

Dispense del corso di Metodi Numerici per l'Ingegneria  
Appunti di teoria sul Metodo agli Elementi Finiti (FEM)

Mario Putti  
Dipartimento di Metodi e Modelli Matematici  
per le Scienze Applicate

16 maggio 2011

# Indice

<b>1</b>	<b>Introduzione</b>	<b>3</b>
<b>2</b>	<b>Equazioni differenziali alle derivate parziali</b>	<b>3</b>
<b>3</b>	<b>Equazioni ellittiche</b>	<b>6</b>
3.1	Il caso monodimensionale . . . . .	6
3.2	Formulazione variazionale . . . . .	8
3.3	Equazioni di Eulero-Lagrange . . . . .	12
3.4	Formulazione agli elementi finiti . . . . .	14
3.4.1	Studio dell'errore e della convergenza del metodo FEM . . . . .	17
3.5	Estensione al caso multidimensionale . . . . .	25
3.5.1	Convergenza del metodo FEM nel caso multidimensionale . . . . .	30
3.5.2	Sulla convergenza e spazi funzionali (spazi di Hilbert) . . . . .	31
3.6	Problema di Neumann: condizioni al contorno naturali e essenziali . . . . .	34
3.7	Tipologia di Elementi finiti . . . . .	37
3.7.1	Elementi isoparametrici . . . . .	38
3.8	Equazione di diffusione e trasporto . . . . .	40
3.8.1	Caso monodimensionale . . . . .	42
<b>A</b>	<b>Appendice A: Discretizzazione alle differenze finite dell'equazione di convezione e diffusione.</b>	<b>46</b>
<b>B</b>	<b>Equazioni paraboliche</b>	<b>48</b>

# 1 Introduzione

## 2 Equazioni differenziali alle derivate parziali

Ci occuperemo in queste note della soluzione numerica di equazioni differenziali alle derivate parziali derivanti da leggi di conservazione. Queste equazioni sono anche chiamate “equazioni in forma di divergenza” dal fatto che l’operatore di divergenza traduce in termini matematici il concetto di conservazione nello spazio. Per esempio, l’equazione di convezione e diffusione, che rappresenta il bilancio di massa di una componente disciolta in acqua che si muove in un campo di moto  $\vec{v}$  si scrive come:

$$\frac{\partial u}{\partial t} = \operatorname{div} D \nabla u - \nabla(\vec{v}u) + f \quad \text{in } \Omega \in \mathbb{R}^3$$

dove  $u$  rappresenta la concentrazione del soluto,  $t$  è il tempo,  $\operatorname{div} = \partial/\partial x + \partial/\partial y + \partial/\partial z$  è l’operatore di divergenza ( $x$ ,  $y$ , e  $z$  sono le 3 direzioni coordinate spaziali,  $D$  è il tensore di dispersione-diffusione,  $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)^T$  è l’operatore di gradiente spaziale (un vettore). Per la risoluzione di tale equazione bisogna specificare le condizioni iniziali e al contorno. Per fare ciò, assumiamo che il contorno  $\Gamma = \partial\Omega$  del dominio  $\Omega$  sia dato dall’unione di tre pezzi  $\Gamma_D$ ,  $\Gamma_N$  e  $\Gamma_C$ , per cui abbiamo:

$$\begin{aligned} u(x, 0) &= u_o(x) & x \in \Omega, & \quad t = 0 & \quad (\text{condizioni iniziali}) \\ u(x, t) &= g_o(x) & x \in \Gamma_D, & \quad t > 0 & \quad (\text{condizioni al contorno di Dirichlet}) \\ D \nabla u(x, t) &= q_N(x) & x \in \Gamma_N, & \quad t > 0 & \quad (\text{condizioni al contorno di Neumann}) \\ \vec{v}u + D \nabla u(x, t) &= q_c(x) & x \in \Gamma_C, & \quad t > 0 & \quad (\text{condizioni al contorno di Cauchy}) \end{aligned}$$

Formalmente questa è un’equazione “parabolica”.

E’ infatti possibile classificare le equazioni differenziali alle derivate parziali (PDE-partial differential equation in analogia con la classificazione delle coniche in geometria di  $\mathbb{R}^n$ . Per fare ciò scriviamo la generica PDE come:

$$F(x, y, z, u, u_x, u_y, u_z, u_{xx}, u_{xy}, u_{yy}, u_{xz}, u_{zz}, u_{yz}) = 0 \quad (1)$$

dove abbiamo indicato con  $u_x$  e  $u_{xx}$  le derivate parziali prime e seconde di  $u(x, y, z)$  lungo la direzione  $x$ , con ovvia estensione alle altre direzioni. Se la  $F$  è una funzione lineare della  $u$  e delle sue derivate, l’equazione si dice lineare, e si può quindi scrivere in  $\mathbb{R}^2$  come:

$$a(x, y, z) + b(x, y, z)u + c(x, y, z)u_x + d(x, y, z)u_y + e(x, y, z)u_{xx} + f(x, y, z)u_{yy} + \dots = 0$$

Tale equazione è a coefficienti variabili, al contrario del caso in cui tutti i coefficienti sono costanti. L’ordine di una PDE è uguale all’ordine della derivata di grado massimo che compare nell’equazione. Ad esempio:



Figura 1: Curva  $\gamma$  con sistema di riferimento locale solidale con la curva

$$\begin{aligned} \Delta u &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 && 2^\circ \text{ grado (eq. di Laplace)} \\ \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} &= 0 && 1^\circ \text{ grado (eq. di trasporto o convezione)} \\ \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} && 2^\circ \text{ grado (eq. di diffusione)} \end{aligned}$$

Prendiamo quindi un'equazione a coefficienti costanti considerando per semplicità un dominio bidimensionale e le derivate di ordine massimo (quelle di ordine inferiore non entrano nella caratterizzazione):

$$au_{xx} + bu_{xy} + cu_{yy} = 0 \quad (2)$$

Cerchiamo una curva  $\gamma(x, y)$  particolare per cui la (2) diventa una equazione alle derivate ordinarie (ODE-ordinary differential equation), e quindi più semplice da risolvere. In realtà vedremo che non in tutti i casi questo procedimento aiuta la soluzione della PDE.

Scriviamo la curva  $\gamma$  in coordinate parametriche definendo una coordinata  $\sigma$  solidale con la curva  $\gamma$  (Fig. 1), che è quindi definita dalle equazioni parametriche:

$$\begin{aligned} x &= x(\sigma) \\ y &= y(\sigma) \end{aligned}$$

Per la regola di derivazione di funzione composta possiamo scrivere le derivate delle quantità sopra definite sul sistema di riferimento locale:

$$\begin{aligned} \frac{du_x}{d\sigma} &= \frac{\partial u_x}{\partial x} \frac{dx}{d\sigma} + \frac{\partial u_x}{\partial y} \frac{dy}{d\sigma} = u_{xx} \frac{dx}{d\sigma} + u_{xy} \frac{dy}{d\sigma} \\ \frac{du_y}{d\sigma} &= \frac{\partial u_y}{\partial x} \frac{dx}{d\sigma} + \frac{\partial u_y}{\partial y} \frac{dy}{d\sigma} = u_{xy} \frac{dx}{d\sigma} + u_{yy} \frac{dy}{d\sigma} \end{aligned}$$

Ricavando  $u_{xx}$  dal sistema precedente e sostituendolo in (2) si ottiene:

$$u_{xy} \left[ a \left( \frac{dy}{dx} \right)^2 - b \frac{dy}{dx} + c \right] - \left( a \frac{du_x}{dx} \frac{dy}{dx} + c \frac{du_y}{dx} + e \frac{dy}{dx} \right) = 0$$

Questa equazione è soddisfatta sulla curva  $\gamma$  (è una riscrittura della PDE su  $\gamma(\sigma)$ ). Quindi, scegliendo  $\gamma(\sigma)$  in maniera tale da azzerare il primo addendo tra parentesi quadra si ottiene una equazione alle derivate ordinarie nelle incognite  $u_x$  e  $u_y$ . Si pone cioè:

$$a \left( \frac{dy}{dx} \right)^2 - b \frac{dy}{dx} + c = 0$$

Quindi, l'equazione della curva  $\gamma(\sigma)$  si ottiene risolvendo l'ODE che si ricava dalla precedente, e cioè:

$$\frac{dy}{dx} = \frac{b \pm \sqrt{b^2 - 4ac}}{2a}$$

Si vede immediatamente che si possono avere diverse famiglie di curve, dette curve caratteristiche, in funzione del segno del discriminante  $b^2 - 4ac$ . In analogia con le coniche in  $\mathbb{R}^n$  si pone:

- $b^2 - 4ac < 0$ : due soluzioni complesse coniugate: equazione di tipo ellittico;
- $b^2 - 4ac = 0$ : una soluzione reale: equazione di tipo parabolico;
- $b^2 - 4ac > 0$ : due soluzioni reali distinte: equazione di tipo iperbolico.

Vediamo subito alcuni esempi di classificazione di PDE:

- Equazione di Laplace:

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

$a = c = 1 \quad b = 0 \Rightarrow b^2 - 4ac < 0$  è un'equazione ellittica.

- equazione delle onde:

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0$$

$a = 1 \quad b = 0 \quad c = -1 \Rightarrow b^2 - 4ac > 0$  è un'equazione iperbolica.

- 

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad 2^\circ \text{ grado (eq. di diffusione)}$$

## 3 Equazioni ellittiche

### 3.1 Il caso monodimensionale

Partiamo con un semplice esempio di un problema di Cauchy monodimensionale, formato da un'equazione differenziale del secondo ordine e due condizioni al bordo del dominio:

**Problema 3.1** (differenziale).

Trovare  $u(x)$  che soddisfa al problema di Cauchy:

$$\begin{aligned} -u''(x) &= f(x), \\ u(0) &= u(1) = 0, \end{aligned} \tag{D}$$

dove abbiamo indicato la derivata prima con  $u' = du/dx$  e quella seconda con  $u'' = d^2u/dx^2$ , e la funzione  $f(x)$  è sufficientemente continua perchè l'equazione abbia senso. Si vede immediatamente che il problema è ben posto (ha soluzione unica): basta integrare due volte e imporre le condizioni al bordo, come mostrato nei seguenti passi:

$$\begin{aligned} - \int u''(x) dx &= \int f(x) dx; \\ -u'(x) &= c_1 + \int f(t) dt; \\ - \int u'(x) dx &= \int c_1 dx + \int \left( \int f(t) dt \right) dx; \\ u(x) &= c_2 + c_1 x - \int_0^x F(t) dt, \end{aligned}$$

dove abbiamo definito il funzionale (funzione di funzione) lineare:

$$F(t) = \int_0^t f(s) ds. \tag{3}$$

Usando le condizioni al bordo si calcolano in maniera univoca le costanti  $c_1$  e  $c_2$ , ottenendo la soluzione:

$$u(x) = x \left( \int_0^1 F(t) dt \right) - \int_0^x F(t) dt,$$

che è evidentemente univocamente determinata, mostrando che il problema di Cauchy (D) è ben posto.

Integrando per parti la (3), si ottiene:

$$\int_0^x F(t) dt = [tF(t)]_0^x - \int_0^x tF'(t) dt = \int_0^x (x-t)f(t) dt,$$

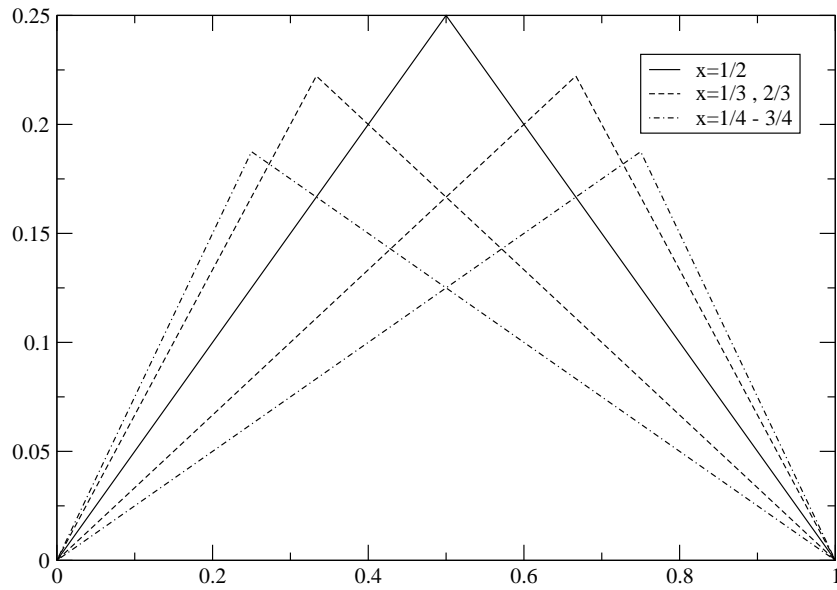


Figura 2: La funzione di Green in corrispondenza a diversi valori di  $x$

da cui la soluzione del problema (D) si può scrivere come:

$$u(x) = x \int_0^1 (1-t)f(t) dt - \int_0^x (x-t)f(t); dt.$$

Definendo la funzione di Green  $G(x, t)$ , data da:

$$G(x, t) = \begin{cases} t(1-x), & \text{se } 0 \leq t \leq x, \\ x(1-t), & \text{se } x \leq t \leq 1. \end{cases} ,$$

la soluzione può essere scritta in forma più compatta come:

$$u(x) = \int_0^1 G(x, t)f(t) dt.$$

La funzione di Green ha le seguenti proprietà:

- è lineare a  $t$  fissato e viceversa;
- è simmetrica, e cioè  $G(x, t) = G(t, x)$ ;
- è continua;
- è non negativa, nulla solo agli estremi dell'intervallo  $[0, 1]$ ;

- $\int_0^1 G(x, t) dt = \frac{1}{2}x(1 - x)$ .

La funzione di Green in questione è rappresentata graficamente in Figura 2 per diversi valori di  $x$ .

Dal punto di vista fisico, tale problema può rappresentare ad esempio una corda elastica vincolata agli estremi e soggetta ad un carico distribuito, nel qual caso  $u(x)$  rappresenta lo spostamento verticale,  $\sigma(x)$  rappresenta la tensione sulla corda, e il coefficiente  $E$  è il modulo di Young. Il problema (D) si può quindi scrivere come:

$$\begin{aligned} \sigma(x) &= Eu'(x) && \text{Legge di Hook;} \\ -\sigma'(x) &= g(x) && \text{Equilibrio elastico;} \\ u(0) &= u(1) = 0 && \text{Condizioni al bordo.} \end{aligned} \tag{4}$$

Oppure si può pensare a  $u(x)$  come la temperatura di una barra soggetta ad una sorgente di calore  $g(x)$ . In questo caso, indicando con  $k$  la conducibilità termica del materiale e con  $q(x)$  il flusso di calore, il modello diventa:

$$\begin{aligned} q(x) &= -ku'(x) && \text{Legge di Fourier;} \\ q(x) &= g(x) && \text{Conservazione dell'energia;} \\ u(0) &= u(1) = 0 && \text{Condizioni al bordo.} \end{aligned} \tag{5}$$

Nello stesso modo si può pensare all'equazione della diffusione di una sostanze disciolta in un liquido, nel qual caso si utilizza la legge di Fick, oppure al flusso dell'acqua in un mezzo poroso, nel qual caso si usa la legge di Darcy, eccetera. Più in generale, si può dire che queste equazioni rappresentano il modello di un moto a potenziale.

