

Introduzione ai metodi numerici alle differenze finite per equazioni di evoluzione¹

Roberto Natalini

Istituto per le Applicazioni del Calcolo “Mauro Picone”
Consiglio Nazionale delle Ricerche

¹Corso di Dottorato “**Metodi Numerici per Problemi di Evoluzione**”,
Dottorato di ricerca in Modelli e Metodi Matematici per la tecnologia e la società,
Università di Roma “La Sapienza”, a.a. 2003/04

Premessa

Queste note costituiscono un tentativo, per la verità abbastanza incompleto, di fornire una prima introduzione ai problemi relativi all'approssimazione numerica alle differenze finite dei problemi differenziali di tipo iperbolico e parabolico.

Ho diviso per semplicità il materiale in tre parti. La prima è dedicata ai problemi di trasporto iperbolico lineare. Tutto il materiale è trattato nel caso a coefficienti costanti e senza termini d'ordine zero o di sorgente per meglio evidenziare le idee essenziali. Analogamente nella seconda parte ci si concentrerà piuttosto sull'equazione del calore, pur accennando ai problemi posti dall'interazione con i termini convettivi.

In entrambi i casi ho dovuto rinunciare per mancanza di spazio alla trattazione dei problemi al bordo, tranne che per l'equazione del calore, e multidimensionali che pure sono di estremo interesse applicativo.

Infine nella terza parte, che copre oltre la metà del materiale trattato, ho cercato di rendere in qualche modo lo spirito con cui vengono oggi affrontati alcuni difficili problemi non lineari.

Ecco infine una lista di referenze bibliografiche per chi fosse interessato ad ulteriori approfondimenti.

Per dei prerequisiti sulle equazioni a derivate parziali, oltre al classico libro di F. John [Jo], un buon testo recente sulla parte lineare è quello di DiBenedetto [Di2].

Per le equazioni iperboliche e paraboliche lineari e non lineari un'ottima introduzione è data dal Kreiss e Lorentz [KL], un testo un po' superato e con qualche imprecisione, ma ricco di idee suggestive è quello di Smoller [Sm]. Un trattato vasto e profondo, ma per questo abbastanza difficile, è quello in tre volumi di Taylor [Ty]. Ci sono poi alcuni testi recenti più specificatamente centrati sulla teoria iperbolica come quelli di Dafermos [Da] e di Serre [Se].

Infine alcuni testi di base di analisi numerica delle equazioni differenziali sono lo Strikwerda [St] e il Richtmyer e Morton [RM], senza trascurare il piccolo volume di Morton e Mayers [MM] da dove ho pescato parecchie figure ed esempi.

A livello più avanzato segnalo senz'altro per i problemi iperbolici non lineari i due volumi di Godlewski e Raviart [GR1, GR2] e il libro recente di Leveque [Le] e, sul versante parabolico, il bel libro di sintesi di Morton [Mo]. Un'introduzione di livello più avanzato ai metodi di rilassamento e ai volumi finiti è contenuta in [Bo].

Indice

1	Equazioni iperboliche	5
1.1	Generalità sulle equazioni iperboliche	5
1.2	Introduzione ai metodi alle differenze finite	8
1.3	Convergenza e consistenza	11
1.4	Stabilità	16
1.5	La condizione di CFL	19
1.6	Stime nella norma del massimo e monotonia	20
1.7	L'analisi di von Neumann	21
1.7.1	Cenni sull'analisi di Fourier	21
1.7.2	L'analisi di von Neumann	23
1.8	Ordine di accuratezza e schema di Lax–Wendroff	26
1.9	Cenni sull'equazione delle onde	29
2	Equazioni paraboliche	35
2.1	Il problema di Cauchy–Dirichlet per l'equazione del calore	35
2.2	Uno schema esplicito per l'equazione del calore	36
2.3	Un metodo implicito	38
2.4	I θ -metodi e lo schema di Crank–Nicolson	40
2.5	Principio di massimo e convergenza	42
2.6	L'equazione di convezione-diffusione	44
3	Leggi di conservazione non lineari	49
3.1	Alcuni esempi di problemi non lineari di evoluzione	51
3.2	Richiami di teoria delle soluzioni deboli	53
3.2.1	Esistenza locale di soluzioni classiche nel caso iperbolico	53
3.2.2	Soluzioni deboli e condizioni di entropia	54
3.2.3	Il problema di Riemann	59
3.2.4	Esistenza e unicità delle soluzioni entropiche	60
3.2.5	Problemi parabolici non lineari	61
3.3	Schemi espliciti per le leggi di conservazione	63
3.3.1	Definizioni principali	64
3.3.2	Esempi di schemi a tre punti	67
3.4	Analisi della convergenza	74
3.4.1	Schemi TVD e monotoni	74

3.4.2	Forma incrementale e viscosità numerica	76
3.4.3	La condizione di entropia	79
3.5	Schemi TVD del secondo ordine	80
3.6	Il caso parabolico	85

Capitolo 1

Equazioni iperboliche

Cominceremo il nostro corso considerando una delle classi più importanti di equazioni differenziali, le equazioni di tipo iperbolico. Noi considereremo inizialmente il puro moto di tipo convettivo lineare per poi includere i sistemi iperboliche e le equazioni a coefficienti variabili. Dopo aver introdotto i concetti fondamentali legati alla definizione degli schemi numerici, viene presentata un'analisi dei concetti di convergenza, consistenza e stabilità, in relazione al risultato più importante per lo studio dell'approssimazione di problemi lineari a coefficienti costanti, il Teorema di equivalenza di Lax-Richtmyer. Il capitolo contiene infine una discussione della condizione di Courant-Friedrichs-Lewy (CFL) e altri argomenti legati alla stabilità delle approssimazioni.

1.1 Generalità sulle equazioni iperboliche

L'esempio base di equazione differenziale di tipo iperbolico è dato dall'equazione del trasporto lineare:

$$u_t + au_x = 0 \tag{1.1}$$

dove a è una costante, t rappresenta un tempo e x è la variabile spaziale. Data la funzione $u = u(x, t)$ ad un istante iniziale $t = 0$, ossia imponendo la condizione

$$u(x, 0) = u_0(x), \tag{1.2}$$

per una certa funzione data u_0 , vogliamo determinare il valore di u per tempi successivi, ossia per $t > 0$. Questo è chiamato un problema ai valori iniziali o di Cauchy. Come è noto, l'unica soluzione di questo problema è data da

$$u(x, t) = u_0(x - at) . \tag{1.3}$$

Dalla formula (1.3) possiamo dedurre alcuni fatti importanti. Innanzitutto la soluzione non è altro che una traslazione (verso destra se $a > 0$ o verso sinistra se $a < 0$) del dato iniziale a velocità a , chiamata quindi velocità dell'onda. Inoltre lungo le rette $\xi = x - at$ la soluzione è costante e vale $u_0(\xi)$. Queste rette

sono chiamate caratteristiche dell'equazione. Possiamo ritrovare la soluzione direttamente utilizzando il concetto di caratteristica. Si faccia il cambiamento di variabili da (x, t) a (ξ, τ) definito da

$$\xi = x - at, \quad \tau = t.$$

La sua inversa è data da

$$x = \xi + a\tau, \quad t = \tau.$$

Possiamo allora definire una nuova funzione $\tilde{u}(\xi, \tau) = u(x, t)$ tramite questa trasformazione. Nelle nuove coordinate l'equazione (1.1) diventa

$$\tilde{u}_\tau = \frac{\partial t}{\partial \tau} u_t + \frac{\partial x}{\partial \tau} u_x = 0.$$

Questa è una semplice equazione differenziale in τ che ha come unica soluzione

$$\tilde{u}(\xi, \tau) = u_0(\xi)$$

che dà esattamente la formula (1.3). Analogamente è possibile dimostrare che la soluzione dell'equazione

$$u_t + au_x + bu = f(x, t) \tag{1.4}$$

con la condizione iniziale (1.2) è data dalla funzione

$$u(x, t) = u_0(x - at)e^{-bt} + \int_0^t e^{-b(t-s)} f(x - a(t-s), s) ds. \tag{1.5}$$

Anche qui si vede che la soluzione è influenzata solo dai valori dei dati u_0 e f lungo la retta caratteristica $\xi = x - a(t - s)$.

Si osservi che nonostante l'equazione richieda la differenziabilità della soluzione u , le formule (1.3) e (1.5) hanno senso per dati iniziali molto generali. Questo permette di trattare soluzioni generalizzate eventualmente discontinue. Questi metodi si possono generalizzare facilmente, benchè in quel caso non ci sia più necessariamente una soluzione esplicita, alle equazioni semilineari

$$u_t + au_x = f(x, t, u). \tag{1.6}$$

Passiamo ora ad esaminare il caso dei sistemi di equazioni iperboliche in una variabile spaziale. L'incognita U è adesso un vettore di dimensioni k .

Definizione 1.1 *Un sistema della forma*

$$U_t + AU_x + BU = F(x, t), \tag{1.7}$$

dove A, B sono matrici $k \times k$ e F è una funzione a valori in \mathbb{R}^k , è detto (fortemente) iperbolico se la matrice A è diagonalizzabile con autovalori reali.

Poichè la matrice A è diagonalizzabile, esiste una matrice non singolare P tale che

$$PAP^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \dots & \\ 0 & & \lambda_k \end{pmatrix} = \Lambda,$$

dove Λ è una matrice diagonale i cui elementi $\lambda_1, \dots, \lambda_k$ sono gli autovalori di A e vengono chiamati le *velocità caratteristiche* del sistema. A questo punto basta effettuare il cambiamento di variabili $W = PU$ per ottenere

$$W_t + \Lambda W_x + PBB^{-1}W = PF.$$

Nel caso $B = 0$, il sistema in W è disaccoppiato e formato da k equazioni scalari che si risolvono come prima con il metodo delle caratteristiche. Il caso generale presenta qualche difficoltà supplementare, ma non altera la natura generale delle soluzioni che sempre si propagano lungo le caratteristiche.

Esempio 1.1 Si consideri per esempio il sistema 2×2

$$\begin{aligned} u_t + u_x + 2v_x &= 0, \\ v_t + 2u_x - 2v_x &= 0, \end{aligned}$$

con dati iniziali

$$u(x, 0) = u_0(x), \quad v(x, 0) = v_0(x),$$

per delle funzioni date u_0 e v_0 , che può essere scritto nella forma (1.7) per $U = (u, v)$ e

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}.$$

Si vede facilmente che gli autovalori di A sono dati da $\lambda_1 = -3$ e $\lambda_2 = 2$ e la matrice P è data da

$$P = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}.$$

Posto quindi $W = (w, z) = PU = (u - 2v, 2u + v)$, si ottiene il sistema

$$\begin{aligned} w_t - 3w_x &= 0, \\ z_t + 2z_x &= 0. \end{aligned}$$

Questo sistema ha soluzione

$$w(x, t) = w_0(x+3t) = (u_0 - 2v_0)(x+3t), \quad z(x, t) = z_0(x-2t) = (2u_0 + v_0)(x-2t).$$

Utilizzando la trasformazione inversa $U = P^{-1}W$ otteniamo il risultato finale

$$u(x, t) = \frac{1}{5}[(u_0 - 2v_0)(x + 3t) - 2(2u_0 + v_0)(x - 2t)]$$

$$v(x, t) = \frac{1}{5}[2(u_0 - 2v_0)(x + 3t) + (2u_0 + v_0)(x - 2t)].$$

Si osservi come la soluzione sia naturalmente decomposta nella somma di un'onda che viaggia alla velocità -3 e una che viaggia alla velocità 2.

Un'ulteriore estensione della teoria è data dalle equazioni e sistemi a coefficienti variabili, eventualmente con termine di sorgente dipendente dall'incognita. Si consideri per esempio l'equazione scalare

$$u_t + a(x, t)u_x = 0 \quad (1.8)$$

dove $a(x, t)$ è una funzione regolare. Se cerchiamo un cambio di variabili che come prima ci riporti ad un'equazione differenziale ordinaria, o visto in maniera geometrica, che "raddrizzi" le caratteristiche, facciamo la trasformazione (x, t) a (ξ, τ) definito da

$$\tau = t,$$

lasciando per ora ξ indeterminata. Possiamo definire una nuova funzione $\tilde{u}(\xi, \tau) = u(x, t)$ tramite questa trasformazione. Nelle nuove coordinate l'equazione (1.8) diventa

$$\tilde{u}_\tau = \frac{\partial t}{\partial \tau} u_t + \frac{\partial x}{\partial \tau} u_x = 0,$$

se imponiamo che

$$\frac{dx}{d\tau} = a(x, \tau).$$

In pratica si procede nel modo seguente. Si risolve il sistema di equazioni differenziali ordinarie

$$\begin{aligned} \frac{d\tilde{u}}{d\tau} &= 0, & \tilde{u}(\xi, 0) &= u_0(\xi), \\ \frac{dx}{d\tau} &= a(x, \tau), & x(0) &= \xi. \end{aligned} \quad (1.9)$$

Essendo $x(0) = \xi$, almeno per piccoli valori di τ , la relazione $x = x(\xi, \tau)$ è invertibile per τ fissato tramite una funzione $\xi = \xi(x, \tau)$. La soluzione del problema (1.8) sarà allora data dalla funzione

$$u(x, t) = \tilde{u}(\xi, \tau) = u_0(\xi(x, t), t).$$

Questa funzione è costante lungo le curve $\xi = \text{cost.}$. Per i sistemi si procederà in maniera analoga richiedendo che la diagonalizzabilità sia uniforme, in un'opportuna topologia, rispetto a (x, t) .

Osserviamo infine che le idee qui esposte non si applicano direttamente al caso in cui le velocità possano dipendere dall'incognita u , poichè, come vedremo in seguito, le caratteristiche possono intersecarsi e generare soluzioni discontinue, che necessitano di un trattamento particolare.

1.2 Introduzione ai metodi alle differenze finite

Cominciamo la nostra discussione sui metodi alle differenze finite definendo una griglia di punti nel piano (x, t) . Fissati un passo di discretizzazione spaziale $\Delta x = h$ e temporale $\Delta t = k$, la griglia sarà data dai punti nodali del piano

$(x_m, t_n) = (mh, nk)$, per valori interi arbitrari m e n . Per una funzione v definita sulla griglia, scriveremo $v_{m,n}$ per il suo valore al nodo (x_m, t_n) , e questa notazione sarà estesa naturalmente anche a funzione definite ovunque per indicare i valori sui nodi.

L'idea di base dei metodi alle differenze finite è quella di sostituire, nell'equazione da approssimare, ad ogni derivata un rapporto incrementale finito. Per esempio, per una funzione regolare u la derivata temporale $\partial_t u(x_m, t_n)$ può essere approssimata sia dalla differenza in avanti (forward)

$$\frac{\Delta_{+t}u}{\Delta t} := \frac{u(mh, (n+1)k) - u(mh, nk)}{k},$$

che dalla differenza all'indietro (backward)

$$\frac{\Delta_{-t}u}{\Delta t} := \frac{u(mh, nk) - u(mh, (n-1)k)}{k},$$

che infine dalla differenza centrata (central)

$$\frac{\delta_t u}{2\Delta t} := \frac{u(mh, (n+1)k) - u(mh, (n-1)k)}{2k}.$$

Analogamente si procede nel caso di una derivata spaziale, ed eventualmente anche nel caso di più variabili spaziali. In particolare la derivata seconda spaziale sarà spesso approssimata dalla differenza seconda centrata

$$\frac{\delta_x^2 u}{(\Delta x)^2} := \frac{u(mh, (n+1)k) - 2u(mh, nk) + u(mh, (n-1)k)}{h^2}.$$

Questi schemi sono dati da un sistema algebrico

$$P_{h,k}v = 0$$

che agisce sullo spazio delle funzioni definite sui nodi della griglia.

Usando diverse approssimazioni per le derivate si ottengono chiaramente schemi numerici differenti. Per esempio per l'equazione (1.1) otteniamo i seguenti schemi.

$$\frac{v_{m,n+1} - v_{m,n}}{k} + a \frac{v_{m+1,n} - v_{m,n}}{h} = 0 \quad (1.10)$$

$$\frac{v_{m,n+1} - v_{m,n}}{k} + a \frac{v_{m,n} - v_{m-1,n}}{h} = 0 \quad (1.11)$$

$$\frac{v_{m,n+1} - v_{m,n}}{k} + a \frac{v_{m+1,n} - v_{m-1,n}}{2h} = 0 \quad (1.12)$$

$$\frac{v_{m,n+1} - v_{m,n-1}}{2k} + a \frac{v_{m+1,n} - v_{m-1,n}}{2h} = 0 \quad (1.13)$$

$$\frac{v_{m,n+1} - \frac{1}{2}(v_{m+1,n} + v_{m-1,n})}{k} + a \frac{v_{m+1,n} - v_{m-1,n}}{2h} = 0 \quad (1.14)$$

Il primo schema (1.10) è un forward-forward o upwind con $a < 0$ (le differenze prendono la direzione delle caratteristiche), perchè in tale approssimazione è usata la differenza in avanti sia in tempo che in spazio. Analogamente lo schema (1.11) è un forward-backward o upwind con $a > 0$ e (1.12) è un forward-central. Lo schema (1.13) è un central-central e viene chiamato “leapfrog” (=saltaranocchio) e il (1.14), che ha una struttura leggermente differente, è chiamato schema di Lax–Friedrichs.

Come si è visto il modo di derivare gli schemi è estremamente semplice e permette in principio di approssimare qualsiasi equazione differenziale, anche a coefficienti variabili o non lineare. Questo non vuol dire che tutti gli schemi così ottenuti siano ugualmente efficienti e veloci, e in effetti non è detto nemmeno che tutti approssimino la soluzione del problema (1.1). Nel seguito ci occuperemo di stabilire quindi dei criteri che assicurino la convergenza, ossia la proprietà di approssimare la soluzione del problema limite quando i passi di discretizzazione spaziale e temporale tendono entrambi a zero, e quindi, tra gli schemi convergenti, cercheremo di selezionare quelli che danno un’approssimazione più accurata con il minor numero di punti.

Una prima osservazione da fare riguardo agli schemi (1.10)–(1.14) è che tutti possono essere riformulati esprimendo $v_{m,n+1}$ come una combinazione lineare dei valori di v al livello n o, nel caso dello schema leapfrog, anche al livello $n - 1$. Per esempio lo schema (1.10) può essere riscritto come

$$v_{m,n+1} = (1 + a\lambda)v_{m,n} - a\lambda v_{m+1,n},$$

avendo posto $\lambda := k/h$. Questa quantità λ , detto anche rapporto di CFL, apparirà spesso nel seguito e il suo valore sarà determinante nello stabilire le proprietà degli schemi. Gli schemi che dipendono dai valori al passo immediatamente precedente sono chiamati schemi ad un passo, altrimenti, come nel caso dello schema leapfrog, si parla di schemi “multistep”. Per gli schemi ad un passo, una volta fissato il dato iniziale ed una sua ragionevole approssimazione sulla griglia di discretizzazione (per esempio il suo valore sui nodi), la soluzione approssimata viene costruita facilmente per ricorrenza. Per gli schemi multistep a k passi, $k > 1$, bisogna invece prescrivere una procedura di inizializzazione per assegnare la soluzione nei livelli k occorrenti. Per fare questo usualmente si utilizza un metodo ad un passo per i primi k passi e poi si continua con il metodo proposto.

Esempio 1.2 Prima di procedere con l’analisi degli schemi a differenze finite, analizziamo alcuni risultati numerici ottenuti utilizzando lo schema di Lax–Friedrichs. Si consideri l’equazione

$$u_t + u_x = 0$$

con il dato iniziale

$$u_0 = \begin{cases} 1 - |x|, & \text{per } |x| \leq 1, \\ 0, & \text{per } |x| > 1, \end{cases}$$

la cui soluzione è ottenuta semplicemente ponendo $u(x, t) = u_0(x - t)$, ossia traslando verso destra a velocità 1 il dato iniziale. Per calcolare la soluzione

noi ci limiteremo al caso $0 \leq t \leq 2$, per cui il supporto della soluzione sarà contenuto solo nell'intervallo $-1 \leq x \leq 3$ e basterà imporre le condizioni al bordo nulle per la v . La formula relativa allo schema di Lax–Friedrichs è

$$v_{m,n+1} = \frac{1}{2}(v_{m+1,n} + v_{m-1,n}) - \frac{1}{2}\lambda(v_{m+1,n} - v_{m-1,n}),$$

eccetto negli estremi destri e sinistri dove imporremo $v = 0$. La soluzione al tempo $t = 1.6$ per $\lambda = 0.8$ appare in figura 1.1, dove la soluzione esatta è data dalla linea continua. Si osserva allora un sostanziale rispetto della soluzione, che viene però smussata negli spigoli a causa del fenomeno della cosiddetta “viscosità numerica”, che verrà discusso nel seguito. Mantenendo fisso il valore di λ e diminuendo il valore di h , e quindi di k , la precisione migliora e possiamo constatare una sostanziale convergenza numerica. Se però usiamo $\lambda = 1.6$, come

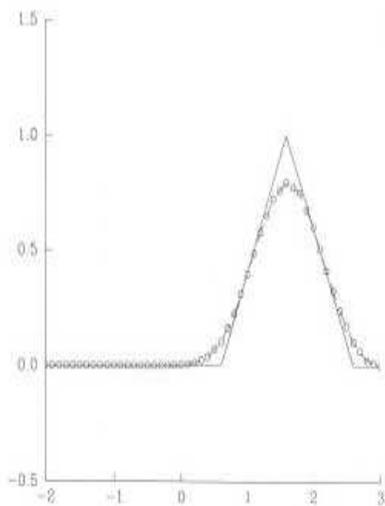


Figura 1.1: Lax–Friedrichs, $\lambda = 0.8$

mostrato nella figura 1.2 per il tempo $t = 0.8$, vediamo che la soluzione numerica si comporta in modo anomalo, oscillando in modo incontrollato intorno alla soluzione esatta. Questo fatto, che verrà ripreso e spiegato in seguito, mostra l'importanza di una corretta scelta del parametro λ .

1.3 Convergenza e consistenza

Come già osservato, la proprietà principale di uno schema è la sua convergenza verso la soluzione esatta al tendere dei passi di discretizzazione h e k verso lo zero. Nel seguito estenderemo la nozione di schema convergente ad operatori più

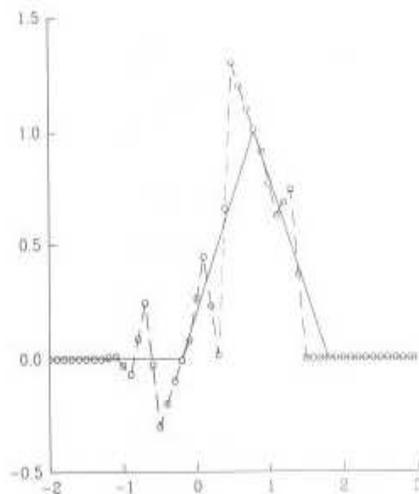


Figura 1.2: Lax–Friedrichs, $\lambda = 1.6$

generali, ma per ora vogliamo limitarci solo alle equazioni scalari del prim'ordine iperboliche.

Definizione 1.2 *Uno schema alle differenze finite ad un passo dato da un operatore discreto $P_{h,k}v = 0$ è convergente per l'equazione differenziale (1.1) se per qualsiasi soluzione u dell'equazione con la condizione iniziale (1.2), data la soluzione $v_{m,n}$ dello schema alle differenze finite, tale che $v_{m,0}$ converge a u_0 quando mh converge a x , allora $v_{m,n}$ converge a $u(x,t)$ se (mh, nk) tende a (x,t) al tendere di h e k a 0.*

Per gli schemi multistep la definizione si estende senza difficoltà specificando la procedura di inizializzazione.

Per confrontare una soluzione discreta v con una definita ovunque u , dobbiamo introdurre delle opportune topologie. Sia Ω un intervallo reale e J_Ω l'insieme dei valori dell'indice m tali che $mh \in \Omega$. I valori di una discretizzazione $v = \{v_{m,n}\}$ al livello temporale n saranno definiti da

$$v^n := \{v_{m,n}, m \in J_\Omega\}.$$

Una notazione analoga sarà utilizzata per la restrizione della funzione u sui nodi della griglia.

Possiamo allora considerare due tipi di norma. Per prima cosa la norma del massimo, che è data da

$$\|v^n\|_{\infty, h} := \max\{|v_{m,n}|, m \in J_\Omega\}. \quad (1.15)$$

Utilizzando la stessa norma per valutare u^n avremo un'approssimazione (dal basso) della norma usuale in L^∞ della funzione $u(\cdot, t_n)$, che in generale non coinciderà con essa.

Un'altra norma largamente utilizzata, specialmente nel contesto dei problemi lineari, sarà la norma discreta L^2

$$\|v^n\|_{2,h} := \left\{ \sum_{m \in J_\Omega} h |v_{m,n}|^2 \right\}^{\frac{1}{2}}. \quad (1.16)$$

La scelta della norma discreta da utilizzare non modifica in modo sostanziale l'analisi degli schemi, se non dal punto di vista tecnico, ma è importante che sia sempre consistente.

Si osservi che queste norme sono state scelte per cercare di mantenere la discussione al livello elementare. Nel Capitolo 3, seguendo la pratica corrente dell'analisi numerica, definiremo le approssimazioni numeriche in tutto lo spazio, per esempio assegnandole valori costanti nei rettangoli della griglia, e le norme scelte saranno allora le norme usuali degli spazi considerati.

Nell'esempio precedente si vedeva che per $\lambda = 0.8$ gli schemi di Lax-Friedrichs e leapfrog erano convergenti, mentre lo schema di Lax-Friedrich mostrava un cattivo comportamento per $\lambda = 1.6$. Dimostrare che uno schema è convergente è spesso difficile per un problema di natura generale e procedendo in maniera diretta. Per le equazioni lineari a coefficienti costanti si può però caratterizzare la convergenza in modo completo utilizzando delle condizioni di facile verifica, la *consistenza* e la *stabilità*.

Definizione 1.3 *L'errore di troncatura T di uno schema alle differenze finite per l'equazione (1.1) $P_{h,k}v = 0$ è definito sostituendo nello schema alla soluzione numerica una soluzione esatta u di (1.1) abbastanza regolare, ossia*

$$T(x, t) = P_{h,k}u(x, t). \quad (1.17)$$

Definizione 1.4 *Uno schema si dice consistente con l'equazione (1.1) se*

$$T(x, t) \rightarrow 0 \text{ se } h, k \rightarrow 0,$$

per ogni (x, t) nel dominio di interesse.

Per alcuni schemi dovremmo dare delle restrizioni al modo in cui h e k tendono a zero. Le soluzioni dovranno essere abbastanza regolari nel senso di avere abbastanza derivate per eseguire i calcoli.

Esaminiamo adesso la consistenza degli schemi proposti

Esempio 1.3 **Lo schema upwind con $a < 0$.** In questo caso sia u soluzione di (1.1), ossia $u_t + au_x = 0$ e

$$P_{h,k}v = \frac{v_{m,n+1} - v_{m,n}}{k} + a \frac{v_{m+1,n} - v_{m,n}}{h} = 0.$$

Facendo lo sviluppo di Taylor di u in un intorno di (mh, nk) otteniamo allora, ponendo $u_{m,n} = u(mh, nk)$:

$$u_{m,n+1} = u_{m,n} + ku_t + \frac{1}{2}k^2u_{tt} + O(k^3),$$

$$u_{m+1,n} = u_{m,n} + hu_x + \frac{1}{2}h^2u_{xx} + O(h^3),$$

le derivate essendo calcolate nel punto stesso (mh, nk) . Da questo si deduce che

$$T = P_{h,k}u = u_t + au_x + \frac{1}{2}k^2u_{tt} + \frac{1}{2}ah^2u_{xx} + O(k^2) + O(h^2), \quad (1.18)$$

che tende a zero se h e k tendono a zero, per cui lo schema è consistente.

Esempio 1.4 Lo schema di Lax–Friedrichs. Per questo schema abbiamo

$$P_{h,k}u = \frac{u_{m,n+1} - \frac{1}{2}(u_{m+1,n} + u_{m-1,n})}{k} + a \frac{u_{m+1,n} - u_{m-1,n}}{2h}.$$

Usando di nuovo gli sviluppi di Taylor si vede che

$$u_{m\pm 1,n} = u_{m,n} \pm hu_x + \frac{1}{2}h^2u_{xx} \pm \frac{1}{6}h^3u_{xxx} + O(h^4).$$

Otteniamo quindi

$$\frac{1}{2}(u_{m+1,n} + u_{m-1,n}) = u_{m,n} + \frac{1}{2}h^2u_{xx} + O(h^4),$$

e

$$\frac{u_{m+1,n} - u_{m-1,n}}{2h} = u_x + \frac{1}{6}h^3u_{xxx} + O(h^4).$$

Sostituendo queste espressioni nello schema otteniamo

$$\begin{aligned} P_{h,k}u &= u_t + au_x + \frac{1}{2}k^2u_{tt} - \frac{1}{2}k^{-1}h^2u_{xx} \\ &\quad + \frac{1}{6}ah^3u_{xxx} + O(h^4 + k^{-1}h^4 + k^2), \end{aligned}$$

che, essendo u una soluzione, tende a zero se h e k tendono a zero con $k^{-1}h^2 \rightarrow 0$.

