

Silvio Fanzon
Annalisa Malusa

APPUNTI di
**EQUAZIONI DIFFERENZIALI
ORDINARIE**

Indice

1 Motivazioni e concetti di base	1
1.1 Modelli ecologici unidimensionali	1
1.2 Modello predatore preda (Lotka-Volterra)	4
1.3 Funi sospese	8
1.4 Il pendolo semplice	10
1.5 Conclusioni	13
2 Prime definizioni e richiami	15
2.1 Equazioni differenziali e problema di Cauchy	15
2.2 Notazioni ed ipotesi	20
2.3 Esistenza ed unicità locale in ipotesi di Lipschitz	24
3 Alcuni metodi di risoluzione	29
3.1 Variabili separabili	29
3.2 1-forme ed equazioni differenziali	34
3.2.1 Richiami sulle 1-forme differenziali	37
3.2.2 Curve soluzione	38
3.2.3 Metodo del fattore integrante	43
3.2.4 Integrali primi	46
3.2.5 Equazioni lineari del primo ordine non omogenee	50
3.3 Equazioni di Bernoulli	52
3.4 Equazioni omogenee	53
4 Teoria generale di base	55
4.1 Il Lemma di Gronwall	55
4.2 Soluzioni approssimate	57
4.3 Funzioni lipschitziane e teorema di unicità locale	59
4.4 Il teorema di Peano	61
4.5 Prolungamenti	67

4.6 Dipendenza dai dati iniziali	76
5 Sistemi di equazioni lineari	79
5.1 Sistemi lineari omogenei	80
5.2 Sistemi lineari non omogenei	92
6 Matrice fondamentale esponenziale	97
6.1 Sistemi lineari a coefficienti costanti	97
6.2 Notazioni e richiami di Algebra Lineare	98
6.3 Operatori lineari e limitati in spazi di Banach	105
6.4 Operatore esponenziale	111
6.5 Matrice fondamentale esponenziale	116
7 Soluzioni di sistemi lineari omogenei autonomi	121
7.1 Esponenziale di matrici diagonalizzabili in \mathbb{C}	121
7.2 Forma canonica di Jordan	132
7.3 Rappresentazione esplicita delle soluzioni	143
8 Analisi qualitativa dei sistemi lineari omogenei autonomi	147
8.1 Proprietà generali delle soluzioni	147
8.2 Studio qualitativo dei sistemi planari	152
9 Equazioni differenziali lineari di ordine superiore	159
9.1 Rappresentazione delle soluzioni	159
9.2 Equazioni del secondo ordine a coefficienti costanti	165
10 Analisi locale di sistemi non lineari autonomi	177
10.1 Il teorema di rettificabilità locale	178
10.2 Stabilità	180

Capitolo 1

Motivazioni e concetti di base

Un'equazione differenziale ordinaria è una relazione

$$F(t, x, x', x'', \dots, x^{(k)}) = 0.$$

che lega una funzione $x : I \rightarrow \mathbb{R}^n$, $I \subseteq \mathbb{R}$ intervallo, ad alcune sue derivate.

In questo capitolo introduttivo vogliamo proporre alcuni esempi che

- ci mostrino perché vale la pena di strutturare una teoria per questo tipo di equazioni,
- ci dicano quali siano le questioni da investigare,
- ci diano qualche idea su come procedere.

1.1 Modelli ecologici unidimensionali

Vogliamo descrivere la dinamica di una popolazione, con particolare interesse all'evoluzione nel tempo della sua numerosità. Se indichiamo con $p(t)$ la popolazione al tempo t , il tasso di crescita è determinato da $p'(t)/p(t)$. Possiamo ragionevolmente supporre che il tasso di crescita dipenda solamente dal tempo e dalla numerosità della popolazione presente, per cui

$$\frac{p'(t)}{p(t)} = r(t, p(t)).$$

Supporremo sempre che il sistema sia chiuso (nessun individuo può migrare altrove o introdursi nel sistema dall'esterno) in modo tale che $r(t, p) = g(t, p) - s(t, p)$ dove g ed s sono funzioni che descrivono, rispettivamente, il tasso di crescita e il tasso di morte della popolazione.

Una volta note (sperimentalmente) le funzioni g ed s , e nota la numero-
sità della popolazione p_0 ad un fissato tempo iniziale t_0 , per determinarne
l'evoluzione, almeno in un intervallo temporale $[t_0, T)$ fissato, dobbiamo saper
risolvere il problema

$$\begin{cases} p'(t) = r(t, p(t))p(t), & \text{per } t \in [t_0, T), \\ p(t_0) = p_0 \geq 0. \end{cases} \quad (1.1)$$

Stiamo quindi cercando una funzione p derivabile nell'intervallo $[t_0, T)$ e tale
che

- i) soddisfi l'equazione differenziale in (1.1) per ogni $t \in [t_0, T)$,
- ii) soddisfi la condizione iniziale.

Primo caso: crescita costante. In prima battuta, vediamo che tipo di
evoluzione può avere la popolazione nel caso in cui il tasso di crescita sia
costante, ossia $r(t, p) = \alpha \in \mathbb{R}$. L'equazione (1.1) diventa quindi

$$\begin{cases} p'(t) = \alpha p(t), & \text{per } t \in [t_0, T), \\ p(t_0) = p_0 \geq 0. \end{cases}$$

Se $p_0 = 0$ (nessun individuo al tempo iniziale), si ottiene la soluzione costante
 $p(t) = 0$ (ovviamente non si possono generare individui dal nulla). Ragione-
volmente, se $p_0 > 0$, la soluzione corrispondente, almeno per un certo tempo
deve essere strettamente positiva (ricordiamo che stiamo richiedendo che la
funzione $p(t)$ sia derivabile, quindi sicuramente continua). Quindi possiamo
determinate queste soluzioni per separazione di variabili (studieremo in detta-
glio questo metodo di risoluzione nel Paragrafo 3.1): dividiamo ambo i membri
dell'equazione differenziale per p ed integriamo tra t_0 e t

$$\int_{t_0}^t \frac{p'(s)}{p(s)} ds = \int_{t_0}^t \alpha ds.$$

Nel primo integrale operiamo il cambio di variabile $z = p(t)$

$$\int_{p_0}^{p(t)} \frac{1}{z} dz = \alpha(t - t_0)$$

e calcoliamo gli integrali, ottenendo

$$\log \left| \frac{p(t)}{p_0} \right| = \alpha(t - t_0).$$

Ne concludiamo che il problema ha come unica soluzione la funzione

$$p(t) = p_0 e^{\alpha(t-t_0)}, \quad p_0 \in \mathbb{R}.$$

Quindi, se $\alpha > 0$, la numerosità della popolazione diverge per $t \rightarrow +\infty$, mentre per $\alpha < 0$ la popolazione si estinguerebbe per $t \rightarrow +\infty$.

Queste previsioni sono chiaramente grossolane e ci dicono che, come potevamo aspettarci, il modello è troppo poco raffinato e che dobbiamo fare ipotesi più ragionevoli sulla funzione r se vogliamo avere previsioni più realistiche.

Secondo caso: modello logistico. Per ovviare al precedente problema di esplosione della popolazione, possiamo, per esempio, introdurre una soglia critica $\xi > 0$ tale che, se la popolazione p supera ξ , il tasso di crescita diventa negativo (a causa di assenza di cibo, spazio, etc.). La funzione r più semplice che modellizzi questa condizione è

$$r(t, p) = \beta(\xi - p), \quad \beta, \xi > 0.$$

Il problema che determina l'evoluzione della popolazione è allora

$$\begin{cases} p' = \beta(\xi - p)p = \alpha p - \beta p^2 \\ p(t_0) = p_0 \geq 0. \end{cases}$$

dove si è posto $\alpha = \beta\xi$. Per determinare le soluzioni dell'equazione differenziale useremo nuovamente il metodo di separazione delle variabili. Questa volta l'equazione ha soluzioni costanti $p \equiv 0$ e $p \equiv \xi$ (dette anche equilibri) che, per motivi che saranno chiari più avanti, non sono valicabili dalle altre soluzioni. In altri termini se $0 < p_0 < \xi$, allora la soluzione corrispondente avrà valori tra 0 e ξ per tutti i tempi e, in maniera analoga, se $p_0 > \xi$ allora la soluzione corrispondente avrà sempre valori superiori a ξ .

Una volta fatte queste osservazioni preliminari, procediamo alla separazione delle variabili ed integriamo tra t_0 e t , ottenendo

$$\int_{t_0}^t \frac{p'(s)}{p(s)(\alpha - \beta p(s))} ds = t - t_0.$$

Cambiando variabile e scomponendo la frazione in fratti semplici si ottiene

$$\int_{p_0}^{p(t)} \left(\frac{1}{\alpha p} + \frac{\beta}{\alpha(\alpha - \beta p)} \right) dp = t - t_0$$

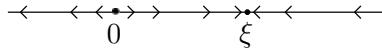


Figura 1.1: Dinamica del modello logistico

e, a conti fatti,

$$\frac{1}{\alpha} \log \left| \frac{p(t)}{p_0} \right| - \frac{1}{\alpha} \log \left| \frac{\alpha - \beta p(t)}{\alpha - \beta p_0} \right| = t - t_0.$$

Infine, usando le proprietà elementari dei logaritmi per esplicitare $p(t)$, otteniamo che l'evoluzione della popolazione è data da

$$p(t) = \frac{\alpha p_0}{\beta p_0 + (\alpha - \beta p_0) e^{-\alpha(t-t_0)}}.$$

Notiamo che, se partiamo da una qualsiasi popolazione iniziale p_0 diversa da 0 o da ξ , si ha

$$\lim_{t \rightarrow +\infty} p(t) = \frac{\alpha}{\beta} = \xi,$$

cioè l'evoluzione tende sempre all'equilibrio non banale per tempi grandi.

La dinamica della popolazione p può essere visualizzata in quello che prende il nome di spazio delle fasi: ci si dimentica della dipendenza dal tempo di p e si associa ad ogni valore di $p \in \mathbb{R}$, il campo di velocità $p' = \alpha p - \beta p^2$. Rappresentando graficamente il campo (Figura 1.1) si può notare che le soluzioni non costanti tendono ad allontanarsi da 0 e ad avvicinarsi a ξ . Si dirà che 0 è un equilibrio instabile, mentre ξ è un equilibrio stabile.

1.2 Modello predatore preda (Lotka-Volterra)

Introduciamo ora un modello per l'evoluzione di due popolazioni che interagiscono mutuamente. Indichiamo con $x(t)$ la popolazione delle prede e con $y(t)$ la popolazione dei predatori. La dinamica viene descritta dal sistema di due equazioni differenziali del primo ordine accoppiate in due incognite

$$\begin{cases} \dot{x} = r_1(t, x, y) x, \\ \dot{y} = r_2(t, x, y) y. \end{cases}$$

Supporremo che la preda sia l'unico mezzo di sussistenza del predatore e che la preda abbia a disposizione una quantità illimitata di cibo.

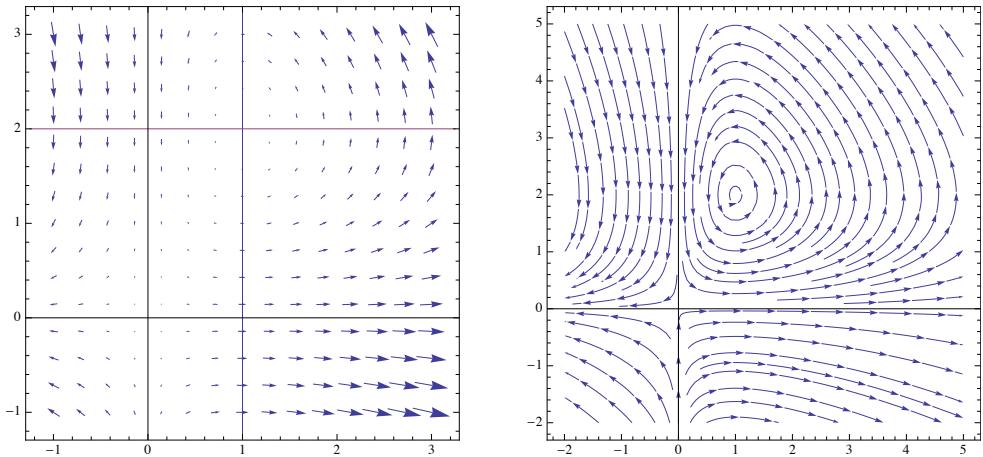


Figura 1.2: Campo di velocità e ritratto di fase del sistema di Lotka-Volterra

Primo caso: crescita costante. In questo primo approccio supponiamo che le prede, in assenza di predatori, abbiano un tasso di crescita costante α . La presenza dei predatori comporta una riduzione delle prede ragionevolmente proporzionale al numero xy di incontri tra prede e predatori. Indicheremo con β il coefficiente di proporzionalità. Viceversa, i predatori hanno un tasso di decadimento costante γ e sarebbero destinati ad estinguersi se non ci fossero le prede. Il loro numero invece aumenta proporzionalmente al numero di incontri. Abbiamo quindi il sistema

$$\begin{cases} \dot{x} = \alpha x - \beta xy, & \alpha, \beta > 0, \\ \dot{y} = -\gamma y + \delta xy, & \gamma, \delta > 0 \\ (x(t_0), y(t_0)) = (x_0, y_0), & x_0, y_0 > 0 \end{cases} \quad (1.2)$$

Se introduciamo li vettore $p(t) = (x(t), y(t))$ e il campo $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f(x, y) = ((\alpha - \beta y)x, (-\gamma + \delta x)y)$ il sistema si scrive in forma compatta come

$$\begin{cases} p' = f(p) \\ p(t_0) = p_0 \end{cases}$$

Invece di tentare di determinare le soluzioni in forma esplicita come nei casi precedenti, cerchiamo di fare direttamente un'analisi qualitativa delle dinamiche che partono da un punto del primo quadrante (le uniche interessanti per il modello).

Innanzi tutto osserviamo che ci sono due soluzioni stazionarie $(0, 0)$ e $(\gamma/\delta, \alpha/\beta)$. Inoltre le traiettorie che partono da dati del tipo $(x_0, 0)$, $x_0 > 0$

(assenza iniziale di predatori) rimangono sull'asse x per tutti i tempi e divergono (le prede crescono esponenzialmente), mentre le traiettorie che partono da dati del tipo $(0, y_0)$, $y_0 > 0$ (assenza iniziale di prede) rimangono sull'asse delle y e tendono a zero (estinzione in tempo infinito delle prede). Per la descrizione qualitativa delle altre traiettorie procediamo nel modo seguente: siamo in grado di individuare, anche se in maniera grossolana, le direzioni del campo di velocità (\dot{x}, \dot{y}) studiando il segno delle componenti (si veda la Figura 1.2 a sinistra). Una volta determinato il campo di velocità, le traiettorie saranno le curve che hanno per tangente quel campo. In particolare, lungo i punti della retta $x = \gamma/\delta$ il campo di velocità risulta essere orizzontale, mentre lungo la retta $y = \alpha/\beta$ il campo risulta essere verticale. Inoltre

$$\begin{cases} \dot{x} < 0, \dot{y} > 0 & \text{se } x > \gamma/\delta, y > \alpha/\beta, \\ \dot{x} < 0, \dot{y} < 0 & \text{se } x < \gamma/\delta, y > \alpha/\beta, \\ \dot{x} > 0, \dot{y} < 0 & \text{se } x < \gamma/\delta, y < \alpha/\beta, \\ \dot{x} > 0, \dot{y} > 0 & \text{se } x > \gamma/\delta, y < \alpha/\beta. \end{cases}$$

Da questa analisi si riesce a concludere che le traiettorie del primo quadrante ruotano intorno al punto stazionario $(\gamma/\delta, \alpha/\beta)$, ma non si riesce a capire se siano delle traiettorie chiuse o delle spirali.

Per capire quale sia la situazione dobbiamo sfruttare una proprietà aggiuntiva di questo sistema (che sarà studiata in dettaglio nel Paragrafo 3.2.4), cioè il fatto che ogni soluzione di (1.2) è confinata su una curva di livello della funzione

$$U(x, y) = \gamma x - \beta y - \delta \log x + \alpha \log y \quad (1.3)$$

(la funzione U prende il nome di integrale primo del sistema di equazioni differenziali). Infatti per la formula di derivazione delle funzioni composte si ha

$$\begin{aligned} \frac{d}{dt} U(x(t), y(t)) &= \langle \nabla U(x(t), y(t)), (\dot{x}(t), \dot{y}(t)) \rangle \\ &= \left(\gamma - \frac{\delta}{x} \right) \dot{x} + \left(\frac{\alpha}{y} - \beta \right) \dot{y} \\ &= \left(\gamma - \frac{\delta}{x} \right) (\alpha - \beta y)x + \left(\frac{\alpha}{y} - \beta \right) (-\gamma + \delta x)y = 0. \end{aligned}$$

Poiché le curve di livello della funzione U nel primo quadrante sono curve chiuse, si può concludere che l'evoluzione è di tipo periodico. Il ritratto di fase in Figura 1.2 a destra descrive le traiettorie della soluzione.

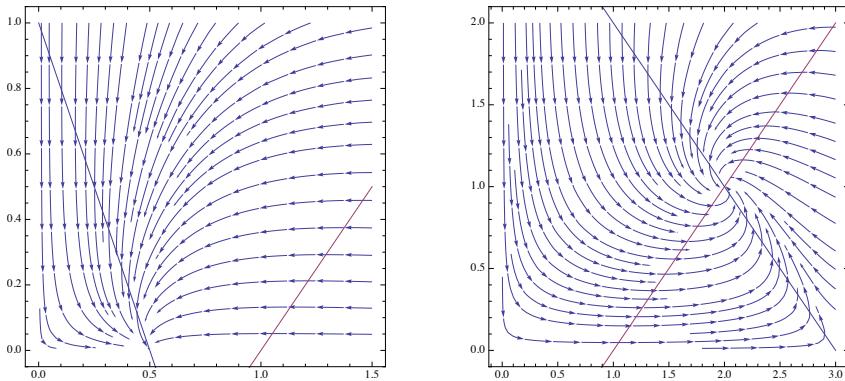


Figura 1.3: Ritratti di fase del sistema di Lotka-Volterra con attrito sociale

In conclusione, anche se non abbiamo scritto esplicitamente l'espressione delle soluzioni, gli argomenti precedenti ci hanno fornito informazioni piuttosto dettagliate sulla dinamica: l'evoluzione è ciclica e il sistema ha un andamento (globalmente) instabile: a piccole variazioni dei dati iniziali corrispondono drastiche variazioni della dinamica per tempi lunghi.

Secondo caso: crescita limitata. Come variante del modello precedente, introduciamo un termine di attrito sociale, ovvero un termine che tenga conto del fatto che troppi elementi di una stessa specie si danneggiano. Come nel caso del modello logistico, un modo semplice di tenere conto di questo fatto è aggiungere nelle equazioni un termine di decadimento proporzionale al quadrato della numerosità della popolazione:

$$\begin{cases} \dot{x} = \alpha x - \beta xy - \lambda x^2, & \alpha, \beta, \lambda > 0, \\ \dot{y} = -\gamma y + \delta xy - \mu y^2, & \gamma, \delta, \mu > 0. \end{cases}$$

Il ritratto di fase delle soluzioni dipende dalla posizione reciproca delle due rette di equazione $\alpha - \beta y - \lambda x = 0$ (lungo la quale il campo è puramente verticale) e $\delta x - \mu y - \gamma = 0$ (lungo la quale il campo è puramente orizzontale). Se le due rette non si intersecano nel primo quadrante, il sistema ha due punti di equilibrio: $(0, 0)$ e $(\alpha/\lambda, 0)$. Le traiettorie sull'asse y entrano nell'origine, mentre quelle sull'asse delle x escono dall'origine ed entrano in $(\alpha/\lambda, 0)$. Diremo che $(0, 0)$ è un equilibrio instabile, mentre $(\alpha/\lambda, 0)$ è un equilibrio stabile e attrattore. Se ne conclude che, in questo caso, i predatori si estinguono sempre, seppur in tempo infinito. In presenza di prede al tempo iniziale, la dinamica tende sempre all'equilibrio $(\alpha/\lambda, 0)$ (si veda la Figura 1.3 a sinistra).

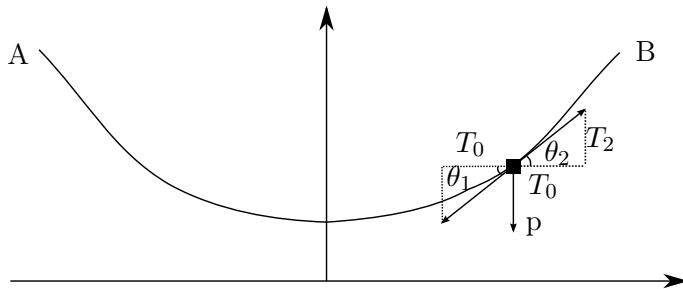


Figura 1.4: Fune sospesa

La situazione è significativamente diversa quando i parametri sono tali da far intersecare le due rette all'interno del primo quadrante. In questo caso il secondo equilibrio del sistema coincide con il punto di intersezione tra le rette e, a conti fatti, le dinamiche che partono con una numerosità di prede strettamente positiva (quindi fuori dall'asse y) risultano essere delle spirali che si avvolgono intorno a tale equilibrio. Quindi in questo caso non abbiamo un'evoluzione ciclica, ma una che tende, con un'alternanza di aumento di predatori e prede, all'equilibrio (si veda la Figura 1.3 a destra).

1.3 Funi sospese

Consideriamo una fune sottile, inestensibile e perfettamente flessibile, appesa nei punti A e B . Supponiamo che la fune sia soggetta solamente alla sua forza peso. Per semplicità, supponiamo che i punti A e B si trovino alla stessa quota. Ci chiediamo quale forma assuma il profilo della fune. Per motivi evidenti, è ragionevole supporre che il profilo della fune sia il grafico di una funzione di una variabile reale $y(x)$. Per ragioni fisiche è anche naturale aspettarsi che il profilo della fune sia simmetrico rispetto all'asse verticale passante per il punto medio tra A e B . Scriviamo una equazione per y facendo un bilancio delle forze che agiscono su un elemento infinitesimale della fune (Figura 1.4). Introduciamo un sistema di riferimento con origine sull'asse di simmetria. Se consideriamo un elemento di fune compreso tra i punti del piano $(x, y(x))$ e $(x+h, y(x+h))$, agli estremi di esso agisce una tensione con pari intensità orizzontale T_0 e con componenti verticali date rispettivamente da T_1 e T_2 . Indicando con θ_1 e θ_2 gli angoli che le tensioni formano con l'asse orizzontale, si ha

$$\frac{T_1}{T_0} = \tan \theta_1 = y'(x), \quad \frac{T_2}{T_0} = \tan \theta_2 = y'(x+h)$$

e quindi, all'equilibrio si ha il bilancio

$$T_0(y'(x+h) - y'(x)) = P$$

dove P indica il peso della fune. L'ipotesi di Galileo è che la configurazione di equilibrio fosse un arco di parabola. In generale questo è falso e vale solamente nel caso di funi molto tese. La giustificazione rigorosa di questa affermazione è basata sullo studio di un'equazione differenziale ordinaria.

Primo caso: funi molto tese. Quando una fune sostiene un grande carico uniforme lungo tutta la sua lunghezza (ad esempio nel caso dei cavi che sostengono l'impalcato dei ponti sospesi) è ragionevole supporre che l'unica componente rilevante della forza peso sia quella orizzontale, cioè che il peso dell'elemento infinitesimo sia $P = ph$. In tal modo il bilancio dell'energia restituisce la condizione

$$\frac{y'(x+h) - y'(x)}{h} = \frac{p}{T_0}$$

e quindi, passando al limite per $h \rightarrow 0$, si trova che il profilo della fune $y(x)$ deve soddisfare l'equazione differenziale ordinaria

$$y'' = \frac{p}{T_0},$$

con condizioni ai limiti $y(-a) = y(a) = H$. In questo caso quindi l'ipotesi di Galileo risulta corretta: il profilo della fune è un arco di parabola (basta procedere ad una doppia integrazione).

Secondo caso: funi poco tese. Questo è il caso di una fune che sostenga solamente il proprio peso. Dunque non possiamo fare nessuna approssimazione e, se indichiamo con p la densità di massa, otteniamo che il peso dell'elemento di fune è dato da

$$P = p dl = p \sqrt{1 + |y'(x)|^2} h.$$

La condizione di equilibrio diventa

$$\frac{y'(x+h) - y'(x)}{h} = \frac{p}{T_0} \sqrt{1 + |y'(x)|^2},$$

e quindi, passando al limite per $h \rightarrow 0$, si trova che, questa volta, il profilo della fune $y(x)$ deve soddisfare l'equazione differenziale ordinaria

$$y'' = \frac{p}{T_0} \sqrt{1 + |y'(x)|^2}$$

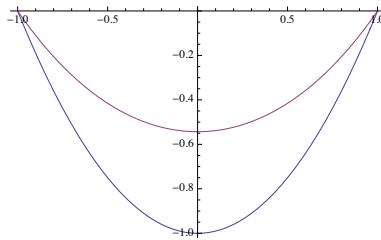


Figura 1.5: Profilo parabolico e catenaria

con condizioni ai limiti $y(-a) = y(a) = H$. Con la sostituzione $z = y'$ otteniamo una equazione differenziale del primo ordine

$$z' = \frac{p}{T_0} \sqrt{1 + z^2}.$$

Separiamo le variabili e integriamo:

$$\begin{aligned} \frac{z'(x)}{\sqrt{1 + z(x)^2}} &= \frac{p}{T_0} \implies \int \frac{1}{\sqrt{1 + z^2}} dz = \frac{p}{T_0} x + \alpha \implies \\ \text{settsinh } z(x) &= \frac{p}{T_0} x + \alpha \implies z(x) = \sinh \left(\frac{p}{T_0} x + \alpha \right). \end{aligned}$$

Ricordando che abbiamo posto $z = y'$, con un’ulteriore integrazione concludiamo che il profilo della fune è il grafico della funzione

$$y(x) = \cosh \left(\frac{p}{T_0} x + \alpha \right) + \beta$$

in cui le costanti α e β vengono determinate tramite le condizioni ai limiti. Quindi in questo caso il profilo è il grafico della funzione coseno iperbolico, che, per questa sua proprietà, prende il nome di curva catenaria.

1.4 Il pendolo semplice

Il pendolo semplice è formato da una massa m appesa ad un filo di lunghezza l fissato, all’altra estremità, in un perno. Per ottenere un modello semplice, supporremo che il filo sia inestensibile, sempre in tensione e di peso trascurabile. Supporremo inoltre assenza di attrito. Introduciamo un sistema di riferimento con origine nel perno del pendolo e asse y lungo la verticale orientato verso l’alto (si veda Figura 1.6).

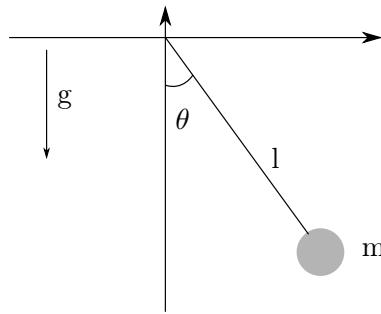


Figura 1.6: Pendolo

Come si può osservare sperimentalmente, il moto è piano e completamente descritto dall'equazione del moto $F = ma$ scritta in termini dell'angolo θ che il filo forma con l'asse delle y :

$$ml\ddot{\theta} = -mg \sin \theta. \quad (1.4)$$

Questa equazione, come tutte quelle che descrivono dinamiche in campi di forze conservativi, rientra in una importante classe di equazioni differenziali ordinarie a cui si può associare una energia che si conserva lungo le traiettorie del moto (sistemi conservativi, appunto). Nel nostro caso il campo di forze è

$$f(\theta) = -\frac{g}{l} \sin \theta$$

e un suo potenziale (cioè una funzione U tale che $f = -\nabla U$) è

$$U(\theta) = -\frac{g}{l} \cos \theta.$$

L'energia in questo caso è

$$E(\theta, \dot{\theta}) = \frac{1}{2}\dot{\theta}^2 - \frac{g}{l} \cos \theta,$$

Infatti si ha

$$\frac{d}{dt} E(\theta(t), \dot{\theta}) = \dot{\theta}\ddot{\theta} + \frac{g}{l} \sin \theta \dot{\theta} = \dot{\theta} \left(\ddot{\theta} + \frac{g}{l} \sin \theta \right) = 0,$$

quindi E si conserva lungo le soluzioni di (1.4). Osserviamo che E coincide, a meno di una costante moltiplicativa, con l'energia meccanica del sistema, data da

$$\frac{1}{2}m(l\dot{\theta})^2 - mgl \cos \theta,$$

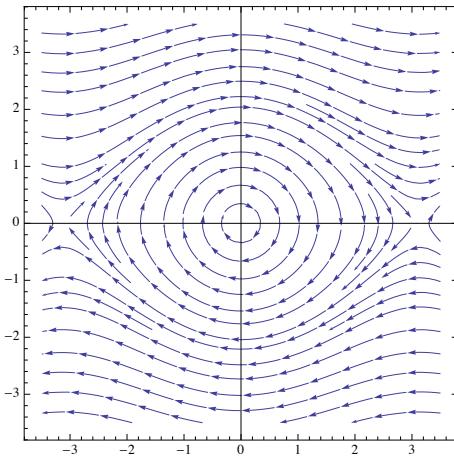


Figura 1.7: Ritratto di fase del pendolo

quindi quanto appena mostrato corrisponde esattamente al principio di conservazione dell'energia in meccanica classica.

Se $\theta(t)$ è soluzione soddisfa quindi l'equazione

$$E(\theta(t), \dot{\theta}) = \frac{1}{2}\dot{\theta}^2(t) - \frac{g}{l} \cos \theta = \text{costante}$$

cioè la soluzione è contenuta negli insiemi di livello del potenziale. Possiamo quindi ottenere informazioni qualitative sul comportamento delle soluzioni di (1.4) nel piano delle fasi $(\theta, \dot{\theta})$ analizzando le curve di livello dell'energia $E = E(\theta, \dot{\theta})$ (si veda Figura 1.7). Il sistema ha traiettorie estremamente differenti al variare del livello di energia che si considera.

- Le traiettorie corrispondenti al valore minimo dell'energia (pari a $-g/l$ e assunto per $(\theta, \dot{\theta}) = (2k\pi, 0)$) sono equilibri stabili, mentre le traiettorie corrispondenti al valore massimo dell'energia (pari a g/l e assunto per $(\theta, \dot{\theta}) = (\pi + 2k\pi, 0)$) sono equilibri instabili.
- Per valori bassi dell'energia abbiamo traiettorie periodiche (oscillazioni del pendolo) con centro in un equilibrio stabile e con periodo che dipende dal livello energetico considerato.
- Esiste un valore critico dell'energia tale per cui le traiettorie smettono di essere periodiche ed entrano in tempo infinito negli equilibri instabili (oscillazioni che tendono alla posizione verticale in tempo infinito).

- Superata la soglia critica le traiettorie non invertono più la loro direzione (rotazioni complete del pendolo).

Piccole oscillazioni Se l'oscillazione è piccola, poiché $\sin \theta = \theta + o(\theta)$, è ragionevole supporre che l'equazione

$$\ddot{\theta} + \omega^2 \theta = 0 \quad \omega = \sqrt{\frac{g}{l}}. \quad (1.5)$$

descriva il moto del pendolo. Questa equazione (soddisfatta anche dal moto dell'oscillatore libero, si veda (2.3)) è il prototipo di equazione lineare del secondo ordine e sarà presa spesso in considerazione nel seguito. Mostreremo che, in questo caso semplificato, si possono determinare esplicitamente le soluzioni e che hanno la forma

$$\theta(t) = A \cos(\omega t) + B \sin(\omega t)$$

al variare di $A, B \in \mathbb{R}$. Abbiamo quindi sempre oscillazioni periodiche la cui ampiezza dipende dalle costanti A e B che vengono fissate una volta assegnate la posizione e velocità iniziali.

1.5 Conclusioni

Pur avendo riportato solo pochissimi esempi di problemi formulabili in termini di equazioni differenziali ordinarie, dovrebbe comunque essersi chiarito quali sono le questioni più interessanti.

Buona posizione Un'equazione differenziale ordinaria gode della proprietà di buona posizione se è possibile garantire l'esistenza e l'unicità della soluzione corrispondente a un opportuno dato iniziale assegnato e se, inoltre, si è in grado di mostrare che tali soluzioni dipendono con continuità dalle condizioni iniziali. Dal punto di vista applicativo queste sono in genere le richieste minimali che devono essere soddisfatte da un modello matematico. Queste questioni saranno affrontate prima nel Capitolo 2, in cui si ricorderanno i risultati tipicamente già considerati nei corsi di base di Analisi Matematica, e, in maniera più completa e generale, nel Capitolo 4.

Metodi risolutivi Chiaramente, avere una rappresentazione esplicita delle soluzioni di un'equazione differenziale garantisce una conoscenza completa

della dinamica. In realtà i casi in cui le soluzioni possono essere definite esplicitamente non sono moltissimi. Abbiamo visto che in alcuni casi (ad esempio, il modello di Lotka–Volterra o il problema del pendolo semplice) si riesce ad ottenere un’informazione implicita sulle soluzioni che, in ogni caso, si può ritenere soddisfacente. Alcuni metodi risolutivi saranno descritti nel Capitolo 3. Una classe particolarmente interessante di sistemi di equazioni differenziali, quelli lineari del primo ordine, sarà studiata in dettaglio nei Capitoli 5, 6, 7 mentre le equazioni differenziali scalari di ordine superiore al primo saranno analizzati nel Capitolo 9.

Analisi Qualitativa In assenza di rappresentazioni, esplicite o implicite, delle soluzioni di un’equazione differenziale, si vorrebbe ottenere il maggior numero di informazioni possibili sulla dinamica attraverso un’analisi qualitativa del campo e delle traiettorie associate. Di questo ci occuperemo nel Capitolo 8 (per quanto riguarda i sistemi lineari omogenei a coefficienti costanti) e nel Capitolo 10 (affrontando solo poche questioni, delle moltissime in letteratura, relative all’analisi qualitativa di sistemi nonlineari).

Capitolo 2

Prime definizioni e richiami

2.1 Equazioni differenziali e problema di Cauchy

Un'equazione differenziale ordinaria è una relazione che lega una funzione

$$x : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}^n$$

con I intervallo, ad alcune sue derivate:

$$F(t, x, x', x'', \dots, x^{(k)}) = 0. \quad (2.1)$$

L'ordine di una equazione differenziale ordinaria è l'ordine massimo di derivazione che compare nell'equazione.

Definizione 2.1.1. Una soluzione dell'equazione differenziale ordinaria (2.1) è una funzione x definita e derivabile k volte in un intervallo $I \subseteq \mathbb{R}$ tale che

1. $(t, x(t), x'(t), \dots, x^{(k)}(t)) \in \text{Dom } F, \forall t \in I,$
2. $F(t, x(t), x'(t), \dots, x^{(k)}(t)) = 0, \forall t \in I.$

Osservazione 2.1.2. Nella definizione di soluzione si richiede esplicitamente che sia definita in un intervallo di \mathbb{R} . Ad esempio, le funzioni $x_1(t) = e^{-1/t}$, $t > 0$, e $x_2(t) = e^{-1/t}$, $t < 0$ sono soluzioni dell'equazione differenziale scalare $x' = x/t^2$, mentre la funzione $x_3(t) = e^{-1/t}$, $t \neq 0$ non lo è.

L'aggettivo ordinaria si riferisce al fatto che la funzione incognita è una funzione di una sola variabile reale. Quando l'incognita è una funzione di più variabili, allora si parla di equazione alle derivate parziali.

Si riesce a fare uno studio sistematico delle equazioni differenziali ordinarie che si presentano in forma normale

$$x^{(n)} = f(t, x, x', \dots, x^{(n-1)}), \quad (2.2)$$

dove $f : A \subseteq \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, cioè equazioni nelle quali è possibile esplicitare la derivata di ordine massimo.

Osservazione 2.1.3. Osserviamo che una equazione scalare di ordine n in forma normale

$$x^{(n)} = f(t, x, x', \dots, x^{(n-1)}), \quad x : I \rightarrow \mathbb{R}$$

si può sempre scrivere, in forma equivalente, come un sistema di equazioni differenziali del primo ordine in forma normale ponendo

$$x = x_1, \quad x' = x_2, \quad \dots, \quad x^{(n-1)} = x_n,$$

e ottenendo

$$\begin{cases} x'_1 = x_2, \\ x'_2 = x_3, \\ \vdots \\ x'_{n-1} = x_n, \\ x'_n = f(t, x_1, x_2, \dots, x_n). \end{cases}$$

In maniera analoga, un sistema di equazioni differenziali ordinarie di ordine superiore al primo può essere ricondotto ad un sistema del primo ordine. Ci limiteremo quindi a sviluppare la teoria per sistemi del tipo

$$\begin{cases} x'_1 = f_1(t, x_1, \dots, x_n) \\ \vdots \\ x'_n = f_n(t, x_1, \dots, x_n) \end{cases}$$

che di solito scriveremo in forma vettoriale compatta come

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

dove

$$f = (f_1, \dots, f_n), \quad x = (x_1, \dots, x_n).$$

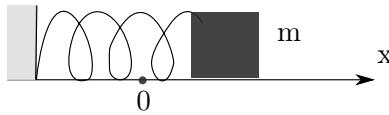


Figura 2.1: Oscillatore armonico

Come è stato evidenziato nel capitolo precedente, la teoria delle equazioni differenziali ordinarie non è fine a se stessa, ma è uno strumento che permette di predire l'evoluzione di un sistema. Il prototipo di equazione differenziale è la seconda legge di Newton

$$F = ma,$$

che scritta in termini della posizione $x(t)$ è

$$x''(t) = \frac{1}{m} F(t, x(t), x'(t)).$$

Vediamo, a titolo di esempio, cosa diventa la legge del moto di un punto materiale di massa m che oscilla su un piano senza attrito grazie all'azione di una molla, le cui caratteristiche materiali sono descritte dalla costante elastica $k > 0$. Se fissiamo nell'origine la posizione di riposo del sistema ed indichiamo con $x(t)$ la posizione del punto materiale al tempo t , la legge di Hooke ci dice che il punto materiale subisce la forza di richiamo $F(t) = -kx(t)$. Quindi, in assenza di attrito, la dinamica dell'oscillatore è descritta dall'equazione differenziale

$$x'' = -\frac{k}{m}x, \quad (2.3)$$

dello stesso tipo di quella ottenuta in (1.4) per le piccole oscillazioni del pendolo e che è soddisfatta da infinite funzioni

$$x(t) = A \cos(\omega t + \phi_0), \quad A, \phi_0 \in \mathbb{R}, \quad (2.4)$$

dove $\omega = \sqrt{k/m}$ è la pulsazione dell'oscillatore. La (2.4) descrive tutti i modi possibili di oscillare intorno alla posizione di equilibrio. Per selezionare una dinamica precisa bisogna fissare l'ampiezza dell'oscillazione A e la fase iniziale ϕ_0 . Un modo di farlo è, ad esempio, assegnare la posizione e la velocità ad un istante iniziale t_0 , anch'esso assegnato.

La versione generale di questa selezione di soluzioni che soddisfano delle assegnate condizioni iniziali prende il nome di Problema di Cauchy.

Definizione 2.1.4 (Problema di Cauchy). Dati $f: A \rightarrow \mathbb{R}^n$, A sottoinsieme aperto di $\mathbb{R} \times \mathbb{R}^n$, e $(t_0, x_0) \in A$ diremo che la funzione x è soluzione del problema di Cauchy

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0. \end{cases} \quad (2.5)$$

se esiste un intervallo non banale $I(x) \subseteq \mathbb{R}$ tale che

- i) $t_0 \in I$ e $x(t_0) = x_0$;
- ii) $(t, x(t)) \in A$ per ogni $t \in I(x)$
- iii) $x : I(x) \rightarrow \mathbb{R}^n$ è derivabile in $I(x)$;
- iv) $x'(t) = f(t, x(t))$, per ogni $t \in I(x)$.

L'insieme delle soluzioni dell'equazione $x' = f(t, x)$ al variare dei dati iniziali prende il nome di integrale generale dell'equazione differenziale.

Nell'esempio dell'oscillatore armonico, possiamo scrivere (2.3) come un sistema

$$\begin{cases} x'_1 = x_2 \\ x'_2 = -\frac{k}{m}x_1 \end{cases}$$

ed il problema di Cauchy si ottiene effettivamente assegnando posizione e velocità iniziali

$$(x_1(t_0), x_2(t_0)) = (x_1^0, x_2^0).$$

Nel caso dell'oscillatore ci aspettiamo che:

- il tipo di dinamica sia completamente determinato dalla legge del moto e dall'assegnazione di una posizione e una velocità iniziali (esistenza e unicità per il problema di Cauchy);
- a piccole variazioni delle condizioni iniziali, corrispondano piccole variazioni della dinamica (dipendenza continua dal dato iniziale).

Se la soluzione soddisfa queste proprietà il problema si dice ben posto. Non tutti i problemi sono ben posti. Per esempio, consideriamo il problema di

trovare una primitiva $x: I(x) \rightarrow \mathbb{R}$ di una fissata funzione f a valori reali

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

Se la funzione f è continua, allora la primitiva esiste sempre e il problema di Cauchy ha un'unica soluzione. Se la funzione f è discontinua, le primitive in generale possono non esistere (si veda l'esempio qui sotto). Nasce il sospetto che per avere buona posizione del problema occorra fare qualche richiesta sulla regolarità della funzione f . Vediamo qualche esempio che ci faccia capire di quali ipotesi abbiamo bisogno.

Esempio 2.1.5 (di non esistenza). Il problema scalare

$$\begin{cases} x' = f(x), \\ x(0) = 0, \end{cases}$$

con f definita da

$$f(x) = \begin{cases} 1 & \text{se } x \leq 0, \\ 0 & \text{se } x > 0, \end{cases}$$

non ammette soluzione. Infatti, se per assurdo esistesse una soluzione x , si dovrebbe avere

$$x'(0) = f(x(0)) = f(0) = 1$$

e quindi in un intorno di 0 si avrebbe

$$\frac{x(t)}{t} > 0.$$

Allora esisterebbe un $\delta > 0$ tale che $x(t) > 0$, per ogni $t \in (0, \delta)$. La f calcolata in punti strettamente positivi è nulla, e quindi $x'(t) = f(x(t)) = 0$ per $t \in (0, \delta)$, cioè x costante in $(0, \delta)$. Ma la condizione iniziale è $x(0) = 0$ e allora per continuità $x \equiv 0$ in $(0, \delta)$. Assurdo perché si aveva $x(t) > 0$ in $(0, \delta)$.

Vedremo che la continuità del campo f garantisce l'esistenza (locale) delle soluzioni (Teorema 4.4.2). Tuttavia non basta per la buona posizione del problema.

Esempio 2.1.6 (non unicità). Consideriamo il problema

$$\begin{cases} x' = 3\sqrt[3]{x^2}, \\ x(0) = 0. \end{cases} \tag{2.6}$$

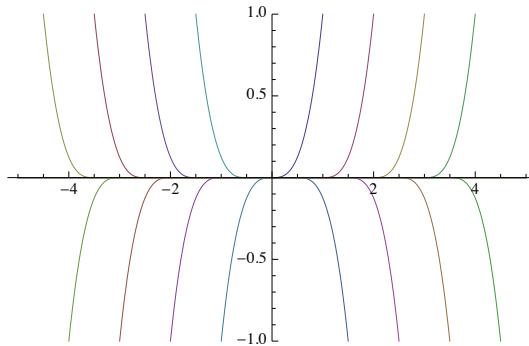


Figura 2.2: Soluzioni del problema di Cauchy (2.6)

Notiamo subito che $x \equiv 0$ è una soluzione del problema. Se poi $k \geq 0$ è un parametro, è facile verificare che anche

$$x_k(t) = \begin{cases} 0 & \text{se } t < k, \\ (t - k)^3 & \text{se } t \geq k, \end{cases}$$

sono soluzioni, ottenibili per esempio per separazione delle variabili (Figura 2.2).

Quale è il problema? Perché non si ha unicità? Vedremo che la mancanza di unicità dipende dal fatto che vicino all'origine la f varia troppo rapidamente (Paragrafo 4.4). Osserviamo che la soluzione del problema di Cauchy relativo ad un dato iniziale $x(0) \neq 0$ risulta essere unica (quindi il problema dell'unicità è di tipo locale).

2.2 Notazioni ed ipotesi

La teoria generale delle equazioni differenziali ordinarie sarà discussa in due tempi: nel prossimo paragrafo richiameremo solo il basilare teorema di esistenza ed unicità in ipotesi di Lipschitzianità, che dovrebbe essere già stato presentato nei corsi di Analisi del primo biennio, mentre nel Capitolo 4 sarà proposta una presentazione più profonda e completa delle proprietà generali delle soluzioni. In questo paragrafo vogliamo fissare tutte le notazioni e introdurre le ipotesi che utilizzeremo nel seguito.

Lo spazio ambiente sarà sempre \mathbb{R}^{n+1} , in cui però terremo separati il ruolo della variabile temporale t da quello della variabile spaziale x , ossia considereremo elementi del tipo $(t, x) \in \mathbb{R} \times \mathbb{R}^n$. Indicheremo con $\|x\|$ la norma Euclidea di $x \in \mathbb{R}^n$ e con $|t|$ il valore assoluto di $t \in \mathbb{R}$.

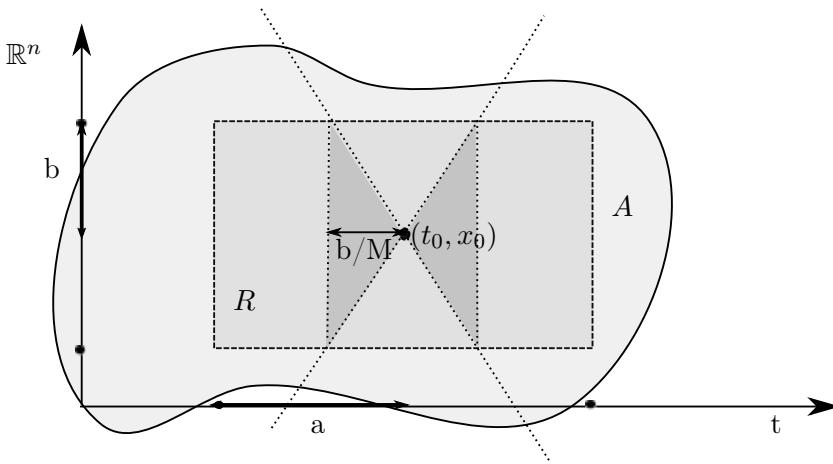


Figura 2.3: Regione di esistenza della soluzione locale

Inoltre assumeremo valide le seguenti ipotesi.

- A è un sottoinsieme aperto di $\mathbb{R} \times \mathbb{R}^n$;
- $f: A \rightarrow \mathbb{R}^n$ è una funzione continua in A (notazione: $f \in C(A, \mathbb{R}^n)$);
- (t_0, x_0) è un punto di A (che determina le condizioni iniziali del problema di Cauchy);

Una volta fissato $(t_0, x_0) \in A$, determiniamo $a, b \in \mathbb{R}$ tali che

$$[t_0 - a, t_0 + a] \times \overline{B}_b(x_0) \subset A, \quad \overline{B}_b(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| \leq b\}.$$

Utilizzeremo le seguenti notazioni

- $R = [t_0 - a, t_0 + a] \times \overline{B}_b(x_0)$ cilindro compatto contenuto in A (sarà la regione intorno ad (t_0, x_0) in cui si cercherà una soluzione locale del problema di Cauchy);
- $M = \max\{\|f(t, x)\| ; (t, x) \in R\}$ (ben definito per il teorema di Weierstrass: f è continua in R sottoinsieme compatto di \mathbb{R}^n).

Infine, dopo quanto abbiamo visto nel paragrafo precedente, ci siamo convinti che per garantire l'unicità per il problema di Cauchy, dovremo richiedere che la f non dilati troppo le distanze in x . L'ipotesi giusta è quella di lipschitzianità.

Definizione 2.2.1. Sia $A \subseteq \mathbb{R} \times \mathbb{R}^n$ un insieme aperto ed $f : A \rightarrow \mathbb{R}^n$. Diremo che la funzione f è localmente lipschitziana in x uniformemente in t (e scriveremo $f \in \text{Lip}_{\text{loc}}^x(A, \mathbb{R}^n)$) se per ogni insieme compatto $K \subset A$ esiste una costante $L > 0$, tale che

$$\|f(t, x) - f(t, y)\| \leq L\|x - y\|, \quad \text{per ogni } (t, x), (t, y) \in K.$$

Diremo che f è lipschitziana in A rispetto ad x uniformemente in t (e scriveremo $f \in \text{Lip}^x(A, \mathbb{R}^n)$) se la costante $L > 0$ può essere scelta indipendente dal compatto K .

Nel seguito, quando supporremo che f sia localmente lipschitziana in x uniformemente in t , una volta fissati t_0, x_0 ed il cilindro R come sopra, indicheremo con L la costante di Lipschitz di f in R , in modo tale risulti

$$\|f(t, x) - f(t, y)\| \leq L\|x - y\|, \quad \forall (t, x), (t, y) \in [t_0 - a, t_0 + a] \times \overline{B}_b(x_0). \quad (2.7)$$

Osservazione 2.2.2. Come conseguenza del teorema della media di Lagrange, la proprietà (2.7) è sicuramente verificata se $f \in C(A, \mathbb{R}^n)$ ha tutte le derivate parziali rispetto alle variabili in x continue, cioè se

$$\frac{\partial f_i}{\partial x_j}(t, x) \in C(A, \mathbb{R}^n), \quad \text{per ogni } i, j = 1, \dots, n.$$

Infatti, se indichiamo con $K = [t_0 - a, t_0 + a] \times \overline{B}_b(x_0)$ il cilindro compatto, e con $D_x f_i$ il gradiente della componente f_i del campo f solo nelle n variabili spaziali, il Teorema di Weierstrass ci garantisce che è ben definito il valore

$$L_i = \max_K \|D_x f_i(t, x)\|, \quad i = 1, \dots, n.$$

Fissiamo ora $(t, x), (t, y)$ in K e applichiamo il Teorema di Lagrange alla funzione $f_i(t, \cdot)$: per ogni $i = 1, \dots, n$ esiste $\xi_i \in K$ tale che

$$f_i(t, x) - f_i(t, y) = \langle D_x f_i(t, \xi_i), x - y \rangle$$

dove $\langle \cdot, \cdot \rangle$ indica l'usuale prodotto scalare in \mathbb{R}^n . Usando questo risultato e la diseguaglianza di Cauchy–Schwarz otteniamo

$$\|f(t, x) - f(t, y)\| \leq \sum_{i=1}^n \|f_i(t, x) - f_i(t, y)\| \leq L \|x - y\|,$$

per ogni $(t, x), (t, y) \in K$, dove $L = \sum_{i=1}^n L_i$.

Esempio 2.2.3. Consideriamo

$$f(t, x) = \frac{|x|^\alpha}{1 + t^2}.$$

Si ha che se

1. $\alpha > 1$ allora f è localmente lipschitziana in \mathbb{R}^2 uniformemente in t ;
2. $\alpha = 1$ allora f è lipschitziana in \mathbb{R}^2 uniformemente in t ;
3. $0 < \alpha < 1$ allora f non è lipschitziana intorno all'origine.

Concludiamo questo paragrafo introducendo la formulazione integrale di un'equazione differenziale. La nuova formulazione ci permetterà di lavorare nello spazio delle funzioni continue senza (apparentemente) ulteriori ipotesi di derivabilità.

Proposizione 2.2.4 (Formulazione integrale). *La funzione $u : I(u) \rightarrow D$ è soluzione del problema di Cauchy*

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

se e solo se valgono

1. $u \in C(I(u), \mathbb{R}^n)$;
2. $u(t) = x_0 + \int_{t_0}^t f(s, u(s)) ds$, per ogni $t \in I(u)$.

Dimostrazione. Il fatto che una soluzione dell'equazione differenziale sia anche soluzione dell'equazione integrale si ottiene semplicemente integrando i due membri dell'equazione e tenendo conto delle condizioni iniziali. Il viceversa è conseguenza diretta del Teorema fondamentale del calcolo integrale. \square

2.3 Esistenza ed unicità locale in ipotesi di Lipschitz

Questo paragrafo è dedicato alla dimostrazione del teorema di esistenza ed unicità in ipotesi di lipschitzianità, i cui ingredienti sono la formulazione integrale e la completezza dello spazio delle funzioni continue.

Teorema 2.3.1 (di esistenza ed unicità in ipotesi di Lipschitz). *Sia A un sottoinsieme aperto di $\mathbb{R} \times \mathbb{R}^n$ e sia $f \in C(A, \mathbb{R}^n) \cap \text{Lip}_{\text{loc}}^x(A, \mathbb{R}^n)$. Allora per ogni $(t_0, x_0) \in A$ esiste $\delta > 0$ e $x \in C^1(I_\delta, \mathbb{R}^n)$, $I_\delta = [t_0 - \delta, t_0 + \delta]$, soluzione di*

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0. \end{cases} \quad (2.8)$$

Tale soluzione è unica, nel senso che ogni altra soluzione y di (2.8) coincide con x in $I_\delta \cap I(y)$. Inoltre, nelle notazioni del Paragrafo 2.2, si può scegliere $\delta = \min(a, b/M)$.

Dimostrazione. Fissato $(t_0, x_0) \in A$, determiniamo a, b, M ed L come nel Paragrafo 2.2 e prendiamo $\delta = \min(a, b/M)$.

Passo 1: Mappa di Picard e sue iterate. Grazie alla formulazione integrale 2.2.4, il problema di determinare una soluzione di (2.8) è riconducibile alla determinazione di un punto fisso dell'operatore di Picard

$$P: C(I_\delta, \overline{B}_b(x_0)) \longrightarrow C(I_\delta, \overline{B}_b(x_0))$$

definito da

$$P[x](t) = x_0 + \int_{t_0}^t f(s, x(s)) ds, \quad t \in I_\delta, \quad x \in C(I_\delta, \overline{B}_b(x_0)). \quad (2.9)$$

Questo problema può essere risolto con l'utilizzo diretto del teorema di punto fisso di Banach-Caccioppoli per le contrazioni (si veda, per esempio, 3) oppure con un metodo più costruttivo di approssimazioni successive.

