

Introduzione alle equazioni differenziali

Fabio Prestipino

8 maggio 2022

Definition 1 (Equazione differenziale) Si dice equazione differenziale di ordine n una funzione del tipo

$$f(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0$$

dove x è la variabile indipendente ed $y = y(x)$ è la funzione incognita che vogliamo determinare.

Trovare la soluzione di un'equazione differenziale o equivalentemente trovarne l' "integrale" corrisponde al trovare la funzione $y(x)$ tale che rispetti le condizioni imposte dalla funzione. Si noti che esistono infinite equazioni che risolvono una stessa equazione differenziale poiché nel trovare la soluzione si opera un processo analogo a quello d'integrazione da cui risultano costanti arbitrarie, per ogni valore di queste costanti l'equazione differenziale è risolta. Per ottenere un'unica soluzione fra le infinite possibili bisogna anche conoscere le costanti.

La forma in cui si presenta un'equazione differenziale lineare di ordine n (le uniche di cui ci occuperemo) è

$$y^{(n)} + a_1(x)y^{(n-1)} + \dots + a_{n-1}(x)y' + a_n(x)y = f(x)$$

Definiamo una siffatta equazione "omogenea" se $f(x) = 0$, se ciò non si avverasse sarebbe non omogenea. L'equazione omogenea associata ad una non omogenea è la stessa equazione ma con $f(x) = 0$.

1 Il wronskiano

Prima di trattare direttamente la risoluzione di equazioni differenziali è bene svolgere alcune considerazioni algebriche preliminari che ci permetteranno di avere una visione più chiara della strategia generale di risoluzione generale di equazioni differenziali.

Proposition 1 Le soluzioni dell'equazione omogenea

$$y^{(n)} + \sum_{j=1}^n a_j(x)y^{(n-j)}$$

sono uno spazio vettoriale di dimensione pari all'ordine dell'equazione.

Proof 1 Cominciamo con il dimostrare che è uno spazio vettoriale dimostrando che è un sottospazio vettoriale delle funzioni: basterà dimostrare che contiene lo zero, che è chiuso rispetto a somma e a prodotto per uno scalare. È immediato notare che lo zero ($y = 0$) è soluzione dell'equazione. Per dimostrare gli altri due punti supponiamo di conoscere n soluzioni (tante quanto è l'ordine dell'equazione) $y_1(x), \dots, y_n(x)$ diverse da zero e linearmente indipendenti e siano $c_1, \dots, c_n \in \mathbb{R}$; mostriamo che la combinazione lineare delle soluzioni (usando gli c_i come

scanali) è anch'essa una soluzione:

$$\begin{aligned}
 y_0 &= \sum_{j=1}^n c_j y_j \\
 \left(\sum_{j=1}^n c_j y_j \right)^{(n)} + a_1(x) \left(\sum_{j=1}^n c_j y_j \right)^{(n-1)} + \dots + a_n(x) \left(\sum_{j=1}^n c_j y_j \right) &= 0 \\
 \begin{cases} (c_1 y_1^{(n)}) + a_1(x)(c_1 y_1^{(n-1)}) + \dots + a_{n-1}(c_1 y_1'(x)) + a_n(x)(c_1 y_1) = 0 \\ (c_2 y_2^{(n)}) + a_1(x)(c_2 y_2^{(n-1)}) + \dots + a_{n-1}(c_2 y_2'(x)) + a_n(x)(c_2 y_2) = 0 \\ \dots \\ (c_n y_n^{(n)}) + a_1(x)(c_n y_n^{(n-1)}) + \dots + a_{n-1}(c_n y_n'(x)) + a_n(x)(c_n y_n) = 0 \end{cases}
 \end{aligned}$$

Sostituendo la soluzione otteniamo una somma di sommatorie che equivale a sommare le n equazioni del sistema, la somma di cose uguali a zero resta uguale a zero dunque anche la c.l. delle soluzioni è soluzione.

Ora vogliamo dimostrare che la dimensione dello spazio vettoriale delle soluzioni è n . Dimostrare ciò equivale a dimostrare che, conoscendo n soluzioni l.i. dell'equazione differenziale, qualsiasi altra soluzione y_{n+1} deve essere c.l. delle n soluzioni note. Fissiamo un punto arbitrario x_0 e calcoliamo le soluzioni note in questo punto

$$y_1(x_0) = b_0, \dots, y_n(x_0) = b_n$$

Allora se la dimensione fosse n la soluzione $n+1$ deve poter essere scritta come

$$c_1 y_1(x_0) + \dots + c_n y_n(x_0) = y_{n+1}(x_0)$$

cioè, devono esistere n numeri reali c_i tali che questa uguaglianza sia vera. Possiamo quindi impostare un sistema in cui deriviamo $n-1$ volte questa condizione

$$\begin{cases} c_1 y_1 + c_2 y_2(x_0) + \dots + c_n y_n(x_0) = y_{n+1}(x_0) \\ c_1 y_1' + c_2 y_2'(x_0) + \dots + c_n y_n'(x_0) = y_{n+1}'(x_0) \\ \dots \\ c_1 y_1^{(n-1)} + c_2 y_2^{(n-1)}(x_0) + \dots + c_n y_n^{(n-1)}(x_0) = y_{n+1}^{(n-1)}(x_0) \end{cases}$$

Se esiste una soluzione di questo sistema (dove le incognite sono i coefficienti c_i) allora è possibile scrivere la soluzione $n+1$ come c.l. delle n soluzioni l.i. note. La risolubilità di questo sistema, per il teorema di Rouchè-Capelli, equivale al fatto che il determinante della matrice del sistema omogeneo associato sia diverso da zero.

$$\det \begin{pmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \dots & \dots & \dots & \dots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{pmatrix} \neq 0$$

Una matrice ha determinante diverso da zero se o le righe o le colonne sono linearmente indipendenti; notiamo che per ipotesi le n soluzioni sono l.i. e sappiamo inoltre che derivate di cose linearmente indipendenti continuano ad essere l.i., ne segue che le colonne sono l.i. quindi il determinante è diverso da zero. Abbiamo dunque dimostrato che qualsiasi soluzione può essere scritta come c.l. di n soluzioni l.i. Per definizione l'insieme di n soluzioni è base dello spazio vettoriale se queste sono l.i. e se il loro span genera lo spazio, il teorema è dimostrato.

Definition 2 (Il wronskiano) Si dice Wronskiano il determinante della matrice $n \times n$ che su ogni colonna ha le derivate crescenti di una soluzione di un'equazione differenziale omogenea

$$\text{wronskiano} = \det \begin{pmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \dots & \dots & \dots & \dots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{pmatrix}$$

Theorem 1 Dato il wronskiano di un'equazione differenziale di ordine n , se esso non è nullo tutte e sole le soluzioni dell'equazione differenziale omogenea sono date dalle combinazioni lineari delle n soluzioni che figurano nel wronskiano.

Inoltre, si dimostra che se il wronskiano non si annulla in un punto non si annulla mai.

Il calcolo del wronskiano in un punto è un metodo utile per verificare che le n soluzioni trovate costituiscono effettivamente tutte le soluzioni possibili.

1.1 Soluzione generale di un'equazione differenziale non omogenea

Siano y, \bar{y} due soluzioni particolari di un'equazione differenziale lineare di ordine n non omogenea. Possiamo sottrarre le equazioni in cui sostituiamo le soluzioni

$$\begin{cases} (y^{(n)} + a_1(x)y^{(n-1)} + \dots + a_{n-1}(x)y' + a_n(x)y = f(x) \\ (\bar{y}^{(n)} + a_1(x)\bar{y}^{(n-1)} + \dots + a_{n-1}(x)\bar{y}' + a_n(x)\bar{y} = f(x) \end{cases}$$

$$(y - \bar{y})^{(n)} + a_1(x)(y - \bar{y})^{(n-1)} + \dots + a_{n-1}(x)(y - \bar{y})' + a_n(x)(y - \bar{y}) = 0$$

Visto che la differenza di soluzioni particolari è soluzione dell'omogenea saranno c.l. delle n soluzioni dell'omogenea:

$$\begin{aligned} \exists c_1, \dots, c_n \in \mathbb{R} : y - \bar{y} &= c_1 y_1 + \dots + c_n y_n \\ \Rightarrow y &= c_1 y_1 + \dots + c_n y_n + \bar{y} \end{aligned}$$

Ne segue che una soluzione particolare di un'equazione differenziale non omogenea è uguale ad una combinazione lineare delle n soluzioni dell'omogenea associata più un'altra soluzione particolare della non omogenea. Si noti che y è una singola soluzione in cui i coefficienti c_i sono fissati. Per ottenere la soluzione generale basta "liberare" quei coefficienti: la soluzione generale dell'equazione differenziale non omogenea si ottiene facendo variare in \mathbb{R} tutti i coefficienti c_i .

$$y = c_1 y_1 + \dots + c_n y_n + \bar{y} \quad \forall c_i \in \mathbb{R}$$

Si noti che in abbiamo n gradi di libertà e la dimensione dello spazio delle soluzioni è n , come l'ordine dell'equazione differenziale.

1.2 La strategia

Per risolvere equazioni differenziali non omogenee la strategia da seguire sarà la seguente:

1. Trovare tutte le soluzioni possibili dell'omogenea associata. Per far ciò bisognerà trovare n soluzioni linearmente indipendenti e combinarle linearmente. Per verificare che queste siano effettivamente tutte le soluzioni possibili basterà calcolare il wronskiano.
2. Trovare una soluzione particolare dell'equazione differenziale non omogenea. Visto che per trovare la soluzione generale di un'equazione differenziale non omogenea basta combinare linearmente la soluzione generale della omogenea associata con una soluzione particolare, basterà trovarne una. Esistono vari metodi per trovare una soluzione particolare, uno di quelli che illustreremo è il **metodo di variazione delle costanti arbitrarie**, anche detto **metodo di Lagrange**.
3. Sommare la soluzione generale dell'omogenea associata con la soluzione particolare. Si noti che le soluzioni dell'omogenea associata sono combinate linearmente fra loro mentre quella particolare è semplicemente sommata, i coefficienti che possono variare in \mathbb{R} sono n , la dimensione dello spazio delle soluzioni è infatti n .

2 Risoluzione di equazioni differenziali

2.1 Omogenee del primo ordine

2.1.1 A variabili separabili

Il caso più semplice è quello in cui è possibile avere tutti i termini in cui compare y da un lato dell'equazione e tutti quelli in cui compare x dall'altro per poi integrare e trovare le soluzioni.

Ciò è possibile solo se l'equazione differenziale è del tipo

$$\begin{aligned}y' &= X(x)Y(y) \\ \frac{y'}{Y(y)} &= X(x) \\ \int \frac{y'}{Y(y)} dx &= \int X(x) dx \\ \int \frac{dy}{dx} \frac{dx}{Y(y)} &= \int \frac{dy}{Y(y)} = \int X(x) dx \\ f(y) &= g(x) + c\end{aligned}$$

Dove $f(y)$ è una primitiva di $\frac{1}{Y(y)}$ e $g(x)$ una primitiva di $X(x)$. Se $f(y)$ è invertibile con inversa $\phi(f(y)) = y$ possiamo trovare un'espressione di y

$$y = \phi(g(x) + c)$$

È chiaro che questo metodo è applicabile solo ad una quantità ristretta di casi, dobbiamo quindi adottare un approccio più generale.

