



Università degli Studi di Cagliari
Dipartimento di Matematica e Informatica

**Dispense di Calcolo Numerico e
Matematica Applicata
per la laurea magistrale in Ingegneria per
l'Ambiente e il Territorio**

prof. Sebastiano Seatzu

(a cura del dottor Pietro Contu)

ultimo aggiornamento 7 aprile 2011

A.A. 2010/2011

Indice

1	Metodi analitici per la risoluzione delle PDE	1
1.1	Introduzione	1
1.2	Metodo degli integrali generali	5
1.2.1	Applicazione alle PDE	10
1.3	Metodo di separazione delle variabili	15
1.4	Temi proposti in prove d'esame	48
2	Metodi alle differenze finite	77
2.0.1	Equazioni ellittiche	80
2.0.2	Esercizi proposti	83
2.0.3	Equazioni paraboliche	85
2.0.4	Esercizi proposti	87
2.0.5	Equazioni iperboliche	88
2.0.6	Esercizi proposti	91
2.0.7	Modelli debolmente non lineari	93
2.0.8	Un'applicazione industriale delle differenze finite	97
3	Metodo agli elementi finiti	101
3.1	Introduzione	101
3.2	Generalità	102
3.3	Risoluzione di ODEs con il FEM	104
3.4	Risoluzione di PDEs con il FEM	109
3.5	Problema modello	122
3.5.1	Calcolo delle funzioni di base $\phi_l(x, y)$	126
3.5.2	Calcolo della stiffness matrix e del load vector	132
3.5.3	Come usare un triangolo di riferimento nel FEM	135

A Serie di Fourier	141
A.1 Funzioni periodiche e polinomi trigonometrici	141
A.2 Serie di Fourier	149
B Problema di Sturm-Liouville	159
C Metodi Iterativi	163
C.1 Introduzione	163
C.2 Metodi iterativi del primo ordine	164
C.2.1 Metodo di Jacobi o delle iterazioni simultanee	165
C.2.2 Metodo di Gauss-Seidel o delle iterazioni successive	167
C.3 Risoluzione numerica di sistemi nonlineari	170
C.3.1 Considerazioni sui metodi di risoluzione	170
C.3.2 Risultati sulla convergenza dei metodi iterativi	173
C.3.3 I metodi della tangente, e della secante	183
C.3.4 Altri metodi iterativi	191

Obiettivi: Far acquisire una conoscenza operativa di alcuni fra i principali metodi analitici e numerici concernenti i modelli differenziali alle derivate parziali del 2° ordine di tipo stazionario ed evolutivo. Per tale motivo l'illustrazione di ogni argomento avviene mediante lo svolgimento di esercizi. Per questo motivo non esiste alcuna distinzione formale tra le ore di lezione e quelle di esercitazione.

Capitolo 1

Metodi analitici per la risoluzione delle PDE

1.1 Introduzione

La trattazione presentata, anche se fondamentalmente riferita alle equazioni alle derivate parziali (PDE) del 2° ordine in 2 variabili, può estendersi in modo del tutto naturale ai modelli in più variabili. Allo scopo di rendere concreta questa affermazione, nel seguito verranno considerati alcuni esempi di PDE in 3 variabili. La generalizzazione a \mathbb{R}^n sarà quindi immediata.

Consideriamo la generica equazione alle derivate parziali lineari del 2° ordine in 2 variabili

$$A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + F u = G \quad (1.1)$$

con A, B, \dots, F e G in generale dipendenti da x, y ma non da u .

L'equazione (1.1) è detta *ellittica*, *parabolica* o *iperbolica* a seconda che la quantità

$$B^2 - 4AC$$

sia minore, uguale o maggiore di zero, rispettivamente.

Inoltre, la Eq. (1.1) si dice *debolmente non lineare* se il termine $F = F(x, y, u)$, con dipendenza lineare da u , mentre è *non lineare* se in almeno uno dei coefficienti A, B, \dots, G compare la non linearità della u e delle sue derivate parziali.

Esempi:

(1) Caso ellittico: (equazione di Laplace)

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

$[u(x, t)$ può rappresentare la temperatura in condizioni stazionarie, il potenziale gravitazionale oppure il potenziale elettrostatico];

(2) Caso parabolico: (equazione del calore)

$$\frac{\partial u}{\partial t} = k \frac{\partial^2 u}{\partial x^2}, \quad k = \frac{K}{\sigma \mu} = \frac{\text{conducibilità termica}}{\text{calore specifico} \times \text{densità}}$$

$[u(x, t)$: temperatura di un punto x di un intervallo $[a, b]$ all'istante t];

(3) Caso iperbolico: (equazione delle onde)

(3') *Corda vibrante*

$$\frac{\partial^2 y}{\partial t^2} = a^2 \frac{\partial^2 y}{\partial x^2}, \quad a^2 = \frac{\tau}{\mu} = \frac{\text{tensione}}{\text{densità}}$$

$[y(x, t)$: spostamento di un generico punto di una corda in una dimensione]

(3'') *Vibrazioni longitudinali di una trave*

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \quad c^2 = \frac{e}{\mu} = \frac{\text{elasticità}}{\text{densità}}$$

$[u(x, t)$: spostamento longitudinale dalla posizione di equilibrio della sezione trasversale della trave].

Se i coefficienti A , B e C sono costanti, la classificazione non dipende dal punto del dominio. Diversamente essa può dipendere dal punto, come mette in evidenza il seguente esempio dovuto a Tricomi:

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

In tal caso, essendo $A(x, y) = y$, $B(x, y) = 0$ e $C(x, y) = 1$, si ha

$$B^2(x, y) - 4A(x, y)C(x, y) = -4y,$$

per cui l'equazione è ellittica se $y > 0$, parabolica se $y = 0$, iperbolica se $y < 0$.

Nel caso generale di n variabili ($n \geq 2$) i modelli differenziali di riferimento, con le condizioni che assicurano l'esistenza e l'unicità della soluzione sono i seguenti:

a. **Modello ellittico**

$$\begin{cases} \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) + a_0 u = f, & x \in \Omega \subset \mathbb{R}^n \\ u = \varphi, & \text{per } x \in \partial\Omega \end{cases}$$

b. **Modello parabolico**

$$\begin{cases} \frac{\partial u}{\partial t} - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) + a_0 u = f, & (x, t) \in \Omega \times [0, T] \\ u = \varphi, & \text{su } \partial\Omega \times [0, T] \\ u(x, 0) = \psi, & x \in \Omega \subset \mathbb{R}^n \end{cases}$$

c. **Modello iperbolico**

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) + a_0 u = f, & (x, t) \in \Omega \times [0, T] \\ u = \varphi, & \text{su } \partial\Omega \times [0, T] \\ u(x, 0) = \psi_0, & x \in \Omega \subset \mathbb{R}^n \\ u_t(x, 0) = \psi_1, & x \in \Omega \subset \mathbb{R}^n \end{cases}$$

Le definizioni introdotte sono valide nell'ipotesi che i coefficienti $\{a_{ij}, a_0\}$ soddisfino le condizioni di ellitticità, ossia che risulti

$$\sum_{i,j=1}^n a_{ij}(x) \xi_i \xi_j \geq a_0 \|\xi\|^2 \text{ per } x \in \Omega \text{ e } \xi \in \mathbb{R}^n$$

essendo $\|\xi\|$ la norma euclidea del vettore ξ ($\|\xi\| = \sqrt{\xi_1^2 + \dots + \xi_n^2}$).

Definizioni.

L'ordine di un'equazione differenziale è identificato dall'ordine massimo delle derivate che in essa compaiono. Per *soluzione* si intende ogni funzione, sufficientemente regolare, che soddisfi l'equazione data.

L' *integrale generale* è la soluzione con un numero di funzioni arbitrarie pari all'ordine dell'equazione stessa.

L' *integrale particolare* è invece una soluzione ottenuta dalla generale con una particolarizzazione delle funzioni arbitrarie.

Esempi:

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

Integrando ordinatamente rispetto alla x e alla y si ottiene la soluzione generale

$$u(x, y) = x^2 y - \frac{1}{2} x y^2 + F(y) + G(x),$$

con F e G funzioni arbitrarie di classe C^2 .

Le funzioni

$$u_1(x, y) = x^2 y - \frac{1}{2} x y^2 + 3y^4 + \sin x - 5$$

e

$$u_2(x, y) = x^2 y - \frac{1}{2} x y^2 + y + \cos^2 x$$

sono invece due soluzioni particolari.

- $$u = x \frac{\partial u}{\partial x} - \left(\frac{\partial u}{\partial x} \right)^2$$

La soluzione generale è

$$u(x, y) = x F(y) - F^2(y), \quad F \text{ arbitraria di classe } C^2.$$

Le funzioni

$$u_1(x, y) = \frac{x^2}{4}$$

e

$$u_2(x, y) = x \sin y - \sin^2 y$$

sono due soluzioni particolari.

- $u_{xy} = \frac{\partial^2 u}{\partial x \partial y} = x \cos y$

La soluzione generale è

$$u(x, y) = \frac{x^2}{2} \sin y + F(y) + G(x),$$

con F e G di classe C^2 .

Le funzioni

$$u_1(x, y) = \frac{x^2}{2} \sin y - \frac{1}{2} y^3 + 3x^4 + \sinh x$$

e

$$u_2(x, y) = \frac{x^2}{2} \sin y - \cosh y + y - \cos^2 x$$

sono invece due soluzioni particolari.

Problema differenziale: equazione differenziale con condizioni al contorno e condizioni iniziali che assicurano l'unicità della soluzione.

1.2 Metodo degli integrali generali

- (1) *integrali generali* da particolarizzare in modo che siano soddisfatte le condizioni al contorno e/o iniziali;
- (2) *integrali particolari* per separazione delle variabili da cui partire per ottenere la soluzione del problema differenziale.

Prima di descrivere il metodo per le equazioni a derivate parziali, ricordiamo come si determina la soluzione generale nel caso di un'equazione differenziale ordinaria (ODE, ordinary differential equation) del 2° ordine. Consideriamo prima il caso di una ODE omogenea, avente cioè termine noto nullo e successivamente affronteremo il caso non omogeneo.

Caso Omogeneo

Prendiamo in considerazione l'equazione differenziale con assegnate condizioni agli estremi

$$\begin{cases} y'' + c^2 y = 0, & 0 < x < 1, \quad c \in \mathbb{R}^+ \\ y(0) = 1, \\ y(1) = 0 \end{cases}$$

In questo caso si cercano soluzioni del tipo $y(x) = e^{\alpha x}$. È immediato verificare che α deve soddisfare l'equazione $\alpha^2 + c^2 = 0$ (detta *equazione caratteristica*), ossia deve risultare $\alpha = \pm ic$, $c > 0$. La soluzione generale è dunque del tipo

$$y(x) = a e^{icx} + b e^{-icx}$$

che, nel campo reale (grazie alle formule di Eulero $e^{\pm icx} = \cos cx \pm i \sin cx$), esprimiamo nella forma

$$y(x) = a \cos cx + b \sin cx$$

con a e b costanti reali arbitrarie.

La soluzione generale contiene dunque 2 costanti arbitrarie, mentre nelle equazioni alle derivate parziali del 2° ordine contiene 2 funzioni arbitrarie. Le costanti vengono poi determinate imponendo che siano soddisfatte le condizioni agli estremi:

$$\begin{cases} y(0) = a = 1, \\ y(1) = \cos c + b \sin c = 0, \end{cases}$$

dalle quali si ricava

$$b = \begin{cases} -\cot c, & \text{se } \cos c \neq 0 \\ 0, & \text{se } \cos c = 0. \end{cases}$$

Più in generale, l'applicazione della tecnica dell'integrale generale alla risoluzione di un'equazione differenziale lineare omogenea (a coefficienti costanti) del 2° ordine con condizione agli estremi comporta il seguente procedimento:

$$\begin{cases} y'' + a_1 y' + a_2 y = 0, & a < x < b \\ y(a) = \alpha, \\ y(b) = \beta \end{cases},$$

dove a_1, a_2, α e β sono costanti reali note.

Ponendo $y(x) \cong e^{cx}$, si trova che c deve soddisfare l'*equazione caratteristica*

$$c^2 + a_1 c + a_2 = 0, \tag{1.2}$$

e dunque $c_{1,2} = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2}}{2}$. Si possono quindi presentare 3 casi a seconda del segno del discriminante $\Delta \stackrel{\text{def}}{=} a_1^2 - 4a_2$:

1. $\Delta > 0$

In questo caso c_1 e c_2 sono soluzioni reali e distinte dell'equazione caratteristica (3.2) e pertanto si trovano due soluzioni esponenziali. L'integrale generale è dato da una loro combinazione lineare:

$$y(x) = d_1 e^{c_1 x} + d_2 e^{c_2 x},$$

con d_1 e d_2 costanti reali.

2. $\Delta = 0$

In questo caso c_1 e c_2 sono soluzioni reali e coincidenti dell'equazione caratteristica (3.2) e si dimostra che l'integrale generale è dato da una combinazione lineare di e^{cx} e $x e^{cx}$, essendo $c \stackrel{\text{def}}{=} c_1 = c_2$:

$$y(x) = (d_1 + d_2 x) e^{cx},$$

dove d_1 e d_2 sono costanti reali.

3. $\Delta < 0$

In questo caso c_1 e c_2 sono soluzioni complesse e coniugate dell'equazione caratteristica (3.2) e si dimostra che l'integrale generale è dato da:

$$y(x) = e^{\frac{a_1}{2} x} \left(d_1 \cos \frac{\sqrt{4a_2 - a_1^2}}{2} x + d_2 \sin \frac{\sqrt{4a_2 - a_1^2}}{2} x \right),$$

con $d_1, d_2 \in \mathbb{R}$.

Le costanti d_1 e d_2 vengono poi determinate imponendo le condizioni agli estremi:

$$\begin{cases} y(a) = \alpha, \\ y(b) = \beta. \end{cases}$$

L'esempio già considerato rientra nel secondo caso con $c_{1,2} = \pm i c$. Consideriamo ora i seguenti esempi.

$$\begin{cases} y'' + 3\sqrt{3}y' + 6y = 0, & 0 < x < 1 \\ y(0) = 1, \\ y(1) = 2. \end{cases}$$

Cerchiamo soluzioni del tipo $y(x) \cong e^{\alpha x}$. E' immediato verificare che α deve soddisfare l'equazione

$$\alpha^2 + 3\sqrt{3}\alpha + 6 = (\alpha + \sqrt{3})(\alpha + 2\sqrt{3}) = 0,$$

per cui l'integrale generale é

$$y(x) = c_1 e^{-\sqrt{3}x} + c_2 e^{-2\sqrt{3}x},$$

con c_1 e c_2 costanti arbitrarie che si ricavano dal sistema

$$\begin{cases} c_1 + c_2 = 1 \\ c_1 e^{-\sqrt{3}} + c_2 e^{-2\sqrt{3}} = 2. \end{cases}$$

Da notare che tale sistema (di equazioni lineari) possiede una sola soluzione poiché il determinante della matrice dei coefficienti é dato da $e^{-2\sqrt{3}} - e^{-\sqrt{3}} \neq 0$.

Esempio 1.2.1

$$\begin{cases} y'' + 2y' + y = 0, & 0 < x < 1 \\ y(0) = 1, \\ y(1) = 2 \end{cases}$$

Si cercano soluzioni del tipo $y(x) \cong e^{\alpha x}$. In questo caso α deve soddisfare l'equazione

$$\alpha^2 + 2\alpha + 1 = (\alpha + 1)^2 = 0,$$

per cui l'integrale generale é

$$y(x) = (c_1 + c_2 x) e^{-x},$$

con c_1 e c_2 costanti arbitrarie che si ricavano dal sistema

$$\begin{cases} c_1 = 1 \\ (c_1 + c_2) e^{-1} = 2, \end{cases}$$

che ha una e una sola soluzione.

Caso Non Omogeneo

Si consideri la seguente EDO non omogenea con condizioni agli estremi:

$$\begin{cases} y'' + a_1 y' + a_2 y = f(x), & a < x < b \\ y(a) = \alpha, \\ y(b) = \beta \end{cases},$$

dove a_1, a_2, α e β sono costanti reali.

L'integrale generale è dato dalla somma della soluzione $\widehat{y}(x) = d_1 y_1(x) + d_2 y_2(x)$ della equazione omogenea con una *soluzione particolare* $\overline{y}(x)$ della equazione non omogenea:

$$y(x) = d_1 y_1(x) + d_2 y_2(x) + \overline{y}(x),$$

con d_1 e d_2 costanti reali. La ricerca di un integrale particolare $\overline{y}(x)$ dipende fortemente dalla forma analitica della $f(x)$ come illustra il seguente esempio.

Esempio 1.2.2

$$\begin{cases} y'' + 3y = x^2 + 1, & 0 < x < 1 \\ y(0) = 1, \\ y(1) = 0 \end{cases}$$

Si trova facilmente che la soluzione dell'equazione omogenea è data da:

$$\widehat{y}(x) = d_1 \cos(\sqrt{3}x) + d_2 \sin(\sqrt{3}x). \quad (1.3)$$

Per trovare l'integrale particolare $\overline{y}(x)$ notiamo che una scelta ragionevole è prendere un polinomio di secondo grado in base alla forma del termine noto $f(x) = x^2 + 1$:

$$\overline{y}(x) = a_0 x^2 + a_1 x + a_2. \quad (1.4)$$

Sostituendo la (3.5) nella ODE si trovano le condizioni per fissare i coefficienti a_0, a_1 e a_2 :

$$2a_0 + 3a_0 x^2 + 3a_1 x + 3a_2 = x^2 + 1,$$

pertanto si ha

$$\begin{cases} 3a_0 = 1 \\ 3a_1 = 0 \\ 2a_0 + 3a_2 = 1 \end{cases} \Rightarrow \begin{cases} a_0 = 1/3 \\ a_1 = 0 \\ a_2 = 1/9 \end{cases}.$$

Di conseguenza la soluzione generale è:

$$y(x) = d_1 \cos(\sqrt{3}x) + d_2 \sin(\sqrt{3}x) + \frac{1}{3}x^2 + \frac{1}{9}, \quad (1.5)$$

e imponendo le condizioni al contorno si ottiene:

$$\begin{cases} y(0) = 1 \\ y(1) = 0 \end{cases} \Rightarrow \begin{cases} d_1 + 1/9 = 1 \\ d_1 \cos(\sqrt{3}) + d_2 \sin(\sqrt{3}) + \frac{1}{3} + \frac{1}{9} = 0 \end{cases} \Rightarrow \begin{cases} d_1 = 8/9 \\ d_2 = -\frac{4}{9} \frac{2 \cos \sqrt{3} + 1}{9 \sin \sqrt{3}} \end{cases}.$$

Si ha quindi:

$$y(x) = \frac{8}{9} \cos(\sqrt{3}x) - \frac{4}{9} \frac{2 \cos \sqrt{3} + 1}{9 \sin \sqrt{3}} \sin(\sqrt{3}x) + \frac{1}{3}x^2 + \frac{1}{9}.$$

Descriviamo ora il metodo relativamente alle equazioni alle derivate parziali (PDE partial differential equations), dando un particolare risalto al metodo di D'Alembert che rappresenta l'estensione naturale del metodo illustrato per le equazioni differenziali ordinarie.

1.2.1 Applicazione alle PDE

Tenuto conto delle finalità operative del corso, il metodo verrà illustrato mediante la sua applicazione diretta a varie equazioni differenziali.

Esempio 1.2.3 Consideriamo l'equazione differenziale del secondo ordine omogenea:

$$3u_{tt} = 10u_{xx}. \quad (1.6)$$

Estendendo il ragionamento adottato per le ODE, cerchiamo una soluzione della forma $u(x, t) = e^{ax+bt}$ che, sostituita nella equazione differenziale, fornisca la seguente relazione tra le costanti reali a e b (metodo di d'Alembert):

$$3b^2 = 10a^2, \Rightarrow b = \pm \sqrt{\frac{10}{3}}a.$$

Questo implica che nella equazione differenziale data (equazione delle onde) la variabile spaziale e temporale sono correlate. Ipotizziamo ora che la soluzione generale della (3.6) è

$$u(x, t) = F\left(x - \sqrt{\frac{10}{3}}t\right) + G\left(x + \sqrt{\frac{10}{3}}t\right), \quad (1.7)$$

con F e G di classe C^2 . Per verificare se la (3.18) soddisfa l'Eq. (3.6) sostituiamo, osservando che

$$u_t = \sqrt{\frac{10}{3}} \left[-F' \left(x - \sqrt{\frac{10}{3}}t \right) + G' \left(x + \sqrt{\frac{10}{3}}t \right) \right],$$

$$u_x = F' \left(x - \sqrt{\frac{10}{3}}t \right) + G' \left(x + \sqrt{\frac{10}{3}}t \right),$$

e analogamente si ha:

$$u_{tt} = \frac{10}{3} \left[F'' \left(x - \sqrt{\frac{10}{3}}t \right) + G'' \left(x + \sqrt{\frac{10}{3}}t \right) \right],$$

$$u_{xx} = F'' \left(x - \sqrt{\frac{10}{3}}t \right) + G'' \left(x + \sqrt{\frac{10}{3}}t \right).$$

Sostituendo nella (3.6) si trova:

$$3u_{tt} = 10u_{xx}.$$

Esempio 1.2.4

$$\begin{cases} \frac{\partial u}{\partial x} + 3 \frac{\partial u}{\partial y} = 0 \\ u(0, y) = 4 \sin y \end{cases}$$

Per la costruzione dell'integrale generale, seguendo il metodo di D'Alembert, si osserva preliminarmente che una funzione del tipo

$$u(x, y) = e^{ax+by}$$

è soluzione dell'equazione differenziale se $a = -3b$.

Si osserva poi, di conseguenza, che l'integrale generale è della forma

$$u(x, y) = F(y - 3x), \quad \text{con } F \text{ differenziabile rispetto a } x \text{ e } y.$$

La condizione al contorno richiede che

$$u(0, y) = F(y) = 4 \sin y$$

il che implica che la soluzione del problema differenziale è

$$u(x, y) = F(y - 3x) = 4 \sin(y - 3x).$$

Esempio 1.2.5

$$\begin{cases} \frac{\partial^2 u}{\partial x \partial y} = x^2 y \\ u(x, 0) = x^2 \\ u(1, y) = \cos y \end{cases}$$

Integrando ordinatamente rispetto alla x e alla y si ottiene l'integrale generale

$$u(x, y) = \frac{1}{6} x^3 y^2 + F(y) + G(x).$$

Per le condizioni al contorno deve essere

$$u(x, 0) = F(0) + G(x) = x^2 \quad \Rightarrow \quad G(x) = x^2 - F(0)$$

e

$$u(1, y) = \frac{1}{6} y^2 + F(y) + G(1) = \cos y \quad \Rightarrow \quad F(y) = \cos y - \frac{1}{6} y^2 - G(1)$$

dalle quali si ottiene

$$u(x, y) = \frac{1}{6} x^3 y^2 + \cos y - \frac{1}{6} y^2 - G(1) + x^2 - F(0), \quad \text{con } G(1) + F(0) = 1.$$

La soluzione del problema differenziale è dunque

$$u(x, y) = \frac{1}{6} x^3 y^2 + \cos y - \frac{1}{6} y^2 + x^2 - 1.$$

Esempio 1.2.6

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

Integrando rispetto alla variabile x si ha

$$t \frac{\partial u}{\partial t} + 2u = \frac{1}{3} x^3 + F(t), \quad F \text{ differenziabile rispetto a } t.$$

Osservando ora che $\left(t^2 \frac{\partial u}{\partial t} + 2u t \right) = \frac{\partial}{\partial t} (t^2 u)$ si ottiene

$$\frac{\partial}{\partial t} (t^2 u) = \frac{1}{3} x^3 t + t F(t)$$

$$t^2 u = \frac{1}{6} x^3 t^2 + \hat{F}(t) + H(x), \quad H \text{ differenziabile rispetto a } x.$$

La soluzione generale dell'equazione è pertanto

$$u(x, y) = \frac{1}{6} x^3 + G(t) + \frac{H(x)}{t^2}.$$

Esempio 1.2.7

$$\begin{cases} 4 \frac{\partial^2 u}{\partial t^2} = 25 \frac{\partial^2 u}{\partial x^2}, & 0 < x < \pi, \quad t > 0 \\ u(x, 0) = \sin 2x \\ u_t(x, 0) = 0 \\ u(0, t) = u(\pi, t) = 0 \end{cases}$$

Si osserva preliminarmente che ogni funzione

$$u(x, y) = e^{ax+bt}$$

è soluzione dell'equazione data se $b = \pm \frac{5}{2} a$.

L'integrale generale è dunque della forma

$$u(x, t) = F(2x + 5t) + G(2x - 5t), \quad F \text{ e } G \text{ differenziabili rispetto a } x \text{ e } t.$$

Dalle condizioni al contorno

$$\begin{cases} u(x, 0) = F(2x) + G(2x) = \sin 2x \\ u_t(x, 0) = 5[F'(2x) - G'(2x)] = 0 \end{cases}$$

segue che

$$\begin{aligned} F(2x) &= \frac{1}{2} \sin 2x + c_1 \\ G(2x) &= \frac{1}{2} \sin 2x + c_2 \end{aligned}$$

purchè sia $c_1 = -c_2$.

Infine, imponendo le condizioni iniziali, si ottiene la soluzione

$$u(x, t) = \frac{1}{2} \sin(2x + 5t) + \frac{1}{2} \sin(2x - 5t).$$

Esempio 1.2.8

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} & \text{con } 0 < x < \pi \text{ e } t > 0 \\ u(0, t) + u(\pi, t) = 0 \\ u(x, 0) = \sin(x) \\ u_t(x, 0) = \cos(x) \end{cases}$$

Per la costruzione dell'integrale generale si osserva che una funzione del tipo

$$u(x, t) = e^{ax+bt}$$

soddisfa l'equazione se $b = \pm ca$. Per cui $u(x, t) = e^{a(x \pm ct)}$.

Quindi l'integrale generale sarà della forma

$$u(x, t) = F(x + ct) + G(x - ct)$$

con F e G derivabili almeno 2 volte rispetto a x e a t .

Poiché $u_t(x, t) = cF'(x+ct) - cG'(x-ct)$, dalle condizioni iniziali troviamo il seguente sistema:

$$\begin{cases} u(x, 0) = F(x) + G(x) = \sin(x) \\ u_t(x, 0) = c[F'(x) - G'(x)] = \cos(x). \end{cases}$$

Derivando la prima equazione si ha $F'(x) + G'(x) = \cos(x)$ per cui, tenuto conto della seconda equazione è possibile ricavare $F'(x)$ e $G'(x)$. Più precisamente si trova $F'(x) = \frac{1}{2} \left(1 + \frac{1}{c}\right) \cos(x)$ e $G'(x) = \frac{1}{2} \left(1 - \frac{1}{c}\right) \cos(x)$. Integrando risulta

$$\begin{aligned} F(x) &= \frac{1}{2} \left(1 + \frac{1}{c}\right) \sin(x) + c_1, \\ G(x) &= \frac{1}{2} \left(1 - \frac{1}{c}\right) \sin(x) + c_2, \end{aligned}$$

dove c_1 e c_2 sono costanti arbitrarie che, per la condizione $F(x) + G(x) = \sin(x)$, devono soddisfare il vincolo $c_1 + c_2 = 0$. Posso quindi scrivere

$$\begin{aligned} F(x + ct) &= \frac{1}{2} \left(1 + \frac{1}{c}\right) \sin(x + ct) + c_1, \\ G(x - ct) &= \frac{1}{2} \left(1 - \frac{1}{c}\right) \sin(x - ct) - c_1, \end{aligned}$$

Perciò l'integrale generale diventa

$$u(x, t) = \frac{1}{2} \left(1 + \frac{1}{c} \right) \sin(x + ct) + \frac{1}{2} \left(1 - \frac{1}{c} \right) \sin(x - ct).$$

Rimane ancora da imporre la condizione $u(0, t) + u(\pi, t) = 0$. Si ha

$$\begin{aligned} u(0, t) + u(\pi, t) &= \frac{1}{2} \left\{ \left(1 + \frac{1}{c} \right) \sin(ct) - \left(1 - \frac{1}{c} \right) \sin(ct) + \right. \\ &\quad \left. \left(1 + \frac{1}{c} \right) \sin(\pi + ct) + \left(1 - \frac{1}{c} \right) \sin(\pi - ct) \right\} \\ &= \frac{1}{2} \left\{ \frac{2}{c} \sin(ct) - \left(1 + \frac{1}{c} \right) \sin(ct) + \left(1 - \frac{1}{c} \right) \sin(ct) \right\} = 0. \end{aligned}$$

Quindi questa condizione è identicamente soddisfatta, per ogni t , dalla soluzione trovata. Rimane da verificare che la soluzione trovata soddisfi il modello iniziale. Essendo

$$\begin{aligned} u_{tt}(x, t) &= -\frac{c^2}{2} \left[\left(1 + \frac{1}{c} \right) \sin(x + ct) + \left(1 - \frac{1}{c} \right) \sin(x - ct) \right], \\ u_{xx}(x, t) &= -\frac{1}{2} \left[\left(1 + \frac{1}{c} \right) \sin(x + ct) + \left(1 - \frac{1}{c} \right) \sin(x - ct) \right] \end{aligned}$$

l'equazione $\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$ è soddisfatta. La verifica delle due condizioni $u(x, 0) = \sin(x)$ e $u_t(x, 0) = \cos(x)$ è infine immediata.

1.3 Metodo di separazione delle variabili

Il metodo consiste nel ricondurre un assegnato problema alle derivate parziali, **dotato di un'unica soluzione**, alla risoluzione di più problemi alle derivate ordinarie. Naturalmente esso non è sempre applicabile. Esso è infatti fondamentalmente utilizzato, anche se non esclusivamente, per risolvere problemi a derivate parziali con coefficienti costanti e con domini regolari. Quando è applicabile è in grado di fornire la soluzione esatta del problema a differenza di quanto avviene con i metodi numerici. Questi ultimi, infatti, sono in grado di determinare soltanto delle approssimazioni più o meno rappresentative della soluzione analitica. Nei casi più complicati il ricorso

ai metodi numerici è comunque inevitabile. Tuttavia, anche in questi casi il ricorso ai metodi analitici è essenziale per verificare l'effettività dei codici di calcolo che si intende utilizzare. A tale scopo tipicamente si procede nel modo seguente:

- (a) si risolvono analiticamente uno o più problemi, il più possibile vicini a quello da risolvere numericamente;
- (b) si risolvono numericamente i problemi semplificati e si confrontano i risultati ottenuti con le soluzioni analitiche;
- (c) si procede alla risoluzione numerica del problema di effettivo interesse con il codice disponibile dopo aver verificato che l'errore commesso nei problemi semplificati sia del tutto accettabile.

I seguenti esempi mostrano l'applicazione del metodo alla risoluzione di vari problemi.

Esempio 1.3.1

$$\begin{cases} \frac{\partial u}{\partial x} = 4 \frac{\partial u}{\partial y} \\ u(0, y) = 8 e^{-3y} \end{cases}, \quad (x, y) \in \mathbb{R}^2$$

Dalla teoria delle equazioni a derivate parziali è noto che il problema dato ammette una e una sola soluzione. La tecnica consiste nel cercare un integrale dell'equazione a variabili separate:

$$u(x, y) = X(x)Y(y), \quad X \text{ e } Y \text{ differenziabili.}$$

Sostituendo e riordinando si ha

$$\begin{aligned} X'Y &= 4XY', \\ \frac{X'}{4X} &= \frac{Y'}{Y}. \end{aligned} \tag{1.8}$$

Il primo membro della (3.19) dipende esclusivamente dalla variabile x , mentre il secondo membro dalla variabile y . Se due quantità dipendenti rispettivamente da due variabili indipendenti hanno sempre lo stesso valore, sono inevitabilmente uguali ad una costante. Indicando con $\lambda \in \mathbb{R}$ il valore della costante, dalla (3.19) si ottiene il sistema

$$\begin{cases} X' - 4\lambda X = 0 \\ Y' - \lambda Y = 0 \end{cases}, \quad \text{da cui} \quad \begin{cases} X = e^{4\lambda x} \\ Y = e^{\lambda y} \end{cases}.$$

Ossia

$$u(x, y) = c e^{\lambda(4x+y)}.$$

Imponendo infine la condizione $u(0, y) = 8 e^{-3y}$ si ha $c = 8$ e $\lambda = 3$, da cui

$$u(x, y) = 8 e^{3(4x+y)}.$$

$$\begin{cases} \frac{\partial u}{\partial t} = 2 \frac{\partial^2 u}{\partial x^2}, & 0 < x < 3, & t > 0, & u(x, t) \text{ limitata} \\ u(0, t) = u(3, t) = 0 \\ u(x, 0) = 5 \sin 4\pi x - 3 \sin 8\pi x + 2 \sin 10\pi x. \end{cases} \quad (1.9)$$

Si tratta dell'equazione del calore in una dimensione spaziale che ammette una e una sola soluzione, visto che sono dati i valori di temperatura agli estremi della sbarra e la distribuzione di temperatura all'istante iniziale lungo tutta la sbarra. Cerchiamo, con il metodo di separazione delle variabili, di trovare una soluzione del tipo:

$$u(x, y) = X(x)T(t), \quad X \text{ e } T \text{ differenziabili.}$$

Sostituendo e riordinando si ha

$$X T' = 2 X'' T \quad \text{ossia} \quad \frac{T'}{2T} = \frac{X''}{X}.$$

Imponendo che entrambi i rapporti siano uguali a un parametro λ , si ottiene il sistema

$$\begin{cases} X''(x) - \lambda X(x) = 0 \\ T'(t) - 2\lambda T(t) = 0 \end{cases}.$$

Utilizzando le condizioni al contorno date agli estremi, si ha

$$\begin{cases} u(0, t) = X(0)T(t) = 0 \\ u(3, t) = X(3)T(t) = 0 \end{cases} \Rightarrow \begin{cases} X(0) = 0 \\ X(3) = 0 \end{cases},$$

perché $T(t) = 0$ per ogni t implicherebbe $u(x, t) = X(x)T(t) = 0$ per ogni (x, t) nel dominio e non è accettabile in quanto non soddisfa la condizione iniziale $u(x, 0) = 5 \sin 4\pi x - 3 \sin 8\pi x + 2 \sin 10\pi x$.

Di conseguenza la $X(x)$ deve soddisfare il problema spettrale:

$$\begin{cases} X''(x) - \lambda X(x) = 0 \\ X(0) = 0 \quad X(3) = 0 \end{cases} \quad x \in [0, 3]. \quad (1.10)$$

Il problema (3.21) è un problema di Sturm-Liouville con condizioni al contorno di Cauchy, in quanto si cerca una soluzione che assume valori nulli agli estremi del dominio. Dalla teoria delle ODE sappiamo che ammette una e una sola soluzione per ogni prefissato λ . Per la sua risoluzione occorre distinguere tra i seguenti casi:

- $\lambda > 0$

In questo caso l'equazione caratteristica ammette due radici reali e distinte con una soluzione esponenziale del tipo:

$$X(x) = ae^{\sqrt{\lambda}x} + be^{-\sqrt{\lambda}x}.$$

Imponendo le condizioni al contorno si trova il seguente sistema lineare omogeneo

$$\begin{pmatrix} 1 & 1 \\ e^{3\sqrt{\lambda}} & e^{-3\sqrt{\lambda}} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Il determinante della matrice vale $e^{-3\sqrt{\lambda}} - e^{3\sqrt{\lambda}} = -2\sinh(3\sqrt{\lambda}) \neq 0 \forall \lambda > 0$. Di conseguenza $a = b = 0$, il che implica la sola soluzione banale ovviamente da scartare perché non soddisferebbe la condizione iniziale $u(x, 0) = 5 \sin 4\pi x - 3 \sin 8\pi x + 2 \sin 10\pi x$.

- $\lambda = 0$

L'equazione caratteristica ammette due soluzioni reali e coincidenti, per cui

$$X(x) = e^{\sqrt{\lambda}x}(a + bx).$$

Applicando le condizioni al contorno, si trova subito che $a = b = 0$, ossia una soluzione banale non accettabile.

- $\lambda < 0$

In questo caso l'equazione caratteristica possiede due soluzioni complesse coniugate e pertanto la soluzione generale:

$$X(x) = a \cos(\sqrt{\lambda}x) + b \sin(\sqrt{\lambda}x),$$

Imponendo che siano soddisfatte le condizioni agli estremi, si ottiene

$$X(0) = X(3) = 0 \Rightarrow \begin{cases} a = 0 \\ b \sin(3\sqrt{\lambda}) = 0 \end{cases} .$$

Di conseguenza si ha una infinità numerabile di autovalori λ_k a cui corrisponde una infinità numerabile di autofunzioni $X_k(x)$ e precisamente:

$$\lambda_k = \left(\frac{k\pi}{3}\right)^2, \quad X_k(x) = \sin\left(\frac{k\pi}{3}x\right), \quad k \in \mathbb{N} = \{1, 2, 3, \dots\}. \quad (1.11)$$

Avendo risolto il problema di Sturm-Liouville per la soluzione $X(x)$, dobbiamo determinare per ogni λ_k la corrispondente soluzione $T_k(t)$ della ODE

$$T_k'(t) - 2\lambda_k T_k(t) = 0. \quad (1.12)$$

La soluzione della (1.12) è

$$T_k(t) = \text{cost } e^{2\lambda_k t} = \text{cost } e^{[2(k\pi)^2/9]t},$$

per cui

$$u_k(x, t) = \text{cost } e^{[2(k\pi)^2/9]t} \sin\left(\frac{k\pi}{3}x\right), \quad k \in \mathbb{N}.$$

La soluzione generale del problema a derivate parziali dato è una combinazione lineare delle $u_k(x, t)$, ossia la serie:

$$u(x, t) = \sum_{k=1}^{\infty} c_k u_k(x, t) = \sum_{k=1}^{\infty} c_k e^{[2(k\pi)^2/9]t} \sin\left(\frac{k\pi}{3}x\right). \quad (1.13)$$

Imponendo infine l'ultima condizione all'istante iniziale $u(x, 0) = 5 \sin 4\pi x - 3 \sin 8\pi x + 2 \sin 10\pi x$, si ha

$$\sum_{k=1}^{\infty} c_k \sin\left(\frac{k\pi}{3}x\right) = 5 \sin 4\pi x - 3 \sin 8\pi x + 2 \sin 10\pi x,$$

e questo implica che soltanto tre addendi della serie sono diversi da zero, con i coefficienti:

$$\begin{cases} c_{12} = 5 \\ c_{24} = -3 \\ c_{30} = 2 \end{cases} .$$

Di conseguenza l'unica soluzione del problema a derivate parziali (1.9) è la seguente

$$u(x, t) = 5e^{[2(12\pi)^2/9]t} \sin 4\pi x - 3e^{[2(24\pi)^2/9]t} \sin 8\pi x + 2e^{[2(30\pi)^2/9]t} \sin 10\pi x.$$

Esempio 1.3.2

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - 2u, & 0 < x < 3, \quad t > 0, \quad u(x, t) \text{ limitata} \\ u(0, t) = u(3, t) = 0 \\ u(x, 0) = 2 \sin \pi x - \sin 4\pi x. \end{cases} \quad (1.14)$$

Questo problema è una variante dell'Es.1.9 che, come sappiamo, ammette una sola soluzione che cerchiamo col metodo della separazione delle variabili. A tale scopo poniamo

$$u(x, t) = X(x)T(t) \Rightarrow \frac{T'}{T} = \frac{X''}{X} - 2 \Rightarrow \begin{cases} X'' - (2 - \lambda)X = 0 \\ T' + \lambda T = 0 \end{cases}$$

Applicando le condizioni agli estremi ($u(0, t) = u(3, t) = 0$) come abbiamo fatto nell'Es.1.9, troviamo il problema di Sturm-Liouville:

$$\begin{cases} X''(x) - (2 - \lambda)X(x) = 0 & x \in [0, 3], \\ X(0) = 0 \quad X(3) = 0 \end{cases} \quad (1.15)$$

la cui soluzione varia a seconda del valore di λ . Come per l'Es.1.9 ci aspettiamo che le uniche soluzioni accettabili siano oscillanti:

$$2 - \lambda < 0 \Rightarrow X(x) = a \cos(\sqrt{\lambda - 2}x) + b \sin(\sqrt{\lambda - 2}x),$$

e

$$\begin{cases} X(0) = 0 \\ X(3) = 0 \end{cases} \Rightarrow \begin{cases} a = 0 \\ b \sin(3\sqrt{\lambda - 2}) = 0 \end{cases}.$$

Di conseguenza si ha una infinità numerabile di valori di autovalori λ_k e di autofunzioni $X_k(x)$:

$$\lambda_k = \left(\frac{k\pi}{3}\right)^2 + 2, \quad X_k(x) = \sin\left(\frac{k\pi}{3}x\right), \quad k \in \mathbb{N}. \quad (1.16)$$

Avendo risolto il problema di Sturm-Liouville per la $X(x)$, possiamo determinare le corrispondenti soluzioni $T_k(t)$, essendo:

$$T'_k(t) + \lambda_k T_k(t) = 0.$$

Si tratta di una ODE del primo ordine la cui soluzione è

$$T_k(t) = \text{cost } e^{-\lambda_k t}, \quad \lambda_k = \left(\frac{k\pi}{3}\right)^2 + 2, \quad k \in \mathbb{N}.$$

Di conseguenza, per ogni $k = 1, 2, 3, \dots$, abbiamo

$$u_k(x, t) = \text{cost } e^{-\lambda_k t} \sin\left(\frac{k\pi}{3}x\right).$$

La soluzione generale del problema a derivate parziali dato è una combinazione lineare delle $u_k(x, t)$ (serie di funzioni) del tipo:

$$u(x, t) = \sum_{k=1}^{\infty} c_k u_k(x, t) = \sum_{k=1}^{\infty} c_k e^{-\lambda_k t} \sin\left(\frac{k\pi}{3}x\right). \quad (1.17)$$

Imponendo infine l'ultima condizione all'istante iniziale $u(x, 0) = 2 \sin \pi x - \sin 4\pi x$, si ottiene

$$\sum_{k=1}^{\infty} c_k \sin\left(\frac{k\pi}{3}x\right) = 2 \sin \pi x - \sin 4\pi x,$$

il che implica che soltanto due addendi della serie siano diversi da zero con:

$$\begin{cases} c_3 = 2 \\ c_{12} = -1 \end{cases}.$$

L'unica soluzione del problema a derivate parziali 1.3.2 è pertanto

$$u(x, t) = 2e^{[2(3\pi)^2/9]t} \sin \pi x - e^{[2(12\pi)^2/9]t} \sin 4\pi x.$$

Esempio 1.3.3

$$\begin{cases} \frac{\partial u}{\partial t} = 2 \frac{\partial^2 u}{\partial x^2} \\ u(0, t) = 10, \quad u(3, t) = 40 \\ u(x, 0) = 6 \sin \pi x + 3(x^2 + 1) \end{cases}, \quad (1.18)$$

con $x \in [0, 3]$ e $t > 0$.

In questo esempio si tratta nuovamente un problema parabolico (equazione

del calore) con condizioni agli estremi non omogenee. In questi casi è sempre possibile convertire il problema dato in uno avente condizioni agli estremi omogenee. È sufficiente effettuare la sostituzione:

$$u(x, t) = v(x, t) + \varphi(x), \quad (1.19)$$

con $\varphi(x) \in C^2([0, 3])$. Sostituendo la (1.19) nel problema (1.18) si ha

$$\begin{cases} \partial_t v = 2\partial_{xx}^2 v + \varphi''(x) \\ v(0, t) + \varphi(0) = 10, \quad v(3, t) + \varphi(3) = 40 \\ v(x, 0) + \varphi(x) = 6 \sin \pi x + 3(x^2 + 1) \end{cases} .$$

Di conseguenza per avere un problema differenziale con soluzione della stessa PDE e con condizioni agli estremi omogenee, è sufficiente scegliere $\varphi(x)$ tale che

$$\begin{cases} \varphi''(x) = 0 \\ \varphi(0) = 10, \quad \varphi(3) = 40 \end{cases} , \quad (1.20)$$

la cui soluzione è $\varphi(x) = 10x + 10$. Il problema (1.18) diventa quindi

$$\begin{cases} \partial_t v = 2\partial_{xx}^2 v \\ v(0, t) = 0, \quad v(3, t) = 0 \\ v(x, 0) = 6 \sin \pi x + 3(x^2 + 1) - 10(x + 1) = 6 \sin \pi x + 3x^2 - 10x - 7 \end{cases} . \quad (1.21)$$

Procedendo per separazione delle variabili, come nell'esempio precedente, si trova che la $v(x, t)$ è rappresentata dalla serie

$$v(x, t) = \sum_{k=1}^{\infty} c_k v_k(x, t) = \sum_{k=1}^{\infty} c_k e^{-\lambda_k t} \sin\left(\frac{k\pi}{3}x\right). \quad (1.22)$$

Imponendo infine l'ultima condizione all'istante iniziale $u(x, 0) = 2 \sin \pi x - \sin 4\pi x$, si ha

$$\sum_{k=1}^{\infty} c_k \sin\left(\frac{k\pi}{3}x\right) = 6 \sin \pi x + 3x^2 - 10x - 7,$$

che implica che un addendo della serie viene fissato dal termine $6 \sin \pi x$:

$$c_3 = 6,$$

mentre le altre costanti c_k con $k \in \mathbb{N} \setminus \{3\}$ devono essere calcolate sfruttando la conoscenza del termine polinomiale $3x^2 - 10x - 7$:

$$\sum_{\substack{k=1 \\ k \neq 3}}^{\infty} c_k \sin k \frac{\pi}{3} x = 3x^2 - 10x - 7. \quad (1.23)$$

La (3.28) è una serie di Fourier e per isolare i coefficienti c_k occorre moltiplicare primo e secondo membro per $\sin(m\pi/3)x$ ($m \in \mathbb{N} \setminus \{0, 3\}$) e integrare su tutto il dominio (si sfrutta l'ortogonalità delle funzioni $\sin(m\pi/3)x$):

$$\int_0^3 \sin m \frac{\pi}{3} x \left(\sum_{\substack{k=1 \\ k \neq 3}}^{\infty} c_k \sin k \frac{\pi}{3} x \right) dx = \int_0^3 \left(\sin m \frac{\pi}{3} x \right) (3x^2 - 10x - 7) dx. \quad (1.24)$$

Nella teoria delle serie di Fourier si dimostra che la serie nelle Eq. (3.29) è uniformemente convergente e pertanto è possibile portare l'integrale dentro il segno di serie:

$$\sum_{\substack{k=1 \\ k \neq 3}}^{\infty} \int_0^3 c_k \left(\sin m \frac{\pi}{3} x \right) \left(\sin k \frac{\pi}{3} x \right) dx = \int_0^3 \left(\sin m \frac{\pi}{3} x \right) (3x^2 - 10x - 7) dx. \quad (1.25)$$

Per calcolare il primo membro della (3.30) si ricorre alle formule di prostaferesi

$$\sin \alpha \sin \beta = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)],$$

ottenendo

$$\begin{aligned} \int_0^3 c_k \left(\sin m \frac{\pi}{3} x \right) \left(\sin k \frac{\pi}{3} x \right) dx &= \frac{c_k}{2} \int_0^3 \left[\cos \frac{\pi}{3} (m - k)x - \cos \frac{\pi}{3} (m + k)x \right] dx \\ &= \frac{c_k}{2} \left[\frac{\sin \frac{\pi}{3} (m - k)x}{\frac{\pi}{3} (m - k)} \Big|_0^3 - \frac{\sin \frac{\pi}{3} (m + k)x}{\frac{\pi}{3} (m + k)} \Big|_0^3 \right] = \frac{3}{2} c_k \delta_{mk}, \end{aligned}$$

dove δ_{mk} è il delta di Kronecker che vale zero se $m \neq k$ e uno se $m = k$.

L'integrale al secondo membro della (3.30) si può risolvere mediante

integrazione per parti:

$$\begin{aligned}
 3 \int_0^3 x^2 \sin\left(\frac{m\pi}{3}x\right) dx &= 3 \left[-x^2 \frac{3}{m\pi} \cos\left(\frac{m\pi}{3}x\right) \right]_0^3 + 2 \frac{3^2}{m\pi} \int_0^3 x \cos\left(\frac{m\pi}{3}x\right) dx \\
 &= \frac{3^4}{m\pi} (-1)^{m+1} - 2 \frac{3^2}{m\pi} \left[-\frac{3}{m\pi} \int_0^3 \sin\left(\frac{m\pi}{3}x\right) dx \right] \\
 &= \frac{3^4}{m\pi} (-1)^{m+1} + 2 \frac{3^3}{(m\pi)^2} [(-1)^m - 1].
 \end{aligned}$$

$$\begin{aligned}
 10 \int_0^3 x \sin\left(\frac{m\pi}{3}x\right) dx &= 10 \frac{3^2}{m\pi} (-1)^{m+1}, \\
 7 \int_0^3 \sin\left(\frac{m\pi}{3}x\right) dx &= 7 \frac{3}{m\pi} [(-1)^{m+1} + 1].
 \end{aligned}$$

Pertanto la (3.30) diventa:

$$\sum_{\substack{k=1 \\ k \neq 3}}^{\infty} \frac{3}{2} c_k \delta_{mk} = 2 \frac{3^3}{(m\pi)^2} [(-1)^m - 1] + \frac{192}{m\pi} (-1)^{m+1} + \frac{21}{m\pi},$$

e poiché

$$\sum_{\substack{k=1 \\ k \neq 3}}^{\infty} \frac{3}{2} c_k \delta_{mk} = \frac{3}{2} c_m$$

si ha che i coefficienti dello sviluppo sono interamente determinati

$$c_m = \frac{2}{3} \left[2 \frac{3^3}{(m\pi)^2} [(-1)^m - 1] + \frac{192}{m\pi} (-1)^{m+1} + \frac{21}{m\pi} \right], \quad (1.26)$$

con $m \in \mathbb{N} \setminus \{0, 3\}$.

Possiamo dunque affermare che la soluzione del problema (1.18) è data da

$$u(x, t) = \sum_{m=1}^{\infty} c_m e^{-\lambda_m t} \sin\left(\frac{m\pi}{3}x\right) + 10(x+1),$$

dove $c_3 = 6$ e i c_m con $m \neq 3$ sono dati dalla (3.31).

Esempio 1.3.4

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \\ u(x, 0) = 4e^{-3x}, & u\left(x, \frac{\pi}{2}\right) = 0 \\ u(0, y) = 4\cos 3y, & u\left(\frac{\pi}{2}, y\right) = 4e^{-\frac{3}{2}\pi} \cos 3y \end{cases}, \quad (1.27)$$

dove $(x, y) \in [0, \frac{\pi}{2}] \times [0, \frac{\pi}{2}]$.

Si tratta di un problema ellittico su un quadrato, e visto che i valori della funzione incognita sono noti lungo la frontiera del dominio, sappiamo che il problema ammette una sola soluzione che possiamo ottenere col metodo della separazione delle variabili:

$$u(x, y) = X(x)Y(y), \quad X \text{ e } Y \text{ differenziabili.}$$

Sostituendo e riordinando si ha

$$\begin{cases} \frac{X''}{X} = -\frac{Y''}{Y} = \lambda \\ X(x)Y(0) = 4e^{-3x}, & X(x)Y\left(\frac{\pi}{2}\right) = 0 \\ X(0)Y(y) = 4\cos 3y, & X\left(\frac{\pi}{2}\right)Y(y) = 4e^{-\frac{3}{2}\pi} \cos 3y \end{cases}. \quad (1.28)$$

Si hanno quindi tre casi a seconda del segno del parametro λ (autovalore).

1. $\lambda > 0$

$$\begin{cases} X''(x) - \lambda X(x) = 0 \\ Y''(y) + \lambda Y(y) = 0 \end{cases} \Rightarrow \begin{cases} X(x) = c_1 e^{\sqrt{\lambda}x} + c_2 e^{-\sqrt{\lambda}x} \\ Y(y) = d_1 \cos(\sqrt{\lambda}y) + d_2 \sin(\sqrt{\lambda}y) \end{cases}.$$

Le condizioni in (3.32) implicano

$$X(x) = 4e^{-3x} \Rightarrow \begin{cases} c_1 = 0 \\ c_2 = 4 \\ \sqrt{\lambda} = 3 \end{cases}$$

e

$$\begin{cases} Y(0) = 1 \\ Y(\pi/2) = 0 \end{cases} \Rightarrow \begin{cases} d_1 = 1 \\ \cos\left(\frac{3}{2}\pi\right) + d_2 \sin\left(\frac{3}{2}\pi\right) = 0 \Rightarrow d_2 = 0 \end{cases}$$

Quindi in questo caso la soluzione del problema risulta essere

$$u(x, y) = 4 e^{-3x} \cos 3y,$$

dato che essa soddisfa tutte le condizioni (1.29).

Poiché la soluzione è unica risulta del tutto inutile analizzare i casi $\lambda = 0$ e $\lambda < 0$.

Esempio 1.3.5

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = 4 \frac{\partial^2 u}{\partial x^2} \\ u(0, t) = u(5, t) = 0 \\ u(x, 0) = 0, \quad u_t(x, 0) = 2 \sin \pi x - \sin 4\pi x, \end{cases} \quad (1.29)$$

dove $x \in [0, 5]$ e $t > 0$.

Si tratta di una problema iperbolico in una dimensione spaziale (equazione delle onde descrivente una corda vibrante) con condizioni iniziali e alla frontiera che garantiscono l'unicità della soluzione, che cerchiamo col metodo di separazione delle variabili:

$$u(x, t) = X(x) T(t), \quad X \text{ e } T \text{ differenziabili.}$$

Sostituendo e riordinando si ha:

$$X T'' = 4 X'' T \Rightarrow \frac{T''}{4T} = \frac{X''}{X} = -\lambda \Rightarrow \begin{cases} X''(x) + \lambda X(x) = 0 \\ T''(t) + 4\lambda T(t) = 0 \end{cases}$$

Applicando le condizioni di Dirichlet agli estremi della corda $u(0, t) = u(5, t) = 0$, si trova il problema di Sturm-Liouville per la $X(x)$:

$$\begin{cases} X'' + \lambda X = 0 \\ X(0) = 0, \quad X(5) = 0 \end{cases}, \quad (1.30)$$

con $x \in [0, 5]$.

Il problema (3.33) ammette diversi casi a seconda del segno dell'autovalore λ ma, in analogia a ciò visto negli esempi precedenti, sappiamo che l'unico caso accettabile è quello con soluzioni oscillanti ($\lambda > 0$):

$$X(x) = a \cos(\sqrt{\lambda}x) + b \sin(\sqrt{\lambda}x), \quad \begin{cases} X(0) = 0 \\ X(5) = 0 \end{cases} \Rightarrow \begin{cases} \lambda_k = \left(\frac{k\pi}{5}\right)^2 \\ X_k(x) = \sin\left(\frac{k\pi}{5}x\right) \end{cases},$$

con $k \in \mathbb{N}$.

