Capitolo 4

Metodi lineari multistep

In questo capitolo trattiamo i metodi lineari multistep *per la risoluzione numerica approssimata di problemi ai valori iniziali per equazioni differenziali ordinarie. Di questi metodi studieremo la convergenza e l'analisi di stabilità lineare, arrivando a enunciare le cosiddette* "barriere di Dahlquist".

Occupiamoci, ora, della risoluzione numerica di problemi ai valori iniziali del primo ordine per equazioni differenziali ordinarie:

$$y'(t) = f(t, y(t)), \qquad t \in [t_0, T], \qquad y(t_0) = y_0. \tag{4.1}$$

Senza perdere in generalità, assumeremo per semplicità che y e f siano funzioni scalari. Tuttavia, tutto quello che vedremo si estende in modo naturale al caso di sistemi di equazioni differenziali. Assumeremo inoltre che la funzione f soddisfi le usuali condizioni di continuità e Lipschitzianità rispetto alla y, ed assumeremo che la soluzione del problema (4.1) esista e sia unica. Dal punto di vista concettuale, la risoluzione numerica del problema (4.1) avviene in tre fasi:

1. definizione di un dominio discreto $\{t_n\}$;
2. definizione di un problema discreto definito su tale dominio;
3. risoluzione del problema discreto.

Per quanto riguarda il primo punto, definiremo un dominio discreto assai semplice:

$$t_n = t_0 + nh, \qquad n = 0, \ldots, N, \qquad h = \frac{T - t_0}{N}, \tag{4.2}$$

che costituisce quella che è usualmente denominata *mesh uniforme* con *passo di discretizzazione (o integrazione)* o, più semplicemente, *passo*, h.

Per quanto riguarda il secondo punto considereremo, definendo y_n la approssimazione a $y(t_n)$ e

$$f_n \equiv f(t_n, y_n), \tag{4.3}$$

problemi discreti nella forma:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = h \sum_{i=0}^{k} \beta_i f_{n+i}, \qquad n = 0, \ldots, N - k, \tag{4.4}$$

dove i coefficienti $\{\alpha_i\}$ e $\{\beta_i\}$ definiscono il particolare *metodo lineare multistep*, o *linear multistep formula (LMF)*, a k passi. Per la (4.4) si supporranno note le *condizioni iniziali*

$$y_0, \ldots, y_{k-1}. \tag{4.5}$$

In altri termini, approssimiamo il problema *continuo* ai valori iniziali del primo ordine (4.1) con il problema *discreto* ai valori iniziali di ordine k (4.4)-(4.5).

Per quanto riguarda il terzo punto, osserviamo che la (4.4) è definita a meno di una costante moltiplicativa nonnulla. Questa può essere determinata, imponendo la normalizzazione di un coefficiente. Nel nostro caso, imporremo la normalizzazione

$$\alpha_k = 1. \tag{4.6}$$

Pertanto, la (4.4) può essere riscritta come

$$y_{n+k} - h\beta_k f_{n+k} = -\sum_{i=0}^{k-1} \alpha_i y_{n+i} + h \sum_{i=0}^{k-1} \beta_i f_{n+i}, \qquad n = 0, \ldots, N-k. \tag{4.7}$$

È evidente che, note le k condizioni iniziali (4.5), si determineranno $y_k, y_{k+1}, \ldots$, dalla (4.7). In particolare, se $\beta_k = 0$, la (4.7) diviene:

$$y_{n+k} = -\sum_{i=0}^{k-1} \alpha_i y_{n+i} + h \sum_{i=0}^{k-1} \beta_i f_{n+i}, \qquad n = 0, \ldots, N-k.$$

Ovvero, si ottiene la nuova approssimazione direttamente mediante delle semplici sostituzioni in avanti. Si parla, in questo caso, di *metodo esplicito.*

Diversamente, ricordando la definizione (4.3) di f_n, y_n si otterrà dalla (4.7) risolvendo, ad ogni passo, un'equazione (generalmente) nonlineare. In questo caso, il metodo si dirà *implicito.*

Osservazione 4.1 *È evidente che l'implementazione di un metodo esplicito è assai più semplice di quella di un metodo implicito.*

Concludiamo questa sezione definendo i seguenti polinomi:

$$\rho(z) = \sum_{i=0}^{k} \alpha_i z^i, \qquad \sigma(z) = \sum_{i=0}^{k} \beta_i z^i, \tag{4.8}$$

denominati, rispettivamente, *primo* e *secondo polinomio caratteristico* associati al metodo LMF (4.4). In questo modo, quest'ultimo può essere rappresentato, in forma più compatta, come:

$$\rho(E)y_n - h\sigma(E)f_n = 0, \qquad n = 0, \ldots, N-k, \tag{4.9}$$

dove E è, al solito, l'operatore di *shift.* Poiché i due polinomi in (4.8) caratterizzano il dato metodo LMF, talora faremo riferimento a quest'ultimo come al *metodo* (ρ, σ).

4.1 Ordine di un metodo LMF

In generale, la soluzione della (4.4) differirà da quella del problema continuo (4.1) (chiaramente, proiettata sulla *mesh* (4.2)). È tuttavia lecito richiedere che i due problemi siano *vicini* nel seguente senso: sia assegnata la soluzione esatta $y(t)$ valutata ai nodi della *mesh* (4.2), $\{y(t_n)\}$, e denotiamo con

$$\hat{f}_n \equiv f(t_n, y(t_n)). \tag{4.10}$$

Definizione 4.1 *Diremo che il metodo ha* ordine p *se, per* $h \to 0^+$,

$$\sum_{i=0}^{k} \alpha_i y(t_{n+i}) - h \sum_{i=0}^{k} \beta_i \hat{f}_{n+i} \equiv \tau_{n+k} = O(h^{p+1}), \qquad n = 0, \ldots, N-k, \tag{4.11}$$

dove τ_n *è detto* errore locale di troncamento. *In particolare, il metodo si dirà* consistente *se* $p \geq 1$.

In altri termini, si richiede che il *residuo* ottenuto inserendo la soluzione continua $\{y(t_n)\}$ nell'equazione discreta sia un infinitesimo di opportuno ordine del passo h. L'ordine di un metodo LMF si traduce in condizioni algebriche sui suoi coefficienti, che andiamo ora ad esaminare. A questo fine, supporremo nel seguito che $y(t)$ sia, per semplicità di esposizione, sviluppabile in serie di Taylor in t_0. Vale dunque il seguente risultato.

Teorema 4.1 *Il metodo LMF (4.4) ha ordine* p *se:*

$$\sum_{i=0}^{k} \alpha_i = 0, \qquad \sum_{i=0}^{k} \left(i^j \alpha_i - j i^{j-1} \beta_i \right) = 0, \qquad j = 1, \ldots, p. \tag{4.12}$$

Dimostrazione. Dalla (4.11) si ottiene, tenendo conto delle (4.1)-(4.2):

$$\begin{aligned} \sum_{i=0}^{k} \alpha_i y(t_{n+i}) - h \sum_{i=0}^{k} \beta_i \hat{f}_{n+i} &= \sum_{i=0}^{k} \alpha_i y(t_{n+i}) - h \sum_{i=0}^{k} \beta_i y'(t_{n+i}) \\ &= \sum_{i=0}^{k} \alpha_i \sum_{j \geq 0} \frac{y^{(j)}(t_n)}{j!} i^j h^j - \sum_{i=0}^{k} \beta_i \sum_{j \geq 1} \frac{y^{(j)}(t_n)}{(j-1)!} i^{j-1} h^j \\ &= y(t_n) \sum_{i=0}^{k} \alpha_i + \sum_{j \geq 1} \frac{y^{(j)}(t_n)}{j!} h^j \sum_{i=0}^{k} \left(i^j \alpha_i - j i^{j-1} \beta_i \right). \end{aligned} \tag{4.13}$$

Imponendo che le potenze di h^j, $j = 0, 1, \ldots, p$, si annullino, si ottengono quindi le condizioni di ordine (4.12). □

Come semplice corollario, si ottiene che il metodo è consistente se ha ordine almeno 1. Ovvero se:

$$\sum_{i=0}^{k} \alpha_i = 0, \qquad \sum_{i=0}^{k} (i\alpha_i - \beta_i) = 0. \tag{4.14}$$

Le condizioni di consistenza (4.14) possono essere riscritte, tenendo conto dei polinomi (4.8), come:

$$\rho(1) = 0, \qquad \rho'(1) = \sigma(1). \tag{4.15}$$

Vale inoltre il seguente risultato.

