_2$ tra i coefficienti delle potenze x^4 e l'uguaglianza $0 = 2B_0 - 5B_1 + 4B_2$ tra i coefficienti delle potenze x^3 in ambo i membri, da cui si ricavano i valori $B_2 = 1/2, B_1 = 5/4, B_0 = 17/4$.

Pertanto la frazione (8.27) si decompone nella somma

$$\frac{x^5 - x^2/4}{2x^3 - 5x^2 + 4x - 1} = \frac{1}{2}\,x^2 + \frac{5}{4}\,x + \frac{17}{4} - \frac{1/16}{x - 1/2} + \frac{3/4}{(x-1)^2} + \frac{3}{x-1}\,,$$

da cui segue dunque l'integrale indefinito

$$\int \frac{x^5 - x^2/4}{2x^3 - 5x^2 + 4x - 1}\, dx =$$

$$= \int \left(\frac{1}{2}x^2 + \frac{5}{4}x + \frac{17}{4}\right) dx - \frac{1}{16}\int \frac{dx}{x - 1/2} + \frac{3}{4}\int \frac{dx}{(x-1)^2} + 3\int \frac{dx}{x-1} =$$

$$= \frac{1}{6}x^3 + \frac{5}{8}x^2 + \frac{17}{4}x - \frac{1}{16}\log\left|x - \frac{1}{2}\right| - \frac{3}{4(x-1)} + 3\log|x-1| + k.$$

Concludiamo questa trattazione sulla decomposizione delle frazioni algebriche finalizzata alla ricerca delle loro funzioni *primitive* illustrando il caso delle tre frazioni algebriche particolari

$$\frac{x}{ax+b}, \qquad \frac{x^2}{ax+b}, \qquad \frac{x^3}{ax+b},$$

alle quali non applichiamo il procedimento generale che abbiamo presentato in precedenza, ma applichiamo un procedimento "semplificato". Ricordando i due prodotti notevoli

$$(C+D)(C-D) = C^2 - D^2 \qquad \text{e} \qquad (C+D)(C^2 - CD + D^2) = C^3 + D^3,$$

otteniamo le formule relative alle tre frazioni algebriche particolari

$$\frac{x}{ax+b} = \frac{1}{a}\left(\frac{ax}{ax+b}\right) = \frac{1}{a}\left[\frac{(ax+b) - b}{ax+b}\right] = \frac{1}{a}\left(1 - \frac{b}{ax+b}\right) = \frac{1}{a} - \frac{b/a}{ax+b},$$

$$\frac{x^2}{ax+b} = \frac{1}{a^2}\left(\frac{a^2x^2}{ax+b}\right) = \frac{1}{a^2}\left[\frac{(a^2x^2 - b^2) + b^2}{ax+b}\right] =$$

$$= \frac{1}{a^2}\left[\frac{(ax+b)(ax-b)}{ax+b} + \frac{b^2}{ax+b}\right] = \frac{x}{a} - \frac{b}{a^2} + \frac{b^2/a^2}{ax+b},$$

$$\frac{x^3}{ax+b} = \frac{1}{a^3}\left(\frac{a^3x^3}{ax+b}\right) = \frac{1}{a^3}\left[\frac{(a^3x^3 + b^3) - b^3}{ax+b}\right] =$$

$$= \frac{1}{a^3}\left[\frac{(ax+b)(a^2x^2 - abx + b^2)}{ax+b} - \frac{b^3}{ax+b}\right] = \frac{x^2}{a} - \frac{b}{a^2}x + \frac{b^2}{a^3} - \frac{b^3/a^3}{ax+b},$$

ovvero abbiamo le relazioni

$$\begin{gathered} 1) \quad \frac{x}{ax+b} = \frac{1}{a} - \frac{b/a}{ax+b}, \qquad 2) \quad \frac{x^2}{ax+b} = \frac{x}{a} - \frac{b}{a^2} + \frac{b^2/a^2}{ax+b}, \\ 3) \quad \frac{x^3}{ax+b} = \frac{x^2}{a} - \frac{b}{a^2}x + \frac{b^2}{a^3} - \frac{b^3/a^3}{ax+b}. \end{gathered} \tag{8.28}$$

8.3.2 Metodo di sostituzione

Fino ad ora il simbolo dx, chiamato *differenziale*, che compare nell'integrale non ha giocato nessun ruolo perché nella scrittura $f(x)dx$, la funzione integranda è solo la $f(x)$.

Poiché il significato e le proprietà formali rigorose del differenziale vengono studiate in quella branca della matematica chiamata *Geometria Differenziale*, che ovviamente esula dagli scopi del presente testo, ci limitiamo a giustificare in modo intuitivo e semplicistico le due proprietà del differenziale che sono indispensabili per l'applicazione di quel metodo di calcolo della primitiva denominato *metodo di sostituzione*. Se indichiamo l'operazione di derivata con il simbolo di Leibniz d/dx e riscriviamo quindi la tesi $G'(x) = f(x)$ del teorema di Torricelli-Barrow con la notazione

$$\frac{dG(x)}{dx} = f(x),$$

rimandando appunto alla *Geometria Differenziale* la dimostrazione che il simbolo della derivata si "comporta" come una frazione ordinaria il cui denominatore sia il differenziale della variabile, otteniamo invertendo $dG(x) = f(x)\,dx$, da cui segue appunto la relazione con gli integrali di ambo i membri

$$G(x) = \int dG(x) = \int f(x)\,dx.$$

Possiamo pertanto esprimere le due proprietà del differenziale necessarie per la corretta applicazione del metodo di sostituzione nel seguente modo:

