Diario del Corso di Analisi - III Unità Didattica

Corsi di Laurea: Matematica, Fisica, Fisica Applicata
Docente: Sisto Baldo

ATTENZIONE: Il presente Diario del Corso vuole essere un riassunto abbastanza dettagliato di quello che è stato detto in aula, e come tale può essere un utile sussidio per chi voglia sistemare i propri appunti, o per chi sia stato assente e voglia ricostruire i contenuti di una lezione. D'altra parte, queste brevi paginette NON possono sostituire completamente un libro di testo, la lezione in aula o un'interazione diretta con il docente o l'esercitatrice: siete quindi invitati a servirvi ANCHE di queste altre opportunità per approfondire le vostre conoscenze!

Lezione del 30/4/2003 (2 ore): Alla fine della seconda unità didattica, c'eravamo posti il problema di enunciare un soddisfacente teorema di *esistenza e unicità locale* per le soluzioni del problema di Cauchy

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0. \end{cases}$$

In sostanza, vogliamo un risultato che ci dica che se la funzione f è *sufficientemente buona*, allora il problema ammette un'unica soluzione, e questa soluzione è definita *in un opportuno intorno* dell'istante iniziale t_0.

Un tipico risultato di questo tipo è il seguente:

TEOREMA (Di esistenza e unicità locale): Sia $f : [a, b] \times [c, d] \to \mathbf{R}$ una funzione di due variabili continua, per cui esiste una costante $L > 0$ tale che

$$|f(t, w) - f(t, z)| \le L|w - z| \quad \forall t \in [a, b],\ \forall w, z \in [c, d]$$

(in altre parole, f è lipschitziana nella seconda variabile, con una costante che non dipende dalla prima).

Siano poi $t_0 \in (a, b)$, $y_0 \in [c, d]$. Allora esiste $\delta > 0$ ed un'unica funzione derivabile $y(t) : [t_0 - \delta, t_0 + \delta] \to \mathbf{R}$ che risolve il problema di Cauchy

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0. \end{cases}$$

Di questo teorema, dimostreremo solo la parte relativa all'unicità: per dimostrare l'esistenza c'è bisogno di alcuni risultati (non difficili) sugli spazi funzionali che esulano dallo scopo di questo corso. Prima di passare alla dimostrazione dell'unicità, facciamo comunque qualche osservazione sulle ipotesi del teorema e sul suo significato.

OSSERVAZIONI:

- Per capire l'enunciato, ci occorre la definizione di funzione continua in due variabili. In realtà, essa non cambia molto rispetto all'analoga definizione in una variabile: se $(x_0, y_0) \in A \subset \mathbf{R}^2$ e $f : A \to \mathbf{R}$, diremo che f è continua in (x_0, y_0) se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che

  $$|f(x, y) - f(x_0, y_0)| < \varepsilon$$

  ogni volta che la distanza tra i punti $(x, y) \in A$ e (x_0, y_0) è minore di δ.

  Esattamente come in una variabile, si verifica facilmente che somma, prodotto, rapporto di funzioni continue è una funzione continua (se non si annulla il denominatore), che composizione di funzioni continue è continua, che i polinomi e le altre funzioni elementari sono continue. Inoltre, vale il teorema di Weierstrass: una funzione continua su un rettangolo chiuso e limitato ammette massimo.

  Torneremo comunque sul concetto di continuità in due variabili nel seguito del corso: ci accorgeremo che i limiti in più variabili possono essere decisamente difficili da calcolare...

- La condizione di lipschitzianità di f serve a garantire l'unicità della soluzione (per l'esistenza basterebbe la continuità): abbiamo visto nella precedente unità didattica che se f è soltanto continua, il problema di Cauchy può avere più di una soluzione.

  Fortunatamente, la lipschitzianità è implicata da codizioni più maneggevoli da verificare: per esempio, è sufficiente che esista la derivata parziale $\frac{\partial f}{\partial y}(t, y)$, e che essa sia una funzione continua in tutto il rettangolo $[a, b] \times [c, d]$. Infatti, in tali condizioni il teorema di Weierstrass ci assicura che la funzione continua $\left|\frac{\partial f}{\partial y}(t, y)\right|$ ammette massimo L sul rettangolo. Siano ora $t \in [a, b]$, $w, z \in [c, d]$. Usando il teorema di Lagrange (in una variabile) otteniamo:

  $$|f(t, w) - f(t, z)| = \left|\frac{\partial f}{\partial y}(t, c)\right| |w - z| \leq L|w - z|,$$

  dove c è un opportuno punto compreso tra w e z.

- Nelle ultime lezioni della seconda unità didattica, abbiamo anche visto che in generale non ci possiamo aspettare che la soluzione sia definita su tutto l'intervallo $[a, b]$: con le ipotesi generali che abbiamo scelto possiamo pretendere soltanto *esistenza locale*. Per avere *esistenza globale* ci sarà bisogno di avere qualche informazione supplementare sulla funzione f: per esempio, abbiamo visto che nel caso delle equazioni lineari si ha sempre esistenza globale (e si riesce anche a scrivere esplicitamente la soluzione!).

Per dimostrare il risultato di unicità avremo bisogno di un lemma:

LEMMA: Sia $y(t) : [a, b] \to \mathbf{R}$ una funzione derivabile. Supponiamo che $y(a) = 0$, e che esista una costante $L > 0$ tale che

$$|y'(t)| \leq L|y(t)|$$

per ogni $t \in [a, b]$. Allora $y(t) = 0$ per ogni $t \in [a, b]$.

DIM.: Sia $\delta > 0$, e denotiamo con M il massimo della funzione $|y(t)|$ sull'intervallo $[a, a + \delta]$. Se $t \in [a, a + \delta]$, grazie al teorema di Lagrange otteniamo

$$|y(t)| = |y(t) - y(a)| = |y'(c)| \, |t - a| \leq L|y(c)|\delta \leq LM\delta,$$

dove c è un punto opportuno tra a e t. Prendendo il massimo del membro di sinistra della nostra disugaglianza otteniamo $M(1-L\delta) \leq 0$: se scegliamo $\delta < 1/A$, quest'ultima stima implica che $M = 0$, cioè $y(t)$ si annulla identicamente tra a e $a + \delta$. Ripetendo questo stesso ragionamento, troveremo che $y(t)$ si annulla anche tra $a + \delta$ e $a + 2\delta$ ... e in un numero finito di passi otterremo la tesi. Q.E.D.

Si noti che il lemma vale, con dimostrazione identica, anche se si sostituisce l'ipotesi $y(a) = 0$ con $y(t_0) = 0$, con t_0 un punto qualunque dell'intervallo $[a, b]$.

DIMOSTRAZIONE DELL'UNICITÀ: Siano $y_1(t), y_2(t) : [t_0 - \delta, t_0 + \delta] \to \mathbf{R}$ due soluzioni del problema di Cauchy, e poniamo $z(t) = y_1(t) - y_2(t)$. Evidentemente, $z(t_0) = 0$. Inoltre:

$$|z'(t)| = |f(t, y_1(t)) - f(t, y_2(t))| \leq L|y_1(t) - y_2(t)| = L|z(t)|.$$

Applicando il Lemma, otteniamo che $z(t)$ si annulla identicamente nell'intervallo $[t_0 - \delta, t_0 + \delta]$, e il teorema è dimostrato. Q.E.D.

Lezione del 5/5/2003 (2 ore): Il risultato di esistenza e unicità locale dimostrato la volta scorsa ha un'importante interpretazione geometrica: se $f(t, y)$ è una funzione di due variabili continua e lipschitziana in $\mathbf{R}^2$, allora per ogni punto del piano ty passa una ed una sola soluzione dell'equazione differenziale $y'(t) = f(t, y(t))$.

Quest'osservazione apparentemente banale, consente spesso di effettuare uno studio qualitativo delle soluzioni di un'equazione differenziale ordinaria: in altre parole, è spesso possibile *disegnare con buona approssimazione l'andamento delle soluzioni senza dover risolvere esplicitamente l'equazione.* Come esempio di questo, in classe ci siamo divertiti a disegnare le soluzioni delle equazioni differenziali

$$y'(t) = ty(t), \qquad y'(t) = y(t)(\alpha - \beta y(t)).$$

Lo studio qualitativo delle soluzioni della seconda di queste equazioni (che si chiama *equazione logistica*, e fornisce un semplice modello di crescita di una colonia di batteri), è descritto in dettaglio nella dispensina[1] che ho scritto per una conferenza destinata agli alunni delle scuole superiori.

Passiamo ora a considerare brevemente che aspetto assume il problema di Cauchy per un *sistema di equazioni ordinarie del primo ordine*: per un sistema di n equazioni in n incognite $y_1(t), \ldots, y_n(t)$ si avrà una scrittura del tipo

$$\begin{cases} y_1'(t) = f_1(t, y_1(t), \ldots, y_n(t)) \\ \vdots \\ y_n'(t) = f_n(t, y_1(t), \ldots, y_n(t)) \\ y_1(t_0) = y_1^0 \\ \vdots \\ y_n(t_0) = y_n^0 \end{cases}$$

dove $f_1, \ldots, f_n$ sono n funzioni note di $n+1$ variabili, definite in un opportuno intorno rettangolare Q del dato iniziale $(t_0, y_1^0, \ldots, y_n^0)$.

Come spesso avviene in matematica, quando si studiano i sistemi di equazioni differenziali l'uso di una buona notazione produce dei benefici quasi miracolosi: innanzitutto, possiamo considerare le nostre n funzioni incognite come *un'unica funzione incognita a valori vettoriali*,

$$y(t) = (y_1(t), \ldots, y_n(t)).$$

(dunque, $y : I \to \mathbf{R}^n$, dove I è un intorno di t_0).

[1] http://degiorgi.science.unitn.it/baldo/aa2002/equadiff.pdf

Analogamente, possiamo ragruppare le n funzioni $f_1, \ldots, f_n$ in un'unica funzione vettoriale $f = (f_1, \ldots, f_n) : Q \to \mathbf{R}^n$, e altrettanto possiamo fare con i dati iniziali definendo $y^0 = (y_1^0, \ldots, y_n^0)$. Utilizzando questa notazione stenografica, il problema di Cauchy diventa

$$\begin{cases} y(t) = f(t, y(t)) \\ y(t_0) = y^0 \end{cases}$$

ossia formalmente identico a quello per un'unica equazione: l'unica differenza è che ora le quantità coinvolte vanno interpretate come vettori!

Questa fortissima somiglianza suggerisce che il Teorema di esistenza e unicità locale valga anche per i sistemi. Ecco infatti l'enunciato:
Se la funzione vettoriale f è continua in tutti i punti $(t, y) \in Q \subset \mathbf{R}^n$ e lipschitziana nella variabile $y \in \mathbf{R}^n$ (con costante indipendente da t), allora esiste un'unica soluzione $y(t)$ del problema di Cauchy per il sistema, definita in un intorno di t_0.

Ovviamente, dobbiamo precisare cosa s'intende per continuità e lipschitzianità di funzioni vettoriali. Per la continuità non c'è problema: una funzione vettoriale è continua se sono continue le sue componenti. Per quanto riguarda la lipschitzianità, chiediamo che valga una disuguaglianza del tipo

$$|f(t, y) - f(t, z)| \leq L|y - z|,$$

formalmente identica a quella che avevamo nel caso scalare: l'unica differenza è che i moduli vanno interpretati come moduli di un vettore. Ricordiamo che se $y = (y_1, \ldots, y_n)$, allora $|y| = (y_1^2 + y_2^2 + \ldots + y_n^2)^{1/2}$.

OSSERVAZIONE: Naturalmente non dimostreremo il teorema di esistenza locale per i sistemi, come non abbiamo dimostrato quello per le equazioni. I più volenterosi possono invece adattare la dimostrazione del teorema di unicità al caso dei sistemi: essa funziona in maniera quasi del tutto identica.

C'è un'unica ma importante differenza nella dimostrazione del Lemma: il teorema di Lagrange è infatti *falso* per le funzioni vettoriali. È vera però una sua versione un po' più debole, che è esattamente quel che ci serve nella dimostrazione del lemma (per il resto identica a quella che abbiamo visto nella scorsa lezione):
Sia $y(t) : [a, b] \to \mathbf{R}^n$ una funzione vettoriale derivabile sull'intervallo chiuso e limitato $[a, b]$. Allora

$$|y(b) - y(a)| \leq L \cdot (b - a),$$

dove L è il massimo di $|y'(t)|$ sull'intervallo $[a, b]$.[2]

La necessità di estendere il teorema di esistenza e unicità locale al caso dei sistemi non sorge da un peregrino desiderio di generalità, ma da un'effettiva necessità applicativa.

Prima di tutto, molti dei problemi della fisica sono intrinsecamente vettoriali: per determinare la traiettoria di un punto materiale, devo trovare le 3 componenti del vettore posizione in funzione del tempo.

Inoltre, le equazioni di ordine superiore al primo si possono scrivere in maniera equivalente come sistemi di equazioni del primo ordine: si capisce subito l'importanza di questo se si pensa che, in fisica, le equazioni che provengono da problemi di dinamica sono del secondo ordine!

Consideriamo infatti il problema di Cauchy per una generica equazione del secondo ordine:

$$(P) \quad \begin{cases} y''(t) = f(t, y(t), y'(t)) \\ y(t_0) = y_0 \\ y'(t_0) = v_0 \end{cases}$$

Introducendo l'incognita ausiliaria $v(t)$, il problema si può tradurre nel seguente (per un sistema di 2 equazioni del primo ordine):

$$(P') \quad \begin{cases} y'(t) = v(t) \\ v'(t) = f(t, y(t), v(t)) \\ y(t_0) = y_0 \\ v(t_0) = v_0 \end{cases}$$

I due problemi di Cauchy sono del tutto equivalenti: saper risolvere il secondo equivale a saper risolvere il primo, e viceversa.

