Soluzioni Analitiche e Numeriche Applicate all'Ingegneria Ambientale

Massimiliano Martinelli
massimiliano.martinelli@gmail.com

Università Politecnica delle Marche, Ancona
Facoltà di Ingegneria

11-12 Marzo 2009

Consideriamo il problema di Dirichlet omogeneo

$$\begin{cases} -u''(x) = f(x) & 0 < x < 1 \\ u(0) = u(1) = 0 \end{cases} \tag{1}$$

che descrive, per esempio, la configurazione di equilibrio di un filo elastico di tensione pari a uno, fissato agli estremi e soggetto a una forza trasversale f.

- Se si considera il caso in cui il carico è concentrato in uno o più punti, la soluzione fisica esiste ed è continua ma non è derivabile
- Se il carico f è una funzione costante a tratti, la soluzione fisica non ha derivata seconda continua

Queste funzioni non possono essere soluzione del problema differenziale, in quanto il problema (1) richiede che $u(x)$ abbia derivata seconda continua

Serve una formulazione del problema alternativa che permetta di descrivere lo stesso modello fisico ma con una richiesta di regolarità minore per la soluzione u

- Moltiplichiamo l'equazione differenziale (1) per una *funzione test* v (le cui proprietà verranno definite più tardi) e integriamo sul dominio in cui è definito il problema

$$-u''v = fv \qquad \Rightarrow \qquad -\int_0^1 u''(x)v(x)\,dx = \int_0^1 f(x)v(x)\,dx$$

- Effettuiamo un'integrazione per parti allo scopo di eliminare la derivata seconda (in modo da richiedere una soluzione con minore regolarità)

$$-\int_0^1 u''(x)v(x)\,dx = \int_0^1 u'(x)v'(x)\,dx - \left[u'(x)v(x)\right]_0^1$$

- Scegliendo solo funzioni test che si annullanno nei punti $x = 0$ e $x = 1$ l'equazione diventa

$$\int_0^1 u'(x)v'(x)\,dx = \int_0^1 f(x)v(x)\,dx$$

- Quali requisiti deve soddisfare lo spazio V delle funzioni test, in modo che le operazioni introdotte abbiano senso?

- Se u e v appartenessero a $C^1([0,1])$, avremmo $u, v \in C^0([0,1])$ e quindi l'integrale a primo membro avrebbe senso
- Però le soluzioni fisiche potrebbero non essere derivabili con continuità $\Rightarrow$ dobbiamo richiedere regolarità inferiore
- Introduciamo la nozione di *spazio* L^p

$$L^p(0,1) = \left\{ v\colon (0,1) \to \mathbb{R} \text{ t.c. } ||v||_{L^p(0,1)} \equiv \left(\int_0^1 |v(x)|^p \, dx \right)^{1/p} < +\infty \right\}$$

- Dato che si vuole che $\int_0^1 u'v' \, dx$ sia ben definito, la richiesta minima è che $u'v'$ appartenga a $L^1(0,1)$
- Ora notiamo che la norma dello spazio $L^2(0,1)$ è indotta dal prodotto scalare

$$\langle \varphi, \psi \rangle_{L^2(0,1)} = \left(\int_0^1 \varphi \psi \, dx \right)^{1/2}$$

perciò, utilizzando la disuguaglianza di Cauchy-Schwarz si ha che

$$\text{Se } \varphi, \psi \in L^2(0,1) \quad \Rightarrow \quad \varphi\psi \in L^1(0,1)$$

- Perché l'integrale $\int_0^1 u'v'\,dx$ abbia senso, occorre allora che u e v appartengano allo spazio (*di Sobolev*)

$$H^1(0,1) = \left\{ v \in L^2(0,1) \text{ t.c. } v' \in L^2(0,1) \right\}$$

- Poiché abbiamo le condizioni al bordo $u(0) = u(1) = 0$ e avevamo scelto v tra le funzioni nulle in $x = 0$ e $x = 1$, dobbiamo scegliere u e v nel seguente sottospazio di $H^1(0,1)$

$$H_0^1(0,1) = \left\{ v \in H^1(0,1) \text{ t.c. } v(0) = v(1) = 0 \right\}$$

- Perché anche l'integrale $\int_0^1 fv\,dx$ abbia senso, dobbiamo inoltre supporre che $f \in L^2(0,1)$

Il problema differenziale (1) viene dunque ricondotto al problema integrale

$$\text{Cercare } u \in H_0^1(0,1) \text{ t.c. } \quad \int_0^1 u'v'\,dx = \int_0^1 fv\,dx \qquad \forall v \in H_0^1(0,1) \qquad (2)$$

- La soluzione $u \in C^2([0,1])$ del problema differenziale (1) è chiamata *soluzione forte*
- La soluzione $u \in H_0^1(0,1)$ del problema integrale (2) è chiamata *soluzione debole*
- Una soluzione forte è soluzione del problema in forma debole, ma non è vero il viceversa (ovvero, una soluzione debole può non essere soluzione del problema differenziale)
- Le funzioni di $H^1(0,1)$ non sono necessariamente derivabili in senso classico. Per esempio le funzioni continue a tratti con raccordi a spigolo appartengono ad $H^1(0,1)$ ma non a $C^1([0,1])$

Problema di Dirichlet non omogeneo

Nel caso non omogeneo le condizioni al bordo in

$$\begin{cases} -u''(x) = f(x) & 0 < x < 1 \\ u(0) = u(1) = 0 \end{cases}$$

sono sostituite da $u(0) = u_0$ e $u(1) = u_1$ con $u_0, u_1 \in \mathbb{R}$ valori assegnati.

