

Capitolo 4

Funzioni da \mathbb{R}^n in \mathbb{R}^m

Studiamo ora le funzioni di più di una variabile; ossia funzioni

$$\mathbf{F}(\mathbf{x})$$

a valori in \mathbb{R}^m , della variabile $\mathbf{x} \in \mathbb{R}^n$, in generale con $n \neq m$. Ovviamente, le prime proprietà da studiare saranno le proprietà relative ai limiti ed alla continuità. Come vedremo, non si incontrano situazioni sostanzialmente diverse da quelle delle funzioni da \mathbb{R} in \mathbb{R} (però le proprietà relative agli estremi saranno tipiche delle funzioni a valori reali). Differenze profonde si troveranno invece nello studio della differenziabilità.

4.1 Limiti e continuità

Ricordiamo che col simbolo $|\mathbf{r}|$ intendiamo la norma di \mathbf{r} , senza esplicitamente indicare la dimensione del vettore. Per esempio,

$$\begin{aligned} \text{se } \mathbf{r} = x \in \mathbb{R} \text{ allora } |\mathbf{r}| &= \sqrt{x^2} = |x|; \\ \text{se } \mathbf{r} = (x, y) \text{ allora } |\mathbf{r}| &= \sqrt{x^2 + y^2}; \\ \text{se } \mathbf{r} = (x, y, z) \text{ allora } |\mathbf{r}| &= \sqrt{x^2 + y^2 + z^2}. \end{aligned}$$

Se vorremo specificare che la norma è quella di \mathbb{R}^n allora scriveremo

$$|\mathbf{r}|_{\mathbb{R}^n}.$$

Sia $\mathbf{F}(\mathbf{r})$ una funzione definita su un insieme Ω di \mathbb{R}^n , a valori in \mathbb{R}^m e sia \mathbf{r}_0 un punto di accumulazione di Ω .

Diremo che

$$\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \mathbf{F}(\mathbf{r}) = \mathbf{L}$$

Se per **ogni** $\epsilon > 0$ **esiste** un intorno $I_\epsilon(\mathbf{r}_0)$ del punto \mathbf{r}_0 tale che

per **ogni** $\mathbf{r} \in \Omega \cap I_\epsilon(\mathbf{r}_0)$, $\mathbf{r} \neq \mathbf{r}_0$ si ha

$$|\mathbf{F}(\mathbf{r}) - \mathbf{L}|_{\mathbb{R}^m} < \epsilon.$$

In simboli:

$$\forall \epsilon > 0 \exists I_\epsilon(\mathbf{r}_0) \forall \mathbf{r} \in \Omega \cap (I_\epsilon(\mathbf{r}_0) - \{\mathbf{r}_0\}) \implies |\mathbf{F}(\mathbf{r}) - \mathbf{L}|_{\mathbb{R}^m} < \epsilon.$$

Se $\mathbf{L} = \mathbf{0}$ allora si dice che la funzione $\mathbf{F}(\mathbf{r})$ è *infinitesima* per $\mathbf{r} \rightarrow \mathbf{r}_0$.
 Se accade che $\mathbf{F}(\mathbf{r})$ è definita in \mathbf{r}_0 e se inoltre $\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \mathbf{F}(\mathbf{r}) = \mathbf{F}(\mathbf{r}_0)$, allora si dice che $\mathbf{F}(\mathbf{r})$ è *continua* in \mathbf{r}_0 .

Si noti che la definizione di limite e quella di continuità sono formalmente analoghe a quelle note per funzioni di una variabile. la differenza importante da tener presente è che ora il simbolo $I_\epsilon(\mathbf{r}_0)$ indica una palla aperta di centro \mathbf{r}_0 . E' un po' diversa la definizione dei limiti infiniti o per \mathbf{r} tendente ad infinito, perché non si definisce il limite "direzionale". Piuttosto, si definisce

Sia Ω un insieme illimitato. Si dice che

$$\lim_{|\mathbf{r}|_{\mathbb{R}^n} \rightarrow +\infty} \mathbf{F}(\mathbf{r}) = \mathbf{L}$$

se per **ogni** $\epsilon > 0$ **esiste** δ tale che se $|\mathbf{r}|_{\mathbb{R}^n} > \delta$ e $\mathbf{r} \in \Omega$ si ha

$$|\mathbf{F}(\mathbf{r}) - \mathbf{L}|_{\mathbb{R}^m} < \epsilon.$$

Sia \mathbf{r}_0 punto di accumulazione di Ω . Si dice che

$$\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} |\mathbf{F}(\mathbf{r})|_{\mathbb{R}^m} = +\infty$$

se per **ogni** $\epsilon > 0$ **esiste** $\delta > 0$ tale che se $\mathbf{r} \in \Omega$, $\mathbf{r} \neq \mathbf{r}_0$ e $|\mathbf{r} - \mathbf{r}_0|_{\mathbb{R}^n} < \delta$ si ha

$$|\mathbf{F}(\mathbf{r})|_{\mathbb{R}^m} > \epsilon.$$

In queste definizioni, per completezza, abbiamo usato gli indici \mathbb{R}^n e \mathbb{R}^m per indicare esplicitamente gli spazi nei quali le norme vanno lette; da ora in poi però ometteremo tali indici.

Osservazione 121 Una funzione $\mathbf{F}(\mathbf{r})$ da \mathbb{R}^n in \mathbb{R}^m è infinitesima (per $\mathbf{r} \rightarrow \mathbf{r}_0$ oppure per $|\mathbf{r}| \rightarrow +\infty$) se e solo se $|\mathbf{F}(\mathbf{r})|$ è infinitesima come funzione da \mathbb{R}^n in \mathbb{R} . ■

4.1. LIMITI E CONTINUITÀ

Si rileggano i teoremi sui limiti delle funzioni di una variabile, identificando quelle dimostrazioni che, grazie alle proprietà della norma, si ripetono senza alcuna modifica nel caso delle funzioni di più variabili. Naturalmente non rientrano tra queste quelle dimostrazioni che dipendono dall'ordine tra i numeri reali, come il teorema delle funzioni monotone e le sue conseguenze. Sottolineiamo che in particolare valgono:

1. Il teorema di unicità del limite;
2. il teorema di limitatezza locale;
3. i teoremi che correlano i limiti e la continuità con le operazioni. Naturalmente, tra funzioni a valori vettoriali non si fanno divisioni; e il teorema del prodotto va applicato al prodotto scalare o anche, quando $m = 3$, al prodotto vettoriale: i prodotti scalari e vettoriali di funzioni continue sono funzioni continue;
4. i teoremi relativi ai limiti di funzioni composte;
5. per funzioni a valori in \mathbb{R} , il teorema di confronto per i limiti e il teorema di permanenza del segno;
6. il teorema di confronto per gli infiniti e gli infinitesimi.

In particolare, vale

Lemma 122 *Sia*

$$\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \frac{\mathbf{F}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_0|} = 0.$$

Sia

$$\lim_{t \rightarrow t_0} \mathbf{r}(t) = \mathbf{r}_0$$

ed esista un intorno di t_0 su cui la funzione $\mathbf{r}(t)$ non prende valore \mathbf{r}_0 . In tal caso vale

$$\lim_{t \rightarrow t_0} \frac{\mathbf{F}(\mathbf{r}(t))}{|\mathbf{r}(t) - \mathbf{r}_0|} = 0.$$

Notiamo un caso particolare del teorema sulla continuità delle funzioni composte, già usato nel corso di Analisi Matematica 1, trattando le equazioni differenziali:

Teorema 123 *Sia $f(\mathbf{x})$ un funzione da \mathbb{R}^n in \mathbb{R}^m , continua su un insieme Ω .*

Sia $\mathbf{x}(t)$ una funzione continua della variabile reale $t \in [a, b]$, a valori in Ω .

La funzione composta $f(\mathbf{x}(t))$ è continua su $[a, b]$.

Come si è detto, una funzione che ha limite zero (per $\mathbf{r} \rightarrow \mathbf{r}_0$) si chiama ancora un *infinitesimo* (per $\mathbf{r} \rightarrow \mathbf{r}_0$); e si può istituire un confronto tra gli infinitesimi, esattamente come nel caso di funzioni di una variabile. Per esempio, il simbolo

$$\mathbf{F} = o(|\mathbf{r} - \mathbf{r}_0|)$$

significa

$$\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \frac{|\mathbf{F}(\mathbf{r})|}{|\mathbf{r} - \mathbf{r}_0|} = 0.$$

Usando il linguaggio degli infinitesimi, possiamo enunciare:

Teorema 124 *Si ha*

$$\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \mathbf{F}(\mathbf{r}) = \mathbf{L}$$

se e solo se la funzione $\mathbf{F}(\mathbf{r}) - \mathbf{L}$ è infinitesima per $\mathbf{r} \rightarrow \mathbf{r}_0$.

Infine, notiamo che le disuguaglianze (3.3) permettono di provare:

Teorema 125 *Siano $F_i(\mathbf{r})$ le componenti della funzione $\mathbf{F}(\mathbf{r})$ da \mathbb{R}^n in \mathbb{R}^m . Si ha $\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \mathbf{F}(\mathbf{r}) = \mathbf{L}$ se e solo se per ogni indice i vale $\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} F_i(\mathbf{r}) = L_i$, ove L_i sono le componenti del vettore \mathbf{L} .*

Si enunci il risultato analogo per i limiti per $|\mathbf{r}| \rightarrow +\infty$. Nonostante che abbiamo insistito sulla completa corrispondenza che intercorre tra definizioni e teoremi in una e più variabili, bisogna sottolineare una **importante dissimmetria**, che illustriamo con riferimento a funzioni da \mathbb{R}^2 in \mathbb{R}^2 . Sia quindi

$$\mathbf{F}(x, y) = f(x, y)\mathbf{i} + g(x, y)\mathbf{j}.$$

Da

$$\begin{aligned} |f(x, y)| &\leq |\mathbf{F}(x, y)| \leq |f(x, y)| + |g(x, y)| \\ |g(x, y)| &\leq |\mathbf{F}(x, y)| \leq |f(x, y)| + |g(x, y)| \end{aligned}$$

si vede che $\mathbf{F}(x, y)$ è infinitesima se e solo se ambedue le sue componenti lo sono; è continua se e solo se ambedue le sue componenti lo sono; ossia, come asserito dal Teorema 125, limiti e continuità possono studiarsi esaminando separatamente le due componenti dei valori assunti dalla funzione, ossia le due componenti $f(x)$ ed $g(x)$ di $\mathbf{F}(x)$. Ciò non accade nello spazio di partenza; ossia le due componenti x ed y non possono trattarsi separatamente. Per vedere questo, basta considerare una sola delle componenti di $\mathbf{F}(x, y)$, per esempio la funzione a valori reali $f(x, y)$. La funzione

$$\phi(x) = \lim_{y \rightarrow y_0} f(x, y) \tag{4.1}$$

può essere definita per ogni x , e può esistere $\lim_{x \rightarrow x_0} \phi(x)$; però tale limite è in generale diverso da $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y)$, come mostra l'esempio seguente:

4.1. LIMITI E CONTINUITÀ

Esempio 126 Sia

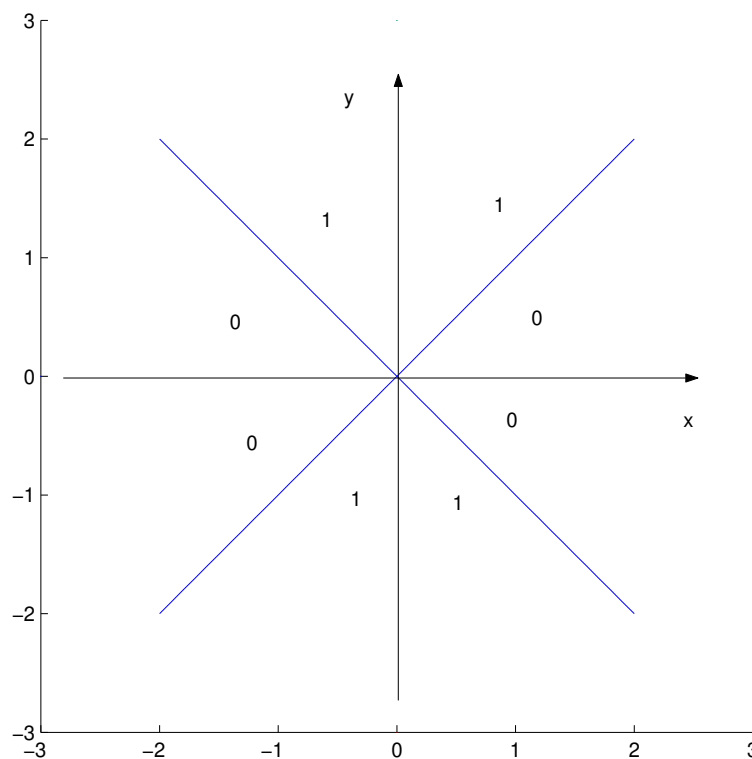
$$f(x, y) = \begin{cases} 1 & \text{se } |x| < |y| \\ 0 & \text{se } |x| \geq |y|. \end{cases}$$

La funzione $f(x, y)$ è priva di limite per (x, y) tendente a zero, mentre

$$\lim_{x \rightarrow 0} \left[\lim_{y \rightarrow 0} f(x, y) \right] = 0, \quad \lim_{y \rightarrow 0} \left[\lim_{x \rightarrow 0} f(x, y) \right] = 1. \quad \blacksquare$$

Si veda la figura 4.1.

Figura 4.1: La funzione dell'esempio 126



Può anche accadere che esista

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y)$$

ma che non esista il limite che definisce $\phi(x)$ in (4.1), come mostra l'esempio seguente:

Esempio 127 Sia

$$f(x, y) = x [\operatorname{sgn}(y)] .$$

Essendo $|f(x, y)| \leq |x|$, il *teorema di confronto dei limiti* mostra che

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0 ;$$

ma,

$$\lim_{y \rightarrow 0} f(x, y)$$

esiste soltanto se $x = 0$. ■

Infine, ricordiamo che successione è il termine che si usa per designare una funzione definita sui numeri naturali. Una successione a valori in \mathbb{R}^m si indicherà col simbolo (\mathbf{r}_n) o, più frequentemente $\{\mathbf{r}_n\}$ ¹. Come nel caso particolare delle successioni a valori reali, le definizioni e i teoremi sui limiti delle successioni sono casi particolari di quelli relativi alle funzioni.

4.1.1 Funzioni continue su insiemi

E' possibile provare l'analogo del Teorema di *Bolzano-Weierstrass*: ogni successione limitata a valori in \mathbb{R}^n ammette sottosuccessioni convergenti. Un insieme Ω si dice compatto, quando ogni successione in Ω ammette sottosuccessioni convergenti a punti di Ω . Si può ancora provare:

Teorema 128 *Un insieme Ω è compatto se e solo se è limitato e chiuso.*

Grazie a ciò, si prova l'analogo del *Teorema di Weierstrass* con la medesima dimostrazione vista per funzioni di una variabile. Diremo che:

1. un punto \mathbf{r}_0 è *punto di minimo*, o di massimo, per una funzione $f(\mathbf{r})$ a valori reali, se vale, rispettivamente,

$$f(\mathbf{r}) \geq f(\mathbf{r}_0) \quad \text{oppure} \quad f(\mathbf{r}) \leq f(\mathbf{r}_0)$$

per ogni $\mathbf{r} \in \Omega$;

2. diremo che la funzione $\mathbf{F}(\mathbf{r})$, da \mathbb{R}^n in \mathbb{R}^m è *uniformemente* continua se **per ogni** $\epsilon > 0$ esiste $\delta > 0$ tale che

$$\mathbf{r} \in \Omega, \quad \mathbf{r}_0 \in \Omega, \quad |\mathbf{r} - \mathbf{r}_0| < \delta \quad \text{implica} \quad |\mathbf{F}(\mathbf{r}) - \mathbf{F}(\mathbf{r}_0)| < \epsilon .$$

¹se si usa il simbolo con la parentesi graffa, si deve fare attenzione a capire quando si indica la successione oppure quando se ne vuole indicare l'insieme immagine.

4.2. LE PROPRIETÀ DI DIFFERENZIABILITÀ

Vale:

Teorema 129 Sia $f(\mathbf{r})$ una funzione a valori reali definita e continua su un compatto $K \subseteq \mathbb{R}^n$. La funzione $f(\mathbf{r})$ ammette in K sia punti di minimo che punti di massimo.

