

Pietro Baldi

Sistemi Dinamici modulo 1

Appunti del corso settembre-dicembre 2019

Premessa

Queste dispense sono la trascrizione delle lezioni di *Sistemi dinamici modulo 1*, corso di ventiquattro lezioni da due ore tenuto da settembre a dicembre 2019 per gli studenti del corso di laurea triennale in matematica.

Il corso tratta di equazioni differenziali ordinarie, calcolo differenziale in spazi di Banach e serie di Fourier in una variabile.

Essendo un corso di primo semestre del terzo anno, i prerequisiti sono gli argomenti di analisi e algebra lineare che gli studenti hanno visto in Analisi 1, Analisi 2 e Geometria 1. In particolare, queste lezioni non fanno uso dell'integrale di Lebesgue ma solo di quello di Riemann, e sono adatte a studenti che hanno poca familiarità con i numeri complessi. Vari risultati già noti vengono comunque richiamati.

La prima parte del corso (lezioni 1-15) contiene i teoremi classici alla base della teoria delle equazioni differenziali ordinarie: Teorema di esistenza e unicità locale per il problema di Cauchy, Teorema di esistenza di Peano (con il Teorema di Ascoli-Arzelà), Teorema di esistenza globale, intervallo massimale di esistenza, criteri di prolungabilità, fuga dai compatti, disuguaglianza di Gronwall, dipendenza continua dai dati iniziali, dipendenza differenziabile dai dati iniziali. Alcuni esercizi sono dedicati all'esistenza globale e all'analisi qualitativa delle soluzioni di equazioni scalari.

Vengono trattati i principali metodi per la soluzione esplicita di equazioni differenziali scalari di forma speciale: equazioni a variabili separabili, equazioni con integrale primo, equazioni lineari del primo ordine, equazioni di Bernoulli e altre equazioni che si possono risolvere per sostituzione.

Una parte è dedicata alle equazioni lineari: spazio vettoriale delle soluzioni, determinante wronskiano, matrice risolvente, metodo di variazione delle costanti, polinomio caratteristico, esponenziale di matrice.

Nel preparare le lezioni 1-15 ho seguito le dispense [6] di Coti Zelati, [5] di Berti, [3] mie, e i libri [9] di Fusco, Marcellini, Sbordone, [1] di Acerbi, Modica, Spagnolo, [7] di De Marco, [10] di Gentile, [11] di Hartman, [12] di Hirsch, Smale.

La seconda parte (lezioni 15-19) è un'introduzione al calcolo differenziale negli spazi di Banach, e tratta il Lemma delle contrazioni di Banach-Caccioppoli, l'invertibilità di operatori lineari per serie di Neumann, e il Teorema della funzione implicita.

Nel preparare le lezioni 15-19 ho seguito le dispense [5] di Berti e i libri [2] di Ambrosetti, Prodi e [7] di De Marco.

La terza parte (lezioni 19-24) è un'introduzione alla teoria delle serie di Fourier, fatta senza l'integrale di Lebesgue: convergenza puntuale, uniforme e in media quadratica, nuclei di Dirichlet e di Fejer, sommabilità alla Cesàro, densità dei polinomi trigonometrici, Lemma di Riemann-Lebesgue, regolarità di una funzione e decadimento dei coefficienti di Fourier, lo spazio ℓ^2 , spazi vettoriali complessi con prodotto scalare, ortogonalità, identità di Parseval, disuguaglianza di Bessel.

Per le lezioni 19-24 ho seguito i libri [13] di Stein, Shakarchi e [4] di Barozzi.

Di alcune dimostrazioni la versione presentata in questi appunti si discosta abbastanza dalle fonti citate.

Esercizi. Esercizi su questi argomenti si possono trovare, per esempio, nell'ultima parte delle dispense [5] di Berti, nelle prove scritte di Sistemi dinamici nelle pagine docenti unina di Coti Zelati e mie, nel Demidovič [8], e soprattutto, per l'analisi qualitativa delle equazioni differenziali, nell'Acerbi, Modica, Spagnolo [1], libro che consiglio di svolgere per intero.

Errori. Ogni segnalazione riguardo a errori in queste dispense è gradita, dai veri errori fino alle virgole fuori posto: scrivere all'indirizzo `pietro.baldi (chiocciola) unina.it`

Pietro Baldi

Napoli, 20 dicembre 2019

1 Lezione 1

La prima parte di questo corso è un'introduzione di base alla teoria delle equazioni differenziali ordinarie (ODE).

La teoria delle equazioni ordinarie tratta le equazioni e i sistemi di equazioni in cui l'incognita è una funzione di una variabile e in cui compaiono le derivate della funzione incognita. Vengono dette "ordinarie" per distinguerle dalle equazioni differenziali "alle derivate parziali" (PDE), in cui l'incognita è funzione di più variabili e compaiono derivate parziali dell'incognita rispetto a variabili diverse (esempi famosi di PDE: equazione delle onde, del calore, di Laplace, di Schrödinger, del trasporto).

Notazioni standard per le ODE: la variabile viene indicata x (classico nome per la variabile reale di Analisi 1), oppure t (pensando $t =$ tempo, cosicché l'equazione descrive l'evoluzione temporale di un sistema), oppure s , ecc.; la funzione incognita viene indicata $y(x)$, $y(t)$, $x(t)$, $u(t)$, ecc. e può essere scalare o vettoriale.

Definizione 1.1. (*Soluzione di un'equazione differenziale*). Sia $\Omega \subseteq \mathbb{R}^{n+2}$, e sia $F : \Omega \rightarrow \mathbb{R}$ una funzione. Sia $I \subseteq \mathbb{R}$ un intervallo, e sia $u : I \rightarrow \mathbb{R}$ una funzione derivabile n volte in I . Diciamo che u è soluzione dell'equazione differenziale

$$F(t, u, u', \dots, u^{(n)}) = 0, \quad (1.1)$$

o che u risolve l'equazione (1.1) su I , se

- (i) per ogni $t \in I$, il vettore $(t, u(t), u'(t), \dots, u^{(n)}(t))$ appartiene all'insieme Ω ;
- (ii) per ogni $t \in I$ si ha l'identità $F(t, u(t), u'(t), \dots, u^{(n)}(t)) = 0$.

Le soluzioni delle equazioni differenziali sono dunque *definite su intervalli*.

Definizione 1.2. (*Ordine di un'equazione differenziale*). Si dice *ordine* (o *grado*) di un'equazione differenziale il massimo ordine di derivazione che compare nell'equazione.

Esempio: $u' = u^2$ è del primo ordine, $u'' - u''^3 + u^2 = 0$ è del secondo ordine.

Definizione 1.3. (*Equazione in forma normale*). Un'equazione di ordine n è in *forma normale* se è del tipo

$$u^{(n)} = f(t, u, u', \dots, u^{(n-1)}).$$

Quindi le equazioni in forma normale di ordine 1 sono del tipo $u' = f(t, u)$; di ordine 2 sono della forma $u'' = f(t, u, u')$; di ordine 3 sono $u''' = f(t, u, u', u'')$.

Definizione 1.4. (*Sistema di equazioni scalari, o equazione vettoriale, del primo ordine*). Un sistema di n equazioni differenziali (scalari) del primo ordine in n incognite (scalari) $u_1(t), \dots, u_n(t)$ è un sistema del tipo

$$\begin{cases} u_1' = f_1(t, u_1, \dots, u_n) \\ u_2' = f_2(t, u_1, \dots, u_n) \\ \dots \\ u_n' = f_n(t, u_1, \dots, u_n). \end{cases} \quad (1.2)$$

Il sistema (1.2) si scrive anche

$$u' = f(t, u) \quad (1.3)$$

dove $u(t) = (u_1(t), \dots, u_n(t))$ è funzione a valori in \mathbb{R}^n , e $f = (f_1, \dots, f_n)$ è definita su un sottoinsieme di \mathbb{R}^{1+n} a valori in \mathbb{R}^n .

Definizione 1.5. (*Campo vettoriale, spazio delle fasi, traiettorie, sistema dinamico n-dimensionale*). Siano u, f come nella Definizione 1.4.

Il sistema (1.2) o (1.3) si dice *sistema dinamico n-dimensionale*.

Lo spazio vettoriale \mathbb{R}^n in cui vive $u(t)$ si dice *spazio delle fasi*.

La funzione $f(t, u)$ si dice *campo vettoriale* del sistema dinamico.

Le soluzioni dell'equazione differenziale si dicono anche *traiettorie del sistema dinamico*.

Definizione 1.6. (*Sistema autonomo*). Quando il campo vettoriale f non dipende da t , ma solo da u , l'equazione diventa $u' = f(u)$ e si dice che il sistema dinamico è *autonomo*.

Esempio: $u' = u^2 + 3$ è un sistema autonomo, $u' = u^2 + 3t$ non lo è.

* * *

Riscaldamento: proviamo a calcolare (o, per ora, "indovinare") le soluzioni di alcune equazioni differenziali semplici.

1) Equazione $u' = u$: qual è la funzione che è uguale alla sua derivata? Certamente e^t , e anche i suoi multipli. L'equazione ha soluzioni $u(t) = ce^t$, definite per ogni t nell'intervallo $I = \mathbb{R}$, dove $c \in \mathbb{R}$ è una qualsiasi costante.

2) $u' = 2u$: ha soluzioni $u(t) = ce^{2t}$, $t \in I = \mathbb{R}$, $c \in \mathbb{R}$.

3) $u' = 1 - u$: soluzioni $u(t) = 1 + ce^{-t}$, $t \in I = \mathbb{R}$, $c \in \mathbb{R}$.

4) $u' = u^2$: soluzione $u(t) = \frac{1}{c-t}$, definita per ogni $t \neq c$; ma il dominio I deve essere un intervallo, quindi per ogni costante c c'è sia una soluzione $u(t) = \frac{1}{c-t}$ su $t \in I = (-\infty, c)$ che una soluzione $u(t) = \frac{1}{c-t}$ su $t \in I = (c, +\infty)$.

5) Trovare una soluzione dell'equazione $u''' - u'' + tu' - u = 2$: la più immediata è $u(t) = -2$ per ogni $t \in I = \mathbb{R}$. È facile da calcolare perché, essendo funzione costante, le sue derivate sono nulle e quindi spariscono dall'equazione differenziale.

6) $u'' + 4u = 0$: soluzioni $u(t) = \cos(2t)$, $u(t) = \sin(2t)$, e anche ogni loro combinazione lineare $u(t) = c_1 \cos(2t) + c_2 \sin(2t)$, su $I = \mathbb{R}$, per ogni $c_1, c_2 \in \mathbb{R}$. Notiamo che per le equazioni degli esempi precedenti, che sono tutte equazioni del primo ordine, le soluzioni sono famiglie a 1 parametro, mentre per questa equazione, che è del secondo ordine, le soluzioni sono una famiglia a 2 parametri.

7) Sistema

$$\begin{cases} u' = v \\ v' = -4u. \end{cases}$$

Se $u' = v$, allora, derivando, si ottiene $u'' = v'$, ma $v' = -4u$, da cui $u'' + 4u = 0$, come nell'esempio 6). Soluzioni:

$$u(t) = c_1 \cos(2t) + c_2 \sin(2t), \quad v(t) = -2c_1 \sin(2t) + 2c_2 \cos(2t),$$

definite per ogni $t \in I = \mathbb{R}$, al variare dei due parametri $c_1, c_2 \in \mathbb{R}$. Osserviamo che il sistema 7) (equazione vettoriale 2-dimensionale del primo ordine, o sistema di 2 equazioni scalari del primo ordine) e l'equazione 6) (equazione scalare del secondo ordine) sono strettamente legati. Prima di enunciare la relazione generale tra equazioni scalari di ordine n ed equazioni vettoriali del primo ordine, vediamo un altro esempio. Consideriamo

8) l'equazione scalare del terzo ordine

$$u''' = t^2 + u^2 + uu' + \sin(u'' - u') \quad (1.4)$$

nell'incognita scalare $u : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$, e

9) il sistema di 3 equazioni scalari del primo ordine

$$\begin{cases} u' = v \\ v' = w \\ w' = t^2 + u^2 + uv + \sin(w - v), \end{cases} \quad (1.5)$$

nell'incognita vettoriale $(u(t), v(t), w(t)) \in \mathbb{R}^3$.

Supponiamo di avere una soluzione $u(t)$ di (1.4), cioè una funzione $u : I \rightarrow \mathbb{R}$ definita su un certo intervallo $I \subseteq \mathbb{R}$, che supponiamo di classe $C^3(I)$, che soddisfa (1.4). Allora la funzione vettoriale

$$I \rightarrow \mathbb{R}^3, \quad t \mapsto \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = \begin{pmatrix} u(t) \\ u'(t) \\ u''(t) \end{pmatrix}$$

è soluzione del sistema (1.5), ed è di classe $C^1(I)$. Viceversa, se abbiamo una soluzione $(u(t), v(t), w(t)) \in C^1(I, \mathbb{R}^3)$ di (1.5), allora la sua prima componente $u(t)$ è soluzione di (1.4), ed è $C^3(I, \mathbb{R})$. Dunque in questo esempio 1 equazione di ordine 3 equivale a 3 equazioni di ordine 1. Questo vale in generale:

Osservazione 1.7. (*Equivalenza tra 1 equazione scalare di ordine n ed n equazioni scalari di ordine 1*). L'equazione differenziale scalare di ordine n

$$u^{(n)}(t) = f(t, u(t), u'(t), \dots, u^{(n-1)}(t)) \quad (1.6)$$

è equivalente al sistema di n equazioni differenziali scalari di ordine 1, cioè un'equazione del primo ordine a valori in \mathbb{R}^n ,

$$\begin{cases} u_1'(t) = u_2(t) \\ \dots \\ u_{n-1}'(t) = u_n(t) \\ u_n'(t) = f(t, u_1(t), u_2(t), \dots, u_n(t)). \end{cases} \quad (1.7)$$

“Equivalente” in questo senso: se $u : I \rightarrow \mathbb{R}$ è soluzione di (1.6), allora la funzione vettoriale $(u_1, u_2, \dots, u_n) = (u, u', \dots, u^{(n-1)}) : I \rightarrow \mathbb{R}^n$ è soluzione di (1.7); viceversa, se $(u_1, \dots, u_n) : I \rightarrow \mathbb{R}^n$ è soluzione di (1.7), allora $u = u_1$ è soluzione di (1.6).

Torniamo all'esempio più semplice, e cerchiamo di rispondere a questa domanda: tra le soluzioni dell'equazione $u' = u$, quali soddisfano anche la condizione $u(0) = 1$? Tra le soluzioni $u(t) = ce^t$ dell'equazione $u' = u$ ce n'è una sola che soddisfa $u(0) = 1$, ed è quella che si ottiene per $c = 1$, cioè $u(t) = e^t$. Domanda simile:

$$\begin{cases} u' = u \\ u(2) = -3 \end{cases}$$

ha soluzione $u(t) = -3e^{t-2}$, ottenuta per $c = -3e^{-2}$. Se la condizione è $u(4) = 0$, la soluzione è quella costante $u(t) = 0$.

Definizione 1.8. (*Problema di Cauchy*). Si chiama *problema di Cauchy* del primo ordine il problema

$$\begin{cases} u' = f(t, u) & \text{(equazione differenziale)} \\ u(t_0) = y_0 & \text{(condizione iniziale),} \end{cases} \quad (1.8)$$

con $t_0 =$ istante iniziale, $y_0 =$ dato iniziale. La soluzione $u(t)$ deve essere definita su un intervallo I che contiene t_0 .

Questa definizione vale anche per le equazioni differenziali in \mathbb{R}^n , con $t_0 \in I \subseteq \mathbb{R}$, $u : I \rightarrow \mathbb{R}^n$, $f(t, u) \in \mathbb{R}^n$, e $y_0 \in \mathbb{R}^n$.

Il problema di Cauchy (1.8) può essere riformulato come equazione integrale (equazione contenente un integrale invece di una derivata). Tale formulazione integrale verrà usata più volte nel seguito, ed è alla base del metodo delle "iterate di Picard" per la costruzione della soluzione. Prima ci serve la definizione di integrale di una funzione vettoriale.

Definizione 1.9. (*Integrale di funzione a valori in \mathbb{R}^N*). Sia $f : [a, b] \rightarrow \mathbb{R}^N$, $f(t) = (f_1(t), \dots, f_N(t))$ una funzione a valori in \mathbb{R}^N , e supponiamo che ogni funzione componente f_k sia integrabile sull'intervallo $[a, b]$. Allora $\int_a^b f(t) dt$ indica il vettore avente $\int_a^b f_k(t) dt$ come componente k -esima,

$$\int_a^b f(t) dt := \begin{pmatrix} \int_a^b f_1(t) dt \\ \vdots \\ \int_a^b f_N(t) dt \end{pmatrix}.$$

Lemma 1.10. (*Formulazione integrale del problema di Cauchy*). Sia $\Omega \subseteq \mathbb{R}^{1+N}$ un aperto, $f : \Omega \rightarrow \mathbb{R}^N$ una funzione continua, $(t_0, y_0) \in \Omega$. Sia $I \subseteq \mathbb{R}$ un intervallo, con $t_0 \in I$, e sia $u : I \rightarrow \mathbb{R}^N$ una funzione. Si ha

$$\begin{cases} u \in C^1(I) \\ (t, u(t)) \in \Omega \quad \forall t \in I \\ u'(t) = f(t, u(t)) \quad \forall t \in I \\ u(t_0) = y_0 \end{cases} \quad (1.9)$$

se e solo se

$$\begin{cases} u \in C^0(I) \\ (t, u(t)) \in \Omega \quad \forall t \in I \\ u(t) = y_0 + \int_{t_0}^t f(s, u(s)) ds \quad \forall t \in I. \end{cases} \quad (1.10)$$

Dimostrazione. Supponiamo che valga (1.9). Essendo u' continua, dal Teorema fondamentale del calcolo integrale (applicato su ciascuna componente $u_k(t)$ del vettore $u(t)$) si ha

$$u(t) - u(t_0) = \int_{t_0}^t u'(s) ds.$$

Per ipotesi, $u(t_0) = y_0$ e $u'(s) = f(s, u(s))$, da cui, sostituendo, segue (1.10).

Viceversa, supponiamo che valga (1.10). Essendo f e u continue per ipotesi, la funzione $f(t, u(t))$ è continua. Dunque la funzione

$$t \mapsto \int_{t_0}^t f(s, u(s)) ds$$

è di classe C^1 , e, ancora dal Teorema fondamentale del calcolo integrale, la sua derivata è $f(t, u(t))$. Quindi la funzione

$$t \mapsto y_0 + \int_{t_0}^t f(s, u(s)) ds$$

(che, per (1.10), è uguale a $u(t)$) è di classe C^1 , la sua derivata è $f(t, u(t))$, e in $t = t_0$ vale y_0 , da cui otteniamo (1.9). \square

2 Lezione 2

Teorema 2.1. (Teorema di esistenza e unicità locale per il problema di Cauchy).

Sia $t_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^N$,

$$I = [t_0 - a, t_0 + a], \quad J = \{y \in \mathbb{R}^N : |y - y_0| \leq b\} = \overline{B(y_0, b)},$$

con $a > 0$, $b > 0$. Sia $f : I \times J \rightarrow \mathbb{R}^N$ una funzione tale che

- f è continua in $I \times J$;
- $f(t, y)$ è lipschitziana in y , uniformemente in t , su $I \times J$, cioè esiste $L \geq 0$ tale che

$$|f(t, y) - f(t, z)| \leq L|y - z| \quad \forall t \in I, \quad \forall y, z \in J. \quad (2.1)$$

Siano M, δ tali che

$$|f(t, y)| \leq M \quad \forall (t, y) \in I \times J \quad (2.2)$$

e

$$0 < \delta \leq a, \quad M\delta \leq b. \quad (2.3)$$

Sia

$$I_0 = [\tau_1, \tau_2] \subseteq [t_0 - \delta, t_0 + \delta], \quad t_0 \in I_0.$$

Allora il problema di Cauchy

$$\begin{cases} u' = f(t, u) \\ u(t_0) = y_0 \end{cases} \quad (2.4)$$

ha una e una sola soluzione definita nell'intervallo I_0 .

Si ha inoltre la stima

$$|u(t) - y_0| \leq M\delta \frac{e^{\delta L} - 1}{\delta L} \quad \forall t \in I_0 \quad (2.5)$$

se $L > 0$; se $L = 0$ si ha $|u(t) - y_0| \leq M\delta$ per ogni $t \in I_0$.

Dimostrazione. Supponiamo che $L > 0$, e rimandiamo alla fine della dimostrazione il caso (banale) in cui $L = 0$.

Sappiamo (Lemma 1.10) che trovare una soluzione del problema di Cauchy (2.4) è equivalente a trovare una soluzione del problema integrale

$$u(t) = y_0 + \int_{t_0}^t f(s, u(s)) ds \quad \forall t \in I_0. \quad (2.6)$$

Costruiamo una soluzione u di (2.6) come limite di una successione di funzioni.

Passo 1: costruzione della successione (u_n) . Definiamo

$$u_0(t) := y_0 \quad \forall t \in I_0.$$

Per $n \geq 0$, vogliamo definire ricorsivamente

$$u_{n+1}(t) := y_0 + \int_{t_0}^t f(s, u_n(s)) ds \quad \forall t \in I_0. \quad (2.7)$$

L'integrale in (2.7) è ben posto se, per ogni $s \in I_0$, la coppia $(s, u_n(s))$ appartiene all'insieme $I \times J$ dove f è definita, e se la funzione composta $f(s, u_n(s))$ è integrabile. Essendo $I_0 \subseteq I$, per ogni $s \in I_0$ si ha $(s, u_n(s)) \in I \times J$ se e solo se $u_n(s) \in J$. Dimostriamo dunque per induzione che:

Per ogni $n \geq 0$ si ha

$$u_n \in C^1(I_0), \quad u_n(t) \in J \quad \forall t \in I_0 \quad (2.8)$$

e, di conseguenza, possiamo definire u_{n+1} come in (2.7).

Per $n = 0$ la (2.8) è banalmente vera perché $u_0(t) = y_0$. Supponiamo che (2.8) valga per un certo $n \geq 0$, e dimostriamo che allora vale anche per $n + 1$. Dall'ipotesi induttiva, $(t, u_n(t)) \in I \times J$ per ogni $t \in I_0$, $f(t, u_n(t))$ è ben definita e continua, quindi integrabile. Dunque definiamo u_{n+1} come in (2.7). Si ha $u_{n+1} \in C^1(I_0)$ perché $f(s, u_n(s))$ è continua. Poiché $|f| \leq M$ in $I \times J$ e $M\delta \leq b$, per ogni $t \in I_0$ si ha

$$|u_{n+1}(t) - y_0| = \left| \int_{t_0}^t f(s, u_n(s)) ds \right| \leq \left| \int_{t_0}^t |f(s, u_n(s))| ds \right| \leq M|t - t_0| \leq M\delta \leq b.$$

Dunque $u_{n+1}(t) \in J$ per ogni $t \in I_0$, e (2.8) vale anche per u_{n+1} .

In questo modo abbiamo definito una successione (u_n) che soddisfa (2.8) e (2.7) per ogni $n \geq 0$.

Passo 2: stima di $|u_n - u_{n-1}|$. Dimostriamo, per induzione, che

$$|u_n(t) - u_{n-1}(t)| \leq \frac{M}{L} \frac{L^n |t - t_0|^n}{n!} \quad \forall t \in I_0, \quad n \geq 1. \quad (2.9)$$

Per ogni $t \in I_0$ si ha

$$|u_1(t) - u_0(t)| = |u_1(t) - y_0| = \left| \int_{t_0}^t f(s, y_0) ds \right| \leq M|t - t_0| = \frac{M}{L} \frac{L^1 |t - t_0|^1}{1!},$$

quindi (2.9) è vera per $n = 1$. Supponiamo che (2.9) valga per un certo $n \geq 1$. Allora per $t \in I_0$, $t \geq t_0$ si ha

$$\begin{aligned} |u_{n+1}(t) - u_n(t)| &= \left| y_0 + \int_{t_0}^t f(s, u_n(s)) ds - y_0 - \int_{t_0}^t f(s, u_{n-1}(s)) ds \right| \\ &= \left| \int_{t_0}^t (f(s, u_n(s)) - f(s, u_{n-1}(s))) ds \right| \\ &\leq \int_{t_0}^t |f(s, u_n(s)) - f(s, u_{n-1}(s))| ds \\ &\leq \int_{t_0}^t L |u_n(s) - u_{n-1}(s)| ds \\ &\leq \int_{t_0}^t L \frac{M}{L} \frac{L^n |s - t_0|^n}{n!} ds = \frac{M}{L} \frac{L^{n+1}}{n!} \int_{t_0}^t (s - t_0)^n ds = \frac{M}{L} \frac{L^{n+1}}{n!} \frac{(t - t_0)^{n+1}}{n+1} \\ &= \frac{M}{L} \frac{L^{n+1} |t - t_0|^{n+1}}{(n+1)!}. \end{aligned}$$

Per $t \in I_0$, $t \leq t_0$, similmente,

$$\begin{aligned}
|u_{n+1}(t) - u_n(t)| &= \left| \int_{t_0}^t \left(f(s, u_n(s)) - f(s, u_{n-1}(s)) \right) ds \right| \\
&= \left| \int_t^{t_0} \left(f(s, u_n(s)) - f(s, u_{n-1}(s)) \right) ds \right| \\
&\leq \int_t^{t_0} |f(s, u_n(s)) - f(s, u_{n-1}(s))| ds \\
&\leq \int_t^{t_0} L |u_n(s) - u_{n-1}(s)| ds \\
&\leq \int_t^{t_0} L \frac{M}{L} \frac{L^n |s - t_0|^n}{n!} ds = \frac{M}{L} \frac{L^{n+1}}{n!} \int_t^{t_0} (t_0 - s)^n ds = \frac{M}{L} \frac{L^{n+1}}{n!} \frac{(t_0 - t)^{n+1}}{n+1} \\
&= \frac{M}{L} \frac{L^{n+1} |t - t_0|^{n+1}}{(n+1)!}.
\end{aligned}$$

Quindi (2.9) vale per $n+1$, per ogni $t \in I_0$. Per induzione, (2.9) vale per ogni $n \geq 1$.

Passo 3: (u_n) converge uniformemente in I_0 . Scriviamo ogni u_n come somma parziale di una serie telescopica,

$$\begin{aligned}
u_n(t) &= u_0(t) + (u_1(t) - u_0(t)) + (u_2(t) - u_1(t)) + \dots + (u_n(t) - u_{n-1}(t)) \\
&= y_0 + \sum_{k=1}^n (u_k(t) - u_{k-1}(t)).
\end{aligned} \tag{2.10}$$

Dalla stima (2.9), ogni termine soddisfa

$$|u_k(t) - u_{k-1}(t)| \leq \frac{M}{L} \frac{L^k |t - t_0|^k}{k!} \leq \frac{M}{L} \frac{L^k \delta^k}{k!} \quad \forall t \in I_0, k \geq 1. \tag{2.11}$$

Quindi

$$\|u_k - u_{k-1}\|_{\infty, I_0} := \sup\{|u_k(t) - u_{k-1}(t)| : t \in I_0\} \leq \frac{M}{L} \frac{(L\delta)^k}{k!} \quad \forall k \geq 1. \tag{2.12}$$

La serie numerica di termine generale $\frac{M}{L} \frac{(L\delta)^k}{k!}$ è convergente, perché

$$\sum_{k=1}^{\infty} \frac{M}{L} \frac{(L\delta)^k}{k!} = \frac{M}{L} \left(\left(\sum_{k=0}^{\infty} \frac{(L\delta)^k}{k!} \right) - 1 \right) = \frac{M}{L} (e^{L\delta} - 1) < \infty.$$

Quindi, dalla (2.12), anche la serie numerica $\sum_{k=1}^{\infty} \|u_k - u_{k-1}\|_{\infty, I_0}$ converge, e la serie di funzioni $\sum_{k=1}^{\infty} (u_k(t) - u_{k-1}(t))$ è totalmente convergente in I_0 . Di conseguenza (criterio di Weierstrass: la convergenza totale implica quella uniforme) la serie di funzioni converge uniformemente in I_0 , cioè la successione delle sue somme parziali converge uniformemente in I_0 . Quindi, per la (2.10), u_n converge uniformemente in I_0 . Chiamiamo u il suo limite, $u(t) := \lim_{(n \rightarrow \infty)} u_n(t)$.

Per ogni $t \in I_0$ sappiamo che $u_n(t) \in J$, cioè $|u_n(t) - y_0| \leq b$; perciò, passando al limite,

$$\lim_{n \rightarrow \infty} |u_n(t) - y_0| = |u(t) - y_0| \leq b,$$

cioè $u(t) \in J$. Dunque $(t, u(t)) \in I \times J$ per ogni $t \in I_0$.

Inoltre per ogni $t \in I_0$ si ha

$$\begin{aligned} |u_n(t) - y_0| &= \left| \sum_{k=1}^n (u_k(t) - u_{k-1}(t)) \right| \leq \sum_{k=1}^n |u_k(t) - u_{k-1}(t)| \\ &\leq \sum_{k=1}^n \frac{M}{L} \frac{(L\delta)^k}{k!} \leq \sum_{k=1}^{\infty} \frac{M}{L} \frac{(L\delta)^k}{k!} = \frac{M}{L} (e^{L\delta} - 1). \end{aligned}$$

Passando al limite per $n \rightarrow \infty$ nella disuguaglianza $|u_n(t) - y_0| \leq \frac{M}{L} (e^{L\delta} - 1)$ appena dimostrata si ottiene la (2.5).

Passo 4: esistenza della soluzione. Abbiamo dimostrato che $u_n \rightarrow u$ uniformemente in I_0 , cioè $\|u_n - u\|_{\infty, I_0} \rightarrow 0$ per $n \rightarrow \infty$. Per ogni $t \in I_0$ si ha

$$|f(t, u_n(t)) - f(t, u(t))| \leq L|u_n(t) - u(t)| \leq L\|u_n - u\|_{\infty, I_0}. \quad (2.13)$$

Passando al sup nella disuguaglianza $|f(t, u_n(t)) - f(t, u(t))| \leq L\|u_n - u\|_{\infty, I_0}$ appena provata otteniamo

$$\sup_{t \in I_0} |f(t, u_n(t)) - f(t, u(t))| \leq L\|u_n - u\|_{\infty, I_0} \rightarrow 0 \quad (n \rightarrow \infty),$$

cioè la funzione $f(t, u_n(t))$ converge alla funzione $f(t, u(t))$ uniformemente in I_0 . Grazie alla convergenza uniforme, si ha il passaggio al limite sotto il segno di integrale,

$$\lim_{n \rightarrow \infty} \left(\int_{t_0}^t f(s, u_n(s)) ds \right) = \int_{t_0}^t f(s, u(s)) ds \quad \forall t \in I_0.$$

Ricordando la (2.7), per ogni $t \in I_0$ abbiamo

$$u(t) = \lim_{n \rightarrow \infty} u_n(t) = \lim_{n \rightarrow \infty} \left(y_0 + \int_{t_0}^t f(s, u_{n-1}(s)) ds \right) = y_0 + \int_{t_0}^t f(s, u(s)) ds,$$

cioè u soddisfa l'equazione integrale (2.6).

Passo 5: unicità della soluzione. Supponiamo che u e v siano due soluzioni su I_0 dell'equazione integrale (2.6), cioè u, v continue in I_0 , con

$$u(t) = y_0 + \int_{t_0}^t f(s, u(s)) ds, \quad v(t) = y_0 + \int_{t_0}^t f(s, v(s)) ds \quad \forall t \in I_0. \quad (2.14)$$

In $t = t_0$ le due soluzioni coincidono; dimostriamo che esse coincidono su tutto l'intervallo $I_0 = [\tau_1, \tau_2]$. In particolare, proviamo che $u = v$ in $[t_0, \tau_2]$ (la dimostrazione per l'intervallo $[\tau_1, t_0]$ è simile). Sia

$$E := \{c \in [t_0, \tau_2] : u(t) = v(t) \forall t \in [t_0, c]\}.$$

L'insieme E è non vuoto perché $t_0 \in E$. Definiamo

$$p := \sup E.$$

Per costruzione, $p \in [t_0, \tau_2]$. Supponiamo, per assurdo, che sia $p < \tau_2$. Fissiamo $\delta_1 > 0$ abbastanza piccolo da soddisfare

$$L\delta_1 \leq \frac{1}{2}, \quad p + \delta_1 \leq \tau_2.$$

In particolare, la seconda disuguaglianza assicura che $[p, p + \delta_1] \subseteq [t_0, \tau_2]$. Sia

$$\lambda := \sup\{|u(t) - v(t)| : t \in [p, p + \delta_1]\}.$$

Usando le identità in (2.14) e il fatto che, per costruzione, $u = v$ su $[t_0, p]$, si ha, per ogni $t \in [p, p + \delta_1]$,

$$\begin{aligned} u(t) - v(t) &= \int_{t_0}^t (f(s, u(s)) - f(s, v(s))) ds = \int_{t_0}^p (\dots) ds + \int_p^t (\dots) ds \\ &= \int_p^t (f(s, u(s)) - f(s, v(s))) ds. \end{aligned}$$

Dunque, per ogni $t \in [p, p + \delta_1]$,

$$\begin{aligned} |u(t) - v(t)| &\leq \int_p^t |f(s, u(s)) - f(s, v(s))| ds \\ &\leq \int_p^t L|u(s) - v(s)| ds \\ &\leq \int_p^t L\lambda ds = L\lambda(t - p) \leq L\lambda\delta_1 \leq \frac{1}{2}\lambda. \end{aligned}$$

Abbiamo provato che $|u(t) - v(t)| \leq \frac{1}{2}\lambda$ per ogni $t \in [p, p + \delta_1]$. Di conseguenza

$$\sup\{|u(t) - v(t)| : t \in [p, p + \delta_1]\} \leq \frac{1}{2}\lambda,$$

cioè $\lambda \leq \frac{1}{2}\lambda$. Questo implica che $\lambda = 0$, da cui $u = v$ in $[p, p + \delta_1]$. D'altra parte, come già osservato, $u = v$ su $[t_0, p]$. Quindi $u = v$ in tutto l'intervallo $[t_0, p + \delta_1]$. Dunque $p + \delta_1 \in E$, ma $p + \delta_1 > p = \sup E$, assurdo. Abbiamo dunque dimostrato che

$$\sup E = p = \tau_2.$$

Poiché $\sup E = \tau_2$, esiste una successione (c_n) di elementi di E che converge a τ_2 . Per definizione di E si ha dunque $u = v$ in $[t_0, c_n]$ per ogni n , quindi $u = v$ in tutto $[t_0, \tau_2]$ (notiamo che u, v coincidono anche nell'estremo: da $u(c_n) = v(c_n)$, passando al limite, si ottiene $u(\tau_2) = v(\tau_2)$).

Analogamente si prova che $u = v$ in $[\tau_1, t_0]$.

Passo 6: caso banale $L = 0$. Concludiamo la dimostrazione trattando il caso in cui L sia nullo. Dall'ipotesi (2.1) con $L = 0$ segue che $f(t, y) = f(t, y_0)$ per ogni $y \in J$. L'equazione in (2.4) diventa dunque $u'(t) = f(t, y_0)$ (campo vettoriale indipendente da u), e, per integrazione, l'unica soluzione è

$$u(t) = y_0 + \int_{t_0}^t f(s, y_0) ds \quad \forall t \in I_0.$$

Essendo $|f(s, y_0)| \leq M$ per ogni $s \in I_0$, si ha

$$|u(t) - y_0| \leq \left| \int_{t_0}^t f(s, y_0) ds \right| \leq M|t - t_0| \leq M\delta,$$

e la dimostrazione è conclusa. □

Osservazione 2.2. (*Iterate di Picard*). La successione $(u_n(t))$ costruita ricorsivamente nella dimostrazione appena conclusa si chiama successione delle *iterate di Picard*.

Osservazione 2.3. (*Problema di punto fisso*). Per ogni funzione $u(t)$ sia

$$\Phi(u)(t) := y_0 + \int_{t_0}^t f(s, u(s)) ds.$$

L'equazione (2.6) si riscrive come

$$u = \Phi(u),$$

che è un *problema di punto fisso*: risolvere il problema di Cauchy nella formulazione integrale (2.6) significa trovare u che sia “punto fisso” di Φ , cioè tale che $\Phi(u) = u$.

In questo tipo di problemi l'approccio classico consiste nel costruire una successione (u_n) definita ricorsivamente ponendo

$$u_{n+1} = \Phi(u_n). \tag{2.15}$$

Se si riesce a dimostrare che u_n converge, in una norma opportuna, ad un certo limite u , e se $\Phi(u_n)$ converge a $\Phi(u)$, allora passando al limite nella (2.15) si ottiene $u = \Phi(u)$, la soluzione cercata. La dimostrazione del Teorema 2.1 segue questo approccio (vedere la definizione ricorsiva (2.7) della successione (u_n) per risolvere il problema (2.6)).

Esercizio 2.4. (*Domanda riguardo alla dimostrazione del Teorema 2.1*). Nella dimostrazione del teorema abbiamo dapprima dimostrato la disuguaglianza (2.9) in cui la differenza $|u_n(t) - u_{n-1}(t)|$ tra un termine della successione e il suo precedente viene stimata in termini di $|t - t_0|^n$, e poi, nella (2.11), abbiamo usato la disuguaglianza (2.9) per stimare nuovamente la stessa differenza $|u_n(t) - u_{n-1}(t)|$, questa volta però maggiorando $|t - t_0|^n$ con δ^n .

Perché non abbiamo maggiorato $|t - t_0|$ con δ già nella (2.9) invece che farlo in un secondo momento? Cosa sarebbe cambiato nel risultato finale?

3 Lezione 3

Esercizio 3.1. (*Convergenza uniforme*). Nella dimostrazione del Teorema 2.1 (vedi (2.13)) abbiamo usato la lipschitzianità di f per dedurre la convergenza uniforme di $f(t, u_n(t))$ a $f(t, u(t))$. Tuttavia la sola continuità di f sarebbe stata sufficiente allo scopo. Dimostrare dunque la seguente proposizione.

Sia $I_0 = [\tau_1, \tau_2]$, $J = \{y \in \mathbb{R}^N : |y - y_0| \leq b\}$. Sia $f : I_0 \times J \rightarrow \mathbb{R}^N$ continua, e sia $u_n : I_0 \rightarrow J$ una successione di funzioni continue che convergono uniformemente alla funzione $u : I_0 \rightarrow J$. Sia $g_n(t) := f(t, u_n(t))$ e $g(t) := f(t, u(t))$. Allora $g_n \rightarrow g$ uniformemente in I_0 .

Svolgimento dell'Esercizio 3.1. Sia $\varepsilon > 0$. La funzione f è continua, l'insieme $I \times J$ è compatto, quindi (Teorema di Cantor) f è uniformemente continua in $I \times J$. Dunque esiste $\delta > 0$ tale che

$$|f(t, y) - f(s, z)| < \varepsilon \quad \forall (t, y), (s, z) \in I \times J, \quad |(t, y) - (s, z)| < \delta.$$

Per ipotesi, $u_n \rightarrow u$ uniformemente in I . Quindi esiste $\bar{n} \in \mathbb{N}$ tale che

$$\|u_n - u\|_{\infty, I} = \sup_{t \in I} |u_n(t) - u(t)| < \delta \quad \forall n \geq \bar{n}.$$

Di conseguenza, per ogni $n \geq \bar{n}$, per ogni $t \in I$ si ha

$$|(t, u_n(t)) - (t, u(t))| = |u_n(t) - u(t)| < \delta$$

e quindi

$$|f(t, u_n(t)) - f(t, u(t))| < \varepsilon \quad \forall t \in I, \quad n \geq \bar{n}.$$

Abbiamo dimostrato che per ogni $\varepsilon > 0$ esiste $\bar{n} \in \mathbb{N}$ tale che $|g_n(t) - g(t)| < \varepsilon$ per ogni $t \in I$, per ogni $n \geq \bar{n}$, cioè che $g_n \rightarrow g$ uniformemente in I . \square

Definizione 3.2. (*Funzione $f(t, y)$ localmente lipschitziana in y uniformemente in t*). Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto. Una funzione $f : \Omega \rightarrow \mathbb{R}^N$ si dice *localmente lipschitziana in y uniformemente in t* se per ogni punto $(t_0, y_0) \in \Omega$ esistono $a > 0$, $b > 0$, $L \geq 0$ tali che

- il “rettangolo” $I \times J = [t_0 - a, t_0 + a] \times \overline{B(y_0, b)}$ è contenuto in Ω ;
- $|f(t, y) - f(t, z)| \leq L|y - z|$ per ogni $t \in I$, per ogni $y, z \in J$.

Osservazione 3.3. (*Esistenza e unicità locale su aperti*). Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ una funzione continua e localmente lipschitziana in y uniformemente in t . Allora per ogni $(t_0, y_0) \in \Omega$ il problema di Cauchy

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

ha una, e solo una, *soluzione locale*, cioè esiste $\delta > 0$ tale che in ogni intervallo $I_0 = [\tau_1, \tau_2] \subseteq [t_0 - \delta, t_0 + \delta]$, con $t_0 \in I_0$, il problema di Cauchy ha un'unica soluzione in I_0 .

In altre parole, le ipotesi di continuità e locale lipschitzianità in y della funzione f assicurano che intorno a ogni punto (t_0, y_0) dell'aperto Ω c'è un rettangolo $I \times J$ centrato in (t_0, y_0) in cui possiamo applicare il Teorema di esistenza locale e unicità per il problema di Cauchy.

Nella prossima proposizione (Proposizione 3.10) osserviamo che se f è di classe C^1 , o anche solo se f è continua e ha, continue, le derivate parziali rispetto alle variabili y , allora f è localmente lipschitziana in y . Questo permette di applicare il Teorema di esistenza locale e unicità senza fare verifiche dirette di lipschitzianità per f (e questo è particolarmente conveniente nello svolgimento degli esercizi).

Prima di enunciare e dimostrare la proposizione, apriamo una parentesi sulle norme delle matrici e ricordiamo come la disuguaglianza di Cauchy-Schwarz in \mathbb{R}^N ,

$$|x \cdot y| \leq |x||y| \quad \forall x, y \in \mathbb{R}^N,$$

si può applicare al prodotto righe per colonne di una matrice contro un vettore e all'integrale di una funzione vettoriale.

Ricordiamo che con $|x|$ indichiamo la norma euclidea di un vettore $x \in \mathbb{R}^N$, cioè

$$|x| = \sqrt{x_1^2 + \dots + x_N^2} = \left(\sum_{k=1}^N x_k^2 \right)^{\frac{1}{2}}, \quad x = (x_1, \dots, x_N) \in \mathbb{R}^N.$$

Definizione 3.4. (*Norma euclidea di una matrice*). Sia A una matrice a n righe e q colonne,

$$A = \begin{pmatrix} A_{11} & \dots & A_{1q} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nq} \end{pmatrix}.$$

La norma euclidea della matrice A è

$$|A| = |A|_E = |A|_{\text{Euclidea}} := \left(\sum_{\substack{k=1, \dots, n \\ j=1, \dots, q}} A_{kj}^2 \right)^{\frac{1}{2}}.$$

Lemma 3.5. (*Disuguaglianza di Cauchy-Schwarz per matrice contro vettore*). Per ogni A matrice $n \times q$, per ogni vettore $x \in \mathbb{R}^q$ si ha

$$|Ax| \leq |A||x|$$

dove $|A|$ è la norma euclidea di A .

Dimostrazione. Sia $y = Ax$, e siano y_k le sue componenti. Allora, per ogni $k = 1, \dots, n$, si ha

$$y_k = (A_{k1} \quad \dots \quad A_{kq}) \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix} = \sum_{j=1}^q A_{kj} x_j = v_k \cdot x, \quad v_k := \begin{pmatrix} A_{k1} \\ \vdots \\ A_{kq} \end{pmatrix}$$

(il vettore v_k è la riga k -esima di A scritta come colonna). Dalla disuguaglianza di Cauchy-Schwarz,

$$|y_k| = |v_k \cdot x| \leq |v_k||x|,$$

da cui

$$\begin{aligned} |Ax| = |y| &= \left(\sum_{k=1}^n y_k^2 \right)^{\frac{1}{2}} \\ &\leq \left(\sum_{k=1}^n |v_k|^2 |x|^2 \right)^{\frac{1}{2}} = \left(\sum_{k=1}^n |v_k|^2 \right)^{\frac{1}{2}} |x| = \left(\sum_{k=1}^n \sum_{j=1}^q A_{kj}^2 \right)^{\frac{1}{2}} |x| = |A||x|. \quad \square \end{aligned}$$

Il Lemma 3.6 è stato dimostrato in Analisi 2; per completezza, ne riportiamo comunque la dimostrazione.

Lemma 3.6. (Disuguaglianza per l'integrale di funzione vettoriale). *Sia $[a, b] \subset \mathbb{R}$ un intervallo, e sia $g : [a, b] \rightarrow \mathbb{R}^N$, $g(t) = (g_1(t), \dots, g_N(t))$ una funzione a valori vettoriali. Supponiamo che le funzioni componenti $g_k : [a, b] \rightarrow \mathbb{R}$, $k = 1, \dots, N$, siano integrabili in $[a, b]$. Allora*

$$\left| \int_a^b g(t) dt \right| \leq \int_a^b |g(t)| dt.$$

Dimostrazione. Sia $c_k := \int_a^b g_k(t) dt$, dove g_k è la funzione componente k -esima di g , e sia c il vettore $(c_1, \dots, c_N) \in \mathbb{R}^N$, cioè $c = \int_a^b g(t) dt$. La tesi è $|c| \leq \int_a^b |g(t)| dt$. Se $c = 0$, la tesi è banalmente vera. Sia dunque $c \neq 0$. Allora

$$\begin{aligned} |c|^2 &= \sum_{k=1}^N c_k^2 = \sum_{k=1}^N \int_a^b g_k(t) dt c_k = \int_a^b \sum_{k=1}^N g_k(t) c_k dt = \int_a^b g(t) \cdot c dt \\ &\leq \int_a^b |g(t)| |c| dt = |c| \int_a^b |g(t)| dt. \end{aligned}$$

Poiché $|c| > 0$, la disuguaglianza $|c|^2 \leq |c| \int_a^b |g(t)| dt$ dà la tesi. \square

Scritta in componenti, la disuguaglianza del Lemma 3.6 diventa

$$\left\{ \sum_{k=1}^N \left(\int_a^b g_k(t) dt \right)^2 \right\}^{\frac{1}{2}} \leq \int_a^b \left(\sum_{k=1}^N g_k^2(t) \right)^{\frac{1}{2}} dt.$$

Definizione 3.7. (*Norma operatoriale di matrice*). Sia A una matrice a n righe e q colonne. La norma operatoriale della matrice A è

$$|A|_{\text{op}} = |A|_{\text{operatoriale}} := \sup \left\{ \frac{|Ax|}{|x|} : x \in \mathbb{R}^q, x \neq 0 \right\}.$$

Esercizio 3.8. (*Formule equivalenti della norma operatoriale*). Dimostrare che per ogni matrice $A \in \text{Mat}_{n \times q}(\mathbb{R})$ si ha

$$\sup_{x \neq 0} \frac{|Ax|}{|x|} = \sup_{|x|=1} |Ax| = \sup_{|x| \leq 1} |Ax|.$$

Esercizio 3.9. (*Equivalenza di norma euclidea e operatoriale*). Dimostrare che la norma euclidea $|A|_{\text{E}}$ e quella operatoriale $|A|_{\text{op}}$ sono equivalenti, cioè dimostrare che esistono due costanti C_1, C_2 (dipendenti da n, q) tali che

$$C_1 |A|_{\text{E}} \leq |A|_{\text{op}} \leq C_2 |A|_{\text{E}}$$

per ogni matrice $A \in \text{Mat}_{n \times q}(\mathbb{R})$. Calcolare inoltre le costanti ottimali C_1, C_2 , cioè calcolare

$$\begin{aligned} \min\{C_2 > 0 : |A|_{\text{op}} \leq C_2 |A|_{\text{E}} \quad \forall A \in \text{Mat}_{n \times q}(\mathbb{R})\}, \\ \max\{C_1 > 0 : C_1 |A|_{\text{E}} \leq |A|_{\text{op}} \quad \forall A \in \text{Mat}_{n \times q}(\mathbb{R})\}. \end{aligned}$$

Proposizione 3.10. (Criterio di lipschitzianità). Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto, sia $f : \Omega \rightarrow \mathbb{R}^N$ una funzione continua, e supponiamo che esistano, continue, le derivate parziali di f rispetto alle variabili y_1, \dots, y_N . Allora f è localmente lipschitziana in y uniformemente in t .

Dimostrazione. Sia $(t_0, y_0) \in \Omega$. Poiché Ω è un aperto, esistono $a > 0$, $b > 0$ tali che $I \times J \subset \Omega$, dove $I = [t_0 - a, t_0 + a]$, $J = \overline{B}(y_0, b) = \{y \in \mathbb{R}^N : |y - y_0| \leq b\}$.

Sia $D_y f(t, y)$ la matrice delle derivate parziali di f rispetto a y (matrice jacobiana della funzione $y \mapsto f(t, y)$, per t fissato), cioè

$$D_y f(t, y) = \begin{pmatrix} \partial_{y_1} f_1(t, y) & \dots & \partial_{y_N} f_1(t, y) \\ \vdots & \ddots & \vdots \\ \partial_{y_1} f_N(t, y) & \dots & \partial_{y_N} f_N(t, y) \end{pmatrix}$$

dove f_1, \dots, f_N sono le funzioni componenti di f . Dal Teorema di Weierstrass, esiste finito il massimo

$$M := \max\{|D_y f(t, y)| : (t, y) \in I \times J\},$$

dove $|D_y f(t, y)|$ è la norma euclidea della matrice. Notiamo che, per ogni $t \in I$ fissato, la funzione $J \rightarrow \mathbb{R}^N$, $y \mapsto f(t, y)$ è differenziabile ed è di classe C^1 (Teorema del differenziale, Analisi 2).

Sia $t \in I$, $y, z \in J$. Poiché la palla J è convessa, il segmento di estremi y, z è contenuto in J , cioè $y + s(z - y) \in J$ per ogni $s \in [0, 1]$. Sia

$$\varphi(s) := f(t, y + s(z - y)), \quad s \in [0, 1].$$

Per ogni $k = 1, \dots, N$, la funzione componente

$$\varphi_k : [0, 1] \rightarrow \mathbb{R}, \quad \varphi_k(s) = f_k(t, y + s(z - y))$$

è di classe C^1 (essendo composizione di funzioni C^1), e ha derivata

$$\varphi'_k(s) = \sum_{j=1}^N (\partial_{y_j} f_k)(t, y + s(z - y)) (z_j - y_j).$$

Quindi il vettore $\varphi'(s)$ è

$$\varphi'(s) = (D_y f)(t, y + s(z - y)) (z - y).$$

Dalla disuguaglianza di Cauchy-Schwarz per il prodotto di una matrice contro un vettore (Lemma 3.5),

$$|\varphi'(s)| \leq |(D_y f)(t, y + s(z - y))| |z - y| \leq M |z - y|.$$

Il Teorema fondamentale del calcolo integrale applicato a ogni funzione componente $\varphi_k : [0, 1] \rightarrow \mathbb{R}$ dà

$$\varphi_k(1) - \varphi_k(0) = \int_0^1 \varphi'_k(s) ds \quad \forall k = 1, \dots, N,$$

quindi i vettori $\varphi(1), \varphi(0)$ soddisfano la stessa formula in \mathbb{R}^N ,

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(s) ds.$$

Dalla disuguaglianza per l'integrale di funzione vettoriale (Lemma 3.6) abbiamo

$$|\varphi(1) - \varphi(0)| = \left| \int_0^1 \varphi'(s) ds \right| \leq \int_0^1 |\varphi'(s)| ds \leq \int_0^1 M|z - y| ds = M|z - y|.$$

Poiché $\varphi(1) = f(t, z)$, $\varphi(0) = f(t, y)$, abbiamo provato che per ogni $t \in I$, per ogni $y, z \in J$, si ha $|f(t, z) - f(t, y)| \leq M|z - y|$. \square

Con gli stessi passaggi si dimostra il seguente teorema.

Teorema 3.11. (Teorema del valor medio in \mathbb{R}^N). *Sia $\Omega \subseteq \mathbb{R}^q$ un aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ una funzione di classe $C^1(\Omega)$. Siano x, y due punti di Ω tali che tutto il segmento che li unisce sia contenuto in Ω , cioè*

$$\{x + s(y - x) : s \in [0, 1]\} \subset \Omega.$$

Allora

$$|f(y) - f(x)| \leq \left(\sup_{s \in [0, 1]} |(Df)(x + s(y - x))| \right) |y - x|,$$

dove Df è la matrice jacobiana di f .

Dimostrazione. La funzione $\varphi : [0, 1] \rightarrow \mathbb{R}^N$, $\varphi(s) := f(x + s(y - x))$ è di classe C^1 , quindi, come nella dimostrazione precedente,

$$|\varphi(1) - \varphi(0)| \leq \int_0^1 |\varphi'(s)| ds \leq \int_0^1 |(Df)(x + s(y - x))| |y - x| ds \leq M|x - y|,$$

dove M è il sup di $|Df|$ lungo il segmento. \square

Osservazione 3.12. (C^1 implica Lipschitz). Se $f(t, y)$ è di classe C^1 , allora, in particolare, f è continua ed esistono, continue, tutte le sue derivate parziali rispetto alle variabili y ; dunque, dalla Proposizione 3.10, f è lipschitziana in y uniformemente in t .

Osservazione 3.13. (Il Teorema di Lagrange non vale in dimensione ≥ 2). Per ottenere una stima di Lipschitz come quella del Teorema del valor medio (Teorema 3.11), per funzioni a valori reali il modo più elementare è usare il Teorema di Lagrange: supponiamo che valgano le ipotesi del Teorema 3.11 con $N = 1$, cioè f funzione scalare. La funzione $\varphi(s) = f(x + s(y - x))$ è funzione reale di una variabile reale, di classe C^1 , quindi, dal Teorema di Lagrange, esiste $\vartheta \in (0, 1)$ tale che

$$\varphi(1) - \varphi(0) = \varphi'(\vartheta)(1 - 0) = \varphi'(\vartheta),$$

cioè

$$f(y) - f(x) = (Df)(x + \vartheta(y - x))(y - x),$$

da cui segue che

$$\begin{aligned} |f(y) - f(x)| &\leq |(Df)(x + \vartheta(y - x))| |y - x| \\ &\leq \left(\sup_{s \in [0, 1]} |(Df)(x + s(y - x))| \right) |y - x|. \end{aligned}$$

Se invece f è a valori vettoriali, con componenti $f = (f_1, \dots, f_N)$, $N \geq 2$, allora il Teorema di Lagrange, in generale, non vale. Per questo motivo nella dimostrazione del Teorema 3.11 abbiamo usato il Teorema fondamentale del calcolo integrale

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(s) ds$$

al posto del Teorema di Lagrange.

Vediamo perché il Teorema di Lagrange non vale in \mathbb{R}^N . Sia $\varphi : [a, b] \rightarrow \mathbb{R}^N$ una funzione di classe C^1 a valori vettoriali, $\varphi(s) = (\varphi_1(s), \dots, \varphi_N(s))$. Ogni funzione componente è di classe C^1 ed è una funzione scalare, quindi il Teorema di Lagrange si applica a ciascuna componente: esistono $\vartheta_1, \dots, \vartheta_N \in (a, b)$ tali che

$$\varphi_1(b) - \varphi_1(a) = \varphi_1'(\vartheta_1)(b - a), \quad \dots \quad \varphi_N(b) - \varphi_N(a) = \varphi_N'(\vartheta_N)(b - a).$$

Il punto è che, in generale, $\vartheta_1, \dots, \vartheta_N$ non coincidono. In altre parole, in generale non esiste alcun $\vartheta \in (a, b)$ che soddisfi *simultaneamente* tutte le uguaglianze

$$\begin{cases} \varphi_1(b) - \varphi_1(a) = \varphi_1'(\vartheta)(b - a), \\ \dots \\ \varphi_N(b) - \varphi_N(a) = \varphi_N'(\vartheta)(b - a), \end{cases}$$

cioè l'identità vettoriale $\varphi(b) - \varphi(a) = \varphi'(\vartheta)(b - a)$.

Esempio. Sia

$$\varphi : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \varphi(s) = \begin{pmatrix} \cos s \\ \sin s \end{pmatrix}.$$

L'identità $\varphi(2\pi) - \varphi(0) = \varphi'(\vartheta)(2\pi - 0)$ vale se

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -\sin \vartheta \\ \cos \vartheta \end{pmatrix} 2\pi,$$

e non esiste alcun ϑ che soddisfa simultaneamente le due equazioni $\sin \vartheta = 0$ e $\cos \vartheta = 0$.

Vediamo qualche esempio di applicazione del Teorema di esistenza locale e unicità.

1) $u' = tu^3$. Il campo vettoriale è $f(t, y) = ty^3$, che è di classe $C^\infty(\Omega)$ nel dominio $\Omega = \mathbb{R}^2$. Quindi f è localmente lipschitziana in y , uniformemente in t , intorno a ogni punto $(t_0, y_0) \in \Omega$. Perciò si può applicare il Teorema di esistenza locale e unicità intorno a ogni dato iniziale. Di conseguenza, per ogni coppia $(t_0, y_0) \in \mathbb{R}^2$, il problema di Cauchy

$$\begin{cases} u' = tu^3 \\ u(t_0) = y_0 \end{cases}$$

ha una e una sola soluzione locale (cioè definita in un intervallo $[t_0 - \delta, t_0 + \delta]$ per un certo $\delta > 0$).

2) $u'' = u'^2 - t + \sin(u+t)$. È un'equazione del secondo ordine: per poter applicare il Teorema di esistenza locale e unicità dobbiamo scriverla come sistema del primo ordine:

$$\begin{cases} u' = v \\ v' = v^2 - t + \sin(u+t). \end{cases}$$

Il campo vettoriale è la funzione

$$f : \Omega = \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad f(t, y_1, y_2) = \begin{pmatrix} y_2 \\ y_2^2 - t + \sin(y_1 + t) \end{pmatrix},$$

che è $C^\infty(\mathbb{R}^3)$. Quindi, ragionando come nell'esempio 1), otteniamo che per ogni dato iniziale $(t_0, u_0, v_0) \in \mathbb{R}^3$ il problema di Cauchy

$$\begin{cases} u' = v, \\ v' = v^2 - t + \sin(u + t), \\ u(t_0) = u_0, \\ v(t_0) = v_0 \end{cases}$$

ha un'unica soluzione locale $(u(t), v(t))$ definita in un certo intervallo $[t_0 - \delta, t_0 + \delta]$. Possiamo a questo punto tornare all'equazione scalare del secondo ordine, e concludiamo che per ogni $(t_0, u_0, v_0) \in \mathbb{R}^3$ il problema di Cauchy

$$\begin{cases} u'' = u'^2 - t + \sin(u + t), \\ u(t_0) = u_0, \\ u'(t_0) = v_0 \end{cases}$$

ha un'unica soluzione locale $u(t)$.

3) $u' = |t|e^{u^2}$. Il campo vettoriale è la funzione $f(t, y) = |t|e^{y^2}$, che è continua in (t, y) in $\Omega = \mathbb{R}^2$, è derivabile in y , con derivata parziale $\partial_y f(t, y) = 2yf(t, y)$, anch'essa continua in Ω . Quindi f è localmente lipschitziana in y , uniformemente in t , in Ω . Di conseguenza per ogni $(t_0, y_0) \in \mathbb{R}^2$ il problema di Cauchy con dato iniziale $u(t_0) = y_0$ ha un'unica soluzione locale. (Notiamo che f non è $C^1(\Omega)$).

4) $u' = |u|$. È un'equazione autonoma. Il campo vettoriale è la funzione $f(t, y) = |y|$, che è continua in $\Omega = \mathbb{R}^2$, ma non è derivabile rispetto a y in $y = 0$. Potremmo restringere il dominio a sottoinsiemi aperti ($y > 0$, $y < 0$) in cui f ammette derivate parziali rispetto a y , e poi ragionare per casi ($y > 0$, $y < 0$, $y = 0$), osservando che $u(t) = 0$ è soluzione costante. Tuttavia, in questo caso è più rapido fare una verifica diretta della lipschitzianità: per ogni $t, y, z \in \mathbb{R}$ si ha

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

e quindi il Teorema di esistenza locale e unicità si può applicare intorno a ogni punto $(t_0, y_0) \in \mathbb{R}^2$.

D'altra parte il problema di Cauchy

$$\begin{cases} u' = |u|, \\ u(t_0) = y_0 \end{cases}$$

è molto semplice, e può essere risolto in modo immediato. La sua soluzione è

$$\begin{aligned} u(t) &= y_0 e^{t-t_0} && \text{se } y_0 > 0, \\ u(t) &= 0 && \text{se } y_0 = 0, \\ u(t) &= y_0 e^{t_0-t} && \text{se } y_0 < 0, \end{aligned}$$

e in tutti i casi la soluzione è definita su tutto \mathbb{R} (soluzione “globale”).

5) $u' = \sqrt{|u|}$. Il campo vettoriale è la funzione $f(t, y) = \sqrt{|y|}$, che è continua in $\Omega = \mathbb{R}^2$, ma non è localmente lipschitziana in Ω . Quindi non possiamo applicare il Teorema di esistenza locale e unicità ad un qualunque dato iniziale (t_0, y_0) . Vedremo nelle prossime lezioni cosa succede al problema di Cauchy in assenza dell'ipotesi di locale lipschitzianità.

Esercizio 3.14. (*La radice quadrata non è lipschitziana*). Dimostrare che per ogni $b > 0$ non esiste alcuna costante $L \geq 0$ tale che

$$|\sqrt{|y|} - \sqrt{|z|}| \leq L|y - z| \quad \forall y, z \in [-b, b].$$

Vediamo un importante corollario del Teorema di esistenza e unicità locale.

Corollario 3.15. (Conseguenza dell'unicità: soluzioni distinte non si toccano). *Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto, sia $f : \Omega \rightarrow \mathbb{R}^N$ continua e localmente lipschitziana in y , uniformemente in t , in Ω . Siano $u(t), v(t)$ due funzioni, definite entrambe in un intervallo I , di classe $C^1(I)$, tali che*

$$\begin{cases} (t, u(t)) \in \Omega, \\ u'(t) = f(t, u(t)), \end{cases} \quad \begin{cases} (t, v(t)) \in \Omega, \\ v'(t) = f(t, v(t)), \end{cases} \quad \forall t \in I.$$

Se esiste $t_0 \in I$ tale che $u(t_0) = v(t_0)$, allora $u(t) = v(t)$ per ogni $t \in I$.

Se esiste $t_0 \in I$ tale che $u(t_0) \neq v(t_0)$, allora $u(t) \neq v(t)$ per ogni $t \in I$.

4 Lezione 4

Il Corollario 3.15 non è una conseguenza immediata del Teorema 2.1 di esistenza e unicità locale: il Teorema dà infatti l'unicità della soluzione nell'intervallo $I_\delta = [t_0 - \delta, t_0 + \delta]$ o nei suoi sottointervalli $[\tau_1, \tau_2]$ contenenti t_0 , mentre il Corollario tratta di un intervallo I che, in generale, potrebbe essere più grande di I_δ ; per esempio, potrebbe essere anche $I = \mathbb{R}$.

Dimostrazione del Corollario 3.15. Sia $u(t_0) = v(t_0)$ per un certo $t_0 \in I$. Supponiamo, per assurdo, che esista $t_1 \in I$, $t_1 > t_0$, tale che $u(t_1) \neq v(t_1)$. Sia

$$E = \{t \in [t_0, t_1] : u(t) = v(t)\}.$$

L'insieme E è non vuoto perché $t_0 \in E$. Inoltre $E \subseteq [t_0, t_1]$, e poi $[t_0, t_1] \subseteq I$ perché $t_0, t_1 \in I$ e I è intervallo. Sia

$$p = \sup E.$$

Per costruzione, $p \in [t_0, t_1]$. Per definizione di p , esiste una successione (s_n) di elementi di E che converge a p . Dunque $u(s_n) = v(s_n)$ per ogni n , e, passando al limite, si trova

$$u(p) = v(p)$$

perché u e v sono funzioni continue. D'altra parte $u(t_1) \neq v(t_1)$, quindi $p \neq t_1$. Di conseguenza, $t_0 \leq p < t_1$.

Sia $y_0 := u(p) = v(p)$, e consideriamo il problema di Cauchy con dati iniziali (p, y_0) . Dal Teorema di esistenza e unicità locale, esiste $\delta > 0$ tale che in ogni intervallo $I_0 = [\tau_1, \tau_2] \subseteq [p - \delta, p + \delta]$, con $p \in I_0$, esiste un'unica soluzione del problema di Cauchy con dati iniziali (p, y_0) . Fissiamo $\delta_1 > 0$, $\delta_1 \leq \delta$ tale che $p + \delta_1 \leq t_1$, e sia $I_0 = [p, p + \delta_1]$. Notiamo che $I_0 \subseteq [t_0, t_1]$.

Sia u che v risolvono l'equazione differenziale in I_0 e soddisfano le condizioni iniziali (p, y_0) . Quindi, per unicità della soluzione, $u(t) = v(t)$ per ogni $t \in [p, p + \delta_1]$. In particolare, $p + \delta_1 \in E$, ma $p + \delta_1 > p = \sup E$, assurdo.

Abbiamo dimostrato che non esiste alcun $t_1 \in I$, $t_1 > t_0$, tale che $u(t_1) \neq v(t_1)$. Di conseguenza, $u(t) = v(t)$ per ogni $t \in I$, $t > t_0$.

In modo simile si dimostra che $u(t) = v(t)$ per ogni $t \in I$, $t < t_0$. Dunque u, v coincidono in tutto I .

Infine, sia $u(t_0) \neq v(t_0)$ per un certo $t_0 \in I$. Se esistesse un punto $t_1 \in I$ in cui $u(t_1) = v(t_1)$, allora, per quanto appena dimostrato, u e v coinciderebbero in tutto l'intervallo I , quindi anche in t_0 , assurdo. \square

Osservazione 4.1. (*Versione topologica della dimostrazione del Corollario 3.15.*) La dimostrazione del Corollario 3.15 è basata sulla seguente idea topologica.

L'insieme $A = \{t \in I : u(t) = v(t)\}$ in cui u e v coincidono è un chiuso di I perché è l'anti-immagine di $\{0\}$ in I tramite la funzione continua $u - v$; d'altra parte, dal Teorema di esistenza e unicità locale, intorno a ogni punto $p \in A$ c'è un intervallo $[p - \delta, p + \delta] \cap I$ in cui $u = v$, quindi A è anche aperto di I . Dunque A è aperto, è chiuso, ed è non vuoto; I è connesso, quindi $A = I$.

Vediamo un esempio di applicazione del Corollario 3.15.

Esercizio 4.2. Sia $u : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ la soluzione del problema di Cauchy

$$\begin{cases} u' = u^2 \sin(u), \\ u(t_0) = y_0 \end{cases}$$

con $0 < y_0 < \pi$. Dimostrare che $0 < u(t) < \pi$ per ogni $t \in I$.

Svolgimento. Il campo vettoriale $f(t, y) = y^2 \sin(y)$ è di classe $C^\infty(\mathbb{R}^2)$, quindi si può applicare il Teorema 2.1 di esistenza e unicità locale e il suo Corollario 3.15.

La funzione costante

$$v : \mathbb{R} \rightarrow \mathbb{R}, \quad v(t) = 0$$

è soluzione dell'equazione differenziale $v' = v^2 \sin(v)$ in tutto \mathbb{R} . All'istante t_0 si ha $v(t_0) = 0 \neq u(t_0)$, quindi, per il Corollario 3.15, $v(t) \neq u(t)$ per ogni $t \in I$. Inoltre, poiché $v(t_0) < u(t_0)$, si ha $v(t) < u(t)$ per ogni $t \in I$: se infatti fosse $v(t_1) > u(t_1)$ in un certo $t_1 \in I$, allora la differenza $h(t) = v(t) - u(t)$ sarebbe una funzione continua con valori $h(t_0) < 0$, $h(t_1) > 0$, e dal Teorema di esistenza degli zeri avremmo $h(t_2) = 0$ per un certo $t_2 \in I$, assurdo. Quindi $0 < u(t)$ per ogni $t \in I$.

Osserviamo poi che anche la funzione costante

$$w : \mathbb{R} \rightarrow \mathbb{R}, \quad w(t) = \pi$$

è soluzione dell'equazione differenziale $w' = w^2 \sin(w)$ in \mathbb{R} ; analogamente, ne deduciamo che $u(t) < \pi$ per ogni $t \in I$. \square

Il Corollario 3.15 impedisce che due soluzioni distinte si tocchino. Quindi la soluzione $u(t)$ dell'esercizio 4.2 è imprigionata tra gli equilibri $v(t) = 0$ e $w(t) = \pi$. Vedremo in seguito (criteri di prolungabilità) che in una situazione del genere la soluzione $u(t)$ è definita per ogni $t \in \mathbb{R}$ (esistenza globale).

Osservazione 4.3. (*Soluzioni distinte non si toccano: caso scalare e vettoriale*). Come visto nell'esercizio 4.2, in dimensione $N = 1$ (caso scalare) il Corollario 3.15 vale con le disuguaglianze: siano u, v due soluzioni come nelle ipotesi del Corollario 3.15; se esiste $t_0 \in I$ tale che $u(t_0) < v(t_0)$, allora $u(t) < v(t)$ per ogni $t \in I$.

In dimensione $N \geq 2$ non c'è una simile proprietà.

Osservazione 4.4. (*Due soluzioni distinte possono visitare gli stessi punti dello spazio delle fasi ma a tempi diversi*). Il Corollario 3.15 dice che due soluzioni di un'equazione differenziale o sono distinte ad ogni istante t , o coincidono ad ogni istante t ; non possono cioè trovarsi in due punti diversi dello spazio delle fasi ad un certo istante, e trovarsi poi in uno stesso punto ad un altro istante di tempo.

Questo non significa che due soluzioni distinte non possano visitare gli stessi punti dello spazio: possono farlo, ma in tempi diversi. Per esempio, le funzioni

$$v(t) = e^t, \quad w(t) = 2e^t, \quad t \in \mathbb{R},$$

sono entrambe soluzioni dell'equazione differenziale $u' = u$ in \mathbb{R} ; sia v che w percorrono tutti i punti $y > 0$; d'altra parte, un punto $y > 0$ viene visitato da v al tempo $t = \log y$, e da w al tempo $t = \log(y/2)$.

Lemma 4.5. (Regolarità delle soluzioni). Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ una funzione di classe $C^k(\Omega)$, $k \geq 1$. Allora ogni soluzione dell'equazione $u' = f(t, u)$ è di classe C^{k+1} nell'intervallo in cui è definita. Se $f \in C^\infty(\Omega)$, allora $u \in C^\infty$.

Dimostrazione. Sia $f \in C^k(\Omega)$ per un certo $k \geq 1$. Poiché f è continua, dal Teorema di esistenza locale e unicità (Teorema 2.1), ricordando anche il Lemma di formulazione integrale del problema di Cauchy (Lemma 1.10), sappiamo che la soluzione u è di classe C^1 .

Poiché f e u sono entrambe di classe C^1 , la loro composizione $f(t, u(t))$ è C^1 . Poi $u'(t) = f(t, u(t))$, da cui $u' \in C^1$. Abbiamo provato che $u \in C^2$.

Iteriamo questo ragionamento: supponiamo di avere dimostrato che $u \in C^n$ per un certo $n \geq 2$. Per ipotesi, $f \in C^k$. Se $n \leq k$, allora sia f che u sono di classe C^n , e la composizione $f(t, u(t))$ è C^n . Poi $u'(t) = f(t, u(t))$, quindi $u' \in C^n$, cioè $u \in C^{n+1}$.

In questo modo, per induzione, abbiamo dimostrato che $u \in C^{n+1}$ per ogni $n \leq k$, cioè che $u \in C^{k+1}$. Notiamo che il ragionamento non può proseguire oltre: $u \in C^{k+1}$, ma per ipotesi $f \in C^k$, quindi la composizione è soltanto C^k (in generale non sappiamo se è C^{k+1}).

Se $f \in C^\infty$, l'argomento induttivo non ha la restrizione $n \leq k$ e quindi si ha $u \in C^n$ per ogni $n \in \mathbb{N}$, cioè $u \in C^\infty$. \square

Osservazione 4.6. (*Bootstrap*). Il tipo di procedimento usato nella dimostrazione del Lemma 4.5 viene chiamato *bootstrap*.

* * *

Come osservato nella prima lezione, non esiste una formula generale che risolve ogni equazione differenziale. In alcuni casi significativi, tuttavia, esistono dei metodi per il calcolo esplicito delle soluzioni. Vediamo ora alcune di queste tecniche risolutive.

Equazioni a variabili separabili. Cominciamo con un esempio: consideriamo l'equazione

$$u' = (u - 1)t^2.$$

Notiamo subito che la funzione costante $u : \mathbb{R} \rightarrow \mathbb{R}$, $u(t) = 1$ è soluzione (equilibrio).

Osserviamo anche che il campo è $f(t, y) = (y - 1)t^2$, definito in $(t, y) \in \Omega = \mathbb{R}^2$, di classe $C^\infty(\mathbb{R}^2)$, quindi vale il Teorema di esistenza locale e unicità.

Come conseguenza dell'unicità, se una soluzione u , diversa dall'equilibrio 1, soddisfa $u(t_0) > 1$ in un certo istante t_0 , allora soddisfa $u(t) > 1$ per ogni t in cui è definita. Analogamente, se una soluzione è minore di 1 in un certo istante, allora rimane minore di 1 in tutto il suo intervallo di esistenza. In conclusione, una soluzione o è sempre = 1, o è sempre > 1, o è sempre < 1.

Cerchiamo le soluzioni $u \neq 1$. Dividiamo per $u - 1$ e riscriviamo l'equazione come

$$\frac{u'(t)}{u(t) - 1} = t^2. \tag{4.1}$$

Dalla regola della derivata di funzione composta e dall'integrale indefinito (primitiva) di Analisi 1

$$\int \frac{1}{x - 1} dx = \log |x - 1| + c,$$

possiamo scrivere la frazione $u'/(u-1)$ come derivata rispetto a t di una funzione:

$$\frac{u'(t)}{u(t)-1} = \frac{d}{dt} \left(\log |u(t)-1| \right).$$

Scriviamo anche t^2 come derivata di una funzione:

$$t^2 = \frac{d}{dt} \left(\frac{t^3}{3} \right).$$

L'equazione (4.1) si riscrive dunque

$$\frac{d}{dt} \left(\log |u(t)-1| \right) = \frac{d}{dt} \left(\frac{t^3}{3} \right),$$

cioè

$$\frac{d}{dt} \left(\log |u(t)-1| - \frac{t^3}{3} \right) = 0,$$

cioè

$$\log |u(t)-1| - \frac{t^3}{3} = \text{costante},$$

cioè

$$\log |u(t)-1| = \frac{t^3}{3} + c,$$

con $c \in \mathbb{R}$ costante (cioè non dipendente da t). Abbiamo trovato una relazione che lega $u(t)$ e t , e non contiene la derivata $u'(t)$ (non è più un'equazione differenziale). Per ricavare u , facciamo l'esponenziale

$$|u(t)-1| = \exp \left(\frac{t^3}{3} + c \right) = e^c e^{t^3/3}.$$

Se $u > 1$ ricaviamo

$$u(t) = 1 + e^c e^{t^3/3},$$

mentre se $u < 1$ si ha

$$u(t) = 1 - e^c e^{t^3/3}.$$

In ogni caso, rinominando c il fattore costante $\pm e^c$, abbiamo trovato la formula

$$u(t) = 1 + ce^{t^3/3}, \quad c \in \mathbb{R}.$$

Notiamo che questa formula, per $c = 0$, ci dà l'equilibrio $u = 1$.

Possiamo verificare immediatamente la validità della formula trovata: se $u(t) = 1 + ce^{t^3/3}$, calcoliamo la sua derivata, che è

$$u'(t) = ce^{t^3/3} t^2,$$

calcoliamo il campo

$$t^2(u-1) = t^2 ce^{t^3/3},$$

e constatiamo che sono uguali.

Consiglio: in esercizi di questo tipo è bene verificare subito la validità della formula trovata, per eliminare eventuali errori di calcolo.

Il procedimento appena illustrato si applica ogni volta che il campo $f(t, y)$ si fattorizza come prodotto di una funzione della sola y per una funzione della sola t :

Definizione 4.7. (*Equazioni a variabili separabili*). Si chiamano *equazioni a variabili separabili* le equazioni della forma

$$u' = \varphi(t)g(u)$$

dove φ e g sono funzioni reali di variabile reale.

Definizione 4.8. (*Equilibrio*). Una soluzione costante di un'equazione differenziale si dice anche un *equilibrio* dell'equazione.

Ripercorriamo il procedimento visto sopra, adattandolo al caso generale. Consideriamo l'equazione

$$u' = \varphi(t)g(u).$$

Per prima cosa cerchiamo gli eventuali zeri di g : se $g(x_0) = 0$ in un certo valore $x_0 \in \mathbb{R}$, allora la funzione costante $u(t) = x_0$ è soluzione. Le soluzioni costanti si dicono anche *equilibri*.

Se g è localmente lipschitziana, allora, per l'unicità della soluzione del problema di Cauchy, ogni soluzione che è diversa dagli equilibri in un certo istante lo è per tutti i tempi in cui è definita. Se u è diversa dagli equilibri, allora $g(u) \neq 0$, e l'equazione si riscrive

$$\frac{u'(t)}{g(u(t))} = \varphi(t).$$

Calcoliamo una primitiva $G(y)$ della funzione $1/g(y)$, cioè risolviamo l'integrale indefinito di Analisi 1

$$\int \frac{1}{g(y)} dy = G(y) + c,$$

cosicché

$$\frac{u'(t)}{g(u(t))} = \frac{d}{dt} (G(u(t))).$$

Calcoliamo una primitiva $\Phi(t)$ della funzione $\varphi(t)$, cosicché $\varphi(t) = \Phi'(t)$. L'equazione differenziale si riscrive dunque

$$\frac{d}{dt} (G(u(t))) = \Phi'(t),$$

cioè

$$\frac{d}{dt} (G(u(t)) - \Phi(t)) = 0,$$

da cui

$$G(u(t)) = \Phi(t) + c$$

con $c \in \mathbb{R}$ costante. A questo punto dobbiamo invertire G per ricavare u : poiché $g(u) \neq 0$, e $G'(y) = 1/g(y) \neq 0$, la funzione G è localmente invertibile in un intorno del valore $u(t)$ (Analisi 2). Quindi

$$u(t) = G^{-1}(\Phi(t) + c), \quad c \in \mathbb{R}, \tag{4.2}$$

famiglia a un parametro di soluzioni.

In alcuni casi la famiglia (4.2) include gli eventuali equilibri dell'equazione, in altri casi no; in generale, l'insieme di tutte le soluzioni dell'equazione $u' = g(u)\varphi(t)$ è dato dall'unione delle soluzioni (4.2) con le soluzioni costanti $u(t) = x_0$, dove $g(x_0) = 0$.

Esercizio 4.9. Si consideri l'equazione differenziale

$$u' = \frac{u^3}{t-1}.$$

(a) Scrivere tutte le soluzioni. (b) Risolvere il problema di Cauchy con dato iniziale (i) $u(0) = 1$; (ii) $u(2) = -1/2$; (iii) $u(2) = 0$.

Equazioni riconducibili ad equazioni a variabili separabili. Alcune equazioni, che non sono a variabili separabili, ammettono un cambio di variabile che trasforma il problema in un'equazione a variabili separabili. Illustriamo il procedimento tramite un esempio.

Esercizio 4.10. Risolvere l'equazione

$$u' = t^2u + t^3 - 1.$$

Svolgimento. L'equazione non è a variabili separabili, perché il campo $f(t, y) = t^2y + t^3 - 1$ non è della forma $\varphi(t)g(y)$. Tuttavia, se portiamo 1 a sinistra e mettiamo t^2 in evidenza, l'equazione diventa

$$u' + 1 = t^2(u + t).$$

Notiamo che $u' + 1$ è la derivata di $u + t$. Quindi usiamo $u + t$ come nuova incognita: facciamo cioè il *cambio di variabile*

$$v(t) := u(t) + t.$$

Calcoliamo $v' = u' + 1$, e riscriviamo l'equazione in termini della nuova incognita v :

$$v' = t^2v.$$

L'equazione per v è a variabili separabili, e dunque la possiamo risolvere esplicitamente. Procedendo come prima, troviamo le soluzioni

$$v(t) = ce^{t^3/3}, \quad c \in \mathbb{R}.$$

Infine ritorniamo all'incognita di partenza u : essendo $u = v - t$, troviamo le soluzioni

$$u(t) = ce^{t^3/3} - t, \quad c \in \mathbb{R}.$$

Il principio è abbastanza simile a quello degli integrali per sostituzione imparati in Analisi 1: un integrale che non sappiamo fare viene trasformato, tramite un cambio di variabile, in un altro che sappiamo risolvere, e come ultimo passaggio riscriviamo il risultato in termini della variabile di partenza.

Formula risolutiva dell'equazione $u' + a(t)u = 0$. Consideriamo l'equazione *lineare omogenea del primo ordine*

$$u' + a(t)u = 0.$$

È a variabili separabili: $u' = -a(t)u$. La soluzione costante $u = 0$ è equilibrio, mentre per $u \neq 0$ l'equazione è

$$\frac{u'(t)}{u(t)} = -a(t).$$

Si ha

$$\frac{u'(t)}{u(t)} = \frac{d}{dt} \left(\log |u(t)| \right).$$

Fissiamo una primitiva A di a , cioè $A'(t) = a(t)$. L'equazione si riscrive

$$\frac{d}{dt} \left(\log |u(t)| \right) = -A'(t),$$

cioè

$$\frac{d}{dt} \left(\log |u(t)| + A(t) \right) = 0,$$

da cui $\log |u(t)| + A(t) = \text{costante}$, cioè

$$\log |u(t)| = -A(t) + c.$$

Quindi

$$|u(t)| = e^{-A(t)+c} = e^c e^{-A(t)},$$

cioè

$$u(t) = \pm e^c e^{-A(t)},$$

e, rinominando la costante $\pm e^c$, troviamo la famiglia a un parametro di soluzioni

$$u(t) = ce^{-A(t)}, \quad c \in \mathbb{R}$$

(per $c = 0$ ritroviamo l'equilibrio $u = 0$). Data l'importanza di questa semplice formula, mettiamola bene in evidenza.

Le soluzioni dell'equazione lineare omogenea del primo ordine

$$u' + a(t)u = 0$$

sono le funzioni

$$u(t) = ce^{-A(t)}, \quad c \in \mathbb{R},$$

dove $A(t)$ è una primitiva di $a(t)$.

Fissato un istante t_0 , le soluzioni si possono anche scrivere come

$$u(t) = c \exp \left(- \int_{t_0}^t a(s) ds \right), \quad c \in \mathbb{R}. \quad (4.3)$$

Formula risolutiva dell'equazione $u' + a(t)u = b(t)$. Ora consideriamo l'equazione *lineare non omogenea del primo ordine*

$$u' + a(t)u = b(t). \quad (4.4)$$

Non è un'equazione a variabili separabili, ma può essere trasformata in un'equazione a variabili separabili tramite un cambio di variabile. Vedremo in seguito che il procedimento descritto qua sotto è un caso particolare del *metodo di variazione delle costanti*.

Consideriamo un cambio di variabile della forma

$$u(t) = h(t)v(t),$$

con $v(t)$ nuova incognita, e $h(t)$ funzione da scegliere nel modo più conveniente. Poiché $u' = h'v + hv'$, l'equazione (4.4) diventa

$$hv' + (h' + ah)v = b. \quad (4.5)$$

Scegliamo h in modo che il coefficiente $h' + ah$ davanti a v sia nullo: fissiamo dunque

$$h(t) := e^{-A(t)}$$

con A primitiva di a , cioè $A'(t) = a(t)$. L'equazione (4.5) diventa

$$e^{-A(t)}v'(t) = b(t),$$

cioè

$$v'(t) = e^{A(t)}b(t). \quad (4.6)$$

Fissiamo una primitiva $F(t)$ di $e^{A(t)}b(t)$, cioè $F'(t) = e^{A(t)}b(t)$. Le funzioni v che soddisfano (4.6) sono quelle della forma $v(t) = F(t) + c$, con $c \in \mathbb{R}$ costante. Ricordando che $u = hv$, troviamo

$$u(t) = e^{-A(t)}(F(t) + c).$$

Come sopra, data l'importanza di questa formula, riscriviamola bene in evidenza.

Le soluzioni dell'equazione lineare non omogenea del primo ordine

$$u' + a(t)u = b(t)$$

sono le funzioni

$$u(t) = e^{-A(t)}(c + F(t)), \quad c \in \mathbb{R},$$

dove $A(t)$ è una primitiva di $a(t)$ e $F(t)$ è una primitiva di $e^{A(t)}b(t)$.

Fissato un istante t_0 , le soluzioni si possono anche scrivere come

$$u(t) = \exp\left(-\int_{t_0}^t a(s) ds\right) \left[c + \int_{t_0}^t \exp\left(\int_{t_0}^s a(\sigma) d\sigma\right) b(s) ds \right], \quad c \in \mathbb{R}. \quad (4.7)$$

5 Lezione 5

Esercizio 5.1. Risolvere

$$u' + \frac{u}{t} = e^{t^2} \quad (t > 0).$$

Esercizio 5.2. Risolvere

$$\begin{cases} u' + tu = t \\ u(1) = 2. \end{cases}$$

Esercizio 5.3. Risolvere

$$\begin{cases} u' + 2 = te^u \\ u(0) = 0 \end{cases}$$

(trasformare l'equazione in un'equazione a variabili separabili usando un opportuno cambio di variabile).

Vediamo altre situazioni in cui si calcolano esplicitamente le soluzioni per sostituzione, cioè tramite cambio di variabile.

Equazioni del tipo $u' = g(u/t)$. Per equazioni della forma

$$u' = g\left(\frac{u}{t}\right),$$

definite per t in un intervallo $I \subseteq (0, \infty)$ oppure $I \subseteq (-\infty, 0)$, si può fare il cambio di variabile

$$u(t) = tv(t).$$

Si ha $u' = v + tv'$. In termini della nuova incognita v , l'equazione diventa

$$v' = \frac{1}{t}(g(v) - v),$$

che è a variabili separabili. Si calcola v , quindi si ottiene u .

Esercizio 5.4. Calcolare le soluzioni dell'equazione

$$u' = \frac{t^2u + u^3}{t^3} \quad (t > 0).$$

Equazioni di Bernoulli. Consideriamo l'equazione

$$u' + a(t)u + b(t)u^p = 0,$$

con $p \in \mathbb{R}$, $p \neq 0, 1$ (se $p = 0$ o $p = 1$ l'equazione è lineare, caso già trattato). Cerchiamo le soluzioni $u \neq 0$: dividiamo per u^p , e l'equazione diventa

$$\frac{u'}{u^p} + a(t)\frac{1}{u^{p-1}} + b(t) = 0.$$

Consideriamo il cambio di variabile

$$v(t) = \frac{1}{u^{p-1}(t)} = u^{1-p}(t).$$

Calcoliamo

$$v' = (1-p)u^{-p}u' = (1-p)\frac{u'}{u^p}.$$

L'equazione diventa

$$\frac{v'}{1-p} + a(t)v + b(t) = 0,$$

cioè

$$v' + (1-p)a(t)v + (1-p)b(t) = 0,$$

che è lineare non omogenea del primo ordine. Calcoliamo v , e quindi $u = v^{1/(1-p)}$.

Esercizio 5.5. Risolvere

$$\begin{cases} 2u' - t^2u + (t^3 + t^2 + 1)u^3 = 0 \\ u(0) = \frac{1}{2}. \end{cases}$$

Equazioni esatte (cioè equazioni con integrale primo). Consideriamo un'equazione

$$u' = f(t, u)$$

in cui il campo vettoriale $f(t, y)$ è dato dal quoziente

$$f(t, y) = -\frac{A(t, y)}{B(t, y)}$$

e il vettore (A, B) è il gradiente di una funzione scalare $W(t, y)$, cioè

$$A(t, y) = \partial_t W(t, y), \quad B(t, y) = \partial_y W(t, y). \quad (5.1)$$

In questo caso la funzione W calcolata lungo una qualunque soluzione dell'equazione differenziale è costante. Infatti, supponiamo che $u(t)$ sia soluzione di $u' = f(t, u)$. Allora

$$\begin{aligned} \frac{d}{dt} (W(t, u(t))) &= (\partial_t W)(t, u(t)) + (\partial_y W)(t, u(t))u'(t) \\ &= (\partial_t W)(t, u(t)) + (\partial_y W)(t, u(t))f(t, u(t)) \\ &= A(t, u(t)) + B(t, u(t))\frac{-A(t, u(t))}{B(t, u(t))} \\ &= 0. \end{aligned}$$

Quindi $W(t, u(t))$ è costante in t . Se u è la soluzione del problema di Cauchy con dato iniziale $u(t_0) = y_0$, allora

$$W(t, u(t)) = W(t_0, u(t_0)) = W(t_0, y_0)$$

per ogni t in cui $u(t)$ è definita. In altre parole, la curva $(t, u(t))$ è tutta contenuta nell'*insieme di livello*

$$\{(t, y) : W(t, y) = W(t_0, y_0)\}.$$

Si dice che W è un *integrale primo* dell'equazione differenziale. L'equazione si dice *esatta*.

Definizione 5.6. (*Integrale primo*). Una funzione che è costante lungo tutte le soluzioni di un'equazione differenziale viene detta un *integrale primo* dell'equazione.

Notiamo (Analisi 2) che (5.1) equivale a dire che la forma differenziale

$$\omega(t, y) := A(t, y)dt + B(t, y)dy$$

è esatta, e che $W(t, y)$ è una sua primitiva. Ricordiamo anche che se la forma differenziale ω è chiusa in un aperto Ω , cioè

$$\partial_y A(t, y) = \partial_t B(t, y) \quad \forall (t, y) \in \Omega,$$

e Ω è semplicemente connesso, allora ω è esatta in Ω , cioè ammette una primitiva.

Vediamo un esempio di risoluzione di un'equazione differenziale esatta.

Esercizio 5.7. Risolvere il problema di Cauchy

$$u' = \frac{e^{2t} - tu^2}{t^2 u}, \quad u(1) = 1$$

con $t > 0$, $u > 0$.

Svolgimento. Il campo vettoriale

$$f(t, y) = \frac{e^{2t} - ty^2}{t^2 y}$$

è dato dal quoziente

$$f(t, y) = -\frac{A(t, y)}{B(t, y)},$$

con

$$A(t, y) = -e^{2t} + ty^2, \quad B(t, y) = t^2 y.$$

La forma differenziale $\omega(t, y) = A(t, y)dt + B(t, y)dy$ è chiusa se

$$\partial_y A(t, y) = \partial_t B(t, y),$$

cioè se $\partial_y(-e^{2t} + ty^2) = \partial_t(t^2 y)$, che è vero. Inoltre il dominio

$$\Omega = \{(t, y) \in \mathbb{R}^2 : t > 0, y > 0\}$$

è semplicemente connesso, quindi ω è esatta in Ω . Cerchiamo una sua primitiva, cioè cerchiamo una funzione scalare $W(t, y)$ tale che

$$\partial_t W(t, y) = A(t, y) = -e^{2t} + ty^2, \quad \partial_y W(t, y) = B(t, y) = t^2 y.$$

La seconda di queste equazioni vale se $W(t, y) = \frac{1}{2}t^2 y^2 + h(t)$ per una funzione $h(t)$ indipendente da y ; la prima equazione diventa dunque

$$ty^2 + h'(t) = -e^{2t} + ty^2,$$

da cui si trova $h(t) = -\frac{1}{2}e^{-2t}$, e

$$W(t, y) = \frac{1}{2}(t^2 y^2 - e^{2t}).$$

Il grafico $\{(t, u(t))\}$ della soluzione del problema di Cauchy con dato iniziale $u(1) = 1$ si trova dunque nell'insieme di livello

$$\{(t, y) \in \Omega : W(t, y) = W(1, 1)\}$$

cioè

$$\left\{ (t, y) \in \mathbb{R}^2 : t > 0, y > 0, \frac{1}{2}(t^2 y^2 - e^{2t}) = \frac{1}{2}(1 - e^2) \right\}.$$

La soluzione $u(t)$ soddisfa allora

$$t^2 u^2(t) - e^{2t} = 1 - e^2,$$

da cui, usando le condizioni $u > 0, t > 0$, ricaviamo la formula esplicita della soluzione

$$u(t) = \frac{\sqrt{e^{2t} + 1 - e^2}}{t}.$$

Il suo intervallo di definizione è

$$I = \{t > 0 : e^{2t} + 1 - e^2 > 0\} = (\log \sqrt{e^2 - 1}, +\infty). \quad \square$$

Esercizio 5.8. Risolvere l'equazione

$$4t^3 - 4tu + 2(u - t^2)u' = 0$$

trovandone un integrale primo.

