EQUAZIONI DIFFERENZIALI ORDINARIE *

A. SOMMARIVA†

Conoscenze richieste. Formula di Taylor. Risoluzione di equazioni nonlineari. Calcolo differenziale. Conoscenza di Matlab/Octave.

Conoscenze ottenute. Discretizzazione con Eulero esplicito ed Eulero implicito. Stabilità assoluta dei due metodi.

1. Problema di Cauchy.

PROBLEMA. 1.1 (Problema di Cauchy). *Si determini la funzione y tale che*

$$\begin{cases} y'(x) = f(x, y(x)), \ x \geq x_0 \\ y(x_0) = y_0 \end{cases} \tag{1.1}$$

dove f è a valori in $\mathbb{R}^n$ e definita in un sottoinsieme Ω di $\mathbb{R} \times \mathbb{R}^n$, con $(x_0, y_0) \in \Omega$.

NOTA. 1.1. *Di seguito supporremo che tale problema abbia una sola soluzione.*

ESEMPIO 1.1 (Equazione differenziale ordinaria).

$$\begin{cases} y'(x) = y(x), \ x \geq 0 \\ y(0) = 1 \end{cases} \tag{1.2}$$

la cui soluzione è $\exp(x)$*, è un problema di Cauchy con $f(x, y(x)) = y(x)$.*

ESEMPIO 1.2 (Sistema di equazioni differenziali ordinarie).

$$\begin{cases} y_1'(x) = -y_2(x), \ x \geq 0 \\ y_2'(x) = y_1(x), \ x \geq 0 \\ y_1(0) = 1, y_2(0) = 0 \end{cases} \tag{1.3}$$

la cui soluzione è $y(x) = (y_1(x), y_2(x)) = (\cos(x), \sin(x))$, è un problema di Cauchy con $y = (y_1, y_2)$, $x_0 = 0$, $y(x_0) = (1, 0)$ e

$$f(x, y(x)) = f(x, y_1(x), y_2(x)) = (-y_2(x), y_1(x)).$$

TEOREMA 1.1 (Teorema di Picard-Lindelöf). *Sia $D = [t_0 - \varepsilon, t_0 + \varepsilon] \times \mathbb{R}$. Dato il problema ai valori iniziali*

$$\begin{cases} y'(t) = f(t, y(t)), \ t \in [t_0 - \varepsilon, t_0 + \varepsilon] \\ y(t_0) = y_0 \end{cases}$$

se $f : D \to \mathbb{R}$ è una funzione lipschitziana in y cioè

$$|f(t, \mathbf{y_1}) - f(t, \mathbf{y_2})| \leq K \|\mathbf{y_1} - \mathbf{y_2}\|, (t, \mathbf{y_1}), (t, \mathbf{y_2}) \in D$$

ed è continua in t allora per qualche $\varepsilon > 0$, allora esiste un'unica soluzione y al problema ai valori iniziali sull'intervallo $[t_0 - \varepsilon, t_0 + \varepsilon]$.

TEOREMA 1.2 (Cauchy in piccolo, [2, p.7]). *Si supponga che*

*Ultima revisione: 22 maggio 2017

†Dipartimento di Matematica, Universitá degli Studi di Padova, stanza 419, via Trieste 63, 35121 Padova, Italia (alvise@euler.math.unipd.it). Telefono: +39-049-8271350.

- *D sia un aperto connesso di $\mathbb{R} \times \mathbb{R}^n$;*
- *$f : D \to \mathbb{R}$ sia una funzione continua in D;*
- *(t_0, y_0) un punto interno a D;*
- *la funzione f verifichi la condizione di Lipschitz*

$$\|f(t, \mathbf{y_1}) - f(t, \mathbf{y_2})\| \leq K\|\mathbf{y_1} - \mathbf{y_2}\|, (t, \mathbf{y_1}), (t, \mathbf{y_2}) \in D.$$

per qualche $K \geq 0$.

Allora esiste un unica funzione $\mathbf{y}$ definita su un intervallo $[t_0 - \alpha, t_0 + \alpha]$, per qualche $\alpha > 0$ tale che

$$\begin{cases} \mathbf{y}'(t) = f(t, \mathbf{y}(t)), \ t \in [t_0 - \alpha, t_0 + \alpha] \\ \mathbf{y}(t_0) = \mathbf{y_0} \end{cases} \tag{1.4}$$

TEOREMA 1.3 (Cauchy in grande, [4, p.331]). *Sia $f : [a, b] \times \mathbb{R}^n \to \mathbb{R}$ tale che*

- *$f(t, \mathbf{y})$, $t \in [a, b]$, $\mathbf{y} \in \mathbb{R}^n$ continua nella prima variabile,*
- *soddisfi rispetto alla seconda variabile la condizione di Lipschitz*

$$\|f(t, \mathbf{y_1}) - f(t, \mathbf{y_2})\| \leq K\|\mathbf{y_1} - \mathbf{y_2}\|, t \in [a, b], \mathbf{y_1}, \mathbf{y_2} \in \mathbb{R}^n.$$

per qualche norma vettoriale $\| \cdot \|$ e $K \geq 0$.

Allora il problema di Cauchy

$$\begin{cases} \mathbf{y}'(t) = f(t, \mathbf{y}(t)), \ t \in [a, b] \\ \mathbf{y}(a) = \mathbf{y_0} \end{cases} \tag{1.5}$$

ha **una e una sola** *soluzione $\mathbf{y}$ in $[a, b]$, per un arbitrario $\mathbf{y_0} \in \mathbb{R}^n$.*

NOTA. 1.2.

- *Nel teorema di Cauchy in piccolo si fornisce un teorema di esistenza e unicità locale.*
- *Nel teorema di Cauchy in grande si fornisce un teorema di esistenza e unicità globale.*

NOTA. 1.3 (Soluzioni multiple). *Esistono casi in cui il problema di Cauchy ha soluzioni multiple. Ad esempio, il problema*

$$\begin{cases} y'(x) = 2\sqrt{y(x)}, \ x > 0 \\ y(x_0) = 0 \end{cases} \tag{1.6}$$

ha per soluzioni, per $\alpha \in \mathbb{R}^+$

$$y(x) = \begin{cases} y(x) = (x - \alpha)^2, x \geq \alpha \\ y(x) = 0, \ \textit{altrimenti.} \end{cases} \tag{1.7}$$

2. Metodo di Eulero esplicito. Indichiamo con $\mathcal{I}(x, \overline{x})$ il più piccolo intervallo aperto contenente x e $\overline{x}$ (cioè $(x, \overline{x})$ oppure $(\overline{x}, x)$). Assumendo che la soluzione sia sufficientemente regolare, abbiamo dalla formula di Taylor per $x \approx \overline{x}$, $\xi \in \mathcal{I}(x, \overline{x})$ e (1.1)

$$\begin{aligned} y(x) &= y(\overline{x}) + y'(\overline{x})(x - \overline{x}) + \frac{y''(\xi)(x - \overline{x})^2}{2} \\ &\approx y(\overline{x}) + y'(\overline{x})(x - \overline{x}) = y(\overline{x}) + f(\overline{x}, y(\overline{x}))(x - \overline{x}) \end{aligned} \tag{2.1}$$

cioè

$$\boxed{y(x) \approx y(\overline{x}) + f(\overline{x}, y(\overline{x}))(x - \overline{x})}$$

Di conseguenza se si desidera calcolare la soluzione nei punti $x_{k+1} = x_0 + kh = x_k + h$ con $k > 0$, ponendo $x = x_{k+1}$, $\overline{x} = x_k$

$$y(x_{k+1}) \approx y(x_k) + h \cdot f(x_k, y(x_k)). \tag{2.2}$$

METODO 2.1 (Eulero esplicito, 1768-1770). *Tale metodo consiste nell'approssimare $y(x_{k+1})$ con u_{k+1} dove*

$$\boxed{u_{k+1} = u_k + h \cdot f(x_k, u_k),\ \ u_0 = y(x_0),} \tag{2.3}$$

ove $h = x_{k+1} - x_k$.

TEOREMA 2.1 (Errore Eulero esplicito). *Se l'unica soluzione y del problema di Cauchy in $[a, b]$ è sufficientemente regolare in $[x_k, x_{k+1}] \subseteq [a, b]$ e $u_k = y(x_k)$ allora*

$$\boxed{y(x_{k+1}) - u_{k+1} = \frac{y''(\xi)(x_{k+1} - x_k)^2}{2}} \tag{2.4}$$

per qualche $\xi \in (x_k, x_{k+1})$.