D'altra parte, il teorema di esistenza e unicità locale per i sistemi del primo ordine ci dice che se f è continua, ed è anche lipschitziana nelle ultime due variabili, allora il problema (P) ammette una soluzione definita in un intorno di t_0, e che tale soluzione è unica.

Lezione del 7/5/2003 (2 ore): In questa lezione, cominceremo lo studio di una classe importante di equazioni differenziali del II ordine: le equazioni lineari.

[2]Dimostrazione:

$$|y(b) - y(a)| = |\int_a^b y'(t)\, dt| \leq \int_a^b |y'(t)|\, dt \leq L\,(b-a).$$

La generica equazione differenziale lineare del II ordine ha la seguente forma:

$$y''(t) + a(t)y'(t) + b(t)y(t) = f(t),$$

dove a, b, f sono funzioni continue note, definite su un intervallo I.

Evidentemente, a questo tipo di equazione si può applicare il discorso fatto la volta scorsa, e si scopre che per il problema di Cauchy associato vale il teorema di esistenza e unicità locale. Anzi, è possibile dimostrare che per le equazioni lineari si ha *esistenza globale*: se fissiamo le condizioni iniziali $y(t_0), y'(t_0)$, esiste un'unica soluzione dell'equazione, ed essa è definita su *tutto* l'intervallo I.

Un'equazione lineare si dice poi *omogenea* se $f(t) = 0$. Le soluzioni di un'equazione lineare omogenea del secondo ordine hanno una struttura particolarmente semplice:

TEOREMA: Siano $a, b : I \to \mathbf{R}$ due funzioni continue. Allora l'insieme $\mathcal{V}$ delle soluzioni dell'equazione lineare omogenea

$$(L)\ y''(t) + a(t)y'(t) + b(t)y(t) = 0$$

formano uno spazio vettoriale su $\mathbf{R}$ di dimensione 2. In altre parole:

(i) Se $y_1(t), y_2(t) : I \to \mathbf{R}$ sono due soluzioni e $c \in \mathbf{R}$, allora anche $y_1(t) + y_2(t)$ e $cy_1(t)$ sono soluzioni (questo è il "principio di sovrapposizione" dei fisici).

(ii) Se $y_1(t)$ e $y_2(t)$ sono due soluzioni che non siano multiple l'una dell'altra, esse formano una base di $\mathcal{V}$. In particolare, ogni altra soluzione dell'equazione sarà della forma

$$y(t) = C_1 y_1(t) + C_2 y_2(t),$$

con C_1 e C_2 costanti reali.

DIM.: Il punto (i) si verifica immediatamente sostituendo nell'equazione: è vero perché l'equazione (L) è lineare, e la derivata pure (la derivata della somma è la somma delle derivate...).

Il punto (ii) si ottiene invece come conseguenza del teorema di esistenza e unicità. Sia infatti $t_0 \in I$ e si considerino i seguenti problemi di Cauchy:

$$(P_1)\begin{cases} y''(t) + a(t)y'(t) + b(t)y(t) = 0 \\ y(t_0) = 1 \\ y'(t_0) = 0 \end{cases} \qquad (P_2)\begin{cases} y''(t) + a(t)y'(t) + b(t)y(t) = 0 \\ y(t_0) = 0 \\ y'(t_0) = 1 \end{cases}$$

Denotiamo $\tilde{y}_1(t)$ l'unica soluzione di (P_1), $\tilde{y}_2(t)$ l'unica soluzione di (P_2). Evidentemente, $\tilde{y}_1$ e $\tilde{y}_2$ non possono essere multiple l'una dell'altra, perché le condizioni iniziali non lo sono: quindi lo spazio delle soluzioni $\mathcal{V}$ ha almeno dimensione 2.

D'altra parte, affermo che $\tilde{y}_1$ e $\tilde{y}_2$ formano una base di $\mathcal{V}$. Sia infatti $y(t)$ una qualunque altra soluzione di (L), e definiamo $C_1 = y(t_0)$, $C_2 = y'(t_0)$. Evidentemente, $y(t)$ è l'unica soluzione del problema di Cauchy

$$\begin{cases} y''(t) + a(t)y'(t) + b(t)y(t) = 0 \\ y(t_0) = C_1 \\ y'(t_0) = C_2 \end{cases}$$

D'altra parte, per come sono state definite $\tilde{y}_1$ e $\tilde{y}_2$, si verifica subito che *anche la funzione* $C_1\tilde{y}_1(t)+C_2\tilde{y}_2(t)$ *è una soluzione dello stesso problema di Cauchy.* Per il teorema di unicità si deve dunque avere

$$y(t) = C_1\tilde{y}_1(t) + C_2\tilde{y}_2(t) \quad \forall t \in I,$$

e siamo riusciti a scrivere $y(t)$ come combinazione lineare degli elementi della base. Q.E.D.

Tra non molto, vedremo che anche l'insieme delle soluzioni di un'equazione lineare non omogenea ha una struttura piuttosto semplice.

Dal punto di vista pratico, il teorema ci dice che per risolvere l'equazione (L) basta saper trovare *due* soluzioni indipendenti della stessa. Non sempre questo è fattibile: è invece molto semplice quando l'equazione è a *coefficienti costanti*, cioè quando a, b sono numeri reali (indipendenti da t).

Se proviamo a risolvere un'equazione lineare ed omogenea *del primo ordine* a coefficienti costanti, vediamo che le soluzioni sono multiple di un esponenziale del tipo $y(t) = e^{kt}$. Questo suggerisce di cercare una soluzione dello stesso tipo anche per l'equazione del secondo ordine

$$(A)\ y''(t) + ay'(t) + by(t) = 0.$$

Sostituendo si trova che
l'esponenziale $y(t) = e^{kt}$ *è soluzione dell'equazione se e solo se* k *soddisfa l'equazione di secondo grado*

$$(B)\ k^2 + ak + b = 0,$$

detta equazione caratteristica.

Evidentemente, si possono presentare 3 casi:

1. *L'equazione caratteristica* (B) *ha due radici reali e distinte* k_1 *e* k_2. In queste condizioni, $e^{k_1 t}$ ed $e^{k_2 t}$ sono due soluzioni indipendenti dell'equazione differenziale (A), per cui il teorema ci assicura che *tutte* le soluzioni possono essere scritte come

$$y(t) = C_1 e^{k_1 t} + C_2 e^{k_2 t}, \quad C_1, C_2 \in \mathbf{R}.$$

2. *L'equazione caratteristica* (B) *ha due radici reali e coincidenti* $\overline{k}$. In tali condizioni, $y_1(t) = e^{\overline{k}t}$ è una soluzione, ma ce ne serve un'altra! Si osservi che il fatto che $\overline{k}$ sia una radice doppia implica che $a = -2\overline{k}$ e $b = \overline{k}^2$, per cui l'equazione (B) si scrive $y'' - 2\overline{k}y' + \overline{k}^2 y = 0$. Con una sostituzione diretta, si verifica che $y_2(t) = te^{\overline{k}t}$ è un'altra soluzione (indipendente dalla prima), per cui la soluzione generale dell'equazione differenziale data sarà

$$y(t) = C_1 e^{\overline{k}t} + C_2 t e^{\overline{k}t}, \quad C_1, C_2 \in \mathbf{R}.$$

3. *L'equazione caratteristica ha due radici complesse e coniugate* $\alpha \pm i\beta$. In questo caso, abbiamo due soluzioni *formali* $\tilde{y}_1(t) = e^{(\alpha+i\beta)t}$ e $\tilde{y}_2(t) = e^{(\alpha-i\beta)t}$, e non sappiamo bene cosa farcene...

   Per fortuna, è possibile dare un senso agli esponenziali complessi, e trasformare queste "inutili" soluzioni formali nella soluzione generale *reale* dell'equazione: scopriremo che tutte le soluzioni dell'equazione data sono del tipo

$$(*)\ y(t) = e^{\alpha t}(C_1 \cos \beta t + C_2 \sin \beta t), \quad C_1, C_2 \in \mathbf{R}.$$

   Definiamo l'esponenziale complesso: se $z \in \mathbf{C}$, poniamo

$$e^z := \sum_{n=0}^{+\infty} \frac{z^n}{n!}.$$

   Non è difficile verificare che la serie esponenziale converge per ogni $z \in \mathbf{C}$ (Come è ragionevole, diremo che $\lim_{n \to +\infty} (x_n + iy_n) = \overline{x} + i\overline{y}$ se $x_n \to \overline{x}$ e $y_n \to \overline{y}$ come successioni reali. A questo punto, diremo che una serie a valori complessi converge se esiste il limite delle sue somme parziali...). Si può anche far vedere che $e^{z_1+z_2} = e^{z_1}e^{z_2}$, per cui il nome "esponenziale complesso" è legittimo.

   Cosa succede se calcoliamo l'esponenziale complesso di un numero immaginario puro? Usando la definizione e ricordando le serie di Taylor di seno e coseno, non è affatto difficile verificare che vale la *formula di Eulero*

$$e^{i\beta} = \cos\beta + i\sin\beta.$$

In conclusione, avremo dunque

$$e^{\alpha+i\beta} = e^{\alpha}(\cos\beta + i\sin\beta).$$

[*In alternativa, si potrebbe prendere questa formula come definizione, per la verità un po' misteriosa, di esponenziale complesso. Si verifichi per esercizio che vale la proprietà dell'esponenziale richiamata sopra, cioè che si ha*

$$e^{(\alpha_1+i\beta_1)+(\alpha_2+i\beta_2)} = e^{\alpha_1+i\beta_1}\cdot e^{\alpha_2+i\beta_2}.]$$

A questo punto, non è più così sorprendente che le soluzioni complesse dell'equazione differenziale conducano alle soluzioni reali (*) scritte sopra: verificheremo questo fatto in dettaglio nella prossima lezione.

Lezione del 14/5/2003 (2 ore): La volta scorsa, era rimasto in sospeso il discorso sulle soluzioni delle equazioni lineari omogenee del secondo ordine a coefficienti costanti

$$(A)\ y'' + ay' + by = 0$$

nel caso in cui l'equazione caratteristica $k^2 + ak + b = 0$ abbia due radici complesse coniugate $k_{1,2} = \alpha \pm i\beta$.

Con la definizione di esponenziale complesso data nell'ultima lezione, abbiamo visto che possiamo scrivere $e^{(\alpha\pm i\beta)t} = e^{\alpha t}(\cos\beta t \pm i\sin\beta t)$. In che senso possiamo dire che queste sono soluzioni dell'equazione differenziale?

Se prendiamo la funzione $e^{i\beta t} = \cos(\beta t) + i\sin(\beta t)$ e la deriviamo (considerando formalmente i come una costante), troviamo

$$(e^{i\beta t})' = i\beta e^{i\beta t},$$

cioè la regola di derivazione per l'esponenziale immaginario coincide formalmente con quella per l'esponenziale reale. Più in generale (derivata del prodotto...), $(e^{(\alpha+i\beta)t})' = (\alpha+i\beta)e^{(\alpha+i\beta)t}$.

Dunque, le due funzioni $\tilde{y}_1(t) = e^{\alpha t}(\cos(\beta t) + i\sin(\beta t))$ e $\tilde{y}_2(t) = e^{\alpha t}(\cos(\beta t) - i\sin(\beta t))$ scritte sopra, sono soluzioni dell'equazione differenziale (A). Siccome vogliamo soluzioni *reali*, ci conviene prendere

$$y_1(t) = \frac{\tilde{y}_1(t) + \tilde{y}_2(t)}{2} = e^{\alpha t}\cos\beta t, \quad y_2(t) = \frac{\tilde{y}_1(t) - \tilde{y}_2(t)}{2i} = e^{\alpha t}\sin\beta t.$$

La soluzione generale dell'equazione nel caso delle radici complesse coniugate sarà dunque
$$y(t) = e^{\alpha t}(C_1 \cos(\beta t) + C_2 \sin(\beta t)).$$

Passiamo ora a considerare un'equazione lineare *completa*

$$(B)\ y'' + a(t)y'(t) + b(t)y(t) = f(t).$$

Come è fatto l'insieme delle soluzioni?

Grazie alla linearità dell'equazione, è immediato verificare che la *differenza di due soluzioni* è una soluzione dell'equazione omogenea (ottenuta prendendo $f(t) = 0$ nell'equazione (B)).

In altre parole, se siamo capaci di trovare una soluzione $y_0(t)$ dell'equazione completa (B), tutte le altre si ottengono sommando a $y_0(t)$ le soluzioni dell'equazione omogenea... Il professore di Geometria concluderebbe che l'insieme delle soluzioni dell'equazione (B) forma uno spazio affine su $\mathbf{R}$ di dimensione 2.

Visto che le soluzioni dell'equazione omogenea siamo capaci di trovarle, almeno nel caso delle equazioni a coefficienti costanti, si capisce come sia importante trovare un metodo generale per trovare una *soluzione particolare* dell'equazione (B).