Il teorema di esistenza degli zeri, equivalentemente quello dei valori intermedi, richiede, per funzioni di una variabile, di lavorare su intervalli. Gli insiemi aperti e connessi sono gli analoghi degli intervalli aperti, nel senso che per essi vale il teorema seguente:

Teorema 130 Sia Ω un insieme aperto e connesso e sia $f(\mathbf{r})$ una funzione a valori reali e continua su Ω . Se \mathbf{r}_0 e \mathbf{r}_1 sono due punti di Ω , la funzione assume ogni valore c compreso tra $f(\mathbf{r}_0)$ e $f(\mathbf{r}_1)$.

Dim. Sia

$$f(\mathbf{r}_0) = a, \quad f(\mathbf{r}_1) = b.$$

Se è possibile congiungere \mathbf{r}_0 e \mathbf{r}_1 con un segmento contenuto in Ω , si considera la restrizione della funzione a tale segmento. Si trova una funzione di una sola variabile, continua e definita su un intervallo, che assume i valori a e b . Dunque assume anche il valore intermedio c . Nel caso che i due punti si congiungano con una poligonale, ossia, con un numero finito di segmenti, si procede in modo analogo sui singoli segmenti. ■ Esistono versioni del teorema precedente anche per funzioni a valori vettoriali. Si veda per esempio il paragrafo 7.5.2.

4.2 Le proprietà di differenziabilità

Per lo studio di queste proprietà, conviene separare lo studio delle funzioni a valori in \mathbb{R} da quello delle funzioni a valori in \mathbb{R}^m con $m > 1$.

4.2.1 Il differenziale delle funzioni a valori reali

Quando la funzione $f(x, y)$ dipende da due variabili, la derivata parziale rispetto ad x è la funzione sia di x che di y definita da

$$\lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h}.$$

Per indicare questa funzione si usa uno dei simboli

$$\frac{\partial}{\partial x} f(x, y), \quad f_x(x, y).$$

In modo analogo si definisce la derivata parziale rispetto ad y e quella rispetto alle ulteriori variabili nel caso che la funzione dipenda da più di due variabili. Si noti esplicitamente che trattando delle equazioni differenziali si è dovuto richiedere però non la sola esistenza delle derivate parziali, ma la loro continuità. In effetti, la sola esistenza delle derivate parziali è un concetto molto debole. Infatti:

Esempio 131 L'esistenza delle derivate parziali in un punto (x_0, y_0) , non implica la continuità della funzione in tale punto, come mostra l'esempio seguente:

$$f(x, y) = \begin{cases} 0 & \text{se } x \cdot y = 0 \\ 1 & \text{altrimenti,} \end{cases} \quad (x_0, y_0) = (0, 0).$$

Per contrasto, si ricordi che l'esistenza della derivata prima in un punto di una funzione di una sola variabile, implica la continuità in tale punto. Ricordiamo che la dimostrazione di ciò segue dalla prima formula degli incrementi finiti e che praticamente tutte le proprietà delle funzioni derivabili di una variabile seguono dalla prima oppure dalla seconda formula degli incrementi finiti. Dunque, se si vuol sperare di ripetere, per le funzioni di più variabili, una teoria simile a quella delle funzioni di una variabile, dovremo dare condizioni per l'esistenza di una formula analoga alla prima formula degli incrementi finiti: ossia, nel caso di funzioni di due variabili, vorremo condizioni perchè valga la formula seguente

$$f(x, y) - f(x_0, y_0) = a(x - x_0) + b(y - y_0) + o(|(x - x_0, y - y_0)|). \quad (4.2)$$

Il risultato che vogliamo provare è:

Teorema 132 *Sia $f(x, y)$ una funzione di due variabili. Supponiamo che essa ammetta ambedue le derivate parziali in ogni punto (x, y) di un intorno di (x_0, y_0) e che queste siano continue in tale intorno. Esistono numeri a, b per i quali vale la formula (4.2), ed è:*

$$a = f_x(x_0, y_0), \quad b = f_y(x_0, y_0).$$

La dimostrazione è in appendice. Passando al limite per $\text{dist}((x, y), (x_0, y_0))$ tendente a zero in (4.2) si vede che $\lim f(x, y) = f(x_0, y_0)$. Si ha quindi in particolare:

Corollario 133 *Se una funzione ha derivate parziali continue un intorno di (x_0, y_0) , essa è continua in (x_0, y_0) .*

Argomenti del tutto analoghi valgono per funzioni di tre o più variabili:

4.2. LE PROPRIETÀ DI DIFFERENZIABILITÀ

Teorema 134 Una funzione di n variabili le cui n derivate prime esistono e sono continue su un aperto Ω è continua su Ω e per essa vale

$$f(\xi^1, \xi^2, \dots, \xi^n) = f(x^1, x^2, \dots, x^n) + \sum_{i=1}^n (\xi^i - x^i) f_{x^i}(x^1, x^2, \dots, x^n) + o(|\mathbf{h}|), \quad \mathbf{h} = ((\xi^1 - x^1), (\xi^2 - x^2), \dots, (\xi^n - x^n)).$$

Si chiama **differenziale** della funzione f in $\mathbf{r} = [x^1 \ x^2 \ \dots \ x^n]'$ la trasformazione che al vettore (y_1, \dots, y_n) associa il numero

$$\sum_{i=1}^n y_i f_{x^i}(x^1, x^2, \dots, x^n).$$

Una funzione dotata di differenziale si chiama **differenziabile**. Il vettore

$$\begin{bmatrix} f_{x^1}(\mathbf{r}) \\ f_{x^2}(\mathbf{r}) \\ \vdots \\ f_{x^n}(\mathbf{r}) \end{bmatrix} \quad \mathbf{r} = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}$$

si chiama il **gradiente** della funzione $f(\mathbf{r})$ e si indica col simbolo $\nabla f(\mathbf{r})$. Il simbolo ∇ si legge “grad”, oppure “del” od anche “nabla”. Usando questo simbolo, la (4.3) si scrive

$$f(\mathbf{r}) - f(\mathbf{r}_0) = \nabla f(\mathbf{r}_0) \cdot (\mathbf{r} - \mathbf{r}_0) + o(|\mathbf{r} - \mathbf{r}_0|).$$

Si chiama *piano tangente* al grafico di $f(\mathbf{r})$ nel punto $\mathbf{r}_0 = (x_0^1, \dots, x_0^n)$ il grafico della funzione

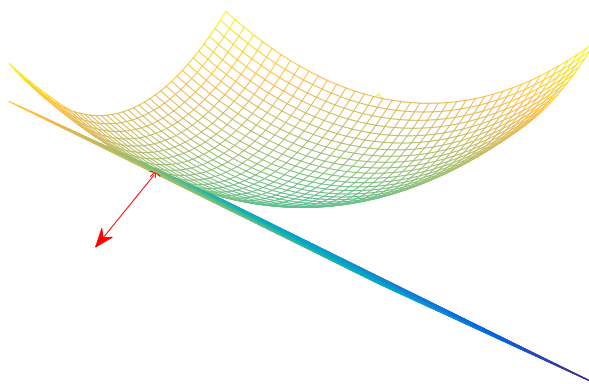
$$\mathbf{r} \rightarrow \mathbf{r}_0 + \nabla f(\mathbf{r}_0) \cdot (\mathbf{r} - \mathbf{r}_0).$$

Dunque, l'equazione del piano tangente è

$$\mathbf{y} = \mathbf{r}_0 + \nabla f(\mathbf{r}_0) \cdot (\mathbf{r} - \mathbf{r}_0).$$

Il vettore $\nabla f(\mathbf{r}_0)$ è ortogonale al piano tangente. Per definizione, si chiama il **vettore normale** al grafico di $f(\mathbf{r})$ nel punto $(\mathbf{r}_0, f(\mathbf{r}_0))$. Si veda la figura 4.2. Una funzione dotata di derivate parziali prime continue su un aperto Ω si chiama una funzione **di classe $C^1(\Omega)$** e si scrive $f \in C^1(\Omega)$. Si noti che, per noi, il **gradiente** è un **vettore colonna**. D'altra parte, il differenziale, come trasformazione da \mathbb{R}^n , dotato della base canonica, ad \mathbb{R} , si rappresenta mediante una matrice $1 \times n$, ossia mediante un **vettore riga**. Il **vettore riga** che rappresenta il differenziale è il **trasposto del gradiente**. Diciamo infine che si chiamano **punti stazionari** di $f(\mathbf{r})$ i punti nei quali si annulla il gradiente $\nabla f(\mathbf{r})$. Tali punti si chiamano anche punti **estremali** o punti **critici**.

Figura 4.2: Piano tangente e vettore normale



4.2.2 Regole di derivazione

Dato che la derivata parziale rispetto ad x di una funzione $f(x, y)$ si calcolano fissando prima il valore di y , e lavorando con la funzione della sola x , si hanno immediatamente le regole seguente:

$$\begin{aligned}\frac{\partial}{\partial x} a f(x, y) &= a \frac{\partial}{\partial x} f(x, y) \quad \forall a \in \mathbb{R}; \\ \frac{\partial}{\partial x} (f(x, y) + g(x, y)) &= \frac{\partial}{\partial x} f(x, y) + \frac{\partial}{\partial x} g(x, y); \\ \frac{\partial}{\partial x} (f(x, y)g(x, y)) &= \left(\frac{\partial}{\partial x} f(x, y) \right) g(x, y) + f(x, y) \left(\frac{\partial}{\partial x} g(x, y) \right); \\ \frac{\partial}{\partial t} f(x(t), y) &= \left(\frac{\partial}{\partial x} f(x(t), y) \right) x'(t).\end{aligned}$$

Però nello studio della funzione composta si incontrano casi più complessi: può essere che sia x che y vengano a dipendere dalla medesima variabile, ossia che si voglia calcolare

$$\frac{d}{dt} f(x(t), y(t)).$$

In questo caso si ha:

4.2. LE PROPRIETÀ DI DIFFERENZIABILITÀ

Teorema 135 Sia $f(x, y)$ differenziabile in ogni punto di una regione $\Omega \subseteq \mathbb{R}^2$ e sia $(x(t), y(t))$ una funzione derivabile di $t \in (a, b)$ a valori in Ω . Allora per ogni t si ha:

$$\frac{d}{dt}f(x(t), y(t)) = f_x(x(t), y(t))\dot{x}(t) + f_y(x(t), y(t))\dot{y}(t). \quad (4.3)$$

Dim. Si fissi un punto t_0 in cui si vuol calcolare la derivata, e si ponga $a = f_x(x(t_0), y(t_0))$, $b = f_y(x(t_0), y(t_0))$. Si ha:

$$\begin{aligned} f(x(t), y(t)) - f(x(t_0), y(t_0)) &= a[x(t) - x(t_0)] + b[y(t) - y(t_0)] \\ &\quad + o(|(x(t) - x(t_0), y(t) - y(t_0))|) \\ &= [a\dot{x}(t_0) + b\dot{y}(t_0)](t - t_0) + o(t - t_0) + o(|(x(t) - x(t_0), y(t) - y(t_0))|). \end{aligned}$$

Notando che

$$\frac{o(|(x(t) - x(t_0), y(t) - y(t_0))|)}{|(x(t) - x(t_0), y(t) - y(t_0))|} \cdot \frac{|(x(t) - x(t_0), y(t) - y(t_0))|}{t - t_0}$$

si vede che

$$o(|(x(t) - x(t_0), y(t) - y(t_0))|) = o(t - t_0).$$

Infatti

$$\frac{o(|(x(t) - x(t_0), y(t) - y(t_0))|)}{|(x(t) - x(t_0), y(t) - y(t_0))|}$$

tende a zero per il Lemma 122 mentre

$$\frac{|(x(t) - x(t_0), y(t) - y(t_0))|}{t - t_0}$$

rimane limitata, come si vede dalla seconda formula degli incrementi finiti applicata sia ad $x(t)$ che ad $y(t)$:

$$\frac{|(x(t) - x(t_0), y(t) - y(t_0))|}{t - t_0} = \sqrt{(\dot{x}(c))^2 + (\dot{y}(d))^2}.$$

Sia ora

$$g(t) = f(x(t), y(t)).$$

L'uguaglianza

$$\begin{aligned} g(t) - g(t_0) &= [a\dot{x}(t_0) + b\dot{y}(t_0)](t - t_0) + o(t - t_0) \\ &\quad + o(|(x(t) - x(t_0), y(t) - y(t_0))|) \\ &= [a\dot{x}(t_0) + b\dot{y}(t_0)](t - t_0) + o(t - t_0) \end{aligned}$$

mostra che $g(t)$ è derivabile per $t = t_0$, con

$$\dot{g}(t) = \frac{d}{dt}f(x(t), y(t)) = a\dot{x}(t_0) + b\dot{y}(t_0). \quad \blacksquare$$

Osservazione 136 Si noti che usando il Lemma 122, si è implicitamente assunto che in un opportuno intorno di t_0 la funzione $(x(t), y(t))$ non prenda valore $(x(t_0), y(t_0))$. Si provi per esercizio che questa condizione può rimuoversi. \blacksquare

Naturalmente, il Teorema 135 si estende al caso di funzioni di n variabili e, ricordando il Teorema 134, possiamo enunciare:

Corollario 137 Sia $f(\mathbf{r}) \in C^1(\Omega)$ e sia $\mathbf{r}(t)$ una funzione di $t \in (a, b)$ a valori in Ω , derivabile. Allora, la funzione composta $f(\mathbf{r}(t))$ è derivabile su (a, b) ed inoltre:

$$\frac{d}{dt}f(\mathbf{r}(t)) = \nabla f(\mathbf{r}(t)) \cdot \dot{\mathbf{r}}(t)$$

(il simbolo \cdot indica il prodotto scalare).

Sia ora $x = x(u, v)$, $y = y(u, v)$. Applicando il teorema precedente alla variabile u , con v fissato, e quindi alla variabile v , con u fissato si trova:

Teorema 138 Sia $f(x, y) \in C^1(\Omega)$. Siano $x = x(u, v)$, $y = y(u, v)$ due funzioni definite su $\tilde{\Omega}$ e di classe $C^1(\tilde{\Omega})$, a valori in Ω . Valgono le uguaglianze

$$\begin{aligned} \frac{\partial}{\partial u}f(x(u, v), y(u, v)) &= f_x(x(u, v), y(u, v))x_u(u, v) + f_y(x(u, v), y(u, v))y_u(u, v) \\ &= \nabla f(x(u, v), y(u, v)) \cdot \begin{bmatrix} x_u(u, v) \\ y_u(u, v) \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial v}f(x(u, v), y(u, v)) &= f_x(x(u, v), y(u, v))x_v(u, v) + f_y(x(u, v), y(u, v))y_v(u, v) \\ &= \nabla f(x(u, v), y(u, v)) \cdot \begin{bmatrix} x_v(u, v) \\ y_v(u, v) \end{bmatrix} \end{aligned}$$

(il punto indica il prodotto scalare di vettori colonna). Questa formula si chiama anche formula della derivazione a catena. Consideriamo ora un caso particolare: sia $x = tv^1$, $y = tv^2$. In questo caso, i punti $(x_0 + tv^1, y_0 + tv^2)$ sono punti di una retta nella direzione del vettore $\mathbf{v} = (v^1, v^2)$, uscente da (x_0, y_0) . La derivata

$$\frac{d}{dt}f(x_0 + tv^1, y_0 + tv^2)$$

4.2. LE PROPRIETÀ DI DIFFERENZIABILITÀ

calcolata per $t = 0$ si chiama la *derivata secondo il* vettore \mathbf{v} della funzione e si indica col simbolo

$$\frac{\partial f}{\partial \mathbf{v}}(x, y).$$

Dal Teorema 135 segue che, se $f(x, y)$ è di classe C^1 ,

$$\frac{\partial f}{\partial \mathbf{v}}(x, y) = f_x(x, y)v^1 + f_y(x, y)v^2 = \nabla f(x, y) \cdot \mathbf{v}.$$

Se \mathbf{v} è un *versore*, allora si parla di *derivata* direzionale nella direzione \mathbf{v} . E' facile estendere gli argomenti precedenti a funzioni di tre o più variabili.