### 3.2 Formulazione variazionale

In questo paragrafo, discuteremo brevemente l'approccio variazionale alla soluzione del problema (D), che forma la base per il metodo agli elementi finiti. Per fare questo, introduciamo lo spazio (di funzioni  $v(x)$ )  $V$ , lineare e normato, definito dalla seguente<sup>1</sup>:

$$\begin{aligned} V([0, 1]) &= \{ v(x) : \text{dove } v(x) \text{ è una funzione continua e limitata nell'intervallo } [0, 1], \\ &\quad v'(x) \text{ è una funzione continua a tratti e limitata nell'intervallo } [0, 1], \\ &\quad \text{e } v(0) = v(1) = 0 \}. \end{aligned}$$

In questo spazio si può definire un prodotto interno (o prodotto scalare) tra i suoi elementi:

$$(v, w) = \int_0^1 v(x)w(x) dx,$$

---

<sup>1</sup>Si confrontino le definizioni seguenti con le analoghe definizioni dell'algebra lineare e degli spazi vettoriali.



che definisce il funzionale (o forma) quadratico  $F : V \rightarrow \mathbb{R}$ :

$$F(v) = \frac{1}{2} (v', v') - (f, v) + c.$$

Si possono definire ora il problema di minimizzazione (M) e quello variazionale (V), che vedremo sono in qualche senso da specificare meglio equivalenti al problema differenziale di partenza (D).

**Problema 3.2** (di minimizzazione).

Trovare  $u \in V$  tale che:

$$F(u) \leq F(v) \quad \forall v \in V. \tag{M}$$

**Problema 3.3** (variazionale).

Trovare  $u \in V$  tale che:

$$(u', v') = (f, v) \quad \forall v \in V. \tag{V}$$

**Osservazione 3.4.** Con riferimento al problema elastico mostrato in (4), si noti che la quantità  $F(v)$  rappresenta l'energia potenziale totale associata allo spostamento ammissibile  $v(x) \in V$ ; il termine  $\frac{1}{2} (v', v')$  è l'energia elastica del sistema e il termine  $(f, v)$  il potenziale delle forze esterne. Da questa osservazione si deduce immediatamente che il problema (M) è la formulazione del noto “principio della minimizzazione dell'energia potenziale”, mentre il problema (V) è la formulazione del “principio dei lavori virtuali”.

**Equivalenza tra le varie formulazioni (D), (M), (V).**

**(D) $\Rightarrow$ (V)** *Dimostrazione* - Dimostriamo che la soluzione di (D) è anche soluzione di (V). Per fare questo basta moltiplicare l'equazione differenziale per una funzione arbitraria  $v \in V$  e integrare sul dominio, ottenendo immediatamente:

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

che può essere riscritta usando la notazione di prodotto scalare:

$$-(u'', v) = (f, v).$$

Integrando per parti il primo membro si ottiene:

$$-(u'', v) = -u'(1)v(1) + u'(0)v(0) + (u', v') = (u', v'),$$

che può essere scritto quindi, notando che  $v(0) = v(1) = 0$ :

$$(u', v') = (f, v) \quad \forall v \in V \tag{6}$$

che dimostra la tesi. □

**(V) ⇔ (M)** *Dimostrazione* - Adesso vogliamo dimostrare che (V) e (M) hanno la stessa soluzione. Supponiamo quindi che  $u(x)$  sia soluzione di (V) e sia  $v(x) \in V$ ; calcoliamo la differenza  $w(x) = u(x) - v(x) \in V$ . Abbiamo facilmente che:

$$\begin{aligned} F(v) &= F(u + w) = \frac{1}{2} (u' + w', u' + w') - (f, u + w) + c \\ &= \frac{1}{2} (u', u) - (f, u) + \frac{1}{2} (u', w') - (f, w) + \frac{1}{2} (w', w') \geq F(u), \end{aligned}$$

poichè da (6)  $(u', w') - (f, w) = 0$  e  $(w', w') \geq 0$ . Quindi, siccome  $w$  è una funzione arbitraria, questo dimostra che  $u$  è punto di minimo del funzionale  $F(u)$  e quindi è soluzione del problema (M). Anche il contrario è vero. Infatti, se  $u$  fosse soluzione di (M), allora per ogni  $v \in V$  e  $\epsilon \in \mathbb{R}$ , si ha:

$$F(u) \leq F(u + \epsilon v),$$

poichè  $u + \epsilon v \in V$ . Definiamo la funzione differenziabile

$$g(\epsilon) := F(u + \epsilon v) = \frac{1}{2} (u', u') + \epsilon (u', v') + \frac{\epsilon^2}{2} (v', v') - (f, u) - \epsilon (f, v),$$

che ha un minimo in  $\epsilon = 0$ , da cui  $g'(0) = 0$ . Facendo i conti si ottiene:

$$g'(0) = (u', v') - (f, v),$$

che dimostra che  $u$  è soluzione di (V).

E' anche facile vedere che la soluzione di (V) è unica. Infatti, se  $u_1 \in V$  e  $u_2 \in V$  sono due soluzioni di (V), allora:

$$\begin{aligned} (u_1', v') &= (f, v) & \forall v \in V; \\ (u_2', v') &= (f, v) & \forall v \in V. \end{aligned}$$

Sottraendo membro a membro e prendendo  $v = u_1' - u_2'$ , si ottiene immediatamente che

$$\int_0^1 (u_1' - u_2')^2 dx = 0,$$

da cui, per la linearità dell'operatore di derivata, risulta che  $(u_1 - u_2)(x) = cost$ , e siccome  $u(0) = u(1) = 0$ , tale costante è nulla, da cui la tesi.  $\square$

**(V)⇒(D).** Per dimostrare la tesi ci serve il seguente (caso particolare del) lemma fondamentale del calcolo delle variazioni:

**Lemma 3.5.** Sia  $g \in C^0([0, 1])$  e

$$\int_0^1 g(x) \cdot \phi(x) dx = 0 \quad \forall \phi(x) \in V([0, 1]),$$

allora  $g(x) = 0$  per ogni  $x \in [0, 1]$ .

*Dimostrazione* - Sia  $g(x) \in C^0([0, 1])$ . Una funzione in  $V([0, 1])$  è continua e ha le derivate continue (a tratti). Prendiamo allora  $r(x)$  una funzione che si annulli in  $x = 0$  e  $x = 1$  e che sia positiva in  $(0, 1)$  (ad es.  $r = x(1 - x)$ ). Prendiamo quindi  $\phi(x) = r(x)g(x)$  una funzione che si annulla agli estremi. Ovviamente  $\phi(x) \in V([0, 1])$ . Si ha quindi:

$$0 = \int_0^1 g(x)\phi(x) dx = \int_0^1 r(x)g^2(x) dx$$

La funzione integranda è non negativa, quindi deve essere nulla. Siccome  $r(x) \neq 0$  per  $x \in (a, b)$ , dovrà necessariamente essere  $g(x) = 0$  per ogni  $x \in [a, b]$ . Si noti che tale dimostrazione si può estendere con uno sforzo minimo anche a funzioni di  $\mathbb{R}^d$ .  $\square$

*Dimostrazione* - Ora, per verificare che (V)⇒(D), assumiamo che  $u \in V$  sia soluzione del problema (V). Allora:

$$\int_0^1 u'v' dx - \int_0^1 fv dx = 0 \quad \forall v \in V.$$

Assumendo ora che  $u''$  esista e sia continua, possiamo integrare per parti:

$$\int_0^1 u'v' dx - \int_0^1 fv dx = [u'v]_0^1 - \int_0^1 u''v dx - \int_0^1 fv dx = 0,$$

da cui, usando le condizioni al contorno omogenee, otteniamo:

$$- \int_0^1 (u'' + f)v dx = 0 \quad \forall v \in V.$$

Ora, essendo  $V$  uno spazio di funzioni che sono  $C_0^\infty$  a tratti (escludendo quindi punti isolati dove si possono avere discontinuità nelle derivate, ma tali discontinuità non contribuiscono agli integrali che dobbiamo calcolare avendo supporto di misura nulla), si può applicare (tratto per tratto) il Lemma 3.5, per cui dovrà essere per forza:

$$-u'' + f = 0.$$

□

Abbiamo quindi dimostrato l'equivalenza del problema variazionale con il problema differenziale. Si noti però che questo è vero sotto l'ipotesi di derivata seconda di  $u$  continua. Si può concludere questo paragrafo osservando che tale ipotesi non è richiesta nel problema variazionale, visto che esso richiede solo l'uso di derivate prime. Infatti, utilizzando l'integrazione per parti abbiamo diminuito l'ordine massimo delle derivate presenti nel nostro problema, richiedendo di fatto una grado di continuità inferiore alla nostra soluzione. Riassumendo, le soluzioni del problema differenziale sono sempre anche soluzioni del problema variazionale. Viceversa, le soluzioni del problema variazionale sono soluzioni del problema differenziale se imponiamo una sufficiente continuità nelle derivate seconde.

### 3.3 Equazioni di Eulero-Lagrange

Si può estendere ad un contesto più generale gli argomenti sopra esposti, per arrivare all'equazione di Eulero-Lagrange del calcolo delle variazioni. Di seguito mostriamo la derivazione di tali equazioni nel caso monodimensionale.

Si vuole trovare una funzione  $u(x)$  che soddisfi alle condizioni  $u(0) = u(1) = 0$  e minimizzi il funzionale:

$$F(u) = \int_0^1 L(x, u(x), u'(x)) dx.$$

Assumiamo che  $L$  sia sufficientemente continuo in modo tale che le derivate parziali fatte rispetto a  $x$ ,  $u$  e  $u'$  esistano. Se  $u(x)$  è punto di minimo, allora ogni sua perturbazione deve aumentare il valore di  $F(u)$ , cioè:

$$F(u) \leq F(u + \epsilon v) \quad \forall \epsilon \in \mathbb{R} \text{ e } \forall v(x).$$

Sia quindi  $w(x) = u(x) + \epsilon v(x)$ . Si noti che  $v(x)$  dovrà essere presa in modo tale da soddisfare le condizioni  $v(0) = v(1) = 0$ . Allora:

$$F(w) = F(\epsilon) = \int_0^1 L(\epsilon, x, w(x), w'(x)) dx.$$

La variazione prima di  $F(w)$  è data da:

$$\frac{dF}{d\epsilon} = \frac{d}{d\epsilon} \int_0^1 L(\epsilon, x, w(x), w'(x)) dx = \int_0^1 \frac{d}{d\epsilon} L(\epsilon, x, w(x), w'(x)) dx.$$

Usando la regola di derivazione della funzione composta si ottiene:

$$\begin{aligned} \frac{dL(\epsilon)}{d\epsilon} &= \frac{\partial L}{\partial x} \frac{dx}{d\epsilon} + \frac{\partial L}{\partial w} \frac{\partial w}{\partial \epsilon} + \frac{\partial L}{\partial w'} \frac{\partial w'}{\partial \epsilon} \\ &= \frac{\partial L}{\partial w} v(x) + \frac{\partial L}{\partial w'} v'(x). \end{aligned}$$

Quindi:

$$\frac{dF}{d\epsilon} = \int_0^1 \left( \frac{\partial L}{\partial w} v(x) + \frac{\partial L}{\partial w'} v'(x) \right) dx.$$

Per  $\epsilon = 0$  si ha che  $w = u$  e quindi  $F(w)|_{\epsilon=0}$  deve essere minima. Quindi:

$$\frac{dF}{d\epsilon}|_{\epsilon=0} = \int_0^1 \left( \frac{\partial L}{\partial w} v(x) + \frac{\partial L}{\partial w'} v'(x) \right) dx = 0.$$

Ora, usando il teorema di integrazione per parti otteniamo:

$$\int_0^1 \frac{\partial L}{\partial w} v(x) dx + v(x) \frac{\partial L}{\partial w'} \Big|_0^1 - \int_0^1 \frac{d}{dx} \frac{\partial L}{\partial w'} v(x) dx = 0.$$

Raccogliendo  $v(x)$ , notando che il termine agli estremi si annulla perchè si  $v(0) = v(1) = 0$ , l'applicazione del lemma fondamentale (Lemma 3.5) porta alla cosiddetta equazione di Eulero-Lagrange:

$$\frac{\partial L}{\partial w} v(x) - \frac{d}{dx} \left[ \frac{\partial L}{\partial w'} \right] = 0.$$

Tale equazione determina la condizione necessaria (non sufficiente) per l'esistenza di un minimo del funzionale  $F(u) = \int_0^1 L(x, u, u') dx$ . se  $L(x, u, u')$  è una funzione convessa di  $u$  e  $u'$ , allora l'equazione di Eulero-Lagrange è anche condizione sufficiente.

**Esempio 3.6.** Come esempio, consideriamo il cosiddetto integrale di Dirichlet:

$$D(x, u, u') = \int_0^1 \frac{1}{2} (u')^2 dx.$$

Cerchiamo il minimo di  $D(u)$  nella classe di funzioni continue conservata continua ( $e \in C^1([0, 1])$ ). L'equazione di Eulero-Lagrange la ricaviamo calcolando le derivate di  $L(x, u, u') = (u')^2/2$ , e cioè:

$$\frac{\partial L}{\partial u} = 0; \quad \frac{d}{dx} \left[ \frac{\partial L}{\partial u'} \right] = \frac{d}{dx} \frac{1}{2} (2u') = u''(x),$$

da cui si ricava subito:

$$u''(x) = 0,$$

e cioè l'equazione di Laplace in una dimensione. Come conseguenza, la soluzione dell'equazione di Laplace è proprio il punto di minimo (il funzionale è convesso) del funzionale  $D(x, u, u')$ .

**Esempio 3.7.** Modifichiamo ora il funzionale di Dirichlet scrivendolo come:

$$D(x, u, u') = \int_0^1 \left[ \frac{1}{2} (u')^2 + fu \right] dx.$$

Derivando opportunamente il funzionale le derivate di  $L(x, u, u') = (u')^2/2 + fu$ , otteniamo:

$$\frac{\partial L}{\partial u} = f(x); \quad \frac{d}{dx} \left[ \frac{\partial L}{\partial u'} \right] = \frac{d}{dx} (2u') = u''(x).$$

L'equazione di Eulero-Lagrange coincide quindi con l'equazione di Poisson monodimensionale:

$$-u''(x) = f(x).$$

### 3.4 Formulazione agli elementi finiti

La costruzione di un metodo numerico che risolva il problema (V) si può ricondurre essenzialmente al problema di trovare un opportuno sottospazio  $V_h \subset V$  di dimensione finita. Per esempio, possiamo scegliere di lavorare con funzioni lineari a tratti che interpolino la soluzione vera  $u(x)$  (incognita) in maniera opportuna. Per fare questo consideriamo una “griglia computazionale o mesh” e cioè una partizione dell'intervallo  $I = [0, 1]$  in sotto intervalli aventi estremi di coordinata  $x_i$ ,  $i = 0, 1, \dots, n + 1$ , quindi tali che il generico  $I_i = [x_i, x_{i-1}]$  sotto intervallo abbia lunghezza  $h_i = x_i - x_{i-1}$ , e sia  $h = \max_i h_i$  la dimensione caratteristica della mesh (Fig. 3). Costruiamo lo spazio  $V_h$  come lo spazio delle funzioni  $v$  lineari a tratti, quindi continue con derivata continua a tratti e che appartengono a  $V$ , tali che  $v(0) = v(1) = 0$ . Ricordando l'interpolazione di Lagrange [1], possiamo costruire queste funzioni utilizzando delle funzioni di base per lo spazio  $V_h$  che scegliamo per comodità con supporto<sup>2</sup> in ciascun  $I_i$ , e lineari a tratti. Queste sono quindi univocamente definite dalla seguente condizione:

$$\phi_j(x_i) = \begin{cases} 1, & \text{se } i = j, \\ 0, & \text{se } i \neq j. \end{cases}$$

La funzione  $v \in V_h$  si può quindi costruire tramite una combinazione lineare delle funzioni di base sui valori nodali:

$$v(x) = \sum_{j=1}^n v_j \phi_j(x), \tag{7}$$

dove il coefficiente  $v_j = v(x_j)$  è appunto il valore di  $v$  in ciascun nodo della mesh. Si osservi che utilizzando una griglia con  $n + 2$  nodi (estremi inclusi) possiamo definire  $n$  funzioni di base, per cui lo spazio  $V_h$  risulta avere dimensione  $n$ , oltre che essere ovviamente uno spazio lineare. Si noti anche che  $V_h = \text{span}(\phi_1, \dots, \phi_n)$ .

