Diario del Corso di Analisi - II 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 22/2/2005 (2 ore): Nella prima unità didattica, abbiamo usato con qualche disinvoltura una proprietà della derivata che non abbiamo dimostrato. Infatti, abbiamo osservato che siccome la derivata corrisponde geometricamente alla "pendenza del grafico", se una funzione ha derivata positiva in un intervallo, essa sarà crescente in esso.

Questo fatto è intuitivamente molto plausibile, perché non si vede come una funzione possa avere il grafico "in salita" in tutti i punti di un intervallo senza essere anche crescente! D'altra parte, non ne abbiamo vista alcuna dimostrazione rigorosa: tale dimostrazione è lo scopo di questa prima lezione.

Cominciamo col ricordare la seguente definizione:

DEFINIZIONE: Sia $f : [a, b] \to \mathbf{R}$. Un punto $x_0 \in [a, b]$ si dice di *massimo relativo* (risp., di minimo relativo) per f se esiste un intorno I_{x_0} di x_0 tale che $f(x) \leq f(x_0)$ (risp., $f(x) \geq f(x_0)$) per ogni $x \in I_{x_0} \cap [a, b]$.

Veniamo ad un primo, semplice risultato: se una funzione è derivabile in un punto di *massimo o minimo relativo interno* all'intervallo di definizione, in quel punto la derivata si deve annullare:

TEOREMA (Principio di Fermat): Sia $f : [a, b] \to \mathbf{R}$, $x_0 \in (a, b)$ un punto di massimo o minimo relativo per f. Se f è derivabile in x_0, allora $f'(x_0) = 0$.

DIMOSTRAZIONE: Supponiamo per fissare le idee che x_0 sia di minimo relativo.

Consideriamo il rapporto incrementale per f in x_0:

$$\frac{f(x_0 + h) - f(x_0)}{h}.$$

Se prendiamo h abbastanza piccolo, in modo che x_0+h appartenga all'intorno I_{x_0} nella definizione di minimo relativo, vediamo subito che il numeratore è maggiore o uguale a 0. Ne consegue che il rapporto incrementale sarà positivo (o nullo) per $h > 0$ abbastanza piccolo, e negativo (o nullo) per $h < 0$ abbastanza piccolo in modulo. Ne segue che

$$f'(x_0) = \lim_{x \to x_0^+} \frac{f(x_0 + h) - f(x_0)}{h} \geq 0,$$

e contemporaneamente

$$f'(x_0) = \lim_{x \to x_0^-} \frac{f(x_0 + h) - f(x_0)}{h} \leq 0.$$

Dunque, $f'(x_0) = 0$, Q.E.D.

Si noti che questo teorema può essere falso per punti di massimo o minimo relativo che siano *agli estremi* dell'intervallo su cui f è definita. Per esempio, si consideri la funzione $f(x) = x$ sull'intervallo $[0, 1]$...

Due conseguenze di questo principio sono i teoremi di Rolle e di Lagrange, che sono tra i risultati più importanti del calcolo differenziale per funzioni di una variabile:

TEOREMA (di Rolle): Sia $f : [a, b] \to \mathbf{R}$ *una funzione continua, derivabile nell'intervallo* (a, b)*. Se* $f(a) = f(b)$*, allora esiste un punto* $c \in (a, b)$ *tale che* $f'(c) = 0$*.*

DIMOSTRAZIONE: Grazie al teorema di Weierstrass, la funzione possiede un punto di massimo assoluto x_M, e uno di minimo assoluto x_m.

Se uno di questi punti appartiene all'interno dell'intervallo, esso è anche di massimo o minimo relativo, e per il principio di Fermat la derivata si deve annullare in quel punto (che sarà dunque il punto c cercato).

In caso contrario, x_m e x_M coincidono con gli estremi a e b dell'intervallo. Allora, per ogni $x \in [a, b]$ si ha

$$f(a) = f(x_m) \leq f(x) \leq f(x_M) = f(b) = f(a),$$

per cui il massimo e il minimo di f coincidono. Ne segue che f è costante, e la sua derivata si annulla in *tutti* i punti dell'intervallo.

TEOREMA (di Lagrange): Sia $f : [a, b] \to \mathbf{R}$ *una funzione continua, derivabile nell'intervallo* (a, b)*. Esiste un punto* $c \in (a, b)$ *tale che* $f'(c) = \frac{f(b) - f(a)}{b - a}$*.*

DIMOSTRAZIONE: Consideriamo la funzione ausiliaria

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Questa è una funzione continua in $[a, b]$ e derivabile in (a, b). Inoltre $g(a) = f(a) = g(b)$, per cui possimo applicare il teorema di Rolle e ottenere un punto $c \in (a, b)$ tale che $g'(c) = 0$. Questo è proprio il punto cercato. Q.E.D.

A lezione, abbiamo visto che le ipotesi di questi due teoremi non possono essere in generale indebolite, e abbiamo discusso il significato geometrico di questi risultati.

Quel che è più importante, è comunque la seguente conseguenza del teorema di Lagrange:

COROLLARIO: Sia $f : [a, b] \to \mathbf{R}$ una funzione derivabile. Se $f'(x) \geq 0$ per ogni $x \in [a, b]$, allora f è crescente in $[a, b]$.

Se f' è invece minore o uguale a 0, la funzione è decresente in $[a, b]$. Infine, se $f' = 0$ in tutto l'intervallo, la funzione è costante.

DIMOSTRAZIONE: Facciamo vedere per esempio che vale la prima delle nostre affermazioni: supponiamo che $f'(x) \geq 0$ per ogni $x \in [a, b]$. Siano poi x_1 e x_2 due punti di $[a, b]$, con $x_1 < x_2$. Applichiamo il teorema di Lagrange a f sull'intervallo $[x_1, x_2]$: troviamo $c \in (x_1, x_2)$ tale che

$$f(x_2) - f(x_1) = (x_2 - x_1)f'(c).$$

Per l'ipotesi sulla derivata, il membro di destra è maggiore o uguale a zero. Q.E.D.

Come conseguenza del teorema di Lagrange, abbiamo anche ottenuto un risultato semplice e utile: sia f una funzione derivabile in un intorno I_{x_0} di x_0, tranne eventualmente in x_0. Se f è continua in x_0, e se esiste $\lim_{x \to x_0} f'(x) = \ell$, allora f è derivabile anche in x_0 e $f'(x_0) = \ell$.

Per ottenere questo risultato è sufficiente applicare il teorema di Lagrange al rapporto incrementale di f in x_0...

Lo stesso risultato si può ottenere anche grazie al Teorema di l'Hôpital, che è utilissimo quando si devono calcolare limiti nella forma indeterminata $0/0$ o ∞/∞:

TEOREMA (l'Hôpital): Siano f, g funzioni derivabili in un intorno di x_0, tranne eventualmente in x_0 (dove possono anche non essere definite). Supponiamo inoltre che $g'(x) \neq 0$ in tale intorno e che

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

oppure

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = \infty.$$

Se esiste (finito o infinito) il limite

$$\lim_{x\to x_0} \frac{f'(x)}{g'(x)},$$

allora esiste anche il limite

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

e questi due limiti sono uguali. L'enunciato del teorema rimane vero anche se $x_0 = \pm\infty$.

A lezione, abbiamo visto un esempio di applicazione di questo teorema. Attenzione: il limite del rapporto $f(x)/g(x)$ (in una delle due forme indeterminate suddette) può benissimo esistere anche se non esiste il limite del rapporto delle derivate. Un esempio è stato visto a esercitazione!

Non vedremo invece la dimostrazione del teorema.

Lezione del 24/2/2005 (1 ora): Nella lezione di oggi ci occuperemo del problema di approssimare una funzione regolare, in un intorno di un punto, mediante polinomi.

Supponiamo di avere una funzione f derivabile in x_0: se ci chiedessero qual è la retta che meglio approssima il grafico di f vicino a x_0, probabilmente risponderemmo tutti che è la retta tangente $y = f'(x_0)(x - x_0) + f(x_0)$: la cosa è ancora più plausibile se facciamo un disegno!

Vediamo però di precisare meglio (in maniera quantitativa) in che senso la retta tangente è quella che approssima meglio f in un intorno di x_0: se $g(x) = ax + b$ è un polinomio di primo grado che approssima f, possiamo scrivere che $f(x) = g(x) + R(x)$, dove $R(x)$ è un "resto" che vogliamo sia il più piccolo possibile quando x si avvicina a x_0.

Siccome $R(x) = f(x) - a(x - x_0) - b$, notiamo subito che

$$\lim_{x\to x_0} R(x) = 0 \Longleftrightarrow b = f(x_0).$$

In questo senso, tutte le rette che passano per il punto $(x_0, f(x_0)$ "approssimano f", nel senso che il resto tende a zero quando x si avvicina a x_0! Perché, dunque, la retta tangente è meglio delle altre?

Perché è l'unica per cui *il resto tende a zero più rapidamente di* $x - x_0$, cioè è l'unica per cui si abbia

$$\lim_{x\to x_0} \frac{R(x)}{(x - x_0)} = 0.$$

Infatti, siccome abbiamo già osservato che deve essere $b = f(x_0)$, la quantità di cui dobbiamo fare il limite diventa:

$$\frac{R(x)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} - a \to f'(x_0) - a.$$

Dunque, il limite è zero se $a = f'(x_0)$, mentre è non nullo in tutti gli altri casi.