DIMOSTRAZIONE. 2.1. *Essendo*

$$y(x_{k+1}) = y(x_k) + y'(x_k)(x - x_k) + \frac{y''(\xi)(x_{k+1} - x_k)^2}{2}$$

da $u_k = y(x_k)$ otteniamo

$$\begin{aligned} u_{k+1} &= u_k + h \cdot f(x_k, u_k) = u_k + h \cdot f(x_k, y(x_k)) \\ &= y(x_k) + y'(x_k)(x_{k+1} - x_k), \end{aligned}$$

e quindi si ha per qualche ξ in (x_k, x_{k+1})

$$\begin{aligned} y(x_{k+1}) - u_{k+1} &= y(x_k) + y'(x_k)(x_{k+1} - x_k) + \frac{y''(\xi)(x_{k+1} - x_k)^2}{2} \\ &\quad - (y(x_k) + y'(x_k)(x_{k+1} - x_k)) \\ &= \frac{y''(\xi)(x_{k+1} - x_k)^2}{2}. \end{aligned}$$

□

3. Metodo di Eulero implicito. Similmente al caso di Eulero esplicito, da

$$y(x) \approx y(\overline{x}) + f(\overline{x}, y(\overline{x}))(x - \overline{x})$$

se poniamo $x = x_k$, $\overline{x} = x_{k+1}$ abbiamo

$$y(x_k) \approx y(x_{k+1}) + f(x_{k+1}, y(x_{k+1}))(x_k - x_{k+1}), \tag{3.1}$$

e quindi

$$y(x_{k+1}) \approx y(x_k) + h \cdot f(x_{k+1}, y(x_{k+1})). \tag{3.2}$$

METODO 3.1 (Metodo di Eulero implicito). *Tale metodo consiste nell'approssimare $y(x_{k+1})$ con u_{k+1} definito da*

$$\boxed{u_{k+1} = u_k + h \cdot f(x_{k+1}, u_{k+1}),\ u_0 = y(x_0)} \tag{3.3}$$

NOTA. 3.1 (Risoluzione equazioni nonlineari). *Evidentemente ad ogni iterazione si richiede di risolvere un'equazione nonlineare nella variabile z del tipo*

$$\boxed{z = u_k + h \cdot f(x_{k+1}, z)}$$

la cui soluzione può essere calcolata utilizzando ad esempio

- *col metodo di Newton o*
- *del punto fisso o*
- *delle secanti.*

NOTA. 3.2 (Metodi di Eulero e sistemi di equazioni differenziali). *Il problema di Cauchy è definito da*

$$\begin{cases} y'(x) = f(x, y(x)),\ x \geq x_0 \\ y(x_0) = y_0 \end{cases} \tag{3.4}$$

dove f è a valori in $\mathbb{R}^n$ e definita in un sottoinsieme Ω di $\mathbb{R} \times \mathbb{R}^n$, con $(x_0, y_0) \in \Omega$.

Nei casi multivariati con $n > 1$, indipendentemente dalla derivazione dovuta alla espansione di Taylor, i metodi

- *Eulero esplicito: $u_{k+1} = u_k + h \cdot f(x_k, u_k)$, $u_0 = y(x_0)$*
- *Eulero implicito: $u_{k+1} = u_k + h \cdot f(x_{k+1}, u_{k+1})$, $u_0 = y(x_0)$,*

possono essere utilizzati, in quanto la sequenza è ben definita.

4. Linear Multistep Methods. DEFINIZIONE 4.1 (Linear Multistep Methods). *Sono i metodi che approssimano $y(x_k)$ con u_k, dove $x_k = x_0 + kh$ e*

$$u_{n+1} = \sum_{j=0}^{p} a_j u_{n-j} + h \sum_{j=-1}^{p} b_j f(x_{n-j}, u_{n-j})$$

NOTA. 4.1. *Si osservi che*

- *Sono metodi per cui se voglio calcolare u_{n+1} necessitano $u_{n-p}, \ldots, u_n$ (si chiamano metodi a $p+1$ passi). Siccome inizialmente si dispone solo di u_0, per calcolare $u_1, \ldots, u_p$, necessari ad innescare il metodo, si utilizzano di solito altre strategie (come ad esempio applicare più volte metodi a un passo, cioè in cui $p = 0$, come ad esempio Eulero implicito).*
- *Se $b_{-1} = 0$ allora il metodo è esplicito, altrimenti è implicito.*

5. Analisi convergenza. NOTA. 5.1. *Osserviamo che il metodo di Eulero esplicito*

$$u_{k+1} = u_k + h \cdot f(x_k, u_k)$$

e Eulero implicito

$$u_{k+1} = u_k + h \cdot f(x_{k+1}, u_{k+1})$$

hanno la forma

$$u_{n+1} = \sum_{j=0}^{p} a_j u_{n-j} + h \sum_{j=-1}^{p} b_j f(x_{n-j}, u_{n-j})$$

ponendo rispettivamente
- *Eulero esplicito:* $p = 0,\ a_0 = 1,\ b_{-1} = 0,\ b_0 = 1$,
- *Eulero implicito:* $p = 0,\ a_0 = 1,\ b_{-1} = 1,\ b_0 = 0$.

DEFINIZIONE 5.1 (Convergenza). *Supponiamo di analizzare il problema di Cauchy nell'intervallo compatto* $I = [x_0, x_{fin}]$. *Siano* $x_s = x_0 + sh$, *con* $Nh_N = x_{fin} - x_0$,
- $y^{(h_N)} = \{y(x_k)\}$ *la soluzione esatta di un fissato problema di Cauchy,*
- $u^{(h_N)} = \{u_k\}$ *dove* u_k *è l'approssimazione di* $y(x_k)$ *fornita nei punti* x_k *da un metodo numerico a* $p+1$ *passi.*

Un metodo Linear Multistep a $p+1$ *passi si dice* **convergente** *se qualora i passi iniziali* $u_0^{(h_N)}, \ldots, u_p^{(h_N)}$ *sono tali che*

$$\eta(h_N) = \max_{0 \leq k \leq p} |y_k^{(h_N)} - u_k^{(h_N)}| \to 0 \textit{ per } h_N \to 0$$

si ha che

$$\|y^{(h_N)} - u^{(h_N)}\|_\infty \to 0 \textit{ per } h_N \to 0.$$

Se qualora $\eta(h_N) \leq C_1 h^q$ *implica* $\|y^{(h_N)} - u^{(h_N)}\|_\infty \leq C_2 h^q$, *con* C_1, C_2 *indipendenti da* h, *allora il metodo si dice convergente con ordine* q.

DEFINIZIONE 5.2 (Errori di troncamento, [2], p.112). *Se un metodo per la soluzione di problemi di Cauchy ha la forma*

$$u_{n+1} = \sum_{j=0}^{p} a_j u_{n-j} + h \sum_{j=-1}^{p} b_j f(x_{n-j}, u_{n-j})$$

ove $u_n \approx y(t_n)$, y *soluzione di (1.1) in* $t_n = x_0 + nh$, *e sia*

$$\boxed{\tau_n(h) = \frac{1}{h} \cdot \left(y(x_{n+1}) - \left(\sum_{j=0}^{p} a_j y(x_{n-j}) + h \sum_{j=-1}^{p} b_j f(x_{n-j}, y(x_{n-j}))\right)\right)}\ .$$

- *La quantità* $\tau_n(h)$ *chiama* *errore locale di troncamento* *del metodo.*
- *La quantità* $\tau(h) = \max_{n=0,\ldots} |\tau_n(h)|$ *si chiama* **errore globale di troncamento** *del metodo.*
- *Un metodo per cui* $\tau(h)$ *tende a* 0*, quando* $h \to 0$ *si dice* **consistente**.