Possiamo ora scrivere la formulazione del seguente metodo agli *elementi finiti* (FEM, Finite Element Method) nei seguenti modi equivalenti:

---

<sup>2</sup>Il supporto di una funzione è quel sottoinsieme del dominio dove la funzione è diversa da zero.

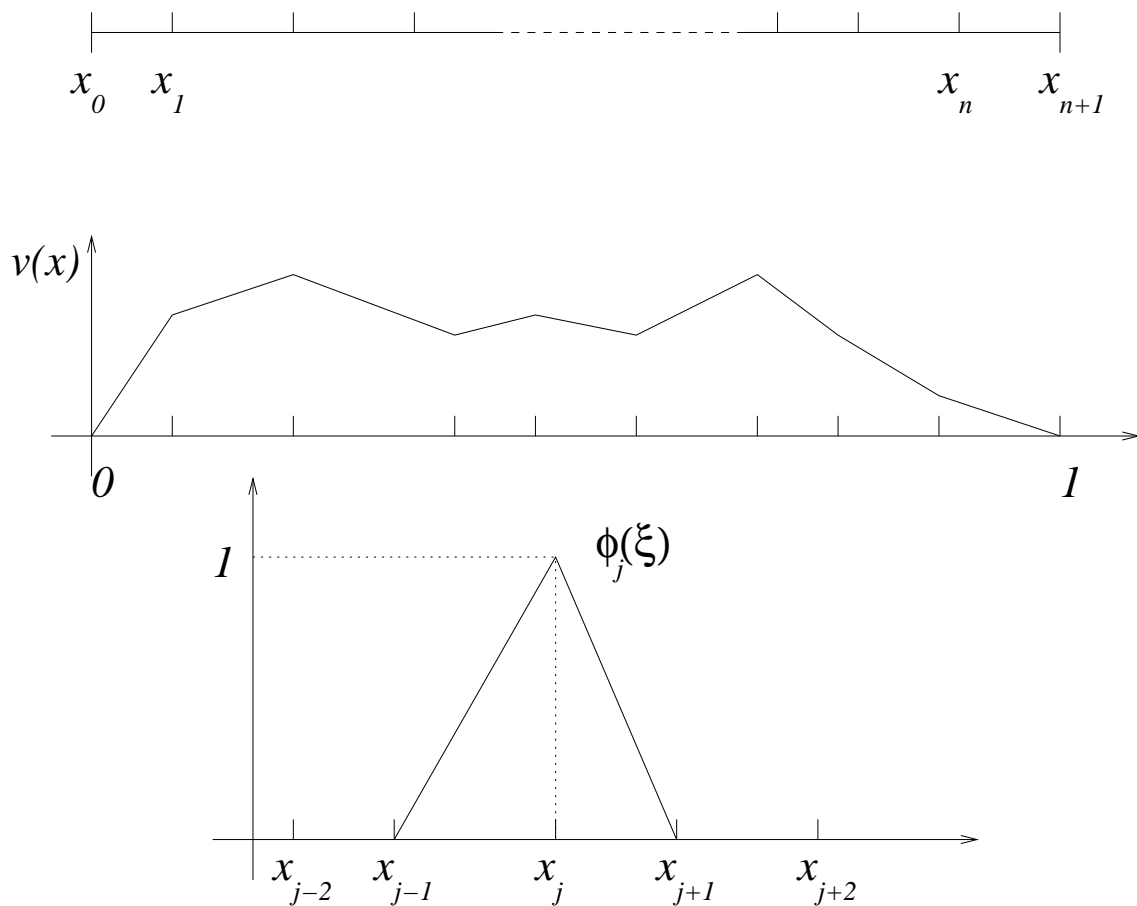


Figura 3: Griglia o mesh computazionale nell'intervallo  $I = [0, 1]$  (in alto); esempio di una funzione  $v(x) \in V_h$  (centro); esempio di una funzione di base  $\phi_j(x)$  (in basso).

**Problema 3.8** (metodo di Ritz).

Trovare  $u_h \in V_h$  tale che:

$$F(u_h) \leq F(v) \quad \forall v \in V_h. \quad (\text{Mh})$$

**Problema 3.9** (metodo di Galerkin).

Trovare  $u_h \in V_h$  tale che:

$$(u'_h, v') = (f, v) \quad \forall v \in V_h. \quad (\text{Vh})$$

Usando la combinazione lineare di funzioni di base per esprimere la generica funzione  $v \in V_h$ , e cioè l'eq. (7, si vede subito che se  $u_h \in V_h$  soddisfa l'equazione (Vh), in particolare soddisfa anche:

$$(u'_h, \phi'_i) = (f, \phi_i) \quad i = 1, \dots, n, \quad (8)$$

e siccome anche  $u_h$  può essere scritta come combinazione lineare delle funzioni di base, e cioè:

$$u_h(x) = \sum_{j=1}^n u_j \phi_j(x) \quad u_j = u_h(x_j), \quad u'_h(x) = \sum_{j=1}^n u_j \phi'_j(x) \quad (9)$$

si ottiene immediatamente:

$$\sum_{j=1}^n (\phi'_i, \phi'_j) u_j = (f, \phi_i) \quad i = 1, \dots, n, \quad (10)$$

che è un sistema lineare  $n \times n$ . In forma matriciale esso può essere scritto come:

$$Au = b$$

dove la matrice  $A_{[n \times n]} = \{a_{ij}\} = \{(\phi'_i, \phi'_j)\}$  è detta matrice di *rigidezza*, il vettore delle incognite è  $u_{[n \times 1]} = \{u_i\}$  e il vettore termine noto è  $b_{[n \times 1]} = \{b_i\} = \{(f, \phi_i)\}$ .

Gli elementi  $a_{ij}$  e  $b_i$  sono facilmente calcolabili. Infatti, si osservi che  $a_{ij} = 0$  per  $|i - j| > 1$ , essendo in tale caso i supporti di  $\phi_i$  e di  $\phi_j$  hanno intersezione nulla, per cui  $\phi_i(x)\phi_j(x) = 0$  e anche  $\phi'_i(x)\phi'_j(x) = 0$ . Quindi per  $i = 1, \dots, n$ :

$$a_{ii} = (\phi'_i, \phi'_i) = \int_{x_{i-1}}^{x_i} \frac{1}{h_i^2} dx + \int_{x_i}^{x_{i+1}} \frac{1}{h_{i+1}^2} dx = \frac{1}{h_i} + \frac{1}{h_{i+1}},$$

e per  $i = 2, \dots, n$ :

$$a_{i,i-1} = a_{i-1,i} = (\phi'_i, \phi'_{i-1}) = (\phi'_{i-1}, \phi'_i) = - \int_{x_{i-1}}^{x_i} \frac{1}{h_i^2} dx = -\frac{1}{h_i}.$$



La matrice  $A$  è quindi tridiagonale e simmetrica. Si può dimostrare anche che  $A$  è anche positiva definita. Infatti, per ogni  $v(x) = \sum_{j=1}^n c_j \phi_j(x)$  si ha immediatamente che:

$$\sum_{i,j=1}^n c_i (\phi'_i, \phi'_j) c_j = \left( \sum_{i=1}^n c_i \phi'_i, \sum_{j=1}^n c_j \phi'_j \right) = (v', v') \geq 0.$$

Nella precedente, si verifica l'uguaglianza solo nel caso in cui  $v'(x) \equiv 0$ , e cioè  $v(x) = \text{cost}$ , ma tale costante risulta nulla perchè  $v(0) = v(1) = 0$ . Quindi, raggruppando in un vettore  $c = c_i$  le costanti della generica combinazione lineare (7), possiamo scrivere:

$$\left( \sum_{i=1}^n c_i \phi'_i, \sum_{j=1}^n c_j \phi'_j \right) = \langle c, Ac \rangle > 0 \quad \forall c \in \mathbb{R}^n, c \neq 0,$$

che dimostra la positiva definizione di  $A$ , e quindi anche che il sistema ha soluzione unica. Si noti che la matrice  $A$  è sparsa, cioè ha solo pochi elementi non nulli. Questa è una caratteristica importante, che dipende in maniera fondamentale dal fatto che le funzioni di base hanno supporto compatto e locale. Nel nostro caso monodimensionale infatti ciascuna di esse è diversa da zero solo in due sotto intervalli contigui. Questa caratteristica dovrà essere mantenuta in tutti gli schemi agli elementi finiti, anche in dimensione spaziale maggiore di uno.

Nel caso speciale di griglia uniforme, e cioè con  $h_i = h = 1/(n+1)$ , e una funzione forzante costante  $f(x) = \text{cost}$  il sistema ha la forma speciale:

$$\frac{1}{h^2} \begin{bmatrix} 2 & -1 & 0 & \dots & \dots & \dots & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & \dots & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 & -1 & 2 & -1 \\ 0 & \dots & \dots & \dots & \dots & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix},$$

con  $b_i = f$ .

### 3.4.1 Studio dell'errore e della convergenza del metodo FEM

**Consistenza, convergenza e stabilità** La convergenza di del metodo FEM <sup>3</sup> parte dai concetti di consistenza e stabilità di uno schema numerico. Si dice che uno schema è consistente

---

<sup>3</sup>Lo studio della convergenza di uno schema di discretizzazione spaziale è fondamentale non solo da un punto di vista teorico, ma anche per procedere a confronti tra i diversi schemi e quindi poter scegliere lo schema ottimale per il problema che si deve risolvere. Inoltre, da un confronto tra la convergenza teorica e quella sperimentale, non solo su problemi sufficientemente semplici per avere una soluzione analitica, necessaria per poter calcolare l'errore, si hanno utilissime indicazioni sulla correttezza dell'implementazione particolare che si sta utilizzando.

se l'errore commesso dallo schema avendo sostituito la soluzione esatta tende a zero al tendere a zero del passo di discretizzazione. Uno schema è stabile se piccole variazioni dei dati portano a piccole variazioni nei risultati numerici.

Nel nostro caso, indicando con  $L(u, f) = 0$  l'equazione alle derivate parziali, dove  $L$  è l'operatore differenziale,  $u$  la soluzione "vera",  $f$  i dati del problema, e con  $L_h(u_h, f_h) = 0$  lo schema numerico, con  $L_h$  l'operatore discreto,  $u_h$  la soluzione numerica, e  $f_h$  l'approssimazione numerica dei dati del problema, si dice che lo schema numerico "converge" se

$$\|u - u_h\| \rightarrow 0 \quad h \rightarrow 0,$$

dove  $\|\cdot\|$  è una norma funzionale opportuna.

Un metodo numerico di discretizzazione si dice "consistente" se

$$L_h(u, g) \rightarrow 0 \quad h \rightarrow 0,$$

e si dice "fortemente consistente" se:

$$L_h(u, g) = 0 \quad \forall h.$$

Spesso non è agevole provare la convergenza di uno schema direttamente, ma si utilizza il risultato fondamentale (teorema di equivalenza) per cui uno schema consistente è convergente se e solo se è anche stabile [4].

D'altro canto, lo studio diretto della convergenza è utile non solo perchè si riesce a verificare che uno schema funziona, ma anche perchè si riesce a quantificare la velocità di convergenza, (in realtà l'ordine con cui l'errore tende a zero al tendere a zero di  $h$ ) e quindi dall'analisi parallela del costo computazionale del metodo si riesce a prevedere i tempi di calcolo necessari per risolvere un dato problema con un prefissato errore.

**Stima dell'errore del metodo FEM** Sia  $u \in V$  la soluzione al problema (D) e  $u_h \in V_h$  la soluzione al problema (Vh). Siccome la (V) vale per ogni  $v \in V$  e  $V_h \subset V$ , la stessa equazione vale anche per le funzioni  $v \in V_h$ . Sostituendo la stessa  $v \in V_h$  nella (V) e nella (Vh), e sottraendo membro a membro si ottiene subito:

$$\begin{aligned} (u', v') &= (f, v) & \forall v \in V_h \\ (u'_h, v') &= (f, v) & \forall v \in V_h \\ ((u' - u'_h), v') &= 0 & \forall v \in V_h, \end{aligned} \tag{11}$$

che dimostra che lo schema è fortemente consistente.

Ora dobbiamo usare la nozione di norma di funzione. Pensando alla norma euclidea di vettori, una semplice sua estensione fornisce la seguente norma per funzioni:

$$\|w\| = (w, w)^{\frac{1}{2}} = \left( \int_0^1 w^2 dx \right)^{\frac{1}{2}}.$$

Si può verificare facilmente che il simbolo  $(v, w) = \int_0^1 vw \, dx$  definisce un prodotto scalare tra due funzioni  $v$  e  $w$  (soddisfa alle proprietà fondamentali), e dà luogo alla norma funzionale sopra definita. Si ricorda in particolare la proprietà definita come disuguaglianza di Cauchy:

$$| (v, w) | \leq \|v\| \|w\|$$

A questo punto è facile dimostrare il seguente risultato:

$$\|(u - u_h)'\| \leq \|(u - v)'\| \quad \forall v \in V_h \quad (12)$$

*Dimostrazione* - Assumiamo  $\|(u - u_h)'\| \neq 0$ . Nel caso la norma fosse nulla, il risultato seguirebbe direttamente.

Sia quindi  $v \in V_h$  una funzione arbitraria e chiamiamo  $w = u_h - v$ , una funzione anch'essa appartenente a  $V_h$  e arbitraria. Sfruttando il fatto che dalla (11) il termine  $((u - u_h)', w') = 0$  e quindi può essere sommato arbitrariamente, si ottiene:

$$\begin{aligned} \|(u - u_h)'\|^2 &= ((u - u_h)', (u - u_h)') + ((u - u_h)', w') \\ &= ((u - u_h)', (u - u_h + w)') = ((u - u_h)', (u - v)') \\ &\leq \|(u - u_h)'\| \|(u - v)'\|. \end{aligned}$$

Il risultato segue dividendo per  $\|(u - u_h)'\|$ , che è stato assunto non nullo. □

Si può anche dimostrare che  $\|v\| \leq \|v'\|$  per ogni  $v \in V_h$ , ovvero:

$$\int_0^1 v^2 \, dx \leq \int_0^1 (v')^2 \, dx \quad \forall v \in V_h.$$

Si noti che questo è vero perchè si è richiesto nella definizione di  $V_h$  che le funzioni assumano valori nulli agli estremi dell'intervallo. Infatti:

$$v(x) = v(0) + \int_0^x v'(t) \, dt = \int_0^x v'(t) \, dt,$$

da cui, usando ancora la disuguaglianza di Cauchy:

$$| v(x) | \leq \int_0^1 | v' | \, dx \leq \left( \int_0^1 1^2 \, dx \right)^{\frac{1}{2}} \left( \int_0^1 | v' |^2 \, dx \right)^{\frac{1}{2}} \leq \left( \int_0^1 | v' |^2 \, dx \right)^{\frac{1}{2}},$$

da cui, integrando tra 0 e 1, si ottiene infine:

$$\int_0^1 | v(x) |^2 \, dx \leq \int_0^1 \left( \int_0^1 | v'(x) |^2 \, dx \right) \, dy = \int_0^1 | v'(x) |^2 \, dx.$$

L'applicazione di tale risultato alla funzione  $v = u - u_h$  dimostra che

$$\|u - u_h\| \leq \|(u - u_h)'\| \leq \|(u - v)'\| \quad \forall v \in V_h \quad (13)$$

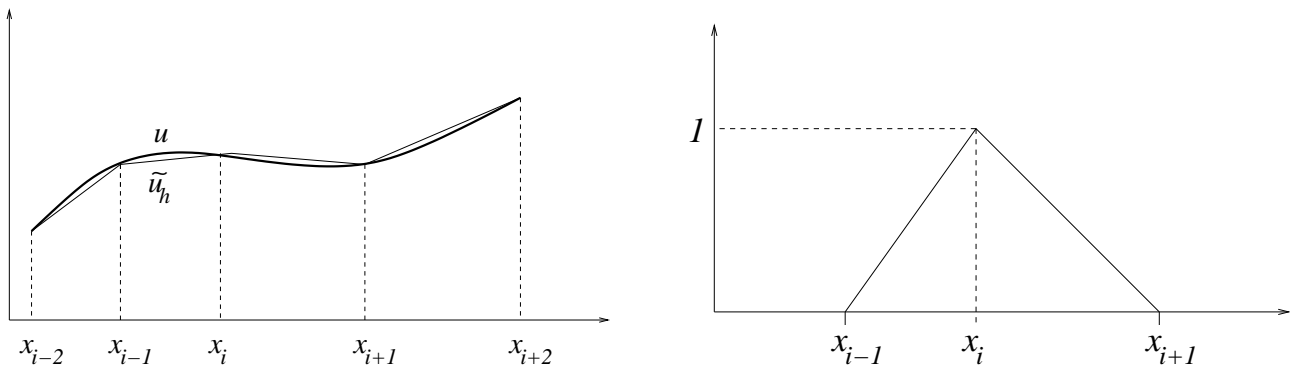


Figura 4: Interpolante  $\tilde{u}_h$  (a sinistra), funzione di base  $\phi_i(x)$  (a destra).