Concludiamo dunque che la retta tangente è la retta di migliore approssimazione intorno a x_0, nel senso che è quella per cui *il resto tende a* 0 *più rapidamente quando* $x \to x_0$!

Nel tentativo di generalizzare quanto appena scoperto, diventa naturale chiedersi qual è il polinomio di grado n che meglio approssima una certa funzione f (che supporremo derivabile quante volte si vuole) in un intorno di x_0. Ci viene il sospetto che sia un polinomio simile alla retta tangente, nel senso che le sue derivate fino alla $n-esima$ nel punto x_0 dovranno coincidere con quelle di f...

Per semplificarci la vita, supponiamo che sia $x_0 = 0$: ci si può sempre ridurre a questa situazione con una traslazione lungo l'asse delle x.

Il *polinomio di Taylor di grado n per f centrato in* 0 è definito da

$$P_n(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k.$$

(Si noti che, ironia della sorte, il polinomio di Taylor "di grado n" ha in realtà grado *minore o uguale a n*...)

LEMMA 1: Sia f una funzione derivabile n volte in x_0. Tra tutti i polinomi $P(x)$ di grado minore o uguale a n, il polinomio di Taylor $P_n(x)$ è l'unico tale che $P(0) = f(x)$, $P'(0) = f'(0)$, $P''(0) = f''(0)$, ..., $P^{(n)}(0) = f^{(n)}(0)$.

DIM.: Il generico polinomio di grado minore o uguale a n sarà della forma $P(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$.

Ora, derivando h volte il monomio x^k si ottiene $\frac{k!}{(k-h)!}x^{k-h}$ se $h \leq k$, mentre si ottiene 0 per $h > k$. Tale funzione è sempre nulla in 0, tranne che nell'unico caso in cui $h = k$, quando vale $k!$.

Dunque, sostituendo nelle relazioni che abbiamo si ottiene $P(0) = a_0 = f(0)$, $P'(0) = a_1 = f'(0)$, $P''(0) = 2a_2 = f''(0)$, e in generale (per $k \leq n$) $P^{(k)}(0) = k!a_k = f^{(k)}(0)$. Di conseguenza, i coefficienti del polinomio devono essere proprio quelli che abbiamo attribuito al polinomio di Taylor di grado n. Q.E.D.

Lezione del 1/3/2005 (2 ore): Proseguiamo la nostra discussione sull'approssimazione di funzioni mediante polinomi. Ci servirà il seguente lemma: *LEMMA 2: Sia $g(x)$ una funzione derivabile $n-1$ volte in un intorno di*

0, con derivata n-esima in 0. Se $g(0) = g'(0) = g''(0) = \ldots = g^{(n)}(0) = 0$, allora abbiamo

$$\lim_{x\to 0} \frac{g(x)}{x^n} = 0.$$

DIM.: Basta applicare $n-1$ volte il teorema di L'Hôpital (grazie alle nostre ipotesi, ad ogni passo abbiamo una forma indeterminata $0/0$) ed infine la definizione di derivata come limite del rapporto incrementale:

$$\lim_{x\to 0} \frac{g(x)}{x^n} = \lim_{x\to 0} \frac{g'(x)}{nx^{n-1}} = \ldots = \lim_{x\to 0} \frac{g^{(n-1)}(x)}{n!x} =$$
$$\lim_{x\to 0} \frac{1}{n!} \frac{g^{(n-1)}(x) - g^{(n-1)}(0)}{x} = \frac{1}{n!} g^{(n)}(0) = 0.$$

Q.E.D.

Osservazione: Se g è come nel lemma, ma qualcuna delle derivate di ordine minore o uguale a n è diversa da 0, il limite *non può* essere 0: rifacendo lo stesso calcolo, si trova che è infinito, oppure è un numero diverso da zero.

Come corollario, otteniamo una prima forma del Teorema di Taylor:

TEOREMA (Di Taylor con resto di Peano): Sia f una funzione derivabile $n-1$ volte in un intorno di 0 con derivata n-esima in 0. Se $P_n(x)$ denota il polinomio di Taylor di grado n centrato in 0, allora

$$\lim_{x\to 0} \frac{f(x) - P_n(x)}{x^n} = 0.$$

In altre parole, abbiamo $f(x) = P_n(x) + o(x^n)$.

Tra tutti i polinomi di grado minore o uguale a n, il polinomio di Taylor è l'unico ad avere questa proprietà.

DIM.: Grazie al Lemma 1, la funzione $g(x) = f(x) - P_n(x)$ soddisfa le ipotesi del Lemma 2, e il teorema risulta dimostrato. L'unicità del polinomio di Taylor rispetto a questa proprietà segue dall'Osservazione fatta dopo la dimostrazione del Lemma 2. Q.E.D.

Osservazione: Grazie a quanto osservato nel Lemma2, il teorema rimane vero se chiediamo che f abbia $n-1$ derivate continue in un intorno di 0, e possieda la derivata n-esima in 0.

Osservazione: Se al posto di 0 prendiamo un generico punto x_0, il polinomio di Taylor di grado n centrato in x_0 sarà

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

In questo caso, il teorema di Taylor dice che se f è derivabile $(n-1)$ volte in un intorno di x_0, e possiede derivata n-esima in x_0, allora

$$f(x) = P_n(x) + R_n(x), \ con \quad \lim_{x \to x_0} \frac{R_n(x)}{(x - x_0)^n} = 0.$$

Applichiamo ora la formula di Taylor con resto di Peano allo studio dei massimi e dei minimi relativi di una funzione. L'idea è che spesso, per capire se un punto in cui la derivata prima si annulla è di massimo o minimo relativo, basta studiare il segno della derivata seconda:

TEOREMA: Sia f una funzione derivabile k volte in x_0 ($k \geq 2$), e supponiamo che le prime $k-1$ derivate esistano in un intorno di x_0. Supponiamo anche che $f^{(k)}(x_0) \neq 0$, mentre $f^{(h)}(x_0) = 0$ per $h = 1, \ldots, k-1$ (in altre parole, la prima derivata che non si annulla è la k-esima).

Allora

- *se k è dispari, x_0 non è né di massimo relativo né di minimo relativo;*
- *se k è pari e $f^{(k)}(x_0) > 0$, allora x_0 è un punto di minimo relativo;*
- *se k è pari e $f^{(k)}(x_0) < 0$, allora x_0 è un punto di massimo relativo.*

DIM.: Si tratta di studiare il segno della funzione $f(x) - f(x_0)$ quando x varia in un intorno sufficientemente piccolo di x_0: se tale funzione è positiva siamo in presenza di un punto di minimo relativo, se è negativa di un massimo (mentre se essa cambia di segno in ogni intorno, comunque piccolo, di x_0, il punto non è né di massimo né di minimo).

Se prendiamo x abbastanza vicino a x_0 in modo che valgano le ipotesi del teorema di Taylor, otteniamo

$$f(x) - f(x_0) = \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_k(x) = [\frac{f^{(k)}(x_0)}{k!} + \frac{R_k(x)}{(x - x_0)^k}](x - x_0)^k.$$

Siccome $R_k(x)/(x - x_0)^k \to 0$, la quantità tra parentesi quadre tende a $f^{(k)}(x_0)/k!$ per $x \to x_0$. Di conseguenza, essa *avrà lo stesso segno* di $f^{(k)}(x_0)$ quando x varia in un intorno sufficientemente piccolo di x_0. Poiché è invece chiaro che il polinomio $(x-x_0)^k$ è sempre positivo per k pari, mentre cambia di segno per k dispari (a seconda che x stia a destra o a sinistra di x_0), la tesi segue immediatamente. Q.E.D.

Osservazione: Il teorema appena dimostrato ci dice che se $f'(x_0) = 0$, $f''(x_0) > 0$, allora $f(x) > f(x_0)$ (con la disuguaglianza stretta) se x è sufficientemente vicino a x_0: in altre parole, il punto in questione *è un punto di minimo relativo stretto, e dunque non può di sicuro essere di massino relativo.*

Questo ci permette di invertire in parte il risultato precedente: se f è derivabile due volte in un intorno di x_0 e x_0 è di massimo relativo, allora $f'(x_0) = 0$ e $f''(x_0) \leq 0$ (La derivata prima si annulla per il principio di Fermat, mentre la seconda non può essere positiva perché se lo fosse il risultato precedente mi darebbe un punto di minimo relativo "stretto", il che sarebbe assurdo!). Analogamente, in un punto di minimo relativo deve essere $f'(x_0) = 0$ e $f''(x_0) \geq 0$.

In molti casi, il nostro teorema permette di stabilire se un punto in cui si annulla la derivata prima corrisponde ad un massimo o minimo relativo: basta trovare la prima derivata diversa da zero (a patto che la funzione sia derivabile abbastanza volte)! Può però capitare che la funzione sia derivabile infinite volte, ma *tutte* le derivate si annullino nel punto che ci interessa:

ESEMPIO: Si consideri la funzione

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

Con un po' di fatica si vede che f è derivabile infinite volte in 0, e che tutte le derivate si annullano (abbiamo visto la dimostrazione a lezione: il punto chiave è far vedere, per esempio per induzione, che per $x \neq 0$ si ha per ogni n

$$f^{(n)}(x) = \frac{g_n(x)}{x^{3n}} e^{-1/x^2},$$

dove $g_n(x)$ è un polinomio... A questo punto, basta passare al limite per $x \to 0$ per concludere che la derivata n-esima esiste ed è nulla nell'origine).

