Corso di Laurea in Fisica

Complementi di Analisi Matematica di Base (II modulo)

Giulio Schimperna

Dipartimento di Matematica, Università di Pavia

Via Ferrata 1, 27100 – PAVIA

E-mail: `giulio@dimat.unipv.it`

Homepage: `http://www-dimat.unipv.it/~giulio/camb03.html`

Versione del **15 Dicembre 2003**

Nota per il lettore. La presente dispensa non si propone di coprire continuativamente il programma svolto durante il modulo, ma solo di completare alcuni argomenti che sono trattati in modo solo parziale sul libro di testo (G. Gilardi, "Analisi Matematica di Base", McGraw-Hill Editore). Per questo motivo, in molti punti del seguito si fa esplicito riferimento al libro di Gilardi (che verrà indicato semplicemente come "testo") e si è cercato di usare le stesse notazioni: questo vale in particolare per il capitolo sull'integrazione secondo Lebesgue; il capitolo su curve e superfici é invece più autonomo. In ogni caso, una lettura della dispensa "da sola" è sconsigliata, proprio perché certe definizioni, o risultati, necessari per la comprensione della teoria e già trattati sul testo non sono stati ripetuti, ma solo citati. Si avvisa inoltre il lettore che gli argomenti trattati nel seguito si possono trovare su diversi testi "tradizionali" di Analisi 2, svolti sicuramente con maggiore dettaglio che nella presente dispensa. Alcuni libri di riferimento possono essere:

N. Fusco, P. Marcellini, C. Sbordone, "Analisi Matematica due", Liguori Editore;

E. Giusti, "Analisi Matematica 2", Bollati Boringhieri;

G. Gilardi, "Analisi Due", McGraw-Hill Editore.

Contenuto

1 Complementi sull'integrazione secondo Lebesgue

Riportiamo in questo capitolo alcune osservazioni sulla teoria di Lebesgue aggiuntive rispetto agli argomenti trattati nel testo. Quanto segue presuppone la conoscenza dei primi 6 paragrafi del Cap. XI, con l'eccezione del Teor. 5.9 e del Cor. 5.10, che possono anche non essere letti dato che nel seguito vedremo dei risultati più generali e di più facile comprensione. In tutto il capitolo, indicheremo con A (più precisamente, avremmo dovuto scrivere $(A, \mathcal{E}, m)\dots$) lo spazio di misura in cui intendiamo ambientare la teoria. Il lettore può comunque pensare che sia $A = \mathbb{R}^N$.

La prima cosa che vogliamo fare è estendere la definizione di integrale di Lebesgue in modo da ammettere che il valore dell'integrale sia $+\infty$ o $-\infty$. Si osservi che, grazie all'Es. XI.6.3.2, se f è una funzione *misurabile*, sono tali anche le sue parti positiva e negativa. Nel seguito chiameremo sempre *sommabile* una funzione che verifica la Def. XI.3.1 del testo (il cui integrale nel senso di Lebesgue dunque esiste *finito*). Al termine "integrabile" verrà invece dato un significato diverso:

Definizione 1.1. *Sia f una funzione misurabile. Dico che $\int_A f = +\infty$ se e solo se f^- è sommabile e f^+* non *è sommabile. Analogamente, dico che $\int_A f = -\infty$ se e solo se f^+ è sommabile e f^-* non *è sommabile. Dico infine che f è* integrabile *se almeno una tra f^+ e f^- è sommabile.*

Dunque, in particolare ogni funzione sommabile è integrabile e, in più, ha integrale finito. Nel seguito, il termine *integrabile* avrà sempre il senso introdotto nella precedente Definizione; tale terminologia è motivata dal fatto che le funzioni integrabili sono la classe più larga di funzioni per le quali si riesce a dare senso a $\int_A f$ nell'ambito della teoria di Lebesgue. Infatti, se f è misurabile, ma né f^+ né f^- sono sommabili, si *rinuncia* a definire l'integrale di f, almeno nel senso della teoria di Lebesgue. Peraltro, vedremo più avanti che in certi casi si riesce a dar senso al simbolo $\int_A f$ anche se f non è integrabile. Tuttavia, il simbolo $\int_A f$ **non denoterà in tali casi l'integrale nel senso di Lebesgue**.

Notiamo anche che per le funzioni integrabili continuano a valere buona parte delle proprietà usuali dell'integrale (del tipo additività, linearità, monotonia, ecc.). Il "buona parte" è dovuto al fatto che possono comparire *forme indeterminate* ed in tal caso certe proprietà cadono. Il lettore provi ad esempio a chiedersi che cosa si può dire sull'integrale di $f + g$ nel caso in cui si abbia $\int_A f = +\infty$ e $\int_A g = -\infty$.

Osserviamo ora che la Def. precedente, come del resto la Def. XI.3.1 del testo, di per sè non è molto utile, perchè non fornisce nessun criterio per calcolare quando le condizione richieste su f^+ e f^- sono verificate nei casi concreti. Dunque, abbiamo bisogno di fornire degli strumenti appropriati che ci permettano di calcolare gli integrali nel senso di Lebesgue, finiti o infiniti che siano. Il primo passo che va in questa direzione è la seguente estensione del Teorema di Beppo Levi, che migliora il Teor. XI.5.3 del testo. L'enunciato che riportiamo è sufficientemente generale da comprendere tutti i casi in cui le cose funzionano; dunque d'ora in poi quando parleremo di "Teorema della convergenza monotona", intenderemo sempre l'enunciato seguente (vedi anche le Osservazioni successive), anziché il Teor. XI.5.3.

Teorema 1.2. *Sia $\{f_n\}$ una successione di funzioni integrabili e non negative e supponiamo che $\{f_n\}$ sia monotona non decrescente in n q.o. in A. Denotando con f il limite (q.o.) della $\{f_n\}$, si ha allora che f è integrabile e*

$$\int_A f = \lim_{n\to\infty} \int_A f_n. \tag{1.1}$$

Prova. Osserviamo innanzitutto che f è misurabile grazie al Teorema XI.6.5 del testo. Sussistono a questo punto due casi: o la successione $n \mapsto \int_A f_n$ è limitata e allora siamo nelle ipotesi del Teor. XI.5.3, grazie al quale concludiamo, oppure $\lim_{n\to\infty} \int_A f_n = +\infty$. In tal caso, dobbiamo mostrare che $\int_A f = +\infty$. Supponiamo per assurdo che ciò non valga. Allora, esiste $L \geq 0$ tale che $\int_A f = L$. Allora, essendo $f_n \leq f$ q.o., grazie alla monotonia dell'integrale si ha che

$$\int_A f_n \leq \int_A f = L \quad \forall\, n;$$

dunque, la successione $n \mapsto \int_A f_n$ è limitata, il che dà l'assurdo. ∎

Osservazione 1.3. La dimostrazione del Teorema precedente, in realtà, non è del tutto rigorosa, in quanto c'è una piccola difficoltà aggiuntiva un po' riposta. Infatti, sotto la sola ipotesi di non decrescenza della $\{f_n\}$, può accadere che la funzione limite f debba essere interpretata come una funzione da A in $(-\infty, +\infty]$. Addirittura potrebbe capitare che f valga $+\infty$ in tutti i punti di A. Tuttavia, si verifica facilmente che il Teorema continua a valere a patto di estendere la nozione di integrale alle funzioni f a valori in $(-\infty, +\infty]$ nel seguente modo naturale: se l'insieme E_∞ in cui f vale $+\infty$ è trascurabile, allora si "cambia valore" alla f in E_∞ (ad esempio, ponendola uguale a 0, ma la scelta è ininfluente) e si prende come integrale di f l'integrale della funzione cosí "aggiustata"; viceversa, se E_∞ non è trascurabile, si pone $\int_A f := +\infty$. Anzi, con questo tipo di convenzione si potrebbe ammettere che anche le $\{f_n\}$ siano funzioni a valori in $(-\infty, +\infty]$.

Osservazione 1.4. Il Teorema precedente può essere ulteriormente generalizzato sostituendo l'ipotesi di nonnegatività della f_n con la seguente:

$$\exists\, \phi \text{ sommabile:} \quad f_n \geq \phi \text{ a.e. in } A \quad \forall\, n \in \mathbb{N} \tag{1.2}$$

(anzi, grazie alla monotonia, basterebbe supporlo per f_1). Si noti che per dimostrare questa estensione basta applicare il Teorema alla successione non negativa $\{f_n - \phi\}$ (ed eventualmente tenere conto dell'Osservazione precedente).

Osservazione 1.5. Invece la proprietà di passaggio al limite sotto integrale in generale *non vale* se non si suppone la (1.2). Per vederlo, si prenda $A = \mathbb{R}$ e $f_n \equiv -1/n$. Ovviamente, $f_n \to f \equiv 0$, addirittura in modo uniforme; tuttavia,

$$\int_{\mathbb{R}} f_n = -\infty \quad \forall\, n, \qquad \int_{\mathbb{R}} f = 0$$

(la prima proprietà, intuitivamente evidente, non è stata ancora dimostrata, ma lo sarà tra breve).

Osservazione 1.6. Notiamo infine che la maggiore generalità del Teorema 1.2 ha un (piccolo) prezzo; quando si ha a che fare con integrali infiniti, infatti, non è più detto che valga la proprietà di "convergenza L^1" ((XI.5.6) del testo), che invece vale se la successione degli integrali è limitata.

Veniamo ora a descrivere degli strumenti per calcolare gli integrali nel senso di Lebesgue. Il risultato seguente (Teor. XI.8.1 del testo, che rienunciamo con le nuove notazioni) ci assicura che, nei casi già noti, le cose continuano a funzionare come sapevamo. Le maggiori difficoltà, invece, riguarderanno le situazioni nuove (integrali di funzioni non limitate, integrali infiniti, ecc.).

Teorema 1.7. *Sia f integrabile secondo Riemann. Allora f è* sommabile *secondo Lebesgue e i due integrali (nel senso di Riemann e nel senso di Lebesgue) coincidono.*

Richiamiamo ora qualche nozione sugli integrali su sottoinsiemi e sulla *misura* di Lebesgue; questo è l'argomento del paragrafo XI.7 del testo; tuttavia, con le nuove notazioni e il Teorema 1.2 a disposizione, le cose si semplificano. Ricordiamo che si può liberamente supporre $A = \mathbb{R}^N$.

Definizione 1.8. *Sia $B \subset A$. Diciamo che B è* misurabile secondo Lebesgue *se χ_B è* integrabile *e poniamo*

$$m(B) := \int_A \chi_B \quad \text{(valore finito o infinito)}. \tag{1.3}$$

Sia B misurabile e $f : A \to \mathbb{R}$ misurabile. Allora, se $f\chi_B$ è integrabile, diciamo che f è integrabile su B e poniamo

$$\int_B f := \int_A f\chi_B \quad \text{(valore finito o infinito)}. \tag{1.4}$$

In particolare, se $f\chi_B$ è sommabile diciamo che f è sommabile su B.