**Osservazione 3.10.** Il risultato precedente ci dice che  $u'_h$  è la migliore approssimazione di  $u'$  in  $V_h$ , essendo l'errore minore rispetto a quello commesso da qualsiasi altra funzione  $v \in V_h$ .

Il risultato precedente ci dice che se riusciamo a dare una maggiorazione della norma della differenza tra  $u$  e una qualsiasi funzione  $v \in V_h$ , questa maggiora anche l'errore del metodo FEM. Quindi si può ricavare una stima quantitativa dell'errore  $\|(u - u_h)'\|$  passando attraverso la stima dell'errore commesso andando a prendere al posto di  $v$  in (12) una funzione opportuna. Per fare questo, scegliamo di lavorare con la funzione  $\tilde{u}_h \in V_h$ , una “interpolante” lineare a tratti che interpola la  $u$  sui punti della griglia. Si dice che la funzione  $\tilde{u}_h$  è una interpolante di  $u(x)$ , ovvero  $\tilde{u}_h$  interpola  $u(x)$  nei nodi  $x_i$ ,  $i = 0, \dots, n + 1$ , se valgono le seguenti relazioni:

$$\tilde{u}_h(x_i) = u(x_i) \quad i = 0, \dots, n + 1.$$

Il concetto è mostrato graficamente in Figura 4, a sinistra.

E' facile definire tale interpolante utilizzando i polinomi di Lagrange [1, 4]. Di seguito, per completezza, ricaviamo tutti i risultati senza ricorrere a tali polinomi. Si vuole dunque interpolare una funzione  $v(x)$  generica tramite un polinomio lineare a tratti. Tale polinomio si può scrivere come:

$$P_1(x) = \sum_{i=1}^n a_i \phi_i(x).$$

La generica funzione di base nel nodo  $i$ -esimo è data da:

$$\phi_i(x) = \begin{cases} \frac{x-x_i}{x_i-x_{i-1}}, & \text{se } x_{i-1} \leq x \leq x_i, \\ \frac{x_{i+1}-x}{x_{i+1}-x_i}, & \text{se } x_i \leq x \leq x_{i+1}. \end{cases}$$

Tale funzione, mostrata in Figura 4 (a destra), è effettivamente una funzione di base dell'interpolazione. Le seguenti proprietà, che mostrano che  $P_1(x)$  è proprio il polinomio cercato, sono

facilmente verificabili:

$$\phi_i(x) = \begin{cases} 1, & \text{se } x = x_i, \\ 0, & \text{se } x = x_j, \quad i \neq j. \end{cases}$$

$$P_1(x_i) = a_i = v(x_i)$$

$$P_1'(x_i) = v'(x_i)$$

Definiamo ora l'errore di interpolazione  $e(x) = v(x) - P_1(x)$ . Si noti che essendo  $P_1(x)$  lineare a tratti,  $P_1''(x) = 0$  in tutto l'intervallo  $[0, 1]$ . E' chiaro che  $e(x_i) = 0$  in tutti i punti di appoggio (i nodi della griglia)  $x_i$ ,  $i = 0, \dots, n+1$ . Quindi, per il teorema di Rolle, esistono  $n$  punti  $\eta_i$ ,  $i = 1, \dots, n$  con  $\eta_i \in [x_i, x_{i+1}]$  dove  $e'(\eta_i) = 0$ . Possiamo quindi scrivere per  $x_i \leq x \leq x_{i+1}$ :

$$e'(x) = \int_{\eta_i}^x e''(t) dt = \int_{\eta_i}^x v''(t) dt,$$

da cui:

$$\begin{aligned} |e'(x)| &\leq \int_{x_i}^{x_{i+1}} |v''(t)| dt = \int_{x_i}^{x_{i+1}} 1 \cdot |v''(t)| dt \leq (\text{per la disuguaglianza di Cauchy}) \\ &\leq \left( \int_{x_i}^{x_{i+1}} 1^2 dt \right)^{\frac{1}{2}} \left( \int_{x_i}^{x_{i+1}} |v''(t)|^2 dt \right)^{\frac{1}{2}} = h^{\frac{1}{2}} \left( \int_{x_i}^{x_{i+1}} |v''(t)|^2 dt \right)^{\frac{1}{2}}, \end{aligned}$$

da cui si ricava subito:

$$|e'(x)|^2 \leq h \left( \int_{x_i}^{x_{i+1}} |v''(t)|^2 dt \right). \quad (14)$$

Integrando la precedente tra  $x_i$  e  $x_{i+1}$ , si ha:

$$\int_{x_i}^{x_{i+1}} |e'(x)|^2 dx \leq h^2 \int_{x_i}^{x_{i+1}} |v''(t)|^2 dt.$$

Per valutare  $e(x)$ , si noti che  $e(x) = \int_{x_i}^x e'(t) dt$ . Quindi, usando la (14) e integrando, si ottiene:

$$|e(x)| \leq h^{\frac{3}{2}} \left( \int_{x_i}^{x_{i+1}} |v''(t)|^2 dt \right)^{\frac{1}{2}},$$

da cui si ricava:

$$\int_{x_i}^{x_{i+1}} |e(x)|^2 dx \leq h^4 \int_{x_i}^{x_{i+1}} |v''(t)|^2 dt.$$

Sommando ora su tutti i sotto intervalli che formano la griglia computazionale si ottengono le seguenti stime dell'errore di interpolazione:

$$\left( \int_0^1 |e(x)|^2 \right)^{\frac{1}{2}} \leq h^2 \left( \int_0^1 |v''(x)|^2 dx \right)^{\frac{1}{2}}$$

$$\left( \int_0^1 |e'(x)|^2 \right)^{\frac{1}{2}} \leq h \left( \int_0^1 |v''(x)|^2 dx \right)^{\frac{1}{2}}$$

ovvero, in termini di norme:

$$\|v - P_1(x)\| \leq h^2 \|v''(x)\|$$

$$\|v' - P_1'(x)\| \leq h \|v''(x)\|$$

Utilizzando ora (12) e (13) si ottengono le seguenti stime dell'errore:

$$\|u - u_h\| \leq h \|u''\| \tag{15}$$

$$\|(u - u_h)'\| \leq h \|u''\| \tag{16}$$

che dimostrano che, se la derivata seconda della soluzione vera ha norma limitata, l'errore dello schema FEM converge a zero al tendere a zero del passo di griglia  $h$ . Si noti che con qualche sforzo in più si può dimostrare che l'errore sulla soluzione  $u(x)$  converge a zero quadraticamente, anzichè linearmente come l'errore sulla derivata, e cioè:

$$\|u - u_h\| \leq h^2 \|u''\|. \tag{17}$$

Di nuovo se la soluzione vera non ha derivata seconda limitata l'ordine quadratico è perso.

**Osservazione 3.11.** Per poter apprezzare quest'ultimo fatto, bisognerebbe ricorrere alla definizione di integrale di Lebesgue. Tale definizione esula dagli ambiti di queste note e si rimanda il lettore interessato a testi più specializzati [5, 3]. Per il momento basti pensare a funzioni continue, con derivate sufficientemente lisce non necessariamente limitate il cui quadrato abbia integrale finito (funzioni di quadrato sommabile), in modo tale che i prodotti scalari e le norme integrali usate siano ben definite.

**Osservazione 3.12.** Dalla stima dell'errore si può ricavare una stima dell'indice di condizionamento spettrale della matrice di rigidezza  $A$  (simmetrica e definita positiva) del metodo FEM. Infatti si ha:

$$\kappa(A) = \frac{\lambda_1}{\lambda_N} = Ch^{-2}$$

dove si è indicato rispettivamente con  $\lambda_1$  e con  $\lambda_N$  gli autovalore massimo e minimo di  $A$ , e la costante  $C$  non dipende da  $h$ . Se si utilizzasse il metodo del gradiente coniugato per risolvere il sistema lineare, sarebbe possibile stimare l'indice di condizionamento e quindi il numero di iterazioni necessarie al metodo del gradiente coniugato per ottenere una soluzione con una prefissata tolleranza. Analogamente, è possibile tramite questo risultato stabilire il variare del numero di iterazioni impiegate dal gradiente coniugato per arrivare alla convergenza al variare della dimensione della mesh.

**Alcuni esempi semplici** Si consideri il problema di Cauchy:

$$\begin{aligned} -u''(x) &= q & x \in [0, 1], \\ u(0) &= u(1) = 0. \end{aligned}$$

Si consideri il funzionale:

$$F(u) = \int_0^1 \left[ \frac{1}{2}(u')^2 - qu \right] dx,$$

con una soluzione approssimata data da:

$$u_n(x) = \sum_{j=1}^n a_j \phi_j(x).$$

La minimizzazione del funzionale (metodo di Ritz) richiede che la soluzione sia un punto di stazionarietà per  $F(u)$ , talchè si ottiene un sistema lineare (uguale a quello che si otterrebbe con l'approccio di Galerkin), la cui  $i$ -esima equazione è data da:

$$\frac{\partial F}{\partial a_i} = \int_0^1 \left[ \left( \sum_{j=1}^n a_j \phi_j'(x) \right) \phi_i'(x) - q \phi_i(x) \right] dx = 0.$$

Dobbiamo ora scegliere le funzioni di base  $\phi_i(x) \in V_h$ .

**Esempio 3.13.** Scegliamo come funzioni di base la base canonica dello spazio dei polinomi di grado  $n$ :

$$\phi_i(x) = x^i \quad i = 0, 1, \dots, n-1.$$

La nostra soluzione numerica può quindi essere scritta:

$$u_n(x) = x(x-1) \sum_{i=1}^n a_i x^{i-1}$$

dove i primi due monomi sono stati inseriti per poter soddisfare le condizioni al contorno. Si noti che si hanno le seguenti funzioni:

$$\begin{aligned} \phi_1(x) &= x(x-1) \\ \phi_1'(x) &= 2x-1 \\ &\dots \quad \dots \\ \phi_i(x) &= x(x-1)x^{i-1} = x^{i-1} - x^i \\ \phi_i'(x) &= (i+1)x^i - ix^{i-1} \\ &\dots \quad \dots \end{aligned}$$

Per  $n = 1$  si ha che  $i = 1$ , da cui:

$$u_n(x) = x(x - 1)a_1$$

$$u'_n(x) = 1(x - 1)a_1$$

$$\begin{aligned} \frac{\partial F}{\partial a_1} &= \int_0^1 [a_1 2(x - 1)^2 - qx(x - 1)] dx \\ &= \int_0^1 [a_1(4x^2 + 1 - 4x) - qx^2 + qx] dx = 0, \end{aligned}$$

da cui si ricava immediatamente  $a_1 = -q/2$ , che sostituito nella soluzione numerica mi fornisce:

$$u_n(x) = -x(x - 1)\frac{q}{2}.$$

Derivando due volte si vede che immediatamente la  $u_n(x)$  soddisfa l'equazione differenziale di partenza, e quindi è la sua soluzione esatta e certamente  $a_2 = a_3 = \dots = a_n = 0$ .

**Esempio 3.14.** Sia

$$u_n(x) = \sum_{i=1}^n a_i \sin(i\pi x)$$

da cui le funzioni di base sono individuate da:

$$\phi_i(x) = \sin(i\pi x) \quad \phi'_i(x) = i\pi \cos(i\pi x).$$

Il sistema lineare (metodo di Ritz) diventa:

$$\frac{\partial F}{\partial a_i} = \int_0^1 \left[ \left( \sum_{j=1}^n a_j \phi'_j(x) \right) \phi'_i(x) - q\phi_i(x) \right] dx = 0,$$

da cui, risolvendo per  $a_1$  nel caso  $n = 1$ , si ottiene:

$$\frac{\partial F}{\partial a_1} = \int_0^1 [a_1 \pi^2 \cos^2(\pi x) - -q \sin(\pi x)] dx = 0,$$

che fornisce la seguente espressione:

$$a_1 = \frac{\int_0^1 q \sin(\pi x) dx}{\int_0^1 \pi^2 \cos(\pi x) dx} = \frac{4}{\pi^3} q.$$

La soluzione numerica è quindi data da:

$$u_n(x) = \frac{4}{\pi^3} q \sin(\pi x)$$

Un confronto tra la soluzione numerica e quella analitica è data nella seguente tabella:



$x$	$u/q$	$u_n/q$
0.00	0.00	0.00
0.25	0.09375	0.09122
0.50	0.125	0.12901
0.75	0.09375	0.09122
1.00	0.00	0.00

**Esempio 3.15.** Sia:

$$u_n(x) = \sum_{i=1}^n a_i \sin(2\pi i x)$$

In questo caso risulta  $a_1 = a_2 = \dots = a_n = 0$ . In realtà lo spazio  $V_h$  generato dalle funzioni di base  $\phi_i(x) = \sin(2\pi i x)$  non contiene la soluzione analitica del problema, e quindi lo schema calcola la soluzione identicamente nulla.

### 3.5 Estensione al caso multidimensionale

Consideriamo ora l'equazione di Poisson nel caso  $d$ -dimensionale, con  $d = 2$  o  $3$ :

**Problema 3.16** (differenziale).

Trovare  $u(x)$  che soddisfa al problema al contorno:

$$\begin{aligned} -\Delta u &= f(x), & x \in \Omega \subset \mathbb{R}^d \\ u(x) &= 0 & x \in \Gamma, \end{aligned} \tag{18}$$

dove  $\Omega \subset \mathbb{R}^d$  è un dominio limitato di  $\mathbb{R}^d = \{x = [x_1, x_2, \dots, x_d], x_i \in \mathbb{R}\}$  avente contorno  $\Gamma = \partial\Omega$ , assunto sufficientemente liscio, e  $\Delta$  è l'operatore Laplaciano definito da:

$$\Delta = \operatorname{div} \nabla = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}.$$

Lo strumento principale che utilizzeremo in questo paragrafo è la formula (o lemma) di Green, ovvero il procedimento di integrazione per parti multidimensionale. Per ricavare la formula di Green si parte dal teorema della divergenza (o di Gauss) che per una funzione vettoriale (detto anche campo vettoriale) generica  $\vec{F}(x) \in \Omega$  prende la forma:

$$\int_{\Omega} \operatorname{div} \vec{F} \, dx = \int_{\Gamma} \vec{F} \cdot \vec{n} \, ds,$$

dove  $\vec{n}$  è il vettore normale unitario esterno a  $\Gamma$ ,  $dx$  denota la misura di volume su  $\Omega$  (in  $\mathbb{R}^d$ ) e  $ds$  la misura di superficie su  $\Gamma$  (in  $\mathbb{R}^{d-1}$ ), e  $\vec{F} \cdot \vec{n}$  indica il prodotto scalare tra due vettori di  $\mathbb{R}^d$ .