i) il differenziale si "comporta" come un fattore moltiplicativo;

ii) il differenziale dx della variabile non modifica la funzione integranda e quindi non gioca nessun ruolo, mentre il differenziale $dh(x)$ di una generica funzione $h(x)$ (derivabile) si sviluppa secondo la relazione (6.2), in cui si ponga $x - x_0 \equiv dx$, ovvero $dh(x) = h'(x)\,dx$, in modo tale quindi che risulti $f(x)\,dh(x) = f(x)h'(x)\,dx$. Il metodo di sostituzione per il calcolo della primitiva dell'integrale

$$\int f(x)\,dx \tag{8.29}$$

consiste nello scegliere una funzione ausiliaria $g(t)$ della variabile t, invertibile nell'intervallo $[a, b]$, in modo tale che con la sostituzione $x = g(t)$, da cui segue lo sviluppo del differenziale $dx = dg(t) = g'(t)\,dt$, l'integrale iniziale (8.29) venga sostituito dal nuovo integrale

$$\int f(x)\,dx = \int [f(g(t))g'(t)]\,dt,$$

in cui la nuova funzione integranda è la funzione $f(g(t))g'(t)$ della nuova variabile t e il nuovo differenziale dt non gioca nessun ruolo perché è il differenziale della variabile.

Poiché la funzione $x = g(t)$ è invertibile, indichiamo la sua inversa con $t = g^{-1}(x)$ e l'insieme delle primitive della nuova funzione integranda $f(g(t))g'(t)$ con $G(t)$. Se quindi sostituiamo la $t = g^{-1}(x)$ nella primitiva $G(t)$ dell'integrale indefinito (8.29), otteniamo

$$\int f(x)\,dx = \int [f(g(t))g'(t)]\,dt = \left. G(t)\right|_{t=g^{-1}(x)} + k = G(g^{-1}(x)) + k.$$

Dalla primitiva $G(g^{-1}(x)) + k$ della $f(x)$, ottenuta con il metodo di sostituzione, ricaviamo quindi l'integrale definito (8.9) della $f(x)$ applicando il teorema di Torricelli-Barrow

$$\int_a^b f(x)\,dx \equiv \int_{g^{-1}(a)}^{g^{-1}(b)} [f(g(t))g'(t)]\,dt = G(g^{-1}(b)) - G(g^{-1}(a)), \tag{8.30}$$

in cui osserviamo che, per l'invertibilità della funzione di sostituzione $x = g(t)$, gli estremi d'integrazione a, b nell'integrale con la variabile x corrispondono agli estremi d'integrazione rispettivamente $g^{-1}(a), g^{-1}(b)$ nell'integrale con la nuova variabile t.

Esempio 1. Per calcolare l'integrale definito

$$\int_1^2 \frac{x+1}{(x^2+2x)^8}\,dx$$

effettuiamo la sostituzione $t = x^2 + 2x$, da cui otteniamo $dt = d(x^2+2x) = 2(x+1)\,dx$ e quindi $(x+1)\,dx = dt/2$. Inserendo tali espressioni nell'integrale, si ottiene

$$\int_1^2 \frac{x+1}{(x^2+2x)^8}\,dx = \int (x^2+2x)^{-8}\,(x+1)\,dx = \frac{1}{2}\int t^{-8}\,dt =$$

$$= \frac{t^{-7}}{2(-7)} + k = -\frac{1}{14t^7} + k = -\frac{1}{14(x^2+2x)^7}\Bigg|_1^2 = \left(-\frac{1}{14\cdot 8^7}\right) - \left(-\frac{1}{14\cdot 3^7}\right) = \frac{1}{14}\left(\frac{1}{3^7} - \frac{1}{8^7}\right).$$

Osserviamo che l'integrale definito può essere risolto anche mediante la formula (8.30) utilizzando la *funzione primitiva* con la variabile t, senza scrivere la *funzione primitiva* finale con la variabile x. A seguito della sostituzione effettuata $t = x^2 + 2x$, abbiamo che l'estremo $x = 1$ diventa $t = 3$, mentre l'estremo $x = 2$ diventa $t = 8$, da cui segue che l'integrale può essere calcolato sostituendo nella *funzione primitiva* con la variabile t gli estremi scritti rispetto alla variabile t stessa, ovvero

$$\int_1^2 \frac{x+1}{(x^2+2x)^8}\,dx = \frac{1}{2}\int_3^8 \frac{dt}{t^8}\,dt = -\frac{1}{14t^7} + k\Bigg|_3^8 = \left(-\frac{1}{14\cdot 8^7}\right) - \left(-\frac{1}{14\cdot 3^7}\right) = \frac{1}{14}\left(\frac{1}{3^7} - \frac{1}{8^7}\right).$$

Esempio 2. Per calcolare l'integrale definito

$$\int_1^4 \frac{e^{\sqrt{x}}}{\sqrt{x}}\,dx$$

effettuiamo la sostituzione $t = \sqrt{x}$, da cui otteniamo $dt = d(\sqrt{x}) = dx/(2\sqrt{x})$ e quindi $dx/\sqrt{x} = 2\,dt$. Inserendo tali espressioni nell'integrale, si ottiene

$$\int_1^4 \frac{e^{\sqrt{x}}}{\sqrt{x}}\,dx = \int_1^2 e^{\sqrt{x}}\,\frac{dx}{\sqrt{x}} = 2\int e^t\,dt = 2e^t + k = \ 2e^{\sqrt{x}}\Bigg|_1^4 = 2e^2 - 2e.$$