2.1.2 Caso generale

Il caso generale delle equazioni differenziali omogenee del primo ordine è

$$y' = a(x)y$$

Questa è a variabili separabili quindi

$$\begin{aligned}\int \frac{dy}{y} &= \int a(x) dx \\ \ln(y) &= \int a(x) dx + C(x) \\ y &= e^{\int a(x) dx + C} = e^{\int a(x) dx} C\end{aligned}$$

Si noti che anche $y = 0$ sarebbe una soluzione ma è stata esclusa nell'applicare il metodo della separazione delle variabili durante la divisione.

2.2 Soluzione particolare non omogenea: metodo di variazione della costante arbitraria

Vogliamo ora trovare una sola soluzione particolare della non omogenea (ovvero una funzione specifica che risolva l'equazione non omogenea).

Imponiamo di voler trovare una soluzione del tipo

$$y = e^{\int a(x) dx + C(x)} = e^{\int a(x) dx} \cdot C(x)$$

questa posizione è giustificata dal fatto che stiamo cercando una funzione che derivata rimanga uguale ma moltiplicata per una costante e sommata ad un'altra funzione; l'unica funzione la cui derivata è se stessa è proprio l'esponenziale, se lo si moltiplica per un'altra funzione ($C(x)$) si ottiene, mediante la derivata di prodotto, la verifica che è proprio questa la forma della soluzione. Andando a ritroso, deriviamo la soluzione posta

$$\begin{aligned}y' &= \frac{d}{dx}(e^{\int a(x) dx} C(x)) = a(x)e^{\int a(x) dx} \cdot C(x) + e^{\int a(x) dx} \cdot C'(x) \\ \Rightarrow a(x)e^{\int a(x) dx} \cdot C(x) + e^{\int a(x) dx} \cdot C'(x) &= a(x)e^{\int a(x) dx} \cdot C(x) + f(x)\end{aligned}$$

In quest'ultimo passaggio abbiamo imposto che la derivata ottenuta rispetti le condizioni imposte dal problema; continuiamo a semplificare

$$\begin{aligned}f(x) &= e^{\int a(x) dx} \cdot C'(x) \\ \Rightarrow C'(x) &= e^{-\int a(x) dx} \cdot f(x) \\ \Rightarrow C(x) &= \int e^{-\int a(x) dx} \cdot f(x) dx\end{aligned}$$

sostituendo l'espressione di $C(x)$ ottenuta nella soluzione che avevamo imposto si ottiene

$$y = e^{\int a(x)dx} \int e^{-\int a(x)dx} \cdot f(x)dx$$

Dove dal primo integrale all'esponente non risulterà nessuna costante, poiché questa era $C(x)$, di cui abbiamo già tenuto conto (in questo caso il simbolo di integrale ha significato "una primitiva di") mentre dal secondo integrale uscirà una costante arbitraria.

2.3 Problema di Cauchy, condizione di risolubilità

Si ipotizzi, ad esempio in un caso fisico, che si voglia risolvere un'equazione differenziale (ad esempio per trovare la legge oraria di un moto a partire dalla seconda legge della dinamica) in cui siano note posizione e velocità iniziali ($y(x_0) = b_0$; $y'(x_0) = b_1$). In questo caso non vorremo trovare tutte le soluzioni possibili, come illustrato fin'ora, ma solamente le soluzioni che rispettino i dati iniziali.

Il problema in cui, data un'equazione differenziale e alcuni dati noti, detti **dati di Cauchy** si voglia trovare la soluzione dell'equazione differenziale che rispetti questi dati, è detto **problema di Cauchy**. I dati di Cauchy (o valori iniziali) sono i valori delle derivate di una soluzione calcolate in un punto x_0 :

$$y(x_0) = b_0; \quad y'(x_0) = b_1, \quad \dots, \quad y^{(n-1)}(x_0) = b_{n-1}$$

dove $b_j \in \mathbb{R}$. Cominciamo con il chiederci se questo problema è risolubile; se lo fosse allora il seguente sistema sarebbe risolubile

$$\begin{cases} (c_1 y_1)(x_0) + \dots + (c_n y_n)(x_0) = b_0 - \bar{y}(x_0) \\ (c_1 y_1')(x_0) + \dots + (c_n y_n')(x_0) = b_1 - \bar{y}'(x_0) \\ \dots \\ (c_1 y_1^{(n-1)})(x_0) + \dots + (c_n y_n^{(n-1)})(x_0) = b_{n-1} - \bar{y}^{(n-1)}(x_0) \end{cases}$$

questo sistema impone solamente che i dati di Cauchy siano c.l. delle soluzioni dell'omogenea associata calcolate in x_0 più una soluzione particolare calcolata in x_0 (portando quest'ultima dall'altro lato dell'uguale). Il problema è risolubile se è possibile determinare tutti i coefficienti c_i che figurano nel sistema. Dal teorema di Rouché-Capelli sappiamo che questo sistema non omogeneo ha una sola soluzione se e solo se le equazioni del sistema omogeneo associato sono linearmente indipendenti, che equivale a dire che il determinante della matrice omogenea associata sia diverso da zero. Riguardandone la definizione, ciò equivale a calcolare il wronskiano in x_0 , se questo è diverso da zero allora è possibile risolvere il problema di Cauchy (ovvero è possibile risolvere l'equazione differenziale ponendo quei valori iniziali).

Osserviamo che visto che se il wronskiano non si annulla in un punto non si annulla mai, basterà verificare che il wronskiano non si annulla in un punto per verificare la risolubilità del problema di Cauchy.

2.3.1 Equazione differenziale del primo ordine non omogenea con dati di Cauchy

Vogliamo risolvere la seguente equazione differenziale non omogenea ($f(x) \neq 0$), imponendo il dato di Cauchy

$$\begin{cases} y' = a(x)y + f(x) \\ y(x_0) = b_0 \end{cases}$$

Notiamo che l'equazione differenziale assomiglia alla derivata di un prodotto. Sia $A(x) = \int_{x_0}^x a(s)ds$, moltiplichiamo ambo i membri per e^{-A} .

$$y' e^{-A} - a(x) e^{-A} y = f(x) e^{-A}$$

Il primo membro è la derivata del prodotto

$$\begin{aligned}(ye^{-A})' &= fe^{-A} \\ \int_{x_0}^x (ye^{-A})' ds &= \int_{x_0}^x fe^{-A(s)} ds \\ y(x)e^{-A(x)} - y(x_0)e^{-A(x_0)} &= \int_{x_0}^x fe^{-A(s)} ds \\ y(x) &= y(x_0)e^{A(x)} + \int_{x_0}^x f(s)e^{A(x)-A(s)} ds = e^{\int_{x_0}^x a(t)dt} [b_0 + \int_{x_0}^x e^{-\int_{x_0}^s a(t)dt} f(s)ds]\end{aligned}$$

Si noti che $A(x_0) = \int_{x_0}^{x_0} a(s)ds = 0$.

Questa è la soluzione generale del problema di Cauchy per equazioni differenziali lineari del primo ordine.

3 Equazioni differenziali lineari a coefficienti costanti

Fin ora ci siamo mantenuti nel caso più generale in cui i coefficienti delle equazioni differenziali sono funzioni, nel caso particolare in cui queste funzioni siano delle costanti reali la risoluzione diventa più semplice. Questo tipo di equazioni differenziali sono molto diffuse in fisica.

3.1 Primo ordine

Vogliamo risolvere un'equazione differenziale del tipo

$$y' = a_1 y + f(x) \quad a_1 \in \mathbb{R}$$

Cominciamo con il risolvere l'equazione omogenea associata; imponiamo di cercare una soluzione del tipo $y_1 = e^{\alpha x}$ con α da determinare

$$\begin{aligned}y_1' &= a_1 y_1 \\ \alpha y_1 &= a_1 y_1 \Rightarrow \alpha = a_1 \\ \Rightarrow y_1 &= e^{a_1 x}\end{aligned}$$

Tuttavia, nel risolvere l'equazione differenziale abbiamo implicitamente integrato dunque dovremmo aggiungere una costante d'integrazione

$$y_1 = e^{a_1 x + C} = e^{a_1 x} C$$

Ricaviamo quindi la soluzione particolare della non omogenea con il metodo di variazione della costante.

$$\bar{y} = e^{a_1 x} \int e^{-a_1 s} f(s) ds$$

Come sappiamo, la soluzione generale è la della soluzione dell'omogenea associata con quella particolare

$$y = y_1 + \bar{y} = e^{a_1 x} [C + \int e^{-a_1 s} f(s) ds]$$

3.2 Secondo ordine

Vogliamo risolvere un'equazione differenziale del tipo

$$y'' + a_1 y' + a_2 y = f(x)$$

con $a_1, a_2 \in \mathbb{R}$. Cominciamo con la risoluzione dell'omogenea associata ponendo che si cerca una soluzione del tipo $y = e^{\alpha x}$

$$\Rightarrow y'' + a_1 y' + a_2 y = e^{\alpha x} (\alpha^2 + a_1 \alpha + a_2) = 0$$

Vogliamo determinare α dunque vogliamo risolvere l'equazione ottenuta dalla posizione. Notiamo che l'esponenziale è sempre maggiore di zero quindi semplifichiamo. Ci rimane solamente il polinomio del secondo grado con incognita in α . Questo è detto **polinomio caratteristico** dell'equazione differenziale.

Definition 3 (Polinomio caratteristico) Data un'equazione differenziale lineare a coefficienti costanti di ordine n

$$y^{(n)} + \dots + a_1 y' + a_2 y = f(x)$$

il polinomio caratteristico ad essa associato è

$$\alpha^n + \dots + a_{n-1} \alpha + a_n = 0$$

Nel caso di ordine due si presentano 3 situazioni possibili rispetto alle soluzioni del polinomio caratteristico:

1. Due soluzioni reali e distinte ($\alpha_1 \neq \alpha_2$)

Ne seguirà che le soluzioni dell'equazione omogenea associata saranno

$$y_1 = e^{\alpha_1 x}; y_2 = e^{\alpha_2 x}$$

Si noti che date queste soluzioni il wronskiano è diverso da zero, ne segue che tutte le possibili soluzioni dell'omogenea associata sono una c.l. di queste due

$$\text{wronskiano} = \det \begin{pmatrix} e^{\alpha_1 x} & e^{\alpha_2 x} \\ \alpha_1 e^{\alpha_1 x} & \alpha_2 e^{\alpha_2 x} \end{pmatrix} = (\alpha_2 - \alpha_1) e^{(\alpha_1 + \alpha_2)x} \neq 0$$

Dunque la soluzione generale dell'omogenea associata è

$$C_1 e^{\alpha_1 x} + C_2 e^{\alpha_2 x}$$

È possibile pervenire alla stessa soluzione interpretando la derivata come un operatore lineare D ; vediamola nel caso più generale in cui $f(x) \neq 0$ per poi tornare al caso omogeneo

$$D^2 y + a_1 D y + a_2 y = f(x)$$

$$(D^2 + a_1 D + a_2) y = f(x)$$

$$(D - \alpha_1)(D - \alpha_2) y = f(x)$$

Quest'ultimo passaggio ha una giustificazione rigorosa che in questo ambito ometteremo, lo si prenda come qualcosa di puramente formale.