Iniziamo col dimostrare che l'operatore di Picard è ben definito, ossia trasforma funzioni continue in I_δ e a valori nella sfera $\overline{B}_b(x_0)$ in funzioni dello stesso tipo. Per ogni $x \in C(I_\delta, \overline{B}_b(x_0))$ si ha che per ogni $t \in I_\delta$ l'elemento $(t, x(t))$ appartiene al cilindro R contenuto nel dominio di f , quindi $P[x](t)$ è

una funzione definita e continua in I_δ . Resta da verificare che $P[x](t) \in \overline{B}_b(x_0)$ per ogni $t \in I_\delta$. Questo segue dalla scelta di δ che abbiamo fatto:

$$\|x_0 - P[x](t)\| \leq \left| \int_{t_0}^t \|f(s, x(s))\| ds \right| \leq M\delta \leq b.$$

Possiamo quindi considerare la successione $x_k: I_\delta \rightarrow \overline{B}_b(x_0)$ definita per ricorrenza:

$$\begin{cases} x_0(t) = x_0 \\ x_{k+1}(t) = x_0 + \int_{t_0}^t f(s, x_k(s)) ds \end{cases} \quad t \in I_\delta.$$

Passo 2: passaggio al limite. Proviamo ora che la successione delle iterate della mappa di Picard converge alla soluzione dell'equazione. Scriviamo

$$x_{k+1} = x_0 + \sum_{j=0}^k (x_{j+1} - x_j)$$

ed mostriamo, per induzione, che vale la stima

$$\|x_{j+1}(t) - x_j(t)\| \leq \frac{ML^j}{(j+1)!} |t - t_0|^{j+1}. \quad (2.10)$$

Il caso $j = 0$ corrisponde alla stima

$$\|x_1(t) - x_0\| \leq M |t - t_0|$$

già provata nel passo precedente. Supponiamo ora che (2.10) valga per $j - 1$ e mostriamo che vale anche per j : ricordando che L è la costante di lipschitzianità di f in R , otteniamo

$$\begin{aligned} \|x_{j+1}(t) - x_j(t)\| &\leq \left| \int_{t_0}^t \|f(s, x_j(s)) - f(s, x_{j-1}(s))\| ds \right| \\ &\leq L \left| \int_{t_0}^t \|x_j(s) - x_{j-1}(s)\| ds \right| \\ &\leq L \frac{ML^{j-1}}{j!} \left| \int_{t_0}^t |s - t_0|^j ds \right| = M \frac{L^j}{j!} \frac{|t - t_0|^{j+1}}{j+1} \end{aligned}$$

dove nell'ultima diseguaglianza abbiamo usato l'ipotesi induttiva. Abbiamo quindi dimostrato che x_{k+1} è la ridotta k -esima della serie di funzioni

$$x_0 + \sum_{j=0}^{+\infty} (x_{j+1} - x_j).$$

che converge totalmente in I_δ , grazie alla (2.10) e al fatto che

$$\lim_{k \rightarrow +\infty} \sum_{j=0}^k \frac{ML^j}{(j+1)!} \delta^{j+1} = \frac{M}{L}(e^{L\delta} - 1).$$

Questo garantisce la convergenza uniforme della successione $\{x_k\}$ in I_δ ad una funzione $x \in C(I_\delta, \overline{B}_b(x_0))$. Inoltre, le funzioni $g_k(t) = f(t, x_k(t))$ convergono uniformemente in I_δ alla funzione $g(t) = f(t, x(t))$. Infatti, per l'ipotesi di locale lipschitzianità della f abbiamo che

$$\|g_k(t) - g(t)\| \leq L \|x_k(t) - x(t)\|, \quad t \in I_\delta.$$

Possiamo allora passare al limite sotto il segno di integrale nell'identità

$$x_{k+1}(t) = x_0 + \int_{t_0}^t f(s, x_k(s)) ds$$

ottenendo che il limite x è soluzione dell'equazione integrale in I_δ e quindi dell'equazione differenziale (2.8).

Passo 3: unicità. Supponiamo che $y : I(y) \rightarrow \mathbb{R}^n$ sia un'altra soluzione di (2.8). Dimostriamo per induzione che vale la stima

$$\|y(t) - x_k(t)\| \leq \frac{ML^k}{(k+1)!} |t - t_0|^{k+1}, \quad \forall t \in I_\delta \cap I(y). \quad (2.11)$$

Infatti per $k = 0$, dal momento che y è soluzione di (2.8) e quindi anche della relativa equazione integrale, abbiamo che

$$\begin{aligned} \|y(t) - x_0\| &= \left\| \int_{t_0}^t f(s, y(s)) ds \right\| \\ &\leq \int_{t_0}^t \|f(s, y(s))\| ds \leq M |t - t_0|. \end{aligned}$$

Se ora supponiamo che (2.11) sia vera per k , otteniamo

$$\begin{aligned} \|y(t) - x_{k+1}(t)\| &= \left\| \int_{t_0}^t f(s, y(s)) - f(s, x_k(s)) ds \right\| \\ &\leq L \int_{t_0}^t \|y(s) - x_k(s)\| ds \\ &\leq L \frac{ML^k}{(k+1)!} \int_{t_0}^t |s - t_0|^{k+1} ds \\ &= \frac{ML^{k+1}}{(k+2)!} |t - t_0|^{k+2} \end{aligned}$$

in cui, nuovamente, abbiamo ustato la lipschitzianità della funzione f e ipotesi induttiva. A questo punto si può passare al limite per $k \rightarrow +\infty$ in (2.11) e si ottiene $y = x$ in $I_\delta \cap I(y)$. \square

Osservazione 2.3.2. Sottolineiamo che $\delta = \min(a, b/M)$ è la richiesta minima per garantirci che la soluzione del problema di Cauchy (2.8) non esca dal cilindro R . Infatti, essendo $x(t_0) = x_0$ e $\|x'\| \leq M$, la soluzione deve essere contenuta all'interno dei coni disegnati nella Figura 2.3 e quindi può uscire da R solo per $|t - t_0| > b/M$ (a patto che questo intervallo sia, a sua volta, minore dell'intervallo temporale del cilindro, ossia che b/M sia più piccolo di a).

Osservazione 2.3.3 (Interpretazione geometrica del Teorema 2.3.1 nel caso $n = 1$). Il fatto che per ogni scelta di $(t_0, x_0) \in A$ esista almeno una soluzione del problema di Cauchy associato ha come controparte geometrica che il dominio A sia ricoperto di traiettorie (i.e. di grafici di soluzioni dell'equazione). In più, l'unicità locale garantisce che i grafici di due soluzioni non si possono intersecare mai. Infatti, se questo succedesse nel punto (\bar{t}, \bar{x}) , il problema di Cauchy con dato iniziale $x(\bar{t}) = \bar{x}$ avrebbe due soluzioni distinte. In particolare due soluzioni di un'equazione differenziale (e, più in generale, di un sistema di equazioni differenziali) che coincidono in un punto, coincidono in tutta l'intersezione dei loro intervalli di esistenza. Il fatto che, in ipotesi di esistenza ed unicità, due soluzioni distinte non possano intersecarsi è particolarmente utile quando si vogliono determinare delle stime a priori sulle soluzioni. Per esempio, nel caso dell'equazione

$$x' = \alpha x,$$

basta osservare che $x(t) = 0$, $t \in \mathbb{R}$, è soluzione per concludere che ogni altra soluzione non si può mai annullare e, in particolare, che la soluzione del problema di Cauchy corrispondente ad un dato iniziale $x(t_0) = x_0 \neq 0$ ha sempre lo stesso segno di x_0 (questa questione era rimasta in sospeso nella descrizione del modello base per l'evoluzione di specie, Sezione 1.1).

Il risultato appena dimostrato ci garantisce l'esistenza e l'unicità della soluzione del problema di Cauchy in un intorno del dato iniziale. Ovviamente, dal punto di vista operativo, ossia quando si vanno a risolvere esplicitamente delle equazioni differenziali, questo risultato determina in modo troppo vago l'insieme di definizione della soluzione. Questa vaghezza viene risolta procedendo nel modo seguente: la soluzione, la cui esistenza ed unicità sono garantite dal Teorema 2.3.1, viene definita nel più grande insieme di definizione possibile, ossia si considera la soluzione massimale del problema di Cauchy. Torneremo

su questa questione nel Paragrafo 4.5, ma abbiamo bisogno, in funzione degli esercizi relativi al prossimo Capitolo dedicato ai metodi risolutivi, di anticipare le seguenti definizioni.

Definizione 2.3.4. Sia $u : I(u) \rightarrow \mathbb{R}^n$ una soluzione di $x' = f(t, x)$. Allora $v : I(v) \rightarrow \mathbb{R}^n$ è un prolungamento di u se

- i) $I(u) \subseteq I(v)$ e v è soluzione di $x' = f(t, x)$ in $I(v)$;
- ii) $u = v$ in $I(u)$.

La funzione u si dice soluzione massimale di $x' = f(t, x)$ se non ammette prolungamenti non banali, cioè se per ogni v prolungamento di u , si ha $I(v) = I(u)$.

Capitolo 3

Alcuni metodi di risoluzione

In questo capitolo introduciamo alcuni metodi esplicativi per la risoluzione di particolari tipi di equazioni differenziali ordinarie.

3.1 Variabili separabili

Consideriamo il problema di Cauchy relativo ad un'equazione differenziale scalare della forma seguente

$$\begin{cases} x' = g(t)h(x) \\ x(t_0) = x_0 \end{cases} \quad (3.1)$$

dove $g \in C(I, \mathbb{R})$ ed $h \in C(J, \mathbb{R})$ sono funzioni assegnate, $I, J \subseteq \mathbb{R}$ intervalli.

Gli zeri della h ci forniscono soluzioni costanti: se x_0 dato iniziale, è tale che $h(x_0) = 0$ allora

$$x(t) = x_0, \quad \text{per ogni } t \in I$$

è soluzione globale di (3.1).

Consideriamo ora un dato iniziale x_0 tale che $h(x_0) \neq 0$. Sia $W \subseteq J$ il massimo intervallo aperto contenente x_0 e tale che $h(\xi) \neq 0$ per ogni $\xi \in W$. W non è altro che la componente connessa dell'insieme $\{h \neq 0\}$ che contiene x_0 e h ha un segno fissato in W .

Se x è soluzione di (3.1), per continuità esiste un intorno J_1 di t_0 tale che $x(t) \in W$ per ogni $t \in J_1$. Allora in J_1 posso separare le variabili ottenendo

$$\frac{x'(t)}{h(x(t))} = g(t), \quad \text{per ogni } t \in J_1.$$

Integrando abbiamo

$$\int_{t_0}^t \frac{x'(s)}{h(x(s))} ds = \int_{t_0}^t g(s) ds$$

che, a seguito della sostituzione $\xi = x(s)$ nel primo integrale, diventa

$$\int_{x_0}^{x(t)} \frac{1}{h(\xi)} d\xi = \int_{t_0}^t g(s) ds, \text{ per ogni } t \in J_1. \quad (3.2)$$

L'identità (3.2) definisce implicitamente la soluzione del problema di Cauchy. Mostriamo che la definisce univocamente.

Teorema 3.1.1. *L'identità integrale (3.2) definisce localmente un'unica soluzione di (3.1) quando x_0 è tale che $h(x_0) \neq 0$.*

Dimostrazione. Indichiamo con $G \in C^1(I)$ la primitiva di g :

$$G(t) = \int_{t_0}^t g(s) ds.$$

Analogamente consideriamo $H : W \rightarrow \mathbb{R}$ definita da

$$H(x) = \int_{x_0}^x \frac{1}{h(\xi)} d\xi.$$

Allora $H \in C^1(W)$ e $H' = \frac{1}{h}$, con segno fissato in W . Quindi $H : W \rightarrow H(W)$ è un diffeomorfismo. Vogliamo, almeno in linea teorica, esplicitare x a partire dalla relazione

$$H(x(t)) = G(t).$$

Poiché H è strettamente monotona, e quindi invertibile, in W , potremo esplorare $x(t)$ per ogni $t \in J_1$ tale che $G(t) \in H(W)$. Poiché $G(t_0) = 0 = H(x_0)$, $H(W)$ è aperto e G è continua, esiste un intorno di t_0 tale che la sua immagine sia contenuta in $H(W)$. Sia dunque J il più grande intorno di t_0 tale che $G(J) \subseteq H(W)$. Su J vale l'identità $H(x(t)) = G(t)$ ed è possibile invertire H , ottenendo

$$x(t) = H^{-1}(G(t)), \text{ per ogni } t \in J.$$

La x così definita è una soluzione di (3.1) nell'intervallo J .

Supponiamo ora che \tilde{x} sia un'altra soluzione di (3.1). Poiché $\tilde{x}(t_0) = x_0$, si ha che $h(\tilde{x}(t_0)) \neq 0$. Per la continuità di \tilde{x} , esiste un intorno \tilde{J} di t_0 tale che $h(\tilde{x}(t)) \neq 0$ per ogni $t \in \tilde{J}$. Allora $\tilde{x}(\tilde{J}) \subseteq W$, per cui in \tilde{J} la \tilde{x} risolve l'identità integrale (3.2), cioè si ha

$$H(\tilde{x}(t)) = G(t), \quad \text{per ogni } t \in \tilde{J}.$$

Poiché H è un diffeomorfismo locale, l'equazione ammette un'unica soluzione, e quindi $x \equiv \tilde{x}$ in $J \cap \tilde{J}$. \square

Esempio 3.1.2. Vogliamo risolvere il problema di Cauchy

$$\begin{cases} x' = 1 + x^2, \\ x(t_0) = x_0. \end{cases}$$

Seguiamo la strategia generale appena esaminata. Si ha che $g(t) = 1$ e $h(x) = 1 + x^2$, che non si annulla mai. Quindi non ci sono soluzioni costanti e possiamo sempre dividere per $1 + x^2$. La (3.2) in questo caso diventa

$$\int_{x_0}^{x(t)} \frac{1}{1 + \xi^2} d\xi = t - t_0,$$

e, integrando,

$$\arctan x(t) = \arctan x_0 + t - t_0.$$

Dobbiamo invertire. L'immagine della funzione arcotangente è $(-\frac{\pi}{2}, +\frac{\pi}{2})$, per cui per invertire si deve avere

$$-\frac{\pi}{2} < \arctan x_0 + t - t_0 < \frac{\pi}{2},$$

cioè la soluzione $x(t) = \tan(\arctan x_0 + t - t_0)$ è definita nell'intervallo

$$J = \left(\alpha - \frac{\pi}{2}, \alpha + \frac{\pi}{2} \right),$$

con $\alpha = t_0 - \arctan x_0$. Osserviamo quindi che per ogni dato iniziale (t_0, x_0) esiste un'unica soluzione e che tale soluzione non è globale nonostante il secondo membro dell'equazione sia regolarissimo su tutto l'asse reale.

Esempio 3.1.3. Si consideri il problema di Cauchy

$$\begin{cases} x' = \operatorname{sign}(x)\sqrt{|x|}, \\ x(t_0) = x_0. \end{cases}$$

Per prima cosa osserviamo che il secondo membro è continuo e nullo in $x = 0$, per cui $x \equiv 0$ è soluzione costante. Osserviamo poi che il secondo membro è dispari, per cui $x(t)$ è soluzione se e solo se $-x(t)$ è soluzione. Per questo ci basta studiare il caso $x_0 > 0$ ed ottenere le altre soluzioni per simmetria. Osserviamo che, essendo $x_0 > 0$, esisterà un intorno di t_0 tale che $x(t) > 0$ in tale intorno. Ma allora, l'equazione differenziale ci dice che $x'(t) = \sqrt{x} > 0$, quindi la funzione x è strettamente crescente in tale intorno e $\operatorname{sign}(x(t)) = 1$. Separiamo le variabili e integriamo:

$$\int_{x_0}^{x(t)} \frac{1}{\sqrt{\xi}} d\xi = t - t_0,$$

da cui

$$\sqrt{x(t)} = \sqrt{x_0} + \frac{t - t_0}{2}.$$

Esplicitiamo la soluzione richiedendo che il membro di destra sia strettamente positivo (la positività serve per poter invertire la radice, in più avevamo richiesto che $x(t)$ non si annullasse per poter separare le variabili), ottenendo

$$x(t) = \left(\sqrt{x_0} + \frac{t - t_0}{2} \right)^2, \text{ per } t > t_0 - 2\sqrt{x_0}. \quad (3.3)$$

Osserviamo che la soluzione ottenuta si può raccordare in maniera C^1 alla funzione identicamente nulla in corrispondenza di $t = t_0 - 2\sqrt{x_0}$, quindi si può costruire la soluzione globale di (3.3) data da

$$x(t) = \begin{cases} \left(\sqrt{x_0} + \frac{t - t_0}{2} \right)^2 & \text{se } t > t_0 - 2\sqrt{x_0}, \\ 0 & \text{se } t \leq t_0 - 2\sqrt{|x_0}|. \end{cases}$$

Ricordando che se $x_0 < 0$, allora $-x(t)$ è soluzione, abbiamo che, per ogni $x_0 \neq 0$

$$x(t) = \begin{cases} \operatorname{sign}(x_0) \left(\sqrt{|x_0|} + \frac{t - t_0}{2} \right)^2 & \text{se } t > t_0 - 2\sqrt{|x_0|}, \\ 0 & \text{se } t \leq t_0 - 2\sqrt{|x_0|}. \end{cases}$$

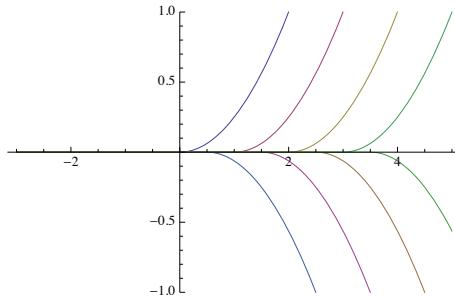


Figura 3.1: Alcune soluzioni dell'esempio 3.1.2

è l'unica soluzione del problema di Cauchy considerato.

Se invece consideriamo il dato iniziale $x_0 = 0$, come già detto, abbiamo la soluzione costante $x \equiv 0$. D'altra parte possiamo anche raccordare in maniera C^1 la soluzione nulla con delle parabole in qualsiasi istante dall'asse delle t e avere ancora una soluzione: abbiamo, ad esempio, tutta una famiglia di soluzioni $\{x_\alpha\}$ con $\alpha \geq t_0$ definita da

$$x_\alpha(t) = \begin{cases} \frac{(t-\alpha)^2}{4} & \text{oppure } -\frac{(t-\alpha)^2}{4} & \text{se } t > \alpha, \\ 0 & & \text{se } t \leq \alpha, \end{cases}$$

(si veda la figura 3.1). Per il dato $x_0 = 0$ abbiamo quindi una perdita di unicità: possiamo esibire infinite soluzioni. Questo non va in conflitto con il Teorema 2.3.1, in quanto la funzione $f(x) = \text{sign}(x)\sqrt{|x|}$ non è Lipschitz in $x = 0$ (la sua derivata esplode in quel punto). Tuttavia è continua. Come vedremo questa condizione, pur non assicurando l'unicità, basta a garantire l'esistenza di soluzioni.

Esempio 3.1.4 (Equazioni lineari omogenee). Consideriamo l'equazione lineare omogenea

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

con $a : I \rightarrow \mathbb{R}$ continua, $I \subseteq \mathbb{R}$ intervallo aperto contenente t_0 . Il secondo membro è localmente lipschitziano perché lineare nella variabile x e quindi vale il Teorema 2.3.1 di esistenza e unicità locale. Inoltre $x \equiv 0$ è soluzione con dato iniziale $x_0 = 0$ e allora ogni altra soluzione mantiene il segno del proprio dato iniziale $x_0 \neq 0$. Separiamo le variabili e integriamo

$$\int_{x_0}^{x(t)} \frac{1}{\xi} d\xi = \int_{t_0}^t a(s) ds \implies \log \left| \frac{x(t)}{x_0} \right| = \int_{t_0}^t a(s) ds \implies$$

$$\left| \frac{x(t)}{x_0} \right| = e^{\int_{t_0}^t a(s) ds}$$

e allora poiché $\text{sign}(x(t)) = \text{sign}(x_0)$ per ogni $t \in I$, otteniamo la soluzione globale

$$x(t) = x_0 e^{\int_{t_0}^t a(s) ds}.$$

3.2 1-forme ed equazioni differenziali

Consideriamo un'equazione differenziale della forma

$$x' = -\frac{b(t, x)}{a(t, x)}. \quad (3.4)$$

Introdurremo un metodo di risoluzione che fa uso della 1-forma differenziale

$$\omega = a(t, x)dx + b(t, x)dt,$$

che penseremo associata a (3.4) e di cui useremo eventuali proprietà per determinare le soluzioni, almeno in forma implicita.

Più precisamente, se ω ammette un potenziale, cioè esiste una funzione U tale che

$$\nabla U(t, x) = (a(t, x), b(t, x)),$$

allora ci aspettiamo che una soluzione $x(t)$ di (3.4) in qualche senso soddisfi la relazione $\omega = 0$ e, di conseguenza,

$$\frac{d}{dt} U(t, x(t)) = 0 \implies U(t, x(t)) = \text{costante}.$$

Quindi associando all'equazione differenziale una 1-forma, se quest'ultima è esatta, otteniamo un integrale primo: le soluzioni dell'equazione giacciono sulle curve di livello del potenziale.

Chiaramente quanto detto finora necessita di alcune precisazioni: cosa significa risolvere $\omega = 0$? In che senso è possibile associare ad un'equazione differenziale una forma differenziale? Nel seguito ci proponiamo di rendere rigoroso questo procedimento:

1. associare ad una equazione differenziale una 1-forma;
2. se la 1-forma è esatta mostrare che le soluzioni dell'equazione differenziale sono contenute negli insiemi di livello delle primitive;

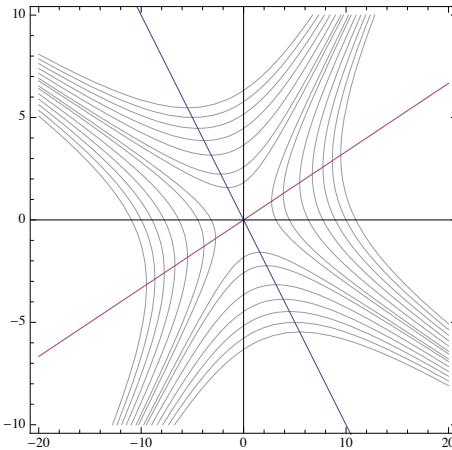


Figura 3.2: Alcune soluzioni dell'Esempio 3.2.1

3. studiare il caso in cui la forma è esatta a meno di un fattore moltiplicativo (metodo del fattore integrante).

Prima di procedere, supponiamo di aver perfezionato già la strategia e proviamo ad applicarla in un caso concreto.

Esempio 3.2.1. Consideriamo il problema di Cauchy

$$\begin{cases} x' = -\frac{t+x}{t-3x}, \\ x(0) = 1. \end{cases}$$

Con gli strumenti introdotti finora l'unico modo di studiare questo sistema è un'analisi qualitativa del campo di velocità. Il segno di x' è facilmente determinabile:

$$\begin{cases} x'(t) < 0 & \text{se } x > -t \& x > t/3 \text{ oppure se } x < -t \& x < t/3, \\ x'(t) > 0 & \text{se } x < -t \& x > t/3 \text{ oppure se } x > -t \& x < t/3. \end{cases}$$

Sulla retta $x = t/3$ le traiettorie arrivano a tangente verticale perché il denominatore di f si annulla; sulla retta $x = -t$ invece arrivano a tangente orizzontale perché il campo è nullo. Siamo così in grado di disegnare, almeno qualitativamente, alcune soluzioni dell'equazione (si veda Figura 3.2). I grafici ottenuti ci suggeriscono che, tra una famiglia di curve simili tra loro e l'altra ci siano delle soluzioni della forma $x = \alpha t$, che otterremo facilmente per sostituzione nell'equazione differenziale:

$$\alpha = -\frac{t+\alpha t}{t-3\alpha t} \implies 3\alpha^2 - 2\alpha - 1 = 0 \implies \alpha = -\frac{1}{3}, 1.$$

Quindi abbiamo ottenuto le due soluzioni esplicite $x = t$ e $x = -1/3t$.

Tramite l'analisi qualitativa siamo arrivati con non poca fatica ad avere solo una descrizione approssimativa delle soluzioni. Vogliamo ora mostrare come il metodo delle 1-forme permetta di avere informazioni molto più precise (in questo caso l'espressione esplicita della soluzione) e con poco sforzo. La 1-forma associata è

$$\omega = (t + x)dt + (t - 3x)dx.$$

Se esiste una primitiva $U = U(t, x)$ di ω , questa per definizione deve soddisfare le condizioni

$$\begin{cases} U_t = t + x \\ U_x = t - 3x. \end{cases}$$

Integrando la prima otteniamo

$$U(t, x) = \frac{t^2}{2} + tx + \phi(x).$$

Per determinare la ϕ deriviamo in x e imponiamo che sia uguale a $t - 3x$, ottenendo l'equazione

$$\phi' = -3x \implies \phi(x) = -\frac{3}{2}x^2$$

e allora

$$U(t, x) = \frac{t^2}{2} + tx - \frac{3}{2}x^2.$$

Abbiamo detto che ci aspettiamo che la soluzione $x(t)$ del problema di Cauchy soddisfi

$$\frac{t^2}{2} + tx(t) - \frac{3}{2}x(t)^2 = C.$$

La costante C è determinata dal dato iniziale: sostituendo $t = 0$ e $x(0) = 1$ nella precedente equazione e otteniamo che $x(t)$ risolve l'equazione di secondo grado

$$3x(t)^2 - 2tx(t) - (t^2 + 3) = 0$$

da cui

$$x(t) = \frac{t + \sqrt{4t^2 + 9}}{3}.$$

3.2.1 Richiami sulle 1-forme differenziali

Ricordiamo che lo spazio duale di \mathbb{R}^n , indicato con $(\mathbb{R}^n)^*$, è l'insieme dei funzionali lineari su \mathbb{R}^n , ossia $L \in (\mathbb{R}^n)^*$ se e solo se $L: \mathbb{R}^n \rightarrow \mathbb{R}$ verifica $L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$ per ogni $\alpha, \beta \in \mathbb{R}$ e $x, y \in \mathbb{R}^n$. Inoltre $(\mathbb{R}^n)^*$ è isomorfo ad \mathbb{R}^n tramite la mappa $T: (\mathbb{R}^n)^* \rightarrow \mathbb{R}$ definita da $T(p)(x) = \langle p, x \rangle$. In particolare $(\mathbb{R}^n)^*$ ha dimensione n ed è possibile associare alla base standard $\{x_1, \dots, x_n\}$ di \mathbb{R}^n la base duale $\{dx_1, \dots, dx_n\}$ caratterizzata da

$$dx_i(x_j) = \delta_{ij} = \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases}$$

Definizione 3.2.2. Sia $A \subseteq \mathbb{R}^n$ un aperto. Una 1-forma differenziale su A è una applicazione

$$\omega : A \rightarrow (\mathbb{R}^n)^*$$

che ad ogni punto $p \in A$ associa $\omega_p \in (\mathbb{R}^n)^*$. Scrivendo ω_p rispetto alla base duale,

$$\omega_p = a_1(p)dx_1 + \cdots + a_n(p)dx_n,$$

si ottengono le funzioni coordinate $a_i : A \rightarrow \mathbb{R}$.

Diremo che la 1-forma ω è C^k in A , $k \in \mathbb{N}$ se le coordinate a_i sono funzioni di classe C^k in A .

Per i nostri scopi ci limiteremo al caso bidimensionale \mathbb{R}^2 con coordinate $\{t, x\}$. La base duale è allora $\{dt, dx\}$ e una 1-forma su un aperto $A \subseteq \mathbb{R}^2$ è una espressione del tipo

$$\omega = a(t, x)dt + b(t, x)dx$$

che tipicamente supporremo essere di classe C^1 nel suo insieme di definizione.

Definizione 3.2.3. La 1-forma $\omega \in C^1(A)$ si dice chiusa in A se

$$\frac{\partial a}{\partial x} = \frac{\partial b}{\partial t}$$

e si dice esatta se esiste una funzione $U : A \rightarrow \mathbb{R}$, $U \in C^1(A)$, tale che

$$\nabla U(t, x) = (a(t, x), b(t, x)), \quad \text{per ogni } (t, x) \in A.$$

Tale U viene detta potenziale oppure primitiva di ω , e scriveremo

$$\omega = dU = U_t dt + U_x dx.$$

Ricordiamo che una 1-forma $\omega \in C^1$ esatta è anche chiusa (come elementare conseguenza del Teorema di Schwartz), mentre il viceversa non è sempre vero. Una condizione sufficiente è data dal seguente Teorema.

Teorema 3.2.4. *Sia $A \subseteq \mathbb{R}^2$ un insieme aperto e semplicemente connesso (cioè A è connesso e ogni curva chiusa contenuta in A è omotopa ad un punto). Allora ω è esatta se e solo se è chiusa. In particolare, ogni forma differenziale chiusa è localmente esatta.*

Per la dimostrazione del teorema e ulteriori approfondimenti sull'argomento rinviamo, ad esempio, al testo [3].

3.2.2 Curve soluzione

Sia ω una 1-forma su A ,

$$\omega = a(t, x)dt + b(t, x)dx.$$

Vogliamo dare una nozione rigorosa di soluzione dell'equazione $\omega = 0$.

Ad ogni curva C^1

$$\varphi(\tau) = (\varphi_1(\tau), \varphi_2(\tau)), \quad \tau \in J$$

$J \subseteq \mathbb{R}$ intervallo, tale che $\varphi(J) \subset A$, è possibile associare ad ω la 1-forma su J definita da

$$\varphi^*\omega = [a(\varphi_1(\tau), \varphi_2(\tau))\varphi'_1(\tau) + b(\varphi_1(\tau), \varphi_2(\tau))\varphi'_2(\tau)] d\tau, \quad \tau \in J$$

è detta pull-back di ω su J tramite φ .

Definizione 3.2.5. Diremo che la curva Γ è soluzione di $\omega = 0$ se esiste $\varphi : J \rightarrow A$ parametrizzazione C^1 di Γ , tale che $\varphi^*\omega = 0$. La soluzione Γ passa per il punto $(t_0, x_0) \in A$ se $(t_0, x_0) \in \varphi(J)$.

Consideriamo ora l'equazione differenziale

$$x' = -\frac{a(t, x)}{b(t, x)} \quad (3.5)$$

con $a, b \in C(A)$, $b \neq 0$ in A , in modo che la 1-forma associata sia proprio ω . Abbiamo il seguente risultato che lega le soluzioni di $\omega = 0$ a quelle di (3.5).

Ricordiamo che una curva Γ si dice regolare se è C^1 e se ammette una parametrizzazione $\varphi : J \rightarrow A$ tale che $\|\varphi'(\tau)\| \neq 0$ per ogni $\tau \in J$.

Proposizione 3.2.6. *Se $x(t)$ è soluzione di (3.5) allora il suo grafico è soluzione di $\omega = 0$.*

Viceversa, se Γ , parametrizzata da $\varphi(t) = (\varphi_1(t), \varphi_2(t))$, $t \in J$, è soluzione di $\omega = 0$ e b non si annulla lungo Γ , allora $x(t) = \varphi_2[\varphi_1^{-1}(t)]$, $t \in \varphi_1(J)$ è soluzione di (3.5).

Dimostrazione. Sia $x \in C^1(I)$ soluzione di (3.5). La corrispondente soluzione di $\omega = 0$ è esattamente il grafico di x . Infatti, posto $\varphi(\tau) = (\tau, x(\tau))$, si ha che φ è una curva regolare di classe C^1 . Inoltre, il pull-back di ω tramite φ è

$$\varphi^*\omega = [a(\tau, x(\tau)) + b(\tau, x(\tau))x'(\tau)] d\tau.$$

Poiché x è soluzione di (3.5) si ha che $\varphi^*\omega = 0$ in I , cioè φ risolve $\omega = 0$.

Viceversa, sia $\varphi \in C^1(J, A)$ un cammino tale che

$$\varphi^*\omega = [a(\varphi_1(\tau), \varphi_2(\tau))\varphi'_1(\tau) + b(\varphi_1(\tau), \varphi_2(\tau))\varphi'_2(\tau)] d\tau = 0,$$

per ogni $\tau \in J$. Osserviamo che $\varphi'_1(\tau) \neq 0$ per ogni $\tau \in J$. Infatti se per assurdo $\varphi'_1(\tau_0) = 0$, essendo b non nulla, necessariamente si dovrebbe avere anche $\varphi'_2(\tau_0) = 0$, contraddicendo la regolarità di φ . Quindi φ_1 e φ_2 sono entrambi diffeomorfismi C^1 . Poniamo $I = \varphi_1(J)$ e

$$x(t) = \varphi_2[\varphi_1^{-1}(t)], \text{ per } t \in I.$$

Tale $x : I \rightarrow A$ è C^1 ed è soluzione di (3.5), infatti

$$x'(t) = \varphi'_2[\varphi_1^{-1}(t)] \frac{1}{\varphi'_1[\varphi_1^{-1}(t)]} = -\frac{a(t, x(t))}{b(t, x(t))}.$$

□

Osservazione 3.2.7. L'equivalenza tra le soluzioni classiche di (3.5) e le soluzioni di $\omega = 0$ sussiste solo se $b \neq 0$ nel dominio A in cui si sta considerando l'equazione. Infatti, mentre tutte le soluzioni di (3.5) sono anche soluzioni di $\omega = 0$, il viceversa in generale non è vero. A titolo di esempio si consideri la Figura 3.2 dell'Esempio 3.2.1: le curve che tagliano la $x = -t$ a tangente verticale sono curve soluzione di $\omega = 0$, ma non risolvono l'equazione differenziale poiché non sono il grafico di una funzione.

Quindi le soluzioni di $\omega = 0$ possono essere determinate anche nel caso $b = 0$, mentre nella equazione (3.5) dobbiamo restringerci al caso delle curve cartesiane. In tale senso, abbiamo ottenuto una generalizzazione del concetto di soluzione.

Supponiamo ora che la forma ω sia esatta in A . Vediamo quali informazioni possiamo ricavare sulle curve soluzione di $\omega = 0$ a partire dal suo potenziale.

Definizione 3.2.8. Sia $F \in C^1(A, \mathbb{R})$. Consideriamo per $c \in \mathbb{R}$ l'insieme di livello

$$F^{-1}(c) = \{(t, x) \in A : F(t, x) = c\}.$$

Diremo che c è un valore regolare per F in A se $\nabla F(t, x) \neq 0$ per ogni $(t, x) \in F^{-1}(c)$.

Gli insiemi di livello del potenziale relativi a valori regolari contengono tutte e sole le curve soluzione.

Proposizione 3.2.9. *Sia ω una forma esatta in A , $F \in C^1(A, \mathbb{R})$ un suo potenziale e c un valore regolare per F . Allora per ogni punto $(t, x) \in F^{-1}(c)$ passa una curva soluzione di $\omega = 0$. Inoltre, il supporto di ogni curva soluzione di $\omega = 0$ è interamente contenuto in un insieme di livello del potenziale F .*

Dimostrazione. Per il Teorema della funzione implicita sappiamo che $F^{-1}(c)$ ha una parametrizzazione locale regolare. Sia $\varphi: J \rightarrow A$ la parametrizzazione di una curva con supporto in $F^{-1}(c)$, per cui $F(\varphi(t)) = c$ per ogni $t \in J$. Differenziando in t otteniamo

$$\begin{aligned} \frac{d}{dt}F(\varphi(t)) &= F_x(\varphi(t))\varphi'_1(t) + F_y(\varphi(t))\varphi'_2(t) \\ &= a(\varphi(t))\varphi'_1(t) + b(\varphi(t))\varphi'_2(t) \end{aligned}$$

Poiché F è costante su $\varphi(t)$, la sua derivata è nulla, e otteniamo

$$\varphi^*\omega = 0,$$

cioè $\Gamma = \varphi(J)$ è soluzione di $\omega = 0$. Viceversa, se Γ è soluzione di $\omega = 0$, leggendo le uguaglianze al contrario otteniamo

$$\frac{d}{dt}F(\varphi(t)) = 0,$$

cioè $F(\varphi(t)) = c$ per ogni $t \in J$ e Γ è contenuta in un insieme di livello di F . \square

Come conseguenza immediata della Proposizione si ha che i tratti regolari degli insiemi di livello di F sono tutte e sole le soluzioni di $\omega = 0$. In particolare, se $b \neq 0$, le curve di livello di F sono tutte cartesiane, per la Proposizione 3.2.6.

Riassumiamo la strategia per integrare (3.5) alla luce del nuovo formalismo. Possiamo associare alla (3.5) la forma differenziale ω : la Proposizione 3.2.6 garantisce l'equivalenza tra le soluzioni di $\omega = 0$ e di (3.5). Siamo interessati allora alle soluzioni di $\omega = 0$. Se ω è esatta la Proposizione 3.2.9 appena provata garantisce che le soluzioni di $\omega = 0$ giacciono sulle curve di livello del potenziale e che, in particolare, i tratti regolari delle curve di livello del potenziale sono tutte e sole le curve soluzione di $\omega = 0$. Otteniamo quindi una relazione implicita soddisfatta dalle soluzioni di (3.5), del tipo

$$U(t, x(t)) = c,$$

dove la costante c è determinata dal dato iniziale. Se riusciamo ad esplicitare questa relazione, otteniamo in forma esplicita le soluzioni di (3.5). Da notare che, per determinare le effettive soluzioni di (3.5), bisognerà tenere in conto la eventuale restrizione imposta da $b \neq 0$.

Facciamo un esempio di come funziona questo metodo risolutivo.

Esempio 3.2.10. Determiniamo in forma implicita la soluzione di

$$x' = -\frac{x \cos t + 2te^x}{\sin t + 2 + t^2e^x}.$$

La forma associata è

$$\begin{aligned}\omega &= a(t, x)dt + b(t, x)dx \\ &= (x \cos t + 2te^x)dt + (\sin t + 2 + t^2e^x)dx\end{aligned}$$

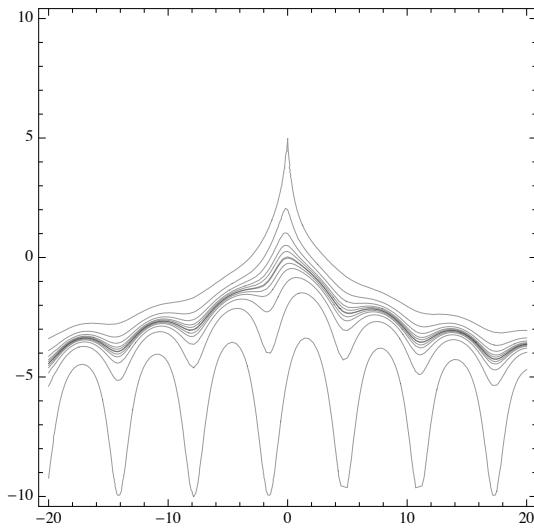


Figura 3.3: Grafico di alcune soluzioni dell'esempio refe:mering

definita su $A = \mathbb{R}^2$ semplicemente connesso. Dunque è esatta se e solo se è chiusa. Controlliamo che lo sia:

$$\partial_x(x \cos t + 2te^x) = \cos t + 2te^x = \partial_t(\sin t + 2 + t^2e^x).$$

Cerchiamo un potenziale $U = U(t, x)$. Se U è potenziale, si deve avere $U_x = a$ e integrando rispetto ad x si ottiene $U = x \sin t + 2x + t^2e^x + \psi(t)$. Affinché U sia primitiva dobbiamo imporre anche $U_t = b$, ottenendo $\psi'(t) = 0$. Si trova allora la primitiva

$$U(t, x) = x \sin t + 2x + t^2e^x.$$

Le soluzioni della nostra equazione si trovano sulle curve di livello di U , cioè se x è soluzione, soddisfa la relazione implicita

$$x(t) \sin t + 2x(t) + t^2e^{x(t)} = c,$$

dove la costante c sarà determinata dal dato iniziale imposto. Osserviamo infine che b non si annulla mai e allora i livelli di U sono tutte curve cartesiane, che sono tutti e soli i grafici delle soluzioni dell'equazione assegnata (si veda la Figura 3.3).

3.2.3 Metodo del fattore integrante

Torniamo all'equazione

$$x' = -\frac{a(t, x)}{b(t, x)} \quad (3.6)$$

con $a, b \in C(A)$ e sia

$$\omega = a(t, x)dt + b(t, x)dx$$

la forma associata. Abbiamo visto che se ω è esatta, le soluzioni di (3.6) giacciono nelle curve di livello del potenziale di ω . Supponiamo ora che ω non sia esatta, ma che esista una funzione $h \in C(A)$ tale che

1. $h(t, x) \neq 0$ in A ;
2. la forma $h\omega = h(t, x)a(t, x)dt + h(t, x)b(t, x)dx$ è esatta.

Una tale h si dice fattore integrante della forma ω . Vediamo quali informazioni riusciamo ad ottenere sulle soluzioni di (3.6). Abbiamo visto che le soluzioni di (3.6) sono curve soluzione di $\omega = 0$ e che, viceversa, le curve cartesiane che risolvono $\omega = 0$ sono soluzioni di (3.6). L'osservazione chiave è che l'equazione $\omega = 0$ ha esattamente le stesse soluzioni di $h\omega = 0$.

Proposizione 3.2.11. *Sia $h \in C(A, \mathbb{R})$, $h \neq 0$ in A . Allora $\omega = 0$ e $h\omega = 0$ hanno le stesse curve soluzione.*

Dimostrazione. Il pull-back di $h\omega$ tramite la curva $\varphi : J \rightarrow A$ è

$$\begin{aligned} \varphi^*(h\omega) &= h(\varphi)a(\varphi)\varphi'_1 + h(\varphi)b(\varphi)\varphi'_2 \\ &= h(\varphi)\varphi^*\omega \end{aligned}$$

quindi φ è soluzione di $\omega = 0$ se e solo se lo è per $h\omega = 0$, poiché h non si annulla A . \square

Ha quindi senso cercare un tale fattore integrante: le soluzioni della nostra equazione (3.6) sono contenute nelle curve di livello del potenziale di $h\omega$.

Quindi, se ω non è esatta, proveremo a determinare un fattore integrante. Supponendo che A sia semplicemente connesso, la condizione necessaria e sufficiente affinché $h\omega$ sia esatta è

$$\frac{\partial}{\partial x}(h(t, x)a(t, x)) = \frac{\partial}{\partial t}(h(t, x)b(t, x)).$$

In generale tuttavia, trovare un fattore integrante, posto che esista, è molto complicato. La richiesta sulla h sarebbe

$$h_x a + h a_x = h_t b + h b_t$$

ovvero

$$h(a_x - b_t) = h_t b - h_x a. \quad (3.7)$$

Un caso in cui le cose si semplificano è per esempio quello in cui

$$\frac{1}{b}(a_x - b_t) = g(t)$$

è funzione solo del tempo. Infatti in questo caso si può scegliere h indipendente da x e la (3.7) si diventa

$$h_t = g(t)h$$

che ha soluzione $h(t) = e^{G(t)}$, con G primitiva di g . Naturalmente abbiamo anche il caso simmetrico in x , cioè se accade che

$$\frac{1}{a}(a_x - b_t) = w(x)$$

è funzione solo di x , allora è possibile scegliere h tale che $h_t = 0$ e ottenere

$$h_x = -w(x)h$$

che ha soluzione $h(x) = e^{-W(x)}$, con W primitiva di w .

Esempio 3.2.12. Integriamo l'equazione

$$x' = \frac{x + tx^2}{t} \quad (3.8)$$

usando la tecnica del fattore integrante. La forma associata è

$$\omega = (x + tx^2)dt - tdx$$

che non è chiusa perché

$$a_x - b_t = \partial_x(x + tx^2) - \partial_t(-t) = 2tx + 2 \neq 0.$$

Tuttavia

$$\frac{1}{a}(a_x - b_t) = \frac{2tx + 2}{x + tx^2} = \frac{2}{x}$$

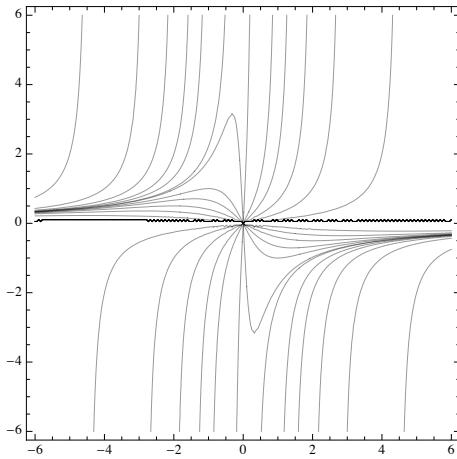


Figura 3.4: Soluzioni di (3.8)

è funzione solamente della x . Per quanto detto in precedenza, il fattore integrante è

$$h(x) = e^{-2 \log|x|} = \frac{1}{x^2}.$$

Dunque la funzione h non è definita per $x = 0$. Osserviamo che $x(t) \equiv 0$ è soluzione stazionaria dell'equazione, quindi possiamo supporre, senza perdita di informazione, che $x \neq 0$. Verifichiamo che effettivamente

$$h\omega = \left(\frac{1}{x} + t \right) dt - \frac{t}{x^2} dt$$

sia esatta. Poiché

$$\partial_x \left(\frac{1}{x} + t \right) = -\frac{1}{x^2} = \partial_t \left(-\frac{t}{x^2} \right)$$

la forma è chiusa. Cerchiamo un potenziale U . Derivando in t si ottiene

$$U_t(t, x) = \frac{1}{x} + t$$

e integrando

$$U(t, x) = \frac{t}{x} + \frac{t^2}{2} + c(x).$$

Derivando in x si ottiene

$$U_x(t, x) = -\frac{t}{x^2} + c'(x) \implies c'(x) = 0$$

cioè

$$U(t, x) = \frac{t}{x} + \frac{t^2}{2}.$$

Le soluzioni di (3.8) soddisfano quindi

$$\frac{t}{x(t)} + \frac{t^2}{2} = c$$

dove la costante c è determinata dal dato iniziale

$$c = \frac{t_0}{x_0} + \frac{t_0^2}{2}.$$

Possiamo esplicitare la x e ottenere la soluzione

$$x(t) = \frac{2t}{2c - t^2},$$

che sarà definita su un'opportuna semiretta (o segmento), in funzione delle condizioni iniziali.

3.2.4 Integrali primi

Il metodo di individuare in forma implicita le soluzioni di un'equazione differenziale ordinaria tramite la determinazione di una funzione che resta costante lungo tali soluzioni è piuttosto generale. Nel paragrafo precedente ne abbiamo visto l'applicazione al caso di equazioni esatte, eventualmente a meno di un fattore integrante. In questo paragrafo ne vogliamo presentare qualche altra applicazione. A tale scopo, introduciamo la nozione di integrale primo di una equazione differenziale.

Definizione 3.2.13. Sia $A \subseteq \mathbb{R}^n$ un aperto e $f : A \rightarrow \mathbb{R}^n$ un campo vettoriale. Una funzione $E : A \rightarrow \mathbb{R}$ di classe C^1 è un integrale primo per il sistema autonomo

$$x' = f(x) \tag{3.9}$$

se E è costante lungo le soluzioni di (3.9), cioè se per ogni soluzione $x \in C^1(I)$ di (3.9) esiste $c \in \mathbb{R}$ tale che $E(x(t)) = c$ per ogni $t \in I$.

Gli integrali primi sono caratterizzati dalla seguente proprietà geometrica:

Proposizione 3.2.14. *Sia f localmente lipschitziana e sia $E \in C^1(A)$. E è un integrale primo di (3.9) se e solo se*

$$\langle \nabla E(y), f(y) \rangle = 0, \text{ per ogni } y \in A. \quad (3.10)$$

Dimostrazione. Per il Teorema di derivazione delle funzioni composte, se la funzione $\varphi: I \rightarrow A$ è di classe C^1 si ha

$$\frac{d}{dt} E(\varphi(t)) = \langle \nabla E(\varphi(t)), \varphi'(t) \rangle, \text{ per ogni } t \in I.$$

Sia $y \in A$ fissato e sia $x \in C^1(I, A)$ l'unica soluzione di (3.9) con dato iniziale $x(0) = y$. Supponiamo che valga (3.10). Allora poiché $x' = f(x)$, si ha

$$\frac{d}{dt} E(x(t)) = 0, \text{ per ogni } t \in I,$$

cioè E è costante su x . Viceversa, se E è un integrale primo, allora $E(x(t)) = c$ per ogni $t \in I$ e in particolare

$$0 = \langle \nabla E(x(0)), x'(0) \rangle = \langle \nabla E(y), f(y) \rangle.$$

□

Esempio 3.2.15 (Sistemi conservativi). Consideriamo l'equazione

$$x'' = f(x) \quad \text{in } \mathbb{R}^n$$

corrispondente al sistema $2n$ -dimensionale

$$\begin{cases} x' = y, \\ y' = f(x). \end{cases}$$

Se f è un campo conservativo ed $U: \mathbb{R}^n \rightarrow \mathbb{R}$ un suo potenziale, cioè una funzione tale che $f = -\nabla U$, allora l'energia meccanica totale

$$E(x, y) = \frac{1}{2} \|y\|^2 + U(x)$$

è un integrale primo. Infatti

$$\nabla E(x, y) = (-f(x), y)$$

e allora

$$\langle \nabla E(x, y), (-f(x), y) \rangle = 0.$$

Integrali primi di sistemi autonomi bidimensionali. La teoria delle forme differenziali può essere utilizzata anche per la risoluzione di sistemi bidimensionali non lineari, come per esempio il Lotka-Volterra.

Consideriamo il sistema autonomo bidimensionale

$$\begin{cases} x' = a(x, y), \\ y' = b(x, y), \end{cases} \quad (3.11)$$

con $a, b \in C(A)$, A aperto di \mathbb{R}^n . Cerchiamo un integrale primo E di (3.11). Per la Proposizione 3.2.14 E deve verificare

$$\langle \nabla E(x, y), f(x, y) \rangle = 0, \quad f(x, y) = (a(x, y), b(x, y))$$

ovvero

$$\nabla E(x, y) = (-b(x, y), a(x, y)).$$

Quindi E è una primitiva della forma differenziale

$$\omega = -b(x, y)dx + a(x, y)dy.$$

Come conseguenza, E esiste se e solo se la forma ω è esatta. In realtà basta molto meno per trovare un integrale primo di (3.11): è sufficiente che ω sia esatta a meno di un fattore integrante. Supponiamo infatti che esista $\lambda : A \rightarrow \mathbb{R}$ di classe C^1 tale che sia non nulla in A e che

$$\tilde{\omega} = \lambda\omega = -\lambda(x, y)b(x, y)dx + \lambda(x, y)a(x, y)dy$$

sia esatta. Se U è una primitiva di $\tilde{\omega}$, allora

$$\nabla U(x, y) = \lambda(x, y)(-b(x, y), a(x, y))$$

e quindi

$$\langle \nabla E(x, y), f(x, y) \rangle = 0$$

cioè U è integrale primo di (3.11).

Esempio 3.2.16 (Lotka-Volterra). Utilizziamo quanto visto per trovare un integrale primo del sistema evolutivo preda-predatore introdotto nel Paragrafo 1.2

$$\begin{cases} x' = (\alpha - \beta y)x \\ y' = (-\gamma + \delta x)y \end{cases} \quad \alpha, \beta, \gamma, \delta > 0. \quad (3.12)$$

Consideriamo il caso interessante $x(t) > 0$, $y(t) > 0$. La forma differenziale associata a (3.12) è

$$\omega = (\gamma - \delta x)ydx + (\alpha - \beta y)x dy.$$

Come è facile verificare, ω non è chiusa, e quindi non è esatta. Cerchiamo un fattore integrante. Affinché $\tilde{\omega} = \lambda(x, y)\omega$ sia esatta nel primo quadrante, è sufficiente che sia chiusa, ovvero che valga

$$\partial_y(\lambda(x, y)(\gamma - \delta x)y) = \partial_x(\lambda(x, y)(\alpha - \beta y)x)$$

cioè

$$\lambda_y(\gamma - \delta x)y + \lambda(\gamma - \delta x) = \lambda_x(\alpha - \beta y)x + \lambda(\alpha - \beta y)$$

e quindi

$$(\gamma - \delta x)(\lambda_y y + \lambda) = (\alpha - \beta y)(\lambda_x x + \lambda). \quad (3.13)$$

Questa è una ricerca complicata in generale. Proviamo per esempio a supporre che $\lambda > 0$ sia tale che

$$\begin{cases} \lambda_y y + \lambda = 0 \\ \lambda_x x + \lambda = 0 \end{cases}$$

in modo tale che la (3.13) sia verificata. La prima equazione è

$$\frac{\lambda_y}{\lambda} = -\frac{1}{y}$$

dalla quale per integrazione rispetto ad y

$$\log \lambda = -\log y + c(x).$$

Affinché λ verifichi anche la seconda equazione, è necessario che

$$\frac{\lambda_x}{\lambda} = c'(x) = -\frac{1}{x}$$

cioè $c(x) = -\log(x)$. Quindi il fattore integrante deve soddisfare

$$\log \lambda = -\log x - \log y$$

da cui

$$\lambda(x, y) = \frac{1}{xy}.$$

La forma $\tilde{\omega}$ è quindi

$$\tilde{\omega} = \left(\frac{\gamma}{x} - \delta \right) dx + \left(\frac{\alpha}{y} - \beta \right) dy$$

per $x, y > 0$. Per costruzione, $\tilde{\omega}$ è chiusa, e il primo quadrante è semplicemente connesso, per cui $\tilde{\omega}$ ammette un potenziale in tale regione, che sarà integrale

primo del sistema (3.12). Integriamo esplicitamente $\tilde{\omega}$. Il potenziale $U = U(x, y)$ è tale che

$$U_x = \frac{\gamma}{x} - \delta \implies U = \gamma \log x - \delta x + \psi(y).$$

Derivando in y otteniamo

$$U_y = \psi'(y) = \frac{\alpha}{y} - \beta \implies \psi(y) = \alpha \log y - \beta y$$

e quindi una primitiva è

$$U(x, y) = \gamma \log x - \delta x + \alpha \log y - \beta y$$

che è l'integrale primo che avevamo dichiarato in (1.3).

