

CAPITOLO VIII

EQUAZIONI DIFFERENZIALI

Consideriamo il seguente problema di Cauchy per i sistemi di equazioni differenziali del primo ordine :

$$\begin{aligned} y'(t) &= f(t,y(t)) \\ y(t_0) &= y_0 \end{aligned} \quad (8.1)$$

dove $f(t,y): [t_0, t_f] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$. La funzione $f(t,y)$ è supposta continua rispetto a t , e lipschitziana rispetto ad y nella striscia illimitata $[t_0, t_f] \times \mathbb{R}^m$,

$$\|f(t,u)-f(t,v)\| < L\|u-v\|, \quad \forall t \in [t_0, t_f] \text{ e } \forall u,v \in \mathbb{R}^m.$$

In tali condizioni è garantita l'esistenza e l'unicità della soluzione nell'intero intervallo di integrazione $[t_0, t_f]$.

Consideriamo una discretizzazione dell'intervallo $[t_0, t_f]$ che, per semplicità di esposizione, supporremo uniforme:

$$t_0 < t_1 < \dots < t_N (=t_f) \quad h = \frac{t_f - t_0}{N}.$$

e, per ogni intervallo $[t_n, t_{n+1}]$, consideriamo l'identità:

$$\begin{aligned} \int_{t_n}^{t_{n+1}} y'(t) dt &= \int_{t_n}^{t_{n+1}} f(t,y(t)) dt \\ y(t_{n+1}) &= y(t_n) + \int_{t_n}^{t_{n+1}} f(t,y(t)) dt. \end{aligned}$$

Ogni formula di quadratura che fa uso dei valori nodali di $y(t)$ può essere utilizzata per creare una formula di integrazione numerica per il problema (8.1). Limitiamoci a considerare alcune formule, trattate nel capitolo precedente, che fanno uso di uno o di entrambi gli estremi dell'integrale. In particolare:

$$\begin{aligned} 1) \quad \int_{t_n}^{t_{n+1}} f(t,y(t)) dt &= hf(t_n, y(t_n)) + \sigma_1(t_n, h) \\ \sigma_1(t_n, h) &= \frac{1}{2} \frac{\partial}{\partial t} f(\xi_n, y(\xi_n)) h^2 = \frac{1}{2} y''(\xi_n) h^2 \quad \xi_n \in (t_n, t_{n+1}) \end{aligned}$$

$$2) \quad \int_{t_n}^{t_{n+1}} f(t, y(t)) dt = hf(t_{n+1}, y(t_{n+1})) + \sigma_2(t_n, h)$$

$$\sigma_2(t_n, h) = -\frac{1}{2} \frac{\partial}{\partial t} f(\xi_n, y(\xi_n)) h^2 = -\frac{1}{2} y''(\xi_n) h^2 \quad \xi_n \in (t_n, t_{n+1})$$

$$3) \quad \int_{t_n}^{t_{n+1}} f(t, y(t)) dt = \frac{1}{2} h (f(t_n, y(t_n)) + f(t_{n+1}, y(t_{n+1}))) + \sigma_3(t_n, h)$$

$$\sigma_3(t_n, h) = -\frac{1}{12} \frac{\partial^2}{\partial t^2} f(\xi_n, y(\xi_n)) h^3 = -\frac{1}{12} y'''(\xi_n) h^3 \quad \xi_n \in (t_n, t_{n+1})$$

Otteniamo così le relazioni:

$$1') \quad y(t_{n+1}) = y(t_n) + hf(t_n, y(t_n)) + \sigma_1(t_n, h)$$

$$2') \quad y(t_{n+1}) = y(t_n) + hf(t_{n+1}, y(t_{n+1})) + \sigma_2(t_n, h)$$

$$3') \quad y(t_{n+1}) = y(t_n) + \frac{1}{2} h (f(t_n, y(t_n)) + f(t_{n+1}, y(t_{n+1}))) + \sigma_3(t_n, h).$$

Trascurando ad ogni passo l'errore $\sigma(t_n, h)$, detto **errore locale di troncamento**, si ottengono le formule ricorsive:

formula di **Eulero Esplicita** $y_{n+1} = y_n + hf(t_n, y_n)$

formula di **Eulero Implicita** $y_{n+1} = y_n + hf(t_{n+1}, y_{n+1})$

formula dei **Trapezi** $y_{n+1} = y_n + \frac{1}{2} h (f(t_n, y_n) + f(t_{n+1}, y_{n+1})).$

dove, per ogni n , y_n è l'approssimazione della soluzione nel punto $t_n = t_0 + nh$ ed y_0 è il valore iniziale assegnato nel problema (8.1).