Osservazione 1.9. Si noti che, se f è integrabile (risp., sommabile) su A, allora f è sicuramente integrabile (risp., sommabile) su B (dimostrare per esercizio). Naturalmente può capitare che f sia integrabile (o sommabile) su B, ma non lo sia su tutto lo spazio A.

Per quanto riguarda le proprietà della misura, si veda il Teor. XI.7.2 del testo; è anche utile svolgere gli Esercizi XI.7.4. Ulteriori proprietà interessanti della misura e dell'integrale di Lebesgue sono anche descritte a partire dal Lemma XI.7.15 fino alla fine del Paragrafo XI.7. Per quanto riguarda invece la parte centrale del Par XI.7 e l'inizio del Par. XI.8, daremo ora un'impostazione alternativa e in parte diversa.

Cominciamo a introdurre una notazione che manterremo nel seguito: sia $B \subset A$ un insieme e siano $\{B_n\}$, $\{C_n\}$ due successioni di insiemi *misurabili* e tali che

$$B = \bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} C_n; \qquad B_i \cap B_j = \emptyset \quad \forall\, i \neq j; \qquad C_i \subset C_{i+i} \quad \forall\, i \in \mathbb{N}. \tag{1.5}$$

In altre parole, B è supposto essere unione della famiglia di insiemi $\{B_n\}$ a due a due *disgiunti* e della famiglia *crescente* di insiemi $\{C_n\}$. Si noti che le due situazioni sono da un certo punto di vista equivalenti; infatti, se B è unione disgiunta dei $\{B_n\}$, allora è unione della famiglia crescente $\{C_n\}$ ove $C_n := \cup_{i=1}^n B_i$; viceversa, se B è unione crescente dei $\{C_n\}$, è anche unione disgiunta dei $\{B_n\}$ ove $B_0 = C_0$, $B_n := C_n \setminus C_{n-1}$ per $n \geq 1$. Anzi, nel seguito supponiamo per semplicità di trattazione che le due famiglie $\{B_n\}$ e $\{C_n\}$ siano legate da questo tipo di relazione. Osserviamo anche che, grazie alle proprietà già note della misura, per ogni $n \in \mathbb{N}$, si ha

$$m(C_n) = m\big(\cup_{i=1}^n B_i\big) = \sum_{i=1}^n m(B_i). \tag{1.6}$$

Analogamente, sappiamo che, almeno quando f è una funzione *sommabile* in A,

$$\int_{C_n} f = \sum_{i=1}^n \int_{B_i} f. \tag{1.7}$$

Tale proprietà segue dalla linearità dell'integrale, dalla Definizione 1.8 e dal fatto che

$$\chi_{C_n} = \sum_{i=1}^n \chi_{B_i}.$$

Ci chiediamo ora se le proprietà (1.6) e (1.7) possano essere generalizzate al caso di unioni *numerabili* anziché finite. Inoltre, vogliamo capire cosa succede se l'ipotesi di sommabilità su f è sostituita dall'ipotesi, più debole, di integrabilità. Presentiamo due risultati, modellati rispettivamente sulla "versione aggiornata" del Teorema di Beppo Levi e sul Teorema di Lebesgue, nei cui enunciati manterremo le ipotesi sugli insiemi B, $\{B_n\}$, $\{C_n\}$ fatte poco sopra. Infine, presenteremo un Corollario che "metterà insieme" i due casi e fornirà la situazione più generale possibile in cui le cose funzionano.

Teorema 1.10. *L'insieme B è misurabile. Inoltre, se $f : A \to \mathbb{R}$ è integrabile e non negativa, si ha che*

$$\int_B f = \sum_{n=1}^\infty \int_{B_n} f = \lim_{n\to\infty} \int_{C_n} f. \tag{1.8}$$

In particolare, prendendo $f \equiv 1$, otteniamo

$$m(B) = \sum_{n=1}^\infty m(B_n) = \lim_{n\to\infty} m(C_n). \tag{1.9}$$

PROVA. Basta applicare il Teorema di Beppo Levi alla successione $\{f_n\}$ data da $f_n = f\chi_{C_n}$, che è non decrescente e non negativa. Si noti che la misurabilità di B (ossia l'integrabilità di χ_B) è parte della tesi quando si scelga $f \equiv 1$. ∎

Teorema 1.11. *Sia $f : A \to \mathbb{R}$ misurabile. Supponiamo che $|f|$ sia sommabile su B oppure che*

$$\text{esista finito il } \lim_{n\to\infty} \int_{C_n} |f|. \tag{1.10}$$

Allora f è sommabile su B e vale la (1.8).

PROVA. Innanzitutto, le due ipotesi date come alternative sono equivalenti grazie al Teorema precedente applicato a $|f|$. A questo punto, osservo che la successione $\{f_n\}$ data da $f_n = f\chi_{C_n}$ verifica $|f_n| \le |f|$ q.o.; dal momento che $|f|$ è sommabile, la tesi segue allora dal Teorema di Lebesgue. ∎

Corollario 1.12. *Sia $f : A \to \mathbb{R}$ integrabile. Allora vale la* (1.8).

PROVA. Dato che f è integrabile, si ha che almeno una tra f^+ e f^- è sommabile, ad esempio f^+. Allora posso applicare uno a scelta tra i Teoremi 1.11 e 1.10 a f^+ ed il Teorema 1.10 a f^- e concludere sommando le relazioni (1.8) scritte per f^+ ed f^- ed usando la linearità dell'integrale e del limite. ∎

I risultati che abbiamo visto hanno due conseguenze importanti: una di carattere teorico e un'altra di carattere operativo. Cominciamo a vedere quella teorica: le relazioni (1.9) e (1.8) ci mostrano due nuove proprietà, rispettivamente della misura e dell'integrale secondo Lebesgue, che non erano valide, in generale, nel caso dell'integrale di Riemann e della misura di Peano-Jordan: entrambe queste proprietà vanno sotto il nome di σ-additività. Nel caso della teoria di Riemann, le cose fallivano già in partenza; infatti, se anche gli insiemi B_n fossero supposti più regolari (ad esempio misurabili secondo Peano-Jordan), non c'è nessuna garanzia che l'insieme B sia anch'esso misurabile secondo P.-J.; dunque, non si può, in generale, costruire un integrale su B nel senso di Riemann.

La conseguenza interessante dal punto di vista del calcolo riguarda, parlando molto grossolanamente, la possibilità di includere gli infiniti nel calcolo degli integrali. Ricordiamo innanzitutto che il *supporto* di una funzione $f : A \to \mathbb{R}$ è definito da

$$\operatorname{supp}(f) := \overline{\{x \in A : f(x) \neq 0\}}. \tag{1.11}$$

Supponendo sempre $A = \mathbb{R}^N$, si ricorda che ogni funzione integrabile secondo Riemann è per ipotesi *limitata*; inoltre, è conseguenza della condizione di integrabilità il fatto che anche il suo supporto è limitato. Orbene, l'uso dei Teoremi 1.10 e 1.11 consente di ricondurre in modo rigoroso il calcolo di integrali (nel senso di Lebesgue) di funzioni non limitate, oppure definite su domini non limitati, al calcolo di integrali di tipo Riemann.

Esempio 1.13. Calcoliamo l'integrale

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

Poichè l'integranda è non negativa, scegliendo ad esempio $C_n := \{(1/n, n)\}$, possiamo applicare la formula (1.8) del Teorema 1.10. Inoltre, una volta che ci siamo ricondotti all'integrale su C_n, grazie al Teorema 1.7 questo può venire inteso nel senso di Riemann e calcolato con le tecniche usuali, in quanto C_n e f sono limitati e f è regolare. A questo punto un semplice calcolo ci dice che l'integrale in esame ha come risultato $+\infty$.

Questo tipo di procedimento funziona sempre quando f ha sempre lo stesso segno; infatti siamo sicuri di poter applicare il Teorema 1.10. Nel caso in cui f assume valori di segno diverso, bisogna essere un pochino più cauti. Infatti, per applicare il Teorema 1.11 (ovvero il Corollario 1.12), bisogna fare una verifica preliminare, ossia escludere il caso in cui né f^+ né f^- siano sommabili. Un modo comodo per fare questo è provare a verificare la condizione (1.10), il che ci consentirebbe di usare il Teorema 1.11. Se questo tentativo fallisce, si può provare ad applicare lo stesso tipo di condizione alle funzioni f^+ e f^-. Se ad esempio si trova che

$$\lim_{n\to\infty} \int_{C_n} f^+ \in \mathbb{R}, \qquad \lim_{n\to\infty} \int_{C_n} f^- = +\infty,$$

allora questo è sufficiente per dire che $\int_A f = -\infty$.

Integrali impropri. Abbiamo insistito sulle condizioni per l'applicabilità della formula (1.8) di σ-additività, dal momento che questa in certi casi non vale. Quello che può succedere è che *esista finito* il limite a destra, ma la funzione f *non sia sommabile* secondo Lebesgue. Il motivo per cui può accadere una cosa del genere è dovuto a fenomeni di compensazione simili a quelli che si verificano per le serie a termini di segno alterno. Anzi, è comodo in quest'ambito fare esempi immediatamente riconducibili alla teoria delle serie.

Esempio 1.14. Sia $B = [0, +\infty)$, $C_n = [0, n)$,

$$f := \sum_{i=1}^{\infty} \frac{(-1)^i}{i} \chi_{[i-1,i)}$$

(serie di funzioni che converge uniformemente). Allora è evidente che

$$\int_{C_n} f = \sum_{i=1}^{n} \frac{(-1)^i}{i}; \qquad \text{dunque} \ \lim_{n\to\infty} \int_{C_n} f \quad \text{esiste finito.}$$

Tuttavia si verifica facilmente che f^+ e f^- non sono sommabili; dunque il simbolo $\int_B f$ *non ha senso* nell'ambito della teoria di Lebesgue.

Il concetto di integrabilità in senso improprio, che ora introduciamo, serve proprio a dar significato a $\int_B f$ in casi come quelli dell'esempio precedente. Tuttavia, per maneggiare i fenomeni di compensazione bisogna procedere con una certa cura in modo da evitare ambiguità e forme indeterminate. Dunque, diamo una definizione che si riferisce a un caso ben specifico e, in particolare, ad una situazione monodimensionale.