- Ci si può ricondurre al caso omogeneo notando che se se u è la soluzione del caso non omogeneo allora la funzione

$$\overset{\circ}{u} = u - \left[(1-x)u_0 + xu_1\right]$$

è soluzione del problema omogeneo.

Problema di Neumann

Definiamo la formulazione debole per il seguente problema di Neumann

$$\begin{cases} -u''(x) + \sigma(x)u(x) = f(x) \qquad 0 < x < 1\,, \quad \sigma(x) > 0 \\ u'(0) = h_0 \\ u'(1) = h_1 \end{cases} \tag{3}$$

- Moltiplicando per una funzione test v e integrando per parti sull'intervallo $(0,1)$ si ottiene

$$\int_0^1 u'(x)v'(x)\,dx + \int_0^1 \sigma uv\,dx - \left[u'v\right]_0^1 = \int_0^1 f(x)v(x)\,dx$$

- Supponiamo $f \in L^2(0,1)$ e $\sigma \in L^\infty(0,1)$ (ovvero una funzione limitata q.o. su $(0,1)$)
- u non è nota al bordo $\Rightarrow$ non si deve richiedere che v si annulli al bordo
- La formulazione debole è quindi

Cercare $u \in H^1(0,1)$ t.c. $\displaystyle\int_0^1 u'v'\,dx + \int_0^1 \sigma uv\,dx = \int_0^1 fv\,dx + h_1 v(1) - h_0 v(0)$

Problema di Dirichlet per l'equazione di Poisson multidimensionale

Il problema consiste nel cercare u tale che

$$\begin{cases} -\nabla^2 u = f & \text{in } \Omega \\ u = 0 & \text{in } \partial\Omega \end{cases} \tag{4}$$

- Per la formulazione debole, procediamo come nel caso monodimensionale, moltiplicando per una funzione test v e integriamo su Ω

$$-\int_\Omega \nabla^2 u v \, d\Omega = \int_\Omega f v \, d\Omega$$

- Ricordando il teorema di Gauss-Green (teorema della divergenza) abbiamo che

$$\int_{\partial\Omega} v \nabla u \cdot \mathbf{n} \, d\Sigma = \int_\Omega \text{div}(\nabla u v) = \int_\Omega \nabla^2 u v \, d\Omega + \int_\Omega \nabla u \cdot \nabla v \, d\Omega$$

- Perciò abbiamo

$$\int_\Omega \nabla u \cdot \nabla v \, d\Omega - \int_{\partial\Omega} v \nabla u \cdot \mathbf{n} \, d\Sigma = \int_\Omega f v \, d\Omega$$

Problema di Dirichlet per l'equazione di Poisson multidimensionale

- Scegliendo le funziondi test nulle al bordo, il termine al contorno dell'equazione precedente sara nullo
- Tenendo conto di questo fatto, giungiamo alla seguente formulazione debole

$$\text{Cercare } u \in H_0^1(\Omega) \text{ t.c. } \int_\Omega \nabla u \cdot \nabla v \, d\Omega = \int_\Omega f v \, d\Omega \quad \forall v \in H_0^1(\Omega) \tag{5}$$

con $f \in L^1(\Omega)$ ed avendo posto

$$H^1(\Omega) \equiv \left\{ v \colon \Omega \to \mathbb{R} \text{ t.c. } v \in L^2(\Omega), \frac{\partial v}{\partial x_i} \in L^2(\Omega), i = 1, \dots \right\}$$

$$H_0^1(\Omega) \equiv \left\{ v \in H^1(\Omega) \text{ t.c. } v = 0 \text{ su } \Omega \right\}$$

Definizione di forma

Dato uno spazio funzionale normato V, si dice *forma* un'applicazione a che associa ad ogni coppia di elementi di V un numero reale, ovvero $a(\cdot,\cdot): V \times V \to \mathbb{R}$.
Una forma si dice

- *bilineare* se è lineare rispetto ad entrambi i suoi argomenti
- *continua* se $\exists M > 0$ tale che

$$|a(u,v)| \leq M||u||_V||v||_V \qquad \forall u, v \in V$$

- *coerciva* se $\exists \alpha > 0$ tale che

$$a(v,v) \geq \alpha||v||_V^2 \qquad \forall v \in V$$

- Il termine $\int_\Omega \nabla u \cdot \nabla v \, d\Omega$ può essere visto come una *forma bilineare* $a(\cdot,\cdot)\colon V \times V \to \mathbb{R}$

$$a(u,v) \equiv \int_\Omega \nabla u \cdot \nabla v \, d\Omega$$

(sia l'integrazione che la differenziazione sono operazioni lineari nei rispettivi argomenti)

