UNIVERSITÀ DEGLI STUDI DI LECCE
FACOLTÀ DI INGEGNERIA

Appunti del corso di

Analisi Matematica I

Angela Albanese, Diego Pallara

a.a. 2006/07

PREFAZIONE

Nel presente fascicolo sono raccolte le nozioni di Analisi matematica presentate nel corso di Analisi Matematica I del primo anno di Ingegneria.

In un altro fascicolo sono raccolte le nozioni presentate nel corso di Analisi Matematica II.

Il pochissimo tempo destinato dai nuovi ordinamenti all'insegnamento della materia non permette alcun approfondimento, ed anzi obbliga ad escludere dai programmi argomenti tradizionalmente ritenuti indispensabili.

Riteniamo però imprescindibile, pur con tale riduzione dei contenuti, conservare intatti l'impianto concettuale e l'impostazione metodologica dell'Analisi, e riteniamo che questo obbiettivo sia conseguibile solo dando enunciati sintetici e precisi, e rifuggendo da espressioni vaghe o poco chiare. Per semplificare un enunciato si può rinunziare alla massima generalità possibile, ma non al rigore della presentazione. Per questa ragione abbiamo ritenuto opportuno, e, speriamo, utile agli studenti, raccogliere in poche pagine le definizioni ed i risultati principali che vengono esposti durante le lezioni. Lo stile degli appunti è volutamente scarno ed avaro di commenti e divagazioni, che restano affidati all'esposizione orale in aula; suggeriamo agli studenti, pertanto, di limitarsi ad appuntare, durante le lezioni, solo le parti meno formali delle lezioni stesse, affidandosi a questa dispensa per gli enunciati che richiedono maggior rigore.

È per altro evidente che questi appunti non hanno la pretesa di sostituire il libro di testo, che resta indispensabile per acquisire una conoscenza dignitosa della materia. La loro funzione è piuttosto, come già detto, quella di sostituire gli appunti di lezione, troppo poco affidabili per tanti motivi, e di indicare il bagaglio *minimo* di conoscenze richieste per affrontare l'esame.

Infine, ringraziamo il collega Raffaele Vitolo per averci fornito il file di stile LATEX usato per la compilazione delle dispense, e dichiariamo in anticipo la nostra gratitudine a tutti i lettori che ci segnaleranno ogni osservazione utile a migliorare il testo.

INDICE

CAPITOLO 1

NUMERI REALI E COMPLESSI

1.1 L'insieme dei numeri reali

L'ambiente in cui si svolgerà la nostra trattazione è quello dei numeri *reali*. Diamo per note le definizioni e le proprietà dei numeri *naturali*, il cui insieme è denotato con $\mathbf{N} = \{0, 1, 2, 3, \ldots\}$, dei numeri *interi*, il cui insieme è denotato con $\mathbf{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ e dei numeri *razionali*, il cui insieme è denotato con $\mathbf{Q} = \{p/q : \ p \in \mathbf{Z}, q \in \mathbf{N}, q \neq 0\}$. Una definizione costruttiva dell'insieme $\mathbf{R}$ dei numeri reali non sarebbe altrettanto immediata, ed in effetti non la daremo in queste note, rinviando ai libri di testo. È essenziale però impadronirsi delle proprietà dell'insieme dei numeri reali, che esprimeremo in forma assiomatica. La costruzione di $\mathbf{R}$ ha il ruolo (fondamentale) di provare che un insieme che gode delle proprietà elencate esiste nell'ambito delle usuali teorie insiemistiche.

Assiomi dei numeri reali.

Assumiamo che esista un insieme $\mathbf{R}$ dotato di due *operazioni binarie*, dette *somma* e *prodotto* e denotate rispettivamente $+$ e $\cdot$, e della *relazione d'ordine* di "maggiore od uguale", denotata $\geq$; chiamiamo gli elementi di tale insieme *numeri reali*, e assumiamo che valgano le seguenti proprietà, per ogni scelta di $a, b, c \in \mathbf{R}$.

Assiomi di campo.

Assioma 1. (Proprietà associative) $(a + b) + c = a + (b + c)$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

Assioma 2. (Proprietà commutative) $a + b = b + a$, $a \cdot b = b \cdot a$.

Assioma 3. (Proprietà distributiva) $(a + b) \cdot c = a \cdot c + b \cdot c$.

Assioma 4. (Elementi neutri) Esistono due numeri reali, denotati 0 e 1, che agiscono come *elementi neutri* rispettivamente dell'addizione e della moltiplicazione, cioè che verificano le uguaglianze: $a + 0 = a$ e $a \cdot 1 = a$ per ogni $a \in \mathbf{R}$.

Assioma 5. (Opposto) Per ogni $a \in \mathbf{R}$ esiste $b \in \mathbf{R}$ tale che $a + b = 0$; tale numero si dice *opposto* di a e si denota $-a$.

Assioma 6. (Inverso) Per ogni $a \in \mathbf{R}$, $a \neq 0$, esiste $b \in \mathbf{R}$ tale che $a \cdot b = 1$; tale numero si dice *inverso* di a e si denota a^{-1}.

Assiomi dell'ordine.

Assioma 7. Per ogni coppia di numeri reali a, b, o vale $a \geq b$ oppure $b \geq a$.

Assioma 8. Se valgono contemporaneamente $a \geq b$ e $b \geq a$ allora $a = b$.

Assioma 9. Se $a \geq 0$ e $b \geq 0$ allora $a + b \geq 0$ e $a \cdot b \geq 0$.

Gli assiomi elencati fin qui contengono, in forma rigorosa e concisa, delle proprietà dei numeri che sono già familiari; da esse si possono dedurre in modo sistematico tutte le proprietà note (per esempio, l'unicità degli elementi neutri) e le usuali regole algebriche e del calcolo letterale (semplificazioni, passaggio da un membro all'altro nelle uguaglianze e nelle diseguaglianze, eccetera). Non procederemo in questo modo, ritenendo che queste regole siano già note. Osserviamo che i primi sei assiomi sono di contenuto puramente algebrico e riguardano le operazioni di somma e moltiplicazione, mentre il 7 e l'8 riguardano la relazione d'ordine e il 9 lega le operazioni algebriche alla relazione d'ordine. Notiamo anche che, ovviamente, i numeri 0 e 1 dell'assioma 4 sono gli stessi degli insiemi $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, che si possono considerare sottoinsiemi di $\mathbf{R}$.

Osservazione 1.1 Osserviamo che i nove assiomi elencati sopra *non definiscono ancora completamente* $\mathbf{R}$. Infatti, essi valgono (per esempio) in $\mathbf{Q}$. Per definire $\mathbf{R}$ occorre un altro assioma che enunceremo fra poco e richiede qualche ulteriore nozione preliminare.

Definizione 1.2 (Valore assoluto) *Per ogni $x \in \mathbf{R}$, si definisce il* valore assoluto di *x, denotato con $|x|$, il numero*

$$|x| = \begin{cases} x & \text{se } x > 0 \\ 0 & \text{se } x = 0 \\ -x & \text{se } x < 0 \end{cases}$$

Proposizione 1.3 (Proprietà del valore assoluto) *Per ogni $r > 0$ vale la seguente equivalenza:*

$$|x| \leq r \qquad \Longleftrightarrow \qquad -r \leq x \leq r. \tag{1.1.1}$$

Inoltre, per ogni $x, y \in \mathbf{R}$:

$$|x| \geq 0, \qquad |x| = 0 \;\Leftrightarrow\; x = 0; \tag{1.1.2}$$

$$|x| = |-x|, \;\; |x \cdot y| = |x| \cdot |y|; \tag{1.1.3}$$

$$|x + y| \leq |x| + |y|; \tag{1.1.4}$$

$$||x| - |y|| \leq |x - y|. \tag{1.1.5}$$

cProviamo prima (1.1.1). Supponiamo dapprima $|x| \leq r$. Se $x \geq 0$ allora $x \geq -r$ e $|x| = x \leq r$; se $x < 0$ allora $x \leq r$ e $|x| = -x \leq r$, da cui $x \geq -r$. Viceversa, supponiamo $-r \leq x \leq r$. Allora, se $x \geq 0$ si ha $|x| = x \leq r$, mentre se $x < 0$ si ha $x = -|x| \geq -r$, da cui $|x| \leq r$.

Le (1.1.2), (1.1.3) sono ovvie conseguenze della definizione di valore assoluto. La (1.1.4), detta *diseguaglianza triangolare*, si può dimostrare usando (1.1.1). Infatti, sommando le relazioni:

$$-|x| \leq x \leq |x|, \quad -|y| \leq y \leq |y|$$

si deduce

$$-(|x|+|y|) \le x+y \le (|x|+|y|)$$

da cui per la (1.1.1) segue $|x+y| \le |x|+|y|$.

Infine, da (1.1.4) si deduce facilmente (1.1.5); infatti, risulta $|x| = |(x-y)+y| \le |x-y|+|y|$, e da qui $|x|-|y| \le |x-y|$; scambiando x con y si ottiene $|y|-|x| \le |x-y|$ e quindi $-|x-y| \le |x|-|y| \le |x-y|$ e la tesi segue da (1.1.1). QED

Definiamo una classe di sottoinsiemi di $\mathbf{R}$ che interverrà in numerose considerazioni.

Definizione 1.4 (Insiemi limitati) *Sia X un sottoinsieme non vuoto di $\mathbf{R}$. Allora:*

1. *si dice che X è* limitato superiormente *se esiste $M \in \mathbf{R}$ tale che $x \le M$ per ogni $x \in X$;*
2. *si dice che X è* limitato inferiormente *se esiste $m \in \mathbf{R}$ tale che $x \ge m$ per ogni $x \in X$;*
3. *si dice che X è* limitato *se è limitato superiormente ed inferiormente.*

Introduciamo una notazione per gli intervalli di $\mathbf{R}$, che sono i sottoinsiemi con cui prevalentemente (ma non *esclusivamente*!) lavoreremo, e per gli intorni di un punto, che useremo per descrivere le proprietà di vicinanza tra numeri reali.

Definizione 1.5 (Intervalli e intorni) *Dati a e b in $\mathbf{R}$, con $a < b$, si dice* intervallo chiuso *di estremi a e b l'insieme*

$$[a,b] = \{x \in \mathbf{R}: \ a \le x \le b\};$$

si dice intervallo aperto *di estremi a e b l'insieme*

$$]a,b[= \{x \in \mathbf{R}: \ a < x < b\};$$

si dice intervallo semiaperto a destra *di estremi a e b l'insieme*

$$[a,b[= \{x \in \mathbf{R}: \ a \le x < b\};$$

si dice intervallo semiaperto a sinistra *di estremi a e b l'insieme*

$$]a,b] = \{x \in \mathbf{R}: \ a < x \le b\}.$$

dato $a \in \mathbf{R}$, si denota con $[a,+\infty[$ l'insieme

$$[a,+\infty[= \{x \in \mathbf{R}: \ a \le x\},$$

e con $]a,+\infty[$ l'insieme

$$]a,+\infty[= \{x \in \mathbf{R}: \ a < x\};$$

analogamente:

$$\begin{aligned}]-\infty,a] &= \{x\in\mathbf{R}:\ x\le a\},\\ ]-\infty,a[&= \{x\in\mathbf{R}:\ x<a\}. \end{aligned}$$

Dati $x_0\in\mathbf{R}$ *e* $r>0$*, si dice* intorno aperto di x_0 di raggio r *l'insieme*

$$I_r(x_0)=\{x\in\mathbf{R}:\ x_0-r<x<x_0+r\}=]x_0-r,x_0+r[.$$

Nella definizione di intervalli con un estremo infinito, non si è dato alcun significato ai simboli $\pm\infty$ fuori dal contesto dell'intera espressione che li contiene. Questo accadrà spesso anche nel seguito.

Molto spesso, parleremo genericamente di *intervallo*; se non viene specificato nulla, s'intende che quanto detto vale per intervalli qualunque (aperti, chiusi, semiaperti limitati, illimitati, indifferentemente). Accanto all'intorno aperto di x_0 di raggio r possiamo considerare l'intorno chiuso

$$\overline{I}_r(x_0)=\{x\in\mathbf{R}:\ x_0-r\le x\le x_0+r\}=[x_0-r,x_0+r].$$

Notiamo inoltre che un insieme X è limitato se e solo se è contenuto in un intervallo limitato, cioè se e solo se esistono $m,M\in\mathbf{R}$ tali che $X\subset[m,M]$. Introduciamo due importanti concetti legati alla limitatezza.

Definizione 1.6 (Maggioranti, minoranti, massimo, minimo) *Sia* X *un sottoinsieme di* $\mathbf{R}$.

Si dice che $M\in\mathbf{R}$ *è un* maggiorante per X *se* $x\le M$ *per ogni* $x\in X$*. Si dice che* M *è il* massimo di X*, e si scrive* $M=\max X$*, se* M *è un maggiorante ed inoltre* M *appartiene ad* X.

Si dice che $m\in\mathbf{R}$ *è un* minorante per X *se* $x\ge m$ *per ogni* $x\in X$*. Si dice che* m *è il* minimo di X*, e si scrive* $m=\min X$*, se* m *è un minorante ed* m *appartiene ad* X.

Le considerazioni che seguono sono tutte conseguenze dirette delle definizioni.

Osservazioni 1.7

1. Un insieme ammette maggioranti se e solo se è limitato superiormente, ed ammette minoranti se e solo se è limitato inferiormente.

2. Se un insieme ammette un maggiorante M allora ne ammette infiniti, poiché ogni numero maggiore di M è ancora un maggiorante. Naturalmente, una considerazione analoga vale per i minoranti.

3. A differenza dei maggioranti, il massimo e il minimo di un insieme, se esistono, sono unici. Infatti, se M_1 ed M_2 sono entrambi massimi di X allora M_1,M_2 appartengono entrambi ad X ed applicando la definizione di massimo prima con $M=M_1$ ed $x=M_2$ e poi con $M=M_2$ e $x=M_1$ si trova $M_2\le M_1$ e poi $M_1\le M_2$, da cui $M_1=M_2$ e l'unicità del massimo. Ovviamente un ragionamento analogo porta all'unicità del minimo.

4. Un insieme limitato può non avere massimo o minimo. Per esempio, l'intervallo $]a, b]$ ha massimo b ma non ha minimo, perché i suoi minoranti sono gli elementi dell'intervallo $]-\infty, a]$, e nessuno di essi appartiene ad $]a, b]$.

5. Ogni insieme costituito da un numero *finito* di numeri reali ha sempre massimo e minimo.

Tenendo conto delle osservazioni precedenti, diamo la seguente definizione.

Definizione 1.8 (Estremo superiore ed inferiore) *Dato $X \subset \mathbf{R}$ limitato superiormente, e detto M_X l'insieme dei suoi maggioranti, diciamo* estremo superiore di X *il più piccolo dei maggioranti, cioè il numero*

$$\sup X = \min M_X.$$

Dato $X \subset \mathbf{R}$ limitato inferiormente, e detto M'_X l'insieme dei suoi minoranti, diciamo estremo inferiore di X *il più grande dei minoranti, cioè il numero*

$$\inf X = \max M'_X.$$

Come osservato, in generale, un insieme, anche limitato, può non avere massimo o minimo. Al contrario, in $\mathbf{R}$ l'insieme dei maggioranti (rispettivamente, dei minoranti) di un insieme dato ha *sempre* minimo (risp. massimo) in $\mathbf{R}$, e questo è ciò che distingue in modo essenziale $\mathbf{R}$ da $\mathbf{Q}$. L'esistenza dell'estremo superiore (inferiore) per un insieme limitato superiormente (inferiormente) completa la nostra descrizione assiomatica dell'insieme dei numeri reali. Notiamo che in tutta la trattazione precedente sono stati usati solo gli assiomi già enunciati, e quindi essa è logicamente coerente, anche se la descrizione di $\mathbf{R}$ non era ancora completa.

Assioma 10. (Completezza) Ogni insieme $X \subset \mathbf{R}$ non vuoto e limitato superiormente ammette estremo superiore $\sup X$ in $\mathbf{R}$.

Osservazioni 1.9

1. Dall'assioma 10 segue subito che ogni sottoinsieme di $\mathbf{R}$ non vuoto e limitato inferiormente ammette estremo inferiore in $\mathbf{R}$.

2. Abbiamo già osservato che il massimo e il minimo di un insieme, se esistono, sono unici. Segue subito quindi dalla definizione l'unicità dell'estremo superiore e dell'estremo inferiore.

3. Conveniamo di porre $\sup X = +\infty$ se $\emptyset \neq X \subset \mathbf{R}$ e X non è limitato superiormente, e $\inf X = -\infty$ se $\emptyset \neq X \subset \mathbf{R}$ e X non è limitato inferiormente. Come prima, non diamo un significato ai simboli $\pm\infty$ isolati dal contesto, ma solo all'intera espressione che li contiene.

4. È evidente che se $\emptyset \neq A \subset B \subset \mathbf{R}$ allora $\sup A \leq \sup B$ e $\inf A \geq \inf B$.

Discutiamo separatamente altre due importanti conseguenze dell'assioma di completezza, la cui dimostrazione è meno immediata.

Osservazione 1.10

1. **(Proprietà archimedea)**: per ogni coppia di numeri reali a e b, con $0 < a < b$, esiste un numero naturale n tale che $na > b$. Segnaliamo anche la seguente conseguenza: se un numero $c \geq 0$ è minore di ε *per ogni* $\varepsilon > 0$ allora $c = 0$. Infatti, se fosse $c > 0$, dato $\varepsilon > 0$ esisterebbe $n \in \mathbf{N}$ tale che $nc > \varepsilon$, ossia $c > \varepsilon/n$, che contraddice l'ipotesi che c sia minore di *ogni* numero strettamente positivo prefissato. Tale risultato è talvolta utile per provare che due numeri reali a e b sono uguali, applicandolo a $c = |a-b|$. Se infatti si riesce a provare che $|a-b| < \varepsilon$ per ogni $\varepsilon > 0$, allora segue $a = b$.

2. **(Densità dei razionali nei reali)**: per ogni coppia di numeri reali a, b, con $a < b$, esiste un numero *razionale* r tale che $a < r < b$. prendiamo prima $a > 0$. Allora per n maggiore del più grande fra i numeri $\frac{1}{a}$ e $\frac{1}{b-a}$ risulta $0 < \frac{1}{n} < a$ e per la proprietà archimedea esiste m tale che $\frac{m}{n} > a$. Se si sceglie m in modo che $\frac{m-1}{n} \leq a$, essendo $\frac{1}{n} < b - a$, si ha anche $\frac{m}{n} < b$. Se $b < a < 0$ si ragiona come prima con $-a$ e $-b$ e poi si cambia di segno il numero trovato. se $a \leq 0$ e $b > 0$ si trova come prima r tra 0 e b, se $a < 0$ e $b = 0$ si trova r tra a e 0.

Si può dare una utile caratterizzazione dell'estremo superiore e dell'estremo inferiore di un insieme.

Proposizione 1.11 (Caratterizzazione del sup e dell'inf) *Sia $X \subset \mathbf{R}$ limitato. Allora valgono le seguenti equivalenze:*

$$(1.1.6)\quad L = \sup X \quad \Longleftrightarrow \quad \begin{cases} x \leq L \text{ per ogni } x \in X \\ \text{per ogni } \varepsilon > 0 \text{ esiste } x_\varepsilon \in X \text{ tale che } x_\varepsilon > L - \varepsilon \end{cases}$$

$$(1.1.7)\quad \ell = \inf X \quad \Longleftrightarrow \quad \begin{cases} x \geq \ell \text{ per ogni } x \in X \\ \text{per ogni } \varepsilon > 0 \text{ esiste } x_\varepsilon \in X \text{ tale che } x_\varepsilon < \ell + \varepsilon \end{cases}$$

DIM. Sia dapprima $L = \sup X$. Allora, $x \leq L$ per ogni $x \in X$ perché L è un maggiorante di X. Inoltre, nessun numero più piccolo di L è un maggiorante di X; poiché ogni numero minore di L si può scrivere nella forma $L - \varepsilon$, con $\varepsilon > 0$, negare che $L - \varepsilon$ sia un maggiorante di X equivale a dire che esiste un $x_\varepsilon \in X$ tale che $x_\varepsilon > L - \varepsilon$.

Viceversa, supponiamo che valgano le due condizioni a destra in (1.1.6); allora, la prima dice che L è un maggiorante di X. La seconda afferma che nessun numero più piccolo di L, espresso nella forma $L - \varepsilon$, con $\varepsilon > 0$ arbitrario, è un maggiorante di X. Segue $L = \sup X$.