D'altra parte, siccome la funzione f è non negativa, è evidente che 0 è un punto di minimo relativo.

Conviene tenere presente la funzione f di questo esempio: tornerà utile tra non molto!

Lezione del 8/3/2005 (2 ore): Dimostriamo ora una versione leggermente più generale del teorema di Lagrange, che ci tornerà utile per trovare subito dopo un'espressione molto precisa del resto nella formula di Taylor.

TEOREMA (di Cauchy): Siano $f, g : [a, b] \to \mathbf{R}$ due funzioni continue, derivabili in (a, b). Supponiamo poi che $g'(x) \neq 0$ per ogni $x \in (a, b)$. Allora esiste $c \in (a, b)$ tale che

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

DIM.: Si noti che risulta $g(b) \neq g(a)$ (altrimenti otterremo un assurdo col teorema di Rolle): di conseguenza, il denominatore del secondo membro nella tesi non si annulla, e il teorema ha perfettamente senso.

Per dimostrarlo, basta applicare il teorema di Rolle alla funzione ausiliaria $h(x) = (f(x) - f(a))(g(b) - g(a)) - (g(x) - g(a))(f(b) - f(a))$. Q.E.D.

TEOREMA (Formula di Taylor con resto di Lagrange): Sia f una funzione derivabile $(n+1)$ volte in un intorno di x_0, e sia x un punto appartenente a tale intorno. Allora esiste un punto c, compreso tra x_0 e x, tale che

$$f(x) = P_n(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1},$$

dove

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

è il polinomio di Taylor di grado n centrato in x_0.

DIM.: Consideriamo la funzione $g(x) = f(x) - P_n(x)$. Abbiamo visto la volta scorsa che essa ha la fondamentale proprietà di avere $g(x_0) = g'(x_0) = \ldots = g^{(n)}(x_0) = 0$. Inoltre, $g^{(n+1)}(x) = f^{(n+1}(x)$ perché la derivata $(n+1)$-esima del polinomio P_n si annulla identicamente.

L'idea è ora di applicare $(n+1)$ volte il teorema di Cauchy, a partire dal rapporto

$$\frac{g(x)}{(x - x_0)^{n+1}} = \frac{g(x) - g(x_0)}{(x - x_0)^{n+1} - (x_0 - x_0)^{n+1}}.$$

Infatti, il teorema di Cauchy applicato a tale espressione ci garantisce l'esistenza di un punto c_1 compreso tra x_0 e x, tale che l'espressione stessa è uguale a

$$\frac{g'(c_1)}{(n+1)(c_1 - x_0)^n}.$$

L'ultima quantità può essere anche riscritta

$$\frac{g'(c_1) - g'(x_0)}{(n+1)(c_1 - x_0)^n - (n+1)(x_0 - x_0)^n},$$

e possiamo riapplicare il teorema di Cauchy... Ripetendo questo passaggio per $(n+1)$ volte troviamo:

$$\frac{g(x)}{(x-x_0)^{n+1}} = \ldots = \frac{g^{(n)}(c_n) - g^{(n)}(x_0)}{(n+1)!(c_n - x_0)} = \frac{g^{(n+1)}(c)}{(n+1)!} = \frac{f^{(n+1)}(c)}{(n+1)!},$$

dove $c_1, \ldots, c_n, c$ sono opportuni punti compresi tra x_0 e x. Q.E.D.

ESERCIZIO: Supponiamo che una funzione f soddisfi le ipotesi del teorema appena dimostrato, e che inoltre la derivata $f^{(n+1)}$ sia continua in x_0. Mostrare che allora

$$\lim_{x \to x_0} \frac{f(x) - P_{n+1}(x)}{(x-x_0)^{n+1}} = 0.$$

(In buona sostanza, avremo ridimostrato il Teorema di Taylor con resto di Peano, anche se con ipotesi leggermente più forti perché chiediamo la continuità della derivata $(n+1)$-esima...)

Mettiamo subito a frutto il teorema dimostrato, usandolo per approssimare alcune importanti funzioni con polinomi.

Cominciamo col prendere $f(x) = e^x$ (e $x_0 = 0$): la formula di Taylor con resto di Lagrange ci dice che esiste un punto c compreso tra 0 e x tale che

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{e^c}{(n+1)!} x^{n+1}.$$

Dimostriamo ora che il resto, qualunque sia $x \in \mathbf{R}$, tende a 0 per $n \to +\infty$:

$$\lim_{n \to +\infty} \frac{e^c}{(n+1)!} x^{n+1} = 0 \qquad \forall x \in \mathbf{R}.$$

Si noti che, anche se c dipende in generale sia da x che da n, la quantità e^c è maggiorata da $e^{|x|}$, per cui ci basterà far vedere che $x^{n+1}/(n+1)! \to 0$ per $n \to +\infty$. Questo fatto, a sua volta, si dimostra immediatamente ricordando la stima

$$n! \geq \left(\frac{n}{3}\right)^n,$$

che abbiamo provato per induzione.

In conclusione, abbiamo fatto vedere che

$$e^x = \lim_{n \to +\infty} \sum_{k=0}^{n} \frac{x^k}{k!}.$$

L'ultimo limite si chiama *serie di Taylor di* e^x, e si denota usualmente con $\sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$:

DEFINIZIONE: Se $\{a_k\}$ è una successione di numeri reali, col simbolo $\sum\limits_{k=0}^{\infty} a_k$ (che si legge *"serie degli a_k"*, si intende per definizione il limite

$$\lim_{n\to+\infty} \sum_{k=0}^{n} a_k,$$

a patto che tale limite esista. Se il limite è finito, si dice che la serie *converge*, se è infinito che *diverge*, se infine non esiste si dice che la serie *è indeterminata.*

Con lo stesso tipo di conti, abbiamo poi verificato che si ha anche

$$\begin{aligned} \sin x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \\ \cos x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \end{aligned}$$

per ogni $x \in \mathbf{R}$.

Si hanno analoghe serie di Taylor per molte delle funzioni elementari, alcune delle quali non convergono su tutto $\mathbf{R}$, ma solo su intervalli più piccoli: ad esempio si ha

$$\log(1+x) = \sum_{k=1}^{n} (-1)^{k+1} \frac{x^k}{k}$$

per $-1 < x \leq 1$ (non abbiamo però dato, per il momento, una dimostrazione della convergenza).

A questo punto, sorge spontanea la seguente questione: se f è una funzione derivabile infinite volte in x_0, è sempre possibile trovare un intorno di x_0 tale che

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

La risposta è negativa:

ESEMPIO: Si consideri la funzione

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

già studiata ieri in un altro contesto: abbiamo visto che essa è derivabile infinite volte in 0, e che tutte le derivate si annullano.

Ne segue che la serie di Taylor centrata in 0 per f è identicamente nulla, e quindi evidentemente essa coincide con f solo per $x = 0$.

Le funzioni la cui serie di Taylor converge alla funzione stessa in un intorno di x_0 si dicono *analitiche in* x_0: la funzione dell'esempio non è analitica in 0, mentre la funzione esponenziale, il seno ed il coseno lo sono.

Come ulteriore applicazione del teorema di Taylor con resto di Lagrange, dimostriamo che *il numero di Nepero e è irrazionale.*

Supponiamo infatti per assurdo che si abbia $e = p/q$, con p e q numeri naturali. Applichiamo il teorema di Taylor con resto di Lagrange alla funzione esponenziale, con $x_0 = 0$ e $x = 1$: per ogni $n \in \mathbf{N}$ troviamo un punto $c \in (0, 1)$ tale che

$$e = \frac{p}{q} = 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \ldots + \frac{1}{n!} + e^c \frac{1}{n+1!}.$$

Prendiamo $n > q$, e moltiplichiamo ambo i membri dell'identità per $n!$. Otteniamo:

$$\frac{p}{q}n! = \left(1 + 1 + \frac{1}{2} + \frac{1}{3!} + \ldots + \frac{1}{n!}\right) \cdot n! + \frac{e^c}{n+1}.$$

Il membro di sinistra dell'uguaglianza è evidentemente un intero, così come il primo pezzo del membro di destra (la quantità tra parentesi tonde moltiplicata per $n!$). Invece, l'ultimo termine si maggiora con $e/(n + 1)$ (perché $c < 1$), e quest'ultima quantità è strettamente minore di 1 per n abbastanza grande: questo è evidentemente assurdo (un intero non può essere uguale ad un intero più una quantità minore di 1).

Conclusi i nostri discorsetti sulla formula di Taylor, cominceremo a studiare una definizione rigorosa di area per certe figure piane delimitate da un contorno curvilineo. Questo ci porterà a dare la definizione di integrale nel senso di Riemann per una funzione $f : [a, b] \to \mathbf{R}$.

Partiamo da un esempio concreto: supponiamo di voler dare un senso all'area della regione limitata del piano delimitata dall'asse delle x, dal grafico della funzione $f(x) = e^x$, dall'asse delle y e dalla retta $x = a$: si tratta di una specie di trapezio rettangolo, solo che il lato obliquo è curvo (è infatti il grafico della funzione e^x.