Sostituiamo $(D - \alpha_2)y = z$

$$(D - \alpha_1)z = z' - \alpha_1 z = f(x)$$

$$z = (D - \alpha_1)^{-1} f(x)$$

Questa eq.differenziale la sappiamo risolvere con il metodo di variazione delle costanti

$$z = e^{\alpha_1 x} \int e^{-\alpha_1 x} \cdot f(x) dx$$

$$\Rightarrow (D - \alpha_2)y = e^{\alpha_1 x} \int e^{-\alpha_1 x} \cdot f(x) dx$$

Abbiamo visto che applicare $(D - \alpha_1)^{-1}$ ad $f(x)$ equivale a

$$f(x) \xrightarrow{(D - \alpha_1)^{-1}} e^{\alpha_1 x} \int e^{-\alpha_1 x} \cdot f(x) dx$$

Dunque, con un ragionamento analogo possiamo reiterare

$$\begin{aligned} & e^{\alpha_1 x} \int e^{-\alpha_1 x} \cdot f(x) dx \xrightarrow{(D - \alpha_2)^{-1}} e^{\alpha_2 x} \int e^{-\alpha_2 x} \cdot e^{\alpha_1 x} \int e^{-\alpha_1 x} \cdot f(x) dx = \\ & e^{\alpha_1 x} \int e^{(\alpha_1 - \alpha_2)x} \int e^{-\alpha_1 x} \cdot f(x) dx \end{aligned}$$

visto che inizialmente volevamo trovare la soluzione omogenea, sostituiamo $f(x) = 0$

$$y = e^{\alpha_1 x} \int e^{(\alpha_1 - \alpha_2)x} \int e^{-\alpha_1 x} \cdot 0 dx =$$

$$e^{\alpha_1 x} \int e^{(\alpha_1 - \alpha_2)x} C_1 dx = \frac{C_1}{\alpha_1 - \alpha_2} e^{\alpha_1 x} + C_2 e^{\alpha_2 x} =$$

$$C_1 e^{\alpha_1 x} + C_2 e^{\alpha_2 x}$$

Eccoci pervenuti alla stessa soluzione

2. Due soluzioni reali e coincidenti ($\alpha_1 = \alpha_2$)

Analogamente a quanto appena fatto, possiamo vedere la derivata come un operatore D da applicare a funzioni. Per ipotesi l'unica soluzione è α_1 ; possiamo esprimere l'equazione differenziale come

$$(D - \alpha_1)^2 y = (D - \alpha_1)(D - \alpha_1)y = f(x)$$

Definiamo $z \equiv (D - \alpha_1)y = Dy - \alpha_1 y$, come prima, applichiamo l'inverso dell'operatore

$$(D - \alpha_1)z = 0$$

$$z = (D - \alpha_1)^{-1} f(x) = e^{\alpha_1 x} \int e^{-\alpha_1 x} \cdot 0 dx = e^{\alpha_1 x} C_1$$

$$(D - \alpha_1)y = e^{\alpha_1 x} C_1$$

$$y = (D - \alpha_1)^{-1} e^{\alpha_1 x} C_1 = e^{\alpha_1 x} \int e^{-\alpha_1 x} \cdot e^{\alpha_1 x} C_1 dx = e^{\alpha_1 x} \int C_1 dx = e^{\alpha_1 x} (C_1 x + C_2) = C_1 x e^{\alpha_1 x} + C_2 e^{\alpha_1 x}$$

3. Due soluzioni complesse coniugate ($\alpha_1 = u + iv$; $\alpha_2 = u - iv$)

Sfruttiamo la relazione di Eulero per scrivere le soluzioni

$$y_1 = e^{\alpha_1 x} = e^{ux} e^{ivx} = e^{ux} (\cos(vx) + i \sin(vx))$$

$$y_2 = e^{\alpha_2 x} = e^{ux} e^{-ivx} = e^{ux} (\cos(vx) - i \sin(vx))$$

Abbiamo dimostrato che una c.l. di soluzioni continua ad essere una soluzione, possiamo quindi scrivere

$$z_1 = \frac{y_1 + y_2}{2} = e^{ux} \cos(vx)$$

$$z_2 = \frac{y_1 - y_2}{2i} = e^{ux} \sin(vx)$$

La soluzione generale dell'omogenea è quindi

$$y = z_1 + z_2 = C_1 e^{ux} \cos(vx) + C_2 e^{ux} \sin(vx)$$

In ognuno dei tre casi precedenti è possibile calcolare il wronskiano e verificare che è diverso da zero, conferma che le soluzioni trovate sono tutte quelle possibili.

Ricordiamo che lo scopo iniziale era quello di calcolare la soluzione generale di un'equazione differenziale non omogenea a coefficienti costanti; per ottenere questa soluzione dovremo fare una c.l. delle soluzioni dell'omogenea associata trovate (a seconda delle soluzioni del polinomio caratteristico) e di una soluzione particolare della non omogenea.

3.3 Estensione ad ordine n

Reiterando il metodo degli operatori e dei loro inversi n volte è possibile ottenere le soluzioni di equazioni differenziali lineari non omogenee a coefficienti costanti di ordine n .

4 Approccio algebrico

È possibile applicare concetti di algebra lineare per ottenere una visione più profonda delle equazioni differenziali e, in particolare, elaborare un metodo di risoluzione alternativo. Tratteremo unicamente il caso a coefficienti costanti per semplicità.

4.1 Equazioni differenziali omogenee come sistemi

Cominciamo studiando un'equazione differenziale omogenea del secondo ordine a coefficienti costanti

$$y'' + a_1 y' + a_2 y = 0 \text{ con } a_1, a_2 \in \mathbb{R}$$

effettuiamo una semplice sostituzione

$$\begin{cases} y' = v \\ y'' = v' = -a_1 y' - a_2 y = -a_2 y - a_1 v \end{cases}$$

Abbiamo ottenuto un sistema 2x2 del primo ordine (figura solo la derivata prima) in due incognite (y e v), si noti che abbiamo ridotto un problema al secondo ordine ad uno del primo. Possiamo ora scrivere questo sistema in forma vettoriale

$$\frac{d}{dx} \begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} y' \\ v' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -a_2 & -a_1 \end{pmatrix} \begin{pmatrix} y \\ v \end{pmatrix}$$

Quanto fatto nel caso di secondo ordine, è applicabile in modo analogo al caso di ordine n: ridurremo un'equazione differenziale di ordine n ad un sistema del primo ordine con n incognite.

$$y^{(n)} + a_1(x)y^{(n-1)} + \dots + a_{n-1}(x)y' + a_n(x)y = 0$$

$$\begin{cases} y' = v_1 \\ v_1' = v_2 \\ \dots \\ v_{n-1}' = -a_n y - a_{n-1} v_1 - \dots - a_1 v_{n-1} \end{cases}$$
$$\frac{d}{dx} \begin{pmatrix} y \\ v_1 \\ \dots \\ v_{n-1} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -a_n & -a_{n-1} & \dots & -a_1 & \dots \end{pmatrix} \begin{pmatrix} y \\ v_1 \\ \dots \\ v_{n-1} \end{pmatrix}$$

Notiamo che le incognite formano un vettore $\vec{x}(t) \in \mathbb{R}^n$ ed il sistema diventa è una matrice $A \in \mathbb{M}_n$. Possiamo riscrivere l'equazione differenziale come

$$\dot{\vec{x}} = A\vec{x}$$

Se A fosse un numero (ovvero una matrice 1x1), la soluzione si riduce a quella del primo ordine omogenea a coefficienti costanti già vista nella sezione 3.1 ($y = e^{tA}C, C \in \mathbb{R}$). Se invece A è una matrice nxn si sarebbe portati ad eseguire il passaggio formale

$$y = e^{tA}\vec{C}(t)$$

Tuttavia questo vorrebbe dire introdurre il concetto di esponenziale di matrice. Si tratta ora di reinterpretare i simboli, definire l'esponenziale di matrice e risolvere tutti i problemi che questa operazione può sollevare.

4.2 Definizione dell'esponenziale di matrice

Preliminarmente, definiamo l'esponenziale di matrice mediante la serie di Taylor

$$e^{tA}\vec{C}(t) \equiv \sum_{j=0}^{\infty} \frac{t^j}{j!} A^j$$

Dove l'argomento della sommatoria sono funzioni di t a valori matrici nxn dunque l'esponenziale di matrice è una somma infinita di matrici.

Per dare significato alla serie bisogna che converga, per verificare la convergenza necessitiamo innanzi tutto di una distanza.

4.2.1 Definizione norma e distanza

Possiamo considerare una matrice come un vettore appartenente a \mathbb{R}^{n^2} , possiamo vedere le matrici come appartenenti a questo insieme. Per definire una distanza definiamo una norma per poi calcolare la distanza come la norma della differenza di matrici. La norma è una funzione che porta un vettore, in questo caso di $X = \mathbb{R}^{n^2}$ in uno di $[0, +\infty)$ e deve rispettare 4 proprietà:

- $\|x\| \geq 0 \quad \forall x \in X$
- $\|x\| = 0 \Leftrightarrow x = 0$
- $\|\lambda x\| = |\lambda| \|x\|, \quad \forall \lambda \in \mathbb{R}$
- $\|x + y\| \leq \|x\| + \|y\|, \quad \forall x, y \in X$

Le prime tre proprietà sono facili da comprendere nella definizione ma l'ultima è da dimostrare e non è immediata. Ad esempio, verrebbe naturale definire la norma come un'estensione di quella euclidea per le matrici, ma sarebbe molto complesso dimostrare la quarta proprietà. Definiamo allora la norma come

$$\|A\| = \sup_{\vec{x} \in \mathbb{R}^n} \frac{|A\vec{x}|}{|\vec{x}|}$$

$$|\vec{x}| = \sqrt{x_1^2 + \dots + x_n^2}$$

Dimostriamo ora la quarta proprietà della norma, Sia $B \in \mathbb{M}_n$. Fissato $\vec{x} \in \mathbb{R}^n$, il prodotto $B\vec{x}$ sarà sempre un vettore in \mathbb{R}^n dunque scriviamo

$$\begin{aligned} \|A\| &\geq \frac{|A\vec{x}|}{|\vec{x}|} \\ \Rightarrow |A(B\vec{x})| &\leq \|A\| |B\vec{x}| \leq \|A\| \|B\| |\vec{x}| \\ \frac{|A(B\vec{x})|}{|\vec{x}|} &\leq \|A\| \|B\| \\ \sup_{\vec{x} \in \mathbb{R}^n} \frac{|(AB)\vec{x}|}{|\vec{x}|} &= \|AB\| \leq \|A\| \|B\| \end{aligned}$$

Da questi ragionamenti risulta ovvia la proprietà

$$\|A^n\| \leq \|A\|^n$$

ottenibile usando A al posto di B nella dimostrazione precedente e reiterando. Possiamo quindi definire la distanza

Definition 4 (Distanza tra matrici) *La serie*

$$\sum_{j=0}^{\infty} \frac{t^j A^j}{j!}$$

converge alla matrice B se

$$s_n = \sum_{j=0}^n \frac{t^j A^j}{j!}$$

converge a B

$$\Leftrightarrow \forall \varepsilon > 0 \exists N : \forall n > N \|s_n - B\| < \varepsilon$$

4.2.2 Dimostrazione completezza $(\mathbb{M}_n, d(A, B))$

Theorem 2 (Completezza dello spazio metrico $(\mathbb{M}_n, d(A, B))$) *Lo spazio metrico $(\mathbb{M}_n, d(A, B))$ con*

$$d(A, B) = \|A - B\|$$

per come definita sopra è completo.

