

Capitolo 3

Schemi numerici per la fluido dinamica: alcuni cenni

3.1 Discretizzazione: generalità

In generale, le equazioni alle derivate parziali dei differenti modelli di flusso non possono essere risolte per via analitica. Infatti, solo alcune soluzioni analitiche per flussi molto semplici sono note. Per risolvere queste equazioni nei casi più generali occorre utilizzare metodi di soluzione di tipo numerico. L'elemento chiave che sta alla base di ciascun metodo numerico è il processo di discretizzazione. Il problema originale, che consiste nel trovare una o più funzioni soluzione del problema differenziale, è trasformato attraverso il processo di discretizzazione in un problema algebrico la cui incognita è una sequenza discreta di valori numerici. La conversione del problema differenziale in un problema algebrico permette di utilizzare nel modo più efficiente i calcolatori elettronici, che sono in grado di eseguire le operazioni algebriche tra numeri a velocità molto elevata.

I metodi numerici utilizzati in fluidodinamica si possono suddividere in base ai tre ai tre processi di discretizzazione più comunemente utilizzati:

- i metodi alle differenze finite,
- i metodi ai volumi finiti,
- i metodi agli elementi finiti.

Nell'ambito del presente corso i processi di discretizzazione ai volumi finiti ed agli elementi finiti saranno descritti in modo sommario, mentre la discretizzazione alle differenze finite sarà analizzata in modo più esteso. In particolare, verrà mostrato come è possibile costruire alcuni semplici schemi numerici alle

differenze finite per l'equazione campione e quali sono le proprietà di questi schemi numerici.

3.2 Discretizzazione alle differenze finite

Prima di procedere alla discretizzazione alle differenze finite dell'equazione campione è utile richiamare alcuni concetti relativi al calcolo delle derivate di una funzione attraverso le formule alle differenze finite.

3.2.1 Discretizzazione di derivate

Si consideri una funzione $g(x)$ ed una successione di punti di ascissa x_i equispaziati con passo Δx . Si indichi con g_i il corrispondente valore della $g(x)$ nel punto di ascissa x_i (vedi figura 3.1).

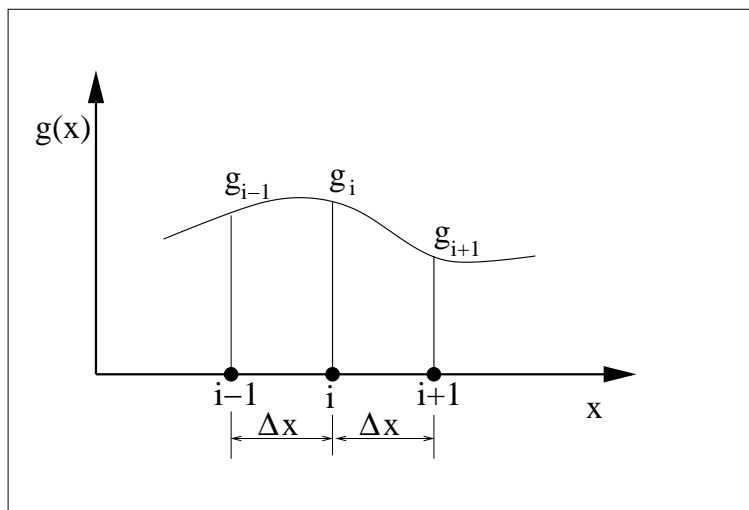


Figura 3.1:

Per costruire le formule alle differenze per il calcolo delle derivate della $g(x)$ si scriva lo sviluppo in serie di Taylor della funzione $g(x)$ nel punto x_{i+1} rispetto al punto x_i

$$g_{i+1} = g_i + \frac{dg}{dx} \Big|_i \Delta x + \frac{d^2g}{dx^2} \Big|_i \frac{\Delta x^2}{2} + \frac{d^3g}{dx^3} \Big|_i \frac{\Delta x^3}{6} + \frac{d^4g}{dx^4} \Big|_i \frac{\Delta x^4}{24} + \dots \quad (3.1)$$

e quello nel punto x_{i-1}

$$g_{i-1} = g_i - \frac{dg}{dx} \Big|_i \Delta x + \frac{d^2g}{dx^2} \Big|_i \frac{\Delta x^2}{2} - \frac{d^3g}{dx^3} \Big|_i \frac{\Delta x^3}{6} + \frac{d^4g}{dx^4} \Big|_i \frac{\Delta x^4}{24} + \dots \quad (3.2)$$