Si osservi che le formule di Eulero Implicita e dei trapezi presentano una maggiore complessità computazionale, rispetto alla formula di Eulero esplicita, poiché l'incognita y_{n+1} si presenta come la risoluzione di un sistema di equazioni, in generale, non lineari.

Per comodità di trattazione, esprimiamo le precedenti formule nella forma generale:

$$y_{n+1} = y_n + h\Phi(t_n, y_n, y_{n+1}) \quad (8.2)$$

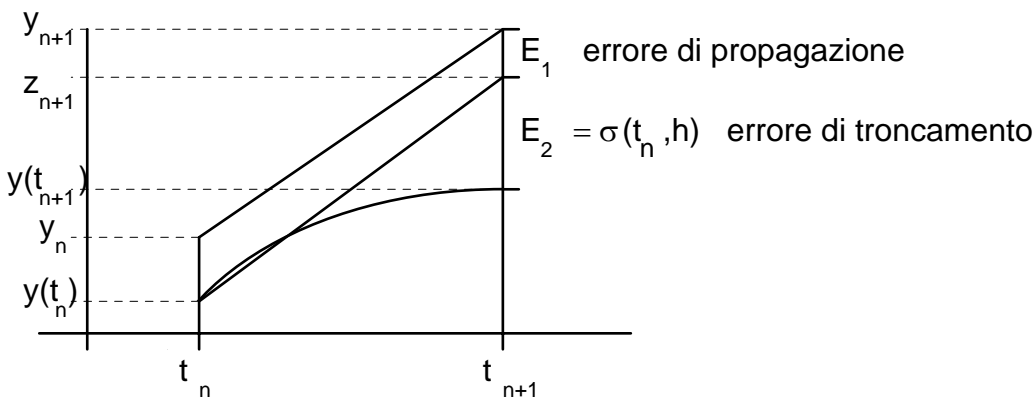
per ciascuna delle quali è immediato verificare la lipschitzianità

$$|\Phi(t, u, v) - \Phi(t, w, z)| < M(|u-w| + |v-z|)$$

come conseguenza della lipchitzianità di f .

Detto $e_n := y_n - y(t_n)$ l'errore accumulato fino al passo n -esimo di integrazione, analizziamo come esso si propaga nel passo successivo. A tale scopo indichiamo con z_{n+1} il valore fornito dalla formula (8.2) qualora essa fosse applicata al punto $y(t_n)$ della traiettoria esatta:

$$z_{n+1} = y(t_n) + h\Phi(t_n, y(t_n), z_{n+1})$$



L'errore totale al passo $n+1$ è quindi dato da:

$$e_{n+1} = y_{n+1} - y(t_{n+1}) = E_1 + E_2 = (y_{n+1} - z_{n+1}) + \sigma(t_n, h)$$

$$\|e_{n+1}\| \leq \|y_{n+1} - z_{n+1}\| + \|\sigma(t_n, h)\| \quad (8.3)$$

dove:

$$y_{n+1} - z_{n+1} = y_n + h\Phi(t_n, y_n, y_{n+1}) - (y(t_n) + h\Phi(t_n, y(t_n), z_{n+1}))$$

$$y_{n+1} - z_{n+1} = e_n + h(\Phi(t_n, y_n, y_{n+1}) - \Phi(t_n, y(t_n), z_{n+1}))$$

$$\|y_{n+1} - z_{n+1}\| \leq \|e_n\| + hM(\|e_n\| + \|y_{n+1} - z_{n+1}\|)$$

$$(1 - hM) \|y_{n+1} - z_{n+1}\| < (1 + hM) \|e_n\|$$

e, per h sufficientemente piccolo,

$$\|y_{n+1} - z_{n+1}\| < \frac{1 + hM}{1 - hM} \|e_n\|.$$

Tornando alla (8.3) si ottiene quindi:

$$\|e_{n+1}\| < \frac{1+hM}{1-hM} \|e_n\| + \|\sigma(t_n, h)\|.$$

Osservato inoltre che $\frac{1+hM}{1-hM} < 1+3hM$, si ha:

$$\|e_{n+1}\| < (1+3hM) \|e_n\| + \|\sigma(t_n, h)\|$$

Maggiorando infine l'errore locale di troncamento $\|\sigma(t_n, h)\|$ in modo uniforme sull'intervallo di integrazione $[t_0, t_f]$

$$\sigma(h) := \max_{t \in [t_0, t_n]} \|\sigma(t, h)\|$$

si ottiene la seguente relazione ricorsiva per l'errore:

$$\|e_{n+1}\| < (1+3hM) \|e_n\| + \sigma(h), \quad n=0,1,\dots,N-1. \quad (8.4)$$

Lemma: Se la successione $\{a_n\}$, $a_n > 0$, soddisfa la relazione ricorsiva

$$a_{n+1} < (1+hQ)a_n + c(h) \quad n=0,1,2,\dots,N$$

con $(1+hQ) > 0$, allora vale la maggiorazione:

$$a_m < (1+hQ)^N a_0 + c(h) \frac{(1+hQ)^N - 1}{hQ} \quad \forall m \leq N.$$

(La dimostrazione è lasciata come esercizio).

Applicando il lemma alla relazione ricorsiva (8.4), tenendo conto che $e_0=0$, si ottiene la maggiorazione:

$$\|e_m\| < \sigma(h) \frac{(1+3hM)^N - 1}{3hM} \quad \forall m \leq N$$

e, tenuto conto della disuguaglianza $(1+3hM) < e^{3hM}$,

$$\|e_m\| < \sigma(h) \frac{e^{3MhN} - 1}{3hM} \quad \forall m \leq N$$

Poichè il numero totale di passi N e l'ampiezza del passo h sono legati dalla relazione $Nh = (t_f - t_0)$, si ottiene

$$\|e_m\| < \sigma(h) \frac{e^{3M(t_f - t_0)} - 1}{3hM} \quad \forall m \leq N. \quad (8.5)$$

L'ultima relazione è fondamentale per l'analisi della convergenza del metodo.

Si dirà che il metodo (8.2) è **convergente** nell'intervallo d'integrazione $[t_0, t_f]$, se $\max_{m \leq N} \|e_m\| \rightarrow 0$ per $N \rightarrow \infty$ e $h \rightarrow 0$

ferma restando la relazione $Nh = (t_f - t_0)$. Si dirà inoltre che il metodo ha **ordine di convergenza** uguale a p se il termine $\max_{m \leq N} \|e_m\|$ è infinitesimo di ordine p .

Dalla relazione (8.5) si deduce immediatamente il seguente teorema di convergenza:

Teorema di convergenza. *Affinchè il metodo*

$$y_{n+1} = y_n + h\Phi(t_n, y_n, y_{n+1})$$

sia convergente di ordine p nell'intervallo $[t_0, t_f]$ è sufficiente che la funzione $\Phi(t, u, v)$ sia lipchitziana rispetto a u e v per ogni $t \in [t_0, t_f]$, e che il rapporto $\frac{\sigma(h)}{h}$ sia infinitesimo di ordine p .

Dalle espressioni dell'errore locale di troncamento si deduce che i metodi di Eulero esplicito ed implicito convergono con ordine $p=1$, mentre il metodo dei trapezi converge con ordine $p=2$.