Applichiamo il teorema di Gauss ad un campo vettoriale opportuno,  $\vec{F} = v\vec{q}$ , dato dal prodotto di un campo vettoriale  $\vec{q}(x)$  per una funzione  $v(x)$ . Utilizzando la regola di derivazione del prodotto dopo aver sviluppato componente per componente il prodotto scalare, si ottiene:

$$\int_{\Omega} \nabla v \cdot \vec{q} \, dx = \int_{\Gamma} v \vec{q} \cdot \vec{n} \, ds - \int_{\Omega} v \operatorname{div} \vec{q} \, dx.$$

Nel caso particolare in cui  $\vec{q} = \nabla w$ , si ottiene la prima identità di Green:

$$\int_{\Omega} \nabla v \cdot \nabla w \, dx = \int_{\Gamma} v \nabla w \cdot \vec{n} \, ds - \int_{\Omega} v \Delta w \, dx,$$

che intuitivamente può essere pensata come una formula di integrazione per parti in domini multidimensionali, notando che  $v$  è una primitiva di  $\nabla v$ , e  $\Delta w = \operatorname{div} \nabla w$  è la derivata di  $\nabla w$ .

La formulazione variazionale per il problema (18) si scrive nel modo seguente:

**Problema 3.17** (variazionale).

Trovare  $u \in V$  tale che:

$$a(u, v) = (f, v) \quad \forall v \in V, \tag{19}$$

dove:

$$\begin{aligned} a(u, v) &= \int_{\Omega} \nabla u \cdot \nabla v \, dx \\ (f, v) &= \int_{\Omega} f v \, dx \\ V &= \left\{ v(x) : v \text{ è continua in } \Omega, \frac{\partial v}{\partial x_i} \text{ sono continue in } \Omega \forall i, \text{ e } v(x) = 0 \text{ per } x \in \Gamma \right\}. \end{aligned}$$

Per vedere come tale formulazione variazionale segue dal problema differenziale di partenza, moltiplichiamo la (18) per una funzione test arbitraria  $v(x) \in V$  e integriamo su  $\Omega$ . Usando la formula di Green si ottiene:

$$(f, v) = - \int_{\Omega} \Delta u \, dx = - \int_{\Gamma} v \nabla u \cdot \vec{n}; \, ds + \int_{\Omega} \nabla u \cdot \nabla v \, dx = a(u, v),$$

dove l'integrale al bordo è nullo perchè  $v(x) = 0$  per  $x \in \Gamma$ . In modo del tutto analogo al caso mono-dimensionale, si vede che:

- la soluzione del problema variazionale è soluzione del problema differenziale se si assume che  $u(x)$  sia sufficientemente regolare;
- il problema variazionale è equivalente al seguente problema di minimizzazione:

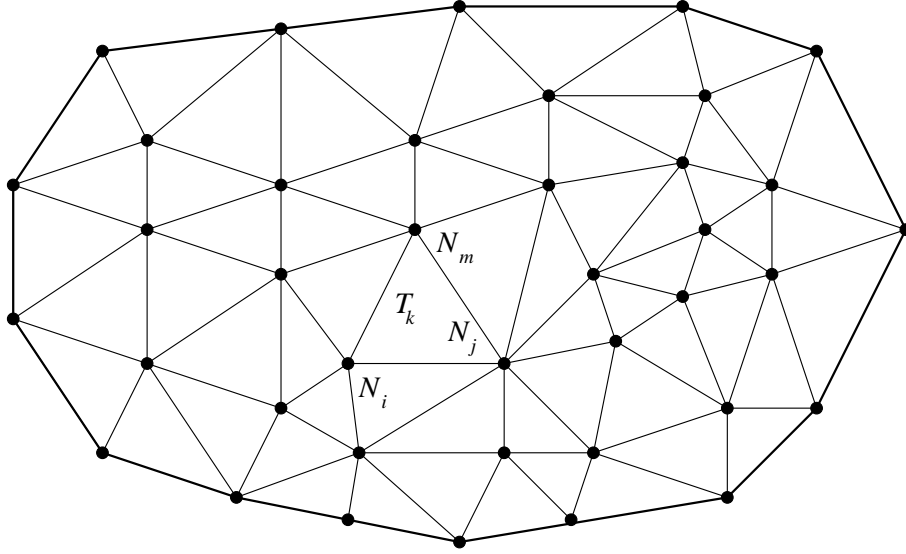


Figura 5: Un esempio di una triangolazione ammissibile di un dominio  $\Omega$ . Il contorno  $\Gamma = \partial\Omega$  è evidenziato con la linea più spessa.

**Problema 3.18** (di minimizzazione).

Trovare  $u \in V$  tale che:

$$F(u) \leq F(v) \quad \forall v \in V, \quad (20)$$

dove:

$$F(v) = \frac{1}{2}a(u, v) - (f, v).$$

Bisogna ora definire opportunamente le funzioni di base, e per fare ciò bisogna prima costruire la griglia computazionale, cioè un opportuno partizionamento del dominio  $\Omega$ . Nel caso bi-dimensionale (i.e.,  $d = 2$ ,  $\Omega \subset \mathbb{R}^2$ ) possiamo procedere definendo una triangolazione che partiziona  $\Omega$  in un insieme  $\mathcal{T}_h$  di triangoli  $T_k$  con le seguenti proprietà:

- $\mathcal{T}_h$  è formata da  $n$  nodi (i vertici dei triangoli, indicati con il simbolo  $N_i$ ,  $i = 1, \dots, n$ , che è sostanzialmente il vettore delle coordinate dell' $i$ -esimo nodo) e  $m$  triangoli (indicati con  $T_k$ ,  $k = 1, \dots, m$ );
- $\Omega = \bigcup_{T_k \in \mathcal{T}_h} T_k = T_1 \cup T_2 \dots \cup T_m$ ;
- $T_i \cap T_j = e_{ij}$ ,  $i \neq j$ , dove  $e_{ij}$  indica il lato in comune ai triangoli  $T_i$  e  $T_j$ ;
- nessun vertice di nessun triangolo giace all'interno di un lato;

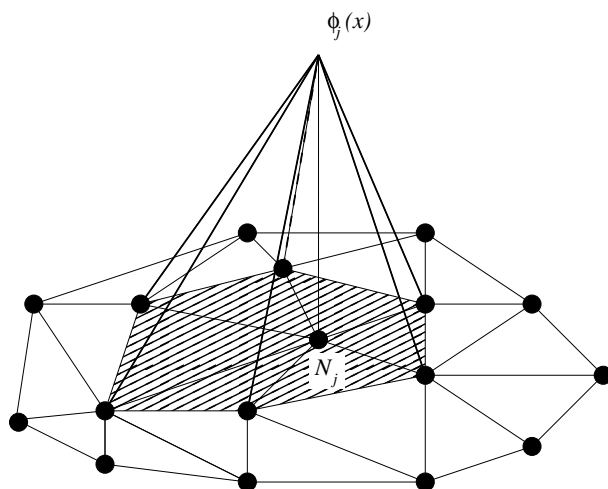


Figura 6: Funzione di base lineare  $\phi_j(x) \in V_h$ .

- i triangoli di “bordo o di contorno” hanno almeno un vertice nel contorno  $\Gamma = \partial\Omega$ .

Un esempio di triangolazione ammissibile è riportato in Figura 5. Si noti come perchè le stime teoriche di convergenza siano efficaci, bisogna che la geometria del contorno non vari al variare della mesh. Per questo motivo si è disegnato un dominio con un contorno formato da una spezzata (lineare a tratti).

Si introduce ora il parametro di mesh  $h$  definito da:

$$h = \max_{T_i \in \mathcal{T}_h} \text{diam}(T_i), \quad (21)$$

dove il diametro del triangolo  $T_i$  indicato con  $\text{diam}(T_i)$  è il lato di lunghezza massima di  $T_i$ . Lo spazio di funzioni finito-dimensionale  $V_h$  è quindi definito da:

$$V_h = \{v(x) : v \text{ è continua in } \Omega, v|_{T_i} \text{ è lineare su ciascun } T_i \in \mathcal{T}_h, v(x) = 0 \text{ per } x \in \Gamma\}.$$

dove  $v|_{T_i}$  è la restrizione della funzione test  $v(x)$  al triangolo  $T_i$ <sup>4</sup>. Si noti che  $V_h \subset V$ . Per usare l'interpolazione lagrangiana consideriamo come punti di appoggio i nodi  $N_i$  della triangolazione escludendo quelli di contorno dove  $v(x) = 0$ . Le funzioni di base  $\phi_i(x)$ ,  $i = 1, \dots, n$  sono quindi definite dalle seguenti condizioni:

$$\phi_i(x_j) = \begin{cases} 1, & \text{se } i = j, \\ 0, & \text{se } i \neq j. \end{cases} \quad i, j = 1, \dots, n$$

Esse sono funzioni piramidali, come mostrato in Figura 6, che hanno come supporto tutti gli elementi che hanno il nodo  $j$  in comune. La generica funzione  $v \in V_h$  può essere rappresentata

<sup>4</sup>La restrizione a  $T_i$  di  $v(x)$  è una funzione definita in  $T_i$  e coincidente con  $v(x)$  in  $T_i$

nel seguente modo:

$$v(x) = \sum_{j=1}^n \eta_j \phi_j(x), \quad \eta_j = v(x_j),$$

e infine si può scrivere il problema agli elementi finiti alla Galerkin come:

**Problema 3.19** (metodo di Galerkin).

Trovare  $u_h \in V_h$  tale che:

$$a(u_h, v) = (f, v) \quad \forall v \in V_h. \quad (22)$$

Sostituendo ora l'espansione in termini delle funzioni di base (esattamente come fatto nella (9) nel caso 1D) si trova il seguente sistema lineare:

$$\sum_{j=1}^n a(\phi_i, \phi_j) u_j = (f, \phi_i) \quad i = 1, \dots, n, \quad (23)$$

che fornisce il sistema lineare degli elementi finiti, che in forma matriciale può essere scritto di nuovo come:

$$Au = b$$

dove ora la matrice di rigidità, il vettore delle incognite e il vettore termini noti sono dati da:

$$A_{[n \times n]} = \{a_{ij}\} \quad a_{ij} = a(\phi_i, \phi_j) = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, d\Omega \quad (24)$$

$$u_{[n \times 1]} = \{u_i\}, \quad b_{[n \times 1]} = \{b_i\} \quad b_i = (f, \phi_i) = \int_{\Omega} f \phi_i \, d\Omega. \quad (25)$$

Notiamo che l'espressione del prodotto scalare ora coinvolge un integrale multidimensionale definito sul dominio  $\Omega$ . Procedendo in modo del tutto analogo al caso 1D, si dimostra che la matrice  $A$  è simmetrica, sparsa e definita positiva.

Nel caso di dominio quadrato discretizzato con triangoli rettangoli aventi i cateti di lunghezza  $h$ , come mostrato in Figura 7, la matrice diventa penta-diagonale, e il sistema assume la seguente forma:

$$\begin{bmatrix} 4 & -1 & 0 & 0 & 0 & -1 & \dots & \dots & \dots & 0 \\ -1 & 4 & -1 & 0 & 0 & 0 & -1 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & -1 & \dots & -1 & 4 & -1 & \dots & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & -1 & \dots & 0 & -1 & 4 & -1 \\ 0 & \dots & \dots & \dots & -1 & 0 & 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_n \end{bmatrix},$$

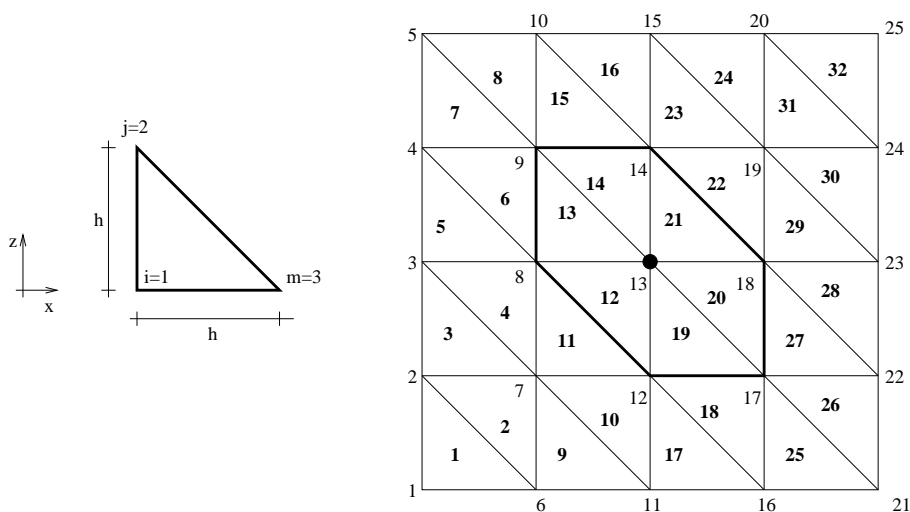


Figura 7: Triangolazione regolare su dominio quadrato.

dove, nel caso  $f = \text{cost}$ , si ha:

$$b_i = fh^2$$

Si riconosce che la matrice di rigidità coincide con la matrice di rigidità del metodo alle differenze finite (FD) del secondo ordine (il cosiddetto “stencil a 5 punti”) [1], mentre con facili calcoli si vede che il termine noto è diverso da quello calcolato da FD.

**Osservazione 3.20.** Si noti che la costruzione della matrice di rigidità può procedere secondo un procedimento che calcola la matrice di rigidità locale di ciascun triangolo e poi assembla i diversi contributi nella matrice di rigidità globale. Tale procedura, chiamata assemblaggio, comune a tutti gli schemi agli elementi finiti, permette di calcolare le quantità necessarie elemento per elemento, e ciò rende il calcolo della matrice  $A$  particolarmente efficiente e flessibile soprattutto in presenza di domini irregolari ed eterogeneità spaziali nei coefficienti della equazione alle derivate parziali da discretizzare. Un altro vantaggio considerevole di questa procedura è quello di rendere facilmente gestibile raffinamenti localizzati della griglia. Talchè è possibile immaginare di costruire una griglia computazionale in cui il passo  $h_j$  della mesh diventa via via più piccolo in corrispondenza a zone del dominio in cui si prevede di avere una derivata seconda della soluzione analitica particolarmente grande (ad esempio vicino a sorgenti puntuali). E ciò per rendere l'errore del metodo il più possibile uniforme spazialmente, utilizzando formule analoghe alla (17) estese al caso multidimensionale. Le griglie triangolari sono particolarmente adatte a questo scopo.

### 3.5.1 Convergenza del metodo FEM nel caso multidimensionale

Gli argomenti riportati nella sezione 3.4.1 possono essere estesi al caso multidimensionale con complicazioni tecniche che esulano dallo scopo di queste note. In particolare, è possibile di-

mostrare, analogamente al caso 1D, che  $u_h \in V_h$  è la miglior approssimazione della soluzione esatta nel senso che:

$$\|\nabla u - \nabla u_h\| \leq \|\nabla u - \nabla v\| \quad \forall v \in V_h,$$

dove la norma è definita qui con:

$$\|\nabla v\| = a(v, v)^{\frac{1}{2}} = \left( \int_{\Omega} u + |\nabla u|^2 d\Omega \right)^{\frac{1}{2}},$$

che dimostra la proprietà di ottimalità della soluzione di Galerkin  $u_h$  rispetto alla norma usata.

Procedendo come prima, possiamo usare le interpolazioni lagrangiane a tratti e usare le stime dell'errore di interpolazione per ottenere:

$$\|\nabla u - \nabla \tilde{u}_h\| \leq Ch.$$

Infine, con dettagli tecnici non trascurabili, è possibile dimostrare che:

$$\|u - u_h\| = \left( \int_{\Omega} (u - u_h)^2 d\Omega \right)^{\frac{1}{2}} \leq Ch^2,$$

che è l'equivalente della (17) del caso 1D. Qui di deve assumere però la “regolarità” della triangolazione al tendere a zero del parametro di mesh  $h$  dato in (21). Più precisamente, si deve richiedere che ogni triangolo non tende a degenerare con il raffinamento progressivo, cioè nessun angolo di nessun triangolo tende a zero per  $h \rightarrow 0$ .