* * *

Riprendiamo lo studio generale del problema di Cauchy

$$u' = f(t, u), \quad u(t_0) = y_0.$$

Cosa succede quando manca l'ipotesi di lipschitzianità per il campo f ? L'esistenza della soluzione rimane intatta (Teorema 6.6 di esistenza di Peano), ma non è più garantita l'unicità della soluzione, come mostra il famoso esempio seguente.

Osservazione 5.9. (*Esistenza senza unicità: il pennello di Peano*). Il problema di Cauchy

$$\begin{cases} u' = \sqrt{|u|} \\ u(0) = 0 \end{cases} \quad (5.2)$$

ha la soluzione $u(t) = 0, t \in \mathbb{R}$. D'altra parte, anche la funzione

$$v(t) = \begin{cases} \frac{1}{4}t^2 & \text{se } t \geq 0, \\ 0 & \text{se } t < 0 \end{cases}$$

è soluzione di (5.2), e, più in generale, per ogni $\lambda > 0$ la funzione

$$v_\lambda(t) = \begin{cases} \frac{1}{4}(t - \lambda)^2 & \text{se } t \geq \lambda, \\ 0 & \text{se } t < \lambda \end{cases}$$

è soluzione di (5.2) (farne la verifica diretta). Il problema di Cauchy (5.2) ha infinite soluzioni.

Per dimostrare il Teorema di esistenza di Peano ci serve il Teorema di Ascoli-Arzelà (Teorema 5.11); cominciamo con la seguente proposizione.

Proposizione 5.10. ($C([a, b], \mathbb{R}^N)$ è spazio di Banach). *Lo spazio vettoriale*

$$C([a, b], \mathbb{R}^N) = \{f : [a, b] \rightarrow \mathbb{R}^N, f \text{ continua}\}$$

munito della sup-norma

$$\|f\|_\infty = \|f\|_{\infty, [a, b]} = \sup\{|f(x)| : x \in [a, b]\}$$

è uno spazio di Banach.

Dimostrazione. Il fatto che $C([a, b], \mathbb{R}^N)$ sia uno spazio vettoriale è immediato, e che $\|\cdot\|_\infty$ sia una norma è una facile verifica (farla per esercizio); resta da dimostrare la completezza.

Sia (f_n) una successione di Cauchy in $C([a, b], \mathbb{R}^N)$, cioè per ogni $\varepsilon > 0$ esiste $\bar{n} \in \mathbb{N}$ tale che $\|f_n - f_m\|_\infty \leq \varepsilon$ per ogni $n, m \geq \bar{n}$. Per ogni punto $x \in [a, b]$ si ha

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty \leq \varepsilon \quad \forall n, m \geq \bar{n}. \quad (5.3)$$

Dunque la successione $(f_n(x))_{n \in \mathbb{N}}$ è di Cauchy in \mathbb{R}^N , ed \mathbb{R}^N è completo; quindi $(f_n(x))_{n \in \mathbb{N}}$ converge in \mathbb{R}^N . Diciamo $f(x)$ il suo limite, $f(x) := \lim_{n \rightarrow \infty} f_n(x)$. Dimostriamo che $f_n \rightarrow f$ uniformemente in $[a, b]$.

Sia $\varepsilon > 0$. Come già osservato, esiste \bar{n} tale che valga (5.3). Sia $n \geq \bar{n}$ e $x \in [a, b]$. Allora $f_m(x) \in \overline{B(f_n(x), \varepsilon)}$ (palla chiusa di centro $f_n(x)$ e raggio ε) per ogni $m \geq \bar{n}$. È successione convergente, quindi anche il suo limite $f(x) = \lim_{m \rightarrow \infty} f_m(x)$ appartiene alla stessa palla, cioè $|f_n(x) - f(x)| \leq \varepsilon$. Poiché questo vale per ogni $x \in [a, b]$, ne segue che

$$\|f_n - f\|_\infty \leq \varepsilon \quad \forall n \geq \bar{n}.$$

Abbiamo dimostrato che per ogni $\varepsilon > 0$ esiste $\bar{n} \in \mathbb{N}$ tale che $\|f_n - f\|_\infty \leq \varepsilon$ per ogni $n \geq \bar{n}$, cioè che $\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0$, cioè che f_n converge a f uniformemente in $[a, b]$.

Infine f è continua in $[a, b]$ in quanto limite uniforme di una successione di funzioni continue in $[a, b]$ (Analisi 2). \square

Teorema 5.11. (Teorema di Ascoli-Arzelà). *Sia $f_n : [a, b] \rightarrow \mathbb{R}^N$, $n \in \mathbb{N}$, una successione di funzioni*

- *equilimitata, cioè esiste $M > 0$ tale che*

$$|f_n(x)| \leq M \quad \forall x \in [a, b], \forall n \in \mathbb{N};$$

- *equicontinua, cioè per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che*

$$|f_n(x) - f_n(y)| \leq \varepsilon \quad \forall x, y \in [a, b], |x - y| \leq \delta, \forall n \in \mathbb{N}.$$

Allora (f_n) ammette una sottosuccessione che converge uniformemente in $[a, b]$.

Osservazione 5.12. (*Notazione per infinite sottosuccessioni estratte ricorsivamente da sottosuccessioni e matrici di infinite righe e infinite colonne*). La dimostrazione del Teorema 5.11 è basata sulla costruzione ricorsiva di sottosuccessioni estratte dalla successione di partenza (f_n) , prendendo

una prima estratta,
 poi un'estratta della prima estratta,
 poi un'estratta dell'estratta della prima estratta,
 poi un'estratta dell'estratta dell'estratta della prima estratta,
 e così via.

Da questa lista infinita di successioni in cui ciascuna è un'estratta della precedente, con l'*argomento diagonale di Cantor* consideriamo la successione “diagonale” formata dal primo elemento della prima sottosuccessione, dal secondo elemento della seconda sottosuccessione, dal terzo elemento della terza sottosuccessione, eccetera.

Se scriviamo gli elementi della prima sottosuccessione uno dopo l'altro in riga, e poi quelli della seconda sottosuccessione anch'essi in riga, sotto la prima riga, e poi la terza estratta a formare una terza riga, eccetera, la nostra lista di infinite estratte forma una matrice di infinite righe e infinite colonne: la riga j -esima è la j -esima estratta. La successione diagonale è quella formata dagli elementi della diagonale della matrice.

Dovendo trattare infinite estratte, la notazione usuale per indicare le sottosuccessioni non ci basta: partendo da $(f_n)_{n \in \mathbb{N}}$, la prima estratta verrebbe indicata $(f_{n_k})_{k \in \mathbb{N}}$, l'estratta dell'estratta $(f_{n_{k_j}})_{j \in \mathbb{N}}$, la terza estratta $(f_{n_{k_{j_i}}})_{i \in \mathbb{N}}$, e chiaramente con questo modo di scrivere le sottosuccessioni non si può arrivare molto lontano. Allora, in analogia con l'idea di matrice infinito dimensionale, indichiamo $n(1, 1), n(1, 2), n(1, 3), \dots$ gli indici che formano la prima estratta (quelli che normalmente indicheremmo n_1, n_2, n_3, \dots); poi indichiamo $n(2, 1), n(2, 2), n(2, 3), \dots$ quelli che formano l'estratta della prima estratta (quelli che normalmente indicheremmo $n_{k_1}, n_{k_2}, n_{k_3}, \dots$); e in generale indichiamo $n(j, k)$, $k \in \mathbb{N}$, gli indici che danno la j -esima sottosuccessione della lista. È una notazione un po' insolita, ma ci permette di scrivere infinite sottosuccessioni in modo abbastanza semplice.

Dimostrazione del Teorema 5.11. Sia $\mathcal{D} = \{x_1, x_2, x_3, \dots\}$ un sottoinsieme denso e numerabile dell'intervallo $[a, b]$, per esempio $\mathcal{D} = [a, b] \cap \mathbb{Q}$.

Dall'ipotesi di equilimitatezza, la successione $(f_n(x_1))$ (la successione ottenuta valutando tutte le f_n nel punto x_1) è limitata, $|f_n(x_1)| \leq M$ per ogni $n \in \mathbb{N}$. Quindi (Bolzano-Weierstrass) esiste una sua sottosuccessione convergente, cioè esiste una successione strettamente crescente di naturali

$$S_1 = (n(1, 1), n(1, 2), n(1, 3), n(1, 4), \dots) \subseteq \mathbb{N}$$

tale che la successione

$$(f_{n(1,k)}(x_1))_{k \in \mathbb{N}} = (f_{n(1,1)}(x_1), f_{n(1,2)}(x_1), f_{n(1,3)}(x_1), f_{n(1,4)}(x_1), \dots),$$

che indichiamo anche $(f_n(x_1))_{n \in S_1}$, converge. Chiamiamo

$$y_1 := \lim_{k \rightarrow \infty} f_{n(1,k)}(x_1),$$

e osserviamo che $|y_1| \leq M$.

Ora consideriamo la successione di funzioni

$$(f_n)_{n \in S_1} = (f_{n(1,k)})_{k \in \mathbb{N}} = (f_{n(1,1)}, f_{n(1,2)}, f_{n(1,3)}, \dots),$$

e valutiamola nel punto x_2 . Anche $(f_n(x_2))_{n \in S_1}$ è una successione limitata, quindi ammette una sottosuccessione convergente, cioè esiste una successione strettamente crescente di naturali

$$S_2 = (n(2, 1), n(2, 2), n(2, 3), n(2, 4), \dots) \subseteq S_1$$

tale che la successione

$$(f_{n(2,k)}(x_2))_{k \in \mathbb{N}} = (f_{n(2,1)}(x_2), f_{n(2,2)}(x_2), f_{n(2,3)}(x_2), f_{n(2,4)}(x_2), \dots)$$

converge. Chiamiamo

$$y_2 := \lim_{k \rightarrow \infty} f_{n(2,k)}(x_2),$$

e osserviamo che $|y_2| \leq M$. Osserviamo inoltre che

$$\lim_{k \rightarrow \infty} f_{n(2,k)}(x_1) = y_1,$$

perché $(f_n(x_1))_{n \in S_1}$ è successione che converge a y_1 , e $(f_n(x_1))_{n \in S_2}$ è una sua sottosuccessione. Ripetiamo, con altre parole: la successione S_2 è un'estratta della successione S_1 , ed entrambe sono successioni crescenti di numeri naturali.

Ora consideriamo la successione $(f_n)_{n \in S_2}$ che abbiamo appena ottenuto, e la valutiamo nel punto x_3 : anche $(f_n(x_3))_{n \in S_2}$ è successione limitata, quindi ammette una sottosuccessione convergente, cioè esiste una successione strettamente crescente di naturali

$$S_3 = (n(3, 1), n(3, 2), n(3, 3), n(3, 4), \dots) \subseteq S_2$$

tale che la successione

$$(f_{n(3,k)}(x_3))_{k \in \mathbb{N}} = (f_{n(3,1)}(x_3), f_{n(3,2)}(x_3), f_{n(3,3)}(x_3), f_{n(3,4)}(x_3), \dots)$$

converge. Chiamiamo

$$y_3 := \lim_{k \rightarrow \infty} f_{n(3,k)}(x_3),$$

e osserviamo che $|y_3| \leq M$. Osserviamo inoltre che

$$\lim_{k \rightarrow \infty} f_{n(3,k)}(x_1) = y_1, \quad \lim_{k \rightarrow \infty} f_{n(3,k)}(x_2) = y_2,$$

perché $(f_n(x_1))_{n \in S_2}$ converge a y_1 e $(f_n(x_1))_{n \in S_3}$ è una sua sottosuccessione, e similmente $(f_n(x_2))_{n \in S_2}$ converge a y_2 e $(f_n(x_2))_{n \in S_3}$ è una sua sottosuccessione. Ripetiamo ancora, S_3 è una successione strettamente crescente di numeri naturali estratti dalla successione S_2 .

Procediamo in questo modo, e otteniamo infinite successioni di numeri naturali

$$S_1 \supseteq S_2 \supseteq S_3 \supseteq S_4 \supseteq \dots,$$

dove, per ogni p , S_{p+1} è una successione strettamente crescente di naturali estratti dalla successione S_p ; gli elementi della successione S_p sono indicati

$$S_p = (n(p, 1), n(p, 2), n(p, 3), n(p, 4), \dots) = (n(p, k))_{k \in \mathbb{N}}, \quad \forall p \geq 1,$$

e formano una successione strettamente crescente, cioè $n(p, 1) < n(p, 2) < \dots$; inoltre la successione $(f_n)_{n \in S_p}$ valutata nei punti x_1, x_2, \dots, x_p converge a y_1, y_2, \dots, y_p rispettivamente, cioè

$$\lim_{k \rightarrow \infty} f_{n(p,k)}(x_1) = y_1, \quad \lim_{k \rightarrow \infty} f_{n(p,k)}(x_2) = y_2, \quad \dots \quad \lim_{k \rightarrow \infty} f_{n(p,k)}(x_p) = y_p.$$

Ora consideriamo la successione “diagonale” di naturali che otteniamo prendendo il primo elemento di S_1 , il secondo elemento di S_2 , il terzo elemento di S_3 , eccetera: definiamo

$$D_0 := (n(1, 1), n(2, 2), n(3, 3), n(4, 4), \dots) = (n(k, k))_{k \in \mathbb{N}}.$$

Osserviamo che

$$(n(1, 1), n(2, 2), n(3, 3), n(4, 4), \dots) \subseteq S_1$$

perché $n(1, 1) \in S_1$, $n(2, 2) \in S_2 \subseteq S_1$, $n(3, 3) \in S_3 \subseteq S_2 \subseteq S_1$, eccetera; inoltre osserviamo che

$$(n(2, 2), n(3, 3), n(4, 4), \dots) \subseteq S_2$$

perché $n(2, 2) \in S_2$, $n(3, 3) \in S_3 \subseteq S_2$, $n(4, 4) \in S_4 \subseteq S_3 \subseteq S_2$, eccetera; similmente,

$$(n(3, 3), n(4, 4), \dots) \subseteq S_3$$

e, più in generale, per ogni $p \geq 1$ si ha

$$(n(p, p), n(p+1, p+1), \dots) = (n(k, k))_{k \geq p} \subseteq S_p.$$

In altre parole, D_0 è definitivamente contenuta in S_p , per ogni $p \geq 1$.

Di conseguenza, per ogni $p \geq 1$, la successione

$$(f_{n(k,k)}(x_p))_{k \geq p} = (f_{n(p,p)}(x_p), f_{n(p+1,p+1)}(x_p), f_{n(p+2,p+2)}(x_p), \dots) \quad (5.4)$$

è una sottosuccessione di $(f_n(x_p))_{n \in S_p}$. Come già osservato, $(f_n(x_p))_{n \in S_p}$ converge a y_p , quindi anche la successione (5.4), essendo sua sottosuccessione, converge a y_p . Perciò, per ogni $p \geq 1$,

$$\lim_{k \rightarrow \infty} f_{n(k,k)}(x_p) = y_p.$$

Abbiamo dimostrato che la successione $(f_n(x_p))_{n \in D_0}$ converge a y_p , per ogni $p \geq 1$.

In altre parole, se indichiamo

$$g_k := f_{n(k,k)},$$

abbiamo dimostrato che $\lim_{(k \rightarrow \infty)} g_k(x_p) = y_p$ per ogni p , cioè (g_k) converge puntualmente in ogni punto dell'insieme $\mathcal{D} = \{x_1, x_2, \dots\}$.

Resta da dimostrare che (g_k) converge uniformemente in $[a, b]$; lo proveremo nella prossima lezione.

6 Lezione 6

Conclusione della dimostrazione del Teorema di Ascoli-Arzelà (Teorema 5.11). Nella scorsa lezione abbiamo dimostrato che la sottosuccessione “diagonale” (g_k) converge puntualmente in ogni punto dell’insieme $\mathcal{D} = \{x_1, x_2, \dots\}$; ora rimane da dimostrare che (g_k) converge uniformemente in $[a, b]$.

Sia $\varepsilon > 0$. Dall’ipotesi di equicontinuit , esiste $\delta > 0$ tale che

$$|f_n(x) - f_n(y)| < \frac{\varepsilon}{3} \quad \forall x, y \in [a, b], \quad |x - y| < \delta, \quad \forall n \in \mathbb{N}.$$

Questa disuguaglianza vale, in particolare, per gli indici $n \in D_0$:

$$|f_{n(k,k)}(x) - f_{n(k,k)}(y)| < \frac{\varepsilon}{3} \quad \forall x, y \in [a, b], \quad |x - y| < \delta, \quad \forall k \in \mathbb{N},$$

cio 

$$|g_k(x) - g_k(y)| < \frac{\varepsilon}{3} \quad \forall x, y \in [a, b], \quad |x - y| < \delta, \quad \forall k \in \mathbb{N}. \quad (6.1)$$

Poich  \mathcal{D}   denso in $[a, b]$, per ogni $x \in [a, b]$ esiste un punto $x_j \in \mathcal{D}$ tale che $|x - x_j| < \delta$, quindi $x \in (x_j - \delta, x_j + \delta)$. Di conseguenza,

$$[a, b] \subseteq \bigcup_{j=1}^{\infty} (x_j - \delta, x_j + \delta).$$

Per compattezza di $[a, b]$, esiste un sottoricoprimento finito, quindi esiste $q \in \mathbb{N}$ tale che

$$[a, b] \subseteq \bigcup_{j=1}^q (x_j - \delta, x_j + \delta).$$

Poich  $(g_k(x_1))$   convergente, esiste $\bar{k}_1 \in \mathbb{N}$ tale che

$$|g_k(x_1) - g_j(x_1)| < \frac{\varepsilon}{3} \quad \forall k, j \geq \bar{k}_1.$$

Similmente, per ogni $\ell = 2, \dots, q$ esiste \bar{k}_ℓ tale che

$$|g_k(x_\ell) - g_j(x_\ell)| < \frac{\varepsilon}{3} \quad \forall k, j \geq \bar{k}_\ell.$$

Sia $\bar{k} := \max\{\bar{k}_1, \dots, \bar{k}_q\}$, cosicch 

$$|g_k(x_\ell) - g_j(x_\ell)| < \frac{\varepsilon}{3} \quad \forall k, j \geq \bar{k}, \quad \forall \ell = 1, \dots, q. \quad (6.2)$$

Ora sia $x \in [a, b]$. Allora $x \in (x_\ell - \delta, x_\ell + \delta)$ per un certo $\ell \in \{1, \dots, q\}$. Quindi, usando (6.1) e (6.2),

$$|g_k(x) - g_j(x)| \leq |g_k(x) - g_k(x_\ell)| + |g_k(x_\ell) - g_j(x_\ell)| + |g_j(x_\ell) - g_j(x)| < \varepsilon \quad \forall k, j \geq \bar{k}.$$

Abbiamo dimostrato che per ogni $\varepsilon > 0$ esiste $\bar{k} \in \mathbb{N}$ tale che $|g_k(x) - g_j(x)| < \varepsilon$ per ogni $k, j \geq \bar{k}$, per ogni $x \in [a, b]$. Dunque la successione (g_k)   uniformemente di Cauchy in $[a, b]$, e quindi (Proposizione 5.10) converge uniformemente in $[a, b]$. \square

Osservazione 6.1. (*Esempio di costruzione delle sottosuccessioni nella dimostrazione di Ascoli-Arzelà*). Riportiamo un esempio numerico per aiutare a seguire la costruzione delle estratte $S_1 \supseteq S_2 \supseteq S_3 \supseteq \dots$ nella dimostrazione del Teorema di Ascoli-Arzelà: partiamo dalla successione $(f_n)_{n \in \mathbb{N}} = (f_1, f_2, f_3, \dots)$, da questa estraiamo la prima sottosuccessione, diciamo, per esempio,

$$(f_n)_{n \in S_1} = (f_{n(1,k)})_{k \in \mathbb{N}} = (f_3, f_7, f_{11}, f_{12}, f_{15}, f_{20}, f_{21}, f_{22}, f_{28}, f_{30}, f_{32}, f_{37}, f_{42}, f_{43}, \\ f_{50}, f_{54}, f_{70}, f_{80}, f_{81}, f_{90}, f_{103}, \dots)$$

cosicché S_1 è la successione di naturali

$$S_1 = (3, 7, 11, 12, 15, 20, 21, 22, 28, 30, 32, 37, 42, 43, 50, 54, 70, \\ 80, 81, 90, 103, \dots),$$

$$\begin{aligned} n(1, 1) &= 3, & n(1, 2) &= 7, & n(1, 3) &= 11, & n(1, 4) &= 12, & n(1, 5) &= 15, \\ n(1, 6) &= 20, & n(1, 7) &= 21, & n(1, 8) &= 22, & n(1, 9) &= 28, & n(1, 10) &= 30, \\ n(1, 11) &= 32, & n(1, 12) &= 37, & n(1, 13) &= 42, & n(1, 14) &= 43, & n(1, 15) &= 50, \\ n(1, 16) &= 54, & n(1, 17) &= 70, & n(1, 18) &= 80, & n(1, 19) &= 81, & n(1, 20) &= 90, \\ n(1, 21) &= 103, & & \dots & & & & & & \end{aligned}$$

Da S_1 estraiamo la seconda successione $S_2 \subseteq S_1$, diciamo

$$S_2 = (7, 12, 20, 28, 30, 37, 43, 70, 80, 90, 103, \dots),$$

$$\begin{aligned} n(2, 1) &= 7, & n(2, 2) &= 12, & n(2, 3) &= 20, & n(2, 4) &= 28, & n(2, 5) &= 30, \\ n(2, 6) &= 37, & n(2, 7) &= 43, & n(2, 8) &= 70, & n(2, 9) &= 80, & n(2, 10) &= 90, \\ n(2, 11) &= 103, & & \dots & & & & & & \end{aligned}$$

$$(f_n)_{n \in S_2} = (f_{n(2,k)})_{k \in \mathbb{N}} = (f_7, f_{12}, f_{20}, f_{28}, f_{30}, f_{37}, f_{43}, f_{70}, f_{80}, f_{90}, f_{103}, \dots).$$

Poi da S_2 estraiamo la terza successione $S_3 \subseteq S_2$, diciamo

$$S_3 = (20, 30, 43, 80, 103, \dots),$$

$$n(3, 1) = 20, \quad n(3, 2) = 30, \quad n(3, 3) = 43, \quad n(3, 4) = 80, \quad n(3, 5) = 103, \dots,$$

$$(f_n)_{n \in S_3} = (f_{n(3,k)})_{k \in \mathbb{N}} = (f_{20}, f_{30}, f_{43}, f_{80}, f_{103}, \dots).$$

Da S_3 estraiamo la quarta successione $S_4 \subseteq S_3$, diciamo

$$S_4 = (20, 30, 80, 103, \dots),$$

$$n(4, 1) = 20, \quad n(4, 2) = 30, \quad n(4, 3) = 80, \quad n(4, 4) = 103, \dots,$$

$$(f_n)_{n \in S_4} = (f_{n(4,k)})_{k \in \mathbb{N}} = (f_{20}, f_{30}, f_{80}, f_{103}, \dots),$$

eccetera. La successione diagonale D_0 è dunque

$$D_0 = (n(1, 1), n(2, 2), n(3, 3), n(4, 4), \dots) = (3, 12, 43, 103, \dots)$$

e

$$(f_n)_{n \in D_0} = (f_{n(k,k)})_{k \in \mathbb{N}} = (f_3, f_{12}, f_{43}, f_{103}, \dots),$$

$$\begin{aligned} g_1 &= f_{n(1,1)} = f_3, & g_2 &= f_{n(2,2)} = f_{12}, & g_3 &= f_{n(3,3)} = f_{43}, \\ g_4 &= f_{n(4,4)} = f_{103}, & \dots & \end{aligned}$$

Osserviamo che

$$(g_1, g_2, g_3, g_4, \dots) = (f_3, f_{12}, f_{43}, f_{103}, \dots) \subseteq (f_n)_{n \in S_1};$$

g_1 non è elemento della successione $(f_n)_{n \in S_2}$, ma da g_2 in poi si ha

$$(g_2, g_3, g_4, \dots) = (f_{12}, f_{43}, f_{103}, \dots) \subseteq (f_n)_{n \in S_2};$$

g_1, g_2 non sono elementi della successione $(f_n)_{n \in S_3}$, ma da g_3 in poi si ha

$$(g_3, g_4, \dots) = (f_{43}, f_{103}, \dots) \subseteq (f_n)_{n \in S_3};$$

g_1, g_2, g_3 non sono elementi della successione $(f_n)_{n \in S_4}$, ma da g_4 in poi si ha

$$(g_4, \dots) = (f_{103}, \dots) \subseteq (f_n)_{n \in S_4};$$

in generale, $(g_k)_{k \geq p} \subseteq (f_n)_{n \in S_p}$. Fine dell'esempio.

Nelle prossime proposizioni osserviamo che la lipschitzianità locale implica la lipschitzianità sui compatti; questo è il “caso lipschitziano” di un comportamento generale dei compatti: quando una certa proprietà vale localmente (cioè in un intorno di ciascun punto), sui compatti vale globalmente. Ne vediamo tre versioni leggermente diverse: la prima è la più semplice, la seconda è quella più standard, la terza è quella più adatta allo studio dei problemi di Cauchy.

Proposizione 6.2. (Lipschitzianità locale implica lipschitzianità sui compatti). *Sia $K \subset \mathbb{R}^d$ un compatto, sia $g : K \rightarrow \mathbb{R}^N$ una funzione limitata tale che per ogni $p \in K$ esistono $r > 0$, $L \geq 0$ tali che $|g(x) - g(y)| \leq L|x - y|$ per ogni $x, y \in K \cap B(p, r)$. Allora g è lipschitziana in K .*

Dimostrazione. Supponiamo per assurdo che non esista alcuna costante $L \geq 0$ tale che $|g(x) - g(y)| \leq L|x - y|$ per ogni $x, y \in K$. Questo significa che per ogni $L \geq 0$ esistono $x, y \in K$ che violano la disuguaglianza. In particolare, per ogni $n \in \mathbb{N}$ esistono $x_n, y_n \in K$ tali che

$$|g(x_n) - g(y_n)| > n|x_n - y_n|. \quad (6.3)$$

Essendo K compatto, la successione (x_n) ammette una sottosuccessione x_{n_k} convergente; diciamo $p = \lim_{(k \rightarrow \infty)} x_{n_k}$. Per ipotesi, g è limitata in K , diciamo $|g(x)| \leq M$ per ogni $x \in K$. Quindi, da (6.3),

$$|x_{n_k} - y_{n_k}| < \frac{|g(x_{n_k}) - g(y_{n_k})|}{n_k} \leq \frac{2M}{n_k} \rightarrow 0 \quad (k \rightarrow \infty),$$

e dunque anche y_{n_k} converge a p . Il punto p appartiene a K . Per ipotesi, esistono $r > 0$, $L \geq 0$ tali che

$$|g(x) - g(y)| \leq L|x - y| \quad \forall x, y \in K \cap B(p, r).$$

Poiché x_{n_k}, y_{n_k} convergono a p , esiste k_0 tale che $x_{n_k}, y_{n_k} \in B(p, r)$ per ogni $k \geq k_0$. Dunque per $k \geq k_0$ si ha

$$n_k|x_{n_k} - y_{n_k}| < |g(x_{n_k}) - g(y_{n_k})| \leq L|x_{n_k} - y_{n_k}|. \quad (6.4)$$

La prima disuguaglianza implica che $|g(x_{n_k}) - g(y_{n_k})| > 0$, quindi, dalla seconda disuguaglianza, anche $|x_{n_k} - y_{n_k}| > 0$. Di conseguenza $n_k < L$ per ogni $k \geq k_0$, ma $n_k \rightarrow \infty$, assurdo. \square

Proposizione 6.3. (Lipschitzianità locale implica lipschitzianità sui compatti/2). Sia $\Omega \subseteq \mathbb{R}^d$ un aperto, e sia $g : \Omega \rightarrow \mathbb{R}^N$ una funzione localmente lipschitziana, cioè per ogni $p \in \Omega$ esistono $r > 0$, $L \geq 0$ tali che $B(p, r) \subseteq \Omega$ e $|g(x) - g(y)| \leq L|x - y|$ per ogni $x, y \in B(p, r)$. Sia $K \subset \Omega$ compatto. Allora g è lipschitziana in K .

Proposizione 6.4. (Locale lipschitzianità in y uniforme in t implica lipschitzianità in y uniforme in t sui compatti). Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto, sia $f : \Omega \rightarrow \mathbb{R}^N$ una funzione localmente lipschitziana in y , uniformemente in t (come nella Definizione 3.2). Sia $K \subset \Omega$ compatto. Allora la funzione $f(t, y)$ è lipschitziana in y , uniformemente in t , nell'insieme K .

Esercizio 6.5. (Locale lipschitzianità). Dimostrare le Proposizioni 6.3 e 6.4 (per esempio, adattare la dimostrazione della Proposizione 6.2).

Vediamo ora il Teorema di Peano che dà l'esistenza (senza unicità) di una soluzione del problema di Cauchy quando il campo è una funzione continua (senza ipotesi di lipschitzianità).

Teorema 6.6. (Teorema di esistenza di Peano). Sia $t_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^N$,

$$I = [t_0 - a, t_0 + a], \quad J = \{y \in \mathbb{R}^N : |y - y_0| \leq b\} = \overline{B(y_0, b)},$$

con $a > 0$, $b > 0$. Sia $f : I \times J \rightarrow \mathbb{R}^N$ una funzione continua. Siano M, δ tali che

$$|f(t, y)| \leq M \quad \forall (t, y) \in I \times J; \quad 0 < \delta \leq a, \quad M\delta \leq b. \quad (6.5)$$

Allora il problema di Cauchy

$$\begin{cases} u' = f(t, u) \\ u(t_0) = y_0 \end{cases} \quad (6.6)$$

ha una soluzione definita nell'intervallo $[t_0 - \delta, t_0 + \delta]$.

Riportiamo due dimostrazioni del Teorema 6.6. La prima, qui di seguito suddivisa in lemmi indipendenti, si appoggia sul Teorema di esistenza e unicità locale e sulla possibilità di approssimare una funzione continua con funzioni lipschitziane; la seconda, soltanto accennata a lezione, e qui inserita in appendice (sezione 26), costruisce una successione risolvendo una versione approssimata dell'equazione differenziale, secondo il metodo di Eulero. In entrambe le dimostrazioni si fa uso del Teorema di Ascoli-Arzelà.

Lemma 6.7. (Se il campo f è approssimabile con funzioni lipschitziane allora esiste una soluzione). Assumiamo le ipotesi del Teorema 6.6. Supponiamo inoltre che esista una successione (f_n) di funzioni $f_n : I \times J \rightarrow \mathbb{R}^N$ tali che

- (i) per ogni n , f_n è continua in $I \times J$;
- (ii) per ogni n , $|f_n(t, y)| \leq M$ per ogni $(t, y) \in I \times J$;
- (iii) per ogni n , f_n è lipschitziana in y , uniformemente in t , in $I \times J$, cioè per ogni n esiste $L_n \geq 0$ tale che $|f_n(t, y) - f_n(t, z)| \leq L_n|y - z|$ per ogni $t \in I$, $y, z \in J$;
- (iv) f_n converge a f uniformemente in $I \times J$ per $n \rightarrow \infty$.

Allora il problema di Cauchy (6.6) ha una soluzione in $I_0 := [t_0 - \delta, t_0 + \delta]$.

Si noti che al punto (iii) non si richiede che la lipschitzianità sia uniforme in n : ogni funzione f_n ha la sua costante di Lipschitz L_n , senza assumere che l'insieme $\{L_n : n \in \mathbb{N}\}$ sia limitato.

Dimostrazione del Lemma 6.7. Per ogni n , il campo $f_n(t, y)$ è continuo, lipschitziano in y , uniformemente in t , e soddisfa $|f_n| \leq M$ in $I \times J$. Quindi, dal Teorema 2.1 di esistenza locale e unicità, nell'intervallo I_0 esiste (unica) la soluzione del problema di Cauchy con campo f_n e condizioni iniziali (t_0, y_0) ; indichiamo $u_n(t)$ tale soluzione, cosicché si ha

$$u_n'(t) = f_n(t, u_n(t)) \quad \forall t \in I_0, \quad u_n(t_0) = y_0.$$

Dalla formulazione integrale del problema di Cauchy, per ogni n si ha dunque

$$u_n(t) = y_0 + \int_{t_0}^t f_n(s, u_n(s)) ds \quad \forall t \in I_0. \quad (6.7)$$

Dalla (6.7), per ogni $t \in I_0$ e ogni $n \in \mathbb{N}$ si ha

$$|u_n(t)| \leq |y_0| + \left| \int_{t_0}^t f_n(s, u_n(s)) ds \right| \leq |y_0| + M|t - t_0| \leq |y_0| + M\delta$$

(e d'altra parte, per definizione di soluzione, $(t, u_n(t)) \in I \times J$ per ogni $t \in I_0$, da cui si ha direttamente $|u_n(t)| \leq |y_0| + b$), quindi la successione (u_n) è equilimitata; inoltre, usando ancora la (6.7),

$$|u_n(t) - u_n(\tau)| = \left| \int_{\tau}^t f_n(s, u_n(s)) ds \right| \leq M|t - \tau|,$$

cioè le funzioni u_n sono tutte lipschitziane con la stessa costante di Lipschitz M . Di conseguenza la successione (u_n) è equicontinua: dato $\varepsilon > 0$, prendiamo $\delta_1 > 0$ tale che $M\delta_1 < \varepsilon$, e abbiamo

$$|u_n(t) - u_n(\tau)| \leq M|t - \tau| \leq M\delta_1 < \varepsilon \quad \forall t, \tau \in I_0, |t - \tau| \leq \delta_1, \forall n \in \mathbb{N}.$$

Dal Teorema 5.11 di Ascoli-Arzelà, la successione (u_n) ha una sottosuccessione (u_{n_k}) che converge uniformemente in I_0 ; chiamiamo u il suo limite, $u = \lim_{(k \rightarrow \infty)} u_{n_k}$. La funzione u è continua in quanto limite uniforme di una successione di funzioni continue; inoltre $u(t) \in J$ per ogni $t \in I_0$ perché $u_{n_k}(t) \in J$ per ogni $t \in I_0$ e ogni $k \in \mathbb{N}$, J è un chiuso, e $u(t) = \lim_{(k \rightarrow \infty)} u_{n_k}(t)$.

Mostriamo che la funzione composta $f_{n_k}(t, u_{n_k}(t))$ converge a $f(t, u(t))$ uniformemente in I_0 per $k \rightarrow \infty$. Sia $\varepsilon > 0$. Per ogni $t \in I_0$, aggiungendo e togliendo il termine $f(t, u_{n_k}(t))$ e usando la disuguaglianza triangolare, si ha

$$\begin{aligned} |f_{n_k}(t, u_{n_k}(t)) - f(t, u(t))| &\leq |f_{n_k}(t, u_{n_k}(t)) - f(t, u_{n_k}(t))| + |f(t, u_{n_k}(t)) - f(t, u(t))| \\ &\leq \|f_{n_k} - f\|_{\infty, I \times J} + |f(t, u_{n_k}(t)) - f(t, u(t))|. \end{aligned}$$

Poiché, per ipotesi, f_n converge a f uniformemente in $I \times J$, e (f_{n_k}) è sottosuccessione di (f_n) , esiste $\bar{k}_1 \in \mathbb{N}$ tale che

$$\|f_{n_k} - f\|_{\infty, I \times J} < \varepsilon/2 \quad \forall k \geq \bar{k}_1.$$

Poiché f è continua e $I \times J$ è compatto, f è uniformemente continua in $I \times J$; quindi esiste $\delta_2 > 0$ tale che

$$|f(t, y) - f(s, z)| < \varepsilon/2 \quad \forall (t, y), (s, z) \in I \times J, |t - s| + |y - z| \leq \delta_2.$$

Poiché $\|u_{n_k} - u\|_{\infty, I_0} \rightarrow 0$, esiste \bar{k}_2 tale che $\|u_{n_k} - u\|_{\infty, I_0} < \delta_2$ per ogni $k \geq \bar{k}_2$.
Dunque

$$|f(t, u_{n_k}(t)) - f(t, u(t))| < \varepsilon/2 \quad \forall t \in I_0, \quad \forall k \geq \bar{k}_2.$$

Per ogni $k \geq \bar{k} := \max\{\bar{k}_1, \bar{k}_2\}$ si ha quindi $|f_{n_k}(t, u_{n_k}(t)) - f(t, u(t))| < \varepsilon$ per ogni $t \in I_0$.

Abbiamo provato che per ogni $\varepsilon > 0$ esiste $\bar{k} \in \mathbb{N}$ tale che $|f_{n_k}(t, u_{n_k}(t)) - f(t, u(t))| < \varepsilon$ per ogni $t \in I_0$ e ogni $k \geq \bar{k}$, cioè che $f_{n_k}(t, u_{n_k}(t))$ converge a $f(t, u(t))$ uniformemente in I_0 per $k \rightarrow \infty$.

L'identità (6.7) vale per ogni n , quindi, in particolare, per ogni n_k , cioè

$$u_{n_k}(t) = y_0 + \int_{t_0}^t f_{n_k}(s, u_{n_k}(s)) ds \quad \forall t \in I_0, \quad \forall k \in \mathbb{N}.$$

La convergenza uniforme permette il passaggio al limite sotto il segno di integrale; passando al limite per $k \rightarrow \infty$ otteniamo

$$u(t) = y_0 + \int_{t_0}^t f(s, u(s)) ds \quad \forall t \in I_0,$$

quindi u è soluzione del problema di Cauchy (6.6) in I_0 . □

Grazie al Lemma 6.7, la dimostrazione del Teorema 6.6 di esistenza di Peano diventa una questione di approssimazione del campo: resta da dimostrare che ogni funzione f continua in $I \times J$ ammette una successione (f_n) come nel Lemma 6.7.

Lemma 6.8. (Approssimazione di funzioni continue con funzioni lipschitziane). *Sia $d, N \geq 1$, sia $\Omega \subseteq \mathbb{R}^d$ un aperto, $g : \Omega \rightarrow \mathbb{R}^N$ una funzione continua, $M \geq 0$, e sia $|g(x)| \leq M$ per ogni $x \in \Omega$. Sia $K \subset \Omega$ un insieme compatto. Allora esiste una successione (g_n) di funzioni $g_n : K \rightarrow \mathbb{R}^N$, continue e lipschitziane in K , con $|g_n(x)| \leq M$ per ogni $x \in K$, tale che g_n converge a g uniformemente in K .*

Prima di dimostrare il Lemma 6.8, proviamo che dai Lemmi 6.8 e 6.7 si ottiene il Teorema di esistenza di Peano.

Dimostrazione del Teorema 6.6 usando i Lemmi 6.8, 6.7. Assumiamo le ipotesi del Teorema 6.6. Definiamo un'estensione continua \tilde{f} di f prolungando i valori che f assume al bordo di $I \times J$ anche al di fuori di $I \times J$. Definiamo dunque

$$\tilde{f} : \mathbb{R}^{1+N} \rightarrow \mathbb{R}^N, \quad \tilde{f}(t, y) := f(h(t), \psi(y))$$

dove $h : \mathbb{R} \rightarrow I$, $\psi : \mathbb{R}^N \rightarrow J$,

$$h(t) = \begin{cases} t & \text{per } |t - t_0| \leq a, \\ t_0 + a & \text{per } t > t_0 + a, \\ t_0 - a & \text{per } t < t_0 - a, \end{cases} \quad \psi(y) = \begin{cases} y & \text{per } |y - y_0| \leq b, \\ y_0 + \frac{y - y_0}{|y - y_0|} b & \text{per } |y - y_0| > b. \end{cases}$$

La funzione \tilde{f} è continua perché f, h, ψ sono continue (ψ è banalmente continua nella palla aperta $|y - y_0| < b$; poi per $y \neq y_0$ si ha $\psi(y) = y_0 + \min\{1, b/|y - y_0|\}(y - y_0)$, quindi ψ è composizione di funzione continue). Si ha $\tilde{f}(t, y) = f(t, y)$ per ogni $(t, y) \in I \times J$; inoltre $|\tilde{f}(t, y)| \leq M$ per ogni $(t, y) \in \mathbb{R}^{1+N}$ perché $|f| \leq M$ in $I \times J$.

Dal Lemma 6.8 (applicato con $d = 1 + N$, $\Omega = \mathbb{R}^{1+N}$, $K = I \times J$, $g = \tilde{f}$), esiste una successione (f_n) di funzioni continue e lipschitziane in $I \times J$, con $|f_n| \leq M$, che converge a f uniformemente in $I \times J$. Sono dunque verificate le ipotesi del Lemma 6.7, da cui si deduce la tesi del Teorema 6.6. □

Ci rimane dunque da provare il Lemma 6.8. La dimostrazione che segue fa uso dell'integrale in più variabili, ma soltanto su domini limitati che sono "rettangoli" (cioè prodotto cartesiano di intervalli) o loro unioni finite, e soltanto per funzioni continue. Sui rettangoli, le "formule di riduzione" sono particolarmente semplici: se $f(x) = f(x_1, \dots, x_d)$ è una funzione continua di $x \in \mathbb{R}^d$ ed E è il rettangolo

$$E = [a_1, b_1] \times \dots \times [a_d, b_d] = \{x \in \mathbb{R}^d : a_i \leq x_i \leq b_i, i = 1, \dots, d\},$$

l'integrale di $f(x)$ sul rettangolo E è

$$\int_E f(x) dx = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \dots \left(\int_{a_d}^{b_d} f(x_1, \dots, x_d) dx_d \right) \dots dx_2 \right) dx_1.$$

Quando il rettangolo è del tipo $[c_1 - r, c_1 + r] \times \dots \times [c_d - r, c_d + r]$, lo chiamiamo "cubo" di centro $c = (c_1, \dots, c_d)$ e lato $2r$.

Dimostrazione del Lemma 6.8. Sia

$$\psi : \mathbb{R} \rightarrow \mathbb{R}, \quad \psi(s) = \begin{cases} 1 - |s| & \text{se } |s| \leq 1, \\ 0 & \text{se } |s| > 1. \end{cases}$$

La funzione ψ è continua e soddisfa

$$0 \leq \psi(s) \leq 1 \quad \forall s \in \mathbb{R}; \quad \psi(s) = 0 \quad \forall s \notin [-1, 1]; \quad \int_{-1}^1 \psi(s) ds = 1$$

(l'integrale si calcola anche come area di un triangolo di base 2 e altezza 1). È inoltre lipschitziana,

$$|\psi(s) - \psi(t)| \leq |s - t| \quad \forall s, t \in \mathbb{R}.$$

Infatti per $s, t \in [-1, 1]$ si ha

$$|\psi(s) - \psi(t)| = |(1 - |s|) - (1 - |t|)| = ||t| - |s|| \leq |t - s|,$$

per $|s| \leq 1, |t| > 1$ si ha

$$|\psi(s) - \psi(t)| = |\psi(s)| = 1 - |s| < |t| - |s| = ||t| - |s|| \leq |t - s|,$$

e per $|s|, |t| > 1$ si ha

$$|\psi(s) - \psi(t)| = 0 \leq |s - t|.$$

Per $\lambda > 0$, definiamo la funzione riscalata

$$\psi_\lambda(s) := \frac{1}{\lambda} \psi\left(\frac{s}{\lambda}\right).$$

Dalle proprietà di ψ si deduce direttamente (farlo per esercizio) che, per ogni $\lambda > 0$, la funzione ψ_λ è continua e soddisfa

$$\begin{aligned} 0 \leq \psi_\lambda(s) \leq \frac{1}{\lambda} \quad \forall s \in \mathbb{R}; \quad \psi_\lambda(s) = 0 \quad \forall s \notin [-\lambda, \lambda]; \\ \int_{-\lambda}^{\lambda} \psi_\lambda(s) ds = 1; \quad |\psi_\lambda(s) - \psi_\lambda(t)| \leq \frac{|s - t|}{\lambda^2} \quad \forall s, t \in \mathbb{R}. \end{aligned} \tag{6.8}$$

Definiamo

$$\rho_\lambda : \mathbb{R}^d \rightarrow \mathbb{R}, \quad \rho_\lambda(x) = \rho_\lambda(x_1, x_2, \dots, x_d) = \psi_\lambda(x_1)\psi_\lambda(x_2) \cdots \psi_\lambda(x_d).$$

Dalle proprietà di ψ_λ deduciamo immediatamente che

$$0 \leq \rho_\lambda(x) \leq \frac{1}{\lambda^d} \quad \forall x \in \mathbb{R}^d; \quad \rho_\lambda(x) = 0 \quad \forall x \notin [-\lambda, \lambda]^d.$$

Infatti non appena anche una sola coordinata x_k è al di fuori dell'intervallo $[-\lambda, \lambda]$, il fattore $\psi_\lambda(x_k)$ è nullo e quindi anche il prodotto $\rho_\lambda(x) = \psi_\lambda(x_1) \cdots \psi_\lambda(x_d)$ è nullo. Si ha poi

$$\int_{[-\lambda, \lambda]^d} \rho_\lambda(x) dx = 1$$

perché, usando le formule di riduzione e la definizione di ρ_λ , l'integrale si fattorizza come prodotto di d integrali

$$\int_{[-\lambda, \lambda]^d} \rho_\lambda(x) dx = \left(\int_{-\lambda}^{\lambda} \psi_\lambda(x_1) dx_1 \right) \left(\int_{-\lambda}^{\lambda} \psi_\lambda(x_2) dx_2 \right) \cdots \left(\int_{-\lambda}^{\lambda} \psi_\lambda(x_d) dx_d \right)$$

ciascuno dei quali vale 1. Infine ρ_λ è lipschitziana in \mathbb{R}^d ,

$$|\rho_\lambda(x) - \rho_\lambda(y)| \leq \frac{\sqrt{d}}{\lambda^{d+1}} |x - y| \quad \forall x, y \in \mathbb{R}^d.$$

Esercizio: dimostrarlo (lo vediamo nella prossima lezione).

7 Lezione 7

Riprendiamo la dimostrazione del Lemma 6.8. Proviamo che la funzione ρ_λ è lipschitziana in \mathbb{R}^d : per ogni $x, y \in \mathbb{R}^d$ si ha

$$\begin{aligned} \rho_\lambda(x) - \rho_\lambda(y) &= \psi_\lambda(x_1)\psi_\lambda(x_2) \cdots \psi_\lambda(x_d) - \psi_\lambda(y_1)\psi_\lambda(y_2) \cdots \psi_\lambda(y_d) \\ &= \left(\psi_\lambda(x_1) - \psi_\lambda(y_1) \right) \psi_\lambda(x_2) \cdots \psi_\lambda(x_d) \\ &\quad + \psi_\lambda(y_1) \left(\psi_\lambda(x_2) - \psi_\lambda(y_2) \right) \psi_\lambda(y_3) \cdots \psi_\lambda(x_d) \\ &\quad + \dots \\ &\quad + \psi_\lambda(y_1) \cdots \psi_\lambda(y_{d-1}) \left(\psi_\lambda(x_d) - \psi_\lambda(y_d) \right), \end{aligned}$$

quindi, usando le stime (6.8),

$$\begin{aligned} |\rho_\lambda(x) - \rho_\lambda(y)| &\leq \frac{1}{\lambda^{d+1}} (|x_1 - y_1| + |x_2 - y_2| + \dots + |x_d - y_d|) \\ &\leq \frac{\sqrt{d}}{\lambda^{d+1}} |x - y|. \end{aligned}$$

Nell'ultimo passaggio abbiamo usato Cauchy-Schwarz: $|x_1 - y_1| + \dots + |x_d - y_d|$ è il prodotto scalare del vettore $v = (|x_1 - y_1|, \dots, |x_d - y_d|)$ con il vettore $u = (1, \dots, 1)$, e poi $|v \cdot u| \leq |v||u| = |x - y|\sqrt{d}$.

Vale inoltre la seguente proprietà elementare di traslazione: per ogni $x, y \in \mathbb{R}^d$, la funzione $\rho_\lambda(x - y)$ è non nulla solo se l'argomento $x - y$ della funzione ρ_λ è nel "cubo" $[-\lambda, \lambda]^d$, cioè

$$-\lambda \leq x_i - y_i \leq \lambda \quad \forall i = 1, \dots, d$$

perché, come già osservato, la funzione ρ_λ vale zero fuori dal cubo $[-\lambda, \lambda]^d$. Quindi, per ogni x fissato, $\rho_\lambda(x - y)$ è non nulla solo se

$$x_i - \lambda \leq y_i \leq x_i + \lambda \quad \forall i = 1, \dots, d,$$

cioè se y appartiene al cubo di centro x e lato 2λ , che indichiamo

$$\begin{aligned} Q(x, \lambda) &:= [x_1 - \lambda, x_1 + \lambda] \times \dots \times [x_d - \lambda, x_d + \lambda] \\ &= \{y \in \mathbb{R}^d : x_i - \lambda \leq y_i \leq x_i + \lambda \quad \forall i = 1, \dots, d\}. \end{aligned}$$

Si ha inoltre, per ogni $x \in \mathbb{R}^d$,

$$\int_{Q(x, \lambda)} \rho_\lambda(x - y) dy = \int_{Q(0, \lambda)} \rho_\lambda(z) dz = 1 \quad (7.1)$$

(basta fare il cambio di variabile $x_i - y_i = z_i$ in ciascun integrale 1-dimensionale).

Siano Ω, K come nelle ipotesi. Poiché Ω è aperto e K compatto, esiste $r > 0$ tale che l'insieme

$$K_r := \{x + z : x \in K, z \in \mathbb{R}^d, |z| \leq r\} \quad (7.2)$$

è contenuto in Ω (esercizio: dimostrare l'esistenza di un tale r , per esempio per assurdo). Poiché K è compatto, anche K_r è compatto (esercizio: dimostrarlo, per esempio provando che K_r è chiuso e limitato).

Consideriamo ora un punto x nel compatto K e un punto y nel cubo $Q(x, \lambda)$. Per definizione di $Q(x, \lambda)$, si ha $|x_i - y_i| \leq \lambda$ per ogni $i = 1, \dots, d$, quindi

$$|x - y| = \sqrt{(x_1 - y_1)^2 + \dots + (x_d - y_d)^2} \leq \sqrt{d\lambda^2} = \sqrt{d}\lambda.$$

Se $\sqrt{d}\lambda \leq r$, allora $|x - y| \leq r$, e dunque $y \in K_r$ perché, posto $z := y - x$, si ha $y = x + z$ con $x \in K$ e $|z| \leq r$.

Abbiamo mostrato che, per $\lambda \leq r/\sqrt{d}$, si ha

$$Q(x, \lambda) \subseteq K_r \subset \Omega \quad \forall x \in K. \quad (7.3)$$

Per $\lambda \leq r/\sqrt{d}$, per ogni $x \in K$ definiamo

$$\varphi_\lambda(x) := \int_{Q(x, \lambda)} g(y) \rho_\lambda(x - y) dy.$$

La definizione è ben posta perché, per (7.3), ogni y nel dominio di integrazione $Q(x, \lambda)$ appartiene a Ω , e quindi $g(y)$ è definita. Inoltre $g(y)$ e $\rho_\lambda(x - y)$ sono funzioni continue di y , quindi $\varphi_\lambda(x)$ è un integrale di funzione continua su un cubo.

Proviamo che $|\varphi_\lambda(x)| \leq M$ per ogni $x \in K$. Dall'ipotesi $|g| \leq M$ in Ω , ricordando che $\rho_\lambda \geq 0$ e usando (7.1), si ha

$$\begin{aligned} |\varphi_\lambda(x)| &\leq \int_{Q(x, \lambda)} |g(y) \rho_\lambda(x - y)| dy \\ &\leq \int_{Q(x, \lambda)} M \rho_\lambda(x - y) dy = M \int_{Q(x, \lambda)} \rho_\lambda(x - y) dy = M. \end{aligned}$$

Proviamo che φ_λ è lipschitziana in K . Siano $x, z \in K$. Indichiamo, in breve, $Q_x := Q(x, \lambda)$ e $Q_z := Q(z, \lambda)$. La funzione $\rho_\lambda(x - y)$ vale zero per ogni y al di fuori del cubo Q_x , quindi in particolare è nulla per y nell'insieme $Q_z \setminus Q_x$. Dunque

$$\varphi_\lambda(x) = \int_{Q_x} g(y) \rho_\lambda(x - y) dy + \int_{Q_z \setminus Q_x} g(y) \rho_\lambda(x - y) dy$$

(l'ultimo integrale vale zero). I due insiemi di integrazione Q_x e $(Q_z \setminus Q_x)$ sono disgiunti, quindi possiamo riscrivere direttamente un unico integrale avente come dominio di integrazione la loro unione. L'unione dei due domini $Q_x \cup (Q_z \setminus Q_x)$ è uguale all'insieme $Q_x \cup Q_z$. Dunque

$$\varphi_\lambda(x) = \int_{Q_x \cup Q_z} g(y) \rho_\lambda(x - y) dy.$$

Ripetendo lo stesso ragionamento per $\varphi_\lambda(z)$, troviamo

$$\varphi_\lambda(z) = \int_{Q_x \cup Q_z} g(y) \rho_\lambda(z - y) dy.$$

Quindi

$$\varphi_\lambda(x) - \varphi_\lambda(z) = \int_{Q_x \cup Q_z} g(y) (\rho_\lambda(x - y) - \rho_\lambda(z - y)) dy.$$

Ora, ρ_λ è lipschitziana in \mathbb{R}^d , con costante di Lipschitz \sqrt{d}/λ^{d+1} . Quindi

$$\begin{aligned} |\varphi_\lambda(x) - \varphi_\lambda(z)| &\leq \int_{Q_x \cup Q_z} |g(y)| |\rho_\lambda(x-y) - \rho_\lambda(z-y)| dy \\ &\leq \int_{Q_x \cup Q_z} M\sqrt{d}\lambda^{-d-1}|x-z| dy \\ &= M\sqrt{d}\lambda^{-d-1}|x-z| \int_{Q_x \cup Q_z} dy \\ &\leq \frac{M}{\lambda} \sqrt{d} 2^{d+1} |x-z| \end{aligned}$$

perché

$$\int_{Q_x \cup Q_z} dy \leq \int_{Q_x} dy + \int_{Q_z} dy = (2\lambda)^d + (2\lambda)^d = 2^{d+1}\lambda^d$$

(“ \leq ” e non “ $=$ ” perché separando gli integrali un’eventuale intersezione tra i due cubi viene contata due volte). Quindi φ_λ è lipschitziana in K .

Infine proviamo la convergenza uniforme di φ_λ a g . Per $x \in K$, usando (7.1), si ha

$$\begin{aligned} \varphi_\lambda(x) - g(x) &= \int_{Q(x,\lambda)} g(y)\rho_\lambda(x-y) dy - \left(\int_{Q(x,\lambda)} \rho_\lambda(x-y) dy \right) g(x) \\ &= \int_{Q(x,\lambda)} (g(y) - g(x))\rho_\lambda(x-y) dy. \end{aligned}$$

Per ipotesi, g è continua; inoltre K_r è compatto, quindi g è uniformemente continua in K_r . Sia $\varepsilon > 0$. Dunque esiste $\delta_\varepsilon > 0$ tale che

$$|g(x) - g(y)| \leq \varepsilon \quad \forall x, y \in K_r, |x-y| \leq \delta_\varepsilon.$$

Come già osservato, per ogni y nel cubo $Q(x, \lambda)$ si ha $|x-y| \leq \sqrt{d}\lambda$. Se poi $\sqrt{d}\lambda \leq \delta_\varepsilon$, allora $|x-y| \leq \delta_\varepsilon$, e di conseguenza $|g(y) - g(x)| \leq \varepsilon$. Dunque (ricordando che $\lambda \leq r/\sqrt{d}$ è la condizione per avere φ_λ ben posta) si ha

$$|g(y) - g(x)| \leq \varepsilon \quad \forall y \in Q(x, \lambda), \quad \forall \lambda \leq \lambda_\varepsilon := \min\{r/\sqrt{d}, \delta_\varepsilon/\sqrt{d}\}.$$

Perciò

$$\begin{aligned} |\varphi_\lambda(x) - g(x)| &\leq \int_{Q(x,\lambda)} |g(y) - g(x)|\rho_\lambda(x-y) dy \\ &\leq \int_{Q(x,\lambda)} \varepsilon\rho_\lambda(x-y) dy = \varepsilon \int_{Q(x,\lambda)} \rho_\lambda(x-y) dy = \varepsilon. \end{aligned}$$

Abbiamo provato che per ogni $\varepsilon > 0$ esiste $\lambda_\varepsilon > 0$ tale che $|\varphi_\lambda(x) - g(x)| \leq \varepsilon$ per ogni $x \in K$ e ogni $\lambda \leq \lambda_\varepsilon$. Questo significa che φ_λ converge a g uniformemente in K per $\lambda \rightarrow 0$.

La successione

$$g_n(x) := \varphi_{\lambda_n}(x), \quad \lambda_n := \frac{r}{n\sqrt{d}}, \quad n = 1, 2, 3, \dots$$

soddisfa tutte le richieste della tesi. □

Osservazione 7.1. (*Convoluzione*). L'operazione che a due funzioni f, g associa la funzione

$$(f * g)(x) = \int f(y)g(x - y) dy$$

si chiama *prodotto di convoluzione*, o semplicemente *convoluzione*. Tra le sue speciali e importanti proprietà c'è questa: $(f * g)(x)$ ha la stessa regolarità della *più regolare* tra $f(x)$ e $g(x)$ (nel caso della dimostrazione appena conclusa, ρ_λ è lipschitziana, g è soltanto continua, e la loro convoluzione $\varphi_\lambda = g * \rho_\lambda$ è lipschitziana). Per esempio, se f è continua (o anche solo integrabile) e g è C^∞ , anche $(f * g)$ è C^∞ .

Questa proprietà non è affatto ovvia: per esempio, la somma $f(x) + g(x)$, il prodotto $f(x)g(x)$ e la composizione $f(g(x))$ hanno la stessa regolarità della *meno regolare* tra f e g . D'altra parte si intuisce dalla formula che, quando si deriva $(f * g)(x)$ (purché la derivata rispetto al parametro x e l'integrale in dy si possano scambiare), la derivata va a finire solo su $g(x - y)$ e non tocca $f(y)$, la quale, dunque, non ha bisogno di essere derivabile: $\partial_x(f * g)(x) = f * (\partial_x g)(x)$.

La convoluzione è uno dei metodi principali alla base dei risultati di approssimazione di funzioni (densità di spazi di funzioni molto regolari in spazi di funzioni meno regolari); è uno strumento classico dell'analisi funzionale, soprattutto nel caso in cui ρ_λ è una funzione C^∞ , con integrale unitario, che si annulla fuori da una palla di raggio λ (si dice in tal caso che (ρ_λ) è una famiglia di *mollificatori*).

Il Lemma 6.8 vale anche con le funzioni g_n di classe C^∞ , non soltanto lipschitziane: basta scegliere $\rho_\lambda \in C^\infty$ invece che solo lipschitziana. D'altra parte, il vantaggio della funzione ρ_λ che abbiamo usato nella dimostrazione sta nel fatto che i calcoli sono tutti espliciti ed elementari.

Possiamo interpretare la convoluzione $(g * \rho_\lambda)(x)$ anche come *media integrale pesata* della funzione g , con i pesi dati da $\rho_\lambda(x - y)$. Per λ molto piccolo, la gran parte del peso viene dato ai punti y molto vicini a x (ai punti y fuori dal cubo $Q(x, \lambda)$ non viene dato alcun peso nel calcolo della media integrale); al limite per $\lambda \rightarrow 0$, tutto il peso viene assegnato al solo punto x : per questa ragione, per $\lambda \rightarrow 0$, $(g * \rho_\lambda)(x)$ converge a $g(x)$ (che è la media integrale di $g(y)$ con il peso tutto concentrato in x : gli altri punti $y \neq x$ hanno peso nullo).

Torneremo a parlare della convoluzione e delle sue proprietà nell'ultima parte del corso, trattando le serie di Fourier.

Esercizio 7.2. (*Iterate di Picard e Teorema di esistenza di Peano*). Assumiamo le ipotesi del Teorema 6.6 di esistenza di Peano.

(i) Dimostrare che, se il campo $f(t, y)$ è continuo in $I \times J$ (e non necessariamente lipschitziano), la successione delle iterate di Picard, cioè la successione definita ricorsivamente da

$$u_0(t) = y_0, \quad u_n(t) = y_0 + \int_{t_0}^t f(s, u_{n-1}(s)) ds, \quad n \geq 1,$$

è una successione equilimitata ed equicontinua in $[t_0 - \delta, t_0 + \delta]$.

(ii) Perché allora non possiamo usare le iterate di Picard insieme al Teorema di Ascoli-Arzelà per dimostrare il Teorema di esistenza di Peano?

Riprendiamo l'analisi generale delle equazioni $u' = f(t, u)$ in \mathbb{R}^N per cui si ha esistenza e unicità locale della soluzione.

Definizione 7.3. (*Prolungamento*). Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ una funzione continua e localmente lipschitziana in y uniformemente in t . Sia $(a, b) \subseteq \mathbb{R}$ un intervallo, e sia $u : (a, b) \rightarrow \mathbb{R}^N$ una soluzione dell'equazione differenziale $u' = f(t, u)$ in (a, b) . Una funzione $v : (a_1, b_1) \rightarrow \mathbb{R}^N$ definita su un intervallo (a_1, b_1) è detta un *prolungamento* di u se

1. $(a, b) \subseteq (a_1, b_1)$,
2. $v = u$ in (a, b) (cioè $v(t) = u(t)$ per ogni $t \in (a, b)$),
3. v è soluzione dell'equazione differenziale in (a_1, b_1) (cioè $(t, v(t)) \in \Omega$ per ogni $t \in (a_1, b_1)$ e $v'(t) = f(t, v(t))$ per ogni $t \in (a_1, b_1)$).

Quando un tale prolungamento di v esiste, diciamo che u è *prolungabile all'intervallo* (a_1, b_1) .

Osservazione 7.4. (*Unicità del prolungamento sull'intersezione*). Nelle ipotesi della Definizione 7.3, se

$$v_1 : (a_1, b_1) \rightarrow \mathbb{R}^N, \quad v_2 : (a_2, b_2) \rightarrow \mathbb{R}^N$$

sono due prolungamenti di $u : (a, b) \rightarrow \mathbb{R}^N$, allora

$$v_1(t) = v_2(t) \quad \forall t \in (a_1, b_1) \cap (a_2, b_2).$$

Dimostrazione. Sia $t_0 \in (a, b)$, e sia $y_0 := u(t_0)$. Sia v_1 che v_2 sono soluzioni del problema di Cauchy

$$v' = f(t, v), \quad v(t_0) = y_0$$

nell'intervallo $I := (a_1, b_1) \cap (a_2, b_2)$, quindi (vedi Corollario 3.15) coincidono in tutto l'intervallo I . \square

Teorema 7.5. (*Esistenza del prolungamento massimale*). Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ continua e localmente lipschitziana in y uniformemente in t . Sia $u : (a, b) \rightarrow \mathbb{R}^N$ una soluzione dell'equazione $u' = f(t, u)$. Sia

$$T_+ := \sup \{b_1 \geq b : u \text{ è prolungabile ad } (a, b_1)\},$$

$$T_- := \inf \{a_1 \leq a : u \text{ è prolungabile ad } (a_1, b)\}.$$

Allora u è prolungabile a (T_-, T_+) .

Detto v il prolungamento di u a (T_-, T_+) , v è prolungamento massimale, cioè v non ammette alcun prolungamento ad un intervallo che contenga strettamente (T_-, T_+) . In altre parole, se $w : I \rightarrow \mathbb{R}^N$ è un prolungamento di v , allora $I = (T_-, T_+)$ e $w = v$.

Una tale soluzione v è detta *soluzione massimale*; il corrispondente intervallo (T_-, T_+) è detto *intervallo massimale*.

Dimostrazione. Definiamo v nell'intervallo (T_-, T_+) nel modo seguente.

Come prima cosa, definiamo $v(t) := u(t)$ per ogni $t \in (a, b)$. Se $T_+ = b$, allora non c'è bisogno di definire v a destra di b . Se invece $T_+ > b$, dobbiamo definire v nell'intervallo $[b, T_+)$. Sia

$$B := \{b_1 \geq b : u \text{ è prolungabile ad } (a, b_1)\},$$

cosicché $T_+ = \sup B$. Sia $t \in [b, T_+)$. Dalla definizione di \sup , esiste $b_1 \in B$ tale che $t < b_1 \leq T_+$. Dalla definizione di B , u è prolungabile ad (a, b_1) , e diciamo w tale prolungamento. Definiamo $v(t) := w(t)$. In questo modo abbiamo definito $v(t)$ per ogni $t \in [b, T_+)$. Questa definizione è ben posta: per l'Osservazione 7.4, qualunque altro prolungamento di u che sia definito in t porta alla stessa definizione di $v(t)$.

A sinistra procediamo in modo analogo, e otteniamo v definita su tutto l'intervallo (T_-, T_+) .

Per costruzione, v è prolungamento di u . Se v ammettesse un prolungamento w a (T_-, β) con $\beta > T_+$, allora w sarebbe anche un prolungamento di u a (T_-, β) , quindi β apparterebbe all'insieme B , contro la definizione di $T_+ = \sup B$, assurdo. Similmente, v non ammette prolungamenti ad (α, T_+) con $\alpha < T_-$. Quindi v è prolungamento massimale. \square

Proposizione 7.6. (Criterio di prolungabilità: limite di successione). *Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ continua e localmente lipschitziana in y uniformemente in t . Sia $u : (a, b) \rightarrow \mathbb{R}^N$ soluzione dell'equazione $u' = f(t, u)$ nell'intervallo (a, b) , e supponiamo che esista una successione $(t_n)_{n \in \mathbb{N}}$ e un punto $v_0 \in \mathbb{R}^N$ tali che*

- $t_n \in (a, b)$ per ogni $n \in \mathbb{N}$,
- $t_n \rightarrow b$ per $n \rightarrow \infty$,
- $u(t_n) \rightarrow v_0$ per $n \rightarrow \infty$

con

$$(b, v_0) \in \Omega.$$

Allora u è prolungabile a destra di b , cioè esiste un prolungamento di u ad (a, b_1) per un certo $b_1 > b$.

Analogamente, se esiste una successione (t_n) in (a, b) tale che $t_n \rightarrow a$ e $u(t_n) \rightarrow v_1$, con $(a, v_1) \in \Omega$, allora u è prolungabile a sinistra di a , cioè esiste un prolungamento di u ad (a_1, b) per un certo $a_1 < a$.

Dimostrazione nella prossima lezione.

8 Lezione 8

Dimostrazione del Criterio di prolungabilità con limite di successione (Proposizione 7.6).
Poiché $(b, v_0) \in \Omega$ e Ω è aperto, esistono $r > 0$, $\rho > 0$ tali che, posto

$$I := [b - r, b + r], \quad J := \overline{B(v_0, \rho)},$$

il rettangolo $I \times J$ sia contenuto in Ω . Poiché f è continua e $I \times J$ è compatto, f è limitata in $I \times J$, diciamo $|f(t, y)| \leq M$ per ogni $(t, y) \in I \times J$. Fissiamo una costante δ tale che

$$0 < \delta \leq r/2, \quad M\delta \leq \rho/2.$$

Per esempio, fissiamo

$$\delta := \min\{r/2, \rho/(2M + 1)\}. \quad (8.1)$$

Definiamo $y_n := u(t_n)$,

$$I_n := [t_n - r/2, t_n + r/2], \quad J_n := \overline{B(y_n, \rho/2)}.$$

Per ipotesi, $t_n \rightarrow b$, $y_n \rightarrow v_0$, quindi esiste $\bar{n}_1 \in \mathbb{N}$ tale che

$$|b - t_n| = b - t_n \leq r/2, \quad |y_n - v_0| \leq \rho/2 \quad \forall n \geq \bar{n}_1.$$

Questo implica che

$$I_n \times J_n \subseteq I \times J \quad \forall n \geq \bar{n}_1.$$

Infatti per ogni $y \in J_n$ si ha $|y - y_n| \leq \rho/2$ per definizione di J_n , e, per $n \geq \bar{n}_1$, si ha anche $|y_n - v_0| \leq \rho/2$, da cui

$$|y - v_0| = |y - y_n + y_n - v_0| \leq |y - y_n| + |y_n - v_0| \leq \rho,$$

cioè $y \in J$; quindi $J_n \subseteq J$ per ogni $n \geq \bar{n}_1$. Analogamente si ha $I_n \subseteq I$ per $n \geq \bar{n}_1$. Di conseguenza

$$|f(t, y)| \leq M \quad \forall (t, y) \in I_n \times J_n, \quad \forall n \geq \bar{n}_1.$$

Ricapitolando, per $n \geq \bar{n}_1$, nel rettangolo $I_n \times J_n$ si ha f continua, lipschitziana in y uniformemente in t , con $|f| \leq M$, e $0 < \delta \leq r/2$, $M\delta \leq \rho/2$. Dunque, per $n \geq \bar{n}_1$, dal Teorema 2.1 di esistenza e unicità locale, il problema di Cauchy

$$\varphi' = f(t, \varphi), \quad \varphi(t_n) = y_n \quad (8.2)$$

ha un'unica soluzione nell'intervallo $[t_n, t_n + \delta]$. Notiamo che δ è stato fissato in (8.1), e non dipende da n .

Ricordando che $t_n \rightarrow b$ per $n \rightarrow \infty$, *fissiamo* un intero $n \geq \bar{n}_1$ tale che

$$|b - t_n| = b - t_n < \delta.$$

Quindi

$$a < t_n < b < t_n + \delta.$$

Sia φ la soluzione del problema di Cauchy (8.2) definita in $[t_n, t_n + \delta]$. Sia φ che u soddisfano l'equazione differenziale nell'intervallo $[t_n, b)$, e coincidono in t_n ; quindi, dal Corollario 3.15,

$$u(t) = \varphi(t) \quad \forall t \in [t_n, b).$$

Sia v la funzione così definita:

$$v(t) = \begin{cases} u(t) & \text{per } t \in (a, t_n), \\ u(t) = \varphi(t) & \text{per } t \in [t_n, b), \\ \varphi(t) & \text{per } t \in [b, t_n + \delta]. \end{cases}$$

La funzione $v(t)$ coincide con $u(t)$ per ogni $t \in (a, b)$, e soddisfa l'equazione differenziale in ogni $t \in (a, t_n + \delta]$ perché

$$u'(t) = f(t, u(t)) \quad \forall t \in (a, b), \quad \varphi'(t) = f(t, \varphi(t)) \quad \forall t \in [t_n, t_n + \delta].$$

Dunque v è un prolungamento di u ed è definita anche a destra di b , nell'intervallo $(b, t_n + \delta]$. La dimostrazione del caso $t_n \rightarrow a$, $u(t_n) \rightarrow v_1$ è analoga. \square

Il corollario che segue è un'immediata conseguenza della Proposizione 8.1.

Corollario 8.1. (Criterio di prolungabilità: limite continuo). *Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ continua e localmente lipschitziana in y uniformemente in t . Sia $u : (a, b) \rightarrow \mathbb{R}^N$ soluzione dell'equazione $u' = f(t, u)$ nell'intervallo (a, b) . Se esiste*

$$\lim_{t \rightarrow b^-} u(t) = v_0, \quad \text{con } (b, v_0) \in \Omega,$$

allora u è prolungabile a destra di b , cioè esiste un prolungamento di u ad (a, b_1) per un certo $b_1 > b$.

Analogamente, se esiste

$$\lim_{t \rightarrow a^+} u(t) = v_1, \quad \text{con } (a, v_1) \in \Omega,$$

allora u è prolungabile a sinistra di a , cioè esiste un prolungamento di u ad (a_1, b) per un certo $a_1 < a$.

Dimostrazione. Supponiamo che $u(t) \rightarrow v_0$ per $t \rightarrow b^-$, con $(b, v_0) \in \Omega$. Sia (t_n) una qualunque successione in (a, b) che converge a b . Allora $u(t_n)$ converge a v_0 , e dalla Proposizione 7.6 otteniamo la tesi. Analogamente per $t \rightarrow a^+$. \square

Teorema 8.2. (Teorema di fuga dai compatti). *Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ continua e localmente lipschitziana in y uniformemente in t . Sia $u : (T_-, T_+) \rightarrow \mathbb{R}^N$ soluzione dell'equazione $u' = f(t, u)$, con (T_-, T_+) intervallo massimale. Per ogni compatto $K \subset \Omega$ esistono $\alpha, \beta \in \mathbb{R}$ con*

$$T_- < \alpha \leq \beta < T_+$$

tali che

$$(t, u(t)) \notin K \quad \forall t \in (T_-, \alpha) \cup (\beta, T_+).$$

In altre parole, per ogni t fuori dall'intervallo $[\alpha, \beta]$ la coppia $(t, u(t))$ si trova al di fuori del compatto K .

Dimostrazione. Supponiamo che non esista alcun β per cui $(t, u(t)) \notin K$ per ogni $t \in (\beta, T_+)$. Questo significa che per ogni $\beta \in (T_-, T_+)$ esiste $t \in (\beta, T_+)$ tale che $(t, u(t)) \in K$. Allora, presa una qualunque successione $(\beta_n) \subset (T_-, T_+)$ con

$\beta_n \rightarrow T_+$, per ogni n esiste $t_n \in (\beta_n, T_+)$ tale che $(t_n, u(t_n)) \in K$. Quindi la successione $(t_n)_{n \in \mathbb{N}}$ soddisfa

$$\begin{cases} t_n \in (T_-, T_+) & \forall n \in \mathbb{N}, \\ t_n \rightarrow T_+ & (n \rightarrow \infty), \\ (t_n, u(t_n)) \in K & \forall n \in \mathbb{N}. \end{cases}$$

Per compattezza, esiste una sottosuccessione $(t_{n_k}, u(t_{n_k}))$ convergente. Poiché $t_n \rightarrow T_+$, anche la sua sottosuccessione t_{n_k} ha lo stesso limite. Sia v_0 il limite di $u(t_{n_k})$, cosicché $(t_{n_k}, u(t_{n_k}))$ converge a (T_+, v_0) . Essendo K chiuso, (T_+, v_0) appartiene a K . Inoltre $K \subset \Omega$, quindi $(T_+, v_0) \in \Omega$. Dal criterio di prolungabilità con successione (Proposizione 7.6), u è prolungabile a destra di T_+ , ma questo è impossibile perché u , per ipotesi, è soluzione massimale. Abbiamo dunque provato l'esistenza di β come nell'enunciato. L'esistenza di α si prova in modo analogo. \square

In altre parole, il Teorema 8.2 dice che, per t che tende a T_+ o a T_- , la coppia $(t, u(t))$ esce definitivamente da qualunque compatto $K \subset \Omega$.

Osservazione 8.3. (*Compatti in spazio-tempo*). Nel Teorema di fuga dai compatti, il compatto K non è un compatto dello spazio delle fasi \mathbb{R}^N (lo spazio “delle y ” nella nostra notazione abituale, dove vive $u(t)$), ma un compatto contenuto in $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$, i cui elementi sono coppie (t, y) di “tempo e spazio”, cioè coppie in cui t è il tempo e y è un vettore dello spazio delle fasi \mathbb{R}^N .

Come conseguenza (Corollario 8.4) del Teorema di fuga dai compatti, se una soluzione $u(t)$ rimane confinata in una regione compatta $G \subset \mathbb{R}^N$ dello spazio delle fasi, allora deve esistere per tutti i tempi, non può cioè cessare di esistere ad un tempo finito. In altre parole, poiché le soluzioni devono uscire definitivamente da ogni compatto, se $u(t)$ rimane in un compatto (dello spazio), deve essere il tempo t a evitare il confinamento in un compatto (del tempo), in modo tale che la coppia $(t, u(t))$ non sia confinata in un compatto (dello spazio-tempo).

Il punto è che le soluzioni delle equazioni differenziali non possono smettere di esistere “senza un motivo”: se una soluzione $u(t)$ è definita per tempi $t \in (a, b)$, ciò che $u(t)$ può fare per $t \rightarrow b^-$ (o per $t \rightarrow a^+$) è scappare all'infinito $|u(t)| \rightarrow \infty$, o accumularsi alla frontiera $\partial\Omega$ del dominio Ω del campo, oppure, se ciò non accade, u continua a esistere anche a destra di b (o a sinistra di a).

Corollario 8.4. (Compatto nello spazio delle fasi implica esistenza globale nel tempo). *Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ continua e localmente lipschitziana in y uniformemente in t . Sia $u : (T_-, T_+) \rightarrow \mathbb{R}^N$ soluzione dell'equazione $u' = f(t, u)$, con (T_-, T_+) intervallo massimale. Sia $t_0 \in (T_-, T_+)$, e sia $G \subset \mathbb{R}^N$ un insieme compatto. Se*

$$[t_0, +\infty) \times G \subset \Omega, \quad u(t) \in G \quad \forall t \in [t_0, T_+),$$

allora $T_+ = +\infty$. Analogamente, se

$$(-\infty, t_0] \times G \subset \Omega, \quad u(t) \in G \quad \forall t \in (T_-, t_0],$$

allora $T_- = -\infty$. In particolare, se

$$\mathbb{R} \times G \subset \Omega, \quad u(t) \in G \quad \forall t \in (T_-, T_+),$$

allora u è soluzione globale, cioè $(T_-, T_+) = \mathbb{R}$.

Dimostrazione. Sia $u(t) \in G$ per ogni $t \in [t_0, T_+)$, e supponiamo, per assurdo, che $T_+ < \infty$. Allora la coppia $(t, u(t))$ appartiene al compatto $K := [t_0, T_+] \times G \subset \Omega$ per ogni $t \in [t_0, T_+)$; questo va contro il Teorema di fuga dai compatti, assurdo. Analoghi gli altri casi. \square

Esercizio 8.5. Si consideri il problema di Cauchy

$$\begin{cases} u' = 2u - u^2, \\ u(0) = y_0. \end{cases}$$

- 1) Per $y_0 \in [0, 2]$ calcolare l'intervallo massimale della soluzione.
- 2) Per ogni $y_0 \in \mathbb{R}$ calcolare la soluzione esplicitamente.
- 3) Disegnare l'andamento qualitativo delle soluzioni.

Svolgimento. 1) Il campo vettoriale $f(t, y) = 2y - y^2 = y(2 - y)$ è definito nell'aperto $\Omega = \mathbb{R}^2$ e ha due zeri, $y = 0$ e $y = 2$. Quindi il sistema ha due equilibri, le soluzioni costanti $u(t) = 0$ e $u(t) = 2$ definite per ogni $t \in \mathbb{R}$.

La funzione f è di classe $C^\infty(\mathbb{R}^2)$, quindi è lipschitziana in y , e dunque, per unicità della soluzione locale, due soluzioni distinte non si possono incontrare (Corollario 3.15). Di conseguenza, la soluzione u del problema di Cauchy con dato iniziale $u(0) = y_0 \in (0, 2)$ rimane confinata tra i due equilibri per tutta la sua esistenza, cioè

$$0 < u(t) < 2 \quad \forall t \in (T_-, T_+)$$

dove (T_-, T_+) è l'intervallo massimale di esistenza della soluzione u .

Poi $\mathbb{R} \times [0, 2]$ è contenuto in $\Omega = \mathbb{R}^2$, $u(t) \in [0, 2]$ per ogni t nel suo intervallo massimale, e quindi, dal Corollario 8.4 (che applichiamo con $G = [0, 2]$), l'intervallo massimale è \mathbb{R} , cioè u è soluzione globale.

Altro modo. Sia $y_0 \in (0, 2)$, e supponiamo per assurdo che $T_+ < \infty$. Poiché $u(t) \in (0, 2)$ per ogni $t \in (T_-, T_+)$, si ha

$$u'(t) = f(t, u(t)) = u(t)(2 - u(t)) > 0 \quad \forall t \in (T_-, T_+).$$

Quindi u è crescente, e per monotonia esiste il suo limite per $t \rightarrow T_+$. Essendo $0 < u(t) < 2$, il limite $v_0 := \lim_{(t \rightarrow T_+)} u(t)$ appartiene all'intervallo $[0, 2]$. Perciò $(T_+, v_0) \in \Omega = \mathbb{R}^2$, e dal Criterio di prolungabilità con limite continuo (Corollario 8.1) u è prolungabile a destra di T_+ , assurdo. Questo dimostra che $T_+ = \infty$. In modo simile si dimostra che $T_- = -\infty$.

2) È un esercizio sulle equazioni a variabili separabili. Con calcoli elementari si trova che per ogni $y_0 \in \mathbb{R}$, $y_0 \neq 0, 2$,

$$u(t) = \frac{2}{1 + ce^{-2t}}, \quad c = \frac{2 - y_0}{y_0}, \quad (8.3)$$

con intervallo massimale

$$(T_-, T_+) = \begin{cases} (t_*, \infty) & \text{per } y_0 > 2, \\ (-\infty, \infty) & \text{per } y_0 \in [0, 2], \\ (-\infty, t_*) & \text{per } y_0 < 0, \end{cases} \quad t_* := \frac{1}{2} \log \left(1 - \frac{2}{y_0} \right).$$

3) Chiaramente si possono disegnare le soluzioni facendo un breve “studio di funzione” a partire dalla formula esplicita (8.3). Tuttavia in questo contesto è più istruttivo dedurre l'andamento delle soluzioni a partire dal campo $f(t, y)$.

Possiamo rappresentare graficamente il campo f in questo modo: tracciamo un piano cartesiano con il tempo t come ascissa e la variabile y (spazio delle fasi) come ordinata; in ogni punto (t, y) del piano rappresentiamo $f(t, y)$ disegnando una freccina che punta verso la direzione positiva del tempo (cioè verso destra) e che ha pendenza positiva se f è positiva, pendenza negativa se f è negativa, e pendenza nulla (freccina orizzontale) se f è nulla. Possiamo essere più precisi, disegnando una freccina poco inclinata, cioè con pendenza piccola, se il valore di f è piccolo, e molto inclinata se f è grande; un valore grandissimo del campo f corrisponde dunque a una freccina quasi verticale.

Poiché il campo è autonomo, si ha $f(t, y) = f(s, y)$ per ogni $t, s \in \mathbb{R}$, quindi le freccine che si trovano su una qualunque retta orizzontale del piano (t, y) avranno tutte la stessa pendenza.

Dopo aver rappresentato il campo $f(t, y)$ nei punti (t, y) del piano tramite le freccine, l'andamento delle traiettorie dell'equazione differenziale diventa evidente: se $u(t)$ è soluzione dell'equazione $u'(t) = f(t, u(t))$, il suo grafico $\{(t, u(t))\}$ è una curva che all'istante t ha pendenza $u'(t)$ pari al valore che ha il campo f in quell'istante t e in quel punto $y = u(t)$; in altre parole, u deve seguire le freccine istante per istante.

Si può dunque completare l'esercizio facendo dapprima la rappresentazione del campo f tramite il disegno delle freccine in vari punti del piano, e poi disegnando le curve che seguono le freccine. Ne risulta un disegno approssimativo del grafico delle soluzioni, che comunque individua qualitativamente il loro andamento (un po' come nel classico "studio di funzione" di Analisi 1). \square

Vediamo in dettaglio un esercizio sullo studio qualitativo delle soluzioni (è l'esercizio 13 di [1]).

Esercizio 8.6. Si consideri il problema di Cauchy

$$y'(t) = \frac{1}{y(t) - t^2}, \quad y(0) = a \quad (a > 0). \quad (8.4)$$

- 1) Dimostrare che esiste un'unica soluzione in un intorno di $t = 0$. Calcolare y' , y'' , y''' in $t = 0$.
- 2) Dimostrare che la soluzione è definita in $[0, \infty)$ e ha limite per $t \rightarrow +\infty$; calcolare tale limite.

Svolgimento. 1) Il campo vettoriale $f(t, y) = 1/(y - t^2)$ è definito nell'insieme $\{(t, y) \in \mathbb{R}^2 : y \neq t^2\}$, che è l'unione di due aperti connessi, le regioni $\{y > t^2\}$ e $\{y < t^2\}$, cioè sopra/sotto la parabola $\mathcal{P} = \{y = t^2\}$. Il dato iniziale $(0, a)$ si trova nella regione $\{y > t^2\}$, quindi consideriamo l'aperto connesso

$$\Omega := \{(t, y) \in \mathbb{R}^2 : y > t^2\}$$

come dominio del campo f .

Poiché $f \in C^\infty(\Omega)$, f è localmente lipschitziana in y uniformemente in t , quindi, dal Teorema di esistenza e unicità locale, esiste, unica, la soluzione del problema di Cauchy in un intorno di $t = 0$.

Per calcolare $y'(0)$ basta usare l'equazione differenziale in $t = 0$: si trova $y'(0) = 1/y(0) = 1/a$. Per calcolare le derivate successive di y deriviamo l'equazione differenziale $y' = (y - t^2)^{-1}$: troviamo

$$\begin{aligned}y'' &= -(y - t^2)^{-2}(y' - 2t), \\y''' &= 2(y - t^2)^{-3}(y' - 2t)^2 - (y - t^2)^{-2}(y'' - 2).\end{aligned}$$

Calcolando in $t = 0$, e usando le identità $y(0) = a$, $y'(0) = 1/a$, troviamo $y''(0) = -1/a^3$ e poi $y'''(0) = 3/a^5 + 2/a^2$.

2) Dal Teorema del prolungamento massimale, possiamo supporre che la soluzione y sia definita sul suo intervallo massimale di esistenza, che indichiamo (α, β) . Dobbiamo dimostrare che $\beta = +\infty$.

Osserviamo che $f > 0$ in Ω . Quindi $y' = f(t, y) > 0$ per ogni $t \in (\alpha, \beta)$. Perciò y è crescente in (α, β) , e dunque ammette limite per $t \rightarrow \beta$; chiamiamo

$$c := \lim_{t \rightarrow \beta} y(t). \tag{8.5}$$

Supponiamo, per assurdo, che $\beta < \infty$. Per ogni $t \in (\alpha, \beta)$ si ha $(t, y(t)) \in \Omega$, cioè

$$y(t) > t^2 \quad \forall t \in (\alpha, \beta).$$

Passando al limite per $t \rightarrow \beta$ si trova dunque $c \geq \beta^2$.

Se $\beta^2 < c < \infty$, allora la coppia (β, c) appartiene a Ω , e per il Criterio di prolungabilità con limite continuo (Corollario 8.1) la soluzione sarebbe prolungabile a destra di β , assurdo. Dunque $c = \beta^2$ oppure $c = +\infty$.

Esercizio: provare che $c = +\infty$; poi provare che $y(t) - t^2 \rightarrow 0$ per $t \rightarrow +\infty$. Questo esercizio verrà completato nella prossima lezione.

9 Lezione 9

Proseguiamo l'esercizio. Supponiamo che $c = \beta^2$. Allora $y(t) - t^2 \rightarrow c - \beta^2 = 0$ per $t \rightarrow \beta$, ed essendo $y(t) - t^2 > 0$ per ogni $t \in (\alpha, \beta)$, si ha

$$\lim_{t \rightarrow \beta} y'(t) = \lim_{t \rightarrow \beta} \frac{1}{y(t) - t^2} = +\infty. \quad (9.1)$$

Mostriamo che questo è incompatibile con il fatto che in $t = \beta$ la curva $y(t)$ incontri la parabola \mathcal{P} provenendo dalla regione al di sopra della parabola. L'idea geometrica è semplice: se la curva $y(t)$ arriva a incontrare la parabola con tangente verticale diretta verso l'alto (come dice il limite (9.1)), allora deve per forza arrivarci dalla regione al di sotto della parabola.

Da (9.1) segue che per ogni $M > 0$ esiste $\delta_0 > 0$ tale che $y'(t) \geq M$ per ogni $t \in (\beta - \delta_0, \beta)$ (M verrà fissata tra poco; per ora lasciamola libera). Quindi per ogni $t, t_1 \in (\beta - \delta_0, \beta)$ con $\beta - \delta_0 < t < t_1 < \beta$ si ha

$$y(t_1) - y(t) = \int_t^{t_1} y'(s) ds \geq M(t_1 - t).$$

Passando al limite per $t_1 \rightarrow \beta$ nella disuguaglianza $y(t_1) - y(t) \geq M(t_1 - t)$, e ricordando che $y(t_1) \rightarrow c = \beta^2$ per $t_1 \rightarrow \beta$, otteniamo

$$\beta^2 - y(t) \geq M(\beta - t), \quad \forall t \in (\beta - \delta_0, \beta).$$

Poi $y(t) > t^2$ per ogni $t \in (\alpha, \beta)$, quindi

$$M(\beta - t) \leq \beta^2 - y(t) < \beta^2 - t^2 \quad \forall t \in (\beta - \delta_0, \beta),$$

da cui

$$M < \frac{\beta^2 - t^2}{\beta - t} = \beta + t < 2\beta \quad \forall t \in (\beta - \delta_0, \beta).$$

La disuguaglianza $M < 2\beta$ è in contrasto con l'arbitrarietà di M : possiamo quindi tornare indietro, fissare $M \geq 2\beta$ (per esempio, scegliamo $M = 2\beta + 1$) e arriviamo alla conclusione che $M < 2\beta$, assurdo. Questo dimostra che non può essere $c = \beta^2$.

Quindi $c = +\infty$, cioè $y(t) \rightarrow +\infty$ per $t \rightarrow \beta$. Ricordando che, per ipotesi, β è finito, si ha

$$\lim_{t \rightarrow \beta} y'(t) = \lim_{t \rightarrow \beta} \frac{1}{y(t) - t^2} = 0.$$

Ma anche questo è impossibile: $y(t) \rightarrow +\infty$ per $t \rightarrow \beta$, con $\beta \in \mathbb{R}$, significa che la funzione $y(t)$ ha la retta $t = \beta$ come asintoto verticale da sinistra, quindi la sua derivata $y'(t)$ non può certo tendere a zero avvicinandosi all'asintoto. Per dimostrarlo rigorosamente possiamo usare un ragionamento simile a prima: $y'(t) \rightarrow 0$ per $t \rightarrow \beta$ significa che per ogni $\varepsilon > 0$ esiste un $\delta_1 > 0$ tale che $|y'(t)| \leq \varepsilon$ per ogni $t \in [\beta - \delta_1, \beta)$. Scegliendo $\varepsilon = 1$, e usando anche il fatto che $y'(t) > 0$ per ogni $t \in (\alpha, \beta)$, si ha che esiste $\delta_1 > 0$ tale che

$$0 < y'(t) \leq 1 \quad \forall t \in [\beta - \delta_1, \beta).$$

Per ogni $t \in [\beta - \delta_1, \beta)$ si ha

$$y(t) - y(\beta - \delta_1) = \int_{\beta - \delta_1}^t y'(s) ds \leq \int_{\beta - \delta_1}^t 1 ds = t - (\beta - \delta_1) \leq \delta_1,$$

da cui

$$y(t) \leq C = y(\beta - \delta_1) + \delta_1 \quad \forall t \in [\beta - \delta_1, \beta),$$

contro il fatto che $y(t) \rightarrow +\infty$ per $t \rightarrow \beta$, assurdo. Abbiamo dimostrato che in nessun caso β può essere finito. Dunque $\beta = +\infty$.

Poiché $y(t) > t^2$ per ogni $t \in (\alpha, \beta) = (\alpha, +\infty)$, e $t^2 \rightarrow +\infty$ per $t \rightarrow +\infty$, si ha

$$\lim_{t \rightarrow +\infty} y(t) = +\infty.$$

Questo conclude l'esercizio. □

Esercizio 9.1. Sia $y(t)$ come nell'esercizio precedente. Dimostrare che la soluzione $y(t)$ è asintotica alla parabola t^2 , cioè che la differenza $y(t) - t^2$ tende a 0 per $t \rightarrow +\infty$.

Svolgimento. Sia $u(t) := y(t) - t^2$ la differenza da studiare. Che equazione differenziale soddisfa u ? Usando l'equazione differenziale dell'esercizio precedente, calcoliamo

$$u'(t) = y'(t) - 2t = \frac{1}{y(t) - t^2} - 2t = \frac{1}{u(t)} - 2t.$$

Quindi $u(t)$ è soluzione del problema di Cauchy

$$u'(t) = \frac{1}{u(t)} - 2t \quad \forall t \in [0, +\infty), \quad u(0) = a,$$

e inoltre $u(t) > 0$ per ogni $t \geq 0$ (perché, dall'esercizio precedente, $y(t) > t^2$).

Il campo

$$F(t, u) := \frac{1}{u} - 2t$$

è positivo se $u < 1/2t$, negativo se $u > 1/2t$ e nullo se $u = 1/2t$. Dunque

- la soluzione $u(t)$ è strettamente crescente quando si trova nella regione al di sotto dell'iperbole $\gamma(t) := 1/2t$, cioè quando $u(t) < \gamma(t)$;
- $u(t)$ è strettamente decrescente quando si trova nella regione al di sopra dell'iperbole, cioè quando $u(t) > \gamma(t)$;
- $u(t)$ ha derivata nulla quando u incontra l'iperbole, cioè quando $u(t) = \gamma(t)$.