Di conseguenza il problema differenziale per la $T(t)$ diventa:

$$\begin{cases} T_k''(t) + 4\lambda_k T_k(t) = 0 \\ T_k(0) = 0 \end{cases}, \quad (1.31)$$

che ha come soluzione

$$T_k(t) = \sin\left(\frac{2k\pi}{5}t\right), \quad k \in \mathbb{N} \setminus \{0\}.$$

Pertanto

$$u_k(x, t) = \sin\left(\frac{k\pi}{5}x\right) \sin\left(\frac{2k\pi}{5}t\right),$$

e

$$u(x, t) = \sum_1^{\infty} a_k \sin\left(\frac{k\pi}{5}x\right) \sin\left(\frac{2k\pi}{5}t\right). \quad (1.32)$$

Possiamo ora applicare la condizione $u_t(x, 0) = 2 \sin \pi x - \sin 4\pi x$ sfruttando il fatto che la serie in (3.36) è uniformemente convergente rispetto alla variabile t e quindi è possibile permutare l'operatore di serie con quello di derivata parziale rispetto al tempo:

$$u_t(x, t) = \sum_1^{\infty} a_k \frac{2k\pi}{5} \sin\left(\frac{k\pi}{5}x\right) \cos\left(\frac{2k\pi}{5}t\right),$$

e quindi si ha

$$\sum_1^{\infty} a_k \frac{2k\pi}{5} \sin\left(\frac{k\pi}{5}x\right) = 2 \sin \pi x - \sin 4\pi x. \quad (1.33)$$

Dalla (3.37) si capisce chiaramente che soltanto i due addendi della serie sono diversi da zero i cui coefficienti a_5 e a_{20} sono:

$$\begin{cases} a_5 \frac{10\pi}{5} = 2 \\ a_{20} \frac{40\pi}{5} = -1 \end{cases} \Rightarrow \begin{cases} a_5 = 1/\pi \\ a_{20} = -1/(8\pi) \end{cases}.$$

La soluzione del problema (1.29) risulta dunque

$$u(x, t) = \frac{1}{\pi} \sin(\pi x) \sin(2\pi t) - \frac{1}{8\pi} \sin(4\pi x) \sin(8\pi t). \quad (1.34)$$

Facendo uso delle formule di prostaferesi la (3.38) può venir scritta nel seguente modo

$$u(x, t) = \frac{\cos \pi(x - 2t) - \cos \pi(x + 2t)}{2\pi} - \frac{\cos 4\pi(x - 2t) - \cos 4\pi(x + 2t)}{16\pi},$$

dove viene messo l'accento sul fatto che nella soluzione della equazione delle onde, le variabili spazio e tempo compaiono sempre come $x \pm ct$ essendo c la velocità di propagazione dell'onda nel mezzo considerato.

Esempio 1.3.6

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \\ u(0, y) = 0 \\ u(0, L) = f(y) \\ u(x, 0) = u(M, 0) = 0 \end{cases} \quad (x, y) \in [0, M] \times [0, L]. \quad (1.35)$$

Si tratta di un problema ellittico definito su un rettangolo. La funzione incognita è nota lungo la frontiera del dominio e pertanto sappiamo che il problema ammette una sola soluzione che cerchiamo col metodo di separazione delle variabili.

$$u(x, y) = X(x)Y(y) \Rightarrow \underbrace{\frac{X''(x)}{X(x)}}_{=\lambda} + \underbrace{\frac{Y''(y)}{Y(y)}}_{=-\lambda} = 0 \Rightarrow \begin{cases} X''(x) - \lambda X(x) = 0 \\ Y''(y) + \lambda Y(y) = 0 \end{cases}.$$

Applicando le condizioni sulla frontiera è facile vedere che per la $Y(y)$ si ottiene un problema di Sturm-Liouville completo, mentre per la $X(x)$ se ne ottiene uno con una sola condizione iniziale:

$$\begin{cases} X'' - \lambda X = 0 \\ X(0) = 0 \end{cases} \quad \begin{cases} Y'' + \lambda Y = 0 \\ Y(0) = Y(M) = 0 \end{cases}.$$

Per trovare autovalori e autofunzioni analizziamo il problema differenziale ordinario completo e in analogia agli esempi già visti, otteniamo la seguente soluzione (spettro):

$$\begin{cases} \lambda_k = \left(\frac{k\pi}{M}\right)^2 \\ Y_k(y) = \sin\left(\frac{k\pi}{M}y\right) \end{cases} \quad k \in \mathbb{N}. \quad (1.36)$$

Facendo uso della (3.39) il problema differenziale per la $X(x)$ diventa

$$\begin{cases} X_k(x)'' - \lambda_k X(x) = 0 \\ X_k(0) = 0 \end{cases}, \quad (1.37)$$

che ammette la seguente soluzione ($\lambda_k > 0 \forall k \in \mathbb{N} \setminus \{0\}$)

$$X_k(x) = ae^{\sqrt{\lambda_k}x} + be^{-\sqrt{\lambda_k}x}, \quad X_k(0) = 0 \Rightarrow a + b = 0,$$

e pertanto la soluzione del (3.40) è (a meno di una costante moltiplicativa)

$$X_k(x) = e^{\sqrt{\lambda_k}x} - e^{-\sqrt{\lambda_k}x} = 2 \sinh(\sqrt{\lambda_k}x). \quad (1.38)$$

Facendo uso delle (3.39) e (3.41) la soluzione del problema differenziale a derivate parziali è dato da una serie di funzioni ortogonali:

$$u(x, y) = \sum_1^{\infty} a_k \sinh\left(\frac{k\pi}{M}x\right) \sin\left(\frac{k\pi}{M}y\right). \quad (1.39)$$

Dalla condizione $u(0, L) = f(y)$ si ottiene

$$\sum_{n=1}^{\infty} a_n \sinh \frac{n\pi L}{M} \sin \frac{n\pi}{M}y = f(y),$$

da cui segue che i coefficienti di Fourier a_n sono dati da

$$a_n = \frac{\frac{2}{M} \int_0^M f(x) \sin \frac{n\pi}{M}y dy}{\sinh \frac{\pi L}{M}},$$

e pertanto la soluzione del problema (1.35) è completamente determinata.

Esempio 1.3.7 (Equazione del calore con estremi isolati)

$$\begin{cases} \frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, & 0 \leq x \leq L, \quad t > 0 \\ u(0, t) = u(L, t) = 0 \\ u(x, 0) = f(x) \end{cases}, \quad (1.40)$$

dove a è una costante reale e positiva.

Il problema (1.40) è un modello parabolico descrivente una sbarra monodimensionale di lunghezza L mantenuta a zero gradi agli estremi e avente

una distribuzione di temperatura iniziale data da $f(x) \in C^1([0, L])$. Dalla teoria sappiamo che il problema (1.40) ammette una e una sola soluzione che cerchiamo col metodo di separazione della variabili.

$$u(x, t) = X(x)T(t) \Rightarrow \underbrace{\frac{T'(t)}{a^2 T(t)}}_{=-\lambda} = \underbrace{\frac{X''(x)}{X(x)}}_{=-\lambda}.$$

Dalle condizioni agli estremi della sbarra ($u(0, t) = u(L, t) = 0$) seguono due condizioni di Dirichlet

$$\begin{cases} X''(x) + \lambda X(x) = 0 \\ X(0) = 0 \quad X(L) = 0 \end{cases}.$$

Questo problema spettrale ammette i seguenti autovalori e autofunzioni:

$$\begin{cases} \lambda_k = \left(\frac{k\pi}{L}\right)^2 \\ X_k(x) = \sin\left(\frac{k\pi}{L}x\right) \end{cases} \quad k \in \mathbb{N}. \quad (1.41)$$

Facendo uso della (1.41), otteniamo la soluzione della EDO per $T_k(t)$:

$$T_k(t) = e^{-\frac{k^2\pi^2 a^2}{L^2}t},$$

e quindi la soluzione $u(x, t)$ sarà

$$u(x, t) = \sum_{k=1}^{\infty} c_k e^{-\frac{k^2\pi^2 a^2}{L^2}t} \sin \frac{k\pi}{L}x.$$

Dalla condizione $u(x, 0) = f(x)$ si ricava

$$c_k = \frac{2}{L} \int_0^L f(x) \sin \frac{k\pi}{L}x \, dx.$$

Esempio 1.3.8 (Equazione del calore con un estremo radiante)

$$\begin{cases} \frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, & 0 \leq x \leq L, \quad t > 0 \\ u(0, t) = 0 \\ u_x(L, t) = -A u(L, t) \\ u(x, 0) = f(x) \end{cases}, \quad (1.42)$$

con a e A costanti reali positive.

Il problema (1.42) modella una sbarra di lunghezza L con una distribuzione di temperatura iniziale data da $f(x) \in C^1([0, L])$, con un estremo mantenuto a zero gradi e con l'altro estremo che irradia calore.

Dalla teoria sappiamo che il problema (1.42) ammette una sola soluzione che cerchiamo separando le variabili.

$$u(x, t) = X(x)T(t) \Rightarrow \frac{T'}{a^2 T} = \frac{X''}{X} = -\lambda.$$

Le condizioni agli estremi della sbarra danno luogo alle seguenti condizioni per la $X(x)$:

$$\begin{cases} u(0, t) = 0 \\ u_x(L, t) = -A u(L, t) \end{cases} \Rightarrow \begin{cases} X(0)T(t) = 0 \\ X'(L)T(t) = -AX(L)T(t) \end{cases} \Rightarrow \begin{cases} X(0) = 0 \\ X'(L) = -AX(L) \end{cases},$$

che generano il seguente problema di Sturm-Liouville con condizioni miste:

$$\begin{cases} X'' + \lambda X = 0 \\ X(0) = 0 \\ X'(L) + AX(L) = 0 \end{cases} \quad (\text{sistema di Sturm-Liouville}). \quad (1.43)$$

Avendo condizioni al contorno miste, il problema (3.44) va discusso al variare del segno dell'autovalore λ .

- $\lambda < 0$

$$\begin{aligned} X(x) &= ae^{\sqrt{|\lambda|x}} + be^{-\sqrt{|\lambda|x}}, \\ X'(x) &= a\sqrt{|\lambda|}e^{\sqrt{|\lambda|x}} - b\sqrt{|\lambda|}e^{-\sqrt{|\lambda|x}}. \end{aligned}$$

Le condizioni al contorno implicano

$$\begin{cases} X(0) = 0 \\ X'(L) + AX(L) = 0 \end{cases} \Rightarrow \begin{cases} a + b = 0 \\ a\sqrt{|\lambda|}e^{\sqrt{|\lambda|L}} - b\sqrt{|\lambda|}e^{-\sqrt{|\lambda|L}} + A(ae^{\sqrt{|\lambda|L}} + be^{-\sqrt{|\lambda|L}}) = 0 \end{cases}$$

Occorre capire se il seguente sistema omogeneo ammette soluzioni non banali:

$$\begin{pmatrix} 1 & 1 \\ (\sqrt{|\lambda|} + A)e^{\sqrt{|\lambda|L}} & (-\sqrt{|\lambda|} + A)e^{-\sqrt{|\lambda|L}} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (1.44)$$

Si trova facilmente che il determinante della matrice dei coefficienti della (3.45) è sempre diverso da zero $\forall \lambda < 0$ perché l'equazione trascendente

$$\tanh(\sqrt{|\lambda|}L) = -\frac{\sqrt{|\lambda|}}{A}$$

è soddisfatta solo per $\lambda = 0$. Di conseguenza $a = b = 0$ e questo caso è da scartare.

- $\lambda = 0$

$$\begin{aligned} X(x) &= ax + b, \\ X'(x) &= a. \end{aligned}$$

Le condizioni al contorno implicano

$$\begin{cases} X(0) = 0 \\ X'(L) + AX(L) = 0 \end{cases} \Rightarrow \begin{cases} b = 0 \\ a + A(aL + b) = 0 \end{cases} ,$$

cioè $a = b = 0$ e quindi anche questo caso è da scartare.

- $\lambda > 0$

$$\begin{aligned} X(x) &= a \cos(\sqrt{\lambda}x) + b \sin(\sqrt{\lambda}x), \\ X'(x) &= -a\sqrt{\lambda} \sin(\sqrt{\lambda}x) + b\sqrt{\lambda} \cos(\sqrt{\lambda}x). \end{aligned}$$

Le condizioni al contorno implicano

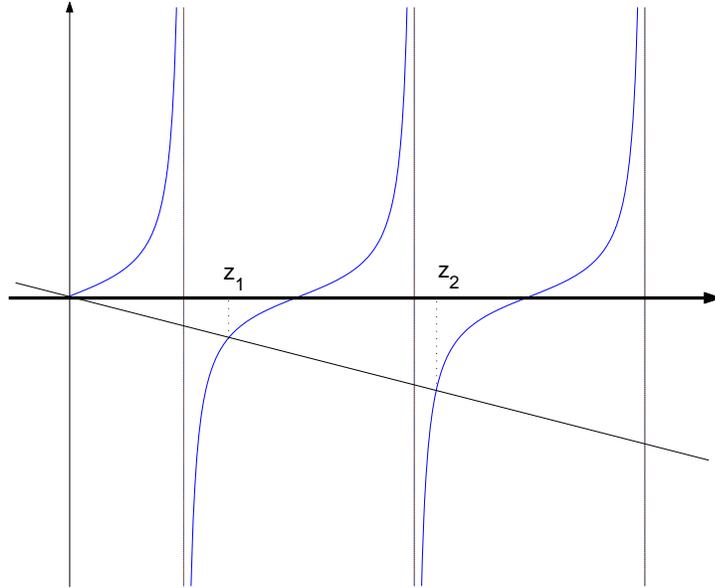
$$\begin{aligned} &\begin{cases} X(0) = 0 \\ X'(L) + AX(L) = 0 \end{cases} \Rightarrow \\ \Rightarrow &\begin{cases} a = 0 \\ -a\sqrt{\lambda} \sin(\sqrt{\lambda}L) + b\sqrt{\lambda} \cos(\sqrt{\lambda}L) + A(a \cos(\sqrt{\lambda}L) + b \sin(\sqrt{\lambda}L)) = 0 \end{cases} , \end{aligned}$$

cioè

$$\tan z = -\frac{z}{AL}, \quad \text{con } z = \sqrt{\lambda}L. \quad (1.45)$$

La (3.46) è una equazione trascendente che possiede una infinità numerabile di soluzioni (Fig. 1.3.8 e Rel. 3.48).

$$(2n - 1)\frac{\pi}{2} \leq z_n \leq n\pi, \quad n = 1, 2, \dots \quad (1.46)$$



Pertanto il problema di Sturm-Liouville (3.44) ammette le seguenti autofunzioni e autovalori

$$\begin{cases} X_n(x) = \sin\left(\frac{z_n}{L}x\right) \\ (2n-1)\frac{\pi}{2} \leq z_n \leq n\pi, \quad n \in \mathbb{N} \setminus \{0\}, \end{cases} \quad (1.47)$$

con $z_n = \sqrt{\lambda_n}L$.

Facendo uso della (3.49) è possibile risolvere l'EDO del primo ordine per la $T(t)$. Si trova facilmente che la soluzione è

$$T_n(t) = e^{-\frac{z_n^2 a^2}{L^2}t},$$

e quindi

$$u(x, t) = \sum_{n=1}^{\infty} c_n e^{-\frac{z_n^2 a^2}{L^2}t} \sin \frac{z_n}{L}x.$$

Dalla condizione $u(x, 0) = f(x)$ si ottiene

$$\sum_{n=1}^{\infty} c_n \sin \frac{z_n}{L}x = f(x),$$

da cui, sfruttando l'ortogonalità delle autofunzioni $\sin \frac{zn}{L}x$, si ha che

$$c_n = \frac{\int_0^L f(x) \sin \frac{zn}{L}x dx}{\int_0^L \sin^2 \frac{zn}{L}x dx}.$$

Esempio 1.3.9

$$\begin{cases} u_t = c^2 u_{xx} + 2u_x, & \text{con } 0 < x < 1 \text{ e } t > 0 \\ u(1, t) = 0 \\ u(0, t) + 4u_x(0, t) = 0 \\ u(x, 0) = f(x), \end{cases}$$

dove c è una costante reale positiva.

Si cercano soluzioni del tipo $u(x, t) = X(x)T(t)$. Sostituendo nell'equazione si trova

$$X(x)T'(t) = c^2 X''(x)T(t) + 2X'(x)T(t)$$

dalle quali si ricava

$$\frac{T'(t)}{T(t)} = c^2 \frac{X''(x)}{X(x)} + 2 \frac{X'(x)}{X(x)} = -\lambda.$$

L'ultima equazione, tenuto conto delle prime due condizioni iniziali del problema e trascurando per comodità l'argomento x , dà origine al seguente *problema spettrale* con condizioni al contorno miste:

$$\begin{cases} c^2 X'' + 2X' + \lambda X = 0 \\ X(1) = 0 \\ X(0) + 4X'(0) = 0. \end{cases}$$

Osserviamo che altre informazioni saranno tratte dall'equazione

$$T'(t) + \lambda T(t) = 0.$$

Soffermiamoci ora sul problema spettrale. Dividendo primo e secondo membro per c^2 , si ha

$$X'' + \frac{2}{c^2} X' + \frac{\lambda}{c^2} X = 0,$$

per la quale si cercano soluzioni del tipo $X(x) \cong e^{\alpha x}$. Semplici calcoli mostrano come α debba soddisfare la seguente *equazione caratteristica*

$$\alpha^2 + \frac{2}{c^2} \alpha + \frac{\lambda}{c^2} = 0,$$

da cui si ricava $\alpha_{1,2} = \frac{-1 \pm \sqrt{1 - \lambda c^2}}{c^2}$. Occorre quindi distinguere tre casi in base al segno del discriminante.

a. Sia $1 - \lambda c^2 > 0$.

In tal caso l'equazione caratteristica possiede due radici reali e distinte

$\alpha_{1,2} = \frac{-1 \pm \sqrt{1 - \lambda c^2}}{c^2}$ e quindi la soluzione generale del problema spettrale è del tipo

$$X(x) = e^{-\frac{1}{c^2}x} \left(a e^{\frac{\sqrt{1-\lambda c^2}}{c^2}x} + b e^{-\frac{\sqrt{1-\lambda c^2}}{c^2}x} \right),$$

con a e b costanti arbitrarie da determinare mediante le condizioni iniziali.

La condizione $X(1) = 0$ comporta $a e^{\frac{\sqrt{1-\lambda c^2}}{c^2}} + b e^{-\frac{\sqrt{1-\lambda c^2}}{c^2}} = 0$, mentre dalla condizione $X(0) + 4X'(0) = 0$ si ricava $(c^2 - 4 + 4\sqrt{1 - \lambda c^2}) a + (c^2 - 4 - 4\sqrt{1 - \lambda c^2}) b = 0$. Occorre quindi risolvere il seguente sistema omogeneo

$$\begin{cases} e^{\frac{\sqrt{1-\lambda c^2}}{c^2}} a + e^{-\frac{\sqrt{1-\lambda c^2}}{c^2}} b = 0 \\ (c^2 - 4 + 4\sqrt{1 - \lambda c^2}) a + (c^2 - 4 - 4\sqrt{1 - \lambda c^2}) b = 0. \end{cases}$$

Esso ammette l'unica soluzione $a = b = 0$ in quanto la matrice dei coefficienti è non singolare. Pertanto questo caso è da scartare perché conduce alla soluzione banale $X(x) = 0$. Infatti una tale soluzione implicherebbe $u(x, t) = X(x)T(t) = 0$ per ogni (x, t) appartenenti al dominio e questa soluzione (banale) non soddisfa la condizione $u(x, 0) = f(x)$ data nel problema iniziale.

b. Sia $1 - \lambda c^2 = 0$.

In tal caso l'equazione caratteristica possiede l'unica soluzione $\alpha_1 =$

$\alpha_2 = \alpha = -\frac{1}{c^2}$ e la soluzione generale della equazione differenziale ordinaria (EDO) è del tipo

$$X(x) = (a + bx) e^{-\frac{1}{c^2}x},$$

con a e b costanti reali arbitrarie da determinare attraverso le condizioni iniziali.

La condizione $X(1) = 0$ implica $a + b = 0$, mentre la condizione $X(0) + 4X'(0) = 0$ comporta $(1 - \frac{4}{c^2}) a + 4b = 0$. Bisogna quindi risolvere il seguente sistema di 2 equazioni lineari omogenee

$$\begin{cases} a + b = 0 \\ (1 - \frac{4}{c^2}) a + 4b = 0. \end{cases}$$

Anche in questo caso si verifica subito che il sistema possiede come unica soluzione quella banale $a = b = 0$, in quanto il determinante della matrice dei coefficienti è sempre diverso da zero. Quindi, come nel caso precedente, la condizione $1 - \lambda c^2 = 0$ è da scartare.

c. Sia $1 - \lambda c^2 < 0$.

In tal caso l'equazione caratteristica possiede le radici complesse coniugate $\alpha_{1,2} = \frac{-1 \pm i\sqrt{\lambda c^2 - 1}}{c^2}$ e la soluzione generale della ODE è

$$X(x) = e^{-\frac{1}{c^2}x} \left(a \cos \frac{\sqrt{\lambda c^2 - 1}}{c^2} x + b \sin \frac{\sqrt{\lambda c^2 - 1}}{c^2} x \right),$$

con a e b costanti reali arbitrarie da determinare imponendo le condizioni iniziali.

La condizione $X(1) = 0$ implica $a \cos \frac{\sqrt{\lambda c^2 - 1}}{c^2} + b \sin \frac{\sqrt{\lambda c^2 - 1}}{c^2} = 0$, mentre la condizione $X(0) + 4X'(0) = 0$ comporta $a + 4 \left(-\frac{a}{c^2} + \frac{\sqrt{\lambda c^2 - 1}}{c^2} b \right) = 0$. Al fine di determinare a e b occorre quindi risolvere il seguente sistema omogeneo

$$\begin{cases} \cos \frac{\sqrt{\lambda c^2 - 1}}{c^2} a + \sin \frac{\sqrt{\lambda c^2 - 1}}{c^2} b = 0 \\ (1 - \frac{4}{c^2}) a + \frac{4}{c^2} \sqrt{\lambda c^2 - 1} b = 0. \end{cases}$$

Per risolvere questo sistema distinguiamo due casi:

1. Se $c^2 = 4$, la seconda equazione del sistema è soddisfatta per a arbitrario e $b = 0$. La prima equazione del sistema diventa quindi $a \cos \frac{\sqrt{\lambda c^2 - 1}}{c^2} = 0$. Tale equazione è soddisfatta (trascuriamo il caso $a = 0$ che conduce alla soluzione non accettabile $X(x) = 0$) se $\frac{\sqrt{4\lambda - 1}}{4} = \frac{\pi}{2}(2k + 1)$, per $k = 0, 1, 2, \dots$. Quindi la soluzione del problema spettrale è data da

$$\left\{ \beta_k = \frac{\sqrt{4\lambda_k - 1}}{4} = \frac{\pi}{2}(2k + 1); \quad e^{-\frac{x}{4}} \cos \frac{\pi}{2}(2k + 1)x \right\}.$$

Tenendo conto che l'altra equazione $T'_k + \lambda_k T_k = 0$ ammette come soluzione $T_k \cong e^{-\lambda_k t}$, si trova

$$u_k(x, t) = X_k(x) T_k(t) = e^{-\lambda_k t} e^{-\frac{x}{4}} \cos \frac{\pi}{2}(2k + 1)x, \quad k = 0, 1, 2, \dots$$

La soluzione generale si ottiene come combinazione lineare di tutte le (infinite) $u_k(x, t)$, ossia

$$u(x, t) = \sum_{k=0}^{\infty} c_k e^{-(\lambda_k t + \frac{x}{4})} \cos \frac{\pi}{2}(2k + 1)x.$$

Rimane da imporre la condizione $u(x, 0) = f(x)$, da cui si ricava

$$\sum_{k=0}^{\infty} c_k e^{-\frac{x}{4}} \cos \frac{\pi}{2}(2k + 1)x = f(x),$$

e, non dipendendo $e^{-\frac{x}{4}}$ da k , si può scrivere

$$\sum_{k=0}^{\infty} c_k \cos \frac{\pi}{2}(2k + 1)x = f(x) e^{\frac{x}{4}}.$$

Il calcolo dei coefficienti c_k si può agevolmente ottenere mediante il metodo di Fourier.

2. Sia ora $c^2 \neq 4$. Per semplicità prendiamo $c^2 = 1$, ma la trattazione è analoga qualunque sia il valore di $c^2 \neq 4$. Prendendo $c^2 = 1$ la soluzione cercata è del tipo

$$X(x) = e^{-x} \left(a \cos \sqrt{\lambda - 1} x + b \sin \sqrt{\lambda - 1} x \right),$$

con a e b costanti arbitrarie da determinare dal seguente sistema:

$$\begin{cases} a \cos \sqrt{\lambda - 1} + b \sin \sqrt{\lambda - 1} = 0 \\ -3a + 4 \sqrt{\lambda - 1} b = 0. \end{cases}$$

Ovviamente, siamo interessati al caso in cui tale sistema possiede soluzioni non banali. Dalla seconda equazione si trova $a = \frac{4}{3}\sqrt{\lambda - 1} b$ da cui, sostituendo nella prima equazione, si ottiene

$$b \left(\frac{4}{3} \sqrt{\lambda - 1} \cos \sqrt{\lambda - 1} + \sin \sqrt{\lambda - 1} \right) = 0.$$

Non essendo interessati a soluzioni banali, non prendiamo in considerazione il caso $b = 0$ in quanto se b si dovesse annullare anche a sarebbe uguale a 0, quindi occorre risolvere la seguente equazione

$$\frac{4}{3} \sqrt{\lambda - 1} \cos \sqrt{\lambda - 1} + \sin \sqrt{\lambda - 1} = 0.$$

Dividendo per $\cos \sqrt{\lambda - 1} \neq 0$ (cosa questa che si può assumere perché se fosse $\cos \sqrt{\lambda - 1} = 0$ allora dall'equazione di sopra si avrebbe $\sin \sqrt{\lambda - 1} = 0$, ma non può capitare che $\sin x$ e $\cos x$ siano nulle simultaneamente), questa diventa

$$\tan \sqrt{\lambda - 1} = \frac{-4}{3} \sqrt{\lambda - 1}.$$

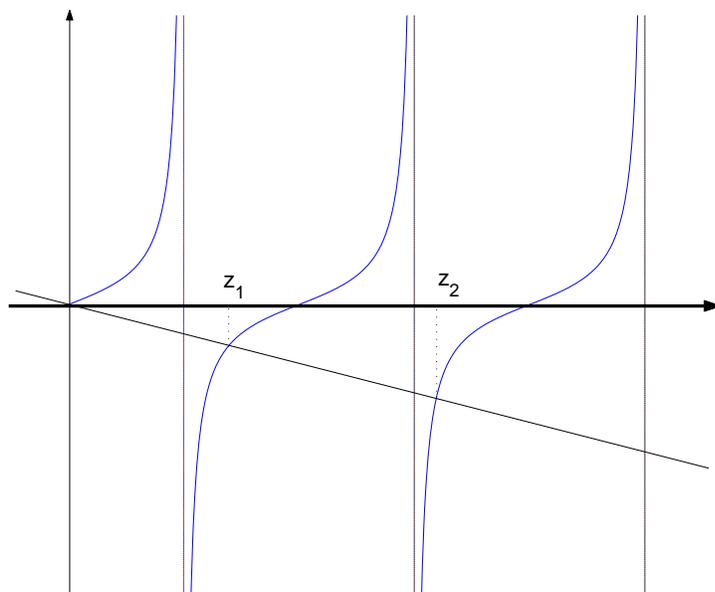
Ponendo $z = \sqrt{\lambda - 1} > 0$, si ottiene $\tan z = -\frac{4}{3} z$. Questa equazione può essere risolta graficamente come indicato nella Fig. 1.3.9.

Dalla 1.3.9 si osserva che le soluzioni dell'equazione sono i

$$z_k : (2k + 1) \pi < z_k < k \pi, \text{ con } k = 1, 2, 3, 4, \dots$$

Quindi autovlori e autofunzioni nel caso $c = 1$ sono dati da

$$\left\{ z_k = \sqrt{\lambda_k - 1}; e^{-x} \left(\frac{4}{3} z_k \cos z_k x + \sin z_k x \right) \right\}.$$



Tenendo conto che l'altra equazione $T_k' + \lambda_k T_k = 0$ ammette come soluzione $T_k = e^{-\lambda_k t}$, si trova

$$u_k(x, t) = X_k(x) T_k(t) = e^{-\lambda_k t} e^{-x} \left(\frac{4}{3} z_k \cos z_k x + \sin z_k x \right),$$

con $k = 0, 1, 2, \dots$. La soluzione generale si ottiene facendo una combinazione lineare di tutte le (infinite) $u_k(x, t)$. Perciò tale soluzione è espressa come:

$$u(x, t) = \sum_{k=1}^{\infty} c_k e^{-(\lambda_k t + x)} \left(\frac{4}{3} z_k \cos z_k x + \sin z_k x \right).$$

Applicando la condizione $u(x, 0) = f(x)$, si ha

$$\sum_{k=1}^{\infty} c_k e^{-x} \left(\frac{4}{3} z_k \cos z_k x + \sin z_k x \right) = f(x),$$

ovvero

$$\sum_{k=1}^{\infty} c_k \left(\frac{4}{3} z_k \cos z_k x + \sin z_k x \right) = f(x) e^x.$$

Si nota che $u(x, 0)$ non è definita da una serie di Fourier, dato che l'argomento di $\sin z_k x$ non è esprimibile nella forma $k\omega x$ al

variare di k . Il calcolo dei coefficienti c_k può comunque essere effettuato sfruttando l'ortogonalità delle autofunzioni (problema di Sturm-Liouville).

Esempio 1.3.10 (Equazione delle onde con velocità iniziale nulla)

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, & 0 \leq x \leq L, \quad t > 0 \\ u(0, t) = u(L, t) = 0 \\ u(x, 0) = f(x) \\ u_t(x, 0) = 0 \end{cases}, \quad (1.48)$$

dove a è una costante reale positiva.

Il problema (1.48) modella una corda vibrante lunga L con estremi fissati, forma iniziale data da $f(x) \in C^1([0, L])$ e con ogni suo punto avente velocità iniziale nulla. Dalla teoria sappiamo che questo problema ammette una e una sola soluzione che cerchiamo col metodo della separazione delle variabili.

La procedura è analoga a quella vista negli esempi precedenti, quindi possiamo procedere in maniera più rapida sorvolando tutta una serie di dettagli già affrontati e discussi precedentemente.

$$u(x, t) = X(x)T(t)$$

$$\begin{cases} X'' + \lambda X = 0 \\ X(0) = X(L) = 0 \end{cases} \Rightarrow \begin{cases} \lambda_n = \frac{n^2 \pi^2}{L^2}, & n = 1, 2, \dots \\ X_n = b_n \sin \frac{n\pi}{L} x \end{cases}$$

$$\begin{cases} T'' + \lambda a^2 T = 0 \\ T(0) = 0 \end{cases} \Rightarrow T_n = a_n \cos \frac{a n \pi}{L} t$$

$$\begin{aligned} u(x, t) &= \sum_{n=1}^{\infty} c_n \sin \frac{n\pi}{L} x \cos \frac{a n \pi}{L} t \\ &= \sum_{n=1}^{\infty} \frac{c_n}{2} \left[\sin \frac{n\pi}{L} (x - at) + \sin \frac{n\pi}{L} (x + at) \right] \end{aligned}$$

Dalla condizione $\sum_{n=1}^{\infty} c_n \sin \frac{n\pi}{L} x = f(x)$ si ricava

$$c_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi}{L} x dx .$$

Esempio 1.3.11 (Equazione delle onde con spostamento iniziale nullo)

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, & 0 \leq x \leq L, \quad t > 0 \quad \text{a costante positiva} \\ u(0, t) = u(L, t) = 0 \\ u(x, 0) = 0 \\ u_t(x, 0) = g(x) \end{cases} \quad (1.49)$$

Si tratta dello stesso problema affrontato nell'Es. (1.48) ma con la differenza che adesso la corda vibrante possiede posizione iniziale nulla ($u(x, 0) = 0$) e ogni suo punto ha una velocità iniziale data da $g(x) \in C^1([0, L])$. In analogia al precedente esempio, si ha:

$$u(x, t) = X(x)T(t)$$

$$\begin{cases} X'' + \lambda X = 0 \\ X(0) = X(L) = 0 \end{cases} \Rightarrow \begin{cases} \lambda_n = \frac{n^2 \pi^2}{L^2}, & n = 1, 2, \dots \\ X_n = b_n \sin \frac{n\pi}{L} x \end{cases}$$

$$\begin{cases} T'' + \lambda a^2 T = 0 \\ T(0) = 0 \end{cases} \Rightarrow T_n = a_n \sin \frac{a n \pi}{L} t$$

$$\begin{aligned} u(x, t) &= \sum_{n=1}^{\infty} c_n \sin \frac{n\pi}{L} x \sin \frac{n\pi}{L} at \\ &= \sum_{n=1}^{\infty} \frac{c_n}{2} \left[\cos \frac{n\pi}{L} (x - at) + \cos \frac{n\pi}{L} (x + at) \right] \end{aligned}$$

Dalla condizione $u_t(x, 0) = \sum_{n=1}^{\infty} c_n \frac{n\pi a}{L} \sin \frac{n\pi}{L} x = g(x)$ si ricava

$$c_n = \frac{2}{n\pi a} \int_0^L g(x) \sin \frac{n\pi}{L} x dx .$$

Esempio 1.3.12 Equazione di Laplace sul cerchio unitario (formula risolutiva di Poisson)

$$\begin{cases} u_{xx} + u_{yy} = 0, & x^2 + y^2 < 1 \\ u(x, y) = f(x, y), & x^2 + y^2 = 1 \end{cases} \quad (1.50)$$

Il problema (1.50) non è risolubile mediante il metodo di separazione delle variabili in quanto il dominio (essendo un cerchio) manca della regolarità necessaria per applicare il metodo. Pertanto l'idea è quella di fare un cambiamento di coordinate al fine di avere (nel nuovo sistema di riferimento) un dominio regolare (quadrato o rettangolo) che ci permetta di separare le variabili nella equazione differenziale trasformata.

Vista la simmetria polare del dominio, trasformiamo il problema da coordinate cartesiane a coordinate polari.

Si ha:

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \end{cases} \Rightarrow \begin{cases} \rho = \sqrt{x^2 + y^2} \\ \theta = \arctan \frac{y}{x} \end{cases}$$

da cui, applicando le regole di derivazione delle funzioni composte,

$$u_x = \frac{\partial u}{\partial \rho} \frac{\partial \rho}{\partial x} + \frac{\partial u}{\partial \theta} \frac{\partial \theta}{\partial x} = \frac{x}{\rho} \frac{\partial u}{\partial \rho} - \frac{y}{x^2 + y^2} \frac{\partial u}{\partial \theta} = \cos \theta \frac{\partial u}{\partial \rho} - \frac{\sin \theta}{\rho} \frac{\partial u}{\partial \theta}$$

$$u_y = \frac{\partial u}{\partial \rho} \frac{\partial \rho}{\partial y} + \frac{\partial u}{\partial \theta} \frac{\partial \theta}{\partial y} = \frac{y}{\rho} \frac{\partial u}{\partial \rho} + \frac{x}{x^2 + y^2} \frac{\partial u}{\partial \theta} = \sin \theta \frac{\partial u}{\partial \rho} + \frac{\cos \theta}{\rho} \frac{\partial u}{\partial \theta}$$

$$\begin{aligned} u_{xx} &= \left[\frac{\partial}{\partial \rho} \left(\cos \theta \frac{\partial u}{\partial \rho} - \frac{\sin \theta}{\rho} \frac{\partial u}{\partial \theta} \right) \right] \frac{\partial \rho}{\partial x} + \left[\frac{\partial}{\partial \theta} \left(\cos \theta \frac{\partial u}{\partial \rho} - \frac{\sin \theta}{\rho} \frac{\partial u}{\partial \theta} \right) \right] \frac{\partial \theta}{\partial x} = \\ &= \left(\cos \theta \frac{\partial^2 u}{\partial \rho^2} + \frac{\sin \theta}{\rho^2} \frac{\partial u}{\partial \theta} - \frac{\sin \theta}{\rho} \frac{\partial^2 u}{\partial \rho \partial \theta} \right) \cos \theta + \\ &\quad + \left(-\sin \theta \frac{\partial u}{\partial \rho} + \cos \theta \frac{\partial^2 u}{\partial \rho \partial \theta} - \frac{\cos \theta}{\rho} \frac{\partial u}{\partial \theta} - \frac{\sin \theta}{\rho} \frac{\partial^2 u}{\partial \theta^2} \right) \left(-\frac{\sin \theta}{\rho} \right) = \\ &= \cos^2 \theta \frac{\partial^2 u}{\partial \rho^2} + 2 \frac{\cos \theta \sin \theta}{\rho^2} \frac{\partial u}{\partial \theta} - 2 \frac{\cos \theta \sin \theta}{\rho} \frac{\partial^2 u}{\partial \rho \partial \theta} + \frac{\sin^2 \theta}{\rho} \frac{\partial u}{\partial \rho} + \frac{\sin^2 \theta}{\rho^2} \frac{\partial^2 u}{\partial \theta^2} \end{aligned}$$

$$\begin{aligned}
u_{yy} &= \left[\frac{\partial}{\partial \rho} \left(\sin \theta \frac{\partial u}{\partial \rho} + \frac{\cos \theta}{\rho} \frac{\partial u}{\partial \theta} \right) \right] \frac{\partial \rho}{\partial y} + \left[\frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial u}{\partial \rho} + \frac{\cos \theta}{\rho} \frac{\partial u}{\partial \theta} \right) \right] \frac{\partial \theta}{\partial y} = \\
&= \left(\sin \theta \frac{\partial^2 u}{\partial \rho^2} - \frac{\cos \theta}{\rho^2} \frac{\partial u}{\partial \theta} + \frac{\cos \theta}{\rho} \frac{\partial^2 u}{\partial \rho \partial \theta} \right) \sin \theta + \\
&\quad + \left(\cos \theta \frac{\partial u}{\partial \rho} + \sin \theta \frac{\partial^2 u}{\partial \rho \partial \theta} - \frac{\sin \theta}{\rho} \frac{\partial u}{\partial \theta} + \frac{\cos \theta}{\rho} \frac{\partial^2 u}{\partial \theta^2} \right) \frac{\cos \theta}{\rho} = \\
&= \sin^2 \theta \frac{\partial^2 u}{\partial \rho^2} - 2 \frac{\cos \theta \sin \theta}{\rho^2} \frac{\partial u}{\partial \theta} + 2 \frac{\cos \theta \sin \theta}{\rho} \frac{\partial^2 u}{\partial \rho \partial \theta} + \frac{\cos^2 \theta}{\rho} \frac{\partial u}{\partial \rho} + \frac{\cos^2 \theta}{\rho^2} \frac{\partial^2 u}{\partial \theta^2} \\
u_{xx} + u_{yy} &= \frac{\partial^2 u}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial u}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 u}{\partial \theta^2} = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial u}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 u}{\partial \theta^2} = 0
\end{aligned}$$

Pertanto, tenendo conto delle condizioni di periodicit  della u in $(1, 0)$ e di limitatezza della soluzione nel cerchio, l'equazione di Laplace   ora esprimibile nella forma seguente:

$$\left\{ \begin{array}{l} \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial u}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 u}{\partial \theta^2} = 0 \\ u(1, 2\pi) = u(1, 0) \\ u_\theta(1, 2\pi) = u_\theta(1, 0) \\ u(1, \theta) = f(\theta) \\ \lim_{\rho \rightarrow 0} u(\rho, \theta) \text{ limitato} \end{array} \right. \quad \begin{array}{l} \text{(condizioni di periodicit )} \\ \text{(condizione al contorno)} \end{array}$$

Adesso $\rho \in [0, 1]$ e $\theta \in [0, 2\pi]$, quindi il dominio   rettangolare e ha senso separare le variabili:

$$U(\rho, \theta) = U(\rho)V(\theta)$$

si ottiene

$$U''V + \frac{1}{\rho}U'V + \frac{1}{\rho^2}UV'' = 0 \quad \Leftrightarrow \quad \rho^2U''V + \rho U'V + UV'' = 0$$

$$\frac{\rho^2U'' + \rho U'}{U} = \frac{V''}{V} = \lambda$$

$$\left\{ \begin{array}{l} V'' + \lambda V = 0 \\ V(0) = V(2\pi) \\ V'(0) = V'(2\pi) \end{array} \right. \Rightarrow \lambda_n = n^2, \quad n \geq 0$$

$$V_n(\theta) = a_n \cos n\theta + b_n \sin n\theta$$

$$\begin{cases} \rho^2 u'' + \rho U' - \lambda U = 0 \\ \lim_{\rho \rightarrow 0} U(\rho), \text{ limitato} \end{cases}$$

$$\rho^2 U'' + \rho U' - n^2 U = 0 \quad ^1$$

Posto $\rho = e^t$ e $z(t) = U(\rho(t))$ si ha $z' = U' \rho' = U' \rho$ e $z'' = U'' \rho^2 + U' \rho$, da cui

$$z'' - n^2 z = 0$$

$$z_n = c_n e^{nt} + d_n e^{-nt}$$

$$U_n(\rho) = c_n \rho^n + d_n \rho^{-n}$$

$$\lim_{\rho \rightarrow 0} U(\rho), \text{ limitato} \quad \Rightarrow \quad d_n = 0$$

$$u(\rho, \theta) = \sum_{n=0}^{\infty} \rho^n (a_n \cos n\theta + b_n \sin n\theta)$$

Dalla condizione $u(1, \theta) = f(\theta)$ si ricavano i coefficienti di Fourier

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_0^{2\pi} f(\tau) d\tau \\ a_n &= \frac{1}{\pi} \int_0^{2\pi} f(\tau) \cos n\tau d\tau \\ b_n &= \frac{1}{\pi} \int_0^{2\pi} f(\tau) \sin n\tau d\tau, \end{aligned}$$

¹Caso particolare dell'equazione di Eulero-Cauchy $x^2 y'' + a x y' + b y = 0$ che si risolve con la sostituzione $x = e^t$, $z(t) = y(e^t)$. Si ha $z' = y' x$, $z'' = y'' x^2 + y' x$, $y'' x^2 = z'' - y' x$ da cui $z'' + (a - 1) z + b z = 0$.

da cui

$$\begin{aligned}
 u(\rho, \theta) &= \frac{1}{2\pi} \int_0^{2\pi} \left[1 + 2 \sum_{n=1}^{\infty} \rho^n (\cos n\tau \cos n\theta + \sin n\tau \sin n\theta) \right] f(\tau) d\tau = \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \left[1 + 2 \sum_{n=1}^{\infty} \rho^n \cos n(\theta - \tau) \right] f(\tau) d\tau = \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \tau)} f(\tau) d\tau \quad \text{Formula risolvante di Poisson.}
 \end{aligned}$$

Esempio 1.3.13 (Equazione delle onde sul rettangolo)

$$\begin{cases}
 \frac{\partial^2 u}{\partial t^2} = a^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), & 0 \leq x \leq L, \quad 0 \leq y \leq M, \quad t > 0 \\
 u(x, 0, t) = u(x, M, t) = 0 \\
 u(0, y, t) = u(L, y, t) = 0 \\
 u(x, y, 0) = f(x, y) \\
 u_t(x, y, 0) = 0
 \end{cases} \quad (1.51)$$

Il problema (1.51) rappresenta la propagazione delle onde su una membrana elastica di forma rettangolare (lunghezza dei lati M e L) con i lati mantenuti fermi, con posizione iniziale data dalla superficie regolare $f(x, y) \in C^1([0, L] \times [0, M])$ e con velocità iniziale nulla.

Vista la regolarità del dominio, è possibile applicare il metodo di separazione delle variabili.

$$u(x, y, t) = X(x)Y(y)T(t)$$

2

$$\begin{aligned}
 1 + 2 \sum_{n=1}^{\infty} r^n \cos n\theta &= \operatorname{Re}(1 + r e^{i\theta})(1 + r e^{i\theta} + r^2 e^{2i\theta} + \dots) = \\
 &= \operatorname{Re} \frac{1 + r e^{i\theta}}{1 - r e^{i\theta}} = \frac{1 - r^2}{1 + r^2 - 2r \cos \theta}.
 \end{aligned}$$

$$\begin{cases} \frac{T''}{a^2 T} - \frac{Y''}{Y} = \frac{X''}{X} = -\lambda \\ \frac{T''}{a^2 T} + \lambda = \frac{Y''}{Y} = -\mu \end{cases}$$

$$\begin{cases} X'' + \lambda X = 0 \\ X(0) = X(L) = 0 \end{cases} \Rightarrow \lambda_n = \frac{n^2 \pi^2}{L^2}, \quad n = 1, 2, \dots$$

$$\begin{cases} Y'' + \mu Y = 0 \\ Y(0) = Y(M) = 0 \end{cases} \Rightarrow \mu_n = \frac{m^2 \pi^2}{M^2}, \quad m = 1, 2, \dots$$

$$T'' + (\lambda + \mu) a^2 T = 0$$

$$T''_{n,m} + \left(\frac{n^2}{L^2} + \frac{m^2}{M^2} \right) a^2 T_{n,m} = 0$$

$$T_{n,m}(t) = a_{n,m} \cos \sqrt{\frac{n^2}{L^2} + \frac{m^2}{M^2}} \pi a t + b_{n,m} \sin \sqrt{\frac{n^2}{L^2} + \frac{m^2}{M^2}} \pi a t$$

$$T'_{n,m}(0) = 0 \quad \Rightarrow \quad b_{n,m} = 0$$

$$\begin{cases} X_n(x) = a_n \sin \frac{n\pi}{L} x \\ Y_m(y) = b_m \sin \frac{m\pi}{M} y \\ T_{n,m} = a_{n,m} \cos \sqrt{\frac{n^2}{L^2} + \frac{m^2}{M^2}} \pi a t \end{cases}$$

$$u_{n,m}(x, y, t) = c_{n,m} \sin \frac{n\pi}{L} x \sin \frac{m\pi}{M} y \cos \sqrt{\frac{n^2}{L^2} + \frac{m^2}{M^2}} \pi a t$$

$$u(x, y, t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} c_{n,m} \sin \frac{n\pi}{L} x \sin \frac{m\pi}{M} y \cos \sqrt{\frac{n^2}{L^2} + \frac{m^2}{M^2}} \pi a t$$

$$c_{n,m} = \frac{4}{LM} \int_0^M \int_0^L f(x, y) \sin \frac{n\pi}{L} x \sin \frac{m\pi}{M} y \, dx \, dy.$$

Come già osservato in casi simili, per la linearità del problema la tecnica è analoga se $u(x, y, 0) = 0$ e $u_t(x, y, 0) = g(x, y)$. Se $u(x, y, 0) = f(x, y)$ e $u_t(x, y, 0) = g(x, y)$ si risolvono separatamente i due casi indicati e si assume come $u(x, y)$ la somma delle due soluzioni.

Osservazione 1.3.14 *Come abbiamo visto negli esempi precedenti, l'applicazione del metodo di separazione delle variabili è notevolmente semplificata nel caso in cui in tutte le condizioni al bordo e iniziali, tranne una, siano assegnate funzioni identicamente nulle. Questa tuttavia, almeno per le equazioni differenziali omogenee, non rappresenta una oggettiva difficoltà in quanto il problema iniziale può sempre essere ricondotto alla risoluzione di uno o più problemi che presentano le caratteristiche desiderate. L'affermazione è una diretta conseguenza della seguente proprietà, immediatamente verificabile.*

Proposizione 1.3.15 *Se $u_1(x, y)$ e $u_2(x, y)$ sono soluzioni di un'equazione omogenea, lo è anche una qualunque loro combinazione lineare.*

Esempio Consideriamo l'equazione

$$u_{tt} = A(x, t) u_{xx} + B(x, t) u_x + C(x, t) u_t + D(x, t) u.$$

Se $u_1(x, t)$ e $u_2(x, t)$ sono soluzioni, lo è anche $u(x, t) = c_1 u_1(x, t) + c_2 u_2(x, t)$, qualunque siano le costanti c_1 e c_2 .

Dimostrazione. Per semplicità ci riferiamo all'equazione presentata nell'esempio. Si ha

$$c_1 (u_1)_{tt} + c_2 (u_2)_{tt} = A(x, t) [c_1 (u_1)_{xx} + c_2 (u_2)_{xx}] + B(x, t) [c_1 (u_1)_x + c_2 (u_2)_x] + C(x, t) [c_1 (u_1)_t + c_2 (u_2)_t] + D(x, t) [c_1 u_1 + c_2 u_2]$$

in quanto, per ipotesi,

$$c_1 (u_1)_{tt} = c_1 [A(x, t) (u_1)_{xx} + B(x, t) (u_1)_x + C(x, t) (u_1)_t + D(x, t) u_1] \text{ e} \\ c_2 (u_2)_{tt} = c_2 [A(x, t) (u_2)_{xx} + B(x, t) (u_2)_x + C(x, t) (u_2)_t + D(x, t) u_2].$$

■

Supponiamo ora di dover risolvere, con il metodo di separazione delle variabili, il seguente problema differenziale a coefficienti costanti:

$$\begin{cases} u_{tt} = A u_{xx} + B u_x + C u_t + D u, & \text{con } a < x < b \text{ e } t > 0 \\ u(a, t) = f_1(t), & u(b, t) = f_2(t) \\ u(x, 0) = g_1(x), & u_t(x, 0) = g_2(x). \end{cases}$$

Supponiamo di saper risolvere la stessa equazione differenziale in ciascuna delle seguenti situazioni

$$(1) f_1(t) \neq 0, f_2(t) \equiv 0, g_1(x) \equiv 0, g_2(x) \equiv 0;$$

$$(2) f_1(t) \equiv 0, f_2(t) \neq 0, g_1(x) \equiv 0, g_2(x) \equiv 0;$$

$$(3) f_1(t) \equiv 0, f_2(t) \equiv 0, g_1(x) \neq 0, g_2(x) \equiv 0;$$

$$(4) f_1(t) \equiv 0, f_2(t) \equiv 0, g_1(x) \equiv 0, g_2(x) \neq 0.$$

Indichiamo con $u_i(x, t)$, $i = 1, 2, 3, 4$ la soluzione dell'equazione differenziale con assegnate le condizioni (1), (2), (3), (4) rispettivamente. E' allora immediato verificare che

$$u(x, t) = u_1(x, t) + u_2(x, t) + u_3(x, t) + u_4(x, t).$$

A tale scopo basta osservare che:

- (a) in conseguenza della linearità dell'equazione differenziale, $u(x, t)$ è una soluzione;
- (b) $u(x, t)$ soddisfa sia le condizioni al bordo sia quelle iniziali, dato che su ciascuna di esse tre soluzioni sono nulle e la quarta assume i valori desiderati.

1.4 Temi proposti in prove d'esame

a) **Sviluppare la funzione $f(x) = e^{-\sqrt{2}x}$ mediante le autofunzioni del seguente problema di Sturm-Liouville:**

$$\begin{cases} \varphi'' + 2\varphi' + (2 + \lambda)\varphi = 0, & \text{con } 0 < x < \frac{\pi}{2} \\ \varphi'(0) + 2\varphi(0) = 0 \\ \varphi'(\frac{\pi}{2}) = 0. \end{cases}$$

Si cercano soluzioni del tipo $\varphi(x) \simeq e^{\alpha x}$ che conducono all'equazione caratteristica

$$\alpha^2 + 2\alpha + (2 + \lambda) = 0.$$

Questa equazione ha come soluzioni $\alpha_{1,2} = -1 \pm \sqrt{-(\lambda + 1)}$. Occorre distinguere tre casi a seconda del segno del discriminante.

1. Sia $\lambda < -1$. In tal caso abbiamo due radici reali e distinte $\alpha_{1,2} = -1 \pm \sqrt{-(\lambda+1)}$.

La soluzione generale è $\varphi(x) = e^{-x} \left(c_1 e^{-\sqrt{-(\lambda+1)}x} + c_2 e^{\sqrt{-(\lambda+1)}x} \right)$.

Poiché:

$$\begin{aligned}\varphi(0) &= c_1 + c_2, \\ \varphi'(0) &= -(c_1 + c_2) + \left(-c_1 \sqrt{-(\lambda+1)} + c_2 \sqrt{-(\lambda+1)} \right) \\ \varphi'\left(\frac{\pi}{2}\right) &= e^{-\frac{\pi}{2}} \left(c_1(-1 - \sqrt{-(\lambda+1)})e^{-\sqrt{-(\lambda+1)}\frac{\pi}{2}} \right. \\ &\quad \left. + c_2(-1 + \sqrt{-(\lambda+1)})e^{\sqrt{-(\lambda+1)}\frac{\pi}{2}} \right),\end{aligned}$$

le costanti c_1 e c_2 devono soddisfare il sistema di equazioni lineari (rispetto a c_1 e c_2) e omogenee

$$\begin{cases} \left(1 - \sqrt{-(\lambda+1)}\right) c_1 + \left(1 + \sqrt{-(\lambda+1)}\right) c_2 = 0 \\ \left(-1 - \sqrt{-(\lambda+1)}\right) e^{-\sqrt{-(\lambda+1)}\frac{\pi}{2}} c_1 + \left(-1 + \sqrt{-(\lambda+1)}\right) e^{\sqrt{-(\lambda+1)}\frac{\pi}{2}} c_2 = 0 \end{cases}$$

Affinché c_1 e c_2 non siano simultaneamente nulli il determinante di questo sistema

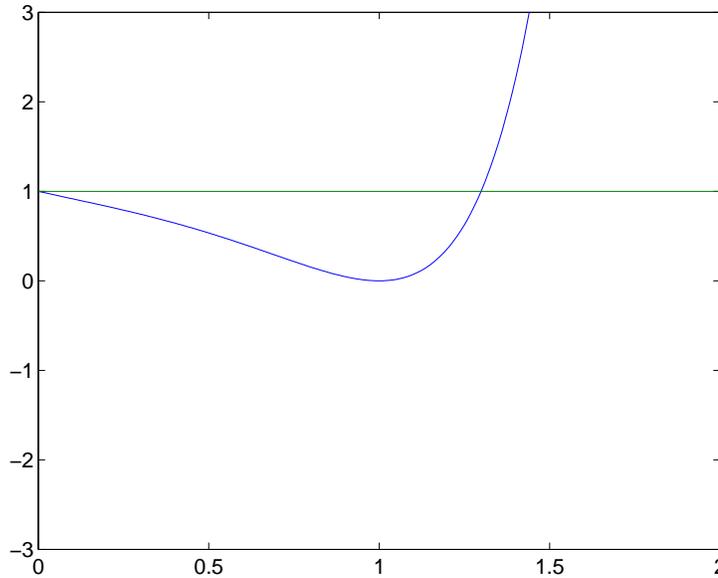
$$-(1 - \sqrt{-(\lambda+1)})^2 e^{\sqrt{-(\lambda+1)}\frac{\pi}{2}} + (1 + \sqrt{-(\lambda+1)})^2 e^{-\sqrt{-(\lambda+1)}\frac{\pi}{2}}$$

deve risultare nullo. Posto $b = \sqrt{-(\lambda+1)} > 0$, si può affermare che il determinante si annulla quando

$$-(1-b)^2 e^{b\frac{\pi}{2}} + (1+b)^2 e^{-b\frac{\pi}{2}} = 0.$$

Riscritta l'ultima equazione come $\left(\frac{1-b}{1+b}\right)^2 e^{\pi b} = 1$, le sue soluzioni sono date dall'intersezione fra la retta di equazione $y = 1$ e la funzione $y(b) = \left(\frac{1-b}{1+b}\right)^2 e^{\pi b}$ per $b > 0$. Occorre quindi rappresentare graficamente la funzione y .

Poiché $y'(b) = (b-1) \frac{e^{\pi b}}{(b+1)^3} (4 + (b^2-1)\pi)$ la funzione è decrescente per $0 < b < 1$, crescente per $b > 1$ e raggiunge un minimo (assoluto) per $b = 1$. Inoltre poiché $y(0) = 1$, $y(1) = 0$ e $\lim_{b \rightarrow +\infty} y(b) = +\infty$



il grafico (tralasciando lo studio della concavità) è del tipo riportato nella figura seguente:

Poiché i due grafici si intersecano in un unico punto dove $\tilde{b} > 1$, esiste un $\tilde{\lambda} < -2$ in cui il determinante del sistema si annulla. La soluzione non banale del sistema sarà proporzionale a $(1, \frac{-(1-\sqrt{-\tilde{\lambda}+1})}{(1+\sqrt{-\tilde{\lambda}+1})})$. Si ottiene quindi la soluzione del nostro problema spettrale (per $\lambda < -1$)

$$\left\{ \tilde{\lambda}; \tilde{\varphi} = e^{-x} \left(e^{-\sqrt{-\tilde{\lambda}+1}x} - \frac{(1 - \sqrt{-\tilde{\lambda}+1})}{(1 + \sqrt{-\tilde{\lambda}+1})} e^{\sqrt{-\tilde{\lambda}+1}x} \right) \right\}.$$

2. Sia $\lambda = -1$. In tal caso abbiamo due radici reali e coincidenti $\alpha_1 = \alpha_2 = -1$.

La soluzione generale è $\varphi(x) = e^{-x} (c_1 + c_2 x)$. Poiché $\varphi(0) = c_1$, $\varphi'(0) = -c_1 + c_2$ e $\varphi'(\frac{\pi}{2}) = e^{-\frac{\pi}{2}} (-c_1 + c_2 - c_2 \frac{\pi}{2})$, le costanti c_1 e c_2 devono soddisfare il sistema di equazioni lineari (rispetto a c_1 e c_2) e omogenee

$$\begin{cases} c_1 + c_2 = 0 \\ -c_1 + (1 - \frac{\pi}{2})c_2 = 0 \end{cases}$$

Poiché il determinante di questo sistema di equazioni lineari è sempre diverso da zero, l'unica soluzione è quella banale. Quindi tale caso non è da prendere in considerazione in quanto conduce alla soluzione $\varphi(x) \equiv 0$.