Teorema 4.2 *L'ordine massimo di un LMF a k passi è $2k$, se il metodo è implicito, o $2k-1$, se il metodo è esplicito.*

Dimostrazione. In virtù delle (4.14), i coefficienti di un metodo di ordine p soddisferanno l'equazione

$$M_p \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_k \\ -\beta_0 \\ \vdots \\ -\beta_k \end{pmatrix} = \mathbf{0},$$

in cui

$$M_p = \begin{pmatrix} 1 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & k & 1 & \dots & 1 \\ 0^2 & \dots & k^2 & 2\cdot 0 & \dots & 2\cdot k \\ \vdots & & \vdots & 3\cdot 0^2 & \dots & 3\cdot k^2 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0^p & \dots & k^p & p\cdot 0^{p-1} & \dots & p\cdot k^{p-1} \end{pmatrix} \in \mathbb{R}^{p+1\times 2k+2}.$$

La tesi discende quindi dal fatto che M_p ha sempre rango massimo e, pertanto, $p \leq 2k$, se il metodo è implicito, avendo fissato la normalizzazione (4.6). Se il metodo è esplicito, invece, $p \leq 2k-1$, perchè dovrà anche valere $\beta_k = 0$. È evidente che i metodi di ordine massimo, rispettivamente $p = 2k$ e $p = 2k-1$, esistono e sono unici. □

Esercizio 4.1 *Determinare i coefficienti del metodo LMF implicito a k passi (4.4)-(4.6) di ordine $2k$.*

Esercizio 4.2 *Determinare i coefficienti del metodo LMF esplicito a k passi (4.4)-(4.6) di ordine $2k-1$.*

Osservazione 4.2 *Dalle (4.11)-(4.12) segue che:*

$$\tau_{n+k} = c_{p+1} y^{(p+1)}(t_n) h^{p+1} + O(h^{p+2}),$$

dove

$$c_{p+1} = \frac{1}{(p+1)!} \sum_{i=0}^{k} \left[i^{p+1}\alpha_i - (p+1) i^p \beta_i \right]$$

è denominato coefficiente principale *dell'errore.*

Osservazione 4.3 *La denominazione* "locale" *dell'*errore di troncamento τ_n *(vedi (4.11)) deriva dal fatto che (si supponga, per semplicità, che il metodo sia esplicito), facendo l'*ipotesi "locale" [1]

$$y_{n+i} = y(t_{n+i}), \qquad i = 0, \dots, k-1, \tag{4.16}$$

e detto $\tilde{y}_{n+k}$ il nuovo punto ottenuto dalla (4.4), allora $y(t_{n+k}) - \tilde{y}_{n+k} = \tau_{n+k}$. In altre parole, τ_{n+k} può essere interpretato come l'errore che si commetterebbe nel singolo passo di

[1] Questo significa supporre esatte le condizioni iniziali del corrente passo di integrazione.

integrazione, *sotto l'*ipotesi locale *(4.16). Nella prossima sezione esamineremo in maggior dettaglio la relazione esistete tra l'*errore globale,

$$e_n = y(t_n) - y_n, \qquad n = 0, \ldots, N,$$

ed i corrispondenti errori locali $\tau_k, \ldots, \tau_N$.

4.2 Convergenza

Osservando che la soluzione discreta $\{y_n\}$ soddisfa la (4.4), mentre la soluzione del problema (4.1), proiettata sulla *mesh* (4.2), $\{y(t_n)\}$, soddisfa la (4.11), è naturale richiedere che la prima *converga* alla seconda, quando $N \to \infty$, ovvero, quando il passo di discretizzazione

$$h = \frac{T - t_0}{N} \to 0. \tag{4.17}$$

Per questo motivo, si da la seguente definizione.

Definizione 4.2 *Il metodo (4.4) si dirà* convergente *se, definito l'*errore [2]

$$e_n = y(t_n) - y_n, \qquad n = 0, \ldots, N, \tag{4.18}$$

si ha:

$$\lim_{N\to\infty} \max_{n=0,\ldots,N} \|e_n\| = 0.$$

Osserviamo che l'errore (4.18) soddisferà l'*equazione dell'errore*, ottenuta sottraendo la (4.4) dalla (4.11):

$$\rho(E)e_n - h\sigma(E)(\hat{f}_n - f_n) = \tau_n, \qquad n = 0, \ldots, N-k. \tag{4.19}$$

Poichè $h \to 0$ e $\tau_n = O(h^{p+1})$, con $p \geq 1$ se il metodo è consitente, ha senso studiare la stabilità della soluzione nulla dell'equazione

$$\rho(E)e_n = 0, \qquad n = 0, \ldots, N-k. \tag{4.20}$$

Questa si ottiene direttamente applicando il metodo (4.4) all'equazione $y' = 0$, la cui soluzione è costante. Osserviamo che $\rho(z)$ non può essere un polinomio di Schur, a causa della proprietà di consistenza (4.15). [3] Può tuttavia essere un polinomio di Von Neumann, il che garantisce la stabilità della soluzione nulla per la (4.20). Si capisce, quindi, il senso della seguente definizione.

Definizione 4.3 *Un metodo LMF* (ρ, σ) *si dirà* 0-stabile *se* $\rho(z)$ *è un polinomio di Von Neumann.*

La proprietà di 0-stabilità è importante, in quanto è possibile dimostrare il seguente risultato.

Teorema 4.3 *Siano* $f \in C^{(p+1)}$ *ed il metodo* (ρ, σ) *0-stabile. Se, inoltre,*

[2] Talora l'errore è denominato anche *errore globale*, per distinguerlo dall'errore *locale* di troncamento τ_n definito in precedenza.
[3] Se così non fosse, si avrebbe $\tau_n = O(h^0)$.

- $\tau_n = O(h^{p+1}) \quad n = k, \ldots, N$, *e*
- $e_i = O(h^r), \quad i = 0, \ldots, k-1$,

allora:

$$|e_n| \leq O(h^{\min\{r,p\}}), \qquad n = k, \ldots, N.$$

Dimostrazione. Si veda la Sezione 4.2.1. □

In altre parole, se un metodo di ordine p è 0-stabile, e le condizioni iniziali hanno accuratezza $O(h^p)$, allora l'*errore globale* e_n sarà a sua volta $O(h^p)$. Come semplice corollario, si ottiene il seguente risultato.

Corollario 4.1 *Un metodo LMF che sia consistente e 0-stabile è convergente.*

Osservazione 4.4 *Evidentemente, il problema continuo (4.1) fornisce solo la condizione iniziale y_0. Le altre condizioni iniziali in (4.5) possono essere approssimate, entro l'accuratezza desiderata, ad esempio mediante sviluppi in serie in $t = t_0$.*[4]

La 0-stabilità è, in virtù del Teorema 4.3, una proprietà essenziale per un metodo LMF. Tuttavia è altresì evidente che richiedere che il polinomio $\rho(z)$ associato ad un metodo sia un polinomio di Von Neumann, introduce dei vincoli sui suoi coefficienti che, pertanto, non potranno essere utilizzati per incrementare l'ordine del metodo. Questo argomento, che abbiamo esposto in modo intuitivo, si concretizza nel segente risultato, di cui si riporta il solo enunciato.