- Anche il termine $\int_\Omega fv \, d\Omega$ può essere visto come un *funzionale lineare* $F(\cdot)\colon V \to \mathbb{R}$

$$F(v) \equiv \int_\Omega fv \, d\Omega$$

(se $f \in L^2(0,1)$ il funzionale lineare F è limitato, e quindi anche continuo)

- Il problema in forma debole (5) allora diventa

$$\text{Cercare } u \in H_0^1(\Omega) \text{ t.c. } a(u,v) = F(v) \quad \forall v \in H_0^1(\Omega)$$

Un problema differenziale lineare $\begin{cases} Lu = f & \text{su } \Omega \\ Bu = g & \text{su } \partial\Omega \end{cases}$ si scrive in forma debole come

$$\text{Cercare } u \in U \text{ t.c. } a(u,v) = F(v) \quad \forall v \in V$$

dove U e V sono due opportuni spazi di funzioni, $a(\cdot,\cdot)\colon U \times V \to \mathbb{R}$ una forma bilineare e $F(\cdot)\colon V \to \mathbb{R}$ è un funzionale lineare

Teorema di esistenza e unicità (Lax-Milgram)

Sia V uno spazio di Hilbert, $a(\cdot,\cdot)\colon V \times V \to \mathbb{R}$ è una forma bilineare continua e coerciva. Sia $F(\cdot)\colon V \to \mathbb{R}$ è un funzionale lineare e continuo. Allora, esiste ed è unica la soluzione del problema

$$\text{Trovare } u \in V \text{ t.c. } a(u,v) = F(v) \quad \forall v \in V$$

Dato un problema lineare in forma debole, si studia l'esistenza e l'unicità attraverso la verifica della continuità e coercività di $a(\cdot,\cdot)\colon V \times V \to \mathbb{R}$ e attraverso la continuità (o limitatezza) di $F(\cdot)\colon V \to \mathbb{R}$

Approssimazione con il metodo di Galerkin

Consideriamo la formulazione debole di un generico problema

$$\text{Trovare } u \in V \text{ t.c. } a(u,v) = F(v) \quad \forall v \in V \tag{6}$$

dove V un opportuno spazio di Hilbert, $a(\cdot,\cdot)$ una forma biulineare continua e coerciva da $V \times V$ in $\mathbb{R}$ e $F(\cdot)$ un funzionale lineare continuo da V in $\mathbb{R}$.

- Il problema di Galerkin per la soluzione approssimata di (6) consiste nel cercare una soluzione approssimata $u_h \in V_h$ dove

$$V_h \subset V, \quad \dim(V_h) = N_h < +\infty$$

Problema di Galerkin

$$\text{Trovare } u_h \in V_h \text{ t.c. } a(u_h,v_h) = F(v_h) \quad \forall v_h \in V_h \tag{7}$$

- $\dim(V_h) = N_h < +\infty \Rightarrow V_h$ ammette una base finita $\{\varphi_j(\mathbf{x}), j = 1, \ldots, \varphi_{N_h}\}$, ovvero $V_h = \text{span}(\varphi_1, \ldots, \varphi_{N_h})$
- Ogni elemento u_h di V_h può essere scritto come combinazione lineare degli elementi della base

$$u_h(\mathbf{x}) = \sum_{j=1}^{N_h} u_j \varphi_j(\mathbf{x})$$

  dove gli u_j sono dei coefficienti incogniti
- Nel problema di Galerkin, v_h è un qualsiasi elemento di V_h (e quindi vale anche per gli elementi della base), perciò il problema diventa

$$\text{Trovare } u_h \in V_h \text{ t.c. } a(u_h, \varphi_i) = F(\varphi_i) \quad i = 1, \ldots, N_h$$

- Per la linearità della forma a abbiamo

$$a(u_h, \varphi_i) = a\left(\sum_{j=1}^{N_h} u_j \varphi_j, \varphi_i\right) = \sum_{j=1}^{N_h} u_j a(\varphi_j, \varphi_i)$$

- Il problema di Galerkin allora diventa

$$\text{Trovare } u_1, \ldots, u_{N_h} \text{ t.c. } \sum_{j=1}^{N_h} u_j a(\varphi_j, \varphi_i) = F(\varphi_i) \quad i = 1, \ldots, N_h \tag{8}$$

- Denotiamo con A_h la matrice (detta *di rigidezza*) i cui elementi a_{ij} sono definiti da

$$a_{ij} \equiv a(\varphi_j, \varphi_i)$$

  e siano $\mathbf{f}_h = \left(F(\varphi_1), \ldots, F(\varphi_{N_h})\right)^T$ e $\mathbf{u}_h = \left(u_1, \ldots, u_{N_h}\right)^T$
- Il problema di Galerkin (8) è equivalente alla risoluzione del sistema lineare

$$A_h \mathbf{u}_h = \mathbf{f}_h$$

- Per un problema ellittico, la matrice $\mathbf{A}_h$ è simmetrica e definita positiva
- La struttura di sparsità di A_h dipende dal supporto delle funzioni di base

Esistenza, unicità e convergenza della soluzione del problema di Galerkin

Se valgono le stesse ipotesi del teorema di Lax-Milgram per $a(\cdot,\cdot)$ e $F(\cdot)$ allora:

- La soluzione u_h esiste ed è unica (equivale all'invertibilità della matrice A_h)
- Il problema di Galerkin è stabile uniformemente rispetto ad h in quanto vale la seguente maggiorazione della soluzione $||u_h||_V \leq \frac{1}{\alpha}||F||_{V'}$
- (*Lemma di Céa*) Il metodo di Galerkin è fortemente consistente, ovvero

$$a(u - u_h, v_h) = 0 \quad \forall v_h \in V_h$$

Il problema di Galerkin è convergente se

$$\lim_{h \to 0} ||u - u_h||_V = 0$$

. Convergenza di ordine α se $||u - u_h||_V = O(h^\alpha)$

Studio della convergenza

- Un aspetto importante dell'approssimazione dii un problema variazionale è la scelta di uno spazio V_h "opportuno"
  - Se ad esempio sappiamo che la soluzione è poco regolare, non conviene scegliere spazi V_h di funzioni troppo regolari (come ad esempio lo spazio dei polinomi definiti su tutto il dominio del problema)
  - Viceversa, gli spazi regolari dovrebbero essere preferiti per l'approssimazione di soluzioni regolari
- In generale, più lo spazio è regolare, meno è sparsa la matrice $A_h \Rightarrow$ maggior costo per il calcolo di A_h e per la risoluzione del sistema lineare
- In compenso, con spazi regolari si ha una convergenza migliore (il limite è dato dalla regolarità della soluzione)

Metodo degli Elementi Finiti

È un approssimazione di Galerkin in cui lo spazio V_h ha le seguenti caratteristiche:

- V_h è associato ad una triangolazione T_h del dominio Ω_h
- Su ogni elemento K di T_h (cioè su ogni triangolo, rettangolo, tetraedro, ...), gli elementi di V_h sono *funzioni polinomiali*
- V_h ammette una base $\{\varphi_j(\mathbf{x}), j = 1, \ldots, \varphi_{N_h}\}$ dove φ_j sono funzioni a *supporto limitato "piccolo"*
- In generale, le funzioni φ_j sono dei polinomi a tratti di grado poco elevato

Risultato di convergenza

Sia $u \in V$ la soluzione esatta del problema variazionale

$$a(u,v) = F(v) \quad \forall v \in V$$

e sia u_h la sua approssimazione ottenuta con il metodo degli elementi finiti di grado r, ovvero la soluzione del problema

$$a(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h$$

in cui

$$V_h = \left\{ v_h \in X_h^r \text{ t.c. } v_h(0) = v_h(1) = 0 \right\}$$

e X_h^r è lo spazio delle funzioni continue sull'intervallo $I = [0,1]$ e polinomiali a tratti di grado r sugli intervalli $I_i = [x_{i-1}, x_i]$. Sia inoltre $u \in H^p(I)$ per un opportuno $p \geq r$. Allora vale la seguente *stima a priori dell'errore*

$$||u - u_h||_V \leq Ch^r |u|_{H^{r+1}(I)}$$

dove $h = \max_i (x_i - x_{i-1})$.

Ordine di convergenza del metodo degli elementi finiti

r	$u \in H^1(I)$	$u \in H^2(I)$	$u \in H^3(I)$	$u \in H^4(I)$	$u \in H^5(I)$
1	converge	$\boxed{h}$	h	h	h
2	converge	h	$\boxed{h^2}$	h^2	h^2
3	converge	h	h^2	$\boxed{h^3}$	h^3
4	converge	h	h^2	h^3	$\boxed{h^4}$

Esempio

Il problema differenziale

$$\begin{cases} -u''(x) + g(x)u(x) = f(x) & x \in (a,b)\,, g(x) > 0 \\ u(a) = u(b) = 0 \end{cases}$$

ha la seguente formulazione debole (o variazionale, o integrale)

Trovare una funzione $u \in H_0^1(a,b)$ tale che

$$\int_a^b u'(x)v'(x)\,dx + \int_a^b g(x)u(x)v(x)\,dx = \int_a^b f(x)v(x)\,dx \qquad \forall v \in H_0^1(a,b)$$

dove $H_0^1(a,b) = \left\{v \in L^2(a,b)\,, v' \in L^2(a,b)\,, v(a) = v(b) = 0\right\}$

Problema discreto:

- *Triangolazione*: si divide l'intervallo $[a,b]$ in $m+1$ sottointervalli $I_i = [x_{i-1}, x_i]$ con $a = x_0 < x_1 < \cdots < x_m < x_{m+1} = b$ (Poniamo $h_i = x_i - x_{i-1}$ e $h = \max_i h_i$)
- *Definizione dello spazio V_h:* V_h è l'insieme delle funzioni lineari su ogni intervallo I_j, continue su $[a,b]$ e a valore nullo in a e b (Si verifica che questo spazio è in $H_0^1(a,b)$)
- *Base di V_h*: $\varphi_i(x)$ funzione lineari a tratti tale

$$\varphi_i(x_j) = \delta_{ij} = \begin{cases} 1 & \text{se } j = i \\ 0 & \text{se } j \neq i, j = 1, \ldots, m \end{cases}$$

