10. Approssimazioni e calcolo differenziale

Daremo ora un'introduzione al calcolo differenziale, nei limiti delle necessità del nostro corso. Arriveremo ai concetti di cui abbiamo bisogno prendendo in esame alcuni dei molti modi di approssimare una funzione data con un'altra più semplice. (S'intende che questo capitolo non sostituisce la trattazione più completa e rigorosa che verrà fatta nel corso di Analisi.) Fino a nuovo avviso, supporremo sempre di aver a che fare con funzioni $\mathbf{R} \to \mathbf{R}$, ossia *funzioni reali di variabile reale.*

Infinitesimi

Iniziamo introducendo delle notazioni che ci riusciranno utili per semplificare molti discorsi. Ricordiamo che una funzione $f : x \mapsto y = f(x)$ si dice *limitata* in un intervallo se esiste un reale positivo K tale che in ogni punto x dell'intervallo $|f(x)| < K$. Viceversa, f non sarà limitata se comunque si scelga K, esiste qualche punto dell'intervallo in cui $|f(x)| \ge K$. In parole, il valore assoluto di una funzione non limitata "diventa grande quanto si vuole."

Notazione 1: Se scriviamo $y = O(x^\alpha)$ $(\alpha > 0)$ intendiamo che esiste un intervallo $(-a, a)$ in cui $y/|x|^\alpha$ *è limitato*, e che α è il più grande reale per cui ciò accade. In queste condizioni diremo che y è *infinitesimo di ordine* α rispetto a x, oppure che y è *infinitesimo di ordine* x^α. In particolare, $y = O(1)$ sta a significare che y è limitato, ma $y/|x|^\beta$ non lo è se $\beta > 0$.

Esempio: $x \cos x = O(x)$; invece $\sqrt{|x|}$ non è $O(x)$, perché $\sqrt{|x|}/|x|$ non è limitato; ma $\sqrt{|x|} = O(x^{1/2})$, ossia $\sqrt{|x|}$ è infinitesimo di ordine $1/2$.

Notazione 2: $y = o(x^\alpha)$ significa invece che $y/|x|^\alpha \to 0$ per $x \to 0$. In particolare, $y = o(1)$ vuol dire che $y \to 0$ per $x \to 0$. Diremo che y è *infinitesimo di ordine superiore a* x^α (detto alla buona, y tende a zero *più rapidamente* di x^α).

Nota: Mentre $y = O(x^\alpha)$ implica $y = o(x^\beta)$ per $\beta < \alpha$, può benissimo accadere che sia $y = o(x^\beta)$, ma che non esista nessun α tale che $y = O(x^\alpha)$. Esempio: $y = x \log|x|$ è $o(x^\beta)$ se $\beta < 1$, ma per ogni $\alpha \ge 1$ si ha $|x|^{1-\alpha} \log|x| \to \infty$ per $x \to 0$.

Approssimazione di una funzione

Affronteremo ora il problema di approssimare una data funzione f con una funzione g, scelta in una certa classe.

Definizione: Diremo *errore dell'approssimazione* di f con g in un intervallo (a, b) il numero

$$\varepsilon = \sup_{t \in (a,b)} |f(t) - g(t)|.$$

In parole: ε è il più piccolo numero tale che la differenza fra f e g non supera ε in nessun punto dell'intervallo (a, b).

Approssimazione con costanti

La scelta più semplice possibile per g è una costante: perciò nell'intervallo $[t_0 - \tau, t_0 + \tau]$ poniamo $g(t) = f(t_0)$ (questa scelta di g verrà indicata con f_0). Sia ε l'errore dell'approssimazione di f con f_0 (che in generale, per un dato t_0, dipende da τ). Se si può migliorare l'approssimazione, ossia ridurre ε quanto si vuole, pur di prendere τ abbastanza piccolo, si dice che f è *continua in* t_0. In altre parole, f è *continua in* t_0 sse $\varepsilon = o(1)$ (fig. 10–1). Questo corrisponde all'idea intuitiva che una funzione continua varia di poco per piccole variazioni della variabile indipendente.