### 3.5.2 Sulla convergenza e spazi funzionali (spazi di Hilbert)

In questa sezione si riporta una breve e incompleta descrizione degli spazi funzionali di interesse per lo studio del metodo agli elementi finiti. Per prima cosa notiamo che, in analogia con l'algebra lineare, dato uno spazio lineare  $V(\Omega)$ , l'operatore  $a(\cdot, \cdot)$  definisce una forma bilineare in  $V \times V$ . Infatti, le seguenti proprietà sono immediatamente verificate:

$$\begin{aligned} a &: V \times V \rightarrow \mathbb{R}, \\ a(u, v) &= a(v, u), \\ a(\alpha u + \beta v, w) &= \alpha a(u, w) + \beta a(v, w), \\ a(u, \alpha v + \beta w) &= \alpha a(u, v) + \beta a(u, w). \end{aligned}$$

La forma bilineare  $a(\cdot, \cdot)$  definisce un prodotto scalare su  $V(\Omega)$  se:

$$a(v, v) > 0 \quad \forall v \in V, \quad v \neq 0.$$

La norma associata con tale prodotto scalare diventa:

$$\|v\|_V = (a(v, v))^{\frac{1}{2}}.$$

Il prodotto scalare soddisfa la disuguaglianza di Cauchy-Schwartz:

$$|a(u, v)| \leq \|u\|_V \|v\|_V.$$

Uno spazio lineare  $V$  dotato di prodotto scalare e norma corrispondente è detto spazio di Hilbert se è completo, cioè se ogni successione di Cauchy<sup>5</sup> converge rispetto a  $\|\cdot\|_V$ .

Per esempio, lo spazio delle funzioni di quadrato sommabile nell'intervallo  $I = [a, b]$ :

$$L_2(I) = \left\{ v(x) : I \rightarrow \mathbb{R} \text{ tali che } \int_a^b v^2 dx < \infty \right\}$$

è uno spazio di Hilbert e può essere dotato del prodotto scalare:

$$(u, v) = \int_a^b u(x)v(x) dx$$

con norma corrispondente data da:

$$\|u\|_{L_2(I)} = \|u\|_2 = \left( \int_a^b [u(x)]^2 dx \right)^{\frac{1}{2}}.$$

**Osservazione 3.21.** In riferimento all'Osservazione 3.11, si noti che lo spazio  $L_2$  è proprio l'insieme di tutte le funzioni a quadrato sommabile menzionate in tale osservazione.

**Esempio 3.22.** La funzione  $v(x) = x^{-\alpha}$ ,  $x \in I = [0, 1]$ , appartiene a  $L_2(I)$  solo per valori di  $\alpha < 1/2$ .

Introduciamo ora lo spazio naturale che contiene la soluzione dei nostri problemi differenziali ellittici: lo spazio di Hilbert  $H^1(I) = \{v : v \text{ e } v' \text{ appartengono a } L_2(I)\}$ . Il prodotto scalare è dato da:

$$(u, v)_{H^1(I)} = \int_a^b [u(x)v(x) + u'(x)v'(x)] dx$$

e norma data da:

$$\|u\|_{H^1(I)} = \left( \int_a^b [u(x)^2 + u'(x)^2] dx \right)^{\frac{1}{2}}$$

Si noti che lo spazio  $V$  definito nel paragrafo 3.2 è uno spazio di Hilbert opportuno, ed è di solito indicato con:

$$V(I) = H_0^1(I) = \{v(x) : \mathbb{R} \rightarrow \mathbb{R} \text{ tali che } v(x) \in L_2(I), v'(x) \in L_2(I) \text{ e } v(0) = v(1) = 0\}$$

---

<sup>5</sup>Una successione di funzioni  $v_i \in V$  è detta di Cauchy se esiste  $\epsilon > 0$  tale che  $\|v_i - v_j\|_V < \epsilon$  per  $i$  e  $j$  sufficientemente grandi. Si dice che  $v_i$  converge a  $v$  se  $\|v - v_i\|_V \rightarrow 0$  per  $i \rightarrow \infty$ .



Si noti che il pedice 0 nel simbolo dello spazio di Hilbert viene usato per denotare il fatto che le funzioni sono nulle al bordo del dominio. L'apice 1 viene usato per denotare l'ordine massimo delle derivate che sono di quadrato sommabile (in questo caso le derivate prime).

Tutte queste nozioni sono facilmente estendibili al caso multidimensionale. Gli spazi sopra definiti per funzioni definite su un dominio limitato  $\Omega \in \mathbb{R}^d$  con contorno  $\Gamma = \partial\Omega$  sufficientemente liscio, sono dati da:

$$L_2(\Omega) = \left\{ v(x) : \mathbb{R}^d \rightarrow \mathbb{R} \text{ tali che } \int_{\Omega} v(x)^2 < \infty \right\}$$

$$H^1(\Omega) = \left\{ v(x) : \mathbb{R}^d \rightarrow \mathbb{R} \text{ tali che } v(x) \in L_2(\Omega) \text{ e } \frac{\partial v(x)}{\partial x_i} \in L_2(\Omega) \text{ per } i = 1, \dots, d \right\}$$

con i seguenti prodotti scalari (e conseguenti norme):

$$(u, v)_{L^2(\Omega)} = \int_{\Omega} uv \, dx \qquad \|u\|_{L^2(\Omega)} = \left( \int_{\Omega} u^2 \, dx \right)^{\frac{1}{2}}$$

$$(u, v)_{H^1(\Omega)} = \int_{\Omega} [uv + \nabla u \cdot \nabla v] \, dx \qquad \|u\|_{H^1(\Omega)} = \left( \int_{\Omega} [u^2 + |\nabla u|^2] \, dx \right)^{\frac{1}{2}}$$

Si definisce poi:

$$H_0^1(\Omega) = \{v(x) \in H^1(\Omega) \text{ tali che } v(x) = 0 \in \Gamma\}.$$

con lo stesso prodotto scalare e la stessa norma definita per  $H^1(\Omega)$ . Si noti che lo spazio di funzioni  $V(\Omega) = H_0^1(\Omega)$

**Ottimalità della soluzione al problema variazionale.** Per apprezzare meglio l'ottimalità della soluzione  $u_h$  dichiarata in precedenza, prendiamo un caso particolare. Applichiamo il metodo di Galerkin alla seguente equazione:

$$\begin{aligned} -\Delta u + u &= f & x \in \Omega \\ u &= 0 & x \in \Gamma = \partial\Omega \end{aligned}$$

Il problema di variazionale diventa:

**Problema 3.23.** Problema variazionale Trovare  $u \in H_0^1(\Omega)$  tale che:

$$a(u, v)_V = \langle f, v \rangle \qquad \forall v \in H_0^1(\Omega) \qquad (26)$$

dove

$$a(u, v)_V = \int_{\Omega} [\nabla u \cdot \nabla v + uv] \, dx$$

Il corrispondente problema agli elementi finiti diventa dunque:

**Problema 3.24.** Problema FEM Trovare  $u_h \in V_h(\Omega) \subset H_0^1(\Omega)$  tale che:

$$a(u_h, v)_V = \langle f, v \rangle \quad \forall v \in V_h(\Omega). \quad (27)$$

Sottraendo (27) da (26), si ottiene l'equazione che mostra la consistenza forte dello schema, e cioè:

$$a(u - u_h, v)_V = 0 \quad \forall v \in V_h(\Omega)$$

che sancisce l'ortogonalità della funzione errore a tutte le funzioni di  $V_h(\Omega)$  rispetto al prodotto scalare  $a(\cdot, \cdot)_V$ . Questo equivale a dire che  $u_h$  è la proiezione ortogonale di  $u$  su  $V_h(\Omega)$  rispetto al prodotto scalare  $a(\cdot, \cdot)$ . In altre parole, notando che il prodotto scalare è proprio quello di  $H^1(\Omega)$ ,  $u_h$  è caratterizzato da una norma  $H^1$  minima in confronto con qualsiasi altra funzione di  $V_h(\Omega)$ , e cioè:

$$\|u - u_h\|_{H^1(\Omega)} \leq \|u - v\|_{H^1(\Omega)} \quad \forall v \in V_h(\Omega)$$

D'ora in avanti si ometteranno i pedici nei simboli di norma o prodotto scalare quando lo spazio di definizione è dettato dal contesto.

### 3.6 Problema di Neumann: condizioni al contorno naturali e essenziali

Si consideri ora il seguente problema al contorno:

$$\begin{aligned} -\Delta u + u &= f && \text{in } \Omega, \\ \nabla u \cdot \vec{n} &= g && \text{in } \Gamma = \partial\Omega. \end{aligned} \quad (28)$$

Moltiplicando la prima equazione per una funzione test  $v \in V$ , e integrando sul dominio, si ottiene:

$$\int_{\Omega} (\Delta uv - uv) d\Omega = \int_{\Omega} fv d\Omega. \quad (29)$$

Applicando il lemma di Green al primo termine nell'integrale del primo membro, si ottiene:

$$\int_{\Omega} uv d\Omega - \int_{\Gamma} \nabla u \cdot \vec{n} v d\Gamma + \int_{\Omega} \nabla u \cdot \nabla v d\Omega = \int_{\Omega} fv d\Omega,$$

dove ancora  $\Gamma = \partial\Omega$  è la frontiera di  $\Omega$ . Notiamo che il primo termine dell'equazione precedente contiene esattamente il termine di Neumann sul bordo:  $\nabla u \cdot \vec{n}$ . Nel caso di condizioni al contorno di Dirichlet omogenee, avevamo richiesto che la funzione  $v$  fosse nulla al bordo con la conseguenza che l'integrale su  $\Gamma$  era nullo. In questo caso, invece, dobbiamo richiedere che le funzioni test siano diverse da zero al bordo. Infatti, possiamo dare la seguente formulazione variazionale:

**Problema 3.25** (variazionale).

Trovare  $u \in V$  tale che:

$$a(u, v) = (f, v) + \langle g, v \rangle \quad \forall v \in V, \quad (30)$$

dove:

$$\begin{aligned} a(u, v) &= \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx \\ (f, v) &= \int_{\Omega} f v \, dx \\ \langle g, v \rangle &= \int_{\Gamma} g v \, dx \\ V &= \left\{ v(x) : v \text{ è continua in } \Omega, \frac{\partial v}{\partial x_i} \text{ sono continue in } \Omega \forall i \right\}, \end{aligned}$$

che si può dimostrare essere equivalente al seguente problema di minimizzazione:

**Problema 3.26** (minimizzazione).

Trovare  $u \in V$  tale che:

$$F(u) \leq F(v) \quad \forall v \in V \quad (31)$$

dove:

$$F(v) = \frac{1}{2} a(v, v) - (f, v) - \langle g, v \rangle.$$

Si noti che, assumendo la  $u$  sufficientemente regolare, e applicando all'indietro il lemma di Green alla (30), si ottiene

$$\int_{\Omega} (-\Delta u + u - f) v \, d\Omega + \int_{\Gamma} (\nabla u \cdot \vec{n} - g) v \, d\Gamma = 0 \quad \forall v \in V. \quad (32)$$

Si noti che le funzioni  $v \in V$  sono non nulle al contorno, per cui possiamo imporre le seguenti due condizioni:

$$\int_{\Omega} (-\Delta u + u - f) v \, d\Omega = 0 \quad \forall v \in V,$$

e

$$\int_{\Gamma} (\nabla u \cdot \vec{n} - g) v \, d\Gamma = 0 \quad \forall v \in V.$$

Variando  $v$  nello spazio  $V$  (dove  $v$  non si annulla in  $\Gamma$ ), si ottiene applicando il lemma 3.5:

$$-\Delta u + u - f = 0 \quad \text{in } \Omega,$$

e

$$\nabla u \cdot \vec{n} - g = 0 \quad \text{in } \Gamma;$$

che ci dice che le condizioni al contorno del problema originale sono soddisfatte.

**Osservazione 3.27.** Le condizioni al contorno di Neumann non appaiono esplicitamente nella formulazione variazionale, che si differenzia dal caso di condizioni al contorno di Dirichlet solo dal fatto che le funzioni test non sono più nulle al bordo. Possiamo quindi dire che le condizioni di Dirichlet vanno imposte esplicitamente (richiedendo appunto l'annullarsi delle funzioni test al bordo di Dirichlet), mentre le condizioni di Neumann sono "naturali" nella formulazione variazionale. Si noti che se fosse  $g = 0$  (flusso nullo al contorno), il termine  $\langle g, v \rangle$  sparirebbe: se non si impone esplicitamente alcuna condizione al bordo nella formulazione variazionale, si assume implicitamente condizioni al bordo di Neumann nulle. E' facile ora ricavare una formulazione variazionale per un problema misto in cui si impongano condizioni al contorno di Dirichlet in una porzione della frontiera, e condizioni di Neumann nella porzione complementare.

Siamo ora in grado di scrivere la seguente formulazione agli elementi finiti:

**Problema 3.28** (Galerkin).

Trovare  $u_h \in V_h$  tale che:

$$a(u_h, v) = (f, v) + \langle g, v \rangle \quad \forall v \in V_h, \quad (33)$$

dove:

$$\begin{aligned} a(u, v) &= \int_{\Omega} (\nabla u_h \cdot \nabla v + u_h v) \, dx \\ (f, v) &= \int_{\Omega} f v \, dx \\ \langle g, v \rangle &= \int_{\Gamma} g v \, dx \\ V_h &= \{v(x) : v \text{ è continua in } \Omega, v|_{T_k} \text{ è lineare } \forall T_k \in \mathcal{T}_h\}. \end{aligned}$$

**Osservazione 3.29.** La formulazione pratica procede quindi come nel caso del problema di Dirichlet omogeneo. Si noti ancora che le condizioni al contorno di Dirichlet, essendo imposte esplicitamente, sono soddisfatte in maniera "forte", mentre quelle di Neumann, essendo imposte in modo variazionale, sono soddisfatte in maniera "debole". Questo si osserva nella pratica quando si vanno a sperimentare le convergenze teoriche, dove si vede che l'errore puntuale nei nodi di Dirichlet è praticamente nullo (inferiore o uguale alla tolleranza usata per la soluzione del sistema lineare), mentre l'errore sulla soluzione nei nodi di Neumann tende a zero come  $h^2$  (si veda la (17)).

**Osservazione 3.30.** Nella pratica l'imposizione delle condizioni al contorno di Dirichlet non omogenee avviene direttamente nella matrice del sistema lineare. Si procede nel modo seguente. Supponiamo che il nodo  $i$ -esimo sia contenuto nel contorno di Dirichlet. Si procede allora alla costruzione della  $i$ -esima riga del sistema lineare senza tenere in considerazione il fatto che essa corrisponde ad un nodo di Dirichlet. Una volta costruita tutta la matrice di rigidezza, si impone direttamente sul nodo  $i$ -esimo che  $u_i$  sia uguale esattamente al valore imposto, chiamiamolo  $\bar{u}_i$ . Questo si può ottenere in due modi.

**I modo.** Si azzerano tutti gli elementi extra diagonali della  $i$ -esima riga della matrice, e si impone pari a uno l'elemento diagonale e pari al valore imposto il corrispondente elemento del termine noto. Facendo così, si sostituisce alla  $i$ -esima equazione l'equazione:

$$u_i = \bar{u}_i$$

però contemporaneamente la matrice di rigidezza non è più simmetrica, e va opportunamente cambiata per mantenere la simmetria.

**II modo, detto di "penalty".** Si sostituisce al termine diagonale della  $i$ -esima riga un valore molto elevato  $\lambda$  e si impone il corrispondente elemento del vettore termini noti pari a  $\lambda\bar{u}_i$ . In questo modo si ottiene la seguente equazione  $i$ -esima:

$$\sum_{k=1}^{i-1} a_{ik}u_k + \lambda u_i + \sum_{k=i+1}^n a_{ik}u_k = \lambda\bar{u}_i.$$

A primo membro, però, tutti i termini extra diagonali (rappresentati dalle due sommatorie) sono effettivamente trascurabili rispetto al termine diagonale se  $\lambda$  è sufficientemente grande. Quindi l'equazione precedente corrisponde in pratica a:

$$\lambda u_i = \lambda\bar{u}_i,$$

che evidentemente impone in maniera corretta la condizione di Dirichlet. Un valore utilizzabile per  $\lambda$  è ad esempio  $\lambda = 10^{40}$ .