Un metodo che *funziona sempre*, anche se spesso conduce a calcoli piuttosto complicati e antipatici, è il seguente:

TEOREMA (Metodo della variazione delle costanti): Data l'equazione (B), supponiamo di conoscere la soluzione generale $y(t) = Ay_1(t) + By_2(t)$ dell'equazione OMOGENEA associata. Allora si possono trovare due funzioni derivabili $A(t)$, $B(t)$ soluzioni del sistema

$$(C)\ \begin{cases} A'(t)y_1(t) + B'(t)y_2(t) = 0, \\ A'(t)y_1'(t) + B'(t)y_2'(t) = f(t). \end{cases}$$

Inoltre, le funzioni $A(t)$, $B(t)$ così trovate hanno la proprietà che

$$(D)\ y_0(t) = A(t)y_1(t) + B(t)y_2(t)$$

è una soluzione particolare dell'equazione non omogenea (B).

DIM.: Si arriva facilmente al sistema (C) cercando una soluzione dell'equazione completa del tipo (D): basta sostituire nell'equazione, imponendo arbitrariamente che valga la prima delle due equazioni del sistema. Si trova che l'equazione differenziale è soddisfatta se e solo se vale la seconda equazione di (C).

A questo punto, affermiamo che esistono due *uniche* funzioni continue $A'(t)$ e $B'(t)$ che soddisfano il sistema (C). Facendone l'integrale, troveremo così le funzioni richieste $A(t)$, $B(t)$.

Nel corso di geometria abbiamo imparato che il sistema (C) ammette una ed una sola soluzione se sappiamo che la matrice

$$\begin{pmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{pmatrix}$$

(nota come *matrice wronskiana*) è invertibile per ogni valore del parametro t (cioè se ha sempre determinante non nullo). D'altra parte, una matrice 2×2 è invertibile se e soltanto se le sue colonne *non* sono multiple l'una dell'altra, e sappiamo anche che le due funzioni $y_1(t)$ e $y_2(t)$ sono solzioni indipendenti dell'equazione omogenea.

Se per assurdo esistessero t_0 e $c \in \mathbf{R}$ tali che $y_1(t_0) = cy_2(t_0)$ e $y_1'(t_0) = cy_2'(t_0)$, il teorema di unicità per il problema di Cauchy ci direbbe che $y_1(t) = cy_2(t)$ *per ogni* t, contro l'indipendenza di y_1 e y_2.

Per convincersi poi che le soluzioni $A'(t)$, $B'(t)$ del sistema sono funzioni continue di t, basta pensare a come si scrive l'inversa di una matrice 2×2...
Q.E.D

Come esempio di applicazione del metodo, abbiamo considerato l'equazione

$$y'' - y = \frac{1}{1 + e^{2t}},$$

e abbiamo scoperto che una soluzione particolare è data da

$$y_0(t) = -\frac{1}{2} + \frac{e^t}{2} \arctan e^{-t} - \frac{e^{-t}}{2} \arctan e^t,$$

per cui la soluzione generale sarà

$$y(t) = Ae^t + Be^{-t} - \frac{1}{2} + \frac{e^t}{2} \arctan e^{-t} - \frac{e^{-t}}{2} \arctan e^t.$$

Fortunatamente per noi, in alcuni casi si può trovare una soluzione particolare di un'equazione lineare a coefficienti costanti *senza bisogno di usare il metodo della variazione delle costanti*: questo succede quando il termine noto $f(t)$ è il prodotto di un polinomio per un esponenziale (eventualmente complesso), cioè se

$$f(t) = e^{\alpha t} P(t) \quad \textit{con } P(t) \textit{ polinomio di grado } n,$$

(oppure

$$f(t) = e^{\alpha t}(a \cos \beta t + b \sin \beta t) P(t) \quad \textit{con } P(t) \textit{ come sopra.})$$

In questi casi *sappiamo a priori che esiste una soluzione particolare di forma "simile" a* $f(t)$, e precisamente:

1. Se α ($\alpha + i\beta$ nel secondo caso) non è una soluzione dell'equazione caratteristica, allora si cerca una soluzione del tipo $Q(t)e^{\alpha t}$ (nel secondo caso, $Q(t)e^{\alpha t}(c_1 \cos \beta t + c_2 \sin \beta t)$, dove Q è un polinomio di grado n.

2. Se α (oppure $\alpha + i\beta$) è una radice dell'equazione caratteristica, le "candidate soluzioni" scritte sopra vanno moltiplicate per t o per t^2, a seconda che si tratti di una radice semplice o di una radice doppia.

ESEMPIO: Se si considera l'equazione dell'*oscillatore armonico forzato*

$$y'' + k^2 y = C \sin \omega t,$$

troviamo che la soluzione generale dell'omogenea è $y(t) = A \cos kt + B \sin kt$. Il "ricettario" appena visto ci dice che nel caso in cui $\omega \neq \pm k$, troveremo una soluzione del tipo $y(t) = c_1 \cos \omega t + c_2 \sin \omega t$, mentre nel caso in cui $\omega = \pm k$ (sistema in risonanza), ci sarà una soluzione particolare del tipo $y(t) = t(c_1 \cos kt + c_2 \sin kt)$.

Lezione del 15/5/2003 (2 ore): Nella prima lezione di questa unità didattica abbiamo anticipato il concetto di *continuità per una funzione di due variabili*: una funzione $f : \mathbf{R}^2 \to \mathbf{R}$ si dice continua in (x_0, y_0) se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che

$$|f(x, y) - f(x_0, y_0)| < \varepsilon$$

per ogni (x, y) tale che $\operatorname{dist}((x, y), (x_0, y_0)) < \delta$.

Ricordiamo che si ha *per definizione*

$$\operatorname{dist}((x, y), (x_0, y_0)) = |(x, y) - (x_0, y_0)| = \sqrt{(x - x_0)^2 + (y - y_0)^2}.$$

Analogamente si definisce il limite di una funzione di due variabili:

DEFINIZIONE: Scriveremo che

$$\lim_{(x,y) \to (x_0, y_0)} f(x, y) = \ell$$

se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che $0 < |(x, y) - (x_0, y_0)| < \delta$ implica $|f(x, y) - \ell| < \varepsilon$.

In buona sostanza, la definizione di continuità (e quella di limite) per una funzione di due variabili è molto simile a quella per una funzione di una variabile. E' anche facile vedere che somme, prodotti, quozienti e composizioni di funzioni continue sono ancora funzioni continue (purché non si annulli il denominatore, nel caso dei quozienti...): in particolare, sono continue le funzioni elementari.

E dunque, tutto bene?

Non proprio: con le funzioni di due variabili possono accadere fatti abbastanza sorprendenti! Per esempio, potremmo pensare che se

$$\lim_{x \to x_0} f(x, y_0) = f(x_0, y_0), \quad \lim_{y \to y_0} f(x_0, y) = f(x_0, y_0),$$

allora f sia continua in (x_0, y_0). Nulla di più sbagliato: la funzione

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

ha questa proprietà ma *non è continua nell'origine* $(0, 0)$, perchè $f(x, x) = 1/2$ (e quindi f *non tende a* 0 se ci si avvicina all'origine lungo la bisettrice del primo e terzo quadrante). Siamo stati troppo ottimisti!

Ma certamente, se $g(x, y)$ tende a $g(x_0, y_0)$ quando ci si avvicina a (x_0, y_0) lungo *tutte le rette passanti per* (x_0, y_0), allora g sarà continua in quel punto?

Anche questo è clamorosamente falso: si consideri la funzione

$$(B)\ g(x, y) = \begin{cases} \frac{x^2 y}{x^4+y^2} & se\ (x, y) \neq (0, 0) \\ 0 & se\ (x, y) = (0, 0) \end{cases}$$

La restrizione di g a qualunque retta passante per l'origine tende a 0 quando ci si avvicina all'origine. In compenso, $g(x, x^2) = 1/2$, e la funzione *non* tende a 0 se ci si avvicina all'origine lungo la parabola $y = x^2$...

Questo ci dice che non è possibile testare la continuità di una funzione di due variabili limitandosi a vedere come si comporta lungo le rette, o lungo una fissata famiglia di curve: è proprio necessario ricorrere alla definizione.

Se questi problemi si presentano per la continuità, possiamo immaginare che siano ancora più macroscopici quando andiamo a studiare il calcolo differenziale per funzioni di due variabili!

In particolare, ci poniamo il seguente problema. Il grafico di una funzione di due variabili rappresenta graficamente una superficie nello spazio $\mathbf{R}^3$: sotto quali condizioni tale grafico è dotato di piano tangente in un certo punto? In altre parole, sotto quali condizioni potremo “approssimare bene” il grafico di una funzione attorno al punto (x_0, y_0) con un piano?

Quanto visto sopra per quanto riguarda la continuità ci suggerisce che probabilmente la sola esistenza delle derivate parziali non sarà sufficiente. Ricordiamo che le derivate parziali di f si definiscono nel modo seguente:

$$\begin{aligned}\frac{\partial f}{\partial x}(x_0, y_0) &= \lim_{h\to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}, \\ \frac{\partial f}{\partial y}(x_0, y_0) &= \lim_{h\to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h},\end{aligned}$$

purché i limiti esistano finiti. Si tratta cioè delle derivate delle restrizioni di f alle rette $y = y_0$ e $x = x_0$ rispettivamente.

Se riprendiamo la funzione f definita nell’equazione (A) sopra, vediamo che essa possiede derivate parziali nulle nell’origine (infatti la funzione si annulla identicamente su entrambi gli assi coordinati): questo ci dice che *la sola esistenza delle derivate parziali non garantisce nemmeno la continuità della funzione*, e tantomeno sarà sufficiente all’esistenza del piano tangente! (Le cose vanno ancor peggio con la funzione g definita sopra in (B): le restrizioni della funzione a qualunque retta passante per l’origine sono derivabili, ma questo non impedisce a g di essere discontinua.)

Per recuperare una decente regolarità, occorre invece definire il piano tangente al grafico di una funzione di due variabili come “piano di migliore approssimazione”, nella maniera seguente:

DEFINIZIONE: Il piano di equazione $z = f(x_0, y_0) + a\,(x - x_0) + b\,(y - y_0)$ si dice *piano tangente* al grafico di f nel punto (x_0, y_0) se e soltanto se

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

cioè se e soltanto se

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x, y) - f(x_0, y_0) - a\,(x - x_0) - b\,(y - y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0.$$

Se il grafico una funzione f ammette piano tangente nel punto (x_0, y_0), si dice che *f è differenziabile in* (x_0, y_0).

A riprova che questa è la definizione “giusta”, verifichiamo che una funzione differenziabile è continua ed anche derivabile parzialmente:

TEOREMA: Se f è differenziabile in (x_0, y_0), essa è anche continua nello stesso punto. f è poi derivabile parzialmente in (x_0, y_0), e l'equazione del piano tangente è data da

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

Se definiamo il vettore gradiente

$$\nabla f(x_0, y_0) := (\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0)),$$

l'equazione del piano tangente può anche essere scritta

$$z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0).$$

DIM.: Si ha

$$\begin{aligned} &\lim_{(x,y)\to(x_0,y_0)} [f(x, y) - f(x_0, y_0)] = \\ &\lim_{(x,y)\to(x_0,y_0)} [a\ (x - x_0) + b\ (y - y_0) + o(|(x - x_0, y - y_0)|)] = 0, \end{aligned}$$

per cui f è continua in (x_0, y_0). Inoltre, ponendo $x = x_0 + h$, $y = y_0$ nella definizione di piano tangente:

$$\frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = \frac{a\ h + o(h)}{h},$$

e passando al limite per $h \to 0$ si trova che f è derivabile parzialmente rispetto a x e $\frac{\partial f}{\partial x}(x_0, y_0) = a$.

In maniera del tutto analoga si mostra che $\frac{\partial f}{\partial y}(x_0, y_0) = b$. Q.E.D.

Evidentemente, la definizione di differenziabilità appena proposta non è molto agevole da verificare, in quanto occorre fare un limite in due variabili, e gli esempi visti sopra dovrebbero essere sufficienti a convincerci che questo non è sempre facile.

Per fare un esempio, in classe ci siamo divertiti (si fa per dire...) a verificare che la funzione $f(x, y) = x^2 - 3y^2$ è differenziabile in tutti i punti (x_0, y_0).... Questo esempio ci preoccupa un po': se è necessario riempire una lavagnata di conti per dimostrare che un polinomio di secondo grado è differenziabile, chissà cosa succederà con funzioni più complicate!

Fortunatamente, esiste un comodo teorema che dice che una funzione derivabile parzialmente in un intorno di (x_0, y_0) e con *derivate parziali continue*, è anche differenziabile in (x_0, y_0).

TEOREMA (Del differenziale totale): Sia $f : \mathbf{R}^2 \to \mathbf{R}$ una funzione derivabile parzialmente in un intorno del punto (x_0, y_0) (cioè in un rettangolo che contiene (x_0, y_0) al suo interno). Se le derivate parziali $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ sono entrambe continue nel punto (x_0, y_0), allora la funzione f è differenziabile nel punto (x_0, y_0).

ESEMPIO: Riprendiamo la funzione $f(x,y) = x^2 - 3y^2$, di cui avevamo verificato "a mano" la differenziabilità. Si ha

$$\frac{\partial f}{\partial x}(x,y) = 2x, \quad \frac{\partial f}{\partial y}(x,y) = -6y.$$

Le derivate parziali esistono ovunque, e sono funzioni continue di (x, y), per cui la funzione è ovunque differenziabile.

In maniera analoga, un qualunque *polinomio di due variabili* è ovunque differenziabile, perché le sue derivate parziali sono ancora polinomi e sappiamo che i polinomi sono funzioni continue... Utilizzando il teorema del differenziale totale, si può così mostrare facilmente che le funzioni elementari sono differenziabili.

Per avere un'idea "grafica" delle questioni trattate in questa lezione, potete consultare una mia pagina web sulle funzioni di più variabili[3] con animazioni interattive.