4.2.3 La direzione del gradiente e la direzione di massima velocità crescita

Sia $f(\mathbf{r})$ una funzione differenziabile a valori reali e sia \mathbf{r}_0 un punto del suo dominio, che supponiamo aperto. Fissiamo un segmento uscente da \mathbf{r}_0 , dato da

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}, \quad t \in (-\epsilon, \epsilon). \quad (4.4)$$

Il numero ϵ si sceglie piccolo, in modo che il segmento sia contenuto nel dominio della funzione, e \mathbf{v} è un versore. La derivata in $t = 0$ della funzione $f(\mathbf{r}_0 + t\mathbf{v})$ rappresenta la velocità di variazione dei valori della restrizione di $f(\mathbf{r})$ al segmento (4.4). E'

$$\frac{d}{dt}f(\mathbf{r}_0 + t\mathbf{v})|_{t=0} = \nabla f(\mathbf{r}_0) \cdot \mathbf{v}.$$

Questo numero rappresenta la componente di $\nabla f(\mathbf{r}_0)$ lungo \mathbf{v} e quindi il suo valore assoluto è sempre minore di $|\nabla f(\mathbf{r}_0)|$; è uguale quando accade che \mathbf{v} è il versore

$$\mathbf{v} = \frac{\nabla f(\mathbf{r}_0)}{|\nabla f(\mathbf{r}_0)|}.$$

In questo caso si ha anche

$$\nabla f(\mathbf{r}_0) \cdot \mathbf{v} = |\nabla f(\mathbf{r}_0)|;$$

ossia, la funzione $f(\mathbf{r})$ ha massima velocità di crescita nella direzione e verso di $\nabla f(\mathbf{r}_0)$; nel verso opposto essa ha massima velocità di decrescita.

4.2.4 Le funzioni definite tramite integrali

Sia $f(x, y)$ una funzione continua su un rettangolo chiuso $[a, b] \times [c, d]$. Per ogni fissata coppia di numeri $x \in [a, b]$ e $y \in [c, d]$ si può definire il numero

$$\phi(x, y) = \int_c^y f(x, s) \, ds.$$

Si trova così una seconda funzione definita su $[a, b] \times [c, d]$. Vale:

Teorema 139 *La funzione $f(x, y)$ sia continua sul rettangolo chiuso R . Allora, la funzione $\phi(x, y)$ è continua in R .*

Dim. Si scriva

$$|\phi(x, y) - \phi(x', y')| \leq |\phi(x, y) - \phi(x', y)| + |\phi(x', y) - \phi(x', y')|.$$

Vogliamo provare che **per ogni** $\eta > 0$ **esiste** $\sigma > 0$ tale che se

$$d((x, y), (x', y')) < \sigma$$

allora vale

$$|\phi(x, y) - \phi(x', y)| < \eta/2, \tag{4.5}$$

$$|\phi(x', y) - \phi(x', y')| < \eta/2. \tag{4.6}$$

Studiamo (4.5):

$$\begin{aligned} |\phi(x, y) - \phi(x', y)| &= \left| \int_c^y f(x, s) \, ds - \int_c^y f(x', s) \, ds \right| \\ &= \left| \int_c^y [f(x, s) - f(x', s)] \, ds \right| \leq \int_c^y |f(x, s) - f(x', s)| \, ds \\ &\leq \int_c^d |f(x, s) - f(x', s)| \, ds. \end{aligned}$$

La funzione $f(x, y)$ è continua sul rettangolo limitato e chiuso R e quindi è uniformemente continua: per ogni $\delta > 0$ esiste $\sigma > 0$ tale che

$$|(x, y) - (x', y')| < \sigma \implies |f(x, y) - f(x', y')| < \delta.$$

La condizione

$$|(x, y) - (x', y')| < \sigma$$

4.2. LE PROPRIETÀ DI DIFFERENZIABILITÀ

vale in particolare se $y = y' \in [c, d]$ e se $|x - x'| < \sigma$. Dunque,

$$|x - x'| < \sigma \implies |f(x, s) - f(x', s)| < \delta \implies \int_c^d |f(x, s) - f(x', s)| \, ds < (d - c)\delta.$$

L'asserò segue scegliendo

$$\delta = \frac{\eta}{2(d - c)}.$$

Studiamo il secondo addendo (4.6). Per fissare le idee sia $y < y'$:

$$|\phi(x', y) - \phi(x', y')| = \left| \int_c^y f(x', s) \, ds - \int_c^{y'} f(x', s) \, ds \right| \leq \int_y^{y'} |f(x', s)| \, ds.$$

La funzione $f(x, y)$ è continua sul rettangolo limitato e chiuso R . Dunque, per il *Teorema di Weierstrass*, è limitata. Sia $|f(x, y)| < M$. Per la *monotonia dell'integrale* si ha

$$|\phi(x', y) - \phi(x', y')| \leq M(y' - y).$$

Per avere questa differenza minore di $\eta/2$ basta imporre la condizione $|y' - y| < \eta/2M$. ■ Di conseguenza, dal *teorema sulla continuità delle funzioni composte*, viene ad essere continua anche la funzione

$$\int_c^{g(x,y)} f(x, s) \, ds,$$

per ogni funzione continua $g(x, y)$. Scegliendo in particolare $g(x, y)$ costantemente uguale a b si trova:

Teorema 140 *Sia $f(x, y)$ continua su $[a, b] \times [c, d]$ e sia*

$$\phi(x) = \int_a^b f(x, y) \, dy.$$

La funzione $\phi(x)$ è continua su $[a, b]$.

Studiamo ora la derivabilità:

Teorema 141 *Il rettangolo chiuso R sia interno ad una regione Ω su cui la funzione $f(x, y)$ ammette derivate parziali continue. Si ha:*

$$\begin{aligned} \frac{\partial}{\partial x} \phi(x, y) &= \frac{\partial}{\partial x} \int_c^y f(x, s) \, ds = \int_c^y f_x(x, s) \, ds \\ \frac{\partial}{\partial y} \phi(x, y) &= \frac{\partial}{\partial y} \int_c^y f(x, s) \, ds = f(x, y). \end{aligned}$$

Dim. La seconda uguaglianza è ovvia perchè il calcolo dell'integrale e quello della derivata parziale rispetto ad y si fanno per ogni x fissata; e quindi di fatto si lavora con funzioni della sola variabile y . Proviamo la prima uguaglianza. Scriviamo

$$\frac{\phi(x+h, y) - \phi(x, y)}{h} - \int_c^y f_x(x, s) \, ds = \int_c^y \left[\frac{f(x+h, s) - f(x, s)}{h} - f_x(x, s) \right] \, ds.$$

Vogliamo provare che questa differenza tende a zero per $h \rightarrow 0$. Fissiamo $s \in [c, d]$ e scriviamo la formula della media per la funzione della sola variabile x :

$$\frac{f(x+h, s) - f(x, s)}{h} - f_x(x, s) = f_x(\tilde{x}, s) - f_x(x, s)$$

ove \tilde{x} dipende sia da h che da s . E' però un punto dell'intervallo $(x, x+h)$. La funzione $f_x(x, y)$ è uniformemente continua su R e quindi, dato $\epsilon > 0$, esiste $\delta_\epsilon > 0$ tale che

$$\text{se } |h| < \delta_\epsilon \quad \text{vale} \quad |f_x(\tilde{x}, s) - f_x(x, s)| < \epsilon;$$

e quindi,

$$\left| \frac{\phi(x+h, y) - \phi(x, y)}{h} - \int_c^y f_x(x, s) \, ds \right| \leq \int_c^y \epsilon \, ds \leq (d-c)\epsilon.$$

Ciò prova che il limite per h tendente a zero è nullo. ■ Torniamo ora a considerare la funzione

$$\phi(x) = \int_c^d f(x, y) \, dy.$$

Essendo questa funzione continua, essa può venir integrata rispetto alla variabile x :

$$\int_a^b \left[\int_c^d f(x, y) \, dy \right] \, dx$$

Si chiama questo *integrale iterato* della funzione $f(x, y)$. Naturalmente, si può anche introdurre un secondo integrale iterato,

$$\int_c^d \left[\int_a^b f(x, y) \, dx \right] \, dy$$

e si pone il problema di sapere se i loro valori coincidano o meno. La risposta affermativa si vedrà nel cap. 7.

4.3 Le derivate di ordine superiore

Ricordiamo che le derivate parziali sono a loro volta funzioni di più variabili, e quindi può accadere che esse siano ulteriormente derivabili. Si possono quindi definire le funzioni

$$f_{xx}(x, y) = \frac{\partial}{\partial x} \left[\frac{\partial}{\partial x} f(x, y) \right], \quad f_{xy}(x, y) = \frac{\partial}{\partial x} \left[\frac{\partial}{\partial y} f(x, y) \right] \quad f_{yy}(x, y) = \frac{\partial}{\partial y} \left[\frac{\partial}{\partial y} f(x, y) \right]$$

e così via. Quando una funzione ammette tutte le derivate parziali continue fino all'ordine k incluso su un aperto Ω , si dice che essa è di classe $C^k(\Omega)$. La notazione

$$\frac{\partial^n}{\partial x^n} f(x, y)$$

indica la derivata n -ma di $f(x, y)$ rispetto alla variabile x , ossia la derivata ottenuta tenendo fisso il valore di y e derivando n volte rispetto ad x . Significato analogo ha il simbolo

$$\frac{\partial^n}{\partial y^n} f(x, y).$$

Le cose sono più complicate se si vogliono le *derivate miste*, ossia le derivate ottenute derivando alcune volte rispetto ad x e anche rispetto ad y perché in generale dovremo tener conto dell'ordine con cui si eseguono le derivate. Per esempio, in generale

$$D_x(D_y f(x, y)) = \frac{\partial}{\partial x} \left[\frac{\partial}{\partial y} f(x, y) \right] \neq \frac{\partial}{\partial y} \left[\frac{\partial}{\partial x} f(x, y) \right] = D_y(D_x f(x, y)).$$

Fortunatamente, ciò non avviene nei casi più interessanti per le applicazioni. Vale infatti:

Teorema 142 (di Schwarz) *Sia $f(x, y)$ una funzione delle due variabili (x, y) , di classe C^1 un intorno V di (x_0, y_0) . Supponiamo che nei punti di V esistano le derivate miste $D_x(D_y f(x, y))$ ed $D_y(D_x f(x, y))$, e supponiamo che queste siano continue. Allora, per ogni (x, y) vale*

$$D_x(D_y f(x, y)) = D_y(D_x f(x, y)).$$

Esaminando la dimostrazione, in appendice, si vede facilmente che il teorema si estende al caso di funzioni di più di due variabili. Grazie a questo teorema, le due derivate miste rispetto ad x ed y di $f(x, y)$ si indicano semplicemente con i simboli

$$f_{x,y}(x, y), \quad \frac{\partial^2}{\partial x \partial y} f(x, y),$$

senza preoccuparsi dell'ordine di derivazione. Il Teorema di Schwarz si estende a funzioni di n variabili, ed a qualsiasi ordine di derivazione, come segue:

Corollario 143 Sia $f(\mathbf{r}) \in C^n(\Omega)$. Si derivi la funzione k_1 volte rispetto alla prima componente di \mathbf{r} ; k_2 rispetto alla seconda ecc., ma con

$$k_1 + k_2 + \cdots + k_n \leq n.$$

La derivata che si ottiene non dipende dall'ordine col quale si eseguono le derivate.

Indicare derivate successive di funzioni di più variabili è alquanto noioso. Un simbolo comodo si ottiene in questo modo. Sia $f(\mathbf{r})$ una funzione di $\mathbf{r} \in \mathbb{R}^n$. Chiamiamo multiindice un vettore $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ le cui componenti sono numeri interi nulli o positivi. Indichiamo con $|\alpha|$ la lunghezza del multiindice:

$$|\alpha| = \alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_n.$$

Col simbolo \mathbf{r}^α intendiamo:

$$\mathbf{r}^\alpha = (r_1^{\alpha_1}, r_2^{\alpha_2}, \dots, r_n^{\alpha_n}).$$

Col simbolo

$$D^\alpha f(\mathbf{r}) = \frac{\partial^{|\alpha|}}{\partial \mathbf{r}^\alpha} f(\mathbf{r})$$

si intende la derivata di $f(\mathbf{r})$ che si ottiene derivando la funzione $f(\mathbf{r})$ nell'ordine, prima α_1 volte rispetto alla prima variabile, poi α_2 volte rispetto alla seconda variabile, ecc. (intendendo che se $\alpha_i = 0$ la corrispondente derivata non si esegue). Il Teorema di Schwarz assicura che se ciascuna derivata di ordine $|\alpha|$ di $f(\mathbf{r})$ è continua allora le derivate miste di ordine $|\alpha|$ non dipendono dall'ordine con cui vengono calcolate. Ciò giustifica la notazione $f \in C^k(\Omega)$, che si usa quando la funzione $f(\mathbf{r})$ ammette tutte le derivate di ordine k su Ω , e queste sono continue.

4.3.1 La formula di Taylor per le funzioni a valori reali

Così come nel caso delle funzioni di una variabile, gli argomenti che hanno condotto alla prima formula degli incrementi finiti possono iterarsi se la funzione ammette continue le successive N derivate. Senza entrare nei dettagli della dimostrazione (simile a quella del Teorema 132), limitiamoci a dare la formula che si ottiene nel caso delle derivate seconde. Quando la funzione dipende da

4.3. LE DERIVATE DI ORDINE SUPERIORE

due variabili, si trova

$$\begin{aligned}
 f(x, y) &= f(x_0, y_0) \\
 &+ [f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)] \\
 &+ \frac{1}{2} \{ f_{xx}(x_0, y_0)(x - x_0)^2 + f_{xy}(x_0, y_0)(x - x_0)(y - y_0) \\
 &+ f_{yx}(x_0, y_0)(x - x_0)(y - y_0) + f_{yy}(x_0, y_0)(y - y_0)^2 \} + R(x, y).
 \end{aligned}$$

e

$$\lim_{d((x,y),(x_0,y_0)) \rightarrow 0} \frac{R(x, y)}{[d((x, y), (x_0, y_0))]^2} = 0, \quad \text{ossia} \quad R(x, y) = o(|(x - x_0, y - y_0)|^2). \quad (4.7)$$

Grazie al Teorema di Schwarz, la formula precedente si può scrivere

$$\begin{aligned}
 f(x, y) &= f(x_0, y_0) \\
 &+ [f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)] \\
 &+ \frac{1}{2} \{ f_{xx}(x_0, y_0)(x - x_0)^2 + 2f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + f_{yy}(x_0, y_0)(y - y_0)^2 \} \\
 &+ R(x, y).
 \end{aligned}$$

Questa formula si chiama ancora *formula di Taylor* arrestata al secondo ordine perché il resto $R(x, y)$ verifica la condizione (4.7). Naturalmente, aggiungendo i termini con le derivate rispetto ad una terza variabile z si trova la formula di Taylor in tre variabili e, in generale, in n variabili. Si immagina facilmente che le formule divengano via via più complesse, in particolare se la funzione dipende da più di due variabili e si vuole scrivere la formula di Taylor di ordine maggiore di 2. E' però possibile scrivere queste formule in modo compatto come segue. Riguardiamo il caso della formula di Taylor di $f(x, y)$ arrestata all'ordine 2. Si noti che il termine

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

si può pensare ottenuto come segue:

$$\frac{1}{1!} \left[(x - x_0) \frac{\partial}{\partial x} + (y - y_0) \frac{\partial}{\partial y} \right] f(x, y) = \frac{1}{1!} [f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)].$$

Consideriamo ora il termine di ordine 2:

$$\frac{1}{2} \{ f_{xx}(x_0, y_0)(x - x_0)^2 + 2f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + f_{yy}(x_0, y_0)(y - y_0)^2 \}$$

si può ottenere con questa regola mnemonica: si “calcola”

$$\frac{1}{2!} \left[(x - x_0) \frac{\partial}{\partial x} + (y - y_0) \frac{\partial}{\partial y} \right]^2$$

scrivendo

$$\frac{1}{2!} \left[(x - x_0)^2 \frac{\partial^2}{\partial x^2} + 2(x - x_0)(y - y_0) \frac{\partial^2}{\partial x \partial y} + (y - y_0)^2 \frac{\partial^2}{\partial y^2} \right].$$

Si “applica” quindi questo alla funzione $f(x, y)$ scrivendo

$$\begin{aligned} & \frac{1}{2!} \left[(x - x_0)^2 \frac{\partial^2}{\partial x^2} + 2(x - x_0)(y - y_0) \frac{\partial^2}{\partial x \partial y} + (y - y_0)^2 \frac{\partial^2}{\partial y^2} \right] f(x, y) \\ &= \frac{1}{2!} \left[(x - x_0)^2 f_{xx}(x_0, y_0) + 2(x - x_0)(y - y_0) f_{xy}(x_0, y_0) + (y - y_0)^2 f_{yy}(x_0, y_0) \right]. \end{aligned}$$

Questa “regola mnemonica” si estende al caso dell’ordine maggiore di 2 e fornisce l’espressione corretta della formula di Taylor arrestata all’ordine k :

$$f(x, y) = \sum_{j=0}^k \frac{1}{j!} \left[(x - x_0) \frac{\partial}{\partial x} + (y - y_0) \frac{\partial}{\partial y} \right]^j f(x, y) + o \left(\sqrt{(x - x_0)^2 + (y - y_0)^2} \right)^k.$$

In questa formula si intende che

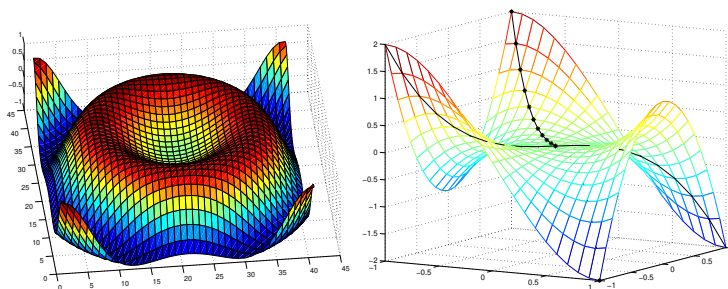
$$\left[(x - x_0) \frac{\partial}{\partial x} + (y - y_0) \frac{\partial}{\partial y} \right]^0 f(x, y) = f(x_0, y_0).$$

L’estensione di questa formula al caso di funzioni dipendenti da più di due variabili è ovvia.

