

UNIVERSITÀ DEGLI STUDI DI FERRARA
FACOLTÀ DI INGEGNERIA

Appunti del corso di
Analisi Matematica II
c.d.l. Ingegneria Civile e Ambientale ¹

Michele Miranda

a.a. 2008-2009

¹Versione aggiornata al 3 marzo 2009

Nel presente fascicolo sono raccolti gli appunti relativi ad alcuni capitoli del corso di Analisi Matematica 2 tenuto presso la facoltà di Ingegneria dell'Università di Ferrara, corso di laurea in Ingegneria Civile ed Ambientale.

Il materiale contenuto in queste note vuole essere semplicemente una guida per gli argomenti trattati durante il corso; è inevitabilmente incompleto, così come è inevitabile che siano presenti errori ed inesattezze. Non si risponde tuttavia degli errori che possono essere contenuti in questo fascicolo, in quanto è cura del lettore rilevare e segnalare eventuali imprecisioni.

Riteniamo imprescindibile, pur con la riduzione dei contenuti dei corsi imposta dal nuovo ordinamento degli studi, conservare intatti l'impianto concettuale e l'impostazione metodologica dell'Analisi, e riteniamo che questo obiettivo sia conseguibile solo dando enunciati sintetici e precisi. Per semplificare un enunciato si può rinunciare 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.

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. Si consiglia pertanto sempre di studiare sui testi di Analisi Matematica esistenti in letteratura, sicuramente più affidabili e corretti; fortunatamente le biblioteche dei nostri Atenei traboccano di ottimi testi.

Il corso è strutturato come segue; i primi capitoli saranno un complemento del corso di Analisi Matematica I. Si tratteranno alcuni argomenti che, causa il poco tempo a disposizione, non sono stati affrontati nel corso di Analisi Matematica I, quali ad esempio le formule di Mac Laurin–Taylor e lo studio dei numeri complessi.

In un successivo capitolo si studiano le equazioni differenziali; si affronteranno equazioni differenziali di ordine al più due, con solo alcuni cenni alla teoria delle equazioni differenziali lineari a coefficienti costanti di ordine superiore.

Nei restanti capitoli si affronta quello che è più propriamente il programma di Analisi Matematica II, e cioè lo studio delle funzioni di più variabili. Si estenderanno quindi alle più variabili i concetti studiati nel primo corso di Analisi, quali la continuità, la derivabilità e l'integrabilità delle funzioni.

Si consiglia infine di prestare attenzione alla data di aggiornamento della presente dispensa, in quanto in continua evoluzione e correzione. In particolare, per la sezione riguardante le Domande, si dovrà prendere come riferimento, per ogni sezione del presente volume, la versione che verrà messa on-line ad inizio Marzo 2009.

Michele Miranda, Ferrara

Indice

1	Approssimazione polinomiale e formula di Taylor	5
1.1	Infinitesimi e il simbolo o	7
1.2	Formula di Mac Laurin-Taylor	8
1.3	Algebra degli o e formule di Taylor di funzioni composte	12
1.4	Applicazioni della formula di Taylor	14
2	I Numeri Complessi	17
2.1	Definizione e prime proprietà	18
2.2	Coniugato e modulo di un numero complesso	20
2.3	Forma polare ed esponenziale	21
2.4	Polinomi e radici n -esime	23
3	Equazioni differenziali	27
3.1	Equazioni del primo ordine	28
3.1.1	Equazioni a variabili separabili	32
3.1.2	Equazioni lineari del primo ordine	33
3.1.3	Equazioni di Bernoulli	35
3.2	Equazioni lineari del secondo ordine	35
3.2.1	Equazioni a coefficienti costanti	38
3.2.2	Oscillatore forzato	41
3.3	Equazioni lineari di ordine superiore	44
4	\mathbb{R}^N, topologia, limiti e funzioni continue	47
4.1	Topologia di \mathbb{R}^N	47
4.2	Successioni in \mathbb{R}^N	50
4.3	Limiti e funzioni continue	51
5	Curve ed integrali curvilinei	55
5.1	Curve e curve regolari	55
5.1.1	Curve nel piano	58
5.2	Lunghezza, ascissa curvilinea e integrali curvilinei	58
5.2.1	Integrali curvilinei	62
6	Dalla continuità alla differenziabilità	63
6.1	Sezioni e insiemi connessi	63
6.2	Derivabilità e differenziabilità per funzioni scalari	65
6.3	Differenziabilità per funzioni vettoriali	70

6.3.1	Cambiamenti di coordinate: polari, cilindriche e sferiche	71
7	Integrali multipli	73
7.1	Integrale di Riemann	73
7.2	Insiemi normali e calcolo degli integrali multipli	75
7.2.1	Integrali doppi	76
7.2.2	Integrali tripli	76
7.3	Cambiamenti di coordinate negli integrali multipli	78
7.4	Applicazioni	79
7.4.1	Solidi di rotazione	80
7.4.2	Baricentri e momenti d'inerzia	80
8	Estremi e punti stazionari	83
8.1	Massimi e minimi	83
8.2	Punti stazionari	85
8.3	Punti stazionari vincolati	86
8.3.1	Parametrizzazione del vincolo	86
8.3.2	Moltiplicatori di Lagrange	86
8.4	Calcolo di massimi e minimi assoluti su compatti	87
8.5	Classificazione dei punti stazionari	88
8.5.1	Derivate seconde e matrice Hessiana	88
8.5.2	Forme quadratiche	89
8.5.3	Classificazione dei punti stazionari	92
9	Domande	95
9.1	Sviluppi di Taylor: domande	95
9.2	Numeri complessi: domande	95
9.3	Equazioni differenziali: domande	96
9.4	\mathbb{R}^N , topologia, limiti e funzioni continue: domande	97
9.5	Curve ed integrali di linea: domande	97
9.6	Calcolo Differenziale in più variabili: domande	98
9.7	Integrali multipli: domande	99
9.8	Massimi e minimi: domande	99

Capitolo 1

Approssimazione polinomiale e formula di Taylor

Uno strumento molto importante in matematica è quello della approssimazione lineare e polinomiale; iniziamo con la prima, detta anche approssimazione del primo ordine.

Data una funzione $f : I \rightarrow \mathbb{R}$ con I intervallo di \mathbb{R} e $x_0 \in I$, dire che f è derivabile, o differenziabile, in x_0 significa richiedere l'esistenza di un numero reale $m \in \mathbb{R}$ tale che

$$(1.1) \quad \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - m(x - x_0)}{x - x_0} = 0,$$

e si pone $m = f'(x_0)$, la derivata o differenziale di f in x_0 . La precedente espressione ci dice che la retta

$$(1.2) \quad f(x) = f(x_0) + m(x - x_0),$$

rappresentante la tangente al grafico di f nel punto $(x_0, f(x_0))$ e ha la proprietà che la differenza

$$(1.3) \quad f(x) - f(x_0) - f'(x_0)(x - x_0)$$

è infinitesima, cioè tende a zero, per x che tende a x_0 ; in altri termini possiamo riscrivere la (1.1) dicendo che

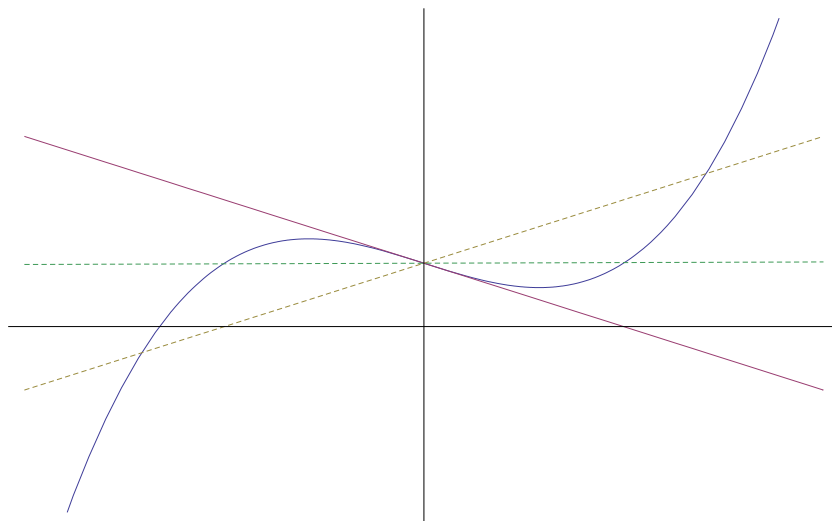
$$(1.4) \quad f(x) - f(x_0) - f'(x_0)(x - x_0) = (x - x_0)E_{x_0}(x)$$

dove $E_{x_0}(x)$ è una quantità che tende a zero per x che tende ad x_0 . L'equazione (1.4) dice anche che la differenza (1.3) è infinitesima di ordine 1 in x_0 , cioè tende a zero se divisa per $(x - x_0)$. La retta (1.2) viene detta anche linearizzazione, o approssimazione del primo ordine di f in x_0 e definisce l'unica retta $ax + b$ per la quale

$$(1.5) \quad \lim_{x \rightarrow x_0} \frac{f(x) - ax - b}{(x - x_0)} = 0;$$

difatti, la precedente espressione, per avere limite finito, deve avere il numeratore che tende a zero in quanto il denominatore tende a zero, cioè

$$\lim_{x \rightarrow x_0} f(x) - ax - b = 0;$$



da cui $b = f(x_0) - ax_0$. Inoltre, dato che (1.5) si presenta a questo punto come una forma indeterminata, applicando il Teorema di De l'Hôpital, la (1.5) diventa equivalente a

$$\lim_{x \rightarrow x_0} f'(x) - a = 0,$$

cioè $a = f'(x_0)$. Alcuni esempi di linearizzazione:

x	per $\sin x$	in $x_0 = 0$,
1	per $\cos x$	in $x_0 = 0$,
$-x + \pi/2$	per $\cos x$	in $x_0 = \pi/2$,
$x - 1$	per $\ln x$	in $x_0 = 1$,
$x + 1$	per e^x	in $x_0 = 0$.

Vediamo ora come andare oltre ed ottenere approssimazioni di ordine successivo; se partiamo dal limite notevole

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}$$

che può equivalentemente essere scritto come

$$\lim_{x \rightarrow 0} \frac{1 - \cos x - x^2/2}{x^2} = 0$$

oppure nella forma

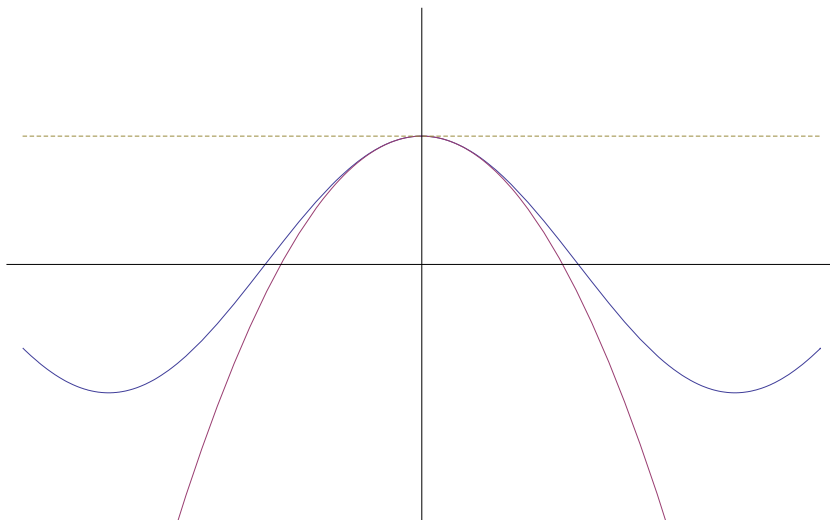
$$\cos x = 1 - \frac{x^2}{2} - x^2 E_0(x)$$

con $E_0(x) \rightarrow 0$ per $x \rightarrow 0$, notiamo che nel sostituire alla funzione $\cos x$ la parabola $1 - \frac{x^2}{2}$ si ottiene un errore, $x^2 E_0(x)$, che è infinitesimo di ordine 2, cioè tende a zero se diviso per x^2 . Un vantaggio nella sostituzione della funzione $\cos x$ non con la sua linearizzazione ma con la parabola $1 - x^2/2$ sta nel fatto che ad esempio si può dedurre che $x_0 = 0$ è un punto di massimo per $\cos x$, in quanto la parabola trovata ha la concavità rivolta verso il

basso (maggiori dettagli su questo verranno dati nella sezione 1.4). Si può in qualche modo affermare che la parabola trovata è la parabola tangente al grafico di $\cos x$ in $x_0 = 0$, nel senso che è l'unica parabola $ax^2 + bx + c$ tale che

$$\lim_{x \rightarrow 0} \frac{\cos x - ax^2 - bx - c}{x^2} = 0.$$

Per formalizzare meglio questi concetti e più in generale il concetto di “polinomio tangente”,



è utile introdurre il simbolo di Landau, anche detto o (leggasi *o piccolo*).

1.1 Infinitesimi e il simbolo o

La trattazione che svilupperemo qui è quella relativa al concetto di infinitesimo e del simbolo di Landau o .

Definizione 1.1 Dato un punto $x_0 \in \mathbb{R} \cup \{-\infty, +\infty\}$ e due funzioni f, g definite in un intorno I di x_0 , $f, g : I \rightarrow \mathbb{R}$, diremo che f è infinitesima rispetto a g in x_0 (oppure che f è un o piccolo di g in x_0 , ossia $f \in o_{x_0}(g)$), se esiste il limite

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0.$$

Nel caso in cui $g(x) = 1$ su I , si dice che f è infinitesima in x_0 , cioè

$$\lim_{x \rightarrow x_0} f(x) = 0$$

e scriveremo $f \in o_{x_0}(1)$. Nel caso in cui $x_0 = 0$, si scrive semplicemente $f \in o(g)$.

Osservazione 1.2 Si può pensare ad $o_{x_0}(g)$ come ad una non meglio precisata funzione, o famiglia di funzioni, con la proprietà

$$\lim_{x \rightarrow x_0} \frac{o_{x_0}(g)}{g(x)} = 0.$$

In relazione a quanto visto in precedenza, si ha che

$$0 = \lim_{x \rightarrow 0} \left(\frac{\sin x}{x} - 1 \right) = \lim_{x \rightarrow 0} \left(\frac{\sin x - x}{x} \right)$$

e quindi $\sin x - x = o(x)$, cioè $\sin x = x + o(x)$, o più in generale, per una funzione derivabile, $f(x) = f(x_0) + f'(x_0)(x - x_0) + o_{x_0}(x - x_0)$. Analogamente, si avrà che $\cos x = 1 + o(1)$ e $\cos x = 1 - \frac{x^2}{2} + o(x^2)$. Un modo per confrontare i due errori dati dalle precedenti formule è quello di confrontare i due infinitesimi $o(1)$ e $o(x^2)$. Abbiamo in generale che, se $a, b \in \mathbb{R}$, allora $x^a = o(x^b)$ se e solo se

$$0 = \lim_{x \rightarrow 0} \frac{x^a}{x^b} = \lim_{x \rightarrow 0} x^{a-b},$$

cioè, se e solo se $a > b$. Quindi ad esempio $x \in o(1)$, $x^2 \in o(x)$, ecc. Nel caso $a > b$ scriveremo anche $o(x^a) \subset o(x^b)$, così ad esempio $o(x^2) \subset o(1)$. Vediamo come utilizzare la nozione di o per definire le approssimazioni polinomiali.

1.2 Formula di Mac Laurin-Taylor

Nella precedente sezione abbiamo dato, per una funzione derivabile, la sua linearizzazione o approssimazione al primo ordine; vediamo ora come arrivare ad approssimazioni di ordine superiore. Per ottenere ciò occorrerà richiedere maggiore regolarità sulla funzione f . Il seguente Teorema fornisce l'approssimazione polinomiale nel punto $x_0 = 0$.

Teorema 1.3 (Formula di Mac Laurin con resto di Peano) *Sia f una funzione derivabile n volte con derivata n -esima continua in $x_0 = 0$; esiste allora un unico polinomio di grado n , denotato con T_n^f o semplicemente T_n , per il quale vale*

$$(1.6) \quad f(x) = T_n(x) + o(x^n).$$

La quantità $o(x^n)$ viene detta resto di Peano ed il polinomio T_n viene detto polinomio di Mac Laurin di grado n ed è determinato dalla formula

$$(1.7) \quad \begin{aligned} T_n(x) &= f(0) + f'(0)x + \dots + \frac{f^{(n)}(0)}{n!}x^n \\ &= \sum_{k=0}^n \frac{f^{(k)}(0)}{k!}x^k \end{aligned}$$

ed è l'unico polinomio di grado n per il quale

$$T_n(0) = f(0), \quad T'_n(0) = f'(0), \quad \dots, \quad T_n^{(n)}(0) = f^{(n)}(0).$$

DIM. Per vedere che il polinomio definito dalla formula (1.7) verifica la (1.6) basta applicare n volte il Teorema di de L'Hôpital; infatti

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{f(x) - T_n(x)}{x^n} &= \lim_{x \rightarrow 0} \frac{f'(x) - T'_n(x)}{nx^{n-1}} = \dots = \lim_{x \rightarrow 0} \frac{f^{(n)}(x) - T_n^{(n)}(x)}{n!} \\ &= \lim_{x \rightarrow 0} \frac{f^{(n)}(x) - f^{(n)}(0)}{n!} = 0 \end{aligned}$$

grazie alla continuità della derivata n -esima. Per dimostrare l'unicità di T_n , si supponga che esista un altro polinomio $p_n(x) = a_0 + a_1x + \dots + a_nx^n$ di grado n che soddisfa la (1.6); allora $f(x) - p_n(x) = o(x^k)$ per ogni $k = 0, \dots, n$. In particolare, per $k = 0$, si ha che

$$\lim_{x \rightarrow 0} (f(x) - p_n(x)) = 0,$$

da cui $p_n(0) = f(0)$; analogamente, applicando nuovamente il Teorema di de L'Hôpital, si ricava che

$$0 = \lim_{x \rightarrow 0} \frac{f(x) - p_n(x)}{x^k} = \lim_{x \rightarrow 0} \frac{f^{(k)}(x) - p_n^{(k)}(x)}{k!},$$

da cui $f^{(k)}(0) = p_n^{(k)}(0) = k!a_k$. Questo implica che

$$a_k = \frac{f^{(k)}(0)}{k!}, \quad \forall k = 0, \dots, n.$$

□

Il primo, e più semplice, esempio che si può fare è la funzione esponenziale; difatti, per $f(x) = e^x$ si ha $f^{(n)}(x) = e^x$ per ogni $n \in \mathbb{N}$, e quindi

$$T_n^{e^x}(x) = 1 + x + \frac{x^2}{2} + \dots + \frac{x^n}{n!},$$

da cui

$$e^x = 1 + x + \frac{x^2}{2} + \dots + \frac{x^n}{n!} + o(x^n).$$

In modo analogo si ottengono gli sviluppi

$$\sin x = x - \frac{x^3}{3!} + \dots + \frac{(-1)^k x^{2k+1}}{(2k+1)!} + o(x^{2k+2}),$$

$$\cos x = 1 - \frac{x^2}{2} + \dots + \frac{(-1)^k x^{2k}}{(2k)!} + o(x^{2k+1}).$$

Non è sorprendente trovare che lo sviluppo di Mac Laurin della funzione coseno contiene solo potenze pari di x , essendo la funzione coseno una funzione pari; allo stesso modo, dato che la funzione seno è dispari, lo sviluppo di Mac Laurin avrà solo potenze dispari.

Come esercizio, si provi a ricavare gli sviluppi di Mac Laurin per le funzioni

$$\ln(1+x), \sqrt{1+x}, (1+x)^\alpha, \sinh(x), \cosh(x).$$

Il seguente Teorema fornisce una stima più precisa dell'errore che si commette approssimando una funzione con il suo polinomio di Mac Laurin; esso fornisce una valutazione della quantità $o(x^n)$.

Teorema 1.4 (Formula di Taylor; resto integrale e di Lagrange) *Sia f una funzione derivabile $n+1$ volte con derivata $(n+1)$ -esima continua; allora il resto della formula di Mac Laurin può essere espresso in forma integrale come*

$$(1.8) \quad o(x^n) = x^{n+1} \int_0^1 \frac{(1-t)^n}{n!} f^{(n+1)}(tx) dt = \int_0^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt \quad (\text{resto integrale})$$

oppure in forma differenziale

$$(1.9) \quad o(x^n) = \frac{f^{(n+1)}(c)}{(n+1)!} x^{n+1} \quad (\text{resto di Lagrange})$$

con c elemento dell'intervallo di estremi x e 0 .

DIM. La dimostrazione della formula con resto integrale si può fare per induzione partendo dalla funzione $g(t) = f(tx)$; notando che $g(0) = f(0)$, $g(1) = f(x)$ e $g'(t) = (x)f'(tx)$, dal Teorema fondamentale del calcolo integrale si ottiene

$$g(1) = g(0) + \int_0^1 g'(t) dt,$$

da cui

$$f(x) = f(0) + x \int_0^1 f'(tx) dt.$$

Si noti che

$$x \int_0^1 f'(tx) dt = o(1)$$

e quindi la formula di Taylor di ordine 0 è dimostrata. Per procedere con il passo induttivo, si integra per parti ottenendo

$$\begin{aligned} \int_0^1 f'(tx) dt &= -(1-t)f'(tx) \Big|_0^1 + x \int_0^1 (1-t)f''(tx) dt \\ &= f'(0) + x \int_0^1 (1-t)f''(tx) dt \end{aligned}$$

o più in generale

$$\int_0^1 \frac{(1-t)^k}{k!} f^{(k+1)}(tx) dt = \frac{f^{(k+1)}(0)}{k!} + x \int_0^1 \frac{(1-t)^{k+1}}{(k+1)!} f^{(k+2)}(tx) dt,$$

da cui la formula

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + x^{n+1} \int_0^1 \frac{(1-t)^n}{(n)!} f^{(n+1)}(tx) dt;$$

si noti che la (1.8) è immediata. Per la (1.9) si applicherà invece il Teorema di Lagrange; esso infatti afferma che esiste c nell'intervallo di estremi x e 0 tale che

$$\frac{f(x) - f(0)}{x} = f'(c),$$

cioè

$$f(x) = f(0) + f'(c)x$$

che è la formula di Taylor di ordine 0; per ottenere la formula generale si procederà per induzione. \square

La formula con il resto di Lagrange permette di dare una stima del tipo

$$|f(x) - T_n(x)| = \frac{|f^{(n+1)}(c)||x|^{n+1}}{(n+1)!}.$$

Ad esempio, se si vuole calcolare \sqrt{e} (sapendo a priori che $e \leq 3$), si può utilizzare lo sviluppo di Mac Laurin per $f(x) = e^x$ con $x = 1/2$: ad esempio, se fissiamo $n = 3$, avremo:

$$T_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

con $T_3(1/2) = \frac{79}{48}$. Inoltre $|\sqrt{e} - T_3(1/2)| = \frac{e^c}{2^4 4!} \leq \frac{\sqrt{3}}{2^4 4!} \sim 0,0045$ in quanto $c \in [0, 1/2]$.

Tutti le considerazioni fatte nel caso del polinomio di Mac Laurin per $x_0 = 0$ possono essere trasferiti ad un generico punto x_0 ; si definisce quindi il polinomio di Taylor di grado n centrato nel punto x_0 nel seguente modo.

Teorema 1.5 (Formula di Taylor) *Sia f una funzione derivabile n volte con derivata n -esima continua in un punto x_0 ; esiste allora un unico polinomio di grado n in $(x - x_0)$, denotato con $T_{x_0,n}^f$ o semplicemente T_n se la funzione f ed il punto x_0 sono noti, per il quale vale*

$$(1.10) \quad f(x) = T_{x_0,n}^f(x) + o((x - x_0)^n). \quad (\text{Formula di Taylor con resto di Peano}).$$

Il polinomio di Taylor è dato da

$$(1.11) \quad \begin{aligned} T_{x_0,n}^f(x) &= f(x_0) + f'(x_0)(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \\ &= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k. \end{aligned}$$

Se poi la funzione f è derivabile $n + 1$ volte in x_0 , allora si ottengono le formule di Taylor con resto integrale e di Lagrange

$$\begin{aligned} f(x) &= T_{x_0,n}^f(x) + (x - x_0)^{n+1} \int_0^1 \frac{(1-t)^n}{n!} f^{(n+1)}(t(x - x_0)) dt \\ &= T_{x_0,n}^f(x) + \int_{x_0}^x \frac{(x - x_0 - t)^n}{n!} f^{(n+1)}(t) dt \\ &= T_{x_0,n}^f(x) + \frac{f^{(n+1)}(c)}{(n+1)!} (x - x_0)^{n+1}, \end{aligned}$$

dove c è un punto compreso nell'intervallo di estremi x_0 e x .

Ad esempio, si possono calcolare i polinomi di Taylor di

$$f(x) = \sin x, x_0 = \frac{\pi}{2}, \quad f(x) = e^x, x_0 = 1, \quad f(x) = \ln(1 + x), x_0 = 2.$$

Osservazione 1.6 Una domanda che ci si potrebbe porre è se l'approssimazione che si ottiene al variare del grado n del polinomio di Taylor possa migliorare per $n \rightarrow +\infty$. Supponiamo quindi che la funzione sia derivabile infinite volte in x_0 ; in questo caso possiamo

costruirci i polinomi $T_{x_0,n}^f$ per ogni grado n . Nel passaggio al limite per $n \rightarrow \infty$ subentrano però due problemi; il primo è che stiamo definendo una serie, quindi bisogna porsi il problema della convergenza della serie in considerazione. Il secondo, ammesso che la serie converga, consiste nel chiedersi se la serie

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k := g(x),$$

che viene detta serie di Taylor associata ad f , coincida o meno con f , cioè se $g(x) = f(x)$. Questo non è sempre vero, come mostra la funzione

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0; \end{cases}$$

tale funzione è derivabile infinite volte in 0 con derivate tutte nulle, quindi la serie associata $g(x)$ definisce la funzione nulla, mentre f non è nulla. Il problema delle serie di Taylor non verrà affrontato in questo corso, ma sarà argomento di corsi più avanzati.

1.3 Algebra degli o e formule di Taylor di funzioni composte

Usando le potenze di x , si può dare la seguente definizione.

Definizione 1.7 Sia $x_0 \in \mathbb{R} \cup \{-\infty, +\infty\}$ e sia f una funzione definita in un intorno I di x_0 , $f : I \rightarrow \mathbb{R}$. Avremo allora le seguenti possibilità:

1. nel caso $x_0 \in \mathbb{R}$, se esiste $\lambda \in \mathbb{R} \setminus \{0\}$ tale che per $a \neq 0$ si abbia

$$\lim_{x \rightarrow x_0} \frac{f(x)}{(x - x_0)^a} = \lambda,$$

diremo che f è infinitesima di ordine a se $a > 0$, mentre f è infinita di ordine a se $a < 0$;

2. nel caso $x_0 = \pm\infty$, se esiste $\lambda \in \mathbb{R} \setminus \{0\}$ ed esiste $a \neq 0$ tale che il limite

$$\lim_{x \rightarrow \pm\infty} \frac{f(x)}{x^a} = \lambda,$$

diremo che f è infinitesima di ordine a se $a < 0$, mentre f è infinita di ordine a se $a > 0$.

In entrambi i precedenti casi si potrà anche dire che f ha ordine a in x_0 e si scriverà $\text{ord}_{x_0}(f) = a$.

Ad esempio, dal limite notevole

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1,$$

si ricava che $\sin x$ è infinitesima di ordine 1 in $x_0 = 0$ o anche che

$$\text{ord}_0(\sin x) = 1,$$

mentre dal limite

$$\lim_{x \rightarrow 0} \frac{2x^2}{1 - \cos x} = 1$$

si deduce che la funzione $\frac{2}{1 - \cos x}$ è infinita di ordine 2 per $x_0 = 0$ o anche che

$$\text{ord}_0 \left(\frac{2}{1 - \cos x} \right) = -2.$$

Abbiamo la seguente proposizione.

Proposizione 1.8 (Algebra degli o e relazioni tra o e ord) *Sia $x_0 \in \mathbb{R} \cup \{-\infty, +\infty\}$, siano f e g due funzioni definite in un intorno I di x_0 , $f, g : I \rightarrow \mathbb{R}$ e sia $\lambda \in \mathbb{R} \setminus \{0\}$. Valgono allora le seguenti proprietà;*

1. $o_{x_0}(f) \pm o_{x_0}(f) = o_{x_0}(f)$;
2. $\lambda o_{x_0}(f) = o_{x_0}(f)$;
3. $o_{x_0}(f)o_{x_0}(g) = o_{x_0}(fg)$;
4. se $f \in o_{x_0}(g)$, allora $o_{x_0}(f) \subset o_{x_0}(g)$, $o_{x_0}(f) + o_{x_0}(g) = o_{x_0}(g)$;
5. se $\text{ord}_{x_0}(f) = a$, allora, se $x_0 \in \mathbb{R}$ si avrà $f \in o_{x_0}((x - x_0)^b)$ per ogni $b < a$, mentre se $x_0 = \pm\infty$ si avrà che $f \in o_{\pm\infty}(x^b)$ per ogni $b > a$.

DIM. Grazie all'Osservazione 1.2, si ha che

$$\lim_{x \rightarrow x_0} \frac{o_{x_0}(f) \pm o_{x_0}(f)}{f(x)} = \lim_{x \rightarrow x_0} \frac{o_{x_0}(f)}{f(x)} \pm \lim_{x \rightarrow x_0} \frac{o_{x_0}(f)}{f(x)} = 0,$$

da cui la 1. Analogamente,

$$\lim_{x \rightarrow x_0} \frac{\lambda o_{x_0}(f)}{f(x)} = \lambda \lim_{x \rightarrow x_0} \frac{o_{x_0}(f)}{f(x)}$$

da cui l'equivalenza 2. in quanto $\lambda \neq 0$. Per la 3. si nota che

$$\lim_{x \rightarrow x_0} \frac{o_{x_0}(f)o_{x_0}(g)}{f(x)g(x)} = \lim_{x \rightarrow x_0} \frac{o_{x_0}(f)}{f(x)} \lim_{x \rightarrow x_0} \frac{o_{x_0}(g)}{g(x)} = 0.$$

Per la proprietà 4. si nota che

$$\lim_{x \rightarrow x_0} \frac{o_{x_0}(f)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} \frac{o_{x_0}(f)}{f(x)} = 0.$$

Infine, dimostriamo la 5. solo nel caso $x_0 \in \mathbb{R}$ (essendo il caso $x_0 \pm \infty$ analogo); dalla definizione di ordine, si ottiene che

$$\lim_{x \rightarrow x_0} \frac{f(x)}{(x - x_0)^b} = \lim_{x \rightarrow x_0} (x - x_0)^{a-b} \frac{f(x)}{(x - x_0)^a} = \lambda \lim_{x \rightarrow x_0} (x - x_0)^{a-b} = 0$$

se e solo se $a > b$. □

Con questi strumenti si può ad esempio calcolare il polinomio di Mac Laurin della funzione $f(x) = e^{\sin x}$ senza calcolare le sue derivate; difatti, sfruttando gli sviluppi di e^x e $\sin x$, si ottiene

$$\begin{aligned} e^{\sin x} &= e^{x - \frac{x^3}{6} + o(x^3)} \\ &= 1 + \left(x - \frac{x^3}{6} + o(x^3)\right) + \frac{1}{2} \left(x - \frac{x^3}{6} + o(x^3)\right)^2 + \\ &\quad + \frac{1}{6} \left(x - \frac{x^3}{6} + o(x^3)\right)^3 + o\left(\left(x - \frac{x^3}{6} + o(x^3)\right)^3\right) \\ &= 1 + x + \frac{x^2}{2} + o(x^3) \end{aligned}$$

da cui

$$T_3^{e^{\sin x}}(x) = 1 + x + \frac{x^2}{2}.$$

1.4 Applicazioni della formula di Taylor

Una prima applicazione della formula di Taylor si può avere, ad esempio, nel calcolo dei limiti. Abbiamo infatti il seguente risultato.

Proposizione 1.9 *Sia $x_0 \in \mathbb{R} \cup \{-\infty, +\infty\}$ e siano f , f_1 e g tre funzioni definite in un intorno I di x_0 , $f, f_1, g : I \rightarrow \mathbb{R}$; allora, se $f_1 \in o_{x_0}(f)$, il limite*

$$\lim_{x \rightarrow x_0} \frac{f(x) + f_1(x)}{g(x)}$$

esiste se e solo se esiste il limite

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$$

e vale

$$\lim_{x \rightarrow x_0} \frac{f(x) + f_1(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}.$$

DIM. La dimostrazione segue semplicemente osservando che

$$\frac{f(x) + f_1(x)}{g(x)} = \frac{f(x)}{g(x)} \left(1 + \frac{f_1(x)}{f(x)}\right)$$

e dal fatto che, essendo $f_1 \in o_{x_0}(f)$,

$$\lim_{x \rightarrow x_0} \left(1 + \frac{f_1(x)}{f(x)}\right) = 1.$$

□

Abbiamo quindi che, se $f(x) = T_{n,x_0}^f(x) + o_{x_0}((x-x_0)^n)$ e $g(x) = T_{m,x_0}^g(x) + o((x-x_0)^m)$ con polinomi non nulli, allora

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{T_{n,x_0}^f(x)}{T_{m,x_0}^g(x)}$$

almeno ogni qualvolta il limite di destra non si presenti in forma indeterminata. Questa osservazione rende il calcolo dei limiti un problema più semplice, in quanto ridotto al limite di una espressione razionale. Così ad esempio

$$\lim_{x \rightarrow 0} \frac{x - \operatorname{sen} x}{x^3} = \lim_{x \rightarrow 0} \frac{x - x + \frac{x^3}{6} + o(x^3)}{x^3} = \lim_{x \rightarrow 0} \frac{\frac{x^3}{6}}{x^3} = \frac{1}{6}.$$

Così, ad esempio, se si volesse studiare la convergenza della serie

$$\sum_{n=1}^{\infty} \left(\frac{1}{n} - \operatorname{sen} \frac{1}{n} \right),$$

il limite precedente ci dice che il termine generale è asintoticamente equivalente a $\frac{1}{6n^3}$, e quindi la serie è assolutamente convergente.