Teorema 4.4 (prima barriera di Dahlquist) *L'ordine massimo di un metodo LMF a k passi che sia 0-stabile è:*

- $k+1$, *se k è dispari;*
- $k+2$, *se k è pari.*

4.2.1 Dimostrazione della convergenza di un metodo LMF

In questa sezione dimostriamo il Teorema 4.3. Per fare questo, riscriviamo il problema discreto (4.3)-(4.4) e quello perturbato (4.10)-(4.11), rispettivamente, in forma vettoriale come segue:

$$A_N \mathbf{y}_N - hB_N \mathbf{f}_N = \boldsymbol{\eta}_N, \qquad A_N \hat{\mathbf{y}}_N - hB_N \hat{\mathbf{f}}_N = \hat{\boldsymbol{\eta}}_N + \boldsymbol{\tau}_N, \tag{4.21}$$

con

$$A_N = \begin{pmatrix} \alpha_k & & & \\ \vdots & \ddots & & \\ \alpha_0 & & \ddots & \\ & \ddots & & \ddots \\ & & \alpha_0 \quad \ldots \quad & \alpha_k \end{pmatrix}, \; B_N = \begin{pmatrix} \beta_k & & & \\ \vdots & \ddots & & \\ \beta_0 & & \ddots & \\ & \ddots & & \ddots \\ & & \beta_0 \quad \ldots \quad & \beta_k \end{pmatrix} \in \mathbb{R}^{N-k+1 \times N-k+1}, \tag{4.22}$$

[4] In alternativa, si possono utilizzare dei metodi *one-step* di opportuno ordine, come vedremo successivamente nel Capitolo 9.

$$\mathbf{y}_N = \begin{pmatrix} y_k \\ \vdots \\ y_N \end{pmatrix}, \ \hat{\mathbf{y}}_N = \begin{pmatrix} y(t_k) \\ \vdots \\ y(t_N) \end{pmatrix}, \ \boldsymbol{\tau}_N = \begin{pmatrix} \tau_k \\ \vdots \\ \tau_N \end{pmatrix} \in \mathbb{R}^{N-k+1}, \tag{4.23}$$

$$\mathbf{f}_N = \begin{pmatrix} f_k \\ \vdots \\ f_N \end{pmatrix}, \ \hat{\mathbf{f}}_N = \begin{pmatrix} \hat{f}_k \\ \vdots \\ \hat{f}_N \end{pmatrix} \in \mathbb{R}^{N-k+1}, \tag{4.24}$$

e, infine,

$$\boldsymbol{\eta}_N = \begin{pmatrix} \sum_{i=0}^{k-1}[h\beta_i f_i - \alpha_i y_i] \\ \vdots \\ h\beta_0 f_{k-1} - \alpha_0 y_{k-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \hat{\boldsymbol{\eta}}_N = \begin{pmatrix} \sum_{i=0}^{k-1}[h\beta_i \hat{f}_i - \alpha_i y(t_i)] \\ \vdots \\ h\beta_0 \hat{f}_{k-1} - \alpha_0 y(t_{k-1}) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{N-k+1}. \tag{4.25}$$

Le matrici A_N e B_N definite in (4.22) sono matrici triangolari inferiori, a banda, di *Toeplitz.*

Definizione 4.4 *Una matrice di Toeplitz ha elementi costanti lungo le sue diagonali:*

$$T = (a_{ij}) \in \mathbb{R}^{N\times N} \text{ è di Toeplitz} \quad \Leftrightarrow \quad a_{ij} = a_{j-i}, \quad \forall i, j = 1, \ldots, N,$$

dove le costanti

$$a_{1-N}, \ldots, a_0, \ldots, a_{N-1}$$

sono gli elementi (costanti) lungo le corrispondenti diagonali, di indice $1-N, \ldots, 0, \ldots, N-1$, *della matrice* T. *In questo caso, si utilizza la convenzione che l'indice denota:*

- *una* sottodiagonale *se negativo,*
- *la* diagonale principale *se nullo,*
- *una* sopradiagonale *se positivo.*

Nella trattazione seguente, avremo altresì bisogno di un'altra tipologia di matrici: le *M-matrici.*

Definizione 4.5 *Una matrice del tipo*

$$M = cI - B \in \mathbb{R}^{n\times n}, \qquad B \geq O, \qquad c > \rho(B),^{5}$$

è detta M-matrice.

[5] Come usuale, $\rho(B)$ denota il *raggio spettrale* della matrice B.

Si verificano facilmente le seguenti proprietà di una M-matrice (farlo per esercizio nel caso, più semplice, in cui B sia strettamente triangolare. Il caso generale potrà essere dimostrato dopo che avremo visto le *funzioni di matrice*):

$$M \text{ è nonsingolare,} \tag{4.26}$$
$$(M^{-1})_{ij} \geq 0, \quad \text{per ogni } i,j, \tag{4.27}$$
$$(M)_{ij} \leq 0, \quad \text{se } i \neq j, \tag{4.28}$$
$$(M)_{ii} > 0. \quad \text{per ogni } i. \tag{4.29}$$

Le M-matrici sono, inoltre, particolari matrici *monotone*:

Definizione 4.6 *Una matrice con inversa non negativa è detta* matrice monotona.

Osservazione 4.5 *Le matrici monotone sono utili nella risoluzione di diseguaglianze lineari. Infatti, se*

$$A\mathbf{x} \leq \mathbf{b},$$

ed $A^{-1} \geq 0$, *allora*

$$\mathbf{x} \leq A^{-1}\mathbf{b}.$$

Inoltre, se A *e* B *sono monotone, allora:*

$$A \leq B \quad \Rightarrow \quad 0 \leq B^{-1} \leq A^{-1}. \tag{4.30}$$

La matrice A_N definita nella (4.22) è nonsingolare, in quanto $\alpha_k \neq 0$ (per la normalizzazione (4.6)). Inoltre, definendo i polinomi (vedi (4.8))

$$\hat{\rho}(z) \equiv z^k \rho(z^{-1}) = \sum_{i=0}^{k} \alpha_i z^{k-i}, \qquad \hat{\sigma}(z) \equiv z^k \sigma(z^{-1}) = \sum_{i=0}^{k} \beta_i z^{k-i},$$

le due matrici A_N e B_N possono essere scritte come:

$$A_N = \hat{\rho}(H_N), \qquad B_N = \hat{\sigma}(H_N), \qquad \text{con} \qquad H_N = \begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix} \in \mathbb{R}^{N-k+1 \times N-k+1}.$$

Osserviamo che H_N^j è una matrice con tutti gli elementi nulli, tranne quelli sulla j-esima sottodiagonale, che sono uguali a 1 (e, quindi, $H_N^{N-k+1} = O$). Pertanto, una matrice di Toeplitz triangolare inferiore di dimensione $N-k+1$ sarà *sempre* scrivibile nella forma $p(H_N)$, per un opportuno polinomio $p \in \Pi_{N-k}$. Valgono le seguenti proprietà.

Lemma 4.1 *Il prodotto di due matrici triangolari inferiori (superiori) di Toeplitz è una matrice triangolare interiore (superiore) di Toeplitz.*

Dimostrazione. Siano $T_N = p(H_N)$ e $U_N = q(H_N)$ due matrici di Toeplitz triangolari inferiori (per quelle superiori è sufficiente considerare le trasposte), con H_N definita come sopra. Si ottiene che $T_N U_N = p(H_N)q(H_N) = c(H_N)$, dove $c = p \cdot q$, che è una matrice di Toeplitz triangolare inferiore. □

Lemma 4.2 *A_N^{-1} è una matrice di Toeplitz triangolare inferiore.*

Dimostrazione. Chiaramente, A_N^{-1} è triangolare inferiore. Dimostriamo che è anche una matrice di Toeplitz. A questo fine, denotiamo

$$x_{1-j}, \ldots, x_0, \ldots, x_{N-k-j+1}$$

le componenti in colonna j di tale matrice. Esse saranno date da

$$x_i = 0, \quad i < 0, \qquad x_0 = \alpha_k^{-1}$$

Le rimanenti componenti si otterranno per ricorrenza come:

$$x_\ell = -\alpha_k^{-1} \sum_{i=0}^{k-1} \alpha_i x_{\ell-k+i}, \qquad \ell \geq 1.$$

La tesi si completa osservando che

$$x_\ell, \qquad \ell = 0, 1, \ldots,$$

altri non è che l'elemento che si trova sulla sottodiagonale ℓ-esima, indipendentemente dall'indice di colonna j considerato.□

Lemma 4.3 *Se $\rho(z)$ è un polinomio di Von Neumann, gli elementi di A_N^{-1} sono uniformemente limitati rispetto a N.*

Dimostrazione. Essendo A_N^{-1} una matrice di Toeplitz, dal precedente Lemma 4.2, è sufficiente considerare gli elementi sulla sua prima colonna. Questi ultimi soddisferanno l'equazione alle differenze

$$\sum_{i=0}^{k} \alpha_i x_{n+i-k} = 0, \qquad n \geq 0,$$

con le condizioni iniziali

$$x_n = 0, \quad n < 0, \qquad x_0 = \alpha_k^{-1}.$$

Poiché il polinomio caratteristico dell'equazione alle differenze è $\rho(z)$, che è un polinomio di Von Neumann, gli elementi della successione $\{x_n\}_{n\geq 0}$ risulteranno essere uniformemente limitati rispetto ad n, essendo l'origine stabile. □