3. Sia $\lambda > -1$. In tal caso abbiamo due radici complesse coniugate $\alpha_{1,2} = -1 \pm i\sqrt{\lambda+1}$.

La soluzione generale è

$$\varphi(x) = e^{-x} \left(c_1 \cos \sqrt{\lambda+1}x + c_2 \sin \sqrt{\lambda+1}x \right).$$

Essendo $\varphi(0) = c_1$, $\varphi'(0) = -c_1 + c_2\sqrt{\lambda+1}$ e

$$\varphi'\left(\frac{\pi}{2}\right) = e^{-\frac{\pi}{2}} \left(-c_1 \cos \sqrt{\lambda+1}\frac{\pi}{2} - c_2 \sin \sqrt{\lambda+1}\frac{\pi}{2} - c_1 \sqrt{\lambda+1} \sin \sqrt{\lambda+1}\frac{\pi}{2} + c_2 \sqrt{\lambda+1} \cos \sqrt{\lambda+1}\frac{\pi}{2} \right),$$

le costanti c_1 e c_2 sono da determinare mediante il sistema di equazioni lineari (rispetto a c_1 e c_2) e omogenee

$$\begin{cases} c_1 + c_2\sqrt{\lambda+1} = 0 \\ c_1(-\cos \sqrt{\lambda+1}\frac{\pi}{2} - \sqrt{\lambda+1} \sin \sqrt{\lambda+1}\frac{\pi}{2}) + \\ + c_2(-\sin \sqrt{\lambda+1}\frac{\pi}{2} + \sqrt{\lambda+1} \cos \sqrt{\lambda+1}\frac{\pi}{2}) = 0 \end{cases}$$

Dalla prima equazione del sistema si ricava $c_1 = -c_2\sqrt{\lambda+1}$, che sostituita nella seconda equazione fornisce

$$c_2 \left[2\sqrt{\lambda+1} \cos \sqrt{\lambda+1}\frac{\pi}{2} + \lambda \sin \sqrt{\lambda+1}\frac{\pi}{2} \right] = 0.$$

Essendo interessati a soluzioni non banali del sistema non consideriamo $c_2 = 0$ che comporta anche $c_1 = 0$. Quindi l'equazione di sopra è soddisfatta quando

$$2\sqrt{\lambda+1} \cos \sqrt{\lambda+1}\frac{\pi}{2} + \lambda \sin \sqrt{\lambda+1}\frac{\pi}{2} = 0.$$

Al fine di risolvere questa equazione si osserva che:

- a. Se $\cos \sqrt{\lambda+1}\frac{\pi}{2} = 0$ allora $\sin \sqrt{\lambda+1}\frac{\pi}{2} \neq 0$ (non esistendo un arco x per cui simultaneamente si abbia $\sin x = \cos x = 0$) e l'equazione è soddisfatta per $\lambda = 0$. Per tale valore di λ si vede facilmente

che il sistema possiede soluzioni non banali proporzionali al vettore $(1, -1)$. Si ottiene così la seguente soluzione del problema spettrale

$$\{\lambda_0 = 0; \varphi_0 = e^{-x}(\cos x - \sin x)\}.$$

- b. Se $\cos \sqrt{\lambda + 1} \frac{\pi}{2} \neq 0$ dividendo primo e secondo membro per tale quantità si perviene alla seguente equazione

$$2\sqrt{\lambda + 1} + \lambda \tan \sqrt{\lambda + 1} \frac{\pi}{2} = 0.$$

Quest'ultima equazione, supposto $\lambda \neq 0$, può risciversi come

$$\tan \sqrt{\lambda + 1} \frac{\pi}{2} = -2 \frac{\sqrt{\lambda + 1}}{\lambda}.$$

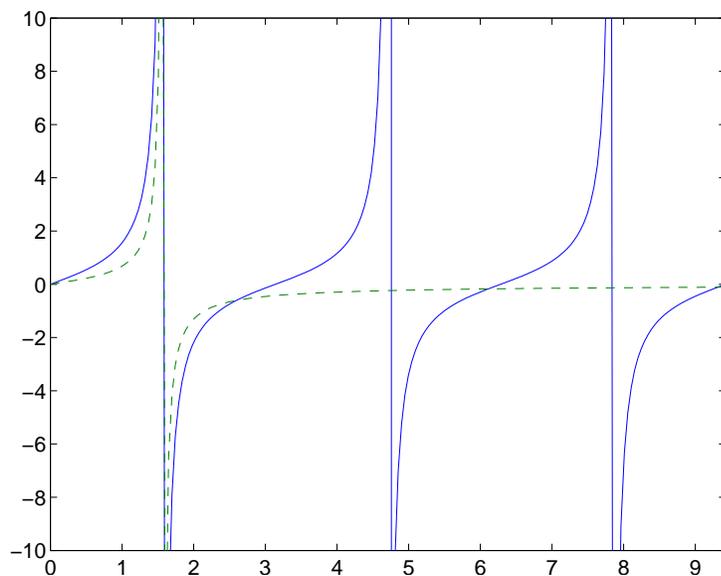
Per risolvere tale equazione poniamo $\sqrt{\lambda + 1} \frac{\pi}{2} = z > 0$, da cui si ricava la seguente equazione

$$\tan z = -\frac{4\pi z}{4z^2 - \pi^2}.$$

Anche in questo caso occorre procedere per via grafica. Nei punti in cui il grafico della funzione $y = \tan z$ interseca quello della funzione $y = -\frac{4\pi z}{4z^2 - \pi^2}$ il determinante del sistema si annulla e quindi per quei valori di λ (ricavabili dall'equazione $\sqrt{\lambda + 1} \frac{\pi}{2} = z$) il sistema possiede soluzioni non banali. Per tracciare il grafico della funzione $y = -\frac{4\pi z}{4z^2 - \pi^2}$ per $z > 0$ osserviamo che $\lim_{z \rightarrow 0^+} y = 0$, $\lim_{z \rightarrow \infty} y = 0$. Inoltre in $z = \frac{\pi}{2}$ la funzione ha un asintoto verticale (più precisamente $\lim_{z \rightarrow \frac{\pi}{2}^+} y = -\infty$ e $\lim_{z \rightarrow \frac{\pi}{2}^-} y = +\infty$) e la derivata prima ($y' = \frac{16\pi z^2 + 4\pi^3}{(4z^2 - \pi^2)^2}$) è strettamente maggiore di zero (dunque la funzione è sempre crescente).

Si vede che i due grafici si intersecano in un insieme numerabile di punti $\{z_k : k = 1, 2, \dots\}$. Le autofunzioni si ottengono in corrispondenza di quei $z_k = \sqrt{\lambda_k + 1} \frac{\pi}{2}$, $k = 1, 2, \dots$ in cui i due grafici si intersecano. In corrispondenza di questi z_k , ricordando che $c_1 = -c_2 \sqrt{\lambda + 1} = -\frac{2}{\pi} z$, si hanno le seguenti autofunzioni

$$\varphi_k(x) = e^{-x} \left(-\frac{2}{\pi} z_k \cos \frac{2}{\pi} z_k x + \sin \frac{2}{\pi} z_k x \right)$$



Perciò la soluzione generale è data da

$$\varphi(x) = \tilde{a}\tilde{\varphi}(x) + a_0\varphi_0(x) + \sum_{k=1}^{\infty} a_k\varphi_k(x),$$

essendo $\tilde{\varphi}(x)$ l'autofunzione relativa all'autovalore $\tilde{\lambda} < -2$, $\varphi_0(x)$ l'autofunzione relativa all'autovalore $\lambda_0 = 0$, $\varphi_k(x)$ le autofunzioni relative agli autovalori $z_k = \sqrt{\lambda_k + 1}\frac{\pi}{2}$, $k = 1, 2, \dots$ con $\lambda_k > -1$ e \tilde{a} , a_0 , a_k costanti da determinare. Rimane da imporre la condizione

$$e^{-\sqrt{2}x} = \tilde{a}\tilde{\varphi}(x) + a_0\varphi_0(x) + \sum_{k=1}^{\infty} a_k\varphi_k(x).$$

Per fare questo osserviamo che l'equazione $\varphi'' + 2\varphi' + (2 + \lambda)\varphi = 0$ può essere riscritta come

$$\frac{d}{dx}(e^{2x}\varphi') + e^{2x}(2 + \lambda)\varphi = 0,$$

che è del tipo di Sturm-Liouville con $r(x) = e^{2x} > 0$. Quindi essendo le autofunzioni ortogonali rispetto al peso $r(x)$ si hanno le seguenti espressioni

dei coefficienti

$$\begin{aligned}\tilde{a} &= \frac{\int_0^{\frac{\pi}{2}} e^{2x} e^{-\sqrt{2}x} \tilde{\varphi}(x) dx}{\int_0^{\frac{\pi}{2}} e^{2x} \tilde{\varphi}(x)^2 dx}, \\ a_0 &= \frac{\int_0^{\frac{\pi}{2}} e^{2x} e^{-\sqrt{2}x} \varphi_0(x) dx}{\int_0^{\frac{\pi}{2}} e^{2x} \varphi_0(x)^2 dx}, \\ a_k &= \frac{\int_0^{\frac{\pi}{2}} e^{2x} e^{-\sqrt{2}x} \varphi_k(x) dx}{\int_0^{\frac{\pi}{2}} e^{2x} \varphi_k(x)^2 dx}, \text{ con } k = 1, 2, \dots\end{aligned}$$

b) Risolvere, mediante la separazione delle variabili, il seguente problema differenziale

$$\begin{cases} u_t = 6u_{xx} + 3u_x + u, & \text{con } 0 \leq x \leq 2, t > 0 \\ u_x(0, t) + u(0, t) = 0 \\ u(2, t) = 0 \\ u(x, 0) = x^2(2 - x) \end{cases}$$

Si cercano soluzioni del tipo $u(x, t) = X(x)T(t)$. Sostituendo nell'equazione si trova

$$\frac{T'}{T} = 6\frac{X''}{X} + 3\frac{X'}{X} + 1 = -\lambda,$$

dove, per comodità, non sono indicate le dipendenze di T da t e di X da x . Si perviene quindi al seguente problema spettrale

$$\begin{cases} 6X'' + 3X' + (\lambda + 1)X = 0, & \text{con } 0 \leq x \leq 2 \\ X'(0) + X(0) = 0 \\ X(2) = 0 \end{cases}$$

Altre informazioni vengono poi ottenute dall'equazione

$$T' = -\lambda T.$$

Iniziamo a risolvere il problema spettrale. Si cercano soluzioni del tipo $X(x) \simeq e^{\alpha x}$ che conducono all'equazione caratteristica

$$6\alpha^2 + 3\alpha + (1 + \lambda) = 0.$$

Questa equazione ha come soluzioni $\alpha_{1,2} = \frac{-3 \pm \sqrt{-(24\lambda + 15)}}{12}$. Occorre distinguere tre casi in base al segno del discriminante.

1. Sia $\lambda = -\frac{5}{8}$.

In tal caso abbiamo due radici reali e coincidenti $\alpha_1 = \alpha_2 = -\frac{1}{4}$.

La soluzione generale è $X(x) = e^{-\frac{1}{4}x}(c_1 + c_2 x)$. Poiché $X(0) = c_1$, $X'(0) = -\frac{1}{4}c_1 + c_2$ e $X(2) = e^{-\frac{1}{2}}(c_1 + 2c_2)$, le costanti c_1 e c_2 devono soddisfare il sistema

$$\begin{cases} \frac{3}{4}c_1 + c_2 = 0 \\ c_1 + 2c_2 = 0 \end{cases}$$

Poiché il determinante di questo sistema di equazioni lineari è sempre diverso da zero ($\det = \frac{3}{2} - 1$), questo sistema ammette come unica soluzione quella banale. Quindi tale caso non è da prendere in considerazione in quanto conduce alla soluzione $X(x) \equiv 0$.

2. Sia $\lambda < -\frac{5}{8}$.

In tal caso abbiamo due radici reali e distinte $\alpha_{1,2} = \frac{-3 \pm \sqrt{-(24\lambda + 15)}}{12}$.

La soluzione generale è $X(x) = e^{-\frac{x}{4}}(c_1 e^{-\mu x} + c_2 e^{\mu x})$, dove si è posto $\mu = \frac{\sqrt{-24\lambda - 15}}{12} > 0$. Poiché $X(0) = c_1 + c_2$, $X'(0) = c_1(-\frac{1}{4} - \mu) + c_2(-\frac{1}{4} + \mu)$ e $X(2) = e^{-\frac{1}{2}}(c_1 e^{-2\mu} + c_2 e^{2\mu})$, le costanti c_1 e c_2 sono da determinare attraverso il sistema di equazioni lineari (rispetto a c_1 e c_2) e omogenee

$$\begin{cases} c_1 e^{-2\mu} + c_2 e^{2\mu} = 0 \\ (\frac{3}{4} - \mu)c_1 + (\frac{3}{4} + \mu)c_2 = 0 \end{cases}$$

Il determinante di questo sistema è

$$(\frac{3}{4} + \mu)e^{-2\mu} - (\frac{3}{4} - \mu)e^{2\mu}.$$

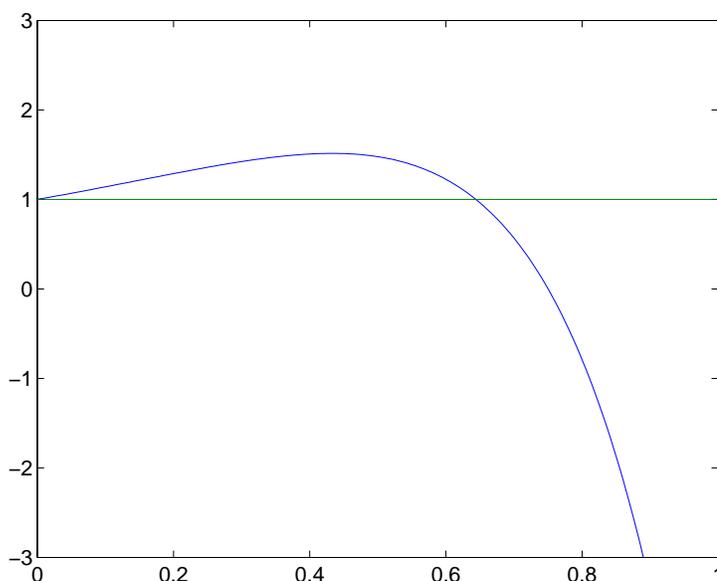
A noi interessano eventuali soluzioni non banali del sistema di equazioni lineari e queste si trovano quando il determinante si annulla, cioè quando

$$(\frac{3}{4} + \mu)e^{-2\mu} - (\frac{3}{4} - \mu)e^{2\mu} = 0.$$

Tenuto conto che $\mu > 0$, $\frac{3}{4} + \mu > 0$, l'ultima equazione può essere riscritta come

$$\left(\frac{3 - 4\mu}{3 + 4\mu}\right) e^{4\mu} = 1.$$

Le soluzioni di questa equazione sono date dall'intersezione fra la retta di equazione $y = 1$ e la funzione $y = \left(\frac{3-4\mu}{3+4\mu}\right) e^{4\mu}$ per $\mu > 0$. Occorre quindi rappresentare graficamente la funzione y . Poiché $y' = \frac{4e^{4\mu}}{(3+4\mu)^2} (-16\mu^2 + 3)$ la funzione è crescente per $0 < \mu < \frac{\sqrt{3}}{4}$, decrescente per $\mu > \frac{\sqrt{3}}{4}$ e raggiunge il massimo (assoluto) per $\mu = \frac{\sqrt{3}}{4}$. Inoltre poiché $\lim_{\mu \rightarrow 0^+} y = 1$, $y(\frac{3}{4}) = 0$ e $\lim_{\mu \rightarrow +\infty} y = -\infty$ il grafico (tralasciando lo studio della concavità) è quello riportato nella figura successiva.



Poiché i due grafici si intersecano in un unico punto di ascissa $\tilde{\mu}$ tale che $\frac{\sqrt{3}}{4} < \tilde{\mu} < \frac{3}{4}$, con l'equazione $\tilde{\mu} = \frac{\sqrt{-24\lambda-15}}{12}$ è possibile determinare il valore $\tilde{\lambda}$ di λ per cui il determinante si annulla. Quindi, in corrispondenza di $\tilde{\mu}$ si trova l'autofunzione

$$\tilde{X}(x) = e^{-\frac{x}{4}}(e^{-\tilde{\mu}x} - e^{-4\tilde{\mu}} e^{\tilde{\mu}x}).$$

In corrispondenza all'autovalore $\tilde{\mu}$ abbiamo inoltre l'equazione

$$\tilde{T}' = -\tilde{\mu}\tilde{T},$$

da cui si ricava $\tilde{T} = e^{-\tilde{\mu}t}$. Otteniamo quindi la soluzione (del problema principale)

$$\tilde{u}(x, t) = e^{-\tilde{\mu}t} e^{-\frac{x}{4}}(e^{-\tilde{\mu}x} - e^{-4\tilde{\mu}} e^{\tilde{\mu}x}).$$

3. Sia $\lambda > -\frac{5}{8}$.

In tal caso abbiamo due radici complesse coniugate $\alpha_{1,2} = \frac{-3 \pm i\sqrt{24\lambda+15}}{12}$. La soluzione generale è $X(x) = e^{-\frac{x}{4}}(c_1 \cos \mu x + c_2 \sin \mu x)$, dove si è posto $\mu = \frac{\sqrt{24\lambda+15}}{12} > 0$. Poiché $X(0) = c_1$, $X'(0) = -\frac{1}{4}c_1 + c_2\mu$ e $X(2) = e^{-\frac{1}{2}}(c_1 \cos 2\mu + c_2 \sin 2\mu)$, le costanti c_1 e c_2 sono da determinare attraverso il sistema di equazioni lineari (rispetto a c_1 e c_2) e omogenee

$$\begin{cases} \frac{3}{4}c_1 + c_2\mu = 0 \\ c_1 \cos 2\mu + c_2 \sin 2\mu = 0. \end{cases}$$

Dalla prima equazione del sistema si ricava $c_1 = -\frac{4}{3}c_2\mu$, che sostituita nella seconda equazione fornisce

$$c_2 \left[-\frac{4}{3}\mu \cos 2\mu + \sin 2\mu \right] = 0.$$

Essendo interessati a soluzioni non banali del sistema non consideriamo $c_2 = 0$ che comporta anche $c_1 = 0$. Quindi l'equazione di sopra è soddisfatta quando

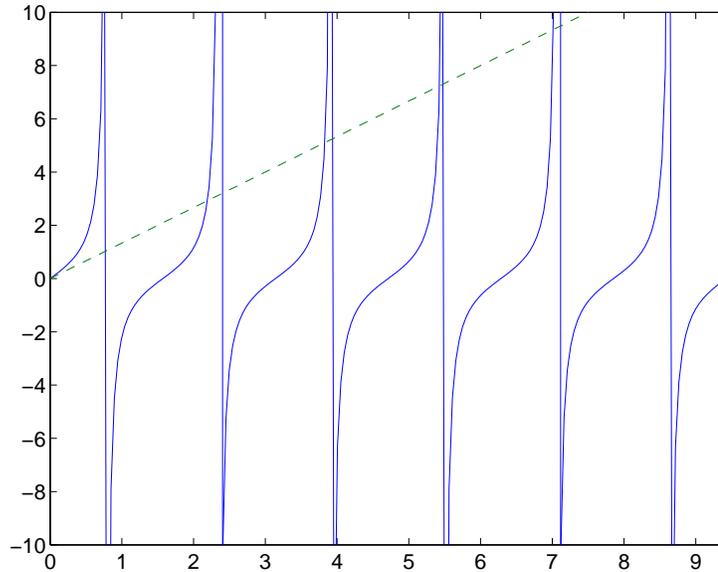
$$-\frac{4}{3}\mu \cos 2\mu + \sin 2\mu = 0.$$

Osserviamo che se fosse $\cos 2\mu = 0$ allora, affinché l'ultima equazione scritta sia soddisfatta, dovrebbe essere $\sin 2\mu = 0$, il che è impossibile. Possiamo allora assumere che $\cos 2\mu \neq 0$. Sotto questa ipotesi, con semplici calcoli si perviene all'equazione

$$\tan 2\mu = \frac{2}{3}(2\mu).$$

Per risolvere questa equazione si procede per via grafica (vedi figura sotto).

I grafici delle funzioni $y = \tan 2\mu$ e $y = \frac{2}{3}(2\mu)$ (deve essere $\mu > 0$!) si intersecano in un insieme numerabile di punti in ciascuno dei quali il sistema di equazioni lineari possiede una soluzione non banale. Dal grafico si vede che le ascisse μ_k dei punti d'intersezione soddisfano la seguente condizione $\frac{\pi}{4} + k\frac{\pi}{2} < \mu_k < \frac{\pi}{4} + (k+1)\frac{\pi}{2}$, $k = 0, 1, 2, \dots$. Un volta noti questi valori di μ_k è possibile determinare i corrispondenti



valori di λ_k per cui il sistema possiede soluzioni non banali attraverso la relazione $\lambda_k = -\frac{5}{8} + 6\mu_k^2$. È facile vedere che la soluzione che si ottiene per $k = 0$ non è accettabile in quanto comporta $\lambda_0 = -\frac{5}{8}$ e questo caso è stato già analizzato (e escluso) in precedenza. Quindi, in corrispondenza agli autovalori $\frac{\pi}{4} + k\frac{\pi}{2} < \mu_k < \frac{\pi}{4} + (k+1)\frac{\pi}{2}$, $k = 1, 2, 3 \dots$ del problema spettrale si trovano le seguenti autofunzioni

$$X_k(x) = e^{-\frac{x}{4}} \left(-\frac{4}{3}\mu_k \cos \mu_k x + \sin \mu_k x \right).$$

In corrispondenza abbiamo anche le equazioni $T'_k + \mu_k T = 0$, che conducono alle soluzioni

$$T_k = e^{-\mu_k t}.$$

In definitiva troviamo le autofunzioni

$$u_k(x, t) = T_k(t)X_k(x) = e^{-\mu_k t} e^{-\frac{x}{4}} \left(-\frac{4}{3}\mu_k \cos \mu_k x + \sin \mu_k x \right).$$

Siamo quindi in grado di scrivere la soluzione del problema, che risulta

$$u(x, t) = \tilde{a} \tilde{u}(x, t) + \sum_{k=1}^{\infty} c_k u_k(x, t) = \tilde{a} e^{-\tilde{\mu}t} e^{-\frac{x}{4}} (e^{-\tilde{\mu}x} - e^{-4\tilde{\mu}} e^{\tilde{\mu}x}) \\ + \sum_{k=1}^{\infty} c_k e^{-\mu_k t} e^{-\frac{x}{4}} \left(-\frac{4}{3} \mu_k \cos \mu_k x + \sin \mu_k x \right),$$

con \tilde{a} e c_k costanti da determinare. A tal fine imponiamo la condizione $u(x, 0) = x^2(2 - x)$. Questa implica

$$\tilde{a} e^{-\frac{x}{4}} (e^{-\tilde{\mu}x} - e^{-4\tilde{\mu}} e^{\tilde{\mu}x}) + \sum_{k=1}^{\infty} c_k e^{-\frac{x}{4}} \left(-\frac{4}{3} \mu_k \cos \mu_k x + \sin \mu_k x \right) = x^2(2 - x).$$

In questa condizione compaiono $\tilde{u}(x, 0)$ e le $u_k(x, 0)$ che sono soluzioni dell'equazione $X'' + \frac{1}{2}X' + \frac{1}{6}(\lambda + 1)X = 0$, la quale può scriversi nella forma canonica di Sturm-Liouville come

$$\frac{d}{dx} (e^{\frac{x}{2}} X') + e^{\frac{x}{2}} \left(\frac{1}{6} + \frac{1}{6} \lambda \right) X = 0.$$

Quindi la funzione $r(x) = \frac{e^{\frac{x}{2}}}{6} > 0$ non cambia segno e questo garantisce che le funzioni $\tilde{u}(x, 0)$ e $u_k(x, 0)$ ($k = 1, 2, 3, \dots$) siano ortogonali rispetto ad essa. Perciò si hanno le seguenti espressioni dei coefficienti

$$\tilde{a} = \frac{\int_0^2 e^{\frac{x}{2}} x^2 (2 - x) \tilde{u}(x, 0) dx}{\int_0^2 e^{\frac{x}{2}} \tilde{u}(x, 0)^2 dx}, \\ c_k = \frac{\int_0^2 e^{\frac{x}{2}} x^2 (2 - x) u_k(x, 0) dx}{\int_0^2 e^{\frac{x}{2}} u_k(x, 0)^2 dx}, \quad k = 1, 2, \dots$$

c) **Risolvere il seguente problema differenziale:**

$$\begin{cases} u_{tt} = 3u_{xx} + u_x + 2x, & \text{con } 0 < x < 1 \text{ e } t > 0 \\ u(0, t) = u(1, t) = 1 \\ u(x, 0) = x(1 - x), \quad u_t(x, 0) = 0. \end{cases}$$

Si cercano funzioni $u(x, t) = v(x, t) + \phi(x)$. Sostituendo nell'equazione si trova

$$v_{tt} = 3(v_{xx} + \phi'') + v_x + \phi' + 2x.$$

Quindi devono essere soddisfatti i problemi differenziali

$$\begin{cases} v_{tt} = 3v_{xx} + v_x, \\ v(0, t) = v(1, t) = 0 \\ v(x, 0) = x(1-x) - \phi(x), \quad v_t(x, 0) = 0. \end{cases} \quad (1.52)$$

e

$$\begin{cases} 3\phi'' + \phi' + 2x = 0 \\ \phi(0) = \phi(1) = 1. \end{cases} \quad (1.53)$$

Iniziamo a risolvere il problema (1.53). A tal fine osserviamo che una soluzione del problema si ottiene aggiungendo una soluzione particolare alla soluzione generale dell'equazione omogenea associata. La soluzione generale dell'equazione omogenea associata si ottiene cercando soluzioni del tipo $\phi_0(x) \simeq e^{\alpha x}$. Si ottiene l'equazione caratteristica $3\alpha^2 + \alpha = 0$, che ha come soluzioni $\alpha = 0$ e $\alpha = -\frac{1}{3}$. Per cui la soluzione generale dell'equazione omogenea è $\phi_0(x) = c_1 + c_2 e^{-\frac{1}{3}x}$. Cerchiamo adesso una soluzione particolare dell'equazione. Essa sarà della forma $\phi_1(x) = x(ax + b)$. Essendo $\phi_1'(x) = 2ax + b$, $\phi_1''(x) = 2a$, Sostituendo nell'equazione si trova $6a + b + 2ax + 2x = 0$ che deve essere identicamente soddisfatta. Questo capita se

$$\begin{cases} 6a + b = 0 \\ 2a + 2 = 0. \end{cases}$$

Da questo sistema si ricava quindi $a = -1$ e $b = 6$. Una soluzione particolare dell'equazione proposta è dunque $\phi_1(x) = x(-x + 6)$. La soluzione generale dell'equazione (1.53) è

$$\phi(x) = c_1 + c_2 e^{-\frac{1}{3}x} + x(-x + 6).$$

Imponendo le condizioni al contorno si vede che i coefficienti c_1 e c_2 devono soddisfare il seguente sistema di equazioni lineari (rispetto a c_1 e c_2)

$$\begin{cases} c_1 + c_2 = 1 \\ c_1 + e^{-\frac{1}{3}}c_2 = -4. \end{cases}$$

E' immediato verificare che il determinante di questo sistema di equazioni lineari è diverso da zero ($\det = e^{-\frac{1}{3}} - 1$) e quindi il sistema possiede un'unica

soluzione. Indicata con $(\tilde{c}_1, \tilde{c}_2)$ la soluzione del sistema, la soluzione del problema (1.53) è

$$\phi(x) = \tilde{c}_1 + \tilde{c}_2 e^{-\frac{1}{3}x} + x(-x + 6). \quad (1.54)$$

Passiamo allora alla risoluzione del problema (1.52). Procediamo con il metodo di separazione delle variabili. Cerchiamo soluzioni del tipo $v(x, t) = X(x)T(t)$. Sostituendo nell'equazione si perviene ai seguenti problemi (agli autovalori)

$$\begin{cases} 3X'' + X' + \lambda X = 0 \\ X(0) = X(1) = 0. \end{cases} \quad (1.55)$$

e

$$\begin{cases} T'' + \lambda T = 0 \\ T'(0) = 0. \end{cases} \quad (1.56)$$

Iniziamo a risolvere il problema (1.55). Cerchiamo soluzioni del tipo $X(x) \simeq e^{\alpha x}$. Si ottiene così l'equazione caratteristica $3\alpha^2 + \alpha + \lambda = 0$ che ammette come soluzioni

$$\alpha_{1,2} = \frac{-1 \pm \sqrt{1 - 12\lambda}}{6}.$$

Occorre quindi distinguere tre casi a seconda che sia $1 - 12\lambda = 0$, $1 - 12\lambda > 0$ o $1 - 12\lambda < 0$.

1. Sia $\lambda = \frac{1}{12}$.

In tal caso l'equazione caratteristica possiede due radici reali coincidenti $\alpha_1 = \alpha_2 = -\frac{1}{6}$.

La soluzione dell'equazione sarà $X(x) = e^{-\frac{1}{6}x}(a + bx)$, con a e b costanti da determinare. A tal fine essendo, $X(0) = a$ e $X(1) = e^{-\frac{1}{6}}(a + b)$, si vede che a e b devono essere ricavate dal sistema di equazioni lineari e omogenee

$$\begin{cases} a & = 0 \\ a + b & = 0. \end{cases}$$

Quest'ultimo ammette come unica soluzione quella banale $a = b = 0$, e quindi questo caso non si deve prendere in considerazione in quanto conduce alla soluzione banale $X(x) \equiv 0$.

2. Sia $\lambda < \frac{1}{12}$.

In tal caso l'equazione caratteristica possiede due radici reali e distinte $\alpha_{1,2} = \frac{-1 \pm \sqrt{1-12\lambda}}{6}$.

La soluzione dell'equazione sarà $X(x) = e^{-\frac{1}{6}x} (a e^{-\mu x} + b e^{\mu x})$, dove si è posto $\mu = \frac{\sqrt{1-12\lambda}}{6} > 0$ e con a e b costanti da determinare dal sistema di equazioni lineari e omogenee

$$\begin{cases} a + b = 0 \\ e^{-\mu} a + e^{\mu} b = 0. \end{cases}$$

Il determinante di questo sistema di equazioni lineari è $\det = e^{-\mu} - e^{\mu}$ che si annulla solo se $\mu = 0$. Ma poiché deve essere $\mu > 0$, si può concludere che $\det \neq 0$. Quindi il sistema ammette come unica soluzione quella banale $a = b = 0$, e anche questo caso non si deve prendere in considerazione in quanto conduce alla soluzione banale $X(x) \equiv 0$.

3. Sia $\lambda > \frac{1}{12}$.

In tal caso l'equazione caratteristica possiede due radici reali e distinte $\alpha_{1,2} = \frac{-1 \pm i\sqrt{-1+12\lambda}}{6}$. La soluzione dell'equazione sarà $X(x) = e^{-\frac{1}{6}x} (a \cos \mu x + b \sin \mu x)$, dove si è posto $\mu = \frac{\sqrt{-1+12\lambda}}{6} > 0$ e con a e b costanti da determinare dal sistema di equazioni lineari e omogenee

$$\begin{cases} a = 0 \\ a \cos \mu + b \sin \mu = 0. \end{cases}$$

Quindi le soluzioni non banali di questo sistema si ricavano dall'equazione $\sin \mu = 0$. Si trova $\frac{\sqrt{-1+12\lambda_k}}{6} = k\pi, k = 1, 2, \dots$. In corrispondenza a questi autovalori si hanno le seguenti autofunzioni $X_k(x) = e^{-\frac{1}{6}x} (\sin k\pi x), k = 1, 2, \dots$. In corrispondenza degli autovalori sopra trovati occorre risolvere

$$\begin{cases} T_k'' + \lambda_k T_k = 0 \\ T_k'(0) = 0. \end{cases}$$

L'ultima equazione ammette come soluzione $T_k(t) = a_k \cos \sqrt{\lambda_k} t + b_k \sin \sqrt{\lambda_k} t$. Poiché $T_k'(0) = b_k \sqrt{\lambda_k} = 0$ implica $b_k = 0$ si ottiene $T_k(t) = a_k \cos \sqrt{\lambda_k} t$.

Quindi, relativamente agli autovalori $\frac{\sqrt{-1+12\lambda_k}}{6} = k\pi, k = 1, 2, \dots$, si trovano le seguenti autofunzioni $v_k(x, t) = e^{-\frac{1}{6}x} (\sin k\pi x) \cos \sqrt{\lambda_k} t$.

Rimane da imporre la condizione $v(x, 0) = x(1 - x) - \phi(x)$ con $\phi(x)$ data da (1.54). A tal fine osserviamo che

$$v(x, t) = \sum_{k=1}^{\infty} c_k e^{-\frac{1}{6}x} \sin k\pi x \cos \sqrt{\lambda_k} t,$$

e quindi $\sum_{k=1}^{\infty} c_k e^{-\frac{1}{6}x} \sin k\pi x = x(1 - x) - \phi(x)$. Questa equazione può essere riscritta come

$$\sum_{k=1}^{\infty} c_k (\sin k\pi x) = e^{\frac{1}{6}x} (x(1 - x) - \phi(x)),$$

dalla quale è possibile ricavare i coefficienti b_k attraverso l'analisi di Fourier (a primo membro compare infatti una serie di Fourier). Infatti, moltiplicando primo e secondo membro per $\sin h\pi x$ e integrando (tenendo conto dell'ortogonalità delle funzioni $\sin k\pi x \sin h\pi x$) si trova

$$c_k = 2 \int_0^1 e^{\frac{1}{6}x} (x(1 - x) - \phi(x)) \sin k\pi x. \quad (1.57)$$

Finalmente si può scrivere la soluzione del problema come

$$v(x, t) = \sum_{k=1}^{\infty} c_k e^{-\frac{1}{6}x} \sin k\pi x \cos \sqrt{\lambda_k} t + \tilde{c}_1 + \tilde{c}_2 e^{-\frac{1}{3}x} + x(-x + 6),$$

con i c_k dati dalle equazioni (1.57).

d) Risolvere, mediante il metodo degli integrali generali, il problema differenziale:

$$\begin{cases} u_{tt} = c^2 u_{xx}, & \text{con } 0 < x < \pi \text{ e } t > 0 \\ u(0, t) + u(\pi, t) = 0 \\ u(x, 0) = \sin x, \quad u_t(x, 0) = \cos x. \end{cases}$$

e) Determinare, mediante il metodo degli integrali generali, la soluzione del seguente problema differenziale:

$$\begin{cases} u_{xy} = x + y \cos x, \\ u(x, 0) = 3x + 2 \\ u(0, y) = 2. \end{cases}$$

f) Risolvere, mediante il metodo della separazione delle variabili, il problema differenziale

$$\begin{cases} u_{tt} = 2u_{xx} + 3u_x + 4u_t + 5u, & \text{con } 0 \leq x \leq 2 \text{ e } t > 0 \\ u(x, 0) = 1 + \frac{x}{2}, & u_t(x, 0) = 0 \\ u(0, t) = 1, & u(2, t) = 2. \end{cases}$$

g) Risolvere, mediante il metodo degli integrali generali, il problema differenziale

$$\begin{cases} u_{tt} = 3u_{xx} + 3u_{xt}, \\ u(x, 0) = x \\ u_t(x, 0) = 1. \end{cases}$$

h) Risolvere, mediante il metodo della separazione delle variabili, il problema differenziale

$$\begin{cases} u_t = 4u_{xx} + 6u_x - 2u, \\ u_x(0, t) + u(0, t) = 0 \\ u(1, t) = 0 \\ u(x, 0) = \sin x. \end{cases}$$

i) Descrivere l'applicazione della tecnica della separazione delle variabili alla risoluzione del seguente problema differenziale:

$$\begin{cases} u_{xx} + 2u_{yy} + u_x + 2u_y = 0, \\ u(x, 0) = u(x, 1) = 0 \\ u(0, y) = y, & u(1, y) = \sin \pi y. \end{cases}$$

l) Risolvere, mediante il metodo della separazione delle variabili, il problema differenziale

$$\begin{cases} u_{tt} = 3u_{xx} + 2u, & \text{con } 0 < x < 1 \text{ e } t > 0 \\ u(0, t) = 2, & u(1, t) = 3 \\ u(x, 0) = x \sin x, & u_t(x, 0) = 0. \end{cases}$$

Prova intermedia di matematica applicata del 17/11/07

1) **Determinare, mediante il metodo degli integrali generali, la soluzione del seguente problema differenziale**

$$\begin{cases} u_{xx} + 2u_{xy} - 3u_{yy} = 0, \\ u(x, 0) = \sin x, \\ u_y(x, 0) = \frac{4}{3} \cos x. \end{cases}$$

Si cercano soluzioni del tipo $u(x, y) \simeq e^{ax+by}$ che, dopo aver derivato tale funzione e sostituito nell'equazione del problema, conduce all'equazione $a^2 + 2ab - 3b^2 = 0$. Dividendo tale equazione per $b \neq 0$ (se fosse $b = 0$ allora sarebbe anche $a = 0$ e tale caso è ovviamente da scartare) si ottiene

$$\left(\frac{a}{b}\right)^2 + 2\left(\frac{a}{b}\right) - 3 = 0.$$

Tale equazione ammette come soluzioni $\left(\frac{a}{b}\right)_1 = -3$ e $\left(\frac{a}{b}\right)_2 = 1$, quindi le soluzioni cercate sono del tipo $u(x, y) \simeq e^{b(-3x+y)}$ e $u(x, y) \simeq e^{x+y}$.

La soluzione generale sarà allora

$$u(x, y) = F(x + y) + G(-3x + y)$$

con le funzioni F e G derivabili almeno 2 volte rispetto a x e y . Dal teorema di derivazione delle funzioni composte si ha $u_y(x, y) = F'(x + y) + G'(-3x + y)$, quindi imponendo le condizioni si trova il seguente sistema

$$\begin{cases} F(x) + G(-3x) = \sin x \\ F'(x) + G'(-3x) = \frac{4}{3} \cos x. \end{cases}$$

Derivando la prima equazione si ottiene

$$\begin{cases} F'(x) - 3G'(-3x) = \cos x \\ F'(x) + G'(-3x) = \frac{4}{3} \cos x, \end{cases}$$

da cui, sottraendo dalla prima equazione la seconda, si trova $G'(-3x) = \frac{1}{12} \cos x$. Quindi $G(-3x) = \frac{-1}{4} \sin\left(\left(-\frac{1}{3}\right)(-3x)\right) + c = -\frac{1}{4} \sin x + c$, e di conseguenza $F(x) = \sin x + \frac{1}{4} \sin x - c = \frac{5}{4} \sin x - c$. Quindi la soluzione generale del problema è

$$u(x, y) = \frac{5}{4} \sin(x+y) - \frac{1}{4} \sin\left(\left(-\frac{1}{3}\right)(y-3x)\right) = \frac{5}{4} \sin(x+y) - \frac{1}{4} \sin\left(\frac{1}{3}(3x-y)\right).$$

È poi facile verificare che tale funzione è la soluzione del problema differenziale proposto.

2) **Risolvere, per separazione delle variabili, il problema differenziale**

$$\begin{cases} u_t = u_{xx} + 2u_x + 3u \\ u_x(0, t) + 2u(0, t) = 0 \\ u(\pi, t) = 3 \\ u(x, 0) = \sin \pi x \end{cases}$$

Si cercano soluzioni del tipo $u(x, t) = v(x, t) + \phi(x)$. Sostituendo nell'equazione si trova

$$\begin{cases} v_t = v_{xx} + \phi'' + 2v_x + 2\phi' + 3v + 3\phi \\ v_x(0, t) + \phi'(0) + 2v(0, t) + 2\phi(0) = 0 \\ v(\pi, t) + \phi(\pi) = 3 \\ v(x, 0) = \sin \pi x - \phi(x), \end{cases}$$

da cui si ottengono i due seguenti problemi

$$\begin{cases} v_t = v_{xx} + 2v_x + 3v \\ v_x(0, t) + 2v(0, t) = 0 \\ v(\pi, t) = 0 \\ v(x, 0) = \sin \pi x - \phi(x), \end{cases} \quad (1.58)$$

e

$$\begin{cases} \phi'' + 2\phi' + 3\phi = 0 \\ \phi'(0) + 2\phi(0) = 0 \\ \phi(\pi) = 3. \end{cases} \quad (1.59)$$

Iniziamo a risolvere il problema (1.59). Si cercano soluzioni del tipo $e^{\alpha x}$, che conducono all'equazione caratteristica $\alpha^2 + 2\alpha + 3 = 0$. Poiché tale equazione ammette come soluzioni $\alpha_{1,2} = -1 \pm i\sqrt{2}$, la soluzione cercata sarà del tipo $\phi(x) = e^{-x} (a \cos \sqrt{2}x + b \sin \sqrt{2}x)$, con a e b costanti da determinare. Poiché

$\phi(0) = a$, $\phi'(0) = -a + \sqrt{2}b$ e $\phi(\pi) = e^{-\pi} (a \cos \sqrt{2}\pi + b \sin \sqrt{2}\pi)$, dalle condizioni iniziali si ottiene il sistema

$$\begin{cases} a + \sqrt{2}b = 0 \\ (a \cos \sqrt{2}\pi + b \sin \sqrt{2}\pi) = 3e^\pi \end{cases}$$

che semplici calcoli mostrano avere un'unica soluzione (\tilde{a}, \tilde{b}) . Quindi la soluzione del problema (1.59) è

$$\phi(x) = e^{-x} \left(\tilde{a} \cos \sqrt{2}x + \tilde{b} \sin \sqrt{2}x \right).$$

Possiamo adesso soffermarci sul problema (1.58). Cerchiamo soluzioni del tipo $u(x, t) = X(x)T(t)$. Sostituendo nell'equazione si trova

$$\frac{T'}{T} = \frac{X''}{X} + 2\frac{X'}{X} + 3 = -\lambda,$$

dove, per comodità, non sono indicate le dipendenze di T da t e di X da x . Si perviene quindi al seguente problema spettrale

$$\begin{cases} X'' + 2X' + (\lambda + 3)X = 0, \\ X'(0) + 2X(0) = 0 \\ X(\pi) = 0 \end{cases}$$

Altre informazioni vengono poi ottenute dall'equazione

$$T' = -\lambda T.$$

Iniziamo a risolvere il problema spettrale. Si cercano soluzioni del tipo $X(x) \simeq e^{\alpha x}$ che conducono all'equazione caratteristica

$$\alpha^2 + 2\alpha + (3 + \lambda) = 0.$$

Questa equazione ha come soluzioni $\alpha_{1,2} = -1 \pm \sqrt{-(\lambda + 2)}$. Occorre distinguere tre casi a seconda che sia $-2 - \lambda = 0$, $-2 - \lambda > 0$ oppure $-2 - \lambda < 0$.

1. Sia $\lambda = -2$.

In tal caso abbiamo due radici reali e coincidenti $\alpha_1 = \alpha_2 = -1$.

La soluzione generale é $X(x) = e^{-x}(c_1 + c_2 x)$. Poiché $X(0) = c_1$,

$X'(0) = -c_1 + c_2$ e $X(\pi) = e^{-\pi}(c_1 + c_2\pi)$, le costanti c_1 e c_2 devono soddisfare il sistema

$$\begin{cases} c_1 + c_2 & = 0 \\ c_1 + c_2\pi & = 0 \end{cases}$$

Poiché il determinante di questo sistema di equazioni lineari è sempre diverso da zero ($\det = \pi - 1$), questo sistema ammette come unica soluzione quella banale. Quindi tale caso non è da prendere in considerazione in quanto conduce alla soluzione $X(x) \equiv 0$.

2. Sia $\lambda < -2$ e, per semplicità, poniamo $\lambda + 2 = -\beta^2$, $\beta > 0$. In tal caso abbiamo due radici reali e distinte $\alpha_{1,2} = -1 \pm \beta$ e la soluzione generale è

$$X(x) = e^{-x}(c_1 e^{\beta x} + c_2 e^{-\beta x}).$$

Poiché $X(0) = c_1 + c_2$, $X'(0) = (\beta - 1)c_1 - (\beta + 1)c_2$ e $X(\pi) = e^{-\pi}(c_1 e^{\beta\pi} + c_2 e^{-\beta\pi})$, le costanti c_1 e c_2 sono da determinare attraverso il sistema di equazioni lineari (rispetto a c_1 e c_2) e omogenee

$$\begin{cases} (\beta + 1)c_1 + (1 - \beta)c_2 = 0 \\ c_1 e^{\beta\pi} + c_2 e^{-\beta\pi} = 0 \end{cases}$$

Il determinante di questo sistema è

$$\beta + 1 - (1 - \beta)e^{2\beta\pi}.$$

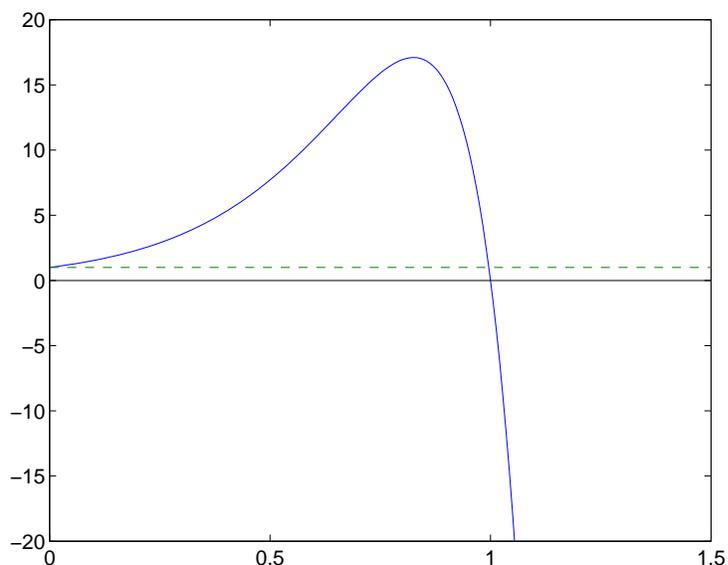
A noi interessano eventuali soluzioni non banali del sistema di equazioni lineari e queste si trovano quando il determinante si annulla, cioè quando

$$\beta + 1 - (1 - \beta)e^{2\beta\pi} = 0.$$

Tenuto conto che per $\beta > 0$ si ha $1 + \beta > 0$ si può riscrivere l'ultima equazione (dividendo ambo i membri per $1 + \beta > 0$) come

$$\left(\frac{1 - \beta}{1 + \beta}\right) e^{2\beta\pi} = 1.$$

Le soluzioni, se esistono, sono date dall'intersezione fra la funzione $y = 1$ e la funzione $y = \left(\frac{1 - \beta}{1 + \beta}\right) e^{2\beta\pi}$. Occorre quindi rappresentare graficamente tali funzioni (vedi figura successiva).



Poiché i due grafici si intersecano in un unico punto di ascissa $\tilde{\beta}$ tale che $0 < \tilde{\beta} < 2$, con l'equazione $\lambda + 2 = -\tilde{\beta}^2$ è possibile determinare il valore $\tilde{\lambda}$ di λ per cui il determinante si annulla. Quindi, in corrispondenza di $\tilde{\beta}$ si trova l'autofunzione

$$\tilde{X}(x) = e^{-x}(e^{\tilde{\beta}(2\pi-x)} + e^{\tilde{\beta}x}).$$

Nell'ultima equazione si è tenuto conto che $c_2 = -c_1 e^{2\beta\pi}$. In corrispondenza all'autovalore $\tilde{\beta}$ abbiamo inoltre l'equazione

$$\tilde{T}' = -\tilde{\beta}\tilde{T},$$

da cui si ricava $\tilde{T} = e^{-\tilde{\beta}t}$. Otteniamo quindi la soluzione (del problema principale)

$$\tilde{v}(x, t) = e^{-\tilde{\beta}t-x}(e^{\tilde{\beta}(2\pi-x)} + e^{\tilde{\beta}x}).$$

3. Sia $\lambda > -2$, e poniamo $\lambda + 2 = \beta^2$, $\beta > 0$.

In tal caso abbiamo due radici complesse coniugate $\alpha_{1,2} = -1 \pm i\beta$ e la soluzione generale è

$$X(x) = e^{-x}(c_1 \cos \beta x + c_2 \sin \beta x).$$

Poiché $X(0) = c_1$, $X'(0) = -c_1 + c_2\beta$ e $X(\pi) = e^{-\pi}(c_1 \cos \beta\pi + c_2 \sin \beta\pi)$, le costanti c_1 e c_2 sono da determinare attraverso il sistema

di equazioni lineari e omogenee

$$\begin{cases} c_1 + c_2\beta = 0 \\ c_1 \cos \beta\pi + c_2 \sin \beta\pi = 0. \end{cases}$$

Dalla prima equazione del sistema si ricava $c_1 = -c_2\beta$, che sostituita nella seconda equazione fornisce

$$c_2 [-\beta \cos \beta\pi + \sin \beta\pi] = 0.$$

Essendo interessati a soluzioni non banali del sistema non consideriamo $c_2 = 0$ che comporta anche $c_1 = 0$. Quindi l'equazione di sopra è soddisfatta quando

$$-\beta \cos \beta\pi + \sin \beta\pi = 0.$$

Osserviamo che se fosse $\cos \beta\pi = 0$ allora, affinché l'ultima equazione scritta sia soddisfatta, dovrebbe essere $\sin \beta\pi = 0$, ma questo è impossibile non esistendo un arco x per cui $\cos x = \sin x = 0$. Possiamo allora assumere che $\cos \beta\pi \neq 0$. Sotto questa ipotesi, con semplici calcoli si perviene all'equazione

$$\tan \beta\pi = \frac{\beta\pi}{\beta}.$$

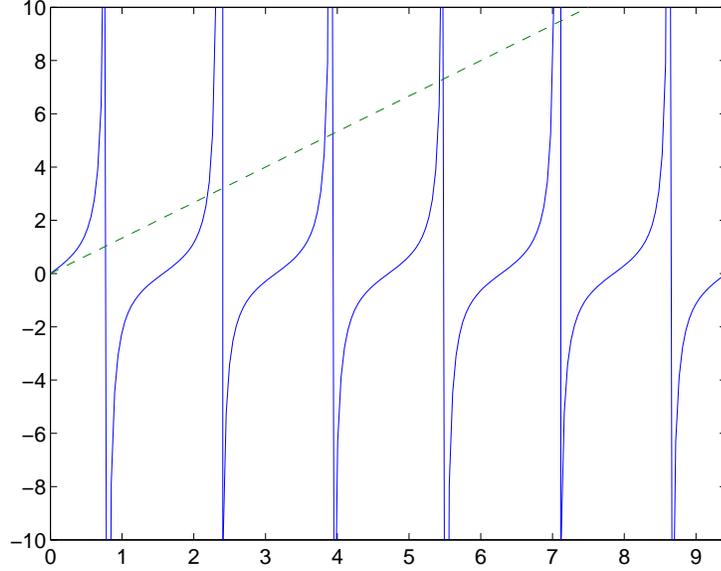
Per risolvere questa equazione si procede per via grafica (vedi figura sotto).

I grafici delle funzioni $y = \tan \beta\pi$ e $y = \frac{\beta\pi}{\beta}$ (deve essere $\beta > 0$) si intersecano in un insieme numerabile di punti in ciascuno dei quali il sistema di equazioni lineari possiede una soluzione non banale. Dal grafico si vede che le ascisse β_k dei punti d'intersezione soddisfano la seguente condizione $k < \beta_k < \frac{2k+1}{2}$, $k = 1, 2, \dots$. Un volta noti questi valori di β_k è possibile determinare i corrispondenti valori di λ_k per cui il sistema possiede soluzioni non banali attraverso la relazione $\lambda_k = \beta_k^2 - 2$. Quindi, in corrispondenza agli autovalori $k < \beta_k < \frac{2k+1}{2}$, $k = 1, 2, 3 \dots$ del problema spettrale si trovano le seguenti autofunzioni

$$X_k(x) = e^{-x} (-\beta_k \cos \beta_k x + \sin \beta_k x).$$

In corrispondenza abbiamo anche le equazioni $T_k' + \beta_k T = 0$, che conducono alle soluzioni

$$T_k = e^{-\beta_k t}.$$



In definitiva troviamo le autofunzioni

$$v_k(x, t) = T_k(t)X_k(x) = e^{-\beta_k t} e^{-x} (-\beta_k \cos \beta_k x + \sin \beta_k x).$$

Siamo quindi in grado di scrivere la soluzione del problema, che risulta

$$\begin{aligned} u(x, t) &= \tilde{a} \tilde{v}(x, t) + \sum_{k=1}^{\infty} c_k v_k(x, t) = \tilde{a} e^{-\tilde{\beta} t - x} (e^{\tilde{\beta}(2\pi - x)} + e^{\tilde{\beta} x}) \\ &+ \sum_{k=1}^{\infty} c_k e^{-\beta_k t} e^{-x} (-\beta_k \cos \beta_k x + \sin \beta_k x), \end{aligned}$$

con \tilde{a} e c_k costanti da determinare. A tal fine imponiamo la condizione $u(x, 0) = \sin \pi x - \phi(x)$. Questa implica

$$\tilde{a} e^{-x} (e^{\tilde{\beta}(2\pi - x)} + e^{\tilde{\beta} x}) + \sum_{k=1}^{\infty} c_k e^{-x} (-\beta_k \cos \beta_k x + \sin \beta_k x) = \sin \pi x - \phi(x).$$

In questa condizione compaiono $\tilde{u}(x, 0)$ e le $u_k(x, 0)$ che sono soluzioni dell'equazione $X'' + 2X' + (\lambda + 3)X = 0$, la quale può scriversi nella forma canonica di Sturm-Liouville come

$$\frac{d}{dx} (e^{2x} X') + e^{2x} (3 + \lambda) X = 0.$$

Quindi la funzione $r(x) = e^{2x} > 0$ non cambia segno e questo garantisce che le funzioni $\tilde{u}(x, 0)$ e $u_k(x, 0)$ ($k = 1, 2, 3, \dots$) siano ortogonali rispetto ad essa. Perciò si hanno le seguenti espressioni dei coefficienti

$$\tilde{a} = \frac{\int_0^\pi e^{2x}(\sin \pi x - \phi(x))\tilde{u}(x, 0)dx}{\int_0^\pi e^{2x}\tilde{u}(x, 0)^2 dx},$$

$$c_k = \frac{\int_0^\pi e^{2x}(\sin \pi x - \phi(x))u_k(x, 0)dx}{\int_0^\pi e^{2x}u_k(x, 0)^2 dx}, \quad k = 1, 2, \dots$$

3) Determinare i problemi differenziali ai quali conduce l'applicazione del metodo della separazione delle variabili al seguente problema:

$$\left\{ \begin{array}{l} u_{tt} = 3u_{xx} + 2u_{yy} + 4u_x + 3u_y, \quad 0 < x, y < \pi \text{ e } t > 0 \\ u(0, y, t) = u(\pi, y, t) = 0 \\ u(x, 0, t) = u(x, \pi, t) = 0 \\ u(x, y, 0) = 0 \\ u_t(x, y, 0) = \sin xy \end{array} \right.$$

Si cercano soluzioni del tipo $u(x, y, t) = X(x)Y(y)T(t)$. Sostituendo nell'equazione si trova $XYT'' = 3X''YT + 2XY''T + 4X'YT + 3XY'T$ da cui la seguente

$$\frac{T''}{T} = 3\frac{X''}{X} + 2\frac{Y''}{Y} + 4\frac{X'}{X} + 3\frac{Y'}{Y}$$

nella quale, per comodità, non sono state indicate le dipendenze di T da t , di X da x e di Y da y . Potendosi scrivere l'ultima equazione come

$$3\frac{X''}{X} + 4\frac{X'}{X} = \frac{T''}{T} - 2\frac{Y''}{Y} - 3\frac{Y'}{Y} = -\lambda,$$

si perviene al seguente problema spettrale

$$\left\{ \begin{array}{l} 3X'' + 4X' + \lambda X = 0 \\ X(0) = 0 \\ X(\pi) = 0 \end{array} \right. \quad (1.60)$$

Inoltre, poiché il rapporto $\frac{T''}{T}$ non dipende da y , si può scrivere

$$2\frac{Y''}{Y} + 3\frac{Y'}{Y} = \frac{T''}{T} + \lambda = -\mu,$$

si ottengono i problemi differenziali

$$\begin{cases} 2X'' + 3Y' + \mu Y = 0 \\ Y(0) = 0 \\ Y(\pi) = 0 \end{cases} \quad (1.61)$$

e

$$\begin{cases} T'' + (\lambda + \mu)T = 0 \\ T(0) = 0 \end{cases} \quad (1.62)$$

Possiamo brevemente discutere la soluzione di ciascuno di questi tre problemi. Cominciamo con il problema (1.60). Si cercano soluzioni del tipo $X(x) \simeq e^{\alpha x}$ che conducono all'equazione caratteristica

$$3\alpha^2 + 4\alpha + \lambda = 0.$$

Questa equazione ha come soluzioni $\alpha_{1,2} = \frac{-2 \pm \sqrt{4 - 3\lambda}}{3}$. Occorre distinguere tre casi a seconda che sia $4 - 3\lambda = 0$, $4 - 3\lambda > 0$ oppure $4 - 3\lambda < 0$.