Controesempio: $f(t) = \lfloor t \rfloor$ (parte intera di t), $t_0 = 1$: allora $\forall \tau$ è sempre $\varepsilon \geq 1$ (fig. 10–2). In casi come questo ci si può accontentare di approssimare a destra o a sinistra: $t \in [t_0, t_0 + \tau)$ oppure $t \in (t_0 - \tau, t_0]$. La funzione $\lfloor t \rfloor$ è continua *a destra*, non a sinistra.

Una funzione può benissimo essere continua in ogni punto di un intervallo, ma a parità di τ l'errore ε in genere dipende da t_0 in un modo imprevedibile. Ad es. la funzione $f(t) = 1/t$ è continua nell'intervallo *aperto* $(0, 1)$: si verifichi per esercizio che $\varepsilon = \tau / t_0(t_0 - \tau)$ (naturalmente occorre prendere $\tau < t_0$). In questo caso a parità di τ l'errore cresce indefinitamente quando t_0 diminuisce: dunque non è possibile approssimare ugualmente bene la funzione in tutti i punti dell'intervallo. Però questa spiacevole situazione non si presenta se la funzione è continua in un intervallo *chiuso*, come dice il

Teorema: *Se f è continua in ogni punto di un intervallo* chiuso $[a, b]$, *l'errore ε, fissato τ, ha un massimo in* $[a + \tau, b - \tau]$. In altre parole: fissato ε, esiste τ tale che $|f(t_1) - f(t_2)| < \varepsilon$ tutte le volte che $t_1, t_2 \in [a, b]$ e $|t_1 - t_2| < \tau$. Si dice che f è *uniformemente continua*. Tralasciamo la dimostrazione.

L'insieme delle funzioni continue in $[a, b]$ s'indica con $C^0[a, b]$.

Approssimazione con funzioni lineari

Dopo la costante, la scelta più semplice per g è una funzione lineare: nell'intervallo $[t_0 - \tau, t_0 + \tau]$ poniamo

$$g(t) = f(t_0) + k\,(t - t_0).$$

Si vede subito che se f è continua in t_0 la g approssima f con errore $o(1)$, comunque si scelga k. Si può fare in modo che l'approssimazione sia "migliore"? Con ciò s'intende scegliere k per avere errore $o(\tau)$. Come vedremo, la risposta non è sempre positiva; se un tale k esiste, lo indicheremo con $f'(t_0)$, e la corrispondente g con f_1.

Teorema: *Se $f'(t_0)$ esiste, è unico.*

Dim.: Sia $g = f(t_0) + k(t - t_0)$, $\bar{g} = f(t_0) + \bar{k}(t - t_0)$ e siano ε, $\bar{\varepsilon}$ i rispettivi errori. Supponiamo che ε ed $\bar{\varepsilon}$ siano entrambi $o(\tau)$. Abbiamo

$$|g(t) - \bar{g}(t)| \leq |f(t) - g(t)| + |f(t) - \bar{g}(t)|$$

e perciò

$$\sup |g(t) - \bar{g}(t)| \le \varepsilon + \bar{\varepsilon}$$

ossia

$$|k - \bar{k}|\,\tau \le \varepsilon + \bar{\varepsilon}.$$

Ora se $k \neq \bar{k}$ il primo membro è $O(\tau)$, mentre il secondo è $o(\tau)$ per ipotesi: dunque $k = \bar{k}$. ∎

Interpretazione grafica: Il grafico di f_1 è la retta tangente a quello di f in t_0 (fig. 10–3); il teorema dice che se il grafico di f ha una tangente in t_0, questa è unica.

Controesempio: $f(t) = t^{1/3}$ è continua in $t_0 = 0$. Però $g(t) = kt$ per τ abbastanza piccolo dà errore $\tau^{1/3} - k\tau$, e $\tau^{-2/3} - k \to \infty$ per $\tau \to 0$, qualunque sia k. Questo succede perché il grafico di f in $t_0 = 0$ ha tangente "verticale" (fig. 10–4).

Definizione: Se f ammette approssimazione lineare $o(\tau)$, si dice che è *derivabile* (o *differenziabile*) in t_0; $f'(t_0)$ è la *derivata* in t_0 e

$$df \stackrel{\text{def}}{=} f_1(t) - f_0(t) = f'(t_0)(t - t_0)$$

si chiama il *differenziale* della f in t_0.