La dimostrazione di (1.1.7) è analoga. $\boxed{QED}$

L'assioma di completezza *non vale* nell'insieme dei numeri razionali $\mathbf{Q}$, e quindi $\mathbf{Q}$ è contenuto *propriamente* in $\mathbf{R}$. I numeri reali non razionali si dicono *irrazionali*.

Esempio 1.12 Sia $X = \{r \in \mathbf{Q} : \ r^2 < 2\}$; allora, X è limitato superiormente, ma $\sup X \notin \mathbf{Q}$; segue che l'assioma di completezza non vale in $\mathbf{Q}$. Per giustificare la nostra affermazione, procediamo in due passi: mostriamo prima che se $L = \sup X$ allora $L^2 = 2$, e poi che se $L^2 = 2$ allora $L \notin \mathbf{Q}$. Per la prima parte, si può ragionare così: supposto $L^2 < 2$, e posto $m = 2 - L^2$, si ha $m > 0$ e si vede che esistono soluzioni positive della seguente disequazione nella variabile ε:

$$\varepsilon^2 + 2L\varepsilon - m < 0;$$

per tali valori di ε risulta che $(L+\varepsilon)^2 < 2$ e quindi per la densità di $\mathbf{Q}$ in $\mathbf{R}$ ci sono elementi di X compresi tra L ed $L+\varepsilon$, sicché, supposto $L^2 < 2$, L non può essere il sup di X. Analogamente, supposto $L^2 > 2$, e posto $m = L^2 - 2$, si ha ancora $m > 0$ e si vede che esistono soluzioni positive della seguente disequazione nella variabile ε:

$$\varepsilon^2 - 2L\varepsilon - m > 0;$$

per tali valori di ε risulta che $(L-\varepsilon)^2 > 2$ e quindi esistono maggioranti di X compresi tra $L - \varepsilon$ ed L, in particolare più piccoli di L. Questo prova che, supposto $L^2 > 2$, L non può essere il sup di X. In definitiva, $(\sup X)^2 = 2$. Proviamo ora che $\sup X \notin \mathbf{Q}$. Supposto vero il contrario, sia $\sup X = p/q$ con la frazione p/q ridotta ai minimi termini. Si vede facilmente che questo porta ad una contraddizione. Infatti:

> $(\frac{p}{q})^2 = 2$ implica $p^2 = 2q^2$, da cui p^2 pari, e quindi p pari, diciamo $p = 2r$ allora $q^2 = 2r^2$ pari, e q pari.

La precedente conclusione è impossibile perché la frazione era supposta ridotta ai minimi termini.

1.2 Funzioni elementari

Prima di affrontare lo studio delle funzioni reali di variabile reale, che sarà l'argomento centrale del corso, richiamiamo alcune nozioni generali sulle funzioni tra insiemi generici.

1.2.a Generalità sulle funzioni

Definizione 1.13 *Siano U un qualunque insieme (non vuoto), che consideriamo come l'universo del nostro discorso.*

1. *Una* **funzione** *è una terna costituita da due sottoinsiemi di U, il primo, che denotiamo con A, detto* **dominio**, *il secondo, denotato con B, detto* **codominio**, *ed una legge di corrispondenza che fa corrispondere ad ogni elemento x di A uno (ed un solo) elemento di B, denotato con $f(x)$. Simbolicamente, scriviamo $f : A \to B$.*

2. *Si dice* **insieme immagine** *di f l'insieme*

$$f(A) = \{y \in B : \ \textit{esiste } x \in A \textit{ tale che } f(x) = y\} \subset B.$$

3. *Si dice* **grafico di** f *l'insieme*

$$G(f) = \{(x,y) \in A \times B : \ x \in A, \ y = f(x)\} \subset A \times B.$$

4. *Data* $f : A \to B$, *e dato un sottoinsieme* $C \subset A$, *si dice* **restrizione di** f *la funzione* $f|_C : C \to B$ *che ha per dominio* C, *per codominio* B *e come legge di corrispondenza la stessa della* f *iniziale.*

5. *Una funzione* $f : A \to B$ *si dice* **iniettiva** *se* $x_1, x_2 \in A$, $x_1 \neq x_2$ *implica* $f(x_1) \neq f(x_2)$.

6. *Una funzione* $f : A \to B$ *si dice* **surgettiva** *se* $f(A) = B$, *cioè se per ogni* $y \in B$ *esiste* $x \in A$ *tale che* $f(x) = y$.

7. *Una funzione* $f : A \to B$ *si dice* **bigettiva** *se è iniettiva e surgettiva.*

8. *Una funzione* $f : A \to B$ *si dice* **invertibile** *se esiste una funzione* $g : B \to A$ *tale che* $y = f(x) \ \Leftrightarrow \ x = g(y)$ *per ogni* $x \in A$, $y \in B$. *Se* f *è invertibile, la funzione* g *suddetta si dice* inversa di f *e si denota con* f^{-1}.

9. *Per ogni insieme* $A \subset U$, *si definisce la* **funzione identità** $id_A : A \to A$ *ponendo* $id_A(x) = x$ *per ogni* $x \in A$.

10. *Date le funzioni* $f : A \to B$ *e* $g : B \to C$, *si definisce la* **funzione composta** $g \circ f : A \to C$ *ponendo* $(g \circ f)(x) = g(f(x))$ *per ogni* $x \in A$.

Osservazioni 1.14 1. Una funzione è invertibile se e solo se è bigettiva.

2. Data $f : A \to B$, la funzione $g : B \to A$ è l'inversa di f se e solo se $g \circ f = id_A$ e $f \circ g = id_B$, cioè se $(g \circ f)(x) = x$ per ogni $x \in A$ e $(f \circ g)(y) = y$ per ogni $y \in B$.

3. Una funzione è sempre surgettiva prendendo come codominio l'insieme immagine $f(A)$, quindi ogni funzione iniettiva è sempre bigettiva da A in $f(A)$.

4. Se $f : A \to B$ è invertibile e $G(f)$ è il suo grafico, allora il grafico della funzione inversa $f^{-1} : B \to A$ è l'insieme

$$G(f^{-1}) = \{(y,x) \in B \times A : \ (x,y) \in G(f)\}.$$

1.2.b Funzioni reali di una variabile

D'ora in poi l'universo del nostro discorso, salvo avviso contrario, sarà l'insieme dei numeri reali, o qualche suo sottoinsieme. Considereremo perciò *funzioni reali di una variabile reale*, cioè funzioni definite in $X \subset \mathbf{R}$ ed a valori in $\mathbf{R}$, simbolicamente $f : X \to \mathbf{R}$. Useremo ovviamente la terminologia introdotta nella Definizione 1.13, ma, in questo caso particolare, accanto alle proprietà generali delle funzioni già viste possiamo segnalarne altre, peculiari delle funzioni reali. Iniziamo dall'importante nozione di monotonia.

Definizione 1.15 *Sia* $f : X \subset \mathbf{R} \to \mathbf{R}$*; si dice che* f *è* crescente *se*

$$x_1, x_2 \in X,\ x_1 < x_2 \quad \Longrightarrow \quad f(x_1) \leq f(x_2);$$

si dice che f *è* strettamente crescente *se*

$$x_1, x_2 \in X,\ x_1 < x_2 \quad \Longrightarrow \quad f(x_1) < f(x_2);$$

si dice che f *è* decrescente *se*

$$x_1, x_2 \in X,\ x_1 < x_2 \quad \Longrightarrow \quad f(x_1) \geq f(x_2);$$

si dice che f *è* strettamente decrescente *se*

$$x_1, x_2 \in X,\ x_1 < x_2 \quad \Longrightarrow \quad f(x_1) > f(x_2);$$

si dice che f *è* monotòna *se è crescente o decrescente, che è* strettamente monotòna *se è strettamente crescente o strettamente decrescente.*

Osservazione 1.16 È chiaro che ogni funzione $f : X \subset \mathbf{R} \to \mathbf{R}$ strettamente monotona è iniettiva, e perciò è invertibile da X su $f(X)$. Useremo sistematicamente questo fatto per studiare l'invertibilità delle funzioni. Inoltre, segue subito dalle definizioni che l'inversa di una funzione crescente è crescente, e l'inversa di una funzione decrescente è decrescente.

Altre proprietà delle funzioni reali corrispondono alle proprietà dei sottoinsiemi di $\mathbf{R}$ visti nella sezione precedente.

Definizione 1.17 (Funzioni limitate) *Sia* $f : X \subset \mathbf{R} \to \mathbf{R}$*. si dice che* f *è limitata se* $f(X)$ *è limitato, cioè se esistono* $m, M \in \mathbf{R}$ *tali che* $m \leq f(x) \leq M$ *per ogni* $x \in X$*.*

Come per la limitatezza, per definire massimo, minimo, estremo superiore ed inferiore di una funzione si fa riferimento all'insieme immagine $f(X)$.

Definizione 1.18 (max, min, sup ed inf di una funzione) *Si definiscono il* massimo*, il* minimo*, l'*estremo superiore *e l'*estremo inferiore *di* f *su* X *ponendo:*

$$\begin{aligned}
\max_X f &= \max f(X) = \max\{f(x) :\ x \in X\}, \\
\min_X f &= \min f(X) = \min\{f(x) :\ x \in X\}, \\
\sup_X f &= \sup f(X) = \sup\{f(x) :\ x \in X\}, \\
\inf_X f &= \inf f(X) = \inf\{f(x) :\ x \in X\}.
\end{aligned}$$

Osservazione 1.19 Come per gli insiemi, una funzione $f : X \to \mathbf{R}$, anche limitata, può non avere massimo o minimo. La funzione f ha massimo se e solo se esiste $x_1 \in X$ tale che $f(x) \leq f(x_1)$ per ogni $x \in X$. In tal caso, x_1 si dice *punto di massimo assoluto per* f *in* X. Analogamente, la funzione f ha minimo se e solo se esiste $x_2 \in X$ tale che $f(x) \geq f(x_2)$ per ogni $x \in X$. In tal caso, x_2 si dice *punto di minimo assoluto per* f *in* X.

Naturalmente, il grafico di una funzione $f : X \subset \mathbf{R} \to \mathbf{R}$ è sempre un sottoinsieme di $\mathbf{R}^2$, prodotto cartesiano della retta reale per sé stessa, in cui supponiamo fissato un riferimento cartesiano ortogonale di assi x (asse delle ascisse) ed y (asse delle ordinate). Conveniamo di rappresentare sull'asse delle ascisse la variabile indipendente e sull'asse delle ordinate la variabile dipendente, sicché il grafico di una funzione $f : X \subset \mathbf{R} \to \mathbf{R}$ sarà l'insieme dei punti del piano che verificano le condizioni $x \in X$ e $y = f(x)$. Dal punto di vista geometrico, i grafici delle funzioni nel piano sono caratterizzati dalla seguente condizione:

> Sia $G \subset \mathbf{R}^2$. Allora, esistono $X \subset \mathbf{R}$ ed $f : X \to \mathbf{R}$ tali che $G = G(f)$ se e solo se l'intersezione di G con ogni retta verticale, cioè del tipo $x =$ costante, contiene al più un punto.

Una condizione geometrica analoga caratterizza l'*iniettività* di una funzione; infatti,

> Sia $f : X \to \mathbf{R}$; allora, f è iniettiva se e solo se l'intersezione di $G(f)$ con ogni retta orizzontale, cioè del tipo $y =$ costante, contiene al più un punto.

1.2.c Funzioni elementari

In questo paragrafo definiamo le più usuali funzioni di una variabile, a partire dalle quali, con le operazioni algebriche e la composizione di funzioni si otterranno la maggior parte degli esempi che incontreremo. Le funzioni che andiamo a considerare saranno definite attraverso *espressioni analitiche*, cioè algoritmi di calcolo che comprendono operazioni algebriche oppure calcolo di estremi superiore od inferiore. Tali algoritmi consistono in una procedura che, dato il numero reale x in un opportuno insieme, prescrive come si debba calcolare il numero $f(x)$, cioè il valore cha la funzione f assume in corrispondenza del valore assegnato alla variabile indipendente. Non bisogna per altro confondere la *funzione* f con la procedura di calcolo di $f(x)$, che chiamiamo *espressione analitica.* Infatti, per dare la funzione f, pur dando per scontato che il suo codominio sia $\mathbf{R}$, bisogna dichiarare quale sia il *dominio* scelto, oltre ad assegnare l'espressione analitica che contiene la legge di corripondenza richiesta per completare la definizione. Ciò non ostante, talvolta il dominio corrispondente ad una certa espressione analitica è taciuto: in tal caso, si assume come dominio il così detto *dominio naturale dell'espressione analitica*, cioè il più grande sottoinsieme di $\mathbf{R}$ in cui tutte le operazioni richieste per il calcolo di $f(x)$ si possono eseguire.

Funzioni razionali

Le funzioni razionali sono quelle definite da espressioni analitiche che contengono solo operazioni algebriche. Esplicitamente, abbiamo i *polinomi* e le *funzioni razionali.* Chiameremo *polinomio di grado* n nella variabile x l'espressione

$$P(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

definita per ogni x reale, dove gli $a_0, \dots, a_n$ sono numeri reali dati, con $a_n \neq 0$, detti *coefficienti del polinomio*. Il grado del polinomio è quindi il massimo esponente della potenza di x con coefficiente non nullo. Tra i polinomi hanno un ruolo particolare le funzioni *affini*, cioè i polinomi di grado 1, $f(x) = ax + b$, quelle *lineari*, $f(x) = ax$ e le funzioni *potenza*, cioè del tipo $f(x) = x^n$, in cui uno solo dei coefficienti è non nullo. Notiamo che le funzioni affini e lineari, con $a \neq 0$, sono strettamente monotone, e perciò invertibili (crescenti per $a > 0$ e decrescenti per $a < 0$). Per quanto riguarda le funzioni potenza, per n dispari esse sono strettamente crescenti su $\mathbf{R}$ (e quindi invertibili), mentre per n pari sono strettamente crescenti le loro restrizioni all'insieme dei numeri positivi $\{x \in \mathbf{R} : \ x \geq 0\}$.

Le altre funzioni razionali sono quelle espresse come rapporto di polinomi. Fra queste, le più semplici sono le funzioni *potenze negative* x^{-n}, con $n \in \mathbf{N}$, $n \geq 1$, definite per $x \neq 0$ come $x^{-n} = \frac{1}{x^n}$. La generica funzione razionale sarà del tipo

$$f(x) = \frac{P(x)}{Q(x)},$$

dove P e Q sono polinomi di grado qualunque. Il dominio naturale di f è in questo caso l'insieme $X = \{x \in \mathbf{R} : \ Q(x) \neq 0\}$, dal momento che, delle operazioni richieste per il calcolo di $f(x)$, l'unica che non si possa eseguire per ogni numero reale è la divisione, in cui è escluso che il denominatore possa essere 0.

Radici aritmetiche e potenze razionali

Per ogni $n \in \mathbf{N}$, con $n \geq 2$, e per ogni $x \geq 0$, poniamo $\sqrt[n]{x} = \sup\{y \in \mathbf{R} : \ y^n < x\}$, funzione detta *radice n-esima aritmetica di x*. Notiamo che per ogni $x \geq 0$ risulta $\sqrt[n]{x} \geq 0$. L'esistenza della radice è assicurata dall'assioma di completezza. Inoltre, segue dalle proprietà dell'estremo superiore che $(\sqrt[n]{x})^n = x$, cioè che la funzione radice n-esima è l'*inversa della restrizione della funzione potenza n-esima all'insieme* $\{x \in \mathbf{R} : \ x \geq 0\}$, che, come già osservato, è strettamente crescente. Poiché le funzioni potenza n-esima per n dispari sono strettamente crescenti su $\mathbf{R}$ e non solo su $\{x \geq 0\}$, si possono estendere le funzioni radice n-esima, per n dispari, ad $\mathbf{R}$, ponendo $\sqrt[n]{x} = -\sqrt[n]{-x}$ per ogni $x < 0$.

Per $r \in \mathbf{Q}$, posto per fissare le idee $r = p/q$ con $p \in \mathbf{Z}$ e $q \in \mathbf{N}$, $q \neq 0$, poniamo $x^r = x^{p/q} = \sqrt[q]{x^p}$ per ogni $x \geq 0$ ($x > 0$ se $p < 0$).

Esponenziali e logaritmi

Avendo definito l'epressione x^r, con $x > 0$ ed $r \in \mathbf{Q}$, possiamo pensarla con x fissato ed r variabile, passando così dalla *funzione potenza* già considerata alla *funzione esponenziale*, per ora con esponente razionale. Siccome $1^r = 1$ per ogni r, considereremo in questo paragrafo una base strettamente positiva e diversa da 1. Per sottolineare che la base della potenza è costante, scriveremo a^r, supponendo fissato il numero reale $a > 0$, $a \neq 1$. Vale il seguente importante risultato.

Proposizione 1.20 *La funzione esponenziale a^r, $r \in \mathbf{Q}$, gode delle seguenti proprietà:*

1. *$a^r > 0$ per ogni $a > 0$ e per ogni $r \in \mathbf{Q}$;*
2. *per $a \in]0,1[$ la funzione a^r è strettamente decrescente;*
3. *per $a > 1$ la funzione a^r è strettamente crescente;*
4. *vale l'uguaglianza $a^{r_1} \cdot a^{r_2} = a^{r_1+r_2}$ per ogni $r_1, r_2 \in \mathbf{Q}$.*

Dalla Proposizione 1.20 segue che la definizione di esponenziale si può estendere al caso di esponenti *reali* qualsiasi. Sia $a > 0$, $a \neq 1$. Poniamo

$$a^x = \inf\{a^r : r \in \mathbf{Q},\ r < x\} \qquad \text{se } a \in]0,1[; \tag{1.2.8}$$

$$a^x = \sup\{a^r : r \in \mathbf{Q},\ r < x\} \qquad \text{se } a > 1; \tag{1.2.9}$$

È molto importante notare che la funzione esponenziale *con esponente reale* appena definita gode ancora delle proprietà elencate nella Proposizione 1.20 ed è surgettiva su $]0, +\infty[$. Segue allora dall'Osservazione 1.16 che è possibile definire la funzione inversa dell'esponenziale.

Per $a > 0$, $a \neq 1$, si dice *logaritmo in base* a, e si denota $\log_a : \{x > 0\} \to \mathbf{R}$, la funzione inversa di a^x; risulta allora

$$\log_a x = y \quad \Longleftrightarrow \quad a^y = x \qquad \forall\, x > 0,\ y \in \mathbf{R}.$$

Sempre per l'Osservazione 1.16 si ha che $\log_a$ è strettamente crescente per $a > 1$ e strettamente decrescente per $a \in]0,1[$.

Tra le funzioni esponenziali e i logaritmi, per motivi che saranno chiari più avanti nel corso, ha un ruolo importantissimo quella la cui base è il numero

$$e = \sup_{n \in \mathbf{N}} \left(1 + \frac{1}{n}\right)^n. \tag{1.2.10}$$

Risulta che e è un numero *irrazionale* e in particolare un suo valore approssimato è $e = 2,71828$.

Definizione 1.21 (Funzione esponenziale e logaritmo naturale) *si dice* funzione esponenziale *la funzione e^x, e* logaritmo naturale *il logaritmo in base e, denotato semplicemente* log.

Notiamo che $e > 1$, quindi sia l'esponenziale e^x che il logaritmo naturale $\log x$ sono funzioni strettamente crescenti.

Infine, il procedimento esposto permette di definire anche le *funzioni potenza con esponente reale.* In altri termini, fissato $\alpha \in \mathbf{R}$, possiamo definire la *funzione potenza di esponente* α ponendo, per ogni $x > 0$, x^α come il valore dato dalle (1.2.8), (1.2.9), secondo i casi. Notiamo che $x^\alpha = e^{\alpha \log x}$.

Funzioni trigonometriche

Non ci soffermeremo sulle (tante) proprietà delle funzioni trigonometriche, che supponiamo note. Ci limitiamo a darne brevissimi cenni, limitati agli aspetti più legati alle applicazioni che seguono. Fissiamo anzitutto un riferimento ortogonale nel piano e consideriamo la circonferenza di centro l'origine O e raggio 1, di equazione $x^2+y^2=1$. I suoi punti P possono essere identificati tramite l'angolo che il raggio OP forma col semiasse positivo dell'asse x. Definiamo ora l'unità di misura degli angoli.