Altra applicazione si può avere nello studio della convessità e nella classificazione dei punti stazionari. Supponiamo ad esempio che $f'' \geq 0$; dalla formula di Taylor con resto di Lagrange si ottiene quindi che

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(c)(x - x_0)^2}{2} \geq f(x_0) + f'(x_0)(x - x_0),$$

quindi il grafico di f si trova al di sopra del grafico della sua retta tangente in $(x_0, f(x_0))$. Inoltre, se x_0 è un punto stazionario e $f''(x_0) > 0$, allora dalla formula di Taylor di ordine 2 si ottiene

$$f(x) = f(x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + o_{x_0}((x - x_0)^2)$$

e quindi

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{(x - x_0)^2} = \frac{f''(x_0)}{2} > 0$$

da cui la positività di $f(x) - f(x_0)$ in un intorno di x_0 , cioè il fatto che x_0 è un punto di minimo.

Usando il polinomio di Taylor di grado superiore, si può enunciare il seguente risultato riguardante la classificazione dei punti stazionari di una funzione.

Proposizione 1.10 *Supponiamo che f sia una funzione derivabile n volte con continuità in x_0 e supponiamo che*

$$f'(x_0) = \dots = f^{(n-1)}(x_0) = 0, \quad f^{(n)}(x_0) \neq 0.$$

Allora:

1. *se n è pari e $f^{(n)}(x_0) > 0$, x_0 è un punto di minimo;*
2. *se n è pari e $f^{(n)}(x_0) < 0$, x_0 è un punto di massimo;*
3. *se n è dispari, x_0 è un punto di flesso, ascendente se $f^{(n)}(x_0) > 0$, discendente se $f^{(n)}(x_0) < 0$.*

Capitolo 2

I Numeri Complessi

In questo capitolo daremo la definizione e le principali proprietà di un nuovo insieme numerico: il campo¹ dei numeri complessi \mathbb{C} .

La motivazione che spinge ad introdurre questo nuovo insieme numerico, quantomeno per il presente corso, viene dalla necessità di risolvere equazioni, principalmente, del secondo ordine, cioè equazioni della forma

$$ax^2 + bx + c = 0;$$

come è ben noto, le soluzioni di tale equazione sono date dalla formula

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Tale formula funziona con i seguenti accorgimenti;

1. se $b^2 - 4ac > 0$, allora la radice quadrata è ben definita e si ottengono due soluzioni reali distinte;
2. se $b^2 - 4ac = 0$, la radice non compare e si ottiene una sola soluzione, o meglio due soluzioni coincidenti.

Resta il problema del caso $b^2 - 4ac < 0$, in cui ci si trova di fronte al problema di dover calcolare

$$\sqrt{b^2 - 4ac} = \sqrt{-|b^2 - 4ac|} = \sqrt{-1} \sqrt{|b^2 - 4ac|};$$

nella precedente espressione basta dare un senso alla radice quadrata del numero negativo -1 per ottenere una buona formula risolutiva di ogni polinomio di secondo grado, senza distinzioni sul discriminante (in realtà l'unica distinzione sarà discriminante nullo o non

¹Ricordiamo che per campo si intende un insieme K sul quale siano definite due operazioni, dette somma e prodotto e denotate con $+$ e \cdot , per le quali valgono le proprietà:

1. $+$ è associativa, ammette elemento neutro, denotato con 0 , e ogni elemento $a \in K$ è invertibile rispetto alla somma con inversa denotata con $-a$;
2. \cdot è associativa, ammette elemento neutro, denotato con 1 , e ogni elemento $a \in K$ con $a \neq 0$ è invertibile rispetto al prodotto con inverso denotato con a^{-1} o $\frac{1}{a}$;
3. le operazioni di somma e prodotto godono della proprietà distributiva.

Se le operazioni di somma e prodotto sono commutative, si parla di campo commutativo o abeliano.

nullo). Si tratta quindi di trovare un insieme numerico in cui l'equazione $x^2 + 1 = 0$ abbia soluzione; come vedremo, risolvere quest'ultima equazione renderà possibile trovare le radici non solo di polinomi di secondo grado, ma di grado arbitrario e a coefficienti non reali (vedi Teorema 2.13).

2.1 Definizione e prime proprietà

In questa sezione daremo la definizione e le principali proprietà dei numeri complessi. Come si è visto nel corso di Analisi Matematica I, si parte dall'insieme numerico \mathbb{N} sul quale sono ben definite le operazioni di somma e prodotto ma nel quale non esistono gli elementi inversi rispetto a queste due operazioni. Si introducono quindi delle estensioni di \mathbb{N} in cui sono ancora definite somma e prodotto, che sono estensioni della somma e prodotto su \mathbb{N} , in modo che esistano gli elementi inversi rispetto alla somma (e si ottiene così l'insieme \mathbb{Z}) e prodotto (ottenendo così \mathbb{Q}). L'introduzione di \mathbb{R} viene fatta in modo che ci sia completezza non tanto rispetto alle operazioni di somma e prodotto, ma rispetto alla convergenza delle successioni di Cauchy, ottenendo in definitiva le seguenti inclusioni $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$. Vogliamo definire qui il campo \mathbb{C} in modo che $\mathbb{R} \subset \mathbb{C}$ e che le operazioni di somma e prodotto su \mathbb{C} non siano altro che estensioni della somma e prodotto su \mathbb{R} .

Esistono vari modi equivalenti di definire \mathbb{C} ; quello che seguiremo è quello di tipo cartesiano.

Definizione 2.1 (Campo complesso) Diremo campo complesso, e lo denoteremo con \mathbb{C} , l'insieme consistente nel piano \mathbb{R}^2 (che prende il nome di piano di Gauss) munito delle operazioni di somma e prodotto

$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

Un generico elemento di \mathbb{C} si indicherà con le ultime lettere dell'alfabeto, z, w, \dots intendendo $z = (a, b)$, ecc. La prima componente di un numero complesso si chiama parte reale, $\operatorname{Re}(z) = \operatorname{Re}(a, b) = a$ e la seconda componente si chiama parte immaginaria, $\operatorname{Im}(z) = \operatorname{Im}(a, b) = b$ (si tenga ben presente che parte reale e parte immaginaria di un numero complesso sono entrambi numeri reali). Un numero complesso z si dirà reale (o reale puro) se $\operatorname{Im}(z) = 0$, mentre si dirà immaginario puro se $\operatorname{Re}(z) = 0$.

Osservazione 2.2 Dato che l'insieme dei numeri complessi è definito tramite una coppia ordinata, l'uguaglianza tra numeri complessi, $z = w$ con $z = (a, b)$ e $w = (c, d)$, si verificherà se e solo se $a = c$ e $b = d$, cioè se e solo se $\operatorname{Re}(z) = \operatorname{Re}(w)$ e $\operatorname{Im}(z) = \operatorname{Im}(w)$.

La terminologia *campo* è motivata dalla seguente Proposizione.

Proposizione 2.3 \mathbb{C} è un campo abeliano.

DIM. Le proprietà di associatività e commutatività della somma sono immediate, mentre quelle per il prodotto sono lasciate come verifica. Si nota quindi che gli elementi $(0, 0)$ e $(1, 0)$ sono elementi neutri per la somma e per il prodotto rispettivamente; si nota infine che l'elemento $-z = -(a, b) = (-a, -b)$ è l'inverso additivo di $z = (a, b)$, mentre, se $z \neq (0, 0)$, allora

$$(2.1) \quad z^{-1} = \frac{1}{z} = \left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2} \right)$$

è l'inverso moltiplicativo di z . Si noti che dire $z \neq (0, 0)$ significa che almeno uno tra a e b è diverso da zero, da cui il fatto che $a^2 + b^2 \neq 0$ e cioè la buona definizione di z^{-1} \square

Vediamo ora in che senso \mathbb{C} è una estensione di \mathbb{R} ; si nota che sull'insieme

$$R = \{(a, 0) : a \in \mathbb{R}\}$$

le operazioni sopra definite si riducono a

$$(a, 0) + (b, 0) = (a + b, 0). \quad (a, 0) \cdot (b, 0) = (ab, 0).$$

ed inoltre

$$-(a, 0) = (-a, 0), \quad (a, 0)^{-1} = (1/a, 0), \quad a \neq 0.$$

Identificando quindi R con \mathbb{R} , avremo che $\mathbb{R} \subset \mathbb{C}$; useremo sempre questa identificazione che ci porta a scrivere la seguente uguaglianza

$$(a, 0) = a.$$

Avremo in particolare che gli elementi neutri rispetto a somma e prodotto in \mathbb{C} sono gli stessi di quelli in \mathbb{R} , essendo $0 = (0, 0)$ e $1 = (1, 0)$.

Osservazione 2.4 Si noti che, mentre \mathbb{R} è un campo ordinato, su \mathbb{C} non abbiamo introdotto nessuna nozione di ordinamento; questo è dovuto al fatto che non c'è un modo naturale per estendere l'ordinamento \leq su \mathbb{C} e non avrà quindi senso per i numeri complessi l'espressione $z \leq w$.

Notiamo inoltre che, con le notazioni appena introdotte, si ha:

$$(0, 1) \cdot (0, 1) = (-1, 0) = -1,$$

cioè l'elemento $i = (0, 1)$ ha la proprietà che $i^2 = -1$; tale elemento verrà chiamato unità immaginaria. In questo modo siamo arrivati a poter scrivere un numero complesso, oltre che con la notazione cartesiana, anche in notazione algebrica. Infatti, abbiamo che

$$(a, b) = (a, 0) + (0, b) = (a, 0) + (0, 1) \cdot (b, 0) = a + ib.$$

La definizione algebrica dei numeri complessi passa quindi attraverso la definizione di un numero *speciale* i con la proprietà che $i^2 = -1$, e definendo numero complesso tutte le possibili combinazioni lineari degli elementi 1 e i , cioè tutti i numeri della forma appunto

$$a + ib, \quad a, b \in \mathbb{R}.$$

In questo modo l'operazione di prodotto diventa più intuitiva; infatti dati due numeri complessi $a + ib, c + id$, abbiamo, tenendo presente che $i^2 = -1$,

$$(a + ib) \cdot (c + id) = a \cdot c + a \cdot id + ib \cdot c + ib \cdot id = ac - bd + i(ad + bc).$$

2.2 Coniugato e modulo di un numero complesso

Sui numeri complessi è definita l'operazione di coniugio; dato cioè un numero complesso $z = a + ib$, si definisce il numero complesso \bar{z} detto coniugato di z tramite

$$\bar{z} = a - ib.$$

Per l'operazione di coniugio abbiamo le seguenti proprietà.

Proposizione 2.5 *Siano $z, w \in \mathbb{C}$; allora*

1. $\bar{\bar{z}} = z$ (proprietà involutiva del coniugio);
2. $\bar{z} = z$ se e solo se z reale;
3. $\overline{z + w} = \bar{z} + \bar{w}$;
4. $\overline{zw} = \bar{z} \cdot \bar{w}$;
5. se $z \neq 0$, allora $\overline{1/z} = 1/\bar{z}$;
6. $z + \bar{z} = 2\operatorname{Re}(z)$, $z - \bar{z} = 2i\operatorname{Im}(z)$.

DIM. Basta scrivere z e w in forma algebrica $z = a + ib$, $w = c + id$ e verificare le identità; vediamo solamente le dimostrazioni di 4. e 5. Dato che $zw = (ac - bd) + i(ad + bc)$, segue che $\overline{zw} = (ac - bd) - i(ad + bc)$, mentre

$$\bar{z} \cdot \bar{w} = (a - ib)(c - id) = (ac - bd) + i(-ad - bc)$$

da cui la 4. Per la 5. si nota che da (2.1) applicata a z e \bar{z} , si ottiene che

$$\frac{1}{z} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2}, \quad \frac{1}{\bar{z}} = \frac{a}{a^2 + b^2} + i \frac{b}{a^2 + b^2},$$

da cui la 5. □

Notiamo che scrivendo $z = a + ib$ si ottiene che $z\bar{z} = a^2 + b^2$, e quindi il fatto che $z\bar{z}$ è un numero reale positivo. Possiamo quindi dare la seguente definizione.

Definizione 2.6 (Modulo in \mathbb{C}) *Dato un numero complesso $z \in \mathbb{C}$, si definisce il modulo di z tramite*

$$|z| = \sqrt{z\bar{z}}.$$

Osservazione 2.7 Non si deve confondere la notazione di modulo di un numero complesso con quella di valore assoluto di un numero reale; tuttavia, le due nozioni coincidono nel caso in cui z sia reale puro, in quanto in questo caso $b = 0$ e quindi

$$|z| = \sqrt{a^2} = |a|.$$

Per il modulo di un numero complesso valgono le seguenti proprietà.

Proposizione 2.8 *Siano $z, w \in \mathbb{C}$; allora*

1. $|z| \geq 0$, $|z| = 0$ se e solo se $z = 0$;

2. $|z| = |\bar{z}|$, $|-z| = |z|$;
3. $|\operatorname{Re}(z)| \leq |z|$, $|\operatorname{Im}(z)| \leq |z|$, $|z| \leq |\operatorname{Re}(z)| + |\operatorname{Im}(z)|$;
4. se $z \neq 0$, $|1/z| = 1/|z|$, $1/z = \bar{z}/|z|^2$;
5. $|zw| = |z||w|$;
6. $|z + w| \leq |z| + |w|$ (disuguaglianza triangolare);
7. $|z + w| \geq ||z| - |w||$ (seconda disuguaglianza triangolare).

DIM. Si scrivono $z = a + ib$ e $w = c + id$ e le proprietà 1., 2. e 3. sono immediate. Notando poi che per $z \neq 0$

$$\left| \frac{1}{z} \right| = \left| \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2} \right| = \sqrt{\frac{a^2}{(a^2 + b^2)^2} + \frac{b^2}{(a^2 + b^2)^2}} = \frac{1}{\sqrt{a^2 + b^2}} = \frac{1}{|z|}$$

da cui la 4. (la seconda proprietà della 4. segue direttamente dalla definizione $|z|^2 = z\bar{z}$). Per la 5., si nota che, dalle proprietà del coniugato, si ottiene:

$$|zw|^2 = zw\bar{z}\bar{w} = zw\bar{z} \cdot \bar{w} = z\bar{z}w\bar{w} = |z|^2|w|^2.$$

Per quanto riguarda la disuguaglianza triangolare, tenendo presente la proprietà 6. della Proposizione 2.5 e la 3. di questa Proposizione, si ha:

$$\begin{aligned} |z + w|^2 &= (z + w)\overline{(z + w)} = z\bar{z} + w\bar{w} + z\bar{w} + \bar{z}w = |z|^2 + |w|^2 + 2\operatorname{Re}(z\bar{w}) \\ &\leq |z|^2 + |w|^2 + 2|z\bar{w}| = |z|^2 + |w|^2 + 2|z| \cdot |w| \\ &= (|z| + |w|)^2, \end{aligned}$$

da cui la 6. Per quanto riguarda la 7., basta notare che

$$|z| = |z - w + w| \leq |z - w| + |w|,$$

da cui $|z| - |w| \leq |z - w|$; analogamente

$$|w| \leq |w - z| + |z| = |z - w| + |z|,$$

da cui $|w| - |z| \leq |z - w|$. Mettendo insieme queste due disuguaglianze, si ottiene la 7. \square

2.3 Forma polare ed esponenziale

Nelle sezioni precedenti abbiamo introdotto le forme cartesiane ed algebrica di un numero complesso; in questa sezione introdurremo la forma polare ed esponenziale. Il vantaggio di queste nuove definizioni è che, mentre la forma cartesiana e algebrica è comoda quando si vogliono sommare due numeri complessi, la forma polare ed esponenziale lo sono nella moltiplicazione.

La definizione della forma polare di un numero complesso segue dall'osservazione che un numero complesso $z = a + ib = (a, b)$ rappresenta un punto nel piano \mathbb{R}^2 ed è quindi determinato dalle sue coordinate polari (ϱ, ϑ) , dove $\varrho = \sqrt{a^2 + b^2}$ rappresenta la distanza

del punto (a, b) dall'origine e coincide con il modulo del numero complesso, $\varrho = |z|$, mentre l'angolo ϑ , detto argomento o anomalia e denotato con $\vartheta = \arg(z)$, rappresenta l'angolo, preso in senso antiorario, formato dal semiasse $\{x = 0, y \geq 0\}$ e la semiretta originata in $(0, 0)$ e passante per (a, b) . Si nota che mentre ϱ è univocamente determinato, l'angolo ϑ è individuato a meno di multipli di 2π (fa eccezione l'origine, che è individuata da $\varrho = 0$ ma non ha un ϑ determinato) ed è univocamente determinato in un intervallo semiaperto di ampiezza 2π ; si parla in questo caso di argomento principale di z e come intervallo si può scegliere $(-\pi, \pi]$ (o $[0, 2\pi)$ a seconda dei casi). Per passare dalla forma algebrica alle coordinate polari del numero $z = a + ib$ si possono usare, nel caso $z \neq 0$, le formule

$$(2.2) \quad \begin{cases} a = \varrho \cos \vartheta \\ b = \varrho \sin \vartheta \end{cases}, \quad \begin{cases} \varrho = \sqrt{a^2 + b^2} \\ \cos \vartheta = \frac{a}{\sqrt{a^2 + b^2}} \\ \sin \vartheta = \frac{b}{\sqrt{a^2 + b^2}} \end{cases}$$

Possiamo quindi dare la seguente definizione.

Definizione 2.9 (Forma polare) Dato $z \in \mathbb{C}$, si chiama forma polare (o trigonometrica) l'espressione di z usata utilizzando le coordinate polari

$$z = \varrho(\cos \vartheta + i \sin \vartheta).$$

Uno dei maggiori vantaggi della rappresentazione polare dei numeri complessi si presenta quando si deve fare il prodotto di due numeri complessi. Supponiamo infatti di avere due numeri complessi $z = \varrho(\cos \vartheta + i \sin \vartheta)$, $w = r(\cos \phi + i \sin \phi)$, otteniamo che

$$(2.3) \quad zw = \varrho r(\cos(\vartheta + \phi) + i \sin(\vartheta + \phi)), \quad \frac{z}{w} = \frac{\varrho}{r}(\cos(\vartheta - \phi) + i \sin(\vartheta - \phi)),$$

cioè la moltiplicazione per il numero w è data da una dilatazione pari a r e una rotazione di un angolo ϕ . In particolare, se $w = z$, si ottiene che

$$z^2 = \varrho^2(\cos 2\vartheta + i \sin 2\vartheta),$$

o più in generale, per ogni $n \in \mathbb{N}$

$$(2.4) \quad z^n = \varrho^n(\cos n\vartheta + i \sin n\vartheta);$$

le formule (2.3) e (2.4) prendono anche il nome di Formule di De Moivre. Dato che la funzione

$$f(\vartheta) = \cos \vartheta + i \sin \vartheta,$$

ha la proprietà che $f(\vartheta_1)f(\vartheta_2) = f(\vartheta_1 + \vartheta_2)$, ha senso la seguente definizione.

Definizione 2.10 Si definisce l'esponenziale immaginario come la funzione $f : \mathbb{R} \rightarrow \mathbb{C}$ data da

$$e^{i\vartheta} = f(\vartheta) = \cos \vartheta + i \sin \vartheta;$$

Più in generale, dato un numero complesso $z = a + ib$, si definisce l'esponenziale complesso $e : \mathbb{C} \rightarrow \mathbb{C}$

$$e^z = e^{a+ib} = e^a e^{ib} = e^a(\cos b + i \sin b).$$

Osservazione 2.11 Si noti che, dalla definizione data, si ha che la funzione $\vartheta \mapsto e^{i\vartheta}$ è 2π -periodica e che per ogni $\vartheta \in \mathbb{R}$ $|e^{i\vartheta}| = 1$.

Definizione 2.12 (Forma esponenziale) Dato un numero complesso $z \in \mathbb{C}$, si chiama forma esponenziale di z la scrittura di z nella forma

$$z = \varrho e^{i\vartheta},$$

dove ϱ e ϑ sono determinate dalle (2.2).

2.4 Polinomi e radici n -esime

In questa sezione tratteremo i polinomi in campo complesso. Ricordiamo che un polinomio complesso di grado n è una funzione $p: \mathbb{C} \rightarrow \mathbb{C}$ definita da

$$p(z) = a_0 + a_1 z + \cdots + a_n z^n, \quad a_0, \dots, a_n \in \mathbb{C}, a_n \neq 0.$$

Ricordiamo anche che un numero complesso z_0 si dice radice del polinomio p se $p(z_0) = 0$; in tal caso il polinomio p è divisibile per $(z - z_0)$ e si potrà scrivere

$$p(z) = (z - z_0)q(z)$$

con q polinomio di grado $n - 1$. Si dice inoltre che z_0 ha molteplicità m se

$$p(z) = (z - z_0)^m q(z)$$

con q polinomio di grado $n - m$ tale che $q(z_0) \neq 0$; in tal caso $p(z)$ è divisibile per $(z - z_0)^m$ ma non per $(z - z_0)^{m+1}$. Conseguenza di questi fatti è che un polinomio di grado n ha al più n radici, contate con le relative molteplicità. In campo complesso vale però anche il viceversa, cioè che le radici sono esattamente n , se contate con le loro molteplicità; vale infatti il seguente Teorema, che non dimostreremo.

Teorema 2.13 (Teorema fondamentale dell'algebra) *Ogni polinomio complesso di grado almeno 1 ammette una radice complessa.*

Il precedente Teorema ha come immediato corollario il seguente risultato.

Teorema 2.14 (Teorema fondamentale dell'algebra) *Ogni polinomio complesso di grado $n \geq 1$ ammette n radici complesse, se si conta ogni radice con la relativa molteplicità.*

I precedenti Teoremi non danno alcuna informazione su come trovare le radici del polinomio considerato; abbiamo però il seguente Teorema, valido per polinomi a coefficienti reali.

Proposizione 2.15 *Se p è un polinomio complesso a coefficienti reali, cioè $a_0, \dots, a_n \in \mathbb{R}$, allora $z_0 \in \mathbb{C}$ è radice di p se e solo se $\overline{z_0}$ lo è e in tal caso z_0 e $\overline{z_0}$ hanno la stessa molteplicità.*

DIM. Basta osservare che, essendo $p(z_0) = 0$, allora

$$\begin{aligned} 0 &= \overline{p(z_0)} = \overline{a_0 + a_1 z_0 + \cdots + a_n z_0^n} = \overline{a_0} + \overline{a_1} \cdot \overline{z_0} + \cdots + \overline{a_n} \cdot \overline{z_0}^n \\ &= a_0 + a_1 \overline{z_0} + \cdots + a_n \overline{z_0}^n = p(\overline{z_0}). \end{aligned}$$

□

Osservazione 2.16 Si noti che, come corollario, si deduce che ogni polinomio reale a coefficienti reali può essere scritto come prodotto di polinomi di grado uno o al massimo due; infatti, visto come polinomio complesso, si ha dal Teorema fondamentale dell'algebra 2.14 che

$$(2.5) \quad p(z) = a_n(z - z_1)^{m_1} \cdots (z - z_k)^{m_k}$$

con $m_1 + \cdots + m_k = n$; ne deriva che in campo complesso i fattori primi dei polinomi sono i polinomi di grado uno, nel senso che ogni polinomio si scompone come prodotto di binomi di primo grado (si pensi all'analogia della decomposizione dei numeri interi in fattori primi). Tornando a vedere il polinomio come reale a coefficienti reali, se z_i è reale, allora abbiamo un fattore di grado uno, mentre se z_i è non reale, allora in (2.5) deve comparire anche \bar{z}_i (Proposizione 2.15), esiste cioè $j \neq i$ tale che $z_j = \bar{z}_i$ e $m_j = m_i$. Notando poi che

$$(z - z_i)(z - \bar{z}_i) = z^2 - 2\operatorname{Re}(z_i)z + |z_i|^2$$

è un polinomio di secondo grado a coefficienti reali, segue l'osservazione.

Consideriamo ora il problema della radice n -esima di un numero complesso. Dato un numero complesso w , si dice che il numero complesso z è una radice n -esima di w se $z^n = w$. Le radici n -esime sono quindi le soluzioni dell'equazione $z^n - w = 0$, cioè sono le radici del polinomio $p(z) = z^n - w$; quindi, per quanto visto sopra, esistono al più n radici n -esime del numero w . La seguente Proposizione afferma che le radici sono esattamente n .

Proposizione 2.17 Dato il numero complesso $w \in \mathbb{C}$, $w \neq 0$, e $n \in \mathbb{N}$, $n \geq 2$, esistono n radici complesse distinte di w date dalla formula

$$z_k = \sqrt[n]{|w|} e^{i\vartheta_k}, \quad \text{con } \vartheta_k = \frac{\arg(w) + 2k\pi}{n}, k = 0, \dots, n-1.$$

Osservazione 2.18 Va osservato che la radice n -esima complessa non definisce una funzione in \mathbb{C} , avendo essa più valori; bisogna quindi fare attenzione che il simbolo $\sqrt[n]{w}$ può avere significati differenti, anche nel caso in cui w sia un numero reale, a seconda che si parli di radice reale o radice complessa.

DIM. Scrivendo $z = \varrho e^{i\vartheta}$ e $w = r e^{i\phi}$, si nota che $z^n = w$ se e solo se

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

tale sistema ha per soluzioni $\varrho = \sqrt[n]{r}$ (radice reale) e

$$(2.6) \quad \vartheta_k = \frac{\phi + 2k\pi}{n}, \quad k \in \mathbb{Z}.$$

Chiaramente abbiamo infiniti valori di ϑ_k ; notiamo però che due differenti valori di k e h definiscono lo stesso numero complesso, cioè $z_h = z_k$ se ϑ_h e ϑ_k differiscono per un multiplo di 2π , cioè se

$$\vartheta_h = \vartheta_k + 2m\pi;$$

la precedente espressione, usando (2.6), è equivalente a

$$h - k = nm,$$

cioè $z_h = z_k$ se e solo se $h - k$ è divisibile per n , o equivalentemente h e k hanno lo stesso resto nella divisione per n . Siccome i possibili resti della divisione per n sono $0, 1, \dots, n-1$, la dimostrazione segue. \square

Esempio 2.1 Consideriamo il numero $w = 1$, numero di modulo uno e argomento zero; nel caso $n = 2$, cioè nel caso delle radici quadrate, si ottengono i numeri

$$z_k = e^{i\vartheta_k}, \quad \vartheta_k = k\pi, k = 0, 1$$

cioè i due numeri $z_0 = 1$ e $z_1 = -1$. Per $n = 3$, si ottengono

$$z_k = e^{i\vartheta_k}, \quad \vartheta_k = \frac{2k\pi}{3}, k = 0, 1, 2$$

cioè i numeri complessi $z_0 = 1$, $z_1 = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ e $z_2 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$. Analogamente, per $n = 4$

$$z_k = e^{i\vartheta_k}, \quad \vartheta_k = \frac{k\pi}{2}, k = 0, 1, 2, 3$$

che producono i numeri $z_0 = 1$, $z_1 = i$, $z_2 = -1$ e $z_3 = -i$. In figura sono rappresentate queste radici e quelle relative al caso $n = 5$ e $n = 6$; si noti quindi che il numero 1 si trova sempre tra le radici per ogni ordine (come nel caso reale) ed anche il numero -1 si trova (sempre come nel caso reale) in ogni radice di ordine pari. Si hanno però altre radici a partire dal caso $n = 3$ che si distribuiscono nel piano complesso secondo triangoli equilateri ($n = 3$), quadrati ($n = 4$), pentagoni ($n = 5$) ed esagoni ($n = 6$), avendo sempre (siccome il numero w è reale) due radici coniugate tra loro (si veda figura 2.1).

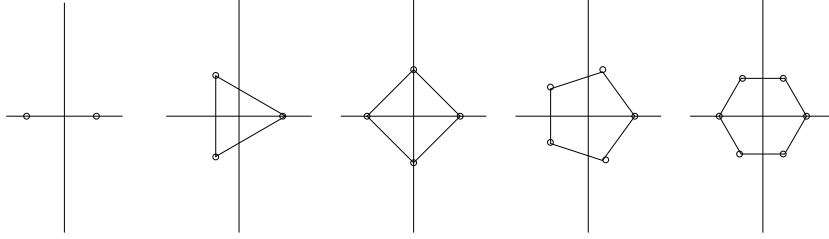


Figura 2.1: Distribuzione nel piano complesso delle radici n -esime di 1, $n = 2, 3, 4, 5, 6$.

Esempio 2.2 Nel caso $w = -1$, il numero di modulo uno e argomento π , si ottiene che nel

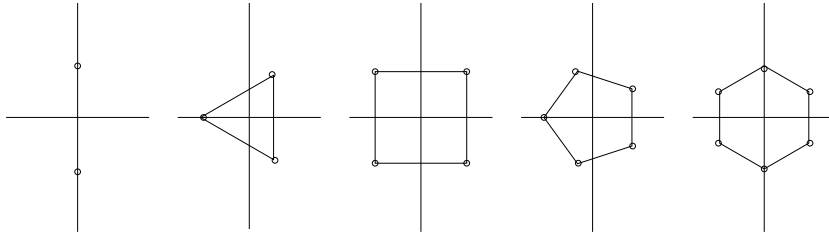


Figura 2.2: Distribuzione nel piano complesso delle radici n -esime di -1 , $n = 2, 3, 4, 5, 6$.

caso $n = 2$,

$$z_k = e^{i\vartheta_k}, \quad \vartheta_k = \frac{\pi}{2} + k\pi, k = 0, 1$$

cioè i due numeri $z_0 = i$ e $z_1 = -i$. Per $n = 3$, si ottengono

$$z_k = e^{i\vartheta_k}, \quad \vartheta_k = \frac{\pi}{3} + \frac{2k\pi}{3}, k = 0, 1, 2$$

cioè i numeri complessi $z_0 = \frac{1}{2} + i\frac{\sqrt{3}}{2}$, $z_1 = -1$ e $z_2 = \frac{1}{2} - i\frac{\sqrt{3}}{2}$. Analogamente, per $n = 4$

$$z_k = e^{i\vartheta_k}, \quad \vartheta_k = \frac{\pi}{4} + \frac{k\pi}{2}, k = 0, 1, 2, 3$$

che producono i numeri $z_0 = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$, $z_1 = -\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$, $z_2 = -\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}$ e $z_3 = \frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}$. Si veda la figura 2.2 per la distribuzione di queste radici e per quelle del caso $n = 5$ e $n = 6$. Come esercizio, si provi a vedere cosa succede nel caso $w = i$ e $w = -i$.

Capitolo 3

Equazioni differenziali

Ricordiamo brevemente alcune nozioni sulle equazioni algebriche; queste sono espressioni del tipo $F(x) = 0$ con $F : \mathbb{R} \rightarrow \mathbb{R}$ (oppure $F : \mathbb{C} \rightarrow \mathbb{C}$ nel caso di equazioni nei numeri complessi) con x incognita. Esempi sono dati dal problema della ricerca delle radici di un polinomio di secondo grado, che si traduce nell'equazione $ax^2 + bx + c = 0$, o anche equazioni del tipo $x + e^{x^2} = 0$; qualora si voglia sottolineare il fatto che nell'equazione la variabile x compare insieme alle sue potenze x^2, \dots, x^n , l'equazione si può anche scrivere nella forma $F(x, x^2, \dots, x^n) = 0$ e si parla in tal caso di equazione di grado n . Per soluzione dell'equazione si intende un numero $x_0 \in \mathbb{R}$ (o \mathbb{C} in caso di equazioni complesse) per il quale vale l'identità $F(x_0) = 0$, mentre per insieme delle soluzioni si intende l'insieme $S = \{x_0 \in \mathbb{R} : F(x_0) = 0\}$.

Un'equazione differenziale è una equazione del tipo $F(y(x)) = 0$, la cui incognita è una funzione $y(x)$ che appare nell'equazione assieme alle sue derivate. Si dice equazione differenziale di ordine n un'equazione del tipo $F(y(x), y'(x), \dots, y^{(n)}(x)) = 0$ o, se si vuole esplicitare anche la dipendenza dalla variabile x un'espressione del tipo

$$F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0$$

(la variabile x viene anche chiamata variabile indipendente, mentre la y viene detta variabile dipendente; solitamente al posto di x si usano anche le variabili t, s , ecc, mentre al posto di y si usano anche $y(t), y(s), x(t), x(s), u(x), u(t), u(s)$, ecc). La forma precedente si chiama forma implicita dell'equazione differenziale, mentre si parla di equazione in forma esplicita o in forma normale quando l'equazione differenziale si presenta nella forma $y^{(n)}(x) = f(x, y(x), \dots, y^{(n-1)}(x))$. Una soluzione sarà una funzione $u : I \rightarrow \mathbb{R}$ con I intervallo di \mathbb{R} e u derivabile n volte per la quale $F(x, u(x), u'(x), \dots, u^{(n)}(x)) = 0$ per ogni $x \in I$. Infine, per insieme delle soluzioni si intende l'insieme di tutte le funzioni n volte derivabili che sono soluzione dell'equazione differenziale e si dirà integrale generale di una equazione differenziale una espressione che riassume tutte le soluzioni.