Se si esplicita la (3.1) rispetto alla derivata $\left. \frac{dg}{dx} \right|_i$ si ottiene la seguente relazione

$$\left. \frac{dg}{dx} \right|_i = \underbrace{\frac{g_{i+1} - g_i}{\Delta x}}_{\text{Differenza finita}} - \underbrace{\left. \frac{d^2g}{dx^2} \right|_i \frac{\Delta x}{2} - \left. \frac{d^3g}{dx^3} \right|_i \frac{\Delta x^2}{6} + \dots}_{\text{Errore } e_T|_i} \quad (3.3)$$

La (3.3) mostra che la derivata prima della funzione $g(x)$ nel punto x_i può essere posta uguale a una differenza finita, il primo termine a secondo membro, più infiniti termini che costituiscono l'errore di troncamento ($e_T|_i$). Poichè i termini dell'errore di troncamento dipendono da Δx , questo errore può essere reso piccolo a piacere scegliendo un Δx sufficientemente piccolo e la derivata di $g(x)$ può essere approssimata con il primo termine a secondo membro. In questo caso e_T rappresenta l'errore che si commette approssimando la derivata con la differenza finita. Il primo termine dell'errore di troncamento, quello caratterizzato dall'esponente di Δx più piccolo, prende il nome di termine dominante dell'errore di troncamento. Infatti se $\Delta x \rightarrow 0$, tutti i termini dell'errore di troncamento tendono a zero, tuttavia il termine dominante tende a zero meno rapidamente degli altri termini. Quindi per Δx sufficientemente piccolo il termine dominante prevale su tutti i rimanenti termini dell'errore di troncamento. Quando si verifica questa condizione, allora si è in convergenza asintotica. In condizione di convergenza asintotica l'errore di troncamento della formula alle differenze (3.3) è dell'ordine di Δx . Per questa ragione, la formula viene detta accurata al primo ordine ed l'intero errore viene indicato con una notazione che mette in evidenza questo aspetto

$$\left. \frac{dg}{dx} \right|_i = \frac{g_{i+1} - g_i}{\Delta x} + O(\Delta x) \quad (3.4)$$

Questa formula è una formula alle differenze in avanti, in quanto coinvolge i nodi i e $i + 1$ nel calcolo della derivata nel nodo i .

Analogamente dalla (3.2) si ottiene una formula alle differenze all'indietro che coinvolge i nodi i e $i - 1$.

$$\left. \frac{dg}{dx} \right|_i = \frac{g_i - g_{i-1}}{\Delta x} + \left. \frac{d^2g}{dx^2} \right|_i \frac{\Delta x}{2} - \left. \frac{d^3g}{dx^3} \right|_i \frac{\Delta x^2}{6} + \dots = \frac{g_i - g_{i-1}}{\Delta x} + O(\Delta x) \quad (3.5)$$

Anche questa formula è accurata al primo ordine ma ha un errore di troncamento leggermente differente da quella considerata in precedenza.

È possibile costruire formule alle differenze per la derivata prima di ordine di accuratezza superiore al primo ordine. Infatti se si sottrae la (3.1) alla (3.2) e si esplicita rispetto alla $\left. \frac{dg}{dx} \right|_i$ si ottiene una formula alle differenze per

la derivata prima accurata al secondo ordine di tipo centrato, in quanto, la formula alle differenze coinvolge i nodi $i + 1$ e $i - 1$ e il termine dominante dell'errore di troncamento è del secondo ordine

$$\left. \frac{dg}{dx} \right|_i = \frac{g_{i+1} - g_{i-1}}{2\Delta x} - \left. \frac{d^3g}{dx^3} \right|_i \frac{\Delta x^2}{6} + \dots = \frac{g_{i+1} - g_{i-1}}{2\Delta x} + O(\Delta x^2) \quad (3.6)$$

Sommando la (3.1) alla (3.2) ed esplicitando rispetto alla $\left. \frac{d^2g}{dx^2} \right|_i$ si ottiene una formula alle differenze per la derivata seconda

$$\left. \frac{d^2g}{dx^2} \right|_i = \frac{g_{i+1} - 2g_i + g_{i-1}}{\Delta x^2} - \left. \frac{d^4g}{dx^4} \right|_i \frac{\Delta x^2}{12} + \dots = \frac{g_{i+1} - 2g_i + g_{i-1}}{\Delta x^2} + O(\Delta x^2) \quad (3.7)$$