Questa analisi del segno del campo $F(t, u)$ ci fa intuire l'andamento della soluzione $u(t)$ in $[0, \infty)$:

- al tempo $t = 0$ la soluzione u parte dal punto $(0, a)$ con pendenza positiva $u'(0) = 1/u(0) = 1/a$;
- per un certo intervallo di tempo u resta al di sotto dell'iperbole γ , e cresce, mentre γ decresce;
- u sale, e dunque $u(t) > u(0) = a$, mentre γ scende, e aspettando un tempo sufficiente γ scende al di sotto della quota a , dunque necessariamente u e γ arrivano ad incontrarsi;
- all'istante dell'incontro $u(t) = \gamma(t)$, u ha tangente orizzontale; subito dopo l'incontro, u avanza con pendenza vicina a zero, mentre γ ha sempre pendenza negativa: quindi, dopo l'incontro, u rimane al di sopra di γ , e dunque comincia a decrescere;

- per un motivo geometrico simile a quello dell'esercizio precedente, non ci può essere un secondo punto di incontro di u e γ : infatti, dove u e γ si incontrano, la pendenza di u deve essere nulla, ma questo, essendo γ decrescente, è possibile solo se u proviene dalla regione al di sotto di γ ed entra nella regione al di sopra di γ , non viceversa;
- non avendo più intersezioni con γ , u rimane per sempre al di sopra di γ , decrescendo. Da qui si arriva alla conclusione.

Dimostriamo tutto ciò.

In $t = 0$ si ha $u(0) = a$. Dalla continuità di u si ha $u(t) \in [a/2, 3a/2]$ per ogni $t \in [0, \delta_0]$, per un certo $\delta_0 > 0$, mentre $\gamma(t) = 1/2t \geq 2a$ per $t \in (0, 1/4a]$. Quindi $u(t) < \gamma(t)$ per ogni t in $(0, \delta_1]$, $\delta_1 = \min\{\delta_0, 1/4a\}$.

Se fosse $u(t) < \gamma(t)$ per ogni $t \in (0, \infty)$, u sarebbe crescente in $(0, \infty)$, da cui $u(t) > u(0) = a$ per ogni $t > 0$. Per $t > 1/2a$ si avrebbe allora $\gamma(t) < a < u(t) < \gamma(t)$, assurdo. Questo prova che non può essere $u < \gamma$ in $(0, \infty)$.

Sia

$$t_* := \sup\{t \in (0, \infty) : u < \gamma \text{ in } [0, t]\}.$$

Da quanto dimostrato, $t_* < \infty$ (e, di più, abbiamo dimostrato che $t_* < 1/2a$). Per definizione di t_* e per la continuità di u e γ si ha che

$$u(t_*) = \gamma(t_*).$$

Infatti, se fosse $u(t_*) > \gamma(t_*)$, allora $u > \gamma$ in un intervallo $[t_* - \delta, t_* + \delta]$, e il sup che definisce t_* sarebbe $\leq t_* - \delta$, assurdo; se fosse $u(t_*) < \gamma(t_*)$, allora $u < \gamma$ in un intervallo $[t_* - \delta, t_* + \delta]$, e il sup sarebbe $\geq t_* + \delta$, assurdo.

Mostriamo che $u > \gamma$ in un intorno destro di t_* . In t_* si ha

$$u(t_*) = \gamma(t_*), \quad u'(t_*) = 0, \quad \gamma'(t_*) < 0.$$

Poiché $u'(t_*) > \gamma'(t_*)$ e u' e γ' sono continue, si ha $u' > \gamma'$ in un certo intervallo $[t_* - \delta, t_* + \delta]$. Quindi per ogni $t \in (t_*, t_* + \delta]$ si ha

$$u(t) = u(t_*) + \int_{t_*}^t u'(s) ds > \gamma(t_*) + \int_{t_*}^t \gamma'(s) ds = \gamma(t),$$

cioè $u > \gamma$ in $(t_*, t_* + \delta]$.

Ora proviamo che $u > \gamma$ in (t_*, ∞) . Sia

$$\xi := \sup\{t > t_* : u > \gamma \text{ in } (t_*, t)\},$$

e supponiamo per assurdo che $\xi < \infty$. Dal fatto che $u > \gamma$ in un certo intervallo $(t_*, t_* + \delta]$ segue che $\xi > t_*$. Dalla definizione di ξ segue che

$$u(t) > \gamma(t) \quad \forall t \in (t_*, \xi), \quad u(\xi) = \gamma(\xi)$$

(come prima, se $u(\xi) > \gamma(\xi)$ o $u(\xi) < \gamma(\xi)$ si arriva a una contraddizione con la definizione di ξ). Per $t \in (t_*, \xi)$ si ha $u(t) - u(\xi) > \gamma(t) - \gamma(\xi)$ e $t - \xi < 0$, da cui

$$\frac{u(t) - u(\xi)}{t - \xi} < \frac{\gamma(t) - \gamma(\xi)}{t - \xi} \quad \forall t \in (t_*, \xi).$$

Passando al limite per $t \rightarrow \xi$ otteniamo

$$u'(\xi) \leq \gamma'(\xi),$$

ma $u'(\xi) = 0$ perché $u(\xi) = \gamma(\xi)$ e il campo vettoriale F è nullo lungo l'iperbole γ , mentre $\gamma'(\xi) < 0$, assurdo. Questo prova che $\xi = \infty$, e dunque $u > \gamma$ in (t_*, ∞) .

Dalla disuguaglianza $u > \gamma = 1/2t$ segue che $u' = 1/u - 2t < 0$ in (t_*, ∞) . Quindi u è decrescente in (t_*, ∞) , e dunque esiste il limite

$$\ell = \lim_{t \rightarrow \infty} u(t).$$

Poiché $u > \gamma > 0$, si ha $\ell \geq 0$. Se $\ell > 0$, allora $1/u(t) \rightarrow 1/\ell$, e $u' = 1/u - 2t \rightarrow -\infty$. Dalla definizione di limite, esiste $t_1 > 0$ tale che $u' \leq -1$ in $[t_1, \infty)$. Dunque

$$u(t) = u(t_1) + \int_{t_1}^t u'(s) ds \leq u(t_1) - (t - t_1) \quad \forall t \in [t_1, \infty),$$

da cui, passando al limite per $t \rightarrow \infty$, si ha $u(t) \rightarrow -\infty$. D'altra parte $u(t) \rightarrow \ell \geq 0$, assurdo. Dunque $\ell = 0$. \square

Vediamo ora due teoremi che, oltre a essere importanti di per sé, sono anche molto utili negli esercizi di analisi qualitativa: il teorema di esistenza globale e quello del confronto.

Teorema 9.2. (Teorema di esistenza globale). *Sia $(\alpha, \beta) \subseteq \mathbb{R}$ un intervallo (sono ammessi anche i casi $\alpha = -\infty$ e/o $\beta = +\infty$). Sia $f : (\alpha, \beta) \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ una funzione*

- (i) *continua,*
- (ii) *localmente lipschitziana in y uniformemente in t ,*
- (iii) *a crescita al più lineare in y : esistono due funzioni $A, B : (\alpha, \beta) \rightarrow \mathbb{R}$, con $A(t) \geq 0, B(t) \geq 0$, entrambe limitate sugli intervalli compatti, tali che*

$$|f(t, y)| \leq A(t)|y| + B(t) \quad \forall t \in (\alpha, \beta), \quad \forall y \in \mathbb{R}^N.$$

Allora le soluzioni dell'equazione differenziale $u' = f(t, u)$ sono globali, cioè definite su tutto l'intervallo (α, β) .

Dimostrazione. Sia u una soluzione dell'equazione $u' = f(t, u)$, e sia (T_-, T_+) il suo intervallo massimale di esistenza, $(T_-, T_+) \subseteq (\alpha, \beta)$. Dobbiamo dimostrare che $(T_-, T_+) = (\alpha, \beta)$, cioè che $T_- = \alpha, T_+ = \beta$.

Supponiamo per assurdo che $T_+ < \beta$. Quindi $T_+ < \infty$ (anche se $\beta = \infty$, comunque T_+ è finito). Fissiamo $T_0 \in (T_-, T_+)$, cosicché $[T_0, T_+] \subset (\alpha, \beta)$. Per ipotesi, $A(t), B(t)$ sono limitate in $[T_0, T_+]$, quindi esistono A_0, B_0 costanti tali che

$$A(t) \leq A_0, \quad B(t) \leq B_0 \quad \forall t \in [T_0, T_+].$$

Fissiamo $c \in (T_0, T_+)$ tale che

$$(T_+ - c)A_0 \leq \frac{1}{2}.$$

Per ogni $p \in (c, T_+)$, stimiamo la soluzione u sull'intervallo $[c, p]$. Ricapitolando, i punti sono in quest'ordine:

$$T_0 < c < p < T_+ < \beta.$$

Indichiamo, in breve,

$$\lambda := \max_{s \in [c, p]} |u(s)|.$$

Per ogni $t \in [c, p]$ si ha

$$\begin{aligned} |u(t)| &= \left| u(c) + \int_c^t u'(s) ds \right| \\ &\leq |u(c)| + \int_c^t |f(s, u(s))| ds \\ &\leq |u(c)| + \int_c^t (A(s)|u(s)| + B(s)) ds \\ &\leq |u(c)| + \int_c^t (A_0\lambda + B_0) ds \\ &= |u(c)| + (t - c)(A_0\lambda + B_0) \\ &\leq |u(c)| + (T_+ - c)A_0\lambda + (T_+ - c)B_0 \\ &\leq \frac{1}{2}\lambda + K, \quad K := |u(c)| + (T_+ - c)B_0. \end{aligned}$$

Dalla disuguaglianza $|u(t)| \leq \lambda/2 + K$, passando al sup su $t \in [c, p]$, otteniamo

$$\lambda \leq \frac{1}{2}\lambda + K,$$

cioè $\lambda/2 \leq K$ (il termine $\lambda/2$ che compare a destra “viene assorbito”),

$$\lambda = \max_{s \in [c, p]} |u(s)| \leq 2K.$$

In particolare, $|u(p)| \leq \lambda \leq 2K$. Abbiamo dimostrato che

$$|u(p)| \leq 2K \quad \forall p \in [c, T_+), \quad (9.2)$$

e K non dipende da p . Quindi per ogni $t \in [c, T_+)$ la coppia $(t, u(t))$ appartiene al compatto $[c, T_+] \times \overline{B(0, 2K)} \subset \Omega$, contro il Teorema 8.2 di fuga dai compatti: assurdo. Questo dimostra che $T_+ = \beta$. Allo stesso modo si prova che $T_- = \alpha$. \square

Altro modo di concludere la dimostrazione. Abbiamo dimostrato (9.2), con K indipendente da p . Sia (t_n) una successione in $[c, T_+)$ che converge a T_+ . Poiché $(t_n, u(t_n))$ è successione limitata, ammette una sottosuccessione $(t_{n_k}, u(t_{n_k}))$ convergente; il suo limite è (T_+, v_0) , per un certo $v_0 \in \mathbb{R}^N$, $|v_0| \leq 2K$. Poiché $(T_+, v_0) \in \Omega = (\alpha, \beta) \times \mathbb{R}^N$, dal criterio di prolungabilità con successione (Proposizione 7.6), u è prolungabile a destra di T_+ , assurdo perché (T_-, T_+) è intervallo massimale. Allo stesso modo si prova che $T_- = \alpha$. \square

Teorema 9.3. (Teorema del confronto). *Sia $\Omega \subseteq \mathbb{R}^2$ un aperto, e siano $f, g : \Omega \rightarrow \mathbb{R}$ due funzioni continue, con g localmente lipschitziana in y uniformemente in t . Sia $I \subseteq \mathbb{R}$ un intervallo, $t_0 \in I$, e siano $u, v : I \rightarrow \mathbb{R}$ due funzioni derivabili, con $(t, u(t)), (t, v(t)) \in \Omega$ per ogni $t \in I$. Sia*

$$u'(t) \leq f(t, u(t)), \quad v'(t) \geq g(t, v(t)) \quad \forall t \in I.$$

Supponiamo inoltre che

$$f(t, u(t)) \leq g(t, u(t)) \quad \forall t \in I.$$

- (i) Se $u(t_0) \leq v(t_0)$, allora $u(t) \leq v(t)$ per ogni $t \geq t_0$, $t \in I$.
(ii) Se $u(t_0) \geq v(t_0)$, allora $u(t) \geq v(t)$ per ogni $t \leq t_0$, $t \in I$.

Il Teorema del confronto verrà dimostrato nella prossima lezione.

Esercizio 9.4. Sia u la soluzione del problema di Cauchy

$$\begin{cases} u' = u^2 + t^2 + e^{ut}, \\ u(0) = 1, \end{cases}$$

e sia (T_-, T_+) il suo intervallo massimale di esistenza. Dimostrare che $T_+ < \infty$.
Suggerimento: confrontare con $\varphi' = \varphi^2$, $\varphi(0) = 1$.

10 Lezione 10

Dimostrazione del Teorema 9.3 del confronto. (i) Indichiamo

$$w(t) := u(t) - v(t).$$

Per ipotesi, $w(t_0) \leq 0$; dobbiamo dimostrare che $w(t) \leq 0$ per ogni $t \geq t_0$, $t \in I$. Supponiamo, per assurdo, che esista $t_1 > t_0$, $t_1 \in I$, tale che $w(t_1) > 0$. Sia

$$E := \{t \in [t_0, t_1] : w(t) \leq 0\}, \quad \xi := \sup E.$$

Notiamo che $E \neq \emptyset$ perché $t_0 \in E$. Inoltre $\xi \in [t_0, t_1]$ perché $E \subseteq [t_0, t_1]$. Si ha poi

$$\xi < t_1.$$

Infatti, dalla definizione di $\xi = \sup E$, esiste una successione (t_n) di elementi di E che converge a ξ ; poiché $w(t_n) \leq 0$ e w è continua, passando al limite si ha $w(\xi) \leq 0$, mentre $w(t_1) > 0$, quindi $\xi \neq t_1$.

Dalla definizione di $\xi = \sup E$ si ha

$$w(t) > 0 \quad \forall t \in (\xi, t_1]. \quad (10.1)$$

Passando al limite per $t \rightarrow \xi^+$, dalla (10.1) si ha $w(\xi) \geq 0$. D'altra parte $w(\xi) \leq 0$, da cui

$$w(\xi) = 0.$$

Quindi u e v coincidono in ξ . Sia

$$c := u(\xi) = v(\xi).$$

Il punto $(\xi, c) = (\xi, u(\xi))$ appartiene a Ω . Quindi, per ipotesi, esistono $a, b > 0$, $L \geq 0$ tali che il rettangolo $[\xi - a, \xi + a] \times [c - b, c + b]$ è contenuto in Ω e

$$|g(t, y) - g(t, z)| \leq L|y - z| \quad \forall t \in [\xi - a, \xi + a], \quad \forall y, z \in [c - b, c + b].$$

Poiché u è continua, esiste $\delta_1 > 0$ tale che

$$|u(t) - u(\xi)| = |u(t) - c| \leq b \quad \forall t \in [\xi - \delta_1, \xi + \delta_1] \cap I,$$

e similmente esiste $\delta_2 > 0$ tale che

$$|v(t) - v(\xi)| = |v(t) - c| \leq b \quad \forall t \in [\xi - \delta_2, \xi + \delta_2] \cap I.$$

Sia

$$\delta_0 := \min\{a, \delta_1, \delta_2, t_1 - \xi\};$$

δ_0 è positivo perché $a, \delta_1, \delta_2, t_1 - \xi$ sono tutti positivi. Dunque $[\xi, \xi + \delta_0] \subseteq [\xi, t_1]$ e

$$u(t), v(t) \in [c - b, c + b] \quad \forall t \in [\xi, \xi + \delta_0].$$

Perciò

$$|g(t, u(t)) - g(t, v(t))| \leq L|u(t) - v(t)| \quad \forall t \in [\xi, \xi + \delta_0].$$

Allora per ogni $t \in [\xi, \xi + \delta_0]$ si ha

$$\begin{aligned} w'(t) &= u'(t) - v'(t) \\ &\leq f(t, u(t)) - g(t, v(t)) \\ &\leq g(t, u(t)) - g(t, v(t)) \\ &\leq |g(t, u(t)) - g(t, v(t))| \\ &\leq L|u(t) - v(t)| = L|w(t)| = Lw(t) \end{aligned}$$

(nell'ultimo passaggio abbiamo usato il fatto che $w \geq 0$ in $[\xi, \xi + \delta_0]$). Dalla disuguaglianza $w' - Lw \leq 0$ in $[\xi, \xi + \delta_0]$ segue che

$$(w(t)e^{-Lt})' = (w'(t) - Lw(t))e^{-Lt} \leq 0 \quad \forall t \in [\xi, \xi + \delta_0].$$

Perciò $w(t)e^{-Lt}$ è decrescente in $[\xi, \xi + \delta_0]$, e

$$w(t)e^{-Lt} \leq w(\xi)e^{-L\xi} = 0 \quad \forall t \in [\xi, \xi + \delta_0].$$

Dunque

$$w(t) \leq 0 \quad \forall t \in [\xi, \xi + \delta_0],$$

contro l'osservazione (vedi (10.1)) che $w > 0$ in $(\xi, t_1]$, assurdo. Questo conclude la dimostrazione di (i).

Il caso (ii) è simile. Rapidamente: definiamo $w := u - v$. L'ipotesi ora è $w(t_0) \geq 0$, la tesi $w(t) \geq 0$ per ogni $t \leq t_0$, $t \in I$. Supponiamo, per assurdo, che esista $t_1 < t_0$, $t_1 \in I$, tale che $w(t_1) < 0$. Sia $E := \{t \in [t_1, t_0] : w(t) \geq 0\}$, e sia $\xi := \inf E$. Si ha $t_1 < \xi$, $w(\xi) = 0$, $w < 0$ in $[t_1, \xi)$. Esiste $\delta_0 > 0$ simile a sopra; si deduce che in $[\xi - \delta_0, \xi]$ si ha $w' = u' - v' \leq \dots \leq L|w| = -Lw$ perché $w \leq 0$ a sinistra di ξ . Quindi $(w(t)e^{Lt})' \leq 0$, $w(t)e^{Lt}$ è decrescente, dunque per ogni $t \in [\xi - \delta_0, \xi]$ si ha $w(t)e^{Lt} \geq w(\xi)e^{L\xi} = 0$. Perciò $w \geq 0$ in $[\xi - \delta_0, \xi]$, assurdo. \square

Osservazione 10.1. (*Da dove viene il fattore e^{-Lt} ?*). Nella dimostrazione appena conclusa la disequazione differenziale $w'(t) \leq Lw(t)$ viene trattata introducendo il fattore e^{-Lt} . A prima vista potrebbe sembrare un'idea tecnicamente efficace ma un po' immotivata; possiamo però ragionare in questo modo: se al posto di " \leq " ci fosse "=", w risolverebbe l'equazione differenziale $w' = Lw$, le cui soluzioni sono $w(t) = ce^{Lt}$, c costante, cioè $w(t)e^{-Lt} = \text{costante}$, cioè $(w(t)e^{-Lt})' = 0$.

Al posto dell'equazione abbiamo la disequazione, ma, imitando il caso dell'equazione, andiamo comunque a vedere come evolve nel tempo il prodotto $w(t)e^{-Lt}$.

Vedremo tra alcune lezioni un'altra dimostrazione del Teorema di esistenza globale, basata sul Lemma di Gronwall; altre dimostrazioni sono per esempio in [5].

Esercizio 10.2. Dimostrare il Teorema di esistenza globale (nel caso scalare $N = 1$) usando il Teorema del confronto.

Vediamo qualche altro esercizio di analisi qualitativa delle soluzioni.

Esercizio 10.3. (*Esercizio 121 di [1]*). Consideriamo l'equazione

$$y'' = e^{-y^2} - \log t, \quad t > 0.$$

Dimostrare che (i) la soluzione è definita in $(0, \infty)$; (ii) $y(t) \rightarrow -\infty$ per $t \rightarrow +\infty$.

Esercizio 10.4. Si consideri l'equazione differenziale

$$u' = f(t, u) = \frac{u^4 + t^2}{u^2 t}, \quad t > 0, \quad u > 0.$$

Studiare l'esistenza globale delle soluzioni.

Esercizio 10.5. (*Soluzioni pari/dispari*). Sia $f : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ continua e localmente lipschitziana in y .

(i) Dimostrare che se $f(-t, y) = -f(t, y)$ per ogni $(t, y) \in \mathbb{R} \times \mathbb{R}^N$, allora per ogni $y_0 \in \mathbb{R}^N$ la soluzione del problema di Cauchy

$$u' = f(t, u), \quad u(0) = y_0$$

è pari, cioè $u(-t) = u(t)$.

(ii) Dimostrare che se $f(-t, -y) = f(t, y)$ per ogni $(t, y) \in \mathbb{R} \times \mathbb{R}^N$, allora la soluzione del problema di Cauchy

$$u' = f(t, u), \quad u(0) = 0$$

è dispari, cioè $u(-t) = -u(t)$.

Esercizio 10.6. Si consideri il problema di Cauchy

$$u' = \frac{u^2}{t^2}, \quad u(t_0) = y_0.$$

Si discuta esistenza globale/blowup per $t \geq t_0$ al variare dei dati iniziali (t_0, y_0) , con $t_0 > 0, y_0 > 0$.

Esercizio 10.7. Come nell'esercizio precedente, ma più in generale: si consideri il problema di Cauchy

$$u' = a(t)u^2, \quad u(t_0) = y_0,$$

e si discuta esistenza globale/blowup per $t \geq t_0$ al variare dei dati iniziali (t_0, y_0) e della funzione (continua) $a(t)$: esiste un criterio generale su $a(t)$ che determina l'esistenza globale o il blowup a tempo finito?

Esercizio 10.8. Studiare l'esistenza globale dell'equazione

$$u' = \frac{u^4 + t^2}{1 + u^2 + t^2}.$$

Esercizio 10.9. Studiare l'esistenza globale delle soluzioni del sistema

$$\begin{cases} u' = -v^3 \\ v' = u + u^5. \end{cases}$$

Suggerimento: mostrare che il sistema ammette un integrale primo $E(u, v)$, e verificare che i suoi insiemi di livello sono compatti di \mathbb{R}^2 .

Nell'esercizio 10.9, l'integrale primo confina ogni soluzione a rimanere su una curva di equazione $E(u, v) = \text{costante}$, che è un insieme compatto, e questo implica l'esistenza globale di ogni soluzione. Questa implicazione vale più in generale:

Proposizione 10.10. (Esistenza globale tramite integrale primo). Sia $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$ continua e localmente lipschitziana. Supponiamo che il sistema autonomo $u' = f(u)$ abbia un integrale primo $E : \mathbb{R}^N \rightarrow \mathbb{R}$, di classe C^1 . Sia $y_0 \in \mathbb{R}^N$. Se l'insieme di livello $\{y \in \mathbb{R}^N : E(y) = E(y_0)\}$ è limitato, allora la soluzione del problema di Cauchy

$$u' = f(u), \quad u(t_0) = y_0$$

è globale, cioè definita per ogni $t \in \mathbb{R}$.

Esercizio 10.11. Dimostrare la Proposizione 10.10 (ragionare come nell'esercizio 10.9).

Esercizio 10.12. (*Sistemi meccanici conservativi*). (Teorema 3.3.2 di [5]) Consideriamo il sistema in \mathbb{R}^{2N}

$$\begin{cases} u' = v, \\ v' = -\nabla P(u), \end{cases} \quad (10.2)$$

dove $P : \mathbb{R}^N \rightarrow \mathbb{R}$ è una funzione scalare di classe C^1 . Sia $E : \mathbb{R}^{2N} \rightarrow \mathbb{R}$ la funzione

$$E(u, v) = \frac{1}{2}|v|^2 + P(u). \quad (10.3)$$

(i) Verificare che E è integrale primo del sistema.

(ii) Dimostrare che, se esiste $c_0 \in \mathbb{R}$ tale che $P(x) \geq c_0$ per ogni $x \in \mathbb{R}^N$, allora tutte le soluzioni del sistema sono soluzioni globali.

I sistemi del tipo (10.2)-(10.3) si dicono *sistemi meccanici conservativi*: $u(t)$ è la posizione, $v(t)$ la velocità, $P(u)$ è l'energia potenziale, e $\frac{1}{2}|v|^2$ l'energia cinetica.

Il prossimo teorema è un risultato di Analisi 1 utile negli esercizi di studio qualitativo di equazioni differenziali.

Teorema 10.13. (Teorema dell'asintoto). (i) Sia $a \in \mathbb{R}$, sia $u : (a, \infty) \rightarrow \mathbb{R}$ una funzione derivabile, e supponiamo che esistano i limiti

$$\ell = \lim_{t \rightarrow +\infty} u(t), \quad m = \lim_{t \rightarrow +\infty} u'(t).$$

Se ℓ è finito, allora $m = 0$.

(ii) Siano $a, b \in \mathbb{R}$, con $a < b$, sia $u : (a, b) \rightarrow \mathbb{R}$ una funzione derivabile, e supponiamo che esistano i limiti

$$\ell = \lim_{t \rightarrow b} u(t), \quad m = \lim_{t \rightarrow b} u'(t).$$

Se $\ell = +\infty$, allora $m = +\infty$. Se $\ell = -\infty$, allora $m = -\infty$.

(iii) Sia $b \in \mathbb{R}$, sia $u : (-\infty, b) \rightarrow \mathbb{R}$ una funzione derivabile, e supponiamo che esistano i limiti

$$\ell = \lim_{t \rightarrow -\infty} u(t), \quad m = \lim_{t \rightarrow -\infty} u'(t).$$

Se ℓ è finito, allora $m = 0$.

(iv) Siano $a, b \in \mathbb{R}$, con $a < b$, sia $u : (a, b) \rightarrow \mathbb{R}$ una funzione derivabile, e supponiamo che esistano i limiti

$$\ell = \lim_{t \rightarrow a} u(t), \quad m = \lim_{t \rightarrow a} u'(t).$$

Se $\ell = +\infty$, allora $m = -\infty$. Se $\ell = -\infty$, allora $m = +\infty$.

Esercizio 10.14. Dimostrare il Teorema dell'asintoto (usare il Teorema di Lagrange di Analisi 1).

Osserviamo che nel Teorema dell'asintoto l'ipotesi che esista il limite di $u'(t)$ non può essere rimossa.

Vediamo ora il Lemma di Gronwall, un risultato simile al Teorema del confronto, basato anch'esso sul confronto con le equazioni lineari del primo ordine non omogenee. L'enunciato standard si potrebbe limitare al punto 2) del lemma seguente, ma anche le altre versioni (in particolare il punto 7)) sono spesso utili sia negli esercizi che nelle dimostrazioni di altri teoremi; quindi, al prezzo di allungare l'enunciato, consideriamo anche alcune sue varianti.

Lemma 10.15. (Lemma di Gronwall). *Sia $I \subseteq \mathbb{R}$ un intervallo, $t_0 \in I$. Indichiamo*

$$I_+ = \{t \in I : t \geq t_0\}, \quad I_- = \{t \in I : t \leq t_0\}.$$

1) *Sia $\varphi : I_+ \rightarrow \mathbb{R}$ una funzione derivabile, siano $a, b : I_+ \rightarrow \mathbb{R}$ due funzioni continue, e sia*

$$\varphi'(t) \leq a(t)\varphi(t) + b(t) \quad \forall t \in I_+.$$

Allora

$$\varphi(t) \leq \exp\left(\int_{t_0}^t a(s) ds\right) \left[\varphi(t_0) + \int_{t_0}^t \exp\left(-\int_{t_0}^s a(\sigma) d\sigma\right) b(s) ds\right] \quad \forall t \in I_+.$$

2) (Versione standard). *Sia $\varphi : I_+ \rightarrow \mathbb{R}$ derivabile, siano $a, b : I_+ \rightarrow \mathbb{R}$ continue, e*

$$a(t) \geq 0, \quad b(t) \geq 0, \quad \varphi'(t) \leq a(t)\varphi(t) + b(t) \quad \forall t \in I_+.$$

Allora

$$\varphi(t) \leq \exp\left(\int_{t_0}^t a(s) ds\right) \left(\varphi(t_0) + \int_{t_0}^t b(s) ds\right) \quad \forall t \in I_+.$$

3) *Sia $\varphi : I_- \rightarrow \mathbb{R}$ derivabile, siano $\alpha, \beta : I_- \rightarrow \mathbb{R}$ continue, e*

$$\varphi'(t) \geq \alpha(t)\varphi(t) + \beta(t) \quad \forall t \in I_-.$$

Allora

$$\varphi(t) \leq \exp\left(\int_{t_0}^t \alpha(s) ds\right) \left[\varphi(t_0) + \int_{t_0}^t \exp\left(-\int_{t_0}^s \alpha(\sigma) d\sigma\right) \beta(s) ds\right] \quad \forall t \in I_-.$$

4) *Sia $\varphi : I_- \rightarrow \mathbb{R}$ derivabile, siano $\alpha, \beta : I_- \rightarrow \mathbb{R}$ continue, e*

$$\alpha(t) \leq 0, \quad \beta(t) \leq 0, \quad \varphi'(t) \geq \alpha(t)\varphi(t) + \beta(t) \quad \forall t \in I_-.$$

Allora

$$\varphi(t) \leq \exp\left(\int_{t_0}^t \alpha(s) ds\right) \left(\varphi(t_0) + \int_{t_0}^t \beta(s) ds\right) \quad \forall t \in I_-.$$

5) *Sia $\varphi : I \rightarrow \mathbb{R}$ derivabile, siano $A, B : I \rightarrow \mathbb{R}$ continue, e*

$$A(t) \geq 0, \quad B(t) \geq 0, \quad |\varphi'(t)| \leq A(t)\varphi(t) + B(t) \quad \forall t \in I.$$

Allora

$$\varphi(t) \leq \exp\left(\int_{I(t_0,t)} A(s) ds\right) \left(\varphi(t_0) + \int_{I(t_0,t)} B(s) ds\right) \quad \forall t \in I$$

dove $I(t_0, t)$ indica l'intervallo di estremi t_0 e t

$$I(t_0, t) := [\min\{t_0, t\}, \max\{t_0, t\}] = \begin{cases} [t_0, t] & \text{se } t \geq t_0, \\ [t, t_0] & \text{se } t \leq t_0. \end{cases}$$

6) Sia $\varphi : I \rightarrow \mathbb{R}$ derivabile, siano A, B costanti, e

$$A \geq 0, \quad B \geq 0, \quad |\varphi'(t)| \leq A\varphi(t) + B \quad \forall t \in I.$$

Allora

$$\varphi(t) \leq e^{A|t-t_0|}(\varphi(t_0) + B|t-t_0|) \quad \forall t \in I.$$

7) (Caso vettoriale). Sia $u : I \rightarrow \mathbb{R}^N$ una funzione (vettoriale) derivabile, siano $A, B : I \rightarrow \mathbb{R}$ due funzioni continue, e sia

$$A(t) \geq 0, \quad B(t) \geq 0, \quad |u'(t)| \leq A(t)|u(t)| + B(t) \quad \forall t \in I.$$

Allora

$$|u(t)| \leq \exp\left(\int_{I(t_0,t)} A(s) ds\right) \left(|u(t_0)| + \int_{I(t_0,t)} B(s) ds\right) \quad \forall t \in I.$$

8) Sia $u : I \rightarrow \mathbb{R}^N$ funzione (vettoriale) derivabile, siano A, B costanti, e sia

$$A \geq 0, \quad B \geq 0, \quad |u'(t)| \leq A|u(t)| + B \quad \forall t \in I.$$

Allora

$$|u(t)| \leq e^{A|t-t_0|}(|u(t_0)| + B|t-t_0|) \quad \forall t \in I.$$

Dimostrazione. Dimostrazione di 1) Sia $g(t, y) := a(t)y + b(t)$, cosicché $\varphi' \leq g(t, \varphi)$ in I_+ . Sia ψ la soluzione del problema di Cauchy

$$\psi' = g(t, \psi), \quad \psi(t_0) = \varphi(t_0).$$

Dal Teorema del confronto si ha allora $\varphi \leq \psi$ in I_+ (applicare il Teorema 9.3, punto (i), con $u = \varphi$, $v = \psi$, $f = g$). Inoltre, ricordando la formula (4.7) delle soluzioni delle equazioni lineari, non omogenee, del primo ordine, si ha

$$\psi(t) = \exp\left(\int_{t_0}^t a(s) ds\right) \left[\varphi(t_0) + \int_{t_0}^t \exp\left(-\int_{t_0}^s a(\sigma) d\sigma\right) b(s) ds\right].$$

Dimostrazione di 2) Poiché $a \geq 0$ in I_+ , per ogni $s \in I_+$ si ha

$$\int_{t_0}^s a(\sigma) d\sigma \geq 0,$$

quindi

$$\exp\left(-\int_{t_0}^s a(\sigma) d\sigma\right) \leq 1.$$

Poiché $b \geq 0$ in I_+ , si ha

$$\exp\left(-\int_{t_0}^s a(\sigma) d\sigma\right)b(s) \leq b(s).$$

Da 1) segue dunque 2).

Dimostrazione di 3) Sia $g(t, y) := \alpha(t)y + \beta(t)$, cosicché $\varphi' \geq g(t, \varphi)$ in I_- . Sia ψ la soluzione di $\psi' = g(t, \psi)$ con dato iniziale $\psi(t_0) = \varphi(t_0)$. Dal Teorema del confronto si ha allora $\varphi \leq \psi$ in I_- (applicare il Teorema 9.3, punto (ii), con $u = \psi$, $v = \varphi$, $f = g$). Infine, ricordando (4.7),

$$\psi(t) = \exp\left(\int_{t_0}^t \alpha(s) ds\right)\left[\varphi(t_0) + \int_{t_0}^t \exp\left(-\int_{t_0}^s \alpha(\sigma) d\sigma\right)\beta(s) ds\right].$$

Dimostrazione di 4) Siano $t, s \in I_-$, con $t \leq s \leq t_0$. Allora

$$-\int_{t_0}^s \alpha(\sigma) d\sigma = \int_s^{t_0} \alpha(\sigma) d\sigma \leq 0$$

perché $\alpha \leq 0$. Quindi

$$0 < \exp\left(-\int_{t_0}^s \alpha(\sigma) d\sigma\right) \leq 1.$$

Moltiplicando per $(-\beta)$, che è ≥ 0 , si ha

$$0 \leq -\beta(s) \exp\left(-\int_{t_0}^s \alpha(\sigma) d\sigma\right) \leq -\beta(s).$$

Dunque

$$\begin{aligned} \int_{t_0}^t \exp\left(-\int_{t_0}^s \alpha(\sigma) d\sigma\right)\beta(s) ds &= \int_t^{t_0} -\beta(s) \exp\left(-\int_{t_0}^s \alpha(\sigma) d\sigma\right) ds \\ &\leq \int_t^{t_0} -\beta(s) ds = \int_{t_0}^t \beta(s) ds. \end{aligned}$$

Dalla 3) segue dunque 4).

Dimostrazione di 5) L'ipotesi $|\varphi'| \leq A(t)\varphi + B(t)$ implica che $A(t)\varphi + B(t) \geq 0$ e che

$$-A(t)\varphi - B(t) \leq \varphi' \leq A(t)\varphi + B(t) \quad \forall t \in I.$$

Il punto 2) con $a = A$, $b = B$ dà

$$\varphi(t) \leq \exp\left(\int_{t_0}^t A(s) ds\right)\left(\varphi(t_0) + \int_{t_0}^t B(s) ds\right) \quad \forall t \in I_+,$$

che è la tesi per $t \in I_+$. Il punto 4) con $\alpha = -A$, $\beta = -B$ dà

$$\begin{aligned} \varphi(t) &\leq \exp\left(\int_{t_0}^t -A(s) ds\right)\left(\varphi(t_0) + \int_{t_0}^t -B(s) ds\right) \\ &= \exp\left(\int_t^{t_0} A(s) ds\right)\left(\varphi(t_0) + \int_t^{t_0} B(s) ds\right) \quad \forall t \in I_-, \end{aligned}$$

che è la tesi per $t \in I_-$.

Dimostrazione di 6) È un caso particolare della 5).

Dimostrazione di 7) Sia $\varepsilon > 0$, e definiamo

$$\varphi(t) := \sqrt{\varepsilon + |u(t)|^2}, \quad t \in I.$$

Si ha allora

$$|u(t)| < \varphi(t), \quad \varphi(t) \geq \sqrt{\varepsilon} > 0 \quad \forall t \in I$$

e

$$\begin{aligned} \varphi'(t) &= \frac{d}{dt} \sqrt{\varepsilon + u_1^2(t) + \dots + u_N^2(t)} \\ &= \frac{1}{2\sqrt{\varepsilon + |u(t)|^2}} \left(2u_1(t)u_1'(t) + \dots + 2u_N(t)u_N'(t) \right) = \frac{u(t) \cdot u'(t)}{\varphi(t)}. \end{aligned}$$

Dalla disuguaglianza di Cauchy-Schwarz, ricordando che $|u(t)| \leq \varphi(t)$, si ha

$$\begin{aligned} |\varphi'(t)| &= \frac{|u(t) \cdot u'(t)|}{\varphi(t)} \leq \frac{|u(t)||u'(t)|}{\varphi(t)} \leq |u'(t)| \leq A(t)|u(t)| + B(t) \\ &\leq A(t)\varphi(t) + B(t). \end{aligned}$$

Dalla 5) si ha allora

$$\varphi(t) \leq \exp \left(\int_{I(t_0,t)} A(s) ds \right) \left(\varphi(t_0) + \int_{I(t_0,t)} B(s) ds \right) \quad \forall t \in I.$$

Poiché $|u(t)| \leq \varphi(t)$, si ha

$$|u(t)| \leq \exp \left(\int_{I(t_0,t)} A(s) ds \right) \left(\sqrt{\varepsilon + |u(t_0)|^2} + \int_{I(t_0,t)} B(s) ds \right) \quad \forall t \in I.$$

Questa disuguaglianza vale per ogni $\varepsilon > 0$; passando all'inf su $\varepsilon > 0$ otteniamo la tesi.

Dimostrazione di 8) È un caso particolare di 7). □

Esercizio 10.16. Dimostrare il Teorema 9.2 di esistenza globale usando il Lemma di Gronwall.

Svolgimento di alcuni esercizi nella prossima pagina.

Svolgimento dell'Esercizio 10.3. (i) È un'equazione scalare del secondo ordine. Per vedere se possiamo applicare il Teorema di esistenza globale, scriviamola come sistema del primo ordine:

$$\begin{cases} y' = v, \\ v' = e^{-y^2} - \log t, \end{cases}$$

cioè

$$u' = f(t, u), \quad u = \begin{pmatrix} y \\ v \end{pmatrix}, \quad f(t, u) = \begin{pmatrix} v \\ e^{-y^2} - \log t \end{pmatrix}.$$

Il campo $f : (0, \infty) \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ è di classe C^∞ . Usando la disuguaglianza elementare

$$\sqrt{a^2 + b^2} \leq |a| + |b| \quad \forall a, b \in \mathbb{R},$$

stimiamo

$$\begin{aligned} |f(t, u)| &= \sqrt{v^2 + (e^{-y^2} - \log t)^2} \leq |v| + |e^{-y^2} - \log t| \\ &\leq |v| + 1 + |\log t| \leq A(t)|u| + B(t) \end{aligned}$$

con $A(t) = 1$ e $B(t) = 1 + |\log t|$. Sono soddisfatte le ipotesi del Teorema di esistenza globale, quindi la soluzione $u = (y, v)$ è definita su tutto $(0, \infty)$.

(ii) Si ha $e^{-y^2} - \log t \leq 1 - \log t \rightarrow -\infty$ per $t \rightarrow +\infty$, quindi $e^{-y^2} - \log t$ è definitivamente minore di qualunque numero $m < 0$; per esempio, scegliendo $m = -1$, si ha

$$e^{-y^2} - \log t \leq 1 - \log t \leq -1 \quad \forall t \geq e^2.$$

Dunque la soluzione $y(t)$ soddisfa

$$y''(t) = e^{-y^2(t)} - \log t \leq -1 \quad \forall t \geq e^2,$$

da cui

$$y'(t) - y'(e^2) = \int_{e^2}^t y''(s) ds \leq \int_{e^2}^t -1 ds = -(t - e^2) \quad \forall t \geq e^2,$$

cioè

$$y'(t) \leq -t + c \quad \forall t \geq e^2, \quad c := y'(e^2) + e^2,$$

e c è indipendente da t . Di conseguenza

$$y(t) - y(e^2) = \int_{e^2}^t y'(s) ds \leq \int_{e^2}^t (-s + c) ds = -\frac{t^2}{2} + \frac{e^4}{2} + c(t - e^2),$$

cioè

$$y(t) \leq -\frac{t^2}{2} + ct + c_0 \quad \forall t \geq e^2,$$

con $c_0 := y(e^2) - ce^2 + e^4/2$. Questo prova che $y(t) \rightarrow -\infty$ per $t \rightarrow \infty$. \square

Svolgimento dell'Esercizio 10.4. Il campo f soddisfa

$$f(t, y) = \frac{y^4 + t^2}{y^2 t} > \frac{y^4}{y^2 t} = \frac{y^2}{t} =: g(t, y).$$

Vogliamo usare il Teorema del confronto: dato $(t_0, y_0) \in \mathbb{R}^2$, con t_0, y_0 positivi, siano u e φ le soluzioni dei problemi di Cauchy $u' = f(t, u)$ e $\varphi' = g(t, \varphi)$ con lo stesso dato iniziale (t_0, y_0) , cioè

$$\begin{cases} u' = f(t, u) = \frac{u^4 + t^2}{u^2 t}, \\ u(t_0) = y_0, \end{cases} \quad \begin{cases} \varphi' = g(t, \varphi) = \frac{\varphi^2}{t}, \\ \varphi(t_0) = y_0. \end{cases}$$

L'equazione $\varphi' = \varphi^2/t$ è a variabili separabili. Riscriviamola come $\varphi'/\varphi^2 = 1/t$, cioè

$$(-1/\varphi)' = (\log t)',$$

cioè $(1/\varphi + \log t)' = 0$. Quindi $1/\varphi + \log t = \text{costante}$, da cui

$$\frac{1}{\varphi(t)} + \log t = \frac{1}{\varphi(t_0)} + \log t_0,$$

cioè, ricordando che $\varphi(t_0) = y_0$,

$$\varphi(t) = \frac{1}{1/y_0 - \log(t/t_0)}, \quad \forall t \in [t_0, c), \quad c := t_0 \exp(1/y_0).$$

Inoltre φ “esplode” al tempo c (*blowup*), cioè

$$\varphi(t) \rightarrow +\infty \quad (t \rightarrow c).$$

Dal Teorema del confronto (Teorema 9.3), poiché $f(t, y) > g(t, y)$, su ogni intervallo $[t_0, t_1]$ in cui u e φ sono entrambe definite si ha

$$u(t) \geq \varphi(t) \quad \forall t \in [t_0, t_1].$$

Questo implica che anche u esplode a tempo finito, e ci dà anche una stima del tempo massimo di esistenza di u : si ha

$$u(t) \rightarrow +\infty \quad (t \rightarrow \beta),$$

per un certo tempo massimale β nell'intervallo $t_0 < \beta \leq c = t_0 \exp(1/y_0)$.

Infatti, supponiamo per assurdo che u sia definita anche a destra di c , cioè $\beta > c$. Allora u , essendo una funzione continua, è limitata nell'intervallo compatto $[t_0, c]$, cioè esiste $K > 0$ tale che

$$|u(t)| \leq K \quad \forall t \in [t_0, c]. \quad (10.4)$$

D'altra parte, per ogni $t_1 \in (t_0, c)$, sia u che φ sono definite nell'intervallo $[t_0, t_1]$, e quindi $u(t) \geq \varphi(t)$ per ogni $t \in [t_0, t_1]$. In particolare, in $t = t_1$ si ha

$$u(t_1) \geq \varphi(t_1) \quad \forall t_1 \in (t_0, c).$$

Passando al limite per $t_1 \rightarrow c^-$ si ottiene $u(t_1) \rightarrow +\infty$ per $t_1 \rightarrow c$, contro (10.4), assurdo. \square

Svolgimento dell'esercizio 10.5. (i) Sia $v(t) := u(-t)$. Dall'ipotesi $-f(-t, y) = f(t, y)$, usando anche l'identità $u'(t) = f(t, u(t))$, calcoliamo

$$v'(t) = -u'(-t) = -f(-t, u(-t)) = -f(-t, v(t)) = f(t, v(t))$$

e $v(0) = u(0) = y_0$. Dunque u e v soddisfano

$$\begin{cases} u' = f(t, u), \\ u(0) = y_0, \end{cases} \quad \begin{cases} v' = f(t, v), \\ v(0) = y_0, \end{cases}$$

sono entrambe soluzione dello stesso problema di Cauchy, e quindi, dove sono definite entrambe, coincidono (vedi Corollario 3.15).

(ii) Sia $v(t) := -u(-t)$. Usando l'ipotesi $f(-t, -y) = f(t, y)$ e $u'(t) = f(t, u(t))$, calcoliamo

$$v'(t) = u'(-t) = f(-t, u(-t)) = f(-t, -v(t)) = f(t, v(t))$$

e $v(0) = -u(0) = 0$. Dunque u e v sono entrambe soluzione dello stesso problema di Cauchy, e quindi coincidono. \square

Svolgimento dell'Esercizio 10.9. Il campo vettoriale $f(u, v) = (-v^3, u + u^5)$ non soddisfa l'ipotesi di crescita al più lineare $|f(u, v)| \leq A|(u, v)| + B$ del Teorema di esistenza globale. D'altra parte, il sistema ha un integrale primo: sia

$$E(u, v) := \frac{u^2}{2} + \frac{u^6}{6} + \frac{v^4}{4}.$$

Se $(u(t), v(t))$ è soluzione del sistema in un certo intervallo I , allora

$$\begin{aligned} \frac{d}{dt} \{E(u(t), v(t))\} &= uu' + u^5 u' + v^3 v' \\ &= (-v^3)(u + u^5) + v^3(u + u^5) = 0. \end{aligned}$$

Quindi E rimane costante lungo le soluzioni. Dunque ogni soluzione $(u(t), v(t))$ rimane confinata in un unico insieme di livello di E per tutta la sua esistenza.

Per ogni $\lambda \geq 0$, l'insieme di livello

$$\mathcal{E}_\lambda := \{(x, y) \in \mathbb{R}^2 : E(x, y) = \lambda\}$$

è compatto: \mathcal{E}_λ è chiuso perché è l'anti-immagine di $\{\lambda\}$ tramite E , che è funzione continua; \mathcal{E}_λ è limitato perché ogni suo punto $(x, y) \in \mathcal{E}_\lambda$ soddisfa

$$\frac{x^2}{2} \leq \frac{x^2}{2} + \frac{x^6}{6} + \frac{y^4}{4} = E(x, y) = \lambda,$$

da cui $|x| \leq \sqrt{2\lambda}$, e

$$\frac{y^4}{4} \leq \frac{x^2}{2} + \frac{x^6}{6} + \frac{y^4}{4} = E(x, y) = \lambda,$$

da cui $|y| \leq (4\lambda)^{1/4}$.

Dal Corollario 8.4 (compatto nello spazio implica globale nel tempo), ogni soluzione $(u(t), v(t))$ è globale nel tempo. \square

Svolgimento dell'Esercizio 10.11. *Dimostrazione della Proposizione 10.10.* Poiché $E(u(t)) = E(u(t_0)) = E(y_0)$ per ogni t , la soluzione $u(t)$ rimane nell'insieme di livello $\{y : E(y) = E(y_0)\}$ per ogni t nel suo intervallo massimale. Essendo chiuso e limitato, l'insieme di livello è compatto. Quindi la soluzione è globale per il Corollario 8.4 della fuga dai compatti. \square

Svolgimento dell'Esercizio 10.12. (i) Lungo ogni soluzione $(u(t), v(t))$ dell'equazione, si ha

$$\begin{aligned} \frac{d}{dt}\{E(u(t), v(t))\} &= \frac{d}{dt}\left(\frac{1}{2}(v_1^2 + \dots + v_N^2) + P(u_1, \dots, u_N)\right) \\ &= v \cdot v' + \nabla P(u) \cdot u' \\ &= v \cdot (-\nabla P(u)) + \nabla P(u) \cdot v \\ &= 0. \end{aligned}$$

Quindi l'energia (10.3) è un integrale primo del sistema.

(ii) Sia (u, v) soluzione di (10.2), e sia $I = (T_-, T_+)$ il suo intervallo massimale di esistenza. Sia $t_0 \in I$ un istante fissato, e sia $(u_0, v_0) := (u(t_0), v(t_0))$ il punto in cui si trova la soluzione all'istante t_0 . Poiché E è costante lungo la soluzione, per ogni $t \in I$ si ha

$$E(u(t), v(t)) = \frac{1}{2}|v(t)|^2 + P(u(t)) = E(u_0, v_0) = \frac{1}{2}|v_0|^2 + P(u_0) =: \lambda.$$

Inoltre, poiché $P \geq c_0$, si ha

$$\lambda = \frac{1}{2}|v(t)|^2 + P(u(t)) \geq \frac{1}{2}|v(t)|^2 + c_0 \geq c_0,$$

da cui $\lambda \geq c_0$ e, anzi, $\lambda \geq \frac{1}{2}|v(t)|^2 + c_0$, cioè

$$|v(t)|^2 \leq 2(\lambda - c_0).$$

Quindi $v(t)$ è limitata,

$$|v(t)| \leq M := \sqrt{2(\lambda - c_0)} \quad \forall t \in I.$$

Di conseguenza per ogni $t \in I$ si ha

$$|u(t)| = \left|u(t_0) + \int_{t_0}^t u'(s) ds\right| \leq |u_0| + \left|\int_{t_0}^t v(s) ds\right| \leq |u_0| + M|t - t_0|.$$

Se T_+ fosse finito, avremmo

$$|v(t)| \leq M, \quad |u(t)| \leq |u_0| + M(T_+ - t_0) =: C \quad \forall t \in [t_0, T_+). \quad (10.5)$$

Quindi esisterebbe una successione $t_n \rightarrow T_+$ con $(t_n, u(t_n), v(t_n))$ convergente ad un certo limite finito (T_+, u_1, v_1) , e allora, dal criterio di prolungabilità per successione (Proposizione 7.6), la soluzione sarebbe prolungabile a destra di T_+ , assurdo.

Altro modo di concludere: se T_+ fosse finito, avremmo (10.5), e $(t, u(t), v(t))$ starebbe nel compatto $[t_0, T_+] \times \{|u| \leq C\} \times \{|v| \leq M\}$ per tutti i tempi $t \in [t_0, T_+)$, contro il Teorema di fuga dai compatti (Teorema 8.2), assurdo.

Dunque $T_+ = +\infty$. In modo analogo si prova che $T_- = -\infty$. \square

Svolgimento dell'Esercizio 10.14. *Dimostrazione del Teorema 10.13 (Teorema dell'asintoto).* (i) Sia $\ell \in \mathbb{R}$. Per ogni $n \in \mathbb{N}$, dal Teorema di Lagrange esiste $t_n \in (n, n+1)$ tale che $u(n+1) - u(n) = u'(t_n)$. Quindi

$$\lim_{n \rightarrow +\infty} u'(t_n) = \lim_{n \rightarrow +\infty} (u(n+1) - u(n)) = \ell - \ell = 0.$$

D'altra parte $u'(t) \rightarrow m$ per $t \rightarrow +\infty$, quindi $m = 0$.

(ii) Sia $\ell = +\infty$, e supponiamo che $\lim_{(t \rightarrow b)} u'(t) = m \in \mathbb{R}$. Dalla definizione di limite, esiste $t_0 \in (a, b)$ tale che $|u'(t) - m| \leq 1$ per ogni $t \in [t_0, b)$. Quindi si ha $|u'(t)| \leq |m| + 1$ per ogni $t \in [t_0, b)$.

Sia ora $t \in (t_0, b)$. Dal Teorema di Lagrange, $u(t) - u(t_0) = u'(s)(t - t_0)$ per un certo $s \in (t_0, t)$, quindi

$$|u(t)| = |u(t_0) + u'(s)(t - t_0)| \leq |u(t_0)| + (|m| + 1)(b - t_0) =: C.$$

Questo prova che $u(t) \leq C$ per ogni $t \in (t_0, b)$, con C indipendente da t . Quindi $\ell \leq C$, assurdo.

Se $m = -\infty$, dalla definizione di limite esiste $t_0 \in (a, b)$ tale che $u'(t) \leq 0$ per ogni $t \in [t_0, b)$. Quindi $u(t) \leq u(t_0)$ per ogni $t \in (t_0, b)$, e dunque $\ell \leq u(t_0)$, assurdo. Questo conclude la dimostrazione del fatto che se $\ell = +\infty$ allora $m = +\infty$.

Tutti gli altri casi dell'enunciato si dimostrano in modo simile.

[Anche così: ponendo $u(t) = -v(t)$ si deduce l'implicazione $\ell = -\infty \Rightarrow m = -\infty$ del caso (ii) dall'implicazione appena provata; ponendo $u(t) = v(-t)$ si deduce (iii) da (i); ponendo $u(t) = v(a + b - t)$ si deduce (iv) da (ii)]. \square

Svolgimento dell'Esercizio 10.16. *Dimostrazione del Teorema 9.2 di esistenza globale tramite Lemma di Gronwall.* Assumiamo le ipotesi del Teorema 9.2. Supponiamo $T_+ < \beta$ (quindi, in particolare, T_+ è finito). Fissiamo un qualunque $t_0 \in (T_-, T_+)$. Si ha dunque $[t_0, T_+] \subset (\alpha, \beta)$. Per ipotesi, esistono costanti $A_0, B_0 \geq 0$ tali che

$$A(t) \leq A_0, \quad B(t) \leq B_0 \quad \forall t \in [t_0, T_+].$$

Allora la soluzione u soddisfa

$$|u'| = |f(t, u)| \leq A(t)|u| + B(t) \leq A_0|u| + B_0 \quad \forall t \in [t_0, T_+].$$

Per Gronwall (Lemma 10.15, punto 8)), per ogni $t \in [t_0, T_+)$ si ha

$$\begin{aligned} |u(t)| &\leq e^{A_0(t-t_0)} \left(|u(t_0)| + B_0(t - t_0) \right) \\ &\leq e^{A_0(T_+ - t_0)} \left(|u(t_0)| + B_0(T_+ - t_0) \right) =: K, \end{aligned} \quad (10.6)$$

e K è una costante indipendente da t . Dunque per ogni $t \in [t_0, T_+)$ la coppia $(t, u(t))$ rimane nel compatto $[t_0, T_+] \times \{|u| \leq K\} \subset \Omega = (\alpha, \beta) \times \mathbb{R}^N$, contro il Teorema 8.2 di fuga dai compatti, assurdo. Perciò $T_+ = \beta$.

In modo simile si prova che $T_- = \alpha$. \square

11 Lezione 11

Della disuguaglianza di Gronwall esiste anche una versione integrale (con le sue varianti, che non riportiamo).

Lemma 11.1. (Lemma di Gronwall, versione integrale). *Sia $I \subset \mathbb{R}$ intervallo, $t_0 \in I$. Sia $w : I \rightarrow \mathbb{R}$ integrabile, e sia*

$$w(t) \leq A \int_{t_0}^t w(s) ds + B \quad \forall t \in I, t \geq t_0,$$

con A, B costanti ≥ 0 . Allora

$$w(t) \leq B(1 + Ae^{A(t-t_0)}(t-t_0)) \quad \forall t \geq t_0.$$

Dimostrazione. Definiamo

$$\varphi(t) := \int_{t_0}^t w(s) ds,$$

cosicché l'ipotesi si riscrive $w(t) \leq A\varphi(t) + B$. Si ha $\varphi(t_0) = 0$ e

$$\varphi'(t) = w(t) \leq A\varphi(t) + B.$$

Quindi, dal Lemma 10.15,

$$\varphi(t) \leq e^{A(t-t_0)}B(t-t_0),$$

e dunque

$$\begin{aligned} w(t) &\leq A\varphi(t) + B \\ &\leq Ae^{A(t-t_0)}B(t-t_0) + B. \end{aligned} \quad \square$$

Teorema 11.2. (Dipendenza continua dai dati iniziali). *Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ una funzione continua e localmente lipschitziana in y , uniformemente in t . Sia $(t_0, y_0) \in \Omega$, e sia $u : I \rightarrow \mathbb{R}^N$ la soluzione del problema di Cauchy*

$$\begin{cases} u' = f(t, u), \\ u(t_0) = y_0, \end{cases} \quad (11.1)$$

definita nel suo intervallo massimale I .

Allora per ogni intervallo compatto $[\alpha, \beta] \subset I$, con $t_0 \in [\alpha, \beta]$, esistono $\delta_0 > 0$, $L \geq 0$, $M \geq 0$ tali che per ogni coppia $(t_1, y_1) \in \Omega$ che soddisfa

$$|t_1 - t_0| \leq \delta_0, \quad |y_1 - y_0| \leq \delta_0, \quad t_1 \in [\alpha, \beta], \quad (11.2)$$

la soluzione v del problema di Cauchy

$$\begin{cases} v' = f(t, v), \\ v(t_1) = y_1 \end{cases} \quad (11.3)$$

è definita su $[\alpha, \beta]$ e soddisfa

$$|v(t) - u(t)| \leq e^{L|t-t_0|}(|y_1 - y_0| + M|t_1 - t_0|) \quad \forall t \in [\alpha, \beta]. \quad (11.4)$$

Di conseguenza v converge a u uniformemente in $[\alpha, \beta]$ per $(t_1, y_1) \rightarrow (t_0, y_0)$.

Dimostrazione. Per definizione di soluzione di un'equazione differenziale, e poiché $[\alpha, \beta] \subset I$, si ha

$$(t, u(t)) \in \Omega \quad \forall t \in [\alpha, \beta]. \quad (11.5)$$

Per $r > 0$, $t \in [\alpha, \beta]$, consideriamo la palla di centro $u(t)$ e raggio r , e costruiamo l'intorno "tubolare" del grafico di u

$$K := \{(t, y) \in \mathbb{R} \times \mathbb{R}^N : t \in [\alpha, \beta], |y - u(t)| \leq r\}.$$

Dimostriamo che esiste $r > 0$ tale che $K \subset \Omega$. Supponiamo, per assurdo, che un tale r non esista. Allora per ogni $r > 0$ esiste un punto di K che non appartiene a Ω , cioè una coppia (t, y) , con $t \in [\alpha, \beta]$, $|y - u(t)| \leq r$, tale che $(t, y) \notin \Omega$. In particolare, per ogni intero positivo n esiste una coppia $(t_n, y_n) \in \mathbb{R} \times \mathbb{R}^N$ tale che

$$t_n \in [\alpha, \beta], \quad |y_n - u(t_n)| \leq \frac{1}{n}, \quad (t_n, y_n) \notin \Omega.$$

La successione (t_n) è limitata, quindi ammette una sottosuccessione (t_{n_k}) convergente, diciamo $t_{n_k} \rightarrow p$, con $p \in [\alpha, \beta]$. Essendo u continua, si ha $u(t_{n_k}) \rightarrow u(p)$. Anche y_{n_k} converge a $u(p)$ perché $|y_{n_k} - u(t_{n_k})| \leq 1/n_k \rightarrow 0$.

Per ogni k , $(t_{n_k}, y_{n_k}) \in (\mathbb{R} \times \mathbb{R}^N) \setminus \Omega$, che è un insieme chiuso. Quindi anche il suo limite $(p, u(p))$ vi appartiene, cioè $(p, u(p)) \notin \Omega$. Ma $p \in [\alpha, \beta]$, quindi, dalla (11.5), $(p, u(p)) \in \Omega$, assurdo. Questo dimostra che esiste $r > 0$ tale che $K \subset \Omega$.

Proviamo che K è compatto. Sia (t_n, y_n) una successione in K che converge a un certo limite $(t_0, y_0) \in \mathbb{R} \times \mathbb{R}^N$. Poiché $t_n \in [\alpha, \beta]$, anche il suo limite t_0 appartiene ad $[\alpha, \beta]$. Essendo u continua, $u(t_n) \rightarrow u(t_0)$. Poiché $|y_n - u(t_n)| \leq r$, passando al limite si ha $|y_0 - u(t_0)| \leq r$. Siccome $t_0 \in [\alpha, \beta]$, $|y_0 - u(t_0)| \leq r$, la coppia (t_0, y_0) appartiene a K . Quindi K è chiuso.

Essendo u continua e $[\alpha, \beta]$ compatto, u è limitata su $[\alpha, \beta]$, diciamo $|u(t)| \leq C$. Quindi per $(t, y) \in K$ si ha $|y| \leq |y - u(t)| + |u(t)| \leq r + C$, quindi K è limitato. Abbiamo provato che K è compatto.

Dal Teorema di Weierstrass e dalla Proposizione 6.4 di lipschitzianità sui compatti, esistono allora due costanti $M, L \geq 0$ tali che

$$|f(t, y)| \leq M, \quad |f(t, y) - f(t, z)| \leq L|y - z| \quad \forall (t, y), (t, z) \in K. \quad (11.6)$$

Passiamo ora al confronto tra le soluzioni u e v dei due problemi di Cauchy. Sia $\delta_0 > 0$, $(t_1, y_1) \in \Omega$, $t_1 \in [\alpha, \beta]$, con

$$|t_1 - t_0| \leq \delta_0, \quad |y_1 - y_0| \leq \delta_0.$$

Stimiamo la differenza tra u e v all'istante t_1 (istante iniziale del problema di Cauchy di v): poiché $|u'(t)| = |f(t, u(t))| \leq M$ per ogni $t \in [\alpha, \beta]$, si ha

$$|u(t_0) - u(t_1)| = \left| \int_{t_0}^{t_1} u'(t) dt \right| \leq M|t_1 - t_0| \leq M\delta_0;$$

poi $u(t_0) = y_0$, quindi $y_1 - u(t_1) = y_1 - y_0 + u(t_0) - u(t_1)$, e

$$\begin{aligned} |y_1 - u(t_1)| &\leq |y_1 - y_0| + |u(t_0) - u(t_1)| \\ &\leq |y_1 - y_0| + M|t_1 - t_0| \\ &\leq \delta_0 + M\delta_0 \\ &\leq \frac{r}{2}, \end{aligned} \quad (11.7)$$

e l'ultima disuguaglianza vale per δ_0 sufficientemente piccolo, cioè per $\delta_0 \leq r/(2 + 2M)$.

Vediamo ora per quali altri $t \geq t_1$ si ha $(t, v(t)) \in K$. Sia

$$E := \{p \in [t_1, \beta] : \text{la soluzione } v \text{ di (11.3) è definita in } [t_1, p], \\ \text{e soddisfa } (t, v(t)) \in K \text{ per ogni } t \in [t_1, p]\}.$$

Dalla (11.7), $t_1 \in E$, che è dunque non vuoto. Sia $p \in E$. Poiché il campo f in K soddisfa (11.6), si ha

$$|v'(t) - u'(t)| = |f(t, v(t)) - f(t, u(t))| \leq L|v(t) - u(t)| \quad \forall t \in [t_1, p],$$

cioè la differenza $w := v - u$ soddisfa $|w'| \leq L|w|$ in $[t_1, p]$. Di conseguenza, dal Lemma di Gronwall,

$$|v(t) - u(t)| \leq e^{L|t-t_1|}|v(t_1) - u(t_1)| \quad \forall t \in [t_1, p].$$

Poiché $|v(t_1) - u(t_1)| \leq (1 + M)\delta_0$ (vedi (11.7)), e $|t - t_1| \leq \beta - \alpha$, si ha

$$e^{L|t-t_1|}|v(t_1) - u(t_1)| \leq e^{L(\beta-\alpha)}(1 + M)\delta_0 \leq \frac{r}{2},$$

dove l'ultima disuguaglianza vale per δ_0 sufficientemente piccolo, cioè per $\delta_0 \leq e^{-L(\beta-\alpha)}r/(2 + 2M)$. Dunque

$$|v(t) - u(t)| \leq \frac{r}{2} \quad \forall t \in [t_1, p], \quad \forall p \in E. \quad (11.8)$$

Sia

$$c := \sup E,$$

e mostriamo che $c = \beta$.

Se $t_1 = \beta$, si ha banalmente $E = \{t_1\} = \{\beta\}$ e $c = \beta$.

Sia dunque $\alpha \leq t_1 < \beta$, e supponiamo, per assurdo, che $c < \beta$. Per definizione di c , esiste una successione (p_n) di elementi di E che converge a c . Poiché $(p_n, v(p_n))$ appartiene a K , che è compatto, esiste una sottosuccessione $(p_{n_k}, v(p_{n_k}))$ che converge in K : $p_{n_k} \rightarrow c$ (essendo sottosuccessione di (p_n)), e $v(p_{n_k}) \rightarrow v_0$ per un certo v_0 , con $(c, v_0) \in K \subset \Omega$. Allora, dal Criterio di prolungabilità con successione (Proposizione 7.6), v è prolungabile a destra di c . Poiché v è continua, si ha $v_0 = v(c)$.

Da (11.8), $|v(p_{n_k}) - u(p_{n_k})| \leq r/2$; passando al limite si trova $|v(c) - u(c)| \leq r/2$. Poiché u, v sono continue, si ha allora $|v - u| < r$ in un intervallo a destra di c .

Dunque esistono elementi di E che sono maggiori di c , assurdo. Questo dimostra che $c = \beta$, cioè v è definita in $[t_1, \beta)$, con $|v - u| \leq r$ in $[t_1, \beta)$. Quindi v è prolungabile a destra di β , e $|v - u| \leq r$ in $[t_1, \beta]$.

In modo simile si dimostra che v è definita in $[\alpha, t_1]$, con $|v - u| \leq r$ in $[\alpha, t_1]$. Dunque v è definita in tutto l'intervallo $[\alpha, \beta]$, e si ha

$$|v(t) - u(t)| \leq e^{L|t-t_1|}(|y_1 - y_0| + M|t_1 - t_0|) \quad \forall t \in [\alpha, \beta].$$

Per ottenere la stima con $L|t - t_0|$ all'esponente, osserviamo che $|(u - v)'| \leq L|u - v|$ in $[\alpha, \beta]$, da cui, per Gronwall,

$$|v(t) - u(t)| \leq e^{L|t-t_0|}|v(t_0) - u(t_0)| \quad \forall t \in [\alpha, \beta].$$

Poi

$$\begin{aligned} |v(t_0) - u(t_0)| &\leq |v(t_0) - v(t_1)| + |v(t_1) - u(t_0)| \\ &= \left| \int_{t_0}^{t_1} f(t, v(t)) dt \right| + |y_1 - y_0| \\ &\leq M|t_1 - t_0| + |y_1 - y_0|, \end{aligned}$$

da cui

$$|v(t) - u(t)| \leq e^{L|t-t_0|}(M|t_1 - t_0| + |y_1 - y_0|)$$

per ogni $t \in [\alpha, \beta]$. □

Osservazione 11.3. (*Il Teorema di dipendenza continua è ottimale*). Il Teorema di dipendenza continua non può, in generale, essere migliorato. Come primo esempio, consideriamo il caso $f(t, y) = y$, $t_0 = t_1 = 0$, $y_0 = 0$, $y_1 \neq 0$: i problemi di Cauchy (11.1) e (11.3) sono

$$\begin{cases} u' = u \\ u(0) = 0, \end{cases} \quad \begin{cases} v' = v \\ v(0) = y_1, \end{cases}$$

e le rispettive soluzioni sono

$$u(t) = 0, \quad v(t) = y_1 e^t,$$

entrambe definite su tutto \mathbb{R} . Si ha $v \rightarrow u = 0$ uniformemente su ogni intervallo $[\alpha, \beta]$ per $y_1 \rightarrow y_0 = 0$ (come dice il teorema), e inoltre la stima della differenza $|v(t) - u(t)|$ è proprio la (11.4) con $L = 1$. Tuttavia v non converge a u uniformemente su \mathbb{R} , e questo accade nonostante \mathbb{R} sia intervallo massimale di entrambe le soluzioni.

Come secondo esempio, consideriamo il caso $f(t, y) = y^2$, $t_0 = t_1 = 0$, $y_0 = 0$, $y_1 > 0$: i problemi di Cauchy (11.1) e (11.3) sono

$$\begin{cases} u' = u^2 \\ u(0) = 0, \end{cases} \quad \begin{cases} v' = v^2 \\ v(0) = y_1, \end{cases}$$

e le rispettive soluzioni sono

$$u(t) = 0, \quad v(t) = \frac{y_1}{1 - ty_1},$$

u definita su \mathbb{R} e v definita su $(-\infty, 1/y_1)$. Anche in questo caso, come dice il teorema, fissato un qualunque intervallo $[\alpha, \beta]$ si ha che v è definita su $[\alpha, \beta]$ per y_1 sufficientemente vicino a $y_0 = 0$, e $v \rightarrow u = 0$ uniformemente in $[\alpha, \beta]$ per $y_1 \rightarrow 0$. Ciononostante, v non converge a u uniformemente nel suo dominio $(-\infty, 1/y_1)$, e, per ogni $y_1 > 0$, v non è definita globalmente anche se u lo è.

Il teorema di dipendenza continua dai dati iniziali ci dice in particolare che la soluzione u del problema di Cauchy

$$\begin{cases} u' = f(t, u) \\ u(t_0) = y_0 \end{cases} \tag{11.9}$$

vista come funzione dei dati iniziali $(t_0, y_0) \in \mathbb{R} \times \mathbb{R}^N$ a valori nello spazio $C^0([\alpha, \beta], \mathbb{R}^N)$ delle funzioni continue su $[\alpha, \beta]$ a valori in \mathbb{R}^N munito della norma del sup

$$\|u\|_{C^0([\alpha, \beta], \mathbb{R}^N)} := \sup\{|u(t)| : t \in [\alpha, \beta]\}$$

è una funzione continua. Se, per evidenziare la dipendenza dai dati iniziali, indichiamo

$$\phi(t, t_0, y_0)$$

la soluzione di (11.9) valutata al tempo t , il teorema dice che

$$\|\phi(\cdot, t_1, y_1) - \phi(\cdot, t_0, y_0)\|_{C^0([\alpha, \beta], \mathbb{R}^N)} \rightarrow 0 \quad \text{per } (t_1, y_1) \rightarrow (t_0, y_0).$$

Vedremo tra alcune lezioni che, quando il campo f è di classe $C^1(\Omega)$, la dipendenza della soluzione dai dati iniziali è non solo continua, ma anche differenziabile.

* * *

Equazioni lineari. Per alcune lezioni vediamo in dettaglio una classe importante di equazioni differenziali, le equazioni *lineari*: ne abbiamo già parlato nel caso scalare ($N = 1$), provando la formula risolutiva (vedi (4.3), (4.7)); ora trattiamo il caso vettoriale generale.

Definizione 11.4. (*Equazioni lineari del primo ordine in dimensione N*). L'equazione $u' = f(t, u)$ si dice *lineare* se il campo f è definito in $\Omega = I \times \mathbb{R}^N$, con $I \subseteq \mathbb{R}$ intervallo, ed è della forma $f(t, y) = A(t)y + b(t)$, con $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$, $b : I \rightarrow \mathbb{R}^N$, A, b funzioni continue. Le equazioni lineari sono dunque le equazioni differenziali della forma

$$u' = A(t)u + b(t); \tag{11.10}$$

per ogni t , $A(t)$ è una matrice, $b(t)$ un vettore.

Osservazione 11.5. (*Funzioni a valori nello spazio delle matrici*). $\text{Mat}_{N \times N}(\mathbb{R})$ indica l'insieme delle matrici a N righe e N colonne a entrate reali. Nella Lezione 3 abbiamo definito due delle principali norme di matrici, quella euclidea (Definizione 3.4) e quella operatoriale (Definizione 3.7), e abbiamo osservato che sono norme equivalenti (Esercizio 3.9).

La definizione di continuità è quella standard: una funzione $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ è continua in $t \in I$ se $\|A(s) - A(t)\|_E \rightarrow 0$ per $s \rightarrow t$. Si ha poi $\|A(s) - A(t)\|_E \rightarrow 0$ se e solo se $\|A(s) - A(t)\|_{op} \rightarrow 0$.

Come accade per le funzioni a valori in \mathbb{R}^N , anche per le funzioni a valori in $\text{Mat}_{N \times N}(\mathbb{R})$ la continuità è equivalente alla continuità di ogni componente: $A(t)$ è funzione continua di t se e solo se ogni entrata $A_{jk}(t)$ è una funzione (scalare) continua di t . La dimostrazione è identica a quella fatta in Analisi 2 per le funzioni a valori in \mathbb{R}^N .

Esercizio 11.6. (*In dimensione finita tutte le norme sono equivalenti*). In realtà *tutte* le norme in $\text{Mat}_{N \times N}(\mathbb{R})$ (e non solo quelle euclidea e operatoriale) sono equivalenti, perché $\text{Mat}_{N \times N}(\mathbb{R})$ è uno spazio vettoriale di dimensione finita. Questa è una proprietà generale degli spazi vettoriali, non ha a che fare, in particolare, con le matrici.

Dimostrare che, dato un qualsiasi spazio vettoriale V su \mathbb{R} di dimensione N , e date due norme qualsiasi $\| \cdot \|_A, \| \cdot \|_B$ su V , esse sono equivalenti.

Suggerimento: fissare una base v_1, \dots, v_N di V ; per $v = \sum x_k v_k \in V$, con $x_1, \dots, x_N \in \mathbb{R}$, definire $\|v\|_{E,V}^2 := \sum x_k^2$; dimostrare che $\| \cdot \|_{E,V}$ è una norma in V ; sia $\| \cdot \|_A$ una qualsiasi norma su V : dimostrare che esiste M tale che $\|v\|_A \leq M\|v\|_{E,V}$ per ogni $v \in V$; dimostrare che l'insieme $J := \{\|v\|_A : v \in V, \|v\|_{E,V} = 1\}$ ha $\inf J = c > 0$; dedurre che $\|v\|_A \geq c\|v\|_{E,V}$ per ogni $v \in V$; quindi ogni norma $\| \cdot \|_A$ è equivalente a $\| \cdot \|_{E,V}$; concludere.

Definizione 11.7. (*Equazioni lineari omogenee/non omogenee*). Per $b = 0$, l'equazione (11.10) diventa

$$u' = A(t)u, \quad (11.11)$$

e si dice equazione lineare *omogenea*. L'equazione (11.10) si dice *non omogenea*; in tal caso, l'equazione (11.11) è la sua equazione *omogenea associata*; $b(t)$ è il *termine non omogeneo* dell'equazione.

Lemma 11.8. (*Esistenza globale per equazioni lineari*). *Sia $I \subseteq \mathbb{R}$ intervallo, $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$, $b : I \rightarrow \mathbb{R}^N$, con A, b continue. Per ogni $(t_0, y_0) \in I \times \mathbb{R}^N$ esiste, ed è unica, e di classe C^1 , la soluzione massimale del problema di Cauchy*

$$u' = A(t)u + b(t), \quad u(t_0) = y_0;$$

tale soluzione è globale, cioè è definita in tutto l'intervallo I .

Dimostrazione. Il campo vettoriale $f(t, y) = A(t)y + b(t)$ soddisfa le ipotesi dei teoremi di esistenza e unicità locale, del prolungamento massimale, e di esistenza globale (si ha infatti $|f(t, y)| \leq |A(t)||y| + |b(t)|$). \square

Equazioni omogenee. Vediamo alcune proprietà delle equazioni omogenee $u' = A(t)u$, basate principalmente sull'algebra lineare.

Teorema 11.9. (*Le soluzioni di un'equazione lineare omogenea formano uno spazio vettoriale*). *Sia $I \subseteq \mathbb{R}$ un intervallo, e sia $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ una funzione continua. Sia*

$$S := \{u \in C^1(I, \mathbb{R}^N) : u'(t) = A(t)u(t) \quad \forall t \in I\}$$

l'insieme delle soluzioni dell'equazione lineare omogenea $u' = A(t)u$. Allora

- (i) *l'insieme S è uno spazio vettoriale su \mathbb{R} , di dimensione N ;*
- (ii) *per ogni $t_0 \in I$, la mappa di "valutazione in t_0 " che a ogni elemento $u \in S$ associa il suo valore all'istante t_0 ,*

$$\delta_{t_0} : S \rightarrow \mathbb{R}^N, \quad u \mapsto \delta_{t_0}(u) := u(t_0),$$

è lineare e biiettiva (isomorfismo di spazi vettoriali);

- (iii) *per ogni $t_0 \in I$, l'inversa di δ_{t_0} è la mappa $\Phi_{t_0} : \mathbb{R}^N \rightarrow S$ che a ogni vettore $y_0 \in \mathbb{R}^N$ associa la soluzione del problema di Cauchy*

$$\begin{cases} u' = A(t)u, \\ u(t_0) = y_0; \end{cases} \quad (11.12)$$

Φ_{t_0} è lineare e biiettiva (isomorfismo di spazi vettoriali), e si ha

$$\Phi_{t_0}(\delta_{t_0}(u)) = u \quad \forall u \in S; \quad \delta_{t_0}(\Phi_{t_0}(y_0)) = y_0 \quad \forall y_0 \in \mathbb{R}^N. \quad (11.13)$$

Dimostrazione. (i) Se $u, v \in S$, $\lambda \in \mathbb{R}$, allora $u+v \in S$, $\lambda u \in S$ (verifica immediata); quindi S è uno spazio vettoriale su \mathbb{R} . Resta da provare che ha dimensione N .

(ii) Se $u, v \in S$, $\lambda \in \mathbb{R}$, si ha banalmente

$$\begin{aligned}\delta_{t_0}(u+v) &= (u+v)(t_0) = u(t_0) + v(t_0) = \delta_{t_0}(u) + \delta_{t_0}(v), \\ \delta_{t_0}(\lambda u) &= (\lambda u)(t_0) = \lambda u(t_0) = \lambda \delta_{t_0}(u),\end{aligned}$$

quindi δ_{t_0} è una mappa lineare. Se $\delta_{t_0}(u) = \delta_{t_0}(v)$, allora u, v sono soluzioni della stessa equazione differenziale e coincidono in t_0 , dunque coincidono in tutto l'intervallo I (Corollario 3.15); quindi δ_{t_0} è iniettiva. Dato $y_0 \in \mathbb{R}^N$, sia u la soluzione del problema di Cauchy (11.12); $u \in S$ (Lemma 11.8) e $\delta_{t_0}(u) = u(t_0) = y_0$, quindi δ_{t_0} è suriettiva.

(iii) Sia $y_0 \in \mathbb{R}^N$; $\Phi_{t_0}(y_0)$ è la soluzione del problema di Cauchy (11.12), e all'istante t_0 vale y_0 : questo prova che $\delta_{t_0}(\Phi_{t_0}(y_0)) = y_0$. Sia $u \in S$. Allora u è l'unica soluzione del problema di Cauchy che all'istante t_0 vale $u(t_0)$, cioè $u = \Phi_{t_0}(\delta_{t_0}(u))$. Abbiamo provato (11.13). Da (11.13) segue che Φ_{t_0} è biiettiva e δ_{t_0} è la sua inversa; da (11.13) e dal fatto che δ_{t_0} sia lineare segue che Φ_{t_0} è lineare (in alternativa, si può anche dimostrare direttamente che Φ_{t_0} è lineare).

(i) Infine S ha dimensione N perché Φ_{t_0} è isomorfismo di spazi vettoriali. \square

Definizione 11.10. (*Funzioni linearmente indipendenti*). Sia $I \subseteq \mathbb{R}$ un intervallo, e siano u_1, \dots, u_n funzioni da I a valori in \mathbb{R}^N . Diciamo che u_1, \dots, u_n sono *linearmente indipendenti* se sono linearmente indipendenti come funzioni in I , cioè se l'unica loro combinazione lineare che dà la funzione identicamente nulla su I è quella a coefficienti tutti nulli, cioè se gli unici numeri reali $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ per cui si ha

$$\lambda_1 u_1(t) + \dots + \lambda_n u_n(t) = 0 \quad \forall t \in I$$

sono $\lambda_1 = \dots = \lambda_n = 0$.

Lemma 11.11. (Soluzioni linearmente indipendenti e valutazione). *Sia $I \subseteq \mathbb{R}$ un intervallo, $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ una funzione continua, $n \in \mathbb{N}$, e siano u_1, \dots, u_n soluzioni dell'equazione lineare omogenea $u' = A(t)u$ in I . Sono equivalenti:*

- (i) u_1, \dots, u_n sono soluzioni linearmente indipendenti;
- (ii) per ogni $t \in I$, $u_1(t), \dots, u_n(t)$ sono vettori linearmente indipendenti di \mathbb{R}^N ;
- (iii) per un certo $t_0 \in I$, $u_1(t_0), \dots, u_n(t_0)$ sono vettori linearmente indipendenti di \mathbb{R}^N .

Dimostrazione. (i) \Rightarrow (ii). Siano $u_1, \dots, u_n \in S$ linearmente indipendenti. Per ogni $t \in I$, la mappa $\delta_t : S \rightarrow \mathbb{R}^N$ è isomorfismo di spazi vettoriali, quindi $\delta_t(u_1), \dots, \delta_t(u_n)$ sono vettori linearmente indipendenti di \mathbb{R}^N .

(ii) \Rightarrow (iii). Ovvio.

(iii) \Rightarrow (i). Siano $u_1(t_0), \dots, u_n(t_0)$ vettori linearmente indipendenti di \mathbb{R}^N . Per ogni i si ha $u_i = \Phi_{t_0}(\delta_{t_0}(u_i)) = \Phi_{t_0}(u_i(t_0))$; la mappa $\Phi_{t_0} : \mathbb{R}^N \rightarrow S$ è isomorfismo di spazi vettoriali, quindi u_1, \dots, u_n sono linearmente indipendenti in S . \square

Esempio. L'equazione lineare (11.11) in dimensione $N = 2$ con matrice

$$A(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (11.14)$$

è il sistema

$$\begin{cases} x' = y, \\ y' = -x \end{cases} \quad (11.15)$$

dove abbiamo indicato $u(t) = (x(t), y(t))$. Come osservato nella Lezione 1, le soluzioni sono le funzioni

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} c_1 \cos t + c_2 \sin t \\ -c_1 \sin t + c_2 \cos t \end{pmatrix}, \quad c_1, c_2 \in \mathbb{R}. \quad (11.16)$$

Le soluzioni

$$u_1(t) := \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}, \quad u_2(t) := \begin{pmatrix} \sin t \\ \cos t \end{pmatrix} \quad (11.17)$$

sono linearmente indipendenti: si ha $\lambda_1 u_1(t) + \lambda_2 u_2(t) = 0$ per ogni $t \in \mathbb{R}$ se e solo se

$$\begin{cases} \lambda_1 \cos t + \lambda_2 \sin t = 0 \\ -\lambda_1 \sin t + \lambda_2 \cos t = 0, \end{cases}$$

cioè se e solo se

$$\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (11.18)$$

e questo è vero se e solo se $(\lambda_1, \lambda_2) = (0, 0)$ perché per ogni $t \in \mathbb{R}$ la matrice in (11.18) è invertibile (ha determinante = 1).

12 Lezione 12

Definizione 12.1. (*Matrice wronskiana, determinante wronskiano*). Sia $I \subseteq \mathbb{R}$ intervallo, $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ funzione continua. Siano u_1, \dots, u_N soluzioni dell'equazione lineare omogenea $u' = A(t)u$ in I . Si chiama *matrice wronskiana* la matrice

$$M(t) = (u_1(t) \quad \dots \quad u_N(t))$$

che ha $u_k(t)$ come k -esima colonna; si chiama *determinante wronskiano* il suo determinante.

Esempio. La matrice in (11.18) è la matrice wronskiana delle soluzioni (11.17); il determinante wronskiano è $= 1$ per ogni $t \in \mathbb{R}$.

Lemma 12.2. (Teorema del wronskiano). *Sia $I \subseteq \mathbb{R}$ un intervallo, e sia $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ una funzione continua. Siano u_1, \dots, u_N soluzioni dell'equazione lineare omogenea $u' = A(t)u$ in I , e sia $W(t)$ il loro determinante wronskiano.*

Se u_1, \dots, u_N sono soluzioni linearmente indipendenti, allora $W(t) \neq 0$ per ogni $t \in I$; se invece u_1, \dots, u_N sono soluzioni linearmente dipendenti, allora $W(t) = 0$ per ogni $t \in I$.

Dimostrazione. Conseguenza diretta del Lemma 11.11. □

Osservazione 12.3. (*Il determinante wronskiano è $C^1(I)$ e ha segno costante*). Il determinante wronskiano $W(t)$ è una somma di prodotti delle componenti dei vettori $u_1(t), \dots, u_N(t)$, che sono funzioni di classe $C^1(I)$, vedi Lemma 11.8); quindi il determinante wronskiano è una funzione $C^1(I)$.

Dal Lemma 12.2 e dalla continuità di $W(t)$ segue allora che $W(t)$ ha segno costante: si ha

$$\begin{aligned} W(t) &> 0 \text{ per ogni } t \in I, \text{ oppure} \\ W(t) &< 0 \text{ per ogni } t \in I, \text{ oppure} \\ W(t) &= 0 \text{ per ogni } t \in I. \end{aligned}$$

Lemma 12.4. (Equazione differenziale della matrice wronskiana). *Sia $I \subseteq \mathbb{R}$ un intervallo, e sia $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ una funzione continua. Siano u_1, \dots, u_N soluzioni dell'equazione lineare omogenea $u' = A(t)u$ in I , e sia*

$$M(t) = (u_1(t) \quad \dots \quad u_N(t))$$

la matrice wronskiana di u_1, \dots, u_N . Allora $M(t)$ è soluzione dell'equazione differenziale (matriciale)

$$M'(t) = A(t)M(t).$$

Questo significa che, detta $M_{jk}(t)$ l'entrata di posto (j, k) ($j = \text{riga}$, $k = \text{colonna}$) della matrice $M(t)$, la sua derivata $M'_{jk}(t)$ è uguale all'entrata di posto (j, k) della matrice prodotto $A(t)M(t)$ (prodotto righe per colonne), cioè

$$M'_{jk}(t) = \sum_{\ell=1}^N A_{j\ell}(t)M_{\ell k}(t) \quad \forall t \in I, \quad \forall j, k = 1, \dots, N.$$

Dimostrazione. Si ha

$$\begin{aligned} M'(t) &= (u_1'(t) \quad u_2'(t) \quad \dots \quad u_N'(t)) \\ &= (A(t)u_1(t) \quad A(t)u_2(t) \quad \dots \quad A(t)u_N(t)) \\ &= A(t) (u_1(t) \quad u_2(t) \quad \dots \quad u_N(t)) \\ &= A(t)M(t). \end{aligned}$$

Abbiamo usato questa semplice identità di algebra lineare: se A, B sono matrici, e b_1, \dots, b_N sono le colonne di B , allora Ab_1, \dots, Ab_N sono le colonne di AB . \square

Esercizio 12.5. (*Equazione differenziale del determinante wronskiano e formula di Liouville*). Si assumano le ipotesi del Lemma 12.4, e sia $W(t) = \det M(t)$ il determinante wronskiano.

Dimostrare che $W(t)$ è soluzione dell'equazione differenziale (scalare)

$$W'(t) = (\text{Tr } A(t)) W(t) \quad (12.1)$$

dove $\text{Tr } A(t) = A_{11}(t) + \dots + A_{NN}(t)$ è la traccia della matrice $A(t)$.

Dalla (12.1), ricordando (4.3), si ottiene immediatamente la *formula di Liouville*

$$W(t) = W(t_0) \exp \left(\int_{t_0}^t \text{Tr } A(s) ds \right).$$

Un suggerimento per dimostrare (12.1): scrivere $W(t)$ usando la formula di Leibniz per il determinante.

Definizione 12.6. (*Matrice risolvente*). Sia $I \subseteq \mathbb{R}$ intervallo, $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ funzione continua. Sia $t_0 \in I$. Per ogni $k = 1, \dots, N$, sia u_k la soluzione del problema di Cauchy

$$\begin{cases} u_k' = A(t)u_k, \\ u_k(t_0) = e_k, \end{cases}$$

dove e_k è il k -esimo vettore della base canonica di \mathbb{R}^N (cioè la k -esima colonna della matrice identità I). La matrice wronskiana delle soluzioni u_1, \dots, u_N si dice *matrice risolvente*, e si indica $R(t, t_0)$.

Osservazione 12.7. (*Problema di Cauchy della matrice risolvente*). La matrice risolvente è l'unica soluzione del problema di Cauchy (matriciale)

$$\begin{cases} M'(t) = A(t)M(t) \\ M(t_0) = I. \end{cases} \quad (12.2)$$

È soluzione dell'equazione perché è una matrice wronskiana; in t_0 coincide con la matrice identità per costruzione. Inoltre la soluzione di (12.2) è unica perché, se $B(t), M(t)$ sono due soluzioni, allora, per ogni $x \in \mathbb{R}^N$, si ha $M(t)x = B(t)x$, essendo entrambe soluzione del problema di Cauchy $u' = A(t)u$, $u(t_0) = x$.

Indicando $R(t, t_0)$ la matrice risolvente, il problema di Cauchy (12.2) si riscrive

$$\begin{cases} \partial_t R(t, t_0) = A(t)R(t, t_0), \\ R(t_0, t_0) = I. \end{cases} \quad (12.3)$$

Il prossimo lemma giustifica il nome della matrice risolvente.

Lemma 12.8. (La matrice risolvente dà la soluzione del problema di Cauchy). *Sia $I \subseteq \mathbb{R}$ intervallo, $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ funzione continua. Sia $R(t, t_0)$ la matrice risolvente (Definizione 12.6). Per ogni $y_0 \in \mathbb{R}^N$, la soluzione u del problema di Cauchy*

$$\begin{cases} u' = A(t)u \\ u(t_0) = y_0 \end{cases} \quad (12.4)$$

è la funzione

$$u(t) = R(t, t_0)y_0,$$

data dal prodotto righe per colonne della matrice risolvente $R(t, t_0)$ per il vettore y_0 .

Dimostrazione. Indichiamo $M(t) = R(t, t_0)$ la matrice risolvente. Sia $u(t) := M(t)y_0$. Allora

$$u'(t) = M'(t)y_0 = A(t)M(t)y_0 = A(t)u(t), \quad u(t_0) = M(t_0)y_0 = Iy_0 = y_0,$$

quindi $u(t) = M(t)y_0$ è la soluzione del problema di Cauchy. \square

Lemma 12.9. (Proprietà della matrice risolvente). *Sia $I \subseteq \mathbb{R}$ intervallo, $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ funzione continua. Per ogni $t, \sigma, s \in I$, la matrice risolvente soddisfa*

$$R(t, \sigma)R(\sigma, s) = R(t, s). \quad (12.5)$$

Per ogni $t, s \in I$, $R(t, s)$ è invertibile e la sua inversa è

$$R(t, s)^{-1} = R(s, t). \quad (12.6)$$

Inoltre

$$R(t, s) = \delta_t \circ \Phi_s \quad (12.7)$$

dove δ_t, Φ_s sono gli isomorfismi definiti nel Teorema 11.9.

Dimostrazione. La funzione $u(t) = \Phi_{t_0}(y_0)(t)$, $t \in I$, è la soluzione del problema di Cauchy (12.4). Per ogni $t_0, t_1 \in I$, dal Lemma 12.8, $R(t_1, t_0)y_0$ è la soluzione del problema di Cauchy (12.4) valutata all'istante t_1 , cioè

$$R(t_1, t_0)y_0 = u(t_1) = \Phi_{t_0}(y_0)(t_1) = \delta_{t_1}(\Phi_{t_0}(y_0)).$$

Questo prova la formula (12.7).

Dalla (12.7), ricordando che $\Phi_\sigma \circ \delta_\sigma$ è l'identità (vedi (11.13)), si ha

$$\begin{aligned} R(t, \sigma)R(\sigma, s)y_0 &= \delta_t(\Phi_\sigma(\delta_\sigma(\Phi_s(y_0)))) \\ &= \delta_t(\Phi_s(y_0)) \\ &= R(t, s)y_0, \end{aligned}$$

cioè (12.5). Prendendo $t = s$ nella (12.5), e ricordando che $R(t_0, t_0) = I$ per ogni t_0 (vedi (12.3)), si ha $R(t, \sigma)R(\sigma, t) = I$ per ogni $t, \sigma \in I$, cioè (12.6). \square

Equazioni non omogenee. Vediamo ora alcune proprietà delle equazioni lineari non omogenee

$$u' = A(t)u + b(t). \quad (12.8)$$

Teorema 12.10. (Le soluzioni di un'equazione lineare non omogenea formano uno spazio affine). Sia $I \subseteq \mathbb{R}$ un intervallo, e siano $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$, $b : I \rightarrow \mathbb{R}^N$ funzioni continue. Sia $u_P : I \rightarrow \mathbb{R}^N$ una soluzione dell'equazione non omogenea (12.8), cioè

$$u'_P = A(t)u_P + b(t) \quad \forall t \in I.$$

Sia S (come definito nel Teorema 11.9) lo spazio vettoriale delle soluzioni dell'equazione omogenea $u' = A(t)u$,

$$S = \{u : I \rightarrow \mathbb{R}^N : u' = A(t)u \quad \forall t \in I\}.$$

Allora le soluzioni dell'equazione non omogenea $u' = A(t)u + b(t)$ sono le funzioni $u = u_P + v$ con v soluzione dell'omogenea associata, cioè

$$\{u : I \rightarrow \mathbb{R}^N : u'(t) = A(t)u(t) + b(t) \quad \forall t \in I\} = \{u_P + v : v \in S\}.$$

Dimostrazione. Sia u soluzione dell'equazione non omogenea, e sia $v := u - u_P$. Allora

$$v' = u' - u'_P = A(t)u + b(t) - A(t)u_P - b(t) = A(t)v,$$

quindi $v \in S$, e $u = u_P + v$.