Le applicazioni principali delle equazioni differenziali vengono dalla fisica, dalle scienze naturali, dalla biologia, dalla finanza, ecc. Ad esempio, una delle formule più importanti della fisica è data dall'equazione di Newton $F = ma$, dove in generale la forza sarà un campo vettoriale di forze \vec{F} e l'accelerazione è data dalla derivata seconda rispetto al tempo della funzione posizione $x(t)$. L'equazione di Newton si può quindi riscrivere $mx''(t) = F$, che è un'equazione differenziale del secondo ordine (la maggior parte delle equazioni provenienti

dalla fisica sono del secondo ordine), con la forza F che può essere in generale funzione del tempo t , della posizione $x(t)$ e della velocità $x'(t)$. Come casi particolari si hanno:

1. caduta di un grave in un campo gravitazionale costante $F = mg$, dove il problema diventa uni-dimensionale, moto lungo la verticale soggetto alla legge $x''(t) = g$; il suo integrale generale, ottenuto con una doppia integrazione, è dato da $x(t) = \frac{gt^2}{2} + c_1t + c_2$;
2. equazione dell'oscillatore armonico, in cui la forza F dipende dalla posizione:

$$F = -mkx(t),$$

dove $k > 0$ è la costante elastica; il moto è governato dall'equazione

$$x''(t) = -kx(t),$$

il cui integrale generale, come vedremo nel paragrafo 3.2.1, è dato da

$$x(t) = c_1 \cos \sqrt{k}t + c_2 \sin \sqrt{k}t;$$

3. equazione dell'oscillatore armonico smorzato, in cui compare anche la forza di attrito, proporzionale alla velocità, cioè $F = -mkx(t) - \mu x'(t)$; in tal caso l'equazione di Newton diventa $x''(t) = -kx(t) - \frac{\mu}{m}x'(t)$, il cui integrale generale è dato da

$$x(t) = c_1 e^{\frac{-\mu + \sqrt{\mu^2 - m^2 k^2}}{2m}t} + c_2 e^{\frac{-\mu - \sqrt{\mu^2 - m^2 k^2}}{2m}t}$$

se $\mu > mk$, altrimenti

$$x(t) = e^{-\frac{\mu}{2m}t} \left(c_1 \cos \frac{t}{2m} \sqrt{m^2 k^2 - \mu^2} + c_2 \sin \frac{t}{2m} \sqrt{m^2 k^2 - \mu^2} \right);$$

4. in generale il campo di forze potrà dipendere anche dal tempo, come ad esempio avviene per una particella carica che si muove in un campo elettro-magnetico variabile (fenomeni della risonanza magnetica).

Si noti che negli esempi presentati, nell'integrale generale ci sono sempre due costanti ad indicare che per l'insieme delle soluzioni ci sono sempre due gradi di libertà; questo, come vedremo, dipende dal fatto che stiamo considerando equazioni del secondo ordine.

3.1 Equazioni del primo ordine

Vediamo ora alcuni esempi di equazioni differenziali del primo ordine, cioè di equazioni del tipo $y'(x) = f(x, y(x))$. L'esempio più semplice è quello in cui $f(x, y(x)) = f(x)$ e l'equazione differenziale diventa

$$y'(x) = f(x),$$

cioè il problema diventa quello della ricerca delle primitive. Si sa, da quanto visto nel corso di Analisi 1 che se f è una funzione continua su di un intervallo I , allora la precedente equazione ha per soluzioni le funzioni

$$u(x) = \int f(x)dx + c$$

definita per tutti gli $x \in I$. Si noti che ci sono infinite soluzioni (una per ogni scelta di $c \in \mathbb{R}$); però, se si fissa un punto $x_0 \in I$ ed un valore $y_0 \in \mathbb{R}$, tra le precedenti soluzioni ne esiste una sola con la condizione $y(x_0) = y_0$, quella data da

$$u(x) = y_0 + \int_{x_0}^x f(t) dt.$$

Un altro esempio importante di equazione differenziale è l'equazione di Malthus, nota anche come equazione della dinamica delle popolazioni o dell'interesse bancario; se ad un dato istante t si è in possesso di un capitale $C(t)$, dopo h giorni tale capitale verrà incrementato di una percentuale p (l'interesse bancario) del capitale stesso moltiplicato per il numero di giorni h in cui il capitale resta depositato in banca, in formule:

$$C(t+h) = C(t) + pC(t)h.$$

Se tale incremento venisse calcolato istantaneamente, si avrebbe la possibilità di considerare il limite per $h \rightarrow 0$ ottenendo quindi l'equazione

$$pC(t) = \lim_{h \rightarrow 0} \frac{C(t+h) - C(t)}{h} = C'(t).$$

Per trovare soluzioni a tale equazioni, si nota anzitutto che $C(t) = 0$ è una soluzione, mentre se C è non nulla, si può dividere per $C(t)$ ed ottenere

$$p = \frac{C'(t)}{C(t)} = \frac{d}{dt} \ln |C(t)|,$$

da cui $|C(t)| = e^{pt+c} = e^c e^{pt} = ce^{pt}$ con l'ultima costante $c > 0$. Eliminando il valore assoluto si ottengono quindi le soluzioni $C(t) = ce^{pt}$ con $c \in \mathbb{R}$ (il valore 0 lo avevamo già considerato all'inizio della discussione). Si noti che anche in questo caso abbiamo infinite soluzioni, una per ogni scelta di $c \in \mathbb{R}$, ma nuovamente, se si fissa un istante $t_0 \in \mathbb{R}$ ed un capitale iniziale $C_0 \in \mathbb{R}$, allora esiste un'unica evoluzione del capitale che all'istante t_0 vale C_0 , ed è dato da $C(t) = C_0 e^{p(t-t_0)}$.

Risulta naturale dagli esempi precedenti aspettarsi unicità delle soluzioni solo quando viene fissato un valore, detto valore iniziale, in un fissato punto, detto punto iniziale. Abbiamo la seguente definizione.

Definizione 3.1 (Problema di Cauchy o ai valori iniziali) *Viene chiamato Problema di Cauchy il problema della ricerca delle soluzioni del sistema*

$$(3.1) \quad \begin{cases} y'(x) = f(x, y(x)) \\ y(x_0) = y_0 \end{cases}$$

o equivalentemente del problema

$$(3.2) \quad y(x) = y_0 + \int_{x_0}^x f(t, y(t)) dt.$$

Una soluzione del Problema di Cauchy è una funzione $u : I \rightarrow \mathbb{R}$ derivabile nell'intervallo I che risolve equivalentemente la (3.1) o la (3.2)

Il motivo per il quale cerchiamo soluzioni definite su intervalli può essere spiegato tramite il seguente esempio.

Esempio 3.1 La soluzione del Problema di Cauchy

$$\begin{cases} y' = \frac{1}{x^2} \\ y(1) = 0 \end{cases}$$

è dato da $u(x) = 1 - \frac{1}{x}$; però anche le funzioni

$$u(x) = \begin{cases} 1 - \frac{1}{x} & x > 0 \\ c - \frac{1}{x} & x < 0 \end{cases}$$

sono soluzioni e sono tante quante sono le costanti $c \in \mathbb{R}$. Rimane tuttavia unica se si cercano le soluzioni definite su intervalli I che contengono il dato iniziale, cioè per le quali $x_0 = 1 \in I$.

Abbiamo il seguente risultato, che fornisce l'esistenza e l'unicità delle soluzioni del Problema di Cauchy.

Teorema 3.2 (Esistenza e unicità per il Problema di Cauchy) Sia $f : I \times J \rightarrow \mathbb{R}$ una funzione con I e J intervalli di \mathbb{R} con la proprietà che

1. la funzione $x \mapsto f(x, y)$ è continua da I in \mathbb{R} per ogni $y \in J$;
2. esiste una costante $L > 0$ tale che per ogni $x \in I$ e per ogni $y_1, y_2 \in J$

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$$

(una funzione f con tale proprietà viene detta *Lipschitziana*).

Allora, per ogni $x_0 \in I$ e $y_0 \in J$ esiste un'unica funzione $u : I \rightarrow J$ soluzione del problema di Cauchy (3.1).

DIMOSTRAZIONE Dato che I è un intervallo, possiamo scrivere $I = [a, b]$ (nel caso I sia un intervallo aperto o un intervallo illimitato si ragionerà su ogni intervallo della forma $[a, b]$ contenuto in I). Si costruisce la successione di funzioni $u_0(x) = y_0$ e

$$u_n(x) = y_0 + \int_{x_0}^x f(t, u_{n-1}(t)) dt.$$

Per tale successione dimostriamo anzitutto che

$$|u_n(x) - u_{n-1}(x)| \leq M \frac{L^{n-1} |x - x_0|^n}{n!}$$

dove

$$M = \max_{x \in [a, b]} |f(x, y_0)|.$$

Si procede per induzione; la base è data da

$$|u_1(x) - u_0(x)| = \left| \int_{x_0}^x f(t, y_0) dt \right| \leq \int_{\min(x_0, x)}^{\max(x_0, x)} |f(t, y_0)| dt \leq M|x - x_0|.$$

Inoltre, supponendo

$$|u_n(x) - u_{n-1}(x)| \leq M \frac{L^{n-1}|x - x_0|^n}{n!}$$

possiamo dare una stima di $|u_{n+1}(x) - u_n(x)|$:

$$\begin{aligned} |u_{n+1}(x) - u_n(x)| &= \left| \int_{x_0}^x f(t, u_n(t)) - f(t, u_{n-1}(t)) dt \right| \\ &\leq \int_{\min(x_0, x)}^{\max(x_0, x)} |f(t, u_n(t)) - f(t, u_{n-1}(t))| dt \\ &\leq L \int_{\min(x_0, x)}^{\max(x_0, x)} |u_n(t) - u_{n-1}(t)| dt \\ &\leq L \int_{\min(x_0, x)}^{\max(x_0, x)} M \frac{L^{n-1}(t - x_0)^n}{n!} dt = M \frac{L^n |x - x_0|^{n+1}}{(n+1)!}. \end{aligned}$$

Si deduce quindi che la successione $(u_n(x))_n$ è una successione di Cauchy in quanto, supposto $n > m$

$$\begin{aligned} |u_n(x) - u_m(x)| &\leq |u_n(x) - u_{n-1}(x)| + |u_{n-1}(x) - u_m(x)| \leq \sum_{k=m}^{n-1} |u_{k+1}(x) - u_k(x)| \\ &\leq \sum_{k=m}^{n-1} M \frac{L^{k-1}|x - x_0|^k}{k!} \leq \frac{M}{L} \sum_{k=m}^{n-1} \frac{L^k(b-a)^k}{k!}. \end{aligned}$$

Essendo

$$\sum_{k=0}^{\infty} \frac{L^k(b-a)^k}{k!} = e^{L(b-a)},$$

siamo in presenza di una serie convergente, se ne deduce che la successione $(u_n(x))_n$ è di Cauchy e quindi converge ad una funzione $u(x)$. Si ha inoltre che

$$u(x) = \lim_{n \rightarrow \infty} u_n(x) = \lim_{n \rightarrow \infty} \left(y_0 + \int_{x_0}^x f(t, u_{n-1}(t)) dt \right) = y_0 + \int_{x_0}^x f(t, u(t)) dt.$$

Per dimostrare l'unicità delle soluzioni, si supponga esistano due funzioni y_1 e y_2 che soddisfano lo stesso problema di Cauchy e si consideri la differenza $g(x) = |y_1(x) - y_2(x)|$; si ottiene che, con $x > x_0$,

$$\begin{aligned} g(x) &= |y_1(x) - y_2(x)| \leq \int_{x_0}^x |f(t, y_1(t)) - f(t, y_2(t))| dt \leq L \int_{x_0}^x |y_1(t) - y_2(t)| dt \\ &= L \int_{x_0}^x g(t) dt. \end{aligned}$$

Dimostriamo quindi che una funzione che soddisfa la precedente relazione con $g(x_0) = 0$ è necessariamente la funzione nulla (questo risultato è noto come Lemma di Gronwall). La precedente espressione è equivalente a

$$L \geq \frac{g(x)}{\int_{x_0}^x g(t) dt} = \frac{d}{dx} \ln \int_{x_0}^x g(t) dt;$$

integrando tra $x_0 + \varepsilon$ e x generico (va considerato $\varepsilon > 0$ altrimenti avremmo il logaritmo di 0), si ottiene che

$$0 \leq \int_{x_0}^x g(t)dt \leq e^{L(x-x_0-\varepsilon)} \int_{x_0}^{x_0+\varepsilon} g(t)dt \rightarrow 0 \text{ per } \varepsilon \rightarrow 0$$

e quindi deve essere $g(t) = 0$ per ogni $t \in [x_0, x]$, da cui l'unicità. \square

Funzioni f che soddisfano le ipotesi del precedente teorema sono date da $f : I \times J \rightarrow \mathbb{R}$ tali che

- 1) $x \mapsto f(x, y)$ è continua in I per ogni $y \in J$;
- 2) $y \mapsto f(x, y)$ è continua con derivata prima continua in J per ogni $x \in I$.

Si può inoltre dimostrare, con una piccola modifica rispetto alla dimostrazione precedente, che se la condizione 2) viene sostituita con la condizione

- 2') $y \mapsto f(x, y)$ continua in J per ogni $x \in I$

allora le soluzioni esistono, ma in generale non è garantita l'unicità.

Esempio 3.2 Se si considera il problema di Cauchy

$$\begin{cases} y' = \frac{3}{2} \sqrt[3]{y} \\ y(0) = 0, \end{cases}$$

allora le due funzioni $y(x) = 0$ e $y(x) = x^{3/2}$ sono entrambe soluzioni. Si nota infatti che $f(x, y) = \frac{3}{2}y^{1/3}$ non dipende da x ed è continua rispetto ad y per $y \geq 0$, ma non derivabile per $y = 0$ e soprattutto la sua derivata rispetto ad y non è limitata per $y \rightarrow 0$.

3.1.1 Equazioni a variabili separabili

I due esempi considerati precedentemente, la ricerca delle primitive e l'equazione di Malthus, sono casi particolari di equazioni a variabili separabili, cioè equazioni in cui

$$y' = f(x, y) = a(x)b(y).$$

L'esistenza e l'unicità per il Problema di Cauchy si hanno come caso particolare del teorema (3.2), quando $a : I \rightarrow \mathbb{R}$ è continua e $b : J \rightarrow \mathbb{R}$ è Lipschitziana, cioè esiste $L > 0$ tale che $|b(y_1) - b(y_2)| \leq L|y_1 - y_2|$; tale condizione è ad esempio soddisfatta quando b è derivabile con derivata continua, avendosi come L il massimo del modulo della derivata di b .

Per ricavare le soluzioni di una equazione a variabili separabili si procede come segue;

1. si cercano anzitutto gli zeri di b ; infatti, se $y_0 \in J$ è tale che $b(y_0) = 0$, allora la funzione costante $u(x) = y_0$ è una soluzione;
2. sotto la condizione $b(y) \neq 0$, si divide per $b(y)$ ottenendo

$$\frac{y'(x)}{b(y(x))} = a(x).$$

Quindi, se A denota una primitiva di a e B una primitiva di $\frac{1}{b}$, con una integrazione si ottiene

$$B(y(x)) = A(x) + c$$

con $c \in \mathbb{R}$. Tale formula fornisce la soluzione generale in forma implicita; se B è invertibile e se la sua inversa è nota, si può ottenere la soluzione in forma esplicita tramite

$$y(x) = B^{-1}(A(x) + c).$$

Esempio 3.3 Si consideri l'equazione $y' = 2x\sqrt{1-y^2}$; $y(x) = \pm 1$ sono soluzioni; nel caso $y \neq \pm 1$, si ottiene

$$\frac{y'}{\sqrt{1-y^2}} = 2x$$

da cui

$$\arcsen(y(x)) = x^2 + c.$$

Notiamo che, visto che il codominio di \arcsen è l'intervallo $[-\pi/2, \pi/2]$, la precedente espressione ha senso fintanto che $x^2 + c \in [-\pi/2, \pi/2]$; in particolare, la costante c , nel caso di un Problema di Cauchy, andrà scelta in modo che $x_0^2 + c \in [-\pi/2, \pi/2]$. Invertendo la precedente espressione, si trova che la soluzione generale è data da $y(x) = \sin(x^2 + c)$. Se si vuole ad esempio risolvere il seguente problema

$$\begin{cases} y' = 2x\sqrt{1-y^2} \\ y(\sqrt{\pi}) = \frac{1}{2}, \end{cases}$$

allora la soluzione è data da $u(x) = \sin(x^2 - 5\pi/6)$, e tale soluzione è definita per

$$x \in [\sqrt{\frac{\pi}{3}}, 2\sqrt{\frac{\pi}{3}}].$$

Esercizio 3.1 Si risolva il seguente Problema di Cauchy

$$\begin{cases} yy' = 1 \\ y(0) = 2. \end{cases}$$

Esercizio 3.2 Si trovi la soluzione generale dell'equazione $y' = ay(1 - by)$ con a e b due numeri reali positivi. In particolare, si risolva l'equazione di Verhulst $y' = \varepsilon y(1 - \frac{y}{k})$

3.1.2 Equazioni lineari del primo ordine

Per equazione differenziale lineare del primo ordine si intende un'equazione della forma

$$y'(x) + a(x)y(x) = f(x);$$

il teorema 3.2 garantisce l'esistenza e l'unicità delle soluzioni nel caso in cui $a, f : I \rightarrow \mathbb{R}$ sono funzioni continue (in questo caso il valore iniziale y_0 può essere un qualsiasi numero reale). Nel caso in cui $f = 0$ si parla di equazione omogenea, mentre con $f \neq 0$ si parla di equazione completa e l'equazione $y' + ay = 0$ si chiama omogenea associata. Abbiamo il seguente risultato.

Teorema 3.3 *La soluzione generale dell'equazione completa è data dalla soluzione generale dell'omogenea associata a cui va aggiunta una soluzione particolare. Inoltre, se a e f sono funzioni continue, per ogni $x_0 \in I$ e $y_0 \in \mathbb{R}$, l'unica soluzione del Problema di Cauchy con dato iniziale in x_0 e valore iniziale y_0 è data da*

$$y(x) = \left(y_0 + \int_{x_0}^x e^{A(t)} f(t) dt \right) e^{-A(x)}$$

dove A è la primitiva di a data da

$$A(x) = \int_{x_0}^x a(t) dt.$$

L'integrale generale infine è dato dall'espressione

$$y(x) = e^{A(x)} \left(c + \int e^{-A(x)} f(x) dx \right),$$

essendo $A(x)$ una qualsiasi primitiva di $a(x)$.

DIMOSTRAZIONE L'esistenza e l'unicità per il Problema di Cauchy segue dal Teorema 3.2 in quanto

$$f(x, y) = f(x) - a(x)y$$

e la continuità in x segue dalla continuità di $f(x)$ e $a(x)$, mentre la funzione $y \mapsto f(x) - a(x)y$ è derivabile con derivata $a(x)$. Se u è la soluzione generale e u_p è una soluzione particolare, allora $u_0 = u - u_p$ soddisfa le condizioni

$$u'_0 + au_0 = u' + au - u'_p - au_p = f - f = 0,$$

cioè u_0 è soluzione dell'omogenea; viceversa, nello stesso modo si dimostra che se u_0 è soluzione dell'omogenea e u_p una soluzione particolare, allora $u = u_0 + u_p$ è soluzione dell'equazione completa. Quindi, per trovare la soluzione generale, si può prima cercare la soluzione generale dell'omogenea e poi cercare una soluzione particolare. Per risolvere l'omogenea associata, moltiplichiamo l'equazione per $e^{A(x)}$ dove A è una primitiva di a , per ottenere

$$0 = e^A y' + a e^A y = \frac{d}{dx}(e^A y)$$

e quindi $e^A y$ è una funzione costante, cioè

$$y(x) = c e^{-A(x)}, \quad c \in \mathbb{R}.$$

Si noti che questa espressione ci dice che l'insieme delle soluzioni dell'equazione omogenea è uno spazio vettoriale di dimensione uno la cui base è data dalla funzione $e^{-A(x)}$.

Per trovare una soluzione particolare, si può utilizzare il metodo detto della variazione delle costanti. Tale metodo consiste nel partire dalla soluzione generale dell'omogenea associata e considerare c non come una costante, ma come una funzione $c(x)$. Cioè, consideriamo

$$y(x) = c(x) e^{-A(x)}$$

in cui ora l'incognita è la funzione $c(x)$. Derivando otteniamo

$$y' = c' e^{-A} - c a e^{-A};$$

se si impone che tale funzione risolva l'equazione completa si ottiene

$$f = y' + ay = c'e^{-A}$$

da cui $c' = fe^A$, cioè, integrando, $c(x) = c + \int e^{A(x)} f(x) dx$. \square

Esercizio 3.3 Si risolva il seguente Problema di Cauchy

$$\begin{cases} y' + \frac{2}{x}y = \frac{1}{x^2} \\ y(-1) = 2. \end{cases}$$

Esercizio 3.4 Si risolva l'equazione che descrive la carica di un condensatore C in un circuito con resistenza R sottoposto ad un campo E

$$CRy' + y = E.$$

Esercizio 3.5 Si risolva l'equazione $y' + ay = f$ con $a > 0$ costante.

3.1.3 Equazioni di Bernoulli

Le equazioni di Bernoulli sono del tipo

$$y' + ay = fy^\alpha$$

con $\alpha \in \mathbb{R}$, $\alpha \neq 0, 1$ (in questi casi siamo ricondotti ad una semplice equazione lineare). Questo tipo di equazioni si può ricondurre ad una equazione lineare; infatti, se si divide per y^α si ottiene

$$y^{-\alpha}y' + ay^{1-\alpha} = f$$

e quindi, ponendo $v = y^{1-\alpha}$ si giunge all'equazione

$$\frac{1}{1-\alpha}v' + av = f.$$

Esempio 3.4 Risolviamo il seguente Problema di Cauchy

$$\begin{cases} y' = xy + x\sqrt{y} \\ y(1) = 1. \end{cases}$$

3.2 Equazioni lineari del secondo ordine

Le equazioni lineari del secondo ordine sono del tipo

$$y''(x) + a(x)y'(x) + b(x)y(x) = f(x)$$

Esempi importanti vengono dalla Fisica e dall'equazione fondamentale della dinamica

$$F = ma;$$

in particolare, la caduta di un grave soggetto alla forza di gravità soddisfa l'equazione $y'' = g$, l'equazione dell'oscillatore armonico è dato da $my'' = -ky$, l'oscillatore armonico con attrito

è dato da $my'' + hy' + ky = 0$ mentre l'oscillatore armonico con attrito e con termine forzante è dato da $my'' + hy' + ky = f$.

Analogamente a quanto detto nel caso di equazione lineare del primo ordine, si dice che l'equazione è omogenea se $f = 0$, completa se $f \neq 0$ mentre nel caso $f \neq 0$ l'equazione $y'' + ay' + by = 0$ si chiama omogenea associata. Il problema di Cauchy nel caso di equazione del secondo ordine si traduce nella determinazione della soluzione di

$$(3.3) \quad \begin{cases} y'' + ay' + by = f \\ y(x_0) = y_0 \\ y'(x_0) = y_1 \end{cases}$$

dove ancora x_0 viene detto dato iniziale e i valori y_0 e y_1 valori iniziali. Abbiamo il seguente risultato.

Teorema 3.4 *Siano $a, b, f : I \rightarrow \mathbb{R}$ funzioni continue con I intervallo di \mathbb{R} e sia $x_0 \in I$: allora, per ogni scelta di $y_0, y_1 \in \mathbb{R}$, il Problema di Cauchy (3.3) ammette una ed una sola soluzione.*

DIMOSTRAZIONE Notiamo anzitutto che un'equazione del secondo ordine può essere ricondotta ad un sistema di equazioni del primo ordine ponendo $V(x) = (V_1(x), V_2(x))$, $V_1(x) = y(x)$, $V_2(x) = y'(x)$. In questo modo il Problema di Cauchy diventa

$$\begin{cases} V'(x) + \vec{a}(x)V(x) + \vec{f}(x) \\ V(0) = V_0 \end{cases}$$

dove

$$\vec{a}(x) = \begin{pmatrix} 0 & -1 \\ b(x) & a(x) \end{pmatrix}, \quad \vec{f}(x) = \begin{pmatrix} 0 \\ f(x) \end{pmatrix}, \quad V_0 = V(x_0) \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}.$$

Quindi si ottiene una equazione lineare del primo ordine nell'incognita $V(x)$, la cui soluzione, così come nel caso di equazione lineare del primo ordine vista nella Sezione 3.1.2, è data dalla formula

$$V(x) = \left(V_0 + \int_{x_0}^x e^{\vec{A}(t)} \vec{f}(t) dt \right) e^{-\vec{A}(x)}$$

con

$$\vec{A}(x) = \begin{pmatrix} 0 & x_0 - x \\ B(x) & A(x) \end{pmatrix}, \quad A(x) = \int_{x_0}^x a(t) dt, B(x) = \int_{x_0}^x b(t) dt$$

e l'esponenziale di una matrice è la matrice data da

$$e^{\vec{A}(x)} = \sum_{k=0}^{\infty} \frac{\vec{A}(x)^k}{k!};$$

si può dimostrare che l'ipotesi di continuità fatta su a e b implica la convergenza della precedente serie. \square

Osservazione 3.5 Nel Teorema precedente abbiamo formalmente una espressione che ci dice come calcolare esplicitamente la soluzione, date le funzioni a , b ed f ed i dati iniziali x_0 , y_0 e y_1 . Il problema fondamentale della precedente formula sta nella difficoltà, in generale, del calcolo esplicito della funzione $e^{\vec{A}(x)}$. Vedremo in seguito che tale calcolo si potrà eseguire sostanzialmente solo quando a e b sono funzioni costanti.

Vediamo ora di studiare le proprietà dell'insieme delle soluzioni di una equazione differenziale lineare del secondo ordine.

Teorema 3.6 *La soluzione generale di una equazione differenziale del secondo ordine è data dalla somma dell'integrale generale dell'equazione omogenea associata e di una soluzione particolare. Inoltre, lo spazio delle soluzioni della omogenea associata è un sottospazio vettoriale di $C^2(I)$ di dimensione 2.*

DIMOSTRAZIONE Come abbiamo già visto per le equazioni differenziali lineari del primo ordine, se y è una soluzione generale e y_p è una soluzione particolare, allora $y_0 = y - y_p$ è soluzione dell'omogenea associata e viceversa, se y_0 è soluzione dell'omogenea e y_p è soluzione particolare, $y = y_0 + y_p$ è soluzione.

Per vedere che l'insieme delle soluzioni dell'omogenea associata è uno spazio vettoriale di dimensione 2 basta fornire una base costituita da due elementi, cioè basta trovare due funzioni linearmente indipendenti u_1 e u_2 tali che per ogni altra soluzione y esistano due numeri $c_1, c_2 \in \mathbb{R}$ tali $y(x) = c_1 u_1(x) + c_2 u_2(x)$. Consideriamo le due funzioni u_1 e u_2 soluzioni dei seguenti problemi di Cauchy

$$\begin{cases} y'' + ay' + by = 0 \\ y(x_0) = 1 \\ y'(x_0) = 0 \end{cases} \quad \begin{cases} y'' + ay' + by = 0 \\ y(x_0) = 0 \\ y'(x_0) = 1 \end{cases}$$

Il teorema di esistenza e unicità garantisce che u_1 ed u_2 esistono e sono uniche. Esse sono linearmente indipendenti in quanto se esistessero $c_1, c_2 \in \mathbb{R}$ tali che

$$u(x) = c_1 u_1(x) + c_2 u_2(x) = 0,$$

valutando in x_0 tale espressione si otterrebbe $0 = c_1 u_1(x_0) + c_2 u_2(x_0) = c_1$, cioè $c_1 = 0$. Inoltre, siccome si ha anche che $0 = u'(x) = c_1 u_1'(x) + c_2 u_2'(x)$, si otterrebbe anche che $0 = u'(x_0) = c_2$, cioè $c_2 = 0$. Presa ora una soluzione generalizzata $y(x)$ e scegliendo $c_1 = y(x_0)$ e $c_2 = y'(x_0)$, abbiamo che la funzione $u(x) = c_1 u_1(x) + c_2 u_2(x)$ ha la proprietà che $u(x_0) = c_1 = y(x_0)$ e $u'(x_0) = c_2 = y'(x_0)$, cioè sia u che y risolvono lo stesso problema di Cauchy, quindi per l'unicità delle soluzioni si deve avere $y(x) = u(x) = c_1 u_1(x) + c_2 u_2(x)$. \square

Come detto in precedenza, il calcolo esplicito delle soluzioni è spesso un problema difficile e fattibile solo in casi molto particolari, come nel seguente esempio.

Esempio 3.5 Trovare l'integrale generale dell'equazione

$$x^2 y''(x) - 3xy'(x) + 3y(x) = 0.$$

In questo caso infatti, quando $x \neq 0$, si ottiene che

$$y''(x) = \frac{3}{x} y'(x) - \frac{3}{x^2} y(x) = \frac{d}{dx} \left(\frac{3}{x} y(x) \right)$$

da cui $y'(x) = \frac{3}{x} y(x) + c_1$. Questa è poi una equazione differenziale lineare del primo ordine con $a(x) = -\frac{3}{x}$ e $f(x) = c_1$, la cui soluzione è data da

$$\begin{aligned} y(x) &= \left(c_2 + \int \frac{c_1}{|x|^3} dx \right) |x|^3 = \left(c_2 + \frac{x}{|x|} \int \frac{c_1}{x^3} dx \right) |x|^3 \\ &= -\frac{c_1 x}{2} + c_2 |x|^3. \end{aligned}$$

3.2.1 Equazioni a coefficienti costanti

È questo il caso in cui a e b sono costanti; cerchiamo prima di tutto le soluzioni dell'equazione omogenea associata. Il vantaggio in questo caso sta nel fatto che il problema può essere ricondotto ad un problema algebrico; infatti, se consideriamo le funzioni $y_\lambda(x) = e^{\lambda x}$, allora y_λ è soluzione dell'equazione omogenea se e solo se $\lambda^2 + a\lambda + b = 0$, cioè se e solo se λ è radice del polinomio $P(x) = x^2 + ax + b$, detto polinomio caratteristico. Siccome si riescono sempre a trovare due radici distinte, almeno nel caso di discriminante diverso da zero, la questione diventa quindi se dati due distinti λ_1 e λ_2 , le funzioni $e^{\lambda_1 x}$ ed $e^{\lambda_2 x}$ siano linearmente indipendenti o meno.

Proposizione 3.7 *Dati $\lambda_1, \lambda_2 \in \mathbb{C}$, le funzioni $e^{\lambda_1 x}$ ed $e^{\lambda_2 x}$ sono linearmente indipendenti per combinazioni complesse. In particolare, nel caso di combinazioni reali, se $\alpha, \beta, c \in \mathbb{R}$ con $\alpha \neq \beta$, le seguenti funzioni sono linearmente indipendenti:*

1. $e^{\alpha x}$ ed $e^{\beta x}$;
2. $\sin(\alpha x)$ e $\sin(\beta x)$, $\cos(\alpha x)$ e $\cos(\beta x)$, $e^{cx} \sin(\alpha x)$ e $e^{cx} \cos(\alpha x)$;
3. x^α e x^β così come $e^{cx} x^\alpha$, $e^{cx} x^\beta$, $\sin(cx) x^\alpha$ e $\sin(cx) x^\beta$.

DIMOSTRAZIONE La dimostrazione segue semplicemente assumendo che esistano $c_1, c_2 \in \mathbb{C}$ tali che

$$u(x) := c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} = 0$$

Siccome $u(0) = c_1 + c_2$, se ne deduce che $c_1 + c_2 = 0$. Inoltre, dato che $u(x) = 0$, anche $u'(x) = 0$, da cui $u'(0) = c_1 \lambda_1 + c_2 \lambda_2 = 0$: abbiamo quindi ottenuto il sistema

$$\begin{cases} c_1 + c_2 = 0 \\ c_1 \lambda_1 + c_2 \lambda_2 = 0 \end{cases}$$

la cui unica soluzione, dato che $\lambda_1 \neq \lambda_2$, è data da $c_1 = c_2 = 0$. Per la restante parte della proposizione, basta per il primo punto considerare $\lambda_1 = \alpha, \lambda_2 = \beta \in \mathbb{R}$ e combinazioni reali, $c_1, c_2 \in \mathbb{R}$, mentre gli altri punti seguono dall'osservazione che se y_1 ed y_2 sono linearmente indipendenti, allora lo sono anche $v_1 = y_1 + y_2$ e $v_2 = y_1 - y_2$. Usando questa osservazione, si considera $\lambda_1 = \alpha + i\beta$ e $\lambda_2 = \alpha - i\beta$ e si ottiene che

$$e^{\alpha x} \cos(\beta x) = \frac{e^{\lambda_1 x} + e^{\lambda_2 x}}{2}, \quad e^{\alpha x} \sin(\beta x) = \frac{e^{\lambda_1 x} - e^{\lambda_2 x}}{2i}.$$

L'ultima parte riguarda l'indipendenza lineare di x^α e x^β ; ma se $u(x) = c_1 x^\alpha + c_2 x^\beta = 0$, allora $0 = u(1) = c_1 + c_2$, così come $0 = u'(1) = c_1 \alpha + c_2 \beta$, e quindi ancora $c_1 = c_2 = 0$. \square

Il precedente risultato può essere usato, assieme a quanto visto per i polinomi in campo complesso, per trovare tutte le soluzioni dell'equazione omogenea. Infatti, per quanto visto, il polinomio caratteristico è dato da $x^2 + ax + b$ con $a, b \in \mathbb{R}$, quindi un polinomio a coefficienti reali. Tale polinomio può avere due radici reali $\lambda_1, \lambda_2 \in \mathbb{R}$, da cui il fatto che $e^{\lambda_1 x}$ ed $e^{\lambda_2 x}$ sono due soluzioni linearmente indipendenti, oppure due radici complesse coniugate tra loro, $\lambda_1 = \alpha + i\beta$ e $\lambda_2 = \alpha - i\beta$, da cui $e^{\alpha x} \cos(\beta x)$ ed $e^{\alpha x} \sin(\beta x)$ sono due soluzioni linearmente indipendenti, oppure ancora una sola radice reale λ di molteplicità due, nel qual caso si ottiene che $e^{\lambda x}$ ed $x e^{\lambda x}$ sono due soluzioni linearmente indipendenti. Abbiamo così dimostrato il seguente risultato.