Anche in questo caso, la formula alle differenze è accurata al secondo ordine ed è di tipo centrato. Si possono ottenere formule alle differenze differenti da quelle già viste. A titolo di esempio si riportano, una formula alle differenze in avanti per la derivata prima accurata al secondo ordine

$$\left. \frac{dg}{dx} \right|_i = \frac{-3g_i + 4g_{i+1} - g_{i+2}}{2\Delta x} + O(\Delta x^2) \quad (3.8)$$

una formula alle differenze all'indietro per la derivata prima accurata al secondo ordine

$$\left. \frac{dg}{dx} \right|_i = \frac{3g_i - 4g_{i-1} + g_{i-2}}{2\Delta x} + O(\Delta x^2) \quad (3.9)$$

ed una formula centrata alle differenze per la derivata seconda accurata al quarto ordine

$$\left. \frac{d^2g}{dx^2} \right|_i = \frac{-g_{i+2} + 16g_{i+1} - 30g_i + 16g_{i-1} - g_{i-2}}{12\Delta x^2} + O(\Delta x^4) \quad (3.10)$$

Nel caso di una funzione di più variabili l'estensione delle formule alle differenze finite per le derivate prime è banale. Si consideri per semplicità una funzione di due variabili indipendenti $g(t, n)$ e si indichi con la notazione g_i^n il corrispondente valore nel punto di coordinate $(n\Delta t, i\Delta x)$ essendo Δt il passo di discretizzazione nel tempo e Δx quello nello spazio. Le formule alle differenze accurate al primo ordine per le derivate rispetto a t e x si ottengono da quelle viste in precedenza operando alcune piccole sostituzioni. Per esempio, la formula accurata al primo ordine in avanti per la derivata rispetto al tempo nel punto n, i si scrive come

$$\left. \frac{\partial g}{\partial t} \right|_i^n = \frac{g_i^{n+1} - g_i^n}{\Delta t} - \left. \frac{\partial^2 g}{\partial t^2} \right|_i^n \frac{\Delta t}{2} + \dots = \frac{g_i^{n+1} - g_i^n}{\Delta t} + O(\Delta t) \quad (3.11)$$

Anche le formule alle differenze finite relative alle derivate seconde, viste in precedenza, si estendono allo stesso modo al caso di funzione di più variabili indipendenti. Tuttavia per completare la casistica delle formule alle differenze per le derivate di ordine superiore al primo sarebbe necessario considerare anche le formule per le derivate miste (per esempio $\frac{\partial^2 g}{\partial t \partial x}$), che non saranno prese in considerazione in questo testo.

Prima di concludere è importante sottolineare alcune questioni.

- Il processo di discretizzazione alle differenze finite ha permesso di sostituire l'operazione di derivazione di una funzione con una sequenza di operazioni algebriche (moltiplicazioni, divisioni, addizioni e sottrazioni) che coinvolgono valori della funzione calcolati in punti vicini. Un calcolatore elettronico è in grado di eseguire queste operazioni algebriche molto rapidamente attraverso semplici programmi di calcolo.
- Più aumenta l'ordine di accuratezza della formula e più nodi contigui sono coinvolti nel calcolo. L'insieme dei nodi adiacenti che vengono utilizzati nel calcolo dalla formula alle differenze definisce la base (o il supporto) della formula.
- Più aumenta l'ordine di accuratezza della formula e più elevato è il costo computazionale. Per esempio si passa dalle 2 operazioni algebriche (una differenza ed una divisione) necessarie per il calcolo della formula accurata al primo ordine in avanti alle 6 operazioni (tre moltiplicazioni, una addizione, una sottrazione e una divisione) della formula in avanti accurata al secondo ordine.
- Utilizzare una formula alle differenze di ordine più accurato non significa che l'errore di troncamento sia per ogni valore del passo di discretizzazione sempre più piccolo di quello di una formula meno accurata.
- Il fatto di essere in convergenza asintotica non assicura che l'errore sia piccolo e trascurabile.

3.2.2 Discretizzazione di equazioni differenziali

Le formule alle differenze finite per le derivate viste nella sezione precedente permettono, non solo di calcolare in modo approssimato il valore di una derivata, ma anche di costruire dei metodi di soluzione approssimata delle equazioni differenziali.