NOTA. 5.2. *La consistenza rende conto di come la soluzione* y *del problema di Cauchy verifichi lo schema discreto del metodo linear multistep.*

NOTA. 5.3. *In queste note ci interesseremo esclusivamente della convergenza e consistenza di metodi linear multistep, ma esistono definizioni che permettono l'analisi di altri metodi non rientranti in questa famiglia.*

NOTA. 5.4. *Un teorema dovuto a Lax/Dahlquist, mostra che un metodo è convergente se e solo se consistente e stabile. Un metodo Linear Multistep a* $p+1$ *passi si dice (zero-)stabile per un certo problema di Cauchy nel compatto* $[x_0, x_{fin}]$ *se per ogni* $\epsilon > 0$ *esiste* h *tale che se* $h_N \leq h$ *e*

$$\max_{k=0,\ldots,p-1} |u_k^{(h_N)} - y(x_0 + kh_N)| \leq \epsilon$$

allora

$$\|u_k^{(h_N)} - y(x_0 + kh_N)\|_\infty \leq K\epsilon$$

con K *indipendente da* h*.*

Di seguito intendiamo mostrare, sotto opportune ipotesi, la **convergenza del metodo di Eulero esplicito**.

Consideriamo il metodo di Eulero esplicito, con $x_n = x_0 + nh$ per un prefissato passo $h = (x_{fin} - x_0)/N$. Sia y soluzione del problema di Cauchy, e

$$\overline{u}_n = y(x_{n-1}) + h \cdot f(x_{n-1}, y(x_{n-1})),$$

$$u_n = u_{n-1} + h \cdot f(x_{n-1}, u_{n-1}).$$

Osserviamo che

- la prima sequenza $\overline{u}^{(h)} = \{\overline{u}_n\}$, ha il generico $\overline{u}_n$ ottenuto applicando il metodo di Eulero esplicito partendo dal valore assunto dalla soluzione y in ogni x_{n-1},
- la seconda sequenza $u^{(h)} = \{u_n\}$, ha il generico u_n ottenuto applicando il metodo di Eulero esplicito partendo dal valore assunto u_{n-1}, che a priori non coincide con $y(x_{n-1})$.

Da

$$e_n = y(x_n) - u_n = (y(x_n) - \overline{u}_n) + (\overline{u}_n - u_n) \tag{5.1}$$

per la disuguaglianza triangolare

$$|e_n| \leq |y(x_n) - \overline{u}_n| + |\overline{u}_n - u_n|.$$

Analizziamo il primo termine $|y(x_n) - \overline{u}_n|$. Se la derivata seconda di y esiste ed è continua allora abbiamo per il teorema di Weierstrass

$$\max_{x \in [x_0, x_{fin}]} |y''(x)| \leq M$$

e quindi per $\xi_n \in (x_{n-1}, x_n)$,

$$|y(x_n) - \overline{u}_n| = (h^2/2) \cdot |y''(\xi_n)| \le Mh^2/2.$$

Di conseguenza,

$$\boxed{h|\tau_n(h)| = \max_{n=0,\ldots,N} |y(x_n) - \overline{u}_n| \le Mh^2/2.}$$

6. Analisi convergenza Eulero esplicito, caso Lipschitziano. Sketch della dimostrazione (supposto y suff. regolare e f continua e L-lipschitziana):

1. $|e_n| \le (1 + (1 + hL) + \ldots + (1 + hL)^{n-1})h|\tau(h)| + (1 + hL)^n |e_0|$;
2. $1+(1+hL)+\ldots+(1+hL)^{n-1} \le \frac{\exp(L(x_{fin}-x_0))}{hL}$, $(1+hL)^n \le \exp(L(x_{fin} - x_0))$;
3. conclusione perchè secondo membro infinitesimo per $h \to 0$ visto che $|\tau(h)| \le Mh^2/2$ e $|e_0| \to 0$ per $h \to 0$.

Se f è L-Lipschitziana (rispetto al secondo argomento) allora

$$|f(x, y_1) - f(x, y_2)| \le L|y_1 - y_2|$$

e otteniamo, ricordando che $e_n := y(x_n) - u_n$,

$$\boxed{|\overline{u}_n - u_n| \le (1 + hL)|e_{n-1}|.}$$

Infatti da

$$\overline{u}_n = y(x_{n-1}) + h \cdot f(x_{n-1}, y(x_{n-1})), \quad u_n = u_{n-1} + h \cdot f(x_{n-1}, u_{n-1})$$

ricaviamo dalla L-lipschitzianità

$$\begin{aligned}
|\overline{u}_n - u_n| &= |(y(x_{n-1}) + h \cdot f(x_{n-1}, y(x_{n-1}))) - (u_{n-1} + h \cdot f(x_{n-1}, u_{n-1}))| \\
&= |(y(x_{n-1}) - u_{n-1}) + h \cdot (f(x_{n-1}, y(x_{n-1})) - f(x_{n-1}, u_{n-1}))| \\
&\le |y(x_{n-1}) - u_{n-1}| + h|f(x_{n-1}, y(x_{n-1})) - f(x_{n-1}, u_{n-1}))| \\
&\le |y(x_{n-1}) - u_{n-1}| + hL|y(x_{n-1}) - u_{n-1}| = (1 + hL)|y(x_{n-1}) - u_{n-1}| \\
&= (1 + hL)e_{n-1}. \qquad (6.1)
\end{aligned}$$

Da

- $e_n := y(x_n) - u_n = (y(x_n) - \overline{u}_n) + (\overline{u}_n - u_n)$,
- $|y(x_n) - \overline{u}_n| \le h|\tau(h)|$,
- $|\overline{u}_n - u_n| \le (1 + hL)|e_{n-1}|$

si deduce

$$\begin{aligned}
|e_n| &= |(y(x_n) - \overline{u}_n) + (\overline{u}_n - u_n)| \\
&\le |y(x_n) - \overline{u}_n| + |\overline{u}_n - u_n| \\
&\le h|\tau(h)| + (1 + hL)|e_{n-1}| \qquad (6.2)
\end{aligned}$$

ovvero

$$\boxed{|e_n| \le h|\tau(h)| + (1 + hL)|e_{n-1}|.}$$

Quindi essendo

- $|e_0| \to 0$ per $h \to 0$,
- $|e_n| \leq h|\tau(h)| + (1+hL)|e_{n-1}|$,

ricaviamo

$$\begin{aligned} |e_n| &\leq h|\tau(h)| + (1+hL)|e_{n-1}| \\ &\leq h|\tau(h)| + (1+hL)(h|\tau(h)| + (1+hL)|e_{n-2}|) \\ &= (1+(1+hL))h|\tau(h)| + (1+hL)^2|e_{n-2}| \\ &\leq \ldots \\ &\leq (1+(1+hL)+\ldots+(1+hL)^{n-1})h|\tau(h)| + (1+hL)^n|e_0| \end{aligned}$$

e cioè

$$\boxed{|e_n| \leq (1+(1+hL)+\ldots+(1+hL)^{n-1})h|\tau(h)| + (1+hL)^n|e_0|.}$$

Ricordando che

$$1 + s + \ldots + s^k = \frac{(1-s^{k+1})}{(1-s)}$$

posto $s = 1 + hL$ deduciamo che

$$1+(1+hL)+\ldots+(1+hL)^{n-1} = \frac{1-(1+hL)^n}{1-(1+hL)} = \frac{(1+hL)^n - 1}{hL}.$$

Notiamo che per $\gamma > 0$

$$\exp(\gamma) = \sum_{k=0}^{\infty} \frac{\gamma^k}{k!} \geq \sum_{k=0}^{1} \frac{\gamma^k}{k!} = 1 + \gamma \tag{6.3}$$

e quindi $(1+\gamma)^n \leq (\exp(\gamma))^n = \exp(n\gamma)$ implica che per $\gamma = hL$ si ha essendo $nh = x_n - x_0$

$$\boxed{(1+hL)^n \leq \exp(nhL) = \exp(L(x_n - x_0)) \leq \exp(L(x_{fin} - x_0)).}$$

Da

- $1+(1+hL)+\ldots+(1+hL)^{n-1} = \frac{1-(1+hL)^n}{1-(1+hL)} = \frac{(1+hL)^n-1}{hL}$
- $(1+hL)^n \leq \exp(L(x_{fin} - x_0))$,

concludiamo che

$$\begin{aligned} 1+(1+hL)+\ldots+(1+hL)^{n-1} = \frac{(1+hL)^n - 1}{hL} &\leq \frac{\exp(L(x_{fin}-x_0)) - 1}{hL} \\ &\leq \frac{\exp(L(x_{fin}-x_0))}{hL} \end{aligned} \tag{6.4}$$

e quindi da $|\tau(h)| \leq Mh^2/2$

$$\begin{aligned} (1+(1+hL)+\ldots+(1+hL)^{n-1})h|\tau(h)| &\leq \frac{\exp(L(x_{fin}-x_0))}{hL} h|\tau(h)| \\ &= \exp(L(x_{fin}-x_0))|\tau(h)| \to 0. \end{aligned}$$

Di conseguenza da

$$
\begin{aligned}
|e_n| &\leq (1 + (1 + hL) + \ldots + (1 + hL)^{n-1})h|\tau(h)| + (1 + hL)^n |e_0| \\
&\leq \exp\left(L(x_{fin} - x_0)\right)|\tau(h)| + \exp\left(L(x_{fin} - x_0)\right)|e_0| \\
&\leq \exp\left(L(x_{fin} - x_0)\right)\frac{Mh^2}{2} + \exp\left(L(x_{fin} - x_0)\right)|e_0| \qquad (6.5)
\end{aligned}
$$

visto che il secondo membro non dipende da n ricaviamo

$$\max_{n=0,\ldots,N} |e_n| \leq \exp\left(L(x_{fin} - x_0)\right)\frac{Mh^2}{2} + \exp\left(L(x_{fin} - x_0)\right)|e_0|$$

e quindi il metodo di Eulero esplicito risulta convergente, visto che il secondo membro è infinitesimo.