1. Sia $\lambda = \frac{4}{3}$.

In tal caso abbiamo due radici reali e coincidenti $\alpha_1 = \alpha_2 = -\frac{2}{3}$.

La soluzione generale é $X(x) = e^{-\frac{2}{3}x}(c_1 + c_2 x)$. Poiché $X(0) = c_1$ e $X(\pi) = e^{-\frac{2}{3}\pi}(c_1 + c_2\pi)$, le costanti c_1 e c_2 devono soddisfare il sistema

$$\begin{cases} c_1 = 0 \\ c_1 + c_2\pi = 0 \end{cases}$$

che ammette come unica soluzione quella banale. Quindi tale caso non è da prendere in considerazione in quanto conduce alla soluzione $X(x) \equiv 0$.

2. Sia $\lambda < \frac{4}{3}$ e, per semplicità, poniamo $\frac{\sqrt{4-3\lambda}}{3} = \beta$.
 In tal caso abbiamo due radici reali e distinte $\alpha_{1,2} = \frac{-2}{3} \pm \beta$ e la soluzione generale é

$$X(x) = e^{-\frac{2}{3}x}(c_1 e^{\beta x} + c_2 e^{-\beta x}).$$

Poiché $X(0) = c_1 + c_2$ e $X(\pi) = e^{-\frac{2}{3}\pi}(c_1 e^{\beta\pi} + c_2 e^{-\beta\pi})$, le costanti c_1 e c_2 sono da determinare attraverso il sistema di equazioni lineari e omogenee

$$\begin{cases} c_1 + c_2 = 0 \\ c_1 e^{\beta\pi} + c_2 e^{-\beta\pi} = 0 \end{cases}$$

che si verifica facilmente possedere come unica soluzione quella banale. Quindi anche questo caso è da scartare.

3. Sia $\lambda > \frac{4}{3}$ e poniamo $\frac{\sqrt{3\lambda-4}}{3} = \beta$.
 In tal caso abbiamo due radici complesse coniugate $\alpha_{1,2} = \frac{-2 \pm i\beta}{3}$ e la soluzione generale è

$$X(x) = e^{-\frac{2}{3}x} (c_1 \cos \beta x + c_2 \sin \beta x).$$

Poiché $X(0) = c_1$ e $X(\pi) = e^{-\frac{2}{3}\pi}(c_1 \cos \beta\pi + c_2 \sin \beta\pi)$, le costanti c_1 e c_2 sono da determinare attraverso il sistema

$$\begin{cases} c_1 = 0 \\ c_1 \cos \beta\pi + c_2 \sin \beta\pi = 0. \end{cases}$$

Essendo interessati a soluzioni non banali del sistema non consideriamo $c_2 = 0$. Quindi l'equazione di sopra è soddisfatta quando

$$\sin \beta\pi = 0.$$

Questa equazione ammette come soluzioni $\beta\pi = k\pi$, $k = 1, 2, \dots$ e quindi $\lambda_k = \frac{4}{3} + 3k^2$. Lo spettro per il problema (1.60) è pertanto

$$\left\{ \lambda_k = \frac{4}{3} + 3k^2, X_k(x) = e^{-\frac{2}{3}x} \sin kx, k = 1, 2, \dots \right\}.$$

Discutiamo ora il problema (1.61). Si cercano soluzioni del tipo $X(x) \simeq e^{\alpha x}$ che conducono all'equazione caratteristica

$$2\alpha^2 + 3\alpha + \mu = 0.$$

Questa equazione ha come soluzioni $\alpha_{1,2} = \frac{-3 \pm \sqrt{9 - 8\mu}}{4}$. Occorre distinguere tre casi a seconda che sia $9 - 8\mu = 0$, $9 - 8\mu > 0$ oppure $9 - 8\mu < 0$.

1. Sia $\mu = \frac{9}{8}$.

In tal caso abbiamo due radici reali e coincidenti $\alpha_1 = \alpha_2 = -\frac{3}{4}$.

La soluzione generale è $Y(y) = e^{-\frac{3}{4}y}(c_1 + c_2 y)$. Poiché $Y(0) = c_1$ e $Y(\pi) = e^{-\frac{3}{4}\pi}(c_1 + c_2\pi)$, le costanti c_1 e c_2 devono soddisfare il sistema

$$\begin{cases} c_1 = 0 \\ c_1 + c_2\pi = 0 \end{cases}$$

che ammette come unica soluzione quella banale. Quindi tale caso non è da prendere in considerazione in quanto conduce alla soluzione $Y(y) \equiv 0$.

In maniera analoga si trova che anche il caso $9 - 8\mu > 0$ è da scartare in quanto conduce alla soluzione $Y(y) \equiv 0$.

2. Sia $9 - 8\mu < 0$ e poniamo $\frac{\sqrt{8\mu-9}}{4} = \beta$.

In tal caso abbiamo due radici complesse coniugate $\alpha_{1,2} = \frac{-3}{4} \pm i\beta$ e la soluzione generale é

$$Y(y) = e^{-\frac{3}{4}y} (c_1 \cos \beta y + c_2 \sin \beta y).$$

Poiché $Y(0) = c_1$ e $Y(\pi) = e^{-\frac{3}{4}\pi}(c_1 \cos \beta\pi + c_2 \sin \beta\pi)$, le costanti c_1 e c_2 sono da determinare attraverso il sistema

$$\begin{cases} c_1 = 0 \\ c_1 \cos \beta\pi + c_2 \sin \beta\pi = 0. \end{cases}$$

Essendo interessati a soluzioni non banali del sistema non consideriamo $c_2 = 0$. Quindi l'equazione di sopra è soddisfatta quando

$$\sin \beta\pi = 0.$$

Questa equazione ammette come soluzioni $\beta\pi = k\pi$, $k = 1, 2, \dots$ e quindi $\mu_n = \frac{9}{8} + 2n^2$. Lo spettro per il problema (1.60) è pertanto

$$\left\{ \mu_n = \frac{9}{8} + 2n^2, Y_n(x) = e^{-\frac{3}{4}y} \sin ny, n = 1, 2, \dots \right\}.$$

Resta da risolvere il problema (1.62). Esso è del tipo

$$\begin{cases} T''_{nk} + (\lambda_k + \mu_n)T_{nk} = 0 \\ T_{nk}(0) = 0, \quad k, n = 1, 2, \dots \end{cases}$$

Poiché $\lambda_k + \mu_n > 0$, $T_{nk} = a_{nk} \cos \sqrt{\lambda_k + \mu_n}t + b_{nk} \sin \sqrt{\lambda_k + \mu_n}t$ e dalla condizione $T_{nk}(0) = 0$ si ricava $a_{nk} = 0$. Lo spettro è quindi

$$\left\{ \mu_n + \lambda_k, T_{nk}(t) = \sin \sqrt{\lambda_k + \mu_n}t, n, k = 1, 2, \dots \right\}.$$

Quindi $u_{kn}(x, y, t) = (e^{-\frac{2}{3}x} \sin kx)(e^{-\frac{3}{4}y} \sin ny)(\sin \sqrt{\lambda_k + \mu_n}t)$ e la soluzione del problema principale sarà

$$u(x, y, t) = \sum_{k, n=1}^{\infty} c_{kn} e^{-(\frac{2}{3}x + \frac{3}{4}y)} \sin kx \sin ny \sin \sqrt{\lambda_k + \mu_n}t.$$

Imponendo la condizione $u_t(x, y, 0) = \sin xy$ si trova

$$\sum_{k, n=1}^{\infty} \sqrt{\lambda_k + \mu_n} c_{kn} e^{-(\frac{2}{3}x + \frac{3}{4}y)} \sin kx \sin ny = \sin xy.$$

Imponendo l'ortogonalità delle $X_k(x)$ rispetto al peso $e^{\frac{4}{3}x}$ si ottiene la relazione $\sum_{n=1}^{\infty} \sqrt{\lambda_k + \mu_n} c_{kn} \left(\int_0^{\pi} \sin^2 kx dx \right) e^{-\frac{3}{4}y} \sin ny = \int_0^{\pi} e^{\frac{2}{3}x} \sin kx \sin xy dx$. Infine imponendo l'ortogonalità delle $Y_n(y)$ rispetto al peso $e^{\frac{3}{2}x}$ si trova

$$\begin{aligned} & \sqrt{\lambda_k + \mu_n} c_{kn} \left(\int_0^{\pi} \sin kx^2 dx \right) \left(\int_0^{\pi} \sin ny^2 dy \right) = \\ & \int_0^{\pi} e^{\frac{3}{4}x} \sin ny \int_0^{\pi} e^{\frac{2}{3}x} \sin kx \sin xy dx dy. \end{aligned}$$

Dall'ultima equazione ricaviamo quindi

$$\frac{1}{4} \sqrt{\lambda_k + \mu_n} c_{kn} = \int_0^{\pi} e^{\frac{3}{4}x} \sin ny \int_0^{\pi} e^{\frac{2}{3}x} \sin kx \sin xy dx dy, \text{ per } k, n = 1, 2, \dots$$

Capitolo 2

Metodi alle differenze finite

Questo capitolo è dedicato ai metodi alle differenze finite, una famiglia di tecniche numeriche per la soluzione di equazioni differenziali alle derivate ordinarie e parziali molto nota e consolidata.

Principali vantaggi:

- facilità di implementazione ;
- buona efficienza computazionale .

Principale difficoltà:

- scarsa capacità di risoluzione in domini con geometria irregolare .

L'ultimo punto è direttamente collegato all'utilizzo di griglie di calcolo *strutturate* da parte dei metodi alle differenze finite. Queste ultime sono griglie i cui nodi possono essere messi in corrispondenza 1-to-1 con una matrice (una matrice 3D per applicazioni tridimensionali). In altre parole, si tratta di griglie i cui nodi sono identificati in modo univoco da una coppia (terna) di indici locali, e quindi possono al più ottenersi da griglie regolari per deformazione continua.

I metodi ad elementi finiti hanno invece la capacità di utilizzare griglie *non strutturate*, in cui i nodi sono liberi di disporsi senza nessun vincolo topologico: questa caratteristica conferisce una flessibilità d'uso notevolmente superiore rispetto alle differenze finite. Un esempio di griglia strutturata e uno di griglia non strutturata per una geometria semplice (un quadrato) sono mostrate nella figura 2.1.

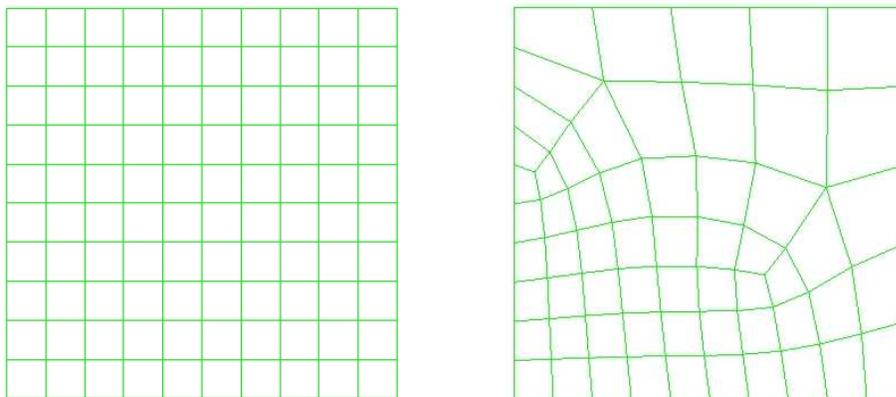


Figura 2.1: Un esempio di griglia strutturata e uno di griglia non strutturata per una semplice geometria di riferimento.

Problema differenziale con valori agli estremi

Risolviamo numericamente il problema differenziale

$$\begin{cases} y''(x) = p(x) y'(x) + q(x) y + r(x) \\ p, q \text{ e } r \text{ funzioni continue in } [a, b] \\ y(a) = \alpha, \quad y(b) = \beta. \end{cases}$$

Supponiamo, per semplicità, di discretizzare l'intervallo $[a, b]$ con punti equidistanti, ponendo pertanto $x_i = a + ih$, dove $i = 0, 1, \dots, n+1$ e $h = \frac{b-a}{n+1}$. Indichiamo inoltre con y_i il valore in x_i , $i = 0, 1, \dots, n+1$, della soluzione del modello discretizzato. Lo schema di discretizzazione è dunque valido se $\max_{i=1, \dots, n} |y_i - y(x_i)|$ è “sufficientemente piccolo”.

Schema di discretizzazione alle differenze centrali del 2° ordine:

$$\begin{aligned} y''(x_i) &\approx \frac{y_{i-1} - 2y_i + y_{i+1}}{h^2}, & y''(x_i) - \frac{y_{i-1} - 2y_i + y_{i+1}}{h^2} &= \mathcal{O}(h^2); \\ y'(x_i) &\approx \frac{y_{i+1} - y_{i-1}}{2h}, & y'(x_i) - \frac{y_{i+1} - y_{i-1}}{2h} &= \mathcal{O}(h^2). \end{aligned}$$

La collocazione in x_i , $i = 1, \dots, n$, dell'equazione differenziale genera il

Theorem 2.0.1 (Gerschgorin) *Nelle ipotesi precedenti, l'errore sulla soluzione possiede una maggiorazione del tipo:*

$$|y(x_i) - y_i| \leq c\tau(h), \quad i = 1, 2, \dots, n,$$

essendo c una costante positiva indipendente da h e $\tau(h) = \max_{a \leq x \leq b} |\tau(x, h)| = \mathcal{O}(h^2)$.

2.0.1 Equazioni ellittiche

Consideriamo ora il problema differenziale ¹

$$\begin{cases} u_{xx} + u_{yy} + p(x, y)u_x + q(x, y)u_y + r(x, y)u + s(x, y) = 0, & (x, y) \in \Omega \\ u(x, y) = f(x, y), & (x, y) \in \partial\Omega, \end{cases}$$

dove $\Omega = [a, b] \times [c, d]$.

Considerando, per semplicità, una reticolazione con nodi equidistanti in ciascuno dei due intervalli $[a, b]$ e $[c, d]$, si ottengono i nodi $\{(x_i, y_j)\}$, essendo

$$x_i = a + ih, \quad i = 0, 1, \dots, n+1, \quad \text{con } h = \frac{b-a}{n+1};$$

$$y_j = c + jk, \quad j = 0, 1, \dots, m+1, \quad \text{con } k = \frac{d-c}{m+1}.$$

La discretizzazione con il metodo alle differenze centrali (schema a 5 punti) e la collocazione dell'equazione differenziale nei punti interni generano il sistema:

$$\begin{aligned} \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{k^2} + p_{i,j} \frac{u_{i+1,j} - u_{i-1,j}}{2h} \\ + q_{i,j} \frac{u_{i,j+1} - u_{i,j-1}}{2k} + r_{i,j}u_{i,j} + s_{i,j} = 0 \end{aligned}$$

per $i = 1, \dots, n$ e $j = 1, \dots, m$.

¹Nel caso di geometria più complessa, è talvolta possibile adottare una efficace strategia di adattamento delle condizioni al contorno nei nodi "prossimi" alla frontiera.

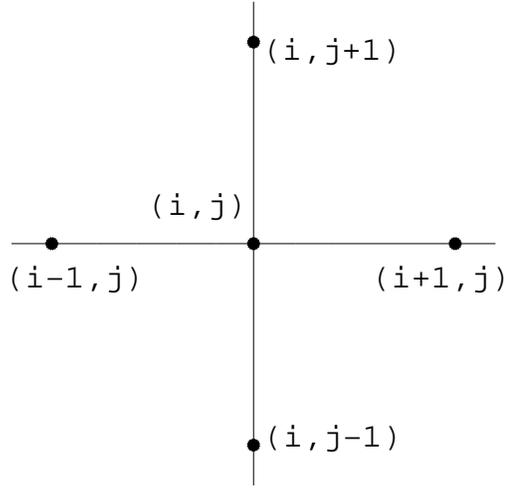


Figura 2.2: Schema di discretizzazione a 5 punti.

Ordinando gli $\{u_{i,j}\}_{i=1,\dots,n, j=1,\dots,m}$ per linee (per j crescente e, a parità di j , per i crescente), la discretizzazione precedente genera il seguente sistema lineare

$$h^2(2 - k q_{i,j})u_{i,j-1} + k^2(2 - h p_{i,j})u_{i-1,j} - 2 [2(h^2 + k^2) - h^2 k^2 r_{i,j}] u_{i,j} \\ + k^2(2 + h p_{i,j})u_{i+1,j} + h^2(2 + k q_{i,j})u_{i,j+1} = -2h^2 k^2 s_{ij}$$

per $i = 1, \dots, n$ e $j = 1, \dots, m$.

Theorem 2.0.2 *Se $r(x, y) \leq 0$ in Ω ed inoltre $ph \leq 2$ e $qk \leq 2$, essendo $|p(x, y)| \leq p$ e $|q(x, y)| \leq q$ per $(x, y) \in \Omega$, il sistema ottenuto con il metodo alle differenze centrali possiede una e una sola soluzione.*

In questo caso, se p , q , r , e s sono funzioni continue in Ω , il residuo differenziale

$$\tau(x, y; h, k) = \mathcal{O}(h^2 + k^2)$$

e inoltre l'errore sulla soluzione, in ogni punto nodale, soddisfa la condizione

$$|u(x_i, y_j) - u_{i,j}| = \mathcal{O}(h^2 + k^2).$$

Metodo upwind Qualora le condizioni $ph \leq 2$ e $qk \leq 2$ del Teorema 2.0.2 risultino molto restrittive, nel senso che h e/o k siano molto piccoli e dunque n e/o m molto grandi, onde evitare la risoluzione di un sistema di grandi dimensioni spesso si ricorre al cosiddetto metodo upwind.

In tal caso, mentre u_{xx} e u_{yy} vengono discretizzate con uno schema alle differenze centrali, u_x e u_y vengono discretizzate nel modo seguente:

$$\begin{aligned} u_x(x_i, y_j) &\approx \frac{u_{i+1,j} - u_{i,j}}{h}, & \text{se } p_{ij} \geq 0 \\ u_x(x_i, y_j) &\approx \frac{u_{i,j} - u_{i-1,j}}{h}, & \text{se } p_{ij} \leq 0 \\ u_y(x_i, y_j) &\approx \frac{u_{i,j+1} - u_{i,j}}{k}, & \text{se } q_{ij} \geq 0 \\ u_y(x_i, y_j) &\approx \frac{u_{i,j} - u_{i,j-1}}{k}, & \text{se } q_{ij} \leq 0. \end{aligned}$$

Questo equivale a porre

$$\begin{aligned} p(x_i, y_j)u_x &\approx \frac{(|p_{ij}| + p_{ij})u_{i+1,j} - 2|p_{ij}|u_{ij} + (|p_{ij}| - p_{ij})u_{i-1,j}}{2h} \\ q(x_i, y_j)u_y &\approx \frac{(|q_{ij}| + q_{ij})u_{i,j+1} - 2|q_{ij}|u_{ij} + (|q_{ij}| - q_{ij})u_{i,j-1}}{2k}. \end{aligned}$$

Tale schema genera il sistema

$$\begin{aligned} &\left[\frac{1}{k^2} + \frac{1}{2k} (|q_{ij}| - q_{ij}) \right] u_{i,j-1} + \left[\frac{1}{h^2} + \frac{1}{2h} (|p_{ij}| - p_{ij}) \right] u_{i-1,j} + \\ &- \left[\frac{2}{h^2} + \frac{2}{k^2} + \frac{|p_{ij}|}{h} + \frac{|q_{ij}|}{k} - r_{ij} \right] u_{ij} + \left[\frac{1}{h^2} + \frac{1}{2h} (|p_{ij}| + p_{ij}) \right] u_{i+1,j} + \\ &+ \left[\frac{1}{k^2} + \frac{1}{2k} (|q_{ij}| + q_{ij}) \right] u_{i,j+1} = -s_{ij} \end{aligned}$$

per $i = 1, \dots, n$ e $j = 1, \dots, m$. Questo sistema è diagonalmente dominante, e dunque univocamente risolvibile per ogni coppia (h, k) , se $r(x, y) \leq 0$.

Questo fatto non implica tuttavia che h e k possano essere presi piccoli, ossia che si debba risolvere un sistema piccolo. Infatti si deve ricordare che, mentre l'errore di discretizzazione del metodo alle differenze centrali è un $\mathcal{O}(h^2 + k^2)$, l'errore di discretizzazione dovuto al metodo upwind è un $\mathcal{O}(h + k)$.

Nelle ipotesi precedenti, posto ad esempio $a = b = 0$ e $b = d = 10$, la scelta di $h = \frac{1}{100}$ e $k = \frac{1}{50}$, comporta la risoluzione di un sistema pentadigonale (diagonalmente dominante) di ordine $n \cdot m = 999 \cdot 499 \simeq 5 \cdot 10^5$.

2.0.2 Esercizi proposti

Discutere la risoluzione numerica, mediante il metodo alle differenze finite, del seguente problema differenziale:

$$\left\{ \begin{array}{l} (x+y)^2 u_{xx} + 5u_{yy} + (\sin \sqrt{x^2 + y^2})u_x + \sqrt{x^2 + y^2}u_y = x^2 + y^2, \\ 0 \leq x, y \leq 5 \\ u(0, y) = f_0(y), \quad u(5, y) = f_1(y) \\ u(x, 0) = g_0(x), \quad u(x, 5) = g_1(x). \end{array} \right.$$

L'equazione é di tipo **ellittico**.

Considerando una reticolazione con nodi equidistanti in ciascuno dei due intervalli $[0, 5]$, si ottengono i nodi $\{(x_i, y_j)\}$, essendo $x_i = ih$, $i = 0, 1, 2, \dots, n+1$ e $y_j = jk$, $j = 0, 1, 2, \dots, m+1$ con $h = \frac{5}{n+1}$ e $k = \frac{5}{m+1}$. Consideriamo inoltre le seguenti approssimazioni

$$u_{xx}(x_i, y_j) \simeq \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2}, \quad u_x(x_i, y_j) \simeq \frac{u_{i+1,j} - u_{i-1,j}}{2h}$$

$$u_{yy}(x_i, y_j) \simeq \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2}, \quad u_y(x_i, y_j) \simeq \frac{u_{i,j+1} - u_{i,j-1}}{2k}.$$

La discretizzazione con il metodo alle differenze centrali (**schema a 5 punti**) e la collocazione dell'equazione differenziale nei punti interni generano il sistema:

$$(x_i + y_j)^2 \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + 5 \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2} + (\sin \sqrt{x_i^2 + y_j^2}) \frac{u_{i+1,j} - u_{i-1,j}}{2h} + \sqrt{x_i^2 + y_j^2} \frac{u_{i,j+1} - u_{i,j-1}}{2k} = x_i^2 + y_j^2,$$

per $i = 1, \dots, n$ e $j = 1, \dots, m$.

Da notare che tale discretizzazione comporta un errore di discretizzazione $\tau = O(h^2 + k^2)$. Ordinando gli $\{u_{i,j}\}_{i=1,\dots,n, j=1,\dots,m}$ per linee (per j crescente e, a

parità di j , per i crescente), la discretizzazione precedente genera il seguente sistema lineare:

$$\begin{aligned} & \left(\frac{5}{k^2} - \frac{\sqrt{x_i^2 + y_j^2}}{2k} \right) u_{i,j-1} + \left(\frac{(x_i + y_j)^2}{h^2} - \frac{\sin \sqrt{x_i^2 + y_j^2}}{2h} \right) u_{i-1,j} \\ & - \left(\frac{2(x_i + y_j)^2}{h^2} + \frac{10}{k^2} \right) u_{i,j} + \left(\frac{(x_i + y_j)^2}{h^2} + \frac{\sin \sqrt{x_i^2 + y_j^2}}{2h} \right) u_{i+1,j} \\ & + \left(\frac{5}{k^2} + \frac{\sqrt{x_i^2 + y_j^2}}{2k} \right) u_{i,j+1} = x_i^2 + y_j^2, \end{aligned}$$

per $i = 1, \dots, n$ e $j = 1, \dots, m$. Nel caso in cui le due condizioni

$$\begin{cases} \frac{5}{k^2} \geq \frac{\sqrt{x_i^2 + y_j^2}}{2k} \\ \frac{|\sin \sqrt{x_i^2 + y_j^2}|}{2} \leq \frac{(x_i + y_j)^2}{h} \end{cases}$$

siano simultaneamente soddisfatte, il sistema é debolmente diagonalmente dominante (e quindi ammette un'unica soluzione) in quanto per $i = 2, \dots, n-1$ e $j = 2, \dots, m-1$

$$\frac{2(x_i + y_j)^2}{h^2} + \frac{10}{k^2} \geq \frac{10}{k^2} + \frac{2(x_i + y_j)^2}{h^2}.$$

Quest'ultima é ovviamente verificata con il segno uguale a patto che le due precedenti condizioni lo siano. La prima condizione impone che $k \leq \sqrt{2}$ (i calcoli sono molto semplici) mentre la seconda é sempre verificata in quanto

$$\frac{|\sin \sqrt{x_i^2 + y_j^2}|}{2} \leq \frac{\sqrt{x_i^2 + y_j^2}}{2} \leq \frac{x_i + y_j}{2} \leq \frac{(x_i + y_j)^2}{h},$$

e dall'ultima di queste disequazioni si ricava $h \leq 2(x_i + y_j) \leq 20(h + k)$ che ovviamente é valida per ogni h e k .

Per $i = j = 1$ poiché i termini $u_{1,0}$ e $u_{0,1}$ sono termini noti, il sistema é diagonalmente dominante se

$$\frac{2(h + k)^2}{h^2} + \frac{10}{k^2} \geq \frac{5}{k^2} + \frac{\sqrt{h^2 + k^2}}{2k} + \frac{(h + k)^2}{h^2} + \frac{\sin \sqrt{h^2 + k^2}}{2h}.$$

Riscrivendo l'equazione come

$$\frac{(h+k)^2}{h^2} + \frac{5}{k^2} \geq \frac{\sqrt{h^2+k^2}}{2k} + \frac{\sin \sqrt{h^2+k^2}}{2h},$$

osserviamo che quest'ultima diseuguaglianza é soddisfatta in quanto:

1. $\frac{(h+k)^2}{h^2} \geq \frac{\sqrt{h^2+k^2}}{2h} \geq \frac{\sin \sqrt{h^2+k^2}}{2h}$, essendo $\frac{(h+k)^2}{h} > \frac{h+k}{2}$, ossia $\frac{h+k}{h} = 1 + \frac{k}{h} > \frac{1}{2}$, e inoltre
2. $\frac{5}{k} \geq \frac{\sqrt{h^2+k^2}}{2}$, in quanto $\frac{5}{k} \geq \frac{\sqrt{50}}{2}$, ossia $k \leq \sqrt{2}$.

Quindi se $k \leq \sqrt{2}$ il sistema é diagonalmente dominante.

2.0.3 Equazioni paraboliche

Consideriamo ora il problema differenziale

$$\begin{cases} u_t = c^2 u_{xx} + p(x,t)u_x + q(x,t)u + r(x,t), & a \leq x \leq b, 0 \leq t \leq T \\ u(a,t) = f_1(t), \quad u(b,t) = f_2(t), & T \geq t \geq 0 \\ u(x,0) = \phi(x), & a \leq x \leq b. \end{cases}$$

Discretizzazione con uno schema a 4 punti:

$$\begin{cases} x_i = a + ih, & i = 0, 1, \dots, n+1, & h = \frac{b-a}{n+1} \\ t_j = jk, & j = 0, 1, \dots, m+1, & k = \frac{T}{m+1}, \end{cases}$$

$$\frac{u_{i,j} - u_{i,j-1}}{k} = c^2 \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + p_{i,j} \frac{u_{i+1,j} - u_{i-1,j}}{2h} + q_{i,j}u_{i,j} + r_{i,j}$$

per $i = 1, \dots, n$ e $j = 1, \dots, m$.

Osservazione Per $j = 1$, le incognite sono $u_{1,1}, \dots, u_{n,1}$, le quali possono essere determinate risolvendo il sistema di ordine n ottenuto per $i = 1, \dots, n$, in quanto le $\{u_{i,0}\}$ sono note. Più in generale, supponendo noti i valori $u_{i,j-1}$

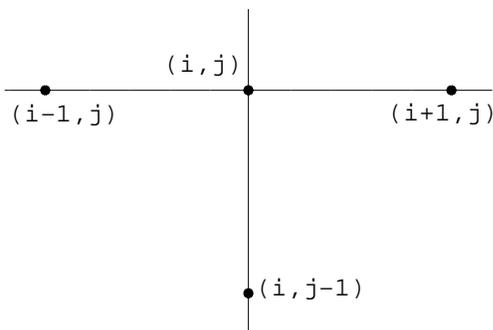


Figura 2.3: Schema di discretizzazione a 4 punti.

al livello $j - 1$, i valori $u_{i,j}$, $i = 1, \dots, n$, al livello j possono essere calcolati risolvendo il sistema di ordine n ottenuto per $i = 1, \dots, n$.

Di conseguenza poichè, assegnato j , il vettore $\{u_{i,j-1}\}$ è noto, lo schema precedente implica che l'avanzamento temporale richiede la risoluzione del sistema:

$$\begin{aligned} \frac{k}{h} \left(\frac{2c^2}{h} - p_{i,j} \right) u_{i-1,j} - 2 \left(1 + 2 \frac{c^2 k}{h^2} - k q_{i,j} \right) u_{i,j} \\ + \frac{k}{h} \left(\frac{2c^2}{h} + p_{i,j} \right) u_{i+1,j} = -2 (u_{i,j-1} + k r_{i,j}) \end{aligned}$$

per $i = 1, \dots, n$, prefissato $j = 1, \dots, m$.

Se $q(x, t) \leq 0$ per $a \leq x \leq b$ e $0 \leq t \leq T$, tale sistema è diagonalmente dominante per $hp \leq 2c^2$, essendo $p = \max_{\substack{a \leq x \leq b \\ 0 \leq t \leq T}} |p(x, t)|$. In tal caso, per ogni prefissato j , il sistema possiede una e una sola soluzione.

Osservazione Poichè $u_t(x_i, t_j)$ è stata approssimata con uno schema alle differenze del 1° ordine in t e uno del 2° in x , l'errore di discretizzazione è un $\mathcal{O}(h^2 + k)$.

Nel caso $a = 0$, $b = 20$ e $T = 10$, con la scelta di $h = \frac{1}{20}$, $k = \frac{1}{100}$, si debbono risolvere $m + 1 = 1000$ sistemi tridiagonali (diagonalmente dominanti), ciascuno di ordine $n = 399$.

2.0.4 Esercizi proposti

Discutere la risoluzione numerica, mediante il metodo alle differenze finite, del seguente problema differenziale:

$$\begin{cases} u_t = x^2 t^2 u_{xx} + 10^{2x+t} (\cos xt) u_x + (xt - 1)u, & 0 \leq x \leq 10, 0 \leq t \leq 5 \\ u(0, t) = f_0(t), \quad u(10, t) = f_1(t) \\ u(x, 0) = g(x). \end{cases}$$

L'equazione é di tipo **parabolico**. Occorre quindi ricorrere a una **discretizzazione con uno schema a 4 punti**:

$$\begin{cases} x_i = ih, & i = 0, 1, \dots, n+1, & h = \frac{b-a}{n+1} \\ t_j = jk, & j = 0, 1, \dots, m+1, & k = \frac{T}{m+1}, \end{cases}$$

Usando il **metodo up-wind** per il termine contenente u_x si dovrà tener conto delle seguenti approssimazioni

$$\begin{aligned} u_{xx}(x_i, t_j) &\simeq \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2}, \\ u_x(x_i, t_j) &\simeq \frac{u_{i+1,j} - u_{i,j}}{h}, \quad \text{se } \cos x_i t_j \geq 0 \\ u_x(x_i, t_j) &\simeq \frac{u_{i,j} - u_{i-1,j}}{h}, \quad \text{se } \cos x_i t_j \leq 0 \\ u_t(x_i, t_j) &\simeq \frac{u_{i,j} - u_{i,j-1}}{k} \end{aligned}$$

La discretizzazione con lo schema a 4 punti e la collocazione dell'equazione differenziale nei punti interni generano il sistema:

$$\begin{aligned} \frac{u_{i,j} - u_{i,j-1}}{k} &= (x_i t_j)^2 \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \\ &+ 10^{2x_i+t_j} \frac{(|\cos x_i t_j| + \cos x_i t_j)u_{i+1,j} - 2|\cos x_i t_j|u_{i,j} + (|\cos x_i t_j| - \cos x_i t_j)u_{i-1,j}}{2h} \\ &+ (x_i t_j - 1)u_{i,j} \end{aligned}$$

In questo caso l'errore di discretizzazione é $\tau = O(h + k)$. Per $j = 1$, le incognite sono $u_{1,1}, \dots, u_{n,1}$, le quali possono essere determinate risolvendo il sistema di ordine n ottenuto per $i = 1, \dots, n$, in quanto le $\{u_{i0}\}$ sono

note. Più in generale, supponendo noti i valori $u_{i,j-1}$ al livello $j - 1$, i valori $u_{i,j}$, $i = 1, \dots, n$, al livello j possono essere calcolati risolvendo il sistema di ordine n ottenuto per $i = 1, \dots, n$. Quindi assegnato j , il vettore $\{u_{i,j-1}\}$ è noto, e lo schema precedente implica che l'avanzamento temporale richiede la risoluzione del sistema:

$$\begin{aligned} & \left[\frac{(x_i t_j)^2}{h^2} + 10^{2x_i+t_j} \frac{|\cos x_i t_j| - \cos x_i t_j}{2h} \right] u_{i-1,j} \\ & - \left[\frac{1}{k} + \frac{2(x_i t_j)^2}{h^2} + 10^{2x_i+t_j} \frac{|\cos x_i t_j|}{h} + (1 - x_i t_j) \right] u_{i,j} \\ & + \left[\frac{(x_i t_j)^2}{h^2} + 10^{2x_i+t_j} \frac{|\cos x_i t_j| + \cos x_i t_j}{2h} \right] u_{i+1,j} = \frac{u_{i,j-1}}{k}, \end{aligned}$$

per $i = 1, \dots, n$, prefissato $j = 1, \dots, m$. Poiché

1. Per $\frac{1}{k} \geq x_i t_j - 1$, che implica $k \leq \frac{1}{49}$, il coefficiente del termine in $u_{i,j}$ è negativo,
2. Il sistema è debolmente diagonalmente dominante se

$$\frac{1}{k} + \frac{2(x_i t_j)^2}{h^2} + 10^{2x_i+t_j} \frac{|\cos x_i t_j|}{h} + (1 - x_i t_j) \geq \frac{2(x_i t_j)^2}{h^2} + 10^{2x_i+t_j} \frac{|\cos x_i t_j|}{h},$$

che implica $k \leq \frac{1}{49}$.

condizione sufficiente affinché il sistema possieda un'unica soluzione è dunque che sia $k \leq \frac{1}{49}$.

2.0.5 Equazioni iperboliche

Consideriamo ora il problema differenziale

$$\begin{cases} u_{tt} = c^2 u_{xx} + p(x, t)u_x + q(x, y)u_t + r(x, t)u + s(x, t), & a \leq x \leq b, t \geq 0 \\ u(a, t) = f_1(t), \quad u(b, t) = f_2(t), & 0 \leq t \leq T \\ u(x, 0) = \phi_1(x), \quad u_t(x, 0) = \phi_2(x), & a \leq x \leq b. \end{cases}$$

Discretizzazione con uno schema a 7 punti:

$$x_i = a + ih, \quad i = 0, 1, \dots, n + 1, \quad h = \frac{b - a}{n + 1}$$

$$t_j = jk, \quad j = 0, 1, \dots, m + 1, \quad k = \frac{T}{m + 1},$$

$$\begin{aligned} \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2} &= c^2 \frac{1}{2} \left[\frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{h^2} \right. \\ &\quad \left. + \frac{u_{i+1,j-1} - 2u_{i,j-1} + u_{i-1,j-1}}{h^2} \right] \\ &\quad + \frac{p_{i,j}}{2} \left[\frac{u_{i+1,j+1} - u_{i-1,j+1}}{2h} + \frac{u_{i+1,j-1} - u_{i-1,j-1}}{2h} \right] \\ &\quad + q_{i,j} \frac{u_{i,j+1} - u_{i,j-1}}{2k} + r_{i,j} u_{i,j} + s_{i,j} \end{aligned}$$

per $i = 1, \dots, n$ e $j = 1, \dots, m$.

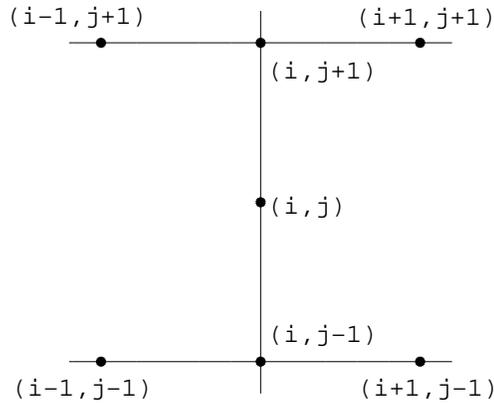


Figura 2.4: Schema di discretizzazione a 7 punti.

Se, in aggiunta agli $u_{i,0}$, fossero noti in valori $u_{i,1}$, lo schema consentirebbe di calcolare gli $u_{i,2}$, $i = 1, \dots, n$, risolvendo un sistema lineare di ordine n . Di conseguenza, iterando il processo, noti i valori in 2 livelli di j , gli n valori del livello superiore verrebbero calcolati risolvendo un semplice sistema lineare di ordine n .

Per il calcolo degli $u_{i,1}$, $i = 1, \dots, n$, viene generalmente utilizzata la formula di Taylor, troncata al termine di 2° ordine. Si pone cioè, tenendo conto delle condizioni iniziali,

$$\begin{aligned} u_{i,1} = u(x_i, t_1) &\approx u_{i,0} + ku_t(x_i, 0) + \frac{k^2}{2}u_{tt}(x_i, 0) \\ &= \phi_1(x_i) + k\phi_2(x_i) + \frac{k^2}{2}u_{tt}(x_i, 0), \end{aligned}$$

da cui, utilizzando l'equazione stessa,

$$\begin{aligned} u_{i1} &\approx \phi_1(x_i) + k\phi_2(x_i) \\ &+ \frac{k^2}{2} [c^2 u_{xx}(x_i, 0) + p_{i0}u_x(x_i, 0) + q_{i0}\phi_2(x_i) + r_{i0}u_{i0} + s_{i0}]. \end{aligned}$$

I valori incogniti di quest'ultima formula, ossia $u_{xx}(x_i, 0)$ e $u_x(x_i, 0)$, possono essere infine valutati mediante lo schema alle differenze centrali

$$\begin{aligned} u_{xx}(x_i, 0) &\approx \frac{u_{i+1,0} - 2u_{i,0} + u_{i-1,0}}{h^2} = \frac{\phi_1(x_{i+1}) - 2\phi_1(x_i) + \phi_1(x_{i-1}))}{h^2} \\ u_x(x_i, 0) &\approx \frac{u_{i+1,0} - u_{i-1,0}}{2h} = \frac{\phi_1(x_{i+1}) - \phi_1(x_{i-1}))}{2h}. \end{aligned}$$

Da notare che procedendo in tal modo, se le funzioni assegnate sono tutte continue, l'errore di discretizzazione è un $\mathcal{O}(h^2 + k^2)$.

Riordinando di conseguenza, ossia tenendo conto del fatto che gli unici valori incogniti sono $u_{i-1,j+1}$, $u_{i,j+1}$ e $u_{i,j+1}$, si ottiene il sistema

$$\begin{aligned} &\frac{1}{2h} \left(\frac{c^2}{h} - \frac{p_{ij}}{2} \right) u_{i-1,j+1} - \left(\frac{1}{k^2} + \frac{c^2}{h^2} - \frac{q_{ij}}{2k} \right) u_{i,j+1} + \frac{1}{2h} \left(\frac{c^2}{h} + \frac{p_{ij}}{2} \right) u_{i+1,j+1} \\ &= \frac{1}{2h} \left(\frac{p_{ij}}{2} - \frac{c^2}{h} \right) u_{i-1,j-1} + \left(\frac{1}{k^2} + \frac{c^2}{h^2} + \frac{q_{ij}}{2k} \right) u_{i,j-1} \\ &- \frac{1}{2h} \left(\frac{p_{ij}}{2} + \frac{c^2}{h} \right) u_{i+1,j-1} - \left(r_{ij} + \frac{2}{k^2} \right) u_{i,j} - s_{ij}, \end{aligned}$$

dove $i = 1, \dots, n$ per ogni prefissato $j = 1, \dots, m$.

Per calcolare $u_{i,j}$ in tutti i punti nodali si debbono dunque risolvere m sistemi lineari, ciascuno di ordine n . Tali sistemi, se $q(x, t) \leq 0$, sono tutti diagonalmente dominanti purchè risulti $hp \leq 2c^2$, essendo $p = \max_{\substack{a \leq x \leq b \\ 0 \leq t \leq T}} |p(x, t)|$.

Inoltre l'errore $y(x_i, t_j) - y_{ij} = \mathcal{O}(h^2 + k^2)$ in quanto la discretizzazione in x e t è basata sulle differenze centrali.

2.0.6 Esercizi proposti

Discutere la risoluzione numerica del seguente problema differenziale:

$$\left\{ \begin{array}{l} (x^2 + 1)u_{tt} = (t^2 + 2)u_{xx} + 3^{x^2+t^2}u_x + (\cos t)u_y + (x + t)^2u, \\ 0 < x < 10, \quad 0 < t < 5 \\ u(0, t) = f_1(t), \quad u(10, t) = f_2(t) \\ u(x, 0) = g_1(x), \quad u_t(x, 0) = g_2(x). \end{array} \right.$$

L'equazione é di tipo **iperbolico**. Occorre quindi ricorrere a una **discretizzazione con uno schema a 7 punti**:

$$\begin{aligned} x_i &= ih, \quad i = 0, 1, \dots, n + 1, & h &= \frac{10}{n + 1} \\ t_j &= jk, \quad j = 0, 1, \dots, m + 1, & k &= \frac{5}{m + 1}, \end{aligned}$$

Usando il **metodo up-wind** per il termine contenente u_x (ossevando che $3^{x^2+t^2} > 0$) si dovrà tener conto delle seguenti approssimazioni

$$\begin{aligned} u_{xx}(x_i, t_j) &\simeq \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{2h^2} + \frac{u_{i+1,j-1} - 2u_{i,j-1} + u_{i-1,j-1}}{2h^2}, \\ u_x(x_i, t_j) &\simeq \frac{u_{i+1,j+1} - u_{i,j+1}}{2h} + \frac{u_{i+1,j-1} - u_{i,j-1}}{2h}, \\ u_t(x_i, t_j) &\simeq \frac{u_{i,j+1} - u_{i,j-1}}{2k}, \\ u_{tt}(x_i, t_j) &\simeq \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2}. \end{aligned}$$

La discretizzazione con lo schema a 7 punti e la collocazione dell'equazione differenziale nei punti interni generano il sistema:

$$\begin{aligned} (x_i^2 + 1) \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2} = \\ (t_j^2 + 2) \frac{1}{2} \left[\frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{h^2} + \frac{u_{i+1,j-1} - 2u_{i,j-1} + u_{i-1,j-1}}{h^2} \right] \\ + 3^{x_i^2+t_j^2} \frac{1}{2} \left[\frac{u_{i+1,j+1} - u_{i,j+1}}{h} + \frac{u_{i+1,j-1} - u_{i,j-1}}{h} \right] + (\cos t_j) \frac{u_{i,j+1} - u_{i,j-1}}{2k} \\ + (x_i + t_j)^2 u_{ij} \end{aligned}$$

In questo caso l'errore di discretizzazione é $\tau = O(h + k^2)$. Riordinando l'equazione precedente, si vede che essa può scriversi come

$$\begin{aligned} & \left(\frac{t_j^2 + 2}{2h^2} \right) u_{i-1,j+1} - \left(\frac{x_i^2 + 1}{k^2} + \frac{t_j^2 + 2}{h^2} + \frac{3^{x_i^2+t_j^2}}{2h} - \frac{\cos t_j}{2k} \right) u_{i,j+1} \\ & + \left(\frac{t_j^2 + 2}{2h^2} + \frac{3^{x_i^2+t_j^2}}{2h} \right) u_{i+1,j+1} = \phi(u_{i-1,j-1}, u_{i,j-1}, u_{i+1,j-1}, u_{i,j}). \end{aligned}$$

Tale sistema risulta essere diagonalmente dominante purché $\frac{x_i^2 + 1}{k^2} > \frac{|\cos t_j|}{2k}$. Questa condizione é soddisfatta se $k < 1$.

Prendendo ad esempio $k = \frac{1}{50}$ e $h = \frac{1}{100}$, si debbono risolvere 250 sistemi tridiagonali e diagonalmente dominanti ciascuno dell'ordine di 1000. Per la determinazione della soluzione al livello 1, ossia per $t = k$ e per ogni x_i , $i = 1, \dots, n$, si procede nel modo seguente:

$$\begin{aligned} u(x_i, k) &= u_{i1} = u(x_i, 0) + ku_t(x_i, 0) + \frac{k^2}{2}u_{tt}(x_i, 0) = g_1(x_i) + kg_2(x_i) \\ &+ \frac{k^2}{2} \left[2\frac{u_{i+1,0} - 2u_{i,0} + u_{i-1,0}}{h^2} + 3^{x_i^2}\frac{u_{i+1,0} - u_{i-1,0}}{2h} + g_1(x_i) + x_i^2g_2(x_i) \right] \\ &= g_1(x_i) + kg_2(x_i) + \frac{k^2}{2} \left[2\frac{g_1(x_{i+1}) - 2g_1(x_i) + g_1(x_{i-1}))}{h^2} \right. \\ &\left. + 3^{x_i^2}\frac{g_1(x_{i+1}) - g_1(x_{i-1}))}{2h} + g_1(x_i) + x_i^2g_2(x_i) \right], \quad i = 1, 2, \dots, n. \end{aligned}$$

In questo modo si calcola la $u(x_i, k)$ con un errore di discretizzazione dell'ordine di k^2 .

2.0.7 Modelli debolmente non lineari

Primo modello. Consideriamo il problema differenziale

$$\begin{cases} y''(x) + p(x)y'(x) + q(x, y) + r(x) = 0, & a \leq x \leq b \\ y(a) = \alpha, \quad y(b) = \beta. \end{cases}$$

Posto $x_i = a + ih$ e $y_i \approx y(x_i)$, per $i = 0, 1, \dots, n+1$ e $h = \frac{b-a}{n+1}$, lo schema alle differenze centrali trasforma il problema iniziale nel sistema non lineare

$$\begin{cases} \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + p_i \frac{y_{i+1} - y_{i-1}}{2h} + q(x_i, y_i) + r_i = 0, & i = 1, \dots, n \\ y_0 = \alpha, \quad y_{n+1} = \beta. \end{cases}$$

Questo sistema è dunque esprimibile nella forma

$$f_i(y_{i-1}, y_i, y_{i+1}) = 0, \quad i = 1, 2, \dots, n, \quad y_0 = \alpha, \quad y_{n+1} = \beta,$$

essendo

$$f_i(y_{i-1}, y_i, y_{i+1}) = \left(1 - \frac{h}{2}p_i\right)y_{i-1} - 2y_i + \left(1 + \frac{h}{2}p_i\right)y_{i+1} + q(x_i, y_i)h^2 + h^2r_i.$$

Condizione sufficiente affinché tale sistema sia univocamente risolubile è che sia $ph \leq 2$, essendo $p = \max_{a \leq x \leq b} |p(x)|$ e $\frac{\partial q}{\partial y} \leq 0$ per $a \leq x \leq b$ e y qualsiasi.

Da notare che se il sistema fosse lineare, ossia $q(x, y) = q(x)y$, questa condizione implicherebbe la dominanza diagonale del sistema, con $q(x) \leq 0$ per $a \leq x \leq b$.

Nelle suddette ipotesi la soluzione può essere ottenuta mediante il metodo iterativo di Newton/Jacobi consistente nel calcolare, per ogni $k = 0, 1, \dots$, gli iterati

$$y_i^{(k+1)} = y_i^{(k)} - \frac{f_i(y_{i-1}^{(k)}, y_i^{(k)}, y_{i+1}^{(k)})}{\frac{\partial f_i}{\partial y_i}(y_{i-1}^{(k)}, y_i^{(k)}, y_{i+1}^{(k)})}, \quad i = 1, \dots, n,$$

dove

$$f_i(y_{i-1}^{(k)}, y_i^{(k)}, y_{i+1}^{(k)}) = \left[1 - \frac{h}{2}p_i\right]y_{i-1}^{(k)} - 2y_i^{(k)} + \left[1 + \frac{h}{2}p_i\right]y_{i+1}^{(k)} + q(x_i, y_i^{(k)})h^2 + h^2r_i,$$

e

$$\frac{\partial f_i}{\partial y_i}(y_{i-1}^{(k)}, y_i^{(k)}, y_{i+1}^{(k)}) = -2 + \frac{\partial q}{\partial y_i}(x_i, y_i^{(k)}).$$

Nelle citate ipotesi il metodo è globalmente convergente, dunque teoricamente indipendente dei valori $y_i^{(0)}$, $i = 1, \dots, n+1$. Poichè $y_1^{(0)} = \alpha$ e $y_{n+1}^{(0)} = \beta$, la scelta più frequente consiste nell'assumere

$$y_i^{(0)} = \alpha + \frac{i}{n+1}(\beta - \alpha), \quad i = 1, \dots, n+1,$$

ossia nell'interpolare linearmente i valori α e β .

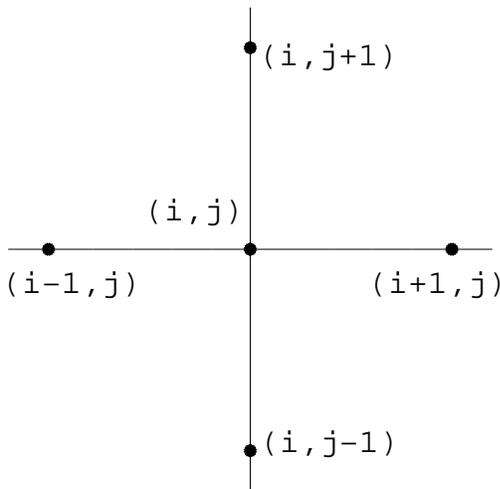
Secondo modello. Consideriamo ora il seguente problema di tipo ellittico

$$\begin{cases} u_{xx} + u_{yy} + p(x, y)u_x + q(x, y)u_y + r(x, y, u) + s(x, y) = 0, & (x, y) \in \Omega \\ u(x, y) = f(x, y), & (x, y) \in \partial\Omega, \end{cases}$$

dove $\Omega \subset [a, b] \times [c, d]$.

Posto

$$\begin{cases} x_i = a + ih, & i = 0, 1, \dots, n+1, & h = \frac{b-a}{n+1} \\ y_j = c + jk, & j = 0, 1, \dots, m+1, & k = \frac{d-c}{m+1} \end{cases}$$



la discretizzazione con il metodo alle differenze centrali genera il sistema debolmente non lineare

$$\begin{aligned} \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2} + p_{i,j} \frac{u_{i+1,j} - u_{i-1,j}}{2h} \\ + q_{i,j} \frac{u_{i,j+1} - u_{i,j-1}}{2k} + r(x_i, y_j, u_{i,j}) + s(x_i, y_j) = 0 \end{aligned}$$

per $i = 1, \dots, n$ e $j = 1, \dots, m$. Tale sistema può essere espresso nella forma

$$\phi_{i,j}(u_{i,j-1}, u_{i-1,j}, u_{i,j}, u_{i+1,j}, u_{i,j+1}) = 0, \quad i = 1, \dots, n, \quad j = 1, \dots, m,$$

essendo

$$\begin{aligned} \phi_{i,j}(u_{i,j-1}, u_{i-1,j}, u_{i,j}, u_{i+1,j}, u_{i,j+1}) = h^2(2 - kq_{ij})u_{i,j-1} + k^2(2 - hp_{ij})u_{i-1,j} \\ - 4(h^2 + k^2)u_{i,j} + k^2(2 + hp_{ij})u_{i+1,j} \\ + h^2(2 + kq_{ij})u_{i,j+1} + 2h^2k^2[r(x_i, y_j, u_{i,j}) + s(x_i, y_j)]. \end{aligned}$$

e con $u_{i,0}$ e $u_{0,j}$ noti per $i = 1, \dots, n$ e $j = 1, \dots, m$.

Condizione sufficiente affinché questo sistema sia univocamente risolubile, nell'ipotesi che sia $\frac{\partial r(x, y, u)}{\partial u} < 0$ per $(x, y) \in \Omega$ e u qualsiasi, è che siano $Ph \leq 2$ e $Qk \leq 2$, essendo $P = \max |p(x, y)|$ e $Q = \max |q(x, y)|$ per $(x, y) \in \Omega$.

In caso di linearità di s rispetto ad u , le condizioni precedenti assicurano che la matrice del sistema sia diagonalmente dominante.

Nelle suddette ipotesi il metodo di Newton/Jacobi applicato al sistema risulta globalmente convergente, per cui, prefissato $l = 0, 1, \dots$, la soluzione può essere ottenuta mediante il processo iterativo

$$u_{i,j}^{(l+1)} = u_{i,j}^{(l)} - \frac{\phi_{i,j}(u_{i,j-1}^{(l)}, u_{i-1,j}^{(l)}, u_{i,j}^{(l)}, u_{i+1,j}^{(l)}, u_{i,j+1}^{(l)})}{\frac{\partial \phi_{i,j}}{\partial u_{i,j}}(u_{i,j-1}^{(l)}, u_{i-1,j}^{(l)}, u_{i,j}^{(l)}, u_{i+1,j}^{(l)}, u_{i,j+1}^{(l)})},$$

per $i = 1, \dots, n$ e $j = 1, \dots, m$, essendo

$$\frac{\partial \phi_{i,j}}{\partial u_{i,j}} = -4(h^2 + k^2) + 2h^2k^2 \frac{\partial r}{\partial u_{i,j}}.$$

Poichè la u è nota al contorno, i valori $u_{ij}^{(0)}$ sono generalmente ottenuti per interpolazione rispetto ai valori assunti sul contorno.