Viceversa, sia $v \in S$ e sia $u = u_P + v$. Allora

$$u' = u'_P + v' = A(t)u_P + b(t) + A(t)v = A(t)u + b(t),$$

cioè u è soluzione dell'equazione non omogenea. □

Osservazione 12.11. (*Soluzione particolare*). Grazie al Teorema 12.10, per scrivere tutte le soluzioni dell'equazione non omogenea $u' = A(t)u + b(t)$ ci basta conoscere una soluzione dell'equazione non omogenea, che negli esercizi viene spesso detta “soluzione particolare” (la u_P del Teorema 12.10), e tutte le soluzioni dell'omogenea associata.

Metodo di variazione delle costanti. Il *metodo di variazione delle costanti* è una tecnica per costruire una soluzione dell'equazione non omogenea a partire da N soluzioni linearmente indipendenti dell'equazione omogenea associata.

Sia $I \subseteq \mathbb{R}$ un intervallo, e siano $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$, $b : I \rightarrow \mathbb{R}^N$ funzioni continue. Supponiamo che u_1, \dots, u_N siano soluzioni linearmente indipendenti dell'equazione omogenea $u' = A(t)u$.

Sia $M(t)$ la matrice wronskiana delle soluzioni u_1, \dots, u_N (vedi Definizione 12.1). Cerchiamo una soluzione u dell'equazione non omogenea $u' = A(t)u + b(t)$ della forma

$$u(t) = M(t)h(t),$$

dove $h : I \rightarrow \mathbb{R}^N$ è una funzione del tempo, da determinare. Se h fosse un vettore costante y_0 , allora $M(t)y_0$ sarebbe una soluzione dell'equazione omogenea; qui invece l'idea è quella di far “variare le costanti”, e considerare $h(t)$ che varia nel tempo, da scegliere in modo che $u' - A(t)u$ non dia zero, ma $b(t)$. Questa tecnica è già stata utilizzata in dimensione $N = 1$ per lo stesso scopo: rivedere l'ultima pagina della Lezione 4.

Sia dunque $u(t) = M(t)h(t)$; ricordando il Lemma 12.4, calcoliamo la derivata:

$$\begin{aligned} u'(t) &= M'(t)h(t) + M(t)h'(t) \\ &= A(t)M(t)h(t) + M(t)h'(t) \\ &= A(t)u(t) + M(t)h'(t). \end{aligned}$$

Cerchiamo $h(t)$ tale che $M(t)h'(t) = b(t)$. Poiché u_1, \dots, u_N sono soluzioni linearmente indipendenti, la matrice wronskiana $M(t)$ è invertibile per ogni $t \in I$ (vedi Lemma 12.2). Quindi l'equazione $M(t)h'(t) = b(t)$ si può riscrivere come $h'(t) = M(t)^{-1}b(t)$. Scegliamo dunque

$$h(t) = \int_{t_0}^t M(s)^{-1}b(s) ds,$$

cosicché $h'(t) = M(t)^{-1}b(t)$, e otteniamo

$$u_P(t) = M(t) \int_{t_0}^t M(s)^{-1}b(s) ds,$$

soluzione dell'equazione non omogenea. La soluzione generale dell'equazione non omogenea si ottiene sommando le soluzioni dell'omogenea associata,

$$u(t) = M(t) \left(c + \int_{t_0}^t M(s)^{-1}b(s) ds \right), \quad c \in \mathbb{R}^N.$$

Se poi $M(t)$ è la matrice risolvete $R(t, t_0)$, allora

$$M(t)M(s)^{-1} = R(t, t_0)R(s, t_0)^{-1} = R(t, t_0)R(t_0, s) = R(t, s),$$

e $u(t_0) = M(t_0)c = R(t_0, t_0)c = c$. Si trova quindi la formula

$$u(t) = R(t, t_0)u(t_0) + \int_{t_0}^t R(t, s)b(s) ds.$$

Abbiamo provato la seguente proposizione.

Proposizione 12.12. (Soluzioni dell'equazione non omogenea tramite variazione delle costanti). *Sia $I \subseteq \mathbb{R}$ intervallo, $A : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$, $b : I \rightarrow \mathbb{R}^N$ funzioni continue. Supponiamo che u_1, \dots, u_N siano soluzioni linearmente indipendenti dell'equazione omogenea $u' = A(t)u$.*

Sia $M(t)$ la matrice wronskiana delle soluzioni u_1, \dots, u_N (vedi Definizione 12.1). Allora le soluzioni dell'equazione non omogenea $u' = A(t)u + b(t)$ sono tutte e sole le funzioni

$$u(t) = M(t) \left(c + \int_{t_0}^t M(s)^{-1}b(s) ds \right), \quad (12.9)$$

con $c \in \mathbb{R}^N$ vettore costante.

Se poi u_1, \dots, u_N sono le soluzioni dell'equazione omogenea $u' = A(t)u$ con dati iniziali $u_k(t_0) = e_k$, allora la matrice wronskiana $M(t)$ è la matrice risolvete $R(t, t_0)$ (vedi Definizione 12.6), e le soluzioni (12.9) si riscrivono

$$u(t) = R(t, t_0)u(t_0) + \int_{t_0}^t R(t, s)b(s) ds. \quad (12.10)$$

Esercizio 12.13. (*Matrice risolvente e variazione delle costanti*). Sia $A(t)$ la matrice (11.14), e si consideri la corrispondente equazione omogenea 2-dimensionale (11.15).

(i) Calcolare la matrice risolvente $R(t, s)$ per ogni $t, s \in \mathbb{R}$.

(ii) Scrivere la formula generale dell'equazione non omogenea $u' = A(t)u + b(t)$, cioè del sistema

$$\begin{cases} x'(t) = y(t) + f(t), \\ y'(t) = -x(t) + g(t), \end{cases}$$

dove $(f(t), g(t))$ sono le componenti del termine non omogeneo $b(t)$.

Equazioni lineari a coefficienti costanti. Vediamo ora un tipo particolarmente importante di equazioni lineari, quelle a *coefficienti costanti*, cioè quelle in cui la matrice $A(t)$ è una matrice costante A , indipendente da t . Per calcolare le soluzioni dell'equazione $u' = Au$, ripartiamo dalla costruzione della soluzione come limite della successione delle iterate di Picard, come nella dimostrazione del Teorema 2.1 di esistenza e unicità locale.

Lemma 12.14. (Soluzione dell'equazione $u' = Au$ tramite iterate di Picard). Sia $A \in \text{Mat}_{N \times N}(\mathbb{R})$ una matrice, $y_0 \in \mathbb{R}^N$. La soluzione del problema di Cauchy

$$\begin{cases} u' = Au, \\ u(t_0) = y_0 \end{cases} \quad (12.11)$$

è definita in \mathbb{R} ed è la somma della serie di funzioni

$$u(t) = \sum_{n=0}^{\infty} \frac{1}{n!} (t - t_0)^n A^n y_0. \quad (12.12)$$

Tale serie converge totalmente in $[t_0 - r, t_0 + r]$ per ogni $r > 0$. La successione delle sue somme parziali è la successione delle iterate di Picard del problema (12.11).

Dimostrazione. La successione (u_n) delle iterate di Picard di (12.11) è definita ricorsivamente come

$$u_0(t) = y_0, \quad u_{n+1}(t) = y_0 + \int_{t_0}^t Au_n(s) ds, \quad n \in \mathbb{N}.$$

In modo diretto calcoliamo

$$u_1(t) = y_0 + (t - t_0)Ay_0, \quad u_2(t) = y_0 + (t - t_0)Ay_0 + \frac{1}{2}(t - t_0)^2 A^2 y_0, \quad \dots$$

e, per induzione,

$$u_n(t) = \sum_{k=0}^n \frac{1}{k!} (t - t_0)^k A^k y_0, \quad \forall n \in \mathbb{N}.$$

Si ha poi $|Ay_0| \leq |A||y_0|$, dove $|A|$ è la norma euclidea della matrice A (vedi Lemma 3.5). Per induzione,

$$|A^k y_0| \leq |A|^k |y_0| \quad \forall k \in \mathbb{N}.$$

Quindi per ogni $r > 0$, per ogni $t \in [t_0 - r, t_0 + r]$ si ha

$$\left| \frac{1}{k!} (t - t_0)^k A^k y_0 \right| \leq \frac{r^k |A|^k}{k!} |y_0|.$$

Dunque $u_n(t)$ è la somma parziale di una serie che converge totalmente, e quindi uniformemente, in $[t_0 - r, t_0 + r]$, per ogni $r > 0$. Sia $u(t)$ il limite di $u_n(t)$; la convergenza uniforme permette il passaggio al limite sotto il segno di integrale, da cui segue che

$$u(t) = y_0 + \int_{t_0}^t Au(s) ds \quad \forall t \in [t_0 - r, t_0 + r], \quad \forall r > 0,$$

cioè u è soluzione del problema di Cauchy (12.11), ed è definita in ogni $t \in \mathbb{R}$. \square

In analogia con il caso scalare, motivati dalla formula (12.12), definiamo l'esponenziale di una matrice come serie di potenze. Prima della definizione, facciamo qualche osservazione sulla norma del prodotto di matrici. Ricordiamo che la norma euclidea $|A|_{\mathbb{E}}$ e quella operatoriale $|A|_{\text{op}}$ di una matrice A sono state definite nelle Definizioni 3.4 e 3.7.

Lemma 12.15. (Norme di matrici). (i) Per ogni $A \in \text{Mat}_{N \times N}(\mathbb{R})$, sia la norma euclidea che quella operatoriale soddisfano la disuguaglianza per il prodotto matrice per vettore

$$|Ax| \leq |A|_{\text{op}}|x|, \quad |Ax| \leq |A|_{\mathbb{E}}|x| \quad \forall x \in \mathbb{R}^N.$$

(ii) Per ogni $A, B \in \text{Mat}_{N \times N}(\mathbb{R})$, sia la norma euclidea che quella operatoriale soddisfano la disuguaglianza per il prodotto di matrici

$$|AB|_{\text{op}} \leq |A|_{\text{op}}|B|_{\text{op}}, \quad |AB|_{\mathbb{E}} \leq |A|_{\mathbb{E}}|B|_{\mathbb{E}}.$$

(iii) Per ogni $A \in \text{Mat}_{N \times N}(\mathbb{R})$, per ogni $n \in \mathbb{N}$, sia la norma euclidea che quella operatoriale soddisfano la disuguaglianza per la potenza di matrice

$$|A^n|_{\text{op}} \leq |A|_{\text{op}}^n, \quad |A^n|_{\mathbb{E}} \leq |A|_{\mathbb{E}}^n.$$

Dimostrazione. (i) La disuguaglianza con la norma euclidea è stata dimostrata nel Lemma 3.5. Per ogni $x \neq 0$ si ha $|Ax|/|x| \leq |A|_{\text{op}}$ per definizione di $|A|_{\text{op}}$; quindi $|Ax| \leq |A|_{\text{op}}|x|$ per $x \neq 0$. Per $x = 0$ la disuguaglianza è banalmente vera.

(ii) Applicando (i) due volte, per ogni x si ha

$$|ABx| \leq |A|_{\text{op}}|Bx| \leq |A|_{\text{op}}|B|_{\text{op}}|x|,$$

da cui

$$\frac{|ABx|}{|x|} \leq |A|_{\text{op}}|B|_{\text{op}} \quad \forall x \neq 0.$$

Passando al sup su $x \neq 0$ si ottiene $|AB|_{\text{op}} \leq |A|_{\text{op}}|B|_{\text{op}}$.

Per provare la disuguaglianza con la norma euclidea, indichiamo b_k la k -esima colonna di B , e osserviamo che Ab_k è la k -esima colonna di AB . Quindi

$$|AB|_{\mathbb{E}}^2 = \sum_k |Ab_k|^2 \leq \sum_k |A|_{\mathbb{E}}^2 |b_k|^2 = |A|_{\mathbb{E}}^2 \sum_k |b_k|^2 = |A|_{\mathbb{E}}^2 |B|_{\mathbb{E}}^2.$$

(iii) Per induzione, usando (ii). \square

Definizione 12.16. (*Esponenziale di matrice*). Sia $A \in \text{Mat}_{N \times N}(\mathbb{R})$. Definiamo *esponenziale della matrice A* la matrice

$$e^A := \exp A := \sum_{n=0}^{\infty} \frac{1}{n!} A^n. \quad (12.13)$$

Osservazione 12.17. (*Convergenza della serie esponenziale*). Dal Lemma 12.15 (iii), il termine generale della serie (12.13) ha norma $|A^n/n!|_{\mathbb{E}} \leq |A|_{\mathbb{E}}^n/n!$, quindi la serie è di Cauchy, e quindi converge in $\text{Mat}_{N \times N}(\mathbb{R})$: la definizione di $\exp A$ è ben posta. Inoltre, per ogni n ,

$$\left| \sum_{k=0}^n \frac{1}{k!} A^k \right|_{\mathbb{E}} \leq \sum_{k=0}^n \left| \frac{1}{k!} A^k \right|_{\mathbb{E}} \leq \sum_{k=0}^n \frac{1}{k!} |A|_{\mathbb{E}}^k,$$

e lo stesso vale con la norma operatoriale; passando al limite, si ha

$$|e^A|_{\mathbb{E}} \leq e^{|A|_{\mathbb{E}}}, \quad |e^A|_{\text{op}} \leq e^{|A|_{\text{op}}}.$$

13 Lezione 13

Lemma 13.1. (La matrice esponenziale è la matrice risolvente di $u' = Au$).

(i) La soluzione del problema di Cauchy (12.11) è

$$u(t) = e^{(t-t_0)A}y_0.$$

(ii) La matrice esponenziale $\exp((t-t_0)A)$ è la matrice risolvente $R(t, t_0)$ dell'equazione omogenea $u' = Au$,

$$e^{(t-t_0)A} = R(t, t_0), \quad \forall t, t_0 \in \mathbb{R}.$$

(iii) La matrice esponenziale soddisfa

$$\frac{d}{dt} e^{(t-t_0)A} = A e^{(t-t_0)A} = e^{(t-t_0)A} A,$$

ed è l'unica soluzione del problema di Cauchy $M'(t) = AM(t)$, $M(t_0) = I$.

Dimostrazione. (i) Conseguenza della formula (12.12) e della definizione di esponenziale di matrice.

(ii) Conseguenza di (i) e del Lemma 12.8.

(iii) La prima identità è conseguenza di (ii) e dell'Osservazione 12.7; la seconda segue dal fatto che per ogni n si ha

$$A \left(\sum_{k=0}^n \frac{1}{k!} (t-t_0)^k A^k \right) = \sum_{k=0}^n \frac{1}{k!} (t-t_0)^k A^{k+1} = \left(\sum_{k=0}^n \frac{1}{k!} (t-t_0)^k A^k \right) A,$$

e, passando al limite per $n \rightarrow \infty$, si ha la tesi; l'ultima affermazione segue dall'Osservazione 12.7. \square

Vediamo ora come calcolare l'esponenziale di una matrice $A \in \text{Mat}_{N \times N}(\mathbb{R})$, partendo dai casi più semplici e arrivando al caso più generale.

Matrice diagonale. Il caso più semplice è quello in cui A è una matrice diagonale,

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}, \quad \lambda_1, \dots, \lambda_N \in \mathbb{R}. \quad (13.1)$$

In tal caso si prova per induzione che

$$A^n = \begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_N^n \end{pmatrix}, \quad n \in \mathbb{N},$$

e si calcola facilmente la somma della serie esponenziale,

$$e^{(t-t_0)A} = \sum_{n=0}^{\infty} \frac{1}{n!} (t-t_0)^n A^n = \begin{pmatrix} e^{(t-t_0)\lambda_1} & & \\ & \ddots & \\ & & e^{(t-t_0)\lambda_N} \end{pmatrix}. \quad (13.2)$$

Calcolo indiretto. Si può arrivare alla formula (13.2) anche in un modo più indiretto, risolvendo il problema di Cauchy invece che calcolando la somma della serie esponenziale: dette u_1, \dots, u_N le componenti di u , l'equazione omogenea $u' = Au$ è il sistema di N equazioni scalari *disaccoppiate*

$$\begin{cases} u'_1 = \lambda_1 u_1 \\ \dots \\ u'_N = \lambda_N u_N, \end{cases}$$

le cui soluzioni sono $u_k(t) = c_k e^{\lambda_k t}$, $c_k \in \mathbb{R}$. Dette y_{01}, \dots, y_{0N} le componenti del vettore $y_0 \in \mathbb{R}^N$, il problema di Cauchy

$$\begin{cases} u' = Au \\ u(t_0) = y_0 \end{cases}$$

corrisponde a N problemi di Cauchy scalari *disaccoppiati*,

$$\begin{cases} u'_1 = \lambda_1 u_1 \\ u_1(t_0) = y_{01} \end{cases} \quad \dots \quad \begin{cases} u'_N = \lambda_N u_N \\ u_N(t_0) = y_{0N} \end{cases}$$

la cui soluzione è

$$u_1(t) = e^{(t-t_0)\lambda_1} y_{01}, \quad \dots \quad u_N(t) = e^{(t-t_0)\lambda_N} y_{0N}.$$

D'altra parte, dal Lemma 12.8, $u(t) = R(t, t_0)y_0$, quindi abbiamo calcolato la matrice risolvete

$$R(t, t_0) = \begin{pmatrix} e^{(t-t_0)\lambda_1} & & \\ & \ddots & \\ & & e^{(t-t_0)\lambda_N} \end{pmatrix}; \quad (13.3)$$

infine, come osservato nel Lemma 13.1, la matrice risolvete coincide con l'esponenziale di matrice, che è dunque la matrice (13.3).

Matrice diagonalizzabile con autovalori reali. Il secondo caso più semplice è quello in cui ci possiamo ricondurre ad una matrice diagonale reale (come in (13.1)) tramite un isomorfismo (cambio di base) dello spazio vettoriale \mathbb{R}^N .

Supponiamo che A sia diagonalizzabile, con autovalori reali: questo significa che esiste una matrice invertibile $V \in \text{Mat}_{N \times N}(\mathbb{R})$ e una matrice diagonale $D \in \text{Mat}_{N \times N}(\mathbb{R})$,

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}, \quad \lambda_1, \dots, \lambda_N \in \mathbb{R}, \quad (13.4)$$

tali che

$$AV = VD. \quad (13.5)$$

Per proseguire con il calcolo della matrice esponenziale ci serve l'osservazione seguente.

Lemma 13.2. (Esponenziale e coniugato). *Siano $M, B \in \text{Mat}_{N \times N}(\mathbb{R})$, con M invertibile. Allora*

$$\exp(M^{-1}BM) = M^{-1} \exp(B)M.$$

Dimostrazione. Per ogni $n \in \mathbb{N}$ si ha

$$(M^{-1}BM)^n = M^{-1}B^nM. \quad (13.6)$$

L'identità (13.6) si prova facilmente per induzione; il punto è che tutti i fattori M, M^{-1} nel prodotto si cancellano eccetto il primo e l'ultimo:

$$\begin{aligned} (M^{-1}BM)^n &= (M^{-1}BM)(M^{-1}BM) \cdots (M^{-1}BM) \\ &= M^{-1}B(MM^{-1})B(MM^{-1})B \cdots (MM^{-1})BM \\ &= M^{-1}B^nM. \end{aligned}$$

Quindi per ogni n si ha

$$\sum_{k=0}^n \frac{1}{k!} (M^{-1}BM)^k = M^{-1} \left(\sum_{k=0}^n \frac{1}{k!} B^k \right) M,$$

e passando al limite per $n \rightarrow \infty$ si ha la tesi. \square

Usando l'identità (13.5) e il Lemma 13.2, si ha

$$\begin{aligned} \exp((t-t_0)A) &= \exp((t-t_0)VDV^{-1}) \\ &= \exp(V(t-t_0)DV^{-1}) \\ &= V \exp((t-t_0)D)V^{-1}, \end{aligned}$$

e l'esponenziale $\exp((t-t_0)D)$ è la matrice diagonale $\text{diag}(e^{(t-t_0)\lambda_1}, \dots, e^{(t-t_0)\lambda_N})$ che abbiamo calcolato in (13.2) e (13.3).

Calcolo indiretto. Anche in questo caso possiamo ottenere $\exp((t-t_0)A)$ calcolando la matrice risolvente dell'equazione $u' = Au$, invece che calcolare la somma della serie esponenziale. Risolviamo l'equazione differenziale per sostituzione: sia

$$u(t) = Vv(t).$$

Dalla (13.5), si ha

$$u'(t) = Vv'(t), \quad Au(t) = AVv(t) = VDv(t),$$

quindi, moltiplicando a sinistra per V^{-1} , l'equazione $u' = Au$ si riscrive

$$v' = Dv. \quad (13.7)$$

La condizione iniziale $u(t_0) = y_0$ diventa

$$v(t_0) = V^{-1}y_0. \quad (13.8)$$

La soluzione del problema di Cauchy (13.7)-(13.8) è

$$v(t) = e^{(t-t_0)D}V^{-1}y_0,$$

da cui

$$u(t) = Vv(t) = Ve^{(t-t_0)D}V^{-1}y_0.$$

Dal Lemma 12.8, questo significa che $V e^{(t-t_0)D} V^{-1}$ è la matrice risolvente dell'equazione $u' = Au$; come osservato nel Lemma 13.1, è anche la matrice esponenziale $\exp((t - t_0)A)$.

Matrice reale 2×2 con autovalori complessi coniugati. Un altro caso importante è quello di dimensione $N = 2$ in cui la matrice A è della forma

$$A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \quad \alpha, \beta \in \mathbb{R}, \quad (13.9)$$

con $\beta \neq 0$ (per $\beta = 0$ torniamo al caso, già trattato, di matrice reale diagonale). I suoi autovalori sono i numeri complessi coniugati $\alpha \pm i\beta$.

Indichiamo $(x(t), y(t))$ le componenti di $u(t)$; l'equazione $u' = Au$ è

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

cioè

$$\begin{cases} x' = \alpha x - \beta y \\ y' = \beta x + \alpha y. \end{cases} \quad (13.10)$$

Un primo modo di risolvere il sistema (13.10) è il seguente: $x' = \alpha x - \beta y$ è un'equazione lineare non omogenea nell'incognita x , con termine non omogeneo $-\beta y$; similmente, $y' = \alpha y + \beta x$ è un'equazione lineare non omogenea nell'incognita y , con termine non omogeneo βx . Possiamo dunque procedere in analogia a quanto visto per le equazioni (scalari) lineari non omogenee, seguendo l'idea della variazione delle costanti: consideriamo il cambio di variabile

$$x(t) = e^{\alpha t} \xi(t), \quad y(t) = e^{\alpha t} \eta(t),$$

cosicché il sistema (13.10) riscritto nelle nuove incognite $\xi(t), \eta(t)$ diventa

$$\begin{cases} \xi' = -\beta \eta \\ \eta' = \beta \xi \end{cases} \quad (13.11)$$

(verificare che, facendo la sostituzione, si ottiene (13.11)). Derivando la prima equazione di (13.11) e usando la seconda equazione per sostituire η' , troviamo

$$\xi'' + \beta^2 \xi = 0. \quad (13.12)$$

Abbiamo osservato nella Lezione 1 che $\sin(\beta t), \cos(\beta t)$ sono soluzioni dell'equazione (13.12); poi $\eta = -\xi'/\beta$, e dunque abbiamo due soluzioni linearmente indipendenti del sistema (13.11):

$$\begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} = \begin{pmatrix} \cos(\beta t) \\ \sin(\beta t) \end{pmatrix}, \quad \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} = \begin{pmatrix} \sin(\beta t) \\ -\cos(\beta t) \end{pmatrix}. \quad (13.13)$$

Sappiamo dal Teorema 11.9 che le soluzioni del sistema (13.11) formano uno spazio vettoriale su \mathbb{R} di dimensione 2, quindi le soluzioni di (13.11) sono tutte e sole le combinazioni lineari delle due soluzioni (13.13). Possiamo dunque tornare a x, y , e otteniamo tutte le soluzioni del sistema (13.10):

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 \begin{pmatrix} e^{\alpha t} \cos(\beta t) \\ e^{\alpha t} \sin(\beta t) \end{pmatrix} + c_2 \begin{pmatrix} e^{\alpha t} \sin(\beta t) \\ -e^{\alpha t} \cos(\beta t) \end{pmatrix}, \quad c_1, c_2 \in \mathbb{R}. \quad (13.14)$$

Un altro modo, istruttivo e più rapido, di risolvere il sistema (13.10) utilizza i numeri complessi. Sia

$$z(t) := x(t) + iy(t), \quad \gamma := \alpha + i\beta.$$

Si ha dunque

$$z' = x' + iy' = (\alpha x - \beta y) + i(\beta x + \alpha y) = (\alpha + i\beta)(x + iy) = \gamma z,$$

e il sistema reale (13.10) diventa la singola equazione complessa

$$z'(t) = \gamma z(t). \quad (13.15)$$

L'esponenziale complesso $e^{\gamma t}$ soddisfa

$$\frac{d}{dt} e^{\gamma t} = \gamma e^{\gamma t} \quad \forall t \in \mathbb{R}, \quad \gamma \in \mathbb{C}. \quad (13.16)$$

L'identità (13.16) si può dimostrare a partire dalla serie che definisce l'esponenziale complesso, $e^z := \sum_{n=0}^{\infty} \frac{1}{n!} z^n$, oppure, più semplicemente, ricordando che

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

e osservando che

$$\begin{aligned} \frac{d}{dt} e^{\gamma t} &= \frac{d}{dt} \{e^{\alpha t}(\cos(\beta t) + i \sin(\beta t))\} \\ &= \alpha e^{\alpha t}(\cos(\beta t) + i \sin(\beta t)) + e^{\alpha t}(-\beta \sin(\beta t) + i\beta \cos(\beta t)) \\ &= (\alpha + i\beta)e^{\alpha t}(\cos(\beta t) + i \sin(\beta t)) \\ &= \gamma e^{\gamma t}. \end{aligned}$$

Dunque le funzioni

$$z(t) = ce^{\gamma t}, \quad c \in \mathbb{C}, \quad (13.17)$$

sono soluzioni dell'equazione (13.15). Torniamo a scrivere tutto con parte reale e immaginaria: $z(t) = x(t) + iy(t)$, $\gamma = \alpha + i\beta$, $e^{\gamma t} = e^{\alpha t}(\cos(\beta t) + i \sin(\beta t))$, $c = a + ib$, e otteniamo la stessa formula (13.14) per le soluzioni reali del sistema (13.10). Abbiamo usato i numeri complessi per fare i calcoli, e siamo arrivati alla formula (13.14) in cui tutto è reale.

La soluzione del problema di Cauchy di equazione (13.10) e dato iniziale

$$\begin{pmatrix} x(t_0) \\ y(t_0) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \quad (13.18)$$

si ottiene imponendo la condizione iniziale (13.18) nella formula (13.14) delle soluzioni: si calcola

$$\begin{aligned} c_1 &= e^{-\alpha t_0} \cos(\beta t_0)x_0 + e^{-\alpha t_0} \sin(\beta t_0)y_0, \\ c_2 &= e^{-\alpha t_0} \sin(\beta t_0)x_0 - e^{-\alpha t_0} \cos(\beta t_0)y_0, \end{aligned}$$

e dunque

$$\begin{aligned}
 x(t) &= e^{\alpha(t-t_0)} [\cos(\beta t) \cos(\beta t_0) + \sin(\beta t) \sin(\beta t_0)] x_0 \\
 &\quad + e^{\alpha(t-t_0)} [\cos(\beta t) \sin(\beta t_0) - \sin(\beta t) \cos(\beta t_0)] y_0 \\
 &= e^{\alpha(t-t_0)} \cos(\beta(t-t_0)) x_0 - e^{\alpha(t-t_0)} \sin(\beta(t-t_0)) y_0, \\
 y(t) &= e^{\alpha(t-t_0)} [\sin(\beta t) \cos(\beta t_0) - \cos(\beta t) \sin(\beta t_0)] x_0 \\
 &\quad + e^{\alpha(t-t_0)} [\sin(\beta t) \sin(\beta t_0) + \cos(\beta t) \cos(\beta t_0)] y_0 \\
 &= e^{\alpha(t-t_0)} \sin(\beta(t-t_0)) x_0 + e^{\alpha(t-t_0)} \cos(\beta(t-t_0)) y_0,
 \end{aligned}$$

cioè

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} e^{\alpha(t-t_0)} \cos(\beta(t-t_0)) & -e^{\alpha(t-t_0)} \sin(\beta(t-t_0)) \\ e^{\alpha(t-t_0)} \sin(\beta(t-t_0)) & e^{\alpha(t-t_0)} \cos(\beta(t-t_0)) \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}. \quad (13.19)$$

Abbiamo dunque calcolato la matrice risolvente, che è anche la matrice esponenziale:

$$e^{(t-t_0)A} = \begin{pmatrix} e^{\alpha(t-t_0)} \cos(\beta(t-t_0)) & -e^{\alpha(t-t_0)} \sin(\beta(t-t_0)) \\ e^{\alpha(t-t_0)} \sin(\beta(t-t_0)) & e^{\alpha(t-t_0)} \cos(\beta(t-t_0)) \end{pmatrix}. \quad (13.20)$$

I calcoli che portano alla (13.20) diventano più semplici usando i numeri complessi (si usano esponenziali e scalari invece di seni, coseni e matrici): ritorniamo alla formula (13.17), cioè $z(t) = ce^{\gamma t}$, delle soluzioni complesse, e imponiamo la condizione iniziale complessa

$$z(t_0) = z_0 \quad (13.21)$$

dove $z_0 := x_0 + iy_0$. Troviamo

$$c = e^{-\gamma t_0} z_0,$$

da cui otteniamo

$$z(t) = e^{\gamma(t-t_0)} z_0. \quad (13.22)$$

Abbiamo calcolato la soluzione complessa $z(t)$ del problema di Cauchy (scalare, complesso) (13.15)-(13.21). A questo punto possiamo scrivere $z(t)$, γ , z_0 come $x(t) + iy(t)$, $\alpha + i\beta$, $x_0 + iy_0$ e, uguagliando parte reale e parte immaginaria nella formula (13.22), otteniamo le soluzioni reali $x(t), y(t)$ in (13.19), e quindi otteniamo la matrice esponenziale (13.20).

Notiamo anche che, rimanendo in \mathbb{C} , da (13.22) si vede che la matrice risolvente è lo scalare complesso $e^{\gamma(t-t_0)}$; se consideriamo la mappa di moltiplicazione complessa

$$\mathbb{C} \rightarrow \mathbb{C}, \quad z_0 \mapsto e^{\gamma(t-t_0)} z_0,$$

la matrice che la rappresenta come applicazione da \mathbb{R}^2 a \mathbb{R}^2 rispetto a parte reale e parte immaginaria è proprio la matrice (13.20); la matrice A in (13.9) rappresenta la moltiplicazione per il numero $\gamma = \alpha + i\beta$.

Blocco di Jordan di autovalore reale. Consideriamo $A \in \text{Mat}_{N \times N}(\mathbb{R})$ della forma

$$A = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}, \quad \lambda \in \mathbb{R}. \quad (13.23)$$

La matrice A è la somma

$$A = \lambda I + Z$$

dove $I \in \text{Mat}_{N \times N}(\mathbb{R})$ è la matrice identità e Z è la matrice che ha 1 sopra la diagonale principale e 0 altrove, cioè nel posto (j, k) (riga j , colonna k) si ha

$$Z_{jk} = \begin{cases} 1 & \text{se } k = j + 1, \\ 0 & \text{se } k \neq j + 1, \end{cases} \quad j, k = 1, \dots, N.$$

Dunque

$$(Z^2)_{jk} = \sum_{\ell=1}^N Z_{j\ell} Z_{\ell k} = \begin{cases} 1 & \text{se } k = j + 2, \\ 0 & \text{se } k \neq j + 2, \end{cases} \quad j, k = 1, \dots, N, \quad (13.24)$$

e, per induzione,

$$(Z^p)_{jk} = \sum_{\ell=1}^N (Z^{p-1})_{j\ell} Z_{\ell k} = \begin{cases} 1 & \text{se } k = j + p, \\ 0 & \text{se } k \neq j + p, \end{cases} \quad j, k = 1, \dots, N. \quad (13.25)$$

Per $p = N - 1$ si ha la matrice Z^{N-1} che ha 1 al posto $(1, N)$ (cioè nell'angolo in alto a destra) e 0 altrove; per $p = N$ si ottiene la matrice nulla, $Z^N = 0$, perché $k < j + N$ per ogni j, k . La matrice Z è dunque *nilpotente* (una sua potenza è la matrice nulla).

L'esponenziale $e^{(t-t_0)A}$ si ottiene facilmente usando la seguente proprietà degli esponenziali di matrice.

Lemma 13.3. (Esponenziale della somma). *Siano $B, C \in \text{Mat}_{N \times N}(\mathbb{R})$. Se le matrici B e C commutano, cioè se*

$$BC = CB,$$

allora il prodotto (righe per colonne) delle matrici esponenziali soddisfa

$$e^B e^C = e^C e^B = e^{B+C}.$$

Vediamo due dimostrazioni.

Prima dimostrazione del Lemma 13.3. Poiché $BC = CB$, per ogni $t \in \mathbb{R}$, per ogni $n \in \mathbb{N}$ si ha

$$\sum_{k=0}^n \frac{1}{k!} t^k B^k C = C \sum_{k=0}^n \frac{1}{k!} t^k B^k,$$

da cui, passando al limite,

$$e^{tB} C = C e^{tB}. \quad (13.26)$$

Dal Lemma 13.1, $e^{t(B+C)}$ è l'unica soluzione del problema di Cauchy (matriciale)

$$\begin{cases} M'(t) = (B + C)M(t), \\ M(0) = I. \end{cases} \quad (13.27)$$

Dalla (13.26), il prodotto di matrici $P(t) := e^{tB}e^{tC}$ soddisfa

$$\begin{aligned} P'(t) &= (e^{tB})'e^{tC} + e^{tB}(e^{tC})' \\ &= Be^{tB}e^{tC} + e^{tB}Ce^{tC} \\ &= Be^{tB}e^{tC} + Ce^{tB}e^{tC} = (B + C)P(t), \end{aligned}$$

e $P(0) = I$. Quindi $P(t)$ è soluzione di (13.27), da cui $e^{t(B+C)} = e^{tB}e^{tC}$ per ogni t . In particolare, per $t = 1$ si ha $e^{B+C} = e^B e^C$. Infine $e^{B+C} = e^{C+B} = e^C e^B$. \square

Seconda dimostrazione del Lemma 13.3. Poiché $BC = CB$, le potenze della matrice $B + C$ soddisfano la classica formula delle potenze del binomio per gli scalari:

$$(B + C)(B + C) = B^2 + BC + CB + C^2 = B^2 + 2BC + C^2,$$

e, per induzione,

$$(B + C)^\ell = \sum_{k=0}^{\ell} \binom{\ell}{k} B^k C^{\ell-k}. \quad (13.28)$$

Per ogni n , riunendo in un'unica parentesi tutti i termini $B^k C^j$ con lo stesso grado $k + j = \ell$ di omogeneità, e ordinando per ℓ crescente, si ha

$$\begin{aligned} &\left(\sum_{k=0}^n \frac{1}{k!} B^k\right) \left(\sum_{j=0}^n \frac{1}{j!} C^j\right) \\ &= \left(I + B + \frac{1}{2}B^2 + \dots + \frac{1}{n!}B^n\right) \left(I + C + \frac{1}{2}C^2 + \dots + \frac{1}{n!}C^n\right) \\ &= I + (C + B) + \left(\frac{1}{2}C^2 + BC + \frac{1}{2}B^2\right) + \left(\frac{1}{3!}C^3 + \frac{1}{2}BC^2 + \frac{1}{2}B^2C + \frac{1}{3!}B^3\right) \\ &\quad + \dots + \left(\frac{1}{n!}C^n + \frac{1}{(n-1)!}BC^{n-1} + \dots + \frac{1}{(n-1)!}B^{n-1}C + \frac{1}{n!}B^n\right) \\ &\quad + \left(\frac{1}{n!}BC^n + \frac{1}{2!(n-1)!}B^2C^{n-1} + \dots + \frac{1}{(n-1)!2!}B^{n-1}C^2 + \frac{1}{n!}B^nC\right) \\ &\quad + \left(\frac{1}{2!n!}B^2C^n + \frac{1}{3!(n-1)!}B^3C^{n-1} + \dots + \frac{1}{(n-1)!3!}B^{n-1}C^3 + \frac{1}{n!2!}B^nC^2\right) \\ &\quad + \dots + \left(\frac{1}{(n-1)!n!}B^{n-1}C^n + \frac{1}{n!(n-1)!}B^nC^{n-1}\right) + \frac{1}{n!n!}B^nC^n, \end{aligned}$$

cioè

$$\begin{aligned} &\left(\sum_{k=0}^n \frac{1}{k!} B^k\right) \left(\sum_{j=0}^n \frac{1}{j!} C^j\right) \\ &= \underbrace{\sum_{\ell=0}^n \left(\sum_{k=0}^{\ell} \frac{1}{k!(\ell-k)!} B^k C^{\ell-k}\right)}_{S_n} + \underbrace{\sum_{\ell=n+1}^{2n} \left(\sum_{k=\ell-n}^n \frac{1}{k!(\ell-k)!} B^k C^{\ell-k}\right)}_{r_n}. \quad (13.29) \end{aligned}$$

Usando la formula (13.28), la somma S_n in (13.29) è

$$S_n = \sum_{\ell=0}^n \frac{1}{\ell!} \left(\sum_{k=0}^{\ell} \frac{\ell!}{k!(\ell-k)!} B^k C^{\ell-k}\right) = \sum_{\ell=0}^n \frac{1}{\ell!} (B + C)^\ell.$$

Usando il Lemma 12.15 e il fatto che, per $\ell = n + 1, \dots, 2n$, si ha $[\ell - n, n] \subset [0, \ell]$, la norma (euclidea o operatoriale) $|r_n|$ della somma r_n in (13.29) si stima con

$$\begin{aligned} |r_n| &\leq \sum_{\ell=n+1}^{2n} \sum_{k=\ell-n}^n \frac{1}{k!(\ell-k)!} |B^k C^{\ell-k}| \\ &\leq \sum_{\ell=n+1}^{2n} \frac{1}{\ell!} \sum_{k=\ell-n}^n \frac{\ell!}{k!(\ell-k)!} |B|^k |C|^{\ell-k} \\ &\leq \sum_{\ell=n+1}^{2n} \frac{1}{\ell!} \sum_{k=0}^{\ell} \frac{\ell!}{k!(\ell-k)!} |B|^k |C|^{\ell-k} \\ &= \sum_{\ell=n+1}^{2n} \frac{1}{\ell!} (|B| + |C|)^{\ell} \leq \sum_{\ell=n+1}^{\infty} \frac{1}{\ell!} (|B| + |C|)^{\ell}, \end{aligned}$$

e l'ultima serie tende a zero per $n \rightarrow \infty$, essendo la differenza tra la serie esponenziale $e^{|B|+|C|} = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} (|B| + |C|)^{\ell}$ e la sua somma parziale $\sum_{\ell=0}^n \frac{1}{\ell!} (|B| + |C|)^{\ell}$. Dunque la (13.29) è

$$\left(\sum_{k=0}^n \frac{1}{k!} B^k \right) \left(\sum_{j=0}^n \frac{1}{j!} C^j \right) = \sum_{\ell=0}^n \frac{1}{\ell!} (B + C)^{\ell} + r_n,$$

con $r_n \rightarrow 0$; passando al limite per $n \rightarrow \infty$ otteniamo $e^B e^C = e^{B+C}$. \square

Torniamo al calcolo dell'esponenziale della matrice $A = \lambda I + Z$ in (13.23). Poiché I e Z commutano (e quindi anche $(t - t_0)\lambda I$ e $(t - t_0)Z$ commutano), dal Lemma 13.3 si ha

$$e^{(t-t_0)A} = e^{(t-t_0)\lambda I} e^{(t-t_0)Z}.$$

La matrice $(t - t_0)\lambda I$ è diagonale, e il suo esponenziale è la matrice diagonale con $e^{(t-t_0)\lambda}$ sulla diagonale (vedi (13.2)). La matrice Z è nilpotente, con $Z^N = 0$, quindi la sua serie esponenziale è la somma finita

$$e^{(t-t_0)Z} = \sum_{n=0}^{\infty} \frac{1}{n!} (t - t_0)^n Z^n = \sum_{n=0}^{N-1} \frac{1}{n!} (t - t_0)^n Z^n,$$

e, ricordando (13.25), si ha

$$e^{(t-t_0)Z} = \begin{pmatrix} 1 & (t-t_0) & \frac{(t-t_0)^2}{2} & \cdots & \frac{(t-t_0)^{N-1}}{(N-1)!} \\ & 1 & (t-t_0) & \ddots & \vdots \\ & & 1 & \ddots & \frac{(t-t_0)^2}{2} \\ & & & \ddots & (t-t_0) \\ & & & & 1 \end{pmatrix}. \quad (13.30)$$

Dunque (scriviamo s al posto di $t - t_0$)

$$e^{sA} = \begin{pmatrix} e^{\lambda s} & s e^{\lambda s} & \frac{s^2}{2} e^{\lambda s} & \cdots & \frac{s^{N-1}}{(N-1)!} e^{\lambda s} \\ & e^{\lambda s} & s e^{\lambda s} & \ddots & \vdots \\ & & e^{\lambda s} & \ddots & \frac{s^2}{2} e^{\lambda s} \\ & & & \ddots & s e^{\lambda s} \\ & & & & e^{\lambda s} \end{pmatrix}. \quad (13.31)$$

Blocco di Jordan reale con coppia di autovalori complessi coniugati.
 Consideriamo $A \in \text{Mat}_{N \times N}(\mathbb{R})$ della forma

$$A = \begin{pmatrix} B & I_2 & & & \\ & B & I_2 & & \\ & & \ddots & \ddots & \\ & & & B & I_2 \\ & & & & B \end{pmatrix} \quad (13.32)$$

dove B, I_2 sono i blocchi 2×2

$$B = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \quad \alpha, \beta \in \mathbb{R}, \quad \beta \neq 0; \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Come sopra, scriviamo A come somma di una parte diagonale (qui diagonale a blocchi) e una parte nilpotente:

$$A = S + H$$

dove

$$S = \begin{pmatrix} B & & & & \\ & B & & & \\ & & \ddots & & \\ & & & B & \\ & & & & B \end{pmatrix}, \quad H = \begin{pmatrix} 0 & I_2 & & & \\ & 0 & \ddots & & \\ & & \ddots & I_2 & \\ & & & & 0 \end{pmatrix};$$

H nel posto (j, k) (riga j , colonna k) è

$$H_{jk} = \begin{cases} 1 & \text{se } k = j + 2, \\ 0 & \text{se } k \neq j + 2, \end{cases} \quad j, k = 1, \dots, N$$

(cioè $H = Z^2$ in (13.24)). Poiché $(t - t_0)S$ e $(t - t_0)H$ commutano, dal Lemma 13.3 si ha

$$e^{(t-t_0)A} = e^{(t-t_0)S} e^{(t-t_0)H}.$$

Le potenze di H soddisfano

$$(H^p)_{jk} = \begin{cases} 1 & \text{se } k = j + 2p, \\ 0 & \text{se } k \neq j + 2p, \end{cases} \quad j, k = 1, \dots, N;$$

in particolare, $H^p = 0$ per $p \geq N/2$ (per come è fatta la matrice A , la sua dimensione N deve essere pari). Quindi

$$e^{(t-t_0)H} = \sum_{n=0}^{\infty} \frac{1}{n!} (t - t_0)^n H^n = \sum_{n=0}^{(N/2)-1} \frac{1}{n!} (t - t_0)^n H^n,$$

cioè (scrivendo s invece di $(t - t_0)$ e m invece di $(N/2) - 1$)

$$e^{sH} = \begin{pmatrix} I_2 & sI_2 & \frac{s^2}{2}I_2 & \dots & \frac{s^m}{m!}I_2 \\ & I_2 & sI_2 & \ddots & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & I_2 & sI_2 \\ & & & & I_2 \end{pmatrix},$$

cioè

$$e^{sH} = \begin{pmatrix} 1 & 0 & s & 0 & \frac{s^2}{2} & 0 & & & \frac{s^m}{m!} & 0 \\ 0 & 1 & 0 & s & 0 & \frac{s^2}{2} & & \ddots & 0 & \frac{s^m}{m!} \\ & & 1 & 0 & s & 0 & \frac{s^2}{2} & 0 & & \\ & & 0 & 1 & 0 & s & 0 & \frac{s^2}{2} & & \\ & & & & 1 & 0 & s & 0 & & \\ & & & & 0 & 1 & 0 & s & & \\ & & & & & & \ddots & \ddots & & \\ & & & & & & & & s & 0 \\ & & & & & & & & 0 & s \\ & & & & & & & & 1 & 0 \\ & & & & & & & & 0 & 1 \end{pmatrix}.$$

Poi

$$e^{(t-t_0)S} = \sum_{n=0}^{\infty} \frac{1}{n!} S^n = \begin{pmatrix} e^{(t-t_0)B} & & & \\ & e^{(t-t_0)B} & & \\ & & \ddots & \\ & & & e^{(t-t_0)B} \end{pmatrix},$$

e la matrice $e^{(t-t_0)B}$ (che è un blocco 2×2) è stata calcolata in (13.20). Quindi

$$e^{tA} = \begin{pmatrix} e^{\alpha t} \cos(\beta t) & -e^{\alpha t} \sin(\beta t) & te^{\alpha t} \cos(\beta t) & -te^{\alpha t} \sin(\beta t) & \dots \\ e^{\alpha t} \sin(\beta t) & e^{\alpha t} \cos(\beta t) & te^{\alpha t} \sin(\beta t) & te^{\alpha t} \cos(\beta t) & \dots \\ & & e^{\alpha t} \cos(\beta t) & -e^{\alpha t} \sin(\beta t) & \ddots \\ & & e^{\alpha t} \sin(\beta t) & e^{\alpha t} \cos(\beta t) & \ddots \\ & & & & \ddots \\ & & & & \ddots \end{pmatrix}. \quad (13.33)$$

Caso generale. Per concludere, trattiamo il caso di una matrice reale A qualunque. Ricordiamo il Teorema di algebra lineare di riduzione alla forma canonica reale di Jordan.

Teorema 13.4. (Forma canonica reale di Jordan). *Sia $A \in \text{Mat}_{N \times N}(\mathbb{R})$ una matrice reale. Consideriamo il suo polinomio caratteristico*

$$p(\lambda) = \det(\lambda I - A),$$

che è un polinomio di grado N a coefficienti reali. Supponiamo che le sue radici siano i numeri reali λ_k , $k = 1, \dots, r$, e le coppie di numeri complessi coniugati $\alpha_k \pm i\beta_k$, $k = 1, \dots, q$, con α_k, β_k reali e $\beta_k \neq 0$. Supponiamo che λ_k abbia molteplicità algebrica f_k , e che $\alpha_k + i\beta_k$, $\alpha_k - i\beta_k$ abbiano ciascuno molteplicità algebrica g_k ; si ha dunque

$$f_1 + \dots + f_r + 2g_1 + \dots + 2g_q = N.$$

Allora esiste una matrice reale invertibile $V \in \text{Mat}_{N \times N}(\mathbb{R})$ che coniuga A ad una matrice reale a blocchi $E \in \text{Mat}_{N \times N}(\mathbb{R})$,

$$A = VEV^{-1},$$

che ha questa struttura:

$$E = \begin{pmatrix} F_1 & & & & & \\ & \ddots & & & & \\ & & F_r & & & \\ & & & G_1 & & \\ & & & & \ddots & \\ & & & & & G_q \end{pmatrix}$$

dove F_1, \dots, F_r sono matrici di Jordan con autovalore reale,

$$F_k = \begin{pmatrix} \lambda_k & 1 & & & \\ & \lambda_k & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_k & 1 \\ & & & & \lambda_k \end{pmatrix}, \quad \dim F_k = f_k, \quad k = 1, \dots, r,$$

e G_1, \dots, G_q sono matrici di Jordan con coppia di autovalori complessi coniugati,

$$G_k = \begin{pmatrix} B_k & I_2 & & & \\ & B_k & I_2 & & \\ & & \ddots & \ddots & \\ & & & B_k & I_2 \\ & & & & B_k \end{pmatrix}, \quad \dim G_k = 2g_k, \quad k = 1, \dots, q,$$

con

$$B_k = \begin{pmatrix} \alpha_k & -\beta_k \\ \beta_k & \alpha_k \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A questo punto possiamo calcolare l'esponenziale $e^{(t-t_0)A}$: per coniugazione con V , ricordando il Lemma 13.2, si ha

$$e^{(t-t_0)A} = V e^{(t-t_0)E} V^{-1}.$$

Per induzione, le potenze di E rimangono a blocchi disaccoppiati, cioè

$$E^n = \begin{pmatrix} F_1^n & & & & & \\ & \ddots & & & & \\ & & F_r^n & & & \\ & & & G_1^n & & \\ & & & & \ddots & \\ & & & & & G_q^n \end{pmatrix}, \quad n \in \mathbb{N},$$

da cui segue che

$$\begin{aligned} e^{(t-t_0)E} &= \sum_{n=0}^{\infty} \frac{1}{n!} (t-t_0)^n E^n \\ &= \begin{pmatrix} e^{(t-t_0)F_1} & & & & & \\ & \ddots & & & & \\ & & e^{(t-t_0)F_r} & & & \\ & & & e^{(t-t_0)G_1} & & \\ & & & & \ddots & \\ & & & & & e^{(t-t_0)G_q} \end{pmatrix}. \end{aligned} \quad (13.34)$$

Infine l'esponenziale di ogni singolo blocco di Jordan (ad autovalore reale o complesso) è stato già trattato nelle pagine precedenti.

Esercizio 13.5. Calcolare e^{tA} , dove $A \in \text{Mat}_{N \times N}$ è la matrice

$$A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ & & \dots & \\ 1 & 1 & \dots & 1 \end{pmatrix}.$$

14 Lezione 14

Equazioni lineari a coefficienti costanti non omogenee. Come commento conclusivo di questa parte sulle matrici costanti, ci limitiamo ad osservare che la formula (12.10) di variazione delle costanti assume una forma particolarmente semplice e utile nel caso in cui A sia matrice costante:

Proposizione 14.1. (Soluzione dell'equazione lineare non omogenea a coefficienti costanti tramite variazione delle costanti). *Sia $A \in \text{Mat}_{N \times N}(\mathbb{R})$ una matrice costante, $I \subseteq \mathbb{R}$ un intervallo, $t_0 \in I$, $b : \mathbb{R} \rightarrow \mathbb{R}^N$ una funzione continua. La soluzione del problema di Cauchy*

$$\begin{cases} u' = Au + b(t) \\ u(t_0) = y_0 \end{cases}$$

è la funzione

$$u(t) = e^{(t-t_0)A}y_0 + \int_{t_0}^t e^{(t-s)A}b(s) ds, \quad t \in I. \quad (14.1)$$

Dimostrazione. Come osservato nel Lemma 13.1, la matrice esponenziale $e^{(t-t_0)A}$ è la matrice risolvante $R(t, t_0)$ dell'equazione omogenea $u' = Au$; applicare questa osservazione nella formula (12.10). \square

Osservazione 14.2. (*Operazioni sul sistema di equazioni*). Per calcolare le soluzioni di un sistema lineare $u' = Au$, invece che lavorare sulla matrice A per portarla in forma di Jordan, si può anche lavorare direttamente sul sistema di equazioni con le usuali operazioni di algebra lineare (combinazioni lineari delle equazioni stesse, per esempio “la prima equazione meno la seconda equazione”, eccetera). Vedi per esempio il prossimo esercizio (Esercizio 14.3).

Esercizio 14.3. Calcolare le soluzioni del sistema lineare

$$\begin{cases} x' = 4x - y \\ y' = 9x - 2y \end{cases}$$

e scrivere la matrice risolvante.

Svolgimento. Chiaramente si può seguire il metodo generale illustrato sopra, calcolando V che coniuga la matrice

$$A = \begin{pmatrix} 4 & -1 \\ 9 & -2 \end{pmatrix}$$

alla sua forma di Jordan, eccetera; vediamo invece come risolvere il sistema lavorando direttamente sulle equazioni stesse.

Consideriamo la somma della prima equazione più un multiplo della seconda,

$$x' + \alpha y' = (4x - y) + \alpha(9x - 2y),$$

con α fattore qualunque, che sceglieremo nel modo più conveniente. Si ha dunque

$$\begin{aligned}(x + \alpha y)' &= 4x - y + 9\alpha x - 2\alpha y \\ &= (4 + 9\alpha)x - y - 2\alpha y \\ &= (4 + 9\alpha)(x + \alpha y) - (4 + 9\alpha)\alpha y - y - 2\alpha y \\ &= (4 + 9\alpha)(x + \alpha y) - (9\alpha^2 + 6\alpha + 1)y.\end{aligned}\tag{14.2}$$

Abbiamo aggiunto e sottratto il termine $(4+9\alpha)\alpha y$ in modo da scrivere un'equazione differenziale nell'incognita $x + \alpha y$ più un resto dipendente da α ; ora scegliamo α in modo che questo resto sia nullo, cioè

$$9\alpha^2 + 6\alpha + 1 = 0.$$

Fissiamo dunque $\alpha = -1/3$, sostituiamo in (14.2) e otteniamo l'equazione

$$\left(x - \frac{1}{3}y\right)' = x - \frac{1}{3}y,$$

le cui soluzioni sono

$$x(t) - \frac{1}{3}y(t) = c_1 e^t, \quad c_1 \in \mathbb{R}.$$

Torniamo al sistema di partenza, scrivendo $\frac{1}{3}y + c_1 e^t$ al posto di x : otteniamo l'equazione

$$y' = y + 9c_1 e^t.$$

La soluzione y si calcola tramite la formula risolutiva delle equazioni scalari lineari non omogenee (4.7), oppure ripercorrendo il metodo di variazioni delle costanti: cerchiamo y della forma

$$y(t) = e^t h(t),$$

cosicché

$$y' = e^t h + e^t h' = y + e^t h';$$

cerchiamo $h(t)$ tale che $e^t h'(t) = 9c_1 e^t$, cioè $h'(t) = 9c_1$, cioè

$$h(t) = 9c_1 t + c_2.$$

Otteniamo dunque

$$y(t) = (9c_1 t + c_2)e^t.$$

Ricaviamo anche x :

$$x(t) = \frac{1}{3}y(t) + c_1 e^t = \left(3c_1 t + \frac{c_2}{3} + c_1\right)e^t.$$

Rinominiamo, per comodità di scrittura, le costanti: $c_1 = \lambda$, $c_2/3 = \mu$. Abbiamo trovato le soluzioni

$$\begin{cases} x(t) = (3\lambda t + \lambda + \mu)e^t, \\ y(t) = (9\lambda t + 3\mu)e^t. \end{cases}\tag{14.3}$$

Possiamo subito verificare se abbiamo fatto bene i calcoli: partiamo dalle funzioni x, y in (14.3), calcoliamo $x', 4x - y, y', 9x - 2y$ e verifichiamo che effettivamente $x' = 4x - y$ e $y' = 9x - 2y$.

Calcoliamo ora la matrice risolvente. La soluzione del problema di Cauchy con istante iniziale $t_0 = 0$ e dati iniziali

$$x(0) = x_0, \quad y(0) = y_0$$

è quella che soddisfa

$$\begin{cases} x(0) = \lambda + \mu = x_0, \\ y(0) = 3\mu = y_0, \end{cases}$$

cioè

$$\lambda = x_0 - \frac{y_0}{3}, \quad \mu = \frac{y_0}{3},$$

cioè, sostituendo λ, μ ,

$$\begin{cases} x(t) = (3x_0 - y_0)te^t + x_0e^t, \\ y(t) = (9x_0 - 3y_0)te^t + y_0e^t. \end{cases} \quad (14.4)$$

Riscriviamo le soluzioni (14.4) come prodotto di una matrice per il vettore dei dati iniziali,

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} (3t+1)e^t & -te^t \\ 9te^t & (1-3t)e^t \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Dunque la matrice risolvente, che è anche la matrice esponenziale di A , è

$$R(t, 0) = e^{tA} = \begin{pmatrix} (3t+1)e^t & -te^t \\ 9te^t & (1-3t)e^t \end{pmatrix}.$$

Per un qualsiasi istante iniziale $t_0 \in \mathbb{R}$, si ha $R(t, t_0) = \exp((t - t_0)A)$, quindi

$$R(t, t_0) = e^{(t-t_0)A} = \begin{pmatrix} (3(t-t_0)+1)e^{t-t_0} & -(t-t_0)e^{t-t_0} \\ 9(t-t_0)e^{t-t_0} & (1-3(t-t_0))e^{t-t_0} \end{pmatrix},$$

e l'esercizio è concluso. □

Equazioni scalari di ordine più alto. Abbiamo osservato nella Lezione 1 che ogni equazione scalare di ordine N si può scrivere come equazione vettoriale del primo ordine in dimensione N ; questo vale in particolare per le equazioni lineari.

Definizione 14.4. (*Equazioni lineari di ordine N*). Un'equazione (in forma normale) di ordine N nell'incognita $u(t)$ si dice *lineare* se è della forma

$$u^{(N)} + a_{N-1}(t)u^{(N-1)} + \dots + a_2(t)u'' + a_1(t)u' + a_0(t)u = f(t) \quad (14.5)$$

per certe funzioni continue $a_k(t), f(t)$. Si dice *omogenea* se $f = 0$, *non omogenea* altrimenti.

Osservazione 14.5. (*Equazione lineare scalare scritta come sistema lineare del primo ordine*). Se la funzione scalare $u(t)$ è una soluzione dell'equazione (14.5), allora la funzione vettoriale $U(t) = (u(t), u'(t), \dots, u^{(N-1)}(t))$ è una soluzione dell'equazione del primo ordine in \mathbb{R}^N

$$U' = A(t)U + b(t) \quad (14.6)$$

dove

$$A(t) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -a_0(t) & -a_1(t) & -a_2(t) & \dots & -a_{N-1}(t) \end{pmatrix}, \quad b(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(t) \end{pmatrix}. \quad (14.7)$$

Viceversa, se $U(t) = (u_1(t), \dots, u_N(t))$ è una soluzione di (14.6), allora $u(t) = u_1(t)$ è soluzione di (14.5).

Brevemente, diciamo che l'equazione (14.5), riscritta come sistema del primo ordine, diventa (14.6), o che l'equazione (14.5) è equivalente al sistema (14.6).

Nella corrispondenza tra (14.5) e (14.6), a volte può essere più conveniente indicare (u_0, \dots, u_{N-1}) le componenti di U , invece che (u_1, \dots, u_N) , cosicché (14.6) in componenti diventa

$$\begin{cases} u'_0 = u_1 \\ u'_1 = u_2 \\ \dots \\ u'_{N-2} = u_{N-1} \\ u'_{N-1} = -a_{N-1}(t)u_{N-1} - \dots - a_1(t)u_1 - a_0(t)u_0 + f(t), \end{cases}$$

e $u = u_0$, $u' = u_1$, $u'' = u_2$, eccetera.

Tutto ciò che abbiamo visto per i sistemi del primo ordine in \mathbb{R}^N vale per le equazioni scalari di ordine N tramite la corrispondenza dell'Osservazione 14.5. Ci limitiamo a riformulare i risultati principali.

Lemma 14.6. (Spazio vettoriale delle soluzioni). *Sia $I \subseteq \mathbb{R}$ un intervallo, e siano $a_0, \dots, a_{N-1} : I \rightarrow \mathbb{R}$ funzioni continue. Per ogni $u \in C^N(I, \mathbb{R})$, indichiamo, in breve,*

$$\mathcal{L}u(t) := u^{(N)}(t) + a_{N-1}(t)u^{(N-1)}(t) + \dots + a_1(t)u'(t) + a_0(t)u(t). \quad (14.8)$$

L'insieme

$$S_{\mathcal{L}} = \{u \in C^N(I, \mathbb{R}) : \mathcal{L}u(t) = 0 \quad \forall t \in I\} \quad (14.9)$$

delle soluzioni dell'equazione lineare omogenea $\mathcal{L}u = 0$ è uno spazio vettoriale su \mathbb{R} di dimensione N .

Dimostrazione. Conseguenza dell'Osservazione 14.5 e del Teorema 11.9. \square

Lemma 14.7. (Funzioni linearmente indipendenti). *Sia $I \subseteq \mathbb{R}$ un intervallo, e siano u_1, \dots, u_n funzioni scalari, derivabili k volte in I . Allora u_1, \dots, u_n (funzioni $I \rightarrow \mathbb{R}$) sono linearmente indipendenti se e solo se*

$$U_1 := \begin{pmatrix} u_1 \\ u'_1 \\ \vdots \\ u_1^{(k)} \end{pmatrix}, \quad \dots, \quad U_n := \begin{pmatrix} u_n \\ u'_n \\ \vdots \\ u_n^{(k)} \end{pmatrix}$$

(funzioni $I \rightarrow \mathbb{R}^{k+1}$) sono linearmente indipendenti.

Dimostrazione. Supponiamo che u_1, \dots, u_n siano linearmente indipendenti, e che $\lambda_1 U_1 + \dots + \lambda_n U_n$ sia identicamente nulla in I . Allora, prendendo la prima componente di ciascuna U_j , si ha $\lambda_1 u_1 + \dots + \lambda_n u_n = 0$ in I , e quindi $\lambda_1, \dots, \lambda_n$ sono nulli. Questo prova che U_1, \dots, U_n sono linearmente indipendenti.

Viceversa, supponiamo che U_1, \dots, U_n siano linearmente indipendenti, e che $\lambda_1 u_1 + \dots + \lambda_n u_n$ sia identicamente nulla in I . Allora sono nulle anche tutte le sue derivate, $\lambda_1 u_1' + \dots + \lambda_n u_n' = 0$, ecc.: quindi $\lambda_1 U_1 + \dots + \lambda_n U_n = 0$ in I . Dunque $\lambda_1, \dots, \lambda_n$ sono tutti nulli. Questo prova che u_1, \dots, u_n sono linearmente indipendenti. \square

Lemma 14.8. (Soluzioni linearmente indipendenti). *Sia $I \subseteq \mathbb{R}$ un intervallo, e siano $a_0, \dots, a_{N-1} : I \rightarrow \mathbb{R}$ funzioni continue. Siano u_1, \dots, u_n soluzioni dell'equazione lineare omogenea $\mathcal{L}u = 0$ in I (\mathcal{L} definito in (14.8)). Siano*

$$U_1(t) := \begin{pmatrix} u_1(t) \\ u_1'(t) \\ \vdots \\ u_1^{(N-1)}(t) \end{pmatrix}, \quad \dots, \quad U_n(t) := \begin{pmatrix} u_n(t) \\ u_n'(t) \\ \vdots \\ u_n^{(N-1)}(t) \end{pmatrix}, \quad (14.10)$$

e sia $A(t)$ la matrice in (14.7). Sono equivalenti:

- (i) U_1, \dots, U_n sono soluzioni linearmente indipendenti dell'equazione $U' = A(t)U$;
- (ii) per ogni $t \in I$, $U_1(t), \dots, U_n(t)$ sono vettori linearmente indipendenti di \mathbb{R}^N ;
- (iii) per un certo $t_0 \in I$, $U_1(t_0), \dots, U_n(t_0)$ sono vettori linearmente indipendenti di \mathbb{R}^N ;
- (iv) u_1, \dots, u_n sono soluzioni linearmente indipendenti dell'equazione $\mathcal{L}u = 0$.

Dimostrazione. (i), (ii) e (iii) sono equivalenti per il Lemma 11.11; (i) e (iv) sono equivalenti per il Lemma 14.7 e l'Osservazione 14.5. \square

Definizione 14.9. (Matrice wronskiana, determinante wronskiano). *Sia $I \subseteq \mathbb{R}$ un intervallo, e siano $a_0, \dots, a_{N-1} : I \rightarrow \mathbb{R}$ funzioni continue. Siano u_1, \dots, u_N soluzioni dell'equazione lineare omogenea $\mathcal{L}u = 0$ in I (\mathcal{L} definito in (14.8)). La matrice wronskiana di u_1, \dots, u_N è la matrice*

$$M(t) = \begin{pmatrix} u_1(t) & u_2(t) & \dots & u_N(t) \\ u_1'(t) & u_2'(t) & \dots & u_N'(t) \\ u_1''(t) & u_2''(t) & \dots & u_N''(t) \\ \vdots & \vdots & \dots & \vdots \\ u_1^{(N-1)}(t) & u_2^{(N-1)}(t) & \dots & u_N^{(N-1)}(t) \end{pmatrix};$$

il determinante wronskiano è il suo determinante.

Lemma 14.10. (Teorema del wronskiano per l'equazione $\mathcal{L}u = 0$). *Sia $I \subseteq \mathbb{R}$ un intervallo, e siano $a_0, \dots, a_{N-1} : I \rightarrow \mathbb{R}$ funzioni continue. Siano u_1, \dots, u_N soluzioni dell'equazione lineare omogenea $\mathcal{L}u = 0$ in I (\mathcal{L} definito in (14.8)). Sia $M(t)$ la matrice wronskiana di u_1, \dots, u_N , e sia $W(t) = \det M(t)$ il determinante wronskiano.*

Se u_1, \dots, u_N sono linearmente indipendenti, allora $W(t) \neq 0$ per ogni $t \in I$; se invece u_1, \dots, u_N sono linearmente dipendenti, allora $W(t) = 0$ per ogni $t \in I$.

Dimostrazione. Conseguenza immediata del Lemma 14.8. □

Teorema 14.11. (Le soluzioni di un'equazione lineare non omogenea formano uno spazio affine). *Sia $I \subseteq \mathbb{R}$ un intervallo, e siano $a_0, \dots, a_{N-1}, f : I \rightarrow \mathbb{R}$ funzioni continue. Per ogni $u \in C^N(I, \mathbb{R})$, sia $\mathcal{L}u$ la funzione definita in (14.8). Sia $u_P \in C^N(I, \mathbb{R})$ una soluzione dell'equazione non omogenea*

$$\mathcal{L}u_P = f,$$

e sia $S_{\mathcal{L}}$ lo spazio vettoriale delle soluzioni dell'equazione omogenea $\mathcal{L}u = 0$ definito in (14.9).

Allora le soluzioni dell'equazione non omogenea $\mathcal{L}u = f$ sono le funzioni $u = u_P + v$ con v soluzione dell'equazione omogenea $\mathcal{L}v = 0$, cioè

$$\{u \in C^N(I, \mathbb{R}) : \mathcal{L}u = f \text{ in } I\} = \{u_P + v : v \in S_{\mathcal{L}}\}.$$

Dimostrazione. Sia u soluzione dell'equazione non omogenea, e sia $v := u - u_P$. Allora

$$\mathcal{L}v = \mathcal{L}(u - u_P) = \mathcal{L}u - \mathcal{L}u_P = f - f = 0,$$

quindi $v \in S_{\mathcal{L}}$, e $u = u_P + v$.

Viceversa, sia $v \in S_{\mathcal{L}}$ e sia $u = u_P + v$. Allora

$$\mathcal{L}u = \mathcal{L}(u_P + v) = \mathcal{L}u_P + \mathcal{L}v = f,$$

cioè u è soluzione dell'equazione non omogenea. □

Osservazione 14.12. (*Soluzione particolare*). Grazie al Teorema 14.11, per scrivere tutte le soluzioni dell'equazione non omogenea $\mathcal{L}u = f$ ci basta conoscere una soluzione dell'equazione non omogenea (spesso detta “soluzione particolare”) e tutte le soluzioni dell'omogenea associata.

Metodo di variazione delle costanti. Ripercorriamo il metodo di variazione delle costanti per risolvere l'equazione non omogenea $\mathcal{L}u = f$ di grado $N \geq 2$ quando si conoscono N soluzioni linearmente indipendenti dell'equazione omogenea associata.

Siano v_1, \dots, v_N soluzioni dell'equazione omogenea $\mathcal{L}v = 0$. Cerchiamo u soluzione di $\mathcal{L}u = f$ della forma

$$u(t) = h_1(t)v_1(t) + \dots + h_N(t)v_N(t) = \sum_{k=1}^N h_k(t)v_k(t),$$

con h_k da determinare. Calcoliamo la derivata prima

$$u' = \sum_{k=1}^N h'_k v_k + \sum_{k=1}^N h_k v'_k;$$

se h_1, \dots, h_N soddisfano

$$\sum_{k=1}^N h'_k v_k = 0$$

in I (è ciò che accadrebbe se h_1, \dots, h_N fossero delle costanti), allora rimane

$$u' = \sum_{k=1}^N h_k v_k'.$$

Di conseguenza, la derivata seconda è

$$u'' = \sum_{k=1}^N h_k' v_k' + \sum_{k=1}^N h_k v_k''.$$

Se h_1, \dots, h_k soddisfano anche

$$\sum_{k=1}^N h_k' v_k' = 0$$

(come farebbero delle costanti), allora rimane

$$u'' = \sum_{k=1}^N h_k v_k''.$$

Procediamo in modo simile: se h_1, \dots, h_N soddisfano

$$\sum_{k=1}^N h_k' v_k^{(n)} = 0 \quad \forall n = 0, \dots, N-2,$$

rimane

$$u^{(n)} = \sum_{k=1}^N h_k v_k^{(n)} \quad \forall n = 0, \dots, N-1.$$

Dunque

$$u^{(N)} = (u^{(N-1)})' = \left(\sum_{k=1}^N h_k v_k^{(N-1)} \right)' = \sum_{k=1}^N h_k' v_k^{(N-1)} + \sum_{k=1}^N h_k v_k^{(N)}.$$

Calcoliamo

$$\begin{aligned} \mathcal{L}u &= u^{(N)} + a_{N-1}u^{(N-1)} + \dots + a_1u' + a_0u \\ &= \sum_{k=1}^N h_k' v_k^{(N-1)} + \sum_{k=1}^N h_k v_k^{(N)} + a_{N-1} \sum_{k=1}^N h_k v_k^{(N-1)} + \dots + a_1 \sum_{k=1}^N h_k v_k' + a_0 \sum_{k=1}^N h_k v_k \\ &= \sum_{k=1}^N h_k' v_k^{(N-1)} + \sum_{k=1}^N h_k (v_k^{(N)} + a_{N-1}v_k^{(N-1)} + \dots + a_1v_k' + a_0v_k) \\ &= \sum_{k=1}^N h_k' v_k^{(N-1)} + \sum_{k=1}^N h_k \mathcal{L}v_k \\ &= \sum_{k=1}^N h_k' v_k^{(N)}. \end{aligned}$$

Se h_1, \dots, h_N soddisfano

$$\sum_{k=1}^N h_k' v_k^{(N)} = f,$$

allora $\mathcal{L}u = f$, che è ciò che volevamo.

Le richieste su h_1, \dots, h_N sono dunque:

$$\sum_{k=1}^N h'_k v_k^{(n)} = 0 \quad \forall n = 0, \dots, N-2; \quad \sum_{k=1}^N h'_k v_k^{(N-1)} = f,$$

cioè

$$\begin{pmatrix} v_1 & v_2 & \dots & v_N \\ v'_1 & v'_2 & \dots & v'_N \\ \vdots & \vdots & \dots & \vdots \\ v_1^{(N-2)} & v_2^{(N-2)} & \dots & v_N^{(N-2)} \\ v_1^{(N-1)} & v_2^{(N-1)} & \dots & v_N^{(N-1)} \end{pmatrix} \begin{pmatrix} h'_1 \\ h'_2 \\ \vdots \\ h'_N \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f \end{pmatrix}.$$

La matrice wronskiana di v_1, \dots, v_N è invertibile perché v_1, \dots, v_N sono linearmente indipendenti, quindi esiste un'unica soluzione per h'_1, \dots, h'_N . Infine h_k si ottiene da h'_k con un'integrazione.

Vediamo un esempio.

Esercizio 14.13. Sia $\omega \neq 0$ una costante, $f(t)$ una funzione continua. Calcolare le soluzioni dell'equazione

$$u'' + \omega^2 u = f(t).$$

Svolgimento. Sappiamo che

$$v_1(t) = \cos(\omega t), \quad v_2(t) = \sin(\omega t)$$

sono soluzioni dell'equazione omogenea $v'' + \omega^2 v = 0$, e sono linearmente indipendenti. Cerchiamo dunque una soluzione u dell'equazione non omogenea della forma

$$u(t) = h_1(t)v_1(t) + h_2(t)v_2(t) = h_1(t)\cos(\omega t) + h_2(t)\sin(\omega t).$$

La derivata è

$$u'(t) = h'_1(t)\cos(\omega t) + h'_2(t)\sin(\omega t) - \omega h_1(t)\sin(\omega t) + \omega h_2(t)\cos(\omega t).$$

Supponiamo che h_1, h_2 soddisfino

$$h'_1(t)\cos(\omega t) + h'_2(t)\sin(\omega t) = 0;$$

allora nella formula di u' rimane

$$u'(t) = -\omega h_1(t)\sin(\omega t) + \omega h_2(t)\cos(\omega t).$$

Deriviamo quest'ultima formula:

$$u''(t) = -\omega h'_1(t)\sin(\omega t) + \omega h'_2(t)\cos(\omega t) - \omega^2 h_1(t)\cos(\omega t) - \omega^2 h_2(t)\sin(\omega t).$$

Ora calcoliamo

$$\begin{aligned} u'' + \omega^2 u &= -\omega h'_1(t)\sin(\omega t) + \omega h'_2(t)\cos(\omega t) - \omega^2 h_1(t)\cos(\omega t) - \omega^2 h_2(t)\sin(\omega t) \\ &\quad + \omega^2 [h_1(t)\cos(\omega t) + h_2(t)\sin(\omega t)] \\ &= -\omega h'_1(t)\sin(\omega t) + \omega h'_2(t)\cos(\omega t). \end{aligned}$$

Supponiamo che h_1, h_2 soddisfino

$$-\omega h_1'(t) \sin(\omega t) + \omega h_2'(t) \cos(\omega t) = f(t).$$

Allora $u'' + \omega^2 u = f(t)$, cioè u è soluzione dell'equazione non omogenea, come volevamo.

Quindi cerchiamo h_1, h_2 che soddisfano queste due equazioni:

$$\begin{cases} h_1'(t) \cos(\omega t) + h_2'(t) \sin(\omega t) = 0, \\ -\omega h_1'(t) \sin(\omega t) + \omega h_2'(t) \cos(\omega t) = f(t). \end{cases}$$

Se moltiplichiamo la prima equazione per $\omega \sin(\omega t)$, la seconda per $\cos(\omega t)$ e sommiamo, otteniamo

$$h_2'(t) = \frac{1}{\omega} f(t) \cos(\omega t).$$

Se invece moltiplichiamo la prima equazione per $\omega \cos(\omega t)$, la seconda per $\sin(\omega t)$ e facciamo la differenza, otteniamo

$$h_1'(t) = -\frac{1}{\omega} \sin(\omega t) f(t).$$

Dunque scegliamo

$$h_1(t) = -\frac{1}{\omega} \int_0^t \sin(\omega s) f(s) ds, \quad h_2(t) = \frac{1}{\omega} \int_0^t \cos(\omega s) f(s) ds,$$

e otteniamo una soluzione dell'equazione non omogenea

$$\begin{aligned} u_P(t) &= h_1(t) \cos(\omega t) + h_2(t) \sin(\omega t) \\ &= -\frac{1}{\omega} \int_0^t \sin(\omega s) f(s) ds \cos(\omega t) + \frac{1}{\omega} \int_0^t \cos(\omega s) f(s) ds \sin(\omega t) \\ &= \frac{1}{\omega} \int_0^t f(s) \left(\sin(\omega t) \cos(\omega s) - \cos(\omega t) \sin(\omega s) \right) ds \\ &= \frac{1}{\omega} \int_0^t f(s) \sin(\omega(t-s)) ds. \end{aligned}$$

La soluzione generale dell'equazione non omogenea è $u = u_P + v$ con v soluzione dell'equazione omogenea. Quindi le soluzioni dell'equazione $u'' + \omega^2 u = f(t)$ sono le funzioni

$$u(t) = \frac{1}{\omega} \int_0^t f(s) \sin(\omega(t-s)) ds + c_1 \cos(\omega t) + c_2 \sin(\omega t), \quad c_1, c_2 \in \mathbb{R}.$$

□

Equazioni a coefficienti costanti. Quando i coefficienti a_0, a_1, \dots, a_{N-1} sono costanti, la matrice $A(t)$ in (14.7) è la matrice costante

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{N-1} \end{pmatrix}, \quad a_k \in \mathbb{R}. \quad (14.11)$$

Lemma 14.14. (Polinomio caratteristico della matrice A). Siano $a_0, a_1, \dots, a_{N-1} \in \mathbb{R}$, e sia A la matrice in (14.11). Allora il polinomio caratteristico

$$p(\lambda) := \det(\lambda I - A)$$

della matrice A è uguale a

$$p(\lambda) = \lambda^N + a_{N-1}\lambda^{N-1} + \dots + a_1\lambda + a_0.$$

Esercizio 14.15. Dimostrare il Lemma 14.14.

Suggerimento: per induzione su N ; sviluppare il determinante $\det(\lambda I - A)$ rispetto alla prima colonna.

Definizione 14.16. (Polinomio caratteristico dell'equazione differenziale). Sia $a_0, a_1, \dots, a_{N-1} \in \mathbb{R}$, e consideriamo l'equazione lineare (scalare, omogenea, a coefficienti costanti) $\mathcal{L}u = 0$, dove \mathcal{L} è definito in (14.8), cioè

$$\mathcal{L}u := u^{(N)} + a_{N-1}u^{(N-1)} + \dots + a_1u' + a_0u.$$

Si dice *polinomio caratteristico dell'equazione $\mathcal{L}u = 0$* il polinomio

$$p(\lambda) := \lambda^N + a_{N-1}\lambda^{N-1} + \dots + a_1\lambda + a_0$$

che si ottiene da $\mathcal{L}u$ sostituendo la k -esima derivata $u^{(k)}$ di u con la potenza k -esima λ^k di λ , per ogni $k = 0, \dots, N$.

Dal Lemma 14.14, il polinomio caratteristico dell'equazione scalare $\mathcal{L}u = 0$ è il polinomio della matrice A in (14.11).

Teorema 14.17. (Soluzioni delle equazioni lineari a coefficienti costanti). Siano $a_0, a_1, \dots, a_{N-1} \in \mathbb{R}$, sia \mathcal{L} definito in (14.8), e sia $p(\lambda)$ il polinomio caratteristico dell'equazione $\mathcal{L}u = 0$. Supponiamo che le radici di p siano i numeri reali λ_k , $k = 1, \dots, r$, e le coppie di numeri complessi coniugati $\alpha_k \pm i\beta_k$, $k = 1, \dots, q$, con α_k, β_k reali e $\beta_k \neq 0$. Supponiamo che λ_k abbia molteplicità algebrica f_k , e che $\alpha_k + i\beta_k$, $\alpha_k - i\beta_k$ abbiano ciascuno molteplicità algebrica g_k ; dunque

$$f_1 + \dots + f_r + 2g_1 + \dots + 2g_q = N.$$

Per ogni radice reale λ_k consideriamo le f_k funzioni

$$t^j e^{\lambda_k t}, \quad j = 0, \dots, f_k - 1; \quad (14.12)$$

per ogni coppia di radici complesse coniugate $\alpha_k \pm i\beta_k$ consideriamo le $2g_k$ funzioni

$$t^j e^{\alpha_k t} \cos(\beta_k t), \quad t^j e^{\alpha_k t} \sin(\beta_k t), \quad j = 0, \dots, g_k - 1. \quad (14.13)$$

Le N funzioni considerate sono soluzioni linearmente indipendenti dell'equazione $\mathcal{L}u = 0$. Quindi sono una base dello spazio vettoriale $S_{\mathcal{L}}$ (definito in (14.9)) delle soluzioni dell'equazione $\mathcal{L}u = 0$.

Dimostrazione. Sia $u \in S_{\mathcal{L}}$. Dall'Osservazione 14.5, la funzione vettoriale $U := (u, u', \dots, u^{(N-1)})$ è soluzione dell'equazione $U' = AU$. Dal Teorema 13.4, $A = VEV^{-1}$ con V matrice invertibile ed E matrice in forma canonica reale di Jordan. Quindi

$$U' = AU = VEV^{-1}U.$$

Di conseguenza la funzione vettoriale $X(t) := V^{-1}U(t)$ è soluzione dell'equazione $X' = EX$. Dunque

$$X(t) = e^{tE}c,$$

con $c = X(0)$ vettore di \mathbb{R}^N . La matrice E ha la struttura diagonale a blocchi descritta nel Teorema 13.4; il suo esponenziale è stato calcolato in (13.34). Dalla formula dell'esponenziale di ciascun blocco di Jordan (vedi (13.31) per i blocchi con autovalore reale, (13.33) per i blocchi con coppia di autovalori complessi coniugati), ogni componente di $X(t)$ è una combinazione lineare delle N funzioni (14.12)-(14.13), che indichiamo $\phi_1(t), \dots, \phi_N(t)$; quindi

$$X_k(t) = b_{k,1}\phi_1(t) + \dots + b_{k,N}\phi_N(t), \quad k = 1, \dots, N,$$

per certi coefficienti $b_{k,j} \in \mathbb{R}$. Per definizione di $X(t)$, si ha

$$U(t) = VX(t).$$

La funzione $u(t)$ è la prima componente di $U(t)$, quindi è il prodotto riga per colonna della prima riga di V per il vettore $X(t)$,

$$u(t) = V_{11}X_1(t) + \dots + V_{1N}X_N(t).$$

Quindi $u(t)$ è una combinazione lineare delle funzioni ϕ_1, \dots, ϕ_N .

Abbiamo dimostrato che $\phi_1(t), \dots, \phi_N(t)$ formano un sistema di generatori dello spazio vettoriale $S_{\mathcal{L}}$. Se fossero linearmente dipendenti, potremmo scrivere una di queste funzioni come combinazione lineare delle altre $N - 1$, quindi basterebbero $N - 1$ funzioni a generare $S_{\mathcal{L}}$. Ma questo è assurdo perché $S_{\mathcal{L}}$ ha dimensione N (vedi Lemma 14.6). \square

Esempio. Cerchiamo le soluzioni dell'equazione omogenea

$$u^{(7)} - 11u^{(6)} + 69u^{(5)} - 255u^{(4)} + 615u^{(3)} - 861u'' + 611u' - 169u = 0.$$

Il polinomio caratteristico è

$$p(\lambda) = \lambda^7 - 11\lambda^6 + 69\lambda^5 - 255\lambda^4 + 615\lambda^3 - 861\lambda^2 + 611\lambda - 169;$$

la sua fattorizzazione in \mathbb{C} è

$$p(\lambda) = (\lambda - 1)^3(\lambda - 2 - 3i)^2(\lambda - 2 + 3i)^2,$$

le sue radici in \mathbb{C} sono

$$\begin{aligned} & 1 \quad \text{con molteplicità } 3, \\ & 2 + 3i \quad \text{con molteplicità } 2, \\ & 2 - 3i \quad \text{con molteplicità } 2. \end{aligned}$$

Le 7 soluzioni linearmente indipendenti dell'equazione date dal Teorema 14.17 sono dunque

$$\begin{aligned} \phi_1(t) &= e^t, & \phi_2(t) &= te^t, & \phi_3(t) &= t^2e^t, \\ \phi_4(t) &= e^{2t} \cos(3t), & \phi_5(t) &= e^{2t} \sin(3t), \\ \phi_6(t) &= te^{2t} \cos(3t), & \phi_7(t) &= te^{2t} \sin(3t). \end{aligned}$$

Le soluzioni dell'equazione sono tutte e sole le combinazioni lineari di ϕ_1, \dots, ϕ_7 , cioè la soluzione generale dell'equazione è

$$u(t) = \sum_{k=1}^7 c_k \phi_k(t), \quad c_k \in \mathbb{R}.$$

Metodo di somiglianza per equazioni non omogenee. Consideriamo l'equazione non omogenea

$$\mathcal{L}u = f(t)$$

in cui

$$\mathcal{L}u = u^{(N)} + a_{N-1}u^{(N-1)} + \dots + a_1u' + a_0u$$

ha coefficienti costanti $a_k \in \mathbb{R}$. Il Teorema 14.17 fornisce N soluzioni linearmente indipendenti dell'equazione omogenea $\mathcal{L}u = 0$; quindi possiamo applicare il metodo di variazione delle costanti per ottenere una soluzione dell'equazione non omogenea.

In alternativa, quando $f(t)$ è una somma di prodotti di esponenziali, seni, coseni e polinomi in t , diciamo

$$f(t) = \sum_{k=1}^n [p_k(t) \cos(\omega_k t) + q_k(t) \sin(\omega_k t)] e^{\mu_k t}$$

con ω_k, μ_k costanti e p_k, q_k polinomi, possiamo sfruttare il fatto che \mathcal{L} preserva questa struttura e cercare una soluzione "particolare" $u(t)$ dell'equazione non omogenea $\mathcal{L}u = f$ dello stesso tipo di f . Vediamo questo metodo, detto *di somiglianza*, tramite qualche esempio (riprendiamo dei casi particolari dell'equazione dell'Esercizio 14.13). Esercizi di questo tipo si trovano anche nei testi di Analisi 2.

Esercizio 14.18. Calcolare le soluzioni dell'equazione

$$u'' + 4u = e^t.$$

Svolgimento. Il termine non omogeneo è e^t ; cerchiamo una soluzione dello stesso tipo, cioè della forma $u(t) = ce^t$, con c costante. Calcoliamo

$$u'' + 4u = ce^t + 4ce^t = 5ce^t.$$

Scegliendo $c = 1/5$ si ha la soluzione particolare $u_P(t) = \frac{1}{5}e^t$. La soluzione generale è dunque

$$u(t) = \frac{1}{5}e^t + c_1 \cos(2t) + c_2 \sin(2t).$$

□

Esercizio 14.19. Calcolare le soluzioni dell'equazione

$$u'' + 4u = \sin(2t) + e^t.$$

Svolgimento. Essendo lineare, il problema si può suddividere in due parti: cerchiamo una soluzione particolare v dell'equazione

$$v'' + 4v = \sin(2t), \tag{14.14}$$

una soluzione particolare w dell'equazione

$$w'' + 4w = e^t, \quad (14.15)$$

e la somma $u := v+w$ sarà quindi soluzione dell'equazione con termine non omogeneo $\sin(2t) + e^t$.

Una soluzione particolare di (14.15) è stata già calcolata nell'Esercizio 14.18, ed è $w(t) = \frac{1}{5}e^t$.

Cerchiamo v soluzione di (14.14). Il termine non omogeneo è $\sin(2t)$. Il primo tentativo che potrebbe venire in mente "per somiglianza" è provare con $v(t) = c \sin(2t)$, $c \in \mathbb{R}$, ma non funziona: infatti $\sin(2t)$ è soluzione dell'equazione omogenea $v'' + 4v = 0$, quindi $v'' + 4v$ non può essere uguale a $\sin(2t)$. Per lo stesso motivo, nemmeno $\cos(2t)$ può funzionare.

Proviamo allora con un fattore t , cioè $v(t) = ct \sin(2t)$: si ha

$$\begin{aligned} v(t) &= ct \sin(2t), \\ v'(t) &= c \sin(2t) + 2ct \cos(2t), \\ v''(t) &= 4c \cos(2t) - 4ct \sin(2t), \end{aligned}$$

da cui

$$v'' + 4v = 4c \cos(2t).$$

Ancora non funziona; siamo partiti dal seno e abbiamo ottenuto un coseno come termine non omogeneo. Come prossimo tentativo partiamo dal coseno: sia $v(t) = ct \cos(2t)$. Si ha

$$\begin{aligned} v(t) &= ct \cos(2t), \\ v'(t) &= c \cos(2t) - 2ct \sin(2t), \\ v''(t) &= -4c \sin(2t) - 4ct \cos(2t), \end{aligned}$$

da cui

$$v'' + 4v = -4c \sin(2t).$$

Scegliamo $c = -1/4$: abbiamo trovato

$$v(t) = -\frac{1}{4} t \cos(2t),$$

soluzione particolare di $v'' + 4v = \sin(2t)$.

La soluzione generale dell'esercizio è dunque

$$u(t) = \frac{1}{5} e^t - \frac{1}{4} t \cos(2t) + c_1 \cos(2t) + c_2 \sin(2t), \quad c_1, c_2 \in \mathbb{R}.$$

In alternativa, per calcolare la soluzione particolare v si può procedere così: cerchiamo v della forma

$$v(t) = P(t) \cos(2t) + Q(t) \sin(2t)$$

con P, Q polinomi da determinare. Si ha (con notazione un po' abbreviata)

$$\begin{aligned} v &= P \cos + Q \sin, \\ v' &= P' \cos - 2P \sin + Q' \sin + 2Q \cos, \\ v'' &= P'' \cos - 4P' \sin - 4P \cos + Q'' \sin + 4Q' \cos - 4Q \sin, \end{aligned}$$

da cui

$$v'' + 4v = (P'' + 4Q') \cos + (Q'' - 4P') \sin.$$

Il termine non omogeneo è uguale a $\sin(2t)$ se P, Q soddisfano

$$\begin{cases} P'' + 4Q' = 0, \\ Q'' - 4P' = 1. \end{cases} \quad (14.16)$$

Se P, Q sono costanti, allora P', P'', Q', Q'' sono nulli e il sistema non ha soluzione; saliamo di grado, e proviamo con P, Q di primo grado:

$$P(t) = P_0 + P_1 t, \quad Q(t) = Q_0 + Q_1 t.$$

Si ha

$$P' = P_1, \quad P'' = 0, \quad Q' = Q_1, \quad Q'' = 0,$$

e il sistema diventa

$$4Q_1 = 0, \quad -4P_1 = 1.$$

Su P_0, Q_0 non ci sono condizioni, quindi li possiamo scegliere nulli. Dunque

$$Q(t) = 0, \quad P(t) = -\frac{1}{4}t.$$

Abbiamo trovato (nuovamente) $v(t) = -\frac{1}{4}t \cos(2t)$.

Per capire il grado dei polinomi, si può anche ragionare così: se P, Q sono di grado n , allora P', Q' sono di grado $n - 1$, e P'', Q'' di grado $n - 2$; quindi $P'' + 4Q'$ e $Q'' - 4P'$ sono di grado $n - 1$. I termini a destra nel sistema (14.16) sono costanti, cioè hanno grado zero; quindi, uguagliando i gradi, si trova $n - 1 = 0$, da cui $n = 1$.

Se, in esercizi simili, le equazioni sui gradi danno risultati diversi (per esempio $n = 1$ da un'equazione, $n = 3$ da un'altra equazione), allora come n scegliamo il grado massimo. \square

* * *

Torniamo ora allo studio delle equazioni $u' = f(t, u)$, e vediamo una “versione differenziabile” del teorema di dipendenza continua dai dati iniziali.

Teorema 14.20. (Teorema di differenziabilità rispetto ai dati iniziali). *Sia $\Omega \subseteq \mathbb{R} \times \mathbb{R}^N$ un aperto, e sia $f : \Omega \rightarrow \mathbb{R}^N$ una funzione continua che ha le derivate parziali $\partial_{y_k} f : \Omega \rightarrow \mathbb{R}^N$, $k = 1, \dots, N$, continue. Sia $(t_0, y_0) \in \Omega$, e sia $u : I \rightarrow \mathbb{R}^N$ la soluzione del problema di Cauchy*

$$\begin{cases} u' = f(t, u), \\ u(t_0) = y_0, \end{cases} \quad (14.17)$$

definita nel suo intervallo massimale I .

Sia $[\alpha, \beta] \subseteq I$, con $t_0 \in [\alpha, \beta]$, e sia $\delta_0 > 0$ la costante che viene fornita dal Teorema di dipendenza continua (Teorema 11.2). Per ogni $h_0 \in \mathbb{R}^N$, $|h_0| \leq \delta_0$, sia v la soluzione del problema di Cauchy

$$\begin{cases} v' = f(t, v), \\ v(t_0) = y_0 + h_0 \end{cases} \quad (14.18)$$

(che è definita in $[\alpha, \beta]$ grazie al Teorema di dipendenza continua). Sia $w : I \rightarrow \text{Mat}_{N \times N}(\mathbb{R})$ la soluzione del problema di Cauchy

$$\begin{cases} w' = (D_y f)(t, u(t))w, \\ w(t_0) = I_N, \end{cases} \quad (14.19)$$

dove $D_y f(t, y)$ è la matrice jacobiana di f rispetto alle variabili y_1, \dots, y_N , cioè

$$D_y f(t, y) = \begin{pmatrix} \partial_{y_1} f_1(t, y) & \dots & \partial_{y_N} f_1(t, y) \\ \vdots & \dots & \vdots \\ \partial_{y_1} f_N(t, y) & \dots & \partial_{y_N} f_N(t, y) \end{pmatrix},$$

e $I_N \in \text{Mat}_{N \times N}(\mathbb{R})$ è la matrice identità.