### 3.7 Tipologia di Elementi finiti

Finora abbiamo visto esclusivamente elementi finiti che ammettono funzioni di base lineari (e.g. triangoli in  $\mathbb{R}^2$ ). E' intuitivo pensare che si possano definire su elementi di questa forma funzioni di base polinomiali a tratti con grado più elevato. Per esempio, è immediato definire in 1-D funzioni di base quadratiche: ogni elemento finito sarà formato da 3 nodi, necessari per valutare le tre costanti che formano una parabola utilizzando la proprietà di interpolazione delle funzioni di base (Vedasi Fig. 8. Evidentemente, un elemento finito quadratico su un triangolo è

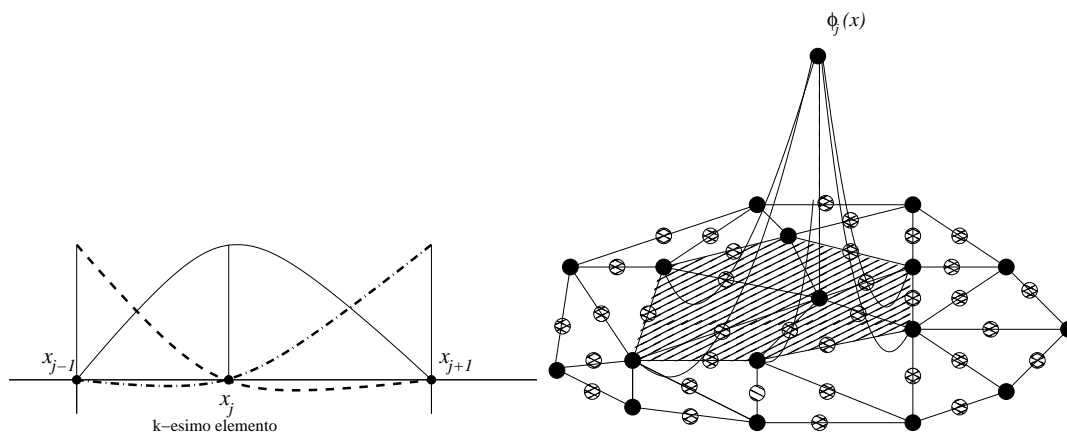


Figura 8: Funzioni di base per elementi finiti quadratici in 1D (a sinistra) e in 2D (a destra)

definito da 6 nodi dove imporre le condizioni di interpolazione per poter definire i 6 coefficienti che definiscono una parabola in  $\mathbb{R}^2$ .

Questa definizione basata su triangolazioni del dominio garantisce naturalmente (cioè senza trasformazioni particolari dell'elemento) la continuità della rappresentazione della soluzione, cioè di  $u_h$  ai bordi degli elementi, e quindi in tutto  $\Omega$ . Ciò non è però più vero nel caso si usino forme geometriche degli elementi diverse dai semplici (sottointervalli, triangoli, tetraedri). Per risolvere questo problema si introduce quindi una trasformazione di ciascun elemento ai fini solo della determinazione delle funzioni di interpolazione della soluzione e da usarsi essenzialmente per il calcolo degli integrali (24).

### 3.7.1 Elementi isoparametrici

Prendiamo in considerazione l'esempio semplificato di elementi finiti di forma quadrata, come l'esempio di Figura 9, dove sono rappresentati due elementi adiacenti. In questo caso, la continuità di  $u_h$  è assicurata se prendiamo delle funzioni di base cosiddette bilineari, e cioè lineari separatamente in  $x$  ed in  $y$ . La loro espressione può genericamente essere scritta come:

$$\phi_i(x, y) = (a_i + b_i x)(c_i + d_i y)$$

Infatti, è facile vedere che al bordo degli elementi, per esempio per  $x = 1$ , la rappresentazione della funzione di base dipende solo da  $y$  ed è lineare, per cui i due nodi del lato in comune sono sufficienti a determinarne i coefficienti, e analogamente per gli altri bordi. I coefficienti da determinare per ciascun elemento sono 4, e 4 sono i nodi su cui imporre la condizione di interpolazione, quindi il loro calcolo è immediato. L'estensione diretta di questo procedimento per funzioni di base formate da polinomi di ordine superiore al primo, ma sempre separatamente lungo  $x$  e lungo  $y$ , è immediata.

Nel caso di elementi quadrangolari ma di forma non quadrata o non rettangolare, quindi con lati non allineati alle direzioni coordinate, la continuità delle funzioni di base bilineari non

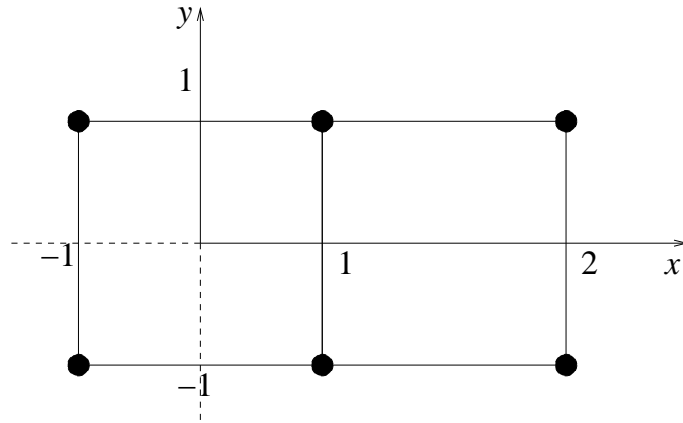


Figura 9: Elemento quadrato con funzioni di forma bilineari

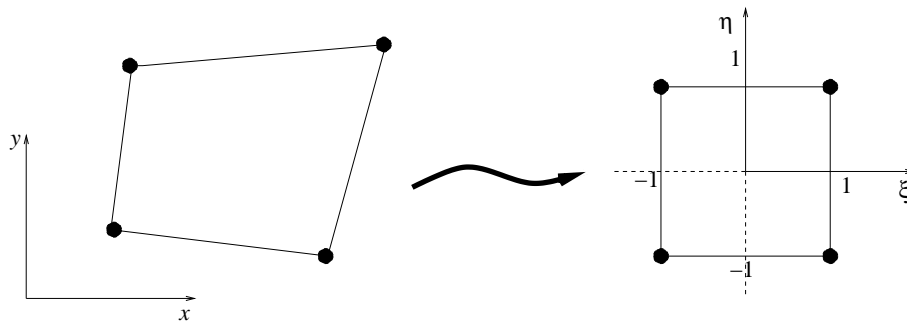


Figura 10: Elemento quadrilatero generico e il suo trasformato

è più verificata. Un modo per ricavare funzioni di base continue al bordo è quello di trasformare ogni elemento quadrilatero in un elemento quadrato di riferimento tramite un cambio locale di coordinate (trasformazione conforme - Fig. 10). Sull'elemento di riferimento è possibile definire funzioni di base bilineare col procedimento sopradescritto e quindi il calcolo degli integrali che formano la (24).

Vediamo un esempio nel caso di funzioni di base bilineari, con riferimento alla figura 10. E' facilmente verificabile che la trasformazione  $(x, y) \rightarrow (\eta, \xi)$  è:

$$\begin{aligned} x &= \frac{1}{4} [(1 - \xi)(1 - \eta)x_i + (1 + \xi)(1 - \eta)x_j + (1 + \xi)(1 + \eta)x_m + (1 - \xi)(1 + \eta)x_k] \\ y &= \frac{1}{4} [(1 - \xi)(1 - \eta)y_i + (1 + \xi)(1 - \eta)y_j + (1 + \xi)(1 + \eta)y_m + (1 - \xi)(1 + \eta)y_k]. \end{aligned}$$

Ricordiamo che si vogliono calcolare gli integrali di tipo (24). Per fare questo dobbiamo calcolare lo Jacobiano della trasformazione:

$$J = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix},$$

e l'integrale della generica funzione  $f(x, y)$  è calcolabile come:

$$\int_{\Omega^e} f(x, y) \det J \, dx dy = \int_{-1}^1 \int_{-1}^1 f(\eta, \xi) \det J \, d\eta d\xi,$$

che può essere calcolate numericamente ad esempio con le formule di Gauss a 4 punti, ricordando comunque che per una funzione scalare  $f(x, y)$  si hanno le seguenti relazioni tra i gradienti (e quindi le derivate):

$$\nabla_{(\xi, \eta)} f(x, y) = J \nabla f(x, y).$$

### 3.8 Equazione di diffusione e trasporto

Siamo ora in grado di affrontare un problema differenziale più complicato. Si consideri l'equazione (ellittica) di diffusione e trasporto seguente:

$$\begin{aligned} -\operatorname{div}(D \nabla u) + \operatorname{div}(\vec{\beta} u) &= f && \text{in } \Omega \\ u &= 0 && \text{in } \Gamma_D \\ D \nabla u \cdot \vec{n} &= g && \text{in } \Gamma_N, \end{aligned} \tag{34}$$

dove  $D$  è il coefficiente di diffusione (scalare e strettamente positivo) e  $\vec{\beta}(x)$  è un campo vettoriale. Dal punto di vista applicativo, questa equazione rappresenta per esempio il trasporto di una sostanza disciolta in un fluido che si muove con il campo di moto  $\vec{\beta}(x)$ .



Procediamo dunque allo sviluppo di una formulazione variazionale per questo problema. Moltiplicando per una funzione test e integrando sul dominio, si ottiene:

$$- \int_{\Omega} \operatorname{div} D \nabla u v \, d\Omega + \int_{\Omega} \operatorname{div}(\vec{\beta} u) v \, d\Omega = \int_{\Omega} f v \, d\Omega.$$

Applicando ora il lemma di Green solo al primo termine del primo membro, si ottiene:

$$- \int_{\Gamma_N} g f \, d\Gamma + \int_{\Omega} D \nabla u \cdot \nabla v \, d\Omega + \int_{\Omega} \operatorname{div}(\vec{\beta} u) v \, d\Omega = \int_{\Omega} f v \, d\Omega,$$

da cui si può ricavare direttamente il seguente metodo agli elementi finiti:

**Problema 3.31** (Galerkin).

Trovare  $u_h \in V_h$  tale che:

$$a(u_h, v) = (f, v) + \langle g, v \rangle \quad \forall v \in V_h, \quad (35)$$

dove:

$$a(u_h, v) = \int_{\Omega} \left( D \nabla u_h \cdot \nabla v + \vec{\beta} \cdot \nabla u_h v \right) dx$$

$$(f, v) = \int_{\Omega} f v \, dx$$

$$\langle g, v \rangle = \int_{\Gamma} g v \, dx$$

$$V_h = \{v(x) : v \text{ è continua in } \Omega, v(x) = 0 \text{ in } \Gamma_D, v|_{T_k} \text{ è lineare } \forall T_k \in \mathcal{T}_h; \}.$$

**Osservazione 3.32.** Si vede immediatamente che  $a(u, v) \neq a(v, u)$ , per cui questo metodo di Galerkin non ha un'equivalente di Ritz. In altre parole, non esiste in questo caso un metodo di minimizzazione, ma solo un metodo di ortogonalizzazione, cioè di Galerkin. Conseguenza di questo fatto è che il sistema lineare derivante dalla discretizzazione agli elementi finiti sarà sparso ma non simmetrico.

Procedendo nel modo consueto, si arriva dunque al sistema lineare seguente:

$$(A + B)u = c,$$

dove  $A$  è la matrice simmetrica di rigidezza, e  $B$  rappresenta la matrice non simmetrica del trasporto, date da:

$$A = \{a_{ij}\} \quad a_{ij} = \int_{\Omega} D \nabla \phi_j \cdot \nabla \phi_i \, d\Omega$$

$$B = \{b_{ij}\} \quad b_{ij} = \int_{\Omega} \operatorname{div}(\vec{\beta} \phi_i) \phi_j \, d\Omega$$

$$c = \{c_i\} \quad c_i = \int_{\Omega} f \phi_i \, d\Omega + \int_{\Gamma_N} g \phi_i \, d\Gamma.$$

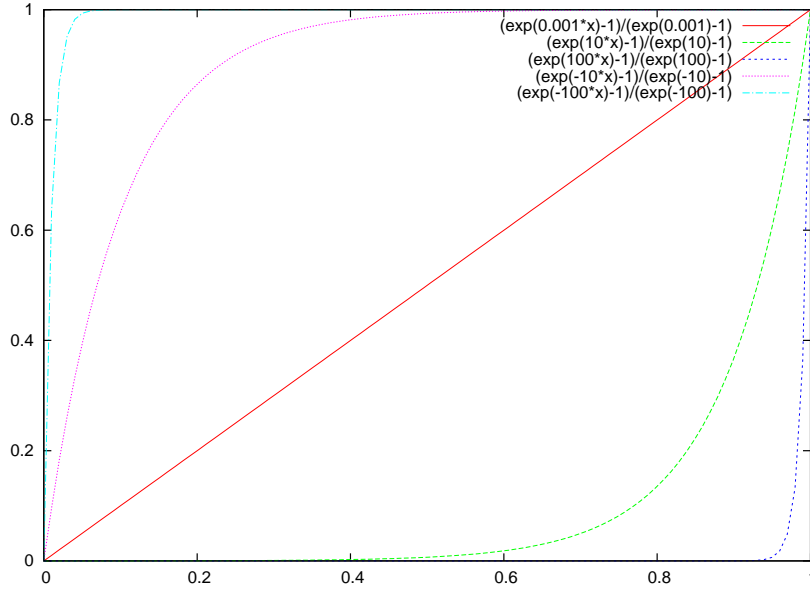


Figura 11: Andamento della soluzione del problema di trasporto (convezione-diffusione) per diversi valori del rapporto  $b/D$  (convezione/diffusione).

### 3.8.1 Caso monodimensionale

Consideriamo il seguente problema monodimensionale:

$$\begin{aligned} -Du'' + bu' &= 0, & 0 < x < 1, \\ u(0) &= 0; & u(1) = 1. \end{aligned} \tag{36}$$

Ricaviamo la soluzione analitica di tale problema. L'equazione caratteristica è data da:

$$-D\lambda^2 + b\lambda = 0,$$

che ha radici pari a  $\lambda_1 = 0$  e  $\lambda_2 = b/D$ , per cui la soluzione generale è data da:

$$u(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} = c_1 + c_2 e^{bx/D}.$$

Imponendo le condizioni al contorno si ottiene immediatamente:

$$u(x) = \frac{e^{\frac{b}{D}x} - 1}{e^{\frac{b}{D}} - 1},$$

che è mostrata nel grafico di Figura 11 per diversi valori di  $b/D$ . Si vede che per valori bassi del rapporto  $b/D$  la soluzione è praticamente lineare, mentre per valori grandi la soluzione presenta un andamento esponenziale marcato, caratterizzato da zone del dominio dove si hanno gradienti spaziali elevati.

Procedendo alla formulazione agli elementi finiti con funzioni test lineari, e ricordando che in un caso monodimensionale con griglia regolare si ha equivalenza tra il metodo FEM e il metodo FD (si veda il paragrafo 3.4, si ricava la seguente equazione alle differenze per il nodo  $i$ -esimo:

$$\frac{D}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) + \frac{b}{2h}(u_{i+1} - u_{i-1}) = 0. \quad (37)$$

Si nota che tale equazione corrisponde ad una discretizzazione “centrata” alle differenze finite sia per la derivata seconda che per la derivata prima. Si veda a tal proposito l’Appendice A.