Consideriamo il metodo di Eulero implicito

$$u_n = u_{n-1} + hf(t_n, u_n),\ u_0 = y_0.$$

Calcoliamo l'errore globale di troncamento. Se $g \in C^2(a, b)$,

$$\int_a^b g(x)dx = (b - a)g(b) - (b - a)^2 g^{(1)}(\xi)/2,\ \xi \in (a, b)$$

e posto $b = x_n$, $a = x_{n-1}$, $b - a = h$, se y è suff. regolare

$$
\begin{aligned}
y_n &= y_{n-1} + \int_{x_{n-1}}^{x_n} y'(x)dx = y_{n-1} + hy'(x_n) - h^2 y''(\xi)/2 \\
&= y_{n-1} + hf(x_n, y_n) - h^2 y''(\xi)/2 \qquad (6.6)
\end{aligned}
$$

e quindi facilmente

$$\tau_n(h) = -(1/h) \cdot (h^2 y''(\xi)/2) = -hy''(\xi)/2.$$

Ricordiamo che $y_n = y(t_n)$ è la soluzione del problema di Cauchy mentre

$$u_n = u_{n-1} + hf(t_n, u_n),\ u_0 = y_0.$$

Definiamo

$$\overline{u}_n := y_{n-1} + hf(t_n, \overline{u}_n),\ \overline{u}_0 = y_0.$$

Come in precedenza

$$y_n - u_n = (y_n - \overline{u}_n) + (\overline{u}_n - u_n)$$

e quindi per la disuguaglianza triangolare

$$|y_n - u_n| \leq |y_n - \overline{u}_n| + |\overline{u}_n - u_n|.$$

Studiamo separatamente i termini al secondo membro.

Osserviamo che

$$|u_n - \overline{u}_n| = |u_n - (y_{n-1} + hf(t_n, \overline{u}_n))| = h|\tau_n(h)|$$

Studiamo $|\overline{u}_n - u_n|$ ricordando che $e_0 = |y_0 - u_0| \to 0$. Dalla lipschitzianità (rispetto la seconda variabile)

$$\begin{aligned} |\overline{u}_n - u_n| &= |y_{n-1} + hf(t_n, \overline{u}_n) - u_{n-1} - hf(t_n, u_n)| \\ &= |y_{n-1} - u_{n-1} + hf(t_n, \overline{u}_n) - hf(t_n, u_n)| \\ &\le |y_{n-1} - u_{n-1}| + h|f(t_n, \overline{u}_n) - f(t_n, u_n)| \\ &\le |y_{n-1} - u_{n-1}| + hL|\overline{u}_n - u_n| \end{aligned} \tag{6.7}$$

da cui facilmente per h sufficientemente piccolo, cosicchè $1 - hL > 0$,

$$|\overline{u}_n - u_n| \le \frac{|y_{n-1} - u_{n-1}|}{1 - hL}.$$

Da

- $|y_n - \overline{u}_n| = h^2|y''(\xi)/2|$,
- $|\overline{u}_n - u_n| \le \frac{|y_{n-1} - u_{n-1}|}{1-hL}$,

se $\|y''\|_\infty \le M$

$$\begin{aligned} |y_n - u_n| &\le |y_n - \overline{u}_n| + |\overline{u}_n - u_n| \\ &\le \frac{h^2|y''(\xi)|}{2} + \frac{|y_{n-1} - u_{n-1}|}{1 - hL} \\ &\le \frac{h^2|y''(\xi)|}{2} + \frac{|y_{n-1} - u_{n-1}|}{1 - hL} \end{aligned} \tag{6.8}$$

Con ragionamenti simili a quelli utilizzati per mostrare la convergenza di Eulero esplicito, abbiamo per h piccolo che $s = 1/(1 - hL) > 1$, posto $\tau(h) = \max_n |\tau_n(h)|$

$$\begin{aligned} |e_n| &\le h|\tau_n(h)| + s|e_{n-1}| \le sh|\tau_n(h)| + s|e_{n-1}| \\ &\le sh|\tau_n(h)| + s(sh|\tau_{n-1}(h)| + s|e_{n-2}|) \\ &\le \dots \\ &\le h\sum_{k=1}^{n} s^k|\tau_{n-k+1}(h)| + s^n|e_0| \\ &\le h(\sum_{k=1}^{n} s^k)|\tau(h)| + s^n|e_0| \end{aligned} \tag{6.9}$$

Notiamo che se $hL < 1$ allora $s = 1/(1 - hL) > 1$ e che

$$\sum_{k=1}^{n} s^k = \frac{s^{n+1} - 1}{s - 1} \le \frac{s^{n+1}}{s - 1}$$

Ora

$$s - 1 = \frac{1}{1 - hL} - 1 = \frac{hL}{1 - hL} > 0 \tag{6.10}$$

e da $(1 + x)^k \le \exp(kx)$ per $x > 0$, abbiamo

$$\begin{aligned} s^{n+1} &= (1 + (s - 1))^{n+1} \le \exp((n + 1)(s - 1)) \\ &= \exp(\frac{hL(n + 1)}{1 - hL}) = \exp(\frac{L(x_{n+1} - x_0)}{1 - hL}) \end{aligned} \tag{6.11}$$

e quindi

$$\sum_{k=1}^{n} s^k \leq \frac{(1-hL)\exp(\frac{L(x_{n+1}-x_0)}{1-hL})}{hL}$$

Così ricapitolando

$$|e_n| \leq h(\sum_{k=1}^{n} s^k)|\tau(h)| + s^n|e_0|,$$

$$\sum_{k=1}^{n} s^k \leq \frac{(1-hL)\exp(\frac{L(x_{n+1}-x_0)}{1-hL})}{hL}$$

$$s^n \leq \exp(\frac{L(x_n - x_0)}{1-hL}),$$

implica finalmente

$$|e_n| \leq h\frac{(1-hL)\exp(\frac{L(x_{n+1}-x_0)}{1-hL})}{hL}|\tau(h)| + \exp(\frac{L(x_n-x_0)}{1-hL})|e_0|.$$

Se $\|y''\| \leq M$ e $x_{\text{fin}} = x_0 + Nh$, $|e_0| \to 0$ per $h \to 0$ abbiamo

$$\begin{aligned}
|e_n| &\leq h\frac{(1-hL)\exp(\frac{L(x_{n+1}-x_0)}{1-hL})}{hL}|\tau(h)| + \exp(\frac{L(x_n-x_0)}{1-hL})|e_0| \\
&\leq h\frac{(1-hL)\exp(\frac{L(x_{n+1}-x_0)}{1-hL})}{hL}\frac{Mh}{2} + \exp(\frac{L(x_{\text{fin}}-x_0)}{1-hL})|e_0| \\
&\leq \frac{(1-hL)\exp(\frac{L(x_{n+1}-x_0)}{1-hL})}{L}\frac{Mh}{2} + \exp(\frac{L(x_{\text{fin}}-x_0)}{1-hL})|e_0|
\end{aligned}$$

e quindi $|e_n| \to 0$ per $h \to 0$, cioè Eulero implicito è convergente.