Lemma 4.4 *Sia data la matrice strettamente triangolare inferiore*

$$F_N = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ \vdots & \ddots & \ddots & \\ 1 & \ldots & 1 & 0 \end{pmatrix} \in \mathbb{R}^{N\times N}. \tag{4.31}$$

Allora, per ogni $j \geq 0$:

$$F_N^j = \frac{1}{(j-1)!} \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 \\ 0^{(j-1)} & 0 & \ldots & \ldots & 0 \\ 1^{(j-1)} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ (N-2)^{(j-1)} & \ldots & 1^{(j-1)} & 0^{(j-1)} & 0 \end{pmatrix}.$$

Dimostrazione. La tesi è ovviamente vera per $j = 1$. Si supponga vera per j e, per $j+1$, segue:

$$
\begin{aligned}
F_N^{j+1} &= F_N F_N^j \\
&= \frac{1}{(j-1)!}\begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ \vdots & \ddots & \ddots & \\ 1 & \dots & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 0^{(j-1)} & 0 & \dots & \dots & 0 \\ 1^{(j-1)} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ (N-2)^{(j-1)} & \dots & 1^{(j-1)} & 0^{(j-1)} & 0 \end{pmatrix} \\
&= \frac{1}{(j-1)!}\begin{pmatrix} 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 & 0 \\ 0^{(j-1)} & 0 & & & \vdots & \vdots \\ 0^{(j-1)} + 1^{(j-1)} & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \sum_{i=0}^{N-3} i^{(j-1)} & \dots & 0^{(j-1)} + 1^{(j-1)} & 0^{(j-1)} & 0 & 0 \end{pmatrix} \\
&= \frac{1}{j!}\begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 0^{(j)} & 0 & \dots & \dots & 0 \\ 1^{(j)} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ (N-2)^{(j)} & \dots & 1^{(j)} & 0^{(j)} & 0 \end{pmatrix}. \square
\end{aligned}
$$

Lemma 4.5 *Sia data la matrice F_N definita in (4.31). Allora, considerando la norma 1 o ∞, per ogni $\mu \in \mathbb{C}$:*

$$\left\| \left(I_N - \frac{\mu}{N} F_N \right)^{-1} \right\| \leq \mathrm{e}^{|\mu|}.$$

Dimostrazione. In virtù del Lemma 4.4, si ha $F_N^N = O$ e, inoltre,

$$\|F_N^j\| = \frac{1}{(j-1)!} \sum_{i=0}^{N-2} i^{(j-1)} = \frac{(N-1)^{(j)}}{j!} = \begin{pmatrix} N-1 \\ j \end{pmatrix}.$$

Pertanto:

$$
\begin{aligned}
\left\| \left(I_N - \frac{\mu}{N} F_N \right)^{-1} \right\| &= \left\| \sum_{j=0}^{N-1} \left(\frac{\mu}{N} \right)^j F_N^j \right\| \leq \sum_{j=0}^{N-1} \left(\frac{|\mu|}{N} \right)^j \|F_N^j\| \\
&= \sum_{j=0}^{N-1} \left(\frac{|\mu|}{N} \right)^j \begin{pmatrix} N-1 \\ j \end{pmatrix} = \sum_{j=0}^{N-1} \frac{|\mu|^j}{j!} \frac{(N-1)^{(j)}}{N^j} \leq \sum_{j=0}^{N-1} \frac{|\mu|^j}{j!} \leq \mathrm{e}^{|\mu|}. \square
\end{aligned}
$$

Lemma 4.6 *Per ogni $\varepsilon \geq 0$ e sufficientemente piccolo, la matrice*

$$M_N(\varepsilon) = \begin{pmatrix} (1-\varepsilon/N) & & & \\ -\varepsilon/N & (1-\varepsilon/N) & & \\ \vdots & \ddots & \ddots & \\ -\varepsilon/N & \dots & -\varepsilon/N & (1-\varepsilon/N) \end{pmatrix} \in \mathbb{R}^{N\times N}$$

è una M-matrice e, inoltre,

$$\|M_N^{-1}(\varepsilon)\| \leq 2\mathrm{e}^{2\varepsilon}, \qquad \forall N \geq 1.$$

Dimostrazione. Dimostriamo la tesi assumendo che $0 \leq \varepsilon \leq \frac{1}{2}$. Pertanto, per ogni $N \geq 1$ risulta:

$$1 - \frac{\varepsilon}{N} \geq \frac{1}{2}. \tag{4.32}$$

La tesi è quindi banalmente vera per $N = 1$. Per $N \geq 2$, evidentemente,

$$M_N = (1-\varepsilon/N)I_N - \varepsilon/N F_N,$$

dove F_N è la matrice definita in (4.31). Inoltre, $F_N \geq 0$ e $\rho(F_N) = 0 < (1-\varepsilon/N)$. Pertanto $M_N(\varepsilon)$ è una M-matrice. Dalla (4.32) si ha quindi che:

$$M_N \geq \frac{1}{2}\left[I_N - \frac{2\varepsilon}{N}F_N\right] \equiv G_N,$$

con G_N a sua volta una M-matrice. Pertanto, dalle (4.27)-(4.30) segue che:

$$0 \leq M_N^{-1} \leq G_N^{-1}.$$

Considerando la norma 1 o la norma ∞, si ottiene quindi che, in virtù del Lemma 4.5:[6]

$$\left\|M_N^{-1}\right\| \leq \left\|G_N^{-1}\right\| \leq 2\mathrm{e}^{2\varepsilon}. \square$$

Utilizzando questi risultati, possiamo ora dimostrare il Teorema 4.3.

Dimostrazione del Teorema 4.3. Dalle (4.21)–(4.25), segue che

$$A_N\mathbf{e}_N - hB_N(\hat{\mathbf{f}}_N - \mathbf{f}_N) = (\hat{\boldsymbol{\eta}}_N - \boldsymbol{\eta}_N) + \boldsymbol{\tau}_N,$$

dove

$$\mathbf{e}_N \equiv \hat{\mathbf{y}}_N - \mathbf{y}_N = \begin{pmatrix} e_k \\ \vdots \\ e_N \end{pmatrix}$$

è il vettore degli errori. Pertanto,

$$\mathbf{e}_N - hA_N^{-1}B_N(\hat{\mathbf{f}}_N - \mathbf{f}_N) = A_N^{-1}\left[(\hat{\boldsymbol{\eta}}_N - \boldsymbol{\eta}_N) + \boldsymbol{\tau}_N\right].$$

Denotando con $|\mathbf{e}_N|$ il vettore le cui componenti sono i moduli di $\mathbf{e}_N$, ed utilizzando la stessa notazione per gli altri vettori e per le matrici, si ottiene:

$$|\mathbf{e}_N| \leq h\left|A_N^{-1}\right| |B_N| |\hat{\mathbf{f}}_N - \mathbf{f}_N| + \left|A_N^{-1}\right| (|\hat{\boldsymbol{\eta}}_N - \boldsymbol{\eta}_N| + |\boldsymbol{\tau}_N|). \tag{4.33}$$

Considerato che:

[6] Ponendo, evidentemente, $\mu = 2\varepsilon$

1. $|\hat{\mathbf{f}}_N - \mathbf{f}_N| \leq L|\mathbf{e}_N|$, essendo L la costante di Lipschitz di f;
2. $\left|A_N^{-1}\right| \leq \alpha C_N$, per il risultato del Lemma 4.3, dove α è una maggiorazione, indipendente da N, per il modulo del generico elemento di A_N^{-1}, e

$$C_N = \begin{pmatrix} 1 & & \\ \vdots & \ddots & \\ 1 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{N-k+1 \times N-k+1};$$