Teorema 3.8 *Data l'equazione differenziale del secondo ordine lineare a coefficienti costanti omogenea*

$$y'' + ay' + by = 0,$$

allora

1. se $a^2 - 4b > 0$, il polinomio caratteristico ammette due radici reali $\lambda_1, \lambda_2 \in \mathbb{R}$ e la soluzione generale è data da

$$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}, \quad c_1, c_2 \in \mathbb{R};$$

2. se $a^2 - 4b < 0$, il polinomio caratteristico ha due radici complesse coniugate tra loro $\lambda_1 = \alpha + i\beta$ e $\lambda_2 = \alpha - i\beta$ e la soluzione generale è data da

$$y(x) = e^{\alpha x} (c_1 \cos(\beta x) + c_2 \sin(\beta x)), \quad c_1, c_2 \in \mathbb{R};$$

3. se $a^2 - 4b = 0$, il polinomio caratteristico ha una sola radice reale λ con molteplicità due e la soluzione generale è data da

$$y(x) = c_1 e^{\lambda x} + c_2 x e^{\lambda x}, \quad c_1, c_2 \in \mathbb{R}.$$

Esercizio 3.6 Risolvere le seguenti equazioni differenziali

1. $y'' - \omega^2 y = 0$;
2. oscillatore armonico $y'' + \omega^2 y = 0$;
3. oscillatore armonico con attrito $y'' - 2\omega y' + \omega^2 y = 0$;
4. con dati iniziali

$$\begin{cases} y'' + 2y' + 3y = 0 \\ y(0) = 1, \quad y'(0) = 2. \end{cases}$$

Vediamo ora come si risolve un'equazione differenziale del secondo ordine completa, cioè come si fa a trovare la soluzione particolare. Si hanno sostanzialmente due metodi.

Metodo per somiglianza Questo metodo si può applicare quando il termine forzante f assume forme particolari, più precisamente quando

$$f(x) = e^{\alpha x} (p_n(x) \cos(\beta x) + q_m(x) \sin(\beta x))$$

con $\alpha, \beta \in \mathbb{R}$ e p_n, q_m due polinomi di grado rispettivamente n ed m . Prima di tutto si considera il numero complesso $\lambda = \alpha + i\beta$ e si pone h uguale alla molteplicità di λ come radice del polinomio caratteristico; in particolare $h = 0$ se $P(\lambda) \neq 0$, $h = 1$ se $P(\lambda) = 0$ ma P ammette una seconda radice diversa da λ , mentre $h = 2$ se $P(x) = (x - \lambda)^2$. Si considera poi $k = \max(n, m)$ (si prende cioè il massimo tra i gradi di p_n e q_m) e si cerca la soluzione particolare nella forma

$$y_p(x) = x^h e^{\alpha x} (\bar{p}_k(x) \cos(\beta x) + \bar{q}_k(x) \sin(\beta x))$$

con \bar{p}_k e \bar{q}_k polinomio da determinare, imponendo che y_p sia una soluzione dell'equazione completa.

Esempio 3.6 Trovare le soluzioni delle seguenti equazioni

1. $y'' + 3y = x + 2 \cos x$;
2. $y'' - \omega^2 y = 1 + x^2$;
3. $y'' + 2y' = x$;
4. $y'' + 2y' + 3y = 2e^{3x}$;
5. $y'' + 2y' - 3y = 2e^{-3x}$;
6. $y'' + 2y' - y = 2e^x \cos(3x)$;
7. $2y'' + y' + 2y = 3 \sin(2x)$.

Vediamo solo la prima equazione; la soluzione generale dell'omogenea associata è data da

$$y_0(x) = c_1 \cos(x\sqrt{3}) + c_2 \sin(x\sqrt{3}),$$

ricavata utilizzando il polinomio caratteristico

$$p(\lambda) = \lambda^2 + 3$$

le cui radici sono $\pm i\sqrt{3}$. Per la soluzione particolare, usiamo un principio detto *principio di sovrapposizione*; siccome il termine forzante $f(x) = x + 2 \cos x$ non è in forma particolare, ma somma di due funzioni $f_1(x) = x$ e $f_2(x) = 2 \cos x$, entrambe in forma particolare, basterà trovare due soluzioni particolari y_1 ed y_2 associate rispettivamente a f_1 e f_2 , per ottenere che la somma $y_p = y_1 + y_2$ sia una soluzione particolare associata a $f = f_1 + f_2$. Nel caso di f_1 , una soluzione particolare va cercata nella forma

$$y_1(x) = (ax + b)$$

in quanto

$$f_1(x) = x = e^{0x}(x \cos(0x) + 0 \sin(0x)),$$

cioè $\alpha = \beta = 0$, $p_1(x) = x$ polinomio di primo grado, $q_0(x) = 0$ polinomio di grado zero e il numero complesso $\lambda = \alpha + i\beta = 0$ non è radice del polinomio caratteristico $p(\lambda)$, cioè $h = 0$. Imponendo $y_1'' + 3y_1 = x$ si ricava $a = 1$ e $b = 0$, cioè $y_1(x) = x$. La seconda soluzione particolare va cercata nella forma

$$y_2(x) = a \cos x + b \sin x$$

in quanto il termine forzante $f_2(x) = 2 \cos x$ ha $\alpha = 0$, $\beta = 1$, $p_0(x) = 2$ e $q_0(x) = 0$ polinomi di grado zero e $\lambda = \alpha + i\beta = i$ non è radice del polinomio caratteristico. Imponendo $y_2'' + 3y_2 = 2 \cos x$ si ricava $a = 1$ e $b = 0$. In definitiva, la soluzione generale dell'equazione differenziale è data da

$$y(x) = c_1 \cos(x\sqrt{3}) + c_2 \sin(x\sqrt{3}) + x + \cos x.$$

Metodo della variazione delle costanti Il metodo è lo stesso visto nel caso dell'equazione lineare del primo ordine; si parte dalla soluzione generale dell'omogenea

$$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}$$

e si considerano $c_1 = c_1(x)$ e $c_2 = c_2(x)$ come funzioni. Derivando si ottiene

$$y'(x) = c_1'(x)e^{\lambda_1 x} + c_2'(x)e^{\lambda_2 x} + c_1(x)\lambda_1 e^{\lambda_1 x} + c_2(x)\lambda_2 e^{\lambda_2 x}$$

e al primo passo si pone

$$c_1'(x)e^{\lambda_1 x} + c_2'(x)e^{\lambda_2 x} = 0.$$

In questo modo la derivata seconda diventa

$$y''(x) = c_1'(x)\lambda_1 e^{\lambda_1 x} + c_2'(x)\lambda_2 e^{\lambda_2 x} + c_1(x)\lambda_1^2 e^{\lambda_1 x} + c_2(x)\lambda_2^2 e^{\lambda_2 x}$$

Imponendo quindi che $y'' + ay' + by = f$ e sapendo che λ_1 e λ_2 sono radici del polinomio caratteristico, si giunge alla condizione

$$c_1'(x)\lambda_1 e^{\lambda_1 x} + c_2'(x)\lambda_2 e^{\lambda_2 x} = f(x).$$

Si tratta quindi di risolvere il sistema, nelle incognite c_1' e c_2'

$$\begin{cases} c_1'(x)e^{\lambda_1 x} + c_2'(x)e^{\lambda_2 x} = 0 \\ c_1'(x)\lambda_1 e^{\lambda_1 x} + c_2'(x)\lambda_2 e^{\lambda_2 x} = f(x). \end{cases}$$

Esercizio 3.7 Risolvere le seguenti equazioni;

1. oscillatore armonico $my'' = -ky$;
2. oscillatore armonico smorzato $y'' + 2\delta y' + \omega^2 y = 0$;
3. $y'' + \omega^2 y = \tan x$.

3.2.2 Oscillatore forzato

In questo paragrafo studiamo l'oscillatore armonico forzato, cioè risolviamo la seguente equazione differenziale

$$y'' + \omega^2 y = B \cos(\omega_0 x).$$

La soluzione generale dell'omogenea associata è data da

$$y(x) = c_1 \cos(\omega x) + c_2 \sin(\omega x).$$

Per la soluzione completa, utilizziamo il metodo della variazione delle costanti; partiamo dalla funzione

$$y(x) = c_1(x) \cos(\omega x) + c_2(x) \sin(\omega x)$$

da cui, derivando

$$y'(x) = c_1'(x) \cos(\omega x) + c_2'(x) \sin(\omega x) - c_1(x)\omega \sin(\omega x) + c_2(x)\omega \cos(\omega x);$$

imponiamo quindi la condizione

$$c_1'(x) \cos(\omega x) + c_2'(x) \sin(\omega x) = 0$$

e deriviamo una seconda volta per ottenere

$$y''(x) = -c'_1(x)\omega \operatorname{sen}(\omega x) + c'_2(x)\omega \cos(\omega x) - c_1(x)\omega^2 \cos(\omega x) - c_2(x)\omega^2 \operatorname{sen}(\omega x).$$

Richiedendo quindi che $y'' + \omega^2 y = B \cos(\omega_0 x)$, si perviene alla condizione

$$-c'_1(x)\omega \operatorname{sen}(\omega x) + c'_2(x)\omega \cos(\omega x) = B \cos(\omega_0 x).$$

Dobbiamo risolvere quindi il sistema

$$\begin{cases} c'_1(x) \cos(\omega x) + c'_2(x) \operatorname{sen}(\omega x) = 0 \\ -c'_1(x)\omega \operatorname{sen}(\omega x) + c'_2(x)\omega \cos(\omega x) = B \cos(\omega_0 x), \end{cases}$$

che ha come soluzioni

$$\begin{cases} c'_1(x) = -\frac{B}{\omega} \cos(\omega_0 x) \operatorname{sen}(\omega x) = -\frac{B}{2\omega} (\operatorname{sen}((\omega_0 + \omega)x) + \operatorname{sen}((\omega - \omega_0)x)) \\ c'_2(x) = \frac{B}{\omega} \cos(\omega_0 x) \cos(\omega x) = \frac{B}{2\omega} (\cos((\omega_0 + \omega)x) + \cos((\omega - \omega_0)x)) \end{cases}$$

Si distinguono quindi due casi, $\omega_0 = \omega$ e $\omega_0 \neq \omega$.

Caso $\omega_0 \neq \omega$. si ottiene

$$\begin{cases} c_1(x) = c_1 + \frac{B}{2\omega} \left(\frac{\cos((\omega_0 + \omega)x)}{\omega_0 + \omega} + \frac{\cos((\omega - \omega_0)x)}{\omega - \omega_0} \right) \\ c_2(x) = c_2 + \frac{B}{2\omega} \left(\frac{\operatorname{sen}((\omega_0 + \omega)x)}{\omega_0 + \omega} + \frac{\operatorname{sen}((\omega - \omega_0)x)}{\omega - \omega_0} \right) \end{cases}$$

Avremo quindi che la soluzione generale sarà data da

$$\begin{aligned} y(x) = & c_1 \cos(\omega x) + c_2 \operatorname{sen}(\omega x) + \\ & - \frac{B}{\omega} \left(\frac{\operatorname{sen}((\omega_0 + \omega)x)}{\omega_0 + \omega} + \frac{\operatorname{sen}((\omega - \omega_0)x)}{\omega - \omega_0} \right) \cos(\omega x) + \\ & + \frac{B}{\omega} \left(\frac{\cos((\omega_0 + \omega)x)}{\omega_0 + \omega} + \frac{\cos((\omega - \omega_0)x)}{\omega - \omega_0} \right) \operatorname{sen}(\omega x) \end{aligned}$$

Se tra queste funzioni cerchiamo la soluzione del Problema di Cauchy

$$\begin{cases} y'' + \omega^2 y = B \cos(\omega_0 x) \\ y(0) = y'(0) = 0 \end{cases}$$

si ottiene la funzione

$$y(x) = \frac{B}{\omega_0^2 - \omega^2} (\cos(\omega x) - \cos(\omega_0 x)) = \frac{2B}{\omega_0^2 - \omega^2} \operatorname{sen} \left(\frac{\omega_0 + \omega}{2} x \right) \operatorname{sen} \left(\frac{\omega_0 - \omega}{2} x \right)$$

il cui grafico è riportato in figura 3.1; si parla per tale soluzione di *fenomeno dei battimenti*.

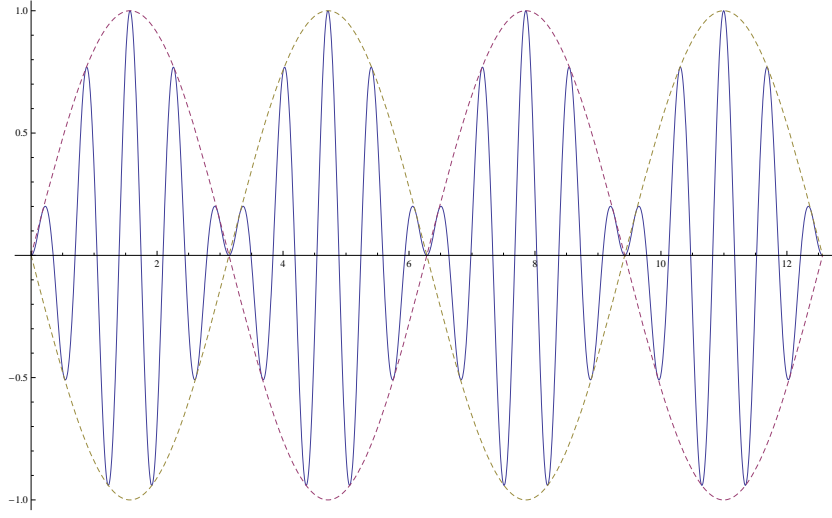


Figura 3.1: Fenomeno dei battimenti, dato con $\omega_0 = 10$ e $\omega = 8$.

Caso $\omega_0 = \omega$. Si ottiene il sistema

$$\begin{cases} c_1'(x) = -\frac{B}{2\omega} \sin(2\omega x) \\ c_2'(x) = \frac{B}{2\omega} (\cos(2\omega x) + 1) \end{cases}$$

le cui soluzioni sono date da

$$\begin{cases} c_1(x) = c_1 + \frac{B}{4\omega^2} \cos(2\omega x) \\ c_2(x) = c_2 + \frac{B}{4\omega} \sin(2\omega x) + \frac{B}{2\omega} x \end{cases}$$

che portano alla soluzione generale

$$y(x) = c_1 \cos(\omega x) + c_2 \sin(\omega x) + \frac{B}{4\omega^2} \cos(\omega x) + \frac{B}{2\omega} x \sin(\omega x).$$

Se tra queste funzioni cerchiamo la soluzione del Problema di Cauchy

$$\begin{cases} y'' + \omega^2 y = B \cos(\omega x) \\ y(0) = y'(0) = 0 \end{cases}$$

si ottiene la funzione

$$y(x) = \frac{B}{2\omega} x \sin(\omega x)$$

il cui grafico è riportato in figura 3.2; si ha a che fare in questo caso con il *fenomeno della risonanza*.

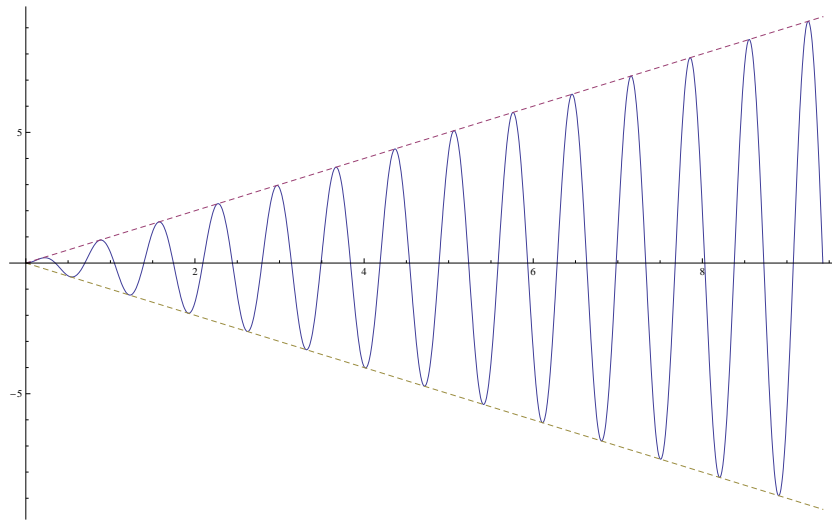


Figura 3.2: Fenomeno della risonanza.

3.3 Equazioni lineari di ordine superiore

I risultati visti nella sezione precedente si possono generalizzare ad equazioni di ordine superiore, cioè al caso dell'equazione

$$y^{(n)}(x) + a_1(x)y^{(n-1)}(x) + \dots + a_n(x)y(x) = f(x)$$

con $a_1, \dots, a_n, f : I \rightarrow \mathbb{R}$ funzioni continue. In particolare si avrà che:

1. la soluzione generale sarà data dalla somma della soluzione generale dell'omogenea associata e di una soluzione particolare;
2. l'insieme delle soluzioni dell'equazione omogenea è un sottospazio vettoriale di dimensione n di $C^n(I)$;
3. nel caso di equazione a coefficienti costanti, $a_1, \dots, a_n \in \mathbb{R}$, le soluzioni dell'equazione omogenea si cercano nella forma $e^{\lambda x}$ con λ radice del polinomio caratteristico $P(x) = x^n + a_1x + \dots + a_n$: la soluzione generale dell'omogenea sarà quindi sempre combinazione lineare di esponenziali e seni e coseni;
4. per l'individuazione della soluzione particolare, si potranno sempre usare i metodi per somiglianza e della variazione delle costanti.

Esempio 3.7 [Equazione della linea elastica o della trave elastica] Si tratta dell'equazione

$$EJy^{(4)}(x) = q$$

dove E è il modulo di Young, J il momento d'inerzia della sezione rispetto all'asse baricentrico ($B = EJ$ viene detto modulo di rigidità a flessione) e q è il carico che viene esercitato

sulla trave, supposto in questo esempio costante per motivi di semplicità. La soluzione generale di tale equazione, nell'ipotesi q costante, è data da

$$y(x) = c_1 x^3 + c_2 x^2 + c_3 x + c_4 + \frac{q}{24B} x^4.$$

Una variante di questo esempio è dato quando la trave è supposta appoggiata ad un supporto con resistenza di tipo elastico; con questo si intende che il profilo della trave soddisfa l'equazione differenziale

$$EJy^{(4)}(x) + ky(x) = q.$$

La soluzione generale è data qui da

$$y(x) = e^{\omega x} (c_1 \cos(\omega x) + c_2 \sin(\omega x)) + e^{-\omega x} (c_3 \cos(\omega x) + c_4 \sin(\omega x)) + \frac{q}{k},$$

dove $\omega = \sqrt[4]{\frac{k}{EJ}}$.

Capitolo 4

\mathbb{R}^N , topologia, limiti e funzioni continue

Fino a questo punto si sono studiate le funzioni $f : I \rightarrow \mathbb{R}$ con I intervallo di \mathbb{R} ; di tali funzioni si sono date le definizioni di continuità, derivabilità e integrabilità, si sono studiati problemi di massimo e minimo e metodi per la loro individuazione. Iniziamo ora lo studio di funzioni $f : A \rightarrow B$ con A e B sottoinsiemi di spazi Euclidei a più dimensioni; si noti che tutte le definizioni (continuità, derivabilità e integrabilità) usano il concetto di limite e che in ultima analisi tale nozione è basata sul concetto di distanza.

Nel prossimo paragrafo inizieremo quindi lo studio delle proprietà dei sottoinsiemi di \mathbb{R}^N , in particolare daremo la nozione di insieme aperto e chiuso (si parla in tal caso di topologia di \mathbb{R}^N) ed estenderemo il concetto di intervallo ad N variabili.

Inizieremo con la definizione di distanza in \mathbb{R}^N ; un problema cui si dovrà fare attenzione quando si trattano spazi euclidei a più dimensioni è il concetto di base. Un punto del piano o dello spazio Euclideo può essere individuato in vari modi, sia tramite le coordinate cartesiane che tramite, ad esempio, le coordinate polari; bisognerà quindi prestare attenzione al fatto che le conclusioni cui si perviene non dipendano dalle coordinate date, cioè che, cambiando coordinate, si riescano ad esempio ad individuare gli stessi punti di massimo, minimo o stazionari. A tal fine dovremo dare definizioni di continuità e soprattutto derivabilità che siano indipendenti dal sistema di riferimento fissato.

4.1 Topologia di \mathbb{R}^N

Iniziamo anzitutto con una precisazione; quando si parla di spazio Euclideo N -dimensionale, si intende uno spazio vettoriale ad N dimensioni con prodotto scalare; l'identificazione di tale spazio con \mathbb{R}^N (N -uple di numeri reali) può essere fatta una volta che si fissi una base dello spazio (la N -upla sarà costruita con i numeri della combinazione lineare). Nel seguito si dovrà fare sempre attenzione al fatto che le nozioni introdotte non dipendano dalla base fissata.

Denotiamo i punti di \mathbb{R}^N come vettori $x = (x_1, \dots, x_N)$ e su \mathbb{R}^N ricordiamo che è definito il prodotto scalare standard

$$x \cdot y = x_1 y_1 + \dots + x_N y_N, \quad \forall x, y \in \mathbb{R}^N.$$

Tale prodotto scalare definisce una norma $\|x\| = \sqrt{x \cdot x}$ per la quale

$$|x_j| \leq \|x\| \leq \sqrt{N} \max_{i=1, \dots, N} (|x_i|), \quad \forall j = 1, \dots, N;$$

la norma $\|\cdot\|$ definisce una distanza $d(x, y) = \|x - y\|$. Le principali proprietà della norma e della distanza sono le seguenti

1. $\|x\| \geq 0$ e $\|x\| = 0$ se e solo se $x = 0$, $d(x, y) \geq 0$ e $d(x, y) = 0$ se e solo se $x = y$;
2. $\|\lambda x\| = |\lambda| \|x\|$ per ogni $\lambda \in \mathbb{R}$, $d(x, y) = d(y, x)$;
3. $\|x + y\| \leq \|x\| + \|y\|$, $d(x, y) \leq d(x, z) + d(z, y)$.

Possiamo dare la seguente definizione

Definizione 4.1 (Intorno sferico) Si chiamano intorno sferico aperto ed intorno sferico chiuso di $x \in \mathbb{R}^N$ con raggio $r > 0$ gli insiemi

$$B_r(x) = \{y \in \mathbb{R}^N : \|x - y\| < r\}, \quad \overline{B}_r(x) = \{y \in \mathbb{R}^N : \|x - y\| \leq r\}$$

mentre viene detta sfera di centro x e raggio $r > 0$ l'insieme

$$S_r(x) = \{y \in \mathbb{R}^N : \|x - y\| = r\}.$$

Si noti che $\overline{B}_r(x) = B_r(x) \cup S_r(x)$.

Esercizio 4.1 Nel caso $N = 1$, si ottiene che $B_r(x) = (x - r, x + r)$, $\overline{B}_r(x) = [x - r, x + r]$ e $S_r(x) = \{x - r, x + r\}$.

Definizione 4.2 (Insieme aperto e chiuso) Dato $A \subset \mathbb{R}^N$ e $x \in A$, si dice che x è punto interno ad A se esiste $r > 0$ tale che $B_r(x) \subset A$; x si dice esterno se x è interno ad $A^c = \mathbb{R}^N \setminus A$, cioè se esiste $r > 0$ tale che $B_r(x) \cap A = \emptyset$. Un punto x che non è né interno né esterno si dice punto di bordo o di frontiera. Si denota con A° l'insieme dei punti interni ad A , con ∂A l'insieme dei punti di frontiera di A e con $\overline{A} = A \cup \partial A$ la chiusura di A . A si dice aperto se $A = A^\circ$, cioè se tutti i suoi punti sono punti interni, mentre A si dice chiuso se $\overline{A} = A$ o equivalentemente se $A^c = \mathbb{R}^N \setminus A$ è aperto.

Esempio 4.1 Se fissiamo $a, b \in \mathbb{R}$ con $a < b$, allora (a, b) è aperto, $[a, b]$ è chiuso e $(a, b]$, $[a, b)$ non sono né aperto né chiusi; in tutti i casi l'insieme dei punti di frontiera è dato da $\{a, b\}$.

Esempio 4.2 Se $A = \mathbb{R}^N$, allora $A^\circ = A = \mathbb{R}^N$ (quindi \mathbb{R}^N è aperto) e $A^c = \partial A = \emptyset$. Se invece $A = \mathbb{Q}$, allora $A^\circ = \emptyset$, $\overline{A} = \partial A = \mathbb{R}$ (quindi \mathbb{Q} non è né aperto né chiuso). Infine, se $A = B_r(x)$, $A^\circ = B_r(x)$, $\overline{A} = \overline{B}_r(x)$ e $\partial A = S_r(x)$. Vediamo come si dimostra che in questo ultimo caso $A^\circ = B_r(x)$: preso $y \in B_r(x)$, dobbiamo trovare $\varepsilon > 0$ tale che $B_\varepsilon(y) \subset B_r(x)$; ma se scegliamo $\varepsilon = r - \|y - x\|$, se $z \in B_\varepsilon(y)$, cioè se $\|z - y\| < r - \|y - x\|$, si ottiene che

$$\|z - x\| = \|z - y + y - x\| \leq \|z - y\| + \|y - x\| < r - \|y - x\| + \|y - x\| = r,$$

cioè $z \in B_r(x)$, che era quanto volevamo provare. Analogamente, se $\|y - x\| > r$, si nota che con la scelta $\varepsilon = \|y - x\| - r$ si ottiene che $B_\varepsilon(y) \cap B_r(x) = \emptyset$, cioè $B_\varepsilon(y) \subset \mathbb{R}^N \setminus \overline{B}_r(x)$, e quindi $\mathbb{R}^N \setminus \overline{B}_r(x)$ è aperto. Infine, se $\|y - x\| = r$, si nota che per ogni $\varepsilon > 0$ la palla $B_\varepsilon(y)$ interseca sia $B_r(x)$ che il suo complementare, e quindi i punti y tali che $\|y - x\| = r$ sono punti di frontiera.

Definizione 4.3 Dato un insieme $A \subset \mathbb{R}^N$, diremo che un punto x_0 è di accumulazione per A se per ogni $r > 0$, la palla $B_r(x_0)$ contiene punti di A diversi da x_0 , cioè se

$$B_r(x_0) \cap (A \setminus \{x_0\}) \neq \emptyset.$$

Un punto di accumulazione x_0 può appartenere o meno all'insieme A , ma ci sono anche punti di un insieme che non sono punti di accumulazione; si pensi ad esempio all'insieme della retta reale $A = (0, 1] \cup \{2\}$, in cui 2 non è un punto di accumulazione, mentre 0 lo è anche se non appartiene ad A . Un punto di A che non sia di accumulazione per A viene detto punto *isolato*.

Nella seguente proposizioni sono elencate le principali proprietà degli aperti e dei chiusi sotto le operazioni di unione e intersezione.

Proposizione 4.4 Supponiamo di avere due famiglie, una \mathcal{A} fatta di insiemi aperti e una \mathcal{C} fatta di insiemi chiusi. Allora

1. l'insieme

$$\bigcup_{A \in \mathcal{A}} A = \{x \in \mathbb{R}^N : x \in A \text{ per qualche } A \in \mathcal{A}\}$$

è aperto, mentre l'insieme

$$\bigcap_{C \in \mathcal{C}} C = \{x \in \mathbb{R}^N : x \in C \text{ per ogni } C \in \mathcal{C}\}$$

è chiuso;

2. dato un numero finito di insiemi $A_1, \dots, A_k \in \mathcal{A}$ e $C_1, \dots, C_m \in \mathcal{C}$, l'insieme $A_1 \cap \dots \cap A_k$ è aperto e l'insieme $C_1 \cup \dots \cup C_m$ è chiuso.

DIMOSTRAZIONE Se $x \in A$ per un qualche A , allora, dato che A è aperto, esiste $r > 0$ tale che

$$B_r(x) \subset A \subset \bigcup_{A \in \mathcal{A}} A$$

da cui il fatto che l'unione qualsiasi di insiemi aperti è aperta. Sia poi $x \in A_1 \cap \dots \cap A_k$; dato che ogni A_j , $j = 1, \dots, k$, è aperto, esistono $r_j > 0$ tali che $B_{r_j}(x) \subset A_j$ per ogni $j = 1, \dots, k$. Scegliendo $r = \min_j r_j$, si ottiene che $B_r(x) \subset A_1 \cap \dots \cap A_k$, e quindi tale intersezione è aperta. Le proprietà sui chiusi si ottengono osservando che se C è chiuso, allora $A = \mathbb{R}^N \setminus C$ è aperto. \square

Osservazione 4.5 Si nota che intersezione arbitraria di insiemi aperti non è in genere aperta; infatti, ad esempio in \mathbb{R} , se si considerano gli insiemi aperti $A_j = (-1/j, 1/j)$, si ottiene che

$$\bigcap_{j \in \mathbb{N}} A_j = \{0\}$$

e quest'ultimo non è un insieme aperto.

Definizione 4.6 (Insiemi limitati e insiemi compatti) Un insieme $A \subset \mathbb{R}^N$ si dice *limitato* se esiste $r > 0$ tale che $A \subset B_r(0)$, cioè se esiste $r > 0$ tale che $\|x\| < r$ per ogni $x \in A$. Un insieme A viene detto *compatto* se è chiuso e limitato.

Osservazione 4.7 La definizione di limitatezza di un insieme A può anche essere data richiedendo che esista $x_0 \in \mathbb{R}^N$ e $r > 0$ tale che $A \subset B_r(x_0)$. Inoltre, la definizione di insieme compatto sarebbe in realtà più topologica, mentre la caratterizzazione dei compatti come insiemi chiusi e limitati è un risultato che vale per gli spazi euclidei finito-dimensionali. La definizione topologica di compattezza dice che: K è compatto se da ogni ricoprimento aperto

$$K \subset \bigcup_{A \in \mathcal{A}} A$$

si può scegliere un sottoricoprimento finito, cioè esistono $A_1, \dots, A_k \in \mathcal{A}$ tali che

$$K \subset A_1 \cup \dots \cup A_k.$$

Esiste anche, come vedremo, una caratterizzazione della compattezza fatta usando le successioni.

4.2 Successioni in \mathbb{R}^N

Una successione in \mathbb{R}^N è semplicemente una funzione $x : \mathbb{N} \rightarrow \mathbb{R}^N$, cioè una famiglia $(x(n) = x_n)_{n \in \mathbb{N}}$ di elementi di \mathbb{R}^N . La successione x_n si dice limitata se l'insieme

$$A = \{x_n : n \in \mathbb{N}\}$$

è limitato in \mathbb{R}^N , cioè se esiste $r > 0$ tale che $\|x_n\| < r$ per ogni $n \in \mathbb{N}$.

Definizione 4.8 Una successione $(x_n)_{n \in \mathbb{N}}$ si dice convergente ad un elemento $x \in \mathbb{R}^N$, e scriveremo

$$\lim_{n \rightarrow \infty} x_n = x,$$

se e solo se per ogni $\varepsilon > 0$ esiste $n_0 \in \mathbb{N}$ tale che $\|x_n - x\| < \varepsilon$ per ogni $n \geq n_0$.

Si nota che, se la successione $(x_n)_{n \in \mathbb{N}}$ è convergente, allora è limitata; inoltre, utilizzando il fatto che, preso un elemento $y = (y_1, \dots, y_N) \in \mathbb{R}^N$

$$|y_j| \leq \|y\| \leq \sqrt{N} \max_{k=1, \dots, N} |y_k|, \quad \forall j = 1, \dots, N,$$

se ne deduce che una successione $(x_n)_{n \in \mathbb{N}}$ è convergente se e solo se le successioni delle sue componenti $(x_{n,j})_{n \in \mathbb{N}}$ convergono a x_j per ogni $j = 1, \dots, N$. Si noti che questo ragionamento implica anche:

1. il limite è unico, dato che il limite di ogni componente è unico;
2. date due successioni $(x_n)_{n \in \mathbb{N}}$ e $(y_n)_{n \in \mathbb{N}}$ convergenti rispettivamente ad x ed y in \mathbb{R}^N , allora la successione $(x_n + y_n)_{n \in \mathbb{N}}$ converge ad $x + y$;
3. se $(x_n)_{n \in \mathbb{N}}$ converge ad $x \in \mathbb{R}^N$ e $\lambda \in \mathbb{R}$, allora $(\lambda x_n)_{n \in \mathbb{N}}$ converge a λx .

Esempio 4.3 Consideriamo la successione in \mathbb{R}^2 data da

$$x_n = \left(\frac{1}{n}, \frac{n+1}{n} \right),$$

dato che la prima componente, $\frac{1}{n}$, converge a 0 e la seconda, $\frac{n+1}{n}$, converge ad 1, si avrà che la successione converge al punto $(0, 1) \in \mathbb{R}^2$. Invece, la successione $x_n = ((-1)^n, \frac{1}{n})$ non può convergere in quanto la successione delle prime componenti, $(-1)^n$, non converge.

La definizione di successione estratta sarà data in modo analogo alla definizione di successione estratta in \mathbb{R} , e cioè tramite una funzione $k : \mathbb{N} \rightarrow \mathbb{N}$ crescente, da cui la successione estratta $(x_{n_k})_{n \in \mathbb{N}}$ è data da $x_{n_k} = x_{k(n)}$.