Lezione del 21/5/2003 (2 ore): Cominciamo col dare la dimostrazione del teorema del differenziale totale, enunciato alla fine della scorsa lezione.

DIMOSTRAZIONE del Teorema del Differenziale Totale: Sia (x, y) un punto sufficientemente vicino a (x_0), in modo che tutto il rettangolo di vertici (x_0, y_0), (x_0, y), (x, y) e (x, y_0) sia contenuto nell'insieme in cui esistono le derivate parziali di f.

Per mostrare che f è differenziabile in (x_0, y_0), dobbiamo far vedere che quando $(x, y) \to (x_0, y_0)$ il seguente rapporto tende a 0:

$$G(x,y) = \frac{f(x,y) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) - \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)}{\sqrt{(x-x_0)^2 + (y-y_0)^2}}$$

Il teorema di Lagrange per funzioni di una variabile mi assicura che esistono dei punti ξ (con ξ compreso tra x_0 ed x) ed η (con η compreso tra y_0 ed y) tali che

$$f(x,y_0) - f(x_0,y_0) \;=\; \frac{\partial f}{\partial x}(\xi,y_0)(x-x_0),$$

[3]http://degiorgi.science.unitn.it/~baldo/divulgazione/Graficiescript/Grafici3D.html

$$f(x,y) - f(x,y_0) \quad = \quad \frac{\partial f}{\partial y}(x,\eta)(y-y_0).$$

Nel rapporto $G(x,y)$, aggiungiamo e togliamo la quantità $f(x,y_0)$ a numeratore, ed applichiamo le due relazioni appena trovate. Possiamo così scrivere

$$\begin{aligned} G(x,y) \quad &= \quad \frac{x-x_0}{\sqrt{(x-x_0)^2+(y-y_0)^2}}\left[\frac{\partial f}{\partial x}(\xi,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)\right] + \\ & \quad\quad \frac{y-y_0}{\sqrt{(x-x_0)^2+(y-y_0)^2}}\left[\frac{\partial f}{\partial y}(x,\eta) - \frac{\partial f}{\partial y}(x_0,y_0)\right]. \end{aligned}$$

Ora, quando $(x,y) \to (x_0,y_0)$ avremo che $\xi \to x_0$ e $\eta \to y_0$, e grazie alla continuità delle derivate parziali in (x_0,y_0) le due espressioni tra parentesi quadre tenderanno a 0. Invece, i rapporti

$$\frac{x-x_0}{\sqrt{(x-x_0)^2+(y-y_0)^2}}, \quad \frac{y-y_0}{\sqrt{(x-x_0)^2+(y-y_0)^2}}$$

sono limitati: precisamente, il modulo di entrambi è minore o uguale a 1. Se ne conclude che

$$\lim_{(x,y)\to(x_0,y_0)} G(x,y) = 0.$$

Q.E.D.

OSSERVAZIONE: Il teorema del differenziale totale ci fornisce solo una condizione *sufficiente* di differenziabilità: una funzione può benissimo essere differenziabile anche se non sono soddisfatte le ipotesi. Ad esempio, la funzione

$$f(x,y) = \begin{cases} 0 & \textit{se } x \in \mathbf{Q} \textit{ oppure } y \in \mathbf{Q} \\ x^2+y^2 & \textit{altrimenti} \end{cases}$$

è differenziabile nell'origine (con piano tangente $z = 0$), ma non è neppure derivabile parzialmente in tutti gli altri punti.

Esattamente come accadeva in una variabile, l'analisi delle derivate parziali è un metodo potentissimo per cercare i *punti di massimo e minimo relativo per una funzione differenziabile di due variabili.* Una prima semplicissima osservazione ci assicura che in un punto di massimo o minimo relativo il piano tangente *deve essere orizzontale*:

OSSERVAZIONE: Se (x_0,y_0) è un punto di massimo o minimo relativo per la funzione differenziabile $f(x,y)$, allora

$$\frac{\partial f}{\partial x}(x_0,y_0) = 0, \quad \frac{\partial f}{\partial y}(x_0,y_0) = 0.$$

Infatti, il punto (x_0, y_0) sarà a maggior ragione un punto di massimo o minimo relativo per la restrizione di f alle rette $y = y_0$ e $x = x_0$, per cui le derivate di tali restrizioni (che sono poi le derivate parziali) si devono annullare nel punto stesso.

I punti in cui il piano tangente è orizzontale si chiamano *punti critici*. Evidentemente, questi non sono tutti punti di massimo o minimo relativo: ci possono essere *punti di sella* (punti corrispondenti a un "passo di montagna", se interpretiamo il grafico della funzione come rappresentazione grafica di un territorio montuoso...).

Cosìcome facevamo in una variabile, ci piacerebbe classificare i punti critici in base al segno della derivata seconda. Un piccolo problemino è dato però dal fatto che in un punto (x_0, y_0) di possibili derivate seconde ce ne sono ben quattro:

$$\frac{\partial^2 f}{\partial x^2}(x_0, y_0),\ \frac{\partial^2 f}{\partial y^2}(x_0, y_0),\ \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0),\ \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0).$$

Per comodità, esse si raggruppano di solito nella *matrice hessiana*

$$Hf(x_0, y_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

(in cui, ovviamente, tutte le derivate parziali sono calcolate in (x_0, y_0)).

Per nostra fortuna, se la funzione f è abbastanza buona le due derivate parziali miste *sono uguali* e la matrice hessiana è una matrice simmetrica: è il contenuto del seguente

TEOREMA (di Schwartz): Se f possiede entrambe le derivate seconde miste

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

in un intorno del punto (x_0, y_0), e queste due derivate sono continue in (x_0, y_0), allora

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0).$$

Omettiamo la dimostrazione del teorema di Schwartz (che comunque non è molto più difficile di quella del teorema del differenziale totale). Osserviamo però che il teorema *può essere falso* se non valgono le ipotesi: esistono esempi di funzioni che hanno le derivate seconde miste diverse!

Purtroppo questa non è una questione di lana caprina: la simmetria della matrice hessiana sarà una proprietà molto importante nel nostro studio dei punti critici[4].

Come per le funzioni di una variabile, lo studio dei punti critici mediante le derivate seconde è strettamente legato ad una *formula di Taylor di ordine 2 (con resto di Peano).* Per poterla dimostrare, ci serve una semplice regola di derivazione delle funzioni composte che è molto interessante anche in se stessa:

PROPOSIZIONE (Regola di derivazione delle funzioni composte): Sia $f(x,y)$ una funzione ovunque differenziabile, e siano $x(t)$, $y(t)$ due funzioni derivabili di una variabile. Allora la funzione composta $F(t) = f(x(t), y(t))$ è derivabile, e si ha

$$F'(t) = \frac{\partial f}{\partial x}(x(t), y(t))\ x'(t) + \frac{\partial f}{\partial y}(x(t), y(t))\ y'(t).$$

DIM: Per dimostrare che f è derivabile in t_0 e che vale la nostra espressione per la derivata, è sufficiente verificare che vale la formula di Taylor del primo ordine:

$$(A)\qquad \begin{aligned} &f(x(t), y(t)) = f(x(t_0), y(t_0)) + \\ &\left(\frac{\partial f}{\partial x}(x(t_0), y(t_0))\ x'(t_0) + \frac{\partial f}{\partial y}(x(t_0), y(t_0))\ y'(t_0)\right)(t - t_0) + o(t - t_0). \end{aligned}$$

A questo scopo, notiamo che la differenziabilità di f implica

$$(B)\qquad \begin{aligned} &f(x(t), y(t)) = f(x(t_0), y(t_0)) + \frac{\partial f}{\partial x}(x(t_0), y(t_0))\ (x(t) - x(t_0)) + \\ &\frac{\partial f}{\partial y}(x(t_0), y(t_0))\ (y(t) - y(t_0)) + o(|(x(t) - x(t_0), y(t) - y(t_0))|), \end{aligned}$$

mentre per le funzioni derivabili $x(t)$ e $y(t)$ abbiamo

$$x(t) - x(t_0) = x'(t_0)(t - t_0) + o(t - t_0), \quad y(t) - y(t_0) = y'(t_0)(t - t_0) + o(t - t_0).$$

[4] In effetti, abbiamo imparato nel corso di geometria che le matrici reali simmetriche sono diagonalizzabili, e in ultima analisi questo è esattamente quel che serve... Siccome però abbiamo solo due variabili, riusciremo a dimostrare il nostro risultato principale in modo elementare, senza ricorrere ai teoremi di algebra lineare che ci servirebbero in dimensione superiore.

Sostituendo queste due ultime relazioni in (B) si ottiene (A), a patto di osservare che $o(|(x(t) - x(t_0), y(t) - y(t_0))|) = o(t - t_0)$.[5] Q.E.D.

Lezione del 22/5/2003 (2 ore): Vediamo ora che aspetto assume il Teorema di Taylor per le funzioni di due variabili. Siccome questo sarà sufficiente per i nostri scopi, ci fermiamo al secondo ordine:

TEOREMA (di Taylor al II ordine): Sia $f(x, y)$ una funzione di due variabili, derivabile due volte con derivate prime e seconde continue in un intorno di (x_0, y_0). Allora, per tutti i punti (x, y) sufficientemente vicini a (x_0, y_0) si ha

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

DIM.: Sia (x, y) sufficientemente vicino a (x_0, y_0), in modo che tutto il segmento che congiunge i due punti sia contenuto nell'insieme in cui esistono le derivate prime e seconde (e sono continue).

Consideriamo la funzione (di una variabile) $F(t) = f(x_0 + t(x - x_0), y_0 + t(y - y_0))$: evidentemente, $F(0) = f(x_0, y_0)$ e $F(1) = f(x, y)$, mentre per valori intermedi di t la funzione f viene calcolata nei punti intermedi del segmento.

Applicando il teorema di derivazione della funzione composta si trova

$$F'(t) = f_x(\ldots)(x - x_0) + f_y(\ldots)(y - y_0),$$

dove l'argomento delle derivate (indicato con i puntini) è ovviamente $(x_0 + t(x - x_0), y_0 + t(y - y_0))$. Derivando ancora:

$$F''(t) = f_{xx}(\ldots)(x - x_0)^2 + 2f_{xy}(\ldots)(x - x_0)(y - y_0) + f_{yy}(\ldots)(y - y_0)^2$$

[5] Infatti, si verifica facilmente che $|(x(t) - x(t_0), y(t) - y(t_0))| \leq C|t - t_0|$ per t sufficientemente vicino a t_0. D'altra parte, possiamo sempre scrivere $o(s) = s\omega(s)$, con $\omega(s) \to 0$ per $s \to 0$. Avremo quindi

$$
\begin{aligned}
&o(|(x(t) - x(t_0), y(t) - y(t_0))|) = \\
&|(x(t) - x(t_0), y(t) - y(t_0))|\omega(|(x(t) - x(t_0), y(t) - y(t_0))|) \leq \\
&C|t - t_0|\omega(|(x(t) - x(t_0), y(t) - y(t_0))|),
\end{aligned}
$$

e quest'ultima quantità è $o(t - t_0)$.

(in cui i puntini hanno ancora lo stesso significato...).

La formula di Taylor per funzioni di una variabile, con resto di Lagrange, dice che è possibile trovare un punto $c \in [0, 1]$ tale che

$$F(1) = F(0) + F'(0) \cdot 1 + \frac{1}{2}F''(c) \cdot 1,$$

e sostituendo le espressioni che abbiamo trovato per le derivate otteniamo

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

dove $a = x_0+c(x-x_0)$, $b = x_0+c(x-x_0)$. Nell'espressione sopra, aggiungiamo e togliamo il polinomio di secondo grado

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

in modo da ottenere

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

dove

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

Ora, è immediato verificare che

$$\lim_{(x,y)\to(x_0,y_0)} \frac{R(x,y)}{(x-x_0)^2 + (y-y_0)^2} = 0.$$

Infatti, se $(x, y) \to (x_0, y_0)$ si ha anche $(a, b) \to (x_0, y_0)$ per cui le quantità tra parentesi quadre tendono a 0. Invece, i rapporti

$$\frac{(x-x_0)^2}{(x-x_0)^2 + (y-y_0)^2}, \quad \frac{2(x-x_0)(y-y_0)}{(x-x_0)^2 + (y-y_0)^2}, \quad \frac{(y-y_0)^2}{(x-x_0)^2 + (y-y_0)^2}$$

sono tutti minori o uguali ad 1 in modulo. Q.E.D.

Siamo ora in grado di enunciare un risultato che ci permette di determinare se un punto critico è di massimo o minimo relativo analizzando le derivate seconde:

TEOREMA: Sia $f(x, y)$ una funzione derivabile due volte in un intorno di (x_0, y_0), con derivate prime e seconde continue. Supponiamo anche che $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$ (cioè che il punto (x_0, y_0) sia un punto critico). Se il determinante della matrice hessiana è positivo,

$$\det Hf(x_0, y_0) = f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - (f_{xy}(x_0, y_0))^2 > 0,$$

allora (x_0, y_0) è un punto di minimo relativo o di massimo relativo a seconda che sia $f_{xx}(x_0, y_0) > 0$ o$f_{xx}(x_0, y_0) < 0$ rispettivamente.

Se viceversa (x_0, y_0) è un punto di minimo (risp. massimo) relativo per f, allora $\det Hf(x_0, y_0) \geq 0$ e $f_{xx}(x_0, y_0) \geq 0$ (risp. $f_{xx}(x_0, y_0) \leq 0$).