Definizione 1.22 *Si dice* misura in radianti di un angolo *la lunghezza dell'arco individuato dalle semirette che determinano l'angolo sulla circonferenza unitaria di centro il punto d'incontro delle semirette stesse.*

Sottolineiamo che la precedente definizione, sebbene intuitiva, non può considerarsi rigorosa, dal momento che non abbiamo precisato come si possa definire e calcolare la lunghezza di un arco di circonferenza. Questo si può fare, sfruttando le proprietà della circonferenza, in modo elementare (cioè seguendo gli *Elementi* di Euclide, dove l'argomento è trattato in modo esauriente), oppure come caso particolare di una trattazione generale del problema della lunghezza delle curve, che viene studiata nel corso di Analisi matematica II e, in particolare, è presentata nel Capitolo 4 delle relative dispense.

Vogliamo ora passare ad una misura *orientata* degli angoli, così come si fa per le misura lineari quando di introduce la nozione di ascissa. Per prima cosa, scegliamo di considerare positivo il verso antiorario di percorrenza della circonferenza, e misuriamo gli angoli a partire dal semiasse positivo delle x.

Ricordiamo che la lunghezza della circonferenza unitaria (per definizione del numero π) ha il valore 2π. Notiamo inoltre che allo stesso punto sulla circonferenza sono associati infiniti valori dell'angolo (positivi e negativi). Fissato un intervallo semiaperto di lunghezza 2π (cioè pari alla lunghezza della circonferenza unitaria), per esempio $]-\pi,\pi]$, uno e uno solo di questi valori appartiene a tale intervallo, e tutti gli altri si ottengono da questo sommando multipli interi di 2π.

Queste proprietà si riflettono nella proprietà delle funzioni trigonometriche di essere *periodiche*, secondo la seguente definizione.

Definizione 1.23 (Funzioni periodiche) *Sia $f:\mathbf{R}\to\mathbf{R}$; diciamo che f è* periodica di periodo T *(o T-*periodica*) se $T>0$ è il piú piccolo numero reale tale che $f(x+T)=f(x)$ per ogni $x\in\mathbf{R}$. Se f è T-periodica, T si dice* periodo *della funzione f.*

Osserviamo che se f è T-periodica allora $f(x+kT)=f(x)$ per ogni $x\in\mathbf{R}$ e per ogni $k\in\mathbf{Z}$.

Passiamo a definire le funzioni trigonometriche.

Definizione 1.24 *Si dicono rispettivamente* seno *e* coseno *del numero $\alpha\in\mathbf{R}$ l'ordinata e l'ascissa del punto della circonferenza unitaria con centro l'origine determinato dalla semiretta per l'origine che forma un angolo di α radianti col semiasse positivo dell'asse x.*

Osservazioni 1.25

1. Nella definizione di seno e coseno abbiamo sottolineato che si deve parlare di seno e coseno di un *numero* e non di un angolo: la costruzione geometrica basata sull'angolo è strumentale (e per altro non è l'unica possibile), ma ciò che viene definito sono il seno e il coseno del numero che esprime la misura in radianti di un angolo, e non dell'angolo stesso. Quest'osservazione è importante al fine di evitare confusioni quando si usino unità diverse dal radiante per misurare gli angoli, Anche per questo, è bene usare solo i radianti per misurare gli angoli. Come per la funzione esponenziale, forse questa scelta può apparire ora innaturale, mentre al contrario, come vedremo, risulterà essere la più naturale possibile.

2. Come abbiamo già osservato, lo stesso punto della circonferenza unitaria è determinato da infiniti valori della misura in radianti dell'angolo: due valori che differiscono per un multiplo intero di 2π, infatti, determinano lo stesso punto.

3. In base alle osservazioni precedenti, la definizione 1.24 definisce le due funzioni sin e cos, aventi dominio $\mathbf{R}$. Come al solito, useremo d'ora in poi la lettera x per denotare la variabile. Inoltre, sono funzioni periodiche di periodo 2π.

4. È chiaro dalla definizione che vale la relazione
$$\sin^2 x + \cos^2 x = 1$$
per ogni $x \in \mathbf{R}$.

Accanto alle funzioni seno e coseno si definisce la funzione *tangente* ponendo
$$\tan : \mathbf{R} \setminus \{\frac{\pi}{2} + k\pi : \ k \in \mathbf{Z}\}, \qquad \tan x = \frac{\sin x}{\cos x}.$$
Geometricamente, detto P il punto sulla circonferenza unitaria determinato dal numero x come al solito, per ogni x nel dominio di tan, la tangente rappresenta l'ordinata del punto d'intersezione della retta $x = 1$ con la retta di origine O passante per P.

È evidente che la funzione tangente è definita per tutti i valori di x in cui il coseno è diverso da zero. Dal punto di vista geometrico, questi valori corrispondono ai punti P della circonferenza unitaria tali che la retta OP sia parallela alla retta $x = 1$.

Osserviamo anche che la funzione tangente è periodica di periodo minimo π e non 2π, malgrado sia rapporto di funzioni 2π-periodiche.

Ricordiamo infine alcuni valori fondamentali delle funzioni seno, coseno e tangente:

$$\begin{array}{lll}
\sin 0 = 0 & \cos 0 = 1 & \tan 0 = 0 \\
\sin \frac{\pi}{2} = 1 & \cos \frac{\pi}{2} = 0 & \nexists \tan \frac{\pi}{2} \\
\sin \frac{\pi}{4} = \frac{\sqrt{2}}{2} & \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2} & \tan \frac{\pi}{4} = 1 \\
\sin \frac{\pi}{3} = \frac{\sqrt{3}}{2} & \cos \frac{\pi}{3} = \frac{1}{2} & \tan \frac{\pi}{3} = \sqrt{3} \\
\sin \frac{\pi}{6} = \frac{1}{2} & \cos \frac{\pi}{6} = \frac{\sqrt{3}}{2} & \tan \frac{\pi}{6} = \frac{\sqrt{3}}{3}.
\end{array}$$

Le proprietà delle funzioni trigonometriche sono numerose ed importanti, ma non ne discutiamo ulteriormente qui perché esse si suppongono note dalla scuola superiore. Ci limitiamo a presentare le funzioni arcsin, *arccos* e *arctan* e a mostrare in che relazione sono con seno, coseno e tangente.

Proposizione 1.26 *La restrizione di* sin *all'intervallo* $[-\pi/2, \pi/2]$ *è strettamente crescente, la restrizione di* cos *all'intervallo* $[0, \pi]$ *è strettamente decrescente, e la restrizione di* tan *all'intervallo* $]-\pi/2, \pi/2[$ *è strettamente crescente.*

Non presentiamo la dimostrazione analitica della proposizione precedente, ma osserviamo che essa è chiara dalla costruzione geometrica delle funzioni trigonometriche. Dalla proposizione precedente e dall'Osservazione 1.16 segue subito che le restrizioni considerate sono invertibili. Diamo pertanto la seguente definizione.

Definizione 1.27 *Si dicono* funzioni trigonometriche inverse *le funzioni:*

- *la funzione* arcoseno, $\arcsin : [-1,1] \to [-\pi/2, \pi/2]$, *inversa della restrizione di* sin *all'intervallo* $[-\pi/2, \pi/2]$, *definita da*

$$x = \arcsin y \quad \Longleftrightarrow \quad y = \sin x \qquad \forall\ x \in [-\pi/2, \pi/2],\ \forall\ y \in [-1,1];$$

- *la funzione* arcocoseno, $\arccos : [-1,1] \to [0, \pi]$, *inversa della restrizione di* cos *all'intervallo* $[0, \pi]$, *definita da*

$$x = \arccos y \quad \Longleftrightarrow \quad y = \cos x \qquad \forall\ x \in [0, \pi],\ \forall\ y \in [-1,1];$$

- *la funzione* arcotangente, $\arctan : \mathbf{R} \to]-\pi/2, \pi/2[$, *inversa della restrizione di* tan *all'intervallo* $]-\pi/2, \pi/2[$, *definita da*

$$x = \arctan y \quad \Longleftrightarrow \quad y = \tan x \qquad \forall\ x \in]-\pi/2, \pi/2[,\ \forall\ y \in \mathbf{R}$$

Dalla Proposizione 1.26 e dall'Osservazione 1.16 segue che arcsin e arctan sono strettamente crescenti e che arccos è strettamente decrescente.

1.3 Numeri complessi

L'introduzione di un altro insieme numerico, oltre ai numeri reali, è giustificata dall'esigenza di trovare un insieme numerico in cui tutte le equazioni algebriche, cioè le equazioni del tipo $P(x) = 0$, con P polinomio di grado arbitrario, abbiano soluzioni. Questo non accade in $\mathbf{R}$; il più semplice esempio di equazione algebrica priva di soluzioni reali è certamente $x^2 + 1 = 0$. È forse sorprendente che, come vedremo, sarà in un certo senso sufficiente "aggiungere" ad $\mathbf{R}$ le soluzioni di quest'equazione per ottenere lo scopo indicato di risolvere *tutte* le equazioni algebriche (vedi il successivo Teorema fondamentale dell'algebra 1.37).

Definizione 1.28 (Campo complesso) *Si dice* campo complesso *l'insieme* $\mathbf{R}^2$ *in cui sono definite le due operazioni di somma e prodotto seguenti:*

$$(a,b)+(c,d) = (a+c, b+d) \tag{1.3.11}$$

$$(a,b)\cdot(c,d) = (ac-bd, bc+ad) \tag{1.3.12}$$

Il campo complesso si denota con $\mathbf{C}$.

Osservazioni 1.29

1. Le operazioni definite in (1.3.11), (1.3.12) godono delle stesse proprietà algebriche indicate negli Assiomi da 1 a 6 per le analoghe operazioni $\mathbf{R}$.

2. Gli elementi neutri per le operazioni di somma e prodotto sono rispettivamente $0 = (0,0)$ e $1 = (1,0)$, mentre l'opposto e l'inverso, quest'ultimo solo per $(a,b) \neq (0,0)$, sono rispettivamente

$$-(a,b) = (-a,-b), \qquad (a,b)^{-1} = \Big(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\Big).$$

3. Alla luce delle considerazioni precedenti, si ha che si possono identificare i numeri reali col sottoinsieme di $\mathbf{C}$ dato da $\{(a,0):\ a \in \mathbf{R}\}$. Ne segue che per $a \in \mathbf{R}$ si ha $a \cdot (b,c) = (ab, ac)$ per ogni $(b,c) \in \mathbf{C}$.

4. A differenza di $\mathbf{R}$, è possibile trovare numeri complessi (a,b) tali che $(a,b)^2 = -1 = (-1,0)$; per esempio, $(0,1)^2 = -1$.

5. In $\mathbf{C}$ non si può definire alcuna relazione d'ordine $\geq$ tale che valgano gli Assiomi 7, 8 e 9 stabiliti per i numeri reali.

6. Avendo definito la somma e la moltiplicazione tra numeri complessi, si possono considerare i *polinomi complessi* nella variabile complessa z nella forma

$$P(z) = a_0 + a_1 z + \ldots + a_n z^n, \quad \text{con } a_0, \ldots a_n \in \mathbf{C}.$$

A proposito dei polinomi e delle equazioni algebriche in campo complesso, dal momento che, come vedremo, essi hanno un ruolo importante nella teoria, ricordiamo le nozioni di *radice* e *molteplicità*.

Definizione 1.30 (Radice di un polinomio e molteplicità) *Sia* $P(z)$ *un polinomio nella variabile complessa* z, *sia* $z_0 \in \mathbf{C}$ *e sia* $m \in \mathbf{N}$, $m \geq 1$. *Si dice che* z_0 *è* radice *di* P *se* $P(z_0) = 0$; *in tal caso,* $P(z)$ *è divisibile per* $(z - z_0)$, *e si dice che* z_0 *ha* molteplicità m *se* $P(z)$ *è divisibile per* $(z-z_0)^m$ *ma non per* $(z-z_0)^{m+1}$.

Una conseguenza della definizione precedente è che un polinomio di grado n non può avere più di n radici (contate con le rispettive molteplicità).

La rappresentazione dei numeri complessi come coppie ordinate non è molto comoda nelle manipolazioni algebriche. Per questo introduciamo la seguente definizione.

Definizione 1.31 (Forma algebrica, Re, Im, modulo, coniugato) *Posto* $i = (0,1)$, *definiamo* forma algebrica del numero complesso $z = (a,b)$ *la scrittura* $z = a + ib$. *Il numero reale* a *si dice* parte reale di z, *denotata* $Re\, z$ *e il numero reale* b *si dice* parte immaginaria di z, *denotata* $Im\, z$. *Si definisce inoltre il* modulo di z *come* $|z| = \sqrt{a^2 + b^2}$ *e il numero* complesso coniugato di z *come* $\bar{z} = a - ib$.

Il modulo complesso gode di proprietà analoghe a quelle del valore assoluto in campo reale.

Proposizione 1.32 (Proprietà del modulo) *Per ogni* $z, w \in \mathbf{C}$ *valgono le seguenti proprietà:*

$$\begin{aligned}
&|z| \geq 0, \qquad |z| = 0 \iff z = 0; \\
&|z| = |-z|, \;\; |z \cdot w| = |z| \cdot |w|; \\
&|z + w| \leq |z| + |w|.
\end{aligned}$$

Segue che per le operazioni tra numeri complessi si possono usare le usuali regole algebriche, trattando il numero i come una variabile letterale, e tenendo conto che $i^2 = -1$. La forma algebrica è basata sulla rappresentazione dei punti del piano $\mathbf{R}^2$ attraverso le coordinate cartesiane ed è molto comoda per le addizioni, non è altrettanto comoda per le moltiplicazioni. Introduciamo perciò nel piano le *coordinate polari*, che ci consentiranno di ottenere un'altra espressione dei numeri complessi, stavolta utile in modo particolare per le moltiplicazioni. L'argomento sarà ripreso nel corso di Analisi Matematica II.

(Coordinate polari) Le coordinate polari (ϱ, ϑ) sono definite come segue: $\varrho = \sqrt{x^2 + y^2}$ rappresenta la distanza del punto generico di coordinate cartesiane (x, y) dall'origine e coincide col modulo, mentre ϑ rappresenta l'angolo formato dalla semiretta di origine $(0,0)$ e passante per (x, y) con il semiasse $\{x \geq 0, y = 0\}$ e si dice *argomento* o anche *anomalia* del numero complesso $z = x + iy$. Ne segue che il punto $(x, y) \neq (0,0)$ è univocamente determinato da una coppia (ϱ, ϑ), con $\varrho \geq 0$ e ϑ che varia in un intervallo semiaperto di ampiezza 2π. Fa eccezione l'origine, che è determinato dal valore $\varrho = 0$, ma non ha un ϑ determinato. In questo contesto è utile scegliere come intervallo di variabilità per l'angolo l'intervallo $]-\pi, \pi]$, e si ha:

$$\left\{ \begin{array}{l} x = \varrho \cos \vartheta \\ y = \varrho \sin \vartheta \end{array} \right. \tag{1.3.13}$$

Definizione 1.33 (Forma trigonometrica) *Ogni* $z \in \mathbf{C}$, *si può esprimere in* forma trigonometrica *usando le coordinate polari; risulta* $z = \varrho(\cos \vartheta + i \sin \vartheta)$.

Vediamo come la forma trigonometrica permette di eseguire facilmente le moltiplicazioni. Dati $z_1 = \varrho_1(\cos\vartheta_1 + i\sin\vartheta_1)$ e $z_2 = \varrho_2(\cos\vartheta_2 + i\sin\vartheta_2)$, risulta:

$$\begin{aligned} z_1 z_2 &= \varrho_1\varrho_2[(\cos\vartheta_1\cos\vartheta_2 - \sin\vartheta_1\sin\vartheta_2) + i(\sin\vartheta_1\cos\vartheta_2 + \sin\vartheta_2\cos\vartheta_1)] \\ &= \varrho_1\varrho_2[\cos(\vartheta_1+\vartheta_2) + i\sin(\vartheta_1+\vartheta_2)], \end{aligned}$$

risultato che si può descrivere dicendo che il prodotto di due numeri complessi ha per modulo il prodotto dei moduli e per argomento la somma degli argomenti. In particolare, per $z = \varrho(\cos\vartheta + i\sin\vartheta)$ ed $n \in \mathbf{N}$ si ha la *formula di De Moivre*:

$$z^n = \varrho^n[\cos(n\vartheta) + i\sin(n\vartheta)]. \tag{1.3.14}$$

Queste osservazioni suggeriscono di dare la seguente definizione.

Definizione 1.34 (Esponenziale complesso e forma esponenziale) *Per $z = x + iy$, si pone*

$$e^z = e^x(\cos y + i\sin y); \tag{1.3.15}$$

in particolare, per $x = 0$ si ottengono i numeri complessi di modulo unitario, sicché si può scrivere ogni numero complesso nella forma esponenziale $z = |z|e^{i\vartheta}$*, dove ϑ è un argomento di z.*

Notiamo che usando la forma esponenziale la formula appena vista per il prodotto di due numeri complessi e la formula di De Moivre si possono riscrivere

$$\varrho_1 e^{i\vartheta_1}\varrho_2 e^{i\vartheta_2} = \varrho_1\varrho_2\ e^{i(\vartheta_1+\vartheta_2)}, \qquad (\varrho e^{i\vartheta})^n = \varrho^n e^{in\vartheta},$$

due uguaglianze che formalmente rispettano le leggi degli esponenti per il prodotto di potenze con la stessa base. Anche se la definizione 1.34 ha delle giustificazioni ben più profonde, questa coerenza formale si può considerare come una prima motivazione. Le notazioni introdotte saranno comode nel calcolo delle radici complesse che ora definiamo.

Definizione 1.35 *Dati $n \in \mathbf{N}$, $n \geq 2$, e $w \in \mathbf{C}$, un numero complesso z si dice* radice complessa n-esima di w *se $z^n = w$.*

Notiamo che un numero complesso non può avere più di n radici complesse distinte, essendo le radici n-esime di w le radici del polinomio $P(z) = z^n - w$. Inoltre, è importante capire che la nozione di radice complessa, a differenza della radice n-esima aritmetica in campo reale, *non definisce una funzione*, anche quando si parta da numeri reali positivi. Per esempio, la *radice quadrata reale* del numero 4 è 2, e infatti in ambito reale si scrive $\sqrt{4} = 2$, mentre le *radici complesse* dello *stesso numero* 4 sono 2 e -2. Di che cosa si stia parlando dev'essere pertanto sempre dichiarato o chiaro dal contesto.

Teorema 1.36 (Radici complesse) *Per ogni $n \in \mathbf{N}$, $n \geq 2$, e per ogni $w \in \mathbf{C}$, $w \neq 0$, esistono n radici complesse distinte di w, date dalla formula*

$$z_k = \sqrt[n]{|z|}e^{i\phi_k}, \quad \textit{con } \phi_k = \frac{\vartheta + 2k\pi}{n} \quad \textit{e } k = 0, \ldots, n-1. \tag{1.3.16}$$

Dim. Posto $w = \varrho e^{i\vartheta}$, cerchiamo i numeri complessi $z = re^{i\phi}$ tali che $z^n = w$. Dalla formula di De Moivre, quest'equazione equivale a $r^n e^{in\phi} = \varrho e^{i\vartheta}$, che a sua volta equivale al sistema

$$\begin{cases} r^n = \varrho \\ \cos(n\phi) = \cos\vartheta \\ \sin(n\phi) = \sin\vartheta \end{cases}$$

nelle incognite reali r e ϕ. Le soluzioni di questo sistema sono

$$r = \sqrt[n]{\varrho}, \qquad \phi_k = \frac{\vartheta + 2k\pi}{n}, \; k \in \mathbf{Z},$$

dove $\sqrt[n]{\ }$ indica la radice aritmetica del numero positivo ϱ. Quindi r, cioè il modulo di z, è unico, mentre per l'argomento abbiamo trovato infinite soluzioni ϕ_k. In realtà, esse non dànno luogo ad infiniti numeri complessi distinti (che sarebbe impossibile, come già osservato), a causa del fatto che l'argomento di un numero complesso è determinato a meno di multipli di 2π. Infatti, $z_k = z_j$ se e solo se $\cos\phi_k = \cos\phi_j$ e $\sin\phi_k = \sin\phi_j$, cioè se e solo se i due indici k e j differiscono per un multiplo intero di n. Ne segue che i numeri $z_0, \ldots, z_{n-1}$ sono tutti distinti tra loro e che ogni altra radice trovata coincide con uno di questi numeri. QED

Come annunciato all'inizio della sezione, la possibilità di trovare soluzioni dell'equazione $w^n = z$, cioè radici complesse di tutti i numeri complessi, si estende a tutte le equazioni algebriche. Infatti, il seguente importante (e difficile!) teorema assicura che ogni polinomio complesso ammette almeno una radice complessa.

