

Metodi Numerici

Daniela Nardin

a.a.2005-06

Si consideri il problema differenziale di Cauchy:

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

con $f : S \equiv \{(x, y) \in R^2 : -\infty < \alpha \leq x \leq \beta < +\infty; -\infty < \gamma \leq y \leq \eta < +\infty\} \rightarrow R$, che soddisfa le seguenti ipotesi: $f(x, y)$ é continua in S e ivi soddisfa la condizione di Lipschitz: $|f(x, y'') - f(x, y')| \leq L|y'' - y'|$ essendo $L \in R$ e $\forall x \in [\alpha, \beta]$; oppure la derivata parziale $\frac{\partial f}{\partial y}$ continua e limitata in S . Allora per il teorema noto esiste ed é unica la soluzione $y(x)$ del problema di Cauchy in un intorno di x_0 di raggio $\delta = \min\{\alpha; \frac{\beta}{M}\}$ con $M = \max|f(x, y)|$ in S . Se invece $f(x, y)$ é definita in una striscia infinita: $S \equiv \{(x, y) \in R^2 : -\infty < \alpha \leq x \leq \beta < +\infty; y \in R\}$ per essere certi della esistenza e unicitá della soluzione in tutta la striscia si deve aggiungere alle ipotesi precedenti la possibilitá di maggiorare la $f(x, y)$ in tutto S cioé $|f(x, y)| \leq A$, $A \in R$, $(x, y) \in S$. Nel contesto appena descritto ci poniamo il problema di determinare in un numero finito di punti $\{x_n\}$ dell'intervallo di interesse $[\alpha, \beta]$ delle approssimazioni $\{y_n\}$ dei valori $y(x_n)$ che la soluzione $y(x)$ assume nei nodi $\{x_n\}$. Ovviamente note le $\{y_n\}$, mediante le tecniche di interpolazione, sará poi possibile dedurre approssimazioni della $y(x)$ in punti diversi dagli $\{x_n\}$. Questi metodi sono raggruppati in due classi principali:

1. Metodi one-step o a un solo passo: il valore y_{n+1} viene calcolato utilizzando unicamente l'approssimazione precedente $\{y_n\}$.
2. Metodi multi step o a piú passi: il valore y_{n+1} viene calcolato utilizzando piú approssimazioni precedenti $y_n, y_{n-1}, \dots, y_{n-k+1}$ mediante una opportuna funzione Ψ quindi:

$$y_{n+1} = \Psi(x_n, y_n, y_{n-1}, \dots, y_{n-k+1}, h) \quad h = h(n)$$

metodo multi step a k passi.

Il generico metodo one-step é il caso particolare con $k = 1$

$$\begin{cases} y_{n+1} = y_n + h_n \Phi(x_n, y_n, h_n) \\ x_{n+1} = x_n + h_n \end{cases}$$

h in generale dipende da n e quindi $h = h(n) = h_n$, quindi con h_n si intende che in generale il passo non è tenuto costante. Il grado di accuratezza delle approssimazioni fornite da un metodo one-step è strettamente legato al comportamento locale del metodo stesso ossia all'errore che il metodo introduce quando opera un singolo passo di integrazione.

1 Comportamento locale dei metodi one-step

Dato un generico punto $(x^*, y^*) \in S$ consideriamo la soluzione $u(x)$ del problema, quindi:

$$\begin{cases} u'(x) = f(x, u(x)) & x^* \leq x \leq \beta \\ u(x^*) = y^* \end{cases}$$

ossia la curva soluzione dell'equazione differenziale data che esce dal punto (x^*, y^*) . Si noti che la curva unica soluzione del problema di Cauchy che passa per (x_0, y_0) è $y(x)$, in genere diversa da $u(x)$:

Il metodo one-step scelto, partendo da (x^*, y^*) avrà come obiettivo la curva $u(x)$ che è la soluzione del problema cui esso viene applicato in quell'istante, cioè cercherà, in generale senza riuscirci, di seguire la $u(x)$.

In realtà il metodo definirà una seconda curva funzione di h $y_h = y^* + h\Phi(x^*, y^*, h)$ che esce dal punto (x^*, y^*) ma che non coincide con $u(x)$.