Definizione 1.15. *Sia $B = [a, b) \subset \mathbb{R}$, ove $b \in (a, +\infty]$, sia $f : B \to \mathbb{R}$ misurabile. Supponiamo che f sia* localmente sommabile *in B, ossia che per ogni intervallo* chiuso e limitato *$I \subset B$, f sia sommabile in I. Se il*

$$\lim_{x\to b^-} \int_a^x f(t)\, dt \quad \textit{esiste finito}, \tag{1.12}$$

allora diciamo che f è integrabile in senso improprio *su B (ovvero che l'integrale improprio di f su B* converge*) e definiamo l'integrale improprio (che, ripetiamo ancora, può non coincidere con quello di Lebesgue) di f su B come il valore del limite. Analogamente, se il limite in* (1.12) *esiste infinito, diciamo che l'integrale improprio di f* diverge *e ancora lo poniamo uguale al valore del limite. Infine, se il limite in* (1.12) *non esiste, diciamo che l'integrale improprio di f* oscilla.

A questo punto può essere opportuno fare un confronto tra le possibilità relative all'integrabilità in senso improprio e quelle relative all'integrabilità nel senso di Lebesgue, facendo anche qualche esempio. Consideriamo dunque una funzione $f : [a, b) \to \mathbb{R}$ misurabile. Può allora capitare che: (a) l'integrale improprio converga, (b) l'integrale improprio diverga, (c) l'integrale improprio oscilli; inoltre, f può essere: (1) sommabile secondo Lebesgue, (2) integrabile con integrale infinito, (3) misurabile ma non integrabile (ossia $\int_B f^+ = \int_B f^- = +\infty$). Abbiamo allora che:

(1) $\Longrightarrow$ (a), grazie al Teorema 1.11;

(2) $\Longrightarrow$ (b), grazie al Cor. 1.12;

(3) non consente di dire nulla sull'integrale improprio. Infatti, nel caso dell'Es. 1.14, questo converge; se invece

$$f := \sum_{i=1}^{\infty} (-1)^i \chi_{[i-1,i)}, \tag{1.13}$$

allora l'integrale improprio oscilla; se infine

$$f := \sum_{i=1}^{\infty} \chi_{[2i-2,2i-1)} - \sum_{i=1}^{\infty} \frac{1}{i} \chi_{[2i-1,2i)}, \tag{1.14}$$

allora f non è integrabile secondo Lebesgue, ma l'integrale improprio diverge (positivamente).

In particolare, in tutti i casi in cui esiste l'integrale nel senso di Lebesgue, esiste anche l'integrale improprio e i due valori coincidono. Vediamo ora cosa succede prendendo invece le condizioni sull'integrale improprio come ipotesi:

(a) $\Longrightarrow$ (1) oppure (3); infatti, se vale (2), allora vale (b) e non (a);

(b) $\Longrightarrow$ (2) oppure (3); infatti, se vale (1), allora vale (a) e non (b);

(c) $\Longrightarrow$ (3); infatti, se vale (1), allora vale (a) e non (c), mentre se vale (2), allora vale (b) e non (c).

Notiamo che nella Def. 1.15, i "problemi" avvenivano al solo estremo b. Di solito, nei casi pratici, questi "problemi" sono di due tipi: o f non è limitata nell'intorno di b o $b = +\infty$ (o entrambe le cose contemporaneamente); in una simile situazione, anche se f è regolare, non si può parlare di integrale nel senso di Riemann e dunque il concetto di integrale improprio può essere di aiuto. Ci si può chiedere inoltre che cosa accade dal punto di vista dell'integrabilità in senso improprio quando si abbia a che fare con una funzione $f : (a, b) \to \mathbb{R}$ localmente sommabile, ma che ha questo tipo di "problemi" a entrambi gli estremi. La risposta è che, mentre il Corollario 1.12 si può comunque applicare con un'opportuna scelta dei $\{C_n\}$ (naturalmente a patto che f ne soddisfi le ipotesi), per l'integrabilità in senso improprio si usa *spezzare* l'integrale nei due contributi $\int_a^x$, $\int_x^b$, ove $x \in (a, b)$, ciascuno dei quali viene studiato separatamente. Naturalmente l'ipotesi di locale sommabilità implica che la scelta del "punto intermedio" x è arbitraria. Il motivo di questo modo di procedere è il desiderio di evitare ulteriori fenomeni di compensazione tra i due limiti agli estremi.

Esempio 1.16. Sia $B = (0, +\infty)$, $f(x) = \log x$. Allora, f è integrabile secondo Lebesgue in B e il suo integrale vale $+\infty$. Per quanto riguarda l'integrabilità di f in senso improprio, si può dire che

$$\int_1^{+\infty} f \quad \text{diverge;} \qquad \text{invece,} \ \int_0^1 f \quad \text{converge.}$$

Osservazione 1.17. Si noti che, per determinare, nei casi concreti, il comportamento di un integrale improprio (ovvero, per decidere se questo converge, diverge o oscilla), è disponibile una famiglia di criteri grosso modo modellati su quelli relativi alle serie. Qui ci limitiamo a citare una sorta di "criterio del confronto asintotico", valido per integrande non negative. Sotto le notazioni della Def. 1.15, se g è un'altra funzione localmente sommabile e non negativa tale che $\lim_{x\to b^-} f/g$ esiste finito e non nullo, allora l'integrale improprio di f converge (risp., diverge) se e solo se converge (risp., diverge) l'integrale improprio di g.

Osservazione 1.18. Ci si può anche chiedere perché nella Def. 1.15 si è preferito prendere il limite rispetto a un parametro "continuo" $x \to b^-$ piuttosto che, se ad esempio $b = +\infty$, richiedere una proprietà del tipo

$$\lim_{n\to\infty} \int_a^n g(t)\, dt \quad \text{esiste finito,} \tag{1.15}$$

che corrisponde esattamente alla (1.8) con la scelta di $C_n = [a, n]$. Il motivo è che la (1.15) è *strettamente più debole* della (1.12), come si vede scegliendo ad esempio $g(t) := f(2t)$, ove f è la funzione definita in (1.13). Si vede infatti che l'integrale improprio di g, ai sensi della Def. 1.15, oscilla; invece' il limite in (1.15) esiste e vale 0.

Notiamo infine che esistono altri concetti di integrale che trattano in modo diverso le compensazioni tra infiniti di segno opposto. Il lettore interessato può trovare un'altra situazione significativa (e utile a volte in analisi funzionale) nel paragrafo 8.14 del testo.

2 Curve e superfici

Scopo di questo capitolo è presentare alcuni concetti fondamentali su curve e superfici che sono trattati in dettaglio sul testo solo in certi casi particolari (vedi [VIII.7]).

Tali nozioni, oltre ad avere un'importanza autonoma, sono necessarie per una corretta comprensione del Teorema delle funzioni implicite, trattato nel prossimo capitolo, e delle sue conseguenze.

2.1 Curve in $\mathbb{R}^N$

Definizione 2.1. *Chiamiamo* curva regolare *in* $\mathbb{R}^N$ *un'applicazione* $\gamma : [a,b] \to \mathbb{R}^N$ *di classe* C^1 *(ove* $[a,b] \subset \mathbb{R}$ *è un intervallo chiuso e limitato) tale che*

$$\gamma'(t) \neq 0 \qquad \forall\, t \in (a,b). \tag{2.1}$$

Chiamiamo sostegno *di* γ *l'insieme immagine* $\gamma([a,b]) \subset \mathbb{R}^N$.

Per brevità indicheremo nel seguito col termine "curva" una curva regolare nel senso della precedente definizione; non siamo infatti interessati a indebolire le ipotesi. Vale la pena invitare il lettore a non confondere la curva γ, che è una funzione, col suo sostegno, che è un sottoinsieme di $\mathbb{R}^N$. Per chiarire il concetto di curva, è opportuno pensare alla variabile $t \in [a,b]$ come a un "tempo di percorrenza" e alla curva γ come alla legge oraria del moto di un punto materiale nello spazio $\mathbb{R}^N$. In quest'ottica è abbastanza naturale richiedere alla mappa γ almeno la continuità. Per capire il perché della regolarità C^1 e soprattutto della condizione (2.1), occorre introdurre il concetto di *velocità istantanea* o, se si preferisce, di *vettore tangente* a γ. Dati $t_0, t \in (a,b)$, sviluppando γ al primo ordine otteniamo

$$\gamma(t) - \gamma(t_0) = \gamma'(t_0)(t - t_0) + o(|t - t_0|) \qquad \text{per } \ t \to t_0.$$

Dal punto di vista geometrico, questa relazione ci dice che, per t vicino a t_0, la differenza $\gamma(t) - \gamma(t_0)$ è "bene approssimata" dal prodotto del vettore $\gamma'(t_0)$ per la quantità scalare $t - t_0$. Chiamiamo dunque $\gamma'(t_0)$ *vettore tangente* a γ nel punto t_0. In effetti, sembra ragionevole interpretare la direzione di $\gamma'(t_0)$ come *direzione tangente* alla curva. Forse leggermente meno evidente è però il significato del *modulo* di $\gamma'(t_0)$. Introduciamo allora una nuova

Definizione 2.2. *Siano* $\gamma : [a,b] \to \mathbb{R}^N$, $\eta : [c,d] \to \mathbb{R}^N$ *curve. Si dice che* γ *ed* η *sono* equivalenti *se esiste un'applicazione* $\delta : [c,d] \to [a,b]$ *di classe* C^1 *con* $\eta = \gamma \circ \delta$ *e tale che* $\delta'(t) > 0$ *per ogni* $t \in [c,d]$ oppure $\delta'(t) < 0$ *per ogni* $t \in [c,d]$. *Nel primo caso* $(\delta' > 0)$ *diciamo anche che* γ *ed* η *hanno lo stesso verso.*

Esercizio 2.3. Dimostrare che la relazione introdotta dalla definizione precedente è effettivamente una relazione di equivalenza nella famiglia delle curve regolari in $\mathbb{R}^N$. Mostrare inoltre che due curve equivalenti hanno lo stesso sostegno. Vale anche il viceversa?

Esercizio 2.4. Confondendo il tempo impiegato con lo spazio percorso (o, se si vuole, la curva col suo sostegno), a volte si parla di $\gamma'(t_0)$ come "vettore tangente in P_0", ove $P_0 = \gamma(t_0)$. Tuttavia non sempre questo è rigoroso (perché?).