Teorema 1.37 (Teorema fondamentale dell'algebra) *Se $P(z)$ è un polinomio complessodi grado almeno 1, esiste almeno un numero complesso z_0 tale che $P(z_0) = 0$.*

È facile trarre dal Teorema fondamentale dell'algebra numerose conseguenze: in particolare, usando la nozione di molteplicità di una radice, si può concludere che ogni polinomio di grado n ha esattamente n radici, pur di contarle con le rispettive molteplicità.

Teorema 1.38 (Teorema fondamentale dell'algebra) *Ogni polinomio complesso di grado n ha esattamente n radici, se si conta m volte una radice di molteplicità m.*

Non abbiamo al momento bisogno di queste ulteriori informazioni. L'argomento sarà ripreso nel corso di Analisi matematica II. Enunciamo invece subito un risultato riguardante i polinomi a coefficienti reali.

Teorema 1.39 (Polinomi a coefficienti reali) *Se $P(z)$ è un polinomio nella variabile complessa z i cui coefficienti sono tutti numeri reali, allora $z_0 \in \mathbf{C}$ è radice di $P(z)$ se e solo se il suo coniugato $\bar{z}_0$ lo è, ed in tal caso z_0 e $\bar{z}_0$ hanno la stessa molteplicità.*

Dim. Basta osservare che se $P(z_0) = 0$ allora

$$P(\bar{z}_0) = \overline{P(z_0)} = 0.$$

QED

CAPITOLO 2

SUCCESSIONI

2.1 Limiti di successioni

2.2 Principio di induzione

2.3 Limiti notevoli

CAPITOLO 3

SERIE NUMERICHE

Trattiamo ora il problema di associare ad una successione (infinita) di numeri reali una procedura che generalizzi il calcolo della somma così come si esegue su un numero finito di termini. Ovviamente, non è possibile *neanche in linea di principio* "sommare fra loro infiniti numeri" e quindi bisognerà fare ricorso ancora una volta al concetto di limite. Iniziamo, come sempre, col definire l'oggetto del nostro discorso, cioè le serie numeriche. Studieremo nel corso di Analisi matematica II il caso più generale delle serie di funzioni.

3.1 Serie, convergenza, convergenza assoluta

Definizione 3.1 *Data la successione reale* (a_k)*, si dice* serie di termine generale a_k *l'operazione che associa alla successione* (a_k) *la successione* (s_n) *definita per ogni* $n \in \mathbf{N}$ *da*

$$s_n = \sum_{k=0}^{n} a_k, \tag{3.1.1}$$

che si dice successione delle somme parziali della serie. *La serie di termine generale* a_k *si denota con la scrittura*

$$\sum_{k=0}^{\infty} a_k.$$

Notiamo che la scrittura introdotta per denotare le serie, facendo uso del simbolo di sommatoria, evoca il problema descritto all'inizio, ma *non denota una somma.*

Osservazione 3.2 La successione delle somme parziali si dice anche *successione delle ridotte*, e si può definire per ricorrenza così

$$s_0 = a_0, \qquad s_n = s_{n-1} + a_n, \quad n \geq 1. \tag{3.1.2}$$

Vediamo ora qual è la proprietà più rilevante delle serie.

Definizione 3.3 *Si dice che la serie di termine generale* a_k *è* convergente *se esiste finito il limite*

$$\lim_{n\to+\infty} s_n.$$

In caso affermativo, il valore di tale limite si dice somma della serie.

Vediamo una prima proprietà delle serie convergenti; intuitivamente, è naturale aspettarsi che, se una serie converge, il termine generale debba diventare sempre più piccolo. Quest'idea è formalizzata nel seguente enunciato.

Proposizione 3.4 *Se la serie* $\sum_{k=0}^{\infty} a_k$ *è convergente allora* $\lim_{k\to\infty} a_k = 0$.

DIM. Per definizione, dire che la serie data converge equivale a dire che la successione (s_n) delle ridotte converge; ma allora, detta S la somma della serie, si ha

$$S = \lim_{n\to+\infty} s_n = \lim_{n\to+\infty} s_{n-1},$$

da cui, essendo il limite finito,

$$\lim_{n\to+\infty} a_n = \lim_{n\to+\infty}(s_n - s_{n-1}) = S - S = 0.$$

QED

La proposizione precedente esprime una condizione solo *necessaria* per la convergenza di una serie. Essa non è però sufficiente, come mostra l'esempio 3.11.

Osservazioni 3.5

1. Notiamo che la condizione di convergenza per la serie di termine generale a_k è l'esistenza del limite di un'altra successione, quella delle ridotte (s_n).

2. Il nome *somma* riservato al limite delle somme parziali, quando esiste finito, evoca di nuovo l'idea di generalizzare la somma tra numeri. Come abbiamo visto, in realtà non si tratta di una *somma* (cosa impossibile, come già osservato) ma di un *limite*.

3. Se la serie di termine generale a_k non è convergente, può accadere che il limite delle somme parziali sia infinito o non esista. Si dirà in tal caso che la serie è

   - *positivamente divergente* se $\lim_{n\to+\infty} s_n = +\infty$;
   - *negativamente divergente* se $\lim_{n\to+\infty} s_n = -\infty$;
   - *indeterminata* se il limite $\lim_{n\to+\infty} s_n$ non esiste.

4. Se la serie di termine generale a_k è convergente, la sua somma si denota con

$$\sum_{k=0}^{\infty} a_k$$

   e lo stesso simbolo si usa anche per indicare che la serie è divergente positivamente, scrivendo

$$\sum_{k=0}^{\infty} a_k = +\infty$$

o negativamente, scrivendo

$$\sum_{k=0}^{\infty} a_k = -\infty.$$

Tale simbolo assume quindi due significati diversi (la serie o la sua somma), ma si vedrà che questo non genera alcuna confusione; infatti, assegnando una serie si pone naturalmente il problema di sapere se essa sia convergente o no, e, in caso affermativo, di determinarne la somma.

5. Si dice *carattere* della serie la sua proprietà di essere convergente, divergente o indeterminata.

6. Se si modifica un numero *finito* di termini della serie $\sum_{k=0}^{\infty} a_k$ non se ne altera il carattere. In altri termini, due serie che differiscono per un numero finito di termini hanno lo stesso carattere. Naturalmente, esse *non hanno in generale la stessa somma.*

7. Dati una serie $\sum_k a_k$ e un numero reale $c \neq 0$, si può considerare la serie $\sum_k (ca_k)$. Ovviamente, questa ha lo stesso carattere della $\sum_k a_k$ e, in caso di convergenza, la sua somma è il prodotto di c per la somma della serie $\sum_k a_k$. In particolare, studieremo le serie a termini positivi, ma in realtà è chiaro che la classe a cui si riferiranno tutti i risultati è quella delle *serie a termini di segno costante* e non importa se tale segno sia $+$ o $-$.

8. Il discorso è più delicato per le serie che sono *somma di due serie.* Per il momento, quello che possiamo dire è che se $\sum_k a_k$ è convergente e $\sum_k b_k$ è convergente, allora anche la serie $\sum_k (a_k + b_k)$ è convergente, e questo segue subito dal teorema sulla somma dei limiti applicato alle successioni delle ridotte. Invece, se $\sum_k a_k$ è convergente e $\sum_k b_k$ *non* è convergente, allora la serie $\sum_k (a_k + b_k)$ *non può convergere.* Infatti, applicando il risultato precedente alle serie $\sum_k (a_k + b_k)$ e $\sum_k (-a_k)$ dedurremmo che $\sum_k b_k$ converge, che è contrario all'ipotesi.

Il seguente risultato segue subito dal criterio di convergenza di Cauchy per le successioni.

Teorema 3.6 (Criterio di Cauchy per le serie) *La convergenza della serie $\sum_k a_k$ è equivalente alla validità della seguente* condizione di Cauchy*:*

$$\textit{per ogni } \varepsilon > 0 \textit{ esiste } \nu > 0 \textit{ tale che } \Big| \sum_{k=\nu+1}^{\nu+p} a_k \Big| < \varepsilon \textit{ per ogni } p \in \mathbf{N}.$$

Dim. Ricordando il criterio di Cauchy per le successioni ??, basta applicarlo alla successione (s_n) delle ridotte, e si ha che (s_n) converge se e solo se per ogni $\varepsilon > 0$ esiste $\nu > 0$ tale che $n > \nu$ implica $|s_{n+p} - s_n| < \varepsilon$ per ogni $p \in \mathbf{N}$. A questo punto, resta solo da osservare che

$$|s_{n+p} - s_n| = \Big| \sum_{k=\nu+1}^{\nu+p} a_k \Big|.$$

$\boxed{QED}$

Studiamo in dettaglio un esempio molto importante che sarà di guida in numerose applicazioni.

Esempio 3.7 Per ogni $x \in \mathbf{R}$, consideriamo la serie di termine generale x^k, cioè la serie

$$\sum_{k=0}^{\infty} x^k;$$

essa è detta *serie geometrica di ragione x*. È uno dei (pochi) casi in cui si riescono a calcolare esplicitamente, al variare di x, le somme parziali s_n. Anzi, è in realtà un lavoro già fatto: nell'esempio ?? abbiamo ottenuto la formula

$$s_n = \frac{1 - x^{n+1}}{1 - x}, \qquad x \neq 1, \tag{3.1.3}$$

(per $x = 1$ è ovvio che $s_n = n + 1$ e la serie è positivamente divergente). Possiamo allora facilmente calcolare il limite della successione (s_n) e determinare il carettere della serie geometrica. Risulta:

$$\lim_{n\to+\infty} s_n \begin{cases} \text{non esiste} & x \leq -1 \\ = \dfrac{1}{1-x} & -1 < x < 1 \\ = +\infty & x \geq 1 \end{cases}$$

e quindi la serie geometrica di ragione x converge se e solo se $-1 < x < 1$, diverge positivamente per $x \geq 1$ ed è indeterminata per $x \leq -1$.

Introduciamo ora la nozione di *convergenza assoluta*, più forte della convergenza semplice.

Definizione 3.8 (Convergenza assoluta) *Si dice che la serie* $\sum_k a_k$ converge assolutamente *se la serie* $\sum_k |a_k|$ *converge.*

Vediamo che la convergenza assoluta implica quella semplice.

Proposizione 3.9 *Se* $\sum_k |a_k|$ *converge allora anche* $\sum_k a_k$ *converge.*

DIM. Usiamo il criterio di convergenza di Cauchy. Per ipotesi, fissato $\varepsilon > 0$, esiste $\nu > 0$ tale che

$$\sum_{k=\nu+1}^{\nu+p} |a_k| < \varepsilon \qquad \forall\, p \in \mathbf{N}.$$

Poiché per la diseguaglianza triangolare (1.1.4) risulta

$$\Big|\sum_{k=\nu+1}^{\nu+p} a_k\Big| \leq \sum_{k=\nu+1}^{\nu+p} |a_k|$$

in corrispondenza dello stesso ν si ha pure

$$\Big|\sum_{k=\nu+1}^{\nu+p} a_k\Big| < \varepsilon \qquad\qquad \forall\, p \in \mathbf{N}$$

e la tesi segue usando l'altra implicazione del criterio di Cauchy. QED

Vedremo nell'Osservazione 3.25.4 che invece esistono serie che convergono semplicemente ma non assolutamente.

3.2 Serie a termini positivi

Tra le serie, hanno un posto di rilievo quelle i cui termini hanno segno costante. Per fissare le idee (vedi Osservazione 3.5.7), supponiamo che i termini delle serie che consideriamo siano tutti positivi. Tali serie, dette per l'appunto *serie a termini positivi*, sono particolarmente importanti anche in relazione alla proprietà espressa dalla Proposizione 3.9: infatti, studiare la convergenza assoluta di una qualunque serie $\sum_k a_k$ consiste nello studiare la convergenza della serie a termini positivi $\sum_k |a_k|$ e, se la serie risulta assolutamente convergente, resta provata anche la convergenza semplice, sicché il problema è completamente risolto. Inoltre, per le serie a termini positivi esistono numerosi criteri di convergenza. Una prima loro proprietà è la seguente.

Proposizione 3.10 *Una serie a termini positivi può essere convergente o positivamente divergente, ma non è mai indeterminata. In particolare, una serie a termini positivi è convergente se e solo se la successione delle sue ridotte è limitata.*

DIM. Data la serie $\sum_k a_k$, se $a_k \geq 0$ per ogni $k \in \mathbf{N}$ allora la successione s_n delle ridotte è monotona crescente, perché

$$s_{n+1} = s_n + a_{n+1} \geq s_n \qquad\qquad \forall\, n \in \mathbf{N}$$

e quindi, per il Teorema fondamentale delle successioni monotone (Teorema ??), ammette limite, pari a $\sup_n s_n$. Segue che la serie data non può essere indeterminata, ed è positivamente divergente se $\sup_n s_n = +\infty$, è convergente se $\sup_n s_n$ è finito, cioè se e solo se la successione delle ridotte è limitata. QED

Studiamo ora una serie di fondamentale importanza, detta *serie armonica*. In particolare, essa è una serie che non converge, malgrado il suo termine generale tenda a 0. Questo, come già sottolineato, mostra che la condizione nella Proposizione 3.4 è *solo necessaria* per la convergenza di una serie, ma non sufficiente.

Esempio 3.11 (Serie armonica) La serie

$$\sum_{k=1}^{\infty} \frac{1}{k},$$

detta *serie armonica*, è positivamente divergente. Infatti, essendo una serie a termini positivi, per la Proposizione 3.10 esiste il limite della successione delle ridotte,

$$S = \lim_{n\to+\infty} s_n = \lim_{n\to+\infty} \sum_{k=1}^{n} \frac{1}{k}.$$

Proviamo che $S = +\infty$. In accordo con l'Osservazione ??, possiamo conoscere S calcolando il limite di una qualsiasi successione estratta da (s_n). Scegliamo l'estratta s_{2^k}, cioè la sottosuccessione della (s_n) ottenuta arrestandosi di volta in volta quando l'indice è una potenza di 2. Allora otteniamo

$$\begin{aligned}
s_{2^k} &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{8} + \cdots + \frac{1}{2^{k-1}} + \frac{1}{2^{k-1}+1} + \cdots + \frac{1}{2^k} \\
&= 1 + \frac{1}{2} + \Big(\frac{1}{3} + \frac{1}{4}\Big) + \Big(\frac{1}{5} + \cdots + \frac{1}{8}\Big) + \cdots + \frac{1}{2^{k-1}} + \Big(\frac{1}{2^{k-1}+1} + \cdots + \frac{1}{2^k}\Big) \\
&\geq 1 + \frac{1}{2} + \Big(\frac{1}{4} + \frac{1}{4}\Big) + \Big(\frac{1}{8} + \cdots + \frac{1}{8}\Big) + \cdots\cdots + \Big(\frac{1}{2^k} + \cdots + \frac{1}{2^k}\Big) \\
&= 1 + \frac{1}{2} + 2\frac{1}{4} + 4\frac{1}{8} + \cdots + 2^{k-1}\frac{1}{2^k} = 1 + \sum_{j=1}^{k} 2^{j-1}\frac{1}{2^j} \\
&= 1 + k\frac{1}{2} \longrightarrow +\infty \qquad \text{per} \quad k \to +\infty.
\end{aligned}$$

Per quanto detto, l'intera successione (s_n) delle ridotte tende a $+\infty$, e quindi la serie armonica diverge positivamente.

Per le serie a termini positivi vale il seguente criterio di confronto.

Teorema 3.12 (Criterio di confronto) *Siano $\sum_k a_k$ e $\sum_k b_k$ due serie a termini positivi, e supponiamo che esista $\nu > 0$ tale che $a_k \leq b_k$ per ogni $k > \nu$. Allora:*

1. *se $\sum_k b_k$ converge allora anche $\sum_k a_k$ converge;*

2. *se $\sum_k a_k$ diverge allora anche $\sum_k b_k$ diverge.*

Dim. Poiché, come abbiamo visto, la convergenza di una serie non cambia modificando un numero finito di termini, possiamo supporre che $0 \leq a_k \leq b_k$ per ogni $k \in \mathbf{N}$ e non solo per $k \geq \nu$. Chiamando (s_n) e (t_n) le successioni delle ridotte di $\sum_k a_k$ e $\sum_k b_k$ rispettivamente, ponendo cioè

$$s_n = \sum_{k=0}^{n} a_k, \qquad t_n = \sum_{k=0}^{n} b_k,$$

entrambe le successioni (s_n) e (t_n) sono crescenti e risulta evidentemente che $s_n \leq t_n$ per ogni n. La tesi allora segue dalla Proposizione 3.10.

Infatti, se $\sum_k b_k$ converge, la successione (t_n) è limitata, quindi lo è anche la (s_n) e la serie $\sum_k a_k$ converge e questo prova 1.

Viceversa, se $\sum_k a_k$ diverge, allora (per definizione) la successione (s_n) diverge, ed ugualmente diverge la (t_n). Segue che $\sum_k b_k$ diverge e questo prova 2. $\boxed{QED}$

Una semplice applicazione della teoria vista fin qui, basata sulla serie geometrica, è nel seguente esempio.

Esempio 3.13 (Sviluppi decimali) Abbiamo tutti familiarità con la rappresentazione decimale (anche con infinite cifre, si pensi a numeri come $\sqrt{2}$, e o π) dei numeri reali, di cui la conoscenza delle serie numeriche dà una completa spiegazione. Un numero reale x si può rappresentare in forma decimale così:

$$x = n, a_1 a_2 a_3 \ldots a_k \ldots \tag{3.2.4}$$

ove $n \in \mathbf{Z}$ è la *parte intera*, le cifre a_k, numeri interi compresi tra 0 e 9, costituiscono la *parte decimale* di x. La scrittura (3.2.4) ha il significato dell'uguaglianza

$$x = n + \sum_{k=1}^{\infty} \frac{a_k}{10^k}, \tag{3.2.5}$$

che è coerente con il linguaggio usuale, secondo cui a_1 sono i *decimi*, a_2 i *centesimi*, a_3 i *millesimi*, eccetera. Notiamo che siccome $a_k \le 9$ per ogni k, risulta $\frac{a_k}{10^k} \le \frac{9}{10^k}$. Quest'ultimo è il termine generale della serie geometrica di ragione $\frac{1}{10} < 1$ (moltiplicato per 9), che converge. Per confronto, *tutte* le serie del tipo (3.2.5) che dànno gli sviluppi decimali convergono. Come applicazione, si possono dedurre senza difficoltà le proprietà degli sviluppi decimali *periodici* in relazione alla rappresentazione frazionaria. Il più semplice esempio è l'uguaglianza $1 = 0,\overline{9}$, che si ottiene sommando la serie corrispondente:

$$0,\overline{9} = \sum_{k=1}^{\infty} \frac{9}{10^k} = 9 \sum_{k=1}^{\infty} \frac{1}{10^k} = 9\Big(\frac{1}{1-\frac{1}{10}} - 1\Big) = 9\Big(\frac{10}{9} - 1\Big) = 1.$$

In particolare, questo mostra che lo sviluppo decimale in generale non è unico. Più in generale, la stessa teoria si applica agli *sviluppi in base b qualunque*: dato un numero naturale b (i più usati in questo contesto sono, oltre a 10, le potenze di 2), la corrisponente rappresentazione in base b è data dalla serie

$$x = n + \sum_{k=1}^{\infty} \frac{a_k}{b^k},$$

dove stavolta le cifre a_k sono interi compresi tra 0 e $b-1$ e la serie converge per confronto con la serie geometrica di ragione $\frac{1}{b} < 1$.

Osservazione 3.14 Siano $\sum_k a_k$ e $\sum_k b_k$ due serie a termini positivi, con $b_k > 0$, e supponiamo che esistano $\nu > 0$ e due costanti $c_1, c_2 > 0$ tali che

$$c_1 \le \frac{a_k}{b_k} \le c_2$$

per $k \geq \nu$. Allora $\sum_k a_k$ e $\sum_k b_k$ hanno lo stesso carattere.