In particolare, un punto critico in cui $\det Hf(x_0, y_0) < 0$ non è né di massimo né di minimo relativo (punto di sella).

Per poter dimostrare il teorema, ci occorrerà la seguente semplice osservazione:

OSSERVAZIONE (Teorema di Weierstrass): Qualche lezione fa, avevamo detto che il teorema di Weierstrass in due variabili afferma che una funzione continua su un rettangolo chiuso e limitato ammette massimo e minimo. Esiste un enunciato un po' più generale, che è spesso utile quando si trattano le funzioni di più variabili:
Una funzione continua su un sottinsieme CHIUSO E LIMITATO di $\mathbf{R}^2$ ammette massimo e minimo.

Un sottinsieme del piano si dice chiuso se contiene tutti i suoi punti di accumulazione, mentre è limitato se esiste un rettangolo (con lati di lunghezza finita) che lo contiene.

DIMOSTRAZIONE DEL TEOREMA:
Siccome (x_0, y_0) è per ipotesi un punto critico, la formula di Taylor di ordine 2 centrata in (x_0, y_0) si riduce a

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

Denotiamo per brevità $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$ la distanza tra i punti (x,y) e (x_0,y_0). Con semplici passaggi algebrici la formula di Taylor diventa

$$
\begin{aligned}
(A) \qquad & f(x,y) - f(x_0,y_0) = \\
& \frac{1}{2}r^2\Big[f_{xx}(x_0,y_0)\frac{(x-x_0)^2}{r^2} + 2f_{xy}(x_0,y_0)\frac{(x-x_0)(y-y_0)}{r^2} + \\
& f_{yy}(x_0,y_0)\frac{(y-y_0)^2}{r^2} + \frac{o(r^2)}{r^2}\Big] = \\
& \frac{1}{2}r^2\Big[Q\left(\frac{x-x_0}{r}, \frac{y-y_0}{r}\right) + \frac{o(r^2)}{r^2}\Big],
\end{aligned}
$$

dove $Q(u,v)$ è il polinomio di secondo grado definito come segue:

$$Q(u,v) = f_{xx}(x_0,y_0)u^2 + 2f_{xy}(x_0,y_0)uv + f_{yy}(x_0,y_0)v^2.$$

Si noti anche che, nell'ultima espressione in (A), il polinomio $Q(u,v)$ è calcolato in un *punto sulla circonferenza unitaria* $u^2+v^2=1$.

Ora, il polinomio di secondo grado $Q(u,v)$ avrà segno costante (uguale al segno di $f_{xx}(x_0,y_0)$) se il suo discriminante è negativo: questa condizione è proprio

$$f_{xx}(x_0,y_0)f_{yy}(x_0,y_0) - (f_{xy}(x_0,y_0))^2 > 0.$$

Se questo è vero, $Q(u,v)$ non si annulla mai (tranne che nell'origine).

Supponiamo per fissare le idee che si abbia $f_{xx}(x_0,y_0) > 0$. In tal caso, il polinomio $Q(u,v)$ ammette un minimo m *strettamente positivo* sull'insieme chiuso e limitato $\{(u,v):\ u^2+v^2=1\}$ (la circonferenza unitaria nel piano). Partendo da (A) possiamo dunque scrivere:

$$f(x,y) - f(x_0,y_0) \geq \frac{1}{2}r^2\Big[m + \frac{o(r^2)}{r^2}\Big].$$

Per r abbastanza piccolo, la quantità tra parentesi quadre è positiva (perché il primo addendo è positivo, mentre il secondo tende a 0). Dunque, per r abbastanza piccolo si ha $f(x,y) - f(x_0,y_0) > 0$, e il punto (x_0,y_0) è di minimo relativo.

In maniera del tutto analoga, se $f_{xx}(x_0,y_0)f_{yy}(x_0,y_0) - (f_{xy}(x_0,y_0))^2 > 0$ e $f_{xx}(x_0,y_0) < 0$, si ha un massimo relativo.

Ci resta da dimostrare il viceversa: se (x_0,y_0) è un punto di minimo (massimo) relativo, dobbiamo mostrare che $\det Hf(x_0,y_0) \geq 0$ e $f_{xx}(x_0,y_0) \geq 0$ ($f_{xx}(x_0,y_0) \leq 0$). Questo lo faremo la prossima volta!

Lezione del 27/5/2003 (2 ore): Dobbiamo concludere la dimostrazione del teorema sui massimi e minimi relativi in più variabili, enunciato la volta scorsa. Ci resta da far vedere che *una funzione (derivabile con derivate prime e seconde continue) che abbia minimo (massimo) relativo in* (x_0, y_0)*, ha la proprietà che* $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$*,* $\det Hf(x_0, y_0) \geq 0$*,* $f_{xx}(x_0, y_0) \geq 0$ *(risp.,* $f_{xx}(x_0, y_0) \leq 0$*).*

Utilizzando la formula di Taylor del secondo ordine con resto di Peano, la volta scorsa abbiamo verificato che se (x_0, y_0) è un punto critico si può scrivere

$$(A) \quad f(x, y) - f(x_0, y_0) = \frac{1}{2} r^2 \Big[Q\left(\frac{x - x_0}{r}, \frac{y - y_0}{r}\right) + \frac{o(r^2)}{r^2}\Big],$$

dove r denota la distanza tra (x, y) e (x_0, y_0), mentre

$$Q(u, v) = f_{xx}(x_0, y_0)u^2 + 2f_{xy}(x_0, y_0)uv + f_{yy}(x_0, y_0)v^2.$$

Se (x_0, y_0) è un punto di minimo relativo, dico che il polinomio $Q(u, v)$ deve essere maggiore o uguale a zero sulla circonferenza unitaria: supponiamo infatti per assurdo che $Q(u, v)$ abbia minimo $m < 0$ nel punto (u_0, v_0) della circonferenza unitaria. Ponendo $x = x_0 + ru_0$, $y = y_0 + rv_0$ nell'equazione (A), e facendo tendere r a zero, troviamo che la quantità tra parentesi quadre tende a m, per cui essa è negativa per r abbastanza piccolo: questo è assurdo perché (x_0, y_0) è un punto di minimo relativo.

Dunque, il polinomio $Q(u, v)$ è non negativo sulla circonferenza. Questo, andando a "risolvere" la disequazione di secondo grado $Q(u, v) \geq 0$, è equivalente a dire che $\det Hf(x_0, y_0) \geq 0$ (discriminante ≤ 0) e $f_{xx}(x_0, y_0) \geq 0$ (primo coefficiente ≥ 0). Q.E.D.

Tutto l'armamentario dell'Hessiano non è veramente necessario se siamo interessati a trovare i massimi e minimi *assoluti* di una funzione di due variabili $F(x, y)$ in un insieme chiuso e limitato A sufficientemente "buono" nel piano $\mathbf{R}^2$.

Tipicamente, A sarà una regione del piano delimitata da una curva regolare, che spesso sarà data in forma implicita (ossia il bordo di A sarà l'insieme degli zeri di una funzione $f(x, y)$). Per fare un esempio concreto, potremmo avere

$$A = \{(x, y) \in \mathbf{R}^2 : \ x^2/4 + y^2 \leq 1\}.$$

In questo caso, l'insieme A è la regione di piano delimitata dall'ellisse $x^2/4 + y^2 - 1 = 0$.

Per il teorema di Weierstrass, sappiamo che F ha sia massimo che minimo in A: si tratta solo di individuarli!

In una sola variabile, è piuttosto semplice trovare i punti di massimo e minimo assoluto di una funzione derivabile su un intervallo chiuso e limitato: basta trovare i punti in cui si annulla la derivata, e confrontare poi i valori assunti dalla funzione in tali punti *e agli estremi dell'intervallo.*

Anche in due variabili, se un punto di massimo o minimo assoluto cade *all'interno di A* [6], il vettore gradiente si deve annullare (cioè un punto di massimo o minimo relativo interno è necessariamente un punto critico per la funzione F). Purtroppo, però, in questo caso abbiamo *infiniti* punti sulla frontiera ∂A di A [7], e non è così chiaro come possiamo confrontare il valore di F su *tutti questi punti*!

Se siamo in grado di parametrizzare esplicitamente il bordo di A (cioè se il bordo di A è l'immagine di una funzione vettoriale di due variabili, $t \mapsto (x(t), y(t))$), trovare il massimo e il minimo di F sul bordo di A corrisponde a trovare il massimo e il minimo della funzione di una variabile $F(x(t), y(t))$, cosa che sappiamo fare! Nell'esempio sopra, una parametrizzazione del bordo è per esempio $(2\cos t, \sin t)$ (con $t \in [0, 2\pi]$), per cui dovremmo massimizzare e minimizzare la funzione composta $F(2\cos t, \sin t)$ al variare di t tra 0 e 2π.

Che fare se invece la frontiera di A è l'insieme degli zeri di una funzione regolare di due variabili?

Per capirlo, sarà necessario studiare come sono fatte le *curve di livello* di una funzione di due variabili $f(x, y)$, cioè gli insiemi

$$S_C = \{(x, y) \in \mathbf{R}^2 : f(x, y) = C\},$$

dove C è una fissata costante.

Se siamo degli appassionati escursionisti, questi oggetti dovrebbero esserci familiari: immaginiamo che il grafico $z = f(x, y)$ della nostra funzione rappresenti il "plastico" di un territorio montuoso. Supponiamo anche che la funzione f sia differenziabile: abbiamo dei pendii, delle colline e degli avvallamenti, ma non sono presenti pareti rocciose verticali a strapiombo. Se un cartografo vuole disegnare una mappa che renda efficacemente l'idea del territiorio e della sua orografia, uno degli espedienti più utili cui può ricorrere consiste nel disegnare le curve che *uniscono punti ad uguale altitudine*, cioè appunto gli *insiemi di livello.*

Ora, la nostra esperienza di lettori di carte topografiche ci suggerisce che, tranne in pochi casi eccezionali, gli insiemi di livello siano *curve regolari*, ossia che essi *localmente coincidano con il grafico di una funzione derivabile (di x o di y).*

---

[6] Un punto interno di A è un punto con la proprietà che esiste un suo intorno rettangolare tutto contenuto in A.

[7] un punto di A si dice di frontiera se non è interno.

Se ci pensiamo un po' vedremo che i casi eccezionali si possono avere quando il piano tangente al grafico di f è orizzontale in alcuni punti dell'insieme di livello: per esempio, la cima di una collina può essere un punto isolato del suo insieme di livello, oppure si può avere una grossa zona piatta, che certo non è una curva. Inoltre, in un punto di sella l'insieme di livello può essere costituito da *due curve che si incrociano*, ed anche in questo caso non è molto regolare.

Quel che ci suggerisce la nostra intuizione è in effetti vero, ed è il contenuto del seguente teorema. In esso, ci occupiamo dell'insieme di livello di f con $C = 0$, cioè dell'insieme dei punti in cui la funzione *si annulla*. Evidentemente, questo non è restrittivo: l'insieme di livello C della funzione f non è altro che l'insieme degli zeri della funzione $g(x, y) = f(x, y) - C$.

TEOREMA (Delle funzioni implicite o del Dini): Sia $f(x, y)$ una funzione derivabile in un intorno di (x_0, y_0), con derivate continue in tale intorno. Supponiamo anche di sapere che $f(x_0, y_0) = 0$, e che $f_y(x_0, y_0) \neq 0$. Allora esistono $\delta > 0$ e $r > 0$ tali che

(i) Per ogni $x \in [x_0 - \delta, x_0 + \delta]$ esiste un unico $y \in [y_0 - r, y_0 + r]$ tale che $f(x, y) = 0$. Grazie all'unicità, possiamo anche scrivere $y = g(x)$.

(ii) La funzione $g : [x_0 - \delta, x_0 + \delta] \to \mathbf{R}$ è continua e derivabile, e si ha

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))}.$$

OSSERVAZIONE: Vediamo di tradurre l'enunciato. Sia

$$Z = \{(x, y) : \ f(x, y) = 0\}$$

l'insieme degli zeri di f. Il teorema dice che *dentro un rettangolo sufficientemente piccolo* $[x_0 - \delta, x_0 + \delta] \times [y_0 - r, y_0 + r]$, centrato nel punto (x_0, y_0) (che appartiene a Z per ipotesi), l'insieme degli zeri coincide con il grafico di una funzione derivabile g. Abbiamo inoltre una simpatica formuletta per calcolare la derivata di g.

Si noti che se invece di avere la condizione $f_y(x_0, y_0) \neq 0$, sapessimo che $f_x(x_0, y_0) \neq 0$, scambiando i ruoli delle due variabili potremmo dire che Z, in un rettangolino attorno a (x_0, y_0), coincide con il grafico di *una funzione della variabile y...* L'unico caso in cui il teorema non ci dice assolutamente nulla, è quello in cui *entrambe* le derivate si annullano: come abbiamo visto prima, questo è appunto il caso in cui l'insieme di livello può avere un aspetto orribile!

DIM. DEL TEOREMA: Supponiamo, per fissare le idee, che si abbia $f_y(x_0, y_0) > 0$. Siccome le due derivate parziali sono continue, per il teorema della permanenza del segno possiamo dire che f_y sarà strettamente positiva in un intero rettangolino centrato in (x_0, y_0): esiste $r > 0$ tale che $f_y(x, y) > 0$ per ogni $(x, y) \in [x_0 - r, x_0 + r] \times [y_0 - r, y_0 + r]$.