Allora per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che

$$\sup_{t \in [\alpha, \beta]} \frac{|v(t) - u(t) - w(t)h_0|}{|h_0|} \leq \varepsilon \quad \forall h_0 \in \mathbb{R}^N, h_0 \neq 0, |h_0| \leq \delta.$$

Dimostrazione nella prossima lezione.

15 Lezione 15

Dimostrazione del Teorema 14.20 di dipendenza differenziabile dai dati iniziali. Sia $[\alpha, \beta]$ fissato come nell'enunciato. Dal Teorema 11.2 di dipendenza continua e dalla sua dimostrazione, esistono $\delta_0, r > 0$ ed $L, M \geq 0$ tali che, per ogni $|h_0| \leq \delta_0$, la soluzione v del problema di Cauchy (14.18) è definita su $[\alpha, \beta]$ e soddisfa

$$|u(t) - v(t)| \leq e^{L|t-t_0|}|h_0| \leq e^{L(\beta-\alpha)}|h_0| \quad \forall t \in [\alpha, \beta], \quad (15.1)$$

il compatto “tubolare”

$$K = \{(t, y) : t \in [\alpha, \beta], |y - u(t)| \leq r\}$$

è contenuto in Ω , in K si ha

$$|f(t, y)| \leq M, \quad |f(t, y) - f(t, z)| \leq L|y - z| \quad \forall (t, y), (t, z) \in K,$$

e inoltre $|v(t) - u(t)| \leq r/2$ per ogni $t \in [\alpha, \beta]$. Anche la soluzione w di (14.19) è definita in $[\alpha, \beta]$, perché è soluzione dell'equazione lineare $w' = (D_y f)(t, u(t))w$.

Sia

$$\psi(t) := v(t) - u(t) - w(t)h_0,$$

cosicché la tesi diventa: per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che

$$|\psi(t)| \leq \varepsilon|h_0| \quad \forall t \in [\alpha, \beta], \quad \forall h_0 \in \mathbb{R}^N \setminus \{0\}, \quad |h_0| \leq \delta.$$

Calcoliamo

$$\begin{aligned} \psi'(t) &= v'(t) - u'(t) - w'(t)h_0 \\ &= f(t, v(t)) - f(t, u(t)) - (D_y f)(t, u(t))w(t)h_0 \\ &= f(t, v(t)) - f(t, u(t)) - (D_y f)(t, u(t))[v(t) - u(t)] + (D_y f)(t, u(t))\psi(t). \end{aligned}$$

Quindi $\psi(t)$ soddisfa l'equazione lineare non omogenea

$$\psi'(t) = a(t)\psi(t) + b(t) \quad (15.2)$$

con

$$\begin{aligned} a(t) &= (D_y f)(t, u(t)), \\ b(t) &= f(t, v(t)) - f(t, u(t)) - (D_y f)(t, u(t))[v(t) - u(t)] \end{aligned} \quad (15.3)$$

(per ogni t , $a(t)$ è una matrice, $b(t)$ un vettore).

Stimiamo i coefficienti $a(t), b(t)$ in (15.3). Poiché $D_y f$ è continua e l'insieme $\{(t, u(t)) : t \in [\alpha, \beta]\}$ è compatto, la funzione $(D_y f)(t, u(t))$ è limitata, cioè

$$|a(t)| = |(D_y f)(t, u(t))| \leq L_1 \quad \forall t \in [\alpha, \beta],$$

per un certo $L_1 \geq 0$. Per stimare $b(t)$, dimostriamo la seguente affermazione:

$$\left\{ \begin{array}{l} \text{Per ogni } \varepsilon_1 > 0 \text{ esiste } \delta_1 > 0 \text{ tale che} \\ |f(t, z) - f(t, y) - D_y f(t, y)(z - y)| \leq \varepsilon_1 |z - y| \\ \forall (t, z), (t, y) \in K, |z - y| \leq \delta_1. \end{array} \right. \quad (15.4)$$

Poiché $D_y f$ è continua e K è compatto, $D_y f$ è uniformemente continua in K . Quindi per ogni $\varepsilon_1 > 0$ esiste $\delta_1 > 0$ tale che

$$|(D_y f)(t, y) - (D_y f)(s, z)| \leq \varepsilon_1 \quad \forall (t, y), (s, z) \in K, \quad |(t, y) - (s, z)| \leq \delta_1.$$

In particolare,

$$|(D_y f)(t, y) - (D_y f)(t, z)| \leq \varepsilon_1 \quad \forall (t, y), (t, z) \in K, \quad |y - z| \leq \delta_1.$$

Siano t, y, z fissati, con $(t, y), (t, z) \in K$ e $|y - z| \leq \delta_1$. Sia

$$\varphi(\lambda) := f(t, y + \lambda(z - y)).$$

Poiché le funzioni $\lambda \mapsto y + \lambda(z - y)$ e $y \mapsto f(t, y)$ sono di classe C^1 , anche φ è di classe C^1 ; la sua derivata è

$$\varphi'(\lambda) = (D_y f)(t, y + \lambda(z - y))(z - y).$$

Quindi

$$\begin{aligned} f(t, z) - f(t, y) &= \varphi(1) - \varphi(0) = \int_0^1 \varphi'(\lambda) d\lambda \\ &= \int_0^1 (D_y f)(t, y + \lambda(z - y))(z - y) d\lambda. \end{aligned}$$

Poi

$$D_y f(t, y)(z - y) = D_y f(t, y)(z - y) \int_0^1 d\lambda = \int_0^1 D_y f(t, y)(z - y) d\lambda,$$

da cui si ha

$$\begin{aligned} &|f(t, z) - f(t, y) - D_y f(t, y)(z - y)| \\ &= \left| \int_0^1 \left((D_y f)(t, y + \lambda(z - y)) - (D_y f)(t, y) \right) (z - y) d\lambda \right| \\ &\leq \int_0^1 |(D_y f)(t, y + \lambda(z - y)) - (D_y f)(t, y)| |z - y| d\lambda \\ &\leq \int_0^1 \varepsilon_1 |z - y| d\lambda = \varepsilon_1 |z - y| \end{aligned}$$

perché $|z - y| \leq \delta_1$. Abbiamo dimostrato (15.4).

Torniamo alla stima di $b(t)$. Come osservato in (15.1), per ogni $|h_0| \leq \delta_0$ si ha

$$(t, v(t)) \in K, \quad |v(t) - u(t)| \leq e^{L(\beta - \alpha)} |h_0| \quad \forall t \in [\alpha, \beta].$$

Sia $\varepsilon_1 > 0$, e sia $\delta_1 > 0$ la corrispondente costante data da (15.4). Se

$$e^{L(\beta - \alpha)} |h_0| \leq \delta_1, \quad |h_0| \leq \delta_0, \tag{15.5}$$

allora $(t, v(t)) \in K$ e $|v(t) - u(t)| \leq \delta_1$ per ogni $t \in [\alpha, \beta]$, e quindi, per (15.4),

$$\begin{aligned} |b(t)| &= |f(t, v(t)) - f(t, u(t)) - (D_y f)(t, u(t))[v(t) - u(t)]| \\ &\leq \varepsilon_1 |v(t) - u(t)| \\ &\leq \varepsilon_1 e^{L(\beta - \alpha)} |h_0| \quad \forall t \in [\alpha, \beta]. \end{aligned}$$

I coefficienti $a(t), b(t)$ soddisfano dunque le stime

$$|a(t)| \leq A := L_1, \quad |b(t)| \leq B := \varepsilon_1 e^{L(\beta-\alpha)} |h_0|, \quad \forall t \in [\alpha, \beta]. \quad (15.6)$$

Poiché ψ soddisfa (15.2), per ogni $t \in [\alpha, \beta]$ si ha

$$\begin{aligned} |\psi'(t)| &\leq |a(t)| |\psi(t)| + |b(t)| \\ &\leq A |\psi(t)| + B. \end{aligned}$$

Dal Lemma di Gronwall, usando il fatto che $|t - t_0| \leq \beta - \alpha$ e $\psi(t_0) = 0$, si ha allora

$$|\psi(t)| \leq e^{A|t-t_0|} (|\psi(t_0)| + B|t - t_0|) \leq e^{A(\beta-\alpha)} B(\beta - \alpha),$$

cioè, dalla definizione (15.6) di A, B ,

$$|\psi(t)| \leq e^{L_1(\beta-\alpha)} \varepsilon_1 e^{L(\beta-\alpha)} |h_0| (\beta - \alpha) = C \varepsilon_1 |h_0| \quad (15.7)$$

per ogni $t \in [\alpha, \beta]$, con

$$C := e^{(L+L_1)(\beta-\alpha)} (\beta - \alpha)$$

(e C è indipendente da ε_1). Abbiamo dimostrato che per ogni $\varepsilon_1 > 0$ esiste $\delta_1 > 0$ tale che, se h_0 soddisfa (15.5), allora $\psi(t)$ soddisfa (15.7).

A questo punto possiamo concludere: dato $\varepsilon > 0$, scegliamo $\varepsilon_1 := \varepsilon/C$, esiste un δ_1 corrispondente, definiamo $\delta := \min\{\delta_0, e^{-L(\beta-\alpha)} \delta_1\}$, cosicché per $|h_0| \leq \delta$ vale la (15.5), da cui segue che $|\psi(t)| \leq C \varepsilon_1 |h_0| = \varepsilon |h_0|$ per ogni $t \in [\alpha, \beta]$. \square

Osservazione 15.1. (*Differenziabilità del flusso rispetto al dato iniziale*). Il Teorema 14.20 dice che

$$v(t) = u(t) + w(t)h_0 + o(|h_0|) \quad \text{in } C^0([\alpha, \beta], \mathbb{R}^N) \quad \text{per } |h_0| \rightarrow 0,$$

cioè che

$$v(t) = u(t) + w(t)h_0 + r(t),$$

dove il resto $r(t) := v(t) - u(t) - w(t)h_0$ ha sup-norma

$$\|r\|_{\infty, [\alpha, \beta]} = \sup\{|r(t)| : t \in [\alpha, \beta]\} = o(|h_0|) \quad \text{per } h_0 \rightarrow 0.$$

Possiamo rinunciare questo risultato in termini di differenziabilità della soluzione rispetto al dato iniziale y_0 . Indichiamo $\phi(t, t_0, y_0)$ il valore al tempo t della soluzione del problema di Cauchy con condizioni iniziali (t_0, y_0) (*flusso* dell'equazione differenziale di campo f). La soluzione u del problema di Cauchy (14.17) è dunque $u(t) = \phi(t, t_0, y_0)$, mentre la soluzione v del problema (14.18) è $v(t) = \phi(t, t_0, y_0 + h_0)$. Consideriamo il flusso come funzione del dato iniziale y_0 , cioè, per t_0 fissato, consideriamo la mappa

$$\mathbb{R}^N \rightarrow C^0([\alpha, \beta], \mathbb{R}^N), \quad y_0 \mapsto \phi(\cdot, t_0, y_0),$$

che a ciascun dato iniziale y_0 associa la soluzione del problema di Cauchy di condizioni iniziali (t_0, y_0) , che è una funzione continua (anzi, C^1) di $t \in [\alpha, \beta]$.

Il Teorema 14.20 dice che tale mappa è differenziabile, e che il suo differenziale è la funzione lineare $h_0 \mapsto w(\cdot)h_0$ che al vettore h_0 associa la funzione $\gamma(t) := w(t)h_0$, soluzione del problema di Cauchy lineare

$$\gamma'(t) = (D_y f)(t, u(t))\gamma(t), \quad \gamma(t_0) = h_0.$$

La differenziabilità equivale alla validità della formula di Taylor di ordine 1 con resto in forma di Peano, dove la norma nello spazio di arrivo $C^0([\alpha, \beta], \mathbb{R}^N)$ è la norma del sup.

Nella seconda parte del programma, che ora cominciamo, tratteremo più in generale la nozione di differenziabilità di una funzione $f : X \rightarrow Y$ tra spazi di Banach X, Y .

Calcolo differenziale in spazi di Banach

In questa parte di corso ci occupiamo di calcolo differenziale in spazi di Banach. Ricordiamo che uno spazio vettoriale normato $(X, \|\cdot\|_X)$ si dice spazio di Banach se è completo, cioè se ogni successione di Cauchy è convergente. Gli spazi di Banach che abbiamo trattato finora sono: \mathbb{R}^N , lo spazio $\text{Mat}_{N \times N}(\mathbb{R})$ delle matrici, e lo spazio $C^0([a, b], \mathbb{R}^N)$ delle funzioni continue da un intervallo $[a, b]$ a valori in \mathbb{R}^N (vedi Proposizione 5.10).

Definizione 15.2. (*Funzioni continue*). Siano X, Y spazi vettoriali normati, $\Omega \subseteq X$ un aperto, $p \in \Omega$, ed $f : \Omega \rightarrow Y$ una funzione.

Diciamo che f è *continua nel punto* p se $f(x) \rightarrow f(p)$ in Y per $x \rightarrow p$ in X , cioè se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che $\|f(x) - f(p)\|_Y \leq \varepsilon$ per ogni $x \in \Omega$, $\|x - p\|_X \leq \delta$.

Diciamo che f è *continua nell'aperto* Ω se è continua in ogni punto $p \in \Omega$.

Notazione. Per le funzioni *lineari*, spesso si usa la notazione Ax al posto di $A(x)$, e ABx al posto di $A(B(x))$ (come si fa per matrici e vettori in \mathbb{R}^N). Quasi sempre scriviamo soltanto $\|Ax\|, \|x\|$ al posto di $\|Ax\|_Y, \|x\|_X$. Usiamo i termini *funzione lineare, operatore lineare, mappa lineare, applicazione lineare* come sinonimi.

Definizione 15.3. (*Norma operatoriale di funzioni lineari*). Siano X, Y spazi vettoriali normati, e sia $A : X \rightarrow Y$ una funzione lineare. Definiamo

$$\|A\|_{\text{op}} := \sup \left\{ \frac{\|Ax\|_Y}{\|x\|_X} : x \in X, x \neq 0 \right\}.$$

Quando lo spazio X ha dimensione finita, ogni funzione lineare $A : X \rightarrow Y$ è continua e ha norma $\|A\|_{\text{op}}$ finita. Quando invece X ha dimensione infinita, esistono funzioni lineari A che non sono continue, per le quali $\|A\|_{\text{op}} = \infty$: vedi Osservazione 15.4 e Proposizione 15.9.

Osservazione 15.4. (*Esempio di funzione lineare con $\|A\|_{\text{op}} = \infty$*). Dall'algebra lineare sappiamo che ogni spazio vettoriale ammette una base (affermazione equivalente al Lemma di Zorn e all'assioma della scelta). Sia X uno spazio vettoriale su \mathbb{R} di dimensione infinita, e sia $(v_i)_{i \in I}$ una base di X ; quindi l'insieme I degli indici ha infiniti elementi. Ogni vettore $x \in X$ si scrive (in modo unico) come combinazione lineare finita di elementi della base, cioè $x = \sum_{i \in F} c_i v_i$ per certi coefficienti $c_i \in \mathbb{R}$, per un certo sottoinsieme finito $F \subset I$; si scrive anche $x = \sum_{i \in I} c_i v_i$, con $c_i = 0$ per ogni $i \notin F$, e si dice che i coefficienti c_i sono "quasi tutti nulli" (cioè sono nulli tutti eccetto un numero finito).

Fissiamo una successione di indici i_1, i_2, \dots in I , tutti distinti tra loro, e consideriamo la funzione lineare $A : X \rightarrow X$ che manda v_{i_1} in v_{i_1} , v_{i_2} in $2v_{i_2}$, v_{i_3} in $3v_{i_3}$, ecc., e manda v_i in zero per tutti gli altri indici $i \in I$, cioè definiamo

$$a_i := \begin{cases} n & \text{se } i = i_n, \\ 0 & \text{se } i \in I \setminus \{i_1, i_2, \dots\}, \end{cases}$$

e

$$A : X \rightarrow X, \quad x = \sum_{i \in I} c_i v_i \quad \mapsto \quad Ax = \sum_{i \in I} c_i a_i v_i.$$

La funzione A è ben posta perché per ogni $x \in X$ solo un numero finito di coefficienti c_i possono essere diversi da zero, quindi la somma che definisce Ax è una somma finita di vettori di X . Con qualsiasi norma in X si ha allora

$$\sup_{x \in X \setminus \{0\}} \frac{\|Ax\|}{\|x\|} \geq \sup_{n=1,2,\dots} \frac{\|Av_{i_n}\|}{\|v_{i_n}\|} = \sup_{n=1,2,\dots} \frac{n\|v_{i_n}\|}{\|v_{i_n}\|} = \infty.$$

Osservazione 15.5. (*Funzione lineare iniettiva e non suriettiva, e viceversa*). Un'altra proprietà notevole delle funzioni lineari tra spazi vettoriali di dimensione infinita è l'esistenza di funzioni lineari iniettive che non sono suriettive, o viceversa. Anche questa è una proprietà un po' sorprendente se abbiamo in mente solo \mathbb{R}^N come esempio di spazio vettoriale.

Consideriamo, per esempio, l'insieme delle successioni limitate di numeri reali,

$$X = \{x = (x_1, x_2, \dots) : x \text{ limitata}\},$$

munito della norma del sup

$$x = (x_1, x_2, \dots) \in X, \quad \|x\| = \sup_{n \geq 1} |x_n|.$$

X è spazio vettoriale normato. La "traslazione a sinistra" $(Ax)_n := x_{n+1}$, cioè

$$A : X \rightarrow X, \quad x = (x_1, x_2, \dots) \mapsto Ax = (x_2, x_3, \dots),$$

è lineare, suriettiva, e non è iniettiva. La "traslazione a destra" $(Bx)_1 := 0$, $(Bx)_{n+1} := x_n$, cioè

$$B : X \rightarrow X, \quad x = (x_1, x_2, \dots) \mapsto Bx = (0, x_1, x_2, \dots),$$

è lineare, iniettiva, e non è suriettiva.

Esercizio 15.6. (*Traslazione a destra/sinistra*). Siano X, A, B come nell'Osservazione 15.5.

- (i) Verificare che $(X, \|\cdot\|)$ è spazio vettoriale normato.
- (ii) Verificare che A, B sono lineari, A è suriettiva, non iniettiva, B viceversa.
- (iii) Calcolare $\|A\|_{\text{op}}, \|B\|_{\text{op}}$.
- (iv) Stabilire se X è spazio di Banach.

Esercizio 15.7. (*Spazio delle successioni definitivamente nulle*). Sia X l'insieme delle successioni di numeri reali definitivamente nulle,

$$X = \{x = (x_1, x_2, \dots) : \text{esiste } n \in \mathbb{N} \text{ tale che } x_k = 0 \text{ per ogni } k \geq n\},$$

sia $\|x\| = \sup_n |x_n|$ la norma del sup, e siano $A, B : X \rightarrow X$ le funzioni lineari che mandano il vettore $x = (x_1, x_2, \dots)$ in

$$Ax = (nx_n)_n = (x_1, 2x_2, 3x_3, \dots), \quad Bx = (x_n/n)_n = (x_1, \frac{x_2}{2}, \frac{x_3}{3}, \dots).$$

- (i) Verificare che $(X, \|\cdot\|)$ è spazio vettoriale normato.
- (ii) Stabilire se X è spazio di Banach.
- (iii) Verificare che A, B mappano X in sé.
- (iv) Verificare che A, B sono lineari e che $ABx = x = BAx$ per ogni $x \in X$.
- (v) Calcolare $\|A\|_{\text{op}}, \|B\|_{\text{op}}$.
- (vi) Sia

$$\|x\|_1 := \sum_{n=1}^{\infty} |x_n|,$$

serie convergente perché, per ogni $x \in X$, $\|x\|_1$ è una somma finita di numeri reali. Verificare che $\|\cdot\|_1$ è una norma su X .

- (vii) Stabilire se $(X, \|\cdot\|_1)$ è di Banach.
- (viii) Calcolare $\|A\|_{\text{op}}, \|B\|_{\text{op}}$ rispetto alla norma $\|\cdot\|_1$.

Esercizio 15.8. (*Spazio dei polinomi*). Sia X l'insieme dei polinomi in una variabile a coefficienti reali,

$$X = \{p(\lambda) = a_0 + a_1\lambda + \dots + a_n\lambda^n : n \in \mathbb{N}, a_0, \dots, a_n \in \mathbb{R}\},$$

sia

$$\|p\| = \max\{|p(\lambda)| : \lambda \in [0, 1]\},$$

e sia $D : X \rightarrow X$ l'operatore di derivazione $p(\lambda) \mapsto p'(\lambda)$.

- (i) Stabilire se $(X, \|\cdot\|)$ è spazio di Banach.
- (ii) Stabilire se D è iniettiva e/o suriettiva.
- (iii) Calcolare $\|D\|_{\text{op}}$.

Proposizione 15.9. (Continuità per funzioni lineari). *Siano X, Y spazi vettoriali normati, $A : X \rightarrow Y$ una funzione lineare. Sono equivalenti:*

- (i) A è continua in ogni punto $p \in X$;
- (ii) A è continua nel punto $p = 0$;
- (iii) $\|A\|_{\text{op}} < \infty$.

Dimostrazione. (i) \Rightarrow (ii). Ovvio.

(ii) \Rightarrow (i). Sia $p \in X$, e sia $\varepsilon > 0$. Per ipotesi, esiste $\delta > 0$ tale che $\|Ah\| \leq \varepsilon$ per ogni $h \in X$, $\|h\| \leq \delta$. Allora per ogni $x \in X$ nella palla $\|x - p\| \leq \delta$ si ha

$$\|Ax - Ap\| = \|A(x - p)\| \leq \varepsilon.$$

Quindi A è continua in p .

(iii) \Rightarrow (ii). Per ipotesi, e per definizione di $\|A\|_{\text{op}}$, si ha $\|Ax\| \leq C\|x\|$ per ogni $x \in X$, con $C = \|A\|_{\text{op}}$. Sia $\varepsilon > 0$. Per ogni $x \in X$ nella palla $\|x\| \leq \delta$, con $\delta := \varepsilon/(1 + C)$, si ha

$$\|Ax\| \leq C\|x\| \leq C\delta \leq \varepsilon.$$

Quindi A è continua in zero.

(ii) \Rightarrow (iii). Per ipotesi, per ogni ε esiste δ tale che $\|Ax\| \leq \varepsilon$ per ogni $\|x\| \leq \delta$. In particolare, scegliendo $\varepsilon = 1$, esiste $\delta > 0$ tale che

$$\|Ax\| \leq 1 \quad \forall x \in X, \|x\| \leq \delta. \quad (15.8)$$

Sia ora $x \in X, x \neq 0$. Definiamo

$$\lambda := \frac{\delta}{\|x\|}, \quad z := \lambda x.$$

Allora $\|z\| = |\lambda|\|x\| = \delta$, quindi, per (15.8),

$$\|Az\| \leq 1.$$

D'altra parte

$$\|Az\| = \|A(\lambda x)\| = \|\lambda Ax\| = |\lambda|\|Ax\| = \frac{\delta}{\|x\|}\|Ax\|,$$

da cui

$$\delta \frac{\|Ax\|}{\|x\|} \leq 1.$$

Questo vale per ogni $x \in X, x \neq 0$, quindi $\|A\|_{\text{op}} \leq 1/\delta < \infty$. \square

Definizione 15.10. (*Spazio degli operatori lineari e continui*). Siano X, Y spazi vettoriali normati. Definiamo

$$\begin{aligned} \mathcal{L}(X, Y) &:= \{A : X \rightarrow Y, A \text{ lineare e continua}\} \\ &= \{A : X \rightarrow Y, A \text{ lineare, } \|A\|_{\text{op}} < \infty\}. \end{aligned}$$

L'insieme $\mathcal{L}(X, Y)$ è uno spazio vettoriale: la somma $A + B$ e il prodotto λA per uno scalare sono

$$(A + B)x := Ax + Bx, \quad (\lambda A)x := \lambda Ax \quad \forall x \in X.$$

Con la norma operatoriale

$$\|A\|_{\mathcal{L}(X, Y)} := \|A\|_{\text{op}},$$

$\mathcal{L}(X, Y)$ è uno spazio vettoriale normato.

Lemma 15.11. ($\mathcal{L}(X, Y)$ è di Banach). *Siano X, Y spazi vettoriali normati. Se Y è spazio di Banach, allora $(\mathcal{L}(X, Y), \|\cdot\|_{\mathcal{L}(X, Y)})$ è spazio di Banach.*

Esercizio 15.12. Dimostrare il Lemma 15.11.

Osservazione 15.13. (*Compattezza della palla unitaria*). Una differenza fondamentale tra gli spazi vettoriali di dimensione finita e quelli di dimensione infinita riguarda la compattezza della palla: negli spazi di dimensione finita, la palla unitaria $\{x \in X : \|x\| \leq 1\}$ è compatta, essendo chiusa e limitata (Heine-Borel); in spazi di dimensione infinita, invece, esiste sempre una successione (x_n) tale che

$$\|x_n\| = 1 \quad \forall n, \quad \|x_n - x_k\| \geq 1 \quad \forall n \neq k$$

(successione limitata che non ha alcuna sottosuccessione convergente), e la palla unitaria è chiusa, è limitata, e non è (mai) compatta. La dimostrazione di questo fatto generale è fuori programma; un esempio si trova nell'Esercizio 15.14.

Esercizio 15.14. (*Palla non compatta*). Per $n = 1, 2, \dots$, sia $f_n : [0, 1] \rightarrow \mathbb{R}$ la funzione continua

$$f_n(x) = \begin{cases} 0 & \text{se } x \in [0, \frac{1}{n+1}], \\ n(n+1)x - n & \text{se } x \in [\frac{1}{n+1}, \frac{1}{n}], \\ 1 & \text{se } x \in [\frac{1}{n}, 1]. \end{cases}$$

- (i) Disegnare il grafico di f_n .
- (ii) Verificare che $\|f_n\|_\infty = 1$ per ogni n , $\|f_n - f_k\|_\infty = 1$ per ogni $n \neq k$, dove $\|\cdot\|_\infty$ è la norma del sup su $[0, 1]$.
- (iii) Dedurre che la palla unitaria dello spazio di Banach $(C^0([0, 1], \mathbb{R}), \|\cdot\|_\infty)$ non è compatta.

16 Lezione 16

Ricordiamo la definizione di convergenza e convergenza totale per le serie di vettori.

Definizione 16.1. (*Serie convergente e totalmente convergente*). Sia X uno spazio vettoriale normato, e sia (x_n) una successione di elementi di X .

La serie $\sum_{n=0}^{\infty} x_n$ si dice *totalmente convergente* se la serie numerica $\sum \|x_n\|$ converge (in \mathbb{R}).

La serie $\sum_{n=0}^{\infty} x_n$ si dice *convergente* se la successione delle sue somme parziali converge (in X), cioè se esiste un elemento $z \in X$ tale che $\|z - \sum_{k=0}^n x_k\| \rightarrow 0$ per $n \rightarrow \infty$.

Proposizione 16.2. (*Completezza e convergenza totale delle serie*). *Sia X uno spazio vettoriale normato. Allora X è completo se e solo se ogni serie totalmente convergente è convergente.*

Esercizio 16.3. Dimostrare la Proposizione 16.2.

Suggerimento: per provare l'implicazione \Leftarrow), che è la più difficile delle due, si ragioni per sottosuccessioni; si può usare il fatto che se una successione di Cauchy ha un'estratta convergente, allora l'intera successione converge allo stesso limite.

Inverso di un operatore lineare e serie di Neumann. In \mathbb{R} possiamo ottenere l'inverso di un numero reale che sia abbastanza vicino a 1 tramite la serie geometrica: se $|p| < 1$, allora l'inverso del numero reale $1 - p$ è la somma della serie geometrica di ragione p ,

$$\frac{1}{1-p} = \sum_{n=0}^{\infty} p^n, \quad p \in \mathbb{R}, |p| < 1.$$

La convergenza delle serie totalmente convergenti (Proposizione 16.2) permette di trasferire agli spazi di Banach (anche di dimensione infinita) l'idea di usare la serie geometrica per costruire l'inverso. La serie geometrica, quando p non è un numero reale, ma un operatore lineare tra spazi di Banach, si chiama *serie di Neumann*.

Lemma 16.4. (*Norma operatoriale*). *Siano X, Y spazi vettoriali normati, e siano $A, B \in \mathcal{L}(X, Y)$. Si ha allora*

- (i) $\|Ax\| \leq \|A\|_{\text{op}} \|x\|$ per ogni $x \in X$;
- (ii) $\|AB\|_{\text{op}} \leq \|A\|_{\text{op}} \|B\|_{\text{op}}$;
- (iii) $\|A^n\|_{\text{op}} \leq \|A\|_{\text{op}}^n$ per ogni $n \in \mathbb{N}$.

Dimostrazione. Identica alla dimostrazione del Lemma 12.15. □

Proposizione 16.5. (*Invertibilità e serie di Neumann*). *Sia X spazio di Banach, $P \in \mathcal{L}(X, X)$, con $\|P\|_{\text{op}} < 1$. Allora $I - P$ è invertibile, e il suo inverso è*

$$(I - P)^{-1} = \sum_{n=0}^{\infty} P^n.$$

Inoltre

$$\|(I - P)^{-1}\|_{\text{op}} \leq \frac{1}{1 - \|P\|_{\text{op}}}.$$

Dimostrazione. Si ha $\|P^n\|_{\text{op}} \leq \|P\|_{\text{op}}^n$ dal Lemma 16.4, e $\sum_{n=0}^{\infty} \|P\|_{\text{op}}^n$ (serie geometrica di ragione $\|P\|_{\text{op}}$) converge. Quindi, per confronto di serie numeriche, la serie numerica $\sum_{n=0}^{\infty} \|P^n\|_{\text{op}}$ converge in \mathbb{R} , cioè la serie $\sum_{n=0}^{\infty} P^n$ converge totalmente.

Dalla Proposizione 16.2, la serie $\sum_{n=0}^{\infty} P^n$ converge in X , cioè la successione (S_n) delle somme parziali $S_n := \sum_{k=0}^n P^k$ converge in X ad un certo limite, che indichiamo $S := \lim_{(n \rightarrow \infty)} S_n = \sum_{k=0}^{\infty} P^k$.

Per ogni n si ha

$$(I - P)S_n = (I - P)(I + P + \dots + P^n) = I - P^{n+1},$$

e anche $S_n(I - P) = I - P^{n+1}$. Passando al limite per $n \rightarrow \infty$, S_n converge a S , P_n converge a zero perché $\|P^n\|_{\text{op}} \leq \|P\|_{\text{op}}^n \rightarrow 0$. Si ottiene dunque

$$(I - P)S = I, \quad S(I - P) = I.$$

Questo prova che S è inverso destro e inverso sinistro di $I - P$.

Per disuguaglianza triangolare, $\|S_n\|_{\text{op}} \leq \sum_{k=0}^n \|P\|_{\text{op}}^k$, e passando al limite si ottiene la stima dell'enunciato. \square

Nella prossima proposizione assumiamo che l'operatore $A : X \rightarrow Y$ sia invertibile, con inverso $A^{-1} : Y \rightarrow X$. Indichiamo $\|\cdot\|_{\text{op}}$ la norma appropriata per ciascun operatore, cioè indichiamo

$$\|A\|_{\text{op}} = \|A\|_{\mathcal{L}(X,Y)}, \quad \|A^{-1}\|_{\text{op}} = \|A^{-1}\|_{\mathcal{L}(Y,X)};$$

simile notazione per $B, A + B, (A + B)^{-1}$.

Proposizione 16.6. (Inverso di $A + B$ per serie di Neumann). *Siano X, Y spazi di Banach, $A, B \in \mathcal{L}(X, Y)$. Sia A invertibile, sia $A^{-1} \in \mathcal{L}(Y, X)$ il suo inverso, e sia*

$$\|A^{-1}\|_{\text{op}} \|B\|_{\text{op}} < 1.$$

Allora $A + B$ è invertibile, $(A + B)^{-1} \in \mathcal{L}(Y, X)$, e si ha

$$\begin{aligned} (A + B)^{-1} &= \sum_{n=0}^{\infty} (-A^{-1}B)^n A^{-1} \\ &= A^{-1} \sum_{n=0}^{\infty} (-BA^{-1})^n \\ &= A^{-1} - A^{-1}BA^{-1} + A^{-1}BA^{-1}BA^{-1} - \dots \end{aligned}$$

Inoltre

$$\|(A + B)^{-1}\|_{\text{op}} \leq \frac{\|A^{-1}\|_{\text{op}}}{1 - \|A^{-1}\|_{\text{op}} \|B\|_{\text{op}}}.$$

In particolare, se $\|A^{-1}\|_{\text{op}} \|B\|_{\text{op}} \leq 1/2$, allora

$$\|(A + B)^{-1}\|_{\text{op}} \leq 2\|A^{-1}\|_{\text{op}}.$$

Dimostrazione. Indichiamo $I_X : X \rightarrow X$ e $I_Y : Y \rightarrow Y$ la funzione identità (in X e in Y rispettivamente). Dal Lemma 16.4 si ha

$$\|A^{-1}B\|_{\text{op}} \leq \|A^{-1}\|_{\text{op}} \|B\|_{\text{op}} < 1, \quad \|BA^{-1}\|_{\text{op}} \leq \|A^{-1}\|_{\text{op}} \|B\|_{\text{op}} < 1.$$

Dalla Proposizione 16.5 con $P = -A^{-1}B$ segue che $I_X - A^{-1}B \in \mathcal{L}(X, X)$ è invertibile, con inverso

$$I_X - A^{-1}B = \sum_{n=0}^{\infty} (-A^{-1}B)^n;$$

analogamente, dalla Proposizione 16.5 con $P = -BA^{-1}$ segue che $I_Y - BA^{-1} \in \mathcal{L}(Y, Y)$ è invertibile, con inverso

$$I_Y - BA^{-1} = \sum_{n=0}^{\infty} (-BA^{-1})^n.$$

Mettendo A in evidenza come fattore a sinistra nella somma $A+B = A(I_X + A^{-1}B)$, si ha

$$\begin{aligned} (A+B)^{-1} &= [A(I_X + A^{-1}B)]^{-1} \\ &= (I_X + A^{-1}B)^{-1}A^{-1} \\ &= \sum_{n=0}^{\infty} (-A^{-1}B)^n A^{-1}; \end{aligned}$$

scrivendo A come fattore a destra nella somma $A+B = (I_Y + BA^{-1})A$, si ha

$$\begin{aligned} (A+B)^{-1} &= [(I_Y + BA^{-1})A]^{-1} \\ &= A^{-1}(I_Y + BA^{-1})^{-1} \\ &= A^{-1} \sum_{n=0}^{\infty} (-BA^{-1})^n. \end{aligned}$$

La stima di $\|(A+B)^{-1}\|_{\text{op}}$ si ottiene per confronto con la serie geometrica, usando la disuguaglianza $\|(-BA^{-1})^n\| \leq \|BA^{-1}\|^n \leq (\|B\|\|A\|)^n$. \square

Corollario 16.7. (Iso(X, Y) è aperto). *Siano X, Y spazi di Banach. L'insieme*

$$\text{Iso}(X, Y) := \{A \in \mathcal{L}(X, Y) : A \text{ invertibile, } A^{-1} \in \mathcal{L}(Y, X)\}$$

è un aperto di $\mathcal{L}(X, Y)$.

Dimostrazione. Sia $A \in \text{Iso}(X, Y)$; sia $r := 1/\|A^{-1}\|_{\text{op}}$. Per ogni $B \in \mathcal{L}(X, Y)$ con norma $\|B\|_{\text{op}} < r$ si ha

$$\|A^{-1}\|_{\text{op}}\|B\|_{\text{op}} < 1.$$

Quindi, dalla Proposizione 16.6, $A+B$ è invertibile, con $(A+B)^{-1} \in \mathcal{L}(Y, X)$, e dunque $A+B \in \text{Iso}(X, Y)$. Perciò la palla in $\mathcal{L}(X, Y)$ di centro A e raggio r , cioè

$$\{T \in \mathcal{L}(X, Y) : \|T - A\|_{\text{op}} < r\} = \{A+B : B \in \mathcal{L}(X, Y), \|B\|_{\text{op}} < r\}$$

è contenuta in $\text{Iso}(X, Y)$. \square

Proposizione 16.8. (Continuità dell'operatore di inversione). *Siano X, Y spazi di Banach, e sia Inv l'operatore di inversione che a ogni A invertibile associa il suo inverso A^{-1} ,*

$$\text{Inv} : \text{Iso}(X, Y) \rightarrow \mathcal{L}(Y, X), \quad \text{Inv}(A) = A^{-1}.$$

Allora Inv è una funzione continua nell'aperto $\text{Iso}(X, Y)$.

Dimostrazione. Sia $A \in \text{Iso}(X, Y)$. Sia $B \in \mathcal{L}(X, Y)$, con $\|B\|_{\text{op}}\|A^{-1}\|_{\text{op}} \leq 1/2$. Dall'ultima affermazione della Proposizione 16.6, $A + B \in \text{Iso}(X, Y)$, con

$$\|(A + B)^{-1}\|_{\text{op}} \leq 2\|A^{-1}\|_{\text{op}}. \quad (16.1)$$

Scriviamo la differenza degli inversi, mettendo in evidenza $(A + B)^{-1}$ a sinistra:

$$\begin{aligned} (A + B)^{-1} - A^{-1} &= (A + B)^{-1} (I - (A + B)A^{-1}) \\ &= (A + B)^{-1} (I - I - BA^{-1}) \\ &= -(A + B)^{-1} BA^{-1}. \end{aligned} \quad (16.2)$$

Quindi stimiamo la differenza degli inversi, usando l'identità (16.2), la proprietà (ii) del Lemma 16.4 e la stima (16.1):

$$\begin{aligned} \|(A + B)^{-1} - A^{-1}\|_{\text{op}} &= \|(A + B)^{-1} BA^{-1}\|_{\text{op}} \\ &\leq \|(A + B)^{-1}\|_{\text{op}} \|B\|_{\text{op}} \|A^{-1}\|_{\text{op}} \\ &\leq 2\|A^{-1}\|_{\text{op}}^2 \|B\|_{\text{op}}. \end{aligned}$$

Questo prova che

$$\|\text{Inv}(A + B) - \text{Inv}(A)\|_{\text{op}} = \|(A + B)^{-1} - A^{-1}\|_{\text{op}} \rightarrow 0 \quad \text{per } \|B\|_{\text{op}} \rightarrow 0,$$

cioè la continuità di Inv nel punto A . □

Differenziabilità. Riprendiamo la definizione di differenziabilità di Analisi 2, valida per funzioni definite su aperti di \mathbb{R}^n , con l'intento di estenderla a funzioni definite su aperti di spazi di Banach. Ricordiamo brevemente la definizione in \mathbb{R}^n : dati un aperto $\Omega \subseteq \mathbb{R}^n$, un punto $p \in \Omega$, e una funzione $f : \Omega \rightarrow \mathbb{R}^m$, diciamo che f è differenziabile in p se esiste una matrice $A \in \text{Mat}_{m \times n}(\mathbb{R})$ tale che

$$\lim_{h \rightarrow 0} \frac{f(p + h) - f(p) - Ah}{|h|} = 0,$$

cioè tale che

$$f(p + h) = f(p) + Ah + o(|h|) \quad \text{per } |h| \rightarrow 0;$$

in tal caso, si prova che la matrice A è la matrice jacobiana di f in p , che si indica $f'(p)$, $Df(p)$, $df(p)$, $J_f(p)$, $\partial f(p)/\partial x$, o simile.

Alla matrice $A \in \text{Mat}_{m \times n}(\mathbb{R})$ corrisponde la *funzione lineare*

$$\mathbb{R}^n \rightarrow \mathbb{R}^m, \quad h \mapsto Ah,$$

e la corrispondenza tra matrici e funzioni lineari è una biiezione: ogni funzione lineare di \mathbb{R}^n in \mathbb{R}^m si rappresenta in modo unico come prodotto riga per colonna per una matrice di $\text{Mat}_{m \times n}(\mathbb{R})$. Quindi la differenziabilità di Analisi 2 si può riformulare anche così: f è differenziabile in p se esiste una funzione lineare $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tale che $f(p + h) = f(p) + Ah + o(\|h\|)$ per $h \rightarrow 0$.

Definizione 16.9. (*Differenziabilità*). Siano X, Y spazi di Banach, sia $\Omega \subseteq X$ un aperto, $p \in \Omega$, e sia $f : \Omega \rightarrow Y$. Diciamo che f è *differenziabile in p* se esiste $A \in \mathcal{L}(X, Y)$ tale che

$$f(p + h) = f(p) + Ah + o(\|h\|_X) \quad \text{per } h \rightarrow 0,$$

cioè tale che

$$\frac{\|f(p+h) - f(p) - Ah\|_Y}{\|h\|_X} \rightarrow 0 \quad \text{per } h \rightarrow 0,$$

cioè tale che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che

$$\frac{\|f(p+h) - f(p) - Ah\|_Y}{\|h\|_X} \leq \varepsilon \quad \forall h \in X, h \neq 0, \|h\|_X \leq \delta,$$

cioè tale che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che

$$\|f(p+h) - f(p) - Ah\|_Y \leq \varepsilon \|h\|_X \quad \forall h \in X, \|h\|_X \leq \delta. \quad (16.3)$$

In tal caso l'operatore $A \in \mathcal{L}(X, Y)$ si indica

$$f'(p), \quad \text{oppure } Df(p), \quad \text{oppure } df(p),$$

e si dice *differenziale di f nel punto p* , oppure *derivata di f in p* , oppure *derivata di Fréchet di f in p* .

Il vettore $Ah \in Y$ si indica

$$f'(p)h, \quad f'(p)[h], \quad Df(p)h, \quad Df(p)[h],$$

o simile, e si chiama *differenziale di f in p in direzione h* oppure *derivata di f in p in direzione h* .

Diciamo che f è *differenziabile in Ω* se è differenziabile in ogni punto $p \in \Omega$.

Se f è differenziabile in Ω , si dice *differenziale di f* la funzione

$$f' : \Omega \rightarrow \mathcal{L}(X, Y), \quad p \mapsto f'(p).$$

Osservazione 16.10. (*Casi $h = 0, h \neq 0$*). A differenza delle formule in cui $\|h\|$ sta al denominatore, nella (16.3) non si richiede che $h \neq 0$ (per $h = 0$ la disuguaglianza (16.3) è banalmente vera).

Osservazione 16.11. (*Codominio del differenziale*). In generale, f ed f' non hanno lo stesso codominio. Se $f : \Omega \subseteq X \rightarrow Y$, il codominio di f' non è Y , ma lo spazio $\mathcal{L}(X, Y)$ delle funzioni lineari e continue di X in Y .

D'altra parte questo è un fatto noto da Analisi 2: se consideriamo il caso di una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ (quindi $X = \mathbb{R}^n, Y = \mathbb{R}^m$), il suo differenziale in p è la matrice jacobiana $Df(p) \in \text{Mat}_{m \times n}(\mathbb{R})$, che non è un elemento dello spazio \mathbb{R}^m , ma una matrice $m \times n$ che rappresenta una funzione lineare e continua da \mathbb{R}^n ad \mathbb{R}^m .

Definizione 16.12. (*Funzioni C^1*). Siano X, Y spazi di Banach, $\Omega \subseteq X$ un aperto, $f : \Omega \rightarrow Y$ una funzione. Diciamo che f è *di classe C^1 nell'aperto Ω* , e scriviamo $f \in C^1(\Omega)$ oppure $f \in C^1(\Omega, Y)$, se f è differenziabile in Ω e il differenziale $f' : \Omega \rightarrow \mathcal{L}(X, Y), p \mapsto f'(p)$ è una funzione continua.

Esercizio 16.13. Sia $X = C^0([a, b], \mathbb{R})$ munito della sup-norma su $[a, b]$.

(i) Sia $f : X \rightarrow X, f(u) = u^2$. Calcolare $f'(u)[h]$ e mostrare che $f \in C^1(X)$.

(ii) Stessa domanda per la funzione $f : X \rightarrow X, f(u)(t) = u(t)u(1-t)$.

(iii) Stessa domanda per la funzione $f : X \rightarrow \mathbb{R}, f(u) = \int_0^1 u(t/2) \sin(u(t)) dt$.

Esercizio 16.14. Siano X, Y spazi di Banach, e sia $f : X \rightarrow Y$, $f(x) = Ax$ con $A \in \mathcal{L}(X, Y)$ (funzione lineare continua). Provare che allora

$$f'(p) = A \quad \forall p \in X,$$

cioè $f'(p)[h] = Ah$ per ogni $p, h \in X$.

Come in dimensione finita, abbiamo le seguenti proprietà.

Proposizione 16.15. *La differenziabilità implica la continuità. La somma, il prodotto per scalari, la composizione di funzioni differenziabili è differenziabile, con le usuali regole di calcolo*

$$\begin{aligned} (f + g)'(p)[h] &= f'(p)[h] + g'(p)[h], \\ (\lambda f)'(p)[h] &= \lambda f'(p)[h], \\ (f \circ g)'(p)[h] &= f'(g(p))[g'(p)[h]]. \end{aligned}$$

Dimostrazione. Adattare agli spazi di Banach le corrispondenti dimostrazioni in \mathbb{R}^n . □

Proposizione 16.16. (Inv è di classe C^1). *Siano X, Y spazi di Banach. Ricordiamo che $\text{Iso}(X, Y)$ è un aperto di $\mathcal{L}(X, Y)$. L'operatore di inversione*

$$\text{Inv} : \text{Iso}(X, Y) \rightarrow \mathcal{L}(Y, X), \quad \text{Inv}(A) = A^{-1}$$

è di classe $C^1(\text{Iso}(X, Y), \mathcal{L}(Y, X))$, e il suo differenziale nel punto A in direzione B è

$$\text{Inv}'(A)[B] = -A^{-1}BA^{-1} = -\text{Inv}(A)B\text{Inv}(A).$$

Dimostrazione. Ricordiamo l'identità (16.2), cioè

$$(A + B)^{-1} - A^{-1} = -(A + B)^{-1}BA^{-1}. \quad (16.4)$$

Portiamo A^{-1} da sinistra a destra,

$$(A + B)^{-1} = A^{-1} - (A + B)^{-1}BA^{-1}. \quad (16.5)$$

Usiamo (16.5) per sostituire $(A + B)^{-1}$ nel termine a destra nella (16.4), e otteniamo

$$\begin{aligned} (A + B)^{-1} - A^{-1} &= -\left(A^{-1} - (A + B)^{-1}BA^{-1}\right)BA^{-1} \\ &= \underbrace{-A^{-1}BA^{-1}}_{L[B]} + \underbrace{(A + B)^{-1}BA^{-1}BA^{-1}}_{Q(B)}. \end{aligned} \quad (16.6)$$

La funzione

$$L : \mathcal{L}(X, Y) \rightarrow \mathcal{L}(Y, X), \quad B \mapsto L[B] := -A^{-1}BA^{-1}$$

è lineare. È anche continua perché

$$\|L[B]\|_{\text{op}} \leq \|A^{-1}\|_{\text{op}}^2 \|B\|_{\text{op}} \quad \forall B \in \mathcal{L}(X, Y),$$

da cui

$$\|L\|_{\text{op}} = \sup_{B \neq 0} \frac{\|L[B]\|_{\mathcal{L}(Y, X)}}{\|B\|_{\mathcal{L}(X, Y)}} \leq \|A^{-1}\|_{\text{op}}^2 < \infty.$$

La funzione

$$Q : \mathcal{L}(X, Y) \rightarrow \mathcal{L}(Y, X), \quad Q(B) := (A + B)^{-1}BA^{-1}BA^{-1}$$

contiene due fattori B ; usando il fatto (Proposizione 16.6) che

$$\|(A + B)^{-1}\|_{\text{op}} \leq 2\|A^{-1}\|_{\text{op}} \quad \text{per } \|B\|_{\text{op}} < 1/\|A^{-1}\|_{\text{op}},$$

si ha

$$\|Q(A, B)\|_{\text{op}} \leq \|(A + B)^{-1}\|_{\text{op}}\|A^{-1}\|_{\text{op}}^2\|B\|_{\text{op}}^2 \leq 2\|A^{-1}\|_{\text{op}}^3\|B\|_{\text{op}}^2,$$

quindi

$$\frac{\|Q(A, B)\|_{\text{op}}}{\|B\|_{\text{op}}} \rightarrow 0 \quad \text{per } B \rightarrow 0.$$

Abbiamo provato che

$$\text{Inv}(A + B) = \text{Inv}(A) + L[B] + o(\|B\|) \quad \text{per } B \rightarrow 0,$$

cioè che Inv è differenziabile in A , e che il suo differenziale nel punto A in direzione B è $L[B]$. \square

Definizione 16.17. (*Derivate parziali*). Siano X, Y, Z spazi di Banach, $\Omega \subseteq X \times Y$ un aperto, $p = (x_0, y_0) \in \Omega$, e sia $f : \Omega \rightarrow Z$ una funzione. Sia $r > 0$ tale che $B_X(x_0, r) \times B_Y(y_0, r) \subseteq \Omega$.

Diciamo che f è *differenziabile in p rispetto alla variabile x* se la funzione

$$f(\cdot, y_0) : B_X(x_0, r) \rightarrow Z, \quad x \mapsto f(x, y_0)$$

è differenziabile in x_0 ; in tal caso indichiamo $\partial_x f(p) \in \mathcal{L}(X, Z)$ il suo differenziale, e, per ogni $h \in X$, chiamiamo $\partial_x f(p)[h]$ la *derivata parziale di f rispetto a x nel punto p in direzione h* .

Analogamente, diciamo che f è *differenziabile in p rispetto alla variabile y* se la funzione

$$f(x_0, \cdot) : B_Y(y_0, r) \rightarrow Z, \quad y \mapsto f(x_0, y)$$

è differenziabile in y_0 ; in tal caso indichiamo $\partial_y f(p) \in \mathcal{L}(Y, Z)$ il suo differenziale, e, per ogni $h \in Y$, chiamiamo $\partial_y f(p)[h]$ la *derivata parziale di f rispetto a y nel punto p in direzione h* .

Lemma 16.18. (La differenziabilità implica l'esistenza delle derivate parziali). *Assumiamo le ipotesi della Definizione 16.17. Se f è differenziabile in $p = (x_0, y_0)$, allora esistono le derivate parziali $\partial_x f(p)$, $\partial_y f(p)$, e si ha*

$$f'(p)[h] = \partial_x f(p)[v] + \partial_y f(p)[w] \quad \forall h = (v, w) \in X \times Y.$$

Esercizio 16.19. Dimostrare il Lemma 16.18, adattando agli spazi di Banach la dimostrazione fatta in Analisi 2.

Teorema del valor medio. Per funzioni a valori in \mathbb{R}^N , con $N \geq 2$, non vale il Teorema di Lagrange (vedi Osservazione 3.13). Per ottenere il Teorema del valor medio in \mathbb{R}^N (Teorema 3.11), invece del Teorema di Lagrange abbiamo usato

il Teorema fondamentale del calcolo integrale, come nella dimostrazione del criterio di lipschitzianità (Proposizione 3.10).

Per estendere gli stessi passaggi agli spazi di Banach, dovremmo definire l'integrale $\int_a^b f(t) dt$ su un intervallo $[a, b] \subset \mathbb{R}$ di una funzione $f : [a, b] \rightarrow Y$ con Y spazio di Banach; definire questo tipo di integrali è possibile, ma abbiamo a disposizione un metodo più semplice che evita gli integrali, ragionando in modo non troppo diverso da quanto fatto nelle varie dimostrazioni di lemmi e teoremi sulle equazioni differenziali.

Proposizione 16.20. (Stima di Lipschitz senza Lagrange e senza integrale). *Sia Y uno spazio vettoriale normato, siano $a, b \in \mathbb{R}$, con $a < b$, e sia $f : [a, b] \rightarrow Y$ una funzione continua in $[a, b]$, differenziabile in (a, b) , con*

$$\|f'(t)\|_{\mathcal{L}(\mathbb{R}, Y)} \leq L \quad \forall t \in (a, b).$$

Allora

$$\|f(b) - f(a)\|_Y \leq L(b - a).$$

Dimostrazione nella prossima lezione.

17 Lezione 17

Osservazione 17.1. (*Derivata di funzione di variabile reale a valori in spazio vettoriale normato*). Sia $(a, b) \subset \mathbb{R}$ un intervallo reale, $p \in (a, b)$, Y uno spazio vettoriale normato, $f : (a, b) \rightarrow Y$ una funzione derivabile in p . Dalla Definizione 16.9 di differenziabilità, con $X = \mathbb{R}$, si ha

$$\frac{\|f(p+h) - f(p) - f'(p)[h]\|_Y}{|h|} \rightarrow 0 \quad \text{per } h \rightarrow 0; \quad (17.1)$$

la “norma $\|h\|_X$ del vettore h ”, in questo caso, è semplicemente il modulo $|h|$ del numero reale $h \in \mathbb{R}$.

Poiché $1/h$ è uno scalare, dalla proprietà $\|\lambda y\| = |\lambda| \|y\|$ della norma si ha

$$\begin{aligned} \frac{\|f(p+h) - f(p) - f'(p)[h]\|_Y}{|h|} &= \left| \frac{1}{h} \right| \|f(p+h) - f(p) - f'(p)[h]\|_Y \\ &= \left\| \frac{1}{h} (f(p+h) - f(p) - f'(p)[h]) \right\|_Y. \end{aligned}$$

Il differenziale $f'(p) \in \mathcal{L}(X, Y) = \mathcal{L}(\mathbb{R}, Y)$ è una funzione lineare, h è uno scalare, quindi

$$f'(p)[h] = f'(p)[h \cdot 1] = hf'(p)[1],$$

da cui

$$\frac{1}{h} (f(p+h) - f(p) - f'(p)[h]) = \frac{f(p+h) - f(p)}{h} - f'(p)[1].$$

La (17.1) si riscrive dunque come

$$\left\| \frac{f(p+h) - f(p)}{h} - f'(p)[1] \right\|_Y \rightarrow 0 \quad \text{per } h \rightarrow 0,$$

e dunque $f'(p)[1]$ è il limite in Y del rapporto incrementale,

$$f'(p)[1] = \lim_{h \rightarrow 0} \frac{f(p+h) - f(p)}{h}. \quad (17.2)$$

Si ha inoltre

$$\begin{aligned} \|f'(p)\|_{\mathcal{L}(\mathbb{R}, Y)} &= \sup_{h \neq 0} \frac{\|f'(p)[h]\|_Y}{|h|} = \sup_{h \neq 0} \frac{\|hf'(p)[1]\|_Y}{|h|} = \sup_{h \neq 0} \frac{|h| \|f'(p)[1]\|_Y}{|h|} \\ &= \|f'(p)[1]\|_Y. \end{aligned} \quad (17.3)$$

Esempio. Consideriamo $Y = \mathbb{R}^3$, $p = 0$,

$$f : \mathbb{R} \rightarrow \mathbb{R}^3, \quad f(t) = \begin{pmatrix} \cos t \\ \sin t \\ 2t + t^3 \end{pmatrix}.$$

Si ha dunque

$$f(p+h) - f(p) = f(h) - f(0) = \begin{pmatrix} 0 \\ h \\ 2h \end{pmatrix} + o(h) \quad \text{per } h \rightarrow 0,$$

quindi

$$f'(0)[h] = \begin{pmatrix} 0 \\ h \\ 2h \end{pmatrix} = hf'(0)[1], \quad f'(0)[1] = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

Con un piccolo abuso di notazione rispetto alla Definizione 16.9, possiamo anche scrivere $f'(0) = (0, 1, 2)$ (come si fa, per esempio, in Analisi 2 parlando di curve in \mathbb{R}^n); tuttavia, a rigore, secondo la Definizione 16.9 $f'(0)$ non è il vettore $(0, 1, 2)$, ma è la funzione lineare che ad ogni $h \in \mathbb{R}$ associa il vettore $(0, h, 2h) \in Y = \mathbb{R}^3$.

Dimostrazione della Proposizione 16.20. Sia $\varepsilon > 0$, sia $c \in (a, b)$, e sia

$$E = \{t \in [c, b] : \|f(t) - f(c)\|_Y \leq (L + \varepsilon)(t - c)\}.$$

E è non vuoto perché $c \in E$. Sia

$$p = \sup E.$$

Per definizione di sup, esiste una successione (t_n) di elementi di E che converge a p ; poiché

$$\|f(t_n) - f(c)\|_Y \leq (L + \varepsilon)(t_n - c) \quad \forall n$$

e f è continua, passando al limite per $n \rightarrow \infty$ si ottiene

$$\|f(p) - f(c)\|_Y \leq (L + \varepsilon)(p - c). \quad (17.4)$$

Supponiamo che sia $c \leq p < b$. Poiché $p \in (a, b)$, f è differenziabile in p . Come abbiamo visto nell'Osservazione 17.1, si ha

$$\|f'(p)\|_{\mathcal{L}(\mathbb{R}, Y)} = \|f'(p)[1]\|_Y,$$

e $f'(p)[1]$ è il limite in Y del rapporto incrementale

$$f'(p)[1] = \lim_{h \rightarrow 0} \frac{f(p+h) - f(p)}{h} = \lim_{t \rightarrow p} \frac{f(t) - f(p)}{t - p}.$$

Per ipotesi,

$$\|f'(p)\|_{\mathcal{L}(\mathbb{R}, Y)} \leq L < L + \varepsilon.$$

Dalla definizione di limite, esiste allora $\delta > 0$ tale che

$$\frac{\|f(t) - f(p)\|_Y}{t - p} = \left\| \frac{f(t) - f(p)}{t - p} \right\|_Y < L + \varepsilon \quad \forall t \in (p, p + \delta), \quad (17.5)$$

con $p + \delta \leq b$. Dunque abbiamo (17.5), (17.4), cioè

$$\begin{cases} \|f(t) - f(p)\|_Y < (L + \varepsilon)(t - p) & \forall t \in (p, p + \delta), \\ \|f(p) - f(c)\|_Y \leq (L + \varepsilon)(p - c), \end{cases}$$

da cui

$$\begin{aligned} \|f(t) - f(c)\|_Y &\leq \|f(t) - f(p)\|_Y + \|f(p) - f(c)\|_Y \\ &< (L + \varepsilon)(t - p) + (L + \varepsilon)(p - c) = (L + \varepsilon)(t - c). \end{aligned}$$

Perciò $t \in E$ per ogni $t \in (p, p + \delta)$, assurdo perché $p = \sup E$.

Questo prova che $p = b$. Quindi, sostituendo p con b nella (17.4),

$$\|f(b) - f(c)\|_Y \leq (L + \varepsilon)(b - c).$$

Questa disuguaglianza vale per ogni $c \in (a, b)$; passando al limite per $c \rightarrow a$, per continuità, si ottiene

$$\|f(b) - f(a)\|_Y \leq (L + \varepsilon)(b - a).$$

Infine, passando al limite per $\varepsilon \rightarrow 0$, si ottiene la tesi. \square

Teorema 17.2. (Teorema del valor medio in spazi vettoriali normati). *Siano X, Y spazi vettoriali normati, $\Omega \subseteq X$ un aperto, $p \in \Omega$, $h \in X$, con $p + th \in \Omega$ per ogni $t \in [0, 1]$. Sia $f : \Omega \rightarrow Y$ una funzione continua in $p + th$ per ogni $t \in [0, 1]$ e differenziabile in $p + th$ per ogni $t \in (0, 1)$. Allora*

$$\|f(p + h) - f(p)\|_Y \leq \left(\sup_{t \in (0, 1)} \|f'(p + th)\|_{\mathcal{L}(X, Y)} \right) \|h\|_X.$$

Dimostrazione. Sia

$$\varphi : [0, 1] \rightarrow Y, \quad \varphi(t) = f(p + th).$$

Proviamo che φ è differenziabile in ogni $t \in (0, 1)$, con

$$\varphi'(t)[1] = f'(p + th)[h]. \tag{17.6}$$

Per ipotesi, per ogni $t \in (0, 1)$, f è differenziabile in $p + th$, cioè

$$f(p + th + v) = f(p + th) + f'(p + th)[v] + o(\|v\|) \quad \text{per } v \rightarrow 0, v \in X.$$

In particolare, per $v = sh$, $s \in \mathbb{R}$ (dove $h \in X$ è fissato dalle ipotesi), si ha

$$f(p + th + sh) = f(p + th) + f'(p + th)[sh] + o(|s|\|h\|) \quad \text{per } s \rightarrow 0, s \in \mathbb{R},$$

cioè

$$\varphi(t + s) = \varphi(t) + sf'(p + th)[h] + o(|s|) \quad \text{per } s \rightarrow 0.$$

Dunque φ è differenziabile in t , e il suo differenziale è

$$\varphi'(t)[s] = sf'(p + th)[h],$$

cioè (17.6). Perciò, ricordando (17.3), per ogni $t \in (0, 1)$ si ha

$$\begin{aligned} \|\varphi'(t)\|_{\mathcal{L}(\mathbb{R}, Y)} &= \|\varphi'(t)[1]\|_Y \\ &= \|f'(p + th)[h]\|_Y \\ &\leq \|f'(p + th)\|_{\mathcal{L}(X, Y)} \|h\|_X \\ &\leq \sup_{\sigma \in (0, 1)} \|f'(p + \sigma h)\|_{\mathcal{L}(X, Y)} \|h\|_X =: L. \end{aligned}$$

Dalla Proposizione 16.20 si ha $\|\varphi(1) - \varphi(0)\|_Y \leq L$, cioè la tesi. \square

Osservazione 17.3. (Nel Teorema del valor medio basta la derivata direzionale). Nella dimostrazione appena conclusa, per poter applicare la Proposizione 16.20 alla funzione $\varphi(t) = f(p + th)$ in realtà non serve che f sia continua e differenziabile, ma bastano la continuità e la differenziabilità di f rispetto alla sola direzione h , cioè basta che sia $\varphi(t)$ ad essere continua e differenziabile.

È possibile definire una nozione meno restrittiva di differenziabilità (detta differenziabilità *alla Gâteaux*) che corrisponde all'esistenza delle derivate direzionali e non richiede la differenziabilità e nemmeno la continuità.

Lemma delle contrazioni. Il Lemma delle contrazioni è uno strumento semplice e importante per risolvere problemi di punto fisso. È ambientato negli spazi metrici completi; per utilizzarlo negli spazi di Banach, osserviamo che ogni chiuso di uno spazio di Banach è uno spazio metrico completo (vedi Osservazione 17.6).

Definizione 17.4. (Contrazione). Dato uno spazio metrico completo (E, d) , una funzione $f : E \rightarrow E$ si dice una *contrazione di E* se è lipschitziana in E con costante di Lipschitz minore di 1.

Teorema 17.5. (Lemma delle contrazioni di Banach-Caccioppoli). Sia (E, d) uno spazio metrico completo. Sia $0 \leq c < 1$, e sia $f : E \rightarrow E$ tale che

$$d(f(x), f(y)) \leq c d(x, y) \quad \forall x, y \in E.$$

Allora esiste un unico punto fisso di f in E , cioè esiste un unico elemento $p \in E$ tale che

$$f(p) = p.$$

Inoltre per ogni $x_0 \in E$ si ha

$$p = \lim_{n \rightarrow \infty} f^n(x_0), \quad d(p, f^n(x_0)) \leq d(f(x_0), x_0) \frac{c^n}{1 - c},$$

dove $f^2(p) = f(f(p))$ e, ricorsivamente, $f^n(p) = f(f^{n-1}(p))$ per ogni n .

Dimostrazione. Sia $x_0 \in E$. Definiamo ricorsivamente

$$x_{n+1} := f(x_n), \quad n \in \mathbb{N}.$$

Per induzione, si dimostra che

$$x_n = f^n(x_0), \quad d(x_{n+1}, x_n) \leq c^n d(x_1, x_0) \quad \forall n \in \mathbb{N}.$$

Dunque per ogni $n, k \in \mathbb{N}$, con $k \geq 1$, si ha

$$\begin{aligned} d(x_{n+k}, x_n) &\leq d(x_{n+k}, x_{n+k-1}) + \dots + d(x_{n+1}, x_n) = \sum_{j=n}^{n+k-1} d(x_{j+1}, x_j) \\ &\leq \sum_{j=n}^{n+k-1} c^j d(x_1, x_0) \\ &\leq \sum_{j=n}^{\infty} c^j d(x_1, x_0) = \frac{c^n}{1 - c} d(x_1, x_0). \end{aligned}$$

Per ogni $\varepsilon > 0$ esiste \bar{n} tale che

$$\frac{c^n}{1-c} d(x_1, x_0) \leq \varepsilon \quad \forall n \geq \bar{n};$$

quindi $d(x_{n+k}, x_n) \leq \varepsilon$ per ogni $n \geq \bar{n}$, per ogni $k \geq 1$. Questo significa che (x_n) è una successione di Cauchy. Dunque (x_n) converge in E ; indichiamo p il suo limite.

Per costruzione, $x_{n+1} = f(x_n)$; passando al limite per $n \rightarrow \infty$, e usando la continuità di f , si ottiene $p = f(p)$.

Torniamo alla stima

$$d(x_{n+k}, x_n) \leq \frac{c^n}{1-c} d(x_1, x_0) \quad \forall n, k \in \mathbb{N}, k \geq 1;$$

passando al limite per $k \rightarrow \infty$ otteniamo

$$d(p, x_n) \leq \frac{c^n}{1-c} d(x_1, x_0) \quad \forall n \in \mathbb{N}.$$

Proviamo infine l'unicità: se p, x sono punti fissi di f in E , allora, poiché $p = f(p)$, $x = f(x)$, si ha

$$d(p, x) = d(f(p), f(x)) \leq cd(p, x),$$

da cui $(1-c)d(p, x) \leq 0$, cioè $d(p, x) = 0$, cioè $p = x$. □

Osservazione 17.6. (*Lemma delle contrazioni in spazi di Banach*). Se X è uno spazio di Banach ed $E \subseteq X$ è un suo sottoinsieme chiuso, allora E è uno spazio metrico completo: infatti E è uno spazio metrico con la metrica indotta dalla norma $d(x, y) = \|x - y\|$; E è completo perché se una successione di elementi di E è di Cauchy, allora converge in X (perché X è spazio di Banach), e il suo limite appartiene a E (perché E è un chiuso).

Quindi il Lemma delle contrazioni in spazi di Banach si può riformulare così:

Sia X spazio di Banach, $E \subseteq X$ un suo sottoinsieme chiuso; sia $0 \leq c < 1$, e sia $f : E \rightarrow E$ una funzione tale che

$$\|f(x) - f(y)\| \leq c\|x - y\| \quad \forall x, y \in E.$$

Allora esiste un unico $p \in E$ tale che $f(p) = p$. Inoltre per ogni $x_0 \in E$ si ha

$$p = \lim_{n \rightarrow \infty} f^n(x_0), \quad \|p - f^n(x_0)\| \leq \frac{c^n}{1-c} \|f(x_0) - x_0\|.$$

Esercizio 17.7. Dimostrare che esiste un'unica funzione continua $f : [0, 1] \rightarrow [0, 1]$ tale che

$$f(x) = \frac{1}{2} \int_0^1 \sin(x + f(t)) dt \quad \forall x \in [0, 1].$$

Suggerimento: considerare $E = \{f \in C^0([0, 1], \mathbb{R}) : \|f\|_\infty \leq R\}$, palla chiusa nello spazio di Banach $C^0([0, 1], \mathbb{R})$, dove $\|\cdot\|_\infty$ è la sup-norma su $[0, 1]$; scegliere R in modo che f mappi $E \rightarrow E$ e sia contrazione in E .

Teorema della funzione implicita. Vediamo ora il Teorema della funzione implicita in spazi di Banach. L'enunciato è molto simile a quello per funzioni di due variabili reali $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ (Teorema di Dini) trattato in Analisi 2; il Teorema 17.8 estende il Teorema di Dini al caso di dimensione finita qualunque $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^q$ e, più in generale, agli spazi di Banach (anche di dimensione infinita).

Teorema 17.8. (Teorema della funzione implicita). Siano X, Y, Z spazi di Banach, $\Omega \subseteq X \times Y$ un aperto, $(x_0, y_0) \in \Omega$, $f : \Omega \rightarrow Z$ una funzione tale che

- (i) f è continua in Ω ;
- (ii) per ogni $(x, y) \in \Omega$ esiste la derivata parziale $\partial_y f(x, y) \in \mathcal{L}(Y, Z)$, e la funzione $\partial_y f : \Omega \rightarrow \mathcal{L}(Y, Z)$ è continua;
- (iii) $f(x_0, y_0) = 0$;
- (iv) $\partial_y f(x_0, y_0) \in \text{Iso}(Y, Z)$.

Allora esistono due costanti $a, b > 0$ e una funzione

$$g : B(x_0, a) \rightarrow B(y_0, b)$$

tali che

- $B(x_0, a) \times B(y_0, b) \subset \Omega$,
- g è continua in $B(x_0, a)$,
- $g(x_0) = y_0$,
- $f(x, g(x)) = 0$ per ogni $x \in B(x_0, a)$,
- se $(x, y) \in B(x_0, a) \times B(y_0, b)$ e $f(x, y) = 0$, allora $y = g(x)$.

Se inoltre $f \in C^1(\Omega)$, allora $g \in C^1(B(x_0, a))$, con

$$g'(x)[h] = -\left(\partial_y f(x, g(x))\right)^{-1} (\partial_x f)(x, g(x))h \quad \forall x \in B(x_0, a), h \in X.$$

Osservazione 17.9. (Premessa: idea della dimostrazione). Prima di cominciare la dimostrazione vera e propria, vediamo l'idea alla base della costruzione della funzione implicita $y = g(x)$ come soluzione di un problema di punto fisso.

Per ogni x vicino a x_0 cerchiamo y tale che $f(x, y) = 0$; in altre parole, vogliamo risolvere l'equazione $f(x, y) = 0$, in cui y è l'incognita e x ha il ruolo di un parametro. Sappiamo inoltre che per $x = x_0$ l'equazione ha la soluzione $y = y_0$.

Se f fosse una funzione affine in y , cioè della forma

$$f(x, y) = u(x) + v(x)y$$

per certi $u(x) \in Z$, $v(x) \in \mathcal{L}(Y, Z)$, allora sapremmo risolvere l'equazione $f(x, y) = 0$: la soluzione sarebbe

$$y = -v(x)^{-1}u(x),$$

pur di sapere che $v(x)$ è invertibile; per ipotesi, $\partial_y f(x_0, y_0) = v(x_0)$ è invertibile, quindi si potrebbe dimostrare che $v(x)$ rimane invertibile per x sufficientemente vicino a x_0 .

In generale $f(x, y)$ non è affine in y . Possiamo però cercare di avvicinarci alla struttura di funzione affine in y : usando lo sviluppo di Taylor del primo ordine rispetto a y , abbiamo infatti

$$f(x, y) = f(x, y_0) + \partial_y f(x, y_0)(y - y_0) + r(x, y), \quad (17.7)$$

con

$$r(x, y) = o(\|y - y_0\|) \quad \text{per } y \rightarrow y_0.$$

Imitando il caso affine, isoliamo il termine lineare in y , cioè $\partial_y f(x, y_0)y$, nella (17.7): se $\partial_y f(x, y_0)$ è invertibile, allora l'equazione $f(x, y) = 0$ si può riscrivere nella forma

$$y = y_0 - (\partial_y f(x, y_0))^{-1}[f(x, y_0) + r(x, y)], \quad (17.8)$$

che è un problema di punto fisso per l'incognita y , e può essere affrontato con il Lemma delle contrazioni. Usando la (17.7) per sostituire $r(x, y)$, i termini a destra nella (17.8) diventano

$$\begin{aligned} & y_0 - (\partial_y f(x, y_0))^{-1}[f(x, y_0) + r(x, y)] \\ &= y_0 - (\partial_y f(x, y_0))^{-1}[f(x, y) - \partial_y f(x, y_0)(y - y_0)] \\ &= y_0 - (\partial_y f(x, y_0))^{-1}f(x, y) + (y - y_0) \\ &= y - (\partial_y f(x, y_0))^{-1}f(x, y). \end{aligned}$$

Quindi l'equazione (17.8) si riscrive come

$$y = y - (\partial_y f(x, y_0))^{-1}f(x, y). \quad (17.9)$$

Possiamo poi semplificare leggermente la formulazione (17.9) del problema, prendendo $(\partial_y f(x_0, y_0))^{-1}$ (con x_0 fissato) invece di $(\partial_y f(x, y_0))^{-1}$ (con x variabile); per vedere a cosa corrisponde questa modifica nella discussione precedente, ritorniamo alla (17.7).

I coefficienti $f(x, y_0)$, $\partial_y f(x, y_0)$ dello sviluppo di Taylor (17.7) dipendono da x ; per $x = x_0$, per ipotesi, sono

$$f(x_0, y_0) = 0, \quad \partial_y f(x_0, y_0) = \text{invertibile}.$$

Consideriamo dunque una formula di Taylor modificata, in cui i coefficienti $f(x, y_0)$, $\partial_y f(x, y_0)$ della (17.7) sono rimpiazzati dai valori $f(x_0, y_0)$, $\partial_y f(x_0, y_0)$ che si ottengono per $x = x_0$: usando l'identità $f(x_0, y_0) = 0$, scriviamo

$$\begin{aligned} f(x, y) &= f(x_0, y_0) + \partial_y f(x_0, y_0)(y - y_0) + R(x, y) \\ &= A(y - y_0) + R(x, y) \end{aligned}$$

dove $A, R(x, y)$ sono definiti come

$$A := \partial_y f(x_0, y_0), \quad R(x, y) := f(x, y) - A(y - y_0).$$

L'equazione $f(x, y) = 0$ si riscrive come $A(y - y_0) + R(x, y) = 0$, cioè

$$y = y_0 - A^{-1}R(x, y),$$

problema di punto fisso per y . Usando la definizione di $R(x, y)$ si ha poi

$$\begin{aligned} y_0 - A^{-1}R(x, y) &= y_0 - A^{-1}(f(x, y) - A(y - y_0)) \\ &= y_0 - A^{-1}f(x, y) + (y - y_0) \\ &= y - A^{-1}f(x, y). \end{aligned}$$

Quindi l'equazione $f(x, y) = 0$ si riscrive come problema di punto fisso

$$y = y - A^{-1}f(x, y). \quad (17.10)$$

È immediato osservare che il problema $f(x, y) = 0$ equivale alla (17.10),

$$f(x, y) = 0 \quad \Leftrightarrow \quad y = y - A^{-1}f(x, y);$$

inoltre tale equivalenza varrebbe anche con qualunque operatore invertibile messo al posto di A^{-1} , e non ha certo bisogno di tutta la discussione precedente. Ciò che invece non è evidente a prima vista è da dove nasce l'idea di riscrivere il problema in questo modo.

La riscrittura del problema $f(x, y) = 0$ nella forma (17.10) è il punto di partenza della dimostrazione del Teorema 17.8.

Dimostrazione del Teorema 17.8. Indichiamo

$$A := \partial_y f(x_0, y_0),$$

che è invertibile per ipotesi. Per $(x, y) \in \Omega$, definiamo

$$\Phi(x, y) := y - A^{-1}f(x, y). \quad (17.11)$$

Poiché A è invertibile, si ha

$$f(x, y) = 0 \quad \Leftrightarrow \quad y = \Phi(x, y).$$

Vogliamo dunque risolvere il problema di punto fisso $y = \Phi(x, y)$.

Calcoliamo la derivata parziale $\partial_y \Phi(x, y)$: usando la definizione (17.11) di $\Phi(x, y)$, per ogni $(x, y) \in \Omega$, per ogni $h \in Y$ si ha

$$\partial_y \Phi(x, y)[h] = h - A^{-1}\partial_y f(x, y)[h],$$

cioè

$$\partial_y \Phi(x, y) = I - A^{-1}\partial_y f(x, y) \quad \forall (x, y) \in \Omega. \quad (17.12)$$

Per ipotesi, $\partial_y f$ è continua in Ω , quindi, dalla formula (17.12), anche $\partial_y \Phi$ lo è. Nel punto (x_0, y_0) si ha

$$\partial_y \Phi(x_0, y_0) = I - A^{-1}\partial_y f(x_0, y_0) = I - A^{-1}A = 0.$$

Dalla continuità di $\partial_y \Phi$ e dal fatto che $\partial_y \Phi(x_0, y_0) = 0$, per ogni $\varepsilon > 0$ esiste un intorno di (x_0, y_0) contenuto in Ω in cui si ha $\|\partial_y \Phi(x, y)\| \leq \varepsilon$. In particolare, esiste $b > 0$ tale che

$$I_b \times J_b \subset \Omega, \quad \|\partial_y \Phi(x, y)\|_{\mathcal{L}(Y, Z)} \leq \frac{1}{2} \quad \forall (x, y) \in I_b \times J_b, \quad (17.13)$$

dove

$$\begin{aligned} I_b &:= \{x \in X : \|x - x_0\| \leq b\}, \\ J_b &:= \{y \in Y : \|y - y_0\| \leq b\}. \end{aligned}$$

Dal Teorema del valor medio (Teorema 17.2), per ogni $x \in I_b$, $y_1, y_2 \in J_b$ si ha

$$\begin{aligned} \|\Phi(x, y_2) - \Phi(x, y_1)\|_Y &\leq \sup_{t \in [0,1]} \|(\partial_y \Phi)(x, y_1 + t(y_2 - y_1))\|_{\mathcal{L}(Y,Y)} \|y_2 - y_1\|_Y \\ &\leq \frac{1}{2} \|y_2 - y_1\|_Y. \end{aligned} \quad (17.14)$$

Dunque per ogni $x \in I_b$ la funzione $\Phi(x, \cdot)$ è lipschitziana in J_b con costante di Lipschitz minore di 1.

Osserviamo che $\partial_y f(x, y)$ è invertibile: infatti, dalla (17.12), si ha

$$\partial_y f(x, y) = A(I - \partial_y \Phi(x, y)) \quad \forall (x, y) \in \Omega;$$

A è invertibile per ipotesi, mentre, dalla stima (17.13), $(I - \partial_y \Phi(x, y))$ è invertibile per serie di Neumann, con

$$\|(I - \partial_y \Phi(x, y))^{-1}\|_{\mathcal{L}(Y,Y)} \leq 2 \quad \forall (x, y) \in I_b \times J_b,$$

quindi

$$(\partial_y f(x, y))^{-1} = (I - \partial_y \Phi(x, y))^{-1} A^{-1},$$

$$\begin{aligned} \|(\partial_y f(x, y))^{-1}\|_{\mathcal{L}(Z,Y)} &\leq \|(I - \partial_y \Phi(x, y))^{-1}\|_{\mathcal{L}(Y,Y)} \|A^{-1}\|_{\mathcal{L}(Z,Y)} \\ &\leq 2 \|A^{-1}\|_{\mathcal{L}(Z,Y)} \quad \forall (x, y) \in I_b \times J_b. \end{aligned} \quad (17.15)$$

La stima (17.15) verrà usata nell'ultima parte della dimostrazione.

Abbiamo provato che la funzione $\Phi(x, \cdot) : y \mapsto \Phi(x, y)$ è lipschitziana con costante $\leq 1/2$ (vedi (17.14)). Per ottenere una contrazione, proviamo che $\Phi(x, \cdot)$ mappa la palla J_b in sé. Stimiamo la distanza tra $\Phi(x, y)$ e y_0 , centro della palla J_b . Per $(x, y) \in I_b \times J_b$ si ha

$$\begin{aligned} \|\Phi(x, y) - y_0\| &\leq \|\Phi(x, y) - \Phi(x, y_0)\| + \|\Phi(x, y_0) - y_0\| \\ &\leq \frac{1}{2} \|y - y_0\| + \|\Phi(x, y_0) - y_0\| \\ &\leq \frac{b}{2} + \|\Phi(x, y_0) - y_0\|. \end{aligned} \quad (17.16)$$

Vogliamo che anche l'ultimo termine sia $\leq b/2$. Per stimarlo, usiamo (17.11),

$$\|\Phi(x, y_0) - y_0\| = \|A^{-1} f(x, y_0)\| \leq \|A^{-1}\| \|f(x, y_0)\|.$$

Poiché f è continua e $f(x_0, y_0) = 0$, esiste $a > 0$, con $a \leq b$, tale che

$$\|f(x, y_0)\| \leq \frac{b}{2(1 + \|A^{-1}\|)} \quad \forall x \in I_a.$$

Quindi per ogni $x \in I_a$ si ha

$$\|\Phi(x, y_0) - y_0\| \leq \|A^{-1}\| \|f(x, y_0)\| \leq \frac{\|A^{-1}\| b}{2(1 + \|A^{-1}\|)} < \frac{b}{2}. \quad (17.17)$$

Da (17.16) e (17.17) segue che

$$\|\Phi(x, y) - y_0\| < b \quad \forall (x, y) \in I_a \times J_b, \quad (17.18)$$

e dunque

$$\Phi(x, y) \in J_b \quad \forall (x, y) \in I_a \times J_b.$$

Questo significa che, per ogni $x \in I_a$, la funzione $\Phi(x, \cdot) : y \mapsto \Phi(x, y)$ mappa J_b in J_b . Quindi $\Phi(x, \cdot)$ è una contrazione dello spazio metrico completo J_b , e dunque ha un unico punto fisso. Dalla (17.18), il punto fisso $y = \Phi(x, y)$ soddisfa $\|y - y_0\| < b$.

Abbiamo dimostrato che esistono $a > 0$, $b > 0$ tali che per ogni $x \in I_a$ esiste un unico elemento y nella palla J_b tale che $y = \Phi(x, y)$, cioè tale che $f(x, y) = 0$. Inoltre $\|y - y_0\| < b$.

Consideriamo in particolare i punti x nella palla aperta $\|x - x_0\| < a$, e chiamiamo $g(x)$ il punto fisso $y = \Phi(x, y)$. Dunque g è una funzione da $B(x_0, a)$ in $B(y_0, b)$ e soddisfa

$$f(x, g(x)) = 0 \quad \forall x \in B(x_0, a).$$

Inoltre, se

$$f(x, y) = 0, \quad (x, y) \in B(x_0, a) \times B(y_0, b),$$

allora $y = g(x)$ per l'unicità del punto fisso (per ogni $x \in I_a$ esiste un unico $y \in J_b$ tale che $y = \Phi(x, y)$, cioè tale che $f(x, y) = 0$). In particolare, si ha

$$y_0 = g(x_0)$$

perché $f(x_0, y_0) = 0$.

La dimostrazione prosegue nella prossima lezione.

18 Lezione 18

Riprendiamo la dimostrazione del Teorema della funzione implicita.

Dimostriamo la continuità della funzione g . Siano $x_1, x_2 \in B(x_0, a)$, e siano $y_1 = g(x_1)$, $y_2 = g(x_2)$ le corrispondenti soluzioni del problema di punto fisso,

$$y_1 = \Phi(x_1, y_1), \quad y_2 = \Phi(x_2, y_2).$$

Si ha

$$\begin{aligned} \|y_2 - y_1\| &= \|\Phi(x_2, y_2) - \Phi(x_1, y_1)\| \\ &\leq \|\Phi(x_2, y_2) - \Phi(x_2, y_1)\| + \|\Phi(x_2, y_1) - \Phi(x_1, y_1)\| \\ &\leq \frac{1}{2}\|y_2 - y_1\| + \|\Phi(x_2, y_1) - \Phi(x_1, y_1)\| \end{aligned} \quad (18.1)$$

perché la funzione $\Phi(x_2, \cdot)$ è lipschitziana con costante $\leq 1/2$, vedi (17.14). Portiamo a sinistra il termine $\frac{1}{2}\|y_2 - y_1\|$ in (18.1) (che si “assorbe” nella disuguaglianza), e otteniamo

$$\|y_2 - y_1\| \leq 2\|\Phi(x_2, y_1) - \Phi(x_1, y_1)\|.$$

Usiamo la formula (17.11) e stimiamo

$$\begin{aligned} \|\Phi(x_2, y_1) - \Phi(x_1, y_1)\| &= \|y_1 - A^{-1}f(x_2, y_1) - y_1 + A^{-1}f(x_1, y_1)\| \\ &\leq \|A^{-1}\| \|f(x_2, y_1) - f(x_1, y_1)\|. \end{aligned}$$

Quindi

$$\begin{aligned} \|g(x_2) - g(x_1)\| &= \|y_2 - y_1\| \\ &\leq 2\|A^{-1}\| \|f(x_2, y_1) - f(x_1, y_1)\| \quad \forall x_1, x_2 \in B(x_0, a), \end{aligned} \quad (18.2)$$

dove $y_1 = g(x_1)$, $y_2 = g(x_2)$. Poiché f è continua,

$$\|f(x_2, y_1) - f(x_1, y_1)\| \rightarrow 0 \quad \text{per } x_2 \rightarrow x_1;$$

quindi $g(x_2) \rightarrow g(x_1)$ per $x_2 \rightarrow x_1$. Questo prova che g è continua in x_1 , per ogni $x_1 \in B(x_0, a)$.

Sia ora $f \in C^1(\Omega)$; dimostriamo che $g \in C^1(B(x_0, a))$.

Sia $x_1 \in B(x_0, a)$, e sia $y_1 = g(x_1)$; il punto (x_1, y_1) appartiene all'aperto $B(x_0, a) \times B(y_0, b)$. Per ipotesi, $\partial_x f$ è continua in (x_1, y_1) . Dunque esiste $r > 0$ tale che

$$\begin{aligned} Q_r &:= B(x_1, r) \times B(y_1, r) \subseteq B(x_0, a) \times B(y_0, b), \\ \|\partial_x f(x, y) - \partial_x f(x_1, y_1)\|_{\mathcal{L}(X, Z)} &\leq 1 \quad \forall (x, y) \in Q_r. \end{aligned}$$

Perciò $\partial_x f$ è limitata in Q_r ,

$$\|\partial_x f(x, y)\|_{\mathcal{L}(X, Z)} \leq \|\partial_x f(x_1, y_1)\|_{\mathcal{L}(X, Z)} + 1 =: C_1 \quad \forall (x, y) \in Q_r. \quad (18.3)$$

Per ogni $x \in B(x_1, r)$, dalla (18.2), applicando il Teorema 17.2 del valor medio e la stima (18.3), si ha

$$\begin{aligned} \|g(x) - g(x_1)\| &\leq 2\|A^{-1}\| \|f(x, y_1) - f(x_1, y_1)\| \\ &\leq 2\|A^{-1}\| \sup_{t \in [0, 1]} \|(\partial_x f)(x_1 + t(x - x_1), y_1)\|_{\mathcal{L}(X, Z)} \|x - x_1\| \\ &\leq 2\|A^{-1}\| C_1 \|x - x_1\|, \end{aligned}$$

cioè

$$\|g(x) - g(x_1)\| \leq C_0 \|x - x_1\| \quad \forall x \in B(x_1, r), \quad (18.4)$$

dove $C_0 = 2\|A^{-1}\|C_1$ (costante che dipende da x_1 ma non dipende da x). In termini di $h = x - x_1$, la (18.4) diventa

$$\|g(x_1 + h) - g(x_1)\| \leq C_0 \|h\| \quad \forall \|h\| < r. \quad (18.5)$$

Sia $\varepsilon > 0$. Definiamo

$$\varepsilon_1 := \frac{\varepsilon}{2\|A^{-1}\|(1 + C_0)}.$$

Poiché f è differenziabile nel punto (x_1, y_1) , esiste $\delta > 0$ tale che

$$\begin{aligned} \|f(x_1 + h, y_1 + k) - f(x_1, y_1) - \partial_x f(x_1, y_1)h - \partial_y f(x_1, y_1)k\| \\ \leq \varepsilon_1 (\|h\| + \|k\|) \quad \forall \|h\|, \|k\| \leq \delta. \end{aligned} \quad (18.6)$$

Consideriamo, in particolare, l'incremento

$$k = g(x_1 + h) - g(x_1).$$

Se $\|h\| < r$ e $C_0\|h\| < \delta$, allora, dalla (18.5),

$$\|k\| = \|g(x_1 + h) - g(x_1)\| \leq C_0 \|h\| < \delta. \quad (18.7)$$

Dunque per h nella palla

$$\|h\| < \delta_1 := \min\{\delta, r, \delta/C_0\}$$

si ha $\|h\|, \|k\| < \delta$, e quindi vale la disuguaglianza in (18.6). Inoltre

$$f(x_1, y_1) = f(x_1, g(x_1)) = 0, \quad f(x_1 + h, y_1 + k) = f(x_1 + h, g(x_1 + h)) = 0$$

perché $f(x, g(x)) = 0$ per ogni $x \in B(x_0, a)$. Quindi, da (18.6) e (18.7), si ha

$$\begin{aligned} \|\partial_x f(x_1, y_1)h + \partial_y f(x_1, y_1)k\| &\leq \varepsilon_1 (\|h\| + \|k\|) \\ &\leq \varepsilon_1 (1 + C_0)\|h\| \quad \forall \|h\| < \delta_1. \end{aligned} \quad (18.8)$$

Dalla (17.15), $\partial_y f(x_1, y_1)$ è invertibile, con

$$\|(\partial_y f(x_1, y_1))^{-1}\| \leq 2\|A^{-1}\|. \quad (18.9)$$

Applicando $(\partial_y f(x_1, y_1))^{-1}$ alla somma che abbiamo stimato nella (18.8) si ha

$$(\partial_y f(x_1, y_1))^{-1} [\partial_x f(x_1, y_1)h + \partial_y f(x_1, y_1)k] = k + \Lambda h, \quad (18.10)$$

dove, in breve, indichiamo

$$\Lambda := (\partial_y f(x_1, y_1))^{-1} \partial_x f(x_1, y_1).$$

Notiamo che Λ è lineare e continuo perché $(\partial_y f(x_1, y_1))^{-1}$, $\partial_x f(x_1, y_1)$ sono funzioni lineari e continue. Dalla formula (18.10) e dalle stime (18.8), (18.9) si ha

$$\begin{aligned} \|k + \Lambda h\| &\leq \|(\partial_y f(x_1, y_1))^{-1}\| \|\partial_x f(x_1, y_1)h + \partial_y f(x_1, y_1)k\| \\ &\leq 2\|A^{-1}\| \varepsilon_1 (1 + C_0)\|h\| \\ &= \varepsilon \|h\| \quad \forall \|h\| < \delta_1. \end{aligned}$$

Ricordando che $k = g(x_1, h) - g(x_1)$, abbiamo dunque provato che per ogni $\varepsilon > 0$ esiste $\delta_1 > 0$ tale che

$$\|g(x_1 + h) - g(x_1) + \Lambda h\| \leq \varepsilon \|h\| \quad \forall \|h\| \leq \delta_1,$$

cioè

$$g(x_1 + h) - g(x_1) - (-\Lambda)h = o(\|h\|) \quad \text{per } h \rightarrow 0.$$

Questo significa che, per ogni $x_1 \in B(x_0, a)$, g è differenziabile in x_1 , e che il suo differenziale in x_1 è $-\Lambda$, cioè

$$g'(x_1)[h] = -(\partial_y f(x_1, y_1))^{-1} \partial_x f(x_1, y_1)h,$$

dove $y_1 = g(x_1)$.

Poiché $g, \partial_x f, \partial_y f, \text{Inv}$ sono funzioni continue, la loro composizione

$$\text{Inv}((\partial_y f)(x_1, g(x_1)))(\partial_x f)(x_1, g(x_1))$$

è una funzione continua; quindi g' è continua in $B(x_0, a)$, e $g \in C^1(B(x_0, a))$. \square

Derivate successive

Come in \mathbb{R} e in \mathbb{R}^n , definiamo la derivata seconda di una funzione come la derivata della derivata; in generale, definiamo $f^{(k)}$ come derivata di $f^{(k-1)}$. Cominciamo con la derivata seconda.

Definizione 18.1. (*Derivata seconda*). Siano X, Y spazi di Banach, sia $\Omega \subseteq X$ un aperto, $p \in \Omega$, e sia $f : \Omega \rightarrow Y$ di classe $C^1(\Omega, Y)$.

Diciamo che f è *differenziabile due volte in p* se il differenziale di f , cioè la funzione

$$f' : \Omega \rightarrow \mathcal{L}(X, Y),$$

è differenziabile in p ; in tal caso indichiamo $f''(p)$, oppure $D^2 f(p)$, oppure $d^2 f(p)$, il differenziale di f' in p .

Diciamo che f è *differenziabile due volte in Ω* se f è differenziabile due volte in ogni punto $p \in \Omega$.

Diciamo che f è di classe $C^2(\Omega, Y)$ se f'' è continua in Ω .