Un punto cruciale del procedimento é quindi relativo alla scelta del vettore d'innescio $u_{ij}^{(0)}$. Si vuole dare un'idea di come tale scelta possa essere fatta usando una particolare tecnica. A tal fine, iniziamo con esplicitare le condizioni al contorno in questo modo:

$$\begin{cases} u(a, y) = f_1(y), & u(b, y) = f_2(y) \\ u(x, c) = \varphi_1(x), & u(x, d) = \varphi_2(x). \end{cases}$$

La scelta del vettore d'innescio puó essere fatta ricorrendo alla funzione di interpolazione bidimensinale

$$F(x, y) = \frac{(x-a)^2(x-b)^2 f(x, y) + (c-y)^2(d-y)^2 \varphi(x, y)}{(x-a)^2(x-b)^2 + (c-y)^2(d-y)^2}.$$

Bisogna notare che

$$\begin{aligned} F(a, y) &= f(a, y) = f_1(y), \\ F(b, y) &= f(b, y) = f_2(y), \\ F(x, c) &= \varphi(x, c) = \varphi_1(x), \\ F(x, d) &= \varphi(x, d) = \varphi_2(x), \end{aligned}$$

cioè la funzione $F(x, y)$ coincide sul bordo con i valori assegnati dal problema. Le funzioni $f(x, y)$ e $\varphi(x, y)$ che compaiono nella definizione di $F(x, y)$ sono ottenute interpolando linearmente, rispettivamente, fra $f_1(y)$ e $f_2(y)$ e $\varphi_1(x)$ e $\varphi_2(x)$. In particolare,

$$f(x_i, y_j) = f_1(x_i, y_j) + \frac{x_i - a}{b - a} [f_2(b, y_j) - f_1(a, y_j)].$$

In tale formula, di volta in volta si fissa j ($j = 1, \dots, m$) e si calcolano le incognite al livello j -esimo facendo variare i . Analogamente definiamo

$$\varphi(x_i, y_j) = \varphi_1(x_i, y_j) + \frac{y_j - c}{d - c} [\varphi_2(x_i, d) - \varphi_1(x_i, c)],$$

e in tal caso si fissa i ($i = 1, \dots, n$) e si calcolano le incognite facendo variare j . Si puó quindi definire il nostro vettore d'innescio $u_{ij}^{(0)}$ come $u_{ij}^{(0)} = F(x_i, y_j)$.

2.0.8 Un'applicazione industriale delle differenze finite

Una delle più note applicazioni industriali delle differenze finite è la *seismic migration*, una metodologia per la ricostruzione della struttura geologica del sottosuolo, orientata alla localizzazione di giacimenti di acqua o idrocarburi. L'industria petrolifera ha sviluppato questa tecnologia a partire dagli anni '70. Camion speciali, dal peso di 40 tonnellate e dotati di piattaforme vibranti generano onde acustiche ed elastiche che si propagano nel sottosuolo, vengono riflesse da eventuali strutture geologiche, e tornano alla superficie dove vengono registrate da appositi ricevitori. I database così costituiti, di grandi dimensioni, vengono opportunamente trattati e utilizzati dagli analisti come strumento di supporto alle decisioni per l'identificazione di eventuali giacimenti. A seconda della profondità e del tipo di applicazione, possono essere usate anche altre sorgenti d'onda, come *air-gun* (cannoni ad aria compressa) o cariche esplosive.

La simulazione della propagazione delle onde acustiche/elastiche è particolarmente efficace con un metodo alle differenze finite, perché il dominio computazionale è estremamente regolare: una porzione di suolo a forma di parallelepipedo, in cui in generale anche la faccia superiore corrispondente alla superficie libera è piana, poiché si trascura la topografia di superficie. In figura 2.5 è mostrato un tipico modello 3D a differenze finite: i confini tra regioni con diverse velocità di propagazione delle onde, se obliqui, vengono discretizzati tramite una struttura a scaletta, con un rapido alternarsi di segmenti orizzontali e verticali; la conseguente perdita di accuratezza rappresenta un ben noto limite del metodo alle differenze finite.

La figura 2.6 mostra il risultato di una *seismic migration* relativa ad un sito nel Nevada (USA), utilizzato per la produzione di energia geotermica: è evidenziato un possibile giacimento di acqua calda intrappolata all'interno di una serie di faglie, ad una profondità di circa 4 km. La *seismic migration* introduce anche possibili effetti spuri, un esempio dei quali è rappresentato con la dicitura "false arcs".

(Immagini del Center for Economic Migration and Tomography.)

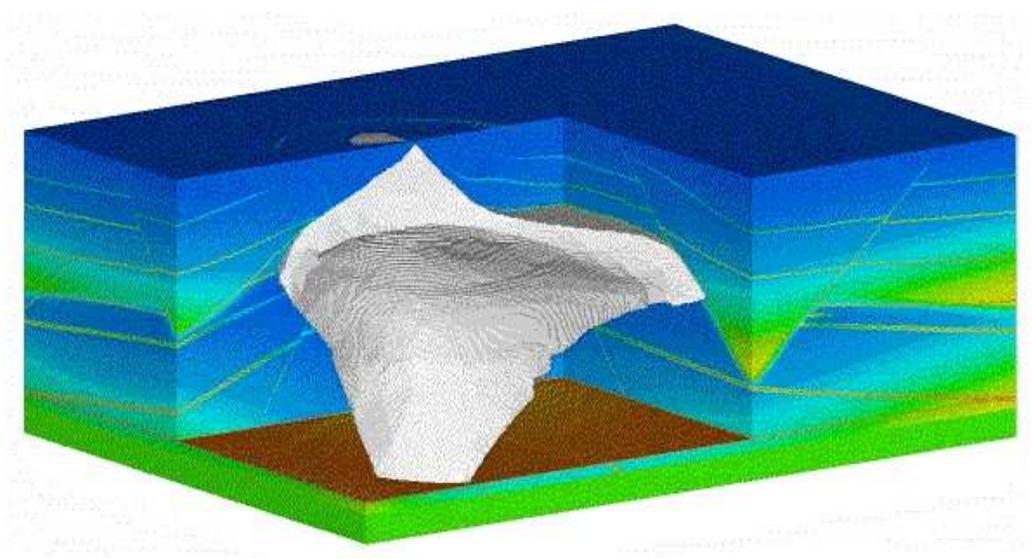


Figura 2.5: Un tipico modello 3D a differenze finite utilizzato per la seismic migration.

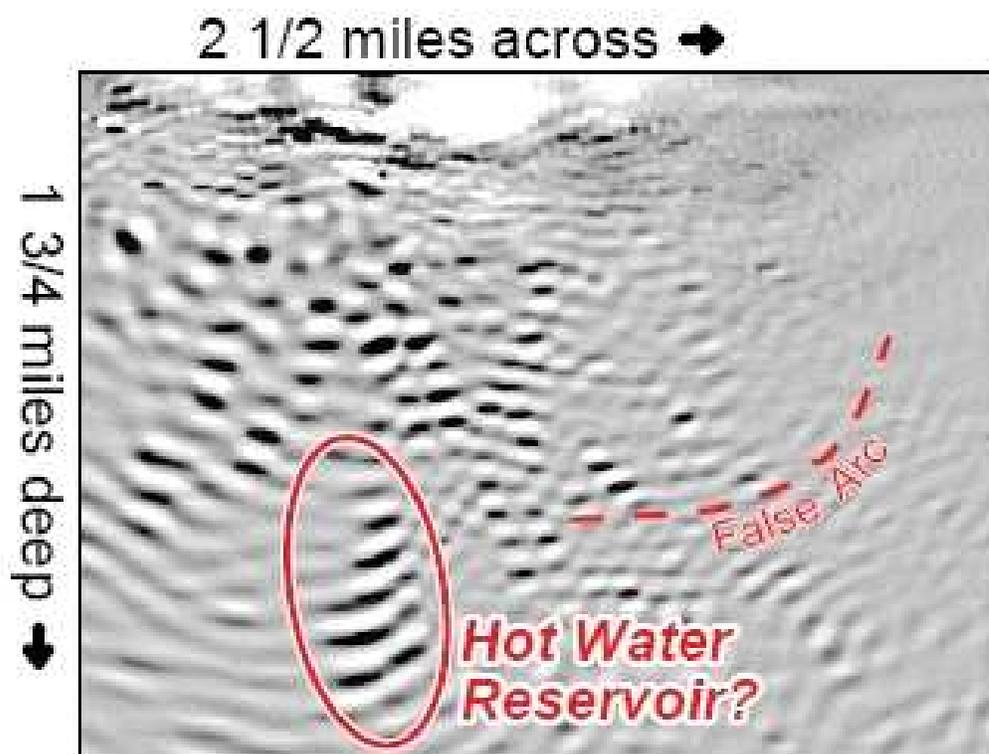


Figura 2.6: Seismic migration relativa ad un sito nel Nevada (USA) utilizzato per la produzione di energia geotermica; è evidenziato un possibile giacimento di acqua calda ad una profondità di circa 4 km.

Capitolo 3

Metodo agli elementi finiti

3.1 Introduzione

Il metodo degli elementi finiti (Finite Element Method, FEM) è una tecnica numerica atta a determinare soluzioni approssimate di problemi descritti da equazioni differenziali ordinarie e alle derivate parziali. A tale scopo il problema iniziale viene trasformato in un sistema di equazioni algebriche la cui soluzione fornisce una soluzione approssimata del problema differenziale iniziale. Benché esso competa, in particolari ambiti, con altre strategie numeriche (metodo delle differenze finite, metodo dei volumi finiti, metodo degli elementi al contorno, metodo delle celle, metodo spettrale, etc.), il metodo FEM mantiene una posizione dominante nel panorama delle tecniche numeriche di approssimazione e rappresenta il kernel di gran parte dei codici di calcolo disponibili in commercio. In generale, il metodo agli elementi finiti si presta molto bene a risolvere equazioni alle derivate parziali quando:

- a) il dominio ha forma complessa (telaio di un'automobile o il motore di un aereo);
- b) il dominio è variabile (reazione di stato con condizioni al contorno variabili);
- c) l'accuratezza richiesta sulla soluzione non è omogenea sul dominio (nel crash test su un'automobile, l'accuratezza richiesta è maggiore in prossimità della zona di impatto);
- d) la soluzione cercata manca di regolarità.

Il metodo degli elementi finiti è stato inizialmente motivato dalla necessità di risolvere problemi complessi di analisi elastica e strutturale nel campo dell'ingegneria civile ed aeronautica. L'origine del metodo può venir fatta risalire agli anni 1930-35 con i lavori di A. R. Collar e W. J. Duncan [1, 2], che introducono una forma primitiva di elemento strutturale nella risoluzione di un problema di aeroelasticità, e agli anni 1940-43 con i lavori di Alexander Hrennikoff e Richard Courant, dove entrambi, benché con differenti approcci, condividevano l'idea di suddividere il dominio del problema in sottodomini di forma semplice [3]. Tuttavia la nascita vera e propria e lo sviluppo del metodo agli elementi finiti si colloca nella seconda metà degli anni 50 con il contributo fondamentale di M. J. (Jon) Turner della Boeing, che formulò e perfezionò il Direct Stiffness Method, il primo approccio agli elementi finiti nel campo del continuo [4]. Il lavoro di Turner trovò diffusione fuori dagli stretti ambiti dell'ingegneria aerospaziale, ed in particolare nell'ingegneria civile, tramite il lavoro di John Argyris presso l'Università di Stoccarda e di Ray W. Clough presso l'Università di Berkeley [5] (che parlò per primo di metodo FEM e la cui collaborazione con Turner aveva dato vita al celebre lavoro[5] che è universalmente considerato come l'inizio del moderno FEM).

3.2 Generalità

Nel caso di equazioni differenziali lineari, esso sostanzialmente riduce un problema differenziale con condizioni alla frontiera (boundary value problem) in un sistema di equazioni lineari. Dato infatti un problema lineare $Af = g$, dove l'operatore differenziale $A : X \rightarrow Y$ con X e Y spazi di Banach infinito dimensionali, il FEM lo trasforma in uno più semplice mediante l'uso di un operatore di proiezione P_n :

$$P_n Af = P_n g,$$

dove $P_n : Y \rightarrow Y_n$, con $Y_n \subseteq Y$ e n un parametro che indica la dimensione **finita** del sottospazio Y_n di Y . Di conseguenza il problema passa dagli spazi infinito dimensionali X e Y a dei rispettivi sottospazi X_n e Y_n di dimensione n , tali che

$$X_1 \subseteq X_2 \subseteq \dots X_n, \quad Y_1 \subseteq Y_2 \subseteq \dots Y_n,$$

e

$$\lim_{n \rightarrow \infty} X_n = X, \quad \lim_{n \rightarrow \infty} Y_n = Y.$$

Ha senso seguire questa procedura quando si esegue la proiezione in sottospazi che consentono di semplificare realmente il problema in esame. È chiaro che un'operazione di proiezione ci porta dal problema originale ad una sua approssimazione, e per cui è fondamentale tenere sotto controllo l'errore, definito come la distanza fra la soluzione f del problema esatto e la soluzione f_n del problema approssimato: $\|f - f_n\|$. Si dimostra, sotto ipotesi opportune, che $\|f - f_n\| \leq \text{Cost} \times \text{dist}(f, X_n)$ e pertanto $\|f - f_n\| \rightarrow 0$ quando $n \rightarrow \infty$ (Fig. 3.1).

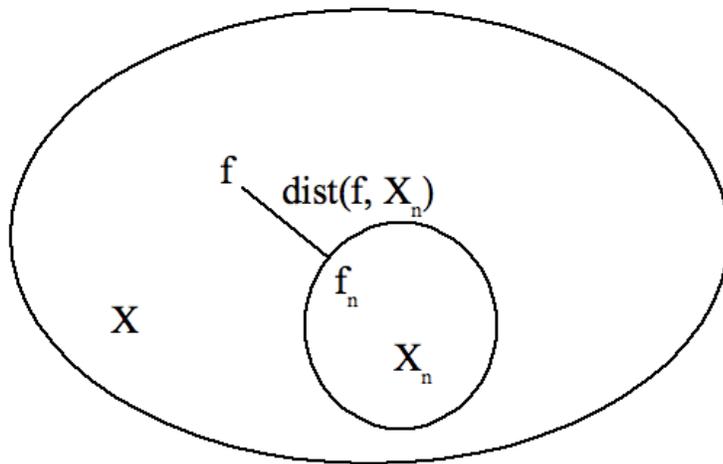


Figura 3.1: X_n è il sottospazio di X di dimensione n nel quale si ricerca la soluzione approssimata f_n . Supponendo che la soluzione esatta f non sia inclusa in X_n , $\text{dist}(f, X_n)$ indica la minima distanza (secondo una data norma naturale) fra soluzione esatta e soluzione approssimata. Al crescere della dimensione n si ha che $X_n \rightarrow X$ e pertanto la distanza tra le due soluzioni diventa sempre più piccola sino ad azzerarsi quando $n \rightarrow \infty$.

A differenza del metodo alle differenze finite, il FEM permette di analizzare anche domini non regolari, dove è richiesta una discretizzazione non uniforme, come può essere il profilo di una costa frastagliata, una porzione del motore di una automobile, l'ala di un aereo, ect. Quando si esegue la discretizzazione del dominio si parla di generazione di una mesh che, potendo essere non regolare, può essere costituita da una famiglia di triangoli, tetraedri, ect di diversa grandezza a seconda che il dominio da discretizzare sia in \mathbb{R}^2 o \mathbb{R}^3 .

3.3 Risoluzione di ODEs con il FEM

Vediamo ora come si esegue in pratica la proiezione del problema differenziale dallo spazio X al sottospazio X_n . Nel caso del FEM, la proiezione consiste nel trasformare inizialmente il problema differenziale in una forma debole infinito dimensionale e nel risolvere successivamente la sua formulazione debole, detta anche variazionale, in uno spazio di dimensione finita. Il seguente esempio illustra come si esegue questa trasformazione e poi come si ottiene la soluzione approssimata in uno spazio di dimensione finita.

Si consideri il seguente problema differenziale ordinario con valori agli estremi:

$$\begin{cases} -\frac{d}{dx} \left(p(x) \frac{dy}{dx} \right) + q(x)y = f(x), & x \in [a, b] \subset \mathbb{R} \\ y(a) = \alpha, & y(b) = \beta, \end{cases} \quad (3.1)$$

dove $p(x) \in C^1([a, b])$ e $q(x), f(x) \in C^0([a, b])$. Per comodità di esposizione consideriamo $\alpha = \beta = 0$.

Questo non è limitativo in quanto se le condizioni agli estremi non sono omogenee, è sempre possibile ricondursi al problema omogeneo con una banale sostituzione $y(x) = z(x) + \varphi(x)$, dove $\varphi(x)$ è una funzione nota con $\varphi(a) = \alpha$ e $\varphi(b) = \beta$. Di conseguenza $z(a) = z(b) = 0$, essendo $z(x)$ soluzione dell'equazione

$$-\frac{d}{dx} [p(x) (z'(x) + \varphi'(x))] + q(x) (z(x) + \varphi(x)) = f(x).$$

Dal momento che la funzione $\varphi(x)$ è nota si ha un problema differenziale per la funzione $z(x)$ con condizioni al contorno omogenee:

$$\begin{cases} -\frac{d}{dx} \left(p(x) \frac{dz}{dx} \right) + q(x)z = g(x), & x \in [a, b] \subset \mathbb{R} \\ z(a) = 0, & z(b) = 0, \end{cases}$$

dove $g(x) = f(x) + \frac{d}{dx} [p(x)\varphi'(x)] - q(x)\varphi(x)$. La scelta più comoda per la funzione φ è la retta passante per i punti (a, α) e (b, β) :

$$\varphi = \alpha + \frac{x - a}{b - a}(\beta - \alpha).$$

Possiamo quindi considerare il problema differenziale (3.1) con condizioni al contorno omogenee, senza perdita di generalità. Tale problema differenziale è detto problema in **forma forte**, perchè la soluzione che si cerca deve appartenere a $C^2([a, b])$ e annullarsi agli estremi del dominio (problema di

Dirichlet). Si cerca pertanto di **indebolire** la richiesta che la soluzione sia due volte derivabile, con derivata seconda continua, procedendo con la proiezione del problema nella sua forma debole. Si considera a tal fine uno spazio di funzioni $v(x)$ differenziabili nel **senso distribuzionale**, tali da annullarsi agli estremi. Fondamentalmente per queste funzioni la derivata ha senso solo se preceduta dal segno di integrale. Lo spazio di questa classe di funzioni si indica con $H_0^1([a, b])$ e si parla di spazi di Sobolev [6]. Per ottenere la forma debole, si moltiplica primo e secondo membro della (3.1) per la generica $v(x) \in H_0^1([a, b])$

$$-\frac{d}{dx} \left(p(x) \frac{dy(x)}{dx} \right) v(x) + q(x)y(x)v(x) = f(x)v(x),$$

e si integra primo e secondo membro lungo il dominio:

$$-\int_a^b \frac{d}{dx} \left(p(x) \frac{dy(x)}{dx} \right) v(x) dx + \int_a^b q(x)y(x)v(x) dx = \int_a^b f(x)v(x) dx. \quad (3.2)$$

Integrando per parti il primo termine della (3.2), si trova

$$-\underbrace{\left[p(x) \frac{dy}{dx} v(x) \right]_a^b}_{=0} + \int_a^b p(x) \frac{dy}{dx} \frac{dv}{dx} dx + \int_a^b q(x)y(x)v(x) dx = \int_a^b f(x)v(x) dx,$$

dove il primo termine è uguale a zero perché $v(x) \in H_0^1([a, b])$ e pertanto si trova:

$$\int_a^b p(x) \frac{dy(x)}{dx} \frac{dv(x)}{dx} dx + \int_a^b q(x)y(x)v(x) dx = \int_a^b f(x)v(x) dx. \quad (3.3)$$

La Eq. (3.3) rappresenta la forma debole del problema, perché la soluzione cercata è in $H_0^1([a, b])$, spazio contenente la classe di funzioni in $H^1([a, b])$ che si annullano agli estremi del dominio. La forma ottenuta è debole perché su tali funzione non è richiesta la differenziabilità in senso ordinario, condizione più restrittiva di quella appena ottenuta.

Osservazione 3.3.1 *Ogni soluzione $y(x)$ del problema (3.1) è soluzione del problema (3.3) e viceversa, ogni soluzione in senso debole sufficientemente regolare lo è anche in senso forte. La prima parte della proposizione è evidente. Supponendo ora che $y(x)$ sia soluzione della (3.3), qualunque sia*

$v(x) \in H_0^1([a, b])$, dobbiamo dimostrare che essa soddisfa anche la (3.1). A tale scopo, procedendo per assurdo, supponiamo che esista un intervallino $[c, d] \subset [a, b]$ nel quale si abbia

$$-\frac{d}{dx} \left(p(x) \frac{dy(x)}{dx} \right) + q(x)y(x) - f(x) > 0.$$

In questo caso possiamo costruire una funzione $v(x) \in H_0^1([a, b])$ anch'essa positiva in $[c, d]$ e nulla altrove in $[a, b]$ (per esempio una spline lineare con valore uno in $(c+d)/2$ e nulla su c e d), contro l'ipotesi che il primo membro della (3.3) sia uguale al secondo membro. Lo stesso ragionamento può essere ripetuto nel caso che esista un intervallo $[c, d] \subset [a, b]$ nel quale il primo membro della (3.1) sia minore del secondo membro.

La soluzione $y(x)$ cercata appartiene comunque ad uno spazio infinito dimensionale e quindi rimane da capire come proiettare il problema (3.2) in uno spazio di dimensione finita. A tale scopo, per prima cosa, si discretizza l'intervallo $[a, b]$ con una funzione, per comodità, a spaziatura costante $x_i = a + ih$, $i = 0, 1, \dots, n+1$, $h = (b-a)/n$. Ad esso si associa uno spazio n -dimensionale, tipicamente rappresentato da n -polinomi a tratti opportunamente raccordati (splines polinomiali). Nel caso più semplice la base è formata da n -splines lineari $\{H_i(x); i = 1, \dots, n\}$, ciascuna delle quali, al fine di tenere conto della omogeneità delle condizioni agli estremi, è nulla in $x_0 = a$ e $x_{n+1} = b$. Il loro tipico andamento è rappresentato nella Fig. 3.2. Analiticamente la i -esima spline, $H_i(x)$, $i = 1, 2, \dots, n$, il cui supporto è l'intervallo $[x_{i-1}, x_{i+1}]$, di ampiezza $2h$, è così definita:

$$H_i(x) = \begin{cases} 0, & a \leq x \leq x_{i-1}, \\ \frac{x-x_{i-1}}{h}, & x_{i-1} < x < x_i, \\ \frac{x_{i+1}-x}{h}, & x_i < x < x_{i+1}, \\ 0 & x_{i+1} \leq x \leq b. \end{cases} \quad (3.4)$$

L'approssimazione n -dimensionale $y_n(x)$ della $y(x)$ in tale spazio, usualmente definito S_0^1 , viene allora fornita dalla interpolazione Lagrangiana

$$y_n(x) \simeq \sum_{j=1}^n y(x_j) H_j(x), \quad (3.5)$$

nella quale, come è evidente, $y_n(x_j) = y(x_j)$ per $j = 1, 2, \dots, n$ e $y_n(a) = y_n(b) = 0$. Le funzioni test vengono anch'esse considerate nello spazio S_0^1 che

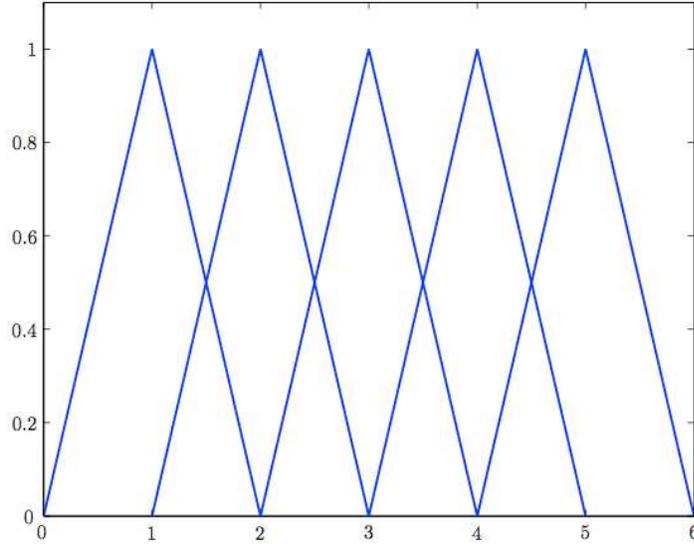


Figura 3.2: Esempio di splines lineari.

è chiaramente un sottospazio di H_0^1 , in quanto ogni funzione H_i è continua in $[a, b]$ e ivi derivabile, tranne in x_{i-1}, x_i e x_{i+1} , dove comunque esistono finite e integrabili in $[a, b]$ la derivata destra e quella sinistra, essendo

$$H_i'(x) = \begin{cases} 0, & a \leq x < x_{i-1}, \\ \frac{1}{h}, & x_{i-1} < x < x_i, \\ -\frac{1}{h}, & x_i < x < x_{i+1}, \\ 0, & x_{i+1} < x \leq b. \end{cases}$$

Sostituendo nella formulazione debole (3.6) del problema modello (3.1), la y con y_n e la funzione test v con la H_i , $i = 1, \dots, n$, si ottiene il seguente sistema lineare n-dimensionale:

$$\sum_j^n \left[\int_a^b p(x) H_j'(x) H_i'(x) dx + \int_a^b q(x) H_j(x) H_i(x) dx \right] y_j = \int_a^b f(x) H_i(x) dx, \quad (3.6)$$

dove $i = 1, 2, \dots, n$. In tale sistema il vettore $\mathbf{y}_n = (y_1, y_2, \dots, y_n)^T$ fornisce una approssimazione della soluzione $y(x)$ nei punti nodali x_1, x_2, \dots, x_n . La sua determinazione consente, tramite la (3.5), di ottenere l'approssimante $y_n(x)$ della $y(x)$ in tutto l'intervallo $[a, b]$. Esprimendo il sistema (3.6) nella

forma

$$A\mathbf{y}_n = \mathbf{b},$$

è immediato osservare che A è simmetrica e tridiagonale, essendo $a_{ij} = a_{ji}$ e $a_{ij} = 0$ per $|i - j| \geq 2$, in quanto le splines lineari sono a supporto compatto e

$$a_{ij} = \int_a^b [p(x)H'_i(x)H'_j(x) + q(x)H_i(x)H_j(x)] dx,$$

per $i, j = 1, 2, \dots, n$. È questo il motivo per cui le splines lineari sono tanto utilizzate nella risoluzione dei problemi differenziali mediante gli elementi finiti. Grazie alla compattezza del supporto delle $H_i(x)$ ($i=1, \dots, n$)

$$a_{ii} = \frac{1}{h^2} \left[\int_{x_{i-1}}^{x_{i+1}} p(x) dx + \int_{x_{i-1}}^{x_i} q(x)(x - x_{i-1})^2 dx + \int_{x_i}^{x_{i+1}} q(x)(x_{i+1} - x)^2 dx \right],$$

$$a_{i,i+1} = a_{i+1,i} = -\frac{1}{h^2} \int_{x_i}^{x_{i+1}} [p(x) + q(x)(x - x_i)(x - x_{i+1})] dx = a_{i+1,i},$$

$$b_i = \frac{1}{h} \left[\int_{x_{i-1}}^{x_i} f(x)(x - x_{i-1}) dx + \int_{x_i}^{x_{i+1}} f(x)(x_{i+1} - x) dx \right],$$

dove $i = 1, \dots, n$ e $j = 1, \dots, n$.

Stima dell'errore

L'errore di risoluzione in $H_0^1[a, b]$ tende a zero con $h \rightarrow 0$. Si dimostra infatti che

$$\text{a) } \|y - y_n\|_{L^2[a,b]} = \left(\int_a^b [y(x) - y_n(x)]^2 dx \right)^{1/2} = O(h^2),$$

$$\text{b) } \|y - y_n\|_{H_0^1[a,b]} = \left(\int_a^b [y(x) - y_n(x)]^2 dx + \int_a^b [y'(x) - y'_n(x)]^2 dx \right)^{1/2} = O(h).$$

Questo significa che l'errore può essere reso piccolo a piacere, pur di decomporre l'intervallo $[a, b]$ con un numero sufficientemente elevato di punti. Problema rilevante, dato che la matrice H è simmetrica tridiagonale.

3.4 Risoluzione di PDEs con il FEM

L'estensione del modello differenziale omogeneo del secondo ordine a modelli di tipo ellittico, in 2 variabili, è il seguente:

$$\begin{cases} -\frac{\partial}{\partial x_1} \left(\kappa(x_1, x_2) \frac{\partial u}{\partial x_1} \right) - \frac{\partial}{\partial x_2} \left(\kappa(x_1, x_2) \frac{\partial u}{\partial x_2} \right) + a_0(x_1, x_2)u = f(x_1, x_2), \\ (x_1, x_2) \in \Omega, \quad u(x_1, x_2) = 0, \text{ per } (x_1, x_2) \in \partial\Omega \text{ (frontiera di } \Omega), \end{cases} \quad (3.7)$$

dove, per l'esistenza e l'unicità in senso forte (puntuale) della soluzione in Ω , si suppone che $\kappa(x_1, x_2)$ sia differenziabile e non negativa, $a_0(x) \geq 0$ e $f(x_1, x_2)$ sia continua. Come nel modello differenziale ordinario, la differenziabilità di κ e la continuità di f in Ω non sono necessarie nell'esistenza della soluzione in senso debole. Qualora le condizioni al bordo non siano omogenee, a differenza di quanto avviene nel modello differenziale ordinario, non è sempre possibile ricondursi al caso omogeneo. Questo è possibile qualora il dominio sia un rettangolo, mediante l'introduzione di una funzione ausiliaria $\varphi(x_1, x_2)$ che al bordo assume i valori assegnati. A tale scopo si può utilizzare la stessa tecnica utilizzata per la determinazione del vettore di innesco nella risoluzione delle equazioni di tipo debolmente nonlineare con il metodo di Newton-Jacobi.

Una volta determinata tale funzione, posto $u(x_1, x_2) = v(x_1, x_2) + \varphi(x_1, x_2)$, si ottiene il seguente modello differenziale con condizioni di omogeneità al bordo:

$$\begin{cases} -\frac{\partial}{\partial x_1} \left(\kappa(x_1, x_2) \frac{\partial v}{\partial x_1} \right) - \frac{\partial}{\partial x_2} \left(\kappa(x_1, x_2) \frac{\partial v}{\partial x_2} \right) + a_0(x_1, x_2)v = g(x_1, x_2), \quad (x_1, x_2) \in \Omega, \\ v(x_1, x_2) = 0, \quad (x_1, x_2) \in \partial\Omega, \end{cases}$$

essendo $g(x_1, x_2) = f(x_1, x_2) - a_0(x_1, x_2)\varphi + \frac{\partial}{\partial x_1} \left(\kappa(x_1, x_2) \frac{\partial \varphi}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left(\kappa(x_1, x_2) \frac{\partial \varphi}{\partial x_2} \right)$ una funzione nota in Ω . Allo scopo di estendere il modello (3.7) al caso n-dimensionale, è necessario premettere le definizioni di gradiente, vettore normale esterno e divergenza, il teorema della divergenza e la prima identità di Green.

Gradiente di una funzione

Indicata con $u(\mathbf{x}) = u(x_1, x_2, \dots, x_n)$ una funzione definita in un dominio Ω , e ivi differenziabile, per gradiente della u in Ω , in simboli, $\nabla u = \text{gradu}$ (nabla)

u uguale gradiente di u), si intende il vettore

$$\nabla u = \left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial u}{\partial x_n} \right)^T.$$

Esempio 3.4.1 Il gradiente della funzione $u(x_1, x_2, x_3) = x_1^2 + x_2^2 + \cos(x_1 x_2 x_3)$, è il vettore

$$\nabla u = \begin{pmatrix} 2x_1 - x_2 x_3 \sin(x_1 x_2 x_3) \\ 2x_2 x_3 \sin(x_1 x_2 x_3) \\ 1 - x_1 x_2 \sin(x_1 x_2 x_3) \end{pmatrix}.$$

Derivata direzionale-gradiente

Supponiamo che $u(x_1, x_2, \dots, x_n)$ e le sue derivate parziali siano continue in una sfera con centro $\mathbf{x}^{(0)}$ e che $\nabla u(\mathbf{x}^{(0)}) \neq 0$. Per derivata direzionale della u in $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T$, lungo la direzione $\mathbf{d} = (d_1, d_2, \dots, d_n)^T$, si intende il vettore

$$\partial_d u(\mathbf{x}^{(0)}) = \nabla u(\mathbf{x}^{(0)}) \cdot \mathbf{d} = \frac{\partial u(\mathbf{x}^{(0)})}{\partial x_1} d_1 + \dots + \frac{\partial u(\mathbf{x}^{(0)})}{\partial x_n} d_n,$$

dove, come mostra la definizione e come usuale, il punto indica il prodotto interno (prodotto scalare) tra i vettori $\nabla u(\mathbf{x}^{(0)})$ e \mathbf{d} . Da tale definizione, ricordando che, per definizione di prodotto tra vettori,

$$\partial_d u(\mathbf{x}^{(0)}) = \|\nabla u(\mathbf{x}^{(0)})\| \|\mathbf{d}\| \cos \theta,$$

essendo θ l'angolo tra i vettori $\nabla u(\mathbf{x}^{(0)})$ e \mathbf{d} , seguono le seguenti osservazioni:

- a) la direzione di massima crescita, ossia la direzione \mathbf{d} , di lunghezza unitaria, lungo la quale, a partire da $\mathbf{x}^{(0)}$ la funzione u cresce maggiormente è rappresentata dal vettore unitario (versore)

$$\mathbf{d} = \frac{\nabla u(\mathbf{x}^{(0)})}{\|\nabla u(\mathbf{x}^{(0)})\|}.$$

b) Per lo stesso tipo di considerazioni, il versore

$$\mathbf{d} = -\frac{\nabla u(\mathbf{x}^{(0)})}{\|\nabla u(\mathbf{x}^{(0)})\|},$$

indica la direzione di massima decrescita, ossia la direzione lungo la quale, a partire da $\mathbf{x}^{(0)}$, la u presenta la più ripida discesa.

Al fine di interpretare geometricamente la considerazione b), consideriamo la famiglia delle curve di livello

$$u(x_1, x_2) = c, \quad c \geq 0,$$

associata ad una funzione differenziabile u in un dominio Ω . Considerato un punto $\mathbf{x}^{(0)}$, la direzione di più ripida discesa, come indicato nella Figura 3.3, rappresenta il versore alla tangente di una curva di livello, in $\mathbf{x}^{(0)}$, rivolto verso l'interno, ossia nel verso delle curve di livello con c decrescenti.

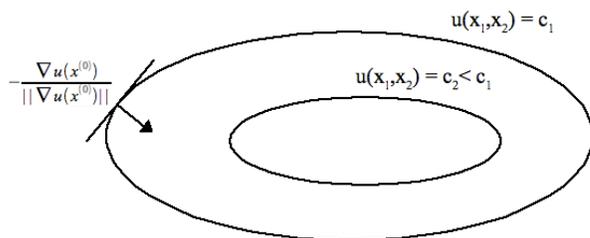


Figura 3.3: La figura mostra due curve di livello della funzione $u(x_1, x_2)$ e il suo versore gradiente.

Esempio 3.4.2 Se $u(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$, il versore di più ripida discesa in $\mathbf{x}^{(0)} = (1, 1, 1)^T$ è $\mathbf{d} = -\frac{1}{\sqrt{3}}(1, 1, 1)^T$. Esso indica infatti che, qualunque sia la superficie sferica di equazione $x_1^2 + x_2^2 + x_3^2 = c$, $c \geq 0$, il versore \mathbf{d} , partendo da $\mathbf{x}^{(0)}$, punta verso l'origine, ossia al centro della sfera caratterizzata da $c = 0$.

Vettore normale esterno

Per vettore normale esterno (vettore esterno di norma 1) ad una superficie, dotata di piano tangente in un punto $\mathbf{x}^{(0)}$, si intende il versore normale al piano tangente, orientato nel verso delle $u(\mathbf{x})$ crescenti. Generalmente indicato

con

$$\mathbf{n} = \frac{\nabla u(\mathbf{x}^{(0)})}{\|\nabla u(\mathbf{x}^{(0)})\|},$$

esso rappresenta, nel punto, la direzione di più rapida crescita.

Esempio 3.4.3 Consideriamo il caso di equazione

$$x_3 = \sqrt{x_1^2 + x_2^2},$$

in cui x_1 e x_2 sono numeri reali qualsiasi. Il versore normale esterno alla superficie $z_3 - \sqrt{x_1^2 + x_2^2} = 0$ in (x_1, x_2, x_3) , è

$$\mathbf{n} = \frac{1}{\sqrt{2}} \left(\frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \frac{x_2}{\sqrt{x_1^2 + x_2^2}}, -1 \right)^T = \frac{1}{\sqrt{2}} \left(\frac{x_1}{x_3}, \frac{x_2}{x_3}, -1 \right)^T.$$

Il versore normale alla suddetta superficie conica, in $(1, 1, \sqrt{2})$ è pertanto

$$\mathbf{n} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -1 \right)^T.$$

Divergenza di un campo vettoriale

Indicato con

$$v(\mathbf{x}) = (v_1(\mathbf{x}), v_2(\mathbf{x}), \dots, v_n(\mathbf{x}))^T, \quad \mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$$

un vettore n -dimensionale differenziabile in un punto $\mathbf{x} \in \Omega$, per divergenza di v in \mathbf{x} , in simboli

$$(\nabla \cdot v)(\mathbf{x}) = (\operatorname{div} v)(\mathbf{x}) \quad (\operatorname{nabla} \cdot v \text{ in } \mathbf{x} = \text{divergenza di } v \text{ in } \mathbf{x}),$$

si intende la funzione

$$(\nabla \cdot v)(\mathbf{x}) = \frac{\partial v_1(\mathbf{x})}{\partial x_1} + \frac{\partial v_2(\mathbf{x})}{\partial x_2} + \dots + \frac{\partial v_n(\mathbf{x})}{\partial x_n}.$$

Esempio 3.4.4 Se $v(\mathbf{x}) = (x_1^2 + x_2^2, x_1x_2x_3, x_3 \sin(x_1x_2))^T$, la divergenza di v in \mathbf{x} è la funzione

$$(\nabla \cdot v)(\mathbf{x}) = 2x_1 + x_1x_3 + \sin(x_1x_2).$$

Se, in particolare, v è una funzione differenziabile due volte in Ω ,

$$\begin{aligned} (\nabla \cdot \nabla v)(\mathbf{x}) &= \frac{\partial}{\partial x_1} \frac{\partial v(\mathbf{x})}{\partial x_1} + \frac{\partial}{\partial x_2} \frac{\partial v(\mathbf{x})}{\partial x_2} + \dots + \frac{\partial}{\partial x_n} \frac{\partial v(\mathbf{x})}{\partial x_n} \\ &= \frac{\partial^2 v(\mathbf{x})}{\partial^2 x_1} + \frac{\partial^2 v(\mathbf{x})}{\partial^2 x_2} + \dots + \frac{\partial^2 v(\mathbf{x})}{\partial^2 x_n} \end{aligned}$$

che, come noto, indica il Laplaciano della v in un punto $\mathbf{x} \in \Omega$. In simboli

$$(\nabla \cdot \nabla v)(\mathbf{x}) = \Delta v(\mathbf{x}) = \nabla^2 v(\mathbf{x}).$$

Nell'esempio precedente ($u(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$)

$$(\nabla \cdot \nabla v)(\mathbf{x}) = (\nabla \cdot \nabla u)(\mathbf{x}) = 3.$$

Teorema della divergenza

Indicato con $\mathbf{f} = (f_1, f_2, \dots, f_n)^T$ un campo vettoriale in un dominio $\Omega \subset \mathbb{R}^n$, le cui componenti $f_1(x), f_2(x), \dots, f_n(x)$ sono ivi differenziabili con continuità, per la $\text{div}(\mathbf{f})$ in Ω vale la seguente importante relazione:

$$\int_{\Omega} \nabla \cdot \mathbf{f} dv = \int_{\partial\Omega} \mathbf{f} \cdot \mathbf{n} d\sigma, \quad (3.8)$$

dove \mathbf{n} è il vettore unitario esterno alla superficie $\partial\Omega$, dv è l'elemento infinitesimo di volume e $d\sigma$ l'elemento infinitesimo di superficie. Questo risultato, oltre che molto utile dal punto di vista del calcolo, è molto importante in fluidodinamica in quanto esso stabilisce che il flusso netto di un liquido in un dominio Ω , ossia quello che attraversa la sua frontiera $\partial\Omega$, è pari all'apporto complessivo delle sue sorgenti meno quello eliminato dai pozzi in esso contenuti.

Verifichiamo ora il teorema su di un esempio. A tale scopo, posto $\mathbf{f}(x, y, z) = (x, y, z)^T$, consideriamo la superficie S , formata dalla superficie S_1 di un cono $z = \sqrt{x^2 + y^2}$, con $x^2 + y^2 \leq 1$, e della superficie piana S_2 formata dal cerchio $x^2 + y^2 \leq 1$ nel piano $z = 1$ (Fig. 3.4).

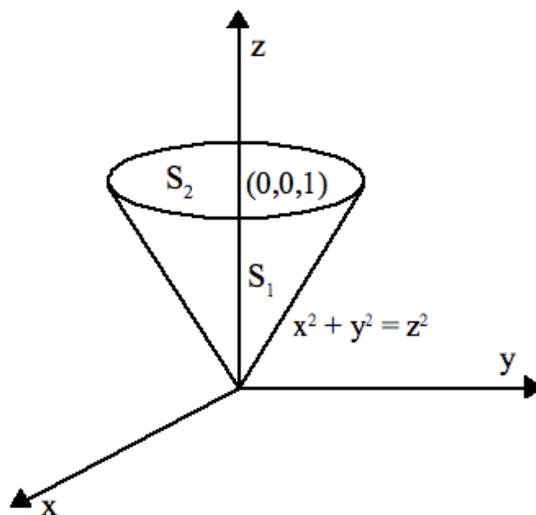


Figura 3.4: La figura mostra la superficie S , formata dalla superficie S_1 di un cono $z = \sqrt{x^2 + y^2}$, con $x^2 + y^2 \leq 1$, e della superficie piana S_2 formata dal cerchio $x^2 + y^2 \leq 1$ nel piano $z = 1$.

Esempio 3.4.5 Come dimostrato in precedenza, il vettore normale esterno (versore normale esterno) in (x, y, z) è $\mathbf{n}_1 = \frac{1}{\sqrt{2}} \left(\frac{x}{z}, \frac{y}{z}, -1 \right)^T$. Di conseguenza, nella superficie S_1 il flusso è nullo, in quanto

$$\mathbf{f} \cdot \mathbf{n}_1 = \frac{1}{\sqrt{2}} \left(\frac{x^2}{z} + \frac{y^2}{z} - z \right) = \frac{1}{\sqrt{2}} \frac{x^2 + y^2 - z^2}{z} = 0.$$

Su S_2 il versore normale esterno è $\mathbf{n}_2 = (0, 0, 1)^T$ per cui, essendo $z = 1$,

$$\int_{S_2} \mathbf{f} \cdot \mathbf{n}_2 d\sigma = \int_{S_2} z d\sigma = \int_{S_2} d\sigma = \pi \quad (\text{area di } S_2).$$

Calcoliamo ora l'integrale di volume. A tale scopo, osservato che

$$\nabla \cdot \mathbf{f} = \frac{\partial}{\partial x} x + \frac{\partial}{\partial y} y + \frac{\partial}{\partial z} z = 3,$$

e indicato con V il volume del cono,

$$\int_V \nabla \cdot \mathbf{f} dv = \int_V 3 dv = \pi \quad (3 \times \text{ il volume del cono}).$$

Il teorema è pertanto verificato nell'esempio. Consideriamo ora un esempio nel quale è tutt'altro che indifferente calcolare il flusso mediante l'integrale di volume o di superficie.

Esempio 3.4.6 Sia $\partial\Omega$ la superficie di un cubo Ω , avente come vertici i punti:

$$(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), (0, 1, 1), (1, 0, 1), (1, 1, 1).$$

Indicato con \mathbf{f} il vettore di flusso $\mathbf{f} = (x^2, y^2, z^2)^T$, vogliamo calcolare il flusso che attraversa le facce del cubo Ω . Esso è dato dall'integrale di superficie

$$I = \int_{\partial\Omega} \mathbf{f} \cdot \mathbf{n} d\sigma$$

che non è immediato calcolare, in quanto occorre valutare il flusso attraverso ciascuna delle sei facce del cubo. Il calcolo dell'integrale di volume della $\text{div}(\mathbf{f})$ è invece immediato in quanto, osservato che $\nabla \cdot \mathbf{f} = 2x + 2y + 2z$, si ha che

$$\begin{aligned} I &= \int_{\Omega} \nabla \cdot \mathbf{f} dv = \int_0^1 \int_0^1 \int_0^1 (2x + 2y + 2z) dx dy dz = \\ &= \int_0^1 \int_0^1 (2x + 2y + 1) dx dy = \int_0^1 (2x + 2) dx = 3. \end{aligned}$$

Dal teorema della divergenza discende la seguente importante relazione, nota come **Prima identità di Green**:

Se u è una funzione dotata di derivata seconda continua in Ω e φ è una funzione ivi differenziabile, indicata con $\frac{\partial u}{\partial n}$ la derivata di u lungo il versore normale esterno \mathbf{n} in $\partial\Omega$, vale la seguente identità:

$$\int_{\Omega} (\Delta u) \varphi dv = \int_{\partial\Omega} \left(\frac{\partial u}{\partial n} \right) \varphi d\sigma - \int_{\Omega} \nabla u \cdot \nabla \varphi dv, \quad (3.9)$$

dove Δu indica il laplaciano della u ($\Delta u = \nabla \cdot \nabla u$).

Per la dimostrazione si osserva, in primo luogo, che

$$\nabla \cdot (\varphi \nabla u) = \nabla u \cdot \nabla \varphi + \varphi \Delta u,$$

da cui

$$\int_{\Omega} \nabla \cdot (\varphi \nabla u) dv = \int_{\Omega} \nabla u \cdot \nabla \varphi dv + \int_{\Omega} \varphi \Delta u dv$$

e inoltre, per il teorema della divergenza applicato all'integrale a sinistra dell'uguale,

$$\int_{\partial\Omega} \varphi \nabla u \cdot \mathbf{n} d\sigma = \int_{\Omega} \nabla u \cdot \nabla \varphi dv + \int_{\Omega} \varphi \Delta u dv.$$

Da tale relazione, ricordando che $\frac{\partial u}{\partial n} = \nabla u \cdot \mathbf{n}$, segue immediatamente l'identità (3.9).

L'estensione n-dimensionale del modello differenziale ellittico (3.7) è la seguente:

$$\begin{aligned} -\nabla \cdot (\kappa(\mathbf{x}) \nabla u) + a_0(\mathbf{x})u &= f(\mathbf{x}), & \mathbf{x} \in \Omega \subset \mathbb{R}^n \\ u(\mathbf{x}) &= 0 & \mathbf{x} \in \partial\Omega \end{aligned} \quad (3.10)$$

nel quale $\kappa(\mathbf{x})$ è differenziabile e nonnegativa in Ω , $a_0(\mathbf{x}) \geq 0$ in Ω e la f è ivi continua. Sotto tali ipotesi, il modello è di tipo ellittico, per cui l'osservazione della u su $\partial\Omega$ garantisce l'esistenza e l'unicità in senso forte (puntuale) della soluzione in Ω . Per la formulazione variazionale (debole) del problema le ipotesi di continuità della f e della u , come alla differenziabilità della κ in Ω , non sono necessarie. Continuano ad essere importanti le condizioni di positività di κ e la nonnegatività di a_0 , in quanto essenziali per la condizione di ellitticità in Ω dell'operatore differenziale:

- (a) $\kappa(\mathbf{x}) \geq c\|\xi\|^2$, qualunque sia $\xi \in \mathbb{R}^n$,
- (b) $a_0(\mathbf{x}) \geq 0$ in Ω .

Esse infatti garantiscono l'esistenza e l'unicità della soluzione debole anche nel caso non esista la soluzione in senso forte del problema (3.10).

Il caso $\kappa(\mathbf{x})$ positivo e $a_0(\mathbf{x}) \leq 0$ (modello modificato) è comunque di notevole interesse nello scattering di onde acustiche, elastiche ed elettromagnetiche, per lo sviluppo delle tecnologie di ricostruzione di immagini non distruttive (imaging). In queste ipotesi è tuttavia molto più complesso dimostrare la buona posizione del problema. Risultati importanti in questo ambito si trovano in [7].

In analogia con quanto visto nel caso ordinario, per ottenere la formulazione debole del modello (3.10) si considera lo spazio $H_0^1(\Omega)$ delle funzioni

nulle in $\partial\Omega$, le cui derivate prime esistono in Ω (nel senso delle distribuzioni) e sono ivi al quadrato integrabili nel senso di Lebesgue e nel quale la norma è così definita:

$$\|u\|^2 = \|u\|_{H_0^1(\Omega)}^2 = \int_{\Omega} \left[u^2 + \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i} \right|^2 \right] d\Omega, \quad u \in H_0^1(\Omega).$$

Questo significa che la soluzione $u \in H_0^1(\Omega)$ come anche le funzioni test $\varphi \in H_0^1(\Omega)$. Si richiede dunque l'esistenza e l'unicità di una funzione $u \in H_0^1(\Omega)$ tale che l'equazione

$$\int_{\Omega} [\nabla(\kappa(\mathbf{x})) \cdot \nabla u + a_0 u] \varphi(\mathbf{x}) dv = \int_{\Omega} f(\mathbf{x}) \varphi(x) dv \quad (3.11)$$

sia soddisfatta qualunque sia $\varphi \in H_0^1(\Omega)$.

A questo punto, allo scopo di poter utilizzare le condizioni di annullamento delle funzioni test su $\partial\Omega$, si utilizza un teorema di Green che fornisce una generalizzazione della prima identità di Green (3.9). Esso stabilisce che, qualunque siano u e $\varphi \in H_0^1(\Omega)$

$$- \int_{\Omega} \nabla(\kappa(\mathbf{x}) \nabla u) \varphi dv = - \int_{\partial\Omega} \left(\kappa(\mathbf{x}) \frac{\partial u}{\partial x} \right) \varphi d\sigma + \int_{\Omega} \kappa(\mathbf{x}) \cdot \nabla \varphi dv$$

dove, al solito, \mathbf{n} rappresenta il versore normale esterno a $\partial\Omega$ in \mathbf{x} . Essendo $\varphi = 0$ su $\partial\Omega$, il problema variazionale (3.11) diventa: determinare $u \in H_0^1(\Omega)$ in modo che, qualunque sia $\varphi \in H_0^1(\Omega)$, sia soddisfatta l'equazione

$$\int_{\Omega} [\kappa(\mathbf{x}) \nabla u \cdot \nabla \varphi + a_0(\mathbf{x}) u \varphi] \varphi(\mathbf{x}) dv = \int_{\Omega} f(\mathbf{x}) \varphi(x) dv. \quad (3.12)$$

Questo modello fornisce la formulazione variazionale debole utilizzata nella risoluzione del problema (3.10) mediante gli elementi finiti. L'esistenza e l'unicità della soluzione di tale formulazione in $H_0^1(\Omega)$ sono garantite dall'ipotesi di ellitticità del modello differenziale (3.10).

Una proprietà basilare per la risoluzione in $H_0^1(\Omega)$ del problema variazionale (3.12), mediante gli elementi finiti, è la sua separabilità. Questo significa che $H_0^1(\Omega)$ possiede un sottoinsieme X numerabile e denso in $H_0^1(\Omega)$. Come ben noto, dire che X è denso in $H_0^1(\Omega)$ vuol dire, prefissato ad arbitrio un numero $\varepsilon > 0$, che qualunque sia la funzione $v \in H_0^1(\Omega)$, esiste una $\varphi \in X$ tale che $\|v - \varphi\| \leq \varepsilon$, dove il simbolo $\|\cdot\|$ indica la norma in $H_0^1(\Omega)$.

Come conseguenza, a prescindere dalla precisione richiesta nei risultati, è sufficiente limitarsi a risolvere il problema variazionale (3.12) in X . Inoltre, non essendo in grado di risolverlo nello spazio infinito dimensionale X , grazie alla sua numerabilità, ci si può accontentare di risolverlo in una sequenza di sottospazi X_h di X , ciascuno di dimensione finita $N = N(h)$, caratterizzato da un parametro $h > 0$ che, al tendere di h a zero, soddisfa le seguenti due proprietà:

- (a) la dimensione $N = N(h)$ tende all'infinito per $h \rightarrow 0$;
- (b) lo spazio X_h diventa denso in $H_0^1(\Omega)$ per $h \rightarrow 0$.

Come vedremo in seguito, la costruzione degli spazi $\{X_h\}$ viene realizzata, reticolando opportunamente il dominio Ω , numerando i punti nodali interni ed associando a ciascuno di essi una funzione che in esso assume il valore 1, mentre assume il valore zero in tutti gli altri punti nodali interni e di frontiera. Una volta costruito lo spazio X_h , si determina in esso l'unica funzione u_h che soddisfa il sistema $N(h)$ -dimensionale

$$\int_{\Omega} [\kappa(\mathbf{x}) \nabla u_h \cdot \nabla \varphi_h + a_0(\mathbf{x}) u_h \varphi_h] dv = \int_{\Omega} f \varphi_h dv \quad (3.13)$$

qualunque sia $\varphi_h \in X_h$. Le ipotesi di ellitticità sull'operatore differenziale garantiscono l'esistenza e l'unicità della soluzione in X_h . In conseguenza della separabilità dello spazio $H_0^1(\Omega)$ e della densità di X_h in $H_0^1(\Omega)$, al tendere di h a zero, vale inoltre il seguente fondamentale risultato sull'errore di approssimazione della soluzione u del problema variazionale (3.12) con la soluzione u_h del problema $N(h)$ -dimensionale (3.13) (Teorema di Ceà [6]):

Prefissato $h \geq 0$, esiste una costante c , indipendente da h , tale che

$$\|u - u_h\| \leq c \inf_{v \in X_h} \|u - v\| \quad (3.14)$$

dove $\inf_{v \in X_h} \|u - v\|$ indica la migliore approssimazione in norma di una funzione u in X_h .

Per risolvere numericamente il sistema (3.13), indicata con $\{\varphi_i\}_{i=1}^{N(h)}$ una base per X_h e posto $u_h = \sum_{j=1}^{N(h)} \alpha_j \varphi_j$, si impone che sia soddisfatto il sistema lineare $N(h)$ -dimensionale

$$\begin{cases} \sum_{j=1}^{N(h)} \alpha_j [\kappa(\mathbf{x}) \nabla \varphi_j \cdot \nabla \varphi_i + a_0(\mathbf{x}) \varphi_j(\mathbf{x}) \varphi_i(\mathbf{x})] dv = \int_{\Omega} f(\mathbf{x}) \varphi_i(\mathbf{x}) dv \\ i = 1, 2, \dots, N(h), \end{cases} \quad (3.15)$$

ottenuto sostituendo in (3.13) u_h con il suo sviluppo e la generica φ con le funzioni di base $\{\varphi_i\}$. La determinazione del vettore $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{N(h)})^T$, soluzione del sistema (3.13), consente di approssimare la soluzione u del problema variazionale (3.12) con u_h , la quale, come evidenziato dalla (3.14), tende alla u per $h \rightarrow 0$. La rapidità con cui questo avviene dipende dall'ordine di approssimazione di una generica funzione u di $H_0^1(\Omega)$ in X_h , come vedremo in seguito. Quanto esposto fornisce 2 dei 3 passi fondamentali del metodo di Garlekin, che rappresenta il più noto dei metodi agli elementi finiti per la risoluzione di un problema di tipo ellittico. Esso si basa infatti sui seguenti 3 passi fondamentali:

- (a) formulazione variazionale (3.13) del problema (3.11);
- (b) generazione di una sequenza opportuna di spazi $\{X_h\}$ con scelta della associata base funzionale $\{\varphi_j\}_{j=1}^{N(h)}$;
- (c) risoluzione numerica del sistema lineare algebrica (3.15).