In particolare, la funzione di una variabile $y \mapsto f(x_0, y)$ è strettamente crescente sull'intervallo $[y_0 - r, y_0 + r]$. Siccome $f(x_0, y_0) = 0$, possiamo dedurne che

$$f(x_0, y_0 - r) < 0, \quad f(x_0, y_0 + r) > 0.$$

Ancora per il teorema della permanenza del segno, la funzione sarà negativa in un intorno di $(x_0, y_0 - r)$, e positiva in un intorno di $(x_0, y_0 + r)$: potremmo così trovare un numero $\delta > 0$ (minore di r) tale che

$$f(x, y_0 - r) < 0, \; f(x, y_0 + r) > 0 \quad \forall x \in [x_0 - \delta, x_0 + \delta].$$

Dunque, per ogni fissato $x \in [x_0 - \delta, x_0 + \delta]$ la funzione $y \mapsto f(x, y)$ è strettamente crescente su $[y_0 - r, y_0 + r]$ (per quanto detto sopra su f_y), è negativa nell'estremo sinistro dell'intervallo e positiva nell'estremo destro. Essa si annullerà dunque in un uno ed un sol punto compreso tra $y_0 - r$ e $y_0 + r$, che è esattamente la parte (i) dell'enunciato.

Il punto (ii) lo dimostreremo domani!

Lezione del 28/5/2003 (2 ore): Dimostriamo il punto (ii) del teorema delle funzioni implicite: per ogni $\overline{x} \in [x_0 - \delta, x_0 + \delta]$ abbiamo definito $g(\overline{x})$ come l'unico numero tra $y_0 - r$ e $y_0 + r$ tale che $f(\overline{x}, g(\overline{x})) = 0$. Se anche $x \in [x_0 - \delta, x_0 + \delta]$ abbiamo

$$\begin{aligned}(B) \qquad & 0 = f(x, g(x)) - f(\overline{x}, g(\overline{x})) = \\ & f(x, g(x)) - f(\overline{x}, g(x)) + f(\overline{x}, g(x)) - f(\overline{x}, g(\overline{x})) = \\ & f_x(\xi, g(x))(x - \overline{x}) + f_y(\overline{x}, \eta)(g(x) - g(\overline{x})),\end{aligned}$$

dove ξ è compreso tra x e $\overline{x}$, mentre η è compreso tra $g(x)$ e $g(\overline{x})$: l'esistenza di valori di ξ ed η che rendano vera l'uguaglianza è garantita dal teorema di Lagrange (in una variabile).

Dunque,

$$(C) \quad g(x) - g(\overline{x}) = -\frac{f_x(\xi, g(x))}{f_y(\overline{x}, \eta)}(x - \overline{x}).$$

La frazione nel membro di destra si mantiene limitata nel nostro rettangolino (possiamo maggiorarla con il massimo di $|f_x|$ diviso il minimo di $|f_y|$: quest'ultimo sarà strettamente positivo perché f_y non si annulla nel rettangolo). Dunque, facendo tendere $x \to \overline{x}$ si ha $g(x) \to g(\overline{x})$, e la funzione g è continua.

A questo punto, riprendiamo ancora una volta (C), e scriviamola nella forma
$$\frac{g(x)-g(\overline{x})}{x-\overline{x}} = -\frac{f_x(\xi, g(x))}{f_y(\overline{x}, \eta)}.$$
Facendo tendere $x \to \overline{x}$ si ottiene la formula per la derivata di g. Q.E.D.

OSSERVAZIONE: Può valer la pena di osservare il seguente fatto geometrico: se f e (x_0, y_0) sono come nel teorema, allora grad $f(x_0, y_0)$ è un vettore *ortogonale* all'insieme Z degli zeri di f (mostrarlo per esercizio). Siccome il vettore gradiente di una funzione punta sempre nella *direzione di massima crescita* della funzione stessa,[8] questo fatto non è sorprendente: infatti, sull'insieme Z la funzione non cresce affatto (è costante).

Il teorema delle funzioni implicite ha come conseguenza un'importante ricetta per *trovare i punti di massimo e minimo di una funzione regolare $F(x, y)$ sull'insieme Z degli zeri di un'altra funzione regolare $f(x, y)$* (tipicamente, Z potrà essere la frontiera di un insieme del piano su cui vogliamo trovare i massimi e i minimi assoluti di f).

TEOREMA (Dei moltiplicatori di Lagrange): Sia $f(x, y)$ una funzione derivabile con derivate parziali continue, e supponiamo che il suo insieme degli zeri
$$Z = \{(x, y) : \ f(x, y) = 0\}$$
abbia la proprietà che grad $f(x, y) \neq (0, 0)$ per ogni $(x, y) \in Z$. Sia $F(x, y)$ un'altra funzione di due variabili, derivabile con derivate continue, di cui siamo interessati a trovare i massimi e i minimi su Z.

Allora, se (x_0, y_0) è un punto di massimo o minimo relativo per la restrizione di F all'insieme Z, esiste un numero reale λ tale che
$$\begin{cases} F_x(x_0, y_0) = \lambda f_x(x_0, y_0) \\ F_y(x_0, y_0) = \lambda f_y(x_0, y_0) \end{cases}$$

[8] Se (u, v) è una direzione (cioè $u^2 + v^2 = 1$), si consideri la funzione $\phi : t \mapsto f(x_0 + tu, y_0 + tv)$, restrizione di f alla retta che passa per (x_0, y_0) ed è parallela alla direzione (u, v). Si vede subito che
$$\phi'(0) = f_x(x_0, y_0)u + f_y(x_0, y_0)v = \text{grad } f(x_0, y_0) \cdot (u, v),$$
ed evidentemente tale derivata è massima quando il vettore gradiente è allineato con la direzione (u, v).

OSSERVAZIONE: Dal punto di vista operativo, i massimi e i minimi di $F(x,y)$ sull'insieme Z sono da ricercare tra i punti (x,y) soluzioni del seguente sistema di *tre* equazioni nelle *tre incognite* x, y, λ:

$$(A)\quad \begin{cases} F_x(x_0,y_0) = \lambda f_x(x_0,y_0) \\ F_y(x_0,y_0) = \lambda f_y(x_0,y_0) \\ f(x,y) = 0. \end{cases}$$

Il numero λ si chiama *moltiplicatore di Lagrange.*

Geometricamente, risolvere il sistema (A) consiste nel cercare punti sulla curva Z in cui i vettori grad F e grad f *abbiano la stessa direzione.* Poiché sappiamo che il vettore grad f è sempre ortogonale a Z, stiamo dunque cercando i punti della curva in cui la funzione da minimizzare *cresce in direzione normale alla curva stessa* (che è come dire che, nella direzione tangente alla curva, la funzione F "ha derivata 0").

DIM. TEOREMA: Supponiamo per fissare le idee che sia $f_y(x_0,y_0) \neq 0$ (se a non annullarsi è l'altra derivata, basta scambiare il ruolo delle due variabili...).

Grazie al teorema delle funzioni implicite, esiste un intorno di (x_0,y_0) entro il quale l'insieme Z coincide con il grafico di una funzione derivabile $g(x)$. Allora, la funzione di 1 variabile $\phi : x \mapsto F(x, g(x))$ è definita in un intorno di x_0, ed ha un punto di massimo o minimo relativo in x_0, per cui $\phi'(x_0) = 0$.

D'altra parte, ricordando la formula per la derivata di g si ottiene subito:

$$0 = \phi'(x_0) = F_x(x_0,y_0) - F_y(x_0,y_0)\frac{f_x(x_0,y_0)}{f_y(x_0,y_0)}.$$

Basta allora porre

$$\lambda = \frac{F_y(x_0,y_0)}{f_y(x_0,y_0)},$$

e si ottiene la tesi. Q.E.D.

OSSERVAZIONE (data come esercizio da fare a casa...): L'ipotesi che il gradiente di f non si annulli mai sull'insieme Z è fondamentale. Si prenda per esempio $f(x,y) = (x-1)^3 - y^2$ e $F(x,y) = x^2 + y^2$. Vogliamo cioè trovare il minimo della funzione F (che rappresenta il quadrato della distanza dall'origine) sulla curva Z di equazione $f(x,y) = 0$... Tale minimo dovrà pur esistere!

Il metodo dei moltiplicatori di Lagrange ci suggerisce (suggerirebbe...) di cercare una soluzione del sistema

$$\begin{cases} 2x = 3\lambda(x-1)^2 \\ 2y = -2\lambda y \\ (x-1)^3 - y^2 = 0 \end{cases}$$

Dalla seconda equazione otteniamo $y = 0$ oppure $\lambda = -1$. Nel primo caso la terza equazione implica $x = 1$, ma questo è incompatibile con la prima equazione che diventa $2 = 0$... Se invece $\lambda = -1$, la prima equazione diventa un'equazione di secondo grado senza radici reali: il sistema dato non ha dunque alcuna soluzione.

D'altra parte, facendo un disegnino ci si rende subito conto che il minimo esiste, ed è raggiunto nel punto $(1, 0)$. Perché il metodo non ha funzionato? Semplicemente perchè grad $f(1, 0) = (0, 0)$, per cui il teorema non è affatto utilizzabile! D'altra parte, il nostro disegnino ci avrà mostrato che nel punto $(1, 0)$ la curva Z ha una cuspide...Si tratta di un punto in cui il luogo degli zeri non è una curva regolare!

Lezione del 4/6/2003 (2 ore): In questa lezione vogliamo rivisitare la nozione di *curva regolare* nel piano, e fare poi una breve carrellata sull'estensione dei risultati che abbiamo visto per funzioni di due variabili *in dimensione superiore*, cioè in n variabili.

CURVE PARAMETRICHE REGOLARI: La nostra idea "prototipo" di curva regolare nel piano, è il grafico di una funzione derivabile di una variabile. Abbiamo visto (Teorema delle funzioni implicite) che sotto certe ipotesi anche il *luogo di zeri di una funzione di due variabili* è una curva regolare, nel senso che esso è *localmente* esprimibile come grafico di funzioni di una variabile.

D'altra parte, la fisica ci ha abituato ad un altro modo di vedere le curve: possiamo pensare ad una funzione vettoriale di una variabile $\gamma : [a, b] \to \mathbf{R}^2$ ($\gamma(t) = (x(t), y(t))$), come alla *legge oraria* che descrive il moto di un punto materiale nel piano. Al variare di t, il nostro punto descriverà una curva nel piano, di cui la funzione γ è detta *parametrizzazione.*

Ora, sarà vero che se $\gamma(t)$ è una funzione derivabile (cioè le due componenti $x(t)$, $y(t)$ sono entrambe funzioni derivabili), allora la curva piana descritta dal punto è una curva regolare nel senso enunciato sopra?

La risposta è in generale negativa: si consideri per esempio la curva $\gamma(t) = (t^2, t^3)$, con $t \in [-1, 1]$. La curva piana descritta da $\gamma(t)$ ha una cuspide nell'origine: non è quindi una curva regolare, anche se γ è una funzione bellissima!

La condizione che garantisce che l'immagine di γ sia una funzione regolare è il non annullarsi del vettore velocità. Vale infatti il seguente risultato:

PROPOSIZIONE: Se $\gamma(t) = (x(t), y(t)) : [a, b] \to \mathbf{R}^2$ *è una funzione derivabile con continuità, e se* $\gamma'(t) = (x'(t), y'(t)) \neq 0$ *per ogni* $t \in [a, b]$*, allora l'immagine della funzione è una curva regolare nel piano, cioè un insieme localmente esprimibile come grafico di una funzione derivabile.*

DIM.: Sia $t_0 \in [a, b]$, e supponiamo per esempio che si abbia $x'(t_0) \neq 0$ (altrimenti $y'(t_0) \neq 0$, e basta scambiare i ruoli di x e y). Poniamo $x_0 = x(t_0)$.

Per il teorema della permanenza del segno, avremo $x'(t) \neq 0$ in un intorno di t_0, per cui in tale intorno la funzione $x(t)$ sarà strettamente crescente o strettamente decrescente. Esisterà allora la *funzione inversa* $g(x)$, derivabile, definita in un intorno I_{x_0} di x_0 ed a valori nel suddetto intorno di t_0. Calcolando la funzione composta $\gamma \circ g$ otteniamo:

$$\gamma(g(x)) = (x(g(x)), y(g(x))) = (x, y(g(x)).$$

Abbiamo così mostrato che l'immagine di γ, in un intorno di $\gamma(t_0)$, non è altro che il grafico $y = y(g(x))$ della funzione derivabile $y \circ g$. Q.E.D.

FUNZIONI DI PIU' VARIABILI: Se abbiamo una funzione di n variabili $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$, le nozioni di derivabilità parziale e di differenziabilità si introducono in maniera del tutto identica al caso delle due variabili. Anche risultati fondamentali come il teorema del differenziale totale ed il teorema di Schwartz, continuano a valere esattamente allo stesso modo.

Vediamo cosa diventano alcuni altri teoremi che abbiamo trattato in dettaglio nel caso delle due variabili.

TEOREMA DELLE FUNZIONI IMPLICITE IN 3 (O PIU') VARIABILI: Se abbiamo una funzione $f(x, y, z)$ derivabile con continuità, cosa possiamo aspettarci dal suo *luogo di zeri* $Z = \{(x, y, z) : \ f(x, y, z) = 0\}$?

Ebbene, se f soddisfa un'ipotesi simile a quella che avevamo in due variabili (precisamente, deve essere vero che grad $f(x, y, z) \neq (0, 0, 0)$ per ogni $(x, y, z) \in Z$), allora Z è una *superficie regolare*.

Più precisamente, Z si può esprimere localmente come grafico di una funzione regolare di *due delle tre variabili*: se per esempio $(x_0, y_0, z_0) \in Z$ e $f_z(x_0, y_0, z_0) \neq 0$, in un intorno di quel punto l'insieme Z coincide con il grafico $z = g(x, y)$ di una funzione differenziabile delle due variabili x, y.