3.2.5 Equazioni lineari del primo ordine non omogenee

Sia $I \subset \mathbb{R}$ un intervallo aperto e $\alpha, \beta \in C(I, \mathbb{R})$. Consideriamo l'equazione differenziale del primo ordine non omogenea

$$x' = \alpha(t)x + \beta(t) \quad (3.14)$$

La forma differenziale associata a (3.14) è

$$\omega = a(t, x)dt + b(t, x)dx = (\alpha(t)x + \beta(t))dt - dx.$$

Poiché

$$a_x - b_t = \partial_x [\alpha(t)x + \beta(t)] - \partial_t [-1] = \alpha(t),$$

si ha che ω è esatta se e solo se $\alpha \equiv 0$, cioè se e solo se il problema (3.14) è quello di trovare una primitiva di β . Per trovare informazioni interessanti per il caso generale, cerchiamo un fattore integrante di ω . Osserviamo che

$$\frac{1}{b}(a_x - b_t) = -\alpha(t)$$

dipende solo da t , per cui il fattore integrante è

$$h(t) = e^{-\int_{t_0}^t \alpha(s) ds},$$

e la forma

$$h\omega = \left[e^{-\int_{t_0}^t \alpha(s) ds} (\alpha(t)x + \beta(t)) \right] dt - e^{-\int_{t_0}^t \alpha(s) ds} dx$$

è esatta. Poniamo

$$A(t) = \int_{t_0}^t \alpha(s) ds.$$

Calcoliamo un potenziale. Se U è potenziale di $h\omega$ allora $U_x = -e^{-A(t)}$ e integrando si ottiene $U = -xe^{-A(t)} + c(t)$. Derivando rispetto a t si ha $U_t = x\alpha(t)e^{-A(t)} + c'(t)$ e imponendolo uguale ad $h\omega$ si ottiene

$$c'(t) = \beta(t)e^{-A(t)} \implies c(t) = \int_{t_0}^t \beta(s)e^{-A(s)} ds$$

da cui

$$U(t, x) = -xe^{-A(t)} + \int_{t_0}^t \beta(s)e^{-A(s)} ds.$$

Tale U essendo primitiva di $h\omega$, è anche integrale primo di (3.14) ed è allora costante lungo le soluzioni. La costante è determinata dal dato iniziale: $U(t, x(t)) = U(t_0, x_0)$ per ogni $t \in I$, dove $x_0 = x(t_0)$ e $t_0 \in I$ è fissato. Esplicitando la x otteniamo

$$x(t) = e^{A(t)} \left(x_0 + \int_{t_0}^t \beta(s)e^{-A(s)} ds \right), \quad A(t) = \int_{t_0}^t \alpha(s) ds \quad (3.15)$$

che è l'unica soluzione del problema di Cauchy per (3.14)

$$\begin{cases} x' = \alpha(t)x + \beta(t), \\ x(t_0) = x_0. \end{cases}$$

L'unicità segue dal Teorema 2.3.1 essendo $f(t, x) = a(t)x + b(t)$ continua in $I \times \mathbb{R}$ e localmente lipschitziana in x uniformemente in t .

Esempio 3.2.17. Consideriamo il problema di Cauchy associato all'equazione lineare del primo ordine non omogenea

$$\begin{cases} x' = x \tan t + \sin t, \\ x(t_0) = x_0, \end{cases} \quad t_0 \neq \mp \frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z}.$$

Sia $k_0 \in \mathbb{Z}$ tale che $t_0 \in I = \left(-\frac{\pi}{2} + k_0\pi, \frac{\pi}{2} + k_0\pi\right)$, in modo tale che $\alpha(t) = \tan t$ sia definita in I , e conseguentemente anche la soluzione.

In I utilizziamo la rappresentazione (3.15):

$$\begin{aligned} A(t) &= \int_{t_0}^t \alpha(s) ds = \int_{t_0}^t \tan s ds = \int_{t_0}^t \frac{\sin s}{\cos s} ds \\ &= -\log \left| \frac{\cos t}{\cos t_0} \right| = -\log \left(\frac{\cos t}{\cos t_0} \right), \end{aligned}$$

$$\begin{aligned} \int_{t_0}^t \beta(s) e^{-A(s)} ds &= \frac{1}{\cos t_0} \int_{t_0}^t \cos s \sin s ds \\ &= \frac{1}{\cos t_0} \left(-\frac{\cos^2 t}{2} + \frac{\cos^2 t_0}{2} \right) \end{aligned}$$

e quindi la soluzione è

$$\begin{aligned} x(t) &= \frac{\cos t_0}{\cos t} \left(x_0 + \frac{1}{\cos t_0} \left(-\frac{\cos^2 t}{2} + \frac{\cos^2 t_0}{2} \right) \right) \\ &= x_0 \frac{\cos t_0}{\cos t} + \frac{1}{2 \cos t} (\cos^2 t_0 - \cos^2 t) \end{aligned}$$

definita per $t \in I = \left(-\frac{\pi}{2} + k_0 \pi, \frac{\pi}{2} + k_0 \pi \right)$.

3.3 Equazioni di Bernoulli

Sia $I \subseteq \mathbb{R}$ un intervallo e $P, Q \in C(I, \mathbb{R})$ funzioni assegnate. Consideriamo per $\alpha \in \mathbb{R}$, $\alpha \neq 0$, $\alpha \neq 1$, l'equazione differenziale di Bernoulli

$$x' = P(t)x + Q(t)x^\alpha. \quad (3.16)$$

Se $\alpha > 0$ si ha che $x \equiv 0$ è soluzione. Per determinare le altre soluzioni non nulle si può dividere per x^α in (3.16) ottenendo

$$x^{-\alpha} x' = P(t)x^{1-\alpha} + Q(t).$$

Posto $z(t) = x^{1-\alpha}(t)$ si ottiene l'equazione lineare del primo ordine non omogenea in z

$$z' = (1-\alpha)P(t)z + (1-\alpha)Q(t).$$

Possiamo allora determinare la z usando (3.15) ed ottenere la soluzione $x = z^{\frac{1}{1-\alpha}}$.

Osserviamo che, se $0 < \alpha < 1$ la funzione $f(t, x) = P(t)x + Q(t)x^\alpha$ non è localmente Lipschitziana in x intorno all'origine, quindi non è garantita l'unicità

della soluzione. Di fatto può accadere che la soluzione non nulla trovata si possa raccordare con la soluzione nulla. Questo non può accadere quando $\alpha > 1$, visto che in questo caso la funzione f è Lipschitziana in x uniformemente in t in ogni sottoinsieme compatto di $I \times \mathbb{R}$.

Esempio 3.3.1 (Equazione logistica a coefficienti variabili). Risolviamo con il metodo appena presentato l'equazione logistica a coefficienti variabili

$$p' = k(t)p - h(t)p^2$$

con $h, k \in C(\mathbb{R})$. Osserviamo che $p \equiv 0$ è soluzione. Come conseguenza del Teorema di esistenza e unicità locale, tutte le altre soluzioni non sono mai nulle. Possiamo cercare le soluzioni non nulle dividendo per p^2 ottenendo

$$p^{-2}p' = k(t)p^{-1} - h(t).$$

Sostituendo $z = p^{-1}$ si ottiene

$$z' = -k(t)z + h(t).$$

Sia ora K una primitiva di k . Allora si ha la soluzione esplicita

$$z(t) = e^{-K(t)} \left(c + \int e^{K(s)} h(s) ds \right)$$

e tornando alla $p = z^{-1}$,

$$p(t) = e^{K(t)} \left(c + \int e^{K(s)} h(s) ds \right)^{-1}.$$

3.4 Equazioni omogenee

Sono equazioni differenziali della forma

$$x' = f\left(\frac{x}{t}\right).$$

In questo caso è utile fare la sostituzione $z(t) = x(t)/t$, in modo tale che $x' = z + tz'$ e l'equazione soddisfatta dalla nuova funzione $z(t)$ sia

$$z' = \frac{f(z) - z}{t}$$

che è a variabili separabili.

Esempio 3.4.1. Torniamo al problema di Cauchy studiato nell'Esempio 3.2.1

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

La f si può scrivere come

$$f(t, x) = -\frac{1 + \frac{x}{t}}{1 - 3\frac{x}{t}}$$

e quindi si ottiene l'equazione per z

$$tz' = \frac{3z^2 - 2z - 1}{1 - 3z}.$$

Separiamo le variabili e integriamo

$$\begin{aligned} \frac{1 - 3z}{3z^2 - 2z - 1} z' &= \frac{1}{t} \implies -\frac{1}{2} \int \frac{6z - 2}{3z^2 - 2z - 1} dz = \log \left| \frac{t}{c} \right| \implies \\ -\frac{1}{2} \log |3z^2 - 2z - 1| &= \log \left| \frac{t}{c} \right| \implies |3z^2 - 2z - 1| = \frac{c^2}{t^2}. \end{aligned}$$

Tornando nella vecchia variabile si ha

$$\left| 3\frac{x^2}{t^2} - 2\frac{x}{t} - 1 \right| = \frac{c^2}{t^2} \implies |3x^2 - 2tx - t^2| = c^2$$

e imponendo il dato iniziale si trova $c^2 = 3$. Come nell'esempio, otteniamo che la soluzione x risolve l'equazione di secondo grado

$$3x^2 - 2tx - (t^2 + 3) = 0.$$

Capitolo 4

Teoria generale di base

4.1 Il Lemma di Gronwall

Iniziamo con un risultato tecnico che è alla base della teoria delle equazioni differenziali ordinarie, il Lemma di Gronwall, che tipicamente serve a trasformare disequazioni differenziali per una funzione u in stime sul suo decadimento. Quasi tutti i risultati quantitativi relativi alle soluzioni di un sistema di equazioni differenziali (convergenza di soluzioni approssimanti, dipendenza continua dai dati iniziali, ecc..) si baseranno su questo tipo di stime.

Teorema 4.1.1 (Lemma di Gronwall). *Sia $I \subseteq \mathbb{R}$ un intervallo, $t_0 \in I$ e siano $u, \alpha, \beta \in C(I, \mathbb{R}^+)$, dove $\mathbb{R}^+ = [0, +\infty)$, tali che*

$$u(t) \leq \alpha(t) + \left| \int_{t_0}^t \beta(s)u(s) ds \right|, \quad \text{per ogni } t \in I. \quad (4.1)$$

Allora

$$u(t) \leq \alpha(t) + \left| \int_{t_0}^t \alpha(s)\beta(s)e^{\left| \int_s^t \beta(\sigma) d\sigma \right|} ds \right|, \quad \text{per ogni } t \in I.$$

Dimostrazione. Introduciamo la funzione

$$v(t) = \int_{t_0}^t \beta(s)u(s) ds, \quad t \in I.$$

Per il Teorema fondamentale del calcolo integrale si ha che

$$v'(t) = \beta(t)u(t)$$

per cui, usando l'ipotesi (4.1), si ottiene una prima stima

$$v'(t) \leq \beta(t) [\alpha(t) + \text{sign}(t - t_0)v(t)], \quad \text{per ogni } t \in I. \quad (4.2)$$

Introduciamo la funzione

$$\gamma(t) = e^{-\left| \int_{t_0}^t \beta(s) ds \right|},$$

in modo tale che $\gamma'(t) = -\text{sign}(t - t_0)\beta(t)\gamma(t)$. Moltiplicando (4.2) per γ si ha

$$\gamma v' \leq \alpha\beta\gamma + \text{sign}(t - t_0)\beta v\gamma = \alpha\beta\gamma - \gamma'v,$$

cioè

$$(\gamma v)' \leq \alpha\beta\gamma.$$

Integriamo quest'ultima diseguaglianza tra t_0 e t :

$$\text{sign}(t - t_0) [v(t)\gamma(t) - v(t_0)\gamma(t_0)] \leq \text{sign}(t - t_0) \int_{t_0}^t \alpha(s)\beta(s)\gamma(s) ds.$$

Ricordando che $v(t_0) = 0$ e dividendo per γ otteniamo

$$\text{sign}(t - t_0)v(t) \leq \text{sign}(t - t_0) \int_{t_0}^t \alpha(s)\beta(s) \frac{\gamma(s)}{\gamma(t)} ds,$$

che, una volta osservato che

$$\frac{\gamma(s)}{\gamma(t)} = e^{\left| \int_s^t \beta(\sigma) d\sigma \right|},$$

ci da la tesi

$$u(t) \leq \alpha(t) + |v(t)| \leq \alpha(t) + \left| \int_{t_0}^t \alpha(s)\beta(s)e^{\left| \int_s^t \beta(\sigma) d\sigma \right|} ds \right|, \quad \text{per ogni } t \in I.$$

□

Enunciamo ora un semplice corollario, che sarà più pratico da utilizzare per i nostri scopi.

Corollario 4.1.2 (Lemma di Gronwall semplificato). *Nelle stesse ipotesi del Lemma di Gronwall, supponiamo anche che α sia della forma*

$$\alpha(t) = \alpha_0(|t - t_0|),$$

con $\alpha_0 : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ monotona crescente. Allora

$$u(t) \leq \alpha(t) e^{\left| \int_{t_0}^t \beta(\sigma) d\sigma \right|}, \text{ per ogni } t \in I.$$

Dimostrazione. Il Lemma di Gronwall garantisce che

$$u(t) \leq \alpha(t) + \left| \int_{t_0}^t \alpha(s) \beta(s) e^{\left| \int_s^t \beta(\sigma) d\sigma \right|} ds \right|$$

D'altra parte la monotonia di α comporta che $\alpha(s) \leq \alpha(t)$ se $|s - t_0| \leq |t - t_0|$ e quindi

$$\begin{aligned} u(t) &\leq \alpha(t) \left(1 + \left| \int_{t_0}^t \beta(s) e^{\left| \int_s^t \beta(\sigma) d\sigma \right|} ds \right| \right) \\ &\leq \alpha(t) \left(1 + \left[e^{\left| \int_s^t \beta(\sigma) d\sigma \right|} \right]_{t_0}^t \right) \\ &= \alpha(t) e^{\left| \int_{t_0}^t \beta(s) ds \right|}. \end{aligned}$$

□

4.2 Soluzioni approssimate

Il risultato di esistenza che dimostreremo nel Paragrafo 4.4 si basa su una tecnica di approssimazione e compattezza. A tal fine è utile introdurre una nozione di soluzione approssimata.

Definizione 4.2.1. Fissato $\varepsilon \geq 0$, una funzione $u : I(u) \rightarrow D$ è una soluzione ε -approssimata di $x' = f(t, x)$ se

1. $I(u) \subseteq I$ è un intervallo compatto e non banale di \mathbb{R} ;

2. $u \in C(I(u), \mathbb{R}^n)$ ed è C^1 a tratti in $I(u)$, ovvero esistono intervalli $\{I_j\}_{j=1}^m$, $I_j \subseteq I(u)$, a due a due disgiunti, tali che $\cup_{j=1}^m I_j = I(u)$ e $u \in C^1(\bar{I}_j, \mathbb{R}^n)$ per ogni $j = 1, \dots, m$;
3. in ogni intervallo $J \subseteq I(u)$ dove u è C^1 deve valere

$$\|u'(t) - f(t, u(t))\| \leq \varepsilon, \text{ per ogni } t \in J.$$

Ovviamente, una soluzione locale $u \in C^1(I(u), \mathbb{R}^n)$ dell'equazione differenziale è una soluzione ε -approssimata con $\varepsilon = 0$. Abbiamo quindi rilassato due condizioni rispetto alla definizione di soluzione classica che abbiamo introdotto in (2.1.4):

- meno regolarità;
- l'equazione non deve essere soddisfatta in maniera esatta, richiediamo solo che l'errore sia piccolo in norma.

Abbiamo visto che la formulazione integrale di un'equazione differenziale risulta particolarmente utile quando si vuole ottenere l'esistenza di una soluzione dell'equazione per approssimazione (nel caso del Teorema di esistenza ed unicità (2.3.1) ha motivato la costruzione delle iterate di Picard, per esempio). Il seguente risultato ci fornisce una formulazione integrale approssimata.

Proposizione 4.2.2. *Se $u : I(u) \rightarrow D$ è una soluzione ε -approssimata di $x' = f(t, x)$, allora*

$$\left\| u(t) - u(t_0) - \int_{t_0}^t f(s, u(s)) ds \right\| \leq \varepsilon |t - t_0|, \text{ per ogni } t, t_0 \in I(u).$$

Dimostrazione. Lavoriamo con $t > t_0$, essendo il caso $t < t_0$ del tutto analogo. Sia quindi u una soluzione ε -approssimata di $x' = f(t, x)$, e sia

$$t_0 = s_0 < s_1 < \cdots < s_m = t,$$

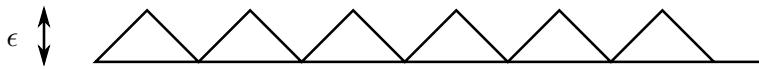


Figura 4.1: Controesempio

una partizione dell’intervallo $[t_0, t] \subseteq I(u)$ tale che u ristretta all’intervallo $[s_i, s_{i+1}]$ appartenga a $C^1([s_i, s_{i+1}], \mathbb{R}^n)$, per ogni $i = 0, 1, \dots, m - 1$. In questi intervalli vale il Teorema fondamentale del calcolo integrale

$$u(s_{i+1}) - u(s_i) = \int_{s_i}^{s_{i+1}} u'(s) \, ds.$$

Quindi si ha che

$$\begin{aligned} u(t) - u(t_0) - \int_{t_0}^t f(s, u(s)) \, ds &= \sum_{i=0}^{m-1} \left[u(s_{i+1}) - u(s_i) - \int_{t_0}^t f(s, u(s)) \, ds \right] \\ &= \sum_{i=0}^{m-1} \int_{s_i}^{s_{i+1}} [u'(s) - f(s, u(s))] \, ds \end{aligned}$$

da cui, utilizzando il fatto che u è una soluzione ε -approssimata, si ricava

$$\left\| u(t) - u(t_0) - \int_{t_0}^t f(s, u(s)) \, ds \right\| \leq \sum_{i=0}^{m-1} \varepsilon (s_{i+1} - s_i) = \varepsilon |t - t_0|.$$

□

Nel caso di soluzioni ε -approssimate la formulazione integrale non equivale alla formulazione differenziale: ad esempio è possibile costruire una soluzione ε -approssimata della formulazione integrale corrispondente a $f = 0$ e dato $u(t_0) = 0$ tale che $\|u'\| = 1$ (si veda Figura 4.1).

4.3 Funzioni lipschitziane e teorema di unicità locale

Sappiamo già che la lipschitzianità della f è un ingrediente fondamentale per garantire la buona posizione del problema di Cauchy. In questo paragrafo otterremo il risultato di unicità della soluzione in ipotesi di Lipschitzianità in maniera del tutto indipendente rispetto al quanto fatto nel Paragrafo 2.3, come conseguenza diretta del Lemma di Gronwall.

In realtà il risultato che dimostreremo è molto più ricco di informazioni perché fornisce una stima precisa di quanto due soluzioni ε -approssimate

dell'equazione differenziale (e quindi, in particolare, due soluzioni) possano allontanarsi una dall'altra in ipotesi di lipschitzianità della f .

Lemma 4.3.1. *Sia $f \in C(A, \mathbb{R}^n) \cap \text{Lip}^x(A, \mathbb{R}^n)$, $A \in \mathbb{R} \times \mathbb{R}^n$ aperto, e sia L la costante di Lipschitz di f in A . Siano $u : I(u) \rightarrow A$, $v : I(v) \rightarrow A$ due soluzioni ε -approssimate di $x' = f(t, x)$ di parametri ε_1 ed ε_2 rispettivamente. Allora per ogni $t_0 \in I(u) \cap I(v)$ vale la stima*

$$\|u(t) - v(t)\| \leq (\|u(t_0) - v(t_0)\| + (\varepsilon_1 + \varepsilon_2)|t - t_0|) e^{L|t-t_0|},$$

per ogni $t \in I(u) \cap I(v)$.

Dimostrazione. Aggiungendo e togliendo i termini opportuni ed utilizzando la diseguaglianza triangolare otteniamo

$$\begin{aligned} \|u(t) - v(t)\| &\leq \left\| u(t) - u(t_0) - \int_{t_0}^t f(s, u(s)) ds \right\| \\ &+ \left\| v(t) - v(t_0) - \int_{t_0}^t f(s, v(s)) ds \right\| \\ &+ \|(u(t_0) - v(t_0))\| + \left\| \int_{t_0}^t [f(s, u(s)) - f(s, v(s))] ds \right\| \end{aligned}$$

Grazie alla Proposizione 4.2.2 e alla stima di Lipschitzianità per f arriviamo alla diseguaglianza

$$\|u(t) - v(t)\| \leq (\varepsilon_1 + \varepsilon_2)|t - t_0| + \|u(t_0) - v(t_0)\| + L \left| \int_{t_0}^t \|u(s) - v(s)\| ds \right|.$$

Applicando il Lemma di Gronwall 4.1.2 a $z(t) = \|u(t) - v(t)\|$, definita per $t \in I(u) \cap I(v)$, $\alpha_0(|t - t_0|) = (\varepsilon_1 + \varepsilon_2)|t - t_0| + \|u(t_0) - v(t_0)\|$ e $\beta = L$, otteniamo la tesi. \square

La stima precedente ha come semplice corollario il risultato di unicità locale in ipotesi di lipschitzianità.

Teorema 4.3.2 (Unicità locale). *Sia $f \in C(A, \mathbb{R}^n) \cap \text{Lip}_{\text{loc}}^x(A, \mathbb{R}^n)$, $A \in \mathbb{R} \times \mathbb{R}^n$ aperto, $(t_0, x_0) \in A$ e siano u e v due soluzioni del problema di Cauchy*

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

Allora $u = v \in I(u) \cap I(v)$.

Dimostrazione. Dal momento che l'asserto da provare è di tipo puntuale, è sufficiente provarlo in un intervallo $I \subseteq I(u) \cap I(v)$ compatto in \mathbb{R} che contenga t_0 . Poiché u e v sono continue, le immagini $u(I)$ e $v(I)$ dell'intervallo compatto I sono anch'esse sottoinsiemi compatti di \mathbb{R}^n , come anche $K = u(I) \cup v(I)$. Poiché $f \in \text{Lip}_{\text{loc}}^x(A, \mathbb{R}^n)$ e $I \times K$ è un compatto, si ha che $f \in \text{Lip}^x(I \times K, \mathbb{R}^n)$ e quindi vale la stima del Lemma 4.3.1 con $\varepsilon_1 = \varepsilon_2 = 0$, e $u(t_0) = v(t_0)$ che implica banalmente che $u = v$ in I . \square

4.4 Il teorema di Peano

Questo paragrafo è dedicato alla dimostrazione del teorema di esistenza di Peano per il sistema di equazioni differenziali $x' = f(t, x)$. La dimostrazione sarà di tipo costruttivo: per ogni $\varepsilon > 0$ costruiremo delle soluzioni ε -approssimate (le spezzate di Eulero) da cui, grazie alle stime a priori, riusciremo ad estrarre una successione che converga uniformemente. La funzione limite risulterà, per passaggio al limite sulle formulazioni integrali approssimate, essere soluzione esatta del sistema di equazioni differenziali.

Data $f \in C(A, \mathbb{R}^n)$, A aperto, e $(t_0, x_0) \in A$ vogliamo costruire una soluzione di

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0. \end{cases} \quad (4.3)$$

Determiniamo a , b ed M come nel Paragrafo 2.2 e fissiamo $\delta = \min(a, b/M)$ (che corrisponderà al raggio dell'intorno di t_0 in cui la soluzione sarà definita).

Osserviamo che, essendo R compatto, per il Teorema di Heine-Cantor la restrizione $f|_E$ è uniformemente continua: fissato $\varepsilon > 0$, esiste un $\sigma > 0$ tale che

$$\|f(t_1, x_1) - f(t_2, x_2)\| < \varepsilon$$

per ogni $(t_1, x_1), (t_2, x_2) \in R$ tali che $|t_1 - t_2| < \sigma$ e $\|x_1 - x_2\| < \sigma$.

Costruzione delle spezzate di Eulero. Fissiamo un $\varepsilon > 0$, consideriamo il σ datoci dalla uniforme continuità di f in E e prendiamo una partizione dell'intervallo I

$$t_0 - \delta < t_{-n} < t_{-n+1} < \cdots < t_{-1} < t_0 < t_1 < \cdots < t_{n-1} < t_n = t_0 + \delta$$

in modo tale che

$$\max_i |t_{i-1} - t_i| \leq \min \left\{ \sigma, \frac{\sigma}{M} \right\}.$$

Definiamo la funzione u_ε ricorsivamente a partire da t_0

$$u_\varepsilon(t) = \begin{cases} u_\varepsilon(t_i) + (t - t_i)f(t_i, u(t_i)), & \text{se } t \in [t_i, t_{i+1}], i \geq 0, \\ u_\varepsilon(t_{i+1}) + (t - t_{i+1})f(t_{i+1}, u(t_{i+1})), & \text{se } t \in [t_i, t_{i+1}], i \leq -1. \end{cases} \quad (4.4)$$

Essendo una spezzata che esce da x_0 e che ha pendenza massima M , al massimo si può allontare da x_0 di $M\delta \leq b$, cioè $u_\varepsilon \in C([t_0 - \delta, t_0 + \delta], \overline{B}_b(x_0))$ e vale la stima di lipschitzianità

$$\|u_\varepsilon(t) - u_\varepsilon(s)\| \leq M |t - s|, \text{ per ogni } s, t \in [t_0 - \delta, t_0 + \delta]. \quad (4.5)$$

Inoltre, per costruzione si ha

$$u'_\varepsilon(t) = f(t_i, u_\varepsilon(t_i)),$$

per ogni $t \in [t_i, t_{i+1}] \cap [t_0, +\infty)$ e $t \in [t_{i-1}, t_i] \cap (-\infty, t_0]$ e anche

$$\|u_\varepsilon(t) - u_\varepsilon(t_i)\| = \|(t - t_i)f(t_i, u_\varepsilon(t_i))\| \leq \min \left\{ \sigma, \frac{\sigma}{M} \right\} M \leq \sigma.$$

Inoltre, dal momento che $|t - t_i| < \sigma$ e dalla (4.5)

$$\|u_\varepsilon(t_i) - u_\varepsilon(t)\| \leq M |t - t_i| \leq M \min \left\{ \sigma, \frac{\sigma}{M} \right\} \leq \sigma.$$

quindi, per la uniforme continuità di f , concludiamo che

$$\|u'_\varepsilon(t) - f(t, u_\varepsilon(t))\| = \|f(t_i, u_\varepsilon(t_i)) - f(t, u_\varepsilon(t))\| < \varepsilon. \quad (4.6)$$

Compattezza e passaggio al limite. La stima (4.6) ci garantisce che la famiglia $\{u_\varepsilon\}$ delle spezzate di Eulero definite da (4.4) è in realtà una famiglia di soluzioni ε -approssimate di $x' = f(t, x)$. L'idea è che questa famiglia converga (almeno su di una sottosuccessione) ad una soluzione del problema di Cauchy (4.3). Per provare questo ci serve un qualche criterio di compattezza. Il criterio giusto nello spazio funzionale in cui ci troviamo è il Teorema di Ascoli-Arzelà (per la dim.

Teorema 4.4.1 (Ascoli-Arzelà). *Sia $[a, b]$ un intervallo compatto in \mathbb{R} e sia $u_k : [a, b] \rightarrow \mathbb{R}^n$ una successione di funzioni con le proprietà*

1. *per ogni $\varepsilon > 0$, esiste un $\delta > 0$ tale che*

$$\|u_k(t) - u_k(s)\| \leq \varepsilon,$$

per ogni $t, s \in [a, b]$ tali che $|t - s| < \delta$, per ogni $k \in \mathbb{N}$ (equicontinuità);

2. *esiste una costante $A > 0$ tale che $\|u_k(t)\| \leq A$, per ogni $t \in [a, b]$, $k \in \mathbb{N}$ (equilimitatezza).*

Allora $\{u_k\}$ ammette una sottosuccessione uniformemente convergente.

Sia $k \in \mathbb{N}$ e indichiamo con u_k la spezzata di Eulero corrispondente a $\varepsilon = \frac{1}{k}$. La stima (4.5) è indipendente da ε , e quindi ci dice che la successione $\{u_k\}$ è equilipschitziana in $[t_0 - \delta, t_0 + \delta]$; in particolare è equicontinua. Inoltre le u_k hanno valore fissato in t_0 : $u_k(t_0) = x_0$ per ogni $k \in \mathbb{N}$. Questo fatto, insieme alla equilipschitzianità, ci garantisce l'equilimitatezza:

$$\begin{aligned} \|u_k(t)\| &\leq \|u_k(t_0)\| + \|u_k(t) - u_k(t_0)\| \leq \|u_k(t_0)\| + M|t - t_0| \\ &\leq \|x_0\| + M\delta \leq \|x_0\| + b. \end{aligned}$$

Siamo allora nelle ipotesi del Teorema di Ascoli-Arzelà: esiste una funzione $u \in C([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$ e una estratta $\{u_{k_j}\}$ che converge uniformemente in $C([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$ ad u . Poiché le u_{k_j} sono soluzioni $\frac{1}{k_j}$ -approssimate, per la Proposizione 4.2.2 si ha

$$\left\| u_{k_j}(t) - x_0 - \int_{t_0}^t f(s, u_{k_j}(s)) ds \right\| \leq \frac{1}{k_j} |t - t_0|.$$

Inoltre la successione di funzioni $g_j(s) = f(s, u_{k_j}(s))$ converge uniformemente $C([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$ alla funzione $g(s) = g_j(s) = f(s, u_{k_j}(s))$, grazie all'uniforme continuità di f in $[t_0 - \delta, t_0 + \delta] \times \overline{B}_b(x_0)$. La convergenza uniforme ci permette di passare al limite sotto il segno di integrale ottenendo

$$\left\| u(t) - x_0 - \int_{t_0}^t f(s, u(s)) ds \right\| = 0,$$

cioè il limite u è soluzione della equazione integrale

$$u(t) = x_0 + \int_{t_0}^t f(s, u(s)) ds.$$

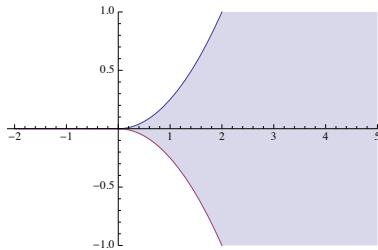


Figura 4.2: Zona coperta dal pennello di Peano nell’Esempio 3.1.2

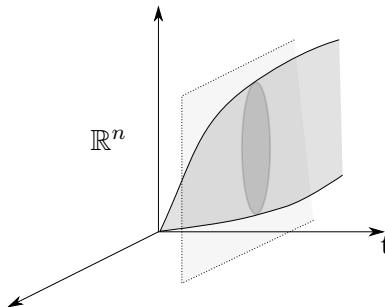
Abbiamo così dimostrato il Teorema di esistenza di Peano.

Teorema 4.4.2 (Teorema di esistenza di Peano). *Se $f \in C(I \times D, \mathbb{R}^n)$, allora il problema di Cauchy*

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

ammette almeno una soluzione $u \in C^1([t_0 - \delta, t_0 + \delta], \overline{B}_b(x_0))$.

Pennello di Peano. Nell’Esempio 3.1.2 abbiamo visto come, in caso di non unicità della soluzione per un’equazione differenziale scalare, possano uscire da un unico punto infinite soluzioni i cui grafici coprono un’intera regione del piano (si veda Figura 4.2). Questo fenomeno, che prende il nome di pennello di Peano, è sistematico. Più precisamente, utilizzando un metodo costruttivo simile a quello delle spezzate di Eulero, si possono costruire due soluzioni del problema di Cauchy (4.3) u_m e u_M che risultano essere rispettivamente la soluzione minima e quella massimale, ossia ogni u soluzione di (4.3) è tale che $u_m \leq u \leq u_M$ in $[t_0, t_0 + \delta]$. È possibile dimostrare che, in realtà, tutta la regione compresa tra i grafici di u_m e u_M viene coperta da grafici di soluzioni: per ogni (t, x) con $t \in [t_0, t_0 + \delta]$ e $u_m(t) \leq x \leq u_M(t)$ passa una soluzione di $x' = f(t, x)$ uscente da (t_0, x_0) . Questo fenomeno continua a valere anche per sistemi di equazioni differenziali: per ogni $t_0 < t_1 < t_0 + \delta$ l’intersezione tra l’iperpiano di \mathbb{R}^n $t = t_1$ e l’insieme delle traiettorie uscenti da (t_0, x_0) è un insieme compatto connesso non vuoto (per le dimostrazioni relative al fenomeno di Peano si veda, ad esempio, [6]).

Figura 4.3: Pennello di Peano n -dimensionale

Ovviamente, se vale l'ipotesi di Lipschitzianità, la soluzione minimale coincide con quella massimale ed il pennello semplicemente collassa in un'unica soluzione.

Con gli strumenti a nostra disposizione, possiamo anche avere una informazione sulla rapidità di convergenza delle spezzate di Eulero alla soluzione.

Teorema 4.4.3. *Se $f \in \text{Lip}^x(R, \mathbb{R}^n)$, allora l'intera successione delle spezzate di Eulero $\{u_k\}$ converge uniformemente in $[t_0 - \delta, t_0 + \delta]$ all'unica soluzione del problema di Cauchy*

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0. \end{cases} \quad (4.7)$$

Vale inoltre la stima

$$\|u_k(t) - u(t)\| \leq \frac{1}{k} |t - t_0| e^{L|t-t_0|},$$

per ogni $t \in [t_0 - \delta, t_0 + \delta]$, dove L è la costante di Lipschitz di $f|_R$.

Dimostrazione. Abbiamo già provato che le spezzate di Eulero hanno una sottosuccessione convergente ad una soluzione del problema di Cauchy (4.7). Dunque se proviamo che l'intera successione è convergente, necessariamente il suo limite sarà soluzione di (4.7). In ipotesi di lipschitzianità per le soluzioni ε -approssimate vale la stima del Lemma 4.3.1. Applichiamola alle $\{u_k\}$ e

otteniamo

$$\begin{aligned}\|u_k(t) - u_n(t)\| &\leq \left(\frac{1}{k} + \frac{1}{n} \right) |t - t_0| e^{L|t-t_0|} \\ &\leq \left(\frac{1}{k} + \frac{1}{n} \right) \delta e^{L\delta},\end{aligned}\tag{4.8}$$

cioè $\{u_k\}$ è una successione di Cauchy nello spazio metrico completo $C([t_0 - \delta, t_0 + \delta], \overline{B}_b(x_0))$, quindi è convergente.

Possiamo dare anche una dimostrazione alternativa del teorema, senza fare uso della completezza. Se infatti $\{u_k\}$ non fosse convergente, avrebbe due sottosuccessioni convergenti uniformemente a due limiti distinti u_1 e u_2 . Ma abbiamo visto che i punti limite delle spezzate di Eulero devono necessariamente essere soluzioni del problema di Cauchy (4.7), che, nelle nostre ipotesi, ammette un'unica soluzione, ottenendo un assurdo. \square

Equazioni scalari di ordine superiore. Abbiamo visto nell'Osservazione 2.1.3 che un'equazione scalare di ordine n del tipo

$$y^{(n)} = F(t, y, y', \dots, y^{(n-1)})\tag{4.9}$$

può essere ricondotta ad un sistema n -dimensionale del primo ordine

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

ponendo

$$x = (y, y', \dots, y^{(n-1)}), \quad f(t, x) = (t, y', \dots, y^{(n-1)}, F(t, y, y', \dots, y^{(n-1)})).$$

Vogliamo illustrare cosa diventa il risultato di esistenza ed unicità locale in ipotesi di Lipschitzianità per le equazioni del tipo (4.9). Innanzi tutto il problema di Cauchy associato corrisponde ad assegnare i valori della soluzione e delle sue prime $n - 1$ derivate in t_0 : dato $x_0 = (y_0, y_0^1, \dots, y_0^{n-1}) \in \mathbb{R}^n$ deve valere

$$x(t_0) = (y(t_0), y'(t_0), \dots, y^{(n-1)}(t_0)) = (y_0, y_0^1, \dots, y_0^{n-1}).$$

A questo punto basta osservare che se F è continua in A ed è localmente lipschitziana in A uniformemente in t (come funzione a valori reali), allora anche la funzione a valori vettoriali f gode delle stesse proprietà. Quindi se ne conclude che per ogni $t_0 \in I$ e per ogni $(y_0, y_0^1, \dots, y_0^{n-1}) \in D$ esiste un'unica

soluzione locale del problema di Cauchy

$$\begin{cases} y^{(n)} = F(t, y, y', \dots, y^{(n-1)}) \\ y(t_0) = y_0 \\ y'(t_0) = y_0^1 \\ \vdots \\ y^{(n-1)}(t_0) = y_0^{n-1} \end{cases}$$

Regolarità della soluzione. Concludiamo questa parte riguardante le soluzioni locali con un risultato di regolarità.

Teorema 4.4.4. *Se $f \in C^k(I \times D, \mathbb{R}^n)$, con $k \in \mathbb{N}$, allora la soluzione del problema di Cauchy (4.7) è di classe $C^{k+1}([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$.*

Dimostrazione. Per induzione su $k \in \mathbb{N}$. Se $k = 0$ il Teorema di Peano ci assicura che la soluzione $u \in C^1([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$. Supponiamo che l'implicazione sia vera per k e proviamo che vale anche per $k + 1$. Se $f \in C^{k+1}$, in particolare $f \in C^k$ e, per ipotesi induttiva, si ha $u \in C^{k+1}$. Ma allora $f(\cdot, u(\cdot)) \in C^{k+1}$, per cui se u' è tale che $u'(t) = f(t, u(t))$, allora $u' \in C^{k+1}$, cioè $u \in C^{k+2}$. \square

Il Teorema di regolarità appena provato vale anche nel caso analitico: se il dato di f è analitico, la soluzione di (4.7) è analitica. Non proveremo questo risultato. Per una dimostrazione, si veda [6], Paragrafo 2.2.

4.5 Prolungamenti

In tutto questo paragrafo ci metteremo nelle ipotesi dei Teoremi 4.3.2 e 4.4.2 in modo da avere esistenza e unicità locale delle soluzioni. Il nostro prossimo obiettivo è quello vedere se e quanto sia possibile estendere le soluzioni locali. Ricordiamo la Definizione 2.3.4 introdotta nel Capitolo 2.

Definizione 4.5.1. Sia $u : I(u) \rightarrow \mathbb{R}^n$ una soluzione di $x' = f(t, x)$. Allora $v : I(v) \rightarrow \mathbb{R}^n$ è un prolungamento di u se

- $I(u) \subseteq I(v)$ e v è soluzione di $x' = f(t, x)$ in $I(v)$;
- $u = v$ in $I(u)$.

La funzione u si dice soluzione massimale di $x' = f(t, x)$ se non ammette prolungamenti non banali, cioè se per ogni v prolungamento di u , si ha $I(v) = I(u)$.

Il prolungamento massimale esiste sempre ed è unico.

Teorema 4.5.2 (Esistenza ed unicità del prolungamento massimale). *Se tutti i problemi di Cauchy per $x' = f(t, x)$ ammettono un'unica soluzione, allora ogni soluzione u ammette un prolungamento massimale univocamente individuato. Inoltre, se v è la soluzione massimale, il suo intervallo di definizione $I(v) = (T_{\min}, T_{\max})$ è aperto.*

Dimostrazione. Sia u soluzione di $x' = f(t, x)$ in $I(u)$. Introduciamo l'insieme dei prolungamenti di u

$$P = \{y: I(y) \rightarrow \mathbb{R}^n; y \text{ soluzione di } x' = f(t, x), I(u) \subseteq I(y)\}.$$

Osserviamo che $P \neq \emptyset$ poiché contiene u come prolungamento banale. Costruiamo il prolungamento massimale di u definendo

$$I(\tilde{u}) = \bigcup_{y \in P} I(y), \quad \tilde{u}(t) := y(t), \text{ se } t \in I(y).$$

Si ha che:

- La \tilde{u} è ben definita perché se $t \in I(y) \cap I(z)$, necessariamente $y(t) = z(t)$, essendo unica la soluzione dei problemi di Cauchy per $x' = f(t, x)$;
- La \tilde{u} è soluzione di $x' = f(t, x)$, perché per $\bar{t} \in I(\tilde{u})$, esiste una $y \in P$ tale che $\tilde{u}(t) = y(t)$ in un intorno di \bar{t} . Poiché, per definizione di P , y è soluzione, avremo $x'(\bar{t}) = f(\bar{t}, x(\bar{t}))$.
- \tilde{u} è prolungamento massimale della u . Infatti, per come è definita \tilde{u} , ogni altro prolungamento y di u ha $I(y) \subseteq I(\tilde{u})$.

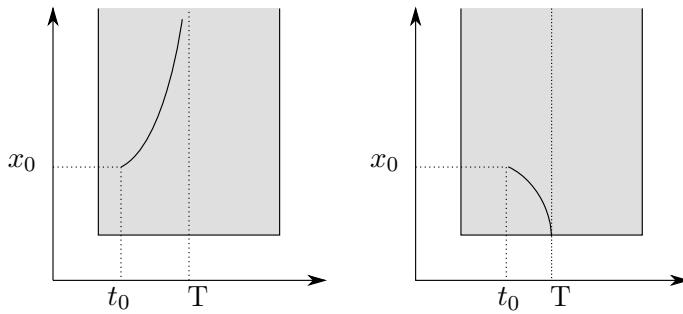


Figura 4.4: Esplosione in tempo finito e uscita dal dominio

Resta da dimostrare che l'intervallo di esistenza massimale $I_{max} = (T_{min}, T_{max})$ è aperto. Per assurdo supponiamo, ad esempio, che $T_{max} \in I_{max}$. Quindi la soluzione massimale \tilde{u} è definita in T_{max} e soddisfa l'equazione in quel punto. D'altra parte, il punto $(T_{max}, \tilde{u}(T_{max}))$ appartiene all'aperto A e quindi il Teorema di esistenza ed unicità assicura l'esistenza di una soluzione v del problema di Cauchy relativo alla condizione iniziale $x(T_{max}) = \tilde{u}(T_{max})$ in un intervallo del tipo $(T_{max} - \delta, T_{max} + \delta)$, $\delta > 0$. La funzione che si ottiene raccordando \tilde{u} con v in T_{max} sarebbe quindi un prolungamento non banale di \tilde{u} , in contraddizione con l'ipotesi di massimalità. \square

Cerchiamo ora di capire quando diventa impossibile prolungare la soluzione. Partiamo dal dato iniziale (t_0, x_0) e ragioniamo solo in avanti per $t \geq t_0$ (all'indietro è analogo). Prolunghiamo la soluzione locale e indichiamo con T il tempo massimale oltre al quale non si riesce ad andare. Cosa può essere successo alla soluzione in T ? Come si vede dalla Figura 4.4, potrebbe accadere per esempio che

- la soluzione tocca il bordo del dominio di definizione della f , e quindi non ha più senso l'equazione;
- la soluzione diverge per $t \rightarrow T^-$, cioè si ha esplosione in tempo finito.

Proveremo che questi sono esattamente tutti e soli i motivi per cui la soluzione non è ulteriormente prolungabile. Vediamo prima qualche esempio.

Esempio 4.5.3. Consideriamo il problema di Cauchy

$$\begin{cases} x' = x^2, \\ x(0) = 1. \end{cases}$$

La funzione $f(t, x) = x^2$ è C^∞ , quindi localmente lipschitziana, il che garantisce esistenza e unicità locale. Inoltre $x \equiv 0$ è soluzione costante, quindi tutte le altre soluzioni sono o strettamente positive o strettamente negative. In particolare, la soluzione del nostro problema sarà positiva. Essendo l'equazione a variabili separabili, riusciamo ad ottenere una soluzione esplicita. Dividiamo per x^2 e integriamo tra 0 e t :

$$\begin{aligned} \frac{x'}{x^2} = 1 &\implies \int_0^t \frac{x'(s)}{x^2(s)} ds = t \implies -\frac{1}{x(t)} + 1 = t \implies \\ &\frac{1}{x(t)} = 1 - t. \end{aligned}$$

Essendo $x > 0$, dobbiamo avere $1 - t > 0$, cioè la soluzione al problema è $x(t) = \frac{1}{1-t}$, definita per $t < 1$. In questo esempio il motivo per cui la soluzione non è ulteriormente prolungabile è l'esplosione in tempo finito.

Esempio 4.5.4. Possiamo generalizzare l'esempio precedente, considerando

$$\begin{cases} x' = x^{1+\varepsilon} \\ x(0) = 1 \end{cases}$$

con $\varepsilon > 0$. Con lo stesso procedimento, possiamo ottenere la soluzione per separazione di variabili e osservare che l'intervallo massimale è limitato a destra. Invece, se $\varepsilon = 0$, la soluzione è $x(t) = e^t$, che è globale. Sorge il sospetto che, per poter avere soluzioni globali, si debba richiedere che f non cresca più che linearmente nelle x (ne ripareremo nel Teorema 4.5.11).

Esempio 4.5.5. Sia D il disco unitario di centro l'origine nel piano

$$D = \{(t, x) \in \mathbb{R}^2 : t^2 + x^2 < 1\}$$

e consideriamo la funzione

$$f(t, x) = \frac{1}{\sqrt{1 - t^2 - x^2}}$$

e il relativo problema di Cauchy

$$\begin{cases} x' = \frac{1}{\sqrt{1 - t^2 - x^2}}, \\ x(0) = 0. \end{cases}$$

Poiché $f \in C^1(D)$, è garantita l'esistenza e l'unicità locale della soluzione. Osserviamo poi che $f > 0$ in D e quindi si ha $x' > 0$, cioè la soluzione è strettamente crescente. Proseguiamo nell'analisi qualitativa calcolandoci la derivata seconda:

$$x'' = f_x(t, x)x' + f_t(t, x) = \frac{xx'}{(1-t^2-x^2)^{\frac{3}{2}}} + \frac{t}{(1-t^2-x^2)^{\frac{3}{2}}}.$$

Il denominatore della frazione è sempre positivo, mentre per il numeratore abbiamo

$$xx' + t > 0, \quad \text{se } t > 0,$$

$$xx' + t < 0, \quad \text{se } t < 0,$$

dove abbiamo usato il fatto che $x' > 0$ sempre e, essendo $x(0) = 0$, $x(t) > 0$ per $t > 0$ e $x(t) < 0$ per $t < 0$. Quindi la soluzione è convessa per $t > 0$ ed è concava per $t < 0$. Grazie a queste informazioni, possiamo avere una descrizione qualitativa della soluzione. A causa della stretta monotonia la soluzione si estinguerebbe nel momento in cui toccherà il bordo di D , poiché la f non è più definita oltre la circonferenza unitaria.

Come anticipato, questi sono tutti e soli i possibili tipi di controsenpi alla prolungabilità delle soluzioni. Si ha infatti il seguente teorema.

Teorema 4.5.6 (Comportamento delle soluzioni massimali). *Sia $u : I(u) \rightarrow D$ una soluzione massimale di*

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

con $f : A \rightarrow \mathbb{R}^n$, dove $A = I \times D$. Sia poi $I(u)^+ = [t_0, T]$. Allora o $T = +\infty$, cioè u è soluzione globale in avanti, oppure

$$\lim_{t \rightarrow T^-} \left(\|u(t)\| + \frac{1}{d((t, u(t)), \partial A)} \right) = +\infty. \quad (4.10)$$

Ricordiamo che se $E \subseteq \mathbb{R}^n$, $z \in \mathbb{R}^n$, la distanza di z da E è definita da

$$d(z, E) = \inf_{w \in E} \|w - z\|,$$

con la convenzione $d(z, \emptyset) = +\infty$ per ogni $z \in \mathbb{R}^n$. Si verifica facilmente che, se $E \neq \emptyset$, la funzione $z \mapsto d(z, E)$ è lipschitziana in \mathbb{R}^n . Dimostriamo il Teorema.

Dimostrazione. Supponiamo $T < +\infty$ e, ragionando per assurdo, supponiamo che il limite in (4.10) esista finito. Esistono allora $M \geq 0$ e una successione di tempi $\{t_k\} \subset [t_0, T]$, con $t_k \rightarrow T^-$, tali che

$$\|u(t_k)\| + \frac{1}{d((t_k, u(t_k)), \partial A)} \leq M.$$

Ma allora si avrebbe

$$\|u(t_k)\| \leq M, \quad d((t_k, u(t_k)), \partial A) > \frac{1}{M}.$$

Essendo $\{u(t_k)\}$ una successione di \mathbb{R}^n limitata, per il Teorema di Bolzano-Weierstrass esiste una sua estratta $u(t_{k_j})$ convergente ad un punto $\bar{x} \in \mathbb{R}^n$. Per continuità della funzione distanza d si ha anche

$$d((T, \bar{x}), \partial A) \geq \frac{1}{M}$$

cioè il punto (T, \bar{x}) è interno ad A . Quindi è possibile determinare a, b tali che il cilindro $R = [T - a, T + a] \times \overline{B}_b(\bar{x})$ sia tutto contenuto in A , sia $M = \max_R |f(t, x)|$ e definiamo $\sigma = \min\{a/2, b/(M + 1)\}$. Fissiamo $k \in \mathbb{N}$ tale che

$$T - t_k < \sigma, \quad \|u(t_k) - \bar{x}\|.$$

Mostriamo che il cilindro $\overline{R} = [t_k - \sigma, t_k + \sigma] \times \overline{B}_{\sigma M}(u(t_k))$ è contenuto in R . Infatti si ha

$$\begin{aligned} |t - T| &\leq |t - t_k| + |t_k - T| \leq 2\sigma < a \quad \forall t \in [t_k - \sigma, t_k + \sigma] \\ \|x - \bar{x}\| &\leq \|x - u(t_k)\| + \|u(t_k) - \bar{x}\| \leq \sigma M + \sigma \leq b \quad \forall x \in \overline{B}_{\sigma M}(u(t_k)) \end{aligned}$$

Poiché si ha $\overline{M} = \max_{\overline{R}} |f(t, x)| \leq M$, la soluzione v del problema di Cauchy

$$\begin{cases} x' = f(t, x), \\ x(t_k) = u(t_k) \end{cases}$$

è definita in tutto l'intervallo $[t_k - \sigma, t_k + \sigma]$. Una volta osservato che $t_k + \sigma > T$, si ottiene che la funzione definita come

$$w(t) = \begin{cases} u(t) & t \in [t_0, t_k] \\ v(t) & t \in (t_k, t_k + \sigma] \end{cases}$$

fornisce un prolungamento di u non banale, in contraddizione con la massimalità di u . \square

Una questione di sicuro interesse è individuare delle condizioni sufficienti affinché una soluzione di un'equazione differenziale sia definita per tutti i tempi.

Definizione 4.5.7. Sia $f \in C(I \times D, \mathbb{R}^n)$, $I \subset \mathbb{R}$ intervallo aperto, D sottoinsieme aperto di \mathbb{R}^n . Una diremo che la funzione u è una soluzione globale di $x' = f(t, x)$ se $I(u) = I$.

L'idea di fondo contenuta nel Teorema 4.5.6 è che una soluzione è prolungabile fino a quando resta limitata e ben contenuta nel dominio di lipschitzianità di f . Come conseguenza immediata otteniamo il seguente risultato di esistenza globale.

Corollario 4.5.8. Sia $f \in C(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n) \cap \text{Lip}_{\text{loc}}^x(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$ e sia u la soluzione massimale del problema di Cauchy

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

Supponiamo che esista una funzione $\alpha \in C(\mathbb{R})$ tale che

$$\|u(t)\| \leq \alpha(t), \quad \text{per ogni } t \in [t_0, T_{\max}).$$

Allora u è globale, cioè $T_{\max} = +\infty$.

Dimostrazione. Poiché il dominio di f è $\mathbb{R} \times \mathbb{R}^n$, si ha che il secondo termine nel limite in (4.10) è nullo. D'altra parte, se per assurdo fosse $T_{\max} < +\infty$, per ipotesi

$$\limsup_{t \rightarrow T_{\max}^-} \|u(t)\| \leq \alpha(T_{\max}) < +\infty$$

quindi deve essere $T_{\max} = +\infty$. □

Esempio 4.5.9. (Sistemi conservativi) Abbiamo già introdotto nell'esempio 3.2.15 i sistemi conservativi

$$x'' = f(x) \tag{4.11}$$

con $f(x) = -\nabla U(x)$, $U \in C^1(\mathbb{R}^n, \mathbb{R})$. Sappiamo che è possibile scrivere l'equazione come sistema del primo ordine

$$\begin{cases} x' = y, \\ y' = f(x) \end{cases} \quad (4.12)$$

e che è possibile associare a questo sistema una funzione, l'energia

$$E(x, y) = \frac{\|y\|^2}{2} + U(x),$$

che risulta essere costante lungo le soluzioni. Siamo ora in grado di dare condizioni sufficienti sul potenziale U affinché le soluzioni di (4.11) siano globali.

Proposizione 4.5.10. *Se esiste una costante c_0 tale che $U(x) \geq c_0$ per ogni $x \in \mathbb{R}^n$, allora tutte le soluzioni di (4.11) sono globali.*

Dimostrazione. Sia x una soluzione massimale di (4.11) e sia $[t_0, T_{max})$ il suo intervallo di definizione in avanti e sia $c \in \mathbb{R}$ tale che

$$\frac{\|x'(t)\|^2}{2} + U(x(t)) = c.$$

Poiché $U(x) \geq c_0$, otteniamo la stima

$$\|x'(t)\| \leq 2\sqrt{c - c_0}$$

e quindi

$$\|x(t)\| = \|x(t_0) + x'(\tau)(t - t_0)\| \leq \|x(t_0)\| + 2\sqrt{c - c_0}(t - t_0).$$

La tesi ora segue dal Corollario 4.5.8. □

Il Corollario 4.5.8 ci dice che se la soluzione rimane limitata nel tempo, allora è globale. A questo punto è naturale domandarsi quali siano delle condizioni sufficienti sulla f affinché la soluzione del problema di Cauchy sia limitata. Abbiamo visto nell'Esempio 4.5.4 che la crescita della f in x (sublineare o superlineare) era un fattore determinante. In generale, vale il seguente risultato.