Introducendo ora il numero di Péclet di griglia, un numero adimensionale che indica il rapporto tra i flussi convettivi e quelli diffusivi, dato dal rapporto:

$$\mathbb{P}e = \frac{|b|h}{D},$$

si arriva alla seguente equazione alle differenze (per  $b > 0$ ):

$$(\mathbb{P}e - 2)u_{i+1} + 4u_i - (\mathbb{P}e + 2)u_{i-1} = 0 \quad i = 1, \dots, n - 1. \quad (38)$$

Si può procedere alla soluzione analitica di tale equazione alle differenze imponendo una soluzione del tipo  $u_i = \lambda^i$ . Sostituendo si ottiene:

$$(\mathbb{P}e - 2)\lambda^2 + 4\lambda - (\mathbb{P}e + 2) = 0,$$

da cui si ricava:

$$\lambda_{1,2} = \frac{-2 \pm \sqrt{2 + \mathbb{P}e^2 - 2}}{\mathbb{P}e - 2} = \begin{cases} (2 + \mathbb{P}e)(2 - \mathbb{P}e), \\ 1. \end{cases}$$

La soluzione generale della (38) è la combinazione lineare:

$$u_i = c_1\lambda_1^i + c_2\lambda_2^i$$

con le costanti che vanno ricavate dall’imposizione delle condizioni al contorno. Questo porta alla fine alla seguente soluzione dell’equazione alle differenze:

$$u_i = \frac{1 - \left(\frac{2+\mathbb{P}e}{2-\mathbb{P}e}\right)^i}{1 - \left(\frac{2+\mathbb{P}e}{2-\mathbb{P}e}\right)^n} \quad i = 0, 1, \dots, n,$$

che fornisce la soluzione del problema discretizzato agli elementi finiti (o alle differenze finite) per ogni nodo della griglia computazionale.

Da questa equazione si vede immediatamente che lo schema risulta oscillante nel caso  $\mathbb{P}e > 2$ , perchè in tal caso il numeratore diventa negativo, e ovviamente assume valori oscillanti a seconda

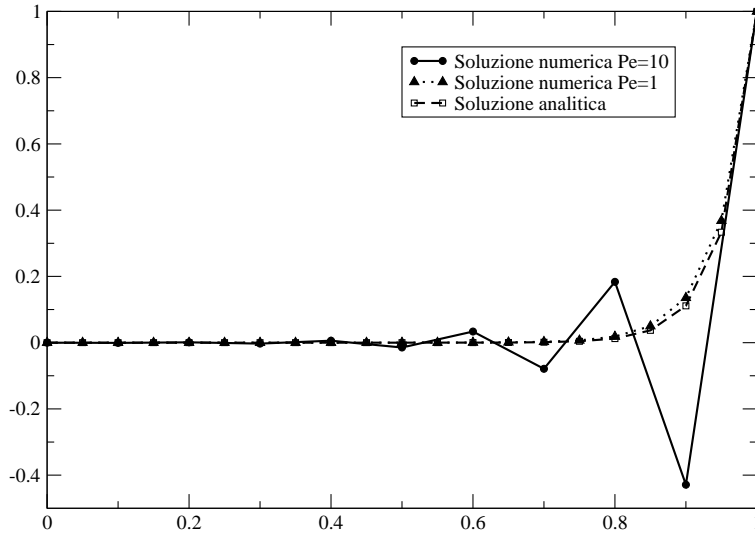


Figura 12: Comportamento dello schema alle differenze finite per la soluzione dell'equazione di convezione e diffusione a confronto con la soluzione analitica nel caso di  $\mathbb{P}e = 0.5$  e  $\mathbb{P}e = 2$ .

che  $i$  sia pari o dispari. Tale comportamento è illustrato in Figura 12, dove si vede che per  $\mathbb{P}e > 2$  si verificano oscillazioni, mentre per valori inferiori lo schema risulta stabile.

Per cercare di correggere la situazione proviamo ad usare una discretizzazione decentrata per la derivata prima. La nuova discretizzazione diventa quindi:

$$\frac{D}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) + \frac{b}{h}(u_{i+1} - u_i) = 0,$$

dove il termine convettivo è ora discretizzato con una differenza del prim'ordine in avanti al posto della differenza centrale del secondo ordine usata in precedenza.

Procedendo nello stesso modo di prima si arriva all'equazione alle differenze:

$$(\mathbb{P}e - 1)u_{i+1} - (\mathbb{P}e - 2)u_i - u_{i-1} = 0 \quad i = 1, \dots, n - 1.$$

La soluzione dell'equazione caratteristica corrispondente è:

$$\lambda_{1,2} = \frac{\mathbb{P}e - 2 \pm \sqrt{(\mathbb{P}e - 2)^2 + 4(\mathbb{P}e - 1)}}{2(\mathbb{P}e - 1)} = \begin{cases} (1)(1 - \mathbb{P}e), \\ 1, \end{cases}$$

che porta alla soluzione:

$$u_i = \frac{1 - \left(\frac{1}{1 - \mathbb{P}e}\right)^i}{1 - \left(\frac{1}{1 - \mathbb{P}e}\right)^n} \quad i = 0, 1, \dots, n,$$

che risulta instabile per  $\mathbb{P}e < 1$ . Se invece si usa la differenza decentrata all'indietro (upwind), si ottiene:

$$\frac{D}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) + \frac{b}{h}(u_i - u_{i-1}) = 0,$$

Rifacendo i conti si vede immediatamente che la soluzione numerica non mostra oscillazioni per nessun valore di  $\mathbb{P}e$  ed è incondizionatamente stabile. Con facili passaggi algebrici, l'equazione alle differenze precedente può essere scritta come:

$$\frac{D}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) + \frac{b}{2h}(u_{i+1} - u_{i-1}) + \frac{bh}{2}\left(\frac{-u_{i-1} + 2u_i - u_{i+1}}{h^2}\right) = 0,$$

da cui si evince che la formulazione “upwind” equivale alla formulazione centrata a cui si è aggiunta una “diffusione numerica” pari a  $h/2$ . In questo caso, il nuovo numero di Péclet diventa:

$$\mathbb{P}e = \frac{bh}{D + bh/2},$$

che è sempre inferiore o uguale a 2 per qualsiasi valore di  $D$  e  $b(> 0)$ . Per stabilizzare lo schema, si ricorre dunque all'aggiunta di un termine spurio (in pratica sto risolvendo una equazione diversa da quella di partenza), che però tende a zero al tendere a zero del parametro di griglia, per cui la consistenza dello schema è assicurata. Questo modo di procedere è un modo usato molto spesso per stabilizzare schemi numerici. In questo caso, quello che si paga è la diminuita accuratezza dello schema a griglia fissata, in questo caso evidenziata da una soluzione molto più regolare di quella analitica, con un fronte più diffuso. Il termine aggiuntivo, in effetti, equivale ad aver risolto un'equazione con il coefficiente di diffusione pari a  $D = D + h/2$ .

Nel caso degli elementi finiti si può pensare di procedere allo stesso modo, di aggiungere cioè al un termine diffusivo opportunamente creato. Per fare questo, aggiungiamo alla nostra forma bilineare un termine sempre bilineare proporzionale a:

$$\int_{\Omega} (\vec{\beta} \cdot \nabla u) (\vec{\beta} \cdot \nabla v) \, d\Omega,$$

che in pratica corrisponde ad un termine di diffusione numerica applicata solo lungo le linee di corrente, cioè solo lungo la direzione del campo di moto  $\vec{\beta}(x)$ . Il metodo agli elementi finiti, chiamato anche SD-FEM (Streamline Diffusion Finite Elements), diventa dunque:

**Problema 3.33** (Streamline Diffusion).

Trovare  $u_h \in V_h$  tale che:

$$a_h(u_h, v) = (f, v) \quad \forall v \in V_h, \tag{39}$$

dove:

$$\begin{aligned}
a_h(u_h, v) &= \int_{\Omega} \left[ D \nabla u_h \cdot \nabla v + \vec{\beta} \cdot \nabla u_h v + \tau \mathbb{P}e_h \left( \vec{\beta} \cdot \nabla u_h \right) \left( \vec{\beta} \cdot \nabla v \right) \right] dx \\
(f, v) &= \int_{\Omega} f v dx \\
V_h &= \{v(x) : v \text{ è continua in } \Omega, v(x) = 0 \text{ in } \Gamma_D, v|_{T_k} \text{ è lineare } \forall T_k \in \mathcal{T}_h; \},
\end{aligned}$$

dove  $\mathbb{P}e_h$  è il numero di Péclet di griglia definito elemento per elemento da:

$$\mathbb{P}e_h = \frac{|\vec{\beta}_k| h_k}{D^{(k)}}$$

con  $D^{(k)}$  il coefficiente di diffusione  $\vec{\beta}_k$  il vettore velocità ancora assunti costanti nell'elemento  $T_k$ , ma potenzialmente variabili da elemento a elemento.

Si riconosce immediatamente che  $a_h(\cdot, \cdot) \rightarrow a(\cdot, \cdot)$  per  $h \rightarrow 0$ , è il termine aggiuntivo di diffusione numerica introdotto lungo la direzione  $\vec{\beta}$  della velocità, e  $\tau$  è un coefficiente che deve essere tarato caso per caso in maniera da ottenere uno schema in cui le oscillazioni numeriche siano minimizzate o annullate, senza contemporaneamente introdurre una quantità eccessiva di diffusione numerica. Le figure 13 e 14 riportano esempi esemplificativi dei tipici andamenti della soluzione nei casi stabili e instabili.

Si noti infine che esistono schemi molto più avanzati dello schema Streamline Upwind, basati essenzialmente nell'introduzione di diffusione numerica in minima quantità e solo quando e dove necessario. Tali algoritmi non vengono trattati in queste note. Il lettore interessato può riferirsi ai testi [2, 5, 3] e alle referenze citate in essi.

## A Appendice A: Discretizzazione alle differenze finite dell'equazione di convezione e diffusione.

Il metodo alle differenze finite per la discretizzazione dell'equazione della convezione e diffusione nel caso monodimensionale procede nel modo. Si consideri una discretizzazione uniforme di passo  $h$  dell'intervallo  $[0, 1]$ , per cui ciascun sotto intervallo è caratterizzato dai suoi estremi indicati con  $x_i$  e  $x_{i+1} = x_i + h$ . Sia  $u_i$  l'approssimazione numerica della soluzione  $u(x)$  nel punto  $x_i$ :  $u_i \approx u(x_i)$ . Usando in maniera sistematica lo sviluppo in serie di Taylor di  $u_i$  con passo  $h$  si ottiene:

$$u_{i+1} = u_i + hu'_i + \frac{h^2}{2}u''_i + \dots \quad (40)$$

$$u_{i-1} = u_i - hu'_i + \frac{h^2}{2}u''_i + \dots \quad (41)$$

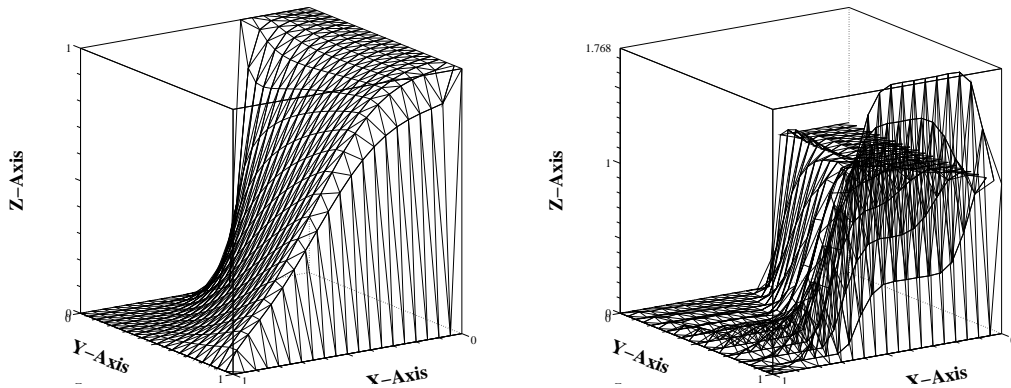


Figura 13: Problema di diffusione e convezione risolto con Galerkin lineare (P1) senza stabilizzazione. Nel grafico a sinistra si mostra il caso con  $D = 0.1$  mentre in quello a destra vi è il caso con  $D = 0.01$ . In ambedue i casi il vettore velocità è  $\vec{\beta} = (1, 3)^T$ , talchè il numero di Péclet di griglia è costante e pari a  $\mathbb{P}e_h = 1$  e  $\mathbb{P}e_h = 10$  nei due casi.

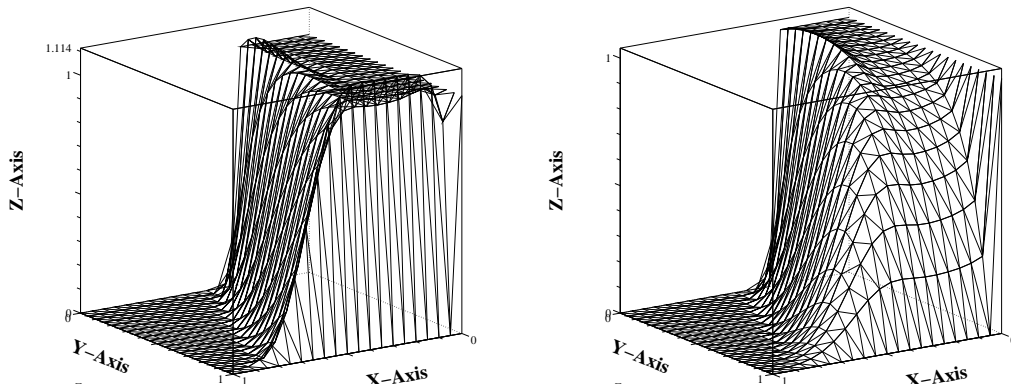


Figura 14: Problema di diffusione e convezione risolto con Galerkin lineare (P1) con stabilizzazione SD nel caso  $D = 0.01$  e  $\vec{\beta} = (1, 3)^T$  ( $\mathbb{P}e_h = 10$ ). A sinistra si riporta il caso con  $\tau = 0.01$  e a destra il caso con  $\tau = 1.0$ . Si noti la diminuzione di oscillazioni e l'aumento della diffusione numerica.

Sommando membro a membro le due equazioni, e trascurando i termini dello sviluppo di ordine superiore, si ottiene la seguente approssimazione della derivata seconda nel nodo  $i$ -esimo:

$$u_i'' \approx \frac{-u_{i-1} + 2u_i - u_{i+1}}{h^2},$$

mentre la derivata prima sempre nel punto  $x_i$  può essere approssimata da:

$$u_i' \approx \frac{u_{i+1} - u_{i-1}}{2h},$$

che si ottiene sottraendo le due espansioni di Taylor sopra riportate. Inserendo le due approssimazioni nella equazione differenziale si ottiene la seguente equazione alle differenze:

$$\frac{D}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) + \frac{b}{2h}(u_{i+1} - u_{i-1}) = 0,$$

che può essere scritta per ogni nodo  $i$  della griglia computazionale. Tale equazione, che coincide con la (37), è caratterizzata dal secondo ordine di approssimazione ( $O(h^2)$ ).

Per la derivata prima si può ricavare un'approssimazione in avanti o una all'indietro esplicitando la (40) o la (41) in maniera opportuna, ottenendo la discretizzazione "upwind" o "downwind":

$$\begin{aligned} \frac{D}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) + \frac{b}{2h}(u_{i+1} - u_{i-1}) &= 0, \\ \frac{D}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) + \frac{b}{2h}(u_{i+1} - u_{i-1}) &= 0. \end{aligned}$$

## B Equazioni paraboliche

### Riferimenti bibliografici

- [1] G. Gambolati. *Lezioni di metodi numerici per ingegneria e scienze applicate*. Cortina, Padova, Italy, 2 edition, 2002. 619 pp.
- [2] R. Peyret and T. D. Taylor. *Computational Methods for Fluid Flow*. Springer-Verlag, New York, 1983.
- [3] A. Quarteroni. *Matematica Numerica per Problemi Differenziali*. Springer-Verlag Italia, Milano, 2003.
- [4] A. Quarteroni, R. Sacco, and F. Saleri. *Matematica Numerica*. Springer-Verlag Italia, Milano, 2008.
- [5] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1994.