Infatti, se $\sum_k a_k$ diverge o $\sum_k b_k$ converge, la tesi segue dal teorema di confronto in virtù delle diseguaglianze

$$a_k \leq c_2 b_k,$$

mentre se $\sum_k a_k$ converge o $\sum_k b_k$ diverge la tesi segue dal teorema di confronto in virtù delle diseguaglianze

$$c_1 b_k \leq a_k.$$

Un risultato che si deduce facilmente dal criterio di confronto è il seguente criterio, basato sul calcolo di un limite.

Teorema 3.15 (Criterio di confronto asintotico) *Siano $\sum_k a_k$ e $\sum_k b_k$ due serie a termini positivi, con $b_k > 0$, e supponiamo che esista*

$$\lim_{k\to\infty} \frac{a_k}{b_k} = c.$$

Allora:

1. *se $0 < c < +\infty$ allora $\sum_k a_k$ e $\sum_k b_k$ hanno lo stesso carattere;*
2. *se $c = 0$ e $\sum_k b_k$ converge allora anche $\sum_k a_k$ converge;*
3. *se $c = 0$ e $\sum_k a_k$ diverge allora anche $\sum_k b_k$ diverge;*
4. *se $c = +\infty$ e $\sum_k a_k$ converge allora anche $\sum_k b_k$ converge;*
5. *se $c = +\infty$ e $\sum_k b_k$ diverge allora anche $\sum_k a_k$ diverge.*

Dim. Per il punto 1, osserviamo che, usando la definizione di limite con $\varepsilon = \frac{c}{2}$, per k abbastanza grande risulta

$$\frac{c}{2} \leq \frac{a_k}{b_k} \leq \frac{3}{2}c$$

e quindi

$$\frac{c}{2} b_k \leq a_k \leq \frac{3}{2} c\ b_k$$

e il risultato segue dall'osservazione 3.14.

Per i punti 2 e 3, fissato un $\varepsilon > 0$ arbitrario, per k abbastanza grande risulta

$$a_k \leq \varepsilon b_k$$

e la tesi segue dal Criterio di confronto.

Analogamente, per i punti 4 e 5, fissato un $M > 0$ arbitrario, per k abbastanza grande risulta

$$a_k \geq M b_k$$

e la tesi segue ancora dal Criterio di confronto. $\boxed{QED}$

Presentiamo un ulteriore criterio di convergenza per le serie a termini positivi, noto come *criterio di condensazione* (in realtà ne esistono molti altri che tralasciamo), di cui vedremo l'utilità nell'esempio 3.17. In un certo senso, è la forma astratta del ragionamento seguito nell'esempio 3.11.

Teorema 3.16 *Sia (a_k) una successione decresente di numeri positivi. Allora le due serie*

$$\sum_{k=0}^{\infty} a_k \qquad e \qquad \sum_{k=1}^{\infty} 2^{k-1} a_{2^k}$$

hanno lo stesso carattere.

Non riportiamo la dimostrazione del criterio, ma lo applichiamo negli esempi che seguono.

Esempio 3.17

1. Consideriamo la serie armonica $\sum 1/k$ e applichiamo il criterio di condensazione. Risulta che essa converge se e solo se la serie

   $$\sum_k 2^{k-1} \frac{1}{2^k} = \sum \frac{1}{2}$$

   converge. Poiché quest'ultima chiaramente non converge (il suo termine generale non tende a 0), non converge neanche la serie armonica.

2. Consideriamo la *serie armonica generalizzata*

   $$\sum_{k=1}^{\infty} \frac{1}{k^\alpha}, \qquad \alpha \in \mathbf{R}.$$

   Se $\alpha \leq 0$ allora la condizione necessaria 3.4 è banalmente violata e la serie non converge. Se $\alpha > 0$ allora si può applicare il criterio di condensazione e concludere che la serie data converge se e solo se $\alpha > 1$:

   $$\sum_{k=1}^{\infty} 2^{k-1} \frac{1}{(2^k)^\alpha} = \frac{1}{2} \sum_{k=1}^{\infty} (2^{1-\alpha})^k,$$

   che converge se e solo se $\alpha > 1$, essendo una serie geometrica di ragione $2^{1-\alpha}$.

3. Consideriamo la serie

   $$\sum_{k=2}^{\infty} \frac{1}{k \log^\alpha k}, \qquad \alpha \in \mathbf{R}.$$

Se $\alpha \le 0$ allora vale la relazione $\frac{1}{k\log^\alpha k} \ge \frac{1}{k}$ e quindi la serie diverge per confronto con la serie armonica. Se $\alpha > 0$ allora si può applicare il criterio di condensazione e concludere che la serie data converge se e solo se $\alpha > 1$:

$$\sum_{k=2}^{\infty} 2^{k-1}\frac{1}{2^k\log^\alpha(2^k)} = \frac{1}{2\log^\alpha 2}\sum_{k=2}^{\infty}\frac{1}{k^\alpha}.$$

4. Consideriamo la serie
$$\sum_{k=2}^{\infty}\frac{1}{(\log k)^{\log k}}$$
Applicando il Criterio di condensazione, se ne può determinare il carattere studiando la serie
$$\sum_{k=2}^{\infty} 2^k\frac{1}{(k\log 2)^{k\log 2}} = \sum_{k=2}^{\infty}\Big(\frac{2}{k^{\log 2}(\log 2)^{\log 2}}\Big)^k$$
che converge per il criterio della radice.

Si può dare un criterio di convergenza basato sul confronto con le serie armoniche generalizzate.

Corollario 3.18 *Data la serie a termini positivi $\sum_k a_k$, se esiste finito*

$$\lim_{k\to\infty} k^\alpha a_k = c > 0$$

allora $\sum_k a_k$ converge se $\alpha > 1$ e diverge se $\alpha \le 1$.

DIM. Il criterio di confronto asintotico (Teorema 3.15) implica che $\sum_k a_k$ e $\sum_k k^{-\alpha}$ hanno lo stesso carattere, quindi il risultato segue subito usando i risultati dell'Esempio 5.46. $\boxed{QED}$

Presentiamo ora due utili criteri di convergenza, basati sul confronto con la serie geometrica.

Teorema 3.19 (Criterio della radice) *Sia $\sum_k a_k$ una serie a termini positivi; se esistono $\nu > 0$ e $h \in [0,1[$ tali che $\sqrt[k]{a_k} \le h$ per ogni $h > \nu$ allora la serie $\sum_k a_k$ converge. Se invece $\sqrt[k]{a_k} \ge 1$ per infiniti valori di k allora la serie non converge.*

DIM. Dall'ipotesi segue immediatamente che $a_k \le h^k$ e la tesi segue per confronto con la serie $\sum_k h^k$, che è convergente perché $h < 1$.

Viceversa, se $\sqrt[k]{a_k} \ge 1$ per infiniti valori di k allora la condizione $a_k \to 0$ è violata e la serie non converge. $\boxed{QED}$

Teorema 3.20 (Criterio del rapporto) *Sia $\sum_k a_k$ una serie a termini strettamente positivi; se esistono $\nu > 0$ e $h \in [0,1[$ tale che $\frac{a_{k+1}}{a_k} \le h$ per ogni $k > \nu$ allora la serie $\sum_k a_k$ converge. Se invece esiste $\nu > 0$ tale che $\frac{a_{k+1}}{a_k} \ge 1$ per $k > \nu$ allora la serie $\sum_k a_k$ diverge.*

Dim. Supponiamo per semplicità che sia $\nu = 0$. Dall'ipotesi segue la diseguaglianza

$$a_{k+1} = \frac{a_{k+1}}{a_k} \cdot \frac{a_k}{a_{k-1}} \cdots \frac{a_1}{a_0} \cdot a_0 \leq h^k a_0$$

e la tesi segue per confronto con la serie $\sum_k h^k$, che è convergente perché $h < 1$.

Se invece $\frac{a_{k+1}}{a_k} \geq 1$ per ogni k allora $a_{k+1} \geq a_k$ ogni k, sicché la condizione necessaria $a_k \to 0$ è violata e la serie non può convergere. $\boxed{QED}$

I due precedenti criteri hanno una versione che è spesso utile ed è basata sul calcolo di un limite, che può risultare più semplice della ricerca di un maggiorante.

Corollario 3.21 *Sia $\sum_k a_k$ una serie a termini positivi, e supponiamo che esista il limite*

$$\lim_{k\to\infty} \sqrt[k]{a_k} = \ell.$$

Allora:

1. *se $0 \leq \ell < 1$ allora la serie converge;*
2. *se $\ell > 1$ allora la serie diverge;*
3. *se $\ell = 1$ allora non si può trarre alcuna conclusione.*

Dim. Se $\ell < 1$ allora, fissato $\ell < h < 1$, esiste $\nu > 0$ tale che per $k > \nu$ risulti $\sqrt[k]{a_k} \leq h$ e la tesi segue dal Teorema 3.19.

Se $\sqrt[k]{a_k} \geq 1$ allora $a_k \geq 1$ per ogni k, sicché la condizione necessaria $a_k \to 0$ è violata e la serie non può convergere.

Infine, per provare il punto 3 basta considerare le serie $\sum_k \frac{1}{k}$ e $\sum_k \frac{1}{k^2}$. Per entrambe $\ell = 1$, ma la prima diverge e seconda converge. $\boxed{QED}$

Corollario 3.22 *Sia $\sum_k a_k$ una serie a termini strettamente positivi, e supponiamo che esista il limite*

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \ell.$$

Allora:

1. *se $0 \leq \ell < 1$ allora la serie converge;*
2. *se $\ell > 1$ allora la serie diverge;*
3. *se $\ell = 1$ allora non si può trarre alcuna conclusione.*

Dim. Se $\ell < 1$ allora, fissato $\ell < h < 1$, esiste $\nu > 0$ tale che per $k > \nu$ risulti $\frac{a_{k+1}}{a_k} \leq h$ e la tesi segue dal Teorema 3.20.

Se $\ell \geq 1$ allora $a_k \geq a_0 > 0$ per ogni k, sicché la condizione necessaria $a_k \to 0$ è violata e la serie non può convergere.

Infine, per provare il punto 3, come per il Corollario 3.21 basta considerare le serie $\sum_k \frac{1}{k}$ e $\sum_k \frac{1}{k^2}$. Per entrambe $\ell = 1$, ma la prima diverge e seconda converge. $\boxed{QED}$

Osservazione 3.23 Sia data la successione (a_k), con $a_k > 0$ per ogni k. Si può dimostrare il seguente risultato:

> Se esiste
> $$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \ell$$
> allora esiste anche il $\lim\limits_{k\to\infty} \sqrt[k]{a_k}$ e vale ℓ.

Di conseguenza, se si deve studiare la serie $\sum_k a_k$ e il criterio del rapporto sotto forma di limite dà $\ell = 1$, è inutile ritentare col criterio della radice, che non può che fornire lo stesso risultato. D'altra parte, segnaliamo che può esistere il limite della radice senza che esista quello del rapporto.

3.3 Serie a termini di segno variabile

Nel paragrafo precedente abbiamo discusso vari criteri di convergenza per serie a termini positivi, che, come abbiamo già osservato, sono anche criteri di convergenza assoluta per serie a termini di segno qualunque. Forse val la pena di spiegare che per "serie a termini di segno qualunque" si intendono serie

$$\sum_{k=0}^{\infty} \varepsilon_k a_k$$

dove $a_k > 0$ per ogni k e gli ε_k possono assumere i valori 1 e -1 con *qualunque* alternanza di segno. Per tali serie non esiste una teoria, né esistono criteri di convergenza *semplice.* Pertanto, tutto ciò che si può fare è studiare caso per caso ciascuna singola serie (il che esula dagli scopi di un corso introduttivo come questo) oppure studiare direttamente la convergenza assoluta con i metodi del paragrafo precedente, ben sapendo che se la serie non è assolutamente convergente resta aperto il problema di stabilire se sia o no semplicemente convergente.

Un caso particolare delle serie a termini di segno variabile è quello delle *serie a segni alterni*, cioè il caso in cui $\varepsilon_k = (-1)^k$. Come dice la locuzione appena introdotta, sono serie a termini di segno variabile, in cui però la legge della variazione di segno ε_k non è arbitraria, come nel caso generale, ma segue una regola semplicissima, che consiste nell'alternarsi di segni + e $-$. Per queste serie esiste un semplice criterio di convergenza, ma prima di presentarlo ci sembra opportuno osservare che, malgrado siano di un tipo particolarissimo, si incontrano molto frequentemente nelle applicazioni (per esempio, come si vedrà, in connessione con le serie di Fourier).

Teorema 3.24 (Criterio di Leibniz) *Sia (a_k) una successione di numeri strettamente positivi, e supponiamo che*

1. $\lim\limits_{k\to\infty} a_k = 0$

2. $a_k \geq a_{k+1}$ *per ogni* $k \in \mathbf{N}$.

Allora la serie a segni alterni

$$\sum_{k=0}^{\infty}(-1)^k a_k$$

è semplicemente convergente. Inoltre, se la serie è convergente, detta S *la sua somma, per ogni* $n \in \mathbf{N}$ *risulta*

$$\left|S - \sum_{k=0}^{n}(-1)^k a_k\right| \leq a_{n+1}. \tag{3.3.6}$$

Dim. Consideriamo le due successioni estratte dalla successione (s_n) delle somme parziali della serie corrispondenti agli indici pari e dispari rispettivamente, e notiamo che sono monotone a causa della decrescenza della successione (a_k):

$$\begin{aligned}
s_{2p} &= \sum_{k=0}^{2p}(-1)^k a_k = \sum_{k=0}^{2p-2}(-1)^k a_k - a_{2p-1} + a_{2p} \\
&\leq \sum_{k=0}^{2p-2}(-1)^k a_k = s_{2p-2} \\
s_{2p+1} &= \sum_{k=0}^{2p+1}(-1)^k a_k = \sum_{k=0}^{2p-1}(-1)^k a_k + a_{2p} - a_{2p+1} \\
&\geq \sum_{k=0}^{2p-1}(-1)^k a_k = s_{2p-1}
\end{aligned}$$

Per il Teorema fondamentale sulle successioni monotone entrambe le estratte (s_{2p}) e (s_{2p+1}) ammettono limite, siano S_p ed S_d rispettivamente. Siccome l'unione degli indici delle due estratte è tutto l'insieme dei numeri naturali, se proviamo che i due limiti sono finiti e che $S_p = S_d$, risulterà provata la convergenza dell'intera successione (s_n) e quindi della serie. Iniziamo a vedere che i due limiti sono finiti. Siccome s_{2p} è decrescente ed s_{2p+1} è crescente, basta trovare un minorante di s_{2p} e un maggiorante di s_{2p+1}; risulta

$$\begin{aligned}
s_{2p} &= s_{2p-1} + a_{2p} \geq s_{2p-1} \geq s_1 \\
s_{2p+1} &= s_{2p} - a_{2p+1} \leq s_{2p} \leq s_0 = a_0
\end{aligned}$$

e quindi sia S_p che S_d sono finiti. D'altra parte, per il teorema sulla somma dei limiti e l'ipotesi 1:

$$0 = \lim_{n\to\infty} a_{2n+1} = \lim_{n\to\infty}(s_{2n+1} - s_{2n}) = \lim_{n\to\infty} s_{2n+1} - \lim_{n\to\infty} s_{2n} = S_d - S_p$$

da cui segue evidentemente che $S_p = S_d$ e che la serie converge. Detta allora S la sua somma, si ha

$$S = S_p = S_d = \inf_p s_{2p} = \sup_p s_{2p+1}$$

e quindi per ogni $p \in \mathbf{N}$

$$s_{2p+1} = s_{2p} - a_{2p+1} \leq S \leq s_{2p} \qquad \Longrightarrow \qquad s_{2p} - S \leq a_{2p+1}$$

e analogamente

$$s_{2p+1} \leq S \leq s_{2p+2} = s_{2p+1} + a_{2p+2} \qquad \Longrightarrow \qquad S - s_{2p+1} \leq a_{2p+2},$$

sicché in ogni caso si ha (3.3.6). $\boxed{QED}$

Osservazioni 3.25 1. Naturalmente, la scelta che i termini pari siano positivi e quelli dispari negativi è una delle due possibili, l'altra essendo ovviamente del tutto equivalente. Inoltre, osserviamo che come al solito le condizioni 1 e 2 del Criterio di Leibniz debbono essere verificate *da un certo indice in poi*, per poter applicare il criterio stesso, e non necessariamente *per tutti gli indici.*

2. L'errore più frequente nell'applicazione del criterio di Leibniz è l'omissione della verifica che la successione (a_k) sia decrescente. Questa verifica è invece essenziale, dal momento che, ovviamente, esistono successioni positive infinitesime *non decrescenti.* Per avere un esempio esplicito, consideriamo la successione (infinitesima ma non decrescente)
$$a_k = \begin{cases} \dfrac{1}{k} & \text{se } k \text{ è pari,} \\[2ex] \dfrac{1}{k^2} & \text{se } k \text{ è dispari.} \end{cases}$$
Per l'Osservazione 3.5.8, la serie a segni alterni $\sum_k (-1)^k a_k$ non può convergere, perché si può scrivere come differenza tra la serie $\sum_k \frac{1}{k}$, che diverge, e la serie $\sum_k \frac{1}{k^2}$, che converge.

3. Si possono aggiungere alcune interessanti considerazioni generali all'esempio precedente (vedi anche il Teorema 3.26 seguente): se la serie a segni alterni $\sum_k (-1)^k a_k$ converge semplicemente ma *non* converge assolutamente, allora né la serie $\sum_k a_{2k}$ dei termini positivi né la serie $\sum_k a_{2k+1}$ dei termini negativi possono convergere. Infatti, una sola di queste ultime non può convergere (l'argomento per provare questa affermazione è quello dell'esempio precedente), e se convergessero entrambe si avrebbe convergenza assoluta della serie di partenza.

4. Ora che abbiamo un criterio di convergenza semplice, possiamo vedere un esempio di serie semplicemente convergente che non è assolutamente convergente. Basta considerare la *serie armonica a segni alterni* $\sum_k (-1)^k \frac{1}{k}$, che converge semplicemente per il criterio di Leibniz, me non converge assolutamente, come abbiamo già visto.

Terminiamo il capitolo con delle considerazioni che dovrebbero far capire quanta distanza ci sia tra le serie e le somme finite. Una delle proprietà più "ovvie" delle somme finite è la proprietà commutativa. È naturale domandarsi se essa valga anche per le serie. Per formulare correttamente il problema bisogna introdurre il concetto di *permutazione dei termini di una serie.* Date la serie $\sum_k a_k$ ed una funzione bigettiva $\pi : \mathbf{N} \to \mathbf{N}$ (permutazione), si dice *serie ottenuta permutando i termini di* $\sum_k a_k$ *secondo* π *la serie* $\sum_k a_{\pi(k)}$. Notiamo che i valori assunti dalla successione $(a_{\pi(k)})$ sono *gli stessi di* (a_k), *e vengono assunti lo stesso numero di volte*, per cui se avessimo a che fare con una somma finita passare da $\sum_k a_k$ a $\sum_k a_{\pi(k)}$ si ridurrebbe a "cambiare l'ordine degli addendi", ed è ben noto che in tal caso "la somma non cambia". Per le serie infinite le cose vanno in modo completamente diverso, a meno che non si abbia convergenza assoluta.

Teorema 3.26 *Sia* (a_k) *una successione convergente a 0. Allora:*

(i) *se la serie* $\sum_{k=0}^{\infty} a_k$ *converge assolutamente e la sua somma è* S*, allora per ogni permutazione* π *la serie* $\sum_{k=0}^{\infty} a_{\pi(k)}$ *converge assolutamente ed ha per somma* S*.*

(ii) *se la serie* $\sum_{k=0}^{\infty} a_k$ *non converge assolutamente, allora nessuna serie permutata converge assolutamente, ed inoltre per ogni* $S \in \mathbf{R}$ *esiste una permutazione* π *tale che la serie permutata* $\sum_{k=0}^{\infty} a_{\pi(k)}$ *converga (semplicemente) ad* S*.*

Non presentiamo la dimostrazione di questo risultato, che per altro non richiede strumenti che vadano oltre la presente trattazione, ma ci limitiamo a sottolineare ancora la differenza tra le somme finite e le serie non assolutamante convergenti: queste, cambiando l'ordine degli addendi, possono dare *qualunque* somma!