Per la funzione $f(t) = t$ ovviamente $df = dt = t - t_0$, e perciò

$$df = f'(t_0)\,dt \qquad f'(t_0) = \frac{df}{dt} \quad \text{(notazione di Leibniz)}.$$

Ancora:

$$f(t) = f(t_0 + dt) = f(t_0) + df + o(dt).$$

Da qui si vede che

$$\frac{df}{dt} = \lim_{dt \to 0} \frac{f(t_0 + dt) - f(t_0)}{dt}$$

che è la classica definizione di derivata.

Nota: L'uso che abbiamo fatto del differenziale è piuttosto scorretto, in quanto non se ne capisce bene la natura: è una funzione di due variabili (t e t_0)? è una nuova variabile indipendente? oppure è una funzione dell'"incremento" $t - t_0$? Quest'ultima è la risposta più vicina all'uso matematico corrente, ma per il momento non insistiamo.

Se f non è derivabile in t_0, può darsi che esista un k che soddisfa la condizione di approssimazione $o(\tau)$ solo per $t > t_0$, oppure solo per $t < t_0$: in tal caso diremo che f è derivabile *a destra* oppure *a sinistra*. Può anche accadere che f sia derivabile tanto a destra quanto a sinistra, ma con derivate distinte: un esempio banale è $f(t) = |t|$ in $t = 0$ (fig. 10–5). Il grafico di una funzione siffatta ha un *punto angoloso* in t_0.

Attenzione: Una funzione derivabile a destra e a sinistra (ma con derivate distinte) *non è derivabile* in senso stretto! Quando diremo "derivabile" senza specificare, intenderemo sempre in senso stretto.

Definizione: Se f è derivabile in tutti i punti di (a, b) si dice *derivabile in* (a, b) e la derivata definisce in (a, b) una nuova funzione f': la *derivata* (*prima*) di f.

Nota: Il fatto che f sia derivabile non implica che f' sia continua. Ad es. $f(t) = t^2 \sin(1/t)$ è derivabile per tutti i t reali, ma la derivata non è continua in $t = 0$. Infatti $f'(0) = 0$, mentre esistono infiniti punti, vicini a 0 quanto si vuole, in cui $f'(t) = 2t \sin(1/t) - \cos(1/t)$ è maggiore per es. di 1/2.

L'insieme delle funzioni derivabili con derivata continua in (a, b) si indica con $C^1(a, b)$.

Primi risultati sulle derivate

Per l'impiego pratico sono essenziali alcuni semplici risultati, che ora esponiamo (il più delle volte senza dimostrazione).

Teorema (banale!): *Una funzione costante in un intervallo ha derivata nulla in ogni punto.* Non è invece affatto banale l'inverso:

Teorema: *Se una funzione ha derivata nulla in un intervallo, essa è costante.* La dimostrazione sarà data nel corso di Analisi.

Teorema: *Se f, g sono derivabili in t_0 lo è $h = af + bg$ (a, b reali), e si ha* $h'(t_0) = af'(t_0) + bg'(t_0)$. In altri termini:

1) *le funzioni derivabili in t_0 formano uno spazio vettoriale $\mathcal{V}$ sui reali*
2) *la derivata è un* omomorfismo *di $\mathcal{V}$ su* $\mathbf{R}$.

La dimostrazione è ovvia.

Teorema: *Se f, g sono derivabili in t_0 lo è $h = fg$, e si ha*

$$h'(t_0) = f'(t_0)\, g(t_0) + f(t_0)\, g'(t_0)$$

(le funzioni derivabili formano un'*algebra* rispetto alla moltiplicazione).

Dim.: Abbiamo

$$\begin{aligned} f(t) &= f(t_0) + f'(t_0)\,(t - t_0) + o(t - t_0) \\ g(t) &= g(t_0) + g'(t_0)\,(t - t_0) + o(t - t_0) \end{aligned}$$

e allora:

$$f(t)\, g(t) = f(t_0)\, g(t_0) + \bigl(f'(t_0)\, g(t_0) + f(t_0)\, g'(t_0)\bigr)(t - t_0) + o(t - t_0)$$

da cui

$$h(t) = h(t_0) + h'(t_0)\,(t - t_0) + o(t - t_0),$$

e da qui la tesi. ■

Osservazione: Questa "regola del prodotto" è caratteristica della derivazione, e può essere usata per definirla in modo astratto.