4.4 Gli estremi

La formula di Taylor, arrestata al primo ordine, ha già dato un’informazione importante. Infatti, ha condotto a provare che ogni funzione le cui derivate parziali sono ovunque continue è essa stessa continua. Daltra parte, nel caso delle funzioni di una sola variabile, la formula di Taylor si usa per dedurre regole che possono condurre ad identificare i punti di massimo o di minimo. Si può immaginare che anche per funzioni di più variabili si possa fare uno studio analogo. Per renderci conto di ciò, ripetiamo la definizione degli estremi:

Figura 4.3: Un punto di minimo e un punto di sella



Definizione 2 Sia $f(\mathbf{r})$ una funzione da $\mathbb{R}^n \rightarrow \mathbb{R}$. Un punto \mathbf{r}_0 interno al dominio di $f(\mathbf{r})$ si dice *punto di minimo* (relativo) per la funzione $f(\mathbf{r})$ se esiste $\nu > 0$ tale che

$$d(\mathbf{r}, \mathbf{r}_0) < \nu \implies f(\mathbf{r}_0) \leq f(\mathbf{r}).$$

In modo analogo si definiscono i *punti di massimo*. ■

La figura 4.3 (a sinistra) mostra un punto di minimo. Quando l'intorno di \mathbf{r}_0 coincide con Ω , si parla di *estremi assoluti* o *estremi globali*. E naturalmente a questo caso ci si può ricondurre, sostituendo Ω con tale intorno. Studiamo ora il caso degli estremi relativi di funzioni di due variabili, che sono interni al dominio. Notiamo che se (x_0, y_0) è punto di minimo, allora la funzione $g(x)$ della sola variabile x ,

$$g(x) = f(x, y_0)$$

ha minimo in x_0 ed x_0 è punto interno al dominio. Dunque, se è derivabile, essa ha derivata nulla in x_0 , per il Teorema di Fermat. Per definizione, $g'(x_0) = f_x(x_0, y_0)$. Se (x_0, y_0) è punto di minimo, esso è anche punto di minimo per la funzione

$$h(y) = f(x_0, y)$$

e quindi $h'(y_0) = 0$; e naturalmente il discorso si ripete se (x_0, y_0) è punto di massimo per $f(x, y)$. Dunque:

Teorema 144 Se la funzione $f(x, y)$ ammette derivate parziali prime nel punto di minimo (o di massimo) (x_0, y_0) interno al dominio, esse sono ambedue nulle:

$$f_x(x_0, y_0) = 0, \quad f_y(x_0, y_0) = 0. \quad \blacksquare$$

Ossia: gli estremi di una funzione vanno ricercati tra i suoi punti stazionari. Naturalmente, un punto stazionario, ossia un punto nel quale si annulla il gradiente della funzione, non è necessariamente un estremo. Una condizione sufficiente per gli estremi si vedrà più avanti. Però, in generale la funzione può avere un comportamento quanto mai vario nei punti nei quali si annulla il gradiente. In particolare definiamo:

Definizione 3 Si dice che \mathbf{r}_0 è punto di sella se è un punto stazionario e se, inoltre, la funzione $f(\mathbf{r}) - f(\mathbf{r}_0)$ cambia segno in ogni intorno \mathbf{r}_0 . ■

La figura 4.3, a destra, illustra un punto di sella. Dunque il piano tangente in un punto di minimo o di massimo o di sella (interno al dominio) è orizzontale e, se la funzione è di classe C^2 , può scriversi:

$$f(x, y) = f(x_0, y_0) + \frac{1}{2} \{ f_{xx}(x_0, y_0)(x - x_0)^2 + 2f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + f_{yy}(x_0, y_0)(y - y_0)^2 \} + R(x, y)$$

con $R(x, y)/[(x - x_0)^2 + (y - y_0)^2]$ infinitesima. Per capire se è possibile ripetere, per le funzioni di più variabili, uno studio analogo a quello delle funzioni di una variabile, consideriamo il caso particolare in cui $x_0 = 0$, $y_0 = 0$, $f_{xx}(0, 0) = 1$, $f_{xy}(0, 0) = 0$, $f_{yy}(0, 0) = 1$. In questo caso,

$$\begin{aligned} f(x, y) &= f(0, 0) + \frac{1}{2} \{ f_{xx}(0, 0)x^2 + 2f_{xy}(0, 0)xy \\ &\quad + f_{yy}(0, 0)y^2 \} + R(x, y) \\ &= f(0, 0) + (x^2 + y^2) \cdot \left(\frac{1}{2} + \frac{R(x, y)}{x^2 + y^2} \right) \end{aligned}$$

Il teorema della permanenza del segno mostra che per $x^2 + y^2$ abbastanza piccolo,

$$1 + \frac{1}{2} \frac{R(x, y)}{x^2 + y^2} > 0$$

e quindi $f(x, y)$ ha, in $(0, 0)$, un punto di minimo relativo, cosiccome la funzione di confronto $g(x, y) = x^2 + y^2$. In generale se $f(x, y)$ è di classe C^2 e le sue derivate prime sono ambedue nulle in un punto (x_0, y_0) dalla formula di Taylor si ha

$$f(x, y) = f(x_0, y_0) + g(x - x_0, y - y_0) + o(|(x - x_0, y - y_0)|^2)$$

e la funzione di confronto $g(x, y)$ ha forma

$$g(x, y) = \frac{1}{2} \{ f_{xx}(x_0, y_0)(x - x_0)^2 + 2f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + f_{yy}(x_0, y_0)(y - y_0)^2 \}.$$

4.4. GLI ESTREMI

Le proprietà che il punto (x_0, y_0) ha per le funzioni

$$g(x, y) = a(x - x_0)^2 + 2b(x - x_0)(y - y_0) + c(y - y_0)^2$$

sono note dai corsi di geometria, caratterizzate mediante la matrice

$$H = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

Con una dimostrazione del tutto analoga a quella vista sopra in un caso particolare, si prova: Se questa matrice è definita **positiva**, il punto è di **minimo**; e tale proprietà è ereditata dalla funzione $f(x, y)$; se la matrice è **indefinita** il punto è di **sella**, e tale proprietà è ereditata dalla funzione $f(x, y)$; se la matrice è definita **negativa**, il punto è di **massimo**; e tale proprietà è ereditata dalla funzione $f(x, y)$. Se invece la matrice è **semidefinita positiva** oppure **semidefinita negativa**, **NIENTE PUÒ DIRSI SULLA FUNZIONE** $f(x, y)$. Ricordando i criteri introdotti nei corsi di geometria, per lo studio della definiteness delle matrici, si ha:

- se $a > 0$ e $\det H > 0$ la matrice H è definita positiva, ed il punto è di minimo.
- se $a < 0$ e $\det H > 0$ si ha un punto di minimo.
- se $ac = 0$ ma $\det H \neq 0$ la matrice è indefinita ed il punto è di sella.
- se $\det H = 0$ la funzione $g(x, y)$ ha per grafico un cilindro parabolico e niente può dirsi della funzione $f(x, y)$.

Questo risultato può estendersi a funzioni di n variabili. Introduciamo per questo la *matrice* hessiana della funzione $f(\mathbf{r})$, $\mathbf{r} = (x^1, \dots, x^n)$:

$$H(\mathbf{r}) = \begin{bmatrix} f_{x^1, x^1}(\mathbf{r}) & f_{x^1, x^2}(\mathbf{r}) & \dots & f_{x^1, x^n}(\mathbf{r}) \\ f_{x^2, x^1}(\mathbf{r}) & f_{x^2, x^2}(\mathbf{r}) & \dots & f_{x^2, x^n}(\mathbf{r}) \\ \vdots & & & \\ f_{x^n, x^1}(\mathbf{r}) & f_{x^n, x^2}(\mathbf{r}) & \dots & f_{x^n, x^n}(\mathbf{r}) \end{bmatrix}.$$

Vale:

Teorema 145 *Siano nulle le derivate parziali prime di $f(\mathbf{r})$ in \mathbf{r}_0 . In tal caso:*

- se la matrice hessiana è definita positiva, il punto è di minimo per $f(\mathbf{r})$;
- se la matrice hessiana è indefinita il punto è di sella per $f(\mathbf{r})$;

- se la matrice hessiana è definita negativa, il punto è di massimo per $f(x, y)$.

Invece, **NIENTE PUÒ DIRSI** se la matrice hessiana è soltanto **semidefinita positiva oppure negativa**.

Osservazione 146 E' possibile dare ulteriori condizioni necessarie che devono essere soddisfatte nei punti di minimo oppure di massimo. Ricordiamo che in un punto di minimo $\mathbf{r}_0 = (x_0, y_0)$ le derivate parziali prime devono essere tutte nulle. Questa condizione **necessaria** si ottiene facilmente considerando la restrizione della funzione ad una delle rette coordinate passanti per il punto; ossia, considerando la funzione $f(x, y_0)$ si trova che deve essere $f_x(x_0, y_0) = 0$. Se esiste la derivata seconda, $f_{xx}(x_0, y_0)$ non può essere negativa, altrimenti la funzione $f(x, y_0)$ avrebbe in x_0 un punto di massimo, invece che di minimo. Dunque, in un punto di minimo (x_0, y_0) si deve avere

$$f_{xx}(x_0, y_0) \geq 0, \quad f_{yy}(x_0, y_0) \geq 0 \quad \text{e quindi anche } f_{xx}(x_0, y_0) + f_{yy}(x_0, y_0) \geq 0.$$

In un punto di massimo deve aversi invece

$$f_{xx}(x_0, y_0) \leq 0, \quad f_{yy}(x_0, y_0) \leq 0 \quad \text{e quindi anche } f_{xx}(x_0, y_0) + f_{yy}(x_0, y_0) \leq 0.$$

Questa condizione si generalizza immediatamente al caso di funzioni dipendenti da più variabili. Un'ulteriore condizione ancora solamente necessaria ma più precisa è la seguente: **In un punto di minimo le derivate parziali prime sono nulle e inoltre la matrice hessiana è definita positiva oppure semidefinita positiva.** Infatti, se la matrice hessiana fosse definita negativa si avrebbe un punto di massimo, se fosse indefinita si avrebbe un punto di sella. Analogamente, **In un punto di massimo le derivate parziali prime sono nulle e inoltre la matrice hessiana è definita negativa oppure semidefinita negativa.** ■

4.5 Il differenziale delle funzioni a valori in \mathbb{R}^m

Sia $\mathbf{F}(\mathbf{r})$ una funzione da \mathbb{R}^n ad \mathbb{R}^m e sia $\mathbf{J}(\mathbf{r})$ una trasformazione lineare da \mathbb{R}^n ad \mathbb{R}^m . Si dice che la trasformazione lineare $\mathbf{J}(\mathbf{r})$ è il differenziale di $\mathbf{F}(\mathbf{r})$ in \mathbf{r}_0 se

$$|\mathbf{F}(\mathbf{r}) - \mathbf{F}(\mathbf{r}_0) - \mathbf{J}(\mathbf{r} - \mathbf{r}_0)| = o(|\mathbf{r} - \mathbf{r}_0|).$$

Ovviamente, la i -ma componente $F_i(\mathbf{r})$ è una funzione a valori reali e dalla disuguaglianza

$$|F^i| \leq |\mathbf{F}| \quad \text{se} \quad \mathbf{F} = (F^1, F^2, \dots, F^m)$$

4.5. IL DIFFERENZIALE DELLE FUNZIONI A VALORI IN \mathbb{R}^M

segue

$$|F^i(\mathbf{r}) - F^i(\mathbf{r}_0) - J^i(\mathbf{r} - \mathbf{r}_0)| = o(|\mathbf{r} - \mathbf{r}_0|).$$

ove $J^i(\mathbf{r})$ è la i -ma componente di $\mathbf{J}(\mathbf{r})$. Segue che una funzione da \mathbb{R}^n in \mathbb{R}^m che è differenziabile ha differenziabile, e quindi continua, ciascuna delle sue componenti:

Teorema 147 *Se la funzione $\mathbf{F}(\mathbf{r})$ è differenziabile in \mathbf{r}_0 , essa è ivi continua.*

Viceversa, dalle disuguaglianze (3.3), esiste un M per cui

$$|\mathbf{F}(\mathbf{r}) - \mathbf{F}(\mathbf{r}_0) - \mathbf{J}(\mathbf{r} - \mathbf{r}_0)| \leq M \sum_{i=1}^m |F^i(\mathbf{r}) - F^i(\mathbf{r}_0) - J^i(\mathbf{r} - \mathbf{r}_0)|.$$

Dunque:

Teorema 148 *La funzione $\mathbf{F}(\mathbf{r})$ è differenziabile in \mathbf{r}_0 se e solo se ciascuna delle sue componenti lo è.*

Ricordiamo che le trasformazioni lineari da \mathbb{R}^n ad \mathbb{R}^m (che rappresentiamo rispetto alle basi canoniche) si rappresentano mediante matrici. Sia J la matrice della trasformazione lineare $\mathbf{J}(\mathbf{r})$, differenziale in \mathbf{r}_0 della funzione $\mathbf{F}(\mathbf{r})$. La sua i -ma riga rappresenta il differenziale della componente $F^i(\mathbf{r})$ e quindi è il trasposto del gradiente della funzione $F^i(\mathbf{r})$; e quindi

$$J(\mathbf{r}_0) = \begin{bmatrix} \frac{\partial}{\partial x^1} F^1(\mathbf{r}_0) & \frac{\partial}{\partial x^2} F^1(\mathbf{r}_0) & \dots & \frac{\partial}{\partial x^n} F^1(\mathbf{r}_0) \\ \frac{\partial}{\partial x^1} F^2(\mathbf{r}_0) & \frac{\partial}{\partial x^2} F^2(\mathbf{r}_0) & \dots & \frac{\partial}{\partial x^n} F^2(\mathbf{r}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x^1} F^m(\mathbf{r}_0) & \frac{\partial}{\partial x^2} F^m(\mathbf{r}_0) & \dots & \frac{\partial}{\partial x^n} F^m(\mathbf{r}_0) \end{bmatrix}$$

La matrice J si chiama la *matrice jacobiana* della funzione $\mathbf{F}(\mathbf{r})$, calcolata in \mathbf{r}_0 . Quando $n = m$, il suo determinante si chiama lo *jacobiano* della trasformazione. Per mezzo della matrice jacobiana, possiamo scrivere

$$\mathbf{F}(\mathbf{r}) = \mathbf{F}(\mathbf{r}_0) + J(\mathbf{r}_0)(\mathbf{r} - \mathbf{r}_0) + o(|\mathbf{r} - \mathbf{r}_0|). \quad (4.8)$$

Quest'uguaglianza si chiama ancora *formula degli incrementi finiti*. Ovviamente, lavorando per componenti, si potrà scrivere anche una formula di Taylor per funzioni da \mathbb{R}^n in \mathbb{R}^m , su cui non indugiamo. Come nel caso delle funzioni a valori in \mathbb{R} , la sola esistenza delle derivate parziali non implica né la continuità né la differenziabilità; ma la **continuità delle derivate parziali in un insieme aperto** implica la differenziabilità, su tale insieme, di ciascuna delle componenti $F^i(\mathbf{r})$ e quindi di $\mathbf{F}(\mathbf{r})$, ossia:

Teorema 149 *Le derivate parziali $\frac{\partial}{\partial x^i} F^j(\mathbf{r}_0)$ per $1 \leq i \leq n$, $1 \leq j \leq m$ esistano continue in un intorno di \mathbf{r}_0 . Allora, la funzione $\mathbf{F}(\mathbf{r})$ è differenziabile e quindi continua in \mathbf{r}_0 .*

Rappresentiamo \mathbb{R}^n rispetto alla base canonica e sia J una trasformazione lineare di \mathbb{R}^n in sé. Col medesimo simbolo J indichiamo anche la matrice che rappresenta la trasformazione. Si sa che la trasformazione J trasforma un parallelepipedo in un altro e l'area del secondo è uguale all'area del primo moltiplicata per $|\det J|$. Consideriamo ora una trasformazione differenziabile da \mathbb{R}^2 in sé:

$$x = x(u, v), \quad y = y(u, v).$$

Fissiamo un punto $\mathbf{r}_0 = (u_0, v_0)$ ed un rettangolo R con un vertice in \mathbf{r}_0 . L'immagine di R è

$$\left\{ \begin{bmatrix} x(u, v) \\ y(u, v) \end{bmatrix}, (u, v) \in R \right\} = \left\{ \begin{bmatrix} x(u_0, v_0) \\ y(u_0, v_0) \end{bmatrix} + J \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} + \mathbf{o} \left(\left\| \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} \right\| \right) \right\}$$

ove J indica la matrice jacobiana della trasformazione. Quest'insieme non è un rettangolo. E' un insieme dalla struttura più complicata, ma approssimato dal parallelogramma

$$\tilde{R} = \left\{ \begin{bmatrix} x(u_0, v_0) \\ y(u_0, v_0) \end{bmatrix} + J \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} \right\}.$$

Dal paragrafo 3.5

$$(\text{area di } \tilde{R}) = |\det J| \cdot (\text{area di } R).$$

Asserto analogo vale per le trasformazioni di \mathbb{R}^3 . Quest'osservazione verrà utilizzata nel paragrafo 7.3.1.