Questi metodi vengono chiamati schemi di integrazione delle equazioni differenziali alle differenze finite. Per illustrare l'idea che sta alla base di questi schemi numerici e le loro proprietà, saranno costruiti tre schemi numerici alle

differenze finite per l'equazione campione e successivamente applicati ai tre problemi differenziali definiti nel capitolo precedente.

In questi problemi differenziali la funzione soluzione $g(t, x)$ deve essere determinata all'interno di uno stesso dominio spazio-temporale limitato dalle condizioni $0 < x < 1$ e $t > 0$ (vedi figura 3.2).

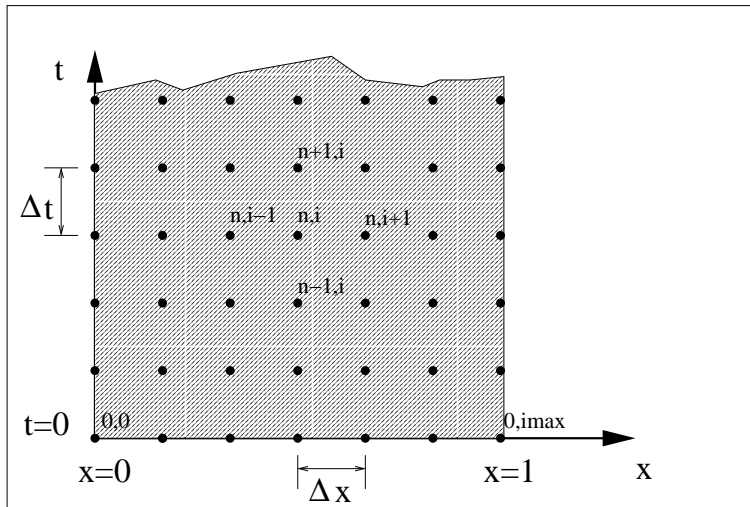


Figura 3.2:

Gli schemi di integrazione numerica alle differenze finite non determinano la funzione in tutti i punti del dominio, ma calcolano il valore approssimato di $g(t, x)$ su un reticolo di nodi costruito all'interno del dominio. Inizialmente si assumerà che il passo di discretizzazione nello spazio (Δx) e nel tempo (Δt) sia costante (vedi figura 3.2).

Schema FTCS (Forward time, centered space)

Lo schema BTCS viene costruito sostituendo nell'equazione campione, le derivate nello spazio con le formule centrate e la derivata nel tempo con una formula alle differenze in avanti. L'equazione campione deve essere soddisfatta in tutti i punti interni del dominio di calcolo e, quindi, in tutti i nodi interni del reticolo di calcolo. Se si considera un generico nodo interno (n, i) di coordinate $(n\Delta t, i\Delta x)$, sostituendo le derivate dell'equazione campione con la formula alle differenze (3.11) e le formule (3.6) e (3.7), modificate per il caso di una funzione a due variabili indipendenti si ottiene

$$\frac{\partial g}{\partial t} + \lambda \frac{\partial g}{\partial x} - \mu \frac{\partial^2 g}{\partial x^2} = \frac{g_i^{n+1} - g_i^n}{\Delta t} + \lambda \frac{g_{i+1}^n - g_{i-1}^n}{2\Delta x} - \mu \frac{g_{i+1}^n - 2g_i^n + g_{i-1}^n}{\Delta x^2} + eT_i^n = 0 \quad (3.12)$$

dove

$$e_{T_i}^n = - \frac{\partial^2 g}{\partial t^2} \Big|_i^n \frac{\Delta t}{2} - \lambda \frac{\partial^3 g}{\partial t^3} \Big|_i^n \frac{\Delta x^2}{6} + \mu \frac{\partial^4 g}{\partial t^4} \Big|_i^n \frac{\Delta x^2}{12} + \dots$$

comprende tutti gli errori di troncamento di ciascuna formula alle differenze. Questo termine per motivi che saranno chiari tra breve prende il nome di errore di troncamento dello schema. Poichè è facile verificare che $e_T \rightarrow 0$ se $\Delta x, \Delta t \rightarrow 0$, allora se Δx e Δt sono sufficientemente piccoli, e_T sarà trascurabile rispetto ai rimanenti termini a secondo membro della (3.12). Se si trascura e_T la (3.12) si trasforma nella seguente equazione alle differenze

$$\frac{\bar{g}_i^{n+1} - \bar{g}_i^n}{\Delta t} + \lambda \frac{\bar{g}_{i+1}^n - \bar{g}_{i-1}^n}{2\Delta x} - \mu \frac{\bar{g}_{i+1}^n - 2\bar{g}_i^n + \bar{g}_{i-1}^n}{\Delta x^2} = 0 \quad (3.13)$$