Supponiamo ora per semplicità $\gamma : [0,1] \to \mathbb{R}^N$ e definiamo $\eta : [0,1/2] \to \mathbb{R}^N$ come $\eta(t) := \gamma(2t)$. La curva η è equivalente a γ, come si vede scegliendo $\delta(t) = 2t$. Cinematicamente, il punto materiale che si muove su η sta percorrendo lo stesso cammino del punto che si muove su γ, ma lo fa in metà tempo (impiega 1/2 anziché 1) e, corrispondentemente, con velocità doppia (in effetti, $\eta'(t) = 2\gamma'(2t)$). Questo suggerisce che il modulo di $\gamma'(t)$ indichi "quanto velocemente" il punto materiale si sta muovendo; si osservi che i due vettori η' e γ' hanno la stessa direzione (e lo stesso verso!). Dunque, da un punto di vista cinematico, la condizione (2.1) ci dice che un punto che si muove su una curva regolare non si può mai "fermare". Da un punto di vista geometrico, ci permette di costruire lo *spazio tangente* a γ al tempo $t \in (a,b)$, definito come

$$T_\gamma(t) := \operatorname{span}\{\gamma'(t)\}, \qquad \text{per } t \in (a,b). \tag{2.2}$$

Anzi, ponendo $P = \gamma(t)$ e $C = \gamma([a,b])$ (sostegno), di solito si preferisce parlare di spazio tangente a C, scrivendo $T_C(P)$ anziché $T_\gamma(t)$ (facendo naturalmente attenzione all'ambiguità generata da funzioni γ non iniettive). In quest'ottica definiamo anche lo *spazio normale*

$$N_\gamma(t) = N_C(P) := T_C(P)^\perp, \qquad \text{per } t \in (a,b), \quad \text{ove } P = \gamma(t), \tag{2.3}$$

che risulta naturalmente essere uno spazio vettoriale di dimensione $N-1$. Dunque, se $N = 2$, è possibile definire, senza ambiguità a parte quella dovuta alla scelta del verso, il *versore normale* a γ al tempo t[1].

Osservazione 2.5. Richiedere la condizione (2.1) ha anche un'altra conseguenza geometrica degna di nota e relativa alla regolarità del sostegno. Si consideri infatti la mappa $\gamma : [-1,1] \to \mathbb{R}^2$ definita da $\gamma(t) := (0,t^2)$ per $t \le 0$ e $\gamma(t) := (t^2,0)$ per $t > 0$. Si vede allora che γ è di classe C^1, ma $\gamma'(0) = 0$. Effettivamente il sostegno di γ presenta uno "spigolo" e non è dunque "liscio" come la regolarità C^1 farebbe sperare. La (2.1) consente di escludere questo tipo di "patologie".

Riferendoci al caso di un moto piano, ossia per $N = 2$, veniamo ora ad introdurre una sottofamiglia particolarmente importante di curve.

Definizione 2.6. *Si dice che una curva $\gamma : [a,b] \to \mathbb{R}^2$ è un* grafico *se esiste una funzione $f : [a,b] \to \mathbb{R}$ di classe C^1 tale che valga una delle seguenti condizioni:*

$$\gamma(t) = (t, f(t)) \quad \forall\, t \in [a,b] \qquad \text{oppure} \qquad \gamma(t) = (f(t), t) \quad \forall\, t \in [a,b]. \tag{2.4}$$

Osservazione 2.7. Normalmente come grafico della funzione f si intende l'insieme $\{(t, f(t)) : t \in [a,b]\} \subset \mathbb{R}^2$ e non la funzione γ introdotta dalla precedente definizione. Dunque, si usa dire che una curva "è" un grafico, ma in realtà è il sostegno ad esserlo. Peraltro questo piccolo abuso di linguaggio viene comunemente accettato.

Denotando con x, y le variabili in $\mathbb{R}^2$, la prima delle (2.4) ci dice grosso modo che γ è il grafico di una funzione $y = f(x)$. Analogamente, la seconda delle (2.4) identifica

[1] peraltro, anche in dimensione N, è possibile identificare nello spazio N_γ un vettore normale "privilegiato" a γ, detto *normale principale* alla curva

γ al grafico di una $x = f(y)$. Vediamo di rimuovere il "grosso modo" e parlare in termini più precisi. Innanzitutto, è immediato verificare che, assegnata $f : [a, b] \to \mathbb{R}$ di classe C^1, definendo $\gamma(t) := (t, f(t))$ (oppure $\gamma(t) := (f(t), t)$) si ottiene una curva regolare; in particolare la (2.1) vale senza alcuna richiesta aggiuntiva su f.

Più interessante è il problema inverso, ossia capire quali curve "sono" grafici. Partiamo, per questo scopo, da un esempio, prendendo la curva $\gamma : [0, 1] \to \mathbb{R}^2$ definita da $\gamma(t) := (2t, 2t)$. È chiaro allora che γ non è un grafico secondo la Def. 2.6. Tuttavia, la curva $\eta : [0, 2] \to \mathbb{R}$, $\eta(t) := (t, t) = \gamma(t/2)$, equivalente a γ, è un grafico (verifica addirittura entrambe le (2.4)) ed ha essenzialmente le stesse proprietà geometriche di γ (salvo il modulo della velocità di percorrenza). Dunque, è più interessante capire quando una data curva $\gamma : [a, b] \to \mathbb{R}$ è *equivalente* a un grafico, ossia può essere ricondotta a un grafico "aggiustando" la rapidità (ed eventualmente il verso) di percorrenza.

Proposizione 2.8. *Sia $\gamma : [a, b] \to \mathbb{R}^2$ una curva regolare, $t_0 \in (a, b)$. Allora γ è equivalente a un grafico almeno in un intorno di t_0.*

PROVA. Siano $x(t), y(t)$ le componenti di $\gamma(t)$. Grazie alla (2.1), si ha che $\gamma'(t_0) \neq 0$. Dunque, almeno una tra $x'(t_0)$ e $y'(t_0)$ è non nulla, supponiamo ad esempio la prima. Anzi, grazie al Teorema della permanenza del segno, si ha che $x'(t) \neq 0$ in un intorno (chiuso) I di t_0 contenuto in $[a, b]$. Dunque, la funzione $t \mapsto x(t)$ è strettamente monotona e invertibile in I. Supponendo ad esempio $x' > 0$ in I, si ha che $x(I) = [\alpha, \beta]$ per qualche $\alpha, \beta \in \mathbb{R}$, con $\alpha < \beta$, ed è definita la funzione inversa $t : [\alpha, \beta] \to I$, $t = t(x)$. Definendo $f : [\alpha, \beta] \to \mathbb{R}$ come $f(r) := y(t(r))$, si vede facilmente che vale la prima delle (2.4) almeno per $t = r \in [\alpha, \beta]$, da cui la tesi. Per $x' < 0$ si procede in modo pressoché analogo. ∎

È interessante analizzare geometricamente, almeno a livello intuitivo, la precedente Proposizione. Ponendo $P_0 = \gamma(t_0)$, questa ci dice che, in un intorno di P_0, il sostegno di γ può essere visto come grafico di una funzione $y = y(x)$ o di una funzione $x = x(y)$. In particolare, supporre $x'(t_0) \neq 0$ significa che la prima componente del vettore tangente in t_0 deve essere non nulla, ossia questo non deve essere verticale nel piano (x, y). In effetti, in questo caso abbiamo rappresentato il sostegno di γ tramite una funzione $y = y(x)$. Se invece $y'(t_0) \neq 0$, ossia se $\gamma'(t_0)$ non è orizzontale otteniamo evidentemente una $x = x(y)$ e, se entrambe le componenti di $\gamma'(t_0)$ sono non nulle, tutte e due le scelte sono possibili.

Ricordando che il grafico di ogni funzione $f : [a, b] \to \mathbb{R}$ di classe C^1 dà luogo a una curva ai sensi della Def. 2.6, possiamo infine affermare che, almeno localmente, ossia nell'intorno di ogni generico punto, gli oggetti introdotti nelle Def. 2.1 e 2.6 sono equivalenti. Invece, ovviamente, non è vero che ogni curva è *globalmente* rappresentabile come un grafico, come si vede considerando la circonferenza $\gamma : [0, 2\pi] \to \mathbb{R}^2$, $\gamma(t) = (\cos t, \sin t)$.

Lunghezza di una curva. Nel paragrafo [VIII.7.4] del testo si suggerisce (con notazioni un po' diverse dalle nostre) che $|\gamma'(t)|$ rappresenta la velocità scalare del punto il cui moto è descritto dalla curva γ, e questo dovrebbe essere stato sufficientemente chiarito dal discorso successivo alla Def. 2.2. Invece, resta meno giustificata la definizione di *lunghezza* della curva γ, data come

$$\text{lungh}(\gamma) := \int_a^b |\gamma'(t)|\,dt, \tag{2.5}$$

sebbene essa sia chiara almeno intuitivamente: infatti, se, come abbiamo visto, $|\gamma'(t)|$ è la velocità scalare puntuale, integrando sul tempo di percorrenza otteniamo la distanza percorsa. Ci proponiamo ora di dare un'ulteriore giustificazione geometrica della definizione (2.5), che parte dal concetto naturale ed intuitivo di lunghezza di un segmento.

Ricordando dunque che una partizione di $[a,b]$ è per definizione una famiglia finita $\mathcal{P} = \{t_0, t_1, \dots, t_n\}$ di istanti tali che $a = t_0 < t_1 < \dots < t_n = b$, definiamo la lunghezza della poligonale associata alla partizione $\mathcal{P}$ ed alla curva γ come

$$\text{lungh}_{\mathcal{P}}(\gamma) := \sum_{i=1}^{n} \big|\gamma(t_i) - \gamma(t_{i-1})\big|. \tag{2.6}$$

D'altronde, come ci si potrebbe aspettare da semplici considerazioni geometriche, per ogni partizione $\mathcal{P}$ si ha che

$$\text{lungh}_{\mathcal{P}}(\gamma) = \sum_{i=1}^{n} \Big| \int_{t_{i-1}}^{t_i} \gamma'(t)\,dt \Big| \le \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} |\gamma'(t)|\,dt = \text{lungh}(\gamma); \tag{2.7}$$

pertanto, prendendo l'estremo superiore al variare di tutte le partizioni $\mathcal{P}$, otteniamo

$$\sup_{\mathcal{P}} \text{lungh}_{\mathcal{P}}(\gamma) \le \text{lungh}(\gamma). \tag{2.8}$$

Dunque, l'oggetto definito in (2.5) e che abbiamo chiamato lunghezza è più grande della lunghezza "naturale" associata a qualunque poligonale inscritta in γ. Meno evidente è che debba valere anche la disuguaglianza inversa, che fornisce una giustificazione geometrica soddisfacente della definizione (2.5).