Vediamo ora come l'errore di troncatura ci possa dare in alcuni esempi elementari una stima dell'errore di approssimazione numerica che stiamo commettendo. Per un dato schema denotiamo con $T_{m,n}$ l'errore di troncatura in (x_n, t_m) e sia data una stima di questo errore su di un intervallo fissato $[0, T_0]$ con

$$|T_{m,n}| \leq \Gamma, \quad (1.19)$$

se $nk \leq T_0$. Si denoti quindi con e l'errore $v - u$ nell'approssimazione; più precisamente sia

$$e_{m,n} = v_{m,n} - u_{m,n}. \quad (1.20)$$

Analizziamo ora per esempio lo schema upwind con $a < 0$. Inserendo la quantità $e_{m,n}$ nello schema (1.10) troviamo

$$e_{m,n+1} = (1 + a\lambda)e_{m,n} - a\lambda e_{m+1,n} - kT_{m,n}. \quad (1.21)$$

Osserviamo allora che se $-1 \leq a\lambda \leq 0$, i coefficienti dei due primi termini a destra sono non negativi e sommati danno 1. Introduciamo quindi la norma del massimo dell'errore

$$E^n := \max\{|e_{m,n}|, m \in J_\Omega\}.$$

A questo punto utilizzando la non negatività dei coefficienti di $e_{m,n}$ in (1.20) si ottiene facilmente

$$E^{n+1} \leq E^n + k\Gamma, \quad (1.22)$$

ossia, essendo $E^0 = 0$,

$$E^n \leq nk\Gamma \leq T_0\Gamma. \quad (1.23)$$

Utilizzando la stima (1.18) abbiamo però che

$$\Gamma \leq \frac{1}{2}k\{M_{tt} + \frac{|a|}{\lambda}M_{xx}\},$$

dove M_{tt} e M_{xx} sono la norma del massimo per le derivate u_{tt} e u_{xx} rispettivamente. Abbiamo quindi che l'errore tende a zero, ovvero l'approssimazione numerica converge, al tendere di $k = \lambda h$ a zero, con λ costante, se la soluzione limite ha le derivate seconde u_{tt} e u_{xx} uniformemente limitate su tutto il dominio. Un ragionamento analogo si applica allo schema di Lax–Friedrichs e ovviamente allo schema upwind con $a > 0$.

In pratica la consistenza e la convergenza sono legati dal fatto che la prima dice che la soluzione esatta è una soluzione approssimata dello schema numerico, mentre la seconda dice che la soluzione dello schema numerico è una soluzione approssimata dell'equazione. Bisogna però stare attenti al fatto che la consistenza da sola non implica necessariamente la convergenza, come si vede nell'esempio successivo.

Esempio 1.5 Lo schema forward-forward non sempre converge. Si riprenda la schema (1.10), però nel caso $a = 1$. Possiamo allora scrivere lo schema come

$$v_{m,n+1} = (1 + \lambda)v_{m,n} - \lambda v_{m+1,n}, \quad (1.24)$$

avendo posto come al solito $\lambda := k/h$. Si prenda come dato iniziale

$$u_0 = \begin{cases} 1, & \text{per } -1 \leq x \leq 0, \\ 0, & \text{altrove.} \end{cases}$$

La soluzione u sarà allora data da una semplice traslazione verso destra con velocità t , ed in particolare ci saranno valori positivi di x per cui $u = 1$. Analogamente per lo schema alle differenze si prenda come dato iniziale

$$v_{m,0} = \begin{cases} 1, & \text{per } -1 \leq mh \leq 0, \\ 0, & \text{altrove.} \end{cases}$$

Da (1.24) si vede che la soluzione numerica dipende solo dai valori di m maggiori o uguali al valore nel punto, per cui avremo

$$v_{m,n} = 0$$

per tutti i valori $m > 0, n \geq 0$. Per questo, essendo l'approssimazione $v_{m,n}$ identicamente nulla per $x > 0$ non può tendere a u nel limite in nessun senso ragionevole.

1.4 Stabilità

Abbiamo visto che la sola condizione di consistenza non è sufficiente in generale a garantire la convergenza. La condizione supplementare da tenere in considerazione è la *stabilità*. Per capire questa condizione si deve pensare che se la successione approssimante converge in una qualche topologia, allora deve essere almeno limitata nella stessa topologia.

Definizione 1.5 *Uno schema alle differenze finite a un passo $P_{h,k}v = 0$ che approssima l'equazione (1.1) è stabile nella norma $\|\cdot\|_{\cdot,h}$ nella regione Ω se per ogni $T > 0$ esiste una costante C_T tale che, per ogni v^0 ,*

$$\|v^n\|_{\cdot,h} \leq C_T \|v^0\|_{\cdot,h}, \quad (1.25)$$

per $0 \leq nk \leq T$.

Ci concentreremo nel seguito sugli schemi ad un passo espliciti per l'equazione (1.1) che possono dunque essere scritti come

$$v^{n+1} = Cv^n, \quad (1.26)$$

dove C è un operatore lineare alle differenze indipendente da n del tipo

$$v_{m,n+1} = (Cv^n)_m = \sum_{k \in J_\Omega} c_{m,k} v_{k,n}. \quad (1.27)$$

I coefficienti $c_{m,k}$ potranno in effetti dipendere da m e k per poter includere in seguito il caso a coefficienti variabili e le possibili condizioni al bordo. Eventuali termini di sorgente potranno anche essere inclusi senza difficoltà e non cambieranno le proprietà essenziali dello schema.

Poichè stiamo trattando il caso lineare, è facile vedere che in generale la nostra richiesta di stabilità equivale ad una condizione sui coefficienti dell'operatore C del tipo

$$\|C^n\| \leq C_T, \quad (1.28)$$

per $0 \leq nk \leq T$, usando un'opportuna norma operatoriale.

Diremo infine che uno schema è convergente nella norma prescelta se

$$\|v^n - u^n\|_{\cdot,h} \rightarrow 0, \quad (1.29)$$

con $\lambda = k/h$ fissato e $k \rightarrow 0$ e per $nk \rightarrow t \in (0, T)$, per ogni dato iniziale u_0 per cui il problema di Cauchy per (1.1) è ben posto nella norma prescelta. Nel caso presente per la norma del massimo serviranno dati limitati e per la norma di tipo L^2 serviranno dati nello stesso spazio.

A questo punto abbiamo tutti gli elementi per enunciare il risultato fondamentale che vale più generalmente per tutti gli schemi alle differenze finite che approssimino problemi di tipo lineare.

Teorema 1.6 (Il Teorema di equivalenza di Lax–Richtmyer) *Uno schema alle differenze finite consistente per un problema lineare ben posto di evoluzione è convergente se e solo se è stabile.*

Cenni sulla dimostrazione. Ci limiteremo a far vedere la sufficienza delle condizioni, per una dimostrazione completa si veda [RM, St]. Si ha che dalla definizione di errore di troncatura, dato uno schema nella forma (1.27), segue che

$$v_{m,n+1} - u_{m,n+1} = \sum_{k=-l}^l c_{m,k}(v_{k,n} - u_{k,n}) - kT_n.$$

Se poniamo $v_{m,0} - u_{m,0}$, abbiamo che per induzione

$$v^{n+1} - u^{n+1} = -k \sum_{r=0}^n C^r T_{n-r}.$$

A questo punto, dalla (1.28) si ha

$$\|v^n - u^n\|_{.,h} \leq kC_T \sum_{r=0}^n \|T_r\|_{.,h}.$$

La convergenza segue dunque semplicemente dalla definizione di consistenza, almeno se la soluzione è abbastanza regolare. Per dati poco regolari si può far vedere che segue dalla buona positura stessa del problema, approssimando i dati iniziali poco regolari con delle funzioni lisce. \square

Il Teorema di equivalenza di Lax–Richtmyer, stabilito per la prima volta da Peter Lax nel 1953, quando aveva soltanto 27 anni, è uno dei migliori esempi di teorema matematico. Riesce ossia a caratterizzare un concetto utile, ma difficile come la convergenza, essenzialmente legata a nozioni di carattere topologico, ad altri criteri più semplici da stabilire come la stabilità e la consistenza. In pratica la consistenza è un requisito abbastanza facile da verificare. La stabilità richiede un po' più di lavoro, ma come vedremo in seguito è possibile operare con manipolazioni essenzialmente di tipo algebrico. Inoltre questa caratterizzazione è ottimale. Ossia non dice soltanto che la consistenza e la stabilità implicano la convergenza, ma anche che uno schema instabile o non consistente non può essere convergente. Questo permette di scartare eventuali candidati con una semplice verifica a priori.

Passiamo ora a considerare la classe degli schemi a tre punti, ossia degli schemi della forma

$$v_{m,n+1} = \alpha v_{m+1,n} + \beta v_{m,n} + \gamma v_{m-1,n}, \quad (1.30)$$

con α , β e γ costanti da fissare. Abbiamo intanto questo primo risultato di stabilità in L^2 .

Proposizione 1.7 *Sia dato uno schema a tre punti della forma (1.30) per l'equazione (1.1). Una condizione sufficiente affinché lo schema risulti stabile rispetto alla norma (1.16) è che valga la disuguaglianza*

$$|\alpha| + |\beta| + |\gamma| \leq 1. \quad (1.31)$$

Dimostrazione. Per semplicità supporremo che l'intervallo di definizione dello schema sia tutta la retta reale, ossia $\Omega = \mathbb{R}$. Abbiamo dunque

$$\begin{aligned} \sum_{m=-\infty}^{\infty} |v_{m,n}|^2 &= \sum_{m=-\infty}^{\infty} |\alpha v_{m+1,n} + \beta v_{m,n} + \gamma v_{m-1,n}|^2 \\ &\leq \sum_{m=-\infty}^{\infty} (|\alpha|^2 |v_{m+1,n}|^2 + |\beta|^2 |v_{m,n}|^2 + |\gamma|^2 |v_{m-1,n}|^2 \\ &\quad + 2|\alpha||\beta| |v_{m+1,n}| |v_{m,n}| + 2|\beta||\gamma| |v_{m,n}| |v_{m-1,n}| + 2|\alpha||\gamma| |v_{m+1,n}| |v_{m-1,n}|) \\ &\leq \sum_{m=-\infty}^{\infty} (|\alpha|^2 + 2|\alpha||\beta| + |\beta|^2 + 2|\beta||\gamma| + |\gamma|^2 + 2|\alpha||\gamma|) |v_{m,n}|^2 \\ &= (|\alpha| + |\beta| + |\gamma|)^2 \sum_{m=-\infty}^{\infty} |v_{m,n}|^2, \end{aligned}$$

avendo usato tra l'altro la relazione elementare $2xy \leq x^2 + y^2$. Questo termina la dimostrazione. \square

Per assicurare la consistenza con l'equazione (1.1) abbiamo la seguente proposizione.

Proposizione 1.8 *Sia il rapporto $\frac{k}{h}$ uguale a una costante fissata $\lambda > 0$. Allora uno schema a tre punti della forma (1.30) è consistente con l'equazione (1.1) se e solo se $\alpha + \beta + \gamma = 1$ e $\gamma - \alpha = \lambda a$.*

Dimostrazione. Dagli sviluppi di Taylor abbiamo al solito

$$\begin{aligned} u_{m,n+1} &= u_{m,n} + k u_t + \frac{1}{2} k^2 u_{tt} + O(k^3), \\ u_{m\pm 1,n} &= u_{m,n} \pm h u_x + \frac{1}{2} h^2 u_{xx} \pm \frac{1}{6} h^3 u_{xxx} + O(h^4). \end{aligned}$$

Si ottiene quindi

$$\begin{aligned} & \frac{1}{k} [u_{m,n+1} - (\alpha u_{m+1,n} + \beta u_{m,n} + \gamma u_{m-1,n})] \\ &= \frac{1-\alpha+\beta+\gamma}{k} u_{m,n} + u_t + (\gamma - \alpha)\lambda^{-1} u_x + \frac{1}{2}k(u_{tt} - \lambda^{-2}(\alpha + \gamma)u_{xx}) + O(k^2). \end{aligned} \quad (1.32)$$

Questo termina la dimostrazione. \square

Si osservi che stiamo in effetti imponendo una limitazione nel modo di convergere delle discretizzazioni spaziali e temporali, richiedendo che λ sia una costante. Da questa proposizione si vede che la classe degli schemi lineari a tre punti che sono consistenti con l'equazione (1.1) forma una famiglia ad un parametro. Ponendo

$$q = \alpha + \gamma, \quad (1.33)$$

possiamo riscrivere tutti questi schemi nella forma cosiddetta “viscosa”

$$v_{m,n+1} = v_{m,n} - \frac{a\lambda}{2}(v_{m+1,n} - v_{m-1,n}) + \frac{q}{2}(v_{m+1,n} - 2v_{m,n} + v_{m-1,n}). \quad (1.34)$$

Il parametro q è legato infatti alla viscosità numerica dello schema, in quanto formalmente lo schema è consistente all'ordine 2, con l'equazione viscosa

$$u_t + au_x = \frac{1}{2}\lambda h \left(\frac{q}{\lambda^2} - a^2 \right) u_{xx}. \quad (1.35)$$

In particolare abbiamo la seguente tabella dei coefficienti.

schema	(α, β, γ)	q	convergenza
upwind con $a < 0$	$(-\lambda a, (1+\lambda a), 0)$	$-\lambda a$	$-1 \leq \lambda a \leq 0$
upwind con $a > 0$	$(0, (1-\lambda a), \lambda a)$	λa	$0 \leq \lambda a \leq 1$
forward-central	$(\frac{-\lambda a}{2}, 1, \frac{\lambda a}{2})$	0	mai
Lax-Friedrichs	$(\frac{1-\lambda a}{2}, 0, \frac{1+\lambda a}{2})$	1	$ \lambda a \leq 1$

Vedremo in seguito come le condizioni di stabilità trovate grazie alla (1.31) siano in questo caso anche necessarie. Osserviamo tuttavia che tra gli schemi esaminati la viscosità massima, in regime di stabilità, sia quella dello schema di Lax-Friedrichs, mentre il forward-central, che come vedremo non è mai stabile, ha viscosità negativa nell'equazione equivalente.

1.5 La condizione di CFL

Prima di passare a discutere più in dettaglio la stabilità dei vari schemi, essenzialmente grazie all'analisi di von Neumann, presentiamo una condizione necessaria per la convergenza di facile verifica e che permette quindi una veloce analisi preliminare. Questa condizione fu proposta nel 1928 da R. Courant,

K.O. Friedrichs e H. Lewy nel loro lavoro fondamentale sui metodi alle differenze finite [CFL] ed è conosciuta oggi come la condizione di CFL. Essa esprime una condizione necessaria di convergenza nei termini del dominio di dipendenza della soluzione.

Prendiamo ad esempio la soluzione del problema (1.1)–(1.2). Questa soluzione è completamente determinata in un punto (x, t) a partire dal valore del dato iniziale nel punto $x - at$. La condizione di CFL stabilisce che per uno schema convergente il dominio di dipendenza dello schema numerico deve contenere il dominio di dipendenza dell'equazione differenziale. Nel caso dell'equazione (1.1), eventualmente anche nel caso disomogeneo (1.5), questo si riduce alla caratteristica retrograda uscente dal punto (x, t) , ossia la retta di equazione $\xi = x - a(t - s)$.

Teorema 1.9 *Se fissiamo il rapporto $\lambda = k/h = \text{const.}$ per uno schema esplicito a tre punti per l'equazione (1.1) della forma (1.30), allora una condizione necessaria per la stabilità è la seguente condizione detta di CFL:*

$$|\lambda a| \leq 1. \quad (1.36)$$

In particolare è facile estendere questo risultato ai sistemi lineari fortemente iperbolici del tipo (1.7) per cui vale quindi la seguente regola generale: *non ci sono schemi espliciti incondizionatamente stabili e consistenti alle differenze finite per sistemi "fortemente iperbolici".* La stabilità incondizionata potrà infatti essere solo recuperata al prezzo di una perdita di accuratezza dello schema per la classe dei cosiddetti schemi *impliciti*.

Dimostrazione. Si consideri il punto $(0, 1)$ nel piano (x, t) . Il valore della soluzione u in quel punto dipende solo dal valore del dato iniziale nel punto $x - at$. Ma per lo schema numerico a tre punti, noi troviamo che $v_{o,n}$ dipende da $v_{m,0}$ per $|m| \leq n$. Sia ora $|\lambda a| > 1$. Se si prende $h = \lambda^{-1}k$ e $kn = 1$, si ha che $mh \leq \lambda^{-1}kn = \lambda^{-1}$. Per cui $v_{m,0}$ dipende su x solo per $|x| \leq \lambda^{-1} < |a|$. Per questo $v_{o,n}$ non può convergere a $u(0, 1)$ quando k tende a zero, poichè basterebbe modificare arbitrariamente il dato iniziale per $x = \pm a$, e quindi il valore di $u(0, 1)$, senza modificare $v_{m,0}$. \square

1.6 Stime nella norma del massimo e monotonia

Un modo semplice per trovare stime della norma del massimo è quello di costruire degli schemi che verifichino la proprietà del confronto monotono, ossia se due dati iniziali discretizzati verificano $v_{m,0} \leq \tilde{v}_{m,0}$ allora per ogni $n > 0$ vale $v_{m,n} \leq \tilde{v}_{m,n}$.

Definizione 1.6 *Uno schema nella forma (1.27) è detto essere monotono se tutti i coefficienti $c_{m,k} \geq 0$.*

Si vede facilmente che lo schema è monotono se e solo se verifica la proprietà del confronto monotono. Abbiamo la seguente stima di limitatezza.

Teorema 1.10 *Sia dato uno schema monotono della forma (1.27). Allora, posto*

$$v_{min} := \min\{v_{m,0}\}; \quad v_{max} := \max\{v_{m,0}\},$$

avremo che

$$v_{min} \leq v_{m,n} \leq v_{max}. \quad (1.37)$$

Dimostrazione. Si prenda per esempio $\tilde{v}_{m,0} = v_{max} = const.$ per dimostrare la disuguaglianza di destra. Analogamente per quella di sinistra. \square

Dal teorema di Lax–Richtmyer abbiamo allora immediatamente la convergenza di tutti gli schemi consistenti e monotoni (per una dimostrazione diretta si può procedere anche stimando direttamente l'errore numerico come in (1.21)–(1.23)). È facile inoltre caratterizzare gli schemi monotoni tra tutti gli schemi consistenti a tre punti della forma (1.30). Per enunciare questa caratterizzazione utilizziamo però la forma viscosa (1.34)

Proposizione 1.11 *Uno schema esplicito a tre punti consistente con l'equazione (1.1), nella forma viscosa (1.34), è monotono se e solo se*

$$\lambda|a| \leq q \leq 1. \quad (1.38)$$

Dimostrazione. Basta osservare che valgono le seguenti uguaglianze

$$\alpha = \frac{1}{2}(q - a\lambda), \quad \beta = 1 - q, \quad \gamma = \frac{1}{2}(q + a\lambda).$$

\square

Vediamo dunque che la condizione di CFL è ancora necessaria per la convergenza. Inoltre gli schemi upwind e Lax–Friedrichs sono monotoni. In particolare lo schema monotono con la più bassa viscosità numerica è lo schema upwind, mentre Lax–Friedrichs è quello con la più alta. Osserviamo infine come la condizione (1.38) sia in effetti equivalente alla condizione (1.31) che avevamo trovato per la norma L^2 .

Sfortunatamente la classe degli schemi monotoni è troppo ristretta per contenere delle approssimazioni più accurate, per cui sarà necessario considerare la stabilità di tipo L^2 , che risulterà essere meno restrittiva.

1.7 L'analisi di von Neumann

1.7.1 Cenni sull'analisi di Fourier

Lo strumento più efficace per l'analisi delle equazioni lineari alle derivate parziali è senz'altro l'uso estensivo della trasformata e delle serie di Fourier. Nel caso a coefficienti costanti questo permette di stabilire in modo diretto la buona positura e le proprietà qualitative delle soluzioni di larghe classi di problemi differenziali. In particolar modo è possibile caratterizzare in modo completo la buona positura in L^2 dei problemi di tipo iperbolico e parabolico, sia nel caso scalare che per i sistemi.

Se si considera una funzione $v = \{v_m\}$ definita su una griglia “spaziale” di passo h , la sua serie di Fourier è data da

$$\hat{v}(\xi) = \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} h e^{-imh\xi} v_m, \quad (1.39)$$

per $\xi \in [-\pi/h, \pi/h]$. Si ha allora che $\hat{v}(-\pi/h) = \hat{v}(\pi/h)$ e la formula di inversione dà

$$v_m = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{imh\xi} \hat{v}(\xi) d\xi. \quad (1.40)$$

Si noti che di solito uno parte con una funzione periodica sull'intervallo $[-\pi/h, \pi/h]$ e mostra che sotto opportune ipotesi vale la formula di rappresentazione (1.39) con i coefficienti dati da (1.40). Una delle conseguenze più importanti di questa definizione è la relazione di Parseval. Sia infatti v limitata nella norma di tipo L^2 (1.16), allora $\hat{v} \in L^2$ e vale la relazione

$$\|\hat{v}\|_h^2 = \int_{-\pi/h}^{\pi/h} |\hat{v}(\xi)|^2 d\xi = \sum_{m=-\infty}^{\infty} |v_m|^2 h = \|v\|_h^2. \quad (1.41)$$

Questa relazione ci permetterà di stabilire nel modo ottimale la relazione (1.25) nel caso L^2 , passando all'analogia disuguaglianza per la funzione trasformata.

Si osservi che una simile relazione non è valida in generale per la norma del massimo e questo ne rende a priori più difficile l'uso nel Teorema di Lax-Richtmyer. Inoltre, nel caso dei sistemi iperbolici, i principi di massimo non sono in generale validi ed il contesto di tipo L^2 risulta più naturale.

Prima di addentrarci ulteriormente nell'analisi degli schemi numerici, vale la pena di dare un cenno alle proprietà della trasformata di Fourier che viene utilizzata per la risoluzione dei problemi differenziali. Sia ad esempio u una soluzione dell'equazione (1.1) e \hat{u} la sua trasformata di Fourier nella variabile spaziale, ossia

$$\hat{u}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega x} u(x) dx. \quad (1.42)$$

Allora \hat{u} risolve l'equazione differenziale ordinaria

$$\hat{u}_t = -ia\omega\hat{u},$$

che ha come unica soluzione

$$\hat{u}(\omega, t) = e^{-iat\omega} \hat{u}_0(\omega).$$

Si noti ancora che la norma L^2 viene conservata nell'evoluzione dell'equazione, grazie ancora alla relazione di Parseval per la trasformata di Fourier, e le proprietà di regolarità della soluzione vengono trasformate in proprietà di decrescenza all'infinito della \hat{u} e viceversa. Questo spiega perchè le soluzioni irregolari presentano “oscillazioni” sulle frequenze più alte, che in questo caso si trovano ad avere grande ampiezza.

1.7.2 L'analisi di von Neumann

Per illustrare il metodo considereremo un caso particolare, lo schema forward-backward (1.11) per l'equazione (1.1):

$$\frac{v_{m,n+1} - v_{m,n}}{k} + a \frac{v_{m,n} - v_{m-1,n}}{h} = 0, \quad (1.43)$$

che riscriviamo in forma compatta come

$$v_{m,n+1} = (1 - a\lambda)v_{m,n} + a\lambda v_{m-1,n}, \quad (1.44)$$

sempre ponendo $\lambda := k/h$. Dalla formula (1.40) abbiamo

$$v_{m,n} = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{imh\xi} \hat{v}^n(\xi) d\xi.$$

Sostituendo quindi nella (1.44) si ottiene la relazione

$$v_{m,n+1} = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{imh\xi} [(1 - a\lambda) + a\lambda e^{-ih\xi}] \hat{v}^n(\xi) d\xi \quad (1.45)$$

da cui si deduce la relazione per ricorrenza

$$\hat{v}^{n+1}(\xi) = g(h\xi) \hat{v}^n(\xi), \quad (1.46)$$

con $g(h\xi) := (1 - a\lambda) + a\lambda e^{-ih\xi}$. Questo mostra come l'avanzamento di un passo temporale dello schema provochi un'amplificazione di un fattore $g(h\xi)$ della trasformata numerica, ottenendo così

$$\hat{v}^n(\xi) = g(h\xi)^n \hat{v}^0(\xi). \quad (1.47)$$

Dalla relazione di Parseval otteniamo allora

$$\begin{aligned} \sum_{m=-\infty}^{\infty} h |v_{m,n}|^2 &= \int_{-\pi/h}^{\pi/h} |\hat{v}^n(\xi)|^2 d\xi \\ &= \int_{-\pi/h}^{\pi/h} |g(h\xi)|^{2n} |\hat{v}^0(\xi)|^2 d\xi. \end{aligned}$$

Per ottenere quindi la stabilità richiesta dalla (1.25) nel caso L^2 dobbiamo dimostrare che $|g(h\xi)|^{2n}$ è opportunamente limitata in n . Poniamo $\chi = h\xi$. Abbiamo allora

$$\begin{aligned} |g(\chi)|^2 &= [(1 - a\lambda) + a\lambda e^{-ih\xi}]^2 \\ &= (1 - 2a\lambda \sin^2(\chi/2))^2 + 4a^2\lambda^2 \sin^2(\chi/2) \cos^2(\chi/2) \\ &= 1 - 4a\lambda(1 - a\lambda) \sin^2(\chi/2). \end{aligned}$$

Vediamo quindi che $|g(\chi)|$ è limitato da 1 se $0 \leq a\lambda \leq 1$, ritrovando quindi il risultato ottenuto con l'analisi diretta. Tuttavia qui abbiamo un elemento supplementare. Infatti se $a\lambda$ non è tra 0 e 1, ossia $|g(\chi)|$ maggiore di 1, faremo vedere che lo schema è instabile, ossia la necessità della condizione.

Daremo adesso la condizione per la stabilità in L^2 per le approssimazioni numeriche dell'equazione (1.1), nonostante questa analisi sia valida più in generale per problemi iperbolici e parabolici che siano ben posti in questa topologia. Seguendo quanto illustrato nell'esempio precedente possiamo affermare che per ogni schema alle differenze ad un passo e su $2l + 1$ punti della forma (1.26), possiamo scrivere la relazione

$$\hat{v}^{n+1}(\xi) = g(h\xi)\hat{v}^n(\xi), \quad (1.48)$$

con $g(h\xi) := \sum_{k=-l}^l c_l e^{ikh\xi}$. La funzione g è detta fattore di amplificazione e contiene tutte le informazioni importanti sullo schema e, cosa più importante, risulta essere uno strumento abbastanza facile da analizzare.

Nell'esempio precedente abbiamo considerato la funzione di amplificazione g che dipendeva solamente da $\chi = h\xi$, ma in generale bisogna considerare una funzione che possa anche dipendere da h e k .

Teorema 1.12 *Uno schema alle differenze ad un passo per l'equazione (1.1) è stabile se e solo se esiste una costante K e alcuni valori fissati k_0 e h_0 tali che*

$$|g(\chi, h, k)| \leq 1 + Kk, \quad (1.49)$$

per tutti i valori di χ , $0 < k \leq k_0$, e $0 < h \leq h_0$. Se inoltre $g(\chi, h, k)$ è indipendente da h e k allora basta prendere $K = 0$ in (1.49).

Si osservi che si può dimostrare che la condizione di von Neumann è necessaria per i sistemi a coefficienti costanti, se applicata agli autovalori della matrice di amplificazione, e continua ad essere anche sufficiente per tutti gli schemi ad un passo per le equazioni scalari con una sola derivata nel tempo.

Dimostrazione. Dalla relazione di Parseval e dalla definizione di g si ha che

$$\|v^n\|_h^2 = \int_{-\pi/h}^{\pi/h} |g(\chi, h, k)|^{2n} |\hat{v}^0(\xi)|^2 d\xi.$$

Quindi se vale la (1.49), si avrà

$$\begin{aligned} \|v^n\|_h^2 &\leq \int_{-\pi/h}^{\pi/h} (1 + Kk)^{2n} |\hat{v}^0(\xi)|^2 d\xi \\ &= (1 + Kk)^{2n} \|v^0\|_h^2. \end{aligned}$$

Essendo $nk \leq T$ si ha che

$$(1 + Kk)^{2n} \leq (1 + Kk)^{2T/k} \leq e^{2KT},$$

da cui segue la (1.25) nella norma L^2 . Passiamo ora a dimostrare la necessità della condizione (1.49). Supponiamo allora che per qualsiasi costante N esista

una terna (χ_N, h_N, k_N) , e quindi per continuità un intorno $\Gamma_N = [\chi_1, \chi_2] \times (0, h_0] \times (0, k_0]$, tale che $|g(\chi, h, k)| \geq 1 + Nk$, per $(\chi, h, k) \in \Gamma_N$. Costruiamo allora una funzione $v_{m,0}$ tale che la sua serie di Fourier sia uguale a

$$\hat{v}_{m,0} = \begin{cases} 0, & \text{se } h\xi \notin [\chi_1, \chi_2], \\ \sqrt{h(\chi_2 - \chi_1)^{-1}}, & \text{se } h\xi \in [\chi_1, \chi_2], \end{cases}$$

Allora

$$\begin{aligned} \|v^n\|_h^2 &\geq \int_{-\pi/h}^{\pi/h} (1 + Nk)^{2n} |\hat{v}^0(\xi)|^2 d\xi \\ &= \int_{-\pi/h}^{\pi/h} (1 + Nk)^{2n} \frac{h}{\chi_2 - \chi_1} d\xi \\ &\geq \frac{1}{2} e^{2TN} \|v^0\|_h^2, \end{aligned}$$

se nk è vicino a T . Prendendo N arbitrariamente grande ne segue la tesi. Il fatto che nel caso in cui $g(\chi, h, k)$ è indipendente da h e k allora basta prendere $K = 0$ in (1.49), segue da un'applicazione degli sviluppi di Taylor. \square

Esempio 1.13 Per lo schema forward-forward (1.10) si verifica facilmente che

$$g(\chi) = 1 + a\lambda(1 - e^{i\chi}),$$

ossia

$$|g|^2 = 1 + 4a\lambda(1 + a\lambda) \sin^2(\chi/2).$$

Se λ è una costante positiva fissata, possiamo usare la condizione (1.49) con $K = 0$ e verificare che se $a > 0$ $|g| > 1$ per $\chi \neq 0$, ossia lo schema è instabile. Altrimenti se $a < 0$ la condizione è ovviamente verificata se e solo se $-1 \leq a\lambda \leq 0$.