Un'idea potrebbe essere la seguente: dividiamo l'intervallo $[0, a]$ in n intervallini uguali (che avranno quindi estremi $0, a/n, 2a/n, 3a/n, \ldots, a$). Per ciascuno di questi intervallini, costruiamo un rettangolo che ha l'intervallino stesso come base, e altezza uguale al valore della funzione esponenziale *nell'estremo di sinistra*: in altre parole, sull'intervallino $[ka/n, (k + 1)a/n]$ costruiamo un rettangolo di altezza $e^{ka/n}$. In questo modo otteniamo una figura “a scala” che è tutta *contenuta* nella figura curvilinea di cui vogliamo calcolare l'area: l'area di questa scala inscritta si calcola subito, e vale

$$\sum_{k=0}^{n-1} \frac{a}{n} e^{ka/n} = \frac{a}{n} \cdot \frac{e^a - 1}{e^{a/n} - 1}.$$

L'ultima espressione, per $n \to +\infty$, tende al numero $(e^a - 1)$, che rappresenta quindi il limite delle nostre approssimazioni per difetto.

In maniera analoga, possiamo costruire una “scala circoscritta” alla regione che ci interessa, prendendo su ciascun intervallino $[ka/n, (k+1)a/n]$ un rettangolo di altezza $e^{(k+1)a/n}$. In tal caso, l'area della scalinata sarà

$$e^{a/n} \frac{a}{n} \cdot \frac{e^a - 1}{e^{a/n} - 1},$$

che tende ancora al limite $e^a - 1$. Possiamo dunque legittimamente affermare che l'area del trapezoide curvilineo sotto la funzione esponenziale tra 0 e a, vale esattamente $e^a - 1$: tale numero si chiama *integrale* della funzione esponenziale tra 0 e a, e si indica

$$\int_0^a e^x \, dx = e^a - 1.$$

Volendo estendere quest'idea in modo sistematico al maggior numero possibile di funzioni, cominciamo col definire le *funzioni a scala* ed il loro integrale:

DEFINIZIONE: Una funzione $\phi : [a, b] \to \mathbf{R}$ si dice *a scala* se esiste una suddivisione di $[a, b]$ in un numero finito di intervallini di estremi $a = x_0 < x_1 < x_2 < \ldots < x_N = b$, in modo tale che ϕ assuma valore costante c_i su ciascun intervallino (x_i, x_{i+1}) (per $i = 0, \ldots, N-1$).

L'*integrale* della suddetta funzione a scala si definisce come

$$\int_a^b \phi(x) \, dx = \sum_{i=0}^{N-1} (x_{i+1} - x_i) c_i.$$

Come si vede, una funzione a scala è caratterizzata dal fatto che il suo grafico è un *istogramma*. Si noti che se ϕ è positiva, l'integrale della funzione a scala è semplicemente l'area dell'unione finita di rettangoli delimitata dal grafico di ϕ, l'asse delle x e le rette verticali $x = a$ e $x = b$. Se ϕ ha anche tratti negativi, l'unica differenza è che gli scalini con “altezza negativa” si contano con “area negativa”.

Le funzioni a scala ed il loro integrale godono delle seguenti proprietà, la cui dimostrazione è lasciata alla riflessione del lettore:

LEMMA: Se ϕ_1 e ϕ_2 sono funzioni a scala su $[a, b]$, e $c \in \mathbf{R}$, allora

- *$c\phi_1$ e $\phi_1 + \phi_2$ sono funzioni a scala;*
- *l'integrale è lineare:*

$$\int_a^b (c\phi_1(x))\,dx = c\int_a^b \phi_1(x)\,dx$$
$$\int_a^b (\phi_1(x) + \phi_2(x))\,dx = \int_a^b \phi_1(x)\,dx + \int_a^b \phi_2(x)\,dx;$$

- *l'integrale è monotono: se $\phi_1(x) \le \phi_2(x)$ per ogni $x \in [a, b]$, allora*

$$\int_a^b \phi_1(x)\,dx \le \int_a^b \phi_2(x)\,dx.$$

Siamo ora in grado di definire l'integrale superiore, l'integrale inferiore ed eventualmente l'integrale di una funzione limitata:

DEFINIZIONE: Sia $f : [a, b] \to \mathbf{R}$ una funzione limitata. L'*integrale superiore* di f si definisce come

$$\overline{\int_a^b} f(x)\,dx = \inf\{\int_a^b \phi(x)\,dx : \ \phi \ a \ scala, \ \phi \ge f \ in \ [a, b]\}.$$

Analogamente, l'*integrale inferiore* di f si definisce come

$$\underline{\int_a^b} f(x)\,dx = \sup\{\int_a^b \psi(x)\,dx : \ \psi \ a \ scala, \ \psi \le f \ in \ [a, b]\}.$$

Se l'integrale superiore e l'integrale inferiore coincidono, diremo che *la funzione f è integrabile secondo Riemann* in $[a, b]$, ed indicheremo il valore comune dei due integrali con

$$\int_a^b f(x)\,dx,$$

integrale secondo Riemann di f su $[a, b]$.

Vale una semplice caratterizzazione dell'integrale di Riemann, che permette tra l'altro di verificare la correttezza di quanto trovato sopra per l'integrale della funzione esponenziale:

PROPOSIZIONE: Sia $f : [a, b] \to \mathbf{R}$ una funzione limitata. f è integrabile secondo Riemann se e solo se per ogni $\varepsilon > 0$ si possono trovare due funzioni a scala ϕ, ψ con $\psi \le f \le \phi$ in $[a, b]$ tali che

$$\int_a^b \phi(x)\,dx - \int_a^b \psi(x)\,dx < \varepsilon.$$

Vedremo la semplice dimostrazione di questo risultato la prossima volta. Vediamo piuttosto un esempio di funzione *non integrabile* secondo Riemann:

ESEMPIO (Funzione di Dirichlet): Si consideri la funzione $f : [0, 1] \to \mathbf{R}$ definita da

$$f(x) = \begin{cases} 1 & se\ x \in \mathbf{Q} \\ 0 & se\ x \in \mathbf{R} \setminus \mathbf{Q} \end{cases}$$

Abbiamo già visto questa funzione durante la prima unità, ed abbiamo osservato che essa è ovunque discontinua. Attualmente, invece, ci preme di osservare che *una funzione a scala maggiore o uguale a* f *sarà ovunque maggiore o uguale a* 1, e che *una funzione a scala minore o uguale a* f *è ovunque minore o uguale a* 0. Questo segue dalla densità dei razionali e degli irrazionali: in ogni "scalino" di qualunque funzione a scala, esistono sia punti razionali in cui la funzione vale 1, che punti irrazionali in cui essa vale 0.

Se ne deduce che

$$\overline{\int_0^1} f(x)\,dx = 1, \qquad \underline{\int_0^1} f(x)\,dx = 0,$$

e la funzione non è integrabile secondo Riemann.

Lezione del 10/3/2005 (2 ore): Dimostriamo la caratterizzazione delle funzioni integrabili secondo Riemann. Se f è integrabile, per definizione di integrale superiore e di integrale inferiore possiamo trovare due funzioni a scala ϕ e ψ, la prima maggiore o uguale e la seconda minore o uguale a f, tali che

$$\int_a^b \phi(x)\,dx < \int_a^b f(x)\,dx + \varepsilon/2,$$
$$\int_a^b \psi(x)\,dx > \int_a^b f(x)\,dx - \varepsilon/2,$$

da cui

$$\int_a^b \phi(x)\,dx - \int_a^b \psi(x)\,dx < \varepsilon.$$

Viceversa, prendiamo $\varepsilon > 0$ e consideriamo le due funzioni a scala ϕ, ψ che ci vengono assicurate dall'ipotesi.

Per definizione di integrale superiore e di integrale inferiore avremo

$$\int_a^b \phi(x)\,dx \geq \overline{\int_a^b} f(x)\,dx,$$
$$\int_a^b \psi(x)\,dx \leq \underline{\int_a^b} f(x)\,dx,$$

per cui

$$0 \leq \overline{\int_a^b} f(x)\, dx - \underline{\int_a^b} f(x)\, dx \leq \int_a^b \phi(x)\, dx - \int_a^b \psi(x)\, dx < \varepsilon.$$

Per l'arbitrarietà di ε, ne deriva che l'integrale superiore e l'integrale inferiore sono uguali. Q.E.D.

Possiamo poi verificare che l'integrale delle funzioni integrabili secondo Riemann gode delle stesse proprietà di linearità dell'integrale delle funzioni a scala (si veda il Lemma enunciato la volta scorsa):

PROPOSIZIONE: Siano $f, g : [a, b] \to \mathbf{R}$ due funzioni integrabili secondo Riemann, $c \in \mathbf{R}$. Allora le funzioni $c \cdot f$ e $f + g$ sono integrabili secondo Riemann e si ha

$$\int_a^b c \cdot f(x)\, dx = c \int_a^b f(x)\, dx$$
$$\int_a^b [f(x) + g(x)]\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx.$$

DIM.: Supponiamo dapprima $c > 0$. L'integrabilità della funzione $c \cdot f$ (e la prima delle due identità nella tesi) segue subito se si osserva che l'insieme delle funzioni a scala maggiori o uguali a $c \cdot f$ coincide con l'insieme delle funzioni a scala maggiori o uguali a f moltiplicate per c. Poiché per le funzioni a scala è lecito "portare la costante c fuori dal segno di integrale" (si veda il lemma enunciato la volta scorsa), ne deduciamo che

$$\overline{\int_a^b} c \cdot f(x)\, dx = c \overline{\int_a^b} f(x)\, dx.$$

Un'identità del tutto analoga vale per l'integrale inferiore, da cui la tesi.