La generica funzione $\varphi_i(x)$ si scrive come

$$\varphi_i(x) = \begin{cases} 0 & x \leq x_{i-1} \\ \dfrac{x - x_{i-1}}{x_i - x_{i-1}} = \dfrac{1}{h_i}(x - x_{i-1}) & x_{i-1} < x \leq x_i \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i} = \dfrac{1}{h_{i+1}}(x_{i+1} - x) & x_i < x < x_{i+1} \\ 0 & x \geq x_{i+1} \end{cases}$$

Problema discreto:

- *Formulazione di Galerkin*: Trovare $u_h \in V_h$ tale che

$$\int_a^b u_h'(x)\varphi_i'(x)\,dx + \int_a^b g(x)u_h(x)\varphi_i(x)\,dx = \int_a^b f(x)\varphi_i(x)\,dx \qquad \forall i = 1,\dots,m$$

- Poiché $u_h(x) = \sum_{j=1}^m u_j\varphi_j(x)$, si ottiene

$$\sum_{j=1}^m u_j\left(\int_a^b \varphi_j'(x)\varphi_i'(x)\,dx + \int_a^b g(x)\varphi_j(x)\varphi_i(x)\,dx\right) = \int_a^b f(x)\varphi_i(x)\,dx \quad \forall i = 1,\dots,m$$

- La derivata prima di $\varphi_i(x)$ è

$$\varphi_i'(x) = \begin{cases} 0 & x \le x_{i-1} \\ \dfrac{1}{h_i} & x_{i-1} < x \le x_i \\ -\dfrac{1}{h_{i+1}} & x_i < x < x_{i+1} \\ 0 & x \ge x_{i+1} \end{cases}$$

$$\int_a^b \varphi_j'(x)\varphi_i'(x)\,dx = \frac{1}{h_i}\int_{x_{i-1}}^{x_i} \varphi_j'(x)\,dx - \frac{1}{h_{i+1}}\int_{x_i}^{x_{i+1}} \varphi_j'(x)\,dx = \begin{cases} -\dfrac{1}{h_i} & \text{se } j = i-1 \\ \dfrac{1}{h_i} + \dfrac{1}{h_{i+1}} & \text{se } j = i \\ -\dfrac{1}{h_{i+1}} & \text{se } j = i+1 \end{cases}$$

- Supponendo che $g(x) = g \in \mathbb{R}$, si ha

$$\int_a^b g(x)\varphi_j(x)\varphi_i(x)\,dx = \frac{g}{h_i}\int_{x_{i-1}}^{x_i} (x - x_i)\varphi_j(x)\,dx + \frac{g}{h_{i+1}}\int_{x_i}^{x_{i+1}} (x_{i+1} - x)\varphi_j(x)\,dx$$

$$= \begin{cases} \dfrac{gh_i}{6} & \text{se } j = i-1 \\ \dfrac{g}{3}\left(h_i + h_{i+1}\right) & \text{se } j = i \\ \dfrac{gh_{i+1}}{6} & \text{se } j = i+1 \end{cases}$$

Arriviamo quindi all'equazione alle differenze

$$u_{i+1}\left(\frac{gh_{i+1}}{6}-\frac{1}{h_{i+1}}\right)+u_i\left[\left(\frac{1}{h_i}+\frac{1}{h_{i+1}}\right)+\frac{g}{3}(h_i+h_{i+1})\right]+u_{i-1}\left(\frac{gh_i}{6}-\frac{1}{h_i}\right)=c_i$$

con $c_i=\int_a^b f(x)\varphi_i(x)\,dx$.

- Se anche $f\in V_h$ si ha

$$c_i=\int_a^b f(x)\varphi_i(x)\,dx=\sum_{j=1}^{m} f_j\int_a^b \varphi_j(x)\varphi_i(x)\,dx=\frac{1}{6}\left[f_{i+1}h_{i+1}+2f_i(h_{i+1}+h_i)+f_{i-1}h_i\right]$$

- Se la griglia è uniforme allora $h_i=h$ per $i=1,\ldots,m+1$, e quindi

$$u_{i+1}\left(\frac{gh}{6}-\frac{1}{h}\right)+u_i\left(\frac{2}{h}+\frac{2gh}{3}\right)+u_{i-1}\left(\frac{gh}{6}-\frac{1}{h}\right)=\frac{h}{6}(f_{i+1}+4f_i+f_{i-1})$$

- Si può fornire una forma esplicita per u_i e studiare la zero-stabilità del metodo con gli stessi metodi usati per le ODE

Problemi parabolici: metodo delle differenze finite

Vogliamo risolvere con il metodo delle differenze finite il problema seguente

$$\begin{cases} u_t - \sigma u_{xx} = 0 & 0 \le x \le 1 \quad t > 0 \\ u(x,0) = u_0(x) & \text{(condizioni iniziali)} \\ u(0,t) = u(1,t) = 0 & \text{(condizioni ai limiti)} \end{cases}$$