Un modo semplice di calcolare la funzione g è quella di sostituire nello schema al posto di $v_{m,n}$ l'espressione $g^n e^{im\chi}$. Questo non vuol dire che stiamo analizzando solo la stabilità delle onde piane nello schema, ma che con questa scorciatoia è possibile ottenere il valore di g che permette l'analisi rigorosa che abbiamo illustrato qui sopra.

Esempio 1.14 Analizziamo lo schema forward-central (1.12)

$$\frac{v_{m,n+1} - v_{m,n}}{k} + a \frac{v_{m+1,n} - v_{m-1,n}}{2h} = 0.$$

Sostituendo $v_{m,n}$ con $g^n e^{im\chi}$ si ottiene

$$\begin{aligned} &\frac{g^{n+1} e^{im\chi} - g^n e^{im\chi}}{k} + a \frac{g^n e^{i(m+1)\chi} - g^n e^{i(m-1)\chi}}{2h} \\ &= g^n e^{im\chi} \left(\frac{g-1}{k} + a \frac{e^{i\chi} - e^{-i\chi}}{2h} \right) = 0 \end{aligned}$$

che dà come fattore di amplificazione

$$g = 1 - ia\lambda \sin \chi.$$

Se λ è una costante allora g è indipendente da h e k e

$$|g(\chi)|^2 = 1 + a^2 \lambda^2 \sin^2 \chi,$$

che implica quindi l'instabilità del metodo.

Esempio 1.15 Analizziamo ora brevemente lo schema di Lax–Friedrichs (1.14)

$$\frac{v_{m,n+1} - \frac{1}{2}(v_{m+1,n} + v_{m-1,n})}{k} + a \frac{v_{m+1,n} - v_{m-1,n}}{2h} = 0.$$

Si vede facilmente che il fattore di amplificazione per questo schema verifica

$$|g(\chi)|^2 = \cos^2 \chi + a^2 \lambda^2 \sin^2 \chi,$$

che raggiunge il suo massimo $a^2 \lambda^2$ per $\chi = \frac{\pi}{2}$. La condizione di stabilità è quindi data, come già avevamo visto, dalla stessa condizione di CFL $|a\lambda| \leq 1$. Per questo nell'esempio 1.2 al valore $\lambda = 1.6$ corrispondeva un'approssimazione non convergente.

Terminiamo infine osservando che l'analisi di von Neumann può essere adattata opportunamente, anche se non direttamente, allo studio delle equazioni a coefficienti variabili, mediante il metodo del “congelamento” dei coefficienti. Per esempio per l'equazione a coefficienti variabili

$$u_t + a(x, t)u_x = 0,$$

varrà ancora la condizione di stabilità $|a(x_m, t_n)\lambda| \leq 1$ per tutti i punti della griglia (x_m, t_n) nel dominio computazionale.

1.8 Ordine di accuratezza e schema di Lax–Wendroff

Fino a questo momento abbiamo considerato gli schemi solamente in base alla loro convergenza o meno, ossia in ultima analisi in base alla loro consistenza e stabilità. Tuttavia dal punto di vista numerico è altrettanto importante stabilire la velocità di convergenza degli schemi. Questo sarà fatto come conseguenza della nozione di accuratezza dello schema che definiamo qui di seguito.

Definizione 1.7 *Uno schema consistente con l'equazione (1.1) si dice accurato di ordine q in spazio e di ordine p in tempo, in breve accurato di ordine (q, p) , se*

$$|T(x, t)| = O(k^p) + O(h^q), \text{ se } h, k \rightarrow 0,$$

per ogni (x, t) nel dominio di interesse. Nel caso in cui $k = K(h)$ diciamo che lo schema è di ordine r se

$$|T(x, t)| = O(h^r), \text{ se } h, \rightarrow 0.$$

In pratica l'accuratezza è una nozione più forte di consistenza, che corrisponde ad un errore di troncatura solamente dell'ordine di $o(1)$.

Cominciamo adesso a verificare l'ordine di accuratezza di alcuni degli schemi conosciuti. Lo schema di Lax-Friedrichs ci dà uno sviluppo del tipo

$$T = \frac{1}{2}ku_{tt} - \frac{1}{2}k^{-1}h^2u_{xx} + \frac{1}{6}ah^2u_{xxx} + O(h^4 + k^{-1}h^4 + k^2).$$

Fissando $k = \lambda h$ per una costante λ fissata, vediamo che lo schema è di ordine 1. Analogamente abbiamo che lo schema upwind con $a < 0$ ci dà

$$T = \frac{1}{2}ku_{tt} + \frac{1}{2}ahu_{xx} + O(k^2) + O(h^2),$$

che corrisponde ad uno schema di ordine (1,1) o, fissando $k = \lambda h$ per una costante λ fissata, ancora uno schema di ordine 1. Più in generale vale la seguente proprietà.

Proposizione 1.16 *Uno schema lineare monotono nella forma (1.27) è al più del primo ordine, a meno che non sia lo schema banale $v_{m,n+1} = v_{m,n}$ con $\lambda|a| = 1$.*

Dimostrazione. Dallo sviluppo (1.32), e utilizzando che $u_{tt} = a^2u_{xx}$, si ha che lo schema è del secondo ordine se e solo se

$$q = (a\lambda)^2. \quad (1.50)$$

D'altra parte lo schema è monotono se e solo se vale la condizione (1.38) $\lambda|a| \leq q \leq 1$. Quindi, essendo $(a\lambda)^2 < \lambda|a| < 1$, si ha la conclusione. \square

Si prenda la forma generale (1.34) di uno schema a tre punti consistente per l'equazione (1.1). Riprendendo la dimostrazione della Proposizione 1.16 si vede che uno schema di questo tipo può essere di ordine (2,2), oppure per $\lambda = k/h$ fissato, di ordine 2, se e solo vale la condizione (1.50). Ossia esiste un solo schema lineare a tre punti consistente che sia di ordine 2. Questo schema, chiamato schema di Lax-Wendroff, è stato proposto per la prima volta nel 1960 in [LW], e si scrive quindi

$$v_{m,n+1} = v_{m,n} - \frac{a\lambda}{2}(v_{m+1,n} - v_{m-1,n}) + \frac{a^2\lambda^2}{2}(v_{m+1,n} - 2v_{m,n} + v_{m-1,n}). \quad (1.51)$$

La sua viscosità numerica è la più bassa possibile per lasciare l'equazione equivalente (1.35) ancora parabolica. Chiaramente, poichè la condizione (1.38) non è verificata, lo schema non è monotono.

Facciamo ora l'analisi di stabilità dello schema di Lax-Wendroff. Il suo fattore di amplificazione si calcola facilmente ed è dato da

$$\begin{aligned} g(\chi) &= 1 - \frac{a\lambda}{2}(e^{i\chi} - e^{-i\chi}) + \frac{a^2\lambda^2}{2}(e^{i\chi} - 2 + e^{-i\chi}) \\ &= 1 - ia\lambda \sin \chi - 2a^2\lambda^2 \sin^2 \frac{\chi}{2}. \end{aligned}$$

Per cui

$$|g|^2 = 1 - 4a^2\lambda^2(1 - a^2\lambda^2)\sin^4\frac{\chi}{2}.$$

Abbiamo quindi che $g(\chi) \leq 1$ se e soltanto se $|a\lambda| \leq 1$. Lo schema risulta quindi L^2 -stabile. In generale vale anzi la seguente caratterizzazione degli schemi a tre punti consistenti con l'equazione (1.1).

Proposizione 1.17 *Uno schema esplicito a tre punti consistente con l'equazione (1.1), nella forma viscosa (1.34), è L^2 -stabile se e solo se*

$$(\lambda a)^2 \leq q \leq 1. \quad (1.52)$$

Vediamo quindi che la stabilità L^2 è meno restrittiva della stabilità nella norma del massimo e permette quindi di trovare almeno uno schema del secondo ordine stabile. Vediamo come si comporta lo schema di Lax–Wendroff nei confronti per esempio dello schema upwind. Si prenda l'equazione (1.1) con il dato iniziale uguale alla funzione caratteristica dell'intervallo $(0.2, 0.4)$. La figura 1.3 mostra il risultato per lo schema upwind, mentre la figura 1.4 per lo schema di Lax–Wendroff. Nello schema monotono si nota uno smussamento degli spigoli e una piccola riduzione dell'altezza, mentre lo schema di Lax–Wendroff mantiene l'altezza ai bordi molto meglio, producendo tuttavia delle piccole oscillazioni. Riducendo h si vede un sostanziale miglioramento, anche se le oscillazioni persistono. Questo è dovuto al fatto che la nostra analisi dell'errore è valida solo per soluzioni abbastanza regolari. Si veda infatti come si comporta lo schema di Lax–Wendroff con un dato molto più regolare (di fatto C^∞) nella figura 1.5.

In pratica, l'instabilità, essendo legata alla crescita delle frequenze elevate, appare prima per dati irregolari, vedi le figure 1.4 e 1.5. Si osservi anche che in generale questa instabilità è di fatto un fenomeno locale. Per esempio nella figura 1.2 si vede che l'instabilità nasce essenzialmente nel punto di discontinuità della derivata e solo dopo si propaga a tutta la soluzione. Questo permette di distinguere tra gli effetti di un errore di programmazione, che spesso sono globali, e l'instabilità numerica dello schema.

Acceniamo infine alla relazione che intercorre tra uno schema di ordine r e la sua accuratezza. Vale il seguente risultato, per la cui dimostrazione e varianti rimandiamo a [St], Cap. 10. In particolare si può dimostrare un analogo risultato nel caso parabolico e più generalmente per tutti i problemi di Cauchy ben posti in L^2 .

Teorema 1.18 *Sia dato uno schema stabile per l'equazione (1.1) che sia accurato di ordine $r \geq 1$. Allora esiste un indice intero non negativo $s \geq r$ tale che se $u_0 \in H^s$, per ogni fissato $T > 0$ esiste una costante C_T tale che*

$$\|u^n - v^n\|_{2,h} \leq C_T h^r \|u_0\|_{H^s},$$

per ogni n tale che $0 \leq kn \leq T$.

In generale per i metodi iperbolici si trova $s = r + 1$ e per quelli parabolici $s = r + 2$.

1.9 Cenni sull'equazione delle onde

Discutiamo brevemente come estendere le idee viste nelle sezioni precedenti al caso dell'equazione delle onde in una variabile spaziale.

Si consideri l'equazione

$$u_{tt} - a^2 u_{xx} = 0, \quad (1.53)$$

dove a è una costante positiva. Il problema di Cauchy sulla retta reale risulta allora determinato quando si assegna un valore iniziale al tempo $t = 0$ per la u e la sua derivata prima temporale

$$u(x, 0) = u_0(x), \quad u_t(x, 0) = u_1(x). \quad (1.54)$$

Come è noto la soluzione di questo problema è data dalla formula di D'Alembert

$$u(x, t) = \frac{1}{2} [u_0(x - at) + u_0(x + at)] + \frac{1}{2a} \int_{x-at}^{x+at} u_1(y) dy.$$

Questa formula mostra come la soluzione sia naturalmente decomposta in due onde che si muovono a velocità $-a$ e a . Si può dimostrare, per esempio utilizzando la soluzione esplicita, oppure con l'analisi di Fourier, che il problema di Cauchy (1.53)–(1.54) è ben posto nelle norme di tipo L^2 . In particolare è possibile stabilire delle stime dell'energia del tipo:

$$\int |u_t|^2 + |u_x|^2 dx \leq \int |u_1|^2 + |(u_0)_x|^2 dx,$$

che assicurano la stabilità negli spazi di Sobolev.

Le definizioni di convergenza, consistenza e ordine di accuratezza rimangono invariate rispetto a quelle date nelle sezioni precedenti. Per quanto riguarda la stabilità bisogna però introdurre qualche modifica. Infatti una soluzione particolare per $u_0 = 0$, $u_1 = 1$ è data da $u(x, t) = t$, per cui bisognerà permettere una crescita nel tempo delle norme.

Definizione 1.8 *Uno schema alle differenze finite $P_{h,k}v = 0$ che approssima l'equazione (1.53) è stabile nella norma $\|\cdot\|_{\cdot,h}$ se esiste un intero L tale che per ogni $T > 0$ esiste una costante C_T per cui*

$$\|v^n\|_{\cdot,h} \leq C_T (1 + n^2)^{\frac{1}{2}} \sum_{l=0}^L \|v^l\|_{\cdot,h}, \quad (1.55)$$

per $0 \leq nk \leq T$.

L'analisi di von Neumann rimarrà invariata, ma bisognerà richiedere che entrambi i fattori di amplificazione g_ν verifichino la condizione (1.49). Con questa definizione di stabilità è anche possibile estendere il teorema di equivalenza di Lax–Richtmyer all'equazione delle onde.

Lo schema più semplice da immaginare è lo schema in cui le derivate seconde spaziali e temporali sono entrambe approssimate dalle differenze centrate seconde:

$$\frac{v_{m,n+1} - 2v_{m,n} + v_{m,n-1}}{k^2} = a^2 \frac{v_{m+1,n} - 2v_{m,n} + v_{m-1,n}}{h^2}. \quad (1.56)$$

Essendo $a > 0$, vogliamo far vedere che questo schema, che è naturalmente del secondo ordine in x e t , risulta anche stabile per $a\lambda \leq 1$.

Applichiamo direttamente l'analisi di von Neumann e troviamo che il fattore di amplificazione dello schema risolve l'equazione

$$g - 2 + g^{-1} = -4a^2\lambda^2 \sin^2\left(\frac{1}{2}\chi\right).$$

Ne segue che le radici di questa equazione sono

$$g_{\pm} = \left(\sqrt{1 - a^2\lambda^2 \sin^2\left(\frac{1}{2}\chi\right)} \pm ia\lambda \sin\left(\frac{1}{2}\chi\right) \right)^2,$$

e sono minori o uguali ad 1 se e solo se $a\lambda \leq 1$. Si osservi che se $\chi = 0$, oppure $a\lambda = 1$ e $\chi = \pi$, allora le due radici coincidono. In questo caso la soluzione dello schema, nelle variabili di Fourier è data da

$$\hat{v}^n = A_+(\xi)g_+^n + nB_+(\xi)g_-^n.$$

La crescita lineare in n essendo permessa nella nostra definizione di stabilità, troviamo che lo schema è stabile anche quando le due radici coincidono.

Concludiamo questa breve discussione con l'analisi dell'inizializzazione dello schema. Come ogni schema multistep, anche questo ha bisogno di una corretta procedura per eseguire il primo passo. Il modo forse più semplice di farlo è quello di utilizzare lo sviluppo di Taylor

$$u(x, k) = u(x, 0) + ku_t(x, 0) + \frac{1}{2}k^2 u_{tt}(x, 0) + O(k^3).$$

I valori di $u(x, 0)$ e $u_t(x, 0)$ sono dati dalle condizioni iniziali e il valore di $u_{tt}(x, 0)$ può essere espresso utilizzando l'equazione, ossia ponendolo uguale a $a^2 u_{xx}(x, 0)$. Possiamo quindi ottenere un'espressione per $v_{m,1}$ che è dello stesso grado di accuratezza del resto dello schema, in questo caso

$$v_{m,1} = v_{m,0} + k(u_t)_m + \frac{1}{2}k^2 a^2 \delta^2 v_{m,0}.$$

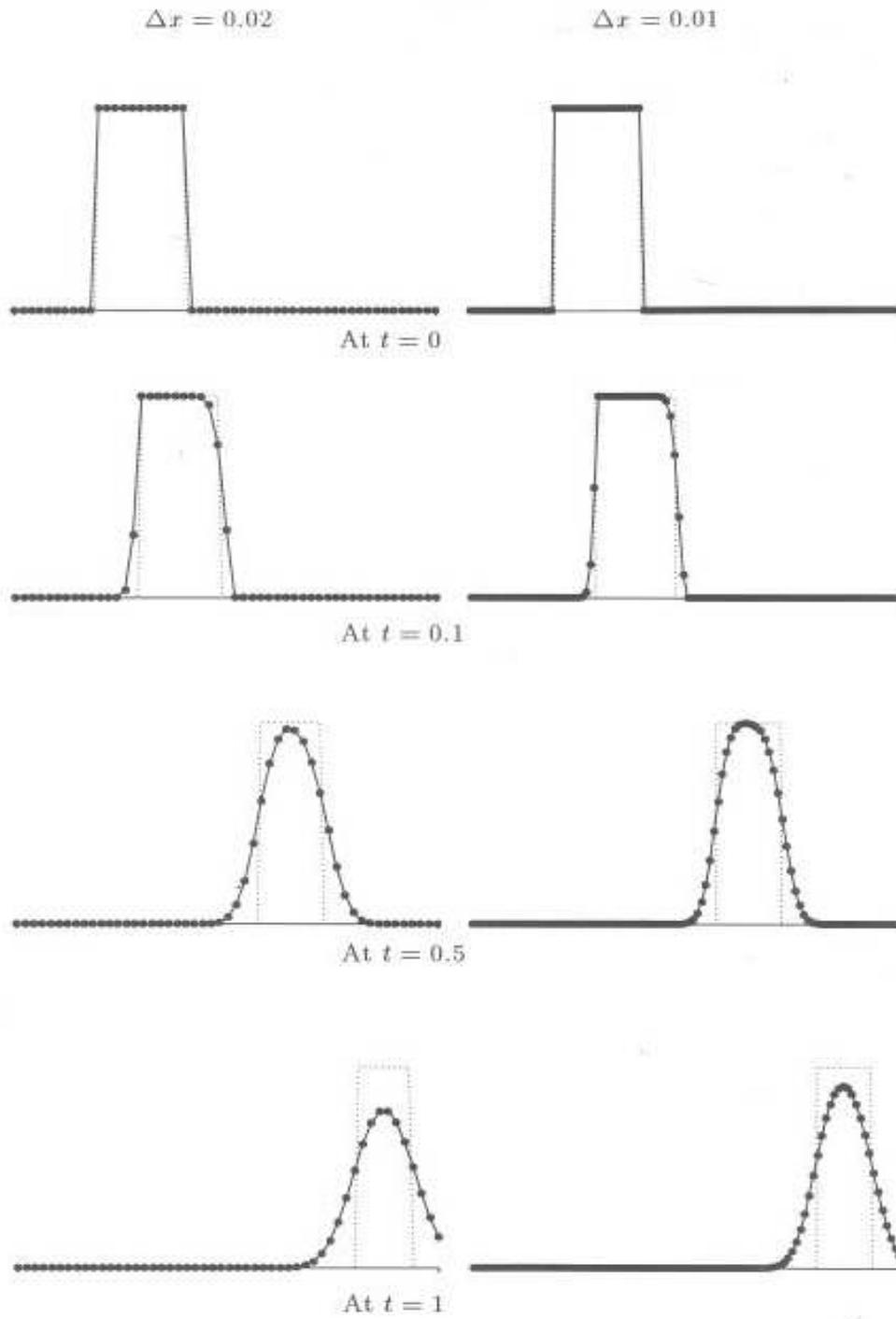


Figura 1.3: Trasporto lineare con il metodo upwind

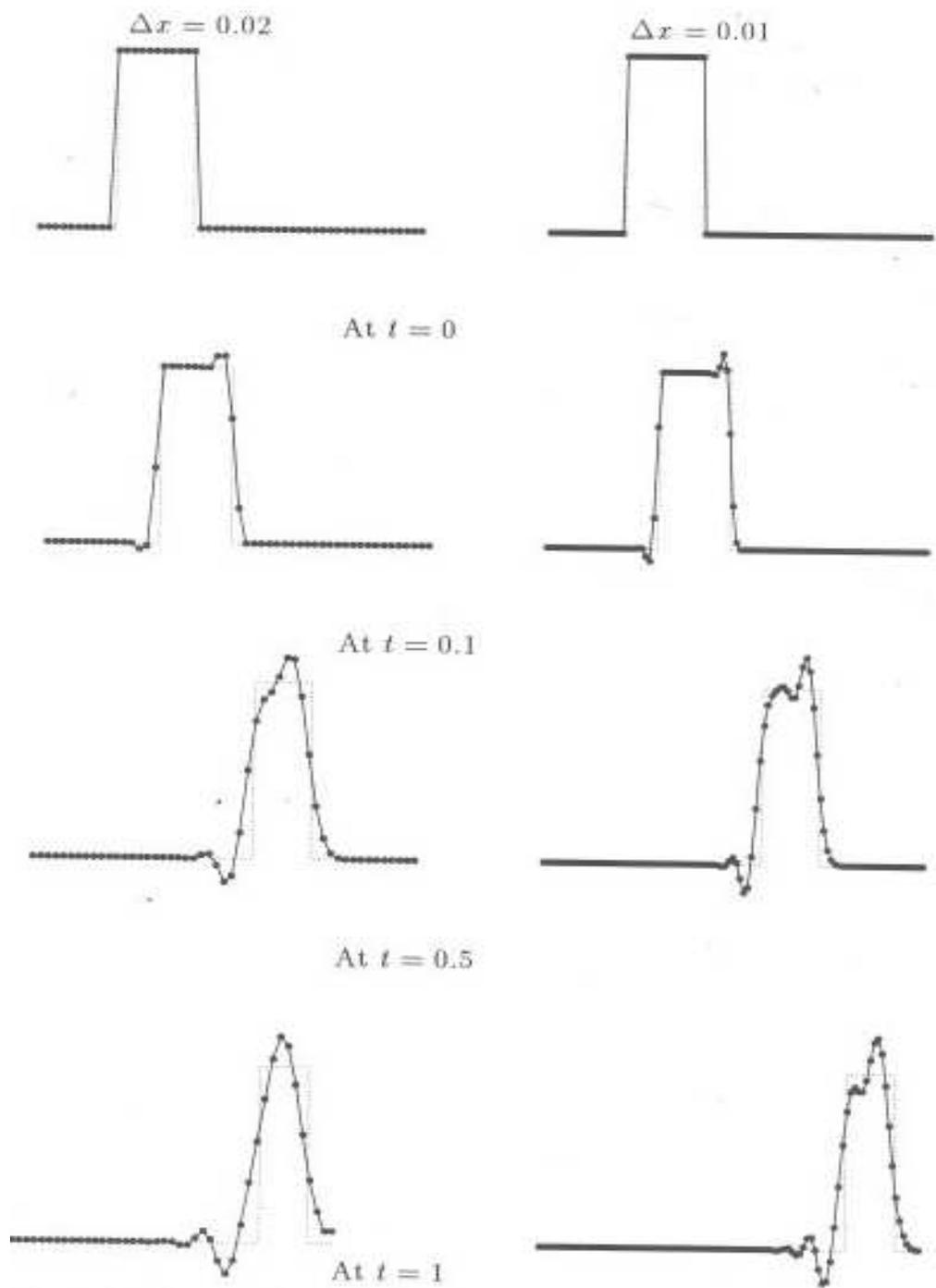


Figura 1.4: Trasporto lineare con il metodo di Lax-Wendroff

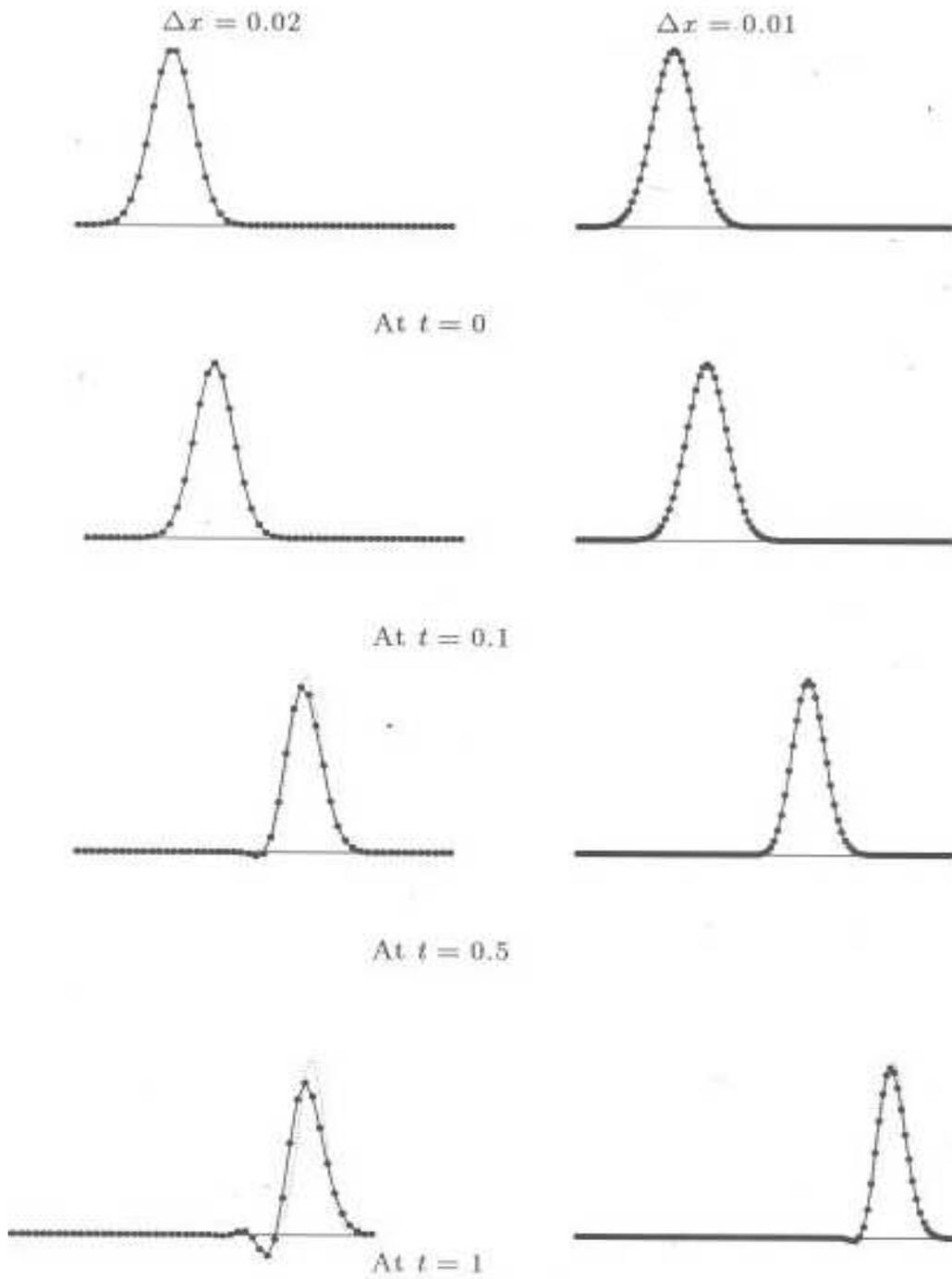


Figura 1.5: Trasporto lineare (dato regolare) con il metodo di Lax-Wendroff

Capitolo 2

Equazioni paraboliche

Tratteremo adesso la risoluzione numerica del più semplice tipo di equazione parabolica, l'equazione del calore in una dimensione spaziale

$$u_t = bu_{xx}, \quad (2.1)$$

con $b > 0$.

Per questo tipo di problemi analizzeremo prima di tutto i metodi espliciti per poi passare ad alcuni metodi più elaborati che serviranno ad ovviare alle condizioni eccessivamente restrittive richieste dai metodi espliciti per ottenere la stabilità

2.1 Il problema di Cauchy-Dirichlet per l'equazione del calore

Innanzitutto consideriamo la risoluzione del problema di Cauchy-Dirichlet sul dominio $[0, 1] \times [0, T)$ con condizione iniziale

$$u(x, 0) = u_0(x), \quad (2.2)$$

e condizioni al bordo omogenee

$$u(0, t) = u(1, t) = 0. \quad (2.3)$$

Per questo semplice problema è possibile adoperare il metodo di separazione delle variabili per trovare una soluzione esatta. Cerchiamo ossia delle soluzioni $u(x, t)$ della forma speciale $f(x)g(t)$; sostituendo in (2.1) troviamo l'equazione differenziale ordinaria

$$\frac{g'}{g} = \frac{f''}{f} = -k^2.$$

Abbiamo scartato valori positivi della costante a secondo membro che avrebbero permesso soluzioni illimitate. Questo dà in ogni modo la soluzione particolare

$$u(x, t) = e^{-m\pi^2 t} \sin m\pi x.$$

Questa soluzione soddisfa le condizioni al bordo (2.3) per tutti i valori (interi) di m . Poichè una combinazione lineare di tali soluzioni ancora verifica l'equazione, possiamo cercare la soluzione generale sotto la forma

$$u(x, t) = \sum_{m=1}^{\infty} a_m e^{-m^2 \pi^2 t} \sin m\pi x.$$

Dobbiamo ora scegliere i coefficienti a_m per soddisfare la condizione iniziale (2.2). Per $t = 0$ abbiamo

$$\sum_{m=1}^{\infty} a_m \sin m\pi x = u_0(x).$$

Questo dimostra che i coefficienti a_m sono esattamente i coefficienti della serie di Fourier della funzione u_0 e sono dati da

$$a_m = 2 \int_0^1 u_0(x) \sin m\pi x dx.$$

In effetti la formula trovata ci dà la soluzione esatta, almeno per dati iniziali per cui la serie di Fourier converge, ma forse andrebbe vista più che altro come una specie di approssimazione numerica. In effetti i coefficienti a_m andranno calcolati in generale con qualche formula di quadratura e si potrà tenere conto solo di un numero finito di addendi nella serie. Questo non toglie che almeno per l'equazione (2.1) questo metodo potrebbe risultare efficace, ma diventerebbe non praticabile in situazioni più complicate come i problemi a coefficienti variabili, in più dimensioni spaziali e non lineari, dove la soluzione esplicita non risulta essere disponibile, almeno con mezzi elementari.