\Leftrightarrow *Le successioni convergenti in \mathbb{M}_n sono tutte e solo di Cauchy, dove una successione \vec{x}_n è di Cauchy se*

$$\forall \varepsilon > 0 \exists n \in \mathbb{N} : \forall n, m > N |\vec{x}_n - \vec{x}_m| < \varepsilon \quad (1)$$

Proof 2 (Se \vec{x}_n è di Cauchy allora è convergente) \mathbb{M}_n è isomorfo a $\mathbb{R}^{n^2} \equiv \mathbb{R}^d$ dunque il teorema è dimostrato se lo spazio metrico $(\mathbb{R}^d, d(A, B))$ è completo. Quando prendiamo una successione di Cauchy in \mathbb{R}^d abbiamo una successione di vettori (le cui d componenti sono funzioni) che converge ad un vettore. La convergenza di un vettore equivale alla convergenza delle sue componenti dunque

$$\vec{x}_n \text{ di Cauchy} \Leftrightarrow \vec{x}_n^{(i)} \text{ convergente } \forall i = 1, \dots, d$$

dove $x_n^{(i)}$ è l'iesima componente del vettore della successione.

Ecco che abbiamo ridotto un problema in \mathbb{R}^d in uno in \mathbb{R} , poiché ogni componente del vettore è una funzione reale. Visto che \mathbb{R} è completo per l' **assioma di completezza**, allora esiste sempre un reale a cui converge la componente iesima del vettore della successione, da cui segue che tutto il vettore converge in \mathbb{R}^d . Si noti che se inizialmente il problema era cominciato usando una distanza in \mathbb{R}^d , passando alla convergenza in \mathbb{R} si è adottata una diversa distanza.

$$\forall i = 1, \dots, d \exists \bar{x}_i : x_n^{(i)} \rightarrow \bar{x}_i \Rightarrow \vec{x}_n \rightarrow \vec{\bar{x}}$$

4.2.3 Dimostrazione che la serie esponenziale è di Cauchy

Da quanto dimostrato segue che qualsiasi serie se è di Cauchy allora converge. Volevamo dimostrare che la serie dell'esponenziale di matrice che costituisce la soluzione della nostra equazione differenziale è convergente: dimostriamo che è di Cauchy. Ricordando la (1), sia $n > m$ per fissare le idee

$$\begin{aligned} s_n - s_m &= \sum_{j=m+1}^n \frac{t^j}{j!} A^j \\ \|s_n - s_m\| &= \left\| \sum_{j=m+1}^n \frac{t^j}{j!} A^j \right\| \leq \sum_{j=m+1}^n \left\| \frac{t^j}{j!} A^j \right\| = \sum_{j=m+1}^n \frac{|t|^j}{j!} \|A^j\| \leq \\ &\leq \sum_{j=m+1}^n \frac{|t|^j}{j!} \|A\|^j = \sum_{j=m+1}^n \frac{(|t| \|A\|)^j}{j!} = |t_n - t_m| \end{aligned}$$

dove nelle disuguaglianze sono state usate, in ordine, la triangolare, la terza proprietà della norma e la proprietà dimostrata $\|A^j\| \leq \|A\|^j$. Infine si è applicata la definizione di esponenziale di matrice dove

$$t_n 0 \sum_{j=0}^n \frac{(|t| \|A\|)^j}{j!} \rightarrow e^{|t| \|A\|}$$

si noti che **t è un reale fissato** e $\|A\|$ è un valore reale dunque l'esponenziale a cui tende t_n è un reale, analogamente per t_m . La differenza di reali è un reale dunque la serie è di Cauchy, quindi converge. La convergenza appena dimostrata è **convergenza puntuale**, per un t fissato. Ora la definizione di esponenziale di matrice ha senso.

4.3 Ricapitolando

Abbiamo trasformato un'equazione differenziale omogenea di ordine n a coefficienti costanti in un sistema del primo ordine con n incognite, siamo passati alla forma vettoriale e abbiamo ottenuto che la soluzione generale (che ora è ben definita) è

$$\vec{x}(t) = e^{tA} \vec{C}$$

dove $\vec{C} \in \mathbb{R}^n$. $\vec{x}(t)$ è una funzione definita mediante una serie di variabile reale (t) a valori matriciali. Vogliamo verificare che questa soluzione soddisfi l'equazione differenziale (la cui espressione è uguale per ogni ordine)

$$\begin{aligned} \dot{\vec{x}} &= A \vec{x} \\ \frac{d}{dt} e^{tA} \vec{C} &= \frac{d}{dt} \sum_{j=0}^{\infty} \frac{t^j}{j!} A^j \vec{C} = \sum_{j=0}^{\infty} \frac{d}{dt} \frac{t^j}{j!} A^j \vec{C} = \sum_{j=0}^{\infty} j \frac{t^{j-1}}{j!} A^j \vec{C} = \\ &= \sum_{j=1}^{\infty} \frac{t^{j-1}}{(j-1)!} A A^{j-1} \vec{C} = A \sum_{j=1}^{\infty} \frac{t^{j-1}}{(j-1)!} A^{j-1} \vec{C} = A e^{tA} \vec{C} \end{aligned}$$

L'equazione differenziale è soddisfatta.

Attenzione: portare il simbolo di derivata è un'operazione formale che non è detto sia giustificata, in seguito approfondiremo questo tema.

4.4 Il polinomio caratteristico dal punto di vista algebrico

Abbiamo visto che all'equazione differenziale omogenea di grado n a coefficienti costanti si associa una matrice (a partire da un sistema)

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & & & & \\ -a_n & -a_{n-1} & \dots & -a_1 & \end{pmatrix}$$

calcolandone il determinante con il metodo iterativo (in cui in ogni passaggio si elimina una riga e una colonna e si calcola il determinante della matrice ridotta) si ottiene

$$\det(A) = (-1)^{n+2}[a_n + \lambda a_{n-1} + \lambda^2 a_{n-2} + \dots + \lambda^n]$$

che è proprio il polinomio caratteristico visto nell'approccio analitico.

4.5 Risoluzione pratica di equazioni differenziali lineari omogenee a coefficienti costanti

Il calcolo del determinante della matrice A è molto utile nella risoluzione pratica di equazioni differenziali perché ci permette di semplificare la matrice diagonalizzandola o jordanizzandola. Se non semplificassimo la serie dell'esponenziale di matrice avremmo come risultato una serie infinita di matrici.

4.5.1 A diagonalizzabile

Supponiamo che dal calcolo del polinomio caratteristico otteniamo n soluzioni distinte ($\lambda_1 \neq \lambda_2 \neq \dots \neq \lambda_n$), la matrice sarà diagonalizzabile quindi

$$\exists U \in \mathbb{M}_n : U^{-1}AU = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 \dots & \lambda_n & \end{pmatrix}$$

Possiamo cambiare variabile $\vec{x} = U\vec{z}$, riscrivendo l'equazione differenziale otteniamo

$$\begin{aligned} \dot{\vec{x}} &= \frac{d}{dt}Uz = U\dot{\vec{z}} = AUz \\ \dot{\vec{z}} &= U^{-1}AU\vec{z} \\ \dot{\vec{z}} &= \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 \dots & \lambda_n & \end{pmatrix} \vec{z} \\ z(t) &= \sum_{j=0}^{+\infty} \frac{t^j}{j!} \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 \dots & \lambda_n & \end{pmatrix} \vec{C} = \begin{pmatrix} e^{t\lambda_1} & 0 & 0 & \dots & 0 \\ 0 & e^{t\lambda_2} & 0 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 \dots & e^{t\lambda_n} & \end{pmatrix} \vec{C} \\ \Rightarrow x(t) &= U \begin{pmatrix} e^{t\lambda_1} \vec{C}_1 \\ e^{t\lambda_2} \vec{C}_2 \\ \dots \\ e^{t\lambda_n} \vec{C}_n \end{pmatrix} \end{aligned}$$

4.5.2 A non diagonalizzabile

trattiamo il caso del secondo ordine per semplicità, i casi più generali si sviluppano con ragionamenti analoghi. Se il polinomio caratteristico ha due soluzioni coincidenti $\lambda_1 = \lambda_2 \in \mathbb{R}$ non è diagonalizzabile, per semplificarlo lo porteremo in forma di Jordan: una matrice diagonale sommata ad una triangolare superiore. Vogliamo trovare una base rispetto alla quale la matrice A sia in forma di Jordan: cominciamo determinando la matrice associata all'equazione differenziale, calcoliamo $\text{Ker}(A - \lambda I)$ e scegliamo un vettore $\vec{v} \in \text{Ker}(A - \lambda I)$ qualsiasi.

$$\begin{aligned}\lambda^2 + a_1\lambda + a_2 &= (\lambda - \lambda_1)^2 \\ \Rightarrow a_1 &= -2\lambda \quad a_2 = \lambda_1^2\end{aligned}$$

La matrice associata è

$$A = \begin{pmatrix} 0 & 1 \\ -\lambda_1^2 & 2\lambda_1 \end{pmatrix}$$

Imponiamo ora che $\vec{v} \in \text{Ker}(A - \lambda I)$

$$\left[\begin{pmatrix} 0 & 1 \\ -\lambda_1^2 & 2\lambda_1 \end{pmatrix} - \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix} \right] \vec{v} = \begin{pmatrix} -\lambda & 1 \\ -\lambda_1^2 & \lambda_1 \end{pmatrix} \vec{v} = 0$$

Fissato un autovettore \vec{v} , cerchiamo \vec{w} autovettore generalizzato tale che $\vec{w} \in \text{Ker}(A - \lambda I)^2$ e $(A - \lambda I)\vec{w} = \vec{v}$. Preliminarmente notiamo che, visto che $\text{vecv} \in \text{Ker}(A - \lambda I)$

$$\begin{aligned}(A - \lambda I)\vec{v} &= \begin{pmatrix} -\lambda_1 v_1 + v_2 \\ -\lambda_1^2 v_1 + \lambda_1 v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ -\lambda_1^2 v_1 + \lambda_1 v_2 &= 0 \Rightarrow v_2 = \lambda_1 v_1\end{aligned}$$

possiamo ora imporre le condizioni che deve rispettare \vec{w}

$$\begin{pmatrix} -\lambda & 1 \\ -\lambda_1^2 & \lambda_1 \end{pmatrix} \vec{w} = \begin{pmatrix} v_1 \\ \lambda_1 v_1 \end{pmatrix}$$

Un sistema siffatto, per il teorema di Rouchè-capelli, visto che $\text{Ker}(A - \lambda I) = 1$, ha ∞^1 soluzioni, ne scegliamo una.