Per dimostrare tale disuguaglianza inversa, sia dato un arbitrario $\varepsilon > 0$ e, dal momento che γ è C^1, in corrispondenza di tale ε, prendiamo $\delta > 0$ tale che $|\gamma'(r) - \gamma'(s)| < \varepsilon$ per ogni coppia di punti $r, s \in [a,b]$ tali che $|r-s| < \delta$. Prendiamo allora una qualunque partizione $\mathcal{P}$ di $[a,b]$ tale che $t_i - t_{i-1} < \delta$ per ogni i. Scegliendo τ_i come un qualunque punto di $[t_{i-1}, t_i]$, otteniamo allora

$$\begin{aligned}
\int_{t_{i-1}}^{t_i} |\gamma'(t)|\,dt &\le \int_{t_{i-1}}^{t_i} \big|\gamma'(t) - \gamma'(\tau_i)\big|\,dt + \int_{t_{i-1}}^{t_i} |\gamma'(\tau_i)|\,dt \\
&\le \int_{t_{i-1}}^{t_i} \big|\gamma'(t) - \gamma'(\tau_i)\big|\,dt + \Big| \int_{t_{i-1}}^{t_i} \gamma'(\tau_i)\,dt \Big| \\
&\le \int_{t_{i-1}}^{t_i} \big|\gamma'(t) - \gamma'(\tau_i)\big|\,dt + \Big| \int_{t_{i-1}}^{t_i} \big(\gamma'(\tau_i) - \gamma'(t) + \gamma'(t)\big)\,dt \Big| \\
&\le 2\int_{t_{i-1}}^{t_i} \big|\gamma'(t) - \gamma'(\tau_i)\big|\,dt + \big|\gamma(t_i) - \gamma(t_{i-1})\big|.
\end{aligned} \tag{2.9}$$

Conseguentemente, sommando per $i = 1, \ldots, n$ otteniamo subito

$$\text{lungh}(\gamma) \le \text{lungh}_{\mathcal{P}}(\gamma) + 2\varepsilon(b-a) \le \sup_{\mathcal{P}} \text{lungh}_{\mathcal{P}}(\gamma) + 2\varepsilon(b-a); \tag{2.10}$$

quindi la proprietà voluta segue dall'arbitrarietà di ε.

Osservazione 2.9. La dimostrazione della proprietà precedente sfrutta pesantemente la regolarità C^1 della funzione γ. Ci si potrebbe chiedere se esistono nozioni di "lunghezza" che si adattano a "curve meno regolari". La risposta è sì, ma la costruzione è decisamente più complicata. Osserviamo invece che, con qualche aggiustamento, la (2.5) si adatta a curve "regolari a tratti", ossia continue e con al più un numero finito di punti di non derivabilità.

Lasciamo al lettore la dimostrazione della seguente semplice (e naturale) proprietà, basata sul Teorema di derivazione delle funzioni composte:

Proposizione 2.10. *Siano γ ed η curve equivalenti secondo la Def. 2.2. Allora* $\text{lungh}(\gamma) = \text{lungh}(\eta)$. ∎

Per concludere il discorso sulla lunghezza di una curva γ, osserviamo che tra le varie curve equivalenti a γ ce ne è una privilegiata, ossia quella che viene percorsa con velocità scalare unitaria. In effetti, se poniamo

$$s(t) := \int_a^t |\gamma'(r)|\, dr, \tag{2.11}$$

grazie alla (2.1) abbiamo che $s(t)$ è strettamente crescente e dunque invertibile. Inoltre, evidentemente, $s([a,b]) = [0, \text{lungh}(\gamma)]$. Denotando con $t = t(s)$ la funzione inversa e ponendo

$$\eta : [0, \text{lungh}(\gamma)] \to \mathbb{R}^N, \qquad \eta(s) := \gamma(t(s)), \tag{2.12}$$

usando la (2.11) e il Teorema di derivazione delle funzioni composte [VIII.1.1] abbiamo che

$$|\eta'(s)| := |\gamma'(t(s))t'(s)| = \frac{|\gamma'(t(s))|}{s'(t(s))} = 1 \qquad \forall\, s \in (0, \text{lungh}(\gamma)). \tag{2.13}$$

La funzione $s(t)$ prende il nome di *lunghezza d'arco*. Parametrizzare la curva γ tramite s, ossia considerare la curva η in (2.12) vuol dire percorrere γ con velocità scalare unitaria. Ovvero, per la curva $\eta(s)$ la distanza percorsa è uguale al tempo impiegato.

Per concludere la trattazione delle curve, osserviamo che la lunghezza definita in (2.5) può essere vista come un elemento di misura sul sostegno C di γ; ossia è possibile vedere C come uno spazio elementare di misura, ove gli insiemi elementari sono le immagini dei sottointervalli di $[a,b]$ tramite γ e la misura è la (2.5). In particolare, usando ancora il Teorema [VIII.7.1] (vedi in particolare la [VIII.(7.3)]) del testo, data una funzione $f : \gamma([a,b]) \to \mathbb{R}$ continua, otteniamo la *formula di integrazione di linea*, che per semplicità si usa scrivere come

$$\int_\gamma f\, ds = \int_a^b f(\gamma(t))|\gamma'(t)|\, dt. \tag{2.14}$$

Notiamo però che, intendendo il secondo membro della (2.14) come una *definizione* dell'oggetto a primo membro, il concetto di *integrale di linea* viene ad acquistare senso anche per funzioni f meno regolari.

2.2 Superfici in $\mathbb{R}^3$

Vogliamo ora introdurre le nozioni fondamentali relative alle superfici in $\mathbb{R}^3$. La limitazione al caso dello spazio tridimensionale è dovuta al fatto che la situazione generale in $\mathbb{R}^N$ è considerevolmente più complessa rispetto al caso delle curve. Invece, in $\mathbb{R}^3$, oltre al vantaggio evidente di poter "disegnare" una superficie, potremo ricalcare quanto detto relativamente alle curve in $\mathbb{R}^2$ che "sono grafici". Per questo motivo, cercheremo di impostare la trattazione imitando quanto più possibile la struttura del paragrafo precedente, enfatizzando le analogie tra curve in $\mathbb{R}^2$ e superfici in $\mathbb{R}^3$.

Per "riprodurre" la Def. 2.1, osserviamo che, mentre una curva è un oggetto monodimensionale, descritto al variare di *una* variabile tempo, geometricamente una superficie è bidimensionale e dunque ci aspettiamo sia descritta dal variare di *due* variabili. Abbiamo però il problema aggiuntivo di dover capire qual è ora l'analogo dell'intervallo $[a,b] \subset \mathbb{R}$ in cui deve essere definita la mappa che dà luogo alla superficie. Sicuramente, vogliamo che tale insieme, detto ad esempio K, sia chiuso e limitato; tuttavia, abbiamo bisogno che esso stesso sia "bidimensionale" e non "contenga" parti degeneri come ad esempio tratti di curve. Un modo di richiedere questo fatto è supporre che K sia chiusura di un aperto connesso Ω. In questo modo, tra l'altro, anche K risulta essere connesso e chiameremo *dominio* in $\mathbb{R}^2$ un insieme in tali condizioni. Diamo ora la definizione:

Definizione 2.11. *Sia dato un dominio $K = \overline{\Omega} \subset \mathbb{R}^2$. Chiamiamo* superficie regolare *in $\mathbb{R}^3$ un'applicazione $\alpha : K \to \mathbb{R}^3$ di classe C^1 verificante le seguenti condizioni:*

$$\alpha_{|\Omega} \qquad \text{è iniettiva;} \tag{S1}$$
$$J\alpha(u,v) \qquad \text{ha rango 2} \quad \forall\,(u,v) \in \Omega, \tag{S2}$$

ove $J\alpha$ indica la matrice Jacobiana della mappa α.

Cerchiamo ora di spiegare la definizione. Posticipando i chiarimenti su (S2) (che vedremo essere l'analogo della (2.1) per le curve e richiede più lavoro), vediamo di dettagliare prima la (S1). Nel caso delle curve, questo tipo di richiesta non c'era e, ad esempio, la $\gamma : [-100, 100] \to \mathbb{R}^2$, $\gamma(t) = (\cos t, \sin t)$ è una curva tutt'altro che iniettiva e descrive un punto che continua a girare per un po' sulla circonferenza unitaria. Nel caso delle superfici, l'interpretazione delle variabili in K come tempo si è persa e prevale invece quella di oggetto geometrico. In un certo senso, la Def. 2.11 definisce la superficie come la mappa α, ma in realtà quello che adesso conta di più è l'immagine $S := \alpha(K) \subset \mathbb{R}^3$, ossia l'analogo di quello che per le curve si era chiamato sostegno. Anzi, in quest'ambito, molto spesso si tende a chiamare "superficie" il "sostegno" S e non la mappa α; anche noi in molti casi accetteremo questa ambiguità e parleremo della superficie S. Tra l'altro, questo fatto è ulteriormente giustificato perché si possono definire oggetti $S \subset \mathbb{R}^3$ che geometricamente hanno tutte le caratteristiche di una superficie, ma che è complicato, o addirittura impossibile, descrivere attraverso la Def. 2.11.

Passiamo ora alla (S2). Per chiarirla fissiamo alcune notazioni che conserveremo per tutto il paragrafo. Chiamiamo (u,v) le variabili in K e (x,y,z) quelle nello spazio $\mathbb{R}^3$ codominio di α. Sia fissato un punto $(u_0,v_0) \in \Omega$ e sia $P_0 = (x_0,y_0,z_0) = \alpha(u_0,v_0)$

il corrispondente punto su S. Facciamo variare la variabile u e teniamo fissa la v. Otteniamo allora, per $t \in I$ intorno di 0 la curva $\varphi^u(t) := (u_0 + t, v_0)$ nel piano (u, v) (anzi, più precisamente, in Ω) tale che $\varphi^u(0) = (u_0, v_0)$. Prendiamo l'immagine di φ^u tramite α, consideriamo cioè $\gamma^u := \alpha \circ \varphi^u$, che è una curva su S tale che $\gamma^u(0) = P_0$. Analogamente, possiamo definire le curve $\varphi^v(t) := (u_0, v_0+t)$ e $\gamma^v := \alpha\circ\varphi^v$ a immagine rispettivamente in Ω e su S. A questo punto, calcoliamo ad esempio il vettore tangente a γ^u. Grazie al Teorema [IV.8.3], abbiamo

$$(\gamma^u)'(0) = \alpha_u(u_0, v_0)(\varphi_1^u)'(0) + \alpha_v(u_0, v_0)(\varphi_2^u)'(0) = \alpha_u(u_0, v_0)$$

e, analogamente, $(\gamma^v)'(0) = \alpha_v(u_0, v_0)$, ove α_u ed α_v denotano le derivate parziali della funzione α e φ_1^u e φ_2^u sono le componenti di φ^u nel piano (u, v).