3. $\left|A_N^{-1}\right| |B_N| \leq \alpha\beta C_N$, per motivi analoghi al precedente punto 2, avendo posto

$$\beta = \sum_{i=0}^{k} |\beta_i| ;$$

4. $\left|A_N^{-1}\right| |\boldsymbol{\tau}_N| \leq O(h^p)$, in virtù del punto 2, considerando che $Nh = T - t_0$;
5. $\left|A_N^{-1}\right| |\hat{\boldsymbol{\eta}}_N - \boldsymbol{\eta}_N| \leq O(h^r)$, essendo i dati iniziali noti con un errore $O(h^r)$;

dalla (4.33) si ottiene

$$M_N|\mathbf{e}_N| \leq \max\{O(h^r), O(h^p)\} = O(h^{\min\{r,p\}}),$$

dove (vedi (4.2))

$$M_N = I_N - (h\alpha\beta)C_N = I_N - \frac{\alpha\beta(T-t_0)}{N} C_N. \tag{4.34}$$

Pertanto, considerando la norma ∞, per $h = (T - t_0)/N$ sufficientemente piccolo, dal Lemma 4.6 si ottiene, tenendo conto del fatto che $M_N^{-1} \geq 0$:

$$\begin{aligned} \|\mathbf{e}_N\| &\leq \left\|M_N^{-1} O(h^{\min\{r,p\}})\right\| \leq \|M_N^{-1}\| O(h^{\min\{r,p\}}) \\ &\leq 2\mathrm{e}^{2\alpha\beta(T-t_0)} O(h^{\min\{r,p\}}) = O(h^{\min\{r,p\}}). \square \end{aligned}$$

4.3 Alcuni esempi di metodi LMF convergenti

Esaminiamo, in questa sezione, alcune classi di metodi che risultano essere 0-stabili e consistenti e, pertanto, convergenti.

I metodi di Adams

Per questi metodi, il polinomio $\rho(z)$ è della forma

$$\rho(z) = z^{k-1}(z-1).$$

Pertanto esso ha $k-1$ radici nello 0 ed una uguale ad 1. Il metodo è perciò 0-stabile per costruzione. I coefficienti del polinomio $\sigma(z)$ sono scelti in modo da ottenere l'ordine massimo, distinguendo, tuttavia, tra metodi espliciti e impliciti.

I metodi di *Adams-Bashforth* sono espliciti (e pertanto $\beta_k = 0$): essi hanno ordine $p = k$. Ecco alcuni esempi:

- $k = 1$: si ottiene il *metodo di Eulero esplicito*,

$$y_{n+1} - y_n = hf_n, \tag{4.35}$$

che ha ordine 1;

- $k = 2$: si ottiene il *metodo di Adams-Bashforth*

$$y_{n+2} - y_{n+1} = \frac{h}{2}(3f_{n+1} - f_n), \tag{4.36}$$

che ha ordine 2.

I metodi di *Adams-Moulton* sono impliciti ed hanno ordine $p = k + 1$. Vediamone alcuni esempi:

- $k = 1$: si ottiene il *metodo dei trapezi*,

$$y_{n+1} - y_n = \frac{h}{2}(f_{n+1} + f_n), \tag{4.37}$$

che ha ordine 2;

- $k = 2$: si ottiene il *metodo di Adams-Moulton* di ordine 3,

$$y_{n+2} - y_{n+1} = \frac{h}{12}(5f_{n+2} + 8f_{n+1} - f_n). \tag{4.38}$$

Formule di Newton-Cotes

Queste formule si ottengono approssimando il teorema fondamentale del calcolo mediante una formula di quadratura. Esse sono del tipo:

$$y_{n+k} - y_n = h\sum_{i=0}^{k} \beta_i f_{n+i}. \tag{4.39}$$

Il polinomio $\rho(z) = z^k - 1$ è, quindi, un polinomio di Von Neumann, essendo le sue radici le radici k-esime dell'unità. Se il metodo ha almeno ordine 1 sarà, dunque, convergente.

Per $k = 1$ si ottiene la formula dei trapezi vista innanzi. Per $k = 2$ si ottiene il *metodo di Simpson*,

$$y_{n+2} - y_n = \frac{h}{3}(f_{n+2} + 4f_{n+1} + f_n), \tag{4.40}$$

che è implicito e si verifica avere ordine 4. Se si utilizza la regola dei rettangoli, si ottiene invece il *metodo del mid-point*,

$$y_{n+2} - y_n = 2hf_{n+1}, \tag{4.41}$$

che è esplicito ed ha ordine 2.

Backward Differentiation Formulae (BDF)

Questi metodi sono nella forma:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = h\beta_k f_{n+k}, \tag{4.42}$$

con i coefficienti $\{\alpha_i\}$ univocamente determinati dall'imporre l'ordine massimo $p = k$. Questi metodi risultano essere impliciti; sono, inoltre, 0-stabili fino a $k = 6$.[7] Vediamone alcuni esempi:

- $k = 1$: si ottiene il *metodo di Eulero implicito*,

$$y_{n+1} - y_n = hf_{n+1}, \tag{4.43}$$

  che ha ordine 1;

- $k = 2$: si ottiene la *BDF di ordine 2*,

$$y_{n+2} - \frac{4}{3}y_{n+1} + \frac{1}{3}y_n = \frac{2}{3}hf_{n+2}, \tag{4.44}$$

  che si verifica essere 0-stabile, in quanto $\rho(z) = (z-1)(z-\frac{1}{3})$.

4.4 Stabilità per h fissato

Nonostante il risultato negativo dato dalla prima barriera di Dahlquist, abbiamo visto che esistono metodi 0-stabili con un ordine di accuratezza arbitrariamente elevato (e, pertanto, convergenti). Nell'analisi su cui si basa la proprietà di 0-stabilità, tuttavia, si è assunto di poter far tendere a 0 il passo h, aumentando il numero N dei punti della *mesh*. Tuttavia, per alcuni problemi, $T \to \infty$. In altri termini, l'intervallo di integrazione è illimitato. Questo significa che, anche facendo tendere N a ∞, non è più possibile assumere il passo h infinitesimo. Inoltre, a questo argomento si aggiungono anche ovvie considerazioni di efficienza computazionale (oltre che di buon senso) che, in ogni caso, escludono la possibilità di effettuare un numero di passi *infinito*. Infatti, oltre all'aumento del tempo di esecuzione, su calcolatore si accumulano, ad ogni passo, gli errori dovuti all'utilizzo dell'aritmetica finita. Un numero maggiore di passi di integrazione si traduce, pertanto, in un accumulo di questi errori, come illustrato dall'esempio seguente.

Esempio 4.1 *Consideriamo il seguente esempio di* equazione logistica,[8]

$$y' = 10y - 2y^2, \qquad t \geq 0, \qquad y(0) = 1, \tag{4.45}$$

la cui soluzione esatta si verifica essere:

$$y(t) = \frac{5}{1 + 4e^{-10t}} \to 5, \qquad t \to \infty.$$

Se utilizziamo il metodo del mid-point (4.41), che abbiamo visto essere 0-stabile e di ordine 2 (e, pertanto, convergente), con passo $h = 0.01$ *sull'intervallo [0,2.5], si ottiene il risultato*

[7] La BDF di ordine 7 è, infatti, 0-instabile.

[8] Esamineremo questa equazione più in dettaglio successivamente in Sezione 8.3.1.