CAPITOLO 4

CALCOLO DIFFERENZIALE

4.1 Limiti di funzioni

4.2 Funzioni continue

4.3 Derivate di una funzione

4.4 Proprietà delle funzioni derivabili

4.5 Grafici di funzioni

CAPITOLO 5

CALCOLO INTEGRALE

Con lo studio del Calcolo integrale ci prefiggiamo di risolvere due problemi apparentemente distinti, ma che in realtà si vedranno essere intimamente legati. Essi sono:

Problema 1 (antiderivazione) Sia $I \subset \mathbf{R}$ un intervallo; data $f : I \to \mathbf{R}$, dire se esiste una funzione G derivabile in I tale che $G' = f$.

Problema 2 (quadratura) Date $f, g : [a, b] \to \mathbf{R}$, con $f(x) \le g(x)$ per ogni $x \in [a, b]$, assegnare un valore numerico che esprima l'*area* dell'insieme

$$E = \{(x, y) \in \mathbf{R}^2 : \ a \le x \le b, \ f(x) \le y \le g(x)\}.$$

Per quanto riguarda il problema 2 si vuole definire l'area di una regione piana in modo che certi (naturali) requisiti siano soddisfatti: per esempio, si vuole che l'area di una regione A contenuta in una B sia più piccola dell'area di B, o che l'area dell'unione di due regioni disgiunte sia la somma delle aree, e così via. Non formalizziamo tutte le richieste, ma è importante avere in mente che una funzione che esprima l'area non può essere arbitraria.

Definizione 5.1 (Primitive) *Date $f, G : I \to \mathbf{R}$, si dice che G è una* primitiva *di f in I se G è derivabile in I e risulta $G'(x) = f(x)$ per ogni $x \in I$.*

Osservazioni 5.2 Le seguenti proprietà delle primitive sono immediate conseguenze della definizione.

1. Se G è una primitiva della funzione f in I, ogni funzione del tipo $G(x) + c$, con $c \in \mathbf{R}$, è ancora una primitiva di f in I. Infatti, $D(G + c) = G' = f$.

2. Se G_1 e G_2 sono primitive della stessa funzione in un intervallo I, allora la funzione differenza $G_1 - G_2$ è costante in I. Infatti, risulta $D(G_1 - G_2) = G_1' - G_2' = 0$ in I, e quindi $G_1 - G_2$ è costante in I per il Corollario ???.

3. Tenendo conto delle precedenti osservazioni, il problema del calcolo di *tutte* le primitive di una funzione f *in un intervallo* si riconduce al calcolo di *una sola* primitiva. Infatti trovatane una, sia G, le funzioni $G + c$, al variare della costante c in $\mathbf{R}$, sono *tutte e sole* le primitive di f in I.

5.1 Funzioni integrabili secondo Riemann

In questo paragrafo considereremo sempre funzioni definite in un intervallo chiuso e limitato che denoteremo $[a,b]$. Daremo una definizione di integrale di una funzione, ma precisiamo che questa non è l'unica definizione possibile per affrontare i problemi su esposti. Le idee presentate in questo capitolo sono state sviluppate dal matematico tedesco Bernhard Riemann (1826-1866), ed è per questo che l'integrale che definiremo viene detto *integrale di Riemann.*

Definizione 5.3 (suddivisione) *Si dice* suddivisione dell'intervallo $[a,b]$ *la scelta di un numero finito di punti* $x_0,\dots,x_n$ *tali che*

$$a = x_0 < x_1 < x_2 \cdots < x_n = b.$$

Per la generica suddivisione di $[a,b]$ *useremo la notazione* $P = \{a = x_0,\dots,x_n = b\}$.

Notiamo che fra le suddivisioni di un intervallo sussiste l'ovvia relazione di inclusione. Costruiremo l'integrale di una funzione con un procedimento di approssimazione basato sulle nozioni introdotte nella seguente definizione.

Definizione 5.4 (Somme integrali) *Sia data* $f : [a,b] \to \mathbf{R}$ *limitata, e fissiamo una suddivisione* $P = \{a = x_0,\dots,x_n = b\}$ *di* $[a,b]$. *Per* $k = 1,\dots,n$ *poniamo*

$$m_k = \inf\{f(x) : \ x_{k-1} \le x \le x_k\} \tag{5.1.1}$$

$$M_k = \sup\{f(x) : \ x_{k-1} \le x \le x_k\} \tag{5.1.2}$$

e definiamo la somma integrale inferiore di f relativa a P *ponendo*

$$s(f,P) = \sum_{k=1}^{n} m_k(x_k - x_{k-1}) \tag{5.1.3}$$

e la somma integrale superiore di f relativa a P *ponendo*

$$S(f,P) = \sum_{k=1}^{n} M_k(x_k - x_{k-1}). \tag{5.1.4}$$

Tutta la costruzione è basata sulle proprietà delle somme integrali esposte nella proposizione seguente.

Proposizione 5.5 *Sia* $f : [a,b] \to \mathbf{R}$ *limitata, e siano* P,Q *suddivisioni di* $[a,b]$, *con* $P \subset Q$. *Allora*

$$s(f,P) \le s(f,Q) \le S(f,Q) \le S(f,P).$$

Dim. Proviamo l'enunciato per Q ottenuta da P aggiungendo un punto, poiché nel caso generale basterà ripetere l'argomento per ogni punto appartenente a Q e non a P. Siano allora $P = \{a = x_0 < \cdots < x_n = b\}$, e $Q = P \cup \{\bar{x}\}$. Se, com'è ovvio, assumiamo che $\bar{x}$ non appartenga a P, esiste un indice $j \in \{1, \ldots, n\}$ tale che $x_{j-1} < \bar{x} < x_j$. Posto

$$\begin{aligned} m'_j &= \inf\{f(x) :\ x_{j-1} \leq x \leq \bar{x}\}, \quad m''_j = \inf\{f(x) :\ \bar{x} \leq x \leq x_j\}, \\ M'_j &= \sup\{f(x) :\ x_{j-1} \leq x \leq \bar{x}\}, \quad M''_j = \sup\{f(x) :\ \bar{x} \leq x \leq x_j\}, \end{aligned}$$

risulta, dall'osservazione 1.9.4, $m'_j \geq m_j$, $m''_j \geq m_j$, $M'_j \leq M_j$, $M''_j \leq M_j$, e quindi

$$\begin{aligned} s(f, Q) - s(f, P) &= m'_j(\bar{x} - x_{j-1}) + m''_j(x_j - \bar{x}) - m_j(x_j - x_{j-1}) \geq 0, \\ S(f, Q) - S(f, P) &= M'_j(\bar{x} - x_{j-1}) + M''_j(x_j - \bar{x}) - M_j(x_j - x_{j-1}) \leq 0. \end{aligned}$$

Poiché la diseguaglianza $s(f, Q) \leq S(f, Q)$ è ovvia dalla definizione per ogni suddivisione Q, la tesi è dimostrata. $\boxed{QED}$

Definizione 5.6 (Funzioni integrabili) *Data* $f : [a, b] \to \mathbf{R}$ *limitata, definiamo il suo* integrale inferiore in $[a, b]$ *ponendo*

$$\underline{\int_a^b} f(x)dx = \sup\{s(f, P) :\ P \text{ suddivisione di } [a, b]\}$$

e il suo integrale superiore in $[a, b]$ *ponendo*

$$\overline{\int_a^b} f(x)dx = \inf\{S(f, P) :\ P \text{ suddivisione di } [a, b]\}.$$

Diciamo che f è integrabile in $[a, b]$ *se*

$$\underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx.$$

In tal caso, il loro comune valore si denota con

$$\int_a^b f(x)dx$$

e si dice integrale definito di f in $[a, b]$.

Notiamo che nella notazione appena introdotta per gli integrali la variabile x è *muta*, può cioè essere sostituita con qualunque altro simbolo senza alterare il significato dell'espressione. In particolare, se f è integrabile in $[a, b]$ allora le scritture $\int_a^b f(x)dx$ e $\int_a^b f(t)dt$ indicano lo stesso numero.

Osservazione 5.7 Mostriamo con un esempio che non tutte le funzioni sono integrabili. Come si vedrà, l'esempio è piuttosto "artificiale", rispetto agli esempi considerati fin qui, che sono sempre stati costruiti usando le funzioni elementari, malgrado sia probabilmente il più semplice possibile. Questo fa capire che in generale le funzioni che incontreremo saranno tutte integrabili negli intervalli chiusi e limitati. D'altra parte, è impornante essere consapevoli che una definizione ha senso solo se qualche oggetto sfugge alla classe che si sta definendo; altrimenti, la definizione è quanto meno inutile. Vediamo l'esempio di funzione non integrabile, che in genere viene chiamata *funzione di Dirichlet*, dal nome di un matematico dell'ottocento. Sia $f : [0,1] \to \mathbf{R}$ definita da

$$f(x) = \begin{cases} 1 & \text{se } x \in [0,1] \cap \mathbf{Q} \\ 0 & \text{se } x \in [0,1] \setminus \mathbf{Q} \end{cases}$$

Poiché fra due qualunque numeri reali ci sono sempre un numero razionale ed un numero irrazionale, per ogni suddivisione $P = \{0 = x_0, \ldots, x_n = 1\}$ di $[0,1]$ e per ogni $k = 1, \ldots, n$ risulta (con la solita notazione) $m_k = 0$ e $M_k = 1$, e quindi $s(f,P) = 0$ e $S(f,P) = 1$. Segue

$$\underline{\int_0^1} f(x)dx = 0, \qquad \overline{\int_0^1} f(x)dx = 1$$

e quindi f non è integrabile.

L'integrale definito gode di alcune semplici proprietà, che valgono in modo ovvio per le somme integrali e seguono facilmente per l'integrale.

Proposizione 5.8 (Proprietà degli integrali) *Siano $f, g : [a,b] \to \mathbf{R}$ integrabili, e siano $\alpha, \beta \in \mathbf{R}$. Allora:*

1. **[Linearità]** *La funzione $\alpha f + \beta g$ è integrabile, e risulta*
$$\int_a^b (\alpha f + \beta g)dx = \alpha \int_a^b f dx + \beta \int_a^b g dx.$$

2. **[Additività rispetto all'intervallo]** *Se $c \in]a,b[$ allora*
$$\int_a^c f dx + \int_c^b f dx = \int_a^b f dx.$$

3. **[Confronto]** *Se $f(x) \le g(x)$ per ogni $x \in [a,b]$ allora*
$$\int_a^b f dx \le \int_a^b g dx.$$

Osservazione 5.9 È utile definire l'integrale anche quando il primo estremo d'integrazione è un numero reale maggiore del secondo estremo. Il modo più coerente di definire l'integrale è il seguente:

$$\int_b^a f dx = -\int_a^b f dx$$

per ogni $f : [a,b] \to \mathbf{R}$ integrabile. Con questa definizione, la proprietà di additività rispetto all'intervallo vale qualunque sia l'ordine dei punti a, b, c.

La verifica dell'integrabilità di una funzione, più che sulla definizione, fa spesso uso del seguente risultato.

Teorema 5.10 (Caratterizzazione delle funzioni integrabili) *Una $f : [a,b] \to \mathbf{R}$ limitata è integrabile in $[a,b]$ se e solo se per ogni $\varepsilon > 0$ esiste una suddivisione P di $[a,b]$ tale che $S(f,P) - s(f,P) < \varepsilon$.*

DIM. Sia f integrabile, sia $I \in \mathbf{R}$ il valore dell'integrale e sia ε fissato. Siccome I è sia il valore dell'integrale superiore che quello dell'integrale inferiore, dalla Proposizione 1.11 (con $\varepsilon/2$ al posto di ε) segue che esistono due suddivisioni P_1 e P_2 tali che

$$S(f,P_1) < I + \frac{\varepsilon}{2}, \qquad s(f,P_2) > I - \frac{\varepsilon}{2}.$$

Per $P = P_1 \cup P_2$ risulta allora, grazie alla Proposizione 5.5,

$$S(f,P) - s(f,P) \le S(f,P_1) - s(f,P_2) < \varepsilon$$

e la prima implicazione è dimostrata.

Viceversa, se per ogni $\varepsilon > 0$ esiste una suddivisione P di $[a,b]$ tale che $S(f,P) - s(f,P) < \varepsilon$, poiché la differenza tra integrale superiore ed integrale inferiore è minore della differenza $S(f,P) - s(f,P)$ per ogni suddivisione, risulta

$$\overline{\int_a^b} f - \underline{\int_a^b} f \le \varepsilon$$

per ogni $\varepsilon > 0$, e, per l'Osservazione 1.10, ciò è possibile solo se $\overline{\int_a^b} f = \underline{\int_a^b} f$, cioè se f è integrabile. $\boxed{QED}$

Usando il precedente risultato, si può dimostrare che ampie classi di funzioni sono integrabili. Vediamo alcuni risultati abbastanza generali.

Teorema 5.11 (Integrabilità delle funzioni continue) *Se $f : [a,b] \to \mathbf{R}$ è continua in $[a,b]$ allora è integrabile in $[a,b]$.*

DIM. Useremo la caratterizzazione dell'integrabilità data nel Teorema 5.10. Fissato $\varepsilon > 0$, poiché f è uniformemente continua in $[a,b]$ in virtù del Teorema di Cantor ??, esiste un $\delta > 0$ tale che $|x-y| < \delta$ implica $|f(x) - f(y)| < \varepsilon$. Di conseguenza, se fissiamo una suddivisione $P = \{x_0, \ldots, x_n\}$ di $[a,b]$ tale che $x_k - x_{k-1} < \delta$ per ogni $k = 1, \ldots, n$, con la solita notazione risulta $M_k - m_k < \varepsilon$, e pertanto

$$S(f,P) - s(f,P) = \sum_{k=1}^{n} (M_k - m_k)(x_k - x_{k-1}) < \varepsilon \sum_{k=1}^{n} (x_k - x_{k-1}) = \varepsilon(b-a).$$

$\boxed{QED}$

Definizione 5.12 (Funzioni continue a tratti) *Diciamo che $f : [a,b] \to \mathbf{R}$ è continua a tratti in $[a,b]$ se è continua per ogni x in $[a,b]$ eccetto un numero finito di punti in cui ha discontinuità di prima specie.*

In modo equivalente, si può dire che f è continua a tratti se è continua per ogni x in $[a,b]$ eccetto un numero finito di punti, siano $x_1, \ldots, x_n$, e le restrizioni di f agli intervalli $[a, x_1]$, $[x_1, x_2], \ldots, [x_n, b]$ sono continue.

Osservazione 5.13 (Integrabilità delle funzioni continue a tratti) Se f è continua a tratti in $[a,b]$ allora è integrabile in $[a,b]$. Infatti, con la notazione appena introdotta, f è integrabile negli intervalli $[a, x_1]$, $[x_1, x_2], \ldots, [x_n, b]$ e risulta

$$\int_a^b f = \int_a^{x_1} f + \int_{x_1}^{x_2} f + \cdots + \int_{x_{n-1}}^{x_n} f + \int_{x_n}^b f$$

Le funzioni monotone in un intervallo possono presentare un numero infinito di punti di discontinuità, ma sono ancora integrabili.

Teorema 5.14 (Integrabilità delle funzioni monotone) *Se $f : [a,b] \to \mathbf{R}$ è monotona in $[a,b]$ allora è integrabile in $[a,b]$.*

Dim. Useremo anche im questo caso la caratterizzazione dell'integrabilità data nel Teorema 5.10. Per fissare le idee, supponiamo f crescente. In tal caso, per ogni suddivisione $P = \{x_0, \ldots, x_n\}$ di $[a,b]$ risulta, con la solita notazione, $M_k = f(x_k)$ e $m_k = f(x_{k-1})$. Dato allora $\varepsilon > 0$, basta scegliere P tale che $x_k - x_{k-1} < \varepsilon$ per ogni $k = 1, \ldots, n$, così che si abbia

$$\begin{aligned} S(f,P) - s(f,P) &= \sum_{k=1}^n (M_k - m_k)(x_k - x_{k-1}) = \sum_{k=1}^n (f(x_k) - f(x_{k-1}))(x_k - x_{k-1}) \\ &< \varepsilon \sum_{k=1}^n (f(x_k) - f(x_{k-1})) = \varepsilon(f(b) - f(a)). \end{aligned}$$

QED

Il seguente risultato ci permetterà di estendere notevolmente la classe delle funzioni integrabili.

Proposizione 5.15 *Siano f integrabile in $[a,b]$ e $g : \mathbf{R} \to \mathbf{R}$ lipschitziana in ogni intervallo limitato. Allora $g \circ f$ è integrabile in $[a,b]$.*

Dim. Poiché f è integrabile in $[a,b]$, f è limitata, diciamo $|f(x)| \le M$ per ogni $x \in [a,b]$. Sia L la costante di Lipschitz di g nell'intervallo $[-M, M]$, sicché risulti

$$|g(y_1) - g(y_2)| \le L|y_1 - y_2| \qquad \forall y_1, y_2 \in [-M, M]. \tag{5.1.5}$$

Dato $\varepsilon > 0$, per ipotesi esiste una suddivisione $P = \{a = x_0, \dots, x_n = b\}$ dell'intervallo $[a, b]$ tale che $S(f, P) - s(f, P) < \varepsilon/L$. Posto, per $k = 1, \dots, n$

$$\begin{aligned}
m_k &= \inf\{f(x) : \ x_{k-1} \le x \le x_k\} \\
M_k &= \sup\{f(x) : \ x_{k-1} \le x \le x_k\} \\
m'_k &= \inf\{g(f(x)) : \ x_{k-1} \le x \le x_k\} \\
M'_k &= \sup\{g(f(x)) : \ x_{k-1} \le x \le x_k\}
\end{aligned}$$

per (5.1.5) risulta $M'_k - m'_k \le L(M_k - m_k)$. Quindi

$$\begin{aligned}
S(g \circ f, P) - s(g \circ f, P) &\le \sum_{k=1}^{n} (M'_k - m'_k)(x_k - x_{k-1}) \\
&\le L \sum_{k=1}^{n} (M_k - m_k)(x_k - x_{k-1}) = L(S(f, P) - s(f, P)) < \varepsilon
\end{aligned}$$

e per l'arbitrarietà di ε la tesi segue dal Teorema 5.10 di caratterizzazione delle funzioni integrabili. QED

Il precedente risultato è molto generale: vediamone alcune semplici conseguenze. È utile introdurre le seguenti notazioni. Poniamo

$$\begin{aligned}
g^+(y) &= \begin{cases} y & \text{se } \ y \ge 0 \\ 0 & \text{se } \ y < 0 \end{cases} \\
g^-(y) &= \begin{cases} 0 & \text{se } \ y \ge 0 \\ -y & \text{se } \ y < 0 \end{cases}
\end{aligned}$$

e, per $f : I \to \mathbf{R}$,

$$\begin{aligned}
f^+(x) &= \begin{cases} f(x) & \text{se } \ f(x) \ge 0 \\ 0 & \text{se } \ f(x) < 0 \end{cases} \\
f^-(x) &= \begin{cases} 0 & \text{se } \ f(x) \ge 0 \\ -f(x) & \text{se } \ f(x) < 0 \end{cases}
\end{aligned}$$

Le funzioni f^+ ed f^- si dicono rispettivamente *parte positiva* e *parte negativa* di f. Notiamo che le funzioni f^+ ed f^- sono entrambe positive e che $f = f^+ - f^-$, mentre $|f| = f^+ + f^-$. Dal momento che risulta $f^+ = g^+ \circ f$, $f^- = g^- \circ f$ e le funzioni g^+, g^- sono lipschitziane, si ha subito che se f è integrabile in $[a, b]$ allora anche $f^+, f^-, |f|$ sono integrabili. Inoltre, vale la diseguaglianza

$$\Big| \int_a^b f(x)dx \Big| \le \int_a^b |f(x)|dx. \tag{5.1.6}$$

Infatti, $-|f| \le f \le |f|$, per la proprietà di confronto, implica

$$- \int_a^b |f| \le \int_a^b f \le \int_a^b |f|,$$

e da (1.1.1) segue (5.1.6). Applicando ancora la Proposizione 5.15 si ottiene che se f è integrabile anche f^2 è integrabile ($f^2 = g \circ f$, con $g(y) = y^2$ lipschitziana sugli intervalli limitati) e, se f e g sono integrabili, lo è anche il prodotto fg; infatti,

$$fg = \frac{1}{4}\Big[(f+g)^2 - (f-g)^2\Big]$$

e, se f e g sono integrabili, sia $f+g$ che $f-g$ lo sono, così come i loro quadrati. Naturalmente, è possibile completare questi risultati, estendendoli al caso di potenze diverse da 2 e al prodotto di più di due funzioni.