I valori g_i^n della (3.12) non sono identici ai valori \bar{g}_i^n della (3.13), poichè nel passaggio dalla (3.12) alla (3.13) è stata operata una approssimazione. Infatti, mentre g_i^n sono valori della $g(t, x)$ soluzione della

$$\frac{\partial g}{\partial t} + \lambda \frac{\partial g}{\partial x} - \mu \frac{\partial^2 g}{\partial x^2} = 0 \quad (3.14)$$

i \bar{g}_i^n sono valori della $\bar{g}(t, x)$ soluzione dell'equazione

$$\frac{\partial \bar{g}}{\partial t} + \lambda \frac{\partial \bar{g}}{\partial x} - \mu \frac{\partial^2 \bar{g}}{\partial x^2} - e_T = 0 \quad (3.15)$$

Appare chiaro che i valori g_i^n potranno coincidere con valori \bar{g}_i^n solo nel caso in cui $e_T = 0$. L'equazione (3.15) prende il nome di equazione modificata dello schema numerico. Questa equazione differisce dall'equazione del problema originale per la presenza dell'errore di troncamento e_T , che, in generale, è differente da zero per $\Delta t, \Delta x \neq 0$.

Lasciando al primo membro della (3.13) il termine al tempo $n+1$ e portando a secondo membro i termini al tempo n si ottiene la seguente espressione

$$\bar{g}_i^{n+1} = (s - \frac{c}{2})g_{i+1}^n + (1 + 2s)g_i^n + (s + \frac{c}{2})g_{i-1}^n \quad (3.16)$$

dove $c = \frac{\Delta t \lambda}{\Delta x}$ e $s = \frac{\Delta t \mu}{\Delta x^2}$ sono due parametri legati alla discretizzazione nello spazio e nel tempo. La (3.16) permette di calcolare il valore \bar{g}_i^{n+1} noti i valori $g_{i+1}^n, g_i^n, g_{i-1}^n$. in generale, questa equazione alle differenze insieme alle condizioni al contorno, che caso dei problemi considerati si scrivono nel seguente modo

$$\begin{aligned} g_1^{n+1} &= g_1^n \\ g_{max}^{n+1} &= g_{max}^n \end{aligned}$$

permette di calcolare la soluzione approssimata al tempo $n + 1$ nota quella al tempo n . Poichè la soluzione per $n = 0$ ($t = 0$) è nota dalle condizioni iniziali, la soluzione può essere calcolata ad ogni istante successivo.

Questo schema di calcolo può essere riscritto nella seguente forma matriciale

$$\bar{g}^{n+1} = B\bar{g}^n \quad (3.17)$$

dove \bar{g} è il vettore contenente lo soluzione ad un certo istante e B è una matrice tridiagonale nella quale appaiono i parametri c ed s

$$\bar{g} = \begin{Bmatrix} \bar{g}_1 \\ \bar{g}_2 \\ \bar{g}_3 \\ \dots \\ \bar{g}_{max} \end{Bmatrix} \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & \dots & 0 \\ (s - \frac{c}{2}) & 1 + 2s & (s + \frac{c}{2}) & 0 & \dots & \dots & 0 \\ 0 & (s - \frac{c}{2}) & 1 + 2s & (s + \frac{c}{2}) & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Questo schema è uno schema esplicito poichè nell'equazione alle differenze (3.12) compare un solo termine $n + 1$ mentre tutti gli altri termini sono al tempo n . Poichè termini dominanti di e_T sono termini dell'ordine di Δt e di Δx^2 lo schema è accurato al primo ordine nel tempo e al secondo ordine nello spazio.