Proposizione 4.9 (Chiusi e compatti per successioni) *Dato un insieme $A \subset \mathbb{R}^N$, si ha che:*

1. *A è chiuso se e solo se per ogni successione $(x_n)_{n \in \mathbb{N}} \subset A$ con x_n convergente ad x , si ha che $x \in A$;*
2. *A è compatto se e solo se da ogni successione $(x_n)_{n \in \mathbb{N}} \subset A$ si può estrarre una sottosuccessione x_{n_k} convergente ad un elemento di $x \in A$.*

Diremo infine che un punto $x \in \mathbb{R}$ è un punto di accumulazione per A se esiste una successione $(x_n)_{n \in \mathbb{N}} \subset A$ con $x_n \neq x$ e x_n convergente ad x .

4.3 Limiti e funzioni continue

Veniamo ora allo studio delle funzioni di più variabili, e più precisamente alla definizione di limite e continuità per funzioni di più variabili.

Definizione 4.10 *Data una funzione $f : A \rightarrow \mathbb{R}^k$ con $A \subset \mathbb{R}^N$ e x_0 punto di accumulazione per A , diremo che $f(x)$ tende ad un vettore $l \in \mathbb{R}^k$ per x che tende ad x_0 , in simboli*

$$\lim_{x \rightarrow x_0} f(x) = l.$$

se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $x \in A$ è tale che $0 < \|x - x_0\| < \delta$, allora $\|f(x) - l\| < \varepsilon$. Nel caso $k = 1$, diremo inoltre che

$$\lim_{x \rightarrow x_0} f(x) = +\infty \quad (-\infty)$$

se per ogni $M > 0$ ($M < 0$) esiste $\delta > 0$ tale che se $x \in A$ soddisfa $0 < \|x - x_0\| < \delta$, allora $f(x) > M$ ($f(x) < M$).

Osservazione 4.11 Il limite di una funzione può anche essere caratterizzato tramite successioni; cioè

$$\lim_{x \rightarrow x_0} f(x) = l$$

se e solo se per ogni successione $(x_n)_n$ che converge ad x_0 , la successione $f(x_n)$ converge ad l . Questa osservazione dimostra anche che la funzione f converge ad l se e solo se le componenti f_i convergono alle componenti l_i , per $i = 1, \dots, k$.

Come visto per le successioni, anche per il limite di funzioni valgono le seguenti proprietà;

1. $f(x) \rightarrow l \in \mathbb{R}^k$ se e solo se per ogni $j = 1, \dots, k$ si ha che $f_j(x) \rightarrow l_j$;
2. se $f(x) \rightarrow l$ e $g(x) \rightarrow l'$, allora $f(x) + g(x) \rightarrow l + l'$;
3. se $f(x) \rightarrow l$ e $\lambda \in \mathbb{R}$, allora $\lambda f(x) \rightarrow \lambda l$.

Definizione 4.12 (Continuità) Una funzione $f : A \rightarrow \mathbb{R}^k$ si dice continua se e solo se per ogni $x_0 \in A$ punto di accumulazione per A , si ha che

$$\lim_{x \rightarrow x_0} f(x) = f(x_0),$$

cioè se e solo se per ogni $x_0 \in A$ punto di accumulazione e per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $x \in A$, con $0 < \|x - x_0\| < \delta$, allora $\|f(x) - f(x_0)\| < \varepsilon$. Diremo inoltre che f è uniformemente continua se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $x, y \in A$ sono tali che $\|x - y\| < \delta$, allora $\|f(x) - f(y)\| < \varepsilon$.

Osservazione 4.13 Così come per i limiti di funzioni, anche la continuità può essere caratterizzata per successioni, cioè

$$\lim_{x \rightarrow x_0} f(x) = f(x_0),$$

se e solo se per ogni successione $(x_n)_n$ convergente ad x_0 , $(f(x_n))_n$ converge a $f(x_0)$. Analogamente, una funzione sarà continua se e solo se ogni sua componente f_i , $i = 1, \dots, k$ è una funzione continua.

La somma di funzioni continue è continua, moltiplicare una funzione continua per uno scalare $\lambda \in \mathbb{R}$ produce una funzione continua, ma anche la moltiplicazione di una funzione continua $f : A \rightarrow \mathbb{R}^k$ con una funzione continua $g : A \rightarrow \mathbb{R}$ produce una funzione continua $gf : A \rightarrow \mathbb{R}^k$ data da $gf(x) = g(x)f(x)$ (moltiplicazione della funzione vettoriale f con la funzione scalare g).

Esempio 4.4 È facile a questo punto rendersi conto che le funzioni $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ date da

$$f(x, y) = x^h, \quad f(x, y) = y^m$$

sono continue (si usi la caratterizzazione per successioni), e quindi ogni polinomio

$$f(x, y) = \sum_{h,m=0}^d a_{hm} x^h y^m$$

è una funzione continua.

Abbiamo inoltre la seguente

Proposizione 4.14 Siano $f : A \rightarrow B$ e $g : B \rightarrow \mathbb{R}^m$ due funzioni continue con $A \subset \mathbb{R}^N$, $f(A) \subset B$ e $B \subset \mathbb{R}^k$; allora la funzione $F : A \rightarrow \mathbb{R}^m$ composizione di f e g , $F = g \circ f$ è continua.

DIMOSTRAZIONE. La dimostrazione è analoga a quella fatta per funzioni di una variabile, usando la definizione di limite; dimostriamo infatti che la funzione $F(x) = g(f(x))$ è continua in ogni punto x_0 di accumulazione per A . Presa una successione x_n convergente ad x_0 , la continuità di f implica che la successione $f(x_n)$ converge ad $f(x_0)$ e quindi, la continuità di g implica che la successione $g(f(x_n))$ converge a $g(f(x_0))$, da cui la continuità di F . \square

Chiaramente ogni funzione uniformemente continua è continua, mentre non vale in generale il viceversa. Si ha tuttavia il seguente Teorema.

Teorema 4.15 (Teorema di Cantor) *Se $F : A \rightarrow \mathbb{R}^k$ è continua e A è compatto, allora F è uniformemente continua su A .*

Esempio 4.5 Esempi di funzioni continue sono dati da

$$f(x, y) = x^2 + y^2, \quad f(x, y) = xy,$$

mentre non è continua la funzione

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{se } (x, y) = (0, 0) \end{cases}$$

anche se sono continue le funzioni $x \mapsto f(x, y)$ e $y \mapsto f(x, y)$. La continuità per funzioni di più variabili è quindi una nozione più riposta rispetto alla continuità per sezioni.

Abbiamo la seguente relazione tra funzioni continue e topologie, che fornisce un modo per costruire insiemi aperti e chiusi tramite funzioni continue.

Proposizione 4.16 *Sia $f : A \rightarrow \mathbb{R}$ una funzione continua, allora:*

1. *vale il teorema di permanenza del segno, cioè se $x_0 \in A$ è tale che $f(x_0) > 0$ (< 0), allora esiste $r > 0$ tale che se $y \in B_r(x_0) \cap A$ vale $f(y) > 0$ (< 0);*
2. *se A è aperto, allora per ogni $c \in \mathbb{R}$ gli insiemi*

$$\{x \in A : f(x) < c\} := \{f < c\}, \quad \{f > c\}$$

sono aperti, mentre l'insieme

$$\{f = c\}$$

è chiuso.

DIMOSTRAZIONE. La permanenza del segno segue direttamente dalla definizione di funzione continua; infatti, dato che f è continua in x_0 , si ha che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $y \in A$ con $\|y - x_0\| < \delta$ allora $|f(y) - f(x_0)| < \varepsilon$ (si noti che stiamo sfruttando il fatto che la norma e il valore assoluto sono la stessa cosa nel caso uni-dimensionale), cioè

$$(4.1) \quad f(x_0) - \varepsilon < f(y) < f(x_0) + \varepsilon.$$

Se prendiamo quindi $\varepsilon = f(x_0)$ ed $r = \delta$, la relazione precedente ci dice esattamente che se $y \in A \cap B_r(x_0)$, allora la disuguaglianza di sinistra di (4.1) implica che $f(y) > 0$. In modo analogo si ragiona nel caso $f(x_0) < 0$. Per vedere che gli insiemi $\{f > c\}$ e $\{f < c\}$ sono aperti, si applica il teorema della permanenza del segno alle funzioni $f(x) - c$ (che sono continue) e, siccome A è aperto, si può prendere $r > 0$ in modo tale che $B_r(x_0) \subset A$. Infine, $\{f = c\}$ risulta essere chiuso dato che

$$\mathbb{R}^N \setminus \{f = c\} = \{f > c\} \cup \{f < c\}$$

e l'unione di aperti è aperta, e da qui segue la chiusura di $\{f = c\}$ in quanto il suo complementare è aperto. \square

Esempio 4.6 Se consideriamo le funzioni $f(x, y) = x^2 + y^2$ e $g(x, y) = xy$ (che sono continue), se ne ricava che $\{x^2 + y^2 < c\}$ è aperto per ogni $c > 0$ (d'altronde coincide con $B_{\sqrt{c}}(0)$), così pure gli insiemi $\{xy > c\}$ sono aperti.

Capitolo 5

Curve ed integrali curvilinei

Iniziamo il nostro studio delle funzioni vettoriali di più variabili con il caso più semplice; tale studio ci permetterà di sviluppare strumenti e concetti utili per un approccio sistematico al caso generale. Cominceremo con lo studiare le curve dello spazio Euclideo, cioè funzioni vettoriali di una sola variabile reale; le curve per noi saranno entità parametrizzate (saranno le funzioni che le definiscono), ma vedremo quali informazioni geometriche sul luogo dei suoi punti (traiettorie) si possono dedurre dalla parametrizzazione.

5.1 Curve e curve regolari

Una curva in \mathbb{R}^N è data da una funzione continua $\varphi : I \rightarrow \mathbb{R}^N$ con I intervallo di \mathbb{R} , limitato o meno (non necessariamente aperto o chiuso). La richiesta della continuità nella definizione di curva serve per sottolineare che la variare di $t \in I$, il punto $\varphi(t)$ si muove nello spazio \mathbb{R}^N con continuità, descrivendo cioè una traiettoria continua. Come abbiamo visto nel capitolo precedente, la curva $\varphi : I \rightarrow \mathbb{R}^N$ sarà continua se tutte le sue componenti $\varphi_i : I \rightarrow \mathbb{R}$ sono continue. Si definisce il supporto o sostegno di φ l'insieme

$$\varphi(I) = \{\varphi(t) : t \in I\} \subset \mathbb{R}^N :$$

il sostegno di una curva non va confuso con il grafico della curva, il quale è un sottoinsieme del prodotto cartesiano $I \times \mathbb{R}^N$.

Esempio 5.1

Retta e Segmento. Dati due punti x e y in \mathbb{R}^N , la retta passante per essi è parametrizzata da $\varphi : \mathbb{R} \rightarrow \mathbb{R}^N$

$$\varphi(t) = x + t(y - x).$$

Se si usa la stessa parametrizzazione ma definita in $[0, 1]$, si ottiene il segmento, che denoteremo con $[x, y]$, congiungente x e y ; tale segmento è un segmento orientato che parte da x (tempo $t = 0$) e termina in y (tempo $t = 1$).

Poligonale. Data una collezione ordinata di punti $x_0, x_1, \dots, x_k \in \mathbb{R}^N$ di punti, si definisce poligonale \mathcal{P} di vertici x_0, \dots, x_k l'unione dei segmenti $[x_{i-1}, x_i]$, $i = 1, \dots, k$ parametrizzata quindi dalla curva $\varphi : [0, 1] \rightarrow \mathbb{R}^N$

$$\varphi(t) = x_{i-1} + t(x_i - x_{i-1}), \quad \frac{i-1}{k} \leq t \leq \frac{i}{k}.$$

Circonferenza nel piano. Nel piano, la circonferenza di raggio $r > 0$ centrata nell'origine può essere parametrizzata dalla funzione $\varphi : [0, 2\pi] \rightarrow \mathbb{R}^2$

$$\varphi(t) = (r \cos t, r \sin t).$$

Ellisse nel piano. L'ellisse del piano $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ può essere parametrizzata dalla funzione $\varphi : [0, 2\pi] \rightarrow \mathbb{R}^2$

$$\varphi(t) = (a \cos t, b \sin t).$$

Si fissi ora un punto $t_0 \in I$ e si consideri, per $t_1 \in I$, la quantità

$$\frac{\varphi(t_1) - \varphi(t_0)}{t_1 - t_0} = \left(\frac{\varphi_1(t_1) - \varphi_1(t_0)}{t_1 - t_0}, \dots, \frac{\varphi_N(t_1) - \varphi_N(t_0)}{t_1 - t_0} \right).$$

Avremo quindi che il limite

$$\lim_{t_1 \rightarrow t_0} \frac{\varphi(t_1) - \varphi(t_0)}{t_1 - t_0} = \varphi'(t_0)$$

esiste se e solo se esistono i limiti

$$\lim_{t_1 \rightarrow t_0} \frac{\varphi_i(t_1) - \varphi_i(t_0)}{t_1 - t_0}$$

per ogni $i = 1, \dots, N$, cioè se e solo se le componenti della curva sono funzioni derivabili. Diremo quindi che la curva è derivabile in t_0 se le sue componenti lo sono, mentre diremo che $\varphi \in C^1(I)$ se le sue componenti sono derivabili con derivata continua in I .

Per quanto abbiamo visto, l'equazione

$$r_{t_1}(t) = \varphi(t_0) + t \frac{\varphi(t_1) - \varphi(t_0)}{t_1 - t_0}$$

rappresenta la retta passante per $\varphi(t_0)$ e $\varphi(t_1)$, cioè la retta secante il sostegno di φ in tali punti. Se passiamo al limite per $t_1 \rightarrow t_0$, tale retta tende alla tangente al supporto di φ nel punto $\varphi(t_0)$, che sarà quindi individuata dalla parametrizzazione

$$r(t) = \varphi(t_0) + t\varphi'(t_0).$$

Il vettore $\varphi'(t_0)$ rappresenta la velocità della curva in t_0 e, nel caso $\varphi'(t_0) \neq 0$, il vettore unitario

$$T_\varphi(t_0) = \frac{\varphi'(t_0)}{\|\varphi'(t_0)\|}$$

rappresenta il versore tangente alla curva in $\varphi(t_0)$.

Definizione 5.1 Data una curva $\varphi : I \rightarrow \mathbb{R}^N$, diremo che:

1. φ è regolare se $\varphi \in C^1(I)$ con $\varphi'(t) \neq 0$ per ogni $t \in I$;
2. φ è regolare a tratti se possiamo decomporre I in sottointervalli disgiunti I_i con $i = 1, \dots, k$ in modo tale che φ è regolare in I_i per ogni $i = 1, \dots, k$;
3. la curva si dice semplice se $t_1 \neq t_2$ implica $\varphi(t_1) \neq \varphi(t_2)$, con almeno uno dei due punti t_1 o t_2 interno ad I ;

4. se $I = [a, b]$, il punto $\varphi(a)$ si dice primo estremo o punto iniziale della curva, mentre $\varphi(b)$ si dice secondo estremo o punto finale della curva;
5. se $I = [a, b]$, la curva si dice chiusa se $\varphi(a) = \varphi(b)$.

Vedremo che si potrà sempre ottenere la condizione $\varphi' \neq 0$ nella definizione di curva regolare: qui è stata posta per una semplice questione di comodità. La definizione di curva semplice può sembrare complicata, ma serve semplicemente per ammettere anche curve chiuse semplici, quali la circonferenza, in cui gli unici due punti di non iniettività della curva sono gli estremi.

Definizione 5.2 (Curve equivalenti) Due curve $\varphi : I \rightarrow \mathbb{R}^N$ e $\psi : J \rightarrow \mathbb{R}^N$ si dicono equivalenti se esiste una funzione $\alpha : I \rightarrow J$ di classe $C^1(I)$ iniettiva e suriettiva tale che $\alpha'(t) \neq 0$ per ogni $t \in I$ e

$$\varphi(t) = \psi(\alpha(t)), \quad \forall t \in I.$$

La mappa α viene anche chiamata cambiamento di parametro o riparametrizzazione ammissibile.

L'equivalenza di due curve implica l'uguaglianza dei due sostegni, mentre il viceversa non è vero; ad esempio le due funzioni $\varphi : [0, 2\pi] \rightarrow \mathbb{R}^2$, $\varphi(t) = (\cos t, \sin t)$ e $\psi : [0, 3\pi] \rightarrow \mathbb{R}^2$, $\psi(t) = (\cos t, \sin t)$ hanno lo stesso sostegno ma non sono equivalenti. Per poter caratterizzare curve equivalenti tramite il sostegno, bisogna restringersi alla classe delle curve semplici. Abbiamo cioè il seguente risultato.

Proposizione 5.3 Siano $\varphi : I \rightarrow \mathbb{R}^N$ e $\psi : J \rightarrow \mathbb{R}^N$ due curve regolari e semplici; allora le due curve sono equivalenti se e solo se hanno lo stesso sostegno, cioè se e solo se

$$\varphi(I) = \psi(J).$$

Si noti che la condizione $\alpha' \neq 0$ implica che $\alpha' > 0$ oppure $\alpha' < 0$; quindi α è sempre strettamente monotona. Nel caso in cui α sia monotona crescente, avremo che due curve equivalenti avranno lo stesso verso di percorrenza, mentre nel caso di α decrescente, il loro verso di percorrenza sarà opposto.

Esempio 5.2

1. Una riparametrizzazione per una curva $\varphi : [a, b] \rightarrow \mathbb{R}^N$ è data ad esempio da

$$\alpha : [a, b] \rightarrow [0, 1], \quad \alpha(t) = \frac{t - a}{b - a}.$$

Non sarà quindi restrittivo supporre sempre una curva definita sull'intervallo $[0, 1]$, almeno nel caso in cui I sia un intervallo chiuso e limitato.

2. L'arco di circonferenza $\varphi : [0, \pi] \rightarrow \mathbb{R}^2$

$$\varphi(t) = (\cos t, \sin t)$$

può essere equivalentemente riparametrizzata da $\psi : [-1, 1] \rightarrow \mathbb{R}^2$

$$\psi(t) = (t, \sqrt{1 - t^2}).$$

5.1.1 Curve nel piano

La geometria del piano permette di vedere le curve in vari modi. Ad esempio, se viene data una funzione $f : I \rightarrow \mathbb{R}$ continua reale di variabile reale, allora possiamo definire in modo canonico due curve parametrizzate $\varphi_x : I \rightarrow \mathbb{R}^2$, $\varphi_y : I \rightarrow \mathbb{R}^2$

$$\varphi_x(t) = (t, f(t)), \quad \varphi_y(t) = (f(t), t)$$

che verranno dette curve cartesiane rispettivamente rispetto ad x e rispetto ad y definite dalla funzione f . In tal caso il sostegno delle curve φ_x e φ_y coinciderà con il grafico della funzione f , grafico rispettivamente in x e in y . Le curve date saranno regolari se e solo se f è di classe $C^1(I)$ ed in tal caso, avendosi

$$\varphi'_x(t) = (1, f'(t)), \quad \varphi'_y(t) = (f'(t), 1)$$

la condizione $\varphi' \neq 0$ è automaticamente verificata.

Viceversa, se si vuole vedere una curva parametrizzata $\varphi : I \rightarrow \mathbb{R}^2$ come una curva cartesiana, si dovrà richiedere che φ sia una curva semplice e che il sostegno di φ sia un grafico, cioè che l'intersezione del sostegno con rette verticali (nel caso si voglia vederla come cartesiana rispetto alla x) contenga al più un punto, mentre l'intersezione con rette orizzontali (nel caso cartesiano rispetto alla y) contenga al più un punto.

Un altro modo per definire una curva nel piano è tramite l'utilizzo delle coordinate polari; in tal caso esistono sostanzialmente due approcci, il metodo esplicito e quello parametrico.

Per metodo esplicito si intende il caso in cui nell'utilizzo delle coordinate polari (r, ϑ) , il raggio r sia una funzione esplicita dell'angolo ϑ , cioè $r = g(\vartheta)$, $g : [\vartheta_1, \vartheta_2] \rightarrow [0, +\infty)$. Se siamo in questo caso, si può definire una curva parametrizzata $\varphi : [\vartheta_1, \vartheta_2] \rightarrow \mathbb{R}^2$ ponendo

$$(5.1) \quad \varphi(\vartheta) = (r \cos \vartheta, r \sin \vartheta) = g(\vartheta)(\cos \vartheta, \sin \vartheta);$$

derivando, si ottiene

$$(5.2) \quad \varphi'(\vartheta) = (g'(\vartheta) \cos \vartheta - g(\vartheta) \sin \vartheta, g'(\vartheta) \sin \vartheta + g(\vartheta) \cos \vartheta).$$

La curva data sarà regolare se e solo se g è di classe $C^1([\vartheta_1, \vartheta_2])$.

Per metodo parametrico si intende il caso in cui sia r che ϑ sono funzioni di un parametro $t \in I$

$$\begin{cases} r = r(t) \\ \vartheta = \vartheta(t). \end{cases}$$

In tal caso si ottiene la curva parametrizzata $\varphi : I \rightarrow \mathbb{R}^2$

$$\varphi(t) = (r(t) \cos \vartheta(t), r(t) \sin \vartheta(t));$$

derivando, si ottiene anche che

$$\varphi'(t) = (r'(t) \cos \vartheta(t) - r(t) \vartheta'(t) \sin \vartheta(t), r'(t) \sin \vartheta(t) + r(t) \vartheta'(t) \cos \vartheta(t)).$$

5.2 Lunghezza, ascissa curvilinea e integrali curvilinei

In questa sezione vogliamo dare la definizione di lunghezza di una curva e fornire un metodo di calcolo, almeno nel caso di curve regolari. La prima osservazione è che, dati due punti $x, y \in \mathbb{R}^N$, la distanza tra essi coincide con la lunghezza del segmento che li unisce, cioè

$$l([x, y]) = \|x - y\|.$$

Analogamente, data una poligonale \mathcal{P} di vertici $x_0, \dots, x_k \in \mathbb{R}^N$, possiamo definire la lunghezza della poligonale ponendo

$$l(\mathcal{P}) = \sum_{i=1}^k \|x_i - x_{i-1}\|.$$

Infine, data una curva $\varphi : [a, b] \rightarrow \mathbb{R}^N$, definiamo, per ogni $k \in \mathbb{N}$, la poligonale \mathcal{P}_k con vertici definiti da

$$x_i = \varphi(t_i), \quad t_i = a + \frac{i}{k}(b-a), i = 0, \dots, k.$$

Definizione 5.4 (Lunghezza di una curva) Data una curva parametrizzata $\varphi : [a, b] \rightarrow \mathbb{R}^N$, si definisce la sua lunghezza come

$$(5.3) \quad l(\varphi, [a, b]) = \lim_{k \rightarrow +\infty} l(\mathcal{P}_k).$$

Nel caso in cui $\varphi : I \rightarrow \mathbb{R}^N$, con I intervallo generico, si procede con un metodo di esaurimento, nel senso che si calcola la lunghezza sugli intervalli $[a, b] \subset I$ e si invade I con tali intervalli.

Diremo che la curva φ è rettificabile in I se ha lunghezza finita, cioè se $l(\varphi, I) < +\infty$.

Osservazione 5.5 La definizione (5.3) non è in realtà molto rigorosa, in quanto bisognerebbe prima di tutto dimostrare che il limite esiste, fatto non ovvio dalla definizione. È però facile rendersi conto che se invece di dividere l'intervallo $[a, b]$ in k parti uguali lo si divide in 2^k parti uguali, allora la monotonia segue. Infatti, se consideriamo la partizione $\tilde{\mathcal{P}}_k$ con vertici

$$x_i = \varphi(t_i), \quad t_i = a + \frac{i}{2^k}(b-a), i = 0, \dots, 2^k,$$

si ottiene che tra i vertici della partizione $\tilde{\mathcal{P}}_{k+1}$ ci sono anche tutti quelli di $\tilde{\mathcal{P}}_k$ ma vengono aggiunti anche i punti intermedi tra t_i e t_{i+1} per ogni i . La disuguaglianza triangolare

$$\|\varphi(t_{i+1}) - \varphi(t_i)\| \leq \left\| \varphi(t_{i+1}) - \varphi\left(\frac{t_{i+1} + t_i}{2}\right) \right\| + \left\| \varphi\left(\frac{t_{i+1} + t_i}{2}\right) - \varphi(t_i) \right\|$$

Ne segue che $l(\varphi, \tilde{\mathcal{P}}_k) \leq l(\varphi, \tilde{\mathcal{P}}_{k+1})$, da cui l'esistenza del limite

$$\tilde{l}(\varphi, [a, b]) = \lim_{k \rightarrow +\infty} l(\varphi, \tilde{\mathcal{P}}_k).$$

In realtà, la definizione più naturale della lunghezza della curva passa attraverso la nozione di estremo superiore, ponendo

$$l'(\varphi, I) = \sup\{l(\mathcal{P}) : \text{su tutte le partizioni } \mathcal{P} \text{ inscritte nella curva}\},$$

in quanto in tal caso non si dovrebbe distinguere tra intervalli chiusi e limitati ed intervalli generici. Inoltre, sembra intuitivo che dove la curva è rettilinea, non c'è necessità di effettuare partizioni dell'intervallo di definizione. La continuità di φ implica poi che le tre definizioni qui date coincidono, cioè

$$l(\varphi, [a, b]) = \tilde{l}(\varphi, [a, b]) = l'(\varphi, [a, b]).$$

Una proprietà importante della definizione della lunghezza di una curva $\varphi : [a, b] \rightarrow \mathbb{R}^N$ è che se $c \in (a, b)$, allora

$$l(\varphi, [a, b]) = l(\varphi, [a, c]) + l(\varphi, [c, b]).$$

In particolare, la funzione $s : [a, b] \rightarrow l(\varphi, [a, b])$ definita da

$$s(t) = l(\varphi, [a, t])$$

è una funzione monotona crescente. Tale funzione prende il nome di *ascissa curvilinea*, oppure *parametro* o *lunghezza d'arco*. Per il calcolo della lunghezza di una curva abbiamo il seguente Teorema, nel caso di curve regolari.

Proposizione 5.6 *Sia $\varphi : I \rightarrow \mathbb{R}^N$ una curva regolare a tratti; allora*

$$l(\varphi, I) = \int_I \|\varphi'(t)\| dt.$$

In particolare, per l'ascissa curvilinea si ha

$$\frac{ds(t)}{dt} = \|\varphi'(t)\|,$$

mentre la curva $\psi : [0, l(\varphi, I)] \rightarrow \mathbb{R}^N$ definita da

$$\psi(s) = \psi(s(t)) := \varphi(t)$$

è una riparametrizzazione equivalente di φ con

$$\left\| \frac{d\psi(s)}{ds} \right\| = 1.$$

DIM. Supponiamo per comodità $I = [a, b]$ e $\varphi : [a, b] \rightarrow \mathbb{R}^N$ regolare; il caso generale si ottiene per additività ed invadendo un intervallo generico I con intervalli chiusi e limitati. Dimostriamo quindi che la funzione lunghezza d'arco

$$s(t) = l(\varphi, [a, t])$$

è derivabile con $s'(t) = \|\varphi'(t)\|$. Se consideriamo i punti t e $t+h$, è facile rendersi conto che

$$(5.4) \quad \|\varphi(t+h) - \varphi(t)\| \leq s(t+h) - s(t) = l(\varphi, [t, t+h]),$$

dato che t e $t+h$ definiscono una particolare partizione dell'intervallo $[t, t+h]$. Prendiamo ora una poligonale generica \mathcal{P} individuata da una partizione t_0, \dots, t_k di $[t, t+h]$; grazie al teorema fondamentale del calcolo integrale, otteniamo

$$\begin{aligned} l(\varphi, \mathcal{P}) &= \sum_{i=1}^k \|\varphi(t_i) - \varphi(t_{i-1})\| = \sum_{i=1}^k \left\| \int_{t_{i-1}}^{t_i} \varphi'(\tau) d\tau \right\| \leq \sum_{i=1}^k \int_{t_{i-1}}^{t_i} \|\varphi'(\tau)\| d\tau \\ &= \int_{t_0}^{t_k} \|\varphi'(\tau)\| d\tau \leq \int_t^{t+h} \|\varphi'(\tau)\| d\tau. \end{aligned}$$

Quindi, passando all'estremo superiore su tutte le partizioni,

$$(5.5) \quad s(t+h) - s(t) = l(\varphi, [t, t+h]) \leq \int_t^{t+h} \|\varphi'(\tau)\| d\tau.$$

Combinando (5.4) e (5.5), si ottiene quindi

$$\left\| \frac{\varphi(t+h) - \varphi(t)}{h} \right\| \leq \frac{s(t+h) - s(t)}{h} \leq \frac{1}{h} \int_t^{t+h} \|\varphi'(\tau)\| d\tau;$$

la continuità di φ' implica la prima parte del teorema. Per la seconda parte, quanto sopra mostra che $s : [a, b] \rightarrow l(\varphi, [a, b])$ è una funzione strettamente monotona crescente, quindi invertibile; definiamo quindi la curva $\psi : [0, l(\varphi, [a, b])] \rightarrow \mathbb{R}^N$ ponendo $\psi(\tau) = \varphi(s^{-1}(\tau))$, cioè $\varphi(t) = \psi(s(t))$. Le curve ψ e φ sono equivalenti con riparametrizzazione data proprio dalla lunghezza d'arco. Infine, dalla derivata della funzione composta, si ottiene

$$\frac{d\psi}{ds}(s) = \frac{dt}{ds} \frac{d}{dt} \psi(s(t)) = \frac{\varphi'(t)}{s'(t)} = \frac{\varphi'(t)}{\|\varphi'(t)\|}.$$

□

Abbiamo anche la seguente proprietà, che ci dice che la lunghezza di una curva non dipende dalla parametrizzazione scelta.

Proposizione 5.7 *Siano $\varphi : I \rightarrow \mathbb{R}^N$ e $\psi : J \rightarrow \mathbb{R}^N$ due curve regolari equivalenti; allora*

$$l(\varphi, I) = l(\psi, J).$$

DIM. Supponiamo per comodità l'insieme $I = [a, b]$ e $J = [c, d]$ e sia $\alpha : [a, b] \rightarrow [c, d]$ una riparametrizzazione ammissibile; supponiamo anche $\alpha' > 0$, da cui si deduce che $\alpha(a) = c$ e $\alpha(b) = d$. Allora, tenendo conto che

$$\varphi'(t) = \frac{d}{dt} \varphi(t) = \frac{d}{dt} \psi(\alpha(t)) = \alpha'(t) \psi'(\alpha(t)),$$

si ottiene, mediante il cambio di variabile $\tau = \alpha(t)$,

$$l(\psi, [c, d]) = \int_c^d \|\psi'(\tau)\| d\tau = \int_a^b \frac{\|\varphi'(t)\|}{|\alpha'(t)|} \alpha'(t) dt = \int_a^b \|\varphi'(t)\| dt$$

in quanto $\alpha' > 0$.

□

Osservazione 5.8 Nel caso di curva $\varphi : [a, b] \rightarrow \mathbb{R}^2$ piana cartesiana rispetto ad esempio ad x , $\varphi(t) = (t, f(t))$ con $f : [a, b] \rightarrow \mathbb{R}$ di classe $C^1([a, b])$, si ottiene che la formula della lunghezza si riduce a

$$l(\varphi, [a, b]) = \int_a^b \sqrt{1 + f'(t)^2} dt$$

che fornisce quindi la formula della lunghezza del grafico della funzione f

Nel caso invece di curva in coordinate polari esplicite, da (5.1) si ottiene la formula

$$l(\varphi, [\vartheta_1, \vartheta_2]) = \int_{\vartheta_1}^{\vartheta_2} \sqrt{g(\vartheta)^2 + g'(\vartheta)^2} d\vartheta,$$

mentre nel caso di curva in coordinate polari parametriche, da (5.2) si ricava

$$l(\varphi, [I]) = \int_I \sqrt{r'(t)^2 + r(t)^2 \vartheta'(t)^2} dt.$$

5.2.1 Integrali curvilinei

Esistono due tipi di integrali curvilinei, a seconda che si voglia integrare funzioni scalari oppure funzioni vettoriali. Nel primo caso, daremo una definizione di integrale curvilineo che sarà indipendente dalla parametrizzazione della curva (e la dimostrazione è identica a quella fatta per l'invarianza della lunghezza di una curva per riparametrizzazioni ammissibili), mentre nel secondo caso avremo un integrale che dipende dall'orientazione della curva e cambierà segno se la curva viene orientata in senso opposto. Questi comportamenti si possono interpretare da un punto di vista fisico, essendo il primo tipo di integrale legato ad applicazioni come il calcolo di masse totali di elementi curvilinei con densità non omogenea, mentre il secondo tipo di integrale ha applicazioni al lavoro di campi vettoriali lungo curve, lavoro che è positivo o negativo a seconda che sia il campo ad effettuare lavoro o meno.

Definizione 5.9 Si definisce integrale curvilineo di primo tipo o per funzioni scalari l'integrale definito per funzioni continue $f : A \rightarrow \mathbb{R}$ con $A \subset \mathbb{R}^N$ e curve regolari $\varphi : I \rightarrow \mathbb{R}^N$ con $\varphi(I) \subset A$ mediante la formula

$$\int_{\varphi} f ds = \int_I f(\varphi(t)) \|\varphi'(t)\| dt.$$

Si definisce invece integrale curvilineo di secondo tipo o per funzioni vettoriali l'integrale definito per funzioni continue $F : A \rightarrow \mathbb{R}^N$ con $A \subset \mathbb{R}^N$ e curve regolari $\varphi : I \rightarrow \mathbb{R}^N$ con $\varphi(I) \subset A$ mediante la formula

$$\int_{\varphi} F \cdot d\vec{s} = \int_I F(\varphi(t)) \cdot \varphi'(t) dt,$$

dove con \cdot abbiamo indicato il prodotto scalare in \mathbb{R}^N .