Per calcolare l'integrale sostituendo gli estremi rispetto alla variabile t nella *funzione primitiva* con la variabile t stessa, secondo la (8.30), osserviamo che, a seguito della sostituzione $t = \sqrt{x}$, l'estremo $x = 1$ diventa $t = 1$, mentre l'estremo $x = 4$ diventa $t = 2$, da cui segue

$$\int_1^4 \frac{e^{\sqrt{x}}}{\sqrt{x}}\,dx = 2\int_1^2 e^t\,dt = \ 2e^t + k\Bigg|_1^2 = 2e^2 - 2e.$$

Esempio 3. Per calcolare l'integrale indefinito

$$\int e^{x^3-3x}\left(x^2-1\right)\,dx$$

effettuiamo la sostituzione $t = x^3 - 3x$, da cui otteniamo $dt = d(x^3-3x) = 3(x^2-1)\,dx$ e quindi $(x^2-1)\,dx = dt/3$, da cui segue

$$\int e^{x^3-3x}\left(x^2-1\right)\,dx = \frac{1}{3}\int e^t\,dt = \frac{1}{3}\,e^t + k = \frac{1}{3}\,e^{x^3-3x} + k.$$

Esempio 4. Per calcolare l'integrale indefinito

$$\int \frac{\log x}{x\sqrt{2+5(\log x)^2}}\,dx$$

sostituiamo $t = 2+5(\log x)^2$, da cui otteniamo $dt = d[2+5(\log x)^2] = 10(\log x)\,dx/x$ e quindi $(\log x)\,dx/x = dt/10$, da cui segue

$$\int \frac{\log x}{x\sqrt{2+5(\log x)^2}}\,dx = \int \frac{1}{\sqrt{2+5(\log x)^2}}\,\frac{(\log x)\,dx}{x} = \frac{1}{10}\int \frac{1}{\sqrt{t}}\,dt = \frac{1}{10}\int t^{-1/2}\,dt =$$

$$= \frac{1}{10}\left(\frac{t^{1/2}}{1/2}\right) + k = \frac{1}{5}\sqrt{t} + k = \frac{1}{5}\sqrt{2+5(\log x)^2} + k.$$

Esempio 5. Per calcolare l'integrale indefinito

$$\int \frac{e^{-x}}{e^x+2}\,dx \tag{8.31}$$

calcoliamo preliminarmente l'integrale

$$\int \frac{dt}{t^3+2t^2}$$

con il metodo illustrato nel paragr. 8.3.1. Poiché il denominatore della frazione algebrica integranda possiede lo *zero semplice* $t=-2$ e lo *zero* $t=0$ di molteplicità algebrica 2, decomponiamo la frazione algebrica integranda ponendo

$$\frac{1}{t^3+2t^2} = \frac{1}{t^2(t+2)} = \frac{a_1}{t+2} + \frac{a_2^{(2)}}{t^2} + \frac{a_2^{(1)}}{t}$$

e applicando le formule (8.17), (8.21). In questo modo otteniamo

$$a_1 = \left.\frac{1}{3t^2+4t}\right|_{t=-2} = \frac{1}{4}, \qquad a_2^{(2)} = \left.\frac{2}{6t+4}\right|_{t=0} = \frac{1}{2}, \qquad a_2^{(1)} = -\left.\frac{12}{3(6t+4)^2}\right|_{t=0} = -\frac{1}{4},$$

da cui segue

$$\begin{aligned}\int \frac{dt}{t^3+2t^2} &= \int \left[\frac{1/4}{t+2} + \frac{1/2}{t^2} - \frac{1/4}{t}\right] dt = \frac{1}{4}\int \frac{dt}{t+2} + \frac{1}{2}\int \frac{dt}{t^2} - \frac{1}{4}\int \frac{dt}{t} = \\ &= \frac{1}{4}\log|t+2| + \frac{1}{2}\int t^{-2}\,dt - \frac{1}{4}\log|t| = \frac{1}{4}\log|t+2| - \frac{1}{2t} - \frac{1}{4}\log|t| + k.\end{aligned} \tag{8.32}$$

Per calcolare l'integrale iniziale (8.31) sostituiamo prima $e^x = t$, da cui segue $x = \log t$ e il differenziale $dx = dt/t$, e quindi il risultato (8.32). Otteniamo dunque

$$\int \frac{e^{-x}}{e^x+2}\,dx = \int \frac{(e^x)^{-1}}{e^x+2}\,dx = \int \left(\frac{t^{-1}}{t+2}\right)\frac{dt}{t} = \int \frac{dt}{t^3+2t^2} =$$

$$= \frac{1}{4}\log|t+2| - \frac{1}{2t} - \frac{1}{4}\log|t| + k = \frac{1}{4}\log(e^x+2) - \frac{e^{-x}}{2} - \frac{1}{4}x + k.$$

Esempio 6. Per calcolare l'integrale indefinito

$$\int \frac{x}{1+2\sqrt{x}}\,dx$$

sostituiamo $t = \sqrt{x}$, da cui otteniamo $x = t^2$ e $dx = d(t^2) = 2t\,dt$. Inserendo tali espressioni nell'integrale ed utilizzando la proprietà di linearità (8.12), si ottiene

$$\int \frac{x}{1+2\sqrt{x}}\,dx = \int \frac{t^2}{1+2t}\,2t\,dt = \int \frac{2t^3}{1+2t}\,dt = \frac{1}{4}\int \frac{8t^3}{1+2t}\,dt = \frac{1}{4}\int \frac{(1+8t^3)-1}{1+2t}\,dt =$$