- Supponiamo di avere una griglia uniforme con passo $\Delta x = \frac{1}{N+1}$ lungo la direzione x e Δt lungo la direzione t
- Notazione: $u_i^n = u(i\Delta x, n\Delta t) = u(x_i, t_n)$
- Discretizzazione alle differenze finite nello spazio:

$$u_{xx}(x_i,t_n) = \frac{u_{i+1}^n - 2u_i^2 + u_{i-1}^n}{\Delta x^2} + O(\Delta x^2)$$

- Discretizzazione alle differenze finite nel tempo:

$$u_t(x_i,t_n) = \frac{u_i^{n+1} - u_i^n}{\Delta t} + O(\Delta t)$$

- Abbiamo perciò lo schema seguente

$$\begin{cases} \dfrac{u_i^{n+1}-u_i^n}{\Delta t} = \sigma\left(\dfrac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2}\right) & \\ u_i^0 = u_0(x_i) & i = 0,1,\ldots,N+1 \\ u_0^n = u_{N+1}^n = 0 & n = 0,1,\ldots \end{cases} \tag{9}$$

- A partire dai dati iniziali e ai limiti, tale schema fornisce in maniera *esplicita* la soluzione approssimata u_i^n, con un errore di troncamento dell'ordine di $O(\Delta t, \Delta x^2)$
- Introducendo i vettori $\mathbf{u}^n = (u_1^n, \ldots, u_N^n)^T$, $\mathbf{u}^{n+1} = (u_1^{n+1}, \ldots, u_N^{n+1})^T$ e la matrice A di discretizzazione del termine $-u_{xx}$, possiamo scrivere

$$\mathbf{u}^{n+1} = \big(I - \sigma \Delta t A\big)\mathbf{u}^n$$

- Studio della precisione: la soluzione dello schema numerico è soluzione (esatta) dell'equazione perturbata

$$u_t - \sigma u_{xx} = -\frac{\Delta t}{2}u_{tt} + \sigma\frac{\Delta x^2}{12}u_{xxxx} + \cdots = \left(-\sigma^2\frac{\Delta t}{2} + \sigma\frac{\Delta x^2}{12}\right)u_{xxxx} + \ldots$$

Problemi parabolici: convergenza

- La soluzione approssimata deve tendere alla soluzione esatta quando Δx e Δt tendono a zero

$$\lim_{\Delta_x \to 0\,,\,\Delta_t \to 0} ||u_i^n - u(x_i, t_n)|| = 0$$

- *Convergenza*: si usa il Teorema di Lax-Richtmyer (dato un problema ben posto e un'approssimazione consistente, la stabilità equivale alla convergenza)
- *Consistenza*: l'errore locale di troncamento tende a zero al tendere a zero dei parametri di discretizzazione Δx e Δt
- *Stabilità*: gli errori non si accumulano

Stabilità iterativa

- Per una *griglia data* (Δx e Δt), se i dati iniziali $\mathbf{u}^0$ sono limitati allora la soluzione $\mathbf{u}^n$ ottenuta al tempo $t_n = n\Delta t$ è *uniformemente limitata* (per n arbitrariamente grande), ovvero $\exists K > 0$ tale che

$$||\mathbf{u}^n|| \leq K||\mathbf{u}^0||$$

- Per lo studio di questo tipo di stabilità si usa l'*analisi di Von Neumann*
- *Schema di approssimazione stabile*: è uno schema tale che, se la soluzione iniziale del problema continuo è limitata, la soluzione numerica resta uniformemente limitata su ogni compatto Ω dello spazio (x,t) per ogni raffinamento della griglia (anche al limite $\Delta t\,, \Delta x \to 0$). Cioè, per ogni Δt, Δx

$$||u_i^n|| \leq C \qquad \forall x_i\,, t^n$$

Schema di approssimazione stabile $\Rightarrow$ Stabilità iterativa

Condizioe sufficiente per la stabilità iterativa

Una condizione sufficiente per la stabilità è la seguente:

$$||\mathbf{u}^{n+1}|| \leq \big(1 + O(\Delta t)\big)||\mathbf{u}^{n}||$$

dove il termine $O(\Delta t)$ è indipendente da Δx per una norma appropriata $||\cdot||$

Condizioe sufficiente per la stabilità iterativa

Una condizione sufficiente per la stabilità è la seguente:

$$||\mathbf{u}^{n+1}|| \leq \big(1 + O(\Delta t)\big)||\mathbf{u}^{n}||$$

dove il termine $O(\Delta t)$ è indipendente da Δx per una norma appropriata $||\cdot||$

Caso particolare: schema lineare costante: $\mathbf{u}^{n+1} = G\mathbf{u}^{n} + \mathbf{f}$

- G è una matrice indipendente da $\mathbf{u}$ (linearità) e indipendente da x e t (schema costante), ma può dipendere da Δt e Δx.
- $\mathbf{f}$ è un vettore con le stesse caratteristiche di G
- Si suppone inoltre che G sia diagonalizzabile: $G = T\Lambda T^{-1}$ con $\Lambda = \text{diag}(g_1, \ldots, g_n)$
- Lo schema è iterativamente stabile se e solo se

$$\rho(G) = \max_{i=1,\ldots,n} |g_i| \leq 1$$

- Questa condizione è solo *necessaria* per avere uno schema di approssimazione stabile