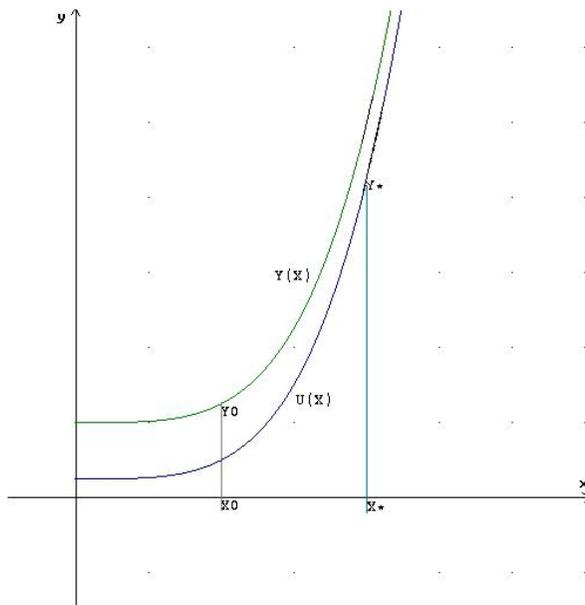
Definizione

Per ogni $(x^*, y^*) \in S$ e $h > 0$, definiamo errore locale unitario di troncamento del metodo Φ nel punto $x^* + h$ la funzione:

$$\tau(x^*, y^*, h) = \frac{u(x^* + h) - u(x^*)}{h} - \frac{y_{n+1} - y_n}{h} = \frac{u(x^* + h) - u(x^*)}{h} - \Phi(x^*, u(x^*), h)$$

(τ rappresenta quanto l'inclinazione della secante a $u(x)$ dista da Φ).

La quantità $\tau(x^*, y^*, h)$ ci da un'indicazione della bontà dello schema discreto $y_h = y^* + h\Phi(x^*, y^*, h)$. Per tale metodo $\tau(x^*, y^*, h)$ corrisponde alla quantità $\frac{u(x^* + h) - y_h}{h}$. Infatti $\Phi = \frac{y_h - y^*}{h}$ quindi $\tau = \frac{u(x^* + h) - u(x^*)}{h} -$



$$\frac{y_h - y^*}{h} = \frac{u(x^* + h) - y_h}{h}.$$

Definizione: Il metodo Φ è detto consistente se per ogni $(x^*, y^*) \in S$ e $\forall f \in C^1(S)$ risulta:

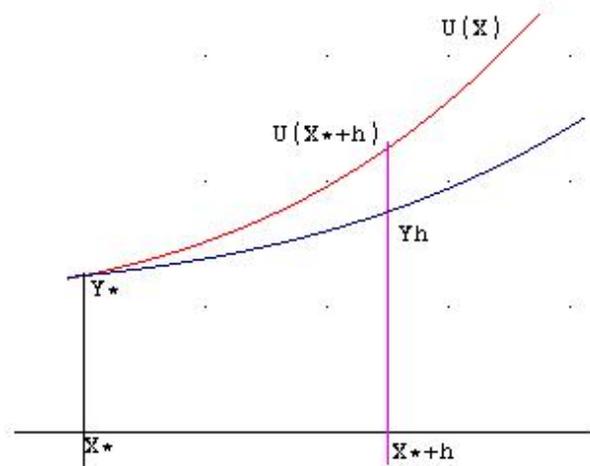
$$\lim_{h \rightarrow 0} \tau(x^*, y^*, h) = 0$$

Poichè $\Phi(x^*, y^*, h)$ costituisce un'approssimazione di $\frac{u(x^* + h) - u(x^*)}{h}$ allora si può anche dire che il metodo è consistente se per ogni $(x^*, y^*) \in S$ e $\forall f \in C^1(S)$ risulta:

$$\lim_{h \rightarrow 0} \Phi(x^*, y^*, h) = f(x^*, y^*)$$

Definizione: Il metodo Φ ha ordine di consistenza $p \in \mathbb{N}$ se per ogni $(x^*, y^*) \in S$ e $\forall f \in C^1(S)$ risulta:

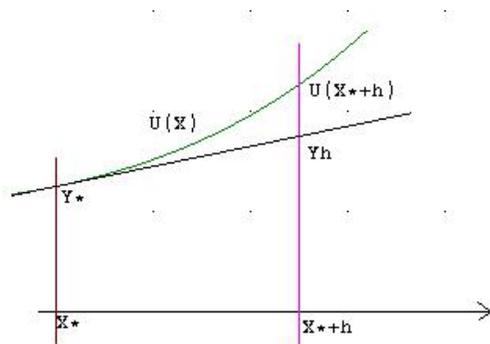
$$\tau(x^*, y^*, h) = O(h^p)$$



per $h \rightarrow 0$ essendo p l'intero più grande per cui sussiste la suddetta relazione.