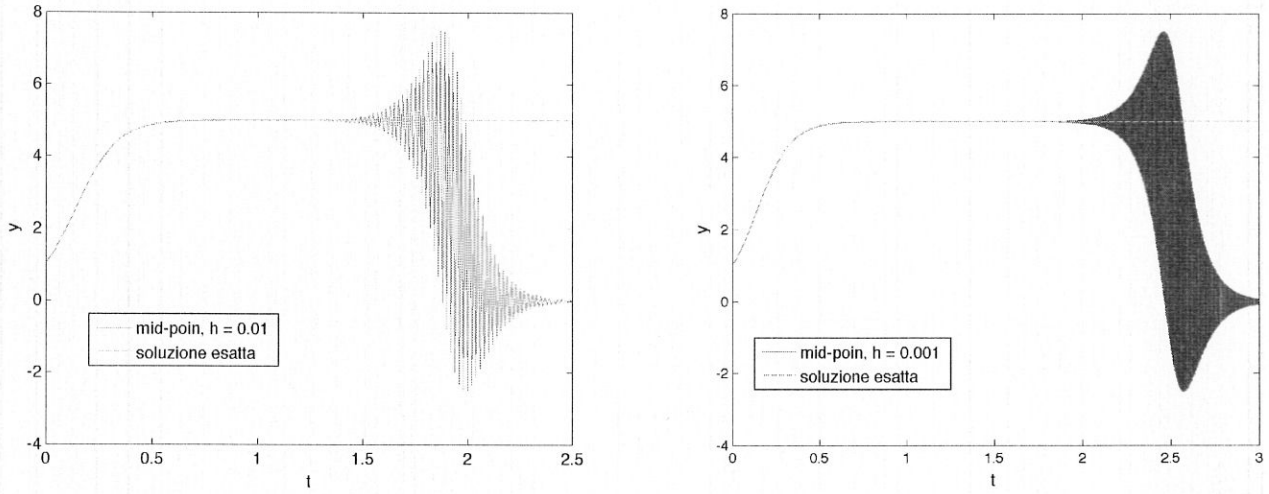

Figura 4.1: Problema (4.45) approssimato con il metodo del mid-point con passo $h = 0.01$ (a sinistra) e con passo $h = 0.001$ (a destra): sono evidenti, in entrambi i casi, le oscillazioni spurie nella approssimazione numerica.

illustrato a sinistra in Figura 4.1: come si vede, per $t \approx 1.5$ *compaiono delle oscillazioni spurie nella soluzione numerica, che si amplificano successivamente. Se riduciamo il passo di integrazione al valore* $h = 0.001$*, lo stesso fenomeno si ripresenta a partire da* $t \approx 2$*, come si evince dal grafico a destra in Figura 4.1. In generale, anche se riducessimo il passo* h *ulteriormente, il problema si ripresenterebbe considerando un intervallo di integrazione più ampio.*

Per quanto su argomentato, ci troviamo quindi a dover discutere l'*equazione dell'errore* (4.19) per h fissato. Questo è, in generale, un compito assai arduo, trattandosi di una equazione alle differenze di ordine k che, in generale, sarà *nonlineare*. Tuttavia, la trattazione diviene notevolmente più semplice nel caso in cui la (4.1) sia data dall'*equazione test*

$$y' = \lambda y, \qquad y(0) = y_0, \qquad \Re(\lambda) < 0, \tag{4.46}$$

in cui $\Re(\lambda)$ denota la parte reale di λ.

Osservazione 4.6 *Nonostante la sua semplicità, il problema (4.46) ha un preciso significato: quest'ultimo sarà compreso appieno dopo la trattazione degli argomenti del Capitolo 7.*

La soluzione di (4.46) è evidentemente

$$y(t) = y_0 \mathrm{e}^{\lambda t} \to 0, \qquad t \to \infty$$

che, pertanto, ha l'origine come punto di equilibrio asintoticamente stabile. Di consequenza, se anche la soluzione discreta, ottenuta risolvendo l'equazione

$$(\rho(E) - q\sigma(E))y_n = 0, \qquad q = h\lambda, \tag{4.47}$$

avesse l'origine come punto di equilibrio asintoticamente stabile, ovvero

$$y_n \to 0, \qquad n \to \infty, \tag{4.48}$$

allora necessariamente anche l'equazione dell'errore (4.19), che si riduce a

$$(\rho(E) - q\sigma(E))e_n = \tau_n,$$

sarebbe tale che

$$e_n \to 0, \qquad n \to \infty.$$

In tal caso, si dirà che il metodo (ρ, σ) è *assolutamente stabile* in $q = h\lambda$. Essendo l'equazione (4.47) lineare a coefficienti costanti, risulta dimostrato il seguente risultato.

Teorema 4.5 *La soluzione dell'equazione (4.47) soddisfa (4.48) se e solo se il polinomio*

$$\pi(z, q) = \rho(z) - q\sigma(z) \tag{4.49}$$

è un polinomio di Schur.

Definizione 4.7 *Il polinomio (4.49) è detto* polinomio di stabilità *del metodo. La regione*

$$\mathcal{D} = \{q \in \mathbb{C} : \pi(z, q) \text{ è un polinomio di Schur}\}, \tag{4.50}$$

è detta regione di assoluta stabilità *del metodo* (ρ, σ).

Richiedendo che (4.48) sia vera per ogni $h > 0$ e per ogni $\lambda \in \mathbb{C}^-$,[9] si capisce il senso della seguente, ulteriore, definizione.

Definizione 4.8 *Un metodo LMF è detto* A-stabile *se* $\mathbb{C}^- \subseteq \mathcal{D}$. *Nel caso in cui* $\mathbb{C}^- \equiv \mathcal{D}$, *il metodo si dirà* perfettamente *(o* precisamente*) A-stabile.*

Osservazione 4.7 *La denominazione* perfettamente A-stabile *deriva dal fatto che, in questo caso, la soluzione discreta ha sempre lo stesso comportamento qualitativo di quella continua. Infatti, la prima risulta avere l'origine asintoticamente stabile se e solo se questo avviene per la seconda, ovvero per* $\Re(\lambda) < 0$.

Chiaramente, la proprietà di A-stabilità è una proprietà che lascia molta libertà nella scelta del passo h e, pertanto, è una proprietà di rilievo per un metodo LMF, come argomenteremo ulteriormente nel seguito. Purtroppo, vale il seguente risultato negativo.

Teorema 4.6 (seconda barriera di Dahlquist) *Non esistono metodi LMF A-stabili espliciti. Inoltre, l'ordine massimo di un LMF A-stabile è 2.*

Osservazione 4.8 *È evidente che, nel caso un metodo non sia A-stabile, una regione di assoluta stabilità più ampia consente l'utilizzo di un passo h maggiore, fissato* $\lambda \in \mathbb{C}^-$, *rispetto ad un metodo con una regione di assoluta stabilità più piccola.*

Calcoliamo ora la regione di assoluta stabilità di qualcuno dei metodi su esposti.

[9] $\mathbb{C}^-$ denota il semipiano complesso a parte reale negativa.

I due metodi di Eulero

Per il metodo di Eulero esplicito (4.35), il polinomio di stabilità è:

$$\pi(z, q) = z - (1 + q).$$

Imponendo che si abbia

$$|z| = |1 + q| < 1,$$

si ottiene che la regione di assoluta stabilità è data dal cerchio (aperto) di centro -1 e raggio 1 del piano complesso, come si vede raffigurato a sinistra in Figura 4.3.

Per il metodo di Eulero implicito (4.43), il polinomio di stabilità è:

$$\pi(z, q) = (1 - q)z - 1.$$

Imponendo che si abbia

$$|z| = |1 - q|^{-1} < 1,$$

si ottiene che la regione di assoluta stabilità è data dal complemento in $\mathbb{C}$ del cerchio chiuso di centro 1 e raggio 1 del piano complesso, come si vede raffigurato a destra in Figura 4.3. Si deduce che il metodo di Eulero implicito è A-stabile.

Il metodo dei trapezi

Per il metodo dei trapezi (4.37), il polinomio di stabilità è:

$$\pi(z, q) = \left(1 - \frac{q}{2}\right) z - \left(1 + \frac{q}{2}\right).$$

Imponendo che si abbia

$$|z| = \frac{|2 + q|}{|2 - q|} < 1,$$

si ottiene $\Re(q) < 0$, ovvero

$$\mathcal{D} \equiv \mathbb{C}^{-}.$$

Pertanto, il metodo è perfettamente A-stabile.

I metodi di Simpson e del *mid-point*

Per il metodo del *mid-point* (4.41), il polinomio di stabilità è

$$\pi(z, q) = z^2 - 2qz - 1.$$

Applicando il primo criterio di Schur,[10] si verifica che la regione di assoluta stabilità non ha punti interni. Infatti, il polinomio di stabilità non potrà in nessun caso essere un polinomio di Schur, in quanto il coefficiente principale ed il termine noto hanno lo stesso modulo. Tuttavia, la regione di assoluta stabilità potrebbe avere una frontiera, laddove $\pi(z, q)$ fosse un polinomio di Von Neumann. Il polinomio ridotto si vede essere dato da:

$$\pi^{(1)}(z, q) = -2(q + \bar{q}) \equiv -4\Re(q).$$

[10] Vedere il Teorema 3.11.