Abbiamo trovato la base di Jordan rispetto alla quale la matrice A è in forma di Jordan (è dimostrabile che sono l.i.). Otteniamo la matrice in forma di Jordan cambiando base

$$\begin{aligned}P &= \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \\ \tilde{A} &= P^{-1}AP = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \lambda_1 I + N\end{aligned}$$

Questo è un caso particolarmente semplice 2x2 in cui è presente un unico autovalore con un unico blocco di Jordan, tuttavia, al variare del numero di blocchi l'espressione generale sarebbe sempre una matrice diagonale sommata con una N dove varierebbero le posizioni degli 1 a seconda della forma di Jordan.

Per esprimere il risultato dell'equazione differenziale vogliamo scrivere la matrice risultante dall'esponenziale della matrice in forma di Jordan appena ottenuta

$$e^{t\tilde{A}} = e^{t(\lambda_1 I + N)} = \sum_{j=0}^{+\infty} \frac{t^j}{j!} (\lambda_1 I + N)^j$$

Per continuare, vorremmo esprimere il quadrato di binomio di matrici mediante il binomio di Newton tuttavia questo è possibile solo se le due matrici in questione commutano (in generale le matrici non godono della proprietà commutativa) poiché in caso contrario non sarebbe possibile avere ad esempio il doppio prodotto e la formula di Newton perderebbe di validità. calcoliamo quindi il **commutatore** delle due matrici

Definition 5 (commutatore) Il commutatore è un operatore matematico che accetta due matrici, svolgendo la seguente operazione

$$[A, B] = AB - BA$$

se risulta uguale a zero le due matrici commutano.

È immediato verificare che il commutatore delle nostre matrici è nullo, possiamo dunque procedere allo svolgimento del binomio di Newton

$$e^{t(\lambda_1 I + N)} = \sum_{j=0}^{+\infty} \frac{t^j}{j!} \sum_{n=0}^j (\lambda_1 I)^{j-n} + N^n = I + \sum_{j=1}^{+\infty} \frac{t^j}{j!} \sum_{n=0}^j (\lambda_1 I)^{j-n} + N^n$$

Notiamo che il **grado d'impotenza** (dopo quanti elevamenti a potenza si annulla una matrice) di N è 2 (e sempre le matrici N prodotte dalla jordanizzazione hanno grado di impotenza finito) dunque la seconda sommatoria ha solo il termine in $k=0$ e 1

$$e^{t(\lambda_1 I + N)} = I + \sum_{j=1}^{+\infty} \frac{t^j}{j!} (\lambda_1^j I + j \lambda_1^{j-1} N) = \begin{pmatrix} e^{t\lambda_1} & t e^{t\lambda_1} \\ 0 & e^{t\lambda_1} \end{pmatrix}$$

Dunque la soluzione generale dell'equazione omogenea è

$$y(t) = e^{t(\lambda_1 I + N)} \vec{C} = c_1 e^{t\lambda_1} + C_2 t e^{t\lambda_1}$$

Che è esattamente quanto visto con il metodo degli operatori.

4.6 Calcolo del wronskiano

Per verificare che la soluzione trovata ($\vec{x} = e^{tA}$) sia quella più generale calcoliamo il wronskiano e vediamo se è diverso da zero. Supponiamo di conoscere due soluzioni \vec{y}_1 ed \vec{y}_2 , e sappiamo anche i loro valori calcolati in 0 $y_1(0)$ ed $y_2(0)$. Essendo soluzioni possiamo scrivere

$$\begin{aligned} \begin{pmatrix} y_1 \\ \dot{y}_1 \end{pmatrix} &= e^{tA} \begin{pmatrix} y_1(0) \\ \dot{y}_1(0) \end{pmatrix} & \begin{pmatrix} y_2 \\ \dot{y}_2 \end{pmatrix} &= e^{tA} \begin{pmatrix} y_2(0) \\ \dot{y}_2(0) \end{pmatrix} \\ \Rightarrow w(t) &= \det \left(e^{tA} \begin{pmatrix} y_1(0) & y_2(0) \\ \dot{y}_1(0) & \dot{y}_2(0) \end{pmatrix} \right) = \\ &= \det(e^{tA}) \det \begin{pmatrix} y_1(0) & y_2(0) \\ \dot{y}_1(0) & \dot{y}_2(0) \end{pmatrix} \end{aligned}$$

Dove nell'ultimo passaggio è stato applicato il teorema di Binet. Sorge però un problema: cosa vuol dire fare il determinante di una serie infinita di matrici?

La strategia è quella di usare la relazione (che però dobbiamo verificare)

$$P^{-1} e^{tA} P = e^{tP^{-1}AP}$$

per poi applicare Binet.

4.6.1 $\det(e^{tA})$: A diagonalizzabile

Usiamo la definizione dell'esponenziale di matrice e vediamo cosa otteniamo dall'elevamento a potenza della matrice $P^{-1}AP$

$$\begin{aligned} e^{tP^{-1}AP} &= \sum_{n=0}^{+\infty} \frac{(P^{-1}AP)^n}{n!} \\ (P^{-1}AP)^2 &= (P^{-1}AP)(P^{-1}AP) = (P^{-1}A^2P) \\ \Rightarrow (P^{-1}AP)^n &= (P^{-1}A^nP) \\ e^{tP^{-1}AP} &= \sum_{n=0}^{+\infty} P^{-1} \frac{A^n}{n!} P = P^{-1} e^{tA} P \end{aligned}$$

che verifica la relazione. Possiamo quindi scrivere

$$\det(e^{tA}) = \det(P^{-1} e^{tA} P) = \det(e^{tP^{-1}AP}) = e^{t\lambda_1} e^{t\lambda_2} \dots e^{t\lambda_n} = e^{t \operatorname{tr}(A)} \neq 0$$

dove si è usato che il determinante di una matrice diagonale è il prodotto degli elementi sulla diagonale.

4.6.2 $\det(e^{tA})$: A non diagonalizzabile

Se la matrice non è diagonalizzabile la jordanizziamo.

$$\det(e^{t(\lambda I+N)}) = \det(e^{t\lambda I})\det(e^{tN})$$

Notiamo che il l'esponenziale di matrice nel determinante del secondo fattore è una triangolare superiore con soli 1 sulla diagonale, ne segue che il determinante è uguale a 1. L'esponenziale di matrice all'interno del primo determinante è una matrice diagonale esattamente come già visto nel caso precedente. Il determinante, anche nel caso in cui la matrice non fosse diagonalizzabile, rimane diverso da zero.

Tornando alla domanda iniziale, in cui volevamo verificare che il wronskiano è diverso da zero, il problema è risolto: abbiamo dimostrato che il determinante dell'esponenziale di matrice è sempre diverso da zero, il determinante della matrice con le soluzioni sulla prima riga e le derivate di questa sulla seconda è diverso da zero perché per ipotesi le soluzioni solo l.i.

5 Continuità, derivabilità e convergenza uniforme di e^{tA}

Nonostante l'esponenziale di matrice abbia assunto significato una volta dimostrata la convergenza, non ne abbiamo ancora dimostrato la derivabilità (e dunque la continuità) necessaria per dar senso allo svolgimento di equazioni differenziali. Per arrivare a questa dimostrazione bisogna passare per alcuni step intermedi che elenco di seguito per non perdere di mente l'obiettivo delle dimostrazioni che verranno svolte

1. Lo spazio delle funzioni continue che portano da un compatto $[a, b]$ a \mathbb{R} è completo rispetto ad una distanza definita è completo.
2. Estensione al caso dello spazio di funzioni continue che portano da un compatto $[a, b]$ ad una matrice.
3. La serie di matrici dell'esponenziale di matrice è totalmente convergente dunque è uniformemente convergente (serve la completezza di $C([a, b]; \mathbb{M}_n)$).
4. Passaggio al limite sotto al segno d'integrale per funzioni reali (serve la convergenza uniforme in $C([a, b]; \mathbb{R})$)
5. Estensione a funzioni in $C([a, b]; \mathbb{M}_n)$ (serve la convergenza uniforme in $C([a, b]; \mathbb{M}_n)$)
6. Se una successione di funzioni che va da un compatto a \mathbb{R} continua a derivata continua e che converge uniformemente ad f ha derivata convergente uniformemente a g allora f è continua con derivata continua e g è la derivata di f . Si estende ad \mathbb{M}_n al posto di \mathbb{R} . (serve la convergenza uniforme il passaggio a limite sotto al segno d'integrale)
7. Se una successione di funzioni che va da un compatto a \mathbb{R} è continua a derivata continua e la serie di questa successione converge uniformemente a g e la derivata della serie converge uniformemente ad h allora h è la derivata di g . (Che equivale a dimostrare la possibilità di entrare la derivata sotto al segno di serie). Si estende ad \mathbb{M}_n al posto di \mathbb{R} .
8. Applicazione di quanto dimostrato per verificare che la serie esponenziale è continua a derivata continua.

I punti di quest'elenco corrispondono al numero della sottosezione.

5.1 Digressione sugli spazi di Banach

Definition 6 (Spazio di Banach) *Uno spazio vettoriale normato (ovvero a cui si associa una norma), completo rispetto alla distanza associata alla norma, si dice **spazio di Banach**.*

Theorem 3 (Completezza dello spazio di Banach) *Lo spazio $X = C([a, b], \mathbb{R})$ (ovvero lo spazio delle funzioni che portano continue reali in reali) con distanza $d(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$ è uno spazio di Banach.*

Proof 3 Per dimostrare che è uno spazio di Banach dobbiamo dimostrarne la completezza, che è equivalente a dimostrare che tutte e solo le funzioni convergenti sono di Cauchy. La dimostrazione è divisa in due punti:

- Se una successione in X converge uniformemente allora è di Cauchy.
Sia $f_n \in X$ uniformemente convergente, ovvero

$$\exists f \in X : d(f_n, f) \rightarrow 0$$

per dimostrare che è di Cauchy dobbiamo verificare che

$$\lim_{n,m \rightarrow +\infty} d(f_n, f_m) = 0$$

Sfruttiamo le proprietà della distanza e scriviamo

$$0 \leq d(f_n, f_m) \leq d(f_n, f) + d(f_m, f)$$

ma per ipotesi i limiti di queste due distanze per $n, m \rightarrow \infty$ sono uguali a zero quindi la condizione di Cauchy è dimostrata.

- Se una successione è di Cauchy allora converge
Questa dimostrazione è ulteriormente divisa in 3 punti:

1. Una successione di Cauchy in X è puntualmente convergente
Scriviamo la definizione di successione di Cauchy in questo caso:

$$\forall \varepsilon > 0 \exists N > 0 : \forall n, m > N, d(f_n - f_m) = \sup_{x \in [a, b]} |f_n - f_m| < \varepsilon$$

per ogni x fissato in $[a, b]$ si ha

$$\begin{aligned} \forall \varepsilon > 0, \exists N > 0 : \forall n, m > N, \forall x \in [a, b], |f_n(x) - f_m(x)| &\leq \sup_{x \in [a, b]} |f_n(x) - f_m(x)| < \varepsilon \\ \Rightarrow \forall \varepsilon > 0 \exists N > 0 : \forall n, m > N, \forall x \in [a, b], |f_n(x) - f_m(x)| < \varepsilon \end{aligned}$$

Se fissiamo x reale la successione di funzioni reali diventa una successione di reali, per la completezza di \mathbb{R} esiste $\lim_{n \rightarrow +\infty} f_n \equiv f(x)$. Si ottiene infine

$$\forall \varepsilon > 0, \exists N > 0 : \forall n, m > N, \forall x \in [a, b], |f_n(x) - f(x)| < \varepsilon$$

dove ricordiamo che ciò vale per ogni x fissato in $[a, b]$.