Possiamo ora dare una soluzione (parziale) al Problema 2 posto all'inizio del capitolo.

Definizione 5.16 (Area di figure piane) *Siano date $f, g : [a,b] \to \mathbf{R}$ integrabili, con $f(x) \le g(x)$ per ogni $x \in [a,b]$. Si dice* area *dell'insieme*

$$E = \{(x,y) \in \mathbf{R}^2 : \ a \le x \le b, \ f(x) \le y \le g(x)\}$$

il numero $\int_a^b (g-f)dx$.

Osserviamo che gli insiemi del tipo descritto nella definizione precedente (cioè gli insiemi di punti compresi tra i grafici di due funzioni integrabili) possono essere usati in situazioni più generali: dato un insieme qualunque del piano, si può infatti tentare di scomporlo nell'unione disgiunta di un numero finito di insiemi del tipo detto, e la sua area sarà data dalla somma delle aree dei singoli sottoinsiemi, ciascuna calcolata come spiegato.

Abbiamo visto che l'integrale è, in un certo senso, un'estensione dell'operazione di somma. Come per un numero finito di numeri $a_1, \ldots, a_n$ si definisce la *media aritmetica* m ponendo

$$m = \frac{1}{n}\sum_{k=1}^{n} a_k,$$

si può definire un concetto di media integrale.

Definizione 5.17 *Sia $f : [a,b] \to \mathbf{R}$ integrabile; si definisce la* media integrale di f in *$[a,b]$ ponendo*

$$m(f) = \frac{1}{b-a}\int_a^b f(x)dx.$$

Teorema 5.18 (Teorema della media integrale) *Sia $f : [a,b] \to \mathbf{R}$ integrabile, e siano*

$$m = \min\{f(x) : \ a \le x \le b\}; \qquad M = \max\{f(x) : \ a \le x \le b\};$$

allora $m \le m(f) \le M$. Se inoltre f è continua in $[a,b]$ allora esiste $x_0 \in [a,b]$ tale che $f(x_0) = m(f)$.

DIM. Per provare la prima affermazione, basta integrare le diseguaglianze $m \leq f(x) \leq M$ in $[a,b]$, tenendo conto della proprietà di confronto. Si ottiene

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a),$$

da cui, dividendo per $b-a$, la tesi.

Per quanto riguarda la seconda affermazione, basta applicare il Teorema ?? dei valori intermedi al numero $m(f) \in [m,M]$. QED

Osserviamo che, se $f(x) \geq 0$, il numero $m(f)$ è l'altezza del rettangolo di base $[a,b]$ che ha la stessa area del *trapeziode di* f definito da $\{(x,y) \in \mathbf{R}^2 : \ a \leq x \leq b, \ 0 \leq y \leq f(x)\}$.

5.2 Teorema fondamentale del calcolo e integrali indefiniti

In questo paragrafo mostriamo come si risolve il Problema 1 del calcolo delle primitive di una funzione, e quali sono i legami tra i due problemi enunciati. Iniziamo osservando che, data una funzione f integrabile in $[a,b]$, per ogni $x \in [a,b]$ si può considerare l'integrale tra a ed x di f, ottenendo un risultato che (ovviamente!) dipende solo da x, e pertanto definsce una nuova funzione con dominio $[a,b]$. Vediamone una prima importante proprietà.

Proposizione 5.19 *Sia* $f : [a,b] \to \mathbf{R}$ *integrabile; definiamo la sua* funzione integrale F *ponendo*

$$F(x) = \int_a^x f(t)dt, \qquad x \in [a,b]. \tag{5.2.7}$$

La funzione F *è lipschitziana in* $[a,b]$.

DIM. Sia $M = \max\{f(x) : \ a \leq x \leq b\}$. Notiamo che per l'Osservazione 5.9 e per la proprietà (5.1.6) risulta

$$|F(x_1) - F(x_2)| = \Big| \int_{x_1}^{x_2} f(t)dt \Big| \leq \Big| \int_{x_1}^{x_2} |f(t)|dt \Big| \leq \Big| \int_{x_1}^{x_2} Mdt \Big| \leq M|x_1 - x_2|.$$

QED

La precedente proposizione mostra che la funzione integrale di f è più regolare della f stessa. Utilizzando F si possono costruire le primitive di f, sotto l'ulteriore ipotesi che f sia continua. È questo il contenuto del seguente importante risultato.

Teorema 5.20 (Teorema fondamentale del calcolo) *Sia* $f : [a,b] \to \mathbf{R}$ *continua in* $[a,b]$. *Allora la funzione*

$$F(x) = \int_a^x f(t)dt, \qquad x \in [a,b]$$

è derivabile in $[a,b]$ *e risulta* $F'(x) = f(x)$ *per ogni* $x \in [a,b]$

DIM. Fissiamo $x_0 \in [a,b]$ e proviamo che $F'(x_0) = f(x_0)$. Per l'arbitrarietà di x_0 questo prova la tesi. Esplicitando la definizione della derivata, dobbiamo mostrare che

$$\lim_{h\to 0} \frac{F(x_0+h) - F(x_0)}{h} = f(x_0),$$

cioè, esplicitando anche la definizione di limite, che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per $|h| < \delta$ risulta

$$\Big|\frac{F(x_0+h) - F(x_0)}{h} - f(x_0)\Big| < \varepsilon. \tag{5.2.8}$$

Sia quindi $\varepsilon > 0$ fissato. Per ipotesi, f è continua in $[a,b]$ e in particolare in x_0, sicché, in corrispondenza del numero ε fissato,

$$\exists \delta > 0 \text{ tale che } |x - x_0| < \delta \qquad \Longrightarrow \qquad |f(x) - f(x_0)| < \varepsilon. \tag{5.2.9}$$

Notiamo che risulta

$$f(x_0) = \frac{1}{h}\int_{x_0}^{x_0+h} f(x_0)dx \tag{5.2.10}$$

(è la media integrale della costante $f(x_0)$ nell'intervallo di estremi x_0 e $x_0 + h$) e che, per l'Osservazione 5.9, possiamo scrivere il rapporto incrementale di F nella forma:

$$\begin{aligned} \frac{F(x_0+h) - F(x_0)}{h} &= \frac{1}{h}\Big[\int_a^{x_0+h} f(x)dx - \int_a^{x_0} f(x)dx\Big] \\ &= \frac{1}{h}\int_{x_0}^{x_0+h} f(x)dx \end{aligned} \tag{5.2.11}$$

Da (5.2.11) e (5.2.10) deduciamo

$$\frac{F(x_0+h) - F(x_0)}{h} - f(x_0) = \frac{1}{h}\int_{x_0}^{x_0+h}\big(f(x) - f(x_0)\big)dx, \tag{5.2.12}$$

Se ora prendiamo $|h| < \delta$, con δ dato da (5.2.9), risulta anche $|x - x_0| < \delta$ per x compreso tra x_0 ed $x_0 + h$, sicché, usando (5.1.6), otteniamo

$$\Big|\frac{F(x_0+h) - F(x_0)}{h} - f(x_0)\Big| \le \frac{1}{h}\int_{x_0}^{x_0+h}\Big|f(x) - f(x_0\Big|dx < \frac{1}{h}\int_{x_0}^{x_0+h} \varepsilon\ dx = \varepsilon,$$

cioè (5.2.8). QED

Osservazione 5.21 È importante notare che nel teorema fondamentale del calcolo non è necessario che l'intervallo I in cui la funzione f è definita sia chiuso e limitato. Infatti, anche se il dominio di f è un intervallo aperto o illimitato (anche $\mathbf{R}$, che anzi è un caso frequentissimo), pur di fissare (arbitrariamente, ma una volta per tutte) un punto $a \in I$, tutti gli integrali che abbiamo scritto sono calcolati su intervalli chiusi e limitati, in accordo con la trattazione precedente.

Il teorema fondamentale del calcolo mostra come si può calcolare una primitiva di una assegnata funzione continua in un intervallo. In realtà, grazie alle Osservazioni 5.2, il procedimento fornisce *tutte* le primitive cercate. Il teorema mostra lo stretto legame che c'è tra i Problemi 1 e 2 enunciati all'inizio del capitolo: il procedimento usato per risolvere il Problema 2, che fornisce l'area delle regioni di piano descritte nel Paragrafo precedente, permette infatti di costruire, facendo variare l'estremo superiore d'integrazione, le primitive di una funzione *continua* data. Notiamo che, com'è evidente dalla dimostrazione, avremmo potuto definire una funzione integrale diversa, ponendo

$$F_c(x) = \int_c^x f(t)dt$$

con c punto *qualunque* di $[a, b]$, ed avremmo ottenuto ancora una primitiva di f (e quindi tutte le primitive). La costruzione appena vista dà ragione della seguente definizione.

Definizione 5.22 (Integrale indefinito) *Sia I un intervallo, e sia $f : I \to \mathbf{R}$; si dice* integrale indefinito di f *l'insieme di tutte le primitive di f. Si denota con*

$$\int f(x)dx.$$

Osservazioni 5.23

1. Non bisogna confondere l'integrale definito con quello indefinito, che sono evidentemente oggetti del tutto diversi: l'integrale di f tra a e b è un numero reale, mentre l'integrale indefinito di f è un insieme di funzioni. Il legame tra i due, che giustifica i nomi, è nella costruzione del secondo, che è basata sulla definizione del primo.

2. Abbiamo dato la definizione di integrale indefinito per una generica funzione f, ma saremo quasi sempre interessati all'integrale di funzioni continue sull'intervallo I. In questo caso possiamo scrivere

   $$\int f(x)dx = \Big\{\int_a^x f(t)dt + c : \ c \in \mathbf{R}\Big\}, \tag{5.2.13}$$

   dove a è un qualunque punto di I. Infatti, la funzione integrale di f è una primitiva di f per il Teorema fondamentale del calcolo, e tutte le altre primitive si ottengono sommando un'arbitraria costante, come spiegato nelle Osservazioni 5.2.

3. Se il dominio di f non è un intervallo, ma, per esempio, l'unione di più intervalli, allora la descrizione dell'integrale indefinito cambia. Per esempio, sapendo che $\log x$ è una primitiva della funzione $1/x$, definita per $x \in]-\infty, 0[\cup]0, +\infty[$, l'integrale indefinito di $1/x$ è dato da tutte le funzioni del tipo

   $$\begin{cases} \log x + c_1 & \text{se } x > 0 \\ \log x + c_2 & \text{se } x < 0 \end{cases}$$

al variare di tutte le possibili scelte di c_1 e c_2 in $\mathbf{R}$, che non c'è alcun motivo di scegliere uguali. In particolare, la descrizione completa dell'integrale indefinito non dipende da una sola costante arbitraria (come nel caso in cui il dominio sia un intervallo) ma da più di una, in questo caso due. La ragione è che nell'Osservazione 5.2.2 abbiamo usato il Corollario ???, che vale per funzioni derivabili *in un intervallo.*

4. Il Teorema fondamentale del calcolo 5.20 afferma che *ogni funzione continua ammette primitive.* Bisogna tener distinta ques'affermazione, che è molto generale, dal problema del *calcolo effettivo* delle primitive di una funzione data. Anzi, anche l'espressione "calcolo effettivo" è da definire con chiarezza. In generale, il problema sarà di determinare le primitive di una *funzione elementare*, che, com'è noto, è sostanzialmente il risultato di (un numero finito di) operazioni algebriche e di composizione sulle funzioni elencate nel Capitolo 1. Anche in questo caso, però, *non è detto che le primitive della funzione data siano ancora funzioni elementari in quest'accezione.* Possiamo dare esempi semplicissimi, come e^{x^2}, $\frac{e^x}{x}$, $\frac{\sin x}{x}$, o le funzioni irrazionali del tipo $x^m(a+bx^n)^p$, con $a, b \in \mathbf{R}$ non nulli, $m, n, p \in \mathbf{Q}$ tali che nessuno dei tre numeri
$$p, \quad \frac{m+1}{n}, \quad p+\frac{m+1}{n}$$
sia intero.

Il legame tra l'integrale indefinito e quello definito è ulteriormente illustrato dal seguente risultato, che è il principale strumento per calcolare integrali definiti senza ricorrere alla definizione.

Teorema 5.24 (Secondo teorema fondamentale del calcolo) *Se $f: I \to \mathbf{R}$ è continua in I e G è una qualunque primitiva di f, allora per ogni intervallo $[a,b] \subset I$ risulta*
$$\int_a^b f(x)dx = G(b) - G(a).$$
DIM. Posto al solito $F(x) = \int_a^x f(t)dt$, per l'Osservazione 5.2.2, esiste $c \in \mathbf{R}$ tale che $G(x) = F(x) + c$, e quindi
$$\int_a^b f(x)dx = F(b) = [G(b) - c] - [G(a) - c] = G(b) - G(a).$$

QED

5.3 Metodi d'integrazione

In questo paragrafo affronteremo il problema del calcolo effettivo degli integrali indefiniti e definiti, cioè della determinazione delle primitive di una funzione data e dell'applicazione di questo al calcolo degli integrali definiti. Dal momento che il teorema fondamentale del calcolo dice sostanzialmente che l'integrazione indefinita è il procedimento inverso della derivazione, è lecito aspettarsi che i metodi di calcolo degli integrali indefiniti siano fondati su un'opportuna elaborazione dei metodi di calcolo delle derivate. Così è, infatti, ed i due principali metodi d'integrazione, detti *per parti* e *per sostituzione*, sono, rispettivamente, i procedimenti inversi del calcolo delle derivate del prodotto e della composizione di funzioni.

Teorema 5.25 (Integrazione per parti) *Siano f e g due funzioni di classe C^1 nell'intervallo I. Allora*

$$\int fg'dx = fg - \int f'g.$$

DIM. Basta osservare che la funzione fg è una primitiva della funzione $fg' + f'g$. QED

Osservazioni 5.26

1. Il teorema precedente può sembrare inutile ai fini del calcolo effettivo di integrali, dal momento che al primo ed al secondo membro appaiono due integrali simili. In realtà, come si vede su esempi concreti, se usata in modo opportuno, la formula può fornire un integrale più semplice di quello di partenza. Per esempio, consideriamo l'integrale
$$\int xe^x dx.$$
Posto $f(x) = x$, $g(x) = e^x$, risulta evidentemente $f'(x) = 1$ e $g'(x) = e^x$, sicché si ha:
$$\int xe^x dx = xe^x - \int e^x dx = xe^x - e^x + c, \qquad c \in \mathbf{R}$$
ed è evidente che al primo passaggio si è ottenuto un integrale più semplice di quello dato.

2. Anche quando dall'integrazione per parti si ottiene un integrale molto simile a quello dato, non è detto che il calcolo sia stato inutile. Ad esempio, consideriamo l'integrale
$$\int \cos^2 x\, dx.$$
Posto $f(x) = g'(x) = \cos x$, risulta $f'(x) = -g(x) = -\sin x$, e si ha:
$$\begin{aligned}\int \cos^2 x\, dx &= -\sin x \cos x + \int \sin^2 x\, dx = \int (1 - \cos^2 x)\, dx \\ &= -\sin x \cos x + x - \int \cos^2 x\, dx\end{aligned}$$
da cui
$$\int \cos^2 x = \frac{1}{2}(x - \sin x \cos x).$$

3. La formula di integrazione per parti si applica ovviamente anche agli integrali definiti. Con la notazione del teorema, se $[a,b] \subset I$, allora
$$\int_a^b f(x)g'(x)dx = [f(b)g(b) - f(a)g(a)] - \int_a^b f'(x)g(x)dx.$$

Teorema 5.27 (Integrazione per sostituzione) *Siano I, J due intervalli, $f : I \to \mathbf{R}$ continua, $\varphi : J \to I$ di classe C^1. Risulta:*

$$\int f(x)\, dx = \int f(\varphi(t))\varphi'(t)\, dt. \tag{5.3.14}$$

Per quanto riguarda gli integrali definiti, risulta

$$\int_{\varphi(c)}^{\varphi(d)} f(x)\, dx = \int_c^d f(\varphi(t))\varphi'(t)\, dt \tag{5.3.15}$$

per ogni intervallo $[c,d] \subset J$.

Dim. Se G è una qualunque primitiva di f, allora $G \circ \varphi$ è una primitiva della funzione $(f \circ \varphi)\varphi'$. Infatti:
$$\frac{d}{dt}(G \circ \varphi) = (G' \circ \varphi)\varphi' = f \circ \varphi)\varphi'$$
e quindi la formula dice semplicemente che l'integrale indefinito di f è $\{G + c\}$ se e solo se l'integrale indefinito di $f \circ \varphi$ è $\{G \circ \varphi + c\}$. QED

Osservazioni 5.28

1. Il modo forse più semplice (almeno dal punto di vista mnemonico) di applicare il Teorema 5.27 è il seguente. Ricordando la notazione $\varphi' = \frac{dx}{dt}$ per la derivata della funzione $x = \varphi(t)$, si può ricavare (formalmente) $dx = \varphi'(t)dt$, e questa sostituzione, assieme alla $x = \varphi(t)$, fornisce l'enunciato corretto del teorema, malgrado non sia giustificata dalle conoscenze presentate fin qui.

2. Il Teorema 5.27 si può usare nei due versi. Per esempio, consideriamo l'integrale
$$\int \sqrt{1-x^2}dx.$$
Ponendo $x = \varphi(t)$ con $\varphi(t) = \sin t$ si ottiene, procedendo come appena indicato, $dx = \cos t\, dt$ e quindi
$$\int \sqrt{1-x^2}dx = \int \sqrt{1-\sin^2 t} \cos t\, dt = \int \cos^2 t\, dt$$
e l'ultimo integrale è stato calcolato nell'Osservazione 5.26.2. Invece, se si considera l'integrale
$$\int \tan t\, dt$$

conviene porre $x = \varphi(t) = \cos t$, sicché $dx = \varphi'(t)dt = \sin t\,dt$ e si ottiene

$$\int \tan t\,dt = \int \frac{\sin t}{\cos t}dt = \int \frac{1}{x}dx = \log|x| + c = \log|\cos t| + c.$$

3. Nel calcolo per sostituzione degli intergali definiti si può procedere in due modi. Dovendo calcolare l'integrale $\int_a^b f(x)dx$, si possono prima calcolare le primitive in termini della variabile x e sostituire alla fine i valori a e b, oppure calcolare le primitive in termini della variabile t e sostituire alla fine i valori c e d tali che $\varphi(c) = a$ e $\varphi(d) = b$. È essenziale però verificare l'applicabilità del teorema *nell'intervallo* $[a, b]$. Per esempio, consideriamo l'integrale

$$\int_0^\pi \frac{1}{1+\cos^2 t}dt.$$

È facile verificare che la sostituzione $x = \tan t$ fornisce

$$\int \frac{1}{1+\cos^2 t}dt = \int \frac{1}{2+\tan^2 t}\frac{1}{\cos^2 t}dt = \int \frac{1}{2+x^2}dx = \frac{1}{\sqrt{2}}\arctan(x/\sqrt{2}) + c.$$

Se ora proviamo a sostituire gli estremi, otteniamo che $t = 0$ dà $x = 0$ e $t = \pi$ dà ancora $x = 0$, e da questi calcoli dedurremmo che l'integrale è nullo, risultato assurdo dal momento che l'integrale di una funzione strettamente positiva è strettamente positivo. L'errore è dovuto al fatto che la funzione tan *non è definita nell'intervallo* di integrazione $[0, \pi]$ perché $\pi/2$ non è nel suo dominio. La sostituzione indicata si può ancora adoperare, ma il calcolo corretto è quest'altro (in cui si tiene conto che $\cos^2(t + \pi/2) = \cos^2 t$):

$$\begin{aligned}\int_0^\pi \frac{1}{1+\cos^2 t}dt &= \int_0^{\pi/2} \frac{1}{1+\cos^2 t}dt + \int_{\pi/2}^\pi \frac{1}{1+\cos^2 t}dt \\ &= 2\int_0^{\pi/2} \frac{1}{1+\cos^2 t}dt = 2\int_0^{+\infty} \frac{1}{2+x^2}dx = \sqrt{2}\pi/4,\end{aligned}$$

dove l'ultimo passaggio è giustificato nell'esempio 5.42 del paragrafo seguente.