Se invece $c < 0$ (il caso $c = 0$ è ovvio!), l'unico cambiamento da fare viene dal fatto che una funzione a scala maggiorante per f moltiplicata per c, darà una funzione a scala minorante per $c \cdot f$.

Veniamo alla somma: osserviamo che

$$\{\Phi :\ \Phi \text{ } a \text{ } scala,\ \Phi \geq f + g\} \supset \{\phi_1 + \phi_2 :\ \phi_1, \phi_2 \text{ } a \text{ } scala,\ \phi_1 \geq f,\ \phi_2 \geq g\},$$

da cui, ricordando la definizione di integrale superiore:

$$\overline{\int_a^b} [f(x) + g(x)]\, dx \leq \overline{\int_a^b} f(x)\, dx + \overline{\int_a^b} g(x)\, dx.$$

(A lezione abbiamo visto un esempio che mostra come, se togliamo l'ipotesi che f e g siano integrabili, possa anche succedere che valga la disuguaglianza stretta).

Analogamente, per gli integrali inferiori si ottiene

$$\underline{\int_a^b}[f(x)+g(x)]\,dx \geq \underline{\int_a^b} f(x)\,dx + \leq \overline{\int_a^b} g(x)\,dx.$$

Mettendo assieme le due disuguaglianze e ricordando l'integrabilità di f e g otteniamo

$$\int_a^b f(x)\,dx + \int_a^b g(x)\,dx \leq \underline{\int_a^b}[f(x)+g(x)]\,dx \leq \overline{\int_a^b}[f(x)+g(x)]\,dx \leq$$
$$\int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$

da cui la tesi. Q.E.D.

A questo punto, diventa importante far vedere che tutte le funzioni "abbastanza decenti" sono integrabili. Per esempio, sono integrabili le funzioni monotone:

TEOREMA: Se $f:[a,b]\to \mathbf{R}$ *è crescente (decrescente), allora è integrabile secondo Riemann.*

DIM.: Per fissare le idee, trattiamo il caso in cui f sia crescente.

Dividiamo l'intervallo $[a,b]$ in n parti uguali (lunghe $(b-a)/n$), di estremi $a=x_0<x_1<x_2<\ldots<x_n=b$. Definiamo una funzione a scala maggiorante ϕ_n e una funzione a scala minorante ψ_n nel modo seguente: nell'i-esimo intervallino $[x_i,x_{i+1})$, poniamo ϕ_n uguale a $f(x_{i+1})$, e ψ_n uguale a $f(x_i)$.

Si ha

$$\int_a^b \phi_n(x)\,dx - \int_a^b \psi_n(x)\,dx = \sum_{i=0}^{n-1}\frac{b-a}{n}[f(x_{i+1})-f(x_i)] =$$
$$\frac{b-a}{n}\sum_{i=0}^{n-1}[f(x_{i+1})-f(x_i)] = \frac{b-a}{n}[f(b)-f(a)],$$

e l'ultima quantità può essere resa piccola a piacere a patto di prendere n abbastanza grande. Q.E.D.

Verifichiamo che sono integrabili anche le funzioni continue. Non sarà altrettanto facile...

Nostro obiettivo è dimostrare il seguente

TEOREMA (Integrabilità delle funzioni continue): Sia $f : [a,b] \to \mathbf{R}$ una funzione continua. Allora f è integrabile secondo Riemann.

La dimostrazione di questo teorema è molto meno semplice di quella che abbiamo dato per le funzioni monotone: in effetti, abbiamo bisogno di un concetto nuovo, quello di *uniforme continuità.*

DEFINIZIONE: Una funzione f si dice *uniformemente continua* su un intervallo I se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni $x, y \in I$ con $|x-y| < \delta$, si ha $|f(x) - f(y)| < \varepsilon$.

In questa definizione chiediamo qualcosa di più della continuità in tutti i punti di I: stiamo infatti pretendendo che δ dipenda da ε, ma non dal punto in cui andiamo a valutare la continuità. Lo stesso δ deve funzionare in tutti i punti dell'intervallo I!

In effetti, se è vero che ogni funzione uniformemente continua in I è anche continua in I, il viceversa può essere falso: per esempio, la funzione $f(x) = x^2$ è continua in tutti i punti della retta reale, ma non è uniformemente continua su $\mathbf{R}$.

Se però consideriamo soltanto intervalli chiusi e limitati, i due concetti coincidono:

TEOREMA (di Heine-Cantor): Una funzione continua $f : [a,b] \to \mathbf{R}$ è anche uniformemente continua.

DIM.: Ragioniamo per assurdo. Se f non fosse uniformemente continua, dovrebbe esistere $\varepsilon_0 > 0$ tale che per ogni $\delta > 0$ esistano $x, y \in [a,b]$ con $|x-y| < \delta$ e tali che $|f(x) - f(y)| \geq \varepsilon_0$.

Poniamo $\delta = \frac{1}{n}$ (δ può essere scelto arbitrariamente...): per quanto appena visto possiamo trovare due punti $x_n, y_n \in [a,b]$ con $|x_n - y_n| < \frac{1}{n}$ e $|f(x_n) - f(y_n)| \geq \varepsilon_0$.

Ora, la successione $\{x_n\} \subset [a,b]$ è limitata, per cui possiamo estrarre una successione $\{x_{n_k}\}$ tale che esiste $\lim\limits_{k\to+\infty} x_{n_k} = \overline{x}$. Evidentemente, visto che si ha $x_{n_k} - \frac{1}{n_k} < y_{n_k} < x_{n_k} + \frac{1}{n_k}$ avremo anche $\lim\limits_{k\to+\infty} y_{n_k} = \overline{x}$.

Per la continuità di f in $\overline{x}$ avremo allora

$$\lim_{k\to+\infty} f(x_{n_k}) = \lim_{k\to+\infty} f(y_{n_k}) = f(\overline{x}),$$

e in particolare

$$\lim_{k\to+\infty} |f(x_{n_k}) - f(y_{n_k})| = 0.$$

Questo è assurdo perché per costruzione la disuguaglianza $|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon_0$ deve valere per ogni k. Q.E.D.

Lezione del 15/3/2005 (2 ore): Dimostriamo che una funzione continua $f : [a, b] \to \mathbf{R}$ è intergrabile. Sia $\varepsilon > 0$. Per il teorema di Heine-Cantor, f è uniformemente continua su $[a, b]$, e possiamo trovare $\delta > 0$ tale che $|f(x) - f(y)| < \varepsilon$ quando $x, y \in [a, b]$ e $|x - y| < \delta$.

In particolare, se I è un qualunque sottointervallo chiuso di $[a, b]$ di lunghezza minore di δ, avremo

$$\max\{f(x) : \ x \in I\} - \min\{f(x) : \ x \in I\} \leq \varepsilon.$$

Scegliamo ora n abbastanza grande, in modo che $(b-a)/n < \delta$, e suddividiamo $[a, b]$ in n parti uguali tramite i punti $a = x_0 < x_1 < x_2 < \ldots < x_n = b$. Definiamo una funzione a scala maggiorante ϕ_n ed una minorante ψ_n nel modo seguente: sull'i-esimo intervallino $[x_i, x_{i+1})$, la funzione ϕ_n vale $M_i = \max\{f(x) : \ x \in [x_i, x_{i+1}]\}$, mentre ψ_n vale $m_i = \min\{f(x) : \ x \in [x_i, x_{i+1}]\}$. Per quanto visto sopra, sarà $M_i - m_i \leq \varepsilon$.

Abbiamo

$$\int_a^b \phi_n(x)\, dx - \int_a^b \psi_n(x)\, dx = \sum_{i=0}^{n-1} \frac{b-a}{n}(M_i - m_i) \leq \varepsilon \sum_{i=0}^{n-1} \frac{b-a}{n} = \varepsilon(b-a),$$

e questa quantità può essere resa piccola a piacere. Q.E.D.

Una proprietà dell'integrale che ci sarà piuttosto utile, è l'additività rispetto all'intervallo di integrazione:

PROPOSIZIONE: Se $f : [a, b] \to \mathbf{R}$ *è integrabile,* $c \in (a, b)$ *si ha*

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx.$$

Lasciamo per esercizio la facile dimostrazione.

Se $f : [a, b] \to \mathbf{R}$ è integrabile secondo Riemann, poniamo poi *per definizione*

$$\int_b^a f(x)\, dx = -\int_a^b f(x)\, dx.$$

Con questa posizione, la proprietà di additività dell'integrale rispetto all'intervallo vale *in generale*, anche se c non è compreso tra a e b (in tal caso, f dovrà essere integrabile sul più grande tra gli intervalli coinvolti...).

Per il calcolo effettivo degli integrali, lo strumento fondamentale risulta il seguente

TEOREMA (teorema fondamentale del calcolo integrale): Sia $f : [a, b] \to \mathbf{R}$ *una funzione continua, e definiamo la funzione integrale*

$$F(x) = \int_a^x f(t)\, dt.$$

Allora F *è derivabile, e* $F'(x) = f(x)$ *per ogni* $x \in [a, b]$.