Teorema 4.5.11 (di esistenza globale). *Supponiamo che esista $c > 0$ tale che*

$$\|f(t, x)\| \leq c(\|x\| + 1), \text{ per ogni } (t, x) \in I \times D.$$

Allora ogni soluzione è limitata. In particolare ogni soluzione massimale è globale, cioè il suo intervallo di definizione coincide con I .

Dimostrazione. Sia $u : I(u) \rightarrow D$ una soluzione. Passiamo alla formulazione integrale:

$$u(t) = u(t_0) + \int_{t_0}^t f(s, u(s)) ds, \text{ per ogni } t \in I(u).$$

Possiamo stimare

$$\begin{aligned} \|u(t)\| &\leq \|u(t_0)\| + \left| \int_{t_0}^t \|f(s, u(s))\| ds \right| \\ &\leq \|u(t_0)\| + c \left| \int_{t_0}^t (\|u(s)\| + 1) ds \right| \\ &= \|u(t_0)\| + c|t - t_0| + c \left| \int_{t_0}^t \|u(s)\| ds \right|. \end{aligned}$$

e applicando il Lemma di Gronwall 4.1.2 a $\|u(s)\|$ otteniamo

$$\|u(t)\| \leq (\|u(t_0)\| + c|t - t_0|) e^{c|t-t_0|}$$

che garantisce la limitatezza della soluzione per tutti i tempi. \square

Osservazione 4.5.12. Naturalmente la condizione di sublinearità della f si guarda bene dall'essere necessaria. Infatti se, a titolo di esempio, consideriamo l'equazione in \mathbb{R}

$$x' = e^x (\sin^2(x) - 1)$$

la funzione f evidentemente non ha crescita sublineare, ma tutte le soluzioni sono limitate. Infatti l'equazione ammette le soluzioni costanti $x \equiv \pi/2 + k\pi$, $k \in \mathbb{Z}$ e, poiché f è regolarissima e quindi siamo in condizioni di esistenza ed unicità per il problema di Cauchy, ogni altra soluzione corrispondente ad un dato iniziale $x(t_0) = x_0$ con $\pi/2 + k_0\pi < x_0 < \pi/2 + (k_0 + 1)\pi$, non potendo valicare le soluzioni costanti, resterà limitata nella striscia. In particolare le soluzioni sono globali.

Esempio 4.5.13 (Esistenza ed unicità globale per sistemi lineari). Si consideri il sistema lineare con $f(t, x) = A(t)x + b(t)$,

$$x' = A(t)x + b(t), \quad (4.13)$$

dove $A(t) = \{a_{ij}(t)\}$ è una matrice di funzioni, $b(t) = (b_1(t), \dots, b_n(t))$ è un vettore di funzioni, tali che $a_{ij}, b_i \in C(I)$, I intervallo aperto. Ogni soluzione massimale di (4.13) è globale.

Dimostrazione. Sia $J \subset I$ un intervallo compatto. Siano

$$M = \max_{t \in J} \|A(t)\| \quad m = \max_{t \in J} \|b(t)\|.$$

Allora f soddisfa alla stima di crescita del Teorema 4.5.11

$$\|f(t, x)\| \leq M \|x\| + m$$

e quindi ogni soluzione massimale in J è globale, cioè definita su tutto J . Per arbitrarietà di J si conclude che ogni soluzione massimale è definita su tutto I . \square

4.6 Dipendenza dai dati iniziali

Continuiamo ad utilizzare le notazioni dei paragrafi precedenti e mettiamoci nelle ipotesi del teorema di esistenza ed unicità per il problema di Cauchy. Possiamo pensare, fissando t_0 , la soluzione del problema di Cauchy al tempo t come funzione del dato iniziale x_0 :

$$x_0 \mapsto x(t, t_0, x_0) \quad (4.14)$$

con $x(t, t_0, x_0)$ soluzione massimale di

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

Rileggendo il Lemma 4.3.1 nel caso specifico di due soluzioni del problema di Cauchy corrispondenti a due dati iniziali distinti, otteniamo che

$$\|x(t, t_0, x_0) - x(t, t_0, x_1)\| \leq \|x_0 - x_1\| e^{L|t-t_0|}, \quad (4.15)$$

per ogni t in un opportuno intorno di t_0 . Quindi vale il seguente risultato.

Teorema 4.6.1. *Nelle ipotesi dei Teoremi 4.3.2 e 4.4.2 di esistenza e unicità locale delle soluzioni la funzione in (4.14) è Lipschitziana.*

Questa è un'informazione particolarmente interessante dal punto di vista applicativo: la dinamica è poco sensibile ad eventuali errori di misurazione sui dati iniziali, ossia piccole variazioni iniziali comportano solo piccole variazioni della dinamica.

In realtà, aggiungendo un po' di regolarità alla f , si ottiene dipendenza dal dato iniziale di classe C^1 .

Teorema 4.6.2. *Nelle ipotesi dei Teoremi 4.3.2 e 4.4.2 di esistenza e unicità locale delle soluzioni, supponiamo anche che tutte le derivate parziali di f rispetto ad x siano continue. Allora la funzione in (4.14) è differenziabile e tutte le derivate seconde miste $\partial^2 x_j / \partial t \partial x_0^k$ sono continue in $I \times D$. Inoltre l'applicazione $t \mapsto D_{x_0} x(t, t_0, x_0)$, definita da I a valori matriciali, è la soluzione calcolata al tempo t del problema di Cauchy (matriciale)*

$$\begin{cases} X' = D_x f(t, x(t, t_0, x_0)) X \\ X(t_0) = I \end{cases} \quad (4.16)$$

Omettiamo la dimostrazione del Teorema (che si trova, ad esempio, in [2]). Possiamo però mostrare che, se tutto è regolare, allora $D_{x_0} x(\cdot, t_0, x_0)$ soddisfa (4.16) (che prende il nome di equazione alle variazioni o equazione linearizzata).

Fissati $j, k = 1, \dots, n$ si ha

$$\begin{aligned} \frac{\partial}{\partial t} \left(\frac{\partial x_j}{\partial x_0^k}(t, t_0, x_0) \right) &= \frac{\partial}{\partial x_0^k} \left(\frac{\partial x_j}{\partial t}(t, t_0, x_0) \right) \\ &= \frac{\partial}{\partial x_0^k} (f_j(t, x(t, t_0, x_0))) \\ &= \sum_{h=1}^n \frac{\partial}{\partial x_h} (f_j(t, x(t, t_0, x_0))) \frac{\partial x_h}{\partial x_0^k}(t, t_0, x_0) \end{aligned}$$

dove la prima uguaglianza è conseguenza del Teorema di Schwarz sulle derivate parziali seconde, nella seconda è stata utilizzata l'equazione differenziale

soddisfatta da x e la terza è la formula di derivazione delle funzioni composte. Quello che abbiamo ottenuto è esattamente la versione per componenti dell'equazione differenziale in (4.16), per $X = D_{x_0}x(\cdot, t_0, x_0)$. La condizione iniziale $X(t_0) = I$ si ottiene semplicemente derivando la relazione $x(t_0, t_0, x_0) = x_0$ rispetto ad x_0 .

Capitolo 5

Sistemi di equazioni lineari

Fissiamo le notazioni e le ipotesi che utilizzeremo in tutto il capitolo, salvo avviso contrario. Con $M_n(\mathbb{R})$ indicheremo lo spazio delle matrici $n \times n$ a coefficienti in \mathbb{R} . Siano $I \subseteq \mathbb{R}$ un intervallo aperto, $A \in C(I, M_n(\mathbb{R}))$ una applicazione continua a valori nello spazio delle matrici, $b \in C(I, \mathbb{R}^n)$ un vettore di componenti continue, con

$$A(t) = \{a_{i,j}(t)\}_{i,j=1}^n, \quad b = \{b_i(t)\}_{i=1}^n.$$

Siamo interessati a studiare il sistema di equazioni differenziali

$$\begin{cases} x'_1 = a_{1,1}(t)x_1 + \cdots + a_{1,n}(t)x_n + b_1(t) \\ \vdots \\ x'_n = a_{n,1}(t)x_1 + \cdots + a_{n,n}(t)x_n + b_n(t) \end{cases}$$

che in forma vettoriale possiamo scrivere come

$$x' = A(t)x + b(t), \tag{5.1}$$

con $x \in C^1(I, \mathbb{R}^n)$ incognita. Abbiamo già dimostrato che il problema di Cauchy

$$\begin{cases} x' = A(t)x + b(t) \\ x(t_0) = x_0 \end{cases} \tag{5.2}$$

ammette unica soluzione globale per ogni dato iniziale $(t_0, x_0) \in I \times \mathbb{R}^n$ (si veda l'Esempio 4.5.13). Ciò vuol dire che l'integrale generale del sistema (5.1) è costituito da infiniti elementi con n gradi di libertà. Il primo problema che ci porremo sarà quello di mostrare che tale insieme gode di ottime proprietà di

struttura che, in diversi casi, ci permetteranno di scrivere in maniera esplicita le soluzioni.

All'analisi dettagliata delle soluzioni dei sistemi di equazioni differenziali lineari saranno dedicati, oltre a questo, i successivi quattro capitoli. L'interesse per questo argomento è dovuto non solo al fatto che siano molto ricorrenti nelle applicazioni, ma anche al fatto che permettano di avere informazioni locali sulle soluzioni dei sistemi non lineari, cosa di cui ci occuperemo nel Capitolo 10.

Nel seguito, quando parleremo della soluzione di (5.2) intenderemo sempre la sua soluzione globale.

5.1 Sistemi lineari omogenei

In questo paragrafo studieremo i sistemi lineari omogenei

$$x' = A(t)x. \quad (5.3)$$

Indicheremo con V l'insieme di tutte le soluzioni di (5.3). Poiché le soluzioni sono tutte globali, V è contenuto nello spazio $C^1(I, \mathbb{R}^n)$ che, ricordiamo per completezza, è uno spazio vettoriale reale con le operazioni

$$(u + v)(t) := u(t) + v(t), \text{ per ogni } t \in I, u, v \in C^1(I, \mathbb{R}^n),$$

$$(\lambda u)(t) := \lambda u(t), \text{ per ogni } t \in I, \lambda \in \mathbb{R}, u \in C^1(I, \mathbb{R}^n).$$

Definizione 5.1.1. Diremo che le funzioni $u_1, \dots, u_k \in C^1(I, \mathbb{R}^n)$ sono linearmente dipendenti se esiste un vettore non nullo $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$ tale che

$$\lambda_1 u_1(t) + \dots + \lambda_k u_k(t) = 0, \text{ per ogni } t \in I,$$

cioè se la combinazione lineare $\lambda_1 u_1 + \dots + \lambda_k u_k$ è la funzione nulla.

Diremo che $u_1, \dots, u_k \in C^1(I, \mathbb{R}^n)$ sono linearmente indipendenti se non sono linearmente dipendenti.

Si verifica immediatamente che la combinazione lineare di soluzioni di (5.3) è ancora soluzione di (5.3). Questo ci dice che V eredita la struttura di spazio vettoriale reale da $C^1(I, \mathbb{R}^n)$ e che, essendo chiuso rispetto alle operazioni di somma e prodotto per uno scalare, ne è un sottospazio vettoriale.

Fissati $(t_0, \xi) \in I \times \mathbb{R}^n$, indichiamo con $u = u(\cdot; t_0, \xi) \in V$ l'unica soluzione globale del problema di Cauchy

$$\begin{cases} x' = A(t)x, \\ x(t_0) = \xi. \end{cases}$$

Allora è ben definita l'applicazione ψ

$$\begin{aligned} \psi: \mathbb{R}^n &\longrightarrow V \\ \xi &\longmapsto u(\cdot; t_0, \xi) \end{aligned} \tag{5.4}$$

che ad ogni vettore ξ associa la soluzione del problema.

Si noti che, in linea di principio, l'applicazione ψ può essere sempre definita sull'insieme V delle soluzioni dei problemi di Cauchy relativi ad un qualsiasi sistema, purché sia garantita l'esistenza e unicità di tali soluzioni. Infatti se siamo in tali ipotesi, ad ogni $\xi \in \mathbb{R}^n$ è sempre possibile associare un'unica soluzione massimale $\psi(\xi) \in V$ (si ricordi il Teorema 4.5.2). Tuttavia, nel caso generale questa applicazione non fornisce informazioni di struttura su V . Invece, nel caso di sistemi lineari omogenei vale il seguente importante risultato.

Teorema 5.1.2 (Struttura dell'integrale generale). *L'applicazione ψ in (5.4) è un isomorfismo di spazi vettoriali tra \mathbb{R}^n e lo spazio V delle soluzioni dei problemi di Cauchy. Pertanto le soluzioni del sistema lineare omogeneo di equazioni differenziali (5.3) formano un sottospazio vettoriale di dimensione n di $C^1(I, \mathbb{R}^n)$.*

Dimostrazione. Innanzi tutto, dimostriamo che ψ è una applicazione lineare. Siano $\xi, \eta \in \mathbb{R}^n$ e $\lambda, \mu \in \mathbb{R}$, dobbiamo dimostrare che

$$\psi(\lambda\xi + \mu\eta) = \lambda\psi(\xi) + \mu\psi(\eta),$$

cioè che

$$u(\cdot; t_0, \lambda\xi + \mu\eta) = \lambda u(\cdot; t_0, \xi) + \mu u(\cdot; t_0, \eta).$$

Ma le due funzioni coincidono in t_0 e sono soluzioni di (5.3), per cui per il Teorema di esistenza e unicità, coincidono su tutto I , il che prova la linearità di ψ .

Inoltre, sempre per il risultato di unicità, $\psi(\xi) = 0$ implica $\xi = 0$, poiché il sistema lineare ammette la soluzione identicamente nulla per $\xi = 0$ e tutte le altre soluzioni non possono mai annullarsi. Quindi ψ , essendo lineare, è iniettiva. Infine, ψ è suriettiva perché se $u \in V$ è soluzione, si può scegliere $\xi = u(t_0)$ ottenendo che $\psi(\xi) = u$. Quindi ψ è un isomorfismo tra \mathbb{R}^n e V . \square

Osservazione 5.1.3. L'isomorfismo ψ introdotto in (5.4) è canonico, cioè non dipende dalla scelta delle basi nei due spazi.

Vista la struttura di spazio vettoriale n -dimensionale di V , sappiamo che per conoscere tutte le soluzioni di (5.3) è sufficiente conoscere una base di V . Il problema della determinazione dell'integrale generale si riduce quindi a trovare n soluzioni linearmente indipendenti di (5.3). Diamo un nome a tali basi.

Definizione 5.1.4. Un sistema di n soluzioni $\{u_1, \dots, u_n\}$ di (5.3) linearmente indipendenti si dice sistema fondamentale di soluzioni. La matrice X che ha per colonne i vettori colonna u_1, \dots, u_n ,

$$X(t) = (u_1(t) | \dots | u_n(t))$$

è detta matrice fondamentale. Se inoltre $X(t_0) = I$ matrice identità di $M_n(\mathbb{R})$, X si dice matrice fondamentale speciale di (5.3) al tempo t_0 .

Useremo la lettera U per denotare la matrice fondamentale speciale.

Osservazione 5.1.5. L'isomorfismo ψ trasforma basi di \mathbb{R}^n in basi di V . Si ha dunque una corrispondenza biunivoca tra basi di \mathbb{R}^n e matrici fondamentali: se $\{\xi_1, \dots, \xi_n\}$ è base di \mathbb{R}^n , allora $X = (\psi(\xi_1) | \dots | \psi(\xi_n))$ è matrice fondamentale; viceversa se $X = (x_1 | \dots | x_n)$ è matrice fondamentale, allora $\{x_1(t_0), \dots, x_n(t_0)\}$ è la corrispondente base di \mathbb{R}^n . La base standard di \mathbb{R}^n , $\{e_1, \dots, e_n\}$, identifica la matrice fondamentale speciale U .

Esempio 5.1.6 (Oscillatore armonico). Ricordiamo che l'equazione del moto di un oscillatore libero è data da

$$x'' = -\omega^2 x, \quad \omega = \frac{k}{m} \tag{5.5}$$

dove m è la massa oscillante e k la costante elastica della molla. Possiamo scrivere l'equazione del secondo ordine come sistema lineare ponendo $x_1 = x$ e

$x_2 = x'$, ottenendo

$$\begin{cases} x'_1 = x_2 \\ x'_2 = -\omega^2 x_1 \end{cases}$$

cioè

$$x' = Ax$$

con $x = (x_1, x_2)$ e

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}.$$

Facciamo esplicitamente i conti su questo esempio per determinare l'insieme di tutte le soluzioni del sistema e l'isomorfismo ψ , per poi esibire la matrice fondamentale speciale. Abbiamo visto che

$$u(t) = A \cos(\omega t) + B \sin(\omega t)$$

è soluzione di (5.5) per ogni scelta di $A, B \in \mathbb{R}$. In termini delle nuove variabili, otteniamo le soluzioni

$$\begin{aligned} x_1(t) &= A \cos(\omega t) + B \sin(\omega t), \\ x_2(t) &= -\omega A \sin(\omega t) + \omega B \cos(\omega t). \end{aligned}$$

Alla luce di quanto detto finora, queste sono tutte le soluzioni del sistema scritte come combinazioni lineari del sistema fondamentale di soluzioni

$$(\cos(\omega t), -\omega \sin(\omega t)) \quad (\sin(\omega t), \omega \sin(\omega t)).$$

Se allora consideriamo i dati iniziali $x_1(0) = \xi_1$, $x_2(0) = \xi_2$ otteniamo $A = \xi_1$ e $B = \xi_2/\omega$. Allora, l'applicazione $\psi : \mathbb{R}^2 \rightarrow V$ manda $\xi = (\xi_1, \xi_2)$ nella soluzione

$$u(t; 0, \xi) = (x_1(t; 0, \xi), x_2(t; 0, \xi))$$

con

$$\begin{aligned} x_1(t; 0, \xi) &= \xi_1 \cos(\omega t) + \frac{\xi_2}{\omega} \sin(\omega t), \\ x_2(t; 0, \xi) &= -\xi_1 \omega \sin(\omega t) + \xi_2 \cos(\omega t). \end{aligned}$$

Dunque, se $\{\xi, \eta\}$ è una base di \mathbb{R}^2 , la matrice fondamentale relativa a questa base è

$$X(t) = (u_1(t) | u_2(t)),$$

con $u_1(t; 0, \xi) = (x_1(t; 0, \xi), x_2(t; 0, \xi))$, $u_2(t; 0, \eta) = (x_1(t; 0, \eta), x_2(t; 0, \eta))$. Alla base standard di \mathbb{R}^2 corrispondono quindi le soluzioni

$$\begin{aligned} u_1(t; 0, e_1) &= (\cos(\omega t), -\omega \sin(\omega t)) \\ u_2(t; 0, e_2) &= \left(\frac{1}{\omega} \sin(\omega t), \cos(\omega t) \right) \end{aligned}$$

e la matrice fondamentale corrispondente è

$$U(t) = \begin{pmatrix} \cos(\omega t) & \frac{1}{\omega} \sin(\omega t) \\ -\omega \sin(\omega t) & \cos(\omega t) \end{pmatrix},$$

che sappiamo essere la matrice fondamentale speciale. Infatti

$$U(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

Verifichiamo, per completezza, che i vettori u_1 e u_2 siano effettivamente linearmente indipendenti, calcolando il determinante di U :

$$\det U(t) = \cos^2(\omega t) + \sin^2(\omega t) = 1, \quad \text{per ogni } t \in \mathbb{R}.$$

Quindi effettivamente le soluzioni sono linearmente indipendenti e formano una base di V .

La matrice fondamentale speciale al tempo $t_0 \in I$ è tutto quello che serve per la descrizione della dinamica ed è, a sua volta, soluzione (matriciale) di un'equazione differenziale, nel senso che vale l'identità

$$X' = A(t)X. \tag{5.6}$$

Mostriamolo nel caso bidimensionale: siano u e v due soluzioni indipendenti di (5.3) e sia

$$X(t) = \begin{pmatrix} u_1(t) & v_1(t) \\ u_2(t) & v_2(t) \end{pmatrix}$$

la relativa matrice fondamentale. Allora si ha

$$\begin{aligned} X'(t) &= \begin{pmatrix} u_1(t)' & v_1(t)' \\ u_2(t)' & v_2(t)' \end{pmatrix} \\ &= \begin{pmatrix} a_{1,1}u_1(t) + a_{1,2}u_2(t) & a_{1,1}v_1(t) + a_{1,2}v_2(t) \\ a_{2,1}u_1(t) + a_{2,2}u_2(t) & a_{2,1}v_1(t) + a_{2,2}v_2(t) \end{pmatrix} \\ &= \begin{pmatrix} a_{1,1}(t) & a_{1,2}(t) \\ a_{2,1}(t) & a_{2,2}(t) \end{pmatrix} \begin{pmatrix} u_1(t) & v_1(t) \\ u_2(t) & v_2(t) \end{pmatrix} = A(t)X(t), \end{aligned}$$

che corrisponde all'identità cercata. Il caso di dimensione n qualsiasi si dimostra in maniera del tutto analoga.

Viceversa, risulta del tutto evidente che se $X \in M_n(\mathbb{R})$ risolve l'equazione matriciale $X' = A(t)X$, allora le sue colonne sono soluzioni di (5.3).

In particolare, la matrice fondamentale speciale di (5.3) U è l'unica soluzione matriciale del problema lineare

$$\begin{cases} X' = A(t)X, \\ X(t_0) = I. \end{cases} \quad (5.7)$$

Di conseguenza, per ogni $\xi \in \mathbb{R}^n$, la funzione $x(t) = U(t)\xi$ è l'unica soluzione del problema di Cauchy

$$\begin{cases} x' = A(t)x, \\ x(t_0) = \xi. \end{cases}$$

Dunque abbiamo ricondotto il problema della risoluzione di (5.3) al problema della ricerca della matrice fondamentale speciale. In realtà, è sufficiente conoscere una qualsiasi matrice fondamentale per ottenere la matrice fondamentale speciale. Infatti, se X è una matrice fondamentale e $C \in M_n(\mathbb{R})$ è una matrice costante invertibile, allora anche XC è una matrice fondamentale (abbiamo semplicemente operato un cambiamento di base in V). In particolare $Y(t) = X(t)X^{-1}(t_0)$ è una matrice fondamentale e $Y(t_0) = I$, quindi $Y(t)$ è la matrice fondamentale speciale $U(t)$.

In particolare, se X e Y sono due matrici fondamentali, valgono simultaneamente le due identità

$$\begin{aligned} U(t) &= X(t)X(t_0)^{-1} \\ U(t) &= Y(t)Y(t_0)^{-1} \end{aligned}$$

da cui segue la relazione

$$Y(t) = X(t)X(t_0)^{-1}Y(t_0).$$

Osserviamo che $X(t_0)$ e $Y(t_0)$ sono le matrici che corrispondono ai cambiamenti di coordinate tra la base standard $\{e_1, \dots, e_n\}$ e le basi $\{x_1(t_0), \dots, x_n(t_0)\}$ e $\{y_1(t_0), \dots, y_n(t_0)\}$, rispettivamente. Quindi il termine di conversione, dato da $X(t_0)^{-1}Y(t_0)$, corrisponde esattamente al cambiamento di coordinate in \mathbb{R}^n che trasforma la base $\{y_1(t_0), \dots, y_n(t_0)\}$ in $\{x_1(t_0), \dots, x_n(t_0)\}$.

Osservazione 5.1.7. Data la matrice fondamentale $X = (x_1 | \dots | x_n)$, possiamo ricavare la relazione $U(t) = X(t)X^{-1}(t_0)$ in maniera più esplicita, seguendo

la strategia proposta nell'Esempio 5.1.6. Partiamo dal fatto che ogni soluzione x di (5.3) si scrive come combinazione lineare

$$x(t) = a_1 x_1(t) + \cdots + a_n x_n(t), \text{ per ogni } t \in I$$

e in particolare in t_0 si ha

$$x(t_0) = a_1 x_1(t_0) + \cdots + a_n x_n(t_0).$$

Poiché $\{x_1(t_0), \dots, x_n(t_0)\}$ è una base di \mathbb{R}^n , al variare dei coefficienti a_j , $j = 1, \dots, n$, il valore di $x(t_0)$ varia su tutto \mathbb{R}^n . A noi interessa trovare soluzioni y_1, \dots, y_n linearmente indipendenti e tali che

$$y_1(t_0) = e_1, \dots, y_n(t_0) = e_n.$$

Queste funzioni sono combinazione lineare delle x_1, \dots, x_n per cui abbiamo

$$\begin{cases} y_1(t) = a_{1,1} x_1(t) + \cdots + a_{1,n} x_n(t) \\ \vdots \\ y_n(t) = a_{n,1} x_1(t) + \cdots + a_{n,n} x_n(t) \end{cases} \quad \text{per ogni } t \in I. \quad (5.8)$$

Osserviamo che al tempo $t \in I$ fissato la k -esima riga di (5.8) non è altro che il sistema lineare $X(t)a_k = y(t)$ nelle incognite $a_k = (a_{k,1}, \dots, a_{k,n})$. In $t = t_0$ si deve avere $y_k(t_0) = e_k$ per cui si ottengono gli n sistemi lineari

$$X(t_0)a_k = e_k, \quad k = 1, \dots, n,$$

che hanno come soluzioni $a_k = X(t_0)^{-1}e_k$. Sostituendo in (5.8) si trova che le y_k soddisfano

$$y_k(t) = X(t)X(t_0)^{-1}e_k,$$

cioè y_k è la k -esima colonna della matrice $X(t)X(t_0)^{-1}$. Dunque

$$Y(t) = (y_1(t) | \dots | y_n(t)) = X(t)X(t_0)^{-1}, \quad t \in I.$$

Riassumiamo quanto provato in una proposizione.

Proposizione 5.1.8. *La matrice fondamentale speciale $U(t)$ al tempo t_0 del sistema (5.3) risolve il problema di Cauchy matriciale*

$$\begin{cases} X' = A(t)X, \\ X(t_0) = I \end{cases}$$

e, di conseguenza, per ogni $\xi \in \mathbb{R}^n$, $x(t) = U(t)\xi$ è la soluzione del sistema lineare omogeneo

$$\begin{cases} x' = A(t)x \\ x(t_0) = \xi. \end{cases}$$

Due matrici fondamentali X e Y di (5.3) sono legate dalla relazione $Y(t) = X(t)X(t_0)^{-1}Y(t_0)$ e in particolare la matrice fondamentale speciale U in t_0 è legata ad una qualsiasi altra matrice fondamentale X dalla relazione $U(t) = X(t)X(t_0)^{-1}$.

Esempio 5.1.9. Consideriamo di nuovo la situazione dell'Esempio 5.1.6. Vogliamo calcolare la matrice fondamentale speciale senza conoscere esplicitamente l'isomorfismo ψ . Supponiamo di conoscere due soluzioni linearmente indipendenti del sistema

$$\begin{cases} x'_1 = x_2, \\ x'_2 = -\omega^2 x_1. \end{cases}$$

Tali soluzioni sono per esempio

$$\begin{aligned} x_1(t) &= (\sin(\omega t), \omega \cos(\omega t)), \\ x_2(t) &= (\cos(\omega t), -\omega \sin(\omega t)). \end{aligned}$$

Dunque conosciamo la matrice fondamentale $X(t) = (x_1(t)|x_2(t))$. Questa non è la matrice fondamentale speciale in $t_0 = 0$ perché

$$X(0) = \begin{pmatrix} 0 & 1 \\ \omega & 0 \end{pmatrix}.$$

Per quanto detto, la matrice fondamentale speciale è $U(t) = X(t)X(0)^{-1}$. Un semplice calcolo ci fornisce la matrice

$$X(0)^{-1} = \begin{pmatrix} 0 & \frac{1}{\omega} \\ 1 & 0 \end{pmatrix}$$

e allora

$$\begin{aligned} U(t) &= X(t)X(0)^{-1} = \begin{pmatrix} \sin(\omega t) & \cos(\omega t) \\ \omega \cos(\omega t) & -\omega \sin(\omega t) \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\omega} \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\omega t) & \frac{1}{\omega} \sin(\omega t) \\ -\omega \sin(\omega t) & \cos(\omega t) \end{pmatrix} \end{aligned}$$

è la matrice fondamentale speciale in $t_0 = 0$, come abbiamo già calcolato.

Introduciamo ora la funzione determinante di una matrice fondamentale per il sistema (5.3).

Definizione 5.1.10. Sia X una matrice fondamentale per il sistema (5.3). La funzione $W : I \rightarrow \mathbb{R}$ definita da $W(t) = \det X(t)$ è detta wronskiano della matrice X .

Sappiamo che la dipendenza lineare in $C^1(I, \mathbb{R}^n)$ richiede che esista una combinazione lineare a coefficienti non tutti nulli che sia nulla puntualmente su tutto l'intervallo I . Nel caso di combinazioni lineari di soluzioni di un sistema lineare omogeneo di equazioni differenziali, questa richiesta equivale all'annullarsi della combinazione lineare in un solo punto dell'intervallo.

Proposizione 5.1.11. Una famiglia di soluzioni $\{x_1, \dots, x_n\}$ di $x' = A(t)x$ è linearmente dipendente in $C^1(I, \mathbb{R}^n)$ se e solo se esiste un $\tau \in I$ tale che $\{x_1(\tau), \dots, x_n(\tau)\}$ sono linearmente dipendenti in \mathbb{R}^n .

Dimostrazione. Supponiamo che esista $\tau \in I$ ed un vettore $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ non nullo tale che

$$\lambda_1 x_1(\tau) + \dots + \lambda_n x_n(\tau) = 0.$$

La funzione $\lambda_1 x_1 + \dots + \lambda_n x_n$, soluzione del sistema, si annulla in τ , quindi, per il risultato di unicità, deve coincidere con la soluzione del sistema corrispondente al dato iniziale $x(\tau) = 0$, cioè con la funzione identicamente nulla. Se ne conclude che le funzioni $\{x_1, \dots, x_n\}$ sono linearmente dipendenti in $C^1(I, \mathbb{R}^n)$. Il viceversa è banale. \square

Come conseguenza otteniamo il seguente risultato.

Corollario 5.1.12. Il wronskiano di una matrice fondamentale per il sistema $x' = A(t)x$ non si annulla mai oppure è identicamente nullo. Una famiglia $\{x_1, \dots, x_n\}$ è un sistema fondamentale di soluzioni per il sistema se e solo se il wronskiano della corrispondente matrice fondamentale è non nullo.

Dimostrazione. Per la Proposizione 5.1.11, se esiste $\tau \in I$ tale che $W(\tau) = 0$, allora $W(t) = 0$ per ogni $t \in I$, da cui segue la tesi. \square

Il Corollario 5.1.12 può essere ottenuto come conseguenza della seguente rappresentazione del wronskiano.

Teorema 5.1.13 (Liouville). *Sia X una matrice fondamentale di $x' = A(t)x$. Allora il wronskiano di X è soluzione dell'equazione lineare*

$$y' = \text{tr}(A(t))y.$$

Quindi vale la rappresentazione (nota come identità di Abel)

$$W(t) = W(t_0)e^{\int_{t_0}^t \text{tr } A(s) ds}, \text{ per ogni } t \in I$$

con $t_0 \in I$ fissato.

Dimostrazione. Ricordiamo che il determinante di una matrice è una funzione multilinear dei suoi vettori colonna, ossia la funzione $c_i \mapsto \det(c_1 | \dots | c_n)$ è una funzione lineare da \mathbb{R}^n in \mathbb{R} per ogni $i = 1, \dots, n$. Inoltre, ogni funzione lineare da \mathbb{R}^n in \mathbb{R} è rappresentabile attraverso un prodotto scalare. In questo caso, ciò corrisponde a dire che per ogni $i = 1, \dots, n$ esiste una funzione $\psi_i: \mathbb{R}^{(n-1)n} \rightarrow \mathbb{R}^n$ tale che $\det(c_1 | \dots | c_n) = \langle \phi_i(\hat{c}_i), c_i \rangle$, dove abbiamo indicato con $\hat{c}_i = (c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_n)$ l'elemento corrispondente a tutte le colonne della matrice tranne l' i -sima. In particolare si ha

$$\frac{d}{dc_{ij}} \det(c_1 | \dots | c_n) = \phi_i^j(\hat{c}_i) \quad i, j = 1, \dots, n$$

Quindi, utilizzando la formula di derivazione delle funzioni composte, otteniamo

$$\begin{aligned} W'(t) &= \frac{d}{dt} \det X(t) = \sum_{i=1}^n \langle \phi_i(\hat{x}_i), x'_i \rangle \\ &= \sum_{i=1}^n \det(x_1(t) | \dots | x'_i(t) | \dots | x_n(t)). \end{aligned} \tag{5.9}$$

Fissato $\tau \in I$, sia $U_\tau(t)$ la matrice fondamentale speciale al tempo τ in modo tale che

$$\frac{d}{dt} \det U(\tau) = \sum_{i=1}^n \det \left(e_1 | \dots | \sum_{j=1}^n a_{ij}(\tau) e_j | \dots | e_n \right) = \sum_{i=1}^n a_{ii}(\tau) = \text{tr}(A(\tau)).$$

D'altra parte, sappiamo che ogni altra matrice fondamentale $X(t)$ del sistema lineare è legata ad U_τ tramite la relazione $X(t) = U_\tau(t)X(\tau)$ e quindi

$$W'(t) = \frac{d}{dt} \det X(t) = \frac{d}{dt} \det(U_\tau(t)X(\tau)) = \det(X(\tau)) \frac{d}{dt} \det U_\tau(t),$$

che, valutata in $t = \tau$, restituisce

$$W'(\tau) = \text{tr}(A(\tau))W(\tau).$$

Per l'arbitrarietà di $\tau \in I$, otteniamo che W è soluzione in I dell'equazione lineare $y' = \text{tr}(A(t))y$ e la conseguente formula di rappresentazione. \square

Nonostante sia possibile, dal punto di vista teorico, ridurre il problema della risoluzione di un sistema lineare omogeneo $x' = A(t)x$ alla sola ricerca di una matrice fondamentale qualsiasi, nella pratica il più delle volte non siamo in grado determinare esplicitamente una base per lo spazio delle soluzioni. Una classe abbastanza ampia di sistemi lineari per cui si trovano esplicitamente tutte le soluzioni sono i sistemi a coefficienti costanti. A questo argomento saranno dedicati i capitoli 6 e 7.

Vediamo un esempio di un'altra particolare classe di sistemi che sono esattamente risolubili.

Esempio 5.1.14 (Sistemi triangolari). Consideriamo il sistema lineare omogeneo bidimensionale

$$\begin{cases} u' = p(t)u \\ v' = q(t)u + r(t)v \end{cases} \quad (5.10)$$

con p, q e r in $C(\mathbb{R})$. In questo caso la matrice associata

$$\begin{pmatrix} p(t) & 0 \\ q(t) & r(t) \end{pmatrix}$$

è triangolare inferiore e questo ci permette di risolvere il sistema a cascata: la prima equazione è lineare del primo ordine nella sola variabile u , quindi siamo in grado di determinarne tutte le soluzioni. A questo punto, sostituendo u nella

seconda equazione, nuovamente siamo in presenza di un'equazione lineare del primo ordine nella sola incognita v che siamo in grado di risolvere.

Procediamo in questa maniera per determinare due soluzioni linearmente indipendenti. Come prima soluzione, possiamo prendere $u \equiv 0$, che è soluzione della prima equazione e, sostituendola nella seconda equazione, ottenere $v' = r(t)v$ che è risolta, ad esempio, da $v(t) = e^{R(t)}$, dove $R(t)$ è una primitiva di $R(t)$. Quindi una prima soluzione di (5.10) è

$$x_1(t) = (x_1^1(t), x_1^2(t)) = (0, e^{R(t)}).$$

Un'altra soluzione della prima equazione è $u(t) = e^{P(t)}$, dove $P(t)$ è una primitiva di $p(t)$. Se la sostituiamo nella seconda equazione, otteniamo

$$v' = q(t)e^{P(t)} + r(t)v$$

equazione lineare non omogenea del primo ordine. Ricordando la formula di rappresentazione (3.15), concludiamo che una possibile soluzione è

$$\begin{aligned} v(t) &= e^{R(t)} \left(\int_0^t q(s)e^{P(s)}e^{-R(s)} ds \right) \\ &= \int_0^t q(s)e^{P(s)+R(t)-R(s)} ds. \end{aligned}$$

Abbiamo così determinato una seconda soluzione del sistema (5.10)

$$x_2(t) = (x_2^1(t), x_2^2(t)) = \left(e^{P(t)}, \int_0^t q(s)e^{P(s)+R(t)-R(s)} ds \right).$$

Verifichiamo che le due funzioni x_1 e x_2 siano un sistema fondamentale di soluzioni per (5.10). Calcolandone il Wronskiano, otteniamo

$$W(t) = \det \begin{pmatrix} x_1^1(t) & x_2^1(t) \\ x_1^2(t) & x_2^2(t) \end{pmatrix} = \det \begin{pmatrix} 0 & e^{P(t)} \\ e^{R(t)} & x_2^2(t) \end{pmatrix} = -e^{P(t)+R(t)}$$

e quindi $W(t) > 0$ per ogni $t \in I$. Abbiamo così ottenuto una matrice fondamentale

$$X(t) = \begin{pmatrix} 0 & e^{P(t)} \\ e^{R(t)} & x_2^2(t) \end{pmatrix}$$

e se sceglieremo le primitive $P(t)$ e $R(t)$ in modo tale che $P(0) = R(0) = 0$, avremo

$$X(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Possiamo ottenere la matrice fondamentale speciale del sistema tramite la relazione $U(t) = X(t)X(0)^{-1}$. Si ha che $X(0)^{-1} = X(0)$ e quindi

$$U(t) = \begin{pmatrix} 0 & e^{P(t)} \\ e^{R(t)} & x_2^2(t) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} e^{P(t)} & 0 \\ x_2^2(t) & e^{R(t)} \end{pmatrix}.$$

Ovviamente, più banalmente, era chiaro che per ottenere la matrice fondamentale speciale fosse sufficiente scambiare le due colonne di X .

5.2 Sistemi lineari non omogenei

Vogliamo ora studiare i sistemi lineari non omogenei

$$x' = A(t)x + b(t), \quad (5.11)$$

con $b \neq 0$. Partiamo da un'osservazione elementare: se $u, v \in C^1(I, \mathbb{R}^n)$ sono soluzioni di (5.11) allora $z = u - v$ è soluzione del sistema omogeneo corrispondente $x' = A(t)x$. Viceversa, se v è soluzione di (5.11) e w è soluzione dell'omogenea, allora $z = v + w$ è ancora soluzione di (5.11). Da questo segue il teorema di struttura dell'insieme delle soluzioni dei sistemi lineari non omogenei.

Teorema 5.2.1 (Struttura dell'integrale generale). *L'insieme delle soluzioni del sistema (5.11) è dato dal sottospazio affine*

$$v + V = \{v + w \mid w \in V\},$$

dove

1. $v \in C^1(I, \mathbb{R}^n)$ è una soluzione qualsiasi del sistema (5.11),
2. V è lo spazio vettoriale n -dimensionale delle soluzioni del sistema omogeneo $x' = A(t)x$.

Quindi il problema della determinazione di tutte le soluzioni di (5.11) si riduce alla ricerca di una sola di esse. Si possono allora tentare strade semplificate, per esempio scegliendo a priori la forma della soluzione da trovare.

In questa ottica, mostreremo che è sempre possibile, almeno a livello teorico, determinare una soluzione particolare di (5.11) attraverso il metodo di variazione delle costanti.

Metodo di variazione delle costanti. Abbiamo visto che per trovare tutte le soluzioni dell'equazione omogenea è necessario e sufficiente determinare una matrice fondamentale X e che la soluzione del problema di Cauchy

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

non è altro che la funzione

$$x(t) = X(t)X^{-1}(t_0)x_0$$

Quindi l'insieme delle soluzioni del sistema omogeneo è descritto da $X(t)c$, al variare di $c \in \mathbb{R}^n$. Proponiamoci di determinare una soluzione particolare dell'equazione (5.11) della forma $x(t) = X(t)c(t)$ con $c : \mathbb{R} \rightarrow \mathbb{R}^n$ da determinare.

Le opportune condizioni da imporre su $c(t)$ si ottengono calcolando la derivata di $x(t) = X(t)c(t)$ ed usando l'equazione matriciale soddisfatta da X . Abbiamo che

$$\begin{aligned} x'(t) &= X'(t)c(t) + X(t)c'(t) = A(t)X(t)c(t) + X(t)c'(t) \\ &= A(t)x(t) + X(t)c'(t) \end{aligned}$$

e quindi, imponendo che x sia soluzione di (5.11), otteniamo che deve essere

$$X(t)c'(t) = b(t) \implies c'(t) = X^{-1}(t)b(t)$$

e, in conclusione,

$$c(t) = \int_{t_0}^t X^{-1}(s)b(s) ds.$$

Quindi la matrice fondamentale X permette anche di determinare una soluzione del sistema non omogeneo data da

$$\begin{aligned} v(t) &= X(t) \int_{t_0}^t X^{-1}(s)b(s) ds \\ &= \int_{t_0}^t X(t)X^{-1}(s)b(s) ds. \end{aligned}$$

Riassumendo quanto fatto finora ed utilizzando il Teorema di struttura 5.2.1, abbiamo ottenuto il seguente risultato.

Teorema 5.2.2 (Formula di Duhamel). *La soluzione del problema di Cauchy*

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

è data da

$$x(t) = X(t)X^{-1}(t_0)x_0 + \int_{t_0}^t X(t)X^{-1}(s)b(s) ds \quad (5.12)$$

dove X è una qualsiasi matrice fondamentale per il sistema omogeneo associato $x' = A(t)x$.

Osservazione 5.2.3. Abbiamo già incontrato la formula di Duhamel nel Paragrafo 3.2.5. Ricordiamo infatti che la soluzione dell'equazione scalare

$$\begin{cases} x' = a(t)x + b(t) \\ x(t_0) = x_0 \end{cases}$$

è data dalla (3.15)

$$\begin{aligned} x(t) &= e^{A(t)} \left(x_0 + \int_{t_0}^t b(s)e^{-A(s)} ds \right) \\ &= e^{A(t)}x_0 + \int_{t_0}^t b(s)e^{A(t)-A(s)} ds \end{aligned} \quad (5.13)$$

con

$$A(t) = \int_{t_0}^t a(s) ds.$$

L'equazione omogenea associata a $x' = a(t)x + b(t)$ è $x' = a(t)x$, che ha come matrice fondamentale la funzione

$$X(t) = e^{A(t)}.$$

Allora $X^{-1}(t) = e^{-A(t)}$ e $X^{-1}(t_0) = 1$, per cui la (5.13) coincide con la formula di Duhamel. In questo senso, la formula di Duhamel è la naturale generalizzazione della formula risolutiva per le equazioni lineari scalari al caso dei sistemi.

Esempio 5.2.4. Determiniamo la soluzione del problema di Cauchy

$$\begin{cases} u' = v + t^4 \\ v' = -\frac{2}{t^2}u + \frac{2}{t}v + t^3 \\ u(2) = 1 \\ v(2) = 4 \end{cases}$$

definita sulla semiretta $t > 0$. In forma matriciale il problema è $x' = A(t)x + b(t)$ con $x = (u, v)$, $b(t) = (t^4, t^3)$ e

$$A = \begin{pmatrix} 0 & 1 \\ -\frac{2}{t^2} & \frac{2}{t} \end{pmatrix}.$$

Per poter usare la formula di Duhamel dobbiamo trovare prima una matrice fondamentale per l'equazione omogenea associata $x' = A(t)x$, cioè due soluzioni linearmente indipendenti di

$$\begin{cases} u' = v, \\ v' = -\frac{2}{t^2}u + \frac{2}{t}v. \end{cases} \quad (5.14)$$

Notiamo che

$$-\frac{2}{t^2}u + \frac{2}{t}v = 2\left(\frac{d}{dt}\left(\frac{1}{t}\right)u + \frac{1}{t}v\right) = 2\left(\frac{d}{dt}\left(\frac{1}{t}\right)u + \frac{1}{t}u'\right) = 2\frac{d}{dt}\left(\frac{u}{t}\right)$$

e il sistema (5.14) può essere scritto come

$$\begin{cases} u' = v, \\ v' = 2\frac{d}{dt}\left(\frac{u}{t}\right). \end{cases}$$

Integrando in t la seconda equazione e sostituendo il risultato nella prima, otteniamo

$$\begin{cases} u' = 2\frac{u}{t} + c \\ v = 2\frac{u}{t} + c \end{cases} \quad c \in \mathbb{R}.$$

La prima equazione differenziale è lineare del primo ordine e possiamo usare la formula di rappresentazione (5.13) determinarne l'integrale generale:

$$A(t) = \int \frac{2}{t} dt = \log t^2$$

$$u(t) = t^2 \left(u_0 + \int_{t_0}^t \frac{c}{s^2} ds \right) = t^2 \left(u_0 - \frac{c}{t} + \frac{c}{t_0} \right),$$

dove $u_0 = u(t_0)$. Scegliamo due valori particolari di c e u_0 in modo che le corrispondenti soluzioni siano linearmente indipendenti. Ad esempio, scegliendo $c = 0$ e $u_0 = 1$ troviamo $u_1(t) = t^2$ e quindi $v_1(t) = 2t$, mentre per $c = -1$ e $u_0 = -1/t_0$ si ottengono $u_2(t) = t$ e $v_2(t) = 1$. La corrispondente matrice

$$X(t) = \begin{pmatrix} u_1(t) & u_2(t) \\ v_1(t) & v_2(t) \end{pmatrix} = \begin{pmatrix} t^2 & t \\ 2t & 1 \end{pmatrix}$$

è una matrice fondamentale per il sistema, poiché il wronskiano

$$W(t) = \det X(t) = -t^2$$

è non nullo per $t > 0$. Per poter usare la formula di Duhamel dobbiamo calcolare l'inversa di X che è

$$X(t)^{-1} = -\frac{1}{t^2} \begin{pmatrix} 1 & -t \\ -2t & t^2 \end{pmatrix}.$$

Imponiamo che valga la condizione iniziale corrispondente a $t_0 = 2$ e $x_0 = (1, 4)$:

$$X(t_0)^{-1}x_0 = \begin{pmatrix} -1/4 & 1/2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 7/4 \\ -3 \end{pmatrix}$$

e determiniamo il primo termine della formula di Duhamel

$$X(t) \begin{pmatrix} 7/4 \\ -3 \end{pmatrix} = \begin{pmatrix} t^2 & t \\ 2t & 1 \end{pmatrix} \begin{pmatrix} \frac{7}{4} \\ -3 \end{pmatrix} = \begin{pmatrix} 7/4t^2 - 3t \\ 7/2t - 3 \end{pmatrix}.$$

Poi procediamo al calcolo dell'integrale: si ha

$$X(s)^{-1}b(s) = \begin{pmatrix} -1/s^2 & 1/s \\ 2/s & -1 \end{pmatrix} \begin{pmatrix} s^4 \\ s^3 \end{pmatrix} = \begin{pmatrix} 0 \\ s^3 \end{pmatrix}$$

da cui

$$X(t) \int_2^t X(s)^{-1}b(s) ds = \begin{pmatrix} t^2 & t \\ 2t & 1 \end{pmatrix} \begin{pmatrix} 0 \\ t^4/4 - 4 \end{pmatrix} = \begin{pmatrix} t^5/4 - 4t \\ t^4/4 - 4 \end{pmatrix}.$$

In conclusione otteniamo che

$$w(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} t^5/4 + 7/4t^2 - 7t \\ t^4/4 + 7/2t - 7 \end{pmatrix}$$

è la soluzione cercata.

Capitolo 6

Matrice fondamentale esponenziale

6.1 Sistemi lineari a coefficienti costanti

Abbiamo osservato nel precedente capitolo che, una volta nota la formula di Duhamel, il problema della risoluzione del sistema $x' = A(t)x + b(t)$ si riduce al problema della determinazione di una matrice fondamentale per il sistema omogeneo associato $x' = A(t)x$.

Mentre non è, in generale, possibile proporre un metodo di costruzione di tale matrice fondamentale, nel caso di sistemi a coefficienti costanti

$$x' = Ax, \quad A \in M_n(\mathbb{R}) \tag{6.1}$$

il problema si risolve completamente utilizzando strumenti di analisi funzionale (operatori lineari e limitati in spazi di Banach) e di algebra lineare (forma canonica di Jordan).

Per capire cosa dobbiamo aspettarci, ragioniamo per un momento in maniera euristica: se U è la matrice fondamentale speciale di (6.1) in $t = 0$, allora risolve il problema di Cauchy matriciale

$$\begin{cases} X' = AX, \\ X(0) = I. \end{cases}$$

Immaginiamo di costruire la soluzione matriciale come limite di iterate di Picard, analogamente a quanto abbiamo fatto nel caso vettoriale (Teorema 2.3.1). Grazie al fatto che l'equazione è lineare a coefficienti costanti, in questo caso

otteniamo che la successione delle iterate ha la forma

$$\begin{aligned} X_0(t) &= I, \\ X_1(t) &= I + \int_0^t AX_0(s) ds = I + At, \\ X_2(t) &= I + \int_0^t AX_1(s) ds = I + A \int_0^t I + As ds = I + At + A^2 \frac{t^2}{2}, \\ &\vdots \\ X_k(t) &= I + \int_0^t AX_{k-1}(s) ds = \sum_{j=0}^k \frac{A^j t^j}{j!}, \end{aligned}$$

ossia coincide con la ridotta k -esima di una serie (a termini matrici) di tipo esponenziale. Immaginando di poter estendere il concetto di limite allo spazio $M_n(\mathbb{R})$ delle matrici a coefficienti reali, è ragionevole affermare che la matrice fondamentale speciale possa essere della forma

$$U(t) = \sum_{j=0}^{+\infty} \frac{A^j t^j}{j!} \quad (6.2)$$

e, per analogia con le serie numeriche, ci verrebbe da concludere che $U(t) = e^{At}$, ossia che la matrice fondamentale speciale è un esponenziale matriciale.

Osserviamo che se $n = 1$, il problema diventa scalare

$$\begin{cases} x' = ax \\ x(0) = 1 \end{cases}$$

ed effettivamente la (6.2) converge alla soluzione $x(t) = e^{at}$. Quindi (6.2) sembra essere la naturale generalizzazione n -dimensionale del caso di dimensione uno.

Effettivamente questa è la strada giusta. Dobbiamo però dare un senso rigoroso alla (6.2). Poiché le matrici $M_n(\mathbb{R})$ possono essere identificate con gli operatori lineari e limitati di \mathbb{R}^n in sé, il contesto giusto per sviluppare la teoria sarà proprio quello degli operatori lineari e limitati in spazi di Banach.

6.2 Notazioni e richiami di Algebra Lineare

Dedicheremo questo paragrafo agli strumenti di algebra lineare che ci serviranno per lo studio dei sistemi lineari di equazioni differenziali. La trattazione

sarà estremamente sintetica. Per le dimostrazioni dei risultati, altri dettagli ed eventuali approfondimenti segnaliamo, ad esempio, [7] e [4].

La nostra attenzione sarà concentrata soprattutto sulle proprietà dello spazio delle matrici $M_n(\mathbb{R})$. Risulterà inoltre molto utile lavorare con vettori e matrici a coefficienti complessi. Quindi in questo paragrafo preferiamo indicare con \mathbb{K} il campo degli scalari, intendendo sempre $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$. Il modulo di $x = a + ib \in \mathbb{C}$ verrà indicato con $|x| = \sqrt{a^2 + b^2}$, mentre il coniugato di x sarà denotato con $\bar{x} = a - ib$. Penseremo \mathbb{K}^n sempre come spazio vettoriale su \mathbb{K} dotato della norma indotta dal prodotto scalare

$$\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i, \text{ se } \mathbb{K} = \mathbb{R},$$

$$\langle x, y \rangle = \bar{x}^T y = \sum_{i=1}^n \bar{x}_i y_i, \text{ se } \mathbb{K} = \mathbb{C}.$$

La base standard di \mathbb{K}^n sarà indicata con $\{e_1, \dots, e_n\}$.

Indicheremo con $M_n(\mathbb{K})$ lo spazio delle matrici $n \times n$ a coefficienti in \mathbb{K} . Ricordiamo che $M_n(\mathbb{K})$ è uno spazio vettoriale ed un’algebra non commutativa su \mathbb{K} con le usuali operazioni di somma, prodotto per uno scalare e prodotto righe per colonne.

Indicheremo con I la matrice identità in $M_n(\mathbb{K})$ ed useremo sistematicamente che una matrice A commuta sia con se stessa che con I .

Diremo che una matrice $A \in M_n(\mathbb{K})$ è invertibile se esiste $A^{-1} \in M_n(\mathbb{K})$ tale che $AA^{-1} = A^{-1}A = I$. A^{-1} prende il nome di matrice inversa di A .

Proposizione 6.2.1. *Le seguenti affermazioni sono equivalenti:*

- i) $A \in M_n(\mathbb{K})$ è invertibile;
- ii) $\det(A) \neq 0$;
- iii) le colonne (o le righe) di A sono vettori linearmente indipendenti.

Se $A \in M_n(\mathbb{K})$, la matrice $A^* \in M_n(\mathbb{K})$ che verifica la proprietà

$$\langle u, Av \rangle = \langle A^*u, v \rangle \quad \text{per ogni } u, v \in \mathbb{K}^n$$

prende il nome di matrice aggiunta di A . In $M_n(\mathbb{R})$ l’aggiunta corrisponde alla matrice trasposta A^T , mentre in $M_n(\mathbb{C})$ corrisponde alla coniugata della trasposta \bar{A}^T .

Una matrice A si dice hermitiana (in $M_n(\mathbb{C})$) o simmetrica (in $M_n(\mathbb{R})$) se $A = A^*$.

Una matrice A si dice unitaria (in $M_n(\mathbb{C})$) o ortogonale (in $M_n(\mathbb{R})$) se è invertibile e $A^{-1} = A^*$.