Esempio 150 La figura 4.4 riporta, a sinistra, il rettangolo $R = [1, 2] \times [1, 3/2]$ del piano (u, v) e a destra la sua immagine mediante la trasformazione

$$x = u + (uv)^2, \quad y = v^2 + uv.$$

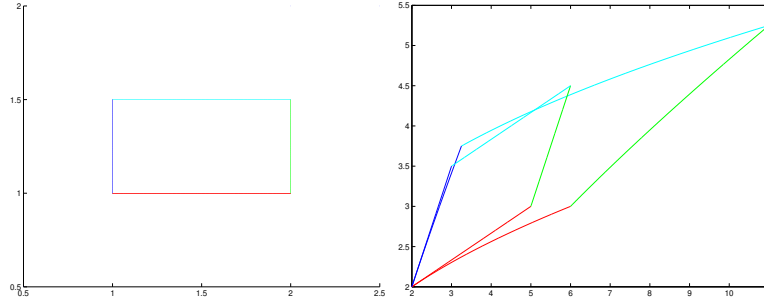
Il parallelogramma a destra è \tilde{R} .

Osservazione 151 Si vede dalla formula degli incrementi finiti che la matrice jacobiana ha, nel caso delle funzioni da \mathbb{R}^n in \mathbb{R}^m , lo stesso ruolo che la derivata ha nel caso $n = m = 1$. Per questa ragione la matrice jacobiana talvolta si indica anche col simbolo

$$\frac{\partial \mathbf{F}(\mathbf{r})}{\partial \mathbf{r}}.$$

4.5. IL DIFFERENZIALE DELLE FUNZIONI A VALORI IN \mathbb{R}^M

Figura 4.4: Un parallelogramma, la sua immagine e il parallelogramma che la approssima



Per esempio, nel caso $n = m = 2$ e $\mathbf{F}(x, y) = [f(x, y) \quad g(x, y)]'$, la matrice jacobiana si trova anche indicata con la notazione

$$\frac{\partial (f(x, y), g(x, y))}{\partial (x, y)} \quad \blacksquare$$

Osservazione 152 (Sulle notazioni) Non esiste una notazione standard per lo jacobiano o per la matrice jacobiana. Nei testi di meccanica del continuo e di scienza delle costruzioni lo jacobiano viene indicato col simbolo ∇ , come il gradiente. Si faccia attenzione però che se lo spazio d'arrivo ha dimensione 1 allora ∇ ="grad" indica un vettore colonna mentre ∇ ="matrice jacobiana" indica un vettore riga². Si faccia anche attenzione a questo: nei testi di meccanica del continuo si considerano funzioni $\mathbf{u} = \mathbf{u}(\mathbf{X}, t)$ con \mathbf{X} e \mathbf{u} in \mathbb{R}^3 e $t \in \mathbb{R}$. La notazione $\nabla \mathbf{u} = \nabla \mathbf{u}(\mathbf{X}, t)$ indica la matrice jacobiana fatta rispetto al vettore \mathbf{X} , ossia

$$\nabla \mathbf{u}(\mathbf{X}, t) = \begin{bmatrix} \frac{\partial u_1(X_1, X_2, X_3, t)}{\partial X_1} & \frac{\partial u_1(X_1, X_2, X_3, t)}{\partial X_2} & \frac{\partial u_1(X_1, X_2, X_3, t)}{\partial X_3} \\ \frac{\partial u_2(X_1, X_2, X_3, t)}{\partial X_1} & \frac{\partial u_2(X_1, X_2, X_3, t)}{\partial X_2} & \frac{\partial u_2(X_1, X_2, X_3, t)}{\partial X_3} \\ \frac{\partial u_3(X_1, X_2, X_3, t)}{\partial X_1} & \frac{\partial u_3(X_1, X_2, X_3, t)}{\partial X_2} & \frac{\partial u_3(X_1, X_2, X_3, t)}{\partial X_3} \end{bmatrix}.$$

I testi di meccanica del continuo usano una notazione interessante per le derivate: le derivate si indicano con indici numerici preceduti da virgola. Per esempio, sia

$$\mathbf{u}(x_1, x_2, x_3) = [u_1(x_1, x_2, x_3) \quad u_2(x_1, x_2, x_3) \quad u_3(x_1, x_2, x_3)]'.$$

²va notato che in certi testi di meccanica del continuo, il gradiente di una funzione a valori in \mathbb{R} si intende essere un vettore riga; ossia in tali libri si confonde il gradiente col vettore che rappresenta il differenziale.

Il simbolo

$$\mathbf{u}_{i,j}$$

indica la derivata rispetto ad x_j della componente u_i ; ossia il numero prima della virgola specifica la componente e quello dopo specifica la variabile rispetto a cui si deriva. Questo simbolo si può iterare e per esempio

$$\mathbf{u}_{2,13} = \frac{\partial^2}{\partial x_1 \partial x_3} u_2(x_1, x_2, x_3). \quad \blacksquare$$

4.5.1 Regole di calcolo della matrice jacobiana

Usiamo il simbolo $J_{\mathbf{F}}(\mathbf{r})$ per indicare la matrice jacobiana della funzione \mathbf{F} nel punto \mathbf{r} . E' immediatamente evidente che

$$\begin{aligned} J_{A\mathbf{F}}(\mathbf{r}) &= AJ_{\mathbf{F}}(\mathbf{r}) && \text{per ogni matrice costante } A; \\ J_{\mathbf{F}+\mathbf{G}}(\mathbf{r}) &= J_{\mathbf{F}}(\mathbf{r}) + J_{\mathbf{G}}(\mathbf{r}). \end{aligned}$$

Supponiamo ora di esprimere $\mathbf{r} \in \mathbb{R}^n$ come funzione

$$\mathbf{r} = \mathbf{r}(\mathbf{u}), \quad \mathbf{u} \in \mathbb{R}^k.$$

Vale

Teorema 153 (della derivazione a catena) Sia $\mathbf{F}(\mathbf{r})$ di classe C^1 in una regione Ω di \mathbb{R}^n ; $\mathbf{r}(\mathbf{u})$ sia di classe C^1 in una regione $\tilde{\Omega} \subseteq \mathbb{R}^k$, ed a valori in Ω . Sia $\mathbf{G}(\mathbf{u}) = \mathbf{F}(\mathbf{r}(\mathbf{u}))$. Si ha

$$J_{\mathbf{G}}(\mathbf{u}_0) = J_{\mathbf{F}}(\mathbf{r}(\mathbf{u}_0))J_{\mathbf{r}}(\mathbf{u}_0).$$

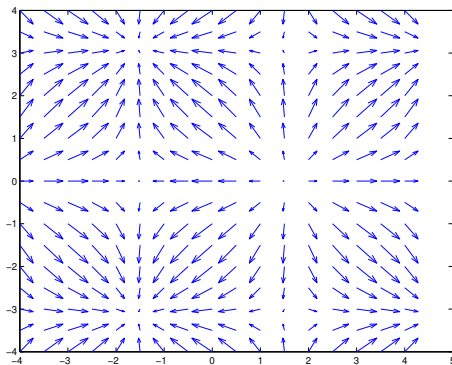
Il prodotto tra matrici si intende come prodotto righe per colonne.

Omettiamo la dimostrazione.

4.6 Campi vettoriali

Nel trattare le funzioni a valori vettoriali, abbiamo sempre specificato “a valori in \mathbb{R}^m ”. Si ricordi che \mathbb{R}^m è uno spazio di **vettori liberi**, equivalentemente, applicati nell'origine. In molte applicazioni si devono considerare funzioni che associano ad un punto \mathbf{r} di \mathbb{R}^n un vettore ad n dimensioni, **applicato nel punto stesso**. Ossia si devono studiare trasformazioni che trasformano \mathbf{r} nella coppia $(\mathbf{r}, \mathbf{V}(\mathbf{r}))$. Trasformazioni di questo tipo si chiamano *campi* vettoriali.

Figura 4.5: un campo vettoriale



E' comune usare la lettera \mathbf{V} per indicare campi vettoriali. Si scrive cioè $\mathbf{V}(\mathbf{r})$ sottintendendo che questo vettore è applicato in \mathbf{r} , invece di scrivere la notazione completa, ma pedante, $(\mathbf{r}, \mathbf{V}(\mathbf{r}))$. La figura 4.5 illustra il modo con cui usualmente si rappresentano i campi vettoriali: da ciascun punto si fa uscire una freccia che indica direzione e verso del vettore. La lunghezza della freccia è proporzionale al modulo. Quando si lavora con campi vettoriali, i concetti di continuità e di differenziabilità sono concettualmente diversi da quelli incontrati per i vettori liberi, perché non si fanno operazioni tra vettori applicati in punti diversi. Per questo la definizione di continuità di un campo vettoriale si definisce confrontando il vettore $\mathbf{V}(\mathbf{r})$, applicato in \mathbf{r} col vettore PARALLELO a $\mathbf{V}(\mathbf{r}')$, anch'esso applicato in \mathbf{r} invece che in \mathbf{r}' . Fortunatamente, la somma di vettori applicati nel medesimo punto si fa operando per componenti, e queste non mutano per spostamenti per parallelismo. Si ha quindi che un campo vettoriale è continuo se e solo se sono continue le sue componenti. Discorso analogo vale per le derivate. Si fissi \mathbf{r}_0 ed un vettore libero \mathbf{v} . Si considerino i vettori

$$\mathbf{r}_0 + t\mathbf{v}$$

che sono i vettori di una retta parallela a \mathbf{v} , passante per \mathbf{r}_0 . Si chiama *derivata secondo il vettore \mathbf{v}* il limite

$$D_{\mathbf{v}}\mathbf{V}(\mathbf{r}_0) = \lim_{t \rightarrow 0} \frac{\mathbf{V}(\mathbf{r}_0 + t\mathbf{v}) - \mathbf{V}(\mathbf{r}_0)}{t}.$$

La differenza al numeratore si calcola supponendo di traslare per parallelismo il vettore $\mathbf{V}(\mathbf{r}_0 + t\mathbf{v})$ applicandolo nel punto \mathbf{r}_0 . Dunque, anche il vettore $D_{\mathbf{v}}\mathbf{V}(\mathbf{r}_0)$ è applicato in \mathbf{r}_0 . Se \mathbf{v} è il versore \mathbf{e}_i , tale derivata si chiama anche *derivata*

direzionale. Il teorema di derivazione della funzione composta immediatamente dà:

Teorema 154 *Sia $\mathbf{V}(\mathbf{r})$ un campo vettoriale le cui componenti sono funzioni di classe C^1 . Esso ammette derivate direzionali in tutte le direzioni, e vale*

$$D_{\mathbf{v}}\mathbf{V}(\mathbf{r}_0) = J_{\mathbf{V}}(\mathbf{r}_0)\mathbf{v} \quad (4.9)$$

In questo teorema, il simbolo $J_{\mathbf{V}}(\mathbf{r}_0)$ indica la matrice jacobiana della trasformazione che al vettore libero \mathbf{r} associa il vettore libero $\mathbf{V}(\mathbf{r})$. Il vettore $D_{\mathbf{v}}\mathbf{V}(\mathbf{r}_0)$ si intende applicato in \mathbf{r}_0 .

4.6.1 Operatori differenziali e campi vettoriali

Abbiamo visto che il gradiente di una funzione a valori scalari è il vettore

$$\nabla f(\mathbf{r}) = \begin{bmatrix} f_{x^1}(\mathbf{r}) \\ f_{x^2}(\mathbf{r}) \\ \vdots \\ f_{x^n}(\mathbf{r}) \end{bmatrix}.$$

Per molte applicazioni, conviene considerare questo un vettore applicato nel punto \mathbf{r} . Dunque il gradiente associa un campo vettoriale ad una funzione a valori scalari. Si noti che

$$[\nabla_{\mathbf{u}}f(\mathbf{r}(\mathbf{u}))]' = [\nabla_{\mathbf{r}}f(\mathbf{r}(\mathbf{u}))]' J_{\mathbf{r}}[\mathbf{u}]$$

ove l'apice indica la trasposizione e l'indice sotto il simbolo ∇ indica le variabili rispetto alle quali si calcola il gradiente.

Introduciamo una notazione comoda: sia $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ la base canonica di \mathbb{R}^n . Scriviamo in modo formale

$$\begin{aligned} \nabla &= \mathbf{e}_1 \frac{\partial}{\partial x^1} + \mathbf{e}_2 \frac{\partial}{\partial x^2} + \dots + \mathbf{e}_n \frac{\partial}{\partial x^n} \\ &= \mathbf{e}_1 \partial_1 + \mathbf{e}_2 \partial_2 + \dots + \mathbf{e}_n \partial_n \\ &= \begin{bmatrix} \partial_1 \\ \partial_2 \\ \vdots \\ \partial_n \end{bmatrix}. \end{aligned}$$

4.6. CAMPI VETTORIALI

Allora il gradiente si ottiene formalmente “moltiplicando” il “vettore” ∇ per lo “scalare” $f(\mathbf{r})$, e eseguendo le operazioni di derivazione. Introdotto il simbolo ∇ , viene naturale introdurre le operazioni

$$\nabla \cdot \mathbf{V}(\mathbf{r}), \quad \nabla \wedge \mathbf{V}(\mathbf{r})$$

operando formalmente come se si trattasse di prodotti scalari o vettoriali (e quindi la seconda si definisce solo in \mathbb{R}^3). La prima dà la divergenza del campo vettoriale:

$$\operatorname{div} \mathbf{V}(\mathbf{r}) = \frac{\partial}{\partial x^1} V^1(\mathbf{r}) + \frac{\partial}{\partial x^2} V^2(\mathbf{r}) + \dots + \frac{\partial}{\partial x^n} V^n(\mathbf{r}).$$

Il risultato della seconda operazione, DEFINITA SOLO IN \mathbb{R}^3 , si chiama rotore: il rotore del campo vettoriale

$$\mathbf{V}(x, y, z) = u(x, y, z)\mathbf{i} + v(x, y, z)\mathbf{j} + w(x, y, z)\mathbf{k}$$

è quindi (sottintendendo le variabili indipendenti) il campo vettoriale

$$\operatorname{rot} \mathbf{V} = \nabla \wedge \mathbf{V} = (w_y - v_z)\mathbf{i} + (u_z - w_x)\mathbf{j} + (v_x - u_y)\mathbf{k}.$$

Si noti che, sviluppando formalmente secondo gli elementi della prima riga,

$$\operatorname{rot} \mathbf{V} = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u & v & w \end{bmatrix}.$$

Sia ora $u(x, y, z)$ una funzione due volte derivabile. Calcoliamo prima il gradiente di $u(x, y, z)$ e poi la divergenza del gradiente:

$$\nabla \cdot (\nabla u(x, y, z)) = u_{xx}(x, y, z) + u_{yy}(x, y, z) + u_{zz}(x, y, z).$$

Usa definire Δ , che si chiama laplaciano, ponendo

$$\Delta u(x, y, z) = u_{xx}(x, y, z) + u_{yy}(x, y, z) + u_{zz}(x, y, z).$$

Formalmente si scrive

$$\Delta = \nabla \cdot \nabla.$$

Si vede facilmente che queste considerazioni valgono per funzioni di qualunque numero di variabili.