2. Una successione di Cauchy in X è uniformemente convergente
Cominciamo sempre scrivendo la definizione di successione di Cauchy

$$\forall \varepsilon > 0 \exists N > 0 : \forall n, m > N, \forall x \in [a, b], d(f_n(x) - f_m(x)) = \sup_{x \in [a, b]} |f_n(x) - f_m(x)| < \varepsilon$$

passiamo al limite $m \rightarrow \infty$ notando che $|f_n(x) - f_m(x)| < \sup_{x \in [a, b]} |f_n(x) - f_m(x)| < \varepsilon$

$$\forall \varepsilon > 0 \exists N > 0 : \forall n, m > N, \forall x \in [a, b], |f_n(x) - f(x)| < \varepsilon$$

Questa è valida per ogni x fissato, fissiamo allora $\sup_{x \in [a, b]}$, otteniamo

$$\lim_{n \rightarrow +\infty} \sup_{x \in [a, b]} |f_n(x) - f(x)| = \lim_{n \rightarrow +\infty} d(f_n(x) - f(x)) = 0$$

che è la definizione di convergenza uniforme.

3. $f(x)$ è continua

Sappiamo che se una funzione è uniformemente continua è anche continua, dunque dimostriamo la continuità uniforme di f_n . Ricordiamo la definizione di continuità uniforme

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall x, y \in [a, b], |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$$

Svolgiamo alcune considerazioni sfruttando le proprietà dei valori assoluti e della distanza

$$\begin{aligned} |f(x) - f(y)| &= |f(x) + f_n(x) - f_n(x) + f_n(x) - f_n(x) - f(y)| \leq \\ &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \leq |f_n(x) - f_n(y)| + 2d(f_n, f) \end{aligned}$$

Per la dimostrazione precedente (punto due) si ha che

$$\forall \varepsilon > 0, \exists N > 0 : \forall n > N d(f_n, f) < \frac{\varepsilon}{3}$$

dunque possiamo scrivere, fissato $n > N$

$$|f(x) - f(y)| \leq |f_n(x) - f_n(y)| + 2d(f_n, f) \leq |f_n(x) - f_n(y)| + \frac{2\varepsilon}{3}$$

Ne segue che f_n è uniformemente continua perché continua sul compatto $[a, b]$ per ipotesi (poiché è una successione che appartiene ad X , lo spazio delle funzioni continue da $[a, b]$ ad \mathbb{R}):

$$\exists \delta > 0 : \forall x, y \in [a, b], |x - y| < \delta \Rightarrow |f_n(x) - f_n(y)| < \frac{\varepsilon}{3}$$

allora possiamo scrivere

$$|f(x) - f(y)| \leq \frac{1}{3}\varepsilon + \frac{2\varepsilon}{3} = \varepsilon$$

Abbiamo dunque dimostrato che se $x \rightarrow y$ allora $f(x) \rightarrow f(y)$, che è la definizione di continuità uniforme. Se è uniformemente continua è anche semplicemente continua.

Visto che lo spazio metrico è completo il sup coincide sempre con il max, possiamo quindi ridefinire la distanza come

$$d(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$$

5.2 Estensione a funzioni reali a valori di matrici

La completezza di $C([a, b], \mathbb{R})$ (ovvero l'essere spazio di Banach) si estende facilmente a $C([a, b], \mathbb{M}_n)$, lo spazio che ci interessa per trattare l'esponenziale di matrice. Preliminarmente dobbiamo adattare la definizione di distanza (prima valida per le funzioni reali), alle matrici: usiamo quella già vista in precedenza

$$d(A, B) = \max_{t \in [a, b]} \|A(t) - B(t)\| = \sup_{t \in [a, b]} \max_{x \neq \mathbb{R}} \frac{|(A(t) - B(t))\vec{x}|}{|\vec{x}|}$$

Per dimostrare che anche questo spazio metrico è di Banach basta ragionare sulle componenti della matrice: ogni componente a_{ij} appartiene a $C([a, b], \mathbb{R})$. Osservando che dire che una matrice è di Cauchy equivale a dire che le sue componenti lo sono, ed equivalentemente per la convergenza, si dimostra immediatamente che se una matrice è di Cauchy allora converge e vice versa. Visto che abbiamo dimostrato che questo spazio è completo possiamo sostituire il sup con il max nella definizione di distanza

$$d(A, B) = \max_{t \in [a, b]} \|A(t) - B(t)\| = \max_{t \in [a, b]} \max_{x \neq \mathbb{R}} \frac{|(A(t) - B(t))\vec{x}|}{|\vec{x}|}$$

5.3 La serie esponenziale di matrice converge uniformemente

Questa dimostrazione risulta particolarmente semplice se per dimostrare al convergenza uniforme si passa da quella totale

Definition 7 (Convergenza totale) $\sum_{j=0}^{\infty} A_n(t)$ converge totalmente se $\sum_{j=0}^{\infty} d(A_n(t), 0)$ (serie a coefficienti reali) converge.

Theorem 4 (Convergenza totale e convergenza) Se una serie converge totalmente allora converge.

Proposition 2 La serie dell'esponenziale di matrice $\sum_{j=0}^{\infty} \frac{t^j}{j!} A^n$ converge uniformemente.

Proof 4 Dimostriamo che converge assolutamente

$$\sum_{j=0}^{\infty} d\left(\frac{t^j}{j!} A^n, 0\right) = \sum_{j=0}^{\infty} \max_{t \in [a, b]} \frac{|t^j|}{j!} \|A^n\| \leq \sum_{j=0}^{\infty} \frac{(\max_{t \in [a, b]} |t|)^j}{j!} \|A\|^n = \sum_{j=0}^{\infty} \frac{(\delta \|A\|)^j}{j!} = e^{\delta \|A\|} \in \mathbb{R}$$

dove $\delta = (\max_{t \in [a, b]} |t|)$. Ecco dimostrato che la serie della distanza tende ad un reale, dunque converge totalmente, dunque converge assolutamente.

5.4 Passaggio al limite sotto al segno d'integrale

Theorem 5 (Passaggio al limite sotto al segno d'integrale per funzioni reali) Sia $f_n \in C([a, b]; \mathbb{R})$ e convergente uniformemente ad f , allora

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow +\infty} f_n(x) dx = \int_a^b f(x) dx$$

Proof 5 Dalla definizione di convergenza uniforme

$$\begin{aligned} \max_{x \in [a, b]} |f_n - f(x)| &\rightarrow 0 \\ \Leftrightarrow \forall \varepsilon > 0, \exists N > 0 : \forall n > N \max_{x \in [a, b]} |f_n(x) - f(x)| &\leq \frac{\varepsilon}{b-a} \\ \Rightarrow \forall \varepsilon > 0, \exists N > 0 : \forall n > N, \forall x \in [a, b] |f_n(x) - f(x)| &\leq \frac{\varepsilon}{b-a} \end{aligned}$$

Integrando:

$$\begin{aligned} \int_a^b |f_n(x) - f(x)| dx &\leq \int_a^b \frac{\varepsilon}{b-a} dx = \varepsilon \\ \Rightarrow \left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| &= \left| \int_a^b (f_n(x) - f(x)) dx \right| \leq \int_a^b |f_n(x) - f(x)| dx \leq \varepsilon \end{aligned}$$

ma questa è la definizione di limite

$$\forall \varepsilon > 0, \exists N > 0 : \forall n > N \left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| \leq \varepsilon \Leftrightarrow \lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx$$

Proposition 3 La funzione $F : C([a, b]; \mathbb{R}) \rightarrow \mathbb{R}$

$$F(f) = \int_a^b f(x) dx$$

è lineare e continua

Proof 6 Per la linearità dell'integrale, F è lineare, per il passaggio al limite sotto al segno d'integrale si ha

$$\lim_{n \rightarrow +\infty} F(f_n(x)) = \lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow +\infty} f_n(x) dx = F(f(x))$$

5.5 Estensione passaggio al limite sotto al segno d'integrale per funzioni a valori matriciali

Come spesso accade, notando che l'integrale della matrice equivale ad integrare ognuna delle sue componenti, che sono funzioni reali, e che il limite della matrice equivale a fare il limite di ognuna delle sue componenti, ci si riconduce al caso di funzioni reali.

5.6 Teorema intermedio

Theorem 6 Sia $f_n \in C^1([a, b]; \mathbb{R})$; se f_n converge uniformemente ad f e Df_n converge uniformemente a g , allora:

$$\begin{cases} f \in C^1([a, b]; \mathbb{R}) \\ g = Df \end{cases}$$

Euristicamente, possiamo immaginare due funzioni che convergono uniformemente come se, isolata una striscia di piano arbitrariamente stretta, queste due ne fossero sempre comprese all'interno e che quindi combaciano.

Proof 7 Se f_n continua tende ad f allora f è continua. Infatti scrivendo la definizione di convergenza uniforme

$$\begin{aligned} \forall \varepsilon > 0, \exists N : n > N \max_{x \in [a, b]} |f_n(x) - f(x)| < \varepsilon \\ \Rightarrow \forall \varepsilon > 0, \exists N : n > N \forall x \in [a, b], |f_n(x) - f(x)| < \varepsilon \\ \lim_{n \rightarrow +\infty} \lim_{x \rightarrow x_0} f_n(x) = \lim_{n \rightarrow +\infty} f_n(x_0) = f(x_0) \end{aligned}$$

Analogamente, se Df_n continua tende a g allora g continua. Abbiamo così dimostrato il primo punto.

Dal teorema fondamentale del calcolo

$$\begin{aligned} f_n(x) &= f_n(a) + \int_a^x Df_n(t) dt \\ \Rightarrow f(x) &= \lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} f_n(a) + \int_a^x \lim_{n \rightarrow +\infty} Df_n(t) dt = f(a) + \int_a^x g(t) dt \end{aligned}$$

Derivando otteniamo

$$\begin{aligned} Df(x) &= Df(a) + D \int_a^x g(t) dt = g(x) \\ \Rightarrow Df(x) &= g(x) \end{aligned}$$

5.7 Entrare la derivata nella serie

Theorem 7 Sia $f_n \in C^1([a, b]; \mathbb{R})$ e $\sum_{j=1}^{\infty} f_n$ converge uniformemente a g e $\sum_{j=1}^{\infty} Df_n$ converge uniformemente ad h , allora

$$h = Dg$$

Si noti che, sostituendo ciò a cui convergono g ed h si ottiene proprio la formula del passaggio della derivata sotto al segno di serie

$$\sum_{j=1}^{\infty} Df_n = D \sum_{j=1}^{\infty} f_n$$

Proof 8 Si noti che se la serie non fosse infinita il teorema sarebbe dimostrato grazie alla linearità della derivata per cui la derivata della somma è uguale alla somma delle derivate. Essendoci il simbolo d'infinito il problema richiede una trattazione più dettagliata. La strategia è quella di ricondurci al teorema precedente.