La proprietà fondamentale dello spazio $M_n(\mathbb{K})$ che ci interessa richiamare è il suo stretto legame con lo spazio delle applicazioni lineari di \mathbb{K}^n .

Definizione 6.2.2. L'applicazione $f: \mathbb{K}^n \rightarrow \mathbb{K}^n$ si dice (\mathbb{K} -)lineare se

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y),$$

per ogni $x, y \in \mathbb{K}^n$, $\lambda, \mu \in \mathbb{K}$.

L'insieme delle applicazioni lineari di \mathbb{K}^n in sé sarà indicato con $L(\mathbb{K}^n)$.

$L(\mathbb{K}^n)$ è uno spazio vettoriale su \mathbb{K} con le usuali operazioni puntuali tra applicazioni:

1. Somma:

$$(f + g)(x) := f(x) + g(x), \text{ per ogni } x \in \mathbb{K}^n;$$

2. Prodotto per uno scalare:

$$(\lambda f)(x) := \lambda f(x), \text{ per ogni } x \in \mathbb{K}^n.$$

Inoltre è un'algebra su \mathbb{K} rispetto all'operazione prodotto data dalla composizione $(fg)(x) = f(g(x))$, $x \in \mathbb{K}^n$.

Data una matrice $A \in M_n(\mathbb{K})$, l'applicazione $x \mapsto Ax$ che associa ad $x \in \mathbb{K}^n$ il prodotto righe per colonne di A con x^T è ovviamente lineare. Il seguente risultato mostra che in realtà queste sono tutte le possibili applicazioni lineari di \mathbb{K}^n in sé.

Proposizione 6.2.3. *L'applicazione*

$$\begin{aligned} \Phi: L(\mathbb{K}^n) &\longrightarrow M_n(\mathbb{K}) \\ A &\longmapsto (Ae_1 | \dots | Ae_n) \end{aligned}$$

è un isomorfismo tra $L(\mathbb{K}^n)$ e $M_n(\mathbb{K})$.

Basandoci su questo risultato, d'ora in poi identificheremo sempre una matrice con l'applicazione lineare corrispondente. Osserviamo che l'isomorfismo Φ dipende dalla base scelta in \mathbb{K}^n ed è interessante vedere come si trasforma la rappresentazione di un'applicazione lineare a seguito di un cambiamento di base. Siano $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_n\}$ due basi di \mathbb{K}^n , siano

$$B = (v_1 | \dots | v_n) \quad B' = (w_1 | \dots | w_n)$$

le matrici che hanno come colonne i vettori di ciascuna base e sia C la matrice delle coordinate dei vettori w_1, \dots, w_n nella base $\{v_1, \dots, v_n\}$, in modo tale che sia $B' = BC$. Per il Teorema di Binet $\det(C) = \det(B')/\det(B) \neq 0$, quindi C è invertibile. Se indichiamo con x e x' le coordinate di un vettore rispettivamente nella base $\{v_1, \dots, v_n\}$ e nella base $\{w_1, \dots, w_n\}$, si ha $Bx = B'x' = BCx'$ e quindi il cambiamento di coordinate associato al cambiamento di base è dato da $x = Cx'$.

Se A e A' sono due matrici in $M_n(\mathbb{K})$ che rappresentano lo stesso operatore di $L(\mathbb{K}^n)$ rispetto alle due basi, il legame tra le due matrici è dato dal diagramma commutativo

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{A'} & \mathbb{K}^n \\ C \downarrow & & \downarrow C \\ \mathbb{K}^n & \xrightarrow{A} & \mathbb{K}^n \end{array}$$

Diremo che due matrici $A, A' \in M_n(\mathbb{K})$ sono equivalenti se rappresentano lo stesso operatore lineare a meno di un cambiamento di base in \mathbb{K}^n , ossia se esiste una matrice C invertibile tale che $A = C^{-1}A'C$. Quindi, se due matrici sono equivalenti, la matrice C che trasforma una nell'altra ha come colonne le coordinate dei vettori nella nuova base (quella in cui la matrice di rappresentazione è A') rispetto alla vecchia base.

Per il Teorema di Binet, $\det(A) = \det(C^{-1}A'C) = \det(A')$ quindi il determinante è un invariante per cambiamenti di base.

Definizione 6.2.4 (Autovalori ed autovettori). Sia $A \in M_n(\mathbb{K})$. Diremo che $\lambda \in \mathbb{K}$ è autovalore di A se esiste $v \in \mathbb{K}^n$, $v \neq 0$, tale che $Av = \lambda v$.

Ogni $v \neq 0$ tale che $Av = \lambda v$ è un autovettore associato all'autovalore λ .

L'insieme di tutti gli autovettori associati ad un autovalore λ è un sottospazio vettoriale di \mathbb{K}^n e prende il nome di autospazio associato a λ .

L'insieme degli autovalori di A si chiama spettro di A e si indica con $\sigma(A)$.

Osservazione 6.2.5. Se $\lambda \in \sigma(A)$ e v è un suo autovettore, allora per ogni $k \in \mathbb{N}$, $k \geq 1$, si ha $\lambda^k \in \sigma(A^k)$ e $A^k v = \lambda^k v$.

Proposizione 6.2.6. Sia $\lambda \in \mathbb{K}$. Le seguenti affermazioni sono equivalenti.

- i) $\lambda \in \sigma(A)$;
- ii) $\ker(A - \lambda I) \neq 0$;
- iii) $\det(A - \lambda I) = 0$.

Questo motiva l'introduzione della funzione

$$\begin{aligned} p_A : \mathbb{K} &\longrightarrow \mathbb{K} \\ \lambda &\longmapsto \det(A - \lambda I) \end{aligned}$$

Si vede che p_A è un polinomio di grado n sul campo \mathbb{K} e per questo prende il nome di polinomio caratteristico di A . Dunque gli autovalori di A su \mathbb{K} sono tutti e soli gli zeri del suo polinomio caratteristico.

Definizione 6.2.7. Sia $\lambda \in \sigma(A)$. Chiameremo

- i) molteplicità algebrica di λ , denotata con $\text{ma}(\lambda)$, la sua molteplicità come radice di p_A ;
- ii) molteplicità geometrica di λ , denotata con $\text{mg}(\lambda)$, la dimensione del nucleo $\ker(A - \lambda I)$ come spazio vettoriale su \mathbb{K} .

In generale si ha che $\text{mg}(\lambda) \leq \text{ma}(\lambda)$ per ogni $\lambda \in \sigma(A)$. Se $\text{ma}(\lambda) = \text{mg}(\lambda)$ diremo che l'autovalore λ è semisemplice.

Poiché $\det(C^{-1}A'C - \lambda I) = \det(C^{-1}(A - \lambda I)C) = \det(A - \lambda I)$ per ogni C invertibile, il polinomio caratteristico è invariante per cambiamento di base e quindi, così come il determinante, può essere pensato come un oggetto intrinseco dell'operatore lineare associato alle matrici equivalenti.

Se $\mathbb{K} = \mathbb{C}$ il teorema fondamentale dell'algebra ci dice che p_A ha esattamente n zeri contati con la loro molteplicità, cioè la matrice A possiede n autovalori.

Osservazione 6.2.8 (Abuso di linguaggio per autovalori di matrici reali). Una matrice $A \in M_n(\mathbb{R})$ può non avere alcun autovalore, come, ad esempio, la matrice

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

il cui polinomio caratteristico $p_A(\lambda) = \lambda^2 + 1$ non ha radici reali. D'altra parte, l'esistenza di n autovalori è garantita se pensiamo la stessa matrice come elemento di $M_n(\mathbb{C})$ ed ambientiamo il problema della ricerca degli autovalori in campo complesso.

D'ora in poi chiameremo autovalori di $A \in M_n(\mathbb{R})$ gli zeri complessi del suo polinomio caratteristico (a coefficienti reali). Le radici reali del polinomio caratteristico corrisponderanno agli autovalori in \mathbb{R} di A : infatti, se $\lambda \in \mathbb{R}$ è un autovalore di A e $z \in \mathbb{C}^n$ è un autovettore associato a λ , allora anche il suo coniugato \bar{z} è autovettore relativo a λ in \mathbb{C}^n . Quindi $2\operatorname{Re} z = z + \bar{z}$ risulta essere un autovettore in \mathbb{R}^n relativo a λ .

Questo nostro abuso di linguaggio ha una spiegazione teorica, basata sulla nozione di complessificazione, che sarà chiarita nell'Osservazione 7.1.4

Osservazione 6.2.9. Dati $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, indicheremo con $\operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$ la matrice diagonale con elementi $a_{ij} = 0$ per $i \neq j$, $a_{ii} = \lambda_i$, $i, j = 1, \dots, n$. Ovviamente, se $\Lambda = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$ allora $\sigma(\Lambda) = \{\lambda_1, \dots, \lambda_n\}$ (con autovalori eventualmente ripetuti) e tutti gli autovalori sono regolari. In particolare, l'elemento e_i della base canonica è un autovettore relativo all'autovalore λ_i . Se $\{\lambda_1, \dots, \lambda_r\}$, $r \leq n$ sono gli autovalori distinti di una matrice diagonale Λ , non è restrittivo supporre (a meno di un riordinamento degli elementi della base canonica) che Λ abbia la forma $\Lambda = \operatorname{diag}\{\lambda_1 I_{m_1}, \dots, \lambda_r I_{m_r}\}$, dove $m_j = \operatorname{ma}(\lambda_j)$ e I_{m_j} è la matrice identità di ordine m_j , $j = 1, \dots, r$. Quindi l'autospazio relativo a λ_1 è generato dai primi m_1 elementi della base, quello relativo a λ_2 dai vettori $\{e_{m_1+1}, \dots, e_{m_2}\}$ e così via. In particolare, lo spazio \mathbb{K}^n si decompone nella somma diretta

$$\mathbb{K}^n = \bigoplus_{j=1}^r \operatorname{Re}(E_{\lambda_j})$$

dove E_{λ_j} è l'autospazio relativo all'autovalore λ_j , ed ogni autospazio è invariante sotto l'azione di Λ , dal momento che $\Lambda v = \lambda_j v$ per ogni $v \in E_{\lambda_j}$.

Quindi, il fatto che una matrice sia diagonale corrisponde geometricamente alla possibilità di decomporre \mathbb{K}^n nella somma diretta di sottospazi invarianti

per Λ , che quindi ne fattorizzano l'azione. Questa proprietà geometrica risulta essere particolarmente utile per i nostri scopi e, come naturale generalizzazione, siamo interessati alle matrici che risultano essere equivalenti ad una matrice diagonale.

Definizione 6.2.10. Una matrice $A \in M_n(\mathbb{K})$ è diagonalizzabile su \mathbb{K} se è equivalente ad una matrice diagonale $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$.

Chiaramente, in questo caso $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$, in cui gli autovalori sono ripetuti secondo la loro molteplicità.

Il risultato generale riguardante la diagonalizzabilità di una matrice è il seguente.

Teorema 6.2.11. Condizione necessaria e sufficiente affinché una matrice $A \in M_n(\mathbb{C})$ sia diagonalizzabile è che tutti i suoi autovalori siano semisemplici. In questo caso, la matrice C tale che $A = C^{-1}\Lambda C$ avrà come colonne gli elementi di una base di \mathbb{C}^n costituita da autovettori.

Inoltre, se $A \in M_n(\mathbb{R})$ ha tutti autovalori reali (pensata come matrice in $M_n(\mathbb{C})$), allora anche la base di autovettori può essere scelta in \mathbb{R}^n .

Infine, se $A \in M_n(\mathbb{R})$ è simmetrica, allora ha tutti autovalori reali ed è sempre diagonalizzabile. Inoltre la base di autovettori in \mathbb{R}^n può essere scelta in modo tale che la matrice di trasformazione C sia ortogonale.

Concludiamo il paragrafo introducendo un'altra classe di matrici che useremo nel seguito.

Definizione 6.2.12. Diremo che una matrice $N \in M_n(\mathbb{K})$ è nilpotente se esiste $k \in \mathbb{N}$ tale che $N^k = 0$, ossia è la matrice con tutti i coefficienti nulli.

Osserviamo che per ogni $k \in \mathbb{N}$

$$(CAC^{-1})^k = (CAC^{-1})(CAC^{-1}) \cdots (CAC^{-1}) = CA^k C^{-1}, \quad (6.3)$$

quindi, in particolare, ogni matrice equivalente ad una matrice nilpotente è nilpotente. Inoltre, se N è nilpotente, $\lambda \in \sigma(N)$ e $v \neq 0$ è un autovettore relativo a λ , si ha $0 = N^k v = \lambda^k v$ e quindi $\sigma(N) = \{0\}$.

6.3 Operatori lineari e limitati in spazi di Banach

Nel paragrafo precedente abbiamo considerato gli spazi $L(\mathbb{K}^n)$, $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$, rappresentandoli tramite matrici. Vogliamo ora introdurre una struttura di spazio normato su $L(\mathbb{K}^n)$. In linea di principio, essendo $L(\mathbb{K}^n)$ uno spazio di dimensione finita, ci potremmo limitare ad introdurre una norma euclidea e lavorare con quella. Preferiamo invece introdurre una nuova nozione di norma operatoriale e sviluppare una teoria che resti valida anche in contesti più generali.

Sia $(X, \|\cdot\|)$ uno spazio vettoriale normato sul campo \mathbb{K} .

Una successione $\{u_k\} \subseteq X$ è di Cauchy se per ogni $\varepsilon > 0$ fissato esiste un indice $\tilde{k} = \tilde{k}(\varepsilon) \in \mathbb{N}$ tale che se $h, k > \tilde{k}$ si ha $\|u_h - u_k\| < \varepsilon$.

La successione $\{u_k\}$ si dice convergente se esiste $u \in X$ tale che $u_k \rightarrow u$ in X , cioè che per ogni $\varepsilon > 0$ esiste un indice $\tilde{k} = \tilde{k}(\varepsilon) \in \mathbb{N}$ tale che se $k > \tilde{k}$ si ha $\|u_k - u\| < \varepsilon$.

Supponiamo che X sia completo rispetto alla norma, cioè che ogni successione $\{u_k\} \subseteq X$ di Cauchy sia convergente. Uno spazio vettoriale normato e completo si dice spazio di Banach.

Definizione 6.3.1. Diremo che un operatore $A: X \rightarrow X$ è \mathbb{K} -lineare se

$$A(\lambda u + \mu v) = \lambda Au + \mu Av, \text{ per ogni } u, v \in X, \lambda, \mu \in \mathbb{K}.$$

Un operatore lineare si dice limitato se esiste una costante $M > 0$ tale che

$$\|Au\| \leq M \|u\|, \text{ per ogni } u \in X.$$

L'insieme di tutti gli operatori lineari e limitati di X in sé verrà indicato con $\mathcal{L}(X)$.

Osservazione 6.3.2. Se A è un operatore lineare limitato, allora

$$\|Au - Av\| = \|A(u - v)\| \leq M \|u - v\|$$

quindi A è continuo in X .

Se $A \in \mathcal{L}(X)$, allora è ben definita la norma operatoriale

$$\|A\|_{\mathcal{L}} := \sup_{\substack{u \in X \\ u \neq 0}} \frac{\|Au\|}{\|u\|}.$$

Si dimostra facilmente che valgono le caratterizzazioni

$$\begin{aligned}\|A\|_{\mathcal{L}} &= \sup_{\substack{u \in X \\ \|u\| \leq 1}} \|Au\| = \sup_{\substack{u \in X \\ \|u\|=1}} \|Au\| \\ &= \inf \{M > 0 : \|Au\| \leq M \|u\|, \text{ per ogni } u \in X\}\end{aligned}$$

dove l'ultima uguaglianza esprime il fatto che la norma dell'operatore è la migliore costante per la stima di limitatezza. In $\mathcal{L}(X)$ possiamo introdurre le tre operazioni:

1. Somma: se $A, B \in \mathcal{L}(X)$ definiamo $A + B \in \mathcal{L}(X)$ come

$$(A + B)(u) := Au + Bu, \text{ per ogni } u \in X;$$

2. Prodotto per uno scalare: se $A \in \mathcal{L}(X)$, $\lambda \in \mathbb{K}$, definiamo $\lambda A \in \mathcal{L}(X)$ come

$$(\lambda A)(u) = \lambda Au, \text{ per ogni } u \in X;$$

3. Prodotto: se $A, B \in \mathcal{L}(X)$ definiamo il prodotto $AB \in \mathcal{L}(X)$ come la composizione

$$(AB)(u) := A(Bu), \text{ per ogni } u \in X.$$

Le prime due operazioni rendono $\mathcal{L}(X)$ uno spazio vettoriale su \mathbb{K} . Viene lasciata per esercizio la verifica che $(\mathcal{L}(X), \|\cdot\|_{\mathcal{L}})$ è uno spazio vettoriale normato.

Il prodotto di composizione è ben definito perché vale la stima

$$\|AB\|_{\mathcal{L}} \leq \|A\|_{\mathcal{L}} \|B\|_{\mathcal{L}}. \quad (6.4)$$

Infatti

$$\|A(Bu)\| \leq \|A\|_{\mathcal{L}} \|Bu\| \leq \|A\|_{\mathcal{L}} \|B\|_{\mathcal{L}} \|u\|$$

e passando al sup sugli $u \in X$ con $\|u\| = 1$ otteniamo (6.4). Il prodotto di composizione rende $\mathcal{L}(X)$ un'algebra su \mathbb{K} . Un'algebra nella quale la norma e il prodotto sono legate dalla (6.4) viene detta algebra normata.

Proposizione 6.3.3. *Con le operazioni definite $(\mathcal{L}(X), \|\cdot\|_{\mathcal{L}})$ è un'algebra normata su \mathbb{K} .*

Il seguente risultato mostra che, se X è uno spazio di Banach, allora $\mathcal{L}(X)$ è un'algebra di Banach, ossia un'algebra normata completa.

Teorema 6.3.4. *Se X è uno spazio di Banach, anche $\mathcal{L}(X)$ è uno spazio di Banach.*

Dimostrazione Teorema 6.3.4. Sia $\{A_k\}$ una successione di Cauchy in $\mathcal{L}(X)$. Per ogni $u \in X$

$$\|A_k u - A_n u\| = \|(A_k - A_n)u\| \leq \|A_k - A_n\|_{\mathcal{L}} \|u\|$$

da cui risulta che la successione $\{A_k u\}$ è di Cauchy in X . Essendo X completo, per ogni $u \in X$ la successione $\{A_k u\}$ è convergente. Questo ci permette di definire l'operatore $A: X \rightarrow X$ tale che

$$Au = \lim_{k \rightarrow +\infty} A_k u, \quad u \in X.$$

Tale limite puntuale A è il candidato operatore limite di A_k in $\mathcal{L}(X)$ (osserviamo che la convergenza degli operatori in norma operatoriale implica la loro convergenza puntuale in X , quindi il limite puntuale è l'unico candidato possibile). Per linearità dell'operazione di limite in X , A è lineare. Inoltre per continuità della norma

$$\|Au\| = \lim_{k \rightarrow +\infty} \|A_k u\| \leq c \|u\|$$

dove abbiamo sfruttato il fatto che una successione di Cauchy è sempre limitata in norma. Allora $A \in \mathcal{L}(X)$. Resta da provare che $A_k \rightarrow A$. Usiamo nuovamente il fatto che $\{A_k\}$ sia una successione di Cauchy: per ogni $\varepsilon > 0$ fissato, esiste un $m = m(\varepsilon) \in \mathbb{N}$ tale che se $k, n > m$ allora $\|A_k - A_n\|_{\mathcal{L}} < \varepsilon$. Ma allora se $k, n > m$,

$$\|A_k u - A_n u\| \leq \|A_k - A_n\|_{\mathcal{L}} \|u\| < \varepsilon \|u\|, \quad u \in X$$

e passando al limite per $n \rightarrow +\infty$ otteniamo

$$\|A_k u - Au\| < \varepsilon \|u\|, \quad u \in X.$$

A questo punto possiamo passare all'estremo superiore sup sugli u tali che $\|u\| = 1$, ottenendo $\|A_k - A\|_{\mathcal{L}} < \varepsilon$. \square

Esempio 6.3.5 ($\mathcal{L}(\mathbb{K}^n) = L(\mathbb{K}^n)$). Il caso che ci interessa è quello di $\mathcal{L}(\mathbb{K}^n)$. Cominciamo ad osservare che ogni operatore $A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ lineare risulta essere limitato. Se infatti fissiamo la base standard di \mathbb{K}^n , per conoscere come agisce A , è sufficiente sapere quanto vale sui vettori della base, per cui

$$\begin{aligned}\|Au\| &= \left\| A \left(\sum_{i=1}^n u_i e_i \right) \right\| = \left\| \sum_{i=1}^n u_i A e_i \right\| \\ &\leq \sum_{i=1}^n |u_i| \|A e_i\| \leq \left(\sum_{i=1}^n \|A e_i\|^2 \right)^{\frac{1}{2}} \|u\|\end{aligned}$$

dove nell'ultimo passaggio è stata usata la ben nota diseguaglianza di Cauchy–Schwartz. Questo ci dice che $\|A\|_{\mathcal{L}}$ è sempre limitata e che vale la stima

$$\|A\|_{\mathcal{L}} \leq \left(\sum_{i=1}^n \|A e_i\|^2 \right)^{\frac{1}{2}}. \quad (6.5)$$

Dunque $\mathcal{L}(\mathbb{K}^n)$ sono gli operatori lineari di \mathbb{K}^n in sé.

Calcoliamo esplicitamente la norma operatoriale di alcune classi di matrici.

Esempio 6.3.6. Osserviamo preliminarmente che, se $A \in M_n(\mathbb{K})$ e $\lambda \in \sigma(A)$, allora

$$|\lambda| \leq \|A\|_{\mathcal{L}}. \quad (6.6)$$

Infatti se $v \in \mathbb{K}^n$ è un autovettore associato a λ , si ha $Av = \lambda v$ e quindi

$$|\lambda| \|v\| = \|Av\| \leq \|A\|_{\mathcal{L}} \|v\|.$$

Poiché gli autovettori sono per definizione non nulli, possiamo dividere per $\|v\|$ e ottenere la (6.6).

Matrici diagonali. Sia $\Lambda \in M_n(\mathbb{K})$ della forma $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$, allora

$$\|A\|_{\mathcal{L}} = \max_{i=1, \dots, n} |\lambda_i|.$$

Infatti per $u \in \mathbb{K}^n$ si ha

$$\|Au\|^2 = \sum_{i=1}^n |\lambda_i|^2 |u_i|^2 \leq \left(\max_{i=1, \dots, n} |\lambda_i| \right)^2 \|u\|^2.$$

Prendendo la radice quadrata e passando all'estremo superiore sugli u con $\|u\| = 1$, otteniamo

$$\|A\|_{\mathcal{L}} \leq \max_{i=1,\dots,n} |\lambda_i|.$$

La diseguaglianza opposta si ricava dalla (6.6).

Matrici ortogonali. Sia $A \in M_n(\mathbb{K})$. Se A è ortogonale (o unitaria), cioè tale che $A^* = A^{-1}$, allora

$$\|A\|_{\mathcal{L}} = \|A^{-1}\|_{\mathcal{L}} = 1.$$

Infatti

$$\|Au\|^2 = \langle Au, Au \rangle = \langle A^*Au, u \rangle = \langle u, u \rangle = \|u\|^2$$

$$\|A^{-1}u\|^2 = \|A^*u\|^2 = \langle A^*u, A^*u \rangle = \langle u, AA^*u \rangle = \langle u, u \rangle = \|u\|^2$$

per cui passando al sup sugli u di norma unitaria otteniamo la tesi.

Matrici simmetriche. Se $A \in M_n(\mathbb{K})$ è una matrice simmetrica allora

$$\|A\|_{\mathcal{L}} = \max_{i=1,\dots,n} |\lambda_i|.$$

Infatti, per il Teorema 6.2.11, A è diagonalizzabile tramite una matrice $C \in M_n(\mathbb{K})$ ortogonale (o unitaria), ossia $A = C^{-1}\Lambda C$ dove $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$, $\lambda_i \in \sigma(A)$ eventualmente ripetuti. Usando la proprietà (6.4) della norma otteniamo

$$\|A\|_{\mathcal{L}} = \|C^{-1}\Lambda C\|_{\mathcal{L}} \leq \|C^{-1}\|_{\mathcal{L}} \|\Lambda\|_{\mathcal{L}} \|C\|_{\mathcal{L}} = \max_{i=1,\dots,n} |\lambda_i|.$$

La diseguaglianza opposta segue dalla (6.6).

Il seguente risultato fornisce un metodo operativo per calcolare la norma operatoriale di una matrice A in termini dello spettro della matrice A^*A .

Proposizione 6.3.7. *Sia $A \in M_n(\mathbb{K})$, allora*

$$\|A\| = \sqrt{\rho(A^*A)}$$

dove $\rho(A^*A) = \max_{\lambda \in \sigma(A^*A)} |\lambda|$.

Dimostrazione. Primo passo: dimostriamo che

$$\|A\|_{\mathcal{L}} = \|A^*\|_{\mathcal{L}}.$$

Infatti, sia $u \in \mathbb{K}^n$. Allora usando la disegualanza di Cauchy-Schwartz

$$\|Au\|^2 = \langle Au, Au \rangle = \langle A^*Au, u \rangle \leq \|A^*Au\| \|u\| \leq \|A^*\|_{\mathcal{L}} \|Au\| \|u\|$$

e quindi dividendo per $\|Au\|$ (non è restrittivo supporre che sia non nullo) e passando all'estremo superiore sugli u di norma 1 otteniamo $\|A\| \leq \|A^*\|$. Poiché $A^{**} = A$, applicando la stima appena fatta ad A^* otteniamo anche la disegualanza opposta.

Secondo passo (identità C^*): dimostriamo che

$$\|A^*A\|_{\mathcal{L}} = \|A\|_{\mathcal{L}}^2. \quad (6.7)$$

Infatti

$$\|A^*A\|_{\mathcal{L}} \leq \|A^*\|_{\mathcal{L}} \|A\|_{\mathcal{L}} = \|A\|_{\mathcal{L}}^2$$

per il passo precedente. Viceversa, sia $u \in \mathbb{K}^n$, allora

$$\|Au\|^2 = \langle Au, Au \rangle = \langle A^*Au, u \rangle \leq \|A^*Au\| \|u\| \leq \|A^*A\|_{\mathcal{L}} \|u\|^2$$

e passando all'estremo superiore sugli u tali che $\|u\| = 1$ otteniamo $\|A\|_{\mathcal{L}}^2 \leq \|A^*A\|_{\mathcal{L}}$.

Conclusione: la matrice A^*A è simmetrica, quindi per il teorema spettrale, ha tutti autovalori reali. Inoltre sono tutti positivi poiché A^*A è definita positiva. Quindi è ben definita la radice quadrata di $\rho(A^*A)$. Essendo A^*A simmetrica, per l'Esempio 6.3.6 si ha $\|A^*A\|_{\mathcal{L}} = \rho(A^*A)$. Per l'identità (6.7) $\|A^*A\|_{\mathcal{L}} = \|A\|_{\mathcal{L}}^2$ da cui la tesi. \square

Concludiamo questa digressione sugli operatori lineari in \mathbb{K}^n con la dimostrazione del fatto che le matrici diagonalizzabili sono dense nelle matrici triangolari superiori.

Proposizione 6.3.8. *Sia $A \in M_n(\mathbb{K})$ una matrice triangolare superiore. Per ogni $\varepsilon > 0$ esiste una matrice $D \in M_n(\mathbb{K})$ diagonalizzabile tale che $\|A - D\|_{\mathcal{L}} < \varepsilon$.*

Dimostrazione. La matrice A è della forma

$$A = \begin{pmatrix} \lambda_1 & a_{1,2} & \dots & a_{1,n} \\ 0 & \lambda_2 & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Fissiamo un $\varepsilon > 0$ e scegliamo degli $\varepsilon_i \in (0, \varepsilon)$ tali che la matrice

$$D = \begin{pmatrix} \lambda_1 - \varepsilon_1 & a_{1,2} & \dots & a_{1,n} \\ 0 & \lambda_2 - \varepsilon_2 & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n - \varepsilon_n \end{pmatrix}$$

abbia tutte entrate distinte sulla diagonale. In tal modo D è diagonalizzabile perché ha tutti autovalori distinti. Dobbiamo calcolare la norma di $A - D$. Ma

$$A - D = \text{diag}\{\varepsilon_1, \dots, \varepsilon_n\}.$$

e allora grazie all'Esempio 6.3.6 otteniamo che

$$\|A - D\|_{\mathcal{L}} = \max\{\varepsilon_1, \dots, \varepsilon_n\} < \varepsilon.$$

□

6.4 Operatore esponenziale

Nello spazio di Banach $\mathcal{L}(X)$ possiamo introdurre la nozione di serie nella maniera usuale: data una successione di operatori $\{A_k\} \in \mathcal{L}(X)$, definiamo le somme parziali

$$S_k = \sum_{j=0}^k A_j.$$

Diciamo che la serie di A_k converge ad $A \in \mathcal{L}(X)$ se la successione delle somme parziali S_k converge ad A nella norma operatoriale. In questo caso scriveremo

$$\sum_{j=0}^{+\infty} A_j = A$$

e chiameremo A somma della serie.

Diremo che la serie di $\{A_k\}$ converge assolutamente se converge la serie reale

$$\sum_{j=0}^{+\infty} \|A_j\|_{\mathcal{L}}.$$

Come in tutti gli spazi di Banach, l'assoluta convergenza di una serie implica la convergenza in norma.

Siamo ora in grado di introdurre rigorosamente l'operatore esponenziale che compariva nell'euristica per la determinazione della matrice fondamentale speciale per i sistemi lineari omogenei a coefficienti costanti (si veda (6.2)). Infatti, se $A \in \mathcal{L}(X)$, applicando ripetutamente la (6.4) si ottiene

$$\|A^k\|_{\mathcal{L}} \leq \|A\|_{\mathcal{L}}^k, \text{ per ogni } k \in \mathbb{N},$$

per cui la serie

$$\sum_{k=0}^{+\infty} \frac{A^k}{k!}$$

converge assolutamente. Infatti, poiché

$$\frac{A^k}{k!} \leq \frac{\|A\|_{\mathcal{L}}^k}{k!}$$

otteniamo

$$\sum_{k=0}^{+\infty} \left\| \frac{A^k}{k!} \right\|_{\mathcal{L}} \leq e^{\|A\|_{\mathcal{L}}}. \quad (6.8)$$

Possiamo quindi dare la seguente definizione.

Definizione 6.4.1 (Operatore esponenziale). Sia $A \in \mathcal{L}(X)$. L'esponenziale di A è l'operatore $e^A \in \mathcal{L}(X)$ definito come la somma della serie

$$e^A = \sum_{k=0}^{+\infty} \frac{A^k}{k!}.$$

La (6.8) ci fornisce anche la maggiorazione

$$\|e^A\|_{\mathcal{L}} \leq e^{\|A\|}.$$

Come è naturale aspettarsi, vista la definizione, l'operatore esponenziale gode delle usuali proprietà.

Proposizione 6.4.2. *L'operatore esponenziale gode delle seguenti proprietà.*

- i) $e^0 = Id$, ovvero $e^0 u = u$ per ogni $u \in X$;
- ii) se $A, B \in \mathcal{L}(X)$ commutano, cioè $AB = BA$, allora $e^{A+B} = e^A e^B$;
- iii) se $A \in \mathcal{L}(X)$, e^A è invertibile con inverso $(e^A)^{-1} = e^{-A}$;
- iv) se $A, C, C^{-1} \in \mathcal{L}(X)$ allora $e^{CAC^{-1}} = C e^A C^{-1}$.

Dimostrazione. L'affermazione del punto i) è evidente, visto che il termine della serie corrispondente a $k = 0$ della serie è esattamente Id .

La dimostrazione del punto ii) è basata sul fatto che, se A e B commutano, allora vale la formula del binomio

$$(A + B)^k = \sum_{j=0}^k \binom{k}{j} A^j B^{k-j}$$

e quindi si ha

$$\sum_{k=0}^n \frac{(A + B)^k}{k!} = \sum_{k=0}^n \frac{1}{k!} \sum_{j=0}^k \binom{k}{j} A^j B^{k-j} = \sum_{k=0}^n \sum_{j=0}^k \frac{A^j}{j!} \frac{B^{k-j}}{(k-j)!}$$

che mostra che le somme parziali relative ad e^{A+B} coincidono con le somme parziali dei prodotti di Cauchy tra le successioni $\{\frac{A^k}{k!}\}$ e $\{\frac{B^k}{k!}\}$. Si può facilmente verificare che anche in questo contesto vale il risultato di convergenza della serie dei prodotti di Cauchy al prodotto delle serie, a patto che almeno una delle serie converga assolutamente (e la serie dei prodotti di Cauchy converge assolutamente se entrambe le serie convergono assolutamente). Quindi, poiché le serie $\{\frac{A^k}{k!}\}$ e $\{\frac{B^k}{k!}\}$ convergono assolutamente, ne concludiamo che

$$\begin{aligned} e^{A+B} &= \sum_{k=0}^{+\infty} \frac{(A + B)^k}{k!} = \sum_{k=0}^{+\infty} \sum_{j=1}^k \frac{A^j}{j!} \frac{B^{k-j}}{(k-j)!} \\ &= \left(\sum_{k=0}^{+\infty} \frac{A^k}{k!} \right) \left(\sum_{j=0}^{+\infty} \frac{B^j}{j!} \right) = e^A e^B. \end{aligned}$$

L'affermazione iii) segue immediatamente da i) e ii).

Infine, per dimostrare iv) basta osservare che

$$(CAC^{-1})^j = CA^j C^{-1}$$

per ogni $j \in \mathbb{N}$ per cui

$$\sum_{j=0}^k \frac{(CAC^{-1})^j}{j!} = C \left(\sum_{j=0}^k \frac{A^j}{j!} \right) C^{-1}$$

da cui segue, per passaggio al limite, la tesi. \square

A questo punto è possibile definire la funzione reale e^{At} a valori nello spazio degli operatori lineari e limitati

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathcal{L}(X) \\ t &\longmapsto e^{At} \end{aligned}$$

Diremo che una funzione $u : \mathbb{R} \rightarrow \mathcal{L}(X)$ è continua in $t_0 \in \mathbb{R}$ se

$$\lim_{t \rightarrow t_0} \|u(t) - u(t_0)\|_{\mathcal{L}} = 0$$

e diremo che è derivabile in $t_0 \in \mathbb{R}$ se esiste $u'(t_0) \in \mathcal{L}(X)$ tale che

$$\lim_{t \rightarrow t_0} \left\| \frac{u(t) - u(t_0)}{t - t_0} - u'(t_0) \right\|_{\mathcal{L}} = 0.$$

Si dimostra facilmente che la derivabilità in t_0 implica la continuità in tale punto.

Proposizione 6.4.3. *Sia $A \in \mathcal{L}(X)$. Allora la funzione $t \in \mathbb{R} \mapsto e^{At} \in \mathcal{L}(X)$ è derivabile per ogni $t \in \mathbb{R}$ e*

$$\frac{d}{dt} e^{At} = A e^{At}.$$

Dimostrazione. Come primo passo, dimostriamo la derivabilità in $t_0 = 0$. Dobbiamo stimare la norma della differenza

$$\frac{e^{At} - I}{t} - A = \frac{1}{t} \left(\sum_{k=0}^{+\infty} \frac{(At)^k}{k!} - I - At \right) = A^2 t \sum_{k=0}^{+\infty} \frac{A^k t^k}{(k+2)!}.$$

Poiché ovviamente si ha $(k+2)! > k!$, ne segue che

$$\left\| \frac{e^{At} - I}{t} - A \right\|_{\mathcal{L}} \leq \|A\|_{\mathcal{L}}^2 |t| \sum_{k=0}^{+\infty} \frac{\|A\|_{\mathcal{L}}^k t^k}{k!} = \|A\|_{\mathcal{L}}^2 |t| e^{\|A\|_{\mathcal{L}} |t|}$$

e l'ultimo termine tende a zero per t che tende a 0. Quindi e^{At} è derivabile in 0 e

$$\left. \frac{d}{dt} (e^{At}) \right|_{t=0} = A. \quad (6.9)$$

Siano ora $t, h \in \mathbb{R}$. Poiché At e Ah commutano, la Proposizione 6.4.2 ii) garantisce che

$$e^{A(t+h)} = e^{At} e^{Ah}$$

per cui

$$\begin{aligned} \left\| \frac{e^{A(t+h)} - e^{At}}{h} - Ae^{At} \right\|_{\mathcal{L}} &= \left\| \left(\frac{e^{Ah} - I}{h} - A \right) e^{At} \right\|_{\mathcal{L}} \\ &\leq \left\| \frac{e^{Ah} - I}{h} - A \right\|_{\mathcal{L}} \|e^{At}\|_{\mathcal{L}} \end{aligned}$$

e, per (6.9), l'ultimo termine tende a zero per t che tende a 0. \square

Come corollario abbiamo che l'operatore e^{At} determina tutte le soluzioni relative a problemi di Cauchy per un'equazione differenziale lineare per funzioni a valori nello spazio di Banach X .

Corollario 6.4.4. *Sia $A \in \mathcal{L}(X)$. Allora per ogni $u_0 \in X$ la funzione $u: \mathbb{R} \rightarrow \mathcal{L}(X)$ definita da $u(t) = e^{At}u_0$ è soluzione del problema di Cauchy*

$$\begin{cases} \frac{du}{dt} = Au, \\ u(0) = u_0, \end{cases}$$

nel senso che $u: \mathbb{R} \rightarrow \mathcal{L}(X)$ è derivabile per ogni $t \in \mathbb{R}$, verifica puntualmente l'equazione differenziale e soddisfa il dato iniziale.

Dimostrazione. Per la Proposizione 6.4.2 i) si ha $u(0) = e^{A0}u_0 = u_0$ e quindi la condizione iniziale è soddisfatta. Inoltre, il rapporto incrementale di u in $t \in \mathbb{R}$ è

$$\frac{u(t+h) - u(t)}{h} - Au(t) = \left(\frac{e^{A(t+h)} - e^{At}}{h} - Ae^{At} \right) u_0$$

per cui

$$\left\| \frac{u(t+h) - u(t)}{h} - Au(t) \right\|_{\mathcal{L}} \leq \left\| \frac{e^{A(t+h)} - e^{At}}{h} - Ae^{At} \right\|_{\mathcal{L}} \|u_0\|$$

e, poiché l'ultimo termine tende a zero per h che tende a zero, la funzione u soddisfa $u'(t) = Au(t)$, $t \in \mathbb{R}$. \square

6.5 Matrice fondamentale esponenziale

Come conseguenza del Corollario 6.4.4 abbiamo il seguente teorema.

Teorema 6.5.1. *Sia $A \in M_n(\mathbb{R})$, allora $e^{A(t-t_0)}$ è la matrice fondamentale speciale per il sistema di equazioni differenziali $x' = Ax$ al tempo $t = t_0$ e quindi per ogni $\xi \in \mathbb{R}^n$ la soluzione di*

$$\begin{cases} x' = Ax \\ x(t_0) = \xi \end{cases}$$

è data da $x(t) = e^{A(t-t_0)}\xi$.

Resta aperto il problema di calcolare esplicitamente l'esponenziale della matrice At , per $t \in \mathbb{R}$ fissato. Iniziamo a vedere alcuni casi in cui si calcola agevolmente.

Esempio 6.5.2 (Esponenziale di matrici diagonali). Se $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ con $\lambda_i \in \mathbb{R}$, allora $\Lambda^k = \text{diag}\{\lambda_1^k, \dots, \lambda_n^k\}$ e

$$e^{At} = \text{diag}\{e^{\lambda_1 t}, \dots, e^{\lambda_n t}\}.$$

Questa situazione banale è particolarmente elementare, visto che il sistema $x' = Ax$ è completamente disaccoppiato

$$\begin{cases} x'_1 = \lambda_1 x_1 \\ \vdots \\ x'_n = \lambda_n x_n \end{cases}$$

e le sue soluzioni sono esattamente $x(t) = e^{At}\xi = (\xi_1 e^{\lambda_1 t}, \dots, \xi_n e^{\lambda_n t})$, al variare di $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$.

Esempio 6.5.3 (Esponenziale di matrici diagonali a blocchi). Una naturale generalizzazione del caso precedente è quello in cui la matrice A si presenta diagonale a blocchi, ossia $A = \text{diag}\{B_1, \dots, B_s\}$ con $B_i \in M_{n_i}(\mathbb{R})$, $\sum_{i=1}^s n_i = n$. Osserviamo che questa forma corrisponde al fatto che lo spazio \mathbb{R}^n risulta fattorizzato nella somma diretta di s sottospazi E_i , $i = 1, \dots, s$, ognuno dei quali è invariante per A (ossia $Au \in E_i$ per tutti e soli gli elementi $u \in E_i$).

Si verifica facilmente che $A^k = \text{diag}\{B_1^k, \dots, B_s^k\}$ e, quindi

$$e^{At} = \text{diag}\{e^{B_1 t}, \dots, e^{B_s t}\},$$

ossia anche la matrice esponenziale ha le stesse proprietà di fattorizzazione sui sottospazi E_i .

Esempio 6.5.4 (Esponenziale di matrici diagonalizzabili). Sia $A \in M_n(\mathbb{R})$ diagonalizzabile su \mathbb{R} . Allora esiste una matrice $C \in M_n(\mathbb{R})$ invertibile tale che $CAC^{-1} = \Lambda$, con $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$, $\lambda_i \in \sigma(A)$ eventualmente con ripetizione. Per la Proposizione 6.4.2 iv) abbiamo

$$\begin{aligned} e^{At} &= e^{C^{-1}\Lambda C t} = C^{-1}e^{\Lambda t}C \\ &= C^{-1} \text{diag}\{e^{\lambda_1 t}, \dots, e^{\lambda_n t}\}C. \end{aligned}$$

Purtroppo non sempre la matrice A è diagonalizzabile. Può accadere che

- gli autovalori di A non siano tutti reali;
- che qualche autovalore di A non sia semisemplice, per cui non si riesce a costruire una base di \mathbb{R}^n fatta di soli autovettori.

La determinazione esplicita della matrice esponenziale in questi casi sarà discussa nel prossimo capitolo.

Esempio 6.5.5. Consideriamo il sistema lineare a coefficienti costanti bidimensionale

$$\begin{cases} x' = x \\ y' = 2x + y \\ x(0) = \xi_1 \\ y(0) = \xi_2 \end{cases} \quad (6.10)$$

nelle incognite (x, y) e con dato iniziale $\xi = (\xi_1, \xi_2)$. La matrice associata al sistema è

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

Il Teorema 6.5.1 ci dice che la matrice fondamentale speciale per questo sistema è $U(t) = e^{At}$. La matrice A non è diagonalizzabile, infatti lo spettro è $\sigma(A) = \{1\}$ e l'autovalore 1 ha molteplicità algebrica 2 e molteplicità geometrica 1. D'altra parte il sistema (6.10) è un sistema triangolare come quelli considerati nell'Esempio 5.1.14. Ne conosciamo quindi la matrice fondamentale speciale:

$$U(t) = \begin{pmatrix} e^t & 0 \\ v(t) & e^t \end{pmatrix}$$

dove

$$v(t) = \int_0^t a(s)e^{\alpha s + \beta(t-s)} ds = \int_0^t 2e^t ds = 2te^t.$$

Poiché sappiamo che l'esponenziale di A è matrice fondamentale speciale, per unicità di quest'ultima si deve avere

$$e^{At} = U(t) = \begin{pmatrix} e^t & 0 \\ 2te^t & e^t \end{pmatrix}.$$

Abbiamo quindi calcolato l'esponenziale di A a posteriori conoscendo l'integrale generale del sistema. Naturalmente a noi interessa il procedimento inverso. A tal proposito osserviamo che la matrice

$$C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

è invertibile con $C^{-1} = C$ ed è tale che, se moltiplicata a sinistra scambia le righe, mentre a destra le colonne. Per cui se cambiamo base tramite C la matrice A diventa

$$B = CAC = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

A questo punto possiamo scrivere $B = I + N$ dove $N = B - I$ è tale che $N^2 = 0$. Inoltre I ed N commutano, per cui

$$e^{Bt} = e^{It+Nt} = e^{It}e^{Nt} = \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix} e^{Nt}.$$

Poiché $N^2 = 0$, tutte le altre iterate sono nulle e la serie esponenziale ha solo due termini non banali

$$e^{Nt} = \sum_{k=0}^{+\infty} \frac{N^k t^k}{k!} = I + Nt = \begin{pmatrix} 1 & 2t \\ 0 & 1 \end{pmatrix}.$$

Allora si ha

$$e^{At} = e^{CBCt} = Ce^{Bt}C = \begin{pmatrix} e^t & 0 \\ 2te^t & e^t \end{pmatrix},$$

trovando lo stesso risultato.

Il procedimento, apparentemente molto artigianale, utilizzato in quest'ultimo esempio per calcolare l'esponenziale della matrice A è stato quello di determinare una base in cui questa avesse una forma adatta a calcolarne le iterate (somma di matrice diagonale e matrice nilpotente). In realtà questo è esattamente il modo generale di procedere e prende il nome di riduzione di A in forma canonica di Jordan (si veda il Paragrafo 7.2).

Capitolo 7

Soluzioni di sistemi lineari omogenei autonomi

7.1 Esponenziale di matrici diagonalizzabili in \mathbb{C}

Abbiamo visto che una possibile ostruzione alla diagonalizzabilità di una matrice $A \in M_n(\mathbb{R})$ è l'eventuale presenza di autovalori complessi. In questo caso, la strategia per trovare comunque una forma semplificata di A che ci permetta un calcolo agevole di e^{At} è quella di pensare A come un elemento di $\mathcal{L}(\mathbb{C}^n)$, diagonalizzarla (se possibile) in questo spazio e poi tornare ad una opportuna matrice reale. Questo processo può essere formalizzato in termine di complessificazione di spazi vettoriali e applicazioni lineari.

Definizione 7.1.1. Sia V uno spazio vettoriale di dimensione n su \mathbb{R} . Il suo complessificato $V^\mathbb{C}$ è lo spazio vettoriale $V \times V$ su \mathbb{C} con le operazioni

$$(v, v') + (w, w') = (v + w, v' + w') , \quad v, v', w, w' \in V , \quad (7.1)$$

$$\lambda(v, v') = (a + ib)(v, v') = (av - bv', bv + av') , \quad \lambda \in \mathbb{C}, v, v' \in V . \quad (7.2)$$

La verifica del fatto che $V^\mathbb{C}$ sia uno spazio vettoriale su \mathbb{C} è elementare.

La (7.2) implica che $(0, v) = i(v, 0)$, $v \in V$, per cui $(v, v') \in V^\mathbb{C}$ può essere scritto come

$$(v, v') = (v, 0) + (0, v') = (v, 0) + i(v', 0)$$

quindi, se indichiamo con

$$V + iV = \{v + iv' \mid v, v' \in V\}$$

possiamo identificare lo spazio vettoriale $V^{\mathbb{C}}$ con $V + iV$. Inoltre l'applicazione

$$\begin{aligned} V &\longrightarrow V^{\mathbb{C}} \\ v &\longmapsto (v, 0) \end{aligned}$$

è \mathbb{R} -lineare, iniettiva e con immagine $V \times \{0\}$, per cui $V \times \{0\}$ è una copia isomorfa di V in $V^{\mathbb{C}}$. D'ora in poi scriveremo il complessificato come la somma formale $V + iV$.

In particolare, avremo che $\mathbb{R}^{\mathbb{C}}$ è isomorfo a \mathbb{C} , così come $(\mathbb{R}^n)^{\mathbb{C}}$ è isomorfo a \mathbb{C}^n e $M_n(\mathbb{R})^{\mathbb{C}}$ è isomorfo a $M_n(\mathbb{C})$.

L'identificazione tra V e $V \times \{0\} \subset V^{\mathbb{C}}$ permette di costruire una base di $V^{\mathbb{C}}$ a partire da una base di V .

Proposizione 7.1.2. *Se $\{v_1, \dots, v_n\}$ è una base di V su \mathbb{R} allora $\{v_1 + i0, \dots, v_n + i0\}$ è base di $V^{\mathbb{C}}$ su \mathbb{C} . In particolare $\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} V^{\mathbb{C}}$.*

Dimostrazione. I vettori $\{v_1 + i0, \dots, v_n + i0\}$ sono linearmente indipendenti in $V^{\mathbb{C}}$. Infatti se consideriamo una loro combinazione lineare complessa

$$a_1(v_1 + i0) + \dots + a_n(v_n + i0) = 0, \quad a_j = x_j + iy_j, \quad x_j, y_j \in \mathbb{R}, \quad j = 1, \dots, n,$$

questa, a conti fatti, diventa

$$\begin{aligned} (x_1 + iy_1)(v_1 + i0) + \dots + (x_n + iy_n)(v_n + i0) &= \\ &= x_1v_1 + iy_1v_1 + \dots + x_nv_n + iy_nv_n \\ &= x_1v_1 + \dots + x_nv_n + i(y_1v_1 + \dots + y_nv_n) = 0 \end{aligned}$$

ossia deve essere

$$x_1v_1 + \dots + x_nv_n = y_1v_1 + \dots + y_nv_n = 0.$$

Poiché $\{v_1, \dots, v_n\}$ sono linearmente indipendenti su \mathbb{R} , questo implica che $x_j = y_j = 0$ per ogni $j = 1, \dots, n$ provando la lineare indipendenza in $V^{\mathbb{C}}$ dei $\{v_1 + i0, \dots, v_n + i0\}$. Per mostrare che generano $V^{\mathbb{C}}$ si usa la stessa strategia. Infatti se $(v, v') \in V^{\mathbb{C}}$ allora $v, v' \in V$ e quindi

$$v = \sum_{j=1}^n \alpha_j v_j, \quad v' = \sum_{j=1}^n \beta_j v_j$$

per opportuni coefficienti reali. Allora

$$v + iv' = \sum_{j=1}^n (\alpha_j v_j + i\beta_j v_j) = \sum_{j=1}^n (\alpha_j + i\beta_j)(v_j + i0).$$

□

Ad un'applicazione \mathbb{R} -lineare $A: V \rightarrow V$ si può associare in maniera canonica un'applicazione \mathbb{C} -lineare sul complessificato $V^\mathbb{C}$.

Definizione 7.1.3. Sia $A: V \rightarrow V$ un'applicazione \mathbb{R} -lineare. L'applicazione lineare $A^\mathbb{C}: V^\mathbb{C} \rightarrow V^\mathbb{C}$ definita da

$$A^\mathbb{C}(v + iv') = Av + iAv'$$

prende il nome di complessificata di A .

Viene lasciata al lettore la verifica che $A^\mathbb{C}$ sia effettivamente \mathbb{C} -lineare.

Osservazione 7.1.4. Sia $A \in M_n(\mathbb{R})$ la matrice che rappresenta un'applicazione lineare $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ nella base $\{v_1, \dots, v_n\}$, per cui ha la forma $A = (f(v_1) | \dots | f(v_n))$. Poiché, per definizione, $f^\mathbb{C}(v_j + i0) = f(v_j)$, la stessa matrice A , pensata come elemento di $M_n(\mathbb{C})$, rappresenta l'operatore complessificato $f^\mathbb{C}: \mathbb{C}^n \rightarrow \mathbb{C}^n$ nella base $\{v_1 + i0, \dots, v_n + i0\}$. Quindi, se $p(\lambda)$ è il polinomio caratteristico di A in \mathbb{R} , lo stesso polinomio, esteso a valori complessi, è il polinomio caratteristico di A pensato come elemento di $M_n(\mathbb{C})$. Questo chiarisce l'abuso di linguaggio sugli autovalori di una matrice reale che abbiamo introdotto nell'Osservazione 6.2.8: noi penseremo sempre $A \in M_n(\mathbb{R})$ come un elemento di $M_n(\mathbb{C})$, identificandola con la sua complessificata, e considereremo il suo spettro in \mathbb{C} .

Esempio 7.1.5. La matrice

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

rappresenta sia l'operatore $A(x, y) = (-y, x)$ in \mathbb{R}^2 che l'operatore $A^\mathbb{C}(z_1, z_2) = (-z_2, z_1)$ in \mathbb{C}^2 . Il vantaggio è che il polinomio caratteristico di A è $p_A(\lambda) = \lambda^2 + 1$ che su \mathbb{R} è irriducibile, mentre su \mathbb{C} ha le radici $\lambda = i, -i$. Per cui A

non è diagonalizzabile su \mathbb{R}^2 ma su \mathbb{C} si diagonalizza attraverso la matrice di trasformazione

$$C = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}.$$

A questo punto, data una matrice $A \in M_n(\mathbb{R})$, possiamo considerarla come elemento di $M_n(\mathbb{C})$ e provare a diagonalizzarla. Se risulta diagonalizzabile in \mathbb{C} , allora avremo a disposizione una base di \mathbb{C}^n di autovettori. A partire da questa base, vogliamo determinare una base per \mathbb{R}^n in cui la matrice A si presenta diagonale a blocchi, in modo da poter utilizzare la fattorizzazione dell'esponenziale descritta nell'Esempio 6.5.3.

Innanzi tutto vediamo cosa possiamo dire dell'autospazio in \mathbb{C}^n relativo ad un autovalore reale $\lambda \in \sigma(A)$. Ci piacerebbe poter dire che questo autospazio è a sua volta il complessificato dell'autospazio in \mathbb{R}^n relativo a λ , in modo tale da poter usare la Proposizione 7.1.2 e determinare una base di autovettori reali per quel sottospazio. Questo sarà possibile perché l'autospazio in \mathbb{C}^n relativo a $\lambda \in \mathbb{R}$ è chiuso rispetto all'operazione di coniugio: se $z \in \mathbb{C}^n$ è un autovettore complesso relativo a λ , allora anche il suo coniugato \bar{z} è un autovettore relativo a λ .