Osservazione 5.10 Si nota che l'integrale di secondo tipo può essere ricondotto ad un integrale di primo tipo se si considera la funzione

$$f(\varphi(t)) = F(\varphi(t)) \cdot T_{\varphi}(t);$$

è importante osservare che tale funzione cambia segno se si riparametrizza la curva in modo che cambi l'orientazione.

Per gli integrali curvilinei valgono proprietà analoghe a quelle che valgono per gli integrali di funzioni reali di variabile reale; elenchiamo qui le principali proprietà, scrivendole solo per l'integrale di primo tipo, ma notando che valgono anche per quelle di secondo tipo.

Siano $f, g : A \rightarrow \mathbb{R}$ due funzioni continue, $a, b \in \mathbb{R}$ e $\varphi : I \rightarrow A$: allora

$$\int_{\varphi} (af + bg) ds = a \int_{\varphi} f ds + b \int_{\varphi} g ds;$$

mentre

$$\left| \int_{\varphi} f ds \right| \leq \int_{\varphi} |f| ds \leq \sup_{t \in I} |f(\varphi(t))| l(\varphi, I).$$

Capitolo 6

Dalla continuità alla differenziabilità

In questo capitolo, utilizzando quanto sviluppato nel capitolo relativo alle curve, passeremo ad uno studio più sistematico delle funzioni di più variabili, investigando con maggiori dettagli la continuità per arrivare ad estendere alle funzioni di più variabili il concetto di derivabilità.

6.1 Sezioni e insiemi connessi

Investighiamo qui la continuità di una funzione data tramite riduzione dimensionale della stessa; tale operazione prende il nome di sezione della funzione. Consideriamo qui funzioni scalari; la continuità di una funzione vettoriale si ottiene sezionando le funzioni scalari che ne definiscono le componenti.

Data una funzione $f : A \rightarrow \mathbb{R}$ con $A \subset \mathbb{R}^N$ e $\varphi : I \rightarrow A$ una curva, dalla Proposizione 4.14 si ha che la funzione di una variabile $f_\varphi : I \rightarrow \mathbb{R}$, $f_\varphi(t) = f(\varphi(t))$ è una funzione continua se f è continua. La funzione f_φ prende anche il nome di sezione di f lungo la curva φ ed ha la proprietà di ridurre la complessità di f facendola diventare funzione di una sola variabile. Un caso particolare di sezioni si ottiene fissando un punto $x_0 \in A$ e considerando la curva $\varphi : I \rightarrow A$, $\varphi(t) = x_0 + te_i$, dove $\{e_1, \dots, e_N\}$ sono gli elementi della base canonica di \mathbb{R}^N . In tal caso le sezioni $f_{e_i}^{x_0}(t) = f(x_0 + te_i)$ vengono chiamate sezioni coordinate i -esime di f uscenti da x_0 .

Abbiamo quindi osservato che se una funzione è continua, allora ogni sua sezione è continua. Ci si può chiedere se è vero il viceversa; notiamo anzitutto che per avere ciò, occorre considerare tutte le possibili curve, ma come vedremo anche questo non sarà sufficiente a meno che non si richieda la connessione per archi del dominio A (vedere Definizione 6.1).

Esempio 6.1

1. Si consideri la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

Si nota che la restrizione di f agli assi coordinati $x = 0$ e $y = 0$ è uguale a 0, mentre la restrizione di f alla curva $\varphi(t) = tv$, con $v = (1, m)$ è data da

$$f_{\varphi}(t) = f(t, mt) = \frac{m}{1 + m^2}$$

e tale limite è non nullo se $m \neq 0$. Quindi la funzione non è continua.

2. L'esempio precedente può essere modificato per mostrare che non bastano neanche le sezioni lungo qualsiasi direzione; infatti la funzione

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 - y^2} & \text{se } |x| \neq |y| \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

Verifichiamo la continuità in $(0, 0)$; le sezioni coordinate sono nulle, mentre le sezioni lungo le curve $\varphi(t) = tv$, con $v = (1, m)$, $m \neq \pm 1$, producono le funzioni

$$f_{\varphi}(t) = \frac{mt}{1 - m^2}$$

e tale funzione tende a zero per $t \rightarrow 0$. Però, se si considera la curva $\varphi(t) = (t, t + t^2)$, si ottiene la sezione

$$f_{\varphi}(t) = \frac{t^3 + t^4}{-2t^3 - t^4} \rightarrow -\frac{1}{2}, \quad \text{per } t \rightarrow 0,$$

da cui la non continuità di f .

Definizione 6.1 (Insieme connesso per archi, convesso e stellato) *Un sottoinsieme A di \mathbb{R}^N si dice connesso per archi se per ogni $x, y \in A$ esiste una curva $\varphi : [a, b] \rightarrow A$ tale che $\varphi(a) = x$ e $\varphi(b) = y$ (diremo che φ connette x e y). L'insieme si dice convesso se l'arco in considerazione è un segmento, cioè se per ogni $x, y \in A$, il segmento $[x, y]$ è contenuto in A . Infine, A si dice stellato in $x_0 \in A$ se per ogni $x \in A$ il segmento $[x_0, x]$ è contenuto in A .*

Osservazione 6.2 Il concetto di connessione può anche essere dato senza l'utilizzo delle curve; in tal caso si ottiene una nozione topologica di connessione. Un insieme A si dice connesso se presi due aperti A_1 ed A_2 con

$$A = (A \cap A_1) \cup (A \cap A_2), \quad A_1 \cap A_2 = \emptyset,$$

allora o $A_1 = \emptyset$ oppure $A_2 = \emptyset$; cioè A è connesso se non lo si può partizionare utilizzando due insiemi aperti. Tuttavia, per insiemi aperti, la connessione e la connessione per archi coincidono.

Grazie alla precedente definizione possiamo enunciare la seguente proposizione.

Proposizione 6.3 *Sia $f : A \rightarrow \mathbb{R}$ con A connesso per archi; allora f è continua in A se e solo se la restrizione di f lungo ogni curva contenute in A è continua.*

Possiamo anche ricavare il Teorema dei valori intermedi per una funzione continua in domini connessi per archi.

Proposizione 6.4 Sia $f : A \rightarrow \mathbb{R}$ una funzione continua e A un sottoinsieme di \mathbb{R}^N connesso per archi; allora, per ogni $x, y \in A$, la funzione f assume tutti i valori compresi tra $f(x)$ ed $f(y)$.

DIM. Fissiamo $x, y \in A$ e consideriamo una curva $\varphi : [a, b] \rightarrow A$ che connette x con y ; basta quindi applicare il teorema dei valori intermedi alla funzione di una variabile f_φ . \square

Osservazione 6.5 La proposizione precedente afferma l'esistenza di un valore intermedio su ogni curva che connette x con y ; in altri termini, se c è un varlore intermedio tra $f(x)$ e $f(y)$, allora l'insieme di livello $E_c = \{f = c\}$ disconnette A , cioè $A \setminus \{f = c\}$ non è più connesso per archi.

6.2 Derivabilità e differenziabilità per funzioni scalari

Introduciamo in questa sezione la nozione di derivabilità per funzioni di più variabili; la definizione che daremo dovrà avere buone proprietà, come ad esempio l'indipendenza dal sistema di riferimento scelto e il fatto che la derivabilità implichi la continuità. Un primo tentativo è quello che utilizza le sezioni della funzione f .

Definizione 6.6 Data una funzione $f : A \rightarrow \mathbb{R}$, $A \subset \mathbb{R}^N$, e $v \in \mathbb{R}^N$ con $\|v\| = 1$, definiamo derivata di f nel punto $x_0 \in A$ in direzione v la seguente quantità, qualora esista finita;

$$\lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t};$$

in tal caso si pone equivalentemente $\frac{\partial f}{\partial v}(x_0)$, $\partial_v f(x_0)$ o $D_v f(x_0)$ la derivata direzionale di f . In particolare scriveremo $\frac{\partial f}{\partial x_i}(x_0)$, $\partial_i f(x_0)$, $D_{x_i} f(x_0)$ o $D_i f(x_0)$ nel caso di derivate nelle direzioni coordinate e_i e parleremo di derivate parziali i -esime. Si definisce infine vettore gradiente di f il vettore le cui componenti sono le derivate parziali di f , che viene denotato usualmente con

$$\nabla f(x_0) = Df(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_N}(x_0) \right).$$

Osservazione 6.7 Nel caso $N = 2$ e $N = 3$ si conviene anche usare le variabili (x, y) e (x, y, z) rispettivamente, in tal caso si indica equivalentemente $\frac{\partial f}{\partial x}$, $\partial_x f$, $\frac{\partial f}{\partial y}$, $\partial_y f$ e $\frac{\partial f}{\partial z}$, $\partial_z f$.

La nozione appena introdotta non implica però in generale la continuità della funzione f , come è facile convincersi col seguente esempio.

Esempio 6.2 Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la funzione definita da

$$f(x, y) = \begin{cases} 1 & \text{se } 0 < y < x^2 \\ 0 & \text{altrimenti.} \end{cases}$$

Un calcolo diretto mostra che tutte le derivate direzionali esistono e sono uguali a 0, ma la funzione è chiaramente non continua.

È quindi facile intuire che serve un'altra nozione di derivata, che tra l'altro non dipenda dal sistema di riferimento scelto (che non necessiti, nella definizione, il dover fissare un sistema di riferimento).

Definizione 6.8 (Differenziale) Data una funzione $f : A \rightarrow \mathbb{R}$, diremo che f è differenziabile nel punto $x_0 \in A$ se esiste un'applicazione lineare $L : \mathbb{R}^N \rightarrow \mathbb{R}$ tale che

$$(6.1) \quad \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} = 0.$$

Il differenziale di f in x_0 viene anche denotato con $df(x_0)$ o con $d_{x_0}f$.

Nella seguente proposizione sono riassunte le principali proprietà delle funzioni differenziabili.

Proposizione 6.9 Sia $f : A \rightarrow \mathbb{R}$, $A \subset \mathbb{R}^N$ una funzione differenziabile in $x_0 \in A$. Allora:

1. f è continua in x_0 ;
2. esistono tutte le derivate direzionali in x_0 e la derivata direzionale ha dipendenza lineare rispetto alla direzione;
3. il differenziale di f coincide con il gradiente di f in x_0 ;
4. il differenziale è unico.

DIM. Per la continuità di f , basta riscrivere la condizione (6.1) in termini di o come segue;

$$f(x) = f(x_0) + L(x - x_0) + o(\|x - x_0\|),$$

da cui segue immediatamente che

$$\lim_{x \rightarrow x_0} f(x) = f(x_0).$$

Per il punto 2., dimostriamo che la derivata direzionale esiste per ogni $v \in \mathbb{R}^N$, $\|v\| = 1$ e vale $\frac{\partial f}{\partial v}(x_0) = L(v)$; infatti, dalla linearità di L si ottiene

$$\lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} - L(v) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0) - L(tv)}{t} = 0,$$

che era quanto si voleva dimostrare. In particolare, per la derivata parziale i -esima si ottiene

$$\frac{\partial f}{\partial x_i} f(x_0) = L(e_i), \quad \forall i = 1, \dots, N,$$

e quindi, in quanto l'applicazione lineare L è univocamente determinata dai valori assunti sugli elementi di una base, L è univocamente determinato dal vettore

$$(L(e_1), \dots, L(e_N)) = \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_N}(x_0) \right) = \nabla f(x_0).$$

Questo argomento mostra in particolare l'unicità del differenziale. □

Abbiamo quindi che la differenziabilità implica l'esistenza di tutte le derivate direzionali nel punto in considerazione, ma, come abbiamo visto, l'esistenza di tutte le derivate direzionali nel punto non implica la differenziabilità. Il seguente Teorema fornisce tuttavia il metodo che usualmente si utilizza per dimostrare la differenziabilità; esso ci garantisce la differenziabilità se le derivate parziali esistono non solo nel punto in considerazione, ma in un intorno del punto e sono ivi continue. Diremo quindi che $f : A \rightarrow \mathbb{R}$ è di classe C^1 in un intorno del punto $x_0 \in A$ se esiste un $r > 0$ tale che $f \in C^1(B_r(x_0))$, cioè se esistono le derivate parziali di f in $B_r(x_0)$ ed in tale insieme sono funzioni continue. Diremo infine che f è di classe C^1 in A ($f \in C^1(A)$) se f è di classe C^1 nell'intorno di ogni punto di A .

Teorema 6.10 (Differenziale totale) Sia $f : A \rightarrow \mathbb{R}$ una funzione di classe C^1 in un intorno del punto $x_0 \in A$; allora f è differenziabile in x_0 .

DIM. Per quanto è stato detto in precedenza, la differenziabilità in x_0 va verificata con $L = \nabla f(x_0)$. Per semplicità di notazione, supponiamo $N = 2$ ed indichiamo con (x_0, y_0) il punto in cui vogliamo verificare la differenziabilità e con (x, y) un punto generico. Dire che f è di classe C^1 in un intorno di (x_0, y_0) significa che le funzioni $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ esistono e sono continue; possiamo quindi applicare gli sviluppi di Taylor per la funzione $y \mapsto f(x, y)$ per ricavare che

$$f(x, y) = f(x, y_0) + \frac{\partial f}{\partial y}(x, y_0)(y - y_0) + o(y - y_0)$$

ed analogamente, la continuità di $x \mapsto \frac{\partial f}{\partial y}(x, y_0)$ per scrivere

$$\frac{\partial f}{\partial y}(x, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) + o(1).$$

Analogamente, applicando la formula di Taylor approssimata al primo ordine alla funzione $x \mapsto f(x, y_0)$ si ottiene che

$$f(x, y_0) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + o(x - x_0).$$

Questi conti mostrano che

$$\frac{f(x, y) - f(x_0, y_0) - \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0)}{\|(x - x_0, y - y_0)\|} = \frac{o(x - x_0) + o(y - y_0)}{\|(x - x_0, y - y_0)\|}.$$

Notiamo che, ad esempio:

$$\begin{aligned} \frac{o(x - x_0)}{\|(x - x_0, y - y_0)\|} &= \frac{o(x - x_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = \frac{|x - x_0|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \frac{o(x - x_0)}{|x - x_0|} \\ &\leq \frac{o(x - x_0)}{|x - x_0|}, \end{aligned}$$

da cui si deduce la differenziabilità di f in (x_0, y_0) . □

Vediamo ora alcune conseguenze della differenziabilità di una funzione; iniziamo con la seguente proposizione.

Proposizione 6.11 Sia $f : A \rightarrow \mathbb{R}$ una funzione di classe $C^1(A)$ e sia $\varphi : I \rightarrow A$ una curva regolare; allora la funzione composta $g = f \circ \varphi : I \rightarrow \mathbb{R}$ è derivabile e vale la seguente formula di derivazione della funzione composta

$$g'(t) = \nabla f(\varphi(t)) \cdot \varphi'(t).$$

DIM. Come per il teorema del differenziale totale, supponiamo $N = 2$ e scriviamo

$$\begin{aligned} f(\varphi_1(t+h), \varphi_2(t+h)) &= f(\varphi_1(t), \varphi_2(t)) + \frac{\partial f}{\partial x}(\varphi_1(t), \varphi_2(t))(\varphi_1(t+h) - \varphi_1(t)) + \\ &\quad + \frac{\partial f}{\partial y}(\varphi_1(t), \varphi_2(t))(\varphi_2(t+h) - \varphi_2(t)) + o(\varphi_1(t+h) - \varphi_1(t)) + \\ &\quad + o(\varphi_2(t+h) - \varphi_2(t)). \end{aligned}$$

L'asserto segue quindi dal rapporto incrementale per la funzione g e passando al limite per $h \rightarrow 0$. □

Teorema 6.12 *Sia $f : A \rightarrow \mathbb{R}$ una funzione di classe $C^1(A)$ e A insieme connesso per archi; se $\nabla f(x) = 0$ per ogni $x \in A$, allora f è costante in A .*

DIM. Basta dimostrare che per ogni $x, y \in A$, $f(x) = f(y)$. Consideriamo quindi un arco $\varphi : [a, b] \rightarrow A$ che connette x con y e la funzione $g(t) = f(\varphi(t))$; $g(t)$ è derivabile e vale

$$g'(t) = \nabla f(\varphi(t)) \cdot \varphi'(t) = 0$$

da cui g costante e quindi $f(x) = f(y)$. \square

Ricaviamo ora l'equazione dello spazio tangente al grafico di una funzione differenziabile; ricordiamo che il grafico di una funzione $f : A \rightarrow \mathbb{R}$ è dato dal luogo di punti

$$G(f) = \{(x, f(x)) : x \in A\} \subset A \times \mathbb{R}$$

e quindi in generale il grafico di una funzione è un sottoinsieme di \mathbb{R}^{N+1} . Se consideriamo una sezione coordinata $f_i(t) = f(x_0 + te_i)$, l'equazione della retta tangente al grafico di f_i nel punto $t = 0$ è data da

$$r_i(t) = f(x_0) + tf'_i(0) = f(x_0) + t \frac{\partial f}{\partial x_i}(x_0).$$

Lo spazio tangente al grafico di f passante per il punto $(x_0, f(x_0))$ sarà quindi descritto dall'equazione

$$x_{N+1} = a \cdot x + b, \quad a \in \mathbb{R}^N, b \in \mathbb{R};$$

imponendo la condizione che tutte le rette r_i siano contenute in tale spazio, si ottiene l'equazione

$$x_{N+1} = f(x_0) + \nabla f(x_0) \cdot (x - x_0).$$

Nel caso $N = 2$ si ottiene l'equazione di un piano nello spazio, chiamato piano tangente al grafico della funzione in $(x_0, f(x_0))$. Il versore

$$\nu(x_0) = \frac{(-\nabla f(x_0), 1)}{\sqrt{1 + \|\nabla f(x_0)\|^2}}$$

è il versore perpendicolare allo spazio tangente, scelto in modo da avere componente positiva lungo l'asse verticale x_{N+1} .

Vediamo ora quale relazione esiste tra il gradiente di una funzione e il suo insieme di livello. Anzitutto, ha senso chiedersi se l'insieme di livello, nel caso $N = 2$, sia una curva o meno; questo segue, almeno localmente, se si richiede $\nabla f \neq 0$ sull'insieme stesso.

Abbiamo il seguente teorema.

Teorema 6.13 (Del Dini o della funzione implicita) *Sia $f : A \rightarrow \mathbb{R}$ con $A \subset \mathbb{R}^2$ una funzione di classe $C^1(A)$ e $(x_0, y_0) \in E_c = \{f = c\}$ con $c \in \mathbb{R}$; se $\nabla f(x_0, y_0) \neq 0$, esiste $\varepsilon > 0$ tale che $E_c \cap B_\varepsilon(x_0)$ è una curva cartesiana regolare.*

DIM. Notiamo anzitutto che, eventualmente sostituendo f con $g = f - c$, possiamo supporre senza ledere in generalità che $c = 0$. Supponiamo si abbia $\frac{\partial f}{\partial y}(x_0, y_0) > 0$; la continuità di $\frac{\partial f}{\partial y}$ implica l'esistenza di $\delta_0 > 0$ tale che $\frac{\partial f}{\partial y}(x, y) > 0$ per ogni $(x, y) \in B_{\delta_0}(x_0, y_0)$. La funzione $y \mapsto f(x_0, y)$ è strettamente monotona crescente in un intorno di y_0 e nulla per $y = y_0$. Indichiamo con $\delta_1 > 0$, $\delta_1 < \delta_0/2$, un numero positivo in modo tale che

$f(x_0, \cdot)$ sia monotona in $(y_0 - \delta_1, y_0 + \delta_1)$. Dalla continuità di f e dal fatto che $f(x_0, y_0 - \delta_1) < 0$, $f(x_0, y_0 + \delta_1) > 0$, deduciamo l'esistenza di un $\delta_2 > 0$, $\delta_2 < \delta_0/2$, tale che per ogni $x \in (x_0 - \delta_2, x_0 + \delta_2)$ valga ancora $f(x, y_0 - \delta_1) < 0$ e $f(x, y_0 + \delta_1) > 0$. Per come sono stati scelti i parametri, abbiamo che per ogni $x \in (x_0 - \delta_2, x_0 + \delta_2)$ la funzione $y \mapsto f(x, y)$ è strettamente monotona crescente, strettamente negativa per $y = y_0 - \delta_1$ e strettamente positiva per $y = y_0 + \delta_1$. Quindi, esiste un unico punto $y = h(x)$ tale che $f(x, h(x)) = 0$ (si veda Figura 6.1). Abbiamo così definito la funzione $h : (x_0 - \delta_2, x_0 + \delta_2) \rightarrow (y_0 - \delta_1, y_0 + \delta_1)$, che prende

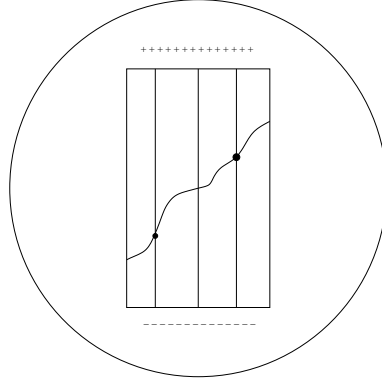


Figura 6.1: Costruzione della funzione implicita.

il nome di funzione implicita, che parametrizza $E_c \cap (x_0 - \delta_2, x_0 + \delta_2) \times (y_0 - \delta_1, y_0 + \delta_1)$. Tale funzione è continua e derivabile; dalla relazione $0 = f(x, h(x))$, se ne deduce che

$$0 = \frac{d}{dx} f(x, h(x)) = \frac{\partial f}{\partial x}(x, h(x)) + \frac{\partial f}{\partial y}(x, h(x))h'(x),$$

cioè la formula

$$h'(x) = -\frac{\frac{\partial f}{\partial x}(x, h(x))}{\frac{\partial f}{\partial y}(x, h(x))}.$$

□

Osservazione 6.14 Il Teorema 6.13 resta valido anche per dimensione arbitraria ed è noto semplicemente come Teorema della funzione implicita. La funzione implicita definisce una ipersuperficie cartesiana regolare, che nel caso $N = 3$ è una superficie cartesiana regolare. La nozione di superficie parametrizzata e superficie cartesiana può essere data in modo da mimare la definizione data per le curve, ma il loro studio va oltre gli obiettivi di questo corso. Si può inoltre vedere che, nel caso di dimensione N generica, se $\frac{\partial f}{\partial x_N}(x_0) \neq 0$, l'insieme di livello è localmente attorno ad x_0 il grafico $x_N = h(x_1, \dots, x_{N-1})$ e per le derivate di h si ha

$$(6.2) \quad \frac{\partial h}{\partial x_i}(x_1, \dots, x_{N-1}) = -\frac{\frac{\partial f}{\partial x_i}(x)}{\frac{\partial f}{\partial x_N}(x)}, \quad i = 1, \dots, x_{N-1}.$$

In particolare, se siamo in un punto $x_0 \in E_c$, avremo che localmente E_c è dato dal grafico di h ; se scriviamo $x_0 = (x'_0, x_{0,N})$, $x'_0 \in \mathbb{R}^{N-1}$, come abbiamo visto, il piano tangente al

grafico di h nel punto x'_0 è individuato dal versore normale

$$\nu = \frac{1}{\sqrt{1 + \|\nabla h(x'_0)\|^2}}(-\nabla h(x'_0), 1)$$

che grazie a (6.2), a meno del segno, coincide con $\frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$.

Tramite il Teorema della funzione implicita si può dedurre la seguente relazione tra gli insiemi di livello e il gradiente.

Proposizione 6.15 *Sia $f : A \rightarrow \mathbb{R}$ una funzione di classe $C^1(A)$ e sia $x_0 \in E_c$, $c \in \mathbb{R}$ con $\nabla f(x_0) \neq 0$. Allora $\nabla f(x_0)$ è ortogonale ad E_c in x_0 , cioè per ogni curva $\varphi : (-\varepsilon, \varepsilon) \rightarrow E_c$ con $\varphi(0) = x_0$ si ha che*

$$\nabla f(x_0) \cdot \varphi'(0) = 0.$$

DIM. Grazie al Teorema della funzione implicita, esistono un $\varepsilon > 0$ ed una curva regolare $\varphi : (-\varepsilon, \varepsilon) \rightarrow E_c$ con $\varphi(0) = x_0$. Siccome tale curva è contenuta in E_c si ha che $c = f(\varphi(t)) = g(t)$ per ogni $t \in (-\varepsilon, \varepsilon)$, quindi $g'(t) = 0$ ed in particolare:

$$\nabla f(x_0) \cdot \varphi'(0) = 0,$$

e quindi, dato che $\nabla f(x_0) \neq 0$ e φ è regolare, ne segue che $\varphi'(0)$ e $\nabla f(x_0)$ sono ortogonali, che era quanto volevamo dimostrare. \square

6.3 Differenziabilità per funzioni vettoriali

Estendiamo qui il concetto di differenziabilità alle funzioni vettoriali.

Definizione 6.16 (Differenziabilità per funzioni vettoriali) *Data una funzione $F : A \rightarrow \mathbb{R}^M$ con $A \subset \mathbb{R}^N$ e $x_0 \in A$, diremo che F è differenziabile in x_0 se esiste un'applicazione lineare $L : \mathbb{R}^N \rightarrow \mathbb{R}^M$ tale che*

$$(6.3) \quad \lim_{x \rightarrow x_0} \frac{F(x) - F(x_0) - L(x - x_0)}{\|x - x_0\|} = 0.$$

In tal caso chiameremo L differenziale di F in x_0 e lo denoteremo con $dF(x_0)$ o $d_{x_0}F$.

Decomponendo la funzione F nelle sue componenti $F(x) = (F_1(x), \dots, F_M(x))$, possiamo riscrivere la (6.3) in modo equivalente richiedendo che

$$\lim_{x \rightarrow x_0} \frac{F_i(x) - F_i(x_0) - L_i(x - x_0)}{\|x - x_0\|} = 0, \quad \forall i = 1, \dots, M,$$

dove con L_i indichiamo la i -esima riga della matrice che rappresenta l'applicazione lineare L . Quindi, richiedere la differenziabilità di F equivale a richiedere la differenziabilità delle sue componenti. Se ne deduce che il differenziale L di F è dato dalla matrice $M \times N$

$$L = DF(x_0) = JF(x_0) = \begin{pmatrix} \nabla F_1(x_0) \\ \vdots \\ \nabla F_M(x_0) \end{pmatrix} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x_0) & \dots & \frac{\partial F_1}{\partial x_N}(x_0) \\ \vdots & \dots & \vdots \\ \frac{\partial F_M}{\partial x_1}(x_0) & \dots & \frac{\partial F_M}{\partial x_N}(x_0) \end{pmatrix}$$

che prende il nome di *Matrice Jacobiana*. Possiamo enunciare il seguente risultato.

Proposizione 6.17 *Siano $F : A \rightarrow B$, $A \subset \mathbb{R}^N$, $B \subset \mathbb{R}^M$ una funzione differenziabile in x_0 e $G : B \rightarrow \mathbb{R}^K$ una funzione differenziabile in $F(x_0)$; allora la funzione $H = G \circ F : A \rightarrow \mathbb{R}^K$ è differenziabile in x_0 e si ha*

$$DH(x_0) = JH(x_0) = D(G \circ F)(x_0) = DG(F(x_0)) \cdot DF(x_0)$$

dove nell'ultimo passaggio col simbolo \cdot si intende il prodotto matriciale righe per colonne.

Chiudiamo la sezione enunciando la versione vettoriale del Teorema del Dini.

Teorema 6.18 (Teorema della funzione implicita) *Sia $F : E \rightarrow \mathbb{R}^M$, $E \subset \mathbb{R}^N \times \mathbb{R}^M$ e sia $(x_0, y_0) \in E$ un punto tale che $F(x_0, y_0) = 0$; se F è di classe C^1 in un intorno del punto (x_0, y_0) e $D_y F(x_0, y_0)$ (matrice $M \times M$ Jacobiana di F ottenuta facendo le derivate solo rispetto alle ultime M variabili), allora esiste un intorno aperto U di x_0 e V di y_0 e una funzione $H : U \rightarrow V$ di classe $C^1(U)$ tale che $F(x, H(x)) = 0$. Lo Jacobiano di H è dato dalla formula*

$$DH(x) = -D_y F(x, H(x))^{-1} \cdot D_x F(x, H(x)).$$

6.3.1 Cambiamenti di coordinate: polari, cilindriche e sferiche

Concludiamo questo capitolo con alcuni richiami sul cambiamento di coordinate, presentando i principali e più importanti esempi.

Definizione 6.19 *Una funzione $F : A \rightarrow B$ con $A, B \subset \mathbb{R}^N$ si definisce cambio di variabili o di coordinate se:*

1. F è iniettiva;
2. F è suriettiva;
3. $\det DF(x) \neq 0$ per ogni $x \in A$.

Coordinate polari nel piano Definiamo la funzione $F : (0, +\infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$ tramite $F(\varrho, \vartheta) = (\varrho \cos \vartheta, \varrho \sin \vartheta)$. Tale funzione consiste nel cambio di variabili

$$\begin{cases} x = \varrho \cos \vartheta \\ y = \varrho \sin \vartheta. \end{cases}$$

Avendo scelto $\varrho > 0$, si nota subito che la funzione F è iniettiva e suriettiva (ammettendo il caso $\varrho = 0$ si avrebbe mancanza di iniettività). Inoltre,

$$DF(\varrho, \vartheta) = \begin{pmatrix} \cos \vartheta & -\varrho \sin \vartheta \\ \sin \vartheta & \varrho \cos \vartheta \end{pmatrix}$$

da cui si deduce che $\det DF(\varrho, \vartheta) = \varrho \neq 0$, e quindi F soddisfa tutti i requisiti di un cambiamento di variabili.

Coordinate cilindriche nello spazio Sia $F : (0, +\infty) \times [0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}^3 \setminus (\{(0, 0)\} \times \mathbb{R})$ definita tramite $F(\varrho, \vartheta, t) = (\varrho \cos \vartheta, \varrho \sin \vartheta, t)$ che corrisponde al cambio di coordinate

$$\begin{cases} x = \varrho \cos \vartheta \\ y = \varrho \sin \vartheta \\ z = t. \end{cases}$$

Tale funzione è iniettiva e suriettiva, mentre

$$DF(\varrho, \vartheta, t) = \begin{pmatrix} \cos \vartheta & -\varrho \sin \vartheta & 0 \\ \sin \vartheta & \varrho \cos \vartheta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

da cui $\det DF(\varrho, \vartheta, t) = \varrho \neq 0$.

Coordinate sferiche nello spazio Sia $F : (0, +\infty) \times [0, 2\pi) \times (0, \pi) \rightarrow \mathbb{R}^3 \setminus (\{(0, 0)\} \times \mathbb{R})$ definita da $F(\varrho, \vartheta, \varphi) = (\varrho \cos \vartheta \sin \varphi, \varrho \sin \vartheta \sin \varphi, \varrho \cos \varphi)$ che corrisponde al cambio di coordinate

$$\begin{cases} x = \varrho \cos \vartheta \sin \varphi \\ y = \varrho \sin \vartheta \sin \varphi \\ z = \varrho \cos \varphi. \end{cases}$$

Tale funzione è iniettiva e suriettiva, mentre

$$DF(\varrho, \vartheta, \varphi) = \begin{pmatrix} \cos \vartheta \sin \varphi & -\varrho \sin \vartheta \sin \varphi & \varrho \cos \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi & \varrho \cos \vartheta \sin \varphi & \varrho \sin \vartheta \cos \varphi \\ \cos \varphi & 0 & -\varrho \sin \varphi \end{pmatrix}$$

da cui $\det DF(\varrho, \vartheta, \varphi) = -\varrho^2 \sin \varphi \neq 0$.

Osservazione 6.20 Abbiamo presentato qui gli esempi più comuni di cambiamento di coordinate; tutte le funzioni che soddisfano le ipotesi delle Definizione 6.19 possono essere considerate cambiamento di coordinate. Così ad esempio, se $A = \{(x, y) \in \mathbb{R}^2 : x, y > 0\}$, la funzione $F : A \rightarrow \mathbb{R}^2$ $F(x, y) = (xy, x/y)$ definisce un cambio di variabili.

Capitolo 7

Integrali multipli

In questo capitolo ci occupiamo della definizione degli integrali multipli, cioè della teoria dell'integrazione per funzioni di più variabili. Considereremo solo funzioni scalari, in quanto l'integrale per funzioni vettoriali sarà semplicemente dato dall'integrale delle sue componenti.

L'integrazione per funzioni di più variabili presenta maggiori complicazioni dell'integrazione in una variabile; una teoria soddisfacente venne sviluppata solo negli anni '40 del secolo scorso da Lebesgue, teoria che passa per una trattazione sistematica della teoria della misura. Non ci occuperemo qui di tale teoria, ma presenteremo solo l'integrazione secondo Riemann, in quanto ci interesserà integrare funzioni continue su insiemi con qualche regolarità.