2 Metodo di Eulero

Nel metodo di Eulero la funzione $\Phi(x^*, y^*, h)$ è scelta uguale al valore della derivata prima della funzione $y(x)$ soluzione del problema di Cauchy quindi $\Phi(x^*, y^*, h) = f(x^*, y^*)$. Questo significa che localmente in ogni punto (x^*, y^*) si approssima la funzione soluzione con la retta tangente al grafico passante per quel punto essendo $f(x^*, y^*)$ il coefficiente angolare della retta tangente.

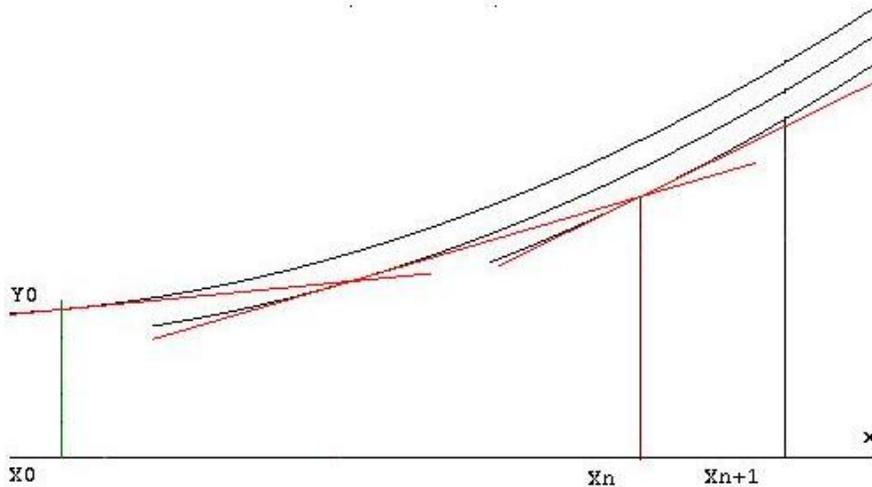


$$y_h = y^* + hf(x^*, y^*)$$

quindi

$$y_{n+1} = y_n + hf(x_n, y_n)$$

La soluzione è approssimata con una spezzata che passa per i punti (x_i, y_i)



Il metodo è consistente; per quanto riguarda l'errore locale unitario di troncamento si ha:

$$\tau(x^*, y^*, h) = \frac{u(x^* + h) - u(x^*)}{h} - f(x^*, y^*)$$

ma $f(x^*, y^*) = u'(x^*)$ Sviluppando in formula di Taylor si ha:

$$u(x^* + h) = u(x^*) + hu'(x^*) + \frac{h^2}{2}u''(\eta^*)$$

da cui segue:

$$\frac{u(x^* + h) - u(x^*)}{h} = u'(x^*) + \frac{h}{2}u''(\eta^*)$$

quindi $\tau(x^*, y^*, h) = \frac{h}{2}u''(\eta^*)$ essendo $x^* < \eta < x^* + h$. Se $f \in C^1(S)$ poichè $u'(x) = f(x, u(x))$ si ha che:

$$u''(x) = f_x(x, u(x)) + f_y(x, u(x))u'(x) = f_x(x, u(x)) + f_y(x, u(x))f(x, u(x))$$

quindi, per le ipotesi fatte su $f(x, y)$ $|u''(x)| \leq M$ pertanto $\tau(x^*, y^*, h) = O(h)$ per $h \rightarrow 0$ il metodo di Eulero ha ordine di consistenza $p = 1$. L'errore