Teorema: *Se f, g sono derivabili in t_0, e se $g(t_0) \neq 0$, anche $h = f/g$ è derivabile, e si ha*

$$h'(t_0) = \frac{f'(t_0)\, g(t_0) - f(t_0)\, g'(t_0)}{[g(t_0)]^2}$$

Dim.: Osserviamo che $f = gh$: allora

$$f'(t_0) = g'(t_0)\, h(t_0) + g(t_0)\, h'(t_0)$$

e basta qualche passaggio per arrivare al risultato voluto. ■

Teorema (di derivazione della funzione inversa): *Se f è derivabile in t_0 con derivata non nulla, ed è strettamente crescente (o decrescente) in un intervallo aperto contenente t_0, la funzione inversa $g : u \mapsto g(u) = f^{-1}(u)$ è derivabile in $u_0 = f(t_0)$, e si ha $g'(u_0) = 1/f'(t_0)$.*

Tralasciamo la dimostrazione.

Teorema (di derivazione delle funzioni composte): *Se f è derivabile in t_0, e g è derivabile in $u_0 = f(t_0)$, la funzione $h = g \circ f$ è derivabile in t_0 e si ha*

$$h'(t_0) = f'(t_0)\, g'(u_0).$$

Tralasciamo la dimostrazione.

Teorema: *La funzione $f : t \mapsto t^\alpha$, $\alpha \in \mathbf{R}$, è derivabile per ogni t* (con l'eccezione di $t = 0$ se $\alpha < 0$). *La derivata vale $\alpha\, t^{\alpha-1}$.*

Dim.: Si procede in più tempi:

1) Se α è intero positivo, abbiamo

$$(t+h)^\alpha = t^\alpha + \alpha\, t^{\alpha-1} h + o(h)$$

e di qui segue immediatamente la tesi.

2) Per α intero negativo, si scrive $t^\alpha = 1/t^{|\alpha|}$ e ci si appoggia sul teorema di derivazione del quoziente.

3) Per $\alpha = m/n$ razionale, si osserva che ponendo

$$g : u \mapsto v = u^n, \qquad h : t \mapsto v = t^m$$

si ha $h = g \circ f$. Basta allora il teorema di derivazione delle funzioni composte per arrivare al risultato.

4) Il caso generale di α reale deriva per continuità dal caso razionale; omettiamo i dettagli. ■

Applicando ripetutamente questi teoremi si possono calcolare le derivate di una larga classe di funzioni, che include tutte le funzioni *algebriche*, ossia quelle ottenute attraverso somme, prodotti, potenze e radici. In particolare i *polinomi* e le funzioni *razionali*, ossia i quozienti di due polinomi. Restano invece escluse molte funzioni *trascendenti* (ossia non algebriche): il caso più ovvio è quello delle funzioni *circolari* (o trigonometriche). Ce ne occuperemo nel prossimo capitolo.

Approssimazione con funzioni quadratiche

Facciamo ora un terzo passo: supposta f derivabile, poniamo

$$g(t) = f_1(t) + k\,(t - t_0)^2.$$

Si vede facilmente che per un k generico l'approssimazione resta $o(\tau)$. Se esiste k tale che l'errore sia $o(\tau^2)$ diremo che f è derivabile (o differenziabile) *due volte*, e indicheremo la corrispondente g con f_2:

$$\begin{aligned}
f(t) &= f_2(t) + o(\tau^2) \\
f_2(t) &= f_0 + f'_0 \cdot (t - t_0) + \tfrac{1}{2} f''_0 \cdot (t - t_0)^2 \qquad \text{(il fattore } \tfrac{1}{2} \text{ è comodo!)} \\
f(t) &= f_0 + df + \tfrac{1}{2} d^2 f + o(dt^2) \qquad f''_0 = \frac{d^2 f}{dt^2}.
\end{aligned}$$