Dalla Definizione 18.1, $f''(p)$ è una funzione lineare e continua da X a $\mathcal{L}(X, Y)$,

$$f''(p) : X \rightarrow \mathcal{L}(X, Y), \quad f''(p) \in \mathcal{L}(X, \mathcal{L}(X, Y)).$$

Quindi per ogni $h \in X$ si ha $f''(p)[h] \in \mathcal{L}(X, Y)$, cioè $f''(p)[h]$ è una funzione lineare e continua da X a Y ,

$$f''(p)[h] : X \rightarrow Y, \quad v \mapsto f''(p)[h][v] \in Y.$$

È conveniente dunque considerare $f''(p)$ come *funzione bilineare* su X , che alla coppia $(h, v) \in X \times X$ associa $f''(p)[h][v] \in Y$.

Definizione 18.2. (*Funzioni bilineari*). Siano X, Y spazi vettoriali. Una funzione $B : X \times X \rightarrow Y$, $(h, v) \mapsto B(h, v)$ si dice *bilineare* se

per ogni $h \in X$ fissato, la funzione $X \rightarrow Y$, $v \mapsto B(h, v)$ è lineare, e
per ogni $v \in X$ fissato, la funzione $X \rightarrow Y$, $h \mapsto B(h, v)$ è lineare.

Nello spazio $X \times X$ consideriamo (per esempio) la norma

$$\|(h, v)\|_{X \times X} = \|h\|_X + \|v\|_X.$$

Per $B : X \times X \rightarrow Y$ bilineare definiamo

$$\|B\|_{\mathcal{L}_2(X, Y)} := \sup_{h, v \in X \setminus \{0\}} \frac{\|B(h, v)\|_Y}{\|h\|_X \|v\|_X}. \quad (18.11)$$

Lemma 18.3. (Continuità per funzioni bilineari). *Siano X, Y spazi di Banach. Una funzione bilineare $B : X \times X \rightarrow Y$, $(h, v) \mapsto B(h, v)$ è continua in $X \times X$ se e solo se*

$$\|B\|_{\mathcal{L}_2(X, Y)} < \infty.$$

Esercizio 18.4. Dimostrare il Lemma 18.3.

Definiamo

$$\begin{aligned} \mathcal{L}_2(X, Y) &:= \{B : X \times X \rightarrow Y, B \text{ bilineare e continua}\} \\ &= \{B : X \times X \rightarrow Y, B \text{ bilineare, } \|B\|_{\mathcal{L}_2(X, Y)} < \infty\}. \end{aligned}$$

Lemma 18.5. (Spazio delle funzioni bilineari continue). *Siano X, Y spazi vettoriali normati. Se Y è di Banach, allora $\mathcal{L}_2(X, Y)$, $\|\cdot\|_{\mathcal{L}_2(X, Y)}$ è spazio di Banach.*

Esercizio 18.6. Dimostrare il Lemma 18.5.

Lemma 18.7. (Isomorfismo di $\mathcal{L}(X, \mathcal{L}(X, Y))$ e $\mathcal{L}_2(X, Y)$). *Siano X, Y spazi di Banach.*

(i) *Sia $E \in \mathcal{L}(X, \mathcal{L}(X, Y))$. Per ogni $h, v \in X$, definiamo $B(h, v) := E[h][v]$. Allora $B \in \mathcal{L}_2(X, Y)$, e si ha*

$$\|E\|_{\mathcal{L}(X, \mathcal{L}(X, Y))} = \|B\|_{\mathcal{L}_2(X, Y)}. \quad (18.12)$$

(ii) *Viceversa, sia $B \in \mathcal{L}_2(X, Y)$. Per ogni $h \in X$, definiamo $E[h]$ come la funzione di X in Y che mappa $v \in X$ in $E[h][v] := B(h, v) \in Y$. Allora, per ogni $h \in X$, $E[h]$ è una funzione lineare e continua di X in Y , cioè $E[h] \in \mathcal{L}(X, Y)$ per ogni $h \in X$. La funzione $E : X \rightarrow \mathcal{L}(X, Y)$, $h \mapsto E[h]$ è lineare e continua, cioè $E \in \mathcal{L}(X, \mathcal{L}(X, Y))$. Si ha inoltre l'uguaglianza (18.12).*

Dimostrazione. (i). Sia $E \in \mathcal{L}(X, \mathcal{L}(X, Y))$, e definiamo $B(h, v) := E[h][v]$ per ogni $h, v \in X$.

Per ogni h , la funzione $E[h] : X \rightarrow Y$ è lineare, ed è la funzione $v \mapsto E[h][v] = B(h, v)$; questo significa che, per ogni h , la funzione $v \mapsto B(h, v)$ è lineare. Per ipotesi, $E : X \rightarrow \mathcal{L}(X, Y)$ è lineare, cioè

$$E[h + k] = E[h] + E[k], \quad E[\lambda h] = \lambda E[h]$$

per ogni $h, k \in X$, $\lambda \in \mathbb{R}$; queste sono funzioni da X in Y , quindi, per ogni $v \in X$,

$$E[h + k][v] = E[h][v] + E[k][v], \quad E[\lambda h][v] = \lambda E[h][v],$$

cioè

$$B(h + k, v) = B(h, v) + B(k, v), \quad B(\lambda h, v) = \lambda B(h, v).$$

Quindi, per ogni v , la funzione $h \mapsto B(h, v)$ è lineare. Dunque B è bilineare.

Proviamo (18.12). Per definizione di norma operatoriale, si ha

$$\begin{aligned} \|E\|_{\mathcal{L}(X, \mathcal{L}(X, Y))} &= \sup_{h \in X, h \neq 0} \frac{\|E[h]\|_{\mathcal{L}(X, Y)}}{\|h\|_X} \\ &= \sup_{h \in X, h \neq 0} \frac{1}{\|h\|_X} \sup_{v \in X, v \neq 0} \frac{\|E[h][v]\|_Y}{\|v\|_X} \\ &= \sup_{h, v \in X, h, v \neq 0} \frac{\|E[h][v]\|_Y}{\|h\|_X \|v\|_X} \\ &= \sup_{h, v \in X, h, v \neq 0} \frac{\|B(h, v)\|_Y}{\|h\|_X \|v\|_X} = \|B\|_{\mathcal{L}_2(X, Y)}. \end{aligned}$$

Infine B è continua perché è bilineare e $\|B\|_{\mathcal{L}_2(X, Y)} < \infty$.

La dimostrazione di (ii) è simile. \square

Il Lemma 18.7 permette di identificare gli spazi $\mathcal{L}(X, \mathcal{L}(X, Y))$ e $\mathcal{L}_2(X, Y)$, e dunque di considerare $f''(p)$ una funzione bilineare,

$$f''(p) \in \mathcal{L}_2(X, Y), \quad f''(p) : X \times X \rightarrow Y, \quad (h, v) \mapsto f''(p)[h, v].$$

Scriviamo $f''(p)[h, v]$ perché, in analogia al caso lineare, usiamo le parentesi quadre per gli argomenti in cui la dipendenza è lineare. Se poi f è differenziabile due volte in ogni punto $p \in \Omega$, si ha

$$f'' : \Omega \rightarrow \mathcal{L}_2(X, Y),$$

e se f'' è continua da Ω a $\mathcal{L}_2(X, Y)$, scriviamo $f \in C^2(\Omega, Y)$.

Lemma 18.8. ($f''(p)$ è simmetrico). *Siano X, Y spazi di Banach, $\Omega \subseteq X$ un aperto, $p \in \Omega$. Se f è differenziabile due volte in p (cioè f è differenziabile in un intorno di p e f' è differenziabile in p), allora la funzione bilineare $f''(p)$ è simmetrica, cioè*

$$f''(p)[h, v] = f''(p)[v, h] \quad \forall h, v \in X.$$

Dimostrazione. Non in programma. Si trova in [2], capitolo 1.3, Teorema 3.4. \square

Esempio. Consideriamo la funzione

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad f(x, y) = \begin{pmatrix} \sin(x + y) \\ e^x - y^3 \\ x^2 y \end{pmatrix}.$$

In questo caso $X = \mathbb{R}^2$, $Y = \mathbb{R}^3$, $\Omega = X$. Il differenziale di f nel punto $(x, y) \in \Omega$ in direzione $(h, k) \in X$ è

$$f'(x, y)[(h, k)] = \begin{pmatrix} \cos(x + y)(h + k) \\ e^x h - 3y^2 k \\ 2xyh + x^2 k \end{pmatrix}.$$

Quindi $f'(x, y)$ è la funzione lineare e continua

$$f'(x, y) : (h, k) \mapsto f'(x, y)[(h, k)]$$

che manda il vettore $(h, k) \in X = \mathbb{R}^2$ nel vettore $f'(x, y)[(h, k)] \in Y = \mathbb{R}^3$, e

$$f'(x, y) \in \mathcal{L}(X, Y) = \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3).$$

Il differenziale f' è la funzione

$$f' : \Omega = \mathbb{R}^2 \rightarrow \mathcal{L}(X, Y) = \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$$

che al punto $(x, y) \in \Omega$ associa $f'(x, y) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$.

La derivata di $f'(x, y)[(h, k)]$ nel punto (x, y) in direzione (s, t) è

$$f''(x, y)[(s, t)][(h, k)] = \begin{pmatrix} -\sin(x+y)(s+t)(h+k) \\ e^x sh - 6ytk \\ 2ysh + 2xth + 2xsk \end{pmatrix}. \quad (18.13)$$

Quindi $f''(x, y)[(s, t)]$ è la funzione lineare e continua

$$f''(x, y)[(s, t)] : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad (h, k) \mapsto f''(x, y)[(s, t)][(h, k)]$$

che manda il vettore $(h, k) \in \mathbb{R}^2$ nel vettore $f''(x, y)[(s, t)][(h, k)] \in \mathbb{R}^3$, dunque

$$f''(x, y)[(s, t)] \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3).$$

Inoltre la funzione

$$f''(x, y) : \mathbb{R}^2 \rightarrow \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3), \quad (s, t) \mapsto f''(x, y)[(s, t)]$$

che al vettore $(s, t) \in \mathbb{R}^2$ associa la funzione $f''(x, y)[(s, t)] \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ è a sua volta lineare, e quindi

$$f''(x, y) \in \mathcal{L}(\mathbb{R}^2, \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)).$$

D'altra parte, la funzione

$$f''(x, y) : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad ((s, t), (h, k)) \mapsto f''(x, y)[(s, t)][(h, k)]$$

è una funzione bilineare,

$$f''(x, y) \in \mathcal{L}_2(\mathbb{R}^2, \mathbb{R}^3),$$

come si vede dal fatto che gli incrementi h, k, s, t nella formula (18.13) compaiono tutti come monomi di grado 2, con dipendenza lineare in (s, t) tenendo fissi (h, k) , e con dipendenza lineare in (h, k) tenendo fissi (s, t) .

Se deriviamo (18.13) in direzione (u, v) troviamo

$$f'''(x, y)[(u, v)][(s, t)][(h, k)] = \begin{pmatrix} -\cos(x+y)(u+v)(s+t)(h+k) \\ e^x ush - 6vtk \\ 2vsh + 2uth + 2usk \end{pmatrix},$$

che è funzione trilineare dei vettori $(u, v), (s, t), (h, k)$ (cioè lineare in ciascuno dei tre, tenendo fissi gli altri due).

Se deriviamo ancora, in direzione (α, β) , troviamo

$$f^{(4)}(x, y)[(\alpha, \beta)][(u, v)][(s, t)][(h, k)] = \begin{pmatrix} \sin(x+y)(\alpha+\beta)(u+v)(s+t)(h+k) \\ e^x \alpha ush \\ 0 \end{pmatrix},$$

funzione 4-lineare dei vettori $(\alpha, \beta), (u, v), (s, t), (h, k)$. E così via.

Vediamo in generale le derivate successive. Poiché f'' è una funzione da Ω a $\mathcal{L}_2(X, Y)$, il suo differenziale in p , cioè $f'''(p)$, appartiene a $\mathcal{L}(X, \mathcal{L}_2(X, Y))$. Come per il Lemma 18.7, si prova che $\mathcal{L}(X, \mathcal{L}_2(X, Y))$ è isomorfo (e isometrico) allo spazio $\mathcal{L}_3(X, Y)$ delle funzioni trilineari continue $T : X^3 \rightarrow Y$, dove $X^3 = X \times X \times X$. Si può quindi considerare

$$f'''(p) \in \mathcal{L}_3(X, Y),$$

e proseguire, definendo la derivata quarta come derivata di f''' , eccetera.

Procediamo per induzione, assumendo di aver definito le funzioni differenziabili $(k-1)$ volte e le funzioni $(k-1)$ -lineari continue.

Definizione 18.9. (*Derivata k -esima*). Siano X, Y spazi di Banach, $\Omega \subseteq X$ aperto, $k \geq 2$ un intero. Supponiamo che $f : \Omega \rightarrow Y$ sia differenziabile $k-1$ volte in Ω . Sia $p \in \Omega$. Diciamo che f è *differenziabile k volte in p* se il differenziale $(k-1)$ -esimo di f , cioè la funzione

$$f^{(k-1)} : \Omega \rightarrow \mathcal{L}_{k-1}(X, Y),$$

è differenziabile in p ; in tal caso indichiamo $f^{(k)}(p)$, oppure $D^k f(p)$, oppure $d^k f(p)$, e chiamiamo *differenziale k -esimo di f in p* , la funzione $(f^{(k-1)})'(p)$, cioè il differenziale di $f^{(k-1)}$ in p . L'insieme $\mathcal{L}_{k-1}(X, Y)$, codominio di $f^{(k-1)}$, è lo spazio delle funzioni $(k-1)$ -lineari e continue.

Diciamo che f è *differenziabile k volte in Ω* se f è differenziabile k volte in ogni punto $p \in \Omega$.

Diciamo che f è di classe $C^k(\Omega, Y)$ se $f^{(k)}$ è continua in Ω .

Dalla Definizione 18.9, $f^{(k)}(p)$ è una funzione lineare e continua da X allo spazio $\mathcal{L}_{k-1}(X, Y)$ delle funzioni $(k-1)$ -lineari e continue,

$$f^{(k)}(p) \in \mathcal{L}(X, \mathcal{L}_{k-1}(X, Y)).$$

Quindi per ogni $h \in X$ si ha $f^{(k)}(p)[h] \in \mathcal{L}_{k-1}(X, Y)$, cioè $f^{(k)}(p)[h]$ è una funzione $(k-1)$ -lineare e continua da X^{k-1} a Y ,

$$f^{(k)}(p)[h] : X^{k-1} \rightarrow Y, \quad (v_1, \dots, v_{k-1}) \mapsto f^{(k)}(p)[h][v_1, \dots, v_{k-1}] \in Y.$$

Come sopra, conviene considerare $f^{(k)}(p)$ come funzione k -lineare su X , che a $(h_1, \dots, h_k) \in X^k$ associa $f^{(k)}(p)[h_1, h_2, \dots, h_k] = f^{(k)}(p)[h_1][h_2, \dots, h_k] \in Y$.

Definizione 18.10. (*Funzioni k -lineari*). Siano X, Y spazi vettoriali. Una funzione

$$M : X^k \rightarrow Y, \quad (h_1, \dots, h_k) \mapsto M(h_1, \dots, h_k)$$

si dice *k -lineare* se per ogni $j = 1, \dots, k$, per h_1, \dots, h_k , eccetto h_j , fissati, la funzione $X \rightarrow Y$, $h_j \mapsto M(h_1, \dots, h_k)$ è lineare. In altre parole, M è k -lineare se è una funzione lineare in ciascuno dei suoi argomenti h_1, \dots, h_k separatamente.

Una funzione si dice *multi-lineare* se è k -lineare per un certo k .

Nello spazio prodotto $X^k = X \times \dots \times X$ consideriamo (per esempio) la norma

$$\|(h_1, \dots, h_k)\|_{X^k} = \|h_1\|_X + \dots + \|h_k\|_X.$$

Per $M : X^k \rightarrow Y$ k -lineare definiamo

$$\|M\|_{\mathcal{L}_k(X,Y)} := \sup_{h_1, \dots, h_k \in X \setminus \{0\}} \frac{\|M(h_1, \dots, h_k)\|_Y}{\|h_1\|_X \cdots \|h_k\|_X}. \quad (18.14)$$

Lemma 18.11. (Continuità per funzioni k -lineari). *Siano X, Y spazi di Banach. Una funzione k -lineare*

$$M : X^k \rightarrow Y, \quad (h_1, \dots, h_k) \mapsto M(h_1, \dots, h_k)$$

è continua in X^k se e solo se

$$\|M\|_{\mathcal{L}_k(X,Y)} < \infty.$$

Dimostrazione. Adattare il caso bilineare. □

Definiamo

$$\begin{aligned} \mathcal{L}_k(X, Y) &:= \{M : X^k \rightarrow Y, M \text{ } k\text{-lineare e continua}\} \\ &= \{M : X^k \rightarrow Y, M \text{ } k\text{-lineare, } \|M\|_{\mathcal{L}_k(X,Y)} < \infty\}. \end{aligned}$$

Lemma 18.12. (Spazio delle funzioni k -lineari continue). *Siano X, Y spazi vettoriali normati. Se Y è di Banach, allora $\mathcal{L}_k(X, Y)$, $\|\cdot\|_{\mathcal{L}_k(X,Y)}$ è spazio di Banach.*

Dimostrazione. Adattare il caso bilineare. □

Lemma 18.13. (Isomorfismo di $\mathcal{L}(X, \mathcal{L}_{k-1}(X, Y))$ e $\mathcal{L}_k(X, Y)$). *Siano X, Y spazi di Banach.*

(i) *Sia $E \in \mathcal{L}(X, \mathcal{L}_{k-1}(X, Y))$. Per ogni $h_1, \dots, h_k \in X$, definiamo*

$$M(h_1, \dots, h_k) := E[h_1][h_2, \dots, h_k].$$

Allora $M \in \mathcal{L}_k(X, Y)$, e si ha

$$\|E\|_{\mathcal{L}(X, \mathcal{L}_{k-1}(X,Y))} = \|M\|_{\mathcal{L}_k(X,Y)}. \quad (18.15)$$

(ii) *Viceversa, sia $M \in \mathcal{L}_k(X, Y)$. Per ogni $h_1 \in X$, definiamo $E[h_1]$ come la funzione di X^{k-1} in Y che mappa $(h_2, \dots, h_k) \in X^{k-1}$ in $E[h_1][h_2, \dots, h_k] := M(h_1, \dots, h_k) \in Y$. Allora, per ogni $h_1 \in X$, $E[h_1]$ è una funzione $(k-1)$ -lineare e continua di X^{k-1} in Y , cioè $E[h_1] \in \mathcal{L}_{k-1}(X, Y)$ per ogni $h_1 \in X$. La funzione $E : X \rightarrow \mathcal{L}_{k-1}(X, Y)$, $h_1 \mapsto E[h_1]$ è lineare e continua, cioè $E \in \mathcal{L}(X, \mathcal{L}_{k-1}(X, Y))$. Si ha inoltre l'uguaglianza (18.15).*

Dimostrazione. Adattare il caso bilineare. □

Dal Lemma 18.13 possiamo identificare gli spazi $\mathcal{L}(X, \mathcal{L}_{k-1}(X, Y))$ e $\mathcal{L}_k(X, Y)$, e dunque consideriamo $f^{(k)}(p)$ una funzione k -lineare,

$$f^{(k)}(p) \in \mathcal{L}_k(X, Y), \quad f^{(k)}(p) : X^k \rightarrow Y, \quad (h_1, \dots, h_k) \mapsto f''(p)[h_1, \dots, h_k].$$

Scriviamo $f^{(k)}(p)[h_1, \dots, h_k]$ perché, in analogia al caso lineare e bilineare, usiamo le parentesi quadre per gli argomenti in cui la dipendenza è lineare. Se poi f è differenziabile k volte in ogni punto $p \in \Omega$, si ha

$$f^{(k)} : \Omega \rightarrow \mathcal{L}_k(X, Y),$$

e se $f^{(k)}$ è continua da Ω a $\mathcal{L}_k(X, Y)$, scriviamo $f \in C^k(\Omega, Y)$.

Lemma 18.14. ($f^{(k)}$ è simmetrica). *Siano X, Y spazi di Banach, $\Omega \subseteq X$ un aperto, $p \in \Omega$. Se f è differenziabile k volte in p (cioè f è differenziabile $k - 1$ volte in un intorno di p e $f^{(k-1)}$ è differenziabile in p), allora la funzione k -lineare $f^{(k)}(p)$ è simmetrica, cioè*

$$f^{(k)}(p)[h_1, \dots, h_k] = f^{(k)}(p)[h_{\sigma(1)}, \dots, h_{\sigma(k)}] \quad \forall h_1, \dots, h_k \in X,$$

per ogni permutazione σ di $\{1, \dots, k\}$.

Dimostrazione. Non in programma. Si trova in [2], capitolo 1.3, Teorema 3.5. \square

Lemma 18.15. *La somma, il prodotto per scalari, la composizione di funzioni di classe C^k è di classe C^k .*

Dimostrazione. Adattare agli spazi di Banach le corrispondenti dimostrazioni in \mathbb{R}^n . \square

Definizione 18.16. (*Funzioni C^∞*). Siano X, Y spazi di Banach, $\Omega \subseteq X$ aperto. Una funzione $f : \Omega \rightarrow Y$ si dice $C^\infty(\Omega, Y)$ se $f \in C^k(\Omega, Y)$ per ogni $k \in \mathbb{N}$.

19 Lezione 19

Proposizione 19.1. (Inv è di classe C^∞). *Siano X, Y spazi di Banach. L'operatore di inversione*

$$\text{Inv} : \text{Iso}(X, Y) \rightarrow \mathcal{L}(Y, X), \quad \text{Inv}(A) = A^{-1}$$

è di classe $C^\infty(\text{Iso}(X, Y), \mathcal{L}(Y, X))$.

Dimostrazione. Nella Proposizione 16.16 abbiamo dimostrato che Inv è di classe C^1 e il suo differenziale nel punto $A \in \text{Iso}(X, Y)$ in direzione $B \in \mathcal{L}(X, Y)$ è

$$\text{Inv}'(A)[B] = -A^{-1}BA^{-1}.$$

Quindi l'operatore lineare

$$\text{Inv}'(A) : \mathcal{L}(X, Y) \rightarrow \mathcal{L}(Y, X), \quad B \mapsto \text{Inv}'(A)[B]$$

è la composizione

$$\text{Inv}'(A) = E(A)M(A), \quad \text{Inv}'(A)[B] = E(A)[M(A)[B]],$$

dove $M(A), E(A)$ sono gli operatori lineari di composizione con A^{-1} a sinistra/destra,

$$\begin{aligned} M(A) : \mathcal{L}(X, Y) &\rightarrow \mathcal{L}(Y, Y), & B &\mapsto M(A)[B] := BA^{-1} = B \text{Inv}(A), \\ E(A) : \mathcal{L}(Y, Y) &\rightarrow \mathcal{L}(Y, X), & H &\mapsto E(A)[H] := -A^{-1}H = -\text{Inv}(A)H. \end{aligned}$$

Procediamo per bootstrap. Poiché Inv (cioè la funzione $A \mapsto \text{Inv}(A)$) è C^1 , anche M, E (cioè le funzioni $A \mapsto M(A), A \mapsto E(A)$) sono C^1 . Quindi la loro composizione $\text{Inv}'(A) = E(A)M(A)$ è una funzione C^1 . Dunque Inv è C^2 .

Supponiamo che Inv sia di classe C^k . Allora E, M sono C^k , quindi $\text{Inv}' = EM$ è C^k , e dunque Inv è C^{k+1} .

Abbiamo provato, per induzione, che $\text{Inv} \in C^k$ per ogni $k \in \mathbb{N}$. □

Proposizione 19.2. (Ulteriore regolarità della funzione implicita). *Supponiamo che valgano le ipotesi del Teorema 17.8 della funzione implicita.*

Se, inoltre, $f \in C^k(\Omega, Z)$, con $k \geq 1$ intero, allora $g \in C^k(B(x_0, a), B(y_0, b))$. Se $f \in C^\infty$, anche $g \in C^\infty$.

Dimostrazione. Nel Teorema 17.8 abbiamo provato che se $f \in C^1$ allora $g \in C^1$, e il differenziale di g nel punto $x \in B(x_0, a)$ è

$$g'(x) = -((\partial_y f)(x, g(x)))^{-1}(\partial_x f)(x, g(x)) \quad \forall x \in B(x_0, a). \quad (19.1)$$

Supponiamo che f sia di classe C^k , con $k \geq 2$ (il caso $k = 1$ è già stato provato nel Teorema 17.8). Come sopra, procediamo per bootstrap.

Poiché $f \in C^k(\Omega)$, le derivate parziali $\partial_x f, \partial_y f$ sono $C^{k-1}(\Omega)$. Supponiamo che g sia di classe $C^n(B(x_0, a))$, con $n \leq k - 1$. La funzione

$$B(x_0, a) \rightarrow \Omega, \quad x \mapsto (x, g(x))$$

è di classe C^n . Le composizioni $(\partial_x f)(x, g(x)), (\partial_y f)(x, g(x))$ sono C^n . Poiché $\text{Inv} \in C^\infty$, anche $\text{Inv}((\partial_y f)(x, g(x)))$ è C^n . Quindi la composizione (19.1) è C^n , e dunque $g \in C^{n+1}$.

Per induzione, questo argomento ($g \in C^n \Rightarrow g \in C^{n+1}$) si può iterare fintantoché $n \leq k - 1$. Si arriva a $g \in C^{n+1}$ con $n = k - 1$, cioè $g \in C^k$, che è la tesi.

Se $f \in C^\infty$, non ci sono limitazioni al passo induttivo, e, per induzione, $g \in C^n$ per ogni $n \in \mathbb{N}$. □

Esercizio 19.3. Sia X lo spazio vettoriale delle successioni di numeri reali $x = (x_1, x_2, \dots)$ limitate, munito della norma del sup $\|x\| = \sup |x_n|$. Sia $f : X^2 \rightarrow X$ la funzione

$$f(x, y) = Ay + P(x, y),$$

dove $A = (A_{jk})_{j,k=1,2,\dots}$ è la matrice a infinite righe e infinite colonne

$$A = \begin{pmatrix} B & & & \\ & B & & \\ & & B & \\ & & & \ddots \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

e $P(x, y)$ è la successione (z_1, z_2, \dots) con

$$z_n = x_n \cos(x_n) + \sin(y_n x_{n+1}).$$

Stabilire se l'equazione $f(x, y) = 0$ definisce una funzione implicita $y = g(x)$ intorno al punto $(0, 0)$. In caso affermativo, calcolare il differenziale $g'(0)[h]$.

Esercizio 19.4. Provare che esistono funzioni $y(x), z(x)$ definite implicitamente dal sistema

$$\begin{cases} x^3 - 3xy^2 + z^3 + 1 = 0, \\ x - 2y^2 - 3z^2 + 4 = 0 \end{cases}$$

che soddisfano $y(1) = z(1) = 1$. Calcolare $y'(1), z'(1)$.

Esercizio 19.5. Si consideri il sistema

$$\begin{cases} (x + y)^2 + u^2 y + 2v - 6 = 0, \\ (x - y)^2 + u^2 v - 2xy - 3 = 0 \end{cases}$$

nelle incognite $(x, y, u, v) \in \mathbb{R}^4$. Dimostrare che il sistema definisce implicitamente delle funzioni $u(x, y), v(x, y)$ che soddisfano

$$u(0, 1) = 1, \quad v(0, 1) = 2.$$

Calcolare $\nabla u, \nabla v$ in $(0, 1)$.

Esercizio 19.6. Siano $a_0, a_1, \dots, a_{n-1} : \mathbb{R} \rightarrow \mathbb{R}$ funzioni di classe $C^k(\mathbb{R})$. Per ogni $t, x \in \mathbb{R}$, sia

$$p(t, x) = x^n + a_{n-1}(t)x^{n-1} + \dots + a_2(t)x^2 + a_1(t)x + a_0(t)$$

polinomio di grado n in x , con coefficienti $a_j(t)$ dipendenti da t . Supponiamo che per ogni $t \in \mathbb{R}$ il polinomio abbia n radici reali distinte

$$x_1(t) < x_2(t) < \dots < x_n(t).$$

Dimostrare che x_1, \dots, x_n sono funzioni di classe $C^k(\mathbb{R})$.

Suggerimento. Per ogni t il polinomio si fattorizza come

$$p(t, x) = (x - x_1(t))(x - x_2(t)) \cdots (x - x_n(t)).$$

Se, per un certo t , $x_j(t)$ fosse radice del polinomio derivato $\partial_x p(t, x)$, allora $x_j(t)$ sarebbe radice di $p(t, x)$ di molteplicità ≥ 2 .

Serie di Fourier

Questa terza e ultima parte del corso è dedicata alle serie di Fourier. Si tratta di un prima, piccola porzione di un vasto campo dell'analisi, detto *analisi di Fourier* (o *analisi armonica*), che ha le sue origini nei primi studi sulle equazioni lineari del calore e delle onde, e ha oggi un ruolo centrale in varie parti dell'analisi, incluse le equazioni alle derivate parziali e la teoria dei numeri.

L'idea alla base delle serie di Fourier è quella di usare seni e coseni, $\sin(nx)$ e $\cos(nx)$, come elementi fondamentali in cui decomporre le funzioni periodiche.

Definizione 19.7. (*Integrale di funzioni a valori complessi*). Sia $f : [a, b] \rightarrow \mathbb{C}$ una funzione di variabile reale a valori complessi. Diciamo che f è *integrabile in* $[a, b]$ se la sua parte reale $\operatorname{Re} f$ e la sua parte immaginaria $\operatorname{Im} f$ sono integrabili in $[a, b]$. Definiamo l'integrale di f in $[a, b]$ con la formula

$$\int_a^b f(x) dx := \int_a^b \operatorname{Re} f(x) dx + i \int_a^b \operatorname{Im} f(x) dx. \quad (19.2)$$

In queste lezioni, come anche nel resto del corso, usiamo l'aggettivo *integrabile* come sinonimo di Riemann-integrabile (come in Analisi 1).

Osservazione 19.8. (*Integrale e primitiva*). Siano $f, h : [a, b] \rightarrow \mathbb{C}$, con f continua, h derivabile, e $h'(x) = f(x)$, cioè

$$(\operatorname{Re} h)'(x) = \operatorname{Re} f(x), \quad (\operatorname{Im} h)'(x) = \operatorname{Im} f(x)$$

per ogni $x \in [a, b]$. Allora dal Teorema fondamentale del calcolo integrale si ha

$$\int_a^b \operatorname{Re} f(x) dx = \operatorname{Re} h(b) - \operatorname{Re} h(a), \quad \int_a^b \operatorname{Im} f(x) dx = \operatorname{Im} h(b) - \operatorname{Im} h(a),$$

e dunque, dalla definizione (19.2),

$$\int_a^b f(x) dx = h(b) - h(a). \quad (19.3)$$

Osservazione 19.9. (*Derivata e integrale degli esponenziali complessi*). Usando la formula dell'esponenziale complesso

$$e^{ix} = \cos x + i \sin x, \quad x \in \mathbb{R},$$

per ogni $\lambda \in \mathbb{R}$ si ha

$$\frac{d}{dx} e^{i\lambda x} = \frac{d}{dx} [\cos(\lambda x) + i \sin(\lambda x)] = -\lambda \sin(\lambda x) + i\lambda \cos(\lambda x) = i\lambda e^{i\lambda x}$$

e dunque, per $\lambda \neq 0$,

$$\frac{d}{dx} \left(\frac{e^{i\lambda x}}{i\lambda} \right) = e^{i\lambda x}.$$

Di conseguenza, dall'Osservazione 19.8,

$$\int_a^b e^{i\lambda x} dx = \left[\frac{e^{i\lambda x}}{i\lambda} \right]_{x=a}^{x=b} = \frac{e^{i\lambda b} - e^{i\lambda a}}{i\lambda}, \quad \lambda \neq 0. \quad (19.4)$$

Esercizio 19.10. Per $\lambda \neq 0$ calcolare

$$\int_a^b \cos(\lambda x) dx + i \int_a^b \sin(\lambda x) dx$$

e verificare che, usando seno e coseno, si ottiene la formula (19.4).

Esercizio 19.11. (*Formula di integrazione per parti per funzioni a valori complessi*). Siano $f, g : [a, b] \rightarrow \mathbb{C}$ funzioni a valori complessi, con $\operatorname{Re} f, \operatorname{Im} f, \operatorname{Re} g, \operatorname{Im} g$ di classe $C^1([a, b], \mathbb{R})$. Allora

$$\begin{aligned} \int_a^b f'(x)g(x) dx + \int_a^b f(x)g'(x) dx &= [f(x)g(x)]_a^b \\ &= f(b)g(b) - f(a)g(a). \end{aligned} \quad (19.5)$$

Scrivendo f, g con parte reale e immaginaria, e usando la formula di integrazione per parti per le funzioni reali (Analisi 1), provare che vale (19.5).

Osservazione 19.12. (*Formule di uso frequente*). Negli esercizi e nei calcoli con i polinomi trigonometrici si usano molto spesso le seguenti identità:

$$e^{im\pi} = (-1)^m \quad \forall m \in \mathbb{Z}, \quad (19.6)$$

$$\int_{-\pi}^{\pi} e^{imx} dx = 0 \quad \forall m \in \mathbb{Z}, m \neq 0. \quad (19.7)$$

Seguono entrambe dall'identità elementare $e^{imx} = \cos(mx) + i \sin(mx)$.

Definizione 19.13. (*Polinomio trigonometrico, serie trigonometrica*). Si dice *polinomio trigonometrico* una somma (finita) di esponenziali complessi del tipo

$$\sum_{|n| \leq N} c_n e^{inx},$$

con coefficienti $c_n \in \mathbb{C}$; $N \in \mathbb{N}$ è il *grado* del polinomio trigonometrico. Si dice *serie trigonometrica* una serie di funzioni del tipo

$$\sum_{n \in \mathbb{Z}} c_n e^{inx},$$

dove l'operazione di limite per $n \rightarrow \pm\infty$ è intesa in senso simmetrico, cioè

$$\sum_{n \in \mathbb{Z}} c_n e^{inx} = \lim_{N \rightarrow +\infty} \sum_{n=-N}^N c_n e^{inx},$$

qualunque sia il tipo di convergenza (puntuale, uniforme, ecc.) che si sta trattando.

Osservazione 19.14. (*Polinomio trigonometrico, serie trigonometrica in seni e coseni*). Ogni polinomio trigonometrico

$$\sum_{|n| \leq N} c_n e^{inx} \quad (19.8)$$

con coefficienti $c_n \in \mathbb{C}$ si può scrivere nella forma

$$a_0 + \sum_{n=1}^N \left(a_n \cos(nx) + b_n \sin(nx) \right) \quad (19.9)$$

con coefficienti $a_n, b_n \in \mathbb{C}$, e viceversa. Infatti, fissati dei coefficienti c_n , il polinomio trigonometrico (19.8) si scrive nella forma (19.9) dove i coefficienti a_n, b_n sono

$$a_0 = c_0, \quad a_n = c_n + c_{-n}, \quad b_n = i(c_n - c_{-n}) \quad \forall n \geq 1. \quad (19.10)$$

Viceversa, fissati dei coefficienti a_n, b_n , il polinomio trigonometrico in seni e coseni (19.9) si riscrive nella forma (19.8), dove i coefficienti c_n sono

$$c_0 = a_0, \quad c_n = \frac{a_n - ib_n}{2}, \quad c_{-n} = \frac{a_n + ib_n}{2} \quad \forall n \geq 1. \quad (19.11)$$

Esercizio 19.15. Dimostrare le affermazioni dell'Osservazione 19.14.

Cominciamo lo studio delle funzioni periodiche.

Lemma 19.16. (Proprietà di traslazione del periodo). *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile in $[0, T]$. Allora*

$$\int_a^{a+T} f(x) dx = \int_0^T f(x) dx \quad \forall a \in \mathbb{R}.$$

Esercizio 19.17. Dimostrare il Lemma 19.16. Per esempio, primo modo: usare il fatto che per ogni $a \in \mathbb{R}$ esiste $m \in \mathbb{Z}$ tale che $mT \leq a < mT + T$. Oppure, secondo modo: separare $\int_a^{a+T} f = \int_a^T f + \int_T^{a+T} f$ e provare che $\int_T^{a+T} f = \int_0^a f$.

Osservazione 19.18. (*Normalizzazione del periodo a 2π*). Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$. Allora la funzione

$$g : \mathbb{R} \rightarrow \mathbb{C}, \quad g(x) := f\left(\frac{T}{2\pi}x\right) \quad (19.12)$$

è periodica di periodo 2π . Tramite il riscaldamento (19.12), ogni risultato enunciato per funzioni periodiche di periodo 2π può essere esteso, con gli opportuni adattamenti riguardo alle costanti, alle funzioni periodiche di periodo $T > 0$ qualunque.

Definizione 19.19. (*Coefficienti di Fourier, serie di Fourier*). Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile in $[0, T]$.

Si dicono *coefficienti di Fourier di f (relativi al periodo T , rispetto agli esponenti complessi)* i numeri complessi

$$\hat{f}(n) := \frac{1}{T} \int_0^T f(x) e^{-in\frac{2\pi}{T}x} dx, \quad n \in \mathbb{Z}$$

(\hat{f} si legge “ f cappuccio”).

Si dice *serie di Fourier di f* la serie trigonometrica di coefficienti $\hat{f}(n)$

$$S(f)(x) := \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{in\frac{2\pi}{T}x}.$$

Si dice *somma parziale N -esima della serie di Fourier di f* la somma parziale sugli indici $n = -N, \dots, N$, cioè il polinomio trigonometrico

$$S_N(f)(x) := \sum_{|n| \leq N} \hat{f}(n) e^{in\frac{2\pi}{T}x}, \quad N \in \mathbb{N}.$$

Osservazione 19.20. Dal Lemma 19.16, si ha

$$\hat{f}(n) = \frac{1}{T} \int_0^T f(x) e^{-in\frac{2\pi}{T}x} dx = \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-in\frac{2\pi}{T}x} dx$$

perché la funzione integranda è T -periodica.

In particolare, per $T = 2\pi$, i coefficienti di Fourier delle funzioni 2π -periodiche sono

$$\hat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \quad n \in \mathbb{Z};$$

per $T = 1$, i coefficienti di Fourier delle funzioni di periodo 1 (come, per esempio, la funzione parte frazionaria di un numero reale) sono

$$\hat{f}(n) = \int_0^1 f(x) e^{-in2\pi x} dx = \int_{-1/2}^{1/2} f(x) e^{-in2\pi x} dx, \quad n \in \mathbb{Z}.$$

Esercizio 19.21. (*Coefficienti di Fourier rispetto a seni e coseni*). Si assumano le ipotesi e la notazione della Definizione 19.19. Verificare che

$$S_N(f)(x) = a_0(f) + \sum_{n=1}^N \left(a_n(f) \cos\left(n\frac{2\pi}{T}x\right) + b_n(f) \sin\left(n\frac{2\pi}{T}x\right) \right)$$

dove

$$a_0(f) = \hat{f}(0) = \frac{1}{T} \int_0^T f(t) dt, \quad (19.13)$$

$$a_n(f) = \hat{f}(n) + \hat{f}(-n) = \frac{2}{T} \int_0^T f(t) \cos\left(n\frac{2\pi}{T}t\right) dt, \quad n \geq 1, \quad (19.14)$$

$$b_n(f) = i(\hat{f}(n) - \hat{f}(-n)) = \frac{2}{T} \int_0^T f(t) \sin\left(n\frac{2\pi}{T}t\right) dt, \quad n \geq 1. \quad (19.15)$$

Esercizio 19.22. Sia $f: \mathbb{R} \rightarrow \mathbb{C}$ la funzione 2π -periodica che nell'intervallo $(-\pi, \pi]$ vale $f(x) = x$. Calcolare i coefficienti di Fourier di f e scrivere la serie di Fourier $S(f)$ di f . Scrivere $S(f)$ anche come serie di seni e coseni.

Svolgimento. Per $n \in \mathbb{Z}$, $n \neq 0$, integrando per parti, calcoliamo

$$\begin{aligned} \hat{f}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} x e^{-inx} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} x \frac{d}{dx} \left(\frac{e^{-inx}}{-in} \right) dx \\ &= \frac{1}{2\pi} \left[x \frac{e^{-inx}}{-in} \right]_{x=-\pi}^{x=\pi} - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-inx}}{-in} dx \\ &= \frac{1}{2\pi} \left(\pi \frac{e^{-in\pi}}{-in} + \pi \frac{e^{in\pi}}{-in} \right) + \frac{1}{2\pi in} \int_{-\pi}^{\pi} e^{-inx} dx \\ &= \frac{(-1)^n}{-in} = \frac{i(-1)^n}{n}. \end{aligned}$$

Abbiamo usato le identità (19.6), (19.7), che si usano spesso in questo tipo di esercizi. Poi

$$\hat{f}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} x dx = 0.$$

Abbiamo calcolato

$$\hat{f}(0) = 0, \quad \hat{f}(n) = \frac{i(-1)^n}{n} \quad \forall n \in \mathbb{Z}, n \neq 0,$$

quindi per ogni $N \in \mathbb{N}$ si ha

$$S_N(f)(x) = \sum_{|n| \leq N, n \neq 0} \frac{i(-1)^n}{n} e^{inx}, \quad (19.16)$$

e la serie di Fourier di f è la serie di funzioni

$$S(f)(x) = \sum_{n \in \mathbb{Z}, n \neq 0} \frac{i(-1)^n}{n} e^{inx}.$$

Dalle formule (19.10) (vedi anche l'Esercizio 19.21) si calcola

$$a_n(f) = 0 \quad \forall n \geq 0, \quad b_n(f) = \frac{2(-1)^{n+1}}{n} \quad \forall n \geq 1,$$

da cui

$$S_N(f)(x) = \sum_{n=1}^N \frac{2(-1)^{n+1}}{n} \sin(nx), \quad S(f)(x) = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n} \sin(nx). \quad (19.17)$$

In alternativa, invece di usare le formule (19.10), torniamo all'identità (19.16) e osserviamo che, per ogni $n \geq 1$, la somma dei due termini di indice $n, -n$ è

$$\begin{aligned} & \frac{i(-1)^n}{n} e^{inx} + \frac{i(-1)^{-n}}{-n} e^{-inx} \\ &= \frac{i(-1)^n}{n} [\cos(nx) + i \sin(nx)] + \frac{i(-1)^n}{-n} [\cos(nx) - i \sin(nx)] \\ &= \frac{i(-1)^n}{n} 2i \sin(nx) \\ &= \frac{2(-1)^{n+1}}{n} \sin(nx), \end{aligned}$$

e otteniamo nuovamente (19.17).

Notiamo che la serie di Fourier $S(f)$ si ottiene come "limite formale" di $S_N(f)$ per $N \rightarrow \infty$ (cioè si scrive ∞ al posto di N nella formula delle somme parziali): per *scrivere* la serie di Fourier basta il calcolo dei suoi coefficienti, senza discuterne la convergenza. \square

Esercizio 19.23. Calcolare i coefficienti di Fourier e scrivere la serie di Fourier della funzione 2π -periodica f che nell'intervallo $[0, 2\pi]$ vale

$$f(x) = \frac{(\pi - x)^2}{4}, \quad x \in [0, 2\pi].$$

Esercizio 19.24. Calcolare i coefficienti di Fourier e scrivere la serie di Fourier della funzione 1-periodica "parte frazionaria"

$$f: \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = x - [x],$$

dove $[x]$ è la parte intera del numero reale x , cioè $[x] = \max\{m \in \mathbb{Z} : m \leq x\}$.

Esercizio 19.25. Sia

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \cos(6x).$$

La funzione f è periodica di periodo $T = \pi/3$, ma anche $T = 2\pi/3$, $T = 2\pi$, e più in generale $T = k\pi/3$ per ogni intero positivo k .

Calcolare i coefficienti di Fourier $\hat{f}(n; T)$ e scrivere la serie di Fourier $S(f; T)(x)$ della funzione f rispetto al periodo T nei casi

$$T = \frac{\pi}{3}, \quad T = \frac{2\pi}{3}, \quad T = \pi.$$

Esercizio 19.26. (Caso generale dell'Esercizio 19.25). Ogni funzione di periodo T è anche periodica di periodo T' , con T' multiplo intero di T . Trovare una formula generale che lega i coefficienti di Fourier rispetto al periodo T a quelli rispetto al periodo $T' = kT$.

Lemma 19.27. (Funzioni pari, dispari, reali). Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile in $[0, T]$, e siano $\hat{f}(n)$ i suoi coefficienti di Fourier.

(i) Se f è pari, allora anche \hat{f} è pari, cioè se $f(-x) = f(x)$ per ogni $x \in \mathbb{R}$, allora

$$\hat{f}(-n) = \hat{f}(n), \quad n \in \mathbb{Z}.$$

(ii) Se f è dispari, allora anche \hat{f} è dispari, cioè se $f(-x) = -f(x)$ per ogni $x \in \mathbb{R}$, allora

$$\hat{f}(-n) = -\hat{f}(n), \quad n \in \mathbb{Z}.$$

(iii) Se f è reale, cioè se $f(x) \in \mathbb{R}$ per ogni $x \in \mathbb{R}$, allora

$$\hat{f}(-n) = \overline{\hat{f}(n)}, \quad n \in \mathbb{Z},$$

dove $\overline{\hat{f}(n)}$ è il complesso coniugato di $\hat{f}(n)$.

Lemma 19.28. (Funzioni pari, dispari, reali con seni e coseni). Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile in $[0, T]$, e siano $a_n(f), b_n(f)$ i suoi coefficienti di Fourier rispetto a seni e coseni, dati dalle formule (19.13)-(19.15).

(i) Se f è pari, allora $b_n(f) = 0$ per ogni $n \geq 1$, e la serie di Fourier di f è una serie di soli coseni.

(ii) Se f è dispari, allora $a_n(f) = 0$ per ogni $n \geq 0$, e la serie di Fourier di f è una serie di soli seni.

(iii) Se f è reale, allora $a_n(f), b_n(f)$ sono reali.

Esercizio 19.29. Dimostrare i Lemmi 19.27 e 19.28.

20 Lezione 20

Ci proponiamo di studiare la convergenza della serie di Fourier, a partire dalla convergenza puntuale.

Supponiamo che $f : \mathbb{R} \rightarrow \mathbb{C}$ sia una funzione periodica, di periodo $T > 0$, integrabile nel periodo. Indichiamo, in breve,

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

Dalla Definizione 19.19, per ogni $N \in \mathbb{N}$, $x \in \mathbb{R}$ si ha

$$\begin{aligned} S_N(f)(x) &= \sum_{|n| \leq N} \hat{f}(n) e^{in\omega x} \\ &= \sum_{n=-N}^N \frac{1}{T} \left(\int_{-T/2}^{T/2} f(y) e^{-in\omega y} dy \right) e^{in\omega x} \\ &= \frac{1}{T} \sum_{n=-N}^N \int_{-T/2}^{T/2} f(y) e^{in\omega(x-y)} dy \\ &= \frac{1}{T} \int_{-T/2}^{T/2} f(y) \left(\sum_{n=-N}^N e^{in\omega(x-y)} \right) dy. \end{aligned} \quad (20.1)$$

Lo scambio di somma e integrale non dà alcun problema perché si tratta di una somma finita. La somma nella formula (20.1) gioca quindi un ruolo di primo piano nello studio del polinomio trigonometrico $S_N(f)$; vediamo le sue proprietà.

Definizione 20.1. (*Nucleo di Dirichlet*). Per $N \in \mathbb{N}$ si chiama N -esimo *nucleo di Dirichlet* la funzione

$$D_N(\vartheta) = \sum_{n=-N}^N e^{in\vartheta}, \quad \vartheta \in \mathbb{R}.$$

Lemma 20.2. (Formula integrale per $S_N(f)$ con il nucleo di Dirichlet). *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile nel periodo. Sia $\omega := 2\pi/T$. Allora*

$$S_N(f)(x) = \frac{1}{T} \int_{-T/2}^{T/2} f(y) D_N(\omega(x-y)) dy \quad \forall x \in \mathbb{R}, N \in \mathbb{N}.$$

Dimostrazione. Appena fatta: riprendere la (20.1) e usare la definizione del nucleo di Dirichlet. \square

Lemma 20.3. (Proprietà del nucleo di Dirichlet). *Per ogni $N \in \mathbb{N}$, il nucleo di Dirichlet $D_N(\vartheta)$ soddisfa le seguenti proprietà:*

- (i) $D_N(-\vartheta) = D_N(\vartheta)$ per ogni $\vartheta \in \mathbb{R}$, cioè D_N è pari;
- (ii) $D_N(\vartheta + 2\pi) = D_N(\vartheta)$ per ogni $\vartheta \in \mathbb{R}$, cioè D_N è 2π -periodica;
- (iii) $D_N(\vartheta) \in \mathbb{R}$ per ogni $\vartheta \in \mathbb{R}$, cioè D_N è reale;
- (iv) $D_N(0) = 2N + 1$;

(v) il suo integrale vale

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(\vartheta) d\vartheta = 1, \quad \frac{1}{2\pi} \int_0^{\pi} D_N(\vartheta) d\vartheta = \frac{1}{2};$$

(vi) vale la formula

$$D_N(\vartheta) = \frac{\sin((N + \frac{1}{2})\vartheta)}{\sin(\frac{1}{2}\vartheta)} \quad \forall \vartheta \in \mathbb{R}, \vartheta \neq 2k\pi, k \in \mathbb{Z}.$$

Dimostrazione. (i), (ii), (iii), (iv). Dalla Definizione (20.1) si ha

$$\begin{aligned} D_N(\vartheta) &= e^{-iN\vartheta} + \dots + e^{-i\vartheta} + 1 + e^{i\vartheta} + \dots + e^{iN\vartheta} \\ &= 1 + (e^{i\vartheta} + e^{-i\vartheta}) + (e^{i2\vartheta} + e^{-i2\vartheta}) + \dots + (e^{iN\vartheta} + e^{-iN\vartheta}) \\ &= 1 + 2 \cos(\vartheta) + 2 \cos(2\vartheta) + \dots + 2 \cos(N\vartheta) \\ &= 1 + 2 \sum_{n=1}^N \cos(n\vartheta). \end{aligned}$$

Dunque D_N è pari, 2π -periodica, reale, e in zero vale $1 + 2N$.

(v). Si ha

$$\int_{-\pi}^{\pi} D_N(\vartheta) d\vartheta = \int_{-\pi}^{\pi} \left(\sum_{n=-N}^N e^{in\vartheta} \right) d\vartheta = \sum_{n=-N}^N \left(\int_{-\pi}^{\pi} e^{in\vartheta} d\vartheta \right) = 2\pi$$

perché (vedi Osservazione 19.12)

$$\int_{-\pi}^{\pi} e^{in\vartheta} d\vartheta = 0 \quad \forall n \neq 0,$$

e l'unico contributo non nullo nella somma è il termine per $n = 0$, cioè l'integrale della costante $e^0 = 1$. Poiché D_N è pari, si ha poi

$$\int_0^{\pi} D_N(\vartheta) d\vartheta = \frac{1}{2} \int_{-\pi}^{\pi} D_N(\vartheta) d\vartheta.$$

(vi). Usiamo la proprietà dell'esponenziale complesso

$$e^{in\vartheta} = (e^{i\vartheta})^n, \quad n \in \mathbb{Z}$$

(che si ottiene per induzione dalla proprietà generale $e^z e^c = e^{z+c}$ per ogni $z, c \in \mathbb{C}$). Indichiamo, per brevità, $p := e^{i\vartheta}$. Si ha

$$\begin{aligned} D_N(\vartheta) &= \sum_{n=-N}^N e^{in\vartheta} = \sum_{n=-N}^N p^n = p^{-N} + \dots + p^N \\ &= p^{-N} (1 + p + p^2 + \dots + p^{2N}). \end{aligned}$$

Poiché

$$(1 - p)(1 + p + \dots + p^{2N}) = 1 - p^{2N+1},$$

per $p \neq 1$ si ha

$$(1 + p + \dots + p^{2N}) = \frac{1 - p^{2N+1}}{1 - p} = \frac{p^{2N+1} - 1}{p - 1}$$

(somma parziale della serie geometrica), da cui

$$D_N(\vartheta) = p^{-N}(1 + p + \dots + p^{2N}) = p^{-N} \frac{p^{2N+1} - 1}{p - 1} = \frac{p^{N+1} - p^{-N}}{p - 1}.$$

Mettiamo in evidenza un fattore $p^{1/2}$ al numeratore e al denominatore,

$$D_N(\vartheta) = \frac{p^{N+1} - p^{-N}}{p - 1} = \frac{p^{\frac{1}{2}}(p^{N+\frac{1}{2}} - p^{-N-\frac{1}{2}})}{p^{\frac{1}{2}}(p^{\frac{1}{2}} - p^{-\frac{1}{2}})} = \frac{p^{N+\frac{1}{2}} - p^{-N-\frac{1}{2}}}{p^{\frac{1}{2}} - p^{-\frac{1}{2}}}.$$

Per ogni $x \in \mathbb{R}$ si ha

$$e^{ix} - e^{-ix} = (\cos x + i \sin x) - (\cos x - i \sin x) = 2i \sin x,$$

quindi

$$p^{N+\frac{1}{2}} - p^{-N-\frac{1}{2}} = e^{i(N+\frac{1}{2})\vartheta} - e^{-i(N+\frac{1}{2})\vartheta} = 2i \sin\left(\left(N + \frac{1}{2}\right)\vartheta\right),$$

$$p^{\frac{1}{2}} - p^{-\frac{1}{2}} = e^{i\frac{1}{2}\vartheta} - e^{-i\frac{1}{2}\vartheta} = 2i \sin\left(\frac{1}{2}\vartheta\right).$$

Perciò

$$D_N(\vartheta) = \frac{2i \sin\left(\left(N + \frac{1}{2}\right)\vartheta\right)}{2i \sin\left(\frac{1}{2}\vartheta\right)} = \frac{\sin\left(\left(N + \frac{1}{2}\right)\vartheta\right)}{\sin\left(\frac{1}{2}\vartheta\right)}.$$

Notiamo che $p \neq 1$, cioè $e^{i\vartheta} \neq 1$, se e solo se $\vartheta \neq 2k\pi$, $k \in \mathbb{Z}$. □

Osservazione 20.4. (*Media integrale pesata*). Il grafico della funzione $D_N(\vartheta)$ nel periodo $[-\pi, \pi]$ ha un massimo in $\vartheta = 0$ a quota $2N + 1$, ha uno zero in $\vartheta = \pi/(N + 1/2)$, e poi varie oscillazioni di minore ampiezza man mano che ϑ si allontana da zero e va verso π ; essendo D_N pari, il grafico in $[-\pi, 0]$ è simmetrico.

La formula del Lemma 20.2 può essere vista come una *media integrale pesata*: il fatto che $D_N(\vartheta)$ abbia un valore molto alto in un intervallo stretto intorno a $\vartheta = 0$ significa che, nella media integrale

$$S_N(f)(x) = \frac{1}{T} \int_{-T/2}^{T/2} f(y) D_N(\omega(x - y)) dy,$$

il maggior peso viene dato ai punti y che si trovano nelle vicinanze di x , cioè i punti y per cui $x - y$ è vicino a zero, dove D_N ha il suo picco.

Quando $N \rightarrow \infty$, il grafico di $D_N(\vartheta)$ ha un picco sempre più alto e stretto vicino a $\vartheta = 0$, e compie oscillazioni sempre più ravvicinate (due zeri successivi di $D_N(\vartheta)$ distano $\pi/(N + 1/2)$, come si vede facilmente dalla formula (vi) del Lemma 20.3). Il Lemma di Riemann-Lebesgue (Lemma 20.6) e le osservazioni che lo seguono riguardano l'effetto di oscillazioni veloci sulla media integrale.

Usando le proprietà del nucleo di Dirichlet, riscriviamo la formula del Lemma 20.2 in un altro modo.

Lemma 20.5. (Formula integrale per $S_N(f)$ sul mezzo periodo con il nucleo di Dirichlet). Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ sia una funzione periodica di periodo $T > 0$, integrabile nel periodo, e sia $\omega := 2\pi/T$. Allora

$$S_N(f)(x) = \frac{1}{T} \int_0^{T/2} (f(x+t) + f(x-t)) D_N(\omega t) dt \quad \forall x \in \mathbb{R}, N \in \mathbb{N}.$$

Dimostrazione. Dal Lemma 20.2 si ha

$$S_N(f)(x) = \frac{1}{T} \int_{-T/2}^{T/2} f(y) D_N(\omega(x-y)) dy.$$

Con il cambio di variabile di integrazione $x - y = s$ si ottiene

$$\frac{1}{T} \int_{-T/2}^{T/2} f(y) D_N(\omega(x-y)) dy = \frac{1}{T} \int_{x-T/2}^{x+T/2} f(x-s) D_N(\omega s) ds.$$

Si ha poi

$$\frac{1}{T} \int_{x-T/2}^{x+T/2} f(x-s) D_N(\omega s) ds = \frac{1}{T} \int_{-T/2}^{T/2} f(x-s) D_N(\omega s) ds$$

dalla proprietà di traslazione del periodo (Lemma 19.16), perché $f(x-s) D_N(\omega s)$ è una funzione T -periodica della variabile s : si ha infatti

$$D_N(\omega(s+T)) = D_N(\omega s + 2\pi) = D_N(\omega s)$$

perché $\omega T = 2\pi$ e il nucleo di Dirichlet D_N è 2π -periodico (vedi Lemma 20.3).

Separiamo il dominio di integrazione $[-T/2, T/2]$ come $[-T/2, 0] \cup [0, T/2]$, e scriviamo l'integrale su $[-T/2, T/2]$ come somma di due integrali,

$$\begin{aligned} S_N(f)(x) &= \frac{1}{T} \int_{-T/2}^{T/2} f(x-s) D_N(\omega s) ds \\ &= \underbrace{\frac{1}{T} \int_{-T/2}^0 f(x-s) D_N(\omega s) ds}_{I_1} + \underbrace{\frac{1}{T} \int_0^{T/2} f(x-s) D_N(\omega s) ds}_{I_2}. \end{aligned}$$

Nell'integrale I_1 facciamo il cambio di variabile di integrazione $s = -t$ e abbiamo

$$I_1 = \frac{1}{T} \int_0^{T/2} f(x+t) D_N(-\omega t) dt,$$

e poi

$$I_1 = \frac{1}{T} \int_0^{T/2} f(x+t) D_N(\omega t) dt$$

perché D_N è pari. Nell'integrale I_2 rinominiamo t la variabile di integrazione. Abbiamo dunque

$$\begin{aligned} S_N(f)(x) &= I_1 + I_2 \\ &= \frac{1}{T} \int_0^{T/2} f(x+t) D_N(\omega t) dt + \frac{1}{T} \int_0^{T/2} f(x-t) D_N(\omega t) dt \\ &= \frac{1}{T} \int_0^{T/2} (f(x+t) + f(x-t)) D_N(\omega t) dt. \end{aligned} \quad \square$$

Lemma 20.6. (Lemma di Riemann-Lebesgue per funzioni Riemann-integrabili).
 Sia $f : [a, b] \rightarrow \mathbb{C}$ una funzione integrabile. Allora

$$\int_a^b f(x) \cos(\lambda x) dx \rightarrow 0, \quad \int_a^b f(x) \sin(\lambda x) dx \rightarrow 0, \quad \int_a^b f(x) e^{i\lambda x} dx \rightarrow 0$$

per $|\lambda| \rightarrow \infty, \lambda \in \mathbb{R}$.

Dimostrazione. Sia $f : [a, b] \rightarrow \mathbb{R}$ reale, e sia $\varepsilon > 0$. Dalla definizione di integrale di Riemann, l'integrale di f in $[a, b]$ è il valore

$$I := \int_a^b f(x) dx = \sup A = \inf B,$$

dove A è l'insieme delle somme integrali inferiori di f su $[a, b]$, e B l'insieme di quelle superiori. Per definizione di $I = \sup A$, esiste un elemento $\alpha \in A$ tale che $I - \varepsilon/2 \leq \alpha \leq I$, da cui $0 \leq I - \alpha \leq \varepsilon/2$, cioè esiste una partizione

$$P = \{x_0, x_1, \dots, x_n\}, \quad a = x_0 < x_1 < \dots < x_n = b$$

dell'intervallo $[a, b]$ tale che la somma inferiore

$$\underline{S}(f, P) = \sum_{k=1}^n (x_k - x_{k-1})c_k, \quad c_k := \inf\{f(x) : x \in [x_{k-1}, x_k]\},$$

soddisfa

$$0 \leq \int_a^b f(x) dx - \underline{S}(f, P) \leq \frac{\varepsilon}{2}.$$

Sia $g : [a, b] \rightarrow \mathbb{R}$ la funzione costante a tratti che vale

$$g(x) = \begin{cases} c_1 & \text{per } x \in [a, x_1], \\ c_2 & \text{per } x \in (x_1, x_2], \\ \dots & \\ c_n & \text{per } x \in (x_{n-1}, b]. \end{cases}$$

Il suo integrale è

$$\begin{aligned} \int_a^b g(x) dx &= \int_a^{x_1} g(x) dx + \int_{x_1}^{x_2} g(x) dx + \dots + \int_{x_{n-1}}^b g(x) dx \\ &= (x_1 - a)c_1 + (x_2 - x_1)c_2 + \dots + (b - x_{n-1})c_n \\ &= \underline{S}(f, P). \end{aligned}$$

Inoltre si ha

$$g(x) \leq f(x) \quad \forall x \in [a, b]$$

perché per ogni $k = 1, \dots, n$, dalla definizione di c_k ,

$$c_k \leq f(x) \quad \forall x \in [x_{k-1}, x_k].$$

Per ogni $\lambda \in \mathbb{R}$ si ha

$$\int_a^b f(x) \cos(\lambda x) dx = \underbrace{\int_a^b (f(x) - g(x)) \cos(\lambda x) dx}_{I_1(\lambda)} + \underbrace{\int_a^b g(x) \cos(\lambda x) dx}_{I_2(\lambda)}.$$

Stimiamo i due integrali separatamente. Per $I_1(\lambda)$ usiamo la disuguaglianza $|\cos(\lambda x)| \leq 1$ e il fatto che $f(x) - g(x) \geq 0$, e abbiamo

$$\begin{aligned}
 |I_1(\lambda)| &= \left| \int_a^b (f(x) - g(x)) \cos(\lambda x) dx \right| \\
 &\leq \int_a^b |f(x) - g(x)| |\cos(\lambda x)| dx \\
 &\leq \int_a^b |f(x) - g(x)| dx \\
 &= \int_a^b (f(x) - g(x)) dx \\
 &= \int_a^b f(x) dx - \int_a^b g(x) dx \\
 &= \int_a^b f(x) dx - \underline{S}(f, P) \\
 &\leq \frac{\varepsilon}{2}
 \end{aligned}$$

per qualsiasi $\lambda \in \mathbb{R}$.

Per stimare $I_2(\lambda)$, $\lambda \neq 0$, scriviamo l'integrale su $[a, b]$ come somma degli integrali sui sottointervalli $[a, x_1], (x_1, x_2], \dots, (x_{n-1}, b]$ nei quali $g(x)$ è costante, e calcoliamo

$$\begin{aligned}
 I_2(\lambda) &= \int_a^b g(x) \cos(\lambda x) dx \\
 &= \sum_{k=1}^n \int_{x_{k-1}}^{x_k} g(x) \cos(\lambda x) dx \\
 &= \sum_{k=1}^n \int_{x_{k-1}}^{x_k} c_k \cos(\lambda x) dx \\
 &= \sum_{k=1}^n c_k \frac{\sin(\lambda x_k) - \sin(\lambda x_{k-1})}{\lambda}.
 \end{aligned}$$

Poiché

$$|\sin(\lambda x_k) - \sin(\lambda x_{k-1})| \leq |\sin(\lambda x_k)| + |\sin(\lambda x_{k-1})| \leq 2,$$

stimiamo

$$\begin{aligned}
 |I_2(\lambda)| &= \left| \sum_{k=1}^n c_k \frac{\sin(\lambda x_k) - \sin(\lambda x_{k-1})}{\lambda} \right| \\
 &\leq \sum_{k=1}^n |c_k| \frac{|\sin(\lambda x_k) - \sin(\lambda x_{k-1})|}{|\lambda|} \\
 &\leq \sum_{k=1}^n |c_k| \frac{2}{|\lambda|} \\
 &= \frac{C}{|\lambda|}, \quad C := 2 \sum_{k=1}^n |c_k|.
 \end{aligned}$$

Si ha dunque

$$|I_2(\lambda)| \leq \frac{C}{|\lambda|} \leq \frac{\varepsilon}{2} \quad \forall \lambda \in \mathbb{R}, |\lambda| \geq \frac{2C}{\varepsilon} =: \lambda_0.$$

Abbiamo provato che per ogni $\varepsilon > 0$ esiste $\lambda_0 \geq 0$ tale che

$$\left| \int_a^b f(x) \cos(\lambda x) dx \right| \leq |I_1(\lambda)| + |I_2(\lambda)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

per ogni $\lambda \in \mathbb{R}$, $|\lambda| \geq \lambda_0$: abbiamo quindi dimostrato che

$$\int_a^b f(x) \cos(\lambda x) dx \rightarrow 0 \quad \text{per } |\lambda| \rightarrow \infty,$$

per ogni f reale integrabile. Per f a valori complessi si ha poi

$$\int_a^b f(x) \cos(\lambda x) dx = \int_a^b \operatorname{Re} f(x) \cos(\lambda x) dx + i \int_a^b \operatorname{Im} f(x) \cos(\lambda x) dx$$

e, da quanto appena dimostrato, ciascuno dei due integrali tende a zero per $|\lambda| \rightarrow \infty$.

Il caso del seno $\sin(\lambda x)$ è del tutto simile; per l'esponenziale complesso $e^{i\lambda x}$ basta separare le componenti reale e immaginaria e applicare quanto già provato per seno e coseno. \square

Osservazione 20.7. (*Oscillazioni ad alta frequenza e cancellazioni*). Abbiamo visto che i due integrali $I_1(\lambda)$, $I_2(\lambda)$ nella dimostrazione del Lemma 20.6 sono entrambi piccoli, ma lo sono per motivi diversi: $I_1(\lambda)$ è piccolo perché l'integrale della differenza $\int_a^b (f - g)$ è piccolo, cioè perché g è stata scelta come buona approssimazione (nel senso dell'integrale) di f ; invece il fatto che $I_2(\lambda)$ sia piccolo per λ grande è un effetto dovuto alle cancellazioni nelle oscillazioni, ed è il punto centrale del Lemma di Riemann-Lebesgue. Un modo di rivederlo è il seguente.

Quando $g(x) = c$ è una funzione costante, il prodotto $g(x) \cos(\lambda x) = c \cos(\lambda x)$ è un coseno, funzione periodica di periodo $\tau = 2\pi/\lambda$, che ha media nulla su ogni periodo. Il suo integrale su un intervallo $[a, b]$ è dato soltanto da eventuali contributi provenienti da periodi incompleti, perché su ogni periodo completo l'integrale del coseno è zero: supponiamo, per esempio, che $[a, b]$ contenga n periodi completi

$$[a, a + \tau], [a + \tau, a + 2\tau], \dots, [a + (n - 1)\tau, a + n\tau],$$

mentre l'ultimo intervallo $[a + n\tau, b]$ non è necessariamente un intero periodo, ma può essere anche più corto, cioè $b - (a + n\tau) \leq \tau$. Allora

$$\begin{aligned} \int_a^b \cos(\lambda x) dx &= \int_a^{a+\tau} \cos(\lambda x) dx + \int_{a+\tau}^{a+2\tau} \cos(\lambda x) dx + \dots + \int_{a+n\tau}^b \cos(\lambda x) dx \\ &= \int_{a+n\tau}^b \cos(\lambda x) dx, \end{aligned}$$

perché

$$\int_s^{s+\tau} \cos(\lambda x) dx = 0 \quad \forall s \in \mathbb{R}.$$

Poiché $|\cos(\lambda x)| \leq 1$, si ha dunque

$$\begin{aligned} \left| \int_a^b \cos(\lambda x) dx \right| &= \left| \int_{a+n\tau}^b \cos(\lambda x) dx \right| \\ &\leq \int_{a+n\tau}^b |\cos(\lambda x)| dx \\ &\leq b - (a + n\tau) \\ &\leq \tau = \frac{2\pi}{\lambda} \rightarrow 0 \quad \text{per } |\lambda| \rightarrow \infty. \end{aligned}$$

Nella dimostrazione del Lemma 20.6 questo effetto di cancellazione su ogni periodo completo si ha su ogni intervallo $(x_{k-1}, x_k]$ in cui $g(x)$ è costante.

Osservazione 20.8. (*Lemma di Riemann-Lebesgue con integrazione per parti*). In alternativa alla dimostrazione vista sopra, per f di classe C^1 possiamo integrare per parti:

$$\begin{aligned} \int_a^b f(x) \cos(\lambda x) dx &= \int_a^b f(x) \frac{d}{dx} \left(\frac{\sin(\lambda x)}{\lambda} \right) dx \\ &= \left[f(x) \frac{\sin(\lambda x)}{\lambda} \right]_a^b - \int_a^b f'(x) \frac{\sin(\lambda x)}{\lambda} dx \\ &= \frac{1}{\lambda} \left\{ f(b) \sin(\lambda b) - f(a) \sin(\lambda a) - \int_a^b f'(x) \sin(\lambda x) dx \right\}. \end{aligned}$$

L'effetto dell'integrazione per parti è la comparsa del fattore λ al denominatore. Poiché $|\sin(\vartheta)| \leq 1$ per ogni $\vartheta \in \mathbb{R}$, si ha

$$\begin{aligned} \left| f(b) \sin(\lambda b) - f(a) \sin(\lambda a) - \int_a^b f'(x) \sin(\lambda x) dx \right| \\ \leq |f(b)| + |f(a)| + \int_a^b |f'(x)| dx =: C_1, \end{aligned}$$

da cui, come sopra,

$$\left| \int_a^b f(x) \cos(\lambda x) dx \right| \leq \frac{C_1}{|\lambda|} \rightarrow 0 \quad (|\lambda| \rightarrow \infty).$$

Se f è soltanto integrabile, si può dimostrare che esiste $g \in C^1$ tale che

$$\int_a^b |f(x) - g(x)| dx \leq \frac{\varepsilon}{2},$$

e quindi usare l'integrazione per parti nell'integrale di $g(x) \cos(\lambda x)$. In altre parole, si possono usare le funzioni C^1 invece di quelle costanti a tratti per approssimare le funzioni integrabili.

D'altra parte, la dimostrazione in cui g è funzione costante a tratti ha il vantaggio di non richiedere niente di più che gli strumenti di Analisi 1.

Come conseguenza immediata del Lemma di Riemann-Lebesgue abbiamo la seguente proposizione.

Corollario 20.9. (I coefficienti di Fourier tendono a zero). Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile nel periodo. Allora i coefficienti di Fourier di f tendono a zero,

$$\hat{f}(n) \rightarrow 0 \quad \text{per } |n| \rightarrow \infty.$$

Dimostrazione. Per ogni $n \in \mathbb{Z}$ si ha

$$\hat{f}(n) = \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-in\omega x} dx = \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{i\lambda x} dx, \quad \lambda = -n\omega = -n \frac{2\pi}{T},$$

e $|\lambda| \rightarrow \infty$ per $|n| \rightarrow \infty$; dunque la tesi segue dal Lemma di Riemann-Lebesgue. \square

Il Corollario 20.9 assicura che $\hat{f}(n) \rightarrow 0$; se inoltre f è di classe C^1 , l'integrazione per parti vista nell'Osservazione 20.8 dà un'informazione quantitativa sulla *velocità* con cui i coefficienti $\hat{f}(n)$ tendono a zero.

Iteriamo questo procedimento: più derivate ha $f(x)$, più volte possiamo integrare per parti, e più potenze di n compaiono al denominatore, cioè più veloce è il decadimento di $\hat{f}(n)$ verso zero.

Lemma 20.10. (Regolarità di $f(x)$ e decadimento di $\hat{f}(n)$). Sia $k \geq 1$ intero, e sia $f : \mathbb{R} \rightarrow \mathbb{C}$ periodica di periodo $T > 0$, di classe $C^k(\mathbb{R})$. Allora

$$\hat{f}(n) = o\left(\frac{1}{|n|^k}\right) \quad \text{per } |n| \rightarrow \infty,$$

cioè $n^k \hat{f}(n) \rightarrow 0$ per $|n| \rightarrow \infty$.

Inoltre i coefficienti di Fourier di f e quelli delle sue derivate f', f'', \dots sono legati dalle formule

$$\widehat{(f^{(j)})}(n) = (in\omega)^j \hat{f}(n), \quad n \in \mathbb{Z}, \quad j = 1, \dots, k, \quad (20.2)$$

dove $\omega = 2\pi/T$.

Dimostrazione. Essendo $f \in C^k$, le sue derivate $f^{(j)}$, $j = 1, \dots, k$ sono funzioni continue, quindi integrabili in $[0, T]$, e i coefficienti di Fourier di $f^{(j)}$ sono tutti ben definiti. Integrando per parti, si ha

$$\begin{aligned} \widehat{(f')}(n) &= \frac{1}{T} \int_0^T f'(x) e^{-in\omega x} dx \\ &= \frac{1}{T} \left[f(x) e^{-in\omega x} \right]_{x=0}^{x=T} - \frac{1}{T} \int_0^T f(x) e^{-in\omega x} (-in\omega) dx \\ &= \frac{1}{T} \left[f(T) e^{-in2\pi} - f(0) \right] + in\omega \frac{1}{T} \int_0^T f(x) e^{-in\omega x} dx \\ &= in\omega \hat{f}(n), \end{aligned} \quad (20.3)$$

che è la (20.2) per $j = 1$. Abbiamo usato le identità

$$\omega T = 2\pi, \quad f(T) = f(0), \quad e^{-in2\pi} = 1.$$

Supponiamo che la (20.2) valga per un certo j , con $1 \leq j < k$. Ripercorrendo la (20.3) con $f^{(j)}$ al posto di f si trova

$$\widehat{(f^{(j+1)})}(n) = in\omega \widehat{(f^{(j)})}(n),$$

e dunque dall'ipotesi induttiva si ottiene la (20.2) per $j + 1$. Per induzione, questo prova la formula (20.2) per $j = 1, \dots, k$.

In particolare, per $j = k$, $n \neq 0$ si ha

$$\hat{f}(n) = \frac{\widehat{(f^{(k)})}(n)}{(in\omega)^k} = C \frac{\widehat{(f^{(k)})}(n)}{n^k},$$

dove $C = 1/(i\omega)^k$ è indipendente da n . Poi

$$\widehat{(f^{(k)})}(n) \rightarrow 0 \quad \text{per } |n| \rightarrow \infty$$

dal Lemma 20.6 di Riemann-Lebesgue, quindi $n^k \hat{f}(n) \rightarrow 0$. □

Corollario 20.11. (Convergenza totale della serie di Fourier per f di classe C^2).
Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ periodica di periodo $T > 0$, di classe $C^2(\mathbb{R})$. Allora la serie di Fourier di f è una serie di funzioni totalmente (e quindi uniformemente) convergente in \mathbb{R} .

Dimostrazione. La serie di Fourier di f è la serie di funzioni

$$S(f)(x) = \sum_{n \in \mathbb{N}} \hat{f}(n) e^{in\omega x} = \lim_{N \rightarrow +\infty} \sum_{n=-N}^N \hat{f}(n) e^{in\omega x}, \quad \omega = 2\pi/T.$$

Poiché $|e^{i\vartheta}| = 1$ per ogni ϑ reale, la sup-norma del termine generale della serie è, per ogni n fissato,

$$\|\hat{f}(n)e^{in\omega x}\|_{\infty, \mathbb{R}} = \sup_{x \in \mathbb{R}} |\hat{f}(n)e^{in\omega x}| = |\hat{f}(n)|.$$

Dal Lemma 20.10 di regolarità e decadimento, $\hat{f}(n) = o(1/n^2)$ per $|n| \rightarrow \infty$, quindi esiste $C > 0$ tale che

$$|\hat{f}(n)| \leq \frac{C}{n^2} \quad \forall n \in \mathbb{Z}, n \neq 0.$$

Poiché la serie armonica generalizzata $\sum_{n=1}^{\infty} 1/n^2$ converge, la serie delle sup-norme $\sum_{n \in \mathbb{Z}} |\hat{f}(n)|$ è convergente. □

21 Lezione 21

Il Corollario 20.11 non ci dice se la serie di Fourier $S(f)$ converge ad f o ad un'altra funzione; riguardo alla convergenza puntuale, usando il Lemma 20.6 di Riemann-Lebesgue possiamo dimostrare il risultato seguente.

Teorema 21.1. (Teorema di convergenza puntuale per serie di Fourier con condizione di Dini per l'integrale di Riemann). *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ periodica di periodo $T > 0$, integrabile sul periodo. Sia*

$$p \in \mathbb{R}, \quad \beta \in \mathbb{C}, \quad 0 < \delta \leq T/2,$$

e supponiamo che la funzione $h : [0, \delta] \rightarrow \mathbb{C}$ che vale $h(0) = 0$,

$$h(t) = \frac{f(p+t) + f(p-t) - 2\beta}{t} \quad \forall t \in (0, \delta],$$

sia limitata e integrabile in $[0, \delta]$. Allora la serie di Fourier di f nel punto p converge, e il suo limite è β , cioè

$$S(f)(p) = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{in\omega p} = \lim_{N \rightarrow \infty} S_N(f)(p) = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} \hat{f}(n) e^{in\omega p} = \beta,$$

dove $\omega = 2\pi/T$. Se, inoltre, f è continua in p , allora

$$S(f)(p) = f(p).$$

Dimostrazione. Dal Lemma 20.5, per ogni N si ha

$$S_N(f)(p) = \frac{1}{T} \int_0^{T/2} (f(p+t) + f(p-t)) D_N(\omega t) dt.$$

Dal Lemma 20.3, e dal cambio di variabile

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

nell'integrale, si ha

$$\frac{1}{T} \int_0^{T/2} D_N(\omega t) dt = \frac{1}{2\pi} \int_0^\pi D_N(\vartheta) d\vartheta = \frac{1}{2}.$$

Quindi

$$\beta = 2\beta \frac{1}{2} = 2\beta \frac{1}{T} \int_0^{T/2} D_N(\omega t) dt = \frac{1}{T} \int_0^{T/2} 2\beta D_N(\omega t) dt,$$

e la differenza $S_N(f)(p) - \beta$ si può scrivere come integrale

$$S_N(f)(p) - \beta = \frac{1}{T} \int_0^{T/2} (f(p+t) + f(p-t) - 2\beta) D_N(\omega t) dt.$$

Separiamo

$$\int_0^{T/2} = \int_0^\delta + \int_\delta^{T/2},$$

cioè scriviamo

$$S_N(f)(p) - \beta = A_N + B_N,$$

dove

$$A_N := \frac{1}{T} \int_0^\delta (f(p+t) + f(p-t) - 2\beta) D_N(\omega t) dt,$$

$$B_N := \frac{1}{T} \int_\delta^{T/2} (f(p+t) + f(p-t) - 2\beta) D_N(\omega t) dt,$$

e studiamo separatamente i due integrali.

Cominciamo con l'integrale B_N . Dal Lemma 20.3,

$$D_N(\omega t) = \frac{\sin((N + \frac{1}{2})\omega t)}{\sin(\frac{1}{2}\omega t)} \quad \forall t \in (0, T/2].$$

La funzione

$$t \mapsto f(p+t) + f(p-t) - 2\beta$$

è integrabile in $[0, T/2]$ (per ipotesi, f è integrabile in ogni periodo). Nell'intervallo $[\delta, T/2]$ la funzione $1/\sin(\frac{1}{2}\omega t)$ è continua, quindi limitata e integrabile, con

$$1 \leq \frac{1}{\sin(\frac{1}{2}\omega t)} \leq \frac{1}{\sin(\frac{1}{2}\omega \delta)} \quad \forall t \in [\delta, T/2].$$

Quindi il loro prodotto

$$g(t) := \frac{f(p+t) + f(p-t) - 2\beta}{\sin(\frac{1}{2}\omega t)}$$

è una funzione integrabile in $[\delta, T/2]$. Si ha allora

$$B_N = \frac{1}{T} \int_\delta^{T/2} \frac{f(p+t) + f(p-t) - 2\beta}{\sin(\frac{1}{2}\omega t)} \sin((N + \frac{1}{2})\omega t) dt$$

$$= \frac{1}{T} \int_\delta^{T/2} g(t) \sin(\lambda t) dt, \quad \lambda := (N + \frac{1}{2})\omega.$$

Poiché $\lambda \rightarrow \infty$ per $N \rightarrow \infty$, dal Lemma 20.6 di Riemann-Lebesgue si ha

$$\lim_{N \rightarrow \infty} B_N = 0.$$

Passiamo ora all'integrale A_N , che, moltiplicando e dividendo per $\frac{1}{2}\omega t$, riscriviamo come

$$A_N = \frac{1}{T} \int_0^\delta \frac{f(p+t) + f(p-t) - 2\beta}{t} \frac{\frac{1}{2}\omega t}{\sin(\frac{1}{2}\omega t)} \frac{1}{\frac{1}{2}\omega} \sin((N + \frac{1}{2})\omega t) dt.$$

La funzione $h(t)$ dell'enunciato è integrabile in $[0, \delta]$ per ipotesi; la funzione

$$\varphi : [0, \delta] \rightarrow \mathbb{R}, \quad \varphi(0) = 1, \quad \varphi(t) = \frac{\frac{1}{2}\omega t}{\sin(\frac{1}{2}\omega t)} \quad \forall t \in (0, \delta] \quad (21.1)$$

è continua e limitata in $[0, \delta]$. Dunque il prodotto $h(t)\varphi(t)$ è integrabile in $[0, \delta]$, e quindi

$$\lim_{N \rightarrow \infty} A_N = 0$$

grazie al Lemma 20.6 di Riemann-Lebesgue. Abbiamo provato che $S_N(f)(p) \rightarrow \beta$ per $N \rightarrow \infty$.

Sia ora f continua nel punto p . Per ipotesi, h è limitata in $[0, \delta]$, cioè esiste $M \geq 0$ tale che $|h(t)| \leq M$ per ogni $t \in [0, \delta]$. Quindi

$$0 \leq |f(p+t) + f(p-t) - 2\beta| \leq Mt \quad \forall t \in (0, \delta],$$

da cui

$$\lim_{t \rightarrow 0^+} |f(p+t) + f(p-t) - 2\beta| = 0.$$

D'altra parte, per la continuità di f in p si ha

$$\lim_{t \rightarrow 0^+} |f(p+t) + f(p-t) - 2\beta| = 2|f(p) - \beta|.$$

Quindi $f(p) = \beta$. □

Il Lemma di Riemann-Lebesgue può essere esteso agli integrali di Riemann generalizzati o impropri.

Lemma 21.2. (Lemma di Riemann-Lebesgue per funzioni Riemann-integrabili in senso generalizzato). *Sia $f : (a, b] \rightarrow \mathbb{C}$ una funzione definita in $(a, b]$. Supponiamo che, per ogni $c \in (a, b]$, nell'intervallo $[c, b]$ la funzione f sia limitata e integrabile. Supponiamo inoltre che esista, finito, il limite*

$$\int_a^b |f(x)| dx := \lim_{c \rightarrow a} \int_c^b |f(x)| dx.$$

Allora per ogni $\lambda \in \mathbb{R}$ esiste finito il limite

$$\int_a^b f(x) \cos(\lambda x) dx := \lim_{c \rightarrow a} \int_c^b f(x) \cos(\lambda x) dx$$

e si ha

$$\lim_{|\lambda| \rightarrow \infty} \int_a^b f(x) \cos(\lambda x) dx = \lim_{|\lambda| \rightarrow \infty} \left(\lim_{c \rightarrow a} \int_c^b f(x) \cos(\lambda x) dx \right) = 0.$$

Lo stesso risultato vale con $\sin(\lambda x)$ oppure $e^{i\lambda x}$ al posto di $\cos(\lambda x)$.

Dimostrazione. Consideriamo la funzione

$$\varphi : (a, b] \rightarrow \mathbb{R}, \quad c \mapsto \varphi(c) := \int_c^b |f(x)| dx.$$

Per ipotesi, φ è ben definita ed esiste, finito, il suo limite

$$\ell := \lim_{c \rightarrow a} \varphi(c) = \lim_{c \rightarrow a} \int_c^b |f(x)| dx = \int_a^b |f(x)| dx \in \mathbb{R}.$$

Per ogni $\lambda \in \mathbb{R}$, definiamo

$$\psi_\lambda : (a, b] \rightarrow \mathbb{C}, \quad c \mapsto \psi_\lambda(c) := \int_c^b f(x) \cos(\lambda x) dx.$$

La funzione ψ_λ è ben posta perché la funzione integranda $f(x) \cos(\lambda x)$ è integrabile (f è integrabile in $[c, b]$, e il coseno è funzione continua). Per ogni $c, t \in (a, b]$ si ha

$$\begin{aligned} \psi_\lambda(c) - \psi_\lambda(t) &= \int_c^b f(x) \cos(\lambda x) dx - \int_t^b f(x) \cos(\lambda x) dx \\ &= \int_c^t f(x) \cos(\lambda x) dx \end{aligned}$$

e similmente

$$\varphi(c) - \varphi(t) = \int_c^b |f(x)| dx - \int_t^b |f(x)| dx = \int_c^t |f(x)| dx.$$

Quindi

$$\begin{aligned} |\psi_\lambda(c) - \psi_\lambda(t)| &= \left| \int_c^t f(x) \cos(\lambda x) dx \right| \\ &\leq \left| \int_c^t |f(x)| dx \right| = |\varphi(c) - \varphi(t)| \quad \forall c, t \in (a, b]. \end{aligned} \quad (21.2)$$

Mostriamo che $\psi_\lambda(c)$ ha limite per $c \rightarrow a$. Sia (c_j) una successione in $(a, b]$ che converge ad a . Dalla (21.2), per ogni $j, k \in \mathbb{N}$ si ha

$$|\psi_\lambda(c_j) - \psi_\lambda(c_k)| \leq |\varphi(c_j) - \varphi(c_k)|. \quad (21.3)$$

Poiché $\varphi(c) \rightarrow \ell$ per $c \rightarrow a$, la successione $(\varphi(c_j))$ è convergente, e quindi è di Cauchy. Di conseguenza, dalla (21.3), la successione $(\psi_\lambda(c_j))$ è di Cauchy, e quindi converge (in \mathbb{C}). Abbiamo dimostrato che

$$\text{per ogni successione } c_j \rightarrow a, \text{ la successione } (\psi_\lambda(c_j)) \text{ converge.} \quad (21.4)$$

Ora proviamo che il limite $\lim_{j \rightarrow \infty} \psi_\lambda(c_j)$ non dipende dalla successione (c_j) , cioè che sostituendo (c_j) con un'altra successione (x_j) convergente ad a si trova lo stesso limite $\lim_{j \rightarrow \infty} \psi_\lambda(x_j) = \lim_{j \rightarrow \infty} \psi_\lambda(c_j)$.

Supponiamo dunque di avere due successioni $c_j \rightarrow a$, $x_j \rightarrow a$. Per la (21.4), $\psi_\lambda(c_j)$, $\psi_\lambda(x_j)$ convergono, diciamo

$$\psi_\lambda(c_j) \rightarrow p, \quad \psi_\lambda(x_j) \rightarrow q.$$

Supponiamo, per assurdo, che $p \neq q$. Definiamo allora la successione (t_j) come

$$(t_1, t_2, t_3, \dots) = (c_1, x_1, c_2, x_2, c_3, x_3, \dots, c_j, x_j, \dots).$$

Si prova facilmente che $t_j \rightarrow a$. La successione $(\psi_\lambda(t_j))$, invece, non ha limite, perché ha due estratte che convergono a limiti diversi: l'estratta degli indici pari

$$\psi_\lambda(t_2), \psi_\lambda(t_4), \psi_\lambda(t_6), \dots$$

converge a q , mentre quella degli indici dispari

$$\psi_\lambda(t_1), \psi_\lambda(t_3), \psi_\lambda(t_5), \dots$$

converge a p . Dunque $t_j \rightarrow a$, $\psi_\lambda(t_j)$ non ha limite, contro la (21.4), assurdo.

Abbiamo dimostrato che per ogni successione $c_j \rightarrow a$ la successione $\psi_\lambda(c_j)$ converge sempre allo stesso limite, che indichiamo $p_\lambda \in \mathbb{C}$; dal “Teorema ponte” di Analisi 1 si ha allora

$$\lim_{c \rightarrow a} \psi_\lambda(c) = p_\lambda \in \mathbb{C}.$$

Questo limite, ricordando la definizione di $\psi_\lambda(c)$, *definisce* l’integrale generalizzato

$$\int_a^b f(x) \cos(\lambda x) dx := \lim_{c \rightarrow a} \int_c^b f(x) \cos(\lambda x) dx = \lim_{c \rightarrow a} \psi_\lambda(c) = p_\lambda, \quad \lambda \in \mathbb{R}.$$

Rimane da dimostrare che $p_\lambda \rightarrow 0$ per $|\lambda| \rightarrow \infty$. Ricordiamo la (21.2), cioè

$$|\psi_\lambda(t) - \psi_\lambda(c)| \leq |\varphi(t) - \varphi(c)| \quad \forall t, c \in (a, b], \lambda \in \mathbb{R}.$$

Per ogni λ, c fissati, passiamo al limite per $t \rightarrow a$, e otteniamo

$$|p_\lambda - \psi_\lambda(c)| \leq |\ell - \varphi(c)| \quad \forall c \in (a, b], \lambda \in \mathbb{R}.$$

Dunque

$$\begin{aligned} |p_\lambda| &\leq |p_\lambda - \psi_\lambda(c)| + |\psi_\lambda(c)| \\ &\leq |\ell - \varphi(c)| + |\psi_\lambda(c)| \quad \forall c \in (a, b], \lambda \in \mathbb{R}. \end{aligned} \quad (21.5)$$

Sia $\varepsilon > 0$. Poiché $\lim_{c \rightarrow a} \varphi(c) = \ell$, esiste $c_0 \in (a, b]$ tale che

$$|\ell - \varphi(c_0)| \leq \varepsilon/2.$$

Dal Lemma 20.6 di Riemann-Lebesgue applicato all’intervallo $[c_0, b]$ si ha

$$\psi_\lambda(c_0) = \int_{c_0}^b f(x) \cos(\lambda x) dx \rightarrow 0 \quad \text{per } |\lambda| \rightarrow \infty,$$

quindi esiste $\lambda_0 \geq 0$ tale che

$$|\psi_\lambda(c_0)| \leq \varepsilon/2 \quad \forall |\lambda| \geq \lambda_0.$$

Dunque, dalla (21.5),

$$|p_\lambda| \leq |\ell - \varphi(c_0)| + |\psi_\lambda(c_0)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \forall |\lambda| \geq \lambda_0.$$

Abbiamo dimostrato che per ogni $\varepsilon > 0$ esiste $\lambda_0 \geq 0$ tale che

$$|p_\lambda| = \left| \int_a^b f(x) \cos(\lambda x) dx \right| \leq \varepsilon \quad \forall |\lambda| \geq \lambda_0,$$

che è la tesi. □

Se usiamo il Lemma di Riemann-Lebesgue 21.2 per gli integrali generalizzati, otteniamo una versione più forte del Teorema 21.1.

Teorema 21.3. (Teorema di convergenza puntuale per serie di Fourier con condizione di Dini per l'integrale di Riemann generalizzato). Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ periodica di periodo $T > 0$, integrabile sul periodo. Sia

$$p \in \mathbb{R}, \quad \beta \in \mathbb{C}, \quad 0 < \delta \leq T/2,$$

e supponiamo che esista, finito, il limite

$$\int_0^\delta \frac{|f(p+t) + f(p-t) - 2\beta|}{t} dt := \lim_{c \rightarrow 0^+} \int_c^\delta \frac{|f(p+t) + f(p-t) - 2\beta|}{t} dt.$$

Allora la serie di Fourier di f nel punto p converge a β , cioè

$$S(f)(p) = \lim_{N \rightarrow \infty} S_N(f)(p) = \beta.$$

Se, inoltre, f è continua nel punto p , allora

$$S(f)(p) = f(p).$$

Dimostrazione. Ripercorriamo la dimostrazione del Teorema 21.1, mantenendo le stesse notazioni.

L'analisi di B_N è identica, e si ha dunque $B_N \rightarrow 0$ per $N \rightarrow \infty$.

Per stimare A_N , osserviamo, come già fatto per il Teorema 21.1, che la funzione $\varphi(t)$ (definita in (21.1)) è continua, e quindi limitata e integrabile, in $[0, \delta]$, mentre la funzione $h(t)$ (definita nell'enunciato del Teorema 21.1) è limitata e integrabile in ogni intervallo $[c, \delta]$ con $0 < c \leq \delta$. Quindi il prodotto $h(t)\varphi(t)$ è limitato e integrabile in $[c, \delta]$ per ogni $c \in (0, \delta]$.