Dunque, le curve γ^u e γ^v tracciate su S si ottengono incrementando separatamente le due "variabili tempo" u e v a partire da (u_0, v_0). Pertanto, è lecito interpretare i vettori $(\gamma^u)'(0) = \alpha_u(u_0, v_0)$, $(\gamma^v)'(0) = \alpha_v(u_0, v_0)$ tangenti a tali curve come vettori tangenti alla superficie. Vedendo α_u ed α_v come vettori colonna, la matrice Jacobiana di α risulta data da

$$J\alpha = \begin{pmatrix}\alpha_u & \alpha_v\end{pmatrix} = \begin{pmatrix} x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{pmatrix};$$

quindi, la (S2) equivale a dire che in ogni punto P_0 di S, esistono due vettori tangenti *linearmente indipendenti*. Questo porta alla

Definizione 2.12. *Dati una superficie S e un punto $P_0 = \alpha(u_0, v_0) \in S$, chiamiamo* piano tangente *ad S in P_0 (o ad α in (u_0, v_0)) lo spazio vettoriale*

$$T_S(P_0) = T_\alpha(u_0, v_0) := \operatorname{span}\{\alpha_u(u_0, v_0), \alpha_v(u_0, v_0)\}; \tag{2.15}$$

chiamiamo vettore normale a S in P_0 il vettore

$$\boldsymbol{n}_S(P_0) := \alpha_u(u_0, v_0) \wedge \alpha_v(u_0, v_0). \tag{2.16}$$

Il versore $\boldsymbol{n}_S/|\boldsymbol{n}_S|$ si dice versore normale *a S e lo spazio $N_S(P_0) := T_S(P_0)^\perp = \operatorname{span}\{\boldsymbol{n}_S(P_0)\}$* spazio normale *a S.*

Si osservi infatti che, per costruzione, $\boldsymbol{n}_S(P_0)$ è ortogonale al piano tangente $T_S(P_0)$.

Osservazione 2.13. Val la pena fare un'ulteriore osservazione di carattere geometrico: consideriamo una *generica* curva φ a sostegno in Ω parametrizzata in modo tale che $\varphi(0) = (u_0, v_0)$ e costruiamo la curva in S data da $\gamma := \alpha \circ \varphi$. Dal momento che

$$\gamma'(0) = \alpha_u(u_0, v_0)\varphi_1'(0) + \alpha_v(u_0, v_0)\varphi_2'(0), \tag{2.17}$$

otteniamo che *il vettore tangente a γ appartiene allo spazio $T_S(P_0)$*, poiché è combinazione lineare di α_u e α_v. Ne segue che T_S *esaurisce* i vettori di $\mathbb{R}^3$ che possono essere costruiti come vettori tangenti di curve a valori in S. Inoltre, (2.17) ci dice che $\gamma'(0)$ dipende da φ solo tramite le derivate delle sue componenti φ_1 e φ_2 calcolate in 0.

Grafici. Imitando la Def. 2.6 abbiamo:

Definizione 2.14. *Dato un dominio $K = \overline{\Omega} \subset \mathbb{R}^2$ e data una funzione $f : K \to \mathbb{R}$ di classe C^1, possiamo interpretare il grafico di f come una superficie ponendo*

$$\alpha(u, v) = (u, v, f(u, v)) \quad \forall\, (u, v) \in K. \tag{2.18}$$

In questo caso, diciamo che la superficie $S = \alpha(K) \subset \mathbb{R}^3$ "è un grafico".

Si noti che nella precedente definizione, abbiamo privilegiato le superfici della forma $z = f(x, y)$ e non abbiamo considerato quelle date da $y = f(z, x)$ o $x = f(y, z)$. Inoltre, abbiamo maggiormente enfatizzato la costruzione di una superficie S a partire da una funzione scalare f di due variabili, piuttosto che il viceversa.

Osservazione 2.15. *(Solo per il lettore particolarmente interessato.)* È peraltro ancora vero, come nel caso delle curve, che *ogni superficie* S è *localmente* grafico nel senso della Def. 2.14. Diamone solo una giustificazione intuitiva, dato che una dimostrazione rigorosa richiede l'uso del Teorema delle funzioni implicite. Sia data una superficie con le notazioni della Def. 2.11 e sia $P_0 = \alpha(u_0, v_0)$ un punto di S. Dato che $T_S(P_0)$ ha dimensione 2, almeno uno dei vettori della base canonica di $\mathbb{R}^3$, ad esempio $\boldsymbol{e}^3 = (0, 0, 1)$, non appartiene a $T_S(P_0)$. Anzi, poiché il piano tangente dipende con continuità da $P \in S$, questo è vero in un intorno Λ di P_0 su S. A questo punto, chiamando H la proiezione di Λ sul piano (x, y), si può mostrare che per ogni $(x, y) \in H$ esiste uno ed un solo $z \in \mathbb{R}$ tale che $(x, y, z) \in \Lambda$ e che la mappa $f : (x, y) \mapsto z$ è di classe C^1.

Si noti che, nel caso S sia un grafico, la (2.18) fornisce un'espressione semplice per i generatori dello spazio tangente:

$$\alpha_u = (1, 0, f_u), \qquad \alpha_v = (0, 1, f_v) \tag{2.19}$$

(pertanto la condizione (S2) è verificata per ogni funzione f) e per il vettore (e il versore) normale

$$\boldsymbol{n}_S = (-f_u, -f_v, 1), \qquad \frac{\boldsymbol{n}_S}{|\boldsymbol{n}_S|} = \frac{(-f_u, -f_v, 1)}{\sqrt{1 + f_u^2 + f_v^2}}. \tag{2.20}$$

Area di una superficie. In analogia con la (2.5), diamo la seguente definizione di *area* di una superficie S:

$$\operatorname{area}(S) := \int_K |\alpha_u(u, v) \wedge \alpha_v(u, v)|\, du\, dv. \tag{2.21}$$

Come ci saremmo aspettati, l'integrale su $[a, b]$ è stato sostituito da un integrale su K. Più indigesta è la sostituzione della velocità scalare $|\gamma'(t)|$ col termine integrando che, ai sensi della (2.16) rappresenta il *modulo del vettore normale*. Giustificare la (2.21) è tutt'altro che semplice e nel seguito elenchiamo solo qualche motivazione "intuitiva" alla base di questa scelta

- Innanzitutto, si noti α_u e α_v rappresentano le "velocità parziali" del punto che si muove su S relative al variare di una delle due distinte "variabili tempo"; il modulo del loro prodotto esterno tiene conto del prodotto dei loro moduli (dunque è tanto più grande quanto le velocità sono grandi) e dell'angolo tra i due vettori.

- Imitando la (2.2) (anche se con qualche complicazione in più) è possibile introdurre il concetto di superfici equivalenti. Analogamente alla Prop. 2.10, si dimostra che due superfici equivalenti hanno la stessa area, se questa è definita dalla (2.21). Pertanto, l'oggetto definito in (2.21) dipende solo dal "sostegno" S e non dalla scelta della parametrizzazione α.

- La giustificazione più profonda della (2.21) si dovrebbe ottenere però, come nel paragrafo precedente, approssimando la mappa α con applicazioni più semplici. Nel caso delle curve avevamo scelto le poligonali inscritte. Ora, osservando che, se α è un'applicazione affine della forma $\alpha(u,v) = P + u\boldsymbol{a} + v\boldsymbol{b}$, ove il punto $P \in \mathbb{R}^3$ e i vettori $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$ linearmente indipendenti sono fissati, si ha che, $\forall\,(u,v) \in K$, $\alpha_u(u,v) = \boldsymbol{a}$ e $\alpha_v(u,v) = \boldsymbol{b}$; pertanto, l'integrando in (2.21) è identicamente uguale a $|\boldsymbol{a} \wedge \boldsymbol{b}|$, ossia l'area del parallelogrammo individuato da $\boldsymbol{a}$ e $\boldsymbol{b}$; dunque, sempre da (2.21) leggiamo $\text{area}(S) = |\boldsymbol{a} \wedge \boldsymbol{b}| m_2(K)$, ove m_2 denota la misura di Lebesgue nel piano. Per rendersi conto che questo è quanto ci aspettavamo, si prendano ad esempio $P = 0$ e $K = [0,1] \wedge [0,1]$, In questo caso, $\text{area}(S) = |\boldsymbol{a} \times \boldsymbol{b}|$ ed in effetti S è proprio il parallelogrammo avente vertici in $0, \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b}$. A questo punto, data una generica superficie α, approssimandola al primo ordine come

$$\begin{aligned} \alpha(u,v) &= \alpha(u_0,v_0) + \alpha_u(u_0,v_0)(u-u_0) + \alpha_v(u_0,v_0)(v-v_0) \\ &\quad + o(|(u-u_0, v-v_0)|) \qquad \text{per } (u,v) \to (u_0,v_0), \end{aligned} \tag{2.22}$$

  vediamo che nell'intorno del punto P_0 la la (2.21) fornisce l'area della *miglior approssimazione affine* della superficie. Il punto che ora è più difficile che nel caso delle curve è la costruzione di una superficie "affine a tratti" che approssima la S e la dimostrazione del fatto che, migliorando l'approssimazione, il valore dell'area approssimata tende proprio al valore definito dalla (2.21).

Interpretando l'area di S come misura ed usando ancora il Teorema [VIII.7.1] (vedi in particolare la [VIII.(7.3)]) del testo, data una funzione $f : S \to \mathbb{R}$ continua, otteniamo la *formula di integrazione su superficie*

$$\int_S f\,dS = \int_K f(\alpha(u,v))|\alpha_u(u,v) \wedge \alpha_v(u,v)|\,du\,dv. \tag{2.23}$$

Anche in questo caso, segnaliamo che è possibile indebolire le ipotesi su f sotto cui ha senso la (2.23).

3 Funzioni implicite

3.1 Il Teorema delle funzioni implicite

Supponiamo che siano dati un aperto $\Omega \subset \mathbb{R}^N$, ove ci limitiamo per semplicità a $N = 2$ oppure $N = 3$, ed una funzione $F : \Omega \to \mathbb{R}$ sufficientemente regolare. Un problema che motiva l'introduzione del teorema delle funzioni implicite è la caratterizzazione degli *insiemi di livello*

$$F_c \subset \mathbb{R}^N, \qquad F_c := \big\{\boldsymbol{x} \in \Omega : F(\boldsymbol{x}) = c\big\} \tag{3.1}$$

al variare di $c \in \mathbb{R}$. Anzi, notando che si può sempre sostituire a F la funzione $F - c$, è sufficiente considerare il caso in cui $c = 0$ e studiare l'insieme F_0 degli *zeri* di F, che indicheremo nel seguito con $Z(F)$. In particolare, vogliamo trovare delle condizioni sulla funzione F che ci garantiscano che, nel caso $N = 2$, $Z(F)$ è (sostegno di) una curva in $\mathbb{R}^2$ ovvero, nel caso $N = 3$, che $Z(F)$ è (sostegno di) una superficie in $\mathbb{R}^3$. Premettiamo alcuni

Esempio 3.1. Sia $N = 2$ e prendiamo $F(x, y) = x^2 + y^2$. Allora, F_c è una curva (una circonferenza) per ogni $c > 0$, mentre $F_0 = \{0\}$ non è una curva ai sensi della Def. 2.1.