Analisi di Von Neumann

- Si considera un problema periodico ($0 \leq x \leq 2\pi$) con condizione iniziale

$$u_0(x) = u(x,0) = \sum_{k=-\infty}^{+\infty} \hat{u}^0(k) e^{ikx}$$

- La soluzione nei nodi della griglia è allora $u_j^n = \sum_{k=-\infty}^{+\infty} \hat{u}^n(k) e^{ikj\Delta x}$
- La soluzione è quindi combinazione lineare dei singoli modi di Fourier $\hat{u}^n(k) e^{ikj\Delta x}$
- Si ha stabilità iterativa se $|\hat{u}^n(k)|$ resta limitato per $t \to +\infty$ e per ogni k
- Se consideriamo il singolo modo di fourier e lo sostituiamo nello schema numerico otteniamo una relazione che lega il modo al tempo t_{n+1} con il modo al tempo t_n

$$\hat{u}^{n+1}(k) = g(k)\hat{u}^n(k)$$

  dove $g(k)$ è il *fattore di amplificazione* del modo associato alla frequenza k
- La condizione di stabilità relativa diventa dunque

$$\max_k |g(k)| \leq 1$$

Analisi di Von Neumann per lo schema esplicito (9)

$$u_j^{n+1} = u_j^n + \sigma \frac{\Delta t}{\Delta x^2} \left(u_{j+1}^n - 2u_j^n + u_{j-1}^n \right)$$

- Consideriamo come condizione iniziale il singolo modo di Fourier alla frequenza k: $u_j^0 = \hat{u}^0(k) e^{ikj\Delta x}$. Procedendo per ricorrenza, al passo $n+1$ si ottiene

$$\begin{aligned}
\hat{u}^{n+1}(k) e^{ikj\Delta x} &= u_j^{n+1} = u_j^n + \sigma \frac{\Delta t}{\Delta x^2} \left(u_{j+1}^n - 2u_j^n + u_{j-1}^n \right) \\
&= \hat{u}^n(k) \left[e^{ikj\Delta x} + \sigma \frac{\Delta t}{\Delta x^2} \left(e^{ik(j+1)\Delta x} - 2e^{ikj\Delta x} + e^{ik(j-1)\Delta x} \right) \right] \\
&= \hat{u}^n(k) e^{ikj\Delta x} \left[1 + \sigma \frac{\Delta t}{\Delta x^2} \left(e^{ik\Delta x} - 2 + e^{-ik\Delta x} \right) \right] \\
&= \hat{u}^n(k) e^{ikj\Delta x} \left[1 + 2\sigma \frac{\Delta t}{\Delta x^2} \left(\cos(k\Delta x) - 1 \right) \right] \\
&= \hat{u}^n(k) e^{ikj\Delta x} \left[1 - 4\sigma \frac{\Delta t}{\Delta x^2} \sin^2\left(\tfrac{k\Delta x}{2} \right) \right] = g(k) \hat{u}^n(k) e^{ikj\Delta x}
\end{aligned}$$

perciò $\hat{u}^{n+1}(k) = g(k)\hat{u}^n(k)$ con $g(k) = 1 - 4\sigma \frac{\Delta t}{\Delta x^2} \sin^2\left(\frac{k\Delta x}{2} \right)$

Analisi di Von Neumann per lo schema esplicito (9)

$$g(k) = 1 - 4\sigma \frac{\Delta t}{\Delta x^2} \sin^2\left(\tfrac{k\Delta x}{2}\right)$$

- La condizione di stabilità iterativa è $\max_k |g(k)| \leq 1$
- Nel nostro caso $g(k) < 1$ per ogni k, perciò resta da vedere sotto quali condizioni di Δx e Δt, $g(k) \geq -1$, ovvero

$$\sigma \frac{\Delta t}{\Delta x^2} \sin^2\left(\frac{k\Delta x}{2}\right) \leq \frac{1}{2} \qquad \forall k$$

- La disuguaglianza è verificata se e solo se $\sigma \frac{\Delta t}{\Delta x^2} \leq \frac{1}{2}$, ovvero $\Delta t \leq \frac{\Delta x^2}{2\sigma}$
- Quindi, per avere stabilità iterativa abbiamo un limite superiore per il passo temporale

Studio della convergenza per lo schema esplicito (9)

La soluzione vera soddisfa

$$u(x_i,t_{n+1}) = \left(1-\tfrac{2\sigma\Delta t}{\Delta x^2}\right)u(x_i,t_n) + \tfrac{\sigma\Delta t}{\Delta x^2}\left[u(x_{i-1},t_n)+u(x_{i+1},t_n)\right] + \underbrace{\tfrac{\Delta t^2}{2}u_{tt}(q) - \Delta t\sigma\tfrac{\Delta x^2}{24}\left[u_{xxxx}(p_1)+u_{xxxx}(p_2)\right]}_{R_i^n}$$

dove il resto R_i^n è maggiorato da

$$\begin{aligned}
|R_i^n| &= \left|\frac{\Delta t^2}{2}u_{tt}(q) - \Delta t\sigma\frac{\Delta x^2}{24}\left[u_{xxxx}(p_1)+u_{xxxx}(p_2)\right]\right| \\
&= \sigma\Delta t\left|\sigma\frac{\Delta t}{2}u_{xxxx}(q) - \frac{\Delta x^2}{24}\left[u_{xxxx}(p_1)+u_{xxxx}(p_2)\right]\right| \\
&\le \frac{\sigma\Delta t}{2}\left|\sigma\Delta t - \frac{\Delta x^2}{6}\right|\sup|u_{xxxx}(x,t)| = C\frac{\sigma\Delta t}{2}\left(\sigma\Delta t + \frac{\Delta x^2}{6}\right)
\end{aligned}$$