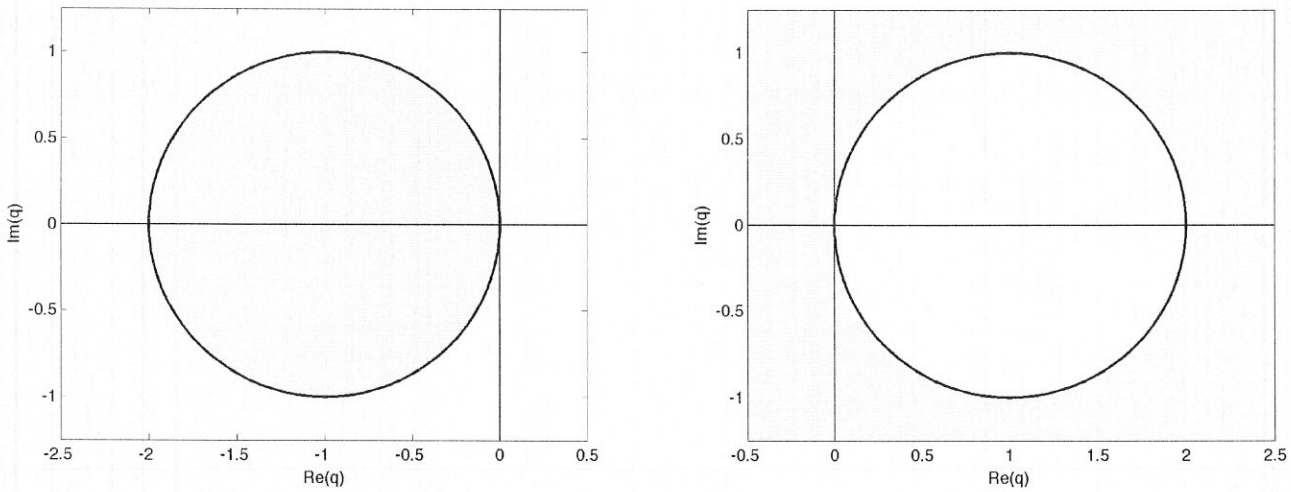

Figura 4.2: Regioni di assoluta stabilità dei metodi di Eulero esplicito (a sinistra) ed implicito (a destra), rispettivamente dati da (4.35) e (4.43).

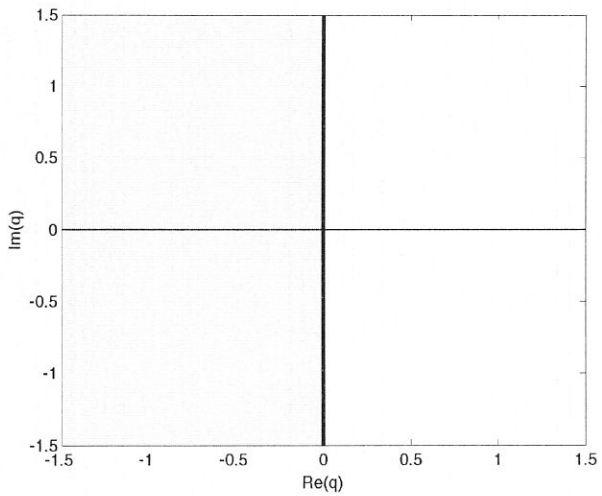

Figura 4.3: Regione di assoluta stabilità del metodo dei trapezi (4.37): si tratta di un metodo *perfettamente* A-stabile.

Imponendo, per il secondo criterio di Schur,[11] che esso sia identicamente nullo, si ottiene $\Re(q) = 0$. Pertanto avremo dei punti dell'asse immaginario. Richiedendo, sempre per il secondo criterio di Schur, anche che

$$\pi'(z, q) = 2(z - q)$$

sia un polinomio di Schur, si ottiene $|q| < 1$. Pertanto, si conclude che

$$\bar{\mathcal{D}} = \{ix : -1 < x < 1\}.$$

Il polinomio di stabilità del metodo di Simpson (4.40) è:

$$\pi(z, q) = \left(1 - \frac{q}{3}\right) z^2 - \frac{4}{3} qz - \left(1 + \frac{q}{3}\right).$$

Usando, anche per esso, argomenti simili a quelli utilizzati per il metodo del *mid-point*, si arriva a concludere che

$$\bar{\mathcal{D}} = \left\{ix : -\sqrt{3} < x < \sqrt{3}\right\}.$$

4.5 Il *boundary locus* di un metodo LMF

Per metodi (ρ, σ) più generali, potrebbe essere difficile determinare la regione di assoluta stabilità, utilizzando direttamente i criteri di Schur, come fatto in precedenza. Per semplificare questo compito, sarà conveniente ricorrere alla tecnica che utilizza il cosiddetto *boundary locus* del metodo. Quest'ultimo è definito come:

$$\Gamma = \left\{ q(\theta) = \frac{\rho(\mathrm{e}^{i\theta})}{\sigma(\mathrm{e}^{i\theta})} : 0 \leq \theta \leq 2\pi \right\}.$$

Infatti, essendo le radici di un polinomio funzioni analitiche dei suoi coefficienti, segue che affinché il numero di radici all'interno del cerchio aperto unitario vari, qualcuna di esse deve attraversare la circonferenza unitaria. Poiché Γ è il luogo dei punti del piano complesso q per cui il polinomio di stabilità ha almeno una radice di modulo 1, segue che Γ dividerà il piano complesso q in regioni connesse, all'interno delle quali il numero di radici nel cerchio unitario aperto è costante. Pertanto, basterà applicare il primo criterio di Schur ad un qualunque punto interno di ciascuna di tali regioni, per stabilire quale (o quali) di esse costituisca la regione di assoluta stabilità del metodo. A titolo di esempio, in Figura 4.4 riportiamo le regioni di assoluta stabilità calcolate con il *boundary locus*, per il metodo di Adams-Bashforth di ordine 2 e per il metodo di Adams-Moulton di ordine 3. Entrambi risultano avere una regione di assoluta stabilità limitata. In Figura 4.5 è raffigurata, invece, la regione di assoluta stabilità della BDF di ordine 2. Quest'ultimo metodo risulta essere A-stabile. Per meglio comprendere questo, andiamo a definire il concetto di *tipo di un polinomio.*

Definizione 4.9 *Dato un polinomio $p(z) \in \Pi_k$, se ne definisce il* tipo *come la terna di interi (k_1, k_2, k_3), con $k_1, k_2, k_3 \geq 0$ e $k_1 + k_2 + k_3 = k$, tale che:*

- *k_1 è il numero delle radici di $p(z)$ (contate con la loro molteplicità) di modulo* minore *di 1;*

[11] Vedere il Teorema 3.12.

- k_2 *è il numero delle radici di* $p(z)$ *(contate con la loro molteplicità) di modulo* uguale *a 1;*
- k_3 *è il numero delle radici di* $p(z)$ *(contate con la loro molteplicità) di modulo* maggiore *di 1.*

Pertanto:

- i polinomi di tipo $(k, 0, 0)$ sono i polinomi di Schur;
- i polinomi di tipo $(k_1, k_2, 0)$, con le k_2 radici di modulo 1 semplici, sono i polinomi di Von Neumann.

Da quanto su esposto, il *boundary locus* di un metodo LMF suddivide il piano complesso q in un numero finito di regioni connesse in cui il tipo del polinomio (che sarà del tipo $(k_1, 0, k - k_1)$ [12]) è costante. L'unione delle regioni in cui il polinomio ha tipo $(k, 0, 0)$ costituiscono la regione di assoluta stabilità del metodo. Se ne deduce che

$$\partial\mathcal{D} \subseteq \Gamma.$$

Andiamo ad esaminare alcune importanti proprietà del *boundary locus* di metodi consistenti e 0-stabili. Infatti, per un metodo consistente e 0-stabile, deve aversi

$$\rho(1) = 0, \qquad \rho'(1) = \sigma(1) \neq 0,$$

e, quindi,

$$q(0) = \frac{\rho(1)}{\sigma(1)} = 0 \in \Gamma.$$

Cioè, l'origine (del piano q) appartiene al *boundary locus* del metodo (nonché alla frontiera della sua regione di assoluta stabilità, perché $\rho(z)$ è un polinomio di Von Neumann). Inoltre, l'asse immaginario risulta essere tangente al *boundary locus.* Infatti,

$$q'(0) = \frac{\rho'(\mathrm{e}^{i\theta})\sigma(\mathrm{e}^{i\theta}) - \rho(\mathrm{e}^{i\theta})\sigma'(\mathrm{e}^{i\theta})}{\sigma(\mathrm{e}^{i\theta})^2} e^{i\theta} i \Big|_{\theta=0} = \frac{\sigma(1)^2}{\sigma(1)^2} i = i.$$

Pertanto,

$$q(\theta) \approx i\theta, \qquad \theta \approx 0.$$

In ultimo, osserviamo che:

$$\overline{q(\theta)} = \overline{\left(\frac{\rho(\mathrm{e}^{i\theta})}{\sigma(\mathrm{e}^{i\theta})}\right)} = \frac{\rho(\mathrm{e}^{-i\theta})}{\sigma(\mathrm{e}^{-i\theta})} = q(-\theta) = q(2\pi - \theta).$$

Ovvero, Γ è una curva simmetrica rispetto all'asse reale.