Supponiamo di dover risolvere il problema di Cauchy

$$\begin{cases} y'(x) = f(x, y(x)), \ x \geq x_0 \\ y(x_0) = y_0 \end{cases} \tag{6.12}$$

Osserviamo che se $x_{n+1} = x_0 + nh$ allora ricordando la formula del trapezio per il calcolo di integrali

$$\begin{aligned}
y(x_{n+1}) &= y(x_n) + \int_{x_n}^{x_{n+1}} y'(x)dx \\
&\approx y(x_n) + (h/2)(y'(x_{n+1}) + y'(x_n)) \\
&= y(x_n) + (h/2)(f(x_n, y(x_n)) + f(x_{n+1}, y(x_{n+1}))).
\end{aligned}$$

NOTA. 6.1. *Utilizzando rispettivamente le formule dei rettangoli,*

$$\int_a^b g(x)dx \approx (b-a)g(a), \quad \int_a^b g(x)dx \approx (b-a)g(b)$$

si possono ottenere similmente i metodi di Eulero esplicito e implicito.
Essendo

$$y(x_{n+1}) \approx y(x_n) + (h/2)(f(x_n, y(x_n)) + f(x_{n+1}, y(x_{n+1}))) \tag{6.13}$$

si introduce il metodo di Crank-Nicolson (1947):

METODO 6.1 (Crank-Nicolson o dei trapezi). *Tale metodo consiste nell'approssimare $y(x_{n+1})$ con u_{n+1} dove*

$$\boxed{u_{n+1} = u_n + (h/2)(f(t_n, u_n) + f(t_{n+1}, u_{n+1}))} \tag{6.14}$$

con $u_0 = y_0$.

Si osservi che ad ogni iterazione, essendo il metodo implicito, bisogna risolvere una equazione nonlineare.

Nell'ambito delle equazioni differenziali ordinarie, esistono vari criteri di stabilitá. Un classico problema è quello di vedere se un metodo è assolutamente stabile. Definito il problema di Cauchy

$$\begin{cases} y'(x) = \lambda y(x), \ x \geq 0 \\ y(0) = 1 \end{cases} \tag{6.15}$$

per un certo $\lambda \in \mathbb{C}$ con $\Re(\lambda) < 0$, visto che l'unica soluzione è $y(x) = \exp(\lambda x)$ si cerca di definire il passo h cosicchè il metodo numerico per la soluzione del problema di Cauchy (1.1) abbia lo stesso comportamento asintotico di $\exp(\lambda x)$.

La regione di assoluta stabilità è composta dagli $h\lambda$, per cui il metodo numerico con passo h è tale da avere lo stesso comportamento asintotico del problema di Cauchy (6.15) relativo al parametro λ (con $\Re(\lambda) < 0$).

NOTA. 6.2 (Facoltativo). *Stabilire la stabilità per il problema di Cauchy*

$$\begin{cases} y'(x) = f(x, y(x)), \ x \geq 0 \\ y(x_0) = y_0 \end{cases} \tag{6.16}$$

è in generale complicato. Se invece consideriamo

$$\begin{cases} y'(x) = \lambda y(x) + g(x), \ x \geq 0 \\ y(0) = y_0 \end{cases} \tag{6.17}$$

la questione è più semplice. Ricordiamo che la soluzione di (6.17) è

$$y(x) = c \cdot \exp(\lambda x) + \int_0^x \exp(\lambda(x-t)) g(t) dt$$

con

$$c = y_0$$

(cf. [1, p.369]).

6.1. *Sia y la soluzione di*

$$\begin{cases} y'(x) = \lambda y(x) + g(x), \ x \geq 0 \\ y(0) = y_0 \end{cases} \tag{6.18}$$

e y_ϵ quella di

$$\begin{cases} y'(x) = \lambda y(x) + g(x), \ x \geq 0 \\ y(0) = y_0 + \epsilon \end{cases} \tag{6.19}$$

Posto $z_\epsilon = y_\epsilon - y$, da (6.18), (6.20),

$$\begin{cases} z'_\epsilon(x) = \lambda z_\epsilon(x), \ x \geq 0 \\ y(0) = \epsilon \end{cases} \tag{6.20}$$

la cui soluzione è $z_\epsilon = \epsilon \cdot \exp(\lambda x)$.

Di solito nelle applicazioni ci si interessa al caso $\lambda < 0$ o complesso con parte reale negativa. In questi casi $z_\epsilon \to 0$ per $n \to \infty$ e quindi l'effetto delle perturbazioni si annulla per valori grandi di x.

Si desidera che il metodo numerico goda delle stesse proprietà e la risposta la si ha nuovamente dallo studio della regione di stabilità.

DEFINIZIONE 6.1 (Problema stiff). *Un problema di Cauchy*

$$\begin{cases} y'(x) = \lambda y(x), \ x \geq 0 \\ y(0) = 1. \end{cases} \tag{6.21}$$

si dice stiff, quando $\Re(\lambda) \ll 0$ (cioè $\Re(\lambda)$ è molto grande in valore assoluto).

Ricordiamo a tal proposito che dalla formula di Eulero, se $z = a + ib = \Re(z) + \Im(z)$ allora

$$\exp(z) = \exp(a) \cdot (\cos(b) + i \sin(b)).$$

Quindi se $\Re(\lambda) < 0$, visto che $|\cos(\Im(\lambda x)) + i \sin(\Im(\lambda x))| = 1$ e $x > 0$ abbiamo

$$\begin{aligned} |\exp(\lambda x)| &= |\exp(\Re(\lambda x))| |\cos(\Im(\lambda x)) + i \sin((\Im(\lambda x)))| \\ &= |\exp(\Re(\lambda x))| \to 0, \ \text{per } x \to \infty. \end{aligned}$$

Visto il comportamento asintotico di $\exp(\lambda x)$, se u_n è l'approssimazione della soluzione in $x_n = x_0 + nh$ fornita da un metodo numerico a passo h, si desidera sia $u_n \to 0$ per $n \to +\infty$.

Nel caso del metodo di Eulero esplicito,

$$u_n = u_{n-1} + hf(x_{n-1}, u_{n-1}) = (1 + h\lambda)u_{n-1}, \ u_0 = 1.$$

Si verifica facilmente che questa equazione alle differenze (di ordine 1) ha quale unica soluzione

$$\boxed{u_n = (1 + h\lambda)^n}$$

e che

$$u_n \to 0 \text{ se e solo se } |1 + h\lambda| < 1$$

il che significa che $h\lambda$ deve stare nel disco del piano complesso di centro -1 e raggio 1.

Di conseguenza, fissato λ, il metodo risulta stabile se e solo se si sceglie un passo sufficientemente piccolo, cioè minore di $1/|\lambda|$.

Nel caso di problemi stiff, cioè con $\Re(\lambda) \ll 0$, si deve scegliere un passo h molto piccolo affinchè la soluzione di Eulero esplicito tenda a 0.

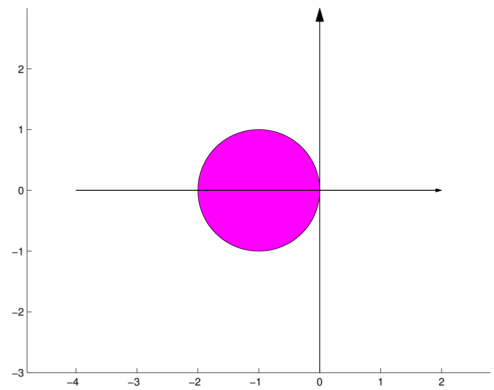

FIGURA 6.1. *In magenta: regione di assoluta stabilità di Eulero esplicito: disco unitario centrato in* $(-1, 0)$.

Nel caso del metodo di Eulero implicito,

$$u_n = u_{n-1} + hf(x_n, u_n) = u_{n-1} + h\lambda u_n, \ u_0 = 1.$$

Portando a primo membro $h\lambda u_n$, dividendo i membri per $1 - h\lambda$

$$u_n = \frac{u_{n-1}}{1 - h\lambda}.$$

Si verifica, ragionando per induzione, che

$$\boxed{u_n = \frac{1}{(1 - h\lambda)^n}}$$

e quindi che

$$u_n \to 0 \text{ se e solo se } \frac{1}{|1 - h\lambda|} < 1.$$

Essendo $\Re(\lambda) < 0$, si vede facilmente che $|\frac{1}{1-h\lambda}| < 1$ per qualsiasi h. Infatti, $h\lambda = a + ib$ con $a < 0$, ciò è vero se e solo se

$$\frac{1}{|1 - h\lambda|} < 1 \Leftrightarrow 1 < |1 - h\lambda| = \sqrt{(1 - a)^2 + b^2}$$

ovvero, elevando ai quadrati entrambi i membri, se e solo se

$$1 < (1 - a)^2 + b^2$$

ovviamente verificata in quanto $a < 0$, $a, b \in \mathbb{R}$. La regione di ass. stabilità è tutto il semipiano negativo $\Re(h\lambda) < 0$.