Evidentemente, se siamo in grado di calcolare la funzione integrale $F(x)$ associata a f, siamo a maggior ragione in grado di calcolare l'integrale $\int_a^b f(x)\, dx$ (che sarà uguale semplicemente a $F(b)$).

Il teorema fondamentale del calcolo integrale dice che la funzione integrale $F(x)$ è da ricercarsi tra le funzioni la cui derivata è data da $f(x)$ (che si chiamano *primitive* di f). Si capisce quindi che se siamo in grado di trovare le primitive di una data funzione $f(x)$, saremo anche in condizione di trovare la funzione integrale ad essa associata: vedremo meglio come fare la prossima volta, dopo aver dimostrato il teorema fondamentale.

Per dimostrare il teorema fondamentale ci servirà il seguente risultato:

TEOREMA (Della media integrale): Sia $f : [a, b] \to \mathbf{R}$ *una funzione continua. Esiste un punto* $c \in [a, b]$ *tale che*

$$\int_a^b f(x)\, dx = f(c) \cdot (b - a).$$

DIM.: Sia $M = \max\{f(x) :\ x \in [a, b]\}$, $m = \min\{f(x) :\ x \in [a, b]\}$.

Per la monotonia dell'integrale, avremo $m(b-a) \leq \int_a^b f(x)\, dx \leq M(b-a)$, per cui la quantità

$$\frac{1}{b-a} \int_1^b f(x)\, dx$$

(che è detta *media integrale*) è compresa tra m e M.

Per il teorema dei valori intermedi, f assume tutti i valori compresi tra m e M, e quindi anche il valore corrispondente alla media integrale! Q.E.D.

DIM. DEL TEOREMA FONDAMENTALE: Consideriamo il rapporto incrementale della funzione F nel punto x. Ricordando l'additività dell'integrale rispetto all'intervallo si ha:

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\left[\int_a^{x+h} f(t)\, dt - \int_a^x f(t)\, dt\right] = \frac{1}{h}\int_x^{x+h} f(t)\, dt.$$

Applicando il Teorema della media all'ultimo integrale troviamo un punto c compreso tra x e $x + h$ tale che

$$\frac{F(x+h) - F(x)}{h} = f(c).$$

Passando al limite per $h \to 0$, avremo $c \to x$ e $f(c) \to f(x)$ (perché f è continua), da cui

$$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = f(x).$$

Q.E.D.

OSSERVAZIONE: Si possono leggermente indebolire le ipotesi del teorema fondamentale del calcolo integrale. Se ad esempio $f : [a, b] \to \mathbf{R}$ è una funzione integrabile secondo Riemann continua *in un certo punto* $x_0 \in [a, b]$, allora la funzione integrale $F(x)$ è derivabile nel punto x_0, e $F'(x_0) = f(x_0)$. Infatti, indichiamo con $M(h)$ e $m(h)$ rispettivamente il sup e l'inf della funzione f nell'intervallo di estremi x_0 e $x_0 + h$. Come visto sopra, si ha

$$\frac{F(x_0+h) - F(x_0)}{h} = \frac{1}{h} \int_{x_0}^{x_0+h} f(t)\, dt,$$

e la media integrale che compare nel membro di destra è compresa tra $m(h)$ e $M(h)$ (vedi dimostrazione del teorema della media integrale...). Grazie alla continuità di f in x_0 si ha che $M(h) \to f(x_0)$ e $m(h) \to f(x_0)$ per $h \to 0$, e la tesi segue dal teorema dei carabinieri.

DEFINIZIONE: Se $f : [a, b] \to \mathbf{R}$, una *primitiva* di f è una funzione $G : [a, b] \to \mathbf{R}$ tale che $G'(x) = f(x)$ per ogni $x \in [a, b]$.

Il teorema fondamentale del calcolo integrale dice che la funzione integrale è una primitiva dell'integranda. Ora, spesso è facile "indovinare" una primitiva di una funzione f (e vedremo presto delle tecniche per trovare in modo più sistematico le primitive di moltissime funzioni elementari). Se ci riusciamo, siamo anche in grado di calcolare l'integrale di f:

PROPOSIZIONE: Sia $f : [a, b] \to \mathbf{R}$ una funzione continua, $G : [a, b] \to \mathbf{R}$ una qualunque primitiva di f. Allora

$$\int_a^b f(x)\, dx = G(b) - G(a).$$

DIM.: Per il Teorema fondamentale, anche la funzione integrale $F(x) = \int_a^x f(t)\, dt$ è una primitiva di F. Ne segue che la funzione $F(x) - G(x)$ ha derivata nulla in $[a, b]$, e quindi è costante: esiste $C \in \mathbf{R}$ tale che

$$F(x) = G(x) + C$$

per ogni $x \in [a, b]$. Siccome $F(a) = \int_a^a f(t)\, dt = 0$, ponendo $x = a$ nell'ultima identità troviamo $C = -G(a)$, da cui $F(x) = G(x) - G(a)$. In particolare, per $x = b$ si ha la tesi. Q.E.D.

ESEMPI:

- Se $f(x) = e^x$, riconosciamo subito che una primitiva è data dalla funzione $G(x) = e^x$, e quindi

$$\int_0^a e^x\, dx = G(a) - G(0) = e^a - 1.$$

  Abbiamo così rapidamente riottenuto un risultato che ci eravamo conquistati con una certa fatica a partire dalla definizione di integrale!

- Se $f(x) = mx$, si vede che una primitiva è $G(x) = mx^2/2$, quindi

$$\int_a^b mx\, dx = m(b^2 - a^2)/2.$$

  Possiamo convincerci della veridicità di questa formula se osserviamo che geometricamente l'integrale appena calcolato rappresenta l'area di un trapezio rettangolo di altezza $(b - a)$ e basi ma e mb.

- Se $f(x) = \sin x$, una primitiva sarà $G(x) = -\cos x$. Quindi

$$\int_0^\pi \sin x\, dx = -cos(\pi) - (-cos(0)) = 2,$$

  risultato che non sarebbe stato facile prevedere in altro modo!

Cerchiamo ora di studiare una maniera un po' più sistematica per trovare le primitive di una data funzione elementare.

Conviene stabilire una notazione per l'insieme di tutte le primitive di una funzione f:

DEFINIZIONE: Se $f : [a, b] \to \mathbf{R}$ è una funzione continua, l'insieme delle primitive $G : [a, b] \to \mathbf{R}$ (cioè l'insieme delle funzioni derivabili la cui derivata è uguale ad f), si denota con il simbolo $\int f(x)\, dx$:

$$\int f(x)\, dx = \{G : [a, b] \to \mathbf{R} : \ G'(x) = f(x)\ \forall x \in [a, b]\}.$$

L'insieme delle primitive di f si chiama talvolta *integrale indefinito* di f.

Grazie al teorema fondamentale del calcolo integrale e a quanto osservato la volta scorsa, sappiamo che la funzione integrale è una primitiva di f, e anche che due primitive diverse definite su uno stesso intervallo differiscono per una costante, dunque:

$$\int f(x)\ dx = \{\int_a^x f(t)\ dt + C :\ C \in \mathbf{R}\}.$$

Con lieve abuso di notazione, se G è una qualunque primitiva di f si scrive $\int f(x)\ dx = G(x) + C$, con C *costante arbitraria*.

Lezione del 16/3/2005 (2 ore): Se prendiamo una tabella delle derivate delle funzioni elementari, e la leggiamo al contrario, troviamo la seguente tabella di integrali indefiniti "immediati":

$f(x)$	$\int f(x)\ dx$
x^a (con $a \neq -1$)	$\frac{1}{a+1}x^{a+1} + C$
$1/x$	$\log\lvert x\rvert + C$
e^x	$e^x + C$
$\sin x$	$-\cos x + C$
$\cos x$	$\sin x + C$
$\frac{1}{\cos^2 x}$	$\tan x + C$
$\frac{1}{1+x^2}$	$\arctan x + C$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x + C$

Altre regole di integrazione vengono direttamente dalle regole di derivazione della somma e del prodotto:

- $\int (f(x) + g(x))\ dx = \int f(x)\ dx + \int g(x)\ dx$ (additività),
- $\int cf(x)\ dx = c \int f(x)\ dx$,
- $\int f(x)g'(x)\ dx = f(x)g(x) - \int f'(x)g(x)\ dx$ (formula di integrazione per parti).

Dalla formula per la derivata della funzione composta viene un'utile formula di *integrazione per sostituzione*: se $F(x)$ è una primitiva di $f(x)$, allora

$$\int f(g(x))g'(x)\ dx = F(g(x)) + C.$$

Se "cambiamo variabile" ponendo $y = g(x)$, l'identità appena scritta diventa $\int f(y)\ dy = \int f(g(x))g'(x)\ dx$. Questa formula può essere ricordata nel seguente modo, poco ortodosso ma efficace: usando la notazione di Leibniz per la derivata, abbiamo

$$\frac{dy}{dx} = \frac{d(g(x))}{dx} = g'(x),$$

e moltplicando ambo i membri per dx (cosa che *ovviamente* non ha alcun senso!) otteniamo l'identità $dy = d(g(x)) = g'(x)\ dx$, che sostituita nell'integrale $\int f(y)\ dy$ ci restituisce la formula voluta...