$\tau(x^*, y^*, h)$ di cui si è parlato viene commesso in ogni singolo nodo; se si divide l'intervallo in n nodi e si eseguono n valutazioni della funzione Φ ci si chiede quale sarà l'errore complessivo perchè non è vero che l'errore complessivo sia la somma degli errori. La formula esatta (nel senso che è l'esatto valore teorico della funzione $y(x)$ soluzione unica del problema di Cauchy) é

$$y(x_n + h) = y(x_n) + h\Phi(x_n, y(x_n), h) + h\tau_{n+1}$$

il valore realmente calcolato è

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

l'errore è

$$e_{n+1} = y(x_n + h) - y_{n+1}$$

quindi

$$e_{n+1} = e_n + h[\Phi(x_n, y(x_n), h) - \Phi(x_n, y_n, h)] + h\tau_{n+1}$$

Supponiamo che la funzione Φ che nel metodo di Eulero è $f(x, y(x))$ sia lipshitziana rispetto a y , quindi

$$|f(x_n, y(x_n)) - f(x_n, y_n)| \leq L|y(x_n) - y_n|$$

si può scrivere:

$$|e_{n+1}| \leq |e_n| + hL|y(x_n) - y_n| + h|\tau_{n+1}|$$

se $\max_i |\tau_i| = \tau < +\infty$ (perchè il metodo è consistente) si ha

$$|e_{n+1}| \leq (1 + hL)|e_n| + h\tau \leq e^{hL}|e_n| + h\tau$$

(perchè $e^{hL} \geq 1 + hL$ in un intorno di $h = 0$.)

$$|e_1| \leq (1 + hL)|e_0| + h\tau$$

$$|e_2| \leq (1 + hL)|e_1| + h\tau = (1 + hL)^2|e_0| + (1 + hL)h\tau + h\tau$$

$$|e_3| \leq (1 + hL)^3|e_0| + (1 + hL)^2h\tau + (1 + hL)h\tau + h\tau$$

$$|e_3| \leq (1 + hL)^3|e_0| + h\tau((1 + hL)^2 + (1 + hL) + 1)$$

....

$$|e_n| \leq (1 + hL)^n|e_0| + \frac{\tau}{L}[(1 + hL)^n - 1]$$

$$|e_n| \leq e^{nhL}|e_0| + \frac{\tau}{L}e^{nhL} = e^{nhL}\left(|e_0| + \frac{\tau}{L}\right)$$

quindi l'errore dipende dall'errore iniziale e da quello di troncamento iniziale locale.

$$|e_n| \leq e^{(\beta-\alpha)L}\left(|e_0| + \frac{\tau}{L}\right)$$

essendo $[\alpha, \beta]$ intervallo nel quale si cerca la soluzione approssimata di $y(x)$; questo errore non dipende da n e, se e_0 è nullo, $\tau \rightarrow 0$ per $h \rightarrow 0$ allora il metodo è convergente. Quindi un metodo a un passo consistente è convergente se la funzione $f(x, y(x))$ dell'equazione differenziale è lipschitziana rispetto a y .

3 Metodi di Heun, Eulero modificato, Runge-Kutta

Per trovare l'ordine di consistenza del metodo di Eulero abbiamo usato lo sviluppo di Taylor della funzione $u(x)$ nel punto x^* e abbiamo approssimato $\frac{u(x^* + h) - u(x^*)}{h}$ con la funzione $f(x, u(x))$ questo fa pensare che se si prendesse come funzione Φ del metodo one step la ridotta dello sviluppo di Taylor di ordine ≥ 2 si potrebbero ottenere metodi più raffinati.

$$u(x^* + h) = u(x^*) + hu'(x^*) + \frac{h^2}{2}u''(x^*) + \frac{h^3}{3!}u'''(\eta)$$

da cui segue:

$$\frac{u(x^* + h) - u(x^*)}{h} = f(x^*, u(x^*)) + \frac{h}{2}[f_x(x^*, u(x^*)) + f_y(x^*, u(x^*))f(x^*, u(x^*))] + \frac{h^2}{3!}u'''(\eta)$$

quindi se scegliamo:

$$\Phi(x^*, y^*, h) = f(x^*, u(x^*)) + \frac{h}{2}[f_x(x^*, u(x^*)) + f_y(x^*, u(x^*))f(x^*, u(x^*))]$$

si ottiene un metodo del secondo ordine. Tuttavia questa formula coinvolge le derivate parziali oltre che $f(x, y)$ ed in ogni passo andrebbero calcolate. I metodi più efficienti di ordine $p \geq 2$ si costruiscono scegliendo come Φ una combinazione lineare di valori della $f(x, y)$. Per esempio nel caso di $p = 2$

$$\begin{cases} \Phi(x^*, y^*, h) = a_1k_1 + a_2k_2 \\ k_1 = f(x^*, y^*) \quad k_2 = f(x^* + b_2h, y^* + b_2hk_1) \end{cases}$$