NOTA. 6.3. *Questa proprietà suggerisce di applicare Eulero implicito invece di Eulero esplicito per risolvere numericamente un problema di Cauchy di tipo stiff.*

Nel caso del metodo di Crank-Nicolson, da $f(x, y) = \lambda y$

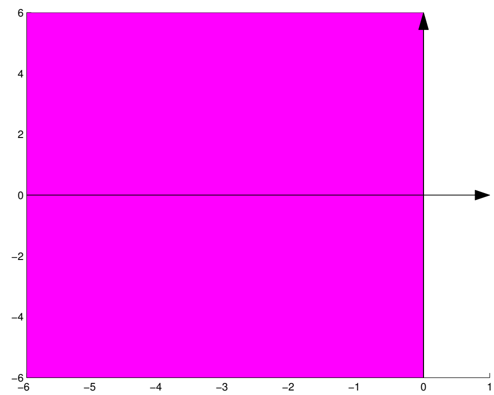

FIGURA 6.2. *In magenta: regione di assoluta stabilità di Eulero implicito, semipiano negativo* $\Re(h\lambda) < 0$.

$$\begin{aligned} u_n &= u_{n-1} + (h/2)(f(x_n, u_n) + f(x_{n-1}, u_{n-1})) \\ &= u_{n-1} + (h/2)\lambda(u_n + u_{n-1}),\ u_0 = 1 \end{aligned} \tag{6.22}$$

Quindi,

$$(1 - \frac{h\lambda}{2})u_n = u_{n-1} + \frac{h\lambda}{2}u_n$$

cioè

$$\frac{2 - h\lambda}{2} \cdot u_n = \frac{2 + h\lambda}{2} \cdot u_{n-1}$$

e di conseguenza

$$u_n = \frac{2 + h\lambda}{2 - h\lambda}u_{n-1}$$

Da

$$u_n = \frac{2 + h\lambda}{2 - h\lambda}u_{n-1}$$

si verifica facilmente che questa equazione alle differenze ha quale unica soluzione

$$\boxed{u_n = \frac{(2 + h\lambda)^n}{(2 - h\lambda)^n}}$$

e che

$$u_n \to 0 \text{ se e solo se } \frac{|2 + h\lambda|}{|2 - h\lambda|} < 1.$$

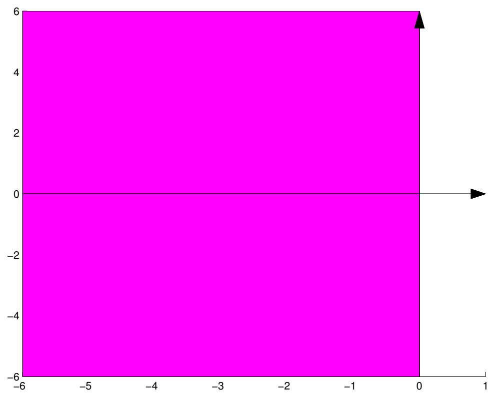

FIGURA 6.3. *In magenta: regione di assoluta stabilità di Crank-Nicolson, semipiano negativo* $\Re(h\lambda) < 0$.

Da $\Re(\lambda) < 0$, $|\frac{2+h\lambda}{2-h\lambda}| < 1$ per $h > 0$. Infatti, se $h\lambda = a + ib$, $a < 0$

$$\frac{|2+h\lambda|}{|2-h\lambda|} = \frac{|2+(a+ib)|}{|2-(a+ib)|} = \frac{|(2+a)+ib|}{|(2-a)+ib|} = \frac{\sqrt{(2+a)^2+b^2}}{\sqrt{(2-a)^2+b^2}} < 1$$

in quanto $(2+a)^2 < (2-a)^2$, qualora $a < 0$.

La regione di ass. stabilità è tutto il semipiano negativo $\Re(h\lambda) < 0$.

Posto $h > 0$, $x_n = x_0 + nh$, un metodo linear multistep è del tipo

$$u_{n+1} = \sum_{j=0}^{p} a_j u_{n-j} + h \sum_{j=-1}^{p} b_j f(x_{n-j}, u_{n-j}), \quad n = p, p+1, \ldots \tag{6.23}$$

noti i valori $u_k \approx y(x_k)$ per $k < p$ e supposto $a_p, b_p \neq 0$.

NOTA. 6.4. *Non è difficile vedere che*

- *Eulero esplicito* ($a_0 = 1$, $b_0 = 1$),
- *Eulero implicito* ($a_0 = 1$, $b_{-1} = 1$),
- *Crank-Nicolson* ($a_0 = 1$, $b_{-1} = 1/2$, $b_0 = 1/2$)

sono metodi che hanno questa struttura.

Da

- $y'(x) = f(x, y(x))$,
- $y(x_{n+1}) = y(x_{n-m}) + \int_{x_{n-m}}^{x_{n+1}} y'(x)\,dx$

ricaviamo che

$$y(x_{n+1}) = y(x_{n-m}) + \int_{x_{n-m}}^{x_{n+1}} y'(x)\,dx = y(x_{n-m}) + \int_{x_{n-m}}^{x_{n+1}} f(x, y(x))\,dx.$$

Se per $\gamma = 0$ oppure $\gamma = -1$ e un numero naturale p, si suppone

$$y(x_{n+\gamma-k}) \approx u_{n+\gamma-k},\ k = \gamma, \ldots, p$$

$$f_{n+\gamma-k} := f(x_{n+\gamma-k}, u_{n+\gamma-k})$$

allora se $\mathcal{P}_p(x) = \sum_k f_{n+\gamma-k} L_k(x)$ è il polinomio che interpola le coppie $(x_{n+\gamma-k}, f_{n+\gamma-k})$ scritto nella forma di Lagrange, si ha

$$y(x_{n+1}) \approx y(x_{n-m}) + \int_{x_{n-m}}^{x_{n+1}} \mathcal{P}_p(x)\,dx = y(x_{n-m}) + \sum_k f_{n+\gamma-k} \int_{x_{n-m}}^{x_{n+1}} L_k(x)dx$$

e quindi facilmente un metodo via quadratura numerica dell'integrale.

Scegliendo i nodi $\{x_n\}$ equispaziati,

- se $m = 0$, $\gamma = 0$ si hanno metodi espliciti (Adams-Bashforth, (1883));
- se $m = 0$, $\gamma = 1$ si hanno metodi impliciti (Adams-Moulton, (1926)).

NOTA. 6.5.

- *Nel caso $m = p - \gamma$ in particolare, la formula di integrazione è in realtá quella ben nota di tipo Newton-Cotes (chiusa o aperta a seconda $\gamma = -1$ o $\gamma = 0$).*
- *Si osservi che altri metodi di tipo* linear multistep *sono ottenibili per altre scelte di m e γ. Esempi sono per $n \geq 1$*
  - *metodo del punto medio:* $u_{n+1} = u_{n-1} + 2hf_n$,
  - *metodo di Milne:* $u_{n+1} = u_{n-1} + (h/3)[f_{n-1} + 4f_n + f_{n+1}]$.
- *Con un approccio diverso si possono ottenere i metodi LM di tipo BDF ove $u_{n+1} = \sum_{j=0}^{p} a_j u_{n-j} + hb_{-1} f_{n+1}$.*
- Adams-Bashforth (p=0): $u_{n+1} = u_n + hf_n$, $\tau_{n+1} = (h/2)y^{(2)}(\xi_n)$.
- Adams-Bashforth (p=1): $u_{n+1} = u_n + \frac{h}{2}(3f_n - f_{n-1})$, $\tau_{n+1} = (5h^2/12)y^{(3)}(\xi_n)$.
- Adams-Bashforth (p=2): $u_{n+1} = u_n + \frac{h}{12}(23f_n - 16f_{n-1} + 5f_{n-2})$, $\tau_{n+1} = (3h^3/8)y^{(4)}(\xi_n)$.
- Adams-Bashforth (p=3): $u_{n+1} = u_n + \frac{h}{24}(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3})$, $\tau_{n+1} = (251h^4/720)y^{(5)}(\xi_n)$.
- Adams-Bashforth (p=4): $u_{n+1} = u_n + \frac{h}{720}(1901f_n - 2774f_{n-1} + 2616f_{n-2} - 1274f_{n-3} + 251f_{n-4})$, $\tau_{n+1} = (95h^5/2888)y^{(5)}(\xi_n)$.