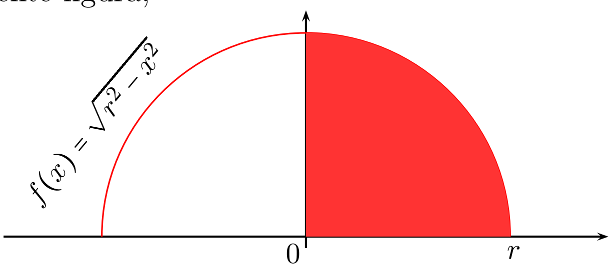
$$= \frac{1}{4}\int \left[\frac{(1+2t)(1-2t+4t^2)}{1+2t} - \frac{1}{1+2t}\right]dt = \frac{1}{4}\int \left[1-2t+4t^2 - \frac{1}{1+2t}\right]dt =$$

$$= \frac{1}{4}\left(t - t^2 + \frac{4}{3}t^3 - \frac{1}{2}\log|1+2t|\right) = \frac{\sqrt{x}}{4} - \frac{x}{4} + \frac{x\sqrt{x}}{3} - \frac{1}{8}\log\left(1+2\sqrt{x}\right) + k.$$

Esempio 7. Se dall'equazione implicita della circonferenza $x^2 + y^2 = r^2$, avente centro nell'origine del sistema di riferimento cartesiano e avente raggio r, esplicitiamo la variabile y, otteniamo la funzione $y = f(x)$ avente espressione esplicita

$$f(x) = \sqrt{r^2 - x^2},$$

riportata nella seguente figura,



il cui integrale

$$\frac{S}{4} = \int_0^r \sqrt{r^2 - x^2}\,dx \tag{8.33}$$

rappresenta dunque l'area $S/4$ della porzione di cerchio, rappresentata in *colore rosso*, che è uguale ad un quarto dell'area S di tutto il cerchio. Per risolvere l'integrale (8.33) effettuiamo la sostituzione $x = r\sin t$, da cui segue $\sqrt{r^2 - x^2} = r\cos t$ e $dx = r\cos t\,dt$.

Poiché con tale sostituzione gli estremi dell'integrale $x = 0, x = r$ relativi alla variabile x si trasformano rispettivamente negli estremi $t = 0$ e $t = \pi/2$ relativi alla t, otteniamo

$$\frac{S}{4} = \int_0^r \sqrt{r^2 - x^2}\,dx = \int_0^{\pi/2} \left(\sqrt{r^2 - r^2\sin^2 t}\right) r\cos t\,dt = \int_0^{\pi/2} \sqrt{r^2\left(1-\sin^2 t\right)}\ r\cos t\,dt =$$

$$= \int_0^{\pi/2} \sqrt{r^2\cos^2 t}\ r\cos t\,dt = \int_0^{\pi/2} r^2\cos^2 t\,dt = \frac{r^2}{2}\int_0^{\pi/2} (1+\cos 2t)\ dt =$$

$$= \frac{r^2}{2}t + \frac{r^2}{4}\sin 2t\Big|_0^{\pi/2} = \frac{\pi r^2}{4},$$

ovvero abbiamo ottenuto che un quarto dell'area del cerchio è $\pi r^2/4$, da cui segue, anche mediante il metodo del calcolo integrale, la "famosa" formula dell'area del cerchio $S = \pi r^2$.

Esempio 8. Per calcolare l'integrale

$$\int x^2 \sqrt{x-3}\,dx$$

effettuiamo la sostituzione $x-3=t^2$, da cui segue $t=\sqrt{x-3}, x=t^2+3, dx=2t\,dt$ e quindi

$$\int x^2\sqrt{x-3}\,dx=\int\left(t^2+3\right)^2 t\,2t\,dt=2\int\left(t^6+6t^4+9t^2\right)dt=\frac{2}{7}\,t^7+\frac{12}{5}\,t^5+6t^3+k=$$

$$=\frac{2}{7}\,(x-3)^3\sqrt{x-3}+\frac{12}{5}\,(x-3)^2\sqrt{x-3}+6(x-3)\sqrt{x-3}+k.$$

Esempio 9. Questo esempio mostra la grande importanza che riveste l'*identità di secondo grado* (4.1), in particolare, per applicare una "certa" sostituzione ad una categoria di funzioni integrande con *radicali*. Se consideriamo il seguente integrale

$$\int\frac{x^2}{\sqrt{x^2-2x+5}}\,dx, \tag{8.34}$$

"solo" l'*identità di secondo grado* (4.1) ci permette di determinare una sostituzione di variabile che "elimina" la radice quadrata e che trasforma la funzione integranda in una somma di integrali immediati. Applicando l'*identità di secondo grado* (4.1), possiamo riscrivere il radicando al denominatore nella forma $x^2-2x+5=(x-1)^2+4$, in virtù della quale effettuiamo la sostituzione tale che la base $x-1$ del quadrato sia uguale a $t-1/t$, ovvero

$$x=1+t-\frac{1}{t},$$

da cui segue

$$x^2=\left(1+t-\frac{1}{t}\right)^2=t^2+\frac{1}{t^2}+2t-\frac{2}{t}-1, \tag{8.35a}$$

$$\sqrt{x^2-2x+5}=\sqrt{\left(1+t-\frac{1}{t}\right)^2-2\left(1+t-\frac{1}{t}\right)+5}=\sqrt{\left(t+\frac{1}{t}\right)^2}=t+\frac{1}{t}=\frac{t^2+1}{t}, \tag{8.35b}$$

$$dx=D\left(1+t-\frac{1}{t}\right)dt=\left(1+\frac{1}{t^2}\right)dt=\frac{t^2+1}{t^2}\,dt. \tag{8.35c}$$