4. A volte è dato da calcolare un integrale definito nella forma

$$\int_a^b f(x)dx$$

e si vuol cercare una funzione φ che ne semplifichi il calcolo. Gli estremi c e d dell'integrale nella nuova variabile t tale che $x = \varphi(t)$ saranno due punti del dominio J di φ tali che $a = \varphi(c)$ e $b = \varphi(d)$ e in generale non sono univocamente determinati, dal momento che nell'enunciato del Teorema 5.27 non abbiamo supposto che φ sia invertibile. Se però φ è invertibile, possiamo scrivere (5.3.15) nella forma

$$\int_a^b f(x)\,dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t))\varphi'(t)\,dt. \tag{5.3.16}$$

A tal proposito, osserviamo che se la funzione φ è decrescente allora l'ordine degli estremi di integrazione viene scambiato, cioè $a < b$ allora $\varphi^{-1}(a) > \varphi^{-1}(b)$. Se si vuole conservare l'ordine crescente negli estremi d'integrazione, detto $[c, d]$ l'intervallo di estremi $\varphi^{-1}(a)$ e $\varphi^{-1}(b)$, la formula (5.3.16) va scritta

$$\int_a^b f(x)\,dx = \int_c^d f(\varphi(t))|\varphi'(t)|\,dt.$$

5.4 Integrali impropri

Finora abbiamo trattato solo integrali di funzioni *limitate* su intervalli *limitati*. Vediamo ora come sia possibile generalizzare la teoria per considerare casi più generali, che fra l'altro sono connessi con la teoria delle serie numeriche studiate nel Capitolo 2. Distinguiamo due casi, quello di funzioni non limitate su intervalli limitati e quello di intervalli non limitati, che si chiamano integrali generalizzati (o impropri) di prima e di seconda specie, rispettivamente. Naturalmente, tutta la trattazione si basa sui procedimenti già visti. Come nel caso delle serie numeriche, esiste una nozione di *convergenza assoluta* accantoa quella di convergenza semplice.

Definizione 5.29 (Integrali generalizzati di 1^a specie) *Sia* $f : [a, b[\to \mathbf{R}$ *integrabile in* $[a, c]$ *per ogni* $c < b$*; diciamo che* f *è* integrabile in senso generalizzato in $[a, b[$ *se esiste finito il limite*

$$\lim_{c\to b^-} \int_a^c f(x)dx.$$

Diciamo che f *è* assolutamente integrabile in senso generalizzato in $[a, b[$ *se esiste finito il limite*

$$\lim_{c\to b^-} \int_a^c |f(x)|dx.$$

Se $f :]a, b] \to \mathbf{R}$ *è integrabile in* $[c, b]$ *per ogni* $c \in]a, b]$*, allora i limiti considerati sopra vanno sostituiti con*

$$\lim_{c\to a^+} \int_c^b f(x)dx, \qquad \lim_{c\to a^+} \int_c^b |f(x)|dx,$$

Sulle relazioni tra integrabilità assoluta e non assoluta torneremo nell'Osservazione 5.36 e nell'Esempio 5.44.

La proprietà che la funzione f sia (assolutamente) integrabile in senso generalizzato in $[a, b[$ dipende da quanto velocemente $f(x)$ diventa grande per x che tende a b da destra. Ritroviamo questa considerazione qualitativa nel fondamentale esempio che segue.

Esempio 5.30 Studiamo l'integrabilità in senso generalizzato nell'intervallo $]0, 1]$ della funzione $f(x) = x^{-\alpha}$ al variare del parametro $\alpha \in \mathbf{R}$. Ovviamente, per $\alpha \leq 0$ la funzione è continua nell'intervallo chiuso $[0, 1]$ e quindi è integrabile, grazie al Teorema 5.11. Il problema dell'integrabilità in senso generalizzato si pone per $\alpha > 0$. Usando la definizione otteniamo:

$$\begin{aligned}\int_0^1 \frac{1}{x^\alpha}dx &= \lim_{c\to 0^+}\int_c^1 \frac{1}{x^\alpha}dx = \begin{cases} \displaystyle\lim_{c\to 0^+} \frac{1-c^{-\alpha+1}}{-\alpha+1} & \alpha \neq 1 \\ \displaystyle\lim_{c\to 0^+} -\log c & \alpha = 1 \end{cases} \\ &= \begin{cases} \dfrac{1}{-\alpha+1} & \alpha < 1 \\ +\infty & \alpha \geq 1 \end{cases}\end{aligned}$$

e quindi x^α è integrabile in senso generalizzato in $]0,1]$ se e solo se $\alpha < 1$.

Come abbiamo accennato prima, una funzione che tenda all'infinito troppo velocemente in un estremo dell'intervallo di definizione non sarà integrabile in senso generalizzato. Infatti, $x^{-\alpha}$ tende all'infinito per $x \to 0$ tanto più velocemente quanto più il parametro α è grande. La soglia per l'integrabilità, come spiegato nell'esempio, è il valore $\alpha = 1$: se α è più piccolo la funzione x^α è integrabile, se α è più grande non lo è. Questo punto di vista è utile per comprendere il seguente teorema di confronto. Trattiamo il caso di una funzione che può essere illimitata in prossimità dell'estremo destro dell'intervallo di definizione. Nel caso in cui il problema dell'integrabilità si ponga nell'estremo sinistro, le modifiche da apportare sono ovvie.

Teorema 5.31 (Criterio di confronto per gli integrali impropri di 1^a specie) *Siano $f, g : [a,b[\to \mathbf{R}$ integrabili in $[a,c]$ per ogni $c < b$, e supponiamo $|f(x)| \leq g(x)$ per ogni $x \in [a,b[$. Se g è integrabile in senso generalizzato in $[a,b[$ allora f è assolutamente integrabile in senso generalizzato in $[a,b[$.*

Osservazione 5.32 Dal teorema precedente segue subito che se $f \geq g \geq 0$ e g *non è* integrabile in senso generalizzato allora neanche f lo è.

Combinando la proposizione e l'osservazione precedenti con l'Esempio 5.30 si ottiene il seguente corollario.

Corollario 5.33 (Criterio d'integrabilità) *Sia $f : [a,b[\to \mathbf{R}$ integrabile in $[a,c]$ per ogni $c \in [a,b[$. Allora:*

(i) se esistono $C > 0$ ed $\alpha < 1$ tali che $|f(x)| \leq C(b-x)^{-\alpha}$ per ogni $x \in [a,b[$ allora f è assolutamente integrabile in senso generalizzato;

(ii) se esistono $C > 0$ ed $\alpha \geq 1$ tali che $|f(x)| \geq C(b-x)^{-\alpha}$ per ogni $x \in [a,b[$ allora f non è assolutamente integrabile in senso generalizzato.

Osservazione 5.34 Quanto visto finora per funzioni definite in un intervallo semiaperto a destra $[a,b[$ si può facilmente riformulare per funzioni definite in un intervallo semiaperto a sinistra del tipo $]a,b]$ o per funzioni definite in un intervallo privato di un punto interno.

Se $f:]a,b] \to \mathbf{R}$ è integrabile in $[c,b]$ per ogni $c \in]a,b]$, le condizioni nel Corollario 5.33 divengono

$$|f(x)| \leq C(x-a)^{-\alpha}, \qquad |f(x)| \geq C(x-a)^{-\alpha};$$

la prima, con $\alpha < 1$ garantisce l'integrabilità, e la seconda, con $\alpha \geq 1$, garantisce la non integrabilità.

Consideriamo ora $c \in]a,b[$, e sia $f: [a,b] \setminus \{c\} \to \mathbf{R}$ integrabile in $[a, c-\delta]$ e in $[c+\delta, b]$ per ogni $\delta > 0$. Allora, f è integrabile in senso generalizzato in $[a,b]$ se è integrabile in senso generalizzato sia in $[a,c]$ che in $[c,b]$, cioè se esistono finiti *entrambi* i seguenti limiti:

$$\lim_{\delta \to 0^+} \int_a^{c-\delta} f(x)dx, \qquad \lim_{\delta \to 0^+} \int_{c+\delta}^{b} f(x)dx.$$

Una definizione analoga vale ovviamente per l'assoluta integrabilità. In altri termini, nel caso di uno (o più) punti singolari *interni* all'intervallo di integrazione, si deve studiare un punto per volta, e ciascuno separatamente da destra e da sinistra.

Passiamo ora a considerare l'integrabilità di funzioni su semirette. Anche in questo caso, studiamo in dettaglio il caso della semiretta $[a,+\infty[$, ed esponiamo poi le modifiche da fare nel caso della semiretta $]-\infty, b]$).

Definizione 5.35 (Integrali generalizzati di 2^a specie) *Sia* $f: [a,+\infty[\to \mathbf{R}$ *integrabile in* $[a,b]$ *per ogni* $b > a$; *diciamo che* f *è* integrabile in senso generalizzato in $[a,+\infty[$ *se esiste finito il limite*

$$\lim_{b \to +\infty} \int_a^b f(x)dx.$$

Diciamo che f *è* assolutamente integrabile in senso generalizzato in $[a,+\infty[$ *se esiste finito il limite*

$$\lim_{b \to +\infty} \int_a^b |f(x)|dx.$$

Come nel caso delle serie numeriche (vedi Proposizione 3.9), la convergenza assoluta implica quella semplice.

Osservazione 5.36 Si può dimostrare che, sia nel caso degli integrali generalizzati di prima che di seconda specie, se f è assolutamente integrabile in senso generalizzato in $[a,b[$ (o $[a,+\infty[$) allora è anche integrabile in senso generalizzato (senza valore assoluto) nello stesso intervallo. Nell'esempio 5.44 vedremo che il viceversa non è vero. Questo tipo di fenomeno è analogo a quanto accade per le serie numeriche, vedi Proposizione 3.9 ed Osservazione 3.25.4.

Esempio 5.37 Studiamo l'integrabilità in senso generalizzato nell'intervallo $[1,+\infty[$ della funzione $f(x)=x^{-\alpha}$ al variare del parametro $\alpha\in\mathbf{R}$. Usando la definizione otteniamo:

$$\begin{aligned}\int_1^{\infty}\frac{1}{x^\alpha}dx &= \lim_{c\to+\infty}\int_1^c\frac{1}{x^\alpha}dx=\begin{cases}\displaystyle\lim_{c\to+\infty}\frac{c^{-\alpha+1}-1}{-\alpha+1} & \alpha\neq 1\\ \displaystyle\lim_{c\to0^+}\log c & \alpha=1\end{cases}\\ &=\begin{cases}\dfrac{1}{\alpha-1} & \alpha>1\\ +\infty & \alpha\leq 1\end{cases}\end{aligned}$$

e quindi $x^{-\alpha}$ è integrabile in senso generalizzato in $[1,+\infty[$ se e solo se $\alpha>1$.

In questo caso, una funzione che tenda a zero troppo lentamente per $x\to+\infty$ non sarà integrabile in senso generalizzato. Infatti, $x^{-\alpha}$ tende a zero per $x\to+\infty$ tanto più velocemente quanto più il parametro α è grande. La soglia per l'integrabilità, anche in questo caso, è il valore $\alpha=1$: stavolta, se α è più grande di 1 la funzione $x^{-\alpha}$ è integrabile, se α è più piccolo non lo è. Vediamo ora un'altra interessante classe di esempi.

Esempio 5.38 Studiamo l'integrabilità in senso generalizzato nell'intervallo $[2,+\infty[$ della funzione $f(t)=\frac{1}{t\log^\alpha t}$ al variare del parametro $\alpha>0$. In base alle proprietà dei logaritmi, non è possibile applicare teoremi di confronto con le funzioni studiate nell'esempio 5.37. Possiamo però usare l'esempio 5.37, dopo aver calcolato l'integrale con la sostituzione $x=\log t$, e otteniamo:

$$\int_2^{\infty}\frac{1}{t\log^\alpha t}dt=\int_{\log 2}^{\infty}\frac{1}{x^\alpha}dx$$

e quindi $\frac{1}{t\log^\alpha t}$ è integrabile in senso generalizzato in $[2,+\infty[$ se e solo se $\alpha>1$.

Anche per gli intervalli illimitati vale un teorema di confronto, analogo al Teorema 5.31.

Teorema 5.39 (Criterio di confronto per gli integrali impropri di 2^a specie) *Siano $f,g:[a,+\infty[\to\mathbf{R}$ integrabili in $[a,b]$ per ogni $b>a$, e supponiamo $|f(x)|\leq g(x)$ per ogni $x\in[a,+\infty[$. Se g è integrabile in senso generalizzato in $[a,+\infty[$ allora f è assolutamente integrabile in senso generalizzato in $[a,+\infty[$.*

Osservazione 5.40 Dal teorema precedente segue subito che se $f\geq g\geq 0$ e g *non* è integrabile in senso generalizzato in $[a,+\infty[$ allora neanche f lo è.

Come prima, combinando la proposizione e l'osservazione precedenti con l'Esempio 5.37 si ottiene il seguente corollario, analogo al Corollario 3.18 valido per le serie.

Corollario 5.41 (Criterio d'integrabilità) *Sia $f:[a,+\infty[\to\mathbf{R}$ integrabile in $[a,b]$ per ogni $b>a$. Allora:*

(i) *se esistono $C > 0$ ed $\alpha > 1$ tali che $|f(x)| \leq Cx^{-\alpha}$ per ogni $x \in [a, +\infty[$ allora f è assolutamente integrabile in senso generalizzato in $[a, +\infty[$;*

(ii) *se esistono $C > 0$ ed $\alpha \leq 1$ tali che $|f(x)| \geq Cx^{-\alpha}$ per ogni $x \in [a, +\infty[$ allora f non è assolutamente integrabile in senso generalizzato in $[a, +\infty[$.*

Il seguente esempio è interessante anche in relazione all'Osservazione 5.28.4.

Esempio 5.42 Consideriamo l'integrale

$$\int_0^{+\infty} \frac{1}{m^2 + x^2} dx,$$

con $m > 0$. Il Corollario 5.41 ci dice subito che l'integrale è convergente, ma è ancora più semplice procedere al calcolo diretto:

$$\int_0^{+\infty} \frac{1}{m^2 + x^2} dx = \frac{1}{m^2} \int_0^{+\infty} \frac{1}{1 + (x/m)^2} dx = \frac{1}{m} \lim_{x \to +\infty} \arctan(x/m) = \frac{\pi}{2m}.$$

Osservazioni 5.43 Consideriamo ora il caso in cui l'intervallo d'interesse sia una semiretta del tipo $]-\infty, b]$ o tutta la retta $]-\infty, +\infty[$. Come nel caso degli integrali di prima specie, sarà sufficiente indicare in breve le modifiche da fare. Distinguiamo i due casi.

1. Sia $f :]-\infty, b] \to \mathbf{R}$ integrabile in $[a, b]$ per ogni $a < b$. I limiti considerati nella Definizione 5.35 vanno sostituiti con

$$\lim_{a \to -\infty} \int_a^b f(x) dx, \qquad \lim_{a \to -\infty} \int_a^b |f(x)| dx,$$

e le condizioni nel Corollario 5.41 con

$$|f(x)| \leq C|x|^{-\alpha}, \qquad |f(x)| \geq C|x|^{-\alpha};$$

la prima, con $\alpha > 1$ garantisce l'integrabilità, e la seconda, con $\alpha \leq 1$, garantisce la non integrabilità in $]-\infty, b]$.

2. Sia ora $f : \mathbf{R} \to \mathbf{R}$ integrabile in $[a, b]$ per ogni $a < b$. Allora, f è integrabile in senso generalizzato in $\mathbf{R}$ se è integrabile in senso generalizzato sia in $[0, +\infty]$ che in $-\infty, 0]$, cioè se esistono finiti *entrambi* i seguenti limiti:

$$\lim_{b \to +\infty} \int_0^b f(x) dx, \qquad \lim_{a \to -\infty} \int_a^0 f(x) dx.$$

Una definizione analoga vale ovviamente per l'assoluta integrabilità. Naturalmente, l'estremo finito d'integrazione nelle formule precedenti è del tutto arbitrario, e si è scelto 0 solo per non introdurre altri parametri. Qualunque numero reale andrebbe bene.

Esempio 5.44 Consideriamo la funzione $\frac{\sin x}{x}$, per $x \in \mathbf{R}$ (com'è noto, la funzione vale 1 per $x = 0$, trattandosi di una discontinuità eliminabile). Allora l'integrale improprio

$$\int_{-\infty}^{+\infty} \frac{\sin x}{x} dx \tag{5.4.17}$$

è convergente (e vale π), mentre l'integrale improprio

$$\int_{-\infty}^{+\infty} \left|\frac{\sin x}{x}\right| dx \tag{5.4.18}$$

non è convergente. Questo prova che esistono funzioni integrabili in senso generalizzato che non sono *assolutamente* integrabili in senso generalizzato. È abbastanza facile vedere la divergenza dell'integrale in (5.4.17), mentre non è altrettanto facile provare la convergenza dell'integrale in (5.4.17). La difficoltà risiede nel fatto che le primitive della funzione $\frac{\sin x}{x}$, che ovviamente esistono per il Teorema fondamentale del calcolo, non sono esprimibili in termini delle funzioni elementari (come già osservato nell'Osservazione 5.23.4), per cui il calcolo esplicito, come negli esempi precedenti, è impossibile. Possiamo mettere in relazione quest'esempio con l'esempio 3.25.4 relativo alle serie a segni alterni, notando che, grazie al criterio di Leibniz, la trattazione relativa alle serie è molto più semplice.

Come accennato all'inizio, ci sono molte somiglianze tra la teoria degli integrali generalizzati e quella delle serie, tanto che sotto opportune condizioni lo studio di un integrale generalizzato e di una serie sono equivalenti, come mostra il seeguente criterio. Esso è particolarmente efficace perché, mentre è in generale molto difficile trovare una formula esplicita per le ridotte di una serie, è spesso possibile determinare le primitive di una funzione usando il teorema fondamentale del Calcolo.

Teorema 5.45 *Sia $f : [0, \infty[\to \mathbf{R}$ una funzione positiva e decrescente. Allora la serie*

$$\sum_{k=0}^{\infty} f(k)$$

converge se e solo se l'integrale improprio

$$\int_0^{\infty} f(x)\, dx$$

converge.

Dim. Siccome f è decrescente, risulta $f(k+1) \le f(x) \le f(k)$ per ogni $x \in [k, k+1]$ ed inoltre f è integrabile in ogni intervallo $[0, x]$, $x > 0$. Poiché f è positiva, la funzione

$$x \mapsto \int_0^x f(t)\, dt$$

è monotona crescente e quindi esiste (finito o no) il

$$\lim_{x\to\infty}\int_0^x f(t)\,dt, \tag{5.4.19}$$

che possiamo calcolare per valori interi di x considerando la successione $(\int_0^n f(t)dt)$. Se per ogni $n \in \mathbf{N}$ consideriamo la suddivisione $P_n = \{0, 1, \ldots, n-1, n\}$ dell'intervallo $[0, n]$ e le relative somme integrali inferiore e superiore, detta (s_n) la successione delle ridotte della serie $\sum_k f(k)$, risulta

$$s(f, P_n) = \sum_{k=1}^{n} f(k) = s_n - f(0) \le \int_0^n f(x)dx \le S(f, P_n) = \sum_{k=0}^{n-1} f(k) = s_{n-1},$$

sicché il limite in (5.4.19) è finito se e solo se la successione (s_n) converge. $\boxed{QED}$

Vediamo come il precedente criterio si possa applicare in due casi particolari importanti, già visti in altro modo.

Esempi 5.46 In quest'esempio discutiamo due famiglie di serie che sono utili anche come riferimento per i teoremi di confronto.

1. Sappiamo dall'Esempio 3.17.2 che la *serie armonica generalizzata*

$$\sum_{k=1}^{\infty}\frac{1}{k^\alpha}, \qquad\qquad \alpha \in \mathbf{R}$$

converge se e solo se $\alpha > 1$. Per ogni $\alpha > 0$ si può applicare il criterio di confronto con l'integrale improprio con $f(x) = \frac{1}{x^\alpha}$ e concludere che la serie data converge se e solo se $\alpha > 1$.

2. Sappiamo dall'Esempio 3.17.2 che la serie

$$\sum_{k=2}^{\infty}\frac{1}{k\log^\alpha k}, \qquad\qquad \alpha \in \mathbf{R}$$

converge se e solo se $\alpha > 1$. Per ogni $\alpha > 0$ si può applicare il criterio di confronto con l'integrale improprio con $f(x) = \frac{1}{x\log^\alpha x}$ e concludere, in base all'esempio 5.38, che la serie data converge se e solo se $\alpha > 1$.