7.1 Integrale di Riemann

Sia $f : A \rightarrow \mathbb{R}$, $A \subset \mathbb{R}^N$ una funzione continua e consideriamo un parallelepipedo $I \subset A$ del tipo $I = [a_1, b_1] \times \dots \times [a_N, b_N]$. Consideriamo una partizione \mathcal{P} di I in parallelepipedi $I_{i_1, \dots, i_N} \in \mathcal{P}$, $i_j = 0, \dots, k_j$ del tipo

$$I_{i_1, \dots, i_N} = [x_{i_1-1}^{(1)}, x_{i_1}^{(1)}] \times \dots \times [x_{i_N-1}^{(N)}, x_{i_N}^{(N)}]$$

con $(x_{i_j}^{(j)})_{i_j=0, \dots, k_j}$ suddivisione dell'intervallo $[a_j, b_j]$, $j = 1, \dots, N$; tale parallelepipedo ha misura data da

$$|I_{i_1, \dots, i_N}| = (x_{i_1}^{(1)} - x_{i_1-1}^{(1)}) \cdot \dots \cdot (x_{i_N}^{(N)} - x_{i_N-1}^{(N)}).$$

Chiameremo inoltre diametro della partizione \mathcal{P} la quantità

$$\text{diam}(\mathcal{P}) = \max\{|x_{i_j}^{(j)} - x_{i_j-1}^{(j)}| : j = 1, \dots, N, i_j = 1, \dots, k_j\}.$$

Definiamo quindi la somma integrale di f relativa a tale partizione come

$$S(f, \mathcal{P}) = \sum_{i_1=1}^{k_1} \dots \sum_{i_N=1}^{k_N} f(x_{i_1}^{(1)}, \dots, x_{i_N}^{(N)}) |I_{i_1, \dots, i_N}|.$$

La continuità di f implica l'esistenza del seguente limite

$$\lim_{\text{diam}(\mathcal{P}) \rightarrow 0} S(f, \mathcal{P}) = \int_I f(x) dx$$

che chiameremo integrale di Riemann di f sul parallelepipedo I .

Se abbiamo due parallelepipedi $I, J \subset A$ con parte interna di $I \cap J$ vuota (con questo ammettiamo che i due parallelepipedi siano adiacenti, ma la loro intersezione avviene solo lungo le superfici laterali), allora l'integrale di f su $I \cup J$ è definito tramite

$$\int_{I \cup J} f(x) dx = \int_I f(x) dx + \int_J f(x) dx.$$

Analogamente, se P è un *plurirettangolo*, cioè $P = I_1 \cup \dots \cup I_k$ con $I_j \subset A$ parallelepipedi con parte interna di $I_i \cap I_j$ vuota per $i \neq j$, allora l'integrale di f su P è definito tramite

$$\int_P f(x) dx = \sum_{j=1}^k \int_{I_j} f(x) dx.$$

La misura del plurirettangolo P è inoltre data da

$$|P| = \sum_{j=1}^k |I_j|.$$

Si usano i plurirettangoli per piastrellare generici insiemi $E \subset A$; diamo le seguenti definizioni.

Definizione 7.1 (Insieme misurabile secondo Peano-Jordan) *Un insieme compatto $E \subset \mathbb{R}^N$ si dice misurabile secondo Peano-Jordan se esiste una successione di plurirettangoli $(P_h)_{h \in \mathbb{N}}$ invadente, cioè tali che $P_h \subset P_{h+1}$ e $E = \bigcup_{h \in \mathbb{N}} P_h$. Si pone in tal caso*

$$(7.1) \quad |E| = \lim_{h \rightarrow +\infty} |P_h|.$$

Un insieme generico $E \subset \mathbb{R}^N$ si dice infine misurabile secondo Peano-Jordan se esiste una successione di insiemi compatti misurabili $(E_h)_{h \in \mathbb{N}}$ invadenti, cioè $E_h \subset E_{h+1}$ tali che $E = \bigcup_{h \in \mathbb{N}} E_h$; in tal caso si pone

$$(7.2) \quad |E| = \lim_{h \rightarrow +\infty} |E_h|.$$

Osservazione 7.2 Si può dimostrare che la definizione di misura data da (7.1) e (7.2) non dipende dall'insieme invadente. Non tutti gli insiemi compatti sono misurabili secondo Peano-Jordan, ma si può vedere che condizione necessaria e sufficiente è che la misura di ∂E sia nulla. Una tale condizione è ad esempio soddisfatta se ∂E è regolare, o regolare a tratti.

Definizione 7.3 (Integrale di Riemann su insiemi misurabili) *Sia $f : A \rightarrow \mathbb{R}$ una funzione continua; se $E \subset A$ è un insieme compatto Peano-Jordan misurabile, allora si definisce l'integrale di Riemann di f su E come*

$$\int_E f(x) dx = \lim_{h \rightarrow +\infty} \int_{P_h} f(x) dx$$

dove $(P_h)_{h \in \mathbb{N}}$ è una successione di plurirettangoli invadenti E . Se $E \subset A$ è un generico insieme Peano-Jordan misurabile, diremo che f è integrabile in senso generalizzato su E se esiste $M > 0$ tale che per ogni successione $(E_h)_{h \in \mathbb{N}}$ invadente E vale

$$\int_{E_h} |f(x)| dx \leq M;$$

in tal caso si pone

$$\int_E f(x)dx = \lim_{h \rightarrow +\infty} \int_{E_h} f(x)dx.$$

Osservazione 7.4 Si può dimostrare che le definizioni date non dipendono dalle particolari successioni invadenti considerate; inoltre, si può estendere in maniera ovvia le definizioni date alle funzioni continue a tratti su A .

Notiamo infine che se $f(x) = 1$, allora l'integrale di f su E determina la misura di E

$$\int_E f(x)dx = \int_E dx = |E|;$$

nel caso $N = 2$, invece di misura di E si parla solitamente di area di E e la si denota con $A(E)$ o $\text{Area}(E)$, mentre per $N = 3$ si parla solitamente di volume di E , denotato con $V(E)$ o $\text{Vol}(E)$.

L'integrale di Riemann gode delle consuete proprietà dell'integrale di una variabile, e più precisamente:

1. se $f, g : A \rightarrow \mathbb{R}$ sono continue, E è misurabile e f e g sono integrabili su E , allora, per ogni $a, b \in \mathbb{R}$, $af + bg$ è integrabile su E e vale

$$\int_E (af(x) + bg(x))dx = a \int_E f(x)dx + b \int_E g(x)dx;$$

2. se $f, g : A \rightarrow \mathbb{R}$ sono continue, E è misurabile, f e g integrabili su E e $f \leq g$ su E , allora

$$\int_E f(x)dx \leq \int_E g(x)dx;$$

3. se $f : A \rightarrow \mathbb{R}$ è continua, E è misurabile e f integrabile su E , allora

$$\left| \int_E f(x)dx \right| \leq \int_E |f(x)|dx \leq |E| \sup_{x \in E} |f(x)|;$$

4. se $f : A \rightarrow \mathbb{R}$ è continua, E_1, E_2 sono misurabili e f integrabile su $E_1 \cup E_2$, allora f è integrabile su E_1 e su E_2 separatamente e vale

$$\int_{E_1 \cup E_2} f(x)dx + \int_{E_1 \cap E_2} f(x)dx = \int_{E_1} f(x)dx + \int_{E_2} f(x)dx.$$

In particolare, se $E_1 \cap E_2$ ha parte interna vuota, allora

$$\int_{E_1 \cup E_2} f(x)dx = \int_{E_1} f(x)dx + \int_{E_2} f(x)dx.$$

7.2 Insiemi normali e calcolo degli integrali multipli

Nella sezione precedente abbiamo dato la definizione di integrale multiplo mentre nella presente sezione ci occupiamo del calcolo degli integrali multipli. Presentiamo qui metodi per il calcolo di integrali doppi e tripli che corrispondono ai casi $N = 2$ ed $N = 3$ rispettivamente; tali metodi potranno estendersi ad integrali multipli generici mediante un processo induttivo.

7.2.1 Integrali doppi

Diamo la seguente definizione.

Definizione 7.5 (Insiemi normali) *Un insieme compatto $E \subset \mathbb{R}^2$ si dice normale rispetto all'asse x se è del tipo*

$$(7.3) \quad E = \{(x, y) : x \in [a, b], \alpha(x) \leq y \leq \beta(x)\}$$

con $\alpha, \beta : [a, b] \rightarrow \mathbb{R}$ funzioni continue con $\alpha(x) \leq \beta(x)$ per ogni $x \in [a, b]$. Si definisce in modo analogo un insieme normale rispetto all'asse y , mentre diremo genericamente che E è un insieme normale se è normale rispetto ad uno dei due assi.

Per insiemi normali abbiamo il seguente metodo di calcolo degli integrali doppi, di cui non daremo la dimostrazione.

Teorema 7.6 *Sia $E \subset \mathbb{R}^2$ un insieme normale rispetto all'asse x come in (7.3); se $f : E \rightarrow \mathbb{R}$ è continua ed E è misurabile, allora*

$$(7.4) \quad \int_E f(x, y) dx dy = \int_a^b \left(\int_{\alpha(x)}^{\beta(x)} f(x, y) dy \right) dx.$$

Osservazione 7.7 La formula (7.4) la scriveremo anche, per comodità, nei seguenti modi

$$\int_E f(x, y) dx dy = \int_a^b dx \int_{\alpha(x)}^{\beta(x)} f(x, y) dy = \int_a^b \int_{\alpha(x)}^{\beta(x)} f(x, y) dy dx.$$

Nel caso di $f = 1$, l'area di E diventa

$$\text{Area}(E) = \int_E dx dy = \int_a^b \int_{\alpha(x)}^{\beta(x)} dy dx = \int_a^b (\beta(x) - \alpha(x)) dx$$

che è l'area della regione del piano compresa tra i due grafici di α e β .

Una formula analoga vale anche nel caso di insieme normale rispetto all'asse y ; può accadere che un insieme sia normale rispetto ad entrambi gli assi, ma anche che non sia normale rispetto a nessuno dei due assi. Nel primo caso si hanno due modi per calcolare l'integrale e la scelta può dipendere semplicemente dalla facilità dei calcoli. Nel secondo caso, generalmente decompone l'insieme in insiemi normali su cui si sa fare i conti, oppure effettua un cambio di coordinate, come vedremo nella Sezione 7.3.

7.2.2 Integrali tripli

Diamo le seguenti definizioni.

Definizione 7.8 (Insiemi normali ed insiemi stratificati) *Un insieme compatto $E \subset \mathbb{R}^3$ si dice normale rispetto al piano xy se è del tipo*

$$(7.5) \quad E = \{(x, y, z) : (x, y) \in D, \phi(x, y) \leq z \leq \psi(x, y)\}$$

con D insieme normale di \mathbb{R}^2 , $\phi, \psi : D \rightarrow \mathbb{R}$ funzioni continue con $\phi(x, y) \leq \psi(x, y)$ per ogni $(x, y) \in D$. Si definisce in modo analogo un insieme normale rispetto al piano yz e xz ,

mentre diremo genericamente che E è un insieme normale se è normale rispetto ad uno dei due assi.

Diremo infine che un insieme compatto $E \subset \mathbb{R}^3$ è stratificato in direzione z se è del tipo

$$(7.6) \quad E = \{(x, y, z) : z \in [a, b], (x, y) \in D_z\}$$

con D_z insieme normale del piano. Definizioni analoghe si hanno per insieme stratificato in direzione x ed y , mentre diremo genericamente che E è stratificato se lo è rispetto ad una delle tre direzioni x , y o z .

Per insiemi normali e stratificati abbiamo i seguenti metodi di calcolo degli integrali tripli, di cui non daremo la dimostrazione.

Teorema 7.9 Sia $E \subset \mathbb{R}^3$ un insieme normale rispetto al piano xy come in (7.5); se $f : E \rightarrow \mathbb{R}$ è continua ed E è misurabile, allora

$$(7.7) \quad \int_E f(x, y, z) dx dy dz = \int_D \left(\int_{\phi(x, y)}^{\psi(x, y)} f(x, y, z) dz \right) dx dy, \quad (\text{integrazione per "fili"});$$

se invece E è stratificato in direzione z come in (7.6), allora

$$(7.8) \quad \int_E f(x, y, z) dx dy dz = \int_a^b \left(\int_{D_z} f(x, y, z) dx dy \right) dz \quad (\text{integrazione per "strati"}).$$

Osservazione 7.10 La formula di integrazione per fili (7.7) la scriveremo anche, per comodità, nei seguenti modi

$$\int_E f(x, y, z) dx dy dz = \int_D dx dy \int_{\phi(x, y)}^{\psi(x, y)} f(x, y, z) dz = \int_D \int_{\phi(x, y)}^{\psi(x, y)} f(x, y, z) dz dx dy,$$

mentre per l'integrazione per strati (7.8) la scriveremo anche

$$\int_E f(x, y, z) dx dy dz = \int_a^b dz \int_{D_z} f(x, y, z) dx dy = \int_a^b \int_{D_z} f(x, y, z) dx dy dz.$$

Nel caso di $f = 1$, se E è normale, il volume di E diventa

$$\text{Vol}(E) = \int_E dx dy dz = \int_D \int_{\phi(x, y)}^{\psi(x, y)} dz dy dx = \int_D (\psi(x, y) - \phi(x, y)) dx dy$$

che è il volume della regione dello spazio compresa tra i due grafici di ϕ e ψ ; se E è stratificato, allora il suo volume diventa

$$\text{Vol}(E) = \int_a^b \text{Area}(D_z) dz.$$

Si osservi infine che se E è normale rispetto al piano xy con $\phi = 0$, allora il suo volume diventa

$$\text{Vol}(E) = \int_D \psi(x, y) dx dy,$$

che descrive quindi l'integrale doppio di una funzione come il volume del suo sottografico.

7.3 Cambiamenti di coordinate negli integrali multipli

Abbiamo già dato nella Sezione 6.3.1 la definizione di cambiamento di coordinate; vediamo qui come si comportano gli integrali multipli quando si cambiano variabili. Tale formula generalizza alle più variabili l'integrazione per sostituzione vista negli integrali di una variabile.

Nella formula di cambiamento di coordinate abbiamo bisogno della nozione di regolarità per un insieme; diremo che $E \subset \mathbb{R}^N$ è regolare se unione finita di insiemi normali di classe C^1 , cioè le funzioni che compaiono nella definizione di insieme normale sono di classe C^1 .

Teorema 7.11 (Cambiamento di variabili negli integrali multipli) *Siano E ed E' sottoinsiemi compatti regolari di \mathbb{R}^N e sia $F : E \rightarrow E'$ un cambiamento di variabili secondo la definizione 6.19. Se $f : E' \rightarrow \mathbb{R}$ è continua, allora*

$$\int_E f(x) |\det DF(x)| dx = \int_{E'} f(F^{-1}(y)) dy.$$

DIM. Per semplicità di notazione, supponiamo $N = 2$; supponiamo in più che E sia un rettangolo $[a, b] \times [c, d]$, il caso generale seguendo da un procedimento di esaustione tramite plurirettangoli. Definiamo la funzione $f : E' \rightarrow \mathbb{R}$, $f(u, v) = f(F^{-1}(u, v))$ in modo si abbia $f(x, y) = \tilde{f}(F(x, y))$. Consideriamo quindi una somma integrale individuato dalla partizione \mathcal{P} determinata dalle suddivisioni $a \leq x_0 \leq \dots \leq x_h \leq b$, $c \leq y_0 \leq \dots \leq y_k \leq d$. Il rettangolo $I_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$ viene mandato tramite F nell'insieme $F(I_{ij})$ la cui area è data da

$$\text{Area}(F(I_{ij})) = \text{Area}(D_{ij}) + o((x_i - x_{i-1})(y_j - y_{j-1}))$$

con D_{ij} quadrilatero di vertici $F(x_{i-1}, y_{j-1})$, $F(x_i, y_{j-1})$, $F(x_i, y_j)$ e $F(x_{i-1}, y_j)$. Ricordia-

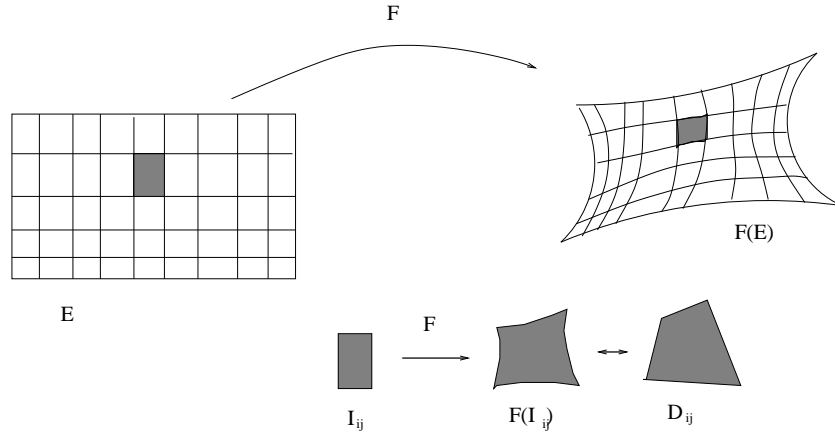


Figura 7.1: Deformazione di E tramite F .

mo la formula che fornisce l'area di un triangolo T di vertici $A = (x_A, y_A)$, $B = (x_B, y_B)$ e $C = (x_C, y_C)$:

$$\text{Area}(T) = \frac{1}{2} \left| \det \begin{pmatrix} x_A & x_B & x_C \\ y_A & y_B & y_C \\ 1 & 1 & 1 \end{pmatrix} \right|.$$

Da tale formula e tramite l'approssimazione di Taylor al primo ordine, si ricava che

$$\begin{aligned} \text{Area}(D_{ij}) &= \frac{1}{2} \left(\left| \det \begin{pmatrix} F_1(x_{i-1}, y_{j-1}) & F_1(x_i, y_{j-1}) & F_1(x_i, y_j) \\ F_2(x_{i-1}, y_{j-1}) & F_2(x_i, y_{j-1}) & F_2(x_i, y_j) \\ 1 & 1 & 1 \end{pmatrix} \right| + \right. \\ &\quad \left. + \left| \det \begin{pmatrix} F_1(x_{i-1}, y_{j-1}) & F_1(x_{i-1}, y_j) & F_1(x_i, y_j) \\ F_2(x_{i-1}, y_{j-1}) & F_2(x_{i-1}, y_j) & F_2(x_i, y_j) \\ 1 & 1 & 1 \end{pmatrix} \right| \right) \\ &= |\det DF(x_i, y_j)| + o(x_i - x_{i-1}) + o(y_j - y_{j-1}). \end{aligned}$$

Si ottiene quindi che

$$\begin{aligned} \tilde{f}(F(x_i, y_j)) \text{Area}(F(I_{ij})) &= f(x_i, y_j) |\det DF(x_i, y_j)| (x_i - x_{i-1})(y_j - y_{j-1}) + \\ &\quad + o((x_i - x_{i-1})(y_j - y_{j-1})). \end{aligned}$$

Sommando su i e j e passando al limite a zero del diametro della partizione, il Teorema segue. \square

Come abbiamo già detto nella Sezione 6.3.1, i cambiamenti di variabili che si incontrano più comunemente sono le coordinate polari nel piano e quelle cilindriche e sferiche nello spazio.

Esempio 7.1 Supponiamo di voler calcolare

$$\int_E x dx dy$$

con $E = \{(x, y) : x^2 + y^2 - 2x \leq 0\}$. Passando alle coordinate polari, considerando cioè la funzione

$$F(\varrho, \vartheta) = (\varrho \cos \vartheta, \varrho \sin \vartheta),$$

avremo che $E' = F^{-1}(E)$ sarà dato dall'insieme (basta sostituire x ed y con $\varrho \cos \vartheta$ e $\varrho \sin \vartheta$ nella definizione di E)

$$E' = \{(\varrho, \vartheta) : -\pi/2 \leq \vartheta \leq \pi/2, 0 \leq \varrho \leq 2 \cos \vartheta\}.$$

Tenendo presente che $DF(\varrho, \vartheta) = \varrho$, si ricava

$$\begin{aligned} \int_E x dx dy &= \int_{E'} \varrho \cos \vartheta \varrho d\varrho d\vartheta = \int_{-\pi/2}^{\pi/2} d\vartheta \int_0^{2 \cos \vartheta} \varrho^2 \cos \vartheta d\varrho \\ &= \frac{8}{3} \int_{-\pi/2}^{\pi/2} \cos^4 \vartheta d\vartheta = \pi. \end{aligned}$$

7.4 Applicazioni

Diamo ora in questa sezione alcuni esempi di applicazioni della teoria dell'integrazione. In particolare vediamo come calcolare il volume dei solidi di rotazione e come si definiscono e calcolano i baricentri e i momenti d'inerzia.

7.4.1 Solidi di rotazione

Un solido di rotazione si ottiene ad esempio partendo da una funzione non negativa $f : [a, b] \rightarrow \mathbb{R}$, $f \geq 0$, ruotando il sottografico rispetto all'asse x . Il solido di rotazione con generatrice f è quindi definito da

$$E = \{(x, y, z) : x \in [a, b], y^2 + z^2 \leq f^2(x)\}.$$

Il volume di tale solido è dato da

$$\text{Vol}(E) = \pi \int_a^b f^2(x) dx \quad (\text{Formula di Pappo}).$$

Tale formula si può ricavare sia notando che E è stratificato in direzione x , sia utilizzando le coordinate cilindriche. Nel primo caso la stratificazione è data dagli insiemi $E_x = \bar{B}_{f(x)}(0)$, e quindi

$$\text{Vol}(E) = \int_E dx dy dz = \int_a^b dx \int_{E_x} dy dz = \int_a^b \text{Area}(E_x) dx = \pi \int_a^b f^2(x) dx.$$

Nel secondo caso si pongono

$$\begin{cases} x = t \\ y = \varrho \sin \vartheta \\ z = \varrho \cos \vartheta \end{cases}$$

con $t \in [a, b]$, $\vartheta \in [0, 2\pi]$ e $\varrho \in [0, f(t)]$; quindi

$$\text{Vol}(E) = \int_E dx dy dz = \int_{E'} \varrho d\varrho d\vartheta dt = \int_a^b dt \int_0^{2\pi} d\vartheta \int_0^{f(t)} \varrho d\varrho = \pi \int_a^b f^2(x) dx.$$

7.4.2 Baricentri e momenti d'inerzia

Data una funzione $f : E \rightarrow \mathbb{R}$, si definisce il suo baricentro come il punto \bar{x} determinato da

$$\bar{x} = \frac{1}{\int_E f(x) dx} \int_E x f(x) dx,$$

cioè il punto di coordinate

$$\bar{x}_i = \frac{1}{\int_E f(x) dx} \int_E x_i f(x) dx.$$

L'interpretazione fisica si ha quando si pensa che E sia un corpo non omogeneo con distribuzione del peso descritta da f (che va in questo caso pensata come densità della massa di E). Allora

$$\int_E f(x) dx$$

rappresenta la massa totale del corpo E e il punto trovato \bar{x} è il baricentro di E . Nel caso particolare di corpo omogeneo con densità di massa $f(x) = c$, si ottiene

$$\bar{x} = \frac{1}{|E|} \int_E x dx,$$

che per $N = 2$ definisce il punto

$$(\bar{x}, \bar{y}) = \left(\frac{1}{\text{Area}(E)} \int_E x dx dy, \frac{1}{\text{Area}(E)} \int_E y dx dy \right),$$

mentre per $N = 3$ il punto

$$(\bar{x}, \bar{y}, \bar{z}) = \left(\frac{1}{\text{Vol}(E)} \int_E x dx dy dz, \frac{1}{\text{Vol}(E)} \int_E y dx dy dz, \frac{1}{\text{Vol}(E)} \int_E z dx dy dz \right).$$

Definiamo infine momento di inerzia di f relativo ad una retta r la quantità

$$I_r = \int_E \text{dist}(x, r)^2 f(x) dx,$$

con $\text{dist}(x, r)$ distanza del punto x dalla retta r . In particolare, se r sono gli assi cartesiani, cioè dalle rette $r(t) = te_i$, si definiscono i momenti di inerzia

$$I_i = \int_E x_i^2 f(x) dx.$$

I momenti di inerzia, se E è un corpo di densità di massa f , rappresentano l'inerzia che ha E nella rotazione attorno all'asse r . La definizione di momento d'inerzia si può anche dare per insiemi E di dimensione più bassa; in particolare, se E è una superficie (una lamina) o una curva (un'asticella), si definiscono i momenti di inerzia rispetto alla retta r come

$$I_r^\Sigma = \int_E \text{dist}(x, r)^2 d\Sigma(x), \quad I_r^s = \int_E \text{dist}(x, r)^2 ds(x),$$

dove con $d\Sigma$ e ds si intendono rispettivamente l'elemento di superficie e di linea.

Capitolo 8

Estremi e punti stazionari

In questo ultimo capitolo ci occuperemo di problemi di massimo e minimo per funzioni di più variabili; in particolare, vedremo come determinare tali punti, detti anche estremi (assoluti e locali), mediante l'individuazione dei punti stazionari della funzione, sia liberi che vincolati. Chiuderemo infine il capitolo con la classificazione dei punti stazionari liberi mediante l'uso delle forme quadratiche.

8.1 Massimi e minimi

Richiamiamo anzitutto alcune definizioni; data una funzione arbitraria $f : E \rightarrow \mathbb{R}$, con $E \subset \mathbb{R}^N$ arbitrario, diremo che $x_0 \in E$ è un *punto di minimo assoluto* (rispettivamente *punto di massimo assoluto*), o semplicemente *minimo* (risp. *massimo*) su E , se $f(x_0) \leq f(x)$ (risp. $f(x_0) \geq f(x)$) per ogni $x \in E$. Tale punto verrà inoltre detto *minimo stretto* (risp. *massimo stretto*) su E se $f(x_0) < f(x)$ (risp. $f(x_0) > f(x)$) per ogni $x \in E$. Diremo poi che $x_0 \in E$ è *punto di minimo locale* (risp. *massimo locale*) su E se esiste $r > 0$ tale che $f(x_0) \leq f(x)$ (risp. $f(x_0) \geq f(x)$) per ogni $x \in E \cap B_r(x_0)$; la definizione di minimo locale e massimo locale stretto è data in modo ovvio. Un punto di massimo o minimo (assoluto o locale) viene anche detto *estremo* (assoluto o locale). Infine, il valore $f(x_0)$ si dirà rispettivamente *valore minimo* o *valore massimo*. Si dirà poi che f ammette minimo su E (risp. ammette massimo) se esiste almeno un punto di minimo (risp. massimo) assoluto su E .

L'esistenza di minimi e massimi assoluti è garantita dalla continuità di f combinata con la compattezza di E ; abbiamo il seguente.

Teorema 8.1 (Weierstrass) *Sia $f : E \rightarrow \mathbb{R}$ una funzione continua con $E \subset \mathbb{R}^N$ compatto; f ammette allora minimo e massimo su E .*

La dimostrazione si basa sul seguente.

Lemma 8.2 *Sia $E \subset \mathbb{R}^N$ un insieme compatto e sia $(x_h)_{h \in \mathbb{N}}$ una successione contenuta in E ; esiste allora una sottosuccessione estratta $(x_{h_k})_{k \in \mathbb{N}}$ convergente ad un elemento di E .*

DIM. Essendo E è compatto, quindi chiuso e limitato, possiamo racchiudere E in un cubo $Q_0 = [-L, L]^N$ di lato $2L$; fissiamo anzitutto $h_0 = 0$ e costruiamo una sottosuccessione

estratta. Dividiamo Q_0 in 2^N cubi chiusi di lato L ottenuti dimezzando i lati di Q_0 ,

$$Q_0 = \bigcup_{i=1}^{2^N} Q_0^i.$$

Definiamo quindi i sottoinsiemi di \mathbb{N} tramite

$$I_0^i = \{h \in \mathbb{N} : x_h \in Q_0^i\};$$

chiaramente almeno uno di questi insiemi deve contenere infiniti elementi. Defiamo Q_1 un cubo tra i Q_0^i per cui $\#I_0^i = +\infty$ (numero degli elementi di I_0^i infinito) e poniamo quindi

$$h_1 = \min\{h > 0 : x_h \in Q_1\}.$$

A questo punto si procede in modo induttivo; cioè, una volta definito Q_k cubo di lato $L/2^k$, lo si suddivide in 2^N cubetti chiusi di lato $L/2^{k+1}$, ottenuti dimezzando i suoi lati

$$Q_k = \bigcup_{i=1}^{2^N} Q_k^i$$

e si definiscono

$$I_k^i = \{h \in \mathbb{N} : x_h \in Q_k^i\}.$$

Prenderemo quindi Q_{k+1} come uno tra i cubetti Q_k^i per cui $\#I_k^i = +\infty$ e si definisce

$$h_{k+1} = \min\{h > h_k : x_h \in Q_{k+1}\}.$$

La successione $(x_{h_k})_{k \in \mathbb{N}}$ sarà quindi di Cauchy, cioè convergente ad un elemento \bar{x} che apparterrà ad E in quanto E è chiuso. \square

DIMOSTRAZIONE DEL TEOREMA DI WEIERSTRASS 8.1. Dimostreremo l'esistenza di un punto di minimo; indichiamo con α il massimo dei minoranti dell'insieme $A = \{f(x) : x \in E\}$, cioè

$$\alpha = \inf_E f,$$

dove si pone $\alpha = -\infty$ se non esistono minoranti. Si considera quindi una successione $(x_h)_{h \in \mathbb{N}}$ di elementi di E tali che la successione $f(x_h)$ converga ad α . La costruzione di tale successione si effettua nel seguente modo: se $\alpha = -\infty$, allora per ogni $h \in \mathbb{N}$ deve esistere un elemento $x_h \in E$ tale che $f(x_h) < -h$. Altrimenti, se $\alpha > -\infty$, per definizione di maggiore tra i minoranti, per ogni $h \in \mathbb{N}$, $\alpha + \frac{1}{h}$ non è più un minorante per A , in altre parole esiste $x_h \in E$ tale che $f(x_h) \leq \alpha + \frac{1}{h}$. La successione $(x_h)_{h \in \mathbb{N}}$ è contenuta in E , quindi ammette una sottosuccessione convergente ad un elemento $\bar{x} \in E$; la continuità di f implica che

$$f(\bar{x}) = \lim_{k \rightarrow +\infty} f(x_{h_k}) = \alpha.$$

Questo mostra simultaneamente che $\alpha > -\infty$ e che x_0 è un punto di minimo. \square

Osservazione 8.3 L'esistenza di massimi e minimi è garantita solo sotto ipotesi di continuità sulla funzione e compattezza del dominio; non occorre che la funzione sia differenziabile. Come vedremo in seguito, il calcolo differenziale aiuta a trovare i punti di massimo e minimo tramite l'individuazione dei punti stazionari; si troveranno però quei punti estremi in cui la funzione (o la sua restrizione all'eventuale vincolo) è derivabile: in mancanza della condizione di derivabilità, occorrerà comportarsi diversamente.

8.2 Punti stazionari

Vediamo qui come il calcolo differenziale aiuta nell'individuazione dei punti di estremo assoluto. Ricordiamo la seguente definizione.

Definizione 8.4 (Punto stazionario) *Data una funzione $f : E \rightarrow \mathbb{R}$ con $E \subset \mathbb{R}^N$, diremo che x_0 è un punto stazionario per f se f è differenziabile in x_0 e $\nabla f(x_0) = 0$.*

Si noti che nella definizione di punto stazionario si richiede la differenziabilità di f e non la sola esistenza delle derivate direzionali; questo rende il concetto di punto stazionario indipendente dal sistema di riferimento. Infatti, se $F : E \rightarrow E'$ è un cambio di variabili, allora x_0 è punto stazionario per f se e solo se $F(x_0)$ è stazionario per $\tilde{f} : E' \rightarrow \mathbb{R}$, $\tilde{f} = f \circ F$. Infatti dalla differenziabilità della funzione composta (Proposizione 6.17) segue che

$$\nabla \tilde{f}(F(x_0)) = DF(x_0) \cdot \nabla f(x_0),$$

cioè $\nabla \tilde{f}(F(x_0)) = 0$ se e solo se $\nabla f(x_0) = 0$ dato che $\det DF(x_0) \neq 0$. Il seguente esempio mostra che la differenziabilità è una condizione necessaria.

Esempio 8.1 La funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definita da

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{se } (x, y) = (0, 0) \end{cases}$$

è continua ed ammette tutte le derivate direzionali nell'origine; esse valgono, se poniamo $v = (\cos \vartheta, \sin \vartheta)$

$$\frac{\partial f}{\partial v}(0, 0) = \cos^2 \vartheta \sin \vartheta.$$

In particolare, se ne deduce che $\nabla f(0, 0) = 0$, ma se ad esempio si ruotano gli assi di un angolo di $\pi/4$, cioè se si considera il cambio di coordinate

$$F(x, y) = (x \cos \vartheta + y \sin \vartheta, -x \sin \vartheta + y \cos \vartheta)$$

con $\vartheta = \pi/4$, si nota che, ponendo $\tilde{f} = f \circ F$,

$$\nabla \tilde{f}(0, 0) = \left(\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4} \right).$$

Tra punti stazionari ed estremi sussiste la seguente relazione.

Proposizione 8.5 *Sia $f : E \rightarrow \mathbb{R}$ ed $x_0 \in E$ un punto estremo locale; se x_0 è punto interno ed f è differenziabile in x_0 , allora x_0 è un punto stazionario per f .*

DIM. Se $x_0 \in E$ è un punto interno, esiste $r > 0$ tale che $B_r(x_0) \subset E$. Quindi, per ogni direzione $v \in \mathbb{R}^N$, $\|v\| = 1$, il segmento $x_0 + tv$, $t \in (-r, r)$ è contenuto in E . La differenziabilità di f in x_0 implica la derivabilità della sezione $f_v^{x_0}$ in $t = 0$; inoltre, dato che x_0 è un punto estremo locale, il punto $t = 0$ è un punto estremo locale per la sezione $f_v^{x_0}$; per quanto visto con le funzioni di una variabile, si deve quindi avere che la derivata di tale funzione deve annullarsi per $t = 0$, cioè

$$0 = \frac{df_v^{x_0}}{dt}(0) = \frac{\partial f}{\partial v}(x_0).$$

Questo deve valere per ogni direzione v , quindi $\nabla f(x_0) = 0$. □

8.3 Punti stazionari vincolati

Altro problema si presenta se il punto estremo locale si trova sul bordo di E ; in tal caso si parla di estremo vincolato. Andiamo ora a descrivere due metodi per l'individuazione di estremi vincolati: parametrizzazione del vincolo e moltiplicatori di Lagrange. Il primo metodo è solitamente il più consigliato, in quanto ha il pregio di abbassare la complessità del problema (difetto di tale metodo è che non sempre risulta chiaro come trovare la parametrizzazione). Il secondo metodo è più pesante nei calcoli, ma ha il pregio di potersi applicare anche quando si riesce a parametrizzare il vincolo.