Esempio 3.2. Prendiamo ora $F(x, y) = (x^2+y^2)(x-1)$. Allora, $Z(F) = \{0\} \cup \{x = 1\}$, abbiamo cioè l'unione di due componenti connesse, una sola delle quali è una curva.

Esempio 3.3. Prendiamo ora $F(x, y) = (x^2 - y^2)$. Allora, $Z(F) = \{x = y\} \cup \{x = -y\}$; dunque, $Z(F)$ non è una curva, ma lo è localmente nell'intorno di ogni punto diverso da $(0, 0)$.

I due esempi mostrano una serie di cose interessanti: innanzitutto, il primo suggerisce che un'alta regolarità non gioca un ruolo decisivo, perché anche per funzioni C^∞ $Z(F)$ può non essere una curva; il secondo e il terzo esempio ci dicono che il problema va considerato da un punto di vista locale, e non globale. La domanda che ci porremo è dunque la seguente: "dato un punto $P \in Z(F)$, sotto quali condizioni esiste un intorno Λ di P tale che $Z(F) \cap \Lambda$ è una curva (o, se $N = 3$, una superficie)?". Anzi, dal momento che, localmente, ogni curva e ogni superficie è rappresentabile come un grafico, potremo direttamente chiederci quando $Z(F)$ è localmente un grafico.

Il Teorema di Dini, o "delle funzioni implicite" risponde proprio a questa domanda; prima di dettagliarne l'enunciato, vediamo però di darne un'ulteriore giustificazione. Ritornando alla (3.1), scritta ad esempio per $N = 2$ e $c = 0$, interpretiamola ora come un'equazione nelle incognite (x, y) di cui vogliamo caratterizzare le soluzioni. Dal momento che abbiamo *una* relazione scalare che lega *due* variabili indipendenti, ci aspettiamo che possano esistere infinite soluzioni. Quello che possiamo fare allora è *fissare* una delle due incognite, ad esempio x e cercare la, o le, soluzioni y dell'equazione, ossia gli zeri della funzione $y \mapsto F(x, y)$. Nel caso fortunato in cui, per ogni x "ammissibile" (non ci curiamo per il momento di essere troppo precisi), esista una e una sola soluzione y, possiamo porre $y := f(x)$ e dalla (3.1) otteniamo che

$$Z(F) = \big\{(x, f(x)) : x \text{ ammissibile}\big\};$$

abbiamo cioè scritto $Z(F)$ come grafico di una funzione di una variabile e ci siamo dunque ricondotti allo stesso problema di prima.

Possiamo finalmente enunciare il Teorema delle funzioni implicite, limitandoci per il momento al caso $N = 2$ che è più intuitivo e semplice geometricamente:

Teorema 3.4. (delle funzioni implicite). *Sia $\Omega \subset \mathbb{R}^2$ un aperto e $F \in C^1(\Omega; \mathbb{R})$. Sia $P_0 = (x_0, y_0) \in \Omega$ con $P_0 \in Z(F)$ e supponiamo che $F_y(P_0) \neq 0$. Allora esistono intorni I di x_0 e J di y_0 ed esiste $f : I \to J$ di classe C^1 e tale che*

$$Z(F) \cap I \times J = \mathrm{graf}(f). \tag{3.2}$$

Inoltre, si ha che

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))} \quad \forall\, x \in I. \tag{3.3}$$

PROVA. Supponiamo ad esempio $F_y(P_0) > 0$. Per il Teorema della permanenza del segno, esistono allora intorni I_0 di x_0 e $J = [y_0 - \varepsilon, y_0 + \varepsilon]$ di y_0 tali che $F_y > 0$ in $I_0 \times J$. In particolare, si ha che $F(x_0, y_0 - \varepsilon) < 0$ e $F(x_0, y_0 + \varepsilon) > 0$. Applicando ancora il Teorema della permanenza del segno alle funzioni $x \to F(x, y_0 \pm \varepsilon)$, si vede facilmente che esiste un intorno $I \subset I_0$ di x_0 tale che $F(x, y_0 - \varepsilon) < 0$ e $F(x, y_0 + \varepsilon) > 0$ per ogni $x \in I$. Dal momento che $F_y > 0$ in $I \times J$, $\forall\, x \in I$, la funzione $y \to F(x, y)$ è strettamente crescente in J e assume valori di segno opposto agli estremi. Per il Teorema degli zeri, $\forall\, x \in I$ esiste allora uno ed un solo $y \in \operatorname{int} J$ tale che $F(x, y) = 0$. Chiamo allora $f(x)$ tale valore y.

Per concludere devo mostrare che la funzione f così ottenuta è di classe C^1 e vale la (3.3). Cominciamo a mostrare la continuità. Siano $(x, f(x))$ e $(x+h, f(x+h))$ due punti distinti di $Z(F) \cap I \times J$. Applicando il Teorema di Lagrange alla funzione F, otteniamo l'esistenza di un punto $(\xi, \eta) \in I \times J$ appartenente al segmento di estremi $(x, f(x))$ e $(x+h, f(x+h))$ e tale che

$$0 = F(x+h, f(x+h)) - F(x, f(x)) = F_x(\xi, \eta)h + F_y(\xi, \eta)\big(f(x+h) - f(x)\big), \tag{3.4}$$

da cui

$$f(x+h) - f(x) = -\frac{F_x(\xi, \eta)}{F_y(\xi, \eta)}h. \tag{3.5}$$

Essendo F di classe C^1, le sue derivate parziali sono uniformemente limitate nel compatto $I \times J$; pertanto, facendo tendere h a zero in (3.5) otteniamo la continuità della funzione f nel generico punto x.

Infine, dividendo ambo i membri di (3.5) per h, mandando ancora h a zero e osservando che in tal modo $(\xi, \eta) \to (x, f(x))$, otteniamo l'espressione (3.3), da cui leggiamo per confronto anche la regolarità C^1 di f; infatti il secondo membro di (3.3) è continuo rispetto a x per la continuità della funzione composta. ■

Il nome del Teorema è giustificato dall'osservazione che l'equazione $F(x, y) = 0$ fornisce una *rappresentazione implicita*, cioè non direttamente calcolabile, della funzione f.

Vediamo ora di esaminare gli Esempi precedenti alla luce del Teorema appena dimostrato. Riferendoci ad esempio alla funzione $F(x, y) = x^2 + y^2 - 1$ (cfr. l'Es. 3.1), notiamo che $F_y = 2y$; dunque gli unici punti di $Z(F)$ in cui F_y si annulla sono $(\pm 1, 0)$. In effetti, disegnando $Z(F)$, si nota che nell'intorno di tali punti esso non è rappresentabile come grafico di una funzione $y = f(x)$. Tuttavia, è chiaramente possibile descrivere $Z(F)$ nell'intorno di $(\pm 1, 0)$ come grafico di una $x = g(y)$; in effetti, modifiche banali alla dimostrazione del Teorema consentono di far vedere che una condizione sufficiente perché ciò avvenga è che sia $F_x(P_0) \neq 0$.

Dunque, ci aspettiamo che i punti "cattivi" di $Z(F)$ siano quelli in cui *entrambe* le derivate parziali di F sono nulle. Questo è avvalorato dagli Esempi 3.2 ed 3.3. In entrambi i casi, si vede che nel punto eccezionale $P = 0$ il gradiente di F si annulla. Possiamo pertanto dire, in generale, che se ∇F non si annulla mai su $Z(F)$, allora tale insieme è *localmente grafico di una funzione di una variabile*. Diciamo anche che $P \in Z(F)$ è *regolare* se $\nabla F(P) \neq 0$ e che è *singolare* altrimenti.

Il passo successivo consiste nell'innalzare la dimensione spaziale. Presentiamo solo l'enunciato del Teorema di Dini in $\mathbb{R}^3$, osservando che la dimostrazione è analoga (fatte salve lievi complicazioni tecniche) al caso bidimensionale ed anzi sarebbe possibile dare anche un enunciato N-dimensionale, a patto di introdurre le notazioni opportune:

Teorema 3.5. (delle funzioni implicite in $\mathbb{R}^3$). *Sia $\Omega \subset \mathbb{R}^3$ un aperto e $F \in C^1(\Omega;\mathbb{R})$. Sia $P_0 = (x_0, y_0, z_0) \in \Omega$ con $P_0 \in Z(F)$ e supponiamo che $F_z(P_0) \neq 0$. Allora esistono intorni I di (x_0, y_0) e J di z_0 ed esiste $f : I \to J$ di classe C^1 e tale che*

$$Z(F) \cap I \times J = \mathrm{graf}(f). \tag{3.6}$$

Inoltre, si ha che

$$f_x(x,y) = -\frac{F_x(x,y,f(x,y))}{F_z(x,y,f(x,y))}, \qquad f_y(x,y) = -\frac{F_y(x,y,f(x,y))}{F_z(x,y,f(x,y))} \quad \forall\,(x,y) \in I. \tag{3.7}$$

Ovviamente, in questo caso, $Z(F)$ risulta essere, nell'intorno di P_0, una *superficie* in $\mathbb{R}^3$, localmente rappresentata come grafico di una funzione regolare $z = f(x,y)$.

Una domanda che il lettore a questo punto potrebbe porsi è se il Teorema di Dini fornisca qualche strumento per *calcolare* l'espressione analitica della funzione f. La risposta, purtroppo, è no: in generale non è possibile determinare la forma esplicita di f. Tuttavia, usando la formula (3.3) (ovvero la (3.7) nel caso tridimensionale), è possibile una rappresentazione qualitativa locale del grafico di f (ossia di $Z(F)$). Per vedere questo, riferendoci per semplicità al caso $\mathbb{R}^2$, cominciamo con l'osservare che se la funzione F è più regolare, ad esempio C^2, ragionando per confronto nella (3.3), otteniamo che anche f è C^2. Più in generale, si mostra per induzione che, se F è C^k, anche f lo è. Derivando allora la (3.3) e usando varie regole di derivazione e la (3.3) stessa, otteniamo che

$$f''(x) = -\frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{F_y^3}, \quad \forall\, x \in I, \tag{3.8}$$

ove tutte le funzioni a secondo membro sono calcolate in $(x, f(x))$. Ovviamente, se la regolarità di F lo consente, si possono ottenere formule per le derivate successive di f semplicemente derivando ripetutamente la (3.8); naturalmente i calcoli sono pesanti. Fermandosi al secondo ordine, comunque, le (3.3) e (3.8) permettono di scrivere lo sviluppo di Taylor del secondo ordine di f centrato in x_0 semplicemente conoscendo le derivate prime e seconde di F in $P_0 = (x_0, y_0)$.