Schema BTCS (Backward time, centered space)

Il secondo schema preso in considerazione è apparentemente molto simile allo schema FTCS. Infatti, l'equazione alle differenze dello schema BTCS è la seguente

$$\frac{\bar{g}_i^{n+1} - \bar{g}_i^n}{\Delta t} + \lambda \frac{\bar{g}_{i+1}^{n+1} - \bar{g}_{i-1}^{n+1}}{2\Delta x} - \mu \frac{\bar{g}_{i+1}^{n+1} - 2\bar{g}_i^{n+1} - \bar{g}_{i-1}^{n+1}}{\Delta x^2} = 0 \quad (3.18)$$

In questo caso lo schema è stato costruito sostituendo le derivate spaziali con le formule centrate valutate al tempo $n + 1$ e la derivata nel tempo con la differenze finita all'indietro rispetto al tempo $n+1$. Come lo schema precedente, anche questo schema è accurato al primo ordine nel tempo e al secondo ordine nello spazio. Tuttavia, al contrario dello schema FTCS, lo schema è implicito. Raggruppando a primo membro i termini \bar{g}^{n+1} e a secondo membro i termini \bar{g}^n si ottiene la seguente equazione alle differenze

$$-(s - \frac{c}{2})\bar{g}_{i+1}^{n+1} + (1 + 2s)\bar{g}_i^{n+1} - (s + \frac{c}{2})\bar{g}_{i-1}^{n+1} = \bar{g}_i^n \quad (3.19)$$

E' evidente che l'equazione alle differenze ha più di un termine al tempo $n + 1$ e che non é possibile calcolare il valore \bar{g}_i^{n+1} senza conoscere anche il valore

di \bar{g}_{i+1}^{n+1} e \bar{g}_{i-1}^{n+1} . Ciò significa che i valori \bar{g}^{n+1} vanno calcolati tutti insieme. Passando alla notazione matriciale, lo schema BTCS si può rappresentare nel seguente modo

$$A\bar{g}^{n+1} = \bar{g}^n \quad (3.20)$$

dove A è la seguente matrice tridiagonale

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & \dots & 0 \\ -(s - \frac{c}{2}) & 1 + 2s & -(s + \frac{c}{2}) & 0 & \dots & \dots & 0 \\ 0 & -(s - \frac{c}{2}) & 1 + 2s & -(s + \frac{c}{2}) & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

nella quale compaiono i parametri c e s . Questo schema numerico è chiaramente più oneroso del precedente in quanto il calcolo del vettore \bar{g}^{n+1} richiede l'inversione della matrice A

$$\bar{g}^{n+1} = A^{-1}\bar{g}^n \quad (3.21)$$

Schema CN (Crank-Nicolson)

L'ultimo schema considerato è lo schema di Crank-Nicolson. Questo schema è costruito sulla seguente equazioni alle differenze

$$\frac{\bar{g}_i^{n+1} - \bar{g}_i^n}{\Delta t} + \lambda \frac{\frac{1}{2}(\bar{g}_{i+1}^{n+1} + \bar{g}_{i+1}^n) - \frac{1}{2}(\bar{g}_{i-1}^{n+1} + \bar{g}_{i-1}^n)}{2\Delta x} - \mu \frac{\frac{1}{2}(\bar{g}_{i+1}^{n+1} + \bar{g}_{i+1}^n) - \frac{2}{2}(\bar{g}_i^{n+1} + \bar{g}_i^n) + \frac{1}{2}(\bar{g}_{i-1}^{n+1} + \bar{g}_{i-1}^n)}{\Delta x^2} = 0 \quad (3.22)$$

Lo schema è costruito utilizzando le formule alle differenze di tipo centrate per le derivate spaziali, ma diversamente dagli schemi precedenti queste derivate sono valutate al tempo $(n + \frac{1}{2})\Delta t$. Rispetto a tale istante temporale la formula alle differenze utilizzata per la derivata nel tempo diventa una formula centrata e, quindi, è accurata al secondo ordine. Questo schema è implicito, come lo schema BTCS ma rispetto a quest'ultimo schema risulta essere accurato al secondo ordine non solo nello spazio ma anche nel tempo. L'espressione matriciale dello schema risulta essere la seguente

$$A\bar{g}^{n+1} = B\bar{g}^n \quad (3.23)$$

Dove A e B sono matrici tridiagonali aventi la seguente espressione

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & \dots & 0 \\ -(s - \frac{c}{2}) & 2 + 2s & -(s + \frac{c}{2}) & 0 & \dots & \dots & 0 \\ 0 & -(s - \frac{c}{2}) & 2 + 2s & -(s + \frac{c}{2}) & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & \dots & 0 \\ (s - \frac{c}{2}) & 2 + 2s & (s + \frac{c}{2}) & 0 & \dots & \dots & 0 \\ 0 & (s - \frac{c}{2}) & 2 + 2s & (s + \frac{c}{2}) & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Anche in questo caso il calcolo della soluzione \bar{g}^{n+1} comporta l'inversione di una matrice