Inoltre, se sommiamo tra loro i primi membri e tra loro i secondi membri delle due uguaglianze

$$x=1+t-\frac{1}{t} \qquad \text{e} \qquad \sqrt{x^2-2x+5}=t+\frac{1}{t},$$

otteniamo la relazione

$$x+\sqrt{x^2-2x+5}=1+2t,$$

in virtù della quale possiamo esprimere

$$t=\frac{x-1+\sqrt{x^2-2x+5}}{2}. \tag{8.36}$$

Se a questo punto sostituiamo le espressioni (8.35) nell'integrale iniziale (8.34), otteniamo

$$\int\frac{x^2}{\sqrt{x^2-2x+5}}\,dx=\int\left(t^2+\frac{1}{t^2}+2t-\frac{2}{t}-1\right)\left(\frac{t}{t^2+1}\right)\left(\frac{t^2+1}{t^2}\right)dt=$$

$$= \int \left(t + \frac{1}{t^3} + 2 - \frac{2}{t^2} - \frac{1}{t} \right) dt = \frac{t^2}{2} - \frac{1}{2t^2} + 2t + \frac{2}{t} - \log|t| + k =$$

$$= \frac{t^2}{2} - \frac{1}{2t^2} + 2\left(t + \frac{1}{t} \right) - \log|t| + k =$$

$$= \frac{\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^2}{8} - \frac{2}{\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^2} +$$

$$+ 2\sqrt{x^2 - 2x + 5} - \log\left(x - 1 + \sqrt{x^2 - 2x + 5}\right) + k,$$

dove abbiamo utilizzato la (8.35b) per sostituire tutta la parentesi tonda $t + 1/t$ e dove abbiamo omesso l'addendo $\log 2$ che viene inglobato nella costante k addizionata.

Poiché la verifica che la derivata della funzione primitiva ottenuta coincida con la funzione integranda risulta, in questo caso, "un po' intricata", mostriamo che appunto la derivata, indicata con d/dx, della funzione primitiva ottenuta coincide effettivamente con la funzione integranda, ovvero che la funzione primitiva ottenuta è corretta.

Utilizzando la derivata

$$\frac{d}{dx}\sqrt{x^2 - 2x + 5} = \frac{d}{dx}(x^2 - 2x + 5)^{1/2} = \frac{1}{2}(x^2 - 2x + 5)^{-1/2}(2x - 2) = \frac{x - 1}{\sqrt{x^2 - 2x + 5}}$$

e ponendo ad un certo punto "per comodità di scrittura" $\sqrt{x^2 - 2x + 5} \equiv \sqrt{R}$, ricaviamo dunque

$$\frac{d}{dx}\left[\frac{\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^2}{8} - 2\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^{-2} + \right.$$

$$\left. + 2\sqrt{x^2 - 2x + 5} - \log\left(x - 1 + \sqrt{x^2 - 2x + 5}\right) + k \right] =$$

$$= \left(\frac{x - 1 + \sqrt{x^2 - 2x + 5}}{4} \right)\left(1 + \frac{x - 1}{\sqrt{x^2 - 2x + 5}} \right) +$$

$$+ \left[\frac{4}{\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^3} \right]\left(1 + \frac{x - 1}{\sqrt{x^2 - 2x + 5}} \right) +$$

$$+ \frac{2x - 2}{\sqrt{x^2 - 2x + 5}} - \left(\frac{1}{x - 1 + \sqrt{x^2 - 2x + 5}} \right)\left(1 + \frac{x - 1}{\sqrt{x^2 - 2x + 5}} \right) =$$

$$= \left(1 + \frac{x - 1}{\sqrt{x^2 - 2x + 5}} \right)\left[\frac{\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^4 + 16}{4\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^3} \right] + \frac{2x - 3}{\sqrt{x^2 - 2x + 5}} =$$

$$= \left[\frac{\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^4 + 16}{4\sqrt{x^2 - 2x + 5}\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^2} \right] + \frac{2x - 3}{\sqrt{x^2 - 2x + 5}} =$$

$$= \frac{\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^4 + 16 + \left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^2 (8x - 12)}{4\sqrt{x^2 - 2x + 5}\left(x - 1 + \sqrt{x^2 - 2x + 5}\right)^2} =$$

$$= \frac{\left(x-1+\sqrt{x^2-2x+5}\right)^2\left[\left(x-1+\sqrt{x^2-2x+5}\right)^2+8x-12\right]+16}{4\sqrt{x^2-2x+5}\left(x-1+\sqrt{x^2-2x+5}\right)^2} =$$

$$= \frac{\left[\left(x^2+x\sqrt{R}-\sqrt{R}\right)-(2x-3)\right]\left[\left(x^2+x\sqrt{R}-\sqrt{R}\right)+(2x-3)\right]+4}{\sqrt{x^2-2x+5}\left(x-1+\sqrt{x^2-2x+5}\right)^2} =$$

$$= \frac{\left(x^2+x\sqrt{x^2-2x+5}-\sqrt{x^2-2x+5}\right)^2-(2x-3)^2+4}{\sqrt{x^2-2x+5}\left(x-1+\sqrt{x^2-2x+5}\right)^2} =$$

$$= \frac{x^2\left(x-1+\sqrt{x^2-2x+5}\right)^2}{\sqrt{x^2-2x+5}\left(x-1+\sqrt{x^2-2x+5}\right)^2} = \frac{x^2}{\sqrt{x^2-2x+5}}.$$