I coefficienti a_1, a_2, b_2 sono scelti in modo tale che lo sviluppo di Taylor di $\tau(x^*, y^*, h)$ nell'intorno (x^*, y^*, h) inizi con la potenza di h di ordine più elevato possibile. Sviluppiamo $\Phi(x^*, y^*, h)$ nell'intorno del punto $(x^*, y^*, 0)$:

$$\begin{aligned} \Phi(x^*, y^*, h) &= a_1f(x^*, y^*) + a_2[f(x^*, y^*) + hb_2f_x(x^*, y^*) + b_2hf_y(x^*, y^*)f(x^*, y^*)] + o(h^2) = \\ &= (a_1 + a_2)f(x^*, y^*) + a_2hb_2[f_x(x^*, y^*) + f_y(x^*, y^*)f(x^*, y^*)] + o(h^2) \end{aligned}$$

Il metodo avrà ordine $p = 2$ se e solo se risulteranno nulli i coefficienti delle potenze h^0 e h^1 dello sviluppo di Taylor dell'errore locale di troncamento

$\tau(x^*, y^*, h)$ quindi confrontiamo lo sviluppo di Taylor di $\Phi(x^*, y^*, h)$ con lo sviluppo della funzione:

$$\frac{u(x^* + h) - u(x^*)}{h} = \frac{1}{h}[u(x^*) + hu'(x^*) + \frac{h^2}{2}u''(x^*) + \frac{h^3}{3!}u'''(\eta) - u(x^*)] =$$

$$u'(x^*) + \frac{h}{2}u''(x^*) + \frac{h^2}{3!}u'''(\eta)$$

Poichè $u''(x^*) = f_x(x^*, y^*) + f_y(x^*, y^*)f(x^*, y^*)$ si ha che:

$$\begin{cases} a_1 + a_2 = 1 \\ a_2 b_2 = \frac{1}{2} \end{cases}$$

Quindi il metodo avrà ordine 2 per tutte le terne a_1, a_2, b_2 che soddisfano il sistema. Una soluzione possibile è: $a_1 = a_2 = \frac{1}{2}$ e $b_2 = 1$. Con questi valori si ottiene il metodo di Heun.

$$\begin{cases} y_{n+1} = y_n + \frac{h}{2}(k_1 + k_2) \\ k_1 = f(x_n, y_n) \quad ; \quad k_2 = f(x_n + h, y_n + hk_1) \end{cases}$$

cioè:

$$y_{n+1} = y_n + \frac{h}{2}[f(x_n, y_n) + f(x_n + h, y_n + hf(x_n, y_n))]$$

Un'altra possibile soluzione è la seguente: $a_1 = 0$ $a_2 = 1$ $b_2 = \frac{1}{2}$ Si ottiene così il metodo di Eulero modificato

$$\begin{cases} y_{n+1} = y_n + hk_2 \\ k_1 = f(x_n, y_n) \quad ; \quad k_2 = f(x_n + \frac{h}{2}, y_n + \frac{h}{2}k_1) \end{cases}$$

cioè:

$$y_{n+1} = y_n + h[f(x_n + \frac{h}{2}, y_n + \frac{h}{2}f(x_n, y_n))]$$

Per costruire metodi di ordine superiore è sufficiente generalizzare quanto fatto e definire Φ nel modo seguente:

$$\left\{ \begin{array}{l} y_{n+1} = y_n + h \sum_{i=1}^r a_i k_i \\ k_1 = f(x^*, y^*) \quad ; \quad k_i = f(x^* + b_i h, y^* + h \sum_{j=1}^{i-1} c_{ij} k_j) \quad i = 2, 3, \dots, r \end{array} \right.$$