La f''_0 si chiama *derivata seconda* di f in t_0.
Interpretazione grafica: Il grafico di f_2 è la parabola *osculatrice* a quello di f in t_0 (fig. 10–6).
Osservazione: Perché la notazione $d^2 f$? Abbiamo

$$\begin{aligned}
f(t_0 + dt) &= f(t_0) + f'(t_0)\,dt + \tfrac{1}{2} f''(t_0)\,dt^2 + o(dt^2) \\
f(t_0 - dt) &= f(t_0) - f'(t_0)\,dt + \tfrac{1}{2} f''(t_0)\,dt^2 + o(dt^2)
\end{aligned}$$

e sommando

$$f(t_0 + dt) + f(t_0 - dt) - 2f(t_0) = f''(t_0)\,dt^2 + o(dt^2).$$

Dunque $d^2 f$ è la "differenza centrale seconda" a meno di $o(dt^2)$.
Teorema: *Se f è derivabile in (a, b) con derivata anch'essa derivabile, allora f è derivabile due volte, e viceversa. Si ha poi $f'' = (f')'$.* Omettiamo la dimostrazione.

Continuando

Si può iterare il procedimento seguito fin qui, tentando d'introdurre approssimazioni via via migliori, e parallelamente una derivata terza, una derivata quarta, ecc. Possono accadere due cose:

a) esiste la derivata n-ma, ma non la successiva

b) esistono derivate di qualunque ordine.

Nel primo caso diremo che la funzione è derivabile n volte, nel secondo caso che è *infinitamente derivabile*.

Esempi: La funzione $x^5 \sin(1/x)$ è derivabile 5 volte in $x = 0$; la funzione x^5 è infinitamente derivabile; così pure $\exp(-1/x^2)$, che ha derivate tutte nulle.

Se una funzione è derivabile n volte in un intervallo (a, b), con derivate tutte continue, diremo che $f \in C^n(a, b)$; se è infinitamente derivabile diremo che $f \in C^\infty(a, b)$.

Osservazione: Tutto quanto detto fin qui non richiede che f sia $\mathbf{R} \to \mathbf{R}$. Infatti per l'insieme immagine si è richiesto che fossero definite:

a) una differenza

b) la moltiplicazione per i reali

c) una "norma" (per dare significato all'errore)

d) l'esistenza dei limiti.

Tutte queste proprietà valgono ad es. per uno spazio vettoriale dotato di prodotto scalare: potremo dunque definire derivate di funzioni a valori vettoriali. Valgono anche nel corpo complesso: possiamo perciò derivare funzioni $\mathbf{R} \to \mathbf{C}$.

Funzioni di più variabili

Consideriamo ora funzioni $\mathbf{R}^n \to \mathbf{R}$ (per semplicità $n = 2$, ma è inessenziale; v. anche l'osservazione alla fine del par. precedente per lo spazio immagine). Le idee e le definizioni introdotte finora si generalizzano senza difficoltà; qui ci limiteremo però a differenziare una sola volta.

Intorno e raggio: Sono possibili infinite definizioni del concetto di intorno, ad es. le seguenti tre: diciamo *intorno* di raggio ϱ del punto (t_0, u_0) l'insieme dei punti (t, u) che soddisfano

$$\begin{aligned}
&a) \qquad |t - t_0| + |u - u_0| < \varrho \\
&b) \qquad \sqrt{(t - t_0)^2 + (u - u_0)^2} < \varrho \\
&c) \qquad \max\left(|t - t_0|, |u - u_0|\right) < \varrho
\end{aligned}$$

(è essenziale notare che le condizioni hanno $<$, non $\leq$). Nel piano (t, u) la prima definizione individua un quadrato coi lati paralleli alle bisettrici degli assi cartesiani; la seconda un cerchio; la terza un quadrato coi lati paralleli agli assi. Le tre definizioni sono equivalenti, nel senso che definiscono *la stessa topologia* in $\mathbf{R}^2$ (v. il corso di Analisi).

Definizione di aperto: Si dice *aperto* un insieme che contiene un intorno di ogni suo punto.