Osservazione 155 (Sulle notazioni) Proseguiamo quanto si è detto all’osservazione 152. Facendo intervenire le notazioni dei campi vettoriali, i testi di meccanica del continuo e di scienza delle costruzioni scrivono la prima formula degli incrementi finiti in questo modo. Sia $\mathbf{u}(\mathbf{x})$ una funzione di tre variabili a valori in \mathbb{R}^3 . Usando la notazione “con la virgola” introdotta all’osservazione 152, la prima formula degli incrementi finiti è

$$\mathbf{u}(\mathbf{x}) = \mathbf{u}(\mathbf{0}) + \begin{bmatrix} u_{1,1}(\mathbf{0}) & u_{1,2}(\mathbf{0}) & u_{1,3}(\mathbf{0}) \\ u_{2,1}(\mathbf{0}) & u_{2,2}(\mathbf{0}) & u_{2,3}(\mathbf{0}) \\ u_{3,1}(\mathbf{0}) & u_{3,2}(\mathbf{0}) & u_{3,3}(\mathbf{0}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + o(\|\mathbf{x}\|).$$

Invece di scrivere in questo modo, viene scritto:

$$\mathbf{u}(\mathbf{x}) = \mathbf{u}(\mathbf{0}) + (\mathbf{x} \cdot \nabla) \mathbf{u}(\mathbf{0}) + o(\|\mathbf{x}\|)$$

ove $(\mathbf{x} \cdot \nabla)$ indica un “prodotto scalare formale” tra i due vettori colonna \mathbf{x} e ∇ ,

$$\mathbf{x} \cdot \nabla = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cdot \begin{bmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{bmatrix} = x_1 \partial_1 + x_2 \partial_2 + x_3 \partial_3.$$

Si trova quindi un operatore differenziale che si applica al campo vettoriale $\mathbf{u}(\mathbf{x})$, applicandolo a ciascuna delle sue componenti:

$$\begin{aligned} (\mathbf{x} \cdot \nabla) \mathbf{u}(\mathbf{x}) &= \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cdot \begin{bmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{bmatrix} \right) \mathbf{u}(\mathbf{x}) \\ &= (x_1 \partial_1 + x_2 \partial_2 + x_3 \partial_3) \mathbf{u}(\mathbf{x}) = \begin{bmatrix} x_1 u_{1,1} + x_2 u_{1,2} + x_3 u_{1,3} \\ x_1 u_{2,1} + x_2 u_{2,2} + x_3 u_{2,3} \\ x_1 u_{3,1} + x_2 u_{3,2} + x_3 u_{3,3} \end{bmatrix} \end{aligned}$$

proprio come si voleva ottenere. ■

4.7 Appendici

4.7.1 Appendice: Rappresentazione di funzioni di due variabile

In quest’appendice mostriamo alcune funzioni di due variabili a valori reali, ed il modo di rappresentarle. Sostanzialmente, ci sono tre modi. Descriviamoli prima di tutto in astratto e poi illustriamoli su opportuni esempi. Sia $f(x, y)$

una funzione. Il modo più ovvio di rappresentarla consiste nel costruirla il grafico, ossia nel costruire l'insieme

$$\{(x, y, z) \mid z = f(x, y)\} \subseteq \mathbb{R}^3.$$

Naturalmente questo richiede la costruzione di un modello per esempio di gesso o di plastica. In pratica rappresenteremo il grafico su un foglio, mediante opportune tecniche di disegno, che diano l'illusione della profondità. Inoltre, talvolta conviene rappresentare, invece della superficie, solamente una famiglia di linee sulla superficie. Il secondo metodo consiste nel tagliare il grafico a quota c , e quindi nel proiettare l'insieme intersezione sul piano (x, y) . Si disegna cioè *sul piano* (x, y) , l'insieme delle soluzioni dell'equazione

$$f(x, y) = c.$$

Quest'insieme si chiama *insieme di livello*. In molti casi³ un insieme di livello è una “curva” e quindi si parla di *curva di livello*. Disegnando “numerosi” curve di livello si può avere un'informazione sul comportamento della funzione: per esempio, le curve di livello sono più fitte dove il grafico della funzione è più ripido. Il terzo metodo consiste nel colorare il piano (x, y) colorandolo con colori “freddi”, per esempio blu, dove i valori della funzione sono più piccoli e con colori caldi, per esempio rosso, dove la funzione prende valori maggiori. Molto spesso conviene combinare questi metodi, attribuendo i colori alla superficie o alle sue curve di livello. Vediamo ora un esempio. Consideriamo la funzione seguente, il cui grafico è un paraboloide di rotazione:

$$f(x, y) = x^2 + y^2.$$

La figura 4.6 mostra vari modi di rappresentare il grafico di questa funzione. La figura 4.7 mostra a sinistra la funzione rappresentata mediante le sue curve di livello (e quindi mediante una rappresentazione sul piano (x, y)) ed a destra il grafico (fatto con fili) della medesima funzione, sovrapposto alle curve di livello. Ciò detto, passiamo ad esaminare alcuni esempi di funzioni.

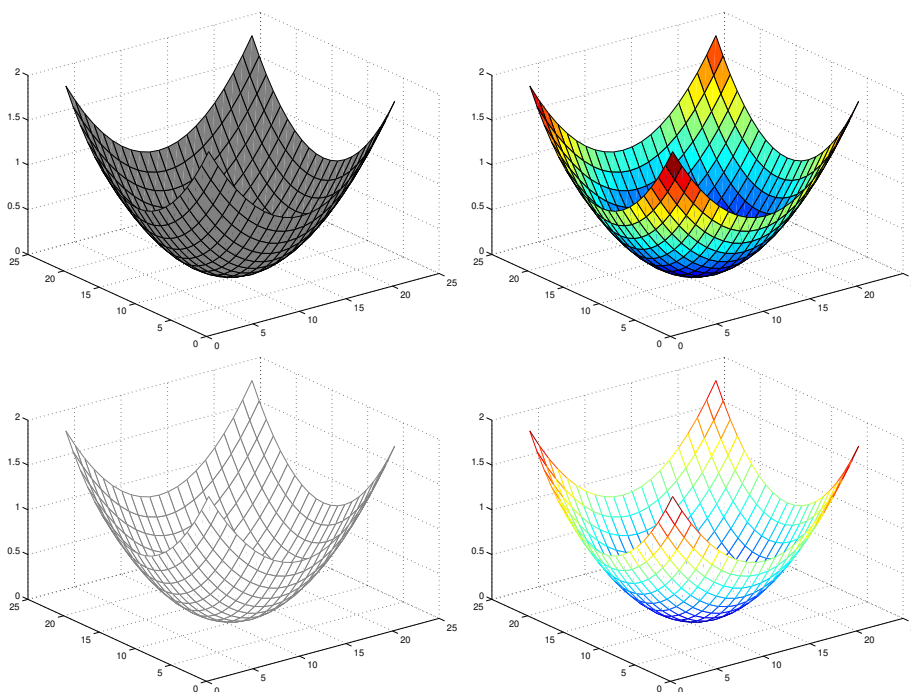
Funzioni costanti rispetto ad una variabile Può accadere che una funzione $f(x, y)$ sia “costante rispetto ad y ”; ossia che valga

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

per ogni scelta di y e di y' , purché la x prenda il medesimo valore a destra ed a sinistra. Vediamo le conseguenze “grafiche” di questa proprietà. La funzione è

³come proveremo al Capitolo 5

Figura 4.6: Paraboloide di rotazione

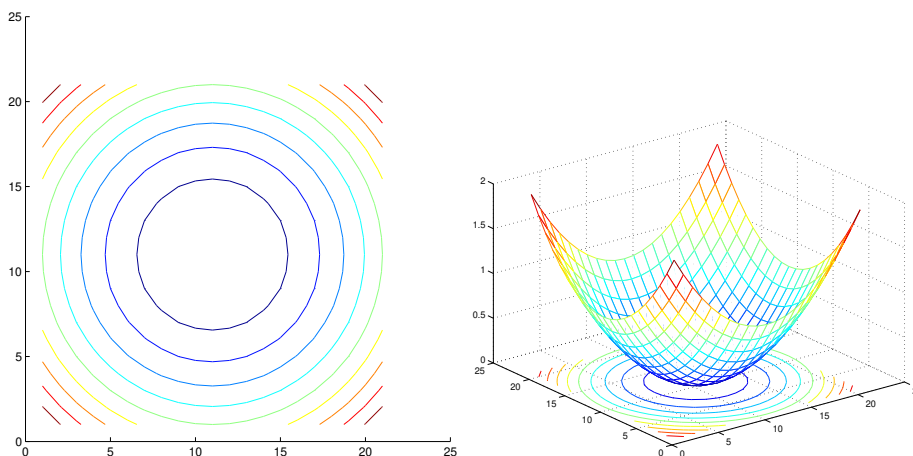


costante sulle rette verticali del piano (x, y) , che pertanto sono tutte contenute in insiemi di livello. Per esempio la funzione

$$f(x, y) = \sin x \quad (4.10)$$

ha gli insiemi di livello rappresentati nella figura 4.8 a sinistra. Rette del medesimo colore corrispondono al medesimo livello e quindi gli insiemi di livello in quest'esempio non sono "curve", ma sono insiemi di (infinite) rette verticali. Dal punto di vista del grafico, il grafico della funzione è unione di rette orizzontali, e quindi tutte parallele tra loro. Il termine *superficie* verrà introdotto più avanti, ma è un fatto che sono "superfici" i grafici di funzioni (differenziabili). Una superficie che è unione di rette tutte tra loro parallele si chiama *cilindro*. Dunque, i grafici di funzioni $f(x, y)$ costanti rispetto ad una variabile sono cilindri. Il cilindro corrispondente alla funzione in esame è nella figura (4.8), a destra.

Figura 4.7: Paraboloide di rotazione e sue curve di livello



Funzioni omogenee Si chiamano *funzioni omogenee* di grado ν (non necessariamente intero) quelle funzioni che verificano l'uguaglianza

$$f(rx, ry) = |r|^\nu f(x, y).$$

Se $\nu = 0$ la funzione è costante sulle rette che escono dall'origine, che pertanto appartengono ad insiemi di livello. Un esempio è la funzione

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}.$$

Dato che il denominatore si annulla in $(0, 0)$, per evitare difficoltà di tipo numerico, consideriamola nel quadrato $0, 1 \leq x \leq 1, 0, 1 \leq y \leq 1$. La figura 4.9, a sinistra, presenta le curve di livello della funzione. Come si è detto, queste sono rette, ma problemi di tipo numerico vicino a $(0, 0)$ (dove si annullano sia il numeratore che il denominatore) provocano le distorsioni visibili in figura. Il valore della funzione varia da retta a retta, ma rimane costante sulle singole rette, come mostra la figura 4.9, a destra. In questa figura abbiamo rappresentato, in verde, anche il piano $z = 0$. Se la funzione è omogenea di grado 1 allora moltiplicando per r sia x che y , il valore di z viene anch'esso moltiplicato per r . Dunque, se un punto (x_0, y_0, z_0) è sul grafico, anche i punti (rx_0, ry_0, rz_0) con $r > 0$ sono sul grafico. Al variare di r , questi descrivono semirette passanti per l'origine: il grafico è unione di semirette passanti per l'origine. Un insieme unione di rette per un medesimo punto si

Figura 4.8: La funzione (4.10)

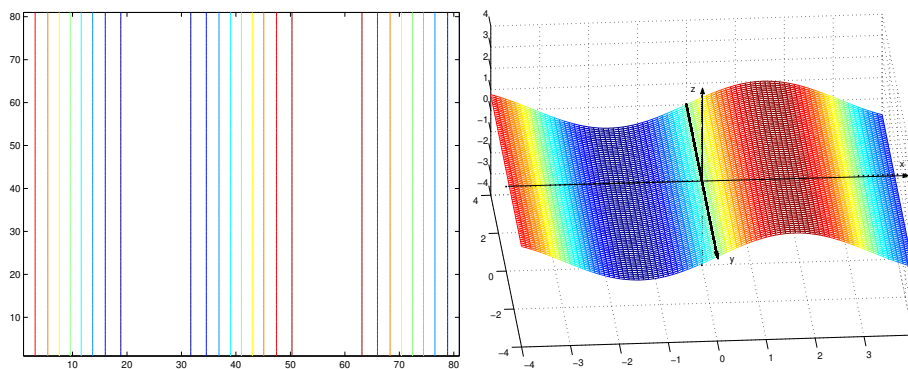
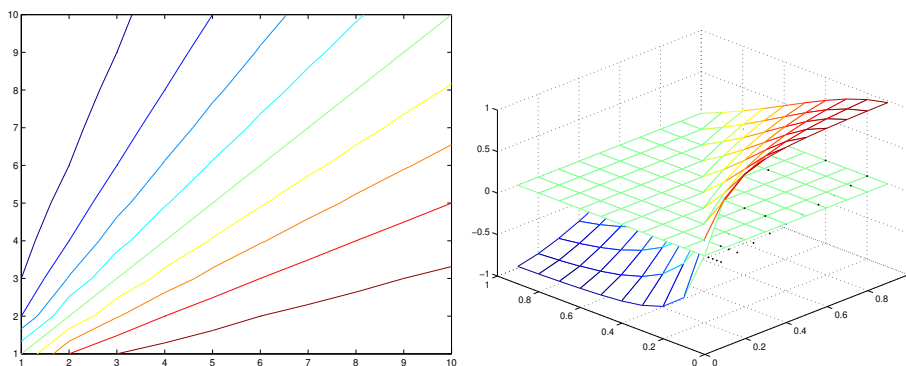


Figura 4.9: Una funzione omogenea



chiama cono. E quindi il grafico di una funzione omogenea di grado 1 è un semicono (con vertice nell'origine). La figura 4.10, a sinistra, illustra la situazione nel caso della funzione

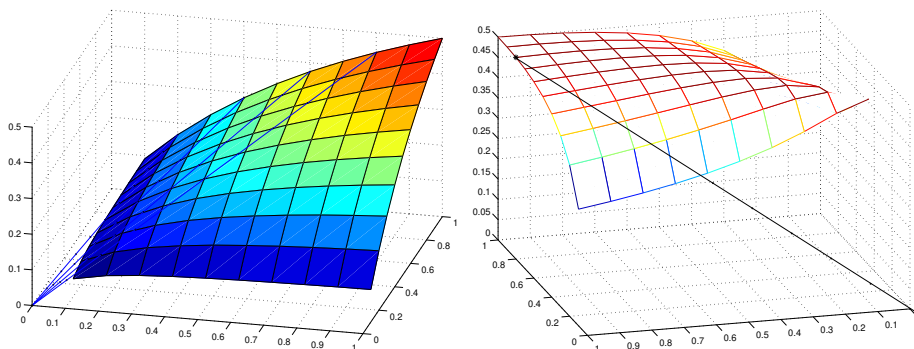
$$f(x, y) = \frac{xy}{x + y}.$$

La figura riporta sia il grafico della funzione sia alcune rette che escono dall'origine e che appartengono al grafico. La figura 4.10, a destra, riporta invece il grafico della funzione

$$f(x, y) = \frac{\sqrt{xy}}{x + y},$$

omogenea di grado $1/2$. La figura riporta anche una retta per l'origine e per un punto del grafico che, come si vede, non giace sul grafico.

Figura 4.10: Funzione omogenea di grado $1/2$



Simmetrie del grafico, funzioni radiali e superfici di rotazione Il grafico di una funzione di più variabili può avere proprietà di simmetria, che estendono quelle note per le funzioni di una variabile. Per esempio, diremo che una funzione $f(x, y)$ è

- *dispari* se $f(x, y) = -f(-x, -y)$;
- *pari* se $f(x, y) = f(-x, -y)$.