La questione generale da porsi è la seguente: dato uno spazio vettoriale V di dimensione finita, quando un sottospazio vettoriale di $V^\mathbb{C}$ è il complessificato di un sottospazio vettoriale di V ? La risposta questa domanda è che $W \subseteq V^\mathbb{C}$ è il complessificato di un sottospazio di V se e solo se contiene i coniugati di tutti i suoi elementi, dove la mappa di coniugio sul complessificato è definita da

$$\begin{aligned} J : V^\mathbb{C} &\longrightarrow V^\mathbb{C} \\ v + iv' &\longmapsto v - iv' \end{aligned} \tag{7.3}$$

Lemma 7.1.6. *Sia W un sottospazio vettoriale di $V^\mathbb{C}$ con dimensione $\dim_{\mathbb{C}} W = k$. Allora esiste $S \subseteq V$ sottospazio vettoriale di dimensione $\dim_{\mathbb{R}} S = k$ tale che $S^\mathbb{C} = W$ se e solo se $J(W) = W$. In questo caso useremo le notazioni $W = S + iS$ e $S = \operatorname{Re} W$.*

Dimostrazione. Supponiamo che esista un sottospazio S di V tale che $W = S^\mathbb{C}$. Allora $J(W) = J(S + iS) = S + iS = W$.

Viceversa, supponiamo $J(W) = W$ e definiamo

$$S = \operatorname{Re} W = \{v \in V \mid v + i0 \in W\}.$$

Mostriamo che $S^{\mathbb{C}} = W$. Se $v + iv' \in S^{\mathbb{C}}$ allora $v, v' \in S$ e per definizione $v + i0, v' + i0 \in W$. Quindi $v + iv' = (v + i0) + i(v' + i0) \in W$, provando $S^{\mathbb{C}} \subseteq W$. Viceversa sia $v + iv' \in W$. Poiché $J(W) = W$, avremo che anche $v - iv'$ appartiene a W . Ma allora

$$\begin{aligned}\frac{1}{2}((v + iv') + (v - iv')) &= v + i0 \in W, \\ \frac{1}{2i}((v + iv') - (v - iv')) &= v' + i0 \in W,\end{aligned}$$

quindi per definizione $v, v' \in S$ cioè $v + iv' \in S^{\mathbb{C}}$. \square

Ovviamente il risultato precedente non si può applicare agli autospazi relativi ad autovalori complessi. In questo caso, però, è vero è che se μ è un autovalore complesso per $A \in M_n(\mathbb{R})$ e $z \in \mathbb{C}^n$ è un suo autovettore, allora anche $\bar{\mu}$ è un autovalore di A e \bar{z} è un autovettore relativo a $\bar{\mu}$. Quindi, in questo caso, la somma diretta dei due autospazi è chiusa per coniugio e, di conseguenza, è il complessificato di un sottospazio di \mathbb{R}^n .

Il risultato generale che descrive questa situazione è il seguente.

Lemma 7.1.7. *Sia $F \subseteq V^{\mathbb{C}}$ un sottospazio di $V^{\mathbb{C}}$ tale che $F \cap J(F) = \{0\}$ e sia $\{v_1 + iv'_1, \dots, v_k + iv'_k\}$ una base di F con $k < n$. Allora $W = F \oplus J(F)$ è tale che $W = J(W)$ e $\{v_1, v'_1, \dots, v_k, v'_k\}$ è una base di $\operatorname{Re} W$.*

Dimostrazione. Poiché $J^2 = Id$, si ha $J(W) = W$. Dal Lemma 7.1.6 segue allora che esiste $\operatorname{Re} W$ sottospazio di V tale che $(\operatorname{Re} W)^{\mathbb{C}} = W$, con $\dim_{\mathbb{R}} \operatorname{Re} W = \dim_{\mathbb{C}} W = 2k$. Quindi basta dimostrare che i vettori $\{v_1, v'_1, \dots, v_k, v'_k\}$ sono linearmente indipendenti su \mathbb{R} . Sia

$$\sum_{j=1}^k (\alpha_j v_j + \beta_j v'_j) = 0$$

una loro combinazione lineare con coefficienti $\alpha_j, \beta_j \in \mathbb{R}$. Il membro di sinistra si scrive come

$$\sum_{j=1}^k \left[\frac{1}{2}(\alpha_j - i\beta_j)(v_j + iv'_j) + \frac{1}{2}(\alpha_j + i\beta_j)J(v_j + iv'_j) \right] = 0$$

cioè una combinazione lineare a coefficienti in \mathbb{C} dei vettori della base di W , per cui i coefficienti sono nulli: $\alpha_j = \beta_j = 0$. \square

A questo punto siamo in grado di trasformare una matrice che ha complessificata diagonalizzabile in una matrice diagonale a blocchi i cui coefficienti dipendono solo dai suoi autovalori (reali e complessi).

Teorema 7.1.8. *Sia $A \in M_n(\mathbb{R})$ con $n \geq 2$. Supponiamo che A abbia $h \leq n$ autovalori reali distinti $\lambda_j, j = 1, \dots, h$ con molteplicità algebrica $\text{ma}(\lambda_j)$ e $2k$ autovalori complessi $\mu_j = \alpha_j + i\beta_j, \bar{\mu}_j = \alpha_j - i\beta_j, j = 1, \dots, k$ con molteplicità algebrica $\text{ma}(\mu_j) = \text{ma}(\bar{\mu}_j)$. Se la complessificata $A^{\mathbb{C}} \in M_n(\mathbb{C})$ è diagonalizzabile, allora esiste una matrice invertibile $C \in M_n(\mathbb{R})$ tale che $A = CBC^{-1}$ con B matrice diagonale a blocchi della forma*

$$B = \begin{pmatrix} \lambda_1 I_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_h I_h & 0 & \cdots & 0 \\ 0 & \cdots & 0 & B_1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & B_k \end{pmatrix}. \quad (7.4)$$

dove $I_j = I_{\text{ma}(\lambda_j)}$ e $B_j \in M_{2\text{ma}(\mu_j)}(\mathbb{R})$ è una matrice diagonale a blocchi della forma

$$B_j = \left(\begin{array}{cc|c|c} \alpha_j & \beta_j & \cdots & 0 \\ -\beta_j & \alpha_j & & \\ \hline \vdots & \ddots & & \vdots \\ \hline 0 & \cdots & \alpha_j & \beta_j \\ & & -\beta_j & \alpha_j \end{array} \right)$$

Dimostrazione. Sia E_{λ} l'autospazio associato all'autovalore λ in \mathbb{C}^n di $A^{\mathbb{C}}$. Si ha che $J(E_{\lambda_j}) = E_{\lambda_j}$ per ogni $j = 1, \dots, h$ e $J(E_{\mu_j}) = E_{\bar{\mu}_j}$ per ogni $j = 1, \dots, k$. Per ipotesi $A^{\mathbb{C}}$ è diagonalizzabile e quindi \mathbb{C}^n si decompone nella

somma diretta degli autospazi

$$\mathbb{C}^n = \bigoplus_{j=1}^h E_{\lambda_j} \oplus \bigoplus_{j=1}^k (E_{\mu_j} + J(E_{\mu_j})) .$$

Per il Lemma 7.1.6 abbiamo che $E_{\lambda_j} = (\operatorname{Re} E_{\lambda_j})^\mathbb{C}$ e per il Lemma 7.1.7, dal momento che gli autospazi sono disgiunti, $E_{\mu_j} \oplus J(E_{\mu_j}) = (\operatorname{Re}(E_{\mu_j} \oplus J(E_{\mu_j})))^\mathbb{C}$. Allora \mathbb{R}^n si decompone nella somma diretta

$$\mathbb{R}^n = \bigoplus_{j=1}^h \operatorname{Re}(E_{\lambda_j}) \oplus \bigoplus_{j=1}^k \operatorname{Re}((E_{\mu_j} + J(E_{\mu_j}))) .$$

Vediamo che forma ha l'operatore A su questa decomposizione. Se un vettore $v \neq 0$ appartiene a $\operatorname{Re}(E_{\lambda_j})$, allora $v + i0 \in E_{\lambda_j}$ per cui si ha $A^\mathbb{C}(v + i0) = \lambda_j(v + i0)$ e per definizione di $A^\mathbb{C}$ si ottiene $Av = \lambda_j v$ essendo $\lambda_j \in \mathbb{R}$. Quindi i sottospazi $\operatorname{Re}(E_{\lambda_j})$ sono gli autospazi reali relativi a λ_j .

Per il Lemma 7.1.7 una base di $\operatorname{Re}(E_{\mu_j} + J(E_{\mu_j}))$ è data dai vettori $\{v_1, v'_1, \dots, v_l, v'_l\}$ corrispondenti alla base $\{v_1 + iv'_1, \dots, v_l + iv'_l\}$ di E_{μ_j} . Quindi basta calcolare A su questa base. Se $v + iv' \in E_{\mu_j}$ allora per definizione di autospazio

$$\begin{aligned} Av + iAv' &= A^\mathbb{C}(v + iv') = \mu_j(v + iv') = (\alpha_j + i\beta_j)(v + iv') \\ &= (\alpha_j v - \beta_j v') + i(\alpha_j v' + \beta_j v) \end{aligned}$$

e quindi

$$\begin{cases} Av = \alpha_j v - \beta_j v', \\ Av' = \beta_j v + \alpha_j v'. \end{cases}$$

Dunque sulla coppia di vettori $\{v_j, v'_j\}$ della base, l'operatore A agisce come il blocco B_j . \square

Esempio 7.1.9. Consideriamo la matrice

$$A = \begin{pmatrix} 5 & 5 & 3 \\ -1 & 3 & 1 \\ 0 & -4 & 0 \end{pmatrix} .$$

Gli autovalori di A sono complessi: $\sigma(A) = \{4, 2+2i, 2-2i\}$. Il Teorema 7.1.8 ci garantisce che in qualche base di \mathbb{R}^3 la matrice A è della forma

$$B = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & -2 & 2 \end{pmatrix} .$$

Vogliamo calcolare esplicitamente tale base, seguendo la dimostrazione del Teorema. Siano allora $\lambda = 4$, $\mu = 2 + 2i$. Hanno entrambi molteplicità algebrica 1. Abbiamo che \mathbb{R}^3 si decomponga nella somma diretta

$$\mathbb{R}^3 = \text{Re}(E_\lambda) \oplus \text{Re}(E_\mu \oplus E_{\bar{\mu}}).$$

Allora è sufficiente scegliere una base di questi sottospazi. Per quanto visto, l'autospazio relativo all'autovalore λ reale è semplice da trattare: una sua base è data da una base di E_λ come sottospazio \mathbb{R}^3 . In tal caso è sufficiente trovare un autovettore relativo a λ , per esempio $u = (2, -1, 1)$.

Una base di $\text{Re}(E_\mu \oplus E_{\bar{\mu}})$ è data dalla coppia di vettori $\{v, v'\}$ tali che $\{v + iv'\}$ è base di E_μ . Allora basta trovare un autovettore di $A^\mathbb{C}$ corrispondente a μ . Un rapido conto fornisce $w = (-1 + i, -1 - i, 2)$, che può essere scritto come $w = v + iv'$ con $v = (-1, -1, 2)$, $v' = (1, -1, 0)$. Nella base $\mathcal{B} = \{u, v, v'\}$ la matrice A assume la forma B , cioè $B = C^{-1}AC$ con $C = (u|v|v')$ matrice di cambiamento di coordinate dalla base \mathcal{B} alla base standard di \mathbb{R}^3 .

A questo punto siamo in grado di calcolare l'esponenziale di una matrice $A \in M_n(\mathbb{R})$ tale che la sua complessificata $A^\mathbb{C}$ sia diagonalizzabile su \mathbb{C}^n . Sia $C \in M_n(\mathbb{R})$ invertibile tale che $A = CBC^{-1}$, con B data dalla (7.4). Quindi l'esponenziale di A è $e^{At} = Ce^{Bt}C^{-1}$ e il problema si riduce a quello di calcolare l'esponenziale di B . Ma ora B è una matrice diagonale a blocchi le cui iterate sono

$$B^m = \begin{pmatrix} \lambda_1^m I_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_h^m I_h & 0 & \cdots & 0 \\ 0 & \cdots & 0 & B_1^m & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & B_k^m \end{pmatrix}.$$

Quindi l'esponenziale di B è la matrice diagonale a blocchi

$$e^{Bt} = \text{diag}\{e^{\lambda_1 t} I_1, \dots, e^{\lambda_h t} I_h, e^{B_1 t}, \dots, e^{B_k t}\}.$$

Restano da calcolare gli esponenziali dei blocchi B_j . La matrice B_j è quadrata di ordine 2 ma(μ_j) e diagonale a blocchi, con blocchi della forma

$$M = \begin{pmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{pmatrix}$$

dove $\mu_j = \alpha_j + i\beta_j$. Osserviamo che

$$M = \alpha_j \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta_j \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

La matrice

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

prende il nome di matrice simplettica fondamentale. Le potenze di J si calcolano facilmente:

$$J^{2k} = (-1)^k I, \quad J^{2k+1} = (-1)^k J$$

e quindi

$$\begin{aligned} e^{\beta_j J t} &= \sum_{k=0}^{+\infty} \frac{\beta_j^k J^k t^k}{k!} \\ &= \left(\sum_{k=0}^{+\infty} \frac{(-1)^k (\beta_j t)^{2k}}{(2k)!} \right) I + \left(\sum_{k=0}^{+\infty} \frac{(-1)^k (\beta_j t)^{2k+1}}{(2k+1)!} \right) J \\ &= \cos(\beta_j t) I + \sin(\beta_j t) J \\ &= \begin{pmatrix} \cos(\beta_j t) & \sin(\beta_j t) \\ -\sin(\beta_j t) & \cos(\beta_j t) \end{pmatrix}. \end{aligned}$$

Poiché le matrici $\alpha_j I$ e $\beta_j J$ commutano, per la Proposizione 6.4.2 si ha

$$\begin{aligned} e^{B_j t} &= e^{\alpha_j I t + \beta_j J t} = e^{\alpha_j I} e^{\beta_j J t} \\ &= e^{\alpha_j t} \begin{pmatrix} \cos(\beta_j t) & \sin(\beta_j t) \\ -\sin(\beta_j t) & \cos(\beta_j t) \end{pmatrix}. \end{aligned}$$

In conclusione, abbiamo ottenuto il seguente risultato.

Teorema 7.1.10. *Sia $A \in M_n(\mathbb{R})$ diagonalizzabile in campo complesso e indichiamo con $\lambda_j, j = 1, \dots, h$, i suoi autovalori reali e con $\mu_j = \alpha_j + i\beta_j$, $\bar{\mu}_j = \alpha_j - i\beta_j$, $j = 1, \dots, k$, i suoi autovalori complessi, $h + 2k = n$. Allora esiste una matrice invertibile $C \in M_n(\mathbb{R})$ tale che la matrice fondamentale speciale U del sistema lineare omogeneo $x' = Ax$ è data da $U = CBC^{-1}$ con B matrice a blocchi della forma*

$$B = \begin{pmatrix} e^{\lambda_1 t} I_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_h} I_h & 0 & \cdots & 0 \\ 0 & \cdots & 0 & B_1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & B_k \end{pmatrix}, \quad (7.5)$$

dove $I_j = I_{\text{ma } \lambda_j}$ e $B_j \in M_{2 \text{ ma } (\mu_j)}(\mathbb{R})$ è una matrice a blocchi della forma

$$B_j = \left(\begin{array}{cc|c|c} e^{\alpha_j t} \cos(\beta_j t) & e^{\alpha_j t} \sin(\beta_j t) & \dots & 0 \\ -e^{\alpha_j t} \sin(\beta_j t) & e^{\alpha_j t} \cos(\beta_j t) & \dots & \vdots \\ \hline \vdots & \ddots & & \vdots \\ \hline 0 & \dots & e^{\alpha_j t} \cos(\beta_j t) & e^{\alpha_j t} \sin(\beta_j t) \\ & & -e^{\alpha_j t} \sin(\beta_j t) & e^{\alpha_j t} \cos(\beta_j t) \end{array} \right).$$

In particolare, le componenti delle soluzioni del sistema lineare $x' = Ax$ sono sempre combinazioni lineari delle funzioni

$$e^{\lambda_1 t}, \dots, e^{\lambda_h t}, \quad (7.6)$$

$$e^{\alpha_1 t} \cos(\beta_1 t), e^{\alpha_1 t} \sin(\beta_1 t), \dots, e^{\alpha_k t} \cos(\beta_k t), e^{\alpha_k t} \sin(\beta_k t). \quad (7.7)$$

Esempio 7.1.11. Determiniamo l'integrale generale del sistema lineare omogeneo

$$\begin{cases} x' = 2x - y - z, \\ y' = x + 2y + 5z, \\ z' = -3z. \end{cases}$$

Per quanto visto finora, si tratta di calcolare l'esponenziale della matrice

$$A = \begin{pmatrix} 2 & -1 & -1 \\ 1 & 2 & 5 \\ 0 & 0 & -3 \end{pmatrix}.$$

Il primo passo è trovare una base di \mathbb{R}^3 in cui A abbia la forma (7.4). Per farlo seguiremo la dimostrazione del Teorema 7.1.8. Gli autovalori di A su \mathbb{C} sono $\sigma(A) = \{-3, 2+i, 2-i\}$. Chiamiamo $\lambda = -3$ e $\mu = 2+i$. Quindi $\mu = \alpha + i\beta$ con $\alpha = 2$ e $\beta = 1$. Lo spazio si decompone come

$$\mathbb{R}^3 = \text{Re}(E_\lambda) \oplus \text{Re}(E_\mu \oplus J(E_\mu))$$

e in una qualsiasi base dei due sottospazi $\text{Re}(E_\lambda)$ e $\text{Re}(E_\mu \oplus J(E_\mu))$ la matrice A assume la forma

$$B = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}.$$

Una base di $\text{Re}(E_\lambda)$ è data da una base dell'autospazio relativo a λ su \mathbb{R} . Poiché tale autospazio ha dimensione 1, è generato da un qualsiasi autovettore di autovalore λ , per esempio $u = (0, -1, 1)$.

Sappiamo poi che una base $\{v, w\}$ di $\text{Re}(E_\mu \oplus J(E_\mu))$ su \mathbb{R} corrisponde alla base $\{v + iw\}$ di E_μ su \mathbb{C} . Quindi, visto che $z = (i, 1, 0)$ è un autovettore complesso di A di autovalore μ e l'autospazio relativo è unidimensionale, i vettori $v = (0, 1, 0), w = (1, 0, 0)$ sono una base per $\text{Re}(E_\mu \oplus J(E_\mu))$. Se $C = (u|v|w)$ è la matrice di cambiamento di base otteniamo $B = C^{-1}AC$. Ora calcoliamo l'esponenziale $e^{At} = e^{CBC^{-1}t} = Ce^{Bt}C^{-1}$. Poiché B è diagonale a blocchi della forma $B = \text{diag}\{-3I_1, B_\mu\}$ con

$$B_\mu = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$$

il suo esponenziale è dato da $e^{Bt} = \text{diag}\{e^{-3t}I_1, e^{B_\mu t}\}$. D'altra parte sappiamo che

$$e^{B_\mu t} = e^{2t} \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

e dunque si ha

$$\begin{aligned} e^{At} &= \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} e^{-3t} & 0 & 0 \\ 0 & e^{2t} \cos(t) & e^{2t} \sin(t) \\ 0 & -e^{2t} \sin(t) & e^{2t} \cos(t) \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} e^{2t} \cos(t) & -e^{2t} \sin(t) & -e^{2t} \sin(t) \\ e^{2t} \sin(t) & e^{2t} \cos(t) & e^{2t} \cos(t) - e^{-3t} \\ 0 & 0 & e^{-3t} \end{pmatrix}. \end{aligned}$$

Ne concludiamo che la soluzione del sistema corrispondente al dato iniziale $x(0) = \xi$, $\xi = (\xi_1, \xi_2, \xi_3)$ è data da

$$\begin{aligned} e^{At}\xi &= \begin{pmatrix} e^{2t} \cos(t) & -e^{2t} \sin(t) & -e^{2t} \sin(t) \\ e^{2t} \sin(t) & e^{2t} \cos(t) & e^{2t} \cos(t) - e^{-3t} \\ 0 & 0 & e^{-3t} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \\ &= \begin{pmatrix} \xi_1 e^{2t} \cos(t) - \xi_2 e^{2t} \sin(t) - \xi_3 e^{2t} \sin(t) \\ \xi_1 e^{2t} \sin(t) + \xi_2 e^{2t} \cos(t) - \xi_3 (e^{2t} \cos(t) - e^{-3t}) \\ \xi_3 e^{-3t} \end{pmatrix}. \end{aligned}$$

Esempio 7.1.12 (Ancora l'oscillatore armonico). Siamo arrivati al punto di poter determinare rigorosamente le soluzioni dell'equazione dell'oscillatore armonico

$$x'' + \omega x = 0,$$

introdotta in (2.3) e richiamata in nell'Esempio 5.1.6. Il sistema lineare è $x' = Ax$ associato a questa equazione del secondo ordine è determinato dalla matrice

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}.$$

Gli autovalori di A sono $\sigma(A) = \{i\omega, -i\omega\}$ per cui essendo distinti $A^{\mathbb{C}}$ è diagonalizzabile su \mathbb{C} . Si ha che $\mathbb{C}^2 = E_{\mu} \oplus E_{\bar{\mu}}$, dove abbiamo posto $\mu = i\omega$. Tale decomposizione di \mathbb{C}^2 induce la decomposizione $\mathbb{R}^2 = \text{Re}(E_{\mu}) \oplus J(E_{\mu})$ e una base di tale spazio è data da $\{v, w\}$ con $\{v + iw\}$ base di E_{μ} su \mathbb{C} . Un autovettore relativo all'autovalore μ è per esempio $z = (-i, \omega)$, per cui $z = v + iw = (0, \omega) + i(-1, 0)$.

I vettori $\{v, w\}$ determinano la matrice $C = (v|w)$ per la quale $A = C^{-1}BC$ dove

$$B = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}.$$

Possiamo allora calcolare e^{At} :

$$\begin{aligned} e^{At} &= C^{-1}e^{Bt}C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix} \begin{pmatrix} 0 & -1 \\ \omega & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\omega t) & \frac{1}{\omega} \sin(\omega t) \\ -\omega \sin(\omega t) & \cos(\omega t) \end{pmatrix} \end{aligned}$$

ritrovando la matrice fondamentale speciale calcolata nell'Esempio 5.1.6.

7.2 Forma canonica di Jordan

Abbiamo risolto completamente il problema di determinare l'esponenziale di $A \in M_n(\mathbb{R})$ nel caso di $A^{\mathbb{C}}$ diagonalizzabile in \mathbb{C}^n . Vogliamo ora discutere il caso generale, in cui $A^{\mathbb{C}}$ può non essere diagonalizzabile in \mathbb{C} perché qualche suo autovalore ha molteplicità geometrica inferiore a quella algebrica, per cui \mathbb{C}^n non si può decomporre come somma diretta di autospazi.

L'osservazione chiave che chiarisce come affrontare il problema è la seguente:

1. $\ker(A - \lambda I) \subseteq \ker(A - \lambda I)^2$;
2. $\ker(A - \lambda I)^2$ è un sottospazio di \mathbb{C}^n invariante sotto l'azione di A .

La prima affermazione è banale. Verifichiamo la seconda: dobbiamo vedere che se $v \in \ker(A - \lambda I)^2$ allora anche $Av \in \ker(A - \lambda I)^2$. Ma

$$(A - \lambda I)^2(Av) = A(A - \lambda I)^2v = A0 = 0$$

perché A commuta con se stessa e con I e quindi commuta anche con $(A - \lambda I)^2$. Per induzione si prova che $\ker(A - \lambda I)^k \subseteq \ker(A - \lambda I)^{k+1}$ per ogni $k \in \mathbb{N}$, e che sono tutti sottospazi A invarianti.

Abbiamo quindi a disposizione una successione di sottospazi di \mathbb{C}^n

$$\{0\} \subseteq \ker(A - \lambda I) \subseteq \ker(A - \lambda I)^2 \subseteq \cdots \subseteq \ker(A - \lambda I)^k \subseteq \cdots \subseteq \mathbb{C}^n. \quad (7.8)$$

Se $\text{mg}(\lambda) = \text{ma}(\lambda)$, sappiamo che il giusto sottospazio per la decomposizione è l'autospazio associato a λ . Proveremo che, quando $\text{mg}(\lambda) < \text{ma}(\lambda)$, tramite queste iterazioni è possibile determinare un sottospazio invariante che abbia dimensione $\text{ma}(\lambda)$ e su cui la matrice $A - \lambda I$ sia nilpotente. Riusciremo a decomporre \mathbb{C}^n come somma diretta di questi sottospazi e, nella nuova base, A risulterà essere la somma di una matrice diagonale ed una nilpotente, il che ci permetterà di calcolare agevolmente la matrice esponenziale.

In questo paragrafo considereremo sempre matrici $A \in M_n(\mathbb{C})$. Indicheremo con $\lambda_j \in \mathbb{C}$, $j = 1, \dots, r$ gli autovalori distinti di A . Se $m_j = \text{ma}(\lambda_j)$ si ha che il polinomio caratteristico fattorizza in

$$p_A(\lambda) = \prod_{j=1}^r (\lambda - \lambda_j)^{m_j}.$$

Definiamo ora gli autovettori generalizzati.

Definizione 7.2.1. Si dice che $v \in \mathbb{C}^n \setminus \{0\}$ è un autovettore generalizzato relativo all'autovalore $\lambda \in \sigma(A)$ se

$$(A - \lambda I)^k v = 0$$

per qualche $k \in \mathbb{N}$. Indicheremo con

$$E_\lambda = \left\{ v \in \mathbb{C}^n \setminus \{0\} \mid \exists k \in \mathbb{N} : (A - \lambda I)^k v = 0 \right\} \cup \{0\}$$

l'autospazio generalizzato associato all'autovalore λ .

Proposizione 7.2.2. *Sia $A \in M_n(\mathbb{C})$, $\lambda \in \sigma(A)$, allora*

- i) *la successione dei nuclei (7.8) si stabilizza, ovvero esiste un $d = d(\lambda) \in \mathbb{N}$ tale che*

$$\ker(A - \lambda I)^{d-1} \subsetneq \ker(A - \lambda I)^d = \ker(A - \lambda I)^k$$

per ogni $k \geq d$. Come conseguenza, se $v \in E_\lambda$ allora $(A - \lambda I)^d v = 0$;

- ii) *gli autospazi generalizzati sono A -invarianti, cioè $A E_\lambda \subseteq E_\lambda$;*

- iii) *la restrizione ad E_λ di $A - \lambda I$ è tale che*

$$((A - \lambda I)|_{E_\lambda})^{d(\lambda)} = (A|_{E_\lambda} - \lambda I_{E_\lambda})^{d(\lambda)} = 0;$$

iv) $\sigma(A|_{E_\lambda}) = \{\lambda\}$.

Dimostrazione. Poiché $\dim(\ker(A - \lambda I)^k) \leq n$ per ogni k , sicuramente deve esistere $k \geq 1$ tale che

$$\ker(A - \lambda I)^k = \ker(A - \lambda I)^{k+1}. \quad (7.9)$$

D'altra parte, se per k vale (7.9), allora necessariamente si ha

$$\ker(A - \lambda I)^{k+r} = \ker(A - \lambda I)^{k+r+1}, \quad \forall r \in \mathbb{N}. \quad (7.10)$$

Infatti, se $v \in \ker(A - \lambda I)^{k+r+1}$, allora $(A - \lambda I)^r v \in \ker(A - \lambda I)^{k+1} = \ker(A - \lambda I)^k$ e quindi $\ker(A - \lambda I)^{k+r+1} \subseteq \ker(A - \lambda I)^{k+r}$. A questo punto basta scegliere $d(\lambda)$ come il più piccolo k per cui vale la (7.9).

Per dimostrare la ii), basta osservare che, dal momento che A commuta con se stessa e con l'identità, commuta anche con $(A - \lambda I)^d$ e quindi si ha

$$(A - \lambda I)^d(Av) = A(A - \lambda I)^d v = A0 = 0, \quad \forall v \in E_\lambda$$

e quindi $Av \in E_\lambda$ per ogni $v \in E_\lambda$.

Chiaramente la proprietà iii) segue da i) ed ii).

Infine, la iii) ci dice che la matrice $(A|_{E_\lambda} - \lambda I_{E_\lambda})$ è nilpotente, quindi $\sigma((A|_{E_\lambda} - \lambda I_{E_\lambda})) = \{0\}$, ossia $\sigma(A|_{E_\lambda}) = \{\lambda\}$. \square

Per dimostrare che \mathbb{C}^n si decomponga nella somma diretta degli autospazi generalizzati abbiamo bisogno di alcuni risultati preliminari.

Lemma 7.2.3 (Triangolazione di Schur). *Sia $A \in M_n(\mathbb{C})$. Allora esiste una matrice unitaria $C \in M_n(\mathbb{C})$ tale che $A = CBC^{-1}$ dove B è una matrice triangolare superiore.*

Dimostrazione. Procediamo per induzione sulla dimensione n . Se $n = 1$ non c'è niente da provare. Supponiamo che la tesi sia vera per $n - 1$. Poiché siamo in \mathbb{C} , la matrice A ammette almeno un autovalore λ_1 . Sia v un autovettore relativo a λ e sia $v_1 = v / \|v\|$ il relativo versore. Indichiamo con $E = \{v \in \mathbb{C}^n \mid \langle v, v_1 \rangle = 0\}$ il sottospazio ortogonale a v_1 . Si ha che $\mathbb{C}^n = \text{span}\{v_1\} \oplus E$ e, rispetto ad una base $\{v_1, u_2, \dots, u_n\}$ relativa a questa decomposizione, A si trasforma nella matrice

$$\begin{pmatrix} \lambda_1 & * & * & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}.$$

D'altra parte, il sottospazio E è invariante sotto l'azione di A_1 , e di conseguenza, per ipotesi induttiva, esiste una base ortonormale $\{v_2, \dots, v_n\}$ di E in cui la matrice A_1 è triangolare superiore. Ne segue che, nella base ortonormale $\{v_1, v_2, \dots, v_n\}$, la matrice A si trasforma in una matrice triangolare superiore. \square

Data una matrice $A \in M_n(\mathbb{C})$, sia p_A il suo polinomio caratteristico, che sappiamo essere fattorizzabile come

$$p_A(\lambda) = \prod_{k=1}^r (\lambda - \lambda_k)^{m_k}$$

con $\lambda_k \in \sigma(A)$, $m_k = \text{ma}(\lambda_k)$. Nell'algebra delle matrici ha perfettamente senso considerare la matrice $p_A(A)$ e, poiché in $p_A(A)$ compaiono solo iterate della matrice A e la matrice identità, tutti elementi che commutano tra loro, dunque vale la fattorizzazione

$$p_A(A) = \prod_{k=1}^r (A - \lambda_k I)^{m_k}.$$

La fattorizzazione di $p_A(A)$ giocherà un ruolo centrale nel seguito, grazie al seguente fondamentale risultato.

Teorema 7.2.4 (Cayley-Hamilton). *Sia $A \in M_n(\mathbb{C})$ e $p_A(\lambda) = \det(A - \lambda I)$ il suo polinomio caratteristico. Allora $p(A) = 0$ in $M_n(\mathbb{C})$.*

Dimostrazione. Denotiamo con

$$p(\lambda) = \det(A - \lambda I) = \sum_{k=0}^n c_k \lambda^k, \lambda \in \mathbb{C}$$

il polinomio caratteristico di A . Quindi la matrice $p(A)$ è

$$p(A) = \sum_{k=0}^n c_k A^k.$$

Supponiamo che A sia diagonale. Allora

$$\begin{aligned} p(A) &= \sum_{k=0}^n c_k \operatorname{diag}\{\lambda_1^k, \dots, \lambda_n^k\} = \operatorname{diag}\left\{\sum_{k=0}^n c_k \lambda_1^k, \dots, \sum_{k=0}^n c_k \lambda_n^k\right\} \\ &= \operatorname{diag}\{p(\lambda_1), \dots, p(\lambda_n)\} = 0. \end{aligned}$$

Quindi la proprietà di annullamento vale per le matrici diagonali e, essendo il polinomio caratteristico invariante per equivalenza, vale anche per tutte le matrici diagonalizzabili.

Infine, nel caso generale, usiamo il fatto che, per il Lemma 7.2.3, ogni matrice A è equivalente ad una matrice triangolare superiore, quindi basta provare il risultato per questo tipo di matrici. Sia allora A della forma

$$A = \begin{pmatrix} \lambda_1 & a_{1,2} & \dots & a_{1,n} \\ 0 & \lambda_2 & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Nella Proposizione 6.3.8 abbiamo provato che esistono $\varepsilon_1, \dots, \varepsilon_n \in (0, \varepsilon)$ tale che

$$A_\varepsilon = \begin{pmatrix} \lambda_1 - \varepsilon_1 & a_{1,2} & \dots & a_{1,n} \\ 0 & \lambda_2 - \varepsilon_2 & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n - \varepsilon_n \end{pmatrix}$$

è diagonalizzabile e $\|A - A_\varepsilon\|_{\mathcal{L}} < \varepsilon$. Osserviamo che il polinomio caratteristico di A_ε ha per radici $\{\lambda_1 - \varepsilon_1, \dots, \lambda_n - \varepsilon_n\}$ e, di conseguenza, $p(\lambda) = p_\varepsilon(\lambda) + O(\varepsilon)$, da cui ricaviamo

$$p(A) = \lim_{\varepsilon \rightarrow 0} p_\varepsilon(A_\varepsilon) = 0.$$

□

Teorema 7.2.5 (della mappa spettrale). *Sia $A \in M_n(\mathbb{C})$ e sia p un polinomio a coefficienti in \mathbb{C} . Allora*

$$p(\sigma(A)) = \sigma(p(A)),$$

cioè gli autovalori della matrice $p(A)$ sono tutti e soli i numeri complessi del tipo $p(\lambda)$, $\lambda \in \sigma(A)$.

Dimostrazione. Sia p un polinomio fissato. Ricordiamo che se $\lambda \in \sigma(A)$ e $v \neq 0$ è un suo autovettore, allora $A^k v = \lambda^k v$ per ogni $k \in \mathbb{N}$ e quindi

$$p(A)v = \sum_{k=0}^r c_k A^k v = \sum_{k=0}^r c_k \lambda^k v = p(\lambda)v,$$

ossia v un autovettore associato all'autovalore $p(\lambda)$ della matrice $p(A)$.

D'altra parte, se $\lambda \in \sigma(p(A))$, allora la matrice $q(A) := p(A) - \lambda I$ non è invertibile. Siano $\{z_1, \dots, z_r\}$ le radici complesse del polinomio q , che quindi si fattorizza in $q(z) = a(z - z_1) \cdots (z - z_r)$. Avremo che

$$q(A) = a(A - z_1 I) \cdots (A - z_r I)$$

Visto che la composizione di operatori invertibili è invertibile e che $q(A)$ non è invertibile, deve esistere $j \in \{1, \dots, r\}$ tale che $A - z_j I$ non sia invertibile, ovvero $z_j \in \sigma(A)$. Infine $0 = q(z_j) = p(z_j) - \lambda$,cioè $\lambda = p(z_j)$, ossia $\lambda \in p(\sigma(A))$. □

Abbiamo tutti gli strumenti necessari per provare che \mathbb{C}^n si decompone nella somma di autospazi generalizzati.

Teorema 7.2.6. *Sia $A \in M_n(\mathbb{C})$ con spettro $\sigma(A) = \{\lambda_1, \dots, \lambda_r\}$ e sia E_{λ_k} l'autospazio generalizzato corrispondente a λ_k per $k = 1, \dots, r$. Allora*

$$\mathbb{C}^n = \bigoplus_{k=1}^r E_{\lambda_k}.$$

Dimostrazione. Sia p_A il polinomio caratteristico di A . Allora

$$p_A(\lambda) = \prod_{k=1}^r (\lambda - \lambda_k)^{m_k}$$

con $m_k = \text{ma}(\lambda_k)$. Introduciamo i polinomi

$$p_k(\lambda) = \prod_{j \neq k} (\lambda - \lambda_j)^{m_j} \quad q(\lambda) = \sum_{k=1}^r p_k(\lambda).$$

Chiaramente $p_k(\lambda_k) \neq 0$ e $p_k(\lambda_j) = 0$ per $j \neq k$ e quindi $q(\lambda) \neq 0$ per ogni $\lambda \in \sigma(A)$. Useremo questi polinomi per costruire le proiezioni di A sugli autospazi generalizzati.

Passo 1. Indichiamo con $E_k = E_{\lambda_k}$, $A_k = A|_{E_k}$ e infine $I = I_{E_k}$ e dimostriamo che

$$(A_k - \lambda_k I)^{m_k} = 0.$$

Per il Teorema di Cayley-Hamilton [7.2.4] abbiamo che

$$0 = p_A(A) = p_k(A)(A - \lambda_k I)^{m_k}.$$

Per la Proposizione [7.2.2] ii) E_k è un sottospazio invariante per A , per cui

$$0 = p(A)|_{E_k} = p(A_k)$$

e, di conseguenza,

$$p_k(A_k)(A_k - \lambda_k I)^{m_k} = 0. \tag{7.11}$$

D'altra parte, il Teorema della mappa spettrale [7.2.5] ci garantisce che

$$\sigma(p_k(A_k)) = p_k(\sigma(A_k)) = \{p_k(\lambda_k)\}$$

dove l'ultima uguaglianza è dovuta al fatto che $\sigma(A_k) = \{\lambda_k\}$, sempre per la Proposizione [7.2.2] iii). Per definizione di p_k si ha che $p_k(\lambda_k) \neq 0$ e allora

$0 \notin \sigma(p_k(A_k))$. Quindi $p_k(A_k)$ è una matrice invertibile e la tesi segue dalla (7.11).

Passo 2. Costruzione delle proiezioni. Osserviamo che $q(A)$ è una matrice invertibile, visto che $q(\lambda_j) \neq 0$ per ogni $\lambda_j \in \sigma(A)$ e quindi, per il Teorema della mappa spettrale, $0 \notin q(\sigma(A)) = \sigma(q(A))$.

Inoltre $q(A)E_k = E_k$ per ogni $k = 1, \dots, r$. Infatti, essendo E_k invariante per A , lo è anche per $q(A)$, che è un polinomio in A . Quindi $q(A): E_k \rightarrow E_k$ ed essendo invertibile, $\dim q(A)E_k = \dim E_k$, cioè $q(A)E_k = E_k$.

Introduciamo le matrici

$$P_k := (q(A))^{-1}p_k(A),$$

che risultano essere operatori di proiezione sugli E_k , dal momento che valgono le seguenti proprietà:

- i) $\sum_{k=1}^r P_k = I$;
- ii) $P_k|_{E_j} = 0$ se $j \neq k$;
- iii) $P_k \mathbb{C}^n = E_k$ e $P_k = I|_{E_k}$.

La i) segue dalla definizione:

$$\sum_{k=1}^r P_k = (q(A))^{-1} \sum_{k=1}^r p_k(A) = (q(A))^{-1}q(A) = I.$$

La ii) è conseguenza del fatto che $P_k|_{E_j} = (q(A))^{-1}p_k(A)|_{E_j}$ e

$$p_k(A)|_{E_j} = \prod_{i \neq k} (A - \lambda_i I)^{m_i}|_{E_j} = 0$$

dove, nell'ultima uguaglianza, abbiamo usato il risultato del primo passo.

Dimostriamo iii). Sia $w = P_k v$ per qualche $v \in \mathbb{C}^n$. Allora

$$w = P_k v = (q(A))^{-1}p_k(A)v \implies q(A)w = p_k(A)v$$

e allora

$$(A - \lambda_k I)^{m_k} q(A)w = (A - \lambda_k I)^{m_k} p_k(A)v = p(A)v = 0.$$

Quindi $q(A)w \in E_k$ o equivalentemente, $w \in (q(A))^{-1}E_k = E_k$.

D'altra parte, se $v \in E_k$, ricordando che $p_k(A)|_{E_j} = 0$ se $k \neq j$,

$$p_k(A)v = \sum_{j=1}^r p_j(A)v = q(A)v \implies P_kv = v$$

che prova contemporaneamente $E_k \subseteq P\mathbb{C}^n$ e $P_k = I_{E_k}$.

Passo 3. Sia $v \in \mathbb{C}^n$ e siano $v_k = P_kv$. Allora, per la proprietà ii), $v_k \in E_k$ e

$$v = Iv = \sum_{k=1}^r P_kv = \sum_{k=1}^r v_k. \quad (7.12)$$

Per concludere che $\mathbb{C}^n = E_1 \oplus \dots \oplus E_r$ resta da verificare che la decomposizione (7.12) è unica. Supponiamo che $v = w_1 + \dots + w_r$, con $w_k \in E_k$ sia un'altra decomposizione. Allora

$$v_k = P_kv = P_k(w_1 + \dots + w_n) = P_k(w_k) = w_k$$

essendo P_k nullo sugli E_j con $j \neq k$ e $P_k|_{E_k} = I_{E_k}$. \square

Siamo così arrivati alla decomposizione di \mathbb{C}^n in sottospazi A invarianti. In una nuova base, ottenuta dall'unione delle basi di tali sottospazi, la matrice A si trasforma in una matrice diagonale a blocchi $\text{diag}\{A_1, \dots, A_r\}$ dove ogni blocco A_k rappresenta l'operatore $A|_{E_k}$.

Osserviamo inoltre che gli operatori $N_k = A_k - \lambda_k I_{E_k}$ sono nilpotenti con $N_k^{d(\lambda_k)} = 0$, per la Proposizione 7.2.2 i).

Proposizione 7.2.7 (Significato geometrico della molteplicità algebrica). *Sia $A \in M_n(\mathbb{C})$, $\sigma(A) = \{\lambda_1, \dots, \lambda_r\}$, $m_k = \text{ma}(\lambda_k)$ ed E_k autosazio generalizzato associato a λ_k , $k = 1, \dots, r$. Allora $m_k = \dim E_k$.*

Dimostrazione. Abbiamo osservato che $N_k = A_k - \lambda_k I_{E_k}$ è un operatore nilpotente, quindi $\sigma(N_k) = \{0\}$, ossia $\sigma(A_k) = \{\lambda_k\}$. Poiché $A_k: E_k \rightarrow E_k$ ha un solo autovalore, per il Teorema fondamentale dell'algebra, questo deve avere molteplicità algebrica $n_k = \dim E_k$, cioè $p_{A_k}(\lambda) = (\lambda - \lambda_k)^{n_k}$.

In ogni sottospazio E_k , in accordo con il Lemma 7.2.3, possiamo scegliere una base in cui il blocco A_k sia triangolare superiore e calcolare il polinomio caratteristico in questa base, ottenendo

$$p_A(\lambda) = \prod_{k=1}^r p_{A_k}(\lambda) = \prod_{k=1}^r (\lambda - \lambda_k)^{n_k}$$

per quanto detto sugli autovalori dei singoli blocchi A_k . Ne concludiamo che $\text{ma}(\lambda_k) = n_k$. \square

In conclusione, abbiamo ottenuto il seguente risultato.

Corollario 7.2.8. *Sia $A \in M_n(\mathbb{C})$, $\sigma(A) = \{\lambda_1, \dots, \lambda_r\}$, ed E_k autospazio generalizzato associato a λ_k per $k = 1, \dots, r$. Allora esiste un cambiamento di base $C \in M_n(\mathbb{C})$ tale che*

$$A = C(D + N)C^{-1}$$

dove D è la matrice diagonale $\text{diag}\{\lambda_1 I_{E_1}, \dots, \lambda_r I_{E_r}\}$ e N è una matrice diagonale a blocchi nilpotente $\text{diag}\{N_1, \dots, N_r\}$ con $N_k = A_k - \lambda_k I_{E_k}$. Inoltre, poiché $DN = ND$, la matrice esponenziale e^{At} ha la forma

$$e^{At} = C \left(e^{Dt} \sum_{k=0}^{n-1} \frac{N^k t^k}{k!} \right) C^{-1}. \quad (7.13)$$

La matrice e^{Dt} è nota (si vedano gli Esempi 6.5.2 e 6.5.4). L'ulteriore proprietà che abbiamo guadagnato scegliendo come nuova base di \mathbb{C}^n una di quelle associate alla decomposizione di \mathbb{C}^n in autospazi generalizzati è che il resto, rispetto alla matrice diagonale D , sia un operatore nilpotente, di cui dobbiamo calcolare solo un numero finito di iterate. Per fare questo, è opportuno selezionare in maniera ancora più raffinata le basi dei singoli autospazi.

Un esempio di matrice nilpotente di cui si possono calcolare le iterate in maniera elementare è il cosiddetto blocco di Jordan.

Definizione 7.2.9. Un blocco di Jordan $k \times k$ è una matrice $J_k \in M_k(\mathbb{R})$ della forma

$$J_k = \begin{pmatrix} 0 & 1 & & & 0 & \\ & 0 & 1 & & & \\ & & 0 & \ddots & & \\ & & & \ddots & 1 & \\ & & & & 0 & 1 \\ 0 & & & & & 0 \end{pmatrix}.$$

Abbiamo che

$$J_k^2 = \begin{pmatrix} 0 & 0 & 1 & & 0 \\ & 0 & 0 & 1 & \\ & & 0 & \ddots & \ddots \\ & & & \ddots & 0 & 1 \\ 0 & & & & 0 & 0 \\ & & & & & 0 \end{pmatrix}$$

e in maniera analoga si trova la forma esplicita delle altre iterate. Quindi la matrice J_k è nilpotente con indice di nilpotenza $k - 1$.

Il seguente teorema, che non dimostreremo, ci assicura che ogni matrice nilpotente si trasforma, a seguito di un opportuno cambiamento di base, in una matrice a blocchi di Jordan (per la dimostrazione rimandiamo, ad esempio, al testo [8]).

Teorema 7.2.10. *Sia W un sottospazio di dimensione k di \mathbb{C}^n e sia N un operatore lineare nilpotente su W . Allora esiste un cambiamento di coordinate C in W tale che $N = C^{-1}JC$ e $J = \text{diag}\{J_{k_1}, \dots, J_{k_r}, Z\}$ dove J_{k_s} sono blocchi di Jordan e Z è un blocco nullo (eventualmente di ordine 0).*

Applicando il teorema di decomposizione di Jordan per operatori nilpotenti ad ogni singolo autospazio generalizzato determinato da una matrice $A \in \mathbb{C}^n$, otteniamo una versione più raffinata del Corollario 7.2.8.

Teorema 7.2.11. *Per ogni $A \in M_n(\mathbb{C})$ esiste un cambiamento di base $C \in M_n(\mathbb{C})$ tale che CAC^{-1} è una matrice diagonale a blocchi, dove i blocchi sono della forma*

$$\begin{pmatrix} \lambda & 1 & & & 0 & & \\ & \lambda & 1 & & & & \\ & & \lambda & \ddots & & & \\ & & & \ddots & 1 & & \\ 0 & & & & \lambda & 1 & \\ & & & & & \lambda & \end{pmatrix} \quad (7.14)$$

con $\lambda \in \sigma(A)$. I blocchi 1×1 sono formati solo dall'autovalore λ .

Dimostrazione. Siano $\lambda_1, \dots, \lambda_r$ gli autovalori di A contati con molteplicità. Per il Teorema 7.2.6, \mathbb{C}^n si decompone nella somma diretta degli autospazi generalizzati

$$\mathbb{C}^n = \bigoplus_{k=1}^r E_k,$$

dove $E_k = E_{\lambda_k}$. In qualsiasi base degli E_k la matrice A sarà diagonale a blocchi $\text{diag}\{A_1, \dots, A_r\}$ dove $A_k = A|_{E_k}$ e gli operatori $N_k = A_k - \lambda_k I_{E_k}$ sono nilpotenti. In particolare, per il Teorema 7.2.10, esiste una base di E_k in cui la matrice di $\lambda_k I_{E_k} + N_k$ è, a sua volta, diagonale a blocchi con blocchi della forma (7.14). \square

7.3 Rappresentazione esplicita delle soluzioni

Sia ora $A \in M_n(\mathbb{R})$. Se, come abbiamo fatto nel Paragrafo 7.1, la pensiamo come elemento di $M_n(\mathbb{C})$, i risultati della sezione precedente ci permettono di determinarne esplicitamente la matrice esponenziale in tale spazio e, conseguentemente, le soluzioni del sistema lineare omogeneo $x' = Ax$ in \mathbb{C}^n .

Teorema 7.3.1. *Per ogni $A \in M_n(\mathbb{R})$ esiste un cambiamento di base $C \in M_n(\mathbb{C})$ tale che $e^{At} = C^{-1}e^{Bt}C$ con e^{Bt} è una matrice diagonale a blocchi, dove i blocchi sono della forma*

$$e^{\lambda t} \begin{pmatrix} 1 & t & t^2/2 & \dots & \dots & t^k/k! \\ 0 & 1 & t & \dots & \dots & t^{k-1}/(k-1)! \\ \vdots & 0 & 1 & t & \dots & t^{k-2}/(k-2)! \\ \vdots & \vdots & 0 & \ddots & t & \vdots \\ \vdots & \vdots & \vdots & 0 & 1 & t \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (7.15)$$

con $\lambda \in \sigma(A)$, e $k \leq \text{ma}(\lambda)$. I blocchi 1×1 hanno come unico elemento $e^{\lambda t}$.

Di conseguenza, se $\sigma(A) = \{\lambda_1, \dots, \lambda_r\}$ e $m_j = \text{ma } \lambda_j$, per ogni $j = 1, \dots, r$ esistono esattamente m_j soluzioni linearmente indipendenti del sistema lineare omogeneo $x' = Ax$ in \mathbb{C}^n della forma

$$u_{j,s}(t) = e^{\lambda_j t} p_{j,s-1}(t), \quad t \in \mathbb{R}, \quad 1 \leq s \leq m_j,$$

dove $p_{j,s-1}(t)$ è un polinomio di grado minore o uguale ad $s - 1$. L'insieme di queste funzioni forma un sistema fondamentale di soluzioni per $x' = Ax$ in \mathbb{C}^n .

Dimostrazione. Basta ricordare la (7.13), rappresentare la matrice A nella base di \mathbb{C}^n determinata nel Teorema 7.2.11 e osservare che, se J è un blocco di Jordan di ordine k , allora $\sum_{k=0}^{n-1} \frac{J^k t^k}{k!}$ coincide con la matrice in (7.15). \square

A questo punto siamo in grado di descrivere completamente le soluzioni del sistema $x' = Ax$ in \mathbb{R}^n per ogni $A \in M_n(\mathbb{R})$.

Corollario 7.3.2. *Sia $A \in M_n(\mathbb{R})$, siano $\lambda_j, j = 1, \dots, h$, i suoi autovalori reali e $\mu_j = \alpha_j + i\beta_j$, $\bar{\mu_j} = \alpha_j - i\beta_j$, $j = 1, \dots, k$, i suoi autovalori complessi, $h + 2k = n$. Allora ogni soluzione del sistema lineare omogeneo di equazioni differenziali $x' = Ax$ ha per componenti una combinazione lineare a coefficienti reali delle funzioni*

$$\begin{aligned} & e^{\lambda_j t}, \quad te^{\lambda_j t}, \dots, \quad t^{\text{ma}(\lambda_j)-1} e^{\lambda_j t}, \quad j = 1, \dots, h, \\ & e^{\alpha_j t} \cos(\beta_j t), \quad te^{\alpha_j t} \cos(\beta_j t), \dots, \quad t^{\text{ma}(\mu_j)-1} e^{\alpha_j t} \cos(\beta_j t), \quad j = 1, \dots, k \\ & e^{\alpha_j t} \sin(\beta_j t), \quad te^{\alpha_j t} \sin(\beta_j t), \dots, \quad t^{\text{ma}(\mu_j)-1} e^{\alpha_j t} \sin(\beta_j t), \quad j = 1, \dots, k. \end{aligned}$$

Dimostrazione. Per il Teorema 7.3.1 ogni soluzione del sistema ha componenti che si ottengono tramite opportune combinazioni lineari a coefficienti complessi delle funzioni

$$\begin{aligned} & e^{\lambda_j t}, \quad te^{\lambda_j t}, \dots, \quad t^{m_j-1} e^{\lambda_j t}, \quad j = 1, \dots, h \\ & e^{\mu_j t}, \quad te^{\mu_j t}, \dots, \quad t^{m_j-1} e^{\mu_j t}, \quad j = 1, \dots, k \\ & e^{\bar{\mu}_j t}, \quad te^{\bar{\mu}_j t}, \dots, \quad t^{m_j-1} e^{\bar{\mu}_j t}, \quad j = 1, \dots, k. \end{aligned}$$

Ricordando l'identità di Eulero

$$e^{\mu_j t} = e^{\alpha_j t} (\cos(\beta_j t) + i \sin(\beta_j t))$$

ne segue che ogni soluzione del sistema ha componenti che si ottengono tramite opportune combinazioni lineari a coefficienti complessi delle funzioni elencate nell'enunciato.

Infine, se una funzione reale è combinazione lineare a valori complessi di funzioni reali, allora è sempre possibile determinare una combinazione lineare a coefficienti reali delle stesse funzioni che la determini (basta prendere la parte reale della combinazione lineare), il che conclude la dimostrazione. \square

Esempio 7.3.3. Determiniamo tutte le soluzioni del sistema lineare omogeneo autonomo

$$\begin{cases} x' = x + y, \\ y' = z, \\ z' = -x - y + 3. \end{cases}$$

La matrice dei coefficienti del sistema è

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & 3 \end{pmatrix}$$

il cui polinomio caratteristico è $p_A(\lambda) = -\lambda(\lambda^2 - 4\lambda + 4)$. Gli autovalori della matrice A sono quindi $\lambda_1 = 0$ con molteplicità algebrica $m_1 = 1$ e $\lambda_2 = 2$ con molteplicità algebrica $m_2 = 2$. Gli autovettori corrispondenti a $\lambda_1 = 0$ devono soddisfare $Au = 0$, ossia, se $u = (u_1, u_2, u_3)$, deve valere

$$\begin{cases} u_1 + u_2 = 0 \\ u_3 = 0 \\ -u_1 - u_2 + 3u_3 = 0 \end{cases}$$

quindi l'autospazio associato a $\lambda_1 = 0$ è generato, ad esempio, dal vettore $u = (1, -1, 0)$.