Questo metodo risulta consistente se

$$\sum_{i=1}^r a_i = 1$$

di solito si impone la condizione:

$$\sum_{j=1}^{i-1} c_{ij} = b_i$$

Il generico metodo di Runge-Kutta a r stadi si identifica anche con la seguente tabella:

0						
b_2	c_{11}					
b_3	c_{31}	c_{32}				
b_r	c_{r1}	c_{r2}		$c_{r(r-1)}$		
	a_1	a_2		a_{r-1}	a_r	

Fissato il numero di stadi $r \geq 1$ i parametri che compaiono nella tabella sono determinati in modo che lo sviluppo di Taylor di τ inizi con il termine di ordine più elevato possibile. Se $p^*(r)$ è l'ordine massimo è stato dimostrato che:

$$p^*(r) = r \quad r = 1, 2, 3, 4$$

$$p^*(r) = r - 1 \quad r = 5, 6, 7$$

$$p^*(r) = r - 2 \quad r = 8, 9$$

$$p^*(r) \leq r - 3 \quad r \geq 10$$

Il metodo di Runge-Kutta più noto è del quarto ordine ed è il seguente:

0				
$\frac{1}{2}$	$\frac{1}{2}$			
$\frac{1}{2}$	0	$\frac{1}{2}$		
1	0	0	1	
	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

cioè:

$$y_{n+1} = y_n + \frac{h}{6}[k_1 + 2k_2 + 2k_3 + k_4]$$

essendo:

$$k_1 = f(x_n, y_n)$$

$$k_2 = f\left(x_n + \frac{h}{2}, y_n + \frac{h}{2}k_1\right)$$

$$k_3 = f\left(x_n + \frac{h}{2}, y_n + \frac{h}{2}k_2\right)$$

$$k_4 = f(x_n + h, y_n + hk_3)$$

4 Considerazioni sugli errori

Nel trattare i metodi di approssimazione numerica one step e in particolare quello di Eulero si era già detto che l'errore è dovuto ad un errore di discretizzazione locale perchè si applica una formula che approssima la funzione cercata e ad un errore trasmesso perchè al passo $n + 1$ -esimo si usa y_n che è a sua volta affetta da un errore. La formula per l'errore di discretizzazione già ricavata era:

$$e_{n+1} = \underbrace{e_n + h[\Phi(x_n, y(x_n), h) - \Phi(x_n, y_n, h)]}_{\text{errore trasmesso}} + \underbrace{h\tau_{n+1}}_{\text{errore di discretizzazione locale}}$$

e nel caso in cui $f(x, y)$ sia Lipshitziana rispetto a y e il metodo sia consistente si è trovato una maggiorazione per l'errore:

$$|e_n| \leq e^{(\beta-\alpha)L}(|e_0| + \frac{\tau}{L})$$

Tale errore non dipende dal numero di passi n e se $e_0 = 0$ cioè se al primo passo non c'è errore e $\tau \rightarrow 0$ se $h \rightarrow 0$, cioè il metodo è consistente, allora il metodo converge, nel senso che si può dopo un numero finito di passi trovare la soluzione con la precisione voluta. Nel considerare il problema degli errori è necessario anche tenere conto del fatto che i computer usano un'aritmetica finita e quindi tutti i valori delle funzioni tabulate sono affetti da errore di arrotondamento quindi il valore y_n al passo n non è esatto ma affetto da un errore di arrotondamento fatto dalla macchina; chiamiamo tale valore Y_n . L'errore totale è allora dato da:

$$|E_n| = |y(x_n) - Y_n| \leq |y(x_n) - y_n| + |y_n - Y_n|$$

L'errore $|y(x_n) - y_n|$ è già stato considerato quindi occupiamoci di $|y_n - Y_n| \equiv U_n$. Troviamo una maggiorazione per l'errore di arrotondamento.

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

$$Y_{n+1} = Y_n + h\Phi(x_n, Y_n, h) + \underbrace{\rho_{n+1}}_{\text{errore di arrotondamento al passo } n+1}$$

$$U_1 \leq \rho_0 + hL\rho_0 + \rho_1 = (1 + hL)\rho_0 + \rho_1$$

$$U_2 \leq (1 + hL)U_1 + \rho_2 = (1 + hL)^2\rho_0 + (1 + hL)\rho_1 + \rho_2$$

iterando si ha:

$$U_n \leq (1 + hL)^n \rho_0 + \frac{\rho}{hL} [(1 + hL)^n - 1] \leq e^{nhL} [\rho_0 + \frac{\rho}{hL}] = e^{(\beta-\alpha)L} [\rho_0 + \frac{\rho}{hL}]$$

essendo $\rho = \max_i \rho_i$.