La generazione degli spazi $X(h)$ è la parte più tecnica, in quanto fortemente basata sulla geometria di Ω . Per essa ci si riferisce, solitamente, a codici di calcolo professionali, per lo più basati sul riempimento di Ω mediante triangoli, nel caso bidimensionale, e di tetraedri, nel caso tridimensionale. Nel caso bidimensionale, come verrà evidenziato successivamente in un caso semplice, si ordinano i punti interni del reticolo. Il parametro h indica il massimo diametro dei triangoli così introdotti (massima distanza tra 2 punti nodali consecutivi). Ad ogni punto nodale interno x_j si associa una funzione $\varphi_j(\mathbf{x})$, richiedendo che essa rappresenti una piramide a faccie piane che assume il valore 1 in x_j e il valore zero in ciascuno dei punti nodali dei triangoli che hanno un vertice in x_j . La base della piramide è evidentemente formata da tali triangolini. A questo punto il dominio della φ_j viene esteso a tutto il reticolo di Ω , richiedendo che la φ_j sia identicamente nulla al di fuori della base della piramide. È evidente che questo procedimento determina, in modo univoco, la φ_j , $j = 1, \dots, N(h)$. Inoltre, essendo $\varphi_j(x_i) = \delta_{ij}$, $i, j = 1, \dots, N(h)$, la funzione u_h può essere rappresentata nella seguente forma Lagrangiana:

$$u_h(\mathbf{x}) = \sum_{j=1}^{N(h)} u_h(x_j) \varphi_j(\mathbf{x}). \quad (3.16)$$

Tale scelta è comoda in quanto consente di identificare il vettore $\alpha = (\alpha_1, \dots, \alpha_{N(h)})^T$ con $\mathbf{u}_h = (u_h(x_1), \dots, u_h(x_n))$. Tenuto conto di questa

identificazione, scriviamo come usuale, il sistema (3.15) nella forma

$$A\mathbf{u}_h = \mathbf{b}, \quad (3.17)$$

dove

$$a_{ij} = a(\varphi_i, \varphi_j) = \int_{\Omega} [\kappa(\mathbf{x}) \nabla \varphi_j \cdot \nabla \varphi_i + a_0(\mathbf{x}) \varphi_i \varphi_j] dv, \quad i, j = 1, \dots, N(h),$$

$$b_i = \int_{\Omega} f(\mathbf{x}) \varphi_i(\mathbf{x}) dv, \quad i = 1, \dots, N(h), \quad \mathbf{u}_h = (u_h(x_1), \dots, u_h(x_{N(h)}))^T.$$

Come evidenziato successivamente con un esempio, la matrice A è sparsa, dato che i supporti di φ_i e φ_j hanno intersezione non vuota soltanto quando i punti nodali x_i e x_j sono entrambi vertici de triangoli appartenenti a ciascuno dei due supporti. Per questo motivo, i metodi numerici che maggiormente si usano nella risoluzione del sistema (3.17) sono di tipo iterativo.

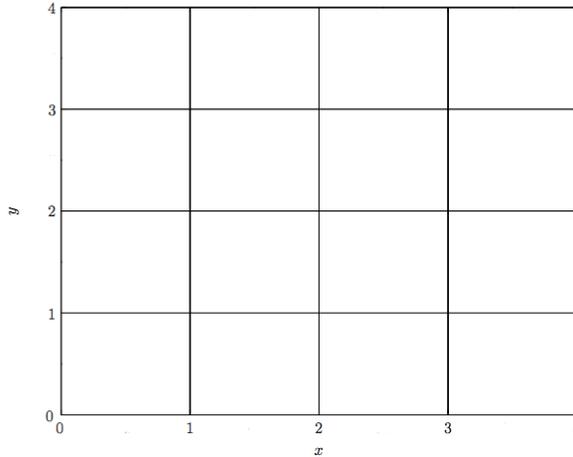


Figura 3.5: La figura mostra il dominio Ω partizionato con $x_i = 1, i = 0, 1, 2, 3, 4$ e $y_j = j, j = 0, 1, 2, 3, 4$.

Allo scopo di illustrare la tecnica, consideriamo la risoluzione del problema (3.11) nel caso particolarmente semplice in cui il dominio sia il quadrato $\Omega = \{(x, y) | 0 \leq x, y \leq 4\}$. Come evidenziato nella Figura 3.5, partizioniamo Ω considerando $x_i = 1, i = 0, 1, 2, 3, 4$ e $y_j = j, j = 0, 1, 2, 3, 4$. Numeriamo ora i punti nodali ottenuti secondo il metodo lessicografico (si procede per valori crescenti di y_j e, a parità di y_j , si ordina per x_i crescente). Si ottengono

così 25 punti nodali di cui 9 interni e 16 di frontiera. Ciascuno dei 16 quadratini ottenuti viene suddiviso in 2 triangolini, seguendo lo schema indicato nella Figura 3.7, ottenendo in tal modo 32 triangolini. Per ciascuno dei 9 punti nodali interni (7, 8, 9, 12, 13, 14, 17, 18, 19) si considera il dominio (esagono) formato dai triangolini avente tale punto come vertice. Per i punti 7, 8 e 12, per esempio, gli esagoni sono quelli indicati nella Figura 3.6.

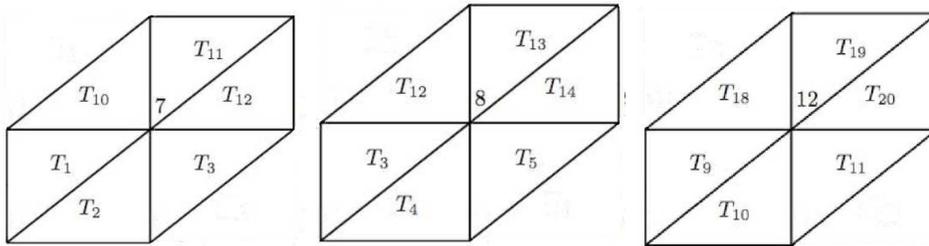


Figura 3.6: Triangoli che hanno un vertice nei punti 7, 8 e 12, rispettivamente.

A questo punto, ad ogni punto interno si associa una funzione di tipo piramidale con base esagonale e vertice nel punto interno. Nel punto interno si assegna il valore 1 e, in ciascuno degli altri punti nodali il valore zero. Essendo le facce piane, per ogni triangolino risulta determinato un piano che assume il valore 1 in un vertice e zero negli altri 2. Le funzioni così costruite sono 9, indicate rispettivamente con $\varphi_7, \varphi_8, \varphi_9, \varphi_{12}, \varphi_{13}, \varphi_{14}, \varphi_{17}, \varphi_{18}$ e φ_{19} . La φ_7 , ad esempio assume il valore 1 nel punto 7 e zero in ciascuno dei punti 1, 2, 8, 3, 12 e 6, mentre φ_{12} assume il valore 1 nel punto 12 e zero nei punti 6, 7, 13, 18, 17 e 11.

Il dominio di ognuna delle 9 funzioni φ_n considerata ($n = 7, 8, 9, 12, 13, 14, 17, 18$ e 19) viene esteso a tutto il dominio Ω (quadrato di lato 4), ponendo $\varphi_h(x, y) = 0$ per (x, y) esterno alla base esagonale della piramide da essa definita. È evidente che ogni $\varphi_h \in H_0^1(\Omega)$, dato che il suo dominio è Ω e il suo supporto è un esagono nel quale la funzione è continua e le sue derivate parziali prime presentano una semplice discontinuità di prima specie lungo i lati dei triangoli che formano l'esagono. Il modello differenziale ellittico (3.11), nel quale si assume nota la u su $\partial\Omega$, è detto di tipo Dirichlet. Nel caso sia invece assegnata su $\partial\Omega$ la derivata normale $\frac{\partial u}{\partial n}$, il problema è detto di tipo Neumann. Il modello è di tipo misto nel caso su una parte Γ_1 di $\partial\Omega$ sia assegnata la u e sulla parte complementare Γ_2 la $\frac{\partial u}{\partial n}$. La teoria garantisce che, in ciascuno di questi 3 casi, nelle ipotesi precedenti su $\kappa(\mathbf{x})$

e $a_0(\mathbf{x})$ per il modello (3.11), esiste ed è unica la soluzione in senso forte del problema.

3.5 Problema modello

Consideriamo il modello differenziale a derivate parziali del secondo ordine:

$$\begin{cases} -\nabla \cdot (\kappa(x)\nabla u) = f(x), & x \in \Omega \subset \mathbb{R}^2 \\ u = 0 & x \in \partial\Omega \end{cases}, \quad (3.18)$$

dove

$$\nabla \cdot (\kappa(x)\nabla u) = \frac{\partial}{\partial x_1} \left(\kappa(x) \frac{\partial u}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left(\kappa(x) \frac{\partial u}{\partial x_2} \right).$$

Poichè il problema è di tipo ellittico, la condizione al contorno sulla frontiera (problema di Dirichlet) assicura l'esistenza e l'unicità della soluzione.

In analogia con quanto visto nell'esempio precedente, anche in questo caso si sceglie una funzione test $v(x) \in H_0^1(\Omega)$ al fine di scrivere il problema differenziale (3.18) nella sua forma debole. Pertanto moltiplicando primo e secondo membro della 3.18 per la funzione test $v(x)$ e integrando in Ω , si richiede che l'equazione:

$$-\int_{\Omega} \nabla \cdot (\kappa(x)\nabla u)v(x)dx = \int_{\Omega} f(x)v(x)dx \quad (3.19)$$

sia soddisfatta qualunque sia $v \in H_0^1(\Omega)$. A questo punto è necessario fare uso del teorema di Green per elaborare l'integrale a primo membro, tecnica equivalente all'integrazione per parti quando le funzioni da integrare sono in più variabili. Questo permette di scrivere l'integrale a primo membro della (3.19) nel seguente modo:

$$-\int_{\Omega} \nabla \cdot (\kappa(x)\nabla u)v(x)dx = -\int_{\partial\Omega} v(x)\kappa(x)\frac{\partial u}{\partial \mathbf{n}} + \int_{\Omega} \kappa(x)\nabla u \cdot \nabla v, \quad (3.20)$$

dove \mathbf{n} rappresenta la normale uscente dalla frontiera di Ω e diretta verso l'esterno. Nel nostro caso, essendo $\Omega \subset \mathbb{R}^2$, si ha che $\partial\Omega$ è una curva e di conseguenza \mathbf{n} non è altro che il vettore normale cambiato di segno del triedro di Frénet.

Dal momento che $v(x) \in H_0^1(\Omega)$ si annulla nella frontiera, si ha che il primo integrale a secondo membro della (3.20) è nullo e pertanto la forma debole del problema differenziale (3.18) diventa:

Date $f(x), v(x) \in H_0^1(\Omega)$, cercare $u(x) \in H_0^1(\Omega)$ tale che

$$a(u, v) = (f, v), \quad (3.21)$$

dove

$$a(u, v) \stackrel{\text{def}}{=} \int_{\Omega} \kappa(x) \nabla u(x) \cdot \nabla v(x) dx, \quad (f, v) \stackrel{\text{def}}{=} \int_{\Omega} f(x) v(x) dx. \quad (3.22)$$

La soluzione del problema (3.21) si trova in uno spazio infinito dimensionale. Per la sua risoluzione numerica occorre discretizzare il sistema dato in modo da approssimare la soluzione $u(x)$ in uno spazio di dimensione finita. Al fine di illustrare la logica del metodo, supponiamo che Ω sia un quadrato. Il primo passo consiste nel costruire su di esso una discretizzazione regolare¹. La Fig. 3.7 illustra il dominio discretizzato e come si può notare la mesh è costituita da 32 triangoli T_r ($r = 1, 2 \dots 32$) e da 25 nodi n_s dove $s = 1, 2 \dots, 25$. Tra i 25 nodi, **nove sono interni** e 16 di frontiera e visto che il problema differenziale considerato in questo esempio è ellittico, si conoscono i valori della soluzione nei 16 nodi di frontiera (in quei punti essa vale zero).

Nel caso della ODE abbiamo usato la base di splines lineari definite dalla (3.4). L'analogo bidimensionale di tali splines è dato dalle **piramidi con base esagonale** di altezza unitaria. Sia $x_i = a + ih$ con $i = 0, \dots, n+1$, $h = (b-a)/(n+1)$, $x_0 = a$, $x_{n+1} = b$ la partizione lungo l'asse x della mesh in Fig. 3.7, e sia $y_j = c + jk$ con $j = 0, \dots, m+1$, $k = (d-c)/(m+1)$, $y_0 = c$, $y_{m+1} = d$ la partizione lungo l'asse y della mesh nella stessa figura. Dal momento che stiamo considerando un dominio quadrato con mesh regolare e che la sua estensione al caso rettangolare non crea alcun problema, assumiamo $n = m$, $h = k$.

Allo scopo di poter evidenziare la struttura del sistema lineare ottenuto e determinare un algoritmo efficace per la sua risoluzione, occorre sviluppare la sua notazione in modo adeguato.

Si inizia con la numerazione dei triangoli, che indicheremo con

$$\mathcal{T}_h = \{T_r : r = 1, \dots, R\}, \quad (\text{nell'esempio } R = 32)$$

e dei nodi della mesh, che indicheremo con

$$\mathcal{N}_h = \{n_s : s = 1, \dots, S\} \quad (\text{nell'esempio } S = 25).$$

¹In termini più tecnici questo passaggio viene detto **costruzione della mesh** che in questo esempio viene considerata regolare soltanto per semplicità di esposizione.

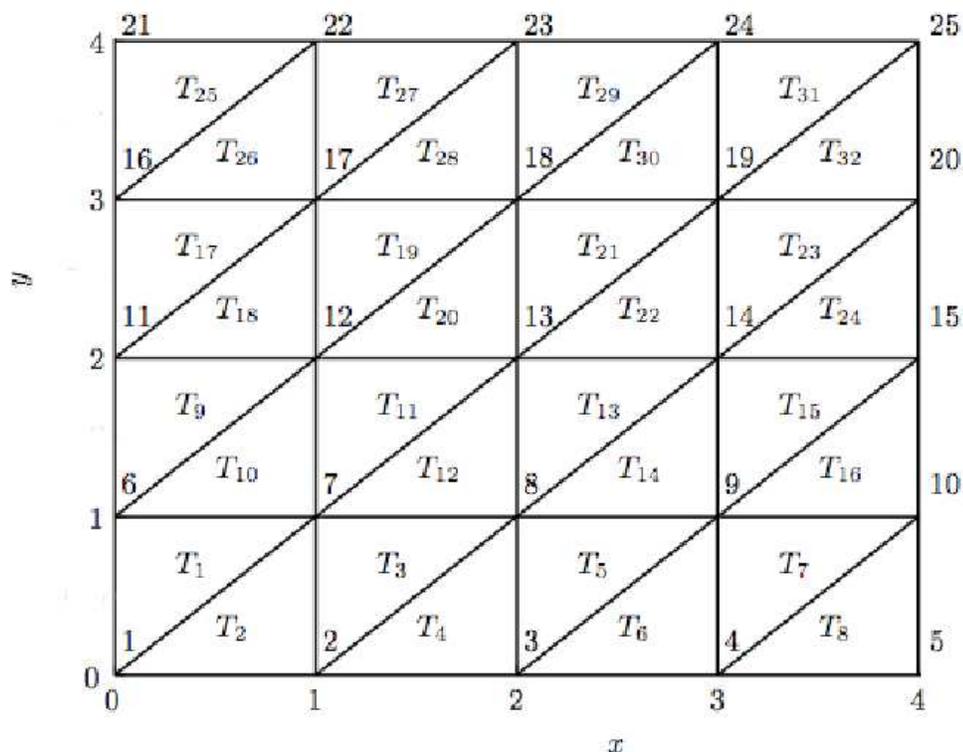


Figura 3.7: La figura rappresenta un esempio di mesh. Il dominio scelto è un quadrato, la mesh è regolare e rappresentata da 32 triangoli e 25 nodi. Dei 25 nodi, 9 sono interni e 16 sono di frontiera.

Quest'ultima notazione è ottenuta introducendo una funzione bivariata che alla coppia di indici (i, j) del generico punto (x_i, y_j) della mesh, associa un numero naturale $n(i, j)$. Nel caso rettangolare la numerazione è ottenuta con la seguente regola:

$$n(i, j) = i + 1 + (n + 2)j, \quad i = 0, 1, \dots, n + 1 \text{ e } j = 0, 1, \dots, m + 1,$$

ossia la numerazione procede dal basso verso l'alto e da sinistra verso destra (nell'esempio specifico $n = m = 3$).

Dall'insieme \mathcal{N}_h occorre estrarre il sottoinsieme dei nodi liberi (free nodes) che indicheremo con

$$\mathcal{N}_f = \{n_{f_1}, \dots, n_{f_L}\} \quad (\text{nell'esempio } L = 9).$$

Nell'esempio specifico: $n_{f_1} = 7$, $n_{f_2} = 8$, $n_{f_3} = 9$; $n_{f_4} = 12$, $n_{f_5} = 13$, $n_{f_6} = 14$; $n_{f_7} = 17$, $n_{f_8} = 18$ e $n_{f_9} = 19$.

Ad ogni nodo interno, identificato dall'indice n_{f_l} , viene associata una funzione $\phi_l(x, y)$ il cui supporto è l'esagono formato dai triangoli che hanno tale punto come loro vertice. Per il nodo caratterizzato da n_{f_1} , ad esempio, l'esagono è formato dai sei triangoli $T_1, T_2, T_3, T_{10}, T_{11}$ e T_{12} e quello identificato da n_{f_2} dai triangoli $T_3, T_4, T_5, T_{12}, T_{13}$ e T_{14} , ecc.

Nel caso più semplice, ma anche il più frequente, $\phi_l(x, y)$ è una box-spline lineare, ossia un polinomio di primo grado in (x, y) in ciascuno dei sei triangoli che definiscono il suo supporto e identicamente nulla altrove. Identificando con n_{f_l} il punto nodale della mesh che gli corrisponde, la funzione ϕ_l è univocamente caratterizzata dalla seguente condizione di cardinalità:

$$\phi_l(n_{f_j}) = \delta_{lj} = \begin{cases} 1, & j = l \\ 0, & j \neq l \end{cases} .$$

Lo span di tutte le ϕ_l , ossia lo spazio vettoriale $S_L = span\{\phi_1, \dots, \phi_L\}$ che ha come base le funzioni ϕ_l , è lo spazio degli elementi finiti associato alla mesh del dominio. Esso rappresenta il sottospazio finito-dimensionale di $H_0^1(\Omega)$, sul quale deve essere proiettata l'equazione (3.21). La forma debole del problema differenziale diventa: trovare $u \in S_L$ tale che $a(u, v) = (f, v)$, per ogni $v \in S_L$. Scriviamo la soluzione $u(x)$ come combinazione lineare degli elementi della base del sottospazio S_L

$$u(x) \approx u_L = \sum_{l=1}^L u_l \phi_l,$$

e otteniamo la proiezione su S_L del problema differenziale:

$$\sum_{l=1}^L a(\phi_l, \phi_j) u_l = (f, \phi_j), \quad j = 1, 2, \dots, L. \quad (3.23)$$

L'equazione (3.23) è un sistema lineare le cui incognite sono rappresentate dalle componenti della soluzione $u(x)$ nel sottospazio S_L . In forma più compatta la (3.23) diventa

$$\mathbf{K}\mathbf{u} = \mathbf{f},$$

dove $\mathbf{K} = a(\phi_l, \phi_j)$ e $\mathbf{f} = (f, \phi_j)$ ($l, j = 1, 2, \dots, L$) sono chiamati **stiffness matrix** e **load vector**, rispettivamente.

A questo punto è necessario scegliere esattamente il tipo di funzioni per scrivere la base $\{\phi\}_{l=1}^L$. In genere si considerano funzioni polinomiali nelle variabili x e y di grado n con $n = 1, 2, 3, \dots$, tra loro opportunamente raccordati. In questo esempio consideriamo $n = 1$ e pertanto il sottospazio finito dimensionale di $H_0^1(\Omega)$ ha come base delle piramidi con base esagonale (Fig. 3.8). Indicando con z_j le coppie di **nodi interni** (x_j, y_j) , si ha:

$$\phi_i(z_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \quad (3.24)$$

dove $i, j = 1, 2, \dots, 9$. Di conseguenza, in questo esempio la base è costituita da nove elementi $\phi_i(z_j)$ aventi ciascuno un esagono² come supporto.

3.5.1 Calcolo delle funzioni di base $\phi_l(x, y)$

Per trovare l'equazione dei piani corrispondenti ad un elemento della base, occorre tenere presente che nella frontiera del dominio Ω gli elementi della base si devono annullare e che in ogni triangolo T_i della mesh (facente parte del supporto di una funzione di base), un piano ha valore uno nel comune vertice e valore zero nei restanti vertici (Fig. 3.8). Pertanto è sempre possibile trovare l'equazione di ogni generico piano su un generico triangolo, visto che per tre punti dello spazio passa uno e un solo piano.

Calcolo di $\phi_1(x, y)$

Gli estremi del triangolo T_1 in Fig. 3.7 sono dati dai punti

$$\begin{cases} x_1 = 0 \\ y_1 = 0 \end{cases}, \quad \begin{cases} x_7 = h \\ y_7 = h \end{cases}, \quad \begin{cases} x_6 = 0 \\ y_6 = h \end{cases},$$

per cui, essendo $z(x, y) = a + bx + cy$ l'equazione di un generico piano in \mathbb{R}^3 , si ha

$$\begin{cases} z(x_1, y_1) = 0 \\ z(x_7, y_7) = 1 \\ z(x_6, y_6) = 0 \end{cases} \Rightarrow \begin{cases} a = 0 \\ a + hb + hc = 1 \\ a + hc = 0 \end{cases} \Rightarrow \begin{cases} a = 0 \\ b = \frac{1}{h} \\ c = 0 \end{cases},$$

per cui l'equazione del piano cercato è $z = x/h$.

²Esso è a sua volta dato dall'unione di sei triangoli T_i della mesh in Fig. 3.7 aventi in comune il vertice nel quale la spline assume il valore 1.

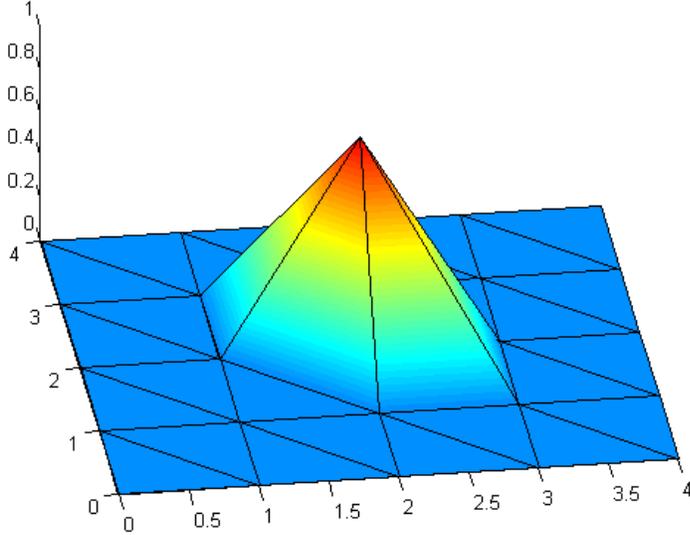


Figura 3.8: La figura rappresenta un esempio di base polinomiale di primo grado in un dominio esagonale. Da notare in particolare come l'unione dei sei piani aventi valore uno nel vertice comune, formano una piramide a base esagonale.

Procedendo in modo analogo, possiamo trovare l'equazione del piano con supporto il triangolo T_2 :

$$\begin{cases} x_1 = 0 \\ y_1 = 0 \end{cases}, \begin{cases} x_2 = h \\ y_2 = 0 \end{cases}, \begin{cases} x_7 = h \\ y_7 = h \end{cases},$$

$$\begin{cases} z(x_1, y_1) = 0 \\ z(x_2, y_2) = 0 \\ z(x_7, y_7) = 1 \end{cases} \Rightarrow \begin{cases} z(0, 0) = 0 \\ z(h, 0) = 0 \\ z(h, h) = 1 \end{cases} \Rightarrow \begin{cases} a = 0 \\ a + hb = 0 \\ a + hb + hc = 1 \end{cases} \Rightarrow \begin{cases} a = 0 \\ b = 0 \\ c = \frac{1}{h} \end{cases}$$

La sua equazione è pertanto $z = y/h$.

Piano con supporto il triangolo T_3 :

$$\begin{cases} x_2 = h \\ y_2 = 0 \end{cases}, \begin{cases} x_8 = 2h \\ y_8 = h \end{cases}, \begin{cases} x_7 = h \\ y_7 = h \end{cases},$$

$$\begin{cases} z(x_2, y_2) = 0 \\ z(x_8, y_8) = 0 \\ z(x_7, y_7) = 1 \end{cases} \Rightarrow \begin{cases} z(h, 0) = 0 \\ z(2h, h) = 0 \\ z(h, h) = 1 \end{cases} \Rightarrow \begin{cases} a + hb = 0 \\ a + 2hb + hc = 0 \\ a + hb + hc = 1 \end{cases}$$

$$\begin{pmatrix} 1 & h & 0 \\ 1 & 2h & h \\ 1 & h & h \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1/h & -1/h \\ -1/h & 0 & 1/h \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

La sua equazione è $z = 1 - x/h + y/h$.

Piano con supporto il triangolo T_{12} :

$$\begin{cases} x_7 = h \\ y_7 = h \end{cases}, \begin{cases} x_8 = 2h \\ y_8 = h \end{cases}, \begin{cases} x_{13} = 2h \\ y_{13} = 2h \end{cases},$$

$$\begin{cases} z(x_7, y_7) = 1 \\ z(x_8, y_8) = 0 \\ z(x_{13}, y_{13}) = 0 \end{cases} \Rightarrow \begin{cases} z(h, h) = 1 \\ z(2h, h) = 0 \\ z(2h, 2h) = 0 \end{cases} \Rightarrow \begin{cases} a + hb + hc = 1 \\ a + 2hb + hc = 0 \\ a + 2hb + 2hc = 0 \end{cases}$$

$$\begin{pmatrix} 1 & h & h \\ 1 & 2h & h \\ 1 & 2h & 2h \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 2 & 0 & -1 \\ -1/h & 1/h & 0 \\ 0 & -1/h & 1/h \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

L'equazione del piano è dunque $z = 2 - x/h$.

Piano con supporto il triangolo T_{11} :

$$\begin{cases} x_7 = h \\ y_7 = h \end{cases}, \begin{cases} x_{13} = 2h \\ y_{13} = 2h \end{cases}, \begin{cases} x_{12} = h \\ y_{12} = 2h \end{cases},$$

$$\begin{cases} z(x_7, y_7) = 1 \\ z(x_{13}, y_{13}) = 0 \\ z(x_{12}, y_{12}) = 0 \end{cases} \Rightarrow \begin{cases} z(h, h) = 1 \\ z(2h, 2h) = 0 \\ z(h, 2h) = 0 \end{cases} \Rightarrow \begin{cases} a + hb + hc = 1 \\ a + 2hb + 2hc = 0 \\ a + hb + 2hc = 0 \end{cases}$$

$$\begin{pmatrix} 1 & h & h \\ 1 & 2h & 2h \\ 1 & h & 2h \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1/h & -1/h \\ -1/h & 0 & 1/h \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Tale piano ha come equazione è $z = 2 - y/h$.

Piano con supporto il triangolo T_{10} :

$$\begin{cases} x_6 = 0 \\ y_6 = h \end{cases}, \begin{cases} x_7 = h \\ y_7 = h \end{cases}, \begin{cases} x_{12} = h \\ y_{12} = 2h \end{cases},$$

$$\begin{cases} z(x_6, y_6) = 0 \\ z(x_7, y_7) = 1 \\ z(x_{12}, y_{12}) = 0 \end{cases} \Rightarrow \begin{cases} z(0, h) = 0 \\ z(h, h) = 1 \\ z(h, 2h) = 0 \end{cases} \Rightarrow \begin{cases} a + hc = 0 \\ a + hb + hc = 1 \\ a + hb + 2hc = 0 \end{cases}$$

$$\begin{pmatrix} 1 & 0 & h \\ 1 & h & h \\ 1 & h & 2h \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ -1/h & 1/h & 0 \\ 0 & -1/h & 1/h \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

La sua equazione è dunque $z = 1 + x/h - y/h$.

Pertanto la funzione di base $\phi_1(x, y)$ è data da (si veda Fig. 3.9):

$$\phi_1(x, y) = \begin{cases} x/h & (x, y) \in T_1 \\ y/h & (x, y) \in T_2 \\ 1 - x/h + y/h & (x, y) \in T_3 \\ 2 - x/h & (x, y) \in T_{12} \\ 2 - y/h & (x, y) \in T_{11} \\ 1 + x/h - y/h & (x, y) \in T_{10} \end{cases}. \quad (3.25)$$

Calcolo di $\phi_2(x, y)$

- Piano con supporto il triangolo T_3 .
Tale piano ha la stessa inclinazione del piano avente come supporto il triangolo T_1 e interseca l'asse z nel punto $(0, 0, 1)$. Pertanto la sua equazione è data da (si veda la (3.25)) $z = 1 + x/h$.
- Piano con supporto il triangolo T_4 .
Tale piano ha la stessa equazione del piano avente come supporto il triangolo T_2 e pertanto (si veda la (3.25)) si ha: $z = y/h$.
- Piano con supporto il triangolo T_5 .
Tale piano ha la stessa inclinazione del piano avente come supporto il triangolo T_3 e interseca l'asse z nel punto $(0, 0, 2)$. Pertanto la sua equazione è data da (si veda la (3.25)) $z = 2 - x/h + y/h$.
- Piano con supporto il triangolo T_{14} .
Tale piano ha la stessa inclinazione del piano avente come supporto il triangolo T_{12} e interseca l'asse z nel punto $(0, 0, 3)$. Pertanto la sua equazione è data da (si veda la (3.25)) $z = 3 - x/h$.
- Piano con supporto il triangolo T_{13} .
Tale piano ha la stessa equazione del piano avente come supporto il triangolo T_{11} e pertanto (si veda la (3.25)) si ha: $z = 2 - y/h$.

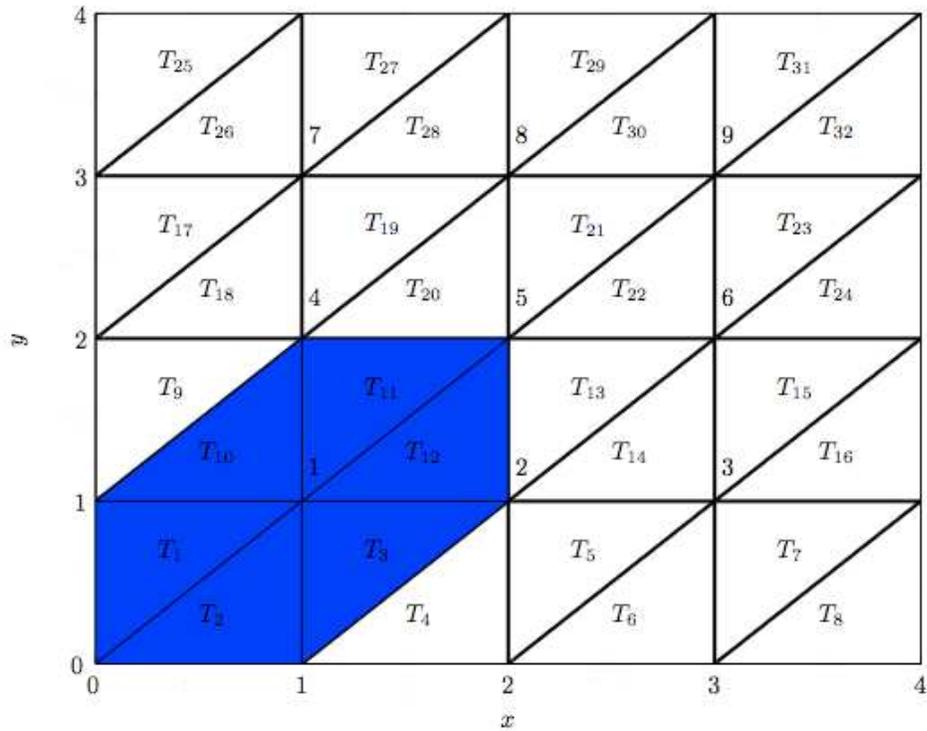


Figura 3.9: La figura rappresenta la mesh considerata in questo esempio dove sono stati rappresentati solo i nodi interni e dove il supporto della funzione base $\phi_1(x, y)$ è evidenziato in blu.

- Piano con supporto il triangolo T_{12} .
Tale piano ha la stessa inclinazione del piano avente come supporto il triangolo T_{10} e interseca l'asse z nel punto $(0, 0, 0)$. Pertanto la sua equazione è data da (si veda la (3.25)) $z = x/h - y/h$.

Pertanto la funzione di base $\phi_2(x, y)$ è data da:

$$\phi_2(x, y) = \begin{cases} 1 + x/h & (x, y) \in T_3 \\ y/h & (x, y) \in T_4 \\ 2 - x/h + y/h & (x, y) \in T_5 \\ 3 - x/h & (x, y) \in T_{14} \\ 2 - y/h & (x, y) \in T_{13} \\ x/h - y/h & (x, y) \in T_{12} \end{cases}. \quad (3.26)$$

Calcolo delle restanti funzioni base

Sfruttando la simmetria della mesh in Fig. 3.7 e la linearità delle funzioni base, si trova facilmente che

$$\phi_3(x, y) = \begin{cases} -2 + x/h & (x, y) \in T_5 \\ y/h & (x, y) \in T_6 \\ 3 - x/h + y/h & (x, y) \in T_7 \\ 4 - x/h & (x, y) \in T_{16} \\ 2 - y/h & (x, y) \in T_{15} \\ -1 + x/h - y/h & (x, y) \in T_{14} \end{cases}. \quad (3.27)$$

$$\phi_4(x, y) = \begin{cases} x/h & (x, y) \in T_9 \\ -1 + y/h & (x, y) \in T_{10} \\ -x/h + y/h & (x, y) \in T_{11} \\ 2 - x/h & (x, y) \in T_{20} \\ 3 - y/h & (x, y) \in T_{19} \\ 2 + x/h - y/h & (x, y) \in T_{18} \end{cases}. \quad (3.28)$$

$$\phi_5(x, y) = \begin{cases} -1 + x/h & (x, y) \in T_{11} \\ -1 + y/h & (x, y) \in T_{12} \\ 1 - x/h + y/h & (x, y) \in T_{13} \\ 3 - x/h & (x, y) \in T_{22} \\ 3 - y/h & (x, y) \in T_{21} \\ 1 + x/h - y/h & (x, y) \in T_{20} \end{cases}. \quad (3.29)$$

$$\phi_6(x, y) = \begin{cases} -2 + x/h & (x, y) \in T_{13} \\ -1 + y/h & (x, y) \in T_{14} \\ 2 - x/h + y/h & (x, y) \in T_{15} \\ 4 - x/h & (x, y) \in T_{24} \\ 3 - y/h & (x, y) \in T_{23} \\ x/h - y/h & (x, y) \in T_{22} \end{cases} \quad (3.30)$$

$$\phi_7(x, y) = \begin{cases} x/h & (x, y) \in T_{17} \\ -2 + y/h & (x, y) \in T_{18} \\ -1 - x/h + y/h & (x, y) \in T_{19} \\ 2 - x/h & (x, y) \in T_{28} \\ 4 - y/h & (x, y) \in T_{27} \\ 3 + x/h - y/h & (x, y) \in T_{26} \end{cases} \quad (3.31)$$

$$\phi_8(x, y) = \begin{cases} -1 + x/h & (x, y) \in T_{19} \\ -2 + y/h & (x, y) \in T_{20} \\ -x/h + y/h & (x, y) \in T_{21} \\ 3 - x/h & (x, y) \in T_{30} \\ 4 - y/h & (x, y) \in T_{29} \\ 2 + x/h - y/h & (x, y) \in T_{28} \end{cases} \quad (3.32)$$

$$\phi_9(x, y) = \begin{cases} -2 + x/h & (x, y) \in T_{21} \\ -2 + y/h & (x, y) \in T_{22} \\ 1 - x/h + y/h & (x, y) \in T_{23} \\ 4 - x/h & (x, y) \in T_{32} \\ 4 - y/h & (x, y) \in T_{31} \\ 1 + x/h - y/h & (x, y) \in T_{30} \end{cases} \quad (3.33)$$

3.5.2 Calcolo della stiffness matrix e del load vector

Per calcolare la stiffness matrix \mathbf{K} occorre fare uso della seguente relazione

$$K_{ij} = a(\phi_i, \phi_j) = \int_{\Omega} k(x, y) \nabla \phi_i(x, y) \cdot \nabla \phi_j(x, y) dx dy. \quad (3.34)$$

Supponendo che il mezzo sia omogeneo ($k(x, y) = k = 1$), si ha

$$\begin{aligned}
K_{11} &= a(\phi_1, \phi_1) = \int_{\Omega} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy \\
&= \int_{T_1} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy + \int_{T_2} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy \\
&+ \int_{T_3} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy + \int_{T_{10}} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy \\
&+ \int_{T_{11}} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy + \int_{T_{12}} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy,
\end{aligned}$$

in quanto il supporto della funzione base $\phi_1(x, y)$ è $T_1 \cup T_2 \cup T_3 \cup T_{10} \cup T_{11} \cup T_{12}$.

Dal momento che $\nabla \phi_1(x, y) = (1/h, 0)^T$ quando $(x, y) \in T_1$, si ha

$$\int_{T_1} \frac{1}{h^2} dx dy = \frac{1}{h^2} \frac{h^2}{2} = \frac{1}{2}.$$

Analogamente

$$(x, y) \in T_2 \Rightarrow \nabla \phi_1(x, y) = \begin{pmatrix} 0 \\ 1/h \end{pmatrix} \Rightarrow \int_{T_2} \frac{1}{h^2} dx dy = \frac{1}{h^2} \frac{h^2}{2} = \frac{1}{2},$$

$$(x, y) \in T_3 \Rightarrow \nabla \phi_1(x, y) = \begin{pmatrix} -1/h \\ 1/h \end{pmatrix} \Rightarrow \int_{T_3} \frac{2}{h^2} dx dy = \frac{2}{h^2} \frac{h^2}{2} = 1.$$

Per simmetria si vede facilmente che

$$\int_{T_{10}} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy = \int_{T_3} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy = 1, \tag{3.35a}$$

$$\int_{T_{11}} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy = \int_{T_2} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy = \frac{1}{2}, \tag{3.35b}$$

$$\int_{T_{12}} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy = \int_{T_1} \nabla \phi_1(x, y) \cdot \nabla \phi_1(x, y) dx dy = \frac{1}{2}, \tag{3.35c}$$

e quindi

$$K_{11} = 2 \left(1 + \frac{1}{2} + \frac{1}{2} \right) = 4. \tag{3.36}$$

Per il calcolo dell'entrata K_{12} notiamo che i supporti delle funzioni base $\phi_1(x, y)$ e $\phi_2(x, y)$ si intersecano nei triangoli T_3 e T_{12} :

$$\text{supp } \phi_1(x, y) \cap \text{supp } \phi_2(x, y) = T_3 \cup T_{12}.$$

Pertanto si ha

$$(x, y) \in T_3 \Rightarrow \begin{cases} \nabla \phi_1(x, y) = \begin{pmatrix} -1/h \\ 1/h \end{pmatrix} \\ \nabla \phi_2(x, y) = \begin{pmatrix} 1/h \\ 0 \end{pmatrix} \end{cases} \Rightarrow \int_{T_3} -\frac{1}{h^2} dx dy = -\frac{1}{h^2} \frac{h^2}{2} = -\frac{1}{2},$$

$$(x, y) \in T_{12} \Rightarrow \begin{cases} \nabla \phi_1(x, y) = \begin{pmatrix} -1/h \\ 0 \end{pmatrix} \\ \nabla \phi_2(x, y) = \begin{pmatrix} 1/h \\ -1/h \end{pmatrix} \end{cases} \Rightarrow \int_{T_{12}} -\frac{1}{h^2} dx dy = -\frac{1}{h^2} \frac{h^2}{2} = -\frac{1}{2},$$

e quindi

$$K_{12} = -1. \quad (3.37)$$

L'entrata $K_{13} = 0$ in quanto

$$\text{supp } \phi_1(x, y) \cap \text{supp } \phi_3(x, y) = \emptyset,$$

mentre le entrate K_{14} e K_{15} possono essere diverse da zero in quanto

$$\text{supp } \phi_1(x, y) \cap \text{supp } \phi_4(x, y) = T_{10} \cup T_{11},$$

$$\text{supp } \phi_1(x, y) \cap \text{supp } \phi_5(x, y) = T_{11} \cup T_{12}.$$

Tenendo conto delle funzioni base e delle loro relazioni di simmetria, si trova facilmente che la stiffness matrix è una matrice sparsa avente la seguente forma

$$\mathbf{K} = \begin{pmatrix} 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 4 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 4 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 4 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 \end{pmatrix}. \quad (3.38)$$

Il calcolo del load vector $\mathbf{f} = (f, \phi_i) \ i = 1, 2, \dots, 9$, dipende fortemente dalla forma del termine noto della PDE (3.18). In ogni caso, si procede in modo del tutto analogo a quanto visto per la stiffness matrix. Per esempio, per ottenere la prima entrata di \mathbf{f} , si procede nel seguente modo

$$\begin{aligned} f_1 = (f, \phi_1) &= \int_{\Omega} f(x, y) \phi_1(x, y) dx dy \\ &= \int_{T_1} f(x, y) \frac{x}{h} dx dy + \int_{T_2} f(x, y) \frac{y}{h} dx dy \\ &+ \int_{T_3} f(x, y) \left(1 - \frac{x}{h} + \frac{y}{h}\right) dx dy + \int_{T_{10}} f(x, y) \left(1 + \frac{x}{h} - \frac{y}{h}\right) dx dy \\ &+ \int_{T_{11}} f(x, y) \left(2 - \frac{y}{h}\right) dx dy + \int_{T_{12}} f(x, y) \left(2 - \frac{x}{h}\right) dx dy, \end{aligned}$$

e facendo uso delle restanti funzioni base $\phi_i(x, y)$ si trovano formule analoghe per il calcolo delle rimanenti entrate del load vector.

3.5.3 Come usare un triangolo di riferimento nel FEM

Nella pratica non si calcolano mai le entrate della stiffness matrix e del load vector facendo direttamente ricorso alla forma esplicita della funzioni base $\{\phi_i(x, y)\}_{i=1}^9$, perché nei problemi concreti la mesh è spesso costituita da un migliaio di elementi e non è computazionalmente comodo calcolare un migliaio di funzioni base da usare per trovare la matrice \mathbf{K} e il vettore \mathbf{f} . Per ottimizzare il procedimento sin'ora visto, si utilizza un triangolo di riferimento T_R avente vertici $(0, 0)$, $(0, 1)$ e $(1, 0)$ e si definisce una trasformazione biunivoca che collega T_R con un generico triangolo T della mesh in Fig. 3.7 avente vertici generici (x_1, y_1) , (x_2, y_2) e (x_3, y_3) nel seguente modo:

$$(0, 0) \rightarrow (x_1, y_1), \quad (0, 1) \rightarrow (x_2, y_2), \quad (1, 0) \rightarrow (x_3, y_3). \quad (3.39)$$

Si tratta quindi di trovare i coefficienti a_1, a_2, a_3, b_1, b_2 e b_3 di una trasformazione lineare:

$$\begin{cases} x = a_1 + a_2 s + a_3 t \\ y = b_1 + b_2 s + b_3 t \end{cases} .$$

Applicando le condizioni (3.39) si trova facilmente che la trasformazione cercata è data da

$$\begin{cases} x = x_1 + (x_2 - x_1)s + (x_3 - x_1)t \\ y = y_1 + (y_2 - y_1)s + (y_3 - y_1)t \end{cases} . \quad (3.40)$$

La trasformazione (3.40) permette mettere in relazione biunivoca una funzione definita su T con la corrispondente funzione definita su T_R : se $f(x, y)$ è definita su un triangolo generico della mesh, allora la sua controparte definita sul triangolo di riferimento T_R è data da

$$g(s, t) = f[x_1 + (x_2 - x_1)s + (x_3 - x_1)t, y_1 + (y_2 - y_1)s + (y_3 - y_1)t],$$

pertanto la funzione $g(s, t)$ ha lo stesso valore della funzione $f(x, y)$ se i punti $(s, t) \in T_R$ e $(x, y) \in T$ sono collegati dalla trasformazione (3.40).

Ai tre vertici di un generico triangolo T della mesh corrispondono tre funzioni base ϕ_i $i = 1, 2, 3$, tali che

$$\phi_i(x_j, y_j) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}, \quad (3.41)$$

e, mediante la (3.40), possiamo scrivere le controparti delle (3.41) su T_R :

$$\gamma_i(s, t) = \phi_i[x_1 + (x_2 - x_1)s + (x_3 - x_1)t, y_1 + (y_2 - y_1)s + (y_3 - y_1)t].$$

Dal momento che i vertici di T_R sono in corrispondenza biunivoca con i vertici di T , anche le funzioni γ_i soddisfano alla condizione (3.41)

$$\gamma_i(s_j, t_j) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}. \quad (3.42)$$

Dal momento che la trasformazione (3.40) e le funzioni $\phi_i(x, y)$ sono lineari, si trova facilmente che

$$\gamma_1(s, t) = 1 - s - t \quad (3.43a)$$

$$\gamma_2(s, t) = s \quad (3.43b)$$

$$\gamma_3(s, t) = t \quad (3.43c)$$

. Di conseguenza, nel calcolo della stiffness matrix \mathbf{K} e del load vector \mathbf{f} possiamo usare le funzioni base (3.43) definite su T_R invece di calcolare le funzioni base $\phi_1(x, y)$, $\phi_2(x, y)$ e $\phi_3(x, y)$ su ogni triangolo della mesh. A questo scopo occorre fare ricorso alle seguenti formule di cambiamento di variabili sotto il segno di integrale:

$$\int_T k \nabla \phi_i \cdot \nabla \phi_j dx dy = \int_{T_R} \tilde{k} (J^{-T} \nabla \gamma_i) \cdot (J^{-T} \nabla \gamma_j) |\det J| ds dt, \quad (3.44)$$

$$\int_T f(x, y) \phi_i(x, y) dx dy = \int_{T_R} \tilde{f}(s, t) \gamma_i(s, t) |\det J| ds dt, \quad (3.45)$$

dove J è la matrice Jacobiana della trasformazione (3.40)

$$J = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}, \quad (3.46)$$

e $J^{-T} = [J^{-1}]^T$.

Calcoliamo ora la stiffness matrix \mathbf{K} facendo uso della formula (3.44) e ipotizzando che il mezzo sia omogeneo ($k(x, y) = 1 \Rightarrow \tilde{k}(s, t) = 1$):

$$\begin{aligned} K_{11} = & \int_{T_1} \nabla \phi_1 \cdot \nabla \phi_1 dx dy + \int_{T_2} \nabla \phi_1 \cdot \nabla \phi_1 dx dy + \int_{T_3} \nabla \phi_1 \cdot \nabla \phi_1 dx dy \\ & + \int_{T_{10}} \nabla \phi_1 \cdot \nabla \phi_1 dx dy + \int_{T_{11}} \nabla \phi_1 \cdot \nabla \phi_1 dx dy + \int_{T_{12}} \nabla \phi_1 \cdot \nabla \phi_1 dx dy. \end{aligned} \quad (3.47)$$

Per calcolare l'integrale esteso al triangolo della mesh T_1 notiamo che i suoi vertici sono

$$\begin{cases} x_1 = 0 \\ y_1 = 0 \end{cases}, \begin{cases} x_2 = h \\ y_2 = h \end{cases}, \begin{cases} x_3 = 0 \\ y_3 = h \end{cases},$$

e la funzione base $\phi_1(x, y)$ vale uno nel vertice (x_2, y_2) , a cui pertanto corrisponde la funzione base $\gamma_2(s, t) = s$. Si ha:

$$J = \begin{pmatrix} h & 0 \\ h & h \end{pmatrix}, \quad J^{-T} = \frac{1}{h^2} \begin{pmatrix} h & -h \\ 0 & h \end{pmatrix}, \quad |\det J| = h^2. \quad (3.48)$$

$$\nabla \gamma_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad J^{-T} \nabla \gamma_2 = \frac{1}{h^2} \begin{pmatrix} h & -h \\ 0 & h \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/h \\ 0 \end{pmatrix}. \quad (3.49)$$

Facendo uso delle relazioni in (3.48) e (3.49), si ha

$$\begin{aligned} \int_{T_1} \nabla \phi_1 \cdot \nabla \phi_1 dx dy &= \int_{T_R} (J^{-T} \nabla \gamma_2) \cdot (J^{-T} \nabla \gamma_2) |\det J| ds dt \\ &= \int_0^1 \int_0^{1-s} \frac{1}{h^2} h^2 ds dt \\ &= \int_0^1 \left(\int_0^{1-s} dt \right) ds = \int_0^1 (1-s) ds = \left(1 - \frac{1}{2} \right) = \frac{1}{2}. \end{aligned}$$

Visto che la trasformazione (3.40) e le funzioni base $\gamma_i(s, t)$ $i = 1, 2, 3$, nella (3.43), sono lineari, il termine $(J^{-T}\nabla\gamma_i) \cdot (J^{-T}\nabla\gamma_j) |\det J|$ non dipende mai dalle variabili (s, t) e pertanto la formula (3.44) diventa:

$$\int_T \nabla\phi_i \cdot \nabla\phi_j dx dy = \frac{1}{2} (J^{-T}\nabla\gamma_i) \cdot (J^{-T}\nabla\gamma_j) |\det J|, \quad (3.50)$$

dove $\int_{T_R} ds dt = 1/2$ perché è l'area del triangolo di riferimento T_R .

Per calcolare l'integrale esteso al triangolo della mesh T_2 nel secondo membro della relazione (3.47), notiamo che i suoi vertici sono

$$\begin{cases} x_1 = 0 \\ y_1 = 0 \end{cases}, \begin{cases} x_2 = h \\ y_2 = 0 \end{cases}, \begin{cases} x_3 = h \\ y_3 = h \end{cases},$$

e la funzione base $\phi_1(x, y)$ vale uno nel vertice (x_3, y_3) , a cui pertanto corrisponde la funzione base $\gamma_3(s, t) = t$. Si ha:

$$J = \begin{pmatrix} h & h \\ 0 & h \end{pmatrix}, \quad J^{-T} = \frac{1}{h^2} \begin{pmatrix} h & 0 \\ -h & h \end{pmatrix}, \quad |\det J| = h^2. \quad (3.51)$$

$$\nabla\gamma_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad J^{-T}\nabla\gamma_3 = \frac{1}{h^2} \begin{pmatrix} h & 0 \\ -h & h \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/h \end{pmatrix}. \quad (3.52)$$

Facendo uso delle relazioni in 3.52 e usando la formula (3.50), si ha

$$\int_{T_2} \nabla\phi_1 \cdot \nabla\phi_1 dx dy = \frac{1}{2} (J^{-T}\nabla\gamma_3) \cdot (J^{-T}\nabla\gamma_3) |\det J| = \frac{1}{2} \frac{1}{h^2} h^2 = \frac{1}{2}.$$

Per calcolare l'integrale esteso al triangolo della mesh T_3 nel secondo membro della relazione (3.47), notiamo che i suoi vertici sono

$$\begin{cases} x_1 = h \\ y_1 = 0 \end{cases}, \begin{cases} x_2 = 2h \\ y_2 = h \end{cases}, \begin{cases} x_3 = h \\ y_3 = h \end{cases},$$

e la funzione base $\phi_1(x, y)$ vale uno nel vertice (x_3, y_3) , a cui pertanto corrisponde la funzione base $\gamma_3(s, t) = t$. Si ha:

$$J = \begin{pmatrix} h & 0 \\ h & h \end{pmatrix}, \quad J^{-T} = \frac{1}{h^2} \begin{pmatrix} h & -h \\ 0 & h \end{pmatrix}, \quad |\det J| = h^2. \quad (3.53)$$

$$\nabla\gamma_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad J^{-T}\nabla\gamma_3 = \frac{1}{h^2} \begin{pmatrix} h & -h \\ 0 & h \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/h \\ 1/h \end{pmatrix}. \quad (3.54)$$

Facendo uso delle relazioni in (3.54) e usando la formula (3.50), si ha

$$\int_{T_3} \nabla \phi_1 \cdot \nabla \phi_1 dx dy = \frac{1}{2} (J^{-T} \nabla \gamma_3) \cdot (J^{-T} \nabla \gamma_3) |\det J| = \frac{1}{2} \frac{2}{h^2} h^2 = 1.$$

Per simmetria sappiamo che valgono le uguaglianze (3.35), ma è istruttivo verificarle usando il metodo del triangolo di riferimento.

Per calcolare l'integrale esteso al triangolo della mesh T_{10} nel secondo membro della relazione (3.47), notiamo che i suoi vertici sono

$$\begin{cases} x_1 = 0 \\ y_1 = h \end{cases}, \begin{cases} x_2 = h \\ y_2 = h \end{cases}, \begin{cases} x_3 = h \\ y_3 = 2h \end{cases},$$

e la funzione base $\phi_1(x, y)$ vale uno nel vertice (x_2, y_2) , a cui pertanto corrisponde la funzione base $\gamma_2(s, t) = s$. Si ha:

$$J = \begin{pmatrix} h & h \\ 0 & h \end{pmatrix}, J^{-T} = \frac{1}{h^2} \begin{pmatrix} h & 0 \\ -h & h \end{pmatrix}, |\det J| = h^2. \quad (3.55)$$

$$\nabla \gamma_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, J^{-T} \nabla \gamma_2 = \frac{1}{h^2} \begin{pmatrix} h & 0 \\ -h & h \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/h \\ -1/h \end{pmatrix}. \quad (3.56)$$

Facendo uso delle relazioni in (3.56) e usando la formula (3.50), si ha

$$\int_{T_{10}} \nabla \phi_1 \cdot \nabla \phi_1 dx dy = \frac{1}{2} (J^{-T} \nabla \gamma_1) \cdot (J^{-T} \nabla \gamma_1) |\det J| = \frac{1}{2} \frac{2}{h^2} h^2 = 1.$$

Per calcolare l'integrale esteso al triangolo della mesh T_{11} nel secondo membro della relazione (3.47), notiamo che i suoi vertici sono

$$\begin{cases} x_1 = h \\ y_1 = h \end{cases}, \begin{cases} x_2 = 2h \\ y_2 = 2h \end{cases}, \begin{cases} x_3 = h \\ y_3 = 2h \end{cases},$$

e la funzione base $\phi_1(x, y)$ vale uno nel vertice (x_1, y_1) , a cui pertanto corrisponde la funzione base $\gamma_1(s, t) = 1 - s - t$. Si ha:

$$J = \begin{pmatrix} h & 0 \\ h & h \end{pmatrix}, J^{-T} = \frac{1}{h^2} \begin{pmatrix} h & -h \\ 0 & h \end{pmatrix}, |\det J| = h^2. \quad (3.57)$$

$$\nabla \gamma_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, J^{-T} \nabla \gamma_1 = \frac{1}{h^2} \begin{pmatrix} h & -h \\ 0 & h \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1/h \end{pmatrix}. \quad (3.58)$$

Facendo uso delle relazioni in (3.60) e usando la formula (3.50), si ha

$$\int_{T_{10}} \nabla \phi_1 \cdot \nabla \phi_1 dx dy = \frac{1}{2} (J^{-T} \nabla \gamma_1) \cdot (J^{-T} \nabla \gamma_1) |\det J| = \frac{1}{2} \frac{1}{h^2} h^2 = \frac{1}{2}.$$

Per calcolare l'integrale esteso al triangolo della mesh T_{12} nel secondo membro della relazione (3.47), notiamo che i suoi vertici sono

$$\begin{cases} x_1 = h \\ y_1 = h \end{cases}, \begin{cases} x_2 = 2h \\ y_2 = h \end{cases}, \begin{cases} x_3 = 2h \\ y_3 = 2h \end{cases},$$

e la funzione base $\phi_1(x, y)$ vale uno nel vertice (x_1, y_1) , a cui pertanto corrisponde la funzione base $\gamma_1(s, t) = 1 - s - t$. Si ha:

$$J = \begin{pmatrix} h & h \\ 0 & h \end{pmatrix}, \quad J^{-T} = \frac{1}{h^2} \begin{pmatrix} h & 0 \\ -h & h \end{pmatrix}, \quad |\det J| = h^2. \quad (3.59)$$

$$\nabla \gamma_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad J^{-T} \nabla \gamma_1 = \frac{1}{h^2} \begin{pmatrix} h & 0 \\ -h & h \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1/h \\ 0 \end{pmatrix}. \quad (3.60)$$

Facendo uso delle relazioni in (3.60) e usando la formula (3.50), si ha

$$\int_{T_{10}} \nabla \phi_1 \cdot \nabla \phi_1 dx dy = \frac{1}{2} (J^{-T} \nabla \gamma_1) \cdot (J^{-T} \nabla \gamma_1) |\det J| = \frac{1}{2} \frac{1}{h^2} h^2 = \frac{1}{2}.$$

Sommando i contributi calcolati, si trova $K_{11} = 4$ come ci aspettiamo (si veda la (3.38)).