dove τ_{n+1} è l'errore locale di troncamento

NOTA. 6.6. *Si noti che sono effettivamente metodi espliciti a $p+1$ passi.*

- Adams-Moulton (p=0): $u_{n+1} = u_n + hf_{n+1}$, $\tau_{n+1} = (-h/2)y^{(2)}(\xi_n)$.
- Adams-Moulton (p=1): $u_{n+1} = u_n + \frac{h}{2}(f_{n+1} + f_n)$, $\tau_{n+1} = (-h^2/12)y^{(3)}(\xi_n)$.
- Adams-Moulton (p=2): $u_{n+1} = u_n + \frac{h}{12}(5f_{n+1} + 8f_n - f_{n-1})$, $\tau_{n+1} = (-h^3/24)y^{(4)}(\xi_n)$.
- Adams-Moulton (p=3): $u_{n+1} = u_n + \frac{h}{24}(9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2})$, $\tau_{n+1} = (-19h^4/720)y^{(5)}(\xi_n)$.
- Adams-Moulton (p=4): $u_{n+1} = u_n + \frac{h}{720}(251f_{n+1} + 646f_n - 264f_{n-1} + 106f_{n-2} - 19f_{n-3})$, $\tau_{n+1} = (-3h^5/160)y^{(6)}(\xi_n)$.

NOTA. 6.7. *Si noti che sono effettivamente metodi impliciti a p passi.*

TEOREMA 6.1 ([5, p.437]). *Un metodo linear multistep*

- *è consistente se e solo se*

$$\boxed{\sum_{j=0}^{p} a_j = 1, \quad -\sum_{j=0}^{p} j a_j + \sum_{j=-1}^{p} b_j = 1,}$$

- *se $y \in C^{q+1}$ per $q \geq 1$, ha ordine di consistenza q se e solo se è consistente e*

$$\boxed{\sum_{j=0}^{p}(-j)^i a_j + i\sum_{j=-1}^{p}(-j)^{i-1}b_j = 1, i = 2,\ldots,q.}$$

TEOREMA 6.2 (Convergenza LM (cf.[1, p.116])). *Si supponga che*

- *i dati iniziali siano calcolati in modo che*

$$\eta(h) = \max_{i=0,\ldots,p} |y(x_i) - u_i| \to 0, \;\; \textit{per } h \to 0;$$

- *il metodo sia consistente;*
- $a_j \geq 0, \;\; j = 0,\ldots,p$*;*
- *il passo di discretizzazione* h *sia tale che* $h \leq 1/(2c)$*, dove* $c = L\sum_{j=-1}^{p}|b_j|$ *e* L *è la costante di Lipschitz di* f*.*

Allora il metodo LM converge ed inoltre esistono C_1*,* C_2 *positive tali che per ogni* n

$$|y(x_n) - u_n| \leq C_1\eta(h) + C_2\tau(h).$$

Se la soluzione è $m+1$ *volte derivabile con continuità, il metodo ha ordine di consistenza* m *e gli errori iniziali soddisfano* $\eta(h) = O(h^m)$ *allora il metodo è di ordine* m*.*

NOTA. 6.8. *Nelle ipotesi del teorema precedente, per ricavare la convergenza chiedavamo:*

- *i dati iniziali siano calcolati in modo che*

$$\eta(h) = \max_{i=0,\ldots,p} |y(x_i) - u_i| \to 0, \;\; \textit{per } h \to 0;$$

- *il metodo sia consistente;*
- $a_j \geq 0, \;\; j = 0,\ldots,p$*;*
- *il passo di discretizzazione* h *sia tale che* $h \leq 1/(2c)$*, dove* $c = L\sum_{j=-1}^{p}|b_j|$ *e* L *è la costante di Lipschitz di* f*.*

1. Se prendiamo un metodo di Adams-Bashforth o Adams-Moulton di quelli esposti, sono tutti consistenti con ordine $p+1$*. Inoltre* $a_0 = 1, a_1 = 0, \ldots, a_p = 0$*. Di conseguenza tutti i metodi esposti sono convergenti.*

2. Se la soluzione è $p+2$ *volte derivabile con continuità, il metodo ha ordine di consistenza* $p+1$ *e gli errori iniziali soddisfano* $\eta(h) = O(h^m)$ *allora il metodo è di ordine* $p+1$*.*

6.2. *In particolare, nelle ipotesi richieste,*

- *il metodo di Eulero esplicito, avendo ordine di consistenza* 1*, è convergente con ordine di convergenza* 1*;*
- *il metodo di Eulero implicito, avendo ordine di consistenza* 1*, è convergente con ordine di convergenza* 1*;*
- *il metodo di Crank-Nicolson, avendo ordine di consistenza* 2*, è convergente con ordine di convergenza* 2*.*

DEFINIZIONE 6.2. *Un metodo numerico è detto* A*-stabile se la sua regione di stabilità assoluta contiene tutto il semipiano negativo.*

Valgono le seguenti barriere di stabilità (Dahlquist, 1963)

TEOREMA 6.3. *Nessun metodo LM esplicito è A-stabile.*

TEOREMA 6.4. *Nessun metodo LM implicito di ordine maggiore di 2 è A-stabile.*

In generale un metodo LM implicito ha proprietà di stabilità migliori rispetto ad un LM esplicito e questo ne suggerisce l'utilizzo.

Purtroppo ad ogni iterazione bisogna risolvere una equazione nonlineare.

Un approccio comunemente utilizzato è quello del predictor-corrector. Consiste nell'utilizzare il metodo di punto fisso con una certa scelta del punto iniziale.

Nel caso di un metodo di Adams implicito (detto corrector), si utilizza ad esempio un metodo di Adams esplicito (detto predictor).

Consideriamo i seguenti metodi di Adams (porremo $f_i = f(x_i, u_i)$).

$$\text{Predictor: } u_{i+1} = u_i + \frac{h}{2}(3f_i - f_{i-1})$$

$$\text{Corrector: } u_{i+1} = u_i + \frac{h}{2}(f_{i+1} + f_i).$$

Posto $f_i^{(\eta)} = f(x_i, u_i^{(\eta)})$ abbiamo

$$\begin{aligned}
&\text{P: } u_{i+1}^{(0)} = u_i^{(1)} + \frac{h}{2}(3f_i^{(1)} - f_{i-1}^{(1)})\\
&\text{E: } f_{i+1}^{(0)} = f(x_{i+1}, u_{i+1}^{(0)})\\
&\text{C: } u_{i+1}^{(1)} = u_i^{(1)} + \frac{h}{2}(f_{i+1}^{(0)} + f_i^{(1)})\\
&\text{E: } f_{i+1}^{(1)} = f(x_{i+1}, u_{i+1}^{(1)})
\end{aligned}$$

Consideriamo un metodo numerico, che diremo di Runge-Kutta di ordine 2, che definisca una sequenza u_n t.c.

$$y(x_{n+1}) \approx u_{n+1} = u_n + hF(x_n, u_n; h)$$

con

$$F(x, y; h) = \gamma_1 f(x, y) + \gamma_2 f(x + \alpha h, y + \beta h f(x, y))$$

e determiniamo i parametri γ_1, γ_2, α, β così da ottenere un metodo del second'ordine cioè

$$\tau(h) = \max_n(\tau_n(h)) = O(h^2)$$

dove

$$\tau_n(h) = \frac{y(x_n)}{\overline{u}_n}, \ \overline{u}_{n+1} = y_n + hF(x_n, y_n; h).$$

Per ottenere questo risultato usiamo la formula di Taylor bivariata. Denotate con f_x, f_y le derivate parziali rispetto al primo e secondo argomento di f, abbiamo

$$f(x_n+\alpha h, y+\beta h f(x_n, y_n)) = f(x_n, y_n)+\alpha h f_x(x_n, y_n)+\beta h f_y(x_n, y_n) f(x_n, y_n)+O(h^2).$$

e quindi

$$\begin{aligned} F(x_n, y_n; h) &= \gamma_1 f(x_n, y_n) + \gamma_2 f(x_n + \alpha h, y_n + \beta h f(x_n, y_n)) \\ &= \gamma_1 f(x_n, y_n) + \gamma_2 (f(x_n, y_n) + \alpha h f_x(x_n, y_n) + \beta h f_y(x_n, y_n) f(x_n, y_n)) \\ &+ O(h^2). \end{aligned} \tag{6.24}$$