2.2 Uno schema esplicito per l'equazione del calore

Consideriamo una discretizzazione standard del dominio $[0, 1] \times [0, T]$, adottando tutte le notazioni già utilizzate per le equazioni iperboliche. Lo schema più semplice consiste nell'approssimare la derivata temporale di u con una differenza forward

$$u_t(x_m, t_n) \approx \frac{\Delta_{+t} u}{\Delta t} := \frac{u(mh, (n+1)k) - u(mh, nk)}{k},$$

e la derivata seconda spaziale con una differenza centrata seconda

$$u_{xx}(x_m, t_n) \approx \frac{\delta_x^2 u}{(\Delta x)^2} := \frac{u(mh, (n+1)k) - 2u(mh, nk) + u(mh, (n-1)k)}{h^2}.$$

Otteniamo quindi uno schema forward-central per l'equazione (2.1)

$$\frac{v_{m,n+1} - v_{m,n}}{k} - b \frac{v_{m+1,n} - 2v_{m,n} + v_{m-1,n}}{h^2} = 0, \quad (2.4)$$

che si scrive anche

$$v_{m,n+1} = (1 - 2b\mu)v_{m,n} + b\mu(v_{m+1,n} + v_{m-1,n}), \quad (2.5)$$

avendo posto $\mu = kh^{-2}$. Si noti come il parametro μ giochi qui un ruolo analogo al parametro λ per le equazioni iperboliche. Utilizzando lo sviluppo di Taylor di una soluzione regolare u e inserendolo nello schema (2.4) possiamo facilmente scrivere l'errore di troncatura

$$Tu = (u_t - bu_{xx}) + \left(\frac{1}{2}ku_{tt} - \frac{b}{12}h^2u_{xxxx}\right) + \dots \quad (2.6)$$

che mostra come lo schema sia di ordine 1 nel tempo e 2 nello spazio.

Analogamente al caso iperbolico possiamo effettuare l'analisi di von Neumann, sostituendo al valore $v_{m,n}$ l'espressione $g^n e^{imx}$ nello schema, sempre ponendo $\chi = h\xi$. Otteniamo quindi

$$\frac{g-1}{k} = b \frac{e^{i\chi} - 2 + e^{-i\chi}}{h^2},$$

che implica

$$g = 1 - 4b\mu \sin^2 \frac{\chi}{2}.$$

Per avere stabilità dobbiamo avere quindi $|g| \leq 1$ ossia $4b\mu \sin^2 \frac{\chi}{2} \leq 2$. Questo è vero per ogni χ se e solo se

$$b\mu \leq \frac{1}{2}. \quad (2.7)$$

Questa condizione è abbastanza restrittiva in pratica perchè se vogliamo dimezzare il passo spaziale dobbiamo prendere un quarto del passo temporale. Questo è il motivo per cui nel seguito altri schemi verranno di proposti.

Vogliamo però prima far vedere come sia facile per questo schema mostrarne la convergenza sotto la condizione (2.7). Procediamo come per il caso iperbolico.

Teorema 2.1 *Sia $v_{m,n}$ l'approssimazione numerica data dallo schema (2.5) e u la soluzione del problema (2.1)–(2.3). Se la norma del massimo della derivata quarta spaziale di u , u_{xxxx} , risulta limitata uniformemente sull'intervallo $[0, 1] \times [0, T_0]$, e $k = \mu h^2$, per una costante positiva $\mu \leq \frac{1}{2b}$, allora la successione $v_{m,n}$ converge alla funzione u nella norma (1.15), quando $h \rightarrow 0$.*

Dimostrazione. Sia $T_{m,n}$ l'errore di troncatura in (x_n, t_m) e si denoti con e l'errore $v - u$ nell'approssimazione,

$$e_{m,n} = v_{m,n} - u_{m,n}. \quad (2.8)$$

Se inseriamo la quantità $e_{m,n}$ nello schema (2.4), troviamo

$$e_{m,n+1} = (1 - 2b\mu)e_{m,n} + b\mu(e_{m+1,n} + e_{m-1,n}) - kT_{m,n}. \quad (2.9)$$

Osserviamo allora che se vale la condizione (2.7), i coefficienti dei due primi termini a destra sono non negativi e sommati danno 1. Introduciamo quindi la norma del massimo dell'errore

$$E^n := \max\{|e_{m,n}|, m \in J_\Omega\}.$$

A questo punto utilizzando la non negatività dei coefficienti di $e_{m,n}$ in (2.9) si ottiene facilmente

$$E^{n+1} \leq E^n + k \max\{|T_{m,n}|, m \in J_\Omega\}, \quad (2.10)$$

ossia, essendo $E^0 = 0$,

$$E^n \leq T_0 \max\{|T_{m,n}|, m \in J_\Omega, nk \leq T_0\}. \quad (2.11)$$

Utilizzando la stima (2.6) e il fatto che $u_{tt} = b^2 u_{xxxx}$ otteniamo allora

$$E^n \leq \frac{1}{2} kbT_0 \left(b + \frac{1}{6\mu}\right) M_{xxxx},$$

dove M_{xxxx} è la norma del massimo per la derivata u_{xxxx} . Abbiamo quindi che l'errore tende a zero, ovvero l'approssimazione numerica converge, al tendere di $h = \sqrt{\mu^{-1}k}$ a zero, con μ costante, se la soluzione limite ha la derivate quarta spaziale uniformemente limitata su tutto il dominio. \square

L'analisi di von Neumann ci dice che, almeno per ottenere la stabilità in L^2 , la condizione (2.7) è necessaria e sufficiente per la stabilità, mentre il Teorema 2.1 ci dice che è almeno sufficiente per la stabilità in L^∞ . Anche in pratica possiamo difficilmente evitarla come si vede dalla figura 2.1 ottenuta per $h = 0.05$ e il dato iniziale

$$u_0 = \begin{cases} 2x, & \text{per } 0 \leq x \leq \frac{1}{2}, \\ 2 - 2x, & \text{per } \frac{1}{2} \leq x \leq 1, \end{cases} \quad (2.12)$$

Vediamo che un piccolo cambiamento nel passo temporale produce in pratica un comportamento molto diverso della soluzione approssimata.

Per ovviare a questo inconveniente il modo più semplice è quello di utilizzare i metodi impliciti.

2.3 Un metodo implicito

Se prendiamo lo schema (2.4), sostituiamo la derivata forward con una derivata backward, lasciando invariata l'approssimazione spaziale, e riarrangiamo gli indici otteniamo lo schema implicito

$$\frac{v_{m,n+1} - v_{m,n}}{k} - b \frac{v_{m+1,n+1} - 2v_{m,n+1} + v_{m-1,n+1}}{h^2} = 0. \quad (2.13)$$

Questo schema dipende da tre punti al livello t_{n+1} e un punto solo al livello t_n ed è quindi meno agevole da implementare del corrispondente schema esplicito. Infatti conoscendo il valore di v^n non è immediato il calcolo di v^{n+1} . Tuttavia

lo schema è ugualmente di ordine (1,2) e dall'analisi di von Neumann risulta essere incondizionatamente stabile poichè

$$\frac{g-1}{k} = bg \frac{e^{i\chi} - 2 + e^{-i\chi}}{h^2},$$

che implica

$$g = \frac{1}{1 + 4b\mu \sin^2 \frac{\chi}{2}},$$

che è sempre compreso tra 0 e 1 per ogni valore di μ . Questo permette di prendere valori abbastanza grandi del passo di discretizzazione temporale k , sempre compatibilmente con il fatto che l'errore di troncatura deve rimanere piccolo. Bisognerà però accertarsi che la somma, su un numero inferiori di passi, dei lavori supplementari necessari a calcolare la soluzione approssimante per ogni passo di tempo sia inferiore al lavoro fatto con lo schema esplicito.

Scriviamo adesso il sistema nella forma

$$(1 + 2b\mu)v_{m,n+1} - b\mu(v_{m+1,n+1} + v_{m-1,n+1}) = v_{m,n}. \quad (2.14)$$

Supponendo che m vari tra $1, 2, \dots, M-1$, abbiamo allora un sistema lineare tridiagonale di $(M-1)$ equazioni in $(M-1)$ incognite, ossia l'equazione numero m contiene solo incognite con i numeri $m-1, m$ e $m+1$. Per risolvere sistemi di questa forma in modo efficiente risulta conveniente usare l'algoritmo di Thomas, che illustriamo brevemente qui di seguito. Riscriviamo il sistema lineare nella forma più generale

$$-a_m v_{m-1} + b_m v_m - c_m v_{m+1} = d_m, \quad m = 1, 2, \dots, M-1, \quad (2.15)$$

con

$$v_0 = v_M = 0. \quad (2.16)$$

L'algoritmo di Thomas permette di ridurre il sistema ricorsivamente ad uno triangolare superiore. Si procede per induzione supponendo che i primi k termini siano già scritti nella forma

$$v_m - e_m v_{m+1} = f_m, \quad m = 1, 2, \dots, k. \quad (2.17)$$

Il termine successivo sia allora nella forma originaria

$$-a_{k+1} v_k + b_{k+1} v_{k+1} - c_{k+1} v_{k+2} = d_{k+1}. \quad (2.18)$$

Possiamo allora eliminare v_k da questa equazione e dall'ultima delle (2.17) e otteniamo

$$v_{k+1} - e_{k+1} v_{k+2} = f_{k+1},$$

con

$$e_{k+1} = \frac{c_{k+1}}{b_{k+1} - a_{k+1} e_k}, \quad f_{k+1} = \frac{d_{k+1} + a_{k+1} f_k}{b_{k+1} - a_{k+1} e_k}.$$

Per induzione possiamo calcolare tutti i coefficienti partendo da $e_0 = f_0 = 0$ e ponendo

$$e_m = \frac{c_m}{b_m - a_m e_{m-1}}, \quad f_m = \frac{d_m + a_m f_{m-1}}{b_m - a_m e_{m-1}}, \quad m = 1, 2, \dots, M-1. \quad (2.19)$$

A questo punto calcoliamo i valori v_m partendo da $v_M = 0$ e tornando indietro usando (2.17) e (2.19).

Si osservi che in generale questo meccanismo di induzione può risultare instabile, ma si dimostra che questo non succede se $|e_m| < 1$. Una condizione sufficiente perchè ciò avvenga è data dalla condizione

$$a_m, b_m, c_m > 0, \quad b_m > a_m + c_m,$$

che è chiaramente soddisfatta per lo schema (2.14). Per risolverlo ad ogni nodo bastano 3 addizioni+3 moltiplicazioni+2 divisioni=8 operazioni, da confrontare alle 2 addizioni+2 moltiplicazioni=4 operazioni per lo schema esplicito. Il metodo implicito risulta quindi più vantaggioso se il numero di passi temporali è ridotto a meno della metà di quelli del metodo esplicito.

2.4 I θ -metodi e lo schema di Crank–Nicolson

Una generalizzazione naturale dei metodi precedenti, gli schemi (2.4) e (2.13), è data da una media pesata dei due che possiamo sinteticamente scrivere

$$v_{m,n+1} - v_{m,n} = b\mu[\theta\delta_x^2 u_{m,n+1} + (1-\theta)\delta_x^2 u_{m,n}], \quad (2.20)$$

con $0 \leq \theta \leq 1$. Si noti che per $\theta = 0$ si ritrova lo schema esplicito, mentre per $\theta = 1$ lo schema totalmente implicito. Altrimenti per $\theta \neq 0$, abbiamo ancora un sistema tridiagonale, che si appoggia genericamente su sei punti totali, tre al tempo t_n e tre al tempo t_{n+1} , da risolvere per gli $u_{m,n}$.

Dall'analisi di von Neumann otteniamo

$$g = \frac{1 - 4b(1-\theta)\mu \sin^2 \frac{\chi}{2}}{1 + 4b\theta\mu \sin^2 \frac{\chi}{2}}.$$

Essendo $\mu > 0$, abbiamo $g \leq 1$. La sola possibilità affinché si creino instabilità è data dal caso $g < -1$, ossia se

$$4b(1-2\theta)\mu \sin^2 \frac{\chi}{2} > 2.$$

In conclusione per $0 \leq \theta < \frac{1}{2}$, il metodo sarà stabile se e solo se

$$\mu \leq \frac{1}{2b}(1-2\theta)^{-1},$$

altrimenti, per $\frac{1}{2} \leq \theta \leq 1$, il metodo risulterà stabile per qualsiasi μ .

Si vede dunque che diventa decisivo l'ordine di accuratezza del metodo per decidere se risulta essere o meno vantaggioso in pratica.

Nel calcolo dell'errore di troncatura è essenziale la scelta del punto dove si deve effettuare l'espansione di Taylor. Per gli schemi espliciti è sempre conveniente effettuare tale sviluppo nel punto (x_m, t_n) , ma nel caso presente è preferibile usare, per ragioni di simmetria, il punto di mezzo dell'insieme dei punti in questione, ossia $(x_m, t_{n+\frac{1}{2}})$, e l'errore di troncatura verrà allora scritto come $T_{m, n+\frac{1}{2}}$. Abbiamo dunque

$$u_{m, n+1} = \left[u + \frac{1}{2}ku_t + \frac{1}{2}\left(\frac{1}{2}k\right)^2u_{tt} + \frac{1}{6}\left(\frac{1}{2}k\right)^3u_{ttt} + \dots \right]_{m, n+\frac{1}{2}}$$

e

$$u_{m, n} = \left[u - \frac{1}{2}ku_t + \frac{1}{2}\left(\frac{1}{2}k\right)^2u_{tt} - \frac{1}{6}\left(\frac{1}{2}k\right)^3u_{ttt} + \dots \right]_{m, n+\frac{1}{2}}.$$

Quindi sottraendo queste due espressioni si ottiene

$$\Delta_t u_{m, n+\frac{1}{2}} = u_{m, n+1} - u_{m, n} = \left[ku_t + \frac{1}{24}k^3u_{ttt} + \dots \right]_{m, n+\frac{1}{2}}.$$

Analogamente si può effettuare l'espansione delle differenze spaziali, ottenendo in conclusione la seguente espressione dell'errore di troncatura:

$$\begin{aligned} T_{m, n+\frac{1}{2}} &= (u_t - bu_{xx}) + \left[\left(\frac{1}{2} - \theta\right)kb u_{xxt} - \frac{b}{12}h^2u_{xxxx} \right] \\ &+ \left[\frac{1}{24}k^3u_{ttt} - \frac{b}{8}k^2u_{xxtt} \right] \\ &+ \left[\frac{b}{12}\left(\frac{1}{2} - \theta\right)kh^2u_{xxxxt} - \frac{2b}{6!}h^4u_{xxxxxx} \right]. \end{aligned}$$

Se u è una soluzione abbastanza regolare, allora lo schema è sempre consistente e accurato di ordine $(1, 2)$, tranne che nel caso $\theta = \frac{1}{2}$, in cui lo schema diventa di ordine $(2, 2)$. Il valore $\theta = \frac{1}{2}$ è speciale e lo schema corrispondente prende il nome di schema di *Crank-Nicolson*. Inoltre essendo implicito possiamo prendere anche k dello stesso ordine di h , ottenendo così uno schema del secondo ordine con una ragionevole scelta del passo temporale. Altre scelte possono risultare convenienti a seconda del tipo di problema affrontato, vedi ad esempio Richtmyer e Morton [RM], dove una quindicina di schemi diversi sono considerati.

Qui ci limitiamo a riportare alcuni grafici degli errori compiuti dai vari metodi con varie scelte dei parametri. Si riprenda il problema test che abbiamo considerato prima, ossia il problema (2.1)–(2.3) con il dato iniziale (2.12) e $b = 1$.

Nell'insieme dei grafici di sinistra della figura 2.2 si vedono i grafici dell'errore nella norma del massimo in funzione del numero dei passi di discretizzazione J , mentre in quello di destra in funzione dei tempi di CPU. Per gli schemi impliciti si considerano sia l'andamento a $\mu = \frac{k}{h^2}$ fissato (linea continua) che quello a $\lambda = \frac{k}{h}$ fissato (linea tratteggiata). In particolare abbiamo i seguenti valori:

- A: $\theta = 0, \mu = \frac{1}{2}$;
- B: $\theta = \frac{1}{2}, \mu = \frac{1}{2}, \lambda = \frac{1}{40}$;
- C: $\theta = \frac{1}{2}, \mu = 5, \lambda = \frac{1}{4}$;
- D: $\theta = 1, \mu = 5, \lambda = \frac{1}{4}$.

Il grafico *A* mostra l'andamento dell'errore come J^{-2} , come ci si aspetta, e lo stesso fa lo schema di Crank–Nicolson per $\mu = \frac{1}{2}$ o $\lambda = \frac{1}{40}$. Il miglior risultato è però ottenuto in *C* per $\lambda = \frac{1}{4}$, mentre per $\mu = 5$ la perdita di accuratezza dovuta ai grandi passi risulta prevalente. Infine in *D* il metodo totalmente implicito mostra, come ci si aspettava, la sua scarsa di accuratezza. Il grafico a destra, che mostra l'errore in funzione del tempo di CPU, dovrebbe dare una misura dell'efficienza dei vari metodi. Il metodo esplicito risulta molto più efficiente dell'implicito, ma Crank–Nicolson, specialmente per $\lambda = \frac{1}{4}$, mostra ancora la migliore prestazione.

2.5 Principio di massimo e convergenza

Una delle proprietà più importanti delle soluzioni dell'equazione del calore $u_t = bu_{xx}$ è il principio di massimo: sia u una soluzione (regolare) del problema (2.1)–(2.3) allora

$$\sup\{u(x, t); x \in [0, 1], t \in [0, T]\} \leq \max\{0, \sup\{u_0(x); x \in [0, 1]\}\}.$$

Un'analogia proprietà vale per l'inf e, di conseguenza per i moduli. Questa proprietà implica, in modo elementare almeno nel presente caso lineare, una proprietà di confronto: se u e v sono le soluzioni corrispondenti ai dati iniziali u_0 e v_0 , abbiamo che se $u_0 \leq v_0$ allora $u \leq v$. Questo segue facilmente dal fatto che $u - v$ è soluzione dello stesso problema e che se $u - v$ è minore di 0 per $t = 0$, allora lo è ancora per tutti i tempi positivi.

Chiaramente noi ci aspettiamo che questo tipo di proprietà sia conservato da una buona approssimazione numerica e anzi in generale saremmo portati a scartare uno schema che non li rispettasse. Abbiamo in questo senso il seguente risultato.

Teorema 2.2 *Se $v_{m,n}$ è la soluzione approssimata generata da un θ -metodo per cui $\mu(1 - \theta) \leq \frac{1}{2b}$, allora, posto*

$$v_{min} := \min\{v_{m,0} \ (m = 1, \dots, M - 1), 0\}$$

e

$$v_{max} := \max\{v_{m,0} \ (m = 1, \dots, M - 1), 0\},$$

avremo che

$$v_{min} \leq v_{m,n} \leq v_{max}. \tag{2.21}$$

Inoltre, se la discretizzazione del dato iniziale è consistente con il dato stesso e la soluzione u del problema (2.1)–(2.3) è abbastanza regolare perchè l'errore di troncamento per lo schema (2.20) tenda a zero, allora la successione $v_{m,n}$ tende a u quando k (e quindi h) tende a 0.

Dimostrazione. Scriviamo lo schema (2.20) nella forma

$$\begin{aligned} (1 + 2\theta b\mu)v_{m,n+1} &= \theta b\mu(v_{m-1,n+1} + v_{m+1,n+1}) \\ &+ (1 - \theta)b\mu(v_{m-1,n} + v_{m+1,n}) + [1 - 2(1 - \theta)b\mu]v_{m,n}. \end{aligned} \quad (2.22)$$

Sotto le ipotesi del teorema tutti i coefficienti a destra sono positivi e la loro somma è uguale a $(1 + 2\theta b\mu)$. Supponiamo ora che v assuma un massimo interno al nodo (x_m, t_{n+1}) , e sia V questo valore. Inoltre sia V^* il massimo dei cinque valori che compaiono a destra in (2.22). Poichè i coefficienti sono positivi abbiamo che $V \leq V^*$, ma essendo V il massimo abbiamo che $V = V^*$ e tutti i valori nei sei punti considerati assumo questo stesso valore. Ragionando allo stesso modo possiamo vedere che possiamo estendere questo argomento fino a toccare un punto del bordo o il valore $m = 0$, dimostrando così la disuguaglianza di destra in (2.21). Quella di sinistra si dimostra in modo analogo. Per dimostrare la convergenza si procede analogamente alla dimostrazione del Teorema 2.1 (che in effetti è solo un caso particolare del presente risultato).□

La condizione $\mu \leq \frac{1}{2b(1-\theta)}$ del precedente teorema è notevolmente più restrittiva della condizione per la stabilità (e quindi convergenza) in L^2 . Per esempio per lo schema di Crank–Nicolson, che è sempre L^2 –stabile, abbiamo che per soddisfare il principio del massimo e quindi avere la convergenza nella norma del massimo, si deve richiedere $b\mu \leq 1$. In effetti questa condizione è ottimale per un numero di nodi arbitrario, ma può essere rilassata se si prende h abbastanza piccolo. Inoltre se uno si limita a richiedere la stabilità in L^∞ , ma non il principio di massimo, allora questa continua a valere sempre per lo schema di Crank–Nicolson (e più generalmente se vale la condizione $b\mu(1 - 2\theta) \leq \frac{1}{2}$).

Un'ultima osservazione riguarda la presenza di oscillazioni. Si osservi la figura 2.3. Qui il solito problema modello è risolto prendendo lo schema di Crank–Nicolson con μ uguale 1 e 2 rispettivamente per un dato iniziale che vale sempre 0 tranne che nel nodo centrale in cui vale 1. Si vede allora che nel caso $\mu = 2$ il principio di massimo non vale e la soluzione approssimata può addirittura diventare negativa. Per $\mu = 1$ il principio di massimo vale, ma la soluzione presenta due picchi distinti dopo il primo passo temporale (al contrario della soluzione esatta che ha sempre un solo punto di massimo al centro). Questo mostra che per usare il nostro schema in presenza di rapide variazioni della soluzione, sembra più conveniente prendere un valore di μ strettamente inferiore al limite dato dalla condizione di stabilità.

2.6 L'equazione di convezione-diffusione

Consideriamo ora uno schema per l'equazione di convezione-diffusione

$$u_t + au_x = bu_{xx}, \quad (2.23)$$

con $b > 0$. Ovviamente se poniamo $y = x - at$ e $w(y, t) = u(y + at, t)$, allora w sarà una soluzione dell'equazione (2.1) nella variabile y e quindi varranno considerazioni analoghe dal punto di vista della costruzione della soluzione. Dal punto di vista dell'approssimazione tuttavia siamo di fronte ad una sorta di concorrenza tra i due termini di natura diversa, la convezione e la diffusione. Cominciamo ad esaminare lo schema più naturale per approssimare l'equazione (2.23)

$$\frac{v_{m,n+1} - v_{m,n}}{k} + a \frac{v_{m+1,n} - v_{m-1,n}}{2h} = b \frac{v_{m+1,n} - 2v_{m,n} + v_{m-1,n}}{h^2}, \quad (2.24)$$

che è del prim'ordine in tempo e del secondo nello spazio e globalmente del secondo ordine essendo la condizione di stabilità data da

$$b\mu \leq \frac{1}{2}. \quad (2.25)$$

Per analizzare la validità del principio di massimo risulta conveniente scrivere lo schema nella forma

$$v_{m,n+1} = (1 - 2b\mu)v_{m,n} + b\mu(1 - \alpha)v_{m+1,n} + b\mu(1 + \alpha)v_{m-1,n}, \quad (2.26)$$

dove $\mu = k/h^2$ e $\alpha := \frac{ha}{2b} = \frac{a\lambda}{2b\mu}$. Abbiamo allora che il principio di massimo per lo schema è valido se e solo se tutti i coefficienti dello schema sono positivi e, assumendo per semplicità per ora che $a \geq 0$, questo vale se e solo se $\alpha \leq 1$.

Se $\alpha > 1$ in generale non varrà più il principio di massimo. Si prenda infatti come dato iniziale

$$v_{m,0} = \begin{cases} 1, & \text{per } m \leq 0, \\ -1, & \text{per } m > 0. \end{cases}$$

Abbiamo allora che $v_{0,1} = 1 + 2b\mu(\alpha - 1)$ che in modulo è maggiore di 1. Si noti che la condizione $\alpha \leq 1$ si scrive anche come

$$h \leq \frac{2b}{a}, \quad (2.27)$$

che costituisce una restrizione sul passo spaziale. Il numero a/b corrisponde a quello che in fluidodinamica viene chiamato il numero di Reynolds ed è un indice del grado di diffusività, in rapporto alla convezione, dell'equazione. Questa restrizione non causa instabilità numerica, poichè è ininfluenza per h abbastanza piccolo, ma può creare lo stesso delle oscillazioni se consideriamo dei problemi con un grande numero di Reynolds con una discretizzazione fissata, come per esempio può capitare per problemi di tipo non lineare, dove un fluido può avere una viscosità molto bassa. Supponiamo per esempio che $a = 0.001$ e

$b = 1$. Allora in questo caso dobbiamo avere che $h \leq 0.002$ e ancora peggio $k \leq 0.000002$.

Un modo per evitare la restrizione (2.27), è quello di utilizzare un'approssimazione di tipo upwind per il termine di ordine 1. Otteniamo allora lo schema

$$\frac{v_{m,n+1} - v_{m,n}}{k} + a \frac{v_{m,n} - v_{m-1,n}}{h} = b \frac{v_{m+1,n} - 2v_{m,n} + v_{m-1,n}}{h^2}, \quad (2.28)$$

se $a > 0$, che implica come condizione di monotonia

$$2b\mu + a\lambda \leq 1, \quad (2.29)$$

che è una specie di media tra la condizione di stabilità dello schema esplicito (2.7) e la condizione di stabilità del metod upwind con $a > 0$, ossia $a\lambda \leq 1$. Questa condizione risulta essere meno restrittiva della (2.27) se $b \leq \frac{1}{4\mu}$, al prezzo di una piccola restrizione sul passo temporale. Si noti tuttavia che lo schema (2.26) è equivalente a

$$\frac{v_{m,n+1} - v_{m,n}}{k} + a \frac{v_{m+1,n} - v_{m-1,n}}{2h} = \left(b + \frac{ah}{2} \right) \frac{v_{m+1,n} - 2v_{m,n} + v_{m-1,n}}{h^2},$$

ossia lo schema (2.24) con la viscosità b sostituita da $b' = b(1 + \alpha)$. Insomma è come se avessimo introdotto una sorta di viscosità artificiale $b\alpha$ nello schema per smussare le oscillazioni.