Esistono molte altre formule del tipo considerato che sono convergenti con vari ordini. A titolo di esempio, una formula esplicita di ordine $p=2$ è la seguente

formula di **Eulero generalizzata**.

$$y_{n+1} = y_n + hf\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}hf(t_n, y_n)\right).$$

Essa deriva dalla relazione:

$$y(t_{n+1}) = y(t_n) + hf(t_n + \frac{1}{2}h, y(t_n) + \frac{1}{2}hf(t_n, y(t_n))) + \sigma(t_n, h) \quad (8.6)$$

nella quale l'errore di troncamento $\sigma(t_n, h)$ risulta essere un infinitesimo di ordine 3 rispetto ad h . Ciò si verifica facilmente sviluppando la funzione $f(t, y)$ in un intorno del punto $(t_n, y(t_n))$:

$$\begin{aligned} f(t_n + \frac{1}{2}h, y(t_n) + \frac{1}{2}hf(t_n, y(t_n))) &= \\ &= f(t_n, y(t_n)) + \frac{1}{2}h f_t(t_n, y(t_n)) + \frac{1}{2}hf(t_n, y(t_n)) f_y(t_n, y(t_n)) + O(h^2) \end{aligned}$$

e sviluppando pure $y(t)$ in un intorno di t_n

$$y(t_{n+1}) = y(t_n + h) = y(t_n) + hy'(t_n) + \frac{1}{2}h^2y''(t_n) + O(h^3).$$

Tenuto infine conto che $y'(t_n) = f(t_n, y(t_n))$, e che

$$y''(t_n) = \frac{\partial}{\partial t} f(t_n, y(t_n)) = f_t(t_n, y(t_n)) + f_y(t_n, y(t_n))y'(t_n),$$

dalla (8.5) si ricava $\sigma(t_n, h) = O(h^3)$.

Propagazione dell'errore e stabilità asintotica.

Abbiamo visto in precedenza per il problema iniziale (8.1), che ad ogni passo l'errore e_n si compone di due parti: l'errore propagato e l'errore di troncamento. Abbiamo altresì visto che gli errori si accumulano durante il processo di integrazione e la stima (8.5) ne rappresenta una limitazione uniforme su tutto l'intervallo $[t_0 - t_f]$. Di fatto, ad ogni passo, applichiamo la formula per risolvere l'equazione data con valore iniziale perturbato y_n anzichè $y(t_n)$.

Se, in particolare, ad ogni passo l'errore propagato $E_2 := y_{n+1} - z_{n+1}$ risulta minore dell'errore accumulato fino al passo precedente, cioè se:

$$\|y_{n+1} - z_{n+1}\| < \|e_n\|,$$

allora si dice che il metodo è **stabile**.

Per i metodi stabili la relazione (8.3) si può sviluppare nel seguente modo:

$$\|e_{n+1}\| \leq \|y_{n+1} - z_{n+1}\| + \|\sigma(t_n, h)\| < \|e_n\| + \|\sigma(t_n, h)\|$$

$$\begin{aligned}
&< \|e_{n-1}\| + \|\sigma(t_{n-1},h)\| + \|\sigma(t_n,h)\| < \dots \\
&< \|e_0\| + \|\sigma(t_0,h)\| + \|\sigma(t_1,h)\| + \dots + \|\sigma(t_n,h)\| \\
&= \|\sigma(t_0,h)\| + \|\sigma(t_1,h)\| + \dots + \|\sigma(t_n,h)\|.
\end{aligned}$$

Poichè, come abbiamo visto, $\sigma(t_k,h) < \sigma(h)$ per ogni k , allora si ottiene :

$$\|e_m\| < m \sigma(h) < N \sigma(h) = \sigma(h) \frac{t_f - t_0}{h} \quad \forall m \leq N.$$

Ciò significa che, per i metodi stabili, la crescita dell'errore è limitata in modo lineare rispetto all'intervallo $[t_0 - t_f]$ anziché in modo esponenziale come indicato dalla (8.5) per un metodo qualunque.

In generale è difficile verificare la stabilità dei metodi per equazioni qualunque, e pertanto ci limiteremo a studiare la stabilità per una classe molto particolare di equazioni test.

Consideriamo dapprima la seguente equazione scalare:

$$\begin{aligned}
y'(t) &= \lambda y(t) \\
y(0) &= 1
\end{aligned} \tag{8.6}$$

dove, per ragioni che vedremo tra poco, il coefficiente λ e la funzione y sono *complessi*.

E' noto che la soluzione è data dalla funzione $y(t) = e^{\lambda t}$.