Questo teorema vale, identico, anche nel caso di n variabili: in tal caso, l'insieme Z sarà un'ipersuperficie (cioè un oggetto esprimibile localmente come grafico di una funzione differenziabile di $(n - 1)$ variabili).

TEOREMA DEI MOLTIPLICATORI DI LAGRANGE: Avendo a disposizione il teorema delle funzioni implicite, è facile capire come si generalizzi anche il teorema dei moltiplicatori di Lagrange. Supponiamo che $Z = \{(x, y, z) : \ f(x, y, z) = 0\}$ sia il luogo di zeri di una funzione che soddisfa le ipotesi del teorema delle funzioni implicite. Siamo interessati a massimizzare o minimizzare la *restrizione a* Z di un'altra funzione derivabile con continuità, $F(x, y, z)$. I punti di massimo o minimo relativo vincolato devono allora soddisfare il *sistema dei moltiplicatori di Lagrange*

$$\begin{cases} F_x(x,y,z) = \lambda f_x(x,y,z) \\ F_y(x,y,z) = \lambda f_y(x,y,z) \\ F_z(x,y,z) = \lambda f_z(x,y,z) \\ f(x,y,z) = 0. \end{cases}$$

Le modifiche da fare se le variabili sono più di tre sono facilmente intuibili.

STUDIO DEI MASSIMI E MINIMI RELATIVI TRAMITE LA MATRICE HESSIANA: In questo paragrafo, vogliamo lanciare un messaggio di avvertimento sul fatto che per studiare i massimi e minimi relativi di una funzione di 3 o più variabili, non è più sufficiente guardare il segno del determinante della matrice hessiana e quello della derivata seconda rispetto alla prima variabile! La condizione è leggermente più complicata, e coinvolge il segno degli *autovalori* della matrice hessiana.

Ovviamente, continua a essere vero che in un punto di massimo e minimo relativo si annulla il gradiente della funzione, cioè il vettore

$$\text{grad}\, f(x_1, \dots, x_n) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right).$$

Vogliamo però capire come diventa la condizione sulla matrice hessiana che caratterizza i punti di massimo e minimo relativo. Nel caso di una funzione di n variabili, derivabile 2 volte con continuità, l'hessiana sarà la matrice quadrata simmetrica $n \times n$

$$Hf = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Per decidere se un punto critico è di massimo o minimo relativo, occorre andare a studiare il segno degli *autovalori della matrice hessiana.* Infatti, sappiamo dal corso di geometria che una matrice reale simmetrica ha autovalori reali. Ebbene, se la matrice hessiana ha gli autovalori *tutti positivi* in un punto critico, il punto è di minimo relativo, se gli autovalori sono tutti *negativi* il punto è di massimo, se infine esistono *sia autovalori positivi che negativi*, il punto è di sella:

TEOREMA: Sia f una funzione di n variabili derivabile due volte con continuità, e supponimo che $\mathbf{x_0}$ sia un punto in cui si annulla il gradiente di f. Allora

1. *Se la matrice hessiana $Hf(\mathbf{x_0})$ è definita positiva, cioè se tutti i suoi autovalori sono strettamente positivi, allora $\mathbf{x_0}$ è un punto di minimo relativo. Se invece $Hf(\mathbf{x_0})$ è definita negativa (tutti gli autovalori sono negativi), avremo un punto di massimo relativo.*

2. *Se $\mathbf{x_0}$ è un punto di minimo relativo (risp. massimo relativo), allora la matrice $Hf(\mathbf{x_0})$ è semidefinita positiva (risp. semidefinita negativa).*

3. *Se la matrice $Hf(\mathbf{x_0})$ ha sia autovalori positivi che autovalori negativi, allora $\mathbf{x_0}$ è un punto di sella.*

Dal punto di vista operativo, quindi, si tratta di studiare il *segno* degli autovalori della matrice hessiana. Dal corso di geometria, sappiamo che λ è un autovalore di H se e soltanto se

$$\det(\lambda I - H) = 0$$

(questo viene dal fatto che vogliamo avere soluzioni non nulle dell'equazione $H\mathbf{v} = \lambda\mathbf{v}$, per cui la matrice $(\lambda I - H)$ deve essere singolare).

Si tratta quindi semplicemente di scrivere l'equazione degli autovalori (un'equazione polinomiale di grado n), e di studiare il segno delle sue n radici reali (sappiamo che sono tutte reali perché la matrice è simmetrica)[9]...

Un utile esercizio consiste nel verificare che il teorema appena visto *si riduce ai criteri che abbiamo a suo tempo enunciato nel caso in cui le variabili siano solo due*!

A causa della prova di evacuazione della Facoltà, che ha interrotto la lezione di oggi, non abbiamo avuto il tempo di dimostrare il teorema (e, tutto sommato, forse non è stato un male!).

[9]Ci sono dei trucchi che consentono di vedere il segno delle radici anche se non si riesce a risolvere esplicitamente l'equazione: questo è estremamente utile perchè non sempre siamo in grado risolvere un'equazione polinomiale di grado elevato! Questi trucchi però esulano dallo scopo di questo corso

Se però qualcuno ha difficoltà a prendere sonno perché vuole vedere la dimostrazione, può trovarla qui sotto:

Dimostrazione del teorema (Facoltativa, sconsigliata a chi ha problemi di digestione): La formula di Taylor del secondo ordine in n variabili si può scrivere

$$
\begin{aligned}
&f(\mathbf{x}) = f(\mathbf{x}_0) + \\
&\text{grad } f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(Hf(\mathbf{x_0})(\mathbf{x} - \mathbf{x_0})) \cdot (\mathbf{x} - \mathbf{x_0}) + o(|\mathbf{x} - \mathbf{x_0}|^2),
\end{aligned}
$$

dove

$$
\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{x_0} = \begin{pmatrix} x_0^1 \\ \vdots \\ x_0^n \end{pmatrix},
$$

$\mathbf{a} \cdot \mathbf{b}$ denota l'ordinario prodotto scalare in $\mathbf{R}^n$, e con la scrittura $Hf(\mathbf{x_0})(\mathbf{x} - \mathbf{x_0})$ indichiamo il prodotto della matrice hessiana con il vettore $(\mathbf{x} - \mathbf{x_0})$.

In un *punto critico*, il termine della formula di Taylor che contiene il gradiente non c'è, e ci rimane soltanto

$$
\begin{aligned}
&f(\mathbf{x}) - f(\mathbf{x}_0) = \\
&\frac{1}{2}(Hf(\mathbf{x_0})(\mathbf{x} - \mathbf{x_0})) \cdot (\mathbf{x} - \mathbf{x_0}) + o(|\mathbf{x} - \mathbf{x_0}|^2) = \\
&\frac{1}{2}|\mathbf{x} - \mathbf{x_0}|^2 \left[\left(Hf(\mathbf{x_0}) \left(\frac{\mathbf{x} - \mathbf{x_0}}{|\mathbf{x} - \mathbf{x_0}|} \right) \right) \cdot \left(\frac{\mathbf{x} - \mathbf{x_0}}{|\mathbf{x} - \mathbf{x_0}|} \right) + \frac{o(|\mathbf{x} - \mathbf{x_0}|^2)}{|\mathbf{x} - \mathbf{x_0}|^2} \right].
\end{aligned}
$$

Esattamente come ci accadeva in due variabili, lo studio dei massimi e minimi relativi sarà allora strettamente legato al *segno* dei massimi e minimi del polinomio omogeneo di secondo grado

$$
Q(\mathbf{v}) = (Hf(\mathbf{x_0})\mathbf{v}) \cdot \mathbf{v} = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x_0}) v_i v_j
$$

sulla *sfera unitaria* $S = \{\mathbf{v} : \ v_1^2 + v_2^2 + \ldots v_n^2 = 1\}$ (si noti che la sfera unitaria è un sottinsieme chiuso e limitato di $\mathbf{R}^n$, per cui il massimo ed il minimo esistono!).

Il nostro enunciato si può allora ricavare immediatamente (con la stessa dimostrazione che abbiamo dato in due variabili) grazie al risultato seguente:

PROPOSIZIONE: Il massimo e il minimo della funzione $Q(\mathbf{v})$ sulla sfera unitaria S non sono altro che il massimo ed il minimo degli autovalori della matrice simmetrica $Hf(\mathbf{x_0})$.

DIM.: Ricordiamo che un numero λ è un autovalore della matrice A se e solo se esiste un vettore non nullo $\mathbf{v}$ tale che $A\mathbf{v} = \lambda\mathbf{v}$. Per brevità, indichiamo con H la matrice hessiana $Hf(\mathbf{x_0})$, e con h_{ij} il suo elemento di posto ij (e quindi $h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x_0})$).

Consideriamo la funzione $g(\mathbf{v}) = Q(\mathbf{v})/|\mathbf{v}|^2$. Si tratta di una funzione definita in tutti i punti diversi dall'origine, e differenziabile nel suo insieme di definizione. Questa funzione è costante sulle semirette uscenti dall'origine e coincide con Q sulla sfera unitaria: in particolare, i massimi e minimi (assoluti) di Q sulla sfera sono anche massimi e minimi per g, e viceversa.

Quindi, se $\mathbf{v_0}$ è per esempio un punto di minimo assoluto per Q su S, dobbiamo avere grad $g(\mathbf{v_0}) = 0$.

Ora, esplicitando la definizione di g si ottiene

$$g(\mathbf{v}) = \frac{\sum_{ij} h_{ij} v_i v_j}{v_1^2 + v_2^2 + \dots v_n^2},$$

da cui si ricava facilmente

$$\frac{\partial g}{\partial v_k}(\mathbf{v}) = 2\frac{\sum_i h_{ik} v_i - g(\mathbf{v}) v_k}{|\mathbf{v}|^2}.$$

Dunque, la condizione di minimalità grad $g(\mathbf{v_0}) = 0$ diventa $H\mathbf{v_0} = g(\mathbf{v_0})\mathbf{v_0} = Q(\mathbf{v_0})\mathbf{v_0}$, cioè *il valore minimo $Q(\mathbf{v_0})$ di Q su S è un autovalore per la matrice hessiana!*.

D'altra parte, se λ è un autovalore di H con autovettore $\mathbf{v}_\lambda$ di norma 1, si ha $Q(\mathbf{v}_\lambda) = (H\mathbf{v}_\lambda) \cdot \mathbf{v}_\lambda = \lambda |\mathbf{v}_\lambda|^2 = \lambda$, per cui *ogni* autovalore di H coincide con un valore assunto da Q sulla sfera unitaria. Se ne deduce che $Q(\mathbf{v_0})$ deve essere proprio il *minimo* degli autovalori.

In maniera del tutto analoga, il massimo di Q su S sarà il massimo autovalore della matrice H. Q.E.D.

Lezione del 11/6/2003 (2 ore): Come ultimo argomento di questo corso, proponiamo un'introduzione elementare alla teoria delle serie di Fourier.

Il problema è in un certo senso analogo a quello delle serie di Taylor: in quel caso si voleva approssimare una funzione regolare con polinomi, mentre ora siamo interessati all'approssimazione di una funzione *periodica* con *polinomi trigonometrici*.

Ricordiamo che una funzione $f : \mathbf{R} \to \mathbf{R}$ è T-periodica se vale che $f(x + T) = f(x)$ per ogni $x \in \mathbf{R}$. Il *periodo* di f è allora il minimo valore di T per cui vale questa relazione.

Evidentemente, data una funzione periodica f non è restrittivo supporre che il suo periodo sia 2π: basta eventualmente comporre con il cambio di variabili $x \mapsto \frac{T}{2\pi}x$. Ora, le più semplici funzioni 2π-periodiche sono le funzioni $\sin x$ e $\cos x$: si tratta di funzioni diffusissime "in natura" (si pensi agli oscillatori armonici, ai circuiti LC, alla proiezione di un moto circolare uniforme sugli assi...). Per ottenere funzioni 2π-periodiche di forma "più complicata", possiamo divertirci a sommare alle funzioni seno e coseno altre oscillazioni sinusoidali di frequenza multipla: otteniamo così i *polinomi trigonometrici*:

DEFINIZIONE: Un *polinomio trigonometrico* è una funzione del tipo

$$\frac{a_0}{2} + \sum_{n=0}^{N} [a_n \cos nx + b_n \sin nx],$$

dove $a_0, \ldots, a_n, b_1, \ldots b_n$ sono coefficienti reali.

L'idea di base delle serie di Fourier è che *ogni funzione sufficientemente regolare* si può scrivere come una serie, la cui somma parziale n-esime è un polinomio trigonometrico di "grado" n (per ogni n): in altre parole, data una "decente" funzione periodica di periodo 2π, $f(x)$, vorremmo poter scrivere

$$(*) \quad f(x) = \frac{a_0}{2} + \sum_{n=0}^{+\infty} [a_n \cos nx + b_n \sin nx],$$

con una scelta opportuna dei coefficienti a_n, b_n.