La funzione

$$(0, \delta] \rightarrow \mathbb{R}, \quad c \mapsto \int_c^\delta |h(t)\varphi(t)| dt$$

è decrescente, quindi, per monotonia, ha limite per $c \rightarrow 0^+$; tale limite è finito perché per ogni $c \in (0, \delta]$ si ha

$$\int_c^\delta |h(t)\varphi(t)| dt \leq \|\varphi\|_\infty \int_c^\delta |h(t)| dt$$

e, passando al limite,

$$\lim_{c \rightarrow 0^+} \int_c^\delta |h(t)\varphi(t)| dt \leq \|\varphi\|_\infty \lim_{c \rightarrow 0^+} \int_c^\delta |h(t)| dt = \|\varphi\|_\infty \int_0^\delta |h(t)| dt < \infty.$$

La funzione $h(t)\varphi(t)$ soddisfa le ipotesi del Lemma 21.2 in $(0, \delta]$. Quindi, grazie a tale lemma, $A_N \rightarrow 0$ per $N \rightarrow \infty$. Questo prova che $S_N(f)(p) \rightarrow \beta$ per $N \rightarrow \infty$.

Infine, sia f continua in p , e supponiamo, per assurdo, che $f(p) \neq \beta$. Dunque

$$\rho := |f(p) - \beta| > 0.$$

Dalla continuità di f in p si ha

$$\lim_{t \rightarrow 0^+} |f(p+t) + f(p-t) - 2\beta| = |2f(p) - 2\beta| = 2\rho.$$

Dalla definizione di limite, esiste $\delta_0 > 0$, con $\delta_0 \leq \delta$, tale che

$$|f(p+t) + f(p-t) - 2\beta| \geq \rho \quad \forall t \in (0, \delta_0].$$

Dunque

$$\begin{aligned} \int_c^\delta \frac{|f(p+t) + f(p-t) - 2\beta|}{t} dt &\geq \int_c^{\delta_0} \frac{|f(p+t) + f(p-t) - 2\beta|}{t} dt \\ &\geq \int_c^{\delta_0} \frac{\rho}{t} dt = \rho \log\left(\frac{\delta_0}{c}\right) \quad \forall c \in (0, \delta_0]. \end{aligned}$$

Poiché $\log(\delta_0/c) \rightarrow \infty$ per $c \rightarrow 0^+$, si ha

$$\lim_{c \rightarrow 0^+} \int_c^\delta \frac{|f(p+t) + f(p-t) - 2\beta|}{t} dt = \infty,$$

contro le ipotesi. Questo prova che $f(p) = \beta$. □

I due risultati che seguono sono altri criteri di convergenza puntuale per la serie di Fourier, conseguenza del Teorema 21.3. Sono anche utili per gli esercizi, soprattutto il secondo, perché in molti casi le sue ipotesi sono semplici da verificare.

Proposizione 21.4. (Criterio di convergenza puntuale per serie di Fourier con condizioni di Hölder). *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile nel periodo. Sia*

$$p \in \mathbb{R}, \quad \beta_1, \beta_2 \in \mathbb{C}, \quad 0 < \delta \leq T/2, \quad \alpha > 0, \quad C \geq 0,$$

e supponiamo che f abbia limiti destro e sinistro

$$\lim_{x \rightarrow p^+} f(x) = \beta_1, \quad \lim_{x \rightarrow p^-} f(x) = \beta_2$$

con condizioni di Hölder

$$|f(p+t) - \beta_1| \leq Ct^\alpha, \quad |f(p-t) - \beta_2| \leq Ct^\alpha \quad \forall t \in (0, \delta].$$

Allora la serie di Fourier $S(f)$ converge nel punto p , e il suo limite è la media aritmetica dei limiti destro e sinistro di f , cioè

$$S(f)(p) = \lim_{N \rightarrow \infty} S_N(f)(p) = \frac{\beta_1 + \beta_2}{2}.$$

Se inoltre f è continua in p , allora

$$S(f)(p) = f(p).$$

Dimostrazione. Sia

$$\beta := \frac{\beta_1 + \beta_2}{2}.$$

Per $t \in (0, \delta]$ si ha

$$\begin{aligned} |f(p+t) + f(p-t) - 2\beta| &= |f(p+t) - \beta_1 + f(p-t) - \beta_2| \\ &\leq |f(p+t) - \beta_1| + |f(p-t) - \beta_2| \leq 2Ct^\alpha. \end{aligned}$$

Quindi, per ogni $c \in (0, \delta]$,

$$\int_c^\delta \frac{|f(p+t) + f(p-t) - 2\beta|}{t} dt \leq \int_c^\delta 2Ct^{\alpha-1} dt = \frac{2C}{\alpha}(\delta^\alpha - c^\alpha) \leq \frac{2C\delta^\alpha}{\alpha}.$$

La funzione

$$(0, \delta] \rightarrow \mathbb{R}, \quad c \mapsto \int_c^\delta \frac{|f(p+t) + f(p-t) - 2\beta|}{t} dt$$

è dunque decrescente e limitata in $(0, \delta]$, quindi ha limite finito per $c \rightarrow 0^+$. La tesi allora segue dal Teorema 21.3. \square

Proposizione 21.5. (Criterio di convergenza puntuale per serie di Fourier con limite destro/sinistro di f, f'). *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile nel periodo. Sia*

$$p \in \mathbb{R}, \quad 0 < \delta \leq T/2,$$

e sia f derivabile in $(p - \delta, p) \cup (p, p + \delta)$. Supponiamo che esistano, finiti, i limiti destro/sinistro di f, f' per $x \rightarrow p^\pm$, cioè

$$\lim_{x \rightarrow p^+} f(x) = \beta_1, \quad \lim_{x \rightarrow p^-} f(x) = \beta_2, \quad \lim_{x \rightarrow p^+} f'(x) = \gamma_1, \quad \lim_{x \rightarrow p^-} f'(x) = \gamma_2$$

con $\beta_1, \beta_2, \gamma_1, \gamma_2 \in \mathbb{C}$. Allora la serie di Fourier $S(f)$ converge nel punto p , e il suo limite è la media aritmetica dei limiti destro/sinistro di f , cioè

$$S(f)(p) = \lim_{N \rightarrow \infty} S_N(f)(x) = \frac{\beta_1 + \beta_2}{2}.$$

Se inoltre f è continua in p , allora

$$S(f)(p) = f(p).$$

Dimostrazione. Dalla regola di de l'Hôpital (che vale per la parte reale e per quella immaginaria di f , e quindi vale per f) si ha

$$\lim_{t \rightarrow 0^+} \frac{f(p+t) - \beta_1}{t} = \lim_{t \rightarrow 0^+} f'(p+t) = \gamma_1 \in \mathbb{C}.$$

La funzione $(f(p+t) - \beta_1)/t$, essendo convergente, è limitata in un intorno destro di 0, cioè esistono $\delta_0 \in (0, \delta]$, $C \geq 0$ tali che

$$\left| \frac{f(p+t) - \beta_1}{t} \right| \leq C \quad \forall t \in (0, \delta_0).$$

Analogamente per il limite sinistro ($t \rightarrow 0^-$). Quindi valgono le condizioni di Hölder della Proposizione 21.4 con esponente $\alpha = 1$. Applichiamo la Proposizione 21.4 e otteniamo la tesi. \square

Nonostante i coefficienti di Fourier $\hat{f}(n)$ di una funzione f siano calcolati tramite integrale su tutto il periodo (cioè un'operazione non locale, che coinvolge i valori della funzione in tutto il suo periodo), il teorema seguente mostra che il comportamento della serie di Fourier di f in un punto p è determinato soltanto dai valori di f vicini al punto p .

Teorema 21.6. (Teorema di localizzazione di Riemann). *Siano $f, g : \mathbb{R} \rightarrow \mathbb{C}$ due funzioni periodiche di periodo $T > 0$, integrabili sul periodo. Sia*

$$p \in \mathbb{R}, \quad 0 < \delta \leq T/2,$$

e supponiamo che

$$f(x) = g(x) \quad \forall x \in [p - \delta, p + \delta].$$

Allora

$$\lim_{N \rightarrow \infty} (S_N(f)(p) - S_N(g)(p)) = 0.$$

Dimostrazione. Sia $h(x) := f(x) - g(x)$. Dalla formula del Lemma 20.5 si ha

$$\begin{aligned} S_N(f)(p) - S_N(g)(p) &= \frac{1}{T} \int_0^{T/2} (f(p+t) + f(p-t)) D_N(\omega t) dt \\ &\quad - \frac{1}{T} \int_0^{T/2} (g(p+t) + g(p-t)) D_N(\omega t) dt \\ &= \frac{1}{T} \int_0^{T/2} (h(p+t) + h(p-t)) D_N(\omega t) dt. \end{aligned}$$

Separiamo

$$\int_0^{T/2} = \int_0^\delta + \int_\delta^{T/2}.$$

Per ipotesi,

$$h(p+t) = 0, \quad h(p-t) = 0 \quad \forall t \in [0, \delta].$$

Perciò l'integrale su $[0, \delta]$ è nullo,

$$\frac{1}{T} \int_0^\delta (h(p+t) + h(p-t)) D_N(\omega t) dt = 0,$$

e rimane l'integrale su $[\delta, T/2]$,

$$\begin{aligned} S_N(f)(p) - S_N(g)(p) &= \frac{1}{T} \int_\delta^{T/2} (h(p+t) + h(p-t)) D_N(\omega t) dt \\ &= \frac{1}{T} \int_\delta^{T/2} \frac{h(p+t) + h(p-t)}{\sin(\frac{1}{2}\omega t)} \sin((N + \frac{1}{2})\omega t) dt. \end{aligned}$$

Dal Lemma 20.6 di Riemann-Lebesgue, questo integrale tende a zero per $N \rightarrow \infty$ (è simile al termine $B_N \rightarrow 0$ nella dimostrazione del Teorema 21.1). \square

Lemma 21.7. (Densità delle funzioni continue in quelle integrabili). (i) *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile sul periodo, con*

$$|\operatorname{Re} f(x)| \leq M, \quad |\operatorname{Im} f(x)| \leq K \quad \forall x \in \mathbb{R}$$

per certe costanti M, K . Allora per ogni $\varepsilon > 0$ esiste una funzione $g : \mathbb{R} \rightarrow \mathbb{C}$, continua, periodica di periodo T , tale che

$$|\operatorname{Re} g(x)| \leq M, \quad |\operatorname{Im} g(x)| \leq K \quad \forall x \in \mathbb{R}$$

e

$$\int_0^T |f(x) - g(x)| dx \leq \varepsilon.$$

(ii) Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile sul periodo. Allora esiste una successione (f_k) di funzioni continue $f_k : \mathbb{R} \rightarrow \mathbb{C}$, periodiche di periodo T , tali che

$$\int_0^T |f(x) - f_k(x)| dx \rightarrow 0 \quad \text{per } k \rightarrow \infty, \quad \|f_k\|_\infty \leq \sqrt{2}\|f\|_\infty \quad \forall k \in \mathbb{N}.$$

Dimostrazione. (i) Supponiamo dapprima che f sia a valori reali. Sia $\varepsilon > 0$. Dalla definizione di integrale di Riemann (vedi anche la dimostrazione del Lemma 20.6), esiste una partizione

$$P = \{x_0, x_1, \dots, x_n\}, \quad 0 = x_0 < x_1 < \dots < x_n = T$$

dell'intervallo $[0, T]$ tale che la somma inferiore

$$\underline{S}(f, P) = \sum_{k=1}^n (x_k - x_{k-1})c_k, \quad c_k := \inf\{f(x) : x \in [x_{k-1}, x_k]\},$$

soddisfa

$$0 \leq \int_0^T f(x) dx - \underline{S}(f, P) \leq \frac{\varepsilon}{2}.$$

Sia $h : \mathbb{R} \rightarrow \mathbb{R}$ la funzione periodica di periodo T , costante a tratti, che nell'intervallo $[0, T)$ vale

$$h(x) = \begin{cases} c_1 & \text{per } x \in [0, x_1), \\ c_2 & \text{per } x \in [x_1, x_2), \\ \dots & \\ c_n & \text{per } x \in [x_{n-1}, T). \end{cases}$$

Il suo integrale sul periodo è

$$\begin{aligned} \int_0^T h(x) dx &= \int_0^{x_1} h(x) dx + \int_{x_1}^{x_2} h(x) dx + \dots + \int_{x_{n-1}}^T h(x) dx \\ &= x_1 c_1 + (x_2 - x_1)c_2 + \dots + (T - x_{n-1})c_n \\ &= \underline{S}(f, P). \end{aligned}$$

Poiché

$$-M \leq f(x) \leq M \quad \forall x \in \mathbb{R},$$

per ogni k , passando all'inf sull'intervallo $[x_{k-1}, x_k]$, si ha

$$-M \leq c_k \leq M \quad \forall k = 1, \dots, n.$$

Quindi

$$-M \leq h(x) \leq M \quad \forall x \in \mathbb{R}.$$

Inoltre si ha

$$h(x) \leq f(x) \quad \forall x \in \mathbb{R}$$

perché per ogni $k = 1, \dots, n$, dalla definizione di c_k ,

$$c_k \leq f(x) \quad \forall x \in [x_{k-1}, x_k].$$

Poiché $f(x) - h(x) \geq 0$, si ha allora

$$\begin{aligned} \int_0^T |f(x) - h(x)| dx &= \int_0^T (f(x) - h(x)) dx \\ &= \int_0^T f(x) dx - \int_0^T h(x) dx \\ &= \int_0^T f(x) dx - \underline{S}(f, P) \leq \frac{\varepsilon}{2}. \end{aligned}$$

La funzione h ha le proprietà richieste dal lemma, eccetto la continuità. Per ottenere una funzione che sia vicina ad h ma sia anche continua, consideriamo un intervallo $[p-r, p+r]$ intorno a ogni punto p in cui h ha una discontinuità di salto, e, in tale intervallo, al posto di h consideriamo una funzione continua che dia un “raccordo” tra i valori di h a sinistra e a destra dell’intervallo $[p-r, p+r]$. Il modo più semplice di fare il raccordo continuo è considerare il tratto di retta che unisce i punti

$$(p-r, h(p-r)), \quad (p+r, h(p+r)).$$

Fissiamo dunque $r > 0$ tale che gli intervalli

$$[-r, r], [x_1-r, x_1+r], [x_2-r, x_2+r], \dots, [x_{n-1}-r, x_{n-1}+r], [T-r, T+r]$$

non abbiano intersezione, cioè

$$0 < r < \frac{x_k - x_{k-1}}{2} \quad \forall k = 1, \dots, n$$

(ricordiamo che $x_0 = 0$, $x_n = T$). Nei punti di $[0, T]$ che non appartengono agli intervalli dedicati al raccordo, definiamo la funzione g come coincidente con h , cioè definiamo

$$g(x) := h(x) \quad \forall x \in [0, T] \setminus \left(\bigcup_{k=0}^n [x_k - r, x_k + r] \right).$$

Nei punti di $[0, T]$ che appartengono agli intervalli di raccordo definiamo g come il tratto di retta descritto sopra, cioè per ogni $k = 0, \dots, n$ definiamo

$$g(x) := \frac{c_{k+1} - c_k}{2r} (x - x_k) + \frac{c_k + c_{k+1}}{2} \quad \forall x \in [0, T] \cap [x_k - r, x_k + r],$$

dove $c_0 := c_n$ e $c_{n+1} := c_1$. In questo modo abbiamo definito la funzione g in $[0, T]$, continua per costruzione: si ha infatti

$$\begin{aligned} g(x_k + r) &= c_{k+1} = h(x_k + r) \quad \forall k = 0, \dots, n-1, \\ g(x_k - r) &= c_k = h(x_k - r) \quad \forall k = 1, \dots, n. \end{aligned}$$

Inoltre

$$g(T) = \frac{c_1 + c_n}{2} = g(0),$$

quindi possiamo definire g su tutto \mathbb{R} , periodica di periodo T , ponendo $g(x+mT) = g(x)$ per ogni $x \in [0, T]$, $m \in \mathbb{Z}$.

Per $x \in [x_k - r, x_k + r]$, i valori $g(x)$ sono compresi tra le costanti c_k, c_{k+1} . Ricordando che $-M \leq c_k \leq M$ per ogni c_k , si ha

$$-M \leq g(x) \leq M \quad \forall x \in \mathbb{R}.$$

Per stimare l'integrale, osserviamo che la differenza $g - h$ è nulla fuori dagli intervalli di raccordo, mentre negli intervalli di raccordo si ha

$$|g(x) - h(x)| \leq |c_{k+1} - c_k| \leq 2M \quad \forall x \in [x_k - r, x_k + r].$$

Quindi

$$\begin{aligned} & \int_0^T |g(x) - h(x)| dx \\ &= \int_0^r |g(x) - h(x)| dx + \int_{x_1-r}^{x_1+r} |g(x) - h(x)| dx + \dots \\ & \quad \dots + \int_{x_{n-1}-r}^{x_{n-1}+r} |g(x) - h(x)| dx + \int_{T-r}^T |g(x) - h(x)| dx \\ & \leq 2M(r + 2r + \dots + 2r + r) = 4Mnr. \end{aligned}$$

Scegliamo r tale che $4Mnr \leq \varepsilon/2$, e abbiamo

$$\int_0^T |f(x) - g(x)| dx \leq \int_0^T |f(x) - h(x)| dx + \int_0^T |h(x) - g(x)| dx \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Sia ora f a valori complessi. Applicando alla parte reale $\operatorname{Re} f$ e a quella immaginaria $\operatorname{Im} f$ ciò che abbiamo dimostrato, per ogni $\varepsilon_1 > 0$ esiste una funzione continua, T -periodica, reale g_1 tale che

$$|g_1(x)| \leq M \quad \forall x \in \mathbb{R}, \quad \int_0^T |\operatorname{Re} f(x) - g_1(x)| dx \leq \varepsilon_1,$$

e per ogni $\varepsilon_2 > 0$ esiste una funzione continua, T -periodica, reale g_2 tale che

$$|g_2(x)| \leq K \quad \forall x \in \mathbb{R}, \quad \int_0^T |\operatorname{Im} f(x) - g_2(x)| dx \leq \varepsilon_2.$$

Definiamo

$$g : \mathbb{R} \rightarrow \mathbb{C}, \quad g(x) = g_1(x) + ig_2(x).$$

Dalla disuguaglianza elementare “(ipotenusa) \leq (somma dei cateti)”

$$\sqrt{a^2 + b^2} \leq a + b \quad \forall a, b \geq 0,$$

si ha

$$|f - g| = \sqrt{|\operatorname{Re} f - g_1|^2 + |\operatorname{Im} f - g_2|^2} \leq |\operatorname{Re} f - g_1| + |\operatorname{Im} f - g_2|,$$

da cui

$$\int_0^T |f(x) - g(x)| dx \leq \int_0^T |\operatorname{Re} f - g_1| dx + \int_0^T |\operatorname{Im} f - g_2| dx \leq \varepsilon_1 + \varepsilon_2.$$

Quindi per ogni $\varepsilon > 0$ scegliamo $\varepsilon_1 = \varepsilon_2 = \varepsilon/2$, definiamo g come descritto, e l'enunciato (i) è dimostrato.

(ii) Data f , per ogni $x \in \mathbb{R}$ si ha $|\operatorname{Re} f(x)| \leq \|f\|_\infty$, $|\operatorname{Im} f(x)| \leq \|f\|_\infty$. Per ogni $k \in \mathbb{N}$, applicando (i) con $\varepsilon = 1/k$, esiste f_k continua, T -periodica, tale che

$$\int_0^T |f(x) - f_k(x)| dx \leq \frac{1}{k}, \quad |\operatorname{Re} f_k(x)| \leq \|f\|_\infty, \quad |\operatorname{Im} f_k(x)| \leq \|f\|_\infty \quad \forall x \in \mathbb{R},$$

da cui

$$|f_k(x)| = (|\operatorname{Re} f_k(x)|^2 + |\operatorname{Im} f_k(x)|^2)^{\frac{1}{2}} \leq \sqrt{2}\|f\|_\infty. \quad \square$$

Osservazione 21.8. (*Raccordo lineare come media integrale e convoluzione*). Il raccordo lineare tra due valori si può anche scrivere come media integrale. Vediamo dapprima il caso più semplice di una funzione costante a tratti con un unico salto: sia

$$h(x) = \begin{cases} \alpha & \text{per } x \leq p, \\ \beta & \text{per } x > p, \end{cases}$$

con $\alpha, \beta, p \in \mathbb{R}$, e sia $g(x)$ la sua media integrale nell'intervallo $[x-r, x+r]$, cioè

$$g(x) = \frac{1}{2r} \int_{x-r}^{x+r} h(t) dt \quad \forall x \in \mathbb{R}.$$

Se $x \leq p-r$, allora $(x-r, x+r) \subset (-\infty, p)$, quindi $h = \alpha$ su $(x-r, x+r)$, e si ha

$$g(x) = \frac{1}{2r} \int_{x-r}^{x+r} \alpha dt = \alpha = h(x) \quad \forall x \leq p-r.$$

Se $x \geq p+r$, allora $(x-r, x+r) \subset (p, \infty)$, quindi $h = \beta$ su $(x-r, x+r)$, e si ha

$$g(x) = \frac{1}{2r} \int_{x-r}^{x+r} \beta dt = \beta = h(x) \quad \forall x \geq p+r.$$

Quindi

$$g(x) = h(x) \quad \forall x \in \mathbb{R}, \quad |x-p| \geq r.$$

Se invece $|x-p| < r$, allora il punto di salto p si trova all'interno dell'intervallo di integrazione $(x-r, x+r)$, e

$$\begin{aligned} g(x) &= \frac{1}{2r} \int_{x-r}^{x+r} h(t) dt \\ &= \frac{1}{2r} \int_{x-r}^p \alpha dt + \frac{1}{2r} \int_p^{x+r} \beta dt \\ &= \frac{1}{2r} \alpha(p-x+r) + \frac{1}{2r} \beta(x+r-p) \\ &= \frac{\alpha + \beta}{2} + \frac{\beta - \alpha}{2r} (x-p) \quad \forall x \in (p-r, p+r), \end{aligned}$$

che è il segmento di retta che unisce i punti $(p-r, \alpha)$ e $(p+r, \beta)$.

Possiamo usare questa osservazione come *modo alternativo di definire* g a partire da h nella dimostrazione del Lemma 21.7. Siano dunque h, r come nella dimostrazione del Lemma 21.7, e definiamo

$$g : \mathbb{R} \rightarrow \mathbb{R}, \quad g(x) := \frac{1}{2r} \int_{x-r}^{x+r} h(t) dt, \quad x \in \mathbb{R}.$$

I punti in cui h ha un salto sono i punti dell'insieme

$$\sigma := \{x_k + mT : k = 1, \dots, n, m \in \mathbb{Z}\}.$$

Se x è al di fuori degli intervalli $(p - r, p + r)$, $p \in \sigma$, allora h non ha punti di salto nell'intervallo di integrazione $(x - r, x + r)$, quindi h è costante in tale intervallo, e di conseguenza h è uguale alla sua media in tale intervallo; dunque

$$g(x) = h(x) \quad \forall x \in \mathbb{R} \setminus \bigcup_{p \in \sigma} (p - r, p + r), \quad \text{cioè } |x - p| \geq r \quad \forall p \in \sigma.$$

Inoltre dalla disuguaglianza

$$-M \leq h(t) \leq M \quad \forall t \in \mathbb{R},$$

integrando su $(x - r, x + r)$ e dividendo per $2r$, si ottiene

$$-M \leq \frac{1}{2r} \int_{x-r}^{x+r} h(t) dt = g(x) \leq M \quad \forall x \in \mathbb{R}.$$

Poiché $g(x), h(x) \in [-M, M]$, si ha

$$|g(x) - h(x)| \leq 2M \quad \forall x \in \mathbb{R}.$$

Usando il fatto che $g - h = 0$ fuori dagli intervalli $(p - r, p + r)$, $p \in \sigma$, si ha

$$\begin{aligned} & \int_0^T |g(x) - h(x)| dx \\ &= \int_0^r |g(x) - h(x)| dx + \int_{x_1-r}^{x_1+r} |g(x) - h(x)| dx + \dots \\ & \dots + \int_{x_{n-1}-r}^{x_{n-1}+r} |g(x) - h(x)| dx + \int_{T-r}^T |g(x) - h(x)| dx \\ & \leq 2M(r + 2r + \dots + 2r + r) = 4Mnr. \end{aligned}$$

La funzione g è periodica di periodo T perché, facendo il cambio di variabile di integrazione $t = s + T$ e usando il fatto che h è T -periodica, si ha

$$\begin{aligned} g(x + T) &= \frac{1}{2r} \int_{(x+T)-r}^{(x+T)+r} h(t) dt \\ &= \frac{1}{2r} \int_{x-r}^{x+r} h(s + T) ds = \frac{1}{2r} \int_{x-r}^{x+r} h(s) ds = g(x) \quad \forall x \in \mathbb{R}. \end{aligned}$$

Per provare la continuità di g , fissiamo $x \in \mathbb{R}$, e consideriamo un punto $y \in (x, x+r)$, cosicché

$$x - r < y - r < x + r < y + r.$$

Allora

$$g(x) - g(y) = \frac{1}{2r} \int_{x-r}^{x+r} h(t) dt - \frac{1}{2r} \int_{y-r}^{y+r} h(t) dt;$$

separando gli integrali

$$\int_{x-r}^{x+r} = \int_{x-r}^{y-r} + \int_{y-r}^{x+r}, \quad \int_{y-r}^{y+r} = \int_{y-r}^{x+r} + \int_{x+r}^{y+r},$$

l'integrale su $(y - r, x + r)$ si cancella, e rimane

$$g(x) - g(y) = \frac{1}{2r} \int_{x-r}^{y-r} h(t) dt - \frac{1}{2r} \int_{x+r}^{y+r} h(t) dt.$$

Quindi

$$\begin{aligned} |g(x) - g(y)| &\leq \frac{1}{2r} \left| \int_{x-r}^{y-r} h(t) dt \right| + \frac{1}{2r} \left| \int_{x+r}^{y+r} h(t) dt \right| \\ &\leq \frac{1}{2r} \int_{x-r}^{y-r} |h(t)| dt + \frac{1}{2r} \int_{x+r}^{y+r} |h(t)| dt \leq 2 \frac{1}{2r} M |x - y|. \end{aligned}$$

Analogamente per $y \in (x - r, x)$ si ottiene la stessa disuguaglianza. Quindi $g(y) \rightarrow g(x)$ per $y \rightarrow x$. Usando la media integrale abbiamo quindi ottenuto la funzione continua che ci serviva nella dimostrazione del Lemma 21.7.

Abbiamo già parlato (vedi Osservazione 7.1) della convoluzione come media integrale pesata e come metodo principale per approssimare una funzione con una funzione più regolare. La media integrale sull'intervallo $[x - r, x + r]$ è anch'essa una convoluzione: sia ψ la funzione costante a tratti

$$\psi(t) = \begin{cases} \frac{1}{2r} & \text{per } |t| \leq r, \\ 0 & \text{per } |t| > r. \end{cases}$$

Allora la media integrale di una qualunque funzione integrabile h (non necessariamente costante a tratti) è la convoluzione di h con ψ :

$$\int h(t) \psi(x - t) dt = \frac{1}{2r} \int_{x-r}^{x+r} h(t) dt.$$

22 Lezione 22

Torniamo a trattare la convoluzione, ora vista come operazione tra funzioni periodiche.

Definizione 22.1. (*Convoluzione di funzioni periodiche*). Siano $f, g : \mathbb{R} \rightarrow \mathbb{C}$ funzioni periodiche di periodo $T > 0$. Diciamo *convoluzione T -periodica*, o semplicemente *convoluzione*, di f e g la funzione

$$(f * g)(x) := \frac{1}{T} \int_0^T f(y)g(x-y) dy, \quad x \in \mathbb{R}.$$

Lemma 22.2. (Proprietà della convoluzione). Siano $f, g, h : \mathbb{R} \rightarrow \mathbb{C}$ funzioni periodiche di periodo $T > 0$, integrabili sul periodo, e sia $c \in \mathbb{C}$. Allora si ha

- (i) *Periodicità:* $(f * g)(x + T) = (f * g)(x)$ per ogni $x \in \mathbb{R}$.
- (ii) *Proprietà associativa:* $(f * g) * h = f * (g * h)$.
- (iii) *Proprietà commutativa:* $f * g = g * f$.
- (iv) *Proprietà distributiva:* $f * (g + h) = f * g + f * h$.
- (v) *Prodotto per le costanti:* $(cf) * g = c(f * g) = f * (cg)$.
- (vi) *Sup-norma:* $\|f * g\|_\infty \leq \|g\|_\infty \frac{1}{T} \int_0^T |f(x)| dx \leq \|g\|_\infty \|f\|_\infty$.
- (vii) *Uniforme continuità:* $f * g$ è uniformemente continua in \mathbb{R} .
- (viii) *Coefficienti di Fourier:* $\widehat{(f * g)}(n) = \hat{f}(n)\hat{g}(n)$ per ogni $n \in \mathbb{Z}$.

Esercizio 22.3. Usando i cambi di variabile negli integrali e la proprietà di traslazione del periodo (Lemma 19.16), dimostrare il Lemma 22.2 nel caso di f, g, h funzioni continue (esercizio da fare prima di leggere la dimostrazione che segue).

Dimostrazione del Lemma 22.2. (i) Siano f, g come nell'enunciato. Essendo g funzione T -periodica, per ogni $x \in \mathbb{R}$ si ha banalmente

$$\begin{aligned} (f * g)(x + T) &= \frac{1}{T} \int_0^T f(y)g(x + T - y) dy \\ &= \frac{1}{T} \int_0^T f(y)g(x - y) dy = (f * g)(x). \end{aligned}$$

(ii) Supponiamo che f, g, h siano continue. Per ogni $x \in \mathbb{R}$ si ha

$$\begin{aligned} ((f * g) * h)(x) &= \frac{1}{T} \int_0^T (f * g)(y)h(x - y) dy \\ &= \frac{1}{T} \int_0^T \left(\frac{1}{T} \int_0^T f(z)g(y - z) dz \right) h(x - y) dy. \end{aligned}$$

Si tratta di un integrale doppio di funzione continua su dominio normale (integrale di $F(y, z) = f(z)g(y - z)h(x - y)$ sul dominio $[0, T]^2$). Come visto in Analisi 2, valgono allora le formule di riduzione, che ci permettono di scambiare l'ordine di integrazione:

$$= \frac{1}{T} \int_0^T f(z) \left(\frac{1}{T} \int_0^T g(y - z)h(x - y) dy \right) dz.$$

Per ogni z fissato, facciamo il cambio di variabile $y - z = t$, $dy = dt$, nell'integrale interno:

$$= \frac{1}{T} \int_0^T f(z) \left(\frac{1}{T} \int_{-z}^{-z+T} g(t)h(x - z - t) dt \right) dz.$$

Poiché $g(t)h(x - z - t)$ è T -periodica come funzione di t , il suo integrale su $[-z, -z+T]$ è uguale a quello su $[0, T]$ (traslazione del periodo, vedi Lemma 19.16):

$$\begin{aligned} &= \frac{1}{T} \int_0^T f(z) \left(\frac{1}{T} \int_0^T g(t)h(x - z - t) dt \right) dz \\ &= \frac{1}{T} \int_0^T f(z)(g * h)(x - z) dz \\ &= (f * (g * h))(x). \end{aligned}$$

Questo prova la proprietà associativa per funzioni continue. Lasciamo in sospeso, per il momento, il caso generale in cui f, g, h non sono necessariamente continue.

(iii) Siano f, g come nell'enunciato. Per ogni $x \in \mathbb{R}$, facendo il cambio di variabile $x - y = z$ e usando la proprietà di traslazione del periodo si ha

$$\begin{aligned} (f * g)(x) &= \frac{1}{T} \int_0^T f(y)g(x - y) dy \\ &= \frac{1}{T} \int_{x-T}^x f(x - z)g(z) dz \\ &= \frac{1}{T} \int_0^T g(z)f(x - z) dz = (g * f)(x). \end{aligned}$$

(iv) Siano f, g, h come nell'enunciato. Per ogni $x, y \in \mathbb{R}$ si ha, banalmente,

$$f(y)(g + h)(x - y) = f(y)g(x - y) + f(y)h(x - y),$$

e, integrando in y , si conclude.

(v) Ovvio.

(vi) Siano f, g come nell'enunciato. Per ogni $x \in \mathbb{R}$ si ha

$$\begin{aligned} |(f * g)(x)| &= \left| \frac{1}{T} \int_0^T f(y)g(x - y) dy \right| \\ &\leq \frac{1}{T} \int_0^T |f(y)| |g(x - y)| dy \\ &\leq \frac{1}{T} \int_0^T |f(y)| \|g\|_\infty dy = \frac{\|g\|_\infty}{T} \int_0^T |f(x)| dx; \end{aligned}$$

quindi

$$\|f * g\|_\infty = \sup_{x \in \mathbb{R}} |(f * g)(x)| \leq \frac{\|g\|_\infty}{T} \int_0^T |f(x)| dx.$$

Poi

$$\frac{1}{T} \int_0^T |f(x)| dx \leq \frac{1}{T} \int_0^T \|f\|_\infty dx = \|f\|_\infty.$$

(vii) Supponiamo che g sia continua. Dal Teorema di Cantor (applicato alla parte reale e alla parte immaginaria di g), g è uniformemente continua in ogni intervallo compatto, e quindi, essendo periodica, g è uniformemente continua in \mathbb{R} . Sia $\varepsilon > 0$, e sia

$$\varepsilon_1 := \frac{\varepsilon}{1 + \|f\|_\infty}.$$

Allora esiste $\delta > 0$ tale che

$$|g(x) - g(z)| \leq \varepsilon_1 \quad \forall x, z \in \mathbb{R}, |x - z| \leq \delta.$$

Per ogni $x, z \in \mathbb{R}$ si ha

$$\begin{aligned} (f * g)(x) - (f * g)(z) &= \frac{1}{T} \int_0^T f(y)g(x - y) dy - \frac{1}{T} \int_0^T f(y)g(z - y) dy \\ &= \frac{1}{T} \int_0^T f(y)(g(x - y) - g(z - y)) dy. \end{aligned}$$

Quindi, per $|x - z| \leq \delta$, si ha

$$\begin{aligned} |(f * g)(x) - (f * g)(z)| &\leq \frac{1}{T} \int_0^T |f(y)| |g(x - y) - g(z - y)| dy \\ &\leq \frac{1}{T} \int_0^T \|f\|_\infty \varepsilon_1 dy = \|f\|_\infty \varepsilon_1 < \varepsilon. \end{aligned}$$

Dunque per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che

$$|(f * g)(x) - (f * g)(z)| \leq \varepsilon \quad \forall x, z \in \mathbb{R}, |x - z| \leq \delta.$$

Questo prova che $f * g$ è uniformemente continua in \mathbb{R} per f integrabile e g continua. Lasciamo in sospenso, per il momento, il caso generale in cui g è soltanto integrabile.

(viii). Supponiamo dapprima che f, g siano continue. Indichiamo $\omega := 2\pi/T$. Per ogni $n \in \mathbb{Z}$ si ha

$$\begin{aligned} \widehat{(f * g)}(n) &= \frac{1}{T} \int_0^T (f * g)(x) e^{-in\omega x} dx \\ &= \frac{1}{T} \int_0^T \left(\frac{1}{T} \int_0^T f(y)g(x - y) dy \right) e^{-in\omega x} dx; \end{aligned}$$

questo è un integrale doppio di funzione continua $F(x, y) = f(y)g(x - y)e^{-in\omega x}$ su dominio normale $[0, T]^2$, quindi valgono le formule di riduzione, cioè possiamo cambiare l'ordine di integrazione:

$$\begin{aligned} &= \frac{1}{T} \int_0^T f(y) \left(\frac{1}{T} \int_0^T g(x - y) e^{-in\omega x} dx \right) dy \\ &= \frac{1}{T} \int_0^T f(y) \left(\frac{1}{T} \int_0^T g(x - y) e^{-in\omega(x - y + y)} dx \right) dy \\ &= \frac{1}{T} \int_0^T f(y) e^{-in\omega y} \left(\frac{1}{T} \int_0^T g(x - y) e^{-in\omega(x - y)} dx \right) dy. \end{aligned}$$

Per ogni y fissato, facciamo il cambio di variabile $x - y = t$, $dx = dt$ nell'integrale interno:

$$= \frac{1}{T} \int_0^T f(y) e^{-in\omega y} \left(\frac{1}{T} \int_{-y}^{-y+T} g(t) e^{-in\omega t} dt \right) dy$$

e poi traslazione del periodo:

$$\begin{aligned} &= \frac{1}{T} \int_0^T f(y) e^{-in\omega y} \left(\frac{1}{T} \int_0^T g(t) e^{-in\omega t} dt \right) dy \\ &= \left(\frac{1}{T} \int_0^T f(y) e^{-in\omega y} dy \right) \left(\frac{1}{T} \int_0^T g(t) e^{-in\omega t} dt \right) \\ &= \hat{f}(n) \hat{g}(n). \end{aligned}$$

Abbiamo provato (viii) per f, g continue; resta in sospeso il caso generale di f, g integrabili.

Ora completiamo la dimostrazione del lemma, provando, per densità, il caso generale di (ii), (vii), (viii).

Conclusione di (vii). Siano f, g come nell'enunciato, e sia (g_k) una successione di funzioni continue come quella fornita dal Lemma 21.7. Allora, dalle proprietà che abbiamo già dimostrato, si ha

$$\begin{aligned} \|f * g_k - f * g\|_\infty &= \|f * (g_k - g)\|_\infty \\ &\leq \|f\|_\infty \frac{1}{T} \int_0^T |g_k(x) - g(x)| dx \rightarrow 0 \quad \text{per } k \rightarrow \infty. \end{aligned}$$

Quindi $f * g_k$ converge a $f * g$ uniformemente in \mathbb{R} .

Per ogni k , la funzione $f * g_k$ è continua in \mathbb{R} , quindi $f * g$ è continua in \mathbb{R} essendo il limite uniforme di una successione di funzioni continue. Inoltre, dal Teorema di Cantor, $f * g$ è uniformemente continua in ogni intervallo compatto; essendo periodica, $f * g$ è uniformemente continua in \mathbb{R} .

Conclusione di (viii). Siano f, g come nell'enunciato, e siano $(f_k), (g_k)$ successioni di funzioni continue come quelle fornite dal Lemma 21.7. Consideriamo la differenza

$$f_k * g_k - f * g = (f_k - f) * g_k + f * (g_k - g).$$

Il primo termine soddisfa

$$\begin{aligned} \|(f_k - f) * g_k\|_\infty &\leq \|g_k\|_\infty \frac{1}{T} \int_0^T |f_k(x) - f(x)| dx \\ &\leq \sqrt{2} \|g\|_\infty \frac{1}{T} \int_0^T |f_k(x) - f(x)| dx \rightarrow 0 \quad \text{per } k \rightarrow \infty, \end{aligned}$$

il secondo soddisfa

$$\|f * (g_k - g)\|_\infty \leq \|f\|_\infty \frac{1}{T} \int_0^T |g_k(x) - g(x)| dx \rightarrow 0 \quad \text{per } k \rightarrow \infty,$$

da cui

$$\|f_k * g_k - f * g\|_\infty \leq \|(f_k - f) * g_k\|_\infty + \|f * (g_k - g)\|_\infty \rightarrow 0 \quad \text{per } k \rightarrow \infty,$$

cioè $f_k * g_k \rightarrow f * g$ uniformemente in \mathbb{R} . Poi

$$\begin{aligned} & (\widehat{f_k * g_k})(n) - (\widehat{f * g})(n) \\ &= \frac{1}{T} \int_0^T (f_k * g_k)(x) e^{-in\omega x} dx - \frac{1}{T} \int_0^T (f * g)(x) e^{-in\omega x} dx \\ &= \frac{1}{T} \int_0^T [(f_k * g_k)(x) - (f * g)(x)] e^{-in\omega x} dx \end{aligned}$$

e

$$\begin{aligned} |(\widehat{f_k * g_k})(n) - (\widehat{f * g})(n)| &\leq \frac{1}{T} \int_0^T |(f_k * g_k)(x) - (f * g)(x)| |e^{-in\omega x}| dx \\ &\leq \frac{1}{T} \int_0^T \|f_k * g_k - f * g\|_\infty dx \\ &= \|f_k * g_k - f * g\|_\infty \rightarrow 0 \quad \text{per } k \rightarrow \infty. \end{aligned}$$

Dunque

$$\lim_{k \rightarrow \infty} (\widehat{f_k * g_k})(n) = (\widehat{f * g})(n) \quad \forall n \in \mathbb{Z}.$$

D'altra parte, per ogni $n \in \mathbb{Z}$, si ha

$$\begin{aligned} \hat{f}_k(n) - \hat{f}(n) &= \frac{1}{T} \int_0^T f_k(x) e^{-in\omega x} dx - \frac{1}{T} \int_0^T f(x) e^{-in\omega x} dx \\ &= \frac{1}{T} \int_0^T (f_k(x) - f(x)) e^{-in\omega x} dx \end{aligned}$$

e

$$|\hat{f}_k(n) - \hat{f}(n)| \leq \frac{1}{T} \int_0^T |f_k(x) - f(x)| dx \rightarrow 0 \quad \text{per } k \rightarrow \infty.$$

Analogamente, per ogni $n \in \mathbb{Z}$ si ha

$$|\hat{g}_k(n) - \hat{g}(n)| \rightarrow 0 \quad \text{per } k \rightarrow \infty,$$

cioè

$$\lim_{k \rightarrow \infty} \hat{f}_k(n) = \hat{f}(n), \quad \lim_{k \rightarrow \infty} \hat{g}_k(n) = \hat{g}(n), \quad \forall n \in \mathbb{Z}.$$

Poiché f_k, g_k sono continue, abbiamo provato che per ogni n si ha

$$(\widehat{f_k * g_k})(n) = \hat{f}_k(n) \hat{g}_k(n).$$

Passando al limite per $k \rightarrow \infty$ si ottiene (viii).

Conclusione di (ii). Siano f, g, h come nell'enunciato, e siano $(f_k), (g_k), (h_k)$ tre successioni come quelle fornite dal Lemma 21.7. Da ciò che abbiamo dimostrato sopra, si ha

$$\|f_k * g_k - f * g\|_\infty \rightarrow 0, \quad \int_0^T |h_k(x) - h(x)| dx \rightarrow 0,$$

da cui

$$\|(f_k * g_k) * h_k - (f * g) * h\|_\infty \rightarrow 0.$$

Analogamente,

$$\|g_k * h_k - g * h\|_\infty \rightarrow 0, \quad \int_0^T |f_k(x) - f(x)| dx \rightarrow 0,$$

da cui

$$\|f_k * (g_k * h_k) - f * (g * h)\|_\infty \rightarrow 0.$$

Poiché

$$(f_k * g_k) * h_k = f_k * (g_k * h_k),$$

passando al limite per $k \rightarrow \infty$ si ottiene (ii). \square

Definizione 22.4. (*Famiglia di buoni nuclei*). Sia $K_n : \mathbb{R} \rightarrow \mathbb{C}$, $n \geq 1$, una successione di funzioni periodiche di periodo $T > 0$. Diciamo che la successione (K_n) è una *famiglia di buoni nuclei* di periodo T se

(i) per ogni $n \geq 1$ si ha

$$\frac{1}{T} \int_{-T/2}^{T/2} K_n(x) dx = 1;$$

(ii) esiste $M > 0$ tale che, per ogni $n \geq 1$,

$$\int_{-T/2}^{T/2} |K_n(x)| dx \leq M;$$

(iii) per ogni $\delta \in (0, T/2]$,

$$\lim_{n \rightarrow \infty} \left(\int_{\delta \leq |x| \leq T/2} |K_n(x)| dx \right) = 0.$$

Nel prossimo teorema usiamo i nuclei $K_n(x)$ come fattori di una convoluzione $f * K_n$; possiamo quindi interpretarli come funzioni che assegnano i pesi ai vari punti dell'intervallo $[-T/2, T/2]$ per calcolare la media pesata di f . La proprietà (iii) dice che, al crescere di n , la gran parte del peso si concentra nei punti intorno all'origine.

Teorema 22.5. (*Teorema dei buoni nuclei*). Sia (K_n) una famiglia di buoni nuclei, di periodo $T > 0$, e sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo T , integrabile nel periodo. Allora in ogni punto $x \in \mathbb{R}$ in cui f è continua si ha

$$\lim_{n \rightarrow \infty} (f * K_n)(x) = f(x).$$

Se f è continua in tutto \mathbb{R} , allora il limite $(f * K_n) \rightarrow f$ è uniforme in \mathbb{R} .

Dimostrazione. Per ogni $x \in \mathbb{R}$, ricordando che la convoluzione è commutativa, si ha

$$(f * K_n)(x) = \frac{1}{T} \int_{-T/2}^{T/2} f(x-y) K_n(y) dy,$$

mentre, dalla proprietà (i) dei buoni nuclei,

$$f(x) = f(x) 1 = f(x) \frac{1}{T} \int_{-T/2}^{T/2} K_n(y) dy = \frac{1}{T} \int_{-T/2}^{T/2} f(x) K_n(y) dy.$$

Dunque la differenza $(f * K_n)(x) - f(x)$ è

$$\begin{aligned}(f * K_n)(x) - f(x) &= \frac{1}{T} \int_{-T/2}^{T/2} f(x-y) K_n(y) dy - \frac{1}{T} \int_{-T/2}^{T/2} f(x) K_n(y) dy \\ &= \frac{1}{T} \int_{-T/2}^{T/2} (f(x-y) - f(x)) K_n(y) dy,\end{aligned}$$

e

$$|(f * K_n)(x) - f(x)| \leq \frac{1}{T} \int_{-T/2}^{T/2} |f(x-y) - f(x)| |K_n(y)| dy. \quad (22.1)$$

Sia $\varepsilon > 0$, e supponiamo che f sia continua in x . Allora, dalla definizione di continuità, esiste $\delta \in (0, T/2]$ tale che

$$|f(x-y) - f(x)| \leq \varepsilon_1 := \frac{T}{M} \frac{\varepsilon}{2} \quad \forall |y| \leq \delta. \quad (22.2)$$

Separiamo dunque l'integrale della (22.1) come

$$\int_{-T/2}^{T/2} = \int_{|y| \leq \delta} + \int_{\delta < |y| \leq T/2},$$

cioè

$$|(f * K_n)(x) - f(x)| \leq A_n(x) + B_n(x)$$

dove

$$\begin{aligned}A_n(x) &:= \frac{1}{T} \int_{|y| \leq \delta} |f(x-y) - f(x)| |K_n(y)| dy, \\ B_n(x) &:= \frac{1}{T} \int_{\delta < |y| \leq T/2} |f(x-y) - f(x)| |K_n(y)| dy,\end{aligned}$$

e stimiamo $A_n(x), B_n(x)$ separatamente.

Usando (22.2) e poi la proprietà (ii) dei buoni nuclei si ha

$$\begin{aligned}A_n(x) &\leq \frac{1}{T} \int_{|y| \leq \delta} \varepsilon_1 |K_n(y)| dy \\ &\leq \frac{\varepsilon_1}{T} \int_{-T/2}^{T/2} |K_n(y)| dy \\ &\leq \frac{\varepsilon_1}{T} M = \frac{\varepsilon}{2} \quad \forall n \in \mathbb{N}.\end{aligned}$$

Usando la disuguaglianza

$$\begin{aligned}|f(x-y) - f(y)| &\leq |f(x-y)| + |f(y)| \\ &\leq \|f\|_\infty + \|f\|_\infty = 2\|f\|_\infty\end{aligned}$$

e la proprietà (iii) dei buoni nuclei si ha

$$\begin{aligned}B_n(x) &\leq \frac{1}{T} \int_{\delta < |y| \leq T/2} 2\|f\|_\infty |K_n(y)| dy \\ &= \frac{2\|f\|_\infty}{T} \int_{\delta < |y| \leq T/2} |K_n(y)| dy \rightarrow 0 \quad \text{per } n \rightarrow \infty.\end{aligned}$$

Quindi esiste $n_0 \in \mathbb{N}$ tale che

$$B_n(x) \leq \frac{\varepsilon}{2} \quad \forall n \geq n_0,$$

da cui

$$|(f * K_n)(x) - f(x)| \leq A_n(x) + B_n(x) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \forall n \geq n_0.$$

Abbiamo provato che per ogni $\varepsilon > 0$ esiste $n_0 \in \mathbb{N}$ tale che

$$|(f * K_n)(x) - f(x)| \leq \varepsilon \quad \forall n \geq n_0,$$

cioè $(f * K_n)(x) \rightarrow f(x)$ per $n \rightarrow \infty$, per ogni punto x in cui f è continua.

Sia ora f continua in \mathbb{R} . Allora f è uniformemente continua in ogni intervallo compatto, e quindi, essendo periodica, f è uniformemente continua in \mathbb{R} .

Sia $\varepsilon > 0$. Per definizione di continuità uniforme, esiste $\delta \in (0, T/2]$ tale che

$$|f(z) - f(x)| \leq \varepsilon_1 := \frac{T}{M} \frac{\varepsilon}{2} \quad \forall x, z \in \mathbb{R}, \quad |x - z| \leq \delta,$$

cioè tale che

$$|f(x - y) - f(x)| \leq \varepsilon_1 \quad \forall x \in \mathbb{R}, \quad \forall |y| \leq \delta. \quad (22.3)$$

Procedendo come sopra, si trova

$$\begin{aligned} A_n(x) &\leq \frac{\varepsilon}{2} \quad \forall n \in \mathbb{N}, \quad \forall x \in \mathbb{R}, \\ B_n(x) &\leq \frac{\varepsilon}{2} \quad \forall n \geq n_0, \quad \forall x \in \mathbb{R}, \end{aligned}$$

dove n_0 non dipende da x perché δ non dipende da x e quindi la successione infinitesima

$$\frac{2\|f\|_\infty}{T} \int_{\delta < |y| \leq T/2} |K_n(y)| dy$$

non dipende da x . Dunque

$$|(f * K_n)(x) - f(x)| \leq \varepsilon \quad \forall n \geq n_0, \quad \forall x \in \mathbb{R}.$$

Questo prova che per ogni $\varepsilon > 0$ esiste $n_0 \in \mathbb{N}$ tale che

$$\|f * K_n - f\|_\infty = \sup_{x \in \mathbb{R}} |(f * K_n)(x) - f(x)| \leq \varepsilon \quad \forall n \geq n_0,$$

cioè che $f * K_n \rightarrow f$ uniformemente in \mathbb{R} per $n \rightarrow \infty$. \square

La successione (D_N) dei nuclei di Dirichlet *non* è una famiglia di buoni nuclei, perché non soddisfa la proprietà (ii): si ha infatti

$$\int_0^\pi |D_N(\vartheta)| d\vartheta \geq \log N \quad \forall N \geq 1. \quad (22.4)$$

Esercizio 22.6. (*I nuclei di Dirichlet non sono buoni nuclei*). Dimostrare la (22.4).

Suggerimento: poiché

$$0 < \sin \alpha < \alpha \quad \forall \alpha \in (0, \pi),$$

si ha

$$\frac{1}{\sin(\frac{1}{2}\vartheta)} > \frac{1}{\frac{1}{2}\vartheta} = \frac{2}{\vartheta} \quad \forall \vartheta \in (0, \pi).$$

Usare questa disuguaglianza e il cambio di variabile $(N + \frac{1}{2})\vartheta = x$.

Nonostante i nuclei di Dirichlet (D_N) non siano buoni nuclei, con un altro procedimento di media, detto *media alla Cesàro*, possiamo ottenere una famiglia di buoni nuclei a partire dai nuclei (D_N). Partiamo dalla definizione generale di media alla Cesàro.

Definizione 22.7. (*Medie alla Cesàro, convergenza alla Cesàro*). Sia (c_0, c_1, \dots) una successione di numeri complessi. Consideriamo la serie di termine generale c_n e le sue somme parziali

$$s_n = \sum_{k=0}^n c_k.$$

La media aritmetica delle prime N somme parziali della serie, cioè

$$\sigma_N := \frac{s_0 + s_1 + \dots + s_{N-1}}{N},$$

si dice *N -esima media alla Cesàro della successione (s_n)* oppure *N -esima somma alla Cesàro della serie $\sum_{n=0}^{\infty} c_n$* .

Se σ_N converge a un limite finito $\ell \in \mathbb{C}$ per $N \rightarrow \infty$, diciamo che la serie $\sum_{n=0}^{\infty} c_n$ è *Cesàro-sommabile ad ℓ* oppure *converge ad ℓ alla Cesàro*.

Esercizio 22.8. Dimostrare che la serie

$$\sum_{n=0}^{\infty} (-1)^n$$

converge alla Cesàro a $1/2$.

Esercizio 22.9. Dimostrare che se una serie numerica converge (in senso classico) a un numero $c \in \mathbb{C}$, allora converge alla Cesàro a c .

Esercizio 22.10. Siano c_n, s_n, σ_N come sopra. Calcolare i coefficienti $w_{k,N}$ tali che

$$\sigma_N = \sum_{k=0}^{N-1} w_{k,N} c_k \quad \forall N \geq 1.$$

Calcoliamo le medie alla Cesàro dei nuclei di Dirichlet.

Definizione 22.11. (*Nuclei di Fejer*). Si dice *N -esimo nucleo di Fejer* l' N -esima media alla Cesàro dei nuclei di Dirichlet

$$F_N(x) := \frac{D_0(x) + D_1(x) + \dots + D_{N-1}(x)}{N}.$$

A differenza dei nuclei di Dirichlet, quelli di Fejer formano una famiglia di buoni nuclei (Lemma 22.12). La differenza principale rispetto ai nuclei di Dirichlet è il fatto che

$$F_N(x) \geq 0 \quad \forall x \in \mathbb{R}$$

(vedi la formula (ii) nel Lemma 22.12). Questo rende la proprietà (ii) dei buoni nuclei una conseguenza immediata della proprietà (i) dei buoni nuclei.

Lemma 22.12. (Proprietà dei nuclei di Fejer). *I nuclei di Fejer soddisfano le seguenti proprietà:*

(i) F_N è reale, 2π -periodica e pari;

(ii) per ogni $N \geq 1$ si ha

$$F_N(x) = \frac{1}{N} \frac{\sin^2(\frac{1}{2}Nx)}{\sin^2(\frac{1}{2}x)} \quad \forall x \in \mathbb{R}, x \neq 2k\pi, k \in \mathbb{Z};$$

(iii) $F_N(0) = N$ per ogni $N \geq 1$;

(iv) (F_N) è una famiglia di buoni nuclei, di periodo 2π .

Dimostrazione. (i) I nuclei di Dirichlet sono reali, 2π -periodici, pari (vedi Lemma 20.3), quindi anche la loro media aritmetica è reale, 2π -periodica, pari.

(ii) Dalla formula per i nuclei di Dirichlet si ha

$$D_0(x) + \dots + D_{N-1}(x) = \sum_{k=0}^{N-1} \frac{\sin((k + \frac{1}{2})x)}{\sin(\frac{1}{2}x)};$$

moltiplichiamo e dividiamo per $2 \sin(\frac{1}{2}x)$,

$$= \frac{1}{2 \sin^2(\frac{1}{2}x)} \sum_{k=0}^{N-1} 2 \sin(\frac{1}{2}x) \sin((k + \frac{1}{2})x);$$

usiamo l'identità trigonometrica $2 \sin \alpha \sin \beta = \cos(\alpha - \beta) - \cos(\alpha + \beta)$ e otteniamo

$$= \frac{1}{2 \sin^2(\frac{1}{2}x)} \sum_{k=0}^{N-1} [\cos(kx) - \cos((k+1)x)];$$

questa è una somma telescopica, e dopo le cancellazioni rimane

$$= \frac{1}{2 \sin^2(\frac{1}{2}x)} (1 - \cos(Nx));$$

ora usiamo l'identità trigonometrica $\sin^2 \alpha = \frac{1}{2}(1 - \cos(2\alpha))$, e otteniamo

$$= \frac{2 \sin^2(\frac{1}{2}Nx)}{2 \sin^2(\frac{1}{2}x)}.$$

(iii) La funzione F_N è continua perché i nuclei di Dirichlet D_0, \dots, D_{N-1} sono funzioni continue (sono somme finite di esponenziali complessi, cioè di seni e coseni). Quindi $F_N(0) = \lim_{x \rightarrow 0} F_N(x)$, e possiamo usare la formula (i) per calcolare il limite.

(iii) Dalla Definizione 22.11 di F_N e dalla proprietà di media integrale

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} D_k(x) dx = 1 \quad \forall k \in \mathbb{N}$$

dei nuclei di Dirichlet (vedi Lemma 20.3), si ha

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} F_N(x) dx &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{N} \sum_{k=0}^{N-1} D_k(x) dx \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} D_k(x) dx \right) = \frac{1}{N} \sum_{k=0}^{N-1} 1 = 1. \end{aligned}$$

Dal fatto (fondamentale) che $F_N(x) \geq 0$ e dal calcolo della media integrale appena fatto segue che

$$\int_{-\pi}^{\pi} |F_N(x)| dx = \int_{-\pi}^{\pi} F_N(x) dx = 2\pi \quad \forall N \geq 1.$$

Infine, dato $0 < \delta \leq \pi$, si ha

$$\frac{1}{\sin^2(\frac{1}{2}x)} \leq \frac{1}{\sin^2(\frac{1}{2}\delta)} \quad \forall x \in [-\pi, -\delta] \cup [\delta, \pi],$$

quindi

$$\begin{aligned} \int_{\delta < |x| \leq \pi} |F_N(x)| dx &= \int_{\delta < |x| \leq \pi} \frac{1}{N} \frac{\sin^2(\frac{1}{2}Nx)}{\sin^2(\frac{1}{2}x)} dx \\ &\leq \int_{\delta < |x| \leq \pi} \frac{\sin^2(\frac{1}{2}Nx)}{N \sin^2(\frac{1}{2}\delta)} dx \\ &\leq \frac{1}{N \sin^2(\frac{1}{2}\delta)} \int_{-\pi}^{\pi} 1 dx = \frac{C}{N} \rightarrow 0 \quad \text{per } N \rightarrow \infty. \end{aligned}$$

I nuclei di Fejer (F_N) soddisfano quindi tutte e tre le condizioni della Definizione 22.4 di famiglia di buoni nuclei. \square

23 Lezione 23

Esercizio 23.1. (*Coefficienti di Fourier dei nuclei di Dirichlet e di Fejer*). Per ogni N , calcolare i coefficienti di Fourier

$$\widehat{D_N}(n), \quad \widehat{F_N}(n) \quad \forall n \in \mathbb{Z}$$

dei nuclei di Dirichlet e di Fejer.

I nuclei di Dirichlet D_N e quelli di Fejer F_N sono funzioni 2π -periodiche. Per $T > 0$ definiamo

$$\begin{aligned} D_{N,T}(x) &:= D_N(\omega x), \\ F_{N,T}(x) &:= F_N(\omega x), \quad \omega := 2\pi/T. \end{aligned} \quad (23.1)$$

Lemma 23.2. (Nuclei di Fejer di periodo T). *Sia $T > 0$, e sia $F_{N,T}$ definito in (23.1). Allora $(F_{N,T})$, $N \geq 1$, è una famiglia di buoni nuclei di periodo T .*

Dimostrazione. Semplice conseguenza del Lemma 22.12 e del cambio di variabile $\omega x = t$ (fare i dettagli per esercizio). \square

Come corollario di quanto visto per i buoni nuclei e per i nuclei di Fejer abbiamo il seguente risultato.

Teorema 23.3. (Sommabilità alla Cesàro per le serie di Fourier). *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione periodica di periodo $T > 0$, integrabile sul periodo. Allora*

$$(f * F_{N,T})(x) \rightarrow f(x) \quad \text{per } N \rightarrow \infty$$

in ogni punto x in cui f è continua, e se f è continua in \mathbb{R} allora la convergenza è uniforme in \mathbb{R} .

Si ha inoltre

$$(f * F_{N,T})(x) = \frac{1}{N} \sum_{k=0}^{N-1} (f * D_{k,T})(x) = \frac{1}{N} \sum_{k=0}^{N-1} S_k(f)(x),$$

*cioè $f * F_{N,T}$ è la media aritmetica dei primi N termini $S_0(f), \dots, S_{N-1}(f)$ della successione delle somme parziali della serie di Fourier di f , cioè l' N -esima media di Cesàro della successione $(S_k(f))$, cioè l' N -esima somma di Cesàro della serie di Fourier di f .*

Possiamo dunque riformulare l'enunciato di convergenza con queste parole: la serie di Fourier di f è Cesàro-sommabile a f in ogni punto in cui f è continua, e se f è continua in \mathbb{R} allora la convergenza è uniforme.

Dimostrazione. Dai Lemmi 22.12 e 23.2, $(F_{N,T})$ è una famiglia di buoni nuclei, quindi possiamo applicare il Teorema 22.5 dei buoni nuclei; otteniamo la convergenza puntuale $(f * F_{N,T})(x) \rightarrow f(x)$ nei punti di continuità di f e la convergenza uniforme se f è continua.

L'identità

$$(f * F_{N,T})(x) = \frac{1}{N} \sum_{k=0}^{N-1} (f * D_{k,T})(x)$$

segue dalla Definizione 22.11 dei nuclei di Fejer come media aritmetica dei nuclei di Dirichlet e dalla (23.1). Si ha poi

$$\frac{1}{N} \sum_{k=0}^{N-1} (f * D_{k,T})(x) = \frac{1}{N} \sum_{k=0}^{N-1} S_k(f)(x)$$

perché

$$\begin{aligned} S_k(f)(x) &= \frac{1}{T} \int_{-T/2}^{T/2} f(y) D_k(\omega(x-y)) dy \\ &= \frac{1}{T} \int_{-T/2}^{T/2} f(y) D_{k,T}(x-y) dy \\ &= (f * D_{k,T})(x) \end{aligned}$$

dal Lemma 20.2, dalla definizione (23.1) di $D_{k,T}$ e dalla Definizione 22.1 di convoluzione T -periodica. \square

Definizione 23.4. (*Somma N -esima di Cesàro di una serie di Fourier*). Indichiamo

$$\sigma_N(f)(x) := \frac{1}{N} \sum_{k=0}^{N-1} S_k(f)(x)$$

la media aritmetica delle prime N somme parziali $S_0(f), \dots, S_{N-1}(f)$ della serie di Fourier di f , cioè l' N -esima media di Cesàro della successione $(S_k(f))$, cioè l' N -esima somma di Cesàro della serie di Fourier $S(f)$ di f .

Dal Teorema 23.3,

$$\sigma_N(f) = f * F_{N,T},$$

e

$$\sigma_N(f) \rightarrow f$$

puntualmente dove f è continua, uniformemente se f è continua in \mathbb{R} .

Dalla Definizione 22.1, la somma parziale della serie di Fourier di f è il polinomio trigonometrico

$$S_N(f)(x) := \sum_{|n| \leq N} \hat{f}(n) e^{in\omega x}. \quad (23.2)$$

Anche $\sigma_N(f)(x)$ è un polinomio trigonometrico; nel prossimo lemma ne calcoliamo i coefficienti.

Lemma 23.5. (*Coefficienti di Fourier della somma di Cesàro*). *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ periodica di periodo $T > 0$, integrabile sul periodo. Allora la somma di Cesàro $\sigma_N(f)$ (vedi Definizione 23.4) è il polinomio trigonometrico*

$$\sigma_N(f)(x) = (f * F_{N,T})(x) = \sum_{|n| \leq N-1} \left(1 - \frac{|n|}{N}\right) \hat{f}(n) e^{in\omega x}.$$

Dimostrazione. Usando la formula (23.2) (cioè la definizione di $S_k(f)(x)$), si ha

$$\begin{aligned}
\sigma_N(f)(x) &= \frac{1}{N} \sum_{k=0}^{N-1} S_k(f)(x) \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \left(\sum_{|n| \leq k} \hat{f}(n) e^{in\omega x} \right) \\
&= \frac{1}{N} \sum_{\substack{k, n \in \mathbb{Z} \\ 0 \leq |n| \leq k \leq N-1}} \hat{f}(n) e^{in\omega x} \\
&= \frac{1}{N} \sum_{|n| \leq N-1} \left(\sum_{k=|n|}^{N-1} \hat{f}(n) e^{in\omega x} \right) \\
&= \frac{1}{N} \sum_{|n| \leq N-1} (N - |n|) \hat{f}(n) e^{in\omega x} \\
&= \sum_{|n| \leq N-1} \left(1 - \frac{|n|}{N} \right) \hat{f}(n) e^{in\omega x}.
\end{aligned}$$

Abbiamo usato il fatto che gli interi k nell'intervallo $|n| \leq k \leq N-1$ sono $N - |n|$. \square

Osservazione 23.6. (*Convoluzione con i nuclei e moltiplicatori di Fourier*). Dalla Definizione 22.1 e dal Lemma 23.5 si ha

$$S_N(f)(x) = \sum_{|n| \leq N} \hat{f}(n) e^{in\omega x}, \quad \sigma_N(f)(x) = \sum_{|n| \leq N-1} \left(1 - \frac{|n|}{N} \right) \hat{f}(n) e^{in\omega x}.$$

L'operazione che alla serie di Fourier

$$S(f)(x) = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{in\omega x}$$

associa la sua somma parziale $S_N(f)(x)$ corrisponde alla moltiplicazione del coefficiente di Fourier $\hat{f}(n)$ per un fattore,

$$\hat{f}(n) \mapsto w(n) \hat{f}(n), \quad w(n) = \begin{cases} 1 & \text{se } |n| \leq N, \\ 0 & \text{se } |n| > N. \end{cases}$$

Analogamente, l'operatore $S(f) \mapsto \sigma_N(f)$ corrisponde alla moltiplicazione dei coefficienti di Fourier

$$\hat{f}(n) \mapsto v(n) \hat{f}(n), \quad v(n) = \begin{cases} 1 - |n|/N & \text{se } |n| \leq N, \\ 0 & \text{se } |n| > N. \end{cases}$$

Si arriva alla stessa conclusione se si considera che

$$S_N(f) = f * D_{N,T}, \quad \sigma_N(f) = f * F_{N,T},$$

si calcolano i coefficienti di Fourier dei nuclei $D_{N,T}$, $F_{N,T}$, e si utilizza la proprietà della convoluzione

$$(f * \widehat{D_{N,T}})(n) = \hat{f}(n) \widehat{D_{N,T}}(n), \quad (f * \widehat{F_{N,T}})(n) = \hat{f}(n) \widehat{F_{N,T}}(n).$$

Gli operatori lineari che mappano una serie di Fourier di coefficienti $\hat{f}(n)$ nella serie di Fourier di coefficienti $\lambda(n)\hat{f}(n)$,

$$\sum_{n \in \mathbb{Z}} \hat{f}(n)e^{in\omega x} \mapsto \sum_{n \in \mathbb{Z}} \lambda(n)\hat{f}(n)e^{in\omega x}$$

per certi fattori moltiplicativi $\lambda(n) \in \mathbb{C}$, sono detti *moltiplicatori di Fourier*.

Per esempio, l'operatore di derivazione $f(x) \mapsto f'(x)$ corrisponde al moltiplicatore di Fourier $\hat{f}(n) \mapsto \lambda(n)\hat{f}(n)$ con $\lambda(n) = in\omega$.

Vediamo una conseguenza del Teorema 23.3.

Lemma 23.7. (Funzioni a coefficienti di Fourier tutti nulli). *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ periodica di periodo $T > 0$, integrabile sul periodo, e supponiamo che tutti i suoi coefficienti di Fourier $\hat{f}(n)$ siano nulli.*

Allora $f(x) = 0$ in ogni punto x in cui f è continua.

Dimostrazione. Poiché $\hat{f}(n) = 0$ per ogni n , le somme parziali $S_N(f)$ sono tutte identicamente nulle, e quindi anche le loro medie aritmetiche $\sigma_N(f)$ sono identicamente nulle. Nei punti x in cui f è continua si ha allora

$$f(x) = \lim_{N \rightarrow \infty} \sigma_N(f)(x) = 0. \quad \square$$

Il prossimo risultato, anch'esso conseguenza quasi immediata del Teorema 23.3, dice che le funzioni periodiche continue sono uniformemente approssimate dai polinomi trigonometrici.

Proposizione 23.8. (Densità dei polinomi trigonometrici nelle funzioni periodiche continue). *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione T -periodica e continua. Allora per ogni $\varepsilon > 0$ esiste un polinomio trigonometrico $p(x)$, T -periodico, tale che*

$$\|f - p\|_\infty = \sup_{x \in \mathbb{R}} |f(x) - p(x)| \leq \varepsilon.$$

Dimostrazione. La funzione f è continua, quindi (Teorema 23.3) f è il limite uniforme della successione $\sigma_N(f)$ delle somme di Cesàro, che sono polinomi trigonometrici. □

Esercizio 23.9. (*Teorema di Weierstrass di approssimazione polinomiale*). Ricordiamo l'enunciato del Teorema di Weierstrass di approssimazione polinomiale:

Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione continua sull'intervallo reale compatto $[a, b]$. Allora per ogni $\varepsilon > 0$ esiste un polinomio $p(x)$ tale che $|f(x) - p(x)| \leq \varepsilon$ per ogni $x \in [a, b]$.

Esercizio (facoltativo): dimostrare il Teorema di Weierstrass di approssimazione polinomiale deducendolo dalla Proposizione 23.8.

Suggerimento: data $f : [a, b] \rightarrow \mathbb{R}$ continua, definire un'estensione di f continua e periodica (in generale, $f(a)$ e $f(b)$ possono essere diversi, quindi il periodo T deve essere maggiore di $b - a$, per esempio $T = 2(b - a)$); applicare la Proposizione 23.8; utilizzare poi il fatto che l'esponenziale $e^{i\theta}$ ha serie di Taylor che converge uniformemente sui compatti, quindi considerare i suoi polinomi di Taylor in un periodo $[a, a + T]$.

Convergenza in media quadratica. Dopo i risultati sulla convergenza puntuale e uniforme, ora vediamo un altro tipo di convergenza per le serie di Fourier, la convergenza in media quadratica. Proveremo che la successione delle somme parziali $S_N(f)$ converge in media quadratica a f per $N \rightarrow \infty$, sotto la sola ipotesi che f sia periodica e integrabile sul periodo.

Cominciamo rivedendo i concetti di spazio vettoriale, prodotto scalare, norma (o seminorma) associata al prodotto scalare, ortogonalità.

Ricordiamo che si parla di “spazio vettoriale reale” o “spazio vettoriale su \mathbb{R} ” se gli scalari sono i numeri reali, e “spazio vettoriale complesso” o “spazio vettoriale su \mathbb{C} ” se gli scalari sono i numeri complessi. Ricordiamo anche la notazione usuale \bar{c} per indicare il coniugato del numero complesso $c \in \mathbb{C}$.

Definizione 23.10. (*Prodotto scalare semidefinito positivo per spazi vettoriali su \mathbb{C}*). Sia V uno spazio vettoriale su \mathbb{C} . Una mappa a valori complessi

$$V \times V \rightarrow \mathbb{C}, \quad (x, y) \mapsto \langle x, y \rangle$$

è un *prodotto scalare semidefinito positivo su V* se

- è hermitiano, cioè

$$\langle x, y \rangle = \overline{\langle y, x \rangle} \quad \forall x, y \in V;$$

- è lineare nella prima variabile e coniugato-lineare nella seconda variabile, cioè

$$\begin{aligned} \langle x + y, z \rangle &= \langle x, z \rangle + \langle y, z \rangle, & \langle cx, y \rangle &= c\langle x, y \rangle & \forall x, y, z \in V, \quad \forall c \in \mathbb{C}, \\ \langle x, y + z \rangle &= \langle x, y \rangle + \langle x, z \rangle, & \langle x, cy \rangle &= \bar{c}\langle x, y \rangle & \forall x, y, z \in V, \quad \forall c \in \mathbb{C} \end{aligned}$$

(lo scalare c , uscendo dalla prima componente, rimane c ; lo stesso scalare c , uscendo dalla seconda componente, diventa \bar{c} , complesso coniugato di c);

- è semidefinita positiva, cioè $\langle x, x \rangle$ è un numero reale non negativo,

$$\langle x, x \rangle \geq 0 \quad \forall x \in V.$$

La condizione di simmetria e quella di linearità nel primo argomento del prodotto scalare implicano la coniugato-linearità nel secondo argomento, che è dunque ridondante nella Definizione 23.10.

Definizione 23.11. (*Seminorma per spazi vettoriali su \mathbb{C}*). Sia V uno spazio vettoriale su \mathbb{C} , e sia $\langle \cdot, \cdot \rangle$ un prodotto scalare semidefinito positivo su V . La *seminorma* indotta dal prodotto scalare su V è la mappa

$$\| \cdot \| : V \rightarrow [0, \infty), \quad \|x\| := \sqrt{\langle x, x \rangle} \quad \forall x \in V.$$

Definizione 23.12. (*Prodotto scalare definito positivo e norma per spazi vettoriali su \mathbb{C}*). Sia V uno spazio vettoriale su \mathbb{C} , sia $\langle \cdot, \cdot \rangle$ un prodotto scalare semidefinito positivo su V , e sia $\| \cdot \|$ la seminorma indotta dal prodotto scalare. Se

$$\|x\| > 0 \quad \forall x \in V, \quad x \neq 0,$$

cioè se l'unico vettore di V di seminorma nulla è il vettore nullo, allora il prodotto scalare si dice *definito positivo*, e la seminorma si dice una *norma*.

Definizione 23.13. (*Spazi di Hilbert su \mathbb{C}*). Sia V uno spazio vettoriale su \mathbb{C} , sia $\langle \cdot, \cdot \rangle$ un prodotto scalare definito positivo su V , e sia $\| \cdot \|$ la norma indotta dal prodotto scalare. Se V è completo, cioè se ogni successione di Cauchy in V converge in V , allora diciamo che V è uno *spazio di Hilbert* complesso.

Esercizio 23.14. (*Lo spazio vettoriale \mathbb{C}^n*). L'insieme $V = \mathbb{C}^n$ è uno spazio vettoriale su \mathbb{C} . Dati due vettori

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in V, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in V, \quad x_k, y_k \in \mathbb{C},$$

consideriamo l'applicazione

$$\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$$

definita da

$$\langle x, y \rangle := x_1 \bar{y}_1 + \dots + x_n \bar{y}_n = \sum_{k=1}^n x_k \bar{y}_k \in \mathbb{C}.$$

Verificare che si tratta di un prodotto scalare complesso definito positivo, con norma indotta

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{|x_1|^2 + \dots + |x_n|^2} = \left(\sum_{k=1}^n |x_k|^2 \right)^{\frac{1}{2}}.$$

Esercizio 23.15. (*Prodotto scalare per funzioni periodiche*). Sia $T > 0$, e sia V l'insieme delle funzioni $f : \mathbb{R} \rightarrow \mathbb{C}$ periodiche di periodo T , integrabili sul periodo. Sia

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}, \quad \langle f, g \rangle := \frac{1}{T} \int_0^T f(x) \overline{g(x)} dx.$$

(i) Verificare che si tratta di un prodotto scalare semidefinito positivo, con seminorma indotta

$$\|f\| = \frac{1}{\sqrt{T}} \left(\int_0^T |f(x)|^2 dx \right)^{\frac{1}{2}}.$$

(ii) Verificare che il prodotto scalare non è definito positivo, e la seminorma non è una norma.

Definizione 23.16. (*Vettori ortogonali*). Sia V uno spazio vettoriale su \mathbb{C} con prodotto scalare semidefinito positivo. Due vettori $x, y \in V$ si dicono *ortogonali* se

$$\langle x, y \rangle = 0,$$

e in tal caso scriviamo $x \perp y$.

Nella proposizione che segue proviamo tre proprietà molto note e importanti senza assumere l'ipotesi che la seminorma sia una norma: prodotti scalari semidefiniti e seminorme sono sufficienti per la dimostrazione.

Proposizione 23.17. (*Proprietà degli spazi vettoriali su \mathbb{C} con prodotto scalare semidefinito positivo*). Sia V uno spazio vettoriale su \mathbb{C} con prodotto scalare semidefinito positivo. Allora si ha

- Teorema di Pitagora: se $x, y \in V$ sono ortogonali, $x \perp y$, allora

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

- Disuguaglianza di Cauchy-Schwarz: per ogni $x, y \in V$ si ha

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

- Disuguaglianza triangolare: per ogni $x, y \in V$ si ha

$$\|x + y\| \leq \|x\| + \|y\|.$$

Dimostrazione. (Teorema di Pitagora) Siano $x, y \in V$ due vettori ortogonali. Dalla definizione di $\| \cdot \|$, dalle proprietà del prodotto scalare e dall'ipotesi $\langle x, y \rangle = 0$ si ha

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \langle x, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2. \end{aligned}$$

(Disuguaglianza di Cauchy-Schwarz) Siano $x, y \in V$. Per ogni $\lambda \in \mathbb{C}$ si ha

$$\begin{aligned} 0 &\leq \|x + \lambda y\|^2 \\ &= \langle x + \lambda y, x + \lambda y \rangle \\ &= \langle x, x \rangle + \langle x, \lambda y \rangle + \langle \lambda y, x \rangle + \langle \lambda y, \lambda y \rangle \\ &= \|x\|^2 + \bar{\lambda} \langle x, y \rangle + \lambda \overline{\langle x, y \rangle} + |\lambda|^2 \|y\|^2. \end{aligned} \tag{23.3}$$

Scriviamo i numeri complessi $\lambda, \langle x, y \rangle$ tramite parte reale e parte immaginaria,

$$\langle x, y \rangle = a + ib, \quad \lambda = t + is, \quad a, b, t, s \in \mathbb{R}.$$

Si ha dunque

$$\begin{aligned} \bar{\lambda} \langle x, y \rangle + \lambda \overline{\langle x, y \rangle} &= (t - is)(a + ib) + (t + is)(a - ib) \\ &= 2(ta + sb), \end{aligned}$$

e la disuguaglianza (23.3) si riscrive

$$\|x\|^2 + 2(ta + sb) + (t^2 + s^2)\|y\|^2 \geq 0 \quad \forall (t, s) \in \mathbb{R}^2. \tag{23.4}$$

Se $\|y\| \neq 0$, l'espressione nella (23.4) è un polinomio di secondo grado in (t, s) , che chiamiamo

$$p(t, s) := \|x\|^2 + 2(ta + sb) + (t^2 + s^2)\|y\|^2.$$

Il gradiente di $p(t, s)$ si annulla per

$$t = -\frac{a}{\|y\|^2}, \quad s = -\frac{b}{\|y\|^2}$$

(punto di minimo di p). Sostituiamo questi valori di t, s nella (23.4) e otteniamo

$$\|x\|^2 \|y\|^2 \geq a^2 + b^2,$$

da cui

$$|\langle x, y \rangle| = |a + ib| = \sqrt{a^2 + b^2} \leq \|x\| \|y\|.$$

Se invece $\|y\| = 0$, la (23.4) diventa

$$\|x\|^2 + 2(ta + sb) \geq 0 \quad \forall (t, s) \in \mathbb{R}^2. \quad (23.5)$$

In particolare, per $s = 0$ troviamo

$$\|x\|^2 + 2at \geq 0 \quad \forall t \in \mathbb{R},$$

da cui segue che $a = 0$: se infatti fosse $a \neq 0$, allora scegliendo, per esempio,

$$t = -\frac{1 + \|x\|^2}{2a}$$

si avrebbe

$$\|x\|^2 + 2at = -1,$$

assurdo (d'altra parte, in un qualunque piano cartesiano, le uniche rette che non contengono punti di ordinata negativa sono rette orizzontali). Analogamente, per $t = 0$ la (23.5) diventa

$$\|x\|^2 + 2sb \geq 0 \quad \forall s \in \mathbb{R},$$

da cui segue che $b = 0$. Quindi

$$\langle x, y \rangle = a + ib = 0, \quad \|x\| \|y\| = 0,$$

e la disuguaglianza di Cauchy-Schwarz vale anche per $\|y\| = 0$.

(Disuguaglianza triangolare) Siano $x, y \in V$. Allora

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} + \|y\|^2. \end{aligned}$$

Come sopra, scriviamo $\langle x, y \rangle = a + ib$, con a, b reali. Allora

$$\langle x, y \rangle + \overline{\langle x, y \rangle} = (a + ib) + (a - ib) = 2a,$$

e

$$\langle x, y \rangle + \overline{\langle x, y \rangle} = 2a \leq 2|a| \leq 2\sqrt{a^2 + b^2} = 2|\langle x, y \rangle| \leq 2\|x\| \|y\|$$

per la disuguaglianza di Cauchy-Schwarz. Dunque

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2, \end{aligned}$$

da cui si ha la tesi. □

Osservazione 23.18. Dalla Proposizione 23.17, negli spazi vettoriali degli Esercizi 23.14, 23.15 valgono il Teorema di Pitagora, la disuguaglianza di Cauchy-Schwarz e la disuguaglianza triangolare.

In particolare, la disuguaglianza triangolare in \mathbb{C}^n scritta in componenti è

$$\left(\sum_{k=1}^n |x_k + y_k|^2\right)^{\frac{1}{2}} \leq \left(\sum_{k=1}^n |x_k|^2\right)^{\frac{1}{2}} + \left(\sum_{k=1}^n |y_k|^2\right)^{\frac{1}{2}} \quad \forall x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{C},$$

quella di Cauchy-Schwarz è

$$\left|\sum_{k=1}^n x_k \overline{y_k}\right| \leq \left(\sum_{k=1}^n |x_k|^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^n |y_k|^2\right)^{\frac{1}{2}} \quad \forall x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{C}.$$

Dopo gli spazi vettoriali degli Esercizi 23.14, 23.15, vediamo ora un altro esempio importante di spazio vettoriale con prodotto scalare.

Notazione. Indichiamo sia a_n che $a(n)$ l'elemento n -esimo di una successione a di numeri complessi.

Definizione 23.19. (*L'insieme $\ell^2(\mathbb{Z}, \mathbb{C})$*). Si indica $\ell^2(\mathbb{Z}, \mathbb{C})$ (e si legge “elle piccolo due di \mathbb{Z} in \mathbb{C} ”) l'insieme delle successioni $(a_n)_{n \in \mathbb{Z}}$ di numeri complessi tali che la serie

$$\sum_{n \in \mathbb{Z}} |a_n|^2 = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} |a_n|^2$$

sia convergente, cioè

$$\ell^2(\mathbb{Z}, \mathbb{C}) := \left\{ a : \mathbb{Z} \rightarrow \mathbb{C}, \sum_{n \in \mathbb{Z}} |a(n)|^2 < \infty \right\}.$$

Scriviamo anche ℓ^2 o $\ell^2(\mathbb{Z})$ al posto di $\ell^2(\mathbb{Z}, \mathbb{C})$.

Lemma 23.20. ($\ell^2(\mathbb{Z}, \mathbb{C})$ è uno spazio vettoriale). Per ogni $a, b \in \ell^2(\mathbb{Z}, \mathbb{C})$, la successione $a + b$ appartiene ad $\ell^2(\mathbb{Z}, \mathbb{C})$, e

$$\left(\sum_{n \in \mathbb{Z}} |a_n + b_n|^2\right)^{\frac{1}{2}} \leq \left(\sum_{n \in \mathbb{Z}} |a_n|^2\right)^{\frac{1}{2}} + \left(\sum_{n \in \mathbb{Z}} |b_n|^2\right)^{\frac{1}{2}}.$$

Di conseguenza $\ell^2(\mathbb{Z}, \mathbb{C})$ è uno spazio vettoriale su \mathbb{C} .

Dimostrazione. Siano $a, b \in \ell^2$. Per ogni $N \in \mathbb{N}$ si ha

$$\left(\sum_{|n| \leq N} |a_n + b_n|^2\right)^{\frac{1}{2}} \leq \left(\sum_{|n| \leq N} |a_n|^2\right)^{\frac{1}{2}} + \left(\sum_{|n| \leq N} |b_n|^2\right)^{\frac{1}{2}}$$

per la disuguaglianza triangolare nello spazio \mathbb{C}^q , $q = 2N + 1$ (vedi Osservazione 23.18). Per $N \rightarrow \infty$ tutte e tre le serie hanno limite, essendo serie a termine generale reale ≥ 0 (le loro somme parziali formano successioni crescenti). Dunque, passando al limite per $N \rightarrow \infty$, otteniamo la disuguaglianza dell'enunciato. Questo prova che $a + b \in \ell^2$.

Poi $\lambda a \in \ell^2$ per ogni $a \in \ell^2$, $\lambda \in \mathbb{C}$. Quindi ℓ^2 è spazio vettoriale su \mathbb{C} . \square

Lemma 23.21. (Convergenza della serie $\sum a_n \overline{b_n}$). Per ogni $a, b \in \ell^2(\mathbb{Z}, \mathbb{C})$, la serie di termine generale $a_n \overline{b_n}$ è assolutamente convergente, con

$$\sum_{n \in \mathbb{Z}} |a_n \overline{b_n}| \leq \left(\sum_{n \in \mathbb{Z}} |a_n|^2\right)^{\frac{1}{2}} \left(\sum_{n \in \mathbb{Z}} |b_n|^2\right)^{\frac{1}{2}}.$$

Di conseguenza la serie

$$\sum_{n \in \mathbb{Z}} a_n \overline{b_n} = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} a_n \overline{b_n}$$

converge in \mathbb{C} .

Dimostrazione. Siano $a, b \in \ell^2$. Per ogni $N \in \mathbb{N}$ si ha

$$\sum_{|n| \leq N} |a_n \overline{b_n}| = \sum_{|n| \leq N} |a_n| |b_n| \leq \left(\sum_{|n| \leq N} |a_n|^2 \right)^{\frac{1}{2}} \left(\sum_{|n| \leq N} |b_n|^2 \right)^{\frac{1}{2}}$$

per la disuguaglianza di Cauchy-Schwarz in \mathbb{R}^q , $q = 2N + 1$. Per $N \rightarrow \infty$ tutte e tre le serie hanno limite, essendo serie a termine generale reale ≥ 0 (le loro somme parziali formano successioni crescenti). Dunque, passando al limite per $N \rightarrow \infty$, otteniamo la disuguaglianza dell'enunciato.

La serie di termine generale $a_n \overline{b_n}$ converge assolutamente, quindi converge in \mathbb{C} (sulla convergenza delle serie vedi sezione 25). \square

Osserviamo anche che se

$$\sum_{n \in \mathbb{Z}} |a_n|^2 = 0,$$

allora, banalmente, $a_n = 0$ per ogni $n \in \mathbb{Z}$: se infatti fosse $a_m \neq 0$ per un certo $m \in \mathbb{Z}$, allora la somma della serie sarebbe maggiore o uguale a $|a_m|^2 > 0$.

Il Lemma 23.21 garantisce che la definizione seguente sia ben posta.

Definizione 23.22. (*Prodotto scalare in ℓ^2 e norma indotta*). Per $a, b \in \ell^2(\mathbb{Z}, \mathbb{C})$ definiamo

$$\langle a, b \rangle := \sum_{n \in \mathbb{Z}} a_n \overline{b_n}, \quad \|a\| := \sqrt{\langle a, a \rangle} = \left(\sum_{n \in \mathbb{Z}} |a_n|^2 \right)^{\frac{1}{2}},$$

prodotto scalare definito positivo e norma di $\ell^2(\mathbb{Z}, \mathbb{C})$.

Quando serve specificare che si tratta del prodotto scalare e della norma di $\ell^2(\mathbb{Z}, \mathbb{C})$ scriviamo anche

$$\langle a, b \rangle_{\ell^2}, \quad \|a\|_{\ell^2}.$$

Lemma 23.23. (ℓ^2 è spazio di Hilbert). *Lo spazio $\ell^2(\mathbb{Z}, \mathbb{C})$ è uno spazio di Hilbert.*

Dimostrazione. Sia a_k , $k = 1, 2, \dots$ una successione di Cauchy di elementi di ℓ^2 , cioè $a_k = (a_{k,n})_{n \in \mathbb{Z}}$ appartiene a ℓ^2 per ogni $k \geq 1$, e per ogni $\varepsilon > 0$ esiste $k_0 \in \mathbb{N}$ tale che

$$\|a_k - a_j\| = \left(\sum_{n \in \mathbb{Z}} |a_{k,n} - a_{j,n}|^2 \right)^{\frac{1}{2}} \leq \varepsilon \quad \forall k, j \geq k_0. \quad (23.6)$$

Per ogni $n \in \mathbb{Z}$ si ha

$$|a_{k,n} - a_{j,n}| \leq \|a_k - a_j\|,$$

quindi per ogni $n \in \mathbb{Z}$ fissato la successione $(a_{k,n})_{k=1,2,\dots}$ è di Cauchy in \mathbb{C} , e dunque converge in \mathbb{C} . Chiamiamo b_n il suo limite,

$$b_n = \lim_{k \rightarrow \infty} a_{k,n}, \quad n \in \mathbb{Z}.$$

Per ogni $N \in \mathbb{N}$, per ogni k, j si ha

$$\sum_{|n| \leq N} |a_{k,n} - a_{j,n}|^2 \leq \sum_{n \in \mathbb{Z}} |a_{k,n} - a_{j,n}|^2,$$

quindi, dalla (23.6),

$$\sum_{|n| \leq N} |a_{k,n} - a_{j,n}|^2 \leq \varepsilon^2 \quad \forall k, j \geq k_0. \quad (23.7)$$

Passiamo al limite per $k \rightarrow \infty$ nella (23.7) (senza alcun problema di convergenza, essendo una somma finita) e otteniamo

$$\sum_{|n| \leq N} |b_n - a_{j,n}|^2 \leq \varepsilon^2 \quad \forall j \geq k_0.$$

Per ogni $j \geq k_0$ fissato, la successione (s_N) , con

$$s_N := \sum_{|n| \leq N} |b_n - a_{j,n}|^2,$$

è una successione di numeri reali crescente e limitata, quindi ha limite, e il limite soddisfa

$$\lim_{N \rightarrow \infty} s_N = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} |b_n - a_{j,n}|^2 = \sum_{n \in \mathbb{Z}} |b_n - a_{j,n}|^2 \leq \varepsilon^2.$$

Questo prova che la successione $b - a_j = (b_n - a_{j,n})_{n \in \mathbb{Z}}$ appartiene a ℓ^2 , con norma

$$\|b - a_j\|^2 = \sum_{n \in \mathbb{Z}} |b_n - a_{j,n}|^2 \leq \varepsilon^2 \quad \forall j \geq k_0.$$

Di conseguenza $b \in \ell^2$ perché è somma di due elementi di ℓ^2 ,

$$b = (b - a_{k_0}) + a_{k_0}.$$

Inoltre, dato $\varepsilon > 0$ qualsiasi, abbiamo trovato $k_0 \in \mathbb{N}$ tale che $\|b - a_j\| \leq \varepsilon$ per ogni $j \geq k_0$, cioè $a_j \rightarrow b$ in ℓ^2 per $j \rightarrow \infty$.

Abbiamo dimostrato che ogni successione di Cauchy in ℓ^2 converge nella norma di ℓ^2 ad un elemento di ℓ^2 . \square

Osservazione 23.24. (*Lo spazio delle funzioni Riemann-integrabili non è completo*). Lo spazio vettoriale dell'Esercizio 23.15 non è completo; il suo completamento rispetto alla seminorma definita nell'esercizio (che diventa una vera norma considerando opportune classi di equivalenza di funzioni integrabili) porta alla definizione dello spazio di Lebesgue $L^2((0, T), \mathbb{C})$.

24 Lezione 24

Definizione 24.1. (*Insieme ortonormale*). Sia V uno spazio vettoriale su \mathbb{C} con prodotto scalare semidefinito positivo. Un sottoinsieme $A \subset V$ si dice *ortonormale* se

$$\|x\| = 1 \quad \forall x \in A, \quad \langle x, y \rangle = 0 \quad \forall x, y \in A, \quad x \neq y.$$

Riprendiamo lo spazio vettoriale dell'Esercizio 23.15, e mostriamo che gli esponenziali complessi che compaiono nelle serie di Fourier formano un insieme ortonormale.

Definizione 24.2. (*Spazio delle funzioni periodiche Riemann-integrabili*). Per ogni $T > 0$ indichiamo \mathcal{R}_T l'insieme delle funzioni $f : \mathbb{R} \rightarrow \mathbb{C}$ periodiche di periodo T , integrabili nel periodo,

$$\mathcal{R}_T := \{f : \mathbb{R} \rightarrow \mathbb{C} : f \text{ è } T\text{-periodica e integrabile in } [0, T]\}.$$

Come osservato nell'Esercizio 23.15, \mathcal{R}_T è uno spazio vettoriale su \mathbb{C} , munito di prodotto scalare semidefinito positivo

$$\langle f, g \rangle = \frac{1}{T} \int_0^T f(x) \overline{g(x)} dx$$

e seminorma indotta dal prodotto scalare

$$\|f\| = \frac{1}{\sqrt{T}} \left(\int_0^T |f(x)|^2 dx \right)^{\frac{1}{2}}.$$

Scriviamo anche

$$\langle f, g \rangle_{\mathcal{R}_T}, \quad \|f\|_{\mathcal{R}_T}$$

per indicare che si tratta del prodotto scalare e della seminorma dello spazio \mathcal{R}_T .