Detto $\lambda = \alpha + i\beta$, si ottiene:

$$y(t) = e^{\lambda t} = e^{(\alpha + i\beta)t} = e^{\alpha t} (\cos \beta t + i \sin \beta t)$$

Il metodo di Eulero esplicito, applicato all'equazione test, è:

$$y_{n+1} = y_n + h\lambda y_n = (1 + h\lambda)y_n$$

ed il corrispondente valore z_{n+1} è dato da:

$$z_{n+1} = y(t_n) + h\lambda y(t_n) = (1 + h\lambda)y(t_n).$$

Si ha quindi, per l'errore propagato:

$$y_{n+1} - z_{n+1} = (1 + h\lambda) e_n.$$

$$\|y_{n+1} - z_{n+1}\| = |(1+h\lambda)| \|e_n\|.$$

In base alla definizione precedente, si osserva che il metodo è stabile per quei valori complessi del prodotto $h\lambda$ per i quali si ha:

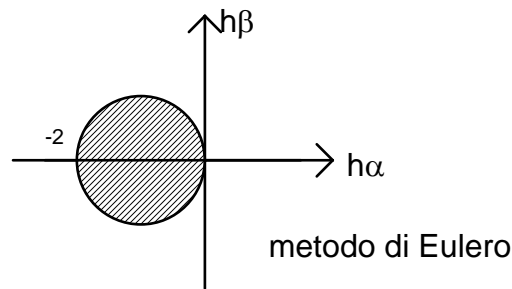
$$|(1+h\lambda)| < 1.$$

La funzione $\varphi(h\lambda) := (1+h\lambda)$ è detta **funzione di stabilità** e varia da metodo a metodo. La regione del piano complesso nella quale si ha:

$$|\varphi(h\lambda)| < 1$$

è detta **regione di assoluta stabilità** del metodo.

Per il metodo di Eulero la regione di assoluta stabilità è tratteggiata nella seguente figura:



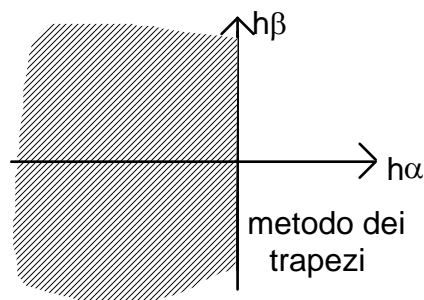
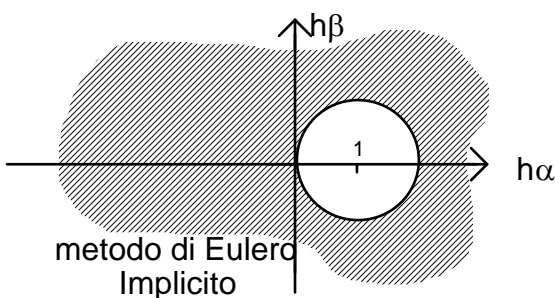
In maniera analoga si trovano le funzioni di stabilità:

$$\varphi(h\lambda) = \frac{1}{1-h\lambda} \quad \text{per il metodo di Eulero implicito}$$

e

$$\varphi(h\lambda) = \frac{1 + \frac{h\lambda}{2}}{1 - \frac{h\lambda}{2}} \quad \text{per il metodo dei trapezi}$$

alle quali corrispondono le seguenti regioni di assoluta stabilità:



Per quanto riguarda l'andamento asintotico della soluzione, osserviamo che, se $\alpha < 0$, allora la soluzione

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

tende a zero per t che tende a infinito. In altre parole la componente reale e immaginaria di $y(t)$ tendono entrambe a zero. In questo caso si dice che la soluzione è **asintoticamente stabile**.

Se invece $\alpha = 0$, allora la soluzione ha modulo costante uguale ad 1, mentre le componenti oscillano periodicamente. Infine, se $\alpha > 0$ la soluzione diverge.