Il calcolo delle altre entrate della stiffness matrix e del load vector sono esattamente analoghi e lo lasciamo come esercizio. Nel caso del load vector la formula (3.45) si semplifica in quanto lo Jacobiano non dipende dalle variabili di integrazione, ma a differenza del calcolo della matrice \mathbf{K} , non è possibile dare una formula esplicita simile alla (3.50), in quanto si hanno vari casi in base alla forma del termine noto $f(x, y)$ della PDE (3.18).

Appendice A

Serie di Fourier

A.1 Funzioni periodiche e polinomi trigonometrici

Definizione A.1.1 Una funzione $f(x)$ definita in un dominio D è detta periodica di periodo T se, qualunque sia $x \in D$, $x + T$ appartiene a D ed inoltre $f(x) = f(x + T)$.

In tal caso T è il periodo e $1/T$ è la frequenza della f . Se $f(x)$ è periodica con periodo T , lo è anche di periodo $2T, 3T, \dots$. Per questo motivo, talvolta si preferisce precisare che T è il *periodo fondamentale*.

Per esempio le funzioni $\sin x$ e $\cos x$ sono periodiche di periodo 2π , mentre $\sin \omega x$ e $\cos \omega x$ ($\omega \neq 0$), sono periodiche di periodo $2\pi/\omega$. Se k è un intero, anche $\sin k\omega x$ e $\cos k\omega x$ sono periodiche con periodo $2\pi/\omega$.

Spesso, in vari contesti applicativi, una funzione definita in un intervallo viene estesa per periodicità a tutto \mathbb{R} . Ad esempio, la funzione

$$f(x) = \begin{cases} x & 0 \leq x < 1 \\ 2 - x & 1 \leq x < 2 \end{cases},$$

può essere estesa per periodicità su \mathbb{R} , ponendo $f(x) = f(x + 2)$.

Armoniche elementari.

Le funzioni

1. $f(x) = a_k \cos k\omega x + b_k \sin k\omega x$, k intero, $\omega \neq 0$

$$2. g(x) = \rho_k \cos(k\omega x + \theta_k), \quad k \text{ intero, } \rho > 0, \omega \neq 0, \theta_k \in \mathbb{R}$$

dette **armoniche elementari**, sono periodiche di periodo $2\pi/\omega$. Esse sono inoltre equivalenti, nel senso che, assegnate le costanti che caratterizzano una forma, è sempre possibile passare dalla (1) alla (2) e viceversa. Per esempio, assegnati ρ_k e θ_k , è immediato calcolare a_k e b_k . Infatti, essendo

$$\begin{aligned} g(x) &= \rho_k \cos(k\omega x + \theta_k) = \rho_k (\cos k\omega x \cos \theta_k - \sin k\omega x \sin \theta_k) \\ &= (\rho_k \cos \theta_k) \cos k\omega x - (\rho_k \sin \theta_k) \sin k\omega x \end{aligned}$$

basta porre $a_k = \rho_k \cos \theta_k$ e $b_k = -\rho_k \sin \theta_k$.

Viceversa, noti a_k e b_k , è immediato determinare ρ_k e θ_k . Infatti, essendo

$$\begin{cases} \rho_k \cos \theta_k = a_k \\ \rho_k \sin \theta_k = -b_k \end{cases},$$

si avrà $\rho_k = \sqrt{a_k^2 + b_k^2} = \arctan\left(-\frac{b_k}{a_k}\right)$, $a_k \neq 0$. In particolare, se $b_k = 0$ si ha $\rho_k = |a_k|$ e $\theta = 0$ o $\theta = \pi$ a seconda che sia $a_k > 0$ o $a_k < 0$, rispettivamente.

Polinomio trigonometrico di ordine n.

$$T_n(x) = a_0 + \sum_{k=1}^n (a_k \cos k\omega x + b_k \sin k\omega x)$$

dove a_0, a_k, b_k e $\omega \neq 0$ sono numeri reali.

T_n è una funzione periodica, in quanto combinazione lineare delle $2n + 1$ funzioni elementari

$$1, \cos \omega x, \sin \omega x, \dots, \cos n\omega x, \sin n\omega x, \quad (\text{A.1})$$

tutte periodiche di periodo $T = \frac{2\pi}{\omega}$.

Le funzioni trigonometriche (A.1) qualunque sia n , sono mutuamente ortogonali in $[0, T]$.

Dimostrazione. La funzione costante $f(x) = 1$ è ortogonale a tutte le altre, in quanto

$$\int_0^T 1 \cdot \cos k\omega x dx = \frac{\sin k\omega x}{k\omega} \Big|_0^T = \frac{1}{k\omega} \sin k\omega T = \frac{1}{k\omega} \sin k2\pi = 0$$

$$\int_0^T 1 \cdot \sin k\omega x dx = -\frac{\cos k\omega x}{k\omega} \Big|_0^T = -\frac{1}{k\omega}(\cos k2\pi - 1) = 0$$

per $k = 1, 2, \dots, n$. Inoltre se $h \neq k$,

$$\begin{aligned} \int_0^T \cos k\omega x \cos h\omega x dx &= \int_0^T \frac{1}{2} [\cos(k-h)\omega x + \cos(k+h)\omega x] dx \\ &= \frac{1}{2} \left[\frac{\sin(k-h)\omega x}{(k-h)\omega} + \frac{\sin(k+h)\omega x}{(k+h)\omega} \right]_0^T = 0, \end{aligned}$$

in quanto $\omega T = 2\pi$. Da notare che se $h = k$,

$$\begin{aligned} \int_0^T \cos^2 k\omega x dx &= \int_0^T \frac{1}{2} [1 + \cos 2k\omega x] dx \\ &= \frac{1}{2} \left[x + \frac{\sin 2k\omega x}{2k\omega} \right]_0^T = \frac{T}{2}. \end{aligned}$$

Pertanto i due risultati possono essere espressi nella forma seguente:

$$\int_0^T \cos k\omega x \cos h\omega x dx = \frac{T}{2} \delta_{hk},$$

essendo δ_{hk} il simbolo di Kroenecker, così definito

$$\delta_{hk} = \begin{cases} 1 & h = k \\ 0 & h \neq k \end{cases}.$$

Analogamente, se $h \neq k$,

$$\begin{aligned} \int_0^T \sin k\omega x \sin h\omega x dx &= \int_0^T \frac{1}{2} [\cos(k-h)\omega x - \cos(k+h)\omega x] dx \\ &= \frac{1}{2} \left[\frac{\sin(k-h)\omega x}{(k-h)\omega} - \frac{\sin(k+h)\omega x}{(k+h)\omega} \right]_0^T = 0, \end{aligned}$$

e, se $h = k$,

$$\begin{aligned} \int_0^T \sin^2 k\omega x dx &= \int_0^T \frac{1}{2} [1 - \cos 2k\omega x] dx \\ &= \frac{1}{2} \left[x - \frac{\sin 2k\omega x}{2k\omega} \right]_0^T = \frac{T}{2}. \end{aligned}$$

Pertanto, indipendentemente dall'essere h e k uguali o diversi,

$$\int_0^T \sin k\omega x \sin h\omega x dx = \frac{T}{2} \delta_{hk}.$$

Infine è immediato osservare che $\sin h\omega x$ e $\cos k\omega x$ sono ortogonali in $[0, T]$, qualunque sia la coppia di valori h, k . Infatti, se $h \neq k$,

$$\begin{aligned} \int_0^T \sin k\omega x \cos h\omega x dx &= \int_0^T \frac{1}{2} [\sin(k-h)\omega x + \sin(k+h)\omega x] dx \\ &= \frac{1}{2} \left[-\frac{\cos(k-h)\omega x}{(k-h)\omega} - \frac{\cos(k+h)\omega x}{(k+h)\omega} \right]_0^T = 0, \end{aligned}$$

e inoltre, se $k = h$,

$$\int_0^T \sin k\omega x \cos h\omega x dx = \int_0^T \frac{\sin 2h\omega x}{2} dx = \frac{1}{2} \left[-\frac{\cos 2h\omega x}{2h\omega} \right]_0^T = 0.$$

■

Proprietà sull'integrazione delle funzioni periodiche.

Indicato con T il periodo di una funzione periodica e integrabile e con a un qualunque numero reale, vale la seguente proprietà:

$$\int_a^{a+T} f(x) dx = \int_0^T f(x) dx.$$

Dimostrazione. Qualunque sia $a \in \mathbb{R}$, esiste un intero n tale che $nT \leq a < (n+1)T$.

$a < (n + 1)T$. Pertanto, essendo $(n + 1)T \leq a + T$,

$$\begin{aligned}
 \int_a^{a+T} f(x)dx &= \int_a^{(n+1)T} f(x)dx + \int_{(n+1)T}^{a+T} f(x)dx \quad \text{posto } x = y + T \text{ nel secondo int.} \\
 &= \int_a^{(n+1)T} f(x)dx + \int_{nT}^a f(y + T)dy \\
 &= \int_a^{(n+1)T} f(x)dx + \int_{nT}^a f(x)dx \\
 &= \int_{nT}^{(n+1)T} f(x)dx \\
 &= \int_0^T f(z + nT)dz \\
 &= \int_0^T f(x)dx.
 \end{aligned}$$

■

Coefficienti di Fourier

Supponiamo che una funzione $f(x)$ sia approssimata in $[0, T]$ mediante un polinomio trigonometrico di ordine n

$$f(x) \sim a_0 + \sum_{k=1}^n (a_k \cos k\omega x + b_k \sin k\omega x), \quad T = \frac{2\pi}{\omega}.$$

È allora naturale richiedere che l'integrazione in $[0, T]$ della $f(x)$ per le funzioni di base

$$\{1, \cos \omega x, \sin \omega x, \dots, \cos k\omega x, \sin k\omega x\}$$

risulti sostanzialmente uguale all'integrazione della funzione approssimante per le stesse funzioni di base. Questo fatto, in conseguenza della ortogonalità

delle funzioni di base, implica che

$$\begin{aligned}\int_0^T f(x) dx &\simeq a_0 \int_0^T 1 dx = a_0 T \\ \int_0^T f(x) \cos k\omega x dx &\simeq a_k \int_0^T \cos^2 k\omega x dx = a_k \frac{T}{2}, \quad k = 1, 2, \dots, n, \\ \int_0^T f(x) \sin k\omega x dx &\simeq b_k \int_0^T \sin^2 k\omega x dx = b_k \frac{T}{2}, \quad k = 1, 2, \dots, n.\end{aligned}$$

Poiché questa osservazione è dovuta a Fourier, i coefficienti

$$\begin{aligned}a_0 &= \frac{1}{T} \int_0^T f(x) dx \\ a_k &= \frac{2}{T} \int_0^T f(x) \cos k\omega x dx, \quad k = 1, 2, \dots, n, \\ b_k &= \frac{2}{T} \int_0^T f(x) \sin k\omega x dx, \quad k = 1, 2, \dots, n,\end{aligned}$$

vengono definiti **coefficienti di Fourier** della f .

Energia di un segnale

Supponendo che una funzione $f(x)$, al quadrato integrabile in $[0, T]$, rappresenti un segnale a valori reali o complessi, si definisce energia del segnale la norma della f così definita

$$\|f\| = \left(\int_0^T |f(x)|^2 dx \right)^{\frac{1}{2}}, \quad |f(x)|^2 = f(x)f(\bar{x}),$$

dove $f(\bar{x})$ indica il complesso coniugato di $f(x)$ se $f(x)$ è complessa, $f(x)$ stessa se $f(x)$ è reale.

Indicato con T_n un polinomio trigonometrico, il quadrato della sua energia

può essere agevolmente calcolato nel modo seguente:

$$\begin{aligned} \|T_n\|^2 &= \int_0^T \left[a_0 + \sum_{k=1}^n (a_k \cos k\omega x + b_k \sin k\omega x) \right]^2 dx, \quad \omega = \frac{2\pi}{T} \\ &= \int_0^T \left[a_0 + \sum_{k=1}^n (a_k^2 \cos^2 k\omega x + b_k^2 \sin^2 k\omega x) \right] dx \\ &= T \left[a_0^2 + \frac{1}{2} \sum_{k=1}^n (a_k^2 + b_k^2) \right], \end{aligned}$$

essendo $\int_0^T \cos^2 k\omega x dx = \int_0^T \sin^2 k\omega x dx = \frac{T}{2}$.

È interessante capire quale sia la distanza minima, in termini di energia, tra una $f(x)$ in $[0, T]$ ed una sua approssimante trigonometrica. Interessa dunque minimizzare la funzione

$$I(a_0, a_1, b_1, \dots, a_n, b_n) = \int_0^T [f(x) - T_n(x)]^2 dx$$

al variare dei coefficienti $a_0, a_1, b_1, \dots, a_n, b_n$ di $T_n(x)$. È possibile dimostrare che, grazie alla ortogonalità delle funzioni di base, tale minimo è determinato dai coefficienti di Fourier.

Dimostrazione. Per brevità limitiamoci a dimostare il risultato per $n = 1$. A tale scopo consideriamo la funzione

$$I(a_0, a_1, b_1) = \int_0^T [f(x) - (a_0 + a_1 \cos \omega x + b_1 \sin \omega x)]^2 dx$$

e indichiamo con \hat{a}_0, \hat{a}_1 e \hat{b}_1 i relativi coefficienti di Fourier

$$\begin{aligned} \hat{a}_0 &= \frac{1}{T} \int_0^T f(x) dx \\ \hat{a}_1 &= \frac{2}{T} \int_0^T f(x) \cos \omega x dx \\ \hat{b}_1 &= \frac{2}{T} \int_0^T f(x) \sin \omega x dx. \end{aligned}$$

Per l'ortogonalità delle funzioni di base e la definizione di \hat{a}_0, \hat{a}_1 e \hat{b}_1

$$\begin{aligned} I(a_0, a_1, b_1) &= \int_0^T f^2(x) dx + T a_0^2 + \frac{T}{2}(a_1^2 + b_1^2) + \\ &\quad - 2 \left[a_0 \int_0^T f(x) dx + a_1 \int_0^T f(x) \cos \omega x dx + b_1 \int_0^T f(x) \sin \omega x dx \right] \\ &= \int_0^T f^2(x) dx + T a_0^2 + \frac{T}{2}(a_1^2 + b_1^2) - 2 \left[T a_0 \hat{a}_0 + \frac{T}{2}(a_1 \hat{a}_1 + b_1 \hat{b}_1) \right], \end{aligned}$$

da cui, aggiungendo e togliendo $T\hat{a}_0^2, \frac{T}{2}\hat{a}_1^2$ e $\frac{T}{2}\hat{b}_1^2$,

$$\begin{aligned} I(a_0, a_1, b_1) &= \int_0^T f^2(x) dx - T a_0^2 + T(a_0 - \hat{a}_0)^2 - \frac{T}{2}\hat{a}_1^2 + \\ &\quad + \frac{T}{2}(a_1 - \hat{a}_1)^2 - \frac{T}{2}\hat{b}_1^2 + \frac{T}{2}(b_1 - \hat{b}_1)^2, \end{aligned}$$

il cui minimo è ovviamente ottenuto per $a_0 = \hat{a}_0, a_1 = \hat{a}_1$ e $b_1 = \hat{b}_1$. ■

In quadrato dell'energia corrispondente alla differenza tra la $f(x)$ e la sua approssimante ottimale $\hat{T}_1(x) = \hat{a}_0 + \hat{a}_1 \cos \omega x + \hat{b}_1 \sin \omega x$ è dunque

$$I(\hat{a}_0, \hat{a}_1, \hat{b}_1) = \int_0^T f^2(x) dx - T\hat{a}_0^2 - \frac{T}{2}(\hat{a}_1^2 + \hat{b}_1^2).$$

Estendendo tali considerazioni al caso generale si ottiene che il minimo di $I(a_0, a_1, b_1, \dots, a_n, b_n)$ è ottenuto per $a_0 = \hat{a}_0, a_k = \hat{a}_k$ e $b_k = \hat{b}_k, k = 1, 2, \dots, n$. In altri termini, prefissato n , i coefficienti di Fourier indentificano il polinomio trigonometrico che, in termini di energia, meglio approssima un segnale, esprimibile con una funzione al quadrato integrabile, in $[0, T]$.

L'estensione delle precedenti considerazioni permette inoltre di affermare che

$$I(\hat{a}_0, \hat{a}_1, \hat{b}_1, \dots, \hat{a}_n, \hat{b}_n) = \int_0^T f^2(x) dx - T\hat{a}_0^2 - \frac{T}{2} \sum_{k=1}^n (\hat{a}_k^2 + \hat{b}_k^2)$$

da cui, essendo $I(\hat{a}_0, \hat{a}_1, \hat{b}_1, \dots, \hat{a}_n, \hat{b}_n) \geq 0$, segue la **diseguaglianza di Bessel**

$$\hat{a}_0^2 + \frac{1}{2} \sum_{k=1}^n (\hat{a}_k^2 + \hat{b}_k^2) \leq \frac{1}{T} \int_0^T f^2(x) dx.$$

Tale disuguaglianza esprime il fatto che l'energia associata al polinomio trigonometrico che meglio approssima un segnale è sempre inferiore a quella del segnale stesso.

Un'altra interessante osservazione è che

$$\|T_{n+1}\|^2 = \|T_n\|^2 + \frac{T}{2}(\hat{a}_{n+1}^2 + \hat{b}_{n+1}^2),$$

ossia: l'energia dell'approssimante T_n è una funzione *monotonamente* crescente rispetto a n e limitata dall'energia del segnale.

A.2 Serie di Fourier

Definizione A.2.1 Sia f una funzione integrabile in $[-L, L]$. Allora ad essa è univocamente associabile la serie di Fourier

$$a_0 + \sum_{n=1}^{\infty} \left(a_n \cos n \frac{\pi}{L} x + b_n \sin n \frac{\pi}{L} x \right). \quad (\text{A.2})$$

dove, per l'ortogonalità delle funzioni di base $\{1, \cos n \frac{\pi}{L} x, \sin n \frac{\pi}{L} x\}$

$$\begin{aligned} a_0 &= \frac{1}{2L} \int_{-L}^L f(x) dx \\ a_n &= \frac{1}{L} \int_{-L}^L f(x) \cos n \frac{\pi}{L} x dx \\ b_n &= \frac{1}{L} \int_{-L}^L f(x) \sin n \frac{\pi}{L} x dx. \end{aligned}$$

Esempio A.2.2 Scrivere la serie di Fourier di $f(x) = x$, $-\pi \leq x \leq \pi$.

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} x dx = \frac{1}{4\pi} [x^2]_{-\pi}^{\pi} = 0; \\ a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} x \cos kx dx = \frac{1}{\pi} \int_{-\pi}^{\pi} x d \frac{\sin kx}{k} = \\ &= \frac{1}{\pi} \left[x \frac{\sin kx}{k} \right]_{-\pi}^{\pi} - \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{\sin kx}{k} dx = \frac{1}{\pi} \left[\frac{\cos kx}{k} \right]_{-\pi}^{\pi} = 0; \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin kx dx = -\frac{1}{\pi} \int_{-\pi}^{\pi} x d \frac{\cos kx}{k} = \\ &= -\frac{1}{\pi} \left[x \frac{\cos kx}{k} \right]_{-\pi}^{\pi} + \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{\cos kx}{k} dx = \frac{2}{k} (-1)^{k+1}, \end{aligned}$$

in quanto $\cos k\pi = \cos(-k\pi) = (-1)^k$.

La serie di Fourier di x su $[-\pi, \pi]$ è dunque la

$$\sum_{k=1}^{\infty} \frac{2}{k} (-1)^{k+1} \sin kx.$$

Esempio A.2.3 Scrivere la serie di Fourier di $f(x) = x^2$, $-\pi \leq x \leq \pi$.

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} x^2 dx = \frac{1}{2\pi} \left[\frac{x^3}{3} \right]_{-\pi}^{\pi} = \frac{\pi^2}{3}; \\ a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 \cos kx dx = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 d \frac{\sin kx}{k} = \\ &= \frac{1}{\pi} \left[x^2 \frac{\sin kx}{k} \right]_{-\pi}^{\pi} + \frac{1}{\pi} \int_{-\pi}^{\pi} 2x d \frac{\cos kx}{k^2} = \\ &= \frac{1}{\pi} \left[2x \frac{\cos kx}{k^2} \right]_{-\pi}^{\pi} - \frac{2}{\pi} \int_{-\pi}^{\pi} \frac{\cos kx}{k^2} dx = \frac{4}{k^2} (-1)^k; \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 \sin kx dx = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 d \frac{\cos kx}{k} = \\ &= \frac{1}{\pi} \left[x^2 \frac{\cos kx}{k} \right]_{-\pi}^{\pi} + \frac{1}{\pi} \int_{-\pi}^{\pi} 2x d \frac{\sin kx}{k^2} = \\ &= \frac{1}{\pi} \left[2x \frac{\sin kx}{k^2} \right]_{-\pi}^{\pi} - \frac{2}{\pi} \int_{-\pi}^{\pi} \frac{\sin kx}{k^2} dx = \frac{2}{k} \left[\frac{\cos kx}{k^2} \right]_{-\pi}^{\pi} = 0. \end{aligned}$$

Pertanto la serie di Fourier di x^2 su $[-\pi, \pi]$ è

$$\frac{\pi^2}{3} + \sum_{k=1}^{\infty} \frac{4}{k^2} (-1)^k \cos kx.$$

Esempio A.2.4 Scrivere la serie di Fourier di

$$f(x) = \begin{cases} 0, & \pi \leq x < 1 \\ 1, & 1 \leq x < 2 \\ 2, & 2 \leq x < \pi \end{cases}.$$

$$\begin{aligned}
a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} x dx = \frac{1}{2\pi} \left[\int_1^2 dx + \int_2^{\pi} 2 dx \right] = \\
&\frac{1}{2\pi} [1 + 2(\pi - 2)] = 1 - \frac{3}{2\pi}; \\
a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} x f(x) \cos kx dx = \frac{1}{\pi} \left[\int_1^2 \cos kx dx + 2 \int_2^{\pi} \cos kx dx \right] = \\
&\frac{1}{k\pi} [(\sin 2k - \sin k) + 2(\sin k\pi - 2 \sin 2k)] = -\frac{1}{k\pi} \sin k(2 + 1); \\
b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx = \left[\int_1^2 \sin kx dx + 2 \int_2^{\pi} \sin kx dx \right] = \\
&-\frac{1}{k\pi} [(\cos 2k - \cos k) + 2(\cos k\pi - 2 \cos 2k)] = \\
&-\frac{1}{k\pi} [-\cos 2k - \cos k + 2(-1)^k].
\end{aligned}$$

Di conseguenza, alla $f(x)$ possiamo associare la serie di Fourier

$$1 - \frac{3}{2\pi} - \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{1}{k} \left\{ [\sin k(\cos 2k + 1)] \cos kx + [2(-1)^k - \cos k \cos 2k] \sin kx \right\}.$$

Esempio A.2.5 Scrivere la serie di Fourier di

$$f(x) = \begin{cases} 0, & -3 \leq x < 0 \\ x, & 0 \leq x < 3 \end{cases}.$$

$$\begin{aligned}
a_0 &= \frac{1}{6} \int_{-3}^3 f(x) dx = \frac{1}{6} \int_0^3 x dx = \frac{3}{4}; \\
a_k &= \frac{1}{3} \int_3^3 f(x) \cos \frac{k\pi x}{3} dx = \frac{1}{3} \int_3^3 x \cos \frac{k\pi x}{3} dx = \frac{3}{k^2 \pi^2} [(-1)^k - 1]; \\
b_k &= \frac{1}{3} \int_3^3 f(x) \sin \frac{k\pi x}{3} dx = \frac{1}{3} \int_3^3 x \sin \frac{k\pi x}{3} dx = -\frac{3}{k\pi} (-1)^k.
\end{aligned}$$

Pertanto, la serie di Fourier di $f(x)$ in $[-3, 3]$ è

$$\frac{3}{4} + \sum_{k=1}^{\infty} \left\{ \frac{3}{k^2 \pi^2} [(-1)^k - 1] \cos \frac{k\pi x}{3} - \frac{3}{k\pi} (-1)^k \sin \frac{k\pi x}{3} \right\}.$$

Funzioni pari e dispari

Sia f una funzione definita in un intervallo $[-L, L] \forall L \in \mathbb{R}^+$. Allora, se per ogni $x \in [-L, L]$, risulta:

$$\begin{aligned} f(x) &= f(-x) && \text{la funzione si dice pari;} \\ f(x) &= -f(-x) && \text{la funzione si dice dispari.} \end{aligned}$$

Se per almeno un valore di x non vale nessuna delle due precedenti condizioni, la f non è né pari né dispari. Di conseguenza, la funzione dell'Es. A.2.2 è dispari, quella dell'Es. A.2.3 è pari, mentre quelle degli esempi A.2.4 e A.2.5 non sono né pari né dispari.

È importante notare che se la funzione $f(x)$ è pari, allora $b_n = 0$ per $n = 1, 2, \dots$ e la sua serie di Fourier si riduce alla serie di soli coseni

$$f(x) \simeq a_0 + \sum_{n=1}^{\infty} a_n \cos n \frac{\pi}{L} x$$

dove

$$a_0 = \frac{1}{L} \int_0^L f(x) dx, \quad a_n = \frac{2}{L} \int_0^L f(x) \cos n \frac{\pi}{L} x dx.$$

Dimostrazione. Tenuto conto dell'ipotesi di parità

$$\begin{aligned} a_0 &= \frac{1}{2L} \int_{-L}^L f(x) dx = \frac{1}{2L} \left[\int_{-L}^0 f(x) dx + \int_0^L f(x) dx \right] \\ &= \frac{1}{2L} \left[- \int_L^0 f(x) dx + \int_0^L f(x) dx \right] = \frac{1}{L} \int_0^L f(x) dx, \end{aligned}$$

dove nel secondo passaggio si è posto $t = -x$ nel primo integrale e si è fatto uso della parità della $f(x)$.

Analogamente

$$\begin{aligned} a_k &= \frac{1}{L} \int_{-L}^L f(x) \cos \frac{k\pi x}{L} dx = \\ &= \frac{1}{L} \left[\int_{-L}^0 f(x) \cos \frac{k\pi x}{L} dx + \int_0^L f(x) \cos \frac{k\pi x}{L} dx \right] \\ &= \frac{2}{L} \int_0^L f(x) \cos \frac{k\pi x}{L} dx, \end{aligned}$$

e infine

$$\begin{aligned} b_k &= \frac{1}{L} \int_{-L}^L f(x) \sin \frac{k\pi x}{L} dx = \\ &= \frac{1}{L} \left[\int_{-L}^0 f(x) \sin \frac{k\pi x}{L} dx + \int_0^L f(x) \sin \frac{k\pi x}{L} dx \right] = 0. \end{aligned}$$

■

Se la funzione $f(x)$ è dispari, allora $a_0 = 0$ e $a_n = 0$ per $n = 1, 2, \dots$ e la sua serie di Fourier si riduce alla serie di soli seni

$$f(x) \simeq \sum_{n=1}^{\infty} b_n \sin n \frac{\pi}{L} x$$

dove

$$b_n = \frac{2}{L} \int_0^L f(x) \sin n \frac{\pi}{L} x dx.$$

Dimostrazione. Si dimostra in modo del tutto analogo, ossia decomponendo l'integrale tra $[-L, L]$ in uno tra $[-L, 0]$, ponendo $t = -x$ nel primo integrale e ricordando che $f(-t) = -f(t)$. ■

Osservazione A.2.6 Non è detto che il valore della f in un generico $x \in [-L, L]$ coincida con il valore della sua serie di Fourier in x . In altri termini, non necessariamente la serie di Fourier di una funzione risulta ad essa convergente in ogni punto. Ad esempio, la serie di Fourier di $f(x) = x$ in $[-\pi, \pi]$ è $\sum_{k=1}^{\infty} \frac{2}{k} (-1)^{k+1} \sin kx$, la quale assume il valore 0 in $x = \pm\pi$ mentre in tali punti $f(x) = \pm\pi$. Inoltre, mentre $f\left(\frac{\pi}{2}\right) = \frac{\pi}{2}$, non è affatto evidente che $\frac{\pi}{2}$ sia il valore della serie in tale punto.

Allo scopo di discutere il problema della convergenza è bene premettere le seguenti definizioni di derivata destra e sinistra.

Definizione A.2.7 Se la $f(x)$ ammette limite destro in x_0 , $f'(x_0^+)$ si definisce **derivata destra** della f in x_0 il limite

$$f'(x_0^+) = \lim_{h \rightarrow 0^+} \frac{f(x_0 + h) - f(x_0^+)}{h},$$

nell'ipotesi che esso esista e sia finito.

Analogamente, se la f ammette limite sinistro in x_0 , $f(x_0^-)$ si definisce **derivata sinistra** della f in x_0 il limite

$$f'(x_0^-) = \lim_{h \rightarrow 0^+} \frac{f(x_0 + h) - f(x_0^-)}{h},$$

nell'ipotesi che esso esista e sia finito.

Esempio A.2.8 Sia

$$f(x) = \begin{cases} 1 + x & -\pi \leq x < 0 \\ x^2 & 0 \leq x < \pi \end{cases},$$

in questo caso esistono derivata destra e sinistra in tutti i punti interni. Per $x \in (-\pi, \pi) \setminus \{0\}$ è evidente, mentre in $x = 0$ essendo $f(0^+) = 0$, $f(0^-) = 1$, risulta

$$\lim_{h \rightarrow 0^+} \frac{h^2 - 0}{h} = 0, \quad \lim_{h \rightarrow 0^-} \frac{1 + h - 1}{h} = 1.$$

Da notare che in $-\pi$ esiste la derivata destra e in π quella sinistra con $f'(-\pi^+) = 1$ e $f'(\pi^-) = 2\pi$.

Definizione A.2.9 La funzione f è regolare a tratti in $[a, b]$ se valgono le seguenti proprietà:

- 1) esiste un numero finito di punti x_1, \dots, x_n con $a < x_1 < \dots < x_n < b$, tale che f è di classe C^1 negli intervalli $[a, x_1], [x_j, x_{j+1}] (j = 1, 2, \dots, n-1)$ e $[x_n, b]$;
- 2) nei punti x_1, \dots, x_n , esistono finite le derivate destra e sinistra;
- 3) nei punti x_1, \dots, x_n , esistono finiti i limiti destro e sinistro.

Teorema A.2.10 (Convergenza della serie di Fourier) Sia f regolare a tratti su $[-L, L]$. Allora:

- 1) se la f è continua in $x_0 \in (-L, L)$, la serie di Fourier assume in x_0 il valore $f(x_0)$;
- 2) se $x_0 \in (-L, L)$ e la f è discontinua in x_0 , la serie di Fourier converge a

$$\frac{1}{2}[f(x_0^+) + f(x_0^-)];$$

3) la serie di Fourier converge a

$$\frac{1}{2}[f(-L^+) + f(L^-)]$$

sia in $-L$ che in L .

La 1) consente, ad esempio, di affermare che $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{k} \sin k \frac{\pi}{2} = \frac{\pi}{2}$. La 3) permette di dimostrare che

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Per convincersene è sufficiente osservare che

$$\pi^2 = \frac{\pi^2}{3} + \sum_{k=1}^{\infty} (-1)^k \frac{4}{k^2} \cos k\pi = \frac{\pi^2}{3} + \sum_{k=1}^{\infty} \frac{4}{k^2}.$$

Teorema A.2.11 (Lemma di Riemann-Lebesgue) Sia f sommabile nel senso che esiste finito l'integrale $\int_{-\infty}^{\infty} |f(x)| dx$. Allora

$$\lim_{n \rightarrow \infty} \int_{-L}^L f(x) \cos n \frac{\pi}{L} x dx = 0, \quad \lim_{n \rightarrow \infty} \int_{-L}^L f(x) \sin n \frac{\pi}{L} x dx = 0,$$

che stabiliscono la convergenza a zero dei coefficienti della serie di Fourier.

Questioni rilevanti sulle serie di Fourier riguardano la sua integrazione e derivazione termine a termine. Come si evince dai due teoremi che seguono, la differenziabilità termine a termine è molto più problematica rispetto alla integrabilità.

Teorema A.2.12 (Integrazione termine a termine) Se la f è regolare a tratti in $[-L, L]$ e ivi sviluppabile in serie di Fourier, essa è integrabile termine a termine, ossia

$$\int_{-L}^x f(t) dt = a_0(x+L) + \frac{L}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \left\{ a_n \sin n \frac{\pi}{L} x - b_n \left[\cos n \frac{\pi}{L} x - \cos n\pi \right] \right\}.$$

Esempio A.2.13 Dall'esempio A.2.2, la serie di Fourier di $f(x) = 2x$ in $[-\pi, \pi]$ è

$$\sum_{k=1}^{\infty} \frac{4}{k} \sin kx.$$

Poiché f è continua in $[-\pi, \pi]$, in virtù del Teorema A.2.12 si può scrivere

$$\begin{aligned} \int_{-\pi}^x 2t dt &= x^2 - \pi^2 = \sum_{k=1}^{\infty} \frac{4}{k} (-1)^{k+1} \int_{-\pi}^x \sin kt dt \\ &= \sum_{k=1}^{\infty} \frac{4}{k} (-1)^{k+1} \left[-\frac{1}{k} (-\cos k\pi) \right] \\ &= \sum_{k=1}^{\infty} \frac{4}{k^2} (-1)^k [-(-1)^k], \end{aligned}$$

da cui, ricordando che $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$, si ottiene, esattamente come nell'esempio A.2.3, che

$$x^2 \sim \frac{\pi^2}{3} + \sum_{k=1}^{\infty} \frac{4}{k} (-1)^k \cos kx.$$

La differenziazione di una serie di Fourier si presenta in modo molto diverso. Infatti, differenziando termine a termine la serie di Fourier di x in $[-\pi, \pi]$ si ottiene

$$\sum_{k=1}^{\infty} \left(\frac{2}{k} (-1)^{k+1} \sin kx \right)' = \sum_{k=1}^{\infty} 2(-1)^{k+1} \cos kx,$$

la quale non converge per alcun valore di x in $(-\pi, \pi)$ e tanto meno risulta $f'(x) = 1$.

Teorema A.2.14 (Differenziazione termine a termine) *Se la f è continua in $[-L, L]$ con $f(-L) = f(L)$ e la f' è regolare a tratti in $[-L, L]$, allora la serie di Fourier è derivabile termine a termine. Ossia, in ogni punto $x \in (-L, L)$ in cui la $f'(x)$ è continua,*

$$f'(x) = \frac{\pi}{L} \sum_{n=1}^{\infty} \left[-na_n \sin n \frac{\pi}{L} x + nb_n \cos n \frac{\pi}{L} x \right].$$

La serie di Fourier di $f(x) = x$ non è dunque derivabile termine a termine in quanto $f(-\pi) \neq f(\pi)$.

Esercizi

1. Risolvere l'equazione differenziale

$$y'' + 8y = f(t), \quad f(t) = \begin{cases} \frac{\pi + 2t}{\pi}, & -\pi \leq t < 0 \\ \frac{\pi - 2t}{\pi}, & 0 \leq t < \pi \\ f(t + 2\pi), & t \in \mathbb{R}. \end{cases}$$

$$f(t) \simeq \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{1 - (-1)^k}{k^2} \cos kt$$

Derivando e sostituendo $y(t) \simeq a_0 + \sum_{k=1}^{\infty} a_k \cos kt$ si ottiene

$$8a_0 + \sum_{k=1}^{\infty} (-k^2 + 8)a_k \cos kt = \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{1 - (-1)^k}{k^2} \cos kt$$

da cui

$$a_0 = 0, \quad (-k^2 + 8)a_k = \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{1 - (-1)^k}{k^2} \cos kt, \quad k = 1, 2, \dots$$

e infine

$$y(t) \simeq \frac{8}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{(2k-1)^2 [8 - (2k-1)^2]} \cos(2k-1)t.$$

2. Calcolare la serie di Fourier di $f(x) = |x|$, $-\pi \leq x \leq \pi$.

Soluzione:

$$f(x) \simeq \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cos(2n-1)x.$$

3. Calcolare la serie di Fourier di $f(x) = \begin{cases} -1, & -\pi \leq x < 0 \\ 1, & 0 \leq x < \pi \end{cases}$.

Soluzione:

$$f(x) \simeq \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)x.$$

Appendice B

Problema di Sturm-Liouville

È un problema spettrale che, nella sua forma canonica, viene rappresentato nel modo seguente:

$$\begin{cases} \frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] + [q(x) + \lambda r(x)] y = 0, & a \leq x \leq b \\ \alpha_1 y(a) + \alpha_2 y'(a) = 0, & \beta_1 y(b) + \beta_2 y'(b) = 0 \end{cases} \quad (\text{B.1})$$

dove $\alpha_1, \alpha_2, \beta_1, \beta_2$ sono costanti assegnate, $p(x) \geq 0$, $q(x)$ e $r(x) \geq 0$ sono funzioni note generalmente derivabili e λ è un parametro indipendente da x . Per il suo ruolo particolare la funzione $r(x)$ è definita funzione peso. I valori $\{\lambda_k\}$ del parametro spettrale λ ai quali corrispondono soluzioni $\{y_k\}$ non identicamente nulle, sono detti autovalori e le corrispondenti y_k sono definite autofunzioni. L'insieme $\{\lambda_k, y_k\}$ rappresenta lo spettro del problema.

Lo spettro è caratterizzato dalle seguenti due proprietà:

- gli autovalori sono reali e formano una infinità numerabile non limitata in quanto $\lim_{k \rightarrow \infty} |\lambda_k| = +\infty$;
- le autofunzioni sono mutuamente ortogonali, rispetto alla funzione peso $r(x)$, nell'intervallo $[a, b]$, ossia $\int_a^b r(x) y_n(x) y_m(x) dx = 0$ per ogni coppia (n, m) con $n \neq m$.

Da notare che ambedue le proprietà valgono anche nell'ipotesi che $r(x) \leq 0$, nel qual caso come peso si assume la funzione $|r(x)|$.

Alcune dimostrazioni.

1) gli autovalori sono reali;

Dim.: supponendo $\alpha_1, \alpha_2, \beta_1, \beta_2, p, q$ ed r reali, si ipotizza l'esistenza di un autovalore λ e della corrispondente autofunzione y . In tal caso λ e y soddisfano il sistema

$$\begin{cases} \frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] + [q(x) + \lambda r(x)] y = 0, & a \leq x \leq b \\ \alpha_1 y(a) + \alpha_2 y'(a) = 0, & \beta_1 y(b) + \beta_2 y'(b) = 0 \end{cases} \quad (\text{B.2})$$

da cui, considerando i complessi coniugati, si ha

$$\begin{cases} \frac{d}{dx} \left[p(x) \frac{d\bar{y}}{dx} \right] + [q(x) + \bar{\lambda} r(x)] \bar{y} = 0, & a \leq x \leq b \\ \alpha_1 \bar{y}(a) + \alpha_2 \bar{y}'(a) = 0, & \beta_1 \bar{y}(b) + \beta_2 \bar{y}'(b) = 0 \end{cases} \quad (\text{B.3})$$

Moltiplicando l'equazione (B.2) per \bar{y} e la (B.3) per y e sottraendo membro a membro si ottiene

$$\frac{d}{dx} [p(x) (y\bar{y}' - \bar{y}y')] = (\lambda - \bar{\lambda}) r(x) y\bar{y} = 0$$

da cui, integrando tra a e b ,

$$p(x) (y\bar{y}' - \bar{y}y') \Big|_a^b = (\lambda - \bar{\lambda}) \int_a^b r(x) y\bar{y} dx = 0$$

in quanto $y(a)\bar{y}'(a) - \bar{y}(a)y'(a) = 0$ e $y(b)\bar{y}'(b) - \bar{y}(b)y'(b) = 0$ (diversamente sarebbero $\alpha_1 = \alpha_2 = 0$ e $\beta_1 = \beta_2 = 0$, rispettivamente).

Di conseguenza $\lambda = \bar{\lambda}$, dato che $\int_a^b r(x) |y|^2 dx > 0$.

2) le autofunzioni corrispondenti ad autovalori diversi sono ortogonali rispetto alla funzione peso $r(x)$;

Dim.: se y_1 e y_2 sono autofunzioni corrispondenti agli autovalori $\lambda_1 \neq \lambda_2$, sono validi i sistemi

$$\begin{cases} \frac{d}{dx} \left[p(x) \frac{dy_1}{dx} \right] + [q(x) + \lambda_1 r(x)] y_1 = 0, & a \leq x \leq b \\ \alpha_1 y_1(a) + \alpha_2 y_1'(a) = 0, & \beta_1 y_1(b) + \beta_2 y_1'(b) = 0 \end{cases} \quad (\text{B.4})$$

$$\begin{cases} \frac{d}{dx} \left[p(x) \frac{dy_2}{dx} \right] + [q(x) + \lambda_2 r(x)] y_2 = 0, & a \leq x \leq b \\ \alpha_1 y_2(a) + \alpha_2 y_2'(a) = 0, & \beta_1 y_2(b) + \beta_2 y_2'(b) = 0 \end{cases} \quad (\text{B.5})$$

Moltiplicando l'equazione (B.4) per y_2 e la (B.5) per y_1 e sottraendo si ottiene

$$\frac{d}{dx} [p(x) (y_1 y_2' - y_2 y_1')] = (\lambda_1 - \lambda_2) r(x) y_1 y_2 = 0$$

Integrando tra a e b e tenendo conto delle condizioni agli estremi si ha

$$(\lambda_1 - \lambda_2) \int_a^b r(x) y_1 y_2 dx = p(x) (y_1 y_2' - y_2 y_1') \Big|_a^b = 0,$$

ossia, essendo, $\lambda_1 \neq \lambda_2$, y_1 e y_2 sono ortogonali rispetto a $r(x)$ in $[a, b]$.

I sistemi di Sturm-Liouville sono spesso generati nella risoluzione dei problemi alle derivate parziali mediante la tecnica di separazione delle variabili.

Appendice C

Metodi Iterativi

C.1 Introduzione

I metodi diretti sono finiti, nel senso che con un numero finito di trasformazioni lineari permettono di risolvere il sistema assegnato. I metodi iterativi sono invece infiniti nel senso che, partendo con un'approssimazione iniziale $x^{(0)}$, si costruisce una successione di vettori $\{x^{(k)}\}$ che, in aritmetica infinita, converge alla soluzione del sistema.

Il numero di moltiplicazioni/divisioni richiesto nei metodi iterativi per passare da un'approssimazione alla successione è notevolmente inferiore ad $n^3/3$ (tipico ordine di grandezza del numero di moltiplicazioni/divisioni dei metodi finiti). Tuttavia, tranne nelle situazioni particolari in cui si verifica una rapida convergenza, i metodi iterativi comportano un numero globale di moltiplicazioni/divisioni superiore ad $n^3/3$.

Nella pratica i metodi iterativi vengono usati soprattutto per risolvere i sistemi lineari ottenuti con la discretizzazione di equazioni differenziali a derivate parziali, in quanto le relative matrici, anche se di notevole dimensione, sono spesso a banda e comunque sparse (solo il 5 – 10% dei loro elementi sono diversi da zero). La loro utilizzazione si giustifica col fatto che, qualora siano soddisfatte le condizioni di convergenza, almeno in aritmetica esatta, ogni vettore di approssimazione è migliore del precedente.

C.2 Metodi iterativi del primo ordine

L'idea di base dei metodi iterativi è la seguente: indicato con

$$Ax = b, \quad (\text{C.1})$$

il sistema non singolare da risolvere, si decompone la matrice A nella forma

$$A = N - P,$$

con l'unica limitazione che N sia non singolare. Il sistema (C.1) diventa pertanto

$$x = N^{-1}Px + N^{-1}b. \quad (\text{C.2})$$

Indicato allora con $x^{(0)}$ il vettore iniziale, si costruisce la successione $\{x^{(k)}\}$ generata dalla relazione di ricorrenza

$$x^{(k)} = N^{-1}Px^{(k-1)} + N^{-1}b. \quad (\text{C.3})$$

Sostituendo $P = N - A$ nella (C.3), si ottiene una relazione di ricorrenza spesso più rapidamente convergente

$$x^{(k)} = x^{(k-1)} - N^{-1}Ax^{(k-1)} + N^{-1}b = (I - N^{-1}A)x^{(k-1)} + N^{-1}b. \quad (\text{C.4})$$

Teorema C.2.1 *Posto $M = N^{-1}P$, la successione $\{x^{(k)}\}$ converge alla soluzione esatta solo se la matrice M ha raggio spettrale inferiore ad uno.*

Dimostrazione. Sia

$$e^{(k)} = x^{(k)} - x, \quad k = 0, 1, 2, \dots,$$

la successione dei vettori-errore generata dal metodo iterativo. Indicato con $e^{(0)}$ il vettore errore iniziale, dalle relazioni (C.2) e (C.3) deriva che

$$e^{(k)} = Me^{(k-1)} = M^2e^{(k-2)} = \dots = M^ke^{(0)}, \quad k = 1, 2, \dots \quad (\text{C.5})$$

Dalla (C.5) segue immediatamente che $e^{(k)} \rightarrow 0$, qualunque sia $x^{(0)}$ se e solo se $\rho(M) < 1$. ■

Quando un metodo iterativo è convergente, come indicatore della velocità di convergenza si considera il numero

$$R = -\ln \rho(M)$$

definito come **indice di convergenza**. La conoscenza di R permette di stabilire il numero minimo di iterazioni necessarie per ridurre la norma di un vettore errore di un prefissato fattore. Infatti, in base alla (C.5), qualunque norma naturale si consideri, per ogni k

$$\|e^{(k)}\| \leq \|M\|^k \|e^{(0)}\|.$$

Pertanto, poiché $\rho(M)$ è l'estremo inferiore delle norme, il numero minimo k di iterazioni richieste per ridurre la norma dell'errore iniziale di un fattore 10^{-m} è così definito:

$$[\rho(M)]^k \leq 10^{-m},$$

ossia

$$k \frac{m}{R},$$

con $R = -\ln \rho(M)$.

C.2.1 Metodo di Jacobi o delle iterazioni simultanee

Il metodo di Jacobi è caratterizzato dalla scelta

$$(N)_{ij} = a_{ij} \delta_{ij} \quad (i, j = 1, \dots, n),$$

dove δ_{ij} è il simbolo di Kronecker. Poiché N deve essere non singolare, il metodo è applicabile solo se tutti gli elementi diagonali di A sono non nulli. In tale ipotesi la matrice di iterazione M è definita dal seguente schema:

$$M = N^{-1}P = I - N^{-1}A,$$

da cui

$$(M)_{ij} = \begin{cases} 0 & \text{per } i = j \\ -\frac{a_{ij}}{a_{ii}} & \text{per } i \neq j \end{cases}.$$

Pertanto, in base alla (C.3), il relativo schema iterativo è definito dalla seguente relazione di ricorrenza:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} x_j^{(k-1)} \right) \quad (i = 1, \dots, n; k = 1, 2, \dots). \quad (\text{C.6})$$

Teorema C.2.2 *Una condizione sufficiente perché il metodo di Jacobi sia convergente è che gli elementi diagonali di A siano diagonalmente dominanti in senso stretto.*

Dimostrazione. La suddetta ipotesi esprime il fatto che

$$\|M\|_\infty = \max_{i=1,\dots,n} \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1.$$

Poiché $\rho(M) \leq \|M\|_\infty$, anche $\rho(M) < 1$, e di conseguenza il metodo è convergente.

Dal momento che $\rho(M) \leq \|M\|_1$, la stessa conclusione vale se $\|M\|_1 < 1$, cioè se la matrice A è strettamente diagonalmente dominante per colonne. ■

Per la semplicità con cui $\|M\|_\infty$ e $\|M\|_1$ possono essere calcolate, oltre che per la convergenza del metodo, $\|M\|_\infty$ e $\|M\|_1$ vengono spesso usate per stimare l'indice di convergenza del metodo. A tale scopo ci si serve della seguente disuguaglianza:

$$R = \ln \frac{1}{\rho(M)} \min \left\{ \ln \frac{1}{\|M\|_\infty}, \ln \frac{1}{\|M\|_1} \right\}.$$

Nel metodo di Jacobi il numero di moltiplicazioni/divisioni per iterazione è dell'ordine di n^2 e pertanto, il metodo converge, per ridurre di un fattore 10^{-m} la norma del vettore errore iniziale, occorrono circa $n^2 \frac{m}{R} \ln 10$ moltiplicazioni/divisioni. Per questo motivo si considera il metodo di Jacobi computazionalmente competitivo soltanto se $\frac{m}{R} \ln 10 \leq \frac{n}{3}$.

Esempio C.2.3 Risolvere con il metodo di Jacobi il seguente sistema:

$$\begin{cases} 10x_1 + x_2 + x_3 = 12 \\ 2x_1 + 10x_2 + x_3 = 13 \\ 2x_1 + 2x_2 + 10x_3 = 14 \end{cases}.$$

Il metodo è convergente perché la matrice dei coefficienti

$$A = \begin{pmatrix} 10 & 1 & 1 \\ 2 & 10 & 1 \\ 2 & 2 & 10 \end{pmatrix}$$

è diagonalmente dominante in senso stretto. Infatti gli elementi di A diagonali sono, in valore assoluto, maggiori della somma dei valori assoluti degli elementi non diagonali lungo le tre righe.

La soluzione esatta del sistema è $x_1 = x_2 = x_3 = 1$. Sia $x^{(0)} = (1.2, 0, 0)^T$ il vettore iniziale. Iterando si ottiene:

$$\begin{cases} x_1^{(1)} = \frac{1}{10}(12 - 0 - 0) = 1.2 \\ x_2^{(1)} = \frac{1}{10}(13 - 2.4 - 0) = 1.06 \\ x_3^{(1)} = \frac{1}{10}(14 - 2.4 - 0) = 1.16 \end{cases}, \quad \begin{cases} x_1^{(2)} = \frac{1}{10}(12 - 1.06 - 1.16) = 0.978 \\ x_2^{(2)} = \frac{1}{10}(13 - 2.4 - 1.16) = 0.972 \\ x_3^{(2)} = \frac{1}{10}(14 - 2.4 - 2.12) = 0.948 \end{cases},$$

$$\begin{cases} x_1^{(3)} = \frac{1}{10}(12 - 0.972 - 0.948) = 1.008 \\ x_2^{(3)} = \frac{1}{10}(13 - 1.956 - 0.948) = 1.0096 \\ x_3^{(3)} = \frac{1}{10}(14 - 1.956 - 1.944) = 1.01 \end{cases}.$$

Iterando ancora ed arrotondando alla quarta cifra decimale si ottiene:

$$\begin{array}{lll} x_1^{(4)} = 0.9980 & x_2^{(4)} = 0.9974 & x_3^{(4)} = 0.9965 \\ x_1^{(5)} = 1.0006 & x_2^{(5)} = 1.0008 & x_3^{(5)} = 1.0009 \\ x_1^{(6)} = 0.9998 & x_2^{(6)} = 0.9998 & x_3^{(6)} = 0.9997 \\ x_1^{(7)} = 1.000 & x_2^{(7)} = 1.0000 & x_3^{(7)} = 1.0000. \end{array}$$

In base alla **norma infinito**, gli errori relativi ai suddetti vettori di iterazione sono:

$$\begin{array}{lll} \|e^{(0)}\| = 1 & \|e^{(1)}\| = 0.2 & \|e^{(2)}\| = 0.052 \\ \|e^{(3)}\| = 0.01 & \|e^{(4)}\| = 0.0035 & \|e^{(5)}\| = 0.0009 \\ \|e^{(6)}\| = 0.0003 & \|e^{(7)}\| = 0. & \end{array}$$

C.2.2 Metodo di Gauss-Seidel o delle iterazioni successive

Il metodo di Gauss-Seidel può essere considerato come una modifica di quello di Jacobi. Infatti è da esso ottenibile con la sola variante che, ogni componente di $x^{(k)}$ ottenuta al passo k viene utilizzata nella stessa iterazione per calcolare le componenti successive. Invece nel metodo di Jacobi le componenti di $x^{(k)}$ calcolate al passo k vengono utilizzate soltanto nella $(k + 1)$ -esima iterazione. Pertanto lo schema iterativo di Gauss-Seidel è così definito:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^n a_{ij}x_j^{(k-1)} \right), \quad (\text{C.7})$$

dove per ogni $k = 1, 2, \dots$, $i = 1, 2, \dots, n$. Si può dimostrare che la decomposizione di A che genera lo schema (C.7) è la seguente:

$$N = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & & 0 \\ \vdots & & & \vdots \\ a_{n1} & \dots & & a_{nn} \end{pmatrix}, \quad P = N - A.$$

Poiché $\det N = \prod_{i=1}^n a_{ii}$, tale decomposizione è utilizzabile soltanto se tutti gli elementi diagonali di A sono non nulli, come avviene nel metodo di Jacobi.

Dalla (C.7) si ricava immediatamente la seguente rappresentazione del k -esimo vettore errore

$$e_i^{(k)} = - \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} e_j^{(k)} - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} e_j^{(k-1)}, \quad (\text{C.8})$$

con $i, k = 1, 2, \dots, n$.

Le (C.7) e (C.8) per $i = 1$ e $i = n$ devono essere così interpretate:

$$\begin{cases} x_1^{(k)} = \frac{1}{a_{11}} \left(b_1 - \sum_{j=2}^n a_{1j} x_j^{(k-1)} \right) \\ e_1^{(k)} = - \sum_{j=2}^n \frac{a_{1j}}{a_{11}} e_j^{(k-1)} \end{cases}, \quad \begin{cases} x_n^{(k)} = \frac{1}{a_{nn}} \left(b_n - \sum_{j=1}^{n-1} a_{nj} x_j^{(k)} \right) \\ e_n^{(k)} = - \sum_{j=1}^{n-1} \frac{a_{nj}}{a_{nn}} e_j^{(k)} \end{cases}.$$

Teorema C.2.4 *Qualunque sia il vettore iniziale $e^{(0)}$, se la matrice A è strettamente diagonalmente dominante, cioè se*

$$r = \max_{1 \leq i \leq n} r_i < 1, \quad r_i = \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right|$$

risulta

$$\|e^{(k)}\|_{\infty} \leq r^k \|e^{(0)}\|_{\infty}, \quad k = 1, 2, \dots, n \quad (\text{C.9})$$

e pertanto il metodo è convergente.

Dimostrazione. La validità della (C.9) può essere accertata per induzione relativamente alle componenti di $e^{(k)}$. Nella suddetta ipotesi, per la

prima componente di $e^{(k)}$ si ha che

$$\begin{aligned} |e_1^{(k)}| &\leq \sum_{j=2}^n \left| \frac{a_{1j}}{a_{11}} \right| |e_j^{(k-1)}| \leq \|e^{(k-1)}\|_\infty \sum_{j=2}^n \left| \frac{a_{1j}}{a_{11}} \right| \\ \|e^{(k-1)}\|_\infty r_1 &\leq r \|e^{(k-1)}\|_\infty \leq \dots \leq r^k \|e^{(0)}\|_\infty. \end{aligned}$$

Supponendo allora $|e_h^{(k)}| \leq r \|e^{(k-1)}\|_\infty$ per $h = 1, \dots, i-1$, si dimostra che anche $|e_i^{(k)}| \leq r \|e^{(k-1)}\|_\infty$ per $i = 1, 2, \dots, n$.