Esercizio da svolgere. Applicando le formule (8.21), determinare il valore dei coefficienti $a_1^{(2)}, a_1^{(1)}, a_2^{(2)}, a_2^{(2)}$ nella decomposizione

$$\frac{t^2}{(t^2-1)^2} = \frac{t^2}{(t-1)^2(t+1)^2} = \frac{a_1^{(2)}}{(t-1)^2} + \frac{a_1^{(1)}}{t-1} + \frac{a_2^{(2)}}{(t+1)^2} + \frac{a_2^{(1)}}{t+1}$$

e quindi risolvere l'integrale

$$\int \sqrt{\frac{x}{x-1}}\, dx$$

effettuando la sostituzione $x/(x-1) = t^2$, da cui segue

$$\frac{x}{x-1} = t^2 \quad\Longrightarrow\quad \frac{x-1}{x} = \frac{1}{t^2} \quad\Longrightarrow\quad x = \frac{t^2}{t^2-1} \quad\Longrightarrow\quad dx = -\frac{2t}{(t^2-1)^2}\, dt.$$

8.3.3 Metodo per parti

Il secondo e ultimo metodo che illustriamo per il calcolo delle primitive è quello che prende il nome di *metodo per parti*. Se utilizziamo la regola di derivazione del prodotto di due funzioni $D[f(x)g(x)] = g(x)Df(x) + f(x)Dg(x)$, in cui D indica l'operazione di derivata che agisce a destra, ed esplicitiamo il secondo addendo del secondo membro

$$f(x)Dg(x) = D[f(x)g(x)] - g(x)Df(x),$$

otteniamo, integrando ambo i membri, l'identità fondamentale del *metodo per parti*

$$\int f(x)Dg(x)\, dx = f(x)g(x) - \int g(x)Df(x)\, dx, \tag{8.37}$$

dove è stata utilizzata la relazione

$$\int D[f(x)g(x)]\, dx = f(x)g(x)$$

senza la costante additiva k perché al secondo membro della (8.37) basta ovviamente aggiungere una sola costante k dopo l'integrale di $g(x)Df(x)$.

Esempio 1. Calcoliamo i seguenti integrali definiti del prodotto della funzione esponenziale e^x per una funzione potenza x^n con $n = 0, 1, 2, 3$.

$$a)\ \int_0^1 e^x\,dx, \qquad b)\ \int_0^1 xe^x\,dx, \qquad c)\ \int_0^1 x^2e^x\,dx, \qquad d)\ \int_0^1 x^3e^x\,dx.$$

Applicando l'identità (8.37) del metodo per parti e gli integrali immediati, otteniamo

$$a) \quad \int_0^1 e^x\,dx = \left. e^x \right|_0^1 = e - 1,$$

$$b) \quad \int_0^1 x\,e^x\,dx = \int x\,D\,e^x\,dx = x\,e^x - \int e^x\,Dx\,dx = x\,e^x - \int e^x\,dx =$$
$$= \left. x\,e^x - e^x + k \right|_0^1 = (e-e) - (-1) = 1$$

e sostituendo il risultato di ogni integrale nell'integrale con la potenza successiva, ricaviamo

$$c) \quad \int_0^1 x^2\,e^x\,dx = \int x^2\,D\,e^x\,dx = x^2\,e^x - \int e^x\,D(x^2)\,dx = x^2\,e^x - 2\int x\,e^x\,dx =$$
$$= x^2\,e^x - 2\left[x\,e^x - e^x\right] + k = \left. x^2\,e^x - 2xe^x + 2e^x + k\right|_0^1 = (e - 2e + 2e) - (2) = e - 2,$$

$$d) \quad \int_0^1 x^3\,e^x\,dx = \int x^3\,D\,e^x\,dx = x^3\,e^x - \int e^x\,D(x^3)\,dx = x^3\,e^x - 3\int x^2\,e^x\,dx =$$
$$= x^3\,e^x - 3\left[x^2\,e^x - 2xe^x + 2e^x\right] + k =$$
$$= \left. x^3\,e^x - 3x^2e^x + 6xe^x - 6e^x + k\right|_0^1 = (e - 3e + 6e - 6e) - (-6) = -2e + 6$$

e così via per tutti gli infiniti integrali successivi di x^ne^x, con $n = 4, 5, 6, \ldots$

E' "facile convincersi" quindi "per induzione" che gli integrali di tutte le funzioni x^ne^x, con qualsiasi esponente intero non negativo $n = 0, 1, 2, 3, \ldots$, sono sempre ovviamente positivi, perché le funzioni integrande x^ne^x sono positive nell'intervallo $[0, 1]$, e dànno sempre risultato

$$0 < \int_0^1 x^ne^x\,dx = k_1e + k_2\,, \tag{8.38}$$

dove i coefficienti interi k_1, k_2 variano al variare dell'esponente n nell'integrale, ma sono sempre numeri interi. Nel prossimo paragrafo dimostreremo l'irrazionalità del numero e in un secondo modo dopo la dimostrazione già illustrata nel paragr. 6.3.2, utilizzando gli integrali calcolati appunto in questo primo esempio.