Evidentemente, abbiamo commesso più di un crimine matematico: la derivata NON è un rapporto, e il simbolo dx nell'integrale NON ha un significato matematico ben definito (se non quello di indicare la variabile indipendente rispetto alla quale si integra). D'altra parte, la notazione per la derivata e l'integrale si è rivelata suggestiva: agendo senza farci troppi scrupoli, abbiamo comunque ottenuto un risultato corretto (infatti lo avevamo già giustificato partendo dalla formula di derivazione di funzioni composte!).

ESEMPI:

1. Si voglia calcolare $\int \cos^2 x\ dx$. Per una nota identità trigonometrica, abbiamo $\cos^2 x = \frac{1+\cos 2x}{2}$ da cui

$$\int \cos^2 x\ dx = \int \frac{1+\cos 2x}{2}\ dx = \frac{x}{2} + \frac{1}{4}\int \cos 2x\ d(2x) = \frac{x}{2} + \frac{\sin 2x}{4} + C$$

2. Calcoliamo poi $\int \sqrt{1-x^2}\ dx$. Poniamo $x = \sin y$, da cui (OK, non è ortodosso ma abbiamo verificato che funziona): $dx = \cos y\ dy$ e

$$\int \sqrt{1-x^2}\ dx = \int \sqrt{1-\sin^2 y}\cos y\ dy = \int \cos^2 y\ dy = \frac{y}{2} + \frac{\sin 2y}{4} + C.$$

   Ora, $y = \arcsin x$ e $\cos y = \sqrt{1-x^2}$, da cui $\sin 2y = 2x\sqrt{1-x^2}$ e l'integrale cercato vale $\frac{1}{2}\arcsin x + \frac{1}{2}x\sqrt{1-x^2} + C$.

   L'integrale definito di questa funzione tra -1 e 1 vale allora $\pi/2$... e questo ha un semplice significato geometrico: quale? Questa è un'ulteriore conferma del fatto che l'integrale di Riemann è l'oggetto giusto per definire l'area di oggetti curvilinei del piano dal contorno abbastanza regolare!

3. Calcoliamo $\int \log x\ dx$. Integrando per parti si ha: $\int \log x\ dx = \int 1 \cdot \log x\ dx = \int (x)' \log x\ dx = x\log x - \int x \cdot \frac{1}{x}\ dx = x\log x - x + C$.

4. Calcoliamo $\int e^x \sin x\ dx$. Integrando due volte per parti si ottiene: $\int e^x \sin x\ dx = e^x \sin x - \int e^x \cos x\ dx = e^x \sin x - e^x \cos x - \int e^x \sin x\ dx$. Portando l'ultimo integrale a primo membro si ottiene subito

$$\int e^x \sin x = \frac{e^x \sin x - e^x \cos x}{2} + C.$$

5. Calcoliamo
$$\int \frac{x}{\sqrt{1-x^2}}\,dx.$$
Poniamo $y = 1 - x^2$, da cui: $dy = -2x\,dx$ e l'integrale diventa
$$-\frac{1}{2}\frac{dy}{\sqrt{y}} = -\sqrt{y} + C = -\sqrt{1-x^2} + C.$$
Oppure, con un uso disinvolto della formula di integrazione per sostituzione:
$$\int \frac{x}{\sqrt{1-x^2}}\,dx = -\frac{1}{2}\int \frac{d(1-x^2)}{\sqrt{1-x^2}} = -\frac{1}{2}\int (1-x^2)^{-1/2} d(1-x^2) = -\sqrt{1-x^2} + C.$$

6. Si voglia calcolare $\int \sin\sqrt{x}\,dx$. Si ottiene subito il risultato ponendo $\sqrt{x} = y$, e integrando per parti.

7. Per calcolare l'integrale $\int \sqrt{1+x^2}\,dx$, si operi la sostituzione $\sqrt{1+x^2} = y - x$...

Lezione del 22/3/2005 (2 ore): Avendo imparato delle efficaci tecniche di integrazione, vogliamo ora cominciare un discorsetto introduttivo sulle equazioni differenziali ordinarie.

Dovremmo essere già sufficientemente motivati allo studio di questo argomento: nel corso di fisica abbiamo visto numerosi esempi di equazioni del secondo ordine che "nascono" dall'applicazione del II principio della dinamica.

Cominciamo col ricordare qualche definizione: un'equazione differenziale è un'equazione *funzionale* (cioè un'equazione in cui l'incognita è una funzione) della forma
$$f(x^{(n)}(t), x^{(n-1)}(t), \ldots, x'(t), x(t), t) = 0,$$
dove $x(t)$ è la funzione incognita da cercare, e f è una funzione nota di $(n+2)$ variabili. L'equazione scritta sopra è *di ordine* n, perché la derivata di ordine massimo che vi compare ha appunto ordine n. In particolare, la più generale equazione differenziale ordinaria del primo ordine sarà del tipo
$$f(x'(t), x(t), t) = 0.$$
Almeno in questa prima fase, ci occuperemo soltanto di equazioni del primo ordine: vedremo infatti che uno studio accorto sulle equazioni e sistemi del primo ordine, ci darà poi un sacco di informazioni utili anche sulle equazioni di ordine superiore.

Notiamo che, in certi casi, un'equazione differenziale del primo ordine si può scrivere esplicitando $x'(t)$ in funzione di $x(t)$ e t:

$$x'(t) = g(x(t), t).$$

In tal caso si dice che l'equazione è scritta *in forma normale*, e sarà di questo tipo di equazioni che ci occuperemo nel seguito.

Se imponiamo il valore della soluzione $y(t)$ in un punto iniziale t_0, otteniamo quello che si chiama *problema di Cauchy*: per una generale equazione del primo ordine in forma normale esso assume la forma

$$\begin{cases} x'(t) = g(x(t), t) \\ x(t_0) = x_0 \end{cases}$$

Ci sarebbe estremamente utile avere un *teorema di esistenza e unicità* che ci dicesse che il problema ha soluzione, e che questa soluzione è unica.

Nel caso in cui l'equazione differenziale rappresenti il modello di una situazione reale, è evidente l'importanza fisica e filosofica di un risultato di esistenza e unicità di questo tipo: se il modello è corretto, il sistema dovrà evolvere in qualche modo (e quindi ci devono essere soluzioni del problema di Cauchy), mentre l'unicità delle soluzioni corrisponde al requisito di *ripetibilità di un esperimento* (il sistema deve reagire nello stesso modo se sono identiche le condizioni iniziali: *determinismo*).

ESEMPIO: Consideriamo il problema di Cauchy

$$\begin{cases} x'(t) = x^2(t) \\ x(0) = 1 \end{cases}$$

Scrivendo l'equazione nella forma $x'(t)/x^2(t) = 1$ e integrando, si trova che l'unica soluzione è $x(t) = 1/(1-t)$. Siccome una soluzione ci aspettiamo che sia continua e derivabile in un intervallo che contiene il punto iniziale, se ne deduce che la soluzione del problema *non è definita su tutta la retta reale, ma solo sulla semiretta* $(-\infty, 1)$.

Dunque, anche se il secondo membro dell'equazione è estremamente regolare, possiamo sperare solo in un teorema di esistenza e unicità *locale*, che ci assicuri l'esistenza di una soluzione in un opportuno *intorno dell'istante iniziale* t_0.

ESEMPIO: Si noti che, affinché la nostra equazione differenziale abbia senso, sembra ragionevole assumere che la funzione $g(x, t)$ sia continua (anche se dovremo precisare cosa significa che una funzione di due variabili è continua).

La sola continuità non è però sufficiente a garantire l'unicità della soluzione del problema di Cauchy: ad esempio il problema

$$\begin{cases} x'(t) = \sqrt{|x(t)|} \\ x(0) = 0 \end{cases}$$

ammette infinite soluzioni. Infatti, si vede subito che $x(t) = 0$ è una soluzione. Inoltre, applicando lo stesso trucco usato sopra scopriamo che la funzione $x(t) = t^2/4$ è una soluzione del problema per $t > 0$... e indoviniamo così che tutte le funzioni del tipo

$$x(t) = \begin{cases} \frac{(t-t_0)^2}{4} & se\ t \geq t_0 \\ 0 & se\ t < t_0 \end{cases}$$

sono soluzioni del problema di Cauchy ($t_0 > 0$ è una costante).

A questo punto, sappiamo che il nostro teorema di esistenza e unicità potrà darci soltanto un *risultato di esistenza locale*, e che per avere l'unicità dovremo chiedere più che la continuità della funzione g...

In compenso, il trucchetto che abbiamo usato prima per risolvere i semplici problemi che abbiamo preso in considerazione, si generalizza a tutta una classe di equazioni differenziali dette *a variabili separabili*: supponiamo che $A(t)$ e $B(x)$ siano due funzioni continue, definite rispettivamente in un intorno di t_0 e di x_0. Supponiamo anche che $B(x_0) \neq 0$. Sia dato il problema di Cauchy

$$\begin{cases} x'(t) = \frac{A(t)}{B(x(t))} \\ x(t_0) = x_0 \end{cases}$$

Supponiamo ora che $x(t)$ sia una soluzione del problema: moltiplichiamo ambo i membri dell'equazione per $B(x(t))$ e integriamo tra t_0 e t. Se denotiamo con $F(t)$ e $G(x)$ delle primitive di $A(t)$ e $B(x)$ rispettivamente, si ottiene $G(x(t)) = F(t) - F(t_0) + G(x_0)$.