Facilmente, dalla formula di Taylor

$$\begin{aligned} y_{n+1} &= y_n + h y^{(1)}(x_n) + (h^2/2) y^{(2)}(x_n) + O(h^3) \\ &= y_n + h f(x_n, y_n) + (h^2/2)(f_x(x_n, y_n) + f_y(x_n, y_n) f(x_n, y_n)) + O(h^3) \end{aligned}$$

Ma è pure

$$\begin{aligned} \overline{u}_{n+1} &= y_n + h F(x_n, y_n; h) = y_n + h(\gamma_1 f(x_n, y_n) + \gamma_2 (f(x_n, y_n) \\ &+ \alpha h f_x(x_n, y_n) + \beta h f_y(x_n, y_n) f(x_n, y_n))) + O(h^3) \\ &= y_n + h(\gamma_1 + \gamma_2) f(x_n, y_n) + h^2 \gamma_2 (\alpha f_x(x_n, y_n) + \beta f_y(x_n, y_n) f(x_n, y_n))) \\ &+ O(h^3) \end{aligned}$$

e quindi affinchè $y_{n+1} - \overline{u}_{n+1} = O(h^3)$, per confronto, basta richiedere

$$\boxed{\gamma_1 + \gamma_2 = 1, \gamma_2 \alpha = 1/2, \gamma_2 \beta = 1/2.}$$

Vediamo alcuni metodi in cui

$$\boxed{\gamma_1 + \gamma_2 = 1, \quad \gamma_2 \alpha = 1/2, \quad \gamma_2 \beta = 1/2.}$$

Metodo di Heun ($\alpha = 1$), (scoperto nel 1900):

$$\boxed{u_{n+1} = u_n + \frac{h}{2}(f(x_n, u_n) + f(x_n + h, u_n + h f(x_n, u_n))).}$$

Metodo di Eulero modificato ($\alpha = 1/2$):

$$\boxed{u_{n+1} = u_n + h f(x_n + \frac{h}{2}, u_n + \frac{h}{2} f(x_n, u_n)).}$$

La scelta particolare di un maggior numero di vincoli permette, con qualche fatica, di calcolare un metodo di ordine 4 (scoperto da Kutta nel 1901). Posto

$$u_{n+1} = u_n + h \cdot \frac{f_1 + 2f_2 + 2f_3 + f_4}{6}$$

ove

$$\begin{aligned} f_1 &= f(t_n, u_n) \\ f_2 &= f(t_n + \frac{h}{2}, u_n + \frac{h \cdot f_1}{2}) \\ f_3 &= f(t_n + \frac{h}{2}, u_n + \frac{h}{2} f_2) \\ f_4 &= f(t_n + h, u_n + h \cdot f_3) \end{aligned} \tag{6.25}$$

si ottiene un metodo del quart'ordine.

Gli argomenti che seguono sono facoltativi. Servono solo a dare una immagine più completa sul tema dell'analisi numerica per la risoluzione di problemi di Cauchy.

Se f non è L-Lipschitziana ma è L-dissipativa, cioè

$$-L \leq \frac{\partial f}{\partial y}(\xi) \leq 0, \quad \xi \in \Omega := (x_0, x_{\mathbf{fin}}) \times \mathbb{R}$$

abbiamo da

- $\overline{u}_n = y(x_{n-1}) + h \cdot f(x_{n-1}, y(x_{n-1}))$,
- $u_n = u_{n-1} + h \cdot f(x_{n-1}, u_{n-1})$,
- $f(x_{n-1}, u_{n-1}) = f(x_{n-1}, y(x_{n-1})) + \frac{\partial f}{\partial y}(\xi)(u_{n-1} - y(x_{n-1}))$,

che per qualche $\xi \in (x_{n-1}, x_n) \times \mathbb{R} \subseteq \Omega$

$$\begin{aligned}
u_n - \overline{u}_n &= (u_{n-1} + h \cdot f(x_{n-1}, u_{n-1})) - (y(x_{n-1}) + h \cdot f(x_{n-1}, y(x_{n-1}))) \\
&= (u_{n-1} - y(x_{n-1})) + h \cdot (f(x_{n-1}, u_{n-1}) - f(x_{n-1}, y(x_{n-1}))) \\
&= (u_{n-1} - y(x_{n-1})) + h \frac{\partial f}{\partial y}(\xi)(u_{n-1} - y(x_{n-1})) \\
&= (1 + h \frac{\partial f}{\partial y}(\xi)) \cdot (u_{n-1} - y(x_{n-1})).
\end{aligned}$$

Da

$$u_n - \overline{u}_n = (1 + h \frac{\partial f}{\partial y}(\xi)) \cdot (u_{n-1} - y(x_{n-1})). \tag{6.26}$$

abbiamo che se $0 < h \leq 2/L$ (non restrittivo per mostrare la conv. visto che si studia il comportamento per $h \to 0$) allora da $-L \leq \frac{\partial f}{\partial y} \leq 0$,

$$-1 = 1 - 2L/L \leq 1 - hL \leq 1 + h \frac{\partial f}{\partial y}(\xi) \leq 1$$

e quindi $|1 + h \frac{\partial f}{\partial y}(\xi)| \leq 1$ da cui

$$\begin{aligned}
|u_n - \overline{u}_n| &= |1 + h \frac{\partial f}{\partial y}(\xi)| \cdot |u_{n-1} - y(x_{n-1})| \\
&\leq |u_{n-1} - y(x_{n-1})| = |e_{n-1}|.
\end{aligned} \tag{6.27}$$

Da

- $e_n := y(x_n) - u_n$,
- $|y(x_n) - \overline{u}_n| \leq h|\tau(h)|$,
- $|\overline{u}_n - u_n| \leq |e_{n-1}|$,
- $|e_n| \leq h\tau(h) + |\overline{u}_n - u_n| \leq h|\tau(h)| + |e_{n-1}|$

ricaviamo, essendo $h|\tau(h)| > 0$, $e_0 = 0$, $n \leq N$, $Nh = x_{\text{fin}} - x_0$

$$
\begin{aligned}
|\overline{u}_n - u_n| &\le |e_{n-1}| \le h|\tau(h)| + |e_{n-1}| \\
&\le h|\tau(h)| + (h|\tau(h)| + |e_{n-2}|) \\
&= 2h|\tau(h)| + |e_{n-2}| \\
&\le 2h|\tau(h)| + (h|\tau(h)| + |e_{n-3}|) \\
&= 3h|\tau(h)| + |e_{n-3}| \le \dots \\
&\le nh|\tau(h)| + |e_0| = nh|\tau(h)| \\
&\le Nh|\tau(h)| = (x_{\text{fin}} - x_0)|\tau(h)| \le \frac{(x_{\text{fin}} - x_0)Mh}{2}.
\end{aligned}
$$

Di conseguenza, visto che $(x_{\text{fin}} - x_0)$ non dipende da n, essendo per ogni n

$$|\overline{u}_n - u_n| \le \frac{(x_{\text{fin}} - x_0)Mh}{2}$$

abbiamo

$$\|u^{(h)} - \overline{u}^{(h)}\|_\infty = \max_n |u_n - \overline{u}_n| \le \frac{(x_{\text{fin}} - x_0)Mh}{2}.$$

Posto $C_2^*(h) = (x_{\text{fin}} - x_0)Mh/2$ abbiamo che

$$\boxed{\|u^{(h)} - \overline{u}^{(h)}\|_\infty \le C_2^*(h)}$$

con $C_2^*(h) \to 0$ qualora $h \to 0$ e quindi il metodo di Eulero esplicito, è convergente anche nel caso in cui f sia L-dissipativa.

RIFERIMENTI BIBLIOGRAFICI

[1] K. Atkinson e W. Han, *Elementary Numerical Analysis*, Wiley, (2004).
[2] K.E. Atkinson, W. Han, D.E. Steward *Numerical Solution of Ordinary Differential Equations*, Wiley, (2009).
[3] V. Comincioli, *Analisi Numerica, metodi modelli applicazioni*, Mc Graw-Hill, 1990.
[4] W. Gautschi, *Numerical Analysis*, Birkhaüser, second edition, 2012.
[5] A. Quarteroni, F. Saleri *Introduzione al Calcolo Scientifico*, Springer, (2002).