D'altra parte, Gli autovettori corrispondenti a $\lambda_2 = 2$ devono soddisfare $Av = 2v$, ossia, se $v = (v_1, v_2, v_3)$, deve valere

$$\begin{cases} v_1 + v_2 = 2v_1 \\ v_3 = 2v_2 \\ -v_1 - v_2 + 3v_3 = 2v_3 \end{cases}$$

da cui si ricava che l'autospazio associato a $\lambda_2 = 2$ è generato, ad esempio, dal vettore $v = (1, 1, 2)$. In particolare l'autovalore $\lambda_2 = 2$ non è semisemplice, quindi la matrice A non è diagonalizzabile. Quindi per determinare

esplicitamente la matrice esponenziale di A , e di conseguenza un sistema fondamentale di soluzioni per il sistema, dobbiamo determinare un terzo vettore $w = (w_1, w_2, w_3)$ in modo tale che, nella base $\{u, v, w\}$ l'azione dell'operatore associato ad A sia rappresentata da una matrice in forma canonica di Jordan

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Poiché, per definizione, $B = (Au|Av|Aw)$, il vettore w deve soddisfare $Aw = v + 2w$, ossia

$$\begin{cases} w_1 + w_2 = 2w_1 + 1 \\ w_3 = 2w_2 + 1 \\ -w_1 - w_2 + 3w_3 = 2w_3 + 2 \end{cases}$$

da cui si ricava che w può essere scelto della forma $w_2v + (-1, 0, 1)$. Fissiamo $w = (-1, 0, 1)$ in modo tale che la matrice di cambiamento di base sia

$$C = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \quad C^{-1} = \frac{1}{4} \begin{pmatrix} 1 & -3 & 1 \\ 1 & 1 & 1 \\ -2 & -2 & 2 \end{pmatrix}.$$

La matrice esponenziale di B è

$$e^{Bt} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{2t} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix}$$

quindi la matrice fondamentale speciale del sistema è data da

$$e^{At} = \frac{1}{4} \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix} \begin{pmatrix} 1 & -3 & 1 \\ 1 & 1 & 1 \\ -2 & -2 & 2 \end{pmatrix}.$$

A conti fatti, otteniamo

$$e^{At} = \frac{1}{4} \begin{pmatrix} (-2e^{2t}t + 3e^{2t} + 1) & (-2e^{2t}t + 3e^{2t} - 3) & (2e^{2t}t - e^{2t} + 1) \\ (-2e^{2t}t + e^{2t} - 1) & (-2e^{2t}t + e^{2t} + 3) & (2e^{2t}t + e^{2t} - 1) \\ -4e^{2t}t & -4e^{2t}t & 4e^{2t}(t + 1) \end{pmatrix}$$

e quindi per ogni $\xi = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^n$ la soluzione $u(t)$ del problema di Cauchy del sistema considerato con dato iniziale $u(0) = \xi$ è data da

$$u(t) = \frac{1}{4} \begin{pmatrix} (e^{2t}(3 - 2t) + 1)\xi_1 + (e^{2t}(3 - 2t) - 3)\xi_2 + (e^{2t}(2t - 1) + 1)\xi_3 \\ (e^{2t}(1 - 2t) - 1)\xi_1 + (e^{2t}(1 - 2t) + 3)\xi_2 + (e^{2t}(2t + 1) - 1)\xi_3 \\ 4e^{2t}(\xi_3 - t(\xi_1 + \xi_2 - \xi_3)) \end{pmatrix}$$

Capitolo 8

Analisi qualitativa dei sistemi lineari omogenei autonomi

8.1 Proprietà generali delle soluzioni

Avere la rappresentazione esplicita delle soluzioni di un sistema lineare autonomo $x' = Ax$ permette di fare una classificazione precisa dei possibili comportamenti asintotici delle traiettorie. Per quanto mostrato nel capitolo precedente, risulta evidente che le proprietà delle soluzioni dipenderanno dalle proprietà dello spettro della matrice A .

Iniziamo col dare una condizione necessaria e sufficiente affinché l'origine (sempre soluzione stazionaria del sistema) attragga asintoticamente tutte le altre soluzioni del sistema. Premettiamo il seguente risultato tecnico che risulta essere interessante di per sé.

Lemma 8.1.1. *Sia $A \in M_n(\mathbb{R})$. Le seguenti proprietà sono equivalenti.*

- i) $\lim_{t \rightarrow +\infty} u(t) = 0$ per ogni soluzione u del sistema $x' = Ax$ in \mathbb{R}^n .
- ii) $\lim_{t \rightarrow +\infty} e^{At} = 0$ in $\mathcal{L}(\mathbb{R}^n)$;
- iii) $\lim_{t \rightarrow +\infty} (e^{At})^{\mathbb{C}} = 0$ in $\mathcal{L}(\mathbb{C}^n)$;
- iv) $\lim_{t \rightarrow +\infty} u(t) = 0$ per ogni soluzione u del sistema $x' = Ax$ in \mathbb{C}^n .

Dimostrazione. Iniziamo col dimostrare che i) equivale a ii). In maniera del tutto analoga si dimostra che iii) equivale a iv). Poiché

$$\|e^{At}y\| \leq \|e^{At}\|_{\mathcal{L}} \|y\|, \quad y \in \mathbb{R}^n,$$

la convergenza dell'operatore a zero in $\mathcal{L}(\mathbb{R}^n)$ implica la convergenza puntuale a zero. A questo punto, ricordando che le soluzioni del sistema $x' = Ax$ in \mathbb{R}^n sono tutte e sole le funzioni della forma $e^{At}y$, $y \in \mathbb{R}^n$ (si veda il Teorema 6.5.1), si ottiene che ii) implica i).

Viceversa, se vale i) e $\{e_1, \dots, e_n\}$ è la base canonica per \mathbb{R}^n , per ogni $\epsilon > 0$ e per ogni $i = 1, \dots, n$ è possibile determinare $t_i > 0$ tale che $\|e^{At_i}e_i\| < \epsilon$ per ogni $t > t_i$. Scelto $t_0 = \max_i t_i$ avremo che per ogni $y \in \mathbb{R}^n$, $y = \sum_{i=1}^n \alpha_i e_i$

$$\|e^{At}y\| = \left\| \sum_{i=1}^n \alpha_i e^{At}e_i \right\| \leq \epsilon \|y\|$$

e, di conseguenza, la convergenza dell'operatore.

Per dimostrare che ii) è equivalente a iii), innanzi tutto si deve ricordare che la matrice A , oltre a rappresentare un operatore lineare in $\mathcal{L}(\mathbb{R}^n)$ rappresenta anche il corrispondente operatore complessificato in $\mathcal{L}(\mathbb{C}^n)$. Inoltre si ha che per il complessificato dell'esponenziale di A

$$(e^{At})^{\mathbb{C}}(x + iy) = e^{At}x + ie^{At}y,$$

vale

$$\begin{aligned} \max\{\|e^{At}x\|^2, \|e^{At}y\|^2\} &\leq \|e^{At}x + ie^{At}y\|^2 = \|e^{At}x\|^2 + \|e^{At}y\|^2 \\ &\leq 2 \max\{\|e^{At}x\|^2, \|e^{At}y\|^2\}, \end{aligned}$$

per ogni $x, y \in \mathbb{R}^n$, il che comporta l'equivalenza tra ii) e iii). \square

Siamo ora in grado di determinare la classe di sistemi lineari per i quali tutte le traiettorie tendono all'equilibrio per t che tende a $+\infty$.

Teorema 8.1.2 (Criterio di stabilità). *Sia $A \in M_n(\mathbb{R})$. Le seguenti proprietà sono equivalenti.*

- i) $\operatorname{Re}(\lambda) < 0$ per ogni $\lambda \in \sigma(A)$;
- ii) $\lim_{t \rightarrow +\infty} u(t) = 0$ per ogni soluzione u del sistema $x' = Ax$ in \mathbb{R}^n .

Dimostrazione. Per il Lemma 8.1.1, possiamo pensare la matrice A in $M_n(\mathbb{C})$ e considerare le soluzioni del sistema $x' = Ax$ in \mathbb{C}^n che sappiamo essere combinazioni lineari di funzioni della forma $t^m e^{\lambda_i t} y$, con $y \in \mathbb{C}^n$. Poiché

$$\left\| t^m e^{\lambda_i t} y \right\| \leq |t|^m e^{\operatorname{Re} \lambda_i t} \|y\|$$

se, per ipotesi, $\operatorname{Re} \lambda_i < 0$, ogni elemento del sistema fondamentale di soluzioni, e quindi ogni soluzione del sistema, tende a zero per t che diverge a $+\infty$.

Resta da dimostrare che se esiste $\lambda \in \sigma(A)$ con $\operatorname{Re} \lambda \geq 0$, allora è sempre possibile costruire almeno una soluzione del sistema di equazioni differenziali in \mathbb{C}^n che non tenda a zero per t che tende a $+\infty$. Per fare questo, basta prendere un autovettore $v \in \mathbb{C}^n$, $v \neq 0$, associato a λ e la funzione $u(t) = e^{\lambda t} v$. Poiché

$$u'(t) = \lambda e^{\lambda t} v = A e^{\lambda t} v$$

la funzione u è soluzione del sistema in \mathbb{C}^n ed inoltre

$$\begin{cases} \lim_{t \rightarrow +\infty} \|u(t)\| = \|v\| \lim_{t \rightarrow +\infty} e^{\operatorname{Re} \lambda t} = +\infty, & \text{se } \operatorname{Re} \lambda > 0, \\ \|u(t)\| = \|v\| & \text{se } \operatorname{Re} \lambda = 0. \end{cases}$$

□

Come conseguenza del risultato precedente otteniamo anche informazioni precise sul comportamento asintotico delle soluzioni per t che tende a $-\infty$.

Corollario 8.1.3. *Sia $A \in M_n(\mathbb{R})$. Le seguenti proprietà sono equivalenti.*

- i) $\operatorname{Re}(\lambda) < 0$ per ogni $\lambda \in \sigma(A)$;
- ii) $\lim_{t \rightarrow -\infty} \|u(t)\| = +\infty$ per ogni soluzione $u \neq 0$ del sistema $x' = Ax$ in \mathbb{R}^n .

Dimostrazione. Sia u soluzione di $x' = Ax$ in \mathbb{R}^n e sia $y \in \mathbb{R}^n$ tale che $u(t) = e^{At} y$. Grazie alla Proposizione 6.4.2 sappiamo che la matrice e^{At} è sempre invertibile e la sua inversa è e^{-At} . Quindi otteniamo

$$\|y\| = \|e^{-At} u(t)\| \leq \|e^{-At}\|_{\mathcal{L}} \|u(t)\| \Rightarrow \|u(t)\| \geq \frac{\|y\|}{\|e^{-At}\|_{\mathcal{L}}}$$

da cui risulta evidente, grazie al Lemma 8.1.1 e al Teorema 8.1.2, che i) implica ii).

Se invece esiste $\lambda \in \sigma(A)$ con $\operatorname{Re} \lambda \geq 0$, abbiamo visto nella dimostrazione del Teorema 8.1.2 che è possibile costruire una soluzione del sistema in \mathbb{C}^n della forma $u(t) = e^{\lambda t}v$, $v \in \mathbb{C}^n$, $v \neq 0$ tale che $\|u(t)\| \leq c$ per ogni $t < 0$. Se denotiamo $v = x + iy$, avremo che

$$\max\{\|e^{At}x\|^2, \|e^{At}y\|^2\} \leq \|e^{At}x + ie^{At}y\|^2 = \|e^{\lambda t}v\|^2 \leq c^2, \quad \forall t < 0$$

quindi $e^{At}x$ e $e^{At}y$ sono soluzioni non simultaneamente nulle del sistema in \mathbb{R}^n che non verificano ii). \square

Un'altra immediata conseguenza del Teorema 8.1.2 è il seguente risultato di instabilità che caratterizza i sistemi lineari per i quali tutte le soluzioni esplodono per t che tende a $+\infty$.

Corollario 8.1.4. *Sia $A \in M_n(\mathbb{R})$. Le seguenti proprietà sono equivalenti.*

- i) $\operatorname{Re}(\lambda) > 0$ per ogni $\lambda \in \sigma(A)$;
- ii) $\lim_{t \rightarrow +\infty} \|u(t)\| = +\infty$ per ogni soluzione $u \neq 0$ del sistema $x' = Ax$ in \mathbb{R}^n .

Dimostrazione. Basta osservare che $\lambda \in \sigma(A)$ se e solo se $-\lambda \in \sigma(-A)$ e che $u(t)$ è soluzione di $x' = Ax$ se e solo se $u(-t)$ è soluzione di $x' = -Ax$. A questo punto l'equivalenza segue dal Corollario 8.1.3. \square

I risultati precedenti mostrano che tutte le soluzioni del sistema lineare $x' = Ax$ hanno un preciso comportamento asintotico per t che tende a $+\infty$ se e solo se la parte reale di tutti gli autovalori della matrice A ha un segno fissato. D'altra parte, grazie alla teoria sviluppata nel Capitolo 7, noi sappiamo che lo spazio \mathbb{R}^n può essere decomposto in somma diretta di sottospazi invarianti per A e tali che la restrizione di A ad uno di tali sottospazi ha come unico autovalore un preciso elemento di $\sigma(A)$. Possiamo quindi immaginare che il comportamento asintotico delle proiezioni delle soluzioni su ciascun sottospazio dipenda solo dal segno della parte reale del relativo autovalore. Questo è effettivamente quello che succede, come mostrato dal seguente importante risultato.

Teorema 8.1.5 (Varietà stabile e instabile). *Sia $A \in M_n(\mathbb{R})$ tale che $\operatorname{Re} \lambda \neq 0$ per ogni $\lambda \in \sigma(A)$. Allora esiste una scomposizione di \mathbb{R}^n come somma diretta $\mathbb{R}^n = E^S \oplus E^I$ tale che per ogni soluzione u del sistema $x' = Ax$ le sue proiezioni u_S e u_I rispettivamente su E^S e E^I hanno le proprietà*

$$\lim_{t \rightarrow +\infty} \|u_S(t)\| = 0, \quad \lim_{t \rightarrow +\infty} \|u_I(t)\| = +\infty.$$

Dimostrazione. Nel processo di riduzione della matrice $A \in M_n(\mathbb{R})$ in forma canonica di Jordan abbiamo decomposto lo spazio \mathbb{R}^n in somma diretta di sottospazi E_j invarianti per A e, conseguentemente, anche per e^{At} . Inoltre gli spazi E_j sono costruiti in modo tale che, se $\sigma(A) = \{\lambda_1, \dots, \lambda_k\}$, allora $\sigma(A|_{E_j}) = \{\lambda_j\}$, $j = 1, \dots, k$. Supponiamo di aver ordinato gli autovalori secondo l'ordine crescente della loro parte reale, in modo tale che sulla parte superiore della matrice compaiano tutti i blocchi relativi alla base del sottospazio invariante $E^S := \bigcup_{\operatorname{Re} \lambda_i < 0} E_i$ e nella parte inferiore quelli relativi alla base del sottospazio invariante $E^I := \bigcup_{\operatorname{Re} \lambda_i > 0} E_i$.

Fissato un dato iniziale $x(0) = x_0$ per il problema di Cauchy relativo al sistema lineare $x' = Ax$, scriviamolo come $x_0 = x_1 + x_2$, $x_1 \in E^S$, $x_2 \in E^I$. Vista l'invarianza dei sottospazi sia rispetto ad A che rispetto ad e^{At} , la soluzione $u(t)$ del problema di Cauchy potrà essere decomposta a sua volta nella somma $u(t) = u_S(t) + u_I(t)$ con $u_S(t) \in E^S$ $u_I(t) \in E^I$ per ogni $t \in \mathbb{R}$ e avremo

$$u'_S = A|_{E^S} u_S, \quad u'_I = A|_{E^I} u_I$$

A questo punto la tesi si ottiene osservando che $\operatorname{Re} \lambda < 0$ per ogni $\lambda \in \sigma(A|_{E^S})$, mentre $\operatorname{Re} \lambda > 0$ per ogni $\lambda \in \sigma(A|_{E^I})$ ed utilizzando il Teorema 8.1.2 ed il Corollario 8.1.3. \square

Concludiamo questo paragrafo con una condizione necessaria e sufficiente affinché tutte le soluzioni di un sistema lineare siano limitate.

Teorema 8.1.6 (Criterio di limitatezza). *Ogni soluzione del sistema lineare $x' = Ax$, $A \in M_n(\mathbb{R})$, è limitata per t che diverge a $+\infty$ se e solo se valgono le seguenti proprietà:*

- i) $\operatorname{Re} \lambda \leq 0$ per ogni $\lambda \in \sigma(A)$;
- ii) ogni autovalore $\lambda \in \sigma(A)$ con $\operatorname{Re} \lambda = 0$ è semisemplice.

Dimostrazione. Dal momento che le soluzioni in \mathbb{R}^n di $x' = Ax$ possono essere pensate come soluzioni dello stesso sistema in \mathbb{C}^n corrispondenti a dati iniziali reali, basta dimostrare il risultato nel caso complesso. In questo caso sappiamo che le soluzioni sono combinazioni lineari di funzioni della forma $t^k e^{\lambda t} y$, $\lambda \in \sigma(A)$, $y \in \mathbb{C}^n$. Chiaramente, se $\operatorname{Re} \lambda < 0$, tali funzioni tendono a zero per t che diverge a $+\infty$. D'altra parte, se $\operatorname{Re} \lambda = 0$ e l'autovalore λ è semisemplice, allora il blocco della matrice A relativo a λ è diagonalizzabile e i corrispondenti elementi della famiglia fondamentale di soluzioni del sistema sono del tipo $e^{i\beta t} y$, $\beta = \operatorname{Im} \lambda$, $y \in \mathbb{C}^n$, e risultano quindi limitati per ogni $t \in \mathbb{R}$. Se invece $\operatorname{Re} \lambda = 0$ e λ non è semisemplice, sicuramente esiste almeno una soluzione della forma $t^k e^{i\beta t} y$, $k > 0$, $y \in \mathbb{C}^n$, che non è limitata per t che diverge a $+\infty$. \square

8.2 Studio qualitativo dei sistemi planari

In questo paragrafo descriveremo tutte le soluzioni dei sistemi lineari omogenei a coefficienti costanti bidimensionali

$$\begin{cases} x' = ax + by \\ y' = cx + dy \end{cases} \quad (8.1)$$

Osserviamo che le soluzioni di questi sistemi sono invarianti per traslazioni temporali (se $(x(t), y(t))$ è soluzione di (8.1), allora per ogni $\tau \in \mathbb{R}$ la funzione $(x(t + \tau), y(t + \tau))$ è ancora soluzione), quindi sono completamente descritte dalla loro proiezione sul piano (x, y) (descrizione nel piano delle fasi o ritratto di fase).

Nel resto del paragrafo indicheremo con B la matrice in forma canonica di Jordan associata alla matrice dei coefficienti A e descriveremo le traiettorie associate a B , che di fatto hanno le stesse proprietà qualitative di quelle associate ad A . Infatti, se C determina il cambiamento di variabili che trasforma A in B , ossia $B = C^{-1}AC$, allora si ha $u' = Au$ se e solo se $v = Cu$ è soluzione di $v' = Bv$.

Indicheremo con λ e μ i due autovalori in \mathbb{C} di A (o, equivalentemente, di B).

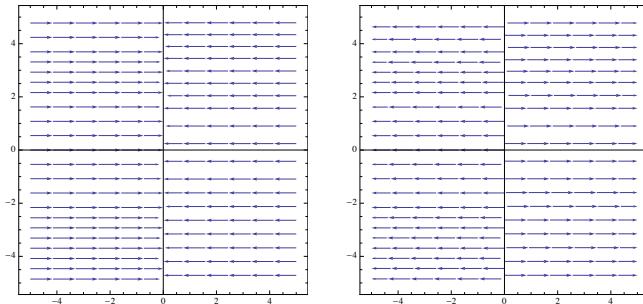


Figura 8.1: Caso degenere a traiettorie orizzontali

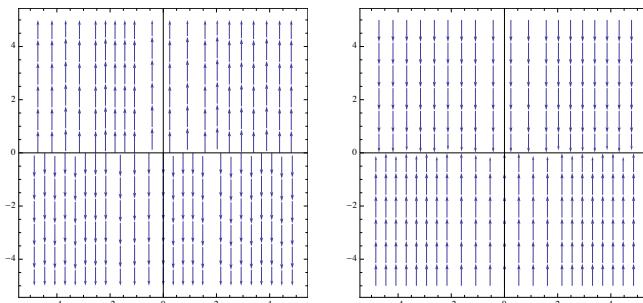


Figura 8.2: Caso degenere a traiettorie verticali

Caso degenere: $\lambda \in \mathbb{R}$, $\lambda \neq 0$, $\mu = 0$. La matrice B ha la forma

$$B = \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix} \quad \text{oppure} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & \lambda \end{pmatrix}.$$

Nel primo caso le traiettorie corrispondenti a dati iniziali del tipo $(x(0), y(0)) = (0, y_0)$, $y_0 \in \mathbb{R}$ sono costanti (soluzioni stazionarie o punti di equilibrio), mentre quelle corrispondenti a dati iniziali del tipo $(x(0), y(0)) = (x_0, y_0)$ con $x_0 \neq 0$ sono tutte orizzontali (la componente y è costante). Se $\lambda > 0$ gli equilibri sono repulsivi, nel senso che altre traiettorie si allontanano dagli equilibri nel tempo, mentre se $\lambda < 0$ gli equilibri sono attrattivi, ossia le altre traiettorie tendono ad avvicinarsi (si veda la Figura 8.1). In maniera del tutto analoga, nel secondo caso abbiamo soluzioni stazionarie lungo tutto l'asse delle ascisse e tutte le altre traiettorie verticali mentre nel secondo caso sono tutte verticali (si veda la Figura 8.2).

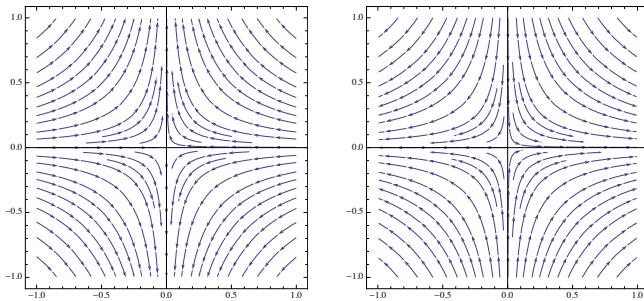


Figura 8.3: Sella corrispondente a $\lambda < 0$ (a sinistra) e a $\mu < 0$ (a destra)

Sella: $\lambda, \mu \in \mathbb{R}, \lambda\mu < 0$. In questo caso le matrici B ed e^{Bt} hanno la forma

$$B = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \quad e^{Bt} = \begin{pmatrix} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{pmatrix}$$

e quindi le soluzioni sono tutte del tipo $v(t) = e^{Bt}c = (c_1e^{\lambda t}, c_2e^{\mu t})$, $c_1, c_2 \in \mathbb{R}$. Se $c_1 = c_2 = 0$ otteniamo l'unico punto di equilibrio $v(t) = (0, 0)$. Per $c_1 = 0$ e $c_2 \neq 0$, abbiamo traiettorie che si trovano sull'asse delle ordinate (che saranno uscenti dall'origine se $\mu > 0$ ed entranti se $\mu < 0$). Analogamente, se $c_1 \neq 0$ e $c_2 = 0$, avremo traiettorie sull'asse delle ascisse, entranti nell'origine se $\lambda < 0$ e uscenti se $\lambda > 0$. Infine, in corrispondenza di $c_1 \neq 0$ e $c_2 \neq 0$, le traiettorie si trovano su curve del tipo $y = c_2(x/c_1)^{\mu/\lambda}$, $\mu/\lambda < 0$. Quindi il ritratto di fase è quello in Figura 8.3.

Nodo: $\lambda, \mu \in \mathbb{R}, \lambda\mu > 0$. Dobbiamo distinguere tre casi:

- i) $\lambda \neq \mu$;
- ii) $\lambda = \mu$ e $m_g(\lambda) = 2$ (stella);
- iii) $\lambda = \mu$ e $m_g(\lambda) = 1$ (nodo improprio).

In tutti i tre casi $v(t) = (0, 0)$ è l'unica soluzione stazionaria, mentre la geometria delle altre traiettorie cambia sensibilmente.

Nel caso i), le matrici B e e^{Bt} hanno la forma

$$B = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \quad e^{Bt} = \begin{pmatrix} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{pmatrix}$$

e quindi le soluzioni sono tutte del tipo $v(t) = (c_1e^{\lambda t}, c_2e^{\mu t})$, $c_1, c_2 \in \mathbb{R}$. Se $c_1 = c_2 = 0$ otteniamo l'unico punto di equilibrio $v(t) = (0, 0)$. Per $c_1 = 0$ e

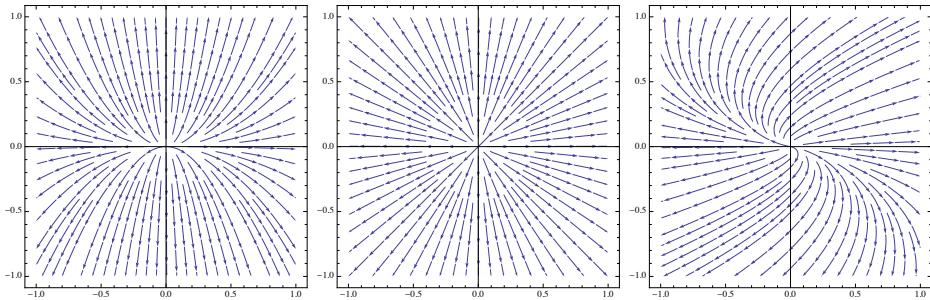


Figura 8.4: Nodo nei casi i) (a sinistra), ii) (al centro) e iii) (a destra)

$c_2 \neq 0$, abbiamo traiettorie che si trovano sull'asse delle ordinate (che saranno uscenti dall'origine se $\mu > 0$ ed entranti se $\mu < 0$.). Analogamente, se $c_1 \neq 0$ e $c_2 = 0$, avremo traiettorie sull'asse delle ascisse, entranti nell'origine se $\lambda < 0$ e uscenti se $\lambda > 0$. Osserviamo che in questo caso λ e μ hanno segno concorde, quindi le traiettorie sugli assi cartesiani sono o tutte entranti o tutte uscenti dall'origine. Infine, in corrispondenza di $c_1 \neq 0$ e $c_2 \neq 0$, le traiettorie si trovano su parabole del tipo $y = c_2(x/c_1)^{\mu/\lambda}$ e risultano entranti nell'origine se gli autovalori sono negativi o uscenti se gli autovalori sono positivi. Nel primo caso tutte le traiettorie tendono ad avvicinarsi nel tempo all'origine (nodo stabile), nel secondo tutte le traiettorie tendono ad allontanarsi nel tempo (nodo instabile). In Figura 8.4 a sinistra è riportato il ritratto di fase nel caso di autovalori positivi. Il ritratto di fase per autovalori negativi si ottiene semplicemente invertendo il verso di percorrenza delle traiettorie.

Nel caso ii), le matrici B ed e^{Bt} hanno la forma

$$B = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \quad e^{Bt} = \begin{pmatrix} e^{\lambda t} & 0 \\ 0 & e^{\lambda t} \end{pmatrix}$$

e quindi le soluzioni sono tutte del tipo $x(t) = (c_1, c_2)e^{\lambda t}$, $c_1, c_2 \in \mathbb{R}$, quindi giacciono sulle rette di equazione cartesiana $c_1y = c_2x$. Il ritratto di fase per $\lambda > 0$ (origine nodo instabile) è riportato in Figura 8.4 al centro. Il ritratto di fase per $\lambda < 0$ (origine nodo stabile) si ottiene semplicemente invertendo il verso di percorrenza delle traiettorie.

Nel caso iii), le matrici B ed e^{Bt} hanno la forma

$$B = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = \lambda I + N \quad e^{Bt} = e^{\lambda t}I e^{Nt} = \begin{pmatrix} e^{\lambda t} & 0 \\ 0 & e^{\lambda t} \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}$$

e quindi le soluzioni sono tutte del tipo $v(t) = e^{Bt}c = (c_1 + c_2t, c_2)e^{\lambda t}$, $c_1, c_2 \in \mathbb{R}$. Per $c_1 = c_2 = 0$ abbiamo l'equilibrio $v(t) = (0, 0)$. Se $c_2 = 0$,

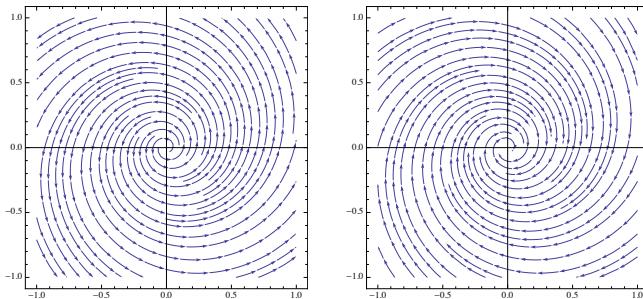


Figura 8.5: Fuoco instabile (a sinistra) e stabile (a destra)

troviamo le traiettorie $(c_1, 0)e^{\lambda t}$, che giacciono sull'asse delle ascisse e sono entranti nell'origine se $\lambda < 0$, uscenti in caso contrario. Infine, se $c_2 \neq 0$, le traiettorie giacciono sulle curve di equazione cartesiana

$$x = \frac{c_1}{c_2}y + \frac{y}{\lambda} \log \left(\frac{y}{c_2} \right)$$

e il ritratto di fase per $\lambda > 0$ (origine nodo instabile) è riportato in Figura 8.4 a destra. Come ormai è chiaro, il ritratto di fase per $\lambda < 0$ (origine nodo stabile) si ottiene semplicemente invertendo il verso di percorrenza delle traiettorie.

Fuoco: $\lambda = \alpha + i\beta$, $\mu = \alpha - i\beta$, $\alpha \neq 0$. In questo caso le matrici B ed e^{Bt} hanno la forma

$$B = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = \alpha I + \beta J \quad e^{Bt} = e^{\alpha t} e^{\beta t J} = e^{\alpha t} \begin{pmatrix} \cos(\beta t) & \sin(\beta t) \\ -\sin(\beta t) & \cos(\beta t) \end{pmatrix}$$

e quindi le soluzioni sono tutte del tipo

$$v(t) = e^{Bt}c = e^{\alpha t}(c_1 \cos(\beta t) + c_2 \sin(\beta t), -c_1 \sin(\beta t) + c_2 \cos(\beta t)), \quad c_1, c_2 \in \mathbb{R}.$$

Tali funzioni soddisfano la relazione

$$\|v(t)\| = \sqrt{c_1^2 + c_2^2} e^{\alpha t},$$

e quindi giacciono su spirali. Se $\alpha < 0$ le traiettorie tendono ad entrare nell'origine (fuoco stabile), mentre se $\alpha > 0$ tendono ad allontanarsene (fuoco instabile). Il ritratto di fase è illustrato in Figura 8.5.

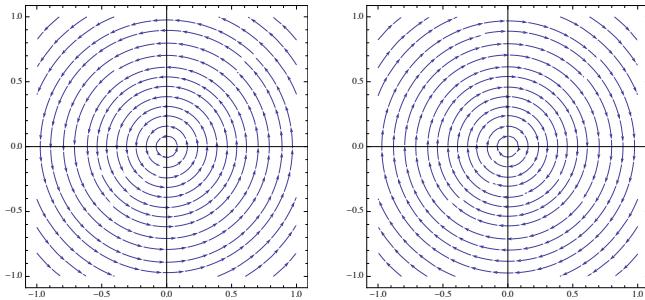


Figura 8.6: Centro

Centro: $\lambda = i\beta$, $\mu = -i\beta$. In questo caso le matrici B ed e^{Bt} hanno la forma

$$B = \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix} = \beta J \quad e^{Bt} = e^{\beta t J} = \begin{pmatrix} \cos(\beta t) & \sin(\beta t) \\ -\sin(\beta t) & \cos(\beta t) \end{pmatrix}.$$

Quindi le soluzioni sono tutte del tipo

$$v(t) = e^{Bt}c = (c_1 \cos(\beta t) + c_2 \sin(\beta t), -c_1 \sin(\beta t) + c_2 \cos(\beta t)), \quad c_1, c_2 \in \mathbb{R}.$$

e soddisfano la relazione

$$\|v(t)\| = \sqrt{c_1^2 + c_2^2}.$$

Si tratta quindi sempre di orbite circolari intorno all'origine; il ritratto di fase è illustrato in Figura 8.6. Osserviamo che, a differenza del caso dei fuochi stabili, in questo caso le traiettorie non si avvicinano asintoticamente all'equilibrio. Tuttavia resta vero che le traiettorie vicine all'origine non se ne allontanano. Questi due diversi fenomeni di stabilità saranno approfonditi nel Capitolo 10.

Per concludere, ricordiamo che per $A \in M_2(\mathbb{R})$ vale la relazione

$$\det(A - \lambda I) = \lambda^2 - \text{tr} A \lambda + \det A$$

dove $\text{tr} A$ è la traccia della matrice (che coincide con la somma degli autovalori). Di conseguenza, le informazioni sugli autovalori possono essere tutte ricavate attraverso il determinante e la traccia di A , essendo

$$\lambda_{1,2} = \frac{1}{2} \left(\text{tr} A \pm \sqrt{(\text{tr} A)^2 - 4 \det A} \right).$$

Si può quindi riassumere la fenomenologia delle soluzioni dei sistemi lineari bidimensionali omogenei a coefficienti costanti con il grafico nel piano traccia-determinante rappresentato nella Figura 8.7. La parabola disegnata corrisponde alla condizione $(\text{tr} A)^2 = 4 \det A$. Vale la pena di osservare che il grafico

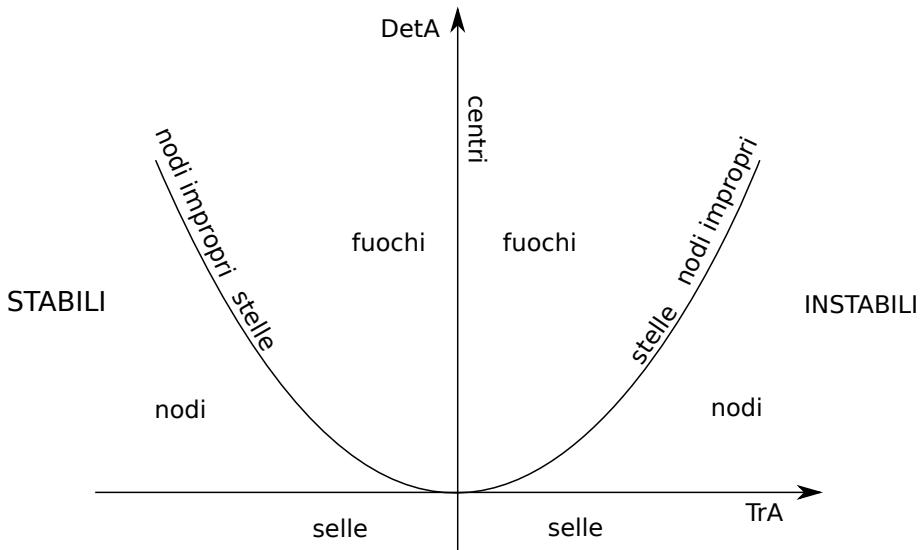


Figura 8.7: Quadro riassuntivo sulle tipologie di traiettorie planari

evidenzia il fatto che, se $\det A > 0$, l'origine risulta essere un equilibrio stabile se e solo se la traccia della matrice A è negativa, che, a conti fatti, corrisponde a richiedere che i due autovalori siano negativi, coerentemente con il risultato generale di stabilità. Inoltre, nel caso dei centri (autovalori complessi con parte reale nulla), le soluzioni sono limitate ma non asintoticamente convergenti all'equilibrio, coerentemente con il risultato di limitatezza.

Capitolo 9

Equazioni differenziali lineari di ordine superiore

9.1 Rappresentazione delle soluzioni

Siamo interessati alla determinazione delle soluzioni di un'equazione differenziale lineare scalare di ordine n del tipo

$$x^{(n)} + a_{n-1}(t)x^{(n-1)} + \cdots + a_1(t)x' + a_0(t)x = b(t) \quad (9.1)$$

con $a_j, b \in C(I)$, $I \subseteq \mathbb{R}$ intervallo. Dal momento che l'equazione può essere scritta come un sistema di equazioni differenziali in \mathbb{R}^n , ponendo $x_0 = x$ e

$$\begin{cases} x'_0 = x_1, \\ x'_1 = x_2, \\ \vdots \\ x'_{n-2} = x_{n-1}, \\ x'_{n-1} = -a_0x_0 - a_1x_1 - \cdots - a_{n-1}x_{n-1} + b, \end{cases}$$

basterà rileggere i risultati già ottenuti per i sistemi del tipo $x' = Ax + b$ in \mathbb{R}^n nel caso particolare

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b \end{pmatrix}.$$

Per brevità di notazione, nel seguito indicheremo con $L(t)$ l'operatore differenziale lineare

$$L(t)x = x^{(n)} + a_{n-1}(t)x^{(n-1)} + \cdots + a_1(t)x' + a_0(t)x, \quad x \in C^n(I). \quad (9.2)$$

Visto il sistema lineare associato all'equazione, le condizioni iniziali naturali da impostare nel problema di Cauchy sono quelle che fissano il valore della funzione e delle sue prime $n - 1$ derivate nello stesso istante $t_0 \in I$.

Definizione 9.1.1 (Problema di Cauchy). Siano a_j , $j = 1, \dots, n - 1$ e b funzioni continue in un intervallo $I \subseteq \mathbb{R}$. Fissati $t_0 \in I$ e $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$, diremo che la funzione x è soluzione del problema di Cauchy

$$\begin{cases} L(t)x = b(t) \\ x(t_0) = \xi_1, \quad x'(t_0) = \xi_2, \dots, x^{n-1}(t_0) = \xi_n. \end{cases} \quad (9.3)$$

se

- i) x è di classe C^n in I ;
- ii) $L(t)x(t) = b(t)$ per ogni $t \in I$;
- iii) $x(t_0) = \xi_1$, $x'(t_0) = \xi_2, \dots, x^{n-1}(t_0) = \xi_n$

Osserviamo che le condizioni iniziali sono quelle corrette in molti contesti; ad esempio, nel caso di equazioni del secondo ordine che descrivono il moto di un corpo in un campo di forze, è assolutamente ragionevole aspettarsi la determinazione univoca della dinamica una volta assegnate la posizione e la velocità iniziali.

Riassumiamo tutti i risultati noti.

Teorema 9.1.2. *Sia $L(t)$ l'operatore differenziale lineare definito in (9.2).*

- *Il problema di Cauchy (9.3) ammette un'unica soluzione per ogni scelta di $t_0 \in I$ e di $\xi \in \mathbb{R}^n$.*
- *L'insieme delle soluzioni di un'equazione differenziale lineare omogenea $L(t)x = 0$ è un sottospazio vettoriale V di dimensione n in $C^n(I)$.*

- L'insieme delle soluzioni di un'equazione differenziale lineare non omogenea $L(t)x = b(t)$ è un sottospazio affine V di $C^n(I)$ della forma $V + v$ dove V è l'insieme di tutte le soluzioni dell'equazione omogenea associata e v è una qualsiasi soluzione dell'equazione non omogenea.

- Le funzioni $u_1, \dots, u_n \in C^n(I)$ sono un sistema fondamentale di soluzioni per l'equazione $L(t)x = 0$ (i.e. sono linearmente indipendenti in $C^n(I)$) e generano tutto il sottospazio V se e solo se

$$W(u_1(t), \dots, u_n(t)) = \det \begin{pmatrix} u_1(t) & \cdots & u_n(t) \\ u'_1(t) & \cdots & u'_n(t) \\ \vdots & & \vdots \\ u_1^{n-1}(t) & \cdots & u_n^{n-1}(t) \end{pmatrix} \neq 0, \quad \forall t \in I.$$

Se $\{u_1, \dots, u_n\}$ è un sistema fondamentale di soluzioni per l'equazione $L(t)x = 0$, la matrice

$$X(t) = \begin{pmatrix} u_1(t) & \cdots & u_n(t) \\ u'_1(t) & \cdots & u'_n(t) \\ \vdots & & \vdots \\ u_1^{n-1}(t) & \cdots & u_n^{n-1}(t) \end{pmatrix}, \quad t \in I,$$

è la matrice fondamentale corrispondente. Sappiamo che la formula di Duhamel (5.12) ci permette di trovare una soluzione particolare del sistema non omogeneo associato della forma

$$y(t) = X(t) \int_{t_0}^t X^{-1}(s)b(s) ds. \quad (9.4)$$

La prima componente di questa funzione vettoriale fornisce una soluzione particolare dell'equazione non omogenea $L(t)x = b(t)$. In questo caso particolare l'integrando $\alpha(t) = X^{-1}(t)b(t)$ in (9.4) può essere scritto in maniera più esplicita. Infatti si tratta di risolvere, per ogni $t \in I$ fissato, il sistema lineare $X(t)\alpha(t) = b(t)$ che, per il Teorema di Cramer, ha come unica soluzione

$$\alpha_j(t) = \frac{\det A_j(t)}{W(t)}, \quad j = 1, \dots, n,$$

dove la matrice

$$A_j(t) = \begin{pmatrix} u_1(t) & \cdots & 0 & \cdots & u_n(t) \\ u'_1(t) & \cdots & 0 & \cdots & u'_n(t) \\ \vdots & & & & \vdots \\ u_1^{n-1}(t) & \cdots & b(t) & \cdots & u_n^{n-1}(t) \end{pmatrix}, \quad t \in I,$$

non è altro che la matrice $X(t)$ in cui è stata sostituita la j -sima colonna con la colonna dei termini noti. Quindi

$$\alpha_j(t) = (-1)^{n+j} b(t) \frac{W_j(t)}{W(t)}, \quad j = 1, \dots, n$$

dove $W_j(t)$ è il determinante della matrice $(n-1) \times (n-1)$ ottenuta da X eliminando l'ultima riga e la j -sima colonna

$$W_j(t) = \det \begin{pmatrix} u_1(t) & \cdots & u_{j-1} & u_{j+1} & \cdots & u_n(t) \\ u'_1(t) & \cdots & u'_{j-1} & u'_{j+1} & \cdots & u'_n(t) \\ \vdots & & & & & \vdots \\ u_1^{n-2}(t) & \cdots & u_{j-1}^{n-2} & u_{j+1}^{n-2} & \cdots & u_n^{n-2}(t) \end{pmatrix}, \quad t \in I. \quad (9.5)$$

La funzione vettoriale $y(t) = X(t) \int_{t_0}^t \alpha(s) ds$ è soluzione del sistema lineare associato, per cui la sua prima componente determina una soluzione particolare dell'equazione non omogenea. Abbiamo quindi dimostrato il seguente risultato.

Proposizione 9.1.3 (Formula di Duhamel per equazioni scalari di ordine superiore). *Una soluzione particolare dell'equazione differenziale lineare non omogenea $L(t)x = b(t)$ è data da*

$$v(t) = \sum_{j=1}^n (-1)^{n+j} u_j(t) \int_{t_0}^t b(s) \frac{W_j(s)}{W(s)} ds \quad t \in I$$

dove $\{u_1, \dots, u_n\}$ è un sistema fondamentale di soluzioni per l'equazione omogenea $L(t)x = 0$, W è il relativo Wronskiano e W_j è definito in (9.5) per ogni $j = 1, \dots, n$.

Vediamo cosa diventa la costruzione della soluzione particolare nel caso di equazioni del secondo ordine

$$x'' + a_1(t)x' + a_0(t)x = b(t) \quad t \in I.$$

Deve essere noto un sistema fondamentale di soluzioni $\{u_1, u_2\}$ e la relativa matrice fondamentale

$$X(t) = \begin{pmatrix} u_1(t) & u_2(t) \\ u'_1(t) & u'_2(t) \end{pmatrix}.$$

in modo tale che siano determinati $W(t) = \det X(t)$ e

$$\begin{aligned} W_1(t) &= \det \begin{pmatrix} 0 & u_2(t) \\ b(t) & u'_2(t) \end{pmatrix} = -u_2(t)b(t) \\ W_1(t) &= \det \begin{pmatrix} u_1(t) & 0 \\ u'_1(t) & \frac{b(t)}{a_2(t)} \end{pmatrix} = u_1(t)b(t). \end{aligned}$$

La soluzione particolare dell'equazione è quindi data da

$$v(t) = \int_{t_0}^t b(s) \left[\frac{u_1(s)u_2(t) - u_1(t)u_2(s)}{W(s)} \right] ds.$$

Sempre utilizzando la teoria sviluppata per i sistemi lineari di equazioni differenziali, ricaviamo che, se l'equazione differenziale ha coefficienti costanti, allora è possibile determinare esplicitamente un sistema fondamentale di soluzioni. Consideriamo quindi un'equazione differenziale lineare di ordine n a coefficienti costanti

$$Lx := x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_1x' + a_0x = b \quad (9.6)$$

con $a_j, b \in \mathbb{R}$.

Ricordando la forma particolare della matrice A dei coefficienti del sistema lineare associato, il relativo polinomio caratteristico sarà dato da

$$P(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} -\lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\lambda & 1 & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & -\lambda & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} - \lambda \end{pmatrix}.$$

Sviluppiamo il determinante secondo l'ultima riga, tenendo conto del fatto che i minori complementari relativi agli elementi dell'ultima riga sono tutte matrici triangolari. Più precisamente, il minore complementare relativo all'elemento a_j , $j = 0, \dots, n-2$, è una matrice B_j triangolare che sulla diagonale ha esattamente j elementi che valgono $-\lambda$ e $(n-1-j)$ elementi che valgono

1. Allo stesso modo, il minore complementare B_n dell'elemento $-a_{n-1} - \lambda$ è triangolare con $(n-1)$ elementi che valgono $-\lambda$. Quindi avremo che

$$\begin{aligned} P(\lambda) &= \sum_{j=0}^{n-2} (-1)^{n+j+1} (-a_j) \det B_j + (-1)^{2n} (-a_{n-1} - \lambda) \det B_n \\ &= (-1)^n \sum_{j=0}^{n-2} (-1)^{j+1} (-a_j)(-\lambda)^j + (a_{n-1} + \lambda) \lambda^{n-1} \\ &= \left(\sum_{j=0}^{n-2} a_j \lambda^j + a_{n-1} \lambda^{n-1} + \lambda^n \right), \end{aligned}$$

ossia il seguente risultato.

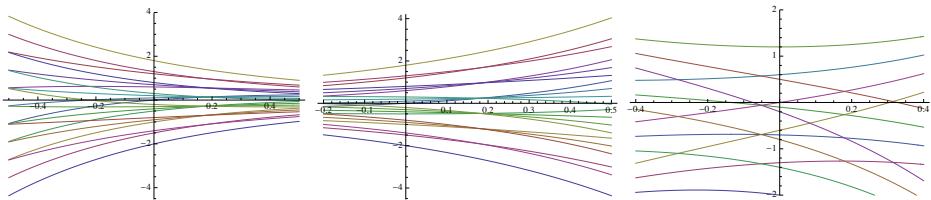
Proposizione 9.1.4. *Il polinomio caratteristico associato all'equazione differenziale lineare omogenea di ordine n a coefficienti costanti (9.6) è dato da*

$$P(\lambda) = \lambda^n + \sum_{j=0}^{n-1} a_j \lambda^j.$$

Osservazione 9.1.5. L'equazione caratteristica $P(\lambda) = 0$ associata all'equazione differenziale viene talvolta ottenuta cercando le condizioni su $\lambda \in \mathbb{R}$ affinché la funzione $x = e^{\lambda t}$ sia soluzione dell'equazione differenziale. Questo approccio però ha diversi limiti: innanzi tutto non chiarisce perché ci si debba aspettare delle soluzioni di tipo puramente esponenziale ed in più nasconde il fatto che i valori di λ che soddisfano l'equazione caratteristica sono autovalori di un'opportuna matrice associata all'equazione.

Dati i coefficienti dell'equazione (9.6) è quindi facilmente ottenibile lo spettro dell'operatore lineare in $\mathcal{L}(\mathbb{R}^n)$ associato. Siano $\lambda_1, \dots, \lambda_r$ gli autovalori reali e $\mu_1 = \alpha_1 \pm i\beta_1, \dots, \mu_s = \alpha_s \pm i\beta_s$ gli autovalori complessi, $r + 2s = n$. Otterremo quindi un sistema fondamentale di soluzioni per l'equazione (9.6) costituito dalle funzioni

$$\begin{aligned} e^{\lambda_i t}, te^{\lambda_i t}, \dots, t^{\text{ma}(\lambda_i)-1} e^{\lambda_i t}, &\quad i = 1, \dots, r, \\ e^{\alpha_j t} \cos(\beta_j t), te^{\alpha_j t} \cos(\beta_j t), \dots, t^{\text{ma}(\mu_j)-1} e^{\alpha_j t} \cos(\beta_j t), \\ e^{\alpha_j t} \sin(\beta_j t), te^{\alpha_j t} \sin(\beta_j t), \dots, t^{\text{ma}(\mu_j)-1} e^{\alpha_j t} \sin(\beta_j t), &\quad j = 1, \dots, s \end{aligned}$$

Figura 9.1: Soluzioni per $\Delta > 0$

9.2 Equazioni del secondo ordine a coefficienti costanti

A titolo di esempio, descriviamo l'insieme delle soluzioni di un'equazione differenziale lineare omogenea

$$ax'' + by' + cy = 0,$$

del secondo ordine a coefficienti costanti $a, b, c \in \mathbb{R}$, $a \neq 0$.

La natura degli autovalori, e conseguentemente del sistema fondamentale di soluzioni, in questo caso dipende banalmente dal segno del discriminante $\Delta = b^2 - 4ac$.

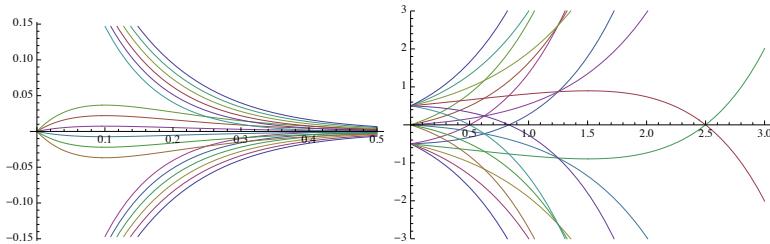
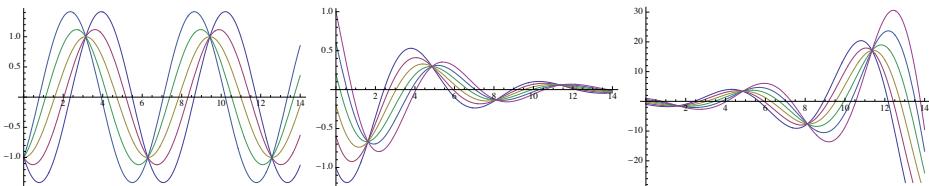
Primo caso: $\Delta > 0$. Abbiamo due autovalori reali e distinti λ_1 e λ_2 . Ne segue che le funzioni $u_1(t) = e^{\lambda_1 t}$ e $u_2(t) = e^{\lambda_2 t}$ formano un sistema fondamentale di soluzioni e che l'insieme di tutte le soluzioni dell'equazione è dato dalle combinazioni lineari

$$x(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}, \quad C_1, C_2 \in \mathbb{R}.$$

Se $\lambda_1 < \lambda_2 \leq 0$ abbiamo un fenomeno di smorzamento esponenziale nel tempo e le soluzioni tendono ad annullarsi (se $\lambda_2 < 0$) o semplicemente a diventare costanti (se $\lambda_2 = 0$) per t che tende a $+\infty$. Se invece $0 \leq \lambda_1 < \lambda_2$ si ha un fenomeno di amplificazione. Quando $\lambda_1 < 0 < \lambda_2$ invece l'equazione ha sia soluzioni che divergono che soluzioni che tendono a zero per t che tende a $+\infty$. Nello spazio delle fasi per il sistema lineare associato, questa situazione corrisponde ad avere un punto stazionario di sella nell'origine. Le tre famiglie di soluzioni sono descritte in figura 9.1

Secondo caso: $\Delta = 0$. Abbiamo unico autovalore reale λ_1 con molteplicità algebrica 2. Ne segue che le funzioni $u_1(t) = e^{\lambda_1 t}$, $u_2(t) = te^{\lambda_1 t}$ formano un sistema fondamentale di soluzioni e che l'insieme di tutte le soluzioni dell'equazione è dato dalle combinazioni lineari

$$x(t) = C_1 e^{\lambda_1 t} + C_2 t e^{\lambda_1 t}, \quad C_1, C_2 \in \mathbb{R}.$$

Figura 9.2: Soluzioni per $\Delta = 0$ Figura 9.3: Soluzioni per $\Delta < 0$

Anche in questo caso tutte le soluzioni hanno andamento esponenziale per t grande, ma alcune di esse possono non essere strettamente monotone vicino all'origine (si veda la figura 9.2).

Terzo caso: $\Delta < 0$. Gli autovalori sono complessi coniugati $\alpha \pm i\beta$, con $\alpha = -b/2a$ e $\beta = \sqrt{-\Delta}/2a$. Ne segue che le funzioni $u_1(t) = e^{\alpha t} \cos(\beta t)$ e $u_2(t) = e^{\alpha t} \sin(\beta t)$ formano un sistema fondamentale di soluzioni e che l'insieme di tutte le soluzioni dell'equazione è dato dalle combinazioni lineari

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

Abbiamo soluzioni oscillanti con frequenza $\beta/2\pi$. Se $\alpha = 0$ allora le soluzioni sono periodiche, se $\alpha > 0$ l'ampiezza delle oscillazioni aumenta esponenzialmente (oscillazioni forzate), se $\alpha < 0$ l'ampiezza delle oscillazioni decade esponenzialmente (oscillazioni smorzate). I grafici di alcune soluzioni sono descritti nella Figura 9.3.

Esempio 9.2.1 (Oscillatore smorzato). Riprendiamo l'esempio dell'oscillatore introdotto all'inizio del Capitolo 2 e consideriamo il problema delle oscillazioni su un piano con attrito. Con buona approssimazione si può supporre che l'attrito sia proporzionale e di verso opposto alla velocità del corpo. Quindi la seconda legge di Newton in questo caso diventa

$$mx''(t) = -kx(t) - hx'(t),$$

con $h > 0$. I risultati che abbiamo ottenuto forniscono tutte le informazioni sulla dinamica in funzione del coefficiente di attrito h e della costante elastica k .