Esempio 2. Per calcolare l'integrale definito

$$\int_1^2 x^2e^{-3x}\,dx$$

applichiamo l'identità (8.37) del metodo per parti e gli integrali immediati, ottenendo

$$\int_1^2 x^2e^{-3x}\,dx = \int_1^2 x^2\,D\left(\frac{e^{-3x}}{-3}\right)dx = -\frac{x^2e^{-3x}}{3} - \left(-\frac{1}{3}\right)\int e^{-3x}Dx^2\,dx =$$

$$= -\frac{x^2e^{-3x}}{3} + \frac{2}{3}\int x\,e^{-3x}\,dx = -\frac{x^2e^{-3x}}{3} + \frac{2}{3}\left[\int x\,D\left(\frac{e^{-3x}}{-3}\right)dx\right] =$$

$$= -\frac{x^2e^{-3x}}{3} + \frac{2}{3}\left[-\frac{xe^{-3x}}{3} - \left(-\frac{1}{3}\right)\int e^{-3x}Dx\,dx\right] = -\frac{x^2e^{-3x}}{3} - \frac{2}{9}xe^{-3x} + \frac{2}{9}\int e^{-3x}\,dx =$$

$$= -\frac{x^2e^{-3x}}{3} - \frac{2}{9}xe^{-3x} - \frac{2}{27}e^{-3x} + k\Big|_1^2 = \left[\left(-\frac{4}{3} - \frac{4}{9} - \frac{2}{27}\right)e^{-6}\right] - \left[\left(-\frac{1}{3} - \frac{2}{9} - \frac{2}{27}\right)e^{-3}\right] =$$

$$= \frac{17e^3 - 50}{27e^6}.$$

Esempio 3. Per calcolare l'integrale definito

$$\int_0^1 x\log(2x+1)\,dx,$$

applichiamo l'identità (8.37) del metodo per parti, la proprietà di linearità (8.12), il secondo degli integrali (8.28) e gli integrali immediati, ottenendo

$$\int_0^1 x\log(2x+1)\,dx = \int \log(2x+1)\,D\left(\frac{x^2}{2}\right)dx = \frac{x^2}{2}\log(2x+1) - \int \frac{x^2}{2}\,D\log(2x+1)\,dx =$$

$$= \frac{x^2}{2}\log(2x+1) - \int \frac{x^2}{2x+1}\,dx = \frac{x^2}{2}\log(2x+1) - \frac{1}{4}\int \frac{(4x^2-1)+1}{2x+1}\,dx =$$

$$= \frac{x^2}{2}\log(2x+1) - \frac{1}{4}\int\left(2x - 1 + \frac{1}{2x+1}\right)dx =$$

$$= \frac{x^2}{2}\log(2x+1) - \frac{x^2}{4} + \frac{x}{4} - \frac{1}{8}\log|2x+1| + k\Big|_0^1 = \frac{3}{8}\log 3.$$

Esempio 4. Per calcolare l'integrale indefinito

$$\int \frac{\log x}{\sqrt{x}}\,dx$$

utilizziamo la derivata $D\sqrt{x} = 1/\left(2\sqrt{x}\right)$ e applichiamo l'identità (8.37) del metodo per parti ottenendo

$$\int \frac{\log x}{\sqrt{x}}\,dx = \int (\log x)\;D\left(2\sqrt{x}\right)dx = 2\sqrt{x}\log x - \int 2\sqrt{x}\,D\log x\,dx =$$

$$= 2\sqrt{x}\log x - \int 2\sqrt{x}\,\frac{1}{x}\,dx = 2\sqrt{x}\log x - 2\int x^{-1/2}\,dx = 2\sqrt{x}\log x - 2\left(\frac{x^{1/2}}{1/2}\right) + k =$$

$$= 2\sqrt{x}\log x - 4\sqrt{x} + k.$$

Esempio 5. Per calcolare l'integrale indefinito

$$\int \left(3\sqrt{x} - 1\right)\log(2x + \sqrt{x} - 3)\,dx, \tag{8.39}$$

calcoliamo preliminarmente l'integrale nella variabile y

$$\int \frac{8y^4 - 2y^3 - y^2}{2y^2 + y - 3}\,dy$$

utilizzando il metodo illustrato nel paragr. 8.3.1. Poiché la differenza tra i gradi dei polinomi al numeratore e al denominatore nella funzione integranda è $4-2=2$ e il denominatore ha i due *zeri semplici* $y=1$ e $y=-3/2$, decomponiamo la funzione ponendo

$$\frac{8y^4-2y^3-y^2}{2y^2+y-3}=B_2y^2+B_1y+B_0+\frac{A_1}{y-1}+\frac{A_2}{y+3/2}\,,$$

da cui otteniamo

$$A_1=\left.\frac{8y^4-2y^3-y^2}{4y+1}\right|_{y=1}=1 \qquad \text{e} \qquad A_2=\left.\frac{8y^4-2y^3-y^2}{4y+1}\right|_{y=-3/2}=-9.$$

Per determinare i tre coefficienti B_2, B_1, B_0 scriviamo l'uguaglianza (8.24b), avente la forma

$$8y^4-2y^3-y^2=(2y^2+y-3)(B_2y^2+B_1y+B_0)$$

e lo sviluppo $8y^4-2y^3-y^2=(2B_2)y^4+(2B_1+B_2)y^3+(2B_0+B_1-3B_2)y^2+\alpha_1x+\alpha_0\,.$

Se uguagliamo i coefficienti delle tre potenze "più alte" y^4, y^3, y^2 in ambo i membri, otteniamo i coefficienti $B_2=4, B_1=-3, B_0=7$ e quindi la decomposizione

$$\frac{8y^4-2y^3-y^2}{2y^2+y-3}=4y^2-3y+7+\frac{1}{y-1}-\frac{9}{y+3/2}\,,$$

da cui segue infine dunque

$$\begin{aligned}\int\frac{8y^4-2y^3-y^2}{2y^2+y-3}\,dy&=\int\left[4y^2-3y+7+\frac{1}{y-1}-\frac{9}{y+3/2}\right]dy=\\&=\frac{4}{3}\,y^3-\frac{3}{2}\,y^2+7y+\log|y-1|-9\log\left|y+\frac{3}{2}\right|+k.\end{aligned}\tag{8.40}$$