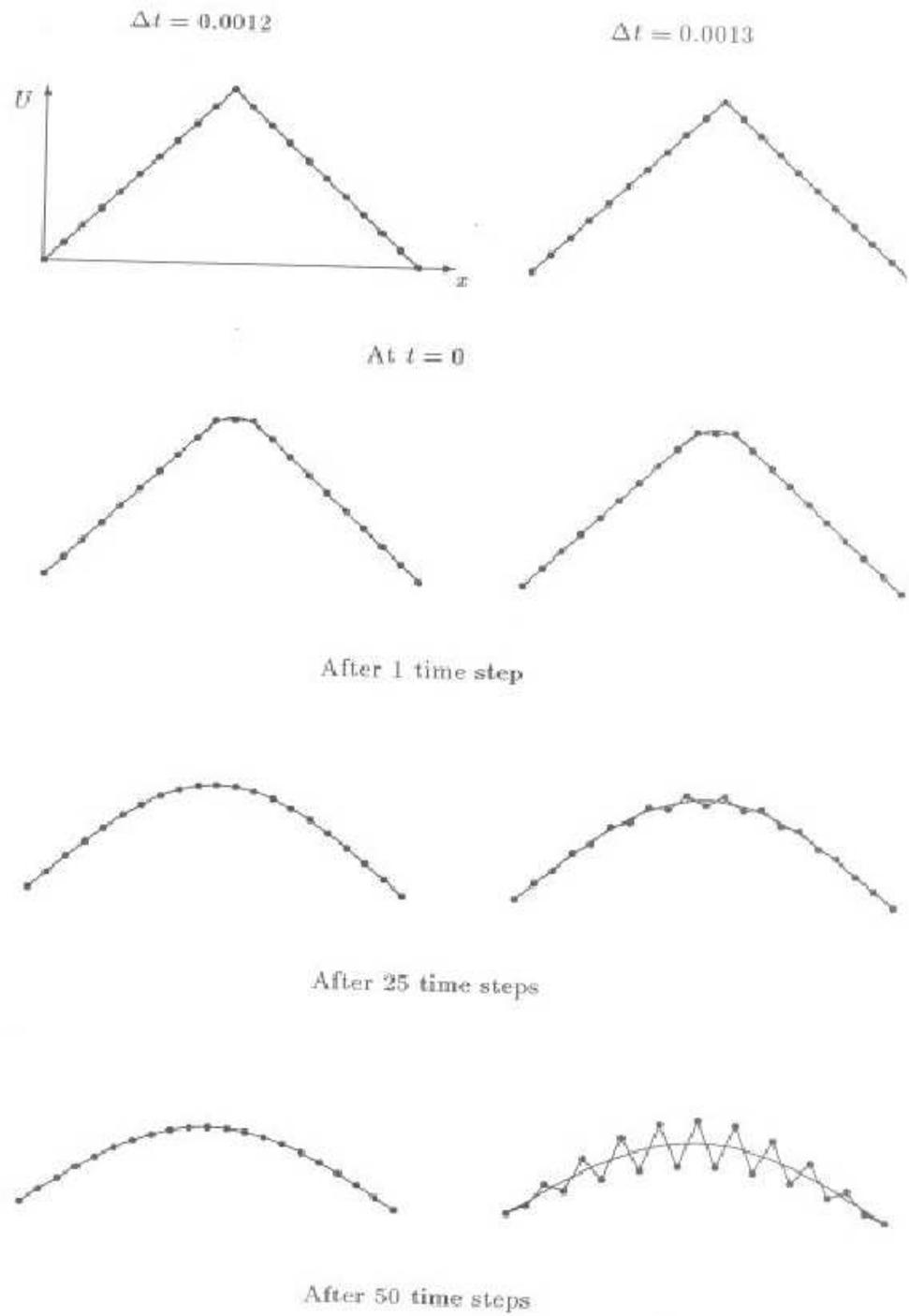


Figura 2.1: Equazione del calore con diverse condizioni di CFL

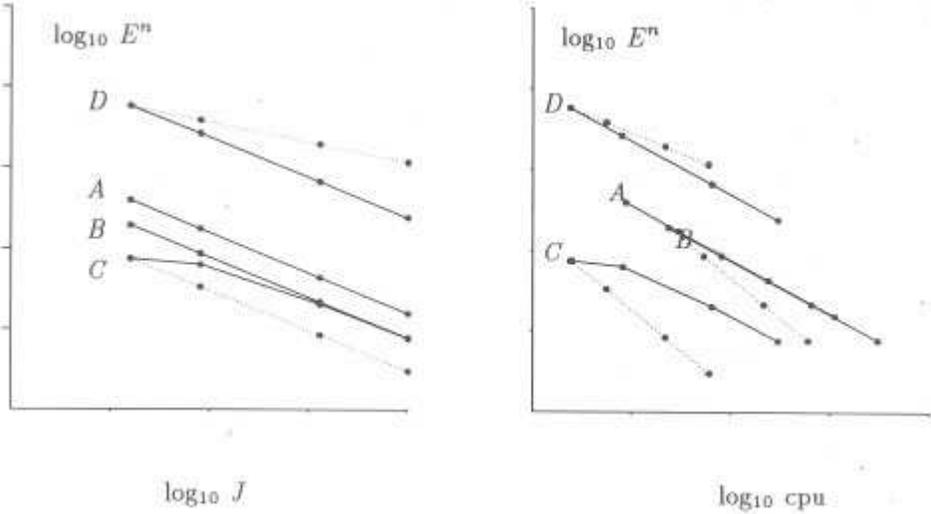


Figura 2.2: Grafici dell'errore

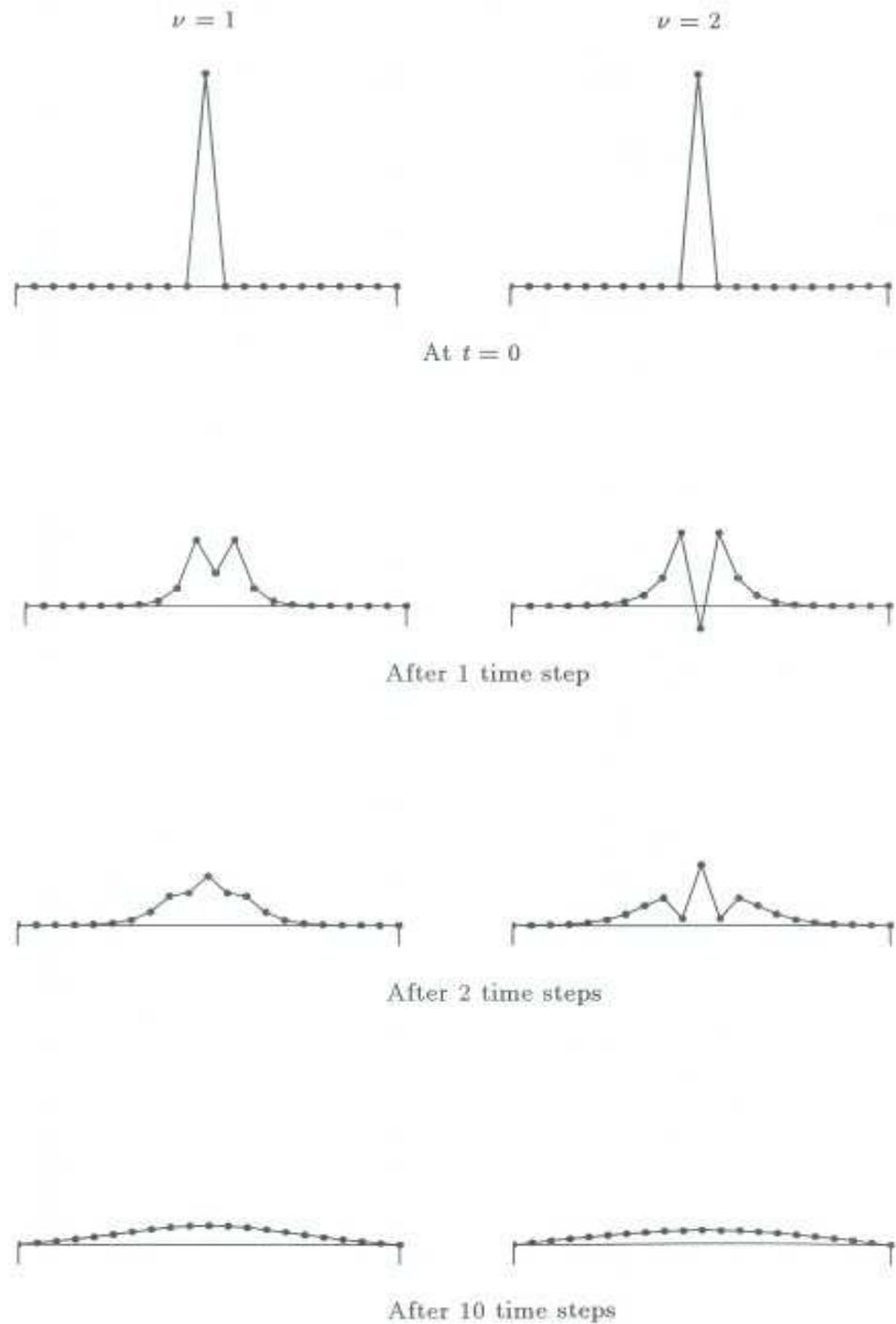


Figura 2.3: Crank-Nicolson con diverse condizioni di CFL

Capitolo 3

Leggi di conservazione non lineari

Molti problemi in meccanica, e in particolare in fluidodinamica ed elasticità, in biologia, in chimica ed in generale ovunque sia possibile fare una modellizzazione macroscopica dell'evoluzione di determinate quantità nel tempo, comportano lo studio di leggi di conservazione lineari iperboliche e paraboliche, che in una dimensione spaziale, e limitandoci per ora al caso scalare, possiamo scrivere come

$$u_t + A(u)_x = B(u)_{xx}, \quad (3.1)$$

dove A e B sono funzioni regolari dell'incognita u e la B sarà sempre non decrescente. La caratteristica principale di questo tipo di equazioni è che anche per dati iniziali regolari, il relativo problema di Cauchy non presenta in generale soluzioni regolari per tutti i tempi.

L'esempio più semplice si ha nel caso dell'equazione di Burgers (non viscosa)

$$u_t + \left(\frac{1}{2}u^2\right)_x = 0. \quad (3.2)$$

Consideriamo le curve caratteristiche relative a questa equazione, come già fatto nel caso a coefficienti variabili (1.9). Supponendo che la soluzione u sia regolare, diciamo C^1 , abbiamo che lungo le curve integrali dell'equazione differenziale ordinaria

$$\frac{dx}{dt} = u(x, t), \quad x(0) = \xi,$$

la u è costante poichè

$$\begin{aligned} \frac{d}{dt}u(x(t), t) &= u_t(x(t), t) + \frac{dx}{dt}u_x(x(t), t) \\ &= (u_t + (\frac{1}{2}u^2)_x)(x(t), t) = 0. \end{aligned}$$

Otteniamo quindi due conseguenze importanti: le caratteristiche sono rette e lungo queste rette la soluzione assume il valore del dato iniziale:

$$u(x(t), t) = u_0(\xi)$$

e

$$x(t) = \xi + tu_0(\xi).$$

Assumiamo ora che ci siano due punti, $\xi_1 < \xi_2$ tali che $u_0(\xi_1) > u_0(\xi_2)$. Questo sarà sempre vero a meno che il dato iniziale non sia monotono crescente su tutta la retta reale. Abbiamo allora che le caratteristiche uscenti dai punti ξ_1 e ξ_2 sono destinate ad incontrarsi esattamente per

$$t_0 = \frac{\xi_2 - \xi_1}{u_0(\xi_1) - u_0(\xi_2)}; \quad \xi_0 = \xi_2 u_0(\xi_1) - \xi_1 u_0(\xi_2).$$

In questo punto ci sarà però un assurdo, perchè la soluzione avrebbe due valori diversi $u_0(\xi_1) \neq u_0(\xi_2)$ portati dalle rispettive caratteristiche. Questo dimostra che non può esistere una soluzione globale regolare per l'equazione (3.2) a meno che il dato iniziale non sia monotono crescente. Questo ci obbligherà, volendo ancora utilizzare queste leggi di conservazione, a passare ad una nozione più debole di soluzione, che permetterà di trattare soluzioni con discontinuità.

Per quanto riguarda i problemi parabolici, se la diffusione è uniformemente positiva, ossia

$$B'(u) \geq \gamma > 0, \tag{3.3}$$

per ogni u nell'intervallo considerato, allora le soluzioni saranno necessariamente classiche, in pratica C^2 in x e C^1 in t , e avranno ottime proprietà di stabilità. La situazione cambia completamente in presenza di valori di u tali che l'equazione degenera, ossia $B'(u) = 0$. In questi punti la soluzione potrà presentare delle singolarità ed in generale si manifesterà la sua natura iperbolica attraverso una velocità finita di propagazione dei supporti.

È questo il caso per il modello più semplice di equazione parabolica degenera, l'equazione dei mezzi porosi

$$u_t = (u^{\gamma+1})_{xx}, \tag{3.4}$$

per $\gamma \geq 1$, per cui si possono esibire fenomeni di formazione di singolarità in tempo finito anche per dati regolarissimi, vedi [Ka, Ar].

Una prevalenza ancora più forte del carattere iperbolico si avrà nel caso fortemente degenera, dove esistono degli intervalli dove la funzione $B'(u) = 0$. In questi intervalli potranno prodursi delle discontinuità che interagiranno in maniera complessa con le regioni diffusive.

Dal punto di vista numerico questo tipo di non linearità pone dei problemi di non facile soluzione, sia per la scarsa regolarità delle soluzioni da approssimare, sia per il delicato bilancio tra fenomeni diffusivi e convettivi.

Nel seguito, dopo una breve presentazione di esempi fisici, faremo alcuni richiami sulla teoria di questo tipo di equazioni. Presenteremo quindi un

panorama dei principali metodi numerici e dei problemi relativi al loro studio qualitativo. Maggiore enfasi verrà data ai problemi di natura iperbolica, dove l'assenza di viscosità e la presenza di soluzioni discontinue pongono i maggiori problemi teorici e computazionali.

3.1 Alcuni esempi di problemi non lineari di evoluzione

Prima di passare a questi aspetti più teorici presenteremo alcuni esempi fisici che ci aiuteranno a capire meglio alcuni aspetti degli argomenti trattati. Ci limiteremo a tre esempi che danno luogo a semplici equazioni scalari. Per un panorama più completo dei modelli di tipo iperbolico e parabolico rimandiamo a [Se, Da, Wh, GR2, BER].

Esempio 3.1 (Flussi di traffico) Consideriamo il flusso delle automobili su di un'autostrada e sia $u = u(x, t)$ la densità di automobili (ossia il numero di automobili per unità di distanza, per esempio centinaia di metri) e sia $q = q(x, t)$ il loro flusso (automobili per unità di tempo, diciamo minuti) nella coordinata spaziale (uno dimensionale) x al tempo t . Fissiamo un certo segmento (a, b) sull'autostrada e due tempi abbastanza vicini $t_1 < t_2$. Poiché stiamo supponendo che le automobili non possano scomparire (o alzarsi in volo), dobbiamo pensare che quelle entrate dal punto a dovranno ad un certo punto uscire dal punto b . Questo vorrà dire che la differenza della quantità totale di automobili nel segmento tra i due istanti considerati

$$\int_a^b u(x, t_2) dx - \int_a^b u(x, t_1) dx$$

deve essere uguale alla differenza del flusso totale agli estremi

$$\int_{t_1}^{t_2} q(a, t) dt - \int_{t_1}^{t_2} q(b, t) dt.$$

Facendo tendere $b-a$ e t_2-t_1 a zero, dopo aver diviso gli integrali per il prodotto di queste stesse quantità, supponendo che le funzioni u e q siano regolari, si ottiene l'equazione differenziale

$$u_t + q_x = 0. \tag{3.5}$$

Risulta abbastanza realistico pensare che il flusso q sia in prima approssimazione solo una funzione della densità u . Più precisamente il flusso sarà espresso come il prodotto uV della densità per la velocità media V . A questo punto facciamo l'ipotesi che la velocità delle macchine sia una funzione (regolare) strettamente decrescente della densità che vale V_{max} per $u = 0$ e 0 per $u = u_{max}$, ossia la densità in cui le macchine sono attaccate l'una all'altra. Sia per esempio

$V(u) = V_{max} \left(1 - \frac{u}{u_{max}}\right)$. Allora

$$q(u) = V_{max} \left(1 - \frac{u}{u_{max}}\right) u.$$

Il flusso è nullo in assenza di automobili o per la densità massima, mentre risulta essere massimo per $u = u_{max}/2$. La presenza di discontinuità è facilmente constatabile se qualcuno frena di colpo all'interno di una fila di macchine. La densità avrà allora una discontinuità che si propagherà all'indietro lungo la fila.

Esempio 3.2 Flusso di un gas in un mezzo poroso Supponiamo di voler studiare la penetrazione di un gas perfetto in un mezzo poroso. Sia ρ la densità di questo gas, \mathbf{v} la sua velocità (vettoriale) e m la porosità del mezzo. Vale allora, come nell'esempio precedente, la legge di conservazione della massa

$$(m\rho)_t + \nabla \cdot (\rho\mathbf{v}) = 0. \quad (3.6)$$

Questa volta però la velocità non dipende direttamente dalla densità, ma a causa di complicati processi microscopici, è una funzione lineare del gradiente della pressione p , come prescritto dalla legge di Darcy

$$\mathbf{v} = -\frac{k}{\mu} \nabla p, \quad (3.7)$$

dove k è la permeabilità del mezzo e μ il coefficiente di viscosità del gas. Per chiudere il sistema di equazioni facciamo l'ipotesi che il gas obbedisca ad una legge politropica della pressione

$$p = c\rho^\gamma, \quad (3.8)$$

dove c è una costante di normalizzazione e $\gamma \in [1, 5/3]$ è la costante caratteristica del gas. Abbiamo allora un'equazione chiusa per la ρ :

$$(m\rho)_t - \nabla \cdot \left(\frac{kC}{\mu} \rho \nabla (\rho^\gamma) \right) = 0. \quad (3.9)$$

Supponendo tutte le quantità fisiche costanti otteniamo finalmente la forma multidimensionale dell'equazione dei mezzi porosi

$$\rho_t = a\Delta(\rho^{\gamma+1}), \quad (3.10)$$

con $a = \frac{kC\gamma}{\mu m(\gamma+1)}$.

Esempio 3.3 Flussi bifasici in mezzi porosi

Presentiamo adesso un modello, alquanto semplificato, di flusso bifasico acqua-petrolio in un giacimento petrolifero. Sia allora s la funzione di saturazione dell'acqua e $1-s$ quella del petrolio. In ipotesi di incompressibilità

il loro spostamento in un mezzo poroso può essere descritto dalle seguenti equazioni (in forma non dimensionale):

$$m(x)s_t + \nabla \cdot (f(s)\mathbf{v} + f_g(s)K\nabla h) = \nabla \cdot (d(x,s)\nabla s) + q(x), \quad (3.11)$$

$$\mathbf{v} = -K\lambda(x,s)(\nabla p - \rho(s)\nabla h), \quad (3.12)$$

$$\nabla \cdot \mathbf{v} = 0. \quad (3.13)$$

Queste equazioni esprimono rispettivamente il bilancio della fase acquifera, la legge di Darcy e la condizione di incomprimibilità. Le incognite sono rappresentate, oltre che dalla saturazione s , anche dalla velocità totale di Darcy \mathbf{v} e dalla pressione globale p . Le varie funzioni che vi compaiono sono: m la porosità della roccia, f , la frazione di flusso, che è una funzione non lineare sperimentale che dipende dalle mobilità relative delle fasi; f_g e ρ sono funzioni delle densità relative delle fasi, K la permeabilità assoluta della roccia, h l'altezza, d la funzione di diffusione, q la funzione di pompaggio, mentre λ è la mobilità totale delle fasi. Si osservi che la funzione d si annulla per $s = 0, 1$ e può essere molto piccola a seconda dei valori della permeabilità del mezzo. Le equazioni (3.11)–(3.13) formano un sistema misto accoppiato non linearmente. Mentre le equazioni per la velocità e la pressione possono essere risolte con le tecniche delle equazioni ellittiche, l'equazione per la saturazione s è un'equazione non lineare che contiene termini di convezione dominanti e di diffusione non lineare, eventualmente degenerare.

3.2 Richiami di teoria delle soluzioni deboli

3.2.1 Esistenza locale di soluzioni classiche nel caso iperbolico

Si consideri adesso l'equazione iperbolica

$$u_t + A(u)_x = 0, \quad (3.14)$$

per $x \in \mathbb{R}$ e $t \geq 0$ con il dato iniziale

$$u(x, 0) = u_0. \quad (3.15)$$

Mostriamo che esistono delle soluzioni regolari per intervalli di tempo positivi, ma in generale queste soluzioni non saranno globali. Supponiamo che $A \in C^2$ e sia u una soluzione C^1 . Una caratteristica per l'equazione sarà allora una curva soluzione dell'equazione differenziale ordinaria

$$\frac{dx}{dt} = A'(u(x, t)), \quad x(0) = \xi,$$

Allora lungo queste curve la soluzione sarà costante poichè

$$\begin{aligned} \frac{d}{dt}u(x(t), t) &= u_t(x(t), t) + \frac{dx}{dt}u_x(x(t), t) \\ &= (u_t + A'(u)u_x)(x(t), t) = 0. \end{aligned}$$

Otteniamo quindi due conseguenze importanti: le caratteristiche saranno rette e lungo queste rette la soluzione assumerà il valore del dato iniziale:

$$u(x(t), t) = u_0(\xi) \quad (3.16)$$

e

$$x(t) = \xi + tA'(u_0(\xi)). \quad (3.17)$$

Sia allora, per t fissato, $F(\xi; t) = \xi + tA'(u_0(\xi))$. Se u_0 è continua lo è anche F e se u_0 è limitata allora $F(\pm\infty; t) = \pm\infty$, per cui dato x esiste sempre ξ tale che $F(\xi; t) = x$. Il punto cruciale consiste nel dimostrare che, almeno su piccoli intervalli di tempo, questo valore di ξ è unico. Possiamo così dimostrare l'esistenza almeno locale di una soluzione classica.

Proposizione 3.4 *Sia $A \in C^2(\mathbb{R})$ e $u_0 \in C^1(\mathbb{R})$, con $|u_0|, |u_0'| \leq M$ per una qualche costante $M > 0$. Sia T^* uguale a ∞ se $A'(u_0)$ è non decrescente, altrimenti poniamo $T^* = -(\inf(A'(u_0))')^{-1}$. Allora esiste una (unica) soluzione classica $u \in C^1(\mathbb{R} \times [0, T^*))$ per il problema (3.14)–(3.15). Inoltre, se $T^* < \infty$ allora*

$$\lim_{t \rightarrow T^*} \|u_x(\cdot, t)\|_{L^\infty} = \infty,$$

e non esiste nessuna soluzione regolare su $\mathbb{R} \times [0, T)$ per nessun $T > T^*$.

Dimostrazione. Abbiamo che se $A'(u_0)$ è non decrescente allora F è sempre crescente strettamente come funzione di ξ , altrimenti

$$F_\xi = 1 + t(A'(u_0(\xi)))' \geq 1 - \frac{t}{T^*} > 0,$$

per ogni $t \in [0, T^*)$. Per ogni x esiste quindi un'unico valore di $\xi = \xi(x, t)$ tale che $F(\xi; t) = x$ che dipende in modo regolare da (x, t) . Si verifica facilmente, dal calcolo delle derivate, che $u(x, t) = u_0(\xi(x, t))$ è soluzione classica di (3.14)–(3.15). Poichè si ha

$$u_x = \frac{u_0'}{1 + t(A'(u_0))'},$$

abbiamo che la soluzione può essere continuata fino a che la derivata u_x rimane limitata per tutti gli $x \in \mathbb{R}$. \square

Il tempo di esistenza massimale della soluzione T^* , se finito, viene chiamato tempo di blow-up del gradiente ed è associato alla formazione di discontinuità nella soluzione. Siamo quindi costretti per continuare tale soluzione al di là di questo tempo, a introdurre il concetto di soluzione debole.

3.2.2 Soluzioni deboli e condizioni di entropia

Definizione 3.1 *Sia $u_0 \in L^\infty(\mathbb{R})$. Allora u è una soluzione debole di (3.14)–(3.15), detta anche soluzione distribuzionale, se $u \in L^\infty(\mathbb{R} \times [0, \infty))$ e*

$$\iint_{\mathbb{R} \times [0, \infty)} u \phi_t + A(u) \phi_x dt dx + \int_{\mathbb{R}} u_0(x) \phi(x, 0) dx = 0, \quad (3.18)$$

per ogni funzione $\phi \in C_0^\infty(\mathbb{R} \times [0, \infty))$.

Le soluzioni che siano regolari a tratti, con curve di discontinuità regolari possono essere caratterizzate completamente dalla condizione di Rankine–Hugoniot. Diremo che una funzione è C^1 a tratti se esiste un numero finito di curve regolari $\Gamma = \{x = s(t)\}$ nel piano (x, t) al di fuori delle quali $u \in C^1$ e attraverso le quali ha una discontinuità di salto. Data una curva di discontinuità Γ ed $\mathbf{n} = (-s'(t), 1)$ la sua normale orientata, porremo allora

$$u_\pm(s(t), t) = \lim_{\varepsilon \rightarrow 0} u((s(t), t) \pm \mathbf{n}).$$

Proposizione 3.5 *Sia u una funzione è C^1 a tratti nel piano (x, t) . Allora u è una soluzione debole di (3.14)–(3.15) se e solo se*

- u è una soluzione regolare di (3.14), nei punti in cui la u è C^1 ;
- vale la condizione

$$(u_+ - u_-)s'(t) = A(u_+) - A(u_-), \quad (3.19)$$

lungo ogni curva di discontinuità.

La dimostrazione, basata sulla formula di Gauss–Green, è elementare e viene omessa, vedi però [GR1, Sm, Da].

Grazie a questa proposizione diventa facile stabilire se una data funzione è soluzione debole o meno dell'equazione (3.14). Questo permette inoltre di far vedere immediatamente la mancanza di unicità di tali soluzioni deboli.

Esempio 3.6 Vogliamo risolvere l'equazione di Burgers (3.2) con il dato iniziale $u_0 \equiv 0$. Naturalmente ci aspettiamo che la soluzione identicamente nulla sia l'unica soluzione (regolarissima quindi) del problema. Si verifica però che è possibile costruire un numero infinito di soluzioni deboli e C^1 a tratti, non identicamente nulle, semplicemente prendendo

$$u(x, t) = \begin{cases} 0, & \text{per } x < -\frac{1}{2}\alpha t, \\ -\alpha, & \text{per } -\frac{1}{2}\alpha t < x < 0, \\ \alpha, & \text{per } 0 < x < \frac{1}{2}\alpha t, \\ 0, & \text{per } x > \frac{1}{2}\alpha t, \end{cases}$$

per ogni $\alpha > 0$.

Più generalmente è possibile vedere come, anche partendo da un dato regolare, nel momento di continuare la soluzione inserendo una discontinuità, non ci sia un modo unico di fare questa scelta.

Questa mancanza di unicità è dovuta dal punto di vista fisico al fatto che i nostri modelli iperbolici, che ricordiamo nascono come approssimazione *macroscopica* di certi processi, non contengono di fatto tutta l'informazione fisica sul livello *microscopico* sottostante. In particolare viene trascurato il secondo principio della termodinamica che di fatto sancisce l'irreversibilità statistica

di tutti i fenomeni con una complessità sufficiente. Per dare un equivalente matematico dell'entropia, che permetta di trattare modelli che non nascono da una modellizzazione di tipo cinetico, risulta vantaggioso utilizzare le leggi supplementari di conservazione. Sia u una soluzione regolare di (3.14). Allora per ogni funzione regolare $\eta = \eta(u)$, sia q una funzione che soddisfa

$$q' = \eta' A', \quad (3.20)$$

ossia

$$q(u) = \int_{u_0}^u \eta'(s) A'(s) ds,$$

per qualche u_0 fissato. Abbiamo allora che u soddisfa la legge supplementare di conservazione

$$\eta_t(u) + q_x(u) = 0. \quad (3.21)$$

Quindi per un'equazione scalare è sempre possibile trovare una legge supplementare soddisfatta da tutte le soluzioni regolari. Si può in effetti vedere che per tutti i sistemi fisicamente rilevanti esiste almeno una legge supplementare di conservazione. In gas-dinamica per esempio, la quantità η è proprio data dall'entropia fisica (cambiata di segno). La coppia (η, q) è chiamata coppia di entropia-flusso di entropia per l'equazione (3.14).

Si osservi però che l'uguaglianza (3.21) non è più soddisfatta in generale per una soluzione debole. Infatti, analogamente a quanto stabilito per la relazione di Rankine-Hugoniot, se u è una soluzione C^1 a tratti che verifica l'equazione (3.21), allora lungo una curva di discontinuità $\Gamma = \{x = s(t)\}$, si avrà

$$(\eta(u_+) - \eta(u_-))s'(t) = (q(u_+) - q(u_-)).$$

In generale questa relazione è incompatibile con la (3.19). Si consideri infatti di nuovo l'equazione di Burgers (3.2) con il dato iniziale

$$u_0(x, 0) = \begin{cases} 0, & \text{per } x < 0, \\ 1, & \text{per } x > 0. \end{cases} \quad (3.22)$$

Una soluzione debole è allora data da

$$u(x, t) = \begin{cases} 0, & \text{per } x < \frac{1}{2}t, \\ 1, & \text{per } x > \frac{1}{2}t, \end{cases} \quad (3.23)$$

essendo la curva di discontinuità data da $s(t) = \frac{1}{2}t$. Sia ora $\eta(u) = \frac{1}{2}u^2$ con $q(u) = \frac{1}{3}u^3$. Lungo la curva di discontinuità si ha allora che

$$(\eta(u_+) - \eta(u_-))s'(t) = \frac{1}{4} < \frac{1}{3} = (q(u_+) - q(u_-)). \quad (3.24)$$

Si osservi che prendendo il dato iniziale con il salto opposto

$$u_0(x, 0) = \begin{cases} 1, & \text{per } x < 0, \\ 0, & \text{per } x > 0. \end{cases} \quad (3.25)$$

e la relativa soluzione

$$u(x, t) = \begin{cases} 1, & \text{per } x < \frac{1}{2}t, \\ 0, & \text{per } x > \frac{1}{2}t, \end{cases} \quad (3.26)$$

la curva di discontinuità non cambia ma la disuguaglianza (3.24) si inverte

$$(\eta(u_+) - \eta(u_-))s'(t) = -\frac{1}{4} > -\frac{1}{3} = (q(u_+) - q(u_-)). \quad (3.27)$$

Il Secondo principio della termodinamica stabilisce che la funzione (concava) di entropia di un sistema tende ad aumentare. Per ragioni matematiche è più conveniente avere a che fare con funzionali decrescenti (tipo Lyapunov), per cui sceglieremo le entropie convesse. Questo ci condurrà alla seguente definizione.

Definizione 3.2 Una soluzione debole u di (3.14)–(3.15) è una soluzione entropica dello stesso problema se, per ogni coppia (η, q) di entropia-flusso di entropia, con η convessa, si ha

$$\iint_{\mathbb{R} \times [0, \infty)} \eta(u)\phi_t + q(u)\phi_x dt dx + \int_{\mathbb{R}} \eta(u_0(x))\phi(x, 0) dx \geq 0, \quad (3.28)$$

per ogni funzione non negativa $\phi \in C_0^\infty(\mathbb{R} \times [0, \infty))$.

Si osservi che la (3.28) esprime semplicemente che l'uguaglianza (3.21), sulle soluzioni entropiche diventa una disuguaglianza (nel senso delle distribuzioni)

$$\eta_t(u) + q_x(u) \leq 0. \quad (3.29)$$

In particolare, ragionando ancora come per la relazione di Rankine–Hugoniot, avremo che per una soluzione regolare a tratti la condizione di ammissibilità di una curva di shock $x = s(t)$, sarà data dalla disuguaglianza

$$(\eta(u_+) - \eta(u_-))s'(t) \geq (q(u_+) - q(u_-)). \quad (3.30)$$

Per esempio la soluzione (3.23) non è ammissibile, mentre lo è la (3.26). Si può verificare inoltre che tutte le soluzioni non nulle nell'esempio 3.6 non sono ammissibili.