Sappiamo che se la successione parziale s_n della serie è convergente allora lo è anche la serie; sappiamo per ipotesi che

$$\begin{aligned} s_n(x) &= \sum_{j=1}^n f_j(x) \in C^1 \quad s_n(x) \rightarrow g \\ Ds_n(x) &= \sum_{j=1}^n Df_j(x) \rightarrow h \end{aligned}$$

Siamo nelle condizioni del teorema precedente dunque

$$\Rightarrow \begin{cases} g \in C^1 \\ h = Dg \end{cases}$$

5.8 La serie esponenziale è continua ed ha derivata continua

Proposition 4 La serie $e^{tA} \in C^1(\mathbb{R}; \mathbb{M}_n)$ e $\frac{d}{dt} e^{tA} = Ae^{tA}$

Proof 9

$$\begin{aligned}
& \sum_{j=0}^{\infty} \left(\frac{t^j}{j!} A^j \right) \in C^\infty(\mathbb{R}; \mathbb{M}_n) \\
& \Rightarrow \frac{d}{dt} \sum_{j=0}^{\infty} \left(\frac{t^j}{j!} A^j \right) = \sum_{j=0}^{\infty} \frac{d}{dt} \left(\frac{t^j}{j!} A^j \right) = \\
& \sum_{j=0}^{\infty} j \frac{t^{j-1}}{j!} A^j = 0 + \sum_{j=1}^{\infty} \frac{t^{j-1}}{(j-1)!} A^j = \\
& A \sum_{j=1}^{\infty} \frac{t^{j-1}}{(j-1)!} A^{j-1} = A e^{tA}
\end{aligned}$$

Una volta dimostrato ciò non resta che dimostrare la convergenza della derivata per poter dimostrare il secondo punto. Dimostriamo la convergenza totale, da cui segue quella uniforme

$$\sum_{j=0}^{\infty} d \left(\frac{t^{j-1}}{(j-1)!} A^j \right) = \sum_{j=0}^{\infty} (\max_{t \in [a, b]} (t))^{j-1} \frac{\|A^j\|}{(j-1)!} = \|A\| \sum_{j=0}^{\infty} \delta^{j-1} \frac{\|A^{j-1}\|}{(j-1)!} = \|A\| e^{\delta \|A\|}$$

Visto che il modulo di A e δ sono dei numeri reali, la derivata della serie di potenza esiste e converge uniformemente. Inoltre, per il teorema della sezione (5.6), visto che sia la serie dell'esponenziale di matrice sia la sua derivata sono continue, allora anche la funzione e^{tA} a cui converge uniformemente e la sua derivata lo saranno: $e^{tA} \in C^1([a, b]; \mathbb{R})$.

6 Teoremi di esistenza e unicità

Un teorema che ci garantisca l'esistenza di soluzioni di un'equazione differenziale è di estrema importanza in ambito fisico: la descrizione dell'evoluzione nel tempo di sistemi fisici è svolta spesso mediante la risoluzione di equazioni differenziali (ad esempio a partire dalla seconda legge della dinamica in meccanica) conoscendo delle condizioni iniziali, che equivale alla risoluzione di un problema di Cauchy. Ad esempio, si ha un'equazione differenziale della forma

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

dove la funzione f è definita in $T = [a, b] \times [\alpha_1, \beta_1] \times \dots \times [\alpha_{n-1}, \beta_{n-1}]$ e si conoscono le condizioni iniziali

$$x_0, y_0, y'_0, \dots, y_0^{(n-1)} \in T$$

e si cerca una soluzione y tale che

$$\begin{cases} y^{(n)} = f(x, y, y', \dots, y^{(n-1)}) \\ y(x_0) = y_0; \dots; y^{(n-1)}(x_0) = y_0^{(n-1)} \end{cases} \quad (2)$$

e ci chiediamo se esista sempre una soluzione.

Cominciamo con l'enunciare (senza dimostrazione) il teorema di esistenza secondo Peano.

Theorem 8 (Teorema di esistenza) Se $f \in C(T; \mathbb{R})$, $\forall x_0, y_0, y'_0, \dots, y_0^{(n-1)} \in T, \exists I_{x_0, y} \in C^n(I; \mathbb{R})$ che soddisfa il sistema.

Si osservi che y è una funzione continua in n variabili che in questo caso equivale a

$$\lim_{(x, y, \dots, y^{(n-1)}) \rightarrow (x_0, y_0, \dots, y_0^{(n-1)})} f(x, y, \dots, y^{(n-1)}) = f(x_0, y_0, \dots, y_0^{(n-1)})$$

Non resta che capire cosa sia la convergenza di un vettore ad un altro

$$\begin{aligned}
& \forall \varepsilon > 0, \exists \delta \in T : (x, y, \dots, y^{(n-1)}) \in T, d \left((x, y, \dots, y^{(n-1)}), (x_0, y_0, \dots, y_0^{(n-1)}) \right) < \delta \\
& \Rightarrow |f(x, y, \dots, y^{(n-1)}) - f(x_0, y_0, \dots, y_0^{(n-1)})| < \varepsilon
\end{aligned}$$

Dove d è una distanza in T .

Tuttavia questo teorema non basta: necessitiamo di un teorema che non garantisca solo l'esistenza ma anche l'unicità. Ad esempio, nella risoluzione di un'equazione differenziale che descrive un moto in fisica vorremmo la garanzia che la traiettoria ottenuta dalla risoluzione del problema sia l'unica possibile altrimenti non avremmo alcuna capacità di predire il comportamento di sistemi fisici. Necessitiamo dunque di un teorema più forte: un **teorema di esistenza ed unicità** della soluzione di equazioni differenziali.

Per giungere a questo risultato bisogna prima passare per alcuni step intermedi.

Definition 8 (Contrazione) Un operatore $K : X \rightarrow X$ è una contrazione se

$$\forall x, y \in X, K(x), K(y) \in X \exists \alpha \in]0, 1[: d(K(x), K(y)) \leq \alpha d(x, y)$$

dove d è una distanza in X rispetto alla quale X è completo.

Theorem 9 (Teorema delle contrazioni (Banach-Caccioppoli)) Sia (X, d) uno spazio metrico completo e $K : X \rightarrow X$ una contrazione. Allora $\exists \bar{x} \in X : K(\bar{x}) = \bar{x}$ dove \bar{x} è detto **punto fisso**.

Proof 10 1. *Unicità*

Sia $\bar{x} : K(\bar{x}) = \bar{x}$, ammettiamo per assurdo che esista $\bar{y} : K(\bar{y}) = \bar{y}$. Allora, visto che K è una contrazione

$$d(K(\bar{x}), K(\bar{y})) \leq \alpha d(\bar{x}, \bar{y}) \quad \alpha \in]0, 1[$$

Ma essendo \bar{x}, \bar{y} punti fissi per ipotesi

$$d(\bar{x}, \bar{y}) \leq \alpha d(\bar{x}, \bar{y}) \quad \alpha \in]0, 1[$$

Dunque se $\bar{x} \neq \bar{y}$ si perviene ad un assurdo, l'unico risultato possibile è $\bar{x} = \bar{y} \Rightarrow d(\bar{x}, \bar{y}) = 0$

2. *Esistenza*

Fisso $x_0 \in X$ a piacere

$$x_1 = K(x_0)$$

$$x_2 = K(x_1) = K(K(x_0)) = K^2(x_0)$$

...

$$x_{n+1} = K(x_n) = K^{n+1}(x_0)$$

abbiamo costruito una successione x_n . Se questa fosse di Cauchy, visto che X è completo rispetto a d , la successione convergerebbe di modo che

$$\begin{aligned} \forall \varepsilon > 0, \exists N : \forall n, m > N, d(x_n - x_m) < \varepsilon \\ \Leftrightarrow \exists \bar{x} \in X : \lim_{n \rightarrow +\infty} x_n = \bar{x} \end{aligned}$$

Inoltre, ricordando le proprietà delle contrazioni

$$\begin{aligned} d(K(x_n), K(\bar{x})) &\leq \alpha d(x_n - \bar{x}) \\ \lim_{n, m \rightarrow +\infty} d(K(x_n), K(\bar{x})) &\leq \lim_{n \rightarrow +\infty} \alpha d(x_n - \bar{x}) \\ \lim_{n, m \rightarrow +\infty} d(K(x_n), K(\bar{x})) &= 0 \\ \Rightarrow \lim_{n, m \rightarrow +\infty} K(x_n) &= K(\bar{x}) \end{aligned}$$

Mettendo insieme quanto ricavato otteniamo la conclusione

$$x_{n+1} = K(x_n) \rightarrow \bar{x} = K(\bar{x})$$

Che è la definizione di punto fisso.

Non ci resta che dimostrare che la successione è di Cauchy. Sia $n < m$ per fissare le idee, per la triangolare della distanza

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_m) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + d(x_{n+2}, x_m) \leq \\ &\dots \leq d(x_n, x_{n+1}) + \dots + d(x_{m-1}, x_m) = \sum_{j=n}^{m-1} d(x_j, x_{j+1}) \end{aligned}$$

Ricordando come abbiamo costruito la successione x_n

$$\begin{aligned}
d(x_j, x_{j+1}) &= d(K^j(x_0), K^{j+1}(x_0)) = d(K(K^{j-1}(x_0)), K(K^j(x_0))) \leq \\
&\leq \alpha d(K^{j-1}(x_0), K^j(x_0)) = \alpha d(K(K^{j-2}(x_0)), K(K^{j-1}(x_0))) \leq \\
&\leq \alpha^2 d(K^{j-2}(x_0), K^{j-1}(x_0)) \leq \dots \leq \alpha^j d(x_0, K(x_0)) \\
&\Rightarrow d(x_j, x_{j+1}) \leq \alpha^j d(x_0, K(x_0)) \\
&\Rightarrow d(x_n, x_m) \leq \sum_{j=n}^{m-1} d(x_j, x_{j+1}) \leq d(x_0, K(x_0)) \sum_{j=n}^{m-1} \alpha^j
\end{aligned}$$

Ne segue che la successione è di Cauchy se $\lim_{n,m \rightarrow +\infty} \sum_{j=n}^{m-1} \alpha^j$ converge

$$\begin{aligned}
\sum_{j=n}^{m-1} \alpha^j &= \sum_{j=0}^{m-1} \alpha^j - \sum_{j=0}^{n-1} \alpha^j \\
\lim_{n,m \rightarrow +\infty} \sum_{j=n}^{m-1} \alpha^j &= \lim_{m \rightarrow +\infty} \sum_{j=0}^{m-1} \alpha^j - \lim_{n \rightarrow +\infty} \sum_{j=0}^{n-1} \alpha^j = \sum_{j=0}^{\infty} \alpha^j - \sum_{j=0}^{\infty} \alpha^j
\end{aligned}$$

ma queste ultime sono serie geometriche che convergono visto che $0 < \alpha < 1$ per ipotesi. La successione è di Cauchy, dunque converge, allora esiste sempre un punto fisso.