3.2 Conseguenze geometriche del Teorema di Dini. Estremi vincolati

Mettiamoci ancora nel caso bidimensionale a andiamo a guardare la curva $C = \mathrm{graf}(f)$ ottenuta nel Teorema 3.4. Parametrizzandola ponendo $\gamma(x) := (x, f(x))$, $x \in I$, otteniamo l'espressione del vettore tangente

$$\gamma'(x) = (1, f'(x)) = \Big(1, -\frac{F_x(x,f(x))}{F_y(x,f(x))}\Big),$$

da cui si vede facilmente che il vettore

$$\boldsymbol{n}(x, f(x)) := (F_x(x, f(x)), F_y(x, f(x)) = \nabla F(x, f(x))$$

risulta essere *normale* alla curva per ogni $x \in I$. Questo fatto ha un'interpretazione geometrica immediata e plausibile: dal momento che $Z(F)$ è la curva di livello zero della F, il vettore ∇F, che rappresenta la direzione di massima variazione di F (vedi anche Oss. [VIII.10.5]), risulta essere ortogonale a tale curva. Ne segue che, se P è un punto regolare di $Z(F)$, allora può essere definito lo spazio normale a $Z(F)$ in P, che viene ad essere generato da $\nabla F(P)$. Analogamente, nel caso tridimensionale, ricordando la (2.20) si vede subito che il vettore ∇F è ortogonale alla superficie $Z(F)$ definita implicitamente.

Prima di parlare di estremi vincolati, occorre fornire una nuova estensione del Teorema di Dini, che risulterà più chiara grazie alle considerazioni geometriche fatte qui sopra. Supponiamo di avere un aperto $\Omega \subset \mathbb{R}^3$ e *due* funzioni $F, G : \Omega \to \mathbb{R}$. Vogliamo descrivere, almeno localmente, l'insieme $Z(F) \cap Z(G) \subset \mathbb{R}^3$ in cui entrambe le funzioni si annullano simultaneamente. Supponendo dunque di avere un punto $P_0 \in Z(F) \cap Z(G)$, la prima cosa da supporre è che P_0 sia regolare sia per F che per G; in questo modo il Teorema ci dice che $Z(F)$ e $Z(G)$ in un intorno di P sono superfici. Vogliamo studiare la loro intersezione; in particolare, vogliamo vedere quando questa è, sempre nell'intorno di P_0, (sostegno di) una *curva* in $\mathbb{R}^3$, come ci aspetteremmo geometricamente. Supponiamo allora che Dini ci dica che $Z(F)$ è localmente rappresentabile come grafico di una funzione $z = f(x, y)$ in un intorno di (x_0, y_0). Dobbiamo allora imporre che $G(x, y, f(x, y)) = 0$. Applicando ancora il Teorema 3.4 alla funzione $\Phi : (x, y) \mapsto G(x, y, f(x, y))$, otteniamo che (x, y) è un punto regolare per tale funzione purché almento una delle due derivate parziali, date da

$$\Phi_x = G_x + G_z f_x = G_x - \frac{G_z F_x}{F_z}, \qquad \Phi_y = G_y + G_z f_y = G_y - \frac{G_z F_y}{F_z} \tag{3.9}$$

(tutto calcolato in $(x, y, f(x, y))$) sia non nulla. Poiché già sappiamo che $F_z \neq 0$, tali condizioni sono equivalenti rispettivamente a

$$\det \begin{pmatrix} F_x & G_x \\ F_z & G_z \end{pmatrix} \neq 0, \qquad \det \begin{pmatrix} F_y & G_y \\ F_z & G_z \end{pmatrix} \neq 0. \tag{3.10}$$

Per interpretare tale condizione, osserviamo che a priori abbiamo supposto che, per quanto riguarda F, sia la variabile z a esplicitarsi come funzione di (x, y). Ripetendo il ragionamento nei casi in cui si esplicitano x oppure y, troviamo altre quattro condizioni alternative a quelle in (3.10). È infine facile vedere che imporre che almeno una delle sei condizioni trovate valga equivale a chiedere che

$$\det \begin{pmatrix} F_x & G_x \\ F_y & G_y \\ F_z & G_z \end{pmatrix} \neq 0, \tag{3.11}$$

ossia che i vettori ∇F e ∇G siano *linearmente indipendenti*.

Anche questo fatto ha una giustificazione geometrica plausibile: dal momento che ∇F e ∇G sono normali rispettivamente a $Z(F)$ e $Z(G)$, la (3.11) significa che le due superfici non devono avere lo stesso versore normale in P_0, in altre parole *non devono essere tangenti* in P_0.

Vediamo infine che questo ragionamento permette di determinare facilmente lo spazio tangente e lo spazio normale alla curva $C := Z(F) \cap Z(G)$. Dal momento infatti che ∇F e ∇G sono ortogonali rispettivamente a $Z(F)$ e $Z(G)$, a maggior ragione sono ortogonali a C. Pertanto, abbiamo trovato due vettori normali a C e linearmente indipendenti, da cui

$$N_C(P_0) = \operatorname{span}\left\{\nabla F(P_0), \nabla G(P_0)\right\}$$

e, corrispondentemente, lo spazio tangente $T_C(P_0)$ si costruisce come spazio ortogonale a $N_C(P_0)$.

Moltiplicatori di Lagrange. Veniamo infine al problema degli estremi vincolati. Chiamiamo *vincolo* una curva in $\mathbb{R}^2$ o in $\mathbb{R}^3$, oppure una superficie in $\mathbb{R}^3$, anche se naturalmente si potrebbero trattare situazioni molto più generali. Usiamo il simbolo V per indicare il vincolo, intendendo che V sia il *sostegno* della curva o della superficie. Supponiamo anche di avere una funzione $f : \Omega \to \mathbb{R}$ di classe C^1, ove Ω è un aperto di $\mathbb{R}^2$ (o di $\mathbb{R}^3$) contenente V. Il problema che ci poniamo è quello di determinare i punti di massimo e di minimo *assoluti* della funzione f sul vincolo V. Poiché f è continua e V è chiuso e limitato, tali punti sicuramente esistono grazie al Teorema di Weierstrass. Distinguiamo due casi.

Vincoli espliciti. Il caso più semplice si ha quando il vincolo è dato in forma esplicita, ossia si è nel caso della Def. 2.1 o della Def. 2.11. Nel primo caso, si considera la funzione composta $f \circ \gamma : [a, b] \to \Omega$ e ci si riconduce a studiare gli estremi di tale funzione di una variabile utilizzando i metodi usuali. Naturalmente il massimo ed il minimo assoluti potrebbero anche essere assunti negli estremi a, b corrispondenti agli estremi della curva V. Analogamente, se V è una superficie in $\mathbb{R}^3$ parametrizzata come in (2.11), si va a prendere la mappa composta $f \circ \alpha : K \to \mathbb{R}^3$ e si studiano gli estremi di tale funzione di due variabili. Anche in questo caso, bisogna fare attenzione al fatto che tali punti di estremo possono essere o all'interno di K o sul bordo ∂K, che è una curva nel piano e a sua volta potrebbe dover essere parametrizzato.

Vincoli impliciti. Può invece capitare che il vincolo V sia definito in forma implicita come luogo degli zeri di una funzione regolare $F : \Omega \to \mathbb{R}$ ovvero, nel caso V sia una curva in $\mathbb{R}^3$, come luogo degli zeri di una funzione $(F, G) : \Omega \to \mathbb{R}^2$. Per studiare il problema in questo caso, si parte ancora una volta da una semplice considerazione geometrica. Supponiamo che P_0 sia un punto di estremo assoluto di f su V e sia $\gamma : [a, b] \to \Omega$ una curva su V, ossia tale che il suo sostegno $\gamma([a, b]) \subset V$. Supponiamo inoltre che per qualche $t_0 \in (a, b)$ sia $\gamma(t_0) = P_0$. Allora è immediato dedurre che t_0 è un punto di estremo assoluto per la funzione composta $f \circ \gamma : [a, b] \to \Omega$. Ne segue in particolare che $(f \circ \gamma)'(t_0) = 0$. Grazie al Teorema di derivazione delle funzioni composte, segue

$$\nabla f(P_0) \cdot \gamma'(t_0) = 0. \tag{3.12}$$

D'altro canto, sappiamo che $\gamma'(t_0) \in T_{P_0}V$; anzi, tutti i vettori dello spazio $T_{P_0}V$ sono della forma $\gamma'(t_0)$ per un'opportuna scelta della curva γ tracciata su V (vedi Oss. 2.13). Deduciamo che $\nabla f(P_0) \perp T_{P_0}V$, ossia che $\nabla f(P_0) \in N_{P_0}V$. A questo punto, ricordando le caratterizzazioni degli spazi normali date nel paragrafo precedente otteniamo, nel caso di un solo vincolo F, la condizione

$$\exists\,\lambda \in \mathbb{R} \setminus \{0\} \quad \text{tale che} \quad \nabla f(P_0) + \lambda \nabla F(P_0) = 0; \tag{3.13}$$

qualora invece i vincoli F, G siano due, abbiamo

$$\exists\,\lambda, \mu \in \mathbb{R} \setminus \{0\} \quad \text{tali che} \quad \nabla f(P_0) + \lambda \nabla F(P_0) + \mu \nabla G(P_0) = 0. \tag{3.14}$$

Le condizioni (3.13), (3.14) vanno intese nel seguente modo: se P_0 è un punto di massimo o di minimo assoluto per f sul vincolo V, allora in P_0 è soddisfatta la (3.13) (o la (3.14)). Da un punto di vista pratico, quello che si fa nei casi concreti è di andare a cercare i punti P_i in cui la (3.13) (o la (3.14)) è verificata, quindi calcolare i valori $f(P_i)$ e vedere a quale punto corrisponde il massimo e a quale corrisponde il minimo. Naturalmente, anche in questo caso può accadere che il vincolo V abbia un bordo, ove, qualora V sia una curva, con "bordo" mi riferisco agli estremi di questa; nel caso V sia invece una superficie, ci si può rifare all'idea intuitiva di bordo, dato che non è del tutto semplice spiegare in modo rigoroso che cosa è una "superficie con bordo" in $\mathbb{R}^3$. In tal caso, nella ricerca dei punti di estremo va naturalmente considerato anche il "bordo" di V, che può dover essere a sua volta opportunamente parametrizzato. Segnaliamo infine che i numeri λ e μ in (3.13), (3.14) vengono normalmente detti *moltiplicatori di Lagrange*.