In tutta questa lezione consideriamo \mathcal{R}_T munito del prodotto scalare e della seminorma della Definizione 24.2.

Lemma 24.3. (*Famiglia ortonormale degli esponenziali*). Sia $T > 0$, $\omega = 2\pi/T$,

$$e_n(x) := e^{in\omega x}, \quad x \in \mathbb{R}, \quad n \in \mathbb{Z}.$$

Allora $e_n \in \mathcal{R}_T$ per ogni $n \in \mathbb{Z}$, e per ogni $n, k \in \mathbb{Z}$ si ha

$$\langle e_n, e_k \rangle_{\mathcal{R}_T} = \delta_{nk} = \begin{cases} 1 & \text{se } n = k, \\ 0 & \text{se } n \neq k, \end{cases}$$

cioè l'insieme $\{e_n : n \in \mathbb{Z}\} \subset \mathcal{R}_T$ è ortonormale.

Dimostrazione. Ogni esponenziale e_n è T -periodico, quindi appartiene a \mathcal{R}_T . Per ogni $n, k \in \mathbb{Z}$, $n \neq k$ si ha

$$\begin{aligned} \langle e_n, e_k \rangle &= \frac{1}{T} \int_0^T e_n(x) \overline{e_k(x)} dx \\ &= \frac{1}{T} \int_0^T e^{in\omega x} e^{-ik\omega x} dx \\ &= \frac{1}{T} \int_0^T e^{i(n-k)\omega x} dx = \frac{1}{T} \left[\frac{e^{i(n-k)\omega x}}{i(n-k)\omega} \right]_0^T = \frac{e^{i(n-k)2\pi} - e^0}{i(n-k)2\pi} = 0 \end{aligned}$$

perché $\omega T = 2\pi$, ed $e^{im2\pi} = 1$ per ogni m intero. Per $k = n$ si ha invece

$$\langle e_n, e_n \rangle = \frac{1}{T} \int_0^T e^{i(n-n)\omega x} dx = \frac{1}{T} \int_0^T 1 dx = 1. \quad \square$$

Lemma 24.4. (Seminorma dei polinomi trigonometrici). *Sia $T > 0$, $\omega = 2\pi/T$, e sia $p(x)$ il polinomio trigonometrico*

$$p(x) = \sum_{|n| \leq N} c_n e^{in\omega x},$$

con coefficienti $c_n \in \mathbb{C}$. La sua seminorma in \mathcal{R}_T è allora

$$\|p\|^2 = \sum_{|n| \leq N} |c_n|^2.$$

Dimostrazione. Poiché

$$p = \sum_{|n| \leq N} c_n e_n,$$

dalle proprietà del prodotto scalare si ha

$$\begin{aligned} \|p\|^2 &= \langle p, p \rangle = \left\langle \sum_{|n| \leq N} c_n e_n, \sum_{|k| \leq N} c_k e_k \right\rangle \\ &= \sum_{|n| \leq N} \left\langle c_n e_n, \sum_{|k| \leq N} c_k e_k \right\rangle \\ &= \sum_{|n| \leq N} \left(\sum_{|k| \leq N} \langle c_n e_n, c_k e_k \rangle \right) \\ &= \sum_{|n| \leq N} \left(\sum_{|k| \leq N} c_n \bar{c}_k \langle e_n, e_k \rangle \right) \\ &= \sum_{|n| \leq N} c_n \left(\sum_{|k| \leq N} \bar{c}_k \delta_{nk} \right) = \sum_{|n| \leq N} c_n \bar{c}_n = \sum_{|n| \leq N} |c_n|^2. \quad \square \end{aligned}$$

Lemma 24.5. (Lemma di ortogonalità). *Sia $T > 0$, $\omega = 2\pi/T$, $e_n(x) = e^{in\omega x}$. Sia $f \in \mathcal{R}_T$. Allora*

(i) (Coefficienti di Fourier) *I coefficienti di Fourier $\hat{f}(n)$ di f soddisfano*

$$\hat{f}(n) = \langle f, e_n \rangle \quad \forall n \in \mathbb{Z}.$$

(ii) (Somma parziale $S_N(f)$) *La somma parziale $S_N(f)$ della serie di Fourier di f soddisfa*

$$S_N(f) = \sum_{|n| \leq N} \langle f, e_n \rangle e_n.$$

(iii) (Ortogonalità) *Per ogni polinomio trigonometrico T -periodico di grado $\leq N$*

$$p(x) = \sum_{|n| \leq N} c_n e_n(x), \quad c_n \in \mathbb{C},$$

si ha

$$f - S_N(f) \perp p.$$

(iv) (Pitagora) Per ogni $N \in \mathbb{N}$ si ha

$$\|f\|^2 = \|f - S_N(f)\|^2 + \sum_{|n| \leq N} |\hat{f}(n)|^2.$$

Dimostrazione. (i) Dalla Definizione 19.19 dei coefficienti di Fourier $\hat{f}(n)$ e dalla Definizione 24.2 del prodotto scalare in \mathcal{R}_T si ha

$$\hat{f}(n) = \frac{1}{T} \int_0^T f(x) e^{-in\omega x} dx = \frac{1}{T} \int_0^T f(x) \overline{e_n(x)} dx = \langle f, e_n \rangle.$$

(ii) Dalla Definizione 19.19 di $S_N(f)(x)$ e da quanto appena osservato si ha

$$S_N(f)(x) = \sum_{|n| \leq N} \hat{f}(n) e^{in\omega x} = \sum_{|n| \leq N} \langle f, e_n \rangle e_n(x).$$

(iii) Sia

$$p = \sum_{|n| \leq N} c_n e_n, \quad c_n \in \mathbb{C}.$$

Dalle proprietà del prodotto scalare si ha

$$\langle f, p \rangle = \langle f, \sum_{|n| \leq N} c_n e_n \rangle = \sum_{|n| \leq N} \overline{c_n} \langle f, e_n \rangle = \sum_{|n| \leq N} \hat{f}(n) \overline{c_n}$$

e, ricordando che $\langle e_n, e_k \rangle = \delta_{nk}$,

$$\langle S_N(f), p \rangle = \left\langle \sum_{|n| \leq N} \hat{f}(n) e_n, \sum_{|k| \leq N} c_k e_k \right\rangle = \sum_{|n| \leq N} \sum_{|k| \leq N} \hat{f}(n) \overline{c_k} \langle e_n, e_k \rangle = \sum_{|n| \leq N} \hat{f}(n) \overline{c_n},$$

quindi

$$\langle f - S_N(f), p \rangle = \langle f, p \rangle - \langle S_N(f), p \rangle = 0.$$

(iv) La proprietà (iii) applicata al polinomio trigonometrico $p = S_N(f)$ dice che

$$f - S_N(f) \perp S_N(f).$$

Quindi f è la somma di due vettori ortogonali,

$$f = (f - S_N(f)) + S_N(f),$$

e dunque, dal Teorema di Pitagora (vedi Proposizione 23.17),

$$\|f\|^2 = \|(f - S_N(f)) + S_N(f)\|^2 = \|f - S_N(f)\|^2 + \|S_N(f)\|^2.$$

Infine

$$\|S_N(f)\|^2 = \sum_{|n| \leq N} |\hat{f}(n)|^2$$

dal Lemma 24.4. □

Lemma 24.6. (Lemma di migliore approssimazione). Sia $T > 0$, $\omega = 2\pi/T$, $e_n(x) = e^{in\omega x}$. Sia $f \in \mathcal{R}_T$. Allora per ogni $N \in \mathbb{N}$ si ha

$$\|f - S_N(f)\| \leq \|f - p\| \quad \forall p = \sum_{|n| \leq N} c_n e_n, \quad c_n \in \mathbb{C},$$

e

$$\|f - S_N(f)\| = \|f - p\| \quad \Leftrightarrow \quad p = S_N(f).$$

In altre parole, tra tutti i polinomi trigonometrici T -periodici di grado $\leq N$, $S_N(f)$ è quello che realizza la minore distanza da f nella norma di \mathcal{R}_T .

Dimostrazione. Sia

$$p = \sum_{|n| \leq N} c_n e_n, \quad c_n \in \mathbb{C}.$$

Scriviamo la differenza $f - p$ come

$$f - p = (f - S_N(f)) + (S_N(f) - p),$$

e osserviamo che $S_N(f) - p$ è un polinomio trigonometrico di grado $\leq N$,

$$S_N(f) - p = \sum_{|n| \leq N} \hat{f}(n) e_n - \sum_{|n| \leq N} c_n e_n = \sum_{|n| \leq N} (\hat{f}(n) - c_n) e_n.$$

Allora, dal punto (iii) del Lemma 24.5 di ortogonalità, si ha

$$(f - S_N(f)) \perp (S_N(f) - p).$$

Di conseguenza, applicando il Teorema di Pitagora (vedi Proposizione 23.17) e poi il Lemma 24.4,

$$\begin{aligned} \|f - p\|^2 &= \|(f - S_N(f)) + (S_N(f) - p)\|^2 \\ &= \|f - S_N(f)\|^2 + \|S_N(f) - p\|^2 \\ &= \|f - S_N(f)\|^2 + \sum_{|n| \leq N} |\hat{f}(n) - c_n|^2. \end{aligned}$$

La tesi segue osservando che la somma

$$\sum_{|n| \leq N} |\hat{f}(n) - c_n|^2$$

è ≥ 0 in ogni caso, ed è $= 0$ solo se $|\hat{f}(n) - c_n| = 0$ per ogni $|n| \leq N$, cioè se $p = S_N(f)$. \square

Teorema 24.7. (Teorema di convergenza in media quadratica per serie di Fourier). Sia $f \in \mathcal{R}_T$. Allora

$$\lim_{N \rightarrow \infty} \|f - S_N(f)\| = 0.$$

Dimostrazione. Sia $\varepsilon > 0$.

Passo 1. Dal Lemma 21.7 di densità delle funzioni continue in quelle integrabili, esiste una funzione g continua, T -periodica, tale che

$$\|g\|_\infty \leq \sqrt{2}\|f\|_\infty, \quad \int_0^T |f(x) - g(x)| dx \leq C\varepsilon^2,$$

dove

$$C = \frac{T}{4(1 + \sqrt{2})(1 + \|f\|_\infty)}.$$

Si ha dunque

$$\begin{aligned} \|f - g\|^2 &= \frac{1}{T} \int_0^T |f(x) - g(x)|^2 dx \\ &\leq \frac{1}{T} \int_0^T |f(x) - g(x)| (|f(x)| + |g(x)|) dx \\ &\leq \frac{1}{T} \int_0^T |f(x) - g(x)| (\|f\|_\infty + \|g\|_\infty) dx \\ &\leq \frac{(1 + \sqrt{2})\|f\|_\infty}{T} \int_0^T |f(x) - g(x)| dx \\ &\leq \frac{(1 + \sqrt{2})\|f\|_\infty}{T} C\varepsilon^2 \\ &\leq \frac{\varepsilon^2}{4}. \end{aligned}$$

Quindi

$$\|f - g\| \leq \frac{\varepsilon}{2}.$$

Passo 2. Poiché g è funzione continua, dalla Proposizione 23.8 di densità dei polinomi trigonometrici nelle funzioni continue esiste un polinomio trigonometrico T -periodico $p(x)$ tale che

$$\|g - p\|_\infty \leq \frac{\varepsilon}{2}.$$

Si ha quindi

$$\|g - p\|^2 = \frac{1}{T} \int_0^T |g(x) - p(x)|^2 dx \leq \frac{1}{T} \int_0^T \|g - p\|_\infty^2 dx \leq \frac{\varepsilon^2}{4},$$

da cui

$$\|g - p\| \leq \frac{\varepsilon}{2}.$$

Perciò

$$\|f - p\| = \|f - g + g - p\| \leq \|f - g\| + \|g - p\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Passo 3. Sia M il grado del polinomio p . Sia $N \geq M$. Dal Lemma 24.6 di migliore approssimazione, $S_N(f)$ è il polinomio trigonometrico che realizza la minore distanza da f tra tutti i polinomi di grado $\leq N$; tra essi c'è anche p (il cui grado è $M \leq N$). Quindi

$$\|f - S_N(f)\| \leq \|f - p\|.$$

Poiché $\|f - p\| \leq \varepsilon$, si ha

$$\|f - S_N(f)\| \leq \varepsilon \quad \forall N \geq M.$$

Abbiamo provato che per ogni $\varepsilon > 0$ esiste $M \in \mathbb{N}$ tale che $\|f - S_N(f)\| \leq \varepsilon$ per ogni $N \geq M$, cioè che $\|f - S_N(f)\| \rightarrow 0$ per $N \rightarrow \infty$. \square

Corollario 24.8. (Identità di Parseval). Sia $f \in \mathcal{R}_T$. Allora la successione $(\hat{f}(n))_{n \in \mathbb{Z}}$ dei suoi coefficienti di Fourier appartiene a $\ell^2(\mathbb{Z}, \mathbb{C})$, e si ha

$$\|f\|^2 = \sum_{n \in \mathbb{Z}} |\hat{f}(n)|^2.$$

Dimostrazione. Dalla (iv) del Lemma 24.5 di ortogonalità si ha

$$\sum_{|n| \leq N} |\hat{f}(n)|^2 = \|f\|^2 - \|f - S_N(f)\|^2 \quad (24.1)$$

e, dal Teorema 24.7 di convergenza in media quadratica,

$$\|f - S_N(f)\| \rightarrow 0 \quad \text{per } N \rightarrow \infty.$$

Quindi passando al limite per $N \rightarrow \infty$ nella (24.1) si ha

$$\sum_{n \in \mathbb{Z}} |\hat{f}(n)|^2 = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} |\hat{f}(n)|^2 = \|f\|^2. \quad \square$$

L'esercizio seguente è uno degli esercizi più classici e famosi sulle serie di Fourier.

Esercizio 24.9. (Somma della serie armonica generalizzata di esponente 2). Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ la funzione 2π -periodica che nell'intervallo $[-\pi, \pi)$ vale $f(x) = x$.

Calcolare i suoi coefficienti di Fourier $\hat{f}(n)$, $n \in \mathbb{Z}$; calcolare $\|f\|^2$; scrivere l'identità di Parseval per f e utilizzarla per dedurre l'uguaglianza

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Se al posto degli esponenziali $e_n(x) = e^{in\omega x}$ consideriamo una famiglia ortonormale qualunque, non abbiamo più l'identità di Parseval, ma vale comunque la seguente disuguaglianza.

Lemma 24.10. (Disuguaglianza di Bessel). Sia $\{v_n : n \in \mathbb{Z}\} \subset \mathcal{R}_T$ un insieme ortonormale, sia $f \in \mathcal{R}_T$, e

$$a_n := \langle f, v_n \rangle, \quad n \in \mathbb{Z}.$$

Allora

$$\sum_{n \in \mathbb{Z}} |a_n|^2 \leq \|f\|^2.$$

Dimostrazione. Sia

$$g_N := \sum_{|n| \leq N} a_n v_n.$$

Per ipotesi, $\langle v_n, v_k \rangle = \delta_{nk}$ per ogni $n, k \in \mathbb{Z}$. Dalle proprietà del prodotto scalare si ha

$$\begin{aligned} \langle f, g_N \rangle &= \left\langle f, \sum_{|n| \leq N} a_n v_n \right\rangle \\ &= \sum_{|n| \leq N} \langle f, a_n v_n \rangle \\ &= \sum_{|n| \leq N} \overline{a_n} \langle f, v_n \rangle \\ &= \sum_{|n| \leq N} \overline{a_n} a_n \\ &= \sum_{|n| \leq N} |a_n|^2 \end{aligned}$$

e

$$\begin{aligned} \langle g_N, g_N \rangle &= \left\langle \sum_{|n| \leq N} a_n v_n, \sum_{|k| \leq N} a_k v_k \right\rangle \\ &= \sum_{|n| \leq N} \left\langle a_n v_n, \sum_{|k| \leq N} a_k v_k \right\rangle \\ &= \sum_{|n| \leq N} a_n \left\langle v_n, \sum_{|k| \leq N} a_k v_k \right\rangle \\ &= \sum_{|n| \leq N} a_n \left(\sum_{|k| \leq N} \langle v_n, a_k v_k \rangle \right) \\ &= \sum_{|n| \leq N} a_n \left(\sum_{|k| \leq N} \overline{a_k} \langle v_n, v_k \rangle \right) \\ &= \sum_{|n| \leq N} a_n \overline{a_n} \\ &= \sum_{|n| \leq N} |a_n|^2. \end{aligned}$$

Quindi

$$\langle f - g_N, g_N \rangle = \langle f, g_N \rangle - \langle g_N, g_N \rangle = 0,$$

cioè

$$f - g_N \perp g_N.$$

Dunque dal Teorema di Pitagora (vedi Proposizione 23.17) si ha

$$\|f\|^2 = \|(f - g_N) + g_N\|^2 = \|f - g_N\|^2 + \|g_N\|^2 \geq \|g_N\|^2 = \sum_{|n| \leq N} |a_n|^2.$$

La disuguaglianza

$$\sum_{|n| \leq N} |a_n|^2 \leq \|f\|^2$$

vale per ogni $N \in \mathbb{N}$, e quindi, passando al limite per $N \rightarrow \infty$, si ha

$$\sum_{n \in \mathbb{Z}} |a_n|^2 = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} |a_n|^2 \leq \|f\|^2.$$

Notiamo che questo limite esiste perché si tratta di una serie reale a termine generale ≥ 0 . \square

Lemma 24.11. (Corrispondenza tra prodotto scalare in \mathcal{R}_T e in ℓ^2). *Siano $f, g \in \mathcal{R}_T$, e siano $\hat{f} = (\hat{f}(n))_{n \in \mathbb{Z}}$, $\hat{g} = (\hat{g}(n))_{n \in \mathbb{Z}}$ le successioni dei rispettivi coefficienti di Fourier. Allora*

$$\langle f, g \rangle_{\mathcal{R}_T} = \langle \hat{f}, \hat{g} \rangle_{\ell^2},$$

cioè

$$\frac{1}{T} \int_0^T f(x) \overline{g(x)} dx = \sum_{n \in \mathbb{Z}} \hat{f}(n) \overline{\hat{g}(n)}.$$

Dimostrazione. Abbiamo già provato (Corollario 24.8) che \hat{f}, \hat{g} appartengono a $\ell^2(\mathbb{Z}, \mathbb{C})$. Per ogni $N \in \mathbb{N}$ separiamo

$$\langle f, g \rangle = \langle f - S_N(f), g \rangle + \langle S_N(f), g \rangle. \quad (24.2)$$

Calcoliamo

$$\begin{aligned} \langle S_N(f), g \rangle &= \left\langle \sum_{|n| \leq N} \hat{f}(n) e_n, g \right\rangle \\ &= \sum_{|n| \leq N} \hat{f}(n) \langle e_n, g \rangle \\ &= \sum_{|n| \leq N} \hat{f}(n) \overline{\langle g, e_n \rangle} \\ &= \sum_{|n| \leq N} \hat{f}(n) \overline{\hat{g}(n)}, \end{aligned}$$

e ricordiamo che questa somma converge in \mathbb{C} per $N \rightarrow \infty$ (vedi Lemma 23.21), cioè

$$\lim_{N \rightarrow \infty} \langle S_N(f), g \rangle = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} \hat{f}(n) \overline{\hat{g}(n)} = \sum_{n \in \mathbb{Z}} \hat{f}(n) \overline{\hat{g}(n)} = \langle \hat{f}, \hat{g} \rangle_{\ell^2} \in \mathbb{C}.$$

Stimiamo l'altro termine della (24.2) con la disuguaglianza di Cauchy-Schwarz (vedi Proposizione 23.17), e poi con il Teorema 24.7 di convergenza in media quadratica, e abbiamo

$$|\langle f - S_N(f), g \rangle| \leq \|f - S_N(f)\| \|g\| \rightarrow 0 \quad \text{per } N \rightarrow \infty.$$

Passando al limite per $N \rightarrow \infty$ nella (24.2) otteniamo la tesi. \square

Abbiamo provato (proprietà (viii) del Lemma 22.2) che la mappa “cappuccio”

$$\mathcal{R}_T \rightarrow \ell^2(\mathbb{Z}, \mathbb{C}), \quad f \mapsto \hat{f}$$

che ad una funzione $f \in \mathcal{R}_T$ associa la successione \hat{f} dei suoi coefficienti di Fourier trasforma la convoluzione di funzioni periodiche nel prodotto (puntuale) di successioni,

$$\widehat{(f * g)}(n) = \hat{f}(n) \hat{g}(n).$$

Vediamo ora (Lemma 24.14) che vale anche il contrario: la mappa “cappuccio” trasforma il prodotto (puntuale) di funzioni periodiche nella convoluzione di successioni,

$$\widehat{(fg)}(n) = (\hat{f} * \hat{g})(n).$$

In sintesi: “cappuccio” trasforma prodotti in convoluzioni e convoluzioni in prodotti.

Definizione 24.12. (*Convoluzione di successioni*). Per ogni $a, b \in \ell^2(\mathbb{Z}, \mathbb{C})$, la convoluzione $a * b$ è la successione definita da

$$(a * b)(n) = \sum_{k \in \mathbb{Z}} a(k)b(n - k), \quad n \in \mathbb{Z}.$$

Lemma 24.13. (La successione $a * b$ è ben posta e limitata). Per ogni $a, b \in \ell^2(\mathbb{Z}, \mathbb{C})$, per ogni $n \in \mathbb{Z}$ si ha

$$\sum_{k \in \mathbb{Z}} |a(k)||b(n - k)| \leq \|a\|_{\ell^2} \|b\|_{\ell^2}.$$

La serie che definisce $(a * b)(n)$ è assolutamente convergente, dunque $(a * b)(n)$ è ben definita e soddisfa

$$|(a * b)(n)| \leq \|a\|_{\ell^2} \|b\|_{\ell^2} \quad \forall n \in \mathbb{Z}.$$

Dimostrazione. Dalla disuguaglianza di Cauchy-Schwarz, per ogni $n \in \mathbb{Z}$ si ha

$$\sum_{k \in \mathbb{Z}} |a(k)||b(n - k)| \leq \left(\sum_{k \in \mathbb{Z}} |a(k)|^2 \right)^{\frac{1}{2}} \left(\sum_{k \in \mathbb{Z}} |b(n - k)|^2 \right)^{\frac{1}{2}}.$$

Con il cambio di indice $n - k = j$ si ha

$$\left(\sum_{k \in \mathbb{Z}} |b(n - k)|^2 \right)^{\frac{1}{2}} = \left(\sum_{j \in \mathbb{Z}} |b(j)|^2 \right)^{\frac{1}{2}} = \|b\|_{\ell^2},$$

quindi

$$\sum_{k \in \mathbb{Z}} |a(k)||b(n - k)| \leq \|a\|_{\ell^2} \|b\|_{\ell^2}.$$

Perciò la serie in k di termine generale $a(k)b(n - k)$ è assolutamente convergente, e dunque è convergente in \mathbb{C} (sulla convergenza delle serie vedere anche l'Appendice 25). \square

Lemma 24.14. (Coefficienti di Fourier del prodotto di funzioni periodiche). Siano $f, g \in \mathcal{R}_T$. Allora i coefficienti di Fourier della funzione prodotto $f(x)g(x)$ sono

$$\widehat{(fg)}(n) = (\hat{f} * \hat{g})(n) = \sum_{k \in \mathbb{Z}} \hat{f}(k)\hat{g}(n - k), \quad n \in \mathbb{Z}.$$

Dimostrazione. Sia $n \in \mathbb{Z}$ fissato. Per ogni $N \in \mathbb{N}$, scrivendo

$$f = S_N(f) + (f - S_N(f))$$

si ha

$$\widehat{(fg)}(n) = \frac{1}{T} \int_0^T f(x)g(x)e^{-in\omega x} dx = A_N + B_N$$

dove

$$A_N := \frac{1}{T} \int_0^T S_N(f)(x)g(x)e^{-in\omega x} dx,$$

$$B_N := \frac{1}{T} \int_0^T [f(x) - S_N(f)(x)]g(x)e^{-in\omega x} dx.$$

Usando la definizione di $S_N(f)$ (vedi Definizione 19.19) calcoliamo

$$\begin{aligned} A_N &= \frac{1}{T} \int_0^T \left(\sum_{|k| \leq N} \hat{f}(k) e^{ik\omega x} \right) g(x) e^{-in\omega x} dx \\ &= \sum_{|k| \leq N} \hat{f}(k) \frac{1}{T} \int_0^T g(x) e^{-i(n-k)\omega x} dx \\ &= \sum_{|k| \leq N} \hat{f}(k) \hat{g}(n-k). \end{aligned}$$

Dal Lemma 24.13, la serie in k di termine generale $\hat{f}(k)\hat{g}(n-k)$ è assolutamente convergente, quindi convergente in \mathbb{C} , e di conseguenza

$$\lim_{N \rightarrow \infty} A_N = \lim_{N \rightarrow \infty} \sum_{|k| \leq N} \hat{f}(k)\hat{g}(n-k) = \sum_{k \in \mathbb{Z}} \hat{f}(k)\hat{g}(n-k) = (\hat{f} * \hat{g})(n).$$

Il termine B_N si può scrivere come prodotto scalare

$$B_N = \frac{1}{T} \int_0^T [f(x) - S_N(f)(x)] \overline{h(x)} dx = \langle f - S_N(f), h \rangle$$

dove h è la funzione

$$h(x) := \overline{g(x)} e^{in\omega x}.$$

Essendo prodotto di funzioni T -periodiche e integrabili sul periodo, h appartiene a \mathcal{R}_T ; inoltre

$$\|h\|_{\mathcal{R}_T}^2 = \frac{1}{T} \int_0^T |h(x)|^2 dx = \frac{1}{T} \int_0^T |g(x)|^2 dx = \|g\|_{\mathcal{R}_T}^2$$

perché $|\overline{g(x)}| = |g(x)|$ e $|e^{in\omega x}| = 1$. Usando la disuguaglianza di Cauchy-Schwarz si ha allora

$$|B_N| = |\langle f - S_N(f), h \rangle| \leq \|f - S_N(f)\|_{\mathcal{R}_T} \|h\|_{\mathcal{R}_T} = \|f - S_N(f)\|_{\mathcal{R}_T} \|g\|_{\mathcal{R}_T}$$

e poi, dal Teorema 24.7,

$$\|f - S_N(f)\|_{\mathcal{R}_T} \rightarrow 0 \quad \text{per } N \rightarrow \infty.$$

Ricapitolando,

$$(\widehat{fg})(n) = A_N + B_N, \quad A_N \rightarrow (\hat{f} * \hat{g})(n), \quad B_N \rightarrow 0 \quad \text{per } N \rightarrow \infty,$$

dunque, passando al limite per $N \rightarrow \infty$, si ha la tesi. \square

Nella sezione 25 (*a carattere facoltativo*) si discute la convergenza delle serie “bilatere”, cioè le serie con indice di somma $n \in \mathbb{Z}$; sono osservazioni destinate agli studenti che si sono posti domande o hanno avuto dubbi nel trattare le serie bilatere e/o le serie di numeri complessi.

25 Appendice A. Convergenza delle serie bilatere

In queste lezioni sulle serie di Fourier abbiamo sempre trattato le serie con indice $n \in \mathbb{Z}$ come limite delle somme parziali sull'intervallo simmetrico $[-N, N]$:

$$\sum_{n \in \mathbb{Z}} = \lim_{N \rightarrow \infty} \sum_{n=-N}^N.$$

È una definizione di serie su \mathbb{Z} molto naturale quando si trattano le serie di Fourier di una variabile, e pienamente adatta ai nostri scopi.

D'altra parte, è naturale anche chiedersi se definire le serie tramite gli intervalli simmetrici $[-N, N] \cap \mathbb{Z}$ è equivalente, o se invece è un caso particolare, rispetto all'operazione di limite

$$\sum_{n \in \mathbb{Z}} = \lim_{\substack{A \rightarrow -\infty \\ B \rightarrow \infty}} \sum_{n=A}^B$$

in cui i punti di accumulazione $\pm\infty$ di \mathbb{Z} sono approssiati in modo indipendente (cioè senza supporre che $A = -B$ né altra relazione che legghi A con B).

Oltre agli intervalli $[A, B] \cap \mathbb{Z}$, si possono più in generale considerare tutti i sottoinsiemi non vuoti $F \subset \mathbb{Z}$ con un numero finito di elementi e definire la serie come limite delle somme finite

$$\sum_{n \in \mathbb{Z}} = \lim_{\substack{F \rightarrow \mathbb{Z} \\ F \subset \mathbb{Z} \\ \#(F) < \infty}} \sum_{n \in F}$$

(dove $\#(F)$, scritto anche $|F|$, indica il numero di elementi dell'insieme $F \subset \mathbb{Z}$).

Vediamo in dettaglio la questione limitatamente ai passaggi che abbiamo fatto nelle dimostrazioni precedenti; in pratica rifacciamo per le serie bilatere alcuni risultati elementari di Analisi 1, rivisti in modo un po' più generale (e adatto a varie generalizzazioni).

Le serie reali a termini non negativi hanno limite

Sia $a_n \geq 0$ per ogni $n \in \mathbb{Z}$. Indichiamo con \mathcal{F} la famiglia di tutti gli insiemi finiti, non vuoti, di numeri interi, cioè

$$\mathcal{F} := \{F : F \subset \mathbb{Z}, F \neq \emptyset, \#(F) < \infty\}. \quad (25.1)$$

Consideriamo l'insieme delle somme finite

$$E := \left\{ \sum_{n \in F} a_n : F \in \mathcal{F} \right\}$$

e indichiamo

$$c := \sup E.$$

Ogni elemento di E è un numero reale non negativo, quindi $0 \leq c \leq \infty$.

Supponiamo che $c < \infty$. Per definizione di $c = \sup E$, per ogni $\varepsilon > 0$ esiste un elemento di E maggiore o uguale a $c - \varepsilon$, cioè esiste un sottoinsieme non vuoto, finito, F_0 di \mathbb{Z} tale che

$$c - \varepsilon \leq \sum_{n \in F_0} a_n \leq c.$$

Poiché i termini a_n sono tutti ≥ 0 , allargando l'insieme degli indici si aggiungono termini ≥ 0 , quindi il valore della somma cresce o resta uguale:

$$\sum_{n \in F} a_n \geq \sum_{n \in F_0} a_n$$

per ogni $F \subset \mathbb{Z}$ non vuoto che contiene F_0 .

D'altra parte $\sum_{n \in F} a_n$ è un elemento di E , quindi è $\leq c$, per ogni F finito. Dunque

$$c - \varepsilon \leq \sum_{n \in F} a_n \leq c \quad (25.2)$$

per ogni $F \in \mathcal{F}$ che contiene F_0 . Abbiamo provato che per ogni $\varepsilon > 0$ esiste $F_0 \in \mathcal{F}$ tale che valga (25.2) per tutti gli insiemi $F \in \mathcal{F}$ che contengono F_0 .

Se invece $c = \infty$, per definizione di $c = \sup E$, per ogni $b \in \mathbb{R}$ esiste un elemento di E maggiore o uguale a b , cioè esiste un sottoinsieme finito F_0 di \mathbb{Z} tale che

$$\sum_{n \in F_0} a_n \geq b.$$

Come sopra,

$$\sum_{n \in F} a_n \geq \sum_{n \in F_0} a_n$$

per ogni $F \in \mathcal{F}$ che contiene F_0 (perché $a_n \geq 0$ per ogni n). Questo mostra che per ogni $b \in \mathbb{R}$ esiste $F_0 \in \mathcal{F}$ tale che

$$\sum_{n \in F} a_n \geq b$$

per tutti gli insiemi $F \in \mathcal{F}$ che contengono F_0 .

Esercizio 25.1. Sia $a_n \geq 0$ per ogni $n \in \mathbb{Z}$, e sia \mathcal{F} la famiglia degli insiemi finiti non vuoti di numeri interi, vedi (25.1). Siano

$$E = \left\{ \sum_{n \in F} a_n : F \in \mathcal{F} \right\}, \quad E_1 = \left\{ \sum_{n=A}^B a_n : A, B \in \mathbb{Z}, A \leq B \right\},$$

$$E_2 = \left\{ \sum_{|n| \leq N} a_n : N \in \mathbb{N} \right\}.$$

Dimostrare che $\sup E = \sup E_1 = \sup E_2$.

Consideriamo ora le seguenti tre definizioni di limite.

Definizione 25.2. (*Limiti di somme finite*). Sia $a_n \in \mathbb{R}$ per ogni $n \in \mathbb{Z}$ (senza ipotesi sul segno di a_n) e sia \mathcal{F} l'insieme dei sottoinsiemi finiti, non vuoti, di \mathbb{Z} (vedi (25.1)).

Per $c \in \mathbb{R}$, diciamo che

$$\lim_{F \rightarrow \mathbb{Z}} \sum_{n \in F} a_n = c$$

se per ogni $\varepsilon > 0$ esiste $F_0 \in \mathcal{F}$ tale che

$$c - \varepsilon \leq \sum_{n \in F} a_n \leq c + \varepsilon \quad \forall F \in \mathcal{F}, F \supseteq F_0;$$

diciamo che

$$\lim_{\substack{A \rightarrow -\infty \\ B \rightarrow \infty}} \sum_{n=A}^B a_n = c$$

se per ogni $\varepsilon > 0$ esistono $A_0, B_0 \in \mathbb{Z}$, con $A_0 \leq B_0$, tali che

$$c - \varepsilon \leq \sum_{n=A}^B a_n \leq c + \varepsilon \quad \forall A, B \in \mathbb{Z}, A \leq A_0, B \geq B_0;$$

diciamo che

$$\lim_{N \rightarrow \infty} \sum_{|n| \leq N} a_n = c$$

se per ogni $\varepsilon > 0$ esiste $N_0 \in \mathbb{N}$ tale che

$$c - \varepsilon \leq \sum_{|n| \leq N} a_n \leq c + \varepsilon \quad \forall N \in \mathbb{N}, N \geq N_0.$$

Per $c = \infty$, diciamo che

$$\lim_{F \rightarrow \mathbb{Z}} \sum_{n \in F} a_n = \infty$$

se per ogni $b \in \mathbb{R}$ esiste $F_0 \in \mathcal{F}$ tale che

$$\sum_{n \in F} a_n \geq b \quad \forall F \in \mathcal{F}, F \supseteq F_0;$$

diciamo che

$$\lim_{\substack{A \rightarrow -\infty \\ B \rightarrow \infty}} \sum_{n=A}^B a_n = \infty$$

se per ogni $b \in \mathbb{R}$ esistono $A_0, B_0 \in \mathbb{Z}$, con $A_0 \leq B_0$, tali che

$$\sum_{n=A}^B a_n \geq b \quad \forall A, B \in \mathbb{Z}, A \leq A_0, B \geq B_0;$$

diciamo che

$$\lim_{N \rightarrow \infty} \sum_{|n| \leq N} a_n = \infty$$

se per ogni $b \in \mathbb{R}$ esiste $N_0 \in \mathbb{N}$ tale che

$$\sum_{|n| \leq N} a_n \geq b \quad \forall N \in \mathbb{N}, N \geq N_0.$$

Analoghe definizioni se $c = -\infty$.

Esercizio 25.3. Sia $a_n \geq 0$ per ogni $n \in \mathbb{Z}$. Abbiamo dimostrato (vedi sopra) che

$$\lim_{F \rightarrow \mathbb{Z}} \sum_{n \in F} a_n = \sup_{F \in \mathcal{F}} \sum_{n \in F} a_n$$

sia nel caso in cui tale sup sia finito che in quello di sup infinito. Dimostrare (sempre con l'ipotesi che $a_n \geq 0$ per ogni $n \in \mathbb{Z}$) che

$$\lim_{\substack{A \rightarrow -\infty \\ B \rightarrow \infty}} \sum_{n=A}^B a_n = \sup_{\substack{A, B \in \mathbb{Z} \\ A \leq B}} \sum_{n=A}^B a_n, \quad \lim_{N \rightarrow \infty} \sum_{|n| \leq N} a_n = \sup_{N \in \mathbb{N}} \sum_{|n| \leq N} a_n$$

nei due casi di sup finito/infinito.

Da quanto osservato (inclusi gli Esercizi 25.1, 25.3) segue che

$$\begin{aligned} \lim_{F \rightarrow \mathbb{Z}} \sum_{n \in F} a_n, & \quad \lim_{\substack{A \rightarrow -\infty \\ B \rightarrow \infty}} \sum_{n=A}^B a_n, & \quad \lim_{N \rightarrow \infty} \sum_{|n| \leq N} a_n, \\ \sup_{F \in \mathcal{F}} \sum_{n \in F} a_n, & \quad \sup_{\substack{A, B \in \mathbb{Z} \\ A \leq B}} \sum_{n=A}^B a_n, & \quad \sup_{N \in \mathbb{N}} \sum_{|n| \leq N} a_n \end{aligned} \quad (25.3)$$

sono tutti coincidenti.

In conclusione: per le serie reali a termini non negativi non c'è niente di restrittivo nell'usare gli intervalli simmetrici $[-N, N]$ per definire la somma della serie: possiamo definire

$$\sum_{n \in \mathbb{Z}} a_n := \lim_{N \rightarrow \infty} \sum_{|n| \leq N} a_n,$$

si ha

$$\lim_{N \rightarrow \infty} \sum_{|n| \leq N} a_n = \sup_{N \in \mathbb{N}} \sum_{|n| \leq N} a_n,$$

e definizioni apparentemente più generali danno lo stesso valore (finito o infinito) per la somma della serie.

Criterio del confronto

Sia $0 \leq a_n \leq b_n$ per ogni $n \in \mathbb{Z}$. Se la serie $\sum_{n \in \mathbb{Z}} b_n$ converge, allora anche la serie $\sum_{n \in \mathbb{Z}} a_n$ converge.

La dimostrazione è immediata: per ogni $N \in \mathbb{N}$ si ha

$$\sum_{|n| \leq N} a_n \leq \sum_{|n| \leq N} b_n \leq \sup_{M \in \mathbb{N}} \sum_{|n| \leq M} b_n = \sum_{n \in \mathbb{Z}} b_n =: b < \infty,$$

quindi $b \in \mathbb{R}$ è un maggiorante dell'insieme delle somme $\sum_{|n| \leq N} a_n$, e dunque

$$\sum_{n \in \mathbb{Z}} a_n = \sup_{N \in \mathbb{N}} \sum_{|n| \leq N} a_n \leq b < \infty.$$

Poiché $a_n \geq 0$, i sei modi di calcolare la somma della serie in (25.3) sono equivalenti, cioè danno tutti e sei lo stesso risultato.

La convergenza assoluta implica la convergenza

Sia ora $a_n \in \mathbb{R}$ (senza assumere che siano non negativi), e supponiamo che la serie $\sum_{n \in \mathbb{Z}} |a_n|$ converga. Definiamo

$$b_n := a_n + |a_n|,$$

e osserviamo che

$$0 \leq b_n \leq 2|a_n|$$

perché $-|a_n| \leq a_n \leq |a_n|$. Per ipotesi,

$$\alpha := \sum_{n \in \mathbb{Z}} |a_n| < \infty,$$

e dunque, dal criterio del confronto, anche la serie di termine generale b_n converge,

$$\beta := \sum_{n \in \mathbb{Z}} b_n < \infty.$$

Poiché $|a_n|, b_n$ sono non negativi, si ottengono gli stessi valori α, β in tutti e sei i modi (vedi (25.3)) di calcolare la somma delle serie.

Mostriamo che

$$\lim_{F \rightarrow \mathbb{Z}} \sum_{n \in F} a_n = \beta - \alpha.$$

Sia $\varepsilon > 0$. Poiché $|a_n| \geq 0$, da quanto abbiamo visto per le serie a termine generale non negativo si ha

$$\alpha = \sum_{n \in \mathbb{Z}} |a_n| = \lim_{F \rightarrow \mathbb{Z}} \sum_{n \in F} |a_n|,$$

quindi esiste $F_1 \in \mathcal{F}$ tale che

$$\alpha - \varepsilon \leq \sum_{n \in F} |a_n| \leq \alpha \quad (25.4)$$

per ogni $F \in \mathcal{F}$ che contiene F_1 . Similmente, poiché $b_n \geq 0$, si ha

$$\beta = \sum_{n \in \mathbb{Z}} b_n = \lim_{F \rightarrow \mathbb{Z}} \sum_{n \in F} b_n,$$

dunque esiste $F_2 \in \mathcal{F}$ tale che

$$\beta - \varepsilon \leq \sum_{n \in F} b_n \leq \beta \quad (25.5)$$

per ogni $F \in \mathcal{F}$ che contiene F_2 . Sia $F_0 := F_1 \cup F_2$. Allora per ogni $F \in \mathcal{F}$ che contiene F_0 valgono (25.4), (25.5), da cui si ottiene

$$\beta - \alpha - \varepsilon \leq \left(\sum_{n \in F} b_n \right) - \left(\sum_{n \in F} |a_n| \right) \leq \beta - \alpha + \varepsilon.$$

Poiché

$$\left(\sum_{n \in F} b_n \right) - \left(\sum_{n \in F} |a_n| \right) = \sum_{n \in F} (b_n - |a_n|) = \sum_{n \in F} a_n,$$

si ha

$$\beta - \alpha - \varepsilon \leq \sum_{n \in F} a_n \leq \beta - \alpha + \varepsilon.$$

Abbiamo dunque provato che per ogni $\varepsilon > 0$ esiste $F_0 \in \mathcal{F}$ tale che

$$\left| \left(\sum_{n \in F} a_n \right) - (\beta - \alpha) \right| \leq \varepsilon$$

per ogni $F \in \mathcal{F}$ che contiene F_0 , cioè

$$\lim_{F \rightarrow \mathbb{Z}} \sum_{n \in F} a_n = \beta - \alpha. \quad (25.6)$$

Esercizio 25.4. Sia $a_n \in \mathbb{R}$ per ogni $n \in \mathbb{Z}$, e sia $\sum_{n \in \mathbb{Z}} |a_n| < \infty$. Abbiamo dimostrato (vedi (25.6)) che la serie di termine generale a_n converge per $F \rightarrow \mathbb{Z}$ in \mathcal{F} . Dimostrare che la serie converge allo stesso limite anche per $A \rightarrow -\infty$, $B \rightarrow +\infty$ in \mathbb{Z} e per $N \rightarrow \infty$ in \mathbb{N} , cioè per tutte e tre le definizioni di limite della Definizione 25.2.

In conclusione: le serie assolutamente convergenti convergono secondo tutte e tre le definizioni di limite della Definizione 25.2, e convergono sempre allo stesso valore.

Senza la convergenza assoluta le tre definizioni di limite delle somme finite non sono equivalenti

Quanto detto sopra per le serie assolutamente convergenti non vale, in generale, per le serie che non convergono assolutamente. Per esempio, consideriamo

$$a_n = \text{sign}(n),$$

cioè $a_n = 1$ per n positivo, $a_n = -1$ per n negativo, e $a_0 = 0$. Si ha

$$\sum_{|n| \leq N} a_n = 0 \quad \forall N \in \mathbb{N},$$

per cui

$$\sup_{N \in \mathbb{N}} \sum_{|n| \leq N} a_n = 0, \quad \lim_{N \rightarrow \infty} \sum_{|n| \leq N} a_n = 0,$$

mentre

$$\sum_{n=A}^B a_n = k \quad \text{per } A = -k, \quad B = 2k, \quad k \in \mathbb{N},$$

quindi

$$\sup_{\substack{A, B \in \mathbb{Z} \\ A < B}} \sum_{n=A}^B a_n \geq \sup \left\{ \sum_{n=A}^B a_n : A = -k, \quad B = 2k, \quad k \in \mathbb{N} \right\} = \sup \{k : k \in \mathbb{N}\} = \infty.$$

D'altra parte

$$\sum_{n=A}^B a_n = -k \quad \text{per } A = -2k, \quad B = k, \quad k \in \mathbb{N},$$

quindi

$$\nexists \lim_{\substack{A \rightarrow -\infty \\ B \rightarrow \infty}} \sum_{n=A}^B a_n$$

perché l'insieme delle somme del tipo $\sum_{n=A}^B a_n$ contiene una successione che tende a $+\infty$, una successione che tende a $-\infty$ e una successione che tende a zero (quella con $A = -B$).

Esercizio 25.5. Prova che se $a_n = \text{sign}(n)$ (esempio qua sopra), per ogni $m \in \mathbb{Z}$ l'insieme delle somme del tipo $\sum_{n=A}^B a_n$ contiene una successione che tende a m (successione costante di valore m).

Serie assolutamente convergenti di numeri complessi

Sia $c_n \in \mathbb{C}$ per ogni $n \in \mathbb{Z}$. Scriviamo parte reale e immaginaria,

$$c_n = a_n + ib_n, \quad a_n, b_n \in \mathbb{R}.$$

Supponiamo che la serie $\sum c_n$ sia assolutamente convergente, cioè che la serie a termini reali non negativi

$$\sum_{n \in \mathbb{Z}} |c_n|$$

sia convergente. Poiché

$$|a_n| \leq \sqrt{a_n^2 + b_n^2} = |c_n|, \quad |b_n| \leq \sqrt{a_n^2 + b_n^2} = |c_n|,$$

dal criterio del confronto le serie a termini non negativi

$$\sum_{n \in \mathbb{Z}} |a_n|, \quad \sum_{n \in \mathbb{Z}} |b_n|$$

convergono (in tutti e sei i modi di calcolare la somma della serie in (25.3)). Quindi le serie

$$\sum_{n \in \mathbb{Z}} a_n, \quad \sum_{n \in \mathbb{Z}} b_n$$

convergono in \mathbb{R} , e convergono allo stesso limite secondo tutte e tre le definizioni di limite della Definizione 25.2. Diciamo

$$a := \sum_{n \in \mathbb{Z}} a_n, \quad b := \sum_{n \in \mathbb{Z}} b_n.$$

Per ogni $F \in \mathcal{F}$ si ha

$$\sum_{n \in F} c_n = \sum_{n \in F} (a_n + ib_n) = \left(\sum_{n \in F} a_n \right) + i \left(\sum_{n \in F} b_n \right).$$

Di conseguenza si ha

$$\lim_{F \rightarrow \mathbb{Z}} \sum_{n \in F} c_n = \lim_{F \rightarrow \mathbb{Z}} \left(\sum_{n \in F} a_n \right) + i \lim_{F \rightarrow \mathbb{Z}} \left(\sum_{n \in F} b_n \right) = a + ib,$$

e si trova lo stesso valore $c = a + ib$ anche secondo le altre due definizioni di limite della Definizione 25.2.

26 Appendice B. Altre dimostrazioni

Riportiamo la dimostrazione del Teorema di Peano basata sul metodo di Eulero di discretizzazione del tempo (dimostrazione soltanto accennata a lezione, senza dettagli).

Dimostrazione del Teorema 6.6 di esistenza di Peano con metodo di Eulero

Costruiamo una successione (u_n) di funzioni lineari a tratti e dimostriamo che converge a una soluzione del problema di Cauchy.

Passo 1: definizione di u_n . Consideriamo l'intervallo $[t_0, t_0 + \delta]$ (nell'intervallo $[t_0 - \delta, t_0]$ la dimostrazione procede in modo simile). Fissiamo n intero positivo, e dividiamo $[t_0, t_0 + \delta]$ in n parti uguali, cioè in n sottointervalli

$$[t_0^{(n)}, t_1^{(n)}], \quad [t_1^{(n)}, t_2^{(n)}], \quad \dots \quad [t_{n-1}^{(n)}, t_n^{(n)}],$$

ciascuno di lunghezza δ/n , dove

$$t_k^{(n)} := t_0 + k \frac{\delta}{n}, \quad k = 0, \dots, n.$$

Sul primo intervallo definiamo

$$u_n(t) := y_0 + f(t_0, y_0)(t - t_0) \quad \forall t \in [t_0, t_1^{(n)}], \quad (26.1)$$

$$y_1^{(n)} := u_n(t_1^{(n)}). \quad (26.2)$$

Dunque

$$y_1^{(n)} = y_0 + f(t_0, y_0) \frac{\delta}{n}, \quad (26.3)$$

e, poiché $|f(t_0, y_0)| \leq M$, si ha

$$|y_1^{(n)} - y_0| = |f(t_0, y_0)| \frac{\delta}{n} \leq M \frac{\delta}{n}.$$

Quindi $|y_1^{(n)} - y_0| \leq M\delta/n \leq M\delta \leq b$, perciò $y_1^{(n)} \in J$, e la coppia $(t_1^{(n)}, y_1^{(n)})$ appartiene al dominio $I \times J$ del campo f . Per $n \geq 2$, gli intervalli in cui definire u_n sono almeno due; essendo $f(t_1^{(n)}, y_1^{(n)})$ ben definito, possiamo “aggiornare” in $t_1^{(n)}$ i valori del dato y e del campo f per definire u_n sul secondo intervallo: definiamo

$$u_n(t) := y_1^{(n)} + f(t_1^{(n)}, y_1^{(n)})(t - t_1^{(n)}) \quad \forall t \in [t_1^{(n)}, t_2^{(n)}],$$

$$y_2^{(n)} := u_n(t_2^{(n)}).$$

Osserviamo che, usando la (26.3), si ha

$$\begin{aligned} y_2^{(n)} &= y_1^{(n)} + f(t_1^{(n)}, y_1^{(n)}) \frac{\delta}{n} \\ &= y_0 + \left(f(t_0, y_0) + f(t_1^{(n)}, y_1^{(n)}) \right) \frac{\delta}{n}. \end{aligned}$$

Quindi, ricordando che $|f(t, y)| \leq M$ per ogni $(t, y) \in I \times J$,

$$|y_2^{(n)} - y_1^{(n)}| \leq M \frac{\delta}{n}, \quad |y_2^{(n)} - y_0| \leq 2M \frac{\delta}{n}. \quad (26.4)$$

Di conseguenza $|y_2^{(n)} - y_0| \leq 2M\delta/n \leq M\delta \leq b$, dunque $y_2^{(n)} \in J$, $f(t_2^{(n)}, y_2^{(n)})$ è ben definito, e, per $n \geq 3$, possiamo “aggiornare” il dato y e il campo f per definire u_n sul terzo intervallo, e così via.

Vediamo in generale il passo induttivo: supponiamo di avere definito u_n sugli intervalli $[t_0, t_1^{(n)}], \dots, [t_{k-1}^{(n)}, t_k^{(n)}]$, per un certo $k \leq n-1$, con

$$u_n(t) := y_j^{(n)} + f(t_j^{(n)}, y_j^{(n)})(t - t_j^{(n)}) \quad \forall t \in [t_j^{(n)}, t_{j+1}^{(n)}], \quad \forall j = 0, \dots, k-1,$$

$$y_j^{(n)} := u_n(t_j^{(n)}) \quad \forall j = 1, \dots, k,$$

e supponiamo di avere osservato che

$$\begin{aligned} y_k^{(n)} &= y_{k-1}^{(n)} + f(t_{k-1}^{(n)}, y_{k-1}^{(n)}) \frac{\delta}{n} \\ &= y_0 + \left(f(t_0, y_0) + f(t_1^{(n)}, y_1^{(n)}) + \dots + f(t_{k-1}^{(n)}, y_{k-1}^{(n)}) \right) \frac{\delta}{n} \end{aligned}$$

e

$$|y_k^{(n)} - y_j^{(n)}| \leq M(k-j) \frac{\delta}{n} \quad \forall j = 0, \dots, k$$

(dove $y_0^{(n)} := y_0$). Allora $|y_k^{(n)} - y_0| \leq Mk\delta/n \leq M\delta \leq b$, e $y_k^{(n)} \in J$. Quindi $f(t_k^{(n)}, y_k^{(n)})$ è ben definito, e possiamo proseguire a definire u_n sull'intervallo successivo: definiamo

$$u_n(t) := y_k^{(n)} + f(t_k^{(n)}, y_k^{(n)})(t - t_k^{(n)}) \quad \forall t \in [t_k^{(n)}, t_{k+1}^{(n)}],$$

$$y_{k+1}^{(n)} := u_n(t_{k+1}^{(n)}).$$

Usando le ipotesi induttive, osserviamo che

$$\begin{aligned} y_{k+1}^{(n)} &= y_k^{(n)} + f(t_k^{(n)}, y_k^{(n)}) \frac{\delta}{n} \\ &= y_0 + \left(f(t_0, y_0) + \dots + f(t_k^{(n)}, y_k^{(n)}) \right) \frac{\delta}{n} \end{aligned}$$

e, per ogni $0 \leq j \leq k$,

$$\begin{aligned} |y_{k+1}^{(n)} - y_j^{(n)}| &\leq |y_{k+1}^{(n)} - y_k^{(n)}| + |y_k^{(n)} - y_j^{(n)}| \\ &\leq M \frac{\delta}{n} + M(k-j) \frac{\delta}{n} = M(k+1-j) \frac{\delta}{n}, \end{aligned}$$

mentre per $j = k+1$ la disuguaglianza $|y_{k+1}^{(n)} - y_j^{(n)}| \leq M(k+1-j)\delta/n$ è banalmente vera. Questo dimostra il passo induttivo. Dunque, per induzione, u_n è ben definita su tutti gli intervalli $[t_0, t_1^{(n)}], \dots, [t_{n-1}^{(n)}, t_n^{(n)}]$, cioè su tutto $[t_0, t_0 + \delta]$.

In conclusione, abbiamo

$$u_n(t) = y_k^{(n)} + f(t_k^{(n)}, y_k^{(n)})(t - t_k^{(n)}) \quad \forall t \in [t_k^{(n)}, t_{k+1}^{(n)}], \quad \forall k = 0, \dots, n-1, \quad (26.5)$$

$$y_k^{(n)} = u_n(t_k^{(n)}) \quad \forall k = 1, \dots, n, \quad (26.6)$$

$$\begin{aligned} y_k^{(n)} &= y_{k-1}^{(n)} + f(t_{k-1}^{(n)}, y_{k-1}^{(n)}) \frac{\delta}{n} \\ &= y_0 + \left(f(t_0, y_0) + f(t_1^{(n)}, y_1^{(n)}) + \dots + f(t_{k-1}^{(n)}, y_{k-1}^{(n)}) \right) \frac{\delta}{n} \quad \forall k = 1, \dots, n, \end{aligned} \quad (26.7)$$

$$|y_k^{(n)} - y_j^{(n)}| \leq M(k-j) \frac{\delta}{n} \quad \forall k, j \in \{0, \dots, n\}, \quad j \leq k. \quad (26.8)$$

Passo 2: applicazione di Ascoli-Arzelà. Per ogni $k = 0, \dots, n-1$, per ogni $t \in [t_k^{(n)}, t_{k+1}^{(n)}]$, usando (26.5) e (26.8), e ricordando che $|f| \leq M$ in $I \times J$, si ha

$$\begin{aligned} |u_n(t)| &\leq |u_n(t) - y_k^{(n)}| + |y_k^{(n)} - y_0| + |y_0| \\ &\leq |f(t_k^{(n)}, y_k^{(n)})(t - t_k^{(n)})| + Mk \frac{\delta}{n} + |y_0| \\ &\leq M \frac{\delta}{n} + Mk \frac{\delta}{n} + |y_0| \leq M\delta + |y_0|. \end{aligned}$$

Dunque

$$|u_n(t)| \leq M\delta + |y_0| \quad \forall t \in [t_0, t_0 + \delta], \quad \forall n \geq 1,$$

cosicché la successione (u_n) è equilimitata in $[t_0, t_0 + \delta]$.

Dimostriamo che (u_n) è anche equicontinua. Siano $t, \tau \in [t_0, t_0 + \delta]$, e sia $n \geq 1$. Se t e τ appartengono a uno stesso intervallo $[t_k^{(n)}, t_{k+1}^{(n)}]$, allora da (26.5) segue che

$$|u_n(t) - u_n(\tau)| = |f(t_k^{(n)}, y_k^{(n)})(t - \tau)| \leq M|t - \tau|.$$

Se invece t e τ non appartengono allo stesso intervallo, si ha

$$\tau \in [t_j^{(n)}, t_{j+1}^{(n)}], \quad t \in [t_k^{(n)}, t_{k+1}^{(n)}]$$

per certi $j, k \in \{0, \dots, n-1\}$, con $j \neq k$. Supponiamo che sia $j < k$. Allora $\tau \leq t_{j+1}^{(n)} \leq t_k^{(n)} \leq t$, e

$$|u_n(t) - u_n(\tau)| \leq |u_n(t) - u_n(t_k^{(n)})| + |u_n(t_k^{(n)}) - u_n(t_{j+1}^{(n)})| + |u_n(t_{j+1}^{(n)}) - u_n(\tau)|.$$

Poiché $t \in [t_k^{(n)}, t_{k+1}^{(n)}]$, da (26.5) segue che

$$|u_n(t) - u_n(t_k^{(n)})| \leq M|t - t_k^{(n)}| = M(t - t_k^{(n)}).$$

Poiché $\tau \in [t_j^{(n)}, t_{j+1}^{(n)}]$, da (26.5) segue che

$$|u_n(t_{j+1}^{(n)}) - u_n(\tau)| \leq M|t_{j+1}^{(n)} - \tau| = M(t_{j+1}^{(n)} - \tau).$$

Da (26.8),

$$|u_n(t_k^{(n)}) - u_n(t_{j+1}^{(n)})| = |y_k^{(n)} - y_{j+1}^{(n)}| \leq M(k-j-1) \frac{\delta}{n} = M(t_k^{(n)} - t_{j+1}^{(n)}).$$

Sommando le stime dei tre termini otteniamo

$$|u_n(t) - u_n(\tau)| \leq M(t - \tau).$$

Perciò in ogni caso si ha

$$|u_n(t) - u_n(\tau)| \leq M|t - \tau| \quad \forall t, \tau \in [t_0, t_0 + \delta], \quad \forall n \geq 1.$$

Tutte le funzioni u_n sono lipschitziane in $[t_0, t_0 + \delta]$ con la stessa costante di Lipschitz M , e dunque la successione (u_n) è equicontinua in $[t_0, t_0 + \delta]$.

Dal Teorema di Ascoli-Arzelà, esiste una sottosuccessione di (u_n) , che con un abuso di notazione indichiamo (u_n) , che converge uniformemente in $[t_0, t_0 + \delta]$. Indichiamo u il suo limite uniforme. Poiché ogni u_n è continua, anche u è continua.

Passo 3: u è soluzione. Dimostriamo che $u_n(t)$ converge uniformemente a

$$y_0 + \int_{t_0}^t f(s, u(s)) ds.$$

Poiché $u_n(t)$ converge a $u(t)$, l'unicità del limite e la formulazione integrale del problema di Cauchy daranno la tesi.

Sia $\varepsilon > 0$. Dalla uniforme continuità di f in $I \times J$ esiste $\delta_0 > 0$ tale che

$$|f(t, y) - f(s, z)| \leq \frac{\varepsilon}{\delta} \quad \forall (t, y), (s, z) \in I \times J, \quad |t - s| + |y - z| \leq \delta_0.$$

Poiché $u_n \rightarrow u$ uniformemente, esiste $\bar{n} \in \mathbb{N}$ tale che

$$(M + 1) \frac{\delta}{n} + \|u_n - u\|_\infty \leq \delta_0 \quad \forall n \geq \bar{n},$$

dove $\|\cdot\|_\infty$ indica la sup-norma su $[t_0, t_0 + \delta]$.

Sia $n \geq \bar{n}$, sia $0 \leq j \leq n - 1$, e sia $s \in [t_j^{(n)}, t_{j+1}^{(n)}]$. Stimiamo la distanza tra i punti $(s, u(s))$ e $(t_j^{(n)}, y_j^{(n)})$:

$$\begin{aligned} |s - t_j^{(n)}| + |u(s) - y_j^{(n)}| &\leq \frac{\delta}{n} + |u(s) - u_n(s)| + |u_n(s) - u_n(t_j^{(n)})| \\ &\leq \frac{\delta}{n} + \|u - u_n\|_\infty + |f(t_j^{(n)}, y_j^{(n)})(s - t_j^{(n)})| \\ &\leq \frac{\delta}{n} + \|u - u_n\|_\infty + M \frac{\delta}{n} \\ &\leq \delta_0. \end{aligned}$$

Di conseguenza per ogni $s \in [t_j^{(n)}, t_{j+1}^{(n)}]$, $j = 0, \dots, n - 1$, $n \geq \bar{n}$ si ha

$$|f(s, u(s)) - f(t_j^{(n)}, y_j^{(n)})| \leq \frac{\varepsilon}{\delta}. \quad (26.9)$$

Sia ora $n \geq \bar{n}$, $k \in \{0, \dots, n - 1\}$, e $t \in [t_k^{(n)}, t_{k+1}^{(n)}]$. Usando (26.5) e (26.7) si ha

$$u_n(t) = y_0 + \left(f(t_0, y_0) + \dots + f(t_{k-1}^{(n)}, y_{k-1}^{(n)}) \right) \frac{\delta}{n} + f(t_k^{(n)}, y_k^{(n)})(t - t_k^{(n)}).$$

Ricordando che $t_{j+1}^{(n)} - t_j^{(n)} = \delta/n$ per ogni j , possiamo riscrivere l'ultima identità come

$$u_n(t) = y_0 + \int_{t_0}^{t_1^{(n)}} f(t_0, y_0) ds + \dots + \int_{t_{k-1}^{(n)}}^{t_k^{(n)}} f(t_{k-1}^{(n)}, y_{k-1}^{(n)}) ds + \int_{t_k^{(n)}}^t f(t_k^{(n)}, y_k^{(n)}) ds.$$

D'altra parte, spezzando l'integrale sui vari intervalli, si ha

$$\int_{t_0}^t f(s, u(s)) ds = \int_{t_0}^{t_1^{(n)}} f(s, u(s)) ds + \dots + \int_{t_{k-1}^{(n)}}^{t_k^{(n)}} f(s, u(s)) ds + \int_{t_k^{(n)}}^t f(s, u(s)) ds.$$

Quindi

$$\begin{aligned} & y_0 + \int_{t_0}^t f(s, u(s)) ds - u_n(t) \\ &= \int_{t_0}^{t_1^{(n)}} \left(f(s, u(s)) - f(t_0, y_0) \right) ds + \dots + \int_{t_{k-1}^{(n)}}^{t_k^{(n)}} \left(f(s, u(s)) - f(t_{k-1}^{(n)}, y_{k-1}^{(n)}) \right) ds \\ & \quad + \int_{t_k^{(n)}}^t \left(f(s, u(s)) - f(t_k^{(n)}, y_k^{(n)}) \right) ds. \end{aligned}$$

Usando la stima (26.9) abbiamo allora

$$\begin{aligned} \left| y_0 + \int_{t_0}^t f(s, u(s)) ds - u_n(t) \right| &\leq \int_{t_0}^{t_1^{(n)}} \frac{\varepsilon}{\delta} ds + \dots + \int_{t_{k-1}^{(n)}}^{t_k^{(n)}} \frac{\varepsilon}{\delta} ds + \int_{t_k^{(n)}}^t \frac{\varepsilon}{\delta} ds \\ &\leq (k+1) \frac{\varepsilon}{\delta} \frac{\delta}{n} \leq \varepsilon. \end{aligned}$$

Abbiamo provato che per ogni $\varepsilon > 0$ esiste $\bar{n} \in \mathbb{N}$ tale che

$$\left| y_0 + \int_{t_0}^t f(s, u(s)) ds - u_n(t) \right| \leq \varepsilon$$

per ogni $t \in [t_0, t_0 + \delta]$, per ogni $n \geq \bar{n}$, cioè che $u_n(t)$ converge uniformemente a $y_0 + \int_{t_0}^t f(s, u(s)) ds$ in $[t_0, t_0 + \delta]$. D'altra parte, $u_n(t)$ converge a $u(t)$, quindi

$$u(t) = y_0 + \int_{t_0}^t f(s, u(s)) ds \quad \forall t \in [t_0, t_0 + \delta],$$

e u è soluzione del problema di Cauchy. □

Il Criterio di prolungabilità con limite continuo (Corollario 8.1) si può ottenere in modo molto semplice anche con dimostrazione diretta (invece che come conseguenza del Criterio di prolungabilità con limite di successione, Proposizione 7.6), nel modo seguente.

Dimostrazione diretta del Corollario 8.1

Poiché $(b, v_0) \in \Omega$, possiamo considerare il problema di Cauchy con dati iniziali (b, v_0) . Dal Teorema di esistenza locale e unicità, esiste $\delta > 0$ tale che esiste una e una sola soluzione $\varphi : [b - \delta, b + \delta] \rightarrow \mathbb{R}^N$ di tale problema di Cauchy nell'intervallo $[b - \delta, b + \delta]$, cioè φ soddisfa

$$\begin{cases} \varphi'(t) = f(t, \varphi(t)) & \forall t \in [b - \delta, b + \delta], \\ \varphi(b) = v_0. \end{cases}$$

Definiamo $v : (a, b + \delta) \rightarrow \mathbb{R}^N$ in questo modo:

$$v(t) := \begin{cases} u(t) & \text{per } a < t < b, \\ v_0 & \text{per } t = b, \\ \varphi(t) & \text{per } b < t < b + \delta. \end{cases}$$

Per ogni $t \in (a, b)$ si ha

$$v'(t) = u'(t) = f(t, u(t)) = f(t, v(t))$$

perché $v = u$ in (a, b) e u è soluzione dell'equazione differenziale. Similmente, per ogni $t \in (b, b + \delta)$ si ha

$$v'(t) = \varphi'(t) = f(t, \varphi(t)) = f(t, v(t))$$

perché $v = \varphi$ in $(b, b + \delta)$ e φ è soluzione dell'equazione differenziale. Quindi l'uguaglianza

$$v'(t) = f(t, v(t))$$

vale per ogni $t \in (a, b) \cup (b, b + \delta)$. Vogliamo dimostrare che vale anche in $t = b$.

Calcoliamo dapprima i limiti sinistro e destro di v per $t \rightarrow b$: per ipotesi,

$$\lim_{t \rightarrow b^-} v(t) = \lim_{t \rightarrow b^-} u(t) = v_0 = v(b),$$

mentre, dalla continuità di φ ,

$$\lim_{t \rightarrow b^+} v(t) = \lim_{t \rightarrow b^+} \varphi(t) = \varphi(b) = v_0 = v(b).$$

Quindi v è continua in b , e dunque è continua in tutto l'intervallo $(a, b + \delta)$.

Calcoliamo i limiti sinistro e destro del rapporto incrementale: dal Teorema di de l'Hôpital,

$$\lim_{t \rightarrow b^-} \frac{v(t) - v(b)}{t - b} = \lim_{t \rightarrow b^-} \frac{u(t) - v_0}{t - b} = \lim_{t \rightarrow b^-} \frac{u'(t)}{1} = \lim_{t \rightarrow b^-} f(t, u(t)) = f(b, v_0)$$

mentre, dalla derivabilità di φ ,

$$\lim_{t \rightarrow b^+} \frac{v(t) - v(b)}{t - b} = \lim_{t \rightarrow b^+} \frac{\varphi(t) - v_0}{t - b} = \lim_{t \rightarrow b^+} \frac{\varphi(t) - \varphi(b)}{t - b} = \varphi'(b) = f(b, \varphi(b)) = f(b, v_0).$$

Quindi v è derivabile in b , e la sua derivata in b vale

$$v'(b) = f(b, v_0) = f(b, v(b)).$$

Dunque v soddisfa l'identità $v'(t) = f(t, v(t))$ per ogni $t \in (a, b + \delta)$. Poiché $v = u$ in (a, b) , v è un prolungamento di u ad $(a, b + \delta)$. \square

Indice

1	Lezione 1	5
2	Lezione 2	10
3	Lezione 3	16
4	Lezione 4	24
5	Lezione 5	32
6	Lezione 6	40
7	Lezione 7	48
8	Lezione 8	54
9	Lezione 9	60
10	Lezione 10	66
11	Lezione 11	79
12	Lezione 12	87
13	Lezione 13	95
14	Lezione 14	108
15	Lezione 15	123
16	Lezione 16	131
17	Lezione 17	139
18	Lezione 18	149
19	Lezione 19	158
20	Lezione 20	166
21	Lezione 21	176
22	Lezione 22	191
23	Lezione 23	202
24	Lezione 24	213
25	Appendice A. Convergenza delle serie bilatere	223
26	Appendice B. Altre dimostrazioni	230

Indice dettagliato

Lezione 1

Definizione 1.1 (Soluzione di un'equazione differenziale)

Definizione 1.2 (Ordine di un'equazione differenziale)

Definizione 1.3 (Equazione in forma normale)

Definizione 1.4 (Sistema di equazioni scalari, o equazione vettoriale, del primo ordine)

Definizione 1.5 (Campo vettoriale, spazio delle fasi, traiettorie, sistema dinamico n -dimensionale)

Definizione 1.6 (Sistema autonomo)

Osservazione 1.7 (Equivalenza tra 1 equazione scalare di ordine n ed n equazioni scalari di ordine 1)

Definizione 1.8 (Problema di Cauchy)

Definizione 1.9 (Integrale di funzione a valori in \mathbb{R}^N)

Lemma 1.10 (Formulazione integrale del problema di Cauchy)

Lezione 2

Teorema 2.1 (Teorema di esistenza e unicità locale per il problema di Cauchy)

Osservazione 2.2 (Iterate di Picard)

Osservazione 2.3 (Problema di punto fisso)

Esercizio 2.4 (Domanda riguardo alla dimostrazione del Teorema 2.1)

Lezione 3

Esercizio 3.1 (Convergenza uniforme)

Definizione 3.2 (Funzione $f(t, y)$ localmente lipschitziana in y uniformemente in t)

Osservazione 3.3 (Esistenza e unicità locale su aperti)

Definizione 3.4 (Norma euclidea di una matrice)

Lemma 3.5 (Disuguaglianza di Cauchy-Schwarz per matrice contro vettore)

Lemma 3.6 (Disuguaglianza per l'integrale di funzione vettoriale)

Definizione 3.7 (Norma operatoriale di matrice)

Esercizio 3.8 (Formule equivalenti della norma operatoriale)

Esercizio 3.9 (Equivalenza di norma euclidea e operatoriale)

Proposizione 3.10 (Criterio di lipschitzianità)

Teorema 3.11 (Teorema del valor medio in \mathbb{R}^N)

Osservazione 3.12 (C^1 implica Lipschitz)

Osservazione 3.13 (Il Teorema di Lagrange non vale in dimensione ≥ 2)

Esempi di applicazione del Teorema di esistenza locale e unicità.

Esercizio 3.14 (La radice quadrata non è lipschitziana)

Corollario 3.15 (Conseguenza dell'unicità: soluzioni distinte non si toccano)

Lezione 4

Dimostrazione del Corollario 3.15

Osservazione 4.1 (Versione topologica della dimostrazione del Corollario 3.15)

Esercizio 4.2 (esempio di applicazione del Corollario 3.15)

Osservazione 4.3 (Soluzioni distinte non si toccano: caso scalare e vettoriale)

Osservazione 4.4 (Due soluzioni distinte possono visitare gli stessi punti dello spazio delle fasi ma a tempi diversi)

Lemma 4.5 (Regolarità delle soluzioni)

Osservazione 4.6 (Bootstrap)

Definizione 4.7 (Equazioni a variabili separabili)

Definizione 4.8 (Equilibrio)

Equazioni riconducibili ad equazioni a variabili separabili.

Formula risolutiva dell'equazione $u' + a(t)u = 0$.

Formula risolutiva dell'equazione $u' + a(t)u = b(t)$.

Esercizi sulle tecniche risolutive.

Lezione 5

Esercizi sulle tecniche risolutive.

Equazioni del tipo $u' = g(u/t)$.

Equazioni di Bernoulli.

Equazioni esatte (cioè equazioni con integrale primo).

Definizione 5.6 (Integrale primo)

Osservazione 5.9 (Esistenza senza unicità: il pennello di Peano)

Proposizione 5.10 ($C([a, b], \mathbb{R}^N)$ è spazio di Banach)

Teorema 5.11 (Teorema di Ascoli-Arzelà)

Osservazione 5.12 (Notazione per infinite sottosuccessioni estratte ricorsivamente da sottosuccessioni e matrici di infinite righe e infinite colonne)

Lezione 6

Conclusione della dimostrazione del Teorema di Ascoli-Arzelà

Osservazione 6.1 (Esempio di costruzione delle sottosuccessioni nella dimostrazione di Ascoli-Arzelà)

Proposizione 6.2 (Lipschitzianità locale implica lipschitzianità sui compatti)

Proposizione 6.3 (Lipschitzianità locale implica lipschitzianità sui compatti/2)

Proposizione 6.4 (Locale lipschitzianità in y uniforme in t implica lipschitzianità in y uniforme in t sui compatti)

Esercizio 6.5 (Locale lipschitzianità)

Teorema 6.6 (Teorema di esistenza di Peano)

Lemma 6.7 (Se il campo f è approssimabile con funzioni lipschitziane allora esiste una soluzione)

Lemma 6.8 (Approssimazione di funzioni continue con funzioni lipschitziane)

Lezione 7

Conclusione della dimostrazione del Lemma 6.8.

Osservazione 7.1 (Convoluzione)

Esercizio 7.2 (Iterate di Picard e Teorema di esistenza di Peano)

Definizione 7.3 (Prolungamento)

Osservazione 7.4 (Unicità del prolungamento sull'intersezione)

Teorema 7.5 (Esistenza del prolungamento massimale)

Proposizione 7.6 (Criterio di prolungabilità: limite di successione)

Lezione 8

Corollario 8.1 (Criterio di prolungabilità: limite continuo)

Teorema 8.2 (Teorema di fuga dai compatti)

Osservazione 8.3 (Compatti in spazio-tempo)

Corollario 8.4 (Compatto nello spazio delle fasi implica esistenza globale nel tempo)

Esercizi 8.5 e 8.6 (esercizi svolti di analisi qualitativa)

Lezione 9

Esercizio 9.1 (esercizio svolto di analisi qualitativa)

Teorema 9.2 (Teorema di esistenza globale)

Teorema 9.3 (Teorema del confronto)

Esercizio 9.4 (su esistenza globale)

Lezione 10

Dimostrazione del Teorema 9.3 del confronto

Osservazione 10.1 (Da dove viene il fattore e^{-Lt} ?)

Esercizio 10.2 su esistenza globale e confronto

Esercizi 10.3, 10.4, 10.6, 10.7, 10.8 su esistenza globale/blowup

Esercizio 10.5 (Soluzioni pari/dispari)

Esercizio 10.9 su esistenza globale e integrale primo

Proposizione 10.10 (Esistenza globale tramite integrale primo)

Esercizio 10.11

Esercizio 10.12 (Sistemi meccanici conservativi)

Esercizio 10.13 (Teorema dell'asintoto)

Esercizio 10.14

Lemma 10.15 (Lemma di Gronwall)

Esercizio 10.16

Svolgimento di alcuni esercizi

Lezione 11

Lemma 11.1 (Lemma di Gronwall, versione integrale)

Teorema 11.2 (Dipendenza continua dai dati iniziali)

Osservazione 11.3 (Il Teorema di dipendenza continua è ottimale)

Equazioni lineari.

Definizione 11.4 (Equazioni lineari del primo ordine in dimensione N)

Osservazione 11.5 (Funzioni a valori nello spazio delle matrici)

Esercizio 11.6 (In dimensione finita tutte le norme sono equivalenti)

Definizione 11.7 (Equazioni lineari omogenee/non omogenee)

Lemma 11.8 (Esistenza globale per equazioni lineari)

Equazioni omogenee.

Teorema 11.9 (Le soluzioni di un'equazione lineare omogenea formano uno spazio vettoriale)

Definizione 11.10 (Funzioni linearmente indipendenti)

Lemma 11.11 (Soluzioni linearmente indipendenti e valutazione)

Esempio.

Lezione 12

Definizione 12.1 (Matrice wronskiana, determinante wronskiano)

Lemma 12.2 (Teorema del wronskiano)

Osservazione 12.3 (Il determinante wronskiano è $C^1(I)$ e ha segno costante)

Lemma 12.4 (Equazione differenziale della matrice wronskiana)

Esercizio 12.5 (Equazione differenziale del determinante wronskiano e formula di Liouville)

Definizione 12.6 (Matrice risolvente)

Osservazione 12.7 (Problema di Cauchy della matrice risolvente)

Lemma 12.8 (La matrice risolvente dà la soluzione del problema di Cauchy)

Lemma 12.9 (Proprietà della matrice risolvente)

Equazioni non omogenee.

Teorema 12.10 (Le soluzioni di un'equazione lineare non omogenea formano uno spazio affine)

Osservazione oss:sol particolare (Soluzione particolare)
Proposizione 12.12 (Soluzioni dell'equazione non omogenea tramite variazione delle costanti)
Esercizio 12.13 (Matrice risolvente e variazione delle costanti)
Equazioni lineari a coefficienti costanti.
Lemma 12.14 (Soluzione dell'equazione $u' = Au$ tramite iterate di Picard)
Lemma 12.15 (Norme di matrici)
Definizione 12.16 (Esponenziale di matrice)
Osservazione 12.17 (Convergenza della serie esponenziale)

Lezione 13

Lemma 13.1 (La matrice esponenziale è la matrice risolvente di $u' = Au$)
Calcolo dell'esponenziale di matrice.
Matrice diagonale.
Matrice diagonalizzabile con autovalori reali.
Lemma 13.2 (Esponenziale e coniugato)
Matrice reale 2×2 con autovalori complessi coniugati.
Blocco di Jordan di autovalore reale.
Lemma 13.3 (Esponenziale della somma)
Blocco di Jordan reale con coppia di autovalori complessi coniugati.
Caso generale.
Teorema 13.4 (Forma canonica reale di Jordan)
Esercizio 13.5

Lezione 14

Proposizione 14.1 (Soluzione dell'equazione lineare non omogenea a coefficienti costanti tramite variazione delle costanti)
Osservazione 14.2 (Operazioni sul sistema di equazioni)
Esercizio 14.3 (esercizio svolto sui sistemi lineari)
Equazioni scalari di ordine più alto.
Definizione 14.4 (Equazioni lineari di ordine N)
Osservazione 14.5 (Equazione lineare scalare scritta come sistema lineare del primo ordine)
Lemma 14.6 (Spazio vettoriale delle soluzioni)
Lemma 14.7 (Funzioni linearmente indipendenti)
Lemma 14.8 (Soluzioni linearmente indipendenti)
Definizione 14.9 (Matrice wronskiana, determinante wronskiano)
Lemma 14.10 (Teorema del wronskiano per l'equazione $\mathcal{L}u = 0$)
Teorema 14.11 (Le soluzioni di un'equazione lineare non omogenea formano uno spazio affine)
Osservazione 14.12 (Soluzione particolare)
Metodo di variazione delle costanti
Esercizio 14.13 (variazione delle costanti)
Equazioni a coefficienti costanti.
Lemma 14.14 (Polinomio caratteristico della matrice A)
Esercizio 14.15 sul polinomio caratteristico della matrice.
Definizione 14.16 (Polinomio caratteristico dell'equazione differenziale)
Teorema 14.17 (Soluzioni delle equazioni lineari a coefficienti costanti)
Esempio.
Metodo di somiglianza per equazioni non omogenee.

Esercizi 14.18, 14.19 sul metodo di somiglianza.

Teorema 14.20 (Teorema di differenziabilità rispetto ai dati iniziali)

Lezione 15

Dimostrazione del Teorema 14.20 di dipendenza differenziabile dai dati iniziali

Osservazione 15.1 (Differenziabilità del flusso rispetto al dato iniziale)

Calcolo differenziale in spazi di Banach

Definizione 15.2 (Funzioni continue)

Definizione 15.3 (Norma operatoriale di funzioni lineari)

Osservazione 15.4 (Esempio di funzione lineare con $\|A\|_{\text{op}} = \infty$)

Osservazione 15.5 (Funzione lineare iniettiva e non suriettiva, e viceversa)

Esercizio 15.6 (Traslazione a destra/sinistra)

Esercizio 15.7 (Spazio delle successioni definitivamente nulle)

Esercizio 15.8 (Spazio dei polinomi)

Proposizione 15.9 (Continuità per funzioni lineari)

Definizione 15.10 (Spazio degli operatori lineari e continui)

Lemma 15.11 ($\mathcal{L}(X, Y)$ è di Banach)

Esercizio 15.12

Osservazione 15.13 (Compattezza della palla unitaria)

Esercizio 15.14 (Palla non compatta)

Lezione 16

Definizione 16.1 (Serie convergente e totalmente convergente)

Proposizione 16.2 (Completezza e convergenza totale delle serie)

Esercizio 16.3

Inverso di un operatore lineare e serie di Neumann

Lemma 16.4 (Norma operatoriale)

Proposizione 16.5 (Invertibilità e serie di Neumann)

Proposizione 16.6 (Inverso di $A + B$ per serie di Neumann)

Corollario 16.7 ($\text{Iso}(X, Y)$ è aperto)

Proposizione 16.8 (Continuità dell'operatore di inversione)

Definizione 16.9 (Differenziabilità)

Osservazione 16.10 (Casi $h = 0$, $h \neq 0$)

Osservazione 16.11 (Codominio del differenziale)

Definizione 16.12 (Funzioni C^1)

Esercizio ex:19.1011.3 (calcolo del differenziale)

Esercizio 16.14 (differenziale di funzione lineare)

Proposizione 16.15

Proposizione 16.16 (Inv è di classe C^1)

Definizione 16.17 (Derivate parziali)

Lemma 16.18 (La differenziabilità implica l'esistenza delle derivate parziali)

Esercizio 16.19

Teorema del valor medio

Proposizione 16.20 (Stima di Lipschitz senza Lagrange e senza integrale)

Lezione 17

Osservazione 17.1 (Derivata di funzione di variabile reale a valori in spazio vettoriale normato)

Dimostrazione della Proposizione 16.20

Teorema 17.2 (Teorema del valor medio in spazi vettoriali normati)

Osservazione 17.3 (Nel Teorema del valor medio basta la derivata direzionale)
Definizione 17.4 (Contrazione)
Teorema 17.5 (Lemma delle contrazioni di Banach-Caccioppoli)
Osservazione 17.6 (Lemma delle contrazioni in spazi di Banach)
Esercizio 17.7 (contrazioni)
Teorema 17.8 (Teorema della funzione implicita)
Osservazione 17.9 (Premessa: idea della dimostrazione)

Lezione 18

Conclusione della dimostrazione del Teorema della funzione implicita.
Derivate successive
Definizione 18.1 (Derivata seconda)
Definizione 18.2 (Funzioni bilineari)
Lemma 18.3 (Continuità per funzioni bilineari)
Esercizio 18.4
Lemma 18.5 (Spazio delle funzioni bilineari continue)
Esercizio 18.6
Lemma 18.7 (Isomorfismo di $\mathcal{L}(X, \mathcal{L}(X, Y))$ e $\mathcal{L}_2(X, Y)$)
Lemma 18.8 ($f''(p)$ è simmetrico)
Esempio
Definizione 18.9 (Derivata k -esima)
Definizione 18.10 (Funzioni k -lineari)
Lemma 18.11 (Continuità per funzioni k -lineari)
Lemma 18.12 (Spazio delle funzioni k -lineari continue)
Lemma 18.13 (Isomorfismo di $\mathcal{L}(X, \mathcal{L}_{k-1}(X, Y))$ e $\mathcal{L}_k(X, Y)$)
Lemma 18.14 ($f^{(k)}$ è simmetrica)
Lemma 18.15
Definizione 18.16 (Funzioni C^∞)

Lezione 19

Proposizione 19.1 (Inv è di classe C^∞)
Proposizione 19.2 (Ulteriore regolarità della funzione implicita)
Esercizi 19.3, 19.4, 19.5, 19.6
Serie di Fourier
Definizione 19.7 (Integrale di funzioni a valori complessi)
Osservazione 19.8 (Integrale e primitiva)
Osservazione 19.9 (Derivata e integrale degli esponenziali complessi)
Esercizio 19.10
Esercizio 19.11 (Formula di integrazione per parti per funzioni a valori complessi)
Osservazione 19.12 (Formule di uso frequente)
Definizione 19.13 (Polinomio trigonometrico, serie trigonometrica)
Osservazione 19.14 (Polinomio trigonometrico, serie trigonometrica in seni e coseni)
Esercizio 19.15
Lemma 19.16 (Proprietà di traslazione del periodo)
Esercizio 19.17
Osservazione 19.18 (Normalizzazione del periodo a 2π)
Definizione 19.19 (Coefficienti di Fourier, serie di Fourier)
Osservazione 19.20
Esercizio 19.21 (Coefficienti di Fourier rispetto a seni e coseni)

Esercizio 19.22 (esercizio svolto su calcolo di serie di Fourier)

Esercizi 19.23, 19.24

Esercizi 19.25, 19.26 (multipli del periodo)

Lemma 19.27 (Funzioni pari, dispari, reali)

Lemma 19.28 (Funzioni pari, dispari, reali con seni e coseni)

Esercizio 19.29

Lezione 20

Definizione 20.1 (Nucleo di Dirichlet)

Lemma 20.2 (Formula integrale per $S_N(f)$ con il nucleo di Dirichlet)

Lemma 20.3 (Proprietà del nucleo di Dirichlet)

Osservazione 20.4 (Media integrale pesata)

Lemma 20.5 (Formula integrale per $S_N(f)$ sul mezzo periodo con il nucleo di Dirichlet)

Lemma 20.6 (Lemma di Riemann-Lebesgue per funzioni Riemann-integrabili)

Osservazione 20.7 (Oscillazioni ad alta frequenza e cancellazioni)

Osservazione 20.8 (Lemma di Riemann-Lebesgue con integrazione per parti)

Corollario 20.9 (I coefficienti di Fourier tendono a zero)

Lemma 20.10 (Regolarità di $f(x)$ e decadimento di $\hat{f}(n)$)

Corollario 20.11 (Convergenza totale della serie di Fourier per f di classe C^2)

Lezione 21

Teorema 21.1 (Teorema di convergenza puntuale per serie di Fourier con condizione di Dini per l'integrale di Riemann)

Lemma 21.2 (Lemma di Riemann-Lebesgue per funzioni Riemann-integrabili in senso generalizzato)

Teorema 21.3 (Teorema di convergenza puntuale per serie di Fourier con condizione di Dini per l'integrale di Riemann generalizzato)

Proposizione 21.4 (Criterio di convergenza puntuale per serie di Fourier con condizioni di Hölder)

Proposizione 21.5 (Criterio di convergenza puntuale per serie di Fourier con limite destro/sinistro di f, f')

Teorema 21.6 (Teorema di localizzazione di Riemann)

Lemma 21.7 (Densità delle funzioni continue in quelle integrabili)

Osservazione 21.8 (Raccordo lineare come media integrale e convoluzione)

Lezione 22

Definizione 22.1 (Convoluzione di funzioni periodiche)

Lemma 22.2 (Proprietà della convoluzione)

Esercizio 22.3

Definizione 22.4 (Famiglia di buoni nuclei)

Teorema 22.5 (Teorema dei buoni nuclei)

Esercizio 22.6 (I nuclei di Dirichlet non sono buoni nuclei)

Definizione 22.7 (Medie alla Cesàro, convergenza alla Cesàro)

Esercizi 22.8, 22.9, 22.10 (convergenza alla Cesàro)

Definizione 22.11 (Nuclei di Fejer)

Lemma 22.12 (Proprietà dei nuclei di Fejer)

Lezione 23

Esercizio 23.1 (Coefficienti di Fourier dei nuclei di Dirichlet e di Fejer)

Lemma 23.2 (Nuclei di Fejer di periodo T)

Teorema 23.3 (Sommabilità alla Cesàro per le serie di Fourier)
 Definizione 23.4 (Somma N -esima di Cesàro di una serie di Fourier)
 Lemma 23.5 (Coefficienti di Fourier della somma di Cesàro)
 Osservazione 23.6 (Convoluzione con i nuclei e moltiplicatori di Fourier)
 Lemma 23.7 (Funzioni a coefficienti di Fourier tutti nulli)
 Proposizione 23.8 (Densità dei polinomi trigonometrici nelle funzioni periodiche continue)
 Esercizio 23.9 (Teorema di Weierstrass di approssimazione polinomiale)
 Convergenza in media quadratica
 Definizione 23.10 (Prodotto scalare semidefinito positivo per spazi vettoriali su \mathbb{C})
 Definizione 23.11 (Seminorma per spazi vettoriali su \mathbb{C})
 Definizione 23.12 (Prodotto scalare definito positivo e norma per spazi vettoriali su \mathbb{C})
 Definizione 23.13 (Spazi di Hilbert su \mathbb{C})
 Esercizio 23.14 (Lo spazio vettoriale \mathbb{C}^n)
 Esercizio 23.15 (Prodotto scalare per funzioni periodiche)
 Definizione 23.16 (Vettori ortogonali)
 Proposizione 23.17 (Proprietà degli spazi vettoriali su \mathbb{C} con prodotto scalare semidefinito positivo)
 Osservazione 23.18
 Definizione 23.19 (L'insieme $\ell^2(\mathbb{Z}, \mathbb{C})$)
 Lemma 23.20 ($\ell^2(\mathbb{Z}, \mathbb{C})$ è uno spazio vettoriale)
 Lemma 23.21 (Convergenza della serie $\sum a_n \bar{b}_n$)
 Definizione 23.22 (Prodotto scalare in ℓ^2 e norma indotta)
 Lemma 23.23 (ℓ^2 è spazio di Hilbert)
 Osservazione 23.24 (Lo spazio delle funzioni Riemann-integrabili non è completo)

Lezione 24

Definizione 24.1 (Insieme ortonormale)
 Definizione 24.2 (Spazio delle funzioni periodiche Riemann-integrabili)
 Lemma 24.3 (Famiglia ortonormale degli esponenziali)
 Lemma 24.4 (Seminorma dei polinomi trigonometrici)
 Lemma 24.5 (Lemma di ortogonalità)
 Lemma 24.6 (Lemma di migliore approssimazione)
 Teorema 24.7 (Teorema di convergenza in media quadratica per serie di Fourier)
 Corollario 24.8 (Identità di Parseval)
 Esercizio 24.9 (Somma della serie armonica generalizzata di esponente 2)
 Lemma 24.10 (Disuguaglianza di Bessel)
 Lemma 24.11 (Corrispondenza tra prodotto scalare in \mathcal{R}_T e in ℓ^2)
 Definizione 24.12 (Convoluzione di successioni)
 Lemma 24.13 (La successione $a * b$ è ben posta e limitata)
 Lemma 24.14 (Coefficienti di Fourier del prodotto di funzioni periodiche)

Appendice A. Convergenza delle serie bilatere

Le serie reali a termini non negativi hanno limite
 Esercizio 25.1
 Definizione 25.2 (Limiti di somme finite)
 Esercizio 25.3
 Criterio del confronto
 La convergenza assoluta implica la convergenza

Esercizio 25.4

Senza la convergenza assoluta le tre definizioni di limite delle somme finite non sono equivalenti

Esercizio 25.5

Serie assolutamente convergenti di numeri complessi

Appendice B. Altre dimostrazioni

Dimostrazione del Teorema 6.6 di esistenza di Peano con metodo di Eulero

Dimostrazione diretta del Corollario 8.1

Riferimenti bibliografici

- [1] E. Acerbi, L. Modica, S. Spagnolo, *Problemi scelti di Analisi Matematica II*, Liguori Editore.
- [2] A. Ambrosetti, G. Prodi, *A primer of nonlinear analysis*, Cambridge University Press, 1993.
- [3] P. Baldi, *Equazioni differenziali*, dispense per il corso di Analisi 2 a Ingegneria (disponibile nella mia pagina docenti nella cartella del materiale didattico di Sistemi Dinamici <https://www.docenti.unina.it/pietro.baldi>)
- [4] G.C. Barozzi, *Matematica per l'Ingegneria dell'Informazione*, Zanichelli, 2001.
- [5] M. Berti, *Note del corso di Sistemi Dinamici*, 16 Dicembre 2011 versione preliminare (disponibile nella mia pagina docenti nella cartella del materiale didattico di Sistemi Dinamici <https://www.docenti.unina.it/pietro.baldi>)
- [6] V. Coti Zelati, dispense per il corso di Sistemi Dinamici, capitoli 1-2 (disponibile nella mia pagina docenti nella cartella del materiale didattico di Sistemi Dinamici <https://www.docenti.unina.it/pietro.baldi>)
- [7] G. De Marco, *Analisi Due/2*, Zanichelli.
- [8] B. Demidovič, *Esercizi e problemi di analisi matematica*, Editori Riuniti.
- [9] N. Fusco, P. Marcellini, C. Sbordone, *Analisi Matematica due*, Liguori Editore.
- [10] G. Gentile, *Introduzione ai Sistemi Dinamici*, <http://www.mat.uniroma3.it/users/gentile/2014-2015/FM410/testo.html>
- [11] P. Hartman, *Ordinary Differential Equations*, SIAM 2002.
- [12] M. Hirsch, S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, 1974.
- [13] E. Stein, R. Shakarchi, *Fourier Analysis: An Introduction*. Princeton Lectures in Analysis, volume I. Princeton University Press, 2003.