Questa è un'equazione che ci fornisce la soluzione $x(t)$ *in forma implicita*: se per caso sapessimo che la funzione G è invertibile in un intorno di x_0 (e tra un attimo mostreremo che questo è proprio vero!), potremmo ottenere esplicitamente la soluzione come $x(t) = G^{-1}(F(t) - F(t_0) + G(x_0))$.

Vediamo di mostrare che G è invertibile in un intorno di x_0. Ora, $G'(x) = B(x)$ è una funzione continua, e per ipotesi $G'(x_0) = B(x_0) \neq 0$. Per il teorema della permanenza del segno, $G'(x)$ deve avere segno costante in un intorno di x_0, per cui G è strettamente monotona ed invertibile in tale intorno. La funzione inversa G^{-1} sarà definita di conseguenza in un intorno di $G(x_0)$, e possiamo scrivere $x(t) = G^{-1}(F(t) - F(t_0) + G(x_0))$.

Si noti che vale anche il viceversa: la funzione $x(t)$ appena trovata è una soluzione del problema di Cauchy (basta ripercorrere a ritroso i passaggi che abbiamo fatto!): in questo caso abbiamo trovato esplicitamente la soluzione del problema, ed abbiamo anche dimostrato che essa è unica.

Un'altra classe di equazioni del primo ordine che si possono risolvere esplicitamente è costituita dalle *equazioni lineari del primo ordine*: esse sono della forma

$$x'(t) + a(t)x(t) = b(t),$$

dove a e b sono date funzioni continue definite su un certo intervallo (a, b).

Se imponiamo anche la condizione iniziale $x(t_0) = x_0$ (con $t_0 \in (a, b)$), mostreremo ora che esiste un'unica soluzione, definita su tutto l'intervallo (a, b): in questo caso particolarmente fortunato abbiamo *esistenza e unicità globale*, e possiamo anche scrivere esplicitamente la soluzione.

Il trucco consiste nel moltiplicare ambo i membri dell'equazione per il *fattore integrante*

$$F(t) = e^{\int_{t_0}^{t} a(s)\,ds}.$$

Facendo questo, a primo membro avremo la derivata del prodotto $x(t)F(t)$, e integrando ambo i membri tra t_0 e t troviamo la soluzione esplicita del problema di Cauchy:

$$x(t) = e^{-\int_{t_0}^{t} a(s)\,ds} \left(x_0 + \int_{t_0}^{t} e^{\int_{t_0}^{t} a(s)\,ds} b(t)\,dt \right).$$

Lezione del 29/3/2005 (2 ore): Dagli esempi visti la volta scorsa, abbiamo capito che quel che vogliamo è 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*

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

(si dice in tal caso che 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 una funzione derivabile $y(t) : [t_0 - \delta, t_0 + \delta] \to \mathbf{R}$ che risolve il problema di Cauchy

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

Questa soluzione è unica, nel senso che ogni altra soluzione del problema coincide con $y(t)$ nell'intervallo in cui sono definite entrambe.

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 loro 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 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.

- Abbiamo 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!).

Come promesso, dimostriamo la parte "unicità" del teorema di esistenza e unicità locale. Ci serve il seguente

LEMMA: Sia $y(t) : [a, b] \to \mathbf{R}$ una funzione derivabile. Supponiamo che $y(a) = 0$ (oppure che $y(t_0) = 0$ per almeno un $t_0 \in [a, b]$), 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 $\delta < 1/L$, 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. Con minime modifiche, la stessa dimostrazione vale anche se l'ipotesi $y(a) = 0$ è sostituita da $y(t_0) = 0$, per qualche $t_0 \in [a, b]$. Q.E.D.

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.

Il risultato di esistenza e unicità locale 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.*

Vedremo un esempio di questo al termine del corso.

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: l'idea è di considerare le nostre n funzioni incognite come *un'unica funzione incognita a valori vettoriali...* Domani vedremo come questo semplifichi enormemente le cose!

Lezione del 31/3/2005 (2 ore): Riscriviamo le n incognite del nostro sistema di equazioni differenziali (vedi lezione precedente) come un'unica funzione incognita vettoriale

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

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

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 $[a, b]$.[1]

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.

Siamo ora in grado di cominciare lo studio di una classe importante di equazioni differenziali del II ordine: le equazioni lineari.

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

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

[1] Dimostrazione:

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

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 sopra, 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).

Vediamo di capire come si può fare: consideriamo un'equazione del tipo

$$y''(t) + ay'(t) + by(t) = 0,$$

con $a, b \in \mathbf{R}$.

La volta scorsa abbiamo visto che l'insieme delle soluzioni di una *qualunque* equazione lineare omogenea del secondo ordine (anche non a coefficienti costanti) è uno spazio vettoriale di dimensione due: in sostanza, ci basta *indovinare* due soluzioni dell'equazione che non siano multiple l'una dell'altra, e tutte le altre si otterranno come combinazione lineare di queste.

Il problema è trovare queste due soluzioni!

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... Non è il caso di spaventarsi: lo vedremo la prossima volta!

Lezione del 5/4/2005 (2 ore): Nella risoluzione delle equazioni differenziali lineari omogenee del II ordine a coefficienti costanti, ci mancava da analizzare il caso in cui l'equazione caratteristica abbia de radici complesse e coniugate: come possiamo dare un senso alle soluzioni formali $\tilde{y}_1(t) = e^{(\alpha+i\beta)t}$ e $\tilde{y}_2(t) = e^{(\alpha-i\beta)t}$?

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

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

Con quello che impareremo sulle serie nella terza unità didattica, non sarà 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: verifichiamo questo fatto in dettaglio.

Dunque, 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*

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

dove $a(t)$, $b(t)$ e $f(t)$ sono funzioni continue definite su un certo intervallo I. 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): $y_0(t) = A(t)y_1(t) + B(t)y_2(t)$. Sostituiamo quest'espressione nell'equazione, imponendo arbitrariamente la condizione aggiuntiva $A'(t)y_1(t) + B'(t)y_2(t) = 0$ (che è la prima equazione del sistema (C)). Si trova che l'equazione è soddisfatta se e solo se si ha $A'(t)y_1'(t) + B'(t)y_2'(t) = f(t)$, che è poi la seconda equazione del sistema.

Dunque, se riusciamo a trovare due funzioni continue $A'(t)$ e $B'(t)$ che soddisfano il nostro sistema, possiamo poi sceglierne due primitive $A(t)$ e $B(t)$ (la loro esistenza è assicurata dal teorema fondamentale del calcolo integrale), ed abbiamo la soluzione desiderata.

Ci rimane da mostrare che il sistema (C) ammette *sempre* una ed una sola coppia $A'(t)$, $B'(t)$ di funzioni soluzione, e che queste sono continue sull'intervallo I dove sono definite le funzioni $a(t)$, $b(t)$ e $f(t)$ che compaiono nell'equazione differenziale. Questo lo vedremo domani.

Lezione del 6/4/2005 (2 ore): Concludiamo la dimostrazione del teorema sul metodo della variazione delle costanti.

Per ogni fissato $t \in I$, abbiamo visto che dobbiamo risolvere un *sistema lineare di due equazioni nelle incognite* $A'(t)$, $B'(t)$: dal corso di geometria sappiamo che il sistema (C) ammette una ed una sola soluzione se sappiamo che la matrice

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

(nota come *matrice wronskiana*) è invertibile per ogni valore del parametro t (cioè se ha sempre determinante non nullo): infatti, il sistema (C) si scrive

$$W(t) \begin{pmatrix} A'(t) \\ B'(t) \end{pmatrix} = \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.$$

D'altra parte, una matrice 2×2 è invertibile se e soltanto se le sue colonne *non* sono multiple l'una dell'altra.

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 permetterebbe di concludere che $y_1(t) = cy_2(t)$ per ogni $t \in I$.

Infatti, sia $y_1(t)$ che $cy_2(t)$ sarebbero soluzione di uno stesso problema di Cauchy, vale a dire

$$\begin{cases} y''(t) + a(t)y'(t) + b(t)y(t) = 0 \\ y(t_0) = y_1(t_0) \\ y'(t_0) = y_1'(t_0) \end{cases}$$

e quindi dovrebbero coincidere. Questo è assurdo perchè per ipotesi y_1 e y_2 generano lo spazio delle soluzioni dell'equazione omogenea, e devono quindi essere indipendenti.

Rimane da verificare che $A'(t)$ e $B'(t)$ sono continue in I. Abbiamo evidentemente

$$\begin{pmatrix} A'(t) \\ B'(t) \end{pmatrix} = W^{-1}(t) \begin{pmatrix} 0 \\ f(t) \end{pmatrix},$$

e l'inversa della matrice wronskiana si scrive esplicitamente come segue:

$$W^{-1}(t) = \frac{1}{\det W(t)} \begin{pmatrix} y_2'(t) & -y_2(t) \\ -y_1'(t) & y_1(t) \end{pmatrix}.$$

Gli elementi della matrice $W^{-1}(t)$ sono evidentemente funzioni continue, in quanto $y_1(t)$ e $y_2(t)$ sono funzioni derivabili due volte, e abbiamo verificato sopra che il determinanete di $W(t)$ è una funzione (evidentemente continua) che non si annulla mai (perché la matrice ha sempre rango due). Q.E.D

Come esempio di applicazione del metodo della variazione delle costanti arbitrarie, 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 $y(t) = t(c_1 \cos kt + c_2 \sin kt)$.