A questo punto risolviamo l'integrale (8.39) applicando prima il metodo di sostituzione, con il cambiamento di variabile $x=y^2$ da cui segue $\sqrt{x}=y, x\sqrt{x}=y^3, dx=2y\,dy$, poi l'identità (8.37) del metodo per parti ed infine il risultato (8.40). Otteniamo quindi

$$\int\left(3\sqrt{x}-1\right)\log(2x+\sqrt{x}-3)\,dx=\int 2y\,(3y-1)\log\left(2y^2+y-3\right)\,dy=$$

$$=\int\left(6y^2-2y\right)\log\left(2y^2+y-3\right)\,dy=\int\log\left(2y^2+y-3\right)\,D\,\left(2y^3-y^2\right)\,dy=$$

$$=\left(2y^3-y^2\right)\,\log\left(2y^2+y-3\right)-\int\left(2y^3-y^2\right)\,D\log\left(2y^2+y-3\right)\,dy=$$

$$=\left(2y^3-y^2\right)\,\log\left(2y^2+y-3\right)-\int\frac{\left(2y^3-y^2\right)(4y+1)}{2y^2+y-3}\,dy=$$

$$=\left(2y^3-y^2\right)\,\log\left(2y^2+y-3\right)-\int\frac{8y^4-2y^3-y^2}{2y^2+y-3}\,dy=$$

$$=\left(2y^3-y^2\right)\,\log\left(2y^2+y-3\right)-\left[\frac{4}{3}\,y^3-\frac{3}{2}\,y^2+7y+\log|y-1|-9\log\left|y+\frac{3}{2}\right|\right]+k=$$

$$=\left(2x\sqrt{x}-x\right)\,\log\left(2x+\sqrt{x}-3\right)-\frac{4}{3}\,x\sqrt{x}+\frac{3}{2}\,x-7\sqrt{x}-\log|\sqrt{x}-1|+9\log\left(\sqrt{x}+\frac{3}{2}\right)+k.$$

Esempio 6. Per calcolare l'integrale

$$\int (\log x)^2\, dx$$

ricordiamo che la derivata di x è 1, ovvero $Dx = 1$, e sviluppiamo

$$\int (\log x)^2\, dx = \int (\log x)^2\, D\, x\, dx = x(\log x)^2 - \int x\, D(\log x)^2\, dx =$$
$$= x(\log x)^2 - \int x\left[2(\log x)\left(\frac{1}{x}\right)\right] dx = x(\log x)^2 - 2\int \log x\, dx =$$
$$= x(\log x)^2 - 2\int \log x\, Dx\, dx = x(\log x)^2 - 2x\log x + 2\int x\, D\log x\, dx =$$
$$= x(\log x)^2 - 2x\log x + 2\int x\left(\frac{1}{x}\right) dx = x(\log x)^2 - 2x\log x + 2\int dx =$$
$$= x(\log x)^2 - 2x\log x + 2x + k.$$

8.3.4 Irrazionalità del numero e (seconda dimostrazione)

Poiché per ogni x appartenente all'intervallo $[0,1]$ la funzione crescente e^x dà risultati compresi tra i valori $e^0 = 1$ e $e^1 = e \leqslant 3$, segue la disuguaglianza $e^x \leqslant 3$ per ogni $x \in [0,1]$ e quindi, in virtù della proprietà (8.13), abbiamo la disuguaglianza tra integrali

$$0 < \int_0^1 x^n e^x\, dx \leqslant \int_0^1 3x^n\, dx = 3\int_0^1 x^n\, dx = 3\left(\frac{x^{n+1}}{n+1}\right)\Bigg|_{x=0}^{x=1} = \frac{3}{n+1}, \tag{8.41}$$

che mostra che per valori "sempre più grandi" di n, il valore dell'integrale

$$\int_0^1 x^n e^x\, dx$$

tende a zero. Per dimostrare che il numero e non è un numero razionale, supponiamo che esso sia uguale ad una frazione indicata con a/b, dove i due numeri a, b siano interi positivi, e mostriamo che dall'uguaglianza $e = a/b$ discende una conseguenza in contraddizione con la disuguaglianza (8.41). Poiché la disuguaglianza (8.41) discende da una proprietà generale degli integrali e non può essere contraddetta, segue che una "certa" ipotesi che conduca ad una contraddizione della (8.41) non può essere valida. Se quindi, come detto, ipotizziamo l'uguaglianza $e = a/b$ e la sostituiamo nei risultati degli integrali (8.38), segue

$$\int_0^1 x^n e^x\, dx = k_1 e + k_2 = k_1\left(\frac{a}{b}\right) + k_2 = \frac{k_1 a + k_2 b}{b} \geqslant \frac{1}{b} > 0, \tag{8.42}$$

perché il numeratore $k_1 a + k_2 b$ è un numero intero positivo e quindi maggiore di 1.

A questo punto è "facile" riconoscere che la disuguaglianza (8.42) contraddice la disuguaglianza (8.41) perché se, come si deduce dalla disuguaglianza (8.41), gli integrali

$$\int_0^1 x^n e^x\, dx$$

tendono a zero per n tendente ad infinito, dobbiamo concludere che gli stessi integrali non possono essere, come invece mostrato dalla disuguaglianza (8.42), tutti maggiori del numero fissato $1/b$ per ogni esponente n. Poiché dunque la disuguaglianza (8.42) non può sussistere e poiché tale disuguaglianza discende dall'aver supposto che il numero e sia uguale ad una frazione, possiamo concludere che il numero e non può essere uguale a nessuna frazione, ovvero che il numero e appartiene all'insieme dei numeri irrazionali.