Un modo rapido per notare che un punto fisso dell'integrale è soluzione dell'equazione differenziale è quello di svolgere le contrazioni (l'integrale), data un'equazione differenziale con un dato di Cauchy, prendendo come base un punto y_0 .

$$\begin{aligned}
&\begin{cases} y'(x) = Ay(x) & A \in \mathbb{R} \\ y(0) = y_0 \end{cases} \\
y(x) &= y_0 + \int_0^x Ay(t)dt \\
y_1(x) &= K(y_1)(x) = y_0 + \int_0^x Ay_0 dt = y_0 + Ay_0 x \\
y_2(x) &= K(y_2)(x) = y_0 + \int_0^x Ay_1(t)dt = y_0 + \int_0^x A(y_0 + Ay_0 t)dt = \\
&y_0 + y_0 Ax + A^2 y_0 \frac{x^2}{2} = (1 + Ax + A^2 \frac{x^2}{2})y_0 \\
&\dots \\
y_n(x) &= \left(\sum_{j=0}^n \frac{x^j}{j!} A^j \right) y_0
\end{aligned}$$

Notiamo che le ripetute contrazioni approssimano, mediante uno sviluppo di Taylor, la soluzione $e^{xA} \cdot y_0$, che è la soluzione che sappiamo dovrebbe risultare. Più contrazioni svolgiamo più ci avviciniamo alla soluzione.

Come visto in precedenza, un'equazione differenziale può essere trasformata in un sistema ed essere espressa in forma vettoriale nel seguente modo:

$$\dot{\vec{x}} = G(\vec{x}, t) \quad (3)$$

$$\vec{x}_0 = (y_0, \dots, y_0^{(n-1)}) \quad (4)$$

dove $G : T \rightarrow [\alpha_1] \times \dots \times [\alpha_{n-1}, \beta_{n-1}]$. Abbiamo ridotto il problema alla dimostrazione che esiste un vettore \vec{x} che soddisfi queste condizioni. Abbiamo visto che unicamente con l'ipotesi di continuità di f avevamo il teorema di esistenza, per avere anche l'unicità bisogna rafforzare le ipotesi assumendo la lipschitzianità di f .

Definition 9 (Lipschitzianità) Una funzione $f : [a, b] \rightarrow \mathbb{R}$ si dice lipschitziana se

$$\exists L > 0 \in \mathbb{R} : |f(x_1) - f(x_2)| \leq L|x_1 - x_2|, \forall x_1, x_2 \in [a, b]$$

Questo concetto si estende a spazi n dimensionali (come lo spazio T del nostro problema) nel modo segue

$$\exists L > 0 \in \mathbb{R} : |f(x_1, \dots, x_n, y) - f(x'_1, \dots, x'_n, y')| \leq L|y - y'|, \forall x_1, x_2 \in [a, b]$$

Si noti che la lipschitzianità implica la continuità, dunque le condizioni sono più restrittive di quelle del teorema di esistenza secondo Peano. In fisica le funzioni che si usano soddisfano spesso questo requisito.

Possiamo dimostrare il teorema di esistenza ed unicità per funzioni che vanno da T ad \mathbb{R} per semplicità, ma l'estensione ad \mathbb{R}^n è immediata.

Theorem 10 (Teorema di esistenza ed unicità) Sia $f \in C(T; \mathbb{R})$ lipschitziana tale che

$$\exists L > 0 : |f(x_0, z) - f(x_0, \bar{z})| \leq L||z - \bar{z}||, \forall (x_0, z), (x_0, \bar{z}) \in T$$

allora il sistema 2 ammette un'unica soluzione ovvero:

se $y' = f(x, y)$ e $y(x_0) = y_0$ la soluzione è $y(x) = y_0 + \int_{x_0}^x f(t, y(t))dt$ e, viceversa, se y è continua e soddisfa la precedente uguaglianza allora è soluzione dell'equazione differenziale.

Proof 11 La strategia di dimostrazione è quella di passare dall'equazione differenziale a quella integrale per poi usare il teorema delle contrazioni.

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases} \Rightarrow y(x) = y_0 + \int_{x_0}^x f(t, y(t))dt$$

Dove possiamo vedere come una contrazione

$$K(f) = y_0 + \int_{x_0}^x f(t, y(t))dt$$

Bisogna quindi dimostrare che l'integrale è una contrazione e che porta funzioni continue in funzioni continue. Una volta fatto questo segue per il teorema delle contrazioni che esiste un punto fisso, cioè una soluzione dell'equazione differenziale.

1. La contrazione trasforma funzioni continue in funzioni continue

$$\Leftrightarrow K : f \in C([\alpha, \beta], \mathbb{R}) \rightarrow g \in C([\alpha, \beta], \mathbb{R})$$

Per ipotesi f è continua, l'integrale porta funzioni continue in funzioni continue. Aggiungendo una costante y_0 la continuità si conserva.

2. K è una contrazione $\Leftrightarrow \forall y, z \in C([\alpha, \beta], \mathbb{R}), d(K(y), K(z)) \leq \alpha d(y, z) \alpha \in]0, 1[$ Dove la distanza fra funzioni è definita come fatto in precedenza.

Fissate $y, z \in C([\alpha, \beta], \mathbb{R})$

$$\begin{aligned} K(y)(x) - K(z)(x) &= \int_{x_0}^x (f(t, y(t)) - f(t, z(t))) dt \\ |K(y)(x) - K(z)(x)| &\leq \int_{x_0}^x |(f(t, y(t)) - f(t, z(t)))| dt \leq \int_{x_0}^x L|y(t) - z(t)| dt \end{aligned}$$

Supponiamo $x \in I = [x_0 - \delta, x_0 + \delta]$

$$\begin{aligned} \max_{x \in I} |K(y)(x) - K(z)(x)| &\leq \max_{x \in I} \int_{x_0}^x L|y(t) - z(t)| dt \leq \\ &\leq \max_{t \in I} |y(t) - z(t)| \cdot \max_{x \in I} (x - x_0) \cdot L = \max_{t \in I} |y(t) - z(t)| \delta \end{aligned}$$

Se è scelto in modo che $\delta \in]0, \frac{1}{L}[$ allora $\delta \cdot L < 1$ l'integrale è una contrazione. Infine, per il teorema delle contrazioni

$$\exists! y \in C^1([x_0 - \delta, x_0 + \delta]) : x \in [x_0 - \delta, x_0 + \delta], \begin{cases} y'(x) = f(x, y(x)) \\ y(x_0) = y_0 \end{cases}$$

dove f è la contrazione ed y il punto fisso (ovvero la soluzione dell'equazione differenziale).

Si noti che l'intervallo $[a, b]$ di partenza è stato ridotto ad $[x_0 - \delta, x_0 + \delta]$, per ottenere un teorema di esistenza ed unicità su tutto l'intervallo dobbiamo estendere quanto dimostrato.

Ipotizziamo che $a = x_0 - \delta$, vogliamo estendere l'intervallo in cui è garantita l'esistenza e l'unicità fino a b . Sappiamo che y è una soluzione generale per tutto l'intervallo $[x_0 - \delta, x_0 + \delta]$, calcoliamo y in $x_0 + \delta$, questo sarà il nostro nuovo dato di Cauchy. Consideriamo ora l'intervallo $[x_0, x_0 + 2\delta]$ ed impostiamo il problema di Cauchy

$$\begin{cases} y'(x) = f(x, y(x)) \\ y(x_0 + \delta) = y_1 \end{cases}$$

Per il teorema di esistenza ed unicità avremo che esiste ed è unica la soluzione dell'equazione differenziale nell'intervallo $[x_0, x_0 + 2\delta]$ ma il dato di Cauchy è quello della precedente equazione dunque la y ottenuta sarà la medesima; abbiamo allargato l'intervallo di δ . Possiamo reiterare il processo fin quando $x_0 + n\delta = b$.

Nel caso generale del teorema di esistenza ed unicità f non è scalare ma vettoriale dunque $f : [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$. Le uniche accortezze da tenere sono quella di ridefinire la lipschitzianità per vettori

$$\|f(x_0, y) - f(x_0, z)\| \leq L\|y - z\|$$

dove è necessario sostituire il valore assoluto (modulo in \mathbb{R}) con un modulo adatto ai vettori di \mathbb{R}^n .

7 Teorema di dipendenza continua dalle condizioni iniziali

Infine, ci aspettiamo che cambiando di poco le condizioni iniziali di un sistema fisico (i dati di Cauchy), l'evoluzione dello stato non è estremamente diversa. Vogliamo dunque un teorema che ci garantisca la dipendenza continua dalle condizioni iniziali (se queste saranno vicine sarà vicina anche la soluzione).

Theorem 11 (Teorema di dipendenza continua dalle condizioni iniziali) Sia $f : [a, b] \times \mathbb{R} \rightarrow \mathbb{R}$ continua. Assumiamo che questa goda di lipschitzianità globale, ovvero

$$\exists L > 0 : |f(x, y_1) - f(x, y_2)| \leq L|y_2 - y_1| \quad \forall x \in [a, b], \quad \forall y_1, y_2 \in \mathbb{R}$$

allora se $y'_j(x) = f(x, y_j(x))$; $y_j(x_0) = y_j$ con $j = 1, 2$,

$$\exists c > 0 : \max_{x \in [a, b]} |y_2(x) - y_1(x)| \leq c|y_2 - y_1|$$

dove c si può stimare come $c \leq e^{L(b-a)}$. Dunque se le condizioni iniziali sono vicine anche le soluzioni lo saranno.

Proof 12 L'idea che segue la dimostrazione è quella di confrontare le soluzioni mediante una disequazione differenziale. Innanzitutto bisogna introdurre una funzione che calcoli la distanza fra le due soluzioni che sia derivabile

$$\begin{aligned} g(x) &= (y_2(x) - y_1(x))^2 \\ g'(x) &= 2(y_2(x) - y_1(x))(y'_2(x) - y'_1(x)) = \\ &= 2(y_2(x) - y_1(x))(f(x, y_2(x)) - f(x, y_1(x))) \leq 2L(y_2(x) - y_1(x))^2 = 2Lg(x) \\ &\Rightarrow \begin{cases} g'(x) \leq 2Lg(x) \\ g(x_0) = (y_2 - y_1)^2 \end{cases} \end{aligned}$$

Come trattare una disequazione differenziale del genere? Possiamo definire una nuova funzione $h(x)$ tale che

$$\begin{aligned} h'(x) &\equiv 2Lh(x) \quad h(x_0) \equiv g(x_0) \\ &\Rightarrow h(x) = e^{2L(x-x_0)}h(x_0) \\ &\Rightarrow h'(x) \geq g'(x) \end{aligned}$$

Ora, interpretando $g'(x)$ ed $h'(x)$ come velocità e $h(x_0)$, $g(x_0)$ come punti di partenza, notiamo che partendo dallo stesso punto, h' è più o ugualmente veloce di g' dunque h sarà sempre maggiore o uguale a g . Sostituendo

$$g(x) \leq h(x)$$

$$(y_2(x) - y_1(x))^2 \leq e^{2L(x-x_0)}(y_2 - y_1)^2$$

Fissati x, x_0 , in qualunque modo li prendiamo in $[a, b]$ abbiamo che $|x - x_0| \leq b - a$

$$\Rightarrow (y_2(x) - y_1(x))^2 \leq e^{2L(b-a)}(y_2 - y_1)^2$$

$$|(y_2(x) - y_1(x))| \leq e^{L(b-a)}|y_2 - y_1|$$

$$\max_{x \in [a, b]} |(y_2(x) - y_1(x))| = d(y_2(x), y_1(x)) = c \cdot |y_2 - y_1|$$

$$c \equiv e^{L(b-a)}$$