$$\bar{g}^{n+1} = A^{-1}B\bar{g}^n \quad (3.24)$$

3.3 Proprietà degli schemi numerici e teorema di Lax

Dopo aver costruito questi tre schemi di integrazione occorre verificare se questi schemi sono in grado di fornire una stima della soluzione esatta caratterizzata da un errore che può essere reso piccolo a piacere riducendo il passo di discretizzazione. Questo problema è meno banale di quanto si possa pensare. Nel calcolo delle derivate con le formule alle differenze, l'errore di troncamento è la differenza che esiste tra il valore esatto della derivata e quello calcolato numericamente. In tale caso se $e_T \rightarrow 0$ allora anche il valore numerico della derivata tenderà necessariamente a quello esatto. Nel caso degli schemi numerici per le equazioni differenziali e_T non coincide con l'errore tra la soluzione numerica e quella esatta.

Nel paragrafo precedente è stato visto che g_i^n sono i valori della funzione $g(t, x)$ soluzione della equazione campione, mentre \bar{g}_i^n sono i valori della funzione $\bar{g}(t, x)$ è soluzione dell'equazione modificata che differisce dall'equazione campione per e_T i cui termini dipendono da Δx e Δt . Tra g_i^n e \bar{g}_i^n si può definire un errore che anche se è legato ad e_T non coincide con e_T . Si definisce l'errore numerico ε_i^n della soluzione numerico nei punti x_i al tempo $n\Delta t$ come

$$\varepsilon_i^n = g_i^n - \bar{g}_i^n \quad (3.25)$$

Se per $\Delta x, \Delta t \rightarrow 0$ si ha $|\varepsilon_i^n| \rightarrow 0$ allora la soluzione dello schema converge alla soluzione esatta e lo schema è detto convergente.

Nei tre schemi analizzati è facile verificare che e_T tende a zero per Δt , $\Delta x \rightarrow 0$. Quando questa proprietà è verificata allora si dice che lo schema è consistente. Per verificare che lo schema sia convergente non è sufficiente che lo schema sia consistente, ma occorre che lo schema goda di un'altra proprietà. Per rendersi conto di ciò occorre analizzare come l'errore numerico è collegato all'errore di troncamento e come questo errore evolve passo dopo passo nel processo di integrazione.

In generale, l'evoluzione nel tempo della soluzione numerica \bar{g} degli schemi numerici analizzati è data da una relazione del tipo

$$\bar{g}^{n+1} = L\bar{g}^n \quad (3.26)$$

dove L è una opportuna matrice che dipende dallo schema numerico e nella quale compaiono c e s . E' facile dimostrare che la soluzione esatta evolve con la seguente legge

$$g^{n+1} = Lg^n + Ce_T^n \Delta t \quad (3.27)$$

dove C è una opportuna matrice dipendente dallo schema, e_T è il vettore dell'errore di troncamento. Sottraendo la (3.27) alla (3.26) ed utilizzando la definizione (3.25) si ottiene una relazione mostra come evolve l'errore numerico passo dopo passo

$$\varepsilon^{n+1} = L\varepsilon^n + Ce_T^n \Delta t \quad (3.28)$$

Questa relazione mostra che l'errore numerico al passo $n+1$ dipende dall'errore di troncamento al passo n , ma anche da come viene modificato dallo schema l'errore prodotto nei passi precedenti. Appare chiaro che perchè si abbia che $\varepsilon^n \rightarrow 0$ per Δt , $\Delta x \rightarrow 0$ occorre che lo schema sia convergente e che lo schema non amplifichi gli errori fatti nei passi precedenti. Perchè si verifichi quest'ultima condizione è necessario che L sia tale che

$$\forall a \quad |La| \leq |a|$$

Questa proprietà prende il nome di stabilità di uno schema. Pertanto condizione necessaria e sufficiente perché uno schema sia convergente (e quindi sia possibile ridurre l'errore numerico riducendo il passo di discretizzazione) è che lo schema sia consistente e stabile. Questo enunciato prende il nome di teorema di Lax.