Quindi l'errore totale é

$$|E_n| \leq e^{(\beta-\alpha)L} (|e_0| + \frac{\tau}{L} + \rho_0 + \frac{\rho}{hL}) = e^{(\beta-\alpha)L} (|e_0| + \rho_0 + \frac{1}{L} (\tau + \frac{\rho}{h}))$$

Quindi se un metodo a un passo é di ordine p l'errore totale é infinitesimo di ordine p ovvero:

$$|E_n| = O(h^p)$$

se e solo se

$$|\rho_0| = O(h^p)$$

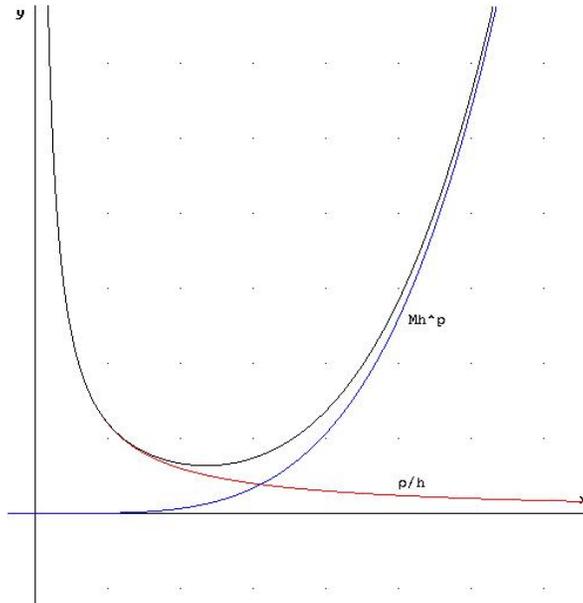
e

$$\rho = O(h^{p+1})$$

Vediamo ora un'analisi grafica dell'errore totale per capire l'andamento. Se il metodo é di ordine p si puó supporre che $\tau = Mh^p$ con M costante opportuna, quindi l'andamento di $|E_n|$ é uguale a quello della funzione:

$$z(h) = Mh^p + \frac{\rho}{h}$$

essendo $p \geq 1$.



Mettendo in evidenza il contributo dei due addendi Mh^p e $\frac{\rho}{h}$ si vede che un valore di h troppo "piccolo" scelto con l'intenzione di ridurre l'errore di troncamento locale può produrre un elevato valore dell'errore di arrotondamento dato da ρ . Quindi fissato il metodo che si vuole usare e quindi l'ordine di consistenza e la precisione con cui si fanno i calcoli e avendo una stima di M si può anche avere una stima del valore ottimale h_M che minimizza $|E_n|$. h_M è molto vicino a \bar{h} ascissa del punto di incontro delle due curve:

$$\bar{h} = \sqrt[p+1]{\frac{\rho}{M}}$$

e

$$h_M = \sqrt[p+1]{\frac{\rho}{Mp}}$$

Si osservi che i due valori coincidono per $p = 1$.

Osservazione La consistenza è una proprietà riguardante la forma del metodo e garantisce che lo schema scelto per approssimare la funzione $f(x, y)$

ovvero la scelta di Φ sia una buona approssimazione del problema. La convergenza invece riguarda il comportamento per $h \rightarrow 0$ della soluzione che il metodo fornisce. il metodo è convergente se per tutti i problemi di Cauchy con $f \in C^1$ con derivate limitate si ha: $\lim_{h \rightarrow 0} e_i = 0$

Osservazione Se un metodo è di ordine p e la precisione aritmetica che si usa è tale che $\rho_0 = O(h^p)$ e $\rho = O(h^{p+1})$ l'errore totale risulta dell'ordine h^p ma rimane tale anche se si aumenta la precisione dei calcoli e quella del valore iniziale, quindi se con lo stesso passo si vuole un risultato più accurato non serve aumentare la precisione ma occorre usare un metodo di ordine maggiore.