4.6 L-stabilità

Vogliamo ora mettere in evidenza una differenza esistente tra i metodi A-stabili finora esaminati. Ovvero, il metodo dei trapezi (4.37) ed i metodi di Eulero implicito (4.43) e BDF

[12] Infatti, le radici di modulo 1 si hanno solo quando q appartiene al *boundary locus* del metodo.

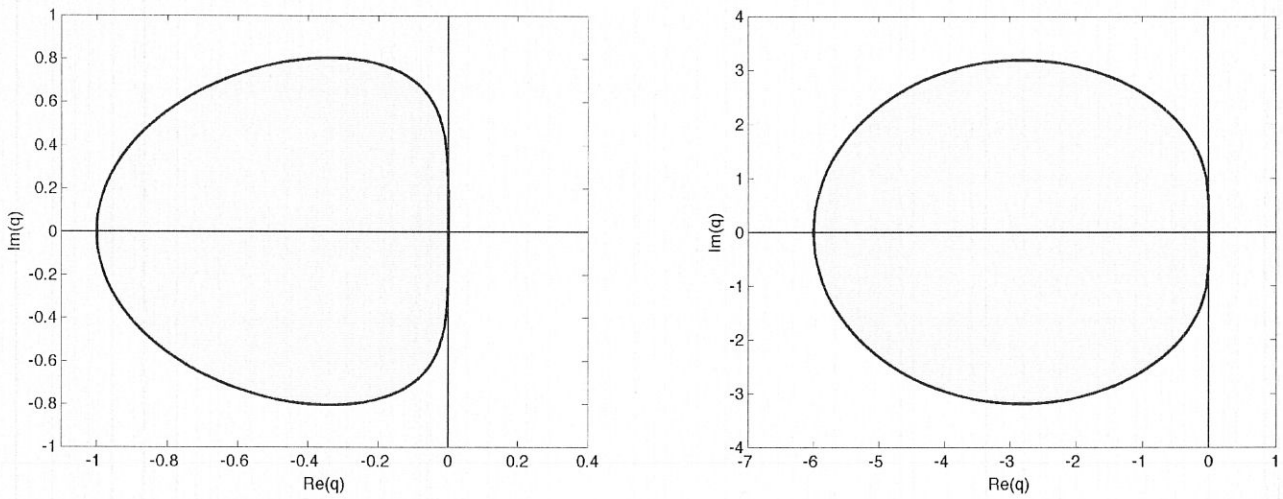

Figura 4.4: Regione di assoluta stabilità dei metodi di Adams-Bashforth (4.36) di ordine 2 (a sinistra) e di Adams-Moulton (4.38) di ordine 3 (a destra).

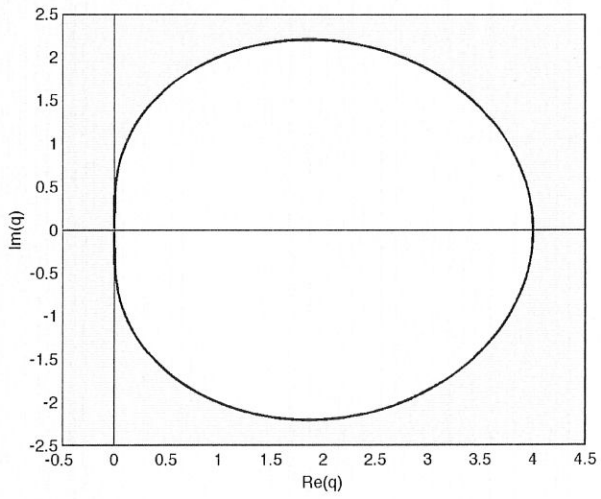

Figura 4.5: Regione di assoluta stabilità della BDF (4.44) di ordine 2.

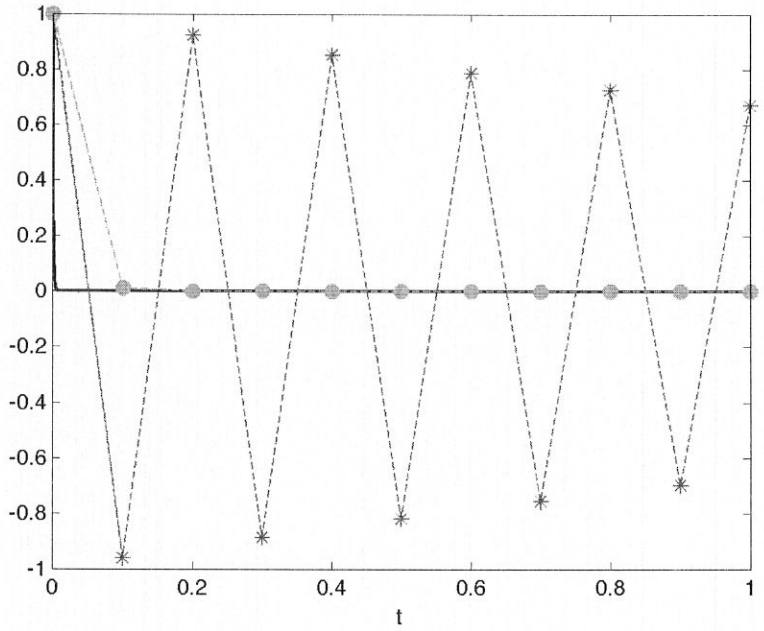

Figura 4.6: Vantaggio di un metodo L-stabile: soluzione esatta (linea continua) e sue approssimazioni ottenute con il metodo dei trapezi (curva tratteggiata e asterischi) ed il metodo di Eulero implicito (curva tratteggiata e cerchietti).

di ordine 2 (4.44). Per il metodo dei trapezi, si ha che, applicato all'equazione test (4.46), si ottiene:

$$y_n(q) = \left(\frac{2+q}{2-q}\right)^n y_0 \to (-1)^n y_0, \qquad q \to \infty.$$

In questo caso, la soluzione discreta tenderà ad oscillare, assumendo segno discorde, senza smorzarsi, quando $q = h\lambda$ tende a ∞. Viceversa, per il metodo di Eulero implicito, si ottiene:

$$y_n(q) = \left(\frac{1}{1-q}\right)^n y_0 \to 0, \qquad q \to \infty.$$

Anche per la BDF di ordine 2 si verifica che

$$y_n(q) \to 0, \qquad q \to \infty. \tag{4.51}$$

Chiaramente, la (4.51) descrive il comportamento qualitativamente corretto. Si capisce, pertanto, il senso della seguente proprietà.

Definizione 4.10 *Se un metodo A-stabile soddisfa la (4.51), allora esso si dirà* L-stabile.

Osservazione 4.9 *Per quanto su esposto, il metodo dei trapezi non è L-stabile, mentre lo sono le BDF di ordine 1 e 2.*

In Figura 4.6 è illustrato un esempio in cui si apprezza la proprietà di L-stabilità di un metodo. Infatti, viene raffigurata l'applicazione del metodo dei trapezi e del metodo di Eulero implicito, utilizzando un passo $h = 0.1$, per approssimare la soluzione del problema

$$y' = -10^3 y, \qquad t \in [0,1], \qquad y(0) = 1.$$

Come si può osservare, nonostante il passo h sia relativamente grande, la soluzione fornita dal metodo di Eulero implicito è qualitativamente corretta, mentre quella ottenuta con il metodo dei trapezi oscilla assumendo alternativamente valori positivi e negativi.