Abbiamo visto che i vari metodi numerici, applicati all'equazione test, assumono la forma:

$$y_{n+1} = \varphi(h\lambda) y_n$$

per cui

$$|y_{n+1}| = |\varphi(h\lambda)| |y_n| = |\varphi(h\lambda)|^2 |y_{n-1}| = \dots = |\varphi(h\lambda)|^{n+1} |y_0|.$$

Tale relazione dice che la soluzione numerica ottenuta con passo h costante ha un comportamento asintotico che dipende da $|\varphi(h\lambda)|$ nel seguente modo:

$$|\varphi(h\lambda)| < 1 \quad \Rightarrow \quad |y_n| \rightarrow 0 \quad \text{per } n \rightarrow \infty$$

$$|\varphi(h\lambda)| = 1 \quad \Rightarrow \quad |y_n| = |y_0| \quad \forall n$$

$$|\varphi(h\lambda)| > 1 \quad \Rightarrow \quad |y_n| \rightarrow \infty \quad \text{per } n \rightarrow \infty.$$

Sono interessanti i metodi che risultano asintoticamente stabili per tutte le equazioni che hanno soluzioni asintoticamente stabili, cioè $\alpha < 0$ o, equivalentemente, λ nel semipiano negativo.

Dalle considerazioni precedenti risulta che il metodo di Eulero esplicito è asintoticamente stabile solo per quei valori del passo h , tali che $h\lambda$ rimane incluso nella regione di assoluta stabilità. Viceversa i metodi di Eulero implicito e dei trapezi risultano asintoticamente stabili per ogni valore del passo h , poichè le loro regioni di assoluta stabilità includono l'intero semipiano negativo.

I metodi le cui regioni di assoluta stabilità includono il semipiano negativo sono detti **assolutamente stabili**.

Si osservi infine che il metodo di Eulero implicito ha una regione di assoluta stabilità più ampia del semipiano negativo. Ciò causa, per certi valori del passo, un andamento asintoticamente stabile del metodo anche per equazioni che hanno $\alpha > 0$, le cui soluzioni esatte divergono. Questa proprietà, nota come *smorzamento numerico* (numerical damping), è un aspetto negativo del metodo.

Un metodo perfetto, da questo punto di vista, è il metodo dei trapezi la cui soluzione ha, in ogni caso, lo stesso andamento qualitativo della soluzione esatta per ogni passo h .

Consideriamo ora, come equazione test, il sistema lineare:

$$\begin{aligned} y'(t) &= Ay(t) \\ y(0) &= u \end{aligned} \tag{8.7}$$

dove $A \in \mathbb{R}^m \times \mathbb{R}^m$ e $u = (1, 1, \dots, 1) \in \mathbb{R}^m$. Si dimostra che la soluzione è asintoticamente stabile se e solo se tutti gli autovalori di A hanno parte reale negativa.

Il metodo di Eulero esplicito applicato al sistema (8.7) è del tipo:

$$y_{n+1} = y_n + hAy_n = (I + hA)y_n$$

E' facile vedere che per ogni metodo si ha:

$$y_{n+1} = \varphi(hA)y_n$$

dove la funzione φ , che ora trasforma matrici in matrici, è proprio la funzione di stabilità precedentemente definita per il caso scalare.

In particolare per il metodo di Eulero implicito è $\varphi(hA) = (I - hA)^{-1}$ e per il metodo dei trapezi $\varphi(hA) = (I - hA)^{-1}(I + hA)$.

Poichè, dalla relazione precedente,

$$\|y_{n+1}\| \leq \|\varphi(hA)\| \|y_n\| \leq \dots \leq \|\varphi(hA)\|^{n+1} \|y_0\|$$

la soluzione numerica sarà asintoticamente stabile se $\|\varphi(hA)\| < 1$, cioè se il raggio spettrale, e quindi il modulo di ogni autovalore, della matrice $\varphi(hA)$ è < 1 .

Ricordando ora che se λ è autovalore di A allora $\varphi(h\lambda)$ è autovalore di $\varphi(hA)$, è sufficiente che sia $|\varphi(h\lambda)| < 1$ per ogni λ autovalore di A . Poichè λ è, in generale, un numero complesso, l'analisi effettuata per l'equazione test scalare (8.6) è sufficiente anche per il

caso vettoriale. Infatti un metodo risulta asintoticamente stabile se $h\lambda$ è incluso nella regione di assoluta stabilità per ogni λ autovalore di A .

In particolare se il metodo è assolutamente stabile, allora la soluzione numerica tende a zero, indipendentemente dal passo, per tutte le equazioni con soluzione asintoticamente stabile.