Nei metodi che andiamo a descrivere compare una dimensione k generica del vincolo; questo è dovuto al fatto che il bordo di un insieme può essere generalmente decomposto in insiemi di varie dimensioni. Si pensi ad esempio ad un cubo, il cui bordo è composto dalle facce laterali (superfici di dimensione due), dagli spigoli (linee di dimensione uno) e dai vertici (elementi di dimensione zero, cioè punti). Nel seguito, indichiamo generalmente con la lettera S una parte del bordo ∂E di E .

8.3.1 Parametrizzazione del vincolo

Supponiamo quindi di avere un insieme $S \subset \mathbb{R}^N$ ed un punto $x_0 \in S$; diremo che S è localmente una varietà parametrizzata di dimensione k in un intorno di x_0 se esiste $r > 0$ ed una funzione $g : D \rightarrow \mathbb{R}^N$, di classe C^1 , rango di $Dg(0)$ pari a k , D aperto di \mathbb{R}^k , $0 \in D$, $g(0) = x_0$ e tale che $g(D) = S \cap B_r(x_0)$. Abbiamo quindi la seguente:

Proposizione 8.6 *Sia $f : S \rightarrow \mathbb{R}$ una funzione, $x_0 \in S$ un punto di estremo locale su S localmente, intorno ad x_0 , varietà di dimensione k con parametrizzazione da $g : D \rightarrow \mathbb{R}^N$, $D \subset \mathbb{R}^k$; se f è differenziabile in x_0 , allora x_0 è un punto stazionario per $\tilde{f} = f \circ g$, cioè $\nabla \tilde{f}(0) = 0$.*

DIM. Essendo D aperto e $0 \in D$, esiste $r > 0$ tale che $B_r(0) \subset D$; fissiamo quindi una direzione $v \in \mathbb{R}^k$, $\|v\| = 1$ e il segmento tv , $t \in (-r, r)$; la funzione di una variabile $\tilde{f}_v(t) = f(g(tv))$ ha un estremo in $t = 0$ ed è ivi derivabile, quindi

$$0 = f'_v(0) = \frac{\partial \tilde{f}}{\partial v}(0);$$

per l'arbitrarietà di v , questo conclude la dimostrazione. \square

8.3.2 Moltiplicatori di Lagrange

Questo metodo si applica quando il vincolo viene visto localmente attorno ad un estremo x_0 come luogo di zeri di una funzione vettoriale $F : \mathbb{R}^N \rightarrow \mathbb{R}^{N-k}$, cioè se esiste $r > 0$ tale che

$$(8.1) \quad S \cap B_r(x_0) = \{F = 0\} \cap B_r(x_0).$$

In tal caso si introducono i moltiplicatori di Lagrange $\lambda = (\lambda_1, \dots, \lambda_{N-k}) \in \mathbb{R}^{N-k}$ e la funzione ausiliaria

$$(8.2) \quad \Phi(x, \lambda) = f(x) - \lambda \cdot F(x),$$

con \cdot prodotto scalare in \mathbb{R}^{N-k} . Si ha il seguente risultato.

Proposizione 8.7 *Sia $f : S \rightarrow \mathbb{R}$, $x_0 \in S$ un punto di estremo locale, S localmente zero, secondo la (8.1) con $F : \mathbb{R}^N \rightarrow \mathbb{R}^{N-k}$ di classe C^1 e con $DF(x_0)$ di rango $(N-k)$; se f è differenziabile in x_0 , allora x_0 è punto stazionario, come variabile sia di x che di λ della funzione ausiliaria Φ definita in (8.2). In particolare, oltre alla condizione $F(x_0) = 0$, esiste un moltiplicatore $\lambda \in \mathbb{R}^{N-k}$ per il quale*

$$\nabla f(x_0) = \lambda \cdot DF(x_0).$$

DIM La dimostrazione segue essenzialmente dal Teorema della funzione implicita 6.18; dato che $DF(x_0)$ ha rango $N-k$, possiamo supporre (al massimo si riordinano le variabili) che $\mathbb{R}^N = \mathbb{R}^k \times \mathbb{R}^{N-k}$, $x_0 = (y_0, z_0)$ e che $D_z F(x_0)$ sia invertibile. Otteniamo quindi una funzione $H : U \rightarrow V$ da un intorno di y_0 ad un intorno di z_0 di classe $C^1(U)$ con la proprietà che possiamo parametrizzare $S \cap B_r(x_0)$ con la funzione $g : U \rightarrow \mathbb{R}^N$, $g(y) = (y, H(y))$; S è cioè dato localmente dall'intersezione di $N-k$ superfici che sono i grafici delle funzioni H_i , $i = 1, \dots, N-k$. Abbiamo quindi che, se x_0 è un estremo, y_0 è un punto stazionario per la funzione $\tilde{f}(y) = f(g(y)) = f(y, H(y))$; per la differenziabilità della funzione composta se ne deduce che

$$(8.3) \quad 0 = \nabla \tilde{f}(y_0) = \nabla f(x_0) \cdot Dg(y_0) = \nabla f(x_0) \cdot \begin{pmatrix} I_k \\ DH(y_0) \end{pmatrix}$$

con I_k matrice identità su \mathbb{R}^k . La condizione (8.3) significa che $\nabla f(x_0)$ si trova nel nucleo dell'applicazione lineare $v \mapsto v \cdot Dg(y_0)$, che si traduce nella condizione che $\nabla f(x_0)$ sia ortogonale ai grafici di ogni funzione H_i . Per quanto visto nell'Osservazione 6.14, l'ortogonale al grafico di H_i è generato dal vettore $\nabla F_i(x_0)$; dire quindi che $\nabla f(x_0)$ sia ortogonale ai grafici H_i equivale a richiedere che $\nabla f(x_0)$ sia combinazione lineare dei vettori $\nabla F_i(x_0)$, cioè che esistano i moltiplicatori di Lagrange $\lambda = (\lambda_1, \dots, \lambda_{N-k})$ tali che

$$\nabla f(x_0) = \sum_{i=1}^{N-k} \lambda_i \nabla F_i(x_0),$$

cioè, per tali λ , la funzione ausiliaria $\Phi(x, \lambda)$ ha derivate nulle in x_0 rispetto alle variabili x . Le derivate di Φ rispetto a λ danno semplicemente la condizione che $F(x_0) = 0$. \square

8.4 Calcolo di massimi e minimi assoluti su compatti

In questa sezione ricapitoliamo il metodo per l'individuazione dei massimi e minimi assoluti sugli insiemi compatti.

Sia quindi $f : E \rightarrow \mathbb{R}$ una funzione continua con $E \subset \mathbb{R}^N$ insieme compatto: si studia anzitutto la differenziabilità di f (mentre gli eventuali punti di non differenziabilità vanno considerati separatamente). I punti di non differenziabilità sono possibili candidati punti di massimo o minimo. Si passa poi a studiare la parte interna di E e si cercano i punti stazionari di f che sono interni ad E . Si passa infine allo studio del bordo di E ; qui serve un pó di regolarità di ∂E , nel senso che deve essere possibile decomporre ∂E in sottoinsiemi che siano superfici localmente parametrizzate di dimensione che vanno da 1 ad $(N-1)$; su tali insiemi si applicano i metodi o di parametrizzazione o dei moltiplicatori di Lagrange e gli eventuali punti stazionari vincolati si aggiungono alla lista dei candidati. Infine, si guardano i punti del bordo che non siano localmente superfici parametrizzate, e si inseriscono questi

punti nella lista dei candidati. Si conclude quindi confrontando i valori f in tutti i punti della lista dei candidati e il massimo sarà il valore massimo, il minimo il valore minimo.

In particolare abbiamo i casi con $N = 2$ ed $N = 3$:

1. nel piano, la parte interna di E è costituita dall'insieme dei punti che hanno due gradi di libertà, mentre i punti di bordo possono essere costituiti da archi di curva regolare (elementi uno-dimensionali) a cui si aggiungono i possibili spigoli (elementi di dimensione zero);
2. nello spazio la parte interna di E è costituita dall'insieme dei punti che hanno tre gradi di libertà, mentre i punti di bordo possono essere costituiti dalle superfici laterali (elementi bi-dimensionali), dagli spigoli, che sono archi di curva regolare (elementi uno-dimensionali) e gli eventuali vertici (elementi di dimensione zero).

Osservazione 8.8 La determinazione di massimi e minimi assoluti su insiemi non compatti è più complicata; in tal caso si possono avere punti del bordo ∂E di E che non appartengono ad E (nel caso in cui E non sia chiuso), come può anche succedere che E non sia limitato. Bisognerebbe quindi calcolare i limiti di f lungo punti che avvicinano $\partial E \setminus E$, ma anche lungo punti che si muovono nelle parti non limitate E .

8.5 Classificazione dei punti stazionari

Concludiamo il capitolo con il problema della classificazione dei punti stazionari. Tale problema è differente dal problema della determinazione dei massimi e minimi assoluti e farne una classificazione significa descrivere un punto stazionario x_0 per una funzione $f : E \rightarrow \mathbb{R}$ come uno dei seguenti possibili casi:

1. punto di minimo locale o minimo locale stretto su E ;
2. punto di massimo locale o massimo locale stretto su E ;
3. punto di sella in E , cioè un punto x_0 per il quale per ogni $r > 0$ la funzione f ammette su $B_r(x_0)$ valori sia maggiori che minori di $f(x_0)$, cioè per ogni $r > 0$, esistono $x_1, x_2 \in B_r(x_0) \cap E$ tali che

$$f(x_1) < f(x_0) < f(x_2).$$

8.5.1 Derivate seconde e matrice Hessiana

La classificazione dei punti stazionari passa, come per le funzioni di una variabile, attraverso l'uso delle derivate seconde della funzione f . Per una funzione di N variabili abbiamo visto che ci sono N derivate parziali; di ognuna di queste possiamo calcolarne la derivata, ottenendo quindi N^2 funzioni che sono le derivate seconde di f .

Definizione 8.9 (Funzioni di classe C^2) Sia $f : E \rightarrow \mathbb{R}$ e $x_0 \in E \subset \mathbb{R}^N$ punto interno; diremo che f è localmente di classe C^2 intorno ad x_0 se esiste $r > 0$ tale che siano definite e continue su $B_r(x_0)$ tutte le possibili derivate seconde di f

$$\frac{\partial^2 f}{\partial x_i \partial x_j}, \quad i, j = 1, \dots, N.$$

Diremo poi che, se E è aperto, f è di classe C^2 se f è localmente di classe C^2 intorno ad ogni punto di E . Si definisce infine matrice Hessiana della funzione f in x_0 la matrice

$$Hf(x_0) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) \right)_{i,j=1,\dots,N}.$$

Osservazione 8.10 La condizione su f di essere di classe C^2 intorno ad un punto x_0 implica in particolare che per ogni $i = 1, \dots, N$, la funzione $\frac{\partial f}{\partial x_i}$ è differenziabile in x_0 ; cioè, la funzione vettoriale $F = \nabla f$ è differenziabile in x_0 con

$$DF(x_0) = D\nabla f(x_0) = Hf(x_0).$$

Le derivate seconde in realtà non sono N^2 , ma $N(N+1)/2$ in quanto vale il seguente.

Teorema 8.11 (Teorema di Schwarz) Sia $f : E \rightarrow \mathbb{R}$ una funzione di classe localmente C^2 intorno ad un punto $x_0 \in E$; allora la matrice Hessiana di f è simmetrica, cioè

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x_0), \quad \forall i, j = 1, \dots, N.$$

Il Teorema di Schwarz 8.11 implica in particolare la simmetria della matrice Hessiana; quindi Hf definisce una forma quadratica su \mathbb{R}^N . Richiamiamo alcune proprietà delle forme quadratiche che ci servono in seguito.

8.5.2 Forme quadratiche

Una forma quadratica $A : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$, definita da una matrice simmetrica $A = (a_{ij})_{i,j=1,\dots,N}$, si dice *semi-definita positiva* (risp. *definita positiva*) se

$$(8.4) \quad Av \cdot v \geq 0 \quad (\text{risp. } Av \cdot v > 0), \quad \forall v \in \mathbb{R}^N,$$

mentre si dice *semi-definita negativa* (risp. *definita negativa*) se

$$(8.5) \quad Av \cdot v \leq 0 \quad (\text{risp. } Av \cdot v < 0), \quad \forall v \in \mathbb{R}^N$$

ed infine *indefinita* se esistono $v_1, v_2 \in \mathbb{R}^N$ tali che

$$(8.6) \quad Av_1 \cdot v_1 < 0, \quad Av_2 \cdot v_2 > 0.$$

Sulla classificazione delle forme quadratiche si ricorda il seguente risultato.

Proposizione 8.12 Sia $A : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$ una forma quadratica. Allora:

1. A è *semi-definita positiva* (risp. *definita positiva*) se e solo se gli autovalori di A sono tutti positivi (risp. strettamente positivi);
2. A è *semi-definita negativa* (risp. *definita negativa*) se e solo se gli autovalori di A sono tutti negativi (risp. strettamente negativi);
3. A è *indefinita* se e solo se esiste un autovalore strettamente positivo ed uno strettamente negativo.

DIM. La dimostrazione si basa sul fatto che una matrice simmetrica ammette una base ortonormale v_1, \dots, v_N di autovettori con rispettivi autovalori $\lambda_1, \dots, \lambda_N$; quindi possiamo scrivere un generico elemento $v \in \mathbb{R}^N$ come combinazione

$$v = \sum_{i=1}^N c_i v_i = c_1 v_1 + \dots + c_N v_N,$$

da cui

$$(8.7) \quad Av \cdot v = \sum_{i=1}^N c_i^2 \lambda_i.$$

Quindi la condizione necessaria segue prendendo nelle (8.4), (8.5) e (8.6) come v i vettori v_1, \dots, v_N , mentre la condizione sufficiente segue dalle condizioni sugli autovalori tramite la (8.7). \square

La determinazione degli autovalori è solitamente un problema facile nel caso $N = 2$ (si tratta di trovare le radici di un polinomio di secondo grado), più complesso ma ancora trattabile per $N = 3$ (radici di un polinomio di terzo grado), mentre generalmente diventa un problema difficile per $N > 3$. Sicuramente più facile, in ogni dimensione, è il calcolo di un determinante; vediamo ora come il segno degli autovalori è legato ai determinanti di A e di matrici ottenute da A . Chiaramente non può bastare il solo determinante di A , essendo questo il prodotto di tutti gli autovalori; è necessario introdurre i determinanti dei minori principali di A . Si definisce *minore principale* di A di ordine k ($k = 1, \dots, N$), la matrice quadrata $k \times k$, denotata con $A^{(k)}$, ottenuta selezionando le prime k righe e k colonne della matrice A . Abbiamo il seguente utile risultato.

Teorema 8.13 (Sylvester) *Data una matrice simmetrica $A \in \mathbb{R}^{N \times N}$, A è semi-definita positiva (risp. definita positiva) se e solo se*

$$(8.8) \quad \det A^{(k)} \geq 0 \quad (\text{risp. } \det A^{(k)} > 0), \quad \forall k = 1, \dots, N.$$

A è semi-definita negativa (risp. definita negativa) se e solo se $-A$ è semi-definita positiva (risp. definita positiva) e, quindi, se e solo se

$$(8.9) \quad (-1)^k \det A^{(k)} \geq 0 \quad (\text{risp. } (-1)^k \det A^{(k)} > 0), \quad \forall k = 1, \dots, N.$$

Possiamo approfondire questo risultato nei casi $N = 2$ ed $N = 3$.

Forme quadratiche nel piano

Supponiamo di avere $A \in \mathbb{R}^{2 \times 2}$ data da

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

con $a_{12} = a_{21}$. I minori principali sono dati da $A^{(1)} = a_{11}$ e $A^{(2)} = A$, quindi il Teorema di Sylvester 8.13 va applicato esaminando i segni di a_{11} e di $a_{11}a_{22} - a_{12}^2$ per avere che:

1. A è semi-definita positiva (risp. definita positiva) se e solo se $a_{11} \geq 0$ e $a_{11}a_{22} - a_{12}^2 \geq 0$ (risp. $a_{11} > 0$ e $a_{11}a_{22} - a_{12}^2 > 0$);

2. A è semi-definita negativa (risp. definita negativa) se e solo se $a_{11} \leq 0$ e $a_{11}a_{22} - a_{12}^2 \geq 0$ (risp. $a_{11} < 0$ e $a_{11}a_{22} - a_{12}^2 > 0$);
3. A è indefinita negli altri casi.

Un altro metodo consiste nel notare che se λ_1 e λ_2 sono i due autovalori di A , allora

$$\det A = \lambda_1 \lambda_2, \quad \operatorname{tr} A = \lambda_1 + \lambda_2;$$

quindi

1. A è semi-definita positiva (risp. definita positiva) se e solo se $\det A \geq 0$ e $\operatorname{tr} A \geq 0$ (risp. $\det A > 0$ e $\operatorname{tr} A > 0$);
2. A è semi-definita negativa (risp. definita negativa) se e solo se $\det A \geq 0$ e $\operatorname{tr} A \leq 0$ (risp. $\det A > 0$ e $\operatorname{tr} A < 0$);
3. A è indefinita se $\det A < 0$.

Forme quadratiche nello spazio

Sia $A \in \mathbb{R}^{3 \times 3}$ data da

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

con $a_{12} = a_{21}$, $a_{13} = a_{31}$ e $a_{23} = a_{32}$. I minori principali di A sono

$$A^{(1)} = a_{11}, \quad A^{(2)} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad A^{(3)} = A.$$

Dovremo controllare i segni di $d_1 = a_{11}$, $d_2 = \det A^{(2)}$ e $d_3 = \det A$ per avere che:

1. A è semi-definita positiva (risp. definita positiva) se e solo se $d_i \geq 0$ (risp. $d_i > 0$), $i = 1, 2, 3$;
2. A è semi-definita negativa (risp. definita negativa) se e solo se $d_1 \leq 0$, $d_2 \geq 0$ e $d_3 \leq 0$ (risp. $d_1 < 0$, $d_2 > 0$ e $d_3 < 0$);
3. A è indefinita negli altri casi.

La classificazione tramite determinante e traccia in questo caso può essere solo parziale, in quanto non basteranno i segni di due numeri per determinare univocamente i segni dei tre autovalori λ_1 , λ_2 e λ_3 . Si nota però che, dati:

$$\det A = \lambda_1 \lambda_2 \lambda_3, \quad \operatorname{tr} A = \lambda_1 + \lambda_2 + \lambda_3,$$

se $\det A < 0$, allora ci deve essere almeno un autovalore negativo e gli altri due di segno concorde; quindi se $\operatorname{tr} A > 0$, i restanti due autovalori sono positivi, altrimenti nulla si può concludere. Se invece $\det A > 0$, un autovalore è sicuramente positivo e gli altri due di segno concorde; quindi se $\operatorname{tr} A < 0$ i restanti due sono negativi, mentre nulla si può concludere altrimenti. Infine, nel caso $\det A = 0$, almeno un autovalore è nullo; se la traccia è nulla, o i due autovalori restanti sono nulli (nel qual caso la matrice A deve avere tutti gli a_{ij} uguali a zero), oppure sono uguali in modulo ma di segno opposto, cioè A è indefinita. Se la traccia è strettamente positiva o strettamente negativa, si può concludere solo che la matrice è non nulla con almeno un autovalore strettamente positivo o strettamente negativo, ma sul terzo autovalore nulla si può dire.

8.5.3 Classificazione dei punti stazionari

Mettiamo ora in relazione queste proprietà con le proprietà di massimo locale, minimo locale e sella di una funzione localmente di classe C^2 . A tal fine, per una funzione di classe localmente C^2 in un intorno di x_0 utilizziamo la formula di Taylor del secondo ordine data da:

$$\begin{aligned}
 f(x) &= f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2} Hf(x_0)(x - x_0) \cdot (x - x_0) + o(\|x - x_0\|^2) \\
 (8.10) \quad &= f(x_0) + \sum_{i=1}^N \frac{\partial f}{\partial x_i}(x_0)(x_i - x_{0,i}) + \\
 &\quad + \frac{1}{2} \sum_{i,j=1}^N \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)(x_i - x_{0,i})(x_j - x_{0,j}) + o(\|x - x_0\|^2).
 \end{aligned}$$

Abbiamo il seguente risultato:

Proposizione 8.14 *Sia $f : E \rightarrow \mathbb{R}$ una funzione localmente di classe C^2 in un intorno del punto x_0 , con x_0 punto stazionario di f . Si hanno le seguenti condizioni necessarie:*

- a1)** *se x_0 è un punto di minimo locale, allora $Hf(x_0)$ è semi-definita positiva;*
- a2)** *se x_0 è un punto di massimo locale, allora $Hf(x_0)$ è semi-definita negativa.*

Condizioni sufficienti sono espresse invece da:

- b1)** *se $Hf(x_0)$ è definita positiva, allora x_0 è un punto di minimo locale stretto;*
- b2)** *se $Hf(x_0)$ è definita negativa, allora x_0 è un punto di massimo locale stretto;*
- b3)** *se $Hf(x_0)$ è indefinita, allora x_0 è un punto di sella.*

DIM. La formula di Taylor (8.10) in un punto stazionario diventa

$$f(x) - f(x_0) = \frac{1}{2} Hf(x_0)(x - x_0) \cdot (x - x_0) + o(\|x - x_0\|^2).$$

Se si fissa una direzione $v \in \mathbb{R}^N$, $\|v\| = 1$, se ne deduce che

$$f(x_0 + tv) - f(x_0) = \frac{t^2}{2} Hf(x_0)v \cdot v + o(t^2);$$

se x_0 punto di minimo, ne segue che

$$0 \leq \frac{f(x_0 + tv) - f(x_0)}{2t^2} = Hf(x_0)v \cdot v + o(t^2)$$

e quindi $Hf(x_0)$ sia semi-definita positiva. In modo analogo si ragiona nel caso di massimo locale.

Per le condizioni **b1)**–**b3)**, se $Hf(x_0)$ è definita positiva, la funzione $g : \partial B_1(0) \rightarrow \mathbb{R}$, $g(v) = Hf(x_0)v \cdot v$ è continua e definita su un insieme compatto, quindi ammette minimo m che deve essere strettamente positivo. Quindi

$$0 < m \leq Hf(x_0) \left(\frac{x - x_0}{\|x - x_0\|} \right) \cdot \left(\frac{x - x_0}{\|x - x_0\|} \right) = \frac{f(x) - f(x_0)}{\|x - x_0\|^2} + \frac{o(\|x - x_0\|^2)}{\|x - x_0\|^2}.$$

Siccome $\frac{o(\|x-x_0\|^2)}{\|x-x_0\|^2} \rightarrow 0$ per $\|x-x_0\| \rightarrow 0$, se ne deduce l'esistenza di $r > 0$ tale che $\|x-x_0\| < r$

$$\frac{f(x) - f(x_0)}{\|x - x_0\|^2} > \frac{m}{2},$$

cioè $f(x) > f(x_0)$ in $B_r(x_0)$. In modo analogo si ragiona nel caso di matrice Hessiana definita negativa. Se la matrice Hessiana è indefinita, basta considerare i segmenti $x_0 + tv_1$ e $x_0 + tv_2$ con v_1 e v_2 dati dalla (8.6) per dedurre che sul primo esistono punti con valori di f minori di $f(x_0)$, mentre sul secondo esistono punti con valori di f maggiori di $f(x_0)$. \square

Osservazione 8.15 Si potrebbe anche considerare il problema della classificazione dei punti stazionari vincolati; si tratta in tal caso di classificare la matrice Hessiana della funzione $f(g(x))$, con $g : D \rightarrow \mathbb{R}^N$ parametrizzazione del vincolo.

Capitolo 9

Domande

9.1 Sviluppi di Taylor: domande

1. Dare la definizione del simbolo o (o piccolo) ed enunciare alcune delle sue principali proprietà, fornendo gli esempi che si ritengono opportuni.
2. Dare la definizione di infinitesimo; tramite esempi, fornire inoltre il concetto di ordine di infinitesimo tramite i monomi x^a , $x \in \mathbb{R}$.
3. Dare la definizione e le principali proprietà del polinomio di Mac Laurin e del polinomio di Taylor; dare alcuni esempi.
4. Dimostrare che, data una funzione f , vale l'identità $o(f) - o(f) = o(f)$ e che in generale $o(f) - o(f) \neq 0$ (fornire almeno un controesempio).
5. Definizione di resto di Peano e resto di Lagrange; dare alcuni esempi.
6. Dare la definizione di linearizzazione di una funzione e fornire alcuni esempi significativi.
7. Discutere il concetto di approssimazione e fornire metodi per la stima dell'errore nelle approssimazioni.
8. Mostrare come si può utilizzare la formula di Taylor del secondo ordine per la classificazione dei punti stazionari di una funzione.

9.2 Numeri complessi: domande

1. Dare la definizione algebrica dei numeri complessi; descrivere inoltre l'operazione di coniugio e modulo di un numero complesso.
2. Equivalenza tra forma algebrica e polare di un numero complesso; definizioni e regole per passare da una all'altra.
3. Definizione di piano di Gauss e rappresentazione cartesiana di un numero complesso; interpretazione geometrica di somma e prodotto tra numeri complessi.

4. Definizione e principali proprietà di coniugio e modulo di un numero complesso.
5. Principio di uguaglianza tra numeri complessi in forma algebrica ed in forma polare.
6. Formula di De Moivre e definizione dell'esponenziale immaginario.
7. Definizione di argomento e modulo di un numero complesso; esempi con numeri reali positivi, negativi, i e $-i$.
8. Definizione di radice n -esime di un numero complesso e loro distribuzione nel piano di Gauss.
9. Dare la definizione di radici n -esime dell'unità, cioè del numero 1, e descrivere le loro principali proprietà.
10. Dare la definizione di radici n -esime di -1 , e descrivere le loro principali proprietà.
11. Enunciato del Teorema Fondamentale dell'algebra nelle sue due forme equivalenti; in che senso le due forme sono equivalenti?
12. Calcolo delle radici di un polinomio complesso in campo complesso.
13. Polinomi complessi a coefficienti reali; proprietà delle radici.

9.3 Equazioni differenziali: domande

1. Dare la definizione di equazione differenziale di ordine n in forma implicita ed esplicita e dire cosa si intende per soluzione; fornire alcuni esempi significativi.
2. Equazioni differenziali del primo ordine; definizione ed esempi di equazioni a variabili separabili.
3. Definizione di Problema di Cauchy ed enunciato di almeno un risultato di esistenza e unicità.
4. Equazioni differenziali lineari del primo ordine; struttura dell'insieme delle soluzioni e formula risolutiva.
5. Metodo della variazione delle costanti; definizione ed alcune applicazioni.
6. Equazione di Bernoulli; definizione e metodo di risoluzione, con discussione riguardo all'esistenza delle soluzioni.
7. Equazioni differenziali del secondo ordine; struttura dell'insieme delle soluzioni.
8. Definizione di polinomio caratteristico ed applicazione alla risoluzione delle equazioni differenziali lineari.
9. Descrizione dei possibili elementi della base dell'insieme delle soluzioni delle equazioni differenziali lineari omogenee a coefficienti costanti.
10. Descrizione di almeno un metodo per la costruzione della soluzione particolare di una equazione differenziale lineare a coefficienti costanti.

11. Descrizione del metodo di somiglianza per le equazioni differenziali.
12. Equazione dell'oscillatore armonico con termine forzante; fenomeno dei battimenti e della risonanza.
13. Descrivere il metodo per la risoluzione di equazioni differenziali lineari a coefficienti costanti di ordine maggiore al secondo; fornire alcuni esempi.

9.4 \mathbb{R}^N , topologia, limiti e funzioni continue: domande

1. Dare la definizione di distanza in \mathbb{R}^N e descriverne le sue principali proprietà.
2. Dare la definizione di intorno sferico, descrivendone le principali proprietà.
3. Dare la definizione di insieme aperto ed insieme chiuso; dare la loro caratterizzazione per successioni.
4. Dare la definizione di parte interna di un insieme e dire quando un insieme è aperto.
5. Dare la definizione di frontiera di un insieme e definire la chiusura di un insieme.
6. Dare la definizione di successione in \mathbb{R}^N ; dire quando una successione è convergente e descrivere le proprietà dei limiti per le successioni.
7. Dare le definizioni di limite per una funzione e di funzione continua.
8. Proprietà dei limiti per funzioni; somma, prodotto e composizione.
9. Dare la definizione di funzione continua e fornire alcuni esempi.
10. Funzioni continue e insiemi aperti; dire come costruire insiemi aperti e chiusi usando le funzioni continue.

9.5 Curve ed integrali di linea: domande

1. Dare la definizione di curva e di lunghezza di una curva.
2. Dare la definizione di curva regolare semplice e fornire la formula per il calcolo della lunghezza di una curva regolare.
3. Dare la definizione di curva regolare chiusa; portare come esempi la circonferenza e l'ellisse.
4. Dare la definizione di sostegno di una curva, di curve regolari equivalenti e dimostrare che la lunghezza di curve equivalenti è la stessa.
5. Dare la definizione di curva regolare φ e descrivere la relazione tra φ' e il sostegno della curva.
6. Dare la definizione di lunghezza di una curva regolare e di ascissa curvilinea.

7. Dare la definizione di curva regolare cartesiana e fornire la formula per il calcolo della sua lunghezza.
8. Dare la definizione di curva regolare in coordinate polari e fornire la formula per il calcolo della sua lunghezza.
9. Dare la definizione di integrale curvilineo di prima specie (per funzioni scalari) e dimostrare che tale integrale è indipendente dalla parametrizzazione della curva.
10. Dare la definizione di integrale curvilineo di seconda specie (per funzioni vettoriali) e descrivere la relazione con gli integrali curvilinei di prima specie (per funzioni scalari).
11. Dare l'interpretazione fisica dell'integrale curvilineo di seconda specie e la definizione di campo conservativo.

9.6 Calcolo Differenziale in più variabili: domande

1. Dare la definizione di insieme connesso, convesso e stellato.
2. Descrivere come e quando è possibile caratterizzare la continuità di una funzione mediante curve.
3. Dare la definizione di derivata direzionale e di derivata parziale.
4. Dare la definizione di differenziale e descrivere le sue relazioni con le derivate direzionali.
5. Dopo aver dato la definizione di differenziale di una funzione f in un punto x_0 , discutere l'unicità del differenziale e la relazione con la continuità di f in x_0 .
6. Discutere la relazione tra derivate parziali e differenziale; Teorema del differenziale totale.
7. Dare la definizione di insieme di livello e descrivere la relazione tra gradiente di una funzione ed insiemi di livello per funzioni differenziabili.
8. Dare la definizione di differenziabilità per funzioni vettoriali e dare la definizione di matrice Jacobiana.
9. Dare la definizione di piano tangente al grafico di una funzione differenziabile.
10. Differenziale della funzione composta; dare alcuni esempi significativi.
11. Dire che cosa si intende per cambio di variabile e dare almeno un esempio di cambio di variabile.

9.7 Integrali multipli: domande

1. Dare la definizione di integrale doppio e descrivere le sue principali proprietà.
2. Dare la definizione di insieme normale nel piano e descrivere i metodi di integrazione doppia su insiemi normali.
3. Dare la definizione di misura di un insieme e descrivere il metodo di calcolo dell'area di un insieme normale nel piano.
4. Dare la definizione di insieme normale nello spazio e descrivere il metodo di integrazione per integrali tripli denominato "integrazione per fili".
5. Descrivere il metodo di integrazione per integrali tripli denominata "integrazione per strati".
6. Scrivere la formula di cambio di variabili negli integrali multipli; descrivere in particolare come passare all'integrazione in coordinate polari nel piano.
7. Scrivere la formula di cambio di variabili negli integrali multipli; descrivere in particolare come passare all'integrazione in coordinate cilindriche nello spazio.
8. Scrivere la formula di cambio di variabili negli integrali multipli; descrivere in particolare come passare all'integrazione in coordinate polari nello spazio.
9. Dare la definizione di insiemi invadenti e dare la definizione di funzione assolutamente integrabile in senso generalizzato; fornire un esempio.
10. Dire cosa si intende per solido di rotazione e scrivere la formula per la determinazione del suo volume.

9.8 Massimi e minimi: domande

1. Dare la definizione di punto di massimo e minimo assoluto e locale; dare condizioni sufficienti per la loro esistenza.
2. Dare la definizione di punto stazionario libero e descrivere la loro relazione con i punti estremi.
3. Dare la definizione di punto stazionario e dimostrare l'indipendenza di tale nozione dal sistema di riferimento scelto.
4. Dare la definizione di punto stazionario vincolato e descrivere la loro relazione con i punti estremi su insiemi compatti.
5. Ricerca di massimi e minimi vincolati; descrivere il metodo mediante parametrizzazione.
6. Ricerca di massimi e minimi vincolati; descrivere il metodo mediante i moltiplicatori di Lagrange.
7. Massimi e minimi su insiemi compatti; fornire un metodo per la loro determinazione.

8. Dire cosa si intende con classificazione dei punti stazionari; fornire alcuni esempi.
9. Dare la definizione di matrice Hessiana e descrivere le sue principali proprietà.
10. Dire come si classifica una forma quadratica mediante lo studio dei suoi autovalori.
11. Enunciare il Teorema di Sylvester e descrivere il suo utilizzo nella classificazione delle forme quadratiche.
12. Descrivere i vari modi particolari con cui si possono classificare le forme quadratiche nel piano.
13. Descrivere i vari modi particolari con cui si possono classificare le forme quadratiche nello spazio.
14. Descrivere la classificazione dei punti stazionari mediante forme quadratiche.
15. Dire come si classificano i punti stazionari nel piano.
16. Dire come si classificano i punti stazionari nello spazio.