Osservazione: L'equivalenza delle tre definizioni di intorno si vede nel fatto che portano agli stessi aperti (lasciamo la verifica per esercizio).

Esempi: Un singolo punto, un segmento, una curva qualsiasi, non sono aperti; non lo è neppure un cerchio o un quadrato, se s'intende che ne facciano parte sia il contorno, sia l'interno; sono invece aperti il cerchio e il quadrato senza contorno; anche l'intero piano è aperto.

Definizione: Una funzione g approssima f con errore ε in un aperto A se

$$\varepsilon = \sup_{(t,u)\in A} |f(t,u) - g(t,u)|.$$

Come si vede, l'unico cambiamento è che si parla di "aperto A" in luogo di "intervallo aperto," come si faceva finora. L'approssimazione con costanti e la definizione di funzione continua e uniformemente continua non richiedono cambiamenti: solo che una funzione è uniformemente continua in un insieme *compatto* (generalizzazione dell'intervallo chiuso: la questione verrà chiarita nel corso di Analisi).

Approssimazione con funzioni lineari

Preso un punto (t_0, u_0) e un intorno di raggio ϱ poniamo

$$g(t,u) = f(t_0,u_0) + h\,(t - t_0) + k\,(u - u_0).$$

Se f è continua, g approssima f con errore $o(1)$ rispetto a ϱ. Se poi esistono h, k tali che l'errore sia $o(\varrho)$, ossia

$$f(t,u) = f(t_0,u_0) + h\,(t - t_0) + k\,(u - u_0) + o(\varrho) \qquad (10\text{–}1)$$

diremo che f è *differenziabile* in (t_0, u_0).

Se fissiamo u al valore u_0, ossia se restringiamo la funzione f a un segmento parallelo all'asse t, possiamo chiederci se la funzione così ottenuta sia derivabile (rispetto all'unica variabile t). In caso affermativo:

Definizione: Tale derivata si chiama *derivata parziale di f rispetto a t*, e s'indica con f_t. Analogamente si definisce la derivata parziale rispetto a u. Si usano anche le notazioni:

$$f_t = \frac{\partial f}{\partial t} \qquad f_u = \frac{\partial f}{\partial u}.$$

Teorema: *Se f è differenziabile in (t_0, u_0), esistono le derivate parziali in (t_0, u_0), e i coefficienti h e k nella* (10–1) *coincidono con f_t, f_u.*

Dim.: Basta guardare la (10–1), per $u = u_0$: ne segue proprio che esiste la derivata parziale rispetto a t. ∎

Attenzione: In più variabili "differenziabile" è *più forte* di "derivabile": quest'ultimo concetto riguarda l'esistenza delle derivate parziali, e si dimostra con controesempi che esistono funzioni derivabili ma non differenziabili. Non ci soffermiamo, perché a noi interesseranno solamente funzioni differenziabili.

Interpretazione grafica: Il grafico di f_1 è il *piano tangente* in (t_0, u_0) al grafico della f, ossia alla superficie di equazione $z = f(t, u)$.
Definizione: L'espressione

$$df \stackrel{\text{def}}{=} f_1(t, u) - f(t_0, u_0) = f_t\,(t - t_0) + f_u\,(u - u_0)$$

si chiama *differenziale* di f. Si può scrivere: $df = f_t\,dt + f_u\,du$ e poi:

$$f(t, u) = f(t_0, u_0) + df + o(\varrho).$$

I differenziali dei fisici e quelli dei matematici

Chi abbia letto questo capitolo con sufficiente spirito critico avrà notato qua e là delle improprietà di linguaggio e anche discorsi un po' confusi, con qualche oscura avvertenza. Ciò dipende da una difficoltà di fondo, che ha radici storiche lontane, e che ora vogliamo guardare un po' più da vicino.