Come nel caso delle funzioni di una variabile, queste proprietà corrispondono a certe simmetrie del grafico rispetto all'origine. Naturalmente, si possono avere altre possibilità che non si incontrano nel caso di funzioni di una variabile (per esempio, simmetrie rispetto ad una retta per l'origine, magari ad uno degli assi coordinati). A noi ora interessa un caso molto particolare di funzione pari di due variabili, ottenuta a partire da funzioni di una sola variabile. Si consideri la funzione $t \mapsto f(t)$ di **una sola variabile** definita su un intervallo⁴ $[a, b]$ e a questa si associ la funzione di **due variabili**

$$(x, y) \mapsto f\left(\sqrt{x^2 + y^2}\right)$$

⁴facciamo il caso di un intervallo chiuso, ma quello che andiamo a dire si adatta in ogni caso, anche nel caso che si tratti di una semiretta.

ossia la funzione

$$z = f\left(\sqrt{x^2 + y^2}\right), \quad \text{dom } f = \left\{(x, y), \sqrt{x^2 + y^2} \in [a, b]\right\}.$$

I grafici di queste funzioni hanno una proprietà particolare. Sia

$$(x_0, y_0) \in \text{dom } f\left(\sqrt{x^2 + y^2}\right) \quad \text{ed} \quad r_0 = \sqrt{x_0^2 + y_0^2}.$$

Il numero r_0 è la distanza di (x_0, y_0) dall'origine. Qualunque altro punto (x, y) che dista r_0 dall'origine, ossia per cui

$$\sqrt{x^2 + y^2} = r_0$$

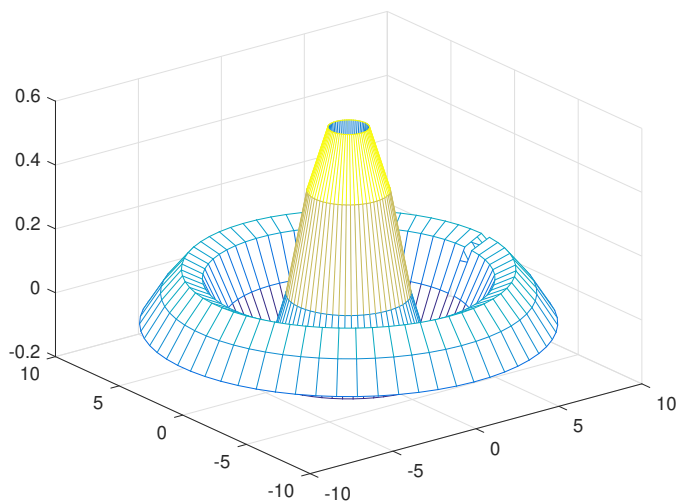
appartiene al dominio della funzione ed inoltre,

$$f\left(\sqrt{x^2 + y^2}\right) = f\left(\sqrt{x_0^2 + y_0^2}\right).$$

Dunque il dominio della funzione è una corona circolare di centro l'origine (un disco se $a = 0$, l'esterno di un disco se $a > 0$ ed $r = +\infty$), e tutte le circonferenze di centro l'origine sono curve di livello della funzione. Questo si interpreta in questo modo: **se il dominio della funzione contiene un punto P , esso contiene tutta la circonferenza di centro O che passa per P . Inoltre su tale circonferenza la funzione è costante. Ossia se il punto P ruota su una circonferenza di centro O il corrispondente punto del grafico ruota su una circonferenza parallela al piano x, y , centrata sull'asse delle quote. Detto in altro modo, ogni piano perpendicolare all'asse delle quote se interseca il grafico della funzione lo interseca lungo una circonferenza (oppure in un solo punto).** Una superficie come quella ora descritta si chiama superficie di rotazione. L'intersezione tra un piano $z = z_0$ e il grafico, se non è vuota, si chiama un parallelo della superficie mentre l'intersezione tra il grafico ed ogni piano passante per l'asse delle quote si chiama meridiano della superficie di rotazione: **la superficie di rotazione può anche pensarsi ottenuta ruotando un meridiano intorno all'asse delle quote.** Il grafico nella figura 4.11 mostra una superficie di rotazione, con indicati alcuni paralleli ed alcuni meridiani.

Osservazione 156 • Una funzione $g(x, y) = f\left(\sqrt{x^2 + y^2}\right)$ che definisce una superficie di rotazione si chiama anche una funzione radiale. Più in generale, si chiama funzione radiale ogni funzione su \mathbb{R}^n il cui valore in ogni \mathbf{x} del suo dominio dipende solo da $|\mathbf{x}|$. Tale funzione si esprime come $f(|x|)$ con f funzione di una variabile reale.

Figura 4.11: Superficie di rotazione



- La funzione $g(x^2 + y^2)$ è una funzione radiale, anche se nella sua espressione la radice non compare. Infatti,

$$g(x^2 + y^2) = g\left(\left(\sqrt{x^2 + y^2}\right)^2\right). \quad \blacksquare$$

Funzioni periodiche Una funzione di due variabili, $f(x, y)$, si dice periodica quando esiste un vettore $\mathbf{v}_0 = (\xi_0, \eta_0)$ tale che

$$f(\mathbf{x}) = f(x + \xi_0, y + \eta_0) = f(x, y)$$

per ogni (x, y) del suo dominio. Un esempio è la funzione

$$f(x, y) = (\sin x)(\sin 2 * y).$$

Se

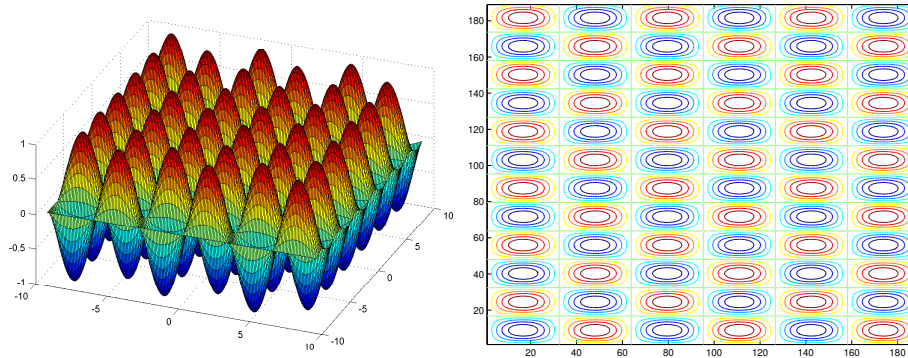
$$\mathbf{v} = 2\pi\mathbf{i} + \pi\mathbf{j}$$

si ha

$$f(\mathbf{x} + \mathbf{v}) = f(x + 2\pi, y + \pi) = [\sin(x + 2\pi)] [\sin(2(y + \pi))] = f(x, y).$$

La figura 4.12 mette a confronto il grafico della funzione $f(x, y)$, a sinistra, con le sue curve di livello, a destra.

Figura 4.12: Una funzione periodica



4.7.2 Appendice: Propagazione ondosa

In molte applicazioni si incontrano funzioni della forma

$$f(\phi(\mathbf{x}) - vt) \quad \text{oppure} \quad f(\phi(\mathbf{x}) + vt)$$

con $v > 0$. In queste applicazioni la variabile t si interpreta come “tempo” mentre la variabile \mathbf{x} indica una posizione nello spazio. Dunque la funzione $f(\phi(\mathbf{x}) \pm vt)$ descrive una configurazione $f(\phi(\mathbf{x}))$ che “si muove” al passare del tempo e ciò suggerisce un linguaggio particolare, che illustriamo con riferimento al caso $\mathbf{x} = x \in \mathbb{R}$ e $\phi(x) = x$ oppure $\phi(x) = mx$. In queste applicazioni, la funzione $f(x)$ si chiama onda. Potremmo pensare al suo grafico come alla configurazione all’istante $t = 0$ di un corpo elastico. Vediamo come si interpretano le due funzioni

$$F_1(x, t) = f(x - vt), \quad F_2(x, t) = f(x + vt)$$

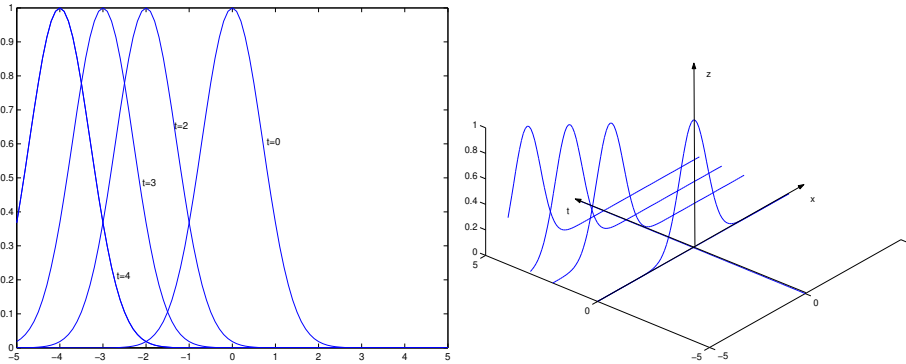
con $v > 0$. Consideriamo la funzione $F_1(x, t)$. Fissiamo un valore t_0 . La funzione $x \rightarrow f(x - vt_0)$ si può interpretare in due modi diversi: come la funzione $f(x)$ rappresentata rispetto ad un sistema di riferimento ottenuto traslando l’origine nella posizione $-vt_0$ dell’asse x . Dunque **in dietro**, perché $v > 0$. Questo modo di intendere le cose è il più comune nelle applicazioni geometriche. Invece, nelle applicazioni fisiche e nell’analisi si incontra più frequentemente un’interpretazione diversa: si pensa ad $f(x)$ ed $f(x - vt_0)$ come a due **diverse** funzioni rappresentate rispetto al **medesimo** sistema di riferimento. In tal caso, il grafico della funzione $f(x - vt_0)$ è ottenuto traslando **in avanti** il grafico di $f(x)$. Se t si interpreta come “tempo”, nell’unità di

tempo il grafico è andato avanti di una quantità v . Dunque, v rappresenta la velocità dello spostamento del grafico. Pensando al grafico di $f(x)$ come ad un'“onda”, si dice che v è la *velocità di propagazione dell'onda*. Le stesse considerazioni si possono ripetere per la funzione $F_2(x, t)$ con la sola differenza che ora il grafico si **sposta in dietro** (grazie al fatto che v si è scelto positivo). Le considerazioni precedenti suggeriscono di chiamare la funzione $F_1(x, t) = f(x - vt)$ *onda progressiva* mentre la funzione $F_1(x, t) = f(x + vt)$ si chiama *onda regressiva* (ricordiamo che $v > 0$). La figura 4.13 riporta i grafici delle funzioni

$$e^{-(x+t)^2} . \tag{4.11}$$

a sinistra sono riportati i grafici delle funzioni $x \rightarrow e^{-(x+t)^2}$ nel medesimo sistema di assi cartesiani (x, y) , per diversi valori di t . A destra è stato costruito il grafico della funzione di due variabili $(x, t) \rightarrow e^{-(x+t)^2}$, nello spazio riferito agli assi cartesiani (x, t, z) . Questo grafico è stato sezionato con i piani $t = 0$, $t = 2$, $t = 3$ e $t = 4$, e sono stati disegnati i grafici risultanti da queste sezioni.

Figura 4.13: La funzione (4.11)

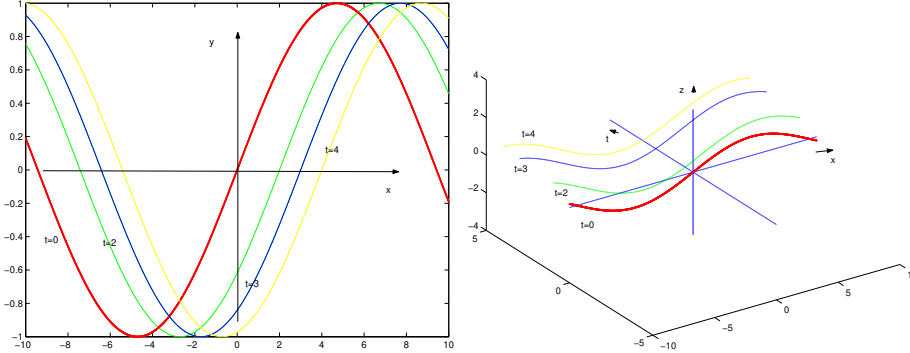


illustrata sopra è particolarmente intuitivo nel caso di funzioni $f(x)$ periodiche. La figura 4.14 riporta grafici analoghi a quelli della figura 4.13, ma questa volta con la funzione

$$f(x) = \sin(mx) . \tag{4.12}$$

Si noti che questa funzione è periodica di periodo $2\pi/m$. Il valore scelto nella figura per il periodo è 3. Il caso in cui $f(x) = \sin(mx)$ oppure $f(x) = \cos(mx)$ è particolarmente importante nelle applicazioni. Le “onde” costruite a partire da queste funzioni si dicono *onde armoniche* e in questo caso si usano notazioni

Figura 4.14: La funzione (4.12)



e definizioni peculiari che verranno illustrate al paragrafo 4.7.2. Consideriamo ora la funzione $F_1(x, t) = f(x - vt)$. Questa funzione prende il valore $f(k)$ in tutti i punti (x, t) nei quali

$$x - vt = k.$$

Per ogni k si viene così a descrivere una retta nei punti della quale l'onda ha il medesimo valore. Ciò suggerisce di dare un nome alla quantità $x - vt$: questa si chiama la fase dell'onda progressiva. Al trascorrere del tempo, si rimane sulla stessa fase k se si passa dal punto x al punto $x + vt$, ossia se si sposta il punto osservato x in avanti con velocità v . Per questa ragione, il numero v si chiama anche (e più frequentemente) velocità di fase. Considerazioni analoghe per le onde regressive si lasciano per esercizio. Si noti che talvolta un'onda viene scritta come

$$f(hx - mt) = f(h(x - (m/h)t)).$$

In tal caso, la velocità di fase è $v = m/h$. Se $\mathbf{x} \in \mathbb{R}^2$ oppure $\mathbf{x} \in \mathbb{R}^3$ si incontra un nuovo concetto, che illustriamo con riferimento a $\mathbf{x} \in \mathbb{R}^2$, $\mathbf{x} = [x \ y]'$. Consideriamo le due onde progressive seguenti:

$$\Phi_1(\mathbf{x}, t) = \Phi_1(x, y, t) = f(ax + by - vt), \quad \Phi_2(\mathbf{x}, t) = \Phi_2(x, y, t) = f(x^2 + y^2 - vt).$$

Ci si può chiedere su quali insiemi del piano (x, y) l'onda ha la medesima fase per ogni fissato valore di t . Lavorando con la funzione $\Phi(\mathbf{x}, t)$, fissato $t = t_0$, la fase vale k nei punti della retta

$$ax + by = k + vt_0;$$

nel caso della funzione $\Phi_2(\mathbf{x}, t)$ fissato $t = t_0$, la fase vale k nei punti della circonferenza

$$x^2 + y^2 = k + vt_0.$$

Per questa ragione, l'onda $\Phi_1(\mathbf{x}, t)$ si chiama onda piana mentre l'onda $\Phi_2(\mathbf{x}, t)$ si chiama onda sferica.

Onde e moto armonico

Ricordiamo che si chiama moto armonico il moto di un punto lungo una circonferenza, se questo avviene con velocità angolare costante. I punti di una circonferenza di raggio A hanno coordinate

$$(A \cos(\omega t + \phi), A \sin(\omega t + \phi)). \quad (4.13)$$

Il numero $(2\pi)/\omega$ è il tempo necessario a descrivere una volta tutta la circonferenza e quindi

$$T = \frac{2\pi}{\omega}$$

è il periodo del moto mentre il suo reciproco

$$f = \frac{\omega}{2\pi}$$

si chiama frequenza del moto. La frequenza indica quanti giri (o parti di giro) percorre il corpo lungo la circonferenza in un'unità di tempo. Il numero di radianti di cui varia la posizione del corpo nell'unità di tempo è ω e per questo il coefficiente ω si chiama velocità angolare. Dunque,

$$\omega = 2\pi f.$$

Dato che $\omega = (2\pi)/T$, la (4.13) si può anche scrivere (si confronti con l'osservazione 82 dove però il periodo è $2T$)

$$\left(A \cos \frac{2\pi}{T} t, A \sin \frac{2\pi}{T} t \right).$$

Questo legame col moto armonico suggerisce di chiamare onde armoniche le onde

$$A \cos(kx - \omega t + \phi), \quad A \sin(kx - \omega t + \phi)$$

(e le analoghe onde regressive, il cui esame si lascia per esercizio). Consideriamo per fissare le idee l'onda progressiva

$$A \sin(kx - \omega t + \phi).$$

Il numero A si chiama l' ampiezza dell'onda e il numero $kx - \omega t + \phi$ si chiama la fase. Le due onde

$$A \sin(kx - \omega t + \phi), \quad A \sin(kx - \omega t + \psi)$$

(con gli stessi valori di ω e di k) si dicono sfasate di $\phi - \psi$ e questo numero si chiama sfasamento tra le due onde⁵. Per semplicità di scrittura, poniamo ora $\phi = 0$. L'onda che si trova all'istante $t = 0$ è l'onda

$$f(x) = \sin kx,$$

periodica di periodo $2\pi/k$. Però, i fisici riservano i termini “periodo” e “frequenza” alla variabile tempo. Il numero

$$\lambda = \frac{2\pi}{k}$$

si chiama invece lunghezza d'onda. Infatti, l'onda prende i medesimi valori nei punti che distano l'uno dall'altro di $2\pi/k$. Questo può visualizzarsi in questo modo: consideriamo la funzione

$$f_e(x) = \sin kx, \quad x \in [0, 2\pi/k) = [0, \lambda).$$

chiamiamola⁶ “onda elementare”. Allora il grafico di $f(x)$ si può pensare ottenuto giustapponendo i grafici di $f_e(x + n(2\pi/k))$ per ogni numero intero n . La figura 4.15 illustra ciò, usando colori diversi per vedere i grafici che si sono giustapposti. Il numero di onde elementari che si ripetono nell'unità di spazio è 1 diviso per la lunghezza dell'onda elementare, che è $k/2\pi$. Il valore

$$\frac{1}{\lambda} = \frac{k}{2\pi}$$

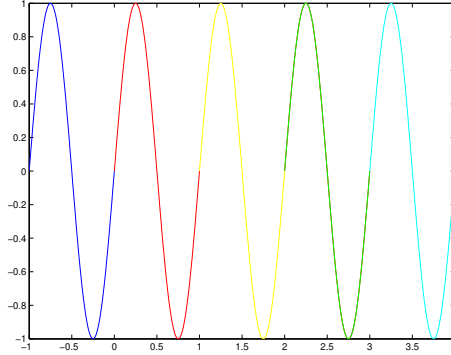
si chiama numero d'onde mentre k si chiama numero d'onde angolare. Come si è detto, i termini “periodo”, “frequenza” e “frequenza angolare” si riservano alla dipendenza dal tempo t dell'onda. Ossia, si riservano alla funzione

$$t \rightarrow A \sin(kx - \omega t + \phi)$$

⁵più comunemente, per sfasamento si prende un numero dell'intervallo $[0, 2\pi)$. Ciò può farsi perché le funzioni $\cos x$ e $\sin x$ sono periodiche di periodo 2π .