Un fondamento abbastanza rigoroso della condizione di entropia è dato dalla teoria dei limiti idrodinamici delle teorie cinetiche, vedi ad esempio [CIP], ma in questo contesto ci limiteremo solo a considerare le leggi di conservazione iperboliche come i limiti di perturbazioni singolari di problemi parabolici. In particolare si consideri una perturbazione viscosa dell'equazione (3.14)

$$u_t + A(u)_x = \varepsilon u_{xx}. \quad (3.31)$$

Avendo aggiunto un termine parabolico (non degenere) la soluzione verrà regolarizzata e, almeno nel caso scalare, risulta facile trovare una soluzione globale regolare del problema di Cauchy (3.31)–(3.15). Si ha allora la seguente caratterizzazione dei limiti per $\varepsilon \rightarrow 0$.

Teorema 3.7 *Sia u^ε una successione regolare di soluzioni del problema di Cauchy (3.31)–(3.15) al variare di $\varepsilon > 0$, tale che, per una certa costante $C > 0$, indipendente da ε ,*

$$\|u^\varepsilon\|_{L^\infty} \leq C; \quad (3.32)$$

e

$$u^\varepsilon \rightarrow u, \quad \text{q.o. in } (x, t). \quad (3.33)$$

Allora il limite u della successione è una soluzione entropica di (3.14)–(3.15).

Dimostrazione. Sia η una qualsiasi entropia convessa regolare. Si moltiplichi l'equazione (3.31) per $\eta'(u^\varepsilon)$, e, utilizzando, la (3.20) si ottiene

$$\eta_t(u^\varepsilon) + q_x(u^\varepsilon) = \varepsilon \eta'(u^\varepsilon) u_{xx}^\varepsilon.$$

Si osserva allora che vale l'uguaglianza

$$\eta'(u^\varepsilon) u_{xx}^\varepsilon = \eta_{xx}(u^\varepsilon) - \eta''(u^\varepsilon) |u_x^\varepsilon|^2,$$

per cui abbiamo, grazie alla convessità di η ,

$$\eta_t(u^\varepsilon) + q_x(u^\varepsilon) \leq \varepsilon \eta_{xx}(u^\varepsilon). \quad (3.34)$$

Utilizzando il Teorema della convergenza dominata di Lebesgue possiamo allora passare al limite nel senso delle distribuzioni in (3.34) ottenendo la (3.28). \square

Si osservi che un'analogia dimostrazione si può fare anche nel caso di sistemi iperbolici multidimensionali che siano dotati di entropie convesse.

Per chiarire meglio il (difficile) concetto di soluzione entropica risulta utile trattare il semplice caso dei flussi convessi. Vale il seguente risultato, dovuto a Lax [La].

Proposizione 3.8 *Sia u una funzione è C^1 a tratti nel piano (x, t) . Allora u è una soluzione entropica di (3.14)–(3.15) se e solo se lungo le curve di shock $x = s(t)$ vale la seguente disuguaglianza*

$$A'(u_-) > s'(t) > A'(u_+). \quad (3.35)$$

Chiaramente dalla (3.19) si ottiene

$$A'(u_-) > \frac{A(u_+) - A(u_-)}{u_+ - u_-} > A'(u_+). \quad (3.36)$$

Possiamo interpretare geometricamente questa condizione dicendo che le caratteristiche della soluzione devono “entrare” nello shock. In gas dinamica questo corrisponde a delle condizioni di compressione del gas. Si osservi che dalla convessità di A la condizione (3.35) è verificata se e solo se

$$u_- > u_+. \quad (3.37)$$

3.2.3 Il problema di Riemann

Possiamo ora presentare una classe particolare di soluzioni esatte del problema (3.14)–(3.15), almeno nel caso elementare dei flussi convessi. Supponiamo che il dato iniziale sia del tipo di Riemann

$$u_0 = \begin{cases} u_s, & \text{per } x < 0, \\ u_d, & \text{per } x > 0, \end{cases} \quad (3.38)$$

per due costanti fissate u_s e u_d . Vogliamo costruire la soluzione entropica del problema (3.14)–(3.15) corrispondente a questo dato iniziale. Come vedremo nel Teorema 3.10, almeno nel caso scalare le soluzioni entropiche sono uniche e quindi basterà esibirne una per trovare quella giusta. Intanto osserviamo che se u è una soluzione debole (entropica) di (3.14), allora pure la soluzione riscalata $u_\lambda = u(\lambda x, \lambda t)$ lo è per ogni $\lambda > 0$. In questo caso particolare, poichè il dato iniziale è invariante per queste dilatazioni abbiamo anche che, sempre a causa dell'unicità,

$$u_\lambda(x, t) = u(x, t),$$

per ogni $\lambda > 0$. Questo implica, prendendo $\lambda = \frac{1}{t}$, che la soluzione è della forma

$$u(x, t) = W\left(\frac{x}{t}\right),$$

per una certa funzione di una variabile W . Ossia stiamo dicendo che la u è costante lungo le rette del tipo $\frac{x}{t} = \text{cost.}$. Se $u_s = u_d$, la soluzione costante sarà l'unica soluzione entropica. Altrimenti sia $u_s > u_d$. Allora dalla condizione di entropia Lax per i flussi convessi (3.37), avremo che la soluzione entropica è data da uno shock isolato che separa i due valori costanti

$$u(x, t) = \begin{cases} u_s, & \text{per } x < ct, \\ u_d, & \text{per } x > ct, \end{cases} \quad (3.39)$$

con

$$c = \frac{A(u_d) - A(u_s)}{u_d - u_s}.$$

Si consideri infine il caso restante $u_s < u_d$. Se provassimo a definire una soluzione data da uno shock come in (3.39), non avremmo una soluzione che soddisfa la condizione di entropia. Osserviamo allora che nelle regioni in cui la u è regolare, la W deve soddisfare l'equazione differenziale ordinaria

$$\left(-\frac{x}{t^2} + \frac{1}{t}A'(W)\right)W' = 0.$$

Questo implica che, dove W' è diversa da 0, allora

$$A'(W(\xi)) = \xi, \quad (3.40)$$

per $\xi = x/t$. Essendo la funzione A strettamente convessa, possiamo definire la funzione inversa di A' e ricavare dalla (3.40)

$$W(\xi) = (A')^{-1}(\xi).$$

Riassumendo una soluzione entropica per questo caso sarà data da

$$u(x, t) = \begin{cases} u_s, & \text{per } x < A'(u_s)t, \\ (A')^{-1}\left(\frac{x}{t}\right), & \text{per } A'(u_s)t < x < A'(u_d)t, \\ u_d, & \text{per } x > A'(u_d)t. \end{cases} \quad (3.41)$$

Questa soluzione è continua e quindi soddisfa automaticamente la condizione di entropia. Fisicamente corrisponde a degli stati espansivi dei fluidi ed è chiamata comunemente onda di “rarefazione”.

3.2.4 Esistenza e unicità delle soluzioni entropiche

Esponiamo di seguito i risultati più importanti riguardanti le soluzioni entropiche e che essenzialmente sono dovuti a S.N. Kružkov [Kr]. Per le dimostrazioni rimandiamo a questo lavoro fondamentale o a una delle numerose esposizioni recenti, per esempio [Se, Da]. Intanto diamo una caratterizzazione diversa della condizione di entropia

Proposizione 3.9 *Una funzione misurabile e limitata su $\mathbb{R} \times [0, T)$ è una soluzione entropica di (3.14)–(3.15) se e solo se soddisfa la relazione*

$$\begin{aligned} \iint_{\mathbb{R} \times [0, \infty)} |u - k| \phi_t + \operatorname{sgn}(u - k)(A(u) - A(k)) \phi_x dt dx \\ + \int_{\mathbb{R}} |u_0(x) - k| \phi(x, 0) dx \geq 0, \end{aligned} \quad (3.42)$$

per ogni $k \in \mathbb{R}$ e ogni funzione non negativa $\phi \in C_0^\infty(\mathbb{R} \times [0, T))$.

In pratica la condizione (3.42) corrisponde a scegliere come entropia-flusso di entropia la sottoclasse

$$\eta_k(u) = |u - k|, \quad q_k(u) = \operatorname{sgn}(u - k)(A(u) - A(k)).$$

Questa classe è sufficiente per assicurare la condizione per tutte le entropie convesse. Vale allora il risultato fondamentale.

Teorema 3.10 *Per ogni funzione misurabile limitata u_0 , esiste un'unica soluzione entropica del problema (3.14)–(3.15) e $u \in C([0, \infty); L_{loc}^1(\mathbb{R}))$.*

Riassumiamo le proprietà principali della soluzione nella proposizione seguente.

Proposizione 3.11 *Siano u_0 e v_0 due funzioni misurabili limitate e u e v le soluzioni entropiche relative. Sia*

$$M := \sup\{|A'(s)|; s \in [\inf(u_0, v_0), \sup(u_0, v_0)]\}.$$

Allora:

- Per qualsiasi $t > 0$ e qualsiasi intervallo $[a, b]$, si ha

$$\int_a^b |v(x, t) - u(x, t)| dx \leq \int_{a-Mt}^{b+Mt} |v_0(x) - u_0(x)| dx;$$

- Se u_0 e v_0 coincidono su di un intervallo (α, β) allora u e v coincidono nel triangolo $\{(x, t); \alpha + Mt < x < \beta - Mt, 0 < t < \frac{\beta - \alpha}{M}\}$;
- Se $u_0 - v_0 \in L^1(\mathbb{R})$ allora per ogni t fissato $u(\cdot, t) - v(\cdot, t) \in L^1(\mathbb{R})$ e

$$\int (v(x, t) - u(x, t)) dx = \int (u_0(x) - v_0(x)) dx,$$

$$\|u(\cdot, t) - v(\cdot, t)\|_{L^1} \leq \|u_0 - v_0\|_{L^1};$$

- Se $u_0 \leq v_0$ per q.o. $x \in \mathbb{R}$ allora $u(x, t) \leq v(x, t)$ per q.o. $(x, t) \in \mathbb{R} \times [0, \infty)$;
- Se $u_0 \in BV(\mathbb{R})$, lo spazio delle funzioni a Variazione Totale limitata, allora per ogni $t > 0$ anche $u(\cdot, t) \in BV(\mathbb{R})$ e

$$TV(u(\cdot, t)) \leq TV(u_0).$$

Si osservi come lo spazio delle funzioni a Variazione Totale limitata, lo spazio BV, è in effetti lo spazio più piccolo (ossia più regolare) che risulti stabile per l'evoluzione delle soluzioni deboli di (3.14). Esistono infatti dati iniziali C_0^∞ che sviluppano un insieme di shock denso in tempo finito, per cui la classe delle funzioni regolari a tratti, benchè genericamente stabile, non risulta adeguata al problema.

3.2.5 Problemi parabolici non lineari

Consideriamo ora un problema non lineare con convezione e diffusione del tipo

$$u_t + A(u)_x = B(u)_{xx}, \quad (3.43)$$

con un dato iniziale

$$u(x, 0) = u_0(x). \quad (3.44)$$

Se vale la condizione (3.3), allora si può applicare la teoria classica dei problemi parabolici, vedi per esempio [LSU]. Altrimenti dobbiamo distinguere due casi, quello semplicemente degenere, dove la funzione B' può annullarsi, ma la funzione B rimane sempre strettamente crescente, dal caso fortemente degenere dove la funzione B' può essere nulla su intervalli.

Nel primo caso, che comprende per esempio l'equazione dei mezzi porosi (3.4), è possibile mostrare che il problema non ammette soluzioni classiche globali, vedi [Ka, Ar]. Tuttavia la soluzione può essere continuata in senso debole.

Definizione 3.3 *Sia B una funzione strettamente crescente. Sia $u_0 \in L^1 \cap L^\infty \cap BV$. Allora u è una soluzione debole di (3.43)–(3.44) se è limitata e misurabile e in BV e valgono*

1. $B(u)_x \in L^2$;
2. per ogni funzione $\phi \in C_0^\infty(\mathbb{R} \times [0, \infty))$

$$\iint_{\mathbb{R} \times [0, \infty)} u \phi_t + [A(u) - B(u)_x] \phi_x dt dx + \int_{\mathbb{R}} u_0(x) \phi(x, 0) dx = 0. \quad (3.45)$$

L'esistenza di tali soluzioni deboli è stata dimostrata, in effetti nel caso multidimensionale, usando la regolarizzazione parabolica da Vol'pert e Hudjaev [VH] nel 1969. L'unicità è stata mostrata da Yin [Yi], senza bisogno di condizioni di entropia. In seguito Zhao [Zh] ha dimostrato che queste soluzioni, almeno in una variabile spaziale, sono sempre continue. Per equazioni speciali, per esempio se le funzioni A e B sono potenze di u , c'è una vasta letteratura che riguarda specialmente il problema della propagazione delle interfacce, vedi [Gi2, Gi1] e i problemi connessi alla regolarità delle soluzioni [Di1], che in generale saranno soltanto Hölderiane.

Il caso delle equazioni fortemente degeneri è stato molto meno studiato, anche se di interesse applicativo in alcuni problemi di sedimentazione [EK1], e solo recentemente si sono ottenuti risultati abbastanza soddisfacenti. Osservando che in particolare le equazioni iperboliche possono essere viste come un caso particolare di questa classe, si capisce immediatamente la necessità di introdurre l'analogo delle condizioni di entropia del caso iperbolico al fine di permettere una trattazione rigorosa delle soluzioni discontinue.

Possiamo introdurre allora una condizione analoga a quella di Kružkov.

Definizione 3.4 *Sia B una funzione non decrescente. Una funzione misurabile e limitata su $\mathbb{R} \times [0, T)$ è una soluzione entropica di (3.43)–(3.44) se è una soluzione debole e soddisfa la relazione*

$$\begin{aligned} \iint_{\mathbb{R} \times [0, \infty)} |u - k| \phi_t + \operatorname{sgn}(u - k) [(A(u) - A(k)) - B(u)_x] \phi_x dt dx \\ + \int_{\mathbb{R}} |u_0(x) - k| \phi(x, 0) dx \geq 0, \end{aligned} \quad (3.46)$$

per ogni $k \in \mathbb{R}$ e ogni funzione non negativa $\phi \in C_0^\infty(\mathbb{R} \times [0, T))$.

Anche in questo caso l'esistenza segue dalla regolarizzazione parabolica, vedi [VH], ma l'unicità è stata dimostrata solo recentemente da Wu e Yin [WY] nel caso monodimensionale e da Carrillo [Ca] nel caso multidimensionale.

In particolare è possibile dare per le soluzioni regolari a tratti delle condizioni di salto di tipo Rankine–Hugoniot e Lax.

Proposizione 3.12 *Sia u una soluzione entropica regolare a tratti di (3.43)–(3.44). Se l'insieme di degenerazione di B si riduce ad un intervallo $[a, b]$, allora un salto della soluzione (shock) u tra due valori u_- (a sinistra) e u_+ (a destra) può avvenire solo se $u_-, u_+ \in [a, b]$ e le seguenti condizioni sullo shock sono soddisfatte:*

1. La velocità s' dello shock è data da

$$(u_+ - u_-)s'(t) = (A(u_+) - (B(u)_x)_+) - (A(u_-) - (B(u)_x)_-);$$

2. vale la seguente disuguaglianza di entropia per ogni $k \in \mathbb{R}$

$$\frac{A(u_+) - A(k) - (B(u)_x)_+}{u_+ - k} \leq s' \leq \frac{A(u_-) - A(k) - (B(u)_x)_-}{u_- - k}.$$

In generale le soluzioni dei problemi parabolici fortemente degeneri esibiranno una struttura molto complessa.

Esempio 3.13 Questo esempio è tratto da [EK1]. Si consideri l'equazione di tipo Burgers

$$u_t + \left(\frac{1}{2}u^2\right)_x = B(u)_{xx}, \quad (3.47)$$

con

$$B'(u) = \begin{cases} 0, & \text{per } u \leq 0.5, \\ 2.5u - 1.25, & \text{per } u \in (0.5, 0.6), \\ 0.25, & \text{per } u > 0.6. \end{cases} \quad (3.48)$$

La funzione B è quindi degenere per $u \leq 0.5$, quadratica nell'intervallo $(0.5, 0.6)$ e lineare per $u > 0.6$. Nella figura 3.1 si vede il grafico del dato iniziale, a sinistra la soluzione che corrisponde all'equazione di Burgers ($B = 0$), e a destra la soluzione corrispondente a (3.47), entrambe per $t = 0.15$. Un metodo numerico convergente, vedi sezione 3.6, è utilizzato per calcolare la soluzione con un passo molto piccolo. Quello che si nota è che la soluzione del problema parabolico, oltre ad avere uno shock a destra, come nel caso iperbolico, ha un nuovo salto a sinistra, nonostante il flusso sia convesso. Inoltre la velocità dei salti non è determinata solamente dal flusso, ma, in accordo con la Proposizione 3.12, anche il termine diffusivo darà un suo contributo.

3.3 Schemi espliciti per le leggi di conservazione

Ci dedicheremo ora all'approssimazione numerica delle soluzioni entropiche del problema di Cauchy (3.14)–(3.15). Ci restringeremo agli schemi espliciti ad un passo che, pur nella loro semplicità permettono uno studio approfondito delle problematiche essenziali che si incontrano in questo tipo di problemi.

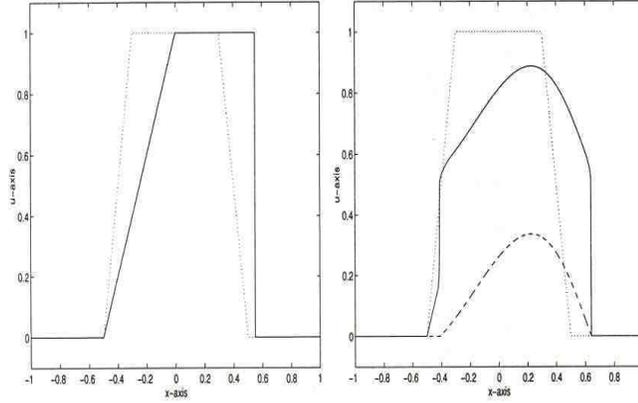


Figura 3.1: Sulla sinistra la soluzione dell'equazione di Burgers non viscosa. Sulla destra la soluzione dell'equazione (3.47) con la funzione di diffusione $B(u)$ tratteggiata. Il dato iniziale è il grafico puntinato.

3.3.1 Definizioni principali

Fissati un passo di discretizzazione spaziale $\Delta x = h$ e temporale $\Delta t = k$, la griglia sarà data dai punti nodali del piano $(x_m, t_n) = (mh, nk)$, per valori interi arbitrari m e n . Per una funzione v definita sulla griglia, scriveremo $v_{m,n}$ per il suo valore al nodo (x_m, t_n) , e questa notazione sarà estesa naturalmente anche a funzione definite ovunque per indicare i valori sui nodi. Porremo inoltre sistematicamente $\lambda = \frac{k}{h}$. Consideriamo allora lo schema esplicito generale a $(2l + 1)$ -punti spaziali e ad un passo

$$v_{m,n+1} = H(v_{m-l,n}, \dots, v_{m+l,n}), \quad (3.49)$$

per m , intero e $n \geq 0$, e dove H è una funzione continua, ma in generale non lineare, dei suoi $(2l + 1)$ argomenti, che in generale scriveremo come $H(v^n)$.

Una prima nozione importante da definire è quella degli schemi conservativi.

Definizione 3.5 *Uno schema del tipo (3.49) può essere messo in forma conservativa se esiste una funzione $g : \mathbb{R}^{2l} \rightarrow \mathbb{R}$ tale che*

$$H(v_{-l}, \dots, v_l) = v_0 - \lambda \{g(v_{-l+1}, \dots, v_l) - g(v_{-l}, \dots, v_{l-1})\}. \quad (3.50)$$

La funzione g è allora chiamata flusso numerico e lo schema (3.49) prende allora la forma

$$v_{m,n+1} = v_{m,n} - \lambda \{g(v_{m-l+1,n}, \dots, v_{m+l,n}) - g(v_{m-l,n}, \dots, v_{m+l-1,n})\}. \quad (3.51)$$

Nel seguito porremo per semplicità

$$g_{m+\frac{1}{2},n} = g(v_{m-l+1,n}, \dots, v_{m+l,n}),$$

e quindi la (3.51) si scriverà nella forma canonica

$$v_{m,n+1} = v_{m,n} - \lambda \{g_{m+\frac{1}{2},n} - g_{m-\frac{1}{2},n}\}. \quad (3.52)$$

Osserviamo che è possibile dimostrare che uno schema della forma (3.49) può essere messo in forma conservativa se e solo se per ogni vettore $v = (v_m)_{m \in \mathbb{Z}} \in L^1(\mathbb{Z})$ tale che $H(v) \in L^1(\mathbb{Z})$ si ha

$$\sum_{m \in \mathbb{Z}} H(v_{m-l}, \dots, v_{m+l}) = \sum_{m \in \mathbb{Z}} v_m. \quad (3.53)$$

Si osservi come la proprietà (3.53) sia l'equivalente discreto della proprietà di conservazione goduta dalle soluzioni entropiche, vedi Proposizione 3.11.

Chiaramente la proprietà di base dello schema (3.52) è però la sua consistenza con l'equazione (3.14). Poichè il rapporto

$$\frac{v_{m,n+1} - v_{m,n}}{k}$$

è una buona approssimazione della derivata temporale u_t e quindi dovremo chiedere che il rapporto

$$\frac{g_{m+\frac{1}{2},n} - g_{m-\frac{1}{2},n}}{h},$$

sia una buona approssimazione della derivata spaziale $A(u)_x$.

Definizione 3.6 *Lo schema conservativo (3.52) è detto consistente con l'equazione (3.14) se*

$$g(v, \dots, v) = A(v)$$

per ogni $v \in \mathbb{R}$ (a meno di una costante additiva).

A differenza di quanto fatto nella trattazione lineare, risulta qui più conveniente introdurre la funzione costante sui rettangoli della griglia numerica v_h , definita da

$$v_h(x, t) = v_{m,n}, \quad x_{m-\frac{1}{2}} < x < x_{m+\frac{1}{2}}, \quad t_n < t < t_{n+1}, \quad (3.54)$$

per $x_{m-\frac{1}{2}} := \frac{x_m + x_{m-1}}{2}$. Imporremo anche nel seguito che il dato iniziale sia discretizzato da

$$v_{m,0} = \frac{1}{h} \int_{x_{m-\frac{1}{2}}}^{x_{m+\frac{1}{2}}} u_0(x) dx. \quad (3.55)$$

Abbiamo allora il seguente Teorema di consistenza debole.

Teorema 3.14 (Lax–Wendroff) *Sia (3.52) uno schema conservativo consistente con l'equazione (3.14) e sia $v_{m,0}$ la discretizzazione del dato iniziale in (3.15), data da (3.55). Si assuma inoltre che esista una sottosuccessione h_l tendente a 0 tale che*

- $\|v_{h_l}\|_{L^\infty} \leq C$;

- v_{h_l} converge in $L^1_{loc}(\mathbb{R} \times (0, \infty))$ e quasi ovunque a una funzione u .

Allora u è una soluzione debole del problema (3.14)–(3.15).

Per una dimostrazione di questo celebre risultato vedi [LW, GR1]. Si osservi che uno schema che non sia in forma conservativa può dare nel limite soluzioni “spurie”, che non verificano l’equazione in senso debole, per esempio con delle discontinuità che non verificano la relazione di Rankine–Hugoniot.

Per meglio precisare la nostra analisi, lungo le linee già seguite nel caso lineare, affronteremo nel seguito i seguenti problemi:

- trovare condizioni sufficienti per la convergenza;
- trovare dei criteri che assicurino che il limite sia una soluzione entropica;
- determinare e migliorare l’ordine di accuratezza degli schemi proposti.

Il primo risultato è esattamente in quest’ultima direzione e si collega ai risultati della Sezione 1.8 del Capitolo 1.

Definizione 3.7 *Uno schema conservativo e consistente con l’equazione (3.14) si dice di ordine r se*

$$T(x, t) = \frac{1}{k} (u(x, t+k) - H(u(x-lh, t), \dots, u(x+lh, t))) = O(h^r),$$

se $h \rightarrow 0$. La quantità T è chiamata errore di troncatura dello schema.

Generalizzando al caso non lineare gli argomenti utilizzati nella Sezione 1.8, abbiamo il seguente risultato, che assicura che gli schemi consistenti sono almeno di ordine uno.

Proposizione 3.15 *Sia dato uno schema conservativo e consistente con l’equazione (3.14) tale che la funzione $H \in C^3$. Allora per ogni soluzione u di (3.14) abbastanza regolare l’errore di troncatura ha la seguente espressione*

$$T(x, t) = -k (\beta(u, \lambda) u_x)_x + O(k^2), \quad (3.56)$$

per

$$\beta(u, \lambda) = \frac{1}{2\lambda^2} \left(\sum_{m=-l}^l m^2 H_{v_m}(u, \dots, u) \right) - \frac{1}{2} (A'(u))^2. \quad (3.57)$$

Dimostrazione. Espandiamo con Taylor la derivata temporale della soluzione u

$$\begin{aligned} u(x, t+k) &= u + k u_t + \frac{1}{2} k^2 u_{tt} + O(k^3) \\ &= u - k A(u)_x + \frac{1}{2} k^2 ((A'(u))^2 u_x)_x + O(k^3). \end{aligned} \quad (3.58)$$

Per espandere il termine $H(u(x-lh, t), \dots, u(x+lh, t))$ abbiamo bisogno di alcune identità. Ovviamente

$$H(u, \dots, u) = u. \quad (3.59)$$

Inoltre, dalla (3.50), si ha, derivando e sommando,

$$\sum_{m=-l}^l m H_{v_m}(u, \dots, u) = -\lambda \sum_{m=-l}^l g_{v_m}(u, \dots, u) = -\lambda A'(u). \quad (3.60)$$

Procedendo analogamente per le derivate seconde si ottiene

$$\begin{aligned} H(u(x-lh, t), \dots, u(x+lh, t)) &= u - kA'(u)u_x \\ &+ \frac{1}{2}h^2 \left(\sum_{m=-l}^l m^2 H_{v_m}(u, \dots, u) \right) + O(h^3). \end{aligned} \quad (3.61)$$

Sommando la (3.58) e la (3.61) si ottiene il risultato. \square

In particolare osserviamo che uno schema può essere di ordine superiore al primo se e solo se

$$\beta(u, \lambda) \equiv 0,$$

per ogni soluzione u regolare.

3.3.2 Esempi di schemi a tre punti

Diamo adesso alcuni esempi di schemi a tre punti conservativi e consistenti. Molti di questi sono direttamente ispirati dal caso lineare che abbiamo studiato nel primo capitolo.

Esempio 3.16 Lo schema di Lax–Friedrichs. Come primo esempio consideriamo un'estensione al caso non lineare dello schema di Lax–Friedrichs. Come nel caso lineare la derivata spaziale $A(u)_x$ sarà approssimata da una differenza centrata, mentre la derivata temporale u_t verrà approssimata da

$$\frac{v_{m,n+1} - \frac{1}{2}(v_{m+1,n} + v_{m-1,n})}{k}.$$

Otteniamo quindi lo schema

$$\frac{v_{m,n+1} - \frac{1}{2}(v_{m+1,n} + v_{m-1,n})}{k} + \frac{A(v_{m+1,n}) - A(v_{m-1,n})}{2h} = 0, \quad (3.62)$$

che ponendo come al solito $\lambda = k/h$, può essere messo nella forma (3.52) con il flusso numerico

$$g^{LF}(u, v) = \frac{1}{2}(A(u) + A(v)) + \frac{1}{2\lambda}(u - v).$$

Essendo in questo caso

$$\beta(u, \lambda) = \frac{1}{2\lambda^2}(1 - (\lambda A'(u))^2),$$

abbiamo che lo schema è del prim'ordine. Naturalmente la stabilità linearizzata si avrà per

$$\lambda |A'(u)| \leq 1.$$

Possiamo scrivere anche lo schema in forma viscosa, analoga alla (1.34):

$$v_{m,n+1} = v_{m,n} - \frac{\lambda}{2}(A(v_{m+1,n}) - A(v_{m-1,n})) + \frac{1}{2}(v_{m+1,n} - 2v_{m,n} + v_{m-1,n}). \quad (3.63)$$

Nel seguito vedremo la successione generata da questo schema converge alla soluzione entropica. Purtroppo la sua alta viscosità è responsabile dell'eccessivo arrotondamento delle discontinuità. Nella figura 3.2 si vede a destra il dato iniziale e a sinistra la soluzione ad un tempo successivo per il problema di Riemann per (3.14)–(3.15) con $A(u) = \frac{1}{2}u^2$ e $u_s = 1$, $u_d = 0$.