Il calcolo differenziale nasce nel '600 (Cavalieri, Fermat, Torricelli), e si sviluppa nel secolo successivo, prima di tutto come strumento della fisica. Le preoccupazioni di rigore sono inizialmente assai scarse, com'è naturale per un campo del tutto nuovo. La sistemazione accurata e la definizione precisa dei concetti sono opera dei matematici dell'800, e in parte non piccola sono proseguite anche in questo secolo. Nel frattempo i fisici continuavano a usare lo strumento matematico nella sua forma più intuitiva e utile (nel senso dello "strumento di pensiero," come abbiamo detto nell'introduzione, e non solo come puro strumento di calcolo) interessandosi ben poco delle esigenze di rigore. La progressiva specializzazione delle competenze, iniziata proprio nell'800 (ancora nel '700 gran parte dei maggiori matematici sono anche grandi fisici, e viceversa) ha portato a separare i problemi: di conseguenza la ricerca si è raramente proposta di cercare soluzioni che venissero incontro a tutte le necessità, sia quelle della costruzione matematica, sia quelle della ricerca fisica.

Il risultato è stato che la matematica ha elaborato una tecnica per trattare tutti i problemi in cui intervengono limiti (la cosiddetta "matematica dell'ε" di Weierstrass) che risolveva le esigenze di rigore, ma era assai poco agile e quindi inadatta a un impiego intuitivo come quello richiesto dai fisici; in contrapposizione, i fisici hanno proseguito sulla strada iniziata nel '700, che si potrebbe definire degli "infinitesimi attuali," interpretando i dx, tanto nelle espressioni differenziali quanto negli integrali, come degli incrementi infinitamente piccoli (allo scopo di sentirsi autorizzati a trascurare i residui di ordine superiore). Solo nel nostro secolo i matematici si sono riavvicinati a questa pratica, generalizzando il concetto di differenziale come funzione lineare dell'incremento della variabile indipendente, e trattando i dx negli integrali come misure (v. al Cap. 34). Un altro approccio moderno al problema degli infinitesimi è stata la cosiddetta "analisi non standard," che però non ha portato innovazioni pratiche nella didattica tradizionale.

Quello che abbiamo fatto in questo capitolo è stato un tentativo, necessariamente ibrido, di tenersi a cavallo dei due atteggiamenti del fisico e del matematico. Nel seguito però dovremo propendere per l'uso fisico; perciò dobbiamo ora trattare di certe regole pratiche per l'uso "disinvolto" dei differenziali.

Uso "disinvolto" dei differenziali

Se $f(t)$ è derivabile in t_0, sappiamo che si può confondere la variazione di f col differenziale, *a meno di termini di ordine superiore* nell'incremento di t. Perciò in tutti i nostri ragionamenti (calcoli, figure ...) ci sentiremo autorizzati a scrivere df, dx, $d\vartheta$, dt tutte le volte che vorremo sottintendere che le relazioni scritte sono valide *a meno di infinitesimi di ordine superiore.* La sola difficoltà è che bisogna prendere pratica con quello che s'intende, caso per caso, per "ordine superiore": la cosa migliore è vedere un esempio.

Esempio: Riprendiamo in esame il calcolo della velocità in coordinate polari, visto nel cap. precedente. Osserviamo la fig. 10–7, che riproduce la (9–9): il fatto che l'angolo è indicato con $d\vartheta$ ci dice che vogliamo trascurare termini di ordine superiore rispetto a quell'angolo, e perciò possiamo identificare la lunghezza della corda PP_1 con quella dell'arco, e la sua direzione con quella della tangente in P:

$$\overrightarrow{\mathrm{PP}}_1 = r\, d\vartheta\, \vec{e}_\vartheta.$$

Allo stesso modo, i due vettori $\overrightarrow{\mathrm{P_1P}}'$ e $\overrightarrow{\mathrm{PP}}_2$ hanno uguale lunghezza, ma direzioni che differiscono per $d\vartheta$: dunque la loro differenza è infinitesima come $d\vartheta$, e anche come dr. Poiché $d\vartheta$ e dr vanno entrambi a zero quando il punto P' si avvicina a P (il che è quanto dire che sono entrambi almeno $O(dt)$), ne segue che a meno di termini di secondo ordine

$$\overrightarrow{\mathrm{P_1P}}' = \overrightarrow{\mathrm{PP}}_2 = dr\, \vec{e}_r$$

e infine

$$d\vec{r} = dr\, \vec{e}_r + r\, d\vartheta\, \vec{e}_\vartheta.$$

Si vede come il ragionamento risulti snellito dall'uso dei differenziali.