Lo schema numerico è definito come $\tilde{u}_i^{n+1} = \left(1-\tfrac{2\sigma\Delta t}{\Delta x^2}\right)\tilde{u}_i^n + \tfrac{\sigma\Delta t}{\Delta x^2}\left(\tilde{u}_{i-1}^n + \tilde{u}_{i+1}^n\right)$, quindi, introducendo l'errore globale di troncamento $e_i^n = u(x_i,t_n) - \tilde{u}_i^n$, si ha

$$e_i^{n+1} = u(x_i,t_{n+1}) - \tilde{u}_i^{n+1} = \left(1-\tfrac{2\sigma\Delta t}{\Delta x^2}\right)e_i^n + \tfrac{\sigma\Delta t}{\Delta x^2}\left(e_{i-1}^n + e_{i+1}^n\right) + R_i^n$$

Studio della convergenza per lo schema esplicito (9)

$$e_i^{n+1} = \left(1 - \tfrac{2\sigma\Delta t}{\Delta x^2}\right) e_i^n + \tfrac{\sigma\Delta t}{\Delta x^2}\left(e_{i-1}^n + e_{i+1}^n\right) + R_i^n$$

- Se $\frac{2\sigma\Delta t}{\Delta x^2} \leq 1$ allora

$$\begin{aligned}
\max_i |e_i^{n+1}| &\leq \left(1 - \frac{2\sigma\Delta t}{\Delta x^2}\right)|e_i^n| + \frac{\sigma\Delta t}{\Delta x^2}\left(|e_{i-1}^n| + |e_{i+1}^n|\right)| + |R_i^n| \\
&\leq \max_i |e_i^n| + C\frac{\sigma\Delta t}{2}\left(\sigma\Delta t + \frac{\Delta x^2}{6}\right) \\
&\leq \max_i |e_i^0| + (n+1)C\frac{\sigma\Delta t}{2}\left(\sigma\Delta t + \frac{\Delta x^2}{6}\right) \\
&= \frac{t_{n+1}C\sigma}{2}\left(\sigma\Delta t + \frac{\Delta x^2}{6}\right) \leq \frac{t_{n+1}C\sigma}{2}\left(\sigma t_{n+1} + \frac{1}{6}\right)
\end{aligned}$$

Quindi l'errore resta uniformemente limitato sul compatto $[0,1] \times [0, t_{n+1}] \Rightarrow$ Lo schema di approssimazione è stabile

$$e_i^{n+1} = \left(1 - \tfrac{2\sigma\Delta t}{\Delta x^2}\right) e_i^n + \tfrac{\sigma\Delta t}{\Delta x^2}\left(e_{i-1}^n + e_{i+1}^n\right) + R_i^n$$

Se $\frac{2\sigma\Delta t}{\Delta x^2} = \nu > 1$ allora

$$\begin{aligned}
\max_i |e_i^{n+1}| &\leq (\nu - 1)|e_i^n| + \frac{\nu}{2}\left(|e_{i-1}^n| + |e_{i+1}^n|\right)| + |R_i^n| \leq (2\nu - 1)\max_i |e_i^n| + |R_i^n| \\
&\leq (2\nu - 1)\max_i |e_i^n| + C\frac{\sigma\Delta t}{2}\left(\sigma\Delta t + \frac{\Delta x^2}{6}\right) \\
&\leq (2\nu - 1)^2 \max_i |e_i^{n-1}| + \left[(2\nu - 1) + 1\right] C\frac{\sigma\Delta t}{2}\left(\sigma\Delta t + \frac{\Delta x^2}{6}\right) \\
&\leq (2\nu - 1)^{n+1} \max_i |e_i^0| + \left[\sum_{k=0}^{n} (2\nu - 1)^k\right] C\frac{\sigma\Delta t}{2}\left(\sigma\Delta t + \frac{\Delta x^2}{6}\right) \\
&= \frac{(2\nu - 1)^{n+1} - 1}{2(\nu - 1)} C\frac{\sigma\Delta t}{2}\left(\sigma\Delta t + \frac{\Delta x^2}{6}\right) \\
&\leq \frac{(2\nu - 1)^{t_{n+1}/\Delta t} - 1}{(\nu - 1)} C\frac{(\sigma\Delta t)^2}{3} \underset{\Delta t \to 0}{\longrightarrow} +\infty
\end{aligned}$$

Quindi sul compatto $[0,1] \times [0, t_{n+1}]$, per $\Delta t \to 0$ l'errore non è limitato $\Rightarrow$ Lo schema di approssimazione non è stabile