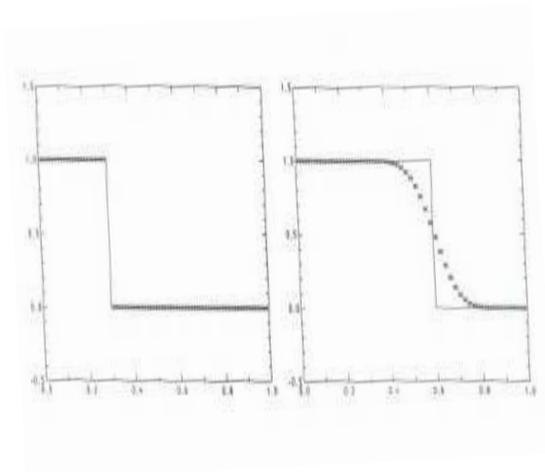


Figura 3.2: Lax–Friedrichs

Esempio 3.17 Lo schema upwind e lo schema di Engquist–Osher. Ricordando lo schema upwind nel caso lineare (1.10) e (1.11), abbiamo una naturale estensione al caso non lineare nel caso di una funzione A *monotona*:

$$v_{m,n+1} = \begin{cases} v_{m,n} - \lambda[A(v_{m,n}) - A(v_{m-1,n})], & \text{per } A' > 0, \\ v_{m,n} - \lambda[A(v_{m+1,n}) - A(v_{m,n})], & \text{per } A' < 0. \end{cases} \quad (3.64)$$

Si può vedere facilmente che lo schema upwind risulta essere consistente con l'equazione (3.14) e in generale avrà meno viscosità dello schema di Lax–Friedrichs. Per trattare il caso generale di flussi non monotoni l'estensione più semplice ed efficace risulta essere lo schema di Engquist–Osher. Si procede nel seguente modo. Siano

$$A^+(u) := \int_0^u \max\{A'(s), 0\} ds, \quad A^-(u) := \int_0^u \min\{A'(s), 0\} ds.$$

Allora, potendo sempre porre $A(0) = 0$ nell'equazione (3.14), si ha $A(u) = A^+(u) + A^-(u)$. Possiamo definire allora il flusso numerico

$$g^{EO}(u, v) = A^+(u) + A^-(v),$$

e lo schema corrispondente, scritto nella forma (3.52) sarà dato da

$$v_{m,n+1} = v_{m,n} - \lambda [A^+(v_{m,n}) + A^-(v_{m+1,n}) - A^+(v_{m-1,n}) - A^-(v_{m,n})]. \quad (3.65)$$

Se il flusso A è Lipschitziano, lo stesso vale per il flusso numerico g^{EO} che sarà anche consistente. La viscosità numerica dello schema è allora data da

$$\beta(u, \lambda) = \frac{|A'(u)|}{2\lambda} (1 - |\lambda A'(u)|),$$

e abbiamo che lo schema è ancora del prim'ordine. Nel seguito dimostreremo anche che converge alla soluzione entropica. La sua viscosità numerica è inferiore a quella dello schema di Lax–Friedrichs, come si vede anche dalla figura 3.3 dove si presenta la risoluzione dello stesso problema dell'esempio precedente.

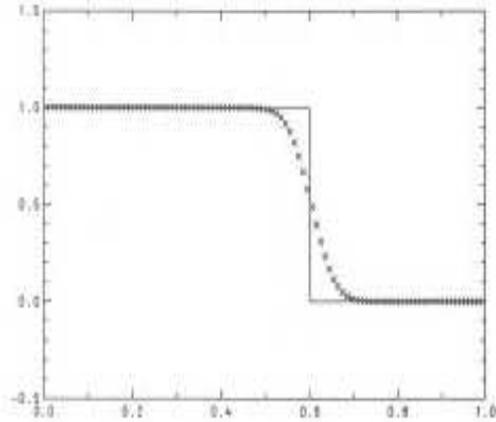


Figura 3.3: Engquist–Osher

Esempio 3.18 Lo schema di Godunov Questo schema è basato sulla soluzione locale del problema di Riemann. Ricordiamo che la soluzione esatta del problema (3.14)–(3.15) con il dato iniziale del tipo di Riemann

$$u_0 = \begin{cases} u_s, & \text{per } x < 0, \\ u_d, & \text{per } x > 0, \end{cases} \quad (3.66)$$

per due costanti fissate u_s e u_d , è sempre, anche nel caso non convesso, della forma autosimilare

$$u(x, t) = W(x/t; u_s, u_d),$$

dove W dipende da A e consiste di due stati costanti u_s e u_d separati da varie onde che partono dall'origine e che viaggiano ad una velocità inferiore al

$$\max\{|A'(s)|; s \text{ compreso tra } u_s \text{ e } u_d\}.$$

Definiamo la soluzione dello schema per ricorrenza, partendo da una soluzione v^n data al tempo $t_n = nk$.

Nella cella $(x_{m-1}, x_m) \times (t_n, t_{n+1})$ risolviamo esattamente il problema di Riemann per (3.14) con la condizione

$$u(x, t_n) = \begin{cases} v_{m-1}, & \text{per } x < x_{m-\frac{1}{2}}, \\ v_m, & \text{per } x > x_{m-\frac{1}{2}}, \end{cases} \quad (3.67)$$

vedi figura 3.4. Questa soluzione la chiamiamo $W(\frac{x-x_{m-\frac{1}{2}}}{k}; u_{m-1}, u_m)$. Dobbiamo assicurarci però che le soluzioni definite su due celle contigue non interagiscano prima del tempo t_{n+1} . Questo sarà sempre vero se è verificata la condizione di tipo CFL

$$\lambda \max\{|A'(s)|; s \text{ compreso tra } u_{m-1} \text{ e } u_m\} \leq \frac{1}{2}, \quad (3.68)$$

per tutti gli $m \in \mathbb{Z}$. Quindi possiamo definire un'unica funzione

$$v(x, t) = W\left(\frac{x - x_{m-\frac{1}{2}}}{k}; u_{m-1}, u_m\right), \text{ per } x \in (x_{m-1}, x_m)$$

nella striscia $\mathbb{R} \times (t_n, t_{n+1})$, data dalla sovrapposizione dei tanti problemi di Riemann con origine in x_m .

Passiamo ora al secondo passo nella costruzione dello schema, la proiezione di questa soluzione "esatta" su di una funzione costante a tratti. Definiamo sull'intervallo $(x_{m-\frac{1}{2}}, x_{m+\frac{1}{2}})$ la nuova v^{n+1} come

$$v_{m,n+1} = \frac{1}{h} \int_{x_{m-\frac{1}{2}}}^{x_{m+\frac{1}{2}}} v(x, t_{n+1}) dx.$$

A questo punto facciamo l'osservazione fondamentale che, essendo la v una soluzione esatta della legge di conservazione (3.14) nella cella $(x_{m-\frac{1}{2}}, x_{m+\frac{1}{2}}) \times (t_n, t_{n+1})$, dalla formula di Gauss–Green abbiamo la seguente espressione

$$\begin{aligned} \int_{x_{m-\frac{1}{2}}}^{x_{m+\frac{1}{2}}} v(x, t_{n+1}) dx &= \int_{x_{m-\frac{1}{2}}}^{x_{m+\frac{1}{2}}} v(x, t_n) dx - \int_{t_n}^{t_{n+1}} A(v(x_{m+\frac{1}{2}}, s)) ds \\ &\quad + \int_{t_n}^{t_{n+1}} A(v(x_{m-\frac{1}{2}}, s)) ds. \end{aligned}$$

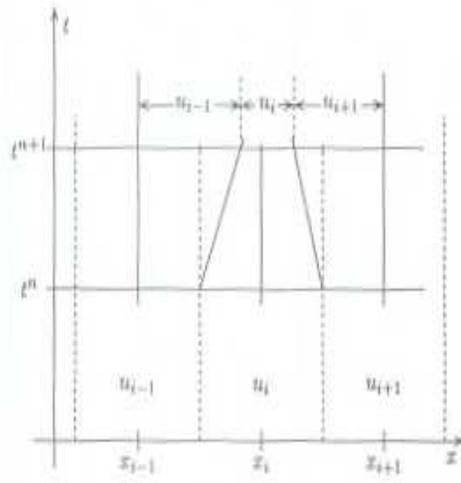


Figura 3.4: Il metodo di Godunov

Essendo la soluzione del problema di Riemann tale che il flusso è costante e continuo lungo la retta $x = x_{m \pm \frac{1}{2}}$, abbiamo il seguente schema numerico

$$v_{m,n+1} = v_{m,n} - \lambda \{A(W(0; u_m, u_{m+1})) - A(W(0; u_{m-1}, u_m))\}. \quad (3.69)$$

Questo schema è in forma conservativa con flusso numerico

$$g^G(u, v) = A(W(0; u, v)).$$

Nel caso particolare in cui il flusso A sia lineare, o semplicemente convesso nelle regioni di monotonia, allora lo schema di Godunov coincide semplicemente con lo schema upwind, come si vede anche dalla figura 3.5 dove si presenta la risoluzione dello stesso problema degli esempi precedenti. Altrimenti, si può lo stesso dare una semplice espressione del flusso numerico ponendo

$$g^G(u, v) = \begin{cases} \min_{w \in [u, v]} A(w) & \text{per } u \leq v, \\ \max_{w \in [v, u]} A(w) & \text{per } v \leq u. \end{cases}$$

Esempio 3.19 Lo schema di Lax–Wendroff È possibile dare una versione non lineare dello schema di Lax–Wendroff, già visto nel caso lineare. Procedendo nello sviluppo di Taylor in modo analogo a quanto fatto nella sezione 1.8 del Capitolo 1, si ha che se u è una soluzione regolare di (3.14), allora

$$u_{m,n+1} = u_{m,n} - kA(u)_x + \frac{1}{2}k^2 (A'(u)A(u)_x)_x + O(k^3).$$

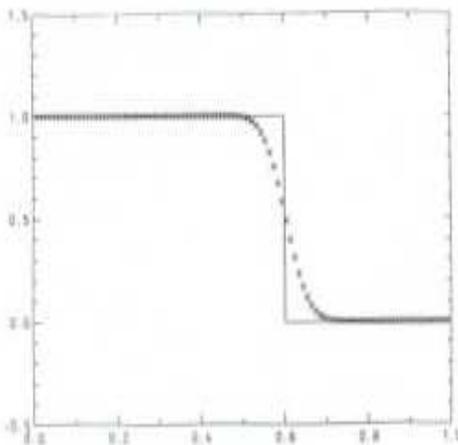


Figura 3.5: Godunov

Poichè

$$A(u)_x = \frac{1}{2h} (A(u_{m+1,n}) - A(u_{m-1,n})) + O(h^2),$$

e, per un qualche θ fissato nell'intervallo $(0, 1)$,

$$\begin{aligned} (A'(u)A(u)_x)_x &= \frac{1}{h^2} [A'(u(x_m + \theta h, t_n)) (A(u_{m+1,n}) - A(u_{m,n})) \\ &\quad - A'(u(x_m - (1 - \theta)h, t_n)) (A(u_{m,n}) - A(u_{m-1,n}))] + O(h). \end{aligned}$$

Troviamo quindi uno schema del secondo ordine della forma conservativa (3.52) con flusso numerico

$$g^{LW}(u, v) = \frac{1}{2}(A(u) + A(v)) - \frac{\lambda}{2} \mathcal{C}(u, v)(A(v) - A(u)), \quad (3.70)$$

dove $\mathcal{C}(u, v)$ è una qualsiasi approssimazione di A' data per esempio da una delle seguenti espressioni

$$\mathcal{C}(u, v) = A'\left(\frac{u+v}{2}\right),$$

oppure

$$\mathcal{C}(u, v) = \begin{cases} \frac{A(u)-A(v)}{u-v} & \text{per } u \neq v, \\ A'(u) & \text{per } u = v. \end{cases}$$

Nonostante si sia visto che questo schema è linearmente stabile di ordine 2, è possibile vedere che è instabile non linearmente vicino ai punti sonici, ossia in cui $A'(u) = 0$. In generale potrà inoltre produrre soluzioni che non rispettano

la condizione di entropia. Si prenda per esempio il solito flusso di Burgers $A(u) = \frac{1}{2}u^2$ e il dato iniziale di tipo Riemann

$$u_0 = \begin{cases} -1, & \text{per } x < 0, \\ 1, & \text{per } x \geq 0, \end{cases} \quad (3.71)$$

Allora l'unica soluzione entropica è di fatto continua (anche per il dato discontinuo, sic!) ed è data da

$$u(x, t) = \begin{cases} -1, & \text{per } x < -t, \\ \frac{x}{t}, & \text{per } -t < x < t, \\ 1, & \text{per } x > t, \end{cases}$$

Lo schema di Lax–Wendroff produce invece erroneamente la soluzione approssimata stazionaria $v_{m,n+1} = v_{m,n}$.

Infine se risolviamo lo stesso problema test che abbiamo visto negli esempi precedenti, vedi la figura 3.6, si osservano delle oscillazioni intorno alla discontinuità, nonostante questa sia risolta su un numero inferiore di punti rispetto a tutti gli schemi precedenti. Questi fenomeni sono entrambi dovuti alla scarsa viscosità numerica di questo schema.

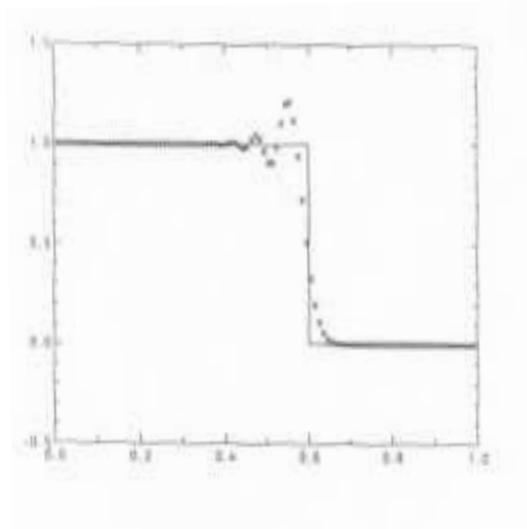


Figura 3.6: Lax–Wendroff

3.4 Analisi della convergenza

3.4.1 Schemi TVD e monotoni

Ci occuperemo ora del problema della convergenza di questi schemi in una opportuna topologia. Poichè le soluzioni limite non saranno in generale continue, dobbiamo aspettarci di avere delle stime molto deboli per passare al limite. L'idea più importante che verrà usata nel seguito è quella di utilizzare esclusivamente le stime di compattezza in L^1 ed in particolare il principio di compattezza di Helly, che assicura che gli insiemi limitati nello spazio BV delle funzioni a variazione limitata risultano precompatti nella topologia di L^1 , vedi per esempio [Gi, Br]. D'altra parte dalla Proposizione 3.11 sappiamo che questo tipo di stime valgono per le soluzioni entropiche, e sono quindi abbastanza naturali anche per gli schemi numerici. Accanto alle norme discrete introdotte nella sezione 3 del Capitolo 1, introduciamo la variazione totale della successione (v_m) :

$$TV(v) = \sum_{m \in \mathbb{Z}} |v_{m+1} - v_m|. \quad (3.72)$$

Definizione 3.8 *Uno schema alle differenze finite della forma (3.52) si dice L^∞ -stabile se esiste una costante $C_1 > 0$ indipendente da n e k tale che*

$$\|v^n\|_{h,\infty} \leq C_1,$$

per ogni $n \geq 0$.

*Uno schema è detto essere a **Variazione Totale limitata** se esiste una costante $C_2 > 0$ indipendente da n e k tale che*

$$TV(v^n) \leq C_2,$$

per ogni $n \geq 0$.

In particolare una classe molto importante di schemi che sono L^∞ -stabili e a variazione totale limitata è dato dalla classe degli schemi **TVD** (Total Variation Diminishing) che sono gli schemi tali che

$$TV(v^{n+1}) \leq TV(v^n), \quad (3.73)$$

per ogni $n \geq 0$. Si può dimostrare che questi schemi non creano nuovi estremi nella soluzione approssimata, impedendo cioè il crearsi di nuove oscillazioni.

Abbiamo allora il seguente risultato dovuto ad Harten [Ha].

Teorema 3.20 Convergenza degli schemi TVD *Sia dato uno schema conservativo nella forma (3.52) con flusso numerico g localmente Lipschitziano. Sia inoltre lo schema TVD e L^∞ -stabile. Sia $u_0 \in BV(\mathbb{R})$. Allora esiste una soluzione debole u di (3.14)–(3.15) e una sottosuccessione h_l , tale che, tenendo $k_l = \lambda h_l$ con λ costante, la successione v_{h_l} converge a u in $L^\infty(0, T; L^1_{loc}(\mathbb{R}))$ per ogni $T > 0$.*

La dimostrazione richiede un minimo di conoscenze di Analisi Funzionale ed è ben descritta in [GR1]. Le condizioni di entropia per queste soluzioni limite verranno discusse nella prossima sezione.

La classe più semplice di schemi TVD è data dagli schemi **monotoni**

Definizione 3.9 *Uno schema numerico della forma (3.49) è detto monotono se la funzione H è monotona non decrescente in tutti i suoi argomenti.*

Abbiamo il seguente importante risultato.

Proposizione 3.21 *Uno schema monotono che può essere messo nella forma conservativa (3.52) è TVD e L^∞ -stabile.*

Dimostrazione. La stabilità per la norma L^∞ segue come nel Teorema 1.10. Per dimostrare la (3.73), poniamo per due numeri reali α e β

$$\alpha \vee \beta := \max(\alpha, \beta); \quad \alpha \wedge \beta := \min(\alpha, \beta).$$

Abbiamo che, dati due vettori $v = (v_m)$ e $w = (w_m)$, poichè

$$|v_m - w_m| = v_m \vee w_m - v_m \wedge w_m,$$

allora dalla proprietà di conservazione (3.53)

$$\sum_{m \in \mathbb{Z}} |v_m - w_m| = \sum_{m \in \mathbb{Z}} H(v_m \vee w_m) - H(v_m \wedge w_m).$$

D'altra parte dalla monotonia abbiamo che

$$H(v_m \vee w_m) \geq H(v_m) \vee H(w_m),$$

e

$$H(v_m \wedge w_m) \leq H(v_m) \wedge H(w_m).$$

Ne segue quindi

$$\sum_{m \in \mathbb{Z}} |v_m - w_m| \geq \sum_{m \in \mathbb{Z}} |H(v_m) - H(w_m)|. \quad (3.74)$$

Prendendo $w_m = v_{m+1}$ otteniamo la proprietà TVD. \square

Osserviamo che negli esempi precedenti, gli schemi di Lax–Friedrichs e di Engquist–Osher risultano essere monotoni sotto la condizione di CFL

$$\lambda \max |A'(v)| \leq 1.$$

Il metodo di Godunov è monotono poichè risulta essere la composizione di due passi monotoni (evoluzione+media integrale) e verifica infatti una condizione tipo CFL, la (3.68). Invece lo schema di Lax–Wendroff non è monotono (di fatto nemmeno TVD), come si vede già nel caso lineare.

In effetti la classe degli schemi monotoni, come già osservato nella sezione 8, Proposizione 1.16 non produce schemi sufficientemente accurati. Abbiamo in effetti il seguente risultato.

Proposizione 3.22 *Uno schema conservativo, consistente e monotono, tale che la funzione H sia C^3 , è al più del prim'ordine.*

Facciamo di seguito due osservazioni che saranno utili nella ricerca di schemi TVD di alto ordine.

- Si può dimostrare che nel caso di uno schema lineare per un'equazione lineare, ossia uno schema della forma (1.27), la condizione di essere TVD equivale alla monotonia. Quindi anche nel caso lineare dovremmo trovare degli schemi che siano non lineari.
- In effetti anche nel caso di schemi non lineari TVD a tre punti abbiamo che, se i coefficienti sono regolari, allora lo schema è ancora al più del prim'ordine. Bisognerà quindi ridurre la regolarità dei coefficienti e aumentare il numero di punti (almeno 5).

3.4.2 Forma incrementale e viscosità numerica

Per poter meglio studiare gli schemi di tipo TVD introduciamo due forme alternative per gli schemi conservativi.

Definizione 3.10 *Uno schema del tipo (3.52) può essere messo in forma viscosa se esiste una funzione Q di $2l$ variabili, chiamata “coefficiente di viscosità numerica”, tale che, posto*

$$\Delta_{m+\frac{1}{2}}v := v_{m+1} - v_m,$$

e

$$Q_{m+\frac{1}{2},n} := Q(v_{m-l+1,n}, \dots, v_{m+l,n}),$$

lo schema può essere riscritto nella forma

$$\begin{aligned} v_{m,n+1} &= v_{m,n} - \frac{\lambda}{2} (A(v_{m+1,n}) - A(v_{m-1,n})) \\ &\quad + \frac{1}{2} \left\{ Q_{m+\frac{1}{2},n} \Delta_{m+\frac{1}{2}}v^n - Q_{m-\frac{1}{2},n} \Delta_{m-\frac{1}{2}}v^n \right\}. \end{aligned} \quad (3.75)$$

Osserviamo che nel caso lineare si ritrova la forma (1.34) con $Q_{m+\frac{1}{2},n} = q$. Inoltre si osserva che dato uno schema nella forma (3.75), allora si può sempre mettere in forma conservativa con flusso numerico

$$g_{m+\frac{1}{2},n} = \frac{1}{2} (A(v_{m+1}) + A(v_m)) - \frac{1}{2\lambda} Q_{m+\frac{1}{2},n} \Delta_{m+\frac{1}{2}}v.$$

Un'altra forma utile per verificare se uno schema è TVD è la forma *incrementale*.

Definizione 3.11 *Uno schema del tipo (3.52) può essere messo in forma incrementale se esistono due funzioni di $2l$ variabili C e D , chiamati coefficienti incrementali, tali che se poniamo*

$$C_{m+\frac{1}{2},n} := C(v_{m-l+1,n}, \dots, v_{m+l,n}), D_{m+\frac{1}{2},n} := D(v_{m-l+1,n}, \dots, v_{m+l,n}),$$

allora lo schema si scrive come

$$v_{m,n+1} = v_{m,n} + C_{m+\frac{1}{2},n} \Delta_{m+\frac{1}{2}} v^n - D_{m-\frac{1}{2},n} \Delta_{m-\frac{1}{2}} v^n. \quad (3.76)$$

Si osservi che se lo schema è in forma viscosa, allora si può sempre mettere in forma incrementale ponendo

$$\begin{aligned} C_{m+\frac{1}{2}} &= \frac{1}{2} \left\{ -\lambda \frac{A(v_{m+1}) - A(v_m)}{\Delta_{m+\frac{1}{2}} v} + Q_{m+\frac{1}{2}} \right\}, \\ D_{m+\frac{1}{2}} &= \frac{1}{2} \left\{ \lambda \frac{A(v_m) - A(v_{m-1})}{\Delta_{m-\frac{1}{2}} v} + Q_{m-\frac{1}{2}} \right\}. \end{aligned} \quad (3.77)$$

Il viceversa non è sempre vero, ma si può dimostrare che almeno nel caso degli schemi a tre punti tutte queste forme sono equivalenti ponendo allora per ottenere la forma viscosa

$$Q_{m+\frac{1}{2}} = \lambda \frac{A(v_{m+1}) + A(v_m) - 2g_{m+\frac{1}{2},n}}{\Delta_{m-\frac{1}{2}} v},$$

e per quella incrementale

$$C_{m+\frac{1}{2}} = \lambda \frac{A(v_m) - g_{m+\frac{1}{2},n}}{\Delta_{m-\frac{1}{2}} v}, \quad D_{m+\frac{1}{2}} = \lambda \frac{A(v_{m+1}) - g_{m+\frac{1}{2},n}}{\Delta_{m-\frac{1}{2}} v}.$$

Abbiamo allora la seguente utile caratterizzazione degli schemi TVD dovuta ad Harten [Ha].

Proposizione 3.23 Caratterizzazione degli schemi TVD

1. Sia dato uno schema conservativo e consistente scritto nella forma viscosa (3.75). Se la viscosità numerica soddisfa

$$\lambda \left| \frac{A(v_{m+1}) - A(v_m)}{\Delta_{m+\frac{1}{2}} v} \right| \leq Q_{m+\frac{1}{2}} \leq 1, \quad (3.78)$$

per ogni $m \in \mathbb{Z}$, allora lo schema è TVD.

2. Sia dato uno schema conservativo e consistente scritto nella forma incrementale (3.76). Se i coefficienti incrementali soddisfano le disuguaglianze

$$C_{m+\frac{1}{2}}, D_{m+\frac{1}{2}} \geq 0, \quad C_{m+\frac{1}{2}} + D_{m+\frac{1}{2}} \leq 1, \quad (3.79)$$

per ogni $m \in \mathbb{Z}$, allora lo schema è TVD.

Dimostrazione. Ci limiteremo a dimostrare soltanto la seconda parte della Proposizione, poichè gli schemi in forma viscosa si possono sempre mettere in forma incrementale utilizzando la (3.77) e la (3.78) segue facilmente dalla (3.79).

Poniamo dunque

$$w = H(v)$$

. Dalla (3.76) segue che

$$\begin{aligned} \Delta_{m+\frac{1}{2}} w &= C_{m+\frac{3}{2}} \Delta_{m+\frac{3}{2}} v \\ &+ (1 - C_{m+\frac{1}{2}} - D_{m+\frac{1}{2}}) \Delta_{m+\frac{1}{2}} v + D_{m-\frac{1}{2}} \Delta_{m-\frac{1}{2}} v. \end{aligned}$$

Poichè dalla (3.79) i coefficienti a destra sono tutti positivi, passando ai moduli e sommando su m otteniamo

$$TV(w) = \sum_m |\Delta_{m+\frac{1}{2}} w| \leq \sum_m |\Delta_{m+\frac{1}{2}} v| = TV(v),$$

che conclude la dimostrazione \square

In particolare per gli schemi a tre punti queste condizioni sono anche necessarie perchè lo schema sia TVD.

Negli esempi precedenti è facile calcolare i coefficienti di viscosità e incrementali. Per esempio lo schema di Lax–Friedrichs è quello con la più alta viscosità numerica ammissibile, essendo

$$Q_{m+\frac{1}{2}}^{LF} = 1.$$

Per lo schema di Engquist–Osher abbiamo

$$Q_{m+\frac{1}{2}}^{EO} = \lambda \frac{1}{\Delta_{m+\frac{1}{2}} v} \int_{v_m}^{v_{m+1}} |A'(s)| ds,$$

e per Godunov

$$Q_{m+\frac{1}{2}}^G = \lambda \max_{s \in I(v_m, v_{m+1})} \left[\frac{A(v_{m+1}) - 2A(s) + A(v_m)}{\Delta_{m+\frac{1}{2}} v} \right].$$

I coefficienti di viscosità di questi due ultimi schemi sono praticamente gli stessi, e minimizzano la disuguaglianza (3.78), al di fuori dei punti sonici.

Se invece calcoliamo la viscosità numerica dello schema di Lax–Wendroff troviamo

$$Q_{m+\frac{1}{2}}^{LW} = \lambda^2 \left(\frac{A(v_{m+1}) - A(v_m)}{\Delta_{m+\frac{1}{2}} v} \right)^2$$

che chiaramente non soddisfa la (3.78). Lo schema tende quindi a creare oscillazioni. In effetti si può dimostrare che tutti gli schemi che si possono scrivere in forma viscosa e che soddisfano la (3.78) sono al più di ordine 1. Un modo per evitare questo sarà quello di ricorrere a delle forme viscosive modificate, o ad usare la forma incrementale per uno schema non a tre punti.

3.4.3 La condizione di entropia

Sia η un'entropia convessa per l'equazione (3.14), con flusso di entropia q .

Definizione 3.12 *Uno schema è consistente con la disuguaglianza di entropia (3.29), e viene detto allora uno "schema entropico", se esiste una funzione continua G di $2l$ variabili tale che:*

1. vale la consistenza con il flusso di entropia q ,

$$G(u, \dots, u) = q(u),$$

2. vale la disuguaglianza di entropia discreta

$$\begin{aligned} \eta(v_{m,n+1}) \leq \eta(v_{m,n}) & - \lambda \{ G(v_{m-l+1,n}, \dots, v_{m+l,n}) \\ & - G(v_{m-l,n}, \dots, v_{m+l-1,n}) \}. \end{aligned} \quad (3.80)$$

Vale allora un risultato analogo al teorema di Lax–Wendroff.

Teorema 3.24 *Se uno schema conservativo e consistente è anche entropico per tutte le entropie convesse per l'equazione (3.14) e la successione di soluzioni approssimanti converge ad una funzione limite u , allora questo limite è l'unica soluzione entropica del problema (3.14)–(3.15).*

Abbiamo allora il seguente risultato, che assicura che almeno una larga classe degli schemi conosciuti converge di fatto alla soluzione entropica.

Teorema 3.25 *Gli schemi monotoni sono entropici.*

Anche qui abbiamo un risultato negativo: tutti gli schemi a tre punti con flusso numerico C^2 che siano consistenti con la condizione di entropia sono al più del primo ordine.

Osserviamo infine che a causa della condizione di entropia non possiamo fare la scelta minimale nella condizione (3.78) di porre, per esempio per uno schema a tre punti,

$$Q(u, v) = \lambda \left| \frac{A(u) - A(v)}{u - v} \right|,$$

che corrisponde al flusso numerico

$$g(u, v) = \frac{1}{2}(A(u) + A(v)) - \frac{1}{2\lambda}Q(u, v)(v - u).$$

Questo schema, chiamato schema di Murman–Roe, risulta chiaramente essere TVD, ma purtroppo viola la condizione di entropia proprio nei punti sonici. Si prenda ancora il flusso di Burgers $A(u) = \frac{1}{2}u^2$ e il dato iniziale di tipo Riemann (3.71). Allora, procedendo come per lo schema di Lax–Wendroff, la soluzione approssimata risulta essere uno shock stazionario, ben diverso dall'onda di rarefazione che risolve il problema. Questo fenomeno è dovuto

alla degenerazione della viscosità numerica per $u = 0$. Si osserva infatti che il coefficiente (3.57) dato dallo sviluppo di Taylor in questo caso vale

$$\begin{aligned}\beta(u, \lambda) &= \frac{1}{2\lambda} (Q(u, u) - (\lambda A'(u))^2) \\ &= \frac{1}{2} |A'(u)| (1 - \lambda A'(u)),\end{aligned}$$

che si annulla per $u = 0$.