Dalla (C.8), in conseguenza della suddetta ipotesi, si ha che

$$\begin{aligned} |e_i^{(k)}| &\leq \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| |e_j^{(k)}| + \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| |e^{(k-1)}| \\ &\leq \|e^{(k-1)}\|_\infty \left\{ \sum_{j=1}^{i-1} r \left| \frac{a_{ij}}{a_{ii}} \right| + \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \right\} \\ &\leq \|e^{(k-1)}\|_\infty \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| = r_i \|e^{(k-1)}\|_\infty < r \|e^{(k-1)}\|_\infty. \end{aligned}$$

La (C.9) rimane dunque dimostrata, essendo

$$\|e^{(k)}\|_\infty < r \|e^{(k-1)}\|_\infty < \dots < r^k \|e^{(0)}\|_\infty.$$

■

Pertanto, quando la matrice del sistema è diagonalmente dominante in senso stretto, risultano convergenti sia il metodo di Jacobi che quello di Gauss-Seidel (si veda Es. C.2.3). Esistono tuttavia degli esempi, con matrici non diagonalmente dominanti, per i quali si ha convergenza con il metodo di Jacobi e non con quello di Gauss-Seidel e viceversa. Quando entrambi i metodi convergono, pur non essendovi alcune dimostrazioni in questo senso, si ritiene generalmente più velocemente convergente il metodo di Gauss-Seidel.

C.3 Risoluzione numerica di sistemi nonlineari

C.3.1 Considerazioni sui metodi di risoluzione

In questo paragrafo verranno fatte alcune considerazioni sui metodi per la valutazione delle radici di un'equazione non lineare

$$f(x) = 0, \quad (\text{C.10})$$

e di un sistema non lineare

$$F(x) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ f_2(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix} = 0. \quad (\text{C.11})$$

Poiché solo in casi molto particolari si riesce a risolvere i suddetti problemi in modo esplicito, la generalità dei metodi numerici è basata su procedimenti iterativi. I metodi iterativi sono incentrati principalmente sulle due fasi seguenti:

- a) ricerca di una funzione di iterazione Φ con la proprietà che ogni soluzione ξ dell'equazione vettoriale $F(x) = 0$ è un punto fisso per la Φ , tale cioè da aversi

$$\Phi(\xi) = \xi$$

e viceversa ogni punto fisso della Φ è una soluzione dell'equazione vettoriale $f(x) = 0$.

- b) generazione di una successione di iterati $x^{(i)}$ a partire da un punto iniziale $x^{(0)}$.

In questo contesto si pongono i seguenti problemi:

- 1) scelta di una funzione di iterazione computazionalmente valida;
- 2) verifica della convergenza della successione $\{x^{(i)}\}$;
- 3) valutazione della velocità di convergenza della successione $\{x^{(i)}\}$.

Scelta della funzione di iterazione

Talvolta è immediato ottenere una funzione di iterazione, come ad es. nell'equazione

$$x - \cos x = 0.$$

È infatti spontaneo pensare ad un metodo iterativo del tipo

$$x^{(i)} = \cos x^{(i-1)} \quad (i = 1, 2, \dots)$$

nel quale $\Phi(x) = \cos x$.

Caso monodimensionale

Più in generale, assegnata un'equazione scalare $f(x) = 0$ e scelta una funzione $\varphi(x)$ con $0 < |\varphi(x)| < \infty$ nel dominio della f , si prende come funzione di iterazione la

$$\Phi(x) = x - \varphi(x)f(x).$$

È infatti immediato osservare che

$$\xi = \Phi(\xi) \iff f(\xi) = 0.$$

Quando la funzione f possiede uno zero ξ ed anche un intorno I_ξ nel quale è derivabile due volte con continuità, è possibile ricavare una funzione di iterazione per la f dalla formula di Taylor d'ordine 2. Infatti, indicato con $x^{(0)}$ un punto fisso di I_ξ , per la formula di Taylor di ordine 2

$$f(\xi) = 0 = f(x^{(0)}) + (\xi - x^{(0)})f'(x^{(0)}) + \frac{(\xi - x^{(0)})^2}{2}f''[\theta(\xi, x^{(0)})],$$

con $\theta(\xi, x^{(0)}) \in (\xi, x^{(0)})$. Qualora il termine di secondo grado sia trascurabile ed $f'(x) \neq 0$ in I_ξ ,

$$\xi \simeq x^{(0)} - \frac{f(x^{(0)})}{f'(x^{(0)})}.$$

Nelle suddette ipotesi, le precedenti considerazioni suggeriscono di scegliere come funzione di iterazione la

$$\Phi(x) = x - \frac{f(x)}{f'(x)}, \quad (\text{C.12})$$

che, come vedremo in seguito, è alla base del metodo della tangente di Newton.

Qualora sia possibile uno sviluppo della f di ordine superiore al secondo, procedendo in modo analogo è facile ricavare funzioni di iterazione che includono derivate di ordine maggiore di uno.

Caso pluridimensionale

Se la f è n -dimensionale ed esiste, in un intorno I_ξ di uno zero ξ della f , uno sviluppo di Taylor almeno di ordine 2, è possibile, analogamente a quanto avviene nel caso monodimensionale, ottenere una funzione di iterazione n -dimensionale $\Phi(x)$. Infatti, in tale ipotesi, indicato con ξ uno zero della f e con $x^{(0)}$ un punto iniziale di I_ξ , in prima approssimazione risulta

$$0 \simeq f(x^{(0)}) + F'(x^{(0)})(\xi - x^{(0)})$$

essendo

$$F'(x^{(0)}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^{(0)}) & \frac{\partial f_1}{\partial x_2}(x^{(0)}) & \dots & \frac{\partial f_1}{\partial x_n}(x^{(0)}) \\ \vdots & & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x^{(0)}) & \frac{\partial f_n}{\partial x_2}(x^{(0)}) & \dots & \frac{\partial f_n}{\partial x_n}(x^{(0)}) \end{pmatrix}$$

la matrice Jacobiana della f in $x^{(0)}$. Pertanto se in I_ξ la matrice Jacobiana $F'(x)$ è non singolare si può introdurre la funzione di iterazione n -dimensionale

$$\Phi(x) = x - [F'(x)]^{-1}f(x) \tag{C.13}$$

detta di Newton in quanto per $n = 1$ si riduce alla (C.12).

Esempio C.3.1 Valutare, a meno di $1/1000$, mediante il metodo iterativo generato dalla funzione di iterazione (C.13), la soluzione positiva del sistema:

$$\begin{cases} f_1(x_1, x_2) = x_1 + 3 \log_{10} x_1 + x_2^2 = 0 \\ f_2(x_1, x_2) = 2x_1^2 - x_1x_2 - 5x_1 + 1 = 0 \end{cases} .$$

Uno studio approssimativo del sistema nel quadrante positivo mette in evidenza che le due funzioni f_1 ed f_2 si annullano simultaneamente in un intorno del punto $(3.4, 2.2)$. Posto pertanto $x^{(0)} = (3.4 \ 2.2)^T$, per precauzione, si

eseguono le iterazioni con 5 cifre decimali. Da tale scelta deriva che

$$f(x^{(0)}) = \begin{pmatrix} 3.4 + 3 \log_{10} 3.4 + (2.2)^2 \\ 2(3.4)1^2 - (3.4)(2.2) - 5(3.4) + 1 \end{pmatrix} \simeq \begin{pmatrix} 0.1544 \\ -0.3600 \end{pmatrix},$$

$$F'(x^{(0)}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^{(0)}) & \frac{\partial f_1}{\partial x_2}(x^{(0)}) \\ \frac{\partial f_2}{\partial x_1}(x^{(0)}) & \frac{\partial f_2}{\partial x_2}(x^{(0)}) \end{pmatrix} \simeq \begin{pmatrix} 1.3832 & -4.4 \\ 6.4 & -3.4 \end{pmatrix},$$

$$\det F'(x^{(0)}) \simeq 23.4571 \neq 0.$$

La matrice $F'(x^{(0)})$ è dunque non singolare e

$$[F'(x^{(0)})]^{-1} = \frac{1}{23.457} \begin{pmatrix} -3.4 & 4.4 \\ -6.4 & 1.3832 \end{pmatrix}.$$

Di conseguenza

$$x^{(1)} = x^{(0)} - [F'(x^{(0)})]^{-1} f(x^{(0)})$$

$$= \begin{pmatrix} 3.4 \\ 2.2 \end{pmatrix} - \frac{1}{23.457} \begin{pmatrix} -3.4 & 4.4 \\ -6.4 & 1.3832 \end{pmatrix} \begin{pmatrix} 0.1544 \\ -0.3600 \end{pmatrix} = \begin{pmatrix} 3.4899 \\ 2.2633 \end{pmatrix}.$$

Iterando nuovamente di ottengono

$$x^{(2)} = \begin{pmatrix} 3.4891 \\ 2.2621 \end{pmatrix}, \quad x^{(3)} = \begin{pmatrix} 3.4875 \\ 2.2621 \end{pmatrix}, \quad \text{con } f(x^{(3)}) = \begin{pmatrix} 0.0005 \\ 0.0000 \end{pmatrix},$$

e pertanto si può considerare $x^{(3)}$ un'approssimazione accettabile della soluzione del sistema.

C.3.2 Risultati sulla convergenza dei metodi iterativi

Nel seguito, sia nel caso monodimensionale sia in quello pluridimensionale, faremo sempre riferimento a metodi iterativi del tipo

$$x^{(i+1)} = \Phi(x^{(i)}) \quad i = 0, 1, \dots, \quad (\text{C.14})$$

dove $x^{(0)}$ è il punto iniziale e Φ è la funzione di iterazione. Supporremo inoltre che la Φ possieda almeno un punto fisso ξ ed indicheremo con I_ξ un suo intorno. Come abbiamo già osservato quest'ultima ipotesi equivale a supporre che ξ sia una soluzione del problema iniziale.

Definizione C.3.2 Indicato con ξ un punto fisso di Φ , con I_ξ un suo intorno e con $x^{(0)}$ un punto di I_ξ , se per la successione (C.14) vale la disuguaglianza

$$\|x^{(i+1)} - \xi\| \leq c \|x^{(i)} - \xi\|^p \quad i = 0, 1, \dots, \quad (\text{C.15})$$

essendo

$$0 \leq c \text{ per } p > 1 \text{ e } 0 \leq c < 1 \text{ per } p = 1, \quad (\text{C.16})$$

il metodo iterativo è **almeno di ordine p** . È esattamente di ordine p se non esiste un valore più elevato di p per cui valgano le (C.15)-(C.16)

Teorema C.3.3 Ogni metodo di ordine p è localmente convergente. Esiste cioè un intorno I_ξ di ξ tale che, qualunque sia $x^{(0)} \in I_\xi$, la successione $x^{(i+1)} = \Phi(x^{(i)})$ è convergente a ξ .

Dimostrazione. Se $p = 1$,

$$\|x^{(i+1)} - \xi\| \leq c \|x^{(i)} - \xi\| \leq \dots \leq c^{i+1} \|x^{(0)} - \xi\|.$$

Poiché $0 \leq c < 1$, $\|x^{(i)} - \xi\| \rightarrow 0 \forall x^{(0)} \in I_\xi$, per cui la successione degli iterati converge a ξ qualunque sia il punto iniziale $x^{(0)} \in I_\xi$.

Se $p > 1$

$$\|x^{(1)} - \xi\| \leq (c \|x^{(0)} - \xi\|^{p-1}) \|x^{(0)} - \xi\| = c' \|x^{(0)} - \xi\|$$

con $c' = c \|x^{(0)} - \xi\|^{p-1} < 1 \forall x^{(0)} \in I_\xi$, essendo I_ξ un intorno sufficientemente piccolo di ξ . Di conseguenza

$$\|x^{(i+1)} - \xi\| \leq (c')^{p^i + p^{i-1} + \dots + 1} \|x^{(0)} - \xi\| \quad i = 0, 1, \dots,$$

e la successione $\{x^{(i)}\}$ converge a ξ ancor più rapidamente che nel caso precedente. ■

Definizione C.3.4 Un metodo iterativo è globalmente convergente quando come intorno I_ξ si può considerare l'intero spazio \mathbb{R}^n , nell'ipotesi che la sua funzione di iterazione sia n -dimensionale.

Esempio C.3.5 Il metodo iterativo $x^{(i+1)} = \cos x^{(i)}$ è globalmente convergente. Infatti, indicato con ξ il punto fisso dell'equazione $x = \cos x$ ed osservato che $\xi \in (0, 1)$, per ogni $x^{(0)} \in \mathbb{R}$ si ha che:

$$\begin{aligned} x^{(1)} &= \cos x^{(0)} \in [-1, 1] \\ x^{(i+1)} &= \cos x^{(i)} = \cos \xi - (x^{(i)} - \xi) \sin \theta^{(i)} \\ \theta^{(i)} &\in (x^{(i)}, \xi) \subset [-1, 1] \quad i = 1, 2, \dots \end{aligned}$$

Pertanto, essendo per ogni i

$$|x^{(i+1)} - \xi| = |x^{(i+1)} - \cos \xi| = |\sin \theta^{(i)}| |x^{(i)} - \xi| < (\sin 1) |x^{(i)} - \xi|,$$

il metodo è globalmente convergente e la sua convergenza è lineare (convergenza ordine 1).

Definizione C.3.6 Una funzione n - dimensionale è contrattiva in un intervallo $I \subset \mathbb{R}^n$, se esiste una costante $\kappa \in [0, 1)$ tale che

$$\|f(x') - f(x'')\| \leq \kappa \|x' - x''\| \quad \forall x', x'' \in I.$$

Esempio C.3.7 Il polinomio $p(x) = x^2 + 2x$ è contrattivo in $[-\frac{5}{4}, -\frac{2}{3}]$.

Infatti

$$|(x')^2 + 2(x') - (x'')^2 - 2x''| = |x' + x'' + 2| |x' - x''|$$

e

$$|x' + x'' + 2| \leq \frac{2}{3}$$

per $x', x'' \in [-\frac{5}{4}, -\frac{2}{3}]$. Lo stesso polinomio non è contrattivo in $[-\frac{5}{4}, -\frac{1}{2}]$, essendo, in questo intervallo, $|x' + x'' + 2| \leq 1$, con l'uguale valido per $x' = x'' = -1/2$.

Teorema C.3.8 Nel caso monodimensionale, condizione sufficiente perché una funzione f , derivabile con continuità in un intervallo $[a, b]$, sia contrattiva è che esista un numero $m \in [0, 1)$ tale che

$$|f'(x)| \leq m < 1 \quad \forall x \in [a, b].$$

Dimostrazione. Per il teorema del valor medio, applicato ad un intervallo $[x', x''] \subset [a, b]$,

$$f(x') - f(x'') = f'(\theta)(x' - x'') \quad \theta \in (x', x'').$$

Pertanto, essendo $|f'(\theta)| \leq m < 1$ per ipotesi,

$$|f(x') - f(x'')| \leq m |x' - x''|$$

con $0 \leq m < 1 \quad \forall x', x'' \in [a, b]$, e quindi la f è contrattiva in $[a, b]$. ■

Esempio C.3.9 Per $a \in [-\frac{1}{2}, \frac{1}{2}]$, la funzione $f(x) = a \log(x+1) - x/3$ è contrattiva in $[0,1]$. Infatti, essendo

$$f'(x) = \frac{a}{x+1} - \frac{1}{3},$$

basta osservare che per $a \in [-\frac{1}{2}, \frac{1}{2}]$ risulta

$$\left| \frac{a}{x+1} - \frac{1}{3} \right| \leq \frac{5}{6} < 1 \quad \forall x \in [0, 1].$$

Teorema C.3.10 *Nel caso n -dimensionale, **condizione sufficiente** perché una funzione f sia contrattiva in un intervallo $I \subset \mathbb{R}^n$ è che esista un numero $\lambda \in [0, 1)$ tale che*

$$\sum_{j=1}^n \left| \frac{\partial f_i(x)}{\partial x_j} \right| \leq \lambda < 1 \quad \text{per } i = 1, \dots, n, \quad \forall x \in I. \quad (\text{C.17})$$

Dimostrazione. Per la formula di Taylor di primo ordine, per $i = 1, 2, \dots, n$ e per ogni coppia di punti $x', x'' \in I$,

$$f_i(x') - f_i(x'') = \sum_{j=1}^n \frac{\partial f_i(x)}{\partial x_j} (x'_j - x''_j)$$

essendo $\theta^{(i,j)}$ un punto interno all'intervallo di estremi x', x'' . Di conseguenza, dato che $\theta^{(i,j)} \in I$, per la (C.17)

$$\begin{aligned} |f_i(x') - f_i(x'')| &\leq \sum_{j=1}^n \left| \frac{\partial \theta^{(i,j)}}{\partial x_j} \right| |x'_j - x''_j| \\ \|x' - x''\|_\infty \sum_{j=1}^n \left| \frac{\partial \theta^{(i,j)}}{\partial x_j} \right| &\leq \lambda \|x' - x''\|_\infty. \end{aligned}$$

Poiché la suddetta disuguaglianza vale per $i = 1, \dots, n$, risulta anche

$$\|f(x') - f(x'')\|_\infty \leq \lambda \|x' - x''\|_\infty$$

e pertanto la f è contrattiva. ■

Esempio C.3.11 Dimostrare che la funzione

$$f(x) = \begin{pmatrix} f_1(x_1, x_2) = \sqrt{[x_1(x_2 + 5) - 1]/2} \\ f_2(x_1, x_2) = \sqrt{3 + \log_{10} x_1} \end{pmatrix}$$

è contrattiva nell'insieme

$$I = \{(x_1, x_2) \text{ con } |3.5 - x_1| \leq 0.1 \text{ e } |2.2 - x_2| \leq 0.1\}.$$

In tale intervallo

$$\begin{aligned} \left| \frac{\partial f_1}{\partial x_1} \right| &\leq \frac{2.3 + 5}{4\sqrt{[3.4(2.1 + 5) - 1]/2}} < 0.54, \\ \left| \frac{\partial f_1}{\partial x_2} \right| &\leq \frac{3.6}{4\sqrt{[3.4(2.1 + 5) - 1]/2}} < 0.27, \\ \left| \frac{\partial f_2}{\partial x_1} \right| &\leq \frac{1 + \frac{3(0.43)}{3.4}}{2\sqrt{3.4 + 2 \log_{10} 3.4}} < 0.42, \\ \left| \frac{\partial f_2}{\partial x_2} \right| &= 0. \end{aligned}$$

Di conseguenza, essendo

$$\begin{aligned} \left| \frac{\partial f_1}{\partial x_1} \right| + \left| \frac{\partial f_1}{\partial x_2} \right| &< 0.81 < 1, \\ \left| \frac{\partial f_2}{\partial x_1} \right| + \left| \frac{\partial f_2}{\partial x_2} \right| &< 0.42 < 1, \end{aligned}$$

la funzione $f(x)$ è contrattiva in I .

Teorema C.3.12 *Supponiamo che in un intervallo chiuso $[a, b]$ esista una sola soluzione ξ di un'equazione monodimensionale $f(x) = 0$. Allora, indicata con $\Phi(x)$ una sua funzione di iterazione, se*

- 1) $\Phi(x)$ è derivabile con continuità in $[a, b]$;
- 2) il punto iniziale $x^{(0)}$ e tutti gli iterati appartengono ad $[a, b]$;
- 3) esiste un numero $m \in [0, 1)$ con $|\Phi'(x)| \leq m$ in $[a, b]$,

la successione degli iterati converge a ξ .

Se in particolare $0 < \Phi'(x) \leq m$ in $[a, b]$, la convergenza è monotona e lineare. Se invece $-1 < m \leq \Phi'(x) < 0$, la convergenza è ancora lineare ma alternata.

Dimostrazione. Per la definizione stessa di funzione di iterazione

$$\begin{aligned} x^{(1)} - \xi &= \Phi(x^{(0)}) - \Phi(\xi) = (x^{(0)} - \xi)\Phi'(\theta^{(0)}) \\ x^{(2)} - \xi &= \Phi(x^{(1)}) - \Phi(\xi) = (x^{(1)} - \xi)\Phi'(\theta^{(1)}) \\ &\vdots \\ x^{(i+1)} - \xi &= \Phi(x^{(i)}) - \Phi(\xi) = (x^{(i)} - \xi)\Phi'(\theta^{(i)}). \end{aligned}$$

Pertanto, per l'ipotesi 3),

$$|x^{(i+1)} - \xi| = |\Phi'(\theta^{(i)})| \dots |\Phi'(\theta^{(0)})| |x^{(0)} - \xi| \leq m^{i+1} |x^{(0)} - \xi|$$

e quindi $x^{i+1} \rightarrow \xi$ linearmente. ■

Osservazione C.3.13 *L'appartenenza degli iterati all'intervallo $[a, b]$ deve essere verificata durante l'esecuzione dei calcoli.*

Teorema C.3.14 *Supponiamo che in un intervallo chiuso $I = \{x \in \mathbb{R}^n \text{ con } a_i \leq x_i \leq b_i \text{ per } i = 1, \dots, n\}$ esiste una sola soluzione ξ dell'equazione n - dimensionale $f(x) = 0$ e quindi un solo punto fisso per una sua funzione di iterazione $\Phi(x)$. Allora, se:*

- 1) *esistono continue in I le derivate parziali della Φ ;*
- 2) *il punto iniziale $x^{(0)}$ e tutti gli iterati successivi appartengono ad I ;*
- 3) *per $i = 1, \dots, n$ valgono in I le disuguaglianze*

$$\left| \frac{\partial \varphi_i}{\partial x_1} \right| + \dots + \left| \frac{\partial \varphi_i}{\partial x_n} \right| \leq m_i < 1$$

essendo $\Phi = (\varphi_1, \dots, \varphi_n)$,

la successione degli iterati converge a ξ .

Esempio C.3.15 Vautare, a meno di $1/1000$, la soluzione positiva del sistema

$$\begin{cases} f_1(x, y) = 2x^2 - xy - 5x + 1 = 0 \\ f_2(x, y) = x + 3 \log_{10} x - y^2 = 0 \end{cases} .$$

Allo scopo di trovare un intervallo nel quale cade la soluzione positiva del sistema, costruiamo le curve $f_1(x, y) = 0$, $f_2(x, y) = 0$. La soluzione positiva, ottenuta graficamente, è

$$x^{(0)} = 3.5 \quad y^{(0)} = 2.2.$$

Per ottenere una funzione di iterazione Φ mettiamo il sistema nella forma seguente:

$$\begin{aligned} x &= \varphi_1(x, y) = \sqrt{\frac{x(y+5) - 1}{2}} \\ y &= \varphi_2(x, y) = \sqrt{3 + \log_{10} x}. \end{aligned}$$

Supponiamo inoltre che l'intervallo I sia così definito:

$$I = \{(x, y) \text{ con } |3.5 - x| \leq 0.1 \text{ e } |2.2 - y| \leq 0.1\}.$$

In tale intervallo

$$\begin{aligned} \left| \frac{\partial \varphi_1}{\partial x_1} \right| &\leq \frac{2.3 + 5}{4\sqrt{[3.4(2.1 + 5) - 1]/2}} < 0.54, \\ \left| \frac{\partial \varphi_1}{\partial x_2} \right| &\leq \frac{3.6}{4\sqrt{[3.4(2.1 + 5) - 1]/2}} < 0.27, \\ \left| \frac{\partial \varphi_2}{\partial x_1} \right| &\leq \frac{1 + \frac{3(0.43)}{3.4}}{2\sqrt{3.4 + 2 \log_{10} 3.4}} < 0.42, \\ \left| \frac{\partial \varphi_2}{\partial x_2} \right| &= 0. \end{aligned}$$

Di conseguenza, essendo

$$\begin{aligned} \left| \frac{\partial \varphi_1}{\partial x_1} \right| + \left| \frac{\partial \varphi_1}{\partial x_2} \right| &< 0.81 < 1, \\ \left| \frac{\partial \varphi_2}{\partial x_1} \right| + \left| \frac{\partial \varphi_2}{\partial x_2} \right| &< 0.42 < 1, \end{aligned}$$

se come punto iniziale di prende la soluzione grafica e se gli iterati successivi rimangono in I (ciò che deve essere verificato durante la esecuzione dei calcoli), il processo iterativo è convergente.

Calcolando gli iterati mediante le formule

$$x^{(i)} = \sqrt{\frac{x^{(i-1)}(y^{(i-1)} + 5) - 1}{2}}$$

$$y^{(i)} = \sqrt{3 + \log_{10} x^{(i-1)}} \quad i = 1, 2, \dots,$$

si costruisce la seguente tabella

$x^{(0)} = 3.5$	$y^{(0)} = 2.2$
$x^{(1)} = 3.479$	$y^{(1)} = 2.259$
$x^{(2)} = 3.481$	$y^{(2)} = 2.260$
$x^{(3)} = 3.484$	$y^{(3)} = 2.261$
$x^{(4)} = 3.486$	$y^{(4)} = 2.261$
$x^{(5)} = 3.487$	$y^{(5)} = 2.262$
$x^{(6)} = 3.487$	$y^{(6)} = 2.262.$

Avendo ottenuto $x^{(6)} = x^{(5)} = 3.487$ e $y^{(6)} = y^{(5)} = 2.262$, si può porre $\xi^{(1)} = 3.487$ e $\xi^{(2)} = 2.262$

Osservazione C.3.16 *L'ipotesi 3) dei teoremi C.3.12 e C.3.14, può essere sostituita dall'ipotesi meno restrittiva della contrattività della f , senza che questo comporti modifiche significative nella dimostrazione dei teoremi. La formulazione più restrittiva è dovuta semplicemente al fatto che, operativamente, la contrattività della funzione di iterazione viene quasi sempre verificata sfruttando la condizione sufficiente espressa dall'ipotesi 3).*

Il seguente teorema è praticamente più utile del teorema C.3.14, in quanto:

- 1) prescinde dalla conoscenza di un intervallo contenente il punto fisso ξ ;
- 2) fornisce una maggiorazione della distanza tra punto fisso ed i -esimo iterato, computazionalmente utilizzabile come criterio di interruzione dei calcoli.

Teorema C.3.17 *Siano:*

- Φ una funzione di iterazione definita in un insieme Ω ,
- $x^{(0)}$ un punto iniziale appartenente da Ω ,
- $I_{x^{(0)},r}$ un intorno circolare con centro $x^{(0)}$ e raggio r .

Allora, se esiste un numero $\kappa \in [0, 1)$ tale che:

- 1) $\|\Phi(x') - \Phi(x'')\| \leq \kappa \|x' - x''\| \quad \forall x', x'' \in I_{x^{(0)},r};$
- 2) $\|x^{(1)} - x^{(0)}\| = \|\Phi(x^{(0)}) - x^{(0)}\| \leq (1 - \kappa)r,$

risultano valide le seguenti affermazioni:

- 1') $x^{(i)} = \Phi(x^{(i-1)}) \in I_{x^{(0)},r}$ per $i = 1, 2, \dots,$
- 2') Φ possiede in $I_{x^{(0)},r}$ un solo punto fisso ξ ,
- 3') gli iterati convergono a ξ secondo la relazione seguente

$$\|x^{(i)} - \xi\| \leq \frac{\kappa^i}{1 - \kappa} \|x^{(1)} - x^{(0)}\|. \quad (\text{C.18})$$

Dimostrazione. La 1') può essere verificata per induzione. Poiché $x^{(1)} \in I_{x^{(0)},r}$ per la 2), resta da dimostrare che se $x^{(1)}, \dots, x^{(i)} \in I_{x^{(0)},r}$ anche $x^{(i+1)} \in I_{x^{(0)},r}$.

Effettivamente in tale ipotesi $x^{(i+1)} \in I_{x^{(0)},r}$ in quanto, essendo

$$\begin{aligned} \|x^{(i+1)} - x^{(i)}\| &= \|\Phi(x^{(i)}) - \Phi(x^{(i-1)})\| \leq \kappa \|x^{(i)} - x^{(i-1)}\| \\ &\leq \kappa^i \|x^{(1)} - x^{(0)}\|, \end{aligned} \quad (\text{C.19})$$

per la disuguaglianza triangolare e per la 2),

$$\begin{aligned} \|x^{(i+1)} - x^{(0)}\| &\leq \|x^{(i+1)} - x^{(i)}\| + \|x^{(i)} - x^{(i-1)}\| + \dots + \|x^{(1)} - x^{(0)}\| \\ &\leq (\kappa^i + \kappa^{i-1} + \dots + 1) \|x^{(1)} - x^{(0)}\| \leq (1 - \kappa^{i+1})^{-1} \|x^{(1)} - x^{(0)}\| < r. \end{aligned}$$

Per verificare la 2') dimostriamo dapprima che la successione $\{x^{(i)}\}$ è di Cauchy. Dalla disuguaglianza (C.19) e dall'ipotesi 2) segue che, per $h = i + j$,

$$\|x^h - x^{(i)}\| \leq \|x^{(h)} - x^{(h-1)}\| + \dots + \|x^{(i+1)} - x^{(i)}\| \leq (\kappa^{(i+j-1)} + \kappa^{(i+j-2)} + \dots + \kappa^i) \|x^{(1)} - x^{(0)}\| \quad (\text{C.20})$$

$$\leq \kappa^i \frac{1 - \kappa^j}{1 - \kappa} \|x^{(1)} - x^{(0)}\| < \frac{\kappa^i}{1 - \kappa} \|x^{(1)} - x^{(0)}\| < \kappa^i \|x^{(1)} - x^{(0)}\|$$

Pertanto, poiché $\{\kappa^i r\}$ è strettamente decrescente a zero, la successione $\{x^{(i)}\}$ è di Cauchy. Poiché in \mathbb{R}^n le successioni di Cauchy sono convergenti, esiste un vettore $\xi \in \mathbb{R}^n$ al quale converge $\{x^{(i)}\}$ ed inoltre, essendo ξ un punto di accumulazione per $I_{x^{(0)}, r}$, ξ appartiene alla chiusura di $I_{x^{(0)}, r}$ ($\xi \in \bar{I}_{x^{(0)}, r}$). Inoltre ξ è un punto fisso per la funzione di interazione, essendo, $\forall \varepsilon \in \mathbb{R}^+$ ed i sufficientemente elevato,

$$\begin{aligned} \|\Phi(\xi) - \xi\| &\leq \|\Phi(\xi) - \Phi(x^{(i)})\| + \|\Phi(x^{(i)}) - \xi\| \\ &\leq \kappa \|\xi - x^{(i)}\| + \|x^{(i+1)} - \xi\| < \varepsilon \end{aligned}$$

per il fatto che $x^{(i)} \rightarrow \xi$.

L'unicità del punto fisso segue immediatamente dall'osservazione che se ξ' fosse un altro punto fisso di avrebbe

$$\|\xi - \xi'\| = \|\Phi(\xi) - \Phi(\xi')\| \leq \kappa \|\xi - \xi'\|$$

e pertanto, dato che $\kappa \in [0, 1)$, deve essere $\xi = \xi'$. Infine per la stima (C.18) dell'errore basta osservare che, in base alla (C.20)

$$\lim_{h \rightarrow \infty} \|x^{(h)} - x^{(i)}\| = \|\xi - x^{(i)}\| \leq \frac{\kappa^i}{1 - \kappa} \|x^{(1)} - x^{(0)}\|$$

■

Esempio C.3.18 Valutare la radice positiva dell'equazione

$$f(x) = x - \sin x - \frac{1}{4} = 0.$$

Per avere una comoda funzione di iterazione si esprime l'equazione nella forma

$$x = \sin x + \frac{1}{4} = \Phi(x).$$

Si può allora osservare graficamente che l'unico punto fisso della funzione di iterazione, e quindi l'unica radice dell'equazione, è compresa nell'intervallo $(1, 2)$. Posto $x^{(0)} = 3/2$ ed $r = 1/2$, è facile verificare che le ipotesi 1) – 2) sono verificate in quanto:

- 1) $|\Phi'(x)| = |\cos x| < \cos 1 < 1$,
- 2) $|x^{(1)} - x^{(0)}| = \left| \sin \frac{3}{2} + \frac{1}{4} - \frac{3}{2} \right| = \frac{5}{4} - \sin \frac{3}{2} < \frac{1}{2}(1 - \cos 1)$.

Il metodo iterativo

$$x^{(i)} = \sin x^{(i-1)} + \frac{1}{4} \quad i = 1, 2, \dots, \quad x^{(0)} = \frac{3}{2}$$

converge pertanto al punto fisso ξ ed inoltre, per la (C.18),

$$|x^{(i)} - \xi| < \frac{(\cos 1)^i}{1 - \cos 1} \left(\frac{5}{4} - \sin \frac{3}{2} \right).$$

C.3.3 I metodi della tangente, e della secante

Il metodo della tangente o di Newton nel caso monodimensionale

Il metodo di Newton è formalmente operativo per la risoluzione di un'equazione monodimensionale $f(x) = 0$, solo nel caso in cui la f è derivabile in un intervallo contenente tutti gli iterati e purché in essi la derivata sia diversa da zero.

Sotto tali ipotesi, prefissato un punto iniziale $x^{(0)}$ nel suddetto intervallo, i successivi iterati sono ottenibili mediante la seguente relazione:

$$x^{(i+1)} = x^{(i)} - \frac{f(x^{(i)})}{f'(x^{(i)})} \quad i = 0, 1, \dots \quad (\text{C.21})$$

Geometricamente il metodo di Newton corrisponde alla sostituzione di un piccolo arco della curva $y = f(x)$ con la tangente alla curva condotta a partire da un punto opportuno.

Teorema C.3.19 *Se in un intervallo $[a, b]$ valgono le seguenti ipotesi:*

- 1) $f(a)f(b) < 0$;
- 2) la f è derivabile due volte con continuità in $[a, b]$;
- 3) $f'(x)$ e $f''(x)$ sono di segno costante in $[a, b]$.

Se inoltre, come punto iniziale $x^{(0)}$, si prende a qualora sia $f(a)f''(a) > 0$ oppure b qualora sia $f(b)f''(b) > 0$, la successione $\{x^{(i)}\}$, ottenuta con l'algoritmo (C.21), converge monotonamente a ξ .

Dimostrazione. Sia, per esempio, $f(a) < 0$, $f(b) > 0$, $f'(x) > 0$, $f''(x) > 0$ per $a \leq x \leq b$ (la dimostrazione nelle altre possibili situazioni è del tutto analoga). Poiché $f(b)f''(b) > 0$, poniamo $x^{(0)} = b$. La 1) assicura l'esistenza di almeno uno zero ξ per la f in $[a, b]$. La non variazione di segno di $f'(x)$ in $[a, b]$ ne assicura l'unicità. Mostriamo ora per induzione che ogni iterato $x^{(i)} > \xi$. Supponendo che $x^{(i)} > \xi$ per $j = 0, 1, \dots, i$, dobbiamo ora dimostrare che $x^{(i+1)} > \xi$. Sviluppando con la formula di Taylor $f(\xi)$, dopo aver posto

$$\xi = x^{(i)} + (\xi - x^{(i)}),$$

si ha

$$f(\xi) = 0 = f(x^{(i)}) + f'(x^{(i)})(\xi - x^{(i)}) + \frac{1}{2}f''(\theta^{(i)})(\xi - x^{(i)})^2$$

con $\xi < \theta^{(i)} < x^{(i)}$. Poiché $f''(x) > 0$, risulta

$$f(x^{(i)}) + f'(x^{(i)})(\xi - x^{(i)}) < 0$$

e di conseguenza, essendo

$$x^{(i+1)} = x^{(i)} - \frac{f(x^{(i)})}{f'(x^{(i)})} > \xi, \quad \xi < x^{(i)} - \frac{f(x^{(i)})}{f'(x^{(i)})} = x^{(i+1)},$$

la dimostrazione è completa. ■

Poiché $f'(x^{(i)}) > 0$ per ipotesi ed $f(x^{(i)}) > 0$ in quanto $x^{(i)} > \xi$ ed $f(\xi) = 0$, la successione $\{x^{(i)}\}$, ottenuta nel modo suddetto, è monotona decrescente e inferiormente limitata da ξ . Esiste dunque un limite $\bar{\xi} = \lim_{i \rightarrow \infty} x^{(i)}$. Passando al limite nell'uguaglianza (C.21) si ha

$$\bar{\xi} = \bar{\xi} - \frac{f(\bar{\xi})}{f'(\bar{\xi})}$$

da cui $f(\bar{\xi}) = 0$ ed infine, per l'unicità degli zeri della f in $[a, b]$, $\bar{\xi} = \xi$.

Talvolta, come è dimostrato nel seguente teorema, è possibile valutare la distanza dell' i -esimo iterato della radice dell'equazione.

Teorema C.3.20 *Se esistono due numeri positivi m_1 ed M_1 tali che:*

$$1) \quad m_1 \leq |f'(x)|, \quad \forall x \in [a, b];$$

2) $|f''(x)| \leq M_2, \forall x \in [a, b];$

tra l' i -esimo iterato di Newton e la radice ξ dell'equazione $f(x) = 0$ vale la seguente relazione:

$$|x^{(i)} - \xi| < \frac{M_2}{2m_1}(x^{(i)} - x^{(i-1)})^2. \quad (\text{C.22})$$

Dimostrazione. Per il teorema del valor medio, applicato ad un generico intervallo $[\bar{x}, \xi] \subset [a, b]$ con $f(\xi) = 0$ ed $f(\bar{x}) \neq 0$,

$$f(\bar{x}) - f(\xi) = (\bar{x} - \xi)f'(\theta) \quad \theta \in (\bar{x}, \xi).$$

Pertanto, se esiste un numero positivo m_1 tale che $|f'(x)| > m_1$ per $x \in [a, b]$, risulta

$$|f(\bar{x})| = |f(\bar{x}) - f(\xi)|m_1|\bar{x} - \xi|$$

e quindi anche

$$|\bar{x} - \xi| \leq \frac{|f(\bar{x})|}{m_1}. \quad (\text{C.23})$$

Inoltre, applicando la formula di Taylor di ordine due,

$$\begin{aligned} f(x^{(i)}) &= f[x^{(i-1)} + (x^{(i)} - x^{(i-1)})] = f(x^{(i-1)}) + f'(x^{(i-1)})(x^{(i)} - x^{(i-1)}) \\ &\quad + \frac{1}{2}f''(\theta^{(i)})(x^{(i)} - x^{(i-1)})^2 \quad \theta^{(i)} \in (x^{(i-1)}, x^{(i)}). \end{aligned}$$

Poiché, per costruzione, nel metodo di Newton

$$f(x^{(i-1)}) + f'(x^{(i-1)})(x^{(i)} - x^{(i-1)}) = 0,$$

dalla formula di Taylor di secondo ordine deriva che

$$|f(x^{(i)})| \leq \frac{M_2}{2}(x^{(i)} - x^{(i-1)})^2, \quad (\text{C.24})$$

essendo

$$|f''(x)| \leq M_2 \quad \text{in } [a, b].$$

Infine, utilizzando la (C.23) e la (C.24), si ha che

$$|x^{(i)} - \xi| < \frac{M_2}{2m_1}(x^{(i)} - x^{(i-1)})^2.$$

■

Esempio C.3.21 Valutare $\sqrt{2}$ con un errore dell'ordine di 10^{-5} , con il metodo di Newton.

Cosideriamo, allo scopo, $f(x) = x^2 - 2$ nell'intervallo $[1, 2]$. Il metodo di Newton è applicabile in quanto:

- 1) $f(a)f(b) = (-1)(2) < 0$;
- 2) $f'(x) = 2x$ è di segno costante in $[1, 2]$;
- 3) $f''(x) = 2 > 0$ in $[1, 2]$.

Poiché $f(2)f''(2) = 4 > 0$, come punto iniziale si prende 2. Pertanto

$$x^{(1)} = x^{(0)} - \frac{f(x^{(0)})}{f'(x^{(0)})} = 2 - \frac{2}{4} = \frac{3}{2},$$

$$x^{(2)} = \frac{17}{12}, x^{(3)} \simeq 1.41422.$$

Essendo, per la (C.22),

$$|x^{(3)} - \sqrt{2}| \leq \frac{M_2}{2m_1}(x^{(3)} - x^{(2)}) \leq \frac{2}{4}(1.41422 - 1.41667)^2 \simeq 3 \times 10^{-6},$$

$x^{(3)}$ rappresenta un'approssimazione valida di $\sqrt{2}$.

Teorema C.3.22 *Se ξ è uno zero semplice per la f ($f(\xi) = 0, f'(\xi) \neq 0$), il metodo di Newton ha convergenza ad secondo ordine. Se ξ è uno zero multiplo ($f(\xi) = \dots = f^{(m-1)}(\xi) = 0, f^{(m)} \neq 0, m > 1$), il metodo di Newton ha convergenza lineare.*

Dimostrazione. • ξ zero semplice

Essendo

$$\Phi'(\xi) = \frac{f(\xi)f''(\xi)}{(f'(\xi))^2} = 0,$$

per la formula di Taylor di ordine due,

$$x^{(i+1)} = \Phi(x^{(i)}) = \Phi(\xi) + \frac{(x^{(i)} - \xi)^2}{2}\Phi''(\theta^{(i)}) \quad \theta^{(i)} \in (x^{(i)}, \xi),$$

da cui, posto $|\Phi''(x)| \leq M_2$ in $[a, b]$,

$$|x^{(i+1)} - \xi| \leq \frac{M_2}{2}|x^{(i)} - \xi|^2$$

e quindi il metodo ha convergenza quadratica.

- **ξ zero multiplo**

Essendo ξ uno zero m -plo, esiste una $g(x)$ con $g(\xi) \neq 0$ tale che

$$f(x) = (x - \xi)^m g(x).$$

Di conseguenza, se $g(x)$ è derivabile,

$$\Phi'(\xi) = 1 - \frac{1}{m} \neq 0.$$

Pertanto, per la formula del valor medio,

$$x^{(i+1)} = \Phi(x^{(i)}) = \Phi(\xi) + (x^{(i)} - \xi)\Phi'(\theta^{(i)}) \quad \theta^{(i)} \in (x^{(i)}, \xi),$$

da cui, posto $|\Phi'(x)| \leq M_1$ in $[a, b]$,

$$|x^{(i+1)} - \xi| \leq M_1|x^{(i)} - \xi|.$$

L'ultima disuguaglianza dimostra che, essendo almeno localmente $M_1 < 1$, il metodo è convergente con convergenza lineare. ■

Metodo di Newton nel caso multidimensionale

Nel caso multidimensionale la successione di Newton è generata dalla relazione ricorsiva

$$x^{(i+1)} = x^{(i)} - [JF(x^{(i)})]^{-1}F(x^{(i)})$$

dove

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad F(x) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix}$$

e $[JF(x^{(i)})]^{-1}$ è la inversa della matrice jacobiana

$$JF(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & \cdots & \frac{\partial f_n(x)}{\partial x_n} \end{pmatrix}.$$

Per una discussione sulle condizioni sufficienti per la convergenza della successione e per l'ordine di convergenza del metodo si rinvia al capitolo cinque del volume di Stoer.

Teorema C.3.23 *Se la matrice Jacobiana $JF(x)$ è nonsingolare nel punto fisso ξ e la F è differenziabile due volte con continuità in un intorno di ξ , il metodo di Newton è del secondo ordine (convergenza quadratica).*

Il metodo della secante o della falsa posizione nel caso monodimensionale

Il metodo della secante, nel caso monodimensionale viene utilizzato per risolvere in un intervallo $[a, b]$ un'equazione $f(x) = 0$, nell'ipotesi che in $[a, b]$ essa abbia una sola radice. Algoritmicamente, essa si presenta in uno dei modi seguenti:

$$x^{(i+1)} = x^{(i)} - \frac{f(x^{(i)})}{f(x^{(i)}) - f(a)}(x^{(i)} - a), \quad (\text{C.25})$$

con $x^{(0)} = b$ e $i = 0, 1, \dots$;

$$x^{(i+1)} = x^{(i)} - \frac{f(x^{(i)})}{f(b) - f(x^{(i)})}(b - x^{(i)}), \quad (\text{C.26})$$

con $x^{(0)} = a$ e $i = 0, 1, \dots$.

Geometricamente corrisponde alla sostituzione, nell'intervallo $[a, b]$, della curva $y = f(x)$ con la corda che unisce i punti $(a, f(a))$ e $(b, f(b))$.

Teorema C.3.24 *Se in un intervallo $[a, b]$ sono soddisfatte le seguenti ipotesi:*

- 1) $f(a)f(b) < 0$;
- 2) $f(x)$ è derivabile due volte con continuità;
- 3) $f'(x)$ ed $f''(x)$ sono di segno costante;

il metodo della secante è convergente all'unica radice ξ dell'equazione $f(x) = 0$ in (a, b) , purché si usi la formula (C.25) se $f(a)f''(a) > 0$, la formula (C.26) se $f(b)f''(b) > 0$.

Dimostrazione. Sia, per fissare le idee, $f''(x) > 0$ in $a \leq x \leq b$ (se questa ipotesi non fosse soddisfatta le considerazioni seguenti cambierebbero solo formalmente). La curva $y = f(x)$ è allora convessa e disposta al di sotto della corda che unisce i punti $(a, f(a))$, $(b, f(b))$. Due casi sono allora possibili:

1) $f(a) > 0$;

2) $f(b) > 0$.

Nel primo caso, l'estremità è fissa e le approssimazioni successive:

$$x^{(0)} = b, \quad x^{(i+1)} = x^{(i)} - \frac{f(x^{(i)})}{f(x^{(i)}) - f(a)}(x^{(i)} - a) \quad i = 0, 1, \dots,$$

formano una successione decrescente e limitata inferiormente da ξ .

Nel secondo caso, l'estremità b è fissa e le approssimazioni successive:

$$x^{(0)} = a, \quad x^{(i+1)} = x^{(i)} - \frac{f(x^{(i)})}{f(b) - f(x^{(i)})}(b - x^{(i)}) \quad i = 0, 1, \dots,$$

formano una successione crescente e limitata da ξ . Poiché, in entrambi i casi, la successione $\{x^{(i)}\}$ è monotona e contenuta in (a, b) , esiste il

$$\lim_{i \rightarrow \infty} x^{(i)} = \bar{\xi} \quad \text{con} \quad a < \bar{\xi} < b.$$

Pertanto, nel primo caso, passando al limite nella (C.25) si ha

$$\bar{\xi} = \bar{\xi} - \frac{f(\bar{\xi})}{f(\bar{\xi}) - f(a)}(\bar{\xi} - a)$$

da cui $f(\bar{\xi}) = 0$ e quindi $\bar{\xi} = \xi$, dato che, nelle ipotesi 1) – 2) del teorema, l'equazione $f(x) = 0$ possiede una sola radice in (a, b) . Procedendo allo stesso modo si trova che, anche nel secondo caso, la successione degli iterati $\{x^{(i)}\}$ converge all'unica radice ξ dell'equazione $f(x) = 0$ in (a, b) . ■

Il seguente teorema consente, sotto opportune ipotesi, di valutare la distanza tra l' i -esimo iterato e la radice in (a, b) dell'equazione $f(x) = 0$.

Teorema C.3.25 *Se esistono due numeri m_1 e M_1 tali che*

$$0 < m_1 \leq |f'(x)| \leq M_1 < \infty$$

per

$$\begin{cases} a \leq x \leq x^{(i)} & \text{nel primo caso,} \\ x^{(i)} \leq x \leq b & \text{nel secondo caso,} \end{cases}$$

tra l' i -esimo iterato $x^{(i)}$ del metodo della secante e l'unica radice ξ dell'equazione $f(x) = 0$ in (a, b) vale la seguente relazione:

$$|x^{(i)} - \xi| \leq \frac{M_1 - m_1}{m_1} |x^{(i)} - x^{(i-1)}|. \quad (\text{C.27})$$

Esempio C.3.26 Valutare, a meno di $1/100$, la radice positiva dell'equazione $f(x) = x^3 - 0.2x^2 - 0.2x - 1.2 = 0$.

Poiché $f(1) = -0.6$ ed $f(2) = 5.6$, la radice cercata appartiene all'intervallo $(1, 2)$. Dato che l'intervallo di indeterminazione è grande, possiamo ridurlo per bisezione calcolando $f(3/2)$. Poiché $f(1/2) = 1.425$, la radice cercata appartiene all'intervallo $(1, 1.5)$.

Applicando ora la relazione di ricorrenza (C.26) otteniamo:

$$\begin{aligned}x^{(1)} &= 1 + \frac{0.6}{1.425 + 0.6}(1.5 - 1) = 1.15, \\x^{(2)} &= 1.15 + \frac{0.173}{1.425 + 0.173}(1.5 - 1.15) = 1.190, \\x^{(3)} &= 1.190 + \frac{0.036}{1.425 + 0.036}(1.5 - 1.190) = 1.198.\end{aligned}$$

Poiché, nell'intervallo $[1.198, 1.5]$,

$$6.073(1.5)^2 - 0.4(1.198) - 0.2f'(x)3(2.298)^2 - 0.4(1.5) - 0.2 = 3.49$$

per la (C.27)

$$0 < \xi - x^{(3)} < \frac{2.58}{6.07}0.008 < \frac{4}{1000}$$

e pertanto $x^{(3)}$ è accettabile come valutazione della radice. Il valore esatto della radice è 1.2 .

Osservazione C.3.27 *Mentre il metodo di Newton, almeno quando la radice ξ dell'equazione $f(x) = 0$ è semplice, ha convergenza quadratica, il metodo della secante ha convergenza lineare. Tuttavia l'importanza del metodo della secante non è trascurabile, soprattutto perché: mentre nel metodo di Newton il calcolo di un nuovo iterato $x^{(i)}$ richiede la valutazione sia della funzione che della sua derivata in $x^{(i-1)}$, il metodo della secante, tranne nel primo passo (nel quale ne sono richieste due) richiede soltanto la valutazione della funzione in $x^{(i-1)}$. Nel calcolo di radici o di logaritmi, ottenuti mediante la risoluzione di equazioni del tipo*

$$\begin{aligned}x^k - \alpha &= 0 & k \in \mathbb{N} \quad \alpha > 0, \\ \log x - \alpha &= 0 & \alpha \in \mathbb{R},\end{aligned}$$

risulta più efficiente il metodo di Newton, anche se si tiene conto del maggior numero di valutazioni funzionali richiesto.

C.3.4 Altri metodi iterativi

Nei metodi iterativi per i sistemi lineari, nei quali la convergenza è tipicamente locale, è di grande importanza la seguente definizione di **punto di attrazione**:

Definizione C.3.28 *un punto $x^* \in \mathbb{R}^n$ è di attrazione per la successione degli iterati $x^{(k+1)} = \Phi(x^{(k)})$ se esiste un intorno I_{x^*} di x^* tale che, qualunque sia il punto iniziale $x^{(0)} \in I_{x^*}$, la successione degli iterati $x^{(k)} \rightarrow x^* = \Phi(x^*)$, ossia alla soluzione del sistema $F(x) = 0$; dato che*

$$x^* = \Phi(x^*) \iff F(x^*) = 0.$$

L'intorno I_{x^*} è detto *dominio di attrazione*. Nel caso particolare in cui $I_{x^*} = \mathbb{R}^n$, x^* è definito *punto di attrazione globale*.

Il risultato più importante è stabilito dal **teorema di Ostrowski**.

Teorema C.3.29 *Se la funzione di iterazione Φ è differenziabile in $x^* = \Phi(x^*)$ e inoltre $\rho(\Phi'(x^*)) < 1$, x^* è un punto di attrazione per gli iterati $x^{(k)}$.*

Supponiamo ora che la funzione di iterazione sia del tipo

$$\Phi(x) = x - [C(x)]^{-1}F(x), \quad (\text{C.28})$$

dove $C(x)$ è una soluzione nonsingolare in un punto fisso $x^* = \Phi(x^*)$, dove $F(x^*) = 0$. Sotto tali ipotesi,

$$\Phi'(x^*) = I - [C(x^*)]^{-1}F'(x^*). \quad (\text{C.29})$$

Da questo risultato segue immediatamente che, ogni punto fisso x^* della funzione di iterazione $\Phi(x) = x - [F'(x)]^{-1}F(x)$ è un punto di attrazione, nell'ipotesi che $F'(x^*)$ sia nonsingolare. Infatti, in tal caso, la matrice $\Phi'(x^*)$ risulta identicamente nulla e dunque $\rho(\Phi'(x^*)) = 0$.

Nel metodo di Newton $C(x) = F'(x)$, ossia la funzione di iterazione è

$$\Phi(x) = x - [F'(x)]^{-1}F(x).$$

Di conseguenza, nell'ipotesi che nel punto fisso x^* la $F'(x)$ sia nonsingolare, la (C.29) implica che $\Phi'(x^*)$ sia identicamente nulla, ossia che $\rho(\Phi'(x^*)) = 0$.

Metodo di Newton-Jacobi

Supponiamo ora di decomporre $F'(x)$ nel modo seguente:

$$F'(x) = D(x) - [L(x) + U(x)]$$

dove $D(x)$ è la diagonale di $F'(x)$, $L(x)$ e $U(x)$ rispettivamente i triangoli strettamente inferiore e superiore di $F'(x)$. Il metodo di Newton-Jacobi è caratterizzato da una funzione di iterazione di tipo (C.28) con

$$C(x) = D(x), \quad \Phi(x) = x - [D(x)]^{-1}F(x)$$

naturalmente nell'ipotesi che tutti gli elementi diagonali di $D(x)$ siano strettamente non nulli. Osserviamo ora che, nel caso lineare, dove $F(x) = Ax - b$, il metodo di Newton-Jacobi coincide con il metodo di Jacobi. In questo caso, essendo $F'(x) = D - (L + U)$,

$$\begin{aligned} \Phi(x) &= x - D^{-1}(Ax - b) = x - D^{-1}\{[D - (L + U)]x - b\} \\ &= D^{-1}[(L + U)x + b] = Hx + c, \quad H = D^{-1}(L + U), \quad c = D^{-1}b. \end{aligned}$$

Supponiamo ora che la F sia differenziabile in tutto un intorno di x^* nel quale $F(x^*) = 0$ e che la matrice Jacobiana F' sia continua in x^* . Se $D(x^*)$ è nonsingolare e inoltre il raggio spettrale

$$\rho\{[D(x^*)]^{-1}[L(x^*) + U(x^*)]\} < 1,$$

la funzione di iterazione $\Phi(x) = x - [D(x)]^{-1}F(x)$ soddisfa, in x^* , il teorema di Ostrowski e di conseguenza x^* è di attrazione per gli iterati $x^{(k)} = \Phi(x^{(k-1)})$. Per la dimostrazione è sufficiente osservare che per la (C.29),

$$\begin{aligned} \Phi'(x^*) &= I - [D(x^*)]^{-1}\{D(x^*) - [L(x^*) + U(x^*)]\} \\ &= [D(x^*)]^{-1}[L(x^*) + U(x^*)]. \end{aligned}$$

Metodo di Newton-Gauss-Seidel

In questo caso la funzione di iterazione è

$$\Phi(x) = x - [D(x) - L(x)]^{-1}F(x)$$

dove $D(x)$ e $L(x)$ hanno il significato precedentemente specificato. Procedendo come nel caso di Newton-Jacobi, è facile dimostrare che, nel caso

$D(x^*) - L(x^*)$ sia nonsingolare, il punto fisso x^* è di attrazione per gli iterati del metodo di Newton-Gauss-Seidel se

$$\rho\{[D(x^*) + U(x^*)]^{-1}U(x^*)\} < 1,$$

Da notare che nel caso lineare, la condizione $\rho(\Phi'(x^*)) < 1$ è necessaria e sufficiente per l'attrattività degli iterati. Nel caso nonlineare essa è sufficiente ma non necessaria. Inoltre, nel caso lineare, il dominio di attrattività è \mathbb{R}^n a differenza del caso nonlineare nel quale è puramente locale.

Bibliografia

- [1] W.J. Duncan, and A.R. Collar. *A method for the solution of oscillation problems by matrices*. Philosophical Magazine (1934); **17**(Series 7):865.
- [2] W.J. Duncan, and A.R. Collar. *Matrices applied to the motions of damped systems*. Philosophical Magazine (1934); **19**(Series 7):197.
- [3] R. Courant. *Variational methods for the solution of problems of equilibrium and vibrations*. Bulletin of the American Mathematical Society (1943); **49**:161.
- [4] M.J. Turner, R.W. Clough, H.C. Martin, and L.C. Topp (1956). *Stiffness and Deflection Analysis of Complex Structures*. Journal of the Aeronautical Sciences **23**: 80582.
- [5] Strang, Gilbert, and George Fix (1973). *An Analysis of the Finite Element Method*. Englewood Cliffs: Prentice-Hall.
- [6] M.S. Gockenbach, *Understanding and Implementing the Finite Element Method*, SIAM, Philadelphia 2006.
- [7] F. Ihlenburg, *Finite Element Analysis of Acoustic Scattering*, Springer Verlag, New York, 1988.