Non è difficile riuscire a "indovinare" come devono essere calcolati i coefficienti a_n e b_n affinché la $(*)$ sia (sperabilmente) vera. Cominciamo con l'osservare che valgono le seguenti *relazioni di ortogonalità*: se $m, n \in \mathbf{N}$

$$(***) \qquad \begin{aligned} \int_{-\pi}^{\pi} \cos nx \cos mx \; dx &= \pi \delta_{mn}, \\ \int_{-\pi}^{\pi} \sin nx \sin mx \; dx &= \pi \delta_{mn}, \\ \int_{-\pi}^{\pi} \sin nx \cos mx \; dx &= 0, \end{aligned}$$

dove δ_{mn} è il simbolo di Kronecker (vale 1 se $m = n$, 0 se $m \neq n$). Procediamo ora in maniera euristica. Supponiamo che valga la $(*)$, moltiplichiamo ambo i membri dell'uguaglianza per $\cos kx$ (o per $\sin kx$) e integriamo tra $-\pi$ e π: se la convergenza della serie è sufficientemente "buona", è ragionevole attendersi che si possa scambiare il segno di integrale con quello di serie. Dalle relazioni di ortogonalità troviamo allora facilmente che deve essere

$$(**) \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \; dx, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx \; dx, \; \forall k \in \mathbf{N}.$$

Si noti che per rendere rigoroso questo ragionamento sarebbe necessario possedere un criterio che ci dica quando è possibile scambiare i simboli di serie ed integrale... Il nostro punto di vista sarà però diverso: data una funzione f per cui sia possibile calcolare i coefficienti di Fourier a_k e b_k con le formule $(**)$, scriviamo la serie di Fourier $(*)$ *con quei coefficienti*, e ci chiediamo sotto quali condizioni essa converge, e converge proprio a $f(x)$.

Prima di enunciare un teorema di convergenza, spendiamo però due parole sull'estrema importanza applicativa delle serie di Fourier. In buona sostanza, esse ci permettono di scomporre un segnale periodico comunque complicato, nella somma di segnali sinusoidali di frequenza multipla di quella del segnale originale. Questo permette, per esempio, di costruire efficaci algoritmi di compressione di un segnale acustico (si pensi allo standard MP3), oppure di immagini (algoritmo JPEG: le immagini NON sono periodiche, ma una funzione definita su un intervallo può sempre essere prolungata periodicamente a tutta la retta reale...).

Fatto ancora più interessante, il nostro orecchio *sostanzialmente* non fa altro che calcolare i coefficienti di Fourier dei segnali acustici che gli arrivano: nella coclea (porzione dell'orecchio interno) ci sono dei gruppi di cellule specializzate, ciascuno dei quali è in grado di entrare in risonanza solo con un ristretto intervallo di frequenze... Utilizzando lo strumento teorico delle serie di Fourier, siamo quindi in grado di capire perché anche un bambino stonato è in grado di percepire un intervallo di ottava o di quinta: nella nota suonata da uno strumento musicale sono presenti (in varia misura, dipendente dal timbro dello strumento) anche le frequenze *multiple* di quella originale. Se suoniamo due note le cui frequenze stanno in rapporti semplici, vi sono "fin da subito" multipli comuni, per cui il nostro cervello è più che disposto a trovare gradevole l'accordo: si noti che la ragione è *fisiologica* e non semplicemente *culturale*...

Un tipico risultato di convergenza per le serie di Fourier è il seguente:

TEOREMA (Convergenza delle serie di Fourier): Sia $f : \mathbf{R} \to \mathbf{R}$ una funzione 2π-periodica, e supponiamo che nell'intervallo $[-\pi, \pi]$ la funzione abbia solo un numero finito di punti di discontinuità $x_1, \ldots, x_k$, in ciascuno dei quali esistono finiti il limite destro e sinistro di f. Supponiamo anche che la funzione sia derivabile con derivata continua in $[-\pi, \pi] \setminus \{x_1, \ldots, x_k\}$,con la possibile eccezione di in un numero finito di altri punti $y_1, \ldots, y_\ell$, e che in tutti i punti eccezionali $x_1, \ldots, x_k, y_1, \ldots, y_\ell$ esistano finiti il limite destro e sinistro di f'. Allora la serie di Fourier di f,

$$\frac{a_0}{2} + \sum_{n=0}^{+\infty} [a_n \cos nx + b_n \sin nx],$$

con i coefficienti a_n e b_n dati da $(**)$, converge a $f(x)$ in ogni punto di continuità di f. Nei punti $\{x_n\}$, la serie converge alla media aritmetica tra il limite destro e il limite sinistro di f.

OSSERVAZIONE: Le funzioni che soddisfano le ipotesi del teorema di convergenza si dicono brevemente *regolari a tratti*.

Si noti che le ipotesi che abbiamo fatto sono piuttosto forti: infatti che per *poter scrivere* i coefficienti di Fourier di f, è sufficiente che la funzione sia limitata ed integrabile secondo Riemann in $[-\pi,\pi]$. D'altra parte, il teorema comprende un'ipotesi sulla *derivata* di f...

In effetti, la sola continuità di f non sarebbe sufficiente a garantire la convergenza: esistono funzioni continue periodiche la cui serie di Fourier *non* converge in moltissimi punti, anche se è possibile (ma MOLTO difficile) dimostrare che vi è sempre convergenza in "quasi tutti" i punti (dove alla parola "quasi" si può dare un ben preciso significato matematico). D'altro canto, il teorema enunciato sopra mostra che vi è convergenza anche per funzioni *discontinue*, a patto di consentire solo discontinuità di salto e di aggiungere un'ipotesi analoga sulle discontinuità della derivata.

ESEMPIO: Consideriamo, sull'intervallo $[-\pi,\pi]$, la funzione $f(x) = |x|$, e la prolunghiamo per periodicità a tutta la retta reale. Otteniamo in questo modo una funzione pari il cui grafico è un'onda triangolare...

La funzione $f(x)$ così definita soddisfa evidentemente le ipotesi del teorema di convergenza. Utilizzando le formule per i coefficienti di Fourier troviamo subito che $b_n = 0$, mentre

$$a_n = \begin{cases} \pi & \text{se } n = 0, \\ 0 & \text{se } n \text{ è pari} > 0, \\ -\frac{4}{\pi n^2} & \text{se } n \text{ è dispari.} \end{cases}$$

di conseguenza possiamo scrivere

$$|x| = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2}\sin((2k+1)x) \quad \forall x \in [-\pi,\pi].$$

In particolare, usando questa identità per $x = 0$ abbiamo scoperto che la somma della serie dei reciproci dei quadrati dei numeri dispari è $\frac{\pi^2}{8}$, e da questo abbiamo poi ricavato che la somma dei reciproci dei quadrati di *tutti* i numeri naturali è $\frac{\pi^2}{6}$.

Lezione del 12/6/2003 (2 ore): Per poter dimostrare il teorema di convergenza, avremo bisogno di due lemmi. La dimostrazione del primo è lasciata come facile esercizio (si può fare per induzione su k, oppure scrivendo le funzioni trigonometriche con gli esponenziali complessi ed usando la formula per la somma della progressione geometrica).

LEMMA 1: Se $k = 1, 2, 3, \ldots$, vale la seguente identità

$$\frac{1}{2} + \cos\alpha + \cos 2\alpha + \cos 3\alpha + \ldots + \cos k\alpha = \frac{\sin((k+\frac{1}{2})\alpha)}{2\sin\frac{\alpha}{2}}.$$

LEMMA2 (Disuguaglianza di Bessel): Sia f una funzione 2π-periodica, limitata ed integrabile secondo Riemann sull'intervallo $[-\pi, \pi]$. Allora, indicati con a_k, b_k i coefficienti di Fourier di f si ha

$$\frac{a_0^2}{2} + \sum_{k=1}^{+\infty}(a_k^2 + b_k^2) \leq \frac{1}{\pi}\int_{-\pi}^{\pi}(f(t))^2\,dt.$$

In particolare, i coefficienti di Fourier a_k e b_k tendono a zero per $k \to +\infty$.

DIM.: Indichiamo con $S_N(x)$ la somma parziale N-esima della serie di Fourier di f:

$$S_N(x) = \frac{a_0}{2} + \sum_{n=1}^{N}[a_n \cos nx + b_n \sin nx].$$

Abbiamo:

$$(A) \qquad 0 \leq \frac{1}{\pi}\int_{-\pi}^{\pi}[f(x) - S_N(x)]^2\,dx =$$
$$\frac{1}{\pi}\int_{-\pi}^{\pi} f^2(x)\,dx + \frac{1}{\pi}\int_{-\pi}^{\pi} S_N^2(x)\,dx - \frac{2}{\pi}\int_{-\pi}^{\pi} f(x)S_N(x)\,dx.$$

Ricordando le relazioni di ortogonalità $(***)$ troviamo:

$$\frac{1}{\pi}\int_{-\pi}^{\pi} S_N^2(x)\,dx = \frac{a_0^2}{2} + \sum_{n=1}^{N}[a_n^2 + b_n^2]$$

(perché gli integrali dei doppi prodotti si annullano tutti...), e tenendo conto anche della definizione $(**)$ dei coefficienti di Fourier

$$-\frac{2}{\pi}\int_{-\pi}^{\pi} f(x)S_N(x)\,dx = -2\left(\frac{a_0^2}{2} + \sum_{n=1}^{N}[a_n^2 + b_n^2]\right).$$

Sostituendo queste due identità in (A) e facendo tendere $N \to +\infty$ si ha subito la disuguaglianza voluta.

Questa, in particolare, dice che le serie numeriche $\sum_{n=1}^{+\infty} a_n^2$ e $\sum_{n=1}^{+\infty} b_n^2$ sono entrambe convergenti. Da questo (per la condizione necessaria di convergenza di una serie) segue che i coefficienti di Fourier devono tendere a 0 per $n \to +\infty$. Q.E.D.

DIMOSTRAZIONE DEL TEOREMA DI CONVERGENZA DELLE SERIE DI FOURIER: Cominciamo a dimostrare il teorema nel caso in cui $f(x)$ è una funzione continua con derivata continua: vedremo poi come adattare la dimostrazione al caso generale.

Se denotiamo con S_N la somma parziale N-esima della serie, ricordando la definizione (**) dei coefficienti di Fourier si ottiene facilmente:

$$\begin{aligned} S_N(x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(u)[\frac{1}{2}+\sum_{n=1}^{N}\cos n(u-x)]\,du = \\ & \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+y)[\frac{1}{2}+\sum_{n=1}^{N}\cos n(y)]\,dy = \\ & \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+y)\frac{\sin(N+\frac{1}{2})y}{2\sin\frac{y}{2}}\,dy, \end{aligned}$$

dove abbiamo usato il cambio di variabili $y = u - x$ ed il Lemma 1.

Siccome si ha evidentemente

$$\frac{1}{\pi}\int_{-\pi}^{\pi}[\frac{1}{2}+\sum_{n=1}^{N}\cos n(y)]\,dy = 1,$$

possiamo scrivere

$$(B) \quad S_N(x) - f(x) = \frac{1}{\pi}\int_{-\pi}^{\pi}\frac{f(x+y)-f(x)}{2\sin\frac{y}{2}}\sin(N+\frac{1}{2})y\,dy.$$

Poniamo allora

$$g(y) = \frac{f(x+y)-f(x)}{2\sin\frac{y}{2}}.$$

Questa è una funzione continua su tutto l'intervallo di periodicità (in particolare, per $y \to 0$ la funzione tende a $f'(x)$), e la (B) diventa:

$$\begin{aligned} S_N(x) - f(x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} g(y)\sin(y/2)\cos Ny\,dy + \\ & \frac{1}{\pi}\int_{-\pi}^{\pi} g(y)\cos(y/2)\sin Ny\,dy. \end{aligned}$$

La disuguaglianza di Bessel ci dice infine che i due integrali nell'ultima formula tendono a 0 per $N \to +\infty$: si tratta infatti dei coefficienti di Fourier delle funzioni (continue) $g(y)\sin(y/2)$ e $g(y)\cos(y/2)$. Questo conclude la dimostrazione del caso particolare del teorema.

Mettiamoci ora nel caso generale. La nostra dimostrazione funziona già per ogni x, tranne che nei punti eccezionali $x_1, \dots, x_k, y_1, \dots, y_\ell$: se f soddisfa le ipotesi del teorema e x è un punto in cui la funzione è continua, derivabile e con derivata continua, la funzione $g(y)$ che abbiamo definito sopra non sarà continua, ma sarà comunque limitata ed integrabile secondo Riemann (perché possiede solo discontinuità di salto). Possiamo dunque applicare la disuguaglianza di Bessel esattamente come prima.

Viceversa, sia x uno dei punti eccezionali, e denotiamo con $f(x^+)$ e $f(x^-)$ i limiti destro e sinistro di f in x (i due limiti coincideranno se x è uno dei punti $y_1, \ldots, y_\ell$). In questo caso, dobbiamo mostrare che

$$S_N(x) \to \frac{f(x^+) + f(x^-)}{2}.$$

Siccome si ha

$$\frac{1}{\pi} \int_0^\pi [\frac{1}{2} + \sum_{n=1}^N \cos n(y)] \, dy = \frac{1}{2},$$

avremo

$$\begin{aligned} S_N(x) - \frac{f(x^+) + f(x^-)}{2} &= \frac{1}{\pi} \int_{-\pi}^0 \frac{f(x+y) - f(x^-)}{2 \sin \frac{y}{2}} \sin(N + \frac{1}{2})y \, dy + \\ &\quad \frac{1}{\pi} \int_0^\pi \frac{f(x+y) - f(x^+)}{2 \sin \frac{y}{2}} \sin(N + \frac{1}{2})y \, dy. \end{aligned}$$

Se guardiamo le frazioni nei due integrali come funzioni di y, esse hanno solo discontinuità di salto (nel punto $y = 0$ questo succede perché esistono finiti i limiti destro e sinistro di f' in x): esse sono dunque limitate e intergrabili secondo Riemann, e la dimostrazione può essere conclusa esattamente come prima. Q.E.D.