⁶Questo non è un termine usuale!

Figura 4.15: Lunghezza d'onda



con x fissato. La terminologia è quella già illustrata nel contesto del moto armonico. Indicando con T il periodo, si ha

$$T = \frac{2\pi}{\omega}, \quad f = \frac{1}{T} = \frac{\omega}{2\pi}$$

Il numero ω , che è la velocità angolare del moto armonico, nello studio delle onde armoniche si chiama frequenza angolare. Studiamo infine la velocità di fase, scrivendo

$$\sin(\omega t + kx) = \sin k \left(\frac{\omega}{k} t + x \right).$$

La velocità di fase, ossia la velocità di spostamento dell'onda, è

$$v = \frac{\omega}{k} = \frac{\lambda}{T}.$$

Onde stazionarie

Le onde si “propagano” nello spazio, al trascorrere del tempo; questa almeno sembra l'informazione che può ricavarsi dalle considerazioni precedenti. Però, in uno stesso mezzo possono propagarsi onde diverse. Consideriamo due onde armoniche con la medesima ampiezza

$$f(x - vt) = A \cos k(x - vt), \quad g(x - vt) = A \cos m(x - vt).$$

Se le due onde hanno il medesimo numero d'onde, $k = m$ e la medesima velocità, $v = \nu$, allora nel mezzo si vede un'onda di ampiezza doppia,

$$2A \cos(kx - vt).$$

Altrimenti, si trova la funzione

$$(x, t) \longrightarrow A [\cos k(x - vt) + \cos m(x - vt)] .$$

Usando le formule di prostaferesi si trova la funzione

$$\begin{aligned} & 2A \cos \left(\frac{(k+m)x - (kv + m\nu)t}{2} \right) \cos \left(\frac{(k-m)x + (m\nu - kv)t}{2} \right) \\ &= 2A \cos \left(\frac{k+m}{2} \left[x - \frac{kv + m\nu}{k+m} t \right] \right) \cos \left(\frac{k-m}{2} \left[x + \frac{m\nu - kv}{k-m} t \right] \right) \end{aligned}$$

Supponiamo ora che la velocità di propagazione sia uguale per le due onde, $v = \nu$. In questo caso si trova il segnale

$$\left[2A \cos \frac{k-m}{2}(x - vt) \right] \cos \frac{k+m}{2}(x - vt)$$

Interpretiamo questa formula: quando i due segnali hanno numero d'onde molto simile, $k \sim m$ e “molto grande”, allora $\epsilon = k - m$ è un numero piccolo relativamente sia a k che ad m . Tenendo fissa la posizione x , al variare del tempo in questa posizione si vede passare l'onda $\cos \frac{k+m}{2}(x - vt) \sim \cos k(x - vt)$, modulata dal segnale di bassa frequenza $2A \cos \frac{k-m}{2}(x - vt)$. Se l'onda in questione è un'onda sonora, nella posizione x si sente un suono di frequenza circa uguale a $kv/2\pi$, ed ampiezza $2A \cos[(k - m)(x - vt)/2]$ lentamente variabile nel tempo, con legge periodica di “piccola” frequenza $(k-m)v/(4\pi)$. Si ricordi che la minima frequenza dei suoni udibili dall'orecchio umano è 16 Herz. Se accade che

$$\frac{(k-m)v}{4\pi} > 16$$

l'orecchio sente comparire un “terzo suono”, diverso dai due suoni emessi dagli strumenti, ed in generale sgradevole. Questo suono si chiama appunto *terzo suono* o *suono di Tartini*. Il caso estremo si ha quando le due onde hanno il medesimo numero d'onde e la medesima velocità, ma si propagano in verso opposto, $v = -\nu$. In tal caso si trova

$$2A \cos kx \cos kvt . \tag{4.14}$$

In questo caso niente si propaga: la somma di onde uguali propagantesi in versi opposti genera una configurazione $A \cos kx$, stazionaria, “modulata” dal segnale sovrapposto $\cos kvt$. Nonostante che in questo caso niente si propaghi, per conservare la memoria dell'origine di questo fenomeno, si chiama ancora “onda” la funzione (4.14) e, per sottolineare che “niente si propaga”, la chiamiamo *onda stazionaria*.

Interferenza

Consideriamo ora due onde uguali, ossia con la medesima ampiezza, numero d'onde e velocità di propagazione, ma sfasate l'una rispetto all'altra. Ossia consideriamo le due onde

$$f(x - vt) = A \cos k(x - vt), \quad g(x - vt) = A \cos k(x - vt + \phi).$$

Se queste due onde si propagano nel medesimo mezzo, il segnale propagantesi nel mezzo è

$$A [\cos k(x - vt) + \cos k(x - vt + \phi)] = \left[2A \cos \frac{k\phi}{2} \right] \cos \left(kx - vt + \frac{\phi}{2} \right).$$

Si ha quindi un segnale di ampiezza $2A \cos k\phi/2$. che può essere nulla. In pratica, la differenza di fase è dovuta alla differente lunghezza del cammino percorso dalle due onde; e quindi in realtà ϕ varia cambiando posizione, $\phi = \phi(x)$. Se l'onda in questione è un'onda sonora, in un punto x fissato si sente un suono uguale a quello descritto da $f(x)$, ma con una diversa intensità. L'intensità varia al variare della posizione d'ascolto e in certe posizioni può annullarsi. Questo fenomeno si chiama interferenza delle due onde.

4.7.3 Appendice: Funzioni omogenee

Una funzione $f(\mathbf{r})$, con $\mathbf{r} \in \mathbb{R}^n$, si chiama *positivamente* omogenea di grado ν (anche non intero) se

$$f(t\mathbf{r}) = |t|^\nu f(\mathbf{r}).$$

Supponiamo che una funzione omogenea di grado ν sia di classe C^1 per $\mathbf{r} \neq 0$. Derivando rispetto a t i due membri e calcolando le derivate per $t = 1$, si trova l'uguaglianza

$$\nabla f(\mathbf{r}) \cdot \mathbf{r} = \nu f(\mathbf{r}) \quad \forall \mathbf{r}. \quad (4.15)$$

Questa relazione si chiama *identità di* Eulero e caratterizza le funzioni omogenee e di classe C^1 (per $\mathbf{r} \neq 0$) di grado ν . Infatti, consideriamo il caso $t > 0$. Derivando la funzione

$$t^{-\nu} f(t\mathbf{r})$$

si trova

$$\begin{aligned} \frac{d}{dt} t^{-\nu} f(t\mathbf{r}) &= -\nu t^{-\nu-1} f(t\mathbf{r}) + t^{-\nu} (\nabla f(t\mathbf{r}) \cdot \mathbf{r}) \\ &= t^{-\nu-1} [-\nu f(t\mathbf{r}) + \nabla f(t\mathbf{r}) \cdot (t\mathbf{r})]. \end{aligned}$$

Per ipotesi la (4.15) vale per ogni \mathbf{r} ; in particolare vale nel punto $t\mathbf{r}$, ossia vale

$$\nu f(t, \mathbf{r}) = \nabla f(t, \mathbf{r}) \cdot (t\mathbf{r});$$

e quindi

$$\frac{d}{dt} t^{-\nu} f(t\mathbf{r}) = 0.$$

Dunque,

$$t^{-\nu} f(t\mathbf{r}) = c, \quad \text{ossia} \quad f(t\mathbf{r}) = ct^\nu$$

ove c è una opportuna costante. Ponendo $t = 1$ si trova $c = f(\mathbf{r})$ e quindi

$$f(t\mathbf{r}) = t^\nu f(\mathbf{r}).$$

In modo analogo si procede per $t < 0$ e quindi si trova che la funzione è omogenea di grado ν .

4.7.4 Appendice: La dimostrazione del teorema 132

Dim. Per semplicità di notazioni, proviamo l'asserto con $x_0 = 0, y_0 = 0$. Fissiamo i punti x ed y e consideriamo la funzione di una sola variabile

$$\phi(t) = f(tx, ty).$$

Si noti che $\phi(1) = f(x, y)$ mentre $\phi(0) = f(0, 0)$. Scriviamo

$$\phi(t) = \{f(tx, ty) - f(tx, 0)\} + \{f(tx, 0) - f(0, 0)\}.$$

Consideriamo la funzione della sola variabile t

$$m(t) = f(tx, 0).$$

In quest'espressione, il numero x si considera come un parametro il cui valore è fissato. Applichiamo ad essa la **seconda** formula degli incrementi finiti: esiste $c \in (0, 1)$ tale che

$$f(x, 0) - f(0, 0) = m(1) - m(0) = \frac{d}{dt} m(t)|_{t=c} = x f_x(cx, 0).$$

Si noti che si sono usate in questo calcolo solamente proprietà delle funzioni di una variabile e che, per applicare la **seconda** formula degli incrementi finiti, si è usata l'esistenza della derivata di $f(tx, 0)$ che, a sua volta, implica la continuità di $f(tx, 0)$. Il numero c dipende dalla particolare funzione $m(t)$ usata e quindi,

4.7. APPENDICI

in definitiva, dipende da x . Dato che $c \in (0, 1)$, si ha $|cx| < |x|$. Usando ciò, stimiamo

$$\frac{|x [f_x(cx, 0) - f_x(0, 0)]|}{\sqrt{x^2 + y^2}} \leq |f_x(cx, 0) - f_x(0, 0)| .$$

La condizione $|cx| < |x|$ e la **continuità della derivata prima** mostrano che il limite per $(x^2 + y^2)$ tendente a zero è nullo. Dunque vale

$$f(x, 0) - f(0, 0) = xf_x(0, 0) + o(x, y) . \quad (4.16)$$

Consideriamo ora la funzione $n(\tau)$, dipendente dai parametri x ed y ,

$$n(\tau) = f(x, \tau y) .$$

Applicando a questa la **seconda** formula degli incrementi finiti, si trova $d \in (0, 1)$ tale che

$$f(x, \tau y) - f(x, 0) = n(1) - n(0) = \frac{d}{d\tau} n(\tau)|_{\tau=\varepsilon} = yf_y(x, dy) .$$

Si noti che il numero d dipende dalla particolare funzione con cui si lavora, e quindi dai parametri x ed y che la identificano. E' però un numero compreso tra 0 ed 1. La precedente uguaglianza vale **per ogni** $\tau \in (0, 1)$, in particolare essa vale se come valore di τ si sceglie 1:

$$f(x, y) - f(x, 0) = yf_y(x, dy) = yf_y(0, 0) + y [f_y(x, dy) - f_y(0, 0)] .$$

L'ultima differenza verifica

$$\frac{|y [f_y(x, dy) - f_y(0, 0)]|}{\sqrt{x^2 + y^2}} \leq |f_y(x, dy) - f_y(0, 0)|$$

e quindi tende a zero per $\text{dist}((x, y), (0, 0))$ tendente a zero, **per la continuità di f_y** e perchè $0 < d < 1$. Dunque si può scrivere

$$f(x, y) - f(x, 0) = yf_y(0, 0) + o(x, y) . \quad (4.17)$$

Sommando la (4.16) e la (4.17) si trova la formula cercata. ■

Osservazione 157 Si noti che per la dimostrazione è sufficiente sapere che la funzione $f(x, y)$ è derivabile ed ha derivate continue in un intorno di (x_0, y_0) . ■

4.7.5 Appendice: la dimostrazione del teorema di Schwarz

La dimostrazione del teorema di Schwarz è importante anche perché fornisce un'ulteriore rappresentazione delle derivate miste. Limitiamoci a provare il teorema assumendo $(x_0, y_0) = (0, 0)$. Mostriamo che, se $f_{x,y}(x, y)$ ed $f_{y,x}(x, y)$ sono continue in un intorno di $(0, 0)$ allora si ha

$$f_{xy}(0, 0) = \lim_{(h,k) \rightarrow (0,0)} \frac{f(h, k) - f(h, 0) - f(0, k) + f(0, 0)}{hk} = f_{yx}(0, 0). \quad (4.18)$$

L'uguaglianza implica in particolare che le derivate miste sono uguali. Introduciamo la funzione $\phi(x) = f(x, k) - f(x, 0)$. Vale:

$$\phi(h) - \phi(0) = f(h, k) - f(h, 0) - f(0, k) + f(0, 0).$$

Dal *Teorema di Lagrange* si ha

$$\phi(h) - \phi(0) = h\phi'(c) = h[f_x(c, k) - f_x(c, 0)].$$

Il numero c dipende sia da h che da k , $c = c(h, k)$ e verifica

$$0 < c(h, k) < k.$$

Si usi ora la continuità della derivata $f_x(x, y)$, l'esistenza in ogni punto della derivata $f_{xy}(x, y)$ e di nuovo il *Teorema di Lagrange* per scrivere

$$\phi(h) - \phi(0) = h[f_x(c, k) - f_x(c, 0)] = hk[f_{xy}(c(h, k), d)].$$

Il punto d dipende da k e da c , quindi in definitiva da h e da k , $d = d(h, k)$, e verifica

$$0 < d(h, k) < k.$$

Dunque, dal *Teorema di confronto* per i limiti,

$$\lim_{(h,k) \rightarrow (0,0)} c(h, k) = 0, \quad \lim_{(h,k) \rightarrow (0,0)} d(h, k) = 0$$

così che, per la continuità della derivata seconda mista

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\phi(h) - \phi(0)}{hk} = \lim_{(h,k) \rightarrow (0,0)} f_{xy}(c(h, k), d(h, k)) = f_{xy}(0, 0).$$

Ciò prova la prima uguaglianza in (4.18). Per provare la seconda uguaglianza si introduca la funzione

$$\psi(y) = f(h, y) - f(0, y).$$

Si procede in modo analogo a quanto fatto per la funzione $\phi(x)$ e si trova:

- $\psi(k) - \psi(0) = f(h, k) - f(0, k) - f(h, 0) + f(0, 0)$
 $= f(h, k) - f(h, 0) - f(0, k) + f(0, 0) = \phi(h) - \phi(0);$
- $\lim_{(h,k) \rightarrow (0,0)} \frac{\psi(k) - \psi(0)}{hk} = f_{yx}(0, 0) = \lim_{(h,k) \rightarrow (0,0)} \frac{\phi(h) - \phi(0)}{hk} = f_{x,y}(0, 0)$

e quindi $f_{xy}(0, 0) = f_{yx}(0, 0)$. ■

Osservazione 158 Si noto che:

- per la dimostrazione del teorema di Schwartz basta che la funzione sia di classe $C^1(V)$, con V intorno di (x_0, y_0) ; che esistono le derivate seconde miste in ogni punto di V e che queste sono continue in (x_0, y_0) . Le derivate $f_{xx}(x, y)$, $f_{yy}(x, y)$ non intervengono.
- Supponiamo che la funzione dipenda da tre o più variabili e che x ed y denotino due delle variabili. Nel fare le derivate rispetto ad x ed y , tutte le altre variabili vengono tenute costanti; e quindi il teorema di Schwarz vale per funzioni di classe C^1 , qualunque sia il numero di variabili da cui la funzione dipende; e per tutte le coppie (x, y) di variabili tali che le rispettive derivate seconde miste siano continue.