3.5 Schemi TVD del secondo ordine

I risultati della sezione precedente mostrano chiaramente la difficoltà di costruire degli schemi numerici che siano al tempo stesso TVD e di ordine 2. A partire dalla fine degli anni 70 questo è stato uno degli obiettivi dell'analisi numerica dei problemi iperbolici non lineari e numerosi risultati sono stati raggiunti con diversi metodi. Qui di seguito ci limiteremo a presentare l'approccio dei limitatori di flusso di Sweby e Davis, che consiste essenzialmente nel convertire uno schema a tre punti del primo ordine in uno schema a cinque punti del secondo modificando opportunamente, e in modo non lineare, i flussi numerici. Per una rassegna di altri metodi ed una trattazione più dettagliata di tutto l'argomento si rinvia comunque a [GR1, GR2, Le, Ta].

In ogni caso una caratteristica comune di tutti gli schemi TVD e anche di ordine più elevato (PPM, ENO) è la procedura di limitazione che serve ad impedire il crearsi di oscillazioni. Questo comporta che la migliore accuratezza raggiunta nelle regioni di regolarità della soluzione sarà eventualmente persa nei punti di estremo.

Inoltre è bene osservare che non esiste nessuno schema TVD del secondo ordine con coefficienti incrementali regolari, che soddisfi la condizione (sufficiente) (3.79). Infatti, se uno schema è scritto nella forma incrementale (3.76), si ha che

$$\begin{aligned}H_{-1}(u, \dots, u) &= D(u, \dots, u), \\ H_1(u, \dots, u) &= C(u, \dots, u), \\ H_m(u, \dots, u) &= 0, \text{ se } m \neq -1, 0, 1.\end{aligned}$$

Quindi, dalla (3.57)

$$\beta(u, \lambda) = \frac{1}{2\lambda^2} (C(u, \dots, u) + D(u, \dots, u) - (\lambda A'(u))^2).$$

Se lo schema è TVD allora

$$C(u, \dots, u), D(u, \dots, u) \geq 0, \quad C(u, \dots, u) + D(u, \dots, u) \leq 1.$$

Inoltre dalla consistenza

$$C(u, \dots, u) - D(u, \dots, u) = -\lambda A'(u).$$

Quindi

$$(\lambda A'(u))^2 = (C(u, \dots, u) - D(u, \dots, u))^2 \leq C(u, \dots, u) + D(u, \dots, u),$$

ossia

$$\beta(u, \lambda) \geq 0,$$

con l'uguaglianza valida solo nel caso banale in cui uno dei due coefficienti valga 1 e l'altro 0. Quindi lo schema sarà in generale solo del primo ordine.

Saremo quindi costretti ad utilizzare dei coefficienti incrementali non regolari. Questo implica che otterremo degli schemi del secondo ordine "formale", richiedendo che il flusso dello schema in questione sia una modifica del secondo ordine dello schema di Lax–Wendroff. Per esempio nell'approccio con i limitatori di flusso chiederemo allo schema di verificare

$$H = H^{LW} + O(h^3).$$

Tuttavia si può dimostrare che nei punti critici non sonici (ossia tali che $A' \neq 0$) uno schema TVD è al più del primo ordine.

Per presentare la costruzione di Sweby, che unifica diversi schemi precedentemente proposti, partiamo da uno schema a tre punti scritto nella forma viscosa

$$\begin{aligned} v_{m,n+1} &= H^3(v^n) = v_{m,n} - \frac{\lambda}{2} (A(v_{m+1,n}) - A(v_{m-1,n})) \\ &\quad + \frac{1}{2} \left\{ Q_{m+\frac{1}{2},n} \Delta_{m+\frac{1}{2}} v^n - Q_{m-\frac{1}{2},n} \Delta_{m-\frac{1}{2}} v^n \right\}. \end{aligned}$$

con flusso numerico

$$g_{m+\frac{1}{2},n}^3 = \frac{1}{2} (A(v_{m+1}) + A(v_m)) - \frac{1}{2\lambda} Q_{m+\frac{1}{2},n} \Delta_{m+\frac{1}{2}} v.$$

Nel confronto con lo schema di Lax–Wendroff otteniamo che

$$g_{m+\frac{1}{2}}^{LW} - g_{m+\frac{1}{2},n}^3 = \frac{1}{2\lambda} (Q_{m+\frac{1}{2}} - \lambda^2 C_{m+\frac{1}{2}}^2) \Delta_{m+\frac{1}{2}} v.$$

Quindi il metodo di Lax–Wendroff si può scrivere come

$$\begin{aligned} H^{LW}(v^n) &= H^3(v^n) - \frac{1}{2} \left[(Q_{m+\frac{1}{2}} - \lambda^2 C_{m+\frac{1}{2}}^2) \Delta_{m+\frac{1}{2}} v \right. \\ &\quad \left. - (Q_{m-\frac{1}{2}} - \lambda^2 C_{m-\frac{1}{2}}^2) \Delta_{m-\frac{1}{2}} v \right]. \end{aligned}$$

Poichè dalla (3.78), si ha che

$$Q_{m+\frac{1}{2}} - \lambda^2 C_{m+\frac{1}{2}}^2 \geq 0,$$

si può pensare al metodo di Lax–Wendroff come ad un metodo a tre punti TVD "meno" una parte viscosa. L'idea di Sweby è semplicemente quella di utilizzare il metodo di Lax–Wendroff nelle regioni in cui non ci siano oscillazioni, per poi passare al metodo a tre punti nelle regioni con forti oscillazioni. Questo può essere fatto introducendo una funzione di "switch" che controlli le oscillazioni.

Qui di seguito ci limiteremo a presentare questa costruzione nel caso lineare, rinviando a [Sw], vedi anche [GR1], per il caso non lineare.

Si consideri quindi il caso dell'equazione (1.1) con $a > 0$ e come schema a tre punti lo schema upwind che si scrive, ponendo $\nu = \lambda a \in (0, 1)$ ed ignorando sistematicamente l'indice temporale n ,

$$H^{up}(v) = v_m - \nu \Delta_{m-\frac{1}{2}} v.$$

Analogamente lo schema di Lax–Wendroff si riduce a

$$H^{LW}(v) = H^{up}(v) - \frac{1}{2} \nu (1 - \nu) \left(\Delta_{m+\frac{1}{2}} v - \Delta_{m-\frac{1}{2}} v \right).$$

Come prima l'ultimo termine sulla destra può essere interpretato come un'antidiffusione che è responsabile della creazione di oscillazioni. Per assicurarci la proprietà TVD consideriamo uno schema numerico modificato

$$H^S(v) = H^{up}(v) - \frac{1}{2} \nu (1 - \nu) \left(\phi_{m+\frac{1}{2}} \Delta_{m+\frac{1}{2}} v - \phi_{m-\frac{1}{2}} \Delta_{m-\frac{1}{2}} v \right), \quad (3.81)$$

dove la funzione ϕ , detta il limitatore, è una funzione continua del rapporto tra due gradienti successivi:

$$\phi_{m+\frac{1}{2}} := \phi(r_{m+\frac{1}{2}}), \quad r_{m+\frac{1}{2}} := \frac{\Delta_{m+\frac{1}{2}} v}{\Delta_{m-\frac{1}{2}} v}.$$

Si osservi che nei punti in cui $r_{m+\frac{1}{2}}$ o $r_{m-\frac{1}{2}}$ si annullano i coefficienti incrementali dello schema possono diventare singolari.

In pratica se ci sono oscillazioni, nel punto di estremo locale avremo che $r_{m+\frac{1}{2}} < 0$, poichè $\Delta_{m+\frac{1}{2}} v$ e $\Delta_{m-\frac{1}{2}} v$ avranno segni opposti. Imponendo $\phi = 0$, lo schema diventerà monotono e eliminerà l'oscillazione. Nelle altre regioni potremo avere ϕ positivo, cercando così di eliminare la viscosità ed al tempo stesso evitare le oscillazioni. Esaminiamo per prima cosa una condizione sufficiente sulla funzione ϕ che garantisca la proprietà TVD.

Teorema 3.26 *Sia verificata la condizione di CFL, $0 \leq \nu \leq 1$. Supponiamo che la funzione $\phi = \phi(r)$ sia uguale a 0 per $r \leq 0$ e sia*

$$0 \leq \phi(r) \leq \min \left(\frac{2r}{\nu}, \frac{2}{1-\nu} \right),$$

per $r \geq 0$. Allora lo schema (3.81) è TVD e L^∞ -stabile.

Dimostrazione. Scriviamo lo schema (3.81) in forma incrementale (3.76) scegliendo

$$C_{m+\frac{1}{2}} = 0$$

$$D_{m-\frac{1}{2}} = \nu \left\{ 1 + \frac{1-\nu}{2} \left(\frac{\phi(r_{m+\frac{1}{2}})}{r_{m+\frac{1}{2}}} - \phi(r_{m-\frac{1}{2}}) \right) \right\}.$$

Allora i coefficienti incrementali devono verificare la (3.79) della Proposizione 3.23, che in questo caso diventa semplicemente che $0 \leq D_{m-\frac{1}{2}} \leq 1$. Essendo la funzione ϕ positiva, ma eventualmente anche nulla, questo equivale a chiedere che, per $r \geq 0$

$$1 + \frac{1-\nu}{2} \frac{\phi(r)}{r} \leq \frac{1}{\nu}$$

e

$$0 \leq 1 - \frac{1-\nu}{2} \phi(r),$$

ossia proprio le ipotesi del Teorema. \square

Questo risultato ci fa vedere come sia possibile scegliere una funzione ϕ positiva e quindi, lontano dalle oscillazioni, passare allo schema di Lax–Wendroff. In effetti quello che si vede è che lo schema di Lax–Wendroff, che corrisponde a $\phi = 1$, non dà nessun problema se i rapporti incrementali hanno segni concordi e quello di sinistra è in valore assoluto più grande di quello di destra. Per esempio per una soluzione decrescente si avranno oscillazioni all’inizio del salto, ma non alla fine, come si vede bene nella figura 3.6.

Cerchiamo adesso di trovare una condizione sufficiente sulla ϕ affinché lo schema sia di ordine 2.

Teorema 3.27 *Sia la funzione ϕ Lipschitziana e tale che per $r = 1$ esistano i limiti destro e sinistro della derivata. Allora, se $\phi(1) = 1$ lo schema (3.81) è del secondo ordine nelle regioni di regolarità della soluzione e al di fuori dei punti critici (ossia tali che $u_x = 0$).*

Dimostrazione. Dal confronto con lo schema di Lax–Wendroff basta far vedere che

$$\begin{aligned} R &= (1 - \phi(r_{m+\frac{1}{2}}))(u(x+h, t) - u(x, t)) \\ &\quad - (1 - \phi(r_{m-\frac{1}{2}}))(u(x, t) - u(x-h, t)) = O(h^3). \end{aligned}$$

Supponiamo che la soluzione $u \in C^2$ e che in un intorno del punto (x, t) valga $u_x \neq 0$. Allora abbiamo i seguenti sviluppi:

$$\begin{aligned} u(x+h, t) - u(x, t) &= hu_x + \frac{1}{2}h^2u_{xx} + O(h^3); \\ u(x, t) - u(x-h, t) &= hu_x - \frac{1}{2}h^2u_{xx} + O(h^3); \\ r_{m+\frac{1}{2}}, r_{m-\frac{1}{2}} &= 1 - h\frac{u_{xx}}{u_x} + O(h^3). \end{aligned}$$

Abbiamo allora

$$R = (\phi'(1\pm)h^2\frac{u_{xx}}{u_x}u_x + O(h^3)) - (\phi'(1\pm)h^2\frac{u_{xx}}{u_x}u_x + O(h^3)),$$

avendo indicato con $\phi'(1\pm)$ i limiti destri e sinistri rispettivi (che dipendono in effetti solo dal segno del rapporto $\frac{u_{xx}}{u_x}$), della derivata prima di ϕ . Questo dimostra il risultato. \square

Nel proporre alcuni esempi di schemi derivati utilizzando il limitatore di flusso osserviamo innanzitutto che per $\phi(r) = r$ si ottiene lo schema a cinque punti di Warming e Beam che può facilmente essere messo nella forma

$$H^{WB}(v) = v_m - \nu \Delta_{m-\frac{1}{2}} v - \frac{1}{2} \nu (1 - \nu) \left(\Delta_{m-\frac{1}{2}} v - \Delta_{m-\frac{3}{2}} v \right),$$

ed è del secondo ordine grazie al Teorema 3.27, ma non verifica le condizioni TVD del Teorema 3.26 per $r > \frac{2}{1-\nu}$. Possiamo quindi pensare di ottenere il nostro schema modificato facendo un'opportuna media dello schema di Lax-Wendroff e di quello di Warming-Beam. In particolare possiamo cercare una ϕ come combinazione lineare della $\phi^{LW} \equiv 1$ e della $\phi^{WB} \equiv r$, prendendo, sempre per $r > 0$,

$$\phi(r) = (1 - \theta(r)) + \theta(r)r = 1 + \theta(r)(r - 1),$$

per una certa funzione Lipschitziana θ . Notiamo che in effetti questo si può in principio sempre fare per ogni ϕ verificante le ipotesi del Teorema 3.27. Seguendo Sweby [Sw], ci limiteremo a prendere solamente delle combinazioni convesse dei due flussi, ossia $\theta \in [0, 1]$, essendo gli altri casi eccessivamente compressivi. Troviamo allora che, utilizzando anche le limitazioni del Teorema 3.26, è sufficiente prendere, per $r > 0$, una funzione ϕ monotona tale che

$$\begin{cases} r \leq \phi(r) \leq \frac{2r}{\nu}, & \text{per } 0 \leq r \leq 1, \\ 1 \leq \phi(r) \leq r, & \text{per } 1 \leq r \leq \frac{2}{1-\nu} \\ 1 \leq \phi(r) \leq \frac{2}{1-\nu}, & \text{per } r > \frac{2}{1-\nu}. \end{cases} \quad (3.82)$$

Un'altra richiesta, che evita delle dissimmetrie nel trattamento delle discontinuità, è quella di chiedere che valga

$$\phi(r) = r\phi\left(\frac{1}{r}\right). \quad (3.83)$$

Ci sono molti esempi di limitatori di flusso che soddisfano queste condizioni. In particolare ricordiamo i seguenti casi, per $r > 0$:

$$\begin{aligned} \phi_1(r) &:= \max\{0, \min\{r, 1\}\} && (\text{minmod}) \\ \phi_2(r) &:= \max\{0, \min\{\frac{2r}{\nu}, 1\}, \min\{r, \frac{2}{1-\nu}\}\} && (\text{superbee di Roe}) \\ \phi_{VL} &:= \frac{2r}{1+r} && (\text{van Leer}) \\ \phi_{VA} &:= \frac{r^2+r}{1+r^2} && (\text{van Albada}) \\ \phi_{CO} &:= \max\{0, \min\{r, \beta\}\}, \quad 1 \leq \beta \leq \frac{2}{1-\nu}, && (\text{Chakravarthy e Osher}) \\ \phi_{AR} &:= \max\{0, \min\{\frac{2r}{\nu}, 1 + \frac{(1+\nu)}{3}(r-1), \frac{2}{1-\nu}\}\} && (\text{Arora e Roe}). \end{aligned}$$

Un esempio di evoluzione con il limitatore *minmod* si vede (nel caso di Burgers, con lo stesso esempio della sottosezione 3.3.2) nella figura 3.7.

Cochudiamo facendo osservare che questi risultati non tengono ovviamente conto delle condizioni di entropia, che però possono essere risolte cercando di aumentare la diffusione nelle zone soniche espansive. Inoltre il peggioramento degli schemi nei punti critici non sonici può essere in qualche modo migliorato ricorrendo ai cosiddetti schemi TVB (Total Variation Bounded) di Shu, [Sh], che in pratica modificano il flusso numerico solo nelle vicinanze di questi punti grazie all'aggiunta di un fattore dell'ordine di h^2 .

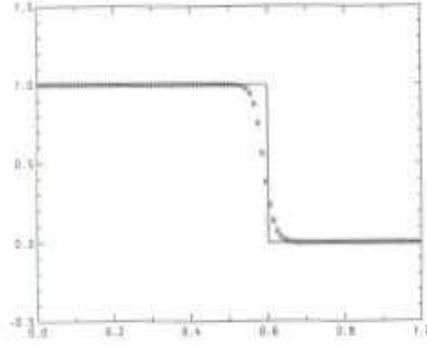


Figura 3.7: Minmod

3.6 Il caso parabolico

L'estensione degli argomenti precedenti al caso parabolico non lineare presenta due difficoltà connesse tra loro: la presenza di termini convettivi, che come nel caso lineare, possono porre dei problemi di stabilità per piccole diffusioni, e la possibilità di punti di degenerazione della viscosità, che induce delle singolarità nella soluzione.

Per approssimare l'equazione (3.43) è possibile sviluppare un formalismo analogo a quello introdotto nel caso iperbolico. In particolare risulta utile considerare degli schemi consistenti, conservativi a $2l + 1$ punti nella forma

$$v_{m,n+1} = v_{m,n} - \lambda \{d_{m+\frac{1}{2},n} - d_{m-\frac{1}{2},n}\}, \quad (3.84)$$

per una funzione di flusso numerico diffusivo

$$d_{m+\frac{1}{2},n} = d(v_{m-l+1,n}, \dots, v_{m+l,n}),$$

consistente con il flusso della (3.43), ossia $A(u) - B(u)_x$. La scelta più semplice per avere questo è di prendere

$$\begin{aligned} d_{m+\frac{1}{2},n} &= g(v_{m-l+1,n}, \dots, v_{m+l,n}) \\ &\quad - \frac{1}{h} (b(v_{m-l+2,n}, \dots, v_{m+l,n}) - b(v_{m-l+1,n}, \dots, v_{m+l-1,n})), \end{aligned}$$

dove g è la funzione generatrice di un flusso numerico consistente con la parte iperbolica, ossia

$$g(v, \dots, v) = A(v).$$

D'altra parte la funzione b è una funzione di $2l - 1$ variabili che deve verificare

$$b(v, \dots, v) = B(v).$$

In particolare per avere uno schema a tre punti, $l = 1$, avremo che la funzione b coincide con la diffusione B e il flusso numerico si scrive allora

$$d(u, v) = g(u, v) - \frac{1}{h}(B(v) - B(u)). \quad (3.85)$$

È possibile estendere il concetto di monotonia anche a questa classe di schemi. Se si prende per esempio come schema iperbolico lo schema di Engquist–Osher

$$g(u, v) = A^+(u) + A^-(v), \quad (3.86)$$

con le definizioni date nell'esempio 3.17, si vede che lo schema (3.84) è monotono se vale la seguente condizione di CFL

$$\max |A'| \frac{k}{h} + 2 \max |B'| \frac{k^2}{h} \leq 1. \quad (3.87)$$

Questa condizione generalizza al caso non lineare la condizione (2.29), trovata per le equazioni lineari di convezione e diffusione. Abbiamo allora il seguente risultato, dimostrato da Evje e Karlsen [EK1].

Teorema 3.28 *Se vale la condizione di CFL (3.87), la successione v_h generata dallo schema (3.84)–(3.86) converge, al tendere di h a zero, in L^1_{loc} , all'unica soluzione entropica BV del problema (3.43)–(3.44) nel senso della definizione 3.4. Inoltre la successione $B(v_h)$ converge uniformemente sui compatti alla funzione $B(u)$.*

È possibile generalizzare questo tipo di risultati usando schemi impliciti per evitare le restrizioni dovute alla condizione di CFL.

Infine osserviamo si può estendere tutta la teoria degli schemi TVD, utilizzando gli schemi iperbolici della Sezione 3.5 accoppiati con la stessa discretizzazione della diffusione, vedi [EK2, ANT, KT].

Osserviamo che la classe di schemi (3.84) discretizza l'equazione direttamente nella sua forma conservativa (3.43), piuttosto che considerare il termine diffusivo (equivalente nel caso regolare) $B'(u)u_x$, come spesso viene fatto nei metodi numerici che agiscono sull'equazione linearizzata, o comunque nel caso non degenero, vedi per esempio [Mo]. Si potrebbe per esempio utilizzare un flusso numerico della forma

$$\tilde{d}_{m+\frac{1}{2},n} = g_{m+\frac{1}{2},n} - \frac{1}{h} B' \left(\frac{v_{m,n} + v_{m+1,n}}{2} \right) \Delta_{m+\frac{1}{2}} v^n. \quad (3.88)$$

Questo metodo, benchè converga al limite ad una funzione, non converge però alla soluzione entropica. Si consideri per esempio di nuovo l'equazione (3.47), con il termine diffusivo dato da (3.48). In figura 3.8 si vede l'evoluzione nel tempo delle soluzioni calcolate con il flusso (3.85)–(3.86) (linea continua) e con il flusso (3.88) (linea tratteggiata). Il dato iniziale è dato dalla linea puntinata. Chiaramente il metodo conservativo produce una soluzione che si muove a velocità non fisica, come si potrebbe vedere utilizzando la Proposizione 3.12.

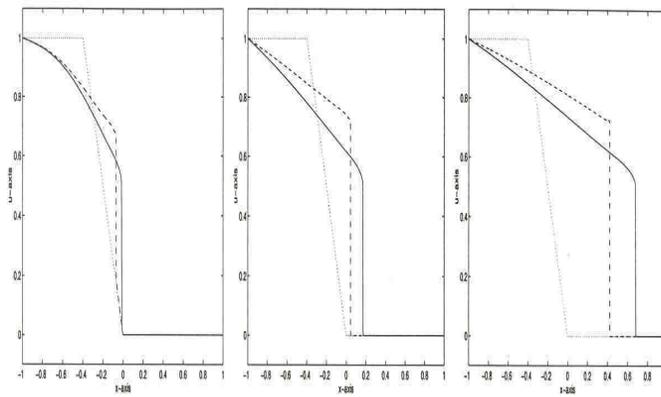


Figura 3.8: Confronto tra il metodo conservativo e quello non conservativo in tempi successivi: $t = 0.0625$ (sinistra), $t = 0.25$ (centro), $t = 1$ (destra).

Bibliografia

- [ANT] D. Aregba-Driollet, R. Natalini, S. Tang, Numerical Study of Diffusive BGK Approximations for Nonlinear Systems of Parabolic Equations, *Math. Comp.* 73 (2004) 63-94.
- [Ar] Aronson, D. G. , *Regularity properties of flows through porous media: A counterexample*. *SIAM J. Appl. Math.* 19 1970 299–307.
- [BER] G.I. Barenblatt, V.M. Entov, V.M. Ryzhik, *Theory of fluid flows through natural rocks*, Kluwer Academic Publishers, Dordrecht, 1990.
- [Bo] F. Bouchut, *Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources*, *Frontiers in Mathematics series*, Birkhäuser, 2004
- [Br] H. Brézis, *Analisi Funzionale*, Liguori, 1988.
- [Ca] Carrillo J., *Entropy solutions for nonlinear degenerate problems*. *Arch. Ration. Mech. Anal.* 147 (1999), no. 4, 269–361.
- [CIP] Cercignani, Carlo; Illner, Reinhard; Pulvirenti, Mario *The mathematical theory of dilute gases*. *Applied Mathematical Sciences*, 106. Springer-Verlag, New York, 1994.
- [CFL] R. Courant, K.O. Friedrichs, H. Lewy, *Über die partiellen differenzgleichungen der mathematischen physik*, *Math. Ann.* **100** (1928), 32.
- [GR1] Godlewski, Edwige; Raviart, Pierre-Arnaud *Hyperbolic systems of conservation laws*. *Mathématiques & Applications [Mathematics and Applications]*, 3/4. Ellipses, Paris, 1991.
- [GR2] Godlewski, Edwige; Raviart, Pierre-Arnaud *Numerical approximation of hyperbolic systems of conservation laws*. *Applied Mathematical Sciences*, 118. Springer-Verlag, New York, 1996.
- [Da] C.M. Dafermos, *Hyperbolic conservation laws in continuum physics*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, 325. Springer-Verlag, Berlin, 2000.

- [Di1] DiBenedetto E., Degenerate parabolic equations. Universitext. Springer-Verlag, New York, 1993.
- [Di2] DiBenedetto E., Partial differential equations. Birkhäuser Boston, 1995.
- [EK1] S. Evje and K.H. Karlsen, Monotone difference approximation of BV solutions to degenerate convection–diffusion equations, *SIAM J. Numer. Anal.* 37 (2000), 1838–1860.
- [EK2] S. Evje and K.H. Karlsen, Second order difference schemes for degenerate convection–diffusion equations, Preprint, University of Bergen, 2000.
- [Gi1] Gilding, B. H. Improved theory for a nonlinear degenerate parabolic equation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 16 (1989), no. 2, 165–224.
- [Gi2] Gilding, B. H. The occurrence of interfaces in nonlinear diffusion-advection processes. *Arch. Rational Mech. Anal.* 100 (1988), no. 3, 243–263.
- [Gi] Giusti, Enrico Minimal surfaces and functions of bounded variation. Monographs in Mathematics, 80. Birkhäuser Verlag, 1984.
- [Ha] Harten, Ami, *High resolution schemes for hyperbolic conservation laws*. *J. Comput. Phys.* 49 (1983), no. 3, 357–393.
- [Ka] Kalašnikov, A. S., *Formation of singularities in solutions of the equation of nonstationary filtration*. (Russian) *Ž. Vychisl. Mat. i Mat. Fiz.* 7 1967 440–444.
- [Kr] S.N. Kružkov, First order quasilinear equations in several independent variables, *Math. USSR Sb.* 10 (1970), 217–243.
- [KL] Kreiss, Heinz-Otto; Lorenz, Jens, Initial-boundary value problems and the Navier-Stokes equations. Pure and Applied Mathematics, 136. Academic Press, 1989.
- [KT] A. Kurganov and E. Tadmor, New high–resolution central schemes for nonlinear conservation laws and convection–diffusion equations, *J. Computational Phys.*, 160 (2000), 214–282.
- [Jo] John, Fritz Partial differential equations. Reprint of the fourth edition. Applied Mathematical Sciences, 1. Springer-Verlag, New York, 1991.
- [LSU] Ladyženskaja, O. A.; Solonnikov, V. A.; Ural' ceva, N. N., Linear and quasilinear equations of parabolic type. (Russian) Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23 American Mathematical Society, 1967.

- [La] Lax, P. D. Hyperbolic systems of conservation laws. II. *Comm. Pure Appl. Math.* 10 1957 537–566.
- [LW] P.D. Lax and B. Wendroff, *Systems of conservation laws*, *Comm. Pure Appl. Math.* **13** (1960), 217–237.
- [Le] LeVeque, Randall J., *Finite Volume Methods for Hyperbolic Problems Applied Mathematics*. Cambridge Texts in Applied Mathematics, Cambridge University Press, 2002, Cambridge.
- [MM] Morton, K. W.; Mayers, D. F. *Numerical solution of partial differential equations. An introduction*. Cambridge University Press, Cambridge, 1994.
- [Mo] Morton, K. W. *Numerical solution of convection-diffusion problems. Applied Mathematics and Mathematical Computation*, 12. Chapman & Hall, London, 1996.
- [RM] Richtmyer, Robert D.; Morton, K. W. *Difference methods for initial-value problems*. Reprint of the second edition. Robert E. Krieger Publishing Co., 1994.
- [Se] Serre, Denis *Systems of conservation laws. 1 and 2. Hyperbolicity, entropies, shock waves*. Cambridge University Press, Cambridge, 1999.
- [Sh] C.W. Shu, TVB uniformly high-order schemes for conservation laws, *Math. Comp.* 49 (1987) 105–121.
- [Sm] Smoller, Joel *Shock waves and reaction-diffusion equations*. Second edition. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, 258. Springer-Verlag, New York, 1994.
- [St] J.C. Strikwerda, *Finite difference schemes and partial differential equations*. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, 1989.
- [Sw] P.K. Sweby, High resolution schemes using flux limiters for hyperbolic equations, *SIAM J. Numer. Anal.* 21 (1984), 995–1011.
- [Ta] Tadmor, Eitan *Approximate solutions of nonlinear conservation laws. Advanced numerical approximation of nonlinear hyperbolic equations (Cetraro, 1997)*, 1–149, *Lecture Notes in Math.*, 1697, Springer, Berlin, 1998.
- [Ty] M.E. Taylor, *Partial differential equations (3 vol.)*. Applied Mathematical Sciences 115–117, Springer-Verlag, New York, 1996.
- [VH] Vol’pert, A. I.; Hudjaev, S. I. The Cauchy problem for second order quasilinear degenerate parabolic equations. (Russian) *Mat. Sb. (N.S.)* 78 (120) 1969 374–396.

- [Yi] Yin, Jing Xue On the uniqueness and stability of BV solutions for nonlinear diffusion equations. *Comm. Partial Differential Equations* 15 (1990), no. 12, 1671–1683.
- [Wh] G. B. Whitham, *Linear and nonlinear waves*. Reprint of the 1974 original. *Pure and Applied Mathematics. A Wiley-Interscience Publication*. John Wiley & Sons, New York, 1999.
- [WY] Wu, Zhuo Qun, Yin, Jing Xue, Some properties of functions in BV_x and their applications to the uniqueness of solutions for degenerate quasilinear parabolic equations. *Dongbei Shuxue* 5 (1989), no. 4, 395–422.
- [Zh] Zhao, Jun Ning Some properties of solutions of quasilinear degenerate parabolic equations and quasilinear degenerate elliptic equations. *Dongbei Shuxue* 2 (1986), no. 3, 281–302.