Se $h^2 > 4km$ (attrito dominante), le soluzioni dell'equazione differenziale hanno tutte la forma $x(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$, dove

$$\lambda_{1,2} = \frac{-h \pm \sqrt{h^2 - 4km}}{2m}$$

e C_1, C_2 sono costanti reali arbitrarie. Notiamo che λ_1 e λ_2 sono entrambi negativi e quindi il moto, che non è più periodico come nel caso dell'oscillatore libero, si smorza in maniera esponenziale.

Se $h^2 = 4km$ (attrito bilanciato dalle forze di richiamo), le soluzioni sono tutte della forma $x(t) = (C_1 + C_2 t) e^{-\frac{h}{2m}t}$, con C_1 e C_2 costanti reali arbitrarie. Quindi l'andamento è simile al caso di attrito dominante, ma il moto si smorza un po' più lentamente.

Se $h^2 < 4km$ (attrito debole), le soluzioni sono tutte della forma

$$x(t) = e^{-\frac{h}{2m}t} (C_1 \cos(qt) + C_2 \sin(qt))$$

dove $q = \frac{\sqrt{4km-h^2}}{2m}$ e C_1, C_2 sono costanti reali arbitrarie. In questo caso l'andamento è oscillante, ma l'ampiezza dell'oscillazione decresce in maniera esponenziale.

Una volta noto il sistema fondamentale di soluzioni per le equazioni omogenee, in linea di principio l'insieme di tutte le soluzioni di un'equazione non omogenea

$$ax'' + bx' + cx = f(t) \quad (9.7)$$

si ricava a partire dalla sola determinazione di una sua soluzione particolare, ottenibile con la formula di Duhamel (5.12).

In alcuni casi risulta più agevole applicare quello che chiameremo **metodo di somiglianza**, cioè la determinazione di una soluzione particolare $v(t)$ che ha una forma simile a quella del termine forzante $f(t)$.

f polinomio: se $f(t)$ è un polinomio di grado n si può tentare di vedere se è possibile determinare una soluzione particolare dell'equazione differenziale che sia anch'essa un polinomio di grado n .

Ad esempio, determiniamo l'integrale generale dell'equazione differenziale

$$x'' + x = t^2.$$

Per prima cosa dobbiamo determinare l'integrale generale dell'equazione omogenea associata. L'equazione caratteristica è $\lambda^2 + 1 = 0$ che non ha soluzioni reali. Siamo nel terzo caso discusso nel paragrafo precedente, con $\alpha = 0$ e $\beta = 1$. L'integrale generale dell'equazione omogenea associata è quindi

$$\bar{x}(t) = C_1 \cos t + C_2 \sin t, \quad C_1, C_2 \in \mathbb{R}.$$

Cerchiamo ora una soluzione particolare della forma $v(t) = at^2 + bt + c$ (generico polinomio di secondo grado). Abbiamo $v'(t) = 2at + b$ e $v''(t) = 2a$. Imponiamo che v sia soluzione dell'equazione:

$$2a + at^2 + bt + c = t^2,$$

che è un'identità solo se

$$a = 1, \quad b = 0, \quad 2a + c = 0,$$

cioè per $a = 1$, $b = 0$ e $c = -2$. Quindi la soluzione particolare è $v(t) = t^2 - 2$ e l'integrale generale è

$$x(t) = C_1 \cos t + C_2 \sin t + t^2 - 2, \quad C_1, C_2 \in \mathbb{R}.$$

Proviamo ora a fare lo stesso con l'equazione differenziale

$$x'' - 3x' = t^2 + 1. \tag{9.8}$$

Questa volta l'equazione caratteristica è $\lambda^2 - 3\lambda = 0$ che ha due soluzioni reali $\lambda_1 = 0$ e $\lambda_2 = 3$. Quindi l'integrale generale dell'equazione omogenea associata è

$$\bar{x}(t) = C_1 + C_2 e^{3t}, \quad C_1, C_2 \in \mathbb{R}.$$

Come prima, cerchiamo la soluzione particolare della forma $v(t) = at^2 + bt + c$. Affinché sia soddisfatta l'equazione dovrà essere

$$2a - 3(2at + b) = t^2 + 1,$$

cosa che chiaramente non sarà mai verificata (abbiamo a sinistra un polinomio di primo grado e vorremmo che coincidesse con un polinomio di secondo grado).

Questa è una difficoltà di ordine generale che si incontra tutte le volte che nell'equazione differenziale non compare la funzione x , ma solo le sue derivate. Infatti, se f è un polinomio di grado k e si cerca una soluzione particolare v che sia un polinomio di grado k , l'assenza del termine x nell'equazione fa in modo che a sinistra dell'uguale appaia sempre un polinomio di grado al più $k - 1$.

In questo caso è quindi necessario cercare la soluzione particolare v come un polinomio di grado $k+1$.

Torniamo all'equazione (9.8) e proviamo a cercare v della forma $v(t) = at^3 + bt^2 + ct + d$. Ora abbiamo $v'(t) = 3at^2 + 2bt + c$ e $v''(t) = 6at + 2b$ e l'equazione sarà verificata se

$$6at + 2b - 3(3at^2 + 2bt + c) = t^2 + 1,$$

che risulta essere un'identità se

$$-9a = 1, \quad 6a - 6b = 0, \quad 2b - 3c = 1,$$

ossia per $a = -1/9$, $b = -1/9$ e $c = -11/27$. Non abbiamo ottenuto nessuna condizione sulla costante d , che può essere scelta arbitrariamente, ad esempio $d = 0$. Quindi l'integrale generale dell'equazione differenziale è

$$x(t) = C_1 + C_2 e^{3t} - \frac{1}{9}t^3 - \frac{1}{9}t^2 - \frac{11}{27}t, \quad C_1, C_2 \in \mathbb{R}.$$

f esponenziale: se abbiamo $f(t) = C_0 e^{\alpha t}$, $C_0 \in \mathbb{R}$, possiamo cercare una soluzione particolare della forma $v(t) = C(t)e^{\alpha t}$ dove la funzione $C(t)$ va determinata attraverso l'equazione differenziale. Infatti, poiché

$$\begin{aligned} v'(t) &= C'(t)e^{\alpha t} + C(t)\alpha e^{\alpha t} \\ v''(t) &= C''(t)e^{\alpha t} + 2C'(t)\alpha e^{\alpha t} + C(t)\alpha^2 e^{\alpha t}, \end{aligned}$$

affinché v sia soluzione di

$$ax'' + bx' + c = C_0 e^{\alpha t},$$

deve essere verificata l'identità

$$[a(C'' + 2\alpha C' + \alpha^2 C) + b(C' + \alpha C) + cC]e^{\alpha t} = C_0 e^{\alpha t},$$

che può essere riscritta, dopo le dovute semplificazioni, come

$$aC'' + (2a\alpha + b)C' + (a\alpha^2 + b\alpha + c)C = C_0.$$

Ora abbiamo tre possibilità:

- 1) α non è soluzione dell'equazione caratteristica (ossia $a\alpha^2 + b\alpha + c \neq 0$).

In questo caso è sufficiente scegliere C costante ($C'(t) = C''(t) = 0$),
 $C = \frac{C_0}{a\alpha^2 + b\alpha + c}$.

- 2) α è soluzione semplice dell'equazione caratteristica (ossia $a\alpha^2 + b\alpha + c = 0$, ma $2a\alpha + b \neq 0$). In questo caso basta prendere $C = C_0/(a\alpha + bt)$, in modo da avere $C''(t) = 0$ e l'identità verificata.
- 3) α è soluzione doppia dell'equazione caratteristica (ossia $a\alpha^2 + b\alpha + c = 0$ e $2a\alpha + b = 0$). In questo caso deve essere verificata l'equazione $aC'' = C_0$, quindi con una doppia integrazione si ottiene $C(t) = \frac{C_0}{a} \frac{t^2}{2}$.

Esempio 9.2.2. Determiniamo la soluzione del problema di Cauchy

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

L'equazione caratteristica è $\lambda^2 - 2\lambda + 1 = 0$ ammette come soluzione solo $\lambda = 1$. Quindi siamo nel secondo caso descritto nel paragrafo precedente e l'integrale generale dell'equazione omogenea associata è

$$\bar{x}(t) = C_1 e^t + C_2 t e^t, \quad C_1, C_2 \in \mathbb{R}.$$

Determiniamo la soluzione particolare. Poiché $\alpha = -1$ non è soluzione dell'equazione caratteristica, possiamo cercare la soluzione della forma $v(t) = Ce^{-t}$ con C costante da determinare. Dal momento che $v'(t) = -Ce^{-t}$ e $v''(t) = Ce^{-t}$, imponendo che v risolva l'equazione otteniamo $C + 2C + C = 3$, quindi $C = 3/4$ e l'integrale generale dell'equazione differenziale è

$$x(t) = C_1 e^t + C_2 t e^t + \frac{3}{4} e^{-t}.$$

Ora (e, attenzione, solo ora) possiamo impostare le condizioni iniziali:

$$\begin{aligned} x(0) &= C_1 + \frac{3}{4} = 0 \\ x'(t) &= C_1 e^t + C_2 e^t + C_2 t e^t - \frac{3}{4} e^{-t} \implies x'(0) = C_1 + C_2 - \frac{3}{4} = 1 \end{aligned}$$

da cui si ottiene che $C_1 = -\frac{3}{4}$, $C_2 = \frac{5}{2}$ e la soluzione del problema di Cauchy è

$$x(t) = -\frac{3}{4} e^t + \frac{5}{2} t e^t + \frac{3}{4} e^{-t}.$$

f trigonometrica: se abbiamo $f(t) = C_0 \cos(\alpha t)$ oppure $f(t) = C_0 \sin(\alpha t)$ $C_0 \in \mathbb{R}$, possiamo cercare una soluzione particolare della forma

$$v(t) = C_1 \cos(\alpha t) + C_2 \sin(\alpha t),$$

dove le costanti C_1 e C_2 vanno determinate attraverso l'equazione differenziale. Se, ad esempio, vogliamo determinare l'integrale generale dell'equazione differenziale

$$x'' - 2x' - 3x = -\sin(3t),$$

come sempre prima determiniamo l'integrale generale dell'equazione omogenea associata che risulta essere

$$\bar{x}(t) = C_1 e^{-t} + C_2 e^{3t}, \quad C_1, C_2 \in \mathbb{R}.$$

Per determinare la soluzione particolare, proviamo con $v(t) = C_1 \cos(3t) + C_2 \sin(3t)$. Abbiamo che

$$v'(t) = -3C_1 \sin(3t) + 3C_2 \cos(3t), \quad v''(t) = -9C_1 \cos(3t) - 9C_2 \sin(3t),$$

e quindi, affinché l'equazione sia verificata, dovrà essere

$$(-9C_1 - 6C_2 - 3C_1) \cos(3t) + (-9C_2 + 6C_1 - 3C_2) \sin(3t) = -\sin(3t),$$

ossia

$$-12C_1 - 6C_2 = 0, \quad -12C_2 + 6C_1 = -1,$$

da cui si ricava $C_1 = -\frac{1}{30}$ e $C_2 = \frac{1}{15}$. Quindi l'integrale generale dell'equazione differenziale di partenza è

$$x(t) = C_1 e^{-t} + C_2 e^{3t} - \frac{1}{30} \cos(3t) + \frac{1}{15} \sin(3t), \quad C_1, C_2 \in \mathbb{R}.$$

Facciamo notare esplicitamente che, se avessimo cercato una soluzione del tipo $v(t) = C_1 \sin(3t)$ (ossia simile al termine $f(t)$) non la avremmo trovata. Questo perché, mentre nei due casi precedenti (polinomi ed esponenziali) avevamo funzioni che, dopo una o più derivazioni mantenevano la stessa struttura (qualsiasi derivata di un polinomio è un polinomio e qualsiasi derivata di un esponenziale è un esponenziale), in questo caso se si parte da una funzione trigonometrica (diciamo il seno) dopo una derivazione otteniamo l'altra (ossia il coseno). Quindi in questo caso il metodo di somiglianza può funzionare a patto di cercare soluzioni che siano combinazioni lineari di seno e coseno. Può tuttavia succedere che non sia possibile trovare una soluzione particolare della forma proposta (si veda l'Esempio 9.2.3), per motivi simili a quelli esposti nei casi precedenti.

Esempio 9.2.3 (Oscillazioni forzate). Consideriamo di nuovo l'oscillatore in assenza di attrito, ma questa volta facciamo intervenire una forza esterna $f(t) = \cos(\gamma t)$, $\gamma > 0$, che spinga la molla. L'equazione differenziale soddisfatta dal moto è

$$x'' + \omega^2 x = \cos(\gamma t).$$

L'integrale generale dell'equazione omogenea associata è stato determinato nell'Esempio 5.1.6 ed è

$$\bar{x}(t) = C_1 \cos(\omega t) + C_2 \sin(\omega t), \quad C_1, C_2 \in \mathbb{R}.$$

Cerchiamo ora una soluzione particolare dell'equazione non omogenea, utilizzando il metodo di somiglianza. Sia $v(t) = C_1 \cos(\gamma t) + C_2 \sin(\gamma t)$. Si ha

$$v'(t) = -\gamma C_1 \sin(\gamma t) + \gamma C_2 \cos(\gamma t), \quad v''(t) = -\gamma^2 C_1 \cos(\gamma t) - \gamma^2 C_2 \sin(\gamma t).$$

Sostituendo nell'equazione si ottiene

$$(\omega^2 - \gamma^2)(C_1 \cos(\gamma t) + C_2 \sin(\gamma t)) = \cos(\gamma t),$$

che risulta risolubile se $\gamma \neq \omega$, con $C_2 = 0$ e $C_1 = 1/(\omega^2 - \gamma^2)$. Quindi se $\gamma \neq \omega$ l'integrale generale è

$$x(t) = C_1 \cos(\omega t) + C_2 \sin(\omega t) + \frac{1}{\omega^2 - \gamma^2} \cos(\gamma t), \quad C_1, C_2 \in \mathbb{R}.$$

Osserviamo che se γ e ω sono molto simili (ossia se la frequenza della forzante è molto simile alla frequenza propria dell'oscillatore), l'ampiezza dell'oscillazione è molto grande.

Si può verificare che nel caso limite $\gamma = \omega$ l'integrale generale è

$$x(t) = C_1 \cos(\omega t) + C_2 \sin(\omega t) + \frac{1}{2\omega^2} \cos(\omega t) + \frac{t}{2\omega} \sin(\omega t),$$

quindi l'ampiezza delle oscillazioni cresce linearmente con il tempo. In questo caso si dice che l'oscillatore è in risonanza.

Il metodo di somiglianza descritto nei casi precedenti continua a funzionare quando $f(t) = P(t)e^{\alpha t}$, $f(t) = P(t) \cos(\alpha t)$, oppure $f(t) = P(t) \sin(\alpha t)$, con P polinomio. Ad esempio, determiniamo la soluzione del problema di Cauchy

$$\begin{cases} x'' - 2x' + 5x = te^{2t} \\ x(0) = 0 \\ x'(0) = 0. \end{cases}$$

L'equazione caratteristica è $\lambda^2 - 2\lambda + 5 = 0$. Abbiamo $\Delta = -16$, $\alpha = 1$ e $\beta = 2$, quindi l'integrale generale dell'equazione omogenea associata è

$$\bar{x}(t) = e^t(C_1 \cos(2t) + C_2 \sin(2t)), \quad C_1, C_2 \in \mathbb{R}.$$

Cerchiamo la soluzione particolare della forma $v(t) = (at + b)e^{2t}$. Si ha che

$$v'(t) = (2at + 2b + a)e^{2t}, \quad v''(t) = 4(at + b + a)e^{2t}.$$

Quindi v verifica l'equazione differenziale se

$$[4(at + b + a) - 2(2at + 2b + a) + 5(at + b)]e^{2t} = te^{2t},$$

ossia, dopo le dovute semplificazioni

$$5at + 2a + 5b = t,$$

verificata se $5a = 1$ e $2a + 5b = 0$, cioè $a = \frac{1}{5}$ e $b = -\frac{2}{25}$. Quindi l'integrale generale dell'equazione differenziale è

$$x(t) = e^t(C_1 \cos(2t) + C_2 \sin(2t)) + \left(\frac{1}{5}t - \frac{2}{25}\right)e^{2t}, \quad C_1, C_2 \in \mathbb{R}.$$

Possiamo ora determinare la soluzione del problema di Cauchy, imponendo le condizioni iniziali:

$$\begin{aligned} x'(t) &= e^t((C_1 + 2C_2)\cos(2t) + (C_2 - 2C_1)\sin(2t)) + \left(\frac{2}{5}t + \frac{1}{25}\right)e^{2t} \\ \implies x(0) &= C_1 - \frac{2}{25} = 0, \quad x'(0) = C_1 + 2C_2 + \frac{1}{25} = 0, \end{aligned}$$

da cui si ricava $C_1 = \frac{2}{25}$ e $C_2 = -\frac{3}{50}$. Quindi la soluzione cercata è

$$x(t) = e^t\left(\frac{2}{25}\cos(2t) - \frac{3}{50}\sin(2t)\right) + \left(\frac{1}{5}t - \frac{2}{25}\right)e^{2t}.$$

Concludiamo con un risultato (valido per equazioni lineari di qualsiasi ordine e dalla dimostrazione elementare) che permette di utilizzare le tecniche viste finora per determinare le soluzioni di altre equazioni differenziali lineari non omogenee.

Proposizione 9.2.4 (Principio di sovrapposizione). *Siano f_1 ed f_2 due funzioni continue sullo stesso intervallo $I \subseteq \mathbb{R}$ e siano x_1 e x_2 rispettivamente soluzioni delle equazioni differenziali*

$$ax_1'' + bx_1' + cx_1 = f_1(t), \quad ax_2'' + bx_2' + cx_2 = f_2(t).$$

Allora la funzione $x(t) = x_1(t) + x_2(t)$ è soluzione dell'equazione differenziale

$$ax'' + bx' + cx = f_1(t) + f_2(t).$$

Esempio 9.2.5. Determiniamo l'integrale generale dell'equazione differenziale

$$x'' - 3x' = e^{-2t} + 5. \tag{9.9}$$

L'equazione caratteristica è

$$\lambda^2 - 3\lambda = 0,$$

che ha due soluzioni reali $\lambda = 0$ e $\lambda = 3$. Quindi l'integrale generale dell'equazione omogenea associata è

$$\bar{x}(t) = C_1 + C_2 e^{3t}, \quad C_1, C_2 \in \mathbb{R}.$$

Per determinare la soluzione particolare $v(t)$ dell'equazione non omogenea, cerchiamo una soluzione particolare delle due equazioni

$$x'' - 3x' = e^{-2t}, \quad x'' - 3x' = 5,$$

e poi utilizziamo il principio di sovrapposizione. Una soluzione particolare della prima equazione ha la forma $v_1(t) = Ce^{-2t}$, con C costante da determinare. Abbiamo $v_1'(t) = -2Ce^{-2t}$ e $v_1''(t) = 4Ce^{-2t}$, che sostituite nell'equazione danno

$$(4C + 6C)e^{-2t} = e^{-2t},$$

verificata per $C = 1/10$. Quindi $v_1(t) = e^{-2t}/10$. Una soluzione particolare della seconda equazione ha la forma $v_2(t) = at + b$, con le costanti a e b da determinare (il termine forzante è un polinomio di grado zero, ma nell'equazione non compare la x e quindi dobbiamo cercare la soluzione particolare come un polinomio di primo grado). Abbiamo $v_2'(t) = a$ e $v_2''(t) = 0$, che sostituite

nell'equazione danno $-3a = 5$. Quindi $v_2(t) = -5t/3$ (la scelta di b è arbitraria e abbiamo scelto $b = 0$). Per il principio di sovrapposizione, una soluzione particolare dell'equazione non omogenea (9.9) è

$$v(t) = v_1(t) + v_2(t) = e^{-2t}/10 - 5t/3,$$

quindi il suo integrale generale è

$$x(t) = C_1 + C_2 e^{3t} + \frac{e^{-2t}}{10} - \frac{5}{3}t, \quad C_1, C_2 \in \mathbb{R}.$$

Capitolo 10

Analisi locale di sistemi non lineari autonomi

In questo capitolo conclusivo ci dedicheremo allo studio qualitativo delle soluzioni di sistemi nonlineari autonomi della forma

$$x' = f(x).$$

Per non avere problemi di regolarità, supporremo sempre che f sia di classe C^1 su tutto \mathbb{R}^n . Dal momento che stiamo considerando sistemi autonomi, tutti i ragionamenti saranno fatti nello spazio delle fasi, ossia identificando le soluzioni che differiscono tra loro solo per una traslazione temporale. Mostreremo che c'è una differenza fondamentale tra il comportamento delle soluzioni del sistema vicino agli equilibri (ossia alle soluzioni costanti del sistema) e quello delle soluzioni che ne stanno lontane.

Definizione 10.0.6. Il punto $x_0 \in \mathbb{R}^n$ è singolare per il campo f se $f(x_0) = 0$. In questo caso la funzione costante $x(t) = x_0$ risulta essere un equilibrio per il sistema $x' = f(x)$.

Dimostreremo che, se x_0 è non singolare per il campo f , allora le soluzioni del sistema vicine alla soluzione costante $x = x_0$ del problema di Cauchy

$$\begin{cases} x' = f(x) \\ x(0) = x_0 \end{cases}$$

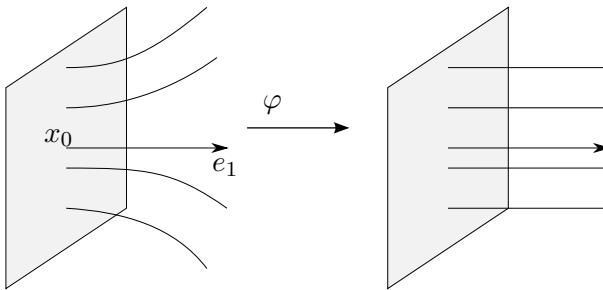


Figura 10.1: Il teorema della scatola di flusso

si trasformano in maniera diffeomorfa in traiettorie rettilinee corrispondenti ad un campo costante (teorema di rettificabilità o della scatola di flusso). Questo risultato è del tutto ragionevole, perché vicino ad un punto non singolare il campo è essenzialmente costante, quindi non c'è motivo di aspettarsi delle variazioni significative delle traiettorie. I punti in cui possono accadere cose interessanti sono i punti singolari, vicino ai quali il campo è piccolo in norma, ma può cambiare verso in maniera imprevedibile. Scopriremo che il comportamento qualitativo delle soluzioni vicino ad un punto singolare (o, meglio, vicino alla soluzione di equilibrio corrispondente) in molti casi può essere determinato per linearizzazione. Dal punto di vista formale, si tratta di pensare che le traiettorie di $x' = f(x)$ vicine all'equilibrio x_0 non siano molto diverse dalle soluzioni del sistema lineare $y' = Df(x_0)y$ (sistema linearizzato in x_0) e, di conseguenza, il comportamento di tali soluzioni sia determinato dagli autovalori della matrice $Df(x_0)$.

10.1 Il teorema di rettificabilità locale

Indichiamo con $x(t; \xi)$ la soluzione del problema di Cauchy

$$\begin{cases} x' = f(x) \\ x(0) = \xi \end{cases} \quad (10.1)$$

Siamo interessati a descrivere le soluzioni del sistema per ξ che varia in un opportuno intorno V di un punto non singolare x_0 .

In generale, un diffeomorfismo $\psi: V \rightarrow W$ trasforma la soluzione $x(t; \xi)$ di (10.1) nella soluzione $y(t, \psi(\xi)) = \psi(x(t; \xi))$ del problema di Cauchy

$$\begin{cases} y' = g(y), & g(y) = D\psi(\psi^{-1}y) f(\psi^{-1}y) \\ y(0) = \psi(\xi). \end{cases} \quad (10.2)$$

Vogliamo costruire un diffeomorfismo che trasformi le soluzioni $x(t; \xi)$, $\xi \in V$ nelle soluzioni di (10.2) con campo g costante (si veda la Figura 10.1). Più precisamente, dal momento che $f(x_0) \neq 0$, possiamo supporre, senza ledere la generalità, che $f(x_0) = e_1$ e cercare di trasformare tramite un'applicazione ψ le soluzioni del sistema (10.1) nelle soluzioni di

$$\begin{cases} y' = e_1, \\ y(0) = \psi(\xi). \end{cases} \quad (10.3)$$

Il fatto che debba valere

$$y(t; \psi(\xi)) = \psi(x(t; \xi)) \quad \forall \xi \in \mathbb{R}^n$$

ci permette di individuare il giusto candidato per la trasformazione. Infatti, se indichiamo con $\phi = \psi^{-1}$, la relazione precedente diventa

$$\phi(y(t; \xi)) = x(t; \phi(\xi)) \quad \forall \xi \in \mathbb{R}^n$$

da cui si ricava, introducendo la notazione $\xi = (\xi_1, \xi')$, $\xi' = (\xi_2, \dots, \xi_n)$, che la mappa ϕ deve soddisfare la proprietà

$$\phi(0, \xi') = y(t; (0, \xi')) = (0, \xi'), \quad \forall \xi' \in \mathbb{R}^{n-1}.$$

Inoltre, osservando che $\xi = y(\xi_1; (0, \xi'))$, $\xi \in \mathbb{R}^n$, dovrà essere

$$\begin{aligned} \phi(\xi) &= \phi(y(\xi_1; (0, \xi'))) = x(\xi_1; \phi(0, \xi')) \\ &= x(\xi_1; (0, \xi')), \quad \forall \xi \in \mathbb{R}^n, \quad \xi = (\xi_1, \xi'). \end{aligned} \quad (10.4)$$

Per il Teorema di esistenza ed unicità, questa applicazione risulta essere ben definita e l'unica cosa da dimostrare sarà il fatto che sia regolare ed invertibile. Gli ingredienti della dimostrazione saranno la dipendenza differenziabile delle soluzioni dai dati ed il Teorema di inversione locale, che ricordiamo per completezza (si veda, ad esempio, [3]).

Teorema 10.1.1 (di inversione locale). *Sia $f \in C^1(A, \mathbb{R}^n)$, A aperto di \mathbb{R}^n e sia $x_0 \in \mathbb{R}^n$ tale che $Df(x_0) \neq 0$. Allora esiste un intorno V di x_0 tale che f risulta invertibile in V , $f^{-1} \in C^1(f(V), \mathbb{R}^n)$ e $D(f^{-1})(f(x)) = (Df)^{-1}(x)$ per ogni $x \in V$.*

Teorema 10.1.2 (di rettificazione locale). *Sia $f \in C^1(A, \mathbb{R}^n)$, A sottoinsieme aperto di \mathbb{R}^n e sia $x_0 \in A$ un punto non singolare per f . Allora esiste un intorno V di x_0 ed un diffeomorfismo ψ tra V ed un opportuno intorno W di 0 tale che per ogni $\xi \in V$ la funzione $t \mapsto \psi(x(t; \xi))$ è l'unica soluzione del problema di Cauchy*

$$\begin{cases} y' = e_1, \\ y(0) = \psi(\xi). \end{cases} \quad (10.5)$$

Dimostrazione. Come abbiamo premesso, supponiamo che $f(x_0) = e_1$ e dimostriamo che la mappa ϕ definita in (10.4) (che corrisponderà all'inversa del diffeomorfismo cercato) è un diffeomorfismo tra un opportuno intorno W dell'origine ed un intorno V in cui $f \neq 0$. La regolarità C^1 della mappa è conseguenza diretta del risultato di differenziabilità delle soluzioni rispetto ai dati iniziali, visto che stiamo supponendo $f \in C^1(A)$ (si veda il Teorema 4.6.2). Resta da dimostrare l'invertibilità della mappa e la regolarità dell'inversa. Grazie al Teorema di inversione locale, queste proprietà seguono dalla invertibilità di $D\phi(0)$. Calcoliamo esplicitamente la matrice. Dalla (10.4) abbiamo che

$$\frac{\partial \phi}{\partial \xi_1} = \frac{\partial}{\partial \xi_1} x(\xi_1; (0, \xi_2, \dots, \xi_n)) = f(x(\xi_1; (0, \xi_2, \dots, \xi_n)))$$

e quindi $\frac{\partial \phi}{\partial \xi_1}(0) = f(x_0) = e_1$. D'altra parte dalla (4.16) ricaviamo

$$\frac{\partial}{\partial \xi_i} x(\xi_1; (0, \xi_2, \dots, \xi_n)) = e_i \quad i = 2, \dots, n$$

e quindi $\frac{\partial \phi}{\partial \xi_i}(0) = e_i$. In conclusione, $D\phi(0) = I$ e $\psi = \phi^{-1}$ è il diffeomorfismo cercato. \square

10.2 Stabilità

Nel paragrafo precedente abbiamo mostrato essenzialmente una equivalenza C^1 tra le traiettorie del sistema autonomo (10.1) vicine ad un punto non singolare e quelle relative ad un campo costante, che si può sintetizzare nella commutatività del seguente diagramma

$$\begin{array}{ccc} \mathbb{R} \times V & \xrightarrow{x(t; \xi)} & V \\ Id \times \psi \downarrow & & \downarrow \psi \\ \mathbb{R} \times W & \xrightarrow{y(t; \xi)} & W \end{array}$$

tra mappe C^1 . Il problema di determinare un campo le cui traiettorie siano equivalenti (nel senso della commutatività del diagramma precedente) a quelle del sistema (10.1) che partono da dati vicini ad un valore x_0 singolare per f è molto più delicato. Quello che si può dimostrare è l'equivalenza con le traiettorie del sistema lineare $y' = Df(x_0)y$, ma solo se tutti gli autovalori di $Df(x_0)$ hanno parte reale non nulla (x_0 punto singolare iperbolico) e, in generale, solo attraverso un omeomorfismo ψ (Teorema di Hartman-Grobman; si veda, ad esempio, [1] o [5]). Noi non dimostreremo questo risultato e ci occuperemo invece di vedere come lo spettro della matrice $Df(x_0)$ fornisca informazioni sul comportamento per tempi lunghi sulle traiettorie vicine ad x_0 (stabilità alla Lyapunov).

Supponiamo che la funzione $f : J \times D \rightarrow \mathbb{R}^n$, $J \subset \mathbb{R}$ intervallo aperto, D aperto di \mathbb{R}^n , soddisfi le ipotesi che garantiscono l'esistenza e l'unicità locale delle soluzioni dei problemi di Cauchy

$$\begin{cases} x' = f(t, x), \\ x(\tau) = \xi. \end{cases} \quad (10.6)$$

Indichiamo con $x(t; \tau, \xi)$ la soluzione locale di (10.6) e con $T_{max}(\tau, \xi)$ l'estremo superiore del suo intervallo massimale di definizione. Supponiamo inoltre $f(t, 0) = 0$, $t \in J$, in modo tale che $x(t; \tau, 0) = 0$ sia soluzione globale di (10.6).

Definizione 10.2.1. Il punto singolare $x = 0$ è stabile (alla Lyapunov) se per ogni suo intorno U e per ogni $\tau \in J$ esiste un intorno V di 0 tale che

$$x(t; \tau, \xi) \in U \quad \forall t \in [\tau, T_{max}(\tau, \xi)), \quad \forall \xi \in V.$$

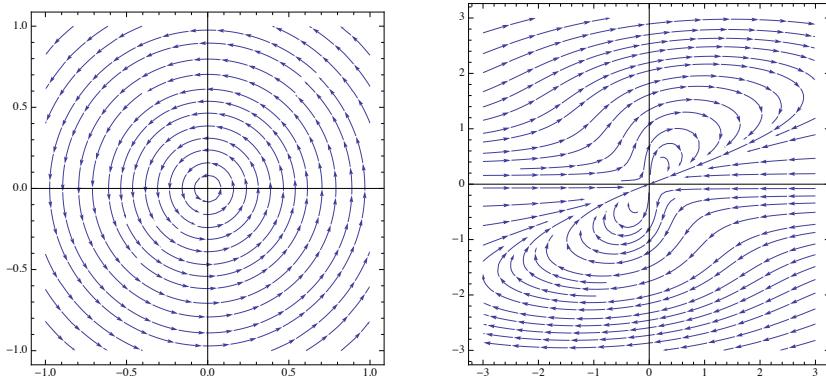


Figura 10.2: Stabilità ed attrattività

Definizione 10.2.2. Il punto singolare $x = 0$ è attrattivo se per ogni $\tau \in J$ esiste un intorno W di 0 tale che per ogni $\xi \in W$ si ha $T_{max}(\tau, \xi) = +\infty$ e $\lim_{t \rightarrow +\infty} x(t; \tau, \xi) = 0$.

Definizione 10.2.3. Il punto singolare $x = 0$ è asintoticamente stabile se è stabile ed attrattivo.

Le proprietà di stabilità e attrattività sono indipendenti: nella Figura 10.2 abbiamo riportato a sinistra un centro che è stabile ma non attrattivo, e a destra il ritratto di fase del sistema non lineare

$$\begin{cases} x' = x^2(y - x) + y^5 \\ y' = y^2(y - 2x) \end{cases}$$

che è attrattivo, ma non stabile.

Quando la scelta degli intorni di stabilità o di attrattività non dipende dal tempo iniziale τ , allora si parla di stabilità o attrattività uniforme.

Definizione 10.2.4. Il punto singolare $x = 0$ è uniformemente stabile se per ogni suo intorno U esiste un intorno V di 0 tale che

$$x(t; \tau, \xi) \in U \quad \forall \tau \in J, \quad \forall t \in [\tau, T_{max}(\tau, \xi)), \quad \forall \xi \in V.$$

Il punto singolare $x = 0$ è uniformemente attrattivo se esiste un intorno W di 0 tale che per ogni $\tau \in J$ per ogni $\xi \in W$ si ha $T_{\max}(\tau, \xi) = +\infty$ e $\lim_{t \rightarrow +\infty} (u; \tau, \xi) = 0$.

Il punto singolare $x = 0$ è uniformemente asintoticamente stabile se è uniformemente stabile ed uniformemente attrattivo.

Osservazione 10.2.5. Studiare la stabilità e l'attrattività nell'origine permette, di fatto, di considerare le analoghe proprietà per qualsiasi soluzione di $x' = f(t, x)$. Infatti, data una soluzione $u(t)$ del sistema, si può introdurre il sistema che misura la deviazione da u

$$y' = f(t, y + u(t)) - f(t, u(t)) \quad (10.7)$$

e che ha $y = 0$ come soluzione globale. A questo punto si può definire u soluzione stabile (attrattiva) se $y = 0$ ha queste proprietà per il sistema (10.7).

In particolare, per i sistemi autonomi diremo che il punto singolare x_0 è stabile (attrattivo) se lo è la soluzione globale $x = x_0$.

Osservazione 10.2.6. Se f è indipendente da t o periodica in t la stabilità e la stabilità asintotica equivalgono rispettivamente all'uniforme stabilità ed all'uniforme stabilità asintotica.

Osservazione 10.2.7 (Confronto tra stabilità e dipendenza continua dai dati). Il risultato di dipendenza continua dai dati visto nel Capitolo 4 afferma che la funzione $\xi \mapsto x(t; \tau, \xi)$ è continua, per τ fissato, uniformemente sugli intervalli compatti contenenti τ (si ricordi la stima (4.15)). Questo garantisce solo che, se $x(t; 0, 0) = 0$, allora le traiettorie che per $\tau = 0$ partono vicine all'origine, restano vicino all'origine per un po' di tempo, ma può tranquillamente succedere che per ogni intorno U dell'origine esista un tempo T tale che tutte le traiettorie che per $t = 0$ si trovavano in U stiano fuori di U per tutti i tempi superiori a T . La proprietà di stabilità è più stringente, richiede che le soluzioni restino nell'intorno dell'origine fino alla loro estinzione.

Osservazione 10.2.8. Per verificare la stabilità (risp. l'attrattività) di $x = 0$ basta, dato un intorno U di 0, determinare l'intorno V (risp. W) di 0 che vada bene per un certo $\tau \in J$ fissato. Infatti a partire da questo si riesce sempre a costruire quello che va bene per ogni altro $\sigma \in J$ nel modo seguente (facciamo i conti per la stabilità, la costruzione dell'intorno nel caso dell'attrattività è del tutto analogo). Fissati $\tau \in J$ ed U intorno dell'origine, sia V tale che

$$x(t; \tau, \xi) \in U \quad \forall (t, \xi) \in [\tau, T_{\max}(\tau, \xi)] \times V.$$

Sia $\sigma < \tau$; per il risultato di dipendenza continua dai dati iniziali (4.15) esiste V' tale che

$$x(t; \sigma, \eta) \in V \quad \forall (t, \eta) \in [\sigma, \tau] \times V'.$$

Quindi, se $\eta \in V'$ abbiamo, una volta osservato che $x(\tau; \tau, \xi) = \xi$ e quindi che $V \subseteq U$,

$$\begin{cases} x(t; \sigma, \eta) \in V \subseteq U & t \in [\sigma, \tau] \\ x(t; \sigma, \eta) = x(t; \tau, x(\tau; \sigma, \eta)) \in U & t \in [\tau, T_{\max} \sigma, \eta] \end{cases}$$

dove nel secondo caso abbiamo usato che per $\eta \in V'$ si ha $x(\tau; \sigma, \eta) \in V$.

Se invece $\sigma > \tau$, allora $x(t; \sigma, \eta) = x(t; \tau, x(\tau; \sigma, \eta))$ e, per la continuità di $x(\tau; \sigma, \cdot)$, l'insieme $V' = \{\eta \in D : x(\tau; \sigma, \eta) \in V\}$ è un aperto. Quindi per ogni $\eta \in V'$, $x(t; \sigma, \eta) \in U$.

Nel caso di sistemi lineari omogenei a coefficienti costanti, la rappresentazione esplicita delle soluzioni ci permette di caratterizzare completamente la stabilità e la stabilità asintotica dell'origine in termini delle proprietà dello spettro della matrice dei coefficienti.

Teorema 10.2.9 (Stabilità lineare). *Sia $A \in M_n(\mathbb{R})$. L'equilibrio $x = 0$ per il sistema $x' = Ax$*

- 1) è stabile se e solo se
 - i) gli autovalori di A hanno parte reale negativa o nulla;
 - ii) ogni autovalore di A con parte reale nulla è semisemplice.
- 2) è asintoticamente stabile se e solo se tutti gli autovalori di A hanno parte reale strettamente negativa.

Dimostrazione. Grazie all'Osservazione 10.2.8, basta verificare cosa succede alle soluzioni con dato assegnato $\xi \in \mathbb{R}^n$ in $\tau = 0$. Sappiamo che tali soluzioni hanno la forma $u(t) = e^{At}\xi$ per cui $x = 0$ risulterà essere stabile se saremo in grado di dimostrare che

$$\omega := \sup_{t \geq 0} \|e^{At}\|_{\mathcal{L}(\mathbb{R}^n)} < +\infty. \quad (10.8)$$

Infatti, se vale (10.8), per ogni $\varepsilon > 0$ si ha che

$$\|e^{tA}\xi\| < \varepsilon, \quad \forall t \geq 0, \quad \forall \xi \text{ tale che } \|\xi\| \leq \frac{\varepsilon}{\omega}.$$

D'altra parte (10.8) equivale a richiedere che tutte le soluzioni del sistema $x' = Ax$ siano limitate. Infine, per il Criterio di limitatezza 8.1.6, la limitatezza delle soluzioni equivale alle richieste i) e ii). Ricapitolando: le richieste i) e ii) sullo spettro di A equivalgono alla richiesta di limitatezza di tutte le soluzioni del sistema, che equivale alla (10.8) che, a sua volta, implica che $x = 0$ sia stabile per il sistema $x' = Ax$.

Viceversa, se una delle proprietà i) o ii) non è verificata, sempre per il Criterio di limitatezza 8.1.6 è possibile determinare $\xi_0 \in \mathbb{R}^n$ tale che $\|e^{At}\xi_0\|$ diverga per t che tende a $+\infty$. Questo ci permette di costruire le funzioni $e^{At}(\varepsilon\xi_0)$, $\varepsilon > 0$, soluzioni del problema di Cauchy con dato iniziale arbitrariamente vicino a zero, ma tali che $\|e^{At}(\varepsilon\xi_0)\|$ diverga per t che tende a $+\infty$, mostrando che l'equilibrio $x = 0$ non è stabile.

In maniera del tutto analoga, utilizzando il Criterio di stabilità 8.1.2, si mostra la condizione necessaria e sufficiente per la stabilità asintotica. \square

Consideriamo ora una situazione apparentemente poco più generale del caso lineare. Sia $g \in \text{Lip}_{\text{loc}}^x(J \times D, \mathbb{R}^n)$, $J \subseteq \mathbb{R}$ intervallo aperto, D aperto di \mathbb{R}^n contenente l'origine, tale che $g(t, 0) = 0$ per ogni $t \in J$ e

$$\lim_{x \rightarrow 0} \frac{g(t, x)}{\|x\|} = 0 \quad \text{uniformemente in } t \in J. \quad (10.9)$$

Data una matrice $A \in M_n(\mathbb{R})$, consideriamo il sistema lineare perturbato

$$x' = Ax + g(t, x). \quad (10.10)$$

Dimostreremo che l'equilibrio $x = 0$ del sistema nonlineare (10.10) ha le stesse proprietà di stabilità asintotica che ha l'equilibrio del problema linearizzato $x' = Ax$.

Questa proprietà si basa sulla seguente semplice osservazione: se $x(t; \tau, \xi)$ è la soluzione di (10.10) con dato iniziale $x(\tau) = \xi$, allora coincide con l'unica soluzione del problema di Cauchy

$$\begin{cases} x' = Ax + g(t, x(t; \tau, \xi)) \\ x(\tau) = \xi \end{cases}$$

corrispondente ad un sistema lineare non omogeneo e quindi può essere rappresentata tramite la formula di Duhamel

$$x(t) = e^{(t-\tau)A}\xi + \int_{\tau}^t e^{(t-s)A}g(s, x(s)) ds \quad t \in I(x). \quad (10.11)$$

Teorema 10.2.10 (Stabilità per sistemi lineari perturbati). *Supponiamo che*

- $A \in M_n(\mathbb{R})$ abbia tutti autovalori con parte reale strettamente negativa;
- $g \in \text{Lip}_{\text{loc}}^x(J \times D, \mathbb{R}^n)$ soddisfi (10.9).

Allora $x = 0$ è un equilibrio uniformemente asintoticamente stabile per il sistema (10.10).

Dimostrazione. Dal fatto che tutti gli autovalori di A hanno parte reale strettamente negativa si deduce facilmente che esistono $\alpha, \beta > 0$ tali che

$$\|e^{At}\|_{\mathcal{L}(\mathbb{R}^n)} \leq \beta e^{-\alpha t}, \quad \forall t \geq 0.$$

Inoltre è evidente che non sia restrittivo supporre $\beta > 1$. Sia $x(t; \tau, \xi)$ una soluzione di (10.10). Dalla formula di rappresentazione (10.11) otteniamo che

$$\|x(t; \tau, \xi)\| \leq \beta e^{-\alpha(t-\tau)} \|\xi\| + \beta \int_{\tau}^t e^{-\alpha(t-s)} \|g(s, x(s; \tau, \xi))\| ds \quad (10.12)$$

per ogni $t \in [\tau, T_{\max}(\tau, \xi))$. Fissiamo $0 < \varepsilon < \alpha$ e sia $0 < \delta < \epsilon$ tale che

$$\|g(t, x)\| \leq \frac{\varepsilon}{\beta} \|x\|, \quad \forall \|x\| \leq \delta, \quad \forall t \geq \tau. \quad (10.13)$$

Per dimostrare la stabilità asintotica basta mostrare che

$$\|x(t; \tau, \xi)\| < \delta < \varepsilon \quad \forall \|\xi\| \leq \frac{\delta}{\beta}, \quad \forall t \in [\tau, T_{\max}(\tau, \xi)).$$

Lo dimostreremo per assurdo: se così non fosse, saremmo in grado di determinare un valore iniziale ξ con $\|\xi\| \leq \delta/\beta$ ed un tempo $\bar{t} \in [\tau, T_{\max}(\tau, \xi))$ tale che

$$\bar{t} = \inf\{t \in [\tau, T_{\max}(\tau, \xi)): \|x(t; \tau, \xi)\| = \delta\}.$$

Osserviamo che $\|x(\tau; \tau, \xi)\| = \|\xi\| < \delta$, quindi $\bar{t} > \tau$. Utilizzando (10.12) e (10.13), otteniamo per $t \in [\tau, \bar{t}]$

$$\|x(t; \tau, \xi)\| \leq \delta e^{-\alpha(t-\tau)} + \varepsilon \int_{\tau}^t e^{-\alpha(t-s)} \|x(s; \tau, \xi)\| ds$$

e, moltiplicando per $e^{\alpha t}$,

$$e^{\alpha t} \|x(t; \tau, \xi)\| \leq \delta e^{-\alpha t} + \varepsilon \int_{\tau}^t e^{\alpha s} \|x(s; \tau, \xi)\| ds.$$

Possiamo applicare la versione semplificata del lemma di Gronwall [4.1.2] alla funzione $e^{\alpha t} \|x(t; \tau, \xi)\|$, ottenendo

$$e^{\alpha t} \|x(t; \tau, \xi)\| \leq \delta e^{\alpha \tau + \varepsilon(t-\tau)}, \quad \forall t \in [\tau, \bar{t}].$$

In particolare abbiamo che

$$\delta = \|x(\bar{t}; \tau, \xi)\| \leq \delta e^{(\varepsilon - \alpha)(\bar{t} - \tau)} < \delta$$

ovviamente impossibile, il che conclude la dimostrazione. \square

Ben più delicata è la dimostrazione del corrispondente criterio di instabilità, che ometteremo (si veda, ad esempio, [1]).

Teorema 10.2.11. *Sia $A \in \mathcal{L}(\mathbb{R}^n)$ tale che almeno uno dei suoi autovalori abbia parte reale positiva e sia g come nel Teorema [10.2.10]. Allora $x = 0$ è un equilibrio instabile per il sistema [10.10].*

Un corollario immediato dei due precedenti risultati è il famoso principio di stabilità linearizzata, valido per i punti singolari di sistemi autonomi.

Teorema 10.2.12 (Principio di stabilità linearizzata). *Sia $f \in C^1(D, \mathbb{R}^n)$, D sottoinsieme aperto di \mathbb{R}^n , e sia $x_0 \in D$ tale che $f(x_0) = 0$.*

- i) *Se tutti gli autovalori di $Df(x_0)$ hanno parte reale strettamente negativa, allora $x = x_0$ è un equilibrio asintoticamente stabile per il sistema autonomo $x' = f(x)$.*
- i) *Se un autovalore di $Df(x_0)$ ha parte reale strettamente positiva, allora $x = x_0$ è un equilibrio instabile per il sistema autonomo $x' = f(x)$.*

Dimostrazione. Basta osservare che x è soluzione di $x' = f(x)$ se e solo se $y = x - x_0$ risolve $y' = Df(x_0)y + g(y)$ con $g(y) = f(y + x_0) - Df(x_0)y$ che verifica le ipotesi dei Teoremi [10.2.10] e [10.2.11]. \square

Osservazione 10.2.13. Il principio di stabilità linearizzata ha carattere locale: le informazioni sugli autovalori di $Df(x_0)$ ci dicono che x_0 è stabile, ma non forniscono alcuna informazione sul suo bacino di attrazione.

Osservazione 10.2.14. Il principio di stabilità linearizzata non fornisce alcuna informazione in caso di autovalori con parte reale nulla. In questo caso la linearizzazione non basta per determinare il carattere del punto stazionario e occorrerebbe guardare termini di ordine superiore al primo. Consideriamo, ad esempio, il sistema in \mathbb{R}^2

$$\begin{cases} x' = -y + x^3 \\ y' = x + y^3 \end{cases} \quad Df(0,0) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad g(x,y) = \begin{pmatrix} x^3 \\ y^3 \end{pmatrix}.$$

$Df(0,0)$ ha autovalori $\pm i$, quindi $(0,0)$ è un centro per il sistema linearizzato. D'altra parte

$$\frac{d}{dt}(x^2 + y^2) = 2(xx' + yy') = 2(x^4 - xy + xy + y^4) = 2(x^4 + y^4)$$

quindi le traiettorie del sistema sono delle spirali che si allontanano dall'origine, che risulta essere un equilibrio instabile per il sistema di partenza. D'altra parte, con ragionamenti analoghi si mostra che l'origine è stabile per il sistema

$$\begin{cases} x' = -y - x^3 \\ y' = x - y^3 \end{cases} \quad Df(0,0) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad g(x,y) = \begin{pmatrix} -x^3 \\ -y^3 \end{pmatrix}$$

che ha lo stesso linearizzato del precedente. In questo caso le traiettorie sono spirali entranti nell'origine.

Per concludere, applichiamo il principio di stabilità linearizzata ai modelli predatore-preda descritti nel Capitolo I. Consideriamo il problema con attrito sociale

$$\begin{cases} \dot{x} = \alpha x - \beta xy - \lambda x^2, & \alpha, \beta, \lambda > 0, \\ \dot{y} = -\gamma y + \delta xy - \mu y^2, & \gamma, \delta, \mu > 0, \end{cases}$$

che ha quattro punti critici

$$\begin{aligned} P_1 &= (0,0), \quad P_2 = \left(\frac{\alpha}{\lambda}, 0\right), \quad P_3 = \left(0, -\frac{\gamma}{\mu}\right), \\ P_4 &= \left(\frac{\alpha\mu + \beta\gamma}{\lambda\mu + \beta\delta}, \frac{\alpha\delta - \lambda\gamma}{\lambda\mu + \beta\delta}\right) = (\xi_0, \eta_0), \end{aligned}$$

dove P_4 è il punto di intersezione tra le due rette $\alpha - \beta y - \lambda x = 0$ e $\delta x - \mu y - \gamma = 0$. Abbiamo

$$f(x, y) = \begin{pmatrix} \alpha x - \beta xy - \lambda x^2 \\ -\gamma y + \delta xy - \mu y^2 \end{pmatrix}, \quad Df(x, y) = \begin{pmatrix} \alpha - \beta y - 2\lambda x & -\beta x \\ \delta y & \delta x - \gamma - 2\mu y \end{pmatrix},$$

e, poiché

$$Df(0, 0) = \begin{pmatrix} \alpha & 0 \\ 0 & -\gamma \end{pmatrix}, \quad Df\left(0, -\frac{\gamma}{\mu}\right) = \begin{pmatrix} \alpha + \frac{\beta\gamma}{\mu} & 0 \\ -\frac{\gamma\delta}{\mu} & \gamma \end{pmatrix},$$

questi due equilibri sono instabili. Invece si ha

$$Df\left(\frac{\alpha}{\lambda}, 0\right) = \begin{pmatrix} -\alpha & -\frac{\alpha\beta}{\lambda} \\ 0 & \frac{\alpha\delta}{\lambda} - \gamma \end{pmatrix}$$

e quindi questo equilibrio è stabile se $\frac{\alpha\delta}{\lambda} - \gamma < 0$ e instabile se $\frac{\alpha\delta}{\lambda} - \gamma > 0$. Per quanto riguarda $P_4 = (\xi_0, \eta_0)$, dopo la sostituzione si ottiene che

$$Df(\xi_0, \eta_0) = \begin{pmatrix} -\lambda\xi_0 & -\beta\xi_0 \\ \delta\eta_0 & -\mu\eta_0 \end{pmatrix}.$$

Quindi il polinomio caratteristico è

$$\det(Df(\xi_0, \eta_0) - \omega I) = \omega^2 + (\mu\eta_0 + \lambda\xi_0)\omega + (\beta\delta + \lambda\mu)\xi_0\eta_0$$

e gli autovalori sono

$$\omega_{1,2} = \frac{1}{2} \left[-(\mu\eta_0 + \lambda\xi_0) \pm \sqrt{(\mu\eta_0 + \lambda\xi_0)^2 - 4(\beta\delta + \lambda\mu)\xi_0\eta_0} \right]$$

Quando $\xi_0 > 0$ e $\eta_0 > 0$ (ossia quando le due rette L e M si intersecano nel primo quadrante), i due autovalori hanno sempre parte reale negativa, quindi l'equilibrio è stabile.

In particolare, per il problema di Lotka–Volterra corrispondente a $\lambda = \mu = 0$, si ritrovano i due equilibri $(0, 0)$ e $(\gamma/\delta, \alpha/\beta)$. Il primo è instabile, mentre

$$Df\left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right) = \begin{pmatrix} 0 & -\frac{\gamma\beta}{\delta} \\ \frac{\alpha\delta}{\beta} & 0 \end{pmatrix}$$

è una matrice con autovalori $\pm i\sqrt{\alpha\gamma}$. Quindi $(\gamma/\delta, \alpha/\beta)$ è un centro per il sistema linearizzato ed il criterio di stabilità linearizzata non può aiutarci nell'analisi locale. Abbiamo già risolto questo problema determinando un integrale primo del sistema.

Bibliografia

- [1] H. Amann. *Ordinary Differential Equations*. De Gruyter, Berlin, 1990.
- [2] A. Bressan e B. Piccoli. *Introduction to the Mathematical Theory of Control*. AIMS, Springfield, 2007.
- [3] N. Fusco, P. Marcellini, e C. Sbordone. *Analisi Matematica due*. Liguori Ed., Napoli, 1996.
- [4] A. I. Kostrikin e Yu. I. Manin. *Linear Algebra and Geometry*. Gordon and Breach Science Publishers, Amsterdam, 1997.
- [5] C. Mascia. *EDO Equazioni Differenziali Ordinarie*. Pitagora, Bologna, 2012.
- [6] L.C. Piccinini, G. Stampacchia, e G. Vidossich. *Equazioni differenziali ordinarie in \mathbb{R}^n* . Liguori Ed., Napoli, 1978.
- [7] E. Sernesi. *Geometria 1*. Bollati Boringhieri, Torino, 1989.
- [8] S. Treil. Linear algebra done wrong. Dispense disponibili online all'indirizzo <http://www.math.brown.edu/~treil/papers/LADW/LADW.html>, 2011.